5' Primer Tags: A cheap way to fluorescently tag PCR products

Travis C. Glenn

Address:	Savannah River Ecology Lab Drawer E Aiken, SC 29802	Department of Biological Sciences University of South Carolina Columbia, SC 29208
Telephone:	(803) 725-5746 (5881 – lab)	(803) 777-4606
Fax:	(803) 725-3309	(803) 777-4002
E-mail:	Glenn@srel.edu	Travis.Glenn@sc.edu

Written: August, 2001 Last Updated: October 2006

Using "automated" DNA sequencers to score microsatellites requires that a fluorescent molecule (i.e., a "tag") be added. The most straightforward way to do this is to add a fluorescent molecule to the 5' end of one of the primers used in PCR. This is simple conceptually, and in practice. Unfortunately, the addition of this tag is somewhat expensive - $60 - $75 for first generation ABI dyes (e.g., FAM), and upwards of $100+ for more recent dyes (NED, VIC, PET). The same is true or even worse for Licor, CEQ, and other platforms. Of course you have to add the cost of the primer itself, so a decent estimate of cost is ~$100 per fluorescently tagged primer.

Unfortunately, when developing loci, many of them don't work for one reason or another – e.g., the primer binding sites occur in multiple places in the genome, the piece of DNA is particularly difficult to amplify, the primers are poor, etc. At ~$100 per locus, the expense associated with primers can add up in a hurry. In the field of molecular ecology and large-scale mapping, this expense can be a substantial portion of the project budget. So, alternative methods are desirable.

A method that has been used extensively since the early 1990’s is to add a known sequence to the 5' end of one of the primers (e.g., refs in Boutin-Ganache et al. 2001). This has the effect of incorporating a known "sequence tag" to one end of the PCR amplicons. By adding a 3rd fluorescently labeled primer with the known tagging sequence to the PCR, amplicons can be generated with a fluorescent molecule Saying this another way, 3 primers are used for PCR – a normal locus specific primer, a locus specific primer with the 5’ tag, and a fluorescently labeled primer that is identical in sequence to the 5’ tag. To see how it works for yourself, you should diagram 3 cycles of PCR. Most workers have used the M13 Forward or M13 Reverse primer sequences for this purpose.

Obviously, you don't gain anything if you are only doing one locus. In fact, it is more expensive, because you are purchasing a longer unlabeled primer in addition to the labeled primer. However, if you are doing many loci, you can simply obtain one labeled (expensive) primer, and many longer unlabeled (cheap) primers. Assuming the cost of additional unlabeled bases is $0.50 per base, adding an additional 20-base "sequence-tag" costs $10. Thus, you can save >$50 per locus. Additionally, Boutin-Ganache et al. (2001) report superior scorability of amplicons from this method vs. the traditional method.

Because any particular sequence you might choose (e.g., M13 Forward) may form substantial secondary structures with the locus-specific primers (see below), it is good to have a small number of fluorescently tagged primers to choose from. In my experience, the M13 Forward sequence forms more secondary structures than the M13 Reverse. Thus, we never use the M13 Forward sequence as a primer tag.

I designed a primer in 2001 that is short, has high primer efficiency, has low 3' stability, and has few dimmers (i.e., it has all the characteristics you want). We have now used this sequence on thousands of primers. It works well on most taxa (note: it is always prudent to try it on ~20 primer pairs with any new organism to ensure it will work well for that critter before ordering large quantities of primers). In practice we now use this sequence almost exclusively.

```
CAGTCGMsatPrimerTag: 5' - CAGTCGGGCGTCATCA - 3'
Tm = 67.4; PE = 446; 3' stability = -6.9; no stems;
no 3' dimers; overall dimer = -3.6
```

The comparable characteristics of the M13 Reverse primer are:

```
M13 Reverse 5' - GGAAACAGCTATGACCATG - 3'
Tm = 62.4; Primer Efficiency (PE) = 351; 3' stability = -8.4;
no stems; 3' dimers = -5.3; overall dimer = -6.3
```

