Ridom EpiCompare

Version 1.0

Contents

Ridom GmbH – Sedanstr. 27 – D-97082 Würzburg, Germany - email: info@ridom.de

1. Introduction

Epidemiological typing has many different applications. Those include surveillance of infectious diseases, outbreak control and others. Often two ore more different typing methods are available. There are several criteria for evaluating the power of typing methods and for comparing different methods. The underlying mathematical formulas are often quite simple, but the handling of the data can get demanding for large data sets.

The program *Ridom EpiCompare* can calculate various criteria for data of different typing techniques. It combines easy calculation with ease of data handling.

Ridom EpiCompare can calculate the following parameters:

- **Typability** is the percentage of distinct microbial strains which can be assigned a positive typing marker.

- **Reproducibility** is the percentage of strains that give the same result on repeated testing.

- **Discriminatory power** of a typing method is its ability to distinguish between unrelated strains.

- **Typing system concordance** helps to compare different independent typing systems.

2. Features

- Import data from files (MS Excel, CSV) and from Clipboard

- Easy handling of data: edit, sort, reorder, print

- Calculation of different performance criteria for data:

  - Typability

  - Reproducibility

  - Discriminatory Index

  - Typing System Concordance

- Export results as HTML-file

- Export via clipboard or files (MS Excel, CSV).

3. Installing *Ridom EpiCompare*

System Requirements

- An installed Java Runtime Environment 1.5 or higher is required to run *Ridom EpiCompare.* Java is freely available from www.java.com.

- An installed PDF Viewer (e.g., Adobe Reader, www.adobe.com) is needed for viewing the manual.

- *Ridom EpiCompare* was tested with Java 1.5 on Windows NT4.0, 2000, XP, and Vista.

Installing Instructions

- Download from www.ridom.de the latest version of *Ridom EpiCompare.*

- Double click the downloaded file and follow the on-screen instructions (please remember, you have to install Java first before you are able to install *Ridom EpiCompare*).

- *Ridom EpiCompare* can now be started by clicking the software's icon (in the installed directory, on the desktop or in the *Start* Menu).

Updating the Software

A download option for on-line updates is integrated into the software (*Help* menu). For disabling the auto-update feature or for proxy-settings, use the Preferences-dialogue from the *File* menu.

Ridom EpiCompare Help

Use the button in the *Help* menu to view this PDF-document (the document includes a Quick Reference with menu and keyboard shortcuts).

4. Getting Started

Keyboard Shortcuts

Frequent occurring actions can be conveniently invoked by keyboard shortcuts. All menu commands can also be invoked (e.g., if no mouse is available) by using the Alt-key in combination with a character as indicated by underscore in the menu (e.g., for the command "Open..." press Alt+F, O).
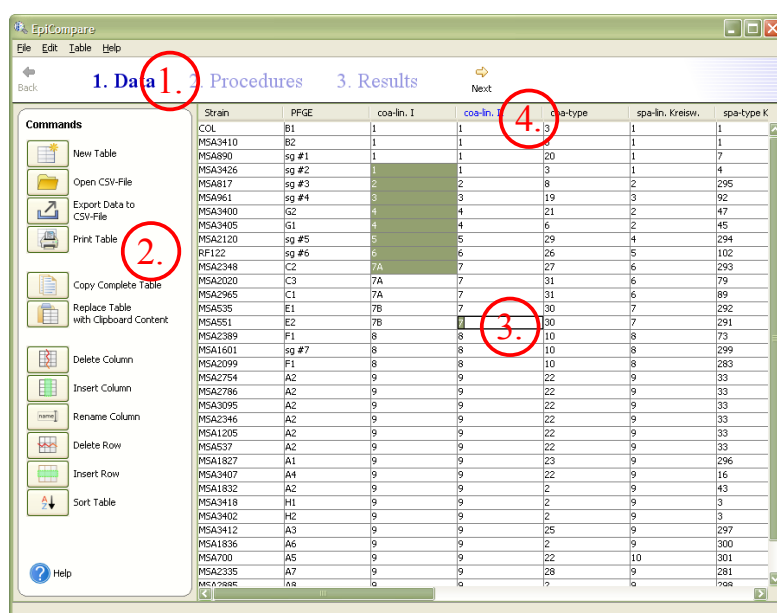
Figure 1: Data View. The buttons for switching views are below the menu bar (1). The command pane (2) offers the currently available commands. The currently edited cell is indicated (3). A blue table head, as shown in (4), marks the currently sorted column.

