←Please use the PDF bookmarks for Navigation

Help File: Look At Kinship

A. Data on Kinship

1. Overview

Queries involving kinship are more complex than queries examining categories of association or modes of attaining eligibility for office. Since the information on kinship for an individual usually contains just a few records, CBDB begins with those records and then looks at the kinship information available for all the kin listed for the initial person. CBDB repeats this search for the kin, the kin of the kin, the kin of the kin of the kin, and so on, until specified criteria are met. First is simply a limit to the number of search iterations to allow. Usually 10 loops are adequate. Second are limits on the distance of the kinship relations being examined:

a. Kinship Distance

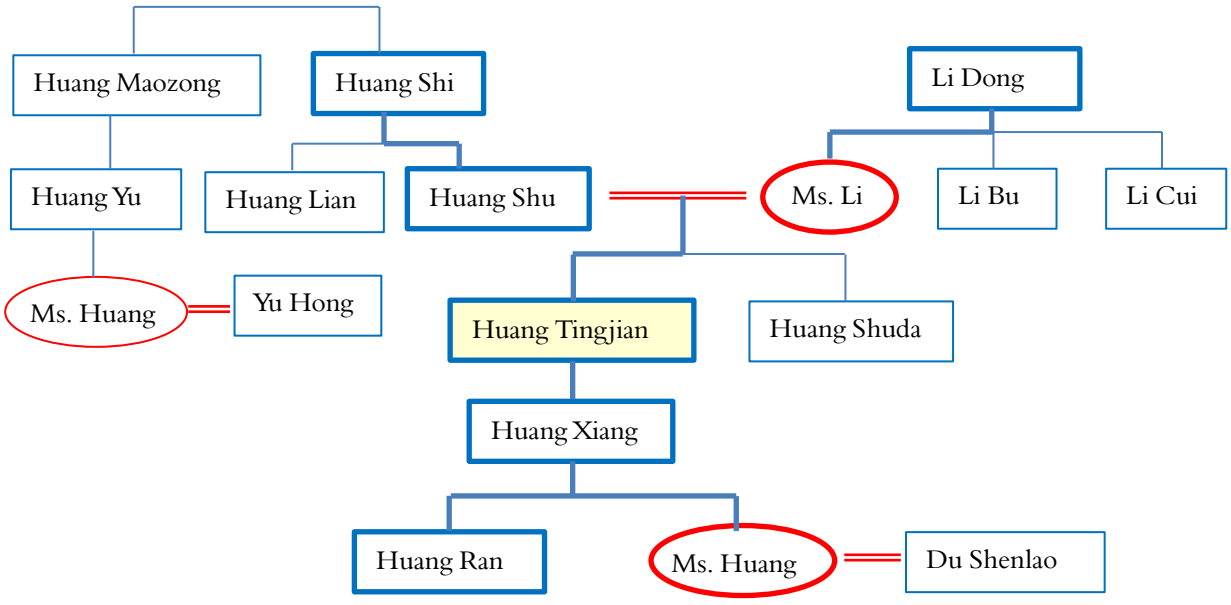
Max. Ancestor Gen. specifies how many generations of ancestors to include. One's father's generation is 1; the grandfather is 2, great-grandfather 3, and so on.

Max. Descend. Gen. specifies how many generations of descendants to include. One's children's generation is 1, grandchildren 2, great-grandchildren 3, and so on.

Max. Collateral Kin limits how many horizontal moves are allowed. For example, one's wife's sister has one unit of "marriage" distance and one unit of "collateral" distance. One's wife's sister's husband's brother has two units of "marriage" distance and two units of "collateral" distance.

Max. Marriage Dist. limits how many links defined by marriage are allowed in the search. One's wife's sister's husband has two units of "marriage" distance.

To visualize these distinctions, consider the partial kinship network:



For Huang Tingjian, the squares and ovals with thick lines show direct lineal descent (fathers and mothers, sons and daughters). The double lines are marriage links. All other single lines mark collateral relations. In the measurement system used in LookAtKinship:

Huang Yu 黃育	is FFBS	(Up = 2, Down = 1, Collateral = 1)
Yu Hong 余宏	is FFBSDH	(Up = 2, Down = 2, Collateral = 1, Marr. = 1)
Li Cui 李萃	is MB	(Up = 1, Collateral = 1, Marr. = 1)
Du Shenlao 杜莘老	is SDH	(Down = 2, Marr. = 1)

b. Kinship Terms

The building-block relations used to describe kinship are 9 basic categories:

e	Ego (the person whose kinship is being explored)
F	Father
M	Mother
B	Brother
Z	Sister
S	Son
D	Daughter
H	Husband
W	Wife
C	Concubine

There are also variations on the nature of the relationship, as well as additional types of notation to represent types of kinship relations beyond the nuclear family:

+	Older (e.g. older brother B+, 兄)
-	Younger (e.g. younger sister Z-， 妹)
⋆	Adopted heir (as in S⋆, adopted son)
°	Adopted
!	Bastard
^	Step- (as in S^ step-son)
½	half- (as in Z½ , half-sister)
~	Nominal (as in M~ , legitimate wife as nominal mother to children of concubine)
%	Promised husband or wife (marriage not completed at time of record)
y	Youngest (e.g., Sy is the youngest known son)
1, 2, 3…	Numbers distinguish sequence (e.g., S1, S2 for first and second sons; W1, W2 for the first and the successor wives)
n	precise generation unknown
G-#, G+#	lineal ancestor (–) or descendant (+) of # generation
G-n, G+n, Gn	lineal kin of an unknown earlier generation (G-n), or unknown later generation (G+n), or unknown generation (Gn)

G-#B, BG+#	a brother of a lineal ancestor of # generation; a brother's lineal descendant of # generation
K, K-#, K+#, Kn	Lineage kin, of the same, earlier (–), later (+) or unknown (n) generation. CBDB uses "lineage kin" for cases where kinship is attested but the exact relationship is not known. Lineage kin are presumably not *lineal* (direct descent) kin.
K–, K+	Lineage kin of the same generation, younger (-) or elder (+).
P, P-#, P+#, Pn	Kin related via father's sisters or mother's siblings, of the same, earlier (–), later (+) or unknown (n) generation. Signified by the term *biao* (表) in Chinese. (CBDB uses these codes only when the exact relationship is not known).
P–, P+	Kin related via father's sisters or mother's siblings, of the same generation, younger (-) or elder (+).
A	Affine/Affinal kin, kin by marriage

