

Identification and Functional Characterization of Oxidosqualene Cyclases from Medicinal Plant *Hoodia gordonii*

Iffat Parveen ^{1,*}, Mei Wang ³, Joseph Lee ¹, Jianping Zhao ¹, Yingjie Zhu ², Amar G. Chittiboyina ¹, Ikhlas A. Khan ^{1,4} and Zhiqiang Pan ^{3,*}

Supplementary Information

Figure S1: Graphical representation of *Hoodia gordonii* HgOSCs conserved domains.

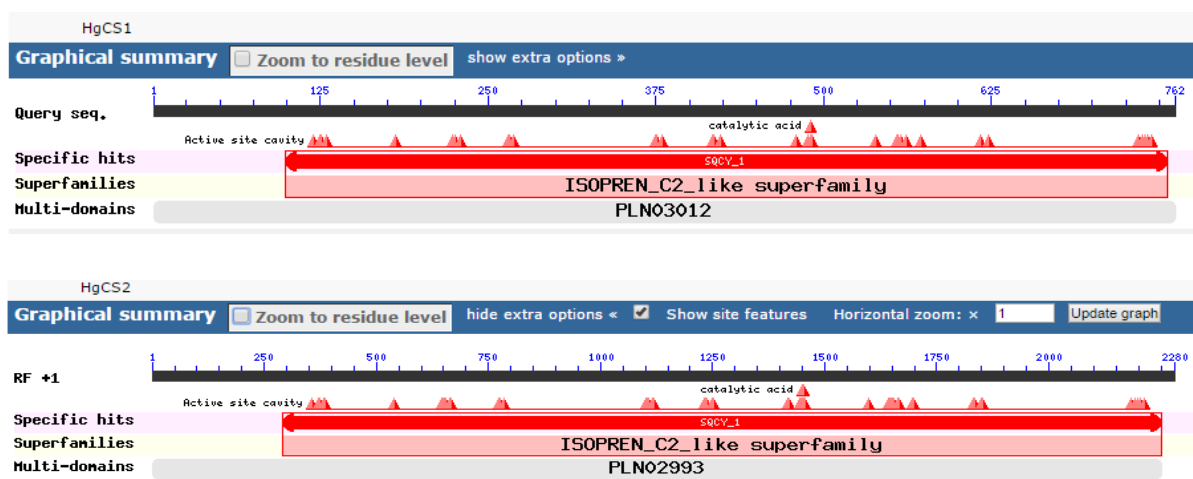


Figure S2. Amino acid alignments of HgOSC1 and HgOSC2 with oxidosqualene cyclases from plants. Four motifs rich motifs (Motif A-D), necessary for the binding of cationic cofactors are shown as boxes.

