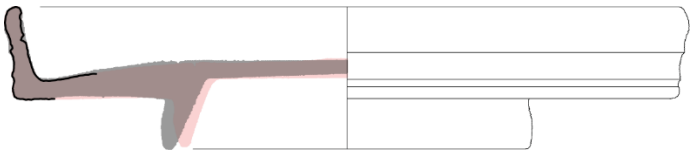
R A C O R D

Computer-Assisted Shape Classification of Archaeological Pottery Fragments



RACORD

Computer-Assisted Shape Classification of Archaeological Pottery Fragments

Credits:
Josef Wilczek, KA UHK Hradec Králové / ARTEHiS Dijon (josef.wilczek@hotmail.com)
Fabrice Monna, ARTEHiS Dijon (Fabrice.Monna@u-bourgogne.fr)
Nicolas Navarro, EPHE Dijon (nicolas.navarro@u-bourgogne.fr)
Carmela Chateau Smith, UFR-SVTE Dijon (carmela.chateau@laposte.net)

Version: 2021/01/29

Installation

1. Support program installation

Download and install following applications, depending on your operational system, from official developer web sites:

- 'R program' (https://www.r-project.org/)
- 'RStudio Desktop' (https://www.rstudio.com/)

2. Application installation

Download and unpack the RACORD.zip file. Copy the unpacked RACORD folder on your computer (for example into C:/RACORD).

3. Installing packages

Run 'RStudio' program and open server.R file (Fig. 1). RACORD uses several libraries for computation and visualisation. In further steps, all necessary libraries must be installed on your computer. Internet connection is required. Two library installation options are possible:

a) Simple click on the 'Install' link (Fig. 1)

b) Copy the following code lines to the R console and press enter (Fig. 2).

```
install.packages(c("DT", "kmlShape", "MASS", "shiny", "sp"))
```

4. Preparing and running the RACORD application

The path to the application must be defined in its source code. Specify the path to the application folder by replacing the code at the beginning of the server.R file by (Fig. 3):

```
WD <- c("C:/RACORD")
```

Note that instead of C:/DACORD in the code, you can specify the path to the folder containing the application.

You can also specify the folder containing '.txt' files including profile coordinates (Fig. 3).

```
dir_profiles <<- c("C:/RACORD/profiles/")
```

The RACORD application is run by clicking on the Run App icon in the RStudio panel (Fig. 4; Fig. 5). The graphical layout of the application can be adjusted at the beginning of the ui.R file, depending on the screen resolution (Fig. 6).

