5. DISCUSSION

One of the major preoccupation of systematics is to determine, by comparison, the unique properties of every species and higher taxon. It also aims to determine the properties certain taxa have in common with each other and the biological causes for the differences or shared characters. Biosystematics is given the rightful importance among biological disciplines, more because of the increasing need for a correct determination of generic and specific complexes, in view of the occurrence of large-scale diversities within species which make bioecological, behavioural as well as related interdisciplinary approaches obligatory for a better understanding of these complexes. Understanding of the differences in feeding strategies, differences in population distribution and success, differences in oviposition period and fecundity, and allometric growth are now emerging as important parameters for differentiating closely related species. Differences in reproductive physiology and feeding behaviour on the same host are major factors helping in the differential diagnosis of closely related taxa. The use of DNA sequence, morphological characters, and ecology to construct phylogenies for taxa of interest, and the use of this information to understand the processes of evolution is also gaining momentum in recent years. By building DNA and morphology based phylogenies, it is possible to identify new species and uncover hidden relationships and patterns between species. The primary goal of this thesis was to differentiate two closely related species of the lygaeid bug, *Spilostethus*, using a biosystematic approach involving morphological, ecological, biochemical and DNA based molecular markers and to develop a phylogenetic tree among the closely related species of Lygaeidae.
5.1 Morphological characters for distinguishing species

*Spilostethus hospes* and *Spilostethus pandurus* are commonly called as the milkweed bugs because of their close association with the milkweed plant *Calotropis gigantea*. Adults of both these bugs are orange red with bold black markings on the head and prothorax and the back has V-shaped markings. The antennae and legs are also black. However, the adults of the two species can be easily identified by the presence of a white circular dot on the membrane of the forewing of *S. pandurus* which is lacking in *S. hospes*. The adult bugs can also be distinguished by their sizes, *S. pandurus* being distinctly larger than *S. hospes*. Colour and body texture of the adults of the two species also are different, *S. hospes* appearing more soft-bodied and clean than *S. pandurus*. However, the problem of identification become explicit when dealing with the nymphs. The nymphs are gregarious and cause the maximum damage to the host plants warranting quick identification for the implementation of management strategies. This aspects makes this study relevant.

Studies on the immature characters of lygaeids facilitate distinguishing them from other Heteroptera as well as from one another and from different instars and from adults. Lygaeid nymphs vary greatly between species in size, colour and shape. There are a number of stable characters that reliably separate whole groups of genera and these remain practically unchanged during ontogenesis. Some of these characters are also shared with the adults. Nymphal characters such as trichobothria, abdominal segments and spiracular arrangements are clearly visible, which in adults are however difficult to be seen.
Studies on the immature characters of lygaeids have been particularly useful in identifying species belonging to a single genus. Mailloux and Streu (1981) used nymphal characters to differentiate six species of lygaeid bugs belonging to the genus *Blissus*. Slater and Baranowski, (1983) described the nymphs of 12 species of *Ozophora* using morphological attributes. In the present study, the two species of *Spilostethus* was distinguished by measuring 10 morphological characters namely the lengths of the foreleg, midleg, hindleg, thorax, abdomen, head, rostrum, antenna, total body length as well as the width of the head. The total body length of the fifth instar of *S. pandurus* measured 12.25 mm while that of *S. hospes* measured only 10.02 mm. Generally lygaeid species range in total body length between 1.5 mm to a little over 13 mm (Judd 1994). The fifth instars of the two species can also be distinguished by their shape; *S. pandurus* is typically oval in shape while *S. hospes* is a little oblong than oval. The shape of *S. hospes* is elongated in one direction having the chief axis considerably longer than the transverse diameter. Dorsally, besides the mesothoracic wing pads, *S. hospes* has two large and two small dark patches representing the dorsal abdominal glands while *S. pandurus* has only two dark patches. Shape is normally constant within subfamilies and related to trophic position (Slater 1976c) wherein the more elongate bodied bugs are ground dwelling and occurs in the litter layer. However *S. hospes* and *S. pandurus* are aposematically coloured and they do not hide in the litter. Their colour and markings are probably a direct consequence of shared selection pressures related to behaviour, habitat and host plant preference.

The morphology of the thorax facilitate distinguishing the two species of lygaeids. A well developed medial ecdysial suture running from the anterior of the pronotum to the apex of the scutellum is distinct in *S. pandurus*, but not very evident
in *S. hospes*. The pronotum is trapezoidal and wider than long in both the species. The mesonotal and metanotal pairs of wing-pads are fleshy and positioned close to each other. Also the nine abdominal segments of the nymphs of *S. hospes* is very distinct compared to *S. pandurus*.

Wing pads for both the species are first visible on the mesothorax in the third instar nymphs. They are longer than the length of the mesonotum in fourth instars, but the tips of the mesothoracic wing-pads do not extend beyond the tips of the metathoracic wing-pads as they do in fifth instars. Similar observations were reported for species of Lygaeidae by Dolling (1991) which make them different from the Coreoidea (Kumar 1966).

Nymphs of *S. hospes* and *S. pandurus* are aposematic, being brightly coloured like their adults and the characteristic adult markings are not evident in the nymphs. However subtle variations exist in the colour intensity and brilliance which facilitate differentiation of the two species. Red, Blue and Green are the major colour spectra. The intensities of these colours for the two species were measured at the region of the Head, Pronotum, Scutellum and Abdomen of the fifth instars of the two species. The study indicates that the intensity of these colours in *S. pandurus* was higher than in *S. hospes*. The combination of these three colours results in the Grey which could be measured using the formula \((R+G+B)/3\) or \(0.299R+0.587G+0.114B\) (Ferreira and Rasband, 2012). The grey values for *S. pandurus* was significantly higher than *S. hospes*. Distinct lygaeine nymphal generic colour patterns are clearly recognisable and are important in the separation of lygaeid genera such as *Horvathiolus* from *Melanocoryphus* (Judd 1994). However,
care must be taken to ensure that adult taxonomic mistakes are not replicated by the creation of a parallel colour identification scheme for nymphs. Colour markings are only used to substantiate morphological differences. Potential confusion in separating species by markings alone is highlighted in *Lygaeus equestris*, where nymphs exhibit intraspecific variation (Sillen-Tullberg (1985). Colour variation in lygaeids is probably a direct result of external stimuli on individual specimens or a genetic response. Climate, geography, population density, host plant association or a combination of factors may be important. Kugelberg (1973a) noted that lighter coloured adults resulted from increased nymphal density, but made no comment on nymphal colouration. Bright aposematic colour has an ecological benefit in terms of escape from predators through exhibition of warning signals. It has been experimentally demonstrated that red *L. equestris* nymphs have a 6.4 times higher rate of survival from bird attack than a mutant, cryptic grey form of this species (Sillen-Tullberg, 1985).

