Digital PCR for determination of cytochrome P450 2D6 and sulfotransferase 1A1 gene copy number variations

Yutaro Motoi\(^1\), Kazufumi Watanabe\(^2\), Hiroyuki Honma\(^1\), Yousuke Tadano\(^1\), Hiroshi Hashimoto\(^2\), Takahiro Kubota\(^1,\)*

\(^1\)Niigata University of Pharmacy and Applied Life Sciences, Niigata, Japan; \(^2\)Hokkaido System Science Co., Ltd., Hokkaido, Japan.

**Summary**

*CYP2D6* and *SULT1A1* occasionally show copy number variations (CNVs), with a larger number generally indicating greater enzymic activity. However, these variations are difficult to calculate using standard methods. With digital PCR, a recently introduced method for CNV analysis, DNA molecules are subjected to limited dilution and separated into nano-scale droplets prior to a PCR assay. Absolute quantitation of copy number can then be performed with high accuracy and sensitivity by determining the number of droplets showing an amplified signal for the target gene. This is the first report of analyses of *CYP2D6* and *SULT1A1* CNVs using a digital PCR method with blood sample from Japanese subjects. Primers and probes were synthesized for the target and reference genes, and copy number calculation was performed using a QX200 Droplet Digital PCR System. Our results showed that the copy numbers in *CYP2D6*\(^{\text{hetero}}\), non-CNV, and *CYP2D6xN* subjects were 1, 2, and 3 to 4, respectively. In addition, in non-CNV and multiplication subjects, the number of copies for *SULT1A1* was 2 and 3 to 6, respectively. We found that the present digital PCR method was useful as well as accurate. In the future, a combined genotyping, allele distinction, and copy number calculation technique will be helpful for analysis of enzymic activity.

**Keywords:** Cytochrome P450 2D6, sulfotransferase 1A1, copy number variation, digital PCR

**1. Introduction**

Cytochrome P450 2D6 (CYP2D6) metabolizes several different drugs, including timolol, propranolol, amitriptyline, propafenone, flecainide, and tamoxifen (1-6). The metabolic ratios of the probe drugs vary, thus patients can be classified into 4 different genotypes; poor metabolizer (PM), intermediate metabolizer (IM), extensive metabolizer (EM), and ultra-rapid metabolizer (UM) (7). Generally, when a PM or IM patient is administered a drug that is inactivated by CYP2D6, the blood concentration and risk of side effects will be increased as compared to EM patients. On the other hand, should a UM patient be administered such a drug, the concentration and therapeutic effect will be lower as compared to EM patients (8-12).

CYP2D6 has been reported to have over 100 gene polymorphisms (13). For example, *CYP2D6*\(^4\), \(^5\), and \(^{14}\) lack enzymic activity, while *CYP2D6*\(^{10}\) and \(^{21}\) show decreased enzymic activity (14-16). Approximately 7% of Caucasians are classified as PM and considered to have *CYP2D6*\(^{4}\) or \(^{5}\), while the frequency of PM in the Japanese population is under 1%, and the main mutations are *CYP2D6*\(^{4}\) or \(^{5}\) and \(^{14}\) (17-23). In contrast, the frequency of IM is high in Japanese at 15%, which can be explained by *CYP2D6*\(^{10}\) (24,25). Since the frequencies of these 3 mutations in Japanese are high, analysis is important to elucidate CYP2D6 activity. In addition, *CYP2D6* occasionally shows copy number variations (CNVs), including gene deletion (*CYP2D6*\(^{4}\)), duplication, and multiplication (*CYP2D6xN*). *CYP2D6*\(^{4}\), \(^{2}\), \(^{4}\), \(^{9}\), \(^{10}\), \(^{17}\), \(^{35}\), \(^{36}\), and \(^{41}\) have been reported as 2-copy alleles (*CYP2D6x2*), among which *CYP2D6*\(^{2}\) is recognized to have 3 to 5 and 13 copies (10,11,26-31). A larger
number of copies generally indicates greater enzymic activity, which induces UM.

The genetic polymorphisms of CYP2D6 are complex and widespread, thus precise estimation of phenotype using a gene test is difficult. Notably, "copy number" and "metabolism activity" are matters for CNV analysis. As for copy number, that is difficult to calculate with existing PCR methods as it is not possible to clearly determine the amount of change in enzymic activity. A more accurate method is needed, as conventional quantitative PCR assays are limited in regard to resolution and can at best distinguish a twofold difference in copy number of a particular gene. Regarding metabolism activity, if the causal genes of PM (e.g., CYP2D6*4 and *36) become duplicated (e.g., CYP2D6*4x2 and *36x2), that also indicates development of PM (27,30,31). Similarly, we previously reported that if CYP2D6*10 (cause of IM) is increased by 2 to 3 copies, there is not a significant difference (29), thus in the case of CYP2D6sN, the activity will be different depending on which allele has the mutation (28). For example, in a comparison of CYP2D6*1x2/*10 and CYP2D6*1/*10x2, the former shows activity nearly to the level of EM, while the latter shows activity lower than EM.

