

Supplementary Data

Systematic evaluation of error rates and causes in short samples in next-generation sequencing

Franziska Pfeiffer, Carsten Gröber, Michael Blank, Kristian Händler, Marc Beyer, Joachim L. Schultze, and Günter Mayer*

For all tables, yellow indicates substitutions, green insertions, and turquoise deletions.

Table S1: Original sequence and 25 most frequent mutations of C12_T_PWO

Sequence	Sequence count
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	982317
G G TTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	4753
T TTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	4240
GTTGGAAGCGACGGGACGGTA G GGCTTGGGCCCAAGGAGTG	3792
GTTGGAAGCGACGGGACGGTAAGGCTT G CCCAAGGAGTG	3722
GTTGGAAGCGAC G ACGGTAAGGCTTGGGCCCAAGGAGTG	3515
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCC C AAGGAGTG	2732
T TTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2226
GTTGGAAGCGACGGGAC G TAAGGCTTGGGCCCAAGGAGTG	1970
GTTGGA A AGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1943
GTTGGAAGCGACGGGACGGT A GCTTGGGCCCAAGGAGTG	1886
GTT G AAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1727
GTT G AAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1711
GTTGGAAGCGACGGG G ACGGTAAGGCTTGGGCCCAAGGAGTG	1688
GTTGG C AAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1641
GTTGGAAGCGACGGGACGGTAAGG C TGGGCCCAAGGAGTG	1590
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCC C CAAGGAGTG	1412
GTTGGAAGCGACGG C CGTAAGGCTTGGGCCCAAGGAGTG	1319
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCC C CAAGGAGTG	1301
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCC C CAAGGAGTG	1291
GTTGGAAGCGACGGGACGGTAA G CTTGGGCCCAAGGAGTG	1239
GTTGGAAGCGACGGGACGGTAAGGCT G GGGCCCAAGGAGTG	1193
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCC C AAGGAGTG	1137
GTTGGAAGCGACGGGACGGT G AGGCTTGGGCCCAAGGAGTG	1123
GTTGGAAGCGACGGG A GGTAAGGCTTGGGCCCAAGGAGTG	1085
GTTGGAAGCGACGGGACGGTAAG G TGGGCCCAAGGAGTG	1059

Table S2: Original sequence and 25 most frequent mutations of C12_T_Taq

Sequence	Sequence count
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2990041
G G TTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	13977
GTTGGAAGCGACGGGACGGTA G GGCTTGGGCCCAAGGAGTG	12302
T TTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	11493
GTTGGAAGCGACGGGACGGTAAGGCTT G CCCAAGGAGTG	11343
GTTGGAAGCGAC G ACGGTAAGGCTTGGGCCCAAGGAGTG	10431

GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	7847
GTCGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	6353
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	6098
GTTGGAAGCGACGGGACGTAAGGCTTGGGCCCAAGGAGTG	5541
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	5406
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	5179
GTTGGAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	5091
GTTGGAAGCGACGGGACGGTAGGCTTGGGCCCAAGGAGTG	5059
GTTGGCAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	4909
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	4721
GTTGGAAGCGACGGGACGGTAAGGCTGGGCCCAAGGAGTG	4541
GTTGGAAGCGACGGGACGGTGAGGCTTGGGCCCAAGGAGTG	4053
GCTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3972
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3948
GTTGGAAGCGACGGGACGGTAAGGCTGGGCCCAAGGAGTG	3914
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3772
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3694
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3635
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3560
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3536

Table S3: Original sequence and 25 most frequent mutations of C12_T_w/o

Sequence	Sequence count
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1640067
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	7565
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	5878
GGTGGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	5444
TTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	4726
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	4683
GTTGGAAGCGACGGGACGGTAGGCTTGGGCCCAAGGAGTG	4033
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3930
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTGC	3777
GTTGGAAGCGACGGGACGGTAAGGCTGGGCCCAAGGAGTG	3700
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3346
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3342
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	3119
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGT	2986
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGATG	2715
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2713
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2659
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2635
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2609
GTTGGAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2500
TTTGGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	2437
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1991
GTTGGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1989
GTCGAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1842
GTTGGCAAGCGACGGGACGGTAAGGCTTGGGCCCAAGGAGTG	1839

GGGGAAAATTTTCCCGGGGAAAC	28885
GGGAAAATTTTCCCGGGGAAAATTTTCC	27528
GGGGAAAATTTTCCCGGGGAAAATTTTCC	26734
GGGGAAAATTTTCCCGGGGAAAATTTTCC	25567
GGGGAAAATTTTCCCGGGGAAATTTTCC	25363
GGGGAAAATTTTCCCGGGAAAATTTTCC	22129
GGGGAAAATTTTCCCGGGGAAATTTTCC	20528
GGGGAAAATTTTCCCGGGGAAAATTTTCC	18488
GGGGAAAATTTTCCCGGGGAAATTTTCC	14789
GGGGAAATTTTCCCGGGGAAAATTTTCC	13789
GGGGAAAATTTTCCCGGGGAAAATTTCC	13618
GGGGAAAATTTTCCCGGGGAAAATTTCC	12188
GGGGAAAATTTTCCCGGGGAAAATTTCC	11202
GGGGAAAATTTTCCCGGGGAAAATTTCC	10651
GGGGAAAATTTTCCCGGGGAAATTTTCC	9828
GGGGAAAATTTTCCCGGGGTAATTTTCC	9817
GGGGAAAATTTTCCCGGGGAAAATTTCC	9097
GGGGAAAATTTTCCCGGGGAAAATTTCC	8683
GGGGAAAATTTTCCCGAGGAAAATTTTCC	8659
GGGGATAATTTTCCCGGGGAAAATTTCC	8512
TGGGAAAATTTTCCCGGGGAAAATTTCC	8250
GGGGAAAATTTTCCCGTGGAAAATTTTCC	8185
GGGGAAAATTTTCCCGGGGATAATTTTCC	8177

Table S8: Original sequence and 25 most frequent mutations of FT2_G4A4T4C4_II

Sequence	Sequence count
GGGGAAAATTTTCCCGGGGAAAATTTTCC	6549166
GGGGAAAATTTTCCCGGGGAAAG	45606
GGGGAAAATTTTCCCGGGGAAAATTTTCC	25874
GGGAAAATTTTCCCGGGGAAAATTTTCC	25002
GGGGAAAATTTTCCCGGGGAAATTTTCC	24447
GGGGAAAATTTTCCCGGGGAAAATTTTCC	23700
GGGGAAAATTTTCCCGGGAAAATTTTCC	20951
GGGGAAAATTTTCCCGGGGAAAC	20084
GGGTAAAATTTTCCCGGGGAAAATTTTCC	19289
GGGGAAAATTTTCCCGGGGAAATTTTCC	18918
GGGGAAAATTTTCCCGGGGAAAATTTTCC	16348
GGGGAAAATTTTCCCGGGGAAAATTTTACC	15108
GGGGAAAATTTTCCCGGGGAAAG	14713
GGGGAAAATTTTCCCGGGGAAATTTTCC	13451
GGGGAAAATTTTCCCGGGGAAAATTTTCC	13426
GGGGAAATTTTCCCGGGGAAAATTTTCC	12803
GGGGAAAATTTTCCCGGGGAAAATTTTCC	11949
GGGGAAAATTTTACC	11730
TGGGAAAATTTTCCCGGGGAAAATTTTCC	8438
GGTGAAAATTTTCCCGGGGAAAATTTTCC	8037
GGGGAAAATTTTCCCGGGTAAAATTTTCC	7069
GGGGATAATTTTCCCGGGGAAAATTTTCC	6921

GGGGAAAATTTTCACCGGGGAAAATTTTCCCC	6831
GGGGAAAATTTTCCCCGGGGAAAATTTTCAACC	6576
GGGGAAATTTTCCCCGGGGAAAATTTTCCCC	6558
GGGGAAAATTTTCCCCGGGGAAAATTTTCCAC	6555

Table S9: Original sequence and 25 most frequent mutations of FT2_G2A2T2C2

Sequence	Sequence count
GGAATTCGGAATTCGGAATTCGGAATTC	2040896
GAATTCGGAATTCGGAATTCGGAATTC	6073
GGAATTCGGAATTCGTAATTCGGAATTC	5990
GGAATTCGGAATTACGGAATTCGGAATTC	5539
GGAATTCGGAATTCGGAATTCGGAATTC	5123
GGAATTCGGAATTCGGAATTCGTAATTC	5085
GGAATTCGGAATTCGGAATTCGGAATTC	5080
GGAATTCGGAATTCGGAATTCGGAATTCAC	4697
GGAATTCGTAATTCGGAATTCGGAATTC	4674
GGAATTCGGAATTCGGAATTCGGAATTC	4548
GGAATTCGGAATTCGGAATTCGGAATTC	4547
GTAATTCGGAATTCGGAATTCGGAATTC	4165
GGAATTCGGAATTCGGAATTCGGAATTC	4057
GGAATTCGGAATTCGAATTCGGAATTC	3949
GGAATTCGGAATTCAGGAATTCGGAATTC	3664
GGAATTCGGAATTCGGAATTCAGGAATTC	3571
GGAATTCGGAATTCGGAATTCGGAATTCAC	3541
TGAATTCGGAATTCGGAATTCGGAATTC	3522
GGAATTCGGAATTCGTAATTCGGAATTC	3473
GGAATTCGTAATTCGGAATTCGGAATTC	3450
GGAATTCGGAATTCGGAATTCGGAATTC	3374
GGAATTCGGAATTCGGAATTCGGAATTC	3215
GGAATTCGGAATTCGGAATTCGGAATTC	3196
GGAATTCGGAATTCGGAATTCGGAATTC	3180
GGAATTCGGAATTCGGAATTCGGAATTC	3161
GGAATTCGGAATTCGGAATTCGGAATTC	3152

Table S10: Original sequence and 25 most frequent mutations of FT2_G3A3T3C3

Sequence	Sequence count
GGGAAATTTCCCGGAAATTTCCCGGAAATT	5535182
GGAAATTTCCCGGAAATTTCCCGGAAATT	21093
GGGAAATTTCCCGGAAATTCGCGGAAATT	19864
GGGAAATTTCCCGGAAATTTCCCGGAAATT	19079
GGGAAATTTCCCGGAAATTTCCCGGAAATT	18549
GGGAAATTTCCCGGAAATTTACCGGAAATT	17688
GGGAAATTTCCCGAAATTTCCCGGAAATT	16754
GGGAAATTTCCAGGGAAATTTCCCGGAAATT	16257
GGGAAATTTCCCGGAAATTTCCAGGGAAATT	16206
GGGAAATTTCCCTGGAAATTTCCCGGAAATT	15353

GGGAAATTTCCCGGGAAATTTCCGGGAAATT	14511
GGTAAATTTCCCGGGAAATTTCCCGGGAAATT	13353
GGGAAATTTCCCGGGAAATTTCCCGGAAATT	12584
TGGAAATTTCCCGGGAAATTTCCCGGGAAATT	12115
GGGAAATTTCCCGGTAAATTTCCCGGGAAATT	12041
GGGAAATTTCCCGGGAAATTTCCCTGGAAATT	11707
GGGAAATTCCTCCGGGAAATTTCCCGGGAAATT	11475
GGGAAATTTCCCGTGAAATTTCCCGGGAAATT	11190
GGGAAATTTCCCGGGAAATTTCCCGGGAATT	10863
GGGAAATTTCACTCCGGGAAATTTCCCGGGAAATT	10839
GGGAAATTTCCCGGGAAATTTCACTCCGGGAAATT	10796
GGGAAATTTCCCGGGAAATTTCCCGTGAAATT	10637
GGGAATTTCCCGGGAAATTTCCCGGGAAATT	10332
GTGAAATTTCCCGGGAAATTTCCCGGGAAATT	10015
GGGAAATTTACTCCGGGAAATTTCCCGGGAAATT	9290
GGGAAATTTCCCGGGAAATTTCCCGGTAAATT	8848

Table S11: Original sequence and 25 most frequent mutations of FT2_TGCA

Sequence	Sequence count
TGCATGCATGCATGCATGCATGCATGCA	6627344
TGCATGCATGCATGCATGCTTGCATGCATGCA	45693
TGCATGCATGCTTGCATGCATGCATGCATGCA	36139
TGCAAGCATGCATGCATGCATGCATGCATGCA	26613
TGCATGCATGCATGCATGCATGCATGCA	24728
TGCATGCATGCATGCATGCATGCATGCATGAA	19114
TGCATGCATGAAATGCATGCATGCATGCATGCA	18958
TGTATGCATGCATGCATGCATGCATGCATGCA	12451
TGCATGCATTCATGCATGCATGCATGCATGCA	11260
TGCATGTATGCATGCATGCATGCATGCATGCA	10812
TGCATGCATGCATGTATGCATGCATGCATGCA	10365
TGCATGCATGTATGCATGCATGCATGCATGCA	10306
TGCATGCATGCATGAATGCATGCATGCATGCA	9163
TGCATGCATGCATGCATGTATGCATGCATGCA	8532
TGCATGCATGCATGCATGCATGTATGCA	7920
TGCATGCATGCATGCATGCATGTATGCATGCA	7487
TGCATGCATGCTGCATGCATGCATGCATGCA	7444
TGCATGCATGCATGCATGCATGCATGCATAA	7208
TGCATGCATCATGCATGCATGCATGCATGCA	7143
TGCATGCATGCATGCATGAAATGCATGCATGCA	7044
TGCATGCATGCATGCATGCATGCATGAAATGCA	6912
TGCATGAAATGCATGCATGCATGCATGCATGCA	6769
TGAAATGCATGCATGCATGCATGCATGCATGCA	6711
TGCATGCATGCATGCATGCATGATGCATGCA	6706
TGCATGCATGCATCATGCATGCATGCATGCA	6549
TGCATGCATGCATGCATGCATGCATGACATGCA	6531

Table S12: Original sequence and 25 most frequent mutations of D3_TGCA

Sequence	Sequence count
TGCATGCATGCATGCATGCATGCATGCATGCA	398611
TGCATGCATGCATGCATGCATGC T TGCATGCATGCA	1654
TGCATGCATGCATGCATGC T TGCATGCATGCATGCA	1560
TGCATG T ATGCATGCATGCATGCATGCATGCA	1190
TGCATGCATGCATGCATGCATGCATGCATGC A	1026
TGCATGCATGCATGCATGCATGCATG T ATGCATGCA	588
TGCATGC T TGCATGCATGCATGCATGCATGCATGCA	555
TGCATGCATGCATGCATGC T ATGCATGCATGCA	545
TGCATGCATGCATGCATGCATGCATGCATGCATGC A CA	530
TG T ATGCATGCATGCATGCATGCATGCATGCATGCA	518
TGCATGCATGCATGCATGCATGCATGCATGCATGC A ATGCA	498
TGCATGCATG T ATGCATGCATGCATGCATGCA	459
TGCATGCAT T CATGCATGCATGCATGCATGCA	455
TG A ATGCATGCATGCATGCATGCATGCATGCA	418
TGCATGCATGCATGCATGC A CATGCATGCATGCA	395
TGCATGCATGCATGCATGC A CATGCATGCATGCA	393
TGCATGCATGCATGCATG T ATGCATGCATGCATGCA	392
TGCATGCATGCATGCATGCATGC A CATGCATGCA	390
TGCATG A ATGCATGCATGCATGCATGCATGCA	379
TGCATG C TGCATGCATGCATGCATGCATGCA	368
TGCATGCATGCATGC A ATGCATGCATGCATGCA	364
TGCATGCATGCATGCATGC T CATGCATGCATGCA	345
TGCATGCATGCATGCATGCATGCATGCATGCATGC A A	344
TGCATGC A CATGCATGCATGCATGCATGCATGCA	335
TGCATGCATGCATGCATGC A ATGCATGCATGCA	332
TGCATGCATGCATGCATGCATGCATGCATGC A ATGCA	306

