The PHANGS-HST Survey: Physics at High Angular Resolution in Nearby Galaxies with the Hubble Space Telescope

Janice C. Lee[1,2] , Bradley C. Whitmore[3], David A. Thilker[4] , Sinan Deger[5] , Kirsten L. Larson[3] , Leonardo Ubeda[3],
Gagandeep S. Anand[3,6] , Médéric Boquien[7] , Rupali Chandar[8] , Daniel A. Dale[9] , Eric Emsellem[10,11] ,
Adam K. Leroy[12] , Erik Rosolowsky[13] , Eva Schinnerer[14] , Judy Schmidt[15] , James Lilly[9], Jordan Turner[9] ,
Schuyler Van Dyk[5] , Richard L. White[3] , Ashley T. Barnes[16] , Francesco Belfiore[17] , Frank Bigiel[16] ,
Guillermo A. Blanc[18,19] , Yixian Cao[20] , Melanie Chevance[21] , Enrico Congiu[19] , Oleg V. Egorov[21,22] ,
Simon C. O. Glover[23] , Kathryn Grasha[24] , Brent Groves[24,25] , Jonathan D. Henshaw[14] , Annie Hughes[26,27] ,
Ralf S. Klessen[23,28] , Eric Koch[13,29] , Kathryn Kreckel[21] , J. M. Diederik Kruijssen[21] , Daizhong Liu[14] ,
Laura A. Lopez[12,30] , Ness Mayker[12,30] , Sharon E. Meidt[31] , Eric J. Murphy[32] , Hsi-An Pan[14,33] , Jérôme Pety[34,35] ,
Miguel Querejeta[36] , Alessandro Razza[19] , Toshiki Saito[14] , Patricia Sánchez-Blázquez[37,38] , Francesco Santoro[14] ,
Amy Sardone[12,30] , Fabian Scheuermann[21], Andreas Schruba[39], Jiayi Sun[12] , Antonio Usero[36] , E. Watkins[21], and
Thomas G. Williams[14]

[1] Gemini Observatory/NSFs NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA; janice.lee@noirlab.edu
[2] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[3] Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21218, USA
[4] Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA
[5] Caltech-IPAC, 1200 E. California Blvd., Pasadena, CA 91125, USA
[6] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Dr., Honolulu, HI 96822, USA
[7] Centro de Astronomía (CITEVA), Universidad de Antofagasta, Avenida Angamos 601, Antofagasta, Chile
[8] University of Toledo, 2801 W. Bancroft St., Mail Stop 111, Toledo, OH 43606, USA
[9] Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA
[10] European Southern Observatory, Karl-Schwarzschild Straße 2, D-85748 Garching bei München, Germany
[11] Univ Lyon, Univ Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France
[12] Department of Astronomy, The Ohio State University, 140 West 18th Ave., Columbus, OH 43210, USA
[13] Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada
[14] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany
[15] Astrophysics Source Code Library, Michigan Technological University, 1400 Townsend Dr., Houghton, MI 49931, USA
[16] Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany
[17] INAF—Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50157, Firenze, Italy
[18] Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101, USA
[19] Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile
[20] Aix Marseille Univ, CNRS, CNES, LAM (Laboratoire d'Astrophysique de Marseille), F-13388 Marseille, France
[21] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany
[22] Sternberg Astronomical Institute, Lomonosov Moscow State University, Universitetsky pr. 13, 119234 Moscow, Russia
[23] Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany
[24] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[25] International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia
[26] CNRS, IRAP, 9 Av. du Colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France
[27] Université de Toulouse, UPS-OMP, IRAP, F-31028 Toulouse cedex 4, France
[28] Universität Heidelberg, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany
[29] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA
[30] Center for Cosmology and AstroParticle Physics, 191 West Woodruff Ave., Columbus, OH 43210, USA
[31] Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium
[32] National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA 22903-2475, USA
[33] Tamkang University, No. 151, Yingzhuan Rd., Tamsui Dist., New Taipei City 251301, Taiwan
[34] Institut de Radioastronomie Millimétrique (IRAM), 300 Rue de la Piscine, F-38406 Saint Martin d'Hères, France
[35] Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, F-75014 Paris, France
[36] Observatorio Astronómico Nacional (IGN), C/Alfonso XII, 3, E-28014 Madrid, Spain
[37] Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, E-28040 Madrid, Spain
[38] Instituto de Física de Partículas y del Cosmos, Universidad Complutense de Madrid, E-28040 Madrid, Spain
[39] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße 1, D-85748 Garching, Germany

Abstract

The PHANGS program is building the first data set to enable the multiphase, multiscale study of star formation across the nearby spiral galaxy population. This effort is enabled by large survey programs with the Atacama Large Millimeter/submillimeter Array (ALMA), MUSE on the Very Large Telescope, and the Hubble Space Telescope (HST), with which we have obtained CO(2–1) imaging, optical spectroscopic mapping, and high-resolution UV–optical imaging, respectively. Here, we present PHANGS-HST, which has obtained NUV–*U*–*B*–*V*–*I* imaging of

the disks of 38 spiral galaxies at distances of 4–23 Mpc, and parallel *V*- and *I*-band imaging of their halos, to provide a census of tens of thousands of compact star clusters and multiscale stellar associations. The combination of HST, ALMA, and VLT/MUSE observations will yield an unprecedented joint catalog of the observed and physical properties of ∼100,000 star clusters, associations, H II regions, and molecular clouds. With these basic units of star formation, PHANGS will systematically chart the evolutionary cycling between gas and stars across a diversity of galactic environments found in nearby galaxies. We discuss the design of the PHANGS-HST survey and provide an overview of the HST data processing pipeline and first results. We highlight new methods for selecting star cluster candidates, morphological classification of candidates with convolutional neural networks, and identification of stellar associations over a range of physical scales with a watershed algorithm. We describe the cross-observatory imaging, catalogs, and software products to be released. The PHANGS high-level science products will seed a broad range of investigations, in particular, the study of embedded stellar populations and dust with the James Webb Space Telescope, for which a PHANGS Cycle 1 Treasury program to obtain eight-band 2–21 μm imaging has been approved.

Unified Astronomy Thesaurus concepts: Star formation (1569); Star clusters (1567); Young star clusters (1833); Surveys (1671); Spiral galaxies (1560); Hubble Space Telescope (761)

1. Introduction

How do stars form from the complex multiphase interstellar medium (ISM) in galaxies? This question lies at the heart of astrophysics, as star formation is a key mechanism governing the evolution of baryons in the universe (e.g., Péroux & Howk 2020). Star formation converts interstellar matter into stars and their planetary systems, depletes galaxies of gas, and feeds back metals, energy, and momentum into the ISM, which may reach the halos of galaxies and beyond. In turn, this feedback, together with galactic-scale inflows and dynamics, impacts the state of the gas and the future of star formation.

Many processes underlie this star formation cycle, and nearly all have been the focus of dedicated study. Star formation typically occurs in molecular clouds (e.g., Blitz 1993; Heyer & Dame 2015; Miville-Deschênes et al. 2017), parts of which can become gravitationally unstable and contract until new stars are born (e.g., Krumholz & McKee 2005; Dobbs et al. 2014; Chevance et al. 2020). This process is controlled by the intricate interplay between self-gravity and various opposing agents, such as supersonic turbulence, magnetic fields, radiation, and gas and cosmic-ray pressure (e.g., Elmegreen 2000; McKee & Ostriker 2007; Girichidis et al. 2020). The flow patterns within the galactic ISM help determine where and at what rate stars form, and are themselves influenced by the energy and momentum input from massive stars (e.g., Mac Low & Klessen 2004; McKee & Ostriker 2007; Hennebelle & Falgarone 2012; Federrath 2013). That is, the local process of stellar birth is impacted by the supply, organization, and stability of cloud-scale natal gas as governed by large-scale galaxy dynamics, including spiral arm features or perturbations from satellite galaxies or accretion of fresh gas from the cosmic web (e.g., Kennicutt 1998; Dobbs et al. 2006; Dobbs 2008; Leroy et al. 2008, 2013; Meidt et al. 2013, 2020). Stellar feedback, in the form of radiation, winds, and supernova explosions, creates a hierarchy of highly nonlinear feedback loops that impact ISM dynamics across a wide range of physical scales (e.g., Hopkins et al. 2014; Lopez et al. 2014; Walch et al. 2015; Gnedin 2016; Rahner et al. 2017; Olivier et al. 2021), thereby determining the chemical and thermal state of the ISM and affecting subsequent star formation (e.g., Klessen & Glover 2016).

We now understand from decades of study and observations across the electromagnetic spectrum that all of these processes that drive, regulate, and extinguish star formation operate together over a vast range of stellar, interstellar, galactic, and circumgalactic scales. Accordingly, we have come to recognize that systematic observations—spanning essential spatial scales and phases of the star formation cycle, across different galactic environments—are necessary for the development of a robust, unified model of star formation and galaxy evolution.

Here, we present the PHANGS−Hubble Space Telescope (HST) Treasury survey, which as part of the Physics at High Angular Resolution in Nearby Galaxies[40] (PHANGS) program, is building a data set for the systematic multiscale, multiphase study of star formation. PHANGS is charting the connections between giant molecular clouds, H II regions, and young stars throughout a diversity of galactic environments in the local universe by combining observations from large surveys with the Atacama Large Millimeter/submillimeter Array (ALMA), MUSE on the Very Large Telescope (VLT), and HST. Supporting data including Very Large Array (VLA) H I and Astrosat far-ultraviolet/near-ultraviolet imaging, as well as the wealth of panchromatic ground- and space-based survey observations obtained for the nearby galaxies in the sample over the past three decades, have also been assembled. There are currently three major components of PHANGS:

PHANGS-ALMA: The foundation of PHANGS has been built with the transformative capabilities of ALMA. Through a Cycle 5 PHANGS-ALMA Large Program (PI E. Schinnerer) and smaller precursor programs, PHANGS has obtained ∼1″ resolution CO(2–1) maps for a sample of 74 massive spiral galaxies at distances of 4–23 Mpc. At these distances, ALMA can detect individual giant molecular clouds with better than 2.4 km s^{-1} velocity resolution and physical resolutions of ∼50–100 pc, while still efficiently covering the star-forming disk (Sun et al. 2018; Leroy et al. 2021a, 2021b).

PHANGS-MUSE and PHANGS-Hα: For 19 of these galaxies, with the VLT Multi Unit Spectroscopic Explorer (MUSE), PHANGS-MUSE (PI E. Schinnerer) has obtained IFU spectroscopy with ∼2.5 Å spectral resolution and ∼0″7 spatial resolution to deliver a 3D view of the ionized (10^4 K) gas, stellar populations, and kinematics via various gas and stellar tracers in the optical from 4800 to 9300 Å (Emsellem et al. 2021; see first results in Kreckel et al. 2018; Ho et al. 2019; Kreckel et al. 2019, 2020). To supplement the MUSE observations, the PHANGS-Hα survey (A. Razza et al. 2021, in preparation) has obtained seeing-limited (∼1″) narrowband Hα imaging to provide star formation rate (SFR) maps and catalogs of ionized nebulae for the full PHANGS-ALMA

---

[40] http://www.phangs.org

sample. The Hα imaging was obtained using WFI on the ESO/MPG 2.2 m telescope at La Silla, and the DirectCCD on the du Pont 2.5 m telescope at Las Campanas Observatory.

PHANGS-HST, the subject of this paper, is a Cycle 26 HST Treasury survey (PI J. C. Lee) that has obtained NUV–U–B–V–I imaging of the disks of 38 galaxies from the parent PHANGS-ALMA sample and parallel V- and I-band imaging of their halos. The HST sample includes all 19 galaxies with MUSE IFU spectroscopy.

The high-resolution capabilities of HST ($\sim 0''.08$) have enabled the study of compact star clusters and associations in galaxies out to distances of several tens of Mpc (e.g., Whitmore et al. 1999; Linden et al. 2017; Adamo et al. 2020). These structures, which typically have half-light radii of a few parsecs (Portegies Zwart et al. 2010; Ryon et al. 2017; Krumholz et al. 2019), have been the focus of much recent work and are not only important to study in their own right (e.g., Whitmore et al. 2007; Chandar et al. 2010a; Kruijssen 2012; Krumholz et al. 2019; Adamo et al. 2020), but also have great utility as "clocks"—effectively single-aged stellar populations that can be age-dated and used to time various star formation and ISM processes. The PHANGS-HST UV–optical imaging enables inventories of young star clusters and associations down to a few thousand solar masses, with age and mass determinations from spectral energy distribution (SED) fitting accurate to a factor of ~ 2 (e.g., Turner et al. 2021) on average.

Altogether, PHANGS will yield an unprecedented sample of $\sim 100,000$ star clusters, associations, H II regions, and molecular clouds in diverse galactic environments to provide answers to the following open questions:

1. What are the timescales for the onset of star formation in clouds, the destruction of clouds, and the removal of gas from young star clusters?
2. How are the mass functions of star clusters/associations related to those of clouds? What are the implied star formation efficiencies?
3. How are star formation and gas organized into multiscale structures? How do their relative spatial distributions evolve with time?

These questions, particularly those examining the relationship between molecular clouds and star clusters, have been posed in the context of the Milky Way (Murray 2011; Lee et al. 2016) and select Local Group and Nearby galaxies (e.g., M51: Hughes et al. 2013; Grasha et al. 2019; NGC 7793: Grasha et al. 2018; NGC 300: Kruijssen et al. 2019). But whether the answers to the questions vary with galactic environment is still unclear, as there has not yet been a systematic study on the cluster scale across a well-defined sample of galaxies spanning a broad range of global properties. PHANGS will provide not only the answers to these questions, but moreover will show how they depend on galactic properties such as the phase balance and physical conditions of the ISM, stellar mass, gas mass, SFR (as well as their surface densities), metallicity, and the presence of dynamical features such as rings, bars, and spiral arms.

Here, we present the design of the PHANGS-HST survey and provide an overview of the data processing pipeline developed to generate the data products required for the investigation of these questions. The remainder of this paper is organized as follows. In Section 2, we describe how galaxies are selected from the parent PHANGS sample for observation

with HST and summarize the global properties of the sample. Our HST imaging observations with WFC3 and ACS are described in Section 3. Section 4 describes the PHANGS-HST pipeline used to produce catalogs of compact star clusters and associations and to measure their observed and physical properties. This high-level description is intended to provide the framework for a series of papers, as summarized in Section 4, that document each of the major components in detail, in particular new methods for selecting star cluster candidates, morphological classification of candidates, and identification of stellar associations over a range of physical scales. Data products resulting from this pipeline that will be released to support community science are described in Section 5. In Section 6, we conclude with a summary and look ahead to upcoming work on dust and young embedded stellar populations with the James Webb Space Telescope (JWST). Through a Cycle 1 Treasury program, we will add a fourth major component of PHANGS and obtain eight-band 2–21 μm imaging for the 19 galaxies with the full suite of PHANGS-ALMA, MUSE, and HST observations.

Magnitudes in this and other PHANGS-HST pipeline papers are given in the Vega system, unless otherwise noted, to facilitate comparison with prior HST studies of resolved stellar populations.

2. Galaxy Sample

Galaxies are chosen for HST observations from the PHANGS-ALMA Large Program sample (PI E. Schinnerer). PHANGS-ALMA has obtained CO(2–1) maps for a complete sample of 74 southern galaxies[41] ($-75° < \delta < 20°$; i.e., ALMA-observable), which were selected to be massive ($M_* \gtrsim 10^{9.75}\ M_\odot$), star-forming, not edge-on to the line of sight, and at distances $\lesssim 17$ Mpc.[42] CO(2–1) observations were obtained for this sample via an ALMA Cycle 5 Large Program (2017.1.00886.L), which builds upon and incorporates several smaller precursor programs in Cycles 2–3.[43] A full description of the PHANGS-ALMA Large Program sample criteria and derivation of integrated properties used for selection, such as stellar mass (M_*), SFR, and integrated CO luminosities, is given in Leroy et al. (2021b).

With HST, we target the galaxies best suited for joint HST-ALMA analysis of resolved young stellar populations and giant molecular clouds. That is, we select galaxies from the PHANGS-ALMA parent sample that (1) have inclinations $i \lesssim 70°$ to minimize source blending and attenuation along the

---

[41] The original PHANGS-ALMA Large Program sample, from which PHANGS-HST targets were selected, consists of 74 galaxies. PHANGS has now extended the sample to 90 galaxies to include additional nearby galaxies with CO mapping available from the ALMA archive, as well as early-type galaxies, as explained in more detail in Leroy et al. (2021b).

