**Figure S1.** Frequency distribution of assembly fragment lengths follow a log-normal distribution. (a) Distribution of genome assembly fragment lengths of Arabidopsis and fitted exponential, log-normal and normal distributions. (b, c, d) Q-Q plots fitting observed and estimated lengths for (b) exponential, (c) log-normal and (d) normal distribution.
Figure S2. Outline of CHERIPIC method

1. Sequence bulks
2. Assembly from bulk or previous assembly
3. Variant calls using assembly as reference
4. Calc. HMES and arrange fragments with high HMES at the center
5. Select candidate variants
Figure S3. Variants selected by CHERIPIC for Arabidopsis sup2 data were presented on all five chromosomes. (a) All variants were depicted; (b) Top five percentile HMES variants were depicted.
Figure S4. All variants selected by CHERIPIC for maize gl3 data were presented on all ten chromosomes.
**Figure S5.** Variants selected by CHERIPIC for barley *mnd* data were presented on all seven chromosomes. (a) All variants were depicted; (b) Top five percentile *HMES* variants were depicted.