D. R. MILLER, Y. BEN-DOV, M. E. GIMPEL, V. GERMAN, G. A. P. GIBSON

Recent enhancements in ScaleNet: a queriable systematic database of the scale insects of the world (Hemiptera: Coccoidea)

Abstract - A number of changes have been implemented in ScaleNet since it was last discussed in 1998 at the International Symposium on Scale Insect Studies. These changes include addition of databases on the following 18 families: Aclerdidae, Beesoniidae, Carayonemidae, Cerococcidae, Dactylopiidae, Grimaldiellidae, Halimococcidae, Inkaidae, Jersicoccidae, Kermesidae, Kerriidae, Kukaspididae, Labiococcidae, Lecanodiaspididae, Micrococcidae, Phenacoleachiidae, Phoenicococcidae, and Stictococcidae making a total of 23 families on line. The bibliographic file has grown from 11,000 to over 14,700 references and a system has been developed to look for variant spellings of author names. A new query has been added that allows a search for a string of letters in a species epithet and very soon a section on on invasive species will be added. The "Scales in a Region /Country/Subunit Query" now allows a search for all of the species in a zoogeographic region, country, or country subunit in a particular scale family or genus. The "Hosts, Natural Enemies and Associates of a Scale Query" allows a search for all of the species that are found on a specified host, have specified natural enemies, or have specified associates in a particular scale family or genus. Significant progress has been achieved on the armored scale, margarodid, and asterolecaniid databases.

Key words: Coccoidea, ScaleNet, information, internet, web.

INTRODUCTION

A detailed explanation of the history and rationale for the development of ScaleNet was presented at the VIII International Symposium on Scale Insect Studies in England in 1998 (Miller et al., in press) and will not be repeated here. The purpose of this paper is to: 1) provide information on progress that has been achieved since the last symposium; and 2) to update the abbreviated user manual that was published in the proceedings of ISSIS VIII.

We are including and/or updating all sections of the ScaleNet manual here so that the most current information will be available in a single document. Please note that detailed query instructions also are provided directly in the "Background" and "Query Descriptions" sections in ScaleNet.

RESULTS

Additions since 1998: The following databases have been added to the ScaleNet system since ISSIS-VIII: Aclerdidae, Beesoniidae, Carayonemidae, Cerococcidae, Dactylopiidae, Grimaldiellidae, Halimococcidae, Inkaidae, Jersicoccidae, Kermesidae, Kerriidae, Kukaspididae, Labiococcidae, Lecanodiaspididae, Micrococcidae, Phenacoleachiidae, Phoenicococcidae, and Stictococcidae making a total of 23 families on line. This includes about 4,700 species and 580 genera. Significant progress has been made on the remaining families (Asterolecaniidae, Diaspididae, and Margarodidae) with more than half of each of these families entered. Although we had hoped to complete the armored scales by the end of the year 2000, the process is tedious and time consuming and has taken considerably more effort than anticipated.

Several important changes have been made to the query system. Those queries that require an author name, have a table of alternative spellings attached so that clients can enter a variant spelling of an author's name and still find all of the papers by that author regardless of the spelling used. This area is a nightmare that just will not quit!!! Names such as: Borchsenius and Borksenius, Jashenko and Jashchenko and Yashenko and Yashchenko; Gómez-Menor and Gómez-Menor Ortega and Ortega and Menor; Kozár and Kozar; Boyer de Fonscolombe and Fonscolombe are used inconsistently in the literature. In many cases it is difficult to make informed decisions concerning the correct spelling. In fact, we now realize that there are multiple spellings of certain names in ScaleNet. Since these various spellings are used throughout the database we have decided to develop a table of alternate spellings to rectify the problem rather than arbitrarily deciding which spelling is correct. This table is accessible from the "References for an Author" and "Scales described by an Author" queries. Also on these same query pages you will note a new button which allows a client to request alternative spellings of the author name. Results will provide the spelling that has been used in ScaleNet if it is different from the one input by the client. Another method of circumventing variable spellings is to enter part of the author's name as a string of letters. This opportunity allows clients to solve some of the problems mentioned above without having to worry about alternate spellings. For example, entry of "senius" (rather than having to decide whether it should be Borchsenius or Borksenius) in the author last name field in any of the queries "Scales Described by an Author", "References by an Author" will give all of the requested information in the database by Borchsenius.

