Detection of epidermal growth factor receptor mutation in lung cancer by droplet digital polymerase chain reaction

Qing Xu1,*, Yazhen Zhu2,*, Yali Bai1, Xiumin Wei1, Xinrun Zheng2, Mao Mao1, Guangjuan Zheng2

1Translational Bioscience and Diagnostics, WuXi AppTec, Shanghai, 2Department of Pathology, Guangdong Provincial Hospital of TCM, Guangzhou University of Chinese Medicine, Guangdong Provincial Academy of Chinese Medical Sciences, Guangzhou, People’s Republic of China

*These authors contributed equally to this work

Background: Two types of epidermal growth factor receptor (EGFR) mutations in exon 19 and exon 21 (ex19del and L858R) are prevalent in lung cancer patients and sensitive to targeted EGFR inhibition. A resistance mutation in exon 20 (T790M) has been found to accompany drug treatment when patients relapse. These three mutations are valuable companion diagnostic biomarkers for guiding personalized treatment. Quantitative polymerase chain reaction (qPCR)-based methods have been widely used in the clinic by physicians to guide treatment decisions. The aim of this study was to evaluate the technical and clinical sensitivity and specificity of the droplet digital polymerase chain reaction (ddPCR) method in detecting the three EGFR mutations in patients with lung cancer.

Methods: Genomic DNA from H1975 and PC-9 cells, as well as 92 normal human blood specimens, was used to determine the technical sensitivity and specificity of the ddPCR assays. Genomic DNA of formalin-fixed, paraffin-embedded specimens from 78 Chinese patients with lung adenocarcinoma were assayed using both qPCR and ddPCR.

Results: The three ddPCR assays had a limit of detection of 0.02% and a wide dynamic range from 1 to 20,000 copies measurement. The L858R and ex19del assays had a 0% background level in the technical and clinical settings. The T790M assay appeared to have a 0.03% technical background. The ddPCR assays were robust for correct determination of EGFR mutation status in patients, and the dynamic range appeared to be better than qPCR methods. The ddPCR assay for T790M could detect patient samples that the qPCR method failed to detect. About 49% of this patient cohort had EGFR mutations (L858R, 15.4%; ex19del, 29.5%; T790M, 6.4%). Two patients with the ex19del mutation also had a naïve T790M mutation.

Conclusion: These data suggest that the ddPCR method could be useful in the personalized treatment of patients with lung cancer.

Keywords: ddPCR, companion diagnostics, L858R, ex19del, T790M

Introduction

Lung cancer is a leading cause of cancer mortality worldwide. More than 85% of lung cancers are non-small cell lung cancer. Adenocarcinoma and squamous cell carcinoma are the two major subtypes of non-small cell lung cancer, comprising 50% and 40% of cases, respectively.1 Genomic studies such as next-generation sequencing and other high-throughput profiling platforms have identified genetic alterations common to both adenocarcinoma and squamous cell carcinoma, including in KRAS, BRAF, epidermal growth factor receptor (EGFR), and ALK.1 Among these mutations or fusion genes, alterations in the EGFR have been identified in about 12%–13% of cases examined so far worldwide,1,2 and are much more frequent in Asian populations (~50%).1 EGFR is a member of the ERBB family of receptor tyrosine kinases involved...
in the pathogenesis of many malignant cancers. Abnormal activity of *EGFR* results in deregulated cell proliferation and growth, which makes *EGFR* an important drug target in cancer cells.\(^5,6\) In fact, small-molecule tyrosine kinase inhibitors (TKIs), such as gefitinib and erlotinib, have been shown to significantly prolong the survival of lung cancer patients with *EGFR*-mutant tumors.\(^5,6\)

Among the TKI-sensitive mutations in *EGFR*, a point mutation in exon 21, which substitutes an arginine for a leucine (L858R), and a small in-frame deletion in exon 19, which removes several amino acids, are the most common activating mutations that confer treatment benefits (together accounting for approximately 90% of TKI-sensitive mutations).\(^5,6\) Despite the therapeutic benefits of *EGFR* TKIs, about 50% of patients with *EGFR*-mutant tumors develop acquired resistance via a second-site mutation in the threonine gatekeeper residue at position 790, ie, T790M.\(^3,6\) Hence, the three mutation biomarkers can be used as genetic determinants for decision-making about treatment and monitoring the treatment course.

