Trends in Peptide and Protein Sciences
2016, Vol. 1, No.2, 61-67

Original Research Article

Development of a Semi-Quantitative Multiplex PCR Method for Detecting Residual *Pichia Pastoris* Host Cell DNA in Biopharmaceuticals

Atefeh Namipashaki [a], Reza Nekouian [b], Zargham Sepehrizadeh [c] and Neda Setayesh [c*]

[a] *Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Tehran University of Medical Sciences, Tehran, Iran.*
[b] *Department of Medical Biotechnology, School of Allied Medicine, Iran University of Medical Sciences, Tehran, Iran.*
[c] *Department of Pharmaceutical Biotechnology, Faculty of Pharmacy and Biotechnology Research Center, Tehran University of Medical Sciences, Tehran 1417614411, Iran.*

Article history:
Received: 29 November 2016
Accepted: 15 December 2016

Keywords:
Residual DNA
Yeast
Pichia pastoris
Multiplex PCR
Gel densitometry

HIGHLIGHTS

- The impurity of residual host cell DNA is an important concern in production of biopharmaceuticals.
- *Pichia pastoris* is an effective and versatile system for the expression of recombinant proteins.
- Quantitative Polymerase Chain Reaction could be used for quantifying residual host-cell DNA.
- We designed a sensitive and valid PCR method for detection and quantification of *Pichia* residual DNA.

ABSTRACT

The use of the methylotrophic yeast, *Pichia pastoris*, as one of the most effective and versatile systems for the expression of heterologous proteins in biopharmaceutical manufacturing has become increasingly popular in recent years. The impurity caused by residual host cell DNA is one of the major concerns in production of recombinant therapeutics. The aim of the present study was to develop a semi-quantitative, multiplex PCR method to determine the level of impurity in biopharmaceuticals produced in *Pichia pastoris* as the host. Primers were designed based on the rDNA repeat region and optimized to achieve the limit of detection in a multiplex PCR following by analyzing with MYImageAnalysis (Thermo Fisher Scientific, USA) software to quantify the concentration of *Pichia pastoris* genomic DNA in pertinent controls and drug samples. The multiplex PCR were able to detect up to 1 femtogram (fg) of genomic DNA under optimized conditions. Moreover, achieved concentration of DNA in controls and samples through relevant standard curve indicates the feasibility of this method in the presence of inhibitory effects. In comparison with other methods such as real-time PCR and Threshold assay, the assay shows acceptable sensitivity, precision and linearity along with ease of use, low equipment costs and analyte flexibility. We thus propose this method to be considered as a useful tool to estimate host cell residual DNA in biopharmaceuticals produced in *Pichia pastoris*.

Introduction

Biopharmaceuticals play a pivotal role in modern medicine. Various expression systems have been employed for the production of therapeutic proteins. Yeast host systems are one of the most widely used approaches for bio-

* *Corresponding Author*:
Email: nsetayesh@tums.ac.ir (N. Setayesh)

manufacturing of recombinant proteins especially those that cannot be actively expressed in other host systems such as *Escherichia coli (E. coli)* due to glycosylation requirements for proper folding and biological activity (Gerngross, 2004; Schmidt, 2004). *Pichia pastoris* (*P. pastoris*) seems to be the second most frequently used yeast strain after *Saccharomyces cerevisiae* (*S. cerevisiae*) for both basic research projects and industrial manufacturing (Macauley-Patrick et al., 2005). Compared to *S. cerevisiae*, *P. pastoris* shows higher cell density cultivation due to its powerful methanol-inducible alcohol oxidase1 (AOX1) promoter. It also has a higher secretion capability for heterologous proteins (Potvin et al., 2012). Other advantages of the *P. pastoris* system over expression systems derived from higher eukaryotes such as insect and mammalian tissue culture cell systems include higher expression levels, simplicity and speed of the protocols and cost effectiveness (Cereghino and Cregg, 2000). The US Food and Drug Administration (FDA) approved the first recombinant protein produced in *P. pastoris* (Kalbitor® by Dyax Corp.) in 2009 while other similar products await the results of clinical trials (Walsh, 2010).