Switching Views

The pane below the menu bar is used for switching different views. There are three views: *Data*, which contains a table for the used data, *Procedures*, which allows selection of calculation procedures and *Results*, where the calculation results are displayed. Click on the text in the pane or use the *Back* and *Forward* buttons to switch the views.

Import Data from a File

Use the button [icon] in the command pane to open a file. Several import-options can be set:

- whether the file contains a header in the first line or not,

- if a CSV-file is chosen, the delimiter can be set. In most cases this will be ";" or ",". Take a look at the preview-table to find out if the delimiter is chosen correctly

- each column in the table needs an unique name. Columns that have a name assigned are displayed in white. If a column has no name assigned, it is displayed in a light red colour. Click on a column and select a predefined name or specify a name by selecting the command *Enter Column Name...* from the context menu.

Import Data from Clipboard

It is also possible to import a table from the clipboard. This feature supports Microsoft Excel clipboard format. This means that you can select a table [icon] (or parts of it) in MS Excel, copy it to the clipboard and paste it into the program. Press the [icon] button to replace the table with the current clipboard content. You will be asked if the first row should be used for selecting column names. If this applies and two or more columns have the same name, numbers are attached to the names to make them unique.

Data View

In the Data View (see Figure 1) each row contains one isolate, and each column contains a specific attribute. The attribute names are unique, this means that two columns cannot have the same name. The table columns can be moved by dragging them with the left mouse button. Table columns can be selected by clicking their column-head. Multiple columns can be selected by pressing the Ctrl-key while clicking the column-head.

To edit the table, double-click a cell. The Tab-key can be used to traverse the table's cells while editing. The *Table* menu contains functions for modifying and sorting of the table. It is also possible to sort the table by double clicking a column head with the left mouse button. The head of the sorted column appears blue.

It is possible to merge two or more columns with the *Merge Columns* button. A new column is created with a name made out of the names of the merged columns and the new column is filled with the concatenated values of these columns. Be careful with cells containing non-typable values. Check that the merged cells contain the correct data for your usage!

Procedures View

Figure 2 shows the view that is used for selecting the procedures. See next chapter for calculation details. A pane allows the input of strings for missing or not non-typable values. The default for missing is empty text and for non-typable values "NT". Change these if your dataset uses other strings.

It is possible to select multiple lines in the lists by using the Ctrl-key while clicking with the mouse-button. Typing system concordance and reproducibility can only be calculated for two attributes at once.
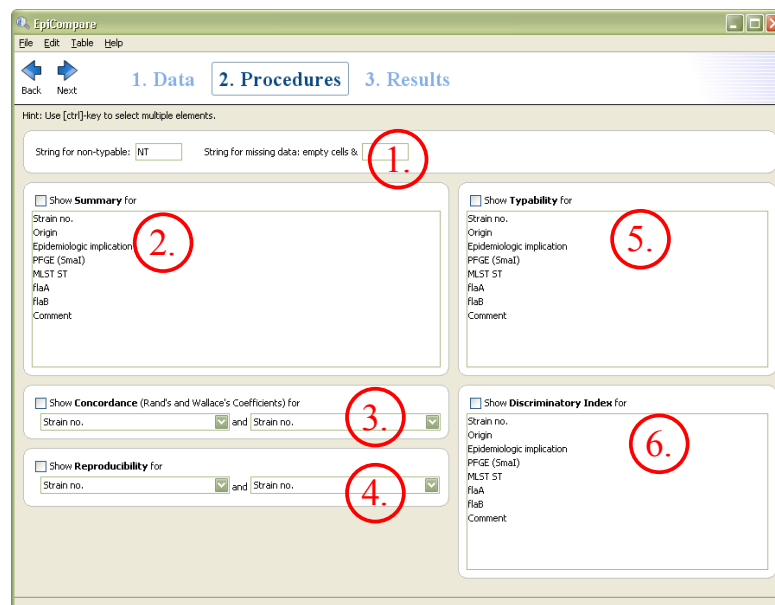

