

Expression of *Shigella flexneri ipaB* Gene in TobaccoMandana Ohadi, Rahimeh Rasouli, Elham Darzi-Eslam, Anis Jafari<sup>\*†</sup>, and Parastoo Ehsani<sup>\*†</sup>

Molecular Biology Unit, Pasteur Institute of Iran, Tehran, Iran

<sup>†</sup> These authors equally contributed to this work

## Abstract

**Background:** Shigellosis is a leading cause of diarrhea in many developing countries and although the disease can be controlled and managed with antibiotics, the constant emergence of resistant species requiring ever newer antibacterial drugs make development of an effective vaccine necessary. The bacteria are highly contagious and since immunity to *Shigella* is serotype-specific a multi-serotype vaccine is required for adequate protection. Proteins encoded by *Shigella* invasion plasmid, which are part of the Type Three Secretion System (TTSS) of this bacteria, are good candidate as vaccine targets since they are both immunogenic and conserved between different *Shigella* species. The advent of molecular farming, which is a low cost system, has opened up new venues for production of recombinant proteins. In view of the difficulties encountered in expressing IpaB in *Escherichia coli* (*E. coli*), the feasibility of the expression of this protein in tobacco has been investigated.

**Methods:** The *ipaB* gene was cloned in place of the Hygromycin gene in pCambia1304 containing GFP as a reporter gene. The vector was then transferred into competent *Agrobacterium tumefaciens* (*A. tumefaciens*) strain LBA4404 which was used for agro-infiltration of *Nicotiana tabacum* (*N. tabacum*) leaves. Transformation was confirmed by expression of GFP. The gene was also cloned in pBAD/genelll A and transformed *E. coli* host containing the construct was induced using different amounts of L-arabinose as inducer. Expression of *IpaB* gene by both hosts was determined by Western blotting using anti-IpaB monoclonal antibody.

**Results:** The data obtained showed that IpaB was expressed in plant leaves but expression in *E. coli* was not detectable.

**Conclusion:** This study showed that *N. tabacum* is capable of expressing this protein without its specific chaperon and in levels detectable by Western blotting.

Avicenna J Med Biotech 2013; 5(2): 118-124

**Keywords:** *Escherichia coli*, Recombinant proteins, Shigellosis, Tobacco**\* Corresponding authors:**

Parastoo Ehsani Ph.D. and Anis Jafari Ph.D., Molecular Biology Unit, Pasteur Institute of Iran, Tehran, Iran

Tel: +98 21 66953311

**E-mail:**

p\_ehsani@yahoo.com

Received: 1 Aug 2012

Accepted: 22 Oct 2012

## Introduction

Diarrhea caused by *Shigella* species is a severe disease with high morbidity and mortality especially in developing countries<sup>1</sup>. It has been estimated that in Asia 91 million *Shigella* episodes and 414,000 deaths occur annually with *Shigella flexneri* being the most

common serotype<sup>2</sup>. As few as 10 organisms can cause illness and the low infective dose makes these bacteria highly contagious with oral-fecal contact being the predominant route of transmission from person to person. *Shigella dysenteriae* (*S. dysenteriae*), *Shigella*

*flexneri*, *Shigella sonnei*, and *Shigella boydii* are the four species that comprise the *Shigellae* genus and are divided into over 45 serotypes depending on the structure of the O antigen component of their outer membrane lipopolysaccharide<sup>1,3</sup>. The Shiga toxin-producing *S. dysenteriae* serotype 1 causes the most severe infections including hemolytic uremic syndrome as well as dysentery epidemics<sup>4</sup>.

Although vaccine development has been a priority for the World Health Organization for many years, no licensed vaccine is as yet available and the emergence of drug resistant strains makes the development of an effective vaccine even more urgent<sup>5-7</sup>. Approaches to vaccine development have included the use of killed<sup>8</sup>, live attenuated<sup>9-12</sup> and recombinant carrier<sup>13</sup> organisms, polysaccharide conjugates<sup>14,15</sup>, and LPS-protein mixtures<sup>16,17</sup>. The shortcomings of these candidate vaccines have been either poor immunogenicity or high reactogenicity when tested in humans. Furthermore, it has been shown that immunity to *Shigella* is serotype-specific, limiting the scope of protection offered by these experimentally developed vaccines and necessitating the development of a multi serotype vaccine for adequate protection<sup>3</sup>.

