



OPEN

FBPP: software to design PCR primers and probes for nucleic acid base detection of foodborne pathogens

Mohamed A. Soliman^{1✉}, Mohamed S. Azab¹, Hala A. Hussein², Mohamed M. Roushdy¹ & Mohamed N. Abu el-naga²

Foodborne pathogens can be found in various foods, and it is important to detect foodborne pathogens to provide a safe food supply and to prevent foodborne diseases. The nucleic acid base detection method is one of the most rapid and widely used methods in the detection of foodborne pathogens; it depends on hybridizing the target nucleic acid sequence to a synthetic oligonucleotide (probes or primers) that is complementary to the target sequence. Designing primers and probes for this method is a preliminary and critical step. However, new bioinformatics tools are needed to automate, specific and improve the design sets to be used in the nucleic acid–base method. Thus, we developed foodborne pathogen primer probe design (FBPP), an open-source, user-friendly graphical interface Python-based application supported by the SQL database for foodborne pathogen virulence factors, for (i) designing primers/probes for detection purposes, (ii) PCR and gel electrophoresis photo simulation, and (iii) checking the specificity of primers/probes.

Foodborne diseases are considered an important cause of mortality and morbidity and a significant obstruction in socioeconomic development worldwide¹. One of the strategies used to lower the costs and incidence of foodborne diseases is the detection of foodborne pathogens to serve a safe supply of food and inhibit foodborne diseases. The nucleic acid base detection method is one of the most rapid and widely used methods in the detection of foodborne pathogens; it depends on hybridization between the target sequence of the nucleic acid and oligonucleotide primers or probes complementary to the specific nucleic acid sequence of the targeted bacteria².

The design of primers and probes is essential, especially for nucleic acid base detection methods, and it is a preliminary and critical step. This task can be challenging and requires the identification of highly unique conserved regions of target sequences and the ability to display high sensitivity, i.e., the ability to amplify its intended target, and specificity, i.e., the ability not to amplify any nontarget. In addition, the biological parameters of the primers, such as GC content, melting temperature, and the formation of secondary structures, which include self-dimers, hairpins and cross dimers, are essential for evaluating the efficient amplification of a target sequence.

There are a number of software tools available to the public to design primers³. However, there are no specific tools for designing primer–probe sets automated for the nucleic acid detection base method. Additionally, these software programs often have limitations in their proficiencies to perform target analysis. Users are typically required to use additional tools for testing primer specificity. They are usually not specific to detect target sequences that contain a significant number of mismatches and are consequently not suitable for application in nucleic acid detection methods. There are several simulation software programs that determine the amplification of amplicons of primer pair targets supplied by users but do not design primers⁴, and some of them are difficult to use without bioinformatics knowledge⁵.

To overcome these issues, we present FBPP, an open-source Python-based application supported with the SQL database for foodborne pathogen virulence factors with the ability to (i) design primers/probes using the modified Primer3 module to be suitable for detection purposes, (ii) perform PCR and gel electrophoresis photo simulations, and (iii) check the specificity of primers/probes. We believe that this application will be particularly useful in nucleic acid base detection methods.

¹Department of Botany and Microbiology, Faculty of Science (Boys), Al-Azhar University, Nasr City, Cairo, Egypt. ²Department of Radiation Microbiology, National Centre for Radiation Research and Technology, Atomic Energy Authority, Nasr City, Cairo, Egypt. ✉email: mohamedsoliman.21@azhar.edu.eg

Method, algorithm and implementation

The FBPP program consists of four modules. The first module is `sqlite3`⁶, which is used to create and access a database for virulence foodborne pathogen genes, while the `Primer3` module⁷, with some little bite modification to be suitable for detection purposes, is used to generate the candidate primer pairs for a given template sequence. Another module for PCR and gel electrophoresis photo simulation is created by using the `Pydna` module⁸, and the last module for specificity checking uses the `BLAST` python module `Bio. Blast` from the `Biopython` project⁹ to look for matches between the primers and targets and not match with any nontarget region.