Table S13: Original sequence and 25 most frequent mutations of FT2_T4G4C4A4

Sequence	Sequence count
TTTTGGGGCCCCAAAATTTTGGGGCCCCAAAA	1745068
TTTTGGGGCCCC A AAAATTTTGGGGCCCCAAAA	9556
TTTTGGGGCCCCAAAAT T GGGGCCCCAAAA	7973
TTTTGGGGCCCCAAA T TTTGGGGCCCCAAAA	7625
TTTTGGGGCCCCAAAATTTTGG C CCCCAAAA	7388
TTTTGGGGCC CA AAAATTTTGGGGCCCCAAAA	6884
TTTTGG GC CCCCAAAATTTTGGGGCCCCAAAA	4855
TTTTGGGGCCCCAAAATTTTGGGGCC CA AAAA	4841
TTTTGGGGC T CCAAAATTTTGGGGCCCCAAAA	4576
TTTTGGGGCC T CAAAAATTTTGGGGCCCCAAAA	4488
TTTTGGGGCCCCAAAATTTTGGGGCCCCAA A	4334
TT TG GGGGCCCCAAAATTTTGGGGCCCCAAAA	4234
TTTTGGGGCCCCAAAATTTT T GGGGCCCCAAAA	4229
TTTTGGGGCC T AAAATTTTGGGGCCCCAAAA	4204
TTTTGGGGCCCCAAAATTTTGGGGC T CCAAAA	3987
TTTTGGGG T CCAAAATTTTGGGGCCCCAAAA	3946
TTTTGGGGCCCCAAAATTTT A GGGGCCCCAAAA	3878
TTTTGGGGCCCCAAAATTTTGG A GCCCCAAAA	3860

TTTTGGGGCCCCAAAATTTTGAAGGCCCAAAA	3606
TTTTGGGGCCCCAAAATTTTGGGAACCCCAAAA	3308
TTTTGGGGCCCCAAAATTTGTGGGCCCAAAA	2968
TTTTGGGGCCAACAAAATTTTGGGGCCCCAAA	2940
TTTTGTGGGCCCAAAAATTTTGGGGCCCCAAA	2649
TTTTGGGGCCCCAAAATTTTGGGGCCCCAAA	2524
TTTTGGGGCCCCAAAATTTTGGGGCCCCAAA	2469
TTTTGGGGCCCCAAAATTTTGGGGCCAACAAA	2430

Table S14: Original sequence and 25 most frequent mutations of D3_T4G4C4A4

Sequence	Sequence count
TTTTGGGGCCCCAAAATTTTGGGGCCCCAAA	5284258
TTTTGGGGCCCCAAAATAATTGGGGCCCCAAA	56226
TTTTGGGGCCCCAAATTTTGGGGCCCCAAA	33501
TTTTGGGGCCCCAAAATTTGGGGCCCCAAA	31991
TTTTGGGGCCCAAAAATTTTGGGGCCCCAAA	19746
TTTTGGGGCCCCAAAATTTTGGGCCCAAAA	18308
TTTTGGGGCCCCAAAATTTTGGGGCCCAAAA	18108
TTTTGGGGCCCAAAAATTTTGGGGCCCCAAA	16172
TTTTGGGGCCCCAAAATTTTGGGGCCCCAAA	16010
TTTTGGGCCCAAAAATTTTGGGGCCCCAAA	15538
TTTTGGGGCCCTAAAATTTTGGGGCCCCAAA	14219
TTTTGGGGCTCCAAAATTTTGGGGCCCCAAA	14016
TTTTGGGGCCCTCAAAAATTTTGGGGCCCCAAA	13073
TTTTGGGGCCCCAAAATTTTGGGGCCCCAA	12625
TTTTGGGGTCCCAAAAATTTTGGGGCCCCAAA	10980
TTTGGGGCCCCAAAATTTTGGGGCCCCAAA	10697
TTTTGGGGCCAACAAAATTTTGGGGCCCCAAA	9072
TTTTGGGGCCCCAAAATTTTGTGGGCCCAAAA	8441
TTTTGTGGGCCCAAAAATTTTGGGGCCCCAAA	8378
TTTTGGGGCCCCAAAATTTTGAAGGCCCAAAA	7925
TTTTGGGGCCCCAAAATTTTGGGGCCAACAAA	7891
TTTTTGGGCCCAAAAATTTTGGGGCCCCAAA	7708
TTTTGGGGCCCCAAAATTTTGGGAACCCCAAAA	7669
TTTTGTGGGCCCAAAAATTTTGGGGCCCCAAA	7171
TTTTGGGGCCCCAAAATTTTGAAGGGCCCCAAA	7018
TTTTGGGGCCCCAAAATTTTGAAGGCCCAAAA	6677

