mgtPrimer User Manual

Hong Xu
Molecular Genomics Core Facility
National Institute of Environmental Health Sciences
National Institute of Health
111 TW Alexander Dr, RTP, NC 27709

June 23, 2015

Abstract

Genetically modified mice are widely used in biomedical research. It's very important to have the mice genotyped correctly. Here we develop mgtPrimer software program to design PCR primers for mouse genotyping assays.

1 Introduction

"mgtPrimer" is a software program to design mouse genotyping primers easily. It designs PCR primers for four genotyping assays: SYBR melting curve, gel electrophoresis & LabChip, universal energy-transfer fluorescent (See Table 1).

Table 1: Mouse genotyping assay

Assay	Pros	Cons	Reference
Gel electrophoresis	Easy to design primer. Some nonspecific PCR products are allowed as long as their sizes are different from target products.	Time cost is high. Manual process cannot be converted to high-throughput protocol.	Gaw et al. 1995[1]
LabChip	Easy to design primer. Some nonspecific PCR products are allowed as long as their sizes are different from target products. Can be semi-automated.	Need new equipment. Reagents cost is high.	Linask et al. 2005[2]
SYBR melting curve	Reagent cost is the lowest.	Difficult to design primer with amplicon Tm constrain. No nonspecific PCR product is allowed. Need more primer testing than other methods.	Ririe et al. 1997[3]
Universal FRET primer	Easy to design primer without Tm or size constrains. Easy to make automatic genotype calling software.	Need two more energy-transfer-fluorescent primers. Cost is higher than SYBR melting curve. No nonspecific amplicon is allowed.	Myakishev et al. 2001[4]

mgtPrimer takes a file of sequences from both wild type and mutant mice as input. It also takes an optional file of SNP rsIDs as input. It then gets SNP information from NCBI dbSNP and maps SNPs to the DNA sequence of wild type mice. It converts sequences into GenBank format with SNPs and sequence fragments as annotated

