

Spitzer Warm Mission Archive Science Opportunities

Lisa J. Storrie-Lombardi*, John R. Stauffer*, Bidushi Bhattacharya*, Sean Carey*, Dave Frayer*, Mark Lacy*, Victoria Meadows*, Alberto Noriega-Crespo*, Luisa Rebull*, Erin Ryan[†], Inseok Song*, Susan Stolovy*, Harry Teplitz*, David Trilling** and Schuyler van Dyk*

**Spitzer Science Center, Caltech 314-6, Pasadena, CA 91125, USA*

[†]*Dept. of Astronomy, U. Minnesota, 116 Church St. SE, Minneapolis, MN 55455, USA*

***Steward Observatory, U. Arizona, 933 N. Cherry Ave., Tucson AZ 85721, USA*

Abstract. The rich data archive from the Spitzer cryogenic mission will be comprised of approximately 25 TB of data. A five-year warm mission would add an additional 15-20 TB. All of these data will be processed and archived to form homogeneous, reliable database to support research for decades after the end of the Spitzer mission. The SSC proposes a robust archival research program during the warm mission phase. A sampling of possible archival programs are described.

Keywords: Spitzer Space Telescope, archives, infrared astronomical observations

PACS: 95.55.Fw, 95.80.+p, 95.85.Hp

1. INTRODUCTION

The Spitzer Space Telescope is approaching the milestone of four years in orbit (at the time of the workshop), with enough cryogen for a 5.5 year nominal mission profile. After the cryogen is exhausted, the 3.6 and 4.5 μm channels of IRAC will still perform well, but the other Spitzer detectors will no longer produce useful data. Because the sensitivity in the two shorter IRAC channels is better than any other existing or planned facility until JWST is launched, we expect that Spitzer will continue to operate for an additional five year period as an IRAC 3.6 and 4.5 μm imaging mission. We refer to this as the Spitzer Warm Mission.

One of the Spitzer Science Center's top priorities after the cryogen is depleted is to do a final reprocessing of all of the cryogenic mission data as quickly possible in order to provide the most homogenous, most reliable database possible for archival usage by the community. The cryogenic archive is expected to comprise of order 25 TB of data. Most of the observing time during the warm mission will be devoted to large, legacy-type programs, which will generate very large volumes of data particularly amenable to archival research.

We believe that the Spitzer combined cryogenic and warm mission archive will be heavily used during (and after) the warm mission time period (2009-2014), and we are planning to support this usage by building a robust, flexible archive interface and by providing a significant amount of funding to support the analysis of Spitzer archival data. Spitzer is able to observe ~ 7000 hours per year which is more than twice as much time as any space or ground-based optical/infrared facility.

In this white paper, we outline the expected order-of-magnitude data volume in the Spitzer final archive; the types of data that will be provided; the planned capabilities of the archive interface; and provide some illustrative examples of the types of science programs we expect would be motivated by the Spitzer data.

2. THE FINAL SPITZER DATA ARCHIVE

The cryo archive contains data from all three instruments, IRAC, MIPS and IRS. Currently, the archive interface (“Leopard”) allows users to obtain raw, basic calibrated data (BCD) and post-BCD data from individual Astronomical Observation Requests (AORs), with their associated calibration files on a per-AOR basis.

Data from Spitzer Legacy programs, processed beyond the level of the current SSC pipelines by the Legacy teams, are currently made available for direct download from the SSC website, and also ingested into the Infrared Science Archive (IRSA). These data include mosaics of multiple AORs, spectral data cubes and band-merged source lists.

The current rate of ingestion into the archive is about 4.5 TB per year, with 4 TB per year from data received and processed by the pipeline, and 0.5 TB per year from the Legacy teams. By the end of the 5.5 year cryo-mission we expect approximately 25 TB of data in the archive. Currently the actual volume stored by the archive is larger than these rates would suggest as 1-2 processing versions are retained on disk before moving them into our long-term tape archive. The warm mission data rate is not expected to be much lower – although IRAC will only be taking data in two channels, the data rates will still be much higher than currently from IRS. Our data rate estimate for the warm mission is approximately 3.5-4 TB per year, but will vary depending on the character of the executed science observations.

After the cryo mission ends, the cryo data will undergo a final processing. This is expected to be completed in 12-18 months after the end of the cryogenic mission. These data will be entered into a “final” Spitzer archive, whose design is underway. The main differences between the final archive and the current one derive from the fact that the structure of the current archive is strongly related to processing requirements, rather than long-term storage and ease of access to the data. The final archive will be designed as a long-term archive, curation of which will be handed over to IRSA at the end of the Spitzer mission.

The main difference between the final archive and the current archive from the user perspective will be that data will be available not only on a per-AOR basis, as currently, but also on a per-BCD basis. The archive will also be fully compatible with interoperability standards (e.g. it will support existing Virtual Observatory protocols). We will ensure that a full provenance of each file is maintained, so that the details of the processing of each product are made clear. Calibration files will also be available online so that users can re-reduce data using different calibrations from those used in the final processing should they feel it is necessary.

As part of the final archive design the SSC is formulating the list of data products expected to be delivered to the final archive. These will, of course, include the existing products generated by the pipelines, but will also include higher-level products which we hope to produce by the end of the cryo mission if time and resources allow. One

example of this type of product is source lists for imaging data.

The final archive may also contain some of the tools currently available to Spitzer users (e.g. Mopex, Spice, GERT, CUBISM and IRSCLEAN). However, it is yet to be determined how maintenance of these tools will be funded past the end of the Spitzer mission, and, consequently, whether they will be served by IRSA.

3. SUPPORT FOR MINING THE SPITZER DATA ARCHIVE

An active archival research program has been supported through the annual calls for proposals since Cycle-1. The Legacy programs executed in the first year of the Spitzer mission provided large, coherent, publicly available data sets for the general community to mine very early in the mission. Additional Legacy programs selected in each subsequent cycle, in addition to the substantial amounts of data becoming public each year, ensure a very rich archive for supporting research programs beyond the science proposed by the original investigators (e.g. Deguchi *et al.* [1]). We have just scratched the surface of the archive so far and we are already seeing an increase in archival research.

The SSC normally recommends that 10% of the available data analysis funds for each cycle be awarded to archive and theory programs (~\$2 million annually). The typical award is \$50,000 to \$100,000. In Cycle-4 the time allocation committee specifically recommended awards of \$2.7 million due to the high quality of the archival programs. Previous time allocation committees and the Spitzer User's Committee have requested that the SSC solicit large archival proposals. The SSC has been unable to add these large programs during the cryo mission due to annual budget cuts which have impacted the User Community funds.

During the warm mission the SSC proposes substantially increasing the funding available for archival programs. The focus on the observing programs will be larger programs and with a smaller dollar/hour for data analysis funding. During the cryo-mission the SSC awards \$20-25 million annually for general observer, archival and theory funding with 90% going to observing data analysis support and 10% for archive/theory support. For the warm mission we propose to award \$20 million annually with 75% going to observing data analysis support and 25% for archive/theory support. In addition to one-year archive proposals we would like to add a new program supporting Legacy archival proposals. These would be multiyear archival programs to tackle problems that can't properly be addressed with a typical one-year archive program.

