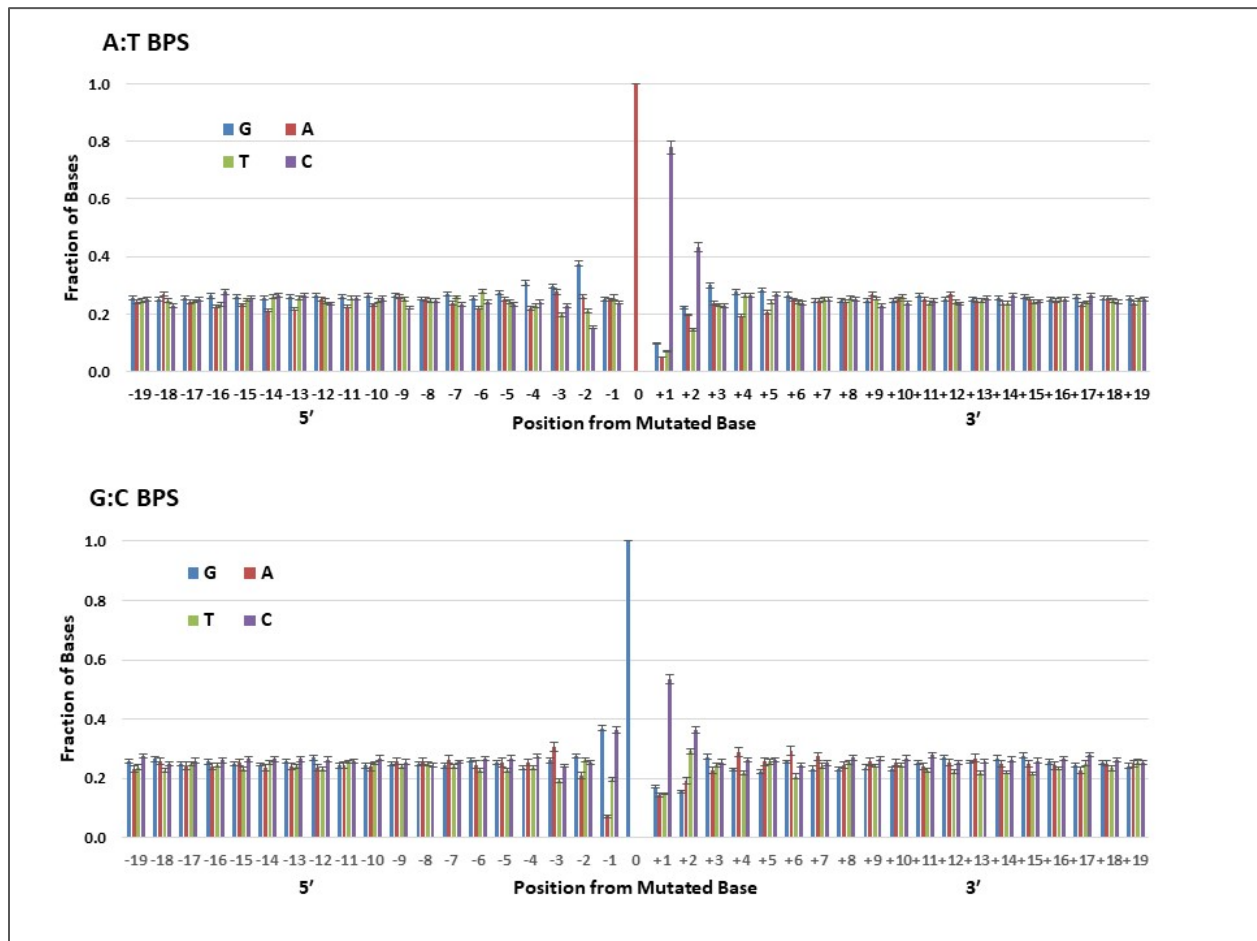**Figure S1. BPS frequencies are influenced by the local sequence context.** The sequences of the 19 bases on each side of each mutated base or their reverse complements were oriented so to flank the purine member of the mutated base pair. The rate at which each base appeared in each of the 38 flanking positions was calculated from all the experiments with MMR defective strains (using the same method as used to calculate the mutation rate; see Materials