

Additional file 12 - Automated annotation based on different BLAST strategies

Summary of the automated annotation method, detailing the different queries (e.g. *Pararge aegeria* translated Open Reading Frames (ORFs), or full nucleotide (nt) sequences), BLAST strategies employed and the publicly available sequence databases used for each of the queries/BLASTs. BLAST outputs were ranked, all hits were pooled and subsequently sorted with a minimum similarity score threshold as explained in the text.

VII* refers to the combined annotated protein databases for 6 insect species (*Drosophila melanogaster*, *Culex quinquefasciatus*, *Aedes aegypti*, *Pediculus humanus corporis*, *Bombyx mori* and *Harpegnathos saltator*) and 1 Arachnid (*Ixodes scapularis*). These databases were specifically consulted for the depth of their annotation.

DB	BLAST+	Group	Query	Taxonomy	Filter	Rank
nr	BLASTp	Alpha	ORF (aa)	<i>Eukaryota</i>	<i>UniProtKB/Swiss-Prot</i>	A
				<i>VII*</i>	-	B
			Mt ORF (aa)	<i>Protostomia</i>	<i>Mitochondrion</i>	C
	BLASTx	Beta	Contig (nt)	<i>Eukaryota</i>	<i>UniProtKB/Swiss-Prot</i>	D
				<i>VII*</i>	-	E
				<i>Protostomia</i>	-	F
nt	BLASTn	Delta	ORF (nt)	<i>Eukaryota</i>	<i>CDS</i>	G
			Contig (nt)	<i>Eukaryota</i>	<i>CDS</i>	H
		Zeta	Contig (nt)	-	-	I