[42] After a recent update to the distance determinations, including the addition of new TRGB distances from our parallel ACS V- and I-band imaging (Section 3.2), we find that galaxies in the PHANGS-HST sample lie at distances between 4.4 and ~ 23 Mpc, with a median of ~ 16 Mpc (Figure 1). Uncertainties in distances that were used in the initial PHANGS galaxy selection, some of which were based on recessional velocities corrected for peculiar motions based on a flow model, led to the inclusion of galaxies that lie beyond the initial ~ 17 Mpc limit. Further discussion of the impact of distance uncertainties on the PHANGS sample selection can be found in Leroy et al. (2021b). Full details on the compilation of best-available distances are provided in Anand et al. (2020) and summarized in Section 3.2.

[43] Cycle 2 program 2013.1.00650.S (PI E. Schinnerer), Cycle 3 program 2013.1.01161.S (PI K. Sakamoto), Cycle 3 program 2015.1.00925.S (PI G. A. Blanc), Cycle 3 program 2015.1.00956.S (PI A. K. Leroy), and additional programs in Table 2, Leroy et al. (2021b).

Table 1
PHANGS-HST Galaxy Sample

Galaxy	α (J2000)	δ (J2000)	b (deg)	D (Mpc)	σ(D) (Mpc)	Method	D(Ref)	i (deg)	T	SFR$_{tot}$ ($M_\odot$ yr^{-1})	SFR ($M_\odot$ yr^{-1})	log M_* (log $M_\odot$)	Σ$_{CO}$ ($M_\odot$ kpc^{-2})
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
IC 1954	03^{h}31^{m}31^s.39	−51°54′17″.4	−51.201	12.8	2.05	NAM+TF	3+4+5	57	3.3	0.36	0.34	9.6	1.1
IC 5332*	23^{h}34^{m}27^s.49	−36°06′03″.9	−71.366	9.01	0.41	TRGB	18	27	6.8	0.41	0.11	9.5	−999
NGC 628*	01^{h}36^{m}41^s.75	+15°47′01″.2	−45.705	9.84	0.63	TRGB	2	9	5.2	1.75	0.93	10.2	1.5
NGC 685	01^{h}47^{m}42^s.81	−52°45′42″.5	−62.304	19.94	2.99	NAM	4+5	23	5.4	0.42	0.37	9.9	0.6
NGC 1087*	02^{h}46^{m}25^s.16	−00°29′55″.1	−51.651	15.85	2.24	Group	6	43	5.2	1.31	1.25	9.9	1.3
NGC 1097	02^{h}46^{m}19^s.05	−30°16′29″.6	−64.681	13.58	2.04	NAM	4+5	49	3.3	4.74	3.85	10.7	2.2
NGC 1300*	03^{h}19^{m}41^s.08	−19°24′40″.9	−55.223	18.99	2.85	NAM	4+5	32	4	1.17	1.02	10.6	1.0
NGC 1317	03^{h}22^{m}44^s.29	−37°06′13″.3	−56.693	19.11	0.84	Group	2	23	0.8	0.48	0.37	10.6	1.7
NGC 1365*	03^{h}33^{m}36^s.37	−36°08′25″.4	−54.598	19.57	0.78	TRGB	2	55	3.2	16.90	12.88	10.8	2.7
NGC 1385*	03^{h}37^{m}28^s.85	−24°30′01″.1	−52.706	17.22	2.58	NAM	4+5	44	5.9	2.09	1.96	9.9	1.3
NGC 1433*	03^{h}42^{m}01^s.55	−47°13′19″.5	−51.195	18.94	0.56	PNLF	7	29	1.5	1.13	0.56	10.4	1.3
NGC 1512*	04^{h}03^{m}54^s.28	−43°20′55″.9	−48.166	17.93	0.88	PNLF	7	43	1.2	1.28	0.59	10.6	1.1
NGC 1559	04^{h}17^{m}35^s.77	−62°47′01″.2	−41.198	19.44	0.44	Mira	8	65	5.9	3.76	3.76	10.2	1.5
NGC 1566*	04^{h}20^{m}00^s.42	−54°56′16″.1	−43.393	17.69	2.00	Group	6	30	4	4.54	3.21	10.7	2.0
NGC 1672*	04^{h}45^{m}42^s.50	−59°14′49″.9	−38.990	19.40	2.91	NAM	4+5	43	3.3	7.60	6.60	10.6	2.1
NGC 1792	05^{h}05^{m}14^s.45	−37°58′50″.7	−36.453	16.20	2.43	NAM	4+5	65	4	3.70	3.23	10.5	1.8
NGC 2775	09^{h}10^{m}20^s.12	+07°02′16″.6	33.988	23.15	3.47	NAM	4+5	41	1.6	0.87	0.70	11.1	1.2
NGC 2835*	09^{h}17^{m}52^s.91	−22°21′16″.8	18.509	12.22	0.94	TRGB	18	41	5	1.24	0.57	9.8	0.9
NGC 2903	09^{h}32^{m}10^s.11	+21°30′03″.0	44.540	9.61	0.39	TRGB	19	67	4	3.08	2.39	10.6	1.7
NGC 3351*	10^{h}43^{m}57^s.70	+11°42′13″.7	56.368	9.96	0.33	TRGB	2	45	3.1	1.32	0.87	10.3	1.4
NGC 3621	11^{h}18^{m}16^s.51	−32°48′50″.6	26.099	7.06	0.28	TRGB	18	66	6.9	0.99	0.78	10.0	1.4
NGC 3627*	11^{h}20^{m}14^s.96	+12°59′29″.5	64.418	11.32	0.48	TRGB	2	57	3.1	3.84	3.31	10.7	1.9
NGC 4254*	12^{h}18^{m}49^s.60	+14°24′59″.4	75.190	13.1	2.8	SCM	14	34	5.2	3.07	2.77	10.3	2.2
NGC 4298	12^{h}21^{m}32^s.76	+14°36′22″.2	75.673	14.92	1.37	TRGB	18	59	5.1	0.46	0.43	9.9	1.4
NGC 4303*	12^{h}21^{m}54^s.90	+04°28′25″.1	66.276	16.99	3.04	Group	6	24	4	5.33	4.25	10.6	2.1
NGC 4321*	12^{h}22^{m}54^s.83	+15°49′18″.5	76.898	15.21	0.49	Cepheid	10	39	4	3.56	2.43	10.7	1.6
NGC 4535*	12^{h}34^{m}20^s.31	+08°11′51″.9	70.641	15.77	0.37	Cepheid	10	45	5	2.16	1.23	10.5	1.2
NGC 4536	12^{h}34^{m}27^s.05	+02°11′17″.3	64.730	16.25	1.13	TRGB	2	66	4.3	3.45	3.12	10.2	1.4
NGC 4548	12^{h}35^{m}26^s.45	+14°29′46″.8	76.830	16.22	0.38	Cepheid	10	38	3.1	0.52	0.34	10.7	1.0
NGC 4569	12^{h}36^{m}49^s.79	+13°09′46″.6	75.623	15.76	2.36	Group	6	70	2.4	1.32	1.22	10.8	1.4
NGC 4571	12^{h}36^{m}56^s.38	+14°13′02″.5	76.654	14.9	1.2	Cepheid	16	33	6.4	0.29	0.26	10.0	0.9
NGC 4654	12^{h}43^{m}56^s.58	+13°07′36″.0	75.889	21.98	1.16	Group	10	56	5.9	3.79	3.06	10.5	1.4
NGC 4689	12^{h}47^{m}45^s.56	+13°45′46″.1	76.607	15.0	2.25	NAM+TF	3+4+5	39	4.7	0.4	0.4	10.1	0.9
NGC 4826	12^{h}56^{m}43^s.64	+21°40′58″.7	84.423	4.41	0.19	TRGB	18	59	2.2	0.2	0.17	10.2	1.5
NGC 5068*	13^{h}18^{m}54^s.81	−21°02′20″.8	41.376	5.20	0.21	TRGB	18	36	6	0.28	0.20	9.4	0.8
NGC 5248	13^{h}37^{m}32^s.02	+08°53′06″.6	68.751	14.87	1.34	Group	6	47	4	2.29	1.66	10.3	1.8
NGC 6744	19^{h}09^{m}46^s.10	−63°51′27″.1	−26.146	9.39	0.43	TRGB	18	53	4	2.41	0.71	10.7	1.0
NGC 7496*	23^{h}09^{m}47^s.29	−43°25′40″.6	−63.801	18.72	2.81	NAM	4+5	36	3.2	2.26	2.00	9.8	1.2

Note. Column 1: Galaxy name. * Indicates PHANGS-MUSE integral field spectroscopy available, and will be observed in the PHANGS-JWST Treasury Survey. Columns 2−3: R.A. and decl. Column 4: Galactic latitude. Columns 5−8: Galaxy distances, uncertainties, and references as follows: (1) Karachentsev et al. (2004), (2) Jacobs et al. (2009), (3) Tully et al. (2016), (4) Shaya et al. (2017), (5) Kourkchi et al. (2020a, 2020b), (6) Kourkchi & Tully (2017), (7) F. Scheuermann et al. (2021, in preparation), (8) Huang et al. (2020), (9) Leonard et al. (2003), (10) Freedman et al. (2001), (11) Olivares E. et al. (2010), (12) Barbarino et al. (2015), (13) Tonry et al. (2001), (14) Nugent et al. (2006), (15) Reid et al. (2019), (16) Pierce et al. (1994), (17) Ruiz-Lapuente (1996), (18) Anand et al. (2020), (19) this paper. Column 9: Galaxy inclination. Following Leroy et al. (2021b) and adopted from Lang et al. (2020). Column 10: Morphological T-type. Column 11: Total galaxy star formation rate. Based on GALEX far-UV and WISE W4 imaging with SFR prescription calibrated to match results from population synthesis modeling of Salim et al. (2016, 2018) as in Leroy et al. (2021b). For NGC 685 and NGC 4689, GALEX far-UV imaging is not available, and the SFRs are based only on WISE W4 data. Column 12: Star formation rate, as computed for column 11, but limited to areas studied by PHANGS-HST, as shown in the footprints illustrated in Figure 3 and provided for the full galaxy sample at https://archive.stsci.edu/hlsp/phangs-hst. Column 13: Galaxy stellar mass. Following Leroy et al. (2021a), based on Spitzer IRAC 3.6 μm when available, or WISE 3.4 μm, and mass-to-light ratio prescription of Leroy et al. (2019) calculated as a function of radius in the galaxy. Column 14: Here, we calculate Σ$_{mol}$ adopting a fixed $\alpha_{CO}^{2-1} = 6$: 25 $M_\odot$ pc^{-2} (K km s^{-1})$^{-1}$, appropriate for a Galactic conversion factor and a typical CO(2−1)/CO(1−0) line ratio (Sun et al. 2018). Thus, the x-axis indicates mean CO surface brightness in units of mass surface density.
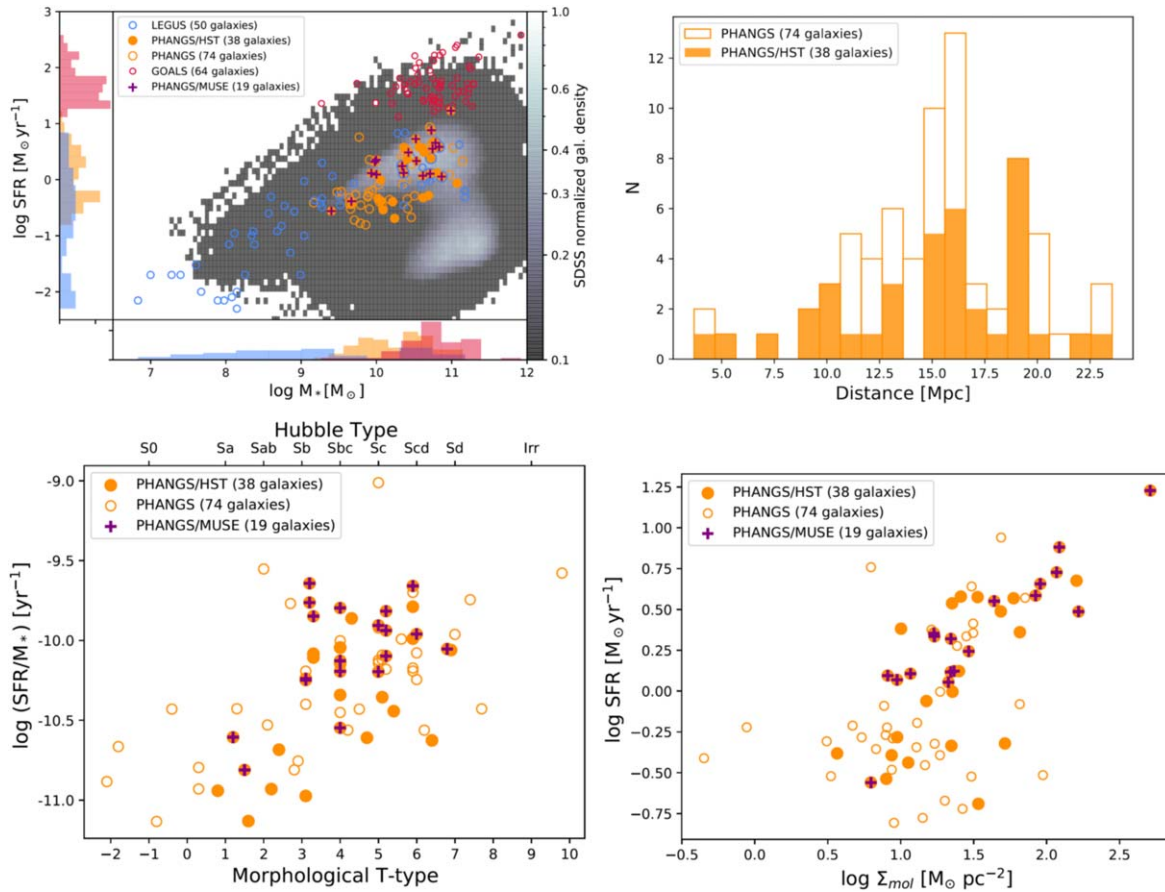
Figure 1. Upper left: coverage of the SFR–M_* plane by galaxies in the parent PHANGS sample ($N = 74$; orange open and filled points) and those targeted for HST observations ($N = 38$; orange filled points). The PHANGS-HST subsample is representative of the $d \lesssim 20$ Mpc massive galaxy population on the local star-forming main sequence, which contains the bulk of molecular gas and present-day star formation. Shown for context are: an SDSS local galaxy sample, which extends over a much larger volume to $z \lesssim 0.3$ (grayscale; Salim et al. 2016); the HST LEGUS sample (blue; Calzetti et al. 2015b), which focuses on the nearest galaxies ($d \lesssim 11$ Mpc) and includes a significant number of lower mass dwarfs; and luminous infrared galaxies within $\sim$350 Mpc from the GOALS sample (red; U et al. 2012). Upper right: distribution of distances of galaxies in the PHANGS parent sample and those selected for HST observations. All distances are provided in Table 1 and are taken from the compilation of Anand et al. (2020), with an update for one new TRGB distance derived from PHANGS-HST observations taken after the publication of that paper (see Section 3.2). Altogether, the PHANGS-HST V- and I-band parallel imaging have yielded eight TRGB measurements, which represent the best-available distances for those galaxies. Bottom panels: PHANGS-HST also provides coverage of a full range of specific SFRs, molecular gas surface densities, and spiral galaxy morphologies. Galaxies with MUSE IFU spectroscopy are indicated ($N = 19$) with crosses in the upper left and bottom plots. These 19 galaxies comprise the sample targeted by the PHANGS-JWST Cycle 1 Treasury program (see Section 6).

line of sight due to dust within the target, (2) avoid the Galactic plane ($|b| > 15°$) to minimize the impact of Milky Way reddening and foreground stars, and (3) are sufficiently active (SFR $\gtrsim 0.3\ M_\odot\ yr^{-1}$) to ensure that wide-spread molecular cloud and star cluster populations are available for joint study. The resulting set of 38 galaxies chosen for HST observations is given in Table 1 along with basic properties (which have been refined and updated since the sample was first compiled Leroy et al. 2021b) relevant to their selection. The 19 galaxies for which VLT/MUSE optical integral field spectroscopy has been obtained are all included by the HST selection criteria.

The 38 PHANGS-HST galaxies probe a full range of global properties covered by the PHANGS-ALMA parent sample, which is itself representative of the overall present-day spiral galaxy population. This is illustrated in Figure 1 (upper left panel), which shows M_* and SFR of the PHANGS sample and the subset targeted with HST, overlaid on the locus occupied by star-forming galaxies from the Sloan Digital Sky Survey. Figure 1 shows that the PHANGS samples provide excellent coverage of the galaxy "main sequence" between stellar masses

of $\sim 10^{9.5}$–$10^{11}\ M_\odot$. Main-sequence galaxies in this stellar mass range are representative of the environments where the bulk of molecular gas and present-day star formation are found (Salim et al. 2007; Saintonge et al. 2017). The bottom panels of Figure 1 illustrate the coverage of specific SFR as a function of morphological type and the SFR as a function of the molecular gas surface density. The PHANGS-HST observations include spiral galaxies with morphological types of Sa through Sd, sSFRs of $\sim 10^{-10.5}$–$10^{-9}\ yr^{-1}$, SFRs of ~ 0.2–$17\ M_\odot\ yr^{-1}$, and Σ_{mol} of $\sim 10^{0.5}$–$10^{2.7}\ M_\odot\ pc^{-2}$.