Compared to other species of lygaeids, the legs of *S. hospes* and *S. pandurus* are short. The arrangement of spines on the prothoracic femur assists in generic and species identification and therefore these details for *S. hospes* and *S. pandurus* were studied. The femora are unarmed in *S. hospes* while it is distinctly spined beneath in *S. pandurus*.

Analysis of the literature indicate full generic status for *Spilostethus* is convincing. Nymphal characters represent an intergrade between *Lygaeus* and *Melanocoryphus*. Meso-lateral longitudinal rows of black spots on terga 2-7 in *Spilostethus* and *Melanocoryphus* are absent in *Lygaeus*. However, unlike
Melanocoryphus, the gland apertures are similarly spaced to Lygaeus. The maximum width of the evaporative area in Spilostethus is larger than in Melanocoryphus relative to the eye width, but is significantly smaller than for Lygaeus. Adults of Cosmopleurus are structurally similar to Spilostethus, suggesting similarity, but the relative position of the dorsal gland apertures, which are close together in each pair in Spilostethus and apart in Cosmopleurus, confirm the generic separation.

5.2 Allometric characters for differentiating species

The subject of allometry is variation in morphometric variables or other features of organisms associated with variation in size. Such variation can be produced by several biological phenomena, and three different levels of allometry are therefore distinguished: 1) static allometry reflects individual variation within a population and age class, 2) ontogenetic allometry is due to growth processes, and 3) evolutionary allometry is the result of phylogenetic variation among taxa (Klingenberg 1996).

Structural changes especially those of the head size and appendage length are subjects of differentiation as the nymphs grow. Therefore mere measurements of the various morphological characters would not suffice for distinguishing species. The importance of allometry which involves studies of different body parts in relation to the whole across developmental stages become important while using it as a parameter to distinguish species. Allometric studies facilitates not only comparing species, but also for an understanding of the growth during ontogeny and evolution.
Allometric growth in *S. hospes* and *S. pandurus* was studied at all three levels viz., ontogenetic, static and evolutionary allometry. Ontogenetic allometry is the growth trajectory of an organ relative to body size during the growth of a single individual. Ontogenetic allometry of the two species of *Spilostethus* was studied by measuring the different body parts as the nymphs grew from the first instar to the fifth instar stage.

Static allometry is the scaling relationship among individuals between one organ and the total body size, or between two organs, after growth has ceased or at a single developmental stage. Static allometrical differences was assessed by comparing the body parts between the two species of *Spilostethus* at the final nymphal instar stage.

Evolutionary allometry is phylogenetic allometry that studies the size relationship between organs across species. Taxonomic errors result when intraspecific variations, such as ontogenetic, sexual dimorphic, or phenotypic differences, are mistaken for interspecific differences. This leads to over split taxa containing species based on minor and inconsistent differences. In order to study the size relationship between organs across species, the data measured for various body parts at the fifth instar stage of the two species of *Spilostethus* was compared to similar measurements of two other species of Lygaeinae namely *Lygaeus equestris* and *L. simulans*; two members of Orsillinae, *Orsillus depressus, Ortholomus punctipennis*; and Ischnorhynchinae, *Kleedocerys resedae* (Betula), *K. truncatulus*; and two members of Rhyparochrominae, *Rhyparochromus pini, Aphanus rolandri*. 

155
5.2.1. Ontogenic allometry

Measurements of ten morphological characters were recorded for both the species of *Spilostethus* during development from the first instar stage to the adult stage while feeding on the common host *C. gigantea*, since host plants are known to have a bearing on the growth of lygaeids (Ananthakrishnan *et al.*, 1982, 1983; Sanjayan and Ananthakrishnan, 1987). The data was subjected to univariate, bivariate and multivariate analyses. A trend line was constructed depicting the growth of each morphological character during the nymphal development. The slope of these lines varied among the ten morphological characters indicating that the growth of all the body parts was not in proportion. The hind leg length of both the species of *Spilostethus* had the maximum slope while the slope of the best fitting line for the head width was closer to the x-axis.

Bivariate analyses was carried out to examine the relations between pairs of variables during growth. The results demonstrated that specimens define a single rectilinear trend line for every bivariate plot. The projection of these trends passes close to the origin of both axes, suggesting that growth relationships approach isometry. For both the species, the allometric growth of different body parts in relation to the TBL progressed at the same rate with the slope of the line increasing in the following hierarchy: HDW, HDL, THORL, RSL, ABDL, ANTL, FLL, MLL, and HLL. Therefore the slope of this bivariate relationship does not appear to be useful in delimiting the species. Some growth relationships show very tight clustering about the trend line; others show a much more scattered distribution, and hence lower correlation coefficients. Wide scatters indicate that the growth controls of some characters are more flexible than of others. For both *S. hospes* and *S.*
Pandurus, very tight clustering about the trend line was observed. The coefficient of determination (R^2) is a measure of the proportion of the variation of one variable that is determined by the variation of the other. Strong positive values for all the coefficients of determination was observed suggesting that overall size is the principal variation within the data set.

