SMP: DNA

Project: National Genotyping Platform (NGP)

Peter Nürnberg – Cologne Center for Genomics , University of Cologne, Köln - nuernberg@uni-koeln.de
Norbert Hübner – Max Delbrück Centre for Molecular Medicine (MDC), Berlin - nhuebner@mdc-berlin.de
Thomas Meitinger – GSF-Nat. Res. Center for Environment and Health, Neuherberg - meitinger@gsf.de
Jochen Hampe – Institut für Klinische Molekularbiologie, CAU, Kiel - j.hampe@mucosa.de
Sascha Sauer – Max Planck Institute for Molecular Genetics, Berlin - sauer@molgen.mpg.de
Markus Nöthen – Life & Brain Center, University of Bonn, Bonn - markus.noethen@uni-bonn.de

Introduction

Positional cloning is now widely used for the identification of gene defects that underlie inherited diseases. A necessary first step for positional cloning is the mapping of the gene locus that co-segregates within families with a particular disease or trait, which allows allocating a specific chromosomal position to the responsible gene. Although mapping was initially developed for monogenetic traits, it has become a widely used strategy to localize genetic factors involved in the aetiology of multifactorial diseases. It is one of the key missions of the National Genome Research Network (NGFN), to identify the genetic factors involved in complex diseases by establishing the link between phenotypic data and the genome of the patients. This approach requires substantial resources in high-throughput genotyping of "short tandem repeat" (STR) and "single nucleotide polymorphism" (SNP) markers. Therefore, a consortium was founded comprising the most potent genotyping centers of Germany to form a National Genotyping Platform (NGP) (**Fig. 1**). This platform acts in a coordinated fashion to meet all genotyping requirements of the projects supported by the NGFN.

For SNP genotyping a plethora of different methods is available each with specific advantages in particular set-ups **[1].** There are two important criteria for cost-efficient genotyping, the number of DNA samples and the number of SNPs under investigation (**Fig. 2**). The National Genotyping Platform (NGP) of NGFN-2 provides efficient technologies for all conceivable scenarios in SNP typing. Array-based procedures, such like the methods supported by Illumina and Affymetrix, are particularly useful when large numbers of SNP markers are to be typed. However, a common problem of all highly parallel microarray-based technologies is their inability to efficiently type subsets of SNPs for stepwise approaches and for replication studies. This opens a window for SNP typing methods working in single or low-plex mode such as MALDI-TOF mass spectrometry and Pyro-sequencing. The SNPlex assay, which is based on a multiplexed version of the oligoligation assay (OLA) with an automated read-out on capillary DNA sequencers, is a powerful alternative technology to all mentioned procedures in cases in which medium numbers of SNPs (fifty to several hundred) and several thousand DNAs have to be genotyped.

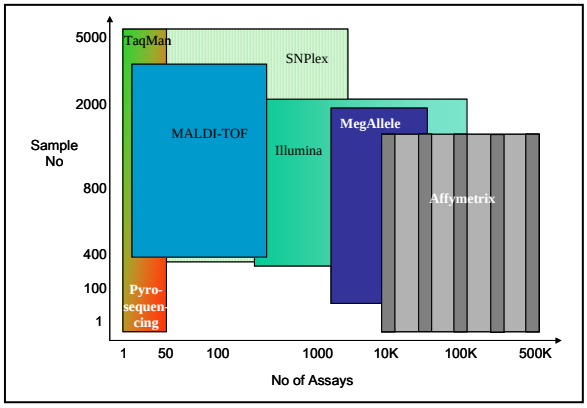

Fig 2: *Choice of SNP typing technology is highly dependent on the numbers of samples and SNPs to be genotyped.*