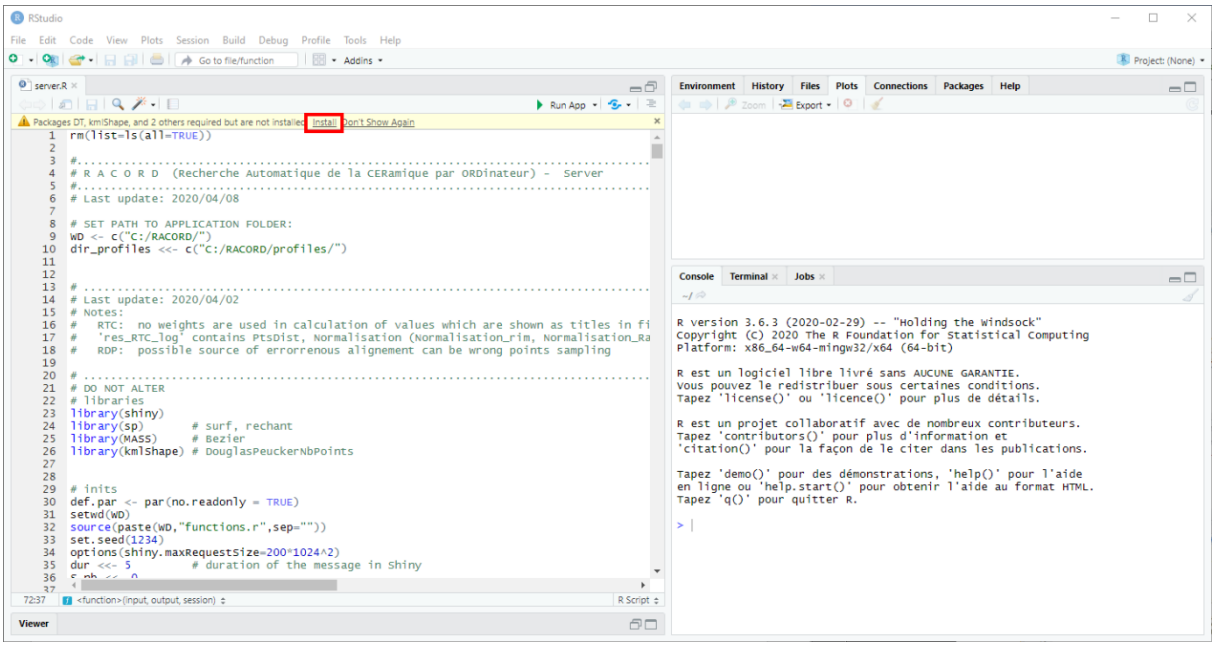


Fig. 1. Packages installation.

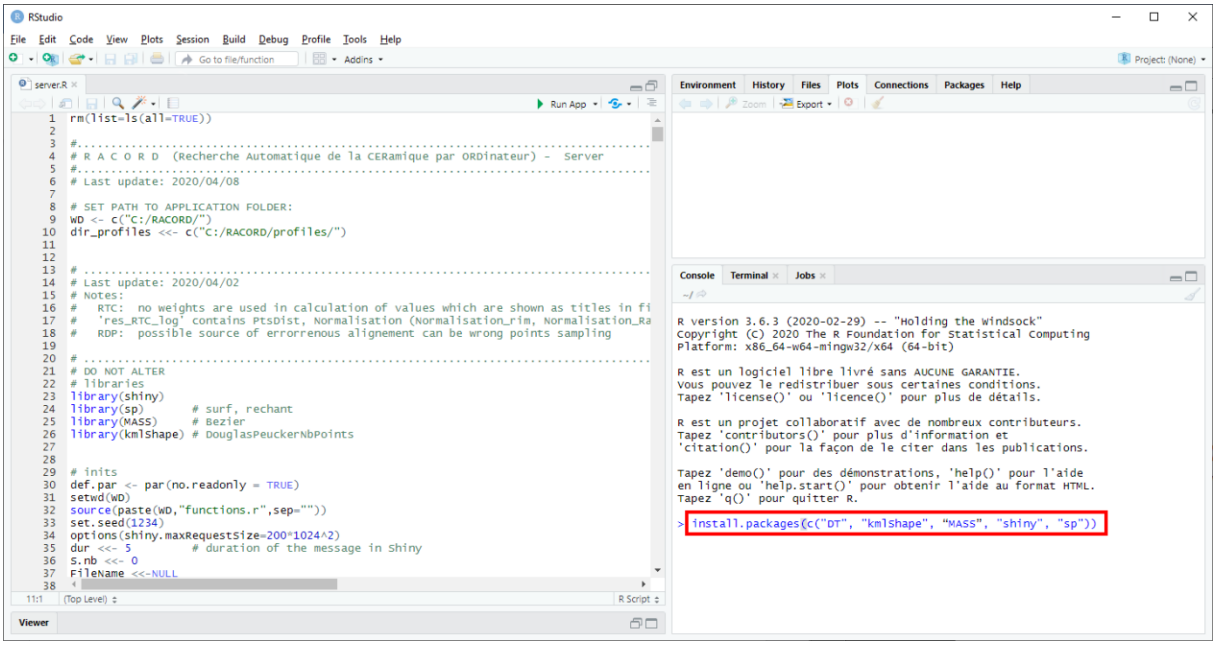


Fig. 2. Packages installation.

