Spitzer Science in the Post Cryogen Era

Patrick J. McCarthy

Carnegie Observatories, 813 Santa Barbara St. Pasadena, CA, 91101, USA

Abstract. I review the broad scientific opportunities offered by the Spitzer warm mission and attempt to put them into context with other priorities and future and ongoing programs in astronomy and astrophysics. The warm mission offers a number of unique opportunities for scientific investigations that are beyond the reach of other facilities or are not practical with Spitzer operating in its current mode. Some possible approaches to making the best use of the facility during a 3-5 year duration are considered. These include undertaking large and ambitious surveys while preserving small programs in fields where they are most effective.

Keywords: Spitzer Space Telescope, infrared astronomical observations
PACS: 95.85.Hp

1. INTRODUCTION

Once the cryogen is depleted on Spitzer it will rapidly lose its great sensitivity at wavelengths longer than about 5μm. The observatory will, however, remain uniquely powerful in the 3-5μm region of the spectrum by providing imaging with sensitivity that will not be matched until JWST flies. In its warm state Spitzer can provide both cutting edge science and prepare a foundation on which JWST, ALMA and other next generation astronomical facilities can build. The goal of this workshop was to identify and articulate these scientific opportunities and consider how to best use the observatory in an efficient manner. As documented throughout this volume, the scientific potential of the warm mission is impressive. In this contribution I will attempt to pull together some of the disparate science goals into a more or less coherent whole and consider some options of how the Spitzer Science Center and the user community might approach the challenge of maximizing the unique opportunities offered by the warm mission.

2. THE POWER OF SPITZER IN ITS WARM STATE

As the temperature in the telescope and instruments rises the spectrometers and Multi-band Imaging Photometer (MIPS) will cease to function, as internal backgrounds will swamp any astronomical signals and saturate the detectors in even the shortest exposures. The two long wavelength channels (5.8 and 8μm bands) of the Infrared Array Camera (IRAC) will also be rendered inoperable. The short wavelength 3.5 and 4.5μm channels will, however, not only remain operable, their sensitivities will be uncompromised. We can put the IRAC channels 1 & 2 sensitivities in some perspective by comparing with ground-based 8m telescopes. In 100 seconds IRAC has a 5σ point-source sensitivity of ~ 3μJy. An 8m telescope working on the ground at K (2.2μm) can achieve this same sensitivity in about 200 seconds, so the 85cm Spitzer telescope is about a factor of ~ 2 faster at L than a ground-based 8m telescope operating at K. At L-band Spitzer is 120 times faster than a ground-based 8m in good conditions, at M-band the speed gain is closer to a factor of 1000. While adaptive optics will improve the comparison with the ground somewhat, the orders of magnitude reduction in background from a cryogenic telescope operating in space produces a gain in sensitivity and speed that will likely never be matched on the ground.

The power of the IRAC instrument on Spitzer arises not just from its sensitivity, but also from its impressive mapping speed. The large pixels of IRAC (1.2″ x 1.2″) provide a field of view just over 5′ x 5′. This large field cannot be easily matched on the ground as the background prohibits such large pixels. The resulting mapping speed for IRAC, while below that of all-sky mapping instruments like WISE, provides a powerful survey capability. Most of the programs discussed at this workshop make use of this unique combination of sensitivity and mapping speed to carry out surveys that are beyond the reach of any currently planned ground- or space-based facilities.

In its warm state the lifetime of Spitzer is not limited by on-board consumables. Rather the Earth-trailing orbit of the spacecraft will ultimately put it beyond the reach of the ground-system needed for control and data transfer. Thus the lifetime of the warm mission is likely to be set by programmatic priorities with NASA and the US astronomical community. For the purpose of this workshop we are considering a lifetime in the 3-5 year range. Spitzer is a highly efficient observatory, much more so than satellites in low Earth orbits, such as HST. A 5-year Spitzer warm mission will yield roughly as much observing time as Hubble has in its 16 years of operations to date. Allowing for the large difference in sensitivity and observing efficiency, the Spitzer warm mission will out strip *3000 years of 3.5μm and 4.5μm observing on a ground-based 8 to 10m-class telescope*. This is a powerful capability that should not be taken lightly.

The power of the warm mission to carry out large surveys is illustrated in Figure 1. I plot the total areal coverage possible using IRAC with exposure times ranging from 30 seconds to several hours in programs with durations that range from 500 hours to the full 5 year lifetime of the mission. I take a year of observing as equal to 7000 hours and assume that one continues to gain in depth like √t, which is probably optimistic for the longest exposure times. I've not plotted below 100nJy as this is likely near the confusion limit and no brighter than 5μJy as programs shallower than this are highly inefficient. The program duration for any combination of depth and area can be inferred from the contours in Figure 1. A useful benchmark is provided by the observation that in the 5-year lifetime of the warm mission one could map 1000 square degrees to 1μJy.

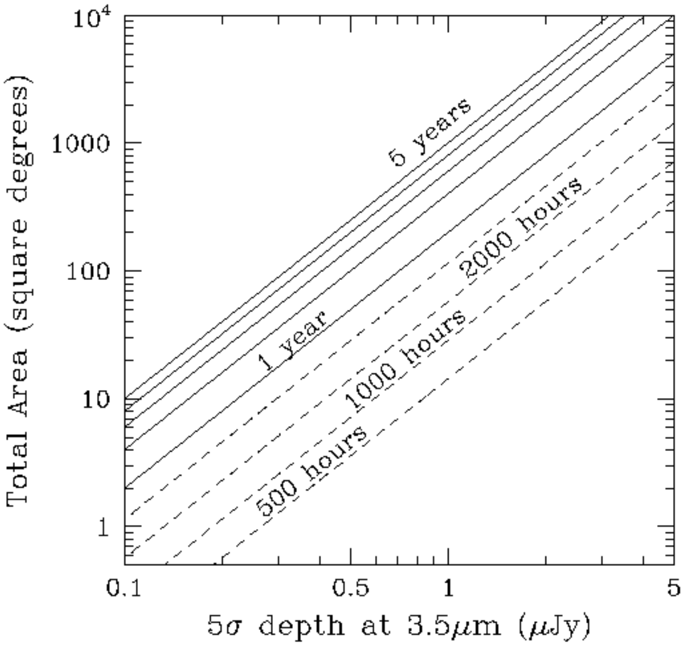

FIGURE 1. Area vs. depth contours for IRAC mapping programs with durations ranging from 500 hours to the life of the mission. A year of observing is taken as 7000 hours of on-source time. Each contour corresponds to exposure times ranging from 30 seconds to several hours. The duration of a program with a particular combination of area and depth at 3.5μm can be read from the graph. A benchmark is provided by the fact that over the entire 5-year mission one could image 1000 square degrees to a 5σ depth of 1μJy.

The NASA explorer program provides an interesting benchmark to put the warm mission in context. The cost cap for the *SMEX* and *MIDEX* missions are currently $105M and $180M, far smaller than the cost of a *Great Observatory* mission, but in the same range as the incremental cost of the warm mission. The *Explorer* class astronomy missions have had lifetimes that range from as short as one year to nearly 20 years, but a 3-5 year lifetime is quite typical for *SMEX* program. These missions tend to have a narrower science focus than a Great Observatory program, but many contain a mix of core survey and guest investigator programs. The Spitzer warm mission thus fits many of the criteria of an *Explorer* class mission with one critical exception – the space craft is in orbit simply waiting for the science mission to be defined by the user community. This is a scientific opportunity that does not often arise – the chance to carry out a cutting edge space astronomy program of the *Explorer* class without the years of suffering that normally accompanies such an endeavor.

3. THE WARM MISSION AND NATIONAL SCIENTIFIC PRIORITIES

While astronomers who are active users of Spitzer, and IRAC in particular, might consider it a foregone conclusion that the warm mission should be supported, the broader community might rightly ask where this fits into our national scientific priorities. There are a number or places where national priorities are considered and tabulated, but the National Academy's Decadal Survey is the most authoritative source for long-range planning in astronomy and astrophysics. The Bachall report from the 1991 survey played a key role in the genesis of the Spitzer observatory. The most recent decadal survey identified a number of scientific questions that are both of fundamental importance and ripe for progress in the coming decade. These include advancing our understanding of the nature and distribution of matter and energy on large scales, identifying the first generation of stars and galaxies, understanding the formation and evolution of black holes, stars and planetary systems and, finally, exploring the impact of the astronomical environment on the Earth. These rather broad goals encompass much of contemporary astrophysics and Spitzer has played an important role in addressing many of these goals. In its warm state Spitzer can continue to have large impacts in most of these areas. As discussed in other contributions to this volume, the warm mission can make unique contributions to our understanding of galaxies in the early universe. IRAC surveys of the earliest clusters of galaxies probe the distribution of matter on very large scales and constrain the expansion history and energy content of the Universe. Even with only its two shortest wavelength channels Spitzer provides powerful probes of star formation and exoplanet physics. It seems clear then that the warm mission fits in well with our national priorities in astronomy and astrophysics and one could argue that it provides a timely and cost effective bridge between the current generation of facilities and future observatories such as the Wide-field Infrared Survey Explorer (WISE) and the James Webb Space Telescope.