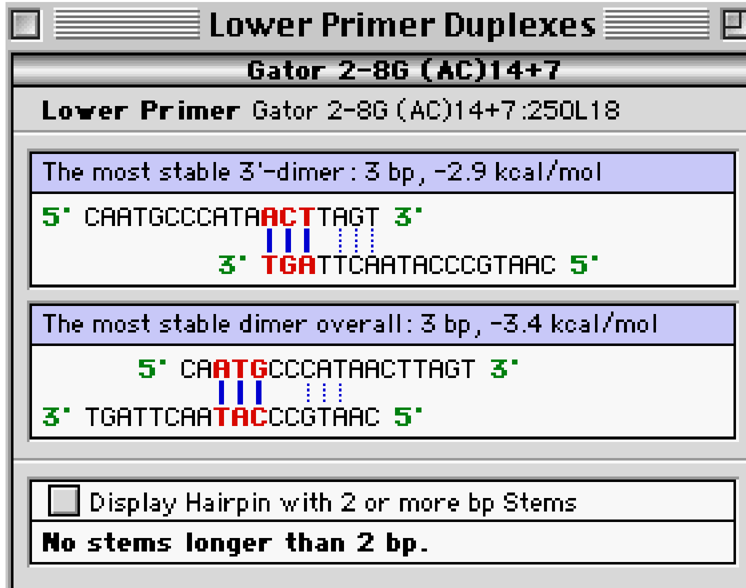
Note that the 3' terminal "CATG" forms a perfect complement with itself. Thus, leaving off the 3' "G" is probably a good idea. That results in the following:

```
M13 Reverse 5' - GGAAACAGCTATGACCAT - 3'
Tm = 59.1; PE = 338; 3' stability = -7.8; no stems;
3' dimers = -3.4; overall dimer = -6.3
```

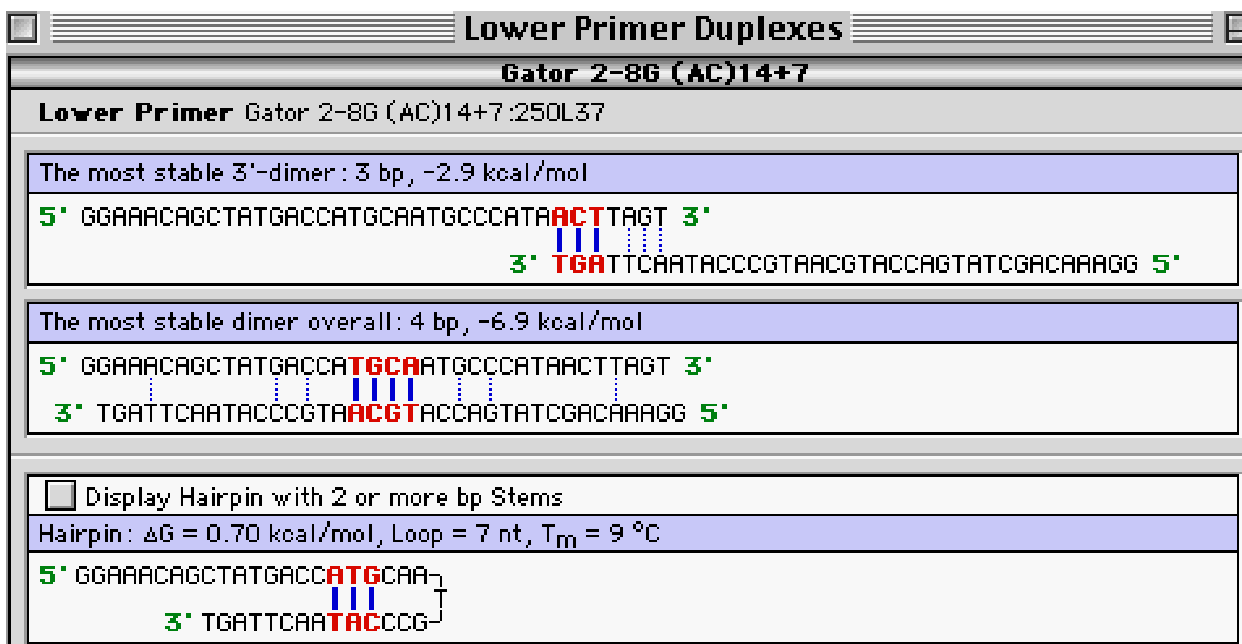
Note that this primer is 18 bases long, has a moderate Tm, a low primer efficiency, and some modest dimers.

