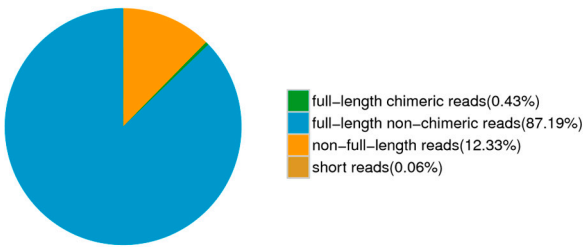
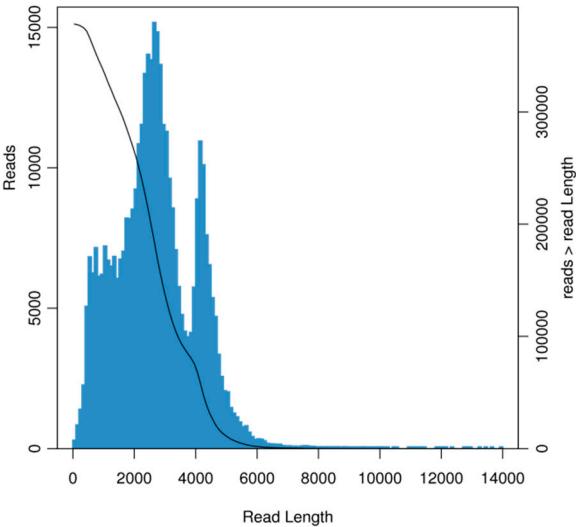


Supplementary Figures

A

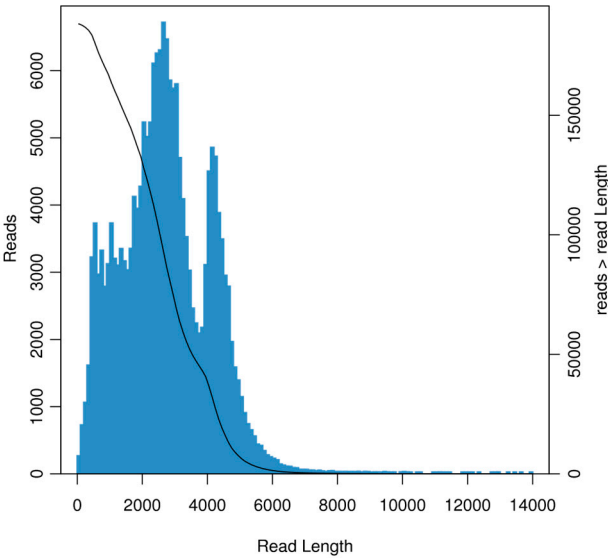


B

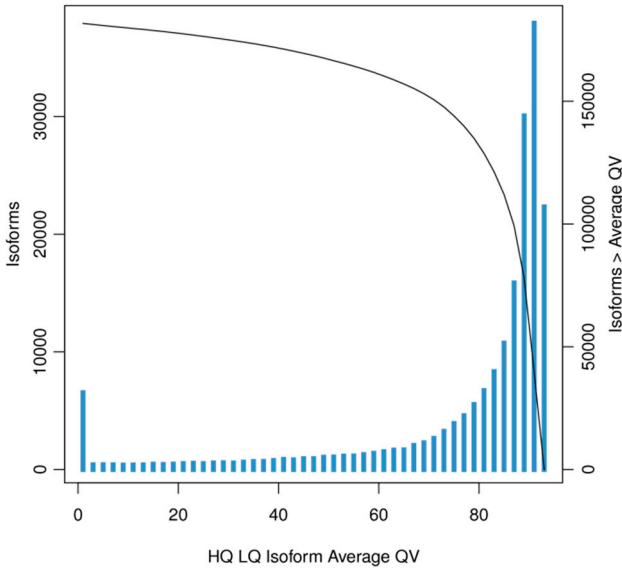


**Figure S 1.** Classification and statistics of CCS reads (A) and length distribution of full-length non-chimeric reads (B) in *C. album* from SMRT data.

