

# Determination of Cytochrome P450 2D6 (CYP2D6) Gene Copy Number by Real-Time Quantitative PCR

Laurent Bodin,<sup>1</sup> Philippe H. Beaune,<sup>1,2</sup> and Marie-Anne Lorient<sup>1,2</sup>

<sup>1</sup>INSERM UMRS 490, Université René Descartes - Paris 5, 45 rue des Saints-Pères, 75270 Paris, France

<sup>2</sup>Assistance Publique des Hôpitaux de Paris, Département de la Biochimie, Hôpital Européen Georges Pompidou, Paris, France

Received 23 November 2004; revised 17 March 2005; accepted 21 March 2005

Gene dosage by real-time quantitative PCR has proved to be accurate for measuring gene copy number. The aim of this study was to apply this approach to the CYP2D6 gene to allow for rapid identification of poor and ultrarapid metabolizers (0, 1, or more than 2 gene copy number). Using the  $2^{-\Delta\Delta C_t}$  calculation method and a duplex reaction, the number of CYP2D6 gene copies was determined. Quantitative PCR was performed on 43 samples previously analyzed by Southern blotting and long PCR including 20 samples with a heterozygous deletion, 11 with normal copy number (2 copies), and 12 samples with duplicated genes. The average ratio ranged from 1.02 to 1.28, 1.85 to 2.21, and 2.55 to 3.30, respectively, for the samples with 1 copy, 2 copies, and 3 copies. This study shows that this method is sensitive enough to detect either a heterozygous gene deletion or duplication.

## INTRODUCTION

Today a variety of assays are available for the quantification of CYP2D6 copy number including Southern blotting, long PCR, DNA invader system [1], and, more recently, TaqMan technology [2]. Southern blotting and long PCR require a large amount of genomic DNA and are costly in terms of reagents, work, and time. Moreover long PCR does not allow for the determination of the exact copy number. To date, real-time quantitative PCR has been widely used to quantify viral copy number, to perform gene expression studies, to diagnose genetic diseases, and to quantify gene copy number in transgenic animals, or to measure oncogene amplification in tumor cells [3]. Indeed, many gene deletions or duplications are responsible for genetic disorder and are now diagnosed by quantitative PCR, such as Charcot-Marie-Tooth I, hereditary neuropathy with liability to pressure palsies [4, 5], proximal spinal atrophy [6],  $\alpha$ -thalassemia [7], and down syndrome [8]. Moreover, gene

dosage of xenobiotic-metabolizing enzymes has already been applied for glutathione S-transferases [9].

Very recently Schaeffeler et al successfully developed TaqMan probes for the determination of CYP2D6 copy number [2]. Because of the larger number of allelic variants among CYP2D6 and the presence in tandem of two inactive pseudogenes (CYP2D8 and CYP2D7), the choice for the localization of the target sequence has become complex. In addition, it is necessary to take into account at the same time the sequence specificity and the constraints required for the TaqMan probes. Schaeffeler et al validated their technique by including 64 samples but with one limitation in the CYP2D6 probe. Thus, a SNP is located in the target sequence, which leads to an undetermined genotype among Asian subjects carrying the CYP2D6\*10 and \*36 alleles.

In this study, a new real-time duplex PCR has been developed, allowing for the determination of the CYP2D6 gene copy number for every sample since no SNP was described in our target sequence. First, TaqMan MGB probes were used to improve the specificity of the reaction. Second, the duplex strategy allowed for the coamplification of CYP2D6 and the referent gene (RNase P) in the same well, which increases reproducibility and decreases the analysis cost.

For this purpose, CYP2D6 copy number was determined by the comparative Ct method [10] in 43 subjects, distributed as follows: 12 duplicated, 20 deleted, and 11 with two copies. The determination of CYP2D6 copy number by real-time quantitative PCR was then validated from previous long PCR/Southern blotting results [11].

---

Correspondence and reprint requests to Philippe H. Beaune, INSERM UMRS 490, Université René Descartes - Paris 5, 45 rue des Saints-Pères, 75270 Paris, France, E-mail: philippe.beaune@univ-paris5.fr

This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

TABLE 1. CYP2D6 genotype of the samples included in this study.