To place the PHANGS-HST survey in further context, nearby galaxies from two other complementary HST imaging programs are also shown in the upper left panel of Figure 1: HST LEGUS (Legacy ExtraGalactic Ultraviolet Survey; Calzetti et al. 2015a), and GOALS (Great Observatories All-Sky LIRG Survey; Armus et al. 2009). As a Cycle 21 Treasury program, LEGUS also obtained five-band UV–optical imaging to study star clusters in a representative sample of 50 galaxies (Adamo et al. 2017), but focused on the nearest late-type systems ($\lesssim 11$ Mpc; Kennicutt et al. 2008; Lee et al. 2011) to

also enable the reconstruction of star formation histories from individually resolved stars (Cignoni et al. 2018, 2019; Sacchi et al. 2018). LEGUS is therefore naturally dominated by lower mass, local volume dwarf and irregular galaxies (e.g., Lee et al. 2009), which comprise about half of its 50 galaxy sample (Cook et al. 2019). PHANGS-HST builds on the groundwork laid by programs such as LEGUS, and earlier transformative work by the Panchromatic Hubble Andromeda Treasury (PHAT; Dalcanton et al. 2012) and the WFC3 Early Release Science (ERS) program.[44] The WFC3 imaging observations obtained by these successive programs, which provide multi-wavelength coverage beginning in the near-UV (with ERS using up to 11 filters including both broad and narrow bands), have enabled the study of the physical properties of star clusters in samples of main-sequence galaxies that have grown from a few to over one hundred in the past decade (e.g., Chandar et al. 2010b, 2014; Johnson et al. 2016, 2017; Adamo et al. 2017; Grasha et al. 2017; Cook et al. 2019). At the edge of the parameter space occupied by nearby star-forming galaxies, GOALS has obtained HST imaging in the B, I, and H filters for $\sim$90 luminous infrared galaxies with thermal IR (8–1000 μm) dust emission greater than $10^{11}\,L_{\odot}$ (Haan et al. 2011; Kim et al. 2013). (A subsample of 64 LIRGS from U et al. 2012 with stellar masses derived from SED fitting is shown in Figure 1.) Such highly active star-forming galaxies are rare in the present-day universe, so the GOALS sample extends over much larger distances compared to LEGUS and PHANGS-HST. GOALS galaxies are located at distances of up to $\sim$350 Mpc ($z \lesssim 0.08$), and hence sample-wide studies of the stellar populations have focused on larger "clumps" ($\sim$90 pc, Larson et al. 2020, but also see Linden et al. 2017). Currently PHANGS is the only program with uniform ALMA CO observations for a significant sample of nearby galaxies, but ultimately, analysis of all of these programs together are needed to fully understand the impact of galactic environment on star formation.

3. HST Observations

Imaging observations for the PHANGS-HST Treasury program (Cycle 26, PID 15654) were conducted from 2019 April to 2021 April with an allocation of 122 orbits. Previous to this program, no HST wide-field UV imaging was available for 80% of the PHANGS-HST sample, and 60% also did not have any optical imaging with either WFC3 or ACS. Thus, PHANGS-HST provides a critical augmentation to the HST archive for the set of nearby spiral galaxies where both star clusters and molecular clouds can be efficiently detected by HST and ALMA over galactic scales.

As discussed in Section 2, the PHANGS-HST sample contains a total of 38 galaxies. No new observations were conducted for four of these galaxies, since sufficient imaging for those targets were previously obtained by the LEGUS program.[45] Altogether, new observations were planned for 43 fields in 34 galaxies. Initial observations for eight fields were corrupted due to guiding failures, but were all rescheduled and successfully observed on the second attempt. These

reobservations were the primary reason the execution of the program spanned over two years instead of one.

To illustrate the PHANGS coverage of each galaxy, footprints of the HST and ALMA observations, together with those for VLT/MUSE when available, are provided at MAST at https://archive.stsci.edu/hlsp/phangs-hst/. Examples of these footprint overlays are given in Figures 2 and 3 for six galaxies chosen to span the range of molecular gas surface densities, specific SFRs, and dynamical features (rings and bars) present in the sample. Figure 2 focuses on the PHANGS-HST prime imaging area in the context of the PHANGS-ALMA CO(2–1) maps. Figure 3 shows a wider areal view of each galaxy with Digitized Sky Survey images,[46] and also includes footprints of the PHANGS-HST parallel, PHANGS-ALMA, and PHANGS-MUSE observations.

3.1. Prime Observations with WFC3

For each of the 43 new fields, we aimed to cover the region mapped in CO(2–1) by ALMA in five filters: F275W (NUV), F336W (U), F438W (B), F555W (V), and F814W (I). F275W is the shortest wavelength filter that avoids the 2175 Å dust feature, and in combination with the U and B bands serves to break the age-extinction degeneracy. The V and I bands are less affected by extinction and variation in the mass-to-light ratio, and serve to constrain the stellar mass.

For 31 fields (in 26 galaxies), new observations with the WFC3 UVIS camera were needed in all five filters. Three exposures with subpixel dithering were taken in each filter (the average pixel size of WFC3 UVIS is 0″04), and all 15 exposures were obtained in a three-orbit visit, yielding total exposure times of $\sim$2200 s (NUV), $\sim$1100 s (U), $\sim$1100 s (B), $\sim$670 s (V), and $\sim$830 s (I). A post flash of 5–10 e^- is applied for each exposure, with larger values in the shorter wavelength filters, to increase the background to 12 e^-, which is the recommended value to mitigate issues due to charge transfer efficiency (CTE) losses. The dither sequence is optimized to cover the WFC3 chip gap and to help recover the undersampled point-spread function (PSF).

For the other 12 pointings (in eight galaxies), suitable WFC3 or ACS data in one or more of these filters were taken by prior programs. The available archival data were obtained from MAST and processed in a consistent manner with our new observations, and new imaging was obtained only for the missing filters within a two-orbit visit.

3.2. Parallel Observations with ACS

While the primary observational goal of the PHANGS-HST survey is to obtain UV–optical imaging of the star-forming disk, we also simultaneously observe the galaxy halo with the Advanced Camera for Surveys Wide Field Channel (ACS/WFC) in "parallel" mode. Such parallel observations can potentially yield measurements of the galaxy distance if the tip of the red giant branch (TRGB) can be identified in a color–magnitude diagram of halo stars. Thus, we have designed our observations so that ACS imaging in F606W and F814W, filters commonly used for TRGB analysis (e.g., McQuinn et al.
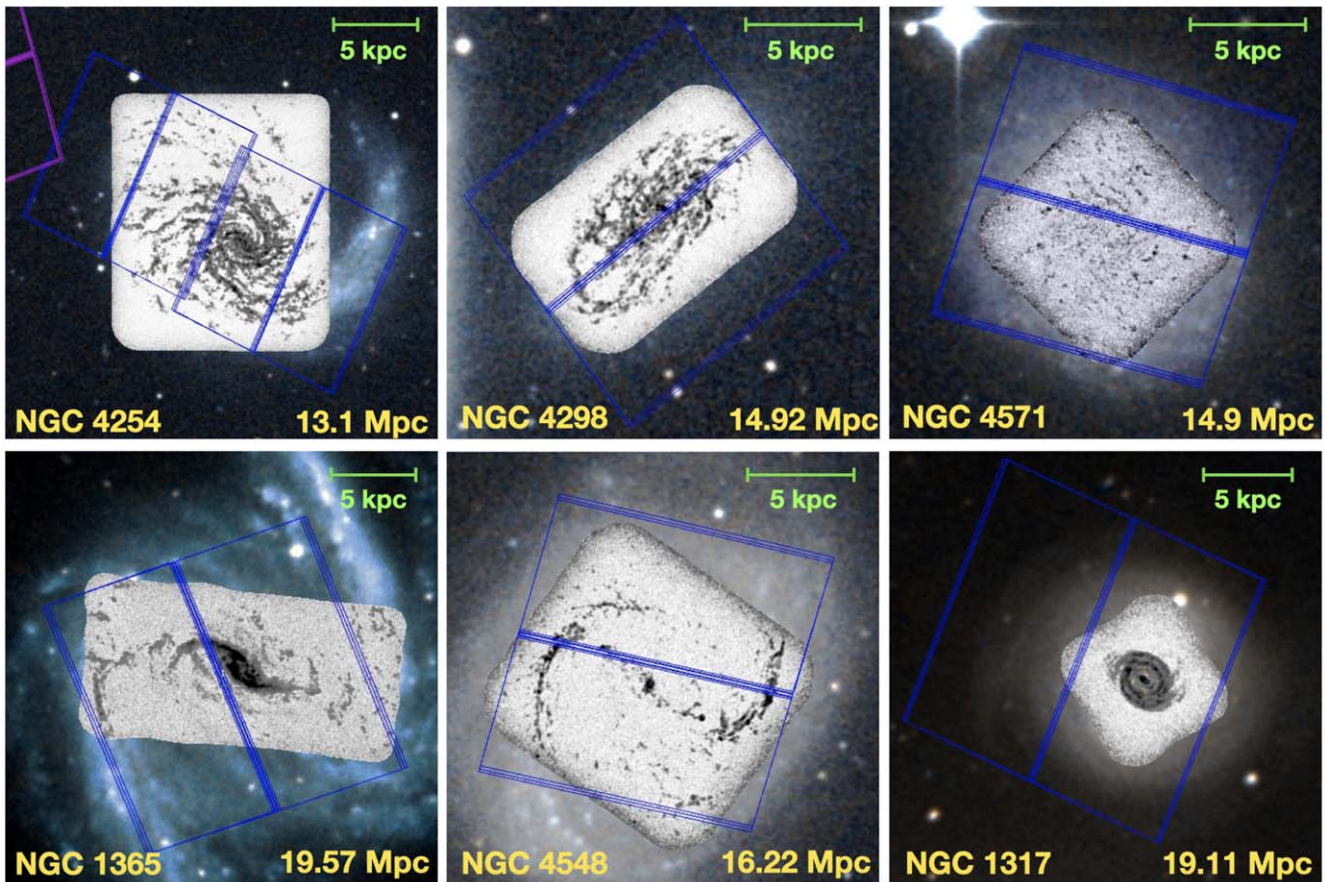

Figure 2. PHANGS-HST WFC3 UVIS footprints ($162'' \times 162''$) overlaid on PHANGS CO(2–1) ALMA maps for six galaxies in our sample of 38, showing the enormous diversity of molecular gas content and morphology in present-day massive star-forming galaxies. CO maps are overlaid on wider field DSS imaging. Scale bars and galaxy distances are shown in the upper and lower right corners, respectively. Top: Targets showing decreasing molecular gas surface density and specific SFR, from left to right. Bottom: impact of dynamical features on the gas distribution; two examples showing differing responses of the gas to the influence of a bar (left, middle) and an example of a ring feature (right). Our WFC3/UVIS observations allow us to find and characterize young stellar clusters and associations over the same area covered by these detailed CO maps, to create the first combined atlas of clouds and clusters across a representative sample of massive main-sequence galaxies in the local universe.

2017; Anand et al. 2018) accompanies each WFC3 "prime" observation.

For the range of distances and angular sizes of the PHANGS-HST galaxies, the ACS field of view generally falls on the halo of the target galaxy when WFC3 is centered on the galaxy itself (Figure 3). Given that the science requirements of PHANGS-HST constrain the positioning of the prime pointings, optimizing placement of the parallel fields (as in a focused TRGB program) is a secondary priority and is restricted by the fixed spatial offset of the two cameras on the focal plane. For galaxies with relatively large angular sizes, the parallel observations may include portions of the outer disk. For smaller galaxies, the parallels may be too far to capture a significant number of halo stars. To the extent possible, ORIENT constraints were imposed to prevent the ACS field from entirely falling on the galaxy disk, on nearby galaxy neighbors, and/or on extremely bright foreground stars. For some targets with large angular sizes where it was not possible to avoid the disk completely, the field was positioned along the major axis to help differentiate between disk and halo stars. In several cases, the desired ORIENT constraints were lifted or relaxed to allow guide stars to fall into the area of the focal plane accessible to the Fine Guidance Sensors.

The five-band prime observations with WFC3/UVIS were sequenced in each orbit to optimize exposure time in the

parallel observations without impacting the primary observations. As discussed in the previous section, WFC3 observations for each pointing required two or three orbits, depending on whether suitable HST archival observations of the galaxy disk were already available. For two-orbit visit pointings, the total exposure times in the ACS parallel V and I images are $\sim$2100 s each, while for the three-orbit visits they are about $\sim$3500 s and 3200 s, respectively. In both cases, three exposures were taken in each filter.

Distance constraints resulting from analysis of the TRGB based on parallel imaging obtained in the first year of the PHANGS-HST program (for 30 galaxies through 2020 July) are presented in Anand et al. (2020). We were able to measure TRGB distances for 10 PHANGS galaxies, 4 of which are the first published TRGB distance measurements for those galaxies (IC 5332, NGC 2835, NGC 4298, and NGC 4321) and 7 of which represent the best-available distances (IC 5332, NGC 2835, NGC 3621, NGC 4298, NGC 4826, NGC 5068, and NGC 6744). Analysis of the remaining six galaxies (with seven parallel fields: IC 1954, NGC 0685, NGC 1097, NGC 2903-N/ S, NGC 5068,[47] and NGC 7496) yields one additional TRGB

---

[47] One of the targets reobserved due to guiding failure. In the first attempt, WFC3 prime observations were corrupted, but most parallel observations were still usable.
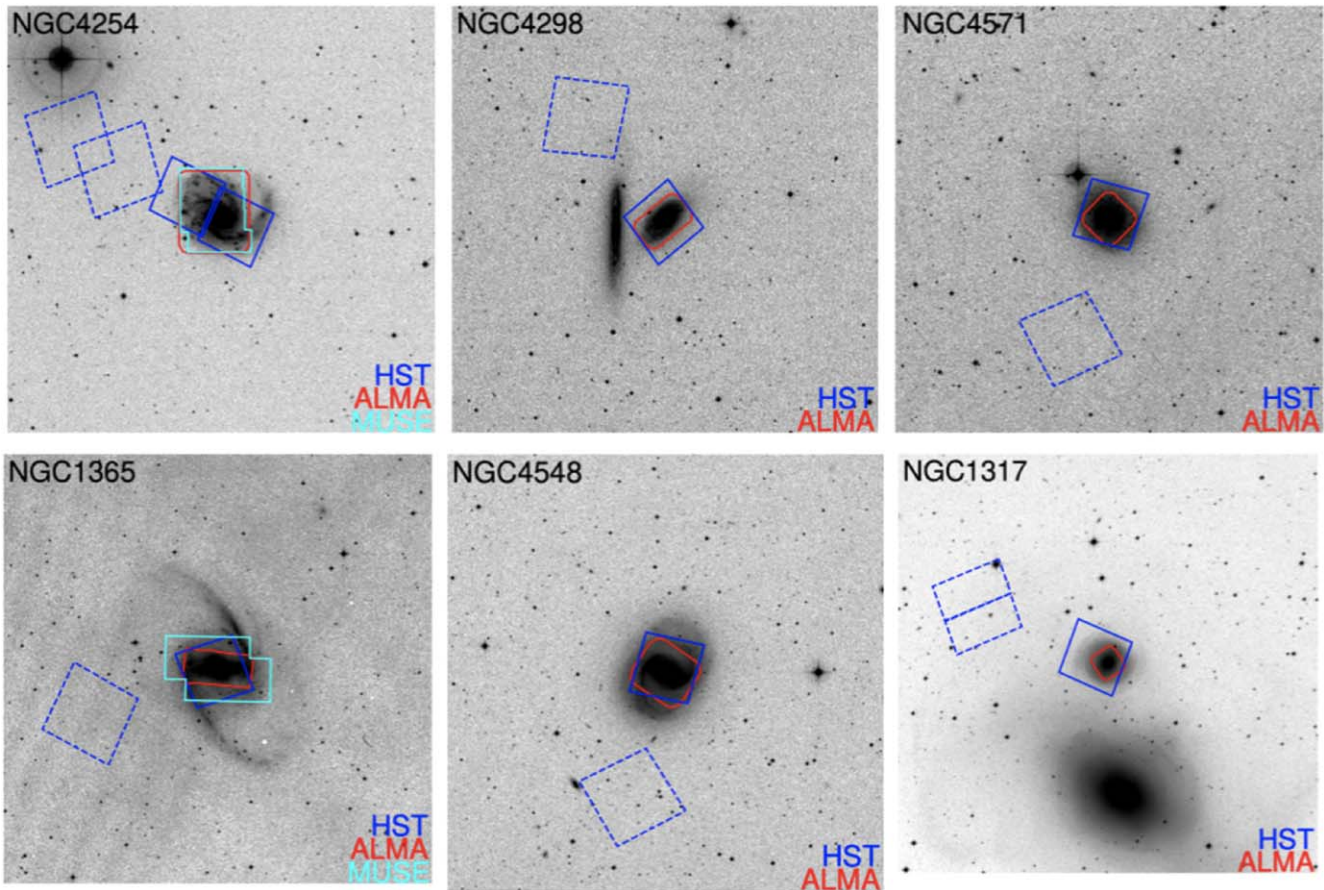
Figure 3. Figures showing the overlap of the PHANGS-HST WFC3 UVIS (blue), ALMA (red), and MUSE (where available; cyan) observation footprints, overlaid on DSS imaging for the same six galaxies as in Figure 2. A larger field (20′×20′) is shown relative to Figure 2 to illustrate the placement of the HST ACS parallel pointing (dashed lines). The WFC3 UVIS field of view is $162'' \times 162''$ and the ACS field of view is $202'' \times 202''$. Such footprint maps for the full PHANGS-HST sample can be found at https://archive.stsci.edu/hlsp/phangs-hst/.

constraint. A distance of 9.61 ± 0.39 Mpc is found for NGC 2903-N, and represents the best-available distance for that galaxy. All PHANGS-HST TRGB measurements and accompanying color–magnitude diagrams will be available from the Extragalactic Distance Database (EDD; Tully et al. 2009; Anand et al. 2021).[48] In total, the PHANGS-HST ACS parallel observations have provided eight new TRGB measurements which are the current best-available distances. The eight measurements span from 4.41 ± 0.19 Mpc (NGC 4826) to 14.9 ± 1.4 Mpc (NGC 4298), and are listed in Table 1 together with other adopted distances from the literature compiled by Anand et al. (2020).