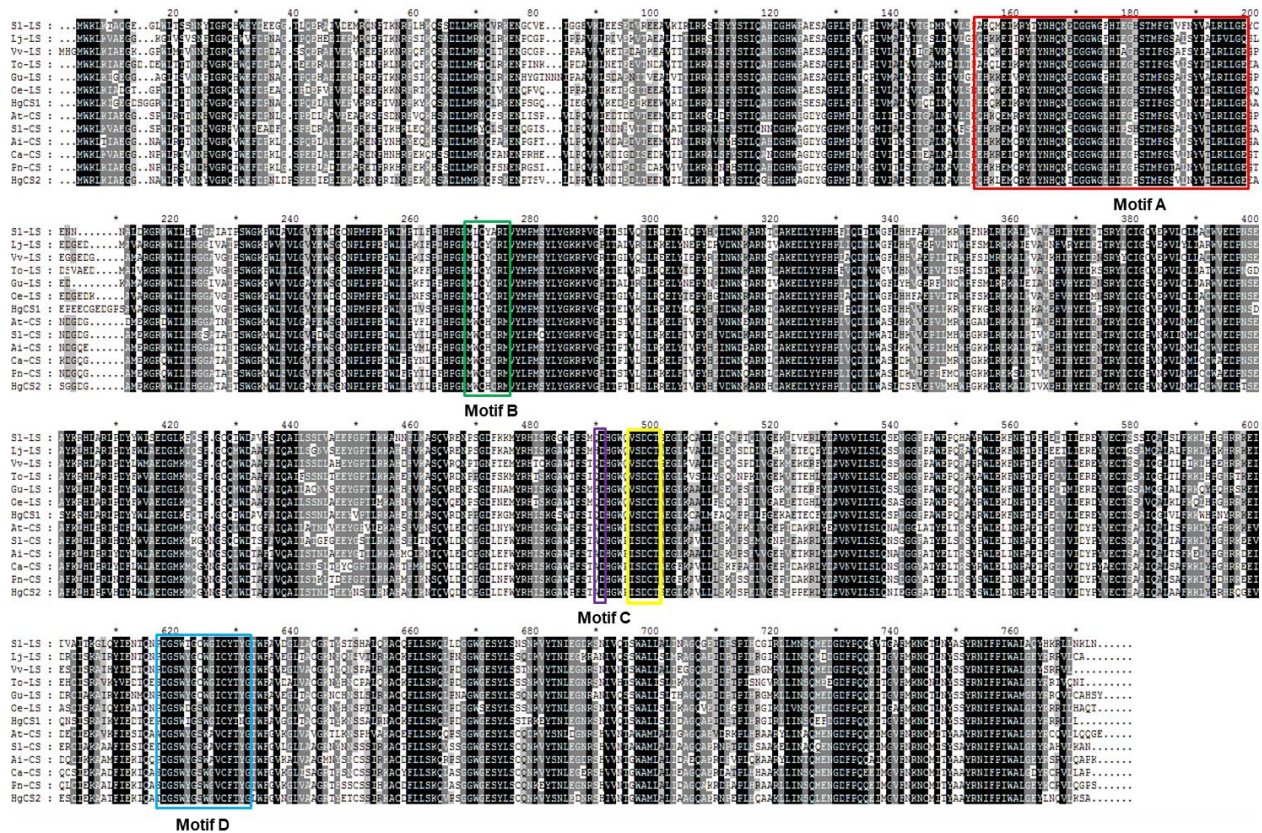


Figure S3: GC-MS mass spectrum of Lupeol

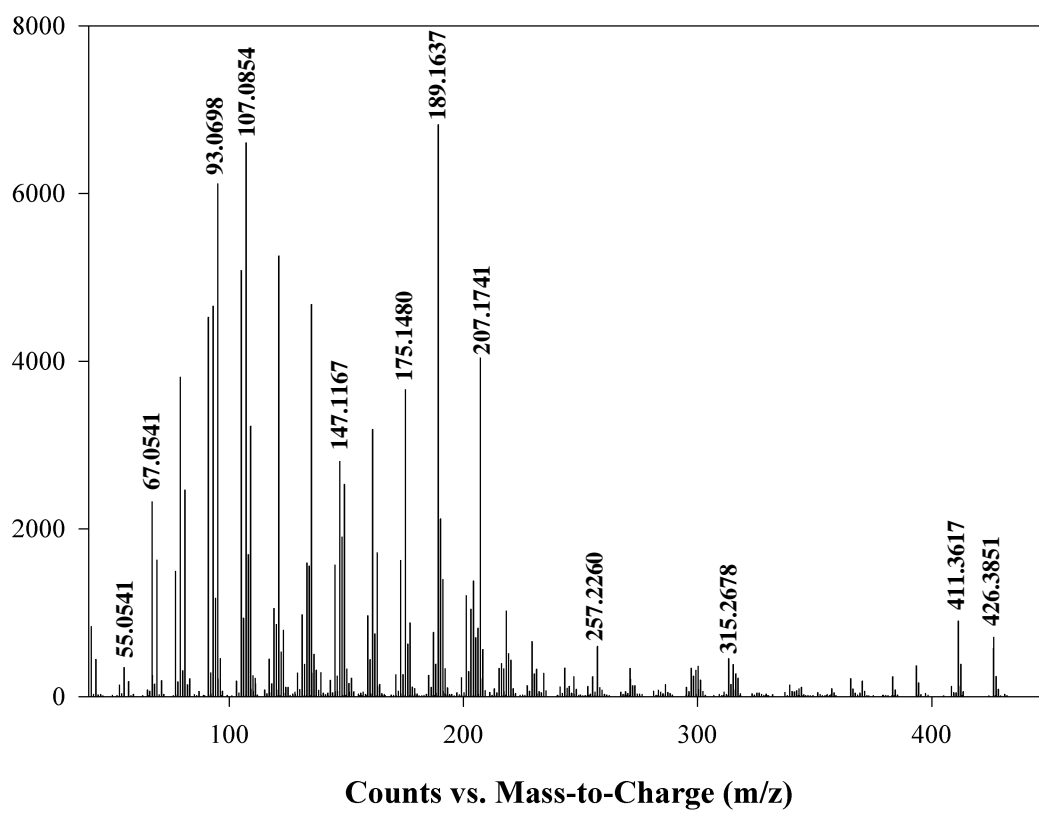


Figure S4. Thin-layer chromatography (TLC) and nuclear magnetic resonance (NMR) spectroscopy analyses. A, Compound extracts were separated on TLC (HG: hexane extract from yeast expressing HgOSC1; R: lupeol reference standard; V: hexane extract from yeast harboring vector as control). B and C, The ^1H (B) and ^{13}C (C) NMR spectra showing signals of the HgOSC1 product and the lupeol standard. The major signals of the ^1H and ^{13}C NMR spectra of the HgOSC1 product (lower panels of B and C) are comparable to those of the lupeol standard (upper panel of B and C).