CYP2D6 copy number	Number of samples	CYP2D6 genotype
1	9	CYP2D6*1/*5
	3	CYP2D6*2/*5
	3	CYP2D6*2B/*5
	3	CYP2D6*4A/*5
	2	Unknown <sup>a</sup>
2	3	CYP2D6*2/*2
	3	CYP2D6*2/*2B
	2	CYP2D6*2/*1
	2	CYP2D6*4 × 2/*5 <sup>b</sup>
	1	Unknown <sup>c</sup>
3 <sup>d</sup>	3	CYP2D6*1/*1
	3	CYP2D6*1/*2
	1	CYP2D6*1/*9
	1	CYP2D6*1/*4
	1	CYP2D6*1/*2d
	1	CYP2D6*1/*28
	1	CYP2D6*1/*2B
	1	Unknown <sup>c</sup>

<sup>a</sup>Carrier of heterozygous CYP2D6\*5 (deletion) and other uncharacterized alleles.

<sup>b</sup>Carrier of both deletion and duplication of CYP2D6 resulting in two final copies of the CYP2D6 gene.

<sup>c</sup>Undetermined genotype, Southern blot, and long PCR only.

<sup>d</sup>Samples with the CYP2D6\*2 × 2, corresponding to the duplication.

## MATERIALS AND METHODS

### DNA samples

DNA samples were obtained from a case/control study where CYP2D6 gene deletion and/or duplication had been previously investigated by long PCR and Southern blotting [11]. DNA was prepared using standard methodology with proteinase K, extracted by phenol-chloroform, and precipitation with absolute ethanol [11]. DNA samples were stored at  $-20^{\circ}\text{C}$ . The DNA concentration was determined by optical density at 260 nm. All samples (except 4) were previously genotyped for CYP2D6 alleles [12] (Table 1). However for these DNA samples, Southern blotting analysis and long PCR were available. For the definition of alleles, see <http://www.imm.ki.se/CYPalleles/cyp2d6.htm>.

### Sequencing

Sequencing was performed on the ABI Prism 9700 Genetic Analyzer System (Applied Biosystems, Courtabœuf, France). Five  $\mu\text{L}$  of purified PCR product (enzymatic purification with ExoSap kit, Amersham Biosciences, Orsay, France) was mixed with 5  $\mu\text{L}$  of sequence mix containing 1  $\mu\text{L}$  of the same forward or reverse primers used for PCR and 4  $\mu\text{L}$  of BDT (DNA sequencing

kit, big dye terminator cycle sequencing; Applied Biosystems). A sequence reaction was then performed. Sequencing products were purified before analysis as indicated by the manufacturer.

### Probe and primer design

Genomic sequences of CYP2D6 (GenBank accession no M33388) and CYP2D7/8 (GenBank accession no M33387) were used to generate the specific target sequence [13]. Primer Express software version 1.0 (Applied Biosystems) was used with manual adjustment to design a set of primers and probe which are specific to CYP2D6 gene. For the detection of gene copy number, the position and length of hybridization probes and primers were optimized by using BLAST program and TaqMan MGB probes. Sequences for the TaqMan MGB probe and PCR primers are listed in Figure 1. The size of the PCR product for CYP2D6 was 78 bp. Primers and probe for the RNase P were obtained as TaqMan endogenous control kits (Applied Biosystems) (sequence and size of PCR product not provided by the manufacturer). Finally, the CYP2D6 probe was labeled at the 5' end with FAM, the RNase P was labeled at 5' end with VIC, and both probes had a nonfluorescent quencher.

### Quantitative real-time PCR conditions

Gene dosage analysis was performed in a quantitative duplex PCR assay. The following reagents were used for amplification in 20  $\mu\text{L}$  : 5  $\mu\text{L}$  of DNA (around 100 ng), 300 nM of each CYP2D6 primer, 10  $\mu\text{L}$  of Master Mix (2X), 250 nM of CYP2D6 probe, 1  $\mu\text{L}$  of RNase P kit (20X) and final volume was adjusted with sterile water. The thermal cycling conditions were first 2 minutes at  $50^{\circ}\text{C}$  and 10 minutes at  $95^{\circ}\text{C}$  and consisted of 40 cycles at  $95^{\circ}\text{C}$  for 15 seconds and  $60^{\circ}\text{C}$  for 1 minute. The PCR was performed in a 384-well clear optical reaction plate 7900 (Applied Biosystems).

