

**Retrieving Treasures From the Data Jungle:
DKFZ and IBM have signed a collaboration agreement on cancer genome analysis**

The International Cancer Genome Project analyzes the complete genomes of thousands of cancer patients. This generates enormous amounts of data. To find those gene segments which are pivotal for cancer development and treatment, intelligent IT systems are needed. The German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ) in Heidelberg and IBM have signed a strategic partnering agreement today at CeBIT 2011. The aim of the agreement is to turn sequencing data into useful information for cancer medicine.

“In the coming years, sequencing of cancer genomes will produce considerable amounts of data. This will fundamentally improve diagnosis and treatment of cancer patients,” says Professor Otmar D. Wiestler, DKFZ’s Scientific Director. “However, in order to actually use the findings from the enormous flood of data, we need intelligent information technology. It will help us to recognize and ultimately use crucial segments. IBM is an ideal partner for us in this great challenge.”

Cancer is a disease of the genes. In every cancer cell, many genes are altered compared to healthy cells. However, by no means are always the same genes mutated – they differ from one type of cancer to the other and even from one patient to the other. In many cases it is still unknown which mutations are really pivotal for the process of cancer and which are more likely to be random. To find this out is the aim of the International Cancer Genome Project: In 50 types of cancer, the genomes of 500 cancer patients for each type will be deciphered letter by letter. Each genome generates 2.4 terabytes, which is equal to 24,000 gigabytes. For comparison: A standard laptop currently has a storage capacity of about 100 gigabytes.

DKFZ is involved in three projects of the International Cancer Genome Consortium (ICGC): It coordinates the PedBrain consortium, which analyzes the genome of childhood brain tumors, as well as the study of the genetic causes of prostate cancer in younger patients, and is partner in the German ICGC network for the analysis of malignant lymphomas.

The partnering agreement between DKFZ and IBM covers three aspects of handling the gigantic masses of data:

- 1) First, data will be compressed, similar to MP3 files in the music industry. To this end, new strategies tailored to genome data will be developed.
- 2) Second, the partners will search for solutions how to transfer the gigantic amounts of data from the storages to the computers and to compare them to identify, for example, frequently altered genes.
- 3) Last, the aim of genome analysis is to use results for tailored cancer treatments. To this end, methods will be developed for aligning genome data with clinical parameters such as disease progression or response to targeted drugs. Efficient data processing on the basis of Watson technology will ensure that cancer medicine physicians can use initially unstructured research data for treatment decisions.

All data of the German ICGC projects will be joined together by Professor Roland Eils, who is head of the Theoretical Bioinformatics Division of DKFZ. To this end, Eils is building one of the world’s largest data storage units for life sciences at the BioQuant Center of Heidelberg

University. The collaborators expect to operate data storages with a capacity of six petabytes for storing the genome sequences.

A picture is available on the Internet at:
www.dkfz.de/de/presse/pressemitteilungen/2011/images/cebit.jpg

Source: Deutsches Krebsforschungszentrum

The German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ), employing over 2,500 staff members, is the largest biomedical research institute in Germany. More than 1,000 scientists are working to investigate the mechanisms of cancer development, identify cancer risk factors and develop new strategies for better cancer prevention, more precise diagnosis and effective treatment of cancer patients. In addition, the staff of the Cancer Information Service (KID) provides information about this widespread disease for patients, their families, and the general public. DKFZ is funded by the German Federal Ministry of Education and Research (90%) and the State of Baden-Wuerttemberg (10%) and is a member of the Helmholtz Association of National Research Centers.

Dr. Stefanie Seltmann
Head of Press and Public Relations
German Cancer Research Center
Im Neuenheimer Feld 280
D-69120 Heidelberg
T: +49 6221 42 2854
F: +49 6221 42 2968
presse@dkfz.de