Jumping the fine LINE between species:

Horizontal transfer and evolution of repetitive elements in eukaryotic species

APPENDICES

By

Atma Maria Ivancevic

THE UNIVERSITY
of ADELAIDE

Department of Genetics and Evolution

School of Biological Sciences

A thesis presented for the degree of Doctor of Philosophy

December 2016

Table of Contents

Appendix A

Supplementary for Chapter 1

Table 1: Examples of known eukaryotic HT cases and proposed vectors

Table A.1: **HT cases:** Extended from Supplemental Table 1 [1] to include the most recent cases. The table is categorised by TE type, name, minimum number of HT events recorded, organisms involved, HT criteria met, vectors (where plausible vectors have been proposed), and references. Cases where the organisms are marked with * indicate cross-phyla HT. The abbreviations for HT criteria follow Schaack *et al.* (2010) [1] and Loreto *et al.* (2008) [2]: ss = sequence similarity; Ks = comparison between the number of synonymous mutations observed ar orthologous genes and the number of synonymous mutations observed in TEs; dN/dS = test for purifying selection; phyl = phylogeny of the TE in-congruent with the phylogeny of the host; pd = patchy taxonomic distribution of the TE.

No	**TE**	**Min. # of HTs**	**Organism**	**HT criteria met**	**Proposed vectors**	**References**
DNA Transposon						
1	*P*	14	*Drosophila*	phyl; pd; ss; dN/dS	*P. regalis*	[3–8]
2	*mariner*	9	Insects	Ks; ss; phyl	Polyd-naviruses	[9–13]
3	*Minos*	6	*Drosophila*	ss; phyl; pd; dN/dS	*D. hydei*	[14, 15]
4	*Bari*	1	*Drosophila*	Ks		[16, 17]
5	*MER46, 6, 30*	4	Animals*	ss		[18]
6	*Hsmar1*	3	Animals*	ss; phyl; pd		[19–21]
7	*PPTN*	1	Vertebrates	ss; phyl; pd		[22]
8	*ITmD37E*	2	Mosquitoes	ss; phyl; pd; dN/dS		[23]
9	*hobo*	3	*Drosophila*	ss; phyl		[24, 25]
10	SPIN	13	Animals*	ss; pd; phyl; dN/dS	*R. prolixus*, *L. stagnalis*	[26–29]
11	*h*AT1	4	Animals*	ss; pd; phyl; dN/dS	*R. prolixus*	[27, 30]
12	ET	3	Animals*	ss; pd; phyl; dN/dS	*R. prolixus*	[27]
13	OC1	12	Animals*	ss; pd; phyl; dN/dS	*R. prolixus*	[27, 30, 31]
14	*h*AT-HT3	3	Animals*	ss; pd; phyl; dN/dS	*R. prolixus*	[30]
15	*Tol2*	1	Medaka	ss		[32]
16	*harrow*	5	*Drosophila*	ss; phyl; pd; Ks		[33]
17	*hosimary*	2	*Drosophila*	ss; pd; dN/dS		[34]

No	TE	Min. # of HTs	Organism	HT criteria met	Proposed vectors	References
18	*IS5*	1	Animals/Bacteria*	phyl		[35]
19	*PiggyBac*	2	Mammals	ss; pd	Baculovirus	[36]
20	*Helitron*	>10	Animals* (vertebrates & invertebrates)	ss; pd; dN/dS	Bracoviruses, *R. prolixus*, *Wolbachia*	[37]
LTR Retrotransposon						
21	*gypsy*	26	*Drosophila*	ss; pd; phyl; dN/dS	Can produce virus-like particles	[38]
22	*Gtwin*	3	*Drosophila*	ss; phyl; dN/dS		[39]
23	*SURL*	7	Sea urchin	phyl; Ks		[40]
24	*HMS-beagle*	2	*Drosophila*	Ks		[16, 17]
25	*opus*	1	*Drosophila*	Ks		[17]
26	*blood*	1	*Drosophila*	Ks		[16, 17]
27	*Burdock*	2	*Drosophila*	Ks		[16]
28	*3S18*	2	*Drosophila*	Ks		[16]
29	*Springer*	1	*Drosophila*	Ks		[16]
30	*diver2*	1	*Drosophila*	Ks		[16]
31	*Max*	2	*Drosophila*	Ks		[16]
32	*mdg1*	1	*Drosophila*	Ks		[16]
33	*Stalker2*	1	*Drosophila*	Ks		[16]
34	*Tabor*	2	*Drosophila*	Ks		[16]
35	*412*	1	*Drosophila*	Ks		[16, 17]
36	*Invader6*	2	*Drosophila*	Ks		[16]
37	*flea*	1	*Drosophila*	Ks		[16]
38	*mdg3*	2	*Drosophila*	Ks		[16]
39	*Micropia*	7	*Drosophila*	Ks; ss; phyl; dN/dS		[16, 38]
40	*297*	1	*Drosophila*	Ks		[16]
41	*Transpac*	1	*Drosophila*	Ks		[16]
42	*Pifo*	1	*Drosophila*	Ks		[16]
43	*copia*	4	*Drosophila*	ss; pd; phyl		[41–43]
44	*Penelope*	11	*Drosophila*	ss; phyl; Ks; pd		[44–46]
Non-LTR Retrotransposon						
45	*Jockey*	1	*Drosophila*	Ks		[16, 17]
46	*doc*	1	*Drosophila*	Ks		[16, 17]
47	*F*	1	*Drosophila*	Ks		[17]
48	RTE	6	Tetrapods	ss; phyl; pd		[47, 48]
49	Smal (SINE)	1	Salmonidae	ss; pd		[49]
50	Rex1	1	Teleost	pd; phyl		[50]
51	CR1	3	Lepidopterans	ss; pd; dS	*mariner* DNA transposons	[51, 52]

No	TE	Min. # of HTs	Organism	HT criteria met	Proposed vectors	References
52	BovB	9	Animals*	pd; phyl; ss	*B. hydrosauri*, *A. limbatum*	[53]

Appendix B

Supplementary for Chapter 2

B.1 Materials and Methods

Table 1: Set of 503 genomes used in this study

Table B.1: **Genomic dataset:** Shows the systematic name, common name, genome version, source and submitter of all the genomes tested for L1 elements. Genomes that were acquired through private collaboration (not publicly available) are marked as 'Private' in the source column. The genomes are listed as they appear in our inferred tree of life (Fig. B.1) with headings indicating the class or species group (e.g. MAMMALIA) and order (e.g. Monotremata). The following abbreviations are used for submitters:
Agencourt Bioscience Corporation = Agen;
Ant Genomics Consortium = AGC;
Aquatic Genome Models = AGM;
Baylor College of Medicine = BCM;
Beijing Genomics Institute = BGI;
Broad Institute = Broad;
California Institute of Technology = Caltech;
Chinese Human Genome Center at Shanghai = CHGC;
Chinese University of Hong Kong = CUHK;
College of Animal Science, Inner Mongolian Agricultural University, China = IMAU;
DOE Joint Genome Institute = JGI;
European Bioinformatics Institute (EMBL-EBI) = EBI;
Genome Reference Consortium = GRC;
Genome Sequencing Platform = GSP;
Genome Sequencing Platform, The Genome Assembly Team = GAT;
Glossina Genomes Consortium = GGC;
i5k Initiative = i5k;
Institute of Molecular and Cell Biology = IMCB;
International Crocodilian Genomes Working Group = ICGWG;
J. Craig Venter Institute = JCVI;
Kazusa DNA Research Institute = KDRI;
Max-Planck Institute = MP;
McGill University = McGill;
modENCODE Project = modENCODE;
Oryza Map Alignment Project = OMAP;
School of Biological & Chemical Sciences, Queen Mary University of London = QMUL;
Seoul National University = SNU;
Texas A&M University = TAMU;
Tokyo Institute of Technology = TokyoTech;
University of Illinois at Urbana-Champaign = UIUC;
University of Lausanne = UNIL;
University of Maryland = Mary;
University of Southern California = USC;
Uppsala University = UU;
Washington University = WashU;
Wellcome Trust Sanger Institute = WTSI;
ZF-screens B.V. = ZF-S.

No	Systematic Name	Common Name	Genome Version	Source	Submitter
MAMMALIA					
Monotremata					
1	*Tachyglossus aculeatus*	Echidna	Tachyglossus	Private	-

No	Systematic Name	Common Name	Genome Version	Source	Submitter
2	*Ornithorhynchus anatinus*	Platypus	ornAna1	UCSC	WashU
Marsupialia					
3	*Monodelphis domestica*	Opossum	monDom5	UCSC	GAT
4	*Macropus eugenii*	Tammar wallaby	Meug_1.1	NCBI	Tammar Wallaby Genome Sequencing Consortium
5	*Sarcophilus harrisii*	Tasmanian devil	sarHar1	UCSC	WTSI
Xenarthra					
6	*Dasypus novemcinctus*	Armadillo	Dasnov3.0	NCBI	BCM
7	*Choloepus hoffmanni*	Sloth	choHof1	UCSC	Broad
Afrotheria					
8	*Chrysochloris asiatica*	Cape golden mole	ChrAsi1.0	NCBI	Broad
9	*Echinops telfairi*	Tenrec	EchTel2.0	NCBI	Broad
10	*Orycteropus afer afer*	Aardvark	OryAfe1.0	NCBI	Broad
11	*Elephantulus edwardii*	Cape elephant shrew	EleEdw1.0	NCBI	Broad
12	*Trichechus manatus latirostris*	Manatee	TriManLat1.0	NCBI	Broad
13	*Procavia capensis*	Rock hyrax	proCap1	UCSC	BCM
14	*Loxodonta africana*	Elephant	LAv4	Private	-
Insectivora					
15	*Erinaceus europaeus*	Hedgehog	EriEur2.0	NCBI	Broad
16	*Sorex araneus*	Common shrew	SorAra2.0	NCBI	Broad
17	*Condylura cristata*	Star-nosed mole	ConCri1.0	NCBI	Broad
Chiroptera					
18	*Pteropus alecto*	Black flying fox	ASM32557v1	NCBI	BGI
19	*Pteropus vampyrus*	Megabat	pteVam1	UCSC	BCM
20	*Eidolon helvum*	Straw-coloured fruit bat	ASM46528v1	NCBI	QMUL
21	*Megaderma lyra*	Greater false vampire bat	ASM46534v1	NCBI	QMUL
22	*Rhinolophus ferrumequinum*	Greater horseshoe bat	ASM46549v1	NCBI	QMUL
23	*Pteronotus parnellii*	Parnells mustached bat	ASM46540v1	NCBI	QMUL
24	*Eptesicus fuscus*	Big brown bat	EptFus1.0	NCBI	Broad
25	*Myotis brandtii*	Brandts bat	ASM41265v1	NCBI	BGI
26	*Myotis davidii*	Mouse-eared bat	ASM32734v1	NCBI	BGI
27	*Myotis lucifugus*	Microbat	Myoluc2.0	NCBI	Broad
Perissodactyla					
28	*Ceratotherium simum simum*	White rhino	CerSimSim1.0	NCBI	Broad
29	*Equus przewalskii*	Przewalski horse	Burgud	NCBI	IMAU
30	*Equus caballus* (Thoroughbred)	Thoroughbred horse	equCab2	UCSC	GAT

No	Systematic Name	Common Name	Genome Version	Source	Submitter
31	*Equus caballus* (Mongolian)	Mongolian horse	Ajinai1.0	NCBI	IMAU
Pholidota					
32	*Manis pentadactyla*	Chinese pangolin	M_pentadactyla-1.1.1	NCBI	WashU
Carnivora					
33	*Felis catus*	Cat	felCat5	UCSC	International Cat Genome Sequencing Consortium
34	*Panthera tigris altaica*	Siberian tiger	PanTig1.0	NCBI	Personal Genomics Institute
35	*Canis lupus familiaris*	Dog	CanFam3.1	NCBI	Dog Genome Sequencing Consortium
36	*Ursus maritimus*	Polar bear	UrsMar_1.0	NCBI	BGI
37	*Ailuropoda melanoleuca*	Panda	ailMel1	UCSC	BGI
38	*Leptonychotes weddellii*	Weddell seal	LepWed1.0	NCBI	Broad
39	*Odobenus rosmarus divergens*	Walrus	Oros_1.0	NCBI	Marine Mammals
40	*Mustela putorius furo*	Ferret	MusPutFur1.0	NCBI	Ferret Genome Sequencing Consortium
Cetartiodactyla					
41	*Camelus dromedarius*	Dromedary	PRJNA234474_ Ca_dromedarius_V1.0	NCBI	King Abdulaziz City for Science and Technology
42	*Camelus ferus*	Bactrian camel	CB1	NCBI	Bactrian Camels Genome Sequencing and Analysis Consortium
43	*Vicugna pacos*	Alpaca	Vicugna_pacos-2.0.1	NCBI	WashU
44	*Sus scrofa* (Duroc)	Duroc pig	Sscrofa10.2	NCBI	Swine Genome Sequencing Consortium
45	*Sus scrofa* (Tibetan)	Tibetan pig	Tibetan_Pig_v1.0	NCBI	Novogene
46	*Sus scrofa* (Ellegaard Gottingen minipig)	Minipig	SscrofaMinipig	NCBI	GlaxoSmithKline
47	*Balaenoptera acutorostrata scammoni*	Minke whale	BalAcu1.0	NCBI	Korea Ocean Research & Development Institute
48	*Physeter catodon*	Sperm whale	Physeter_ macrocephalus-2.0.2	NCBI	AGM
49	*Lipotes vexillifer*	Baiji (Chinese dolphin)	Lipotes_vexillifer_v1	NCBI	BGI
50	*Tursiops truncatus*	Bottlenose dolphin	Ttru_1.4	NCBI	BCM
51	*Orcinus orca*	Killer whale	Oorc_1.1	NCBI	Marine Mammals
52	*Pantholops hodgsonii*	Tibetan antelope	PHO1.0	NCBI	BGI
53	*Capra hircus*	Goat	CHIR_1.0	NCBI	International Goat Genome Consortium
54	*Ovis aries* (Texel)	Domestic sheep	oviAri3	UCSC	International Sheep Genome Consortium
55	*Ovis aries musimon*	Wild sheep	Oori1	NCBI	EBI

No	Systematic Name	Common Name	Genome Version	Source	Submitter
56	*Bubalus bubalis*	Water buffalo	UMD_CASPUR_ WB_2.0	NCBI	Mary
57	*Bison bison bison*	Bison	Bison_UMD1.0	NCBI	Mary
58	*Bos mutus*	Yak	BosGru_v2.0	NCBI	BGI
59	*Bos indicus*	Zebu	Bos_indicus_1.0	NCBI	Genoa Biotecnologia SA
60	*Bos taurus*	Cow	bosTau6	UCSC	Mary
Lagomorpha					
61	*Ochotona princeps*	American pika	OchPri3.0	NCBI	Broad
62	*Oryctolagus cuniculus*	Rabbit	oryCun2	UCSC	GSP
Rodentia					
63	*Ictidomys tridecemlineatus*	Squirrel	SpeTri2.0	NCBI	Broad
64	*Heterocephalus glaber*	Naked molerat	HetGla_female_1.0	NCBI	Broad
65	*Fukomys damarensis*	Damaraland molerat	DMR_v1.0	NCBI	BGI
66	*Cavia aperea*	Brazilian guinea pig	CavAp1.0	NCBI	Leibniz Institute for Zoo and Wildlife research
67	*Cavia porcellus*	Guinea pig	cavPor3	UCSC	GSP
68	*Chinchilla lanigera*	Chinchilla	ChiLan1.0	NCBI	Broad
69	*Octodon degus*	Degu	OctDeg1.0	NCBI	Broad
70	*Dipodomys ordii*	Kangaroo rat	dipOrd1	UCSC	BCM
71	*Jaculus jaculus*	Lesser Egyptian jerboa	JacJac1.0	NCBI	Broad
72	*Nannospalax galili*	Blind molerat	S.galili_v1.0	NCBI	BGI
73	*Mesocricetus auratus*	Golden hamster	MesAur1.0	NCBI	Broad
74	*Cricetulus griseus*	Chinese hamster	CriGri_1.0	NCBI	BGI
75	*Microtus ochrogaster*	Prairie vole	MicOch1.0	NCBI	Broad
76	*Peromyscus maniculatus bairdii*	Deer mouse	Pman_1.0	NCBI	BCM
77	*Rattus norvegicus*	Rat	Rnor_5.0	NCBI	Rat Genome Sequencing Consortium
78	*Mus musculus*	Mouse	GRCm38.p2	NCBI	GRC
Scandentia					
79	*Tupaia belangeri*	Tree shrew	tupBel1	UCSC	Broad
80	*Tupaia chinensis*	Chinese tree shrew	TupChi_1.0	NCBI	BGI
Primates					
81	*Galeopterus variegatus*	Flying lemur	G_variegatus-3.0.2	NCBI	WashU
82	*Otolemur garnettii*	Bushbaby	OtoGar3	NCBI	Broad
83	*Microcebus murinus*	Mouse lemur	micMur1	UCSC	GSP
84	*Tarsius syrichta*	Tarsier	Tarsius_syrichta-2.0.1	NCBI	WashU
85	*Callithrix jacchus*	Marmoset	Callithrix_jacchus-3.2	NCBI	WashU
86	*Saimiri boliviensis boliviensis*	Squirrel monkey	SaiBol1.0	NCBI	Broad
87	*Rhinopithecus roxellana*	Snub-nosed monkey	Rrox_v1	NCBI	Novogene

No	Systematic Name	Common Name	Genome Version	Source	Submitter
88	*Nasalis larvatus*	Proboscis monkey	Charlie1.0	NCBI	Proboscis Monkey Functional Genome Consortium
89	*Chlorocebus sabaeus*	Green monkey	Chlorocebus_sabeus 1.1	NCBI	Vervet Genomics Consortium
90	*Macaca fascicularis*	Crab-eating macaque	Macaca_fascicularis_ 5.0	NCBI	WashU
91	*Macaca mulatta*	Rhesus macaque	rheMac3	UCSC	BGI
92	*Papio anubis*	Baboon	Panu_2.0	NCBI	BCM
93	*Nomascus leucogenys*	Gibbon	nomLeu3	UCSC	Gibbon Genome Sequencing Consortium
94	*Pongo abelii*	Sumatran orangutan	P_pygmaeus_2.0.2	NCBI	Orangutan Genome Sequencing Consortium
95	*Gorilla gorilla gorilla*	Gorilla	gorGor3.1	NCBI	WTSI
96	*Pan paniscus*	Bonobo	panpan1	NCBI	MP
97	*Pan troglodytes*	Chimp	panTro4	UCSC	Chimpanzee Sequencing and Analysis Consortium
98	*Homo sapiens*	Human	GRCh37 (hg19)	NCBI	GRC
SAUROPSIDA					
Testudines					
99	*Apalone spinifera*	Spiny softshell turtle	ASM38561v1	NCBI	WUGSC
100	*Pelodiscus sinensis*	Chinese softshell turtle	PelSin_1.0	NCBI	P. sinensis Genome Project Consortium
101	*Chelonia mydas*	Green sea turtle	CheMyd_1.0	NCBI	BGI
102	*Chrysemys picta bellii*	Painted turtle	chrPic1	UCSC	Painted Turtle Genome Sequencing Consortium
Struthioniformes					
103	*Struthio camelus australis*	Ostrich	ASM69896v1	NCBI	BGI
Tinamiformes					
104	*Tinamus guttatus*	Tinamou	ASM70537v2	NCBI	BGI
Anseriformes					
105	*Anas platyrhynchos*	Mallard	BGI_duck_1.0	NCBI	State Key Laboratory for Agrobiotechnology, China Agricultural University, Beijing
Galliformes					
106	*Lyrurus tetrix tetrix*	Black grouse	tetTet1	NCBI	UU
107	*Gallus gallus*	Chicken	galGal4	UCSC	International Chicken Genome Consortium
108	*Coturnix japonica*	Japanese quail	Coja_1.0	NCBI	Tokyo University of Agriculture
109	*Meleagris gallopavo*	Turkey	melGal1	UCSC	Turkey Genome Consortium
110	*Colinus virginianus*	Bobwhite	NB1.1	NCBI	TAMU
Passeriformes					

No	Systematic Name	Common Name	Genome Version	Source	Submitter
111	*Acanthisitta chloris*	Rifleman	ASM69581v1	NCBI	BGI
112	*Manacus vitellinus*	Manakin	ASM69201v1	NCBI	BGI
113	*Zonotrichia albicollis*	White-throated sparrow	Zonotrichia_albicollis-1.0.1	NCBI	White-throated Sparrow Consortium
114	*Geospiza fortis*	Medium ground finch	geoFor1	UCSC	BGI
115	*Serinus canaria*	Atlantic canary	SCA1	NCBI	MP
116	*Taeniopygia guttata*	Zebra finch	taeGut1	UCSC	WashU
117	*Ficedula albicollis*	Collared flycatcher	FicAlb1.5	NCBI	UU
118	*Pseudopodoces humilis*	Ground tit	PseHum1.0	NCBI	BGI
119	*Corvus brachyrhynchos*	American crow	ASM69197v1	NCBI	BGI
120	*Corvus cornix cornix*	Hooded crow	Hooded_Crow_genome	NCBI	UU
Psittaciformes					
121	*Ara macao*	Scarlet macaw	SMACv1.1	NCBI	TAMU
122	*Amazona vittata*	Puerto Rican parrot	AV1	NCBI	Puerto Rican Parrot Genome Project
123	*Melopsittacus undulatus*	Budgerigar	melUnd1	UCSC	WashU
124	*Nestor notabilis*	Kea	ASM69687v1	NCBI	BGI
Falconiformes					
125	*Falco cherrug*	Saker falcon	F_cherrug_v1.0	NCBI	BGI
126	*Falco peregrinus*	Peregrine falcon	F_peregrinus_v1.0	NCBI	BGI
Cariamiformes					
127	*Cariama cristata*	Seriema	ASM69053v1	NCBI	BGI
Coraciiformes					
128	*Merops nubicus*	Bee eater	ASM69184v1	NCBI	BGI
Piciformes					
129	*Picoides pubescens*	Woodpecker	ASM69900v1	NCBI	BGI
Bucerotiformes					
130	*Buceros rhinoceros silvestris*	Hornbill	ASM71030v1	NCBI	BGI
Trogoniformes					
131	*Apaloderma vittatum*	Trogon	ASM70340v1	NCBI	BGI
Leptosomiformes					
132	*Leptosomus discolour*	Cuckoo roller	ASM69178v1	NCBI	BGI
Accipitriformes					
133	*Haliaeetus albicilla*	White-tailed eagle	ASM69140v1	NCBI	BGI
134	*Haliaeetus leucocephalus*	Bald eagle	Haliaeetus_leucocephalus-4.0	NCBI	The Bald Eagle Consortium
135	*Aquila chrysaetos canadensis*	Golden eagle	Aquila_chrysaetos-1.0.2	NCBI	WashU
136	*Cathartes aura*	Turkey vulture	ASM69994v1	NCBI	BGI
Strigiformes					
137	*Tyto alba*	Barn owl	ASM68720v1	NCBI	BGI
Coliiformes					
138	*Colius striatus*	Mousebird	ASM69071v1	NCBI	BGI

No	Systematic Name	Common Name	Genome Version	Source	Submitter
Charadriiformes					
139	*Charadrius vociferus*	Killdeer	ASM70802v2	NCBI	BGI
Gruiformes					
140	*Balearica regulorum gibbericeps*	Grey crane	ASM70989v1	NCBI	BGI
141	*Chlamydotis macqueenii*	MacQueen’s bustard	ASM69519v1	NCBI	BGI
Cuculiformes					
142	*Cuculus canorus*	Common cuckoo	ASM70932v1	NCBI	BGI
Procellariiformes					
143	*Fulmarus glacialis*	Fulmar	ASM69083v1	NCBI	BGI
Sphenisciformes					
144	*Aptenodytes forsteri*	Emperor penguin	ASM69914v1	NCBI	BGI
145	*Pygoscelis adeliae*	Adelie penguin	ASM69910v1	NCBI	BGI
Pelecaniformes					
146	*Phalacrocorax carbo*	Black cormorant	ASM70892v1	NCBI	BGI
147	*Pelecanus crispus*	Pelican	ASM68737v1	NCBI	BGI
148	*Nipponia nippon*	Ibis	ASM70822v1	NCBI	BGI
149	*Egretta garzetta*	Egret	ASM68718v1	NCBI	College of Medicine and Forensics, Xi’an Jiaotong University
150	*Phaethon lepturus*	Tropicbird	ASM68728v1	NCBI	BGI
Gaviiformes					
151	*Gavia stellata*	Loon	ASM69087v1	NCBI	BGI
Musophagiformes					
152	*Tauraco erythrolophus*	Turaco	ASM70936v1	NCBI	BGI
Opisthocomiformes					
153	*Opisthocomus hoazin*	Hoatzin	ASM69207v1	NCBI	BGI
Columbiformes					
154	*Columba livia*	Rock dove	Cliv_1.0	NCBI	BGI
Pteroclidiformes					
155	*Pterocles gutturalis*	Sandgrouse	ASM69924v1	NCBI	BGI
Apodiformes					
156	*Calypte anna*	Anna’s hummingbird	ASM69908v1	NCBI	BGI
157	*Chaetura pelagica*	Chimney swift	ChaPel_1.0	NCBI	BGI
Caprimulgiformes					
158	*Caprimulgus carolinensis*	Chuck-will’s-widow	ASM70074v1	NCBI	BGI
Eurypygiformes					
159	*Eurypyga helias*	Sunbittern	ASM69077v1	NCBI	BGI
Mesitornithiformes					
160	*Mesitornis unicolor*	Mesite	ASM69576v1	NCBI	BGI
Podicipediformes					
161	*Podiceps cristatus*	Grebe	ASM69954v1	NCBI	BGI
Phoenicopteriformes					

No	Systematic Name	Common Name	Genome Version	Source	Submitter
162	*Phoenicopterus ruber ruber*	Flamingo	ASM68726v1	NCBI	BGI
Crocodilia					
163	*Alligator mississippiensis*	American alligator	allMis1	UCSC	ICGWG
164	*Alligator sinensis*	Chinese alligator	ASM45574v1	NCBI	BGI
165	*Crocodylus porosus*	Saltwater crocodile	Cpor_2.0	NCBI	ICGWG
166	*Gavialis gangeticus*	Gharial crocodile	ggan_v0.2	NCBI	ICGWG
Squamata					
167	*Pogona vitticeps*	Bearded dragon	Pogona_vitticeps.male	Private (Terry)	-
168	*Anolis carolinensis*	Anole lizard	anoCar2	UCSC	Broad
169	*Vipera berus berus*	Common European viper	Vber.be_1.0	NCBI	BCM-HGSC
170	*Crotalus mitchellii pyrrhus*	Pit viper	CrotMitch1.0	NCBI	Reed College
171	*Ophiophagus hannah*	King cobra	OphHan1.0	NCBI	Naturalis Biodiversity Center
172	*Python bivittatus*	Burmese python	Python_molurus_bivittatus-5.0.2	NCBI	The Consortium for Comparative Genomics, UC Denver
AMPHIBIA					
Anura					
173	*Nanorana parkeri*	Tibetan frog	ASM93562v1	NCBI	BGI
174	*Xenopus tropicalis*	Western clawed frog	Xtropicalis_v7	NCBI	JGI
NEOPTERYGII					
Lepisosteiformes					
175	*Lepisosteus oculatus*	Spotted gar	LepOcu1	NCBI	Broad
Anguilliformes					
176	*Anguilla anguilla*	European eel	Anguilla_anguilla_v1_09_nov_10	NCBI	ZF-S
177	*Anguilla japonica*	Japanese eel	japanese_eel_genome_v1_25_oct_2011_japonica_c401b400k25m200_sspacepremium k3a02n24_extra.final.scaffolds	NCBI	ZF-S
Cypriniformes					
178	*Danio rerio*	Zebrafish	danRer7	UCSC	WTSI
Characiformes					
179	*Astyanax mexicanus*	Mexican tetra	Astyanax_mexicanus-1.0.2	NCBI	AGM
Beloniformes					

No	Systematic Name	Common Name	Genome Version	Source	Submitter
180	*Oryzias latipes*	Medaka	ASM31367v1	NCBI	Medaka genome sequencing project
Cyprinodontiformes					
181	*Poecilia formosa*	Amazon molly	Poecilia_formosa-5.1.2	NCBI	AGM
182	*Xiphophorus maculatus*	Southern platyfish	Xiphophorus_ maculatus-4.4.2	NCBI	The Genome Institute, Washington University at St. Louis
183	*Fundulus heteroclitus*	Mummichog	Fundulus_heteroclitus-3.0.2	NCBI	WashU
Tetraodontiformes					
184	*Takifugu flavidus*	Yellowbelly pufferfish	version 1 of Takifugu flavidus genome	NCBI	IOCAS
185	*Takifugu rubripes*	Fugu	fr3	UCSC	The Fugu Genome Sequencing Consortium
186	*Tetraodon nigroviridis*	Tetraodon	tetNig2	UCSC	Genoscope
Pleuronectiformes					
187	*Cynoglossus semilaevis*	Tongue sole	Cse_v1.0	NCBI	BGI
Perciformes					
188	*Haplochromis burtoni*	Burton's haplo	AstBur1.0	NCBI	Broad
189	*Pundamilia nyererei*	Flameback cichlid	PunNye1.0	NCBI	Broad
190	*Maylandia zebra*	Zebra mbuna	MetZeb1.1	NCBI	Broad
191	*Neolamprologus brichardi*	Fairy cichlid	NeoBri1.0	NCBI	Broad
192	*Oreochromis niloticus*	Nile tilapia	oreNil2	UCSC	Broad
193	*Sebastes nigrocinctus*	Tiger rockfish	Snig1.0	NCBI	USC
194	*Sebastes rubrivinctus*	Flag rockfish	SRub1.0	NCBI	USC
195	*Gasterosteus aculeatus*	Stickleback	gasAcu1	UCSC	Broad
Gadiformes					
196	*Gadus morhua*	Atlantic cod	gadMor1	UCSC	Genofisk
CHONDRICHTHYES					
Chimaeriformes					
197	*Callorhinchus milii*	Australian ghostshark	Callorhinchus_milii-6.1.3	NCBI	IMCB
Carcharhiniformes					
198	*Carcharhinus brachyurus*	Copper shark	shark_ass	Private (Terry)	-
ECDYSOZOA					
Ephemeroptera					
199	*Ephemera danica*	Green drake	Edan_1.0	NCBI	i5k
Odonata					
200	*Ladona fulva*	Dragonfly	Lful_1.0	NCBI	i5k
Phthiraptera					
201	*Pediculus humanus corporis*	Body louse	JCVI_LOUSE_1.0	NCBI	JCVI
Thysanoptera					

No	Systematic Name	Common Name	Genome Version	Source	Submitter
202	*Frankliniella occidentalis*	Flower thrips	Focc_1.0	NCBI	i5k
Hemiptera					
203	*Diaphorina citri*	Asian citrus psyllid	Diaci psyllid genome assembly version 1.1	NCBI	International Psyllid Genome Consortium
204	*Pachypsylla venusta*	Petiolegall psyllid	Pven_1.0	NCBI	i5k
205	*Acyrthosiphon pisum*	Pea aphid	Acyr_2.0	NCBI	BCM
206	*Nilaparvata lugens*	Brown planthopper	NilLug1.0	NCBI	Nilaparvata lugens Genome Consortium
207	*Oncopeltus fasciatus*	Milkweed bug	Ofas_1.0	NCBI	BCM-HGSC i5k
208	*Rhodnius prolixus*	Assassin bug	Rhodnius_prolixus-3.0.1	NCBI	WashU
209	*Cimex lectularius*	Bed bug	Clec_1.0	NCBI	i5k
Coleoptera					
210	*Onthophagus taurus*	Taurus scarab	Otau_1.0	NCBI	i5k
211	*Agrilus planipennis*	Emerald ash borer	Apla_1.0	NCBI	BCM-HGSC i5k
212	*Tribolium castaneum*	Red flour beetle	Tcas_3.0	NCBI	BCM
213	*Anoplophora glabripennis*	Asian long-horned beetle	Agla_1.0	NCBI	i5k
214	*Leptinotarsa decemlineata*	Colorado potato beetle	Ldec_1.5	NCBI	i5k
215	*Dendroctonus ponderosae*	Mountain pine beetle	DendPond_male_1.0	NCBI	The Tria Project: Mountain Pine Beetle System Genomics
Strepsiptera					
216	*Mengenilla moldrzyki*	Twisted-wing parasite	Memo_1.0	NCBI	Zoologisches Forschungsmuseum Alexander Koenig
Diptera					
217	*Aedes aegypti*	Yellow fever mosquito	AaegL2	NCBI	TIGR
218	*Culex quinquefasciatus*	Southern house mosquito	CulPip1.0	NCBI	Broad
219	*Anopheles albimanus*	-	Anop_albi_ALBI9_A_V1	NCBI	Broad
220	*Anopheles arabiensis*	-	Anop_arab_DONG5_A_V1	NCBI	Broad
221	*Anopheles atroparvus*	-	Anop_atro_EBRO_V1	NCBI	Broad
222	*Anopheles christyi*	-	Anop_chri_ACHKN1017_V1	NCBI	Broad
223	*Anopheles culicifacies*	-	Anop_culi_species_A-37_1_V1	NCBI	Broad
224	*Anopheles darlingi*	-	A_darlingi_v1	NCBI	Laboratorio Nacional de Computacao Cientifica
225	*Anopheles dirus*	-	Anop_diru_WRAIR2_V1	NCBI	Broad

No	Systematic Name	Common Name	Genome Version	Source	Submitter
226	*Anopheles epiroticus*	-	Anop_epir_ epiroticus2_V1	NCBI	Broad
227	*Anopheles farauti*	-	Anop_fara_FAR1_V2	NCBI	Broad
228	*Anopheles funestus*	-	Anop_fune_ FUMOZ_V1	NCBI	Broad
229	*Anopheles gambiae*	African malaria mosquito	anoGam1	UCSC	International Anopheles Genome Project
230	*Anopheles maculatus*	-	Anop_macu_ maculatus3_V1	NCBI	Broad
231	*Anopheles melas*	-	Anop_mela_ CM1001059_ A_V2	NCBI	Broad
232	*Anopheles merus*	-	Anop_meru_MAF_V1	NCBI	Broad
233	*Anopheles minimus*	-	Anop_mini_ MINIMUS1_V1	NCBI	Broad
234	*Anopheles quadriannulatus*	-	Anop_quad_ QUAD4_A_V1	NCBI	Broad
235	*Anopheles sinensis*	-	AS2	NCBI	Nanjing Medical University
236	*Anopheles stephensi*	-	ASM30077v2	NCBI	Virginia Tech
237	*Mayetiola destructor*	Hessian fly	Mdes_1.0	NCBI	BCM
238	*Lutzomyia longipalpis*	Sand fly	Llon_1.0	NCBI	BCM
239	*Phlebotomus papatasi*	-	Ppap_1.0	NCBI	WashU
240	*Ceratitis capitata*	Med fly	Ccap_1.0	NCBI	i5k
241	*Drosophila albomicans*	-	DroAlb_1.0	NCBI	Kunming Institute of Zoology, Chinese Academy of Sciences
242	*Drosophila ananassae*	-	droAna3	UCSC	Agen
243	*Drosophila biarmipes*	-	Dbia_2.0	NCBI	BCM
244	*Drosophila bipectinata*	-	Dbip_2.0	NCBI	BCM
245	*Drosophila elegans*	-	Dele_2.0	NCBI	BCM
246	*Drosophila erecta*	-	droEre2	UCSC	Agen
247	*Drosophila eugracilis*	-	Deug_2.0	NCBI	modENCODE
248	*Drosophila ficusphila*	-	Dfic_2.0	NCBI	BCM
249	*Drosophila grimshawi*	-	droGri2	UCSC	Agen
250	*Drosophila kikkawai*	-	Dkik_2.0	NCBI	BCM
251	*Drosophila melanogaster*	-	Release 6 plus ISO1 MT	NCBI	The FlyBase Consortium/Berkeley Drosophila Genome Project/Celera Genomics
252	*Drosophila miranda*	-	DroMir_2.2	NCBI	University of California, Berkeley
253	*Drosophila mojavensis*	-	droMoj3	UCSC	Agen
254	*Drosophila persimilis*	-	droPer1	UCSC	Broad

No	Systematic Name	Common Name	Genome Version	Source	Submitter
255	*Drosophila pseudoobscura pseudoobscura*	-	Dpse_3.0	NCBI	BCM
256	*Drosophila rhopaloa*	-	Drho_2.0	NCBI	modENCODE
257	*Drosophila sechellia*	-	droSec1	UCSC	Broad
258	*Drosophila simulans*	-	dsim_caf1	NCBI	WashU
259	*Drosophila suzukii*	-	Dsuzukii.v01	NCBI	BGI
260	*Drosophila takahashii*	-	Dtak_2.0	NCBI	BCM
261	*Drosophila virilis*	-	droVir3	UCSC	Agen
262	*Drosophila willistoni*	-	dwil_caf1	NCBI	JCVI
263	*Drosophila yakuba*	-	dyak_caf1	NCBI	FlyBase
264	*Musca domestica*	House fly	Musca_domestica-2.0.2	NCBI	GGC
265	*Glossina austeni*	Tsetse fly	Glossina_austeni-1.0.3	NCBI	GGC
266	*Glossina brevipalpis*	-	Glossina_brevipalpis_1.0.3	NCBI	GGC
267	*Glossina fuscipes fuscipes*	-	Glossina_fuscipes-3.0.2	NCBI	GGC
268	*Glossina morsitans morsitans*	-	ASM107743v1	NCBI	WTSI
269	*Glossina pallidipes*	-	Glossina_pallidipes-1.0.3	NCBI	GGC
Trichoptera					
270	*Limnephilus lunatus*	Caddis fly	Llun_1.0	NCBI	i5k
Lepidoptera					
271	*Papilio glaucus*	Tiger butterfly	pgl_assembly_v1	NCBI	UT Southwestern Medical Center
272	*Papilio polytes*	Mormon butterfly	Ppol_1.0	NCBI	TokyoTech
273	*Papilio xuthus*	Swallowtail butterfly	Pxut_1.0	NCBI	TokyoTech
274	*Heliconius melpomene melpomene*	Postman butterfly	ASM31383v2	NCBI	Heliconius Genome Sequencing Consortium
275	*Melitaea cinxia*	Glanville fritillary	MelCinx1.0	NCBI	University of Helsinki
276	*Danaus plexippus*	Monarch butterfly	DanPle_1.0	NCBI	The Reppert Lab
277	*Bombyx mori*	Silkworm	ASM15162v1	NCBI	The International Silkworm Genome Sequencing Consortium
278	*Manduca sexta*	Tobacco hornworm	Msex_1.0	NCBI	BCM
279	*Plutella xylostella*	Diamondback moth	DBM_FJ_V1.1	NCBI	Plutella xylostella Genome Consortium
Hymenoptera					
280	*Athalia rosae*	Turnip sawfly	Aros_1.0	NCBI	i5k
281	*Cephus cinctus*	Wheat stem sawfly	Ccin1	NCBI	UIUC
282	*Orussus abietinus*	Parasitic wood wasp	Oabi_1.0	NCBI	i5k
283	*Ceratosolen solmsi marchali*	Pollinating wasp	CerSol_1.0	NCBI	Ceratosolen solmsi Genome Consortium
284	*Nasonia giraulti*	Jewel wasp	Ngir_1.0	NCBI	BCM

No	Systematic Name	Common Name	Genome Version	Source	Submitter
285	*Nasonia longicornis*	-	Nlon_1.0	NCBI	BCM
286	*Nasonia vitripennis*	-	Nvit_2.1	NCBI	BCM
287	*Copidosoma floridanum*	Looper parasitoid wasp	Cflo_1.0	NCBI	i5k
288	*Trichogramma pretiosum*	-	Tpre_1.0	NCBI	i5k
289	*Microplitis demolitor*	-	Mdem1	NCBI	UIUC
290	*Megachile rotundata*	Alfalfa leafcutter bee	MROT_1.0	NCBI	University of Maryland
291	*Apis dorsata*	Giant honey bee	Apis dorsata 1.3	NCBI	Cold Spring Harbor Laboratory
292	*Apis florea*	Dwarf honey bee	Aflo_1.0	NCBI	BCM
293	*Apis mellifera*	Western honey bee	Amel_4.5	NCBI	Human Genome Sequencing Center
294	*Bombus impatiens*	Common Eastern bumblebee	BIMP_2.0	NCBI	Biotechnology Center, University of Illinois
295	*Bombus terrestris*	Buff-tailed bumblebee	Bter_1.0	NCBI	BCM
296	*Linepithema humile*	Argentine ant	Lhum_UMD_V04	NCBI	AGC
297	*Camponotus floridanus*	Florida carpenter ant	CamFlo_1.0	NCBI	BGI
298	*Acromyrmex echinatior*	Panamanian leafcutter ant	Aech_3.9	NCBI	BGI
299	*Atta cephalotes*	Leafcutter ant	Attacep1.0	NCBI	WashU
300	*Solenopsis invicta*	Red imported fire ant	Si_gnG	NCBI	UNIL
301	*Pogonomyrmex barbatus*	Red harvester ant	Pbar_UMD_V03	NCBI	AGC
302	*Harpegnathos saltator*	Jumping ant	HarSal_1.0	NCBI	BGI
303	*Cerapachys biroi*	Clonal raider ant	CerBir1.0	NCBI	BGI
Blattodea					
304	*Blattella germanica*	German cockroach	Bger_1.0	NCBI	i5k
Isoptera					
305	*Zootermopsis nevadensis*	Eusocial termite	ZooNev1.0	NCBI	BGI
Diplostraca					
306	*Daphnia pulex*	Water flea	V1.0	NCBI	JGI
Calanoida					
307	*Eurytemora affinis*	Copepod	Eaff_1.0	NCBI	i5k
Amphipoda					
308	*Hyalella azteca*	Scud	Hazt_1.0	NCBI	BCM
Geophilomorpha					
309	*Strigamia maritima*	European centipede	Smar_1.0	NCBI	BCM
Araneae					
310	*Stegodyphus mimosarum*	Social spider	Stegodyphus_ mimosarum_ v1	NCBI	BGI
311	*Latrodectus hesperus*	Black widow	Lhes_1.0	NCBI	i5k
312	*Parasteatoda tepidariorum*	Common house spider	Ptep_1.0	NCBI	i5k
Prostigmata					

No	Systematic Name	Common Name	Genome Version	Source	Submitter
313	*Tetranychus urticae*	Red spider mite	ASM23943v1	NCBI	Spider Mite Consortium
Astigmata					
314	*Dermatophagoides farinae*	House dust mite	Dfarinae1.0	NCBI	CUHK
315	*Sarcoptes scabiei type canis*	Dog mite	SarSca1.0	NCBI	Wright State University
Oribatida					
316	*Achipteria coleoptrata*	-	SM98876v1	NCBI	UNIL
317	*Hypochthonius rufulus*	-	ASM98884v1	NCBI	UNIL
318	*Platynothrus peltifer*	Oribatid mite	ASM98890v1	NCBI	UNIL
319	*Steganacarus magnus*	-	ASM98888v1	NCBI	UNIL
Ixodida					
320	*Ixodes ricinus*	Castor bean tick	ASM97304v1	NCBI	Luxembourg Institute of Health
321	*Ixodes scapularis*	Deer tick	JCVI_ISG_i3_1.0	NCBI	JCVI
322	*Rhipicephalus microplus*	Cattle tick	CCG_Rmi_1.0	NCBI	USDA-ARS
Mesostigmata					
323	*Metaseiulus occidentalis*	Western predatory mite	Mocc_1.0	NCBI	BCM
324	*Varroa destructor*	Honeybee mite	BRL_Vdes_1.0	NCBI	Varroa Genome Sequencing Consortium
Scorpiones					
325	*Centruroides exilicauda*	Baja california bark scorpion	Cexi_1.0	NCBI	i5k
326	*Mesobuthus martensii*	Chinese scorpion	M_martensii_Version_1	NCBI	Institute of Plant Physiology & Ecology, Shanghai Insitutes for Biology Sciences, CAS
Xiphosura					
327	*Limulus polyphemus*	Atlantic horseshoe crab	Limulus_polyphemus-2.1.2	NCBI	WashU
Trichocephalida					
328	*Trichinella spiralis*	Pork worm	Trichinella spiralis-3.7.1	NCBI	WashU
Ascaridida					
329	*Ascaris suum*	Pig roundworm	AscSuum_1.0	NCBI	BGI
Spirurida					
330	*Elaeophora elaphi*	Red deer nematode	EEL001	NCBI	WTSI
331	*Onchocerca volvulus*	-	OVOC001	NCBI	WTSI
Rhabditida					
332	*Steinernema monticolum*	-	S_monti_v1	NCBI	Caltech
333	*Panagrellus redivivus*	Microworm	Pred3	NCBI	Caltech
334	*Haemonchus contortus*	Wireworm	HCON	NCBI	WTSI
335	*Necator americanus*	Human hookworm	N_ americanus_v1	NCBI	WashU
336	*Heterorhabditis bacteriophora*	Beneficial nematode	Heterorhabditis_ bacteriophora-7.0	NCBI	WashU

No	Systematic Name	Common Name	Genome Version	Source	Submitter
337	*Caenorhabditis angaria*	-	ps1010rel4	NCBI	California Institute of Technology, Division of Biology
338	*Caenorhabditis brenneri*	-	C_brenneri-6.0.1b	NCBI	Caenorhabditis brenneri Sequencing and Analysis Consortium
339	*Caenorhabditis briggsae*	-	ASM455v1	NCBI	The C.briggsae Sequencing Consortium
340	*Caenorhabditis elegans*	-	WBcel235	NCBI	C. elegans Sequencing Consortium
341	*Caenorhabditis japonica*	-	C_japonica-7.0.1	NCBI	WashU
342	*Caenorhabditis sp. 11 MAF-2010*	-	Caenorhabditis_sp11_JU1373-3.0.1	NCBI	WashU
Priapulimorphida					
343	*Priapulus caudatus*	Cactus worm	Priapulus_caudatus-4.0.1	NCBI	WashU
ROTIFERA					
Bdelloidea					
344	*Adineta vaga*	Rotifer	AMS_PRJEB1171_v1	NCBI	Genoscope CEA
PLATYHELMINTHES					
Strigeidida					
345	*Schistosoma curassoni*	-	S_curassoni_Dakar	NCBI	WTSI
346	*Schistosoma haematobium*	-	SchHae_1.0	NCBI	CHGC
347	*Schistosoma japonicum*	-	ASM15177v1	NCBI	CHGC
348	*Schistosoma mansoni*	-	ASM23792v2	NCBI	Schistosoma Genome Network
349	*Schistosoma margrebowiei*	-	S_margrebowiei_Zambia	NCBI	WTSI
350	*Schistosoma mattheei*	-	S_mattheei_Denwood	NCBI	WTSI
351	*Schistosoma rodhaini*	-	S_rodhaini_Burundi	NCBI	WTSI
Opisthorchiida					
352	*Clonorchis sinensis*	Chinese liver fluke	C_sinensis-2.0	NCBI	Department of Parasitology, Zhongshan School of Medicine, Sun Yat-sen University
Cyclophyllidea					
353	*Echinococcus granulosus*	Hyper tapeworm	EGRAN001	NCBI	WTSI
354	*Echinococcus multilocularis*	Fox tapeworm	EMULTI001	NCBI	WTSI
355	*Hymenolepis microstoma*	Rodent tapeworm	HMIC001	NCBI	WTSI
ANNELIDA					
Scolecida					
356	*Capitella teleta*	Polychaete worm	Capca1	NCBI	JGI

No	Systematic Name	Common Name	Genome Version	Source	Submitter
Rhynchobdellida					
357	*Helobdella robusta*	Leech	Helobdella robusta v1.0	NCBI	JGI
MOLLUSCA					
Ostreoida					
358	*Crassostrea gigas*	Pacific oyster	oyster_v9	NCBI	BGI
Gastropoda					
359	*Lottia gigantea*	Owl limpet	Helro1	NCBI	JGI
360	*Aplysia californica*	California sea hare	AplCal3.0	NCBI	Broad
361	*Biomphalaria glabrata*	Freshwater snail	ASM45736v1	NCBI	WashU
CNIDARIA					
Actiniaria					
362	*Nematostella vectensis*	Starlet sea anemone	ASM20922v1	NCBI	JGI
Anthoathecata					
363	*Hydra vulgaris*	Freshwater polyp	Hydra_RP_1.0	NCBI	JCVI
TENTACULATA					
Lobata					
364	*Mnemiopsis leidyi*	Warty comb jelly	MneLei_Aug2011	NCBI	National Human Genome Research Institute, National Institutes of Health
PLACOZOA					
365	*Trichoplax adhaerens*	Placozoan	v1.0	NCBI	JGI
PORIFERA					
Haplosclerida					
366	*Amphimedon queenslandica*	Sea sponge	v1.0	NCBI	JGI
VIRIDIPLANTAE					
Mamiellales					
367	*Micromonas pusilla CCMP1545*	-	Micromonas pusilla CCMP1545 v2.0	NCBI	Micromonas Genome Consortium
368	*Micromonas sp. RCC299*	-	ASM9098v2	NCBI	Micromonas Genome Consortium
369	*Ostreococcus lucimarinus CCE9901*	-	ASM9206v1	NCBI	JGI
370	*Ostreococcus tauri*	-	version 050606	NCBI	Laboratoire Arago, France
Chlamydomonadales					
371	*Chlamydomonas reinhardtii*	-	v3.0	NCBI	JGI
Volvocales					
372	*Volvox carteri f. nagariensis*	-	v1.0	NCBI	JGI
Chlorellales					
373	*Chlorella variabilis*	-	v 1.0	NCBI	JGI

No	Systematic Name	Common Name	Genome Version	Source	Submitter
374	*Auxenochlorella protothecoides*	Green microalga	ASM73321v1	NCBI	BGI
375	*Helicosporidium sp. ATCC 50920*	-	Helico_v1.0	NCBI	University of British Columbia
Chlorococcales					
376	*Coccomyxa subellipsoidea C-169*	-	Coccomyxa subellipsoidae v2.0	NCBI	JGI
Klebsormidiales					
377	*Klebsormidium flaccidum*	-	ASM70883v1	NCBI	TokyoTech
Funariales					
378	*Physcomitrella patens*	-	V1.1	NCBI	Moss Genome Consortium
Selaginellales					
379	*Selaginella moellendorffii*	Spikemoss	v1.0	NCBI	Selaginella Consortium
Pinales					
380	*Pinus taeda*	Loblolly pine	PtaedaFosmidLib.0.8	NCBI	UC Davis
Amborellales					
381	*Amborella trichopoda*	Amborellaceae	AMTR1.0	NCBI	Amborella Genome Sequencing Project
Alismatales					
382	*Spirodela polyrhiza*	Greater duckweed	Spirodela_polyrhiza_v01	NCBI	Waksman Institute
Arecales					
383	*Phoenix dactylifera*	Date palm	DPV01	NCBI	Joint Center for Genomics Research
384	*Elaeis oleifera*	American oil palm	EO8	NCBI	Orion Genomics
Zingiberales					
385	*Ensete ventricosum*	Ethiopian banana	v1.1	NCBI	University of Exeter
386	*Musa acuminata subsp. malaccensis*	-	ASM31385v1	NCBI	Genoscope/IG/CEA
Poales					
387	*Sorghum bicolor*	Sorghum	Sorbi1	NCBI	Sorghum Consortium
388	*Zea mays*	Maize	B73 RefGen_v3	NCBI	Maize Genome Sequencing Project
389	*Setaria italica*	Foxtail millet	Setaria V1	NCBI	JGI
390	*Brachypodium distachyon*	Purple false brome	v1.0	NCBI	JGI and The International Brachypodium Initiative
391	*Leersia perrieri*	-	Lperr_V1.4	NCBI	Arizona Genomics Institute
392	*Oryza barthii*	Rice	O.barthii_v1.3	NCBI	Oryza Chr3 Short Arm Comparative Sequencing Project

No	Systematic Name	Common Name	Genome Version	Source	Submitter
393	*Oryza brachyantha*	-	Oryza_brachyantha. v1.4b	NCBI	The Institute of Genetics and Developmental Biology
394	*Oryza glumipatula*	-	Oryza_glumaepatula_ v1.5	NCBI	OMAP
395	*Oryza longistaminata*	-	O_longistaminata_v1.0	NCBI	BGI
396	*Oryza meridionalis*	-	Oryza_meridionalis_ v1.3	NCBI	OMAP
397	*Oryza nivara*	-	Oryza_nivara_v1.0	NCBI	OMAP
398	*Oryza punctata*	-	Oryza_punctata_v1.2	NCBI	OMAP
399	*Oryza sativa Japonica Group*	-	Build 4.0	NCBI	International Rice Genome Sequencing Project
400	*Zizania latifolia*	Manchurian wild rice	Zizania_latifolia_v01	NCBI	Zhejiang University
401	*Aegilops tauschii*	Tausch’s goatgrass	ASM34733v1	NCBI	BGI
402	*Triticum urartu*	Wild einkorn	ASM34745v1	NCBI	BGI
Proteales					
403	*Nelumbo nucifera*	Sacred lotus	Chinese Lotus 1.1	NCBI	UIUC
Fabales					
404	*Lupinus angustifolius*	Blue lupin	Lupin genome scaffold	NCBI	Department of Agriculture and Food, WA government, Australia
405	*Phaseolus vulgaris*	Common bean	PhaVulg1_0	NCBI	JGI
406	*Cajanus cajan*	Pigeon pea	Cajanus_cajan_ Asha_ver1.0	NCBI	AKI-PGI
407	*Vigna angularis var. angularis*	Adzuki bean	Vigna angularis	NCBI	SNU
408	*Vigna radiata var. radiata*	-	Vradiata_ver6	NCBI	SNU
409	*Glycine max*	Soybean	V1.1	NCBI	JGI
410	*Glycine soja*	-	W05v1.0	NCBI	CUHK
411	*Cicer arietinum*	Chickpea	ASM33114v1	NCBI	BGI
412	*Medicago truncatula*	Barrel clover	MedtrA17_3.5	NCBI	International Medicago Genome Annotation Group
413	*Trifolium pratense*	Red clover	Tp1.0	NCBI	Masaryk University
414	*Lotus japonicus*	Birdsfoot trefoil	ASM18111v1	NCBI	KDRI
Rosales					
415	*Malus x domestica*	Apple	MalDomGD1.0	NCBI	IASMA research center
416	*Pyrus x bretschneideri*	Chinese white pear	Pbr_v1.0	NCBI	Nanjing Agricultural University
417	*Prunus mume*	Japanese apricot	P.mume_V1.0	NCBI	BGI
418	*Prunus persica*	Peach	Prupe1_0	NCBI	JGI
419	*Fragaria iinumae*	Japanese strawberry	FII_r1.1	NCBI	KDRI

No	Systematic Name	Common Name	Genome Version	Source	Submitter
420	*Fragaria nubicola*	Himalayan strawberry	FNU_r1.1	NCBI	KDRI
421	*Fragaria orientalis*	Wild asian strawberry	FOR_r1.1	NCBI	KDRI
422	*Fragaria vesca subsp. vesca*	Woodland strawberry	FraVesHawaii_1.0	NCBI	Virginia Bioinformatics Institute
423	*Fragaria x ananassa*	Garden strawberry	FANhybrid_r1.2	NCBI	KDRI
424	*Morus notabilis*	Mulberry tree	ASM41409v2	NCBI	BGI
425	*Cannabis sativa*	Cannabis	canSat3	NCBI	University of Toronto
Fagales					
426	*Castanea mollissima*	Chinese chestnut	ASM76360v1	NCBI	Clemson University Genomics Institute
427	*Betula nana*	Dwarf birch	ASM32700v1	NCBI	QMUL
Cucurbitales					
428	*Cucumis melo*	Muskmelon	ASM31304v1	NCBI	Genetica Molecular, Centre de Recerca en Agrigenomica
429	*Cucumis sativus*	Cucumber	CucSat_1.0	NCBI	The Cucumber Genome Initiative
430	*Citrullus lanatus*	Watermelon	CiLa_1.0	NCBI	National Engineering Research Center for Vegetables, Beijing Academy of Agriculture and Forestry Sciences
431	*Lagenaria siceraria*	Calabash	Bottle_gourd	NCBI	Institute of Vegetables, Zhejiang Academy of Agricultural Sciences
Malpighiales					
432	*Populus euphratica*	Desert poplar	PopEup_1.0	NCBI	Lanzhou University
433	*Populus trichocarpa*	Black cottonwood	Poptr2_0	NCBI	JGI
434	*Jatropha curcas*	Physic nut	JatCur_1.0	NCBI	Chinese Academy of Sciences
435	*Manihot esculenta subsp. flabellifolia*	-	MW_v2d	NCBI	The Cassava Genome Consortium
436	*Ricinus communis*	Castor oil plant	JCVI_RCG_1.1	NCBI	JCVI
437	*Linum usitatissimum*	Flax	LinUsi_v1.1	NCBI	TUFGEN
Myrtales					
438	*Eucalyptus camaldulensis*	River red gum	EUC_r1.0	NCBI	KDRI
439	*Eucalyptus grandis*	Rose gum	Egrandis1_0	NCBI	Geneglob
Brassicales					
440	*Carica papaya*	Papaya	Papaya1.0	NCBI	The Papaya Genome Sequencing Consortium
441	*Arabidopsis halleri subsp. gemmifera*	-	Ahal_1.0	NCBI	TokyoTech

No	Systematic Name	Common Name	Genome Version	Source	Submitter
442	*Arabidopsis lyrata subsp. lyrata*	-	v.1.0	NCBI	JGI
443	*Arabidopsis thaliana*	-	TAIR10	NCBI	The Arabidopsis Information Resource
444	*Camelina sativa*	Camelina	Cs	NCBI	Agriculture & AgriFood Canada
445	*Capsella rubella*	Capsella	Caprub1_0	NCBI	JGI
446	*Brassica napus*	Rapeseed	Brassica_napus_ assembly_v1.0	NCBI	BGI
447	*Brassica oleracea var. oleracea*	-	BOL	NCBI	CanSeq
448	*Brassica rapa*	Field mustard	Brapa_1.0	NCBI	Brassica rapa Genome Sequencing Project, BraGSP
449	*Raphanus raphanistrum subsp. raphanistrum*	-	ASM76984v1	NCBI	Michigan State University
450	*Raphanus sativus*	Radish	ASM80110v1	NCBI	Myongji University
451	*Aethionema arabicum*	-	VEGI_AA_v_1.0	NCBI	McGill
452	*Arabis alpina*	Alpine rockcress	A_alpina_V4	NCBI	TRANSNET
453	*Eutrema parvulum*	-	Eutrema_parvulum_ v01	NCBI	University of Illinois
454	*Eutrema salsugineum*	Saltwater cress	Eutsalg1_0	NCBI	JGI
455	*Sisymbrium irio*	London rocket	VEGI_SI_v_1.0	NCBI	McGill
456	*Leavenworthia alabamica*	Alabama gladecress	VEGI_LA_v_1.0	NCBI	McGill
457	*Tarenaya hassleriana*	Pink queen	ASM46358v1	NCBI	BGI
Malvales					
458	*Gossypium arboreum*	Tree cotton	Gossypium_arboreum_ v1.0	NCBI	BGI
459	*Gossypium raimondii*	New world cotton	Graimondii2_0	NCBI	JGI
460	*Theobroma cacao*	Cacao	Theobroma_cacao_ 20110822	NCBI	Cacao Genome Consortium
461	*Aquilaria agallochum*	-	Aquilaria_agallocha_v1	NCBI	Academia Sinica
Sapindales					
462	*Azadirachta indica*	Indian lilac	AzaInd2.0	NCBI	Centre for Cellular and Molecular Plateforms
463	*Citrus clementine*	Clementine	Citrus_clementina_v1.0	NCBI	International Citrus Genome Consortium
464	*Citrus sinensis*	Orange	Csi_valencia_1.0	NCBI	China Sweet Orange Genome Project
Vitales					
465	*Vitis vinifera*	Common grape vine	12X	NCBI	International Grape Genome Program
Caryophyllales					

No	Systematic Name	Common Name	Genome Version	Source	Submitter
466	*Amaranthus hypochondriacus*	Prince's feather	AHP 1.0	NCBI	Institute of Bioinformatics and Applied Biotechnology
467	*Amaranthus tuberculatus*	Tall waterhemp	ASM18065v1	NCBI	University of Illinois
468	*Beta vulgaris subsp. vulgaris*	-	RefBeet-1.2.1	NCBI	Beta vulgaris Resource
469	*Spinacia oleracea*	Spinach	Viroflay-1.0.1	NCBI	BeetSeq
470	*Dianthus caryophyllus*	Clove pink	DCA_r1.0	NCBI	KDRI
Ericales					
471	*Actinidia chinensis*	Kiwi fruit	Kiwifruit_v1	NCBI	Boyce Thompson Institute
472	*Vaccinium macrocarpon*	Large cranberry	ASM77533v1	NCBI	Rutgers University
473	*Diospyros lotus*	Caucasian persimmon	ASM77412v1	NCBI	University of California Davis
474	*Primula veris*	Cowslip	ASM78844v1	NCBI	University of Zurich
Solanales					
475	*Solanum arcanum*	-	Soarc10	NCBI	PRI
476	*Solanum habrochaites*	-	Sohab10	NCBI	PRI
477	*Solanum lycopersicum*	Tomato	SL2.40	NCBI	Solanaceae Genomics Project
478	*Solanum melongena*	Eggplant	SME_r2.5.1	NCBI	KDRI
479	*Solanum pennellii*	-	Sopen10	NCBI	PRI
480	*Solanum pimpinellifolium*	Currant tomato	Sol_pimpi_v1.0	NCBI	Sol Genomics Network
481	*Solanum tuberosum*	Potato	SolTub_3.0	NCBI	Potato Genome Sequencing Consortium
482	*Capsicum annuum*	Hot pepper	PGAv.1.5	NCBI	SNU
483	*Nicotiana sylvestris*	Woodland tobacco	Nsyl	NCBI	Philip Morris International R&D
484	*Nicotiana tomentosiformis*	-	Ntom_v01	NCBI	Philip Morris International R&D
Lamiales					
485	*Fraxinus excelsior*	European ash	BATG-0.4	NCBI	QMUL
486	*Penstemon centranthifolius*	Scarlet bugler	ASM73743v1	NCBI	Ohio State University
487	*Penstemon grinnellii*	Grinnell's beardtongue	ASM73742v1	NCBI	Ohio State University
488	*Sesamum indicum*	Sesame	S_indicum_v1.0	NCBI	Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences
489	*Genlisea aurea*	Corkscrew plant	GenAur_1.0	NCBI	Lomonosov Moscow State University, Evolutionary Genomics Laboratory

No	Systematic Name	Common Name	Genome Version	Source	Submitter
490	*Mimulus guttatus*	Common monkey-flower	Mimgu1_0	NCBI	JGI
Asterales					
491	*Conyza canadensis*	Horseweed	ASM77593v1	NCBI	University of Tennessee
ECHINOIDEA					
Temnopleuroida					
492	*Lytechinus variegatus*	Green sea urchin	Lvar_0.4	NCBI	BCM
Echinoida					
493	*Strongylocentrotus purpuratus*	Purple sea urchin	Spur_3.1	NCBI	BCM
ASTEROIDEA					
Valvatida					
494	*Patiria miniata*	Bat star	Pmin_1.0	NCBI	Sea Urchin Genome Sequencing Consortium
ENTEROPNEUSTA					
495	*Saccoglossus kowalevskii*	Acorn worm	Skow_1.1	NCBI	BCM
TUNICATA					
Enterogona					
496	*Ciona intestinalis*	Sea squirt	KH	NCBI	Organization: Department of Zoology, Graduate School
497	*Ciona savignyi*	-	ASM14926v1	NCBI	Broad
Pleurogona					
498	*Botryllus schlosseri*	Star ascidian	356a-chromosome-assembly	NCBI	Stanford University
Copelata					
499	*Oikopleura dioica*	Larvaceans	ASM20953v1	NCBI	Genoscope CEA
LEPTOCARDII					
Amphioxiformes					
500	*Branchiostoma floridae*	Lancelet	Version 2	NCBI	JGI
CEPHALASPIDOMORPHI					
Petromyzontiformes					
501	*Lethenteron camtschaticum*	Arctic lamprey	LetJap1.0	NCBI	IMCB
502	*Petromyzon marinus*	Sea lamprey	petMar2	UCSC	WashU
SARCOPTERYGII					
Coelacanthiformes					
503	*Latimeria chalumnae*	Coelacanth	latCha1	UCSC	Broad

Figure 1: Phylogenetic representation of the genomic dataset

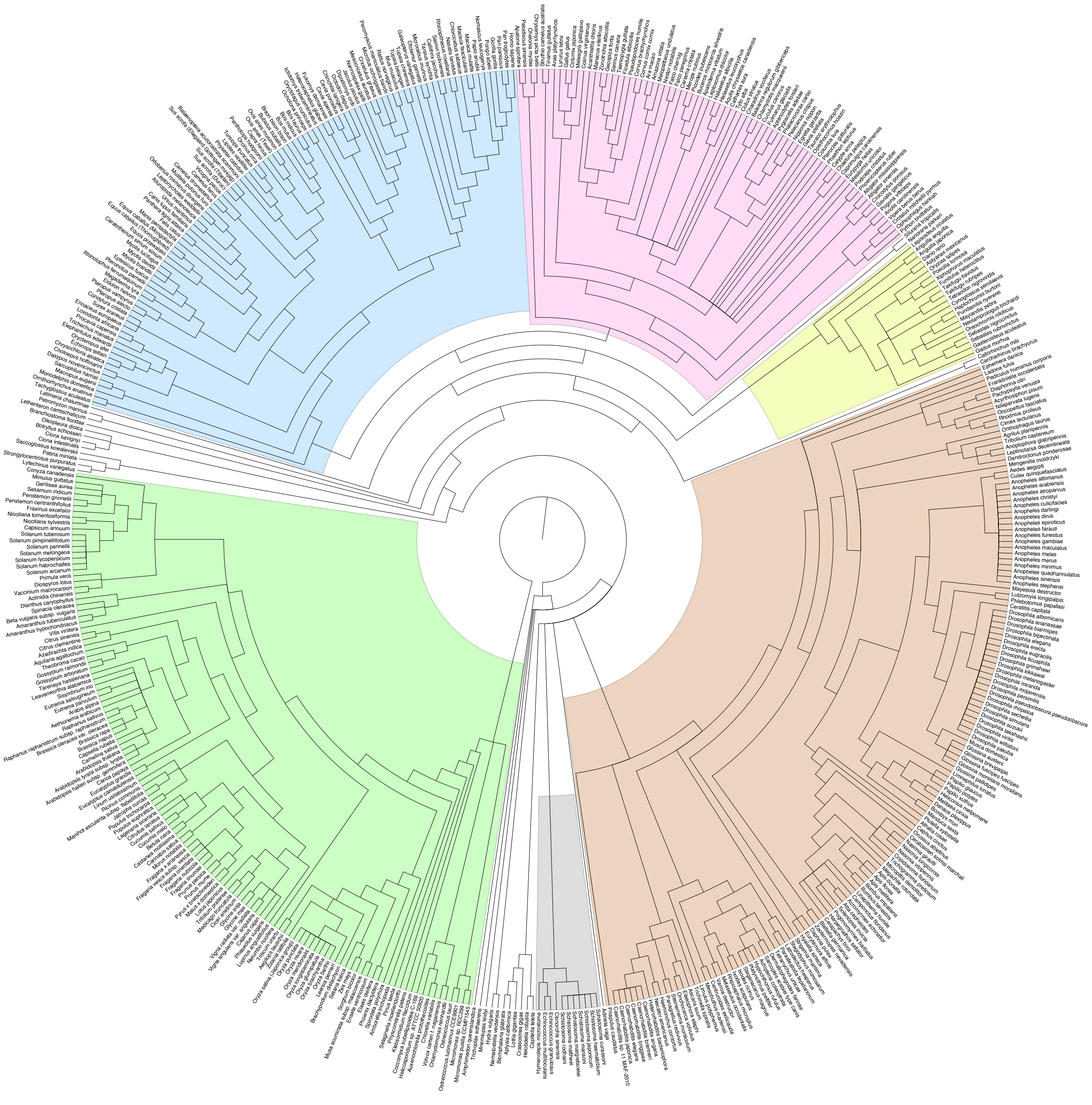


Figure B.1: **Inferred tree of life:** Phylogenetic inference of the genomic dataset representing the eukaryotic tree of life. This tree was built using Archaeopteryx to download the Tree of Life (tolweb.org) topology for all Eukaryota (node identifier 3, about 76,000 species). The tree was extended to include required descendant species, extract the 503 species of interest, and update ambiguous branches based on the most recent literature.

Table 2: Genome quality check - assembly statistics

Table B.2: **Assembly statistics:** Shows the systematic name, total sequence length (i.e. genome size, including bases and gaps), scaffold N50 (i.e. scaffold length at which 50% of the total bases in the assembly are in scaffolds of that length or greater), contig N50 and assembly level (complete genome, chromosome, scaffold or contig). Species are listed in the same order as Table B.1. Statistics for the publicly available genomes can be found on NCBI (www.ncbi.nlm.nih.gov/ → Assembly → look up the genome of interest → GenBank FTP site → *_assembly_stats.txt file).