Figure S1

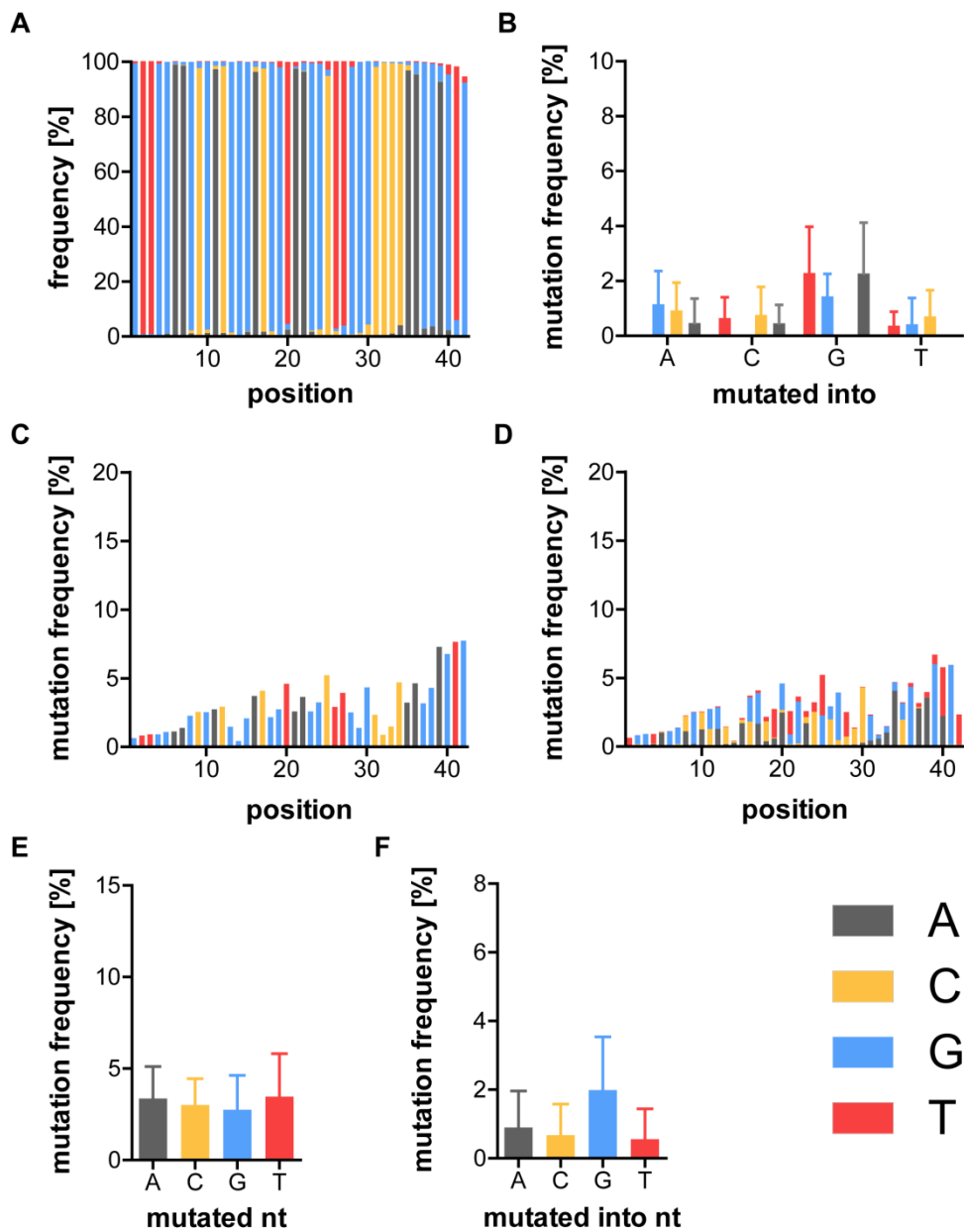


Figure S1: Mutation analysis of C12_T_PWO

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S2

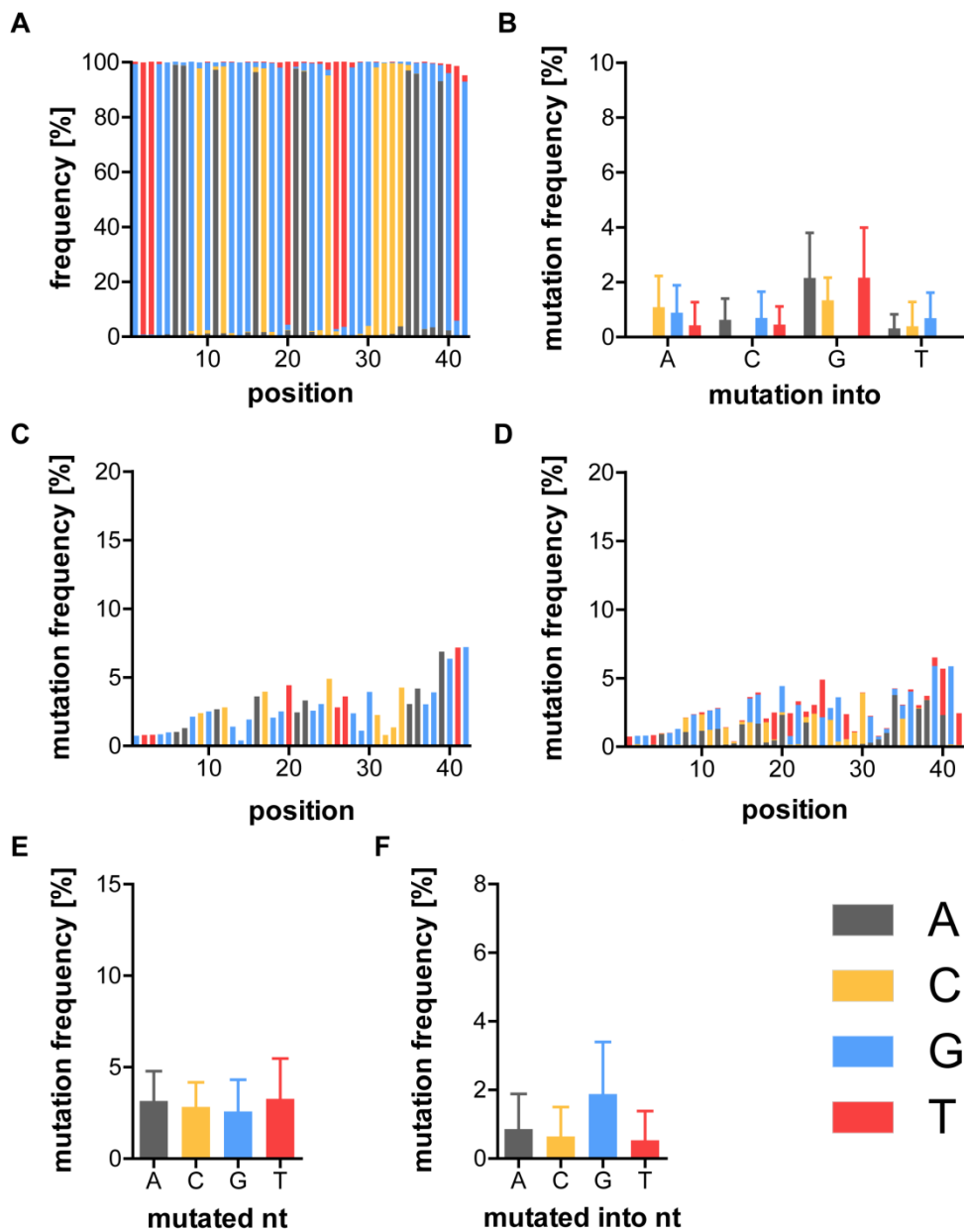


Figure S2: Mutation analysis of C12_T_Taq

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S3

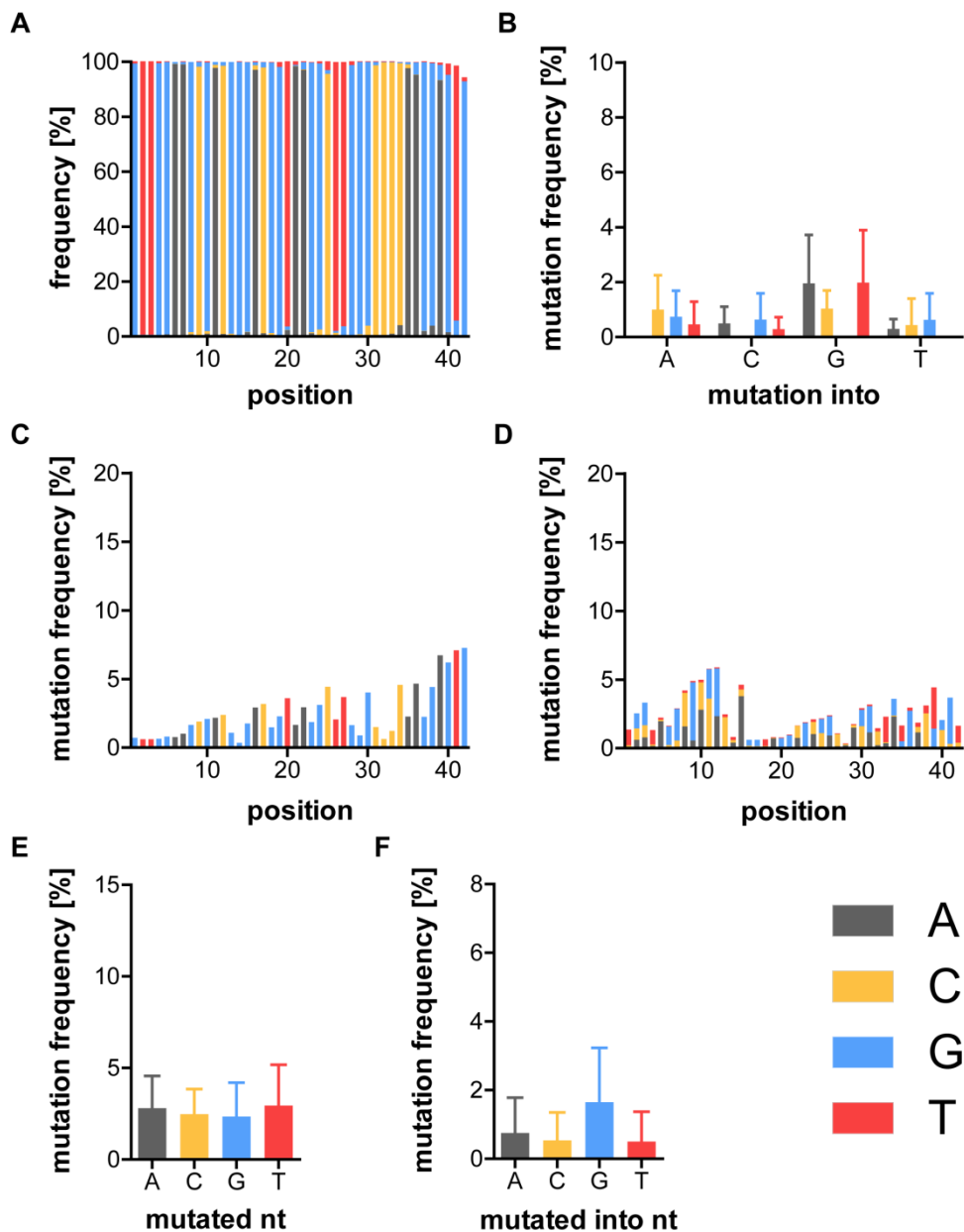


Figure S3: Mutation analysis of C12_T_w/o

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S4

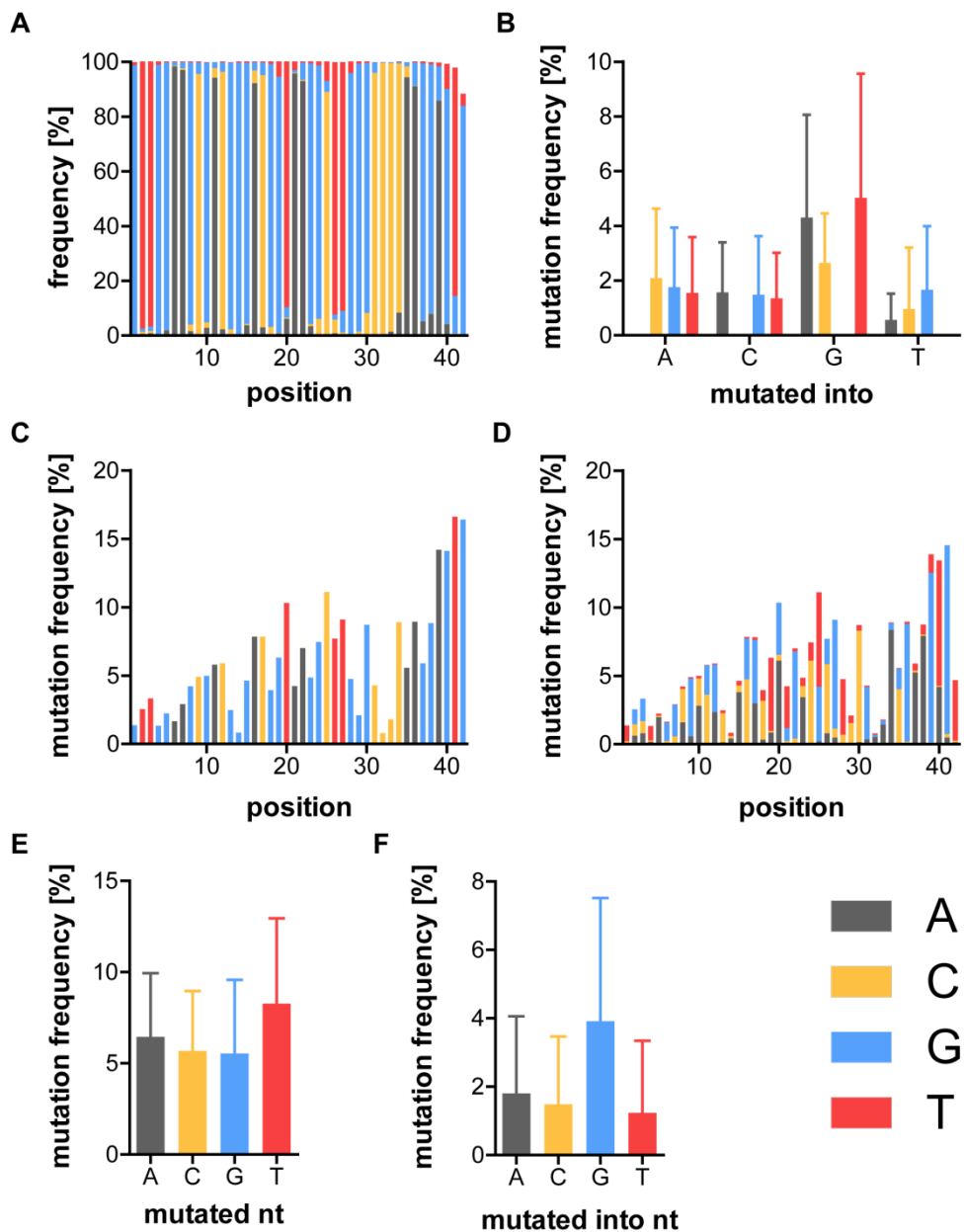


Figure S4: Mutation analysis of C12_EdU

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S5

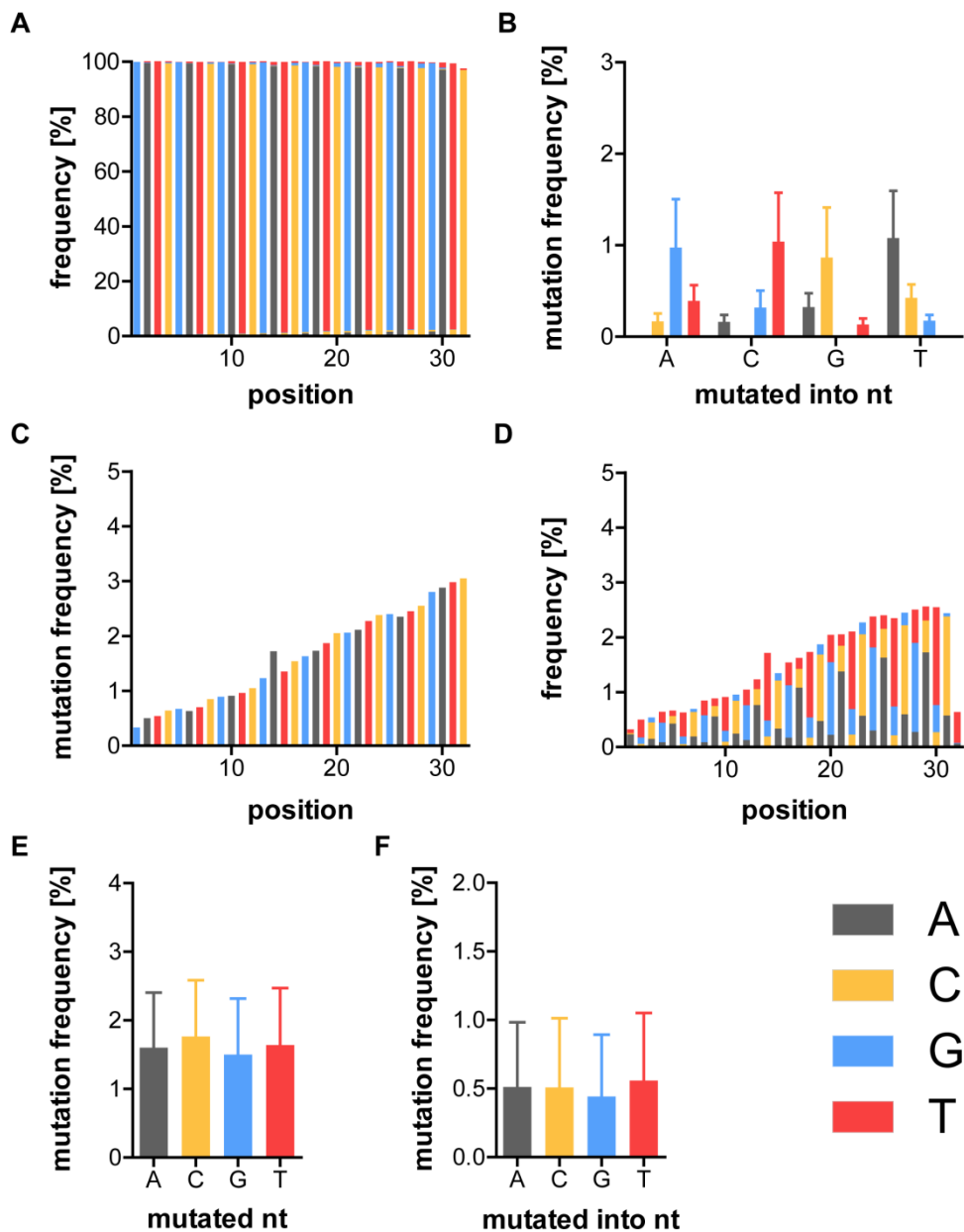


Figure S5: Mutation analysis of FT2_GATC

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S6

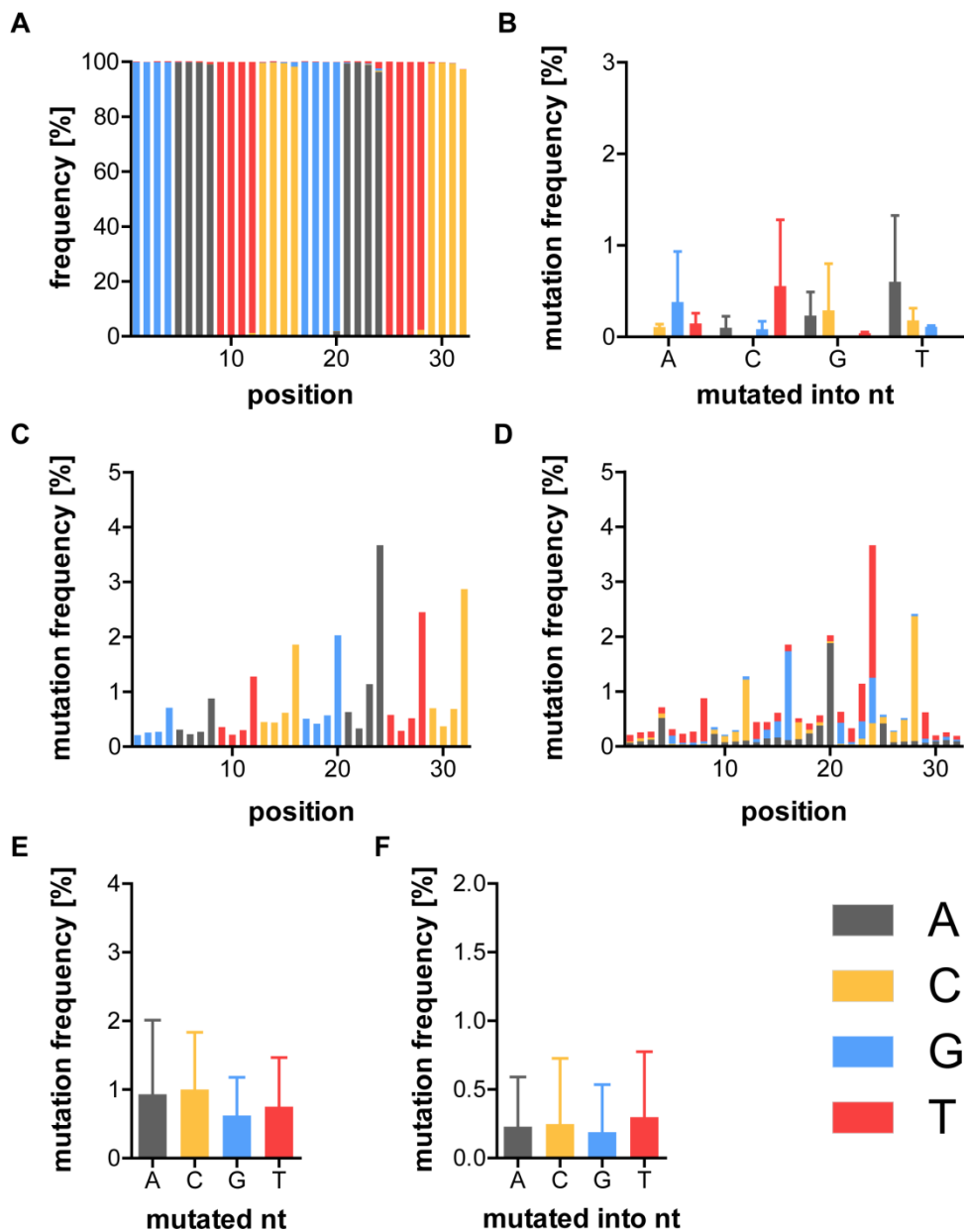