and Methods). The resulting numbers were then normalized to the number of that base in the genome, and then divided by the overall sum of all 4 bases. The resulting fraction of each base in each position is shown with the 95% CL.
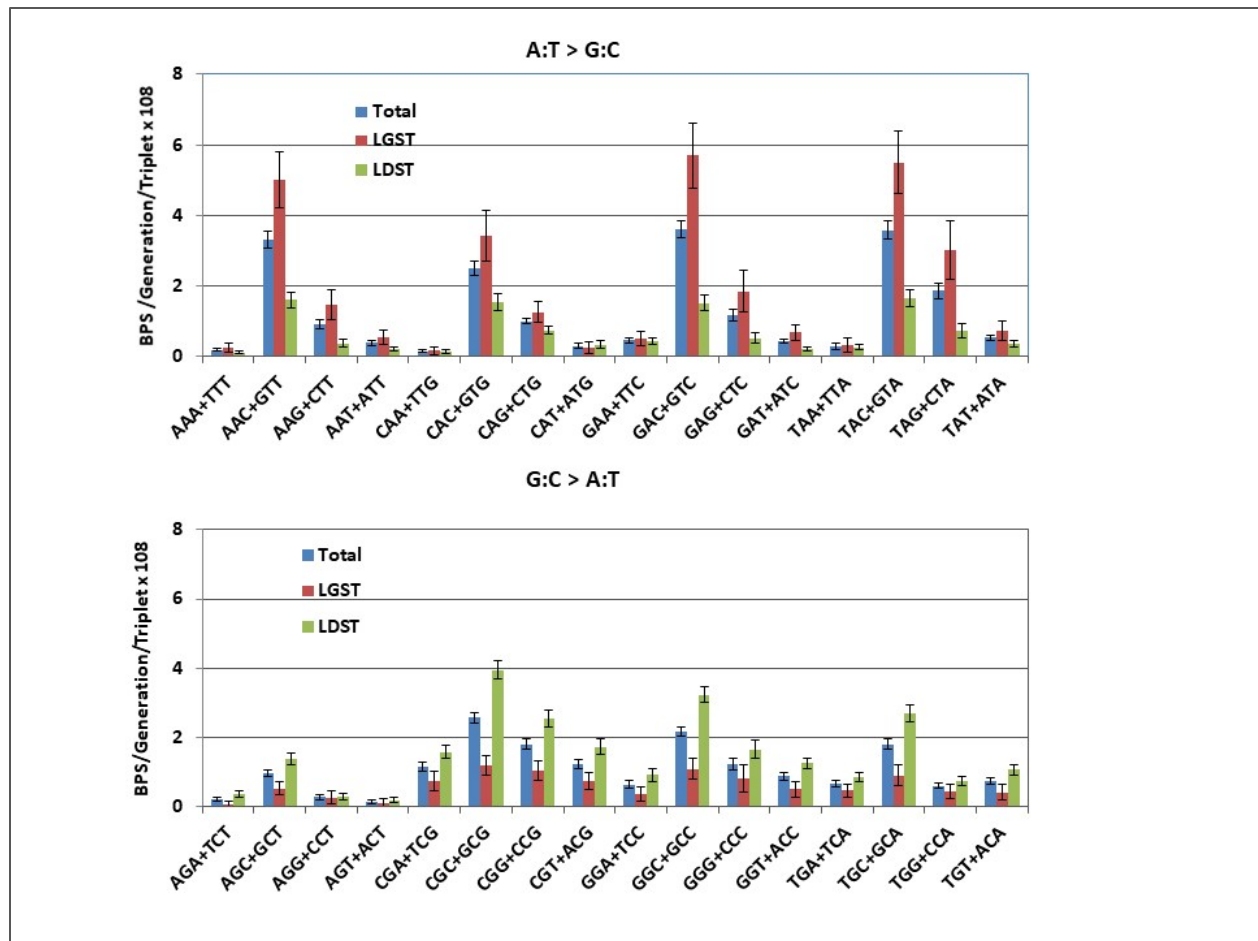
Figure S2. 5'NAC3'+5'GTN3' triplets are hotspots for transitions when MMR-defective strains are grown on minimal medium. Two MA experiments, one with the *mutS* mutant strain PFM343 and one with the *mutL* mutant