The second enhancement involves the "Scales on a Specified Host, Scale Natural Enemy or Scale Associate" and "Scale in a Region/Country/ or Subunit." Since these queries have been available, they always have been frustrating to use because to find all of the species in a scale family or scale genus on a particular host or in a particular geographic areas, it was necessary to look through a results table that included all of the scales that occurred on the host or in the area. We now have added an enhancement that allows a client to specify a family of scale or genus of scale that occurs on a particular host plant (or is fed upon by a particular natural enemy or is affiliated with

a particular associate) or occurs in a particular area. The results tables include only species in the specified scale family or scale genus that occur on the host or in the area of interest. For example, previously if a client wanted to know all of the mealybugs in the genus *Pseudococcus* that occur on *Ficus*, it was necessary to wade through a results table that included 380 records in 7 different scale families. Now it is possible to specify the target host and scale genus and to receive a results table that includes only the 18 records of *Pseudococcus* species that occur on *Ficus*. The same is true for the "Scales in a Region/Country/Subunit" query. Previously if a client wanted to know all of the mealybugs in the genus *Pseudococcus* that occurred in Italy, the results table included 198 records of all scales in 9 families from Italy. With the improvement of being able to request just those species of *Pseudococcus* that occur in Italy, the results table includes only the 6 species known from this country with no extraneous information.

Improvement number 3 involves distribution data within ScaleNet and impinges on two queries, i.e., "Distribution of a Scale" and "Scales in a Region." It has become increasingly obvious that there is a serious need for information on the native range and invasiveness of scale species. This issue was not considered when developing ScaleNet, but we now realize that information on the subject could be very valuable to a wide range of clients. Although we are still in the process of putting the finishing touches to the programming of this area, it is apparent that a species' invasive status will be part of the distribution table. When a results table is produced it will include three kinds of information about a species relative to its invasive status in a country or country subunit, i.e, if the species is invasive, native, or if there is no indication of its invasive status. These designations will be based entirely on the literature. Thus, if Evelyna Danzig wrote a paper on the soft scales of the Urals and indicated that *Eulecanium douglasi* (_ulc) is introduced in this area and that *Rhodococcus spiraeae* is native, this information would be indicated in the results table. If Evelyna did not make a statement about the invasive status of the soft scales of the Urals, the invasive field would be left blank or would be indicated as unknown. The issue for those of us entering data in this field is that it is difficult to go back to the nearly 15,000 references that currently are in ScaleNet and enter this information so that it is current.

Enhancement number 4 involves a new and powerful query called "Search for a String of Letters in the Species/Subspecies Epithet." It allows entry of any consecutive letters in a name (i.e., a string of letters) and gives a resulting table with all of the names containing the string. For more information see the query description given in the user manual that follows.

ScaleNet User's Manual: In order to access a significant portion of the information available in ScaleNet, it is essential to know the valid or correct name of the taxon in question. This information can be obtained using the following query. Clients do not need to know the valid name, but only need to know a combination or synonym that has been used in the literature. To use this query click on **"Find a Valid Name and Catalogue"** on the query page. Clients will be asked to provide a family, genus, and