All biological products originating from complex living systems are mixed with a large number of other macromolecules and potential impurities such as residual host cell DNA and host cell proteins. Because of safety reasons, downstream processing is performed to confirm that such impurities are reduced to acceptable levels in the final drug product (Lokteff et al., 2001; Forcic et al., 2005). Since the residual DNA could theoretically induce oncogenic events and/or reduce efficacy in recipients of the biological products, its quantity determination should be considered as an essential part of the product safety assessment (Sheng-Fowler et al., 2009; Yang et al., 2010; Shen et al., 2013).

In the 1987 World Health Organization (WHO) guideline, the maximum accepted amount of total rDNA is determined to be 100 picograms per dose, which is the cutoff value also used by the FDA . More recently, the WHO limit was changed to 10 nanograms per dose, which is also required by the European Union (EU) (Lebron et al., 2006). There are different approaches for quantifying levels of host-cell DNA in biopharmaceutical products such as PicoGreen analysis, the Threshold assay, hybridization techniques and quantitative Polymerase Chain Reaction (qPCR).These methods detect cellular DNA based on total DNA or species-specific DNA (Durrant et al., 1995; Strachan and Read., 1999). High throughput and low cost are the two main advantages of the PicoGreen analysis but low sensitivity is the main limitation of this technique (Ikeda et al., 2009).

The Threshold assay suffers from low throughput and relatively high cost though it offers a highly sensitive and standardized protocol. Lack of sensitivity and relatively long testing time are the major drawbacks of hybridization techniques, which has reduced their usage in recent years (Wang et al., 2012). Recently, PCR based methods, specifically real-time PCR, have been widely used for the determination of residual DNA in biological and biopharmaceutical manufacturing (Nissom, 2007; Lee et al., 2010; Cai et al., 2011). In PCR methods amplification of a specific sequence in template DNA allows for detection of extremely low levels of contaminating DNA in less than two hours.

To our knowledge, there is no published data on the quantification of host cell residual DNA in biopharmaceuticals produced in *P. pastoris*. The aim of this study is to design a sensitive and valid method for detection and quantification of *Pichia* residual DNA. Effectiveness of this method for residual DNA analysis on a sample of marketed biopharmaceutical will be evaluated.

Materials and Methods

Materials

Pichia pastoris G115 was obtained from Pichia Expression Kit (Invitrogen, Thermo Fisher Scientific, USA). PrimePrepTM Genomic DNA Isolation Kit was purchased from GeNet Bio (Korea) and nuclease-free water (Cat No: R0582) was provided from Fermentase (Thermo Fisher Scientific, USA). Custom synthetic oligonucleotide primers were obtained from SinaClon (Iran). PCR Master Mix 2X (Cat No: K0171) was also supplied from Thermo Fisher Scientific (USA).

Preparation of P. pastoris genomic DNA

P. pastoris G115 was grown in YPD liquid media consisted of 1% (w/v) yeast extract, 2% (w/v) peptone and 2% (w/v) D-glucose in distilled water. Genomic DNA was extracted using the PrimePrepTM Genomic DNA Isolation Kit. The initial concentration of *P. pastoris* genomic DNA was 90 ng/µl. To prepare the DNA standard curve, 10-fold serial dilutions of *P. pastoris* genomic DNA were prepared using nuclease-free water to generate concentrations between 20 pg/µl and 0.2 fg/µl.