In this study, the *EGFR* mutation status (L858R, ex19del, or T790M) of clinical samples from patients with lung adenocarcinoma was evaluated by two methods, ie, quantitative polymerase chain reaction (qPCR) and droplet digital polymerase chain reaction (ddPCR). On the basis of the differences observed between the two platforms, we consider that the ddPCR method can accurately determine *EGFR* mutation status without using standard curves or internal controls, has a wider dynamic range, and has the advantage of absolute quantification, thereby alleviating diagnostic ambiguity in the clinical setting.

**Materials and methods**

**Cell lines**

*EGFR* wild-type A549 cells, *EGFR*-L858R and *EGFR*-T790M mutation-carrying H1975 cells, and *EGFR* ex19del (Glu746-Ala750) mutation-carrying PC9 cells were obtained from the American Type Culture Collection (Manassas, VA, USA). Cells were cultured in their respective medium in T150 flasks until confluency and harvested with trypsin. Cell pellets were used to prepare genomic DNA using the QIAamp\(^\text{®}\) DNA mini kit (Qiagen, Valencia, CA, USA) following the manufacturer’s instructions.

**Normal genomic DNA**

Human blood was collected from healthy volunteers using the standard venipuncture procedure. Blood samples were centrifuged in BD tubes (BD Biosciences, Franklin Lakes, NJ, USA) at 2,500\(^\times\) g for 10 minutes at room temperature. The plasma portion was carefully pipetted into a new 15 mL tube and stored at −80°C. The buffy coat and red cells were lysed in red blood cell lysis buffer. The white cell pellets were resuspended in phosphate-buffered saline and genomic DNA was prepared using the QIAamp DNA mini kit. The concentrations of all the DNA samples were assessed using a NanoDrop 8000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

**Patient samples**

Lung tumor specimens were obtained from Guangdong Provincial Hospital of Traditional Chinese Medicine with previous fully informed consent from the patients and with approval from the hospital. The formalin-fixed, paraffin-embedded (FFPE) sections were used to prepare DNA using a QIAamp DNA FFPE tissue kit (Qiagen).

**Real-time PCR**

Quantitative PCR was conducted with a human *EGFR* gene mutation detection kit (PCR fluorescence probe) kit (YQ Biomed, Shanghai, People’s Republic of China) according to the manufacturer’s instructions. Real-time PCR was performed using a ViiA\(^\text{™}\) 7 instrument (Life Technologies, Carlsbad, CA, USA). PCR conditions were 94°C for 3 minutes to activate the DNA polymerase followed by 40 cycles at 94°C for 15 seconds and 40 cycles at 60°C for 1 minute. The difference in the cycle threshold (Ct) value of the mutant versus internal control (ΔCt) was used to determine the mutation status of each sample.

**Droplet digital PCR**

Design of the primers and probes for L858R, ex19del, and T790M was based on Oxnard et al.\(^10\) The duplex probes for ex19del detection cover the K749-S752 region and the A755-D761 region, respectively, and in principle any deletion in the two areas can be detected. PCR primers and TaqMan probes (FAM-labeled or VIC-labeled) were obtained from Life Technologies. Next, 900 nM probes and 250 nM primers were mixed with 2× Droplet PCR Supermix (Bio-Rad Laboratories, Hercules, CA, USA), 4 µL of template DNA, and H\(_2\)O to become 20 µL for each reaction. The reaction mixture was applied to the droplet generator with 70 µL of oil to form droplets in approximately 35 µL of oil-in-water mixture. The mixture was transferred to a 96-well PCR plate and heat-sealed. The plate was placed in a C1000 Touch thermal cycler (Bio-Rad Laboratories), and amplified to the endpoint PCR. Thermal cycling conditions were 95°C ×10 minutes (one
cycle), 40 cycles at 94°C for 30 seconds, 62°C × 1 minute, and 98°C × 10 minutes (one cycle), with a ramp rate of 1°C per second. After PCR, the 96-well PCR plate was read on a QX-200 droplet reader (Bio-Rad Laboratories). The data were analyzed with Quantalife software. Briefly, a threshold line was drawn for channel 1 and channel 2, respectively, to separate the two clusters of negative and positive droplets. The threshold line was determined by the positive control samples, ie, genomic DNA from H1975 or PC-9. The Quantalife software calculates the copy number of both mutant and wild-type DNA based on the Poisson statistic.

