



HAL
open science

SequenceViewer: Visualization of Genes Sequences

Arnaud Sallaberry, Nicolas Pecheur, Sandra Bringay, Mathieu Roche,
Maguelonne Teisseire

► **To cite this version:**

Arnaud Sallaberry, Nicolas Pecheur, Sandra Bringay, Mathieu Roche, Maguelonne Teisseire. SequenceViewer: Visualization of Genes Sequences. MedInfo: International Congress on Medical Informatics, 2010, Cape Town, South Africa. 2010, Demonstration (long abstract). lirmm-00588556

HAL Id: lirmm-00588556

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00588556v1>

Submitted on 26 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SequencesViewer Visualization of gene sequences

Arnaud Sallaberry^a, Nicolas Pecheur^b, Sandra Bringay^{b,c}, Mathieu Roche^b, Maguelonne Teisseire^d

^a LaBRI Laboratory of research of informatic, INRIA Bordeaux Sud-ouest, Pikko, France

^b LIRMM, Laboratory of Informatics, Robotics, and Microelectronics, University of Montpellier 2, Montpellier, France

^c Mathematic and Informatics Department, University of Montpellier 3, Montpellier, France

^d CEMAGREF - Earth Observation and Geoinformation for Environment and Land Management research Unit, Montpellier, France

INTRODUCTION

Techniques for extracting knowledge from huge volumes of biological data, obtained from DNA microarrays analysis, allow the **discovery of previously unknown knowledge**.

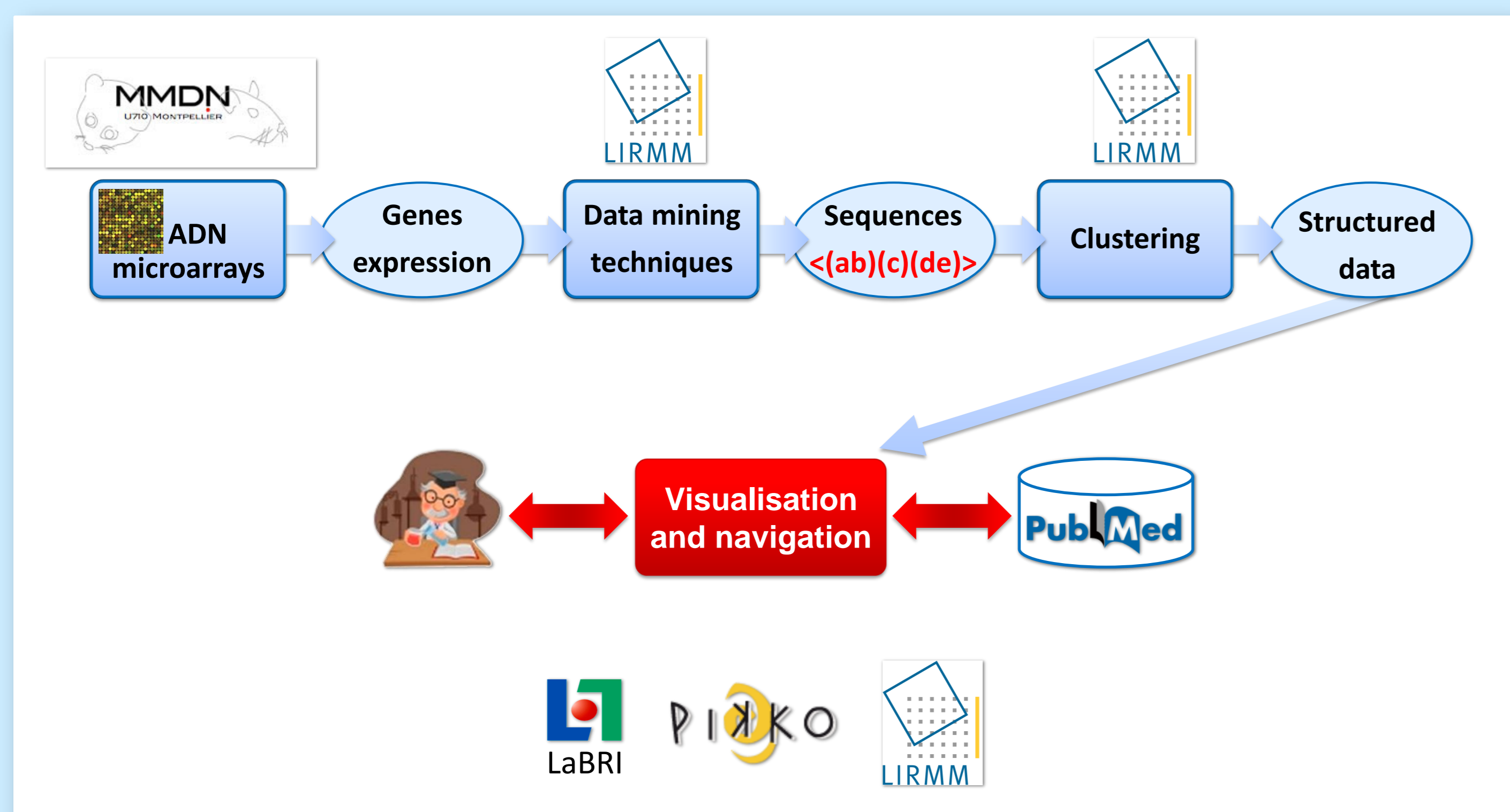
However, these techniques generally produce **many results not easily actionable by the experts**.