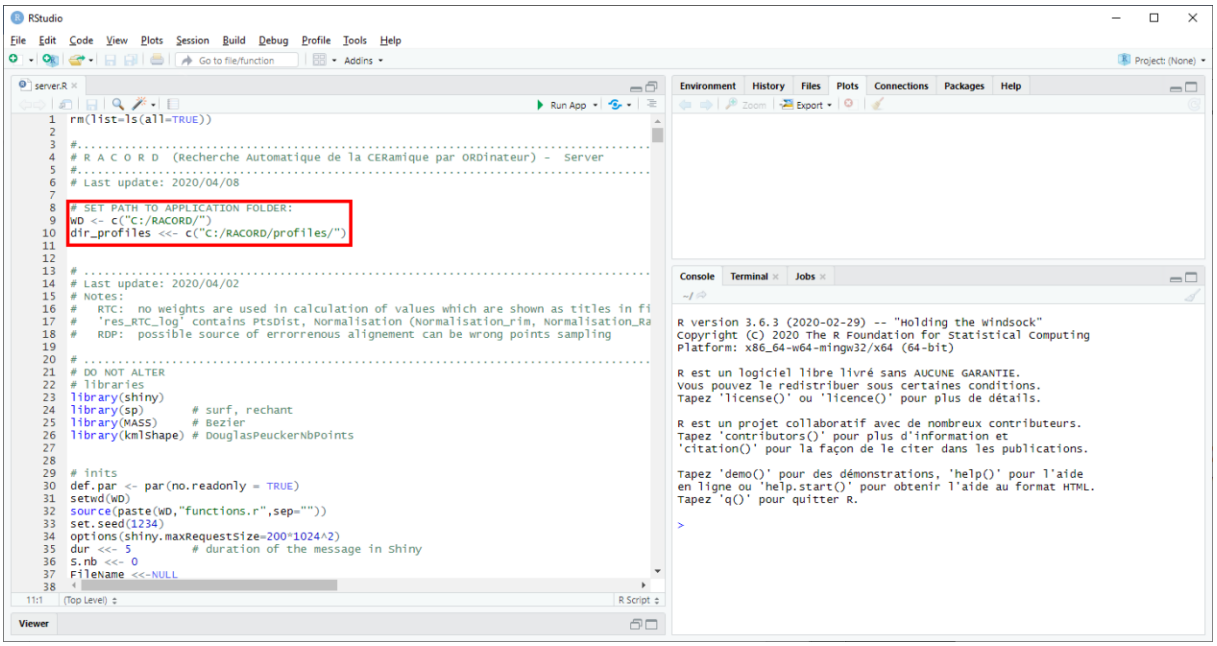


Fig. 3. Setting path to the source code and to the folder containing '.txt' files with profile coordinates.

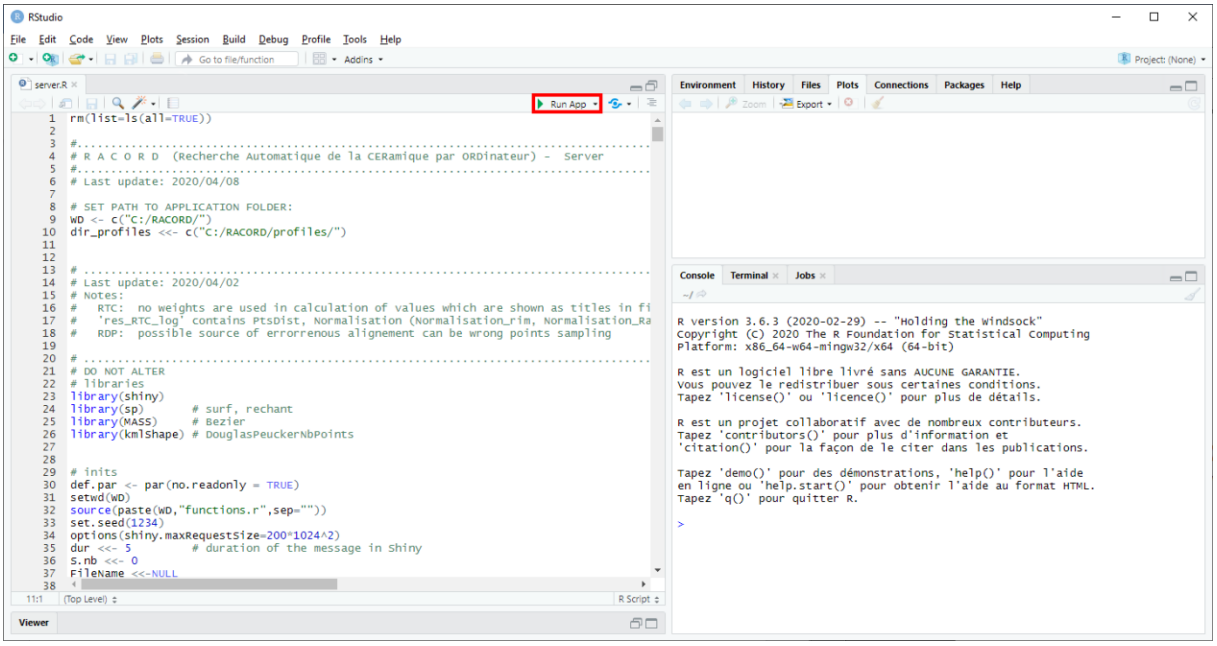


Fig. 4. Running the application.