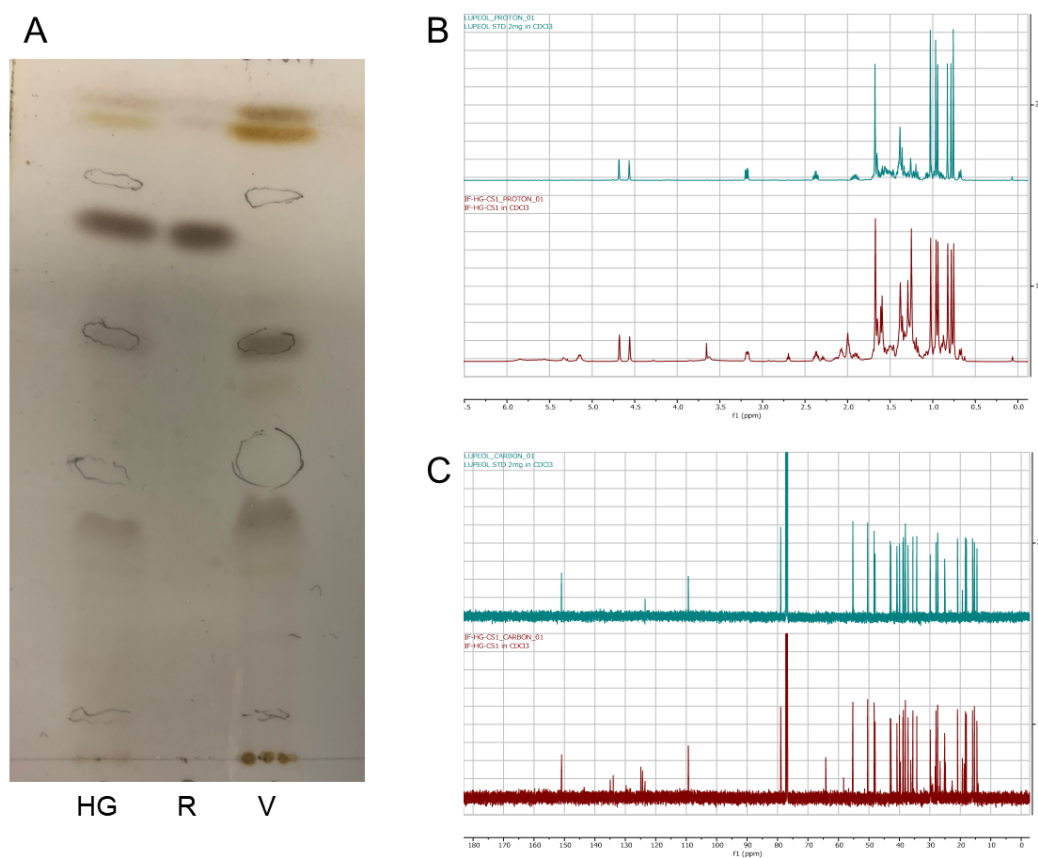
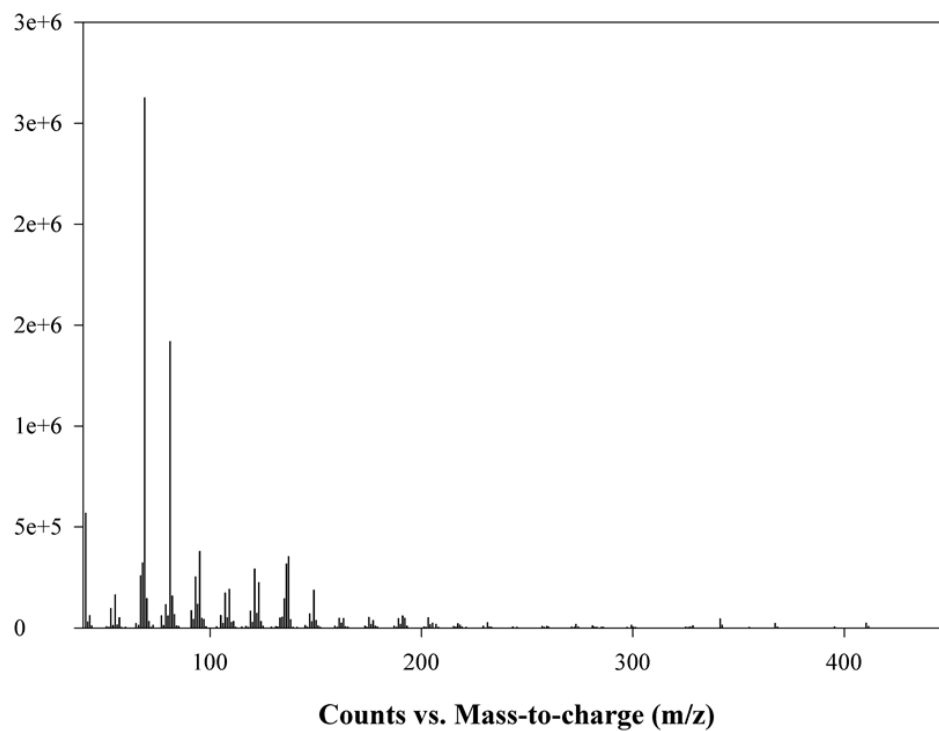


Figure S5. DNA sequence of *HgOSC2* synthesized for optimization of codon usage for expression in yeast.