*Shigellae* species, like many gram-negative bacteria, rely on a TTSS as an essential virulence component which is present at a density of 50 to 100 per bacterial cell and resembles a molecular needle and syringe. The needle tip complex is composed of the invasion plasmid proteins, IpaB, IpaC and IpaD which are required for invasion of epithelial cells as well. These proteins are conserved between different *Shigella* species and serotypes making them desirable for use as candidate vaccines. The protective efficacy of antibodies against IpaB and IpaD (especially IpaB) was demonstrated recently in a mouse model of intranasal immunization and pulmonary challenge with homologous and heterologous strains<sup>3</sup>.

Hydrophobic IpaB protein is encoded by 1743 bp gene with an apparent molecular weight of 57 kDa and its production in *E. coli*

has been difficult. Expression of IpaB in *E. coli* has been only achieved fused with thio-redoxin or complexed with its chaperon IpgC<sup>18,3</sup> which in both cases require additional steps to purify the recombinant protein.

Plants have been increasingly used for production of genetically engineered biological products in recent years and the term "Molecular farming" has been coined to describe this process<sup>19</sup>. Several advantages are attributable to this eukaryotic host such as low cost of production, ease of scale up, ability to perform eukaryotic modification of the product post-translationally<sup>20</sup>. *N. tabacum* has been widely used in molecular farming as a non-food and non-forage product that is easy to manipulate and contains high fresh leaf weight and seed<sup>21</sup>. This is the first study on production of IpaB protein in plants.

## Materials and Methods

### Constructs for expression in tobacco and *E. coli*

The *IpaB* coding region from *Shigella flexneri* for expression in tobacco plants and *E. coli* was amplified using the following primers, which were designed in this study:

ipaB-tob-F: 5'-aata **ctc gag** gcc gcc acc atg cat aat gta agc acc aca ac-3'

ipaB-tob-R: 5'-atat **ctc gag** tca tag ctc atc ttt ctc aga *gtg gtg gtg gtg gtg* agc agt agt ttg ttg caa att g-3'

ipaB-eco-F: 5'-ata gca **cca tgg** gac ata atg taa gca cca ca-3'

ipaB-eco-R: 5'-agc **tct aga** gta gtt tgt tgc aaa att g-3'

Forward primer for cloning the PCR product in pCambia1304 (a gift from Dr. Rajabi Memari; Shahid Chamran University of Ahvaz), contained restriction site for *XhoI* (bold) and Kozak sequence (underlined). In the reverse primer a restriction site for *XhoI* (bold), a KDEL retention signal (underlined), and His-tag sequence (italic) were included. Restriction sites for enzymes *NcoI* and *XbaI* were incorporated in forward and reverse primers used for amplification of the gene for *E. coli* system, respectively.

The chemicals were purchased from Sigma and Merck companies, the enzymes and molecular weight markers were purchased from Fermentas, Lithuania.

The molecular biology techniques were done according to the gene cloning manual<sup>22</sup>. The gene for use in both plant and bacterial system was amplified according to the Enzyme manufacturer protocol (Fermentas, Lithuania), briefly in aliquots of 50  $\mu$ l containing 500 ng of template, 1.5 U High Fidelity Enzyme Mix, 2 mM MgSO<sub>4</sub>, 200  $\mu$ M dNTP mixture 1  $\mu$ M of each primer (Gene Fanavaran, Tehran). The mixture was subjected to an initial denaturation at 94°C for 3 min, followed by 35 cycles of denaturation at 95°C for 45 s, annealing at 60°C for 45 s and extension 72°C for 4 min and a final extension of 72°C for 10 min. PCR products were visualized by UV and eluted by gel extraction kit (Fermentas, Lithuania) according to manufacturer's instructions. The purified fragments were cloned in pTZ57T according to the supplier's instructions (Fermentas, Lithuania) and sequencing was performed by a commercial facility (Gene Fanavaran, Tehran).

