An integrated system supports calling genotype and translating haplotype to phenotype for PGx implementation

**Genotype Phenotype translation system for PGx implementation**

# **BACKGROUND ART**

The single base primer extension method (SBE) is an effective and sensitive approach that can cover over 30 loci in a single reaction. As a cost-effective approach, this method has been widely used in many applications such as forensics, disease diagnosis, and pharmacogenetics (PGx).

Pharmacogenomics (PGx) play an essential part of precision medicine. Several studies confirmed the benefit of panel-based PGx tests in clinical settings. Many important PGx genes encode metabolizing enzymes such as cytochrome P540 (CYPs) and transporter; and these genes highly vary in the population. The variants often referred as star-alleles (e.g., CYP2D6\*10B). Star allele can be a single genetics variant or combination of multiple variants across the gene (haplotype). Therefore, identifying star-allele from customed panels plays a role in PGx implementation.

Many tools have been developed such as PharmCAT, Stargazer, Astrolabe, and Aldy to identify star allele from sequencing data; and all of them focus on analyzing Next-Generation Sequencing data. Moreover, genotyping software (e.g., GeneMapper) does not support identifying haplotypes from experimental data. Therefore, identifying haplotypes of customed PGx test panels remains challenging.

Conventionally, with custom PGx panel results, PGx experts often have two options (Figure 1A): (1) they manually identify the haplotypes from raw intensity data (chromatogram plots); or (2) they identify genotype results using a genotyping software and upload the result to other tools to match haplotypes. Then, they must manually manage the data and analytical results, and history for further steps. This way raises several following risks: manual processes increase human biases; it is hard to record and track the history of the whole process; and it is difficult to scale to many testing samples because it requires more effort to manage all those data and processes.

**DEFINITIONS**

Haplotype (a.k.a star allele) is a speciific allele that inheried from on patent. In this work, we use star allele term for reducing the confusion.

Diplotype is a specific combination of two haplotypes

Phenotype is composite of an organism’s observable characteristics or trait.

Metabolizer phenotype of can be a strengthen of a particular enzyme with a specific substrate, for example: Normal metabolizer, Intermediate Metabolizer, or Poor Metabolizer.

**SUMMARY OF INVENTION**

This system of present invention is a could computational method for automated identifying star allele from a custom PGx panel and translating those result into phenotype. The system comprising of: a) a module to load and mange the customed pre-defined marker positions (allele definition and following bin range), and pre-defined star allele nomenclatures; b) a module that allow user to identify peaks from experimental intensity from single based extended method by adjusting bin range and height of intensity; c) a module to identify star allele, diplotype, and phenotype from peak detection results and pre-defined nomenclature.

[In this work, we developed **a cloud system that integrates all processes to overcome these challenges** (Figure 1B). Once PGx experts upload the raw intensity data to the system, they can do all further steps: identifying genotypes, quality controlling, and matching haplotypes/phenotypes. All activity logs have also been recorded for any further tracking steps. Figure 2 demos how the system implemented on a website application and Figure 3 shows haplotype phenotype translation as example of CYP2D6.]