The forefront questions in astrophysics increasingly require observational data from a wide range of the electromagnetic spectrum. The days of pure optical, radio, or infrared astronomers are passing and the youngest generation of observers is multi-wavelength oriented in its approach. In this context the warm mission has a great deal of scientific synergy with facilities operating at other wavelengths, and even those that will be soon observing overlapping regions of the spectrum. The synergy between HST and Spitzer is well documented; a number of the key legacy programs, such as GOODS, use both HST and Spitzer to sample the optical and IR regions of the spectrum. Similar synergies are likely to come to light, although in a time staggered manner, when ALMA and JWST begin operations.

4. SCIENTIFIC OPPORTUNITIES IN THE WARM MISSION

Spitzer has had a large impact over a wide range of science to date. The potential impact of the warm mission is curtailed somewhat by its restricted wavelength range. The potential for very large programs, however, opens up new possibilities that have been impractical during the heavily over subscribed early cycles. Before discussing detailed science programs it is instructive to ask where a 3-4 micron imager sits in relation to various astronomical phenomena and astrophysical processes.

4.1 IRAC Channels 1 & 2 in the Global Context

In Figure 2 we consider the location of the Spitzer IRAC channels 1 & 2 bands in the context of the extragalactic background light (EBL). The EBL has two broad peaks, one due primarily to direct radiation from the photospheres of stars and accretion disks. The energy density of this component peaks around ~ 1µm, suggesting that much of it comes from redshifts > 0.5 or so. The secondary peak at ~ 100µm is reprocessed radiation that has been thermalized by interstellar and circumstellar dust. This secondary peak has been the focus of much activity in infrared astronomy and IRAS, Spitzer and other missions have had an enormous impact here. IRAC channels 1 & 2 sample the long wavelength tail of the direct radiation peak. At 3.5 and 4.5µm IRAC samples the peak of the spectral energy distributions of evolved stars at intermediate redshifts, 1 < z < 4, and the peak of the energy distributions of actively star forming galaxies at z ~ 5 and higher. Not surprisingly, these two areas – the study of red galaxies at z > 1 and young galaxies at z > 5 are fields in which IRAC has made major, and in some cases quite unanticipated, discoveries.

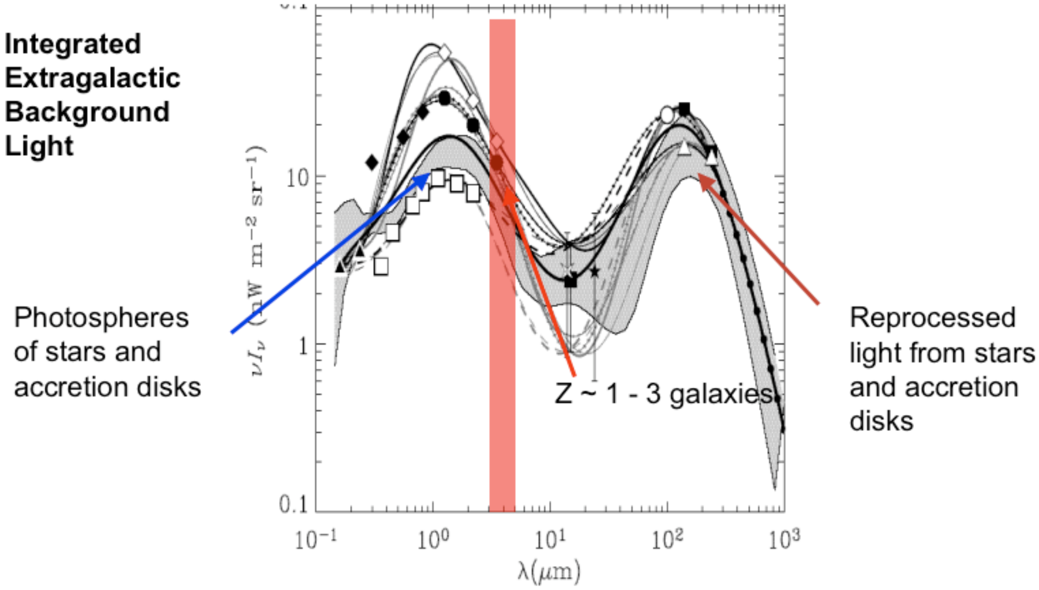

FIGURE 2. The integrated extragalactic background light in the UV to mid-IR range. The spectral range sampled by IRAC channels 1 & 2 is shown as the shaded region. The UV/visible and mid-IR peaks in the background arise from direct and reprocessed emission, respectively.

4.2 Studies of Old Galaxies at Intermediate Redshifts

The sensitivity of the IRAC bands to evolved stellar populations is illustrated in Figures 3, 4, and 5. For redshifts between about 1 and 3 the IRAC bands sample the broad H- induced peak in the spectral energy distributions of galaxies with ages greater than roughly 1 Gyr. This allows one to discriminate red galaxies at z ~ 1.5 – 3 from the foreground population in moderately deep IRAC images with ease. While K-band imaging has been used to identify red galaxies to z ~ 3 extremely long exposure times on 8m telescopes are required. IRAC can detect such sources with greater contrast in far shorter exposures.

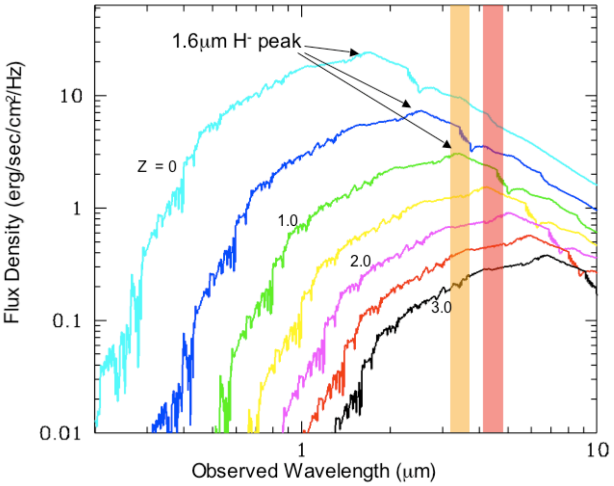

FIGURE 3. Spectral energy distributions, in F_ν units, of evolved galaxies at redshifts from 0 to 3. The IRAC channel 1 and 2 band-passes are shown as shaded regions At redshifts between ~ 1 and 3 the IRAC bands sample the peak of the spectral energy distribution for stellar populations with ages greater than about 1 Gyr.

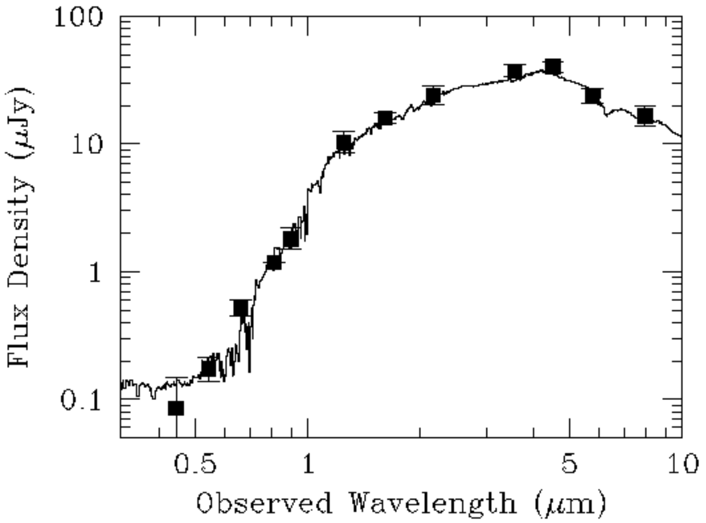

FIGURE 4. An example spectral energy distribution for a massive red galaxy at z = 1.5. The twelve bands cover B (0.4μm) through IRAC channel 4 (8μm). The broad peak of the spectral energy distribution at a rest-frame wavelength of 1.6μm falls in the IRAC 4.5μm band at this redshift.

