Chemometrics and Intelligent Laboratory Systems 179 (2018) 39–45

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

# An advanced model for the detection of short DNA sequences by mass spectrometry based on exonuclease III assisted recycling amplification

Cai-Xia Shi, Qing Liu, Zeng-Ping Chen *, Ru-Qin Yu

*State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha, Hunan, 410082, PR China*

ARTICLE INFO

*Keywords:*
Quantification analysis
Short DNA sequence
Exonuclease III
Target recycling amplification
Spectral shape deformation quantitative theory
Mass spectrometry

ABSTRACT

Exonuclease III assisted target recycling amplification strategy can be used to enhance the sensitivity of mass spectrometry for the detection of short DNA sequences. However, the distribution pattern of DNA fragments produced by exonuclease III assisted target recycling amplification is generally different for samples with different concentrations of the target DNA sequence, which hinders the extraction of both qualitative and quantitative information of the target DNA from mass spectral measurements using traditional univariate or multivariate models. In this contribution, an advanced model was derived based on a reasonable assumption for the qualitative and quantitative analysis of the mass spectral measurements of DNA fragments produced by exonuclease III-assisted target recycling amplification. Experimental results demonstrated that the integration of exonuclease III assisted target recycling amplification, mass spectrometry and the advanced model could achieve sensitive and accurate quantitative results for a target short DNA sequence in complex biological medium with a detection limit of 50 pM and a mean recovery rate within the range of 89.5%–106.7%. More interestingly, the proposed model could unambiguously identify single nucleotide polymorphisms based on the distribution patterns of residual DNA fragments. Therefore, with the aid of the proposed model, mass spectrometry based on exonuclease III assisted recycling amplification has great potential for the reliable, sensitive, selective, and relatively low-cost detection and quantification of short DNA sequences in clinical diagnosis and biomedical research.

## 1. Introduction

The development of reliable methods for sensitive and selective detection of DNA is of great importance in the field of biological studies, clinical diagnostics, and biomedical development [1,2]. Among the methods developed for the detection of DNA [3–10], autocatalytic amplification strategies based on exonuclease III (Exo-III) have been receiving extensive attention in the field of biosensing technology for its exquisite sensitivity, easy operation, isothermal reaction, and more importantly, no requirement of a specific recognition site [11,12]. Exo-III can catalyze the stepwise removal of mononucleotides from 3′-hydroxyl termini of double-stranded DNA with blunt or recessed 3′-termini. While its catalytic ability to single-stranded DNA and duplex DNAs with 3′ overhangs of four or more bases is inactivated or limited [2,11–13]. Hence, Exo-III provides a versatile biosensing platform and has been integrated with various signal detection techniques such as chemiluminescence [13,14], fluorescence [15–18], electrochemistry [19–21], colorimetry [22] and surface-enhanced Raman spectroscopy [2] for amplification detection of DNA.

Most of the biosensors based on Exo-III or other nucleases involve labeling of certain types of bioprobes [11,12,15–18]. Moreover, almost all the existing biosensors based on Exo-III are claimed to be able to identify single nucleotide polymorphisms (SNPs), however, they are actually only capable of equivocally discriminating the perfect complementary target DNA sequences and the base mismatched DNA sequences based on the differences in signal intensities [2,11–14,18] which can be caused by other factors besides SNPs such as concentration variations of the perfect complementary target DNA sequences, or even matrix effects. As a result, though these biosensors might have exquisite sensitivity, they are liable to produce false positive results.