Figure S6: Mutation analysis of FT2_G4A4T4C4

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S7

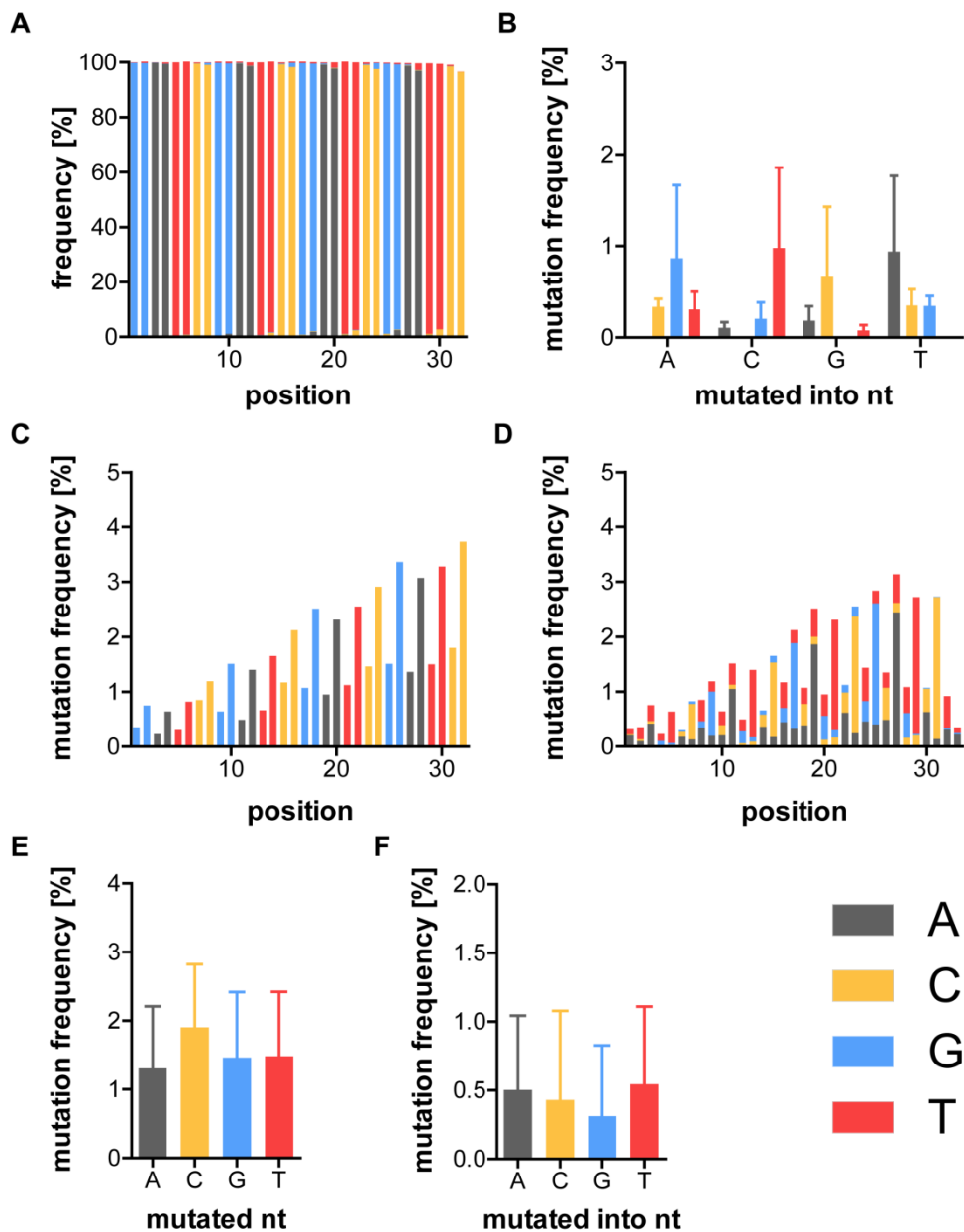


Figure S7: Mutation analysis of FT2_G2A2T2C2

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S8

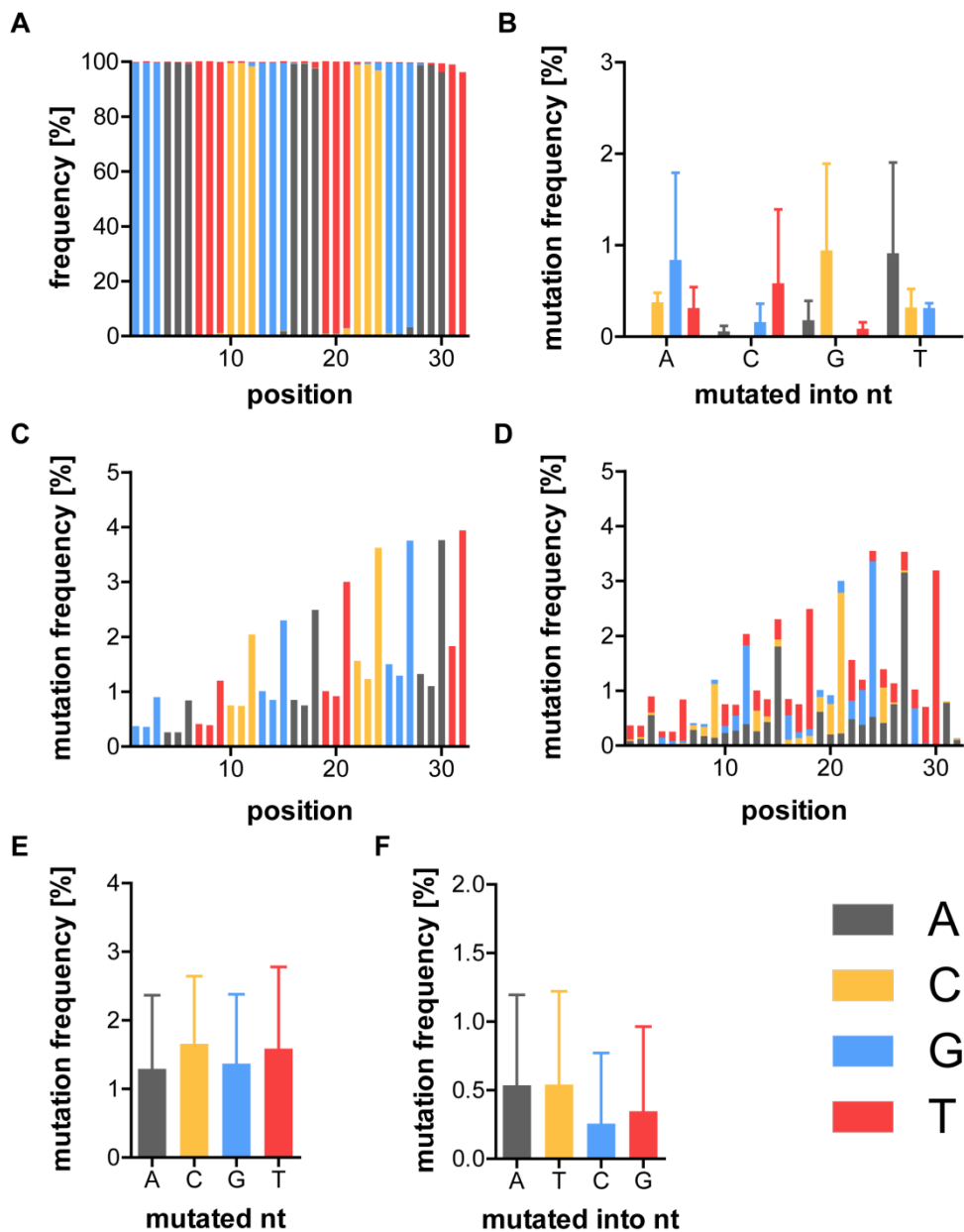


Figure S8: Mutation analysis of FT2_G3A3T3C3

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S9

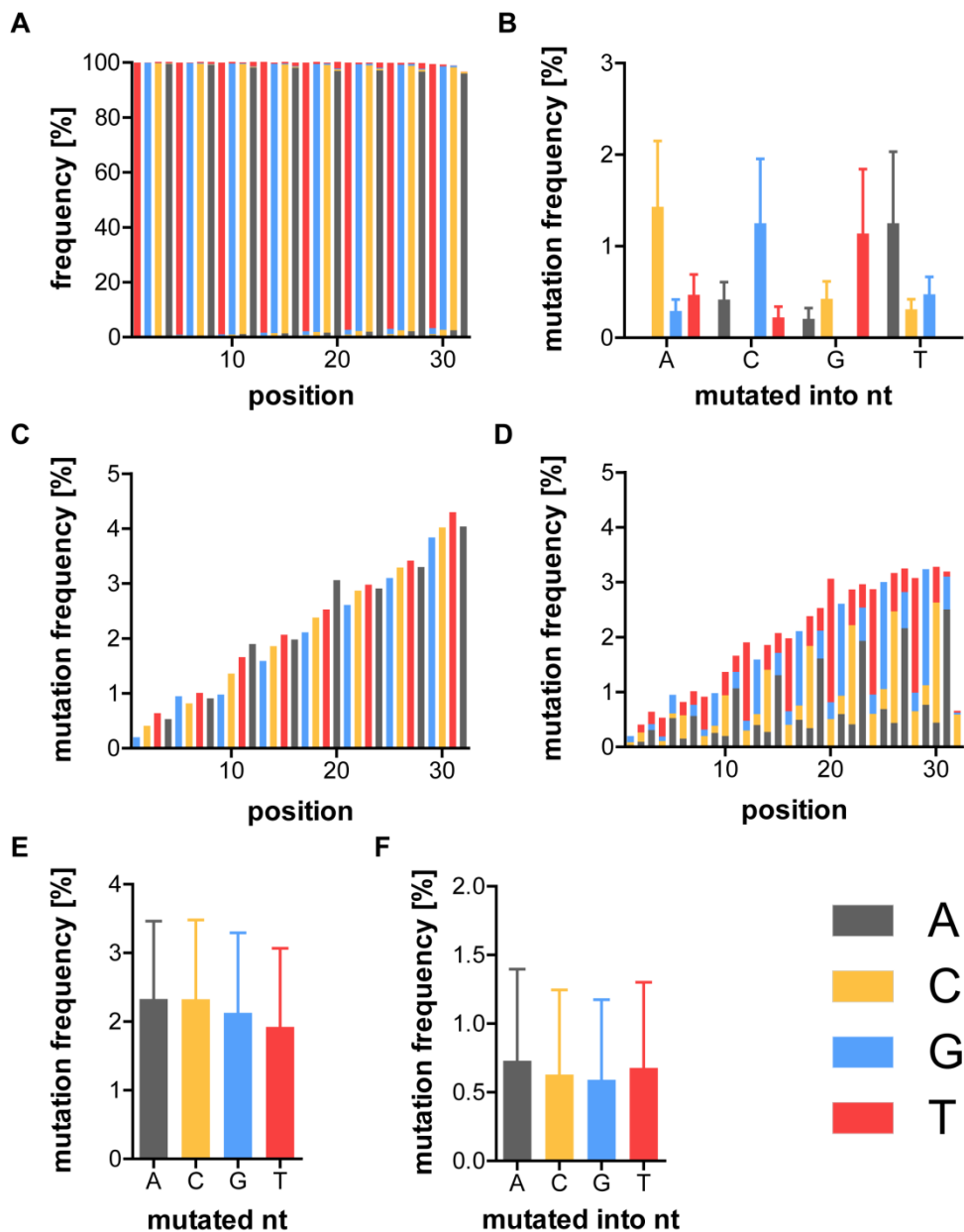


Figure S9: Mutation analysis of FT2_TGCA

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S10

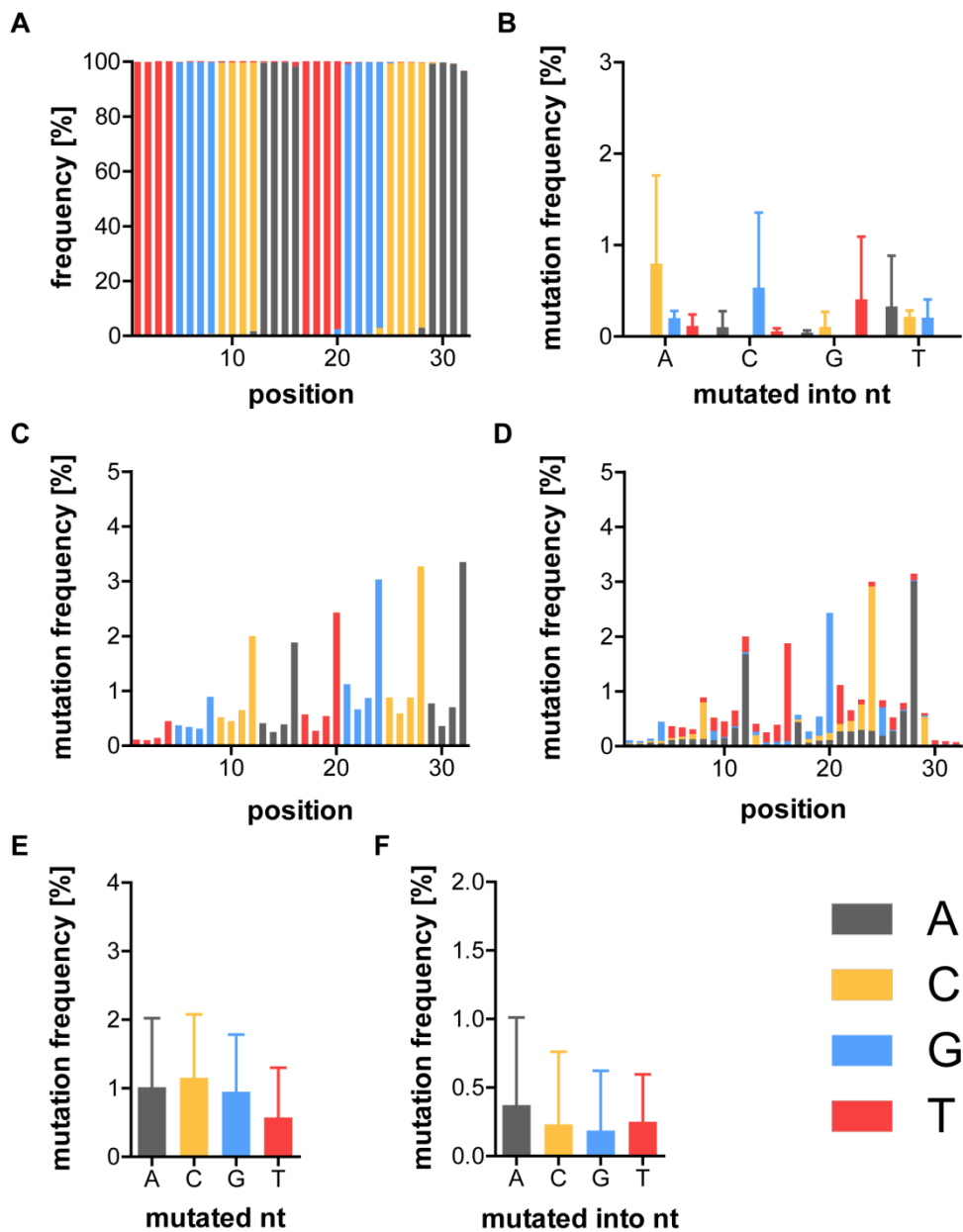


Figure S10: Mutation analysis of FT2_T4G4C4A4

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S11

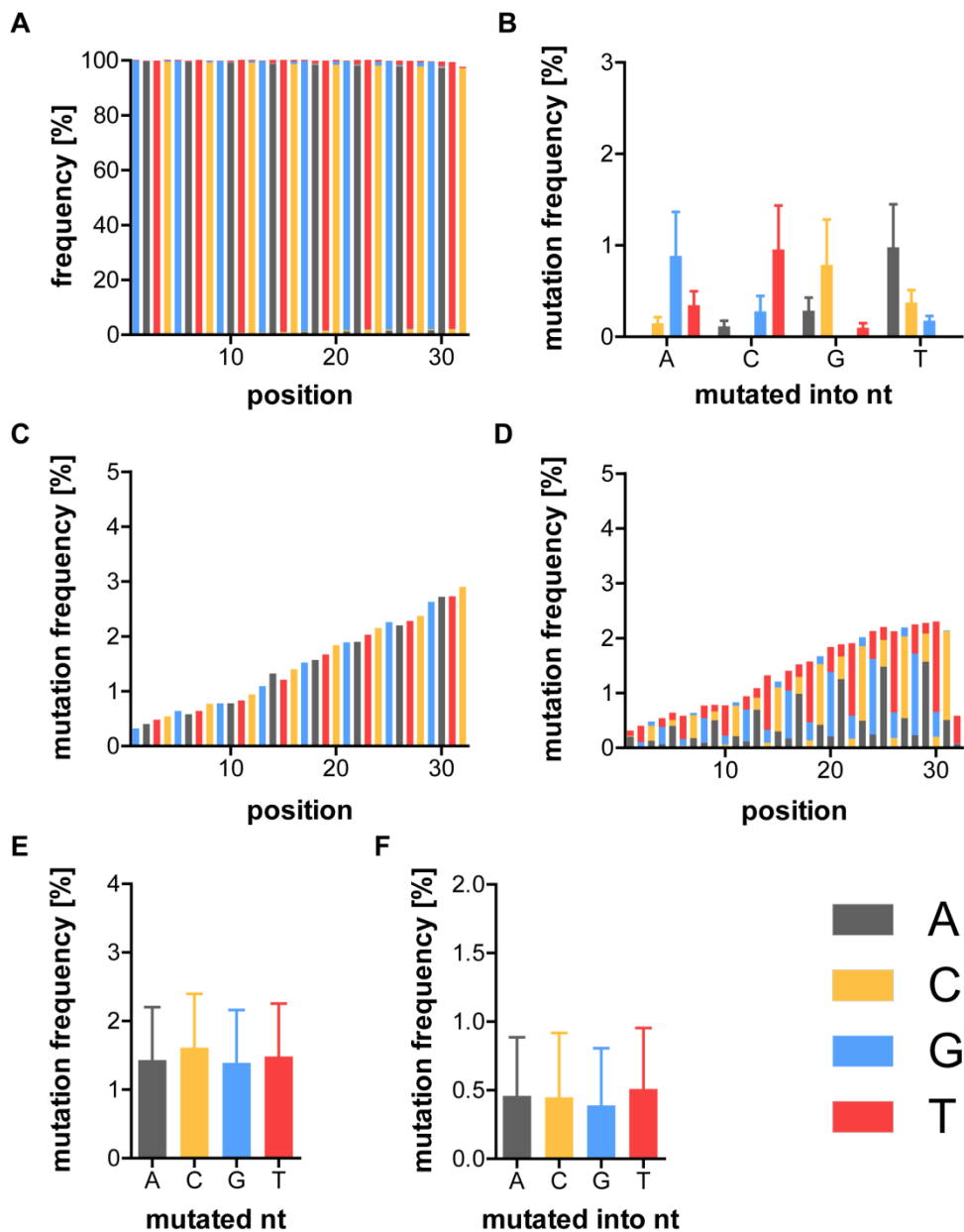


Figure S11: Mutation analysis of FT2_GATC_II

