VOLUME 10 NOMOR 2 DESEMBER 2023 ISSN 2548 – 611X

JURNAL
BIOTEKNOLOGI & BIOSAINS INDONESIA

Homepage Jurnal: http://ejournal.brin.go.id/JBBI/index

DEVELOPMENT OF PLASMID-BASED FOR EXTERNAL CONTROL MATERIALS OF CYP2D6*10 (rs1065852) GENE PCR-BASED DETECTION

Jekmal Malau[1,7], Aliya Azkia Zahra[1], Ahsanal Kasasiah[1,6], Ratika Rahmasari[4], Muhareva Raekiansyah[5], Siti Rohmah[1], Nanda Diva Meilani[1], Annisa Frastica Septi[1], Aurora Fatimatuz Zahro[1], Fathina Annajla[1], Anastasia Aliesa Hermosaningtyas[2,3], Indah Laily Hilmi[1*]

[1]Department of Pharmacy, Faculty of Health Science,
Universitas Singaperbangsa Karawang, West Java, Indonesia
[2]Laboratory of Pharmaceutical Biology and Biotechnology, Department and Division of Practical Cosmetology and Skin Disease Prophylaxis, Poznan University of Medical Sciences, Poznan, Poland
[3] Doctoral School, Poznan University of Medical Sciences, Bukowska 70, 60-812 Poznan, Poland
[4]Microbiology and Biotechnology Laboratory, Faculty of Pharmacy, Universitas Indonesia,
West Java, Depok, 16424, Indonesia
[5]Helix Laboratory and Clinic, Depok, 16424, Indonesia
[6]Microbial Biotechnology Laboratory, UPA Laboratorium Terpadu,
Universitas Singaperbangsa Karawang, West Java, 41361, Indonesia
[7]Biology Laboratory, UPA Laboratorium Terpadu, Universitas Singaperbangsa Karawang, West Java, 41361, Indonesia
**Email: indah.laily@fkes.unsika.ac.id*

ABSTRACT

*Reliable clinical diagnosis of Single Nucleotide Polymorphisms (SNPs) is necessary for personalizing tamoxifen medication according to CYP2D6*10 genetic variations. Our research aimed to create a recombinant plasmid for external control material with a molecular size of 3812 bp. The recombinant plasmid was achieved by cloning an 838 bp gene insert of CYP2D6*10 rs1065852 into a 2974 bp pJET1.2 plasmid into Escherichia coli DH10B and selection on ampicillin agar medium. Isolated E. coli recombinants provided the plasmid molecules for analysis. Bi-directional sequencing and Real-Time PCR confirmed the presence of wild-type and mutant rs1065852 DNA fragments in the plasmid, namely homozygote CC and TT. The conclusion is that we have successfully introduced a novel recombinant plasmid developed by cloning the SNP rs1065852, which carries the 100C>T mutation, using pJET 1.2/blunt system, which could significantly enhance the accuracy of clinical SNP diagnostics for personalized medicine in breast cancer treatment.*

Keywords: *Plasmid recombinant; external control material; Single Nucleotide Polymorphisms (SNPs); CYP2D6*10 rs1065852; Clinical diagnostic*

ABSTRAK

Diagnosis klinis yang tepat dan akurat terhadap *Single Nucleotide Polymorphisms* (SNPs) rs1065852 merupakan hal penting dalam pengobatan individu terapi tamoxifen berdasarkan variasi genetik gen CYP2D6 alel 10. Penelitian ini bertujuan untuk pengembangan kontrol eksternal pengujian diagnostik berbasis plasmid rekombinan yang memiliki ukuran 3812 bp. Hal ini dilakukan dengan kloning fragmen DNA target sepanjang 838 bp yang mengandung variasi genetik rs1065852 ke dalam plasmid pJET1.2 yang memiliki panjang 2974 bp. Selanjutnya dilakukan transformasi ke dalam *Escherichia coli* DH10B dan kemudian diseleksi pada medium

***Received**: 18 November 2023* ***Accepted**: 19 December 2023* ***Published**: 29 December 2023*

agar yang mengandung antibiotik ampicillin. Isolat *E. coli* rekombinan berhasil diisolasi untuk mendapatkan molekul plasmid yang dianalisis lebih lanjut pada uji konfirmasi. Hasil uji konfirmasi yang dilakukan menggunakan metode sanger sekuensing F/R dan Real-Time PCR menunjukkan adanya keberadaan variasi DNA rs1065852 yaitu tipe liar (homozigot CC) dan mutan (homozigot TT) dalam plasmid. Kesimpulannya, kami telah berhasil mengembangkan plasmid rekombinan baru dengan kloning SNP rs1065852 yang membawa mutasi 100C>T, menggunakan sistem pJET 1.2/blunt. Hasil penelitian ini diharapkan dapat meningkatkan akurasi diagnosis SNP rs1065852 pada uji klinis untuk penegakkan diagnostik terapi tamoxifen kanker payudara melalui pengobatan individu.

Kata Kunci: Rekombinan plasmid; kontrol eksternal; *Single Nucleotide Polymorphisms* (SNPs); CYP2D6**10* rs1065852; diagnostik klinis

INTRODUCTION

Breast cancer, which originates in mammary tissue, represents one of the most common neoplastic diseases among women globally, including in Indonesia. Breast cancer is a significant health concern worldwide, with an estimated 2.3 million new cases, which constitute 11.7% of all cancer cases, making it the fifth leading cause of death globally, with a death rate of 685,000 (Sung et al. 2021). In 2020, Indonesia reported 68,858 new cases of breast cancer, which constituted 16.6% of the total cancer cases (Ministry of Health 2022). The prevalence of breast cancer continues to increase annually.

Tamoxifen acts as an estrogen receptor antagonist, inhibiting breast cancer cell proliferation (Jordan 2015; Abderrahman and Jordan 2019). The effectiveness of tamoxifen in treating breast cancer is significantly affected by genetic variations, especially the single nucleotide polymorphisms (SNPs) in the *CYP2D6* gene, which are responsible for drug metabolism (Sutiman et al. 2016). Several studies have investigated the relationship between the CYP2D6 genotype and tamoxifen metabolism rate, leading to diverse cumulative phenotypes (Shah and Smith 2015). Up to this point, over 100 DNA variations of the CYP2D6 gene have been reported and categorized by phenotypic characteristics connected to different metabolic processes (Dean 2014). One of the genetic variations that frequently observed in Asian populations, especially in Indonesia, is the CYP2D6**10* rs1065852 allele, which has a moderate impact on tamoxifen metabolism (Goetz et al. 2008; Byeon et al. 2018; Puaprasert et al. 2018; Chan et al. 2019; Yenny et al. 2019; Kisoi et al. 2020; Han et al. 2021; Bakar 2021). Dose adjustments tailored to individual patients are required for those with intermediate-to-poor phenotype conditions. The determination of this dosage typically can be derived from the results of molecular diagnostic DNA profiling of a patient (Abdullah-Koolmees et al. 2020).

Diagnostic testing plays a crucial role in identifying genetic variants for guidelines of dose adjustments in CYP2D6**10* rs1065852 SNPs. One primary technique in detecting CYP2D6 genetic variants is PCR and DNA sequencing, which have shown promising results in improving the efficacy of tamoxifen therapy (Holden et al. 2011). The precision of PCR diagnostics for *CYP2D6* gene variation detection depends significantly on incorporating external control material. This material guarantees assay validity, recognizes procedural or operational errors, and needs to be used for cross-laboratory proficiency testing (Kalman et al. 2016; Niraula and Bataju 2020).

