SUPPLEMENTARY MATERIAL

Pyrosequencing Oligodeoxyribonucleotides
iC = 5-methylisocytidine; iG = isoguanosine

Common Primer
5'-ATGGCACCGACGGCTGGCAG-3'

ODNs in Figure 2
A 3'-TACCGTGGCTGCCGACCGTCCTATGTATC-5'
B,C 3'-TACCGTGGCTGCCGACCGTCCATAiCATAC-5'
D,E 3'-TACCGTGGCTGCCGACCGTCCTGTiGTGTC-5'
F,G 3'-TACCGTGGCTGCCGACCGTCCTGTiCTGTC-5'

ODNs in Figure S1
A,B 3'-TACCGTGGCTGCCGACCGTCTCACiGCATC-5'
C,D 3'-TACCGTGGCTGCCGACCGTCATCAiGACTC-5'
E,F 3'-TACCGTGGCTGCCGACCGTCCATAiGATAC-5'
G,H 3'-TACCGTGGCTGCCGACCGTCCTATiGTATC-5'
I,J 3'-TACCGTGGCTGCCGACCGTCTCGCiGCGTC-5'
K,L 3'-TACCGTGGCTGCCGACCGTCCGTGiGGTGC-5'

ODNs in Figure S2
A,B 3'-TACCGTGGCTGCCGACCGTCTCACiCCATC-5'
C,D 3'-TACCGTGGCTGCCGACCGTCATCAiCACTC-5'
E,F 3'-TACCGTGGCTGCCGACCGTCTCGCiCCGTC-5'
G,H 3'-TACCGTGGCTGCCGACCGTCGTCGiCGCTC-5'
I,J 3'-TACCGTGGCTGCCGACCGTCCGTGiCGTGC-5'
K,L 3'-TACCGTGGCTGCCGACCGTCCTATiCTATC-5'

Dye Terminator ODNs
Dye terminator ODN templates consist of a 170 nucleotide common region, including a primer binding sequence, constructed from a ligation reaction. A variable sequence region, appended to the common region in the ligation reaction, is the template region shown in the data presented in the figures.
iC = 5-methylisocytidine; iG = isoguanosine; P = phosphate

Common Primer
5'-ATGGCACCGACGGCTGGCAG-3'

ODNs to construct 170 nucleotide common template region
3'-TACCGTGGCTGCCGACCGTCATCCGCAAGATCGTGTCTGTGGTTCAGTTGC
ACGCTCGCGTCAP-5'
3'-GGACCCTCGCTTCCTTCATCTGCACCTTCTGCACTTAGGCCGCTATACACA
CAAGGCATP-5'

3’-CTGGAGTTCAACCCGAATCTCAAGCTCATCGTCACTGTTACTGTTGCTGCP-5’

Variable ODN region in Figure 3, Figure S4, and Figure S5
3’-TTCGTGCiGTiGAACiCATGiCCGCiGAiCTGATTTTTCiGTiGAACiCATGiCCGCiGAiCTGACATCTA-5’

Variable ODN regions in Figure 4
A,B 3’-TGAAAGCiCATGTCAGCiCCTGTCAGCiCGTGTCAGCiCTTTGTCAG-5’
C,D 3’-TGAAAGAiGATGTCAGAiGCTGTCAGAiGGTGTCAGAiGTTGTCAG-5’

Variable ODN regions in Figure S6
A,B 3’-TGAAAGAiCATGTCAGAiCCTGTCAGAiCGTGTCAGAiCTTTGTCAG-5’
C,D 3’-TGAAAGGiCATGTCAGGiCCTGTCAGGiCGTGTCAGGiCTTTGTCAG-5’
E,F 3’-TGAAAGTiCATGTCAGTiCCTGTCAGTiCGTGTCAGTiCTTTGTCAG-5’

Variable ODN regions in Figure S7
A,B 3’-TGAAAGCiGATGTCAGCiGCTGTCAGCiGGTGTCAGCiGTTGTCAG-5’
C,D 3’-TGAAAGGiGATGTCAGGiGCTGTCAGGiGGTGTCAGGiGTTGTCAG-5’
E,F 3’-TGAAAGTiGATGTCAGTiGCTGTCAGTiGGTGTCAGTiGTTGTCAG-5’

Variable ODN regions in Figure S8
A 3’-TGAAAGAiGATGTCAGAiGCTGTCAGAiGGTGTCAGAiGTTGTCAG-5’
B 3’-TGAAAGTiCATGTCAGTiCCTGTCAGTiCGTGTCAGTiCTTTGTCAG-5’