Results
Sensitivity and specificity of detection of EGFR mutation by ddPCR
To determine the limit of detection of mutant DNA on the confounding background of wild-type DNA, we serially diluted 20,000 copies of H1975 or PC9 genomic DNA into 20,000 copies of normal human blood DNA (about 66 ng) to one copy, and measured the actual number of copies using ddPCR. All three assays showed good linearity ($R^2 > 0.999$) and a wide dynamic range (1–20,000 copies, Figure 1). According to the statistical considerations for rare detection experimental design from the Droplet Digital™ PCR Applications Guide (Bio-Rad Laboratories), to guarantee with 95% confidence that at least one mutant molecule will be screened in the sample, statistics dictate that at least three times the number of expected wild-type target molecules must be screened (ie, number of molecules to screen = 3× background molecules). The sensitivity of these assays can be calculated to be 0.02% based on the formula (these detection limits are comparable with other digital methods, such as BEAMing [0.01%], but better than qPCR-based methods such as the Scorpion amplification refractory mutation system [0.1%], wild-type quenching PCR [0.1%], and the Sequenom MassARRAY platform [0.3%]).

To determine the background level of the ddPCR assays, we measured mutant copy numbers in A549 cells and in 92 normal human blood genomic DNA samples from healthy volunteers. No mutant copies were detected in 20,000 copies of A549 genomic DNA. For both the L858R and ex19del assays, no mutant copies were detected in 20,000 copies of

![Figure 1](https://www.dovepress.com/)

**Figure 1** Limit dilution test to determine the sensitivity of the three EGFR mutation detection assays by the ddPCR approach.

**Notes:** (A) L858R assay carried out using 66 ng of H1975 genomic DNA diluted in 20,000 copies of normal human genomic DNA. (B) The ex19del assay carried out using 66 ng of PC9 genomic DNA diluted in 20,000 copies of normal human genomic DNA. (C) T790M assay carried out using 66 ng of H1975 genomic DNA diluted in 20,000 copies of normal human genomic DNA. The lowest copy detected for all three assays was 1.4, and thus the limit of detection was 0.015% considering the sampling errors.

**Abbreviations:** EGFR, epidermal growth factor receptor; ddPCR, digital droplet polymerase chain reaction.
Twelve of the samples were detected as L858R-positive, and 23 were detected as ex19del-positive (Table 2). Interestingly, two patient samples that were positive for the ex19del allele were also positive for T790M mutation (data not shown), consistent with the observation of naive mutation of T790M in patients with lung cancer.3

To determine the background level of ddPCR assays in the clinical samples, we compared samples with known L858R and ex19del mutation status based on their mutual exclusivity. For the L858R assay, no copies of mutant alleles were detected in any of the 23 ex19del-positive samples (Figure 2A); and for the ex19del assay, no copies of mutant alleles were detected in any of the 12 L858R-positive samples (Figure 2B). This result is consistent with the high specificity result obtained with 20,000 copies of normal human blood genomic DNA described above, suggesting that the two assays were highly specific in the FFPE specimens.

An advantage of ddPCR is that the allele frequency of the mutant allele can be determined quantitatively by measuring both mutant and wild-type alleles in a single PCR reaction. The allele frequency can be calculated as the mutant allele copy number divided by the sum of the mutant allele copy number and the wild-type copy number. The lowest allele frequency in the L858R sample set was 4% (Figure 3A), whereas in the ex19del sample set it was 18% (Figure 3B). There appeared to be no correlation between mutant allele frequency and wild-type copy number for each sample in both the L858R and ex19del assays, with a Pearson correlation of $R^2=0.05$ and $R^2=0.4$, respectively (Figure 3). Both sample sets contained samples with allele frequencies greater than 50%, suggesting that there may be amplifications of mutant alleles in these samples or that the tumor cell content was very high in the sample (eg, sample 15 in Figure 3B). In addition, the wild-type copy number may serve as a quality control for the individual testing sample.