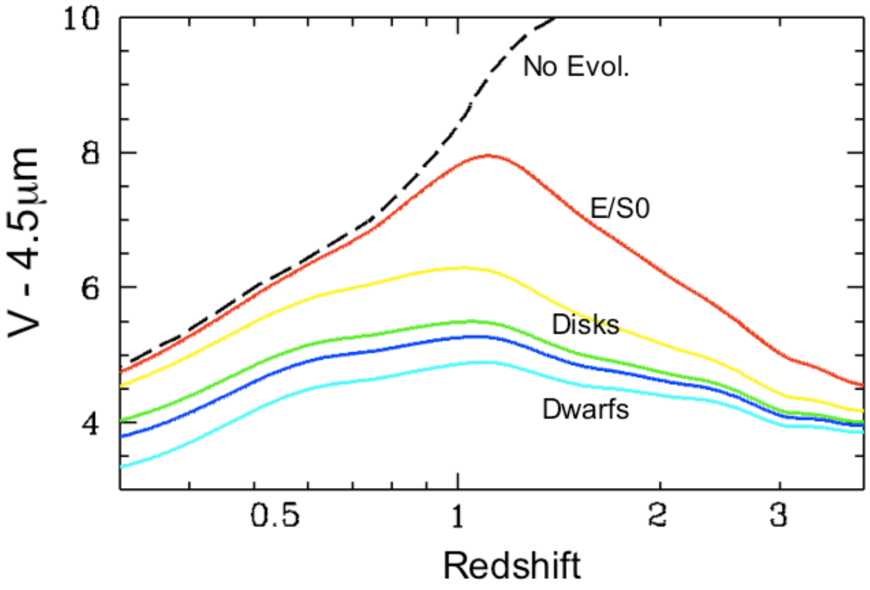

FIGURE 5. Optical V-band minus IRAC Channel 2 colors versus redshift for a variety of evolving galaxy models. The dashed black line shows the color trajectory for a non-evolving 13Gyr stellar population for reference. The other tracks show the color evolution for galaxies with exponentially decaying star formation rates with e-folding times of 1, 2, 3 and 5 Gyrs along with a constant star formation rate model. These are intended to illustrate the color evolution of elliptical/S0 galaxies, disk galaxies and dwarf galaxies in order of increasing e-folding times. The V-4.5μm color is a particularly powerful discriminator for early type galaxies at z ~ 1-2.

4.3 Galaxies at $z \sim 6$ and Beyond

One of the more surprising results from deep IRAC imaging programs was the detection of continuum emission from galaxies at $z > 5$. At a redshift of 6 the Universe is just under 1Gyr old. One might naturally expect that galaxies at this redshift would be extremely young and thus should have essentially flat (in F_ν units) spectral energy distributions. Most of the currently know galaxies at $z \sim 5$-6 have flux densities of ~200-300nJy in the visible and near-IR and thus should be below the detection limit of even the deepest IRAC images. Thus it came as a surprise that many of the $z \sim 6$ galaxies in GOODS, the Hubble Ultra-Deep Field and other deep surveys were detected with IRAC. The very red K – 3.5μm and rather flat 3.5μm – 4.5μm colors reveal a strong spectral break, which has been attributed to the Balmer continuum break in stellar populations dominated by stars with spectral classes from early B to mid-F. The best fitting ages for these galaxies derived from spectral synthesis models have been reported to be ~500Myr and the inferred stellar masses are quite large, reaching to a few x $10^{10} M_{sun}$ and perhaps higher. In Figure 6 I show an example from Eyles et al. [1], but there are also examples shown in Yan et al. [2] and others. The examples from Yan et al. and Eyles et al. are drawn from samples with high confidence redshifts based on Lyα emission. Mobasher et al. [3] and Wiklind et al. [4] have identified a number of candidate galaxies at $z \sim 6$ and higher appearing to have very large stellar masses. Secure spectroscopic redshifts are not yet available for these objects and it is possible that some may be galaxies at lower redshifts with unusually red colors.

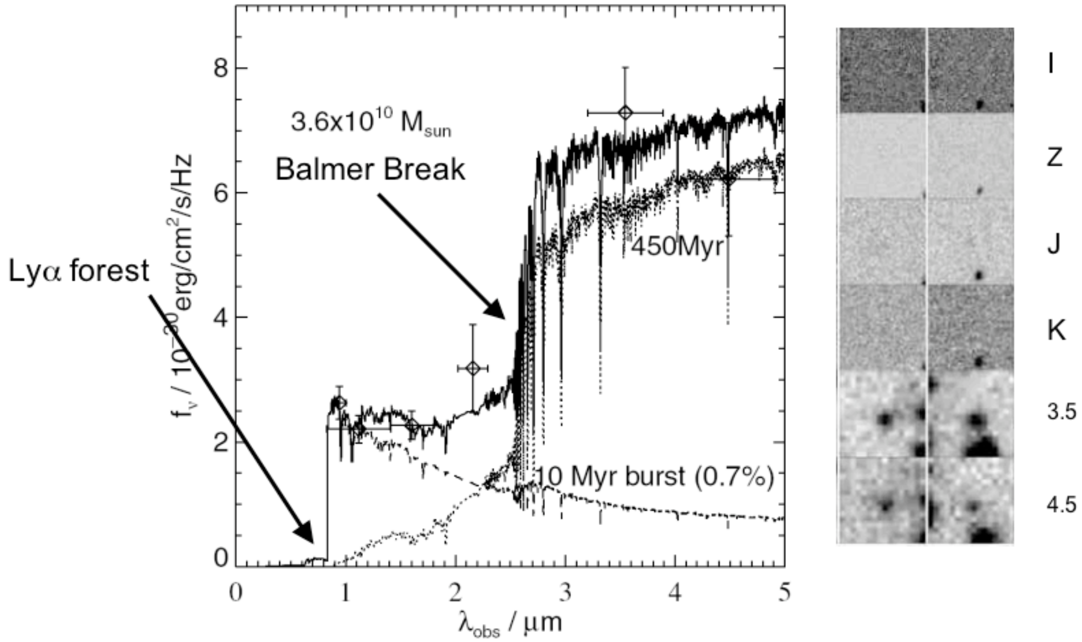

FIGURE 6. The spectral energy distribution of a $z = 5.7$ I-band drop out from Elyes et al. [1]. The grey-scale images on the right show detections in IRAC bands 1&2 and non-detections in all of the shorter wave-bands. The best-fit model SEDs is combination of a massive intermediate age population and a recent star burst. The total stellar mass is estimated to be more than 3 x 10^{10} solar masses.

The precise implications of the large breaks in the z ~ 6 galaxies detected by Spitzer are unclear. Recent spectral synthesis models by Maraston et al. [5] show that AGB stars can make fairly significant contributions, even at rather short wavelengths for ages of ~ 0.3 – 1Gyr. These models may lead to somewhat younger ages and significantly lower stellar masses for the z ~ 6 galaxies than first thought. Time will be needed before these issues are settled, but there is little doubt that Spitzer has provided a new and important view of galaxies in the early universe. It is particularly interesting to note that these discoveries arise entirely from channels 1 and 2 on IRAC, the capability that remain unchanged in the warm mission. Even in the cold state channels 3 and 4 lack the sensitivity required for these very faint objects.

4.4 Cool Stars and Hot Planets

The intrinsic L′ - M colors of main sequence and giant stars are of little diagnostic value as they are all near zero, being essentially Rayleigh-Jeans, for stars warmer than about L5. In the latest L dwarfs, T dwarfs and giant planets CH_4 absorption strongly impact the L-band while the M-band remains mostly clean. Thus L′ – M color provides a sensitive temperature indicator on its own. This can be seen in model atmospheres calculations (e.g. Burrows et al. [6]) as illustrated in Figure 7 below, and in the 3.6μm – 4.5μm vs. 3.6μm color magnitude diagram presented in Knapp et al., this volume.