We propose a tool dedicated to the support of these experts in the process of **appropriation and exploitation of the knowledge** obtained after the extraction process.

This tool is based on **3 visualization techniques** (Clouds, Solar systems and Treemap) that allow biologists to capture a large amount of patterns (ordered sequences of genes)..

Keywords

Visualization, Gene sequences, Data mining

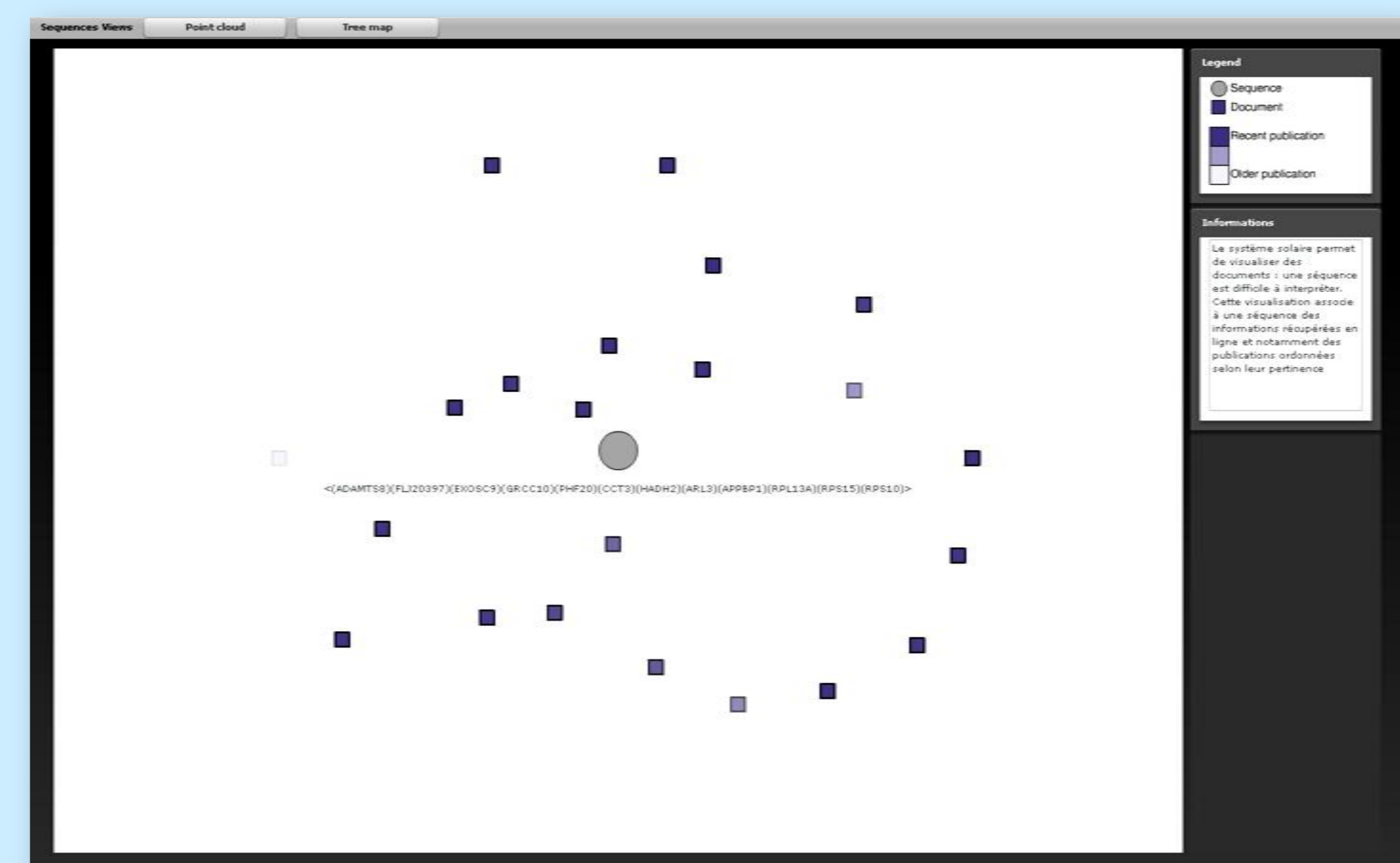


Clouds representation

The clouds allow to visualize **groups of sequences**. Each group is identified by a red center.