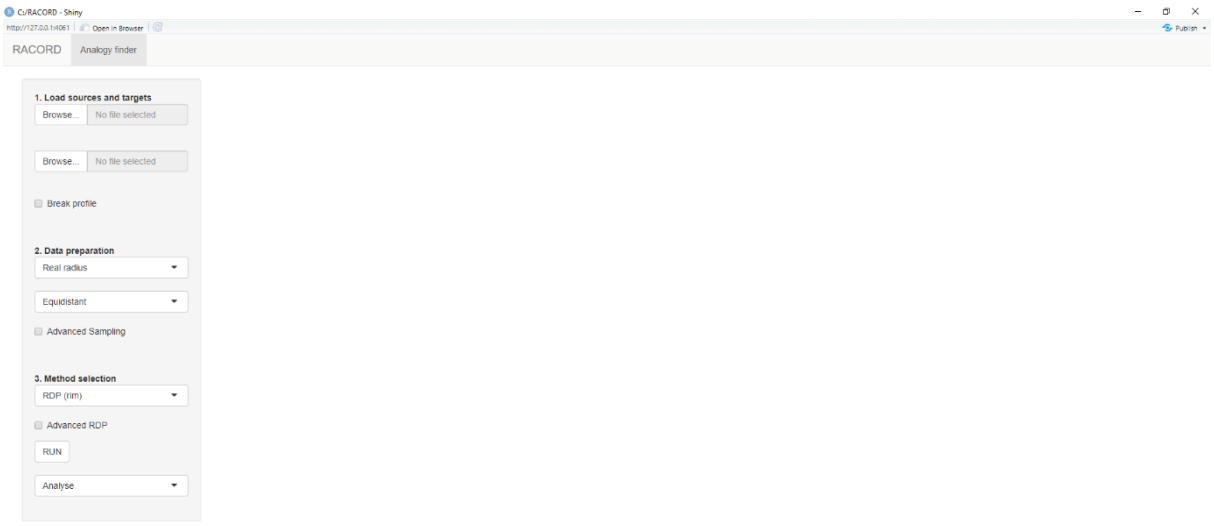


Fig. 5. Running the application.

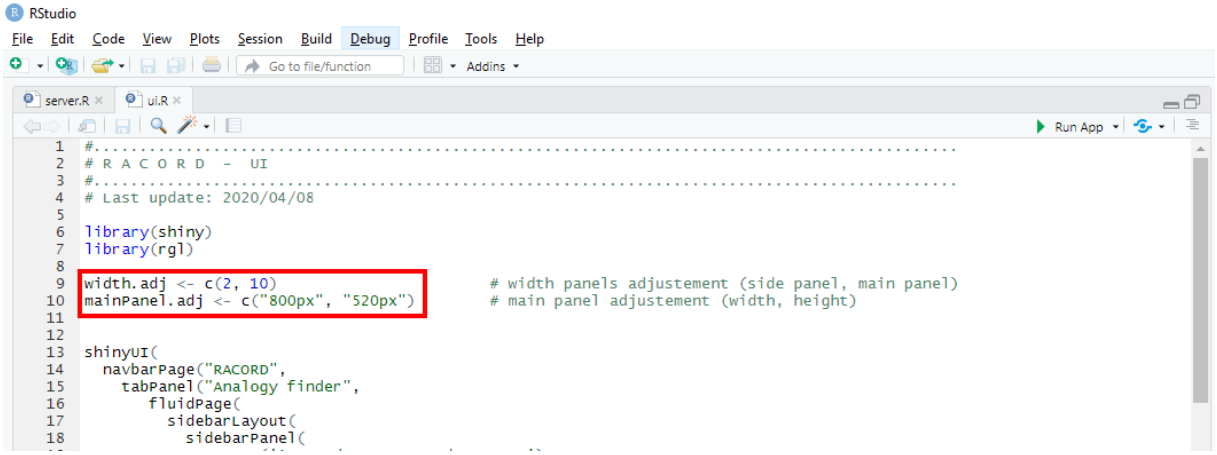


Fig. 6. Adjusting graphical layout of the application.