By using these two primers, I can almost always achieve a primer design with a 5' tag that is acceptable. However, it is still important to check & see which one will work best. Here is an example where adding the M13Reverse is OK, but adding the CAGTCG tag is bad.

Ami-μ215; clone 2-8G, 187 bp, (AC)14+7 repeat, Tm's – 61, 57
*Upper: GCCACTTCATGCTATCTACT
*Lower: GGAAACAGCTATGACCATG CAATGCCCATAACTTAGT



Here is the same primer pair, but adding the M13Reverse primer tag:
Lower: GGAAACAGCTATGACCATG CAATGCCCATAACTTAGT



Here is the same primer but adding the CAGTCG primer tag:
Lower: CAGTCGGGCGTCATCA CAATGCCCATAACTTAGT

Lower Primer Duplexes

Gator 2-8G (AC)14+7

Lower Primer Gator 2-8G (AC)14+7:250L34

The most stable 3'-dimer: 3 bp, -2.9 kcal/mol

```
5' CAGTCGGGCGTCATCACAATGCCCATAACTTAGT 3'
                                  |||  :::
                             3' TGATTCAATACCCGTAACACTACTGCGGGCTGAC 5'
```

The most stable dimer overall: 4 bp, -9.3 kcal/mol

```
      5' CAGTCGGGCGTCATCACAATGCCCATAACTTAGT 3'
              ||||  :   : :    :  ::::
3' TGATTCAATACCCGTAACACTACTGCGGGCTGAC 5'
```

Display Hairpin with 2 or more bp Stems

Hairpin: ΔG = -4.80 kcal/mol, Loop = 11 nt, T_m = 85 °C

```
      5' CAGTCGGGCGTCAT┐
              ||||       C
3' TGATTCAATACCCGTAACA┘
```

Both primers give you a secondary structure where 4-bases form a stem-loop. However, you can see that the Tm for the last set is much higher, and will interfere with most PCRs.

My general approach is to try to add the CAGTCG sequence to the shorter of the upper or lower (or forward or reverse – depending on your nomenclature). If that forms a bad secondary structure, then I attempt to add the CAGTCG sequence to other primer. Note that if your locus specific primer has a 5' base or bases that are identical the 3' end of the CAGTCG or M13R sequence, you can omit them from the primer tag. Also, note that you can add any additional base or bases you want after the 3' end of the primer tag sequence and before the 5' base of the locus specific bases. I wouldn't overdo this, but adding a base or two can do remarkable things to reduce the secondary structure of some primers.

So, you are probably wondering, "What are the downsides of this strategy?"

1) A slightly more difficult process for primer design (see above) & PCR optimization. For the PCR, we usually follow this recipe (uses a 10:1 labeled to unlabeled primer ratio). To optimize it, you may need to modify that ratio.
The 20 µL reaction recipe (per reaction) is:

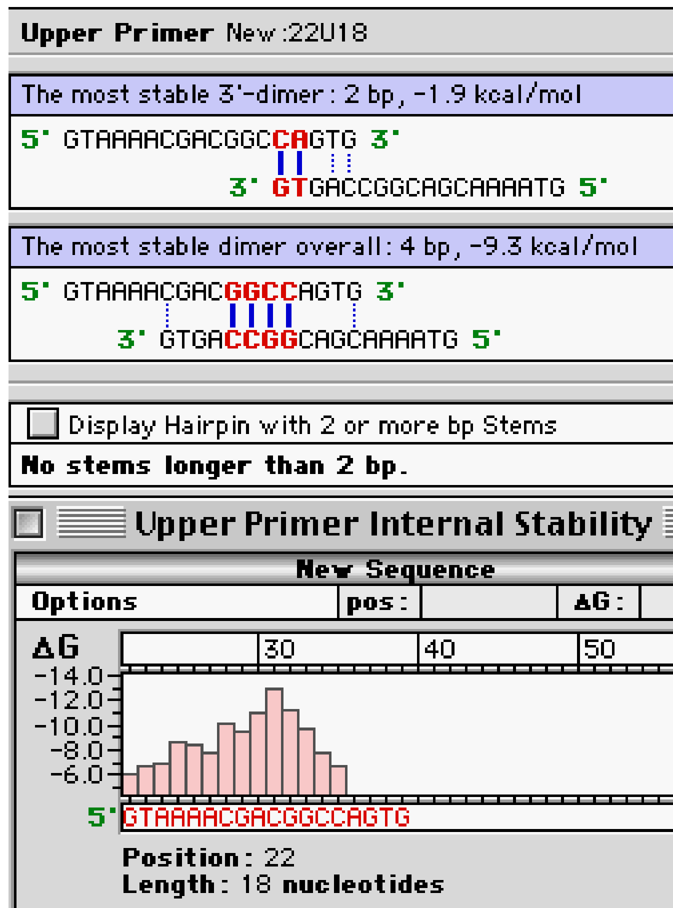
2 µL	250µg/mL BSA (Bovine Serum Albumin)
2 µL	10X PCR reaction buffer
1 µL	5 µM short locus specific primer
1 µL	5 µM Fluorescent universal primer (e.g., M13R or CAG)
1 µL	0.5 µM M13 reverse primer
1.2 µL	25 mM $MgCl_2$
1.2 µL	2.5mM dNTP's (2.5 mM each)
8.5 µL	dH_2O
0.1 µL	Taq DNA Polymerase
2 µL	DNA template (20 ng/µL)

2) You can only multiplex different sized products. You can't do multiplex PCRs with differing fluorescent primers. Of course, you can combine the amplicons after PCR, but before electrophoresis. Thus, this strategy is a good idea for development, small-scale projects, and projects where lots of loci, but only a modest number of individuals (i.e., a few hundred or less) will be investigated. It is not a good approach for large-scale studies where a great numbers of samples/individuals will be investigated (i.e., thousands or more), especially for a modest number of loci – where the efforts to produce multiplex PCR panels really pays off.

From the IDT web site http://www.idtdna.com/program/catalog/ReadyMade_Primers.asp

```
M13 Forward (-20)   GTAAAACGACGGCCAGTG
M13 Forward (-41)   GGTTTTCCCAGTCACGAC
M13 Reverse (-27)   GGAAACAGCTATGACCATG
M13 Reverse (-48)   AGCGGATAACAATTTCACAC
```

Here is how they look by themselves (homodimers, stems and internal stability):

M13 Forward –20: 

Upper Primer New:1U18

The most stable 3'-dimer: 3 bp, -2.9 kcal/mol

5' GGTTTTCCCAGTCACGAC 3'
3' CAGCACTGACCCTTTTGG 5'

The most stable dimer overall: 2 bp, -3.6 kcal/mol

5' GGTTTTCCCAGTCACGAC 3'
3' CAGCACTGACCCTTTTGG 5'

Display Hairpin with 2 or more bp Stems

No stems longer than 2 bp.

Upper Primer Internal Stability

New Sequence

Options | **pos:** 14 | **ΔG:** -7.8

ΔG -14.0 -12.0 -10.0 -8.0 -6.0

1 10 20 30

5' GGTTTTCCCAGTCACGAC

Position: 1
Length: 18 **nucleotides**

M13 Forward –41:

Upper Primer New:13U19

The most stable 3'-dimer: 2 bp, -1.3 kcal/mol

5' CACGACGTTGTAAAACGAC 3'
3' CAGCAAAATGTTGCAGCAC 5'

The most stable dimer overall: 4 bp, -6.8 kcal/mol

5' CACGACGTTGTAAAACGAC 3'
3' CAGCAAAATGTTGCAGCAC 5'

Display Hairpin with 2 or more bp Stems

Hairpin: ΔG = -2.30 kcal/mol, Loop = 4 nt, T_m = 59 °C

5' CACGACGTTGT
3' CAGCAAAA

Upper Primer Internal Stability

New Sequence

Options | **pos:** 27 | **ΔG:** -7.8

ΔG -14.0 -12.0 -10.0 -8.0 -6.0

20 30 40

5' CACGACGTTGTAAAACGAC 3

Position: 13
Length: 19 **nucleotides**

M13 Forward (from Marni [Licor]):