After the sequence verification, the fragment to be cloned in pCambia1304 containing Kanamycin resistance gene was digested with *Xho*I and ligated with similarly digested vector which had been treated with shrimp alkaline phosphatase in place of hygromycin. Orientation of the cloned gene was determined by digestion with *Eco*RI.

The vector containing *ipaB* gene with correct orientation was transformed into competent *Agrobacterium tumefaciens* strain LB4404 (kindly provided by Dr. Salmanian, National Iranian Genetic Engineering Center) by freeze and thaw method<sup>23</sup>. Transformed bacteria were cultured on LB plates containing kanamycin (50 mg/ml) and incubated at 28°C for 2 to 4 days.

Agro infiltration of leaves of tobacco plant *N. tabacum* var. *samsun* (kindly provided by Dr. Rajabi Memari, (Shahid Chamran University of Ahvaz) was performed using the method described by D. Aoust *et al*<sup>24</sup>. Briefly, a

single colony of *Agrobacterium* LBA4404 containing recombinant plasmid was cultured in 5 ml LB medium containing kanamycin (50 mg/ml) at 120 rpm at 28°C for one day. Four hundred  $\mu$ l of the culture (OD=0.8) was centrifuged for 20 min at 4°C (2500  $\times$ g). The pellet was dissolved in 1 ml of resuspension solution (10 Mm MgCl<sub>2</sub>, 10 mM 4-morpholino ethanesulfonic acid (MES), 20 g/l Sucrose, 100  $\mu$ L Acetosyringone (100  $\mu$ M) and 1% Tween 20 in MS media) and incubated for 1 hr at 28°C at 120 rpm. Tobacco leaves were immersed in bacterial suspension and transferred to a glass container attached to a vacuum pump (0.5 mbar) for 4 min. Breaking the pressure in container facilitates entry of bacteria into the leaves. The infiltrated leaves were then rinsed by sterile water and put in transparent containers and incubated at 28°C under continuous illumination for 7 days.

Expression of GFP in the leaves was investigated under UV light (350 nm), Image machine 440 CF (Kodak UK). A week after agro infiltration, the leaves were used for total protein extraction<sup>25</sup>. Briefly, tobacco leaves were ground using mortar and pestle and liquid nitrogen and 1 g of leaf-powder was added to 1 ml of protein extraction buffer containing phenylmethylsulfonyl fluoride (2 mM). The plant extract was loaded to a 12% SDS-acrylamide gel<sup>26</sup> and after electrophoresis the proteins were transferred to nitrocellulose membrane (Schleider and Schull, Germany) for Western blotting using semi-dry blotting unit according to manufacturer's instructions. The membranes were blocked with 0.5% skim milk in Tween-PBS and probed with anti-IpaB MAb (1:1000) a kind gift from Dr. Fariida Nato, Pasteur Institute of Paris. Goat anti-mouse antibody (Sigma, USA) was used as the secondary antibody (1:8000) and the membranes were developed using DAB (3-3' Diaminobenzidine Sigma, USA). The HRP conjugated anti-5xHis antibody was purchased from Qiagen and added in 1/10000 dilution.

The sequence of the gene for IpaB amplified with primers containing *Nco*I and *Xba*I

restriction sites was verified (Gene Fanavaran, Tehran) and after double digestion with these enzymes and gel purification was cloned in similarly digested and purified pBAD/gIIIa (Invitrogen, USA).

Competent *E. coli* Top 10 F' (Invitrogen, USA) was transformed with the construct and the transformed bacteria were selected on LB plates containing 50 µl ampicilin. The transformed bacteria were induced with different amounts of L-arabinose according to the supplier's specifications. SDS-PAGE and Western blotting was performed as described above with HRP-conjugated anti-myc (Invitrogen, USA).

### Results

The two amplified fragments are shown in figure 1 and the slight variation in size between the products is due to the additional sequences required for expression in tobacco.