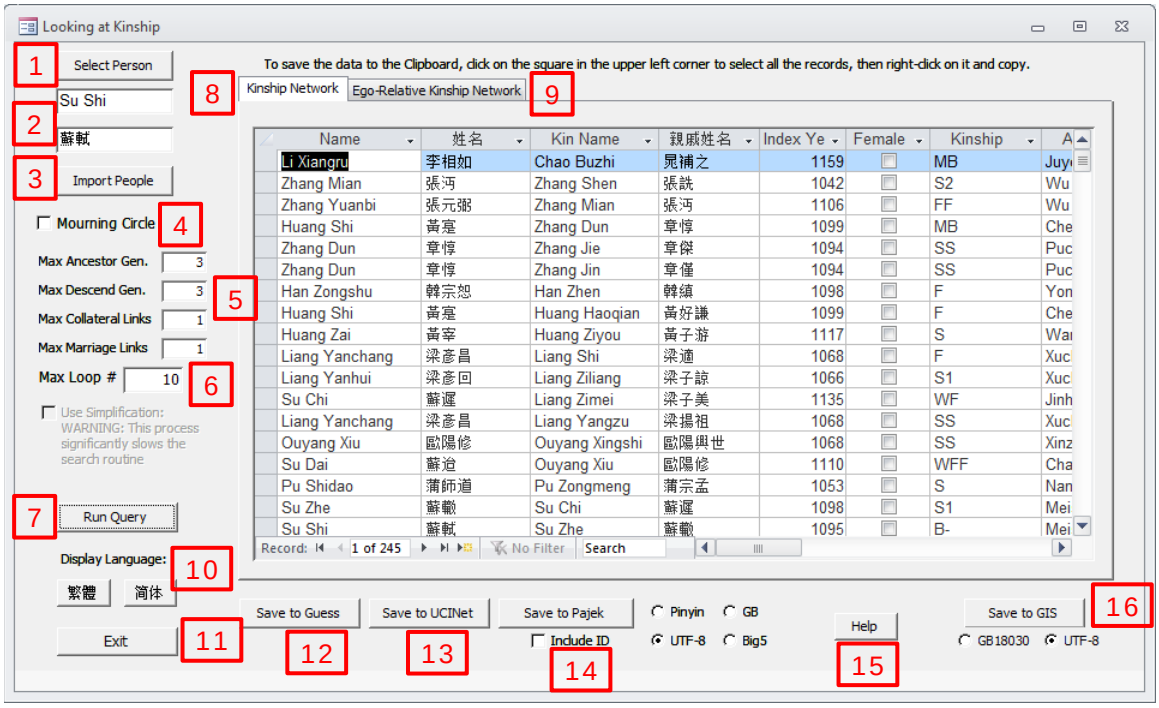
2. Structure of the Kinship Data

Kinship has the simplest data structure of all CBDB entities. It includes:

Person ID
Kinship code
Kin ID
Source information

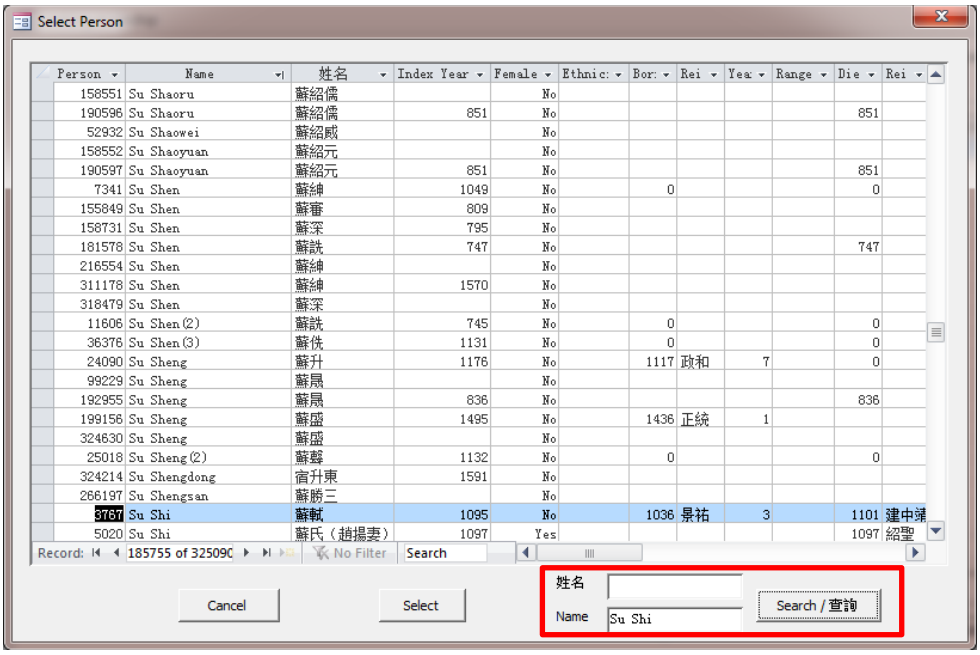
That is it. There are 477 codes for different kinship relationships at present, but in the future, those will be reorganized through codes that will cluster them better.

B. The Interface



1. ***Select Person*** *Command Button*

The first step in using the **LookAtKinship** form is to select a person. When one clicks on the **Select Person** command button, this opens a form:



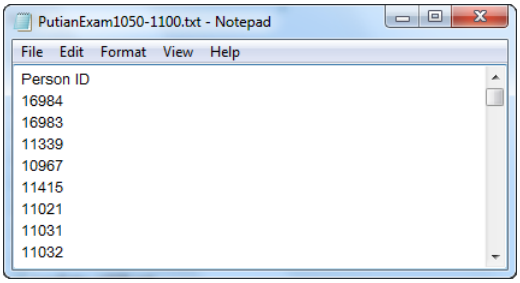
One can search for a person using either Chinese characters or *pinyin*. (Note that sometimes, for a search on a second use of the form, the form finds an unrelated person. It finds a correct person if one simply clicks **Search** again to reset the results.)

2. ***Selected Person*** *Text Boxes*

The **LookAtKinship** form displays the name of the selected person in Chinese and *pinyin* in these two text boxes. The user cannot directly type a person's name into these boxes.