Several external control materials have been developed for *CYP2D6* gene variant analysis, including residual patient specimens and cell lines (Pratt et al. 2010; Pratt et al. 2016; Lin et al. 2017; Gaedigk et al. 2019). Nevertheless, the use of cell lines could be improved by high cost and complex cultivation techniques (Lin et al. 2017). This indicates the necessity for devising substitute external control materials, including those relying on recombinant plasmids (van der Straaten et al. 2008; Lin et al. 2017)

This study focuses on constructing recombinant plasmid as external control encompassing SNP rs1065852 into plasmid pJET1.2/blunt (Lahijani et al. 1996). An external control, incorporating SNP rs1065852 produced in *E. coli* DH10B transformants, was validated using bi-directional sequencing. Furthermore, the assessment of its applicability via Real-Time PCR, a predominant technique employed thus far, was conducted to enhance diagnostic accuracy.

To the best of our knowledge, this is the first study conducted to develop an external control material using a recombinant plasmid carrying SNP rs1065852 and using *E. coli* DH10B as the production system. The findings of this research offer promising insights into the utilization of plasmid-based external control materials for routine PCR-based genetic testing. Such an approach could substantially refine the precision and accuracy of molecular diagnostic processes, essential to advancing personalized medicine.

MATERIALS AND METHODS

Location and Time

The research was conducted in The Microbiology Laboratory of Pharmacy and The Microbial Biotechnology Laboratory at Universitas Singaperbangsa Karawang, West Java, from September until November 2023.

Samples and Reagents

This study used bacterial cultures of recombinant *E. coli* DH10B. Some of the reagents included Presto™ Mini Plasmid Kit (Geneaid, Taiawan), CloneJET PCR Cloning Kit (Thermo Scientific, USA), Primers and probe (IDT, USA), PCR MyTaq™ master mix (Bioline, USA), Thunderbird™ Probe qPCR Mix (Toyobo, China). Gblock DNA Synthetic SNP rs1065852 (IDT, USA), Nuclease Free Water (Himedia, India), TE Buffer (Himedia, India), Agarose Gel (Vivantis, Malaysia) and GelRed (Biotium, California).

In-silico Construction of Recombinant Plasmids

For *in-silico* plasmid construction, the pJET 1.2/blunt plasmid with a sequence length of 2974 base pairs was used as the backbone sequence, obtained from Addgene (www.addgene.org). The sequence of the *CYP2D6* gene (NCBI accession ID: NG.008376.4) was used as the reference sequence. Plasmid construction process was performed using SnapGene software (www.snapgene.com). A flanking sequence containing the target SNP of left and right regions was inserted into the pJET 1.2/blunt plasmid. The inserted gene is placed at a location that has an EcoRV restriction enzyme recognition sequence (5'-GAT|ATC-3'), as shown in Figure 1.

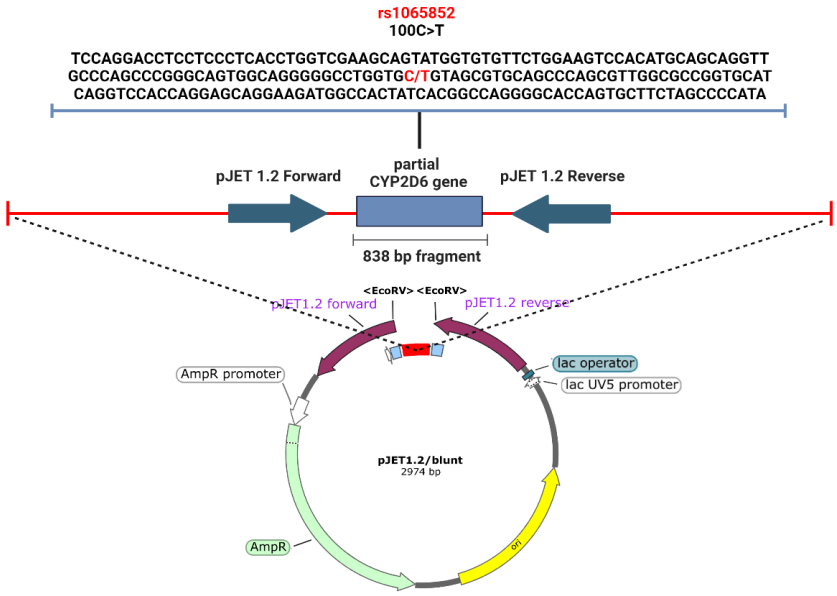

Figure 1. Plasmid map and DNA sequence annotation of pJET1.2/blunt with partial insertion of the cytochrome P450 2D6 (CYP2D6) gene on allele 10, including the flanking DNA containing the rs1065852 (C100T) polymorphism.

Cloning and Transformation in *E. coli* DH10B

An insert DNA fragment consisting of a flanking sequence encompassing the target SNP rs1065852, with a length of 838 base pairs, was utilized. The flanking sequences for the rs1065852 target SNP were synthesized by Integrated DNA Technologies (IDT, USA). Subsequently, a linear gBlock DNA fragment was obtained, which is planned to be used as the insert during cloning. The cloning of the insert DNA fragment was achieved by following the previously described in-silico plasmid construction. Cloning was performed using plasmid pJET1.2 with a blunt end system according to the manufacturing CloneJET PCR Cloning Kit protocol (Catalog No. K1231, Thermo Scientific, USA). The transformation was carried out on *E. coli* DH10B competent cells using the transformation kit (Catalog No. EC0113, Thermo Scientific, USA) by the heat-shock method. *E. coli* transformants were then grown on Luria-Bertani (LB) agar medium and incubated overnight at 37°C. Transformed *E. coli* carrying the target of interest were selected using ampicillin and confirmed using PCR.

Recombinant Plasmid DNA Extraction

Recombinant *E. coli* DH10B bacteria were cultured on Luria-Bertani agar media. Before plasmid DNA extraction, the isolates were subcultured on Luria broth media (Himedia Laboratories, India) supplemented with 100 µg mL^{-1} ampicillin. The cultures were incubated at 37°C and 180 rpm overnight on a shaker incubator. Furthermore, plasmid DNA extraction and purification were performed using Presto™ Mini Plasmid Kit Quick Protocol (Geneaid, China) and with heat modification at 60°C for 2 minutes before the elution. DNA concentration was determined with the Nanodrop EPOCH2 2 (BioTek, USA). DNA of extracted recombinants plasmid were separated by electrophoresis at 70 volts for 45 minutes (Mupid-exu, Japan), using 1% agarose gel (Vivantis, Malaysia). The band visualization was obtained using the Accuris-Benchmark Scientific UV transluminator (USA).

Primers and Probe Selection

A primer-probe set specifically targeting the flanking DNA of rs1065852 (C100T) in the *CYP2D6* gene was used for detecting SNPs with a reference sequence obtained from NCBI (accession ID: NG.008376.4). The primer and TaqMan probe sequences are presented in Table 1 (Hoseini and Sauer 2015; Puaprasert et al. 2018). Integrated DNA Technologies (IDT, USA) synthesized all primers and probes.

Primers and probe set are expected to generate amplicon length 956 and 73 bp by PCR and Real-Time PCR amplification. The representative primers-probe attachment map is shown in Figure 2.