4. Data Processing Pipeline

To enable the joint HST-ALMA-MUSE study of star formation in basic units of star clusters, associations, molecular clouds, and H II regions, we have developed an extensive HST data processing pipeline that produces inventories of stars (point sources), compact star clusters, and stellar associations across multiple physical scales in each galaxy. The pipeline yields aligned, mosaicked images in all five filters as well as catalogs of observed (e.g., photometry and morphological parameters) and physical properties (stellar masses, ages, and reddenings, derived through SED fitting). Ultimately, the

PHANGS-HST star cluster and association catalogs will be cross-correlated with the PHANGS-ALMA molecular cloud catalogs (Rosolowsky et al. 2021, A. Hughes et al. 2021, in preparation) and PHANGS-MUSE H II region catalogs (Santoro et al. 2021). Python packages and high-level science products from the HST pipeline are being publicly released and are described in Section 5.

Here, we summarize the overall strategy of the pipeline and provide a framework for subsequent papers that document each of the major components in detail. The key steps in the PHANGS-HST pipeline are:

1. Image drizzling, mosaicking, and astrometric calibration (this paper, Section 4.1).
2. Source detection and aperture photometry (Thilker et al. 2022, Section 4.2).
3. Identification of bright, isolated star clusters for determining aperture corrections (Deger et al. 2021, Section 4.3).
4. Derivation of aperture corrections for star clusters (Deger et al. 2021, Section 4.3).
5. Selection of candidate star clusters (Thilker et al. 2022, Section 4.4).
6. Morphological classification of candidate star clusters (Wei et al. 2020; Whitmore et al. 2021, Section 4.5).
7. SED fitting (Turner et al. 2021, Section 4.6).
8. Identification and photometry of stellar associations (K. Larson et al. 2021, in preparation, Section 4.7).

---

[48] https://edd.ifa.hawaii.edu/

Our processing workflow reflects common practice for the production of catalogs of compact star clusters in nearby galaxies (e.g., Whitmore et al. 2010; Adamo et al. 2017), with the following key augmentations. First, selection criteria are based on measurement of a series of concentration indices (CI; the difference in photometry measured with circular apertures of two different radii) rather than a single concentration index (Section 4.4). Second, we inject model star clusters into the HST imaging to aid the definition of selection criteria to separate candidate clusters from point sources and other interlopers; this also provides a foundation to estimate completeness in future work (Section 4.4). Third, we utilize convolutional neural network (CNN) models, as discussed in Wei et al. (2020), to supplement human visual inspection, with the goal of eventually automating morphological classification of candidate star clusters, as this has been a limiting step in past cluster studies (e.g., Adamo et al. 2017; Whitmore et al. 2021). Fourth, we separate the process for selecting multipeaked stellar associations from single-peaked compact stellar clusters, by applying a watershed algorithm to identify associations at physical scales from 8 pc to 64 pc (Section 4.7). This produces a far more complete inventory of young stellar populations, which is crucial for robust comparisons with molecular clouds.

A flowchart illustrating the steps in the pipeline as summarized below is provided in Figure 5.

4.1. Drizzling, Mosaicking, Astrometric Calibration

The process used to drizzle and mosaic the HST imaging data follows current standard procedures.

Data acquired for PHANGS-HST are first obtained from MAST, along with other suitable archival data taken by previous programs. These "FLT" exposures have been processed through the standard Pyraf/STSDAS CALACS or CALWFC3 software in the archive, which performs initial data quality flagging, bias subtraction, gain correction, bias stripe removal, correction for CTE losses, dark current subtraction, flat-fielding, and photometric calibration, resulting in "FLC" FITS files for each ACS/WFC and WFC3/UVIS exposure.

The PHANGS-HST pipeline is based on the STScI-supported software package DRIZZLEPAC to combine exposures and improve sampling of the PSF and is used for the prime observations targeting the galaxy disk in a two-step drizzle procedure. The parallel observations are treated separately as discussed in Anand et al. (2020).

The pipeline takes the FLC FITS files retrieved from MAST as input to produce combined images for each filter, which are all aligned and drizzled onto a common grid with a pixel scale of $0\farcs04$ (the native WFC3 pixel scale), with astrometry calibrated with GAIA DR2 sources (Gaia Collaboration et al. 2018; Lindegren et al. 2018). The latter is essential for the proper alignment of the HST imaging with ALMA and VLT/MUSE data and to facilitate a joint study of the three data sets.

The V-band imaging (WFC3 F555W) is used as the reference for the positioning of the images in all other filters (NUV, U, B, and I) and to define the common pixel grid. Using the F555W FLC files from MAST for each pointing, the sky positions of the centers and corners of the images are calculated to define a search area to query the ESA DR2 GAIA catalog. The TWEAKREG routine then matches the GAIA sources to the objects detected in the F555W drizzled image and calculates average shifts (with an accuracy typically better than 0.1 pixel) to correct the astrometric solution. The number

of GAIA sources found in a given F555W HST pointing varies from as few as 15 sources to a maximum of 317 with an average of 40 sources. The TWEAKBACK routine is then used to propagate the corrected WCS solution back to the original F555W FLC images. Finally, TWEAKREG and TWEAKBACK are used again, but now with the drizzled images for the other four filters to find sources in common with those detected in F555W and to align to the F555W image.

The final drizzle combination is carried out with sky subtraction using the "globalmin+match" method for sky calculation. ASTRODRIZZLE first finds a minimum "global" sky value for each chip/image extension in all input images and then uses the "match" method to compute differences in sky values between images in common sky regions and to equalize the sky values between images. The final DRC FITS files are in units of e^- s^{-1} and are registered with north up and east left as usual. "EXP" (exposure time) and "ERR" (error) weight (WHT) maps, calculated from the contribution of all input exposures to a given output pixel, are also produced. The ERR maps includes all noise sources from detector and sky and are used to compute uncertainties in photometry downstream in the pipeline.

Color composites of the drizzled, mosaicked HST imaging are shown in Figure 4 for the same six galaxies featured in Figures 2 and 3.

4.2. Source Detection and Photometry

The principal method of source detection in the PHANGS-HST pipeline is provided by the DOLPHOT photometry package (v2.0; Dolphin 2016), which is based on PSF fitting and operates on the FLC files to detect and deblend sources in HST imaging. We elect to use DOLPHOT for source detection in order to provide a common starting point for the identification of both single-peaked compact star clusters, and multipeaked stellar associations over a range of physical scales. The drizzled V-band image is used as the positional reference and sources are detected to 3.5σ with PSF fitting performed at the same image positions in all bands.

Compact star clusters have effective radii between 0.5 pc to about 10 pc (Portegies Zwart et al. 2010; Ryon et al. 2017). At the distances of the galaxies in the PHANGS-HST sample, such objects will have angular sizes close to the HST WFC3 resolution (i.e., 2 pixels, $\sim0\farcs08$; from 1.7 to 9 pc for the range of distances for the galaxy sample) and will appear sufficiently point-like to be captured by DOLPHOT. To ensure that the catalogs include star clusters with light profiles that may be more extended than the sources detected by DOLPHOT, source detection is also performed on the V-band DRC image with DAOStarFinder (the Python implementation of DAOPHOT within the photutils astropy-affiliated package), with a kernel FWHM of 2.5 pixels and an effective signal-to-noise ratio (S/N) consistent with the detection threshold used for DOLPHOT. The merged DOLPHOT and DAOStarFinder catalogs provide the source lists from which compact star cluster candidates are identified, while the branch of the pipeline that identifies multipeaked stellar associations (Section 4.7) relies only upon the DOLPHOT catalogs, as indicated in the flowchart in Figure 5.

Following source detection, the star cluster pipeline then uses the positions from the merged DOLPHOT-DAOStar-Finder catalog to perform aperture photometry in all five filters. For all galaxies photometry is measured in circular
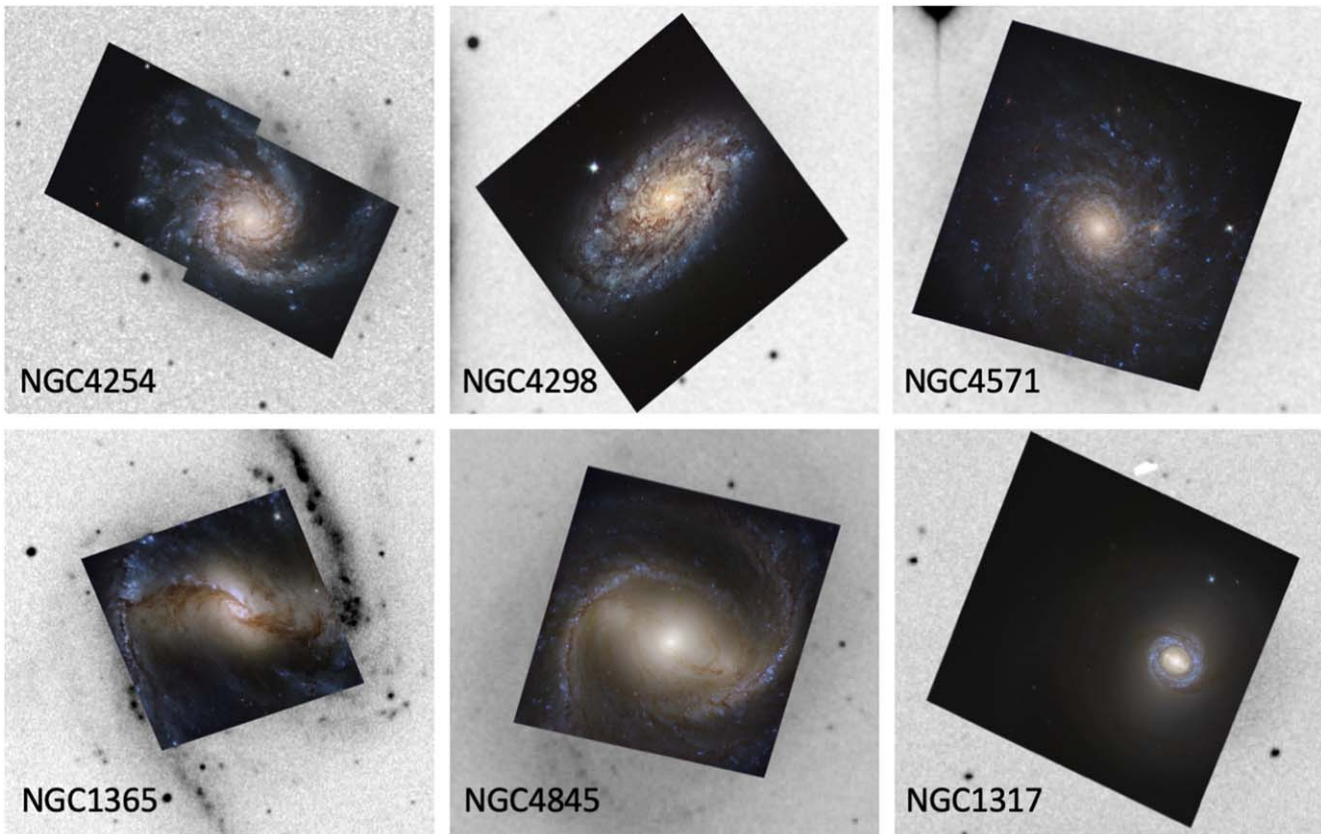
Figure 4. Color composites of PHANGS-HST imaging (Red: WFC3/UVIS F814W, green: WFC3/UVIS F555W, blue: WFC3/UVIS F438W+F336W+F275W), overlaid on DSS imaging for the same six galaxies as in Figures 2 and 3.

apertures with a 4 pixel radius, with the background determined in an annulus between 7 and 8 pixels around the aperture. An aperture correction (Section 4.3) is applied that yields the total fluxes for sources ultimately identified as compact star clusters. Photometry is also measured in a series of circular apertures (with radii from 1 to 5 pixels) to compute concentration indices that are used to distinguish star cluster candidates from stars and other sources/artifacts (Section 4.4). Catalogs with these observed parameters are produced for each galaxy, which on average contain about ∼500,000 sources. The contribution of nonredundant detections from `DAOStarFinder` to these catalogs is less than ∼1% of objects with a V-band aperture photometry S/N > 10, the limit subsequently applied to select star cluster candidates as discussed in Section 4.4.

Thilker et al. (2022) discuss in detail the parameter choices for both the `DOLPHOT` and `DAOStarFinder` routines, and the procedures for obtaining aperture photometry and computing errors.

4.3. Star Cluster Aperture Corrections

Star clusters, in particular those that are young, are generally found in crowded regions. Direct, accurate measurement of the total flux is often not possible since the outer light profile is frequently contaminated by other sources. Thus, HST photometry of compact star clusters in galaxies beyond the Local Group is typically measured with a limited aperture that captures ∼50% of the total flux, and then a correction, determined from bright, isolated clusters, is applied (e.g., Chandar et al. 2010b; Adamo et al. 2017; Cook et al. 2019).

Deger et al. (2021) present a detailed discussion of procedures used to determine average aperture corrections and uncertainties for each galaxy. In summary, the V-band images are visually inspected to identify a few dozen well-detected, isolated, compact clusters, and these objects are used to compute an average correction for each field. Fixed offsets, based on the change in the encircled energy distributions of the WFC3 PSF with wavelength (which has a minimum FWHM of $0\rlap{.}{''}067$ in the V band, and increases to ∼$0\rlap{.}{''}075$ in the NUV and I bands)[49,50] are used to calculate the corresponding corrections for photometry in the NUV, U, B, and I bands, since direct measurements from growth curves in those filters for bright sources in the V band that are very red or blue can be noisy (Cook et al. 2019; Deger et al. 2021). By construction, the resulting V-band corrections are ∼0.75 mag (for the adopted photometric aperture radius of 4 pixels, i.e., $0\rlap{.}{''}16$). The corrections are larger by 0.19, 0.12, 0.03, and 0.12 mag for the NUV, U, B, and I bands, respectively.