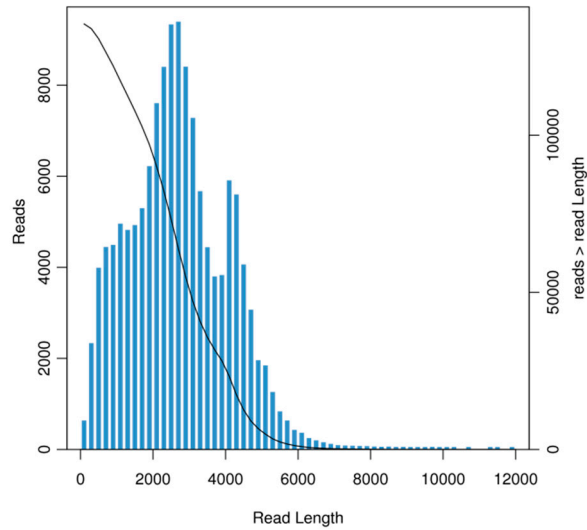
A



B

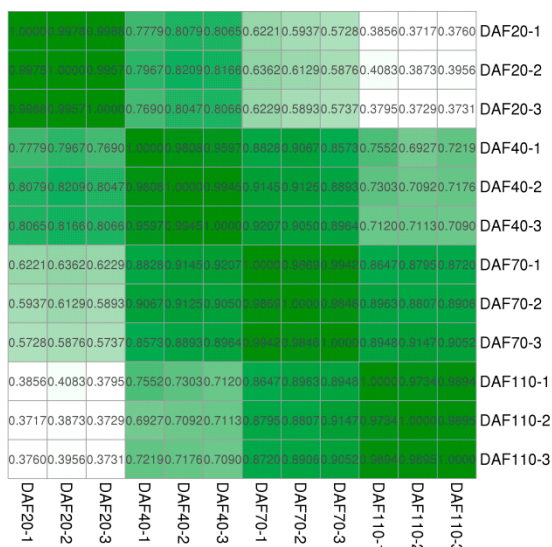


**Figure S 2.** Length (A) and quality (B) distribution of consensus isoforms in *C. album* from SMRT data.

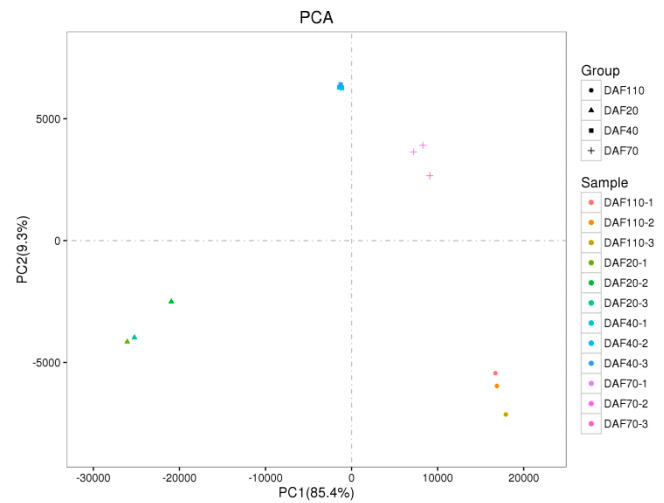


**Figure S 3.** Length distribution of isoform sequences in *C. album* from SMRT data.

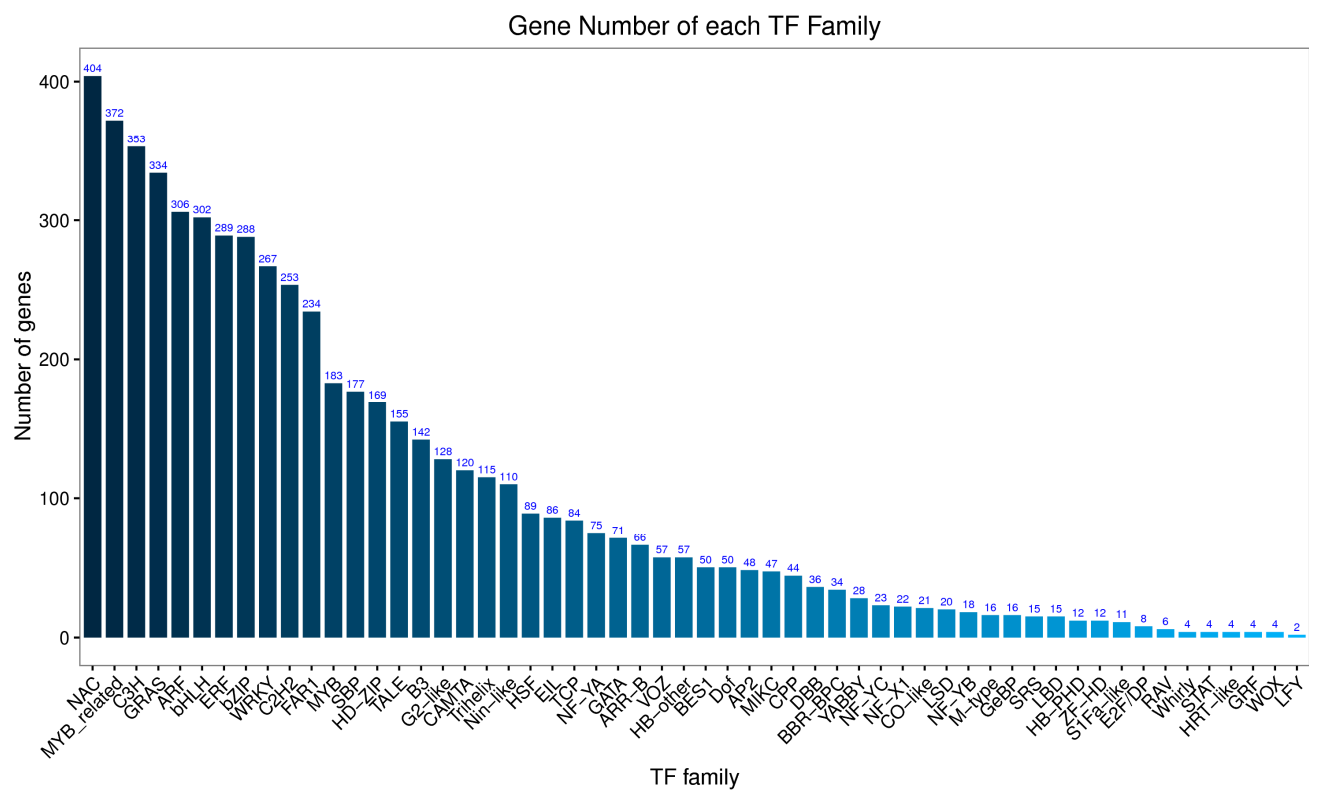
A



B



**Figure S 4.** Correlation analysis (A) and principal component analysis (PCA) (B) among samples. (A) Heat map of the coefficient correlation between two samples. The closer the correlation coefficient gets to 1 indicated the better the repeatability between two parallel experiments. (B) We use R package models to carry out PCA and draw two-dimensional coordinates according to the numerical values of each sample in the first principal component (PC1) and the second principal component (PC2). The value in parentheses of the axis label represents the percentage of variance explained by the principal component of the population.



**Figure S 5.** TF family distribution in *C. album* according to SMRT data.