

Computer-aided analysis of protein-protein-interactions

Rainer Stotzka¹, Jorge Silva¹, Peter Uetz², Jens Jäkel³

- 1 Forschungszentrum Karlsruhe, Institute for Data Processing and Electronics, PO 3640, 76021 Karlsruhe, Germany, phone: +49 7247 82-4738, email: rainer@stotzka.de
- 2 Forschungszentrum Karlsruhe, Institute of Toxicology and Genetics
- 3 Forschungszentrum Karlsruhe, Institute for Applied Computer Science

The two-hybrid system is an experimental procedure to analyze protein-protein interactions [1,2]. At present each yeast interaction screen has to be analyzed by an "expert" individually.

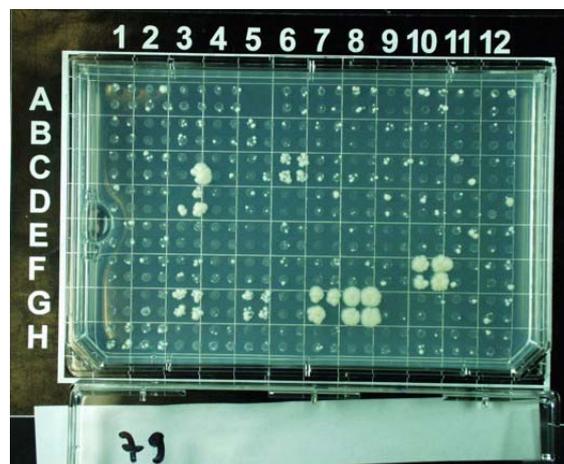


Figure1: Two-hybrid interaction screen. This array consists of 96 positions with 4 yeast colonies each. The 96 positions correspond to 96 different proteins from Human Herpes virus 8 (KSHV) that are expressed in these yeast colonies. In other words: in this particular screen 1 KSHV protein is tested for interactions with 96 different KSHV ("79") proteins. If two proteins interact this interaction allows the cell to grow. The resulting yeast colonies can be recognized as white spots. The size of the spots gives a semi-quantitative measure of interaction strength, i.e. protein 79 interacts strongly with the proteins expressed in G8, F10, and G7 but only weakly with the proteins expressed in C6, G3 or G5.

At the Forschungszentrum Karlsruhe we are developing a computer-aided analysis tool to analyze such two-hybrid screens automatically. The interaction screens are imaged by a digital camera connected to the client's PC in the laboratory. Using a WWW form the image is sent to a Java Servlet [3,4] running on a Linux server (figure 2).

The servlet controls the image processing and pattern recognition process consisting of different steps:

White balance: During imaging the two-hybrid screens the quality of the images depends on the illumination and on the camera settings. An automatic preprocessing and white balance algorithm implemented in Khoros [5] maps the image to a standardized color space.

Template registration: Due to different handling of the camera by the photographer the two-hybrid screens can be shifted, rotated and scaled within the image. For the automatic mapping of the growing spots to the location a standardized position of the screens is needed. A rigid registration process is implemented in ITK [6] which maps the image to a standardized grid template.

Segmentation: The image is now converted to a binary image by simple threshold operations. Due to the standardized color space the thresholds can be determined easily. To get rid of noise and white lines erosion and dilation algorithms are applied successively. Now a white pixel represents yeast and a black one the background. The segmentation is implemented in Khoros.

Classification: According to the number of white pixels at the certain grid locations the growth pattern of the yeast is classified as “_none_”, “_weak_” and “_strong_”.

Formatting result: The intermediate results of the processing steps described above are documented in XML, translated into HTML and presented to the operator for inspection. After validation the resulting growth pattern can be annotated and stored in a database.

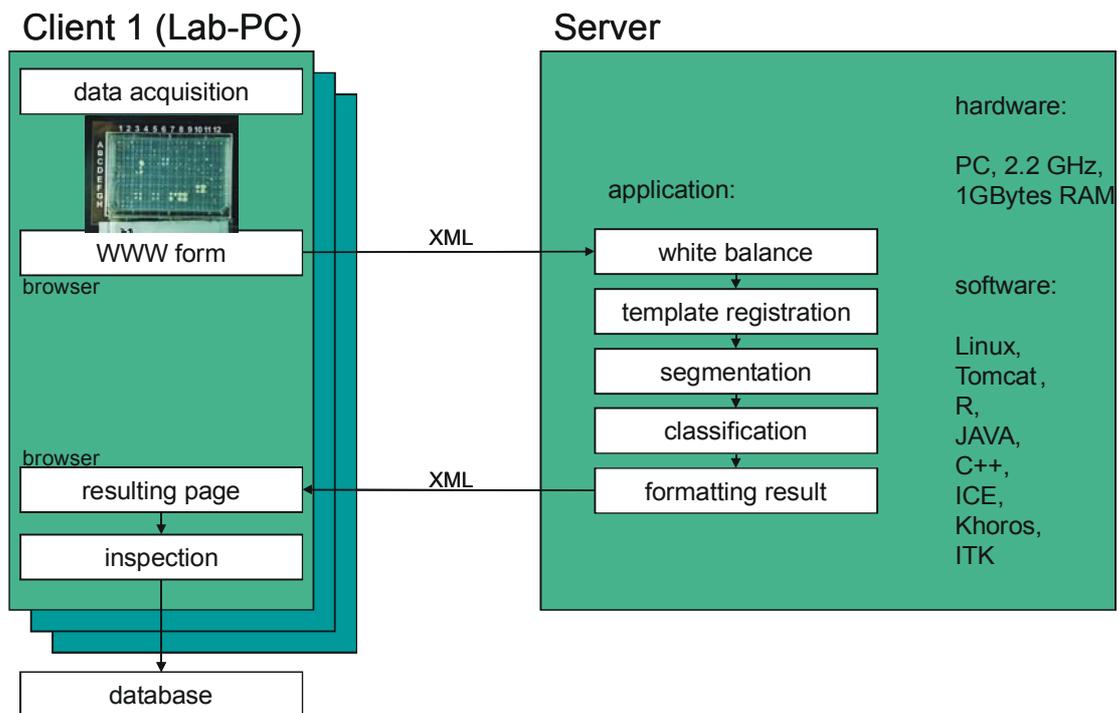


Figure 2: Software structure of the analysis tool. The two-hybrid screens are imaged with a digital camera and recorded on the Lab-PC. For analysis the images are loaded into a WWW form and send to a central server. On the server a Java Servlet ‘application’ controls the pattern recognition process. The results are formatted in XML and send back to the Lab-PC.

The client-server structure allows to upgrade the software and the server hardware without affecting the biologist’s Lab-PC. Furthermore the results of the different processing steps can be monitored permanently for further enhancement of the algorithms.

We expect our first results in September 2003 and we will be able to measure the quality of the pattern recognition process. We plan to use the software routinely in the laboratory by early 2004.

KEYWORDS:

protein-protein interactions, two-hybrid screens system, computer-aided analysis, image processing, registration, segmentation

[1] Uetz P., Vollert C.S.: Protein-Protein Interactions. Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine, Springer Verlag (in press), 2003.
 [2] Vollert C.S., Uetz P.: The Two-hybrid System. Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine, Springer Verlag (in press), 2003.
 [3] Jakarta Tomcat Documentation, <http://jakarta.apache.org/tomcat/tomcat-3.2-doc/index.html>
 [4] Java Technology and Web Services, <http://java.sun.com/webservices/>, 2003
 [5] Khoros Research Inc., <http://www.khoros.com>
 [6] NLM Insight: Segmentation and registration toolkit, <http://www.itk.org/index.htm>