Tamoxifen, a drug developed for treating breast cancer, is metabolized by CYP2D6, while its metabolite, 4-hydroxy tamoxifen, is metabolized by sulfotransferase 1A1 (SULT1A1) (32,33). In other words, CYP2D6 activates tamoxifen and SULT1A1 is involved in its inactivation. The major mutations of SULT1A1 are SULT1A1*2 and *3, whose frequencies are approximately 10% and under 1%, respectively, in Asian individuals (34). SULT1A1*2 shows a lower level of activity than the wild type (35). Moreover, SULT1A1 occasionally shows CNVs (1 or 3-6 copies), thus it is important to calculate the copy number as with CYP2D6. These are important, as such mutations cause individual differences in regard to the effects of tamoxifen in patients administered the drug (36,37).

Digital PCR, in which DNA molecules are subjected to limiting dilution and separated into nano-scale droplets prior to the assay, was recently introduced as a method for CNV analysis. With this technique, absolute quantitation of copy number can be performed with high accuracy and sensitivity by determining the number of droplets with an amplified signal for the target gene. In the present study, we calculated CYP2D6 and SULT1A1 copy numbers using a digital PCR method.

### 2. Materials and Methods

#### 2.1. Samples

Blood samples from randomly selected unrelated Japanese subjects who participated in a previously reported study were obtained (25,29,37). The CYP2D6 and SULT1A1 genotypes were previously analyzed in these subjects. Approval for these studies were obtained from the local Institutional Review Board (25,29) and Chiba Institute of Science (37), separately. The sequences of the primers and probes used in the present study are shown in Table 1.

#### 2.2. Preprocessing for human genome DNA sample

A single DNA molecule was separated to a droplet prior to performing digital PCR. When the target gene has multiple copies, they are closely connected on the same molecule, thus they behave as a single molecule and the target gene copy number is underestimated. Accordingly, preprocessing of the samples was performed using restriction enzymes (New England Biolabs Japan, Tokyo, Japan), with Nde I used for CYP2D6 and Pst I used for SULT1A1 (Figure 1). A 20-µL reaction mix containing 2 µL of 10x buffer, 4 IU of restriction enzymes, and 200 ng of genome DNA was utilized. Incubation was performed at 37°C for 120 minutes.

---

**Table 1. Sequences of primers and probes used for digital and long PCR**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2D6_Dup_F1 (38)</td>
<td>CTT CAC CTC CCT GCT GCA G</td>
</tr>
<tr>
<td>CYP2D6_Dup_R1 (38)</td>
<td>TCA CCA GGA AAG CAA AGA CA</td>
</tr>
<tr>
<td>CYP2D6_Dup_P1 (38)</td>
<td>FAM-CCG GCC CAG CCA CCC TGG-BHQ</td>
</tr>
<tr>
<td>SULT1A1_Dup_F1</td>
<td>AAA GGA TGT GGC AGT TTC CT</td>
</tr>
<tr>
<td>SULT1A1_Dup_R1</td>
<td>CAC ACT TTC CTT CCT CCC AT</td>
</tr>
<tr>
<td>SULT1A1_Dup_P1</td>
<td>FAM-CTC AGG GTG CAC CTT GGC CA-BHQ</td>
</tr>
<tr>
<td>RPPH1_F1</td>
<td>GGG AGG TGA GTT CCC AGA G</td>
</tr>
<tr>
<td>RPPH1_R1</td>
<td>CTT GAG TCT GTT CCA AGT TC</td>
</tr>
<tr>
<td>RPPH1_P1</td>
<td>HEX-CTG CCC AGT CTG ACC TCG CG-BHQ</td>
</tr>
<tr>
<td>P2=21 (39)</td>
<td>GCC ACC ATG GTG TCT TGT CTT TC</td>
</tr>
<tr>
<td>P2=22 (39)</td>
<td>ACC GGA TTC CAG CTT GGA AAT G</td>
</tr>
<tr>
<td>CYP13 (40)</td>
<td>ACC GGG CAC CTG TAC TCC TCA</td>
</tr>
<tr>
<td>CYP24 (40)</td>
<td>GCA TGA GCT AAG GCA CCC AGA C</td>
</tr>
</tbody>
</table>

FAM: 5-Carboxyfluorescein hydrate; HEX: 6-Carboxy-4,7,2',4',5',7'-hexachloro-fluorescein-3’,6’-dipivaloate; BHQ: Black Hole Quencher™.
2.3. Digital PCR for CYP2D6

Digital PCR for CYP2D6 was performed using specific primers and probes, as previously described by Qin J, et al. (38). Briefly, a 20-μL PCR mix containing 10 μL of 2× ddPCR supermix for probes, 0.625 μM each of primers for CYP2D6 and the reference gene (RPPH1), 0.5 μM each of the probes for those, and 20 ng of genome DNA cut with Nde I was prepared (Figure 1). The PCR mix was partitioned into a discrete water-in-oil emulsion using a QX200 Droplet Generator (Bio-Rad Laboratories, Inc. Tokyo, Japan) prior to the PCR assay. The thermal cycling conditions consisted of an initial denaturation step at 95°C for 10 minutes, followed by 45 cycles of denaturation at 94°C for 15 seconds, then annealing and extension at 60°C for 1 minute, with a final inactivation step performed at 98°C for 10 minutes. Following amplification, fluorescence of the samples was analyzed using a QX200 Droplet Reader (Bio-Rad Laboratories, Inc. Tokyo, Japan) to calculate the copy numbers.