Primer design

Primers targeting the rDNA repeat region from the GS115 strain of *P. pastoris* (GenBank accession number: FN392325) were designed using the online software program Primer3Plus. The sequences of the amplification primers used were 5’-CCCTCCAATTGTTCCTCGT-3’ (forward) and 5’-GGGGAATAAGGGTTCGATTC-3’ (reverse) for the 18S rDNA region and 5’-TCGCTATGAACGCTTGACTG-3’ (forward) and 5’-GGGGCTCATGAAGAACAGAA-3’ (reverse) for the

Table 1. Thermal cycling conditions used for different reactions.

Program names	PCR conditions
A	95 ºC for 3 min, 30x [95 ºC for 20 s, 57 ± 7 ºC for 10 s and 72 ºC for 20 s], 72 ºC for 8 min
B, C, D, E	95 ºC for 3 min, 30x [95 ºC for 20 s, 54 ºC for 10 s and 72 ºC for 20 s], 72 ºC for 8 min

26S rDNA region. These primer sets amplify 189-bp and 192-bp fragments from the 18srDNA and 26srDNA repeat regions, respectively. Custom synthetic oligonucleotide primers were obtained from Sigma.

Determination of optimal annealing temperature, concentration and limit of detection for each primer pair

After DNA isolation, gradient PCR reactions were performed to determine the optimum annealing temperature (program name: A) and concentration (program name: B) for each primer pair. The limit of detection was subsequently determined for each primer pair by performing a PCR reaction using optimized conditions with serial dilutions of template DNA (from 100 pg to 1 fg) (program names: C and D). Tables 1 and 2 show the thermal cycling conditions and the reaction contents for each program.

Preparation of spiked control, negative controls and samples

To evaluate the efficiency of multiplex PCR in the presence of potential inhibitory ingredients from commercial recombinant therapeutics, spiked controls were prepared. A known amount of *P. pastoris* genomic DNA was spiked into a 1/10 dilution of the recombinant hepatitis B surface antigen vaccine, which was made in the host *P. pastoris*, to generate final concentrations of 1, 10 and 100 fg. To reduce the inhibitory effects of vaccine ingredients, the vaccine was diluted 10-fold. An un-spiked control was also analyzed to determine the level of *P. pastoris* genomic DNA already presented in the recombinant hepatitis B surface antigen vaccine. In addition, to exclude false results from the PCR amplification, a negative control was included in each run. The negative control contained nuclease-free water in place of the DNA template. All spiked, un-spiked and negative controls were prepared in duplicate.

Two samples generated in a *Saccharomyces cerevisiae* host were also prepared for testing: recombinant insulin and recombinant hepatitis B antigen vaccines. A recombinant hepatitis B antigen vaccine produced in the yeast *Hansenula polymorpha* (syn. *Pichia angusta*) was also tested. All samples were diluted 10-fold.

Multiplex PCR amplification

Multiplex PCR reactions for triplicate 10-fold serially diluted DNA standards, duplicate quality controls, the negative control and samples from the hosts described above were performed using the E PCR program (Table1).

Quantification of PCR products using agarose gel densitometry

Amplified DNA was analyzed using gel electrophoresis with a 1% agarose gel. Eight microliters of each sample were loaded onto the gel, and the results were visualized using a UV transilluminator. The images were saved for

Table 2. Reaction contents for each program used in this study.

Program*	genomic DNA stock concentration (5 µL)	Forward / Reverse primer stock concentration (2 µL)	Types of Primers
A	20 pg/µL	20 µM	26srRNA or 18srRNA forward and reverse primer
B	20 pg/µL	5,10,15,20 µM	26srRNA or 18srRNA forward and reverse primer
C	20, 2, 0.2, 0.02, 0.002, 0.0002 pg/µL	5 µM	26srRNA forward and reverse primer
D	20, 2, 0.2, 0.02, 0.002, 0.0002 pg/µL	5 µM	18srRNA forward and reverse primer
E	20, 2, 0.2, 0.02, 0.002, 0.0002 pg/µL	5 µM	18srRNA and 26srRNA forward and reverse primer

* 12.5 µL of PCR Master Mix (2X) and 5.5 µL of nuclease free water were added to each reactions. Final reaction volume was 25 µL.