The sequences are shaded. The centers are placed according to their distance the one from the others. The sequences in a group are placed according to their distance from the center.

The user can move, zoom, display information about centers or sequences. He can also filter the sequences according to a gene which appears in green.



Solar system representation

The solar system allows to view a sequence with all the **associated documents**. The sequence is in the center.

The documents are represented by squares and are positioned around the sequence according to their proximity. The color of the squares expresses the year of publication.

The user can click on the sequence or on the documents to display information or access to the document.

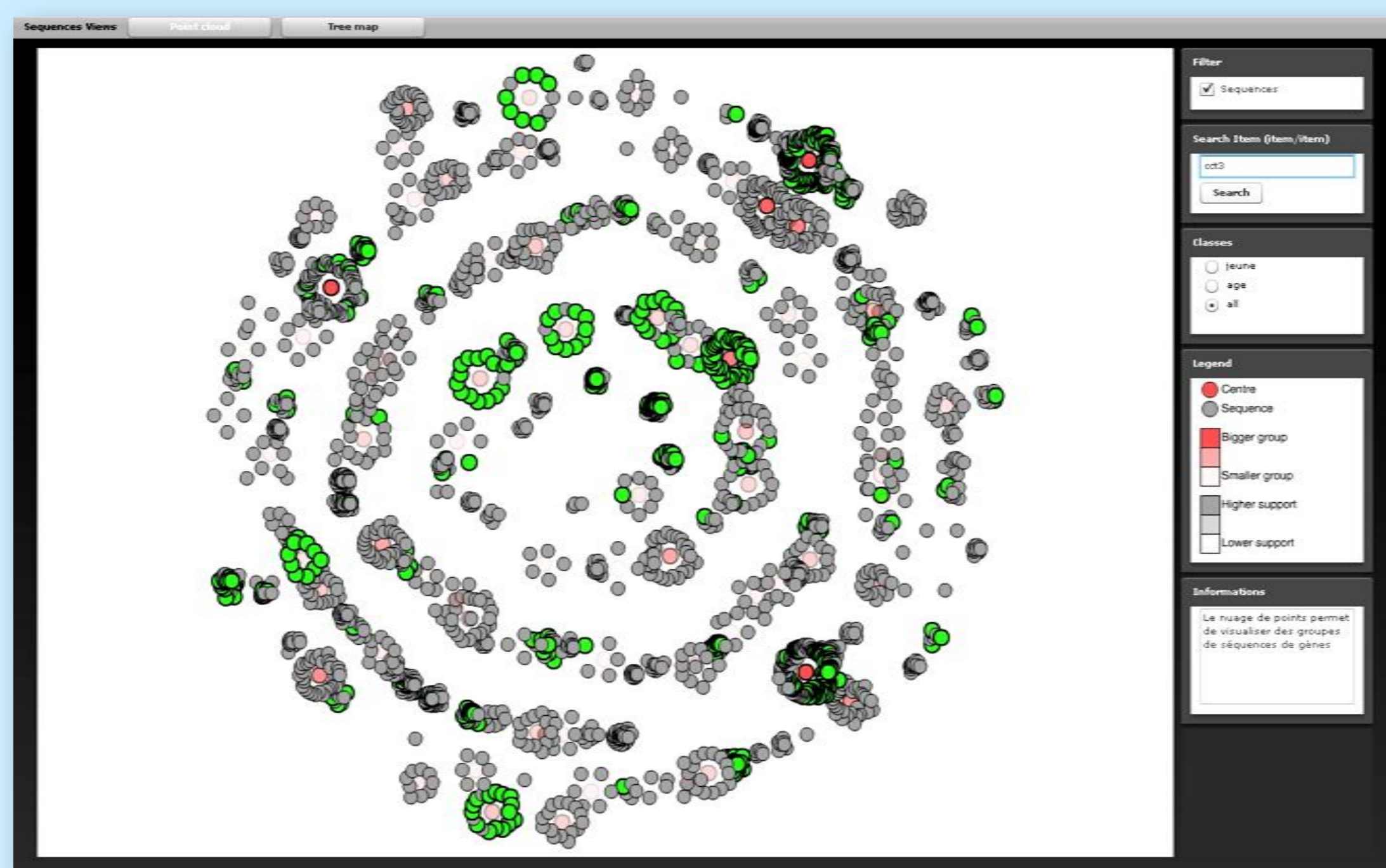


Treemap representation

The Treemap shows the **sequence hierarchy**.

At the beginning, the two first hierarchy levels are displayed according to their class proportion. On click, the user drills down the next level.

The path in the tree is displayed above the Treemap, and helps to return to higher levels.



Collaboration with PIKKO Compagny

SequencesViewer has been developed by the **PIKKO compagny**.

It is a web application designed for biologists, enabling them to visualize a large amount of gene sequences.