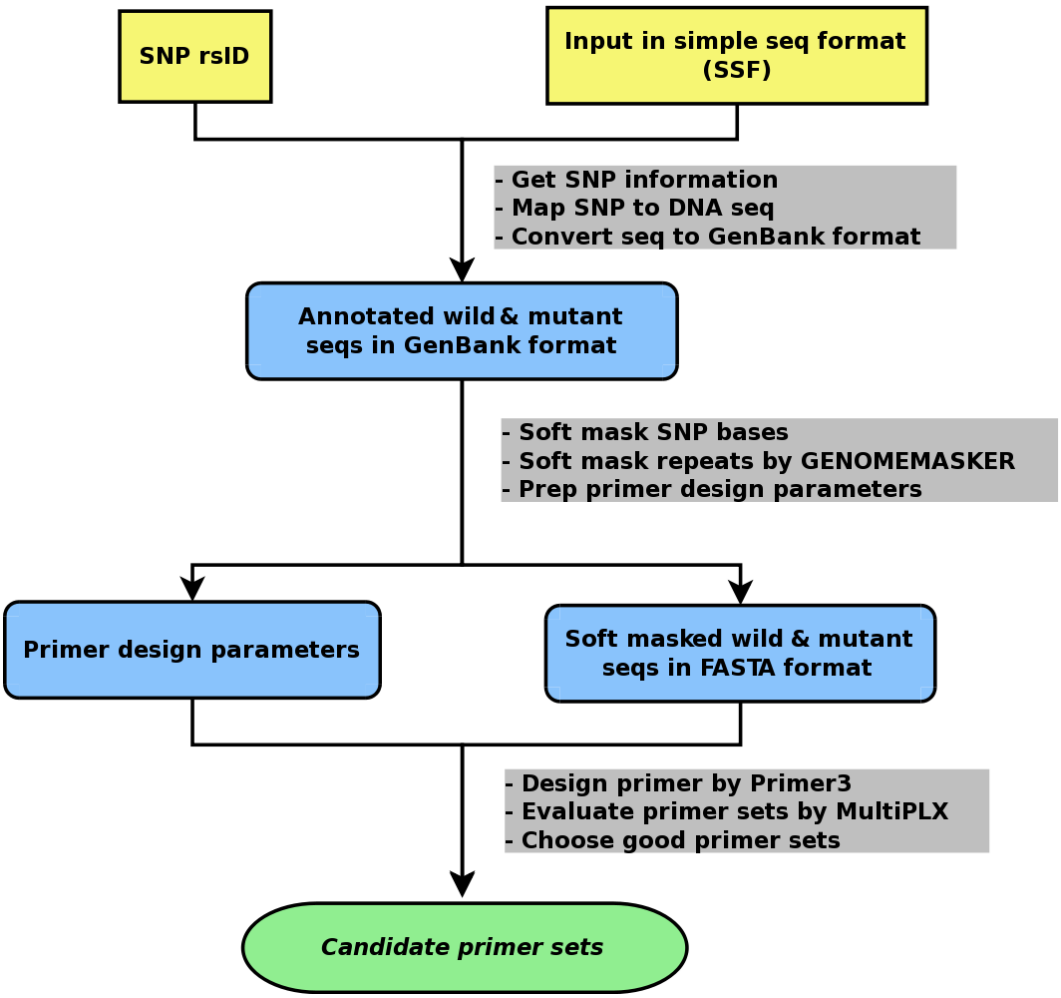

Figure 1: The Workflow of mgtPrimer

features. Before primers are designed, repeat sequences in the DNA are soft-masked by GENOMEMASKER[5]. SNPs are soft-masked as well. Based on DNA modification type, mgtPrimer calls Primer3 [6] software package to design primers automatically. After both wild type and mutant primers are generated, mgtPrimer calls MultiPLX [7] to evaluate compatibility between different wild type and mutant primer sets. The final primer result is output in tab-delimited text file for easy importing into Excel file (See Figure 1).

2 User input

The interface of mgtPrimer is very simple. It requires only three inputs:

- Email: the user email address for sending notice of primer design result when it is ready.
- Seq file: a file contains both wild type and mutant sequences in simple sequence format (SSF).
- SNP file: an optional file contains a list of SNP IDs.

2.1 Seq in simple sequence format

Simple **S**equence **F**ormat (SSF)(See Example 1) is developed for mgtPrimer program. It has two sequence entries, one for wild type DNA, the other for mutant DNA. Each sequence entry has one name field, one type field, several sequence fragment fields, and one end field. Each field is one tab-delimited line.

Example 1. *Example of SSF file*

```
Entry.name        Lyz2-Cre        wild
Entry.type        replace

Sequence.left        CTTGGGCTGCCAGAATTTCTCTCATCACATAAATGAAGAAGGAAGATCAAGTGCTGAAGTCCATAGATCGGTAG
Sequence.change        aagactctcctgactctgggactcctcctgctttctgtcactgctcaggccaaggtctatgaacgttgtgagtt

Entry.end


Entry.name        Lyz2-Cre        mutant
Entry.type        replace

Sequence.left        CTTGGGCTGCCAGAATTTCTCTCATCACATAAATGAAGAAGGAAGATCAAGTGCTGAAGTCCATAGATCGGTAG
Sequence.change        cccaagaagaagaggaaggtgtccaatttactgaccgtacaccaaaatttgcctgcattaccggtcgatgcaac

Entry.end
```

- **Entry.name** has two values - mutant mouse name (prefer JAX official name) and mouse strain (*wild* or *mutant*) .
- **Entry.type** is the mutation type, which can have one value from *deletion*, *insertion*, *flox*, *point*, *replace* (See Figure 2).
- Sequence fragments are listed in the sequencing order from DNA upstream to downstream. They can be any one of the following fragment types:
  - **Sequence.left** is the upstream fragment that is common between wild type and mutant mice. It is a required first fragment.
  - **Sequence.right** is the downstream fragment that is common between wild type and mutant mice. It will be the last fragment if available.
  - **Sequence.change** represents sequence different between wild type and mutant mice. The sequence change can be from *deletion*, *insertion*, or *replace* mutation. It should be located between **Sequence.left** and **Sequence.right**.
  - **Sequence.allele** represents point different between wild type and mutant mice. It has to be from *point* mutation. It should be located between **Sequence.left** and **Sequence.right**.
  - **Sequence.exclude** is any sequence region that is not good for primer picking. For example, the duplicate region introduced into mouse genome when gene modification is performed.
- **Entry.end** indicates the ending of a sequence entry.

2.2 SNP

SNP file is a text file (see Example 2) of a list of SNP rs IDs. Each line has one rs ID.

Example 2. *Example of SNP file*

```
rs30925746
rs30925748
rs30925750
rs30925752
rs30926644
rs30926646
```

3 Program output

It usually takes 2 10 minutes for mgtPrimer to design the primers, depending on the size of DNA sequences and how busy the server is. Users are encouraged to contact the system administer if they cannot get the result in two hours after submitting the request.