species name (if the family is unknown use "ALL"). For example, if they wish to know the correct name of *Pseudococcus citri,* they type *"Pseudococcus"* in the genus field and *"citri"* in the species field, and choose "all" in the family field. The response will be *"Planococcus citri* (Risso) (Family Pseudococcidae) is the valid name." This query is useful for clients who need to know if a name used in the literature is still valid or if they need to know the author or family name of the taxon in question. Below the valid name statement are three options to retrieve a taxonomic catalogue of: 1) the species, 2) the genus, 3) all of the species in the genus. If the first option is selected, the output will be a catalog of *Planococcus citri* that includes the following sections: Nomenclature (Synonymy), Common Names (if any), Foes (if any), Associates (if any), Hosts, Distribution, Biology, General Remarks, Structure, Systematics, Economic Importance and Control, Keys (references to any time that the species was used in a key), and Citations. Selection of the second option gives a catalogue of just the genus by itself. The third option gives a catalogue of the genus and all of the included species. The "Valid Name and Catalogue" query also provides information on homonyms or misidentifications so that the client has the option of selecting the valid name that is pertinent to their query. For example, if clients wanted information on the name *Opisthoscelis globosa*, based on a literature record or identification, they would type *"Opisthoscelis"* in the genus field, *"globosa"* in the species field, and Eriococcidae (or "ALL") in the family field. They would receive the following reply: *"Opisthoscelis globosa* Rübsaamen (Family Eriococcidae) is the valid name (originally described from AUSTRALIA: New South Wales, on *Eucalyptus*) BUT there is another species that once had the same name *Opisthoscelis globosa* Froggatt (Family Eriococcidae) (originally described from AUSTRALIA: Victoria, on *Eucalyptus* sp., by C. French and New South Wales, by Honsby) and its valid name is *Opisthoscelis rubsaameni."* If clients were looking for information about *Opisthoscelis globosa* Froggatt, they would need to locate a catalogue of *Opisthoscelis rubsaameni* not *O. globosa* Rübsaamen. Catalogue options for the valid name of the homonym or misidentification can be accessed directly from the valid name statement without the need to run another query by simply clicking on the hot linked junior homonym name. If a request is made for a valid subspecies, and the subspecies is valid, clients will be given four catalogue options including a catalogue of the subspecies.

After the valid name of the taxon in question is known, several specific queries can be made such as **"Hosts of a Scale"**, **"Distribution of a Scale"**, **"References for a Scale"**, **"Synonymy for a Scale"**, **"Remarks for a Scale"** (remarks on systematics, structure, biology, economic importance, or general remarks). These queries are straight forward, requiring clients to fill in the genus field and species field, and to choose a name from the choices in the family pick list. Output is all of the known hosts, distribution, references, synonymy, or remarks for a scale. The known hosts and distribution queries include links to one or more references that validate each record.

The **"Scales in a Family/Genus"** query requires input of a valid family or valid family and genus name and the output is a table that is a checklist of the valid species

in the taxon requested. This query can be a very valuable tool since it often is difficult to obtain a comprehensive and up-to-date checklist of the component species of higher taxa. Because the species names in the results table are hot linked, it is easy to obtain a catalogue for any species of interest by simply clicking on the species name in the table.

There also is a series of queries that does not require a valid name. For the **"Scales on a Specified Host, Scale Natural Enemy or Scale Associate"** query we will first discuss host information and then will mention natural enemy and associates data later in the paragraph. The client has the option of entering one or more of host family, host genus, and host species. Clients can enter only a Host Family (and be given a table including all of the scales on the selected family), or a Host Genus alone (and be given a table of all of the scales on the selected genus), or a Host Genus and Host Species (and be given a table with all of the scales that occur on the selected host species). The resulting table has column headings for: Scale Family, Genus, Species, Subspecies, and Validation Source (reference where the host record was published). The validation source is hot linked so that the complete reference and associated annotations can be retrieved. Each species name is hot linked so that the client can immediately obtain a catalogue for any of the scale species in the table. Queries can be restricted by the client specifying a family of scale or genus of scale that occurs on a particular host plant. The results table includes only species in the specified scale family or scale genus that occur on the host of interest. Similar information can be obtained for natural enemies or for associates (associates are organisms such as ants that are associated with the scale but are not natural enemies). For example, if clients enter *"Crematogaster"* in the genus field, they will be given a table that includes all of the scale species recorded as associated with *Crematogaster* ants in ScaleNet. A similar search can be made using a natural enemy name. For example, entry of *"Anagyrus"* in the genus field gives a table that includes 37 records and 30 different scale species attacked by *Anagyrus* parasites. Entry of "Encyrtidae" gives 219 records and 134 species. We should point out that the natural enemy and associates areas of the ScaleNet system are far from complete. We have not actively pursued natural enemy and associates information but have included data when they were found.