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
MAMMALIA					
1	*Tachyglossus aculeatus*	-	-	-	Contig
2	*Ornithorhynchus anatinus*	1,995,607,322	958,970	11,554	Chromosome
3	*Monodelphis domestica*	3,598,443,077	59,809,810	108,014	Chromosome
4	*Macropus eugenii*	3,075,184,024	36,602	2,602	Scaffold
5	*Sarcophilus harrisii*	3,174,693,010	1,847,106	20,139	Scaffold
6	*Dasypus novemcinctus*	3,631,522,711	1,687,935	26,277	Scaffold
7	*Choloepus hoffmanni*	2,458,927,620	9,667	2,343	Scaffold
8	*Chrysochloris asiatica*	4,210,110,458	13,470,186	19,631	Scaffold
9	*Echinops telfairi*	2,947,024,286	45,764,842	20,428	Scaffold
10	*Orycteropus afer afer*	4,444,080,527	7,875,832	17,683	Scaffold
11	*Elephantulus edwardii*	3,843,982,861	15,011,382	24,219	Scaffold
12	*Trichechus manatus latirostris*	3,103,808,406	14,442,683	37,750	Scaffold
13	*Procavia capensis*	2,985,258,999	24,297	3,379	Scaffold
14	*Loxodonta africana*	-	-	-	Chromosome
15	*Erinaceus europaeus*	2,715,720,925	3,264,618	21,359	Scaffold
16	*Sorex araneus*	2,423,158,183	22,794,405	22,623	Scaffold
17	*Condylura cristata*	1,769,662,895	55,520,359	46,163	Scaffold
18	*Pteropus alecto*	1,985,975,446	15,954,802	31,841	Scaffold
19	*Pteropus vampyrus*	1,996,076,410	124,060	8,527	Scaffold
20	*Eidolon helvum*	1,837,754,460	27,684	12,668	Scaffold
21	*Megaderma lyra*	1,735,931,796	16,881	7,043	Scaffold
22	*Rhinolophus ferrumequinum*	1,926,439,238	21,151	11,659	Scaffold
23	*Pteronotus parnellii*	1,960,317,893	22,675	9,502	Scaffold
24	*Eptesicus fuscus*	2,026,629,342	13,454,942	21,392	Scaffold
25	*Myotis brandtii*	2,107,242,811	3,225,832	23,289	Scaffold
26	*Myotis davidii*	2,059,799,708	3,454,484	15,182	Scaffold
27	*Myotis lucifugus*	2,034,575,300	4,293,315	64,330	Scaffold
28	*Ceratotherium simum simum*	2,464,367,180	26,277,727	92,960	Scaffold
29	*Equus przewalskii*	2,395,954,271	513,800	57,610	Scaffold
30	*Equus caballus* (Thoroughbred)	2,474,929,062	46,749,900	112,381	Chromosome
31	*Equus caballus* (Mongolian)	2,377,506,909	56,256	40,738	Scaffold
32	*Manis pentadactyla*	2,204,732,179	117,920	28,718	Scaffold

No	**Species**	**Total Sequence Length**	**Scaffold N50**	**Contig N50**	**Assembly Level**
33	*Felis catus*	2,641,342,258	18,072,971	45,189	Chromosome
34	*Panthera tigris altaica*	2,391,082,183	8,860,407	30,032	Scaffold
35	*Canis lupus familiaris*	2,410,976,875	45,876,610	267,478	Chromosome
36	*Ursus maritimus*	2,301,379,344	15,940,661	46,506	Scaffold
37	*Ailuropoda melanoleuca*	2,299,492,210	1,281,781	39,886	Scaffold
38	*Leptonychotes weddellii*	3,156,902,762	904,031	23,664	Scaffold
39	*Odobenus rosmarus divergens*	2,400,150,193	2,616,778	89,951	Scaffold
40	*Mustela putorius furo*	2,410,879,678	9,335,154	44,823	Scaffold
41	*Camelus dromedarius*	2,004,063,690	4,188,677	69,131	Scaffold
42	*Camelus ferus*	2,009,194,609	2,005,940	90,263	Scaffold
43	*Vicugna pacos*	2,172,191,320	7,263,804	24,617	Scaffold
44	*Sus scrofa* (Duroc)	2,808,525,991	576,008	69,474	Chromosome
45	*Sus scrofa* (Tibetan)	2,489,155,924	1,054,539	45,705	Scaffold
46	*Sus scrofa* (Ellegaard Gottingen minipig)	2,358,017,222	-	22,008	Contig
47	*Balaenoptera acutorostrata scammoni*	2,431,687,698	12,843,668	22,690	Scaffold
48	*Physeter catodon*	2,280,727,784	427,290	35,257	Scaffold
49	*Lipotes vexillifer*	2,429,212,129	2,419,148	31,902	Scaffold
50	*Tursiops truncatus*	2,551,418,184	116,287	11,821	Scaffold
51	*Orcinus orca*	2,372,919,875	12,735,091	70,300	Scaffold
52	*Pantholops hodgsonii*	2,696,886,330	2,772,860	18,674	Scaffold
53	*Capra hircus*	2,635,848,900	14,391,519	18,934	Chromosome
54	*Ovis aries* (Texel)	2,619,054,388	100,079,507	40,376	Chromosome
55	*Ovis aries musimon*	2,589,834,840	2,217,029	39,721	Scaffold
56	*Bubalus bubalis*	2,836,166,969	1,412,388	21,938	Scaffold
57	*Bison bison bison*	2,828,031,685	7,192,658	19,971	Scaffold
58	*Bos mutus*	2,645,161,911	1,407,960	22,822	Scaffold
59	*Bos indicus*	2,673,949,103	-	28,375	Chromosome
60	*Bos taurus*	2,670,422,299	6,380,747	96,951	Chromosome
61	*Ochotona princeps*	2,229,840,584	26,863,993	42,347	Scaffold
62	*Oryctolagus cuniculus*	2,737,445,565	35,972,871	64,648	Chromosome
63	*Ictidomys tridecemlineatus*	2,478,393,770	8,192,786	44,137	Scaffold
64	*Heterocephalus glaber*	2,618,204,639	20,532,749	47,778	Scaffold
65	*Fukomys damarensis*	2,333,892,479	5,314,287	44,830	Scaffold
66	*Cavia aperea*	2,716,396,567	-	1,039	Scaffold
67	*Cavia porcellus*	2,723,219,641	27,942,054	80,583	Scaffold
68	*Chinchilla lanigera*	2,390,868,971	21,893,125	61,105	Scaffold
69	*Octodon degus*	2,995,889,303	12,091,372	19,847	Scaffold
70	*Dipodomys ordii*	2,158,502,098	36,427	4,288	Scaffold
71	*Jaculus jaculus*	2,835,250,225	22,080,993	15,675	Scaffold
72	*Nannospalax galili*	3,061,424,619	3,618,479	30,353	Scaffold

No	**Species**	**Total Sequence Length**	**Scaffold N50**	**Contig N50**	**Assembly Level**
73	*Mesocricetus auratus*	2,504,925,039	12,753,307	22,511	Scaffold
74	*Cricetulus griseus*	2,399,786,748	1,147,233	39,361	Scaffold
75	*Microtus ochrogaster*	2,287,340,943	17,270,019	21,250	Chromosome
76	*Peromyscus maniculatus bairdii*	2,630,541,020	3,760,915	36,367	Scaffold
77	*Rattus norvegicus*	2,909,698,938	2,178,346	52,491	Chromosome
78	*Mus musculus*	2,798,785,524	52,589,046	32,273,079	Chromosome
79	*Tupaia belangeri*	2,137,225,476	-	2,974	Contig
80	*Tupaia chinensis*	2,846,580,235	3,670,124	25,938	Scaffold
81	*Galeopterus variegatus*	3,187,660,572	245,189	20,752	Scaffold
82	*Otolemur garnettii*	2,519,724,550	13,852,661	27,100	Scaffold
83	*Microcebus murinus*	2,902,270,736	140,884	3,511	Scaffold
84	*Tarsius syrichta*	3,453,864,774	401,181	38,165	Scaffold
85	*Callithrix jacchus*	2,914,958,544	5,167,444	29,273	Chromosome
86	*Saimiri boliviensis boliviensis*	2,608,588,537	18,744,880	38,823	Scaffold
87	*Rhinopithecus roxellana*	2,899,552,139	1,549,224	77,151	Scaffold
88	*Nasalis larvatus*	3,011,966,170	-	13,295	Chromosome
89	*Chlorocebus sabaeus*	2,789,656,328	81,825,804	90,449	Chromosome
90	*Macaca fascicularis*	2,946,843,737	88,649,475	86,040	Chromosome
91	*Macaca mulatta*	2,969,971,616	1,660,975	12,841	Chromosome
92	*Papio anubis*	2,948,397,226	528,927	40,262	Chromosome
93	*Nomascus leucogenys*	2,962,060,179	52,956,880	35,148	Chromosome
94	*Pongo abelii*	3,441,244,233	747,460	15,648	Chromosome
95	*Gorilla gorilla gorilla*	3,035,660,144	913,458	11,661	Chromosome
96	*Pan paniscus*	2,869,190,071	10,124,892	66,749	Chromosome
97	*Pan troglodytes*	3,323,267,922	8,925,874	50,656	Chromosome
98	*Homo sapiens*	3,137,144,693	46,395,641	38,508,932	Chromosome
SAUROPSIDA					
99	*Apalone spinifera*	1,931,078,847	2,306,994	4,693	Scaffold
100	*Pelodiscus sinensis*	2,202,483,752	3,350,749	21,993	Scaffold
101	*Chelonia mydas*	2,208,410,377	3,864,108	29,240	Scaffold
102	*Chrysemys picta bellii*	2,589,745,704	5,212,367	11,852	Chromosome
103	*Struthio camelus australis*	1,225,041,896	3,593,425	34,997	Scaffold
104	*Tinamus guttatus*	1,047,056,493	246,268	29,773	Scaffold
105	*Anas platyrhynchos*	1,105,052,351	1,233,631	26,114	Scaffold
106	*Lyrurus tetrix tetrix*	657,025,294	-	1,029	Contig
107	*Gallus gallus*	1,046,932,099	12,877,381	279,750	Chromosome
108	*Coturnix japonica*	531,959,683	108,813	1,473	Scaffold
109	*Meleagris gallopavo*	1,061,817,103	857,645	12,520	Chromosome
110	*Colinus virginianus*	1,171,855,925	-	6,061	Scaffold
111	*Acanthisitta chloris*	1,035,876,403	64,469	20,602	Scaffold
112	*Manacus vitellinus*	1,145,871,783	2,558,866	43,697	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
113	*Zonotrichia albicollis*	1,052,600,561	4,866,725	112,748	Scaffold
114	*Geospiza fortis*	1,065,292,181	5,255,844	30,521	Scaffold
115	*Serinus canaria*	1,152,100,110	17,815,079	53,884	Scaffold
116	*Taeniopygia guttata*	1,232,135,591	8,236,790	38,639	Chromosome
117	*Ficedula albicollis*	1,118,343,587	6,542,656	410,964	Chromosome
118	*Pseudopodoces humilis*	1,042,997,632	16,337,386	165,265	Scaffold
119	*Corvus brachyrhynchos*	1,091,312,783	6,953,989	29,093	Scaffold
120	*Corvus cornix cornix*	1,049,964,851	16,358,221	94,375	Scaffold
121	*Ara macao*	1,204,700,227	15,974	4,399	Scaffold
122	*Amazona vittata*	1,175,404,042	19,239	6,904	Scaffold
123	*Melopsittacus undulatus*	1,117,373,619	10,614,383	55,633	Scaffold
124	*Nestor notabilis*	1,053,559,886	61,475	26,546	Scaffold
125	*Falco cherrug*	1,174,811,715	4,154,532	31,327	Scaffold
126	*Falco peregrinus*	1,171,973,431	3,935,757	28,645	Scaffold
127	*Cariama cristata*	1,132,245,425	55,197	24,645	Scaffold
128	*Merops nubicus*	1,062,961,556	48,089	24,675	Scaffold
129	*Picoides pubescens*	1,167,323,935	2,093,929	24,809	Scaffold
130	*Buceros rhinoceros silvestris*	1,065,782,791	53,203	14,587	Scaffold
131	*Apaloderma vittatum*	1,070,836,417	56,673	28,882	Scaffold
132	*Leptosomus discolour*	1,136,244,952	62,640	24,735	Scaffold
133	*Haliaeetus albicilla*	1,133,549,865	57,319	25,143	Scaffold
134	*Haliaeetus leucocephalus*	1,178,409,481	9,145,499	105,493	Scaffold
135	*Aquila chrysaetos Canadensis*	1,192,743,076	9,230,743	172,329	Scaffold
136	*Cathartes aura*	1,152,571,117	36,359	15,248	Scaffold
137	*Tyto alba*	1,120,143,088	52,818	17,226	Scaffold
138	*Colius striatus*	1,075,931,597	46,063	25,860	Scaffold
139	*Charadrius vociferus*	1,219,859,583	3,657,050	39,278	Scaffold
140	*Balearica regulorum gibbericeps*	1,127,622,302	52,178	23,331	Scaffold
141	*Chlamydotis macqueenii*	1,086,566,339	45,221	21,641	Scaffold
142	*Cuculus canorus*	1,153,894,225	2,989,832	38,137	Scaffold
143	*Fulmarus glacialis*	1,141,395,646	47,208	25,926	Scaffold
144	*Aptenodytes forsteri*	1,254,347,440	5,071,598	31,730	Scaffold
145	*Pygoscelis adeliae*	1,216,617,519	5,118,896	22,195	Scaffold
146	*Phalacrocorax carbo*	1,138,967,842	48,427	17,343	Scaffold
147	*Pelecanus crispus*	1,160,924,693	43,364	21,679	Scaffold
148	*Nipponia nippon*	1,223,863,029	5,211,696	29,116	Scaffold
149	*Egretta garzetta*	1,206,501,934	3,067,157	29,019	Scaffold
150	*Phaethon lepturus*	1,152,958,507	47,896	22,941	Scaffold
151	*Gavia stellata*	1,129,694,867	45,523	24,321	Scaffold
152	*Tauraco erythrolophus*	1,155,540,733	56,334	22,885	Scaffold
153	*Opisthocomus hoazin*	1,203,712,246	2,937,227	28,179	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
154	*Columba livia*	1,107,989,085	3,148,738	26,579	Scaffold
155	*Pterocles gutturalis*	1,069,324,295	49,530	26,448	Scaffold
156	*Calypte anna*	1,105,676,412	4,052,191	26,738	Scaffold
157	*Chaetura pelagica*	1,119,188,094	3,841,852	30,757	Scaffold
158	*Caprimulgus carolinensis*	1,119,683,066	46,345	22,156	Scaffold
159	*Eurypyga helias*	1,088,019,637	47,243	24,402	Scaffold
160	*Mesitornis unicolor*	1,087,290,853	47,102	22,740	Scaffold
161	*Podiceps cristatus*	1,134,922,578	30,087	17,412	Scaffold
162	*Phoenicopterus ruber ruber*	1,132,184,511	38,071	38,071	Scaffold
163	*Alligator mississippiensis*	2,174,259,888	508,966	36,484	Scaffold
164	*Alligator sinensis*	2,270,567,745	2,188,296	23,412	Scaffold
165	*Crocodylus porosus*	2,120,573,303	204,532	32,735	Scaffold
166	*Gavialis gangeticus*	2,188,353,730	100,817	14,195	Scaffold
167	*Pogona vitticeps*	-	-	-	Contig?
168	*Anolis carolinensis*	1,799,143,587	4,033,265	79,867	Chromosome
169	*Vipera berus berus*	1,532,390,814	126,452	11,725	Scaffold
170	*Crotalus mitchellii pyrrhus*	1,126,791,715	-	4,103	Scaffold
171	*Ophiophagus hannah*	1,594,074,654	-	5,201	Scaffold
172	*Python bivittatus*	1,435,052,152	213,970	10,658	Scaffold
AMPHIBIA					
173	*Nanorana parkeri*	2,053,849,526	1,069,101	32,798	Scaffold
174	*Xenopus tropicalis*	1,437,530,879	124,127,367	72,201	Scaffold
NEOPTERYGII					
175	*Lepisosteus oculatus*	945,878,036	6,928,108	68,337	Chromosome
176	*Anguilla anguilla*	1,018,701,900	-	2,544	Scaffold
177	*Anguilla japonica*	1,151,137,423	52,849	3,215	Scaffold
178	*Danio rerio*	1,371,719,383	2,181,225	1,258,148	Chromosome
179	*Astyanax mexicanus*	1,191,242,572	1,775,308	14,739	Scaffold
180	*Oryzias latipes*	869,818,208	6,078,048	9,628	Chromosome
181	*Poecilia formosa*	748,923,461	1,574,226	57,472	Scaffold
182	*Xiphophorus maculatus*	729,664,433	1,303,070	22,273	Scaffold
183	*Fundulus heteroclitus*	1,021,898,560	1,252,252	16,688	Scaffold
184	*Takifugu flavidus*	378,032,400	315,240	8,233	Scaffold
185	*Takifugu rubripes*	391,484,725	928,938	52,883	Chromosome
186	*Tetraodon nigroviridis*	342,403,326	-	29,054	Contig
187	*Cynoglossus semilaevis*	470,199,494	509,861	27,008	Chromosome
188	*Haplochromis burtoni*	831,411,547	1,194,190	21,886	Scaffold
189	*Pundamilia nyererei*	830,133,247	2,525,540	22,622	Scaffold
190	*Maylandia zebra*	849,595,388	3,702,874	19,949	Scaffold
191	*Neolamprologus brichardi*	847,910,432	4,430,025	13,047	Scaffold
192	*Oreochromis niloticus*	927,696,114	2,766,223	29,493	Chromosome
193	*Sebastes nigrocinctus*	687,549,873	12,083	8,171	Scaffold
194	*Sebastes rubrivinctus*	756,296,653	30,046	13,541	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
195	*Gasterosteus aculeatus*	446,610,788	-	83,204	Contig
196	*Gadus morhua*	824,311,139	393,166	2,311	Scaffold
CHONDRICHTHYES					
197	*Callorhinchus milii*	974,498,586	4,521,921	46,577	Scaffold
198	*Carcharhinus brachyurus*	-	-	-	Contig?
ECDYSOZOA					
199	*Ephemera danica*	475,911,277	527,117	12,200	Scaffold
200	*Ladona fulva*	1,158,111,285	978,724	16,482	Scaffold
201	*Pediculus humanus corporis*	110,781,312	497,057	34,097	Scaffold
202	*Frankliniella occidentalis*	415,781,860	948,890	6,159	Scaffold
203	*Diaphorina citri*	485,705,082	109,898	34,407	Scaffold
204	*Pachypsylla venusta*	701,763,251	157,704	1,853	Scaffold
205	*Acyrthosiphon pisum*	541,692,442	518,546	28,199	Scaffold
206	*Nilaparvata lugens*	1,140,786,310	356,597	23,518	Scaffold
207	*Oncopeltus fasciatus*	1,098,671,531	339,960	4,047	Scaffold
208	*Rhodnius prolixus*	702,645,054	847,873	27,483	Scaffold
209	*Cimex lectularius*	650,477,627	7,172,596	23,511	Scaffold
210	*Onthophagus taurus*	270,533,845	251,443	15,427	Scaffold
211	*Agrilus planipennis*	353,546,665	910,924	6,314	Scaffold
212	*Tribolium castaneum*	210,264,614	975,455	43,511	Chromosome
213	*Anoplophora glabripennis*	707,712,193	658,851	16,551	Scaffold
214	*Leptinotarsa decemlineata*	1,170,241,964	414,170	4,913	Scaffold
215	*Dendroctonus ponderosae*	252,847,629	628,732	7,451	Scaffold
216	*Mengenilla moldrzyki*	155,727,465	-	4,038	Contig
217	*Aedes aegypti*	1,383,957,531	1,547,048	82,618	Scaffold
218	*Culex quinquefasciatus*	579,042,118	486,756	28,546	Scaffold
219	*Anopheles albimanus*	170,508,315	18,068,499	200,121	Scaffold
220	*Anopheles arabiensis*	246,567,867	5,604,218	74,117	Scaffold
221	*Anopheles atroparvus*	224,290,125	9,206,694	53,613	Scaffold
222	*Anopheles christyi*	172,658,580	9,057	8,461	Scaffold
223	*Anopheles culicifacies*	202,998,806	22,320	13,414	Scaffold
224	*Anopheles darlingi*	136,935,538	-	115,072	Contig
225	*Anopheles dirus*	216,307,690	6,906,475	76,300	Scaffold
226	*Anopheles epiroticus*	223,486,714	366,526	88,996	Scaffold
227	*Anopheles farauti*	183,103,254	12,895,223	161,110	Scaffold
228	*Anopheles funestus*	225,223,604	671,960	60,888	Scaffold
229	*Anopheles gambiae*	265,027,044	12,309,988	85,548	Chromosome
230	*Anopheles maculatus*	141,894,015	3,841	2,984	Scaffold
231	*Anopheles melas*	224,162,116	18,103	11,312	Scaffold
232	*Anopheles merus*	288,048,996	1,489,982	48,115	Scaffold
233	*Anopheles minimus*	201,793,324	10,313,149	123,292	Scaffold
234	*Anopheles quadriannulatus*	283,828,998	1,641,272	39,415	Scaffold
235	*Anopheles sinensis*	220,777,669	814,231	30,137	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
236	*Anopheles stephensi*	221,324,304	1,591,355	36,511	Scaffold
237	*Mayetiola destructor*	185,827,756	756,041	14,032	Scaffold
238	*Lutzomyia longipalpis*	154,229,266	85,093	7,481	Scaffold
239	*Phlebotomus papatasi*	363,767,980	27,956	5,795	Scaffold
240	*Ceratitis capitata*	484,789,472	4,060,615	45,711	Scaffold
241	*Drosophila albomicans*	253,560,284	23,589	3,826	Scaffold
242	*Drosophila ananassae*	230,993,012	4,599,533	95,247	Scaffold
243	*Drosophila biarmipes*	169,378,599	3,386,121	474,639	Scaffold
244	*Drosophila bipectinata*	167,263,958	663,995	149,088	Scaffold
245	*Drosophila elegans*	171,267,669	1,714,184	212,818	Scaffold
246	*Drosophila erecta*	152,712,140	18,748,788	453,616	Scaffold
247	*Drosophila eugracilis*	156,942,009	976,885	224,458	Scaffold
248	*Drosophila ficusphila*	152,439,475	1,050,541	275,894	Scaffold
249	*Drosophila grimshawi*	200,467,819	8,399,593	91,192	Scaffold
250	*Drosophila kikkawai*	164,292,578	903,682	209,056	Scaffold
251	*Drosophila melanogaster*	143,726,002	-	21,485,538	Chromosome
252	*Drosophila miranda*	136,728,780	-	77,651	Chromosome
253	*Drosophila mojavensis*	193,826,310	24,764,193	121,517	Scaffold
254	*Drosophila persimilis*	188,374,079	1,869,541	20,311	Scaffold
255	*Drosophila pseudoobscura pseudoobscura*	152,696,384	12,541,198	202,541	Chromosome
256	*Drosophila rhopaloa*	197,375,704	45,514	19,484	Scaffold
257	*Drosophila sechellia*	166,592,095	2,123,299	42,955	Scaffold
258	*Drosophila simulans*	137,843,219	857,818	15,131	Chromosome
259	*Drosophila suzukii*	232,923,092	388,966	24,954	Scaffold
260	*Drosophila takahashii*	182,106,768	387,676	126,259	Scaffold
261	*Drosophila virilis*	206,026,697	10,161,210	120,091	Scaffold
262	*Drosophila willistoni*	235,516,348	4,511,350	180,217	Scaffold
263	*Drosophila yakuba*	165,709,965	21,770,863	115,562	Chromosome
264	*Musca domestica*	750,403,944	226,573	11,807	Scaffold
265	*Glossina austeni*	370,264,922	812,585	45,924	Scaffold
266	*Glossina brevipalpis*	315,360,362	1,209,507	62,009	Scaffold
267	*Glossina fuscipes fuscipes*	374,774,708	561,190	64,354	Scaffold
268	*Glossina morsitans morsitans*	363,107,242	-	49,769	Contig
269	*Glossina pallidipes*	357,332,231	1,038,751	167,200	Scaffold
270	*Limnephilus lunatus*	1,345,859,447	54,650	2,103	Scaffold
271	*Papilio glaucus*	374,815,656	-	12,225	Scaffold
272	*Papilio polytes*	227,021,014	3,672,263	47,768	Scaffold
273	*Papilio xuthus*	243,890,167	6,198,915	128,246	Scaffold
274	*Heliconius melpomene melpomene*	273,786,188	194,302	51,027	Scaffold
275	*Melitaea cinxia*	389,907,520	119,328	13,489	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
276	*Danaus plexippus*	272,853,388	52,968	50,721	Scaffold
277	*Bombyx mori*	481,819,406	4,008,358	15,510	Scaffold
278	*Manduca sexta*	419,424,057	664,006	40,396	Scaffold
279	*Plutella xylostella*	393,470,562	737,182	18,517	Scaffold
280	*Athalia rosae*	163,837,890	1,366,867	51,418	Scaffold
281	*Cephus cinctus*	162,226,239	622,163	44,905	Scaffold
282	*Orussus abietinus*	201,220,334	2,372,050	54,038	Scaffold
283	*Ceratosolen solmsi marchali*	277,059,195	9,558,897	74,702	Scaffold
284	*Nasonia giraulti*	283,606,953	759,431	1,971	Scaffold
285	*Nasonia longicornis*	285,726,340	758,407	1,876	Scaffold
286	*Nasonia vitripennis*	295,780,872	708,988	18,840	Chromosome
287	*Copidosoma floridanum*	555,044,930	1,037,125	14,521	Scaffold
288	*Trichogramma pretiosum*	195,087,592	3,706,225	78,655	Scaffold
289	*Microplitis demolitor*	250,525,215	318,766	13,540	Scaffold
290	*Megachile rotundata*	272,660,569	1,699,680	64,153	Scaffold
291	*Apis dorsata*	230,340,171	732,052	8,422	Scaffold
292	*Apis florea*	230,467,781	2,863,240	24,704	Scaffold
293	*Apis mellifera*	250,287,000	997,192	45,688	Chromosome
294	*Bombus impatiens*	249,185,056	1,399,493	58,885	Scaffold
295	*Bombus terrestris*	248,654,244	3,506,793	76,043	Chromosome
296	*Linepithema humile*	219,500,750	1,402,257	35,858	Scaffold
297	*Camponotus floridanus*	232,685,334	451,320	19,487	Scaffold
298	*Acromyrmex echinatior*	295,944,863	1,110,580	80,630	Scaffold
299	*Atta cephalotes*	317,671,980	5,154,485	14,798	Scaffold
300	*Solenopsis invicta*	396,024,718	558,018	14,677	Scaffold
301	*Pogonomyrmex barbatus*	235,645,958	819,605	11,605	Scaffold
302	*Harpegnathos saltator*	294,465,601	601,965	38,321	Scaffold
303	*Cerapachys biroi*	212,825,769	1,350,650	34,211	Scaffold
304	*Blattella germanica*	2,037,201,033	1,056,071	12,125	Scaffold
305	*Zootermopsis nevadensis*	485,009,472	751,105	22,794	Scaffold
306	*Daphnia pulex*	197,206,209	642,089	49,250	Scaffold
307	*Eurytemora affinis*	494,890,867	862,645	5,738	Scaffold
308	*Hyalella azteca*	1,178,848,281	987,977	5,445	Scaffold
309	*Strigamia maritima*	176,210,797	139,451	24,745	Scaffold
310	*Stegodyphus mimosarum*	2,738,704,917	480,636	40,146	Scaffold
311	*Latrodectus hesperus*	1,137,104,656	13,889	2,223	Scaffold
312	*Parasteatoda tepidariorum*	1,443,909,906	465,572	10,149	Scaffold
313	*Tetranychus urticae*	90,815,494	2,993,488	212,780	Scaffold
314	*Dermatophagoides farinae*	53,545,338	186,342	8,538	Scaffold
315	*Sarcoptes scabiei type canis*	56,262,437	-	11,197	Scaffold
316	*Achipteria coleoptrata*	88,443,889	-	3,360	Scaffold
317	*Hypochthonius rufulus*	172,365,132	-	3,126	Scaffold
318	*Platynothrus peltifer*	100,532,934	-	1,236	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
319	*Steganacarus magnus*	113,565,242	-	1,617	Scaffold
320	*Ixodes ricinus*	391,986,416	-	1,638	Scaffold
321	*Ixodes scapularis*	1,765,382,190	76,228	2,942	Scaffold
322	*Rhipicephalus microplus*	144,692,238	-	827	Contig
323	*Metaseiulus occidentalis*	151,724,311	896,831	200,706	Scaffold
324	*Varroa destructor*	294,133,941	-	2,262	Contig
325	*Centruroides exilicauda*	926,399,729	342,549	5,173	Scaffold
326	*Mesobuthus martensii*	925,546,267	-	45,228	Contig
327	*Limulus polyphemus*	1,828,271,751	254,089	11,441	Scaffold
328	*Trichinella spiralis*	63,525,422	6,373,445	76,707	Scaffold
329	*Ascaris suum*	269,559,681	413,062	22,226	Scaffold
330	*Elaeophora elaphi*	1,480,825	1,480,825	99,433	Scaffold
331	*Onchocerca volvulus*	96,427,779	25,485,961	577,378	Scaffold
332	*Steinernema monticolum*	89,158,954	11,563	2,456	Scaffold
333	*Panagrellus redivivus*	65,093,147	262,414	19,707	Scaffold
334	*Haemonchus contortus*	319,757,902	99,131	23,747	Scaffold
335	*Necator americanus*	244,075,060	211,861	5,429	Scaffold
336	*Heterorhabditis bacteriophora*	76,974,349	312,328	101,980	Scaffold
337	*Caenorhabditis angaria*	79,761,545	9,453	1,196	Scaffold
338	*Caenorhabditis brenneri*	190,369,721	381,961	31,887	Scaffold
339	*Caenorhabditis briggsae*	108,478,526	14,512,975	45,155	Chromosome
340	*Caenorhabditis elegans*	100,286,401	-	17,493,829	Chromosome
341	*Caenorhabditis japonica*	166,256,191	94,149	10,166	Scaffold
342	*Caenorhabditis sp. 11 MAF-2010*	79,321,433	20,921,866	22,519	Scaffold
343	*Priapulus caudatus*	420,178,403	126,878	4,487	Scaffold
ROTIFERA					
344	*Adineta vaga*	217,933,776	260,259	94,665	Scaffold
PLATYHELMINTHES					
345	*Schistosoma curassoni*	344,202,215	-	7,099	Scaffold
346	*Schistosoma haematobium*	375,894,156	317,484	22,446	Scaffold
347	*Schistosoma japonicum*	402,743,189	176,869	6,121	Scaffold
348	*Schistosoma mansoni*	364,538,298	-	76,721	Chromosome
349	*Schistosoma margrebowiei*	367,396,528	-	28,534	Scaffold
350	*Schistosoma mattheei*	340,818,150	-	4,667	Scaffold
351	*Schistosoma rodhaini*	343,294,136	-	2,714	Scaffold
352	*Clonorchis sinensis*	547,288,241	417,486	233,037	Scaffold
353	*Echinococcus granulosus*	112,350,571	6,536,549	219,470	Scaffold
354	*Echinococcus multilocularis*	7,371,474	277,614	275,497	Scaffold
355	*Hymenolepis microstoma*	126,773,784	651,751	78,513	Scaffold
ANNELIDA					
356	*Capitella teleta*	333,283,208	188,402	21,930	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
357	*Helobdella robusta*	235,376,169	3,060,193	52,195	Scaffold
MOLLUSCA					
358	*Crassostrea gigas*	557,735,934	401,685	31,239	Scaffold
359	*Lottia gigantea*	359,505,668	1,870,055	96,027	Scaffold
360	*Aplysia californica*	927,310,431	917,541	9,586	Scaffold
361	*Biomphalaria glabrata*	916,388,084	48,059	7,298	Scaffold
CNIDARIA					
362	*Nematostella vectensis*	356,613,585	472,588	19,835	Scaffold
363	*Hydra vulgaris*	852,170,992	96,317	10,112	Scaffold
TENTACULATA					
364	*Mnemiopsis leidyi*	155,865,547	187,314	11,936	Scaffold
PLACOZOA					
365	*Trichoplax adhaerens*	105,631,681	5,978,658	204,191	Scaffold
PORIFERA					
366	*Amphimedon queenslandica*	166,699,561	120,365	11,817	Scaffold
VIRIDIPLANTAE					
367	*Micromonas pusilla CCMP1545*	22,000,071	1,183,541	83,658	Scaffold
368	*Micromonas sp. RCC299*	21,109,336	-	-	Complete Genome
369	*Ostreococcus lucimarinus CCE9901*	13,204,888	-	-	Complete Genome
370	*Ostreococcus tauri*	12,572,254	-	15,097	Chromosome
371	*Chlamydomonas reinhardtii*	120,404,952	1,695,175	44,607	Scaffold
372	*Volvox carteri f. nagariensis*	137,684,403	1,491,501	43,981	Scaffold
373	*Chlorella variabilis*	46,159,512	1,469,606	27,941	Scaffold
374	*Auxenochlorella protothecoides*	22,924,637	285,543	35,091	Scaffold
375	*Helicosporidium sp. ATCC 50920*	12,373,820	-	3,036	Contig
376	*Coccomyxa subellipsoidea C-169*	48,826,616	-	1,959,569	Contig
377	*Klebsormidium flaccidum*	104,209,716	134,930	72,848	Scaffold
378	*Physcomitrella patens*	477,947,608	1,331,933	74,043	Scaffold
379	*Selaginella moellendorffii*	212,501,708	1,749,879	119,796	Scaffold
380	*Pinus taeda*	265,480,119	-	16,205	Contig
381	*Amborella trichopoda*	706,332,640	4,927,027	29,332	Scaffold
382	*Spirodela polyrhiza*	132,009,443	-	14,533	Contig
383	*Phoenix dactylifera*	556,480,649	335,289	10,936	Scaffold
384	*Elaeis oleifera*	1,402,725,009	333,109	8,445	Scaffold
385	*Ensete ventricosum*	172,241,963	13,657	8,724	Scaffold
386	*Musa acuminata subsp. malaccensis*	472,235,617	1,311,088	28,326	Chromosome
387	*Sorghum bicolor*	739,150,314	33,610,934	195,424	Chromosome

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
388	*Zea mays*	2,067,622,303	217,928,451	41,305	Chromosome
389	*Setaria italica*	405,737,341	47,253,416	126,273	Scaffold
390	*Brachypodium distachyon*	272,058,505	59,328,898	347,842	Chromosome
391	*Leersia perrieri*	266,687,832	-	50,248	Chromosome
392	*Oryza barthii*	308,272,304	-	18,926	Chromosome
393	*Oryza brachyantha*	259,907,595	21,479,432	20,448	Chromosome
394	*Oryza glumipatula*	372,860,283	-	31,921	Chromosome
395	*Oryza longistaminata*	326,442,508	93,178	4,894	Scaffold
396	*Oryza meridionalis*	335,668,232	-	9,149	Chromosome
397	*Oryza nivara*	337,950,324	-	37,688	Chromosome
398	*Oryza punctata*	393,816,603	-	43,035	Chromosome
399	*Oryza sativa Japonica Group*	382,778,125	-	7,131,322	Chromosome
400	*Zizania latifolia*	603,989,347	604,864	13,068	Scaffold
401	*Aegilops tauschii*	3,313,650,219	68,369	6,204	Scaffold
402	*Triticum urartu*	3,747,047,519	85,725	5,782	Scaffold
403	*Nelumbo nucifera*	804,647,751	3,435,397	38,888	Scaffold
404	*Lupinus angustifolius*	523,297,923	15,485	6,485	Scaffold
405	*Phaseolus vulgaris*	521,076,696	50,367,376	39,053	Chromosome
406	*Cajanus cajan*	510,809,477	-	4,525	Contig
407	*Vigna angularis var. angularis*	291,823,841	21,697	13,270	Scaffold
408	*Vigna radiata var. radiata*	463,085,359	-	48,831	Scaffold
409	*Glycine max*	973,779,290	47,781,076	189,608	Chromosome
410	*Glycine soja*	863,568,428	404,776	24,258	Scaffold
411	*Cicer arietinum*	530,893,862	697,963	34,102	Chromosome
412	*Medicago truncatula*	314,477,977	1,268,006	377,846	Chromosome
413	*Trifolium pratense*	304,979,311	-	2,431	Contig
414	*Lotus japonicus*	147,812,252	-	1,327	Contig
415	*Malus x domestica*	1,874,765,253	1,893,947	12,414	Chromosome
416	*Pyrus x bretschneideri*	508,550,595	535,028	33,767	Scaffold
417	*Prunus mume*	234,030,239	636,887	32,274	Chromosome
418	*Prunus persica*	227,251,827	26,807,724	213,042	Scaffold
419	*Fragaria iinumae*	199,627,509	3,309	3,302	Scaffold
420	*Fragaria nubicola*	203,686,407	1,291	1,288	Scaffold
421	*Fragaria orientalis*	214,184,023	722	721	Scaffold
422	*Fragaria vesca subsp. vesca*	214,373,013	27,879,571	28,051	Chromosome
423	*Fragaria x ananassa*	173,229,572	5,137	2,235	Scaffold
424	*Morus notabilis*	320,378,613	405,448	40,438	Scaffold
425	*Cannabis sativa*	757,438,891	16,337	2,893	Scaffold
426	*Castanea mollissima*	833,240,550	32,186	17,495	Scaffold
427	*Betula nana*	564,011,153	-	4,884	Contig
428	*Cucumis melo*	374,928,489	4,278,129	18,084	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
429	*Cucumis sativus*	244,817,324	1,146,197	20,182	Scaffold
430	*Citrullus lanatus*	321,046,825	-	26,400	Contig
431	*Lagenaria siceraria*	176,727,258	782	537	Scaffold
432	*Populus euphratica*	496,032,534	482,055	40,490	Scaffold
433	*Populus trichocarpa*	417,286,671	18,835,763	242,226	Chromosome
434	*Jatropha curcas*	318,527,106	746,835	27,229	Scaffold
435	*Manihot esculenta subsp. flabellifolia*	390,835,892	14,635	7,381	Scaffold
436	*Ricinus communis*	350,621,860	496,528	21,409	Scaffold
437	*Linum usitatissimum*	282,201,865	-	21,193	Contig
438	*Eucalyptus camaldulensis*	478,318,461	-	5,922	Scaffold
439	*Eucalyptus grandis*	691,429,809	53,892,272	67,273	Scaffold
440	*Carica papaya*	369,781,828	1,089,885	10,609	Scaffold
441	*Arabidopsis halleri subsp. gemmifera*	221,139,660	-	2,864	Scaffold
442	*Arabidopsis lyrata subsp. lyrata*	206,667,935	24,464,547	227,391	Scaffold
443	*Arabidopsis thaliana*	119,667,750	-	11,194,537	Chromosome
444	*Camelina sativa*	641,356,059	-	32,728	Chromosome
445	*Capsella rubella*	133,063,876	15,040,190	134,357	Scaffold
446	*Brassica napus*	930,507,853	41,855,496	39,845	Chromosome
447	*Brassica oleracea var. oleracea*	488,954,160	-	21,938	Chromosome
448	*Brassica rapa*	284,129,391	1,960,303	46,175	Chromosome
449	*Raphanus raphanistrum subsp. raphanistrum*	253,833,977	-	10,186	Contig
450	*Raphanus sativus*	341,476,849	-	24,475	Contig
451	*Aethionema arabicum*	192,487,774	123,806	10,972	Scaffold
452	*Arabis alpina*	308,032,609	13,663,266	9,916	Chromosome
453	*Eutrema parvulum*	137,073,209	9,613,411	5,288,824	Chromosome
454	*Eutrema salsugineum*	243,110,105	13,441,892	222,588	Scaffold
455	*Sisymbrium irio*	245,550,082	144,321	11,849	Scaffold
456	*Leavenworthia alabamica*	173,431,618	71,084	15,066	Scaffold
457	*Tarenaya hassleriana*	249,929,577	1,600,628	22,358	Scaffold
458	*Gossypium arboreum*	1,560,831,891	-	71,965	Contig
459	*Gossypium raimondii*	761,565,430	62,175,169	119,844	Chromosome
460	*Theobroma cacao*	345,993,675	4,345,749	84,396	Chromosome
461	*Aquilaria agallochum*	726,710,453	128,399	16,831	Scaffold
462	*Azadirachta indica*	261,457,759	-	3,491	Contig
463	*Citrus clementine*	301,364,702	31,410,901	109,915	Scaffold
464	*Citrus sinensis*	327,829,540	1,778,813	49,898	Chromosome
465	*Vitis vinifera*	486,196,837	3,426,264	102,851	Chromosome

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
466	*Amaranthus hypochondriacus*	502,147,592	42,518	644	Scaffold
467	*Amaranthus tuberculatus*	4,347,977	-	241	Contig
468	*Beta vulgaris subsp. vulgaris*	566,567,364	2,013,180	26,532	Chromosome
469	*Spinacia oleracea*	474,077,292	20,184	6,184	Scaffold
470	*Dianthus caryophyllus*	567,661,845	60,730	16,727	Scaffold
471	*Actinidia chinensis*	604,217,145	-	58,864	Contig
472	*Vaccinium macrocarpon*	414,621,889	-	4,278	Scaffold
473	*Diospyros lotus*	1,104,189	-	1,864	Contig
474	*Primula veris*	309,692,940	-	13,157	Scaffold
475	*Solanum arcanum*	665,186,956	-	31,288	Contig
476	*Solanum habrochaites*	724,284,658	-	37,085	Contig
477	*Solanum lycopersicum*	781,508,772	16,467,796	86,952	Chromosome
478	*Solanum melongena*	833,080,760	64,530	15,025	Scaffold
479	*Solanum pennellii*	720,458,090	-	28,135	Contig
480	*Solanum pimpinellifolium*	688,247,059	-	5,714	Contig
481	*Solanum tuberosum*	705,934,411	1,344,915	31,933	Scaffold
482	*Capsicum annuum*	3,063,642,317	2,472,394	29,995	Scaffold
483	*Nicotiana sylvestris*	2,221,987,173	79,727	26,746	Scaffold
484	*Nicotiana tomentosiformis*	1,688,468,039	82,598	34,053	Scaffold
485	*Fraxinus excelsior*	875,243,685	98,766	22,012	Scaffold
486	*Penstemon centranthifolius*	4,471,593	-	752	Contig
487	*Penstemon grinnellii*	3,663,524	-	780	Contig
488	*Sesamum indicum*	275,059,498	2,060,396	52,214	Chromosome
489	*Genlisea aurea*	43,357,795	5,786	5,761	Contig
490	*Mimulus guttatus*	322,166,964	1,123,783	45,503	Scaffold
491	*Conyza canadensis*	326,165,195	-	20,748	Contig
ECHINOIDEA					
492	*Lytechinus variegatus*	951,759,914	42,623	6,207	Scaffold
493	*Strongylocentrotus purpuratus*	936,580,645	402,043	13,455	Scaffold
ASTEROIDEA					
494	*Patiria miniata*	811,028,858	52,614	9,466	Scaffold
ENTEROPNEUSTA					
495	*Saccoglossus kowalevskii*	775,840,395	245,820	10,074	Scaffold
TUNICATA					
496	*Ciona intestinalis*	115,227,500	3,102,162	37,096	Chromosome
497	*Ciona savignyi*	587,352,817	601,536	22,563	Scaffold
498	*Botryllus schlosseri*	579,633,380	7,220	6,924	Scaffold
499	*Oikopleura dioica*	70,471,451	395,387	24,932	Scaffold
LEPTOCARDII					
500	*Branchiostoma floridae*	521,895,125	2,586,727	27,925	Scaffold

No	Species	Total Sequence Length	Scaffold N50	Contig N50	Assembly Level
CEPHALASPIDOMORPHI					
501	*Lethenteron camtschaticum*	1,030,662,718	1,051,965	9,240	Scaffold
502	*Petromyzon marinus*	885,534,757	184,619	13,108	Scaffold
SARCOPTERYGII					
503	*Latimeria chalumnae*	2,860,591,921	924,513	12,671	Scaffold

Table 3: Genome quality check - assembly method and coverage

Table B.3: **Assembly method and coverage:** Shows the systematic name, assembly method, sequencing technology and estimated genome coverage for the 503 genomes used in this study. Species are listed in the same order as Table B.1. For most genomes, this information can be found on NCBI (www.ncbi.nlm.nih.gov/ → Assembly → look up the genome of interest → WGS Project file). However, the information is incomplete in some genomes because the NCBI Assembly database only contains the information provided by the submitters.

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
MAMMALIA				
1	*Tachyglossus aculeatus*	-	-	-
2	*Ornithorhynchus anatinus*	PCAP	WGS plasmid, fosmid end and BAC end sequences	6x
3	*Monodelphis domestica*	ARACHNE2+	Sanger	6.8x
4	*Macropus eugenii*	?	ABI 3730; Sanger; SOLiD	2x
5	*Sarcophilus harrisii*	Phusion2 v. 1.0	Illumina HiSeq2000	85x
6	*Dasypus novemcinctus*	Celera Assembler v. 6.0; Atlas-Link; Atlas-Gap-Fill	Sanger	6x
7	*Choloepus hoffmanni*	?	?	2.18x
8	*Chrysochloris asiatica*	Allpaths v. R42316 HAPLOIDIFY=True	Illumina HiSeq	66x
9	*Echinops telfairi*	Allpaths v. R37599	Illumina HiSeq	78x
10	*Orycteropus afer afer*	Allpaths v. R40776 LITTLE_HELPS_BIG=False	Illumina HiSeq	44x
11	*Elephantulus edwardii*	Allpaths v. R42301 HAPLOIDIFY=True	Illumina HiSeq	62x
12	*Trichechus manatus latirostris*	AllPaths v. R38542	Illumina HiSeq	150x
13	*Procavia capensis*	Arachne v. before 2009	Sanger	2.41x
14	*Loxodonta africana*	-	-	-
15	*Erinaceus europaeus*	Allpaths v. R41008	Illumina Hi-Seq	79x
16	*Sorex araneus*	Allpaths v. R41070	Illumina Hi-Seq	120x
17	*Condylura cristata*	AllPaths v. 2012	Illumina HiSeq	113.1x
18	*Pteropus alecto*	SOAPdenovo v. 1.06	Illumina HighSeq 2000	110x
19	*Pteropus vampyrus*	?	?	2.89x
20	*Eidolon helvum*	CLC NGS Cell v. 3.20.50819; SOAPdenovo v. 1.05	Illumina HiSeq	18.0x
21	*Megaderma lyra*	CLC NGS Cell v. 3.20.50819; SOAPdenovo v. 1.05	Illumina HiSeq	18.0x
22	*Rhinolophus ferrumequinum*	CLC NGS Cell v. 3.20.50819; SOAPdenovo v. 1.05	Illumina HiSeq	17.0x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
23	*Pteronotus parnellii*	CLC NGS Cell v. 3.20.50819; SOAPdenovo v. 1.05	Illumina HiSeq	17.0x
24	*Eptesicus fuscus*	ALLPATHS v. R40371	Illumina Hi-Seq	84x
25	*Myotis brandtii*	SOAP de novo v. 2.04	Illumina HiSeq 2000	120x
26	*Myotis davidii*	SOAPdenovo v. 1.06	Illumina HighSeq 2000	110x
27	*Myotis lucifugus*	Arachne v. Dev. Jan 2010	Sanger	7x
28	*Ceratotherium simum simum*	ALLPATHS v. R41370	Illumina Hi-Seq	91x
29	*Equus przewalskii*	SOAPdenovo v. 1.05	Illumina HiSeq	85.63x
30	*Equus caballus* (Thoroughbred)	ARACHNE 2.0	?	6.8x
31	*Equus caballus* (Mongolian)	SOAPdenovo v. 1.05	Illumina HiSeq	90.57x
32	*Manis pentadactyla*	SOAPdenovo v. May 2014	Illumina	61x
33	*Felis catus*	Celera Assembler v. 6.1	Sanger; 454 Titanium	2x Sanger; 12x 454
34	*Panthera tigris altaica*	SOAPdenovo v. 08-Jul-2011	Illumina HiSeq 2000	99x
35	*Canis lupus familiaris*	Arachne v. April 2010	Sanger	7x plus >90Mb finished sequence
36	*Ursus maritimus*	SOAPdenovo v. 1.05	Illumina Genome Analyzer II	101x
37	*Ailuropoda melanoleuca*	SOAPdenovo	Illumina GA	60x
38	*Leptonychotes weddellii*	allpaths v. R41856 HAPLOIDIFY=True	Illumina HiSeq	82x
39	*Odobenus rosmarus divergens*	AllPaths v. 41070; Atlas-link v. 1.0; Atlas-gapfill v. 2.2	Illumina	200.0x
40	*Mustela putorius furo*	ALLPATHS-LG v. R36167	Illumina GA II; HiSeq 2000	162x
41	*Camelus dromedarius*	SOAPdenovo v. 1.05	Illumina HiSeq2000	65x
42	*Camelus ferus*	SOAPdenovo v. V1.05	Illumina GAIIx; 454 GS-FLX Titanium; SOLid 3	30x
43	*Vicugna pacos*	Newbler v. May-2012	Roche 454; ABI 3730	22x
44	*Sus scrofa* (Duroc)	SOAPdenovo v. 2.20 and Cortex v. 1.1	Illumina GAII	24x
45	*Sus scrofa* (Tibetan)	SOAPdenovo v. 1.06	Illumina HiSeq	131.0x
46	*Sus scrofa* (Ellegaard Gottingen minipig)	SOAPdenovo v. September 2011	Illumina HiSeq	79.0x
47	*Balaenoptera acutorostrata scammoni*	SOAPdenovo v. 16-Mar-2012	Illumina HiSeq 2000	92x
48	*Physeter catodon*	AllPaths-LG v. January 2013	Illumina	75x
49	*Lipotes vexillifer*	SOAPdenovo v. 2.01	Illumina Hiseq 2000	115x
50	*Tursiops truncatus*	Atlas, Atlas-Link, Atlas-GapFill v. June 2011	Sanger; 454 FLX; Illumina HighSeq	2.5x Sanger; 3.5x 454; 30x Illumina

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
51	*Orcinus orca*	AllPaths v. 41070; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	200.0x
52	*Pantholops hodgsonii*	SOAPdenovo v. 2.1	Illumina	67.0x
53	*Capra hircus*	SOAPdenovo v. 1.03	Illumina GA IIx	86.64x
54	*Ovis aries* (Texel)	SOAPdenovo v. 1.03	Illumina GAII; 454	142x Illumina; 4x 454
55	*Ovis aries musimon*	?	?	?
56	*Bubalus bubalis*	MaSuRCA v. 1.8.3	Illumina GAIIx; Illumina HiSeq; 454	70.0x
57	*Bison bison bison*	MaSuRCA v. 2.1.0	454; Illumina HiSeq	60.0x
58	*Bos mutus*	SOAPdenovo v. 1.14	Illumina HiSeq; Illumina GA	130x
59	*Bos indicus*	BWA v. 0.5.9	SOLiD	52x
60	*Bos taurus*	UMD Overlapper v. 2009; additional processing	Sanger	9x
61	*Ochotona princeps*	Allpaths v. R41206	Illumina Hi-Seq	103x
62	*Oryctolagus cuniculus*	?	ABI	7.48X
63	*Ictidomys tridecemlineatus*	ALLPATHS v. R37591	Illumina HiSeq	495.1x
64	*Heterocephalus glaber*	AllPaths v. R39605	Illumina HiSeq	90x
65	*Fukomys damarensis*	SOAPdenovo v. 1.06	Illumina Hiseq 2000	140x
66	*Cavia aperea*	GCA_000151735.1	Illumina HiSeq	333.0x
67	*Cavia porcellus*	unknown v. before 2008	Sanger	6.8x
68	*Chinchilla lanigera*	ALLPATHS v. R40776	Illumina Hi-Seq	87x
69	*Octodon degus*	allpaths v. R40507 LITTLE_HELPS_BIG=False	Illumina Hi-Seq	80x
70	*Dipodomys ordii*	?	?	2.02X
71	*Jaculus jaculus*	ALLPATHS v. R41689 HAPLOIDIFY=True	Illumina Hi-Seq	78x
72	*Nannospalax galili*	SOAPdenovo v. 2.04.4	Illumina HiSeq 2000	86x
73	*Mesocricetus auratus*	allpaths v. R44683	Illumina HiSeq	115x
74	*Cricetulus griseus*	SOAPdenovo v. 1.05	Illumina GA IIx	130x
75	*Microtus ochrogaster*	allpaths v. R41985 HAPLOIDIFY=True SPRS_WRITE_SEPS=True	Illumina Hi-Seq	94x
76	*Peromyscus maniculatus bairdii*	Newbler v. 2.3 and 2.5; AllPaths v. 41070; ATLAS-gapfill v. 2.2; ATLAS-link v. 1.0	FLX 454; Illumina HiSeq	110.0x
77	*Rattus norvegicus*	Newbler v. 2.0.0-PreRelease-01162009 paired with Phrap v. 0. 990329 for Sanger reads; CLC bio for Solid reads	Sanger; SOLiD	3x BAC; 6x WGS ABI Sanger reads

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
78	*Mus musculus*	?	?	?
79	*Tupaia belangeri*	?	?	2x
80	*Tupaia chinensis*	SOAPdenovo v. 1.05	Illumina HiSeq 2000	80x
81	*Galeopterus variegatus*	SOAPdenovo2 v. April 2014	Illumina	55x
82	*Otolemur garnettii*	ALLPATHS-LG v. R35710	Illumina GAIIx	137x
83	*Microcebus murinus*	?	?	1.9x
84	*Tarsius syrichta*	Celera v. DEC-2012	Sanger; 454; Illumina	48x
85	*Callithrix jacchus*	PCAP v. 2009	ABI 3730	6.6x
86	*Saimiri boliviensis boliviensis*	AllPaths v. R37830	Illumina HiSeq	80x
87	*Rhinopithecus roxellana*	SOAPdenovo v. 2013.04	Illumina HiSeq	53.7x
88	*Nasalis larvatus*	GCF_000002255.3; Macaca mulatta NCBI build 1.2	454; Illumina HiSeq	290.0x
89	*Chlorocebus sabaeus*	ALLPATHS and Newbler v. 13-Feb-2013	454 Titanium; Illumina HiSeq; ABI	95x
90	*Macaca fascicularis*	SOAPdenovo v. 1.0.5, SRPRISM v. 2.4; ARGO v. 0.1	Illumina HiSeq	68x
91	*Macaca mulatta*	SOAPdenovo v. 1.0	Illumina GA	50x
92	*Papio anubis*	CABOG v. 6.1; ATLAS-LINK v. 1.0; ATLAS-GAPFILL v. 2.0	Sanger 3730; 454 FLX; Illumina	2.5x Sanger; 4.5x 454; 85x Illumina
93	*Nomascus leucogenys*	Arachne Assembler, assisted with Human NCBI build 36 (UCSC hg18); Phrap for mitochondrial sequences	Sanger	5.6x in Q20 bases
94	*Pongo abelii*	?	?	?
95	*Gorilla gorilla gorilla*	?	?	?
96	*Pan paniscus*	Celera Assembler v. 5.4.3	454 GS FLX; 454 GS FLX Titanium	26x
97	*Pan troglodytes*	PCAP	?	6x
98	*Homo sapiens*	?	?	?
SAUROPSIDA				
99	*Apalone spinifera*	SOAPdenovo v. 1.0.5	Illumina PE	33.4x
100	*Pelodiscus sinensis*	SOAPdenovo v. 1.05	Illumina HiSeq 2000	105x
101	*Chelonia mydas*	SOAPdenovo v. 1.05	Illumina HiSeq2000	110x
102	*Chrysemys picta bellii*	Newbler v. 2.6	Sanger; 454; Illumina	15x
103	*Struthio camelus australis*	SOAPdenovo v. 1.6	Illumina HiSeq	85x
104	*Tinamus guttatus*	SOAPdenovo v. 1.6	Illumina HiSeq	100x
105	*Anas platyrhynchos*	SOAPdenovo Release v. 1.03	Solexa	60x
106	*Lyrurus tetrix tetrix*	SOAPdenovo+BWA v. 2012-10	SOLiD	50.0x
107	*Gallus gallus*	Celera Assembler v. 5.4	Sanger; 454	12x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
108	*Coturnix japonica*	Soapdenovo v. 1.0.5b; bwa v. 0.5.9; SSPACE v. 1.2	Illumina	172x
109	*Meleagris gallopavo*	Celera v. 5.3 (modified)	454-Titanium; Illumina GAII	17x
110	*Colinus virginianus*	CLC de novo assembler v. 4.9	Illumina HiSeq	77x
111	*Acanthisitta chloris*	SOAPdenovo v. 1.6	Illumina HiSeq	29x
112	*Manacus vitellinus*	SOAPdenovo v. 1.6	Illumina HiSeq	110x
113	*Zonotrichia albicollis*	Allpaths-LG v. Feb-2013	Illumina	63x
114	*Geospiza fortis*	SOAPdenovo v. 2.01	Illumina HiSeq 2000	115x
115	*Serinus canaria*	?	?	?
116	*Taeniopygia guttata*	PCAP v. 2008	Sanger	5.5x
117	*Ficedula albicollis*	SOAPdenovo v. 1.05	Illumina GAII; Illumina HiSeq	60.0x
118	*Pseudopodoces humilis*	SOAPdenovo v. 1.5	Illumina HiSeq	96.25x
119	*Corvus brachyrhynchos*	SOAPdenovo v. 1.6	Illumina HiSeq	80x
120	*Corvus cornix cornix*	AllPaths v. Allpaths-LG version 41687	Illumina HiSeq	152.0x
121	*Ara macao*	CLC Genomics Workbench v. 4.9	454 GS FLX Titanium; Illumina GAIIx; Illumina HiSeq 2000	26x
122	*Amazona vittata*	Ray software v. 3	Illumina	24.0x
123	*Melopsittacus undulatus*	Celera v. 6.1	454; Illumina	16x 454; 7x Illumina
124	*Nestor notabilis*	SOAPdenovo v. 1.6	Illumina HiSeq	32x
125	*Falco cherrug*	SOAPdenovo v. 1.4	Illumina HiSeq	147.0x
126	*Falco peregrinus*	SOAPdenovo v. 1.4	Illumina HiSeq	137.6x
127	*Cariama cristata*	SOAPdenovo v. 1.6	Illumina HiSeq	24x
128	*Merops nubicus*	SOAPdenovo v. 1.6	Illumina HiSeq	37x
129	*Picoides pubescens*	SOAPdenovo v. 1.6	Illumina HiSeq	105x
130	*Buceros rhinoceros silvestris*	SOAPdenovo v. 1.6	Illumina HiSeq	35x
131	*Apaloderma vittatum*	SOAPdenovo v. 1.6	Illumina HiSeq	28x
132	*Leptosomus discolour*	SOAPdenovo v. 1.6	Illumina HiSeq	32x
133	*Haliaeetus albicilla*	SOAPdenovo v. 1.6	Illumina HiSeq	26x
134	*Haliaeetus leucocephalus*	SOAPdenovo2 v. May 2014	Illumina	103x
135	*Aquila chrysaetos Canadensis*	AllPaths-LG v. August 2013	Illumina	88x
136	*Cathartes aura*	SOAPdenovo v. 1.6	Illumina HiSeq	25x
137	*Tyto alba*	SOAPdenovo v. 1.6	Illumina HiSeq	27x
138	*Colius striatus*	SOAPdenovo v. 1.6	Illumina HiSeq	27x
139	*Charadrius vociferus*	SOAPdenovo v. 1.6	Illumina HiSeq	100x
140	*Balearica regulorum gibbericeps*	SOAPdenovo v. 1.6	Illumina HiSeq	33x
141	*Chlamydotis macqueenii*	SOAPdenovo v. 1.6	Illumina HiSeq	27x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
142	*Cuculus canorus*	SOAPdenovo v. 1.6	Illumina HiSeq	100x
143	*Fulmarus glacialis*	SOAPdenovo v. 1.6	Illumina HiSeq	33x
144	*Aptenodytes forsteri*	SOAPdenovo v. 1.6	Illumina HiSeq	60x
145	*Pygoscelis adeliae*	SOAPdenovo v. 1.6	Illumina HiSeq	60x
146	*Phalacrocorax carbo*	SOAPdenovo v. 1.6	Illumina HiSeq	24x
147	*Pelecanus crispus*	SOAPdenovo v. 1.6	Illumina HiSeq	34x
148	*Nipponia nippon*	SOAPdenovo v. 1.6	Illumina HiSeq	105x
149	*Egretta garzetta*	SOAPdenovo v. 1.6	Illumina HiSeq	74x
150	*Phaethon lepturus*	SOAPdenovo v. 1.6	Illumina HiSeq	39x
151	*Gavia stellata*	SOAPdenovo v. 1.6	Illumina HiSeq	33x
152	*Tauraco erythrolophus*	SOAPdenovo v. 1.6	Illumina HiSeq	30x
153	*Opisthocomus hoazin*	SOAPdenovo v. 1.6	Illumina HiSeq	100x
154	*Columba livia*	SOAPdenovo v. 2.0	Illumina HiSeq 2000	60x
155	*Pterocles gutturalis*	SOAPdenovo v. 1.6	Illumina HiSeq	25x
156	*Calypte anna*	SOAPdenovo v. 1.6	Illumina HiSeq	110x
157	*Chaetura pelagica*	SOAPdenovo v. 1.6	Illumina GA II	106x
158	*Caprimulgus carolinensis*	SOAPdenovo v. 1.6	Illumina HiSeq	30x
159	*Eurypyga helias*	SOAPdenovo v. 1.6	Illumina HiSeq	33x
160	*Mesitornis unicolor*	SOAPdenovo v. 1.6	Illumina HiSeq	29x
161	*Podiceps cristatus*	SOAPdenovo v. 1.6	Illumina HiSeq	30x
162	*Phoenicopterus ruber ruber*	SOAPdenovo v. 1.6	Illumina HiSeq	33x
163	*Alligator mississippiensis*	Allpaths v. R41313; manual scaffolding	Illumina	68x
164	*Alligator sinensis*	SOAPdenovo v. 1.05	Illumina GA II	109x
165	*Crocodylus porosus*	AllPaths v. R41313	Illumina	74.0x
166	*Gavialis gangeticus*	SOAPdenovo v. 2.04.3	Illumina	81.0x
167	*Pogona vitticeps*	-	Illumina HiSeq 2000	83x
168	*Anolis carolinensis*	Arachne v. 3.0.0	ABI	7.10x
169	*Vipera berus berus*	AllPaths LG v. v41070	Illumina	121.0x
170	*Crotalus mitchellii pyrrhus*	CLC NGS Cell v. June 2013	Illumina HiSeq	40.0x
171	*Ophiophagus hannah*	CLC NGS Cell v. 2011	Illumina HiSeq	28.0x
172	*Python bivittatus*	Soap deNovo v. March 2012	Illumina; 454	20x
AMPHIBIA				
173	*Nanorana parkeri*	SOAPdenovo v. 1.06	Illumina Genome Analyzer II	83x
174	*Xenopus tropicalis*	ARACHNE v. 20071016_modified	ABI 3739	7.44x
NEOPTERYGII				
175	*Lepisosteus oculatus*	AllPaths v. R38293	Illumina HiSeq	677.5x
176	*Anguilla anguilla*	C bio Genomics Workbench v. 3.6.5.	Illumina GAIIx	20.0x
177	*Anguilla japonica*	CLC NGS Cell v. 3.2; CLC NGS Cell v. 4.0beta	Illumina GAII	35.0x
178	*Danio rerio*	?	?	?

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
179	*Astyanax mexicanus*	AllPathsLG v. Jan-2013	Illumina	70x
180	*Oryzias latipes*	?	?	?
181	*Poecilia formosa*	AllPaths-LG v. July 2013	Illumina	82x
182	*Xiphophorus maculatus*	PCAP v. 3/30/09; Newbler v. MapAsmResearch-02/17/2010	454; ABI 3730	19.6x
183	*Fundulus heteroclitus*	AllPaths-LG v. November 2012	Illumina	75x
184	*Takifugu flavidus*	HAPs v. 0.2.2	SOLiD	131.0x
185	*Takifugu rubripes*	?	?	?
186	*Tetraodon nigroviridis*	?	?	?
187	*Cynoglossus semilaevis*	SOAPdenovo v. April-2011	Illumina HiSeq2000	117x (male); 94x (female)
188	*Haplochromis burtoni*	ALLPATHS-LG v. R35951	Illumina HiSeq	131x
189	*Pundamilia nyererei*	ALLPATHS-LG v. R37016	Illumina HiSeq	126x
190	*Maylandia zebra*	AllPaths v. R37043	Illumina HiSeq	129.7x
191	*Neolamprologus brichardi*	ALLPATHS-LG v. R36800	Illumina HiSeq	171x
192	*Oreochromis niloticus*	AllPaths-LG v. R35292	Illumina	269x
193	*Sebastes nigrocinctus*	SOAPdenovo v. 1.05	Illumina HiSeq	71.0x
194	*Sebastes rubrivinctus*	SOAPdenovo v. 1.05	Illumina HiSeq	87.0x
195	*Gasterosteus aculeatus*	?	?	?
196	*Gadus morhua*	?	?	?
CHONDRICHTHYES				
197	*Callorhinchus milii*	Celera v. 6.1	454 Titanium; Sanger ABI 3730	19.25x
198	*Carcharhinus brachyurus*	SOAPdenovo	Illumina 100 PE	-
ECDYSOZOA				
199	*Ephemera danica*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	207.7x
200	*Ladona fulva*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	131.2x
201	*Pediculus humanus corporis*	?	?	8x
202	*Frankliniella occidentalis*	AllPaths LG v. 44620; Atlas Link v. 1.0; Atlas GapFill v. 2.2	Illumina	158.7x
203	*Diaphorina citri*	PBJelly v. 2013	Illumina HiSeq; PacBio	180.0x
204	*Pachypsylla venusta*	AllPaths LG v. 44620; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina	181.4x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
205	*Acyrthosiphon pisum*	Newbler v. 2.3-PreRelease-09142009; phrap version 1.080812; atlas-link0.5; atlas-overlapper-1.70	Sanger; 454	15x
206	*Nilaparvata lugens*	SOAPdenovo v. 1.05	Illumina HiSeq 2000; Illumina GA-II	133.27x
207	*Oncopeltus fasciatus*	AllPaths LG v. 44620; Atlas Link v. 1.0; Atlas GapFill v. 2.2	Illumina	106.9x
208	*Rhodnius prolixus*	Celera assembler v. 6.0	ABI 3730; 454	8.1x
209	*Cimex lectularius*	AllPaths LG v. 35218; Atlas Link v. v1.0; Atlas GapFill v. v2.2	Illumina	163.4x
210	*Onthophagus taurus*	AllPaths LG v. 44620; Atlas Link v. v1.0; Atlas GapFill v. v2.2	Illumina	86.0x
211	*Agrilus planipennis*	AllPaths LG v. 44620; Atlas Link v. 1.0; Atlas GapFill v. 2.2	Illumina	124.7x
212	*Tribolium castaneum*	?	?	7.3x
213	*Anoplophora glabripennis*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	121.0x
214	*Leptinotarsa decemlineata*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.0	Illumina HiSeq	52.3x
215	*Dendroctonus ponderosae*	ABySS v. 1.3.0	Illumina GAII	443.0x
216	*Mengenilla moldrzyki*	Newbler v. 2.3	454 GS-FLX Titanium (Roche)	16x
217	*Aedes aegypti*	?	?	7.6x
218	*Culex quinquefasciatus*	?	?	6.14x
219	*Anopheles albimanus*	allpaths v. R43019	Illumina	95.0x
220	*Anopheles arabiensis*	allpaths v. R43436	Illumina	100.0x
221	*Anopheles atroparvus*	allpaths v. R46504	Illumina	98.0x
222	*Anopheles christyi*	allpaths v. R44024	Illumina	35.0x
223	*Anopheles culicifacies*	allpaths v. R46449	Illumina	31.0x
224	*Anopheles darlingi*	Celera Assembler v. 6.1	454	20.0x
225	*Anopheles dirus*	allpaths v. R43500	Illumina	184.0x
226	*Anopheles epiroticus*	allpaths v. R43500	Illumina	49.0x
227	*Anopheles farauti*	allpaths v. R47616	Illumina	233.0x
228	*Anopheles funestus*	allpaths v. R43460	Illumina	135.0x
229	*Anopheles gambiae*	?	?	10x
230	*Anopheles maculatus*	allpaths v. R46504	Illumina	25.0x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
231	*Anopheles melas*	allpaths v. R47956	Illumina	147.0x
232	*Anopheles merus*	allpaths v. R47616	Illumina	147.0x
233	*Anopheles minimus*	allpaths v. R43460	Illumina	211.0x
234	*Anopheles quadriannulatus*	allpaths v. R43436	Illumina	93.0x
235	*Anopheles sinensis*	Celera Assembler v. 6.1	454	30.0x
236	*Anopheles stephensi*	Newbler v. 2.8	454; PacBio; Illumina GAII	16.0x
237	*Mayetiola destructor*	?	?	?
238	*Lutzomyia longipalpis*	Celera Assembler v. 6.1	454	38.9x
239	*Phlebotomus papatasi*	Newbler v. 2.6RC02	454; Sanger	15.1x
240	*Ceratitis capitata*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	152.5x
241	*Drosophila albomicans*	SOAPdenovo v. 1.04	Illumina GAII	100.0x
242	*Drosophila ananassae*	?	?	8.9x
243	*Drosophila biarmipes*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	Illumina; 454	186.9x
244	*Drosophila bipectinata*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina	266.3x
245	*Drosophila elegans*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATKv. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina	204.9x
246	*Drosophila erecta*	?	?	10.6x
247	*Drosophila eugracilis*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina	218.1x
248	*Drosophila ficusphila*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina	208.7x
249	*Drosophila grimshawi*	?	?	7.9x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
250	*Drosophila kikkawai*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina	181.7x
251	*Drosophila melanogaster*	?	?	?
252	*Drosophila miranda*	Newbler v. 2.6	Illumina HiSeq; 454	95.0x
253	*Drosophila mojavensis*	?	?	8.2x
254	*Drosophila persimilis*	?	?	4.1x
255	*Drosophila pseudoobscura pseudoobscura*	PBJelly v. 12.8.2; Atlas genome assembly	Sanger; PacBio	9.7x Sanger; 24x PacBio
256	*Drosophila rhopaloa*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina GAIIx	214.5x
257	*Drosophila sechellia*	?	?	4.9x
258	*Drosophila simulans*	?	?	?
259	*Drosophila suzukii*	SOAPdenovo v. 2	Illumina GA IIx	195x
260	*Drosophila takahashii*	Celera Assembler v. 6.1; BWA v. 0.6.0; Samtools v. 0.1.14; GATK v. 1.1-9; Indel_call_and_upgrade.pl v. 1.0	454; Illumina GAIIx	242.1x
261	*Drosophila virilis*	?	?	8.0x
262	*Drosophila willistoni*	?	?	?
263	*Drosophila yakuba*	?	?	9.1x
264	*Musca domestica*	AllPathsLG v. September 2012	Illumina	76x
265	*Glossina austeni*	ALLPATHS-LG v. August 2013	Illumina	62x
266	*Glossina brevipalpis*	ALLPATHS-LG v. August 2013	Illumina	81x
267	*Glossina fuscipes fuscipes*	ALLPATHS-LG v. September 2013	Illumina	120x
268	*Glossina morsitans morsitans*	?	?	?
269	*Glossina pallidipes*	ALLPATHS-LG v. August 2013	Illumina	78x
270	*Limnephilus lunatus*	AllPaths LG v. version 35218; Atlas Link v. v1.0; Atlas GapFill v. v2.2	Illumina	80.1x
271	*Papilio glaucus*	platanus v. MAY-2014	Illumina HiSeq	72x
272	*Papilio polytes*	Platanus v. 1.2.1	Illumina HiSeq2000	595x

No	**Species**	**Assembly Method**	**Sequencing Technology**	**Genome Coverage**
273	*Papilio xuthus*	Platanus v. 1.2.1	Illumina HiSeq2500	300x
274	*Heliconius melpomene melpomene*	?	?	?
275	*Melitaea cinxia*	SOAPdenovo v. 1.05 ; Newbler v. 2.6 ; MIP Scaffolder v. MARCH-2012	Illumina; 454; SOLiD	70.0x
276	*Danaus plexippus*	CLC bio's de novo assembler SSPACE v. 1.0; GapCloser	454 GS-FLX Titanium; Illumina GAIIx	74.7x
277	*Bombyx mori*	?	?	?
278	*Manduca sexta*	Newbler v. 2.3	454	80.7x
279	*Plutella xylostella*	BAC by BAC v. JUN-2011 (Rabbit software)	Illumina Hiseq2000	141x (10X fosmid+131X WGS)
280	*Athalia rosae*	AllPaths v. v35218; ATLAS-link v. v1.0; ATLAS-gapfill v. v2.2	Illumina HiDrw	467.2x
281	*Cephus cinctus*	SOAPdenovo v. Apr-2012	Illumina GAII; Illumina GAIIx; Illumina HiSeq	56.0x
282	*Orussus abietinus*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	467.1x
283	*Ceratosolen solmsi marchali*	SOAPdenovo v. 2	Illumina HS2000	93x
284	*Nasonia giraulti*	?	Sanger; Illumina Solexa	1x Sanger; 10x Solexa
285	*Nasonia longicornis*	?	Sanger; Illumina Solexa	1x Sanger; 10x Solexa
286	*Nasonia vitripennis*	Atlas-overlapper v. 2008; Phrap; Atlas-scaffold	ABI	6x
287	*Copidosoma floridanum*	AllPaths LG v. Version 35218; Atlas Link v. v1.0; Atlas GapFill v. v2.2	Illumina	139.0x
288	*Trichogramma pretiosum*	AllPaths v. v44620; Atlas Link v. v1.0; Atlas GapFill v. v2.2	Illumina	232.7x
289	*Microplitis demolitor*	SOAPdenovo v. December 2012	Illumina HiSeq	26.0x
290	*Megachile rotundata*	SOAPdenovo Assembler v. 1.05	Illumina GA IIx	300x
291	*Apis dorsata*	SOAPdenovo v. 1.05	Illumina GAII	60.0x
292	*Apis florea*	Newbler v. 2.3-PreRelease-10/19/2009 and Phrap	454	20.5x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
293	*Apis mellifera*	Atlas assembly system v. before 2011	Sanger; SOLiD; 454	8x
294	*Bombus impatiens*	SOAPdenovo v. 1.05	Illumina GAIIx	127x
295	*Bombus terrestris*	Newbler v. 2.3-PreRelease-10/19/2009 and Phrap	454	21.4x
296	*Linepithema humile*	CABOG v. 5.3	454; Illumina	30x
297	*Camponotus floridanus*	SOAPdenovo v. 1.0	Illumina GA	100x
298	*Acromyrmex echinatior*	SOAPdenovo v. 1.04	Illumina HiSeq 2000	107x
299	*Atta cephalotes*	454 GS De novo assembler software R&D v. 03/06/2010	454 Titanium including shotgun, 8-kbp and 20-kbp paired-end sequencing datasets	18-20x
300	*Solenopsis invicta*	SOAPdenovo v. 1.04; Newbler v. 2.3	454; Illumina	15x 454; 30x Illumina
301	*Pogonomyrmex barbatus*	CABOG v. 5.3	454	10.5-12x
302	*Harpegnathos saltator*	SOAPdenovo v. 1.0	Illumina GA	100x
303	*Cerapachys biroi*	SOAPdenovo v. V2.05	Illumina HiSeq; 454	123.0x
304	*Blattella germanica*	AllPaths LG v. 44620; Atlas Link v. 1.0; Atlas GapFill v. 2.2	Illumina	158.5x
305	*Zootermopsis nevadensis*	SOAPdenovo v. 2.04	Illumina HiSeq	98.0x
306	*Daphnia pulex*	JAZZ v. 1.0 (4/11/2006)	Sanger	8.7x
307	*Eurytemora affinis*	AllPaths LG v. v44620; ATLAS-link v. v1.0; ATLAS-gapfill v. v2.2	Illumina	75.0x
308	*Hyalella azteca*	AllPaths LG v. 48744; Atlas Link v. 1.0; Atlas GapFill v. 2.2	Illumina	105.0x
309	*Strigamia maritima*	Celera Assembler v. 6.1	454	23x
310	*Stegodyphus mimosarum*	SOAPdenovo v. 2	Illumina Hiseq 2000	70x
311	*Latrodectus hesperus*	AllPaths LG v. 35218; Atlas Link v. 1.0; Atlas GapFill v. 2.2	Illumina	48.0x
312	*Parasteatoda tepidariorum*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	96.9x
313	*Tetranychus urticae*	?	?	?
314	*Dermatophagoides farinae*	AllPaths v. 39075; Velvet v. 1.1.03; SOAPdenovo v. 1.05	Illumina HiSeq	436.0x
315	*Sarcoptes scabiei type canis*	Minia v. 1.6088; SSPACE Standard v. 3.0	Illumina	174.0x
316	*Achipteria coleoptrata*	Platanus v. 1.2.1	Illumina HiSeq	90.0x
317	*Hypochthonius rufulus*	Platanus v. 1.2.1	Illumina HiSeq	70.0x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
318	*Platynothrus peltifer*	ABySS v. 1.3.5	Illumina GAIIx	19.0x
319	*Steganacarus magnus*	ABySS v. 1.3.5	Illumina GAIIx	17.0x
320	*Ixodes ricinus*	CLC NGS Cell v. 7.4	Illumina HiSeq	47.0x
321	*Ixodes scapularis*	?	?	?
322	*Rhipicephalus microplus*	Newbler v. 2.0.00.20; Cap3 v.d. 05/04/15	454	3.41x
323	*Metaseiulus occidentalis*	Celera v. 6.1	454 XLR	17.7x
324	*Varroa destructor*	Celera Assembler v. 5.2	454	5x
325	*Centruroides exilicauda*	AllPaths v. 35218; ATLAS-link v. 1.0; ATLAS-gapfill v. 2.2	Illumina HiSeq	181.1x
326	*Mesobuthus martensii*	Velvet v. 1.1.04	Illumina HiSeq; Illumina GAIIx; 454	200.0x
327	*Limulus polyphemus*	Newbler v. May 2013	Roche 454	18x
328	*Trichinella spiralis*	PCAP v. January 12, 2007	ABI 3730	34.6x
329	*Ascaris suum*	SOAPdenovo v. 1.04	Illumina HiSeq 2000	90x
330	*Elaeophora elaphi*	?	?	?
331	*Onchocerca volvulus*	?	?	?
332	*Steinernema monticolum*	Velvet v. 1.2.07	Illumina GAIIx	61.0x
333	*Panagrellus redivivus*	Velvet v. 1.2.07	Illumina HiSeq	100.0x
334	*Haemonchus contortus*	?	?	?
335	*Necator americanus*	Newbler v. MapAsmResearch-04/19/2010-patch-08/17/2010	454	26.15x
336	*Heterorhabditis bacteriophora*	Celera assembler v. 6.0	454; ABI 3730	26.1x
337	*Caenorhabditis angaria*	Velvet v. 0.7.56	Illumina	170x
338	*Caenorhabditis brenneri*	PCAP v. 9/3/04	ABI	9.5x
339	*Caenorhabditis briggsae*	?	?	?
340	*Caenorhabditis elegans*	?	?	?
341	*Caenorhabditis japonica*	Celera assembler v. 6.0	ABI; 454	22x
342	*Caenorhabditis sp. 11 MAF-2010*	Celera assembler v. 6.0	454 Titanium	20.2x
343	*Priapulus caudatus*	AllPaths-LG v. February 2013	Illumina	129x
ROTIFERA				
344	*Adineta vaga*	?	?	?
PLATYHELMINTHES				
345	*Schistosoma curassoni*	?	?	?
346	*Schistosoma haematobium*	SOAPdenovo v. 1.04	Illumina HiSeq 2000	74x
347	*Schistosoma japonicum*	?	?	?
348	*Schistosoma mansoni*	?	?	?

