Category: **Science & Research**

Project: **TranScape VR**

What was the challenge?

In bioinformatics, researcher perform RNA sequencing to understand which gene is specific to a certain tissue (e.g. in a plant). These multi-dimensional data sets cover more than thousands of genes over many different tissues and conditions. Getting an overview of this massive amount of data is important but challenging and standard data analysis methods using 2D, non-interactive visualization are limiting and daunting.

What was the solution?

To upgrade to state-of-the-art technology, we developed TranScape VR, a new tool for visualizing the gene landscapes in an immersive, explorative, holistic and entertaining Virtual Reality (VR) space.
In search for special plant carnivory-related genes, we represent the data as an interactive network, mapping genes to each tissue core-node, forming gene clusters. Biological parameters (e.g, gene expression level, tissue specificity) are mapped to visual parameters such as colour, size, distance, and gene position within the 3D space. The visualization offers different modes which depict different properties of the data set. All modes can be interactively explored in order to form and test hypothesis in real-time.

What was the effect?

The VR visualization combined several different analytical plots in one representation, which allows researchers to quickly test their hypotheses in VR. The app was presented at the international conference ISMB ECCB 2021 where it was well received.

Contact:

name: Annika Kreikenbohm
company/organisation: Immersive Design
e-mail: hello@annok.de
website: https://annok.de/transcapevr/
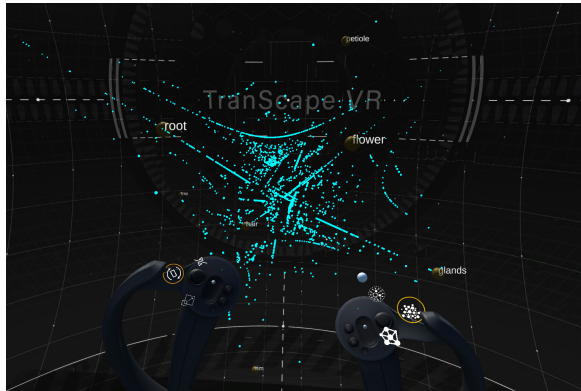
The gene data cloud: Tissues are depicted as yellow spheres (e.g., flower or root), where the size is defined by the number of genes associated to them. The larger the sphere, the more genes are hosted by that tissue. Genes are shown as a blue point cloud. They are mapped by specificity, i.e. genes that are closer to a tissue node are more specific to that certain tissue. By clicking in a tissue, colored lines show which genes were associated to the tissue by a computer algorithm
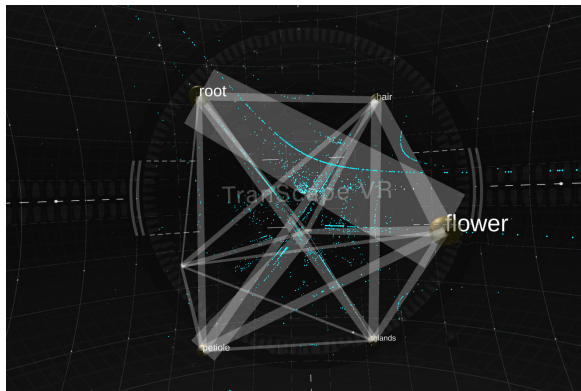
The tissue network: the links between two tissues represent the number of shared genes. The more genes are shared by a pair of tissues, the thicker the lines. We immediately recognize the thick line between root and flower as well as between petiole and rim. Such observation is typically analysed in standard Principle Component Analysis as well as Pearson correlation analysis.
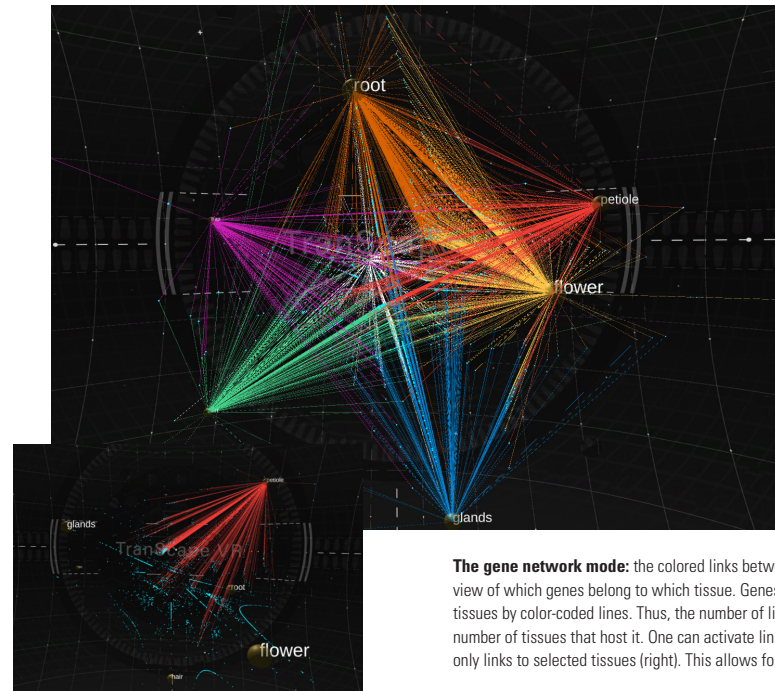
The gene network mode: the colored links between genes and tissues provide a clear view of which genes belong to which tissue. Genes are connected to each of their host tissues by color-coded lines. Thus, the number of links connected to a gene represent the number of tissues that host it. One can activate links to all tissues at once (left) or activate only links to selected tissues (right). This allows for investigation of shared genes between tissues.

Visual Exploration: The user can visually explore and analyse the multidimensional dataset by re-scaling, interacting, retrieving genes' descriptions as well as filtering or highlighting genes based on keywords.

_ Scale, rotate and move the gene cluster for different (over) views

_ Rearrange tissue nodes to investigate different mapping constellations in real-time (e.g., panel D)

_ Activate tissue links to find associated genes and activate gene details for quantitative data and description (e.g., panels A, B)

_ Filter parameters (e.g., expression level, specificity level) to define which genes are displayed and search for (groups of) genes using keywords (e.g, panel C)