Table 3. Density of DNA standards, spikes and sample product bands.

Concentration of template DNA(fg)	Mean density of product band (intensity/pixel)	Concentration of template DNA(fg)	Mean density of product band (intensity/pixel)
DNA standard 100 pg	66959	DNA spike 100 fg	72253
DNA standard 10 pg	61513	DNA spike 10 fg	65002
DNA standard 1 pg	52296	DNA spike 1 fg	61692
DNA standard 100 fg	49403	Sample(*vaccine produced in P. pasroris*)	34043
DNA standard 10 fg	42242		
DNA standard 1 fg	37582		

further analysis.

The intensity of the bands was analyzed using MY Image Analysis Software, version 1.1 (Thermo Fisher Scientific). A standard curve was subsequently generated by plotting the log of the primary template DNA amount against the density of the corresponding band, i.e., mean intensity per pixel. The residual DNA concentration of the samples and the spiked controls was estimated by comparing the band intensity to the standard curve.

Results

Primer efficiency and process optimization

Gradient PCRs demonstrated that selecting the minimum primer concentration (5 µM) and an annealing temperature of 54°C for both primer pairs would maximize the amplification of products and minimize primer dimer formation, resulting in the highest limit of detection for residual DNA.

PCR reactions were performed to determine the limit of detection for each primer pair. The results indicated that the 18s rRNA and 26s rRNA primer pairs could amplify *P. past*oris genomic DNA at the amount of 1 pg and 100 fg respectively (Data not shown).

Multiplex PCR assay and data analysis

Multiplex PCR using serially diluted genomic DNA indicated that the amplification of *P. pastoris* genomic DNA was enhanced by approximately 100 times when two efficient primers were used together, facilitating the detection of genomic DNA at a concentration of 1 fg (Fig. 1).

The *P. pastoris* genomic DNA standard curve displayed a positive correlation between primary template DNA concentration and band density (Fig. 2). The concentration of DNA in the spiked controls and the un-spiked control (i.e., the sample from the recombinant hepatitis B antigen vaccine produced in a *P. pastoris* host) was determined using the standard curve (Table 3) (Fig. 3). The band density of the spiked controls was higher than that of the corresponding standard DNA, likely due to the presence of residual DNA in the vaccine sample, which indicates

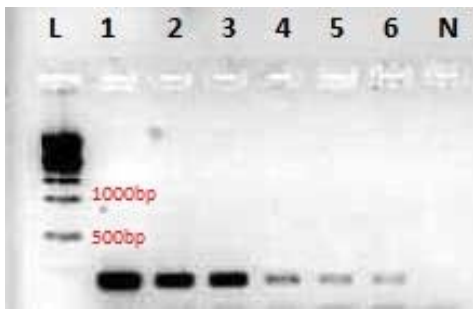

Figure. 1. Multiplex PCR assay of serially diluted genomic DNA (L: Ladder, 1: 100 pg genomic DNA, 2: 10 pg genomic DNA, 3: 1 pg genomic DNA, 4: 100 fg genomic DNA, 5: 10 fg genomic DNA, 6: 1 fg genomic DNA, N: negative control).