----

http://bio.licor.com/Accessories/BioAccess8.html

Licor IR800 Dye-labeled, ready made primers. M13 Forward can be obtained in large quantities. This is the same sequence used by Boutin-Ganache (2001)

M13 Forward (-29)19-mer: 5'- CACGACGTTGTAAAACGAC -3

LicorSeqs

Upper Primer LicorSeqs:1U19

The most stable 3'-dimer: 2 bp, -1.3 kcal/mol

```
5' CACGACGTTGTAAAACGAC 3'
         3' CAGCAAAATGTTGCAGCAC 5'
```

The most stable dimer overall: 4 bp, -6.8 kcal/mol

```
5' CACGACGTTGTAAAACGAC 3'
    3' CAGCAAAATGTTGCAGCAC 5'
```

Display Hairpin with 2 or more bp Stems

Hairpin: ΔG = -2.30 kcal/mol, Loop = 4 nt, T_m = 59 °C

```
5' CACGACGTTGT┐
    3' CAGCAAAA┘
```

M13 Reverse 20-mer: 5'- GGATAACAATTTCACACAGG -3'

LicorSeqs

Upper Primer LicorSeqs:23U20

No 3'-terminal dimer formation

The most stable dimer overall: 4 bp, -5.3 kcal/mol

```
   5' GGATAACAATTTCACACAGG 3'
3' GGACACACTTTAACAATAGG 5'
```

Display Hairpin with 2 or more bp Stems

No stems longer than 2 bp.

T3 Promoter 18-mer: 5'- AATTAACCCTCACTAAAG -3'

LicorSeqs

Upper Primer LicorSeqs:46U18

The most stable 3'-dimer: 2 bp, -1.6 kcal/mol

```
5' AATTAACCCTCACTAAAG 3'
           ||       : :
      3' GAAATCACTCCCAATTAA 5'
```

The most stable dimer overall: 4 bp, -5.3 kcal/mol

```
               5' AATTAACCCTCACTAAAG 3'
                  ||||
3' GAAATCACTCCCAATTAA 5'
```

Display Hairpin with 2 or more bp Stems

Hairpin: ΔG = 1.20 kcal/mol, Loop = 8 nt

```
5' AATTAACCC┐
      |||   │
3' GAAATCACT┘
```

T7 Promoter 20-mer: 5'- TAATACGACTCACTATAGGG -3'

LicorSeqs

Upper Primer LicorSeqs:67U20

No 3'-terminal dimer formation

The most stable dimer overall: 6 bp, -6.7 kcal/mol

```
5' TAATACGACTCACTATAGGG 3'
           : |||||| :
      3' GGGATATCACTCAGCATAAT 5'
```

Display Hairpin with 2 or more bp Stems

Hairpin: ΔG = 1.60 kcal/mol, Loop = 8 nt

```
   5' TAATACGAC┐
        |||    │
3' GGGATATCACT┘
```

SP6 Promoter 19-mer: 5'- GATTTAGGTGACACTATAG -3'

LicorSeqs

Upper Primer LicorSeqs:90U19

The most stable 3'-dimer: 6 bp, -6.7 kcal/mol

```
5' GATTTAGGTGACACTATAG 3'
                ||||||
         3' GATATCACAGTGGATTTAG 5'
```

The most stable dimer overall: 6 bp, -6.7 kcal/mol

```
5' GATTTAGGTGACACTATAG 3'
                ||||||
         3' GATATCACAGTGGATTTAG 5'
```

Display Hairpin with 2 or more bp Stems

Hairpin: ΔG = 1.70 kcal/mol, Loop = 6 nt

```
5' GATTTAGGTG-┐
       |||    |
 3' GATATCACA-┘
```

Thus, the T3 promotor is probably your best bet for LiCor tagged primers available in bulk (their M13 reverse primer is also fine, but 2 bases longer; so it would be a good alternative).

References

Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF. 2001. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. BioTechniques 31:24-28.