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S12

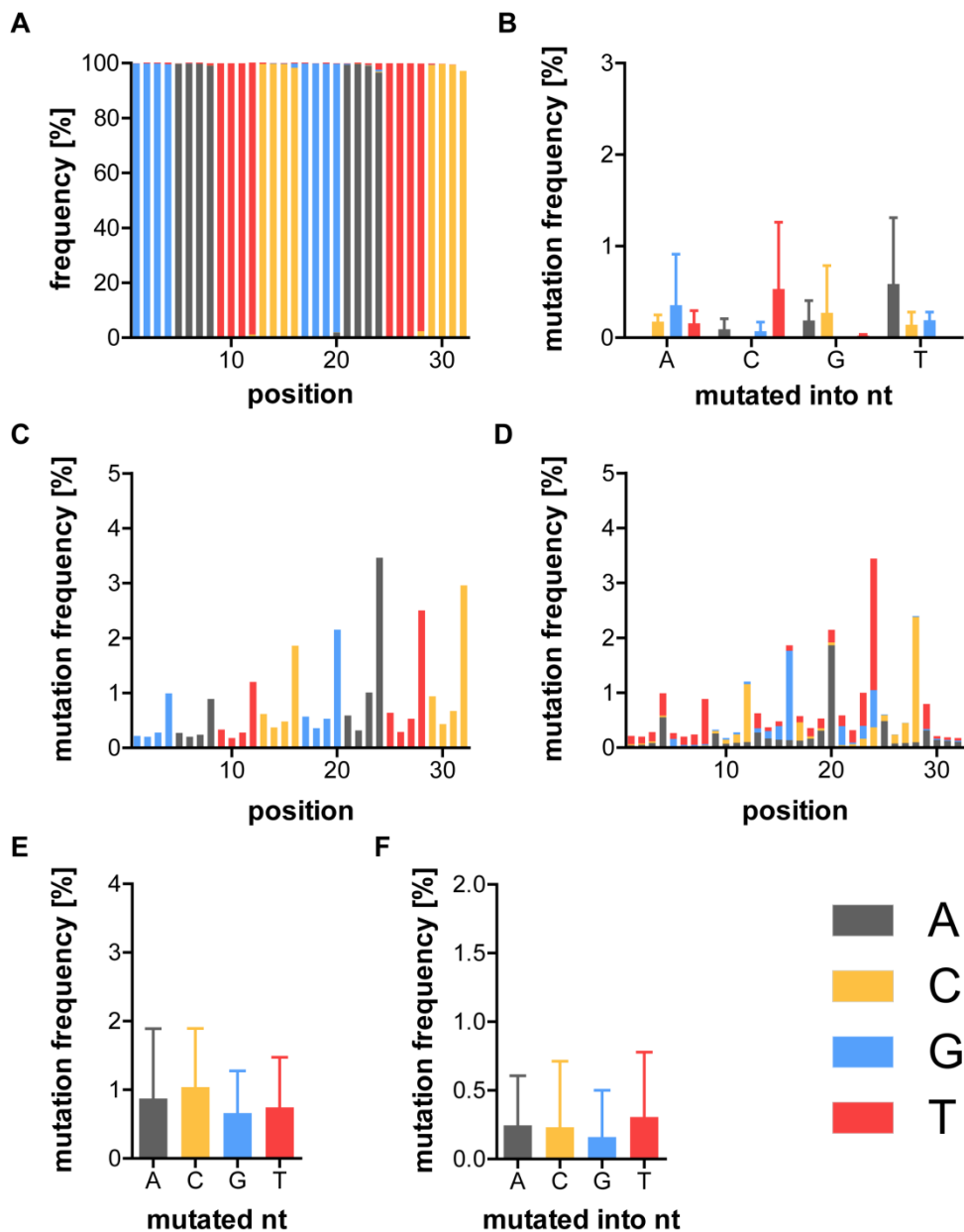


Figure S12: Mutation analysis of FT2_G4A4T4C4_II

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S13

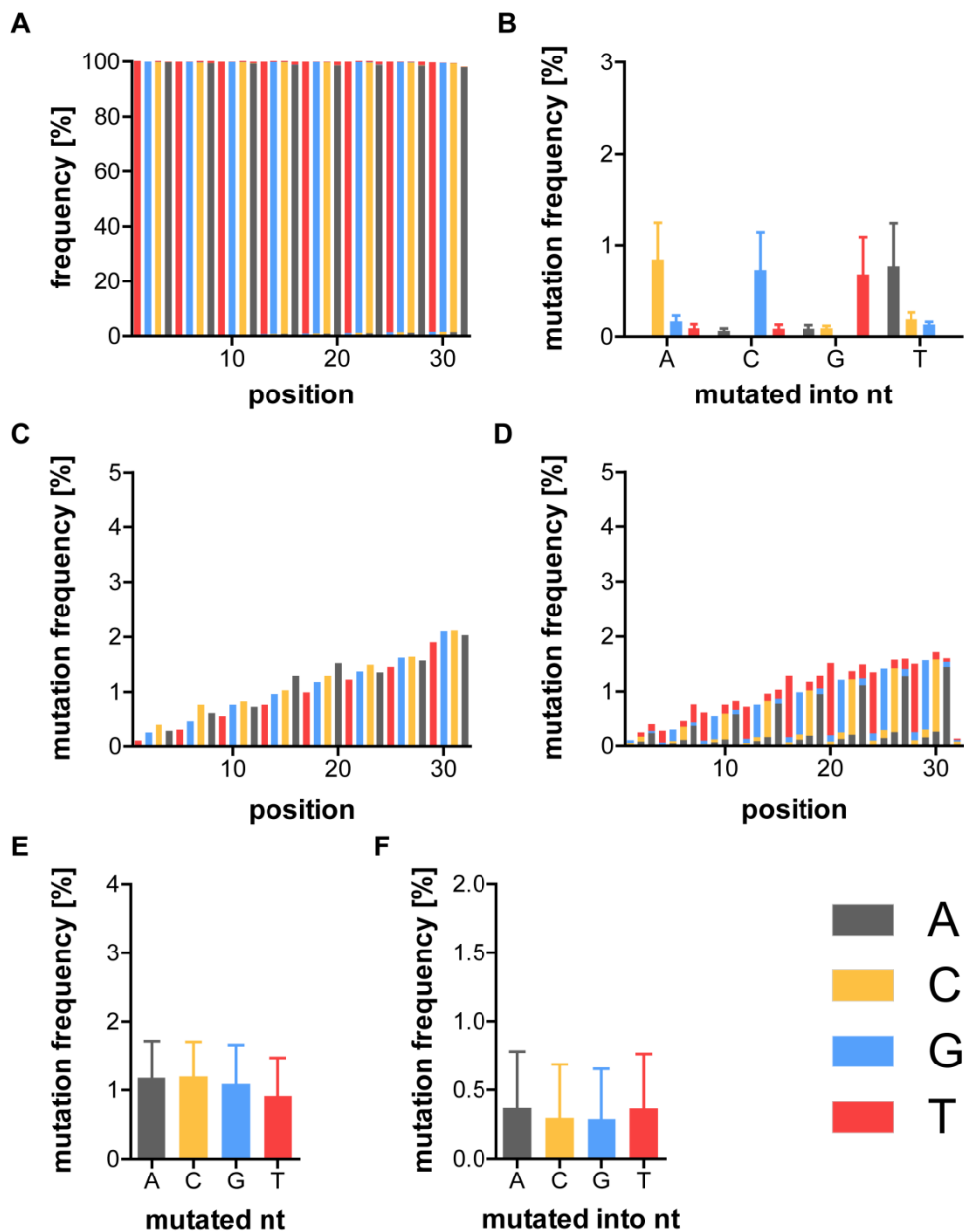


Figure S13: Mutation analysis of D3_TGCA

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S14

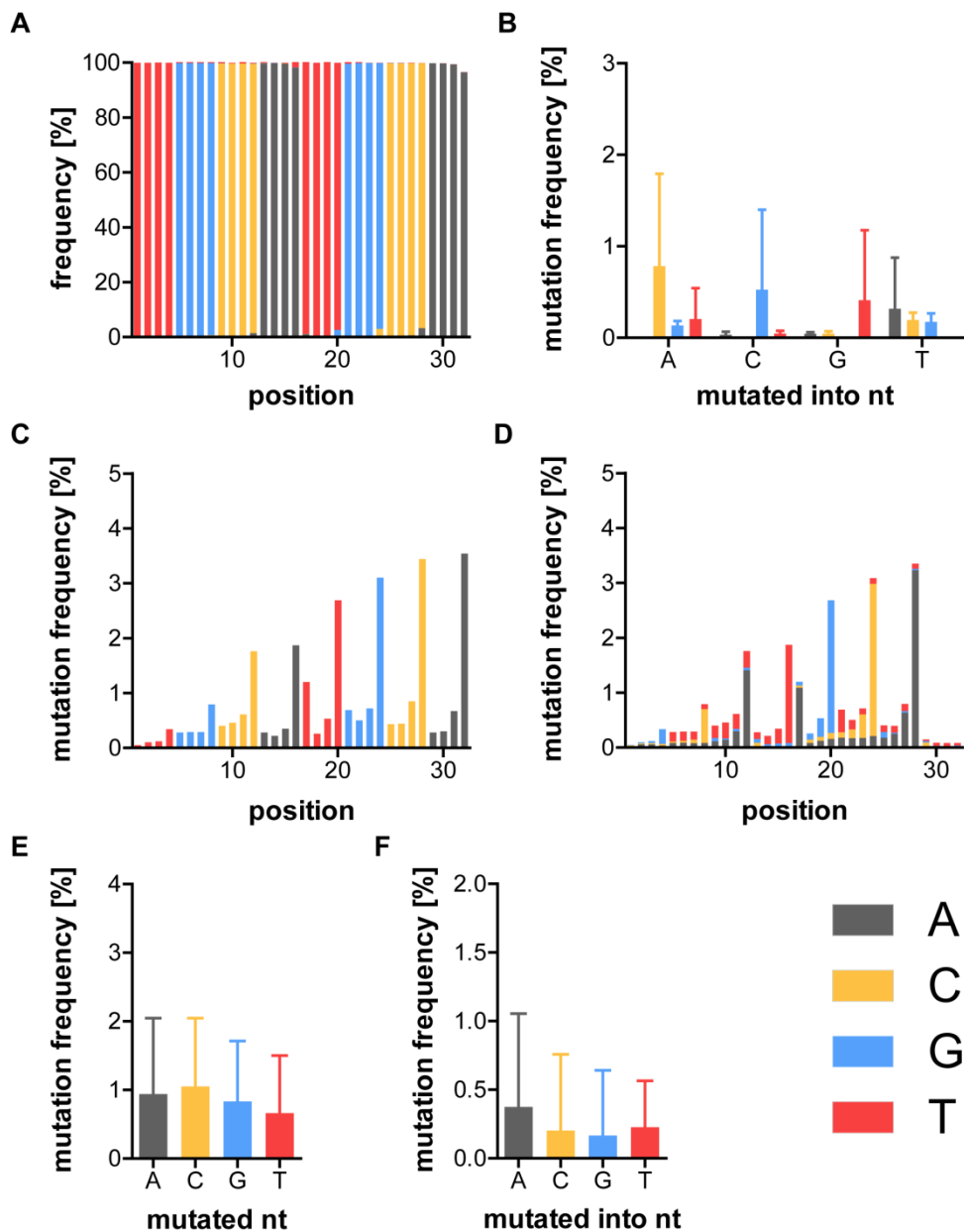


Figure S14: Mutation analysis of D3_T4G4C4A4

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S15

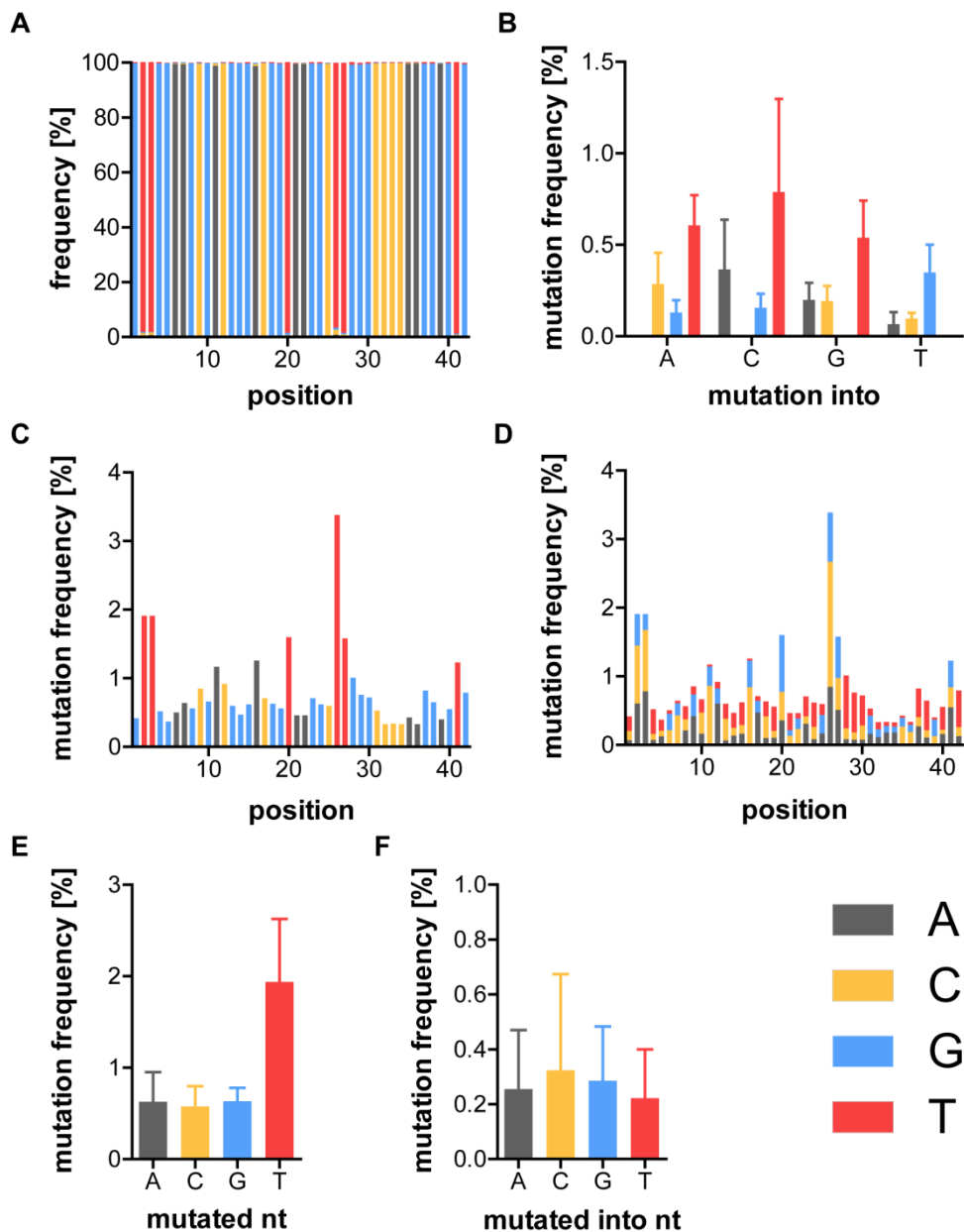


Figure S15: Mutation analysis of C12_EdU after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S16

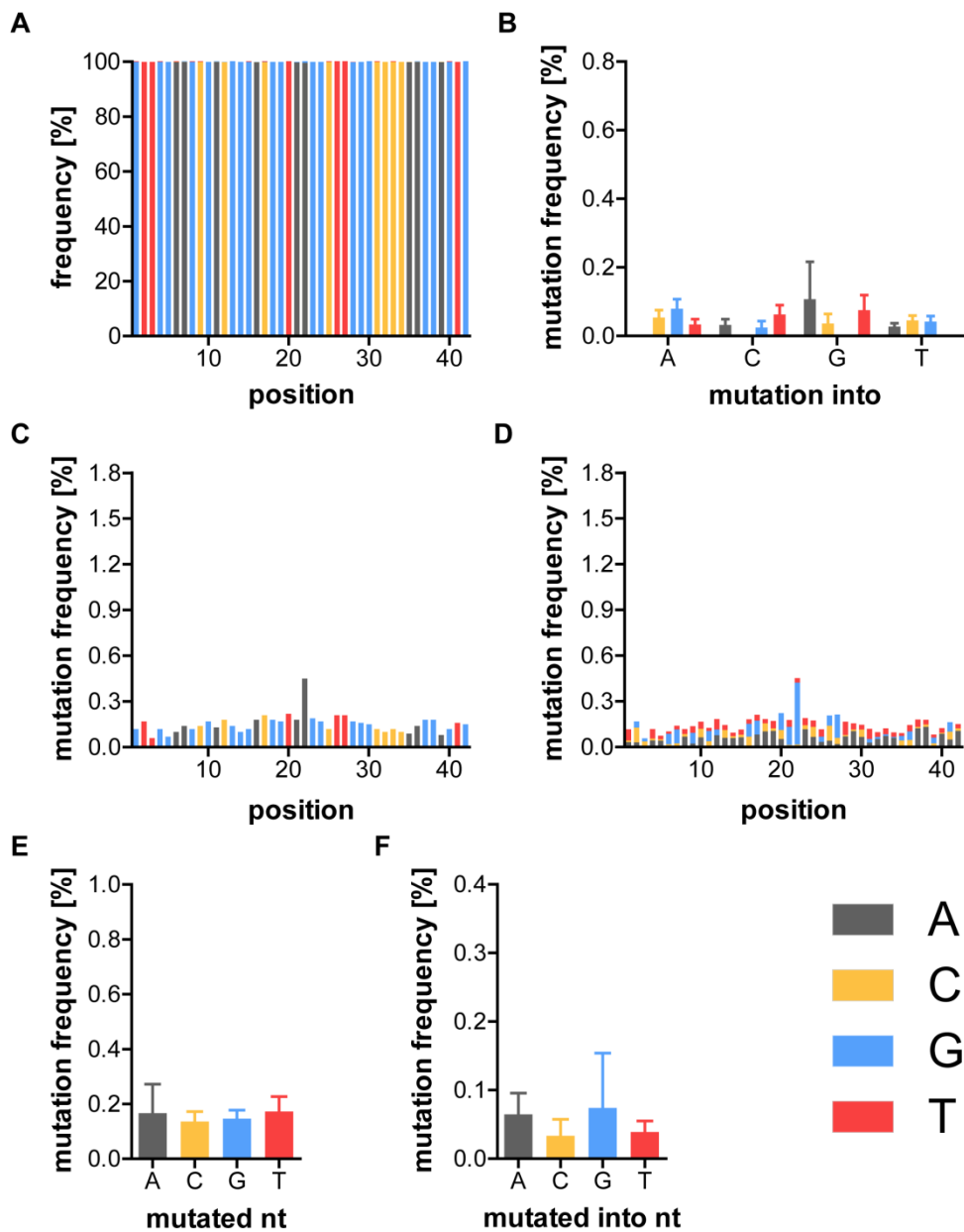


Figure S16: Mutation analysis of C12_T_PWO after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S17