4.4. Star Cluster Candidate Selection

The PHANGS-HST pipeline identifies cluster candidates from the source lists described in Section 4.2 using photometric and morphological properties measured in the V band. In previous work, the concentration index (computed as the difference between photometry measured in circular apertures with radii of 1 and 3 pixels, CI_{13} or simply CI) has been
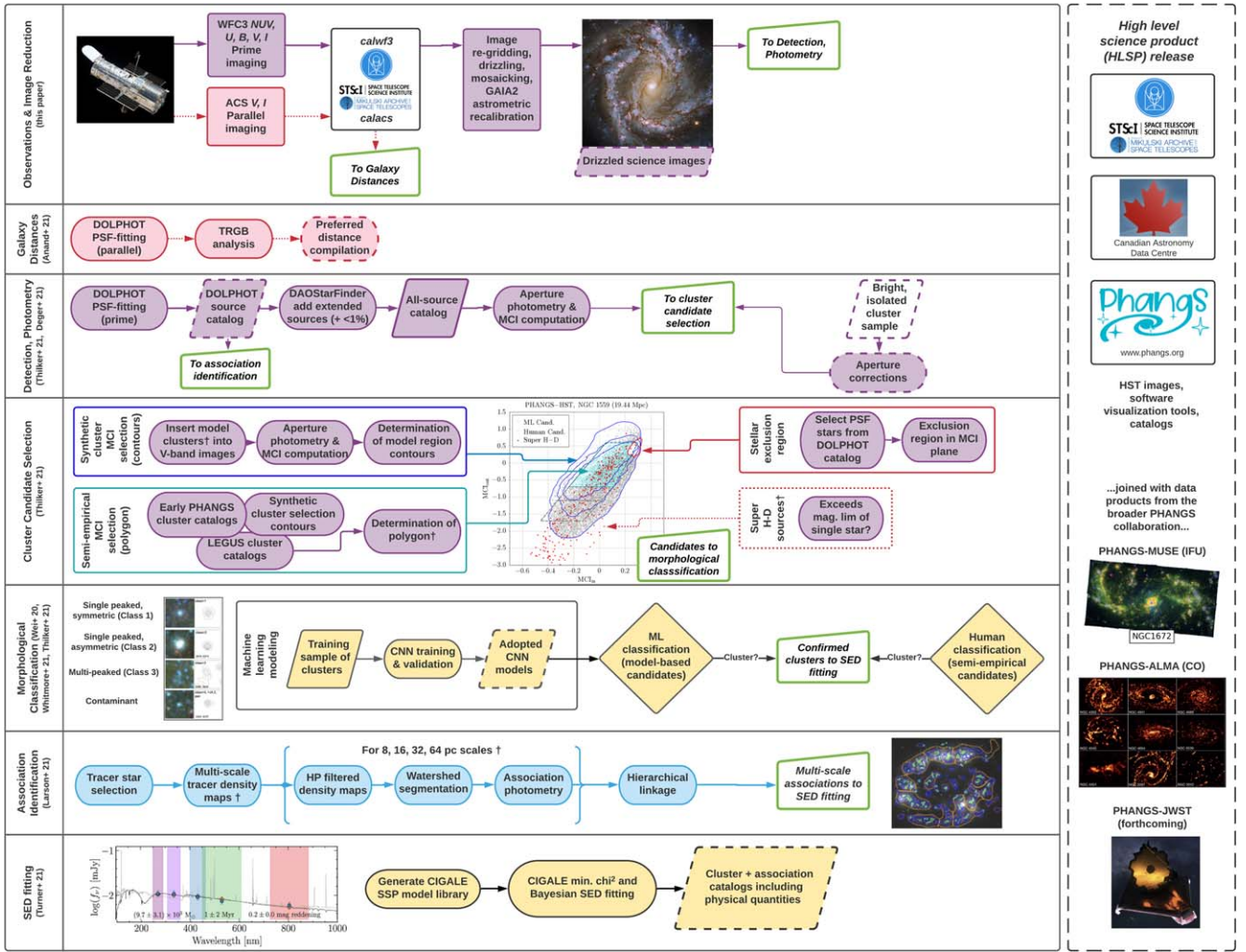

Figure 5. Flowchart illustrating the PHANGS-HST star cluster and multiscale stellar association catalog pipeline. The chart begins at the upper left corner with the acquisition of HST imaging and proceeds from left to right along the rows, which represent the major components of the pipeline, and concludes with the production of catalogs at the bottom right of the chart. Dashed lines indicate high-level science products for public release via MAST, CADC, and the PHANGS collaboration website https://sites.google.com/view/phangs/home. † indicates that the galaxy distance is used in this step.

generally used to remove sources likely to be stars from consideration (e.g., Whitmore et al. 2010; Adamo et al. 2017; Cook et al. 2019). To determine the threshold to separate stars from cluster candidates, CIs have been typically measured for a few dozen isolated, bright point sources in an image. These measurements are then compared with the CI distribution for an analogous sample of compact star clusters (identified through visual inspection and also used to derive aperture corrections).

For PHANGS-HST, Thilker et al. (2022) build upon this method to develop candidate selection criteria based on multiple concentration indices (MCI), rather than a single concentration index. We measure fluxes in circular apertures with radii of 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, and 5.0 pixels, and we define two metrics to characterize the light profiles between 1 and 2.5 pixels ($\mathrm{MCI_{in}}$) and 2.5–5 pixels ($\mathrm{MCI_{out}}$), where

$$\mathrm{MCI} \equiv \frac{1}{3}(\mathrm{NCI}_{ab} + \mathrm{NCI}_{bc} + \mathrm{NCI}_{cd}), \quad (1)$$

where a, b, c, and d represent the radii of the apertures, and

$$\mathrm{NCI}_{ij} \equiv 1 - \frac{\mathrm{CI}_{ij}}{\mathrm{CI}_{ij,\mathrm{fiducial}}} \quad (2)$$

is a CI normalized by $\mathrm{CI}_{ij,\mathrm{fiducial}}$, defined to be the CI of a relatively compact cluster; i.e., a PSF-convolved Moffat/EFF87 function (Moffat 1969; Elson et al. 1987, EFF87) with a FWHM of 2 pixels, and a power-law slope of 3 describing the surface brightness profile of the extended halo. The choice of normalization is arbitrary, and enables the various CIs to be meaningfully averaged. With these definitions, clusters measured on HST optical images generally have values of $-0.5 \lesssim \mathrm{MCI_{in}} \lesssim 0.2$, $-2 \lesssim \mathrm{MCI_{out}} \lesssim 0.6$ (see Figures 7 and 6). $\mathrm{MCI_{in}}$ is anticorrelated with the standard CI_{13}, as would be expected from their definitions; i.e., more compact sources are characterized by larger MCI values (i.e., more concentrated), which is the opposite of the sense of the standard CI_{13}. By construction, the fiducial cluster lies at the origin of the $\mathrm{MCI_{in}}$–$\mathrm{MCI_{out}}$ plane.

We define selection regions in two ways on the $\mathrm{MCI_{in}}$–$\mathrm{MCI_{out}}$ plane (Figure 6) to generate two types of candidate lists. (1) The first set of selection regions are contours in the MCI plane, based on the loci of synthetic star clusters inserted into the V-band imaging for each individual target galaxy. This selection strategy yields up to several thousand candidates per
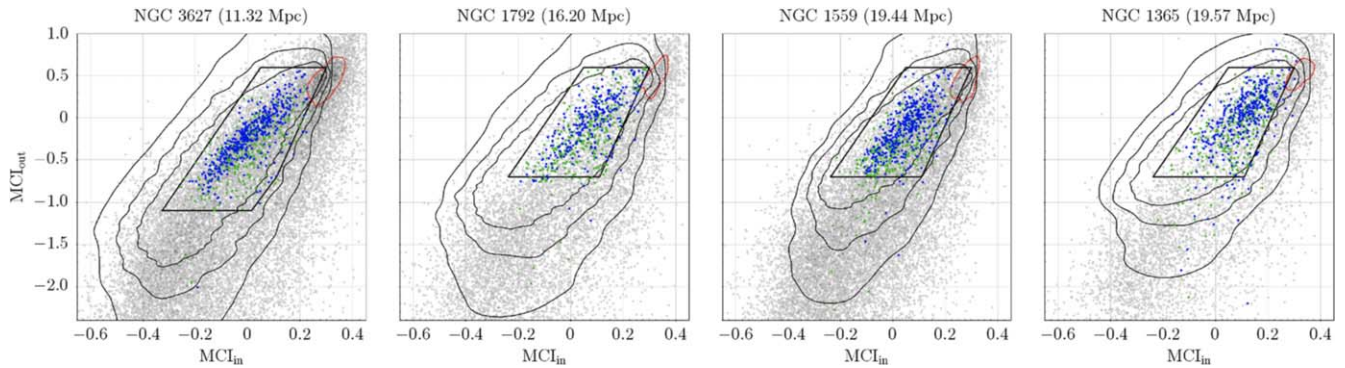
Figure 6. Star cluster candidate selection regions defined in the multiple concentration index plane (MCI; see Equation (1)) based on synthetic star clusters (large contours) for four galaxies at increasing distances from left to right (NGC 3627, NGC 1792, NGC 1559, and NGC 1365). Selection regions are also defined semi-empirically (polygon; see also Figure 7). Candidates are selected from detections that satisfy basic signal-to-noise ratio criteria (gray), and exclude those in the stellar (point-source dominated) region (small red contours). Candidates within the empirical selection region are slated for visual inspection to a V-band magnitude limit of ~24 mag. The classification of the much larger samples of candidates identified with outermost model contours is automated using convolutional neural network models. In all panels, visually classified class 1 (blue) and class 2 (green) clusters, are shown. Some clusters appear outside the polygon—these result from ad hoc human inspection to confirm that the density of clusters rapidly declines outside this selection region, as well as from the inspection of sources brighter than the Humphreys–Davidson limit.

target, and CNN models are used for their classification. (2) The second are smaller polygonal regions, the same set of which are used for all galaxies. Definition of these polygons is based on the loci of human inspected and verified clusters, together with the inner synthetic star cluster contours (which encloses the highest concentration of models) for the first few PHANGS-HST galaxies studied. This second selection produces smaller samples, more suitable for time-intensive, expert human classification; it is intended to yield an average of ~1000 candidates per field, so that inspection over the full set of 38 PHANGS-HST galaxies is manageable. Once defined, it provides a simpler, faster way of selecting cluster candidates for inspection by bypassing the synthetic star cluster analysis required to generate contours for a given galaxy. More specifically:

1. Selection criteria in the MCI plane are defined based on model star clusters. Synthetic clusters are generated with Moffat profiles, which again are parameterized by the effective radius (in units of the intrinsic/pre-PSF-convolution FWHM), and the power-law slope describing the surface brightness profile of the extended halo. The modeling includes 216 distinct Moffat profiles, spanning 0.5 pix $\leqslant$ FWHM $\leqslant 7$ pix and power-law slopes from 0.75 to 4. Models are generated for each of these profile types with a distribution of ~4000 apparent magnitudes. The magnitudes are computed from the distance of the galaxy and V-band luminosities based on solar metallicity, single-aged stellar population models of Bruzual & Charlot (2003), for a grid of masses (10^3–10^5 $M_\odot$), ages (1–1000 Myr), and extinctions ($0 \leqslant A_V \leqslant 0.5$ mag). The synthetic clusters are randomly inserted into the V-band images 200 at a time, and aperture photometry is performed with the same procedure used to measure real sources. In total, $\sim4 \times 10^6$ clusters are inserted. The MCI values from these synthetic clusters are plotted on the MCI plane to form a 2D histogram with a bin size of 0.01, and a contour is derived. Contours are also generated for MCI bin sizes of 0.02 and 0.04, resulting in three nested model selection regions for each field (Figure 6). Typically, the inner model region contains the majority of candidates (~50%–70%), while the two

larger contours each add comparable numbers of candidates (20%–30%). Currently, the outer contour (i.e., corresponding to the largest bin size) is used to produce candidate lists for automated classification by CNNs (see Section 4.5).

2. Selection criteria in the MCI plane are also defined semi-empirically. For the first few PHANGS-HST galaxies studied (e.g., NGC 1559, Wei et al. 2020), sources were inspected over a wide swath of the MCI plane (essentially spanning the outer model contour). The loci of human verified clusters were used in combination with the tightest model contour to define the left, right, and top edges of the polygon. Catalogs of visually inspected clusters published by the HST LEGUS program for 34 HST fields in 30 galaxies[51] were then used to verify that the majority of LEGUS clusters would be captured by this selection (Figure 7). The lower bound of MCI_{out} increases as the distance of the host galaxies increases—clusters will be less resolved and appear more compact at larger distances, and both the model contours and the loci of the human verified cluster populations do indeed shrink with increasing distance. Using the ensemble of cluster populations identified in the first few PHANGS-HST galaxies studied, plus those in the LEGUS catalogs, we adopt limits of $MCI_{out} = -1.7$, -1.1, and -0.7 for galaxies with distances $\leqslant 8$ Mpc, 8 Mpc $< d < 14$ Mpc, and $\geqslant 14$ Mpc, respectively (Figure 7). This initial analysis over a wide swath of the MCI plane for PHANGS-HST sources suggests that the density of clusters rapidly drops beyond the boundaries of the polygon, but a more careful investigation of such completeness issues will be the subject of future work. These semi-empirical polygon selection regions are used to produce candidate lists for human inspection and classification (see Section 4.5 and Whitmore et al. 2021) to a total V-band limit of ~24 mag (corresponding to absolute V-band limits between about -5.5 mag and -8 mag for the distances of the galaxies in the sample). The exact value of the magnitude limit depends on the number of candidates, since, again, the primary purpose
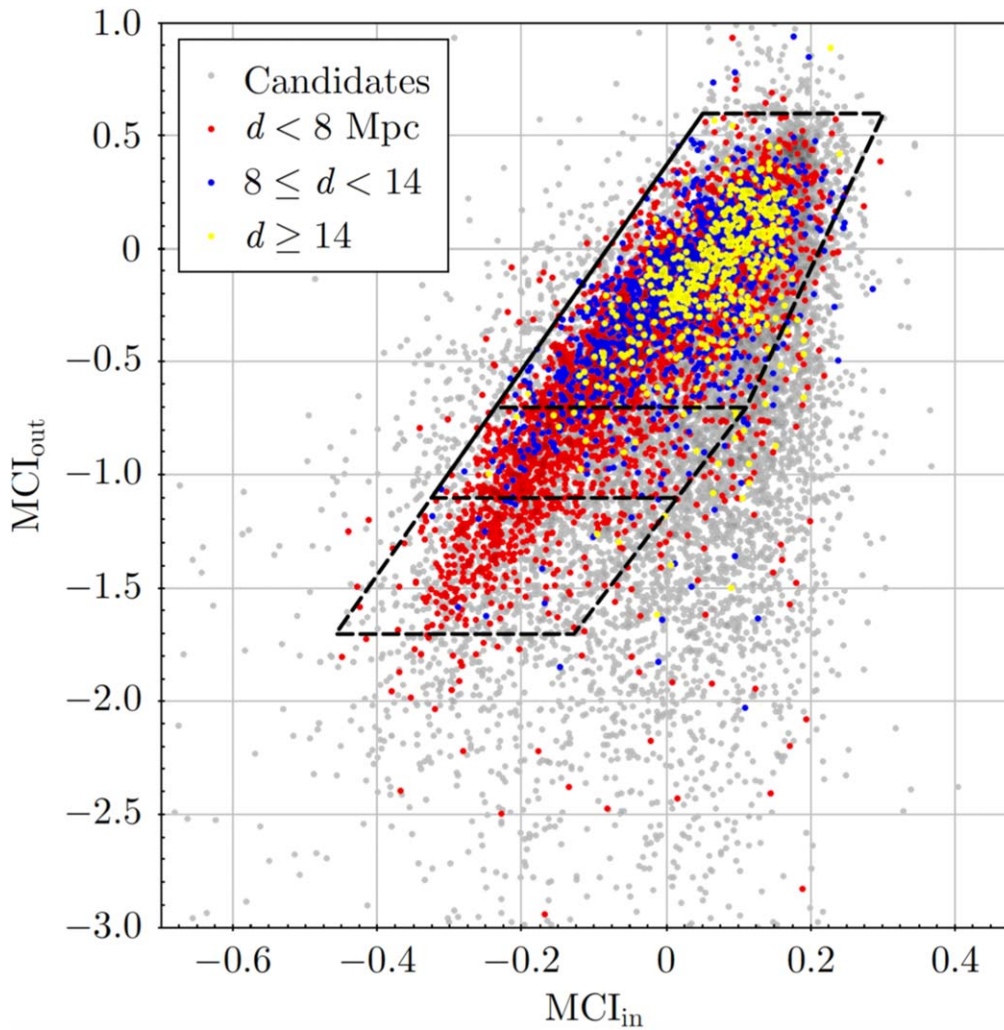
---

[51] https://archive.stsci.edu/prepds/legus/dataproducts-public.html

Figure 7. Semi-empirically defined selection regions for compact star cluster candidates in the multiple concentration index (MCI; see Equation (1)) plane. This selection is devised to produce candidate sample sizes amenable for human inspection, while maximizing the yield of true clusters. We verify that the selection region includes the majority of visually verified class 1 and 2 clusters from 30 galaxies studied by the LEGUS program (Calzetti et al. 2015a; Adamo et al. 2017, colored points; https://archive.stsci.edu/prepds/legus/dataproducts-public.html). The larger set of candidates from LEGUS are also shown (gray points). Clusters are color coded based on the distance of their host galaxy. The lower bound of MCI$_{out}$ increases as the distance of the host galaxy increases (see also Figure 6), since the physical resolution decreases and clusters appear more compact. We adopt different lower bounds for the three different distance ranges indicated in the legend.

of this second selection is to reduce the number of candidates for visual inspection to a manageable level (i.e., ≲1500 per galaxy) while maximizing the yield of true clusters. So, if the total number of candidates is low (several hundred) then the limit is fainter, whereas if the number of candidates is large, the limit is somewhat brighter.