### Gene dosage analysis

The RNase P gene was coamplified with the CYP2D6 gene and served as an internal standard (two copies present in subject with or without duplication/deletion of the CYP2D6 gene). The strategy is based on the relative quantification of the target sequence (CYP2D6) and a reference sequence (RNase P) and requires that the PCR efficiencies of the CYP2D6 gene and the RNase P PCR reactions be close to the identical. For TaqMan PCR calibration, we selected DNA samples from six control individuals with two copies of the CYP2D6 (ie, individuals previously characterized by Southern blotting and long PCR). During the log-linear phase, amplification can be described by  $N = \text{No}(1 + E)^n$ , where  $N$  is the number of amplified molecules,  $\text{No}$  is the initial number of molecules,  $E$  is the amplification efficiency, and  $n$  is the number of cycles. If the amplification

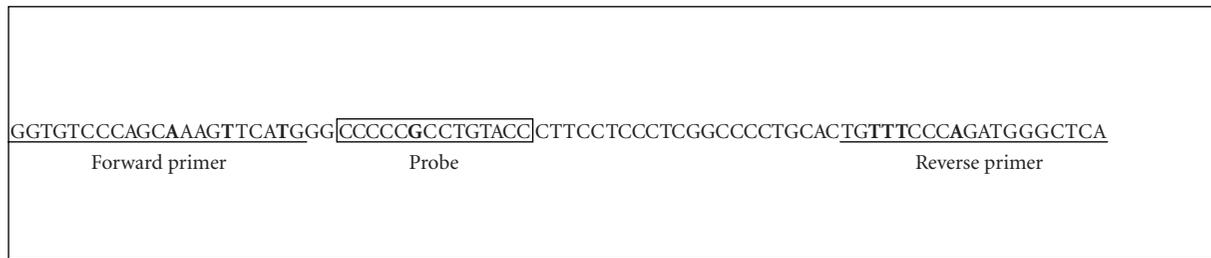


FIGURE 1. Probe and primer design. The forward and reverse primers, as well as the probe, were drawn using the Primer Express software. The target sequence is located in intron 6. The forward primer has a temperature of 58°C for a length of 19 bases including 4 discriminating ones (in bold) for the CYP2D6 gene related to the CYP2D7/8 sequences; the reverse primer has a temperature of 58°C for a length of 21 bases, including 3 nucleotides specific to the CYP2D6 gene (in bold). The MGB probe corresponds to the sequence framed on the diagram with one discriminating nucleotide (in bold). The G base (in bold) is another discriminating nucleotide present in the amplicon.

efficiency is similar for the two reactions, the initial concentration of the sample is calculated on the basis of the above formula by using the comparative delta Ct method as described by Livak et al [10] and the gene copy number was given by the formula  $2^{-(\Delta\Delta Ct + / - SD)}$ , where  $\Delta\Delta Ct = (Ct \text{ RNase P}_{\text{calibrator}} - Ct \text{ CYP2D6}_{\text{calibrator}}) - (Ct \text{ RNase P}_{\text{sample}} - Ct \text{ CYP2D6}_{\text{sample}})$ . A Ct is defined as the point at which the fluorescence level rises above a baseline. The effect of DNA concentration on PCR efficiency was determined using a control DNA in a dilution series from 25, 12.5, 6.25, and 3.125 ng per reaction. The primer concentration for the CYP2D6 gene was titrated from a 3-by-3 primer matrix (eg, combinations of 100, 300, and 900 nM of each forward and reverse primer).

All the experiments, which involved 11 samples with 2 copies of CYP2D6, 20 with a heterozygous deletion, and 12 with a duplication, were performed three times, and in each experiment the samples were analyzed in triplicate.