Figure S1. Additional pyrosequencing results of ODNs containing disoG. All primer binding regions were identical and a 9-nucleotide template region was varied. Nucleotide dispensations are indicated on the x-axis and relative peak height of light emitted from pyrophosphate release is the y-axis. Consecutive dispensations of a single nucleotide are cumulatively tallied in a single bar. At least one negative control dispensation of a non-complementary nucleotide was made during pyrosequencing of each template. These negative control dispensations resulted in no pyrophosphate release and have been omitted for clarity. Template sequences are the reverse complements of the dispensation nucleotides on the x-axis. All templates containing disoG suffered from incomplete incorporation of d^{Me}isoC nucleotide opposite disoG, as seen in templates **A**) 3’-TCACiGCATC-5’, (**C**) 3’-ATCAiGACTC-5’, (**E**) 3’-CATAiGATAC-5’, (**G**) 3’-CTATiGTATC-5’, (**I**) 3’-TCGCiGCGTC-5’ and (**K**) 3’-CGTGiGGTGC-5’. The amount of d^{Me}isoC nucleotide incorporated varied with sequence (compare (**G**) and (**I**)). Consecutive dispensations of d^{Me}isoCTP opposite disoG template positions improved incorporation only slightly, if at all, in the same series of templates in (**B**), (**D**), (**F**), (**H**), (**J**), and (**L**).

Figure S2. Additional pyrosequencing of ODNs containing d^{Me}isoC. All primer binding regions were identical and a 9-nucleotide template region was varied. Nucleotide dispensations are indicated on the x-axis and relative peak height of light emitted from

pyrophosphate release is the y-axis. Consecutive dispensations of a single nucleotide are cumulatively tallied in a single bar. At least one negative control dispensation of a non-complementary nucleotide was made during pyrosequencing of each template. These negative control dispensations resulted in no pyrophosphate release and have been omitted for clarity. Template sequences are the reverse complements of the dispensation nucleotides on the x-axis. All templates containing d^{Me}isoC directed complete incorporation of disoG nucleotide, as seen in templates (**A**) 3'-TCACiCCATC-5', (**C**) 3'-ATCAiCACTC-5', (**E**) 3'-TCGCiCCGTC-5', (**G**) 3'-GTCGiCGCTC-5', (**I**) 3'-CGTGiCGTGC-5', and (**K**) 3'-CTATiCTATC-5'. However, further extension was compromised by the d^{Me}isoC-disoG pair and incomplete incorporation was observed at subsequent template positions. Consecutive dispensations of the same complementary nucleotide at subsequent template positions usually improved incorporation in the same series of templates, as in (**B**), (**D**), (**F**), (**H**), (**J**) and (**L**). Template sequences with dT following disoG demonstrated misincorporation of disoG nucleotide opposite template dT following correct incorporation opposite template d^{Me}isoC, as in (**K**) and (**L**).

Figure S3. Control dispensations in pyrosequencing. Unprocessed signals for two of the ODNs in the pyrosequencing experiments presented in Figures 2, S1, and S2. Signals corresponding to control dispensations are included. Dispensations of disoGTP and d^{Me}isoCTP are uncorrected for residual pyrophosphate. Signals corresponding to sequential dispensations of the same nucleotide are numbered and shown individually, instead of in a cumulative bar as in Figures 2, S1, and S2. The data used to create (**A**) Figure S1L with template 3'-CGTGiGGTGC-5'and (**B**) Figure S2D with template 3'-ATCAiCACTC-5' are shown as examples. For both templates the first two signals (red) are negative control dispensations of out-of-sequence nucleotides. The remaining signals (blue) are from nucleotide dispensations that are complementary to the template ODN sequence. Negative control dispensations of natural nucleotides resulted in essentially no observable signal in all cases, reinforcing that no significant misincorporation occurs among the natural nucleobases in pyrosequencing. The first four template positions always yielded relative signal heights typical of pyrosequencing with all natural nucleobases and form a baseline for comparison.

A negative control dispensation of d^{Me}isoCTP was made in (**A**). A signal of roughly the same magnitude was seen in d^{Me}isoCTP dispensations with all additional templates examined, independent of template sequence. A signal of similar magnitude was also observed for dispensations in control reactions in the absence of template. Thus, d^{Me}isoCTP dispensations always contained a portion of signal from a very small amount of contaminating residual pyrophosphate. The signal height corresponding to contaminating pyrophosphate was subtracted from d^{Me}isoCTP dispensations to obtain the signal associated with d^{Me}isoCTP incorporation (Figure S1L).

An analogous control dispensation of disoGTP was made in (**B**). The signal height indicates the presence of more contaminating pyrophosphate in disoGTP than was observed in d^{Me}isoCTP. The magnitude of signals observed for control disoGTP dispensations with all templates examined or in the absence of template was similar, with the exception of dispensation of disoGTP opposite template dT. Dispensation of disoGTP

opposite dT produced a much larger signal, indicating that a substantial amount of disoGTP was misincorporated opposite T. The signal height corresponding to contaminating pyrophosphate was subtracted from disoGTP dispensations to obtain the signal associated with d^{Me}isoCTP incorporation (Figure S2D).

