

## **Appendix S3 - Methods of Data collection and analysis**

We used a short questionnaire and a structured questionnaire to collect information from experts. These documents have consisted of a series of questions, which aim to collect relevant data on riparian genetic resources in Europe. These contained questions regarding data on current status, researchers' perception, general practice and strengths and weaknesses of the projects.

The short questionnaire consisted of 10 questions regarding relevant contact on riparian genetic conservation, knowledge of riparian genetics in the participant's country and relevant titles related to genetic conservation of riparian species in the participant's country.

The structured questionnaire consisted of 22 questions regarding the past experiences of the researchers, where the focus is on the questions regarding genetic conservation and the researchers' experience and perception of the subject. We were able to carefully read, analyze and interpret the information regarding genetic conservation practices in each country.

After analyzing responses provided by researchers, we separated the most common strengths and weaknesses in practice which are extracted from specific questions in the structured questionnaire.

There were 30 documents with 4 answers per each document, which equals 120 answers. For each answer, there were 2 coded lines which give in total 240 lines of code.

First we started the analysis with separating numerical and text information. The numerical analysis included the frequency of all the questions that could be answered "yes" or "no".

We have carefully separated those by country and put them in the table format in Microsoft Excel.

Table for numerical analysis was done through 5 sheets. The first sheet contained all the information about the experts who filled in the structured questionnaire. The second sheet was filled with information regarding genetic conservation; the best approaches, the main benefits of genetic conservation, and possible changes regarding genetic conservation in that particular country. In the third sheet, we have put the status on riparian genetics in that

particular country. The fourth sheet was filled with information about the strengths of genetic conservation and on the final fifth sheet, there was information on weaknesses of riparian genetics. All the answers were marked as affirmative or negative. Then we were using frequencies, where we determined the percentile of practice repeating for a specific question. To visualize the results, Microsoft Excel was used, by making charts and graphs to show results. For this analysis software that was used was Microsoft Excel.

All the open-ended questions and answers were put in another Microsoft Excel table. That document contained a set of 5 sheets as well. The first sheet contains information regarding experts. The second sheet contains information about the specificities of riparian genetic resources. The third sheet is filled with explanations regarding the strengths of the riparian genetic conservations, while the fourth sheet is focused on weaknesses. The fifth sheet is text answers regarding ideal practice by the opinion of each expert.

For text analysis, we used NVivo 12 software. It is a tool for text analysis, which includes coding. This software helped us to analyze and have a better insight into responses from open-ended questions.

There were two stages of text analysis:

1. Modeler - helped us in building validate predictive models
2. Visualization designer – helped us to make representative visual presentations for writing the paper.

Before we started the first stage, we had to name each relevant category for each question. Each category has a subcategory, which was more specified by the open-ended answers. Naming those subcategories was based on the answers of each interviewee. Categories and subcategories were created by core researchers participating in these analyses, and they were based on the most used answers of the participants and in which area they can best fit.

All of the coding had to be done manually. Therefore, we had to do it by each category. First, we divided all the answers by country in a separate Microsoft Word document. Then, after making categories, we had to go through each structured interview and code each part of the answer. Each of those codes we have assigned to one of the categories to which it best belongs.

After all the answers were coded, we went through the Modeler stage, where we were building validate predictive models based on those codes and answers from the interview. During the process, we were going through all 232 lines that were coded. While going through those coded lines, we were making sure that all the answers were in the right category and subcategory and that all the models for the subcategories were relevant to this paper. In this case, models were corresponding to subcategories. For this part, we used cross-tabulation, where we determined the relationships and practices between many different results.

After stage one was over we have proceeded to stage two - to make representative visual presentations for the paper by visualizing the results through the process of making charts and graphs, which was done in Microsoft Excel. The results of this stage can be seen in the paper.