Table 1. Primers and probes for PCR and Real-Time PCR

Primers and Probe	DNA sequences of oligonucleotides	Amplicon size (bp)	References
pJET1.2 F	5'-CGACTCACTATAGGGAGAGCGGC-3'	956	Hoseini and Sauer 2015 with amplicon size modification in this study
pJET1.2 R	5'-AAGAACATCGATTTTCCATGGCAG-3'		
F_2D6C100T	5′-CCTGGTGGACCTGATGCA-3′	73	Puaprasert et al., 2018 with dye modification in this study
R_2D6C100T	5′-CCCGGGCAGTGGCA-3′		
Probe Wild-Type 2D6C100T_WT	5′-[HEX]-CCTGGTG**G**GTAGCGTG-[NFQ-MGB]-3′		
Probe Mutant 2D6C100T_MT	5′-[FAM]-CCTGGTG**A**GTAGCGTG-[NFQ-MGB]-3′		

Figure 2. Map of Forward/Reverse/Probe attachment position on partial CYP2D6 gene which produces a 73 bp amplification size

PCR and Direct Sequencing

SNP genotyping analysis was performed by conventional PCR and bi-direct sequencing methods. The primers, pJET1.2 Forward and pJET1.2 Reverse, amplified the DNA flanking region and SNP points. PCR amplification was performed with a total of 10 µl of the optimized reaction mix, comprising 5 µl of MyTaq™ master mix (Bioline, USA), 1 µl of DNA template (50 ng.µl^{-1}) and 0.4 µl (10µM working concentration) of forward and reverse primer. The mixture was then made up to the desired final volume with nuclease-free water. The amplification protocol began with a 3-minute initial denaturation at 95°C, followed by 35 cycles of denaturation at 95°C for 15 seconds, annealing at 60°C for 30 seconds, extension at 72°C for 45 seconds, and a final extension at 72°C for 2 minutes. Each run included a no template control, and all runs were performed in triplicate.PCR amplification was conducted using the T100 PCR System (Bio-Rad, USA). Subsequently, the presence of PCR products with the appropriate size was separated by electrophoresis at 70 volts for 45 minutes (Mupid-exu, Japan) using 1.5% agarose gel (Vivantis, Malaysia). Band visualization was performed using a UV transilluminator (Accuris-Benchmark Scientific, USA). Single-band PCR products were then subjected to sequencing using capillary electrophoresis Sanger sequencing (Macrogen, Korea). SNP calling was evaluated using Snapgene viewer software.

Real-Time PCR SNP Genotyping

SNP genotyping was conducted using the CFX96 Touch Deep Well real-time PCR detection system (Bio-rad, USA). The final reagent volume was 20 µl, comprising of 10 µl Thunderbird Probe qPCR Mix master mix (Toyobo, China), 1 µl DNA template (20 ng.µl^{-1}), and 1 µl forward and reverse primers at a final concentration of 500 nM. Additionally, 0.8 µl TaqMan MGB Probe wild-type was included with a final concentration of 400 nM and 0.5 µl TaqMan MGB Probe mutant was added at a final concentration of 250 nM. The amplification protocol consisted of a pre-denaturation at 60°C for 30 seconds, followed by a denaturation at 95°C for 5 minutes. Subsequently, there were 40 cycles of 95°C denaturation for 15 seconds and annealing extension at 60°C for 45 seconds, followed by a post-read at 60°C for 30 seconds. Fluorescent signals were detected during the annealing extension and post-read stages for allele discriminations. Amplification results and SNP clustering patterns were obtained using Bio-Rad CFX Maestro V2.0 software (Bio-rad, USA).

Data Analysis

Real-Time PCR data were analyzed using CFX Maestro 2.0 software (Bio-Rad, USA). The amplification curve and the Ct/Cq value formed were used for the follow-through analysis. The evaluation was conducted by observing the allele discrimination pattern (Bars-Cortina et al. 2019).

Sequencing data was analyzed to ensure the designed assay and recombinant plasmid amplified the SNPs contained in the plasmid construct. Data analysis was performed by employing the SnapGene software to observe the result of cycle sequencing for base-calling and trimming in the sequencing. The peak quality was employed for analyzing the sequencing data quality. The peak quality under 20 was trimmed on the electropherogram right and left of the sequence (Koh et al. 2021). The following step was to generate consensus. The sequence results were assembled employing BioEdit software, and then the SNPs calling were analyzed (Logan et al. 2014).

RESULT

Research Design Guidelines

Our initial step was selecting the reference sequence of the *CYP2D6* gene at allele 10 that carries the SNP rs1065852 (100C>T). Additionally, we conducted in silico analysis of the primer sets and probes. The third step involves cloning and transforming recombinant pJET, which carries the rs1065852 (100C>T) SNP gene insert into the *E. coli* DH10B host. Subsequently, the SNP rs1065852 (100C>T) in *E. coli* DH10B transformants is confirmed and validated in the final step. All of the above steps, flanking DNA map, and recombinant plasmid construction are presented in Figure 3.

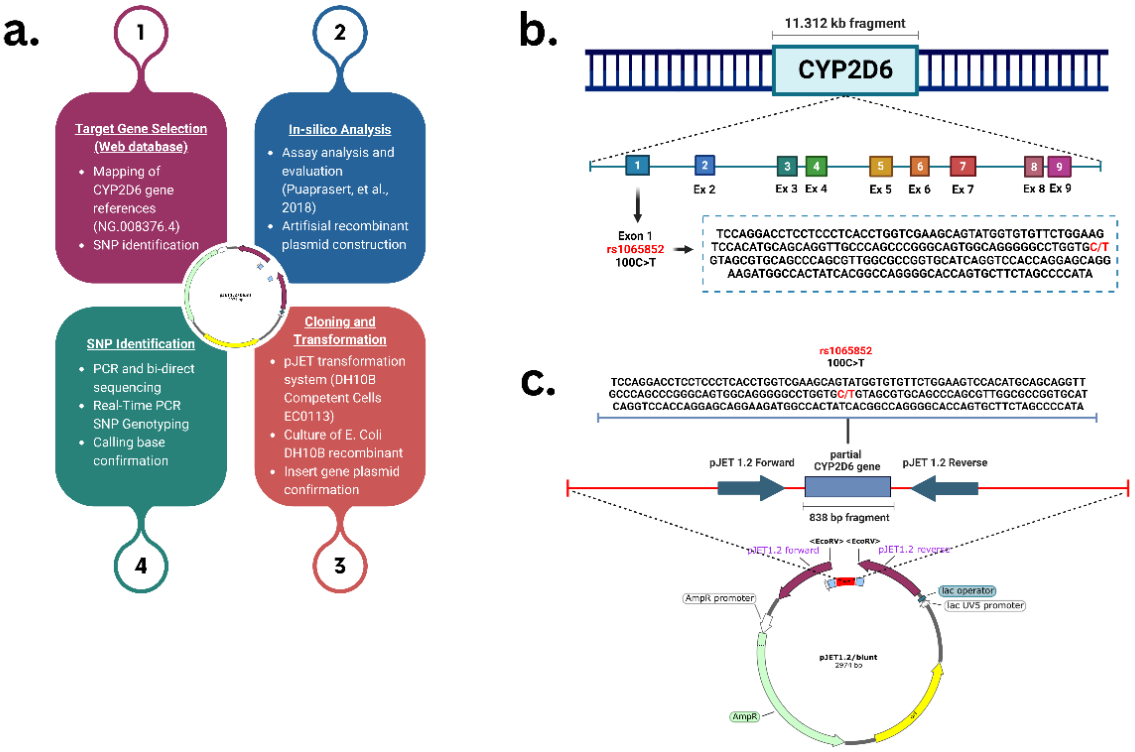

Figure 3. Map schematic and experimental flow diagram of the development of external control materials CYP2D6*10 rs1065852 plasmid-based. (a) Four-step guidelines for recombinant plasmid construction and validation. (b) The target gene was selected from genomic databases, and SNP rs1065852 identification was evaluated using a partial CYP2D6 gene. (c) In-silico analysis of assays and artificial recombinant plasmid construction.

