Simulated Pharmacogenomics Exercises for the Cybertory^{TM} Virtual Molecular Biology Laboratory

Camille Bodley Troup, Bill Martin, Carl E. McMillan, Robert M. Horton
Attotron Biotechnologies Corporation
camitroup@comcast.net, rmhorton@attotron.com

Abstract

The emerging discipline of pharmacogenomics applies genomic technologies to predict individuals' responses to therapeutic drugs based on the genetic sequences of drug targets and enzymes involved in drug metabolism. The first diagnostic test for genotyping two important drug metabolizing enzymes (CYP450 2D6 and 2C19) has been FDA approved. This assay involves PCR amplification and identification of functionally relevant SNPs using a DNA microarray. We have developed a simulation of PCR and microarray analysis of these important enzymes using the open-source Cybertory(TM) Virtual Molecular Biology Laboratory (www.cybertory.org).

Genomic sequences are instantiated from these succinct genotype descriptions by substituting allele sequences onto a framework of the full human reference genome. PCR products can be used as probes for microarray hybridization. Using the Cybertory (TM) microarray image generator[1], we have designed a virtual "SNP chip" to distinguish alleles of CYP450 2D6 based on signal intensities from perfect match and mismatched probe sets.

1. Introduction

The introduction of molecular diagnostic tests to genotype the CYP450 enzymes 2D6 and 2C19 will have important consequences for healthcare providers involved in prescribing drugs which are substrates for these enzymes [2, 3]. Although clinical research to determine dosing regimens based on pharmacogenetic and pharmacogenomic data is ongoing, there is a critical need to develop education materials for students and professional healthcare providers who will need to interpret the results of these tests.

2. Materials and Methods

Three "patients" were constructed to be homozygous for CYP2D6*1A, 4A, and 10A, respectively, by substituting the appropriate allelic sequences into a framework of sequence from the human reference genome. Simulated PCR was used to amplify a 5093 bp PCR fragment containing the entire coding region of the 2D6 locus from each patient using primers 2D6-F CCAGAAGGCTTTGCAGGCTTCA and 2D6-R ACTGAGCCCTGGGAGGTAGGTA [4].

Because a minor product of 303 bp was produced at about 3% of the molar concentration of the major band even at high stringency, the major band was isolated for use as a microarray probe.

Table 1. Virtual "Patients" with CYP2D6 alleles representing major allelic variants and phenotypes

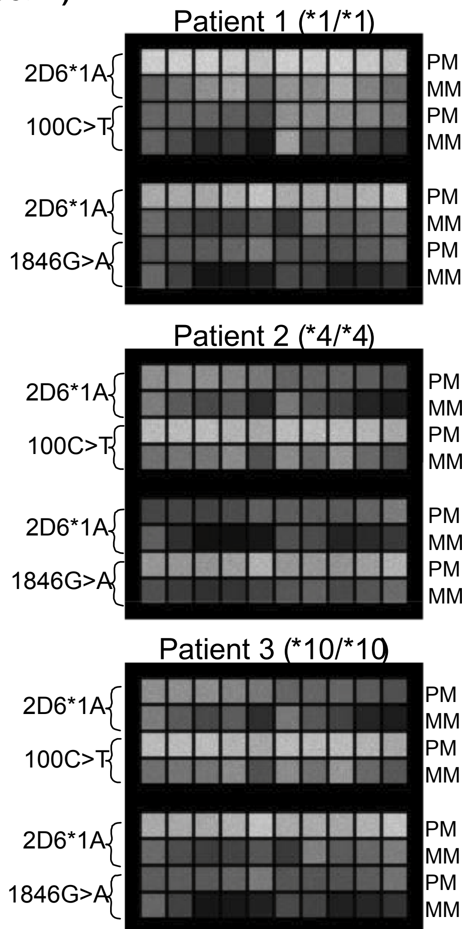
Patient	CYP2D6 Genotype	CYP2D6 Phenotype
1	*1/*1	EM
2	*4/*4	PM
3	*10/*10	PM/IM
4	*5/*1	IM
5	*17/*1	IM/EM
6	*10/*1	IM/EM
7	*1/*2x2	UM

Allelic variants of CYP2D6 are from http://www.imm.ki.se/CYPalleles/. UM = ultra-metabolizer, more than 2 functional 2D6 gene copies; EM = extensive metabolizer 2, functional genes; IM = intermediary metabolizer, 1 functional gene; PM = poor metabolizer, no functional gene

PCR products are genotyped by simulated hybridization to an Affymetrix-style microarray [5]. This array tests for two single nucleotide polymorphisms (SNPs), 100C>T and 1846G>A. Each polymorphism is represented by 4 rows of ten reporters, for a total of 80 reporters per array. Each reporter is 25 nucleotides long. The first set of

reporters represents perfect matches to the standard sequence, 2D6*1A, in a sequence window around the location of the 100C>T SNP. Five match the top strand, and five match the bottom strand. Reporters within a set of five differ in the position of the 25-mer window relative to the polymorphic site. The next ten reporters each contain a single artificial mismatch (MM) with respect to the standard 2D6*1A sequence. The third set represents perfect matches (PM) to the sequence containing the 100C>T polymorphism. Reporters in the fourth set contain an additional artificial base substitution relative to the 100C>T sequence (thus two differences from the standard.) A similarly designed collection of forty reporters is used to determine the variant present at position 1846.

Figure 1. Simulated microarray images for three patients showing two functionally important SNP's in CYP2D6; (100C>T) and (1846G>A).



Hybridization was simulated by comparing the PCR product to a BLAST database of reporter sequences. Each HSP was converted to an intensity by modeling equilibrium hybridization at 45 degrees using nearest neighbor thermodynamic estimations. The overall intensity of a feature represents the sum of intensity contributions from all HSPs matching its reporter.

3. Results and Discussion

A set of individuals are represented in Table 1. They represent common allelic variants of CYP2D6 and CYP2C19 found in various ethnic groups in the human population

In Figure 1, two functionally important SNP's were chosen to generate a virtual SNP chip to compare patients representing two common CYP2D6 variant alleles, *10 and *4 compared to the *1 allele. Patient 1 has two functional CYP2D6 alleles. Patient 2 is homozygous for the *4 allele, which has a splice site mutation (1846G>A) abolishing all enzyme activity, in addition to the 100 C>T SNP. The *10 allele contains the 100C>T SNP, which results in reduced enzyme activity of CYP2D6.

4. Acknowledgements

This work is funded by NIH SBIR grant 2R44RR013645-02A2 to Attotron Corporation. Open source software is available under the GNU GPL at www.cybertory.org.

5. References

[1] B. Martin and R.M. Horton "A Java Program to Create Simulated Microarray Images" *CSB2004 Conference Proceedings*, pp. 564-5.
[2] M. Ingelman-Sundberg "Genetic Polymorphisms of Cytochrome p450 2D6 (cyp2D6), Clinical Consequences, Evolutionary Aspects and Functional Diversity" *The Pharmacogenomics Journal* (2005) 5:6-13.
[3] J. Kirchheiner, K. Nickchen, M. Bauer, M.L., Wong, J. Licinio, I. Roots and J. Brockmoller "Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response" *Molecular Psychiatry* (2004) 9:442-473.
[4] Hajime Matsuzaki, *et al*. "Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array" *Genome Research* (2004) 14:414-425.
[5] E. Lundqvist, I. Johansson and M. Ingelman-Sundberg "Genetic mechanisms for duplication and multiduplication of the human *CYP2D6* gene and methods for detection of duplicated *CYP2D6* genes" *Gene* (1999) 226:327-338