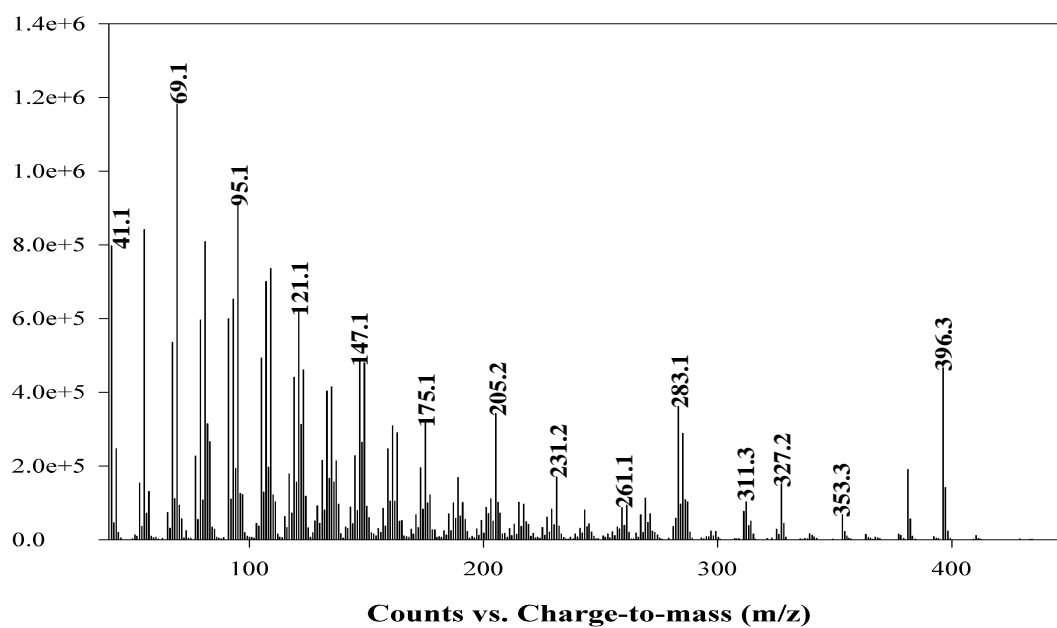
ATGTGGAGATTGAAAATTGCTGAAGGTGGTAACGCATGGTTGAGAAGCTGTTAACTATGTTGGTAGACAATTCTGGGAATTCGATCCAAATT
TGGATCCATCTCCAGAAGAAATCGAAGAAATCGAAAAAGCTAGAGAAAACCTTCAGAATTAATAGATTCGAAAAGAAACATTCAGCAGATTGGA
TGATGAGAATCCAATTTTCTAAGGAAAACCAACTTCAGTTTTGTTGCCAAGAGTTGAAGTTAACGATATCGAAGATTGACAGAAGAAAATGT
TACTTTGACATTGAAGAGAGCTATTAATTTCTATTCTACTTTACAAGGTCATGATGGTCATTGGGCTGGTGACTACGGTGGTCCAATGTTTTTGT
ACCAGGTTTAGTTATTGCTTTGCTATTACAGGTGCTTTGAACGCAGTTTGTCAAAGCAACATAAGTTGGAAATGTGTAGATATTGTACAACC
ATCAAAACATCGATGGTGGTGGGGTTTGCATATTGAAGGTCCATCTACTATGTTTCGGTTCAGTTTAACTACGTTACATTGAGATTGTTAGGTG
AAGAAGCTTCTGGTGGTGACGGTGCAATGGAAAAAGGTAGAAAATGGATTTTGGATCATGGTGGTGCTACTGCAATTACATCTGGGGTAAAA
TGTGGTTATCAGTTTTGGGTGCTTATGAATGGTCTGGTAATAATCCATTACCACCAGAAAATTTGGTTGTTACCATACTTGTACCATTTCATCCAGG
TAGAATGTGGTGCATTGTAGAATGGTTTATTGCCAATGTCTTATTGTACGGTAAAAGATTTCGTTGGTCCAATCACTCCAACAATCTTGTCATTG
AGAAGAGAATTGTACACAGTTCATACCATGAAGTTGATTGGAATAAGGCTAGAAACCATTGTGCAAAGGAAGATTGTACTACCCACATCCAT
TGATCCAAGATATCTTGTGGGCTTCTTTGGATTCAATTGTTGAACAGTTATGATGCATTGGCCAGGTAAAAAGTTGAGAGAAAAGGCATTGAG
AACTGTTATGGAACATATCCATTACGAAGATGAAAACACAAGATACATCTGTATCGGTCCAGTTAATAAGGTTTTGAACATGTTGTGTTGTTGGG