4. SPITZER WARM MISSION SCIENCE: CONSTRAINTS, AREAL COVERAGE AND LIKELY DATA GENERATION RATE

A companion white paper entitled "The Spitzer Warm-Mission Science Prospects" describes in detail the expected capabilities of Spitzer after the cryogen runs out, and our plans for operating the observatory during the warm mission. That white paper also describes several possible large legacy-type programs that could be carried out during the warm mission. Here, we provide a brief review of that material in order to indicate

the scope and character of the data from the warm mission period that will be available for archival usage.

The Spitzer warm mission has a natural and well-defined beginning point, which is the point when the cryogen is exhausted. The best estimate for the end of cryogen is March-April 2009. The warm mission also has a natural and well-defined ending point based on geometry. Spitzer is slowly drifting away from the Earth. In 2014 the spacecraft reaches a distance such that we will be unable to recover from a safe mode using the low-gain antenna. This constraint provides a warm mission duration of order five years.

Based on our current plans and knowledge, the Spitzer observing efficiency should be unchanged in the warm mission, implying that there will be of order 6500-7000 hours of time available each year for science observations. The sensitivities and noise-characteristics of the two shorter-wavelength IRAC channels should also be unchanged from the cryo-mission. Because only two of the four channels will be operating, however, the data generation rate will be of order half that of the cryo-mission. From an extrapolation of the IRAC campaigns executed to date, we expect a data generation rate in the warm mission period of order 3.7 TB/year.

Both as a necessary means to reduce operating costs and as a means to encourage different types of science programs from those carried out during the cryogenic mission, there will be a strong preference for large and very-large programs during the warm mission. IRAC warm mission programs will occupy a niche in sensitivity-area space between that for WISE (all-sky but comparatively very shallow) and JWST-NIRCAM (very deep but pencil-beam). This suggests that most of the warm mission observing programs are likely to cover regions of the sky from tens to hundreds of square degrees, with sensitivities of order $1 - 10 \mu\text{Jy}$, 5σ . If **all** of the observing time in the warm mission were devoted to a relatively shallow survey with $10 \mu\text{Jy}$, 5σ sensitivity, of order 1/4 of the sky could be covered. Other possible programs could image the same region of the sky two or more times for proper motion purposes or to detect variable sources. In some cases, in order to address the primary science goals, it will be necessary for the proposing team to obtain complementary optical or near-IR imaging of the surveyed region, and those data would also be expected to become part of the Spitzer archive.

A direct consequence of allocating the majority of the warm mission observing time to very large programs is that archival science will take on a significantly greater importance. By their nature, programs that cover large areas of the sky uniformly can address multiple science topics. In order to encourage full exploitation of the archive we intend to allocate a greater percentage of the user community funding for archival science during the warm mission time period, compared to the cryo-mission.

5. EXAMPLES OF LARGE ARCHIVAL PROJECTS - GALACTIC SCIENCE

5.1. MIPS $24\mu\text{m}$ Archival Search for Exo-Zodiacal Dust

A number of programs used Spitzer to detect debris disks around nearby dwarfs in order to help understand the processes involved with planet formation. At $24 \mu\text{m}$,

even relatively short integration observations can detect small amounts of warm dust to a hundred parsecs; therefore, any scan-map observation could detect such debris disks around field stars that happen to fall within the mapped area. An analysis of appropriate archival MIPS data could thus add significantly to the statistical study of the frequency and evolution of debris disks around main sequence stars. An automated spectral energy distribution (SED) fitting could be performed on all stars in MIPS 24 μm images with integration times per point on the sky > 24 seconds; the MIPS photometry could be combined with optical/near-IR photometric data from Tycho-2, SuperCOSMOS, USNO-B1, NOMAD, and 2MASS catalogs. A SED fitting technique with synthetic stellar spectra can determine stellar photospheric flux at 24 μm with a 3σ uncertainty of $\lesssim 20$ mJy. This high fidelity SED fitting allows identifications of stars with very small warm IR excess, analogous to our solar system’s asteroidal belt or zodiacal dust grains at somewhat younger ages. A very short lifetime of warm grains around these stars requires either continuous, violent, multiple collisions among planetesimals (i.e., heavy bombardment) or a catastrophic huge collision between two planets (e.g., BD+20 307, Song *et al.* [2]) The latter mimics the collision which created our own Moon. These stars may represent a population of adolescent stars currently undergoing a heavy bombardment in their inner planetary zone. Using about $\sim 8,000$ MIPS 24 μm images from GO Cycle-1, First Look Survey, and Legacy data, about 50 mid-IR excess stars have already been identified. From the entire 5 year Spitzer cryo mission data, we anticipate a couple of hundred warm excess stars would be detected allowing better estimates of the occurrence rate of heavy bombardment at other stars (or the rate of catastrophic planet-planet collisions).

5.2. Y-Dwarf Search from IRAC Archival Images

IRAC archival images provide a large data set to search for late-T and Y-dwarfs. Late T-dwarfs occupy a distinct region in the IRAC color-color diagram and are well-separated from essentially all other stars (even PMS stars with primordial disks). The data shown in Fig. 1 are from an archival image (AOR 3954944) of a known T8 dwarf (2MASSI 041519-0935). A GTO observation by G. Fazio targeted this specific T8 dwarf. Figure 1 demonstrates that late T-dwarfs like this one can be easily identified from an IRAC archival search. The T8 dwarf is designated with horizontal and vertical error bars while other background sources were plotted as dots without error bars. A criterion to be plotted in this figure was $S/N > 10$ detection in all 4 bands. Although no Y-dwarfs have been identified, contemporary theoretical spectral models of these very cool dwarfs predict that at least early Y dwarfs should occupy a similar region of the IRAC color-color space as late T-dwarfs. It is almost certainly true that Y dwarfs will occupy a very distinct region of color space because of their very cool temperatures and presumably exotic molecular features. High redshift quasars can have similar colors to late T dwarfs and so a survey with IRAC or with IRAC plus ground-based near-IR photometry could be contaminated by such quasars – but those are interesting objects also, and their detection is valuable in their own right. Any Y dwarf candidate that can be identified with a wide-area IRAC survey will be near enough to Earth to have relatively large, and

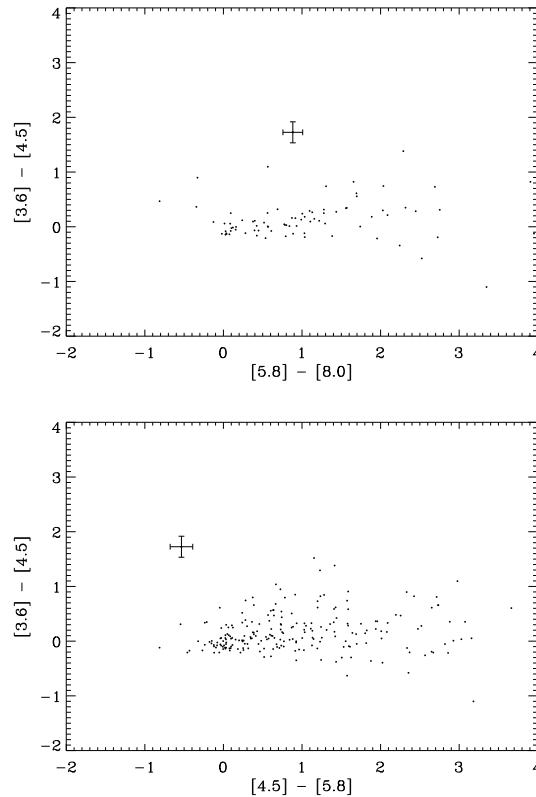


FIGURE 1. A known T8 dwarf (2MASSI 041519-0935) is designated with horizontal and vertical error bars while other background sources were plotted as dots without error bars. A criterion to be plotted in this figure was $S/N > 10$ detection in all 4 bands.

