

Supplemental Table 2: High throughput sequencing summary

Snake #	Tissue sequenced	Barcode	Read pairs (a)	Unique read pairs (b)	Read pairs post filtering (c)	Fraction post filtering	Initial nidovirus BLASTx hit pairs (d)	Actual mapping nidovirus read pairs (e)	Average pairwise % identity to assembled genome (f)
1	lung	TACGCTT	869821	91038	42403	0.05	0	21	99.9%
2	lung	TCCGTAT	5769942	396912	256392	0.04	92	12194	99.6%
5	<i>Not sequenced</i>								
6	lung/liver/spleen	GTGTCCT	2737996	502585	125051	0.05	0	49	99.0%
7	liver/lung/spleen	TCGTTGG	197203	13659	2210	0.01	6	597	99.4%
8	lung/brain	GAAGGAG	2685821	300998	30592	0.01	3	152	95.7%
8	liver/kidney/spleen/pancreas	TACAACG	3942193	135701	13498	0.00	0	0	
9	lung/trachea/liver	AAGCATC	4128179	198614	42020	0.01	1	148	95.2%
10	lung	TGAGATC	5632581	432622	59348	0.01	0	137	96.9%
11	lung	ACACAAG	151283	27544	12220	0.08	0	126	97.0%

Notes

- (a) Number of initial read pairs
- (b) Number of read pairs after removing low quality and low complexity sequences, and collapsing non-unique sequences
- (c) Number read pairs remaining after filtering host-derived sequence, as described in text
- (d) Number of pairs where the best hit from a BLASTx search of nr database for both reads was to a nidovirus sequence w/ E-value < 1e-3
- (e) Number of read pairs from fully-filtered datasets mapping to ball python nidovirus assembly, determined using Bowtie2 alignment tool
- (f) Average % pairwise nucleotide identity of mapping read pairs to ball python nidovirus assembly (assembled from snake #2 reads)