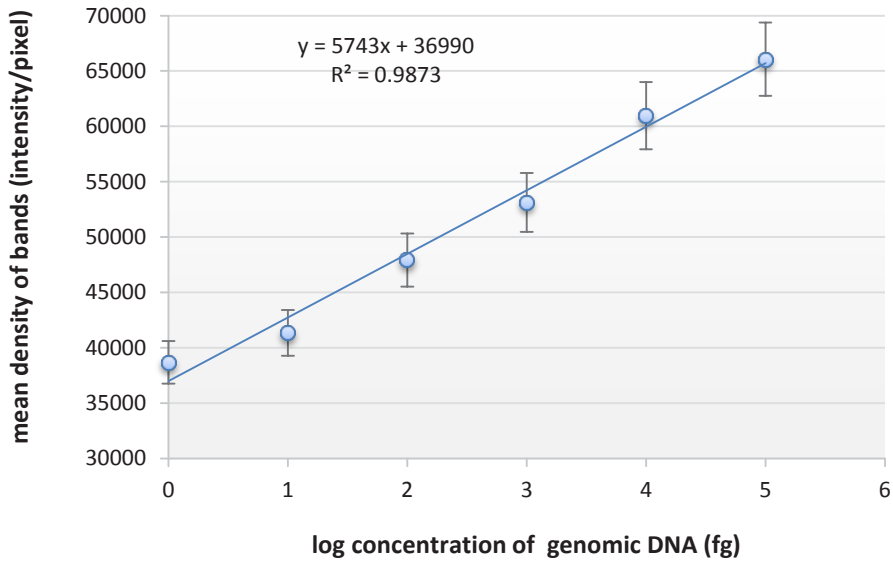

Figure. 2. Standard curve of *P. pastoris* genomic DNA generated by plotting the density of product bands against the log of the template DNA amount.

no inhibitory effect due to the 10-fold dilution of the vaccine.After analyzing the primers using the NCBI blast program, it was observed that these primers could also amplify a 192-bp fragment from the rDNA repeat region of *S. cerevisiae*, which is also used as an expression host.

The efficacy of this system for the amplification of *S. cerevisae* was evaluated. Although the genome sequence of *H. polymorpha* has not yet been published, it is expected that these primer sets could also amplify a similar fragment from this yeast (Fig. 4).

Discussion

This study was performed to develop a method for the detection of residual DNA from the host *P. pastoris* in biopharmaceuticals. A desirable sensitivity level for the detection of residual DNA was achieved using the following strategies: 1) designing two sets of primers that both efficiently amplify a similar size fragment from the rDNA repeat region, and 2) developing a detection method that uses a multiplex PCR assay. Since the difference between the amplified fragment sizes is not

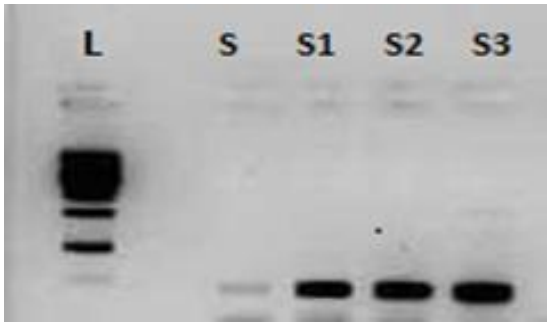

Figure. 3. Multiplex PCR assay of spiked controls and un-spiked control sample (L: Ladder, S: sample, S1: 1 fg spike, S2: 10 fg spike, S3: 100 fg spike).

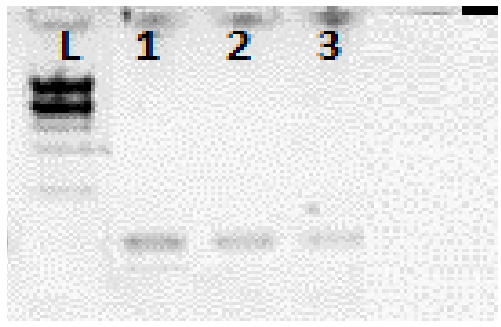

Figure. 4. Multiplex PCR assay of samples produced by other hosts (L: Ladder, 1: recombinant insulin, 2: recombinant hepatitis B antigen produced in *S. cerevisiae*, 3: recombinant hepatitis B antigen produced in *H. polymorpha*).

distinguishable on agarose gel electrophoresis, using these two sets of primers led to achieve increased sensitivity. In addition, choosing suitable image analysis software for calculating the intensity of amplified bands could make the results sufficiently accurate and sensitive for this method to become a standard method for determining whether to accept or reject a lot or batch of biopharmaceuticals.