Designing primers and probes in FBPP is accomplished in four stages: first, the target sequences are identified; in most situations, they are selected from the program database or through the input section. Next, the modified `Primer3` module generates many candidate primer pairs according to the desired primer properties, which are identified by the user or using default primer properties. Then, the candidate primer pairs were subjected to PCR and gel electrophoresis simulation to calculate all amplicons from large lists of primers and ensure that the bands were not false negative results for the primers. Finally, the specificity checking process of success primers in the previous step, by default, uses the `BLAST` module with specific parameters that ensure high sensitivity such that it can detect a target that contains up to 35% mismatches to the primer sequence to exclude nonspecific primers to avoid false positives. The entire search process can be very long if each pair is searched with `BLAST` individually. To solve this problem, we saved the `BLAST` result in the entry database for each primer to avoid repeating the process in every run to the same target sequences, and in case unsaved primer goes to online `BLAST`, also to avoid being an updated program, all `BLAST` result for all primer will clear every three months to research again to online update `BLAST`.

FBPP is implemented in Python and runs on an IDLE 3.7.4 platform. It accesses an SQL database over `SQLite` 3.10.1. The program was tested on a Windows 10 Home Server with an Intel (R) Core (TM) Processor i5-4210 U with 2.40 GHz and 4 GB RAM. moreover, an EXE file is available so that the user can run the application without the need for a python interpreter, python packages, modules, or any related program such as `SQLite`.

Result

User interface

The interface consists of two tabs; the first tab is “Add New Gene”, where users can insert new virulence genes (Fig. 1-A), and the second tab, “select primer”, has several sections where users can input the gene template, option process and other user-adjustable parameters (Fig. 1-B). In the first tab, there are two ways to enter the virulence gene information into the `SQLite` database: either by browsing and uploading the `GenBank` file from the computer or fetching it from `NCBI` using the accession number.

In the second tab, the user has two options to predict the primer or probe: either by selecting the virulence gene from the program or by entering the sequence of the desired gene. When selecting the gene from the program, the gene information appears on a separate screen. Next, identify the desired properties of individual primers and pairs of primers. These properties include, for example, amplicon (PCR product) length, primer size, melting temperature and GC content. In addition to other options, users can also constrain the complementarity properties of primers to ensure that chosen primer candidates do not bind to themselves and that, in selected primer pairs, the forward primer does not bind to the reverse primer. For convenience, all default choices of parameters have been pretested and are likely to work well in detection applications.

Users have three options: generate the candidate primer pairs for a given template, generate the candidate primer pairs with PCR & gel electrophoresis simulation or the specificity checking module uses `BLAST` in case the template is used from the program. Other options include the compare button to compare the template or gene with other sequences in the `NCBI` database. can be found under the “PCR Template” section.

Presentation of results

According to the user choice, three result reports were obtained. The report of the first option presents the specificity of the generated primers, including sequence, length, position on the consensus sequence for each oligo, T_m , GC content, self-complementarity, and self-3' complementarity (Fig. 2). The report of the second choice (Fig. 3) shows the same information of the first report plus the simulation of PCR process and figure of the observed band. The result of the third option reports the additional statement of specificity checking (Fig. 3).

FBPP offers several features that are not found in other software tools. It is the only tool that specifies the prediction of primer/probe for foodborne pathogenic nucleic acid detection-based methods. In addition, it contains a database for most foodborne pathogen virulence genes; however, the program shows result simulation to avoid false negative results in detection tools.

Another advantage of the FBPP program is the ability to check the specificity of the prime/probe and the number of mismatches that a specific primer pair must have to unintended targets and a custom 3' end region where a certain number of mismatches must be present.

Discussion

There are several tools available for the design of primers and probes; however, some of them have limitations that can be mentioned. One of these limitations is that some tools, such as `Primer3` (Rozen and Skaletsky, 2000), cannot check specificity and make simulations, which can be accomplished by our program. Although it is a powerful and online tool for designing primers based on a single, short and conserved sequence, the same limitation was found with other software, such as `BatchPrimer3`¹⁰, `Primaclade`¹¹, `PrimerIdent`¹², `GeneFisher`¹³, `Gemi`⁵ and `PRISE2`¹⁴.