Data were collected and analyzed using sequence detection system software (version 2.0; Applied Biosystems). Amplification plots were reviewed after performing and the baseline setting was adjusted manually between 3 to 18 cycles. Reaction wells with PCR reaction failure were excluded from further data analysis according to Grubbs's test. Thus, the analysis was performed with at least two amplification results (duplicate).

## RESULTS AND DISCUSSION

### Assessment of assay specificity

The first step of this study was to design primers and probe highly specific to the CYP2D6 gene and to avoid coamplification of pseudogenes. The PCR MGB probes used here allow for a discriminative quantification of the CYP2D6 gene related to the CYP2D7/8 gene. Because of their higher melting temperature, they are more sensitive than conventional probes, that is, they create less

background noise and offer greater specificity. The region that was chosen for the target sequence was located at intronic sequence of the CYP2D6 gene (intron 6). No SNP has been reported to be located within this region according to the databases (NCBI). Previously, Schaeffeler et al designed CYP2D6 probe that cannot be used with Asian subjects carrying the CYP2D6\*10 and \*36 alleles. The method that is presented here is an alternative method for the determination of CYP2D6 copy number. Thus, the above indeterminations was avoided. The availability of two validated methods would be useful for the routine determination of CYP2D6 gene copy number especially when checking that a result is necessary or when genotyping Asian individuals.

Specificity for CYP2D6 amplification depends on the sequences of the two primers and the probe, which contained 7 and 1 discriminating nucleotides, respectively (Figure 1). The amplification product was sequenced. Moreover, to assess the specificity of the reaction, an assay was performed with a DNA sample from a subject homozygous for the CYP2D6 deletion gene (data not shown). The copy number of CYP2D6 calculated from the difference in Ct values between CYP2D6 and RNase P corresponded, for this individual, to a value lower than 0.05 copies. These results demonstrated that both sets of primers and probe were highly specific to the CYP2D6 gene.

The sequences of CYP2D6 and RNase P gene were coamplified within the same tube with variable amounts of genomic DNA (25, 12.5, 6.25, and 3.125 ng, resp). When the Ct values of CYP2D6 and RNase P were plotted versus DNA amounts, the PCR efficiency could be calculated for each gene in the multiplex based on the slope of the lines using the formula  $10^{-1/\text{slope}} - 1$ . The calculated efficiencies were similar, yielding to values of 0.82 and 0.80, respectively, for CYP2D6 and RNase P, indicating that the  $2^{-\Delta\Delta Ct}$  calculation was valid. Moreover, no variations were observed in the difference of Ct values between the target gene and the reference gene, demonstrating that the DNA

TABLE 2. Global analysis of CYP2D6 gene quantification based on three independent experiments. The analysis of gene copy number was performed using the formula  $2^{2^{-\Delta\Delta Ct} \pm SD}$ . SD stands for standard deviation and CV coefficient of variation. The three groups are significantly different with  $P < .05$  (Mann-Whitney U test). Means are obtained from the averages of the final results of each run after  $2^{2^{-\Delta\Delta Ct} \pm SD}$  calculation. The borderlines are well delimited in I, II, and III. Means are given by the averages of the results from each group. Range is defined by the lower and upper limits for each group. SD is the average of sd of the mean of the triplicate.

Number of patients	Expected copy number*	Global mean value of CYP2D6 copy**	Number of assay	Mean value $\pm$ SD of CYP2D6 copy**	Range
20	1	1.14	I	1.10 $\pm$ 0.18	0.76–1.57
			II	1.20 $\pm$ 0.12	1.04–1.47
			III	1.12 $\pm$ 0.10	0.95–1.33
11	2	1.99	I	1.87 $\pm$ 0.17	1.60–2.17
			II	2.09 $\pm$ 0.18	1.82–2.45
			III	2.03 $\pm$ 0.17	1.86–2.31
12	3	2.86	I	2.87 $\pm$ 0.39	2.38–3.57
			II	2.91 $\pm$ 0.23	2.64–3.26
			III	2.80 $\pm$ 0.30	2.41–3.40

\* The values obtained after Southern blotting and long PCR.

\*\* The values obtained calculated from real-time quantitative PCR.