A query for **"References Between Two Dates"** was designed for clients to find all of the scale-insect literature published between two dates. Click on this choice and clients are given a screen that allows entry of a "Begin Date" and "End Date." If the latter is not filled in, the default is the current year. There is a restriction of no more than a 10 year span. The results of this query are a list of all of the references in the database between the years selected, including the associated annotations for each reference.

The **"Reference for a Citation Code"** query was programmed to allow clients to find a complete reference for any citation code that is encountered in ScaleNet. Each reference in the BASIS database is designated by a unique code that is formed from the names of the authors and the date. For example, the citation code for Cockerell and Bueker 1930a is CockerBu1930a. In most query outputs this code is hot-linked

to the reference so that a client can simply click on the code to receive a complete citation and associated annotations. In a few instances, however, direct hot-linking was not possible, so a query was developed to allow clients to determine the complete reference citation for any unknown citation code. If a citation code is found that is not already hot-linked to a reference, clients can either use the Internet browser's "copy" option to copy it, or "<Ctrl> C." Then go to the "Reference for a Citation Code" selection. When this option is selected, the next screen is a blank field to enter a citation code, in which the clients paste the copied citation code with the browser "paste" option or "<Ctrl> V". Thus, if "CockerBu1930a" is entered in the "Reference for a Citation Code" query, ScaleNet will return the complete reference and any associated annotations. Please be aware that when entering citation codes, they are case sensitive. That is, you must have capital letters in the correct positions. This is another reason for "copying" and "pasting" the codes. No diacritical marks are used in citation codes.

When clients click on the **"References for an Author"** option the resulting screen provides blank fields for the name of an author. When entering this information be certain that the capitals, hyphens, and diacritical marks are in the correct place. Directions for obtaining diacritical marks are given in the instruction section for the query. The output for this query is a list of references by the author selected, including junior author papers. The client is given the option of including begin and end dates. The list includes the complete reference and associated annotations and gives papers of all authors with the selected last name. Clients should be aware that multiple authors with the same last name will be obtained in this query unless initials are used. Therefore, in the blank author field enter "the last name" then "comma" then "a space" then "the first initial" then "a period" then "a second initial" then "a period", etc. There is no space between the initials but there is a space between the comma after the last name of the author and the first initial. For example, if a request was submitted for papers by Williams (without any initials), the resulting output would include papers by D.F. Williams, D.J. Williams, J.R. Williams, K.S. Williams, L. Williams, M.L. Williams, P. Williams, and R.N. Williams. To obtain all of the papers by Douglas J. Williams enter "Williams, D.J." A table for alternate spellings of authors names has been developed to assist clients in finding references for authors regardless of how they are spelled in ScaleNet. To invoke the alternate author spelling table, clients must click on the "yes" button for alternate author spelling. If the spelling used by the client is not one used in ScaleNet, but is one in the alternate author spelling table, all of the references for the author will be provided and a note at the top will indicate the spelling used in ScaleNet and the possible alternatives. For example, if the client enters "Yashchenko," which isn't the spelling used in ScaleNet, and clicks on the "yes" button for alternate author spelling, a list of all of the papers by Jashenko is provided. In addition a note at the top of the page is provided indicating: "The results of your query for **Yashchenko** are as follows: Author: **Jashenko** (alternate spellings Jashchenko, Yashenko, Yashchenko)." It also is possible to examine the alternate author spelling table directly by clicking on "Alternate author spellings and the preferred spelling used

in ScaleNet" at the bottom of the query page. This query will find all authors that have a specified string of letters in their name. This tool can be used to avoid diacritical marks and can assist in finding alternate spellings of a name. Using the Jashenko example, if the client enters "henko" in the author name field the results of the query include all Jashenko references in addition to papers by Kititchenko, Kharchenko, Ishchenko, etc.