TTGAAGATCCAACCTCTGAAGCTTTAAATTGCATATCCCAAGAGTTCATGATTATTTGTGGTTGGCAGAAGATGGTATGAAAATGCAAGGTTAC
AATGGTTTCACAATTATGGGATACAGCTTTCGCAGTTCAAGCTATCATCTCTACTAATTTGACAGAAGAATATAATTCAACTTTGAGAAACGCTCAT
GCATACATTAAGAATACACAAGTTCAAGATGATTGTCTGGTGACTTGGATTCTGGTACAGACATATCTCTAAGGGTGCTTGGCCATTTTCAACT
GCAGATCATGGTTGGCCAATCTCTGATTGTACATCAGAAGGTTGAAGGCTGTTTTGTTGTTGCTAAGTTGTACCAGAATTGGTTGGTGAATC
TTTGCCACCAGAAAAGTTGTACGATGCTGTTAAGCTTATCTTATCATTTGCAAAAATATTGAAGGTGGTTTCGCAACTTACGAATTGACAAGATCTTA
CTCATGGTTAGAATTGATTAATCCAGCTGAAAACCTTTCGGTGACATCGTTATCGATTACCCATACGTTGAATGTACATCTGCTGCAATTCAAGCATTG
GCTGCTTTTAAAGAAATTGTACCAAGACATAGACAAGGTGAAGTTGAATCATGTATCGAAAAGGCTGCAACTTTTATTGAAAAGATTCAAGCTA
CAGATGGTTCTTGGTATGGTTTCATGGGGTGTGTTTTTACTTACGGTTTGTTGGTTTGGTGTAAATGGTTTAGTTGCTGCAGGTAGAACTTACGAA
ACATGTTCTTCAATCAGAAAGGCATGTGATTCTTGTGTTGCTAAACAAGTTCCATCAGGTGGTTGGGGTGAATCTTATTTGTCATGTCAAAATAA
GGTTTACTCTAATTGGAAGATAACAGATCAGATATCGTTAATACTGGTTGGGCTATGTTAGCATTGATTGCTGCAGGTCAAGCTGAAAGAAATC
CAGAACCATTGCATCAAGCTGCAAAGTTGTTGATTAATTCTCAATTAGAAAATGGTGACTTCCACAACAAGAAATTATGGGTGTTTTTAATAAG
AACTGTATGATCATACGCTGCATACAGAAACATCTTCCAATCTGGGCTTTGGGTGAATATTGAATCAAGTTTTAAATCTGCATAA