Figure 2: Procedures View. The upper pane (1) contains the text for non-typable and missing isolates. Selecting attributes in (2) displays a summary statistic for this attribute. Select (3) – (6) for calculation of: (3) Typing System Concordance, (4) Reproducibility, (5) Typability and (6) Discriminatory Index.

Results View

The results view (see Figure 3) displays the calculation results of the selected procedures. References for the different procedures are given.

The results can be saved and printed. In addition, they can also be exported to the clipboard and used in word-processor programs, like Microsoft Word.
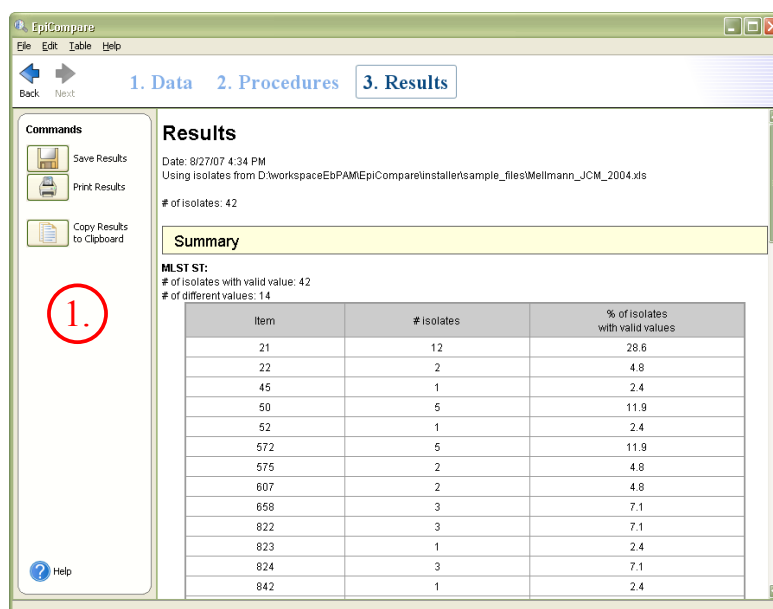

Figure 3: Results View. Use the command pane (1) to export or print the results.

Send Bug Report

Ridom EpiCompare includes a feedback dialogue to provide the user with a convenient way of sending messages directly to the developers such as features requests and bug reports. This dialogue can be opened in the *Help* menu.

If an program error occurs a window opens automatically. The report already contains the internal description of the error that occurred together with some system information (*Ridom EpiCompare* version, OS and Java version). If the error occurred after a specific user transaction it would be of help to us in solving the problem, if a description of the circumstances leading to the problem is given in the comment field.

5. Calculation Details

Typability

The typability is the proportion of strains that are assigned a type by the typing system. The formula reads:

$$T = \frac{N_t}{N}$$

with:

N_t : the number of isolates assigned a type and N : the number of isolates tested.

Ridom EpiCompare uses:
N = Number of isolates that show a non-missing value
N_t = Number of isolates that show a non-missing, typable value

See [Hunter, Gaston 1988], [Struelens, 1996].

Reproducibility

Reproducibility is the ability of a typing system to assign the same type to a strain tested on independent, separate assays. The formula reads:

$$R = \frac{N_r}{N}$$

with
N_r : the number of isolates assigned the same type on repeated testing and
N : the number of isolates tested.