measurable proper motions – allowing the separation of the Y dwarf candidates from the distant quasars with suitable two-epoch data. A few test studies of several square degrees of IRAC survey data from First Look Survey and Legacy programs have been performed in order to find these very cool dwarfs, but have so far not found any. In the combined cryogenic and warm mission archive, we expect several hundred square degrees of imaging capable of detecting Y dwarfs to 50 pc or more - and in such a large area, we would expect to find a few Y dwarfs (combining the IRAC observations with moderate depth J band imaging for the warm mission survey data).

5.3. A Census of Star-Cloud Collisions Using IRAC

The disk of the Milky Way is composed of a mixture of stars and gas. The stars come in a wide range of mass from > 100 solar masses to below 0.01 solar mass, with the study of the stellar IMF being a flourishing subfield of its own. The clouds of gas in the galactic disk also come with a wide range of masses. Radio surveys provide data from which the mass function of these clouds can be derived. Cloud census surveys are

good at finding the most massive clouds, and to some extent in measuring structures within large cloud complexes. However, it is much more difficult to make a census of small, relatively isolated clouds. Therefore their frequency and properties are much less well characterized. One of the best ways to identify such clouds is when they pass near luminous stars. The nebulae surrounding the Pleiades is an example of the collision of a small (~ 10 solar mass) molecular cloud with a star system. Examples of individual field A stars also encountering small clouds have been reported in the literature (e.g. Kalas *et al.* [3]; Gaustad and van Buren [4]). Observations with IRAC channel 4 ($8\mu\text{m}$) provide a particularly sensitive means to detect such encounters because of the strong PAH emission in that band. As shown by Stauffer *et al.* [5], some of the Pleiades stars without evidence of interaction with the Merope cloud at optical wavelengths do have easily detected $8\mu\text{m}$ excesses (and spatially resolved $8\mu\text{m}$ emission at Pleiades distance). Figure 2 shows one of these star-cloud interactions as observed by IRAC. A survey of the entire cryogenic IRAC database for other examples of cloud-star interactions via the signature of excess or spatially resolved $8\mu\text{m}$ excesses would better constrain the volume filling factor of small molecular or atomic clouds, and in favorable cases allow one to estimate the properties of those clouds.

5.4. Cleaning the IRAS PSC and FSC

The IRAS point source catalog (PSC; Beichman *et al.* [6], IRAS) and faint source catalog (FSC; Moshir *et al.* [7]) provide extremely rich resources for all-sky searches for interesting objects such as stars with circumstellar disks or AGN. However, the structure of the background emission at any one location can provide a significant source of confusion to the large-aperture IRAS measurements. A Spitzer archive project to compare IRAS and Spitzer data over large scales could create a much higher-quality catalog of such interesting objects, weeding out knots of nebulosity or spurious sources (such as asteroids), and finding objects that break into multiples when viewed with MIPS at $24\mu\text{m}$ as illustrated here in Fig. 3

In a 10.6 square degree MIPS survey of the Perseus molecular cloud Rebull *et al.* [8] assessed how successful the IRAS survey was in identifying point sources in a potentially confusing region - in this cloud, complex extended emission is present at all three MIPS bands (and the IRAC bands for that matter). Simply examining the PSC and FSC objects in this region finds objects that are listed as detections at 60 or 100 μm , with only upper limits at 12 and $25\mu\text{m}$; even without Spitzer data, one might suspect that such sources correspond to texture in the extended emission. Though the PSC and FSC do contain warning flags for cirrus the Spitzer data comparison clearly show that significantly fewer spurious sources appear in the FSC than in the PSC. Overall, 20-30% of the “point sources” at 12 and $25\mu\text{m}$ resolve into knots of nebulosity. There are many more “point sources” reported at 60 and 100 μm (than at 12 and $25\mu\text{m}$) in the PSC. A much lower fraction of these objects are recovered; many more of these are completely missing or fall apart into nebulosity when viewed with MIPS. More than half of the 100 μm point sources are completely missing, and 25-40% of the 100 and 60 μm point sources are clearly confused by nebulosity. Overall, a much larger

HII 1234

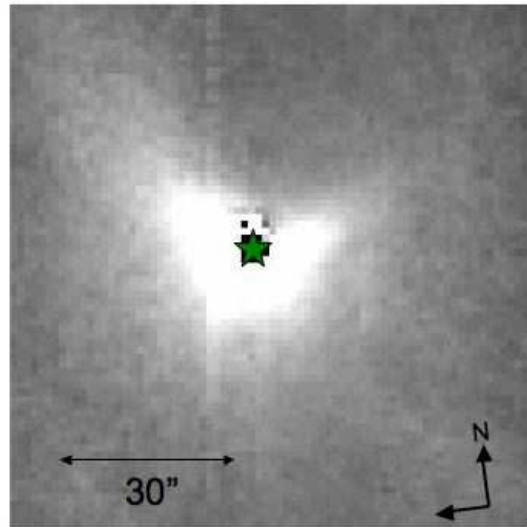


FIGURE 2. Example of an A star impacting a small interstellar cloud. This is an IRAC 8 μm image of HII1234, an A dwarf member of the Pleiades. We have subtracted an empirical PSF from the image to enhance the PAH emission from the star-ISM interaction. The bow-shock shaped nebular emission is due to a shell of material created by radiation pressure from the star sculpting the material in the cloud as the Pleiades moves through the cloud.

fraction of FSC objects are recovered by MIPS than PSC objects. Surprisingly few IRAS sources are resolved into multiple objects by Spitzer, but those that are resolved are often “interesting”(see discussion in Rebull *et al.* [8]).

A Spitzer archival project to do a similar comparison of IRAS and Spitzer maps over the 3750+ hours of MIPS scan observations (very roughly 1000+ square degrees distributed over the whole sky) in the archive could create a vastly improved IRAS PSC that weeds out the 20-40% of objects that are simply knots of nebulosity, and finds the interesting objects that break into multiples. For the Perseus analyses, the comparisons of the IRAS catalog and the MIPS images were done by hand, but an automated routine could be developed to look for true MIPS point sources within the IRAS error ellipse for every IRAS point source.