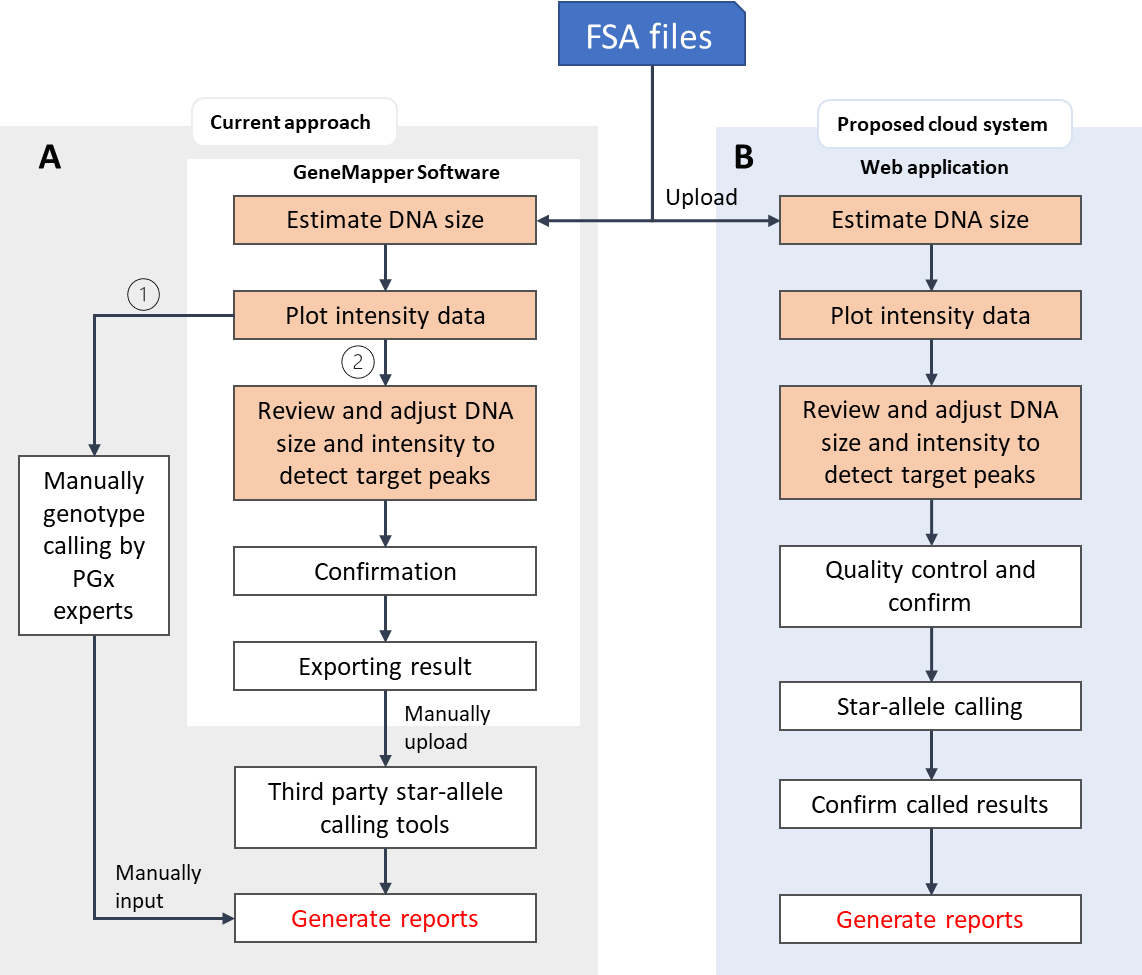


Figure 1. Overview of current approach and purpose system

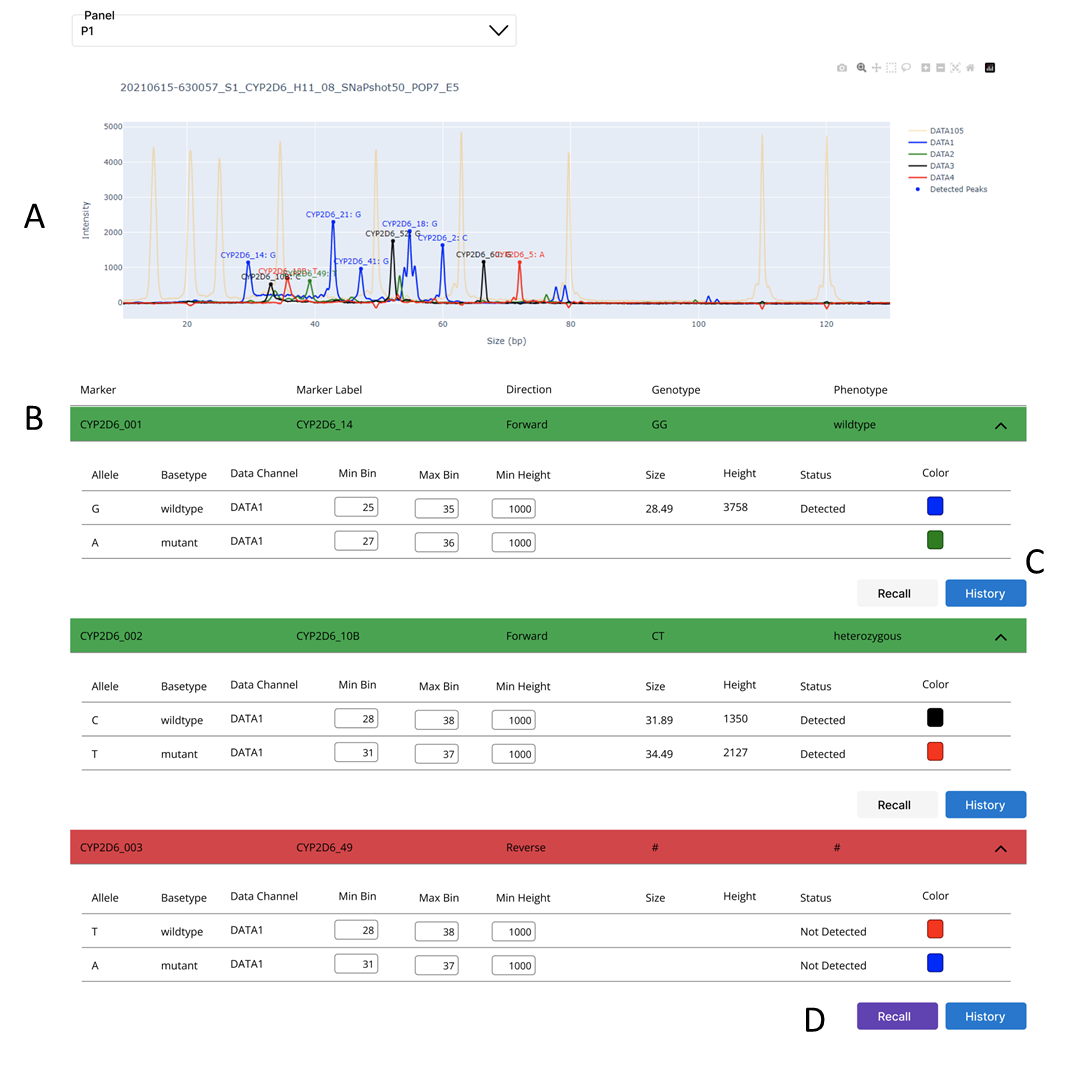


Figure 2. Prototype of the system implemented on web application; (A) raw intensity plot with detected markers; (B) genotype calling results of each marker in a panel; (C) history of each marker; (D) user can adjust data and re-identify genotype for a specific marker.

Graphical user interface, application, Teams

Description automatically generated

Figure 3. Haplotype and phenotype of target gene have been identified