Figure S4. Dye terminator sequencing of ODNs containing d^{Me}isoC and disoG using the BigDye 3.0 kit. Template disoG (black arrows) and d^{Me}isoC (red arrows) are indicated. All other template positions are the complements of the terminators identified. The concentration of d^{Me}isoCTP was fixed at 100 μM and concentrations of disoGTP were varied. More ddA terminator was incorporated opposite d^{Me}isoC positions (red arrows) as the concentration of disoG nucleotide was decreased from (**A**) 200 μM disoGTP, to (**B**) 100 μM disoGTP, to (**C**) 25 μM disoGTP. Incorporation of terminators opposite template disoG positions (black arrows) was unchanged while the concentration of d^{Me}isoCTP was maintained at 100 μM. Signal attenuation caused by unwanted strand termination was increased in (**A**) at 200 μM disoGTP.

Figure S5. Dye terminator sequencing of ODNs containing d^{Me}isoC and disoG using the BigDye 3.0 kit. Template disoG (black arrows) and d^{Me}isoC (red arrows) are indicated. All other template positions are the complements of the terminators identified. The concentration of disoGTP was fixed at 100 μM and concentrations of d^{Me}isoCTP were varied. More ddT terminator was incorporated opposite disoG positions (black arrows) as the concentration of d^{Me}isoC nucleotide was decreased from (**A**) 200 μM d^{Me}isoCTP, to (**B**) 100 μM d^{Me}isoCTP, to (**C**) 25 μM d^{Me}isoCTP. Incorporation of terminators opposite template d^{Me}isoC positions (red arrows) was unchanged while the concentration of disoGTP was maintained at 100 μM. Signal attenuation caused by unwanted strand termination was increased in (**A**) at 200 μM d^{Me}isoCTP.

Figure S6. Dye terminator sequencing of three ODNs used to examine d^{Me}isoC in all possible natural nearest neighbor contexts (see Figure 4 for an additional sequence). ODNs containing four d^{Me}isoC positions (red arrows) were sequenced in a BigDye 3.1 reaction in the presence of (**A**), (**C**), (**E**) 300 μM disoGTP and 0 μM d^{Me}isoCTP or (**B**), (**D**), (**F**) 0 μM disoGTP and 0 μM d^{Me}isoCTP. Very large ddA terminator signals were always observed opposite d^{Me}isoC in the absence of disoGTP. Little, if any, signal attenuation caused by unwanted strand termination was visible upon encountering isolated d^{Me}isoC positions, either in the presence or absence of complement disoGTP.

Figure S7. Dye terminator sequencing of three ODNs used to examine disoG in all possible natural nearest neighbor contexts (see Figure 4 for an additional sequence). ODNs containing four disoG positions (black arrows) were sequenced in a BigDye 3.1 reaction in the presence of (**A**), (**C**), (**E**) 100 μM d^{Me}isoCTP and 0 μM disoGTP or (**B**), (**D**), (**F**) 0 μM d^{Me}isoCTP and 0 μM disoGTP. Terminator ddT was always incorporated opposite disoG in the absence of d^{Me}isoCTP. Additionally, a noticeable signal that may correspond to the polymerase skipping over a fraction of template disoG positions was visible opposite disoG. Extremely large terminator signals were observed at the position following disoG for all conditions. Little, if any, signal attenuation caused by unwanted

strand termination was visible upon encountering isolated disoG positions, either in the presence or absence of complement d^{Me}isoCTP.

Figure S8. Dye terminator sequencing of two ODNs used to verify that significant self-pairing of disoG and d^{Me}isoC does not occur. (**A**) An ODN containing four disoG positions was sequenced using a BigDye 3.1 kit in the presence of 0 μM d^{Me}isoCTP and 100 μM disoGTP. The terminator pattern is identical to Figure 4B, the analogous reaction in the absence of d^{Me}isoCTP disoGTP. (**B**) An ODN containing four d^{Me}isoC positions was sequenced using a BigDye 3.1 kit in the presence of 100 μM d^{Me}isoCTP and 0 μM disoGTP. The terminator pattern is identical to Figure S6F, the analogous reaction in the absence of d^{Me}isoCTP disoGTP.