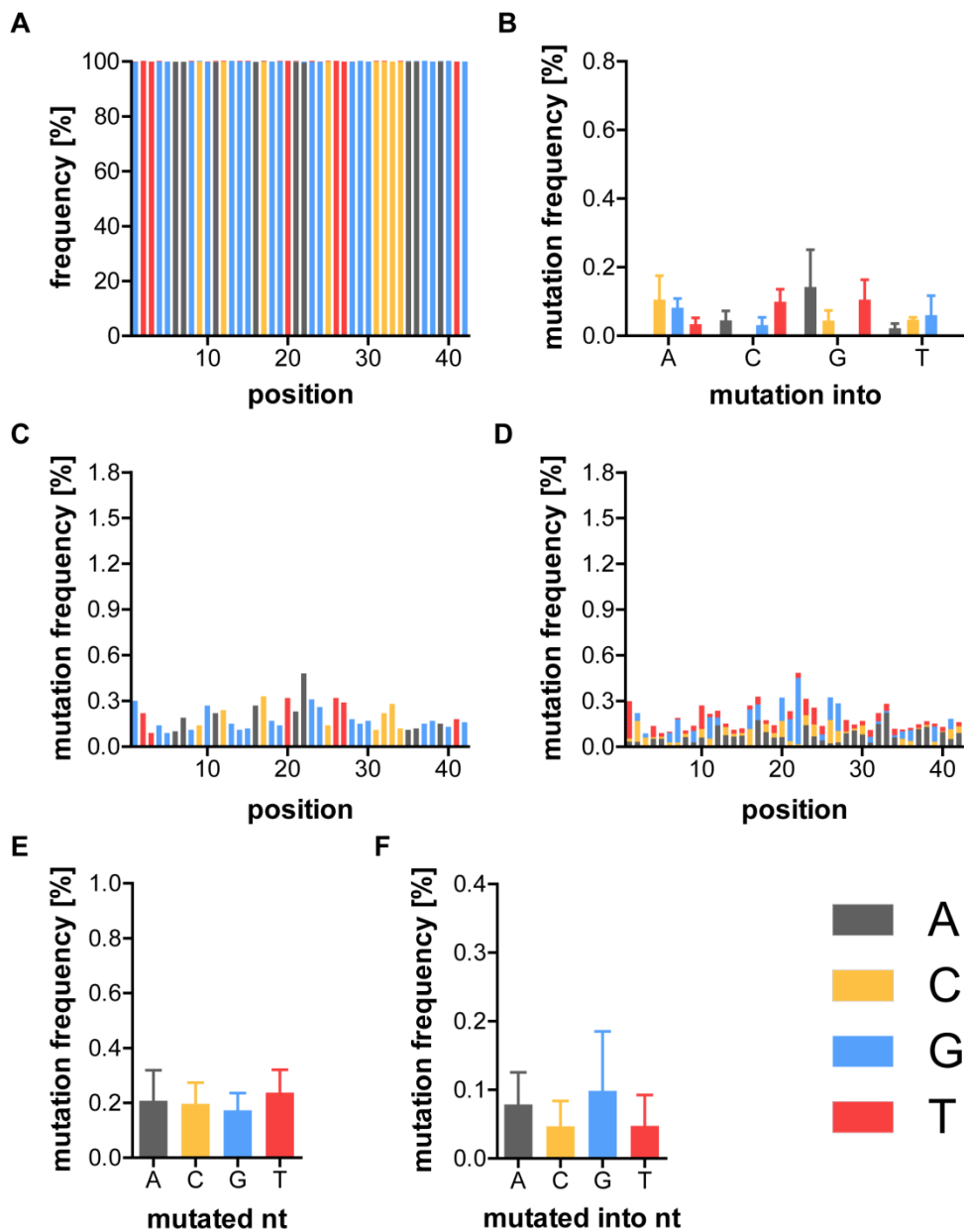


Figure S17: Mutation analysis of C12_T_PWO after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S18

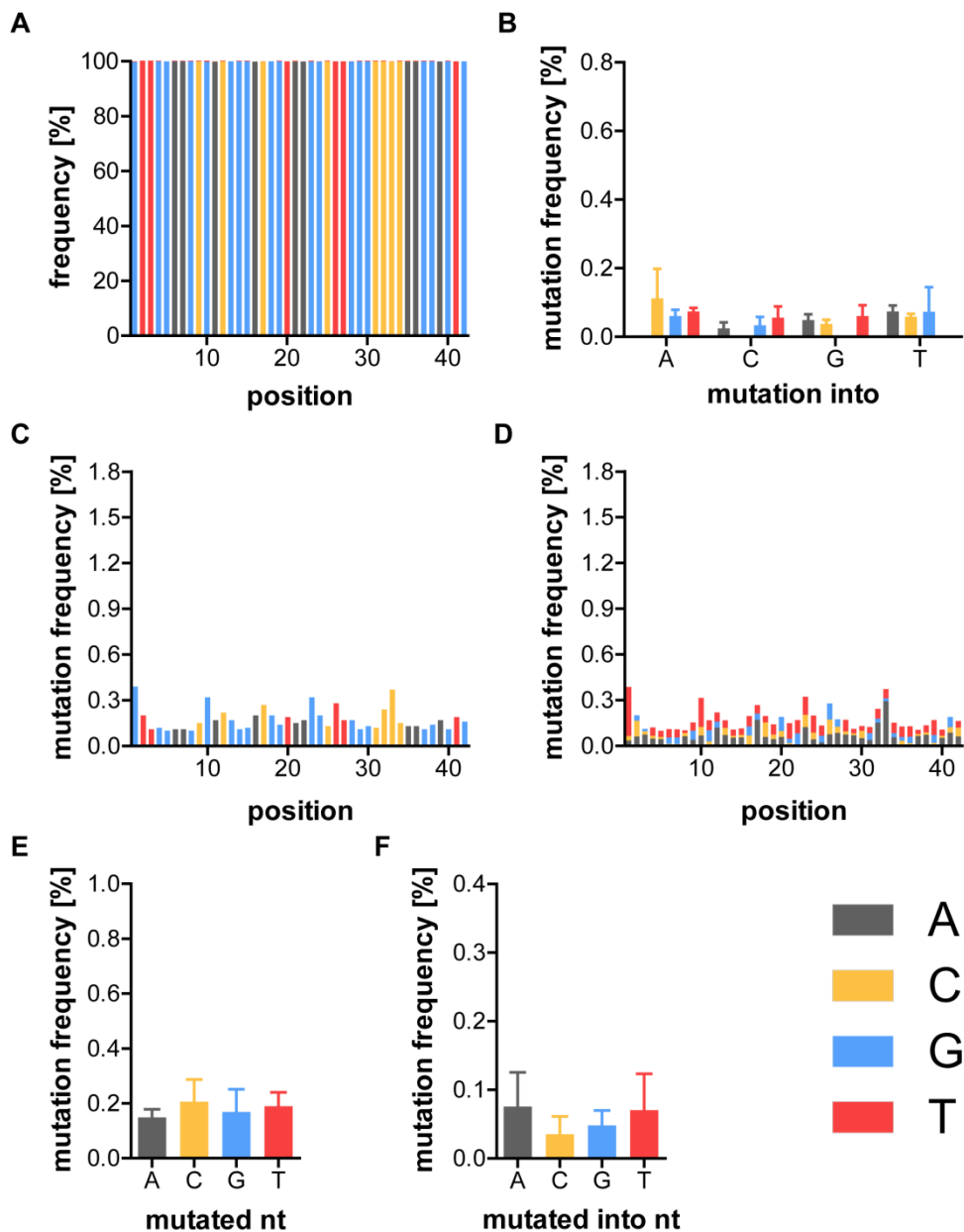


Figure S18: Mutation analysis of C12_T_w/o after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S19

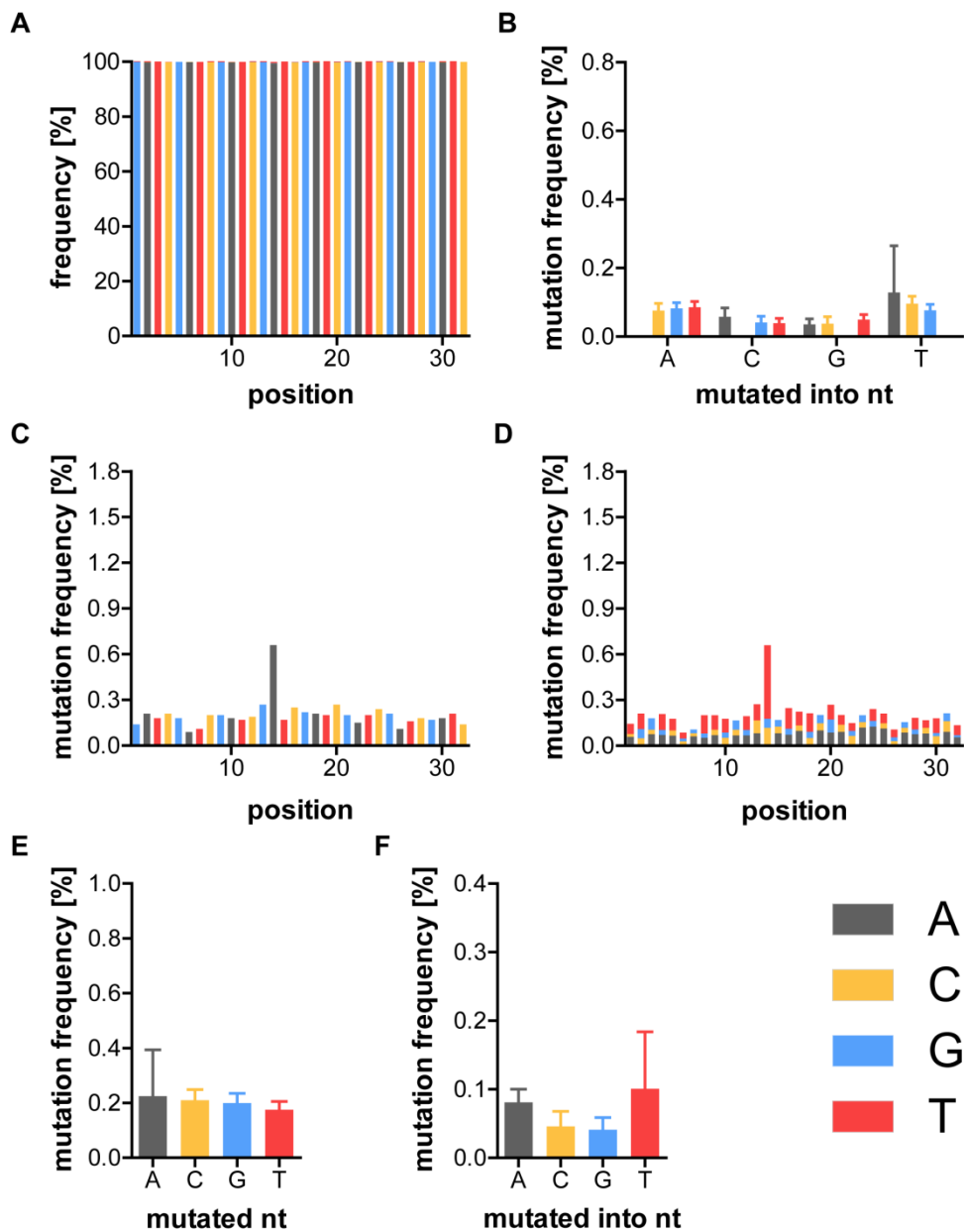


Figure S19: Mutation analysis of FT2_GATC after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S20

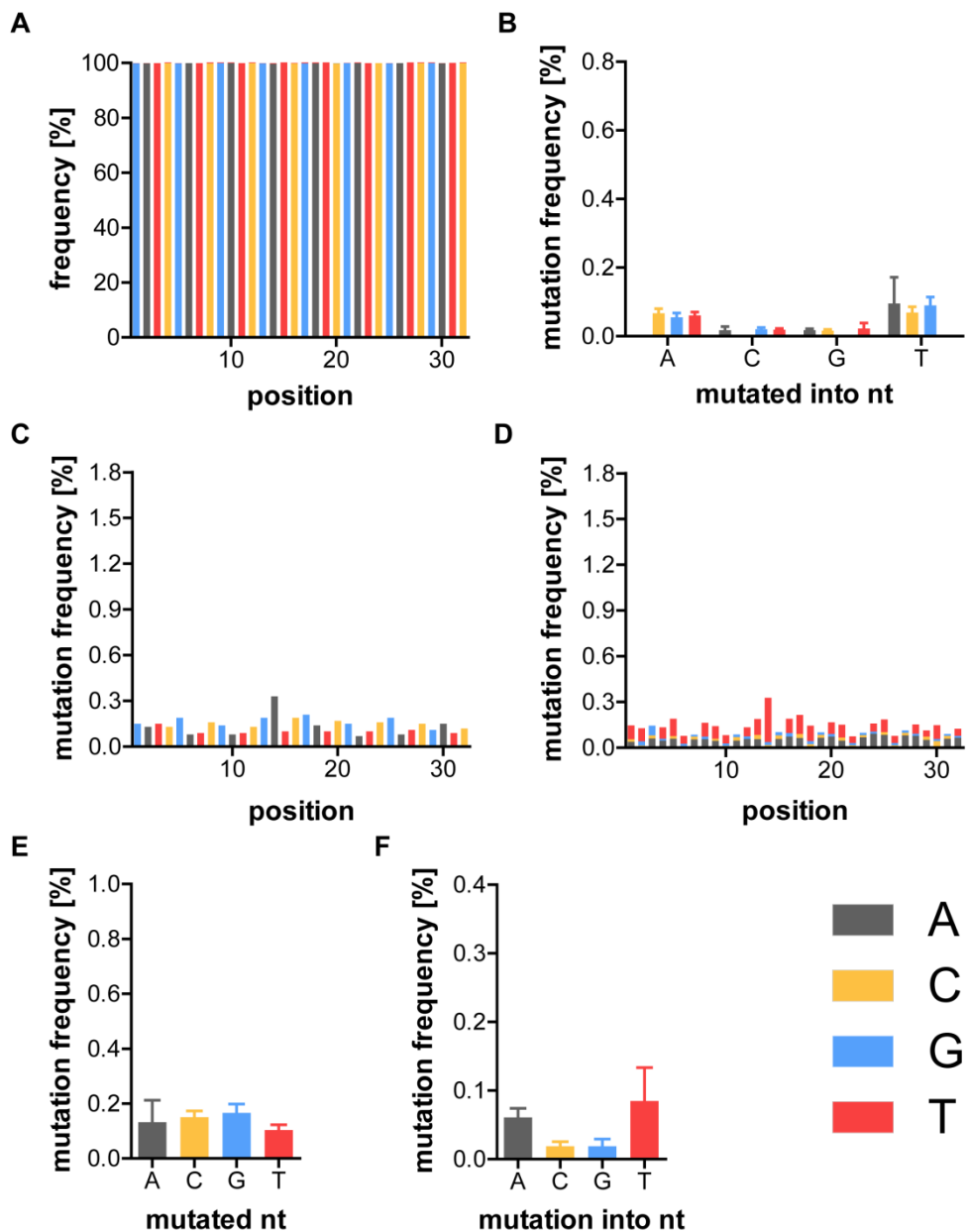


Figure S20: Mutation analysis of FT2_GATC_II after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S21