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
349	*Schistosoma margrebowiei*	?	?	?
350	*Schistosoma mattheei*	?	?	?
351	*Schistosoma rodhaini*	?	?	?
352	*Clonorchis sinensis*	Celera assembler v. 6.1	Illumina	72x
353	*Echinococcus granulosus*	?	?	?
354	*Echinococcus multilocularis*	?	?	?
355	*Hymenolepis microstoma*	?	?	?
ANNELIDA				
356	*Capitella teleta*	JGI Jazz Assembler v. 10/16/2006	Sanger	7.9x
357	*Helobdella robusta*	JGI Jazz v. 01/29/2007	Sanger	7.92x
MOLLUSCA				
358	*Crassostrea gigas*	SOAPdenovo v. 1.05	Illumina HiSeq 2000	100x
359	*Lottia gigantea*	JGI Jazz v. 08/24/2006	Sanger	8.87x
360	*Aplysia californica*	allpaths v. R40582	Illumina Hi-Seq	66x
361	*Biomphalaria glabrata*	Newbler v. July 2011	454 FLX Titanium	27.5x
CNIDARIA				
362	*Nematostella vectensis*	?	?	?
363	*Hydra vulgaris*	Ringer-Phrap	?	around 4x per haplotype
TENTACULATA				
364	*Mnemiopsis leidyi*	Phusion v. 1.02	454 GS-FLX Titanium; Illumina GA IIX	12x
PLACOZOA				
365	*Trichoplax adhaerens*	?	?	8.06x
PORIFERA				
366	*Amphimedon queenslandica*	PHRAPATTACK v. Feb 2007	Sanger	8x
VIRIDIPLANTAE				
367	*Micromonas pusilla CCMP1545*	?	ABI 3730	8x
368	*Micromonas sp. RCC299*	?	?	?
369	*Ostreococcus lucimarinus CCE9901*	?	?	?
370	*Ostreococcus tauri*	?	?	?
371	*Chlamydomonas reinhardtii*	?	?	12.8x
372	*Volvox carteri f. nagariensis*	?	?	?
373	*Chlorella variabilis*	JAZZ v. May 25, 2007	Sanger	8.9x
374	*Auxenochlorella protothecoides*	Newbler v. 2.6; Amos v. 3.1; SOAP denovo v. 2.01	454 GS FLX Titanium; Illumina HiSeq 2000	145x
375	*Helicosporidium sp. ATCC 50920*	Ray v. 2.0.0-rc8	Illumina HiSeq; Illumina GAIIx	62x

No	**Species**	**Assembly Method**	**Sequencing Technology**	**Genome Coverage**
376	*Coccomyxa subellipsoidea C-169*	Arachne v. 20071016	Sanger	8.16x + targeted finishing
377	*Klebsormidium flaccidum*	Newbler v. 2.6	454; Illumina GAIIx	40x
378	*Physcomitrella patens*	?	?	8.1x
379	*Selaginella moellendorffii*	Arachne v. 20070201	Sanger	7.0x
380	*Pinus taeda*	SOAPdenovo v. 2	Illumina GAIIx	Unknown
381	*Amborella trichopoda*	Newbler v. 2.6 build version 20110517_1502	454	30x
382	*Spirodela polyrhiza*	Newbler v. 2.6	454; ABI3730	22.0x
383	*Phoenix dactylifera*	Newbler v. 2.5.3; HAPS v. 0.1.200	454; SOLiD; ABI3730	139.0x
384	*Elaeis oleifera*	Newbler v. 2.6	454	16.0x
385	*Ensete ventricosum*	SOAPdenovo v. 1.05	Illumina HiSeq	40.0x
386	*Musa acuminata subsp. malaccensis*	?	?	?
387	*Sorghum bicolor*	?	Sanger	8x
388	*Zea mays*	?	?	?
389	*Setaria italica*	Arachne v. 2007101641HA	Sanger	8.29x
390	*Brachypodium distachyon*	Arachne v. 200710164HA	Sanger	9.43x
391	*Leersia perrieri*	AllPaths v. 44849	Illumina HiSeq	150.0x
392	*Oryza barthii*	AllPaths v. 44849; Newbler v. 2.8	Illumina GAIIx; 454	110.0x
393	*Oryza brachyantha*	SOAPdenovo v. 1.04	Illumina GA II	104x
394	*Oryza glumipatula*	AllPaths v. 44849	Illumina HiSeq	135.0x
395	*Oryza longistaminata*	SOAPdenovo v. 2.2	Illumina HiSeq	52.5x
396	*Oryza meridionalis*	ABySS v. 1.3.6	Illumina HiSeq	166.0x
397	*Oryza nivara*	AllPaths v. 44849	Illumina HiSeq	102.0x
398	*Oryza punctata*	AllPaths v. 44849	Illumina HiSeq; 454	130.0x
399	*Oryza sativa Japonica Group*	?	?	?
400	*Zizania latifolia*	AllPaths v. allpathslg-45489	Illumina HiSeq	140.0x
401	*Aegilops tauschii*	SOAPdenovo v. 1.05	Illumina GA IIx; HiSeq2000	90x
402	*Triticum urartu*	SOAPdenovo v. 1.05	Illumina Hiseq2000	90x
403	*Nelumbo nucifera*	AllPaths v. 40102	Illumina HiSeq; 454	100.0x
404	*Lupinus angustifolius*	?	?	?
405	*Phaseolus vulgaris*	ARACHNE v. 20071016_modfied	ABI 3730; Roche 454 FLX; Illumina GAII	21.02x
406	*Cajanus cajan*	Newbler v. 2.5.3	454 GS FLX phase D; 454 FLX Paired-end Titanium	10x
407	*Vigna angularis var. angularis*	ABySS v. 1.3.5	Illumina HiSeq	65.0x
408	*Vigna radiata var. radiata*	Allpath-LG v. 42179	Illumina HiSeq	300.0x

No	**Species**	**Assembly Method**	**Sequencing Technology**	**Genome Coverage**
409	*Glycine max*	Arachne v. 2007101641HA	Sanger	8.02x
410	*Glycine soja*	SOAPdenovo v. 2010	Illumina GAII	63.1x
411	*Cicer arietinum*	SOAPdenovo v. 2.03	Illumina Hiseq 2000	120x
412	*Medicago truncatula*	?	?	?
413	*Trifolium pratense*	ABySS v. 1.3.3	Illumina HiSeq	58.0x
414	*Lotus japonicus*	?	?	?
415	*Malus x domestica*	?	?	?
416	*Pyrus x bretschneideri*	BAC to BAC v. unpublished	Illumina HS2000	7.6x BAC; the average coverage depth of each BAC is 86x; additional 107.6x WGS
417	*Prunus mume*	SOAPdenovo v. 1.05	Illumina Hiseq2000	180x
418	*Prunus persica*	ARACHNE v. 20071016_modified	ABI 3739	8.47x
419	*Fragaria iinumae*	SOAPdenovo v. 1.05	Illumina HiSeq1000	321.3x
420	*Fragaria nubicola*	SOAPdenovo v. 1.05	Illumina HiSeq1000	351.5x
421	*Fragaria orientalis*	SOAPdenovo v. 1.05	Illumina HiSeq1000	203.4x
422	*Fragaria vesca subsp. vesca*	Celera Assembler (CA) v. 5.3	454; Solexa	49x
423	*Fragaria x ananassa*	Newbler v. 2.7; SOAPdenovo v. 1.05	454 GS FLX+; Illumina GAIIx; Illumina HiSeq1000	402.6x
424	*Morus notabilis*	SOAPdenovo v. 1.05	Illumina HiSeq2000	236x
425	*Cannabis sativa*	SOAPdenovo v. 1.0.5	Illumina GA II; HiSeq	130x
426	*Castanea mollissima*	Newbler v. 2.8	454; Illumina	46.0x
427	*Betula nana*	?	?	?
428	*Cucumis melo*	?	?	?
429	*Cucumis sativus*	?	ABI and Megabase; Solexa GA	about 72x (3.8x Sanger sequences; 68x Solexa GA sequences)
430	*Citrullus lanatus*	SOAPdenovo v. 1.05	Illumina GA; Illumina HiSeq	108.6x
431	*Lagenaria siceraria*	SOAPdenovo v. 2.01	Illumina HiSeq	8.9x
432	*Populus euphratica*	BAC by BAC v. Oct-2011	Illumina Hiseq2000	115x (4.4x fosmid+111x WGS)
433	*Populus trichocarpa*	ARACHNE v. 2007101641HA	ABI 3739	7.45X
434	*Jatropha curcas*	SOAPdenovo v. 1.0.4	Illumina GAII; Illumina HS	189x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
435	*Manihot esculenta subsp. flabellifolia*	SOAPdenovo v. 1.05; Newbler v. 2.5.3; GATE v. 1.0	Illumina GAII; Illumina HiSeq; 454	103.0x
436	*Ricinus communis*	?	?	4.5x
437	*Linum usitatissimum*	SoapDeNovo v. 01-Sep-2009	Illumina GA	94x
438	*Eucalyptus camaldulensis*	WGS v. 6.0	454 GS FLX; ABI 3730xl	4.2x
439	*Eucalyptus grandis*	ARACHNE v. 20071016_HA	ABI 3739	6.73x
440	*Carica papaya*	?	?	3x
441	*Arabidopsis halleri subsp. gemmifera*	SOAPdenovo v. 1.05	Illumina HiSeq2000	167.9x
442	*Arabidopsis lyrata subsp. lyrata*	Arachne v. before 2009	Sanger	8x
443	*Arabidopsis thaliana*	BAC by BAC	BAC physical map then Sanger sequencing of BACs	?
444	*Camelina sativa*	SOAPdenovo v. 2.01	Illumina HiSeq; 454	100.0x
445	*Capsella rubella*	ARACHNE v. 20071016_modified	ABI 3739; Roche 454FLX	22.35x
446	*Brassica napus*	SOAPdenovo v. 2.01	Illumina Hiseq2000	150x
447	*Brassica oleracea var. oleracea*	SOAPdenovo v. 1.05	Illumina GAII; Illumina HiSeq; 454	94.0x
448	*Brassica rapa*	SOAPdenovo v. 1.04	Illumina GA	72x
449	*Raphanus raphanistrum subsp. raphanistrum*	ABySS v. 1.2.5; Newbler v. 2.5.3; Celera Assembler v. 6.1; Minimus2/AMOS v. 3.1.0	Illumina GAII; 454	47.0x
450	*Raphanus sativus*	Newbler v. 2.6; SMRTanalysis v. 1.3.1	454; PacBio; Illumina HiSeq; ABI3730	267.0x
451	*Aethionema arabicum*	Ray v. 1.7; SoapDeNovo v. 1.05	Illumina	60x
452	*Arabis alpina*	Newbler v. 2.9; SSPACE v. 1.2; BAMBUS v. 2.33	454; Illumina GAIIx; Sanger	140.0x
453	*Eutrema parvulum*	Newbler v. 2.5; ABySS v. 1.2.1; minimus2	454 GS FLX Titanium; Illumina GA II	6x
454	*Eutrema salsugineum*	ARACHNE v. 20071016_modfied	ABI 3730	7.78x
455	*Sisymbrium irio*	Ray v. 1.7; SoapDeNovo v. 1.05	Illumina	60x
456	*Leavenworthia alabamica*	Ray v. 1.7; SoapDeNovo v. 1.05	Illumina	60x
457	*Tarenaya hassleriana*	SOAPdenovo v. 2.2	Illumina Hiseq	139x
458	*Gossypium arboreum*	ARACHNE v. 20071016_modified	ABI3730	19.7x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
459	*Gossypium raimondii*	ARACHNE v. 20071016_modified	ABI3730	19.7x
460	*Theobroma cacao*	Jeremy Schmutz v. 2011-08-01	454; Illumina; Sanger	15.58x
461	*Aquilaria agallochum*	String Graph Assembler v. 0.10.10	Illumina HiSeq	196.0x
462	*Azadirachta indica*	Velvet v. 1.2.07; MIRA v. 3.4.0.2; cd-hit-est v. 4.5.4	Illumina HiSeq	21.0x nuclear; 196.0x chloroplast; 587.0x mitochondrion
463	*Citrus clementine*	ARACHNE v. 2007101641HA	ABI 3739	6.96x
464	*Citrus sinensis*	SOAPdenovo v. April 2011	Illumina	214.0x
465	*Vitis vinifera*	?	Sanger	12x
466	*Amaranthus hypochondriacus*	SOAPdenovo v. 1.05	Illumina GAIIx	105.0x
467	*Amaranthus tuberculatus*	?	454	6x
468	*Beta vulgaris subsp. vulgaris*	Newbler v. 2.6	454; Illumina GAIIx; Illumina HiSeq; Sanger	30.0x
469	*Spinacia oleracea*	SOAPdenovo v. 1.05	Illumina HiSeq	60.1x
470	*Dianthus caryophyllus*	SOAPdenovo v. 1.05	454 GS FLX+; Illumina HiSeq1000	604x
471	*Actinidia chinensis*	AllPaths v. LG	Illumina	150.0x
472	*Vaccinium macrocarpon*	CLC NGS Cell v. CLC Genomic Workbench v6.5.1	Illumina GAIIx	20.0x
473	*Diospyros lotus*	Price v. Beta; cap3 v. December-2013	Illumina HiSeq	23.0x
474	*Primula veris*	AllPaths v. 48579	Illumina HiSeq; Illumina MiSeq; PacBio	116.0x
475	*Solanum arcanum*	?	?	?
476	*Solanum habrochaites*	?	?	?
477	*Solanum lycopersicum*	Newbler v. 2.3-PostRelease-01/11/2010	454; Sanger; Illumina; SOLiD	27x
478	*Solanum melongena*	Newbler v. 2.7; SOAPdenovo v. 1.05; PCAP.rep	454 GS FLX+; Illumina HiSeq2000	112x
479	*Solanum pennellii*	?	?	?
480	*Solanum pimpinellifolium*	ABySS v. 1.1.2	Illumina	20x
481	*Solanum tuberosum*	SOAPdenovo v. 1014	Illumina GA2	114x
482	*Capsicum annuum*	SOAPdenovo v. 1.5; SSPACE v. 1.0	Illumina HiSeq; Illumina GAII	186.6x
483	*Nicotiana sylvestris*	SOAPdenovo v. 1.05	Illumina HiSeq	94.0x
484	*Nicotiana tomentosiformis*	SOAPdenovo v. 1.05	Illumina HiSeq	146.0x

No	Species	Assembly Method	Sequencing Technology	Genome Coverage
485	*Fraxinus excelsior*	?	?	?
486	*Penstemon centranthifolius*	Newbler v. 2.8	454	0.005x
487	*Penstemon grinnellii*	Newbler v. 2.8	454	0.005x
488	*Sesamum indicum*	SOAPdenovo v. 1.06	Illumina Hiseq 2000	152.7x
489	*Genlisea aurea*	CLC Genomics Workbench v. 5.5	Illumina HiSeq	300.0x
490	*Mimulus guttatus*	ARACHNE v. 2007101641HA	ABI3730	6.43x
491	*Conyza canadensis*	Newbler v. 2.6; SOAPdenovo v. 1; CLC NGS Cell v. Mar-2013	454; Illumina HiSeq; PacBio	350.0x
ECHINOIDEA				
492	*Lytechinus variegatus*	CABOG; Newbler; ATLAS-Link and ATLAS-GapFill	454; Illumina	13x 454; 21x Illumina
493	*Strongylocentrotus purpuratus*	Atlas v. WGS for Sanger Assembly, Atlas-Link and Atlas-GapFill for SOLiD and Illumina improvement	Sanger; SOLiD; Illumina	8.3x Sanger; 18x SOLiD; 40x Illumina
ASTEROIDEA				
494	*Patiria miniata*	CABOG v. 6.1; Newbler v. 2.3; ATLAS-LINK v. 1.0; ATLAS-GAPFILL v. 2.0	454; Illumina	15.0x 454; 70x Illumina
ENTEROPNEUSTA				
495	*Saccoglossus kowalevskii*	?	ABI	7.0x
TUNICATA				
496	*Ciona intestinalis*	?	?	?
497	*Ciona savignyi*	?	?	?
498	*Botryllus schlosseri*	Celera Assembler v. 7.0; Velvet v. 1.2.03	Illumina HiSeq	400.0x
499	*Oikopleura dioica*	?	?	?
LEPTOCARDII				
500	*Branchiostoma floridae*	?	?	?
CEPHALASPIDOMORPHI				
501	*Lethenteron camtschaticum*	Newbler v. 2.7	454	20.0x
502	*Petromyzon marinus*	Arachne v. 3.2	ABI 3730	5.0x
SARCOPTERYGII				
503	*Latimeria chalumnae*	AllPaths v. R36819	Illumina HiSeq	77.5x

Figure 2: Pipeline for L1 sequence retrieval from full genome data

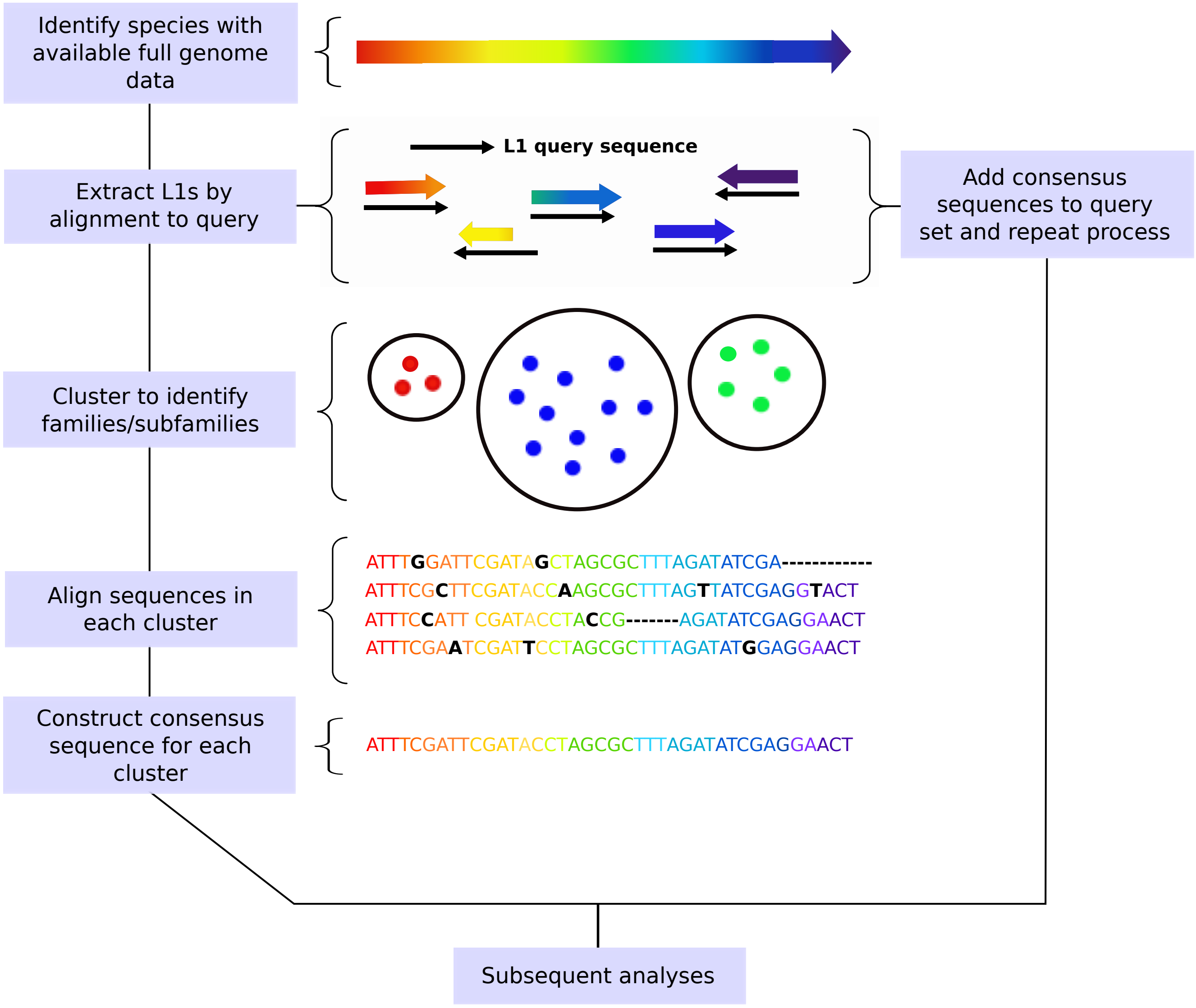


Figure B.2: **L1 extraction pipeline:** Consists of an iterative query-driven method based on sequence similarity. Query L1 sequences from different species are concatenated into one file. This is used to extract similar sequences from each of the 503 genomes. For each species, the L1 hits are clustered and then aligned to generate a consensus sequence for each cluster (subfamily). These consensus sequences are then added to the query file and the process is repeated. This pipeline was initially repeated three times; then, the independent TBLASTN approach was used to generate species-specific queries; then, this pipeline was repeated a final two times: first with the species-specific queries and then with all individual L1s (more than 3 million sequences) .

Figure 3: Categorisation of L1 elements based on ORFs

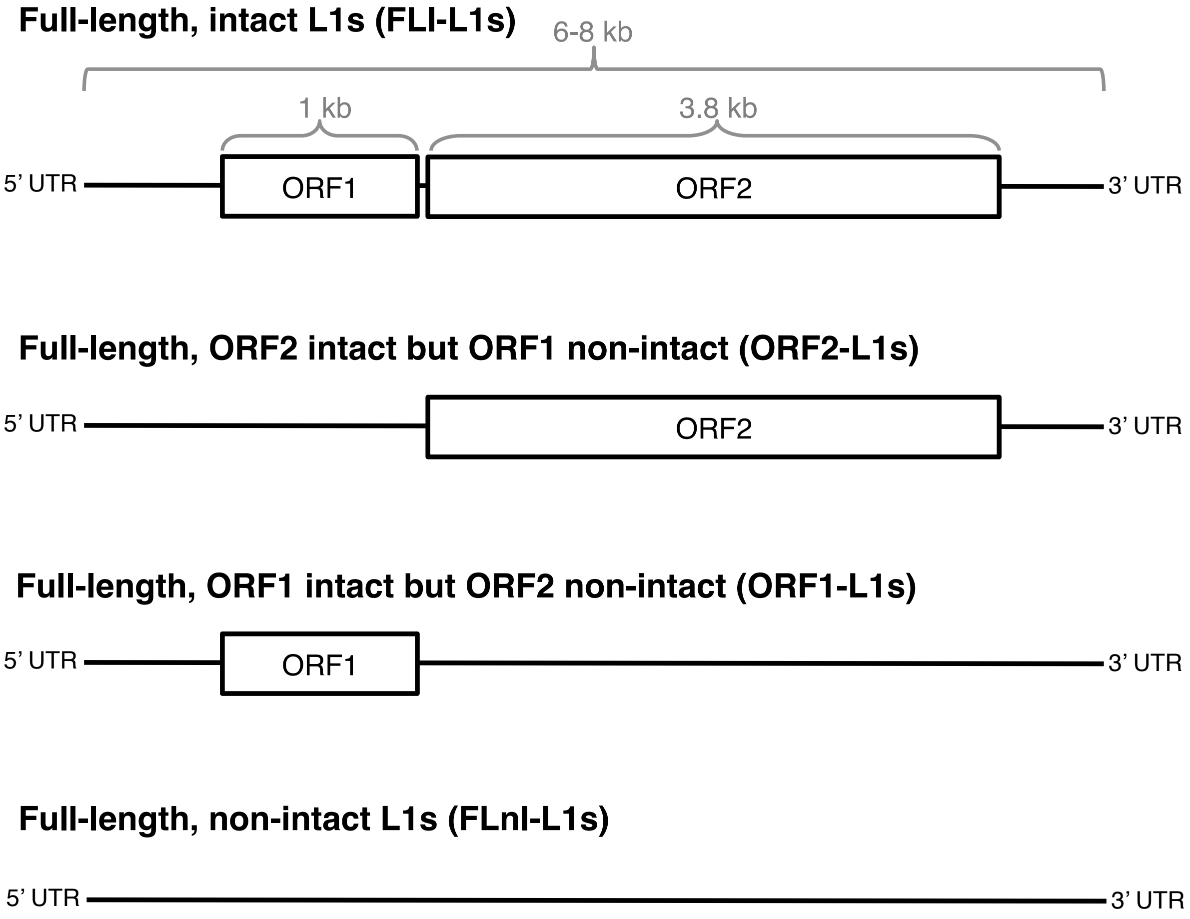


Figure B.3: **Types of L1s:** L1s can be separated into four groups based on the state of their open reading frames, as shown above. An L1 is defined as an 'active' candidate if it is intact in the ORF responsible for reverse transcription, ORF2 (so FLI-L1s and ORF2-L1s are considered active, while ORF1-L1s and FLnI-L1s are not).

Figure 4: Requirements for determining whether an ORF is intact

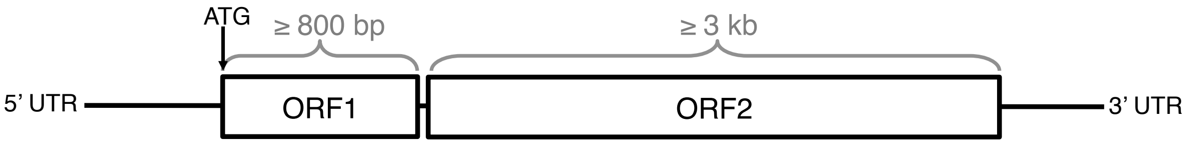


Figure B.4: **Intact ORFs:** An ORF is considered intact if it is at least 80% of the expected length, with a start codon, stop codon, and no debilitating mutations in between (e.g. premature stop codons, large stretches of ambiguous nucleotides). ORF1 has the additional requirement of starting with a methionine start codon (ATG). L1 sequences were extended by 1kb either side during identification of the ORFs, in case the ORF start/stop codons were outside of the extracted L1 sequence. ORF1 and ORF2 were also confirmed by alignment and tested for similarity to known domains, as described in the main text.

B.2 Results

Table 4: L1s in nucleotide nr/htgs databases, found using TBLASTN

Table B.4: **TBLASTN results:** Shows the results for the top hit found in each species. TBLASTN search parameters were default except the e-value was changed to 1e-5. Input was the concatenated ORF1p and ORF2p from 13 full-length L1 elements from Repbase: L1HS (*Homo sapiens*, Hs), TX1 (*Xenopus laevis*, Xl), L1-1_Acar (*Anolis carolinensis*, Ac), L1-3_Dr (*Danio rerio*, Dr), Ag-L1_5 (*Anopheles gambiae*, Ag), L1-1_HM (*Hydra vulgaris*, Hv), ATLINE1 (*Arabidopsis thaliana*, At), Zepp (*Chlorella vulgaris*, Cv), Tx1-3_Spur (*Strongylocentrotus purpuratus*, Sp), Tx1-12_SK (*Saccoglossus kowalevskii*, Sk), L1-1a_Cis (*Ciona savignyi*, Cs), Tx1-1_BF (*Branchiostoma floridae*, Bf), L1-3_LCh (*Latimeria chalumnae*, Lc). These queries span all (available) orders/clades, and consist of both the typical mammalian L1 group and diverse Tx1 group. Databases searched were the NCBI non-redundant nucleotide collection (nr) and high thoroughput genomic sequences (htgs). The table below shows the number of hits that were produced from each query and the statistics for the top hit. bitscore = max score; qcovs = query coverage; evalue = e-value; pident = percentage identity; qstart = start coordinate of the hit on the query sequence; qend = query end coordinate (this is important because true hits are more likely to overlap the reverse-transcriptase domain in ORF2, which is thought to be the most conserved and vital for retrotransposition).

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
MAMMALIA								
1	*Tachyglossus aculeatus* (taxid:9261)	44 Lc, 25 Cv, 22 Ag, 18 Sp, 14 Hs, 13 At, 12 Ac, 9 Cs, 3 Bf, 1 Sk, 1 Xl, 1 Hv	102	19	1e-21	28.40	852	1171
2	*Ornithorhynchus anatinus* (taxid:9258)	403 Lc, 229 Cv, 221 Hs, 191 Ag, 150 Ac, 135 Cs, 129 Sp, 114 At, 6 Bf, 2 Sk, 2 Dr, 2 Hv	110	20	4e-23	28.49	838	1174
3	*Monodelphis domestica* (taxid:13616)	14806 Hs, 4866 Ac, 4658 Lc, 2448 Dr, 2251 Hv, 2240 Bf, 2050 Sp, 2031 Xl, 1917 Cs, 1910 Sk, 1486 At, 1321 Cv, 1214 Ag	1425	91	0.0	53.28	339	1602
4	*Macropus eugenii* (taxid:9315)	4346 Hs, 767 Ac, 633 Lc, 295 Dr, 288 Sp, 226 Hv, 196 Bf, 184 Cs, 160 Xl, 159 Sk, 158 At, 118 Cv, 91 Ag	1442	90	0.0	53.25	342	1602
5	*Sarcophilus harrisii* (taxid:9305)	775 Hs, 227 Ac, 166 Lc, 65 Dr, 37 At, 36 Cs, 34 Sk, 33 Xl, 32 Sp, 32 Bf, 23 Hv, 20 Cv, 15 Ag	1059	91	0.0	56.62	341	1239
6	*Dasypus novemcinctus* (taxid:9361)	15995 Hs, 3588 Lc, 2319 Ac, 936 Hv, 910 Sp, 899 Dr, 814 Bf, 783 Xl, 769 Sk, 679 Cs, 630 At, 453 Ag, 362 Cv	1932	100	0.0	57.31	1	1610

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
7	*Choloepus hoffmanni* (taxid:9358)	10120 Hs, 2477 Lc, 1766 Ac, 985 Hv, 971 Dr, 941 Sk, 893 Bf, 835 Xl, 797 Sp, 797 Cs, 732 At, 647 Ag, 515 Cv	1941	99	0.0	57.76	4	1610
8	*Chrysochloris asiatica* (taxid:185453)	9 Hs, 7 Lc	154	6	2e-42	66.67	1141	1239
9	*Echinops telfairi* (taxid:9371)	2590 Hs, 463 Lc, 362 Ac, 186 Sp, 152 Dr, 141 Bf, 125 Hv, 116 Xl, 114 At, 111 Sk, 111 Cs, 80 Ag, 49 Cv	1801	96	0.0	53.01	54	1609
10	*Orycteropus afer afer* (taxid:1230840)	1 Hs	60.5	2	5e-09	63.41	969	1009
11	*Elephantulus edwardii* (taxid:28737)	7 Hs, 5 Lc	138	6	7e-37	61.22	1142	1239
12	*Trichechus manatus latirostris* (taxid:127582)	5 Hs	73.6	4	5e-13	60.61	1195	1260
13	*Procavia capensis* (taxid:9813)	10631 Hs, 2836 Lc, 2326 Ac, 1110 Dr, 1085 Sp, 1053 Hv, 1031 Xl, 987 Bf, 962 Sk, 922 Cs, 801 At, 671 Ag, 426 Cv	1660	94	0.0	59.86	324	1604
14	*Loxodonta africana* (taxid:9785)	13775 Hs, 5002 Lc, 3725 Ac, 2985 Dr, 2783 Sp, 2550 Xl, 2350 Hv, 2161 Cs, 2079 Bf, 1857 Sk, 1749 At, 1434 Ag, 803 Cv	1696	98	0.0	62.29	339	1610
15	*Erinaceus europaeus* (taxid:9365)	41 Hs, 7 Ac, 4 Lc, 3 Sp, 3 Bf, 3 Sk, 3 Dr, 3 Cs, 2 Cv, 2 Xl, 2 Hv, 2 Ag, 2 At	687	65	0.0	58.48	1056	1609
16	*Sorex araneus* (taxid:42254)	4496 Hs, 908 Ac, 801 Lc, 377 Dr, 370 Cs, 329 Bf, 328 Sp, 319 Hv, 308 Xl, 275 At, 272 Sk, 248 Ag, 226 Cv	1879	99	0.0	56.59	11	1612
17	*Condylura cristata* (taxid:143302)	92 Hs, 8 Lc	111	12	3e-24	38.91	1291	1495
18	*Pteropus alecto* (taxid:9402)	207 Hs, 2 Lc, 1 Ac	172	24	9e-44	33.33	1136	1523
19	*Pteropus vampyrus* (taxid:132908)	4953 Hs, 585 Lc, 526 Ac, 203 Hv, 201 Bf, 182 Dr, 170 Sk, 168 Sp, 160 Cs, 153 Xl, 137 At, 77 Ag, 65 Cv	804	99	0.0	52.99	837	1612

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
20	*Eidolon helvum* (taxid:77214)	2 Hs	62.0	5	1e-12	40.40	1401	1495
21	*Megaderma lyra* (taxid:9413)	No significant similarity found						
22	*Rhinolophus ferrumequinum* (taxid:59479)	8716 Hs, 1304 Lc, 1156 Ac, 470 Dr, 448 Bf, 434 Hv, 430 Xl, 422 Sk, 396 Sp, 367 Cs, 339 At, 281 Ag, 252 Cv	1808	98	0.0	67.51	345	1612
23	*Pteronotus parnellii* (taxid:59476)	No significant similarity found						
24	*Eptesicus fuscus* (taxid:29078)	3 Hs, 1 Lc	144	13	8e-35	40.67	112	319
25	*Myotis brandtii* (taxid:109478)	173 Hs, 1 Lc, 1 Dr, 1 At	147	18	5e-35	35.89	36	318
26	*Myotis davidii* (taxid:225400)	122 Hs, 1 Lc	145	17	9e-35	35.92	36	318
27	*Myotis lucifugus* (taxid:59463)	4484 Hs, 693 Lc, 627 Ac, 184 Dr, 182 Hv, 181 Sk, 177 Sp, 161 Xl, 153 Bf, 146 Cs, 138 At, 121 Ag, 97 Cv	1975	92	0.0	63.00	113	1612
28	*Ceratotherium simum simum* (taxid:73337)	72 Hs	103	5	9e-25	55.17	1201	1287
29	*Equus przewalskii* (taxid:9798)	733 Hs, 14 Lc, 12 Ac, 3 Cv, 3 Sp, 3 Bf, 3 Sk, 3 Cs, 3 At, 2 Dr, 2 Hv, 1 Xl, 1 Ag	465	39	0.0	67.61	970	1287
30	*Equus caballus* (Thoroughbred) (taxid:9796)	8022 Hs, 1583 Lc, 1570 Ac, 559 Dr, 509 Sp, 499 Bf, 490 Hv, 471 Sk, 458 Cs, 446 Xl, 376 At, 267 Ag, 237 Cv	1932	98	0.0	58.40	18	1613
31	*Equus caballus* (Mongolian) (taxid:9796)	As above (same taxid)	"	"	"	"	"	"
32	*Manis pentadactyla* (taxid:143292)	No significant similarity found						
33	*Felis catus* (taxid:9685)	13549 Hs, 2551 Lc, 2138 Ac, 798 Sp, 783 Dr, 774 Hv, 723 Sk, 718 Bf, 691 Xl, 621 Cs, 532 At, 419 Ag, 377 Cv	1897	91	0.0	61.96	139	1613
34	*Panthera tigris altaica* (taxid:74533)	116 Hs, 1 Lc	155	18	1e-37	36.15	22	312

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
35	*Canis lupus familiaris* (taxid:9615)	25106 Hs, 6695 Lc, 5932 Ac, 2726 Dr, 2352 Xl, 2327 Sk, 2258 Hv, 2248 Cs, 2183 Sp, 2011 Bf, 1732 At, 1535 Ag, 1386 Cv	1900	93	0.0	60.53	112	1612
36	*Ursus maritimus* (taxid:29073)	339 Hs, 4 Lc, 1 Ac	152	24	3e-57	35.88	1296	1584
37	*Ailuropoda melanoleuca* (taxid:9646)	610 Hs, 17 Lc, 13 Ac, 4 Hv, 3 Bf, 3 Dr, 2 Cv, 2 Sp, 2 Sk, 2 Xl, 2 Cs, 2 Ag, 2 At	853	73	0.0	52.33	142	963
38	*Leptonychotes weddellii* (taxid:9713)	20 Hs, 1 Lc	148	18	2e-36	36.72	22	321
39	*Odobenus rosmarus divergens* (taxid:9708)	11 Hs, 2 Lc	149	19	6e-36	36.14	6	321
40	*Mustela putorius furo* (taxid:9669)	4945 Hs, 918 Lc, 797 Ac, 381 Dr, 316 Xl, 312 Hv, 291 Bf, 290 Cs, 283 Sk, 268 Sp, 219 Ag, 219 At, 195 Cv	1805	94	0.0	66.33	343	1613
41	*Camelus dromedarius* (taxid:9838)	77 Hs	117	5	2e-29	65.85	1531	1612
42	*Camelus ferus* (taxid:419612)	69 Hs, 1 Lc	94.0	9	1e-18	35.15	1415	1570
43	*Vicugna pacos* (taxid:30538)	5666 Hs, 592 Ac, 588 Lc, 166 Dr, 161 Hv, 148 Sp, 131 Bf, 131 Xl, 122 Cs, 111 Sk, 99 At, 79 Ag, 64 Cv	1766	92	0.0	64.54	327	1612
44	*Sus scrofa* (Duroc) (taxid:9823)	21633 Hs, 5697 Lc, 4748 Ac, 2334 Dr, 2209 Sp, 2040 Hv, 1987 Xl, 1938 Bf, 1900 Cs, 1823 Sk, 1711 At, 1382 Ag, 1152 Cv	1837	97	0.0	67.87	343	1612
45	*Sus scrofa* (Tibetan) (taxid:9823)	As above (same taxid)	"	"	"	"	"	"
46	*Sus scrofa* (Ellegaard Gottingen minipig) (taxid:9823)	As above (same taxid)	"	"	"	"	"	"
47	*Balaenoptera acutorostrata scammoni* (taxid:310752)	180 Hs, 5 Lc, 2 Ac	133	6	8e-31	60.38	1474	1577

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
48	*Physeter catodon* (taxid:9755)	145 Hs, 11 Lc, 11 Ac, 1 Cs	281	11	2e-85	68.06	1044	1234
49	*Lipotes vexillifer* (taxid:118797)	9 Hs, 1 Sp, 1 Lc, 1 Ac	194	18	2e-53	55.00	856	1048
50	*Tursiops truncatus* (taxid:9739)	18 Hs	132	5	1e-31	65.22	1408	1499
51	*Orcinus orca* (taxid:9733)	17 Hs	136	14	3e-40	47.50	1328	1482
52	*Pantholops hodgsonii* (taxid:59538)	59 Hs, 1 Lc, 1 Cs, 1 Ac	498	22	2e-156	63.23	1255	1613
53	*Capra hircus* (taxid:9925)	228 Hs, 23 Sp, 20 Lc, 16 Ac, 12 Bf, 9 Cs, 8 Xl, 8 Hv, 6 Dr, 4 Sk, 3 Ag, 3 At, 2 Cv	534	92	0.0	59.44	1184	1612
54	*Ovis aries* (Texel) (taxid:9940)	2033 Hs, 354 Sp, 339 Ac, 284 Lc, 256 Bf, 230 Cs, 215 Dr, 135 Xl, 115 Hv, 69 Sk, 61 Ag, 52 At, 16 Cv	786	63	0.0	56.10	907	1612
55	*Ovis aries musimon* (taxid:9938)	1220 Hs, 58 Lc, 32 Ac, 15 Dr, 8 Sp, 5 Bf, 4 Sk, 4 Xl, 4 Cs, 4 Hv	206	57	2e-108	41.27	645	1058
56	*Bubalus bubalis* (taxid:89462)	579 Hs, 37 Sp, 37 Ac, 26 Bf, 24 Lc, 24 Dr, 19 Cs, 16 Xl, 14 Hv, 12 Sk, 9 At, 5 Ag, 2 Cv	170	67	1e-98	47.58	896	1116
57	*Bison bison bison* (taxid:43346)	558 Hs, 12 Lc, 4 Ac, 1 Sk	204	41	9e-53	36.19	757	1253
58	*Bos mutus* (taxid:72004)	89 Hs, 8 Lc, 3 Ac, 1 Sp, 1 Xl, 1 Cs, 1 Hv, 1 Ag	236	14	2e-80	65.88	1363	1532
59	*Bos indicus* (taxid:9915)	10 Hs	90.5	14	2e-19	33.90	1094	1322
60	*Bos taurus* (taxid:9913)	16949 Hs, 6654 Sp, 6325 Ac, 4847 Lc, 4284 Cs, 4177 Bf, 4142 Dr, 3800 Xl, 3408 Hv, 2534 Sk, 2528 Ag, 2066 At, 1338 Cv	2023	92	0.0	63.53	113	1612
61	*Ochotona princeps* (taxid:9978)	1302 Hs, 166 Ac, 155 Lc, 57 Bf, 52 Hv, 51 Dr, 50 Cs, 47 Sk, 45 Sp, 41 At, 40 Xl, 39 Cv, 37 Ag	1844	99	0.0	55.35	7	1610
62	*Oryctolagus cuniculus* (taxid:9986)	6268 Hs, 1271 Lc, 1106 Ac, 536 Dr, 492 Hv, 461 Sk, 441 Xl, 425 Bf, 420 Cs, 409 Sp, 315 At, 297 Ag, 232 Cv	2068	97	0.0	62.19	43	1613

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
63	*Ictidomys tridecemlineatus* (taxid:43179)	5252 Hs, 775 Lc, 595 Ac, 286 Dr, 259 Hv, 232 Sk, 228 Sp, 226 Bf, 224 Cs, 214 Xl, 208 At, 160 Ag, 130 Cv	1972	99	0.0	57.68	3	1610
64	*Heterocephalus glaber* (taxid:10181)	60 Hs, 1 Lc, 1 Ac	157	57	3e-59	35.67	723	1061
65	*Fukomys damarensis* (taxid:885580)	153 Hs, 2 Lc	148	57	3e-35	33.08	170	557
66	*Cavia aperea* (taxid:37548)	No significant similarity found						
67	*Cavia porcellus* (taxid:10141)	9718 Hs, 2049 Lc, 1749 Ac, 841 Dr, 819 Bf, 784 Hv, 712 Sp, 683 Cs, 658 Sk, 652 Xl, 550 At, 440 Ag, 416 Cv	1843	97	0.0	56.98	47	1610
68	*Chinchilla lanigera* (taxid:34839)	248 Hs, 9 Lc	152	12	7e-41	51.83	790	953
69	*Octodon degus* (taxid:10160)	9 Hs	110	11	3e-24	34.67	189	369
70	*Dipodomys ordii* (taxid:10020)	24 Hs, 1 Lc, 1 Ac	92.0	6	3e-21	43.86	762	872
71	*Jaculus jaculus* (taxid:51337)	1 Hs	50.1	14	9e-06	22.76	340	577
72	*Nannospalax galili* (taxid:1026970)	27 Hs, 1 Lc	132	12	6e-31	38.28	114	321
73	*Mesocricetus auratus* (taxid:10036)	63 Hs, 6 Lc, 4 Ac, 3 Xl, 3 Hv, 2 Sk	291	14	2e-88	60.94	585	817
74	*Cricetulus griseus* (taxid:10029)	170 Hs, 40 Ac, 38 Lc, 25 Sp, 24 Hv, 20 Dr, 19 Bf, 19 Sk, 19 Cs, 17 Xl, 16 At, 12 Ag, 10 Cv	449	41	5e-128	55.31	935	1347
75	*Microtus ochrogaster* (taxid:79684)	467 Hs, 85 Ac, 81 Lc, 30 Dr, 27 Sp, 25 Sk, 22 Bf, 22 Hv, 21 Cs, 17 Ag, 17 At, 16 Xl, 12 Cv	928	91	0.0	62.13	930	1612
76	*Peromyscus maniculatus bairdii* (taxid:230844)	4 Hs	67.4	4	1e-10	46.75	1528	1604
77	*Rattus norvegicus* (taxid:10116)	18946 Hs, 9104 Lc, 7388 Ac, 5445 Dr, 4065 Sp, 4034 Hv, 3797 Xl, 3674 Sk, 3627 Bf, 3531 Cs, 3288 At, 2764 Ag, 2563 Cv	1890	92	0.0	58.75	113	1612

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
78	*Mus musculus* (taxid:10090)	25459 Hs, 7113 Ac, 6745 Lc, 3856 Dr, 3624 Sp, 3032 Sk, 2966 Bf, 2950 Hv, 2920 Cs, 2832 Xl, 2653 At, 2357 Ag, 1957 Cv	2676	100	0.0	94.99	260	1613
79	*Tupaia belangeri* (taxid:37347)	4584 Hs, 616 Lc, 494 Ac, 210 Dr, 182 Sp, 178 Bf, 175 Hv, 157 Sk, 155 Xl, 144 Cs, 142 At, 103 Cv, 102 Ag	1959	99	0.0	58.45	1	1607
80	*Tupaia chinensis* (taxid:246437)	37 Hs, 1 Lc	177	10	3e-45	54.76	113	280
81	*Galeopterus variegatus* (taxid:482537)	55 Hs, 6 Sp, 6 Lc, 6 Dr, 6 Ac, 1 Cv	289	11	2e-88	69.11	1044	1234
82	*Otolemur garnettii* (taxid:30611)	12768 Hs, 2627 Lc, 2550 Ac, 1077 Sp, 961 Bf, 955 Hv, 954 Dr, 827 Sk, 810 Cs, 804 Xl, 747 At, 667 Ag, 558 Cv	1958	100	0.0	57.50	1	1608
83	*Microcebus murinus* (taxid:30608)	11443 Hs, 2135 Lc, 1934 Ac, 855 Bf, 806 Hv, 794 Dr, 787 Sp, 748 Cs, 671 Sk, 661 At, 640 Xl, 504 Ag, 482 Cv	2067	99	0.0	60.80	1	1608
84	*Tarsius syrichta* (taxid:9478)	3 Hs	56.6	8	3e-07	36.69	1440	1574
85	*Callithrix jacchus* (taxid:9483)	26093 Hs, 6413 Lc, 5613 Ac, 3014 Hv, 2667 Dr, 2655 Bf, 2397 Xl, 2207 Sp, 2169 Sk, 2120 Cs, 2109 At, 1644 Cv, 1429 Ag	2712	99	0.0	80.20	1	1612
86	*Saimiri boliviensis boliviensis* (taxid:39432)	9043 Hs, 1416 Lc, 1210 Ac, 467 Hv, 461 Bf, 454 Dr, 443 Sp, 402 At, 390 Cs, 338 Xl, 332 Sk, 254 Cv, 210 Ag	2359	99	0.0	80.03	1	1433
87	*Rhinopithecus roxellana* (taxid:61622)	220 Hs, 8 Lc, 1 Ac	240	21	1e-65	69.71	1259	1432
88	*Nasalis larvatus* (taxid:43780)	7 Hs, 1 Ac	160	4	3e-46	88.75	1418	1497
89	*Chlorocebus sabaeus* (taxid:60711)	82 Hs, 2 Lc, 2 Ac	224	22	9e-68	45.71	1305	1613

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
90	*Macaca fascicularis* (taxid:9541)	44100 Hs, 16722 Lc, 11509 Ac, 4206 Dr, 4120 Hv, 4024 Sp, 3976 Bf, 3584 Cs, 3504 At, 3103 Xl, 2854 Sk, 1300 Cv, 1051 Ag	3006	100	0.0	89.17	1	1613
91	*Macaca mulatta* (taxid:9544)	22956 Hs, 6008 Lc, 5503 Ac, 2877 Hv, 2625 Dr, 2613 Bf, 2426 Sp, 2399 Xl, 2389 Cs, 2283 Sk, 2162 At, 1658 Cv, 1351 Ag	3019	100	0.0	89.54	1	1613
92	*Papio anubis* (taxid:9555)	17213 Hs, 3397 Lc, 2965 Ac, 1396 Dr, 1367 Hv, 1285 Bf, 1240 Sp, 1159 Cs, 1131 Xl, 1082 Sk, 1079 At, 750 Cv, 614 Ag	3019	100	0.0	89.60	1	1613
93	*Nomascus leucogenys* (taxid:61853)	14492 Hs, 2904 Lc, 2692 Ac, 1181 Hv, 1125 Dr, 1105 Sp, 1069 Bf, 982 Cs, 958 Xl, 936 Sk, 902 At, 682 Cv, 534 Ag	3104	100	0.0	92.91	1	1613
94	*Pongo abelii* (taxid:9601)	27703 Hs, 6718 Lc, 6214 Ac, 3125 Hv, 3054 Dr, 2854 Bf, 2770 Sp, 2622 Xl, 2563 Cs, 2508 Sk, 2363 At, 1893 Cv, 1502 Ag	3199	100	0.0	95.66	1	1613
95	*Gorilla gorilla gorilla* (taxid:9595)	11864 Hs, 2445 Lc, 2246 Ac, 954 Dr, 950 Bf, 937 Hv, 901 Sp, 801 Xl, 794 Cs, 759 At, 751 Sk, 556 Cv, 457 Ag	3221	100	0.0	96.76	1	1613
96	*Pan paniscus* (taxid:9597)	973 Hs, 77 Lc, 60 Ac, 22 Sp, 13 Dr, 12 Hv, 12 At, 11 Cs, 10 Bf, 9 Cv, 9 Xl, 9 Ag, 8 Sk	3205	100	0.0	96.45	1	1613
97	*Pan troglodytes* (taxid:9598)	23666 Hs, 6383 Lc, 6104 Ac, 2870 Hv, 2750 Dr, 2644 Bf, 2432 Cs, 2430 Sp, 2401 Sk, 2390 Xl, 2221 At, 1617 Cv, 1355 Ag	3225	100	0.0	96.94	1	1613
98	*Homo sapiens* (taxid:9606)	15381 Hs, 3318 Lc, 3209 Ac, 1790 Xl, 1691 Hv, 1672 At, 1647 Sp, 1569 Sk, 1558 Bf, 1558 Dr, 1500 Cs, 1418 Cv, 1397 Ag	3275	100	0.0	98.41	1	1613
SAUROPSIDA								

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
99	*Apalone spinifera* (taxid:55534)	No significant similarity found						
100	*Pelodiscus sinensis* (taxid:13735)	24 Xl, 18 Sp, 18 Bf, 16 Sk, 16 Hs, 16 Lc, 16 Hv, 14 Cs, 14 Ac, 13 Dr, 11 Cv, 7 Ag, 7 At	241	8	2e-63	61.50	1326	1503
101	*Chelonia mydas* (taxid:8469)	29 Xl, 9 Sk, 4 Bf, 3 Sp, 3 Hv, 3 Ac, 2 Hs, 2 Dr, 2 Cs, 2 At, 1 Cv, 1 Lc, 1 Ag	142	16	1e-33	45.62	813	971
102	*Chrysemys picta bellii* (taxid:8478)	44 Ac, 41 Xl, 40 Sp, 36 Hs, 35 Bf, 35 Lc, 28 Cs, 24 Cv, 18 Sk, 14 Dr, 13 Ag, 13 At, 10 Hv	160	8	2e-38	45.81	876	1053
103	*Struthio camelus australis* (taxid:441894)	2 Cv, 2 Bf, 2 Hs, 2 Ac, 1 Sp, 1 Xl, 1 Lc, 1 Dr, 1 Cs, 1 Ag, 1 At	77.8	18	6e-14	24.68	852	1152
104	*Tinamus guttatus* (taxid:94827)	3 Sp, 3 Bf, 3 Hs, 3 Lc, 3 Cs, 3 Ac, 2 Cv, 2 Sk, 2 Xl, 2 Dr, 2 Hv, 2 Ag, 1 At	104	26	3e-22	28.84	775	1185
105	*Anas platyrhynchos* (taxid:8839)	7 Lc, 6 Bf, 3 Sp, 3 Hs, 3 Ac, 2 Cv, 1 Dr, 1 Cs, 1 Hv, 1 At	92.0	13	1e-17	27.95	935	1154
106	*Lyrurus tetrix tetrix* (taxid:162951)	No significant similarity found						
107	*Gallus gallus* (taxid:9031)	935 Bf, 934 Lc, 877 Sp, 790 Ac, 787 Hs, 641 Cs, 385 Dr, 327 Cv, 225 At, 200 Sk, 161 Ag, 149 Xl, 39 Hv	838	92	0.0	61.81	341	992
108	*Coturnix japonica* (taxid:93934)	11 Sp, 6 Lc, 5 Cv, 4 Bf, 1 Hs, 1 Dr, 1 Cs, 1 Ac, 1 At	73.2	12	6e-14	29.76	852	1052
109	*Meleagris gallopavo* (taxid:9103)	16 Sp, 16 Lc, 15 Bf, 15 Cs, 13 Cv, 13 Hs, 13 Ac, 11 At, 10 Sk, 10 Dr, 10 Ag, 8 Xl, 3 Hv	107	17	2e-22	28.67	830	1120
110	*Colinus virginianus* (taxid:9014)	No significant similarity found						
111	*Acanthisitta chloris* (taxid:57068)	1 Bf	51.6	8	9e-07	27.14	931	1068
112	*Manacus vitellinus* (taxid:328815)	3 Sp, 3 Bf, 3 Hs, 3 Lc, 3 Ac, 2 Cs, 1 Cv, 1 Dr, 1 Ag	80.1	14	3e-15	27.76	817	1057
113	*Zonotrichia albicollis* (taxid:44394)	2 Bf, 1 Lc, 1 Cs, 1 Ac	68.9	17	5e-11	24.91	658	924
114	*Geospiza fortis* (taxid:48883)	No significant similarity found						

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
115	*Serinus canaria* (taxid:9135)	3 Sp, 3 Bf, 2 Cv, 2 Xl, 2 Hs, 2 Lc, 2 Dr, 2 Cs, 2 Ac, 2 Ag, 2 At, 1 Sk, 1 Hv	105	23	5e-23	25.46	658	1021
116	*Taeniopygia guttata* (taxid:59729)	3 Hs, 1 Sp, 1 Bf, 1 Xl, 1 Lc, 1 Dr, 1 Cs, 1 Ac	203	11	5e-61	54.50	137	325
117	*Ficedula albicollis* (taxid:59894)	1 Lc, 1 Cs	56.6	8	3e-12	29.45	836	977
118	*Pseudopodoces humilis* (taxid:181119)	No significant similarity found						
119	*Corvus brachyrhynchos* (taxid:85066)	3 Ac, 2 Hs, 2 Lc, 2 Cs, 1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Dr, 1 At	93.2	13	4e-19	32.88	836	1052
120	*Corvus cornix cornix* (taxid:932674)	9 Hs, 3 Lc, 2 Cs, 2 Ac, 1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Dr, 1 At	95.9	13	3e-19	34.23	836	1052
121	*Ara macao* (taxid:176014)	No significant similarity found						
122	*Amazona vittata* (taxid:241585)	No significant similarity found						
123	*Melopsittacus undulatus* (taxid:13146)	5 Sp, 5 Bf, 5 Hs, 5 Lc, 5 Ac, 4 Cv, 3 Cs, 3 At, 2 Xl, 1 Dr	70.9	13	8e-12	26.29	931	1154
124	*Nestor notabilis* (taxid:176057)	2 Hs, 2 Cs, 2 Ac, 1 Sp, 1 Bf, 1 Xl, 1 Lc, 1 Dr, 1 At	82.4	22	7e-16	21.41	669	1021
125	*Falco cherrug* (taxid:345164)	1 Ac	52.4	12	4e-06	22.92	669	856
126	*Falco peregrinus* (taxid:8954)	No significant similarity found						
127	*Cariama cristata* (taxid:54380)	4 Hs, 3 Bf, 3 Lc, 3 Ac, 2 Sp, 2 Cs	84.3	12	2e-16	27.70	847	1057
128	*Merops nubicus* (taxid:57421)	3 Sp, 3 Bf, 3 Hs, 3 Ac, 2 Lc, 2 Cs, 1 Cv, 1 At	85.5	15	2e-17	29.02	809	1057
129	*Picoides pubescens* (taxid:118200)	6 Sp, 5 Bf, 5 Lc, 5 At, 4 Cv, 4 Sk, 4 Xl, 4 Hs, 4 Dr, 4 Cs, 4 Hv, 4 Ac, 4 Ag	107	26	5e-24	23.57	658	1061
130	*Buceros rhinoceros silvestris* (taxid:175836)	1 Sp, 1 Bf, 1 Lc, 1 Ac	64.3	13	9e-11	26.05	835	1049
131	*Apaloderma vittatum* (taxid:57397)	4 Sp, 4 Bf, 4 Hs, 4 Lc, 3 Ac, 2 Cv, 2 Sk, 2 Xl, 2 Dr, 2 Cs, 2 At, 1 Hv, 1 Ag	77.4	12	3e-15	29.47	848	1052
132	*Leptosomus discolor* (taxid:188344)	4 Bf, 3 Sp, 3 Lc, 2 Hs, 2 Cs, 2 Ac, 2 Ag, 1 Cv, 1 Xl, 1 At	79.7	21	6e-15	20.23	684	1021

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
133	*Haliaeetus albicilla* (taxid:8969)	3 Lc, 3 Ac, 2 Sp, 2 Bf, 2 Hs, 2 Cs, 1 Cv, 1 Ag	70.9	11	9e-13	30.89	860	1048
134	*Haliaeetus leucocephalus* (taxid:52644)	3 Sp, 3 Ac, 2 Bf, 2 Sk, 2 Hs, 2 Cs, 2 Ag, 2 At, 1 Xl, 1 Lc, 1 Dr, 1 Hv	100	19	6e-21	25.40	719	1018
135	*Aquila chrysaetos canadensis* (taxid:216574)	3 Ac, 2 Sp, 2 Cs, 1 Cv, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 At	95.5	25	5e-19	23.50	658	1057
136	*Cathartes aura* (taxid:43455)	No significant similarity found						
137	*Tyto alba* (taxid:56313)	3 Sp, 2 Bf, 2 Hs, 2 Lc, 1 Cv, 1 Sk, 1 Dr, 1 Cs, 1 Ac, 1 Ag	69.7	18	9e-12	25.96	669	947
138	*Colius striatus* (taxid:57412)	2 Cv, 2 Bf, 2 Lc, 1 Sp, 1 Xl, 1 Hs, 1 Dr, 1 Cs, 1 Ac	75.5	24	7e-14	20.57	694	1093
139	*Charadrius vociferus* (taxid:50402)	1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 Cs, 1 Hv, 1 Ac, 1 Ag	89.4	22	1e-17	22.52	845	1208
140	*Balearica regulorum gibbericeps* (taxid:100784)	1 Hs, 1 Cs, 1 Ac	52.4	9	3e-07	26.75	879	1032
141	*Chlamydotis macqueenii* (taxid:187382)	4 Sp, 4 Lc, 3 Bf, 3 Hs, 2 Cv, 2 Cs, 1 Sk, 1 Xl, 1 Dr, 1 Hv, 1 Ac, 1 Ag	81.6	14	7e-16	28.23	817	1057
142	*Cuculus canorus* (taxid:55661)	7 Sp, 7 Bf, 6 Cv, 6 Sk, 6 Hs, 6 Lc, 6 Dr, 6 Cs, 6 Hv, 6 Ac, 6 At, 5 Xl, 5 Ag	97.4	25	6e-21	24.15	658	1057
143	*Fulmarus glacialis* (taxid:30455)	3 Bf, 2 Ac, 1 Sp, 1 Lc, 1 Cs	63.9	9	4e-10	26.32	752	903
144	*Aptenodytes forsteri* (taxid:9233)	3 Sp, 3 Bf, 3 Hs, 3 Lc, 3 Ac, 2 Dr, 2 Cs, 1 Cv, 1 Sk, 1 Xl, 1 Hv, 1 At	87.0	13	7e-17	28.76	836	1057
145	*Pygoscelis adeliae* (taxid:9238)	6 Sp, 6 Bf, 6 Hs, 6 Ac, 5 Cv, 5 Lc, 5 Cs, 4 At, 2 Sk, 2 Xl, 2 Dr	94.0	20	4e-19	22.46	624	947
146	*Phalacrocorax carbo* (taxid:9209)	No significant similarity found						
147	*Pelecanus crispus* (taxid:36300)	2 Sp, 2 Bf, 2 Hs, 2 Lc, 2 Ac, 1 Cv, 1 Xl, 1 Dr, 1 Cs, 1 At	77.0	15	4e-14	23.23	833	1085
148	*Nipponia nippon* (taxid:128390)	11 Bf, 11 Lc, 10 Sp, 9 Hs, 5 Ac, 3 Cv, 2 Cs, 2 Ag, 1 Xl, 1 Dr, 1 At	80.9	16	2e-15	25.96	930	1199
149	*Egretta garzetta* (taxid:188379)	2 Bf, 2 Hs, 2 Lc, 2 Cs, 2 Ac, 1 Cv, 1 Sp, 1 Sk, 1 Xl, 1 Dr, 1 At	84.0	24	3e-16	20.68	845	1244

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
150	*Phaethon lepturus* (taxid:97097)	No significant similarity found						
151	*Gavia stellata* (taxid:37040)	5 Sp, 5 Bf, 5 Ac, 4 Hs, 4 Lc, 3 Cs, 2 Cv, 2 Ag, 1 Sk, 1 Xl, 1 Dr, 1 At	92.0	14	1e-19	31.02	813	1052
152	*Tauraco erythrolophus* (taxid:121530)	1 Cs	55.8	7	6e-08	27.69	914	1040
153	*Opisthocomus hoazin* (taxid:30419)	6 Lc, 6 Ac, 5 Sp, 5 Hs, 4 Bf, 4 Cs, 3 Cv, 3 Xl, 3 Dr, 3 Ag, 3 At, 1 Sk, 1 Hv	95.1	19	4e-20	24.75	660	954
154	*Columba livia* (taxid:8932)	No significant similarity found						
155	*Pterocles gutturalis* (taxid:240206)	2 Ac, 2 Hs, 1 Xl, 1 Dr, 1 At	70.1	22	3e-12	22.75	673	1021
156	*Calypte anna* (taxid:9244)	1 Bf, 1 Lc	58.9	7	2e-08	25.41	958	1074
157	*Chaetura pelagica* (taxid:8897)	No significant similarity found						
158	*Caprimulgus carolinensis* (taxid:279965)	3 Sp, 3 Hs, 3 Cs, 3 Ac, 2 Bf, 2 Lc, 1 Cv, 1 Sk, 1 Xl, 1 Dr	89.0	12	1e-18	30.73	847	1057
159	*Eurypyga helias* (taxid:54383)	No significant similarity found						
160	*Mesitornis unicolor* (taxid:54374)	2 Cv, 2 Sp, 2 Bf, 2 Xl, 2 Hs, 2 Lc, 2 Dr, 2 Cs, 2 Hv, 2 Ac, 2 Ag, 1 Sk	88.2	10	3e-18	35.39	881	1049
161	*Podiceps cristatus* (taxid:345573)	No significant similarity found						
162	*Phoenicopterus ruber ruber* (taxid:9218)	No significant similarity found						
163	*Alligator mississippiensis* (taxid:8496)	55 Lc, 50 Xl, 46 Bf, 45 Cs, 45 Ac, 40 Sp, 39 Hs, 38 Sk, 28 At, 24 Cv, 24 Dr, 23 Ag, 19 Hv	221	15	4e-57	39.75	1085	1399
164	*Alligator sinensis* (taxid:38654)	22 Xl, 18 Lc, 17 Sk, 12 Ac, 11 Sp, 11 Bf, 10 Cs, 10 Hv, 8 Hs, 7 Dr, 5 Ag, 4 Cv, 4 At	206	13	2e-56	45.80	1170	1407
165	*Crocodylus porosus* (taxid:8502)	13 Cs, 12 Hs, 12 Dr, 11 Lc, 10 Bf, 10 Ac, 9 Sp, 9 Xl, 8 Cv, 8 Sk, 8 At, 7 Ag, 6 Hv	94.0	26	8e-21	23.04	658	1063
166	*Gavialis gangeticus* (taxid:94835)	No significant similarity found						

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
167	*Pogona vitticeps* (taxid:103695)	37 Ac, 35 Hs, 33 Lc, 31 Sp, 29 Dr, 26 Bf, 26 Xl, 23 Ag, 20 Cs, 19 Hv, 18 Cv, 15 Sk, 6 At	129	66	7e-31	27.44	729	1102
168	*Anolis carolinensis* (taxid:28377)	99 Ac, 87 Hs, 72 Lc, 37 Bf, 37 Sk, 35 Xl, 31 Hv, 28 Sp, 27 Dr, 25 Ag, 20 Cs, 15 At, 14 Cv	533	76	2e-157	29.64	337	1568
169	*Vipera berus berus* (taxid:31156)	No significant similarity found						
170	*Crotalus mitchellii pyrrhus* (taxid:384069)	No significant similarity found						
171	*Ophiophagus hannah* (taxid:8665)	2 Hs, 2 Ac	52.4	7	1e-10	34.21	1406	1481
172	*Python bivittatus* (taxid:176946)	5 Lc, 5 Ac, 3 Hs, 2 Sp, 2 Sk, 2 Dr, 2 Cs, 1 Cv, 1 Bf, 1 Xl, 1 Hv, 1 Ag, 1 At	142	12	1e-33	35.71	763	958
AMPHIBIA								
173	*Nanorana parkeri* (taxid:125878)	No significant similarity found						
174	*Xenopus tropicalis* (taxid:8364)	1122 Lc, 982 Hs, 884 Ac, 584 Sp, 580 Bf, 455 Sk, 452 Dr, 446 Cs, 437 Cv, 425 Ag, 411 Xl, 374 At, 308 Hv	1909	88	0.0	84.18	774	1879
NEOPTERYGII								
175	*Lepisosteus oculatus* (taxid:7918)	7 Lc, 5 Hs, 4 Sp, 4 Bf, 3 Sk, 3 Dr, 3 Hv, 3 Ac, 2 Cv, 2 Xl, 2 Cs, 2 Ag, 1 At	106	18	4e-23	27.68	732	1020
176	*Anguilla anguilla* (taxid:7936)	2 Xl, 2 Lc, 1 Sp, 1 Hv	49.7	8	2e-07	28.93	847	967
177	*Anguilla japonica* (taxid:7937)	1 Cv, 1 Sp, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Cs, 1 Hv, 1 Ac, 1 Ag, 1 At	85.9	20	1e-17	24.86	794	1124
178	*Danio rerio* (taxid:7955)	3052 Lc, 2857 Hs, 2498 Dr, 2025 Xl, 1933 Ac, 1877 Sk, 1577 Bf, 1565 Sp, 1472 Hv, 1403 Ag, 1347 Cs, 1177 Cv, 1066 At	2897	100	0.0	93.13	1	1520
179	*Astyanax mexicanus* (taxid:7994)	11 Lc, 7 Xl, 7 Hs, 7 Ac, 6 Bf, 6 Sk, 4 Dr, 3 Hv, 3 Ag, 2 Sp, 2 At, 1 Cv, 1 Cs	137	7	1e-32	53.54	927	1053

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
180	*Oryzias latipes* (taxid:8090)	3835 Lc, 3052 Hs, 2539 Ac, 1867 Sp, 1684 Bf, 1673 Sk, 1655 Xl, 1349 Cs, 1348 Dr, 1337 Hv, 1260 Cv, 1079 At, 980 Ag	832	71	0.0	36.77	778	2067
181	*Poecilia formosa* (taxid:48698)	18 Lc, 14 Hs, 9 Hv, 8 Cv, 8 Sk, 8 Cs, 7 Xl, 7 Ac, 4 Sp, 4 Bf, 4 Dr, 4 Ag, 2 At	240	31	7e-64	37.63	271	639
182	*Xiphophorus maculatus* (taxid:8083)	30 Lc, 24 Hs, 23 Xl, 23 Hv, 21 Sk, 19 Cv, 19 Cs, 15 Dr, 13 Sp, 10 Bf, 10 Ac, 8 At, 7 Ag	201	10	1e-51	47.91	1143	1356
183	*Fundulus heteroclitus* (taxid:8078)	60 Lc, 56 Hs, 44 Ac, 29 Ag, 25 Xl, 24 Bf, 23 Sk, 21 Sp, 21 Cs, 21 At, 13 Dr, 13 Hv, 5 Cv	427	86	3e-174	33.16	865	1625
184	*Takifugu flavidus* (taxid:433684)	No significant similarity found						
185	*Takifugu rubripes* (taxid:31033)	60 Lc, 51 Ac, 49 Hs, 43 Xl, 42 Cs, 35 Sp, 34 Cv, 33 Sk, 29 Bf, 23 Dr, 22 Hv, 22 Ag, 22 At	648	68	0.0	33.68	840	2069
186	*Tetraodon nigroviridis* (taxid:99883)	37 Hs, 33 Sp, 31 Xl, 31 Ac, 31 Ag, 29 At, 26 Bf, 22 Hv, 21 Sk, 21 Lc, 21 Cs, 18 Dr, 16 Cv	843	76	0.0	35.27	673	2076
187	*Cynoglossus semilaevis* (taxid:244447)	8 Hs, 8 Lc, 8 Ac, 6 Cv, 6 Cs, 5 Sp, 5 Sk, 5 Xl, 5 Hv, 4 Dr, 4 Ag, 4 At, 2 Bf	154	32	3e-39	23.95	760	1285
188	*Haplochromis burtoni* (taxid:8153)	54 Lc, 38 Hs, 30 Ac, 20 Sk, 18 Bf, 16 Dr, 15 Sp, 15 At, 14 Xl, 11 Hv, 8 Cv, 7 Ag, 6 Cs	656	72	0.0	34.61	417	1608
189	*Pundamilia nyererei* (taxid:303518)	23 Lc, 21 Ac, 17 Hs, 9 Sp, 8 Bf, 6 At, 5 Sk, 5 Xl, 5 Cs, 4 Dr, 4 Ag, 3 Cv, 3 Hv	255	13	2e-68	45.26	1127	1411
190	*Maylandia zebra* (taxid:106582)	28 Lc, 23 Hs, 14 Ac, 13 Bf, 12 Ag, 11 Xl, 10 Sk, 9 Cs, 8 Sp, 6 Cv, 6 Hv, 6 At, 4 Dr	355	35	8e-113	47.76	882	1258
191	*Neolamprologus brichardi* (taxid:32507)	5 Hs, 4 Lc, 3 Bf, 1 Sp, 1 Sk, 1 Ac	69.3	20	4e-11	21.02	740	1057

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
192	*Oreochromis niloticus* (taxid:8128)	91 Lc, 63 Hs, 55 Ac, 43 Sp, 41 Bf, 34 Sk, 33 Ag, 29 Cv, 29 Xl, 29 Dr, 29 Cs, 29 Hv, 22 At	429	48	4e-121	33.33	918	1628
193	*Sebastes nigrocinctus* (taxid:72089)	No significant similarity found						
194	*Sebastes rubrivinctus* (taxid:72099)	No significant similarity found						
195	*Gasterosteus aculeatus* (taxid:69293)	58 Lc, 57 Ac, 50 Hs, 44 Bf, 43 Sk, 42 Xl, 35 At, 32 Sp, 32 Dr, 29 Ag, 27 Cv, 26 Cs, 13 Hv	449	54	0.0	41.39	931	1503
196	*Gadus morhua* (taxid:8049)	3 Lc, 2 Bf, 1 Cv	61.6	9	5e-10	21.33	1479	1628
CHONDRICHTHYES								
197	*Callorhinchus milii* (taxid:7868)	96 Lc, 46 Ac, 41 Sp, 36 Ag, 35 Hs, 34 Cv, 29 Bf, 17 Cs, 8 Sk, 7 At, 6 Xl, 6 Drp	97.1	19	2e-19	26.16	659	954
198	*Carcharhinus brachyurus* (taxid:671158)	No significant similarity found						
ECDYSOZOA								
199	*Ephemera danica* (taxid:1049336)	No significant similarity found						
200	*Ladona fulva* (taxid:123851)	No significant similarity found						
201	*Pediculus humanus corporis* (taxid:121224)	15 Lc, 14 Ac, 13 Hs, 11 At, 10 Cv, 10 Sp, 10 Bf, 10 Sk, 10 Xl, 10 Cs, 10 Hv, 10 Ag, 9 Dr	85.9	23	1e-16	24.67	966	1404
202	*Frankliniella occidentalis* (taxid:133901)	No significant similarity found						
203	*Diaphorina citri* (taxid:121845)	38 Lc, 37 Cs, 37 Ac, 34 Sp, 33 Hs, 30 Xl, 29 Sk, 28 Cv, 27 Ag, 24 At, 22 Hv, 21 Bf, 18 Dr	132	30	2e-31	28.69	307	665
204	*Pachypsylla venusta* (taxid:38123)	No significant similarity found						
205	*Acyrthosiphon pisum* (taxid:7029)	79 Ac, 68 Cs, 56 Hs, 50 Lc, 49 Sp, 44 Sk, 42 Xl, 40 At, 32 Bf, 30 Ag, 27 Cv, 21 Hv, 17 Dr	120	12	1e-26	32.34	1223	1486