The sequence of *ipaB* was verified after amplification and also after final cloning in plant expression vector. Prior to sequencing the orientation of the cloned gene in pCambia 1304 was determined by *EcoR* I (Figure 2) which has one site in the vector and another in the *ipaB* gene located at position 353 bp. Therefore, gene cloned with correct orientation produced a 1413 bp fragment after digestion with *EcoR* I.

The transformed *A. tumefaciens* were then selected on plates containing kanamycin and transformation was further verified by colony PCR (Figure 3).

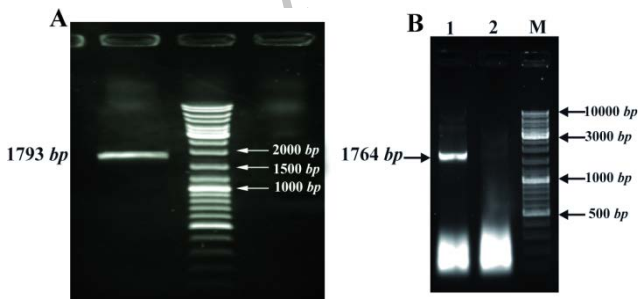


Figure 1. Amplification of *ipaB* gene by PCR with specific primers for plant. A) and *E. coli*; B) cloning. A) Lane 1; *ipaB* gene, Lane 2; MW Lane 3 negative control. B) Lane 1; *ipaB* gene, Lane 2; negative control, Lane 3; MW

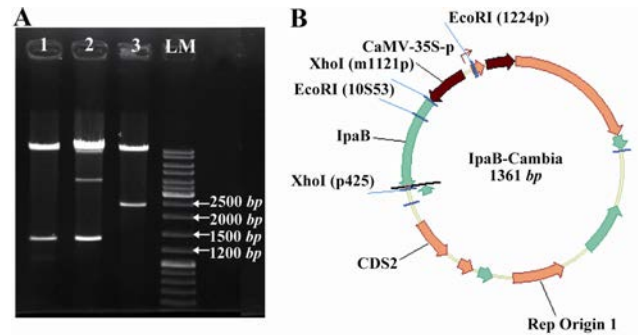


Figure 2. A) Restriction enzyme analysis of colonies of recombinant pCambia1304/ *ipaB* vector using *EcoR* I. (1-3) Lane 1: colony no. 1 containing *ipaB* in correct orientation, Lane 2: unrelated colony, Lane 3: colony no. 3 with reverse orientation; B) Schematic design of the vector containing *IpaB* gene

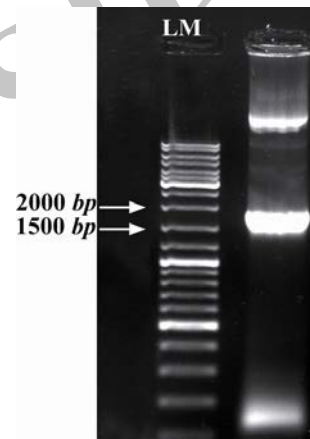


Figure 3. Colony PCR for confirmation of *Agrobacteria* grown on selective media. Lanes 1; positive clone, LM; molecular weight marker

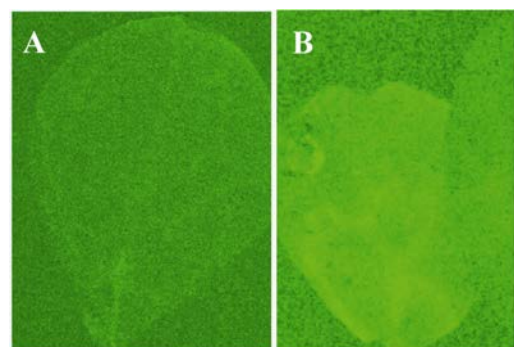


Figure 4. Verification of transfer of the expression cassette containing GFP to tobacco leaves; A) the control leaf; B) the agroinfiltrated leaf that received vector and expresses GFP

Expression of GFP was detected by UV showing that *Agrobacterium* expression cassette containing *IpaB* and *GFP* genes had been successfully transferred (Figure 4).