Fig. S1

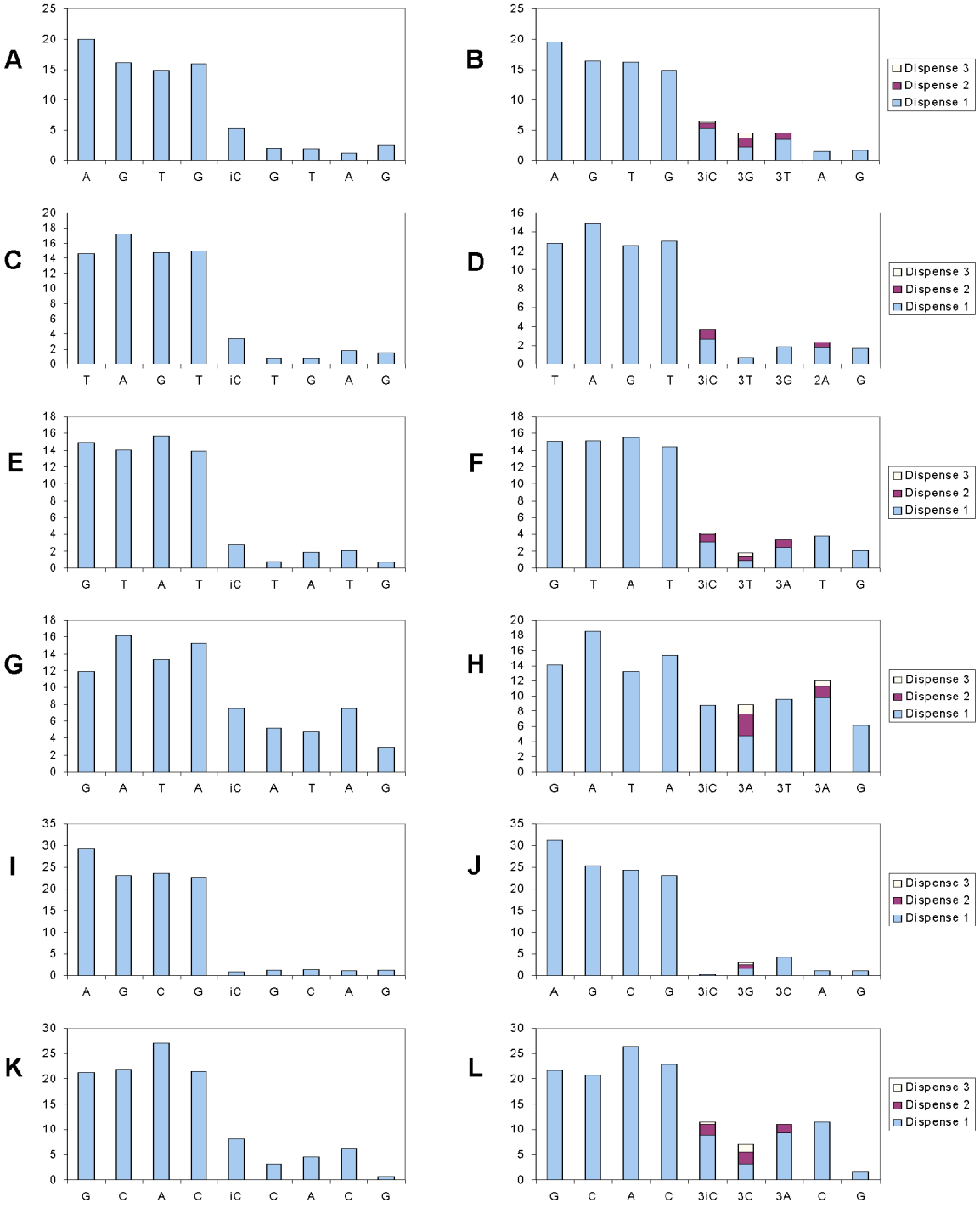

Fig. S2

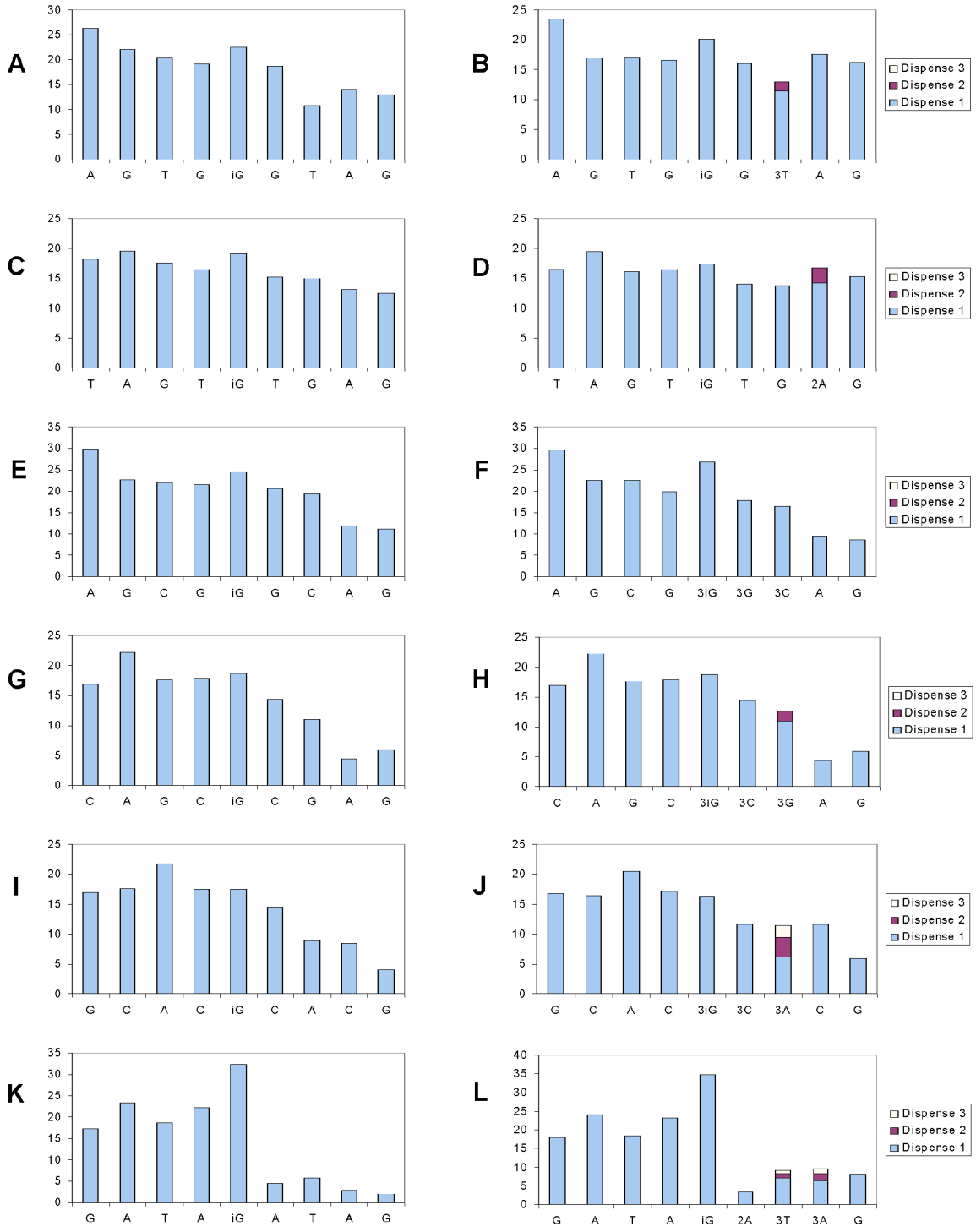
A
A G T G iG G T A G
B
A G T G iG G 3T A G
Dispense 3
Dispense 2
Dispense 1
C
T A G T iG T G A G
D
T A G T iG T G 2A G
E
A G C G iG G C A G
F
A G C G 3iG 3G 3C A G
G
C A G C iG C G A G