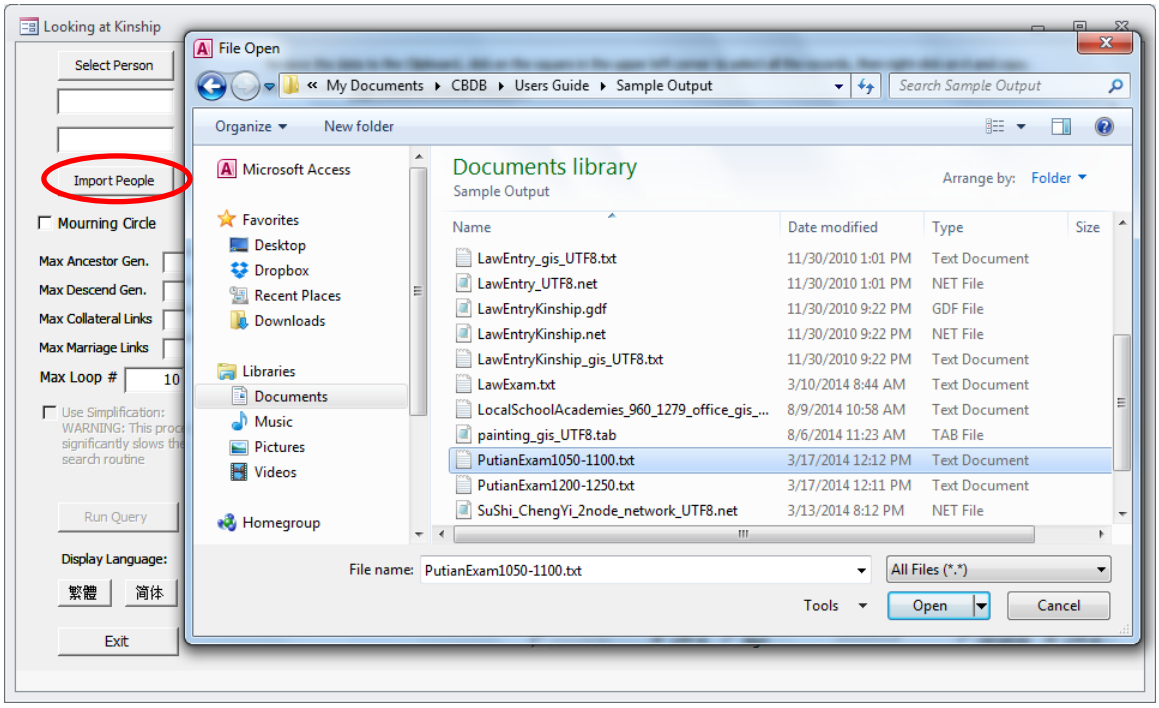
3. ***Import People*** *Command Button*

At times, one want to look at clusters of kinship relations to explore a connections between people who are parts of groups selected through other criteria. For example, one can look at the kinship relations between all the men from Putian in Fujian who attained the *jinshi* degree over particular periods of time. One can use **LookAtEntry** to identify lists of *jinshi* degree holders from Putian and export the list to a *text file*:

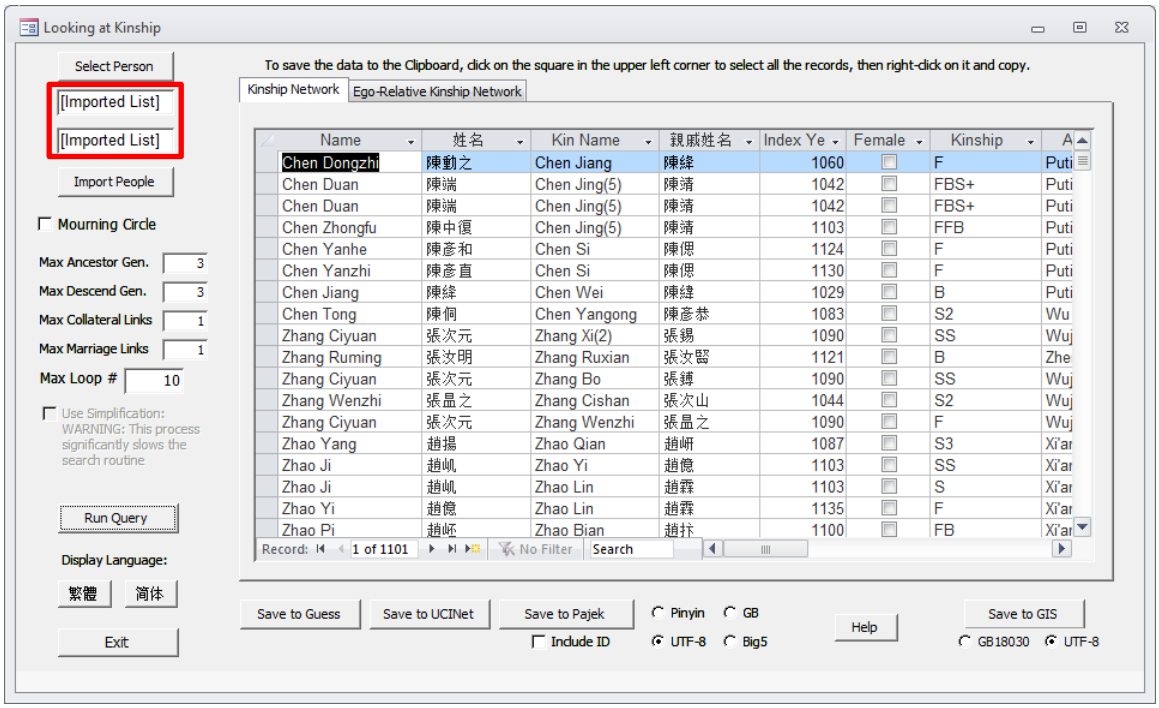


Note that CBDB ignores the first line, “Person ID” because it does not begin with a number. CBDB looks only at numbers starting in the first column: once can include the names of the people in the file, so long as they follow the IDs.

Once one has a list of people, one clicks on the **Import People** command button to open a Windows form:



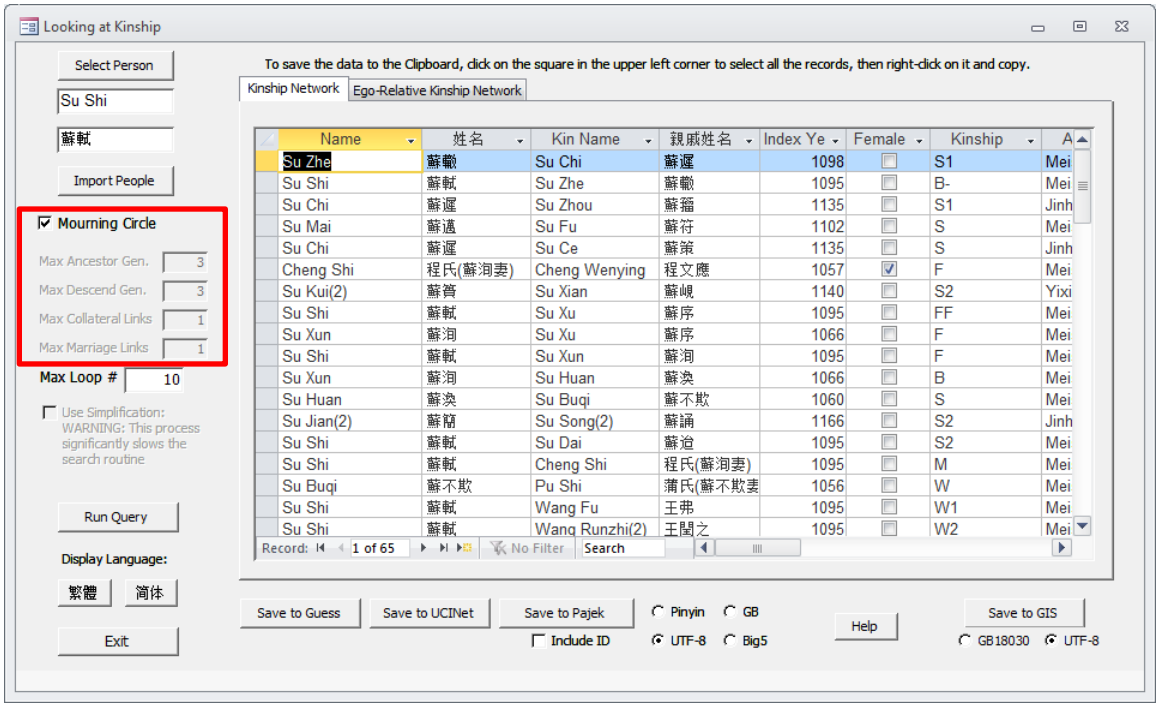
One selects the file and clicks on the **Open** command button. **LookAtKinship** imports that file and displays “[Imported List]” in the **Person Name** text boxes:



*4. **Mourning Circle** Check Box*

The *mourning circle* 五服 is an important ritual concept in pre-modern China. It is the group of 159 kinship relations for which a person has a duty to mourn when the kin dies. The degree of mourning—the type of clothes, the length of mourning, etc.—differ as the kinship distance grows greater, but nonetheless those within a mourning circle form a ritual community for a person. CBDB allows a user to identify as many people within that community as possible. Note that the algorithm is not perfect because it has some *path dependency*. That is, CBDB attempts to identify all of a person's close kinship relations from the available data, but if, for example, a person's uncle appears in the data only as the great-grandson (SSS) of the person's great-great-grandfather (FFFF), CBDB may not conclude that the relationship FFFFSSS (for whom one does not have a mourning duty) is in fact "uncle" FB.

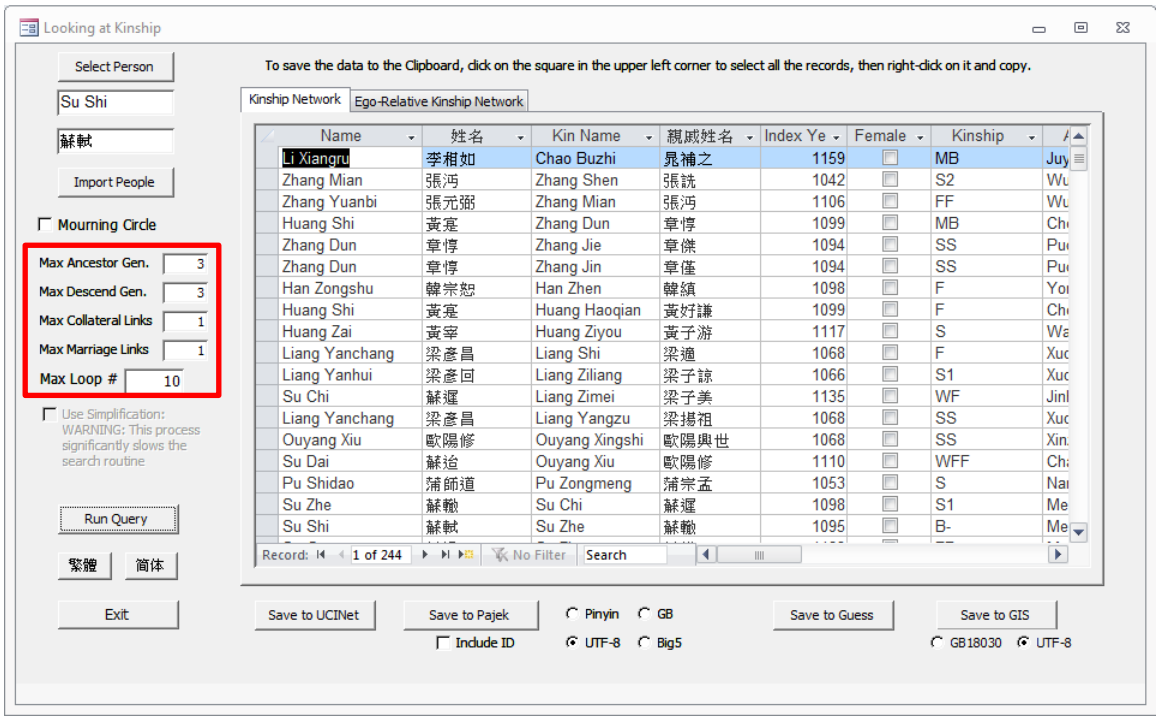
When one clicks on the **Mourning Circle** check box, the form does not allow one to further specify the kinship distances (see below):



*5. **Kinship Distances** Text Boxes*

Because LookAtKinship keeps looking through a very large table of kinship relations until the distance limits are reached, the kinship table produced by the search can grow very large. Therefore please note:

> **WARNING: searching for extended degrees of collateral and marriage distance may result in a very large dataset.**



For example, using 3 generations up, 3 generations down, 1 collateral, and 1 marriage produced 244 kinship relations for Su Shi. Increasing the marriage distance to 2 produces 438, and then increasing the collateral distance to 2 yields 591. Because the searches are usually fairly quick, one can afford to experiment with the parameters.

6. ***Maximum Number of Loops*** *Text Box*

This parameter controls how many times (at most) CBDB repeats the search looking for kin of kin. To be safe, the number always should be at least the sum of all the distance parameters. However, because the search routine also stops when a new search does not produce any additional kinship records, the **Maximum Loop** sets the upper limit, but the search may not perform that many iterations.

7. ***Run Query*** *Command Button*

Once the user has selected a person or successfully imported a list of people, the **Run Query** command button becomes enabled. One clicks on this button to run the query.

8. ***Kinship Network*** *Table*

(a) Overview

The search routine produces two different tables of kinship records. The first is the **Kinship Network**. This table lists all the relevant kinship records as they appear in the kinship data source table. It begins by listing all those records that connect the

target person (people) with kin. It then lists the relations of those initial kin with other kin, and so on.