In addition to the two positive T790M samples, we have observed eight additional samples with >0 copies of mutant T790M alleles. Using the threshold line of six mutant copies out of 20,000 wild-type copies defined by the normal blood

| Table 1 Clinical characteristics of patients |
|-----------------|-----------------|
| Characteristics | Patients (n) |
| Sex             |                 |
| Male (%)        | 43 (55)         |
| Female (%)      | 35 (45)         |
| Age, years      |                 |
| Total median (range) | 62.5 (37–91) |
| Male median (range)  | 66 (37–84)     |
| Female median (range) | 57 (37–91)     |
| Smoking status (female) |          |
| Smoker          | 34 (4)          |
| Never-smoker    | 40 (30)         |
| N/A*            | 4 (1)           |
| Pathology       |                 |
| Adenocarcinoma  | 78              |
| Stage           |                 |
| I–III           | 27              |
| IV              | 51              |
| EGFR mutation by ddPCR (L858R and ex19del) (smoker/non-smoker) | |
| Male            | (5/6)           |
| Female          | (3/21)**        |

Notes: *N/A, smoking status of patient is uncertain; **P=0.03 by $\chi^2$ test.

Abbreviations: ddPCR, digital droplet polymerase chain reaction; EGFR, epidermal growth factor receptor.

normal human DNA, suggesting both assays are highly specific. In the T790M assay, we detected 1–5.2 mutant copies in 15 of 92 normal samples. Using the mean plus 3 standard deviations as the cutoff, the threshold line for false positive calls can be determined to be six copies when measuring 20,000 copies of input DNA (ie, 0.03%).

Detection of EGFR mutation status in clinical specimens

To determine if ddPCR assays can be used to detect EGFR mutations in patients with lung cancer, we screened 78 samples from resected FFPE lung tissues. The characteristics of the patients are summarized in Table 1. All the samples were categorized as adenocarcinoma. The patients were balanced in terms of sex and smoking status. The median age of the female patients was significantly lower than that of the male patients (Mann–Whitney U-test, one-tailed, $P=0.023$). About two thirds of the patients had advanced (stage IV) lung cancer.

| Table 2 Concordance of L858R, ex19del, and T790M mutation detection by ddPCR with qPCR |
|-----------------|-----------------|-----------------|
| ddPCR | L858R | ex19del | T790M |
| Positive | Negative | Total | Positive | Negative | Total | Positive | Negative | Total |
| Positive | 12 | 0 | 12 | 23 | 0 | 23 | 2 | 3 | 5 |
| Negative | 0 | 66 | 66 | 0 | 55 | 55 | 0 | 73 | 73 |
| Total | 12 | 66 | 78 | 23 | 55 | 78 | 2 | 76 | 78 |

Abbreviations: ddPCR, digital droplet polymerase chain reaction; qPCR, quantitative polymerase chain reaction.
genomic DNA (or 0.03%), we could rule out three samples as false positives (data not shown). When examining the two-dimensional plot of the data, we found that two samples only showed droplets in the second quadrant (Figure 4A). Based on the sample partitioning principle of ddPCR, it is not possible that these droplets are from a true mutant allele. Instead, they are most likely from cross-hybridization of the mutant probe with the wild-type DNA in the sample. We then repeated the remaining three samples and found that they consistently showed a mutant allele frequency >0.03% (Figure 4B). When examining the two-dimensional plot, they showed positive droplets mainly in the first quadrant (Figure 4C). These three samples, although with a low mutant allele frequency, are very likely to be true positives.

Figure 2 Background of L858R and ex19del assays in formalin-fixed, paraffin-embedded genomic DNA.
Notes: (A) Twenty-three samples positive in ex19del showed no copies of the L858R allele. (B) Twelve samples positive in L858R showed no copies of the ex19del allele.

Figure 3 WT copy number and mutant allele frequency distribution of L858R and ex19del assays.
Notes: The WT copy number was obtained directly from each assay reaction readout. The frequency of the mutant allele was calculated by the copy number of the mutant allele/(copy number of mutant allele + copy number of WT allele). (A) L858R assay and (B) ex19del assay.
Abbreviations: WT, wild-type; mt, mutant.
To understand how accurately ddPCR methods can capture mutation status in patients with lung cancer, we compared the results with qPCR results previously obtained in the pathology laboratory when patients were diagnosed with lung cancer. The qPCR data were obtained with commercial kits based on the Scorpion amplification refractory mutation system method.\textsuperscript{12,15} The L858R and ex19del assays showed 100% concordance with qPCR results in detecting positive and negative samples; however, three samples determined as positive in the ddPCR assay results for T790M were negative in the qPCR assay results (Table 2). The concordance rate of the T790M assay was 96%. These data suggest that the ddPCR assay for T790M is more sensitive than the qPCR assay in clinical samples.