H
C A G C 3iG 3C 3G A G
I
G C A C iG C A C G
J
G C A C 3iG 3C 3A C G
K
G A T A iG A T A G
L
G A T A iG 2A 3T 3A G

Fig. S3

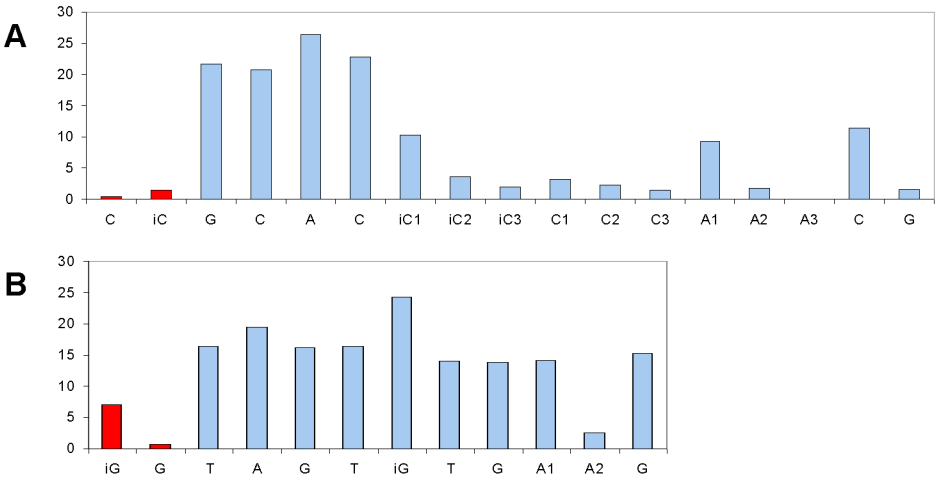
A
30
25
20
15
10
5
0
C
iC
G
C
A
C
iC1
iC2
iC3
C1
C2
C3
A1
A2
A3
C
G
B
30
25
20
15
10
5
0
iG
G
T
A
G
T
iG
T
G
A1
A2
G