Plasmid DNA Construction

Plasmid pJET 1.2/blunt is a linear cloning vector that accepts DNA inserts ranging in size from 6 bp to 10 kbp (ThermoFisher 2009; Hoseini and Sauer 2015; Astuti et al. 2019). The 5' end of the vector is phosphorylated, thus negating the

need for additional PCR primer phosphorylation. This feature expedites the ligation of flat-end PCR products synthesized by DNA polymerases with proofreading activity into the vector in a rapid process that takes approximately 5 minutes (ThermoFisher 2023). The backbone plasmid pJET1.2/blunt has a total length of 2974 bp. An 838 bp gene insert containing mutant and wild-type DNA sequences rs1065852 (100C>T) in the form of a DNA gBlock was successfully produced and inserted through the EcoRV cutting site located between the pJET1.2 forward and reverse primers. As a result, two recombinant plasmids, pJET1.2 (100C>T) WT and pJET1.2 (100C>T) MT, were obtained with 3812 bp of length (Figure 4). The obtained results were then used as a starting point for the cloning and transformation process at a later stage.

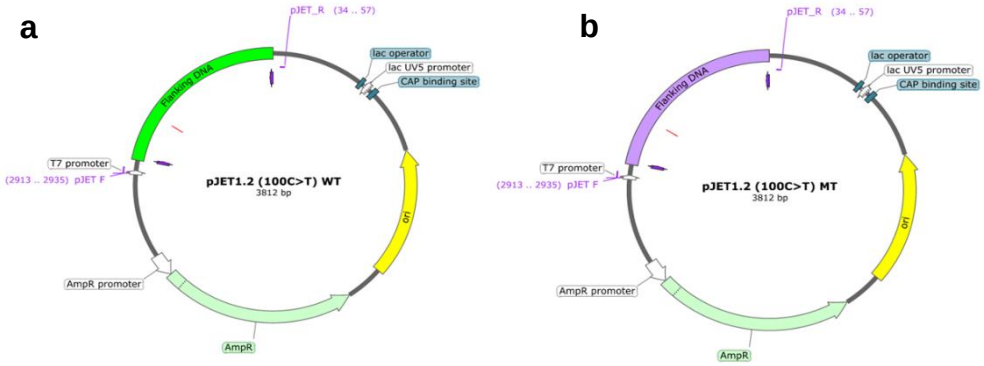

Figure 4. The construction maps of two plasmids (a) pJET1.2 containing rs1065852 (100C>T) WT, (b) pJET1.2 containing rs1065852 (100C>T) MT using partial CYP2D6 gene.

Cloning and Transformation

The DNA cloning was executed per the established *in-silico* plasmid construction protocol. This process involved the utilization of an 838 bp gene insert containing mutant and wild-type DNA sequences in the form of a linear DNA synthetic gBlock and pJET1.2 plasmid, designed explicitly for blunt-end cloning. This system facilitates the efficient and precise integration of DNA fragments, ensuring high fidelity and accuracy in the cloning procedure (Nawawi et al. 2022). In the context of recombinant plasmid DNA propagation for the production of external control materials, the determination of success depends critically on the transformation results in *E. coli* transformants, defined as positive or negative (Nawawi et al. 2023). Thus, it is pivotal to ensure the effectiveness of the transformation process. Based on the presence of bacterial growth observed, the results showed that two *E. coli* DH10B transformants containing the recombinant plasmids pJET1.2 (100C>T) WT and pJET1.2 (100C>T) MT, respectively, were successfully grown on Luria-Bertani (LB) agar supplemented with 100 µg.mL^{-1} ampicillin (Figure 5). These isolates were then subjected to confirmatory testing to verify the presence of the intended SNP DNA insert fragments, since this early screen step only verified colonies through ampicillin resistance.

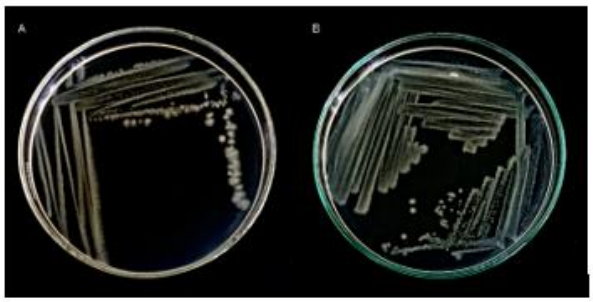

Figure 5. Isolates of *E. coli* DH10B transformant (A) with pJET1.2 (100C>T) WT and (B) with pJET1.2 (100C>T) MT on Luria-Bertani agar supplemented with 100 µg.mL^{-1} ampicillin incubated at 37°±2°C for 12-18 h

Recombinant Plasmid Determination

Prior to plasmid DNA extraction, the isolates were subcultured on Luria broth media supplemented with 100 µg.mL^{-1} ampicillin. The cultures were incubated at 37±2°C and 180 rpm overnight on a shaker incubator. Extraction and purification of plasmid DNA followed the Presto™ Mini Plasmid Kit Quick Protocol, slightly modified by heating the elution buffer at 60˚C for 2 minutes before the elution. We found that the data derived from the gel electrophoresis showed plasmid bands with a length of 3812 bp, following the expected results based on the data obtained from plasmid *in-silico* construction. Detailed information on the gel electrophoresis and visualization results is provided in Figure 6. The results of measuring plasmid DNA concentration are shown in Table 2. The data confirmed that the purity meets the required values, which should equal or exceed 1.8-2.0 (Sambrook et al. 1989; Kirby 1990).

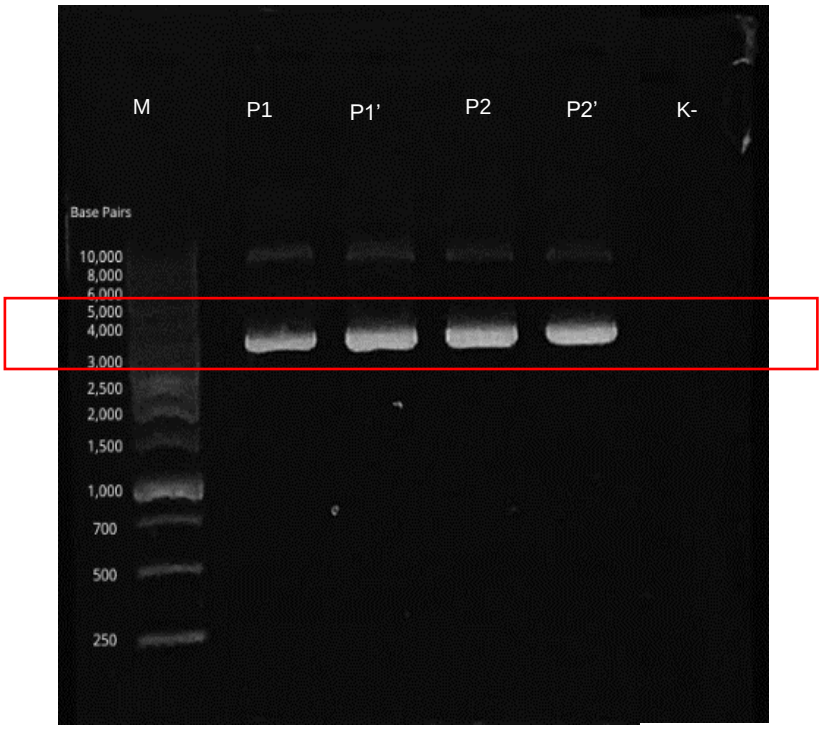

Figure 6. Recombinant plasmid gel electrophoresis result. M= DNA Marker 10kb; P1-P1'= pJET1.2/blunt carrying rs1065852 wild-type, amplicon length = ± 3800 bp; P2-P2'= pJET1.2/blunt carrying rs1065852 mutant, amplicon length = ± 3800 bp; K- = negative control