`Primer-BLAST`⁴ is one of most famous primer design web browsers provided by `NCBI`. It combines `Primer3` and `BLAST` along with the Needleman–Wunsch (NW) global alignment algorithm, which overcomes the short

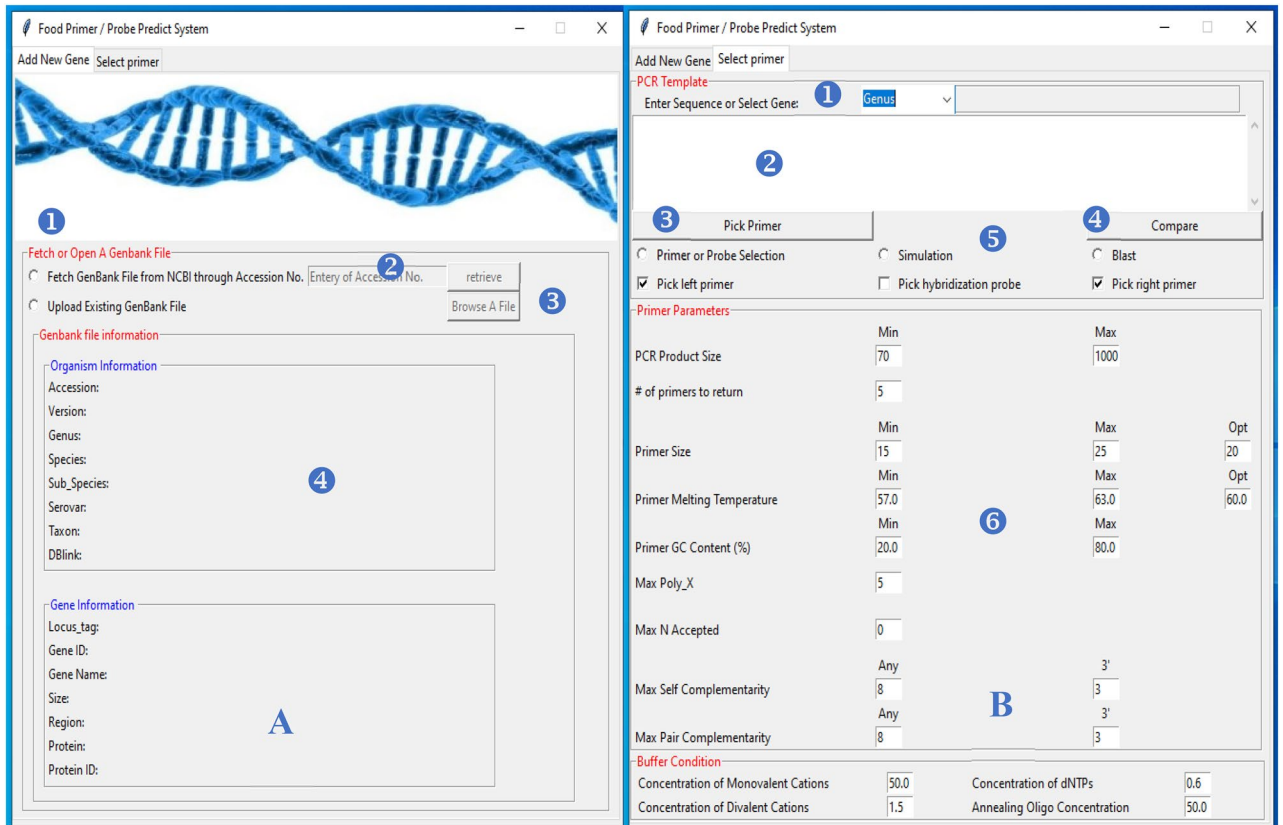


Figure 1. (A): Screenshot of “Add New Gene Tab” in the FBPP program: (1) Switch from fetch to upload gene bank file, (2) entry of the accession numbers, (3) browse for the input gene bank file, and (4) the output information of file. (B): Screenshot of “Select Primer Tab” in the FBPP program. (1) Select virulence gene from Database, (2) entry of the gene Sequence, (3) click to design the primers or probe, (4) click to compare between template or gene with other sequence in NCBI database, (5) switch between the run options, (6) contains parameters specific to the selected primers and their PCR products or probe, such as the Tm of the PCR product, the primer length, the primer GC content, GC clamps at the 3'-end of the primer, and the PCR buffer conditions.