For morphological characteristics, allometries can be best visualized as plots of the size of a trait against the size of the body. When these plots are made from measurements of conspecific individuals at the same life stage, the relationship is called a static allometry. Even a cursory survey of static allometries reveals considerable variation in their slope. Slopes vary between species for the same trait and between traits for the same species (Shingleton et al., 2007). Often, morphological traits scale proportionally with the body, a condition called isometry, so that the relative size of the trait is independent of body size. However, slopes can be very steep, such that traits become relatively larger with increasing body size, or very shallow or flat, such that traits become relatively smaller with increasing body size. In the present study, the slope of the line for the HDW was very flat, while the slope of the line for the HLL was very steep in both the species of Spilostehtus. Slope can even be negative, such that traits become absolutely smaller with increasing body size (Shingleton et al., 2007). None of the traits measured in the present study had negative values. The shape of allometries are often modelled using the allometric equation, which can be applied to traits that scale linearly on a Log – Log scale. However, static allometries need not be linear on a Log – Log scale, or even linear on any scale. They can be sigmoidal or discontinuous, depending on the trait, the species and the unit of measurement. (Emlen and Nijhout 2000). Clearly all the traits studied here showed only linear relationships.
Another outcome of the bivariate allometric studies is with relevance to the corresponding growth rate of different body parts of the five instars in relation to the total body length. The data obtained in the present study indicates that generally the maximum growth of the instars took place during the first three instars. Also growth during the third to fifth instar did not progress equally as with the earlier instars.

Many texts of multivariate statistics introduce principal component analysis (PCA) as a technique for summarizing most of the variation in a multivariate data set in fewer dimensions (Pimentel, 1979; Jolliffe, 1986; Flury, 1988; Flury and Riedwyl, 1988; Johnson and Wichern, 1988; Jackson, 1990; Jobson, 1992). The first principal component (PCI) is the linear combination that accounts for the maximum variance. Geometrically, it corresponds to the direction of the longest axis through the scatter of data points. Subsequent principal components take up maximal variance, subject to being orthogonal to all preceding component axes. Principal component analyses (PCA) was performed combining the data of both the species to elucidate the components accounting for as much as possible of the variance in the multivariate data. These new components are linear combinations of the original variables. PCA was used for reduction of the data set to only two variables, the first two components, for plotting purposes.

Through the PCA routine the eigenvalues and eigenvectors of the variance-covariance matrix or the correlation matrix was computed. The eigenvalues provides a measure of the variance accounted for by the corresponding eigenvectors (components). Through these analysis it appears that the total body length and the length of the hind legs are important characters for comparison.
Two other multivariate techniques were used to analyze each of the two species. The first technique was a discriminant analysis (DA), a classification technique that establishes whether each of the species could be statistically recognized. The second technique, non-metric multidimensional scaling (MDS), is an exploratory technique to assess the morphospace occupied by each variable in multidimensional space. Both techniques assume data normality. Discriminant analysis can be used for visually confirming or rejecting the hypothesis that two species are morphologically distinct. Using a cut off point at zero (the midpoint between the means of the discriminant scores of the two groups), a classification into two groups was performed. A 83.33 percentage of items were classified correctly. Most of the overlaps were seen with regard to the I instar nymphs of the two species. Equality of the means of the two groups was tested by a multivariate test, similar to the t-test, called Hotelling's T-squared test. The Hotelling's $T^2$ value was 83.042 with F value of 7.236 and $p = 2.075 \times 10^{-7}$. These values show significant difference between the two species of the lygaeids.

The non-metric multidimensional scaling was also carried out with a view to visualize the morphospace occupied by each variable for the two species of *Spilostethus* in a multidimensional space. This data also pointed out to a difference between *S. hospes* and *S. pandurus*.

Fitting a line to a bivariate cloud of data would seem a relatively simple and fundamental procedure in data analysis. However, there has been lively debate in the literature concerning which method is appropriate in what situation (Ricker, 1973, 1982; Jolicoeur, 1975, 1990; Sprent and Dolby, 1980; Sokal and Rohlf, 1995; Carroll and Ruppert, 1996), and some of the issues discussed have never completely
been resolved. Authors have offered distinctly different reasons for using one method instead of another (Sokal and Rohlf, 1995; Carroll and Ruppert, 1996), and have advocated different methods (McArdle, 1988; Isobe et. al., 1990; Jolicoeur, 1990). Allometry is a discipline in which alternatives to linear regression are routinely required, because lines are usually fitted to estimate how one variable scales against another, rather than to predict the value of one variable from another. Since in both the species of Spilostehtus, the slope of the allometric growth line for the different traits was more or less similar, they could not be used for the delimiting of the species. Therefore further statistical analysis was necessary to see if the elevation of the slope could be used for distinguishing the species based on allometry. Therefore, in addition to the Ordinary Least Square (OLS), the Major Axis (MA) and the Standardized Major Axis (SMA) calculations were carried out. The OLS is used only when there is a need to predict the value of one variable when the other is known which was not the objective in the present study. On the other hand, if the statistic of primary interest is the slope, then MA or SMA is appropriate rather than linear regression (Warton et al., 2006). In terms of efficiency, Isobe et al., (1990) and Jolicoeur (1990) favours the use of SMA since SMA lines are estimated with greater precision (standard error of the slope is smaller). Anderson (1984) and Jolliffe (2002) advocates MA for some complex problems. In practice, these two methods give similar results if the variances of the two variables are similar (within a factor of 1.2) or if correlation is high, in which case it does not actually matter which method is used. In fact, the methods are identical for tests of whether a slope is equal to ±1 or not, which is commonly the test of interest in allometry.
5.2.2. Static allometry

The maximum nymphal growth is attained during the final instar stage while at the adult stage no further growth is possible. Therefore, these two stages were selected for studying the patterns of static allometry for the two species of *Spilostethus*. The scaling relationship among individuals of a stage was studied between the total body length and measurement of one of the several morphological measurements. The SMATR freeware programme was used for fitting bivariate lines to data and for making inferences about the lines. The line was fitted using the standard major axis (SMA), major axis (MA) and the ordinary least square (OLS) regression techniques. Values of the slope $\beta$ in both the MA and SMA methods estimates the line best describing the bivariate scatter of the two variables and also tests if it equals to 1. The results with the fifth instar data indicate that the scaling exponent between the TBL and other body parts were $>1$ and the values were significant for the Mid leg length, Head length and Rostrum length. The strongest positive allometric scaling was observed between the TBL and the head length as well as between TBL and rostrum length.