strain, PFM5, each grown on minimal glucose medium, have been combined to give 3043 BPSs (see Supplemental Tables 3&4). Bars represent the mean mutation rate at each triplet, and error-bars are 95% CLs. The X-axis labels are the 34 non-redundant triplets displayed 5' to 3' with the target base in the middle. LGST, the target purine as displayed is on the lagging strand template; LDST, the target purine as displayed is on the leading strand template.
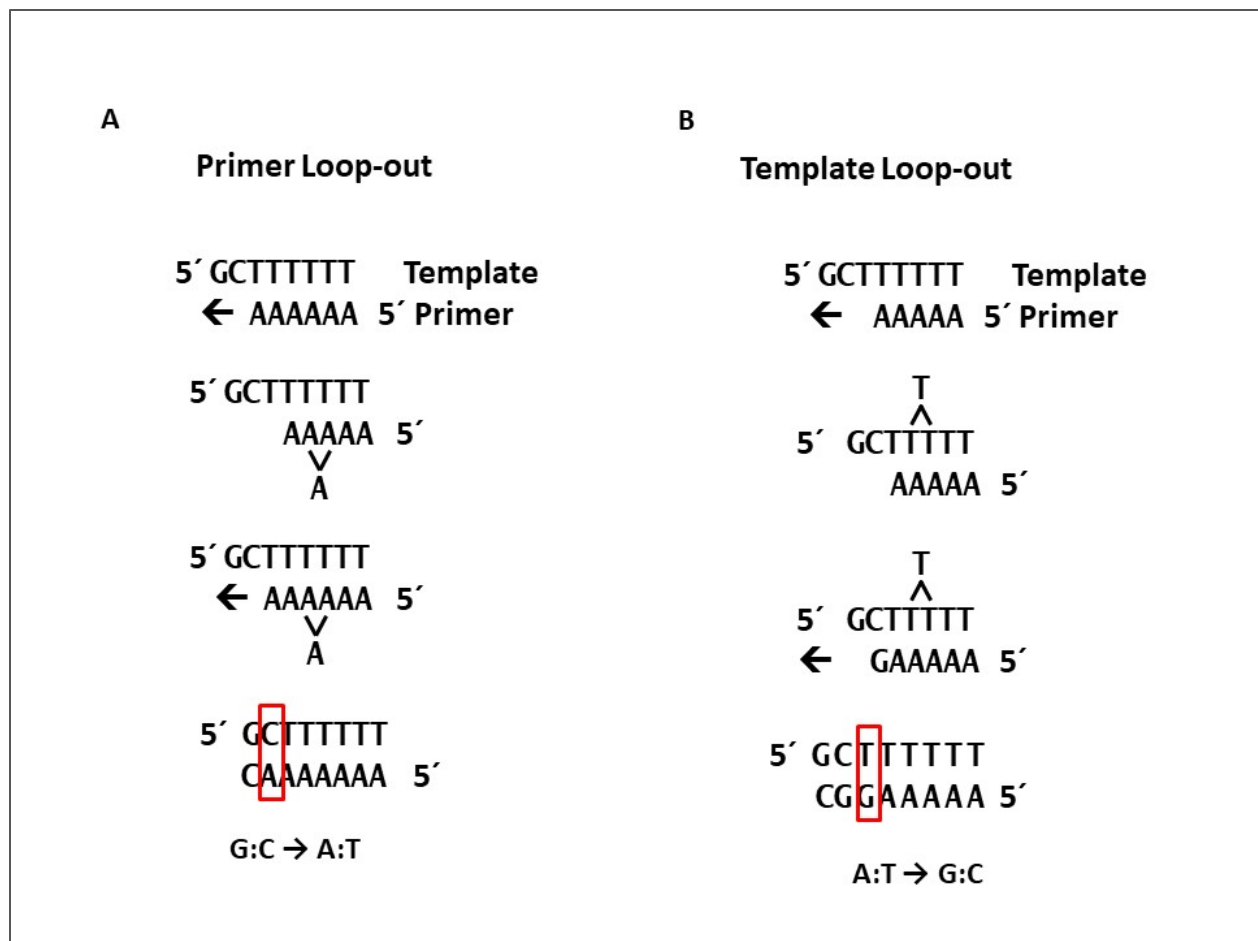
A