Manual

The goal of the application is the retrieval of the most similar analogy to a given **source** fragments from a dataset containing **reference** complete vessel profiles. The procedure consists of three steps: 1) Load sources and targets, 2) Prepare data, 3) Select a method and view results. The application therefore contains three sections dedicated to the corresponding tasks (Fig. 5). The Grafeusenque dataset which was used in the study for method testing is provided as examples.

1. Load sources and targets

Two Browse buttons serve to load files containing information about sources and targets stored in the form of tables. Each table has a header (Fig. 7:A). Table rows of the table correspond to fragments to be classified (for sources) and to complete vessels (for targets), table columns correspond to three descriptors (ID, diameter, type) separated by a semicolon symbol (Fig. 7:B):

- ID: fragment/vessel identification.
- Diameter: rim diameter of the fragment/vessel (in cm); comma is used for decimals
- Type: type of the fragment/vessel.

Note that the ID of each fragment/vessel must be the same as 'txt' file containing its profile coordinates (Fig. 7:C). This 'txt' file needs to be stored in the 'profile' folder which was specified earlier (Fig. 7:D; see also Figure 3 for setting the path to this folder).

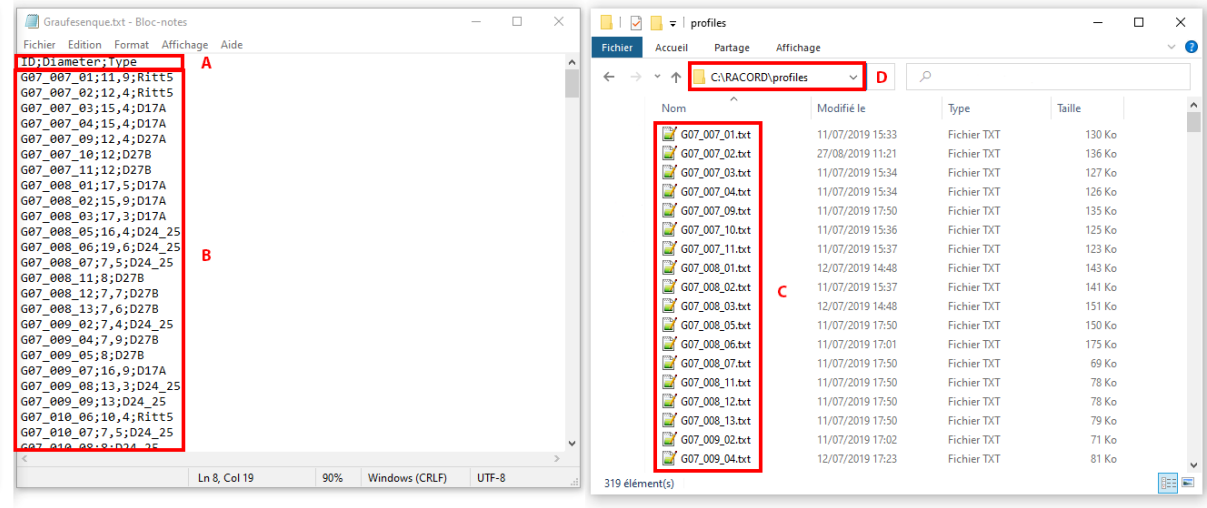


Fig. 7. Structure of source/target file. Note that the ID of vessels must refer to txt files containing corresponding profile coordinates.

Once loaded, both tables appear on the screen. Sources are on left, targets on right. Items selected from these tables by left click mouse can be visualised together on the right part of the screen (Fig. 8).