In some cases, sensitivity can be increased by using a target with a higher copy number, such as a sequence from a repeat region within the genome. However, high copy number sequences, such as transposon elements, are less stable in the genome and may undergo significant changes in copy number. This instability could obscure the data concerning the presence of residual DNA if the target copy number changed during production (Lovatt, 2002). In addition, target genes with higher copy number do not provide the highest measurement or reveal the largest clearance of residual host cell DNA from purified samples. These findings suggest that different DNA sequences may be cleared or degraded at different rates and highlight unexpected factors that must be considered when choosing a PCR gene target for QPCR assays (Verardo et al., 2012). In light of these findings, we focused on the rDNA repeat region, which does not change copy number as frequently as high copy number transposing elements but meets the sensitivity requirements for this type of analysis.

Although real-time PCR quantifies DNA more precisely than the method presented here (Nissom, 2007), the aim of residual DNA detection is to determine whether the residual DNA in a dose of recombinant therapeutic is below the permissible limits set by regulatory guidelines. Thus, the exact amount of residual DNA doesn't necessarily seem to be quantified. In addition, the cost of analysis can be significant and should be considered when evaluating an assay.

The price per sample analysis in the multiplex PCR method is much lower than that of quantitative real-time PCR, which could improve cost management when manufacturing biopharmaceuticals. Besides, ease of execution, robustness and good enough sensitivity are additional advantages of this method.

The detection of a similar product from other yeasts suggests that these primer sets could also be used to test for the presence of residual DNA from other yeast hosts, including *S. cerevisiae* and *H. polymorpha*. Although this could reduce the specificity of the test, it also provides a high-throughput system for analyzing the residual DNA of all yeast hosts with a universal primer set. However, a little more carefulness may be required to prevent possible contamination and false results could effectively be excluded by running negative controls. It is also worth noting that while the same primer set can be used for multiple hosts, each host requires its own genomic DNA standard curve for quantification.

Conclusion

In conclusion, this work presents a convenient and reliable method for detecting residual *P. pastoris* DNA in purified protein samples. This study provides a low-cost and sensitive alternative screening tool for assessing the batches quality control of purified protein samples. However, it would be recommended to analyze the size of residual *P. pastoris* DNA in addition to its quantity.

Acknowledgements

Authors would like to thank M. Razavi for his useful comments on this manuscript. This study was supported financially by the grant number of 16235 from the Research Council of the Tehran University of Medical Sciences.

Competing Interests

The authors declared that there are no competing interests.

References

Cai, H., Gu, X., Scanlan, M. S. and C. R. Lively, (2011). ″Development of a quantitative PCR assay for residual mouse DNA and comparison of four sample purification methods for DNA isolation.″ *Journal of Pharmaceutical and Biomedical Analysis*, **55**: 71-77.

Cereghino, J. L. and J. M. Cregg, (2000). ″Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*.″ *FEMS Microbiology Reviews*, **24**: 45-66.

Durrant, I., Brunning, S., Eccleston, L., Chadwick, P. and M. Cunningham, (1995). ″Fluorescein as a label for non-radioactive in situ hybridization.″ *The Histochemical Journal*, **27**: 94–99.

Forcic, D., Cakanic, K. B., Ivancic, J., Jug, R., Barut, M., Strancar, A. and R. Mazuran, (2005). ″Chromatographic detection of residual cellular DNA on short monolithic columns.″ *Analytical Biochemistry*, **336**: 273-278.

Gerngross, T. U. (2004). ″Advances in the production of human therapeutic proteins in yeasts and filamentous fungi.″ *Nature Biotechnology*, **22**: 1409-1414.