Analysis of the scaling relationship between the TBL and other body parts of the adults of *S. hospes* and *S. pandurus* showed high correlation only for the Mid leg length and the hind leg length with the TBL for *S. hospes* while for *S. pandurus* only the thorax length showed significant correlation with the TBL. The scaling exponent between TBL and Foreleg length only was $<1$ while all the other variables showed values $>1$. A strong negative allometric scaling was observed between the TBL and Abdomen length of *S. pandurus*. Teuscher *et al.*, (2009) studied the allometry between leg and body length of insects and opined that leg length is not only
affected by the rugosity of the environment, but also by functional adaptations, phylogeny, lifestyle, the type of insect development whether holo or hemimetabolous and constrains of gas exchange. Kaspari and Weiser (1999) predicted isometry between leg and body length of flying insects, although it is not obvious to us why evolution should \textit{a priori} favour isometry. In fact, a perusal of the literature shows that deviations from isometry dominate most scaling relationships (Peters, 1983).

Both the SMA and the OLS analyses indicated that the fitting of a common slope was justified as the p value was greater than 0.5 making room for further continuation of the analyses to check for whether the fitted slopes share a common elevation. The results indicate that the SMA data for the Head length did not share a common elevation for the two species of \textit{Spilostethus}. Therefore the head length appears to be a useful parameter for distinguishing the two species. Since the results for the common slope and shift in elevation between the groups were justified for all the other parameters, the test for shift along common slope was done and the results indicate there was significant shift along the common slope for the lengths of foreleg, mid leg, hind leg and rostrum length.

The length of the wing buds of the \textit{S. hospes} and \textit{S. pandurus} were significantly different to separate the species since the slopes were different (Shift A). The length of the thorax and abdomen had a shift in elevation although the slopes were common (Shift B). The data of the lengths of the rostrum and antenna had a shift along common slope (Shift C). The lengths of the forelegs, midlegs and hindlegs had a shift along the common slope as well as a shift in elevation (Shift D) and therefore the best parameter for delimiting the two species.
5.2.3. Evolutionary allometry

Evolutionary allometry reflects covariation among changes in different traits along the branches of a phylogeny. It is concerned with character co-variation among contemporaneous species sharing a common ancestor or among fossil members of an evolutionary lineage. Evolutionary processes leading to the associations between trait changes presumably do not differ depending on whether these changes occur within one lineage successively or in different lineages giving rise to sister groups. It is important to use specimens in comparable ontogenetic stages to avoid confounding evolutionary and ontogenetic variation. This is straightforward in organisms with determinate growth, such as birds or insects (Livezey, 1989; Klingenberg and Zimmermann, 1992a). Besides *S. hospes* and *S. pandurus*, the body measurements of two other species each of Lygaeinae, Orsillinae, Ischnorhynchinae, and Rhyparochrominae were compared for an analysis of the changes in body growth during the process of evolution. Measurements of the fifth instar stage of these species were used in performing the cluster analysis. The cladograms of the evolutionary tree drawn based on multivariate morphometric data using paired algorithm and Euclidean similarity measure indicated that the two species of *Spilostethus* were clubbed together and were distinctly separated from the other two members of the subfamily Lygaeinae. Members of the Orsillinae and Ischnorhynchinae formed a clade among themselves, while the Rhyparochrominae members were grouped with *Lygaeus equestris* and *L. simulans*. 
5.2.4. Dyar's rule

Dyar's rule assumes a geometric progression of size measures, where succeeding growth ratios (i.e. postmoult size/premoult size; also termed Dyar's coefficient) or percentage increments are constant (Dyar, 1890; Hutchinson and Tongring, 1984; Sehnal, 1985). Originally, such relations were used to determine the number of instars (e.g. Dyar, 1890; Gaines and Campbell, 1935). In most of these studies, the width of the head capsule was taken as a measure of larval size (Cole, 1980). However, head capsule width can increase slightly, but significantly, within stadia (Bliss and Beard, 1954) or even exhibit apparently continuous growth (Hemmingsen, 1965). Therefore, some authors included a variety of measurements of rigidly sclerotized structures which lack significant growth within stadia (Hemmingsen, 1965; Brown and Davies, 1972). Several authors already mentioned the relation between Dyar's rule and allometry (Teissier, 1960; Matsuda, 1964; Hemmingsen, 1965; Brown and Davies, 1972). However, Dyar's rule is not simply a special case of allometry (contra Matsuda, 1964), but covers an entirely different aspect of growth. Allometry pertains to the relations between different body parts, whereas Dyar's rule describes the stepwise growth of any single part in successive moults (or, alternatively, of a multivariate measure of general size). In other words, allometry indicates the ontogenetic trajectory, i.e. the path in morphospace along which an organism moves during ontogeny, whereas Dyar's rule is concerned with the discrete steps of growth along that path.
Dyar's rule, the constancy of the ratios of postmoult/premoult size measures in successive larval ecdyses, was not strictly fulfilled for both *S. hospes* and *S. pandurus* although growth ratios varied only within a limited range. Similar results were reported in earlier studies on various insects (Brown and Davies, 1972; Sehnal, 1985; Loudon, 1988; Savopoulou-Soultani and Tzanakakis, 1990), and also in crustaceans (Freeman, 1990). Hutchinson and Tongring (1984) argued that Dyar's rule might result from a maximization of growth efficiency, assuming that the size of the first instar, the number of instars and the arithmetic mean of growth ratios are predetermined. However, the approximate constancy of growth ratios can as well be seen as resulting from the physiological base of moulting (Nijhout, 1981; Sehnal, 1985). In a series of experiments on the bug *Rhodnius prolixus*, Bennet-Clark (1971) showed that the old cuticle plays an important role in determining the size of the successive instar, and he argued that the cuticle works as a template for the formation of the epicuticle of the following instar (Freeman, 1990). If successive moults are controlled by the same physiological mechanism, constancy of growth ratios will result. Therefore, it is preferred to view Dyar's rule as a base of comparison against which specific adaptive hypotheses can be tested, rather than to search for an adaptive explanation for the rule itself, as Hutchinson and Tongring (1984) attempted to do.