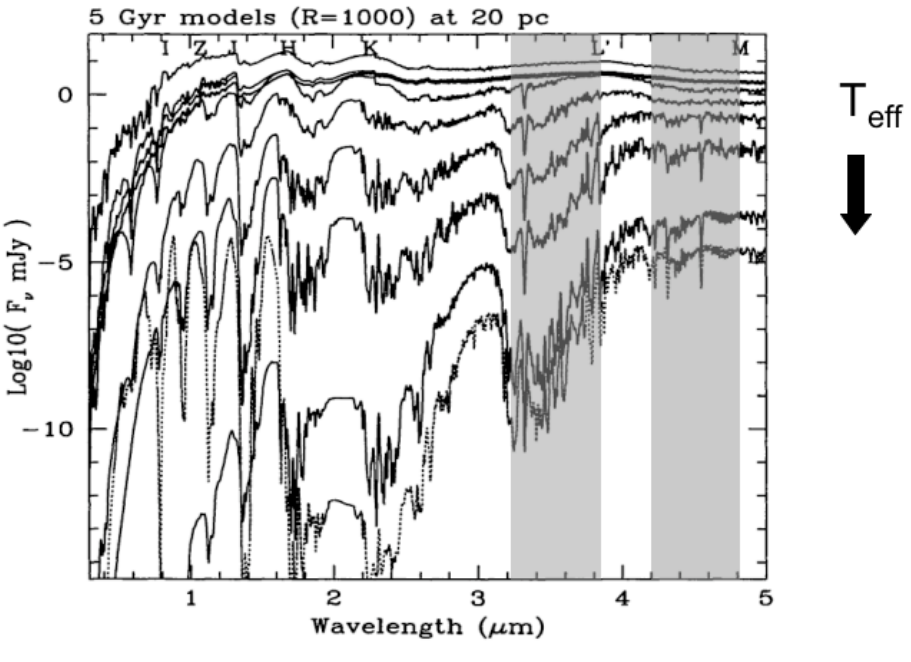

FIGURE 7. Infrared spectra from model atmospheres for old brown dwarfs and giant planets from Burrows et al. [6]. The IRAC channel 1 and 2 band-passes are shown by the grey shaded areas. The strong impact of methane absorption in the cooler objects and monotonic evolution of 3.6μm – 4.5μm color with temperature is evident.

The strong temperature sensitivity of the 3.6μm – 4.5μm color for cool objects suggests that an efficient survey for brown dwarfs and free-floating planets could be carried out with Spitzer in its warm phase. Knapp has suggested that this could be coupled to other imaging surveys at intermediate galactic latitude to leverage their scientific value. The WISE satellite has brown dwarf surveys as part of its core science mission. The faintest brown dwarfs in the WISE all sky survey will be single band detections. These will be confused with some number of artifacts, spurious detections and moving objects. There have been suggestions that Spitzer/IRAC channels 1 and 2 could provide efficient discrimination between genuine cool dwarfs and artifacts. This is an example of the scientific synergy discussed in the previous section.

4.5 Exoplanets

One might not think of an 85cm telescope as an ideal instrument for ultra-high precision photometry. Transit photometry is often a photon starved exercise and larger apertures on the ground benefit from both higher photon rates and reduced impact from scintillation. Spitzer, however, has become the instrument of choice for high precision exoplanet transit work. Knutson et al. [7] used IRAC at 8μm to measure the transit and secondary eclipse of HD 189733b with a precision of a few hundred micro-magnitudes with a time sampling of only 0.4seconds. This allowed them to determine the planet radius with a precision of 0.5% and map the brightness temperature distribution for the day and night side of the planet with uncertainties of only 10-30K. The great thermal stability of Spitzer is the key to its precision in repeated measurements. The heliocentric Earth trailing orbit of Spitzer frees it from the variable heating from Earthshine that impacts other spacecraft.

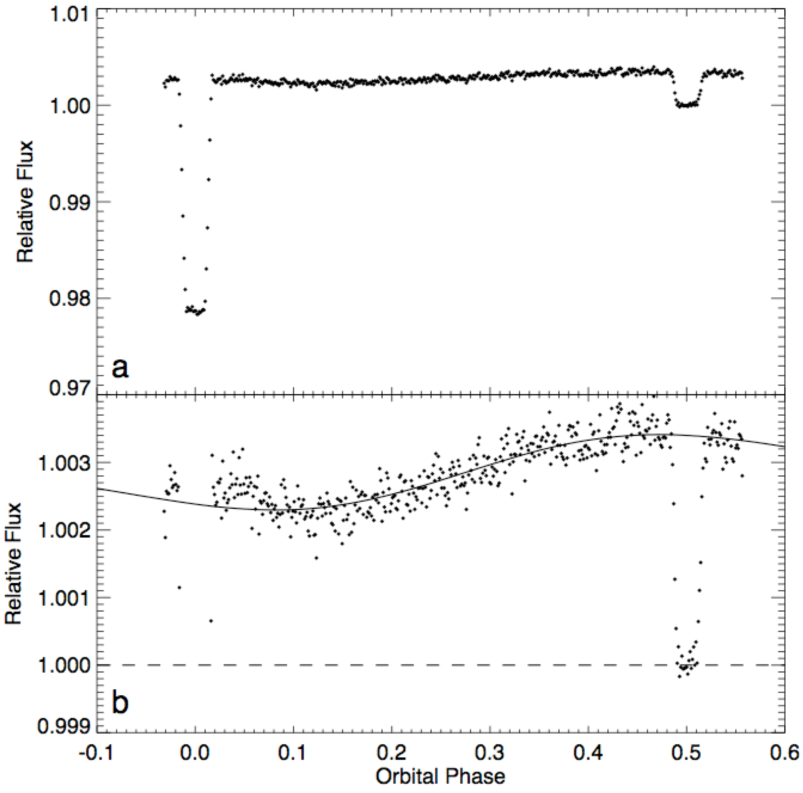

FIGURE 8. IRAC 8μm transit photometry of the exoplanet HD 189733b. The remarkable precision of the measurements (~ 500μ-magnitudes) allows not only detection of the primary transit and secondary eclipse, but leads to inferences regarding the distribution of brightness temperature on the face of the planet. Knutson et al. [7] used the slow variation in the flux between eclipses to derive a brightness temperature variation of ~ 250K due to heating of the planet from the central star. From Knutson et al. [7].

4.6 Active Galactic Nuclei

AGN are among the most broad-band, in terms of spectral energy distributions, of any discrete astronomical sources. The most extreme objects are detected from high-energy gamma rays to meter-long wavelengths. The wide range of physical scales and emission mechanisms in AGN are an important part of this broad spectral coverage. In Figure 9 I show schematic spectral energy distributions for AGN based on the composite energy distributions compiled by Elvis et al. [8]. The dominant contributors in the visible to mid-IR in broad-lined objects (type 1) are thermal accretion disk emission (the “blue bump”), non-thermal synchrotron emission and reprocessed thermalized UV continuum. In the narrow-lined or “type-II” objects the blue bump is either not present or, more likely, is highly obscured. If the type I and II objects are intrinsically similar but differ as the result of heavy obscuration, they should show the hot dust emission in comparable strengths. Thus the mid-IR is a prime testing ground for theories that unify broad and narrow-lined AGN.

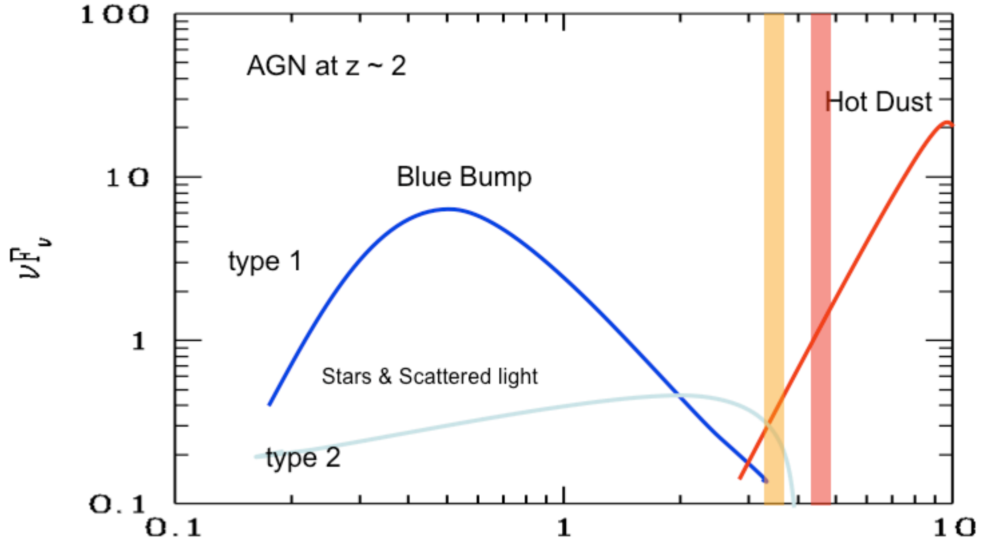

FIGURE 9. Schematic spectra of type I & II age in the visible and mid-IR. Type I, or broad-lined, objects show strong optical/UV emission that peaks near 1000A in the rest-frame, while Type II objects have weaker continuum that is often dominated by starlight. At mid-IR wavelengths both types of AGN show emission from dust with a range of temperatures. This emission arises from material heated by the central engine and provides a means of testing models that unify different types of AGN through orientation biases. The bands available on Spitzer in the warm era fall in on the short wavelength end of the dust spectrum for AGNs at z ~ 2, the peak of the quasar epoch. This figure is based on the composite spectral energy distribution for quasars compiled by Elvis et al. [8].