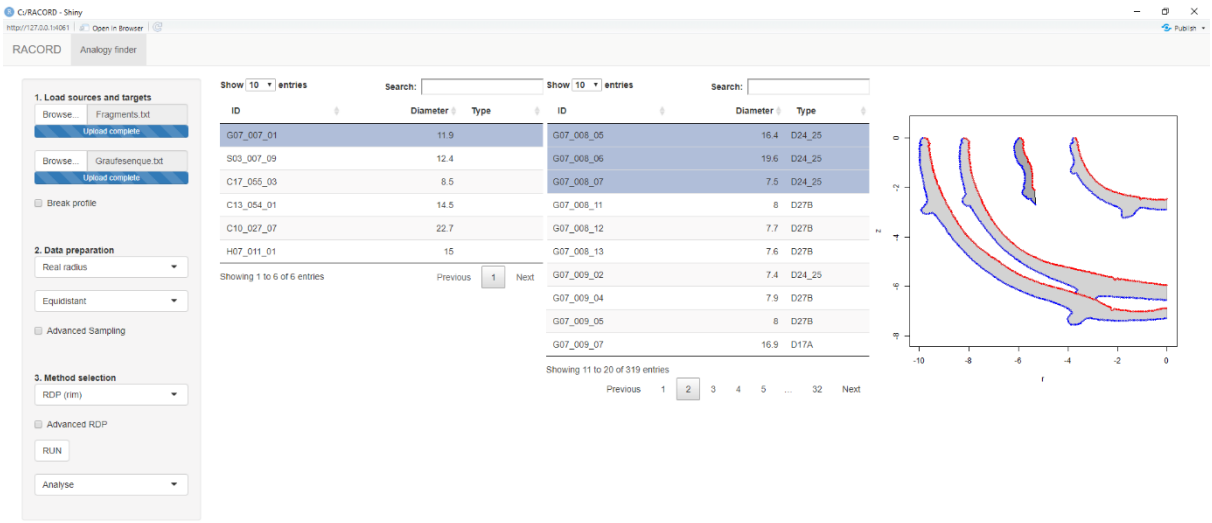


Fig. 8. One selected source fragment (G07_007_01; dark grey polygon) is plotted with three selected target complete vessels (G07_008_05, G07_008_06, and G07_008_07; light grey polygon).

Break profile allows to generate source fragments by virtually breaking (i.e. segmenting) complete vessels. Fragments are generating by selection of a complete source vessel (Fig. 9:A), checking Break profile option (Fig. 9:B), and clicking on Break profile button (Fig. 9:C). Newly created fragments then appear in the sources table (Fig. 9:D).

Four parameters can be additionally specified:

- Number of fragments specify number of generated fragments.
- Minimal % size of one fragment sets a minimal size of one fragment.
- Set seed allows to specify pseudorandom number used in fragment generation.
- Extract only rim allows the generation only of one rim fragment.

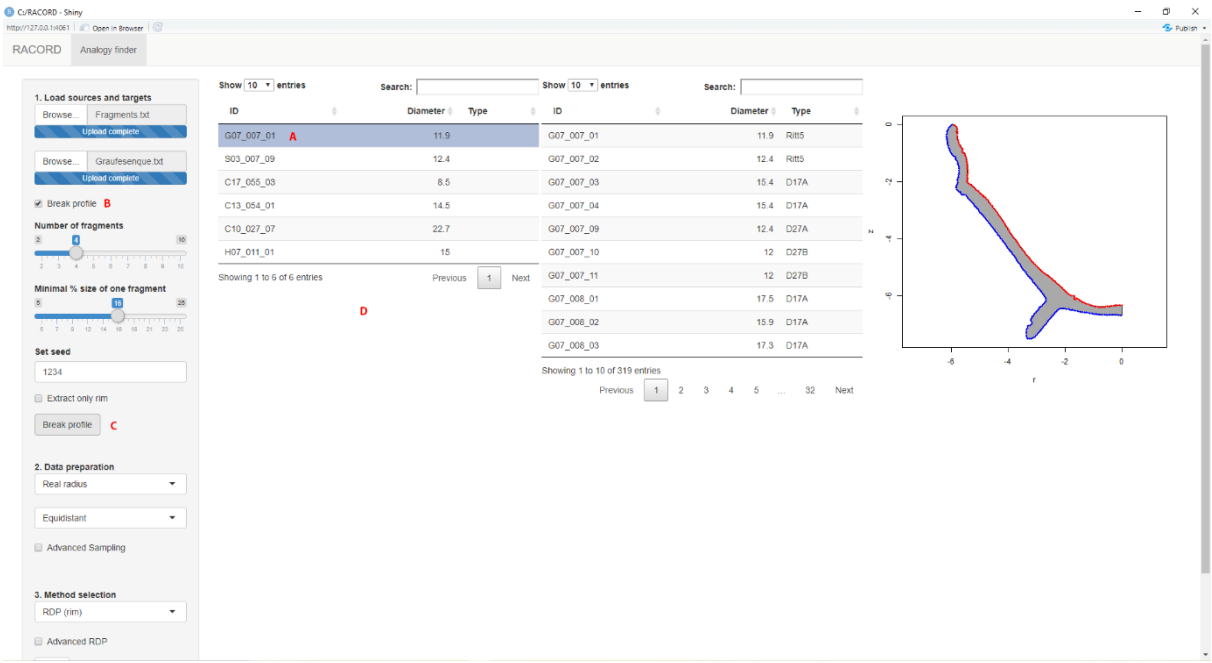


Fig. 9. Generating source fragments by virtually breaking (i.e. segmenting) complete vessel.