Table 2. The concentration and purity of DNA plasmid extracted

No	Samples ID	DNA Conc (ng.µl^{-1})	DNA Purity (260/280)
1	P1 (HsADME.rs1065852. A.11.1 WT 1)	138,48	1,999
2	P1' (Hs ADME.rs1065852. A.11.1 WT 2)	207,89	1,947
3	P2 (Hs ADME.rs1065852. A.11.2 MT 1)	174,33	1,952
4	P2' (Hs ADME.rs1065852. A.11.2 MT 2)	115,28	1,955

SNP Genotyping Verification with Conventional PCR

SNP genotyping was analyzed using conventional polymerase chain reaction (PCR) and bi-directional sequencing. The pJET1.2 Forward and pJET1.2 Reverse primers amplified the DNA regions adjacent to and encompassing the SNP loci to produce an amplicon fragment with a size of ± 800 bp (Figure 7a). Single-band PCR products were then subjected to Sanger sequencing for base-calling verification, as shown in Figure 7b. Homozygous wild-type is reported as CC, whereas homozygous mutant is

called TT. The results indicated that recombinant plasmid as an external control material demonstrated robust performance with perfect concordance in base calling.

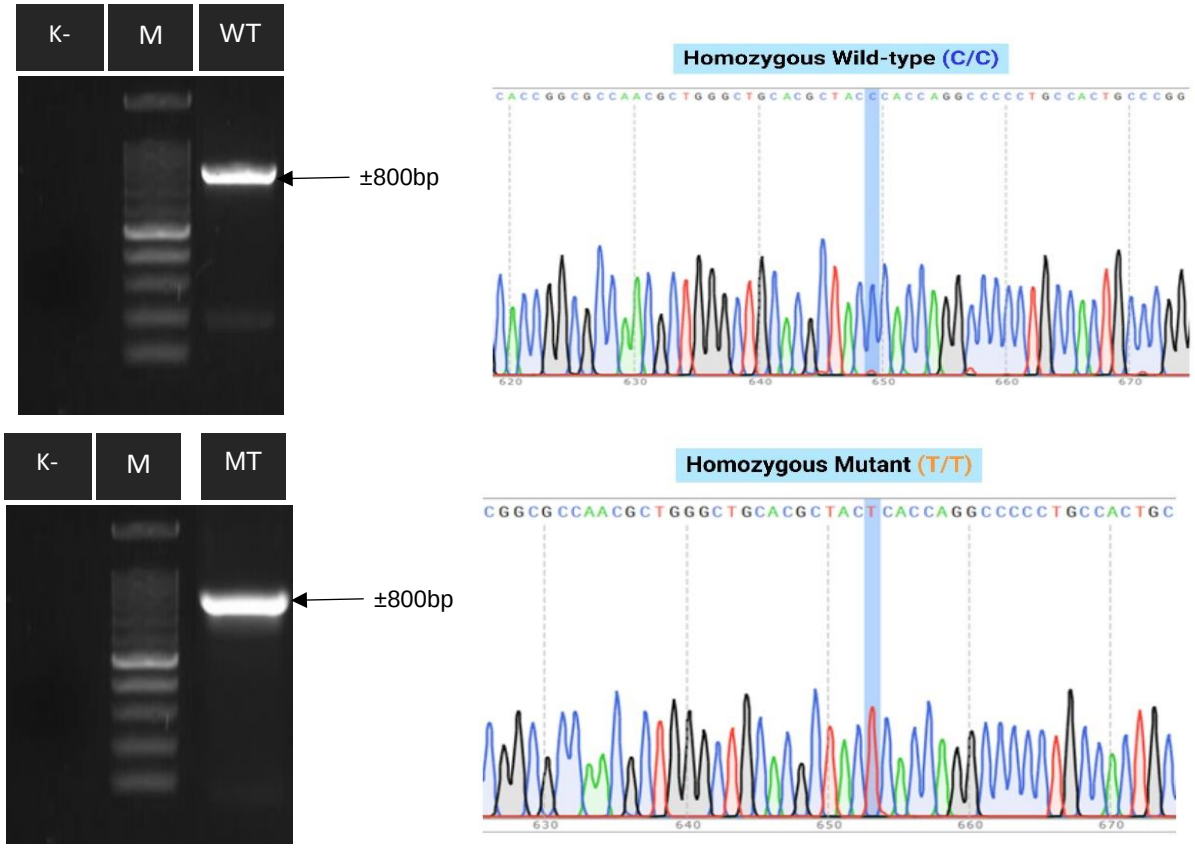

Figure 7. (A) PCR amplification from *E. coli* DH10B transformant carrying rs1065852 (100C>T) SNP using pJET1.2 Forward and pJET1.2 Reverse, amplicon length = ± 800 bp. M =DNA marker 10 kb; WT = pJET1.2/blunt carrying rs1065852 wild-type; MT= pJET1.2/blunt carrying rs1065852 mutant; K- = negative control. (B) Representative electroferogram of rs1065852 wild-type and mutant.

Assessment of Real-Time PCR Allelic Discrimination

Assessment of its applicability of external control material recombinant plasmid via Real-Time PCR, a predominant CYP2D6**10* rs1065852 diagnostic test employed thus far, was performed using 20 ng. μl^{-1} of P1 and P2 samples. The CFX Deep Well Real-Time PCR System with Maestro Software v2.0 (Bio-Rad, USA) enabled an auto-call analytical approach that automatically generated allele discrimination plots. This provided well-separated clusters for genotype calling, and the amplification curve was consistently reliable. Replicate assays of individual samples showed 100% identity in variant calls, as evidenced by their clustering in the same region on the scatterplot. In addition, the genotyping results from PCR sequencing and real-time PCR with TaqMan probe assays were in perfect agreement for both wild-type and mutant genotypes in the sample pairs tested (Figure 8). This data reported the success of the developed external control material for detecting CYP2D6**10* rs1065852.

