Genotyping Chip
A spotted array-based open mutation testing system

Virginie Aris[1,2,3] Deanna Streck[2], Gokce A. Toruner[1,2,3], Patricia Soteropoulos[2,3] and James J. Dermody[1,2]

[1]Dept. of Microbiology & Molecular Genetics, NJ Medical School, UMDNJ, [2]Center for Human and Molecular Genetics, NJ Medical School, UMDNJ, [3]Center for Applied Genomics, Public Health Research Institute, Newark NJ

Introduction

The application of microarray technology for diagnostic mutation detection testing has the potential to change molecular genetics and personalized medicine. The simultaneous analysis of multiple mutations is expected to not only decrease the turn over time in the lab but also decrease the cost per genotype.

The aim of this study was to develop a spotted array-based mutation testing system which is compatible with most of the commercially available microarray scanners and reagents.

Materials and Methods

Samples

Cell lines used in the training set were obtained from the Coriell Institute for Medical Research. Samples used for validation were unlabelled patient samples obtained from the UMDNJ hospital with some of them known to have one of the mutations studied. DNA was isolated using Gentra Puregene DNA isolation kit (Gentra Systems).

Microarray design

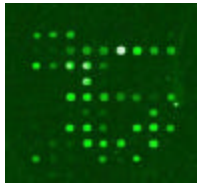

Figure 1: *Example of an array*

Oligonucleotides were designed for the wild type and mutant sequences for each of the 27 alleles tested (table 1). Eight replicate sets of oligonucleotides were printed onto SuperAldehyde coated glass slides (TeleChem International, Inc.) using an Omnigrid 100 robot (Genomic Solutions) (Figure 1).

Labeling and Hybridization

Three multiplexed PCR reactions were performed, yielding Cy3 labeled amplicons of approximately 100bp-1kb. The PCR reactions were pooled and used to probe for the microarray.

Slides were hybridized on a GeneTac Hybridization Station (Genomic Solutions) for 16 hours. Following hybridization, slides were rinsed with sterile water and dried with Nitrogen gas. Slides were scanned using a GenePix 4000B scanner (Molecular Devices) using the GenePix 4.0 software.

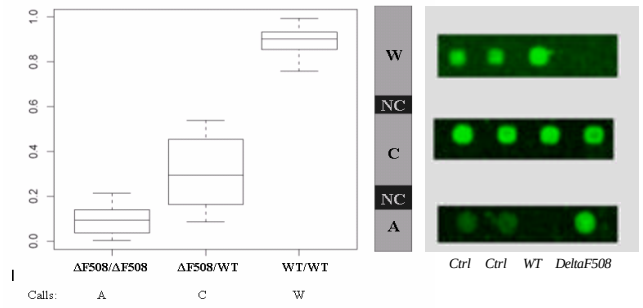

Figure 2: Box plot of the relative ratio values for the training set of the Cystic Fibrosis Delta F508 mutation. On the right you see the spot intensity and the calls the linear discriminant analysis would make for the different relative ratios.

Disease	Mutation
Canavans	693 C -A
Connexin 26	167delT, 35delG
Cystic Fibrosis	delta F508, W1282, G542, N1303K
Familial Dysautonomia	2390 G-C, 2507 +6 T-C
Fanconi Anemia C	IVS 4+4 A-T
Gaucher	1226 A-G, 84gg, 1448 T-C
Hemochromatosis	C282Y, H63D
MTHFR	1298 A-C, 677
Mucolipidosis	IVS3-2 A-G
NPD	fsP330delC, delta R608, L302P
Prothrombin	G20210A
Sickle Cell	HbS
Sickle Cell HbC	HbC
Tay Sachs	Exon 7, Intron 12
Torsion Dystonia	946 delGAG

Table 1: List of the mutations in the panel.

Analysis

A filter was set to eliminate unreliable low signal and high saturated spots, thus both mutant and wild type spots must have intensities above 50 and below 64000. For each mutation, we then computed the relative wild type allele expression level:

$$r_i = \frac{WT_i}{WT_i + Mut_i}$$

If r_i is close to 1 then the genotype is homozygote Wild Type, if it is close to 0 the genotype is homozygote mutant and if r_i is close to 0.5 then the genotype would be heterozygote for the mutation. The relative ratio distributions were estimated independently for each mutation as hybridization affinity and cross-hybridization are sequence dependant.

Forty two samples of known genotypes were hybridized on the genotyping chip and the relative ratio was computed for each mutation. Data was plotted on box plots to see how well the categories (Wild Type, Homozygote Affected and Heterozygote) were separated (Figure 1). Outliers were removed at this point form the training set and testing sets.

With an equal prior probability, we used the median relative ratio for each mutation on the validation set to classify the 26 testing samples using Linear Discriminant Analysis (R software). We then set a threshold of posteriori probability of a sample belonging to a group of 0.7 in order to make a call. An example of the calls based on the relative ratio and the training set are given in figure 2.

For more infromation visit www.cag.icph.org or email virginie.aris@umdnj.edu

Results and Conclusions

The test provided for accurate results for most of the 26 independent samples. Overall, 6.4% of the spot pairs (mutant and wild types) failed under the low intensity filter due to either failure of the PCR in the multiplex or a poor hybridization. After classification we observed a 9.6% no call rate (when the ratio is in between two categories), 2.1% false positives rate, and no false negatives. This is a very conservative result as none of the carrier or affected individuals would be missed in this screening test.

In this study we were successful in taking the first steps towards developing a clinically relevant mutation detection assays in an open and flexible microarray platform which can be easily modified for inclusion of additional mutations and SNPs or the creation of new mutation detection panels. This genotyping chip allows for fast, accurate, cost effective, mutation screening in a diagnostic laboratory environment.

References

Dong J, Katz, et.al. Mol Genet Metab 2001;73:160-3.

Gilbert F. Genet Test 2001;5:83-85.

Strom CM, et.al. Genet Med 2004; 145-152.

Venables, W. N. and Ripley, B.D. (2002) Modern Applied Statistics with S. Fourth ed. Springer.

Acknowledgments

This work is supported by the Center for Human Molecular Genetics, Center for Applied Genomics and by a grant to P. Soteropoulos from the New Jersey Commission on Science and Technology.