### 5.3 Biological characters

*S. hospes* and *S. pandurus* are polyphagous and mainly frequent the open flowers and seeds of the milkweed plant *Calotropis*. Laboratory observations on feeding showed a significantly higher percentage of the insects of *S. pandurus*
preferring to feed on the seeds than *S. hospes*. The number of insects of *S. hospes* that frequented the leaf and flowers of *Calotropis* was more than those of *S. pandurus*. There appears to be a resource partitioning among *S. hospes* and *S. pandurus* as far as utilization of the common host *C. gigantea* is concerned. Also the rate of post embryonic development was faster for *S. hospes* than *S. pandurus*. Ananthakrishnan *et al.*, (1983) studied the impact of different hosts plants on the growth and fecundity of *S. hospes* and indicated that these parameters were significantly effected. The present study appears to be the first observation made to compare the growth of *S. hospes* and *S. pandurus* simultaneously in the laboratory under similar conditions and report that the two species show variation in these biological parameters. Root and Chaplin (1976) observed *Oncopletus cingulifer* and *O. unifasciatellus* to occur on the same milkweed species, *Asclepias curassavaca*, with the adults utilizing the host in the same way except that *O. unifasciatellus* feeds somewhat more on seeds. Nymphs of *O. unifasciatellus* require milkweed seeds to develop whereas *O. cingulifer* can mature slowly while feeding on vegetative tissues of milkweeds.

In many phytophagous insects, different populations of a species may specialise on different food plants locally, potentially resulting in the formation of host races or biotypes (Mopper and Strauss, 1998; Funk *et al.*, 2002). As host race formation may be an initial step towards speciation, this topic has received much attention from evolutionary biologists. Several recent studies on food plant use of herbivorous insects revealed that fitness of insects can be enhanced by secondary endosymbionts. Symbiotic bacteria have been known to facilitate colonization of new feeding niches of insects, allowing specialisation on a broad range of diets.
Bacterial endosymbionts may either provide essential nutrients that are lacking in the diet or aid in digestion and detoxification of food. The two species of *Spilostethus* were examined for presence of bacterial symbionts in their gut. It was interesting to note that though both the species inhabit the same host plant in the habitat, the bacterial complement in the gut were uniquely different in both the species.

### 5.4 Ecological characters

*S. hospes* is distributed from Iran to Japan in the north to New Zealand and Tasmania in the south and is widely distributed throughout most of Australia (Slater 1985). The life history of *S. hospes* in China has been detailed by Hoffmann (1932). Hoffmann listed adults and nymphs being taken on *Solanum nigrum* L and adults feeding on *Physalis peruviana* and other hosts. The species has also been recorded on cotton, *Gossypium hirsutum* in Papua New Guinea (Ballard, 1927). Bergroth's (1916) record of *S.pacificus* (Boisduval, 1835) from Ayers Rock, Central Australia may refer to *S. hospes* (Slater, 1985).

*S. pandurus* is distributed in the Afrotropical regions, Australia, New Zealand, Papua New Guinea, Burma, China, Hong Koong, India, Indonesia, Malaysia, Pakistan, Philippines, Thailand, Iran, Japan (Slater, 1964).

Thangavelu (1980) observed that *S. pandurus* was abundant during the hotter months (March-August) and contrastingly, *S. hospes* was numerous during the colder seasons (January-February and October- December. He recorded maximum temperature to be the most important limiting factor governing the population level and distribution of the species. In the present study, both the species of bugs were
observed to be more prevalent during the hotter parts of year, with populations of *S. hospes* colonizing earlier than *S. pandurus*. Also the populations of *S. hospes* was more than that of *S. pandurus*. It was interesting to observe that the peak occurrence of insect population coincided with the peak in the availability of the pods. The number of pods available as food was significantly correlated with the population of the bugs. There appears to be a difference in the preference of individual plants of *C. gigantea* among the two species of *Spilostethus*. While *S. hospes* preferred colonizing on plants closer to residential areas of the habitat, *S. pandurus* preferred isolated plants. *S. pandurus* was more frequently observed on another weed, *Vernonia cinerea* than *S. hospes*. A behavioural difference was also observed between the two species of bugs. *S. pandurus* could be easily spotted on the host plants between 8-10 hrs in the morning and 16-18 hrs in the evening and were seldom seen during noon, while *S. hospes* could be observed in good numbers throughout the day and behaviourally, these bugs prefer feeding on shrubs which are below 4ft tall. Both the species of insects were observed feeding on the leaves, flowers and seeds of open pods. These bugs were also observed feeding on *Solanum lycopersicum* that was available in the vicinity.

**5.5 Biochemical characters**

**5.5.1 Cardenolides**

Three biochemical characteristics of *S. hospes* and *S. pandurus* viz. cardenolide content, protein profile of eggs and GST activity, were studied to differentiate one from the other. Both these species, being brightly coloured, have the ability to sequester cardenolides from their common host, *C. gigantea*. This
efficiency of sequestering the cardenolides was used as a criteria for species delimiting. It is known that lygaeids belonging to the subfamily Lygaeinae, utilize the cardenolides of host plants for self protection from vertebrate predators (Scudder and Duffey, 1972; Abushama and Ahmed 1976). They sequester these chemicals even during the early instar, although the maximum sequestering takes place during the final instar stage with very little during the adult stage. Therefore, by the time the insects reach the adults, they have successfully sequestered the cardenolides. Therefore adults of both the species were collected from the field and immediately processed for spectrophotometric determination of the cardenolide content. Distinct differences in the cardenolide content in the two species was observed, with the content being higher is \textit{S. pandurus} than in \textit{S. hospes}. Isman (1977) reported that several ecological parameters, such as differences in reproductive phenologies and morphologies of different host species, may interfere with the acquisition of cardenolides. In the present study, the two species of \textit{Spilostethus} were collected from the same habitat, so the ecological aspects gets nullified and differences in the cardenolide content in the two species could be attributed to the differences in the physiology of the insects.