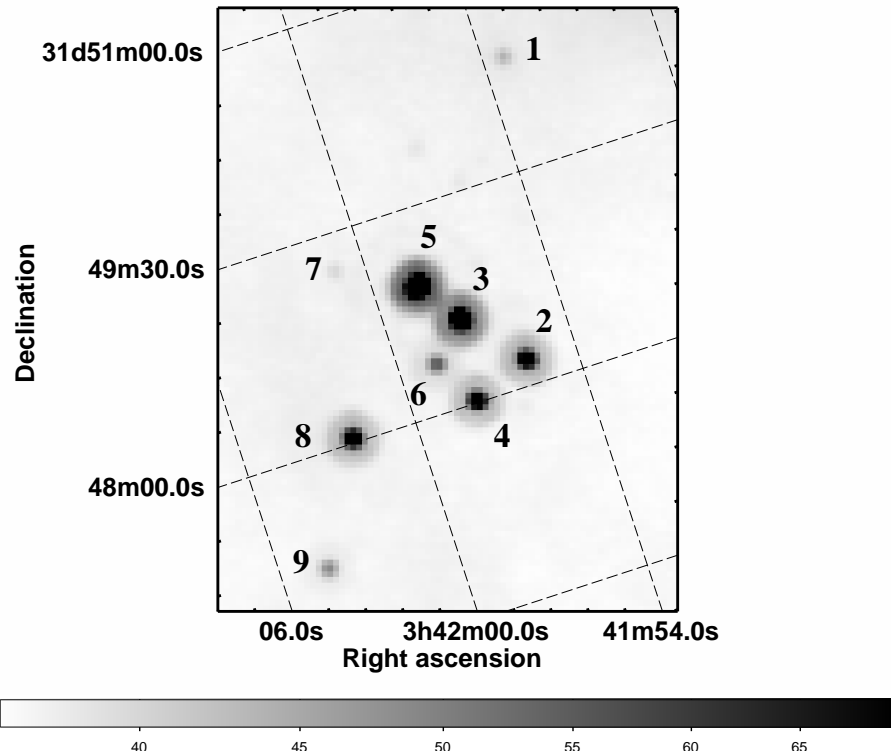


FIGURE 3. Example of an “object” identified as a single point source with IRAS data, which is resolved by MIPS ($24\mu\text{m}$) into a small cluster. Reprinted from Rebull *et al.* [8].

5.5. Identifying “Missing” Evolved Stars

Massive stars profoundly influence the evolution of galaxies, through energy input and chemical enrichment of the interstellar medium via stellar winds and supernovae. Stars with $M_{\text{init}} > 20 M_{\odot}$ are believed to evolve, depending on M_{init} , as O-type ZAMS \rightarrow Be/B[e] \rightarrow (red supergiant ?? \rightarrow) luminous blue variable (Be/B[e]/Of) \rightarrow Wolf-Rayet (WR) star \rightarrow supernova/collapsar. The high luminosities ($L \sim 10^6 L_{\odot}$) and short duration (~ 0.5 Myr) of the WR phase, and short lifetimes (< 4 Myr) of WR precursors make WRs excellent markers and chronometers of very recent star formation in galaxies, as well as sensitive metallicity indicators through their observed relative line strengths. WRs also potentially serve as probes of the IMF in galaxies. The number of observationally identified WRs in the Galaxy so far (approximately 300) is much lower than model predictions by factors of 4–10. This discrepancy motivates the search in the infrared (where, e.g., $A_K/A_V \approx 10\%$ for WRs likely “missing,” or “hidden” optically by the rapid increase of line-of-sight A_V with distance ($0.5\text{--}2$ mag kpc^{-1}) in the Galactic plane (almost all of the known WRs have $|b| < 5^\circ$). All WR stars exhibit infrared excesses, due in large part to the free-free emission originating from electron scattering in the neighborhood of (mainly) He^{++} ions in the outer regions of their dense winds; circumstellar heated

dust shells can also contribute, especially for the late WC stars. What we have found is that using the combined near-IR/mid-IR color space as a ‘sieve’ provides a very successful means of revealing hidden WRs—the near- to mid-IR colors well sample the large infrared excess all WR stars exhibit (with $f_\lambda \propto \lambda^{-2.7}$ to -3.2), compared to normal photospheres (with $f_\lambda \propto \lambda^{-4}$). From a GLIMPSE+2MASS combination we have found 21 new WRs, as well as about 180 new evolved, high-mass stars in Of, Be, B[e], LBV, and Of/WN, i.e., *pre-WR* transitional, phases. We will be continuing this search in Cycle-4 using GLIMPSE-II, GLIMPSE-3D, MIPS GAL, and MIPS GAL-II archival data.

We will not be able to find *all* of the missing evolved stars in the Galaxy, but over the long term, primarily the IRAC and MIPS observations from any program targeting the Galactic plane during the cryo mission will provide a large reservoir of data from which to identify new candidates. The data would help in both characterizing the massive star formation throughout the plane *and* testing more accurately the stellar evolutionary scenarios for massive stars in the Galaxy. This is essential for understanding galaxy evolution both in the nearby Universe and at high redshift.

5.6. Systematic Study of Infrared Extinction

The high-quality mid-infrared photometric surveys (e.g. GLIMPSE, MIPS GAL, C2D) from Spitzer can be used to investigate the mid-infrared extinction law as a function of environment and line of sight. Until recently, extinction in the mid-infrared had only been measured for the Galactic center and work is ongoing to compare observed extinction laws for particular lines of sight to current grain models (Román-Zúñiga *et al.* [9]). Initial results (Indebetouw *et al.* [12]) suggest an extinction curve in rough accord with current models (Weingartner and Draine [11]) but suggest that the mid-IR extinction has a flatter wavelength dependence (see Fig. 4). A key question to address is whether the mid-IR extinction curve is universal or has a strong dependence on line of sight. A comprehensive investigation of the extinction law along many lines of sight will permit modeling the law as a function of environment and will constrain the grain size distribution, grain growth and other physical processes.

One method of determining the extinction to a particular star would be to simultaneously fit the available 1.2-24 μm photometric points with an model spectral energy distribution and extinction law Robitaille *et al.* [13]. Point source catalogs for the entire inner Galactic plane ($|l| < 60^\circ$, > 260 square degrees) exist from the 2MASS, GLIMPSE and MIPS GAL surveys. This project would also make use of existing high quality photometry from the c2D legacy program. A wide range of extinctions could be probed ($A_v \sim 1$ to > 100). The extinction determination can be augmented by archival IRS data of available targets.

With a well parameterized extinction curve, the total dust column can be estimated even for very large columns ($A_v > 100$). The derived column map then can be used in investigating Galactic structure by helping to constrain a deconvolution of far-IR through mm continuum emission into phases of the ISM (warm ionized, warm neutral, cold neutral). In addition, the three-dimensional distribution of the extinction can be modelled using appropriate stellar populations models and the existing photometry

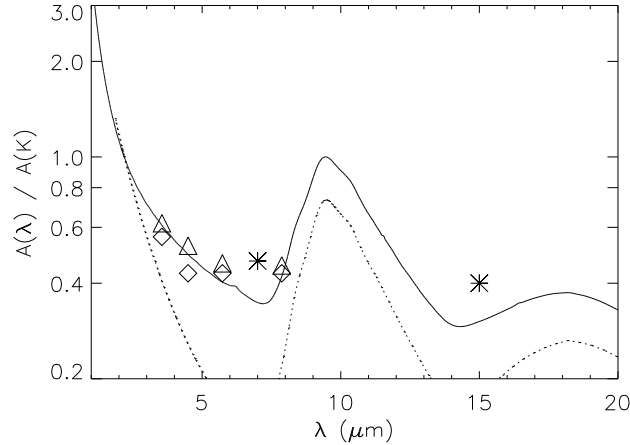


FIGURE 4. Measured mid-infrared extinction curves from Indebetouw *et al.* [12] (diamonds), Román-Zúñiga *et al.* [9] (triangles) and Jiang *et al.* [10] (asterisks). The $R_v = 5.5$ model extinction law from Weingartner and Draine [11] is plotted for comparison.