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
206	*Nilaparvata lugens* (taxid:108931)	2 Ac	50.4	5	4e-08	32.58	775	860
207	*Oncopeltus fasciatus* (taxid:7536)	No significant similarity found						
208	*Rhodnius prolixus* (taxid:13249)	1 Sp, 1 Sk, 1 Xl, 1 Hs, 1 Dr, 1 Cs, 1 Ac, 1 Ag	64.3	12	4e-12	22.58	864	1039
209	*Cimex lectularius* (taxid:79782)	22 Hs, 14 Ac, 12 Lc, 12 Cs, 7 Sp, 6 Bf, 6 Xl, 5 Dr, 5 Hv, 5 Ag, 5 At, 4 Cv, 4 Sk	100	13	5e-21	29.22	785	1001
210	*Onthophagus taurus* (taxid:166361)	No significant similarity found						
211	*Agrilus planipennis* (taxid:224129)	No significant similarity found						
212	*Tribolium castaneum* (taxid:7070)	23 Sp, 22 Xl, 22 Lc, 20 Bf, 15 Hs, 14 Ac, 11 Hv, 11 Ag, 10 At, 8 Cs, 7 Sk, 6 Cv, 2 Dr	132	28	2e-30	24.67	737	1169
213	*Anoplophora glabripennis* (taxid:217634)	No significant similarity found						
214	*Leptinotarsa decemlineata* (taxid:7539)	No significant similarity found						
215	*Dendroctonus ponderosae* (taxid:77166)	No significant similarity found						
216	*Mengenilla moldrzyki* (taxid:1155016)	No significant similarity found						
217	*Aedes aegypti* (taxid:7159)	104 Lc, 104 Ac, 103 Hs, 102 Bf, 101 Sp, 93 Cs, 87 Ag, 85 At, 79 Hv, 77 Cv, 75 Xl, 68 Dr, 59 Sk	164	61	6e-49	28.07	828	1345
218	*Culex quinquefasciatus* (taxid:7176)	42 Ag, 4 Cv, 4 Bf, 4 Lc, 4 Dr, 4 Cs, 4 Hv, 3 Sp, 3 Sk, 3 Xl, 3 Hs, 3 Ac, 3 At	176	15	6e-49	40.09	1008	1224
219	*Anopheles albimanus* (taxid:7167)	6 Ac, 3 Sk, 3 Hs	53.5	5	1e-08	28.92	733	815
220	*Anopheles arabiensis* (taxid:7173)	3 Lc, 2 Hs, 2 Ac, 1 Cv, 1 Sp, 1 Sk, 1 Xl, 1 Dr, 1 Cs, 1 Hv	72.8	10	5e-15	29.89	846	1020
221	*Anopheles atroparvus* (taxid:41427)	No significant similarity found						
222	*Anopheles christyi* (taxid:43041)	No significant similarity found						

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
223	*Anopheles culicifacies* (taxid:139723)	No significant similarity found						
224	*Anopheles darlingi* (taxid:43151)	No significant similarity found						
225	*Anopheles dirus* (taxid:7168)	No significant similarity found						
226	*Anopheles epiroticus* (taxid:199890)	No significant similarity found						
227	*Anopheles farauti* (taxid:69004)	No significant similarity found						
228	*Anopheles funestus* (taxid:62324)	No significant similarity found						
229	*Anopheles gambiae* (taxid:7165)	24 Hs, 24 Lc, 22 Ac, 12 Sp, 12 At, 10 Ag, 8 Cv, 8 Bf, 8 Sk, 8 Cs, 6 Xl, 5 Dr, 5 Hv	818	90	0.0	41.12	325	1380
230	*Anopheles maculatus* (taxid:74869)	No significant similarity found						
231	*Anopheles melas* (taxid:34690)	No significant similarity found						
232	*Anopheles merus* (taxid:30066)	1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 Cs, 1 Hv, 1 Ac, 1 At	78.6	18	4e-17	22.74	662	955
233	*Anopheles minimus* (taxid:112268)	No significant similarity found						
234	*Anopheles quadriannulatus* (taxid:34691)	No significant similarity found						
235	*Anopheles sinensis* (taxid:74873)	5 Sp, 5 Lc, 5 Cs, 3 Bf, 3 Xl, 2 Cv, 2 At	60.5	7	8e-12	28.24	907	1037
236	*Anopheles stephensi* (taxid:30069)	12 Sp, 11 Lc, 10 Hs, 10 Ac, 8 Ag, 6 Xl, 5 Bf, 2 Cv, 2 Cs, 1 Sk, 1 At	67.0	11	1e-13	31.34	447	578
237	*Mayetiola destructor* (taxid:39758)	1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Hs, 1 Dr, 1 Cs, 1 Ac, 1 Ag, 1 At	90.5	24	2e-19	23.42	681	1052
238	*Lutzomyia longipalpis* (taxid:7200)	1 Ac, 1 Sp	50.1	5	3e-08	40.70	810	892
239	*Phlebotomus papatasi* (taxid:29031)	11 Ac, 10 Lc, 10 Ag, 9 Bf, 9 Hs, 8 Sp, 8 Xl, 7 Sk, 7 Cs, 6 Dr, 6 Hv, 6 At, 2 Cv	136	23	5e-33	30.48	839	1197
240	*Ceratitis capitata* (taxid:7213)	3 Sp, 3 Sk, 3 Hs, 3 Ac, 3 Ag, 2 Bf, 2 Lc, 1 Xl, 1 Dr, 1 Cs, 1 Hv, 1 At	81.6	14	6e-15	30.90	724	951

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
241	*Drosophila albomicans* (taxid:7291)	No significant similarity found						
242	*Drosophila ananassae* (taxid:7217)	27 Cs, 25 Hs, 25 Lc, 25 Ac, 24 Cv, 24 Sp, 24 Sk, 24 Xl, 24 Dr, 24 Hv, 24 At, 23 Bf, 23 Ag	116	26	5e-26	26.52	783	1181
243	*Drosophila biarmipes* (taxid:125945)	No significant similarity found						
244	*Drosophila bipectinata* (taxid:42026)	No significant similarity found						
245	*Drosophila elegans* (taxid:30023)	No significant similarity found						
246	*Drosophila erecta* (taxid:7220)	No significant similarity found						
247	*Drosophila eugracilis* (taxid:29029)	No significant similarity found						
248	*Drosophila ficusphila* (taxid:30025)	No significant similarity found						
249	*Drosophila grimshawi* (taxid:7222)	No significant similarity found						
250	*Drosophila kikkawai* (taxid:30033)	No significant similarity found						
251	*Drosophila melanogaster* (taxid:7227)	1886 Ac, 1537 Hs, 1511 Lc, 518 Sp, 487 Cs, 353 Ag, 279 Bf, 208 Cv, 187 Dr, 185 Xl, 173 Sk, 148 At, 107 Hv	151	33	2e-63	45.27	468	708
252	*Drosophila miranda* (taxid:7229)	26 Ac, 21 Lc, 19 Hs, 19 Cs, 15 Sp, 15 Sk, 15 At, 14 Bf, 12 Cv, 12 Dr, 11 Xl, 10 Ag, 9 Hv	104	28	8e-23	25.60	611	1056
253	*Drosophila mojavensis* (taxid:7230)	No significant similarity found						
254	*Drosophila persimilis* (taxid:7234)	2 Ac, 1 Hs	53.1	11	1e-06	22.83	790	973
255	*Drosophila pseudoobscura pseudoobscura* (taxid:46245)	No significant similarity found						

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
256	*Drosophila rhopaloa* (taxid:1041015)	No significant similarity found						
257	*Drosophila sechellia* (taxid:7238)	9 Ac, 8 Hs, 5 Lc, 3 Cs, 1 Cv, 1 Bf	107	21	3e-23	29.94	730	1056
258	*Drosophila simulans* (taxid:7240)	8 Hs, 8 Ac, 7 Lc, 4 Cs, 3 Sk, 3 Ag, 2 Cv, 2 Bf, 2 Dr, 1 Sp, 1 At	105	21	2e-22	30.52	730	1056
259	*Drosophila suzukii* (taxid:28584)	No significant similarity found						
260	*Drosophila takahashii* (taxid:29030)	No significant similarity found						
261	*Drosophila virilis* (taxid:7244)	10 Ac, 9 Hs, 7 Lc, 3 Cv, 3 Sp, 3 Cs, 3 Ag, 2 Bf, 2 Sk, 2 Dr, 2 At	100	22	1e-20	26.93	743	1112
262	*Drosophila willistoni* (taxid:7260)	2 Sp, 2 Sk, 2 Xl, 2 Lc, 2 Dr, 2 Cs, 2 Hv, 2 Ac, 2 Ag, 1 Cv, 1 Bf, 1 Hs, 1 At	118	41	8e-27	23.66	447	1114
263	*Drosophila yakuba* (taxid:7245)	8 Hs, 8 Ac, 7 Lc, 6 Cs, 4 Ag, 2 Bf, 1 At	105	18	4e-22	29.80	730	1017
264	*Musca domestica* (taxid:7370)	19 Hs, 10 Lc, 6 Ac, 6 Ag, 4 Xl, 2 Bf, 2 Cs, 2 Hv	78.2	20	4e-14	27.41	748	1063
265	*Glossina austeni* (taxid:7395)	No significant similarity found						
266	*Glossina brevipalpis* (taxid:37001)	No significant similarity found						
267	*Glossina fuscipes fuscipes* (taxid:201502)	No significant similarity found						
268	*Glossina morsitans morsitans* (taxid:37546)	No significant similarity found						
269	*Glossina pallidipes* (taxid:7398)	No significant similarity found						
270	*Limnephilus lunatus* (taxid:1218281)	No significant similarity found						
271	*Papilio glaucus* (taxid:45779)	No significant similarity found						
272	*Papilio polytes* (taxid:76194)	9 Ac, 7 Lc, 7 At, 6 Xl, 6 Hs, 6 Cs, 5 Sp, 5 Sk, 5 Dr, 5 Ag, 3 Bf, 3 Hv, 2 Cv	144	35	1e-34	28.23	252	665

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
273	*Papilio xuthus* (taxid:66420)	36 Hs, 33 Ac, 29 Lc, 25 At, 24 Sp, 19 Sk, 18 Xl, 18 Hv, 17 Cv, 17 Cs, 16 Ag, 9 Bf, 7 Dr	110	18	3e-24	31.60	781	1065
274	*Heliconius melpomene melpomene* (taxid:171917)	No significant similarity found						
275	*Melitaea cinxia* (taxid:113334)	No significant similarity found						
276	*Danaus plexippus* (taxid:13037)	No significant similarity found						
277	*Bombyx mori* (taxid:7091)	322 Ac, 311 Lc, 310 Hs, 262 Cs, 232 Sp, 221 Ag, 214 Sk, 204 Cv, 185 At, 162 Xl, 157 Bf, 135 Dr, 110 Hv	184	40	2e-46	27.58	448	1093
278	*Manduca sexta* (taxid:7130)	3 Hs, 3 Lc, 3 Ac, 2 Sp, 2 Sk, 2 Cs, 2 At, 1 Cv, 1 Bf, 1 Xl, 1 Dr, 1 Hv, 1 Ag	92.0	37	3e-23	28.91	438	693
279	*Plutella xylostella* (taxid:51655)	74 Ac, 72 Hs, 69 Lc, 57 Sp, 55 Bf, 55 At, 54 Ag, 44 Cs, 41 Sk, 40 Cv, 40 Xl, 33 Dr, 30 Hv	138	45	2e-32	22.06	352	1077
280	*Athalia rosae* (taxid:37344)	2 Ac, 1 Cv, 1 Sp, 1 Bf, 1 Hs, 1 Lc, 1 Hv, 1 Ag, 1 At	100	18	8e-21	27.22	845	1151
281	*Cephus cinctus* (taxid:211228)	18 Lc, 17 Ac, 14 Hs, 14 At, 13 Ag, 12 Sk, 4 Cs, 3 Xl, 2 Sp, 2 Bf, 1 Dr	88.2	16	5e-17	29.28	729	989
282	*Orussus abietinus* (taxid:222816)	4 Hs, 4 Ac, 3 Lc, 2 Sp, 2 Sk, 2 Xl, 2 Cs, 2 Hv, 2 Ag, 2 At, 1 Cv, 1 Bf, 1 Dr	95.5	19	2e-19	25.78	728	1029
283	*Ceratosolen solmsi marchali* (taxid:326594)	1 Bf, 1 Hs	74.7	14	3e-18	26.54	341	571
284	*Nasonia giraulti* (taxid:7426)	No significant similarity found						
285	*Nasonia longicornis* (taxid:7427)	No significant similarity found						
286	*Nasonia vitripennis* (taxid:7425)	31 Ac, 28 Bf, 24 Lc, 22 Hs, 20 Cs, 19 Sp, 19 At, 15 Hv, 14 Cv, 14 Dr, 13 Sk, 12 Xl, 11 Ag	114	29	5e-25	24.59	743	1210

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
287	*Copidosoma floridanum* (taxid:29053)	15 Lc, 14 Cs, 13 Sp, 13 Ac, 11 Ag, 9 Dr, 8 Bf, 8 Xl, 8 Hs, 7 Hv, 7 At, 6 Sk, 4 Cv	106	21	2e-23	26.57	719	1050
288	*Trichogramma pretiosum* (taxid:7493)	16 Hs, 13 Ac, 11 Sp, 11 Lc, 10 Ag, 7 Sk, 5 Xl, 4 Cs, 3 Bf, 3 Dr, 3 At, 1 Cv, 1 Hv	120	17	7e-27	27.05	919	1208
289	*Microplitis demolitor* (taxid:69319)	14 Sp, 13 Cs, 13 Ac, 11 Bf, 11 Hs, 9 Sk, 8 Lc, 8 Ag, 6 Dr, 5 Cv, 5 Xl, 3 Hv, 3 At	114	20	3e-25	28.91	786	1112
290	*Megachile rotundata* (taxid:143995)	10 Ac, 9 Hs, 6 Lc, 5 Cs, 4 Sp, 4 Xl, 3 At, 2 Cv, 1 Bf, 1 Sk, 1 Hv	97.1	17	1e-19	26.82	787	1074
291	*Apis dorsata* (taxid:7462)	No significant similarity found						
292	*Apis florea* (taxid:7463)	No significant similarity found						
293	*Apis mellifera* (taxid:7460)	1 Ac	60.8	7	2e-08	33.07	736	856
294	*Bombus impatiens* (taxid:132113)	4 Ac, 1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 Cs, 1 Hv, 1 Ag, 1 At	103	14	1e-22	29.39	733	959
295	*Bombus terrestris* (taxid:30195)	10 Hs, 9 Lc, 4 Xl, 4 Ac, 3 Sp, 3 Ag, 2 Dr, 2 Cs, 1 Cv	79.7	10	2e-17	31.87	732	889
296	*Linepithema humile* (taxid:83485)	11 Hs, 7 Ac, 7 At, 6 Lc, 3 Cs, 2 Sp, 2 Bf, 2 Sk, 2 Xl, 2 Ag, 1 Cv, 1 Dr, 1 Hv	89.0	18	2e-17	27.36	777	1063
297	*Camponotus floridanus* (taxid:104421)	7 Ac, 6 Lc, 5 Hs, 5 Cs, 4 Bf, 3 Sp, 3 Xl, 3 At, 1 Cv, 1 Sk, 1 Dr, 1 Hv, 1 Ag	87.0	20	1e-16	22.75	733	1052
298	*Acromyrmex echinatior* (taxid:103372)	6 Hs, 5 Lc, 4 Ac, 2 At, 1 Hv	73.9	17	7e-13	25.68	847	1129
299	*Atta cephalotes* (taxid:12957)	1 Hs	57.4	14	3e-08	23.05	341	577
300	*Solenopsis invicta* (taxid:13686)	32 Ac, 24 Hs, 19 Xl, 18 Lc, 16 Sp, 16 Dr, 11 Ag, 11 At, 10 Cs, 9 Hv, 8 Bf, 5 Sk, 1 Cv	108	17	2e-23	30.85	822	1101
301	*Pogonomyrmex barbatus* (taxid:144034)	6 Ac, 5 Sp, 5 Xl, 5 Cs, 3 Bf, 3 Hs, 3 Lc, 3 Ag, 2 Cv, 1 Dr	85.1	20	4e-16	23.75	733	1057
302	*Harpegnathos saltator* (taxid:610380)	18 Hs, 13 Lc, 10 Ac, 7 At, 5 Cv, 3 Ag, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Dr	99.0	20	2e-20	23.53	725	1048

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
303	*Cerapachys biroi* (taxid:443821)	13 Hs, 8 Ac, 6 Lc, 5 Bf, 5 Cs, 4 Sk, 4 Ag, 4 At, 3 Cv, 3 Sp, 3 Xl, 2 Dr, 2 Hv	122	23	2e-27	27.42	653	1020
304	*Blattella germanica* (taxid:6973)	5 Lc, 5 Ac, 4 Cs, 3 Bf, 3 Hs, 3 Dr, 3 Hv, 3 Ag, 3 At, 2 Cv, 2 Sp, 2 Sk, 2 Xl	380	84	2e-110	29.01	335	1337
305	*Zootermopsis nevadensis* (taxid:136037)	No significant similarity found						
306	*Daphnia pulex* (taxid:6669)	5 Sp, 5 Hs, 5 Lc, 5 Cs, 5 Ag, 4 Cv, 4 Sk, 4 Xl, 4 Dr, 4 Ac, 3 Bf, 3 At, 2 Hv	239	56	8e-66	25.61	345	1197
307	*Eurytemora affinis* (taxid:88015)	No significant similarity found						
308	*Hyalella azteca* (taxid:294128)	No significant similarity found						
309	*Strigamia maritima* (taxid:126957)	No significant similarity found						
310	*Stegodyphus mimosarum* (taxid:407821)	No significant similarity found						
311	*Latrodectus hesperus* (taxid:256737)	5 Lc, 4 Hs, 3 Ac, 2 Xl, 2 Ag, 1 Dr, 1 Cs, 1 Hv, 1 At	84.3	21	9e-18	25.62	790	1140
312	*Parasteatoda tepidariorum* (taxid:114398)	No significant similarity found						
313	*Tetranychus urticae* (taxid:32264)	No significant similarity found						
314	*Dermatophagoides farinae* (taxid:6954)	No significant similarity found						
315	*Sarcoptes scabiei type canis* (taxid:52283)	No significant similarity found						
316	*Achipteria coleoptrata* (taxid:229769)	No significant similarity found						
317	*Hypochthonius rufulus* (taxid:66581)	No significant similarity found						
318	*Platynothrus peltifer* (taxid:128015)	No significant similarity found						
319	*Steganacarus magnus* (taxid:52000)	No significant similarity found						

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
320	*Ixodes ricinus* (taxid:34613)	No significant similarity found						
321	*Ixodes scapularis* (taxid:6945)	99 Sp, 93 Lc, 86 Hs, 85 Ac, 76 Ag, 72 Cs, 69 Bf, 55 Sk, 54 Dr, 53 Cv, 47 Xl, 47 At, 33 Hv	318	69	6e-88	26.27	339	1395
322	*Rhipicephalus microplus* (taxid:6941)	30 Ac, 29 Sp, 27 Sk, 27 Hs, 25 Cs, 24 Bf, 23 Lc, 21 Xl, 21 Ag, 18 Dr, 14 Hv, 13 At, 12 Cv	150	46	3e-37	24.42	367	1112
323	*Metaseiulus occidentalis* (taxid:34638)	16 Ac, 13 Hs, 10 Cs, 10 At, 9 Lc, 8 Bf, 7 Cv, 7 Sp, 7 Sk, 6 Xl, 5 Hv, 3 Ag, 2 Dr	123	25	1e-28	26.56	660	1057
324	*Varroa destructor* (taxid:109461)	No significant similarity found						
325	*Centruroides exilicauda* (taxid:6879)	1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Ac, 1 At	60.5	8	4e-13	33.33	830	962
326	*Mesobuthus martensii* (taxid:34649)	No significant similarity found						
327	*Limulus polyphemus* (taxid:6850)	8 Ac, 6 Lc, 6 Dr, 6 Hv, 5 Sp, 5 Xl, 5 Hs, 5 Ag, 3 Cv, 3 Sk, 3 Cs, 3 At, 2 Bf	365	51	2e-102	30.19	438	1280
328	*Trichinella spiralis* (taxid:6334)	11 Xl, 9 Sp, 9 Ac, 8 Sk, 8 Cs, 7 Bf, 7 Lc, 7 Hv, 6 Ag, 5 Dr, 3 Hs, 1 Cv, 1 At	98.2	20	1e-20	25.32	1207	1637
329	*Ascaris suum* (taxid:6253)	1 Cv, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Cs, 1 Ac, 1 At	58.2	9	9e-10	26.04	1042	1210
330	*Elaeophora elaphi* (taxid:1147741)	No significant similarity found						
331	*Onchocerca volvulus* (taxid:6282)	No significant similarity found						
332	*Steinernema monticolum* (taxid:90984)	No significant similarity found						
333	*Panagrellus redivivus* (taxid:6233)	No significant similarity found						
334	*Haemonchus contortus* (taxid:6289)	No significant similarity found						

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
335	*Necator americanus* (taxid:51031)	51 Lc, 46 Xl, 44 Sp, 44 Hs, 43 Bf, 43 Sk, 42 Dr, 41 Ac, 39 Hv, 37 Ag, 35 Cv, 33 Cs, 31 At	110	20	3e-25	25.15	816	1153
336	*Heterorhabditis bacteriophora* (taxid:37862)	No significant similarity found						
337	*Caenorhabditis angaria* (taxid:860376)	No significant similarity found						
338	*Caenorhabditis brenneri* (taxid:135651)	No significant similarity found						
339	*Caenorhabditis briggsae* (taxid:6238)	14 Sp, 11 Lc, 9 Bf, 9 Sk, 8 Ac, 7 Cv, 7 Xl, 7 Cs, 6 Hs, 6 At, 5 Ag, 2 Dr, 2 Hv	79.7	17	1e-14	25.44	749	1021
340	*Caenorhabditis elegans* (taxid:6239)	196 Lc, 192 Hs, 191 Ac, 184 Sp, 152 Cs, 147 Bf, 129 Ag, 122 Sk, 122 At, 102 Xl, 101 Cv, 98 Dr, 94 Hv	1258	39	0.0	95.89	1	635
341	*Caenorhabditis japonica* (taxid:281687)	7 Bf, 6 Cv, 6 Lc, 6 Cs, 6 At, 5 Sp, 5 Sk, 5 Hs, 5 Ac, 5 Ag, 4 Hv, 3 Dr, 2 Xl	122	25	7e-29	29.47	677	1049
342	*Caenorhabditis sp. 11 MAF-2010* (taxid:1561998)	No significant similarity found						
343	*Priapulus caudatus* (taxid:37621)	8 Sp, 8 Ac, 7 Bf, 7 Hs, 6 Sk, 6 Lc, 6 Cs, 5 Dr, 5 Ag, 4 Cv, 4 Xl, 4 At, 3 Hv	114	47	3e-25	23.53	349	1115
ROTIFERA								
344	*Adineta vaga* (taxid:104782)	18 Sp, 18 Xl, 18 Hs, 18 Lc, 18 Cs, 18 Hv, 18 Ac, 18 Ag, 18 At, 13 Dr, 5 Cv, 5 Sk, 4 Bf	110	28	4e-25	24.52	809	1275
PLATYHELMINTHES								
345	*Schistosoma curassoni* (taxid:6186)	600 Sp, 592 Ac, 585 At, 576 Hs, 574 Hv, 573 Cs, 571 Lc, 570 Dr, 560 Sk, 548 Bf, 540 Xl, 536 Cv, 530 Ag	120	17	4e-26	29.29	750	1025
346	*Schistosoma haematobium* (taxid:6185)	4 Sp, 4 Lc, 3 Bf, 3 Sk, 3 Xl, 3 Hs, 2 Hv, 2 Ac, 2 Ag, 1 Cs	65.1	6	5e-12	39.19	512	585

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
347	*Schistosoma japonicum* (taxid:6182)	54 Sp, 52 Hs, 50 Ac, 48 Hv, 47 Cs, 46 Bf, 43 Xl, 43 Dr, 42 Lc, 41 Sk, 37 At, 36 Ag, 31 Cv	149	45	5e-36	22.61	344	1080
348	*Schistosoma mansoni* (taxid:6183)	5010 Sp, 4703 Hs, 4391 Ac, 4193 Lc, 3951 Hv, 3892 Bf, 3801 Sk, 3770 Cs, 3571 Dr, 3554 Xl, 3313 At, 3170 Cv, 3140 Ag	143	52	5e-33	23.60	348	1011
349	*Schistosoma margrebowiei* (taxid:48269)	1017 Sp, 901 Hs, 872 Lc, 853 At, 850 Hv, 843 Ac, 825 Cs, 821 Bf, 784 Sk, 754 Dr, 749 Cv, 747 Xl, 705 Ag	142	46	1e-32	22.46	333	1080
350	*Schistosoma mattheei* (taxid:31246)	602 Sp, 600 Ac, 581 Dr, 580 Sk, 578 Hs, 574 Cs, 574 Hv, 572 Lc, 560 Bf, 547 Xl, 520 Ag, 482 At, 366 Cv	376	49	3e-109	57.84	1053	1339
351	*Schistosoma rodhaini* (taxid:6188)	627 Sp, 608 Ac, 602 Hs, 602 Lc, 592 At, 591 Cs, 587 Hv, 586 Sk, 573 Dr, 559 Xl, 535 Bf, 517 Cv, 513 Ag	723	45	0.0	63.60	509	1057
352	*Clonorchis sinensis* (taxid:79923)	No significant similarity found						
353	*Echinococcus granulosus* (taxid:6210)	No significant similarity found						
354	*Echinococcus multilocularis* (taxid:6211)	No significant similarity found						
355	*Hymenolepis microstoma* (taxid:85433)	No significant similarity found						
ANNELIDA								
356	*Capitella teleta* (taxid:283909)	1 Dr	43.1	1	5e-06	41.46	238	269
357	*Helobdella robusta* (taxid:6412)	23 Hs, 17 Lc, 17 Ac, 15 Cs, 14 Cv, 14 Bf, 12 At, 10 Sp, 9 Xl, 9 Ag, 7 Sk, 6 Dr, 6 Hv	95.9	26	2e-19	21.36	611	1020
MOLLUSCA								
358	*Crassostrea gigas* (taxid:29159)	69 Bf, 64 Sp, 50 Hs, 46 Cs, 45 Dr, 44 Hv, 44 Ac, 40 Sk, 38 Lc, 36 Cv, 35 Xl, 35 Ag, 33 At	481	73	4e-138	29.58	405	1610

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
359	*Lottia gigantea* (taxid:225164)	58 Dr, 16 Bf, 15 Hs, 15 Hv, 11 Xl, 11 Lc, 10 Sp, 10 Sk, 10 Ac, 7 Cs, 6 Cv, 2 Ag, 1 At	169	12	1e-46	39.25	74	286
360	*Aplysia californica* (taxid:6500)	46 Ac, 44 Hs, 41 Sp, 38 Lc, 36 Cs, 35 Bf, 30 Sk, 30 Xl, 29 Dr, 29 Ag, 22 Hv, 14 Cv, 13 At	302	29	9e-85	37.76	405	885
361	*Biomphalaria glabrata* (taxid:6526)	47 Lc, 47 Ac, 43 Hs, 35 Sp, 34 Cs, 29 Sk, 28 Bf, 26 Hv, 24 Xl, 24 Ag, 21 At, 20 Cv, 19 Dr	338	62	1e-93	28.75	344	1289
CNIDARIA								
362	*Nematostella vectensis* (taxid:45351)	11 Dr, 9 Hs, 7 Ac, 6 Sp, 5 Sk, 5 Xl, 4 Bf, 4 Lc, 4 Hv, 3 Cv, 3 Cs, 1 Ag, 1 At	164	21	1e-43	30.77	346	674
363	*Hydra vulgaris* (taxid:6087)	100 Sp, 95 Bf, 91 Lc, 90 Ac, 84 Cs, 80 Hs, 60 Cv, 56 Sk, 55 Xl, 37 Hv, 32 Dr, 25 Ag, 14 At	182	11	8e-48	81.82	642	760
TENTACULATA								
364	*Mnemiopsis leidyi* (taxid:27923)	2 Sp, 2 Bf, 2 Sk, 2 Hs, 2 Lc, 2 Cs, 2 Ac, 2 Ag, 1 Cv, 1 Xl, 1 Dr, 1 Hv	96.3	23	2e-21	24.67	725	1098
PLACOZOA								
365	*Trichoplax adhaerens* (taxid:10228)	No significant similarity found						
PORIFERA								
366	*Amphimedon queenslandica* (taxid:400682)	6 Bf, 6 Hs, 5 Cv, 5 Sp, 5 Lc, 5 Dr, 5 Cs, 5 Ac, 2 Hv, 2 At, 1 Ag	95.9	19	5e-20	24.84	723	1028
VIRIDIPLANTAE								
367	*Micromonas pusilla CCMP1545* (taxid:564608)	1 Bf	48.9	2	1e-05	50.00	201	242
368	*Micromonas sp. RCC299* (taxid:296587)	No significant similarity found						
369	*Ostreococcus lucimarinus CCE9901* (taxid:436017)	1 Dr	47.4	1	8e-06	52.94	238	269
370	*Ostreococcus tauri* (taxid:70448)	1 Dr	47.8	2	8e-06	48.65	238	272

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
371	*Chlamydomonas reinhardtii* (taxid:3055)	7 Sp, 7 Xl, 7 Lc, 7 Hv, 6 Cv, 6 Sk, 6 Cs, 5 Bf, 4 Hs, 4 Ac, 4 Ag, 4 At, 3 Dr	164	34	4e-40	28.16	698	1219
372	*Volvox carteri f. nagariensis* (taxid:3068)	51 Sk, 50 Hs, 50 Hv, 47 Xl, 47 Cs, 46 Lc, 45 Sp, 45 Ac, 41 Bf, 41 At, 40 Dr, 40 Ag, 32 Cv	258	57	5e-69	25.29	344	1208
373	*Chlorella variabilis* (taxid:554065)	6 Ac, 4 Sp, 4 Bf, 4 Hs, 4 Lc, 4 Dr, 4 Hv, 4 Ag, 3 Cs, 2 Cv, 2 Sk, 2 Xl, 2 At	91.7	20	2e-18	25.29	787	1109
374	*Auxenochlorella protothecoides* (taxid:3075)	1 Dr	48.9	7	4e-06	27.97	546	668
375	*Helicosporidium sp. ATCC 50920* (taxid:1291522)	No nr/htgs data for this txid id						
376	*Coccomyxa subellipsoidea C-169* (taxid:574566)	9 Cv, 4 Lc, 3 Sp, 3 Ac, 2 Cs, 2 Hv, 2 Ag, 1 Sk, 1 Xl, 1 Hs, 1 Dr, 1 At	150	4	1e-39	70.83	628	723
377	*Klebsormidium flaccidum* (taxid:3175)	No significant similarity found						
378	*Physcomitrella patens* (taxid:3218)	10 Xl, 7 Bf, 7 Sk, 5 Sp, 4 Cv, 4 Lc, 4 At, 3 Hs, 3 Dr, 3 Cs, 3 Hv, 2 Ag	88.2	11	1e-22	42.73	839	948
379	*Selaginella moellendorffii* (taxid:88036)	5 Xl, 4 Sk, 1 Hs, 1 Ag, 1 At	61.6	9	4e-09	26.58	829	1033
380	*Pinus taeda* (taxid:3352)	293 At, 94 Lc, 91 Sp, 89 Sk, 89 Hv, 89 Ac, 87 Hs, 82 Bf, 79 Xl, 77 Dr, 76 Ag, 75 Cv, 70 Cs	595	80	5e-174	30.58	531	1868
381	*Amborella trichopoda* (taxid:13333)	179 At, 47 Lc, 43 Sp, 40 Xl, 40 Hv, 38 Sk, 35 Cs, 34 Hs, 33 Dr, 31 Cv, 27 Bf, 27 Ac, 24 Ag	317	85	6e-96	38.03	998	1486
382	*Spirodela polyrhiza* (taxid:29656)	8 At, 3 Dr, 3 Hv, 2 Cv, 2 Sp, 2 Bf, 2 Sk, 2 Xl, 2 Hs, 2 Lc, 2 Cs, 2 Ac, 2 Ag	422	82	1e-120	27.72	660	1870
383	*Phoenix dactylifera* (taxid:42345)	372 At, 78 Hs, 68 Lc, 54 Sp, 52 Cv, 52 Ac, 49 Dr, 48 Hv, 47 Xl, 43 Sk, 43 Cs, 42 Bf, 29 Ag	520	61	5e-157	30.07	691	1869

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
384	*Elaeis oleifera* (taxid:80265)	No significant similarity found						
385	*Ensete ventricosum* (taxid:4639)	No significant similarity found						
386	*Musa acuminata subsp. malaccensis* (taxid:214687)	10 At	72.4	6	4e-12	31.30	104	218
387	*Sorghum bicolor* (taxid:4558)	486 At, 151 Hv, 146 Sk, 144 Ac, 142 Hs, 141 Sp, 139 Xl, 137 Lc, 134 Bf, 127 Cs, 111 Dr, 105 Cv, 98 Ag	447	58	1e-126	29.14	526	1640
388	*Zea mays* (taxid:4577)	1403 At, 1030 Hs, 999 Hv, 976 Sp, 968 Lc, 898 Bf, 886 Ac, 878 Sk, 867 Xl, 848 Cs, 821 Dr, 785 Ag, 766 Cv	1212	60	0.0	89.68	694	1352
389	*Setaria italica* (taxid:4555)	40 At, 15 Hs, 15 Hv, 15 Ac, 14 Lc, 12 Bf, 10 Sp, 10 Xl, 9 Sk, 9 Dr, 9 Cs, 9 Ag, 8 Cv	358	62	2e-99	25.67	531	1721
390	*Brachypodium distachyon* (taxid:15368)	214 At, 96 Hs, 95 Lc, 89 Bf, 75 Hv, 69 Sp, 40 Xl, 40 Ac, 39 Sk, 38 Cs, 36 Cv, 34 Dr, 33 Ag	442	55	3e-129	28.81	750	1802
391	*Leersia perrieri* (taxid:77586)	5 Cv, 5 Bf, 5 At, 4 Sp, 4 Sk, 4 Xl, 4 Hs, 4 Lc, 4 Dr, 4 Cs, 4 Hv, 4 Ac, 4 Ag	308	38	5e-91	29.29	975	1715
392	*Oryza barthii* (taxid:65489)	No significant similarity found						
393	*Oryza brachyantha* (taxid:4533)	40 At, 14 Lc, 13 Sp, 13 Hs, 13 Hv, 12 Cv, 12 Bf, 12 Sk, 11 Xl, 11 Dr, 11 Cs, 11 Ac, 9 Ag	417	57	6e-122	27.95	531	1632
394	*Oryza glumipatula* (taxid:40148)	5 At, 3 Hv, 2 Lc, 2 Dr, 1 Cv, 1 Sp, 1 Xl, 1 Hs, 1 Cs, 1 Ac	74.3	9	2e-14	30.07	997	1144
395	*Oryza longistaminata* (taxid:4528)	No significant similarity found						
396	*Oryza meridionalis* (taxid:40149)	No significant similarity found						
397	*Oryza nivara* (taxid:4536)	23 At, 8 Lc, 8 Hv, 7 Hs, 6 Sp, 6 Bf, 6 Ac, 5 Dr, 5 Cs, 5 Ag, 4 Cv, 4 Xl, 3 Sk	385	63	1e-108	27.48	521	1731
398	*Oryza punctata* (taxid:4537)	150 At, 65 Lc, 62 Hs, 61 Ac, 59 Xl, 58 Hv, 56 Cv, 56 Bf, 55 Sp, 52 Cs, 51 Sk, 46 Dr, 43 Ag	406	87	2e-120	32.16	862	1609

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
399	*Oryza sativa Japonica Group* (taxid:39947)	3412 At, 1664 Lc, 1651 Hs, 1651 Hv, 1628 Sp, 1550 Ac, 1495 Bf, 1448 Xl, 1443 Sk, 1434 Dr, 1390 Cv, 1384 Cs, 1292 Ag	642	81	0.0	31.65	627	1868
400	*Zizania latifolia* (taxid:58934)	No significant similarity found						
401	*Aegilops tauschii* (taxid:37682)	382 At, 157 Hs, 156 Lc, 151 Ac, 148 Sp, 135 Hv, 134 Sk, 134 Dr, 128 Bf, 118 Xl, 118 Cs, 116 Cv, 106 Ag	573	71	5e-167	32.11	528	1639
402	*Triticum urartu* (taxid:4572)	24 At, 8 Sp, 8 Bf, 8 Xl, 8 Hs, 8 Lc, 8 Hv, 8 Ac, 7 Cv, 7 Sk, 7 Dr, 7 Cs, 7 Ag	305	58	3e-106	28.92	531	1353
403	*Nelumbo nucifera* (taxid:4432)	212 At, 47 Sp, 43 Hv, 39 Hs, 38 Lc, 34 Bf, 33 Sk, 32 Xl, 30 Cs, 30 Ag, 27 Cv, 25 Dr, 22 Ac	379	55	9e-118	30.17	804	1705
404	*Lupinus angustifolius* (taxid:3871)	4 At, 2 Lc, 2 Ac, 1 Hv	84.3	27	6e-33	23.54	909	1349
405	*Phaseolus vulgaris* (taxid:3885)	213 At, 88 Sp, 74 Lc, 69 Hv, 63 Bf, 63 Hs, 63 Cs, 61 Dr, 59 Ac, 58 Ag, 57 Cv, 57 Xl, 56 Sk	429	56	7e-121	29.76	632	1705
406	*Cajanus cajan* (taxid:3821)	No significant similarity found						
407	*Vigna angularis var. angularis* (taxid:157739)	84 At, 8 Xl, 6 Bf, 6 Lc, 6 Hv, 4 Sk, 4 Hs, 4 Cs, 4 Ac, 3 Sp, 3 Dr, 2 Cv, 2 Ag	125	44	3e-27	33.33	1007	1222
408	*Vigna radiata var. radiata* (taxid:3916)	9 At, 1 Sk, 1 Cs, 1 Hv	110	8	8e-24	36.42	1738	1904
409	*Glycine max* (taxid:3847)	1342 At, 538 Lc, 527 Sp, 502 Hv, 492 Hs, 445 Sk, 440 Bf, 416 Ac, 407 Cs, 404 Xl, 403 Dr, 378 Cv, 350 Ag	508	77	1e-146	29.25	536	1904
410	*Glycine soja* (taxid:3848)	6 At, 3 Sp, 3 Bf, 3 Sk, 3 Lc, 3 Cs, 3 Hv, 2 Cv, 2 Xl, 2 Hs, 2 Dr, 2 Ac, 2 Ag	420	53	6e-120	30.42	536	1415
411	*Cicer arietinum* (taxid:3827)	64 At, 18 Sp, 18 Bf, 18 Cs, 17 Hv, 14 Hs, 12 Sk, 12 Dr, 12 Ac, 11 Xl, 11 Ag, 10 Lc, 7 Cv	140	18	1e-37	29.97	1060	1406

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
412	*Medicago truncatula* (taxid:3880)	2001 At, 864 Hs, 857 Lc, 856 Ac, 824 Sp, 814 Bf, 791 Hv, 775 Dr, 773 Cv, 771 Cs, 761 Sk, 760 Xl, 718 Ag	902	70	0.0	37.72	570	1906
413	*Trifolium pratense* (taxid:57577)	2 At	117	13	2e-28	33.72	1648	1906
414	*Lotus japonicus* (taxid:34305)	1909 At, 818 Sp, 807 Hs, 807 Lc, 777 Hv, 773 Ac, 768 Bf, 746 Dr, 737 Cv, 736 Xl, 734 Sk, 731 Cs, 702 Ag	1041	94	0.0	34.80	71	1880
415	*Malus x domestica* (taxid:3750)	354 At, 115 Hs, 113 Lc, 108 Hv, 105 Xl, 104 Ac, 100 Sk, 98 Bf, 97 Sp, 95 Cv, 94 Dr, 94 Cs, 88 Ag	724	66	0.0	40.76	539	1477
416	*Pyrus x bretschneideri* (taxid:225117)	66 At, 27 Lc, 21 Sp, 21 Hv, 20 Sk, 20 Xl, 19 Cs, 18 Bf, 18 Hs, 17 Ac, 17 Ag, 16 Cv, 15 Dr	930	71	0.0	38.22	534	1904
417	*Prunus mume* (taxid:102107)	181 At, 22 Hv, 21 Xl, 21 Hs, 21 Lc, 21 Ac, 20 Sk, 20 Cs, 19 Sp, 18 Cv, 15 Bf, 14 Dr, 13 Ag	617	69	0.0	30.40	531	1848
418	*Prunus persica* (taxid:3760)	371 At, 80 Lc, 76 Hs, 73 Hv, 69 Xl, 69 Ac, 66 Sp, 63 Sk, 60 Cv, 60 Cs, 57 Ag, 56 Bf, 56 Dr	620	77	0.0	31.02	593	1887
419	*Fragaria iinumae* (taxid:64939)	No significant similarity found						
420	*Fragaria nubicola* (taxid:60188)	No significant similarity found						
421	*Fragaria orientalis* (taxid:101013)	No significant similarity found						
422	*Fragaria vesca subsp. vesca* (taxid:101020)	230 At, 52 Hv, 51 Xl, 41 Hs, 37 Lc, 35 Ag, 31 Sp, 29 Sk, 28 Bf, 27 Cs, 26 Ac, 23 Cv, 23 Dr	501	69	1e-149	27.68	579	1905
423	*Fragaria x ananassa* (taxid:3747)	No significant similarity found						
424	*Morus notabilis* (taxid:981085)	22 At, 6 Hs, 3 Sk, 2 Sp, 2 Lc, 2 Hv, 1 Xl, 1 Dr	123	12	7e-28	35.00	1243	1477
425	*Cannabis sativa* (taxid:3483)	15 At, 5 Xl, 5 Hs, 5 Cs, 4 Sp, 4 Bf, 4 Lc, 3 Cv, 3 Hv, 2 Ac, 2 Ag, 1 Sk, 1 Dr	285	41	9e-79	30.34	870	1448

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
426	*Castanea mollissima* (taxid:60419)	No significant similarity found						
427	*Betula nana* (taxid:216990)	No significant similarity found						
428	*Cucumis melo* (taxid:3656)	3117 At, 1190 Lc, 1032 Hv, 1029 Sp, 947 Cs, 894 Dr, 877 Xl, 868 Hs, 867 Sk, 855 Bf, 817 Ac, 550 Cv, 548 Ag	442	57	7e-124	27.43	531	1634
429	*Cucumis sativus* (taxid:3659)	14 At, 3 Sk, 2 Hs, 1 Cv, 1 Sp, 1 Bf, 1 Xl, 1 Lc, 1 Dr, 1 Cs, 1 Hv, 1 Ac, 1 Ag	295	35	1e-79	29.72	967	1635
430	*Citrullus lanatus* (taxid:3654)	8 At, 2 Hs, 2 Lc, 2 Cs, 2 Ag, 1 Sp, 1 Sk, 1 Xl, 1 Dr, 1 Hv	86.7	8	6e-18	33.33	1268	1423
431	*Lagenaria siceraria* (taxid:3668)	No significant similarity found						
432	*Populus euphratica* (taxid:75702)	78 At, 22 Hs, 19 Lc, 17 Hv, 16 Dr, 13 Sp, 13 Sk, 13 Xl, 13 Ac, 12 Cv, 12 Bf, 12 Cs, 12 Ag	395	48	1e-110	28.87	531	1415
433	*Populus trichocarpa* (taxid:3694)	895 At, 271 Hs, 258 Lc, 254 Ac, 249 Hv, 231 Sp, 223 Xl, 219 Bf, 207 Sk, 198 Cv, 195 Ag, 186 Dr, 185 Cs	967	93	0.0	39.10	531	1904
434	*Jatropha curcas* (taxid:180498)	175 At, 47 Hs, 43 Lc, 41 Ac, 40 Sk, 40 Xl, 39 Hv, 37 Sp, 34 Cs, 34 Ag, 33 Bf, 30 Cv, 28 Dr	525	67	2e-157	30.01	589	1873
435	*Manihot esculenta subsp. flabellifolia* (taxid:478097)	No significant similarity found						
436	*Ricinus communis* (taxid:3988)	48 At, 7 Lc, 5 Sk, 5 Ac, 4 Xl, 4 Hs, 4 Dr, 3 Cs, 2 Cv, 2 Sp, 2 Hv, 2 Ag, 1 Bf	145	11	2e-37	32.16	1111	1336
437	*Linum usitatissimum* (taxid:4006)	40 At, 6 Xl, 5 Lc, 5 Dr, 4 Sk, 4 Hs, 3 Sp, 3 Bf, 2 Cv, 2 Hv, 2 Ac, 1 Cs, 1 Ag	172	67	3e-43	37.74	1215	1475
438	*Eucalyptus camaldulensis* (taxid:34316)	No significant similarity found						
439	*Eucalyptus grandis* (taxid:71139)	501 At, 121 Hs, 115 Sp, 108 Xl, 106 Bf, 105 Hv, 104 Sk, 101 Lc, 91 Cs, 89 Cv, 82 Dr, 80 Ac, 74 Ag	595	74	1e-179	29.48	548	1817

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
440	*Carica papaya* (taxid:3649)	263 At, 105 Sp, 75 Lc, 75 Hv, 72 Hs, 59 Bf, 59 Sk, 57 Cs, 55 Ac, 54 Cv, 47 Xl, 43 Dr, 41 Ag	454	63	0.0	45.60	975	1484
441	*Arabidopsis halleri subsp. gemmifera* (taxid:63677)	No significant similarity found						
442	*Arabidopsis lyrata subsp. lyrata* (taxid:81972)	12 At, 1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 Cs, 1 Hv, 1 Ac, 1 Ag	341	52	6e-94	29.38	866	1651
443	*Arabidopsis thaliana* (taxid:3702)	2163 At, 829 Lc, 796 Hs, 787 Sp, 707 Bf, 703 Ac, 693 Hv, 685 Sk, 679 Xl, 659 Cs, 641 Cv, 580 Ag, 573 Dr	3608	100	0.0	97.43	1	1809
444	*Camelina sativa* (taxid:90675)	756 At, 316 Hs, 309 Lc, 288 Sp, 282 Ac, 276 Hv, 252 Xl, 250 Sk, 226 Bf, 216 Cs, 207 Cv, 192 Dr, 188 Ag	1203	67	0.0	49.96	616	1906
445	*Capsella rubella* (taxid:81985)	99 At, 16 Sp, 16 Bf, 16 Hs, 16 Lc, 15 Cv, 13 Cs, 12 Sk, 12 Hv, 12 Ac, 8 Xl, 8 Dr, 8 Ag	1334	52	0.0	63.85	900	1906
446	*Brassica napus* (taxid:3708)	696 At, 536 Sp, 534 Hs, 534 Lc, 533 Bf, 531 Ac, 527 Hv, 524 Xl, 515 Sk, 492 Cs, 463 Dr, 458 Cv, 446 Ag	2301	99	0.0	59.09	5	1906
447	*Brassica oleracea var. oleracea* (taxid:109376)	738 At, 335 Ac, 332 Hs, 324 Lc, 315 Hv, 312 Sp, 278 Bf, 278 Sk, 271 Xl, 242 Cs, 223 Dr, 219 Ag, 216 Cv	521	78	5e-153	28.87	531	1845
448	*Brassica rapa* (taxid:3711)	1377 At, 808 Lc, 807 Hs, 783 Sp, 770 Ac, 760 Hv, 756 Sk, 740 Bf, 739 Cs, 734 Xl, 727 Dr, 722 Cv, 693 Ag	2309	99	0.0	59.39	6	1904
449	*Raphanus raphanistrum subsp. raphanistrum* (taxid:109997)	No significant similarity found						
450	*Raphanus sativus* (taxid:3726)	300 At, 129 Sp, 122 Xl, 122 Lc, 121 Hv, 119 Hs, 119 Ac, 117 Sk, 114 Cs, 113 Cv, 110 Bf, 107 Ag, 100 Dr	548	61	9e-159	29.50	531	1698

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
451	*Aethionema arabicum* (taxid:228871)	No significant similarity found						
452	*Arabis alpina* (taxid:50452)	8 At, 2 Cv, 2 Sp, 2 Bf, 2 Hs, 2 Dr, 2 Hv, 1 Sk, 1 Xl, 1 Lc, 1 Cs, 1 Ac, 1 Ag	698	81	0.0	63.10	789	1368
453	*Eutrema parvulum* (taxid:98039)	No significant similarity found						
454	*Eutrema salsugineum* (taxid:72664)	52 At, 2 Sk, 1 Sp, 1 Bf, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 Cs, 1 Hv, 1 Ac	176	6	6e-45	69.42	41	161
455	*Sisymbrium irio* (taxid:3730)	1 At	65.1	8	2e-11	29.82	1735	1897
456	*Leavenworthia alabamica* (taxid:310722)	4 At, 2 Cv, 2 Sp, 2 Bf, 2 Hs, 2 Lc, 2 Hv, 1 Sk, 1 Xl, 1 Ac	109	23	7e-54	37.42	1165	1327
457	*Tarenaya hassleriana* (taxid:28532)	129 At, 51 Hs, 51 Lc, 50 Sp, 48 Xl, 47 Bf, 45 Hv, 45 Ac, 41 Sk, 41 Dr, 41 Cs, 37 Cv, 35 Ag	428	71	5e-122	27.80	530	1643
458	*Gossypium arboreum* (taxid:29729)	35 At, 7 Xl, 6 Sp, 6 Lc, 6 Hv, 5 Sk, 5 Hs, 5 Dr, 5 Ac, 4 Cv, 4 Bf, 4 Ag, 2 Cs	199	83	8e-61	31.58	669	1085
459	*Gossypium raimondii* (taxid:29730)	933 At, 247 Lc, 209 Xl, 195 Sk, 191 Hv, 178 Hs, 159 Cv, 151 Sp, 134 Ac, 117 Bf, 108 Ag, 106 Dr, 101 Cs	482	44	3e-145	31.78	834	1690
460	*Theobroma cacao* (taxid:3641)	625 At, 88 Lc, 73 Hs, 71 Sp, 68 Hv, 64 Xl, 63 Bf, 63 Ac, 58 Sk, 55 Cs, 54 Dr, 44 Cv, 39 Ag	557	77	6e-163	29.65	547	1868
461	*Aquilaria agallochum* (taxid:1472648)	No nr/htgs data for this txid id						
462	*Azadirachta indica* (taxid:124943)	No significant similarity found						
463	*Citrus clementina* (taxid:85681)	23 At, 2 Hs	116	13	4e-27	29.96	22	272
464	*Citrus sinensis* (taxid:2711)	289 At, 44 Hs, 37 Hv, 36 Sk, 36 Xl, 34 Sp, 34 Lc, 28 Bf, 28 Ac, 27 Cv, 27 Ag, 26 Dr, 26 Cs	779	68	0.0	37.25	750	1904

No	Species and taxon identifier	# hits from each query (top hit red)	Top hit					
			bitscore	qcovs	evalue	pident	qstart	qend
465	*Vitis vinifera* (taxid:29760)	1576 At, 805 Sp, 776 Bf, 746 Xl, 728 Hs, 719 Dr, 694 Lc, 691 Ac, 690 Sk, 687 Hv, 636 Cs, 625 Ag, 622 Cv	540	69	3e-157	29.49	552	1873
466	*Amaranthus hypochondriacus* (taxid:28502)	5 At, 3 Hv, 2 Sp, 2 Bf, 2 Sk, 2 Hs, 2 Lc, 2 Cs, 2 Ac, 2 Ag, 1 Cv, 1 Xl, 1 Dr	88.2	4	1e-20	46.67	1110	1199
467	*Amaranthus tuberculatus* (taxid:277990)	No significant similarity found						
468	*Beta vulgaris subsp. vulgaris* (taxid:3555)	696 At, 391 Hv, 382 Hs, 365 Ac, 361 Lc, 352 Xl, 335 Sk, 334 Bf, 327 Sp, 325 Cs, 300 Dr, 298 Cv, 293 Ag	989	87	0.0	39.52	529	1904
469	*Spinacia oleracea* (taxid:3562)	1 At	57.8	4	1e-09	40.66	1815	1904
470	*Dianthus caryophyllus* (taxid:3570)	2 At	288	35	3e-116	34.08	1381	1905
471	*Actinidia chinensis* (taxid:3625)	No significant similarity found						
472	*Vaccinium macrocarpon* (taxid:13750)	20 At, 6 Hs, 6 Lc, 4 Bf, 4 Sk, 4 Xl, 3 Cv, 3 Sp, 3 Cs, 3 Hv, 3 Ac, 3 Ag, 2 Dr	188	69	2e-57	45.45	1102	1309
473	*Diospyros lotus* (taxid:55363)	No significant similarity found						
474	*Primula veris* (taxid:170927)	No significant similarity found						
475	*Solanum arcanum* (taxid:376710)	No significant similarity found						
476	*Solanum habrochaites* (taxid:62890)	3 Sk, 3 Lc, 3 Cs, 3 Ag, 3 At, 2 Sp, 2 Hs, 2 Dr, 2 Hv, 1 Cv, 1 Bf, 1 Xl, 1 Ac	229	61	1e-126	31.12	688	1097
477	*Solanum lycopersicum* (taxid:4081)	7088 At, 2074 Sp, 2069 Lc, 1984 Hv, 1776 Hs, 1752 Sk, 1612 Ac, 1565 Bf, 1523 Xl, 1436 Cs, 1155 Ag, 1132 Cv, 1064 Dr	533	86	2e-152	28.60	529	1888
478	*Solanum melongena* (taxid:4111)	2 Cv, 2 Hs, 2 Lc, 2 Ac, 2 Ag, 1 Sp, 1 Bf, 1 Xl, 1 Hv, 1 At	73.2	14	3e-14	26.69	934	1169

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
479	*Solanum pennellii* (taxid:28526)	6096 At, 1653 Lc, 1650 Sp, 1608 Hv, 1411 Sk, 1367 Hs, 1272 Ac, 1256 Bf, 1209 Xl, 1198 Cs, 917 Ag, 872 Cv, 862 Dr	514	86	2e-146	27.50	528	1905
480	*Solanum pimpinellifolium* (taxid:4084)	1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Hs, 1 Lc, 1 Cs, 1 Ag, 1 At	101	17	2e-22	28.10	859	1199
481	*Solanum tuberosum* (taxid:4113)	2068 At, 806 Hs, 799 Hv, 797 Sp, 772 Lc, 728 Ac, 724 Sk, 705 Bf, 683 Xl, 618 Cs, 567 Cv, 557 Dr, 523 Ag	689	82	0.0	33.78	531	1642
482	*Capsicum annuum* (taxid:4072)	32 At, 12 Hs, 11 Sp, 10 Cv, 10 Hv, 8 Bf, 8 Lc, 8 Ac, 7 Cs, 6 Sk, 3 Xl, 3 Dr, 3 Ag	186	47	2e-47	24.79	1198	1870
483	*Nicotiana sylvestris* (taxid:4096)	795 At, 267 Hs, 262 Lc, 258 Sp, 242 Hv, 235 Ac, 206 Xl, 205 Sk, 198 Bf, 189 Cs, 180 Cv, 166 Dr, 158 Ag	402	46	3e-114	30.22	595	1482
484	*Nicotiana tomentosiformis* (taxid:4098)	747 At, 242 Sp, 235 Hs, 227 Lc, 216 Hv, 202 Xl, 197 Ac, 196 Bf, 189 Sk, 166 Cs, 150 Cv, 133 Ag, 118 Dr	416	66	3e-168	39.19	859	1399
485	*Fraxinus excelsior* (taxid:38873)	No significant similarity found						
486	*Penstemon centranthifolius* (taxid:69924)	No significant similarity found						
487	*Penstemon grinnellii* (taxid:388155)	No significant similarity found						
488	*Sesamum indicum* (taxid:4182)	240 At, 99 Hv, 73 Xl, 73 Hs, 72 Bf, 67 Lc, 63 Sp, 60 Sk, 50 Ac, 49 Cs, 45 Cv, 32 Dr, 29 Ag	528	63	6e-160	29.69	561	1768
489	*Genlisea aurea* (taxid:192259)	No significant similarity found						
490	*Mimulus guttatus* (taxid:4155)	225 At, 86 Hs, 73 Lc, 73 Hv, 70 Sp, 69 Sk, 66 Bf, 65 Ac, 64 Cs, 63 Xl, 63 Dr, 62 Cv, 56 Ag	578	67	6e-177	30.56	587	1879
491	*Conyza canadensis* (taxid:72917)	No significant similarity found						
ECHINOIDEA								

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
492	*Lytechinus variegatus* (taxid:7654)	16 Bf, 14 Hs, 14 Lc, 13 Ac, 12 Sp, 12 Xl, 11 Sk, 11 Dr, 11 Cs, 9 Cv, 9 Hv, 8 Ag, 6 At	473	74	8e-146	30.78	487	1640
493	*Strongylocentrotus purpuratus* (taxid:7668)	1930 Sp, 1817 Bf, 1634 Hs, 1619 Lc, 1382 Sk, 1334 Dr, 1320 Ac, 1299 Xl, 1263 Hv, 1197 Cs, 1092 Ag, 1020 Cv, 930 At	2730	100	0.0	89.85	128	1574
ASTEROIDEA								
494	*Patiria miniata* (taxid:46514)	No significant similarity found						
ENTEROPNEUSTA								
495	*Saccoglossus kowalevskii* (taxid:10224)	91 Sk, 86 Sp, 80 Bf, 71 Hs, 67 Xl, 66 Lc, 60 Dr, 56 Hv, 56 Ac, 48 Cs, 38 Ag, 33 At, 30 Cv	1451	99	0.0	56.67	539	1774
TUNICATA								
496	*Ciona intestinalis* (taxid:7719)	12 Cs, 10 Ag, 9 Sp, 8 Bf, 8 Sk, 8 Xl, 8 Hs, 8 Lc, 8 Dr, 8 Hv, 8 Ac, 8 At, 7 Cv	482	74	1e-141	29.28	405	1624
497	*Ciona savignyi* (taxid:51511)	2 Cs, 1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Hs, 1 Lc, 1 Dr, 1 Hv, 1 Ac, 1 Ag, 1 At	773	79	0.0	37.95	469	1626
498	*Botryllus schlosseri* (taxid:30301)	1 Cv, 1 Sp, 1 Bf, 1 Sk, 1 Xl, 1 Dr, 1 Cs, 1 Ac, 1 Ag, 1 At	83.2	16	5e-17	28.29	742	989
499	*Oikopleura dioica* (taxid:34765)	No significant similarity found						
LEPTOCARDII								
500	*Branchiostoma floridae* (taxid:7739)	39 Bf, 35 Ac, 34 Lc, 32 Sp, 31 Sk, 28 Hs, 26 Cs, 25 Dr, 24 Ag, 23 Cv, 23 Xl, 20 At, 19 Hv	545	70	2e-159	32.30	407	1569
CEPHALASPIDOMORPHI								
501	*Lethenteron camtschaticum* (taxid:980415)	2 Sp, 2 Bf, 2 Hs, 2 Lc, 2 Ac, 2 Ag, 1 Cv, 1 Sk, 1 Xl, 1 Dr, 1 Cs, 1 Hv, 1 At	103	16	7e-23	25.35	1138	1485
502	*Petromyzon marinus* (taxid:7757)	15 Sp, 15 Bf, 14 Dr, 14 Ac, 14 At, 13 Hs, 13 Lc, 10 Sk, 10 Xl, 10 Cs, 10 Ag, 8 Cv, 8 Hv	122	28	2e-28	25.16	728	1180
SARCOPTERYGII								

			Top hit					
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**qcovs**	**evalue**	**pident**	**qstart**	**qend**
503	*Latimeria chalumnae* (taxid:7897)	286 Lc, 137 Hs, 105 Ac, 86 Xl, 77 Sk, 72 Sp, 71 Hv, 64 Bf, 62 Cs, 60 Cv, 56 Dr, 51 At, 43 Ag	1331	93	0.0	72.08	377	1300

Table 5: L1s in the genome, found using LASTZ

Table B.5: **Presence of L1:** Shows which species contain evidence of L1 elements based on query-driven iterative similarity searches with LASTZ. Any hits found had to satisfy a 'reciprocal best hit' check: they were screened with CENSOR against the Repbase library of known repeats, and kept only if the best hit was an L1 element (not some other repeat like BovB elements). Overlapping hits were merged to produce a non-redundant (unique) set of L1s for each genome. The Notes column highlights interesting observations or additional information about the L1 hits, particularly in species that have been previously studied. Species are listed in the same order as Table B.1, for easy reference to the common name. L1 hits that are thought to be due to contamination are marked as such, and were excluded from further analyses.

