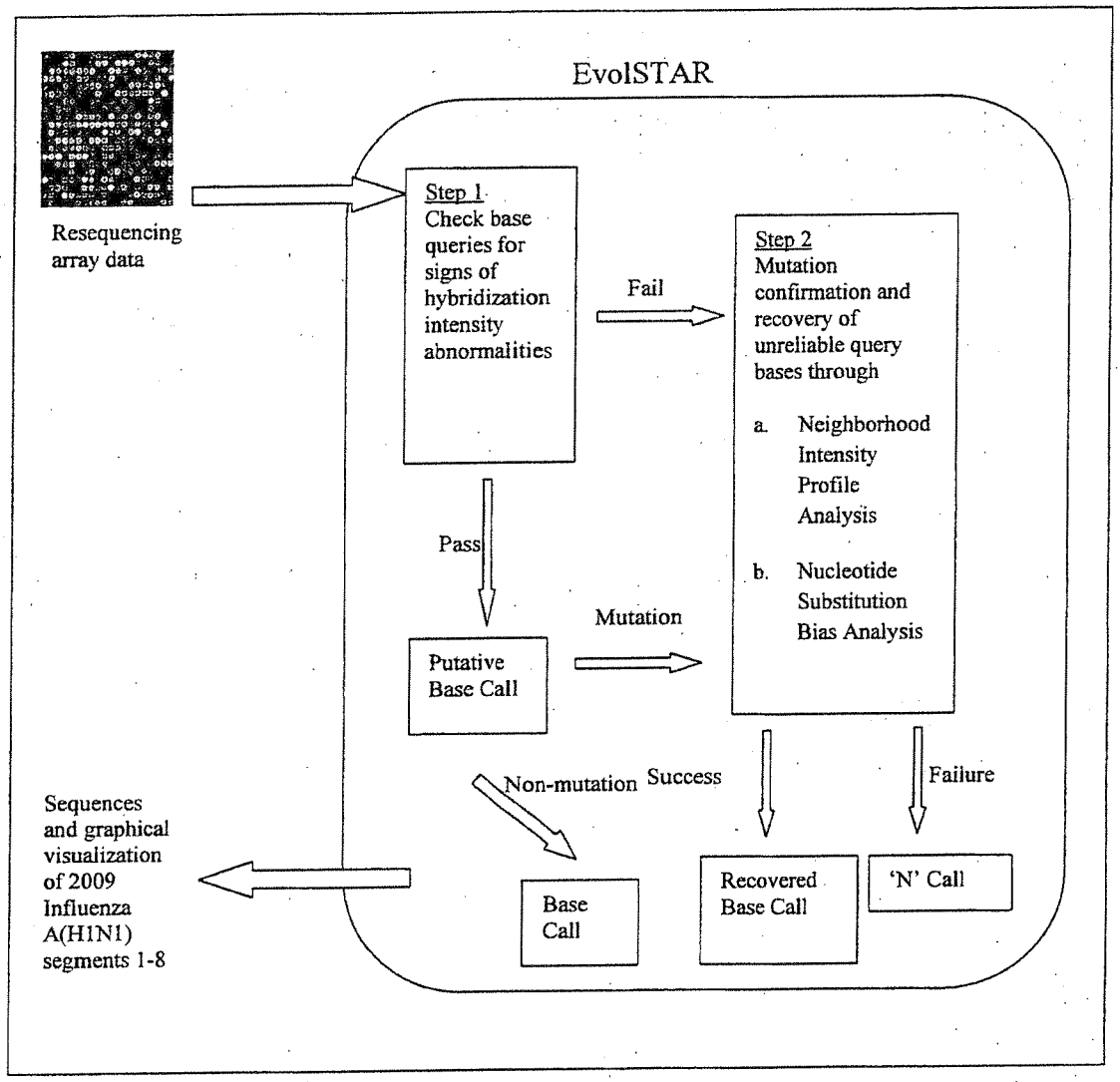
EvolSTAR
Resequencing array data
Step 1
Check base queries for signs of hybridization intensity abnormalities
Fail
Step 2
Mutation confirmation and recovery of unreliable query bases through
a. Neighborhood Intensity Profile Analysis
b. Nucleotide Substitution Bias Analysis
Pass
Mutation
Putative Base Call
Non-mutation
Success
Failure
Sequences and graphical visualization of 2009 Influenza A(H1N1) segments 1-8
Base Call
Recovered Base Call
'N' Call