(Marshall *et al.* [14]). Extinction measurements provide a temperature independent way of estimating the column of big interstellar grains. This method is complementary to the future millimeter/submillimeter observations of interstellar dust where the emission is a convolution of the column density with the temperature distribution and emissivity.

5.7. Measurement of Galactic Star Formation and Cataloguing of Infrared-Dark Clouds

High quality source lists and images of the Galactic plane have been produced by the GLIMPSE and MIPS GAL surveys at wavelengths of 3.6 to $70\mu\text{m}$. The > 260 square degrees these surveys cover encompass most of the Galactic disk and encapsulate most of the star forming regions of the Galaxy. These data are complemented by existing spectral line surveys and will be augmented by upcoming surveys of the Galactic plane in the millimeter/submillimeter. Models of protostellar SEDs exist and are currently being refined (Robitaille *et al.* [13]). A very ambitious synthesis of the available multi-wavelength data of the Galactic plane and theory of star formation would be to catalog all detectable protostars in the Galactic plane and estimate the current star formation rate and initial mass function of the Galaxy. Distances of protostars can be estimated from millimeter spectral line data and association of the diffuse emission/absorption around a protostar with the morphology of the molecular gas as traced by spectral line mapping. One class of star forming objects for which this method has proven fruitful is infrared-dark clouds (IRDCs; Fig. 5). With the multi-wavelength infrared mapping of the Galactic plane, IRDCs should be easily identified using contrast-color filters as has been done previously using the MSX Galactic plane data (Simon *et al.* [15]). A catalog of these objects alone will provide a current snapshot of the earliest stage of star

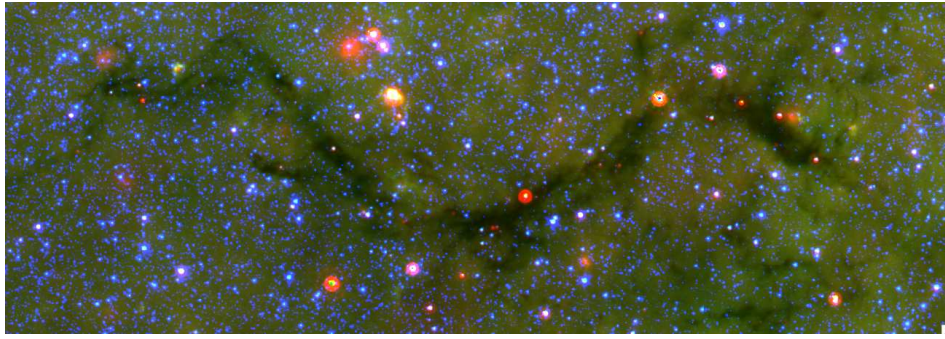


FIGURE 5. Sample infrared-dark cloud at 3.6, 8.0 and 24.0 μm . The extinction lane is well correlated with millimeter wave line emission; therefore, a kinematic distance to the object can be well estimated (3.6 kpc in this case). The majority of red point sources in the image are class I and II protostars.

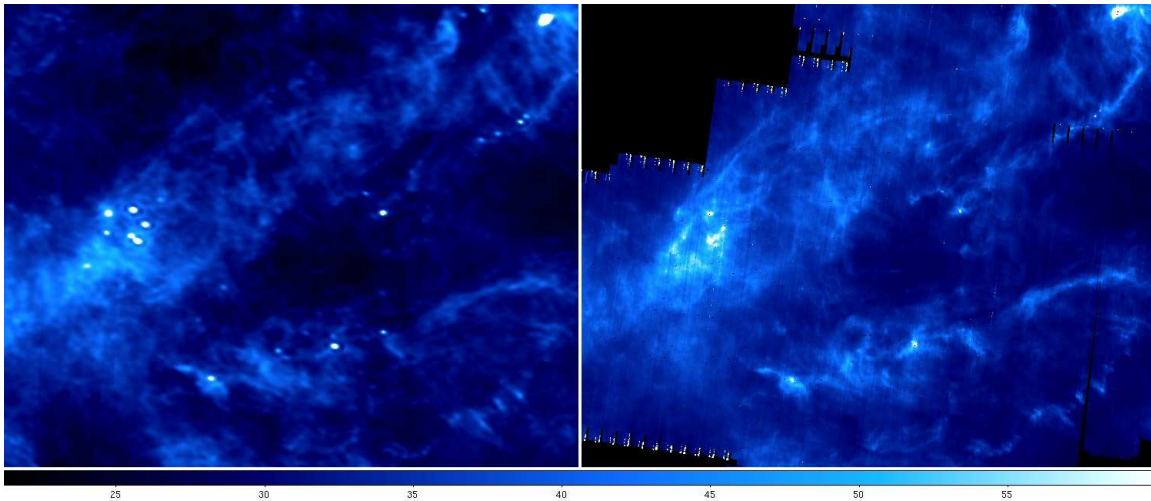


FIGURE 6. The Taurus Molecular Cloud (TMC) has been surveyed with IRAC and MIPS (Padgett *et al.* 2007; in preparation). The image shows a nearly four degree field of the TMC, comparing an IRAS HiRes map at 100 μm (left; $\sim 1.5'$ resolution) and the MIPS 160 μm image (right; $\sim 40''$ resolution). The structure of the cold dust in this classical molecular cloud displays a fascinating and complex morphology.

formation in the Galaxy.

The protostellar identification and extinction studies are complementary and can be done jointly with existing source catalogs and using the existing spectral energy distribution models. Much of the analysis can be done in an automated fashion through least-squares fitting of SEDs and extinction models to the photometry.

5.8. Turbulence and The Structural statistics of the ISM

There is a wealth of research on the morphology and structure of molecular clouds, based on the analysis of two-dimensional images obtained through several tracers. These

emission maps of the interstellar medium (ISM), either in ^{12}CO , or dust extinction or IRAS infrared emission, show a remarkable property: they are self-similar on a wide range of scales, from a few hundreds of parsecs down to ~ 0.02 pc (see e.g. Bensch *et al.* [16], Falgarone *et al.* [17]).

A way to describe this self-similarity is through the angular power spectrum, which is a function of inverse scale size (wavenumber) with an exponent ranging from $\beta \sim -3.6$ to $\beta \sim -2.5$. It is this self-similarity that has allowed us to make the connection between morphology and turbulence. The classical cascading energy dissipated from large to small scales is also a power-law with an index closely related to that of the density or velocity dispersion (through their emission and absorption tracers). The connection between structure or morphology of the gas and dust, either cirrus-like or molecular, with that of the feedback processes that sculpt their shape and the evolution of the ISM, is one of the most fertile and exciting fields in Galactic astronomy.