For both model-based (contours) and semi-empirical (polygon) selection methods:

1. The regions in the MCI plane dominated by stars (point sources) as detected by DOLPHOT, are defined for each field and used to exclude objects. Figure 6 shows both the synthetic and semi-empirical cluster selection regions for four galaxies at a range of distances together with their stellar exclusion regions.
2. Candidates must also satisfy basic criteria: V-band photometry measured in a 4 pixel radius must have S/N ≥ 10, and the source must also be detected in at least two other bands with photometric error ≤0.3 mag. The

faintest sources in the resulting candidate lists for fields with the standard V-band exposure time of 670 s have total V-band magnitudes of ∼24.6, which corresponds to absolute magnitudes between −4.2 and −7.8 for the distance range spanned by the PHANGS-HST galaxies.
3. All sources with total, absolute V-band magnitudes brighter than −10 mag and $-0.55 \leqslant \mathrm{MCI_{in}} \leqslant 0.45$ (i.e., MCI$_{in}$ values plausible for real objects) are kept as candidates to help ensure high completeness for the brightest clusters. Sources brighter than −10 mag exceed the Humphreys–Davidson (HD) limit, the observed maximum luminosity of individual stars in the LMC, thought to be due to a modified Eddington Limit (Humphreys & Davidson 1979; Lamers & Levesque 2017). These "super HD sources" also undergo classification to remove interlopers such as saturated stars and background galaxies.

The number of candidates identified using the empirical MCI selection regions varies from many hundred to several thousand sources for each field, with a median of ∼1000 (Figure 8), and
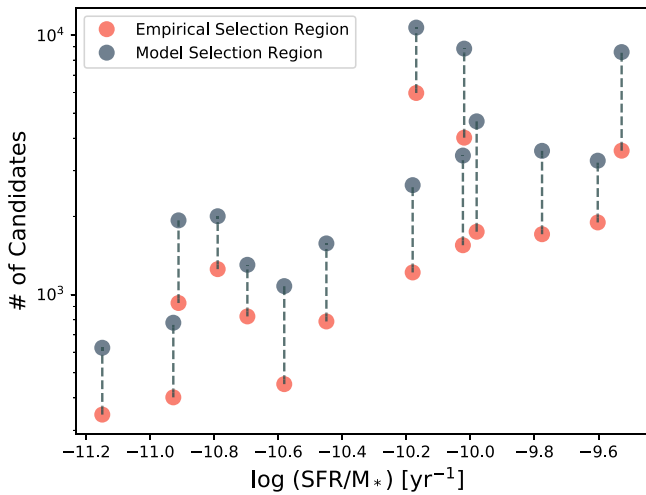
Figure 8. The number of compact star cluster candidates found in the first 15 galaxies processed through the PHANGS-HST pipeline, shown as a function of the sSFR. Candidates identified using the semi-empirical selection (polygon region) are shown in red, while those resulting from the larger model-based selection contours are shown in gray. The cluster candidates undergo a process of inspection to further remove contaminants and sort the objects into three morphological categories as summarized in Section 4.5, and described in detail in Wei et al. (2020) and Whitmore et al. (2021). The smaller samples resulting from the semi-empirical (polygon) selection are designed for human inspection, while the larger samples from the model contours are fed to convolutional neural networks for automated classification.

should ultimately yield sample sizes of compact star clusters similar to those studied in previous work. The variation in candidate sample sizes reflects the variation in global sSFRs for the PHANGS-HST galaxies that span a factor of $\sim$10 (Figure 1). The synthetic cluster MCI selection casts a wider net for potential clusters (Figure 6) and yields candidate samples about a factor two larger than the empirical selection (Figure 8). These larger samples enable analysis of potential incompleteness in previous star cluster studies, in particular for more diffuse clusters that appear to be rare.

These new selection methods, based on the measurement of multiple CI and the use of model star clusters, provides a solid foundation for quantitative investigation of structural properties including: which model clusters actually exist in nature, whether certain clusters are likely to be bound or unbound, and how their morphologies evolve with time. Our model grid of clusters also facilitates future work to characterize cluster completeness levels (e.g., Johnson et al. 2015). More discussion on the utility of the approach to cluster selection is provided in Thilker et al. (2022).

4.5. Star Cluster Candidate Inspection and Morphological Classification

The candidate star clusters then undergo a process of inspection to further remove contaminants and sort the clusters into different morphological categories. A combination of human visual inspection and automated inspection with CNN models is performed as follows:

1. The smaller lists of candidates resulting from the semi-empirical MCI selection region (polygons) are visually inspected by co-author B.C.W. Objects with a total V-band magnitude to a limit of $\sim$24 mag receive visual classifications.

2. The samples of candidates identified in the largest synthetic cluster MCI selection region (contours) are classified by CNN models, as described below. Candidates identified using the semi-empirical selection region generally lie within this synthetic cluster selection area (Figure 6), and hence will have both neural network and human visual classifications.

3. Human visual inspection of fainter candidates and those beyond the boundaries of the empirical selection region are performed on an ad hoc basis to evaluate and monitor the performance of the neural network models.

For PHANGS-HST, we adopt the general classification scheme used by LEGUS as described in Adamo et al. (2017) and Cook et al. (2019):

1. Class 1: compact star cluster—single peak, circularly symmetric, but radial profile more extended relative to point source.
2. Class 2: compact star cluster—similar to Class 1, but elongated or asymmetric.
3. Class 3: compact stellar association—asymmetric, multiple peaks.
4. Class 4: not a compact star cluster or compact stellar association (e.g., image artifacts, background galaxies, individual stars or pairs of stars).

Examples of Class 1, 2, and 3 objects are shown in Figure 9. Whitmore et al. (2021) provide a detailed description of the process of visual inspection and morphological classification, and discuss differences in the application of this scheme for Class 3 objects relative to the LEGUS project (i.e., we require evidence of four or more peaks within a radius of 5 pixels in PHANGS-HST). A brief history of star cluster classification is also given there (also see Wei et al. 2020, Section 2). While we continue to include Class 3 objects in our compact cluster catalogs, we note that this is mainly for historical continuity with the LEGUS project. The PHANGS-HST pipeline is optimized to identify single-peaked compact clusters, and this leads to a high level of incompleteness for multipeaked stellar associations. Instead, we introduce a new identification process for stellar associations, based on a watershed algorithm, which provides a far more complete inventory of young stellar populations and the star formation hierarchy at multiple physical scales, as will be summarized in Section 4.7 and presented in K. Larson et al. (2021, in preparation).

We note that it is debated whether such classifications for a given object can distinguish between gravitationally bound clusters and unbound associations, which may form and evolve under distinct conditions (Gieles & Portegies Zwart 2011; Kruijssen 2012; Ryon et al. 2017; Krumholz et al. 2019; Ward et al. 2020). Statistically however, it is expected that Class 1 should contain the highest percentage of bound clusters, and Class 3 should have the highest percentage of unbound associations (see Whitmore et al. 2021, for further discussion).

Ultimately, PHANGS-HST will generate up to $\sim$80,000 star cluster candidates for inspection and classification. In previous large studies of star clusters, the process of visual inspection has been a limiting step, which motivated the investigation of automated machine-learning techniques (Messa et al. 2018; Grasha et al. 2019; Perez et al. 2021). In Wei et al. (2020), we studied the application of deep transfer learning techniques to train CNN to classify star cluster candidates according to the scheme above. Deep transfer learning involves the tuning of a
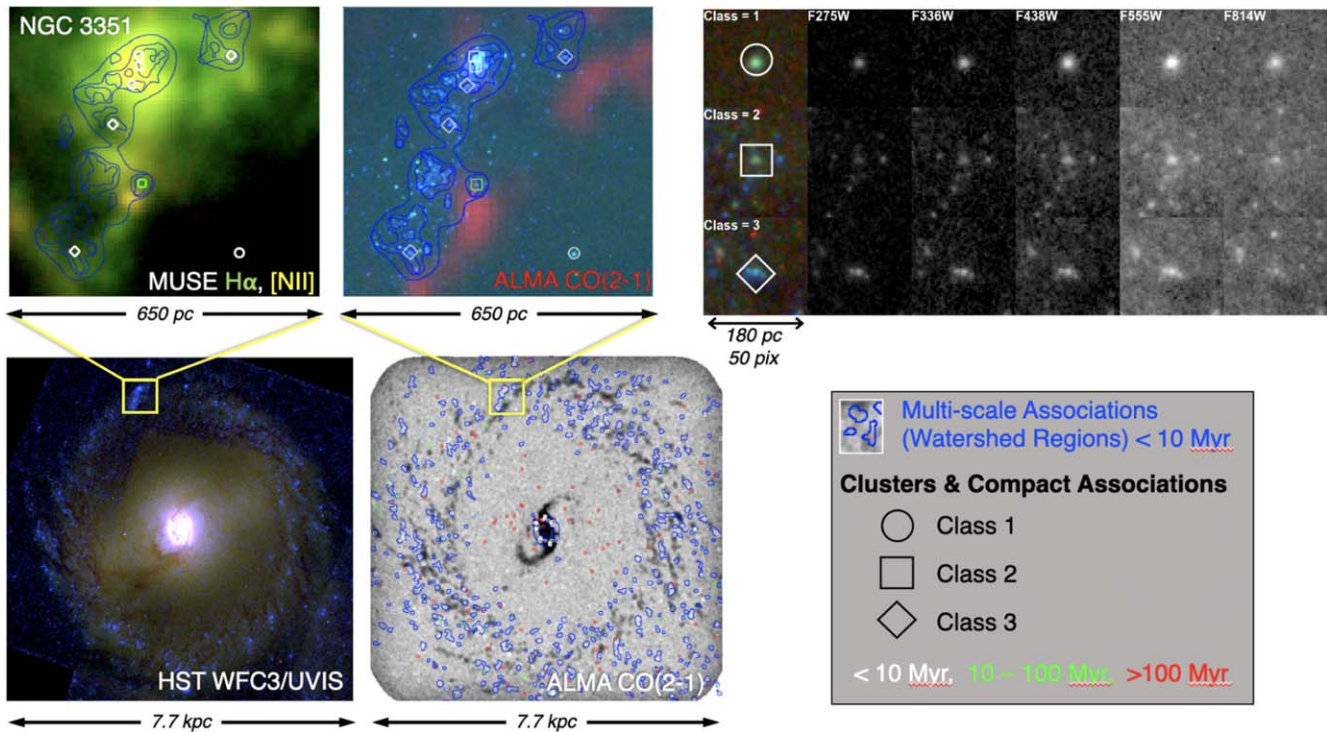
Figure 9. Structures across the physical scales of the star formation hierarchy in NGC 3351, identified by the PHANGS-HST pipeline, from single-peaked compact star clusters, the densest structures, to larger scale multipeaked stellar associations. Bottom left: color composite of WFC3/UVIS F275W+F336W (blue), WFC3/UVIS F435W+F555W (green), and WFC3/UVIS F814W (red). Bottom right: young stellar associations (<10 Myr) traced by the watershed-based method of K. Larson et al. (2021, in preparation, blue contours), together with all compact clusters and associations with human visual classifications (Class 1: circles, Class 2: squares, Class 3: diamonds; color coded by age as indicated), overlaid on the PHANGS-ALMA CO(2−1) map. A 650 pc section of the outer ring (yellow box) is shown in more detail in the top left and middle panels. All three classes of compact clusters and associations are represented in the selected section, and the magnified view allows all four levels traced by the watershed method (64 pc, 32 pc, 16 pc, and 8 pc) to be clearly shown. Top left: magnified view using an Hα map constructed from the VLT/MUSE IFU data cube. Top middle: magnified view using a color-composite image where CO is now shown in red. Top right: further magnification of 180 pc areas centered on examples of the three classes of compact clusters and associations found in the selected 650 pc section of the outer ring in all PHANGS-HST filters.

pretrained network, for example, based on the ImageNet library of everyday objects.[52] In principle, this approach enables CNNs to be successfully trained with relatively small samples (i.e., hundreds to a thousand images), which is the current size of HST star cluster samples with visual classifications. It provides an alternative to the process of training all network layers from scratch, which requires samples that are an order of magnitude larger. The results of Wei et al. (2020) were encouraging, as the prediction accuracies (70%, 40%, 40%–50%, and 50%–70% for Class 1, 2, and 3 star clusters and Class 4 nonclusters, respectively) were found to be competitive with the classification consistency between different human classifiers. The neural network models presented in Wei et al. (2020) provide a starting point for automated classification of the PHANGS-HST and other HST star cluster candidate samples, which can continue to be optimized. Whitmore et al. (2021) present results of the current Wei et al. (2020) models applied to clusters candidates in five PHANGS-HST galaxies, which also have classifications published by the LEGUS project. The Whitmore et al. (2021) analysis includes a detailed comparison between human and automated classifications, from overall prediction accuracy to differences in the distribution of ages, *UBVI* color–color diagrams, and stellar mass functions, and also includes a discussion of additional future work to improve performance.

4.6. SED Fitting

To derive ages, masses, and reddenings for the sources classified as star clusters and associations, we use a modified version of CIGALE[53] (Code Investigating GALaxy Emission; Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019), a publicly available SED-fitting package developed for galaxies. Our modifications, which support the fitting of single-age populations and provide modeling options to facilitate comparison to prior SED modeling results for star clusters, are available in dedicated branches, `SSP` and `SSPmag` respectively, of the public `git` repository[54] of CIGALE. Turner et al. (2021) reports on these modifications and the analyses performed to validate the code.

SED fitting with CIGALE is performed on the five-band photometry (NUV–*U*–*B*–*V*–*I*) for both compact star clusters and stellar associations. The fitting is based on the simple (single-aged) population synthesis models of Bruzual & Charlot (2003), assuming solar metallicity and a Chabrier (2003) IMF (with standard mass limits of 0.1–100 $M_{\odot}$), and no addition of nebular emission. The Cardelli et al. (1989) extinction curve with $R_V = 3.1$ is used and a maximum $E(B − V) = 1.5$ mag is imposed. The reasoning for these choices is discussed in Turner et al. (2021).

---

[52] http://www.image-net.org/

[53] https://cigale.lam.fr
[54] https://gitlab.lam.fr/cigale/cigale.git

When CIGALE is run with the same assumptions and theoretical models as used in LEGUS, Turner et al. (2021) find very similar cluster ages and masses. Quantitatively, fits to identical cluster photometry yield (logarithmic) medians in the ratios of ages and masses for the two surveys of 0.001 ± 0.017 dex and 0.003 ± 0.011 dex, respectively.

4.7. Stellar Association Identification

The majority of star formation occurs in stellar associations (Lada & Lada 2003; Ward & Kruijssen 2018; Ward et al. 2020; Wright 2020, and references therein). Compact star clusters, the focus of the previous sections, are formed only in the densest peaks of the star formation hierarchy (Elmegreen 2008; Kruijssen 2012) and contain between $1\% \sim 50\%$ of the total star formation in galaxies (Kruijssen 2012; Adamo et al. 2015, 2020; Johnson et al. 2016; Chandar et al. 2017; Krumholz et al. 2019). To produce catalogs of stellar associations, methods distinct from those used to identify single-peaked compact clusters are needed to segment the light distribution over larger physical scales and to probe further into the star formation hierarchy. Development of such methods is particularly important for obtaining complete inventories of the youngest stellar populations ($\lesssim 10$ Myr), which are a requirement for a robust joint analysis with molecular clouds and H II regions.

For PHANGS-HST, K. Larson et al. (2021, in preparation) develop a technique to produce catalogs of stellar associations spanning scales from 8 to 64 pc. The technique builds upon the watershed routine in the scikit-image Python package (SKIMAGE.SEGMENTATION.WATERSHED, van der Walt et al. 2014), which is based on the concept of geological watersheds. The routine identifies regions by "flooding" an image, given a set of markers as the starting points. In addition to the image on which associations are to be identified (i.e., requiring segmentation), the inputs needed by WATERSHED are a list of marker positions and an image mask that defines the areas over which regions are allowed to grow.