2. Data preparation

Data preparation allows to normalise the size of profiles and to choose the strategy used for sampling points on profile outlines.

Size normalisation

- Real radius allows to use real metrics (i.e. cm) in matching.
- Rim units allows to standardise outlines to the same rim unit. If selected, only the shape (and not size) of profiles is used in profile comparison.

Sampling specification

- Segments method split a profile onto a specified number of segments. This method should be used only for ICP algorithm. Number of segments can be specified by checking Advanced sampling option.
- Equidistant points method generates equally spaced points along the outline, starting from the most upper point. It assumes the homology of segments sampled along the curvilinear abscissa. Distance between points can be specified by checking Advanced sampling option.

3. Method selection and Viewing results

Once the data prepared, five different methods and their variants can be used for fragment matching:

(1) Bezier (rim) – Bezier Polynomials

(2) DCT (rim) – Discrete Cosine Transform

(3) RDP (rim) – Ramer-Douglas-Peucker algorithm

(4) RTC (rim) – Radius, Tangent and Curvature

(5) ICP (rim) – Iterative Closes Point. No translation, size and rotation allowed. Function thus calculates solely Euclidean distances between profiles.

(6) ICPz (all) – Iterative Closes Point. Only translation along the rotational axis is allowed.

(7) ICP (all) – Iterative Closes Point – translation, size and rotation are allowed.

(8) All methods (all) – All rim methods (1, 2, 3, 4, 5) executed simultaneously.

The choice of the most suitable method depends on the fragment part (rim *vs* no matter which part), and the feature the user chooses to focus on.

The calculation parameters used in matching can be fine-tuned in Advanced options (numbers in parentheses correspond to matching methods):

Harmonics number (2) indicates the number of harmonics obtained with Discrete Cosine Transform (20, by default).

Number of points (3) indicates number of points generated by Ramer-Douglas-Peucker algorithm (20, by default).

Function weights (Rs, Ts, Ks) (4) indicate weights of Radius, Tangent and Curvature functions as proposed by Karasik and Smilansky (2011).

Translation parameter (z) (6) indicates the translation step along the rotational axis used in searching.

Optimisation parameters (r, theta, size) (7) indicates the limits of r-translation (in cm or radius units), rotation (in degrees), and size (magnitude) parameters used in optimisation.

SAN parameters (MaxIter, InitTemp, MaxTemp) (7) indicate the maximum number of iterations, the initial temperature, and the number of evaluations at each temperature used in Simulated Annealing optimisation calculation (2000, 100, and 40, by default).

The matching process is started by selecting sources fragments to be searched and by pressing the Run button.

Once finished, the source table changes its appearance (Fig. 10). Columns of the table now indicates which fragment(s) was/were analysed, and by which method.

The best match or matches is/are then visualised by selecting analysed source fragment (Fig. 11:A), switching Analyse selection to Results (Fig. 11:B), sorting target complete vessels increasingly according to the RMSD value (Fig. 11:C), and selecting target vessel or vessels for which the RMSD value is/are minimal (Fig. 11:D).