FIGURE 1

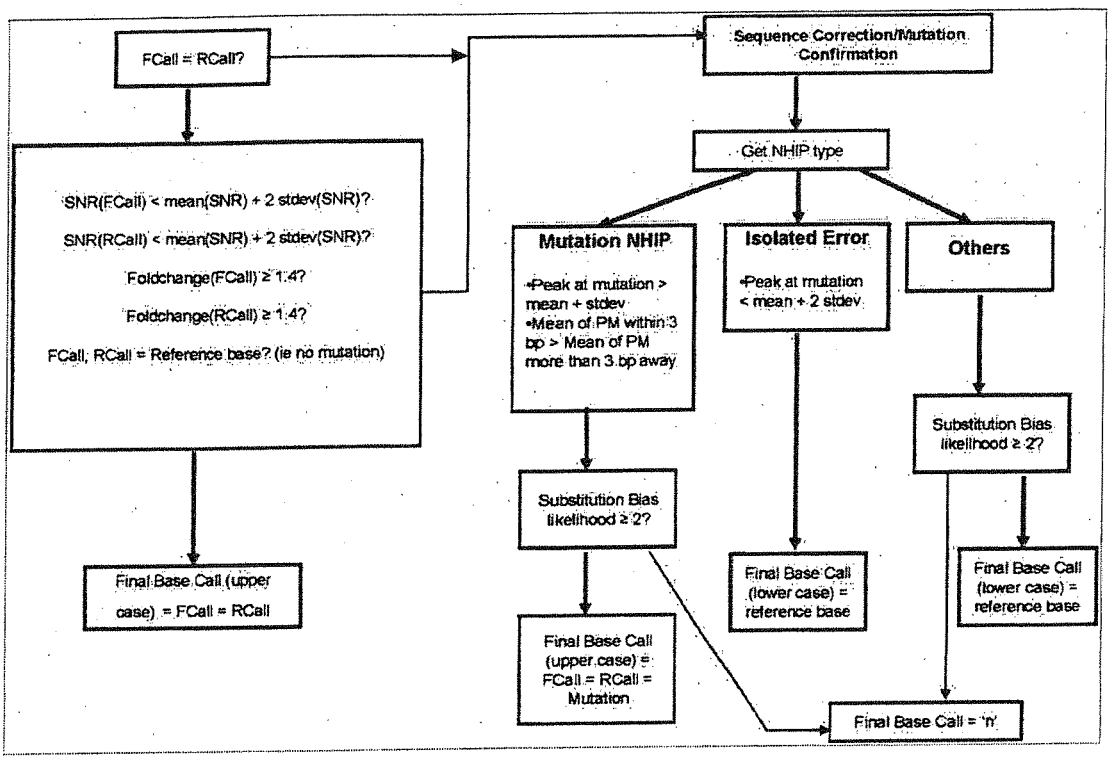
FCall = RCall?
SNR(FCall) < mean(SNR) + 2 stdev(SNR)?
SNR(RCall) < mean(SNR) + 2 stdev(SNR)?
Foldchange(FCall) ≥ 1.4?
Foldchange(RCall) ≥ 1.4?
FCall, RCall = Reference base? (ie no mutation)
Final Base Call (upper case) = FCall = RCall
Sequence Correction/Mutation Confirmation
Get NHIP type
Mutation NHIP
•Peak at mutation > mean + stdev
•Mean of PM within 3 bp > Mean of PM more than 3 bp away
Isolated Error
•Peak at mutation < mean + 2 stdev
Others
Substitution Bias likelihood ≥ 2?
Substitution Bias likelihood ≥ 2?
Final Base Call (lower case) = reference base
Final Base Call (upper case) = FCall = RCall = Mutation
Final Base Call (lower case) = reference base
Final Base Call = 'n'

FIGURE 2

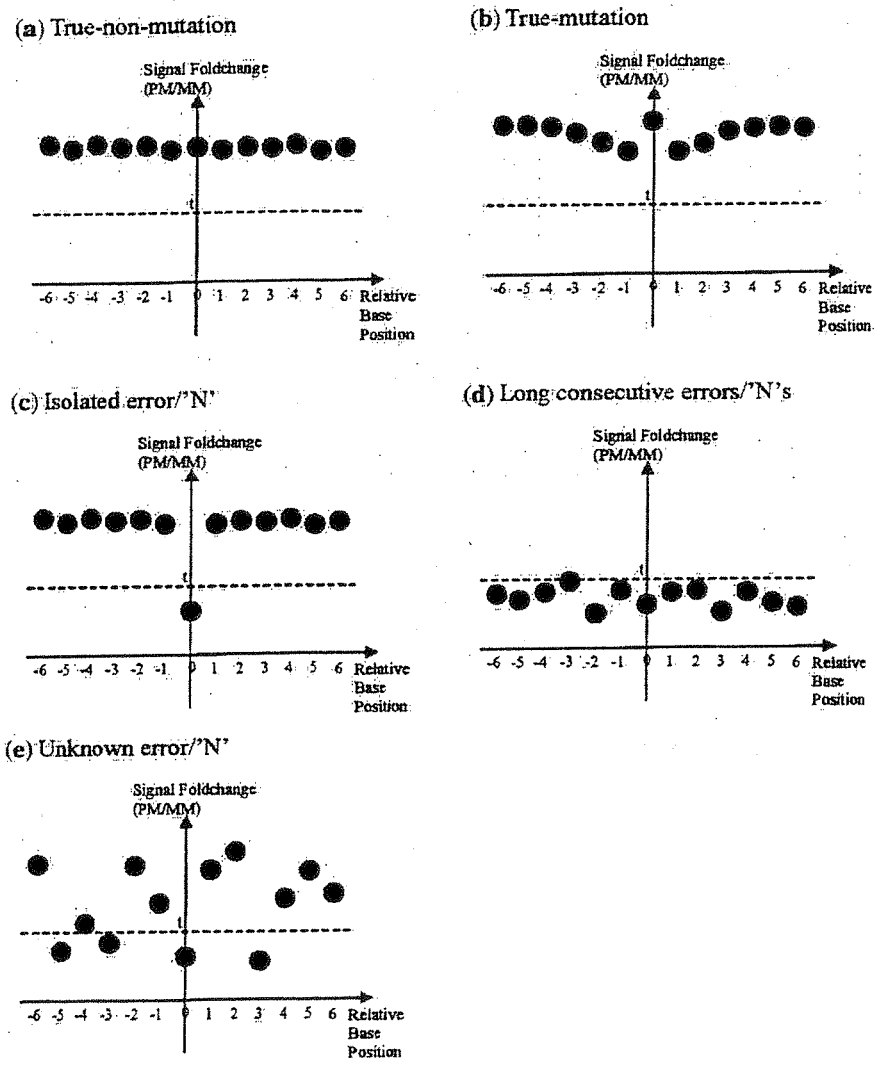
(a) True-non-mutation
Signal Foldchange (PM/MM)
t
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
Relative Base Position
(b) True-mutation
Signal Foldchange (PM/MM)
t
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
Relative Base Position
(c) Isolated error/'N'
Signal Foldchange (PM/MM)
t
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
Relative Base Position
(d) Long consecutive errors/'N's
Signal Foldchange (PM/MM)
t
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
Relative Base Position
(e) Unknown error/'N'
Signal Foldchange (PM/MM)
t
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
Relative Base Position