Fig. S4

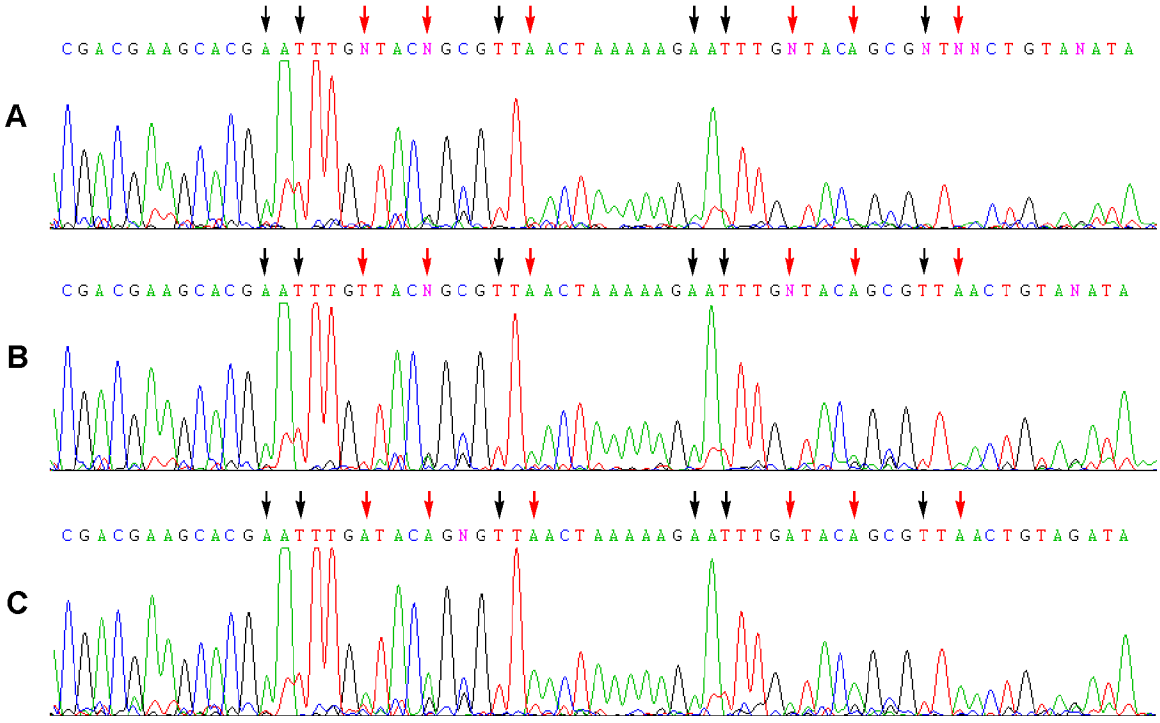
A
CGACGAAGCACGAATTTGNTACNGCGTTAACTAAAAAAGAATTTGNTACAGCGNTNNCTGTANATA
B
CGACGAAGCACGAATTTGTTACNGCGTTAACTAAAAAAGAATTTGNTACAGCGTTAACTGTANATA
C
CGACGAAGCACGAATTTGATACAGNGTTAACTAAAAAAGAATTTGATACAGCGTTAACTGTAGATA