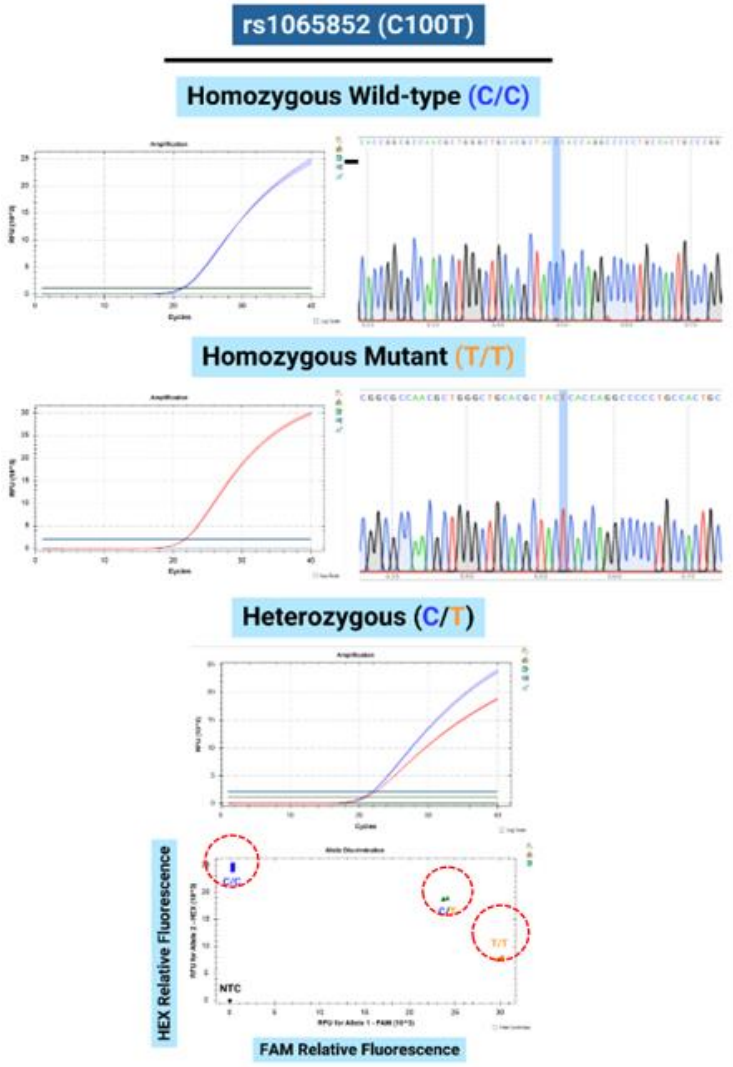

Figure 8. Comparison of the two genotyping platforms. PCR Sanger sequencing (electropherograms) and Real-Time PCR TaqMan probe (amplification plot) detected the presence of wild-type and mutant in the CYP2D6 gene C100T. The X-axis of the amplification plot shows the relative fluorescence for wild-type alleles (blue curve) and mutant alleles (red curve), respectively. Clusters of the homozygous wild-type, heterozygous and homozygous mutant are also shown in the allelic discrimination plots (NTC = no template control). The Real-Time PCR TaqMan probe genotyping results followed the results obtained by PCR Sanger sequencing.

DISCUSSION

Early-stage detection is crucial for the successful treatment of breast cancer. The therapeutic approach typically involves a combination of surgical resection, radiation therapy, and pharmacological interventions (including hormonal therapy, chemotherapy, and targeted biological agents) (Adams 2008). The US Food and Drug Administration has authorized Tamoxifen for a range of therapeutic and preventive uses in breast cancer. Its pharmacological effects are contingent upon the drug-metabolizing enzyme cytochrome P450 2D6 (*CYP2D6*) activating it metabolically, leading to the production of active metabolites such as 4-hydroxytamoxifen and endoxifen. The effectiveness of Tamoxifen is affected by genetic variation in the CYP2D6 gene known as Single Nucleotide Polymorphisms (SNPs). The impaired function of the CYP2D6 enzyme due to reduced activity can result in lower levels of the endoxifen metabolite and reduced

therapeutic effect (Hoskins et al. 2009). Individualized dose modifications are necessary for patients with intermediate-to-poor phenotypic disorders. The dose adjustments are usually obtained from the findings of the DNA profile of a patient through molecular diagnostic results (Abdullah-Koolmees et al. 2020).

Success in clinical diagnostic testing is inseparable from using external control materials during assay development and validation. In routine testing, this also ensures the accuracy and reliability of the test. It will certainly ensure that DNA variations in the CYP2D6 gene found in clinical samples accurately represent the patient's DNA profile, which will greatly help to achieve the efficacy of personalized treatment of drugs such as tamoxifen as part of breast cancer therapy in patients.

Related to this, the idea addressed to help characterize clinically relevant genetic variations in the CYP2D6 gene is the development of cost-effective external control material for guarantees assay validity recognizes procedural or operational errors in routine clinical molecular diagnostics (Kalman et al. 2016; Niraula and Bataju 2020). To identify the existence of the CYP2D6*10 rs1065852 gene variant, we employed two approaches, namely PCR-directional sequencing and Real-Time PCR.

Various external control materials, such as residual patient specimens and cell lines, have been created to investigate CYP2D6 gene variants (Pratt et al. 2010; Pratt et al. 2016; Lin et al. 2017; Gaedigk et al. 2019). However, the utilization of cell lines can be enhanced by addressing the issues of high expenses and intricate cultivation methods (Lin et al. 2017). This highlights the need to create alternative external control materials, such as those based on recombinant plasmids using *E. coli* DH10B as the production system.

In this study, we initially designed and successfully constructed an in-silico plasmid of pJET1.2 harboring the rs1065852 (100C>T) WT, along with pJET1.2 containing the rs1065852 (100C>T) MT, as depicted in Figure 4. The DNA insert was derived from a flanking region of approximately 400 nucleotides upstream and downstream of the rs1065852 SNP. This methodology aligns with the work of (Hoseini and Sauer 2015; Astuti et al. 2019), who utilized pJET.1.2 as the backbone for plasmid cloning. The gene insert was strategically integrated into the EcoRV site, which is situated between the pJET1.2 forward and reverse primers on the plasmid and is located 371 base pairs from the 5' end of the plasmid.

The utilization of the pJET1.2/blunt plasmid as a backbone in the in-silico construction of the CYP2D6 SNP rs1065852 (C100T) plasmid is predicated on several advantages, namely (1) the presence of the eco47IR gene, which encodes the Eco 47I restriction enzyme (Nawawi et al. 2022); (2) contains pJET1.2 forward and reverse primers, enabling PCR or sequencing to verify the presence and orientation of DNA insertions (ThermoFisher 2023); (3) Existence of the EcoRV restriction site for DNA insertion; (4) pJET 1.2/blunt plasmid demonstrated a 100% cloning efficiency (Nawawi et al. 2022). Consequently, employing this plasmid in *in-silico* plasmid construction can enhance the efficiency and accuracy of the cloning process.

Furthermore, the evaluation of the success of in silico plasmid construction was proven by the successful cloning and transformation of recombinant plasmids into *E. coli* DH10B. Linear DNA fragment in the form of a gBlock with a length of 838 bp containing the CYP2D6 SNP rs1065852 gene was successfully cloned into plasmid pJET 1.2/blunt and subsequently transformed into *E. coli* DH10B-competent cells. This observation confirms the successful isolation of two recombinant isolates, each carrying the pJET1.2 (100C>T) wild-type and mutant, respectively (Figure 4). Screening of transformant colonies was based on the ability of the isolates growing on LA+Ampicillin media (Manna et al. 2013). This preconditioning step aims to multiply the desired cloning vector source (Oertel et al. 2015; Astuti et al. 2019). According to the plasmid extraction results, the size of the DNA insert was consistent with expectations, approximately ±3500 bp (Figure 5). The pJET1.2/blunt cloning plasmid is typically employed for storing and duplicating desired gene insert (Thermo Fisher Scientific, patent publication: US 2009/0042249 A1, Genbank

accession number EF694056.1). The use of pJET1.2 as a cloning vector has been widely reported, including *E. coli* (Hoseini and Sauer 2015), wheat plant (Cengiz et al. 2013), MSP1 gene of *Plasmodium falciparum* (Mawardi et al. 2018), *Albuca rautanenii* (Zablocki et al. 2014), *Bacillus subtilis* (Abdel-Salam et al. 2018), *Citrullus lanatus* (Hernandez et al. 2021).

The recombinant plasmid was confirmed by PCR, which showed positive results with clear bands on gel electrophoresis. The approximate molecular size of the bands was around ±800 base pairs. The sequencing result confirms the presence of both wild-type and mutant rs1065852 DNA fragments within the pJET1.2 (100C>T) plasmid. The PCR was confirmed using the pJET1.2 forward and reverse primers derived from the plasmid's backbone. These primers are often efficient in amplifying gene areas with a maximum length of 1800 base pairs (Hoseini and Sauer 2015).