FIGURE 3

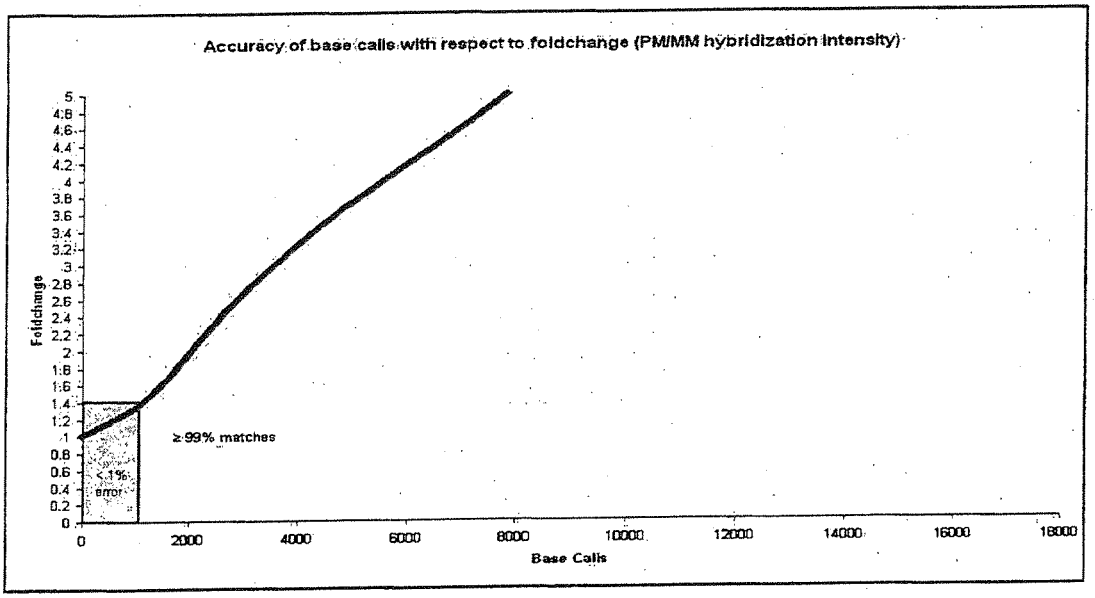

FIGURE 4

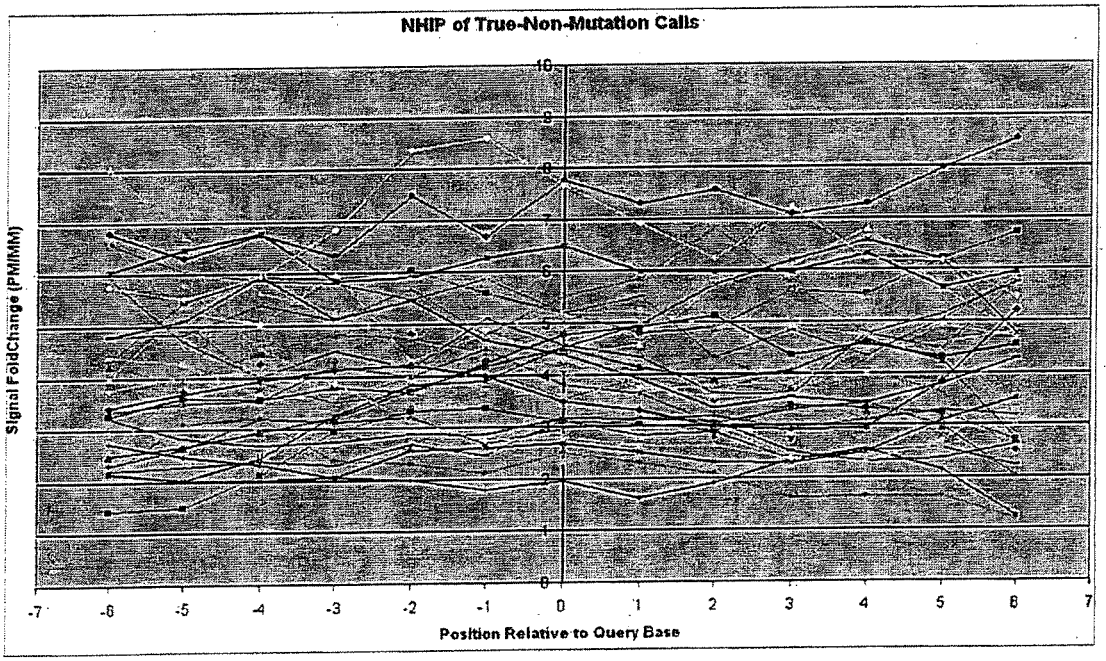

FIGURE 5

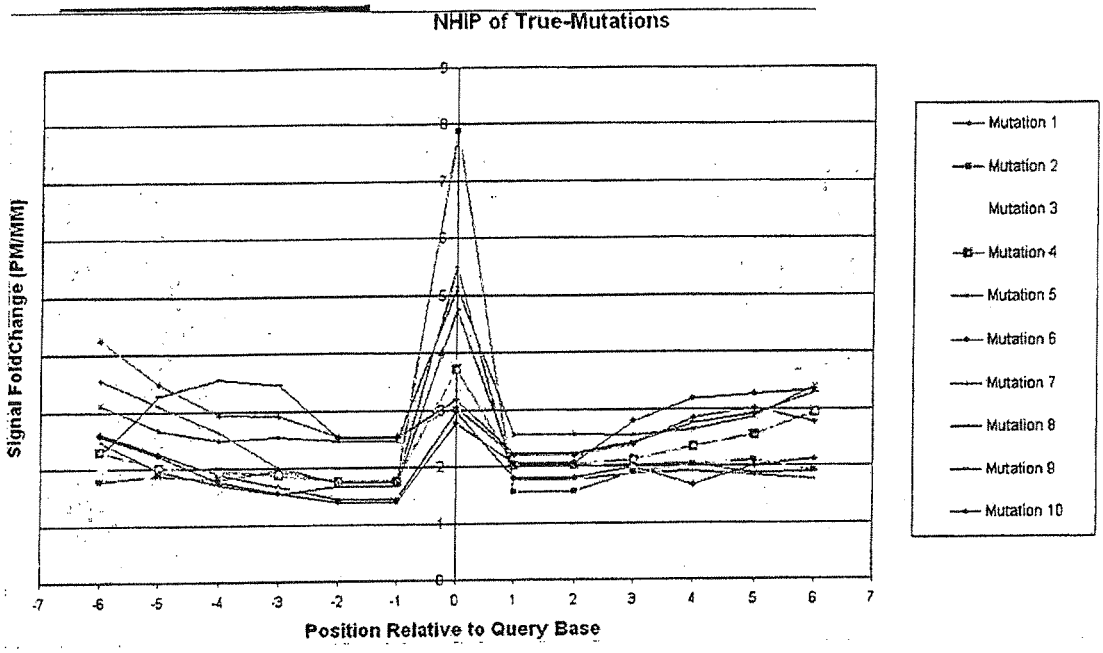

FIGURE 6

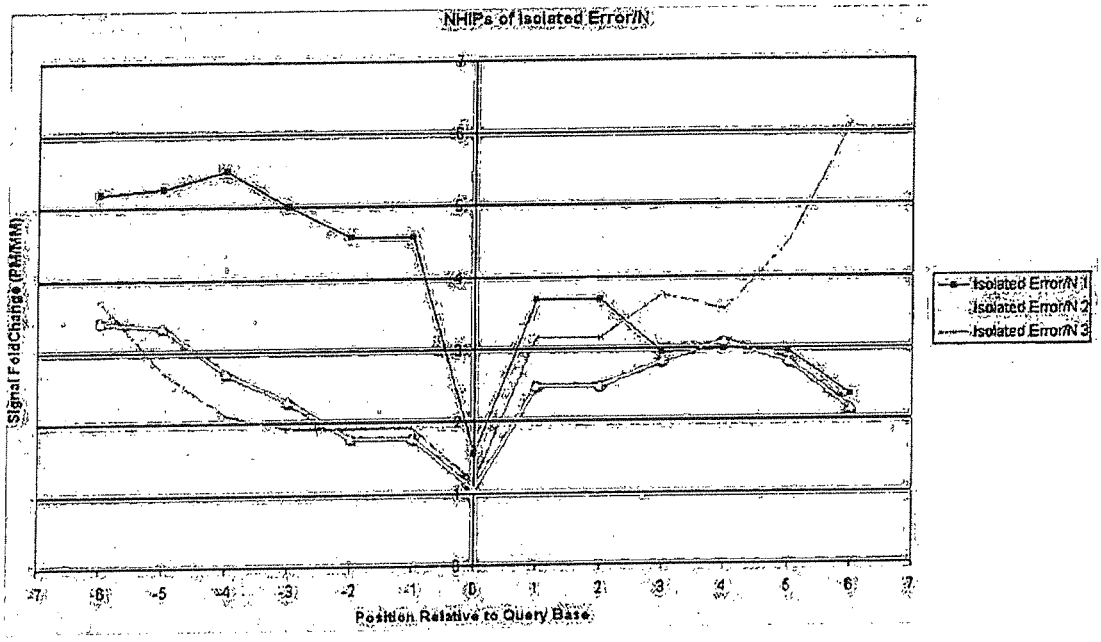

FIGURE 7

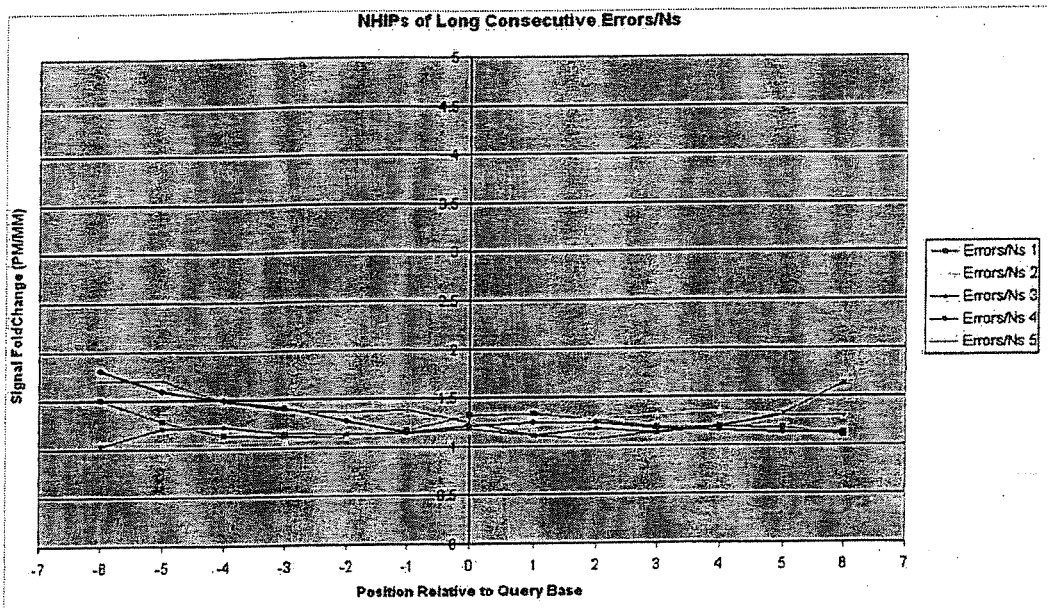

FIGURE 8

