

基于 PrimerHunter 的多重 PCR 引物 设计系统的构建与应用

李美婧¹, 施江程², 李建伟²

(1. 首都师范大学附属回龙观育新学校, 北京 102208; 2. 河北工业大学 计算机科学与软件学院, 天津 300401)

摘要:PrimerHunter 是一款多重 PCR 引物设计软件, 该软件能够高效地设计出具有特异性的多重引物, 但是只能在 Linux 系统的命令行中运行, 而且没有考虑引物二聚体的问题。针对这些问题, 文章建立了基于 PrimerHunter 的多重 PCR 引物设计系统, 实现了 PrimerHunter 的界面化, 完成了 C/S 模式向 B/S 模式的转换, 同时增加了原软件不具备的引物二聚体检测模块, 能够有效检测引物对之间的二聚体情况, 为相关生物学领域研究提供了一种高效的多重 PCR 引物设计工具。

关键词:多重 PCR 引物设计系统; PrimerHunter 软件; 二聚体检测

中图分类号:TP319 **文献标志码:**A **文章编号:**1672-349X(2017)06-0014-04

DOI:10.16160/j.cnki.tsxyxb.2017.06.003

The Development and Application of Multiplex PCR Primers Design System Based on PrimerHunter

LI Mei-jing¹, SHI Jiang-cheng², LI Jian-wei²

(1. Huilongguan Yuxin School Attached to Capital Normal University, Beijing 102208, China; 2. School of Computer Science and Software, Hebei University of Technology, Tianjin 300401, China)

Abstract: PrimerHunter is a multiplex PCR design software, which can efficiently design multiplex primers specificity, but it only runs in the command line of the Linux system, and it doesn't take into consideration the problem of primer dimers. To remedy these defects, the authors of this paper developed a multiplex PCR primers design system based on PrimerHunter, turning PrimerHunter into interface, changing from C/S mode to B/S mode, and adding a new primer dimers detect module. It can effectively detect the dimers between the primer pairs and function as an effective tool for the designing of multiplex PCR primers in the related field of biology.

Key Words: PCR; PrimerHunter; multiplex primers; dimer detect

基金项目:国家自然科学基金项目(81672113)

作者简介:李美婧(2000—), 女, 北京人, 高中生, 主要从事生物信息学研究。

0 引言

聚合酶链式反应(Polymerase Chain Reaction, PCR)是一种体外扩增DNA序列的技术^[1],它能够在短时间内根据极微量的DNA模板序列扩增出大量特异性DNA片段。这一技术可以应用于多个领域,如亲子鉴定、人类基因组计划以及对犯罪嫌疑人DNA的鉴别都依赖这一技术,PCR技术是现代分子生物学中最有价值的技术之一。

多重PCR(multiplex PCR),又称多重引物PCR或复合PCR,它是在同一PCR反应体系中加入两对以上的引物,同时扩增出多个目标片段的技术,它比单一PCR反应更高效,但是设计起来也更有难度。多重PCR技术主要用于多种病原微生物的同时检测或鉴定,以及某些病原微生物、某些遗传病和癌基因的分型鉴定。在应用中多重PCR技术较PCR技术可以节省大量的时间、金钱和实验材料,具有巨大的经济和时间效益。

PrimerHunter^[2]是一款多重PCR引物设计软件,该软件能够高效地设计出具有特异性的多重引物。但是只能在Linux系统的命令行中运行,操作难度较大,且没有考虑引物二聚体的问题。针对该软件的优势和缺陷,我们课题组开发出了基于PrimerHunter的多重PCR引物设计系统。用户可通过网页访问使用该系统,能更便捷地进行多重特异性引物的设计,这为生物学领域的研究提供了更高效的技术支持。

1 PrimerHunter 工具介绍

PrimerHunter是通过聚合酶链式反应设计用于鉴别病毒亚型的高灵敏度和特异性引物的工具。它可为指定的多条目标DNA序列设计出特异性强的引物,允许用户输入一系列目标DNA序列和非目标DNA序列。PrimerHunter根据一系列参数的设置,设计出能够与目标DNA序列进行特异性匹配且不会与非目标DNA序列匹配的多重引物。