2.4. Digital PCR for SULT1A1

We designed primers and probes specific for the SULT1A1 gene. A 20-μL PCR mix containing 10 μL of 2× ddPCR supermix for probes, 0.9 μM each of primers for SULT1A1 and RPPH1, 0.5 μM each of the probes for those, and 10 ng of genome DNA cut with Pst I was prepared (Figure 1). The PCR mix was partitioned into a discrete water-in-oil emulsion using a QX200 Droplet Generator prior to the PCR assay. Thermal cycling conditions consisted of an initial denaturation step at 95°C for 10 minutes, followed by 45 cycles of denaturation at 94°C for 15 seconds, then annealing and extension at 54°C for 1 minute, with a final inactivation step performed at 98°C for 10 minutes. Following amplification, fluorescence of the samples was analyzed using a QX200 Droplet Reader (Bio-Rad Laboratories, Inc. Tokyo, Japan) to calculate the copy numbers.

2.5. Detection of CYP2D6xN and CYP2D6*5

Detection of CYP2D6xN and CYP2D6*5 was performed using long PCR testing, as previously described by Johansson I, et al. (39) and Steen VM, et al. (40), respectively.

3. Results and Discussion

For analysis of CYP2D6, we categorized the samples into 3 groups (CYP2D6*5 hetero, non-CNV, CYP2D6xN) based on known genotypes and compared those with the digital PCR data obtained in the present study. The copy numbers of those were 1, 2, and 3 to 4, respectively, thus there were no contradictions (Figure 2). However, the primers used in this study targeted exon 9 of CYP2D6 and cannot be used to analyze copy numbers that have gene conversion to CYP2D7 in exon 9 (e.g. CYP2D6*36 or *36x2). Thus, it is necessary to design primers for other regions in CYP2D6.

For analysis of SULT1A1, the samples were categorized into 2 groups (non-CNV, and multiplication) based on known genotypes, then compared using our
digital PCR findings. Their copy numbers were 2, 3 to 6, respectively (Figure 3). In our previous study, we did not find samples with an SULT1A1 gene deletion (37). In the present study, none of the samples had 0 or 1 copy, thus there were no contradictions. In that previous study, we were only able to obtain ambiguous copy number results, which were noted as "4 or more", while the present method allowed us to more clearly determine copy numbers.

To the best of our knowledge, this is the first report of analyses of CYP2D6 and SULT1A1 CNVs using a digital PCR method with clinical samples. Copy numbers are primarily presented as a single integer, though some are shown as an ambiguous number (e.g., 3.4). We rounded off the numbers obtained and considered them as final copy numbers. We speculated that an imperfect reaction of the restriction enzymes was the primary reason of the ambiguous numbers, while use of clinical samples instead of cell lines may have also been related to that result.

CYP2D6 and SULT1A1 are involved in the main metabolism pathway of tamoxifen, with the former having effects on activation and the latter on inactivation. Accordingly, the effects and side effects of tamoxifen are increased in CYP2D6*N patients, whereas patients with SULT1A1 multiplication will not see sufficient effects from the drug.

Several reports regarding the relationships between the genotypes of enzymes that metabolize tamoxifen and effects of the drug have been presented. Xu Y, et al. (41) found that the CYP2D6*10 mutation had effects on tamoxifen efficacy in Chinese patients, while Gjerde J, et al. (42) conducted a study in Norway and reported that the CYP2D6 genotype influenced conversion of tamoxifen to potent hydroxylated metabolites. On the other hand, investigations performed by Wegman P, et al. (43) in Sweden and Okishiro M, et al. (44) in Japan found no relationship of CYP2D6 or SULT1A1 genotypes with survival time. Similarly, Lum DWK, et al. (45) analyzed the CYP2D6 genotype and tamoxifen response by meta-analysis, and found no association, while Motamedi S, et al. (46) performed a study in Iran, and reported that there was no significant relationship between CYP2D6 copy number and tamoxifen resistance in their patients.

Estimation of response to tamoxifen given for breast cancer is extremely difficult because CYP2D6 mutations include many variations and high complexity, and the drug is related to some genetic polymorphisms of enzymes such as SULT1A1. Few studies have performed analyses using a combination of multiple mutations, or combined genotyping and CNV analysis. We found that the present digital PCR method was able to calculate CYP2D6 and SULT1A1 copy numbers. In the future, a combined method for genotyping, distinguishing alleles, and calculating copy numbers will be helpful for estimating the effects and side effects of drugs such as tamoxifen.

References

15. Steen VM, Molven A, Aarskog NK, Gulbrandsen AK. Homologous unequal cross-over involving a 2.8 kb direct repeat as a mechanism for the generation of allelic


(Received October 27, 2017; Revised November 10, 2017; Re-revised November 12, 2017; Accepted December 13, 2017)