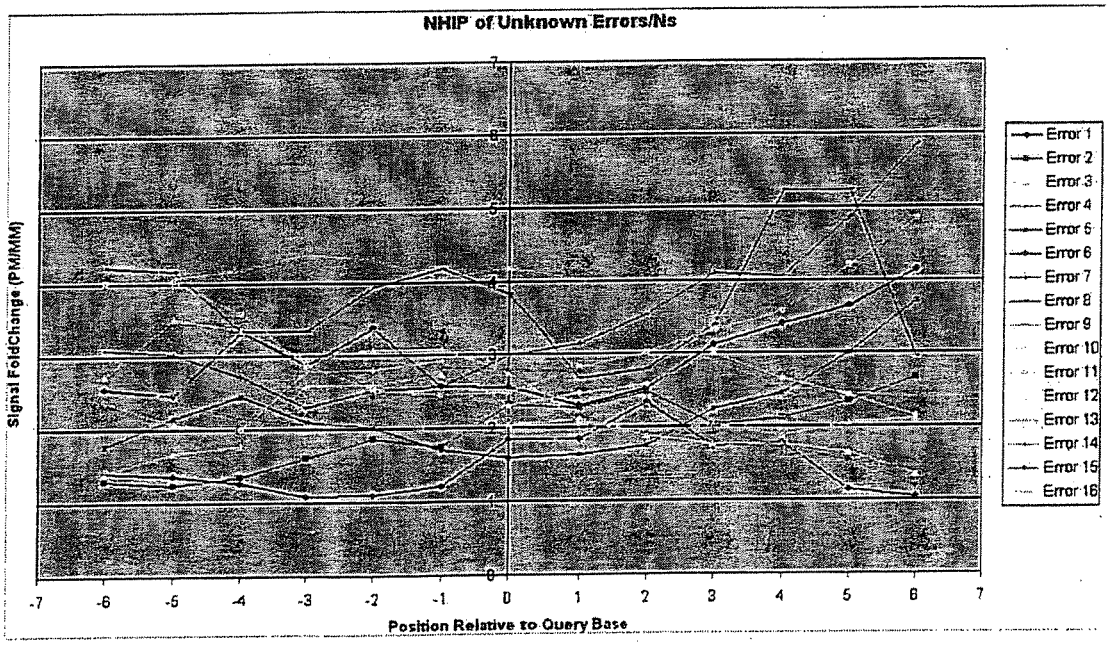

FIGURE 9

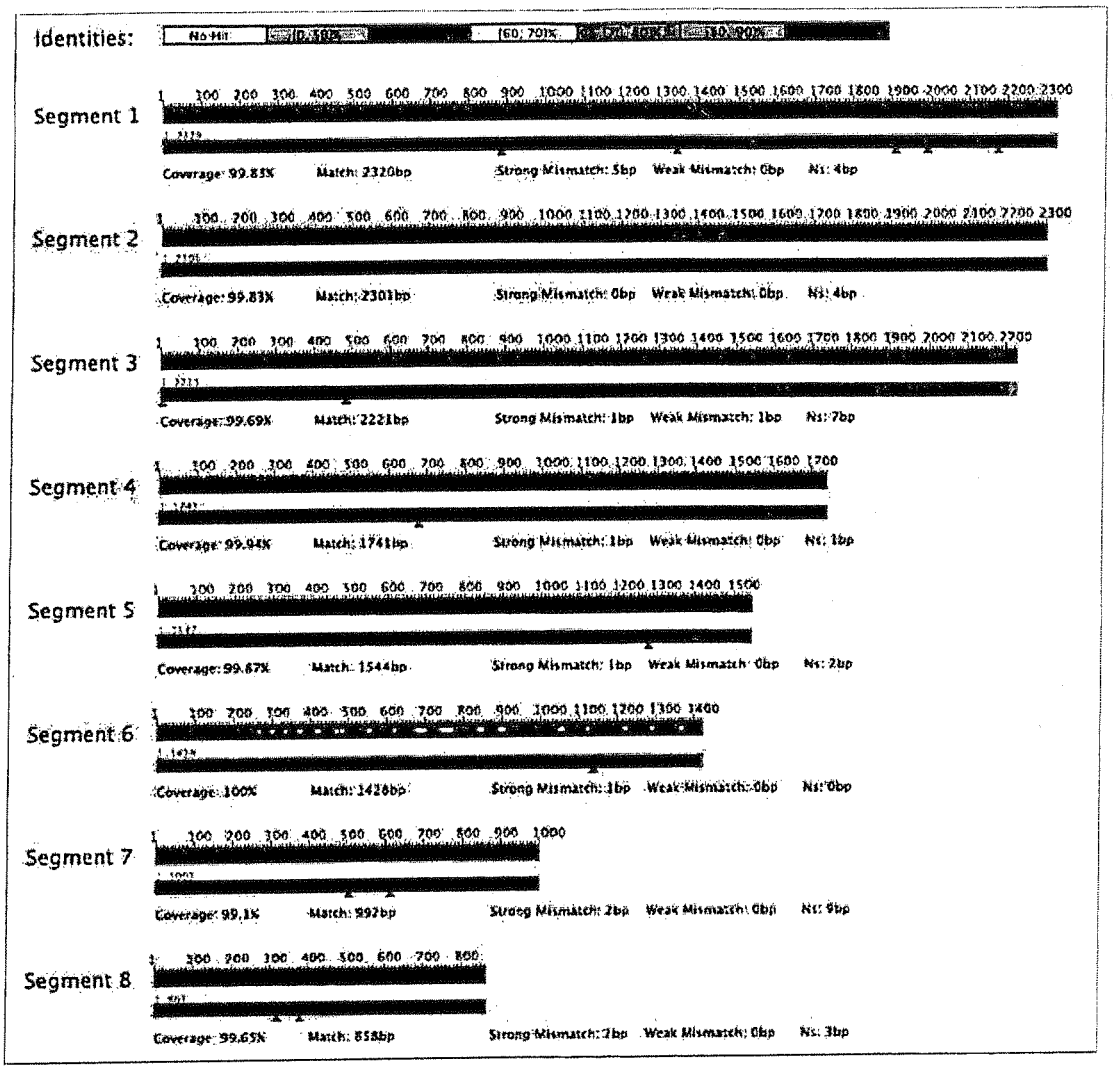
Identities:
No Hit
[60; 70]%
Segment 1
Coverage: 99.83%
Match: 2320bp
Strong Mismatch: 5bp
Weak Mismatch: 0bp
Ns: 4bp
Segment 2
Coverage: 99.83%
Match: 2301bp
Strong Mismatch: 0bp
Weak Mismatch: 0bp
Ns: 4bp
Segment 3
Coverage: 99.69%
Match: 2221bp
Strong Mismatch: 1bp
Weak Mismatch: 1bp
Ns: 7bp
Segment 4
Coverage: 99.94%
Match: 1741bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 1bp
Segment 5
Coverage: 99.87%
Match: 1544bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 2bp
Segment 6
Coverage: 100%
Match: 1428bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 0bp
Segment 7
Coverage: 99.1%
Match: 992bp
Strong Mismatch: 2bp
Weak Mismatch: 0bp
Ns: 9bp
Segment 8
Coverage: 99.65%
Match: 858bp
Strong Mismatch: 2bp
Weak Mismatch: 0bp
Ns: 3bp