The spectral energy distributions shown in Figure 9 have been redshifted to z = 2, near the peak of the quasar epoch. At this redshift the IRAC channel 1 and 2 band-passes fall near the minimum between the direct and reprocessed emission and sample only the tail of the hot dust emission. IRAC channels 3 and 4 are better placed to sample the hot dust emission and a number of groups have used this fact to great advantage in identifying samples of highly obscured AGN. An example of the power of the four IRAC bands in separating obscured AGN from stars and galaxies in shown in Figure 10 from Stern et al. [9]. The two-color diagram in Figure 10 cleanly separates the broad-lined type-I AGN from other objects. The narrow-line type-II AGN are also separated, although not as cleanly as some fall within the region of color-color space occupied primarily by galaxies at modest redshift. Channels 1 and 2 alone would provide some discrimination, particularly between AGN and stars, but the galaxies and AGNs would be significantly more confused than in the case of the two-color approach using all four bands. It is likely that one could recover some of this sensitivity to obscured AGN during the warm mission by combining IRAC channels 1 & 2 with colors at shorter wavelengths where galaxies have spectral shapes that depart strongly from those of AGN. The drawback is that at shorter wavelengths obscured AGN are increasingly dominated by emission from the host galaxies and, particularly from the ground, are difficult to separate from normal galaxies. As one can see from Figures 4 and 9, ordinary galaxies have fairly flat 3.5 – 4.5μm colors over a wide range of redshifts, making redshift and spectral type discrimination difficult on the basis of these two bands alone.

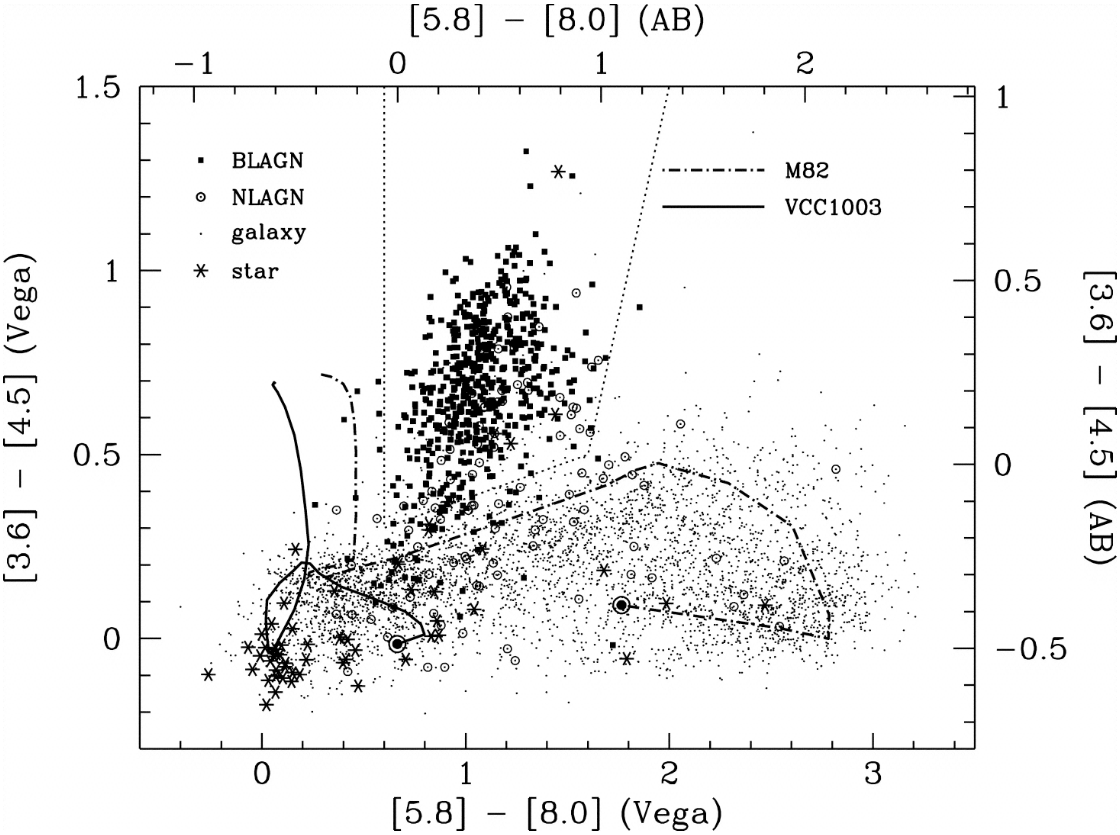

FIGURE 10. A two-color diagram comprised of IRAC channels 1 through 4 from Stern et al. [9]. The two-color approach cleanly separates galaxies at a range of redshifts from stars and AGN. The broad-lined type-I AGN very clearly separate from the galaxies while the narrow-line type-II objects are more widely spread in color, some falling within the galaxy locus.

The schematic spectral energy distributions shown in Figure 9 suggest that IRAC in its warm state is not all that well suited to studies of AGN physics. IRAS, ISO, and Spitzer have made important contributions to testing unification models and improving our understanding of the distribution of the obscuring material. Most of these advances have been made from observations at wavelengths beyond 5µm and from spectroscopy in the mid-IR. A good overview of the contributions of Spitzer to this field is provided by papers in the ApJ Spitzer special issue and the ASP conference proceedings from Spitzer workshops (Armus and Reach [10]). Similar arguments can be made regarding ultra-luminous galaxies where much of the interest is in using spectral diagnostics, such as PAH features, as probes of the energetics of the obscured sources.

5. WHAT SHOULD WE OBSERVE DURING THE WARM MISSION?

The previous section gives us some ideas regarding which areas of science are well suited to observations with Spitzer in its warm state. While there is little doubt that in its warm state Spitzer can still make important contributions to many, and perhaps all, areas of contemporary astrophysics, its fair to say that there are some areas that are more fruitful than others. Below I attempt to make a tabulation of astrophysical phenomenon that would make good targets for a warm Spitzer and those that are likely to be less compelling. There are undoubtedly other classes of objects that should be considered and exceptions to the broad classes listed below. Not surprisingly, the objects in the left column of table 1 tend to have their peak spectral output close to the IRAC 3.5 and 4.5μm bands, or have particularly favorable contrast compared to nearby objects (e.g. central star) at these wavelengths, while the objects in the right column have their peak output at shorter (e.g. hot stars) or at considerably longer wavelengths (e.g. ULIRGS). This harkens back to Figure 1 where we saw that the Spitzer warm bands fall between the direct photospheric emission and reprocessed peaks in the integrated background emission. The sources that will make good targets for Spitzer in the warm state either have been redshifted into this dark region of the spectrum or are just warm enough to naturally benefit from the reduced impact of hotter and cooler sources.

Table 1. What should we observe with warm Spitzer?

Good Targets	**Less Good Targets**
Exoplanets, Planets and small bodies	Hot stars
Cool stars and brown dwarfs	Star forming galaxies
Galaxy Clusters at $z > 1$	ULIRGS
Galaxies at $1 < z < 6$	Quasars & Seyferts (for AGN physics)

6. THE BIG QUESTIONS

Understanding which class of objects make the best targets is a good first step, but if we are going to make wise choices regarding the scope and structure of the science program in the warm phase, we need to identify key scientific question that can be addressed with precision imaging in the 3.5 and 4.5μm bands. Below I list ten of the big picture science questions that came out of the discussions at the workshop.

1. How are galaxies assembled from their constituent components of dark matter, stars and gas?
2. How does the halo mass distribution evolve as a function of time and environment?
3. When did the red sequence of galaxies form and how is it related to the collapse of groups and clusters?
4. What is the balance between star formation and accretion onto black holes in the global radiative luminosity density?
5. How are galactic disks structured, how are they built and what truncates them?

6. Do we understand the components and structure of the Milky Way?
7. What are the total current and past star formation rates in the Milky Way?
8. What is the stellar and sub-stellar mass distribution as a function of metalicity and environment?
9. How do disks around protostars evolve?
10. What sets the equilibrium radii of giant planets and how are they inflated?

The warm mission can make key contributions to addressing each of these questions by using one or more of its unique capabilities. I briefly consider each in turn.

How are galaxies assembled from their constituent components of dark matter, stars and gas? Spitzer is uniquely capable of weighing the stellar mass of galaxies. As shown in Figures 1 and 2, at 3.5 and 4.5μm IRAC samples the peak and long wavelength tail of the spectral energy distribution of red giants. The mass-to-light ratio in these wave-bands is close to unity and, more importantly, varies quite slowly with age. While there remains some uncertainty regarding the impact of AGB stars, this should ultimately be resolved leading to secure stellar masses and mass-to-light ratios for large stellar-mass-selected samples over a wide range of redshifts.