In the infrared, in particular, the data from the Spitzer Galactic First Look Survey (GFLS) (Burgdorf *et al.* [18]) demonstrated how far one can push the limits. In a section of the Gum Molecular cloud, the IRAC and MIPS images have permitted study of the turbulent density field to scales of $\sim 5''$, and shown that HI and dust power spectra, traced by the $24\mu\text{m}$ & $70\mu\text{m}$ emission, are essentially identical down to spatial scales ~ 0.01 pc (Ingalls *et al.* [19]). In this regime, excitation, ionization and extinction are key players in interpreting the observations. The MIPS $24\mu\text{m}$ images offer more than an order of magnitude improvement in the spatial resolution of turbulent processes over other infrared surveys (eg. IRAS).

A similar analysis to that of the GFLS is being currently undertaken on some areas of the MIPS GAL survey and Taurus molecular cloud (Fig. 6). These data sets are but a small fraction of what is available in the Spitzer archive. Data from the C2D and GLIMPSE legacies are obvious candidates to work with, but cirrus-like emission or structures have also been detected in the SWIRE and Extra-GFLS (for example).

An archival project to study the structural statistics of the ISM (diffuse, clumpy, molecular) will allow a better understanding of turbulent density fields and to test our ideas of energy transfer and dissipation, not only at interesting angular scales, but also in the most diverse environments.

6. EXAMPLES OF LARGE ARCHIVAL PROJECTS - EXTRA-GALACTIC SCIENCE

6.1. Spitzer's Combined Extragalactic $16\mu\text{m}$ Survey

The IRS provides science-quality $16\mu\text{m}$ imaging with its Peak-Up (PU) camera. While the number of dedicated PU imaging programs is small ($\sim 5\%$ of IRS time), short low (SL) spectra and PU data are taken together on a single detector. There are about 170,000 SL exposures in the Spitzer archive. These include almost a square degree of low-background PU imaging with exposure times of at least 100 sec; about 100 square arcminutes of these have exposure times more than 15 minutes. These data will include 5–10 thousand galaxies with $f_{16} > 120\mu\text{Jy}$. No pointed Spitzer survey could achieve

this areal coverage due to the small PU field of view, so the archival data represent previously unexplored phase space for the mission.

These PU images can greatly improve the $16\ \mu\text{m}$ source counts. The flux limit of the archival data is comparable to the deepest ISOCAM surveys, and has better spatial resolution. At the limit these counts will be dominated by LIRGs and ULIRGs at redshifts approaching unity. Proper measurement on the extragalactic source counts are required across all available passbands in order to constrain the shape of emission (PAH) and absorption (silicate) features in template SEDs. Current analysis of $24\ \mu\text{m}$ source counts by different groups still favor disparate SEDs; additional wavelengths may resolve the issue.

Many, if not most, SL spectra are taken of objects in fields also observed by IRAC and MIPS. Thus, the $16\ \mu\text{m}$ photometry of many sources can be used together with other Spitzer data to constrain the MIR slope of extragalactic sources. Kasliwal *et al.* [20] used archival $16\ \mu\text{m}$ images to show that the $16/24\ \mu\text{m}$ ratio can be used to identify “silicate dropouts” (Takagi and Pearson [21]), that is $z \sim 1.5$ objects with strong silicate absorption in the MIPS-24 passband, in order to identify the most strongly extinguished starbursts and AGN. Teplitz *et al.* [22] found that $16\ \mu\text{m}$ excess is also indicative of PAH emission at $z \sim 1$.

6.2. Spitzer’s Combined Extragalactic $70\ \mu\text{m}$ Survey

All of the MIPS- $70\ \mu\text{m}$ extragalactic survey data could be reprocessed using optimized offline data reduction scripts. From the ultra-deep programs of GOODS-N and GOODS-S (PI: Frayer, PID3325+20147) and the S-COSMOS (PI: Sanders PID20070+30143, TC: Frayer) and FIDEL (PI: Dickinson, PID30948, TC: Frayer) Legacy programs, it has been demonstrated that offline processing can improve the rms sensitivity by 20-25% over the SSC online products (Frayer *et al.* [23]). More importantly, the techniques clean up the non-Gaussian wings of the noise distribution, enabling the extraction of sources which are about 1.5-2 times fainter at the same level of reliability (e.g., >99.5%). These techniques would be difficult to implement within the online SSC system since they involve iterative reduction steps that would require a complete re-design of the SSC downlink infrastructure. The data from all programs (including targeted GO observations) would be optimally combined for all of the popular survey fields [e.g., xFLS, SWIRE, Bootes, GOODS, EGS, IRAC-deep, Lockman-deep, GTO, and COSMOS] to produce the deepest possible data over each field.

An improvement of a factor of 1.5-2 in depth in the derived source catalogs would have a wide-range of science applications. For instance, one could produce the definitive $70\ \mu\text{m}$ source counts from Spitzer. The wide-area xFLS, SWIRE, and Bootes programs can be used to measure the bright source counts, while the deep S-COSMOS, FIDEL, GTO surveys, and the publicly available GO programs (e.g., IRAC-deep field and Lockman) can be used to measure the turn-over ($\sim 5\text{-}12\ \text{mJy}$) of the Euclidean-normalized differential source counts. Accurate source counts severely constrain the models of galaxy evolution and their spectral-energy-distributions (SEDs). Both the evolutionary models and the underlying SEDs require modification to fit the observed $70\ \mu\text{m}$ counts while

maintaining consistency with the counts at other wavelengths. The enhanced $70\ \mu\text{m}$ data would also be used to constrain the far-infrared properties of BzK galaxies, X-ray sources, radio galaxies, as well as other $24\ \mu\text{m}$ and IRAC populations of high-redshift sources being uncovered by Spitzer. Having a large set of optimally, uniformly reduced $70\ \mu\text{m}$ data is crucial to address these and other similar questions.

Currently, the Spitzer $70\ \mu\text{m}$ archived data are under-utilized by the public. Unlike MIPS $24\ \mu\text{m}$ and IRAC, most MIPS-Ge science programs have relied on just a few experts for their data reduction. Having catalogs and enhanced mosaic images available to the general user would trigger much wider usage of these data, overcoming the typical user's concern about the tractability of the Ge data.