Mass spectrometry has also been applied to the detection of DNA [23–28]. Compared with spectroscopic-based biosensing techniques, mass spectrometry has the advantages of excellent specificity and no requirement for the labeling of bioprobes. However, the detection limit of mass spectrometry for direct detection of DNA is generally in the level of tens of nanomolar [24], several orders of magnitude higher than those of biosensors in combination with various spectroscopic techniques especially fluorescence spectrometry. Amplifying the target DNA by the

* Corresponding author.
*E-mail address:* zpchen@hnu.edu.cn (Z.-P. Chen).

https://doi.org/10.1016/j.chemolab.2018.06.004
Received 29 April 2018; Received in revised form 14 June 2018; Accepted 15 June 2018
Available online 18 June 2018

polymerase chain reaction (PCR) can overcome the sensitivity problem of mass spectrometry [23,27,28]. The genomic DNA extracted from tissues can first be digested by restricted endonucleases to produce short DNA sequences [29]. The concentrations of the targeted short DNA sequences are then amplified to micromolar levels by PCR techniques to satisfy the concentration requirement of mass spectrometry [28]. However, the presence of nonspecific amplification products and excess primers in the PCR reaction mixture would lead to rather messy mass spectral data and hence ambiguous results. Therefore, it is critical to stringently purify the target DNA amplification product before mass spectrometric analysis [23], which greatly hinders the application of mass spectrometry for quantitative analysis of PCR reaction mixtures.

Another possible approach to enhance the sensitivity of mass spectrometry for DNA detection is to integrate mass spectrometry with Exo-III assisted target recycling amplification strategy. Huang et al. explored the feasibility of integrating mass spectrometry with an amplification strategy based on Exo-III [30]. But the mass spectral response in their method was the signal of rhodamine 6G enhanced in the presence of short olignonucleotides. It still has the same specificity problem as the biosensors based on various spectroscopic techniques do. Instead of the target DNA sequence and some nonspecific amplification products, the amplification products of the Exo-III assisted target recycling amplification strategy are the DNA fragments of the probe DNA sequence which is designed to be perfectly complementary to the target DNA sequence. So if mass spectrometry is used to analyze the reaction products of Exo-III assisted target recycling amplification, the recorded mass spectral data would be relatively neat and clean. However, as far as we know, until now, no one tries to quantify short DNA sequences by directly detecting the DNA fragments produced during the Exo-III assisted target recycling amplification process using mass spectrometry. One of the main reasons is that the target short DNA sequence produced by directly digesting of the genomic DNA with restricted endonucleases cannot be amplified by the Exo-III assisted target recycling amplification alone to a concentration level satisfying the requirement of mass spectrometry. Intuitively, the combination of PCR and the Exo-III assisted target recycling amplification can solve both the problem of messy mass spectral data of PCR and the amplification efficiency problem of the Exo-III assisted target recycling amplification. The target short DNA sequence produced by directly digesting of the genomic DNA with restricted endonucleases can first be amplified by PCR with fewer cycles to a concentration level of hundreds of picomolar which cannot be detected by mass spectrometry. With the presence of the target DNA sequence with concentration of hundreds of picomolar, the Exo-III assisted target recycling amplification is then capable of digesting the probe DNA into specific DNA fragments with concentrations of micromolar levels which can be detected by mass spectrometry. It seems that the combination of PCR and the Exo-III assisted target recycling amplification can realize the detection of short DNA sequences by mass spectrometry. However, another problem hindering the direct detection of short DNA sequences using mass spectrometry is that the distribution pattern of DNA fragments produced by Exo-III assisted target recycling amplification is generally different for samples with different concentrations of the target DNA (see section 4.2 for evidences). It is rather difficult to infer the concentration of the target DNA sequence from the mass spectral intensity of certain DNA fragments. So how to extract the qualitative and quantitative information of the target DNA sequence from the mass spectral distribution pattern of DNA fragments produced by Exo-III assisted target recycling amplification is the bottleneck in the detection of short DNA sequences by mass spectrometry. In this contribution, an attempt was made to derive an advanced model for qualitative and quantitative analysis of the mass spectral data of reaction products of Exo-III assisted target recycling amplification. The proposed model can not only extract quantitative information of the target DNA sequences from the mass spectral distribution pattern of DNA fragments produced by the Exo-III assisted target recycling amplification, but also permits unambiguous identification of single nucleotide polymorphisms. The combination of the proposed model, mass spectrometry and the Exo-III assisted target recycling amplification can therefore be developed into a reliable, sensitive, selective, and relatively low-cost assay platform for short DNA sequences.