"References with Words in Article Title/Journal Title/Notes" is a powerful query, but is a bit complicated. Click on this selection and clients will be given a screen with the option of entering from one to five words to search for in the article title, journal title, or associated annotations of the references. Clients also have the opportunity to limit the search with the last name of an author and the begin and end year of publication. Thus, if the first word field is filled in with "apple" the output will be a list of approximately 270 references and the associated annotations. Several of these will have hits on the word "pineapple." This problem can be alleviated by putting a space before the word apple, i.e., if " apple" is submitted, the output will include about 180 hits. Adding new words in the word fields will restrict the number of hits. If the client adds the word "perniciosus" to the second line of the query, the resulting list of publications will include about 71 hits. Please be aware that this query is case sensitive, i.e., if all papers with apple in them are required, separate searches with "Apple" and " apple" must be made. For the former there are 5 hits, for the latter there are 180. The alternative is to search for "pple" but this will add other inappropriate references with words like "supplement." If clients wanted to find literature on the pink hibiscus mealybug (*Maconellicoccus hirsutus*) a search using the words "Maconellicoccus" and "hirsutus" would give a list of 129 references. A query using "ealybug" for clients interested in the literature on mealybugs gives more than 1,400 references and a search using "ealybug" with a year restriction of 1980 to 1990 gives 398 references. Use of "itrus" and "ealybug" for someone interested in the citrus mealybug gives 186 references. A colleague recently asked about a publication on the male of a kermesid. I didn't remember details of the paper so went to this query and put " male" (note the space before male) in the first field and "Kermes" in the second field. The results of the query gave 10 references including the one that was needed.

"Scales Described by an Author" is a query designed to list all of the taxa described by a particular author. If clients choose this selection, they will be given a screen that requires the last name of an author or a string of letters that is part of the name. When entering this information it is important to put capital letters, hyphens, and diacritical marks in the correct place. The query can be restricted by including a begin and end date. If this isn't included, the default is to provide all of the scales described by the author. By clicking on the appropriate circle, clients are given the option of receiving the generic names described by the author, the species names, or the subspecies names. Clients also have the option of invoking use of the alternate author spelling program as described in instructions for the "References for an Author" query. Simply click on the "yes" botton to activate the program. The output for this query is a list of scales described by the selected author. The list includes the cited

name, year, current status of the name, and the valid name and is hot linked to the catalogue of the taxon selected. This query does not distinguish between authors with the same last name so a list might include species described by more than one person. If a client put "Dantsig" in the "Author (last name)" field, clicked on the genus button and clicked the "yes' button for "check alternate author spellings?" the response would be a list of 10 genera described by Evelyna Danzig and there is a note at the top of the screen indicating that the spelling used in ScaleNet is "Danzig" but alternate spellings are "Dansig and Dantsig." If a client put just "Dan" in the author field, the output would be the same 10 genera described by Danzig.

   "Scales in a Region/Country/or Country Subunit" provides a checklist of all of the species of scale insects that occur in a specified geographical area. Click on this selection and clients are given the option of choosing any of the geographic units listed in a picklist. The choices and results will include any of the zoogeographic regions, countries within the selected zoogeographic region, and subunits within the selected country. Clients also have the option of choosing a scale family or scale family and scale genus to restrict the query further. The results give all of the scale species in one or "ALL" scale families that occur in a particular zoogeographic region, country, or country subunit. Thus, if clients submit a request for all of the scales in the "Pseudococcidae," in the genus "Balanococcus," from the "Palearctic Region," in "Italy," they will be given a table listing 3 species of *Balanococcus*.

   "Find a Scientific Name of a Scale Known only by a Common Name" is a query designed to assist clients who want more information about a scale insect but know only a common name. Selection of this query gives a screen with five common name fields that are filled in by the client. Capital letters should only be used when a proper name is part of the common name, e.g., Putnam scale, Florida wax scale; small letters should be used when there are no proper names, e.g., white peach scale, brown soft scale. Clients may wish to know all of the scales that have "scale" in their name. They simply type "scale" in the first common name field and are given a table with a list of 251 common names. If a client wanted to know all of the common names that have "green" in the common name they would enter "reen" (to avoid the problem with any species described for E.E. Green) and they would be given a table with a list of 8 common names with the word "green" or "Green" in them. Submittal of "Green" gives only 2 common names. The client is allowed to enter up to five words in separate fields. Thus, if a client wanted to know the scientific name of the "New Zealand flax mealybug" it would be possible to enter all or any of the four parts of the common name and be given a table containing a list of species including the scientific name of the New Zealand flax mealybug (it also would include other species if only the word "New" was used). Output from this query is one or more valid scientific names. These names are hot-linked and a catalogue for the species is displayed when the scientific name is clicked. Currently, most common names in ScaleNet are in English, but names in other languages are being added. A scale-insect species may have several different common names. For example, *Saissetia oleae* (Olivier) is called "the black scale" in North America and is "the olive scale" in Mediterranean areas.