Our adopted technique deploys WATERSHED on a smoothed, filtered map of the positions of point sources (rather than directly on the HST images) and uses a two-parameter procedure to determine the marker positions and to produce the image mask from these position maps. The smoothed, filtered, position maps are produced as follows. Point sources are selected from the DOLPHOT catalogs that satisfy basic requirements on signal-to-noise ratio, sharpness, and data quality. The positions of these point sources are used to create maps with the same pixel grid as the PHANGS-HST DRC images, containing values of 1's corresponding to the DOLPHOT positions, and 0's otherwise. The maps are then smoothed with Gaussian profiles with FWHM of 2^n pc for $n = 3, 4, 5,$ and 6 (i.e., 8, 16, 32, and 64 pc), computed given the distance of the galaxy. Finally, a high-pass filter is applied by subtracting a map that has been smoothed with a kernel that is four times larger.

Two parameters are then defined that are tied to the characteristics of a single object on these smoothed, filtered position maps. The peak threshold parameter is the level above which markers (local maxima) are identified; it is currently defined to be 1.5 times the maximum value for a single object so that the resulting regions are multipeaked. The edge threshold parameter is the minimum surface "brightness" level beyond which the regions are not allowed to expand further,

and used to create a mask image for WATERSHED; it is defined to be the surface "brightness" at the FWHM of a single object.

The smoothed, filtered, position maps enable the identification of structures on the physical scales over which the maps have been smoothed. The 8 pc smoothed maps allow associations that overlap in size with sources in the compact cluster catalog to be studied, and the maps that are smoothed over larger scales enable the greater star formation hierarchy to be traced. Structures are identified based on the NUV and V-band DOLPHOT point-source catalogs. The resulting NUV-band selected association catalogs will predominantly contain young structures ($\lesssim 100$ Myr). In comparison, V-band selected catalogs will include structures over a larger range of ages and will facilitate comparison with the compact cluster samples, which have also been V-band selected.

K. Larson et al. (2021, in preparation) demonstrate the validity of this technique for identification of stellar associations in the PHANGS-HST galaxy sample based on an analysis of NGC 3351 and NGC 1566.

Figure 9 shows the watershed-identified stellar associations in NGC 3351 together with visually classified objects from the compact star cluster pipeline (Section 4.4). Three different magnifications of the galaxy are shown to illustrate the full disk and star-forming ring (bottom panels), the multiscale associations traced by the watershed method (top left and middle panels), and individual compact star clusters and associations (series of postage stamps at top right). The structures are overlaid on color composites of the HST imaging, ALMA CO(2–1) map, and a color composite of the Hα and [N II] $\lambda 6583$ maps from MUSE. Only the youngest (<10 Myr) watershed associations are shown to illustrate the correlation with the blue starlight, molecular clouds, and H II regions. Ages are derived using CIGALE as summarized in Section 4.6 and documented in Turner et al. (2021). Details on the procedures used to measure fluxes in the regions are presented in K. Larson et al. (2021, in preparation).

From examination of the properties of the resultant watershed structures, we find:

1. Sample sizes of several hundred up to a few thousand associations in each galaxy. For associations identified on physical scales comparable to the aperture sizes used for selecting compact clusters candidates (a 4 pixel radius, which corresponds to 8 pc at a distance of 10 Mpc; Section 4.4), the numbers of associations identified are about $\sim$2–4 times larger than the numbers of visually classified compact clusters.

2. The process is shown to successfully identify structures at the defined scale. The size distributions are well defined, and approximately log-normal with medians near the FWHM of the smoothing kernel.

3. Fluxes computed within the boundaries of regions identified on the 8 and 16 pc smoothed images yield colors that are consistent with single-aged stellar population tracks on the $UBVI$ color–color diagram. The loci are similar to previously studied samples of compact star cluster and associations (Figure 10). The similarity between two panels of Figure 10 demonstrates that these structures, which have been identified with a somewhat complicated algorithm, do behave as one might expect for groups of stars that are physically associated and born close in time. The associations may not strictly be singled-aged populations, but rather composite
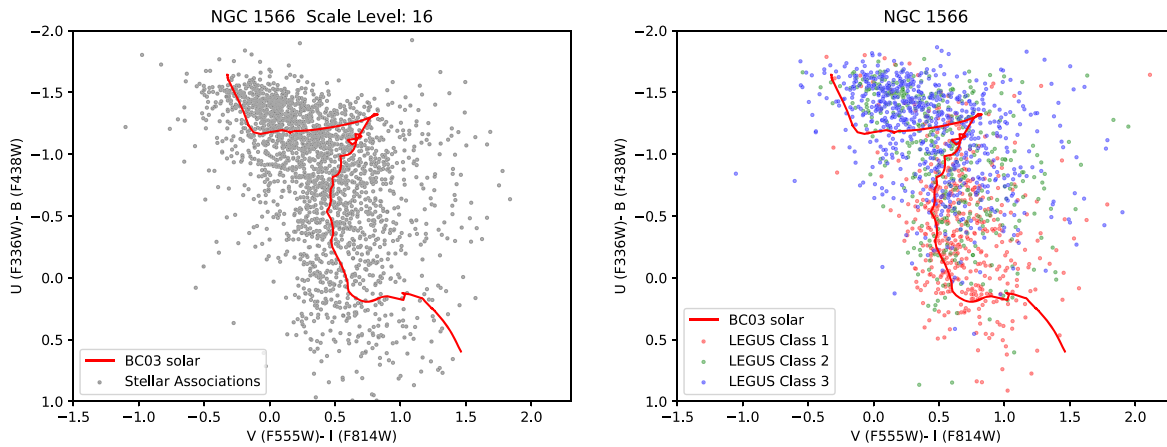
Figure 10. Comparison of *UBVI* color–color diagrams for multipeak stellar associations identified with the PHANGS-HST watershed-based procedure (left panel), and compact star clusters and associations from the LEGUS program (right panel). The stellar associations are based on a *V*-band map of point-source positions smoothed with a 16 pc FWHM Gaussian kernel.
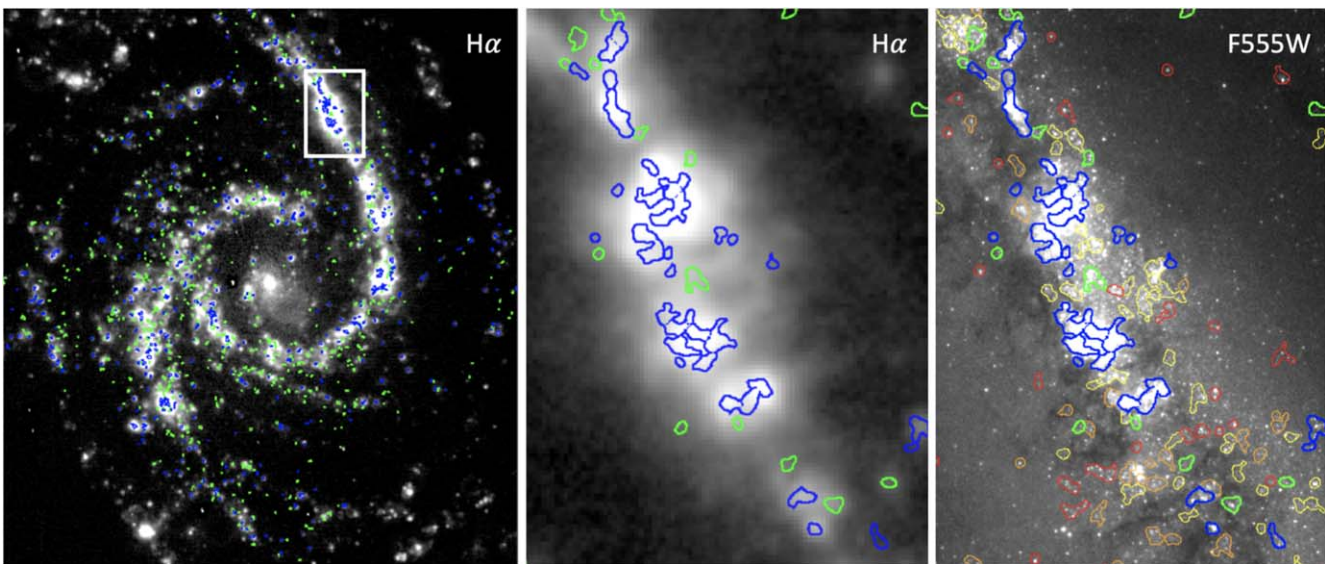

Figure 11. Stellar associations in NGC 1566 color coded by SED-fit age. Associations were identified based on a map of F555W point-source positions smoothed with a 32 pc FWHM Gaussian kernel, and SED fitting was performed with CIGALE assuming a single-aged stellar population. Left: the youngest associations ($\leqslant$3 Myr: blue, 3–5 Myr: green) overlaid on PHANGS ground-based Hα narrowband imaging (A. Razza et al. 2021, in preparation). The HST data do not extend to the corners of the imaging, and associations have not been traced in those areas. Middle: same as the left panel, but expanded to show greater detail in a portion of the northern spiral arm. Right: the same region as the middle panel, but now overlaid on the F555W (*V*-band) imaging and including older associations (5–10 Myr: yellow, 10–60 Myr: orange, >60 Myr: red). Age ranges are exclusive of the lower bound and inclusive of the upper bound.

populations that can be approximated by a single-aged model. The level of deviation from a single-aged population, particularly as a function of the size of the region, is an issue that will be the topic of future work. Such work will be facilitated by the PHANGS-HST pipelines use of CIGALE, which allows for the self-consistent modeling of both single-aged and composite stellar populations as discussed in Turner et al. (2021).

4. Maps of the youngest associations (<5 Myr), show excellent correspondence to H II regions observed in narrowband Hα imaging. As the age of the regions increases, they become more anticorrelated with the H II regions, as would be expected (Figure 11). This provides some evidence that the age-dating of the youngest structures appears to produce reasonable results, despite the simplifying assumption of a single-aged population in the SED modeling and the complexities of the overall process to define the regions and measure their photometry.

Initial testing of our watershed-based methods for identifying stellar association have been performed with NGC 3351 and NGC 1566. These two galaxies were chosen as they span a significant range of distances in the sample (from ~10–18 Mpc), and because catalogs of clusters and compact associations are available from LEGUS for comparison. As with our new methods for identifying clusters, the completeness of our samples of multiscale associations, particularly possible systematics as a function of distance, will need to be studied with simulations.

5. Data Products

The PHANGS-HST data set will enable science extending well beyond the primary goals of the PHANGS collaboration.

To enable the research community to make full use of the PHANGS-HST data, high-level science products from our star cluster and association catalog pipeline are being released. The following will be available through the PHANGS homepage at Mikulski Archive for Space Telescopes (MAST)[55] with doi:10.17909/t9-r08f-dq31.

5.1. Imaging

For imaging of the star-forming disk in NUV–U–B–V–I bands for the 38 PHANGS-HST galaxies:

1. FLC FITS file for each exposure with astrometric solutions updated based on GAIA DR2 sources.
2. Combined DRC FITS images of individual pointings in each filter, each drizzled onto a common pixel grid defined for the galaxy target, also with astrometric solutions based on calibration to GAIA DR2 sources.
3. Mosaicked DRC FITS images in each filter for 13 galaxies covered by multiple pointings (NGC 628, NGC 1097, NGC 1300, NGC 1512, NGC 1672, NGC 2903, NGC 3351, NGC 3621, NGC 3627, NGC 4254, NGC 4321, NGC 4536, an NGC 6744).
4. ERR and EXP weight FITS images for individual pointings as well as mosaics.

5.2. Catalogs

1. `DOLPHOT` catalogs with five-band PSF-fitting photometry.
2. Compact star cluster and stellar association candidate catalogs, including position, five-band aperture photometry, stellar mass, age, reddening, CNN morphological classification, visual classifications for a subset of candidates in the empirical selection region, multiple concentration indices (MCI), and standard concentration index (CI) values.
3. Catalogs of stellar associations detected at 8, 16, 32, and 64 pc scales, including five-band region photometry, stellar mass, age, reddening, and effective radius, together with DS9 region files providing peak position and boundaries of regions, and FITS masks of regions.

5.3. Software

1. The Python routines that constitute the PHANGS-HST compact star cluster and association pipeline will be released at https://github.com/PhangsTeam.
2. CIGALE augmentations for SED fitting of single-aged stellar populations are available in dedicated branches, `SSP` and `SSPmag`, respectively, of the public `git` repository[56] of CIGALE.
3. CNN models for cluster candidate classification as described in Wei et al. (2020) and Whitmore et al. (2021). An annotated Python notebook containing scripts to run the models will be provided. Future updates shown to be improvements over the current models will be released as they are developed.

5.4. ALMA CO and MUSE Data Products

PHANGS-ALMA data have been released for the full PHANGS parent sample through the ALMA Archive and the Canadian Astronomy Data Centre (CADC[57]), and also linked to the PHANGS portal at MAST. The PHANGS-ALMA products include the $^{12}CO(2–1)$ spectral line data cubes, signal masks, and derived products such as the integrated intensity, line-of-sight velocity estimate, and spectral line widths. A description of the data reduction, products, and release is provided in Leroy et al. (2021a).

Likewise, PHANGS-MUSE integral field spectrograph data of the 19 galaxies targeted in the course of VLT Large Programme (ESO 1100.B-0651) has been released via the ESO Science Archive[58] and CADC, and also linked to the PHANGS portal at MAST. The released MUSE data include reduced and fully mosaicked data cubes as well as a series of two-dimensional maps associated with the gas and stellar tracers: broadband reconstructed images, emission line distribution and kinematics, stellar kinematics, star formation histories (mass and light-weighted age and metallicity maps), extinction maps from Balmer decrement, and stellar continuum fitting. The details of the data reduction and analysis processes are provided in Emsellem et al. (2021).

Links to the archive locations for all released PHANGS products are available at the survey webpage (http://phangs.org/data).

6. Summary

For decades, investigations of extragalactic molecular clouds and young resolved stellar populations have proceeded independently, and an integrated analysis has been performed only for case studies of select nearby galaxies. With the transformative capabilities of ALMA and HST working in concert, PHANGS will help bridge the fields of star formation and galaxy evolution by investigating how small-scale physics, which creates the basic quanta of star formation, may depend on the physical conditions of the greater galactic environment, and conspire to produce the scaling relationships that characterize the global properties of galaxies.

With five-band NUV–U–B–V–I imaging of the disks of 38 spiral galaxies at distances of 4–23 Mpc, and parallel V- and I-band imaging of their halos, PHANGS-HST is providing a census of tens of thousands of compact star clusters and associations, which will be combined with PHANGS-ALMA giant molecular clouds (and PHANGS-MUSE H II regions for 19 galaxies in the sample). Previous to this program, no HST wide-field UV imaging existed for 80% of the PHANGS-HST sample, and 60% did not have optical imaging with either WFC3 or ACS. Thus, PHANGS-HST provides a critical augmentation to the HST archive for nearby spiral galaxies in which both star clusters and molecular clouds can be efficiently detected by HST and ALMA over galactic scales. Altogether, PHANGS will provide an unprecedented joint catalog of the observed and physical parameters for ~100,000 star clusters, associations, H II regions, and molecular clouds.

In this paper, we have described the ensemble global properties of the 38 galaxies targeted for HST observations, and how they were selected from the parent PHANGS-ALMA

sample of nearby massive galaxies on the star-forming main sequence. The acquisition and processing of the HST observations to produce aligned, drizzled, science-ready images were described in detail. An overview of the major components of the pipeline developed to produce catalogs of single-peak compact star clusters, and a parallel pipeline for multiscale stellar associations, was provided as a framework for forthcoming detailed papers on each of those components. We highlight new methods involving multiple concentration index (MCI) parameters, synthetic star clusters, and CNN models for cluster candidate selection and morphological classifications, as well as a watershed algorithm based procedure for identifying stellar associations from smoothed, filtered maps of point-source positions. We described the data products to be released via MAST at https://archive.stsci.edu/hlsp/phangs-hst/, including imaging, catalogs, and software, which will enable community science beyond the main goals of PHANGS.

These data products and the PHANGS census of star clusters, associations, H II regions, and molecular clouds will provide the context needed for meaningful study of the earliest phases of dust enshrouded star formation and ISM physics with JWST. Molecular clouds and UV bright clusters/associations are the precursors and descendants of the youngest dusty clusters in nearby galaxies to be uncovered by JWST through their infrared emission. With HST-matched resolution in the near-IR (PSF FWHM $0\rlap{.}''066$ at $2\,\mu$m) and order-of-magnitude improved resolution compared to Spitzer in the mid-IR (PSF FWHM of $0\rlap{.}''665$ at $21\,\mu$m), molecular clouds with embedded sources can be identified, enabling a key test of our census of inactive clouds, and measurement of time to star formation onset.