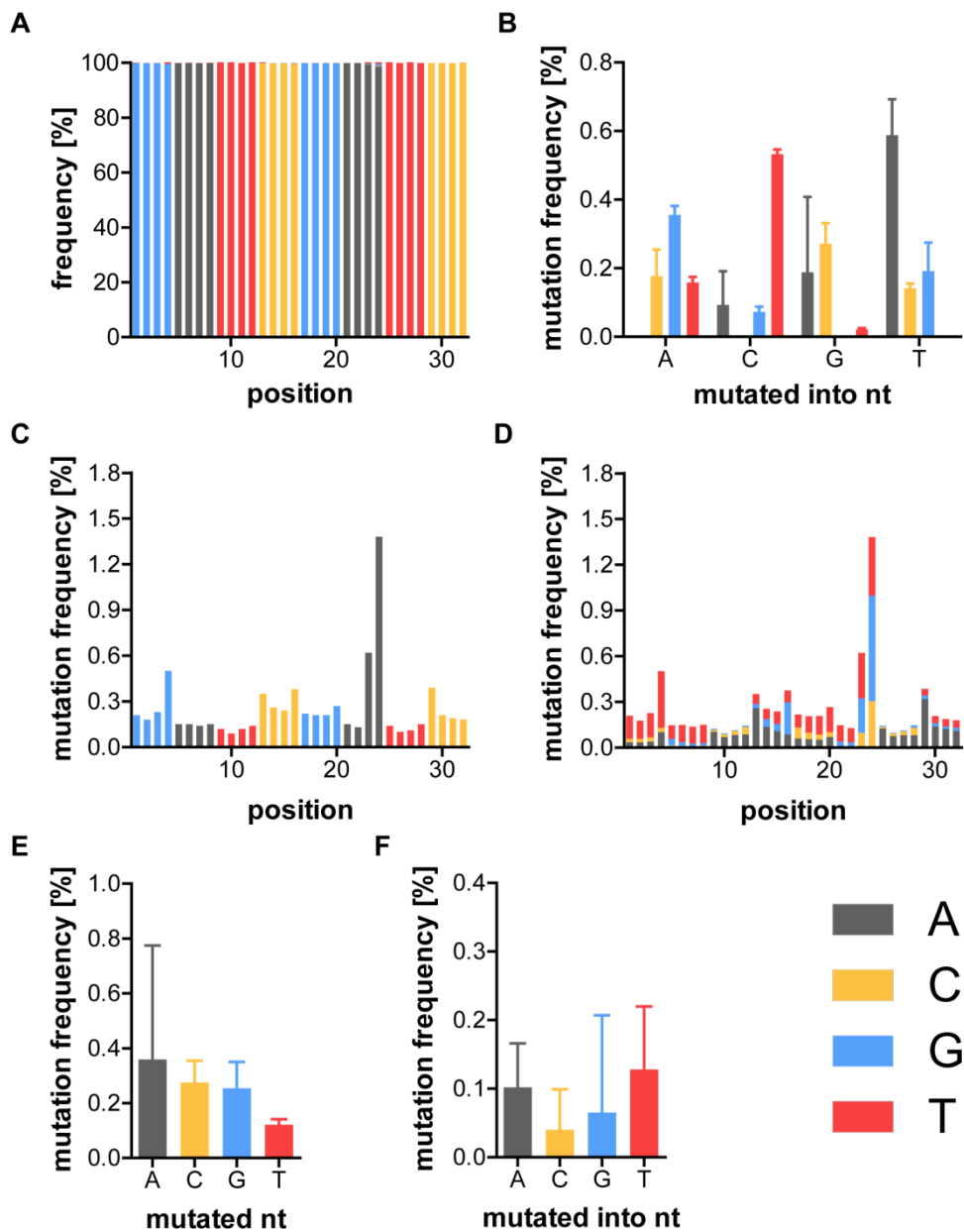


Figure S21: Mutation analysis of FT2_G4A4T4C4 after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S22

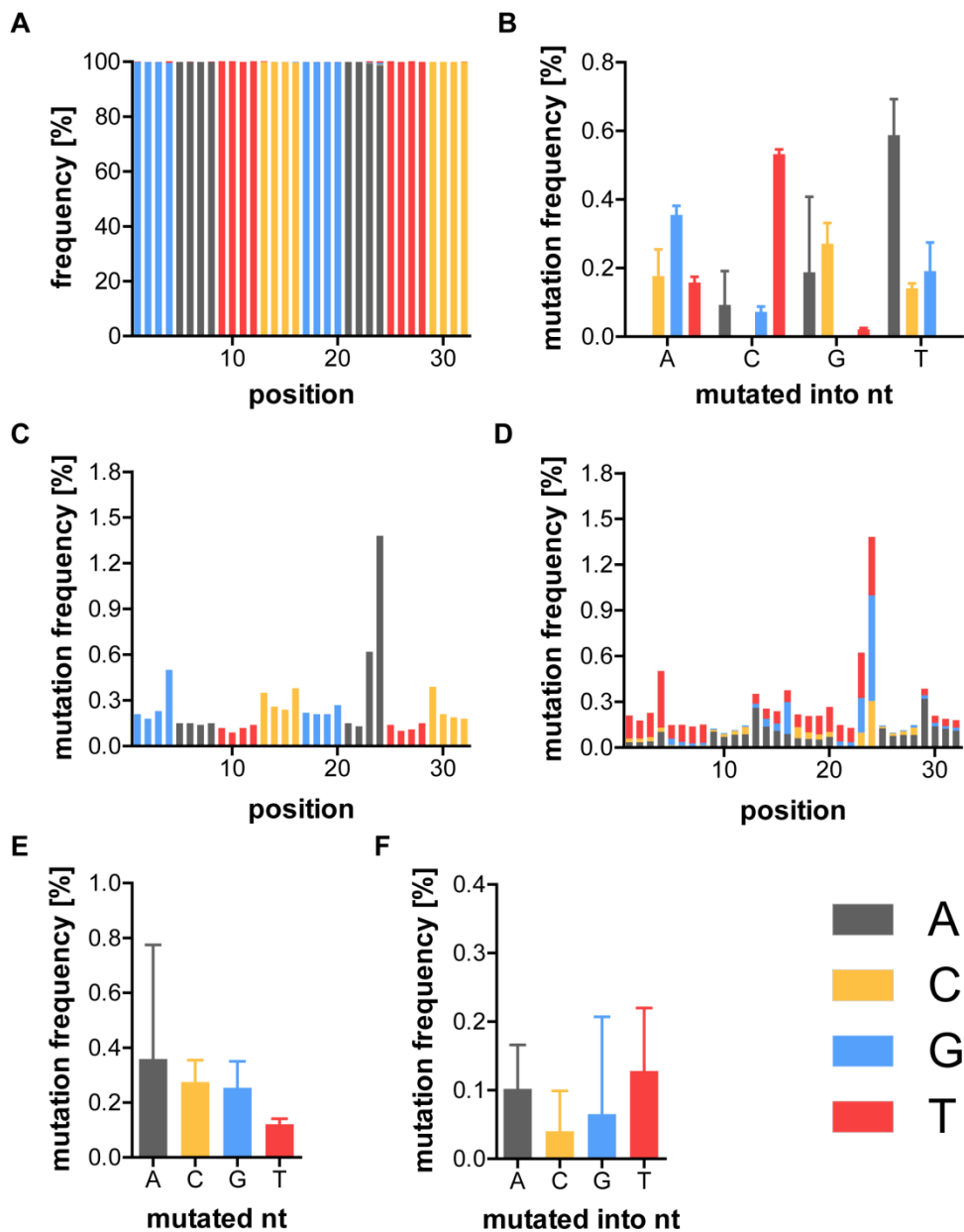


Figure S22: Mutation analysis of FT2_G4A4T4C4_II after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S23

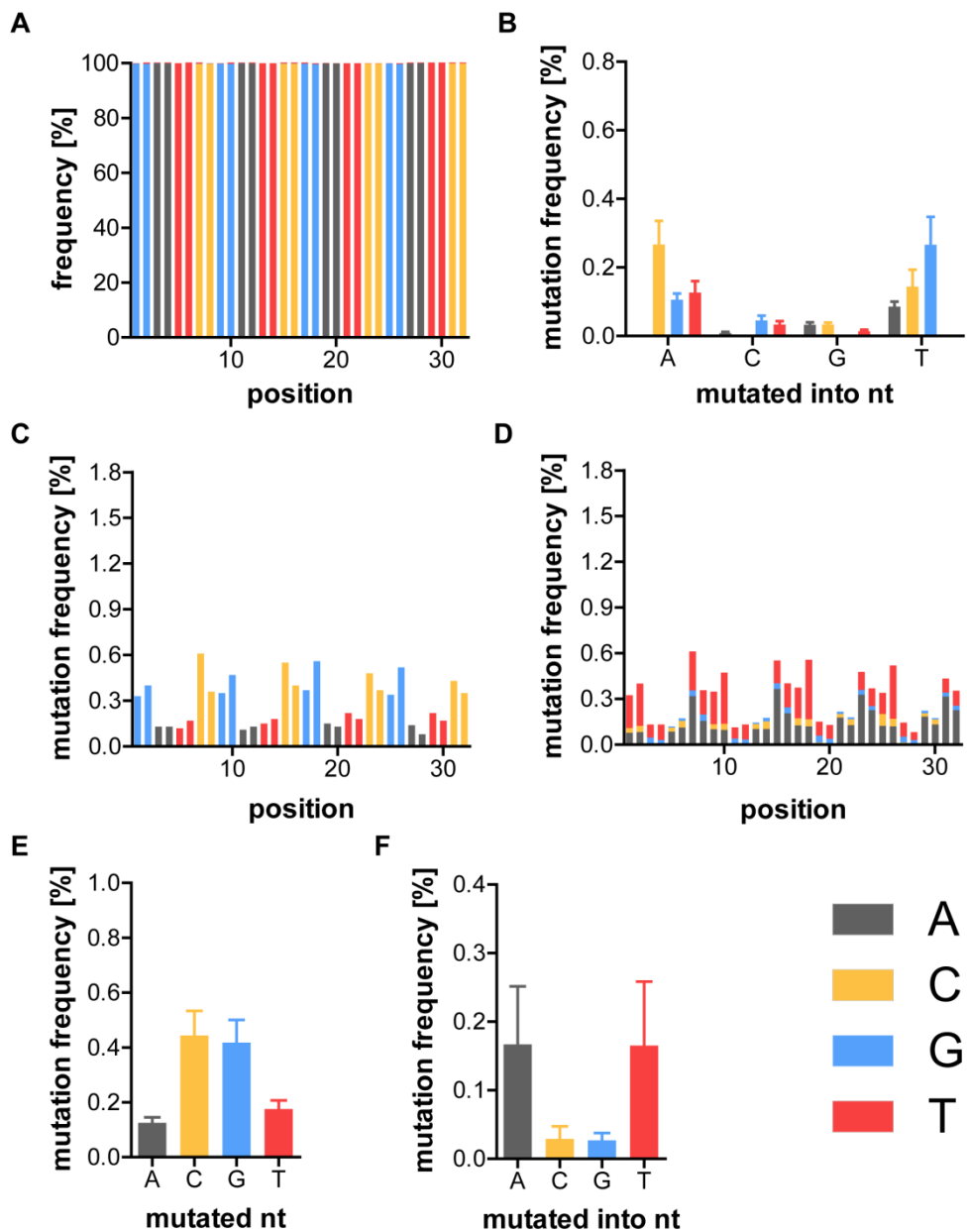


Figure S23: Mutation analysis of FT2_G2A2T2C2 after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S24

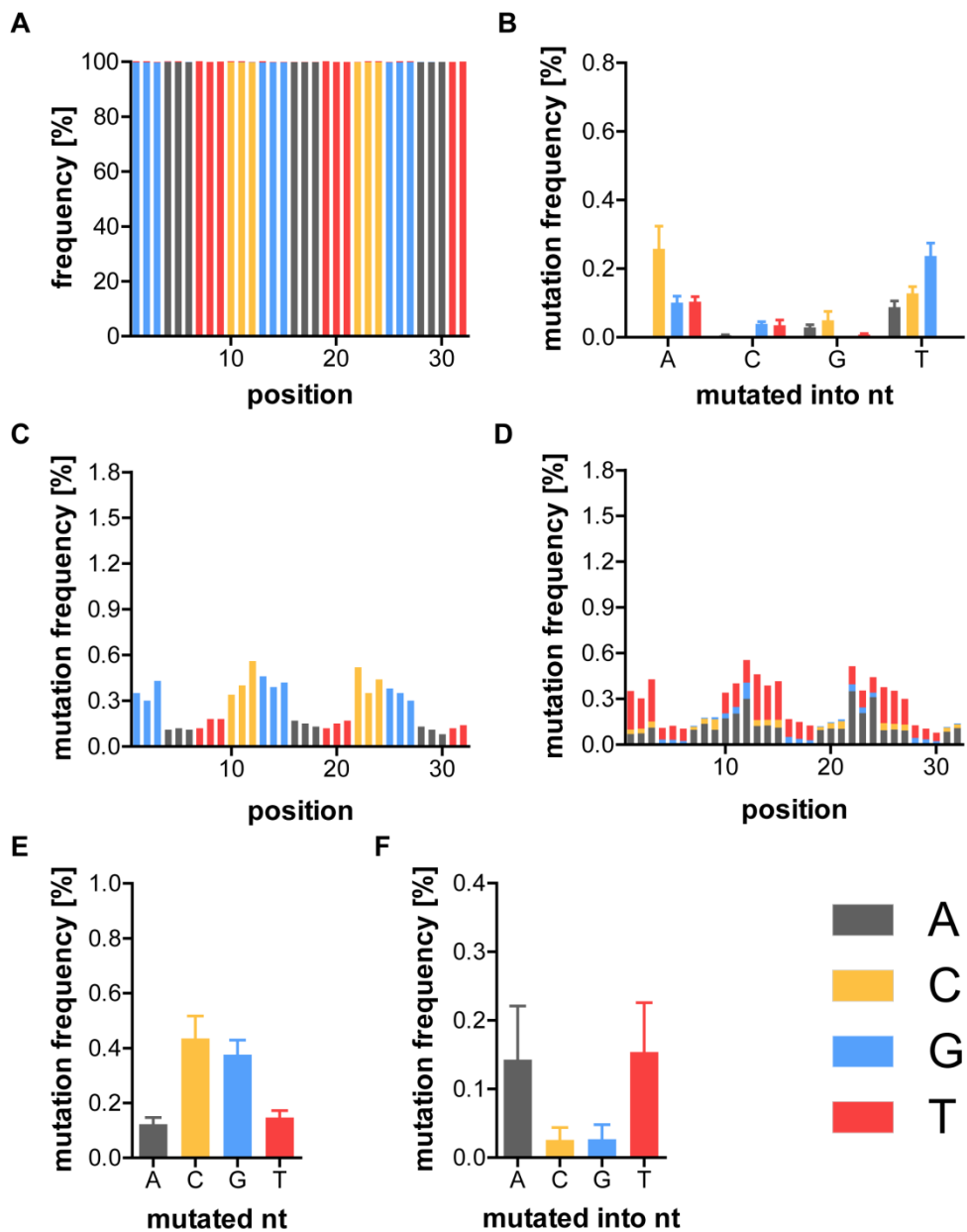


Figure S24: Mutation analysis of FT2_G3A3T3C3 after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S25