**Figure 4** Determination of true positives for T790M in formalin-fixed, paraffin-embedded genomic DNA. 
**Notes:** (A) Example of two-dimensional plot of a false positive sample in the T790M assay using formalin-fixed, paraffin-embedded genomic DNA. The positive signal comes from the second quadrant only. (B) Three reproducible samples showing allele frequency above the 0.03% threshold (dotted line), which were determined as negative in the quantitative polymerase chain reaction assay. (C) Example of a two-dimensional plot of a true positive sample in the T790M assay using formalin-fixed, paraffin-embedded genomic DNA. The main positive signal comes from the first quadrant.
Quantitative comparison of ddPCR with qPCR

Digital PCR has the advantage of measuring absolute copy numbers without using the standard curve. In the qPCR-based method, a standard curve is necessary to measure the copy numbers of test samples, but can introduce extra efforts and errors. In addition, commercial assays sometimes only use the relative quantitation method, ie, the ΔCt with the reference gene to determine the sample mutation status as either positive or negative. To find out whether the absolute copy numbers determined by ddPCR can recapitulate the mutation level measured by the qPCR method, we compared the copy numbers in log scale with the ΔCt measured for each patient sample. The two values showed relatively good correlation when the mutant copy numbers were relatively low (Figure 5). The $R^2$ value was 0.42 if the high copy number points were not considered. However, the ΔCt method showed a plateau with the increase in mutant copy number, suggesting that the ddPCR method has a better dynamic range than the qPCR method when a standard curve is not applied (Figure 5).

**EGFR** mutation spectrum in the lung cancer patient cohort

To test if the ddPCR methods could capture the **EGFR** mutation status in lung cancer patients, we grouped the samples into **EGFR** mutant and **EGFR** wild-type classes. About 49% of samples from this patient cohort contained one or two of the three mutations (Figure 6), which is consistent with the results of previous large-scale studies of the **EGFR** mutation frequency in Asian patients with lung cancer (36%–41%).

To further understand the mutation frequency of the individual mutation, we analyzed the **EGFR** mutant subset of patients and found that L858R and ex19del mutations were mutually exclusive and their mutation frequencies were 32% and 60%, respectively (Figure 6). Although this result is different from the findings of the two earlier studies in the Asian population, it is consistent with observations elsewhere that the ex19del mutation is more frequent than the L858R mutation.

Two patients who harbored the ex19del mutation also had the T790M mutation, and three patients harbored the T790M mutation alone. This low frequency (5/78, 6.4%) is consistent with the rare event of germline mutation of this predisposing allele, or sometimes linked to primary resistance, occurring together with a sensitizing mutation prior to drug exposure. It is worth noting that this primary mutation may have a different incidence in our particular patient cohort; for instance, more than 50% of pretreatment T790M mutations have been observed in certain lung cancer patient cohorts.8,12

---

**Figure 5** Correlation between ΔCt and copy numbers measured by ddPCR. Thirty-five patients positive for **EGFR** mutations had both ΔCt from the qPCR assay and the absolute copy number from ddPCR assay. A good correlation is observed when the copy number is relatively small.

**Abbreviations:** ddPCR, digital droplet polymerase chain reaction; qPCR, quantitative polymerase chain reaction; **EGFR**, epidermal growth factor receptor.

**Figure 6** **EGFR** mutation spectrum in this 78-patient cohort. Fifty-one percent of patients (**EGFR** wild-type, 43/78) showed no mutation in any of the L858R, ex19del, or T790M alleles, while 49% of patients (**EGFR** mutant, 35/78) showed **EGFR** mutation in the L858R, ex19del, or T790M alleles; among these, 32% were L858R, 60% were ex19del, and 13% were T790M mutations. Two patients harbored both ex19del and T790M mutations (5%).

**Abbreviation:** **EGFR**, epidermal growth factor receptor.
Thirty of the 40 never-smokers in this cohort were women. Among the 35 patients positive for the L858R or ex19del mutation, 21 were female never-smokers (Table 1). In this group of patients, the EGFR mutation was significantly more common in female never-smokers ($\chi^2$ test, $P=0.03$), which is consistent with previous findings.