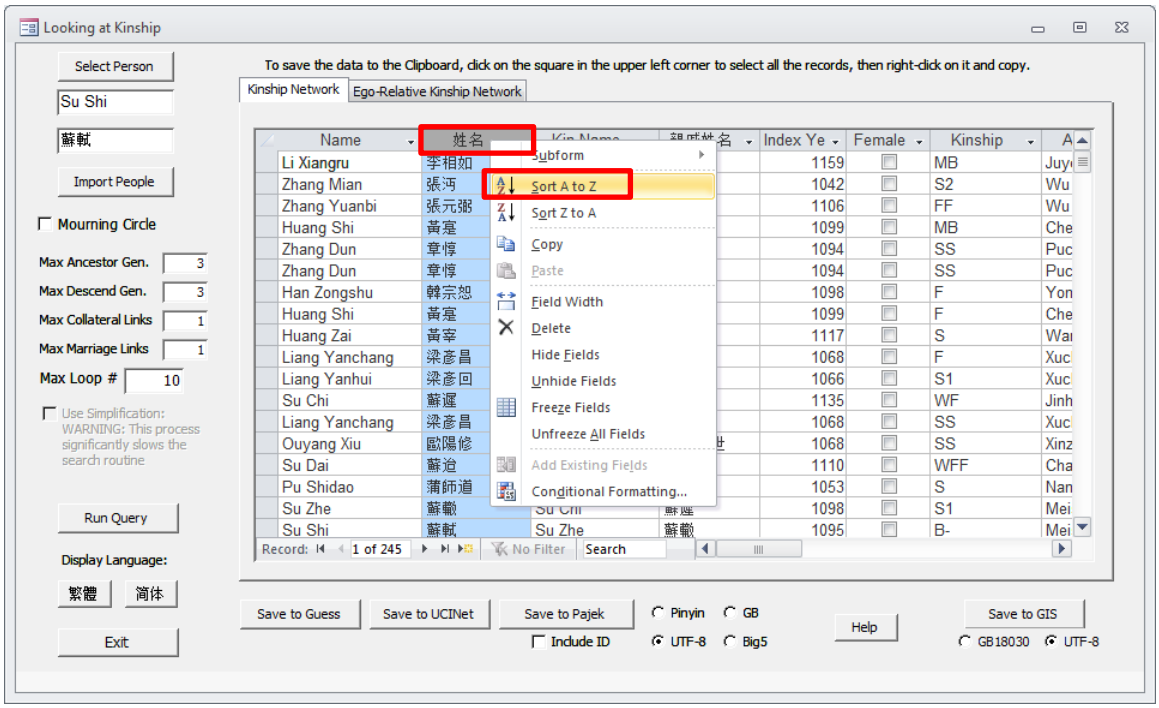
(b) Structure of the Table

This table in **Kinship Network** has nineteen columns:

Name (*pinyin* and Chinese)
Kin Name (*pinyin* and Chinese)
Index Year of Kin
Sex of Kin
Kinship Relation
Address of Person (*pinyin* and Chinese)
Address Type
XY-Coordinates of Address
Address of Kin (*pinyin* and Chinese)
XY-Coordinates of Kin Address
Address Type of Kin Address
Notes
Distance (great-circle arc distance in kilometers between the addresses)

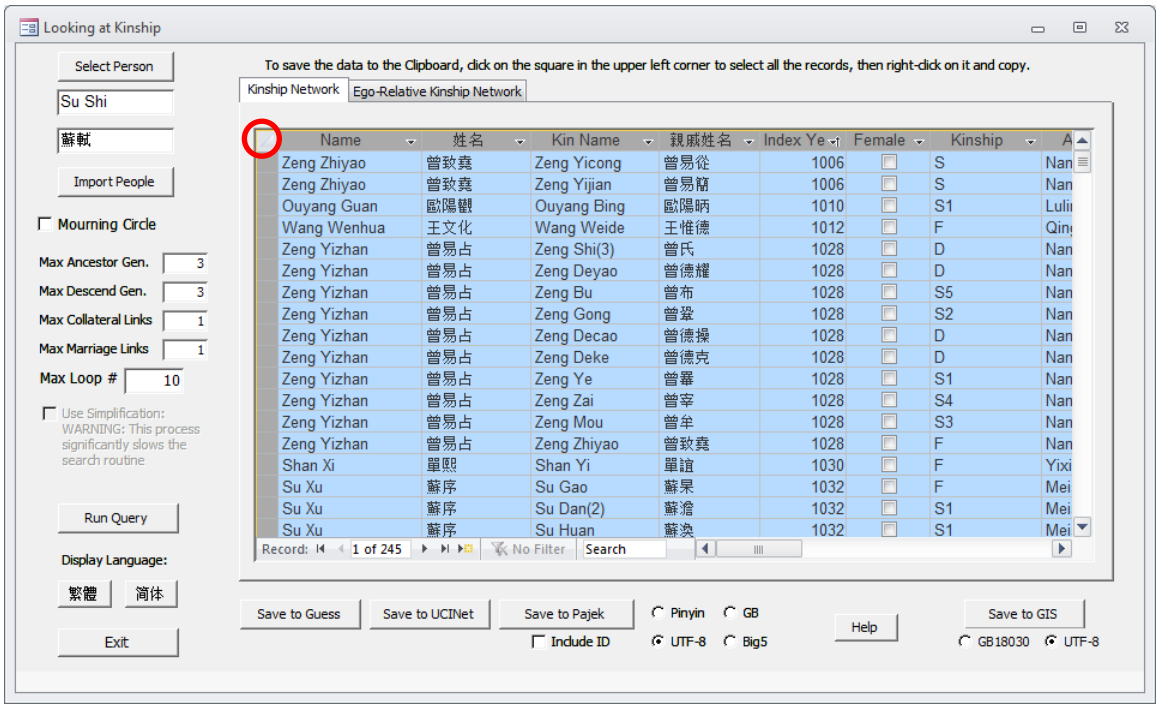
(c) Sorting the Table

One can **sort** the table using any of these columns. "姓名" (Name in Chinese) may be useful. *Left-Click* on the column name to select the column and then *Right-Click* to choose the sorting option (either smallest to largest or largest to smallest):