设计多重PCR引物,首先要回答什么是合适的引物问题,以及这些引物的相互干扰是否

在可接受的范围内。因此,在多重PCR引物的设计问题中需要充分考虑到生物学参数分为两大部分:一是要确定引物设计的各种约束参数,根据这些约束参数来区分设计出的引物的优劣;二是要合理确定引物之间的各种约束参数,并根据这些约束参数判断设计出的引物之间的干扰是否在一个可接受的范围内。

PrimerHunter对引物设计考虑了多种约束参数,包括引物长度、引物序列、引物起始位置、引物退火温度、引物GC含量、寡核苷酸长度和GC夹等^[3]。引物间的约束参数有:引物长度差、退火温度差、发夹结构、产物长度和产物长度差。

由此可见,PrimerHunter有众多约束条件需要设置。如果在命令行中一个个地进行设置,会非常繁琐。PrimerHunter主要是由C++语言编写而成,其程序经过编译器编译、链接生成可执行文件后,用户可在命令行中调用。执行的指令如下(采用默认参数约束):`./primerhunter-tf sampleTarget.txt-nf sampleNonTarget.txt`。如果用户要改变默认的约束参数,需重新输入繁琐的参数,因此实现PrimerHunter操作的界面化是非常必要的。

在功能上,PrimerHunter只考虑了引物自身的发夹结构的检测,即引物通过自身回折使得互补的碱基对相遇,形成氢键结合而丧失引物自身的作用,并没有考虑引物之间的二聚体的检测,即引物和引物之间发生相互作用而丧失引物的作用。PrimerHunter的这一缺陷会导致设计多重引物的效率下降。

2 多重引物设计系统的功能结构

将PrimerHunter软件与Web页面结合起来,对其进行可视化编程处理,并在其基础上进行功能的完善,添加了引物二聚体的检测功能,给用户展现出一个界面友好、简单易用的多重引物设计系统。

本系统主要分为两个主要功能模块:一个是PrimerHunter引物设计模块,该模块分为两个模式,分别是引物设计模式和引物调优模式;另一个是功能模块,即二聚体检测模块,用于检

测设计出的引物之间能否形成二聚体。系统的功能结构如图 1 所示。

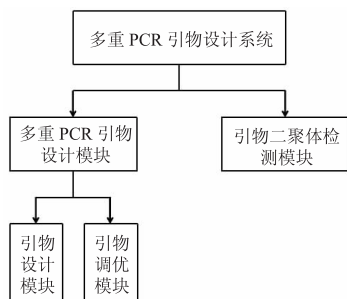


图 1 系统的功能结构图

在基于 PrimerHunter 的多重 PCR 引物设计模块中,有两种运行模式可供用户选择。

在第一种模式下,用户可以提交 Target 文件和 NonTarget 文件。例如,Target 文件中一般存放的是同一个病毒(如 H5N1 型高致病性禽流感病毒)的多条基因序列,而 NonTarget 文件中一般存放的是这个病毒的亚型(如 H9N2, H5N2 型低致病性禽流感病毒)的多条基因序列。系统能够根据用户提交的这两个文件设计出与所有 Target 文件中的目标序列都能进行杂交的引物,且保证设计出的引物没有一条能和 NonTarget 文件中的序列杂交。在这种模式下,用户可将实验中的各种参数设置的相对宽松一些,这样就可以设计出较多的引物对。

在第二种模式下,PrimerHunter 会以第一种模式下的结果文件作为输入,用户通过修改实验中的各种参数,制定更为严格的实验条件,来从第一种模式筛选出来的引物对中进行更加严格的筛选,这是对引物对的进一步调优,从而保证了引物对的设计质量。

在引物二聚体检测模块中,用户可以选择在引物设计模块中生成的引物对作为二聚体检测模块的输入,系统通过匹配算法对输入的引物序列进行两两匹配,分析它们之间有没有形成二聚体的可能。如果有可能形成二聚体,计算它们之间结合所需要的自由能,若自由能过小,则说明引物之间越容易结合成为二聚体,对于这样的情况,则需要摒弃掉。用户最终选择

的是不能结合成二聚体的引物对。如果都有可能结合成二聚体,则选择结合自由能最大的、最不稳定的引物对。

3 系统的实现及结果分析

本系统是在 LAMP(Linux-Apache-Mysql-PHP)网站架构中搭建的^[4-6],采用了 B/S 模式,即浏览器/服务器模式对用户进行展示,如图 2 所示。通过这种方式实现并优化了 PrimerHunter 的功能,又方便了用户操作,提高了设计效率。实现的系统的主界面如图 3 所示。