FIGURE 10

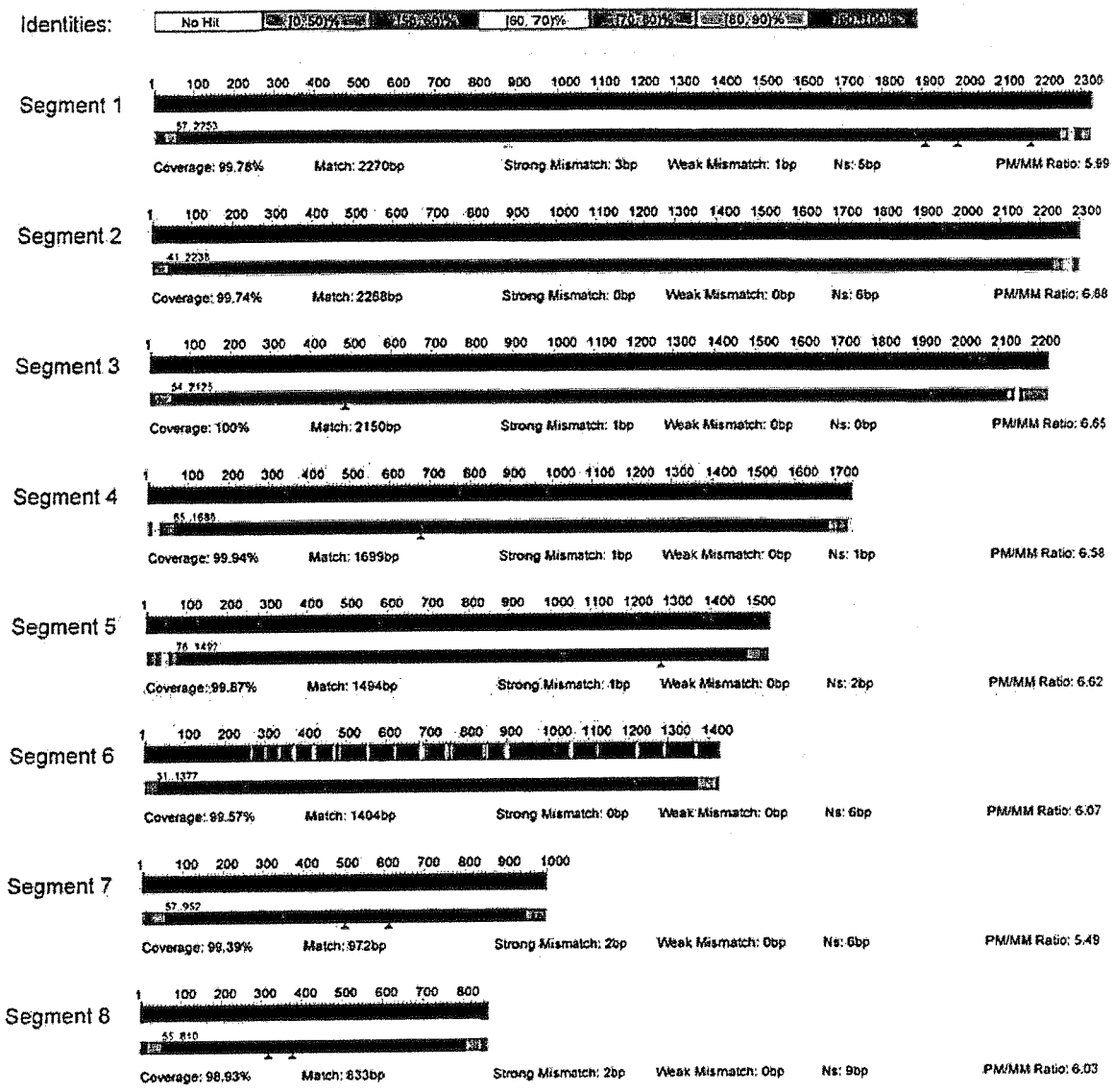
Identities:
No Hit
[60, 70)%
Segment 1
Coverage: 99.76%
Match: 2270bp
Strong Mismatch: 3bp
Weak Mismatch: 1bp
Ns: 5bp
PM/MM Ratio: 5.99
Segment 2
Coverage: 99.74%
Match: 2268bp
Strong Mismatch: 0bp
Weak Mismatch: 0bp
Ns: 6bp
PM/MM Ratio: 6.68
Segment 3
Coverage: 100%
Match: 2150bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 0bp
PM/MM Ratio: 6.65
Segment 4
Coverage: 99.94%
Match: 1699bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 1bp
PM/MM Ratio: 6.58
Segment 5
Coverage: 99.87%
Match: 1494bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 2bp
PM/MM Ratio: 6.62
Segment 6
Coverage: 99.57%
Match: 1404bp
Strong Mismatch: 0bp
Weak Mismatch: 0bp
Ns: 6bp
PM/MM Ratio: 6.07
Segment 7
Coverage: 99.39%
Match: 972bp
Strong Mismatch: 2bp
Weak Mismatch: 0bp
Ns: 6bp
PM/MM Ratio: 5.49
Segment 8
Coverage: 98.93%
Match: 833bp
Strong Mismatch: 2bp
Weak Mismatch: 0bp
Ns: 9bp
PM/MM Ratio: 6.03