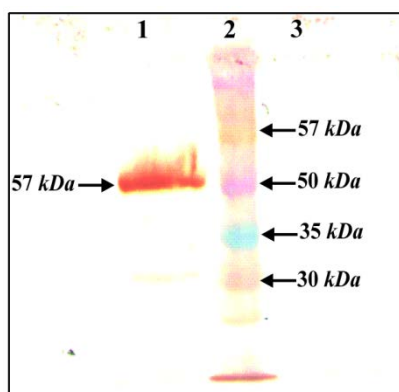


Figure 5. Western blot of leaf extract containing IpaB. Using anti His antibody, 1) infiltrated leaves, 2) MW, 3) plant negative control

Western blot analysis showed 57 kDa protein corresponding to IpaB protein that is absent in extracts of non-transformed leaves using anti His antibody (Figure 5).

After subcloning *ipaB* in pBAD/gIII expression vector and its transformation into *E. coli* (Top10), IpaB failed to be synthesized under any of the conditions used. Co-expression of the gene in the presence of plasmids containing non-specific chaperones such as DnaK, groE was also attempted, but did not result in production of the protein (data not shown).

### Discussion

Recent studies have highlighted the importance of the *Shigellae* invasion proteins especially IpaB in inducing immunity against these bacteria which are a frequent cause of dysentery in areas of poor hygiene<sup>3</sup>. These proteins are not produced by these bacteria in high abundance and are not readily purified from bacterial cultures<sup>18</sup>. Therefore, to produce and purify the IpaB protein, recombinant technology was used. However, our attempt to express this protein in *E. coli* (Top10 expression host recommended by Invitrogen) using pBAD/gIII did not result in production of IpaB protein.

It had been shown that this protein could be expressed in *E. coli* fused to thioredoxin (Trx1) a small cytoplasmic protein which has been successfully employed as a fusion part-

ner to increase the expression level and solubility of heterologous proteins in *E. coli* and seems to act mainly as a chaperone<sup>25</sup>. Therefore, it seems that expression of IpaB in *E. coli* requires the presence of a chaperone either as a fusion partner, as was demonstrated by Picking *et al* who used Trx1, or IpgC which is specific for this and IpaC protein<sup>18</sup>. Both of these strategies have disadvantages of requiring additional steps and time for cloning, separation and purification of IpaB, which lengthens the purification process and increases cost. IpaC also has been expressed in transgenic *Arabidopsis*. This protein is essential for bacterial entry into epithelial cells, capable of eliciting protective antibody responses in animal models. The expression level was 0.2% total plant protein<sup>27</sup>.

Plants can produce proteins that are difficult or impossible to produce in bacteria, because unlike bacteria, inclusion bodies are not formed in plants eliminating the need for renaturation of product, thus avoiding any loss of activity. In addition, it is possible to lead the toxic proteins to intracellular compartments to protect cells from the toxic effects of the products. Transient expression also provides this advantage in addition to fast and higher expression level of recombinant proteins. A variety of genes have been expressed using agroinfiltration including interferons, growth hormone, antibodies and vaccine candidates<sup>25</sup>. In addition, because of using non transgenic plants and short production period in controlled area, transient expression of the recombinant proteins does not produce any genetically modified foods and preventing social concerns<sup>28</sup>.

The objective of this study was transient expression of *ipaB* using agro-infiltration technique. As efficiency of agro-infiltration depends on the ability of bacterial penetration inside the leaf tissue<sup>29</sup>, the infiltration time and vacuum force is important. We found that 4 min infiltration at 0.5 mbar is sufficient for the transfer of *Agrobacteria* into tobacco leaves. Our findings also showed that in younger leaves the amount of recombinant

proteins synthesized is higher indicating that the age of leaves affects the gene expression level (data not shown).

