

Supplementary material

$$H' = \sum_{i=1}^n p_i \log_2 p_i \rightarrow (1)$$

Where n is the total number of SNP/indel states (10) pi= proportion of ESTs in the ith type of SNP/indel state. The calculated value is divided by the log₂10 to get uniformity.

Result	Abnormal apex	Normal apex	Female inflorescence	Male inflorescence	Immature zygotic embryo	Mesocarp Tissue	Total
Total No of ESTs	998	313	349	625	126	3039	5452
Total sequences analysed	242	32	38	103	40	121	576
No. of contigs	86	13	16	44	9	31	199
Total SNPs detected	309	62	103	239	7	597	1317
Total consensus size (bp)	35034	5296	5601	19790	4766	20534	91021
Frequency of SNP per 100 bp	0.88	1.17	1.84	1.21	0.15	2.91	1.36
Transitions							
C/T	111	11	25	90	3	149	389
G/A	74	12	32	48	0	135	301
Transversions							
A/T	24	5	7	19	0	38	93
C/G	40	12	13	33	2	61	161
G/T	20	2	15	13	0	50	100
A/C	27	12	9	30	1	57	136
Indels							
A	5	2	0	1	0	26	34
C	4	1	0	0	0	18	23
G	1	3	0	3	0	42	49
T	3	2	2	2	1	21	31
Shannon index	0.02	0.60	0.43	0.16	0.31	0.16	0.03

Table 1: Summary of SNPs and indels detected in the oil palm EST libraries