6.3. Searches for Galaxy Clusters at $z > 1$

Spitzer data, when combined with wide-field optical/near-IR imaging, is a powerful tool for finding distant galaxy clusters in the redshift range $z \sim 1 - 2$, the so-called "cluster desert". This range is particularly important as it spans the gap between the clusters of galaxies found in optical surveys (e.g. the RCS, Gladders *et al.* [24]) and the non-virialized high- z protoclusters found, e.g., through Lyman break galaxy searches at $z \sim 2 - 5$. It is thus the redshift range at which we expect to see clusters virialize, and the cluster red sequence and morphology-density relation to be established. However, until recently, few clusters had been found in the cluster desert. The low areal density of rich, high- z galaxy clusters makes them hard to find in both ground-based near-infrared surveys, and space-based X-ray surveys. Spitzer/IRAC, with its highly efficient mapping AOT and low background is an excellent cluster-hunting tool. Using Spitzer data, in combination with the ground-based $9\ \text{deg}^2$ NOAO deep wide survey, (Stanford *et al.* [25]) recently discovered a $z = 1.41$ galaxy cluster, the second highest redshift cluster known in the cluster desert. The ongoing Sparcs project (Wilson *et al.* [26]) uses the $50\ \text{deg}^2$ SWIRE survey to search for clusters, using $z' - [3.6]$ colors to identify the cluster red sequence. We expect that cluster surveys will constitute a significant part of the archive use. Including all the wide area surveys such as SWIRE, the IRAC shallow survey and smaller surveys such as the XFLS and COSMOS, about $70\ \text{deg}^2$ in total are available for this work, much of which already has high quality, ground-based imaging data. Although some surveys will already be complete before the end of the cryogenic mission, better cluster finding algorithms will continue to be developed, and better ground-based data will become available to help with photometric redshift estimates. High- z cluster finding is also well-suited to the warm mission, as IRAC channels 1 and 2 are most useful for these surveys, and the warm mission archive will contain a large amount of suitable data.

6.4. Searches for Dust Obscured AGN and Quasars

AGN and quasars obscured by dust make up 1/2-3/4 of the total AGN population, and dominate the $24\ \mu\text{m}$ source counts at $S_{24} > 1\text{mJy}$ (e.g. Brand *et al.* [27]). The areal

density of these objects is $\approx 100 \text{ deg}^{-2}$ in shallow-medium Spitzer surveys, thus, in the total area of Spitzer surveys there are about 7000 objects. Efficient mechanisms to identify AGN on the basis of Spitzer broad-band photometry at 3.6-24 μm have been developed for shallow and medium Spitzer surveys (Lacy *et al.* [28], Sajina *et al.* [29], Stern *et al.* [30]). Importantly, Spitzer surveys are able to find the Compton thick AGN that are almost impossible to find in X-ray surveys, yet which comprise about half of the local AGN population, and probably a similar fraction of the hard X-ray background (Gilli *et al.* [31]). Finding large numbers of obscured AGN and quasars will allow us to quantify the fraction of obscured objects, and how this changes with redshift and AGN luminosity. It will also allow us to investigate the relationship between the obscured and unobscured quasar populations.

6.5. Stacking Analyses of Galaxy Populations

Many galaxy populations are well-represented in the wide area Spitzer surveys, and well detected at the shorter Spitzer wavelengths, but typically undetected at longer wavelengths. Dole *et al.* [32] show that stacking 70 μm and 160 μm images at the positions of faint galaxies detected at 24 μm results in statistical detections an order of magnitude fainter than the individual source detection limits in those bands. Eyles *et al.* [33] stack $z \sim 6$ galaxies detected in IRAC bands 1 and 2 in IRAC bands 3 and 4, and show for the first time that the average $z \sim 6$ galaxy probably has a sub-solar metallicity. Other examples of other possible uses of this sort of analysis would include: stacking of IRAC channel 4 and MIPS data for Lyman break galaxies to establish their AGN activity (using 8 and 24 μm data) and star formation activity (using 70 and 160 μm data), and stacking of 70 and 160 μm data on quasars detected at 24 μm to investigate their far-IR SED as a function of quasar luminosity and redshift.

7. EXAMPLES OF LARGE ARCHIVAL PROJECTS - SOLAR-SYSTEM SCIENCE

Asteroids are remnants of the early solar system and provide clues to its origin and evolution. Unlike planets and satellites, they are not subject to appreciable geological evolution, have remained relatively pristine since their formation and can provide insight into the primordial solar nebula. Their current surface characteristics and size distributions serve as markers of the physical and orbital evolution and of the effects of long-term space weathering on objects within our solar system. As asteroids are routinely present in Spitzer observations near the ecliptic plane, characterizing their prevalence is of interest to non-planetary astronomers as well, who desire to distinguish nearby objects in their mid-infrared data from extrasolar sources they are studying.

The solar system's Main Belt asteroid (MBA) population is characterized via Size Frequency Distribution functions (SFD), taxonomies and surface properties. To date, this work has been carried out using primarily ground-based data. Models such as the Statistical Asteroid Model (Tedesco *et al.* [34]), are limited by observational constraints to defining properties for asteroids with diameters $> 1 \text{ km}$. Ground-based observations

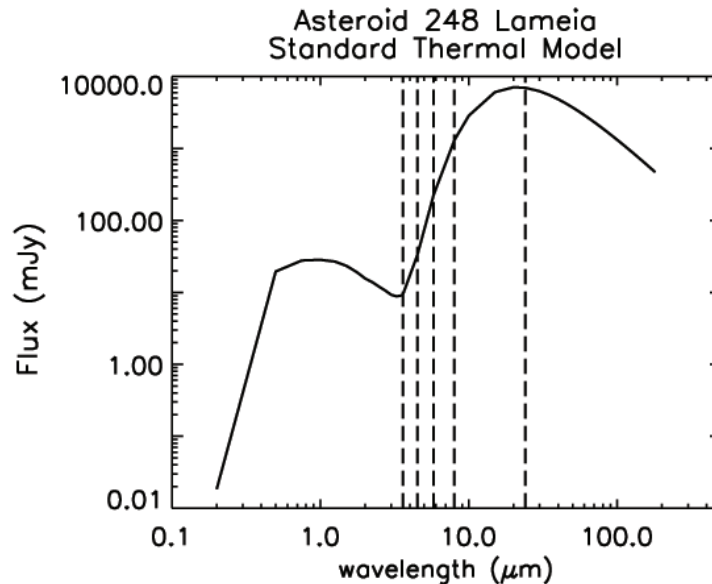


FIGURE 7. A typical main belt asteroid spectral energy distribution. The vertical dashed lines indicate the IRAC 3.6, 4.5, 5.8, and 8.0 μm and the MIPS 24 μm detection bands. The IRAC 8 μm and MIPS 24 μm detectors are well suited to detect MBA thermal emission. For warmer temperatures, the SED is shifted to shorter wavelengths, allowing for detection of NEOs.

are biased towards higher albedo, larger MBAs. Number counts based on traditional surveys are vulnerable to selection effects from intense study of known asteroids, leading to ‘clustering’ of new discoveries near previously known objects. Mid-infrared measurements with Spitzer have provided, for the first time, a detailed look at the sub-km MBA population. In addition, the Spitzer archive provides a wealth of information about larger, serendipitously observed known asteroids. Spitzer is highly sensitive and can measure thermal emission from previously undetected, small asteroids. As shown in Fig. 7, where a sample spectral energy distribution (SED) for a typical X-type MBA is provided, peak thermal flux is emitted within the Spitzer band passes, making these objects ideal candidates for study in the mid-infrared.

Spitzer asteroid surveys have been carried out using targeted observing programs as well as with archival surveys. Although both types of programs have examined only a small portion of the sky, they clearly demonstrate the feasibility of using IRAC and MIPS to characterize MBAs. The number of objects found, particularly out of the ecliptic plane, where smaller asteroids appear to be prevalent (Xenos *et al.* [35]), is far greater than predicted by empirically-based models such as the Statistical Asteroid Model (Tedesco *et al.* [34]). The First Look Survey Ecliptic Plane Component (Meadows *et al.* [36]) scanned portions of the sky in and near the ecliptic plane and demonstrated Spitzer’s sensitivity to MBAs. The survey found double the number of asteroids expected at $\beta = +5^\circ$. Analysis of data in the SWIRE XMM-LSS field (Ryan and Woodward [37]) at $\beta = -17^\circ$ found 39 asteroids in the field and only 11 of these asteroids have ground

based optical detections. In addition, Xenos *et al.* [35] have found previously unknown asteroids as far away from the ecliptic plane as $\beta = 26^\circ$.