If a client fails to retrieve information under one common name, it may be necessary to try another one, if known.

The query **"Search for a String of Letters in the Species/Subspecies Epithet"** allows entry of any consecutive letters (i.e., a string of letters) in a species or subspecies name and gives a resulting table with all of the names containing the string. For example, entry of the string "dan" gives a table containing 22 species names including "DANzigae" three times, "panDANi" four times and "DANielaferreroae" once. When a client knows the approximate spelling of a species name but is not certain how it is spelled in ScaleNet, this query can be very useful. For example, it often is a problem to know if an adjectival species name ends in an "a", "us", "um" etc. Use of this query through entry of only a portion of the name will help the client find the name and the needed information. If a person was looking for a species and recalled that it was named after Penny Gullan but couldn't remember the genus or the family designation, it would be simple to enter the string of letters "gullan" and retrieve all of the species with that set of consecutive letters. In many cases this may be a quicker way of retrieving a catalogue for a species. It often is faster and more efficient to enter a string of letters than to use the "Find a Valid Name and Catalogue" query where the client must pick the family from the family pick list, enter the genus name, and enter the species name.

DISCUSSION

ScaleNet is a valuable and easily accessible information resource on scale insects. Our current estimate of a date to add the last family database is in the fall of 2002. In the earlier paper on ScaleNet (Miller *et al.,* in press) we estimated that we would finish the first round of data entry by the end of 2000, but that was overly optimistic. The greatest challenge is the Diaspididae which we have been entering since 1998. When the first round of data entry is complete, we will do our best to update the system as frequently as possible. There are many possibilities for future enhancements of the system including addition of images that include an illustration or photograph of each species; identification information, descriptions, and interactive keys; detailed data on natural enemies; more precision in host names; and inclusion of unpublished specimen label data from major collections. The primary issue, of course, is funding. Since the initial grant from BARD (see below) our attempts at finding additional financial support have been almost completely fruitless.

ACKNOWLEDGMENTS

Blacksburg, Virginia, USA for her extensive efforts as the person responsible for all aspects of the reference information in ScaleNet. It is a tedious and difficult part of the project and is of major importance to its success. To Jennifer Read, Biological Resources Program, Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Ottawa, Canada we also extend our deepest gratitude for being responsible for making the BASIS database work and for making changes that allow enhancements to the overall system. Thanks to both of them! We also are grateful to the following individuals for reading and commenting on the manuscript: Avas Hamon, Florida Department of Agriculture and Consumer Services, Gainesville, FL, USA; Alexander Konstantinov and Ronald Ochoa, Systematic Entomology Laboratory, Agricultural Research Service, United States Department of Agriculture, Washington, D.C., USA.

REFERENCES

MILLER, D. R., Y. BEN-DOV, G. A. P. GIBSON., 1999 (2001) - ScaleNet: a Searchable Information System on Scale Insects. - Proc. of the ISSIS XIII, International Symposium on Scale insect studies. Entomologica, 33: 37-46.

DOUGLASS R. MILLER - Systematic Entomology Laboratory, Agricultural Research Service, U. S. Department of Agriculture, Beltsville, Maryland 20705 United States. E-mail: dmiller@sel.barc.usda.gov.
YAIR BEN-DOV - Department of Entomology, ARO, The Volcani Center, Bet Dagan 50250, Israel. E-mail: yairbd@netvision.net.il.
MAREN E. GIMPEL - Systematic Entomology Laboratory, Agricultural Research Service, U. S. Department of Agriculture, Beltsville, Maryland 20705 United States. E-mail: mgimpel@sel.barc.usda.gov.
VIKTORIA GERMAN - Department of Entomology, ARO, The Volcani Center, Bet Dagan 50250, Israel. E-mail: yairbd@netvision.net.il.
GARY A. P. GIBSON - Biological Resources Program, Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada. E-mail: gibsong@em.agr.ca