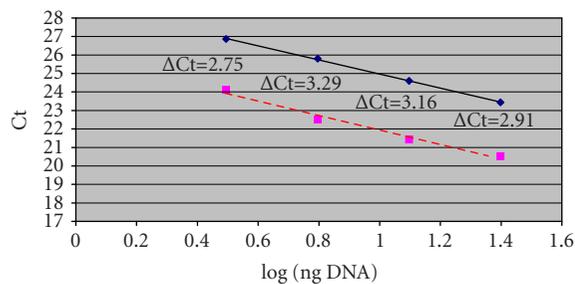


FIGURE 2. Multiplex and the  $2^{-\Delta\Delta Ct}$  method validation: determination and comparison of the PCR efficiency for each gene. The straight side (dotted line) of the PCR of the referent gene (RNase P) with a slope =  $-3.9$  (eg,  $E = 80\%$ ). The straight side (continuous line) of the PCR of the CYP2D6 gene with a slope =  $-3.8$  (eg,  $E = 82\%$ ). After optimizing in simplex (matrix of primers) and carrying out a limitation of primers, this graph shows, on the one hand, that the effectiveness of each PCR is practically identical and, on the other hand, that the use of the  $\Delta\Delta Ct$  calculation method is validated. Variable amounts of genomic DNA, 25, 12.5, 6.25, and 3.125 ng, respectively, were used to draw these curves.

concentration was not a critical parameter in this method (multiplex validation) (Figure 2).

### Validation of the quantitative PCR

In each experiment, the quantitative results of real-time PCR agreed with those of the reference methods (Southern blotting and long PCR) from 11 samples with two copies (including the six samples used as calibrators), 20 samples with 1 copy, and 12 samples with 3 copies. No overlap between the value range corresponding to each group (1, 2, 3 copies of CYP2D6 gene) was observed. Table 2 includes the average of three independent assays, where the copy number of the deleted, normal, and duplicated

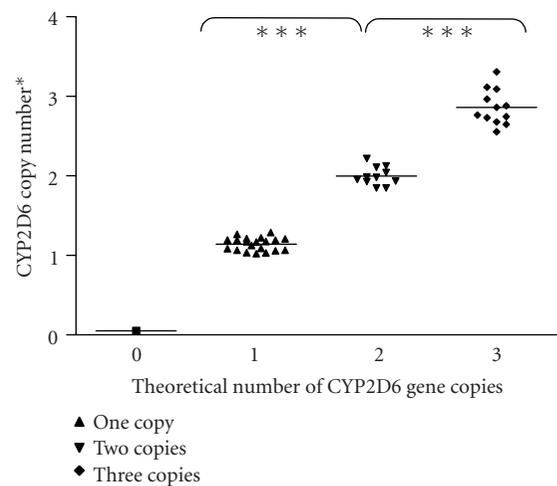


FIGURE 3. Global analysis of CYP2D6 gene quantification in 44 DNA samples based on three independent experiments. The analysis of gene copy number was performed using the formula  $2^{2^{-\Delta\Delta Ct} \pm SD}$ . The three groups are significantly different with  $P < .001$  (ANOVA). Means are realized from the averages of the final results of each run after  $2^{2^{-\Delta\Delta Ct} \pm SD}$  calculation. \* denotes gene copy number after Southern blotting and long PCR analysis.

samples was significantly separated with  $P < .05$ . As shown in Figure 3 the mean values of gene quantification data into the three groups 1, 2, and 3 of CYP2D6 gene were 1.14 (range: 1.02–1.29), 1.99 (range: 1.84–2.21), and 2.86 (range: 2.55–3.30), respectively, in the subjects carrying 1, 2, or 3 copies of the CYP2D6 gene (resp,  $P < .001$ , ANOVA). These experiment results could be interpreted properly because boundaries did not overlap. Figure 4 illustrates representative examples of real-time PCR results for CYP2D6 quantification.