Deep Spitzer observations over multiple observing epochs are particularly well-suited for examining the scale height behavior and size and albedo distributions of previously unknown, small asteroids. Bhattacharya *et al.* [38] have examined the S-COSMOS field and found over a hundred asteroids in a 0.17 deg^2 field. Other programs that may provide a wealth of information include SWIRE and TAUPROP. Additional data can be found in multi-epoch GO programs and in future Legacy programs. The studies described above have worked with small data sets and relied on a significant level of non-automated identification of asteroids. Studies of the archive as a whole will require automated software, such as that currently being developed by Trilling [39] (PID 30213).

Asteroid observations taken throughout Spitzer's cryogenic and warm mission will also enhance the scientific return of future projects such as Pan-STARRS. This ground-based project will identify up to tens of millions of asteroids, a significant fraction of which will be new. The Spitzer archive will provide linkages for MBAs observed during the cryo mission and for Near Earth Objects (NEOs) observed during the warm mission, allowing for further characterization of a large number of small bodies. Continuation of these efforts through the warm mission and beyond, as well as extending their technique to identifying unknown asteroids, is critical.

The solar elongation constraints inherent in Spitzer observations limits asteroidal motions to a range of approximately 10-40 arcseconds per hour. Thus 2 epochs of data are sufficient to detect motion on the sky by the asteroids in survey data. Known asteroids can be identified using a single epoch, if their orbits are well defined. For surveys with large area coverage it is also possible to link IRAC and MIPS observations of known asteroids obtained with a one-month timescale. Use of only the SCOSMOS, SWIRE and TAUPROP data constitutes 350 Gigabytes of data over 56 square degrees which will yield an estimated asteroid catalog of 1000-2000 objects. This will double the number of asteroids for which there are known albedos and diameters based on IRAS detections and extend the size distribution past the IRAS reliability limit of 10 km.

8. SUMMARY

The programs outlined in the previous section are by no means an exhaustive description of the possible archival programs that will likely be undertaken using the Spitzer cryo and warm mission data. In particular, such a large archive lends itself to serendipitous discoveries of rare and unique objects. The archive and interfaces to it are being constructed to facilitate use by Spitzer experts as well as those new to the mission. The archive programs approved in the first four cryogenic cycles can be found on the SSC Web Pages at <http://ssc.spitzer.caltech.edu/approvdprog/>.

ACKNOWLEDGMENTS

This work is based (in part) on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology

under contract with NASA. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

REFERENCES

1. Deguchi, S., Nakashima, J., Kwok, S., and Koning, N., 2007, ApJ, 664, 1130
2. Song, I. *et al.*, 2005, Nature, 436, 363
3. Kalas, P., Graham, J. R., Beckwith, S. V. W., Jewitt, D. C., and Lloyd, J. P., 2002, ApJ, 567, 999
4. Gaustad, J. and van Buren, D. 1993, PASP, 105, 1127
5. Stauffer, J.R. *et al.*, 2007, ApJS, in press (astro-ph/0704:1832)
6. Beichman, C., Neugebauer, G., Habing, H., Clegg, P., Chester, T., 1988, IRAS Catalog and Explanatory Supplement
7. Moshir, M., Kopman, G., Conrow, T., 1992, IRAS Faint Source Survey and Explanatory Supplement
8. Rebull, L. *et al.*, 2007, ApJS, 171, 447
9. Román-Zúñiga, C. G., Lada, C. J., Muench, A., and Alves, J. 2007, astro-ph/07043203
10. Jiang, B., Gao, J., Omont, A., Schuller, F., and Simon, G., 2006, A&A, 446, 551
11. Weingartner, J. C., and Draine, B. T. 2001, ApJ, 548, 296
12. Indebetouw, R. *et al.*, 2005, ApJ, 619, 931
13. Robitaille, T. P., Whitney, B. A., Indebetouw, R., and Wood, K. 2007, ApJS, 169, 328
14. Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., and Picaud, S. 2006, A&A, 453, 635
15. Simon, R., Jackson, J. M., Rathborne, J. M., and Chambers, E. T., 2006, ApJ, 639, 227
16. Bensch, F., Stutzki, J., Ossenkopf, V., 2001, A&A, 366, 636
17. Falgarone, E., *et al.*, 2005 in Star Formation in the Interstellar Medium, ASPC, 323, 185
18. Burgdorf, M. *et al.*, 2005, AdSpR, 36, 1050
19. Ingalls, J. *et al.*, 2004, ApJS, 154, 281
20. Kasliwal, M. M., Charmandaris, V., Weedman, D., Houck, J. R., Le Floch, E., Higdon, S. J. U., Armus, L., and Teplitz, H. I. 2005, ApJL, 634, L1
21. Takagi, T. and Pearson, C. P. 2005, MNRAS, 357, 165
22. Teplitz, H. I., Charmandaris, V., Chary, R., Colbert, J. W., Armus, L., and Weedman, D. 2005, ApJ, 634, 128
23. Frayer, D. *et al.*, 2006, ApJL, 647, L9
24. Gladders, M. *et al.*, 2007, ApJ, 655, 128
25. Stanford, S.A. *et al.*, 2005, ApJL, 634, L129
26. Wilson, G. *et al.*, 2007, To appear in the proceedings of "Infrared Diagnostics of Galaxy Evolution", 2007 (astro-ph/0604289)
27. Brand, K. *et al.*, 2006, ApJ, 644, 143
28. Lacy, M. *et al.*, 2004, ApJS, 154, 166
29. Sajina, A., Lacy, M., Scott, D., 2005, ApJ, 621, 256
30. Stern, S.A. *et al.*, 2005, ApJ, 631, 163
31. Gilli, R., Comastri, A., Hasinger, G., 2007, A&A, 463, 79
32. Dole, H. *et al.*, 2006, A&A, 451, 417
33. Eyles, W., Bunker, A. J., Ellis, R. S., Lacy, M., Stanway, E. R., Stark, D. P., and Chiu, K., 2007, MNRAS, 374, 910
34. Tedesco, E. T., Cellino, A., and Zappalá, V., 2005, AJ, 129, 2869
35. Xenos, S., Bhattacharya, B., Meadows, V., and Puget, J., 2006, DPS 38, 58.03
36. Meadows, V.S. *et al.*, 2004, ApJS, 154, 469
37. Ryan, E. and Woodward, C., 2006, DPS, 38, 5914
38. Bhattacharya, B. Salvato, M., Aussel, H., Ilbert, O., Sanders, D. B., and Scoville, N., 2006, DPS 38, 59.19
39. Trilling, D., Spitzer 30213 abstract, <http://ssc.spitzer.caltech.edu/geninfo/ar/abs-ar3/30213.txt>