Ridom EpiCompare uses the following data:
N = Number of isolates that show a non-missing value (N includes non-typable values).
N_r = Number of isolates that show a non-missing value and that have the same value on repeated testing. If both results are non-typable it is also counted here.

See [Hunter, Gaston 1988], [Struelens, 1996].

Discriminatory Index

Discriminatory index is calculated with the following formula:

$$D = 1 - \frac{1}{N(N-1)} \sum\nolimits_{j=1}^{s} n_j(n_j-1)$$

where

N is the total number of strains in the sample population, s is the total number of types described, and n_j is the number of strains belonging to the jth type.

A 95% Confidence Interval (CI) can be calculated for D with the following formula:

$$CI = [D - 2\sqrt{(\sigma^2)}, D + 2\sqrt{(\sigma^2)}] \text{ with } \sigma^2 = \frac{4}{N}\left[\sum \pi_j^3 - (\sum \pi_j^2)^2\right] \text{ and } \pi_j = \frac{n_j}{N}.$$

Ridom EpiCompare uses all isolates that have no missing data and are typable for calculation of discriminatory index and 95% confidence interval. Isolates with missing data or isolates that are non-typable are ignored in this calculation.

See [Simpson, 1949], [Hunter, Gaston 1988], [Grundmann et. al. 2001].

Typing System Concordance

Use typing system concordance to compare the results of different typing systems. For calculation, pairs of the results of 2 typing systems $(x, y): x, y \in Isolates$ of all isolates are created. All pairs are arranged in a 2x2 table, according to their type in both typing systems:

	System 1, +	*System 1, -*
System 2, +	**A** # of pairs (x,y) for which x and y have the same type in System 1 and in System 2	**B** # of pairs (x,y) for which x and y have the same type in System 2 and a different type in System 1
System 2, -	**C** # of pairs (x,y) for which x and y have the same type in System 1 and a different type in System 2	**D** # of pairs (x,y) for which x and y have a different type in System 1 and in System 2

The ratio $\dfrac{(A+D)}{A+B+C+D}$ is the frequently called in publications the typing system concordance. This value is also known as Rand's coefficient. However, if comparing two random data sets, concordance calculation does not approach the value 0. Therefore, when calculating the Adjusted Rand coefficient, a correction factor is used to take into account the presence of chance agreement:

$$Adjusted\ Rand = \frac{A+D-n_c}{A+B+C+D-n_c}\ ,\ with$$

$$n_c = \frac{n(n^2+1)-(n+1)\sum n_i^2-(n+1)\sum n_j^2+2\sum\sum\dfrac{n_i^2 n_j^2}{n}}{2(n-1)}$$

where *n* represents the total number of isolates and n_i, n_j represent the numbers of isolates with values *i* and *j* for System 1 and System 2.

Wallace's coefficients are calculated by

$$W_1 = \frac{A}{A+B}\ \ and$$

$$W_2 = \frac{A}{A+C}$$

They represent the probability that a pair of isolates that have the same type in System 1 also have the same type in System 2 and vice versa.

See [Carriço et. al. 2006].

6. References

Carriço J.A., Silva-Costa C., Melo-Cristino J., Pinto F.R., de Lencastre H., Almeida J.S., Ramirez M.: Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant Streptococcus pyogenes, J. Clin. Microbiol. 44:2524-2532, 2006

Grundmann, H., Hori, S., Tanner, G.: Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms, J. Clin. Microbiol., 39:4190-4192, 2001

Hunter, P.R., and Gaston, M.A.: Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity, J. Clin. Microbiol. 26:2465–2466, 1988

Simpson, E.H.: Measurement of diversity, Nature 163:688, 1949

Struelens M.J.: Consensus guidelines for appropriate use and evaluation of microbial epidemiologic typing systems, Clin. Microbiol. Infect. 2:2-11, 1996

7. Example Files

There are two example datasets included (as MS Excel and CSV-files). They contain example data from the following publications:

- `Mellmann_JCM_2004`
  Mellmann, A., Mosters, J., Bartelt, E., Roggentin, P., Ammon, A., Friedrich, A., Karch, H. and Harmsen, D.: Sequence-based typing of *flaB* is a more stable screening tool than typing of *flaA* for monitoring of *Campylobacter* populations, J. Clin. Microbiol. 42:4840–4842, 2004 (you will note a small calculation error in this publication. Sorry for the error, but we did not had *Ridom EpiCompare* available at that time :-)

- `Koreen_JCM_2004`
  Koreen, L., Ramaswamy, S.V., Graviss, E.A., Naidich, S., Musser, J.M. and Kreiswirth, B.N.: spa typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation, J. Clin. Microbiol. 42:792-799, 2004

8. Quick Reference

Menu Functions

Menu Item	Icon	Shortcut	Description
File-Menu			
New Table...		Ctrl+N	Replaces the current table with a new one
Open File...		Ctrl+O	Opens a file (MS Excel or CSV)
Export Data...		Ctrl+E	Exports the current table to a file (MS Excel or CSV)
Print Table...		Ctrl+P	Prints the current table
Save Results As...		Ctrl+S	Saves the results of the calculation as a HTML file
Print Results...		Ctrl+P	Prints the results of the calculation
Preferences...			Preferences for automatic updates and proxies
Exit		Ctrl+Q	The preferred way to exit the program
Edit-Menu			
Copy Complete Table		Ctrl+Shift+C	Copies table to clipboard (in MS Excel format)
Replace Table with Clipboard Content...		Ctrl+V	Replaces the table with clipboard content. Table in clipboard must be in MS Excel format
Copy Results to Clipboard		Ctrl+C	Copies results of calculation to clipboard. Clipboard content can be used further e.g. in word processor programs
Table-Menu			
Delete Column		Alt+Del	Deletes the selected table columns
Insert Column...		Alt+Enter	Inserts a table column after the selected column or, if no column is selected, before the first column
Rename Column...		Alt+R	Changes the name of the selected column
Merge Columns		Alt+M	Merges two or more columns
Delete Row		Ctrl+Del	Deletes the selected table rows
Insert Row		Ctrl+Enter	Inserts a row after the first selected row, or, if no row is selected, before the first row
Sort Table...			Sorts the table according to a chosen column

Menu Item	Icon	Shortcut	Description
Help-Menu			
EpiCompare User Guide (PDF)		F1	Opens the program manual PDF in an external viewer (e.g., Adobe Reader)
Software Update - Online Update...			Searches for available software updates at the Ridom web server
Software Update - Import Update File...			Imports a software update file that was manually downloaded from the Ridom web server
Ridom Support - Send Bug Report / Feature Request...			Sends a bug report or a request of new feature
Ridom Support - Support Home Page			Opens the Ridom support home page with the default WWW browser
About...			Opens the Ridom EpiCompare version info (includes also a system overview)

9. *Ridom EpiCompare* Frequently Asked Questions (FAQ)

Can Ridom EpiCompare open Microsoft Excel Files?

Microsoft Excel Files (with a file name typically ending with .xls) can be read directly with *Ridom EpiCompare*. For other spreadsheet programs, you must convert the data into CSV-files (comma separated values).

The procedure "xyz" is missing! Will it be included in future releases?

Maybe. But first, we must know that someone wants the procedure "xyz" to be included. If you are interested in new procedures or new features, you can write an email to info@ridom.de or use the Send *Bug Report / Feature Request...* command from the help-menu to contact Ridom GmbH.

10. End User Licence Contract for *Ridom EpiCompare*

Permission is granted to anyone to use this software FREELY for any purpose, subject to the following restrictions:

I. General
a) This End User Licence Contract (EULC) forms a legally binding contract between you, the licensee, and Ridom GmbH, the licensor for the Ridom EpiCompare software and related media, materials and documentation (the software product).
b) If you do not agree to these conditions, you are not entitled to use the software.
c) The person who concludes this EULC assures and makes warranty that he is entitled to declare a legally binding contract on behalf of his employer as the licensee.
d) The software product is licensed not given away.