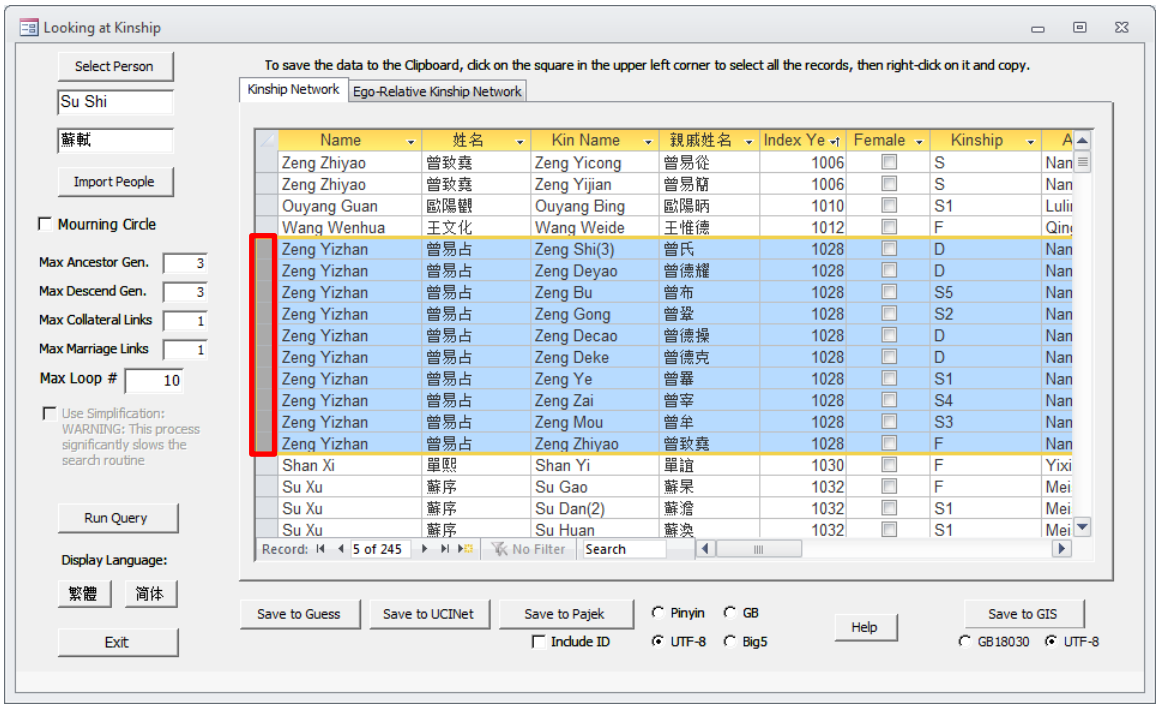


(d) Exporting the Search Results

If one wishes to export the results of a search, one can click on the square in the upper left-hand corner of the table to select all the records in the table. One can then copy the table in the usual manner (**Ctrl-c**), open **Excel** (or any other compatible *spreadsheet* software) or create a new *text* file and open it with **Notepad** (or any preferred software) and paste the results (**Ctrl-v**).

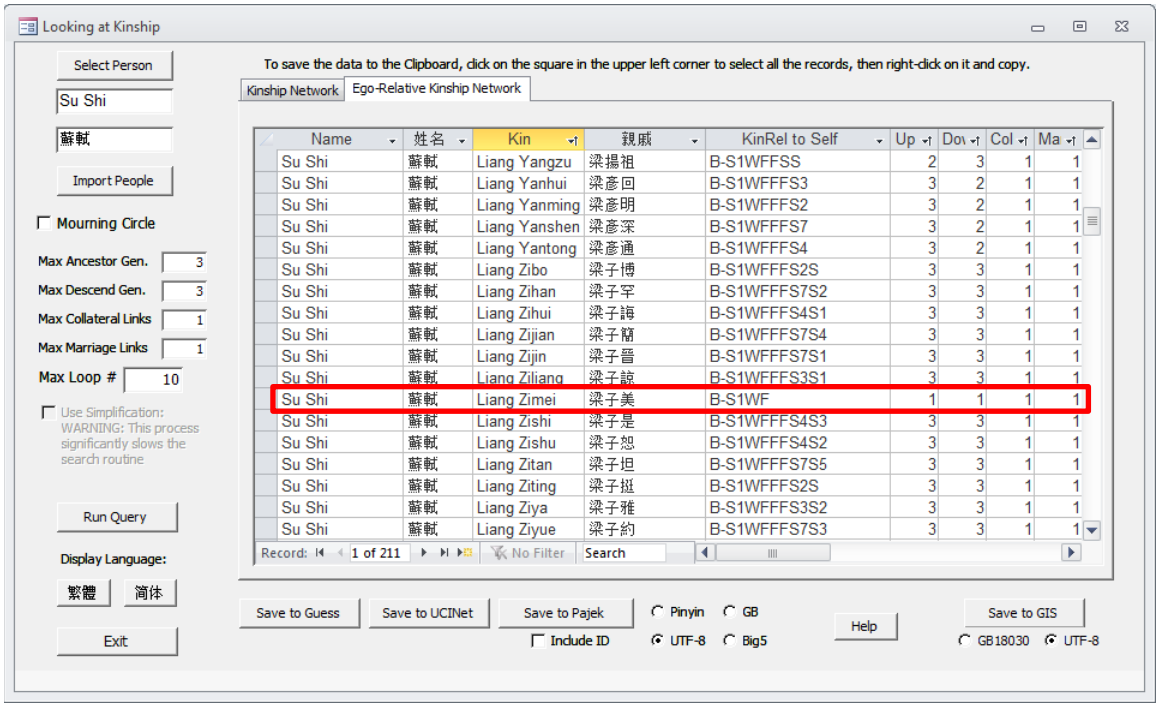


One also can select a specific block of records to copy by left-clicking on the left-hand column of the table, holding the mouse-button down and dragging to copy the block:



9. ***Ego-Relative Kinship*** *Table*

The second table of results, **Ego-Relative Kinship**, describes the kinship relation between each person in the first table and *the person selected at the very beginning*:

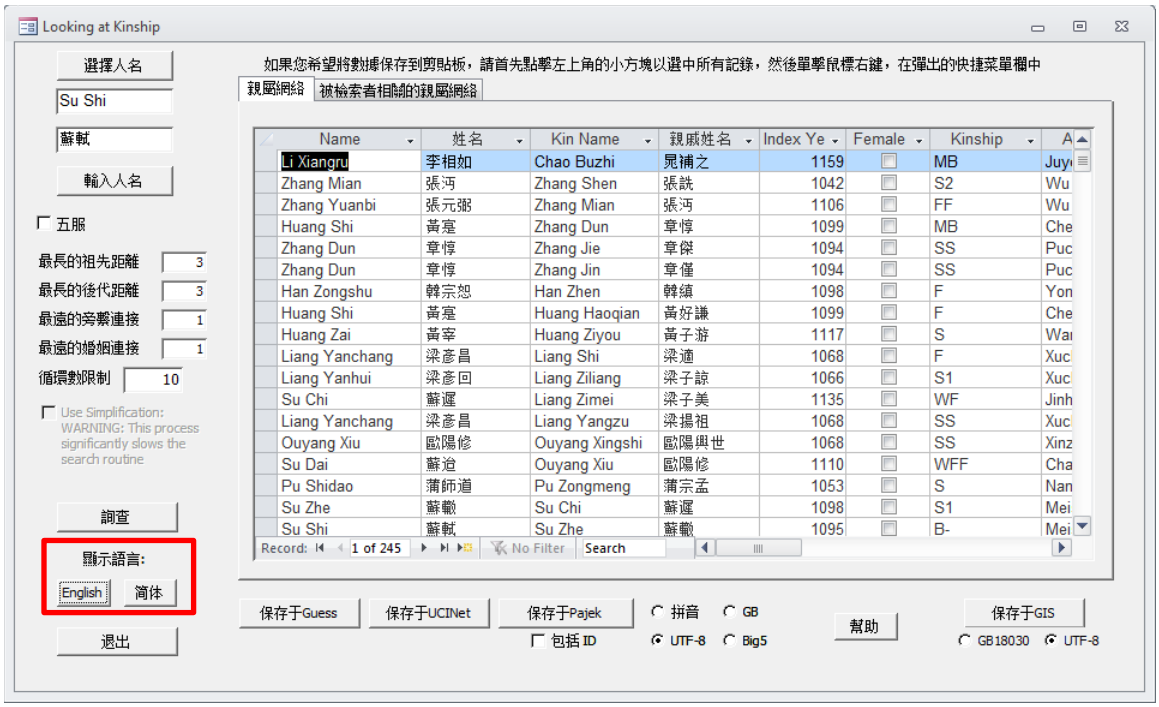


For example, Liang Zimei 梁子美 is Su Shi's younger brother's son's wife's father. The path one traverses to reach Liang Zimei requires one to go over one collateral link (B-), down one generational link (S), over one marriage link (W), and finally up one generation (F). Thus the kinship distance metrics are {1,1,1,1}. If one looks for Liang Zimei in the **Kinship Network** table, one finds him as Su Chi's 蘇遲 wife's father.

As is true for all the other tables in all the other forms, if one clicks on the upper left corner of either table in this form, once selects all the records in the table, which then can be cut and pasted into other programs. Also, right-clicking on any of the column headings will allow one to sort on that column.

10. ***Display Language*** *Command Buttons*

Note that all of the forms have the option to switch the display labels among English, traditional or simplified Chinese. When one click on the "繁體" command button, all the form labels switch to traditional Chinese text (except those that are parts of bilingual pair, like "Name" and "姓名" in the table of results), and the button gives one the option to return to English:



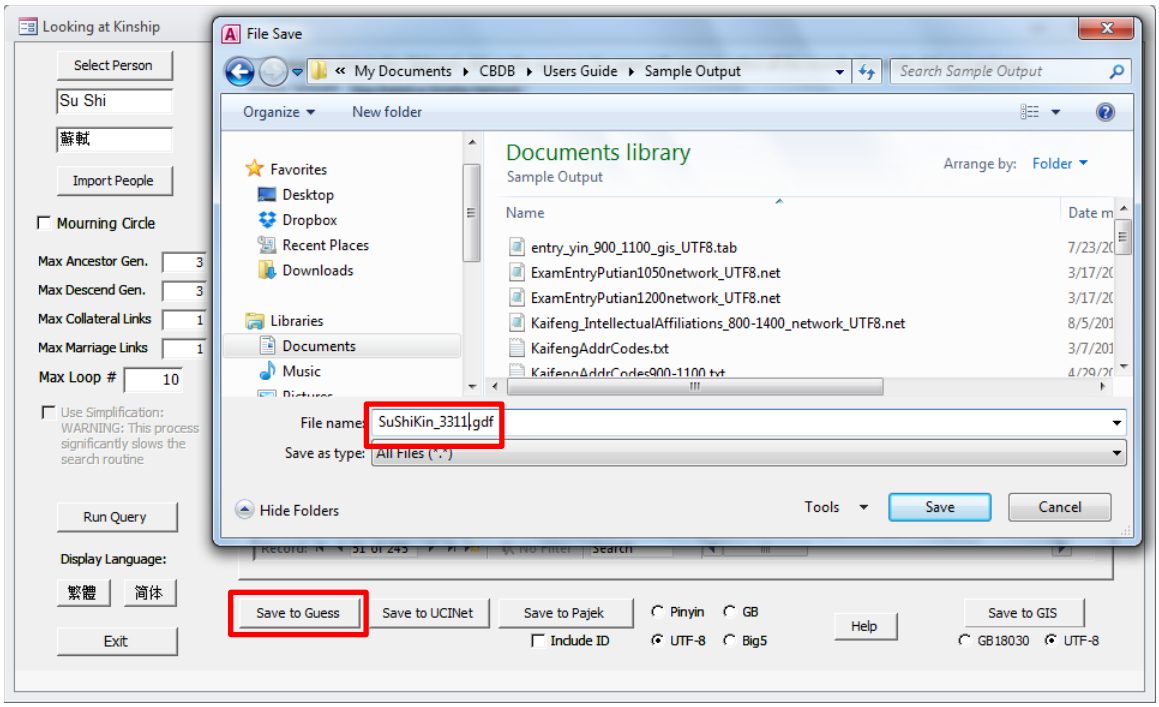
*11. **Exit** Command Button*

To close the form, one clicks on the **Exit** command button (or on the X in the upper right corner)

*12. **Save to Guess** Command Button*

Since association data provides an implicit social network linking the groups of people connected by the category of association being examined, one can save the network for analysis in the Guess format by clicking on the **Save to Guess** button. The Guess format is one standard format for visualization in *social network analysis* (SNA), and other programs can read it and convert it into their formats.

Clicking on the **Save to Guess** button opens a standard Windows "File Save" window. When one has decided where to save the file and what to call it (note that the default file extension is **.gdf**) and clicked Save, CBDB writes the file.



The GUESS file includes the following fields:

For the node (the person)

Name (*pinyin* and Chinese)
Index year
Sex
Kinship distance metrics from selected person
Node color (based on the sum of the distance metrics)

For the edge (the relationship between people)

Kinship relation
Edge color (based on the kinship distance of relation)

The color coding of kinship for nodes is:

White = the target node;
Blue = nodes at a summed kinship distance of 1
Green = nodes at a summed kinship distance of 2
Orange = nodes at a summed kinship distance of 3
Yellow = nodes at a summed kinship distance of 4
Red = nodes at a summed kinship distance of 5 or more

*13. **Save to UCINet** Command Button*

UCINet is another social network analysis (SNA) program like GUESS and Pajek. The procedure for saving the kinship relations for a person in UCINet form is the same as for GUESS: clicking on the **Save to UCINet** button opens a Windows "File Save" window. One selects the location and provides an file name (the default

