Supplementary material

Table 1: Discovered SNP results of Zingiber officinale.

Tissue Name	Sequences analysed	No of contigs	Consensus size	SNPs detected	Transitions (Ts)	Transversions (Tv)	Indels	Ts / Tv	SNP frequency
Leaves	11307	2660	2555383	16046	6805	7258	1983	0.94	0.63/100 bp
Rhizomes	10123	2921	2668324	28809	13433	12249	3127	1.10	1.08/100 bp
Root	10475	2972	2332434	19171	6845	10402	1924	0.66	0.82/100 bp
Total	31905	8553	7556141	64026	27083	29909	7034	0.90	0.84/100 bp

A total of 31905 consensus EST sequence are used to predict the SNP site from Ginger species, which made 8553 cluster groups and found 64026 SNP/indel sites.

Table 2: Summary of SNPs and indels detected in the Ginger EST libraries

Results	Leaves	Rhizomes	Root	Total
Total No. of ESTs	13274	12763	12092	38129
Total sequences analysed	11307	10123	10475	31905
No. of contigs	2660	2921	2972	8553
Total SNPs detected	16046	28809	19171	64026
Total consensus size (bp)	2555383	2668324	2332434	7556141
Frequency of SNP per 100bp	0.63	1.08	0.82	0.84
Transitions				
G/A	3595	6901	3278	13774
C/T	3210	6532	3567	13309
Tranversions				
A/C	1822	3200	2617	7639
A/T	1678	2550	2704	6932
G/T	1917	3065	2553	7535
G/C	1841	3434	2528	7803
Indel				
A	564	818	623	2005
G	504	772	547	1823
T	433	743	361	1537
C	482	794	393	1669
Shannon Index	0.150	0.152	0.164	0.052

$$H' = \sum_{i=1}^{n} piLog_2Pi$$

Where n is the total number of SNP/indel states (10) pi= proportion of ESTs in the i^{th} type of SNP/indel state. The calculated value is divided by the $\log_2 10$ to get uniformity.