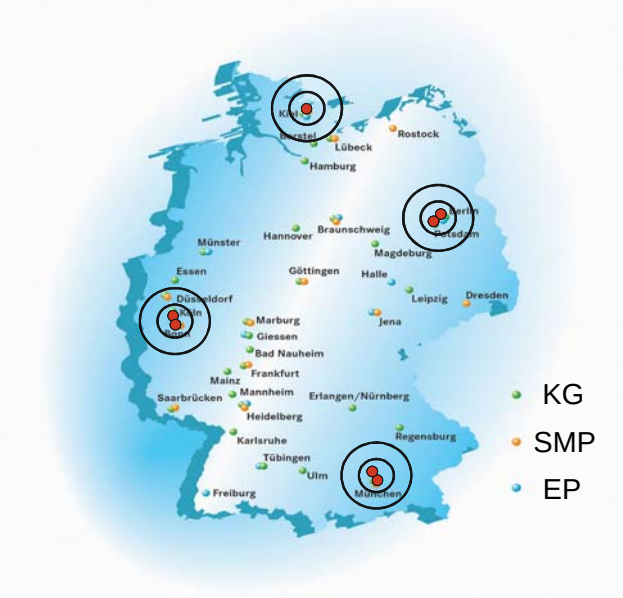

Fig 1: *Geographic distribution of the genotyping centers representing the National Genotyping Platform of the NGFN.*

Results/Project Status

The participating genotyping centers have developed different expertise with the different technologies and therefore closely cooperate during project management (**Table 1**).

Tab 1: *Genotyping technologies provided by the centers.*

Method	Berlin	Bonn	Kiel	Köln	München
STR marker	•			•	
Pyrosequencing	•			•	•
TaqMan	•	•	•	•	•
MALDI-TOF	•	•			•
SNPlex	•		•		
Illumina		•			•
MegAllele	•			•	
Affymetrix	•			•	•

STR markers

Microsatellite or STR marker typing at high-throughput level is technically demanding and needs special expertise. The NGP provides this service at the Gene Mapping Center in Berlin and the Cologne Center for Genomics. Although STRs are being replaced by SNPs step by step, there is still a need for STR marker-based genome scans in model organisms as long as sufficient SNP resources are not available.

Furthermore, due to their highly polymorphic nature, STR markers are very useful for supplementing SNP scans, especially if reliable haplotype information is necessary. Marker sets for human, mouse, rat, and rhesus are available within the NGP.

Single and low-plex SNP technologies

In candidate gene approaches in general rather small numbers of SNPs have to be genotyped and high flexibility with respect to the sample size is required. Highest flexibility is achieved with Pyrosequencing™. In addition to its high flexibility Pyrosequencing™ is also superior for any kind of quantitative analysis such as the analysis of DNA methylation, pooled DNA samples, or allelic expression imbalance. Unfortunately, this technique is the costliest one. Therefore, other methods such like Applied Biosystem's TaqMan™ assays or SEQUENOM's mass spectrometers protocols may be used whenever sufficient samples are available to completely fill 384-well microtitre plates (MTPs). In addition to SEQUENOM's protocols, the MALDI-TOF-based GOOD assay is available, which was developed at the MPIMG in Berlin-Dahlem. The GOOD assay has demonstrated excellent performance leading to a significant cost reduction in a number of genotyping studies.

SNPlex technology

In order to support large scale genotyping projects, such as comprehensive association studies, at competitive prices, a SNPlex platform has been established in Kiel. By using Applied Biosystem's SNPlex™ technology, it is now possible to simultaneously detect 48 SNPs utilizing less than 1 ng of gDNA per SNP genotype. Typically, this DNA is generated by a whole-genome amplification reaction (WGA), thus allowing >3000 SNPlex genotypes to be generated from 20ngs of genomic DNA.
SNPlex is based upon multiplex OLA/PCR and capillary electrophoresis for high throughput genotyping. The principle of the oligonucleotide ligation reaction (OLA), which is the allele-discriminating step, is based upon a method developed around the 1990's: the ligation chain reaction (LCR). In contrast to LCR, SNPlex uses a normal PCR step to exponentially amplify the ligation products. Furthermore, oligonucleotides are designed and synthesized for only one DNA strand by ABI. Hence, SNPlex takes advantage of ligation's specificity and PCR's increase in sensitivity.
At the moment ABI provides 48plex assays but 96plex and 192plex assays will be available in the future. This will increase the throughput at the Kiel facility as well. SNPlex technology design pipeline accepts both AB validated assays and custom SNPs taken from any source. SNPs can be submitted for design of the multiplex panels in the form of a SNP ID (rs#, hCV#) or of a FASTA file. In addition to the existing Tecan Genesis robots, two Tecan te-mo robots (see **Fig. 3**) were installed. Both have a 96 needle pipetting head and can process a maximum of 5 plates at once.