## 2. Quantitative model for the detection of short DNA sequences by mass spectrometry based on Exo-III assisted recycling amplification

The schematic diagram in Fig. 1 illustrates the principal of mass spectrometry-based DNA quantification with the aid of Exo-III assisted recycling signal amplification. If a sample contains the target short DNA sequence, the probe DNA sequence designed to be perfectly complementary to the target DNA sequence will hybridize with the target DNA sequence to form double-strand DNA with a blunt 3′ end. By virtue of the ability to mediate a blunt or recessed 3′ end cleavage of double-stranded DNA [3,11–13], Exo-III can selectively digest the probe DNA in the double-stranded DNA in a 3′ to 5′ direction and release the target DNA sequence which will hybridize with a secondary probe DNA. A target DNA sequence can then be recycled hundreds of times to repeatedly initiate the digestion of the probe DNA by Exo-III. After termination of the digestion reaction and the removal of saline ions from the mixture through dialysis, the residual fragments of the probe DNA and its uncleaved form are detected by mass spectrometry. The higher the concentration of the target DNA in the sample to be analyzed is, the fewer the uncleaved probe DNA will remain, and the more residual fragments of the probe DNA will be produced. So, theoretically, the amount of the target DNA can then be deduced from the mass spectral measurements of the residual fragments of the probe DNA and its uncleaved form. Since the cleavage of double-stranded DNA with a blunt or recessed 3' end by Exo-III is sequence-independent, the above DNA quantification strategy based on Exo-III assisted recycling signal amplification can be applied to an arbitrary target short DNA sequence.

For quantitative analysis of the target DNA sequence in a sample, the mass spectral responses of the multiply charged ions of the residual DNA fragments were assembled to construct a virtual mass spectrum. The virtual mass spectrum ($\mathbf{x}_k$) of the $k$-th standard sample can be decomposed according to the following equation (Eq. (1)).

$$\mathbf{x}_k = \sum_{j=4}^{J} p_k \cdot c_{j,k} \cdot \mathbf{r}_j + \mathbf{d}_k \tag{1}$$

Where $c_{k,j}$ is the concentration of the $j$-mer residual fragment of the probe DNA ($J$ bases) hydrolyzed by Exo III in the presence of the target DNA sequence with concentration of $c_{\text{targ},k}$. $\mathbf{r}_j$ represents the mass spectral response of the $j$-mer fragment per unit concentration (It was found in experiment that the 4-mer DNA fragment was the shortest residual fragment of the probe DNA hydrolyzed by Exo-III, see Section 4.2 for details); The multiplicative parameter $p_k$ explicitly accounts for the multiplicative confounding effects on mass spectral intensities caused by

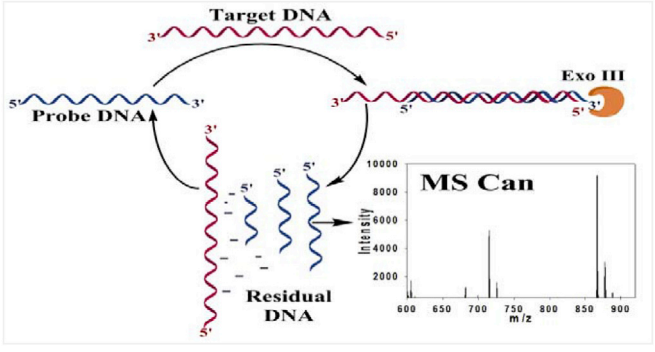


**Fig. 1.** Schematic representation of Exo-III assisted signal amplification strategy for DNA quantification using mass spectrometry.