FIGURE 11

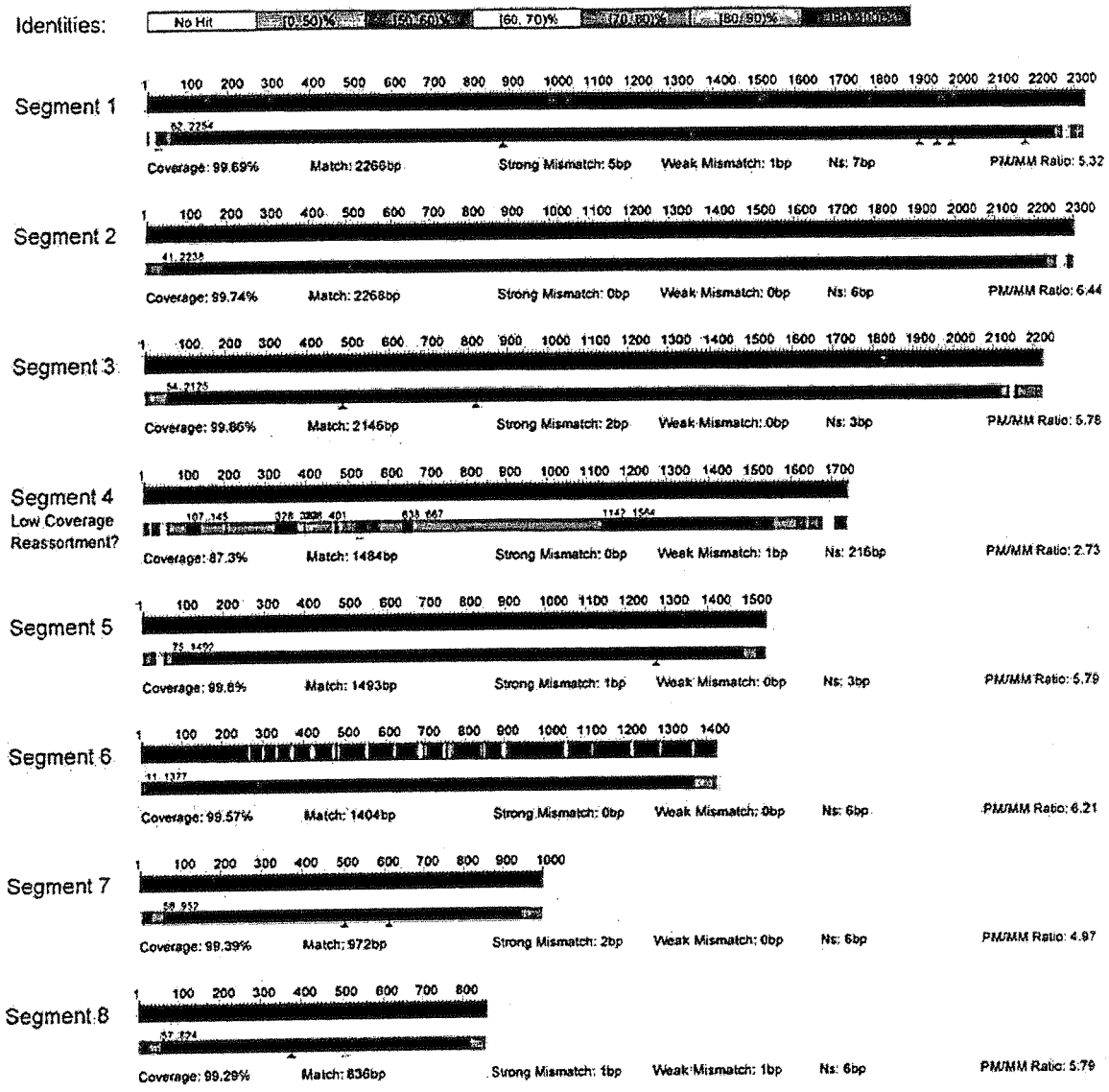
Identities:
No Hit
[0, 50)%
[50, 60)%
[60, 70)%
[70, 80)%
[80, 90)%
[90, 100]%
Segment 1
Coverage: 99.69%
Match: 2266bp
Strong Mismatch: 5bp
Weak Mismatch: 1bp
Ns: 7bp
PM/MM Ratio: 5.32
Segment 2
Coverage: 99.74%
Match: 2268bp
Strong Mismatch: 0bp
Weak Mismatch: 0bp
Ns: 6bp
PM/MM Ratio: 6.44
Segment 3
Coverage: 99.86%
Match: 2146bp
Strong Mismatch: 2bp
Weak Mismatch: 0bp
Ns: 3bp
PM/MM Ratio: 5.78
Segment 4
Low Coverage
Reassortment?
Coverage: 87.3%
Match: 1484bp
Strong Mismatch: 0bp
Weak Mismatch: 1bp
Ns: 216bp
PM/MM Ratio: 2.73
Segment 5
Coverage: 99.8%
Match: 1493bp
Strong Mismatch: 1bp
Weak Mismatch: 0bp
Ns: 3bp
PM/MM Ratio: 5.79
Segment 6
Coverage: 98.57%
Match: 1404bp
Strong Mismatch: 0bp
Weak Mismatch: 0bp
Ns: 6bp
PM/MM Ratio: 6.21
Segment 7
Coverage: 99.39%
Match: 972bp
Strong Mismatch: 2bp
Weak Mismatch: 0bp
Ns: 6bp
PM/MM Ratio: 4.97
Segment 8
Coverage: 99.29%
Match: 836bp
Strong Mismatch: 1bp
Weak Mismatch: 1bp
Ns: 6bp
PM/MM Ratio: 5.79

FIGURE 12