We used Kozak and KDEL sequences for enhancing the production of recombinant proteins. Kozak sequence lies within the 5' untranslated region of eukaryotic genes and directs translation of mRNA by causing the ribosome to pause and recognize the ATG codon. KDEL is a tetra amino acid sequence which keeps the protein from secreting out of endoplasmic reticulum, thus protecting the recombinant proteins from cytoplasmic proteases as well as preventing the addition of complex type glycosides<sup>31</sup>.

pCambia1304 has been used to produce transgenic lines and contains GFP/GUS as reporter genes and hygromycin as antibiotic marker. Since in transient expression there is no need for plant selective marker we were able to use the selective marker site for expression of *ipaB* and GFP as a marker to determine the area of infiltration of the vector and the possibility of expression. To the best of our knowledge the gene replacement in this vector and the use of non-fused GFP as a reporter gene for successful transfer of *Agrobacterium* into tobacco leaves has not been previously documented.

Transient expression of recombinant proteins using agro infiltration is also more convenient for developing countries that do not have biosafety regulations for releasing the transgenic plants.

### Acknowledgement

We gratefully acknowledge the technical help on detection of GFP given by Ms. Torkashvand from the Biotechnology Department of the Pasteur Institute of Iran.

### References

1. Niyogi SK. Shigellosis. *J Microbiol* 2005;43(2): 133-143.
2. WHO. Shigellosis: disease burden, epidemiology and case management. *Wkly Epidemiol Rec* 2005; 80:93-100.
3. Martinez-Becerra FJ, Kissmann JM, Diaz-McNair J, Choudhari SP, Quick AM, Mellado-Sanchez G, et al. Broadly protective Shigella vaccine based on type III secretion apparatus protein. *Infect Immun* 2012;80(3):1222-1231.
4. Fontaine A, Arondel J, Sansonetti PJ. Role of Shiga toxin in the pathogenesis of bacillary dysentery, studied by using a Tox<sup>-</sup> mutant of Shigella dysenteriae 1. *Infect Immun* 1988;56(12):3099-3109.
5. WHO Initiative for Vaccine Research (IVR). Diarrhoeal diseases. 2009. Available from: [http://www.who.int/vaccine\\_research/diseases/diarrhoeal/en/index.html](http://www.who.int/vaccine_research/diseases/diarrhoeal/en/index.html).
6. Wong MR, Reddy V, Hanson H, Johnson KM, Tsoi B, Cokes C, et al. Antimicrobial resistance trends of Shigella serotypes in New York City, 2006-2009. *Microb Drug Resist* 2010;16(2):155-161.
7. Ye C, Lan R, Xia S, Zhang J, Sun Q, Zhang S, et al. Emergence of a new multidrug-resistant serotype X variant in an epidemic clone of Shigella flexneri. *J Clin Microbiol* 2010;48(2):419-426.
8. McKenzie R, Walker RI, Nabors GS, Van De Verg LL, Carpenter C, Gomes G, et al. Safety and immunogenicity of an oral, inactivated, whole-cell vaccine for Shigella sonnei: preclinical studies and a Phase I trial. *Vaccine* 2006;24(18):3735-3745.
9. Barnoy S, Jeong KI, Helm RF, Suvarnapunya AE, Ranallo RT, Tzipori S, et al. Characterization of WRs2 and WRs3, new second-generation virG (icsA)-based Shigella sonnei vaccine candidates with the potential for reduced reactogenicity. *Vaccine* 2010;28(6):1642-1654.
10. Katz DE, Coster TS, Wolf MK, Trespalacios FC, Cohen D, Robins G, et al. Two studies evaluating the safety and immunogenicity of a live, attenuated Shigella flexneri 2a vaccine (SC602) and excretion of vaccine organisms in North American volunteers. *Infect Immun* 2004;72(2):923-930.
11. Kotloff KL, Pasetti MF, Barry EM, Nataro JP, Wasserman SS, Sztein MB, et al. Deletion in the Shigella enterotoxin genes further attenuates Shigella flexneri 2a bearing guanine auxotrophy in a phase I trial of CVD 1204 and CVD 1208. *J Infect Dis* 2004; 190:1745-1754.
12. Rahman KM, Arifeen SE, Zaman K, Rahman M, Raqib R, Yunus M et al. Safety, dose, immunogenicity, and transmissibility of an oral live attenuated Shigella flexneri 2a vaccine candidate (SC602) among healthy adults and school children in Matlab, Bangladesh. *Vaccine* 2011;29(6):1347-1354.
13. Kotloff KL, Losonsky AG, Nataro JP, Wasserman SS, Hale TL, Taylor ND, et al. Evaluation of the safety, immunogenicity, and efficacy in healthy