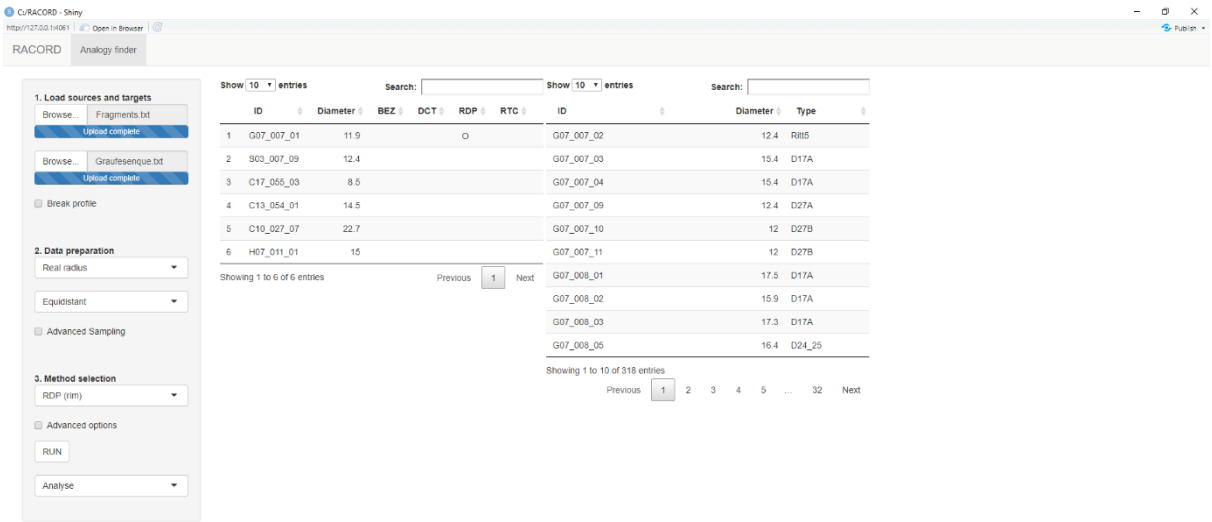


Fig. 10. Example of the screen with finished matching. The source fragment 'G07_007_01' was analysed by 'RDP' method.

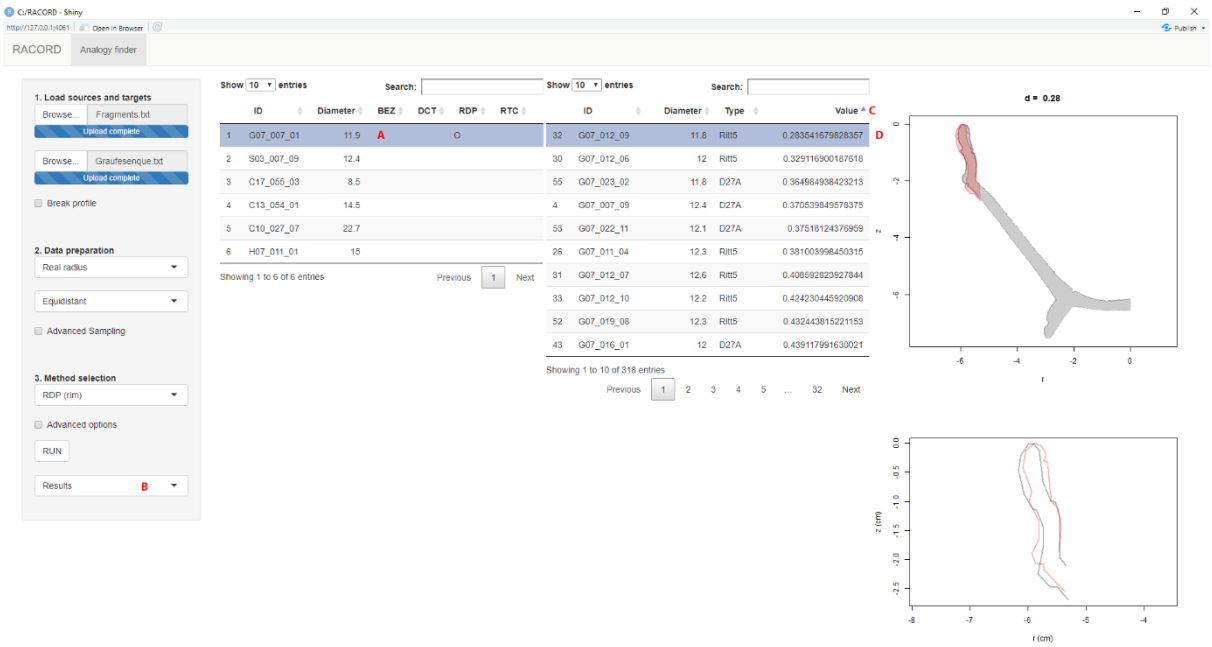


Fig. 11. Example of viewing the best match for the 'G07_007_01' fragment analysed by 'RDP' method.