How does the halo mass distribution evolve as a function of time and environment? The great survey speed of IRAC and the potential for large observing programs during the warm mission offer the opportunity to determine the stellar ages, masses and spatial clustering amplitudes over a range of environments and epochs. The clustering lengths can be connected to the halo mass in a fairly robust manner (e.g. Giavalisco et al. [11]). The coupling of large redshift surveys with 3.5 and 4.5μm imaging surveys is a unique opportunity provided only by Spitzer via the legacy programs and future warm mission surveys.

When did the red sequence of galaxies form and how is it related to the collapse of groups and clusters? Understanding the origin of the massive and passive galaxies that form the red sequence today is a major goal of empirical and theoretical studies of galaxy evolution. The sensitivity to stellar mass and ability to survey large areas, as described above, provide an approach to overcoming the luminosity and large-scale structure biases that have impacted large surveys to date.

What is the balance between star formation and accretion onto black holes in the global radiative luminosity density? The classic Madau diagram provides a measure of the evolving global UV luminosity density. There are a number of uncertainties involved in converting this to a global star formation rate density. These include the impact of reddening and the unknown mix of contributions from hot stars versus AGN. Long-wavelength observations can help address both of these problems as they can recover the radiation that has been reprocessed by dust and, from the dust temperature distribution, distinguish between heating from stellar sources as opposed to AGN and disks with spectral energy distributions that extend to higher energies. The short wavelength channels on IRAC are particularly sensitive to the latter effect as they probe hot dust in the inner regions of accretion disks. IRAC studies coupled with deep and large area X-ray surveys can help determine the contribution of accretion to the global energy budget.

How are galactic disks structured, how are they built and what truncates them? Recent deep observations of the outer regions of galactic disks, particularly in HI and in the vacuum UV, have revealed that many disks extend much further than previously thought (e.g. Gil de Paz [12]). This has potentially profound impact on our understanding of how disk galaxies formed, how they evolve and how they are shaped by interaction with close neighbors. The importance of and interplay between tidal effects, disk flaring and ionization at the edges of galactic disks are not well understood. IRAC has outstanding sensitivity to low surface brightness features. The rather small aperture of Spitzer does not impact these surface-brightness limited studies. With its relatively wide field and fast survey speed the warm Spitzer will be a particularly powerful tool for this problem.

Do we understand the components and structure of the Milky Way? The big picture questions considered above treat galaxies as either test particles in large statistical ensembles or as integral systems to be probed from outside. Just as the theory of stellar evolution must be consistent with observations of the sun, our understanding of the structure and evolution of galaxies should be grounded in a firm understanding of the structure of the Milky Way. Our ability to produce a complete inventory of the contents of the galaxy, along with ages, compositions and dynamics, is hampered by our location in the plane. Long wavelength studies that penetrate the extinction (e.g. COBE) often lack the resolution required to study individual stars and clusters. The GLIMPSE program provided a powerful demonstration of the power of Spitzer to map in the inner region of the galaxy and the dense regions galactic plane. The outer regions of the galaxy and the vertical extent of the disk need further mapping, as both of these are critical to our understanding of the MW in the context of other galaxies.

What are the total current and past star formation rates in the Milky Way? It is now possible to reconstruct accurate and moderately precise star formation histories for the local group dwarf galaxies from color-magnitude diagrams. A detailed star formation history of the MW, while more difficult to produce, would be of great value. Star clusters provide a set of well-understood clocks and Spitzer could survey hundreds of these in the galactic plan, providing input for color-magnitude diagrams, age, reddening and distance determinations.

What is the stellar and sub-stellar mass distribution as a function of metalicity and environment? The slope of the bottom end of the stellar IMF and the transition to sub-stellar and planetary mass objects hold clues to the formation process and are vital to properly understanding the mass evolution of galaxies. Spitzer has unique capabilities to discover large numbers of cool low mass objects. The large gain in sensitivity compared to 2MASS or WISE opens an important niche for Spitzer. While lacking the full sky coverage of survey missions, Spitzer could nonetheless survey enough area to yield samples of hundreds of T-dwarfs and a few Y-dwarfs, as described by Knapp in this volume. Targeted observations of binary systems and clusters can also probe the coolest stars. Selecting clusters with a range of metalicities and densities should allow one to examine the role of environment in shaping the bottom of the IMF.

How do disks around protostars evolve? Disks play an important role in star and planet formation. Surveys of disks around young stars to date have been directed primarily at the youngest star forming regions. With the larger programs envisioned during the warm mission it should be practical to survey older star forming associations to assemble a more complete picture of disk frequency and evolution over a wider range of ages and derive a better understanding of disk lifetimes as a function of mass and environment.

What sets the equilibrium radii of giant planets and how are they inflated? Accurate determinations of the equilibrium sizes and gas giants is important not only for improving models of planet structure and formation, they also inform strategies for imaging exoplanets with the next generation of facilities. As we have seen above, Spitzer has an extraordinary power as a precise photometer. The number of known and potential transiting exoplanet systems should increase greatly in the next few years and the warm mission provides an opportunity for long stretches of uninterrupted observing campaigns to derive the structural parameters of both exoplanetary systems and giant exoplanets themselves.

6.1 Some Possible Key Science Programs

Throughout the course of the workshop we heard suggestions for a number of large programs aimed at key scientific questions. These were developed in varying level of detail in the pre-meeting white papers and most are discussed in these proceedings. I list several of these here in an attempt to provide an overview of some of the large programs under consideration and to set the stage for some of the discussion in the following sections.

- Complete survey(s) of the galactic plane
- Surveys of galactic open clusters
- Surveys of the structure and morphology of disk galaxies
- Studies of Exoplanet transits and eclipses
- Surveys of small bodies in the solar system
- Searches for T and Y dwarf stars
- Studies of IR excesses in white dwarf stars
- Ultra-deep survey of the end of the dark ages
- A Spitzer deep survey for galaxy and structure building
- Ultra-wide survey for galaxy clusters at $z > 1$

Some of these programs can be done in parallel, which is to say that multiple programs can be accommodated with a single data set. The T and Y dwarf survey, for example, could be combined with the galaxy cluster survey if an acceptable range of galactic latitudes could be agreed upon. Similarly, the galactic structure and open star cluster programs might be coordinated in a somewhat looser fashion that would still yield improved efficiencies.

7. OPPORTUNITIES FOR SURVEY PROGRAMS

One of the most attractive aspects of the warm mission is that its combination of restricted observing modes and high efficiency naturally lend themselves to large surveys. The largest Spitzer Legacy programs, while ambitious compared to typical GO programs, still fall short of enabling much of the science outlined at the workshop. The warm mission allows us to think in terms of thousands, rather than hundreds, of hours per program.

A number of ideas for survey programs have been suggested at this meeting. Most are aimed at moderately narrow science goals and nearly all are purely galactic or extragalactic in focus. It is potentially instructive to look at these as a whole and see what common threads can be found to link them into a few larger coherent programs. It is also instructive to consider where the proposed surveys lie compared to the Legacy and large GO and GTO programs. In Figure 11 I consider some of extant and proposed extragalactic surveys in a depth versus area projection. Existing programs are shown as dark rectangles, and suggested surveys are shown as light rectangles. The primary science thrust in various combinations of depth and area are shown as shaded ellipses. The range of surveys that are profitable is bounded on one side by efficiency considerations, as the spacecraft overheads become intolerably large for observations shallower than 5-10μJy. WISE will explore this region of parameter space efficiently. The all sky survey planned for WISE is expected to reach a 5σ point source sensitivity of ~ 100μJy at 3.5μm. At the other extreme one reaches the confusion limit somewhere below 100nJy where further integration fails to yield additional depth, or does so at a rate slower than √t. The sub-100nJy region of parameter space is a key part of the niche for JWST. A modest investment to push into the confusion limit with Spitzer might be useful both for immediate science return as well to aid in planning for observations with NIRCAM on JWST.

In Figure 11 I highlight regions of parameter space that address particular science areas. The boundaries are quite arbitrary and imprecise and there is significant overlap between adjacent areas. The ultra-deep survey discussed by van Dokkum in this volume would probe the stellar content of galaxies at $z > 5$ to a deeper level, and in larger numbers, than GOODS and other deep surveys. This program also directly probes the mass evolution of galaxies in the growth era from $3 < z < 5$ as well. The prime galaxy assembly epoch, from $1 < z < 3$, when the Hubble sequence appears and galaxies acquired many of the properties that differentiate them today is best probed with larger areas. Surveys of ~0.5 – 2 sq. degrees probe enough massive objects and a wide enough range of environments to allow one to address many of the critical issues related to feed-back and merging. The highest density regions can only be sampled with surveys covering several square degrees and large samples of massive clusters at intermediate and high redshift will require samplings on the order of 100 square degrees or more. Thus the three basic areas identified in Figure 11 map to the rough characteristics of the programs outlined in the contribution by the faint galaxy group in this workshop.