Primer Loop-out

5´ GCTTTTTT   Template
  ← AAAAAA 5´ Primer

5´ GCTTTTTT
    AAAAA 5´
      ∨
      A

5´ GCTTTTTT
  ← AAAAAA 5´
        ∨
        A

5´ GCTTTTTT
  CAAAAAAA 5´

G:C → A:T

B

Template Loop-out

5´ GCTTTTTT   Template
  ← AAAAA 5´ Primer

        T
        ∧
5´ GCTTTTT
      AAAAA 5´

        T
        ∧
5´ GCTTTTT
  ← GAAAAA 5´

5´ GCTTTTTT
  CGGAAAAA 5´

A:T → G:C

Figure S3. Base-pair substitution by transient misalignment. The red box indicates the mispair resulting from realignment of the slipped DNA strand. Note that in the case of primer loop-out (3A), the mispair is templated by the run, whereas in the case of template loop-out (3B), the mispair is templated by the base adjacent to the run. (Adapted from Fowler *et al.* 1974; Kunkel and Soni 1988).
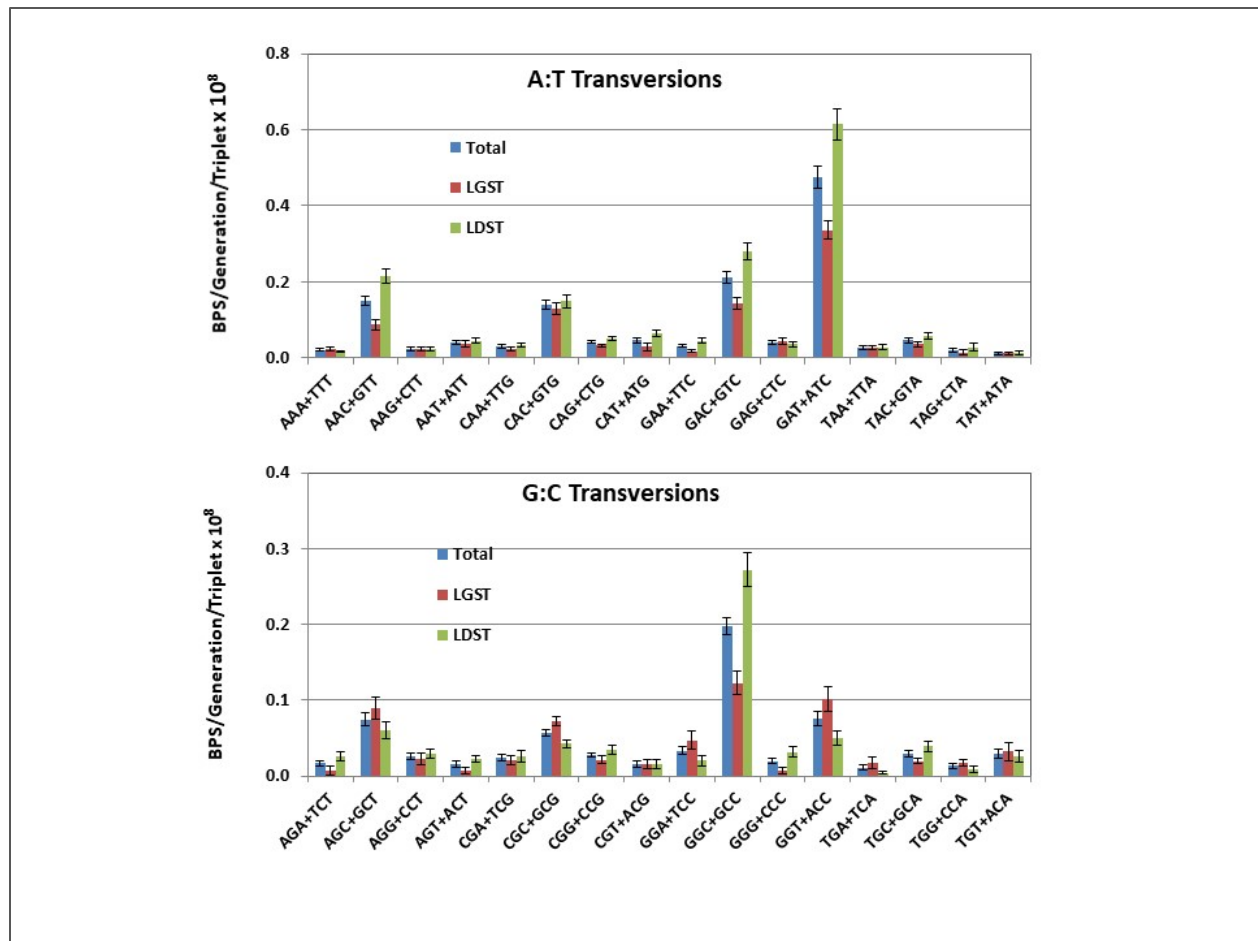