\textbf{5.5.2. Egg protein profile}

Insect eggs represent a self-sustaining system which provides the raw materials for building the larval body and the energy reserves to cover all of the energy demands during embryogenesis (Sander \textit{et al.}, 1985). The two species of \textit{Spilostethus} also differed in the nature of the protein profiles of the just laid eggs. As many as 16 proteins that ranged in molecular weight from 6.43 to 184.12 were
recorded in the eggs, but differences in the individual protein profile was evident. Only 5 proteins were seen in both the species. Four high molecular weight proteins ranging from 70.91 to 115.93 were present in *S. hospes* that was not present in *S. pandurus*. Salkeld (1969) studied the egg protein pattern from 23 insect species representing seven order using polyacrylamide gel electrophoresis, and reported that the patterns were species specific and highly reproducible. Those for congeneric species were very similar and family resemblances were apparent. He regarded the egg-protein patterns to be useful in phylogenetic studies and in the analysis of closely related species. In the present investigation, it was apparent that the two species could be delimited by studying the egg-protein profiles.

### 5.5.3 GST activity

Glutathione S-transferases (GSTs) are a major family of detoxification enzymes found in most organisms. They help to protect cells from oxidative stress and chemical toxicants by aiding the excretion of electrophilic and lipophilic compounds from the cell (Hayes and Pulford, 1995). Interest in insect GSTs is focused on the role of these enzymes in insecticide resistance. Elevated GST activity has been detected in strains of insects resistant to organophosphates (Fournier *et al.*, 1992) and organochlorines (Grant and Hammock, 1992) and this enzyme family has recently been implicated in resistance to pyrethroid insecticides (Kostaropoulos *et al.*, 2001, Vontas, *et al.*, 2001). The continued use of insecticides to combat insect pests in agro ecosystem, has led to the development of resistance among the insect pest. Evidence of the development of resistance could be obtained by observation on the level of activity of detoxifying enzymes, notably esterases and transferases.
The level of activity of GST in the gut of the two species of lygaeids, *S. hospes* and *S. pandurus*, fed on *Calotropis* was analysed. The activity of GST in *S. hospes* was higher than in *S. pandurus*, indicating that the species was developing more resistance. Resistant populations of insects have an advantage in the process of natural selection and can thus lead to the evolution of new races and even species.

### 5.6 Molecular characters

Molecular genetic methods provide taxonomy with a powerful tool to re-identify or describe species irrespective of developmental stage, sex, or body part and offer universal, quantifiable characters. Molecular taxonomy is particularly effective in combination with other methods, usually with morphology, for correct species identification. With DNA barcoding, the foundation is laid for automated and accelerated taxon identification. Here, the amount of DNA that is similar in different species is examined. Then specific genes or proteins can be used as molecular clocks. These clocks help to determine the divergence from a common ancestor, or more simply, the relatedness of the species in question. The advent of molecular taxonomic techniques offered a solution for many problems, which were not available for classical taxonomic approaches. Currently, the methods of construction of phylogenetic tree based on molecular data are widely used not only in systematic and comparative biology, but also in ecology, sociobiology and epidemiology (Hampl *et al.*, 2001).

The rapid evolution of mitochondrial DNA sequences has often been used to investigate the relationships of populations within species (Avise, 2000) and the relationships of closely related species (Sperling, 2003). Recently, the use of
mtDNA sequence to delimit species has been advocated (Templeton, 2001; Wiens and Penkrot 2002, Wahlberg et. al., 2003)

The 28S, 18S, and 5.8S molecules are formed by the processing of a single primary transcript from a cluster of identical copies of single gene. It is suitable to select these sequences for evolutionary relationships analysis. The rRNA is the most conserved (least variable) gene in all cells. For this reason, genes that encode the rRNA (rDNA) have been sequenced to identify an organism’s group, calculate related groups, and estimate rates of species divergence (Wuyts et. al., 2002).

With a view to characterize and separate the two species of *Spilostethus*, three genes, namely 16S rRNA, 28S rRNA and COI, were used in the present study. The characteristic difference between the two species of lygaeids with respect to these three genes is summarized in table 60. The two species had differences in the number of nucleotide base pairs for all the three genes. For both *S. hospes* and *S. pandurus*, the percentage composition of A+T was higher than G+C for the 16S rRNA and COI genes. However, the 28S rRNA gene had the G+C percentage was higher in both the species. The sequence similarity analysis conducted using Blastn programme indicated that for the 16S rRNA sequence there was 97.86% identity between *S. hospes* and *S. pandurus* and 98% identity for the 28S rRNA data and 86% identity for the COI sequence. These values indicate that the cytochrome oxidase subunit I gene fragment is a potentially superior tool for molecular identification and separation of the two species of *Spilostethus*. Several previous studies advocated the use of mtDNA sequence as aids in identifying closely related species (Sperling and Hickey, 1994; Sperling et. al., 1995; Cognato et. al., 1999; Kruse and Sperling, 2000), and many studies explicitly stated the appropriateness of
mtDNA in resolving the relationships of closely related species (Brunto and Hurst 1998; Caterina and Sperling 1999; Rand et al., 2000; Monteiro and Pierce 2001). The usefulness of a gene for delimiting species is also influenced by the length of the sequence analysed.

Many studies looking at closely related species sequence only 400-600 bp per specimen and find few informative characters between species (Pedersen 1996; Walton et al., 1997, Brunton and Hurst 1998Brunton 1998, Brown et al., 1999l Nice and Shapiro 1999,2001). Wahlberg et al., (2003) used a relatively long sequence (1450 bp) and showed that there can be much information in mtDNA despite low pair wise divergence values. This added information may be able to resolve polytomies evident in studies using shorter sequences. The 28S rRNA sequences in the present study for S. hospes had 1070 bp and for S. pandurus there were 1081 bp.

5.6.1. Relative synonymous codon usage (RSCU)

Due to the degeneracy of genetic code, most amino acids are coded by more than one codon (synonymous codon). RSCU is a simple measure of non-uniform usage of synonymous codons in a coding sequence (Sharp et al., 1986, Nei and Kumar, 2000). RSCU values are the number of times a particular codon is observed, relative to the number of times that the codon would be observed for a uniform synonymous codon usage (i.e. all the codons for a given amino-acid have the same probability). In the absence of any codon usage bias, the RSCU values would be 1.00. A codon that is used less frequently than expected will have an RSCU value of less than 1.00 and vice versa for a codon that is used more frequently than expected. Studies of the synonymous codon usage can reveal information about the molecular
evolution of individual genes. The RSCU in the 16S rRNA showed differences between *S. hospes* and *S. pandurus*. The highest characteristic codons used in *S. hospes* 16S rRNA gene sequence were UUA, AUU, ACU UAU UAA, CAU, AAU, AGC, AGA, AGG, GGA, and GGG while in *S. pandurus* 16S rRNA gene sequence the codons were UUU, UUA, AUU, GUU, UCA, CCC, ACU, AGA, AGG, and GGU. However, for the 28S rRNA gene, there were several codons commonly used in high numbers for both *S. hospes* and *S. pandurus*. The highest characteristic codons used in *S. hospes* COI gene sequence was AUU, while for *S. pandurus* it was UUU and AUU.