No	Species	# unique L1 hits	Length distribution (bp)	Notes
MAMMALIA				
1	*Tachyglossus aculeatus*	0	-	
2	*Ornithorhynchus anatinus*	0	-	The few hits found were dismissed as contamination due to very high similarity to wallaby L1s. Both monotreme genomes seem to be missing L1s, yet contain an abundance of L2s.
3	*Monodelphis domestica*	90570	min 32, avg 1858, med 1297, max 12259	Has been shown to have active full-length L1s (Gu et al, 2007; Mikkelsen et al, 2007). Gallus 2015 found more the 10,000 near full-length L1 copies, but only 500 were potentially active.
4	*Macropus eugenii*	119716	min 33, avg 936, med 376, max 10898	Few full-length L1s could be assembled in the initial assembly.
5	*Sarcophilus harrisii*	117881	min 32, avg 755, med 234, max 8834	Jurka et al, 2011: most recently active element was L1-1_SH (6676bp length consensus). Gallus et al, 2015: found 384 L1 copies >6kb, which were either full-length or near full-length; all inactive due to mutations. Screening a second Tas devil genome also found only inactive elements.
6	*Dasypus novemcinctus*	200433	min 34, avg 1153, med 559, max 13923	
7	*Choloepus hoffmanni*	166824	min 35, avg 870, med 609, max 11920	
8	*Chrysochloris asiatica*	58332	min 35, avg 853, med 97, max 7157	
9	*Echinops telfairi*	22557	min 36, avg 1324, med 731, max 7224	
10	*Orycteropus afer afer*	123833	min 34, avg 946, med 134, max 10293	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
11	*Elephantulus edwardii*	64503	min 33, avg 740, med 70, max 7108	
12	*Trichechus manatus latirostris*	142979	min 32, avg 1066, med 502, max 14101	
13	*Procavia capensis*	156194	min 32, avg 1132, med 737, max 11303	
14	*Loxodonta africana*	114985	min 33, avg 2142, med 1203, max 14036	
15	*Erinaceus europaeus*	35717	min 36, avg 1090, med 89, max 10024	
16	*Sorex araneus*	61801	min 33, avg 800, med 370, max 10424	
17	*Condylura cristata*	21240	min 35, avg 765, med 91, max 7199	
18	*Pteropus alecto*	95017	min 33, avg 836, med 333, max 12178	All full-length L1s are extinct (Cantrell et al, 2008; Zhang et al, 2013)
19	*Pteropus vampyrus*	85330	min 35, avg 785, med 301, max 9856	All full-length L1s are extinct (Cantrell et al, 2008; Zhang et al, 2013)
20	*Eidolon helvum*	67370	min 34, avg 749, med 212, max 8961	
21	*Megaderma lyra*	57559	min 32, avg 751, med 203, max 9765	
22	*Rhinolophus ferrumequinum*	75265	min 33, avg 664, med 207, max 9717	
23	*Pteronotus parnellii*	68630	min 33, avg 809, med 179, max 9062	
24	*Eptesicus fuscus*	66233	min 35, avg 559, med 210, max 11254	
25	*Myotis brandtii*	85947	min 37, avg 628, med 244, max 10288	
26	*Myotis davidii*	74106	min 36, avg 540, med 211, max 11658	
27	*Myotis lucifugus*	82275	min 33, avg 790, med 306, max 12207	
28	*Ceratotherium simum simum*	112606	min 32, avg 1016, med 304, max 12095	
29	*Equus przewalskii*	114836	min 35, avg 1129, med 348, max 15366	
30	*Equus caballus* (Thoroughbred)	86806	min 37, avg 1534, med 481, max 13504	
31	*Equus caballus* (Mongolian)	113546	min 34, avg 1076, med 317, max 15282	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
32	*Manis pentadactyla*	130736	min 32, avg 840, med 180, max 10267	
33	*Felis catus*	82057	min 33, avg 1312, med 488, max 11372	
34	*Panthera tigris altaica*	124302	min 33, avg 961, med 366, max 13604	
35	*Canis lupus familiaris*	91247	min 34, avg 1321, med 510, max 11672	Known to contain active L1s.
36	*Ursus maritimus*	114732	min 36, avg 998, med 330, max 13373	
37	*Ailuropoda melanoleuca*	111345	min 35, avg 883, med 256, max 11844	Gallus et al, 2015: 3'-end analysis of L1 copies found 8000 3' copies, 1811 of which were 100% identical.
38	*Leptonychotes weddellii*	123675	min 34, avg 809, med 282, max 10738	
39	*Odobenus rosmarus divergens*	131771	min 35, avg 905, med 337, max 11279	
40	*Mustela putorius furo*	115201	min 35, avg 912, med 312, max 10819	
41	*Camelus dromedarius*	108860	min 34, avg 709, med 281, max 11206	
42	*Camelus ferus*	106834	min 34, avg 677, med 278, max 11795	
43	*Vicugna pacos*	140145	min 34, avg 597, med 298, max 10598	
44	*Sus scrofa* (Duroc)	106518	min 33, avg 1034, med 411, max 11196	
45	*Sus scrofa* (Tibetan)	153974	min 34, avg 828, med 388, max 13253	
46	*Sus scrofa* (Ellegaard Gottingen minipig)	143073	min 34, avg 820, med 372, max 11185	
47	*Balaenoptera acutorostrata scammoni*	160363	min 35, avg 931, med 389, max 10933	
48	*Physeter catodon*	154909	min 34, avg 920, med 406, max 12918	
49	*Lipotes vexillifer*	161501	min 34, avg 1566, med 510, max 15715	
50	*Tursiops truncatus*	204226	min 33, avg 581, med 294, max 14298	
51	*Orcinus orca*	163509	min 34, avg 1001, med 430, max 12828	
52	*Pantholops hodgsonii*	103273	min 32, avg 711, med 250, max 10045	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
53	*Capra hircus*	80532	min 31, avg 723, med 280, max 11295	
54	*Ovis aries* (Texel)	68523	min 33, avg 979, med 332, max 11832	
55	*Ovis aries musimon*	94579	min 34, avg 641, med 244, max 12652	
56	*Bubalus bubalis*	148873	min 31, avg 637, med 372, max 16417	
57	*Bison bison bison*	165700	min 31, avg 612, med 364, max 12472	Lots of small hits <1kb
58	*Bos mutus*	106670	min 31, avg 690, med 264, max 12276	
59	*Bos indicus*	73472	min 31, avg 993, med 337, max 12582	
60	*Bos taurus*	80521	min 31, avg 1019, med 376, max 12797	
61	*Ochotona princeps*	22298	min 36, avg 783, med 298, max 7183	
62	*Oryctolagus cuniculus*	54371	min 36, avg 1591, med 798, max 13252	
63	*Ictidomys tridecemlineatus*	77077	min 34, avg 953, med 405, max 8995	Only has extinct L1s (full-length but none active) (Platt and Ray, 2012)
64	*Heterocephalus glaber*	68845	min 34, avg 1233, med 263, max 9489	
65	*Fukomys damarensis*	60536	min 35, avg 1764, med 932, max 10470	
66	*Cavia aperea*	82452	min 35, avg 979, med 500, max 12977	
67	*Cavia porcellus*	91003	min 37, avg 2032, med 1210, max 13197	
68	*Chinchilla lanigera*	60066	min 36, avg 1337, med 524, max 9883	
69	*Octodon degus*	54941	min 36, avg 1463, med 341, max 7751	
70	*Dipodomys ordii*	35919	min 31, avg 886, med 92, max 7728	
71	*Jaculus jaculus*	32853	min 34, avg 1493, med 1275, max 7358	
72	*Nannospalax galili*	55684	min 36, avg 1756, med 1729, max 7868	
73	*Mesocricetus auratus*	55501	min 35, avg 812, med 206, max 9681	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
74	*Cricetulus griseus*	67154	min 36, avg 970, med 249, max 11338	
75	*Microtus ochrogaster*	30787	min 36, avg 778, med 195, max 6404	
76	*Peromyscus maniculatus bairdii*	46108	min 33, avg 852, med 247, max 9389	
77	*Rattus norvegicus*	95516	min 34, avg 1601, med 848, max 11866	
78	*Mus musculus*	99565	min 34, avg 1820, med 873, max 18858	Known to contain active L1s.
79	*Tupaia belangeri*	67395	min 36, avg 750, med 459, max 10269	
80	*Tupaia chinensis*	73595	min 34, avg 987, med 468, max 10640	
81	*Galeopterus variegatus*	166885	min 33, avg 791, med 146, max 9846	
82	*Otolemur garnettii*	100949	min 34, avg 1003, med 375, max 8925	
83	*Microcebus murinus*	105998	min 34, avg 637, med 281, max 10808	
84	*Tarsius syrichta*	168419	min 34, avg 1365, med 356, max 14075	Tarsier has the most L1 hits out of all the primates.
85	*Callithrix jacchus*	145866	min 33, avg 1293, med 682, max 12300	
86	*Saimiri boliviensis boliviensis*	126736	min 34, avg 1154, med 516, max 11681	
87	*Rhinopithecus roxellana*	161163	min 34, avg 1220, med 475, max 15293	
88	*Nasalis larvatus*	100412	min 34, avg 1079, med 548, max 10255	
89	*Chlorocebus sabaeus*	114335	min 34, avg 1339, med 606, max 12077	
90	*Macaca fascicularis*	114193	min 34, avg 1403, med 646, max 12442	
91	*Macaca mulatta*	108459	min 33, avg 936, med 376, max 10898	
92	*Papio anubis*	123791	min 34, avg 1279, med 629, max 12138	
93	*Nomascus leucogenys*	126230	min 33, avg 1417, med 692, max 13430	
94	*Pongo abelii*	157381	min 33, avg 1268, med 638, max 13087	
95	*Gorilla gorilla gorilla*	127050	min 34, avg 1187, med 603, max 16253	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
96	*Pan paniscus*	149841	min 33, avg 1353, med 591, max 14715	
97	*Pan troglodytes*	130898	min 33, avg 1404, med 665, max 16238	
98	*Homo sapiens*	118667	min 34, avg 1562, med 716, max 18773	Thought to be about 6000-7000 full-length L1s per human genome, of which only 80-100 are proven to be retrotranspositionally-competent in cell culture (Lander et al, 2001; Brouha et al, 2003; Penzkofer et al, 2005; Khan et al, 2006).
SAUROPSIDA				
99	*Apalone spinifera*	2236	min 82, avg 512, med 357, max 5911	Ancient L1s?
100	*Pelodiscus sinensis*	1623	min 84, avg 754, med 477, max 7157	
101	*Chelonia mydas*	2819	min 81, avg 839, med 462, max 8320	
102	*Chrysemys picta bellii*	2343	min 76, avg 975, med 524, max 8573	
103	*Struthio camelus australis*	47	min 75, avg 384, med 248, max 3825	
104	*Tinamus guttatus*	19	min 83, avg 332, med 281, max 1197	
105	*Anas platyrhynchos*	29	min 98, avg 518, med 208, max 4578	
106	*Lyrurus tetrix tetrix*	13	min 72, avg 333, med 280, max 780	
107	*Gallus gallus*	15	min 87, avg 561, med 240, max 4412	
108	*Coturnix japonica*	3	min 82, avg 168, med 143, max 280	
109	*Meleagris gallopavo*	7	min 86, avg 752, med 279, max 4020	
110	*Colinus virginianus*	121	min 70, avg 511, med 376, max 4090	
111	*Acanthisitta chloris*	33	min 90, avg 695, med 289, max 4493	
112	*Manacus vitellinus*	31	min 102, avg 600, med 279, max 4602	
113	*Zonotrichia albicollis*	30	min 65, avg 348, med 207, max 4512	
114	*Geospiza fortis*	20	min 114, avg 715, med 296, max 4489	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
115	*Serinus canaria*	381	min 61, avg 152, med 102, max 4516	Some hits show high similarity to human L1s - there were dismissed as contamination. Hits that showed more similarity to turtle or reptile L1s were kept.
116	*Taeniopygia guttata*	22	min 84, avg 638, med 316, max 4354	
117	*Ficedula albicollis*	28	min 108, avg 523, med 273, max 4569	
118	*Pseudopodoces humilis*	34	min 68, avg 503, med 211, max 4312	
119	*Corvus brachyrhynchos*	45	min 77, avg 459, med 248, max 4515	
120	*Corvus cornix cornix*	32	min 76, avg 542, med 257, max 4586	
121	*Ara macao*	46	min 107, avg 382, med 283, max 2083	
122	*Amazona vittata*	58	min 72, avg 378, med 272, max 3513	
123	*Melopsittacus undulatus*	34	min 73, avg 540, med 293, max 3737	
124	*Nestor notabilis*	47	min 95, avg 515, med 267, max 3928	
125	*Falco cherrug*	49	min 101, avg 489, med 280, max 4512	
126	*Falco peregrinus*	47	min 110, avg 490, med 291, max 3806	
127	*Cariama cristata*	54	min 125, avg 564, med 270, max 4604	
128	*Merops nubicus*	33	min 121, avg 652, med 282, max 4075	
129	*Picoides pubescens*	29	min 72, avg 220, med 208, max 641	
130	*Buceros rhinoceros silvestris*	35	min 71, avg 468, med 289, max 4513	
131	*Apaloderma vittatum*	28	min 117, avg 787, med 295, max 3858	
132	*Leptosomus discolour*	44	min 128, avg 500, med 267, max 3646	
133	*Haliaeetus albicilla*	52	min 116, avg 585, med 282, max 4593	
134	*Haliaeetus leucocephalus*	56	min 85, avg 562, med 274, max 4600	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
135	*Aquila chrysaetos Canadensis*	53	min 127, avg 570, med 301, max 4567	
136	*Cathartes aura*	66	min 108, avg 498, med 305, max 4211	
137	*Tyto alba*	90	min 75, avg 391, med 269, max 4540	
138	*Colius striatus*	33	min 85, avg 664, med 281, max 4391	
139	*Charadrius vociferus*	51	min 106, avg 558, med 287, max 4520	
140	*Balearica regulorum gibbericeps*	51	min 111, avg 601, med 295, max 4608	
141	*Chlamydotis macqueenii*	46	min 69, avg 609, med 270, max 4065	
142	*Cuculus canorus*	33	min 119, avg 526, med 252, max 3744	
143	*Fulmarus glacialis*	58	min 70, avg 550, med 280, max 4628	
144	*Aptenodytes forsteri*	67	min 76, avg 586, med 269, max 4724	
145	*Pygoscelis adeliae*	59	min 112, avg 628, med 286, max 4834	
146	*Phalacrocorax carbo*	45	min 109, avg 580, med 292, max 3731	
147	*Pelecanus crispus*	50	min 121, avg 606, med 279, max 4581	
148	*Nipponia nippon*	57	min 76, avg 562, med 277, max 4564	
149	*Egretta garzetta*	50	min 105, avg 614, med 282, max 4598	
150	*Phaethon lepturus*	43	min 151, avg 633, med 301, max 4615	
151	*Gavia stellata*	60	min 88, avg 567, med 275, max 4577	
152	*Tauraco erythrolophus*	41	min 115, avg 623, med 289, max 4514	
153	*Opisthocomus hoazin*	65	min 146, avg 607, med 304, max 4184	
154	*Columba livia*	46	min 109, avg 541, med 279, max 4523	
155	*Pterocles gutturalis*	50	min 106, avg 617, med 279, max 4508	
156	*Calypte anna*	24	min 137, avg 640, med 286, max 3924	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
157	*Chaetura pelagica*	24	min 100, avg 596, med 247, max 4553	
158	*Caprimulgus carolinensis*	62	min 83, avg 482, med 265, max 4374	
159	*Eurypyga helias*	33	min 102, avg 525, med 267, max 4514	
160	*Mesitornis unicolor*	34	min 130, avg 584, med 279, max 4484	
161	*Podiceps cristatus*	43	min 127, avg 485, med 286, max 4730	
162	*Phoenicopterus ruber ruber*	62	min 124, avg 540, med 285, max 4626	
163	*Alligator mississippiensis*	2914	min 72, avg 1505, med 611, max 13520	
164	*Alligator sinensis*	3107	min 68, avg 1381, med 563, max 8097	
165	*Crocodylus porosus*	2660	min 82, avg 1234, med 534, max 7709	
166	*Gavialis gangeticus*	2841	min 73, avg 1272, med 528, max 13044	
167	*Pogona vitticeps*	3592	min 74, avg 549, med 395, max 4963	
168	*Anolis carolinensis*	2058	min 82, avg 2400, med 766, max 9160	Novick et al (2009) found 170 full-length and 626 truncated elements, making up 20 distinct L1 families each with low copy number. Their cutoff length for FL L1s was 5.25kb.
169	*Vipera berus berus*	4171	min 69, avg 823, med 542, max 5375	
170	*Crotalus mitchellii pyrrhus*	3142	min 76, avg 515, med 378, max 4019	
171	*Ophiophagus hannah*	7777	min 80, avg 634, med 334, max 5153	
172	*Python bivittatus*	2490	min 74, avg 979, med 549, max 5737	
AMPHIBIA				
173	*Nanorana parkeri*	1886	min 74, avg 1321, med 1113, max 5110	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
174	*Xenopus tropicalis*	1305	min 694, avg 2356, med 1836, max 7029	Kordis et al (2006) found 126 diverse L1 families that contain structurally intact L1s. More than 70 diverse families belong to L1 group C, although they are small in size, e.g. 1-5 members/family found.
NEOPTERYGII				
175	*Lepisosteus oculatus*	176	min 88, avg 1426, med 759, max 5363	
176	*Anguilla anguilla*	749	min 72, avg 790, med 312, max 5766	
177	*Anguilla japonica*	864	min 74, avg 925, med 471, max 5066	
178	*Danio rerio*	1566	min 58, avg 2428, med 1872, max 8441	Furano et al (2004) found over 30 distinct L1 lineages. Kordis et al (2006) found 59 L1 families.
179	*Astyanax mexicanus*	361	min 69, avg 699, med 434, max 4962	
180	*Oryzias latipes*	1722	min 62, avg 937, med 690, max 5918	Kordis et al (2006) found 17 L1 families. Previously analysed Swimmer1 L1 element (Duvernell and Turner, 1998) is only one of the 17 diverse L1 families that exist here.
181	*Poecilia formosa*	476	min 72, avg 1474, med 947, max 5610	
182	*Xiphophorus maculatus*	502	min 69, avg 1106, med 742, max 5655	
183	*Fundulus heteroclitus*	1068	min 64, avg 1085, med 731, max 5485	Duvernell et al (2004) used southern blot analyses and PCR to find low copy no, ancient but active L1s. Three lineages: A1 had 16/20 copies with intact ORFs, A2 had 14/17, B had 9/10.
184	*Takifugu flavidus*	330	min 77, avg 979, med 591, max 5147	
185	*Takifugu rubripes*	263	min 69, avg 1143, med 568, max 5497	Kordis et al (2006) found 1 remaining full-length L1
186	*Tetraodon nigroviridis*	58	min 73, avg 1725, med 1014, max 5421	Kordis et al (2006) found 0 L1s
187	*Cynoglossus semilaevis*	81	min 73, avg 780, med 459, max 4385	
188	*Haplochromis burtoni*	783	min 67, avg 1173, med 893, max 6063	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
189	*Pundamilia nyererei*	802	min 65, avg 1150, med 919, max 5895	
190	*Maylandia zebra*	897	min 66, avg 1189, med 939, max 6900	
191	*Neolamprologus brichardi*	789	min 75, avg 1124, med 833, max 5431	
192	*Oreochromis niloticus*	1223	min 67, avg 1302, med 949, max 6544	
193	*Sebastes nigrocinctus*	467	min 79, avg 852, med 578, max 6241	
194	*Sebastes rubrivinctus*	459	min 74, avg 835, med 547, max 6040	
195	*Gasterosteus aculeatus*	179	min 67, avg 1034, med 488, max 4992	Kordis et al (2006) found 0 L1s.
196	*Gadus morhua*	913	min 55, avg 630, med 336, max 5389	
CHONDRICHTHYES				
197	*Callorhinchus milii*	113	min 81, avg 335, med 353, max 1149	
198	*Carcharhinus brachyurus*	2426	min 49, avg 275, med 137, max 5019	
ECDYSOZOA				
199	*Ephemera danica*	0	-	
200	*Ladona fulva*	0	-	
201	*Pediculus humanus corporis*	0	-	
202	*Frankliniella occidentalis*	0	-	
203	*Diaphorina citri*	0	-	
204	*Pachypsylla venusta*	0	-	
205	*Acyrthosiphon pisum*	0	-	
206	*Nilaparvata lugens*	0	-	
207	*Oncopeltus fasciatus*	0	-	
208	*Rhodnius prolixus*	0	-	
209	*Cimex lectularius*	0	-	
210	*Onthophagus taurus*	0	-	
211	*Agrilus planipennis*	0	-	
212	*Tribolium castaneum*	0	-	
213	*Anoplophora glabripennis*	0	-	
214	*Leptinotarsa decemlineata*	0	-	
215	*Dendroctonus ponderosae*	0	-	
216	*Mengenilla moldrzyki*	0	-	
217	*Aedes aegypti*	1048	min 71, avg 2579, med 2518, max 5629	
218	*Culex quinquefasciatus*	387	min 100, avg 1975, med 734, max 5491	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
219	*Anopheles albimanus*	39	min 88, avg 350, med 176, max 3144	
220	*Anopheles arabiensis*	60	min 72, avg 399, med 212, max 2169	
221	*Anopheles atroparvus*	64	min 74, avg 378, med 163, max 2800	
222	*Anopheles christyi*	44	min 89, avg 409, med 210, max 1675	
223	*Anopheles culicifacies*	40	min 72, avg 454, med 214, max 3791	
224	*Anopheles darlingi*	33	min 71, avg 344, med 170, max 3128	
225	*Anopheles dirus*	80	min 88, avg 485, med 359, max 2924	
226	*Anopheles epiroticus*	51	min 85, avg 593, med 239, max 4128	
227	*Anopheles farauti*	58	min 72, avg 765, med 489, max 4674	
228	*Anopheles funestus*	41	min 70, avg 819, med 306, max 4459	
229	*Anopheles gambiae*	67	min 75, avg 1000, med 403, max 4880	Biedler and Tu (2003) found 5 divergent L1 families, 2 of which have multiple full-length copies (4 copies of Ag-L1-5, 2 copies of Ag-L1-2). Presumably the other 3 families only have 1 full-length representative (so 9 full-length L1s total).
230	*Anopheles maculatus*	33	min 102, avg 846, med 605, max 2325	
231	*Anopheles melas*	43	min 86, avg 450, med 212, max 3158	
232	*Anopheles merus*	94	min 79, avg 500, med 236, max 4176	
233	*Anopheles minimus*	38	min 107, avg 674, med 271, max 3995	
234	*Anopheles quadriannulatus*	62	min 77, avg 395, med 196, max 2472	
235	*Anopheles sinensis*	39	min 92, avg 902, med 213, max 4275	
236	*Anopheles stephensi*	51	min 91, avg 530, med 372, max 3220	
237	*Mayetiola destructor*	0	-	
238	*Lutzomyia longipalpis*	0	-	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
239	*Phlebotomus papatasi*	0	-	
240	*Ceratitis capitata*	102	min 70, avg 512, med 325, max 7274	
241	*Drosophila albomicans*	50	min 73, avg 260, med 136, max 1699	
242	*Drosophila ananassae*	23	min 68, avg 240, med 179, max 694	
243	*Drosophila biarmipes*	25	min 78, avg 387, med 225, max 1759	
244	*Drosophila bipectinata*	18	min 68, avg 508, med 269, max 1503	
245	*Drosophila elegans*	32	min 63, avg 243, med 186, max 1241	
246	*Drosophila erecta*	24	min 82, avg 269, med 200, max 956	
247	*Drosophila eugracilis*	25	min 89, avg 337, med 227, max 1472	
248	*Drosophila ficusphila*	64	min 88, avg 974, med 342, max 3911	
249	*Drosophila grimshawi*	70	min 76, avg 190, med 114, max 882	
250	*Drosophila kikkawai*	33	min 87, avg 289, med 221, max 1311	
251	*Drosophila melanogaster*	16	min 97, avg 307, med 134, max 2001	
252	*Drosophila miranda*	29	min 76, avg 307, med 148, max 4515	
253	*Drosophila mojavensis*	113	min 68, avg 316, med 210, max 1345	
254	*Drosophila persimilis*	36	min 67, avg 196, med 122, max 837	
255	*Drosophila pseudoobscura pseudoobscura*	34	min 70, avg 207, med 157, max 953	
256	*Drosophila rhopaloa*	18	min 104, avg 248, med 179, max 1143	
257	*Drosophila sechellia*	17	min 92, avg 253, med 225, max 629	
258	*Drosophila simulans*	19	min 77, avg 194, med 169, max 531	
259	*Drosophila suzukii*	37	min 86, avg 471, med 174, max 5555	
260	*Drosophila takahashii*	25	min 96, avg 257, med 180, max 1660	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
261	*Drosophila virilis*	67	min 69, avg 213, med 125, max 2055	
262	*Drosophila willistoni*	41	min 72, avg 213, med 152, max 1050	
263	*Drosophila yakuba*	22	min 66, avg 359, med 182, max 2240	
264	*Musca domestica*	83	min 84, avg 382, med 269, max 1622	
265	*Glossina austeni*	143	min 71, avg 265, med 175, max 1413	
266	*Glossina brevipalpis*	180	min 81, avg 307, med 226, max 1358	
267	*Glossina fuscipes fuscipes*	135	min 69, avg 218, med 146, max 1268	
268	*Glossina morsitans morsitans*	124	min 72, avg 231, med 151, max 1306	
269	*Glossina pallidipes*	122	min 75, avg 228, med 164, max 1505	
270	*Limnephilus lunatus*	0	-	
271	*Papilio glaucus*	0	-	
272	*Papilio polytes*	0	-	
273	*Papilio xuthus*	0	-	
274	*Heliconius melpomene melpomene*	0	-	
275	*Melitaea cinxia*	0	-	
276	*Danaus plexippus*	0	-	
277	*Bombyx mori*	0	-	
278	*Manduca sexta*	0	-	
279	*Plutella xylostella*	0	-	
280	*Athalia rosae*	0	-	
281	*Cephus cinctus*	0	-	
282	*Orussus abietinus*	20	min 89, avg 299, med 234, max 1082	
283	*Ceratosolen solmsi marchali*	0	-	
284	*Nasonia giraulti*	0	-	
285	*Nasonia longicornis*	0	-	
286	*Nasonia vitripennis*	0	-	
287	*Copidosoma floridanum*	0	-	
288	*Trichogramma pretiosum*	0	-	
289	*Microplitis demolitor*	69	min 88, avg 676, med 799, max 1396	
290	*Megachile rotundata*	0	-	
291	*Apis dorsata*	0	-	
292	*Apis florea*	0	-	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
293	*Apis mellifera*	0	-	
294	*Bombus impatiens*	0	-	
295	*Bombus terrestris*	0	-	
296	*Linepithema humile*	0	-	
297	*Camponotus floridanus*	0	-	
298	*Acromyrmex echinatior*	0	-	
299	*Atta cephalotes*	0	-	
300	*Solenopsis invicta*	0	-	
301	*Pogonomyrmex barbatus*	0	-	
302	*Harpegnathos saltator*	0	-	
303	*Cerapachys biroi*	0	-	
304	*Blattella germanica*	955	min 72, avg 429, med 223, max 3360	
305	*Zootermopsis nevadensis*	29	min 87, avg 476, med 231, max 2199	
306	*Daphnia pulex*	27	min 83, avg 1241, med 357, max 4805	
307	*Eurytemora affinis*	277	min 85, avg 439, med 310, max 3719	
308	*Hyalella azteca*	137	min 83, avg 281, med 225, max 1526	
309	*Strigamia maritima*	0	-	
310	*Stegodyphus mimosarum*	0	-	
311	*Latrodectus hesperus*	78	min 74, avg 530, med 618, max 1100	
312	*Parasteatoda tepidariorum*	0	-	
313	*Tetranychus urticae*	0	-	
314	*Dermatophagoides farinae*	0	-	
315	*Sarcoptes scabiei type canis*	0	-	
316	*Achipteria coleoptrata*	0	-	
317	*Hypochthonius rufulus*	0	-	
318	*Platynothrus peltifer*	0	-	
319	*Steganacarus magnus*	0	-	
320	*Ixodes ricinus*	446	min 70, avg 1137, med 1042, max 3194	
321	*Ixodes scapularis*	1106	min 72, avg 1101, med 975, max 3447	
322	*Rhipicephalus microplus*	88	min 182, avg 923, med 771, max 2946	
323	*Metaseiulus occidentalis*	34	min 96, avg 604, med 371, max 1334	
324	*Varroa destructor*	0	-	
325	*Centruroides exilicauda*	65	min 71, avg 640, med 492, max 2400	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
326	*Mesobuthus martensii*	0	-	
327	*Limulus polyphemus*	145	min 65, avg 712, med 409, max 3393	
328	*Trichinella spiralis*	14	min 79, avg 466, med 160, max 1326	
329	*Ascaris suum*	0	-	
330	*Elaeophora elaphi*	0	-	
331	*Onchocerca volvulus*	0	-	
332	*Steinernema monticolum*	0	-	
333	*Panagrellus redivivus*	0	-	
334	*Haemonchus contortus*	0	-	
335	*Necator americanus*	0	-	
336	*Heterorhabditis bacteriophora*	0	-	
337	*Caenorhabditis angaria*	0	-	
338	*Caenorhabditis brenneri*	0	-	
339	*Caenorhabditis briggsae*	0	-	
340	*Caenorhabditis elegans*	0	-	
341	*Caenorhabditis japonica*	0	-	
342	*Caenorhabditis sp. 11 MAF-2010*	0	-	
343	*Priapulus caudatus*	0	-	
ROTIFERA				
344	*Adineta vaga*	23	min 111, avg 415, med 371, max 999	
PLATYHELMINTHES				
345	*Schistosoma curassoni*	170	min 52, avg 356, med 63, max 2248	
346	*Schistosoma haematobium*	151	min 50, avg 229, med 61, max 2059	
347	*Schistosoma japonicum*	83	min 58, avg 824, med 526, max 6036	
348	*Schistosoma mansoni*	96	min 54, avg 181, med 61, max 2924	
349	*Schistosoma margrebowiei*	136	min 50, avg 171, med 61, max 3942	
350	*Schistosoma mattheei*	120	min 51, avg 394, med 61, max 5065	
351	*Schistosoma rodhaini*	94	min 51, avg 316, med 62, max 4927	
352	*Clonorchis sinensis*	213	min 104, avg 216, med 208, max 2207	
353	*Echinococcus granulosus*	0	-	
354	*Echinococcus multilocularis*	0	-	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
355	*Hymenolepis microstoma*	0	-	
ANNELIDA				
356	*Capitella teleta*	10	min 89, avg 252, med 193, max 452	
357	*Helobdella robusta*	360	min 70, avg 335, med 270, max 1371	
MOLLUSCA				
358	*Crassostrea gigas*	371	min 77, avg 2068, med 1888, max 6771	
359	*Lottia gigantea*	81	min 226, avg 760, med 634, max 3303	
360	*Aplysia californica*	1609	min 78, avg 277, med 207, max 2722	
361	*Biomphalaria glabrata*	532	min 80, avg 350, med 198, max 3280	
CNIDARIA				
362	*Nematostella vectensis*	35	min 109, avg 1664, med 938, max 5309	
363	*Hydra vulgaris*	317	min 73, avg 1523, med 336, max 5900	
TENTACULATA				
364	*Mnemiopsis leidyi*	0	-	
PLACOZOA				
365	*Trichoplax adhaerens*	0	-	
PORIFERA				
366	*Amphimedon queenslandica*	0	-	
VIRIDIPLANTAE				
367	*Micromonas pusilla CCMP1545*	0	-	
368	*Micromonas sp. RCC299*	0	-	
369	*Ostreococcus lucimarinus CCE9901*	0	-	
370	*Ostreococcus tauri*	0	-	
371	*Chlamydomonas reinhardtii*	72	min 69, avg 1438, med 1100, max 6259	
372	*Volvox carteri f. nagariensis*	109	min 75, avg 1777, med 1239, max 8376	
373	*Chlorella variabilis*	9	min 72, avg 516, med 280, max 1388	
374	*Auxenochlorella protothecoides*	0	-	
375	*Helicosporidium sp. ATCC 50920*	0	-	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
376	*Coccomyxa subellipsoidea C-169*	60	min 71, avg 2429, med 404, max 8994	Very similar to L1-like Zepp element from *Chlorella vulgaris*
377	*Klebsormidium flaccidum*	0	-	
378	*Physcomitrella patens*	47	min 82, avg 464, med 277, max 875	
379	*Selaginella moellendorffii*	55	min 82, avg 4027, med 1177, max 17033	
380	*Pinus taeda*	1083	min 66, avg 2105, med 1298, max 5806	
381	*Amborella trichopoda*	2219	min 91, avg 1781, med 938, max 7437	
382	*Spirodela polyrhiza*	48	min 121, avg 2080, med 1705, max 3958	
383	*Phoenix dactylifera*	2657	min 46, avg 1445, med 1034, max 4620	
384	*Elaeis oleifera*	1125	min 80, avg 1924, med 1367, max 4405	
385	*Ensete ventricosum*	8	min 631, avg 1122, med 1141, max 1448	
386	*Musa acuminata subsp. malaccensis*	212	min 85, avg 1222, med 1246, max 2648	
387	*Sorghum bicolor*	1189	min 80, avg 2607, med 2291, max 7857	
388	*Zea mays*	1891	min 50, avg 2749, med 2605, max 7283	
389	*Setaria italica*	937	min 62, avg 2412, med 2348, max 7323	
390	*Brachypodium distachyon*	774	min 129, avg 2171, med 2033, max 7282	
391	*Leersia perrieri*	567	min 194, avg 2008, med 1900, max 6583	
392	*Oryza barthii*	705	min 46, avg 1876, med 1405, max 7342	
393	*Oryza brachyantha*	180	min 94, avg 1862, med 1422, max 5902	
394	*Oryza glumipatula*	673	min 79, avg 1720, med 1215, max 7736	
395	*Oryza longistaminata*	929	min 50, avg 1390, med 835, max 7116	
396	*Oryza meridionalis*	602	min 70, avg 1631, med 1324, max 7197	
397	*Oryza nivara*	747	min 74, avg 1760, med 1362, max 7342	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
398	*Oryza punctata*	604	min 96, avg 2028, med 1657, max 7473	
399	*Oryza sativa Japonica Group*	823	min 82, avg 2019, med 1552, max 7409	
400	*Zizania latifolia*	716	min 88, avg 1748, med 1772, max 5345	
401	*Aegilops tauschii*	8443	min 78, avg 2139, med 2087, max 7683	
402	*Triticum urartu*	8617	min 78, avg 2134, med 2050, max 8407	
403	*Nelumbo nucifera*	4701	min 109, avg 1691, med 1046, max 7648	
404	*Lupinus angustifolius*	1104	min 116, avg 2082, med 1711, max 5768	
405	*Phaseolus vulgaris*	1585	min 58, avg 1764, med 1570, max 5861	
406	*Cajanus cajan*	629	min 87, avg 1361, med 1049, max 5410	
407	*Vigna angularis var. angularis*	133	min 89, avg 404, med 353, max 2875	
408	*Vigna radiata var. radiata*	207	min 76, avg 384, med 302, max 3013	
409	*Glycine max*	1755	min 73, avg 2328, med 1810, max 6483	
410	*Glycine soja*	2146	min 83, avg 1995, med 1171, max 11768	
411	*Cicer arietinum*	459	min 81, avg 884, med 467, max 5029	
412	*Medicago truncatula*	355	min 106, avg 2092, med 1632, max 7369	
413	*Trifolium pratense*	1496	min 113, avg 778, med 438, max 5762	
414	*Lotus japonicus*	819	min 82, avg 1181, med 1076, max 3912	
415	*Malus x domestica*	520	min 82, avg 2278, med 1954, max 6604	
416	*Pyrus x bretschneideri*	676	min 94, avg 2119, med 1364, max 6700	
417	*Prunus mume*	361	min 72, avg 1554, med 920, max 5952	
418	*Prunus persica*	280	min 97, avg 1819, med 1151, max 5749	
419	*Fragaria iinumae*	506	min 76, avg 821, med 554, max 6012	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
420	*Fragaria nubicola*	744	min 81, avg 692, med 502, max 4773	
421	*Fragaria orientalis*	1046	min 72, avg 558, med 443, max 5421	
422	*Fragaria vesca subsp. vesca*	474	min 82, avg 1298, med 862, max 6199	
423	*Fragaria x ananassa*	581	min 90, avg 591, med 355, max 6013	
424	*Morus notabilis*	679	min 89, avg 485, med 266, max 3999	
425	*Cannabis sativa*	1950	min 106, avg 1321, med 1091, max 5132	
426	*Castanea mollissima*	2591	min 74, avg 1850, med 1680, max 4424	
427	*Betula nana*	2085	min 79, avg 939, med 528, max 4513	
428	*Cucumis melo*	655	min 74, avg 1364, med 1122, max 3798	
429	*Cucumis sativus*	515	min 95, avg 1006, med 853, max 3316	
430	*Citrullus lanatus*	589	min 70, avg 1040, med 790, max 4358	
431	*Lagenaria siceraria*	521	min 74, avg 472, med 351, max 2144	
432	*Populus euphratica*	448	min 64, avg 2509, med 2200, max 6319	
433	*Populus trichocarpa*	528	min 76, avg 1983, med 1518, max 6523	
434	*Jatropha curcas*	422	min 121, avg 1715, med 1260, max 5478	
435	*Manihot esculenta subsp. flabellifolia*	630	min 74, avg 1155, med 748, max 5083	
436	*Ricinus communis*	144	min 90, avg 1089, med 672, max 3638	
437	*Linum usitatissimum*	427	min 69, avg 762, med 471, max 5346	
438	*Eucalyptus camaldulensis*	3599	min 83, avg 1330, med 933, max 6573	
439	*Eucalyptus grandis*	2590	min 75, avg 1851, med 1421, max 5972	
440	*Carica papaya*	910	min 89, avg 1584, med 1290, max 6410	
441	*Arabidopsis halleri subsp. gemmifera*	2489	min 69, avg 799, med 146, max 7434	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
442	*Arabidopsis lyrata subsp. lyrata*	827	min 89, avg 2534, med 2698, max 6932	
443	*Arabidopsis thaliana*	260	min 77, avg 2468, med 2324, max 6834	
444	*Camelina sativa*	3335	min 85, avg 2315, med 2368, max 6430	
445	*Capsella rubella*	272	min 120, avg 3028, med 3152, max 6249	
446	*Brassica napus*	3927	min 75, avg 2830, med 2963, max 8479	
447	*Brassica oleracea var. oleracea*	3041	min 82, avg 1915, med 1807, max 6121	
448	*Brassica rapa*	1853	min 73, avg 1787, med 1321, max 6125	
449	*Raphanus raphanistrum subsp. raphanistrum*	1419	min 76, avg 2388, med 2265, max 5996	
450	*Raphanus sativus*	1595	min 77, avg 2945, med 3131, max 6103	
451	*Aethionema arabicum*	1063	min 70, avg 2004, med 1872, max 5667	
452	*Arabis alpina*	1528	min 102, avg 2624, med 2829, max 6298	
453	*Eutrema parvulum*	300	min 84, avg 2741, med 2690, max 6073	
454	*Eutrema salsugineum*	863	min 80, avg 3026, med 3232, max 6363	
455	*Sisymbrium irio*	991	min 79, avg 2459, med 2455, max 6116	
456	*Leavenworthia alabamica*	306	min 100, avg 2175, med 2169, max 5965	
457	*Tarenaya hassleriana*	239	min 91, avg 1731, med 1562, max 4214	
458	*Gossypium arboreum*	2831	min 70, avg 1508, med 1099, max 5838	
459	*Gossypium raimondii*	1571	min 95, avg 1545, med 1183, max 5724	
460	*Theobroma cacao*	574	min 136, avg 1601, med 952, max 6934	
461	*Aquilaria agallochum*	635	min 93, avg 1191, med 961, max 3700	
462	*Azadirachta indica*	206	min 82, avg 1261, med 911, max 5194	
463	*Citrus clementine*	767	min 77, avg 1658, med 1044, max 5614	

No	**Species**	**# unique L1 hits**	**Length distribution (bp)**	**Notes**
464	*Citrus sinensis*	880	min 74, avg 1543, med 951, max 5711	
465	*Vitis vinifera*	2505	min 80, avg 2821, med 2758, max 6488	
466	*Amaranthus hypochondriacus*	1901	min 94, avg 659, med 472, max 4869	
467	*Amaranthus tuberculatus*	102	min 167, avg 254, med 229, max 948	
468	*Beta vulgaris subsp. vulgaris*	2302	min 64, avg 2311, med 2317, max 6948	
469	*Spinacia oleracea*	1623	min 112, avg 1650, med 1250, max 6337	
470	*Dianthus caryophyllus*	1242	min 106, avg 1604, med 1378, max 4045	
471	*Actinidia chinensis*	839	min 70, avg 1488, med 1100, max 5954	
472	*Vaccinium macrocarpon* 1	1035	min 93, avg 894, med 597, max 3606	
473	*Diospyros lotus*	1	min 1950, avg 1950, med 1950, max 1950	
474	*Primula veris*	231	min 143, avg 1488, med 1318, max 3967	
475	*Solanum arcanum*	2319	min 65, avg 2193, med 1442, max 16602	
476	*Solanum habrochaites*	2635	min 64, avg 2162, med 1354, max 14525	
477	*Solanum lycopersicum*	2026	min 52, avg 2154, med 1456, max 28999	
478	*Solanum melongena*	2627	min 61, avg 1733, med 1263, max 10878	
479	*Solanum pennellii*	2383	min 60, avg 2084, med 1363, max 16789	
480	*Solanum pimpinellifolium*	3875	min 43, avg 1327, med 800, max 16052	
481	*Solanum tuberosum*	3069	min 64, avg 1743, med 1130, max 16344	
482	*Capsicum annuum*	7907	min 65, avg 1952, med 1708, max 6394	
483	*Nicotiana sylvestris*	5887	min 55, avg 1991, med 1739, max 6385	
484	*Nicotiana tomentosiformis*	4522	min 55, avg 1978, med 1586, max 6290	
485	*Fraxinus excelsior*	1257	min 77, avg 1662, med 1348, max 4845	

No	Species	# unique L1 hits	Length distribution (bp)	Notes
486	*Penstemon centranthifolius*	14	min 207, avg 945, med 799, max 2734	
487	*Penstemon grinnellii*	14	min 120, avg 748, med 614, max 1452	
488	*Sesamum indicum*	1190	min 64, avg 1457, med 1304, max 4552	
489	*Genlisea aurea*	27	min 121, avg 1325, med 1211, max 2901	
490	*Mimulus guttatus*	414	min 86, avg 2057, med 1900, max 5583	
491	*Conyza canadensis*	316	min 99, avg 1276, med 1260, max 3276	
ECHINOIDEA				
492	*Lytechinus variegatus*	1544	min 73, avg 655, med 437, max 4376	
493	*Strongylocentrotus purpuratus*	2711	min 42, avg 660, med 347, max 6094	Kordis et al (2006) showed they have more diversity than *Ciona* species
ASTEROIDEA				
494	*Patiria miniata*	68	min 72, avg 1592, med 1548, max 3988	
ENTEROPNEUSTA				
495	*Saccoglossus kowalevskii*	629	min 76, avg 1554, med 1016, max 6255	
TUNICATA				
496	*Ciona intestinalis*	9	min 104, avg 1354, med 749, max 4593	Kordis et al (2006)
497	*Ciona savignyi*	671	min 128, avg 2967, med 1814, max 11990	Kordis et al (2006) found 5-6 low copy number families
498	*Botryllus schlosseri*	14	min 109, avg 724, med 253, max 3617	
499	*Oikopleura dioica*	0	-	
LEPTOCARDII				
500	*Branchiostoma floridae*	174	min 71, avg 1628, med 1098, max 5842	40 diverse families (Kordis et al, 2006)
CEPHALASPIDOMORPHI				
501	*Lethenteron camtschaticum*	262	min 74, avg 1367, med 911, max 6536	
502	*Petromyzon marinus*	137	min 65, avg 2554, med 1617, max 7409	Kordis et al (2006) found complete absence of L1 retrotransposons here
SARCOPTERYGII				
503	*Latimeria chalumnae*	3721	min 66, avg 1680, med 827, max 12520	

Table 6: ORF content of L1s in the genome

Table B.6: **L1 open reading frame content:** For each species that exhibited L1 presence (407 species out of 503), we categorised the L1s based on whether they had both ORFs intact, only ORF1 intact, only ORF2 intact, or no ORFs intact (‘intact-ness’ is defined in Fig B.4). Both full-length elements and fragments were screened for ORFs because some species appear to have fragment L1s which still contain intact ORF1 or ORF2 regions. Both ‘confirmed ORF1’ and ‘probable ORF1’ (see explanation in main text) are included in the ORF1 calculations. Note that many of the sequences without ORFs are still full-length - however, the ORF1 and/or ORF2 regions could not be confirmed as ‘intact’, based on our criteria.

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
MAMMALIA				
Monodelphis domestica	194	4459	1607	84310
Macropus eugenii	0	360	0	119356
Sarcophilus harrisii	0	141	0	117740
Dasypus novemcinctus	42	3636	57	196698
Choloepus hoffmanni	4	7898	39	158883
Chrysochloris asiatica	0	122	0	58210
Echinops telfairi	0	94	0	22463
Orycteropus afer afer	0	437	0	123396
Elephantulus edwardii	0	337	0	64166
Trichechus manatus latirostris	0	533	0	142446
Procavia capensis	2	2926	13	153253
Loxodonta africana	228	3835	451	110471
Erinaceus europaeus	0	185	0	35532
Sorex araneus	0	466	0	61335
Condylura cristata	0	114	0	21126
Pteropus alecto	0	106	0	94911
Pteropus vampyrus	0	148	0	85182
Eidolon helvum	0	6	0	67364
Megaderma lyra	0	1	0	57558
Rhinolophus ferrumequinum	0	6	0	75259
Pteronotus parnellii	0	18	0	68612
Eptesicus fuscus	0	248	0	65985
Myotis brandtii	261	3211	201	82274
Myotis davidii	0	196	0	73910
Myotis lucifugus	75	3031	65	79104
Ceratotherium simum simum	0	828	0	111778
Equus przewalskii	27	709	82	114018
Equus caballus (Thoroughbred)	191	1588	132	84895
Equus caballus (Mongolian)	0	129	0	113417
Manis pentadactyla	14	1240	34	129448
Felis catus	1	1074	1	80981

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Panthera tigris altaica	423	943	258	122678
Canis lupus familiaris	424	2786	174	87863
Ursus maritimus	738	1622	343	112029
Ailuropoda melanoleuca	1200	2401	477	107267
Leptonychotes weddellii	0	283	0	123392
Odobenus rosmarus divergens	0	653	1	131117
Mustela putorius furo	0	706	0	114495
Camelus dromedarius	326	677	186	107671
Camelus ferus	1	494	49	106290
Vicugna pacos	0	519	0	139626
Sus scrofa (Duroc)	193	2737	209	103379
Sus scrofa (Tibetan)	196	1635	258	151885
Sus scrofa (Ellegaard Gottingen minipig)	20	829	140	142084
Balaenoptera acutorostrata scammoni	4036	2443	1338	152546
Physeter catodon	0	870	0	154039
Lipotes vexillifer	972	2663	372	157494
Tursiops truncatus	0	458	0	203768
Orcinus orca	0	1087	1	162421
Pantholops hodgsonii	909	1895	433	100036
Capra hircus	0	434	0	80098
Ovis aries (Texel)	784	3146	516	64077
Ovis aries musimon	0	417	0	94162
Bubalus bubalis	0	840	0	148033
Bison bison bison	0	1732	7	163961
Bos mutus	103	1550	207	104810
Bos indicus	167	2454	140	70711
Bos taurus	49	2741	98	77633
Ochotona princeps	0	65	0	22233
Oryctolagus cuniculus	219	2894	164	51094
Ictidomys tridecemlineatus	0	123	0	76954
Heterocephalus glaber	0	291	0	68554
Fukomys damarensis	415	635	199	59287
Cavia aperea	72	6288	78	76014
Cavia porcellus	517	6973	214	83299
Chinchilla lanigera	0	200	0	59866
Octodon degus	0	198	0	54743
Dipodomys ordii	31	1654	22	34212
Jaculus jaculus	0	125	0	32728
Nannospalax galili	0	476	0	55208
Mesocricetus auratus	0	221	0	55280
Cricetulus griseus	14	358	18	66764

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Microtus ochrogaster	0	158	0	30629
Peromyscus maniculatus bairdii	0	240	0	45868
Rattus norvegicus	841	6208	529	87938
Mus musculus	3774	8200	629	86962
Tupaia belangeri	3	1623	3	65766
Tupaia chinensis	186	1189	72	72148
Galeopterus variegatus	0	1422	11	165452
Otolemur garnettii	0	496	0	100453
Microcebus murinus	18	1564	25	104391
Tarsius syrichta	0	1620	4	166795
Callithrix jacchus	13	3435	14	142404
Saimiri boliviensis boliviensis	0	508	0	126228
Rhinopithecus roxellana	2549	2377	714	155523
Nasalis larvatus	1	806	3	99602
Chlorocebus sabaeus	12	808	1	113514
Macaca fascicularis	29	3149	107	110908
Macaca mulatta	96	1494	68	106801
Papio anubis	35	3835	105	119816
Nomascus leucogenys	76	3660	173	122321
Pongo abelii	48	4308	102	152923
Gorilla gorilla gorilla	2	2763	15	124270
Pan paniscus	0	1593	7	148241
Pan troglodytes	60	2932	87	127819
Homo sapiens	266	2918	260	115223
SAUROPSIDA				
Apalone spinifera	0	4	0	2232
Pelodiscus sinensis	0	5	0	1618
Chelonia mydas	0	31	0	2788
Chrysemys picta bellii	2	69	0	2272
Struthio camelus australis	0	0	0	47
Tinamus guttatus	0	0	0	19
Anas platyrhynchos	0	0	0	29
Lyrurus tetrix tetrix	0	0	0	13
Gallus gallus	0	0	0	15
Coturnix japonica	0	0	0	3
Meleagris gallopavo	0	0	0	7
Colinus virginianus	0	6	0	115
Acanthisitta chloris	0	0	0	33
Manacus vitellinus	0	0	0	31
Zonotrichia albicollis	0	0	0	30
Geospiza fortis	0	0	0	20
Serinus canaria	0	0	0	381

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Taeniopygia guttata	0	1	0	21
Ficedula albicollis	0	1	0	27
Pseudopodoces humilis	0	1	0	33
Corvus brachyrhynchos	0	0	0	45
Corvus cornix cornix	0	0	0	32
Ara macao	0	0	0	46
Amazona vittata	0	0	0	58
Melopsittacus undulatus	0	0	0	34
Nestor notabilis	0	0	0	47
Falco cherrug	0	0	0	49
Falco peregrinus	0	0	0	47
Cariama cristata	0	0	0	54
Merops nubicus	0	0	0	33
Picoides pubescens	0	0	0	29
Buceros rhinoceros silvestris	0	0	0	35
Apaloderma vittatum	0	0	0	28
Leptosomus discolour	0	0	0	44
Haliaeetus albicilla	0	0	0	52
Haliaeetus leucocephalus	0	0	0	56
Aquila chrysaetos Canadensis	0	0	0	53
Cathartes aura	0	0	0	66
Tyto alba	0	0	0	90
Colius striatus	0	0	0	33
Charadrius vociferus	0	0	0	51
Balearica regulorum gibbericeps	0	0	0	51
Chlamydotis macqueenii	0	0	0	46
Cuculus canorus	0	0	0	33
Fulmarus glacialis	0	0	0	58
Aptenodytes forsteri	0	0	0	67
Pygoscelis adeliae	0	0	0	59
Phalacrocorax carbo	0	0	0	45
Pelecanus crispus	0	0	0	50
Nipponia nippon	0	0	0	57
Egretta garzetta	0	0	0	50
Phaethon lepturus	0	0	0	43
Gavia stellata	0	0	0	60
Tauraco erythrolophus	0	0	0	41
Opisthocomus hoazin	0	0	0	65
Columba livia	0	0	0	46
Pterocles gutturalis	0	0	0	50
Calypte anna	0	0	0	24
Chaetura pelagica	0	0	0	24

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Caprimulgus carolinensis	0	0	0	62
Eurypyga helias	0	0	0	33
Mesitornis unicolor	0	0	0	34
Podiceps cristatus	0	0	0	43
Phoenicopterus ruber ruber	0	0	0	62
Alligator mississippiensis	0	37	0	2877
Alligator sinensis	0	35	0	3072
Crocodylus porosus	0	23	0	2637
Gavialis gangeticus	0	22	0	2819
Pogona vitticeps	0	24	0	3568
Anolis carolinensis	138	385	74	1461
Vipera berus berus	0	99	2	4070
Crotalus mitchellii pyrrhus	0	6	0	3136
Ophiophagus hannah	0	40	0	7737
Python bivittatus	0	5	0	2485
AMPHIBIA				
Nanorana parkeri	2	65	12	1807
Xenopus tropicalis	102	175	102	926
NEOPTERYGII				
Lepisosteus oculatus	0	1	2	173
Anguilla anguilla	2	2	1	744
Anguilla japonica	0	13	4	847
Danio rerio	182	173	104	1107
Astyanax mexicanus	0	11	0	350
Oryzias latipes	1	27	2	1692
Poecilia formosa	4	34	1	437
Xiphophorus maculatus	1	12	2	487
Fundulus heteroclitus	4	29	3	1032
Takifugu flavidus	0	8	0	322
Takifugu rubripes	0	10	0	253
Tetraodon nigroviridis	0	8	0	50
Cynoglossus semilaevis	0	0	0	81
Haplochromis burtoni	0	16	3	764
Pundamilia nyererei	0	25	1	776
Maylandia zebra	0	31	3	863
Neolamprologus brichardi	0	23	0	766
Oreochromis niloticus	4	57	3	1159
Sebastes nigrocinctus	0	5	0	462
Sebastes rubrivinctus	0	8	0	451
Gasterosteus aculeatus	0	22	8	149
Gadus morhua	1	20	1	891
CHONDRICHTHYES				
Callorhinchus milii	0	0	0	113
Carcharhinus brachyurus	0	0	0	2426

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
ECDYSOZOA				
Aedes aegypti	142	356	45	505
Culex quinquefasciatus	54	147	9	177
Anopheles albimanus	0	5	0	34
Anopheles arabiensis	0	5	0	55
Anopheles atroparvus	0	8	0	56
Anopheles christyi	0	3	0	41
Anopheles culicifacies	0	5	0	35
Anopheles darlingi	0	2	1	30
Anopheles dirus	0	9	0	71
Anopheles epiroticus	0	8	0	43
Anopheles farauti	0	6	0	52
Anopheles funestus	2	5	0	34
Anopheles gambiae	3	9	3	52
Anopheles maculatus	0	2	0	31
Anopheles melas	0	3	0	40
Anopheles merus	0	5	0	89
Anopheles minimus	0	3	0	35
Anopheles quadriannulatus	0	3	0	59
Anopheles sinensis	0	5	1	33
Anopheles stephensi	0	5	0	46
Ceratitis capitata	0	5	0	97
Drosophila albomicans	0	1	0	49
Drosophila ananassae	0	5	0	18
Drosophila biarmipes	0	1	0	24
Drosophila bipectinata	0	3	0	15
Drosophila elegans	0	7	0	25
Drosophila erecta	0	5	0	19
Drosophila eugracilis	0	5	0	20
Drosophila ficusphila	0	4	2	58
Drosophila grimshawi	0	6	0	64
Drosophila kikkawai	0	5	0	28
Drosophila melanogaster	0	2	0	14
Drosophila miranda	0	7	0	22
Drosophila mojavensis	0	2	0	111
Drosophila persimilis	0	5	0	31
Drosophila pseudoobscura pseudoobscura	0	6	0	28
Drosophila rhopaloa	0	4	0	14
Drosophila sechellia	0	3	0	14
Drosophila simulans	0	2	0	17
Drosophila suzukii	0	6	0	31
Drosophila takahashii	0	2	0	23
Drosophila virilis	0	12	0	55

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Drosophila willistoni	0	4	0	37
Drosophila yakuba	0	3	0	19
Musca domestica	0	8	0	75
Glossina austeni	0	6	0	137
Glossina brevipalpis	0	2	0	178
Glossina fuscipes fuscipes	0	6	0	129
Glossina morsitans morsitans	0	4	0	120
Glossina pallidipes	0	6	0	116
Orussus abietinus	0	2	0	18
Microplitis demolitor	0	4	0	65
Blattella germanica	0	6	4	945
Zootermopsis nevadensis	0	1	0	28
Daphnia pulex	0	1	0	26
Eurytemora affinis	0	2	0	275
Hyalella azteca	0	1	0	136
Latrodectus hesperus	0	2	0	76
Ixodes ricinus	2	5	2	437
Ixodes scapularis	0	5	0	1101
Rhipicephalus microplus	0	2	0	86
Metaseiulus occidentalis	0	3	0	31
Centruroides exilicauda	0	0	1	64
Limulus polyphemus	0	0	0	145
Trichinella spiralis	0	1	0	13
ROTIFERA				
Adineta vaga	0	2	0	21
PLATYHELMINTHES				
Schistosoma curassoni	0	0	0	170
Schistosoma haematobium	0	0	0	151
Schistosoma japonicum	0	1	0	82
Schistosoma mansoni	0	1	0	95
Schistosoma margrebowiei	0	0	0	136
Schistosoma mattheei	0	3	0	117
Schistosoma rodhaini	0	1	0	93
Clonorchis sinensis	0	0	0	213
ANNELIDA				
Capitella teleta	0	0	0	10
Helobdella robusta	0	5	0	355
MOLLUSCA				
Crassostrea gigas	12	7	21	331
Lottia gigantea	0	12	0	69
Aplysia californica	0	14	1	1594
Biomphalaria glabrata	0	9	0	523
CNIDARIA				

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Nematostella vectensis	1	4	0	30
Hydra vulgaris	1	6	2	308
VIRIDIPLANTAE				
Chlamydomonas reinhardtii	1	1	3	67
Volvox carteri f. nagariensis	1	4	5	99
Chlorella variabilis	0	0	0	9
Coccomyxa subellipsoidea C-169	13	0	2	45
Physcomitrella patens	0	0	0	47
Selaginella moellendorffii	7	10	3	35
Pinus taeda	3	48	2	1030
Amborella trichopoda	0	21	0	2198
Spirodela polyrhiza	0	2	0	46
Phoenix dactylifera	1	9	23	2624
Elaeis oleifera	0	0	0	1125
Ensete ventricosum	0	0	0	8
Musa acuminata subsp. malaccensis	0	4	0	208
Sorghum bicolor	29	101	68	991
Zea mays	64	147	140	1540
Setaria italica	14	24	31	868
Brachypodium distachyon	1	32	4	737
Leersia perrieri	1	19	2	545
Oryza barthii	3	29	5	668
Oryza brachyantha	1	3	1	175
Oryza glumipatula	0	33	4	636
Oryza longistaminata	12	25	23	869
Oryza meridionalis	0	29	2	571
Oryza nivara	0	30	6	711
Oryza punctata	0	23	1	580
Oryza sativa Japonica Group	8	40	7	768
Zizania latifolia	0	3	2	711
Aegilops tauschii	10	362	82	7989
Triticum urartu	20	411	120	8066
Nelumbo nucifera	0	50	1	4650
Lupinus angustifolius	0	42	3	1059
Phaseolus vulgaris	0	18	27	1540
Cajanus cajan	0	40	1	588
Vigna angularis var. angularis	0	0	0	133
Vigna radiata var. radiata	0	0	0	207
Glycine max	40	152	52	1511
Glycine soja	37	179	40	1890

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Cicer arietinum	0	0	0	459
Medicago truncatula	16	20	10	309
Trifolium pratense	7	48	11	1430
Lotus japonicus	0	12	0	807
Malus x domestica	10	25	10	475
Pyrus x bretschneideri	5	27	8	636
Prunus mume	2	17	3	339
Prunus persica	5	18	3	254
Fragaria iinumae	0	6	0	500
Fragaria nubicola	0	11	0	733
Fragaria orientalis	0	3	0	1043
Fragaria vesca subsp. vesca	1	19	2	452
Fragaria x ananassa	2	4	1	574
Morus notabilis	0	2	0	677
Cannabis sativa	12	130	33	1775
Castanea mollissima	0	44	10	2537
Betula nana	0	4	11	2070
Cucumis melo	0	4	2	649
Cucumis sativus	0	1	3	511
Citrullus lanatus	0	1	0	588
Lagenaria siceraria	0	0	0	521
Populus euphratica	51	60	21	316
Populus trichocarpa	9	53	18	448
Jatropha curcas	1	21	10	390
Manihot esculenta subsp. flabellifolia	1	10	1	618
Ricinus communis	0	3	0	141
Linum usitatissimum	0	0	1	426
Eucalyptus camaldulensis	5	114	16	3464
Eucalyptus grandis	27	217	39	2307
Carica papaya	0	3	0	907
Arabidopsis halleri subsp. gemmifera	37	107	23	2322
Arabidopsis lyrata subsp. lyrata	43	82	91	611
Arabidopsis thaliana	9	40	13	198
Camelina sativa	37	347	94	2857
Capsella rubella	29	39	29	175
Brassica napus	188	362	544	2833
Brassica oleracea var. oleracea	33	180	97	2731
Brassica rapa	90	100	213	1450
Raphanus raphanistrum subsp. raphanistrum	93	133	126	1067

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
Raphanus sativus	125	190	233	1047
Aethionema arabicum	21	89	32	921
Arabis alpina	18	154	51	1305
Eutrema parvulum	0	20	2	278
Eutrema salsugineum	19	133	35	676
Sisymbrium irio	10	74	10	897
Leavenworthia alabamica	8	27	16	255
Tarenaya hassleriana	4	21	8	206
Gossypium arboreum	0	45	1	2785
Gossypium raimondii	1	52	1	1517
Theobroma cacao	3	8	14	549
Aquilaria agallochum	0	6	2	627
Azadirachta indica	0	1	0	205
Citrus clementine	7	76	13	671
Citrus sinensis	4	80	13	783
Vitis vinifera	39	165	80	2221
Amaranthus hypochondriacus	0	46	0	1855
Amaranthus tuberculatus	0	0	0	102
Beta vulgaris subsp. vulgaris	69	172	181	1880
Spinacia oleracea	19	129	31	1444
Dianthus caryophyllus	33	92	94	1023
Actinidia chinensis	5	32	3	799
Vaccinium macrocarpon	0	4	2	1029
Diospyros lotus	0	0	0	1
Primula veris	0	5	3	223
Solanum arcanum	28	68	5	2218
Solanum habrochaites	10	62	3	2560
Solanum lycopersicum	10	63	8	1945
Solanum melongena	0	18	0	2609
Solanum pennellii	1	58	2	2322
Solanum pimpinellifolium	9	37	2	3827
Solanum tuberosum	8	153	16	2892
Capsicum annuum	0	36	1	7870
Nicotiana sylvestris	1	187	6	5693
Nicotiana tomentosiformis	0	115	1	4406
Fraxinus excelsior	3	25	14	1215
Penstemon centranthifolius	0	0	0	14
Penstemon grinnellii	0	0	0	14
Sesamum indicum	0	58	8	1124
Genlisea aurea	0	1	0	26
Mimulus guttatus	10	32	47	325
Conyza canadensis	0	4	5	307

Species	**Both ORFs**	**ORF1 only**	**ORF2 only**	**No ORFs**
ECHINOIDEA				
Lytechinus variegatus	0	6	0	1538
Strongylocentrotus purpuratus	3	16	16	2676
ASTEROIDEA				
Patiria miniata	0	0	2	66
ENTEROPNEUSTA				
Saccoglossus kowalevskii	2	42	4	581
TUNICATA				
Ciona intestinalis	0	0	0	9
Ciona savignyi	30	179	13	449
Botryllus schlosseri	0	0	0	14
LEPTOCARDII				
Branchiostoma floridae	6	7	10	151
CEPHALASPIDOMORPHI				
Lethenteron camtschaticum	2	13	2	245
Petromyzon marinus	1	15	0	121
SARCOPTERYGII				
Latimeria chalumnae	2	265	15	3439

Table 7: L1 status of each species

Table B.7: **L1 status of each species:** The union of LASTZ and TBLASTN results was used to determine the most likely status of each species: L1- (no L1s found), L1+ (L1s present), or L1* (L1s present and potentially active, based on the presence of intact ORF2 satisfying our criteria). This was done to control for differences in genome assembly quality and quantity of available nucleotide data.

		LASTZ		**TBLASTN**		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
MAMMALIA						
1	*Tachyglossus aculeatus*	N	N	N	N	L1-
2	*Ornithorhynchus anatinus*	N	N	N	N	L1-
3	*Monodelphis domestica*	Y	Y	Y	Y	L1*
4	*Macropus eugenii*	Y	N	Y	Y	L1*
5	*Sarcophilus harrisii*	Y	N	Y	N	L1+
6	*Dasypus novemcinctus*	Y	Y	Y	Y	L1*
7	*Choloepus hoffmanni*	Y	Y	Y	Y	L1*
8	*Chrysochloris asiatica*	Y	N	Y	N	L1+
9	*Echinops telfairi*	Y	N	Y	Y	L1*
10	*Orycteropus afer afer*	Y	N	Y	N	L1+
11	*Elephantulus edwardii*	Y	N	Y	N	L1+
12	*Trichechus manatus latirostris*	Y	N	Y	N	L1+
13	*Procavia capensis*	Y	Y	Y	Y	L1*
14	*Loxodonta africana*	Y	Y	Y	Y	L1*
15	*Erinaceus europaeus*	Y	N	Y	N	L1+
16	*Sorex araneus*	Y	N	Y	Y	L1*
17	*Condylura cristata*	Y	N	Y	N	L1+
18	*Pteropus alecto*	Y	N	Y	N	L1+
19	*Pteropus vampyrus*	Y	N	Y	N	L1+
20	*Eidolon helvum*	Y	N	Y	N	L1+
21	*Megaderma lyra*	Y	N	N	N	L1+
22	*Rhinolophus ferrumequinum*	Y	N	Y	Y	L1*
23	*Pteronotus parnellii*	Y	N	N	N	L1+
24	*Eptesicus fuscus*	Y	N	Y	N	L1+
25	*Myotis brandtii*	Y	Y	Y	N	L1*
26	*Myotis davidii*	Y	N	Y	N	L1+
27	*Myotis lucifugus*	Y	Y	Y	Y	L1*
28	*Ceratotherium simum simum*	Y	N	Y	N	L1+
29	*Equus przewalskii*	Y	Y	Y	N	L1*
30	*Equus caballus* (Thoroughbred)	Y	Y	Y	Y	L1*

		LASTZ		**TBLASTN**		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
31	*Equus caballus* (Mongolian)	Y	N	Y	Y	L1*
32	*Manis pentadactyla*	Y	Y	N	N	L1*
33	*Felis catus*	Y	Y	Y	Y	L1*
34	*Panthera tigris altaica*	Y	Y	Y	N	L1*
35	*Canis lupus familiaris*	Y	Y	Y	Y	L1*
36	*Ursus maritimus*	Y	Y	Y	N	L1*
37	*Ailuropoda melanoleuca*	Y	Y	Y	N	L1*
38	*Leptonychotes weddellii*	Y	N	Y	N	L1+
39	*Odobenus rosmarus divergens*	Y	Y	Y	N	L1*
40	*Mustela putorius furo*	Y	N	Y	Y	L1*
41	*Camelus dromedarius*	Y	Y	Y	N	L1*
42	*Camelus ferus*	Y	Y	Y	N	L1*
43	*Vicugna pacos*	Y	N	Y	Y	L1*
44	*Sus scrofa* (Duroc)	Y	Y	Y	Y	L1*
45	*Sus scrofa* (Tibetan)	Y	Y	Y	Y	L1*
46	*Sus scrofa* (Ellegaard Gottingen minipig)	Y	Y	Y	Y	L1*
47	*Balaenoptera acutorostrata scammoni*	Y	Y	Y	N	L1*
48	*Physeter catodon*	Y	N	Y	N	L1+
49	*Lipotes vexillifer*	Y	Y	Y	N	L1*
50	*Tursiops truncatus*	Y	N	Y	N	L1+
51	*Orcinus orca*	Y	Y	Y	N	L1*
52	*Pantholops hodgsonii*	Y	Y	Y	N	L1*
53	*Capra hircus*	Y	N	Y	N	L1+
54	*Ovis aries*	Y	Y	Y	N	L1*
55	*Ovis aries musimon*	Y	N	Y	N	L1+
56	*Bubalus bubalis*	Y	N	Y	N	L1+
57	*Bison bison bison*	Y	Y	Y	N	L1*
58	*Bos mutus*	Y	Y	Y	N	L1*
59	*Bos indicus*	Y	Y	Y	N	L1*
60	*Bos taurus*	Y	Y	Y	Y	L1*
61	*Ochotona princeps*	Y	N	Y	Y	L1*
62	*Oryctolagus cuniculus*	Y	Y	Y	Y	L1*
63	*Ictidomys tridecemlineatus*	Y	N	Y	Y	L1*
64	*Heterocephalus glaber*	Y	N	Y	N	L1+
65	*Fukomys damarensis*	Y	Y	Y	N	L1*
66	*Cavia aperea*	Y	Y	N	N	L1*

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
67	*Cavia porcellus*	Y	Y	Y	Y	L1*
68	*Chinchilla lanigera*	Y	N	Y	N	L1+
69	*Octodon degus*	Y	N	Y	N	L1+
70	*Dipodomys ordii*	Y	Y	Y	N	L1*
71	*Jaculus jaculus*	Y	N	Y	N	L1+
72	*Nannospalax galili*	Y	N	Y	N	L1+
73	*Mesocricetus auratus*	Y	N	Y	N	L1+
74	*Cricetulus griseus*	Y	Y	Y	N	L1*
75	*Microtus ochrogaster*	Y	N	Y	N	L1+
76	*Peromyscus maniculatus bairdii*	Y	N	Y	N	L1+
77	*Rattus norvegicus*	Y	Y	Y	Y	L1*
78	*Mus musculus*	Y	Y	Y	Y	L1*
79	*Tupaia belangeri*	Y	Y	Y	Y	L1*
80	*Tupaia chinensis*	Y	Y	Y	N	L1*
81	*Galeopterus variegatus*	Y	Y	Y	N	L1*
82	*Otolemur garnettii*	Y	N	Y	Y	L1*
83	*Microcebus murinus*	Y	Y	Y	Y	L1*
84	*Tarsius syrichta*	Y	Y	Y	N	L1*
85	*Callithrix jacchus*	Y	Y	Y	Y	L1*
86	*Saimiri boliviensis boliviensis*	Y	N	Y	Y	L1*
87	*Rhinopithecus roxellana*	Y	Y	Y	N	L1*
88	*Nasalis larvatus*	Y	Y	Y	N	L1*
89	*Chlorocebus sabaeus*	Y	Y	Y	N	L1*
90	*Macaca fascicularis*	Y	Y	Y	Y	L1*
91	*Macaca mulatta*	Y	Y	Y	Y	L1*
92	*Papio anubis*	Y	Y	Y	Y	L1*
93	*Nomascus leucogenys*	Y	Y	Y	Y	L1*
94	*Pongo abelii*	Y	Y	Y	Y	L1*
95	*Gorilla gorilla gorilla*	Y	Y	Y	Y	L1*
96	*Pan paniscus*	Y	Y	Y	Y	L1*
97	*Pan troglodytes*	Y	Y	Y	Y	L1*
98	*Homo sapiens*	Y	Y	Y	Y	L1*
SAUROPSIDA						
99	*Apalone spinifera*	Y	N	N	N	L1+
100	*Pelodiscus sinensis*	Y	N	Y	N	L1+
101	*Chelonia mydas*	Y	N	Y	N	L1+
102	*Chrysemys picta bellii*	Y	Y	Y	N	L1*
103	*Struthio camelus australis*	Y	N	N	N	L1+
104	*Tinamus guttatus*	Y	N	N	N	L1+

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
105	*Anas platyrhynchos*	Y	N	Y	N	L1+
106	*Lyrurus tetrix tetrix*	Y	N	N	N	L1+
107	*Gallus gallus*	Y	N	Y	N	L1+
108	*Coturnix japonica*	Y	N	N	N	L1+
109	*Meleagris gallopavo*	Y	N	N	N	L1+
110	*Colinus virginianus*	Y	N	N	N	L1+
111	*Acanthisitta chloris*	Y	N	N	N	L1+
112	*Manacus vitellinus*	Y	N	N	N	L1+
113	*Zonotrichia albicollis*	Y	N	N	N	L1+
114	*Geospiza fortis*	Y	N	N	N	L1+
115	*Serinus canaria*	Y	N	N	N	L1+
116	*Taeniopygia guttata*	Y	N	Y	N	L1+
117	*Ficedula albicollis*	Y	N	N	N	L1+
118	*Pseudopodoces humilis*	Y	N	N	N	L1+
119	*Corvus brachyrhynchos*	Y	N	N	N	L1+
120	*Corvus cornix cornix*	Y	N	N	N	L1+
121	*Ara macao*	Y	N	N	N	L1+
122	*Amazona vittata*	Y	N	N	N	L1+
123	*Melopsittacus undulatus*	Y	N	N	N	L1+
124	*Nestor notabilis*	Y	N	N	N	L1+
125	*Falco cherrug*	Y	N	N	N	L1+
126	*Falco peregrinus*	Y	N	N	N	L1+
127	*Cariama cristata*	Y	N	N	N	L1+
128	*Merops nubicus*	Y	N	N	N	L1+
129	*Picoides pubescens*	Y	N	N	N	L1+
130	*Buceros rhinoceros silvestris*	Y	N	N	N	L1+
131	*Apaloderma vittatum*	Y	N	N	N	L1+
132	*Leptosomus discolor*	Y	N	N	N	L1+
133	*Haliaeetus albicilla*	Y	N	N	N	L1+
134	*Haliaeetus leucocephalus*	Y	N	N	N	L1+
135	*Aquila chrysaetos canadensis*	Y	N	N	N	L1+
136	*Cathartes aura*	Y	N	N	N	L1+
137	*Tyto alba*	Y	N	N	N	L1+
138	*Colius striatus*	Y	N	N	N	L1+
139	*Charadrius vociferus*	Y	N	N	N	L1+
140	*Balearica regulorum gibbericeps*	Y	N	N	N	L1+
141	*Chlamydotis macqueenii*	Y	N	N	N	L1+
142	*Cuculus canorus*	Y	N	N	N	L1+

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
143	*Fulmarus glacialis*	Y	N	N	N	L1+
144	*Aptenodytes forsteri*	Y	N	N	N	L1+
145	*Pygoscelis adeliae*	Y	N	N	N	L1+
146	*Phalacrocorax carbo*	Y	N	N	N	L1+
147	*Pelecanus crispus*	Y	N	N	N	L1+
148	*Nipponia nippon*	Y	N	N	N	L1+
149	*Egretta garzetta*	Y	N	N	N	L1+
150	*Phaethon lepturus*	Y	N	N	N	L1+
151	*Gavia stellata*	Y	N	N	N	L1+
152	*Tauraco erythrolophus*	Y	N	N	N	L1+
153	*Opisthocomus hoazin*	Y	N	N	N	L1+
154	*Columba livia*	Y	N	N	N	L1+
155	*Pterocles gutturalis*	Y	N	N	N	L1+
156	*Calypte anna*	Y	N	N	N	L1+
157	*Chaetura pelagica*	Y	N	N	N	L1+
158	*Caprimulgus carolinensis*	Y	N	N	N	L1+
159	*Eurypyga helias*	Y	N	N	N	L1+
160	*Mesitornis unicolor*	Y	N	N	N	L1+
161	*Podiceps cristatus*	Y	N	N	N	L1+
162	*Phoenicopterus ruber ruber*	Y	N	N	N	L1+
163	*Alligator mississippiensis*	Y	N	Y	N	L1+
164	*Alligator sinensis*	Y	N	Y	N	L1+
165	*Crocodylus porosus*	Y	N	Y	N	L1+
166	*Gavialis gangeticus*	Y	N	N	N	L1+
167	*Pogona vitticeps*	Y	N	Y	N	L1+
168	*Anolis carolinensis*	Y	Y	Y	Y	L1*
169	*Vipera berus berus*	Y	Y	N	N	L1*
170	*Crotalus mitchellii pyrrhus*	Y	N	N	N	L1+
171	*Ophiophagus hannah*	Y	N	Y	N	L1+
172	*Python bivittatus*	Y	N	Y	N	L1+
AMPHIBIA						
173	*Nanorana parkeri*	Y	Y	N	N	L1*
174	*Xenopus tropicalis*	Y	Y	Y	Y	L1*
NEOPTERYGII						
175	*Lepisosteus oculatus*	Y	Y	Y	N	L1*
176	*Anguilla anguilla*	Y	Y	N	N	L1*
177	*Anguilla japonica*	Y	Y	N	N	L1*
178	*Danio rerio*	Y	Y	Y	Y	L1*
179	*Astyanax mexicanus*	Y	N	Y	N	L1+

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
180	*Oryzias latipes*	Y	Y	Y	Y	L1*
181	*Poecilia formosa*	Y	Y	Y	N	L1*
182	*Xiphophorus maculatus*	Y	Y	Y	N	L1*
183	*Fundulus heteroclitus*	Y	Y	Y	N	L1*
184	*Takifugu flavidus*	Y	N	N	N	L1+
185	*Takifugu rubripes*	Y	N	Y	Y	L1*
186	*Tetraodon nigroviridis*	Y	N	Y	Y	L1*
187	*Cynoglossus semilaevis*	Y	N	Y	N	L1+
188	*Haplochromis burtoni*	Y	Y	Y	Y	L1*
189	*Pundamilia nyererei*	Y	Y	Y	N	L1*
190	*Maylandia zebra*	Y	Y	Y	N	L1*
191	*Neolamprologus brichardi*	Y	N	Y	N	L1+
192	*Oreochromis niloticus*	Y	Y	Y	N	L1*
193	*Sebastes nigrocinctus*	Y	N	N	N	L1+
194	*Sebastes rubrivinctus*	Y	N	N	N	L1+
195	*Gasterosteus aculeatus*	Y	Y	Y	N	L1*
196	*Gadus morhua*	Y	Y	Y	N	L1*
CHONDRICHTHYES						
197	*Callorhinchus milii*	Y	N	Y	N	L1+
198	*Carcharhinus brachyurus*	Y	N	N	N	L1+
ECDYSOZOA						
199	*Ephemera danica*	N	N	N	N	L1-
200	*Ladona fulva*	N	N	N	N	L1-
201	*Pediculus humanus corporis*	N	N	N	N	L1-
202	*Frankliniella occidentalis*	N	N	N	N	L1-
203	*Diaphorina citri*	N	N	N	N	L1-
204	*Pachypsylla venusta*	N	N	N	N	L1-
205	*Acyrthosiphon pisum*	N	N	N	N	L1-
206	*Nilaparvata lugens*	N	N	N	N	L1-
207	*Oncopeltus fasciatus*	N	N	N	N	L1-
208	*Rhodnius prolixus*	N	N	N	N	L1-
209	*Cimex lectularius*	N	N	N	N	L1-
210	*Onthophagus taurus*	N	N	N	N	L1-
211	*Agrilus planipennis*	N	N	N	N	L1-
212	*Tribolium castaneum*	N	N	N	N	L1-
213	*Anoplophora glabripennis*	N	N	N	N	L1-
214	*Leptinotarsa decemlineata*	N	N	N	N	L1-
215	*Dendroctonus ponderosae*	N	N	N	N	L1-
216	*Mengenilla moldrzyki*	N	N	N	N	L1-
217	*Aedes aegypti*	Y	Y	Y	N	L1*

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
218	*Culex quinquefasciatus*	Y	Y	Y	N	L1*
219	*Anopheles albimanus*	Y	N	N	N	L1+
220	*Anopheles arabiensis*	Y	N	N	N	L1+
221	*Anopheles atroparvus*	Y	N	N	N	L1+
222	*Anopheles christyi*	Y	N	N	N	L1+
223	*Anopheles culicifacies*	Y	N	N	N	L1+
224	*Anopheles darlingi*	Y	Y	N	N	L1*
225	*Anopheles dirus*	Y	N	N	N	L1+
226	*Anopheles epiroticus*	Y	N	N	N	L1+
227	*Anopheles farauti*	Y	N	N	N	L1+
228	*Anopheles funestus*	Y	Y	N	N	L1*
229	*Anopheles gambiae*	Y	Y	Y	Y	L1*
230	*Anopheles maculatus*	Y	N	N	N	L1+
231	*Anopheles melas*	Y	N	N	N	L1+
232	*Anopheles merus*	Y	N	N	N	L1+
233	*Anopheles minimus*	Y	N	N	N	L1+
234	*Anopheles quadriannulatus*	Y	N	N	N	L1+
235	*Anopheles sinensis*	Y	Y	N	N	L1*
236	*Anopheles stephensi*	Y	N	N	N	L1+
237	*Mayetiola destructor*	N	N	N	N	L1-
238	*Lutzomyia longipalpis*	N	N	N	N	L1-
239	*Phlebotomus papatasi*	N	N	N	N	L1-
240	*Ceratitis capitata*	Y	N	Y	N	L1+
241	*Drosophila albomicans*	Y	N	N	N	L1+
242	*Drosophila ananassae*	Y	N	N	N	L1+
243	*Drosophila biarmipes*	Y	N	N	N	L1+
244	*Drosophila bipectinata*	Y	N	N	N	L1+
245	*Drosophila elegans*	Y	N	N	N	L1+
246	*Drosophila erecta*	Y	N	N	N	L1+
247	*Drosophila eugracilis*	Y	N	N	N	L1+
248	*Drosophila ficusphila*	Y	Y	N	N	L1*
249	*Drosophila grimshawi*	Y	N	N	N	L1+
250	*Drosophila kikkawai*	Y	N	N	N	L1+
251	*Drosophila melanogaster*	Y	N	Y	N	L1+
252	*Drosophila miranda*	Y	N	N	N	L1+
253	*Drosophila mojavensis*	Y	N	N	N	L1+
254	*Drosophila persimilis*	Y	N	N	N	L1+
255	*Drosophila pseudoobscura pseudoobscura*	Y	N	N	N	L1+

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
256	*Drosophila rhopaloa*	Y	N	N	N	L1+
257	*Drosophila sechellia*	Y	N	N	N	L1+
258	*Drosophila simulans*	Y	N	N	N	L1+
259	*Drosophila suzukii*	Y	N	N	N	L1+
260	*Drosophila takahashii*	Y	N	N	N	L1+
261	*Drosophila virilis*	Y	N	N	N	L1+
262	*Drosophila willistoni*	Y	N	N	N	L1+
263	*Drosophila yakuba*	Y	N	N	N	L1+
264	*Musca domestica*	Y	N	Y	N	L1+
265	*Glossina austeni*	Y	N	N	N	L1+
266	*Glossina brevipalpis*	Y	N	N	N	L1+
267	*Glossina fuscipes fuscipes*	Y	N	N	N	L1+
268	*Glossina morsitans morsitans*	Y	N	N	N	L1+
269	*Glossina pallidipes*	Y	N	N	N	L1+
270	*Limnephilus lunatus*	N	N	N	N	L1-
271	*Papilio glaucus*	N	N	N	N	L1-
272	*Papilio polytes*	N	N	N	N	L1-
273	*Papilio xuthus*	N	N	N	N	L1-
274	*Heliconius melpomene melpomene*	N	N	N	N	L1-
275	*Melitaea cinxia*	N	N	N	N	L1-
276	*Danaus plexippus*	N	N	N	N	L1-
277	*Bombyx mori*	N	N	N	N	L1-
278	*Manduca sexta*	N	N	N	N	L1-
279	*Plutella xylostella*	N	N	N	N	L1-
280	*Athalia rosae*	N	N	N	N	L1-
281	*Cephus cinctus*	N	N	N	N	L1-
282	*Orussus abietinus*	Y	N	Y	N	L1+
283	*Ceratosolen solmsi marchali*	N	N	N	N	L1-
284	*Nasonia giraulti*	N	N	N	N	L1-
285	*Nasonia longicornis*	N	N	N	N	L1-
286	*Nasonia vitripennis*	N	N	N	N	L1-
287	*Copidosoma floridanum*	N	N	N	N	L1-
288	*Trichogramma pretiosum*	N	N	N	N	L1-
289	*Microplitis demolitor*	Y	N	Y	N	L1+
290	*Megachile rotundata*	N	N	N	N	L1-
291	*Apis dorsata*	N	N	N	N	L1-
292	*Apis florea*	N	N	N	N	L1-
293	*Apis mellifera*	N	N	N	N	L1-