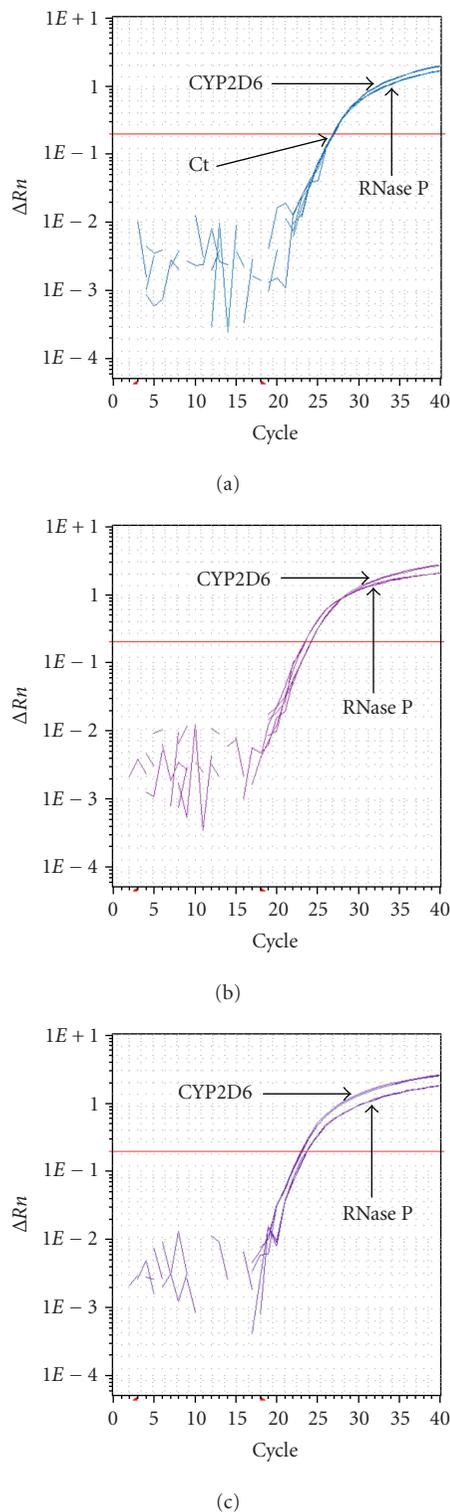


FIGURE 4. Amplification plots of CYP2D6 and RNase P genes. The  $\Delta Rn$  (reflecting the quantity of fluorescent probe degraded) is plotted versus cycle number. (a) Samples with 2 copies of CYP2D6. In (c) samples with a duplication, the Ct value of CYP2D6 decreased compared to RNase P whereas in (b) samples with deletion, the Ct value of CYP2D6 increased compared to RNase P.

### Variability of the quantitative PCR

The reproducibility of these experiments is reflected by the coefficient of variation (CV). All data were reported as the mean, SD, and CV of the average from the results of three different assays. The intra-assay reproducibility was determined from triplicate measurements within the same run and the mean of CV for the three experiments was comprised between 7.1% and 8.9% (data not shown). The interassay variation ranged from 1.50% to 22% for the heterozygous deletion samples, from 2.6% to 19% for the 2 copy samples, and from 3% to 15% for the duplicated samples indicating an acceptable level of reproducibility and allowing for discrimination between the different groups. However, a careful examination of the results showed that the mean of gene dosage for the three groups was distinct although some values could be close. Indeed, in the first assay (experiment I in Table 2), the upper range of the one-gene copy group is very close to the lower range of the two-gene copy group (1.57 and 1.60, resp). For these results, the intra-assay was over 20%, strongly suggesting the necessity for low intra-assay variation to avoid misclassification and minimize indeterminism.

### CONCLUSION

Compared to other methods (Southern blotting, long PCR), quantitative PCR has several advantages: (i) only a small amount of DNA is needed for this assay and a large number of samples can be analyzed simultaneously; (ii) the results can be obtained quickly and submitted to statistical analysis (the use of replicates is needed to increase the power assay); (iii) based on the fluorescence appearance during the exponential phase of the amplification, quantification is also more accurate; (iv) real time PCR assays obviate post-PCR manipulation preventing carry-over contaminations; (v) the multiplex reduces the cost of the analysis compared to a single-well assay [14]; and (vi) when reactions are carried out in 384-well plates, the reaction volume can be reduced ( $15 \mu\text{L}$ ) decreasing the reagent cost [5].