***Fig 3** (left): Tecan te-mo with a high precision head and cooling unit for MTPs. (right): Tecan te-mo with standard head and PW384 for washing MTPs.*

Running one 384 MTP on the 3730xl capillary sequencer takes app. 1 hour. Afterwards the raw data is automatically analyzed by using GeneMapper Analysis Software. Stringent settings for clustering and calling ensure a high quality of individual genotypes. Importing the data to a database has been automated as well while quality checks, such as looking for Mendel errors, are not neglected.
Genotyping one 384 MTP (368 gDNA samples plus 16 control wells) with SNPlex will yield 17,664 genotypes (considering the ideal case). The current throughput is app. 150,000 genotypes per day. In addition to the increase of throughput, prices for genotyping decreased significantly.

Array-based SNP technologies

SNP typing at ultra-high throughput level is performed with array-based technologies such like GeneChip™ microarrays from Affymetrix or BeadChip™ arrays from Illumina. They allow the simultaneous interrogation of 100,000 – 500,000 SNPs in a single experiment from as little as 250-750 ng DNA of a sample. High-speed mapping pipelines based on Affymetrix technology (**Fig. 4**) have been first established by the Gene Mapping Center in Berlin and the Cologne Center for Genomics. At present, similar pipelines are under construction in Bonn and Munich.

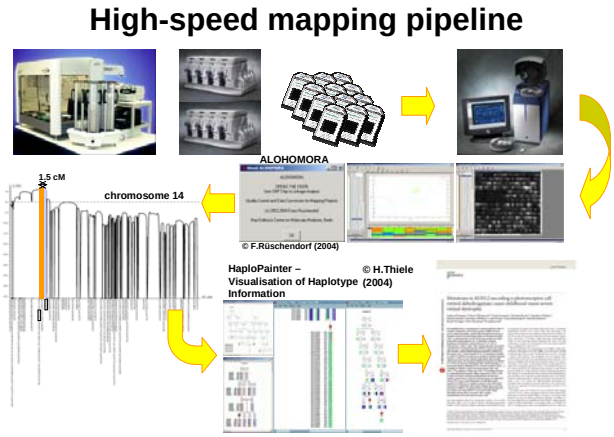

***Fig 4:** SNP typing on arrays accelerates mapping projects.*

In addition to the standard arrays with a fixed SNP content custom-designed arrays of smaller size (1500-20,000 SNPs) are available through the GoldenGate™ assay from Illumina or the MegAllele™ assay from ParAllele/Affymetrix.

Outlook

As suggested already in the middle of the 1990s, high density SNP mapping should be a method, that enables identification of susceptibility genes for complex disorders by linkage disequilibrium mapping either on a regional or genome wide basis **[2]**. Nowadays the advances in genotyping technology allow to analyse markers at densities required for such experiments, i.e. densities that allow resolution of haplotype block structure, even in a genome-wide approach. This will give rise to a number of new success stories in complex disease genetics such like the identification of the chromosome 6 susceptibility gene in sarcoidosis **[3]**. Thus, high density LD mapping with SNPs may become one of the canonical approaches to the identification of complex disease genes in the future.

*Lit.: **1.** Sauer S, et al. Miniaturization in functional genomics and proteomics. Nat Rev Genet 2005 Jun;6(6):465-76. **2.** Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. Science 273:1516-1517. **3.** Valentonyte R, Hampe J, Huse K, Rosenstiel P, Albrecht M, Stenzel M, Nagy M, Gaede KI, Jenisch S, Platzer M, Koch A, Lengauer T, Seegert D, Schwinger S, Krawczak M, Müller-Quernheim J, Schürmann M, Schreiber S. Sarcoidosis is associated with a truncating splice site mutation in the BNTL2 gene. Nature Genetics 2005. 37(4):357-64.*