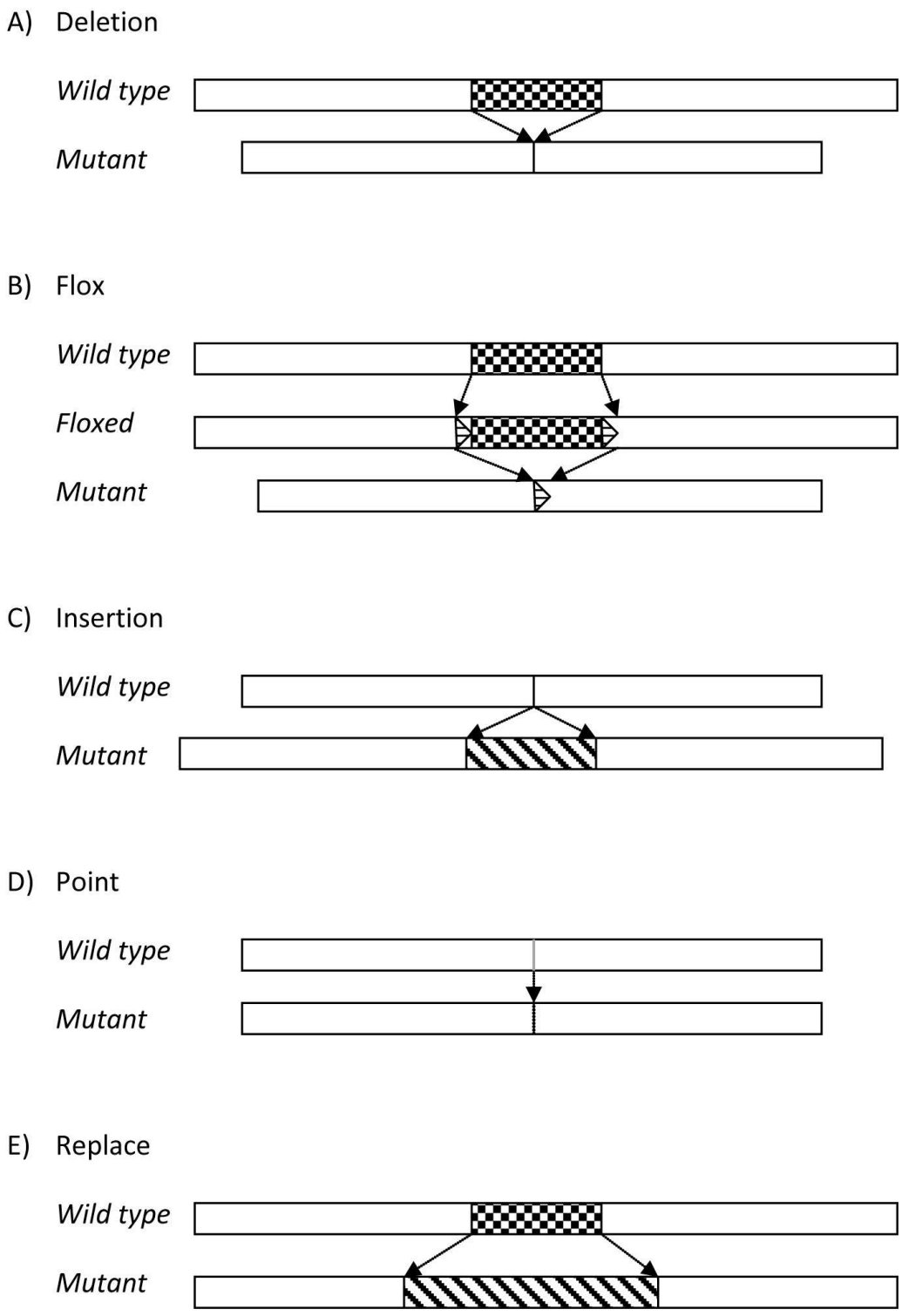

Figure 2: Mouse mutation types

3.1 Result file

The primer design result file is a tab-delimited text file, which can be viewed by Excel program. The result file usually contains three groups of primers, except that point mutation has no size different primer group. The three primer groups are:

- **Melting curve primer** is used for SYBR melting curve PCR assay. Usually Tm difference of 3 is required between wild type and mutant PCR products. If this Tm difference cannot be achieved, a GC-rich tail is added to the PCR product with the high Tm.
- **Size different primer** can be used for both gel electrophoresis or LabChip analysis. The average size difference between wild type and mutant PCR product is at least 100bp.
- **Universal energy-transfer-labelled primer** is for PCR amplification followed by fluorescence emission assay. The wild type PCR product is incorporated with fluoresein-labelled primer, which emits green fluorescence light. The mutant PCR product is incorporated with sulforhodamine-labelled primer, which emits red fluorescence light.

References

[1] Gaw, A., Mancini, F. P., and Ishibashi, S. October 1995 *Laboratory Animals* **29(4)**, 447–449.

[2] Linask, K. L. and Lo, C. W. February 2005 *BioTechniques* **38(2)**, 219–223 PMID: 15727128.

[3] Ririe, K. M., Rasmussen, R. P., and Wittwer, C. T. February 1997 *Analytical biochemistry* **245(2)**, 154–160 PMID: 9056205.

[4] Myakishev, M. V., Khripin, Y., Hu, S., and Hamer, D. H. January 2001 *Genome research* **11(1)**, 163–169 PMID: 11156625.

[5] Andreson, R., Reppo, E., Kaplinski, L., and Remm, M. (2006) *BMC bioinformatics* **7**, 172 PMID: 16566824.

[6] Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S. G. June 2012 *Nucleic acids research* PMID: 22730293.

[7] Kaplinski, L., Andreson, R., Puurand, T., and Remm, M. April 2005 *Bioinformatics* **21(8)**, 1701–1702.