Finally, real-time PCR validation has also assessed the recombinant plasmid's applicability. The selection of real-time PCR for this validation test is based on its extensive utilization in the field, where it is frequently used in molecular clinical diagnostics, particularly for pharmacogenomic and genetic testing (Broeders et al. 2014; Gokduman et al. 2016; Soga et al. 2022; Nuraeni et al. 2023). Based on the analysis of base calling and allelic discrimination, the genotyping results obtained from the Real-time PCR platform followed those obtained from PCR Sanger sequencing.

CONCLUSION

This study introduces a novel recombinant plasmid developed by cloning the SNP rs1065852, which carries the 100C>T mutation, using the pJET 1.2/blunt system. We have successfully cloned essential genetic variations relevant for diagnostic testing by inserting a partial gene of CYP2D6**10* into a plasmid and subsequently conducting a sequencing analysis of the inserted genetic material. In addition, the recombinant plasmids were assessed using Sanger sequencing and Real-Time PCR TaqMan-based, revealing a perfect agreement rate of 100%. This result provides a new direction in studying quality control materials for standardizing clinical laboratory tests. However, we also agreed since this is only the proof of principle. Plasmid characterization and external validation by an independent laboratory are needed to demonstrate the suitability of this plasmid for use as a reference material for external control.

ACKNOWLEDGEMENT

This research was funded by National Competitive Research, Indonesian Ministry of Education, Culture, Research, and Technology (grant number: 255/SP2H/UN64.10/PP/2023). We also would like to express our gratitude to Universitas Singaperbangsa Karawang for providing research facilities and infrastructure.

REFERENCES

Abdel-Salam MS, Ameen HH, Kassab ASM, Mahgoob AEA, Elkelany US (2018) Enhancement of nematicidal potential through cloning and expression of chitinase gene from *Bacillus subtilis* subsp. *subtilis* BTN7A strain. J Genet Eng Biotechnol 16:305–310. https://doi.org/10.1016/j.jgeb.2018.05.006

Abderrahman B, Jordan VC (2019) A novel strategy to improve women's health: selective estrogen receptor modulators. pp 189–213

Abdullah-Koolmees H, van Keulen AM, Nijenhuis M, Deneer VHM (2020) Pharmacogenetics guidelines: overview and comparison of the DPWG, CPIC, CPNDS, and RNPGx guidelines. Front Pharmacol 11:595219. https://doi.org/10.3389/fphar.2020.595219

Adams JU (2008) Pharmacogenomics and personalized medicine. In: Nat. Educ. https://www.nature.com/scitable/topicpage/pharmacogenomics-and-personalized-medicine-643/

Astuti NT, Darsono N, Widyaningrum S, Sawitri WD, Astuti SP, Darmanto W (2019) Expression and purification of recombinant coat protein of sugarcane mosaic virus from

Indonesian isolate as an antigen for antibody production. Indones J Biotechnol 24:57. https://doi.org/10.22146/ijbiotech.45551

Bakar NS (2021) Pharmacogenetics of common SNP affecting drug metabolizing enzymes: comparison of allele frequencies between European and Malaysian/Singaporean. Drug Metab Pers Ther 36:173–181. https://doi.org/10.1515/dmpt-2020-0153

Bars-Cortina D, Riera-Escamilla A, Gou G, Piñol-Felis C, Motilva M-J (2019) Design, optimization and validation of genes commonly used in expression studies on DMH/AOM rat colon carcinogenesis model. PeerJ 7:e6372. https://doi.org/10.7717/peerj.6372

Broeders S, Huber I, Grohmann L, Berben G, Taverniers I, Mazzara M, Roosens N, Morisset D (2014) Guidelines for validation of qualitative Real-Time PCR methods. Trends Food Sci Technol 37:115–126. https://doi.org/10.1016/j.tifs.2014.03.008

Byeon J-Y, Kim Y-H, Lee C-M, Kim S-H, Chae W-K, Jung E-H, Choi C-I, Jang C-G, Lee S-Y, Bae J-W, Lee YJ (2018) CYP2D6 allele frequencies in Korean population, comparison with East Asian, Caucasian and African populations, and the comparison of metabolic activity of CYP2D6 genotypes. Arch Pharm Res 41:921–930. https://doi.org/10.1007/s12272-018-1075-6

Cengiz M, Battal A, Aydin G, Eroglu A (2013) Vector construction strategies for transformation of wheat plant. J Anim Plant Sci 23(3):906–912

Chan W, Li MS, Sundaram SK, Tomlinson B, Cheung PY, Tzang CH (2019) CYP2D6 allele frequencies, copy number variants, and tandems in the population of Hong Kong. J Clin Lab Anal 33. https://doi.org/10.1002/jcla.22634

Dean L (2014) Tamoxifen Therapy and CYP2D6 Genotype. In: Medical Genetics Summaries. National Center for Biotechnology Information (US)

Gaedigk A, Turner A, Everts RE, Scott SA, Aggarwal P, Broeckel U, McMillin GA, Melis R, Boone EC, Pratt VM, Kalman L V (2019) Characterization of reference materials for genetic testing of CYP2D6 alleles: a GeT-RM collaborative project. J Mol Diagn 21:1034–1052. https://doi.org/10.1016/j.jmoldx.2019.06.007

Goetz MP, Kamal A, Ames MM (2008) Tamoxifen pharmacogenomics: The role of CYP2D6 as a predictor of Drug Response. Clin Pharmacol Ther 83:160–166. https://doi.org/10.1038/sj.clpt.6100367

Gokduman K, Dilek Avsaroglu M, Cakiris A, Ustek D, Candan Gurakan G (2016) Recombinant plasmid-based quantitative Real-Time PCR analysis of *Salmonella enterica* serotypes and its application to milk samples. J Microbiol Methods 122:50–58. https://doi.org/10.1016/j.mimet.2016.01.008

Han N, Oh JM, Kim I-W (2021) Combination of genome-wide polymorphisms and copy number variations of pharmacogenes in Koreans. J Pers Med 11:33. https://doi.org/10.3390/jpm11010033

Hernandez RN, Isakeit T, Al Rwahnih M, Villegas C, Alabi OJ (2021) First report of watermelon crinkle Leaf-Associated Virus 1 (WCLaV-1) and WCLaV-2 infecting watermelon (*Citrullus lanatus*) in the United States. Plant Dis 105:2025. https://doi.org/10.1094/PDIS-02-21-0249-PDN

Holden MJ, Madej RM, Minor P, Kalman L V (2011) Molecular diagnostics: harmonization through reference materials, documentary standards and proficiency testing. Expert Rev Mol Diagn 11:741–755. https://doi.org/10.1586/erm.11.50

Hoseini S, Sauer MG (2015) Molecular cloning using polymerase chain reaction, an educational guide for cellular engineering. J Biol Eng 9:2. https://doi.org/10.1186/1754-1611-9-

2

Hoskins JM, Carey LA, McLeod HL (2009) CYP2D6 and tamoxifen: DNA matters in breast cancer. Nat Rev Cancer 9:576–586. https://doi.org/10.1038/nrc2683

Jordan VC (2015) The new biology of estrogen-induced apoptosis applied to treat and prevent breast cancer. Endocr Relat Cancer 22:R1–R31. https://doi.org/10.1530/ERC-14-0448

Kalman L V., Datta V, Williams M, Zook JM, Salit ML, Han J-Y (2016) Development and characterization of reference materials for genetic testing: focus on public partnerships. Ann Lab Med 36:513–520. https://doi.org/10.3343/alm.2016.36.6.513