Fig. S5

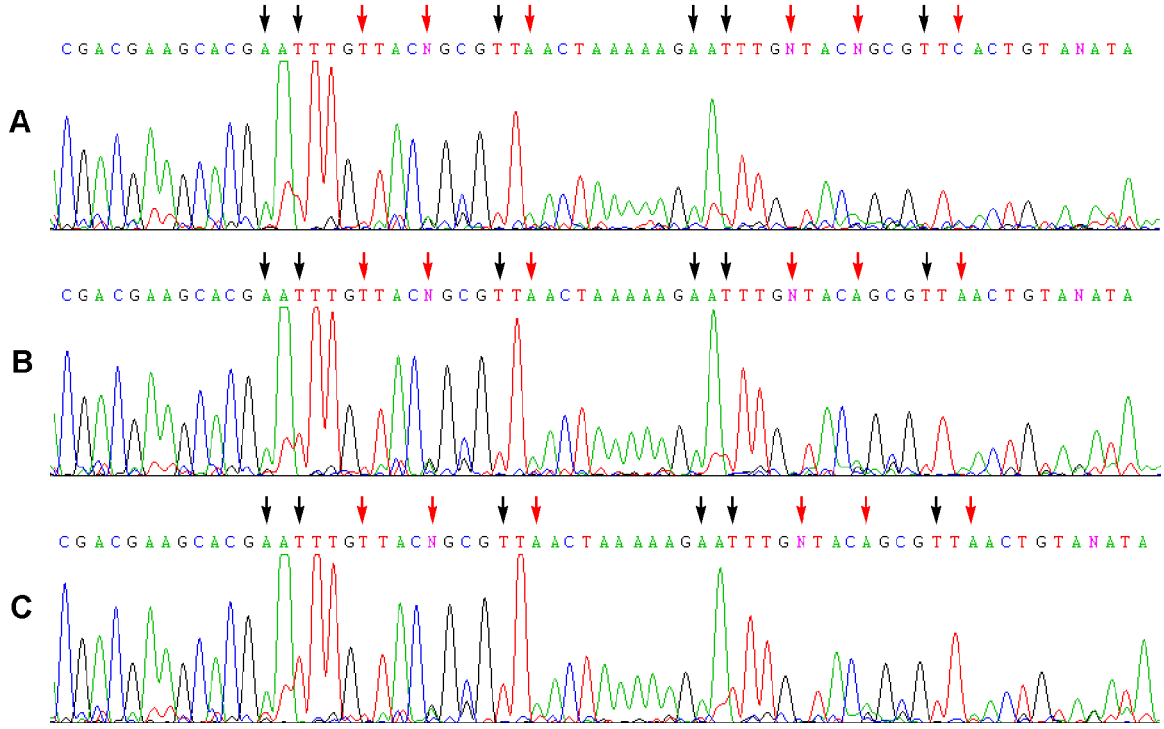

Fig. S6

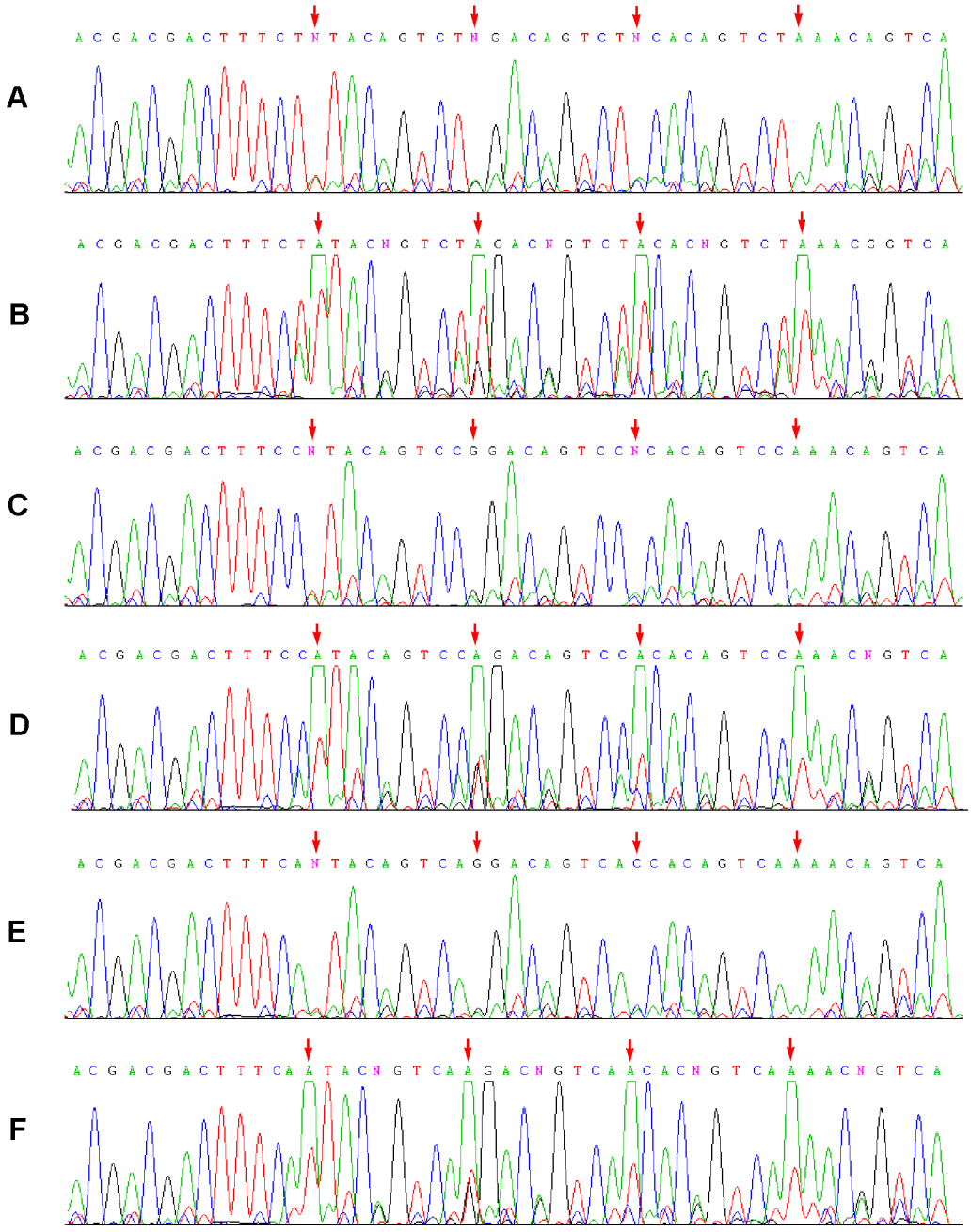

Fig. S7

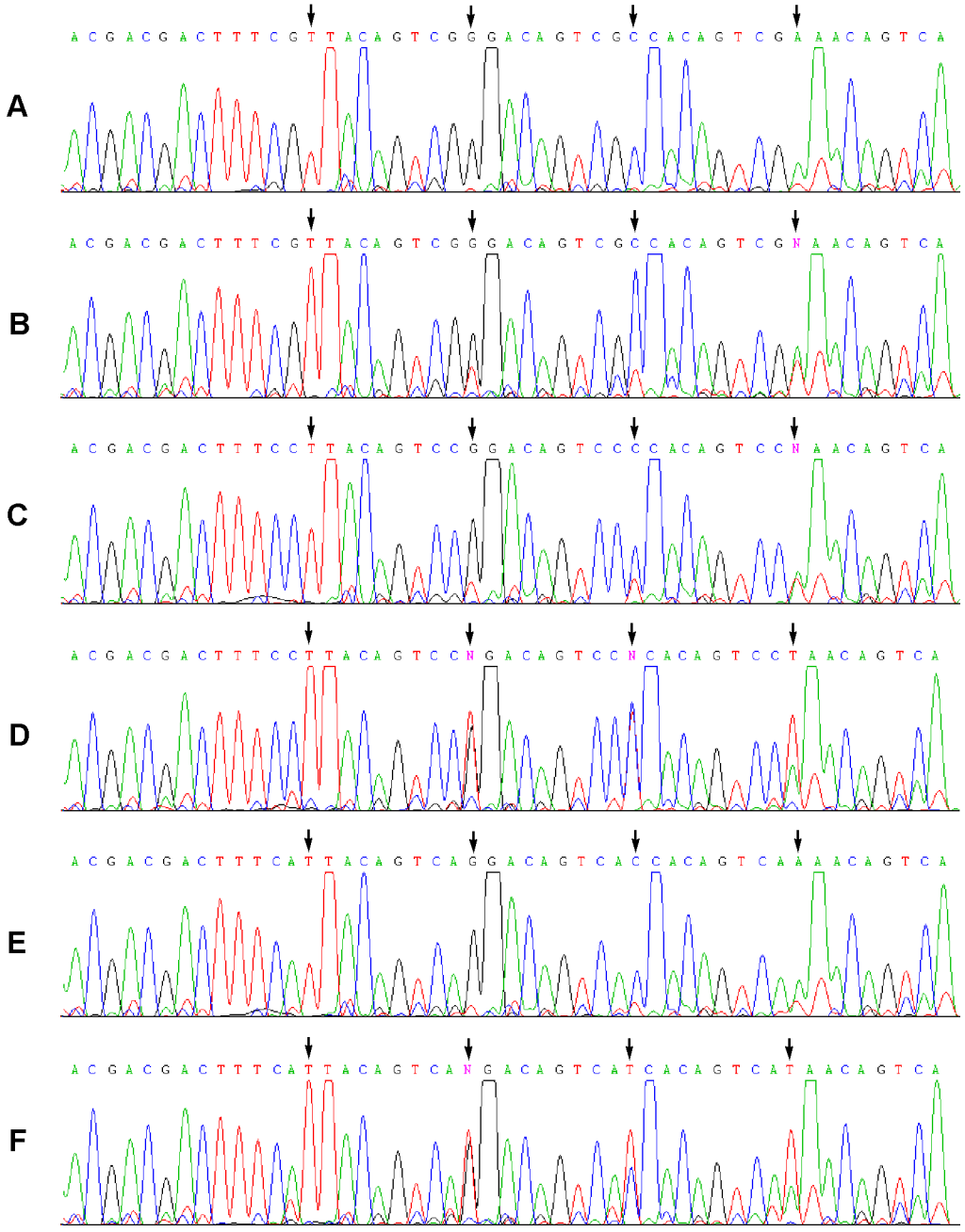
A
A C G A C G A C T T T C G T T A C A G T C G G G A C A G T C G C C A C A G T C G A A A C A G T C A
B
A C G A C G A C T T T C G T T A C A G T C G G G A C A G T C G C C A C A G T C G N A A C A G T C A
C
A C G A C G A C T T T C C T T A C A G T C C G G A C A G T C C C C A C A G T C C N A A C A G T C A
D
A C G A C G A C T T T C C T T A C A G T C C N G A C A G T C C N C A C A G T C C T A A C A G T C A
E
A C G A C G A C T T T C A T T A C A G T C A G G A C A G T C A C C A C A G T C A A A A C A G T C A
F
A C G A C G A C T T T C A T T A C A G T C A N G A C A G T C A T C A C A G T C A T A A C A G T C A

Fig. S8