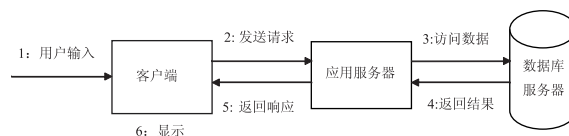


图 2 B/S 模式示意图

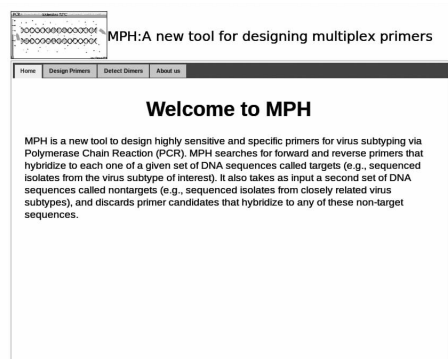


图 3 系统的主界面

3.1 引物设计模块

在系统的引物设计模块下,分别提交内容为目标 DNA 序列的 Target 文件和非目标 DNA 序列的 NonTarget 文件,文件内容均为 FASTA 标准格式,先在页面中进行相应的参数设置,然后提交即可,系统运行得到引物设计的结果文件。

在测试该模块时,我们使用了禽流感病毒的 H9 亚型的 DNA 序列作为目标 DNA 序列,将 H2 亚型的 DNA 序列作为非目标 DNA 序列,在设置完参数提交后,经过 PrimerHunter 算法的设计与筛选,最终生成了包含 15 对引物对的结果文件。

3.2 引物调优模块

如果在引物设计模块中设计出的引物对较

多,用户可将实验参数进一步地严格约束一下,然后再次将引物设计模块生成的结果文件作为该模块的输入文件,提交给系统,系统就会进行引物的调优。

把禽流感病毒亚型设计的15对引物作为该模块的输入文件,将实验参数中的温差由40℃改为2℃,再次提交给系统后,经过引物调优,最终生成的引物对的个数变成了12对,相对于引物设计模块的结果减少了3对。

3.3 引物二聚体检测模块

在引物二聚体检测模块中,用户可以输入在引物调优模块中生成的引物对,选择感兴趣的多对引物进行二聚体的检测,即检测这些引物之间能否发生相互作用。系统会对输入的引物两两之间进行相互作用可能性的检测,并将配对的引物根据碱基互补配对原则形成的二聚体展示出来。若有二聚体产生,用户则需要重新调整引物对的组成,去掉能产生二聚体的引物对,从而保证实验更有效地进行。我们在上述设计得到的12对引物中,选择较好的4对引物进行二聚体检测,发现各个引物彼此间的二聚体现象不是十分严重,可作为多重引物备选。

4 结论

基于PrimerHunter建立了多重PCR引物设计系统,将PrimerHunter软件实现了界面化,完成了从C/S模式到B/S模式的转换,使得用户操作更为方便,用户体验良好。在PrimerHunter引物设计模块的基础上,增加了引物二聚体检测模块,完善了PrimerHunter的功能。每个功能模块的运行均有结果文件生

成,方便用户浏览保存,提高了设计效率。该系统为广大生物研究人员提供了一种更高效、便捷的设计特异性强的多重PCR引物的工具,为生物学研究提供了更好的技术支持。

参考文献:

- [1] Mullis K, Faloona F, Scharf S, et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction [J]. Cold Spring Harb Symp Quant Biol, 1986, 51: 263 - 273.
- [2] Yamada T, Soma H, Morishita S. PrimerStation: a highly specific multiplex genomic PCR primer design server for the human genome [J]. Nucleic Acids Res, 2006, 34: 665 - 669.
- [3] Duitama J, Kumar D M, Hemphill E, et al. PrimerHunter: a primer design tool for PCR-based virus subtype identification [J]. Nucleic Acids Research, 2009, 37(8): 2483 - 2492.
- [4] Rychlik W. OLIGO 7 primer analysis software [J]. Methods Mol Biol, 2007, 402: 35 - 60.
- [5] 汪颖懿. LAMP环境下代码交流网站的设计与实现 [J]. 信息技术与信息化, 2015(3): 254 - 257.
- [6] 温谦. HTML+CSS网页设计与布局从入门到精通 [M]. 北京: 人民邮电出版社, 2008: 217 - 248.

(责任编辑:李秀荣)