Kirby LT (1990) DNA Fingerprinting. Palgrave Macmillan UK, London

Kisoi M, Imai M, Yamamura M, Sakaguchi Y, Murata S, Ichikawa A, Kinoshita K (2020) Unique genotyping protocol of CYP2D6 allele frequency using Real-Time Quantitative PCR from Japanese healthy women. Biol Pharm Bull 43:904–907. https://doi.org/10.1248/bpb.b19-00512

Koh DW-S, Chan K-F, Wu W, Gan SK-E (2021) Yet Another Quick Assembly, Analysis and Trimming Tool (YAQAAT): A server for the automated assembly and analysis of Sanger sequencing data. J Biomol Tech jbt.2021-3202-003. https://doi.org/10.7171/jbt.2021-3202-003

Lahijani R, Hulley G, Soriano G, Horn NA, Marquet M (1996) High-yield production of pBR322-derived plasmids intended for human gene therapy by employing a temperature-controllable point mutation. Hum Gene Ther 7:1971–1980. https://doi.org/10.1089/hum.1996.7.16-1971

Lin G, Zhang K, Han Y, Li J (2017) Quality control materials for pharmacogenomic testing in the clinic. Clin Chem Lab Med 55:926–933. https://doi.org/10.1515/cclm-2016-0755

Logan G, Freimanis GL, King DJ, Valdazo-González B, Bachanek-Bankowska K, Sanderson ND, Knowles NJ, King DP, Cottam EM (2014) A universal protocol to generate consensus level genome sequences for foot-and-mouth disease virus and other positive-sense polyadenylated RNA viruses using the Illumina MiSeq. BMC Genomics 15:828. https://doi.org/10.1186/1471-2164-15-828

Manna S, Harman A, Accari J, Barth C (2013) Altering the selection capabilities of common cloning vectors via restriction enzyme-mediated gene disruption. BMC Res Notes 6:85. https://doi.org/10.1186/1756-0500-6-85

Mawardi A, Aisoi LE, Lefaan PN (2018) Kloning dan analisis bioinformatika gen MSP1 *Plasmodium falciparum* isolat kota Jayapura. J Biol Papua 10:1–10. https://doi.org/10.31957/jbp.126

Ministry of Health (2022) Kanker Payudara Paling Banyak di Indonesia, Kemenkes Targetkan Pemerataan Layanan Kesehatan. https://sehatnegeriku.kemkes.go.id/. Accessed 23 Oct 2023

Nawawi O, Abdullah MP, Yusuf CYL (2022) A strategy for in-house production of a positive selection cloning vector from the commercial pJET1.2/blunt cloning vector at minimal cost. 3 Biotech 12:216. https://doi.org/10.1007/s13205-022-03289-x

Nawawi O, Abdullah MP, Yusuf CYL (2023) A streamlined strategy for self-production of a commercial positive selection vector, the pJET1.2/blunt cloning vector, using common laboratory E. coli strains. 3 Biotech 13:224. https://doi.org/10.1007/s13205-023-03647-3

Niraula A, Bataju M (2020) External quality assessment practices in clinical biochemistry laboratory: what is the need? Kathmandu Univ Med J 18:86–92.

https://doi.org/10.3126/kumj.v18i1.34660

Nuraeni U, Malau J, Astuti RT, Dewantoro A, Apriori D, Lusiana ED, Prasetya B (2023) Droplet digital PCR versus Real-Time PCR for in-house validation of porcine detection and quantification protocol: an artificial recombinant plasmid approach. PLoS One 18:e0287712. https://doi.org/10.1371/journal.pone.0287712

Oertel W, Wichard T, Weissgerber A (2015) Transformation of *Ulva mutabilis* (Chlorophyta) by vector plasmids integrating into the genome. J Phycol 51:963–979. https://doi.org/10.1111/jpy.12336

Pratt VM, Everts RE, Aggarwal P, Beyer BN, Broeckel U, Epstein-Baak R, Hujsak P, Kornreich R, Liao J, Lorier R, Scott SA, Smith CH, Toji LH, Turner A, Kalman L V (2016) Characterization of 137 genomic DNA reference materials for 28 pharmacogenetic genes: a GeT-RM collaborative project. J Mol Diagn 18:109–23. https://doi.org/10.1016/j.jmoldx.2015.08.005

Pratt VM, Zehnbauer B, Wilson JA, Baak R, Babic N, Bettinotti M, Buller A, Butz K, Campbell M, Civalier C, El-Badry A, Farkas DH, Lyon E, Mandal S, McKinney J, Muralidharan K, Noll L, Sander T, Shabbeer J, Smith C, Telatar M, Toji L, Vairavan A, Vance C, Weck KE, Wu AHB, Yeo K-TJ, Zeller M, Kalman L (2010) Characterization of 107 genomic DNA reference materials for CYP2D6, CYP2C19, CYP2C9, VKORC1, and UGT1A1: a GeT-RM and association for molecular pathology collaborative project. J Mol Diagn 12:835–46. https://doi.org/10.2353/jmoldx.2010.100090

Puaprasert K, Chu C, Saralamba N, Day NPJ, Nosten F, White NJ, Dondorp AM, Imwong M (2018) Real-Time PCR detection of common CYP2D6 genetic variants and its application in a Karen population study. Malar J 17:427. https://doi.org/10.1186/s12936-018-2579-8

Sambrook J, Fritsch ER, Maniatis T (1989) Molecular Cloning: A Laboratory Manual (2nd ed.), 2nd ed. Cold Spring Harbor Laboratory Press, New York

Shah RR, Smith RL (2015) Addressing phenoconversion: the Achilles' heel of personalized medicine. Br J Clin Pharmacol 79:222–240. https://doi.org/10.1111/bcp.12441

Soga K, Nakamura K, Egi T, Narushima J, Yoshiba S, Kishine M, Mano J, Kitta K, Takabatake R, Shibata N, Kondo K (2022) Development and validation of a new robust detection method for low-content DNA using ΔΔCq-based Real-Time PCR with optimized standard plasmids as a control sample. Anal Chem 94:14475–14483. https://doi.org/10.1021/acs.analchem.2c03680

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71:209–249. https://doi.org/10.3322/caac.21660

Sutiman N, Lim JSL, Muerdter TE, Singh O, Cheung YB, Ng RCH, Yap YS, Wong NS, Ang PCS, Dent R, Schroth W, Schwab M, Khor CC, Chowbay B (2016) Pharmacogenetics of UGT1A4, UGT2B7 and UGT2B15 and their influence on tamoxifen disposition in Asian breast cancer patients. Clin Pharmacokinet 55:1239–1250. https://doi.org/10.1007/s40262-016-0402-7

ThermoFisher (2009) Patent Application Publication: Vectors For Cloning. US2009/0042249 A1

ThermoFisher (2023) CloneJET PCR Cloning KIT #K1231, #K1232

van der Straaten T, Swen J, Baak-Pablo R, Guchelaar H-J (2008) Use of plasmid-derived external quality control samples in pharmacogenetic testing. Pharmacogenomics 9:1261–1266. https://doi.org/10.2217/14622416.9.9.1261

Yenny, Panigoro SS, Purwanto DJ, Hidayat A, Louisa M, Andalusia R, Setiabudy R (2019) Association of CYP2D6*10 (c. 100 C>T) genotype with Z-END concentration in patients with breast cancer receiving tamoxifen therapy in Indonesian population. Endocrine, Metab Immune Disord - Drug Targets 19:1198–1206. https://doi.org/10.2174/1871530319666190306094617

Zablocki ODJ, Rybicki EP, Cowan DA (2014) First Report of a Potyvirus Infecting Albuca rautanenii in the Namib Desert. Plant Dis 98:1749–1749. https://doi.org/10.1094/PDIS-07-14-0737-PDN