Pick Primer									
Primer pair 1									
	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTTGATTGAAGCCGATGCCG	Plus	20	273	292	59.97	55.00	0.00	0.00
Reverse primer	TCATCGCACCCGTCARAAGAA	Minus	20	577	558	59.97	50.00	0.00	0.00
Primer pair 2									
	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTTGATTGAAGCCGATGCCG	Plus	20	273	292	59.97	55.00	0.00	0.00
Reverse primer	TTCATCGCACCCGTCARAAGAA	Minus	20	578	559	59.97	50.00	0.00	0.00
Primer pair 3									
	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTTGATTGAAGCCGATGCCG	Plus	20	273	292	59.97	55.00	0.00	0.00
Reverse primer	CGTCATGATATCCGCCCA	Minus	20	822	803	59.97	55.00	0.00	0.00
Primer pair 4									
	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTTGATTGAAGCCGATGCCG	Plus	20	273	292	59.97	55.00	0.00	0.00
Reverse primer	TCAGTCCTAACGACGACCT	Minus	20	1051	1032	59.96	55.00	0.00	0.00
Primer pair 5									
	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTTGATTGAAGCCGATGCCG	Plus	20	273	292	59.97	55.00	0.00	0.00
Reverse primer	CTGGTTTGTAGTTTGGCGGC	Minus	20	995	976	60.04	55.00	0.00	0.00

Figure 2. Example results of designing target-specific primers for the *invA* gene (virulence gene for salmonella), which contains a summary of basic properties for returned primer pairs.

References

1. Organization, W. H. WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007–2015 (World Health Organization, 2017).
2. Zhao, X., Lin, C.-W., Wang, J. & Oh, D. H. Advances in rapid detection methods for foodborne pathogens. *J. Microbiol. Biotechnol.* **24**, 297–312 (2014).
3. Abd-El Salam, K. A. Bioinformatic tools and guideline for PCR primer design. *Afr. J. Biotechnol.* **2**, 91–95 (2003).
4. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform.* **13**, 1–11 (2012).
5. Sobhy, H. & Colson, P. Gemi: PCR primers prediction from multiple alignments. *Comparative and functional genomics* **2012**, 783138 (2012).
6. Horstmann, C. *Python for Everyone* (Wiley, 2013).
7. Rozen, S. & Skaletsky, H. in *Bioinformatics Methods and Protocols* 365–386 (Springer, 2000).
8. Pereira, F. *et al.* Pydna: A simulation and documentation tool for DNA assembly strategies using python. *BMC Bioinform.* **16**, 1–10 (2015).
9. Chang, J. *et al.* Biopython tutorial and cookbook. *Update*, 15–19 (2010).
10. You, F. M. *et al.* BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinform.* **9**, 1–13 (2008).
11. Gadberry, M. D., Malcomber, S. T., Doust, A. N. & Kellogg, E. A. Primaclade—a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics* **21**, 1263–1264 (2005).
12. Pessoa, A. M., Pereira, S. & Teixeira, J. PrimerIdent: A web based tool for conserved primer design. *Bioinformatics* **5**, 52 (2010).
13. Giegerich, R., Meyer, F. & Schleiermacher, C. in *ISMB*. 68–77.
14. Huang, Y.-T., Yang, J.-I., Chrobak, M. & Borneman, J. PRISE2: Software for designing sequence-selective PCR primers and probes. *BMC Bioinform.* **15**, 1–8 (2014).
15. Arvidsson, S., Kwasniewski, M., Riaño-Pachón, D. M. & Mueller-Roeber, B. QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinform.* **9**, 1–15 (2008).
16. Xu, D., Li, G., Wu, L., Zhou, J. & Xu, Y. PRIMEGENS: Robust and efficient design of gene-specific probes for microarray analysis. *Bioinformatics* **18**, 1432–1437 (2002).
17. Srivastava, G. P., Guo, J., Shi, H. & Xu, D. PRIMEGENS-v2: Genome-wide primer design for analyzing DNA methylation patterns of CpG islands. *Bioinformatics* **24**, 1837–1842 (2008).
18. Higgins, M. *et al.* PrimedRPA: Primer design for recombinase polymerase amplification assays. *Bioinformatics* **35**, 682–684 (2019).
19. Kalendar, R., Khassenov, B., Ramankulov, Y., Samuilova, O. & Ivanov, K. I. FastPCR: An in silico tool for fast primer and probe design and advanced sequence analysis. *Genomics* **109**, 312–319 (2017).

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed M.A.S., M.S.A., H.A.H., M.M.R., M.N.A. The first draft of the manuscript was written by M.A.S. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024