Probe-free allele-specific copy number detection and analysis of tumors

Ailin Zhu [a, b, 1], Xiaowei Guan [a, 1], Xinbin Gu [b], Guiqin Xie [a, *]

[a] Department of Physiology, Nanjing Medical University, Nanjing 210029, China
[b] Howard University, Washington, DC 20059, USA

ABSTRACT

Cancer development and progression frequently involve nucleotide mutations as well as amplifications and deletions of genomic segments. Quantification of allele-specific copy number is an important step in characterizing tumor genomes for precision medicine. Despite advances in approaches to high-throughput genomic DNA analysis, inexpensive and simple methods for analyzing complex nucleotide and copy number variants are still needed. Real-time polymerase chain reaction (PCR) methods for discovering and genotyping single nucleotide polymorphisms are becoming increasingly important in genetic analysis. In this study, we describe a simple, single-tube, probe-free method that combines SYBR Green I-based quantitative real-time PCR and quantitative melting curve analysis both to detect specific nucleotide variants and to quantify allele-specific copy number variants of tumors. The approach is based on the quantification of the targets of interest and the relative abundance of two alleles in a single tube. The specificity, sensitivity, and utility of the assay were demonstrated in detecting allele-specific copy number changes critical for carcinogenesis and therapeutic intervention. Our approach would be useful for allele-specific copy number analysis or precise genotyping.

© 2015 Elsevier Inc. All rights reserved.

Genetic alterations are the causative events of carcinogenesis. Detection of genetic variants is one of the keys to understanding the mechanisms of cancer development. Somatic genetic variants that render an individual sensitive or resistant to therapeutic agents have been identified [1]. Therefore, detection of genetic alterations is important for therapeutic interventions. Detection of these variants is becoming a crucial step in personalized medicine [2].

Each person has two copies of the genome, and single nucleotide polymorphisms are abundant [3]. Structural variants, such as segmental duplication or deletion, are frequently found in human populations [4], and each person may have more than 1000 structural variants [5]. Allele-specific variants inherited from parents may contribute to different phenotypes and vulnerability to different types of cancers. During carcinogenesis, allele-specific variants in somatic cells may also arise during the development

of disease [6]. Cancer genomes have various somatic sequence alterations compared with their normal individual genomes, ranging in size from single-base changes as point mutations, insertions, or deletions of a few nucleotides to insertions or deletions of large chromosomal fragments and even whole-genome duplications [7]. In addition, the genomes of tumor cells are heterogeneous due to diversified populations of tumor cells with different genetic alterations in different clonal lineages and stages of progression. Tumor cells often are mixed with different numbers of normal cells. Therefore, the genomes of tumor cells often deviate from a diploid state, making it challenging to correctly discover and interpret cancer genomes.

Allele-specific copy number aberrations were often detected by high-density single nucleotide polymorphism microarrays through measuring the quantity of both alleles at heterozygous loci, whereas genotyping arrays allow for estimation of the copy numbers of each allele [8]. With the advance of DNA sequencing technology, whole-genome and whole-exome sequencing are commonly used to quantify DNA copy number and to detect structural variants [9]. The number of sequencing reads for both alleles at variant heterozygous loci allows allele-specific copy numbers to be determined. Other approaches for the detection of

---

single nucleotide polymorphisms and copy number variants include restriction fragment length polymorphism methods [10], allele-specific polymerase chain reaction (PCR) [11], and real-time PCR methods [12]. Real-time PCR is one of the most commonly used methods to determine single nucleotide polymorphisms and copy number variants. Recently, digital PCR has been used to detect single nucleotide polymorphisms and copy number variants [13,14]. However, both real-time PCR and digital PCR need labeled dye probes to determine single nucleotide polymorphisms and copy number variants. The requirement for sequence-specific labeled probes considerably increases both the effort involved in and the cost of making and using variant detection assays.

Besides probe-based real-time PCR variant analysis, quantitative real-time PCR can be carried out in the presence of DNA-binding dyes such as SYBR Green I. In the presence of SYBR Green I, PCR fragment peaks can be distinguished due to their differential melting temperatures. For the detection of single nucleotide variants, SYBR Green-based approaches commonly rely on allele-specific PCR. In the PCR, two allele-specific primers (one each for the wild-type and mutant alleles) and one common primer are used [15]. SYBR Green I-based PCR can be used to monitor the amplification of any double-stranded DNA sequence. Because no probe is required, the SYBR Green I-based assay simplifies assay setup and reduces the cost of performing the assay.

