

Table Appendix S5: Illumina paired-end read statistics, and alignment summaries for sequence data generated from the high-throughput approach, sequences retrieved from GenBank, and the combined total dataset for each gene region.

Gene Region	Primer Pair Name	Number of Raw Reads	Number of Reads After Screen	% Raw Reads Retained	Number of Unique Species Accessions in High-throughput Dataset	Alignment Length (bp) High-throughput (pre-clean)	Number of Unique Species Accessions from GenBank	Alignment Length (bp) GenBank Dataset (pre-clean)	Number of Accessions Total Data Dataset	Alignment Length (bp) Total Data Dataset (pre-clean)	Alignment Length (bp) Total Data Dataset (cleaned)	% Gaps in Total Data Alignment (cleaned)
ITS	ITS	843710	120753	14.31	100	814	103	778	122	1040	530	17.15
	<i>atpB</i>				107	870	9	1491	108	1491	588	27.43
matK	<i>atpB1</i>	38048	1787	4.70	88	1041	71	2095	111	2476	489	4.7
	<i>atpB2</i>	4178	4178	100.00								
ndhF	<i>matK1</i>	213242	53767	25.21	142	1827	23	2262	130	2341	489	0.13
	<i>matK2</i>	234128	16834	7.19								
rbcL	<i>ndhF1</i>	226816	40966	18.06	62	534	47	1419	88	1419	534	2.67
	<i>ndhF2</i>	186198	5674	3.05								
	<i>ndhF3</i>	219686	3291	1.50								
trnTLLF	<i>rbcL</i>	172606	8695	5.04	81	1821	84	1173	106	2616	560	21.44
	<i>trn_ab</i>	234262	6162	2.63								
Concatenated	<i>trn_cd</i>	641278	35825	5.59	145		121		152		3193	37.11