**Figure S1.** Illustration of compound heterozygosity. (A) The state of possessing two identical mutations in a particular gene (homozygosity) may lead to a recessive phenotype. (B) The presence of two different mutations at the same gene locus, both inside coding regions, one on each chromosome (compound heterozygote), may lead to the same recessive phenotype. (C) An example of generalized CH, where the second mutation occurs inside a non-coding region.

****

**Algorithm S1. G**enome shifting algorithm.

|  |
| --- |
| Input: bed\_file\_in (input bed file), k (window size)Output: bed\_file\_out\_1, bed\_file\_out\_2, ..., bed\_file\_out\_k |
|  01: for each i in {sequence from 1 to k}: 02: stream1 <- open a reading stream on bed\_file\_in 03: stream2 <- open a reading stream on bed\_file\_in 04: stream1 skip the first 3 bytes 05: stream2 skip the first (3 + i \* (number of bytes per SNP)) bytes 06: stream3 <- open a writing stream for bed\_file\_out\_i 07: for each j in {sequence from 0 to (total number of SNPs - 1)}: 08: buffer\_collapsed <- initialize a new byte array 09: buffer1 <- read genotypes of one SNP from stream1 10: if j >= (number\_of\_SNPs - i): 11: fill buffer\_collapsed with NA 12: else: 13: buffer2 <- read genotype of one SNP from stream2 14: for each genotypes\_byte1 in buffer1, each genotypes\_byte2 in buffer2: 15: coerce genotypes\_byte1 and genotypes\_byte2 into integers 16: collapsed\_byte <- look up collapsing\_byte\_array at coordinates (genotypes\_byte1, genotypes\_byte2) 17: add collapsed\_byte to buffer\_collapsed 18: end for 19: end if 20: write buffer\_collapsed to stream3 21: end for 22: close all streams 23: end for |