Ikeda, Y., Iwakiri, S. and T. Yoshimori, (2009). ″Development and characterization of a novel host cell DNA assay using ultra-sensitive fluorescent nucleic acid stain ″PicoGreen″.″ *Journal of Pharmaceutical and Biomedical Analysis*, **49**: 997-1002.

Lebron, J. A., Troilol, P. J., Pacchione, S., Griffiths, T. G., Harper, L. B., Mixson, L. A., Jackson, B.E., Michna, L., Barnum, A. B., Denisova, L., Johnson, C. N., Maurer, K. L., Morgan-Hoffman, S., Niu, Z., Roden, D. F., Wang, Z., Wolf, J. J., Hamilton, T. R., Laux, K. M., Soper, K. A. and B. J. Ledwith, (2006). ″Adaptation of the WHO guideline for residual DNA in parenteral vaccines produced on continuous cell lines to a limit for oral vaccines.″ *Developments in Biologicals*, **123**: 35-44.

Lee, D. H., Bae, J. E., Lee, J. H, Shin, J. S and I. S. Kim, (2010). ″Quantitative Detection of Residual E. coli Host Cell DNA by Real-Time PCR.″ *Journal of Microbiology and Biotechnology*, **20**:1463-1470.

Lokteff, M., Klinguer-Hamour, C., Julien, E., Picot, D., Lannes, L., Nguyen, T., Bonnefoy, J. Y. and A. Beck, (2001). ″Residual DNA quantification in clinical batches of BBG2Na, a recombinant subunit vaccine against human respiratory syncytial virus.″ *Biologicals*, **29**:

123–132.

Lovatt, A. (2002). ″Applications of quantitative PCR in the biosafety and genetic stability assessment of biotechnology products.″ *Journal of Biotechnology*, **82**: 279-300.

Macauley-Patrick, S., Fazenda, M.L., McNeil, B. and L. M. Harvey, (2005).″ Heterologous protein production using the Pichia pastoris expression system. ″ *Yeast*, **22**: 249-270.

Nissom, P. M. (2007). ″Specific detection of residual CHO host cell DNA by real-time PCR.″ *Biologicals*, **35**: 211-5.

Potvin, G., Ahmad, A. and Z. Zhang, (2012). ″Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review.″ *Biochemical Engineering Journal*, **64**: 91-105.

Schmidt, F. R. (2004). ″Recombinant expression systems in the pharmaceutical industry.″ *Applied Microbiology and Biotechnology*, **65**: 363-372.

Shen, X., Chen, X., Tabor, D. E., Liu, Y., Albarghouthi, M., Zhang, Y. F. and M. S. Galinski, (2013). ″Size analysis of residual host cell DNA in cell culture-produced vaccines by capillary gel electrophoresis.″ *Biologicals*, **41**: 201-8.

Sheng-Fowler, L., Lewis, J. R. and K. Peden, (2009). ″Issues associated with residual cell-substrate DNA in viral vaccines.″ *Biologicals*, **37**: 190-5.

Strachan, T. and A. P. Read, (1999). ″Nucleic acid hybridization assays″. In: *Human Molecular Genetics*. 2nd ed., Wiley-Liss, Inc., New York, pp. 95–119.

Verardo, M. L., Carvalho, J. G., Delgado, D. N. and S. T. Kuhns, (2012). ″Accuracy and sensitivity of residual DNA detection by QPCR is not predicted by target copy number.″ *Biotechnology Progress*, **28**: 428-34.

Walsh, G. (2010). ″Biopharmaceutical benchmark.″ *Nature Biotechnology*, **28**: 917-24.

Wang, X., Morgan, D. M., Wang, G. and N. M. Mozier, (2012). ″Residual DNA analysis in biologics development: review of measurement and quantitation technologies and future directions.″ *Biotechnology and Bioengineering*, **109**: 307-17.

Yang, H., Zhang, L. and M. Galinski, (2010). ″A probabilistic model for risk assessment of residual host cell DNA in biological products.″ *Vaccine*, **28**: 3308-3311.