- adults of four doses of live oral hybrid *Escherichia coli-Shigella flexneri* 2a vaccine strain EcSf2a-2. *Vaccine* 1995;13(5):495-502.
14. Cohen D, Ashkenazi S, Green M, Lerman Y, Slepov R, Robin G, et al. Safety and immunogenicity of investigational *Shigella* conjugate vaccines in Israeli volunteers. *Infect Immun* 1996;64(10):4074-4077.
  15. Passwell JH, Harlev E, Ashkenazi S, Chu C, Miron D, Ramon R, et al. Safety and immunogenicity of improved *Shigella* O-specific polysaccharide-protein conjugate vaccines in adults in Israel. *Infect Immun* 2001;69(3):1351-1357.
  16. Fries LF, Montemarano AD, Mallett CP, Taylor DN, Hale TL, Lowell GH. Safety and immunogenicity of a proteosome-*Shigella flexneri* 2a lipopolysaccharide vaccine administered intranasally to healthy adults. *Infect Immun* 2001;69(7):4545-4553.
  17. Tribble D, Kaminski R, Cantrell J, Nelson M, Porter C, Baqar S, et al. Safety and immunogenicity of a *Shigella flexneri* 2a Invaplex 50 intranasal vaccine in adult volunteers. *Vaccine* 2010;28(37):6076-6085.
  18. Picking WL, Mertz JA, Marquart ME, Picking WD. Cloning, expression, and affinity purification of recombinant *shigella flexneri* invasion plasmid antigens IpaB and IpaC. *Protein Expr Purif* 1996;8(4):401-408.
  19. Daniell H, Khan MS, Allison L. Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. *Trends Plant Sci* 2002;7(2):84-91.
  20. Mett V, Farrance CE, Green BJ, Yusibov V. Plants as biofactories. *Biologicals* 2008;36(6):354-358.
  21. Daniell H. Medical molecular farming: expression of antibodies, biopharmaceuticals and edible vaccines in plants. London: Kluwer Academic Publishers; 2003.
  22. Sambrook J, Fritsch ER, Maniatis T. Molecular cloning, a laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2003.
  23. Sheerman S, Bevan MW. A rapid transformation method for *Solanum tuberosum* using binary *Agrobacterium tumefaciens* vectors. *Plant Cell Rep* 1987;7(1):13-16.
  24. D'Aoust MA, Lavoie PO, Belles-Isles J, Bechtold N, Martel M, Vézina LP. Transient expression of antibodies in plants using syringe agroinfiltration. *Methods Mol Biol* 2009;483:41-50.
  25. Jurado P, de Lorenzo V, Fernandez LA. Thioredoxin fusions increase folding of a single chain Fv antibodies in the cytoplasm of *Escherichia coli*: evidence that chaperone activity is the prime effect of thioredoxin. *J Mol Biol* 2006;357(1):49-61.
  26. Lammeli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227(5259):680-685.
  27. MacRae AF, Preiszner J, Ng S, Bolla RI. Expression of His-tagged *Shigella* IpaC in *Arabidopsis*. *J Biotechnol* 2004;112(3):247-253.
  28. Maghari BM, Ardekani AM. Genetically modified foods and social concerns. *Avicenna J Med Biotech* 2011;3(3):109-117.
  29. Ma JK, Drake WMP, Christou P. The production of recombinant pharmaceutical proteins in plants. *Nat Rev Genet* 2003;4(10):794-805.
  30. Fischer R, Schillberg S. Molecular farming, plant made pharmaceuticals and technical proteins. Weinheim: Wiley-VCH Verlag GmbH & Co; 2004.
  31. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 1987;15(20):8125-8148.