**Discussion**

Since its discovery about a decade ago, the relationship between the effectiveness of gefitinib in lung cancer patients and EGFR mutation has become a paradigm for the success of targeted therapy in solid tumors. Many methods have since been developed, aiming to select patients based on the somatic mutation status of the EGFR gene.\textsuperscript{15,23–27} Direct sequencing and real-time PCR-based methods are most commonly used in the pathology departments of hospitals or clinical testing laboratories. However, despite their routine use, the direct sequencing method has limited utility, including low sensitivity and dependence on tumor cell content to detect mutations, and real-time PCR-based methods are semiquantitative in that a standard curve and reference control genes are required to determine the mutant copies. The digital PCR concept was proposed a while ago as a way of circumventing these limitations,\textsuperscript{28} and many studies have shown that digital PCR technology is superior for detection of rare mutation sequences on the background of confounding normal sequences in an absolute quantification manner.\textsuperscript{10,11,29}

Among the major platforms currently available for digital PCR,\textsuperscript{10} we chose the Bio-Rad ddPCR platform to investigate the clinical utility in the detection of EGFR mutation. The L858R and ex19del assays can detect as low as one mutant copy among 20,000 copies of normal DNA, and no mutant signal showed up when 92 normal DNA samples are examined, suggesting that this is a highly sensitive and specific assay. In addition, the ex19del assay uses a principle of the “indirect” nature of mutant calling,\textsuperscript{10} and this nature brings a possibility of detecting mutation or deletion in the place where the wild-type probe locates (ie, EGFR aa 756N-761D). In fact, we found one particular FFPE sample having such a mutation, which was confirmed by Sanger sequencing (data not shown).

The T790M assay shows good sensitivity, but the background noise is relatively high. By examining the two-dimensional plot of normal DNA and pure H$_2$O data, we think the background signal may come from non-specific amplification of normal DNA by the mutant probe or from primer/probe interference. By setting a threshold line of six copies per 20,000 normal copies (or 0.03%), and constructing a two-dimensional plot review, we could eliminate the false positives in the clinical samples tested. In fact, we observed such a background “mutant” signal in eight of 78 samples, with five samples being able to be clearly ruled out as false positives by the two criteria above (three samples $<0.03\%$, two samples showing droplets in quadrant 2). For the remaining three samples, we repeated the tests multiple times and found they were true positive samples that were not detected with the qPCR method, suggesting that the ddPCR assay is more sensitive in low abundance mutation detection. The technical background threshold of 0.03% based on the fresh normal genomic DNA samples may be an underestimate of the FFPE clinical samples in terms of the T790M assay; the reproducible clinical sample with a low allele frequency may be a better estimate of the background. For example, sample 2 in Figure 4 consistently showed an allele frequency above 0.05%, which may be a more clinically meaningful cut-off for true positives in patients.

The concordance of L858R and ex19del assays between qPCR and ddPCR is 100%, and is 96% for the T790M assay, suggesting that ddPCR can be a reliable alternative to the qPCR method. In addition, when comparing the $\Delta$Ct with the absolute copy number of the same sample, we found that ddPCR has a better dynamic range than qPCR. This is important because the $\Delta$Ct method of qPCR often gives ambiguous results when the mutant DNA is in low abundance in the test sample. Moreover, absolute copy number determination by ddPCR is accurate and convenient in terms of mutation calling in the clinical settings. The essence of a duplex assay can easily determine the abundance of mutant genes in a specimen, which can not only use smaller amounts of the clinical samples, but can also be used as a quantitative surrogate biomarker for the benefit of targeted therapeutics.\textsuperscript{31}

In summary, we demonstrated that the ddPCR method was easier to interpret without standard curves and internal controls than the qPCR method when used to detect three EGFR mutations in lung cancer patients. Using a cohort of 78 Chinese patients with lung cancer, we recapitulated previous findings with regard to frequency of EGFR L858R, ex19del, and T790M mutations\textsuperscript{13,17–19} using both qPCR and ddPCR methods. The ddPCR method can identify low abundance mutation events than qPCR. These results suggest that ddPCR can be used reliably in the clinical setting to guide patient stratification.

**Acknowledgments**

This research was supported by the National Natural Science Foundation of China (81202960). We thank Dr Yuhong Shen of WuXi AppTec and her team for providing the cell lines used in this study. We thank Ms Louise Mao for proofreading the manuscript. We also appreciate the volunteers at WuXi.
AppTec for providing the normal blood DNA samples used in this research.

Disclosure
The authors report no conflicts of interest in this work.

References