**5.6.2. Nucleotide substitution among the sequences**

Population genetics has always played a central role in evolutionary biology as it deals with the mechanisms by which evolution occurs within populations and species, the ultimate basis of all evolutionary change. Population genetics is concerned with the origin, amount, and distribution of genetic variation present in populations of organisms and the fate of this variation through space and time. The analysis is concerned with describing the motive forces for evolution; mutation and the genetic drift. MEGA 5 software was used for evaluating mutational rate through analysis of the sequences together to determine substitution matrix. In the study of molecular evolution, it is important to know the number of nucleotide substitutions per site (d) between DNA sequences. Two important factors that are considered in the estimation of d are the inequality of the rates of transitional and transversional nucleotide substitution (transition-transversion bias) and the deviation of the G+C content from 0.5 (G+C-content bias). Transitions refer to the substitution of a purine (A or G) by another purine or the substitution of a pyrimidine (T or C) by another
pyrimidine; transversions are the substitutions of a purine by a pyrimidine or a pyrimidine by a purine. Zhang and Gerstein (2003) observed that the nucleotide transitions are more common than transversions, by roughly a factor of two. Moreover, the substitution rates amongst possible nucleotide pairs are not homogeneous as they are affected by the type of immediately neighboring nucleotides and the overall local G+C content. They reported that deletions are about three times more common than insertions. The ratio of Transitional Pairs versus Transversional pairs for the 16S rRNA data was 1.50 while for the 28S rRNA it was 1.25 and for COI it was 1.55. Habeeb et.al., (2011) observed an average ratio to be 0.467 for the 16S rRNA of insects they had studied. When two DNA sequences are derived from a common ancestral sequence, the descendant sequences gradually diverge by nucleotide substitution. A simple measure of the extent of sequence divergence is the proportion of nucleotide sites at which the two sequences are different. This is estimated as the p-distance for nucleotide sequences. It is useful to know the frequencies of different nucleotide pairs between the two sequences. Since there are four nucleotides, there are 16 different types of nucleotide pairs. There are four pairs of identical nucleotides (AA,TT, CC,GG represented as O), four transition-type pairs (AG,GA,TC,CT represented as P) and remaining 8 transversion-type pairs (represented as Q). The p distance for nucleotide sequence, given by the relationship p=P + Q was calculated to be 5 (ie. 3+2). If nucleotide substitution occurs at random, Q is expected to be about two times higher than P when p is small which was not the case in the present investigation. In general, transition usually occur more frequently than transversions. Therefore P may be greater than Q. When the extent of divergence is low, the ratio (R) of transitions to
transversions can be estimated from the observed values of P and Q. R is usually 0.2-2 in many nuclear genes, but in mitochondrial DNA it can be as high at 15 (Vigilant et.al., 1991). In the present study the value of R was 1.5. The analysis of the p value indicate that no synonymous substitution occurs in the first three codons ( p for 1st codon could not be calculated, 2nd codon is 0.02 and 3rd codon is 0.03).

5.6.3. Best fit DNA model

All methods of phylogenetic inference depend on their underlying models. To have confidence in inferences it is necessary to have confidence in the models (Goldman 1993). Because of this all methods based on explicit models of evolution should explore which is the model that fits the data best. Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best. The GTR+G had the least BIC scores and therefore considered the best model for the 16S rRNA data. The T92+G model was the best for the 28S rRNA while the T92 model was the best for the COI data in the present study. Another way of selecting the most appropriate model for a data set is to use the Akaike information criterion (AIC) (Akaike 1974), which can be thought of as the amount of information lost when a particular model is used to approximate reality. The AIC implements best-fit model selection by calculating the likelihood of proposed models, and imposing a penalty based on the number of model parameters. Parameter-rich models incur a larger penalty than more simple models so that fitting an excessively complex model is not likely. The best fitting model is the one with the smallest AIC value. The GTR+G had the smallest AIC value for the 16S rRNA data while the T92+G had the smallest value for the 28S rRNA and T92 had the smallest value for the COI data.
5.6.4. Distance matrix

The concept of descent tells us that organisms sharing a recent common ancestor should, on average, be more similar to each other than organisms whose last common ancestor was more ancient. Therefore, it should be possible to infer evolutionary relationships from the patterns of similarity among organisms. This is the principle that underlies the various distance methods of phylogenetic reconstruction. First, a distance matrix, which is a table of “evolutionary distances” between each pair of taxa, is generated. The resultant matrix is then used to generate a phylogenetic tree.

The number of base substitutions per site from between sequences were conducted using the Maximum Composite Likelihood model. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The analysis involved 12 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA5. The evolutionary divergence between S. hospes and S. pandurus was at a distance of only 0.11 for the 16S rRNA data and 0.005 and 0.175 for the 28S rRNA and COI data respectively.