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
294	*Bombus impatiens*	N	N	N	N	L1-
295	*Bombus terrestris*	N	N	N	N	L1-
296	*Linepithema humile*	N	N	N	N	L1-
297	*Camponotus floridanus*	N	N	N	N	L1-
298	*Acromyrmex echinatior*	N	N	N	N	L1-
299	*Atta cephalotes*	N	N	N	N	L1-
300	*Solenopsis invicta*	N	N	N	N	L1-
301	*Pogonomyrmex barbatus*	N	N	N	N	L1-
302	*Harpegnathos saltator*	N	N	N	N	L1-
303	*Cerapachys biroi*	N	N	N	N	L1-
304	*Blattella germanica*	Y	Y	Y	Y	L1*
305	*Zootermopsis nevadensis*	Y	N	N	N	L1+
306	*Daphnia pulex*	Y	N	Y	N	L1+
307	*Eurytemora affinis*	Y	N	N	N	L1+
308	*Hyalella azteca*	Y	N	N	N	L1+
309	*Strigamia maritima*	N	N	N	N	L1-
310	*Stegodyphus mimosarum*	N	N	N	N	L1-
311	*Latrodectus hesperus*	Y	N	Y	N	L1+
312	*Parasteatoda tepidariorum*	N	N	N	N	L1-
313	*Tetranychus urticae*	N	N	N	N	L1-
314	*Dermatophagoides farinae*	N	N	N	N	L1-
315	*Sarcoptes scabiei type canis*	N	N	N	N	L1-
316	*Achipteria coleoptrata*	N	N	N	N	L1-
317	*Hypochthonius rufulus*	N	N	N	N	L1-
318	*Platynothrus peltifer*	N	N	N	N	L1-
319	*Steganacarus magnus*	N	N	N	N	L1-
320	*Ixodes ricinus*	Y	Y	N	N	L1*
321	*Ixodes scapularis*	Y	N	Y	Y	L1*
322	*Rhipicephalus microplus*	Y	N	Y	N	L1+
323	*Metaseiulus occidentalis*	Y	N	Y	N	L1+
324	*Varroa destructor*	N	N	N	N	L1-
325	*Centruroides exilicauda*	Y	Y	N	N	L1*
326	*Mesobuthus martensii*	N	N	N	N	L1-
327	*Limulus polyphemus*	Y	N	Y	N	L1+
328	*Trichinella spiralis*	Y	N	Y	N	L1+
329	*Ascaris suum*	N	N	N	N	L1-
330	*Elaeophora elaphi*	N	N	N	N	L1-
331	*Onchocerca volvulus*	N	N	N	N	L1-

		LASTZ		**TBLASTN**		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
332	*Steinernema monticolum*	N	N	N	N	L1-
333	*Panagrellus redivivus*	N	N	N	N	L1-
334	*Haemonchus contortus*	N	N	N	N	L1-
335	*Necator americanus*	N	N	N	N	L1-
336	*Heterorhabditis bacteriophora*	N	N	N	N	L1-
337	*Caenorhabditis angaria*	N	N	N	N	L1-
338	*Caenorhabditis brenneri*	N	N	N	N	L1-
339	*Caenorhabditis briggsae*	N	N	N	N	L1-
340	*Caenorhabditis elegans*	N	N	N	N	L1-
341	*Caenorhabditis japonica*	N	N	N	N	L1-
342	*Caenorhabditis sp. 11 MAF-2010*	N	N	N	N	L1-
343	*Priapulus caudatus*	N	N	N	N	L1-
ROTIFERA						
344	*Adineta vaga*	Y	N	N	N	L1+
PLATYHELMINTHES						
345	*Schistosoma curassoni*	Y	N	Y	N	L1+
346	*Schistosoma haematobium*	Y	N	N	N	L1+
347	*Schistosoma japonicum*	Y	N	Y	N	L1+
348	*Schistosoma mansoni*	Y	N	Y	N	L1+
349	*Schistosoma margrebowiei*	Y	N	N	N	L1+
350	*Schistosoma mattheei*	Y	N	Y	N	L1+
351	*Schistosoma rodhaini*	Y	N	N	N	L1+
352	*Clonorchis sinensis*	Y	N	N	N	L1+
353	*Echinococcus granulosus*	N	N	N	N	L1-
354	*Echinococcus multilocularis*	N	N	N	N	L1-
355	*Hymenolepis microstoma*	N	N	N	N	L1-
ANNELIDA						
356	*Capitella teleta*	Y	N	N	N	L1+
357	*Helobdella robusta*	Y	N	Y	N	L1+
MOLLUSCA						
358	*Crassostrea gigas*	Y	Y	Y	Y	L1*
359	*Lottia gigantea*	Y	N	Y	N	L1+
360	*Aplysia californica*	Y	Y	Y	N	L1*
361	*Biomphalaria glabrata*	Y	N	Y	Y	L1*
CNIDARIA						
362	*Nematostella vectensis*	Y	Y	Y	N	L1*
363	*Hydra vulgaris*	Y	Y	Y	N	L1*

		LASTZ		**TBLASTN**		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
TENTACULATA						
364	*Mnemiopsis leidyi*	N	N	N	N	L1-
PLACOZOA						
365	*Trichoplax adhaerens*	N	N	N	N	L1-
PORIFERA						
366	*Amphimedon queenslandica*	N	N	N	N	L1-
VIRIDIPLANTAE						
367	*Micromonas pusilla CCMP1545*	N	N	N	N	L1-
368	*Micromonas sp. RCC299*	N	N	N	N	L1-
369	*Ostreococcus lucimarinus CCE9901*	N	N	N	N	L1-
370	*Ostreococcus tauri*	N	N	N	N	L1-
371	*Chlamydomonas reinhardtii*	Y	Y	Y	N	L1*
372	*Volvox carteri f. nagariensis*	Y	Y	Y	N	L1*
373	*Chlorella variabilis*	Y	N	Y	N	L1+
374	*Auxenochlorella protothecoides*	N	N	N	N	L1-
375	*Helicosporidium sp. ATCC 50920*	N	N	N	N	L1-
376	*Coccomyxa subellipsoidea C-169*	Y	Y	Y	N	L1*
377	*Klebsormidium flaccidum*	N	N	N	N	L1-
378	*Physcomitrella patens*	Y	N	Y	N	L1+
379	*Selaginella moellendorffii*	Y	Y	Y	N	L1*
380	*Pinus taeda*	Y	Y	Y	Y	L1*
381	*Amborella trichopoda*	Y	N	Y	N	L1+
382	*Spirodela polyrhiza*	Y	N	Y	N	L1+
383	*Phoenix dactylifera*	Y	Y	Y	Y	L1*
384	*Elaeis oleifera*	Y	N	N	N	L1+
385	*Ensete ventricosum*	Y	N	N	N	L1+
386	*Musa acuminata subsp. malaccensis*	Y	N	Y	N	L1+
387	*Sorghum bicolor*	Y	Y	Y	Y	L1*
388	*Zea mays*	Y	Y	Y	Y	L1*
389	*Setaria italica*	Y	Y	Y	Y	L1*
390	*Brachypodium distachyon*	Y	Y	Y	Y	L1*
391	*Leersia perrieri*	Y	Y	Y	N	L1*
392	*Oryza barthii*	Y	Y	N	N	L1*

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
393	*Oryza brachyantha*	Y	Y	Y	N	L1*
394	*Oryza glumipatula*	Y	Y	Y	N	L1*
395	*Oryza longistaminata*	Y	Y	N	N	L1*
396	*Oryza meridionalis*	Y	Y	N	N	L1*
397	*Oryza nivara*	Y	Y	N	N	L1*
398	*Oryza punctata*	Y	Y	Y	N	L1*
399	*Oryza sativa Japonica Group*	Y	Y	Y	Y	L1*
400	*Zizania latifolia*	Y	Y	N	N	L1*
401	*Aegilops tauschii*	Y	Y	Y	N	L1*
402	*Triticum urartu*	Y	Y	Y	N	L1*
403	*Nelumbo nucifera*	Y	Y	Y	N	L1*
404	*Lupinus angustifolius*	Y	Y	Y	N	L1*
405	*Phaseolus vulgaris*	Y	Y	Y	Y	L1*
406	*Cajanus cajan*	Y	Y	N	N	L1*
407	*Vigna angularis var. angularis*	Y	N	Y	N	L1+
408	*Vigna radiata var. radiata*	Y	N	Y	N	L1+
409	*Glycine max*	Y	Y	Y	Y	L1*
410	*Glycine soja*	Y	Y	Y	N	L1*
411	*Cicer arietinum*	Y	N	Y	N	L1+
412	*Medicago truncatula*	Y	Y	Y	Y	L1*
413	*Trifolium pratense*	Y	Y	Y	N	L1*
414	*Lotus japonicus*	Y	N	Y	Y	L1*
415	*Malus x domestica*	Y	Y	Y	Y	L1*
416	*Pyrus x bretschneideri*	Y	Y	Y	Y	L1*
417	*Prunus mume*	Y	Y	Y	Y	L1*
418	*Prunus persica*	Y	Y	Y	Y	L1*
419	*Fragaria iinumae*	Y	N	N	N	L1+
420	*Fragaria nubicola*	Y	N	N	N	L1+
421	*Fragaria orientalis*	Y	N	N	N	L1+
422	*Fragaria vesca subsp. vesca*	Y	Y	Y	Y	L1*
423	*Fragaria x ananassa*	Y	Y	N	N	L1*
424	*Morus notabilis*	Y	N	Y	N	L1+
425	*Cannabis sativa*	Y	Y	Y	N	L1*
426	*Castanea mollissima*	Y	Y	N	N	L1*
427	*Betula nana*	Y	Y	N	N	L1*
428	*Cucumis melo*	Y	Y	Y	Y	L1*
429	*Cucumis sativus*	Y	Y	Y	N	L1*
430	*Citrullus lanatus*	Y	N	Y	N	L1+

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
431	*Lagenaria siceraria*	Y	N	N	N	L1+
432	*Populus euphratica*	Y	Y	Y	N	L1*
433	*Populus trichocarpa*	Y	Y	Y	Y	L1*
434	*Jatropha curcas*	Y	Y	Y	Y	L1*
435	*Manihot esculenta subsp. flabellifolia*	Y	Y	N	N	L1*
436	*Ricinus communis*	Y	N	Y	N	L1+
437	*Linum usitatissimum*	Y	Y	Y	N	L1*
438	*Eucalyptus camaldulensis*	Y	Y	N	N	L1*
439	*Eucalyptus grandis*	Y	Y	Y	Y	L1*
440	*Carica papaya*	Y	N	Y	N	L1+
441	*Arabidopsis halleri subsp. gemmifera*	Y	Y	N	N	L1*
442	*Arabidopsis lyrata subsp. lyrata*	Y	Y	Y	N	L1*
443	*Arabidopsis thaliana*	Y	Y	Y	Y	L1*
444	*Camelina sativa*	Y	Y	Y	Y	L1*
445	*Capsella rubella*	Y	Y	Y	Y	L1*
446	*Brassica napus*	Y	Y	Y	Y	L1*
447	*Brassica oleracea var. oleracea*	Y	Y	Y	Y	L1*
448	*Brassica rapa*	Y	Y	Y	Y	L1*
449	*Raphanus raphanistrum subsp. raphanistrum*	Y	Y	N	N	L1*
450	*Raphanus sativus*	Y	Y	Y	Y	L1*
451	*Aethionema arabicum*	Y	Y	N	N	L1*
452	*Arabis alpina*	Y	Y	Y	N	L1*
453	*Eutrema parvulum*	Y	Y	N	N	L1*
454	*Eutrema salsugineum*	Y	Y	Y	N	L1*
455	*Sisymbrium irio*	Y	Y	Y	N	L1*
456	*Leavenworthia alabamica*	Y	Y	Y	N	L1*
457	*Tarenaya hassleriana*	Y	Y	Y	Y	L1*
458	*Gossypium arboreum*	Y	Y	Y	N	L1*
459	*Gossypium raimondii*	Y	Y	Y	Y	L1*
460	*Theobroma cacao*	Y	Y	Y	Y	L1*
461	*Aquilaria agallochum*	Y	Y	N	N	L1*
462	*Azadirachta indica*	Y	N	N	N	L1+
463	*Citrus clementina*	Y	Y	Y	N	L1*
464	*Citrus sinensis*	Y	Y	Y	Y	L1*
465	*Vitis vinifera*	Y	Y	Y	Y	L1*

		LASTZ		**TBLASTN**		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
466	*Amaranthus hypochondriacus*	Y	N	Y	N	L1+
467	*Amaranthus tuberculatus*	Y	N	N	N	L1+
468	*Beta vulgaris subsp. vulgaris*	Y	Y	Y	Y	L1*
469	*Spinacia oleracea*	Y	Y	Y	N	L1*
470	*Dianthus caryophyllus*	Y	Y	Y	N	L1*
471	*Actinidia chinensis*	Y	Y	N	N	L1*
472	*Vaccinium macrocarpon*	Y	Y	Y	N	L1*
473	*Diospyros lotus*	Y	N	N	N	L1+
474	*Primula veris*	Y	Y	N	N	L1*
475	*Solanum arcanum*	Y	Y	N	N	L1*
476	*Solanum habrochaites*	Y	Y	Y	N	L1*
477	*Solanum lycopersicum*	Y	Y	Y	Y	L1*
478	*Solanum melongena*	Y	N	N	N	L1+
479	*Solanum pennellii*	Y	Y	Y	Y	L1*
480	*Solanum pimpinellifolium*	Y	Y	Y	N	L1*
481	*Solanum tuberosum*	Y	Y	Y	Y	L1*
482	*Capsicum annuum*	Y	Y	Y	N	L1*
483	*Nicotiana sylvestris*	Y	Y	Y	N	L1*
484	*Nicotiana tomentosiformis*	Y	Y	Y	N	L1*
485	*Fraxinus excelsior*	Y	Y	N	N	L1*
486	*Penstemon centranthifolius*	Y	N	N	N	L1+
487	*Penstemon grinnellii*	Y	N	N	N	L1+
488	*Sesamum indicum*	Y	Y	Y	Y	L1*
489	*Genlisea aurea*	Y	N	N	N	L1+
490	*Mimulus guttatus*	Y	Y	Y	Y	L1*
491	*Conyza canadensis*	Y	Y	N	N	L1*
ECHINOIDEA						
492	*Lytechinus variegatus*	Y	N	Y	Y	L1*
493	*Strongylocentrotus purpuratus*	Y	Y	Y	Y	L1*
ASTEROIDEA						
494	*Patiria miniata*	Y	Y	N	N	L1*
ENTEROPNEUSTA						
495	*Saccoglossus kowalevskii*	Y	Y	Y	Y	L1*
TUNICATA						
496	*Ciona intestinalis*	Y	N	Y	Y	L1*
497	*Ciona savignyi*	Y	Y	Y	Y	L1*
498	*Botryllus schlosseri*	Y	N	N	N	L1+

		LASTZ		TBLASTN		
No	**Species**	**Verified L1s in the genome**	**Intact ORF2**	**Verified L1s for the taxon**	**Intact ORF2**	**L1 status (L1-, L1+ or L1*)**
499	*Oikopleura dioica*	N	N	N	N	L1-
LEPTOCARDII						
500	*Branchiostoma floridae*	Y	Y	Y	Y	L1*
CEPHALASPIDOMORPHI						
501	*Lethenteron camtschaticum*	Y	Y	Y	N	L1*
502	*Petromyzon marinus*	Y	Y	N	N	L1*
SARCOPTERYGII						
503	*Latimeria chalumnae*	Y	Y	Y	N	L1*

Figure 5: L1 activity superimposed on the inferred tree of life

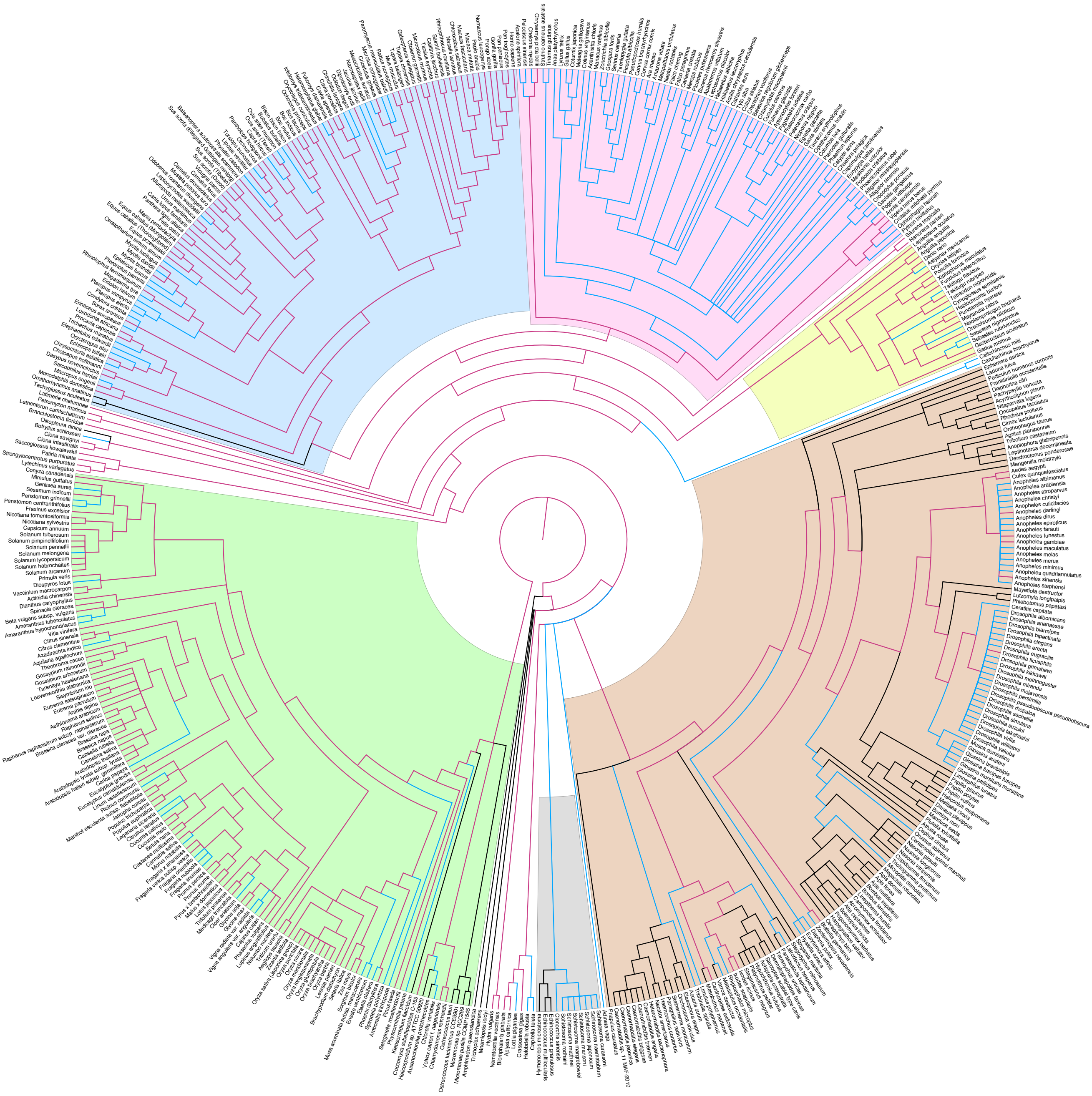


Figure B.5: **Active versus extinct L1s:** The inferred eukaryotic tree of life with branches coloured to indicate the presence of active or ORF2-intact L1s (magenta) versus extinct L1s (blue). Species with complete absence of L1 elements are unchanged (i.e. black), as in Fig. B.1. Since we can only observe the L1 status (active/extinct) at the tree tips, interior branches were coloured based on the most parsimonious explanation being that a loss-of-function is more likely than a gain.

Table 8: Active proportion of full-length L1 elements

Table B.8: **Percentage of L1s active in the genome:** 206 species contain potentially active ORF2-intact L1s (67 mammals, 47 non-mammalian animals, 92 plants). For each species, we calculated the proportion of full-length (or near full-length) L1s that are active. The L1s had to be long enough to contain both ORFs. A range of different cut-off lengths were tested (e.g. 3.8 - 4.5kb at 100bp intervals), and the active/total percentage (column 4) was calculated each time, as well as the average and standard deviation. This, along with the L1 length distributions, helped determine the best cut-off length for each species (e.g. 4.5kb for most mammals, reduced to 3.8 - 4kb for some non-mammals and plants). The percentages did not change significantly regardless of the cut-off used. These sequences are labelled 'near full-length' because they may not have complete 5' and 3' ends. Species where the TBLASTN analysis showed active L1s but the LASTZ analysis did not (e.g. *Macropus eugenii*) are marked as 'TBLASTN only' and do not have a percentage (since 0 intact ORF2 were found in the genome).

Species	**# L1s (near full-length)**	**# active L1s**	**# active L1s/ # total L1s (%)**
MAMMALIAN SPECIES			
Monodelphis domestica	15714	1540	9.80018
Macropus eugenii	572	TBLASTN only	-
Dasypus novemcinctus	13067	69	0.528048
Choloepus hoffmanni	9584	13	0.135643
Echinops telfairi	1325	TBLASTN only	-
Procavia capensis	3762	8	0.212653
Loxodonta africana	22244	538	2.41863
Sorex araneus	1722	TBLASTN only	-
Rhinolophus ferrumequinum	1073	TBLASTN only	-
Myotis brandtii	1659	362	21.8204
Myotis lucifugus	3185	120	3.76766
Equus przewalskii	7812	60	0.768049
Equus caballus (Thoroughbred)	10921	288	2.63712
Equus caballus (Mongolian)	7078	TBLASTN only	-
Manis pentadactyla	5837	37	0.633887
Felis catus	6162	2	0.032457
Panthera tigris altaica	5737	582	10.1447
Canis lupus familiaris	7772	518	6.66495
Ursus maritimus	5253	869	16.5429
Ailuropoda melanoleuca	4411	1307	29.6305
Odobenus rosmarus divergens	4752	1	0.0210438
Mustela putorius furo	5134	TBLASTN only	-
Camelus dromedarius	2273	411	18.0818
Camelus ferus	2187	27	1.23457
Vicugna pacos	1696	TBLASTN only	-
Sus scrofa (Duroc)	5596	251	4.48535
Sus scrofa (Tibetan)	4773	322	6.74628

Species	**# L1s (near full-length)**	**# active L1s**	**# active L1s/** **# total L1s (%)**
Sus scrofa (Ellegaard Gottingen minipig)	4779	110	2.30174
Balaenoptera acutorostrata scammoni	8012	5006	62.4813
Lipotes vexillifer	23066	1191	5.16344
Orcinus orca	9293	1	0.01076
Pantholops hodgsonii	2493	1109	44.4846
Ovis aries (Texel)	3117	1042	33.4296
Bison bison bison	2149	3	0.1396
Bos mutus	2145	174	8.11189
Bos indicus	3480	259	7.44253
Bos taurus	3649	90	2.46643
Ochotona princeps	352	TBLASTN only	-
Oryctolagus cuniculus	6001	280	4.66589
Ictidomys tridecemlineatus	3155	TBLASTN only	-
Fukomys damarensis	8423	461	5.47311
Cavia aperea	2816	42	1.49148
Cavia porcellus	14285	651	4.55723
Dipodomys ordii	1162	41	3.5284
Cricetulus griseus	3416	21	0.614754
Rattus norvegicus	10678	1092	10.2266
Mus musculus	18280	4143	22.6641
Tupaia belangeri	802	4	0.498753
Tupaia chinensis	3262	232	7.1122
Galeopterus variegatus	7565	2	0.0264375
Otolemur garnettii	4030	TBLASTN only	-
Microcebus murinus	1100	33	3
Tarsius syrichta	16086	2	0.0124332
Callithrix jacchus	9148	18	0.196764
Saimiri boliviensis boliviensis	6163	TBLASTN only	-
Rhinopithecus roxellana	11115	2954	26.5767
Nasalis larvatus	4217	1	0.0237135
Chlorocebus sabaeus	7864	13	0.16531
Macaca fascicularis	9655	89	0.921802
Macaca mulatta	3947	128	3.24297
Papio anubis	8205	82	0.999391
Nomascus leucogenys	9990	178	1.78178
Pongo abelii	9466	103	1.0881
Gorilla gorilla gorilla	6802	8	0.117612
Pan paniscus	11833	4	0.0338038
Pan troglodytes	10521	115	1.09305
Homo sapiens	13224	480	3.62976

Species	**# L1s (near full-length)**	**# active L1s**	**# active L1s/ # total L1s (%)**
NON-MAMMALIAN ANIMAL SPECIES			
Chrysemys picta bellii	141	1	0.70922
Anolis carolinensis	694	187	26.9452
Vipera berus berus	146	2	1.36986
Nanorana parkeri	487	11	2.25873
Xenopus tropicalis	580	198	34.1379
Lepisosteus oculatus	16	2	12.5
Anguilla anguilla	68	3	4.41176
Anguilla japonica	100	4	4
Danio rerio	590	268	45.4237
Oryzias latipes	45	2	4.44444
Poecilia formosa	41	5	12.1951
Xiphophorus maculatus	18	3	16.6667
Fundulus heteroclitus	47	7	14.8936
Takifugu rubripes	33	TBLASTN only	-
Tetraodon nigroviridis	15	TBLASTN only	-
Haplochromis burtoni	48	3	6.25
Pundamilia nyererei	42	1	2.38095
Maylandia zebra	45	3	6.66667
Oreochromis niloticus	98	6	6.12245
Gasterosteus aculeatus	24	8	33.3333
Gadus morhua	30	2	6.66667
Aedes aegypti	519	184	35.4528
Culex quinquefasciatus	128	61	47.6562
Anopheles darlingi	1	1	100
Anopheles funestus	3	2	66.6667
Anopheles gambiae	10	6	60
Anopheles sinensis	7	1	14.2857
Drosophila ficusphila	8	2	25
Blattella germanica	52	4	7.69231
Ixodes ricinus	27	4	14.8148
Ixodes scapularis	88	TBLASTN only	-
Centruroides exilicauda	5	1	20
Crassostrea gigas	97	30	30.9278
Aplysia californica	10	1	10
Biomphalaria glabrata	4	TBLASTN only	-
Nematostella vectensis	8	1	12.5
Hydra vulgaris	83	3	3.61446
Lytechinus variegatus	25	TBLASTN only	-
Strongylocentrotus purpuratus	149	19	12.7517
Patiria miniata	8	2	25
Saccoglossus kowalevskii	122	6	4.91803

Species	**# L1s (near full-length)**	**# active L1s**	**# active L1s/ # total L1s (%)**
Ciona intestinalis	2	TBLASTN only	-
Ciona savignyi	311	42	13.5048
Branchiostoma floridae	48	16	33.3333
Lethenteron camtschaticum	38	4	10.5263
Petromyzon marinus	46	1	2.17391
Latimeria chalumnae	921	17	1.84582
PLANT SPECIES			
Chlamydomonas reinhardtii	9	4	44.4444
Volvox carteri f. nagariensis	11	6	54.5455
Coccomyxa subellipsoidea C-169	15	15	100
Selaginella moellendorffii	25	10	40
Pinus taeda	324	5	1.54321
Phoenix dactylifera	307	22	7.16612
Sorghum bicolor	358	61	17.0391
Zea mays	744	165	22.1774
Setaria italica	289	34	11.7647
Brachypodium distachyon	165	3	1.81818
Leersia perrieri	106	2	1.88679
Oryza barthii	130	5	3.84615
Oryza brachyantha	39	2	5.12821
Oryza glumipatula	95	2	2.10526
Oryza longistaminata	118	26	22.0339
Oryza meridionalis	77	2	2.5974
Oryza nivara	128	4	3.125
Oryza punctata	93	1	1.07527
Oryza sativa Japonica Group	176	14	7.95455
Zizania latifolia	102	2	1.96078
Aegilops tauschii	1579	38	2.40659
Triticum urartu	2040	86	4.21569
Nelumbo nucifera	905	1	0.11050
Lupinus angustifolius	271	2	0.738007
Phaseolus vulgaris	295	17	5.76271
Cajanus cajan	52	1	1.92308
Glycine max	523	81	15.4876
Glycine soja	494	65	13.1579
Medicago truncatula	95	24	25.2632
Trifolium pratense	49	12	24.4898
Lotus japonicus	10	TBLASTN	-
Malus x domestica	131	20	15.2672
Pyrus x bretschneideri	176	13	7.38636
Prunus mume	49	5	10.2041

Species	**# L1s (near full-length)**	**# active L1s**	**# active L1s/ # total L1s (%)**
Prunus persica	47	8	17.0213
Fragaria vesca subsp. vesca	46	3	6.52174
Fragaria x ananassa	13	2	15.3846
Cannabis sativa	93	20	21.5054
Castanea mollissima	381	3	0.787402
Betula nana	96	2	2.08333
Cucumis melo	66	2	3.03030
Cucumis sativus	11	3	27.27273
Populus euphratica	139	52	37.4101
Populus trichocarpa	104	17	16.3462
Jatropha curcas	73	5	6.84932
Manihot esculenta subsp. flabellifolia	35	2	5.71429
Linum usitatissimum	58	1	1.72414
Eucalyptus camaldulensis	256	11	4.29688
Eucalyptus grandis	512	49	9.57031
Arabidopsis halleri subsp. gemmifera	267	50	18.7266
Arabidopsis lyrata subsp. lyrata	366	103	28.1421
Arabidopsis thaliana	99	20	20.202
Camelina sativa	1175	106	9.02128
Capsella rubella	147	53	36.0544
Brassica napus	1929	565	29.2898
Brassica oleracea var. oleracea	831	84	10.1083
Brassica rapa	543	228	41.989
Raphanus raphanistrum subsp. raphanistrum	535	182	34.0187
Raphanus sativus	866	287	33.1409
Aethionema arabicum	361	51	14.1274
Arabis alpina	689	53	7.69231
Eutrema parvulum	129	2	1.55039
Eutrema salsugineum	480	52	10.8333
Sisymbrium irio	357	16	4.48179
Leavenworthia alabamica	103	18	17.4757
Tarenaya hassleriana	45	8	17.7778
Gossypium arboreum	341	1	0.293255
Gossypium raimondii	152	2	1.31579
Theobroma cacao	115	16	13.913
Aquilaria agallochum	134	2	1.49254
Citrus clementina	144	16	11.1111
Citrus sinensis	131	14	10.687

Species	**# L1s (near full-length)**	**# active L1s**	**# active L1s/ # total L1s (%)**
Vitis vinifera	1134	111	9.78836
Beta vulgaris subsp. vulgaris	671	187	27.8689
Spinacia oleracea	227	33	14.5374
Dianthus caryophyllus	63	18	28.5714
Actinidia chinensis	94	6	6.38298
Vaccinium macrocarpon	13	1	7.69231
Primula veris	21	1	4.7619
Solanum arcanum	420	33	7.85714
Solanum habrochaites	422	12	2.8436
Solanum lycopersicum	467	18	3.85439
Solanum pennellii	443	3	0.677201
Solanum pimpinellifolium	364	11	3.02198
Solanum tuberosum	470	20	4.25532
Capsicum annuum	1006	1	0.0994036
Nicotiana sylvestris	1124	6	0.533808
Nicotiana tomentosiformis	1055	1	0.0947867
Fraxinus excelsior	157	12	7.64331
Sesamum indicum	60	1	1.66667
Mimulus guttatus	114	38	33.3333
Conyza canadensis	30	4	13.3333

Table 9: Master versus multiple lineage models

Table B.9: **Mammalian L1 lineages:** Shows the predicted lineage model for L1* mammalian species, based on the clustering and dendrogram results. A master lineage is characterised by a single dominant cluster which contains the majority of the active L1s in the genome with high pairwise identity to each other. The L1s from each mammalian species were initially clustered at 70% identity - the following results show the number of active L1s in the dominant cluster(s) at this percent identity. 31 mammalian species seem to adhere to the master lineage model. However, there are 12 species which have several distinct active clusters - indicative of a 'multiple lineage' model. Finally, some species do not have enough active L1s to discern the model type (listed below as 'low active copy number species').

Species	**Dominant active cluster(s)**	**Active proportion (# active L1s/ # L1s in cluster)**
MASTER lineage species		
Monodelphis domestica	cluster_1	1292/3752
Dasypus novemcinctus	cluster_0	68/5203
Loxodonta africana	cluster_9	439/3670
Canis lupus	cluster_4	465/2684
Ursus maritimus	cluster_2	847/1532
Ailuropoda melanoleuca	cluster_0	1166/1224
Camelus dromedarius	cluster_4	377/530
Sus scrofa (Duroc)	cluster_13	250/1458
Sus scrofa (Tibetan)	cluster_14	157/341
Sus scrofa (minipig)	cluster_327	92/134
Balaenoptera acutorostrata	cluster_8	1932/2171
Lipotes vexillifer	cluster_0	581/10452
Pantholops hodgsonii	cluster_6	1094/1438
Ovis aries (Texel)	cluster_1	990/1629
Bos mutus	cluster_441	58/76
Bos indicus	cluster_0	234/1960
Bos taurus	cluster_2	83/1953
Fukomys damarensis	cluster_0	447/1366
Rattus norvegicus	cluster_0	664/4209
Mus musculus	cluster_0	2388/4712
Microcebus murinus	cluster_1	33/775
Callithrix jacchus	cluster_0	14/1853
Rhinopithecus roxellana	cluster_4	1570/2602
Chlorocebus sabaeus	cluster_0	12/1878
Macaca fascicularis	cluster_0	72/3009
Macaca mulatta	cluster_91	94/276
Papio anubis	cluster_5	61/1835
Nomascus leucogenys	cluster_2	117/3187
Pongo abelii	cluster_3	59/514
Pan troglodytes	cluster_7	75/3982
Homo sapiens	cluster_0	409/2725

Species	**Dominant active cluster(s)**	**Active proportion (# active L1s/ # L1s in cluster)**
Possible MULTIPLE lineage species		
Myotis brandtii	cluster_1 cluster_140 cluster_39	11/128 250/339 97/196
Myotis lucifugus	cluster_197 cluster_2 cluster_328 cluster_7	9/83 57/836 37/218 15/250
Equus przewalskiii	cluster_13 cluster_189 cluster_290	57/289 1/8 2/8
Equus caballus (Thoroughbred)	cluster_1 cluster_142 cluster_55	41/420 200/788 46/1436
Manis pentadactyla	cluster_0 cluster_49	27/151 10/655
Panthera tigris	cluster_1 cluster_1745 cluster_628 cluster_78	221/460 131/236 24/29 195/275
Oryctolagus.cuniculus	cluster_0 cluster_1 cluster_424	98/1685 157/1807 14/168
Cavia aperea	cluster_0 cluster_8	16/421 26/1609
Cavia porcellus	cluster_0 cluster_1 cluster_18 cluster_2 cluster_3	214/1396 179/2194 31/597 222/2350 3/3204
Dipodomys ordii	cluster_0 cluster_1 cluster_2	27/536 11/229 3/282
Cricetulus griseus	cluster_0 cluster_74	20/331 1/22
Tupaia chinensis	cluster_129 cluster_1391 cluster_9	149/196 40/40 38/84
LOW ACTIVE COPY NUMBER species		
Choloepus hoffmannis	-	-
Procavia capensis	-	-
Felis catus	-	-
Odobenus rosmarus	-	-

Species	Dominant active cluster(s)	Active proportion (# active L1s/ # L1s in cluster)
Camelus ferus	-	-
Orcinus orca	-	-
Bison bison	-	-
Tupaia belangeri	-	-
Galeopterus variegatus	-	-
Tarsius syrichta	-	-
Nasalis larvatus	-	-
Gorilla gorilla	-	-
Pan paniscus	-	-

Table 10: Plant-specific reverse transcriptase domains

Table B.10: **Reverse transcriptase domains found in plants:** L1 ORF2p are known to encode an apurinic endonuclease and reverse transcriptase. RVT_1 (pfam.xfam.org) is the typical ORF2p reverse transcriptase, and it was found in every single ORF2-intact species. However, most plant species had an additional reverse transcriptase RVT_3, often accompanied by a ribonuclease H domain. Sometimes RVT_3 was included in the ORF2 after RVT_1; other times it was found in a separate, third ORF. The following table shows the variety of reverse transcriptase-like domains found in plant L1s (not restricted to just the ORF2 region).

Plant species	Reverse transcriptase	Ribonuclease
Chlamydomonas reinhardtii	RVT_1	-
Volvox carteri f. nagariensis	RVT_1	-
Chlorella variabilis	RVT_1	-
Coccomyxa subellipsoidea C-169	RVT_1	-
Physcomitrella patens	RVT_1	-
Selaginella moellendorffii	RVT_1	RNH
Pinus taeda	RVT_1, RVT_3	RNH
Amborella trichopoda	RVT_1, RVT_3	RNH
Spirodela polyrhiza	RVT_1, RVT_3	RNH
Phoenix dactylifera	RVT_1, RVT_3	RNH
Elaeis oleifera	RVT_1, RVT_3	-
Ensete ventricosum	-	-
Musa acuminata subsp. malaccensis	RVT_1	-
Sorghum bicolor	RVT_1, RVT_3	RNH
Zea mays	RVT_1, RVT_3	RNH
Setaria italica	RVT_1, RVT_3	RNH
Brachypodium distachyon	RVT_1, RVT_3	RNH
Leersia perrieri	RVT_1, RVT_3	RNH
Oryza barthii	RVT_1, RVT_3	RNH
Oryza brachyantha	RVT_1, RVT_3	RNH
Oryza glumipatula	RVT_1, RVT_3	RNH

Plant species	**Reverse transcriptase**	**Ribonuclease**
Oryza longistaminata	RVT_1, RVT_3	RNH
Oryza meridionalis	RVT_1, RVT_3	RNH
Oryza nivara	RVT_1, RVT_3	RNH
Oryza punctata	RVT_1, RVT_3	RNH
Oryza sativa Japonica Group	RVT_1, RVT_3	-
Zizania latifolia	RVT_1	-
Aegilops tauschii	RVT_1, RVT_3	RNH
Triticum urartu	RVT_1, RVT_3	RNH
Nelumbo nucifera	RVT_1, RVT_3	RNH
Lupinus angustifolius	RVT_1, RVT_3	RNH
Phaseolus vulgaris	RVT_1, RVT_3	RNH
Cajanus cajan	RVT_1, RVT_3	RNH
Vigna angularis var. angularis	-	-
Vigna radiata var. radiata	-	-
Glycine max	RVT_1, RVT_3	RNH
Glycine soja	RVT_1, RVT_3	RNH
Cicer arietinum	-	-
Medicago truncatula	RVT_1, RVT_3	RNH
Trifolium pratense	RVT_1, RVT_3	RNH
Lotus japonicus	RVT_1, RVT_3	RNH
Malus x domestica	RVT_1, RVT_3	RNH
Pyrus x bretschneideri	RVT_1, RVT_3	RNH
Prunus mume	RVT_1, RVT_3	RNH
Prunus persica	RVT_1, RVT_3	RNH
Fragaria iinumae	RVT_1, RVT_3	RNH
Fragaria nubicola	RVT_1, RVT_3	RNH
Fragaria orientalis	RVT_1, RVT_3	RNH
Fragaria vesca subsp. vesca	RVT_1, RVT_3	RNH
Fragaria x ananassa	RVT_1, RVT_3	RNH
Morus notabilis	RVT_1, RVT_3	-
Cannabis sativa	RVT_1, RVT_3	RNH
Castanea mollissima	RVT_1, RVT_3	RNH
Betula nana	RVT_1, RVT_3	-
Cucumis melo	RVT_1	-
Cucumis sativus	RVT_1	-
Citrullus lanatus	RVT_1	-
Lagenaria siceraria	RVT_1	-
Populus euphratica	RVT_1, RVT_3	RNH
Populus trichocarpa	RVT_1, RVT_3	RNH
Jatropha curcas	RVT_1, RVT_3	RNH
Manihot esculenta subsp. flabellifolia	RVT_1, RVT_3	RNH
Ricinus communis	RVT_1	-
Linum usitatissimum	RVT_1	-

Plant species	**Reverse transcriptase**	**Ribonuclease**
Eucalyptus camaldulensis	RVT_1, RVT_3	RNH
Eucalyptus grandis	RVT_1, RVT_3	RNH
Carica papaya	RVT_1	-
Arabidopsis halleri subsp. gemmifera	RVT_1, RVT_3	RNH
Arabidopsis lyrata subsp. lyrata	RVT_1, RVT_3	RNH
Arabidopsis thaliana	RVT_1, RVT_3	RNH
Camelina sativa	RVT_1, RVT_3	RNH
Capsella rubella	RVT_1, RVT_3	RNH
Brassica napus	RVT_1, RVT_3	RNH
Brassica oleracea var. oleracea	RVT_1, RVT_3	RNH
Brassica rapa	RVT_1, RVT_3	RNH
Raphanus raphanistrum subsp. raphanistrum	RVT_1, RVT_3	RNH
Raphanus sativus	RVT_1, RVT_3	RNH
Aethionema arabicum	RVT_1, RVT_3	RNH
Arabis alpina	RVT_1, RVT_3	RNH
Eutrema parvulum	RVT_1, RVT_3	RNH
Eutrema salsugineum	RVT_1, RVT_3	RNH
Sisymbrium irio	RVT_1, RVT_3	RNH
Leavenworthia alabamica	RVT_1, RVT_3	RNH
Tarenaya hassleriana	RVT_1, RVT_3	-
Gossypium arboreum	RVT_1, RVT_3	-
Gossypium raimondii	RVT_1, RVT_3	-
Theobroma cacao	RVT_1, RVT_3	RNH
Aquilaria agallochum	RVT_1, RVT_3	RNH
Azadirachta indica	RVT_1, RVT_3	-
Citrus clementina	RVT_1, RVT_3	RNH
Citrus sinensis	RVT_1, RVT_3	RNH
Vitis vinifera	RVT_1, RVT_3	-
Amaranthus hypochondriacus	RVT_1, RVT_3	-
Amaranthus tuberculatus	-	-
Beta vulgaris subsp. vulgaris	RVT_1, RVT_3	RNH
Spinacia oleracea	RVT_1, RVT_3	RNH
Dianthus caryophyllus	RVT_1, RVT_3	RNH
Actinidia chinensis	RVT_1, RVT_3	-
Vaccinium macrocarpon	RVT_1, RVT_3	-
Diospyros lotus	-	-
Primula veris	RVT_1, RVT_3	-
Solanum arcanum	RVT_1, RVT_3	RNH
Solanum habrochaites	RVT_1, RVT_3	RNH
Solanum lycopersicum	RVT_1, RVT_3	RNH
Solanum melongena	RVT_1, RVT_3	RNH
Solanum pennellii	RVT_1, RVT_3	RNH

Plant species	**Reverse transcriptase**	**Ribonuclease**
Solanum pimpinellifolium	RVT_1, RVT_3	RNH
Solanum tuberosum	RVT_1, RVT_3	RNH
Capsicum annuum	RVT_1, RVT_3	RNH
Nicotiana sylvestris	RVT_1, RVT_3	RNH
Nicotiana tomentosiformis	RVT_1, RVT_3	RNH
Fraxinus excelsior	RVT_1, RVT_3	-
Penstemon centranthifolius	-	-
Penstemon grinnellii	-	-
Sesamum indicum	RVT_1, RVT_3	RNH
Genlisea aurea	RVT_1	-
Mimulus guttatus	RVT_1, RVT_3	RNH
Conyza canadensis	RVT_1, RVT_3	RNH

Figure 6: ORF2p associated domain network diagrams

The following figures show network graphs of known ORF2p domains (e.g. reverse transcriptase RVT_1) and their strongly associated domains, for each designated order of taxa: Mammalia, Sauropsida, Amphibia, Neopterygii, Ecdysozoa, Other (e.g. ’primitive’ organisms) and Viridiplantae. For every ORF2p in every L1 (in each group of species), the HMM top hit was ranked first (this was always RVT_1) and other domains next to the top hit were ranked afterwards by decreasing score. This was used to generate a .csv file and visualise the corresponding network in Gephi. Nodes are the domain hits, and edges are weighted according to the strength of the association (i.e. how frequently they appear in that group of species). Note that edges have been rescaled to allow easy visualisation.

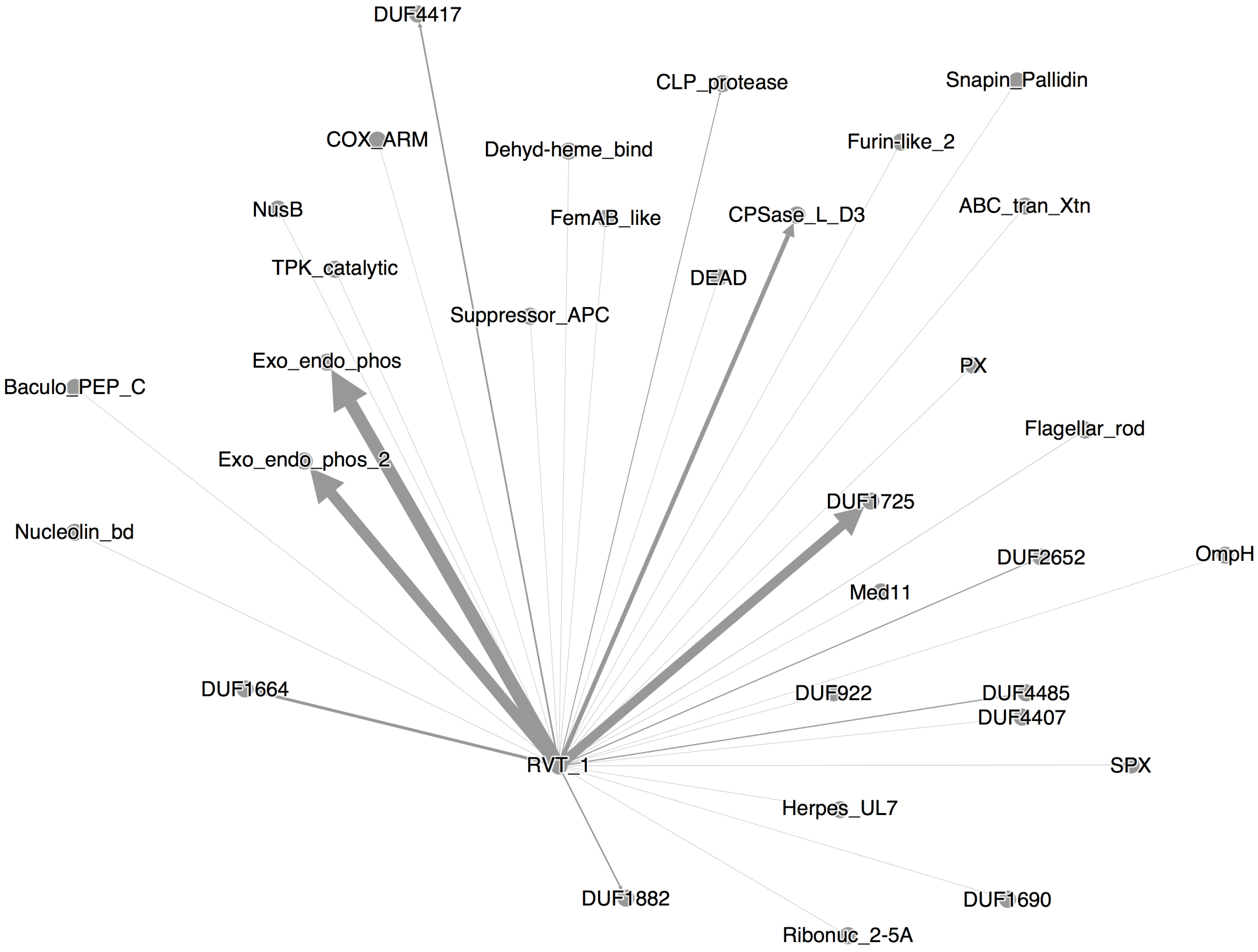


Figure B.6a: **Mammalia ORF2p domains**

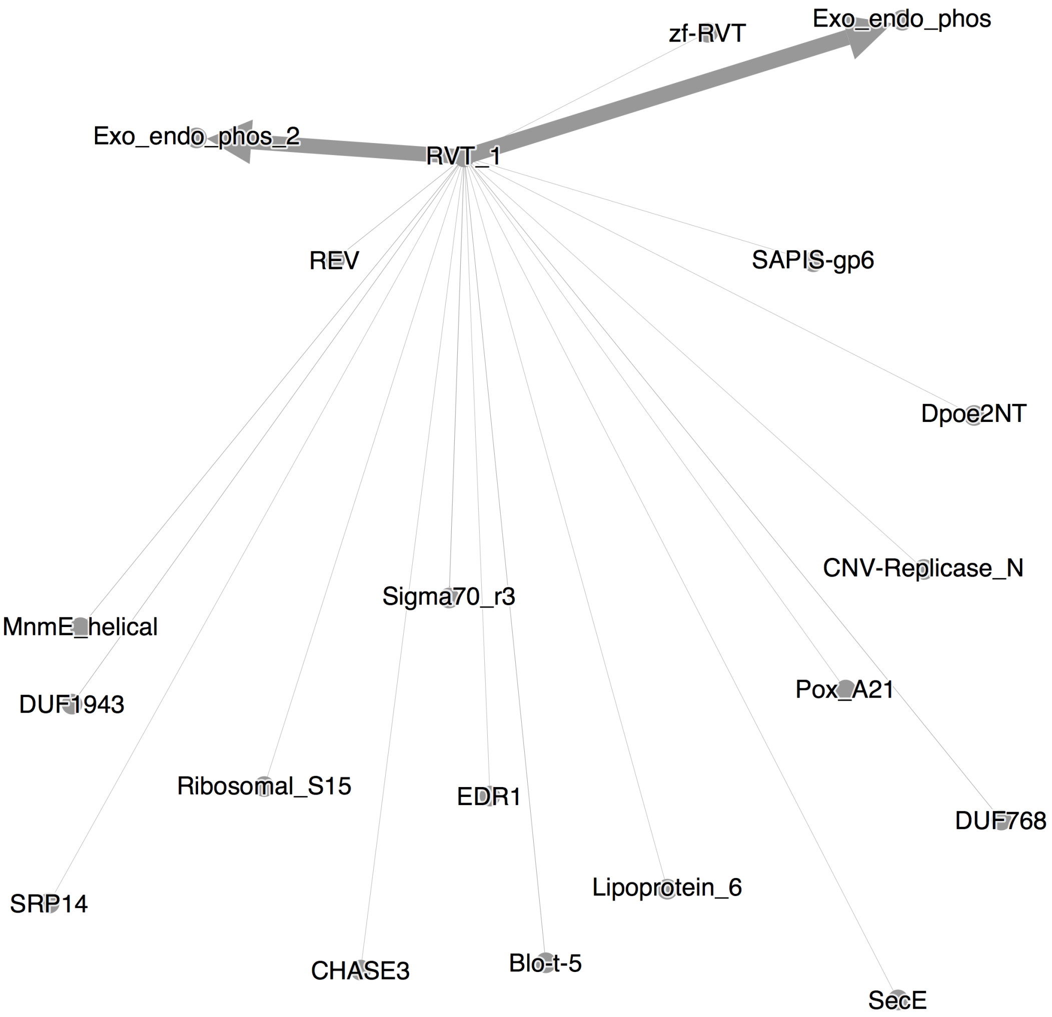


Figure B.6b: **Sauropsida ORF2p domains**

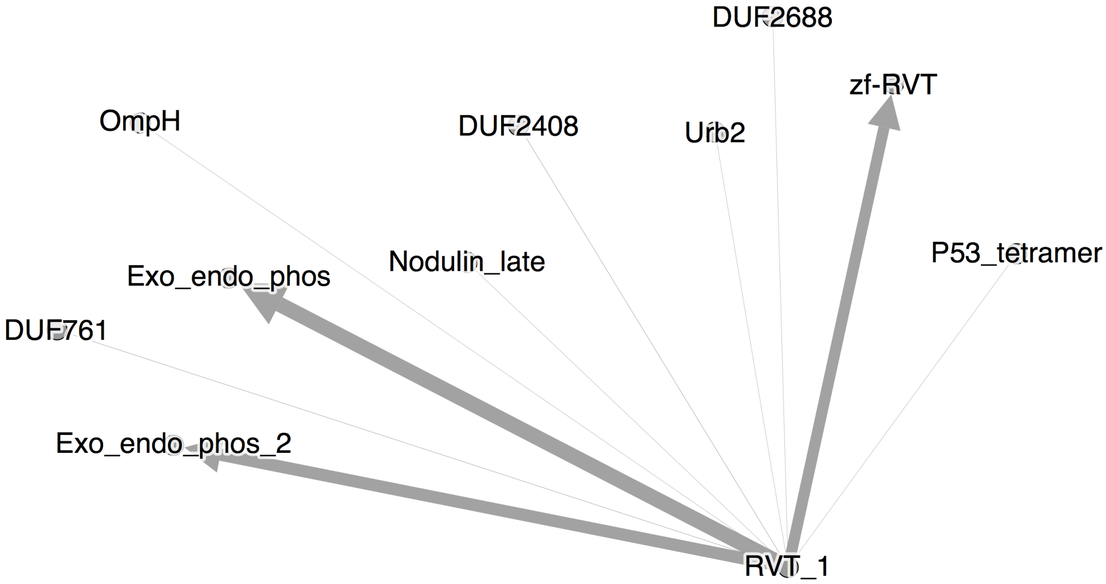


Figure B.6c: **Amphibia ORF2p domains**

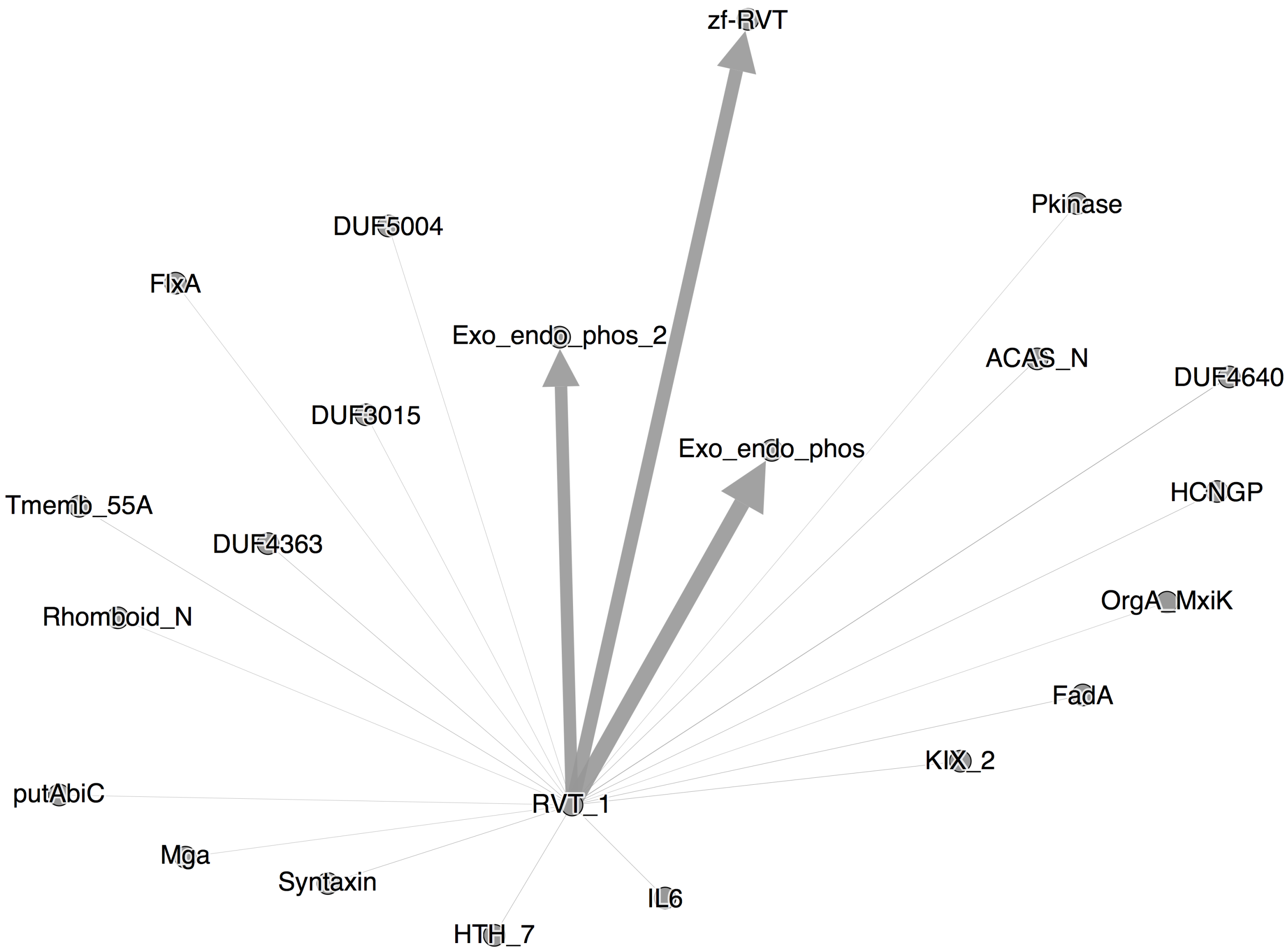


Figure B.6d: **Neopterygii ORF2p domains**

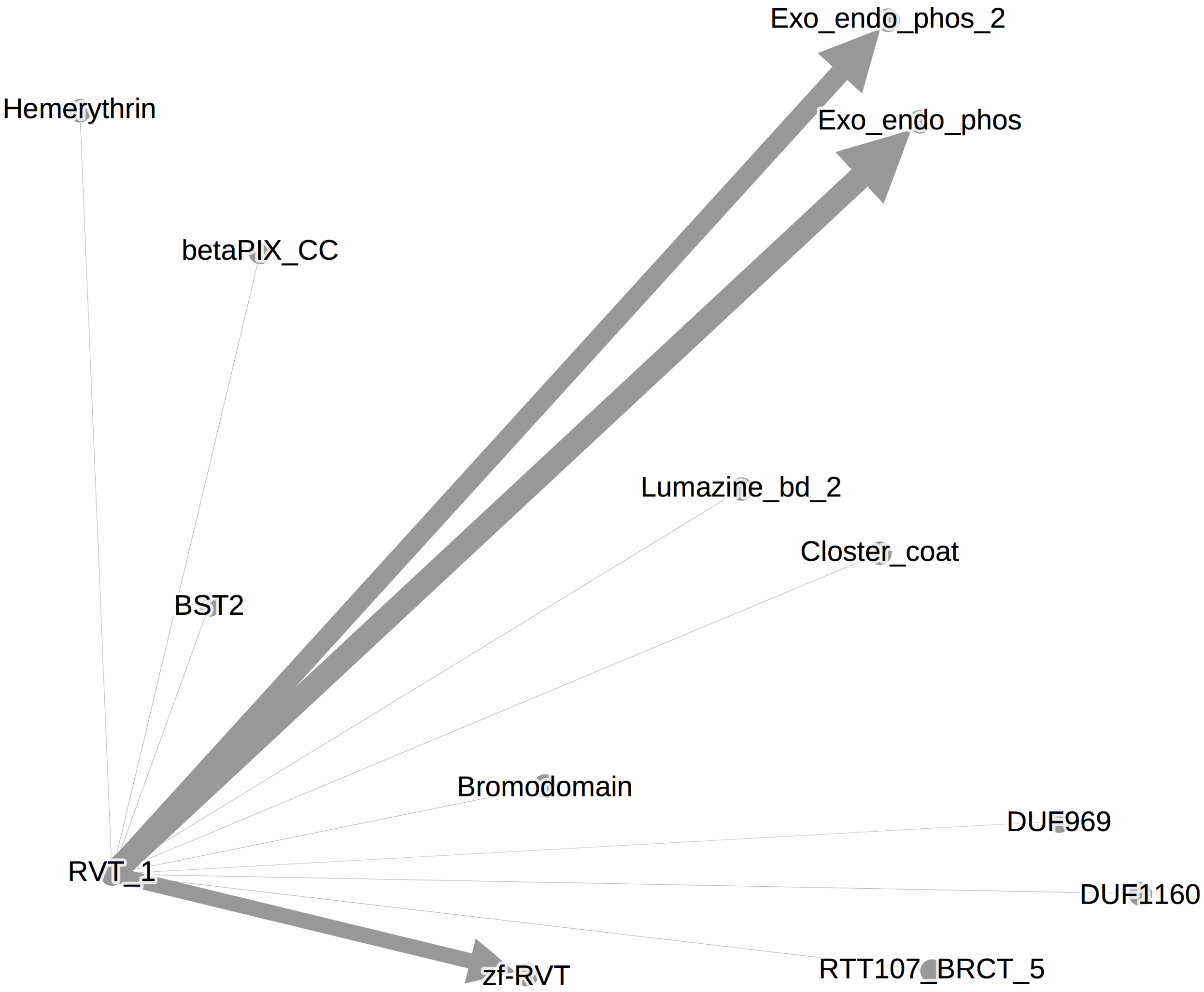


Figure B.6e: **Ecdysozoa ORF2p domains**

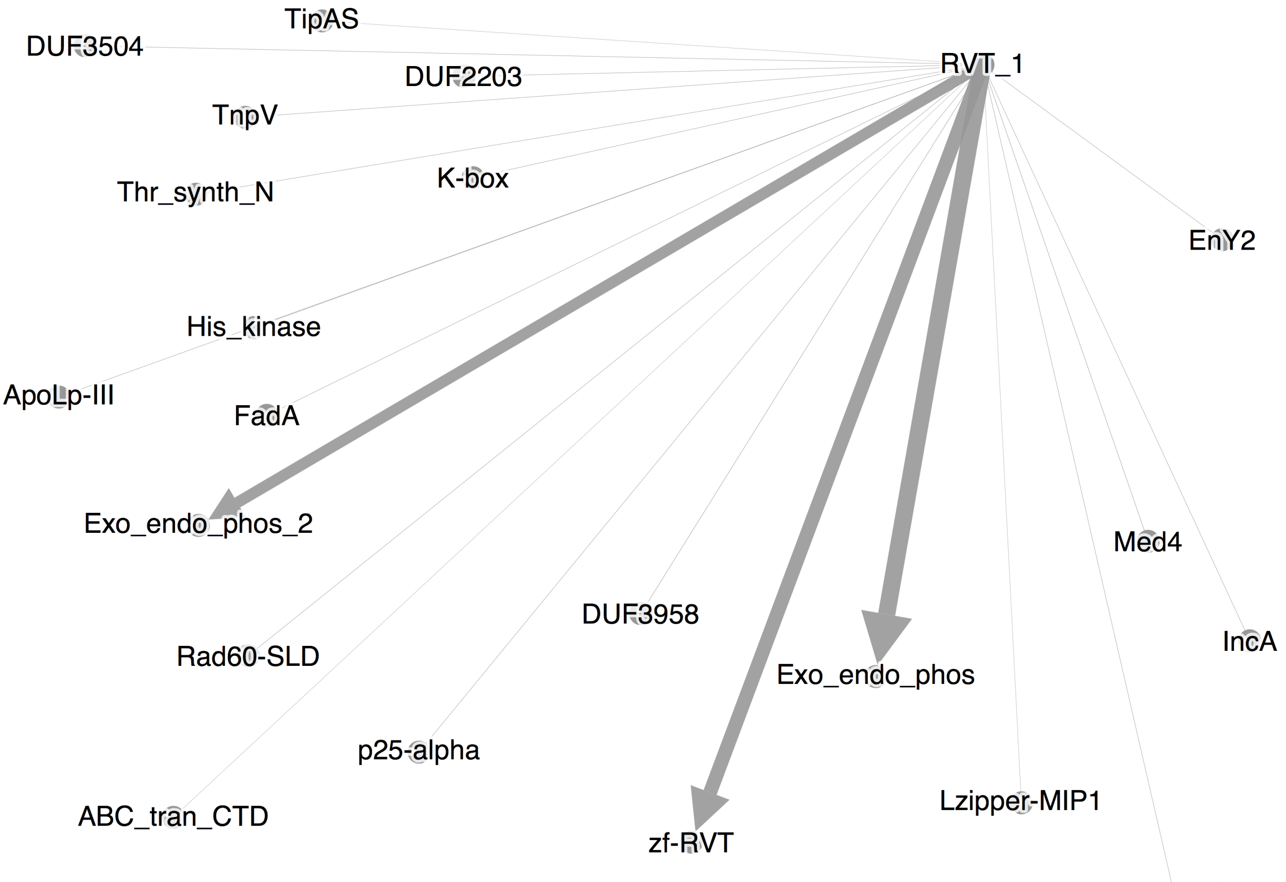


Figure B.6f: **Other, primitive non-mammalian species - ORF2p domains**

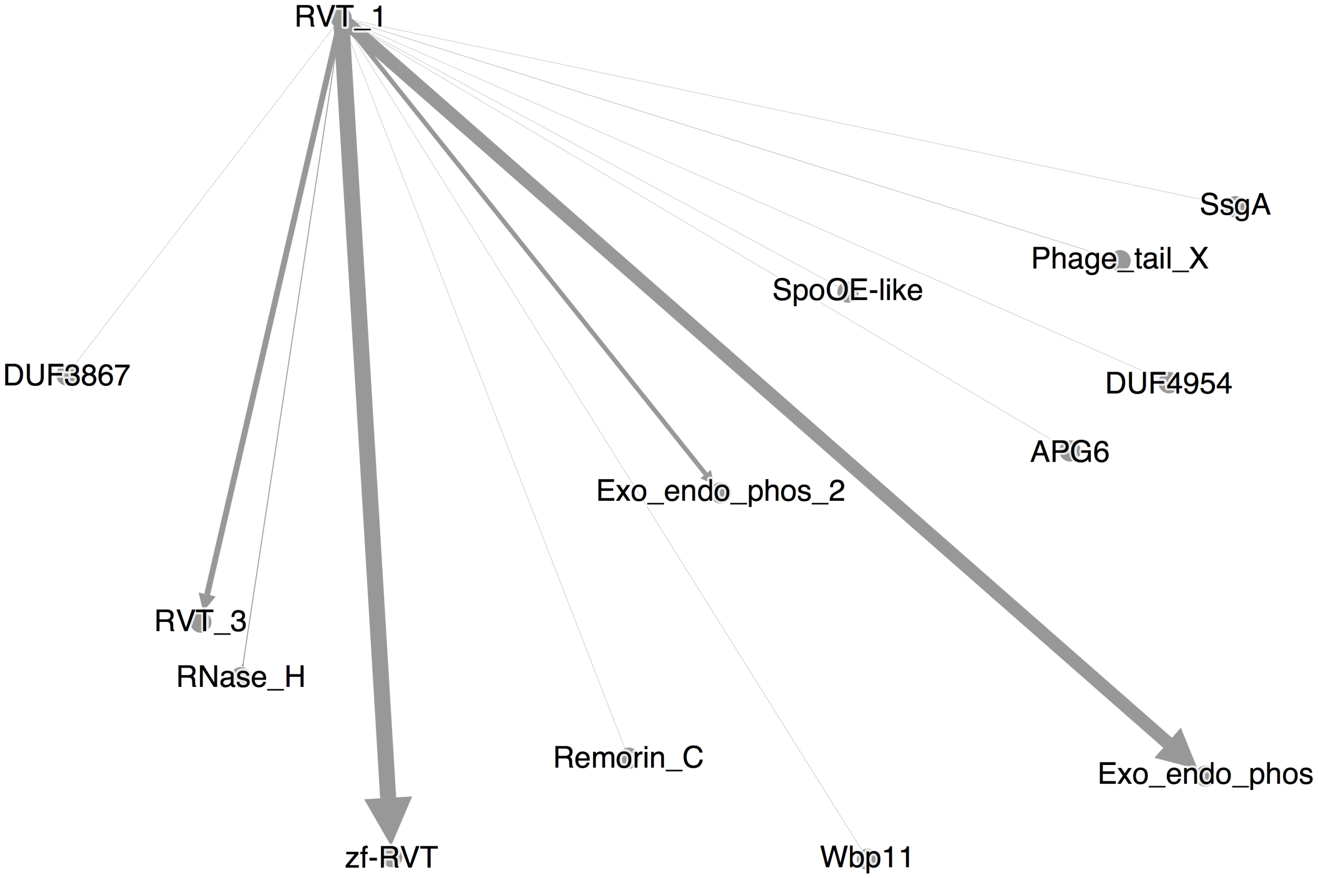


Figure B.6g: **Viridiplantae ORF2p domains**

Table 11: Domains found within ORF1p sequences

Table B.11: **Domains found within ORF1p sequences:** Known domains associated with L1 ORF1p include Transposase_22 (vertebrates) and RRM/zf-CCHC (plants, diverse species). This table summarises the known domains seen in each species, and key unknown domains which appear frequently (DUF4283 in plants) or with very high support (HTH_1 in *Coccomyxa subellipsoidea*). Note that even some mammals contain diverse L1s with RRM/zf-CCHC combinations.