II. Copyright
a) The software product distributed by Ridom GmbH is protected both by copyright laws and international copyright treaties, and also by other laws and agreements concerning intellectual property.
b) Ownership and copyright of the software product (including but not restricted to images, photographs, animations, video, audio, text and applets which are included in the software product), the accompanying printed material and each copy of the software product are the property of Ridom GmbH. All rights and intellectual property rights in and to content which can be accessed with the aid of this software product are the property of the respective owners of the content, and may be protected by applicable intellectual property rights laws and other laws and agreements on intellectual property. This EULC does not grant the licensee the right to use such content.

III. Scope of Licence
a) The licensee is entitled to install the software product, the demo-version or an earlier version on several computers, workstations, portable PCs or another digital electronic device.
b) The software product is licensed as a unitary product. The licensee is not entitled to separate its components for use.
c) The licensee is not entitled to reverse develop, to decompile or to disassemble the software product. No changes to the existing source code, neither any re-use of parts of it are permitted.
d) The licensee is not entitled to resell the software product or in any other way to transfer the software product for a counter-value. The licensee is not entitled to rent the software product, to lease it or to sell it.
e) Without prejudice to its other rights, Ridom GmbH is entitled to terminate this EULC insofar as the licensee contravenes the provisions of this EULC. In such a case, the licensee is obliged to destroy all copies of the software.
f) Ridom GmbH reserves all rights not expressly granted. The individual provisions of this licence agreement can only be changed by Ridom GmbH. The written form is required for this, without any exception.

IV. Warranty and assurances
a) This software is provided 'as-is', without any express or implied warranty. In no event will the Ridom GmbH be held liable for any damages (e.g. hardware damages, loss of data) or miscalculations arising from the use of this software.
b) The software product is supplied in accordance with the current development status. Ridom GmbH draws attention to the fact that is not possible in the current state of technology to create computer software such that it works

without errors in all applications and combinations. The subject of this contract is therefore only software which is in principle usable in the sense of the program descriptions and user manual.

c) Statements contained in brochures, advertisements and similar documents represent only descriptions, and do not contain any assurance of properties or guarantees.

V. Limitation of liability

a) The licensee bears the responsibility for the selection and for the consequences of use of the software product, together with the results thereby intended or achieved.

b) Liability for damages of any kind (included without restriction are damages from loss of profit, loss of business information or from any other financial loss) which arises as a result of the use of this product or of the inability to use this product is expressly excluded.

c) For damage or loss which involve death or physical injury, Ridom GmbH is in no case liable for aggravated compensation for damages or penalty compensation for damages, for any specific indirect, coincidental damage or loss or consequential loss or damage arising the refrom (this applies particularly, but not exclusively, for loss of use, loss of data, loss of profit and for loss of savings and business). This applies regardless of the basis of the claim or grievance.

d) Ridom GmbH is not responsible for and makes no warranty for the product running without interruption or without error.

VI. Place of judgement and applicable law

a) Place of fulfilment for all obligations arising from the contract relationship is, for both contract partners, Würzburg, the base of Ridom GmbH. This applies even if the residence of the customer is unknown or is outside Germany.

VII. Final provisions

a) Should any provision of this EULC be or become ineffective or contestable for any reason, its content is not otherwise disturbed thereby. Rather, it is to be performed in accordance with the intent.

b) Instead of the ineffective or contestable provision, an appropriate regulation is to apply which, as far as legally permissible, comes as close as possible to that which the parties who concluded the contract intended or would have intended in accordance with the sense and purpose of this contract if they had considered the point when concluding this contract. The same applies in the event of a lacuna in this contract. This also applies even if the ineffectiveness of a provision rests in any way on a measure of the performance and the time (period or date) prescribed in the contract. A legally permissible measure of the performance or time (date or period) which comes as close as possible to that which was intended shall then be deemed to be agreed.

Ridom EpiCompare, license version 1.1 (July 2005)