Comparing Figure 11 with Figure 1 shows that most of the surveys under discussion at the workshop require between 1000 – 2000 hours of spacecraft time. Survey more ambitious than these but potentially still well motivated, for example a 10 square degree survey to 350nJy to bridge galaxy assemble and large scale structure studies more effectively than the suggested 2 square degree survey, still require only ~ 10% of the warm mission duration.

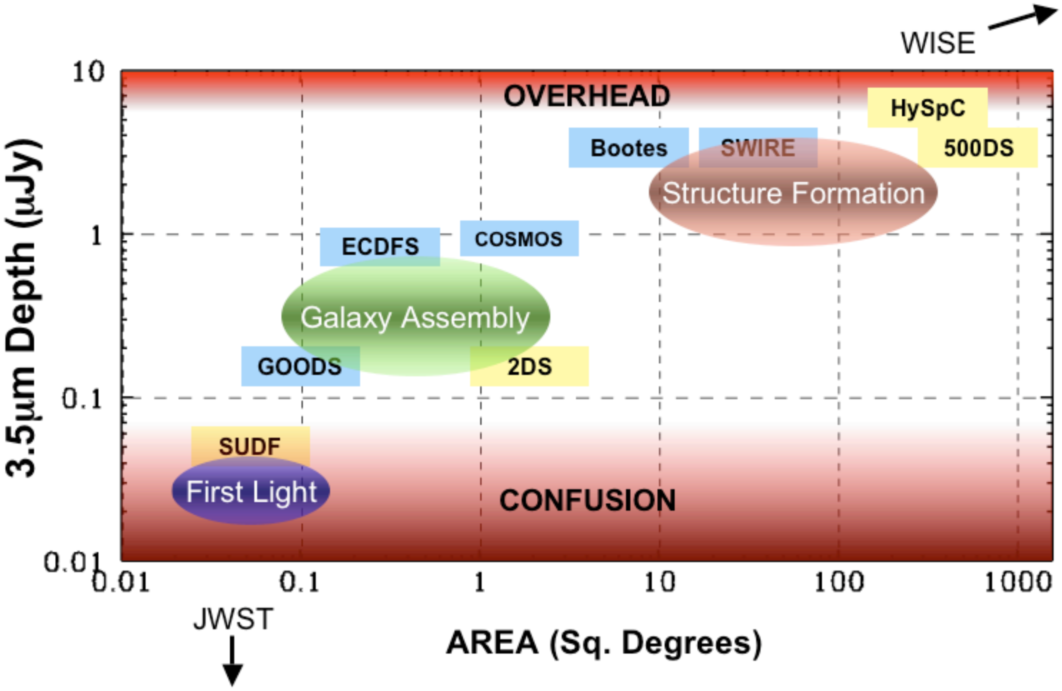

FIGURE 11. Depth versus area projection for current and proposed surveys with Spitzer/IRAC. The vertical axis is the 5σ point source limiting depth at 3.5µm in micro-Janskys. The rough limits imposed by spacecraft overheads and source confusion are shown at the top and bottom of the figure, respectively. Some extant surveys are shown as dark rectangles. These include GOODS, extended CDFS, COSMOS, NDWFS/Bootes and SWIRE. Straw-person surveys suggested at the workshop are shown as light rectangles. These include an ultra-deep field (SUDF), a deep two-square degree survey (2DS), a 500 square degree cluster survey (500DS) and an IRAC survey of the Hyper-Suprime-Cam Subaru/Princeton survey. Three basic science themes – first light, galaxy assembly and structure formation are shown as ellipses in the appropriate regions of the diagram. The suggested surveys typically require 500-2000 hours of spacecraft time, as can be seen by comparing this figure with the contours in Figure 1.

Comparing surveys in the galactic plane is somewhat more complex as all directions are not of equal scientific interest and the confusion limit can be a strong function of latitude and longitude. In Figure 12 I show the depths from a number of all sky or galactic plane surveys for wavelengths shorter than 10µm. This is adapted from a number of sources, including Bob Benjamin's contribution and Ned Wright's WISE web page. The GLIMPSE legacy survey has a depth well matched to 2MASS in the near-IR.

The UKIDSS survey of the galactic plane will cover 1800 square degrees to a K-band depth of 19.0 (Vega) and comparable depths in J and H. All of the plane easily visible from Hawaii will be covered over a period of ~ 7 years. This is quite a bit deeper than 2MASS. As discussed in the contribution by Benjamin in this volume, the UKIDSS survey is one of the motivations for a deeper 3.5 and 4.5µm survey of the galactic plane. WISE will have a raw sensitivity that is a factor of ~ 2 better than GLIMPSE, but the large pixels on WISE will make confusion a serious problem in high-density regions. The smaller pixels on IRAC, while still larger than one might like, offer a distinct advantage at the lowest latitudes. One suggested warm Spitzer galactic plane survey, GLIMPSE360, would reach depths of ~ 10 and 20μJy at 3.5 and 4.5μm, respectively.

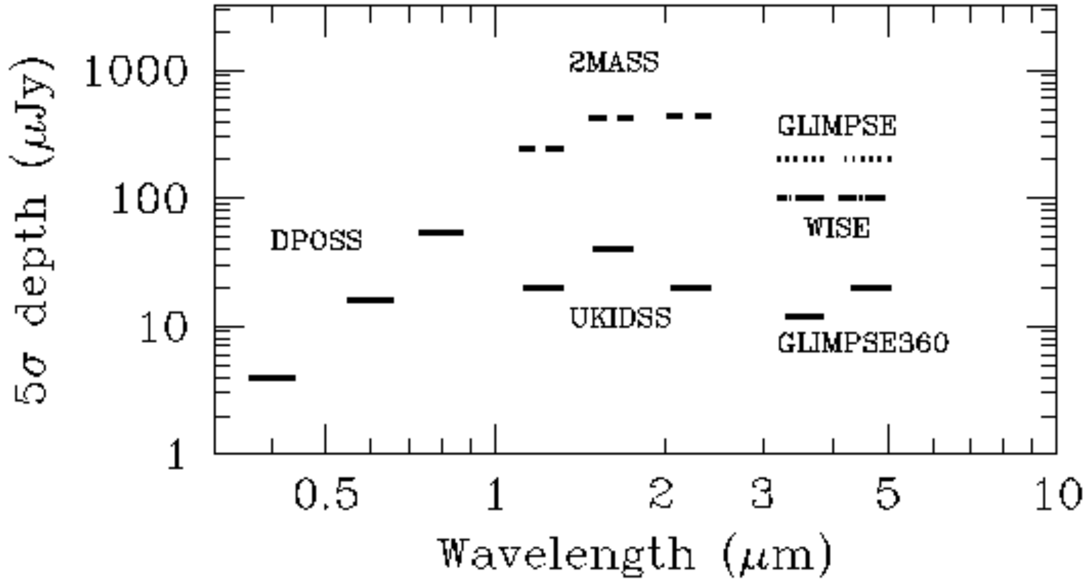

FIGURE 12. Depths of various all sky and galactic plane surveys at visible and infrared wavelengths below 10μm. The UKIDSS survey will miss key regions of the plane, as they are not within reach of observatories in Hawaii. WISE will provide full sky coverage, but crowding is likely to be an issue in the heart of the plane and near the galactic center. GLIMPSE360 offers a combination of depth, area, and sampling that is well matched to the next generation of optical and near-IR surveys. This figure is adapted from figures on the WISE home page and the contribution from Benjamin in this volume.

8. HOW SHOULD WE OBSERVE?

To this point we have considered what targets and combinations of area and depth might be of interest for large programs in the warm era. One of the goals of the workshop was to take a fresh perspective on how the community might plan to use the observing and data resources enabled by a warm mission in a more efficient manner. To date Spitzer has been used by the community at large in two rather different modes. In the Legacy Programs large teams carry out surveys with well-defined deliverables in the form of high-level data products and ancillary data. The teams produce scientific papers from these data, as does the community at large. These survey programs spawn follow-up programs and truly provide a legacy for the observatory. The Legacy programs have varied in size, the largest have been ~ 600 hours, more typical programs are ~100 – 200 hours in duration. One the order of ~ 1/3 of the telescope time have been devoted to these large programs. The remainder of the public time goes towards smaller programs that typically have more limited scope. The GO programs vary widely in scope, but many are 10-20 hours or less in duration. These lead directly to science papers and often to follow-up programs and observations. While data from the small programs are in the Spitzer archive, the PIs of small GO programs are not required to produce high-level data products for the community.