The rapid quantification of the CYP2D6 copy number has become a significant investment as regards the large number of drugs metabolized by the CYP2D6 gene. This new real-time quantitative PCR could be adapted for high-throughput determination of CYP2D6 genotyping.

### ACKNOWLEDGMENTS

We thank Dr Isabelle Stücker and Professor Franck Broly for providing DNA samples and Dr Michel Barrois for helpful discussions. This work was supported by Région Ile de France, SESAME. Laurent Bodin received financial support from the French "Ministère de la Recherche."

## REFERENCES

- [1] Nevilie M, Selzer R, Aizenstein B, et al. Characterization of cytochrome P450 2D6 alleles using the Invader system. *Biotechniques*. 2002;(suppl):34–38, 40–43.
- [2] Schaeffeler E, Schwab M, Eichelbaum M, Zanger UM. CYP2D6 genotyping strategy based on gene copy number determination by TaqMan real-time PCR. *Hum Mutat*. 2003;22(6):476–485.
- [3] Bieche I, Olivi M, Champeme MH, Vidaud D, Lidereau R, Vidaud M. Novel approach to quantitative polymerase chain reaction using real-time detection: application to the detection of gene amplification in breast cancer. *Int J Cancer*. 1998;78(5):661–666.
- [4] Aarskog NK, Vedeler CA. Real-time quantitative polymerase chain reaction. A new method that detects both the peripheral myelin protein 22 duplication in Charcot-Marie-Tooth type 1A disease and the peripheral myelin protein 22 deletion in hereditary neuropathy with liability to pressure palsies. *Hum Genet*. 2000;107(5):494–498.
- [5] Thiel CT, Kraus C, Rauch A, Ekici AB, Rautenstrauss B, Reis A. A new quantitative PCR multiplex assay for rapid analysis of chromosome 17p11.2-12 duplications and deletions leading to HMSN/HNPP. *Eur J Hum Genet*. 2003;11(2):170–178.
- [6] Anhuf D, Eggermann T, Rudnik-Schoneborn S, Zerres K. Determination of SMN1 and SMN2 copy number using TaqMan technology. *Hum Mutat*. 2003;22(1):74–78.
- [7] Chan V, Yip B, Lam YH, Tse HY, Wong HS, Chan TK. Quantitative polymerase chain reaction for the rapid prenatal diagnosis of homozygous alpha-thalassaemia (Hb Barts hydrops fetalis). *Br J Haematol*. 2001;115(2):341–346.
- [8] Hu Y, Zheng M, Xu Z, Wang X, Cui H. Quantitative real-time PCR technique for rapid prenatal diagnosis of Down syndrome. *Prenat Diagn*. 2004;24(9):704–707.
- [9] Shi MM, Myrand SP, Bleavins MR, de la Iglesia FA. High-throughput genotyping method for glutathione S-transferase T1 and M1 gene deletions using TaqMan probes. *Res Commun Mol Pathol Pharmacol*. 1999;103(1):3–15.
- [10] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  method. *Methods*. 2001;25(4):402–408.
- [11] Stucker I, Cosme J, Laurent P, et al. CYP2D6 genotype and lung cancer risk according to histologic type and tobacco exposure. *Carcinogenesis*. 1995;16(11):2759–2764.
- [12] Legrand-Andreolletti M, Stucker I, Marez D, et al. Cytochrome P450 CYP2D6 gene polymorphism and lung cancer susceptibility in Caucasians. *Pharmacogenetics*. 1998;8(1):7–14.
- [13] Kimura S, Umeno M, Skoda RC, Meyer UA, Gonzalez FJ. The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene. *Am J Hum Genet*. 1989;45(6):889–904.
- [14] Olney RC, Mougey EB, Wang J, Shulman DI, Sylvester JE. Using real-time, quantitative PCR for rapid genotyping of the steroid 21-hydroxylase gene in a north Florida population. *J Clin Endocrinol Metab*. 2002;87(2):735–741.