Figure S6: The corresponding mass spectra of (a) Squalene; (b) Cycloartenol; (c) Ergosterol and (d) Cycloartenol acetate from GC-MS analysis.

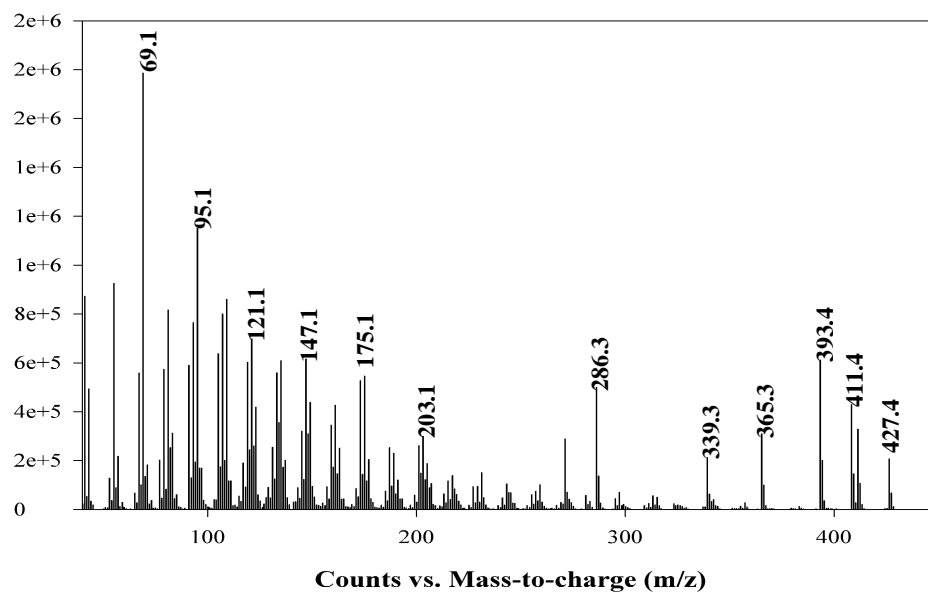
(a)