One limitation of both the distance and parsimony methods is that although they may select one tree over another on the basis of some criterion, it is not possible to say how much more probable one tree is than another. Likelihood and Bayesian methods have been designed to provide such a statistical framework for phylogenetic reconstruction.
5.6.5 Molecular clock analysis and Evolutionary tree

‘Molecular clock’ refers to the approximately steady rate of accumulation of changes in DNA (or protein amino acid) sequences over evolutionary time. A consequence of this rate constancy is a nearly linear relationship between evolutionary distance and time of species divergence. As more molecular sequence data became available, the initial optimism about the universality of the molecular clock was questioned by reports of significant difference in evolutionary rates among species in some genes and in some lineages. Fitch (1976) proposed a test for statistically examining whether the observed difference in evolutionary rates between two sequences is significantly greater than that expected by chance. This test used sequences of a gene from three species. For two of these species, the molecular clock was tested (A and B) and the third was used as an outgroup (C). As far as the data of the 16S rRNA, 28S rRNA and COI is concerned, the null hypothesis is accepted indicating that evolution of the two species of Spilostethus is at equal rates. The evolution of living organisms is the consequence of two processes. First, evolution depends on the genetic variability generated by mutations which continuously arise within populations. Second, it also relies on changes in the frequency of alleles within populations over time. That is why the Tajima's test was performed for inspection of whether occurrence of neutrality between such two processes performed the forcing mechanisms for evolutionary change or it would be the role of natural selection by mutation. A positive value of D computed here for all the three genes indicates an excess of intermediate frequency (polymorphic) alleles. There was equivalence (confirmed by nearly equivalence between $\Theta$ and $\pi$) between genetic drift and mutation. The equivalence between such two values and the positive value of D considered the null hypothesis to be held.
The Tajima’s D test is a widely used test of neutrality in population genetics. This statistic illustrates the allele frequency distribution of nucleotide sequence and is based on the difference between two estimators of (the population mutation rate): (1) Tajima’s estimator, which is based on the average number of pairwise difference between sequences, and (2) Watterson’s estimator, which is based on the number of segregating sites in the sample. Tajima’s estimator takes into account allele frequency when comparing pairwise differences between sequences whereas Watterson's estimator does not. A segregating site counts as any point where there are differing nucleotides between sequences in the data set, independent of the actual number of differences and hence independent of allele frequency. The difference between these two estimators of is scaled by the standard deviation of their difference. A positive value of D as in this study indicates an excess of intermediate frequency (polymorphic) alleles, while a negative value indicates an excess of rare alleles. There was equivalence (confirmed by nearly equivalence between $\Theta$ and $\pi$) between genetic drift and mutation. The equivalence between such two values and the positive value of D considered the null hypothesis to be held.

The results of the phylogenetic relationships among the taxa using 16 S rRNA sequence indicate a topology consistent with the generally accepted relationship. The two species of *Spilostethus* share a most recent common ancestor and are clustered at 70% (ML method) and 87% (Neighbor-Joining method). This clade is a sister group to *Lygaeus equestris* and *Oncopeltus fasciatus*, both being members of Lygaeinae. Species of the subfamily Rhyparochrominae, *Dieuches* sp and *Laryngodus* sp have been positioned distantly. The 28S rRNA data showed the two species of *Spilostethus* were clustered together at 99% and this clade was
clustered with *Dysdercus koenigii* at 80%. A 20% separation between *S. hospes* and *S. pandurus* collected from Chennai was observed with regard to the COI data.

One of the fundamental changes in systematic practice during the past 25 years is the broader application of molecular approaches. The first molecular systematic study with focus on Heteroptera was by Wheeler *et al.*, (1993), who combined 31 morphological characters with partial 18S rDNA sequence data (670 bp) and proposed the phylogeny. After a gap of several years, Gerromorpha became the focus of several molecular systematic studies. Muraji and Tachikawa (2000) analyzed 16S and 28S rDNA for 30 terminals of Gerroidea. Hua *et al.*, (2008) was the first study in Heteroptera to use mitochondrial genomes for cladistic analysis. They analyzed complete or nearly complete mitochondrial genomes of 16 taxa of Geocorisae with emphasis on Pentatomomorpha and found the infraorder to be monophyletic, with Aradoidea as the sister taxon to the Trichophora. This analysis recovered monophyletic Pentatomoidea (three taxa included), Pyrrhocoroidea (two taxa), Lygaeoidea (four taxa), and Coreoidea (three taxa). It has been proposed that in understanding the phylogenetic relationship between taxa, it would be worthwhile to use multiple genes in the analysis.

5.7 Concluding remark

Currently, COI has been selected as a standard barcode gene for animal groups. However, the rationale of selection of COI as standard barcode is subject to debate, and with the increase in barcoded taxa, from algae, fungi, bacteria and plants to invertebrates and vertebrates, scientists have found its less effective in some taxon groups (Hebert *et al.*, 2003; Meier *et al.*, 2006; Waugh *et al.*, 2007). The search for the most suitable gene for species identification is not over, with several recent
studies testing the efficiencies of different genes, using part of, or the whole of mtDNA genome to look for the optimal DNA barcode gene (Luo et al., 2011). On the other hand, empiricists have also proposed other gene segments as candidate DNA barcode loci, such as the nuclear ITS regions (ITS1, ITS2) (Chen et al., 2010). ITS - Internal Transcribed Spacer (ribosomal DNA repeating unit), which is a commonly used DNA biomarker, was suggested and examined in several plant groups and fungi (http://www.boldsystems.org/views/projectmenu.php). This widely used genetic marker might be suitable as a DNA barcode due to its highly variability. This is especially the case for groups composed of closely related species, where the rate of successful species identification with COI is relatively low (less than 70%) (e.g., fly, (Meier et al., 2006). Unlike groups of distantly related species, where the existence of large genetic divergence between species makes discrimination easy, groups of closely related species offer greater challenges for phylogenetic reconstruction and clear species identification. The utility of mtDNA on its own in assessing the boundaries of traditionally recognized species (Wiens and Penkrot, 2002) is suspect. In the present study, although S. hospes and S. pandurus could be delimited by studying the COI profile alone, in the larger perspective of species identification across families, multigene barcoding is suggested. One must combine all possible knowledge, including morphological, ecological and molecular, to understand the species boundaries of groups of very closely related species. Perhaps the most significant limitation of rRNA for phylogenetic analysis is that even the most rapidly evolving regions generally do not evolve fast enough for this molecule to be used to study relatively recent evolution (e.g., relationships among species within genera or within species). When such
recent evolutionary events are being studied, regions or genes that evolve much more rapidly are needed. One possibility is to use protein-coding genes, because the degeneracy of the genetic code means that even when a protein is completely conserved, the DNA sequence encoding that protein can vary. Or, if protein-coding regions do not vary enough, introns, or pseudogenes, or intergenic spacer regions can be used. However, protein-coding genes have a major advantage—even when the DNA sequence is not highly conserved, the amino acid sequence sometimes is, making alignments easier to create.