Figure S1  Timing of inter-genomic conversion. The star marks AD allotetraploidization occurring ~1-2 million years ago. (A) The numbers (proportion) of At, Dt neutral mutations (in intergenic region) and At to Dt conversions in both HC and LC are shown in each evolutionary time scale. (B) The numbers (proportion) of At, Dt neutral mutations (in intergenic region) and At to Dt conversions in HC are shown in each evolutionary time scale.
Figure S2  Frequency distribution of the length of conversion tracks measured by number of continuous converted alleles. Red line is expected distribution (exponential) of the length of continuous random mutations. Green line is the distribution of total number of conversions in the genome. Blue line is the distribution of conversions in HC.
Chr 3

Chr 5

H. Guo et al.
**Figure S3** Genome distribution of biased conversion, heterozygosity and GC-content. See Fig. 2 for detailed legend.
Figure S4  Number of each transposon types in HC and LC. Transposon types are: RLX, LTR-retrotransposon; RLG, Gypsy; TXX, transposon (either transposon or retrotransposon); RLC, Copia; RXX, retrotransposons; DTM, mutator; DTC, CACTA; DXX, DNA transposons; RIX, LINE; Non, others; DTA, hAT; DTT, Tc1-Mariner; DHH, Helitron.
Figure S5  Phylogenetic distribution of the ratio of non-synonymous to synonymous conversions. Barplot shows logarithm of the ratio for each cell in the above table.
Figure S6  Identification of converted alleles. A-genome and D-genome are diploid progenitor genomes. At- and Dt-subgenome are tetraploid genomes with At-genome derived from A-genome and Dt-genome derived from D-genome. Solid line represents cotton reference genome and broken line indicates re-sequenced genomes. Sites with number “0” represents no allele changes in the tetraploid genomes, likewise, “1”: At to Dt conversion; “2”: Dt to At conversion; “3”: Dt mutation; “4”: At mutation.
A

Unchanged  At to Dt  Dt to At  Dt to novel allele  A to novel allele  A2/other  both/other  Others

Read coverage

B

A2D  D2A  no change  D2other  A2other  both/other  others

Original  Deletions filtered

C

H. Guo et al.

13 SI
Figure S7  (A) Average mapped read coverage for different mutation categories in four cotton species. (B-E) Effects of deletion. Blue bar shows number of sites in each mutation category. Red bar shows number of sites after removal of the ones with reads coverage less than or equal to half of the average read coverage of each chromosome. (B) Acala Maxxa, (C) GA120R183, (D) Yucatanense, (E) G. mustelinum
Figure S8 Allelic changes in polyploidy cotton using relaxed read editing distance (0.8). (A) Allelic changes with all other parameters unchanged. (B) Allelic changes with reduced frequency threshold to call heterozygous genotype.
Figure S9  Genome distribution of conversion bias using relaxed read editing distance (0.8).
Figure S10  Correlation between number of allele changes and nucleotide divergence between A and D genomes. Genome is divided into non-overlap 10Kb bins. For each bins, the number of At to Dt converted alleles and the nucleotide divergence between the A and D genomes are calculated. (A) The number of At to Dt conversions and nucleotide divergence between A and D genomes. (B) The number of Dt to At conversions and nucleotide divergence between A and D genomes. (C) The number of unchanged alleles and nucleotide divergence between A and D genomes.