Figure S4. **Transversion mutations are only weakly DNA-strand biased**. Bars represent the mean rates of transversions accumulated in 10 experiments with MMR-defective strains (see Materials and Methods). Mutation rates per generation at each triplet were divided by the number of that triplet in the genome. Error bars are 95% CLs. The X-axis labels are the 32 sets of non-redundant triplets read 5' to 3' with the target base in the center of each triplet. . Note the change in scale between the transversions at A:Ts and at G:Cs. LGST, the target purine as displayed was on the lagging strand template; LDST, the target purine as displayed was on the leading strand template.
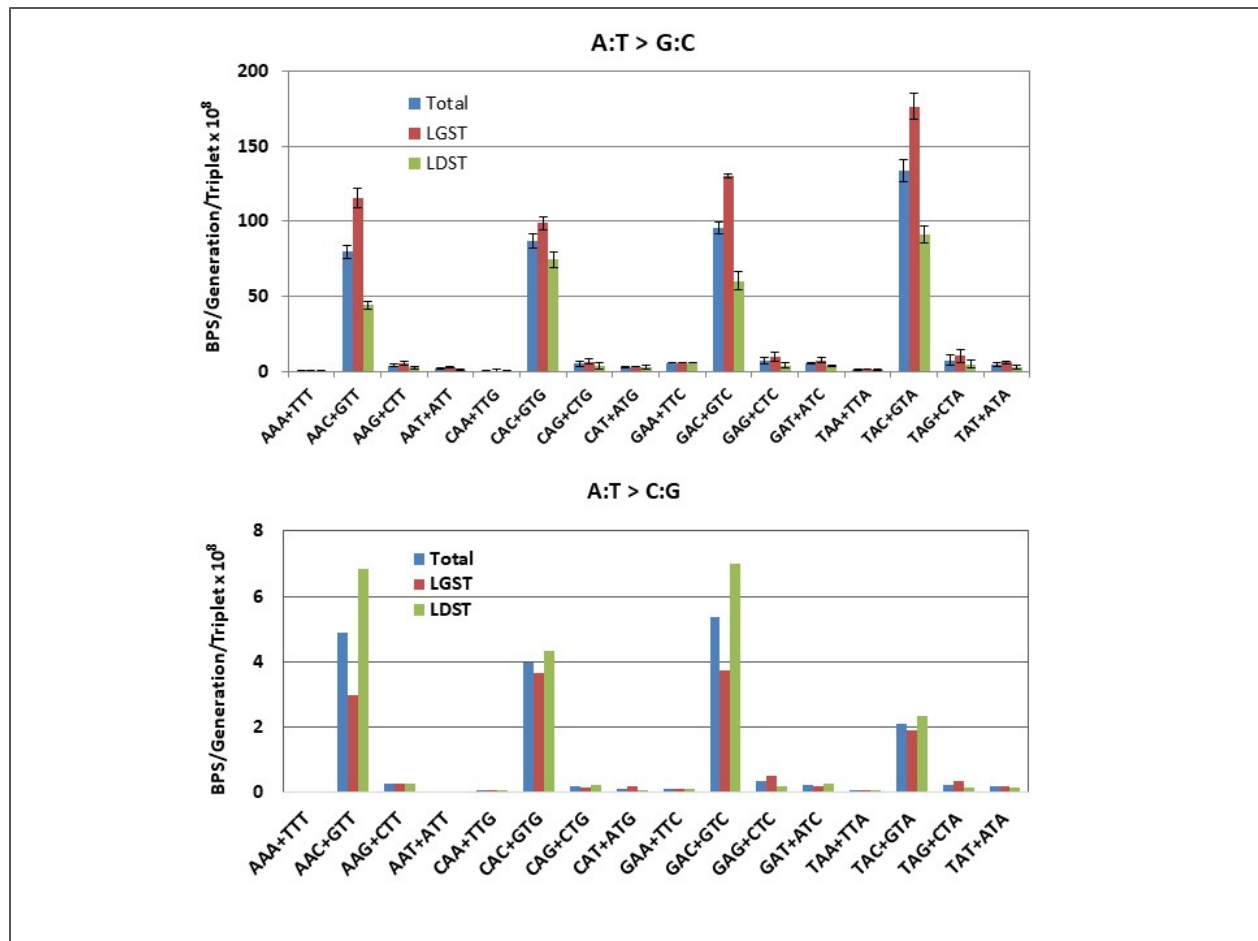
Figure S5. **5'NAC3'+5'GTN3' triplets are hotspots for insertion of C opposite A in the *mutL ndk*
mutant strain**. Bars represent the mean rates of BPS in the *mutL ndk* mutant strain PFM666.
Mutation rates per generation at each triplet were divided by the number of that triplet in the
genome. Error bars are 95% CLs; too few A:T > C:G mutations were accumulated to allow