As we think about the warm mission we may wish to reconsider the balance between large and small programs. By the time the cryogen is exhausted Spitzer will have been in operation for roughly five years and many high priority small programs will have been carried out. Many of the important projects that will not been completed at that time will be those whose scope was outside even that of the legacy programs. Hence there is a strong argument for considering a scale of program that make use of the unique aspects of the warm mission while also reducing the level of support needed per hour of spacecraft time.

Many, if not all, of the most successful large survey projects in astronomy, and those involving observation from orbit, have arisen organically from single principal investigators or from a small group of scientists. The alternative, planning science from the top down by committee is fraught with risks. Meshing top-level science goals and grass roots science priorities is a fine balancing act that requires regular updating. This approach has been quite successful in setting national priorities for large missions and it may be helpful, at some level, in the case of the warm mission where a unique resource is available over a limited time interval.

8.1 Striking The Right Balance

One of the questions raised at the start of the workshop related to the balance between large and small programs. While I have attempted to make the case above for large ambitious survey programs, it is clear that there is no single prescription that works for all areas of science. It seems fairly clear that studies of distant extragalactic targets are best served by large and very large survey programs. Many of the outstanding questions require large databases and in recent years the power of large programs (e.g. SDSS, DEEP2) in breaking long-standing impasses in our understanding of galaxy evolution has been made abundantly clear. This remains true whether one defines large programs in terms of areal coverage (e.g. SDSS) or by shear number of hours thrown at a single deep field (e.g. Hubble Ultra-Deep Field). Both wide and deep surveys have important roles to play in the far extragalactic science.

The near-field extragalactic programs are also transitioning from studies of single objects to large statistical programs. The legacy programs aimed at nearby galaxies have been highly successful in large part because they allow one to identify underlying properties of galaxies that become clear with samples large enough to beat down the “noise” associated with the individual character of each object. Further advances in this area will require large, but not enormous, programs and may well work best in a staged approach where ~500 hours programs are planned and carried out in series and build on results and lessons learned as they progress.

Galactic programs have some of the characteristics of both the distant and near extragalactic programs cited above. Many of the interesting targets are single stars or star clusters that are widely distributed on the sky. While large samples are needed, it is not clear that a new paradigm for observing is called for. Galactic structure programs, on the other hand, typically require areal coverage that is large by the standards of extragalactic programs. A blend of survey programs aimed at mapping large contiguous areas, preferably in concert with similar programs at other wavelengths, and statistical samplings of individual objects is likely to be the way forward here.

Lastly, the planetary and exoplanet science programs appear to be working very well in the current mode. While one can argue that in the galactic and extragalactic areas we know fairly well how to best use the cameras in a technical sense and we have a fairly clear understanding of the appropriate targets, this may not be true for the exoplanet science. Techniques and target lists are evolving on a time scale that is short compared to the lifetime of the warm mission. One might argue that the timescale for the current proposal and scheduling process is a bit too slow for this field.

8.2 A Different Approach to Time Allocation?

Observatory scheduling is a slow and inefficient process. The time lapse between when a proposal is written and when observations are executed is typically 6–18 months. Each year some ~350 person-years are expended globally in writing and reviewing astronomical observing proposals that are rejected. This is appropriate because observing time is a scarce and expensive resource and science is a highly competitive marketplace of ideas. The expectation of declining support resources suggests that we might reexamine this process. If funds supporting the proposal process and those supporting technical activities are indeed fungible, we might think carefully about how to balance our priorities. It has been suggested that the proposal review process at Spitzer costs the equivalent of 2-4 FTEs and that this could be reduced by ~ 1-2 FTE without harm.

Most observing proposals are roughly even combinations of scientific justification and experimental design. Investigators describe and motivate the science they wish to address and layout a set of observations in terms of targets, wavelengths, filters, spectral resolution, and depths that they will use to advance their science. Proposal writing for Spitzer in the warm era will be a somewhat different process. There will be no choice of instrument, no options on filters, and only one spectral mode – R ~ 5 imaging at 3.5 and 4.5μm. The only experimental design choices left to the observers are: where to point the telescope and how long to expose. Reviewers will consider whether the proposing teams made a strong scientific case for their project and if they are pointing in a sensible place and exposing long enough. Technical considerations are likely to be of relatively minor importance in evaluating proposals. Rather, proposing teams are likely to stress their ability to extract science from the data efficiently and to bring ancillary data to bear on the problem in question. There will likely be multiple teams proposing to carry out the same basic science, much of which is discussed in some detail in this volume. It would not be difficult to image that reviewers will be faced with numerous proposals for which the only substantive difference is the composition of the proposing teams.

One could consider alternative approaches to scheduling Spitzer in the warm era. A number of key science areas could be identified as high priorities for the observatories and some guidelines issued for programs that address these. An open call for proposals would invite interested parties to propose 1) where to point, 2) how deep to image, and 3) which ancillary data are useful to addressing the particular problem. The teams could be invited to deliver high-level data products and tools to the Spitzer archive in exchange for financial and technical support from the SSC. This process could enable coordination with large programs and data sets at other wavelengths, both on the ground and in orbit. Naturally individuals or teams would be allowed, and encouraged, to propose for programs outside of the identified key science areas.

Much of the effort described above and the person-years of effort devoted to it involve relatively small requests for telescope time. The typical observing proposal for an 8m class telescope on the ground, or one of the great observatories in orbit, requests ~ 10 - 20 hours of actual on-source integration time. Most of the effort involved in processing and reviewing observing proposals is devoted to these small programs. In the case of HST, for example, roughly 90% of the effort is spent reviewing the small and medium programs. This is sensible for an observatory with several functioning instruments each of which has multiple modes. It may not make sense for Spitzer in the warm era.

One alternative approach is to decouple the review process for small and legacy class proposals. Programs below a predefined threshold, say 50 hours, could be reviewed by a standing committee that need not meet in person. This proposal track could operate on a different cadence than the large proposals and could be more responsive to rapid developments in fast moving fields. Staggering the large and small programs could also enhance pilot programs and follow-up programs in a way that improves the success rate and impact of the legacy class programs. The large proposals would continue to be reviewed by the external panels that meet face to face on a roughly annual basis.

9. SUMMARY

The scientific potential of Spitzer post-cryogen is not only clear, it is exciting and has the potential to energize the community by enabling a scale of projects that is usually only possible with dedicated missions. The science enabled by large programs using IRAC channels 1 & 2 is well matched to our community's priorities and offers important synergy with current and future facilities. The primary challenges may lie in choosing which of the many interesting programs to carry forward, maintaining the balance between large and small programs, and finding the most efficient way to operate the facility without compromising the vital competitiveness that keeps our field alive and vital.

ACKNOWLEDGMENTS

I'd like to thank Lisa Storrie-Lombardi and her team for organizing and supporting such an interesting workshop. I also thank all of the members of the steering committee and contributing authors for their dedicated effort in producing the pre-meeting white papers and the contributions to these proceedings.

REFERENCES

1. L. Eyles, A. Bunker, E. Stanway, M. Lacy, R. Ellis, M. Doherty, *MNRAS*, **364**, 443-454, (2005).
2. H. Yan, M. Dickinson, M. Giavalisco, D. Stern, P. Eisenhardt, H. Ferguson, *ApJ*, **651**, 24-40, (2005).
3. B. Mobasher, et al. *ApJ,* **635,** 843-844, (2005).
4. T. Wiklind et al., “Massive Galaxies at z > 5”, in *Galaxy Evolution Across the Hubble Time*, Proceedings of IAUS 253, Cambridge University Press, 2007, pp. 368-372.
5. C. Maraston, *MNRAS*, **362**, 799-825 (2005).
6. A. Burrows, D. Sudarsky, J. Lunine, ApJ, 596, 587-596 (2003).
7. H. Knutson, et al. *Nature,* **447**, 183-186 (2007).
8. M. Elvis, B. Wilkes, J. McDowell, R. Green, J. Bechtold, S. Willner, M. Oey, E. Polomski, and R. Cutri, ApJS, 95, 1-68 (1994)
9. D. Stern, et al. *ApJ*, **631**, 163-168 (2005).
10. L. Armus and W. Reach, *The Spitzer Space Telescope: New Views of the Cosmos”,* ASP Conference Proc. 357, 2006, Astron. Soc. Pacific.
11. M. Giavalisco, C. Steidel, K. Adelberger, M. Dickinson, M. Petteni, and M. Kellog, ApJ, 503, 543-552. (1998).
12. A. Gil de Paz, et al., *ApJ*, **627,** L29-L32, (2005).