(b)



(c)



(d)

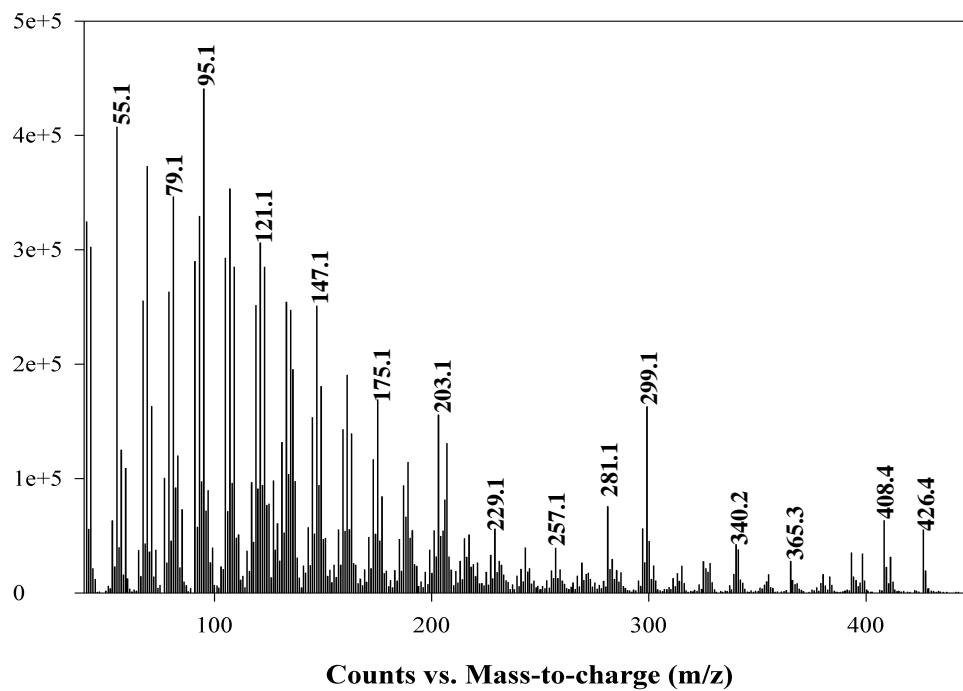


Table S1: Primers used for 5'- and 3' rapid amplification of cDNA ends (RACE).

Primer Code	Locus	Sequence (5'-3')
P711	hga_locus_5882_iso_1_len_575 (F)	TCCTGATGATGGTGATGGTGCTGTT
P712	hga_locus_5882_iso_1_len_575 (Rev)	AGGACATTGGCATATAGGTGGTTCGAC
P713	hga_locus_5300_iso_4_len_427 F	TCTACCTCCTCTGGTGATGGCGTTGT
P714	hga_locus_5300_iso_4_len_427 R	CCACATTCTTCAGGCTCTGCTGCTTCT
P715	hga_locus_3813_iso_1_len_390 F	TGTTGGACGCATAACTGAGCTTGT
P716	hga_locus_3813_iso_1_len_390 R	AGCATATCTTGAGGTAAAGGATGAGGGT
P717	hga_locus_11949_iso_1_len_446 F	AAGCCCTTCGTACTGTGATGGAACA
P718	hga_locus_11949_iso_1_len_446 R	CTCAGTCAGGTTTGTGCGAAATAATCGCC
P719	hga_locus_2117_iso_1_len_589 F	CTGAAGATAGCGGAGGGAGGGAA
P720	hga_locus_2117_iso_1_len_589 R	TCGTGGTAACAATACACTGGTTGGG
P721	hga_locus_5625_iso_3_len_1014 F	AGGCATCGTCAAGGAGAGGTGGAAAG
P722	hga_locus_5625_iso_3_len_1014 R	GTCCAGCAGCAATGAGAGCCAGCA
P723	hga_locus_5748_iso_1_len_268 F	TTGAAGATACACAGGAACCAGATGGC
P724	hga_locus_5748_iso_1_len_268 R	TACCACAGGCAGTCAATCCTCCTAC