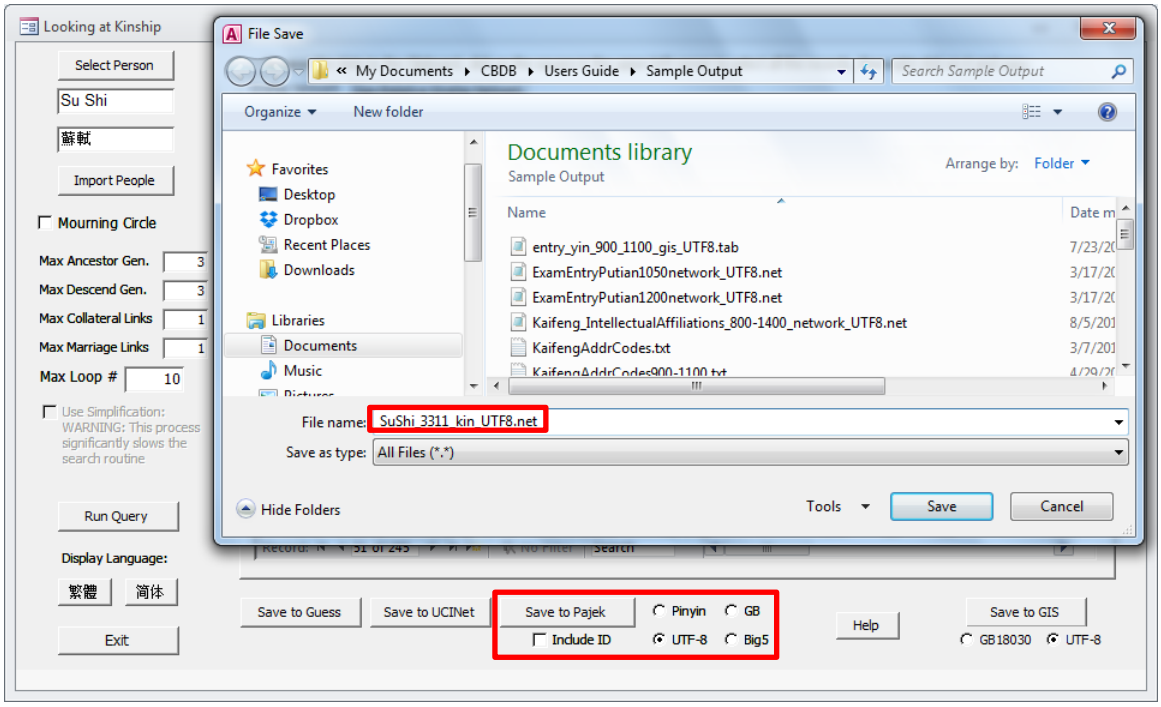
extension is **.vna**) and clicks on the Save button. For UCINet, CBDB provides the following information:

For nodes (people)
- CBDB Person ID
- Name in *pinyin*
- Index year
- Sex
- Kinship distance from target person
- XY coordinates
- Node color (based on kiship distance)

For edges (the kinship relationship)
- Kinship relation

*14. **Save to Pajek** Command Button*

Pajek (along with UCINet) is perhaps the most broadly used SNA file format. CBDB allows files for Pajek to be saved in different text encodings to enable the use of Chinese characters. Depending on whether the user is in a Chinese Windows environment or an English will determine whether to use GB18030 (Chinese) or UTF-8 (English) encoding for the Chinese characters. Note that there also is an option to include the *Person ID* with the node information in the Pajek files. Once the user has selected the appropriate options and clicks on **Save to Pajek**, a window will open for the user to choose the name and location of the Pajek file. The default extension is ".net:"



The information saved in a Pajek file is more limited than for either GUESS or UCINet:

For the nodes (people)

Name in Chinese **or** *pinyin*, depending on one's selection
Node color based on kinship distance from target person
CBDB Person ID, if selected

For the edges (relationships)

Kinship relation
Edge color based on kinship distance of relationship between nodes

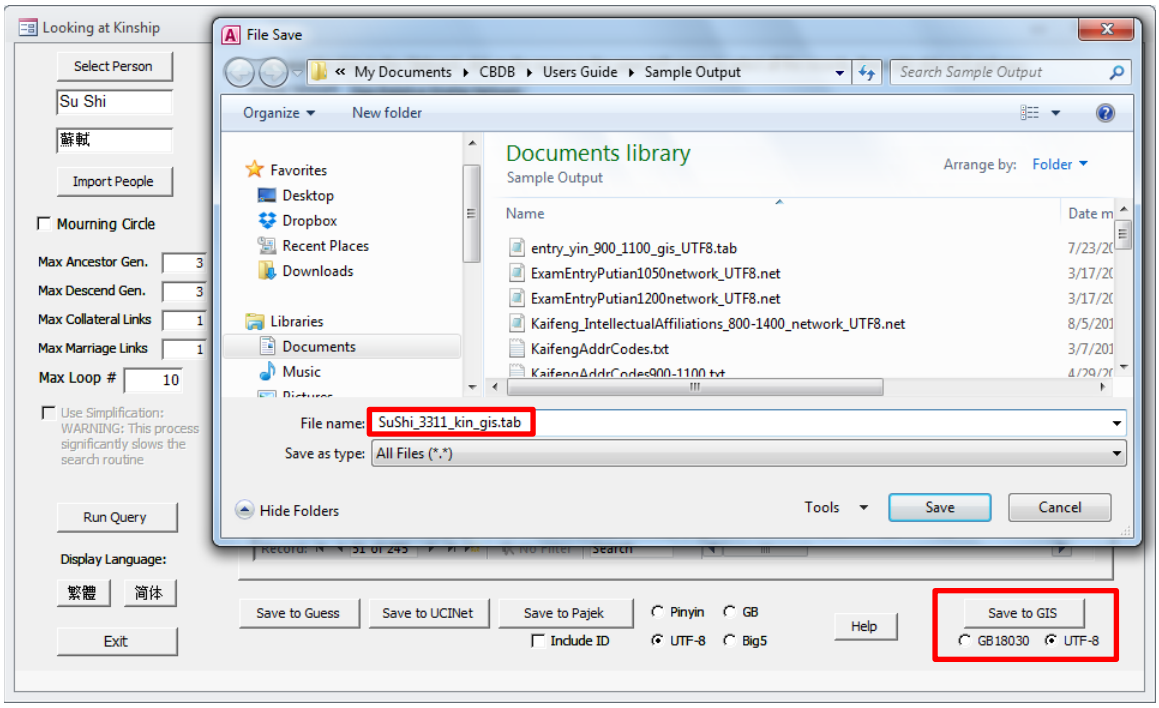
*15. **Help** Command Button*

The **Help** command button opens this file.

*16. **Save to GIS** Command Button*

Data on associations have a geographic component and may reveal significant geographic patterns of distribution that shift over time. In order to allow the user to explore this geographic component, CBDB exports the GIS (Geographic Information Systems) data to a commonly used file format (tab-delimited text) that can be imported into such programs as ArcGIS and QGIS.

Depending on whether the user is in a Chinese Windows environment or an English will determine whether to use GB18030 (Chinese) or UTF-8 (English) encoding for the Chinese characters. Or, if one does not want *any* characters, one can simply select *Pinyin*. To save the data to a GIS file, the user selects the appropriate encoding and clicks on **Save to GIS**. This will open a standard Windows "File Save" window. Note that the default extension is ".tab" and if the user gives the file a different extension, CBDB will automatically append ".tab" to the file name:



When **LookAtKinship** saves the file, the record for each person includes the following fields:

Name (Pinyin and Chinese)
Male/Female
Index Year
Kinship Relation (with the target person)
Kinship Distance Metrics (with the target person)
Place Name (Pinyin and Chinese)
XY coordinates
XY count (the number of people in the table who share the same coordinates)