



Jonah Ventures NGS Report Files Explained

Summary data

Exact Sequence Variant unique identifier

Consensus Taxonomy
Taxonomic ranking from the library matched to each sequence.

The % of base pairs in the sequence that match to the species in the library.

The detected sequence.

Sample identifiers
Numbers after the decimal point represent lab replicates of the same sample.

ESV	Family	Genus	Species	% match	# Spp	Sequence	S10001.1	S10002.1	S10003.1	S10003.2
ESV_000031	Catostomidae	Catostomus	Catostomus commersonii	100	1	AAAAAG...	3066	1340	5780	3462
ESV_009176	Moronidae	Morone	Morone americana	100	1	AAAAAG...	0	2582	249	0
ESV_009177	Ictaluridae	Ictalurus	Ictalurus furcatus	100	1	GAAAAG...	584	101	0	1039
ESV_007836	Clupeidae	Alosa	NA	100	2	GAAAAG...	1328	0	0	0
ESV_009179	Clupeidae	Dorosoma	Dorosoma cepedianum	99	1	AAAAAG...	902	0	0	278
ESV_009181	Catostomidae	Erimyzon	Erimyzon oblongus	100	1	ATAAAG...	341	124	282	381
ESV_009182	Moronidae	Morone	Morone saxatilis	100	1	AAAAAG...	126	0	0	994

"NA" at a taxonomic ranks means multiple species present in the sample differ in that taxonomic rank

The number in each cell is the absolute number of times a given sequence was read by the sequencer.

ESV detail

Exact Sequence Variant

Index identification number. Each sample has an individual index number.

Species Taxonomy

The % of base pairs in the sequence that match to the species. 100% indicates all base pairs matched

ESV	GB_ID	Phylum	Order	Family	Genus	Species	%match
ESV_000031	MG570424	Chordata	Cypriniformes	Catostomidae	Catostomus	Catostomus commersonii	100
ESV_000031	MG570423	Chordata	Cypriniformes	Catostomidae	Catostomus	Catostomus commersonii	100
ESV_009176	KU641485	Chordata	Perciformes	Moronidae	Morone	Morone americana	100
ESV_009176	LC104592	Chordata	Perciformes	Moronidae	Morone	Morone americana	100
ESV_007836	HQ331537	Chordata	Clupeiformes	Clupeidae	Alosa	Alosa sapidissima	100
ESV_007836	LC091583	Chordata	Clupeiformes	Clupeidae	Alosa	Alosa sapidissima	100
ESV_007836	AP009131	Chordata	Clupeiformes	Clupeidae	Alosa	Alosa alosa	100

The sequence represented by an ESV is generated from a denoising algorithm. ESVs are generated by examining the most abundant sequences and then removing sequences that have random differences in them as opposed to consistent differences. Those with consistent differences, even just one base difference, become their own ESV instead of being lumped together.

100% indicates all base pairs matched
99% is roughly 1-2 mismatches per sequence
95% is about 7 base pair mismatches.