Despite these approaches, it is still difficult to detect cancer genomes. Cancer genomes have varieties of genetic alterations in oncogenes. BRAF is one of the most frequently mutated genes in human tumors, including melanoma and thyroid cancer [16]. Mutations of KRAS genes are frequently detected in different cancers, including lung cancer, pancreatic cancer, melanoma, and thyroid cancer [17]. Mutations of the KRAS gene may occur at two different sites, most frequently at the 12th code of the gene. In mice, the $Kras^{G12D}$ mutation is sufficient to initiate carcinogenesis for pancreatic cancer and lung cancer [18–20]. During progression, the wild-type allele will be lost, promoting tumor progression [21]. PTEN is a tumor suppressor. Germline loss of the PTEN gene renders an individual vulnerable to tumorigenesis [22]. In individuals with germline PTEN loss, all somatic cells have only one functional PTEN allele. The remaining PTEN allele may then be lost during carcinogenesis [1,23]. Due to the heterogeneity of tumors, enrichment of tumors with advanced cancer features is necessary to detect copy number changes [21].

In this study, we describe a simple single-tube method that combines SYBR Green I-based quantitative real-time PCR and quantitative melting curve analysis both to detect nucleotide variants and to quantify allele-specific copy number variants in tumors. The approach is based on the quantification of the targets of interest and relative abundance of the two alleles in a single tube. The specificity, sensitivity, and utility of the assay were demonstrated for allele-specific copy number changes of the Pten gene. Our approach would be a useful alternative for either allele-specific copy number analysis or individual precise genotyping of tumor tissues.

Materials and methods

Genomic DNA samples

Genomic DNA from the SK-MEL-1 melanoma cell line and frozen mouse thyroid tumor tissue samples was extracted using the Qiagen DNeasy Blood and Tissue Kit according to the manufacturer's protocol (Qiagen, Hilden, Germany). DNA concentration was measured using a NanoDrop 2000c spectrophotometer (Thermo Scientific, Wilmington, DE, USA). DNA samples were stored in a −20 °C freezer.

Primer design

Primers were designed using the web-based Primer3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). The following primers were used in this study. For the BRAF gene:

BRAF-A, 5′-gcgggcCCACTCCATCGAGATTTCA-3′;
BRAF-T, 5′- gcgggcagggcggcCCACTCCATCGAGATTTCT-3′;
BRAF-common, 5′-CATGAAGACCTCACAGTAAA-3′.

For the mouse insertion locus:

Forward primer, 5′-CGATGGCATCAGGTCCTAAA-3′;
Reverse primer, 5′-GGGCAAGTCCTTACCTGGAT-3′.

For the mouse Pten gene:

Common primer, 5′-TTGCACAGTATCCTTTTGAAG-3′;
Mutant allele primer, 5′-ACGAGACTAGTGAGACGTGC-3′;
Wild-type allele primer, 5′-GTCTCTGGTCCTTACTTCC-3′.

For the reference transferrin receptor (Tfrc) gene:

Forward primer, 5′-TGCCGACAATAACATGAAGG-3′;
Reverse primer, 5′-AGTGCAATAGCTGCAAAGCA-3′.

Real-time PCR methods

Real-time PCR was performed in a total volume of 20 μl using 10 μl of 2 × QuantiTect SYBR Green PCR Master Mix (Qiagen, Valencia, CA, USA), 1 μl of 5 μM primer for each primer per reaction, 2 μl of the genomic DNA dilution (10 ng/μl), and water. The PCR cycling on the 7900HT Fast Real-Time PCR System (Life Technologies, Grand Island, NY, USA) was as follows. An initial denaturation step (95 °C for 15 min) was followed by 45 cycles of two-step amplification and fluorescence detection (94 °C for 15 s, 60 °C for 60 s) and a melting curve measurement step. Melting curve steps were performed in a single cycle of 95 °C for 15 s, 60 °C for 15 s, and 95 °C for 15 s while continuously measuring the change in fluorescence intensity.

Determination of total Pten copy number and wild-type Pten copy number

Genomic DNA from thyroid tissues of wild-type and heterozygous mice was quantified and used to construct standard curves for the Pten and Tfrc genes. PCR amplicons of the wild-type and mutant Pten alleles were quantified and used to generate a standard curve between the peak area ratio of the wild-type allele to the mutant allele and the percentage of the wild-type allele. The numbers of the Pten and Tfrc gene copies in genomic DNA from mouse thyroid tissues and tumors were calculated from each cycle threshold using the standard curves. Each diploid cell from mouse thyroid tissue was assumed to have 5.84 pg of DNA from two copies of the genome with a size of 2.8E + 9 bp using the formula $2 \times (2.8E + 9) \times 650 \times (1E + 12)/(6.23E + 23)$. The total Pten copy number per cell was calculated by dividing the number of Pten copies by the number of Tfrc copies in the genomic DNA sample and then multiplying by 2. To determine the allele-specific wild-type Pten copy number, a melting curve analysis was performed using an ABI Prism 7900HT sequence detection system. The melting analysis data were exported as a text file. The melting analysis values were imported into Prism 6.0 software to analyze the melting curve peaks of the wild-type and mutant Pten alleles. Scatterplot graphs