meaningful error bars. The X-axis labels are the 32 sets of non-redundant triplets read 5' to 3'
with the target base in the center of each triplet. Note the change in scale between the A:T>G:C
and A:T>C:G mutations. LGST, the target purine as displayed was on the lagging strand
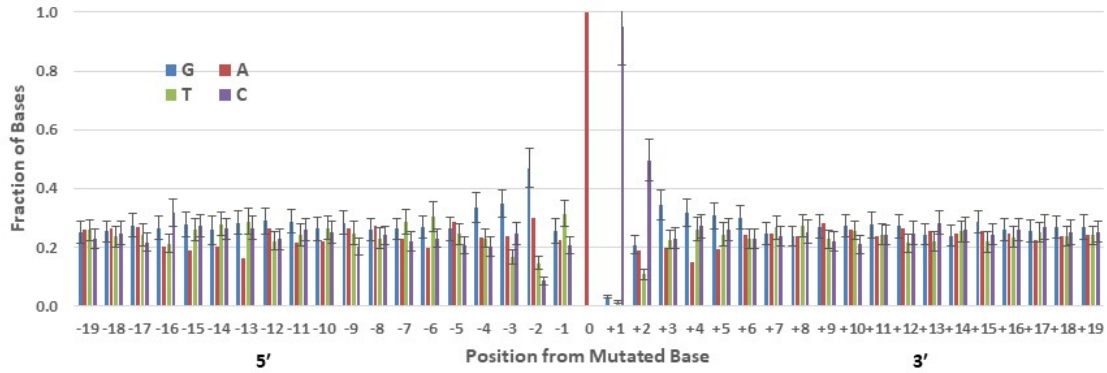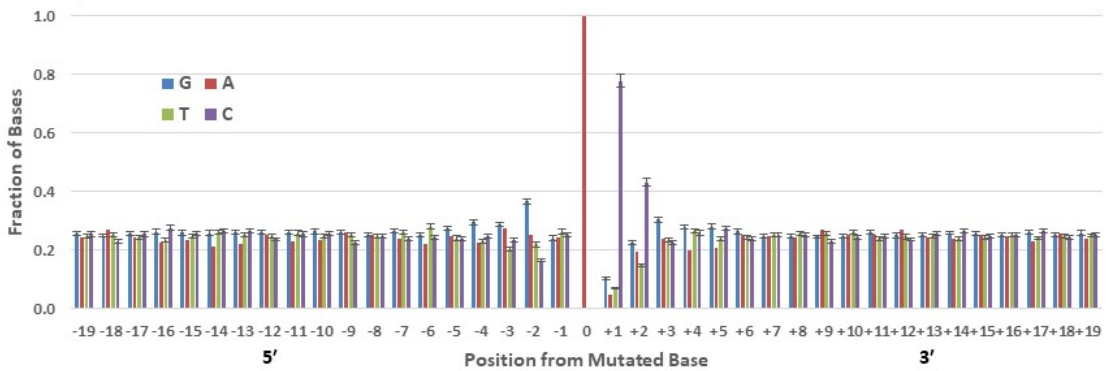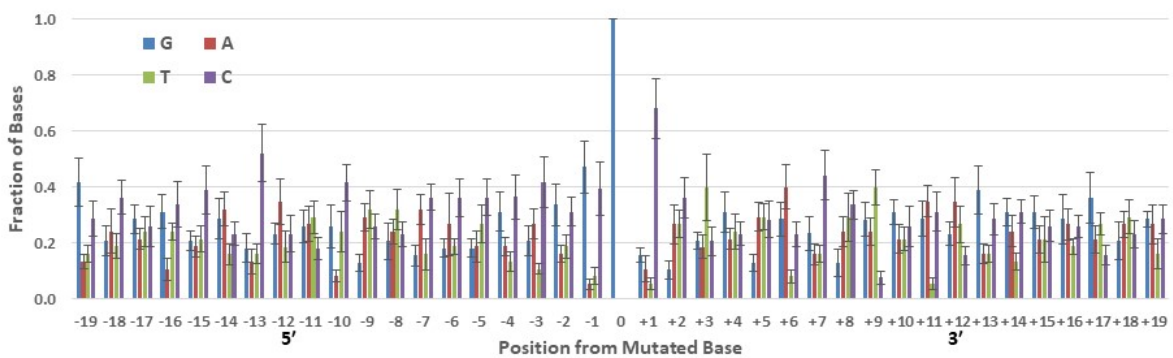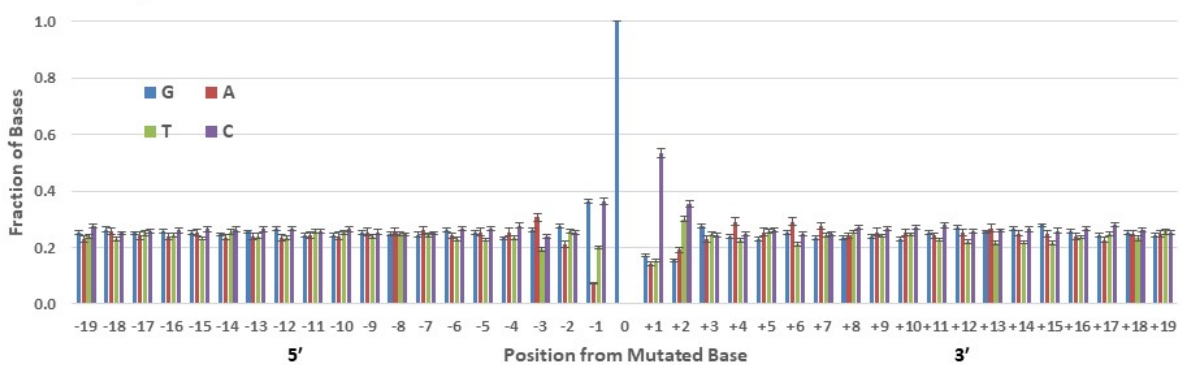template; LDST, the target purine as displayed was on the leading strand template

.

A

Repeated A:T BPS minus runs >3

Non-repeated A:T BPS minus runs >3

B

Repeated G:C BPS minus runs >3

Non-repeated G:C BPS minus runs >3

Figure S6. The sequence context of repeated BPSs does not differ from that of non-repeated BPSs. The sequences of the 19 bases 3' and 5´ to each mutated base or their reverse complements were oriented so to flank the purine member of the mutated base pair. The rate at which each base appeared in each position was used to determine the average rate and its variance over all the experiments (using the same method as used to calculate the mutation rate; see Materials and Methods). The resulting numbers were then normalized to the number of that base in the genome, and then divided by the overall sum of all 4 bases. The resulting fraction of each base in each position is shown with the 95% CLs. To compare the potentially hottest to less hot sequence contexts, instances of repeated mutations at the same base pair were separated from instances of single mutations. Because base pairs associated with mononucleotide runs are themselves hotspots (see Figure 3), these were also eliminated. The resulting sequences suggest that there is no significant difference in the surrounding sequences between repeated and non-repeated BPS. A, BPSs at A:T sites; B, BPSs at G:C sites