	Known ORF1p domains			Novel ORF1p domains	
Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
MAMMALIAN SPECIES					
Monodelphis domestica	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3		
Macropus eugenii	Transposase_22				
Sarcophilus harrisii	Transposase_22				
Dasypus novemcinctus	Transposase_22		zf-CCHC_2		
Choloepus hoffmanni	Transposase_22		zf-CCHC_2		
Chrysochloris asiatica	Transposase_22				
Echinops telfairi	Transposase_22				
Orycteropus afer afer	Transposase_22				
Elephantulus edwardii	Transposase_22	RRM_1, RRM_6	zf-CCHC		
Trichechus manatus latirostris	Transposase_22		zf-CCHC, zf-CCHC_2		
Procavia capensis	Transposase_22				
Loxodonta africana	Transposase_22				
Erinaceus europaeus	Transposase_22				
Sorex araneus	Transposase_22				
Condylura cristata	Transposase_22				
Pteropus alecto	Transposase_22				
Pteropus vampyrus	Transposase_22				
Eidolon helvum	Transposase_22				
Rhinolophus ferrumequinum	Transposase_22				
Pteronotus parnellii	Transposase_22				
Eptesicus fuscus	Transposase_22				
Myotis brandtii	Transposase_22				
Myotis davidii	Transposase_22				
Myotis lucifugus	Transposase_22	RRM_1, RRM_6, RRM_7	zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Ceratotherium simum simum	Transposase_22	RRM_4			
Equus przewalskii	Transposase_22	Nup35_RRM			
Equus caballus (Thoroughbred)	Transposase_22	Nup35_RRM			
Equus caballus (Mongolian)	Transposase_22	Nup35_RRM			
Manis pentadactyla	Transposase_22				

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Felis catus	Transposase_22				
Panthera tigris altaica	Transposase_22				
Canis lupus familiaris	Transposase_22	RRM_1, RRM_5, RRM_6, RRM_7			
Ursus maritimus	Transposase_22				
Ailuropoda melanoleuca	Transposase_22				
Leptonychotes weddellii	Transposase_22				
Odobenus rosmarus divergens	Transposase_22	Nup35_RRM_2, RBM1CTR, RRM_1, RRM_3, RRM_5, RRM_6, RRM_7, RRM_occluded			
Mustela putorius furo	Transposase_22				
Camelus dromedarius	Transposase_22				
Camelus ferus	Transposase_22				
Vicugna pacos	Transposase_22				
Sus scrofa (Duroc)	Transposase_22	RRM_5			
Sus scrofa (Tibetan)	Transposase_22				
Sus scrofa (Ellegaard Gottingen minipig)	Transposase_22				
Balaenoptera acutorostrata scammoni	Transposase_22	Nup35_RRM_2, RRM_1, RRM_5, RRM_6, RRM_7			
Physeter catodon	Transposase_22				
Lipotes vexillifer	Transposase_22	RRM_4			
Tursiops truncatus	Transposase_22				
Orcinus orca	Transposase_22	RRM_1, RRM_3, RRM_5, RRM_6			
Pantholops hodgsonii	Transposase_22				
Capra hircus	Transposase_22				
Ovis aries (Texel)	Transposase_22				
Ovis aries musimon	Transposase_22	RRM_1, RRM_5, RRM_6, RRM_7, RRM_occluded			
Bubalus bubalis	Transposase_22				
Bison bison bison	Transposase_22				
Bos mutus	Transposase_22				
Bos indicus	Transposase_22				
Bos taurus	Transposase_22				
Ochotona princeps	Transposase_22				
Oryctolagus cuniculus	Transposase_22				
Ictidomys tridecemlineatus	Transposase_22				
Heterocephalus glaber	Transposase_22				

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Fukomys damarensis	Transposase_22				
Cavia aperea	Transposase_22				
Cavia porcellus	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_5, zf-CCHC_3, zf-CCHC_4		
Chinchilla lanigera	Transposase_22				
Octodon degus	Transposase_22				
Dipodomys ordii	Transposase_22				
Jaculus jaculus	Transposase_22	RRM_5			
Nannospalax galili	Transposase_22	RRM_5	zf-CCHC_3, zf-CCHC_5		
Mesocricetus auratus	Transposase_22				
Cricetulus griseus	Transposase_22		zf-CCHC_5, zf-CCHC, zf-CCHC_3, zf-CCHC_2		
Microtus ochrogaster	Transposase_22				
Peromyscus maniculatus bairdii	Transposase_22				
Rattus norvegicus	Transposase_22		zf-CCHC, zf-CCHC_4, zf-CCHC_5		
Mus musculus	Transposase_22		zf-CCHC		
Tupaia belangeri	Transposase_22	RRM_1			
Tupaia chinensis	Transposase_22	RRM_1, RRM_6			
Galeopterus variegatus	Transposase_22	RRM_occluded			
Otolemur garnettii	Transposase_22				
Microcebus murinus	Transposase_22	RRM_5, RRM_6	zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Tarsius syrichta	Transposase_22				
Callithrix jacchus	Transposase_22	Nup35_RRM_2, RRM_1, RRM_5, RRM_6, RRM_7			
Saimiri boliviensis boliviensis	Transposase_22				
Rhinopithecus roxellana	Transposase_22				
Nasalis larvatus	Transposase_22				
Chlorocebus sabaeus	Transposase_22	RRM_1, RRM_6, RRM_7			
Macaca fascicularis	Transposase_22				

Species	Transposase	RRM	zf-CCHC	DUF4283	HTH
Macaca mulatta	Transposase_22	RRM_1, RRM_5, RRM_6, RRM_7			
Papio anubis	Transposase_22				
Nomascus leucogenys	Transposase_22				
Pongo abelii	Transposase_22	RRM_6			
Gorilla gorilla gorilla	Transposase_22	RRM_1, RRM_5, RRM_6, RRM_7			
Pan paniscus	Transposase_22	RRM_1, RRM_6, RRM_7			
Pan troglodytes	Transposase_22				
Homo sapiens	Transposase_22				
NON-MAMMALIAN ANIMAL SPECIES					
Sauropsida					
Apalone spinifera			zf-CCHC, zf-CCHC_4		
Pelodiscus sinensis			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Chelonia mydas			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Chrysemys picta bellii	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Alligator mississippiensis		RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Alligator sinensis		RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		
Crocodylus porosus	Transposase_22	RRM_1, RRM_6, RRM_7	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Gavialis gangeticus	Transposase_22	RRM_1, RRM_6	zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		
Anolis carolinensis	Transposase_22		zf-CCHC		
Vipera berus berus	Transposase_22				
Crotalus mitchellii pyrrhus	Transposase_22				
Ophiophagus hannah	Transposase_22				
Amphibia					
Xenopus tropicalis	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		
Nanorana parkeri	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6		
Neopterygii					
Lepisosteus oculatus			zf-CCHC, zf-CCHC_2, zf-CCHC_6		
Anguilla anguilla	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Anguilla japonica	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		
Danio rerio	Transposase_22	RRM_1, RRM_6, RRM_7	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Astyanax mexicanus	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Oryzias latipes	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Poecilia formosa	Transposase_22	RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_4		
Xiphophorus maculatus	Transposase_22				
Fundulus heteroclitus	Transposase_22		zf-CCHC, zf-CCHC_2		
Takifugu flavidus			zf-CCHC, zf-CCHC_2, zf-CCHC_6		
Takifugu rubripes			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Tetraodon nigroviridis		RRM_6, RRM_7	zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Haplochromis burtoni	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_6		
Pundamilia nyererei	Transposase_22		zf-CCHC, zf-CCHC_6		
Maylandia zebra	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4		
Neolamprologus brichardi	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Oreochromis niloticus	Transposase_22	Nup35_RRM, RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6		
Sebastes nigrocinctus	Transposase_22				

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Sebastes rubrivinctus	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Gasterosteus aculeatus		RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_5		
Gadus morhua	Transposase_22	RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Ecdysozoa					
Aedes aegypti		Nup35_RRM_2, RRM_1, RRM_3, RRM_5, RRM_6, RRM_7, RRM_occluded	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4		
Culex quinquefasciatus		Nup35_RRM_2, RRM_1, RRM_5, RRM_6, RRM_occluded	zf-CCHC, zf-CCHC_4, zf-CCHC_5		
Anopheles arabiensis		RRM_1, RRM_6			
Anopheles atroparvus		RRM_1			
Anopheles culicifacies		RRM_occluded			
Anopheles dirus		RRM_1, RRM_6			
Anopheles epiroticus		RRM_1, RRM_6, RRM_occluded			
Anopheles farauti		RRM_1, RRM_5			
Anopheles funestus		RRM_6, RRM_occluded			
Anopheles gambiae		RRM_1, RRM_6, RRM_occluded			
Anopheles maculatus		RRM_6			
Anopheles sinensis		RRM_1			
Drosophila ananassae	RRM_1, RRM_6, RRM_7				
Drosophila virilis		RRM_5			
Ixodes ricinus	Transposase_22		zf-CCHC, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Ixodes scapularis	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_5		
Rhipicephalus microplus	Transposase_22				
Other					
Crassostrea gigas			zf-CCHC, zf-CCHC_2, zf-CCHC_3		
Lottia gigantea			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4		
Aplysia californica			zf-CCHC, zf-CCHC_3, zf-CCHC_4		
Nematostella vectensis	Transposase_22		zf-CCHC, zf-CCHC_3		
Strongylocentrotus purpuratus			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4		
Saccoglossus kowalevskii	Transposase_22	RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5		
Ciona savignyi			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6		
Branchiostoma floridae			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		
Lethenteron camtschaticum			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6		
Petromyzon marinus			zf-CCHC, zf-CCHC_2, zf-CCHC_4		

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Latimeria chalumnae	Transposase_22		zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6		
PLANT SPECIES					
Coccomyxa subellipsoidea C-169					HTH_1
Selaginella moellendorffii			zf-CCHC, zf-CCHC_4	DUF4283	
Pinus taeda			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Amborella trichopoda			zf-CCHC_4	DUF4283	
Phoenix dactylifera			zf-CCHC_4, zf-CCHC_6	DUF4283	
Musa acuminata			zf-CCHC, zf-CCHC_2	DUF4283	
Sorghum bicolor			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Zea mays			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Setaria italica			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Brachypodium distachyon			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Leersia perrieri			zf-CCHC, zf-CCHC_4	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Oryza barthii			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Oryza brachyantha			zf-CCHC, zf-CCHC_3, zf-CCHC_4	DUF4283	
Oryza glumipatula			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Oryza longistaminata			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Oryza meridionalis			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Oryza nivara			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Oryza punctata			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Oryza sativa Japonica Group			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Aegilops tauschii			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Triticum urartu			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Nelumbo nucifera			zf-CCHC_4	DUF4283	
Lupinus angustifolius		RRM_7	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Phaseolus vulgaris			zf-CCHC_4	DUF4283	
Cajanus cajan			zf-CCHC, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Glycine max		RRM_1, RRM_5, RRM_6, RRM_7	zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Glycine soja		RRM_1, RRM_5, RRM_6, RRM_7	zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Medicago truncatula		RRM_1	zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Trifolium pratense		RRM_1, RRM_5	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Lotus japonicus			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Malus x domestica			zf-CCHC, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Pyrus x bretschneideri			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	

Species	Transposase	RRM	zf-CCHC	DUF4283	HTH
Prunus mume			zf-CCHC, zf-CCHC_2, zf-CCHC_4	DUF4283	
Prunus persica			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6	DUF4283	
Fragaria iinumae			zf-CCHC_4, zf-CCHC_6	DUF4283	
Fragaria nubicola		Nup35_RRM_2	zf-CCHC, zf-CCHC_4, zf-CCHC_5	DUF4283	
Fragaria orientalis			zf-CCHC, zf-CCHC_4	DUF4283	
Fragaria vesca subsp. vesca			zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Fragaria x ananassa			zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Morus notabilis			zf-CCHC, zf-CCHC_3, zf-CCHC_4	DUF4283	
Cannabis sativa			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Castanea mollissima			zf-CCHC, zf-CCHC_4	DUF4283	
Betula nana			zf-CCHC_3, zf-CCHC_4	DUF4283	
Cucumis melo			zf-CCHC_4	DUF4283	
Cucumis sativus			zf-CCHC_2, zf-CCHC_4, zf-CCHC_6	DUF4283	
Citrullus lanatus			zf-CCHC, zf-CCHC_4	DUF4283	
Populus euphratica			zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Populus trichocarpa		RRM_1	zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Jatropha curcas			zf-CCHC, zf-CCHC_4	DUF4283	
Manihot esculenta			zf-CCHC, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Ricinus communis			zf-CCHC, zf-CCHC_4, zf-CCHC_5	DUF4283	
Eucalyptus camaldulensis		RRM_6	zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Eucalyptus grandis		Nup35_RRM_2, RRM_6	zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Carica papaya				DUF4283	
Arabidopsis halleri subsp. gemmifera			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Arabidopsis lyrata subsp. lyrata			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Arabidopsis thaliana			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Camelina sativa		Nup35_RRM_2	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Capsella rubella			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Brassica napus			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Brassica oleracea var. oleracea			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Brassica rapa			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Raphanus raphanistrum subsp. raphanistrum			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Raphanus sativus			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Aethionema arabicum		RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Arabis alpina		RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Eutrema parvulum			zf-CCHC_4	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Eutrema salsugineum			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Sisymbrium irio			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Leavenworthia alabamica			zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6	DUF4283	
Tarenaya hassleriana			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Gossypium arboretum		RRM_5	zf-CCHC, zf-CCHC_4	DUF4283	
Gossypium raimondii			zf-CCHC, zf-CCHC_4	DUF4283	
Theobroma cacao			zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Aquilaria agallochum			zf-CCHC, zf-CCHC_4	DUF4283	
Azadirachta indica			zf-CCHC_4	DUF4283	
Citrus clementine		RRM_5, RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_5, zf-CCHC_6	DUF4283	
Citrus sinensis		RRM_5, RRM_6	zf-CCHC, zf-CCHC_2, zf-CCHC_4, zf-CCHC_6	DUF4283	
Vitis vinifera		RRM_5, RRM_6		DUF4283	
Amaranthus hypochondriacus		RRM_occluded	zf-CCHC, zf-CCHC_3, zf-CCHC_4	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Beta vulgaris subsp. vulgaris		Nup35_RRM_2, RRM_1, RRM_5, RRM_6, RRM_7	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Spinacia oleracea		Nup35_RRM_2, RRM_1, RRM_2, RRM_5, RRM_6, RRM_7	zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_5	DUF4283	
Dianthus caryophyllus		RRM_3	zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Actinidia chinensis			zf-CCHC, zf-CCHC_4, zf-CCHC_5	DUF4283	
Vaccinium macrocarpon			zf-CCHC_4	DUF4283	
Primula veris			zf-CCHC_2, zf-CCHC_4	DUF4283	
Solanum arcanum			zf-CCHC_3, zf-CCHC_4	DUF4283	
Solanum habrochaites		Nup35_RRM_2	zf-CCHC, zf-CCHC_3, zf-CCHC_4	DUF4283	
Solanum lycopersicum		Nup35_RRM_2	zf-CCHC_4	DUF4283	
Solanum melongena			zf-CCHC_4	DUF4283	
Solanum pennellii			zf-CCHC_4	DUF4283	
Solanum pimpinellifolium			zf-CCHC, zf-CCHC_2, zf-CCHC_4	DUF4283	
Solanum tuberosum			zf-CCHC, zf-CCHC_3, zf-CCHC_4	DUF4283	
Capsicum annuum			zf-CCHC_4	DUF4283	
Nicotiana sylvestris		RRM_5	zf-CCHC, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Nicotiana tomentosiformis		RRM_5	zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	

Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
Fraxinus excelsior			zf-CCHC, zf-CCHC_2, zf-CCHC_3, zf-CCHC_4, zf-CCHC_6	DUF4283	
Sesamum indicum			zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Mimulus guttatus			zf-CCHC, zf-CCHC_4, zf-CCHC_6	DUF4283	
Conyza canadensis				DUF4283	

Table 12: ORF1p domain summary

Table B.12: **ORF1p domain summary:** Used to construct the bar plot in Fig. 9(a) of the main text.

	Known ORF1p domains			Novel ORF1p domains	
Species	**Transposase**	**RRM**	**zf-CCHC**	**DUF4283**	**HTH**
MAMMALS	99850	167	29	0	0
NON-MAMMALIAN ANIMALS	943	624	1127	0	0
PLANTS	0	258	5784	6491	13

Figure 7: ORF1p associated domain network diagrams

As with the ORF2p network graphs - shows strongly associated domains with known ORF1p domains (Transposase_22, RRM or zf-CCHC). The graph for Viridiplantae (plants) is shown as Figure 9c in the main text.

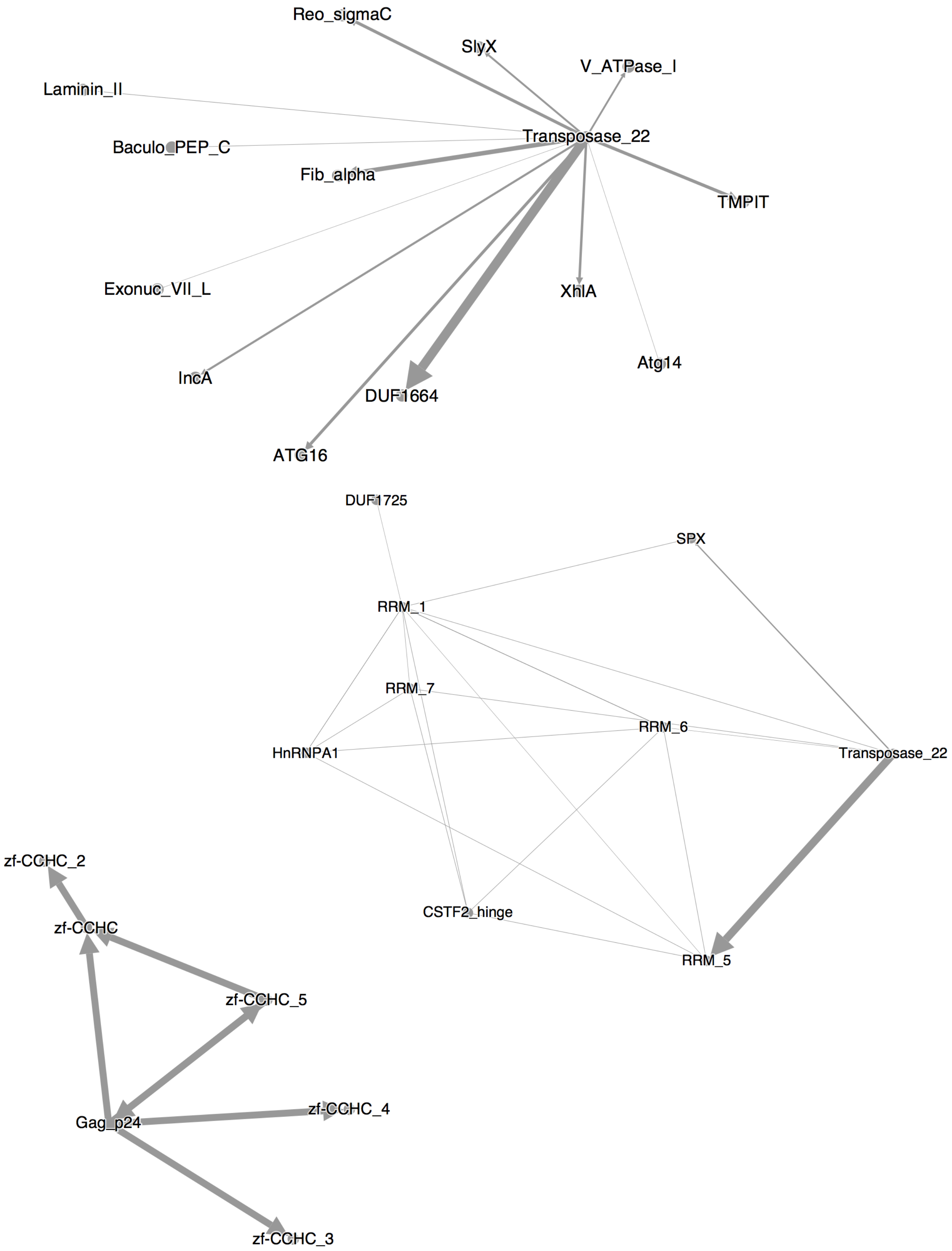


Figure B.7a: **Mammalia ORF1p domains**

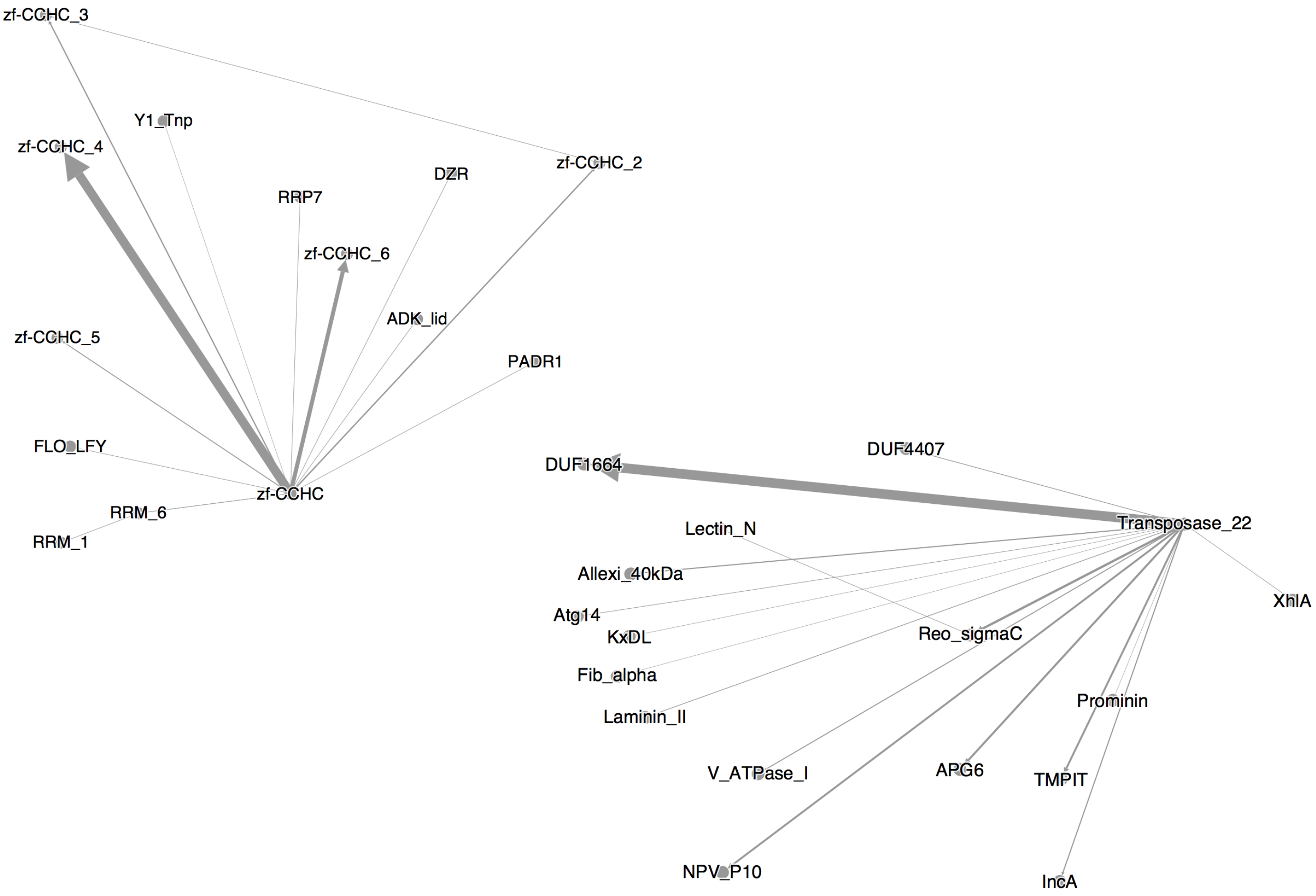


Figure B.7b: **Sauropsida ORF1p domains**

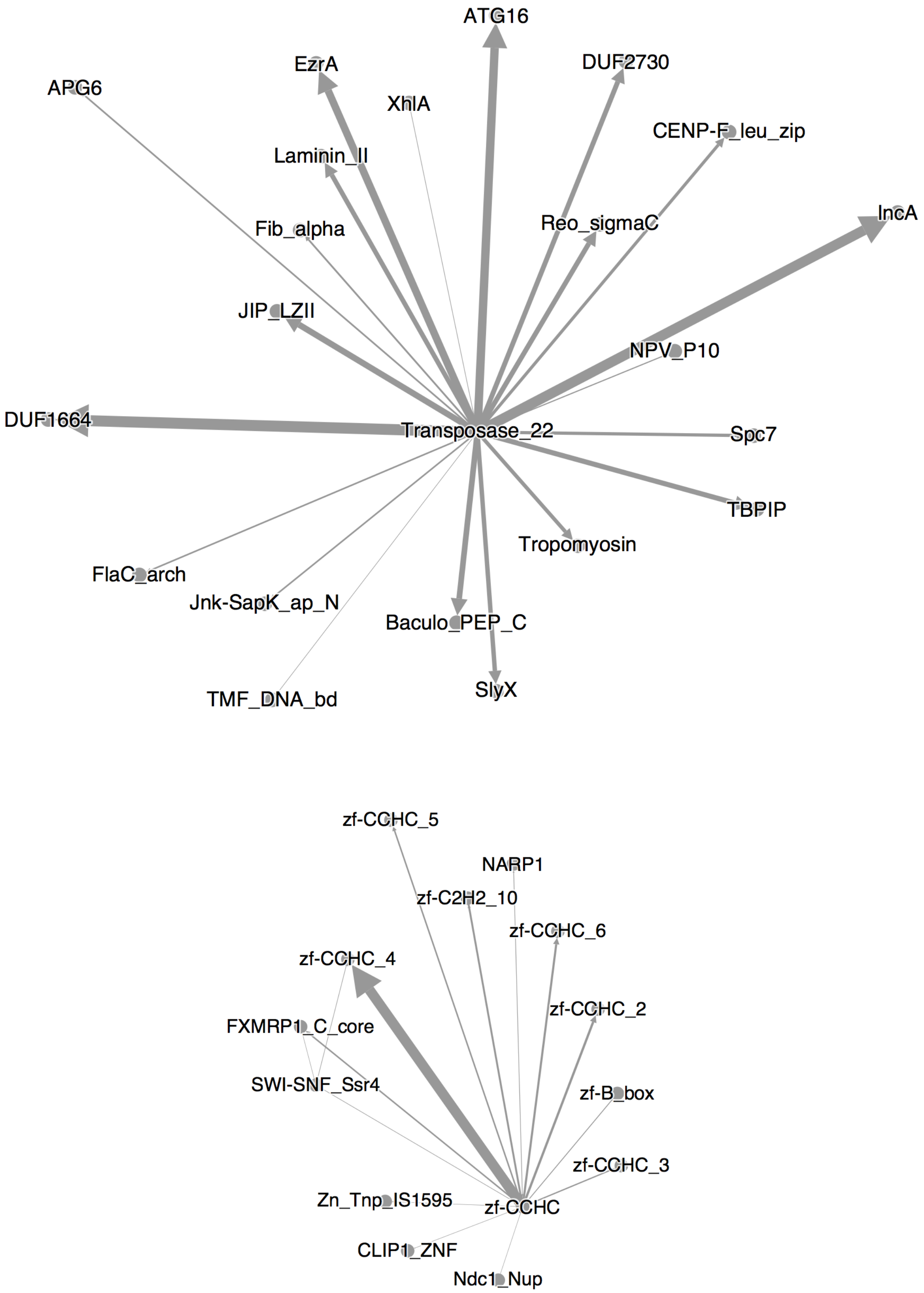


Figure B.7c: **Amphibia ORF1p domains**

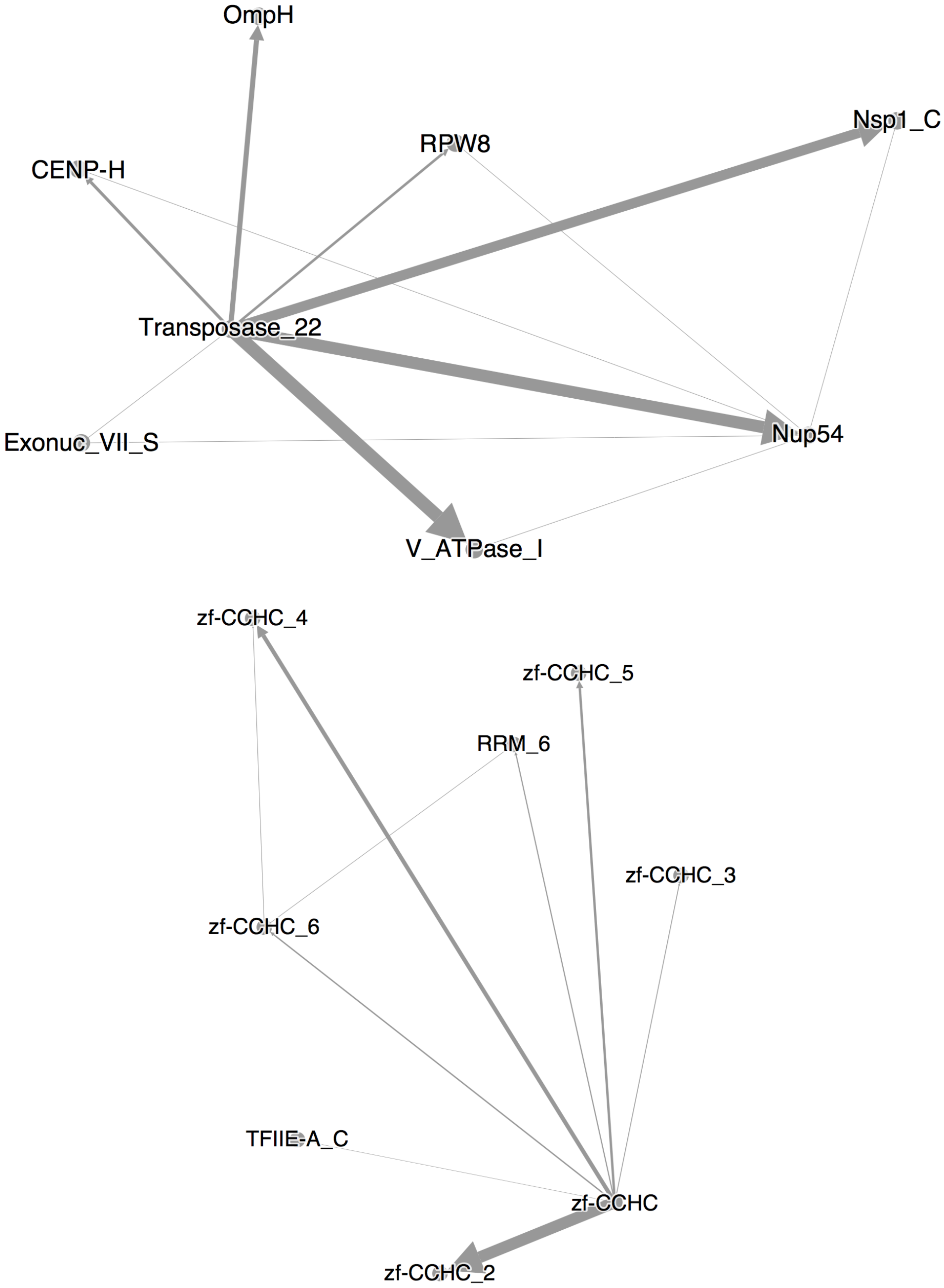


Figure B.7d: **Neopterygii ORF1p domains**

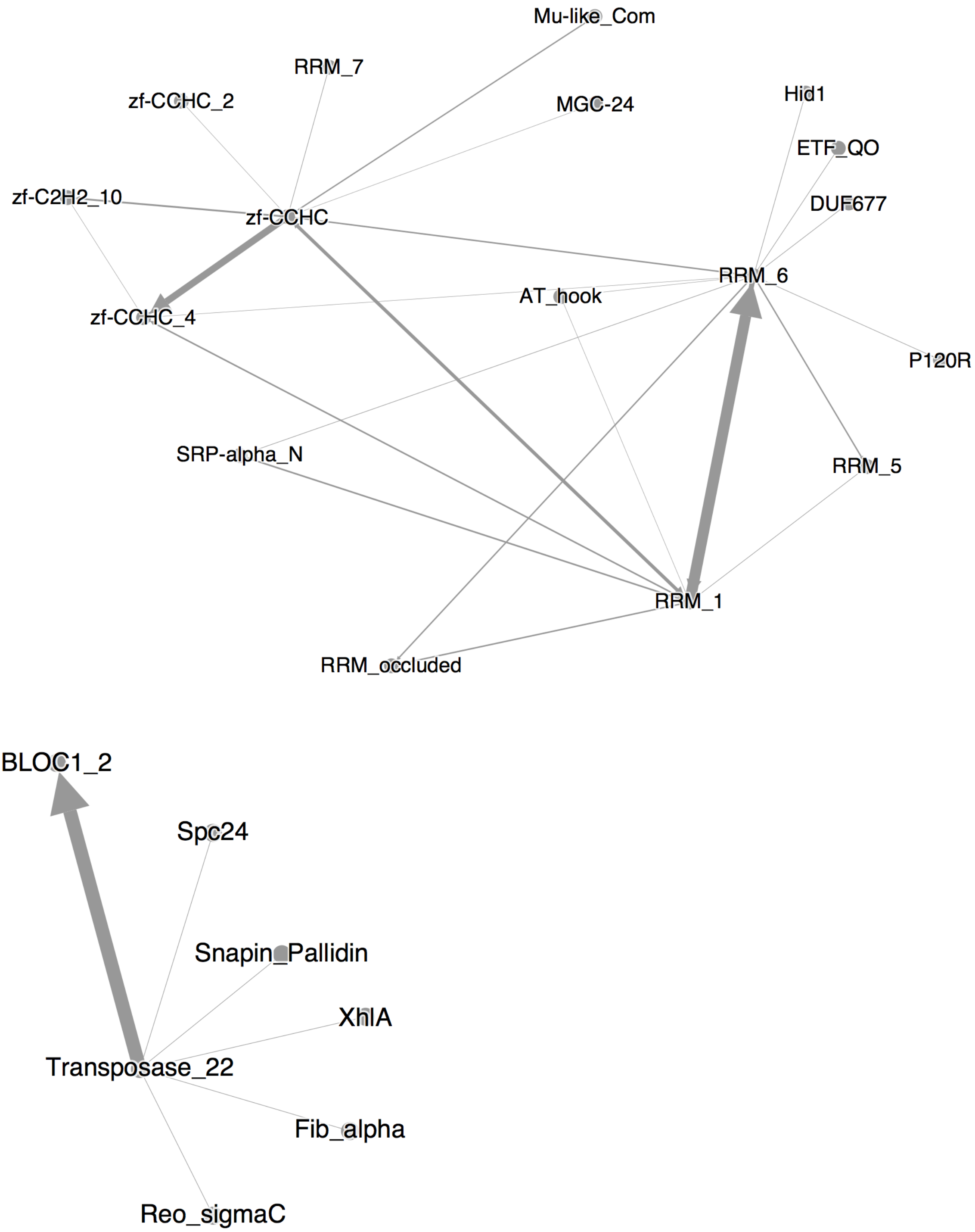


Figure B.7e: **Ecdysozoa ORF1p domains**

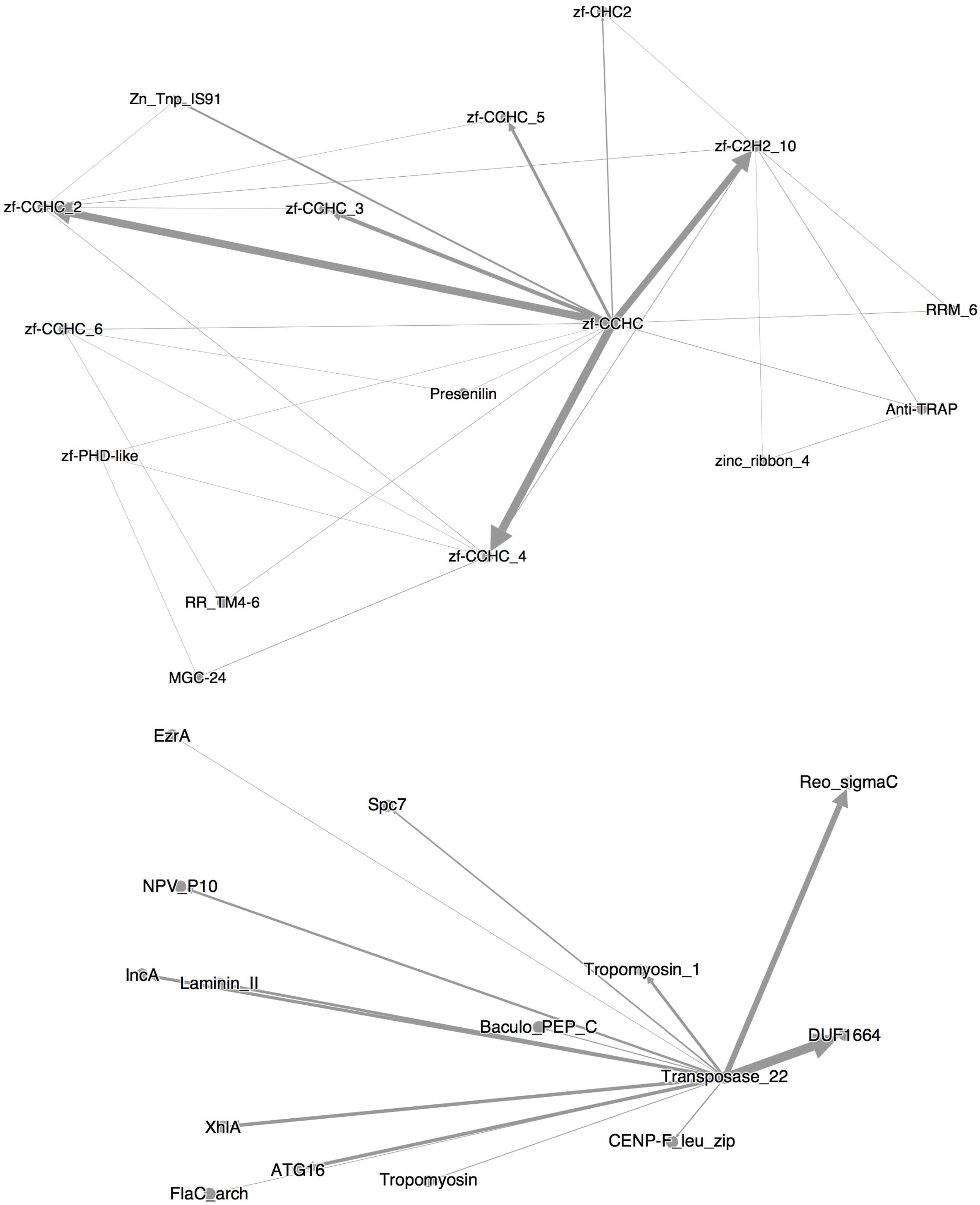


Figure B.7f: **Other, primitive non-mammalian species - ORF1p domains**

Appendix C

Supplementary for Chapter 3

C.1 Materials and Methods

Extraction of L1 and BovB retrotransposons from genome data

To extract the retrotransposons of interest, we used the methods and genomes previously described in Ivancevic *et al.* (2016) [54]. Briefly, this involved downloading 499 publicly available genomes (and acquiring 4 more from collaborations), then using two independent searching strategies (LASTZ [55] and TBLASTN [56]) to identify and characterise L1 amd BovB elements. A third program, CENSOR [57], was used with the RepBase library of known repeats [58] to verify hits with a reciprocal best-hit check. The raw L1 results have been previously published [54]; the BovB results are included below.

Extraction and clustering of conserved amino acid residues

Starting with BovBs, USEARCH [59] was used to find open readings frames (ORFs), with function -fastx_findorfs and parameters -aaout (for amino acid output) and -orfstyle 7 (to allow non-standard start codons). HMMer [60] was used to identify reverse transcriptase (RT) domains within the ORFs. RT domains were extracted using the envelope coordinates from the HMMer domain hits table (-domtblout), with minimum length 200 amino acid residues. The BovB RT domains from all species were collated into one file and clustered with UCLUST [59]. This was done as an initial screening to detect potential horizontal transfer candidates. The process was then repeated with L1 elements.

Clustering of nucleotide sequences to build one consensus per species

The canonical BovB retrotransposon is 3.2 kb in length [53, 58], although this varies slightly between species. In this study, we classified BovB nucleotide sequences =>2.4kb and <=4kb as full-length. We wanted to construct a BovB representative for each species. Accordingly, for each species, UCLUST [59] was used to cluster full-length BovB sequences at varying identities between 65-95%. A consensus sequence of each cluster was generated using the UCLUST -consout option.

The ideal cluster identity was chosen based on the number and divergence of sequences in a cluster. E.g. for species with few BovBs, a lower identity was allowed; whereas for species with thousands of

BovBs, a higher identity was needed to produce an alignable cluster. The final clustering identity and cluster size for each species is shown below in Table 1. Note that the bat species are not included in this table - they were clustered separated, due to the high level of divergence between BovBs.

This method was tested on L1 retrotransposons, but the results were not ideal; most species simply had too many L1 sequences. Other methods tested on both BovBs and L1s included using centroids instead of consensus sequences (this gave better alignments but was less representative of the cluster), and using the same clustering identity for all species (e.g. 80% - this did not work well for species with less than 100 elements in the genome).

Table 1: Clustering identities for BovB consensus sequences

Table C.1: **Clustering summary:** Clustering identity used to generate a single representative consensus per species, for the single consensus tree used to distinguish RTEs from BovBs and for Fig 2a in the text. Cluster size means the number of sequences in that cluster.

Species	**Clustering identity (%)**	**Dominant cluster**	**Cluster size**
MAMMALIA			
Tachyglossus aculeatus	70	Cluster4	11
Ornithorhynchus anatinus	70	Cluster6	16
Monodelphis domestica	65	Cluster9	7
Macropus eugenii	65	Cluster0	19
Sarcophilus harrisii	64	Cluster38	5
Chrysochloris asiatica	88	Cluster3771	88
Echinops telfairi	70	Cluster39	96
Orycteropus afer afer	88	Cluster10752	135
Elephantulus edwardii	87	Cluster430	135
Trichechus manatus latirostris	87	Cluster382	192
Procavia capensis	80	Cluster264	103
Loxodonta africana	90	Cluster2684	94
Ceratotherium simum simum	70	Cluster0	23
Equus przewalskii	70	Cluster4	20
Equus caballus (Thoroughbred)	70	Cluster2	18
Equus caballus (Mongolian)	70	Cluster4	17
Pantholops hodgsonii	88	Cluster61	115
Capra hircus	83	Cluster250	142
Ovis aries (Texel)	90	Cluster184	190
Ovis aries musimon	88	Cluster1388	108
Bubalus bubalis	89	Cluster563	111
Bison bison bison	88	Cluster459	242
Bos mutus	84	Cluster433	106
Bos indicus	90	Cluster279	156
Bos taurus	90	Cluster179	137

Species	**Clustering identity (%)**	**Dominant cluster**	**Cluster size**
SAUROPSIDA			
Apalone spinifera	70	Cluster0	11
Pelodiscus sinensis	70	Cluster2	5
Chelonia mydas	70	Cluster1	29
Chrysemys picta bellii	70	Cluster0	32
Alligator mississippiensis	70	Cluster0	3
Alligator sinensis	70	Cluster0	2
Crocodylus porosus	70	Cluster0	5
Gavialis gangeticus	70	Cluster0	4
Pogona vitticeps	87	Cluster15	175
Anolis carolinensis	78	Cluster2	104
Vipera berus berus	70	Cluster4	85
Crotalus mitchellii pyrrhus	70	Cluster9	2
Ophiophagus hannah	70	Cluster14	25
Python bivittatus	70	Cluster18	27
NEOPTERYGII			
Lepisosteus oculatus	70	Cluster0	3
Danio rerio	-	Singleton	1
Astyanax mexicanus	-	No full-length seqs	0
Oryzias latipes	-	No full-length seqs	0
Poecilia formosa	-	No full-length seqs	0
Xiphophorus maculatus	-	No full-length seqs	0
Fundulus heteroclitus	-	No full-length seqs	0
Takifugu flavidus	-	No full-length seqs	0
Takifugu rubripes	-	No full-length seqs	0
Tetraodon nigroviridis	-	No full-length seqs	0
Cynoglossus semilaevis	-	Singleton	1
Sebastes rubrivinctus	-	No full-length seqs	0
Gadus morhua	70	Cluster0	2
ECDYSOZOA			
Ladona fulva	70	Cluster0	40
Diaphorina citri	90	Cluster1	4
Pachypsylla venusta	-	Singleton	1
Acyrthosiphon pisum	-	No full-length seqs	0
Nilaparvata lugens	90	Cluster0	9
Cimex lectularius	-	Singleton	1
Agrilus planipennis	70	Cluster0	8
Papilio glaucus	-	Singleton	1
Papilio polytes	90	Cluster0	3
Papilio xuthus	-	Singleton	1
Heliconius melpomene melpomene	-	Singleton	1
Danaus plexippus	70	Cluster0	2
Bombyx mori	-	Singleton	1

Species	**Clustering identity (%)**	**Dominant cluster**	**Cluster size**
Manduca sexta	-	Singleton	1
Plutella xylostella	90	Cluster2	2
Linepithema humile	-	Singleton	1
Camponotus floridanus	-	No full-length seqs	0
Acromyrmex echinatior	-	No full-length seqs	0
Atta cephalotes	-	No full-length seqs	0
Solenopsis invicta	85	Cluster1	2
Pogonomyrmex barbatus	-	Singleton	1
Harpegnathos saltator	-	No full-length seqs	0
Cerapachys biroi	-	No full-length seqs	0
Blattella germanica	90	Cluster0	3
Zootermopsis nevadensis	-	Singleton	1
Ixodes ricinus	-	No full-length seqs	0
Ixodes scapularis	-	No full-length seqs	0
Rhipicephalus microplus	-	No full-length seqs	0
Metaseiulus occidentalis	-	No full-length seqs	0
Centruroides exilicauda	70	Cluster3	7
Mesobuthus martensii	75	Cluster5	34
Limulus polyphemus	70	Cluster3	7
Priapulus caudatus	-	Singleton	1
ROTIFERA			
Adineta vaga	70	Cluster0	8
PLATYHELMINTHES			
Schistosoma curassoni	70	Cluster2	45
Schistosoma haematobium	70	Cluster0	39
Schistosoma japonicum	-	No full-length seqs	0
Schistosoma mansoni	70	Cluster0	145
Schistosoma margrebowiei	75	Cluster1	127
Schistosoma mattheei	70	Cluster0	24
Schistosoma rodhaini	70	Cluster1	49
ANNELIDA			
Helobdella robusta	90	Cluster0	2
MOLLUSCA			
Crassostrea gigas	90	Cluster0	2
Aplysia californica	70	Cluster1	58
Biomphalaria glabrata	-	No full-length seqs	0
CNIDARIA			
Nematostella vectensis	85	Cluster0	8
ECHINOIDEA			
Lytechinus variegatus	-	No full-length seqs	0
Strongylocentrotus purpuratus	85	Cluster0	4
ASTEROIDEA			
Patiria miniata	-	Singleton	1

Species	Clustering identity (%)	Dominant cluster	Cluster size
ENTEROPNEUSTA			
Saccoglossus kowalevskii	-	No full-length seqs	0
TUNICATA			
Ciona savignyi	70	Cluster0	8
Botryllus schlosseri	-	No full-length seqs	0
LEPTOCARDII			
Branchiostoma floridae	70	Cluster1	3
CEPHALASPIDOMORPHI			
Lethenteron camtschaticum	-	No full-length seqs	0
Petromyzon marinus	-	No full-length seqs	0
SARCOPTERYGII			
Latimeria chalumnae	70	Cluster0	32

Inferring a phylogeny from consensus sequences

Consensus sequences were aligned with MUSCLE [61]. The multiple alignment was processed with Gblocks [62] to extract conserved blocks, with default parameters except min block size: 5, allowed gaps: all. FastTree [63] was used to infer a maximum likelihood phylogeny using a general time reversible (GTR) model and gamma approximation on substitution rates. Geneious Tree Builder [64] was used to infer a second tree using the neighbor-joining method with 1000 bootstrap replicates.

Distiguishing RTEs from BovBs

All sequences which identified as BovB or RTE were kept and labelled accordingly to their closest RepBase classification [58]. However, there appeared to be numerous discrepancies with the naming: e.g. some RTE sequences shared >90% identity to BovBs, and vice versa. BovB retrotransposons are a subclass of RTE, and they were only discovered relatively recently. It is likely that several so-called RTE sequences are actually BovBs.

To determine which species had BovB sequences, and which only had RTEs, we used the species consensus approach to build a BovB/RTE phylogeny (see Figure 1 below). This effectively separated BovB-containing species from RTE-containing species. The RTE sequences were discarded from further analyses.

Clustering of nucleotide BovB sequences from bats and *Xenopus*

A reliable BovB consensus could not be generated for any of the 10 bat species because the sequences were too divergent and degraded. Some bat BovBs seemed similar to equid BovBs; others did not. Likewise, the single full-length BovB from frog *Xenopus tropicalis* was very different to canonical BovBs, sharing highest identity with the bats.

In an effort to characterise these BovBs into families, we grouped all full-length BovB sequences from the bats, frog, equids and white rhino into a single file. We also added two RepBase equid sequences (RTE-1_EC and BovB_Ec) and 1 RepBase bat sequence (BovBa-1_EF) [58]. After clustering, we expected to find one family of equid BovBs, the equid RTE sequence as an outlier, and numerous families containing bat and frog BovBs.

The actual findings are described in the manuscript (Fig. 2b). We used UCLUST [59] to cluster the sequences (function -cluster_fast with parameters -id, -uc, -clusters). The highest identity at which there were only 2 clusters/families was 40%. At higher identities, the equid BovBs stayed together but the bat and frog BovBs were lost as singletons.

To confirm the clustering, we also used MUSCLE to align all the sequences and FastTree to infer a phylogeny (see Figure 2 below).

HT candidate identification - BovBs and L1s

We compiled all confirmed BovB and L1 sequences into separate multi-fasta databases (316,017 and 1,048,478 sequences respectively). The length cutoff for BovBs was =>2.4kb and <=4kb; for L1s, =>3kb. BovBs were analysed first to identify characteristics of horizontal transfer events.

To detect HT candidates, we used the all-against-all clustering strategy described in El Baidouri *et al.* [65]. Briefly, this method use a nucleotide BLAST [56] to compare every individual sequence in a database against every other sequence; hence the term all-against-all. BLAST parameters were as follows: -r 2, -e 1e-10, -F F, -m 8 (for tabular output). The SiLiX program [66] is then used to filter the BLAST output and produce clusters or families that meet the designated identity threshold.

For BovB sequences, we tested identities of 40-90%. High identity thresholds were useful for finding very recent HT events (e.g. over 90% identity between the bed bug and snakes). However, the majority of clusters contained several copies of the same BovB family from a single species - indicative of vertical inheritance. Using a lower identity threshold was more informative for capturing ancient HT events. At 50% identity, the clustering preserved the recent, high-identity HT events while also finding the ancient, lower-identity HT events. We concluded that this was the best identity to use for our particular dataset of species, considering it includes widely divergent branches of Eukaryota.

Clusters were deemed HT candidates if they contained BovB elements belonging to at least two different species. To reduce the number of possible HT clusters, we went one step further and kept only the clusters which demonstrated cross-Order transfer (e.g. BovBs from Monotremata and Afrotheria in the same cluster). All potential HT candidates were validated by checking that the they were not located on short, isolated scaffolds or contigs in the genome. The flanking regions of each HT candidate pair were extracted and checked (via pairwise alignment) to ensure that the high sequence identity was restricted to the BovB region. This was done to check for contamination or orthologous regions. Phylogenies of HT candidate clusters were inferred using maximum likelihood and neighbour-joining methods (1000 bootstraps).

The same procedure was performed to screen for nucleotide L1 HT candidates. As an extra step for L1s, we also used all ORF1 and ORF2 amino acid sequences from a previous analysis [54] to conduct similar all-against-all BLAST searches. However, the amino acid clusterings did not produce any possible HT candidates.

C.2 Results

BovB presence across the eukaryotic tree of life

Table 2: BovBs in nucleotide nr/htgs databases, found using TBLASTN

Table C.2: **TBLASTN results:** Shows the results for the top hit found in each species. Note that a TBLASTN hit does not necessarily mean that the genome contains BovBs - a lot of false positives were screened out with later steps. TBLASTN search parameters were default except the e-value was changed to 1e-5. Input was the ORF protein from 5 full-length BovB elements from Repbase: BovB (*Bos taurus*, Bt), BovB_ACo (*Agkistrodon contortrix*, Ac), BovB_PMo (*Python molurus*, Pm), BovB_Ta (*Tachyglossus aculeatus*, Ta), BovB_VA (*Vipera ammodytes*, Va). Databases searched were the NCBI non-redundant nucleotide collection (nr) and high throughput genomic sequences (htgs). The table below shows the number of hits that were produced from each query and the statistics for the top hit. bitscore = max score; evalue = e-value; pident = percentage identity; qstart = start coordinate of the hit on the query sequence; qend = query end coordinate.

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
MAMMALIA							
1	*Tachyglossus aculeatus* (taxid:9261)	36 Bt, 27 Ac, 33 Pm, 44 Ta, 36 Va	424	1e-159	86.83	440	682
2	*Ornithorhynchus anatinus* (taxid:9258)	524 Bt, 345 Ac, 418 Pm, 685 Ta, 526 Va	410	2e-143	77.37	440	712
3	*Monodelphis domestica* (taxid:13616)	1677 Bt, 1523 Ac, 1672 Pm, 1482 Ta, 1677 Va	343	2e-121	47.26	309	764
4	*Macropus eugenii* (taxid:9315)	443 Bt, 319 Ac, 452 Pm, 379 Ta, 436 Va	477	0.0	46.87	109	750
5	*Sarcophilus harrisii* (taxid:9305)	38 Bt, 31 Ac, 33 Pm, 34 Ta, 36 Va	128	1e-30	27.22	150	479
6	*Dasypus novemcinctus* (taxid:9361)	307 Bt, 352 Ac, 333 Pm, 366 Ta, 259 Va	138	3e-32	28.01	177	575
7	*Choloepus hoffmanni* (taxid:9358)	528 Bt, 559 Ac, 536 Pm, 537 Ta, 491 Va	150	1e-36	28.70	171	582
8	*Chrysochloris asiatica* (taxid:185453)	No significant similarity found					
9	*Echinops telfairi* (taxid:9371)	5297 Bt, 2736 Ac, 4844 Pm, 5821 Ta, 6043 Va	321	8e-93	44.39	394	783
10	*Orycteropus afer afer* (taxid:1230840)	No significant similarity found					
11	*Elephantulus edwardii* (taxid:28737)	No significant similarity found					
12	*Trichechus manatus latirostris* (taxid:127582)	No significant similarity found					
13	*Procavia capensis* (taxid:9813)	7628 Bt, 3964 Ac, 6782 Pm, 7883 Ta, 8088 Va	523	2e-162	40.66	47	787

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
14	*Loxodonta africana* (taxid:9785)	11941 Bt, 6303 Ac, 10669 Pm, 12164 Ta, 12708 Va	972	0.0	50.35	1	1008
15	*Erinaceus europaeus* (taxid:9365)	1 Ta	51.2	5e-06	28.42	457	551
16	*Sorex araneus* (taxid:42254)	533 Bt, 404 Ac, 579 Pm, 482 Ta, 481 Va	239	1e-63	26.56	197	983
17	*Condylura cristata* (taxid:143302)	No significant similarity found					
18	*Pteropus alecto* (taxid:9402)	13 Bt, 8 Ac, 13 Pm, 21 Ta, 10 Va	124	1e-28	41.00	557	755
19	*Pteropus vampyrus* (taxid:132908)	94 Bt, 84 Ac, 104 Pm, 86 Ta, 72 Va	228	1e-63	37.37	124	505
20	*Eidolon helvum* (taxid:77214)	No significant similarity found					
21	*Megaderma lyra* (taxid:9413)	No significant similarity found					
22	*Rhinolophus ferrumequinum* (taxid:59479)	291 Bt, 269 Ac, 291 Pm, 266 Ta, 264 Va	261	2e-73	45.73	525	812
23	*Pteronotus parnellii* (taxid:59476)	No significant similarity found					
24	*Eptesicus fuscus* (taxid:29078)	No significant similarity found					
25	*Myotis brandtii* (taxid:109478)	7 Bt, 1 Ac, 7 Pm, 7 Ta, 7 Va	214	3e-58	28.12	105	670
26	*Myotis davidii* (taxid:225400)	6 Ta, 6 Va	72.0	4e-12	48.48	932	997
27	*Myotis lucifugus* (taxid:59463)	105 Bt, 107 Ac, 113 Pm, 141 Ta, 88 Va	158	8e-39	34.41	212	515
28	*Ceratotherium simum simum* (taxid:73337)	3 Bt, 1 Ac, 3 Pm, 2 Ta, 3 Va	61.2	1e-08	70.27	649	685
29	*Equus przewalskii* (taxid:9798)	6 Bt, 3 Ac, 7 Pm, 10 Ta, 9 Va	146	5e-39	56.90	383	497
30	*Equus caballus* (Thoroughbred) (taxid:9796)	287 Bt, 280 Ac, 302 Pm, 270 Ta, 221 Va	239	3e-65	53.54	563	787
31	*Equus caballus* (Mongolian) (taxid:9796)	As above (same taxid)					
32	*Manis pentadactyla* (taxid:143292)	No significant similarity found					
33	*Felis catus* (taxid:9685)	408 Bt, 354 Ac, 383 Pm, 379 Ta, 271 Va	139	3e-32	23.48	160	983

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
34	*Panthera tigris altaica* (taxid:74533)	No significant similarity found					
35	*Canis lupus familiaris* (taxid:9615)	1177 Bt, 1184 Ac, 1186 Pm, 1144 Ta, 1020 Va	129	5e-30	26.76	145	479
36	*Ursus maritimus* (taxid:29073)	No significant similarity found					
37	*Ailuropoda melanoleuca* (taxid:9646)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	134	2e-31	30.96	263	575
38	*Leptonychotes weddellii* (taxid:9713)	No significant similarity found					
39	*Odobenus rosmarus divergens* (taxid:9708)	No significant similarity found					
40	*Mustela putorius furo* (taxid:9669)	188 Bt, 191 Ac, 182 Pm, 172 Ta, 176 Va	136	1e-31	31.29	289	575
41	*Camelus dromedarius* (taxid:9838)	No significant similarity found					
42	*Camelus ferus* (taxid:419612)	No significant similarity found					
43	*Vicugna pacos* (taxid:30538)	58 Bt, 68 Ac, 69 Pm, 78 Ta, 52 Va	144	2e-34	29.48	199	575
44	*Sus scrofa* (Duroc) (taxid:9823)	1240 Bt, 1259 Ac, 1247 Pm, 1367 Ta, 1062 Va	141	4e-32	28.10	119	575
45	*Sus scrofa* (Tibetan) (taxid:9823)	As above (same taxid)					
46	*Sus scrofa* (Ellegaard Gottingen minipig) (taxid:9823)	As above (same taxid)					
47	*Balaenoptera acutorostrata scammoni* (taxid:310752)	No significant similarity found					
48	*Physeter catodon* (taxid:9755)	No significant similarity found					
49	*Lipotes vexillifer* (taxid:118797)	No significant similarity found					
50	*Tursiops truncatus* (taxid:9739)	1 Bt, 1 Pm, 1 Ta, 1 Va	152	9e-41	32.99	71	355
51	*Orcinus orca* (taxid:9733)	No significant similarity found					
52	*Pantholops hodgsonii* (taxid:59538)	109 Bt, 7 Ac, 40 Pm, 46 Ta, 70 Va	531	6e-173	76.23	43	363

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
53	*Capra hircus* (taxid:9925)	597 Bt, 91 Ac, 277 Pm, 246 Ta, 403 Va	903	0.0	80.10	61	626
54	*Ovis aries* (Texel) (taxid:9940)	3323 Bt, 1587 Ac, 2517 Pm, 2593 Ta, 2987 Va	1428	0.0	93.78	212	966
55	*Ovis aries musimon* (taxid:9938)	687 Bt, 539 Ac, 559 Pm, 543 Ta, 599 Va	531	1e-176	77.43	48	363
56	*Bubalus bubalis* (taxid:89462)	337 Bt, 116 Ac, 208 Pm, 195 Ta, 245 Va	1623	0.0	81.89	23	1048
57	*Bison bison bison* (taxid:43346)	349 Bt, 57 Ac, 132 Pm, 145 Ta, 219 Va	528	5e-171	77.43	48	363
58	*Bos mutus* (taxid:72004)	199 Bt, 17 Ac, 84 Pm, 97 Ta, 139 Va	529	6e-172	77.43	48	363
59	*Bos indicus* (taxid:9915)	64 Bt, 48 Ac, 52 Pm, 58 Ta, 60 Va	191	1e-77	81.95	916	1048
60	*Bos taurus* (taxid:9913)	20204 Bt, 12883 Ac, 18593 Pm, 19836 Ta, 19655 Va	2017	0.0	95.90	1	1048
61	*Ochotona princeps* (taxid:9978)	34 Bt, 36 Ac, 33 Pm, 36 Ta, 34 Va	132	1e-30	25.69	341	797
62	*Oryctolagus cuniculus* (taxid:9986)	211 Bt, 232 Ac, 230 Pm, 269 Ta, 223 Va	132	3e-30	28.35	222	575
63	*Ictidomys tridecemlineatus* (taxid:43179)	119 Bt, 144 Ac, 131 Pm, 151 Ta, 117 Va	143	1e-33	24.59	160	913
64	*Heterocephalus glaber* (taxid:10181)	No significant similarity found					
65	*Fukomys damarensis* (taxid:885580)	No significant similarity found					
66	*Cavia aperea* (taxid:37548)	No significant similarity found					
67	*Cavia porcellus* (taxid:10141)	357 Bt, 364 Ac, 353 Pm, 321 Ta, 292 Va	125	3e-29	27.32	120	479
68	*Chinchilla lanigera* (taxid:34839)	No significant similarity found					
69	*Octodon degus* (taxid:10160)	No significant similarity found					
70	*Dipodomys ordii* (taxid:10020)	No significant similarity found					
71	*Jaculus jaculus* (taxid:51337)	No significant similarity found					
72	*Nannospalax galili* (taxid:1026970)	No significant similarity found					
73	*Mesocricetus auratus* (taxid:10036)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
74	*Cricetulus griseus* (taxid:10029)	12 Bt, 13 Ac, 12 Pm, 14 Ta, 10 Va	110	6e-24	27.27	294	575
75	*Microtus ochrogaster* (taxid:79684)	10 Bt, 10 Ac, 11 Pm, 11 Ta, 8 Va	112	2e-25	25.42	146	479
76	*Peromyscus maniculatus bairdii* (taxid:230844)	No significant similarity found					
77	*Rattus norvegicus* (taxid:10116)	2580 Bt, 2485 Ac, 2425 Pm, 2803 Ta, 2249 Va	315	1e-108	79.29	476	673
78	*Mus musculus* (taxid:10090)	2111 Bt, 2054 Ac, 2120 Pm, 2217 Ta, 1938 Va	856	0.0	92.89	571	1048
79	*Tupaia belangeri* (taxid:37347)	75 Bt, 81 Ac, 85 Pm, 75 Ta, 64 Va	124	1e-28	27.35	218	575
80	*Tupaia chinensis* (taxid:246437)	No significant similarity found					
81	*Galeopterus variegatus* (taxid:482537)	No significant similarity found					
82	*Otolemur garnettii* (taxid:30611)	474 Bt, 513 Ac, 540 Pm, 601 Ta, 431 Va	137	4e-32	30.31	178	575
83	*Microcebus murinus* (taxid:30608)	378 Bt, 395 Ac, 412 Pm, 442 Ta, 327 Va	143	1e-33	28.87	177	576
84	*Tarsius syrichta* (taxid:9478)	No significant similarity found					
85	*Callithrix jacchus* (taxid:9483)	1503 Bt, 1445 Ac, 1511 Pm, 1475 Ta, 1289 Va	136	3e-31	23.20	160	921
86	*Saimiri boliviensis boliviensis* (taxid:39432)	221 Bt, 216 Ac, 215 Pm, 239 Ta, 172 Va	138	7e-32	25.07	158	831
87	*Rhinopithecus roxellana* (taxid:61622)	No significant similarity found					
88	*Nasalis larvatus* (taxid:43780)	No significant similarity found					
89	*Chlorocebus sabaeus* (taxid:60711)	No significant similarity found					
90	*Macaca fascicularis* (taxid:9541)	1198 Bt, 1190 Ac, 1258 Pm, 1721 Ta, 805 Va	143	7e-33	24.16	158	921
91	*Macaca mulatta* (taxid:9544)	1326 Bt, 1310 Ac, 1334 Pm, 1547 Ta, 1155 Va	149	1e-34	24.44	158	921
92	*Papio anubis* (taxid:9555)	652 Bt, 661 Ac, 673 Pm, 764 Ta, 572 Va	148	8e-35	24.78	158	921
93	*Nomascus leucogenys* (taxid:61853)	545 Bt, 541 Ac, 572 Pm, 662 Ta, 477 Va	144	1e-33	24.32	167	971

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
94	*Pongo abelii* (taxid:9601)	1591 Bt, 1556 Ac, 1608 Pm, 1815 Ta, 1384 Va	146	4e-34	24.44	158	921
95	*Gorilla gorilla gorilla* (taxid:9595)	445 Bt, 447 Ac, 475 Pm, 574 Ta, 373 Va	139	2e-32	23.44	160	921
96	*Pan paniscus* (taxid:9597)	7 Bt, 7 Ac, 6 Pm, 8 Ta, 7 Va	132	2e-30	23.36	160	864
97	*Pan troglodytes* (taxid:9598)	1393 Bt, 1381 Ac, 1394 Pm, 1490 Ta, 1207 Va	145	2e-33	23.91	160	921
98	*Homo sapiens* (taxid:9606)	1229 Bt, 1488 Ac, 1344 Pm, 1572 Ta, 1127 Va	1127	0.0	94.28	9	585
SAUROPSIDA							
99	*Apalone spinifera* (taxid:55534)	No significant similarity found					
100	*Pelodiscus sinensis* (taxid:13735)	56 Bt, 31 Ac, 52 Pm, 46 Ta, 49 Va	222	1e-59	29.15	162	730
101	*Chelonia mydas* (taxid:8469)	9 Bt, 7 Ac, 9 Pm, 6 Ta, 6 Va	108	1e-24	29.28	291	584
102	*Chrysemys picta bellii* (taxid:8478)	85 Bt, 62 Ac, 88 Pm, 76 Ta, 82 Va	207	5e-55	41.97	105	369
103	*Struthio camelus australis* (taxid:441894)	1 Bt, 1 Ac, 1 Pm, 3 Ta, 1 Va	103	3e-23	25.43	218	498
104	*Tinamus guttatus* (taxid:94827)	3 Bt, 3 Ac, 3 Pm, 3 Ta, 3 Va	128	3e-31	30.04	480	73
105	*Anas platyrhynchos* (taxid:8839)	7 Bt, 7 Ac, 7 Pm, 5 Ta, 7 Va	86.7	5e-17	26.94	213	429
106	*Lyrurus tetrix tetrix* (taxid:162951)	No significant similarity found					
107	*Gallus gallus* (taxid:9031)	1091 Bt, 1106 Ac, 1069 Pm, 1000 Ta, 1021 Va	121	2e-26	25.61	399	795
108	*Coturnix japonica* (taxid:93934)	12 Bt, 12 Ac, 11 Pm, 12 Ta, 10 Va	71.2	1e-11	29.30	308	459
109	*Meleagris gallopavo* (taxid:9103)	16 Bt, 16 Ac, 16 Pm, 17 Ta, 14 Va	112	2e-24	22.11	132	670
110	*Colinus virginianus* (taxid:9014)	No significant similarity found					
111	*Acanthisitta chloris* (taxid:57068)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	63.2	8e-11	24.89	493	710
112	*Manacus vitellinus* (taxid:328815)	20 Bt, 18 Ac, 22 Pm, 21 Ta, 22 Va	145	3e-37	28.07	110	422
113	*Zonotrichia albicollis* (taxid:44394)	8 Bt, 8 Ac, 8 Pm, 8 Ta, 4 Va	86.3	2e-16	24.17	393	734

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
114	*Geospiza fortis* (taxid:48883)	No significant similarity found					
115	*Serinus canaria* (taxid:9135)	3 Bt, 3 Ac, 3 Pm, 3 Ta, 3 Va	125	6e-30	24.59	437	907
116	*Taeniopygia guttata* (taxid:59729)	3 Bt, 3 Ac, 3 Pm, 3 Ta, 3 Va	75.9	8e-14	24.90	130	360
117	*Ficedula albicollis* (taxid:59894)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	65.5	7e-11	33.98	439	539
118	*Pseudopodoces humilis* (taxid:181119)	No significant similarity found					
119	*Corvus brachyrhynchos* (taxid:85066)	14 Bt, 14 Ac, 13 Pm, 17 Ta, 11 Va	125	3e-30	24.82	101	492
120	*Corvus cornix cornix* (taxid:932674)	15 Bt, 17 Ac, 9 Pm, 10 Ta, 10 Va	107	5e-23	27.95	550	801
121	*Ara macao* (taxid:176014)	No significant similarity found					
122	*Amazona vittata* (taxid:241585)	No significant similarity found					
123	*Melopsittacus undulatus* (taxid:13146)	5 Bt, 5 Ac, 5 Pm, 5 Ta, 5 Va	105	4e-25	33.33	516	690
124	*Nestor notabilis* (taxid:176057)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	93.2	1e-19	25.32	419	723
125	*Falco cherrug* (taxid:345164)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	57.0	3e-08	24.89	202	417
126	*Falco peregrinus* (taxid:8954)	No significant similarity found					
127	*Cariama cristata* (taxid:54380)	5 Bt, 5 Ac, 5 Pm, 4 Ta, 4 Va	83.2	5e-17	27.76	409	641
128	*Merops nubicus* (taxid:57421)	4 Bt, 4 Ac, 4 Pm, 4 Ta, 4 Va	86.3	2e-18	25.57	197	413
129	*Picoides pubescens* (taxid:118200)	6 Bt, 6 Ac, 5 Pm, 6 Ta, 5 Va	134	3e-34	28.02	152	498
130	*Buceros rhinoceros silvestris* (taxid:175836)	2 Bt, 2 Ac, 2 Pm, 1 Ta, 1 Va	58.9	1e-09	21.67	199	429
131	*Apaloderma vittatum* (taxid:57397)	4 Bt, 4 Ac, 4 Pm, 4 Ta, 4 Va	88.2	1e-18	26.42	127	417
132	*Leptosomus discolor* (taxid:188344)	4 Bt, 4 Ac, 4 Pm, 3 Ta, 3 Va	98.6	3e-21	24.50	132	513
133	*Haliaeetus albicilla* (taxid:8969)	5 Bt, 6 Ac, 5 Pm, 4 Ta, 6 Va	81.3	9e-17	29.94	333	509

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
134	*Haliaeetus leucocephalus* (taxid:52644)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	80.9	1e-15	26.87	194	414
135	*Aquila chrysaetos canadensis* (taxid:216574)	3 Bt, 3 Ac, 3 Pm, 1 Ta, 3 Va	87.0	2e-17	25.73	192	491
136	*Cathartes aura* (taxid:43455)	No significant similarity found					
137	*Tyto alba* (taxid:56313)	3 Bt, 3 Ac, 3 Pm, 2 Ta, 2 Va	63.2	1e-10	26.51	557	763
138	*Colius striatus* (taxid:57412)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	74.7	2e-14	22.80	118	430
139	*Charadrius vociferus* (taxid:50402)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	84.3	1e-16	25.77	269	584
140	*Balearica regulorum gibbericeps* (taxid:100784)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	58.9	9e-10	36.46	292	387
141	*Chlamydotis macqueenii* (taxid:187382)	4 Bt, 4 Ac, 4 Pm, 4 Ta, 4 Va	96.7	5e-21	24.93	421	773
142	*Cuculus canorus* (taxid:55661)	7 Bt, 7 Ac, 7 Pm, 7 Ta, 7 Va	118	3e-28	26.95	429	784
143	*Fulmarus glacialis* (taxid:30455)	4 Bt, 5 Ac, 4 Pm, 4 Ta, 4 Va	66.2	3e-11	33.61	195	314
144	*Aptenodytes forsteri* (taxid:9233)	3 Bt, 3 Ac, 4 Pm, 3 Ta, 3 Va	93.6	2e-19	26.42	422	715
145	*Pygoscelis adeliae* (taxid:9238)	6 Bt, 6 Ac, 6 Pm, 6 Ta, 6 Va	97.8	3e-21	27.05	213	488
146	*Phalacrocorax carbo* (taxid:9209)	No significant similarity found					
147	*Pelecanus crispus* (taxid:36300)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	80.5	1e-15	26.46	330	547
148	*Nipponia nippon* (taxid:128390)	11 Bt, 11 Ac, 11 Pm, 15 Ta, 10 Va	91.3	2e-19	26.35	198	488
149	*Egretta garzetta* (taxid:188379)	3 Bt, 3 Ac, 3 Pm, 3 Ta, 3 Va	96.7	5e-21	25.16	194	498
150	*Phaethon lepturus* (taxid:97097)	1 Bt, 1 Ac, 1 Pm, 1 Va	53.1	9e-07	38.10	585	664
151	*Gavia stellata* (taxid:37040)	6 Bt, 6 Ac, 6 Pm, 6 Ta, 6 Va	85.5	1e-17	27.75	557	763
152	*Tauraco erythrolophus* (taxid:121530)	2 Bt, 2 Ac, 1 Pm, 2 Ta, 2 Va	68.6	1e-12	23.51	546	787
153	*Opisthocomus hoazin* (taxid:30419)	6 Bt, 6 Ac, 6 Pm, 6 Ta, 6 Va	103	2e-23	25.14	292	648