Through a PHANGS Cycle 1 Treasury program[59] with an allocation of 107 hours, we will obtain NIRCAM and MIRI imaging in eight bands from 2 to $21\,\mu$m for the 19 galaxies with the full set of PHANGS-ALMA, MUSE, and HST observations. Imaging in the F200W, F300M, and F360M filters will provide a low obscuration view of stellar photospheric emission (with some contribution from hot dust to F300M and F360M). F335M, F770W, and F1130W will capture polycyclic aromatic hydrocarbon (PAH) emission, tracing a combination of size and charge, with the F300M and F360M filters enabling continuum subtraction for the $3.3\,\mu$m PAH feature. F1000W and F2100W will provide measurements of the warm dust continuum (with some contribution to the F1000W band by silicate absorption).

By resolving the infrared emission across these 19 morphologically diverse galaxies into individual regions and clusters (5–50 pc scales), the PHANGS-JWST observations will complete the inventory of star formation activity in our targets, and reveal the physical state of the small dust grains that heat the ISM. We will calculate mass functions, and spatial distributions for young embedded sources, study their relation to those of other populations, and more clearly identify the conditions that ignite star formation. The combination of uniform, systematic observations from JWST combined with those already in hand from HST, ALMA, and VLT/MUSE will significantly advance our understanding of the multiscale process of star formation, and the progression from clouds to visible stars in a galactic context.

---

[59] https://www.stsci.edu/jwst/phase2-public/2107.pdf

ORCID iDs

Janice C. Lee ⓘ https://orcid.org/0000-0002-2278-9407
David A. Thilker ⓘ https://orcid.org/0000-0002-8528-7340
Sinan Deger ⓘ https://orcid.org/0000-0003-1943-723X
Kirsten L. Larson ⓘ https://orcid.org/0000-0003-3917-6460
Gagandeep S. Anand ⓘ https://orcid.org/0000-0002-5259-2314
Médéric Boquien ⓘ https://orcid.org/0000-0003-0946-6176
Rupali Chandar ⓘ https://orcid.org/0000-0003-0085-4623
Daniel A. Dale ⓘ https://orcid.org/0000-0002-5782-9093
Eric Emsellem ⓘ https://orcid.org/0000-0002-6155-7166
Adam K. Leroy ⓘ https://orcid.org/0000-0002-2545-1700
Erik Rosolowsky ⓘ https://orcid.org/0000-0002-5204-2259
Eva Schinnerer ⓘ https://orcid.org/0000-0002-3933-7677
Judy Schmidt ⓘ https://orcid.org/0000-0002-2617-5517
Jordan Turner ⓘ https://orcid.org/0000-0003-2261-5746
Schuyler Van Dyk ⓘ https://orcid.org/0000-0001-9038-9950
Richard L. White ⓘ https://orcid.org/0000-0002-9194-2807
Ashley T. Barnes ⓘ https://orcid.org/0000-0003-0410-4504
Francesco Belfiore ⓘ https://orcid.org/0000-0002-2545-5752
Frank Bigiel ⓘ https://orcid.org/0000-0003-0166-9745
Guillermo A. Blanc ⓘ https://orcid.org/0000-0003-4218-3944
Yixian Cao ⓘ https://orcid.org/0000-0001-5301-1326
Melanie Chevance ⓘ https://orcid.org/0000-0002-5635-5180
Enrico Congiu ⓘ https://orcid.org/0000-0002-8549-4083
Oleg V. Egorov ⓘ https://orcid.org/0000-0002-4755-118X
Simon C. O. Glover ⓘ https://orcid.org/0000-0001-6708-1317
Kathryn Grasha ⓘ https://orcid.org/0000-0002-3247-5321
Brent Groves ⓘ https://orcid.org/0000-0002-9768-0246
Jonathan D. Henshaw ⓘ https://orcid.org/0000-0001-9656-7682
Annie Hughes ⓘ https://orcid.org/0000-0002-9181-1161
Ralf S. Klessen ⓘ https://orcid.org/0000-0002-0560-3172
Eric Koch ⓘ https://orcid.org/0000-0001-9605-780X
Kathryn Kreckel ⓘ https://orcid.org/0000-0001-6551-3091
J. M. Diederik Kruijssen ⓘ https://orcid.org/0000-0002-8804-0212
Daizhong Liu ⓘ https://orcid.org/0000-0001-9773-7479
Laura A. Lopez ⓘ https://orcid.org/0000-0002-1790-3148
Ness Mayker ⓘ https://orcid.org/0000-0002-5993-6685
Sharon E. Meidt ⓘ https://orcid.org/0000-0002-6118-4048
Eric J. Murphy ⓘ https://orcid.org/0000-0001-7089-7325
Hsi-An Pan ⓘ https://orcid.org/0000-0002-1370-6964
Jérôme Pety ⓘ https://orcid.org/0000-0003-3061-6546
Miguel Querejeta ⓘ https://orcid.org/0000-0002-0472-1011
Alessandro Razza ⓘ https://orcid.org/0000-0001-7876-1713
Toshiki Saito ⓘ https://orcid.org/0000-0002-2501-9328
Patricia Sánchez-Blázquez ⓘ https://orcid.org/0000-0003-0651-0098
Francesco Santoro ⓘ https://orcid.org/0000-0002-6363-9851
Amy Sardone ⓘ https://orcid.org/0000-0002-5783-145X
Jiayi Sun ⓘ https://orcid.org/0000-0003-0378-4667
Antonio Usero ⓘ https://orcid.org/0000-0003-1242-505X
Thomas G. Williams ⓘ https://orcid.org/0000-0002-0012-2142

References

Adamo, A., Hollyhead, K., Messa, M., et al. 2020, MNRAS, 499, 3267
Adamo, A., Kruijssen, J. M. D., Bastian, N., Silva-Villa, E., & Ryon, J. 2015, MNRAS, 452, 246
Adamo, A., Ryon, J. E., Messa, M., et al. 2017, ApJ, 841, 131
Anand, G. S., Lee, J. C., Van Dyk, S. D., et al. 2020, MNRAS, 501, 3621
Anand, G. S., Rizzi, L., & Tully, R. B. 2018, AJ, 156, 105
Anand, G. S., Rizzi, L., Tully, R. B., et al. 2021, AJ, 162, 80
Armus, L., Mazzarella, J. M., Evans, A. S., et al. 2009, PASP, 121, 559
Barbarino, C., Dall'Ora, M., Botticella, M. T., et al. 2015, MNRAS, 448, 2312
Blitz, L. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson, AZ: Univ. Arizona Press), 125

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Bradley, L., Sipocz, B., Robitaille, T., et al. 2019, astropy/photutils: v0.6 Zenodo, doi:10.5281/zenodo.2533376
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Calzetti, D., Johnson, K. E., Adamo, A., et al. 2015a, ApJ, 811, 75
Calzetti, D., Lee, J. C., Sabbi, E., et al. 2015b, AJ, 149, 51
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chabrier, G. 2003, PASP, 115, 763
Chandar, R., Fall, S. M., & Whitmore, B. C. 2010a, ApJ, 711, 1263
Chandar, R., Fall, S. M., Whitmore, B. C., & Mulia, A. J. 2017, ApJ, 849, 128
Chandar, R., Whitmore, B. C., Calzetti, D., & O'Connell, R. 2014, ApJ, 787, 17
Chandar, R., Whitmore, B. C., Kim, H., et al. 2010b, ApJ, 719, 966
Chevance, M., Kruijssen, J. M. D., Vazquez-Semadeni, E., et al. 2020, SSRv, 216, 50
Cignoni, M., Sacchi, E., Aloisi, A., et al. 2018, ApJ, 856, 62
Cignoni, M., Sacchi, E., Tosi, M., et al. 2019, ApJ, 887, 112
Cook, D. O., Lee, J. C., Adamo, A., et al. 2019, MNRAS, 484, 4897
Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012, ApJS, 200, 18
Deger, S., Lee, J., & Whitmore, B. 2021, MNRAS, submitted
Dobbs, C. L. 2008, MNRAS, 391, 844
Dobbs, C. L., Bonnell, I. A., & Pringle, J. E. 2006, MNRAS, 371, 1663
Dobbs, C. L., Krumholz, M. R., Ballesteros-Paredes, J., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 3
Dolphin, A. E. 2016, DOLPHOT: Stellar photometry, Astrophysics Source Code Library, ascl:1608.013
Elmegreen, B. G. 2000, ApJ, 530, 277
Elmegreen, B. G. 2008, ApJ, 672, 1006
Emsellem, E., Schinnerer, E., Santoro, F., et al. 2021, A&A, submitted, (arXiv:2110.03708)
Elson, R. A. W., Fall, S. M., & Freeman, K. C. 1987, ApJ, 323, 54
Federrath, C. 2013, MNRAS, 436, 3167
Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, ApJ, 553, 47
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gieles, M., & Portegies Zwart, S. F. 2011, MNRAS, 410, L6
Girichidis, P., Offner, S. S. R., Kritsuk, A. G., et al. 2020, SSRv, 216, 68
Gnedin, N. Y. 2016, Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality (Berlin: Springer), 1
Grasha, K., Calzetti, D., Adamo, A., et al. 2017, ApJ, 840, 113
Grasha, K., Calzetti, D., Adamo, A., et al. 2019, MNRAS, 483, 4707
Grasha, K., Calzetti, D., Bittle, L., et al. 2018, MNRAS, 481, 1016
Haan, S., Surace, J. A., Armus, L., et al. 2011, AJ, 141, 100
Hennebelle, P., & Falgarone, E. 2012, A&ARv, 20, 55
Heyer, M., & Dame, T. M. 2015, ARA&A, 53, 583
Ho, I. T., Kreckel, K., Meidt, S. E., et al. 2019, ApJL, 885, L31
Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Huang, C. D., Riess, A. G., Yuan, W., et al. 2020, ApJ, 889, 5
Hughes, A., Meidt, S. E., Colombo, D., et al. 2013, ApJ, 779, 46
Humphreys, R. M., & Davidson, K. 1979, ApJ, 232, 409
Jacobs, B. A., Rizzi, L., Tully, R. B., et al. 2009, AJ, 138, 332
Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2015, ApJ, 802, 127
Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2016, ApJ, 827, 33
Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2017, ApJ, 839, 78
Karachentsev, I. D., Karachentseva, V. E., Huchtmeier, W. K., & Makarov, D. I. 2004, AJ, 127, 2031
Kennicutt, R. C. J., Lee, J. C., Funes, J. G., et al. 2008, ApJS, 178, 247
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kim, D. C., Evans, A. S., Vavilkin, T., et al. 2013, ApJ, 768, 102
Klessen, R. S., & Glover, S. C. O. 2016, Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality (Berlin: Springer), 85
Kourkchi, E., Courtois, H. M., Graziani, R., et al. 2020a, AJ, 159, 67
Kourkchi, E., & Tully, R. B. 2017, ApJ, 843, 16
Kourkchi, E., Tully, R. B., Anand, G. S., et al. 2020b, ApJ, 896, 3
Kreckel, K., Faesi, C., Kruijssen, J. M. D., et al. 2018, ApJL, 863, L21
Kreckel, K., Ho, I. T., Blanc, G. A., et al. 2019, ApJ, 887, 80
Kreckel, K., Ho, I. T., Blanc, G. A., et al. 2020, MNRAS, 499, 193
Kruijssen, J. M. D. 2012, MNRAS, 426, 3008
Kruijssen, J. M. D., Schruba, A., Chevance, M., et al. 2019, Natur, 569, 519
Krumholz, M. R., & McKee, C. F. 2005, ApJ, 630, 250
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Lamers, H. J. G. L. M., & Levesque, E. M. 2017, Understanding Stellar Evolution (Bristol: IOP Publishing)

Lang, P., Meidt, S. E., Rosolowsky, E., et al. 2020, ApJ, 897, 122
Larson, K. L., Díaz-Santos, T., Armus, L., et al. 2020, ApJ, 888, 92
Leonard, D. C., Kanbur, S. M., Ngeow, C. C., & Tanvir, N. R. 2003, ApJ, 594, 247
Lee, E. J., Miville-Deschênes, M.-A., & Murray, N. W. 2016, ApJ, 833, 229
Lee, J. C., Gil de Paz, A., Kennicutt, R. C. J., et al. 2011, ApJS, 192, 6
Lee, J. C., Kennicutt, R. C., Jr., Funes, S. J. J. G., Sakai, S., & Akiyama, S. 2009, ApJ, 692, 1305
Leroy, A. K., Hughes, A., Liu, D., et al. 2021a, ApJS, 255, 19
Leroy, A. K., Sandstrom, K. M., Lang, D., et al. 2019, ApJS, 244, 24
Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021b, ApJS, 257, 43
Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782
Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, AJ, 146, 19
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Linden, S. T., Evans, A. S., Rich, J., et al. 2017, ApJ, 843, 91
Lopez, L. A., Krumholz, M. R., Bolatto, A. D., et al. 2014, ApJ, 795, 121
Mac Low, M.-M., & Klessen, R. S. 2004, RvMP, 76, 125
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565
McQuinn, K. B. W., Skillman, E. D., Dolphin, A. E., Berg, D., & Kennicutt, R. 2017, AJ, 154, 51
Meidt, S. E., Glover, S. C. O., Kruijssen, J. M. D., et al. 2020, ApJ, 892, 73
Meidt, S. E., Schinnerer, E., García-Burillo, S., et al. 2013, ApJ, 779, 45
Messa, M., Adamo, A., Östlin, G., et al. 2018, MNRAS, 473, 996
Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, ApJ, 834, 57
Murray, N. 2011, ApJ, 729, 133
Moffat, A. F. J. 1969, A&A, 3, 455
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
Nugent, P., Sullivan, M., Ellis, R., et al. 2006, ApJ, 645, 841
Olivares E., F., Hamuy, M., Pignata, G., et al. 2010, ApJ, 715, 833
Olivier, G. M., Lopez, L. A., Rosen, A. L., et al. 2021, ApJ, 908, 68
Perez, G., Messa, M., Calzetti, D., et al. 2021, ApJ, 907, 100
Péroux, C., & Howk, J. C. 2020, ARA&A, 58, 363
Pierce, M. J., Welch, D. L., McClure, R. D., et al. 1994, Natur, 371, 385
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Rahner, D., Pellegrini, E. W., Glover, S. C. O., & Klessen, R. S. 2017, MNRAS, 470, 4453
Reid, M. J., Pesce, D. W., & Riess, A. G. 2019, ApJL, 886, L27
Rosolowsky, E., Hughes, A., Leroy, A. K., et al. 2021, MNRAS, 502, 1218
Ruiz-Lapuente, P. 1996, ApJL, 465, L83
Ryon, J. E., Gallagher, J. S., Smith, L. J., et al. 2017, ApJ, 841, 92
Sacchi, E., Cignoni, M., Aloisi, A., et al. 2018, ApJ, 857, 63
Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22
Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11
Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, ApJS, 227, 2
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Santoro, F., Kreckel, K., Belfiore, F., et al. 2021, arXiv:2111.09362
Shaya, E. J., Tully, R. B., Hoffman, Y., & Pomarède, D. 2017, ApJ, 850, 207
Science Software Branch at STScI 2012, PyRAF: Python alternative for IRAF, Astrophysics Source Code Library, ascl:1207.011
STSCI Development Team 2012, DrizzlePac: HST image software, Astrophysics Source Code Library, ascl:1212.011
Sun, J., Leroy, A. K., Schruba, A., et al. 2018, ApJ, 860, 172
Thilker, D. A., Whitmore, B. C., Lee, J. C., et al. 2022, MNRAS, 509, 4094
Tonry, J. L., Dressler, A., Blakeslee, J. P., et al. 2001, ApJ, 546, 681
Tully, R. B., Courtois, H. M., & Sorce, J. G. 2016, AJ, 152, 50
Tully, R. B., Rizzi, L., Shaya, E. J., et al. 2009, AJ, 138, 323
Turner, J. A., Dale, D. A., Lee, J. C., et al. 2021, MNRAS, 502, 1366
U, V., Sanders, D. B., Mazzarella, J. M., et al. 2012, ApJS, 203, 9
van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, arXiv:1407.6245
Walch, S., Girichidis, P., Naab, T., et al. 2015, MNRAS, 454, 238
Ward, J. L., & Kruijssen, J. M. D. 2018, MNRAS, 475, 5659
Ward, J. L., Kruijssen, J. M. D., & Rix, H.-W. 2020, MNRAS, 495, 663
Wei, W., Huerta, E. A., Whitmore, B. C., et al. 2020, MNRAS, 493, 3178
Whitmore, B., Lee, J., Chandar, R., & Thilker, D. 2021, MNRAS, 506, 5294
Whitmore, B. C., Chandar, R., & Fall, S. M. 2007, AJ, 133, 1067
Whitmore, B. C., Chandar, R., Schweizer, F., et al. 2010, AJ, 140, 75
Whitmore, B. C., Zhang, Q., Leitherer, C., et al. 1999, AJ, 118, 1551
Wright, N. J. 2020, NewAR, 90, 101549