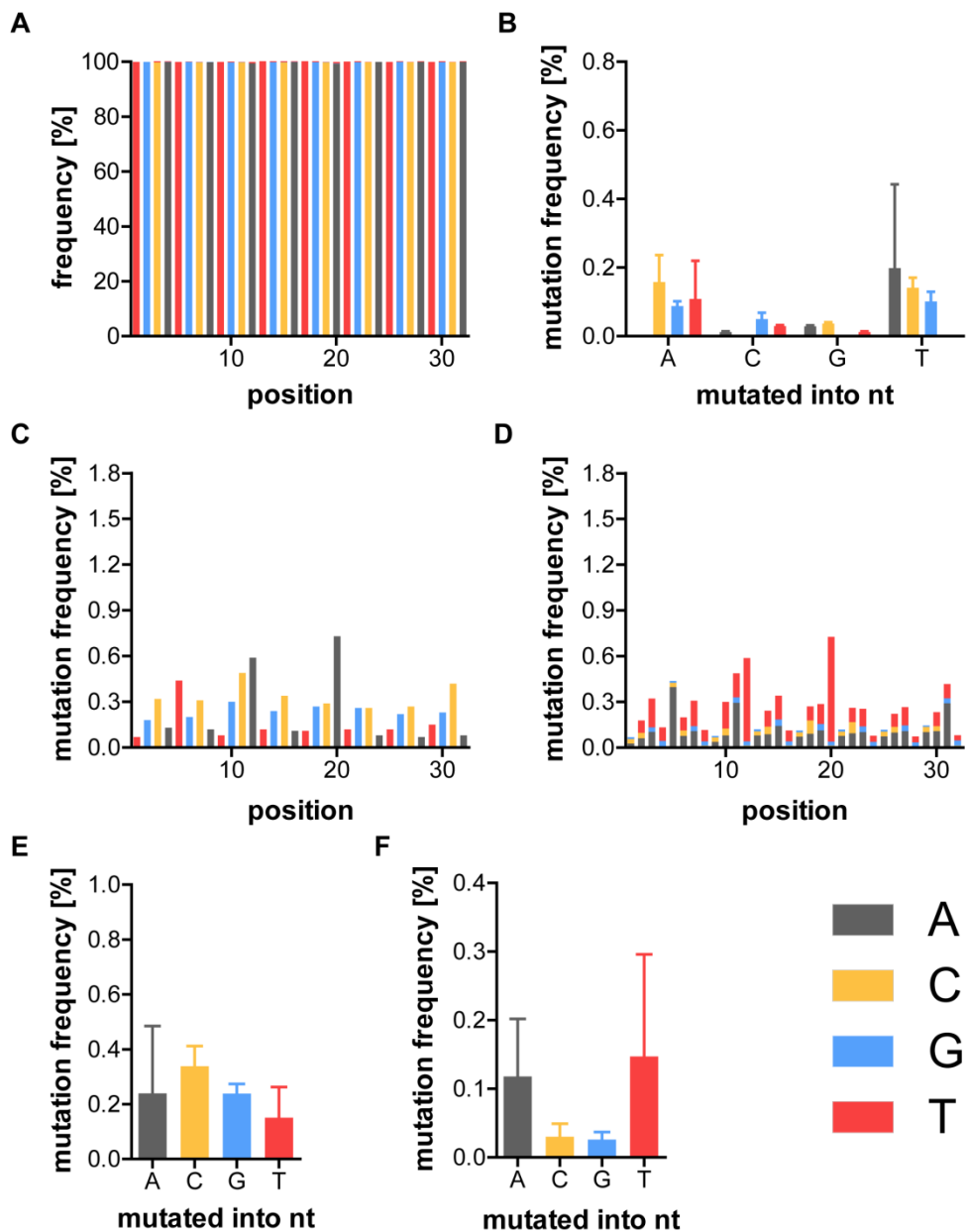


Figure S25: Mutation analysis of FT2_TGCA after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S26

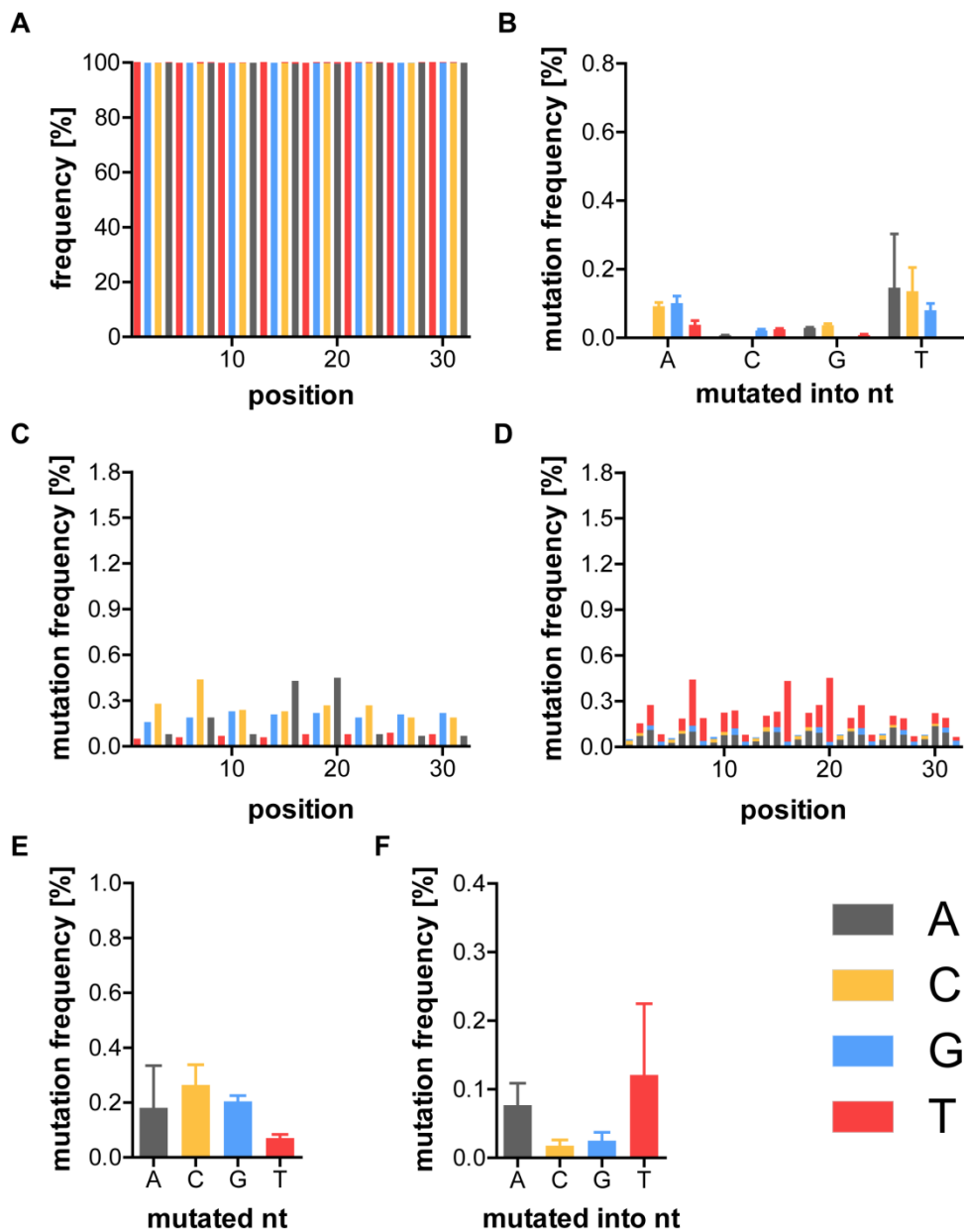


Figure S26: Mutation analysis of D3-TGCA after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S27

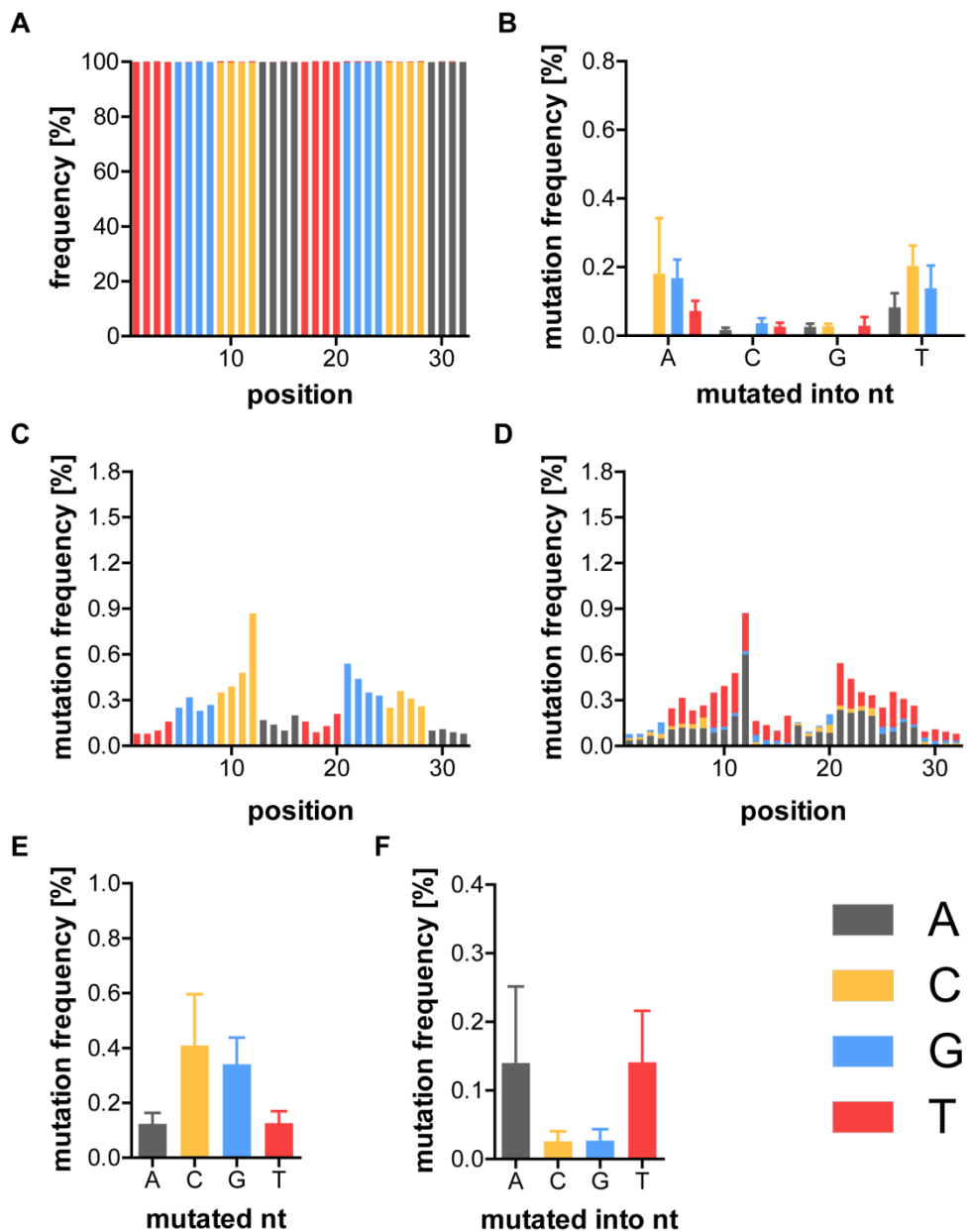


Figure S27: Mutation analysis of FT2-T4G4C4A4 after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

Figure S28

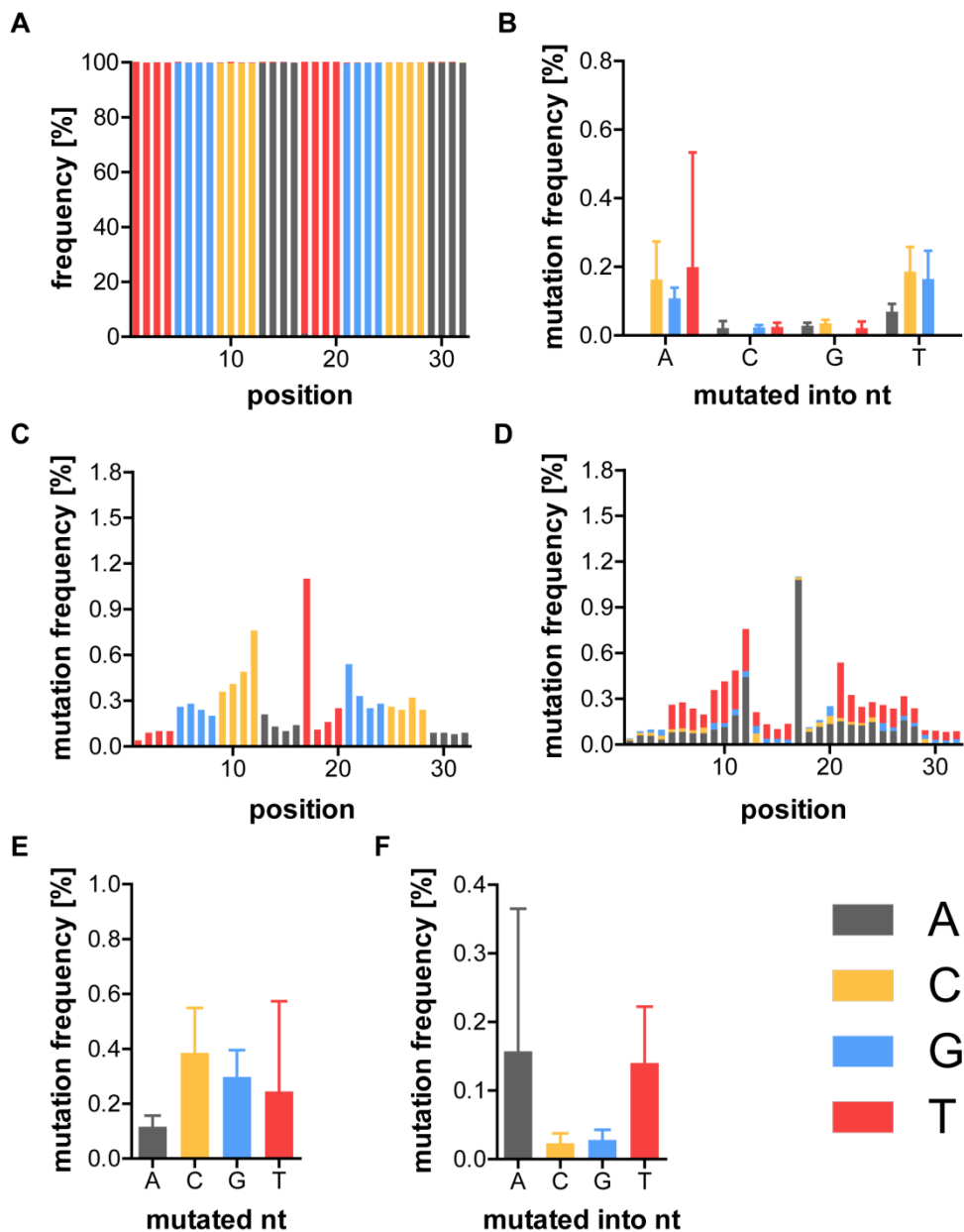


Figure S28: Mutation analysis of D3-T4G4C4A4 after omission of shortened sequences

(A) Frequency of the different nucleotides at all positions of the random region. (B) Frequencies with which one nucleotide (denoted by colour) mutates into another nucleotide (denoted on the x-axis). Depicted is the mean and SD. Mutation frequency at all positions of the random region of (C) the original nucleotide and (D) the nucleotide that has been mutated into. Mutation frequencies of the original nucleotides (E) and the nucleotides that has been mutated into (F). Depicted is the mean and SD.