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
154	*Columba livia* (taxid:8932)	No significant similarity found					
155	*Pterocles gutturalis* (taxid:240206)	3 Bt, 3 Ac, 3 Pm, 2 Ta, 2 Va	74.3	7e-14	25.55	421	641
156	*Calypte anna* (taxid:9244)	1 Bt, 1 Ac, 1 Pm	53.1	3e-07	23.48	182	417
157	*Chaetura pelagica* (taxid:8897)	4 Bt, 4 Ac, 3 Pm, 1 Ta, 2 Va	65.9	8e-11	27.91	546	715
158	*Caprimulgus carolinensis* (taxid:279965)	4 Bt, 3 Ac, 3 Pm, 3 Ta, 3 Va	100	1e-22	27.59	545	799
159	*Eurypyga helias* (taxid:54383)	No significant similarity found					
160	*Mesitornis unicolor* (taxid:54374)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	129	3e-33	37.70	244	424
161	*Podiceps cristatus* (taxid:345573)	No significant similarity found					
162	*Phoenicopterus ruber ruber* (taxid:9218)	No significant similarity found					
163	*Alligator mississippiensis* (taxid:8496)	57 Bt, 53 Ac, 54 Pm, 51 Ta, 50 Va	126	8e-30	27.16	110	413
164	*Alligator sinensis* (taxid:38654)	26 Bt, 25 Ac, 25 Pm, 25 Ta, 24 Va	123	5e-28	24.56	567	823
165	*Crocodylus porosus* (taxid:8502)	10 Bt, 10 Ac, 10 Pm, 9 Ta, 10 Va	99.4	1e-22	24.60	124	488
166	*Gavialis gangeticus* (taxid:94835)	No significant similarity found					
167	*Pogona vitticeps* (taxid:103695)	158 Bt, 90 Ac, 140 Pm, 124 Ta, 168 Va	1153	0.0	78.81	308	1048
168	*Anolis carolinensis* (taxid:28377)	165 Bt, 103 Ac, 149 Pm, 196 Ta, 172 Va	1382	0.0	68.88	5	999
169	*Vipera berus berus* (taxid:31156)	No significant similarity found					
170	*Crotalus mitchellii pyrrhus* (taxid:384069)	No significant similarity found					
171	*Ophiophagus hannah* (taxid:8665)	No significant similarity found					
172	*Python bivittatus* (taxid:176946)	23 Bt, 12 Ac, 20 Pm, 22 Ta, 24 Va	229	9e-68	48.12	133	363
AMPHIBIA							
173	*Nanorana parkeri* (taxid:125878)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
174	*Xenopus tropicalis* (taxid:8364)	525 Bt, 574 Ac, 606 Pm, 441 Ta, 489 Va	155	4e-39	28.69	144	486
NEOPTERYGII							
175	*Lepisosteus oculatus* (taxid:7918)	12 Bt, 10 Ac, 13 Pm, 8 Ta, 10 Va	404	2e-130	42.57	59	553
176	*Anguilla anguilla* (taxid:7936)	4 Bt, 4 Ac, 4 Pm, 2 Ta, 4 Va	90.9	1e-22	39.42	263	399
177	*Anguilla japonica* (taxid:7937)	1 Bt, 1 Ac, 1 Pm, 2 Ta, 1 Va	96.3	3e-21	24.20	314	609
178	*Danio rerio* (taxid:7955)	1199 Bt, 1214 Ac, 1242 Pm, 1088 Ta, 1176 Va	493	0.0	46.42	405	1006
179	*Astyanax mexicanus* (taxid:7994)	2 Bt, 3 Ac, 2 Pm, 1 Ta, 2 Va	82.4	8e-16	27.11	189	406
180	*Oryzias latipes* (taxid:8090)	2703 Bt, 2523 Ac, 2798 Pm, 2456 Ta, 2461 Va	402	1e-116	29.89	132	980
181	*Poecilia formosa* (taxid:48698)	28 Bt, 30 Ac, 30 Pm, 16 Ta, 25 Va	153	8e-39	30.50	419	822
182	*Xiphophorus maculatus* (taxid:8083)	29 Bt, 29 Ac, 30 Pm, 31 Ta, 30 Va	238	1e-66	25.79	24	776
183	*Fundulus heteroclitus* (taxid:8078)	14 Bt, 16 Ac, 18 Pm, 10 Ta, 9 Va	110	4e-24	27.37	456	728
184	*Takifugu flavidus* (taxid:433684)	No significant similarity found					
185	*Takifugu rubripes* (taxid:31033)	77 Bt, 70 Ac, 74 Pm, 70 Ta, 71 Va	327	2e-92	27.78	106	1032
186	*Tetraodon nigroviridis* (taxid:99883)	47 Bt, 52 Ac, 48 Pm, 45 Ta, 46 Va	268	5e-73	28.16	168	834
187	*Cynoglossus semilaevis* (taxid:244447)	22 Bt, 14 Ac, 19 Pm, 15 Ta, 20 Va	338	2e-96	27.30	109	1032
188	*Haplochromis burtoni* (taxid:8153)	38 Bt, 36 Ac, 43 Pm, 27 Ta, 27 Va	182	8e-51	35.74	82	340
189	*Pundamilia nyererei* (taxid:303518)	15 Bt, 15 Ac, 14 Pm, 12 Ta, 14 Va	243	7e-69	31.34	105	599
190	*Maylandia zebra* (taxid:106582)	22 Bt, 19 Ac, 28 Pm, 26 Ta, 20 Va	216	3e-58	26.75	303	1033
191	*Neolamprologus brichardi* (taxid:32507)	5 Bt, 6 Ac, 6 Pm, 2 Ta, 3 Va	99.0	2e-25	30.13	767	990
192	*Oreochromis niloticus* (taxid:8128)	72 Bt, 67 Ac, 70 Pm, 65 Ta, 59 Va	206	5e-57	30.46	527	994
193	*Sebastes nigrocinctus* (taxid:72089)	No significant similarity found					
194	*Sebastes rubrivinctus* (taxid:72099)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
195	*Gasterosteus aculeatus* (taxid:69293)	63 Bt, 59 Ac, 65 Pm, 60 Ta, 61 Va	296	2e-86	27.43	204	1032
196	*Gadus morhua* (taxid:8049)	1 Bt, 1 Ac, 1 Pm, 1 Va	49.3	1e-06	27.64	561	679
CHONDRICHTHYES							
197	*Callorhinchus milii* (taxid:7868)	33 Bt, 39 Ac, 31 Pm, 58 Ta, 16 Va	108	2e-23	24.50	192	664
198	*Carcharhinus brachyurus* (taxid:671158)	No significant similarity found					
ECDYSOZOA							
199	*Ephemera danica* (taxid:1049336)	No significant similarity found					
200	*Ladona fulva* (taxid:123851)	No significant similarity found					
201	*Pediculus humanus corporis* (taxid:121224)	11 Bt, 11 Ac, 11 Pm, 11 Ta, 11 Va	105	3e-23	25.76	414	795
202	*Frankliniella occidentalis* (taxid:133901)	No significant similarity found					
203	*Diaphorina citri* (taxid:121845)	137 Bt, 79 Ac, 132 Pm, 100 Ta, 132 Va	414	3e-132	38.40	82	716
204	*Pachypsylla venusta* (taxid:38123)	No significant similarity found					
205	*Acyrthosiphon pisum* (taxid:7029)	122 Bt, 72 Ac, 110 Pm, 100 Ta, 116 Va	339	6e-98	27.92	59	913
206	*Nilaparvata lugens* (taxid:108931)	No significant similarity found					
207	*Oncopeltus fasciatus* (taxid:7536)	No significant similarity found					
208	*Rhodnius prolixus* (taxid:13249)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	99.0	1e-23	28.51	302	514
209	*Cimex lectularius* (taxid:79782)	18 Bt, 24 Ac, 25 Pm, 28 Ta, 27 Va	640	0.0	88.83	293	667
210	*Onthophagus taurus* (taxid:166361)	No significant similarity found					
211	*Agrilus planipennis* (taxid:224129)	1 Bt, 1 Pm, 1 Ta, 1 Va	90.1	9e-23	37.61	727	835
212	*Tribolium castaneum* (taxid:7070)	23 Bt, 26 Ac, 26 Pm, 25 Ta, 23 Va	112	1e-24	23.06	186	918
213	*Anoplophora glabripennis* (taxid:217634)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
214	*Leptinotarsa decemlineata* (taxid:7539)	No significant similarity found					
215	*Dendroctonus ponderosae* (taxid:77166)	No significant similarity found					
216	*Mengenilla moldrzyki* (taxid:1155016)	No significant similarity found					
217	*Aedes aegypti* (taxid:7159)	205 Bt, 164 Ac, 204 Pm, 184 Ta, 211 Va	431	3e-128	32.11	69	996
218	*Culex quinquefasciatus* (taxid:7176)	7 Bt, 6 Ac, 8 Pm, 6 Ta, 10 Va	110	2e-25	30.86	199	451
219	*Anopheles albimanus* (taxid:7167)	No significant similarity found					
220	*Anopheles arabiensis* (taxid:7173)	12 Bt, 2 Ac, 12 Pm, 2 Ta, 13 Va	103	2e-23	32.08	87	295
221	*Anopheles atroparvus* (taxid:41427)	No significant similarity found					
222	*Anopheles christyi* (taxid:43041)	No significant similarity found					
223	*Anopheles culicifacies* (taxid:139723)	No significant similarity found					
224	*Anopheles darlingi* (taxid:43151)	No significant similarity found					
225	*Anopheles dirus* (taxid:7168)	No significant similarity found					
226	*Anopheles epiroticus* (taxid:199890)	No significant similarity found					
227	*Anopheles farauti* (taxid:69004)	No significant similarity found					
228	*Anopheles funestus* (taxid:62324)	No significant similarity found					
229	*Anopheles gambiae* (taxid:7165)	41 Bt, 35 Ac, 36 Pm, 29 Ta, 40 Va	130	6e-30	34.05	87	313
230	*Anopheles maculatus* (taxid:74869)	No significant similarity found					
231	*Anopheles melas* (taxid:34690)	No significant similarity found					
232	*Anopheles merus* (taxid:30066)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	97.1	5e-24	29.22	195	413
233	*Anopheles minimus* (taxid:112268)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
234	*Anopheles quadriannulatus* (taxid:34691)	1 Bt, 1 Pm	55.8	9e-10	35.71	719	816
235	*Anopheles sinensis* (taxid:74873)	4 Bt, 4 Ac, 4 Pm, 2 Ta, 4 Va	59.3	2e-12	30.08	272	402
236	*Anopheles stephensi* (taxid:30069)	13 Bt, 10 Ac, 14 Pm, 11 Ta, 9 Va	72.8	1e-14	31.36	302	419
237	*Mayetiola destructor* (taxid:39758)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	96.7	2e-21	24.46	406	769
238	*Lutzomyia longipalpis* (taxid:7200)	3 Bt, 3 Ac, 3 Pm, 1 Ta, 3 Va	55.1	7e-11	37.21	624	706
239	*Phlebotomus papatasi* (taxid:29031)	12 Bt, 12 Ac, 14 Pm, 9 Ta, 13 Va	335	7e-97	27.42	105	1034
240	*Ceratitis capitata* (taxid:7213)	8 Bt, 4 Ac, 7 Pm, 6 Ta, 5 Va	106	3e-23	27.72	390	676
241	*Drosophila albomicans* (taxid:7291)	No significant similarity found					
242	*Drosophila ananassae* (taxid:7217)	35 Bt, 26 Ac, 35 Pm, 26 Ta, 35 Va	353	2e-106	29.49	205	1036
243	*Drosophila biarmipes* (taxid:125945)	4 Ac, 1 Pm, 4 Ta	53.5	6e-07	31.76	399	545
244	*Drosophila bipectinata* (taxid:42026)	2 Bt, 1 Ac, 2 Pm, 2 Ta, 2 Va	147	1e-36	25.46	110	526
245	*Drosophila elegans* (taxid:30023)	3 Bt, 3 Ac, 4 Pm, 1 Ta, 3 Va	60.8	5e-09	31.19	439	544
246	*Drosophila erecta* (taxid:7220)	No significant similarity found					
247	*Drosophila eugracilis* (taxid:29029)	No significant similarity found					
248	*Drosophila ficusphila* (taxid:30025)	3 Bt, 2 Ac, 3 Pm, 3 Ta, 3 Va	153	3e-38	26.87	120	592
249	*Drosophila grimshawi* (taxid:7222)	2 Bt, 2 Ac, 2 Pm	36.6	9e-07	28.57	569	651
250	*Drosophila kikkawai* (taxid:30033)	No significant similarity found					
251	*Drosophila melanogaster* (taxid:7227)	1731 Bt, 1686 Ac, 1659 Pm, 1688 Ta, 1618 Va	119	9e-26	25.31	270	663
252	*Drosophila miranda* (taxid:7229)	39 Bt, 40 Ac, 38 Pm, 18 Ta, 28 Va	103	6e-22	26.25	428	826
253	*Drosophila mojavensis* (taxid:7230)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
254	*Drosophila persimilis* (taxid:7234)	No significant similarity found					
255	*Drosophila pseudoobscura pseudoobscura* (taxid:46245)	No significant similarity found					
256	*Drosophila rhopaloa* (taxid:1041015)	2 Bt, 2 Ac, 3 Pm, 2 Ta, 1 Va	71.6	3e-14	32.76	337	452
257	*Drosophila sechellia* (taxid:7238)	8 Bt, 8 Ac, 8 Pm, 5 Ta, 8 Va	98.2	7e-21	27.19	428	818
258	*Drosophila simulans* (taxid:7240)	8 Bt, 8 Ac, 8 Pm, 7 Ta, 8 Va	107	3e-23	26.47	422	808
259	*Drosophila suzukii* (taxid:28584)	No significant similarity found					
260	*Drosophila takahashii* (taxid:29030)	No significant similarity found					
261	*Drosophila virilis* (taxid:7244)	7 Bt, 7 Ac, 7 Pm, 7 Ta, 7 Va	102	2e-21	26.97	458	724
262	*Drosophila willistoni* (taxid:7260)	6 Bt, 6 Ac, 6 Pm, 4 Ta, 6 Va	384	1e-120	32.80	38	826
263	*Drosophila yakuba* (taxid:7245)	9 Bt, 9 Ac, 9 Pm, 8 Ta, 9 Va	116	4e-27	27.96	440	834
264	*Musca domestica* (taxid:7370)	7 Bt, 10 Ac, 9 Pm, 9 Ta, 8 Va	90.1	6e-18	29.41	565	801
265	*Glossina austeni* (taxid:7395)	No significant similarity found					
266	*Glossina brevipalpis* (taxid:37001)	No significant similarity found					
267	*Glossina fuscipes fuscipes* (taxid:201502)	No significant similarity found					
268	*Glossina morsitans morsitans* (taxid:37546)	2 Bt, 2 Ac, 3 Pm, 2 Ta, 2 Va	90.1	3e-31	32.31	209	338
269	*Glossina pallidipes* (taxid:7398)	No significant similarity found					
270	*Limnephilus lunatus* (taxid:1218281)	No significant similarity found					
271	*Papilio glaucus* (taxid:45779)	No significant similarity found					
272	*Papilio polytes* (taxid:76194)	25 Bt, 17 Ac, 26 Pm, 20 Ta, 23 Va	363	2e-111	35.76	110	684
273	*Papilio xuthus* (taxid:66420)	52 Bt, 41 Ac, 54 Pm, 58 Ta, 51 Va	270	1e-77	29.86	295	957

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
274	*Heliconius melpomene melpomene* (taxid:171917)	No significant similarity found					
275	*Melitaea cinxia* (taxid:113334)	No significant similarity found					
276	*Danaus plexippus* (taxid:13037)	No significant similarity found					
277	*Bombyx mori* (taxid:7091)	515 Bt, 366 Ac, 485 Pm, 476 Ta, 500 Va	746	0.0	42.13	91	1032
278	*Manduca sexta* (taxid:7130)	3 Bt, 3 Ac, 5 Pm, 5 Ta, 3 Va	140	2e-36	31.36	272	507
279	*Plutella xylostella* (taxid:51655)	146 Bt, 134 Ac, 152 Pm, 124 Ta, 152 Va	582	0.0	36.83	110	980
280	*Athalia rosae* (taxid:37344)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	107	2e-23	24.54	431	789
281	*Cephus cinctus* (taxid:211228)	18 Bt, 21 Ac, 21 Pm, 18 Ta, 18 Va	127	2e-30	27.33	209	520
282	*Orussus abietinus* (taxid:222816)	3 Bt, 2 Ac, 2 Pm, 3 Ta, 2 Va	147	1e-37	30.94	174	478
283	*Ceratosolen solmsi marchali* (taxid:326594)	No significant similarity found					
284	*Nasonia giraulti* (taxid:7426)	No significant similarity found					
285	*Nasonia longicornis* (taxid:7427)	No significant similarity found					
286	*Nasonia vitripennis* (taxid:7425)	49 Bt, 46 Ac, 52 Pm, 44 Ta, 49 Va	397	1e-116	30.16	54	996
287	*Copidosoma floridanum* (taxid:29053)	23 Bt, 25 Ac, 26 Pm, 21 Ta, 21 Va	170	3e-46	34.17	196	507
288	*Trichogramma pretiosum* (taxid:7493)	26 Bt, 26 Ac, 24 Pm, 25 Ta, 23 Va	132	5e-32	30.43	211	482
289	*Microplitis demolitor* (taxid:69319)	17 Bt, 18 Ac, 18 Pm, 19 Ta, 17 Va	169	2e-43	27.65	420	794
290	*Megachile rotundata* (taxid:143995)	11 Bt, 11 Ac, 13 Pm, 10 Ta, 8 Va	117	2e-28	27.76	181	483
291	*Apis dorsata* (taxid:7462)	No significant similarity found					
292	*Apis florea* (taxid:7463)	No significant similarity found					
293	*Apis mellifera* (taxid:7460)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
294	*Bombus impatiens* (taxid:132113)	4 Bt, 5 Ac, 4 Pm, 5 Ta, 4 Va	118	1e-27	31.54	547	785
295	*Bombus terrestris* (taxid:30195)	4 Bt, 4 Ac, 5 Pm, 4 Ta, 3 Va	77.0	4e-15	25.51	216	455
296	*Linepithema humile* (taxid:83485)	19 Bt, 9 Ac, 11 Pm, 15 Ta, 13 Va	167	3e-45	33.55	69	356
297	*Camponotus floridanus* (taxid:104421)	13 Bt, 10 Ac, 11 Pm, 8 Ta, 11 Va	187	1e-49	27.37	94	715
298	*Acromyrmex echinatior* (taxid:103372)	3 Bt, 6 Ac, 5 Pm, 3 Ta, 2 Va	75.9	5e-14	51.47	307	374
299	*Atta cephalotes* (taxid:12957)	2 Bt, 2 Ac, 2 Pm, 1 Ta, 2 Va	79.3	1e-17	39.20	321	441
300	*Solenopsis invicta* (taxid:13686)	49 Bt, 41 Ac, 50 Pm, 42 Ta, 46 Va	603	0.0	37.05	67	998
301	*Pogonomyrmex barbatus* (taxid:144034)	6 Bt, 8 Ac, 8 Pm, 6 Ta, 7 Va	96.3	7e-20	25.18	415	92
302	*Harpegnathos saltator* (taxid:610380)	23 Bt, 19 Ac, 22 Pm, 12 Ta, 17 Va	117	8e-28	25.95	419	751
303	*Cerapachys biroi* (taxid:443821)	14 Bt, 14 Ac, 10 Pm, 13 Ta, 13 Va	164	9e-43	32.01	197	533
304	*Blattella germanica* (taxid:6973)	3 Bt, 3 Ac, 4 Pm, 4 Ta, 4 Va	442	2e-140	30.78	110	1045
305	*Zootermopsis nevadensis* (taxid:136037)	No significant similarity found					
306	*Daphnia pulex* (taxid:6669)	6 Bt, 7 Ac, 6 Pm, 6 Ta, 5 Va	106	2e-25	24.38	437	823
307	*Eurytemora affinis* (taxid:88015)	No significant similarity found					
308	*Hyalella azteca* (taxid:294128)	27 Bt, 20 Ac, 28 Pm, 24 Ta, 26 Va	216	2e-60	29.46	105	576
309	*Strigamia maritima* (taxid:126957)	No significant similarity found					
310	*Stegodyphus mimosarum* (taxid:407821)	No significant similarity found					
311	*Latrodectus hesperus* (taxid:256737)	2 Bt, 1 Ac, 3 Pm, 3 Ta, 2 Va	58.5	7e-19	23.76	433	694
312	*Parasteatoda tepidariorum* (taxid:114398)	27 Bt, 19 Ac, 27 Pm, 23 Ta, 27 Va	253	1e-72	32.32	301	859

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
313	*Tetranychus urticae* (taxid:32264)	4 Bt, 4 Ac, 4 Pm, 1 Ta, 4 Va	103	3e-22	24.88	428	815
314	*Dermatophagoides farinae* (taxid:6954)	No significant similarity found					
315	*Sarcoptes scabiei type canis* (taxid:52283)	No significant similarity found					
316	*Achipteria coleoptrata* (taxid:229769)	No significant similarity found					
317	*Hypochthonius rufulus* (taxid:66581)	No significant similarity found					
318	*Platynothrus peltifer* (taxid:128015)	No significant similarity found					
319	*Steganacarus magnus* (taxid:52000)	No significant similarity found					
320	*Ixodes ricinus* (taxid:34613)	No significant similarity found					
321	*Ixodes scapularis* (taxid:6945)	146 Bt, 153 Ac, 149 Pm, 126 Ta, 132 Va	134	8e-32	28.66	413	734
322	*Rhipicephalus microplus* (taxid:6941)	38 Bt, 35 Ac, 35 Pm, 32 Ta, 30 Va	350	1e-101	29.38	110	984
323	*Metaseiulus occidentalis* (taxid:34638)	16 Bt, 17 Ac, 17 Pm, 10 Ta, 16 Va	132	2e-32	27.44	182	485
324	*Varroa destructor* (taxid:109461)	No significant similarity found					
325	*Centruroides exilicauda* (taxid:6879)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	54.7	6e-12	28.97	223	364
326	*Mesobuthus martensii* (taxid:34649)	No significant similarity found					
327	*Limulus polyphemus* (taxid:6850)	44 Bt, 27 Ac, 45 Pm, 38 Ta, 46 Va	184	7e-52	36.98	71	328
328	*Trichinella spiralis* (taxid:6334)	11 Bt, 12 Ac, 12 Pm, 13 Ta, 11 Va	116	2e-28	28.66	307	603
329	*Ascaris suum* (taxid:6253)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	93.2	2e-21	29.22	236	478
330	*Elaeophora elaphi* (taxid:1147741)	No significant similarity found					
331	*Onchocerca volvulus* (taxid:6282)	No significant similarity found					
332	*Steinernema monticolum* (taxid:90984)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
333	*Panagrellus redivivus* (taxid:6233)	No significant similarity found					
334	*Haemonchus contortus* (taxid:6289)	2 Bt, 1 Ac, 2 Pm, 2 Ta, 2 Va	88.2	5e-19	26.10	505	778
335	*Necator americanus* (taxid:51031)	173 Bt, 80 Ac, 164 Pm, 78 Ta, 152 Va	178	1e-50	32.71	149	514
336	*Heterorhabditis bacteriophora* (taxid:37862)	No significant similarity found					
337	*Caenorhabditis angaria* (taxid:860376)	No significant similarity found					
338	*Caenorhabditis brenneri* (taxid:135651)	No significant similarity found					
339	*Caenorhabditis briggsae* (taxid:6238)	13 Bt, 13 Ac, 13 Pm, 8 Ta, 11 Va	94.7	1e-19	24.29	422	799
340	*Caenorhabditis elegans* (taxid:6239)	239 Bt, 243 Ac, 249 Pm, 246 Ta, 264 Va	300	2e-83	25.89	43	998
341	*Caenorhabditis japonica* (taxid:281687)	6 Bt, 7 Ac, 8 Pm, 6 Ta, 6 Va	260	2e-72	26.90	108	980
342	*Caenorhabditis sp. 11 MAF-2010* (taxid:1561998)	No significant similarity found					
343	*Priapulus caudatus* (taxid:37621)	20 Bt, 15 Ac, 22 Pm, 18 Ta, 22 Va	311	1e-92	37.26	519	980
ROTIFERA							
344	*Adineta vaga* (taxid:104782)	4 Bt, 18 Ac, 4 Pm, 14 Ta, 17 Va	100	3e-23	27.02	185	521
PLATYHELMINTHES							
345	*Schistosoma curassoni* (taxid:6186)	1361 Bt, 844 Ac, 1301 Pm, 1079 Ta, 1334 Va	341	9e-97	31.38	110	856
346	*Schistosoma haematobium* (taxid:6185)	49 Bt, 22 Ac, 34 Pm, 20 Ta, 49 Va	120	1e-27	40.52	503	655
347	*Schistosoma japonicum* (taxid:6182)	191 Bt, 114 Ac, 196 Pm, 133 Ta, 177 Va	441	1e-137	32.39	110	984
348	*Schistosoma mansoni* (taxid:6183)	15600 Bt, 9579 Ac, 15032 Pm, 12415 Ta, 15767 Va	509	2e-153	33.98	103	1042
349	*Schistosoma margrebowiei* (taxid:48269)	2361 Bt, 1302 Ac, 2223 Pm, 1901 Ta, 2282 Va	461	3e-137	32.41	110	1039

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
350	*Schistosoma mattheei* (taxid:31246)	1146 Bt, 791 Ac, 1050 Pm, 919 Ta, 1127 Va	360	3e-104	29.79	110	982
351	*Schistosoma rodhaini* (taxid:6188)	1394 Bt, 888 Ac, 1306 Pm, 1122 Ta, 1364 Va	343	2e-97	28.95	112	950
352	*Clonorchis sinensis* (taxid:79923)	No significant similarity found					
353	*Echinococcus granulosus* (taxid:6210)	No significant similarity found					
354	*Echinococcus multilocularis* (taxid:6211)	No significant similarity found					
355	*Hymenolepis microstoma* (taxid:85433)	No significant similarity found					
ANNELIDA							
356	*Capitella teleta* (taxid:283909)	No significant similarity found					
357	*Helobdella robusta* (taxid:6412)	233 Bt, 43 Ac, 188 Pm, 67 Ta, 259 Va	265	1e-72	28.72	103	836
MOLLUSCA							
358	*Crassostrea gigas* (taxid:29159)	70 Bt, 48 Ac, 68 Pm, 61 Ta, 56 Va	224	2e-64	35.78	110	439
359	*Lottia gigantea* (taxid:225164)	10 Bt, 9 Ac, 12 Pm, 10 Ta, 8 Va	60.1	1e-10	28.89	669	797
360	*Aplysia californica* (taxid:6500)	174 Bt, 100 Ac, 167 Pm, 100 Ta, 164 Va	265	1e-76	32.42	465	973
361	*Biomphalaria glabrata* (taxid:6526)	191 Bt, 96 Ac, 184 Pm, 127 Ta, 184 Va	365	4e-105	26.70	110	980
CNIDARIA							
362	*Nematostella vectensis* (taxid:45351)	37 Bt, 20 Ac, 39 Pm, 32 Ta, 36 Va	192	7e-55	41.13	110	351
363	*Hydra vulgaris* (taxid:6087)	115 Bt, 119 Ac, 115 Pm, 125 Ta, 110 Va	112	2e-25	29.05	244	584
TENTACULATA							
364	*Mnemiopsis leidyi* (taxid:27923)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	82.8	1e-17	27.43	278	509
PLACOZOA							
365	*Trichoplax adhaerens* (taxid:10228)	No significant similarity found					
PORIFERA							

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
366	*Amphimedon queenslandica* (taxid:400682)	5 Bt, 6 Ac, 4 Pm, 6 Ta, 5 Va	102	9e-23	22.72	119	608
VIRIDIPLANTAE							
367	*Micromonas pusilla CCMP1545* (taxid:564608)	No significant similarity found					
368	*Micromonas sp. RCC299* (taxid:296587)	No significant similarity found					
369	*Ostreococcus lucimarinus CCE9901* (taxid:436017)	No significant similarity found					
370	*Ostreococcus tauri* (taxid:70448)	5 Pm	48.5	7e-06	25.99	535	814
371	*Chlamydomonas reinhardtii* (taxid:3055)	3 Bt, 3 Ac, 3 Pm, 4 Ta, 3 Va	94.0	2e-19	23.90	315	584
372	*Volvox carteri f. nagariensis* (taxid:3068)	47 Bt, 51 Ac, 52 Pm, 48 Ta, 45 Va	153	1e-37	27.44	436	819
373	*Chlorella variabilis* (taxid:554065)	3 Bt, 3 Ac, 10 Pm, 8 Ta, 3 Va	127	6e-30	28.35	432	801
374	*Auxenochlorella protothecoides* (taxid:3075)	No significant similarity found					
375	*Helicosporidium sp. ATCC 50920* (taxid:1291522)	No significant similarity found					
376	*Coccomyxa subellipsoidea C-169* (taxid:574566)	3 Bt, 4 Ac, 5 Pm, 3 Ta, 2 Va	120	6e-29	28.14	205	453
377	*Klebsormidium flaccidum* (taxid:3175)	1 Bt, 1 Ac, 1 Pm	52.8	7e-09	22.81	255	481
378	*Physcomitrella patens* (taxid:3218)	2 Bt	52.0	9e-08	30.00	556	645
379	*Selaginella moellendorffii* (taxid:88036)	1 Ta	47.8	8e-06	29.29	299	396
380	*Pinus taeda* (taxid:3352)	56 Bt, 59 Ac, 60 Pm, 58 Ta, 40 Va	135	4e-32	28.24	426	794
381	*Amborella trichopoda* (taxid:13333)	17 Bt, 20 Ac, 19 Pm, 17 Ta, 7 Va	86.7	8e-17	23.25	501	836

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
382	*Spirodela polyrhiza* (taxid:29656)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	94.0	3e-21	25.69	174	518
383	*Phoenix dactylifera* (taxid:42345)	22 Bt, 21 Ac, 22 Pm, 14 Ta, 16 Va	102	2e-21	26.14	418	711
384	*Elaeis oleifera* (taxid:80265)	No significant similarity found					
385	*Ensete ventricosum* (taxid:4639)	No significant similarity found					
386	*Musa acuminata subsp. malaccensis* (taxid:214687)	No significant similarity found					
387	*Sorghum bicolor* (taxid:4558)	71 Bt, 76 Ac, 72 Pm, 63 Ta, 56 Va	104	2e-23	26.39	201	522
388	*Zea mays* (taxid:4577)	981 Bt, 739 Ac, 854 Pm, 861 Ta, 829 Va	342	7e-96	28.50	110	1037
389	*Setaria italica* (taxid:4555)	12 Bt, 11 Ac, 12 Pm, 10 Ta, 10 Va	98.2	3e-20	33.33	179	350
390	*Brachypodium distachyon* (taxid:15368)	33 Bt, 24 Ac, 20 Pm, 24 Ta, 17 Va	254	2e-71	29.97	110	699
391	*Leersia perrieri* (taxid:77586)	4 Bt, 4 Ac, 4 Pm, 2 Ta, 4 Va	83.6	5e-19	26.10	424	709
392	*Oryza barthii* (taxid:65489)	No significant similarity found					
393	*Oryza brachyantha* (taxid:4533)	4 Bt, 5 Ac, 4 Pm, 2 Ta, 3 Va	81.6	9e-16	23.37	170	485
394	*Oryza glumipatula* (taxid:40148)	No significant similarity found					
395	*Oryza longistaminata* (taxid:4528)	No significant similarity found					
396	*Oryza meridionalis* (taxid:40149)	No significant similarity found					
397	*Oryza nivara* (taxid:4536)	4 Bt, 4 Ac, 4 Pm, 2 Ta, 3 Va	105	4e-24	27.33	193	485
398	*Oryza punctata* (taxid:4537)	26 Bt, 32 Ac, 34 Pm, 15 Ta, 21 Va	109	1e-24	27.04	193	485
399	*Oryza sativa Japonica Group* (taxid:39947)	1057 Bt, 1097 Ac, 1082 Pm, 984 Ta, 890 Va	119	3e-26	28.16	179	488
400	*Zizania latifolia* (taxid:58934)	No significant similarity found					
401	*Aegilops tauschii* (taxid:37682)	102 Bt, 89 Ac, 91 Pm, 77 Ta, 75 Va	270	3e-74	30.19	110	754

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
402	*Triticum urartu* (taxid:4572)	6 Bt, 6 Ac, 6 Pm, 5 Ta, 4 Va	85.1	1e-17	22.71	426	818
403	*Nelumbo nucifera* (taxid:4432)	8 Bt, 9 Ac, 9 Pm, 9 Ta, 7 Va	84.7	5e-16	25.79	506	800
404	*Lupinus angustifolius* (taxid:3871)	4 Bt, 4 Ac, 4 Pm, 3 Ta, 4 Va	99.0	8e-22	24.85	417	740
405	*Phaseolus vulgaris* (taxid:3885)	29 Bt, 32 Ac, 32 Pm, 43 Ta, 30 Va	110	5e-24	23.32	293	822
406	*Cajanus cajan* (taxid:3821)	No significant similarity found					
407	*Vigna angularis var. angularis* (taxid:157739)	1 Bt, 1 Ac, 1 Pm, 1 Va	67.4	2e-10	30.60	242	415
408	*Vigna radiata var. radiata* (taxid:3916)	No significant similarity found					
409	*Glycine max* (taxid:3847)	318 Bt, 291 Ac, 312 Pm, 269 Ta, 297 Va	294	3e-81	27.77	110	1037
410	*Glycine soja* (taxid:3848)	3 Bt, 3 Ac, 2 Pm, 2 Ta, 3 Va	85.9	5e-18	24.62	416	734
411	*Cicer arietinum* (taxid:3827)	9 Ac, 8 Pm	50.1	3e-06	28.10	304	451
412	*Medicago truncatula* (taxid:3880)	664 Bt, 634 Ac, 676 Pm, 646 Ta, 490 Va	192	2e-49	30.36	294	760
413	*Trifolium pratense* (taxid:57577)	No significant similarity found					
414	*Lotus japonicus* (taxid:34305)	630 Bt, 627 Ac, 634 Pm, 618 Ta, 497 Va	128	5e-29	26.09	420	789
415	*Malus x domestica* (taxid:3750)	56 Bt, 48 Ac, 55 Pm, 57 Ta, 44 Va	218	9e-57	27.54	110	696
416	*Pyrus x bretschneideri* (taxid:225117)	11 Bt, 10 Ac, 11 Pm, 12 Ta, 10 Va	94.0	4e-19	27.76	285	577
417	*Prunus mume* (taxid:102107)	5 Bt, 5 Ac, 5 Pm, 3 Ta, 1 Va	85.9	1e-16	23.52	311	711
418	*Prunus persica* (taxid:3760)	28 Bt, 29 Ac, 30 Pm, 36 Ta, 9 Va	91.7	9e-19	24.44	291	710
419	*Fragaria iinumae* (taxid:64939)	No significant similarity found					
420	*Fragaria nubicola* (taxid:60188)	No significant similarity found					
421	*Fragaria orientalis* (taxid:101013)	No significant similarity found					
422	*Fragaria vesca subsp. vesca* (taxid:101020)	12 Bt, 11 Ac, 14 Pm, 9 Ta, 8 Va	92.4	3e-19	28.39	261	484

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
423	*Fragaria x ananassa* (taxid:3747)	No significant similarity found					
424	*Morus notabilis* (taxid:981085)	7 Bt, 5 Ac, 9 Pm, 9 Ta, 9 Va	87.8	5e-18	27.47	184	410
425	*Cannabis sativa* (taxid:3483)	2 Bt, 2 Ac, 2 Pm, 2 Ta	85.1	7e-18	23.91	242	679
426	*Castanea mollissima* (taxid:60419)	No significant similarity found					
427	*Betula nana* (taxid:216990)	No significant similarity found					
428	*Cucumis melo* (taxid:3656)	132 Bt, 137 Ac, 146 Pm, 143 Ta, 36 Va	115	6e-27	29.17	595	829
429	*Cucumis sativus* (taxid:3659)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	73.9	5e-13	22.95	416	716
430	*Citrullus lanatus* (taxid:3654)	No significant similarity found					
431	*Lagenaria siceraria* (taxid:3668)	No significant similarity found					
432	*Populus euphratica* (taxid:75702)	9 Bt, 9 Ac, 9 Pm, 9 Ta, 9 Va	107	7e-23	26.38	416	738
433	*Populus trichocarpa* (taxid:3694)	136 Bt, 138 Ac, 138 Pm, 130 Ta, 107 Va	115	4e-25	23.46	392	795
434	*Jatropha curcas* (taxid:180498)	18 Bt, 19 Ac, 17 Pm, 20 Ta, 14 Va	95.9	3e-20	25.84	132	510
435	*Manihot esculenta subsp. flabellifolia* (taxid:478097)	No significant similarity found					
436	*Ricinus communis* (taxid:3988)	No significant similarity found					
437	*Linum usitatissimum* (taxid:4006)	1 Bt, 1 Ac, 1 Pm, 1 Ta	58.2	3e-09	26.90	347	513
438	*Eucalyptus camaldulensis* (taxid:34316)	No significant similarity found					
439	*Eucalyptus grandis* (taxid:71139)	47 Bt, 54 Ac, 51 Pm, 47 Ta, 36 Va	108	1e-23	21.88	145	813
440	*Carica papaya* (taxid:3649)	5 Bt, 9 Ac, 9 Pm, 13 Ta, 6 Va	79.7	6e-15	24.44	251	582
441	*Arabidopsis halleri subsp. gemmifera* (taxid:63677)	No significant similarity found					

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
442	*Arabidopsis lyrata subsp. lyrata* (taxid:81972)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	89.4	8e-18	21.77	218	676
443	*Arabidopsis thaliana* (taxid:3702)	389 Bt, 406 Ac, 400 Pm, 396 Ta, 265 Va	144	4e-35	30.39	209	483
444	*Camelina sativa* (taxid:90675)	108 Bt, 115 Ac, 113 Pm, 134 Ta, 68 Va	102	8e-22	24.64	146	522
445	*Capsella rubella* (taxid:81985)	8 Bt, 8 Ac, 8 Pm, 6 Ta, 3 Va	83.6	2e-16	24.53	139	489
446	*Brassica napus* (taxid:3708)	319 Bt, 345 Ac, 332 Pm, 342 Ta, 238 Va	114	7e-25	27.61	294	581
447	*Brassica oleracea var. oleracea* (taxid:109376)	126 Bt, 133 Ac, 133 Pm, 126 Ta, 83 Va	111	2e-24	27.18	247	599
448	*Brassica rapa* (taxid:3711)	655 Bt, 662 Ac, 657 Pm, 658 Ta, 545 Va	121	1e-26	23.94	112	835
449	*Raphanus raphanistrum subsp. raphanistrum* (taxid:109997)	No significant similarity found					
450	*Raphanus sativus* (taxid:3726)	92 Bt, 90 Ac, 87 Pm, 88 Ta, 65 Va	114	2e-25	24.57	294	835
451	*Aethionema arabicum* (taxid:228871)	No significant similarity found					
452	*Arabis alpina* (taxid:50452)	2 Bt, 2 Ac, 2 Ta, 1 Va	52.0	4e-10	21.28	146	366
453	*Eutrema parvulum* (taxid:98039)	No significant similarity found					
454	*Eutrema salsugineum* (taxid:72664)	No significant similarity found					
455	*Sisymbrium irio* (taxid:3730)	No significant similarity found					
456	*Leavenworthia alabamica* (taxid:310722)	No significant similarity found					
457	*Tarenaya hassleriana* (taxid:28532)	34 Bt, 29 Ac, 39 Pm, 43 Ta, 32 Va	95.9	1e-19	23.38	244	665
458	*Gossypium arboreum* (taxid:29729)	24 Bt, 28 Ac, 19 Pm, 26 Ta, 16 Va	92.4	4e-19	25.54	208	478
459	*Gossypium raimondii* (taxid:29730)	50 Bt, 54 Ac, 53 Pm, 63 Ta, 28 Va	90.1	1e-17	23.97	331	677
460	*Theobroma cacao* (taxid:3641)	271 Bt, 288 Ac, 309 Pm, 227 Ta, 197 Va	117	4e-25	22.56	110	935

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
461	*Aquilaria agallochum* (taxid:1472648)	No significant similarity found					
462	*Azadirachta indica* (taxid:124943)	No significant similarity found					
463	*Citrus clementina* (taxid:85681)	No significant similarity found					
464	*Citrus sinensis* (taxid:2711)	13 Bt, 13 Ac, 13 Pm, 17 Ta, 9 Va	91.3	2e-18	24.65	247	581
465	*Vitis vinifera* (taxid:29760)	570 Bt, 579 Ac, 576 Pm, 604 Ta, 554 Va	114	2e-24	27.89	262	581
466	*Amaranthus hypochondriacus* (taxid:28502)	1 Bt, 1 Ac	43.9	1e-06	24.47	342	435
467	*Amaranthus tuberculatus* (taxid:277990)	No significant similarity found					
468	*Beta vulgaris subsp. vulgaris* (taxid:3555)	227 Bt, 230 Ac, 231 Pm, 224 Ta, 163 Va	185	4e-51	29.27	310	781
469	*Spinacia oleracea* (taxid:3562)	No significant similarity found					
470	*Dianthus caryophyllus* (taxid:3570)	No significant similarity found					
471	*Actinidia chinensis* (taxid:3625)	No significant similarity found					
472	*Vaccinium macrocarpon* (taxid:13750)	2 Bt, 3 Ac, 2 Pm, 2 Va	46.2	2e-06	26.14	341	484
473	*Diospyros lotus* (taxid:55363)	No significant similarity found					
474	*Primula veris* (taxid:170927)	No significant similarity found					
475	*Solanum arcanum* (taxid:376710)	No significant similarity found					
476	*Solanum habrochaites* (taxid:62890)	No significant similarity found					
477	*Solanum lycopersicum* (taxid:4081)	554 Bt, 535 Ac, 557 Pm, 580 Ta, 233 Va	105	3e-21	27.03	424	789
478	*Solanum melongena* (taxid:4111)	2 Bt, 2 Ac, 2 Pm, 2 Ta, 2 Va	110	5e-26	25.75	433	784
479	*Solanum pennellii* (taxid:28526)	446 Bt, 474 Ac, 442 Pm, 458 Ta, 205 Va	110	9e-23	26.25	424	789

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
480	*Solanum pimpinellifolium* (taxid:4084)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	59.7	2e-10	31.25	358	543
481	*Solanum tuberosum* (taxid:4113)	412 Bt, 401 Ac, 411 Pm, 421 Ta, 316 Va	303	6e-84	26.50	110	981
482	*Capsicum annuum* (taxid:4072)	208 Bt, 62 Ac, 183 Pm, 79 Ta, 174 Va	278	3e-76	26.82	201	1036
483	*Nicotiana sylvestris* (taxid:4096)	307 Bt, 148 Ac, 290 Pm, 179 Ta, 242 Va	189	2e-49	25.12	168	771
484	*Nicotiana tomentosiformis* (taxid:4098)	230 Bt, 131 Ac, 223 Pm, 144 Ta, 188 Va	152	2e-37	26.70	168	674
485	*Fraxinus excelsior* (taxid:38873)	No significant similarity found					
486	*Penstemon centranthifolius* (taxid:69924)	No significant similarity found					
487	*Penstemon grinnellii* (taxid:388155)	No significant similarity found					
488	*Sesamum indicum* (taxid:4182)	17 Bt, 21 Ac, 9 Pm, 16 Ta, 9 Va	89.0	4e-18	23.21	153	519
489	*Genlisea aurea* (taxid:192259)	No significant similarity found					
490	*Mimulus guttatus* (taxid:4155)	35 Bt, 35 Ac, 35 Pm, 37 Ta, 13 Va	99.0	1e-20	23.65	500	945
491	*Conyza canadensis* (taxid:72917)	No significant similarity found					
ECHINOIDEA							
492	*Lytechinus variegatus* (taxid:7654)	10 Bt, 10 Ac, 10 Pm, 11 Ta, 10 Va	135	6e-33	26.78	484	937
493	*Strongylocentrotus purpuratus* (taxid:7668)	2340 Bt, 1717 Ac, 2132 Pm, 2165 Ta, 2227 Va	819	0.0	92.41	484	931
ASTEROIDEA							
494	*Patiria miniata* (taxid:46514)	No significant similarity found					
ENTEROPNEUSTA							
495	*Saccoglossus kowalevskii* (taxid:10224)	89 Bt, 87 Ac, 104 Pm, 93 Ta, 91 Va	189	1e-51	29.93	115	507
TUNICATA							
496	*Ciona intestinalis* (taxid:7719)	9 Bt, 8 Ac, 9 Pm, 9 Ta, 7 Va	172	1e-47	36.36	69	298

			Top hit				
No	**Species and taxon identifier**	**# hits from each query** (top hit red)	**bitscore**	**evalue**	**pident**	**qstart**	**qend**
497	*Ciona savignyi* (taxid:51511)	1 Bt, 1 Ac, 1 Pm, 1 Ta, 1 Va	121	1e-28	25.11	481	913
498	*Botryllus schlosseri* (taxid:30301)	1 Bt, 3 Ac, 3 Pm, 3 Ta, 1 Va	90.1	1e-19	29.20	433	680
499	*Oikopleura dioica* (taxid:34765)	No significant similarity found					
LEPTOCARDII							
500	*Branchiostoma floridae* (taxid:7739)	47 Bt, 36 Ac, 47 Pm, 45 Ta, 45 Va	414	5e-122	31.75	195	980
CEPHALASPIDOMORPHI							
501	*Lethenteron camtschaticum* (taxid:980415)	6 Bt, 7 Ac, 10 Pm, 8 Ta, 7 Va	202	7e-54	28.26	205	768
502	*Petromyzon marinus* (taxid:7757)	22 Bt, 19 Ac, 25 Pm, 19 Ta, 19 Va	325	3e-96	29.51	110	830
SARCOPTERYGII							
503	*Latimeria chalumnae* (taxid:7897)	96 Bt, 67 Ac, 92 Pm, 85 Ta, 89 Va	289	1e-81	27.79	190	970

Figure 1: Distiguishing RTEs from BovBs

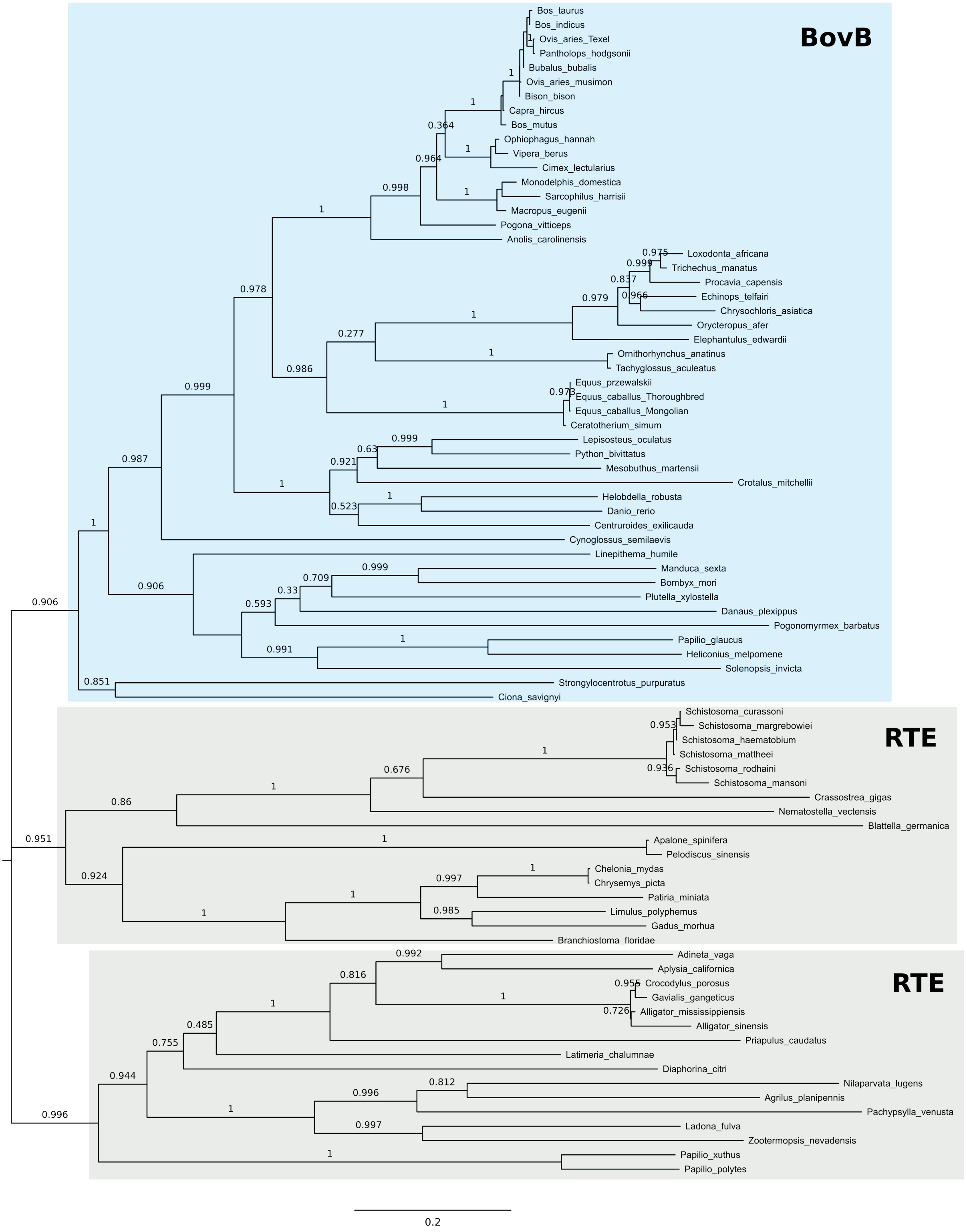


Figure C.1: **BovB vs RTE clades.** Maximum likelihood tree inferred from 82 full-length consensus nucleotide sequences. For each species, USEARCH was used to extract sequences between 2.4-4kb in length; UCLUST was used to cluster and generate a consensus of the dominant clusters; MUSCLE was used to align consensus sequences; Gblocks was used to select conserved blocks; and FastTree was used to infer a phylogeny. Local support values are shown. The BovB group is clearly distinct from the two RTE groups.

Table 3: BovBs in the genome, found using LASTZ

Table C.3: **Presence of BovB:** Shows the length distirbution of hits for the 60 species with confirmed BovBs. These species contain evidence of BovB elements based on query-driven iterative similarity searches with LASTZ, and separated into the BovB (not RTE) clade after clustering (see Figure 1 above). Any hits found had to satisfy a 'reciprocal best hit' check: they were screened with CENSOR against the Repbase library of known repeats, and kept only if the best hit was a BovB element (not some other repeat). Overlapping hits were merged to produce a non-redundant set of L1s for each genome. The Notes column highlights unusual observations.

Species	# BovB hits	Length distribution (bp)	Notes
MAMMALIA			
Tachyglossus aculeatus	1913	min 41, median 427, max 3794	
Ornithorhynchus anatinus	2063	min 37, median 446, max 3849	
Monodelphis domestica	703	min 48, median 282, max 3408	
Macropus eugenii	978	min 39, median 346, max 3497	
Sarcophilus harrisii	404	min 50, median 482, max 3788	
Chrysochloris asiatica	82797	min 43, median 1463, max 6526	
Echinops telfairi	2663	min 39, median 176, max 3603	
Orycteropus afer afer	130793	min 42, median 1805, max 8015	Lots of chimeric/nested BovB elements.
Elephantulus edwardii	59957	min 42, median 1407, max 5564	
Trichechus manatus latirostris	74526	min 37, median 1547, max 6219	
Procavia capensis	23044	min 39, median 1176, max 5482	
Loxodonta africana	110767	min 34, median 1500, max 6505	
Pteropus alecto	87	min 84, median 397, max 3182	
Pteropus vampyrus	66	min 176, median 355, max 3200	
Eidolon helvum	71	min 87, median 509, max 3195	
Megaderma lyra	53	min 82, median 316, max 3092	
Rhinolophus ferrumequinum	83	min 157, median 540, max 3469	
Pteronotus parnellii	60	min 74, median 314, max 3363	
Eptesicus fuscus	63	min 61, median 411, max 3554	
Myotis brandtii	88	min 87, median 343, max 3373	The BovBs in bats seem very divergent from each other.
Myotis davidii	64	min 58, median 438, max 3615	
Myotis lucifugus	79	min 88, median 401, max 3917	
Ceratotherium simum simum	227	min 88, median 679, max 3668	
Equus przewalskii	240	min 46, median 677, max 3737	
Equus caballus (Thoroughbred)	145	min 159, median 677, max 3751	
Equus caballus (Mongolian)	248	min 154, median 672, max 3750	
Pantholops hodgsonii	367462	min 33, median 466, max 6792	
Capra hircus	324991	min 33, median 520, max 6967	
Ovis aries (Texel)	286618	min 48, median 633, max 7098	
Ovis aries musimon	293260	min 34, median 592, max 7130	

Species	**# BovB hits**	**Length distribution (bp)**	**Notes**
Bubalus bubalis	414623	min 32, median 557, max 7930	
Bison bison bison	423522	min 34, median 561, max 8187	
Bos mutus	348770	min 33, median 544, max 8623	
Bos indicus	278488	min 35, median 820, max 8107	
Bos taurus	275419	min 43, median 867, max 7738	Species with lots of active BovBs seem to generate numerous nested BovB regions in the genome (which is why the max BovB length is 7738 nt).
SAUROPSIDA			
Pogona vitticeps	26841	min 37, median 455, max 3658	
Anolis carolinensis	411	min 103, median 1205, max 3562	
Vipera berus berus	6641	min 46, median 449, max 3490	
Crotalus mitchellii pyrrhus	424	min 50, median 345, max 3360	
Ophiophagus hannah	1206	min 45, median 412, max 3667	
Python bivittatus	1386	min 48, median 497, max 3793	
AMPHIBIA			
Xenopus tropicalis	12	min 231, median 280, max 2623	
NEOPTERYGII			
Lepisosteus oculatus	20	min 91, median 513, max 3079	
Danio rerio	20	min 196, median 622, max 3079	
Cynoglossus semilaevis	8	min 292, median 469, max 2463	
ECDYSOZOA			
Cimex lectularius	41	min 258, median 439, max 2625	
Papilio glaucus	9	min 241, median 476, max 2981	
Heliconius melpomene melpomene	2	min 2829, median 2829, max 2872	
Danaus plexippus	2	min 2450, median 2450, max 2583	
Bombyx mori	165	min 62, median 382, max 2902	
Manduca sexta	3	min 2302, median 2504, max 2891	
Plutella xylostella	7	min 2348, median 2667, max 2974	
Linepithema humile	1	min 2861, median 2861, max 2861	
Solenopsis invicta	8	min 897, median 2285, max 3090	
Pogonomyrmex barbatus	1	min 2713, median 2713, max 2713	
Centruroides exilicauda	39	min 1985, median 2632, max 3251	
Mesobuthus martensii	124	min 200, median 2666, max 3737	
ANNELIDA			
Helobdella robusta	65	min 138, median 667, max 2959	
ECHINOIDEA			
Strongylocentrotus purpuratus	106	min 185, median 445, max 3373	
TUNICATA			
Ciona savignyi	10	min 736, median 2752, max 2852	

Table 4: BovB status of each species from our 503 genome dataset

Table C.4: **BovB status:** The union of LASTZ and TBLASTN results was used to determine the most likely status of each species: BovB- (no BovB elements found using either method) or BovB+ (BovBs found in full-length or fragment form, using either TBLASTN or LASTZ). This was used as a control of sorts for genome quality. The last column (Status) is left blank if the species does not contain BovBs (i.e. blank = BovB-). Some genomes contain BovB-like RTE sequences (e.g. see Figure 1); these are labelled RTE+.

No	**Species**	**Verified BovBs in the genome (LASTZ)**	**Verified BovBs for the taxon (TBLASTN)**	**Status (blank=BovB-)**
MAMMALIA				
1	*Tachyglossus aculeatus*	Y	Y	BovB+
2	*Ornithorhynchus anatinus*	Y	Y	BovB+
3	*Monodelphis domestica*	Y	Y	BovB+
4	*Macropus eugenii*	Y	Y	BovB+
5	*Sarcophilus harrisii*	Y	Y	BovB+
6	*Dasypus novemcinctus*	N	N	
7	*Choloepus hoffmanni*	N	N	
8	*Chrysochloris asiatica*	Y	N	BovB+
9	*Echinops telfairi*	Y	Y	BovB+
10	*Orycteropus afer afer*	Y	N	BovB+
11	*Elephantulus edwardii*	Y	N	BovB+
12	*Trichechus manatus latirostris*	Y	N	BovB+
13	*Procavia capensis*	Y	Y	BovB+
14	*Loxodonta africana*	Y	Y	BovB+
15	*Erinaceus europaeus*	N	N	
16	*Sorex araneus*	N	N	
17	*Condylura cristata*	N	N	
18	*Pteropus alecto*	Y	Y	BovB+
19	*Pteropus vampyrus*	Y	Y	BovB+
20	*Eidolon helvum*	Y	N	BovB+
21	*Megaderma lyra*	Y	N	BovB+
22	*Rhinolophus ferrumequinum*	Y	Y	BovB+
23	*Pteronotus parnellii*	Y	N	BovB+
24	*Eptesicus fuscus*	Y	N	BovB+
25	*Myotis brandtii*	Y	Y	BovB+
26	*Myotis davidii*	Y	Y	BovB+
27	*Myotis lucifugus*	Y	Y	BovB+
28	*Ceratotherium simum simum*	Y	N	BovB+
29	*Equus przewalskii*	Y	Y	BovB+
30	*Equus caballus* (Thoroughbred)	Y	Y	BovB+
31	*Equus caballus* (Mongolian)	Y	Y	BovB+
32	*Manis pentadactyla*	N	N	

No	**Species**	**Verified BovBs in the genome (LASTZ)**	**Verified BovBs for the taxon (TBLASTN)**	**Status (blank=BovB-)**
33	*Felis catus*	N	N	
34	*Panthera tigris altaica*	N	N	
35	*Canis lupus familiaris*	N	N	
36	*Ursus maritimus*	N	N	
37	*Ailuropoda melanoleuca*	N	N	
38	*Leptonychotes weddellii*	N	N	
39	*Odobenus rosmarus divergens*	N	N	
40	*Mustela putorius furo*	N	N	
41	*Camelus dromedarius*	N	N	
42	*Camelus ferus*	N	N	
43	*Vicugna pacos*	N	N	
44	*Sus scrofa* (Duroc)	N	N	
45	*Sus scrofa* (Tibetan)	N	N	
46	*Sus scrofa* (Ellegaard Gottingen minipig)	N	N	
47	*Balaenoptera acutorostrata scammoni*	N	N	
48	*Physeter catodon*	N	N	
49	*Lipotes vexillifer*	N	N	
50	*Tursiops truncatus*	N	N	
51	*Orcinus orca*	N	N	
52	*Pantholops hodgsonii*	Y	Y	BovB+
53	*Capra hircus*	Y	Y	BovB+
54	*Ovis aries* (Texel)	Y	Y	BovB+
55	*Ovis aries musimon*	Y	Y	BovB+
56	*Bubalus bubalis*	Y	Y	BovB+
57	*Bison bison bison*	Y	Y	BovB+
58	*Bos mutus*	Y	Y	BovB+
59	*Bos indicus*	Y	Y	BovB+
60	*Bos taurus*	Y	Y	BovB+
61	*Ochotona princeps*	N	N	
62	*Oryctolagus cuniculus*	N	N	
63	*Ictidomys tridecemlineatus*	N	N	
64	*Heterocephalus glaber*	N	N	
65	*Fukomys damarensis*	N	N	
66	*Cavia aperea*	N	N	
67	*Cavia porcellus*	N	N	
68	*Chinchilla lanigera*	N	N	
69	*Octodon degus*	N	N	
70	*Dipodomys ordii*	N	N	
71	*Jaculus jaculus*	N	N	
72	*Nannospalax galili*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
73	*Mesocricetus auratus*	N	N	
74	*Cricetulus griseus*	N	N	
75	*Microtus ochrogaster*	N	N	
76	*Peromyscus maniculatus bairdii*	N	N	
77	*Rattus norvegicus*	N	N	
78	*Mus musculus*	N	N	
79	*Tupaia belangeri*	N	N	
80	*Tupaia chinensis*	N	N	
81	*Galeopterus variegatus*	N	N	
82	*Otolemur garnettii*	N	N	
83	*Microcebus murinus*	N	N	
84	*Tarsius syrichta*	N	N	
85	*Callithrix jacchus*	N	N	
86	*Saimiri boliviensis boliviensis*	N	N	
87	*Rhinopithecus roxellana*	N	N	
88	*Nasalis larvatus*	N	N	
89	*Chlorocebus sabaeus*	N	N	
90	*Macaca fascicularis*	N	N	
91	*Macaca mulatta*	N	N	
92	*Papio anubis*	N	N	
93	*Nomascus leucogenys*	N	N	
94	*Pongo abelii*	N	N	
95	*Gorilla gorilla gorilla*	N	N	
96	*Pan paniscus*	N	N	
97	*Pan troglodytes*	N	N	
98	*Homo sapiens*	N	N	
SAUROPSIDA				
99	*Apalone spinifera*	N	N	RTE+
100	*Pelodiscus sinensis*	N	N	RTE+
101	*Chelonia mydas*	N	N	RTE+
102	*Chrysemys picta bellii*	N	N	RTE+
103	*Struthio camelus australis*	N	N	
104	*Tinamus guttatus*	N	N	
105	*Anas platyrhynchos*	N	N	
106	*Lyrurus tetrix tetrix*	N	N	
107	*Gallus gallus*	N	N	
108	*Coturnix japonica*	N	N	
109	*Meleagris gallopavo*	N	N	
110	*Colinus virginianus*	N	N	
111	*Acanthisitta chloris*	N	N	
112	*Manacus vitellinus*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
113	*Zonotrichia albicollis*	N	N	
114	*Geospiza fortis*	N	N	
115	*Serinus canaria*	N	N	
116	*Taeniopygia guttata*	N	N	
117	*Ficedula albicollis*	N	N	
118	*Pseudopodoces humilis*	N	N	
119	*Corvus brachyrhynchos*	N	N	
120	*Corvus cornix cornix*	N	N	
121	*Ara macao*	N	N	
122	*Amazona vittata*	N	N	
123	*Melopsittacus undulatus*	N	N	
124	*Nestor notabilis*	N	N	
125	*Falco cherrug*	N	N	
126	*Falco peregrinus*	N	N	
127	*Cariama cristata*	N	N	
128	*Merops nubicus*	N	N	
129	*Picoides pubescens*	N	N	
130	*Buceros rhinoceros silvestris*	N	N	
131	*Apaloderma vittatum*	N	N	
132	*Leptosomus discolor*	N	N	
133	*Haliaeetus albicilla*	N	N	
134	*Haliaeetus leucocephalus*	N	N	
135	*Aquila chrysaetos canadensis*	N	N	
136	*Cathartes aura*	N	N	
137	*Tyto alba*	N	N	
138	*Colius striatus*	N	N	
139	*Charadrius vociferus*	N	N	
140	*Balearica regulorum gibbericeps*	N	N	
141	*Chlamydotis macqueenii*	N	N	
142	*Cuculus canorus*	N	N	
143	*Fulmarus glacialis*	N	N	
144	*Aptenodytes forsteri*	N	N	
145	*Pygoscelis adeliae*	N	N	
146	*Phalacrocorax carbo*	N	N	
147	*Pelecanus crispus*	N	N	
148	*Nipponia nippon*	N	N	
149	*Egretta garzetta*	N	N	
150	*Phaethon lepturus*	N	N	
151	*Gavia stellata*	N	N	
152	*Tauraco erythrolophus*	N	N	
153	*Opisthocomus hoazin*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
154	*Columba livia*	N	N	
155	*Pterocles gutturalis*	N	N	
156	*Calypte anna*	N	N	
157	*Chaetura pelagica*	N	N	
158	*Caprimulgus carolinensis*	N	N	
159	*Eurypyga helias*	N	N	
160	*Mesitornis unicolor*	N	N	
161	*Podiceps cristatus*	N	N	
162	*Phoenicopterus ruber ruber*	N	N	
163	*Alligator mississippiensis*	N	N	RTE+
164	*Alligator sinensis*	N	N	RTE+
165	*Crocodylus porosus*	N	N	RTE+
166	*Gavialis gangeticus*	N	N	RTE+
167	*Pogona vitticeps*	Y	Y	BovB+
168	*Anolis carolinensis*	Y	Y	BovB+
169	*Vipera berus berus*	Y	N	BovB+
170	*Crotalus mitchellii pyrrhus*	Y	N	BovB+
171	*Ophiophagus hannah*	Y	N	BovB+
172	*Python bivittatus*	Y	Y	BovB+
AMPHIBIA				
173	*Nanorana parkeri*	N	N	
174	*Xenopus tropicalis*	Y	Y	BovB+
NEOPTERYGII				
175	*Lepisosteus oculatus*	Y	Y	BovB+
176	*Anguilla anguilla*	N	N	
177	*Anguilla japonica*	N	N	
178	*Danio rerio*	Y	Y	BovB+
179	*Astyanax mexicanus*	N	N	RTE+
180	*Oryzias latipes*	N	N	RTE+
181	*Poecilia formosa*	N	N	RTE+
182	*Xiphophorus maculatus*	N	N	RTE+
183	*Fundulus heteroclitus*	N	N	RTE+
184	*Takifugu flavidus*	N	N	RTE+
185	*Takifugu rubripes*	N	N	RTE+
186	*Tetraodon nigroviridis*	N	N	RTE+
187	*Cynoglossus semilaevis*	Y	N	BovB+
188	*Haplochromis burtoni*	N	N	
189	*Pundamilia nyererei*	N	N	
190	*Maylandia zebra*	N	N	
191	*Neolamprologus brichardi*	N	N	
192	*Oreochromis niloticus*	N	N	
193	*Sebastes nigrocinctus*	N	N	
194	*Sebastes rubrivinctus*	N	N	RTE+

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
195	*Gasterosteus aculeatus*	N	N	
196	*Gadus morhua*	N	N	RTE+
CHONDRICHTHYES				
197	*Callorhinchus milii*	N	N	
198	*Carcharhinus brachyurus*	N	N	
ECDYSOZOA				
199	*Ephemera danica*	N	N	
200	*Ladona fulva*	N	N	RTE+
201	*Pediculus humanus corporis*	N	N	
202	*Frankliniella occidentalis*	N	N	
203	*Diaphorina citri*	N	N	RTE+
204	*Pachypsylla venusta*	N	N	RTE+
205	*Acyrthosiphon pisum*	N	N	RTE+
206	*Nilaparvata lugens*	N	N	RTE+
207	*Oncopeltus fasciatus*	N	N	
208	*Rhodnius prolixus*	N	N	
209	*Cimex lectularius*	Y	Y	BovB+
210	*Onthophagus taurus*	N	N	
211	*Agrilus planipennis*	N	N	RTE+
212	*Tribolium castaneum*	N	N	
213	*Anoplophora glabripennis*	N	N	
214	*Leptinotarsa decemlineata*	N	N	
215	*Dendroctonus ponderosae*	N	N	
216	*Mengenilla moldrzyki*	N	N	
217	*Aedes aegypti*	N	N	
218	*Culex quinquefasciatus*	N	N	
219	*Anopheles albimanus*	N	N	
220	*Anopheles arabiensis*	N	N	
221	*Anopheles atroparvus*	N	N	
222	*Anopheles christyi*	N	N	
223	*Anopheles culicifacies*	N	N	
224	*Anopheles darlingi*	N	N	
225	*Anopheles dirus*	N	N	
226	*Anopheles epiroticus*	N	N	
227	*Anopheles farauti*	N	N	
228	*Anopheles funestus*	N	N	
229	*Anopheles gambiae*	N	N	
230	*Anopheles maculatus*	N	N	
231	*Anopheles melas*	N	N	
232	*Anopheles merus*	N	N	
233	*Anopheles minimus*	N	N	
234	*Anopheles quadriannulatus*	N	N	
235	*Anopheles sinensis*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
236	*Anopheles stephensi*	N	N	
237	*Mayetiola destructor*	N	N	
238	*Lutzomyia longipalpis*	N	N	
239	*Phlebotomus papatasi*	N	N	
240	*Ceratitis capitata*	N	N	
241	*Drosophila albomicans*	N	N	
242	*Drosophila ananassae*	N	N	
243	*Drosophila biarmipes*	N	N	
244	*Drosophila bipectinata*	N	N	
245	*Drosophila elegans*	N	N	
246	*Drosophila erecta*	N	N	
247	*Drosophila eugracilis*	N	N	
248	*Drosophila ficusphila*	N	N	
249	*Drosophila grimshawi*	N	N	
250	*Drosophila kikkawai*	N	N	
251	*Drosophila melanogaster*	N	N	
252	*Drosophila miranda*	N	N	
253	*Drosophila mojavensis*	N	N	
254	*Drosophila persimilis*	N	N	
255	*Drosophila pseudoobscura pseudoobscura*	N	N	
256	*Drosophila rhopaloa*	N	N	
257	*Drosophila sechellia*	N	N	
258	*Drosophila simulans*	N	N	
259	*Drosophila suzukii*	N	N	
260	*Drosophila takahashii*	N	N	
261	*Drosophila virilis*	N	N	
262	*Drosophila willistoni*	N	N	
263	*Drosophila yakuba*	N	N	
264	*Musca domestica*	N	N	
265	*Glossina austeni*	N	N	
266	*Glossina brevipalpis*	N	N	
267	*Glossina fuscipes fuscipes*	N	N	
268	*Glossina morsitans morsitans*	N	N	
269	*Glossina pallidipes*	N	N	
270	*Limnephilus lunatus*	N	N	
271	*Papilio glaucus*	Y	N	BovB+
272	*Papilio polytes*	N	N	RTE+
273	*Papilio xuthus*	N	N	RTE+
274	*Heliconius melpomene melpomene*	Y	N	BovB+
275	*Melitaea cinxia*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
276	*Danaus plexippus*	Y	N	BovB+
277	*Bombyx mori*	Y	Y	BovB+
278	*Manduca sexta*	Y	N	BovB+
279	*Plutella xylostella*	Y	Y	BovB+
280	*Athalia rosae*	N	N	
281	*Cephus cinctus*	N	N	
282	*Orussus abietinus*	N	N	
283	*Ceratosolen solmsi marchali*	N	N	
284	*Nasonia giraulti*	N	N	
285	*Nasonia longicornis*	N	N	
286	*Nasonia vitripennis*	N	N	
287	*Copidosoma floridanum*	N	N	
288	*Trichogramma pretiosum*	N	N	
289	*Microplitis demolitor*	N	N	
290	*Megachile rotundata*	N	N	
291	*Apis dorsata*	N	N	
292	*Apis florea*	N	N	
293	*Apis mellifera*	N	N	
294	*Bombus impatiens*	N	N	
295	*Bombus terrestris*	N	N	
296	*Linepithema humile*	Y	N	BovB+
297	*Camponotus floridanus*	N	N	RTE+
298	*Acromyrmex echinatior*	N	N	RTE+
299	*Atta cephalotes*	N	N	RTE+
300	*Solenopsis invicta*	Y	Y	BovB+
301	*Pogonomyrmex barbatus*	Y	N	BovB+
302	*Harpegnathos saltator*	N	N	RTE+
303	*Cerapachys biroi*	N	N	RTE+
304	*Blattella germanica*	N	N	RTE+
305	*Zootermopsis nevadensis*	N	N	RTE+
306	*Daphnia pulex*	N	N	
307	*Eurytemora affinis*	N	N	
308	*Hyalella azteca*	N	N	
309	*Strigamia maritima*	N	N	
310	*Stegodyphus mimosarum*	N	N	
311	*Latrodectus hesperus*	N	N	
312	*Parasteatoda tepidariorum*	N	N	
313	*Tetranychus urticae*	N	N	
314	*Dermatophagoides farinae*	N	N	
315	*Sarcoptes scabiei type canis*	N	N	
316	*Achipteria coleoptrata*	N	N	
317	*Hypochthonius rufulus*	N	N	
318	*Platynothrus peltifer*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
319	*Steganacarus magnus*	N	N	
320	*Ixodes ricinus*	N	N	RTE+
321	*Ixodes scapularis*	N	N	RTE+
322	*Rhipicephalus microplus*	N	N	RTE+
323	*Metaseiulus occidentalis*	N	N	RTE+
324	*Varroa destructor*	N	N	
325	*Centruroides exilicauda*	Y	N	BovB+
326	*Mesobuthus martensii*	Y	N	BovB+
327	*Limulus polyphemus*	N	N	RTE+
328	*Trichinella spiralis*	N	N	
329	*Ascaris suum*	N	N	
330	*Elaeophora elaphi*	N	N	
331	*Onchocerca volvulus*	N	N	
332	*Steinernema monticolum*	N	N	
333	*Panagrellus redivivus*	N	N	
334	*Haemonchus contortus*	N	N	
335	*Necator americanus*	N	N	
336	*Heterorhabditis bacteriophora*	N	N	
337	*Caenorhabditis angaria*	N	N	
338	*Caenorhabditis brenneri*	N	N	
339	*Caenorhabditis briggsae*	N	N	
340	*Caenorhabditis elegans*	N	N	
341	*Caenorhabditis japonica*	N	N	
342	*Caenorhabditis sp. 11 MAF-2010*	N	N	
343	*Priapulus caudatus*	N	N	RTE+
ROTIFERA				
344	*Adineta vaga*	N	N	RTE+
PLATYHELMINTHES				
345	*Schistosoma curassoni*	N	N	RTE+
346	*Schistosoma haematobium*	N	N	RTE+
347	*Schistosoma japonicum*	N	N	RTE+
348	*Schistosoma mansoni*	N	N	RTE+
349	*Schistosoma margrebowiei*	N	N	RTE+
350	*Schistosoma mattheei*	N	N	RTE+
351	*Schistosoma rodhaini*	N	N	RTE+
352	*Clonorchis sinensis*	N	N	
353	*Echinococcus granulosus*	N	N	
354	*Echinococcus multilocularis*	N	N	
355	*Hymenolepis microstoma*	N	N	
ANNELIDA				
356	*Capitella teleta*	N	N	

No	**Species**	**Verified BovBs in the genome (LASTZ)**	**Verified BovBs for the taxon (TBLASTN)**	**Status (blank=BovB-)**
357	*Helobdella robusta*	Y	N	BovB+
MOLLUSCA				
358	*Crassostrea gigas*	N	N	RTE+
359	*Lottia gigantea*	N	N	
360	*Aplysia californica*	N	N	RTE+
361	*Biomphalaria glabrata*	N	N	RTE+
CNIDARIA				
362	*Nematostella vectensis*	N	N	RTE+
363	*Hydra vulgaris*	N	N	
TENTACULATA				
364	*Mnemiopsis leidyi*	N	N	
PLACOZOA				
365	*Trichoplax adhaerens*	N	N	
PORIFERA				
366	*Amphimedon queenslandica*	N	N	
VIRIDIPLANTAE				
367	*Micromonas pusilla CCMP1545*	N	N	
368	*Micromonas sp. RCC299*	N	N	
369	*Ostreococcus lucimarinus CCE9901*	N	N	
370	*Ostreococcus tauri*	N	N	
371	*Chlamydomonas reinhardtii*	N	N	
372	*Volvox carteri f. nagariensis*	N	N	
373	*Chlorella variabilis*	N	N	
374	*Auxenochlorella protothecoides*	N	N	
375	*Helicosporidium sp. ATCC 50920*	N	N	
376	*Coccomyxa subellipsoidea C-169*	N	N	
377	*Klebsormidium flaccidum*	N	N	
378	*Physcomitrella patens*	N	N	
379	*Selaginella moellendorffii*	N	N	
380	*Pinus taeda*	N	N	
381	*Amborella trichopoda*	N	N	
382	*Spirodela polyrhiza*	N	N	
383	*Phoenix dactylifera*	N	N	
384	*Elaeis oleifera*	N	N	
385	*Ensete ventricosum*	N	N	
386	*Musa acuminata subsp. malaccensis*	N	N	
387	*Sorghum bicolor*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
388	*Zea mays*	N	N	
389	*Setaria italica*	N	N	
390	*Brachypodium distachyon*	N	N	
391	*Leersia perrieri*	N	N	
392	*Oryza barthii*	N	N	
393	*Oryza brachyantha*	N	N	
394	*Oryza glumipatula*	N	N	
395	*Oryza longistaminata*	N	N	
396	*Oryza meridionalis*	N	N	
397	*Oryza nivara*	N	N	
398	*Oryza punctata*	N	N	
399	*Oryza sativa Japonica Group*	N	N	
400	*Zizania latifolia*	N	N	
401	*Aegilops tauschii*	N	N	
402	*Triticum urartu*	N	N	
403	*Nelumbo nucifera*	N	N	
404	*Lupinus angustifolius*	N	N	
405	*Phaseolus vulgaris*	N	N	
406	*Cajanus cajan*	N	N	
407	*Vigna angularis var. angularis*	N	N	
408	*Vigna radiata var. radiata*	N	N	
409	*Glycine max*	N	N	
410	*Glycine soja*	N	N	
411	*Cicer arietinum*	N	N	
412	*Medicago truncatula*	N	N	
413	*Trifolium pratense*	N	N	
414	*Lotus japonicus*	N	N	
415	*Malus x domestica*	N	N	
416	*Pyrus x bretschneideri*	N	N	
417	*Prunus mume*	N	N	
418	*Prunus persica*	N	N	
419	*Fragaria iinumae*	N	N	
420	*Fragaria nubicola*	N	N	
421	*Fragaria orientalis*	N	N	
422	*Fragaria vesca subsp. vesca*	N	N	
423	*Fragaria x ananassa*	N	N	
424	*Morus notabilis*	N	N	
425	*Cannabis sativa*	N	N	
426	*Castanea mollissima*	N	N	
427	*Betula nana*	N	N	
428	*Cucumis melo*	N	N	

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
429	*Cucumis sativus*	N	N	
430	*Citrullus lanatus*	N	N	
431	*Lagenaria siceraria*	N	N	
432	*Populus euphratica*	N	N	
433	*Populus trichocarpa*	N	N	
434	*Jatropha curcas*	N	N	
435	*Manihot esculenta subsp. flabellifolia*	N	N	
436	*Ricinus communis*	N	N	
437	*Linum usitatissimum*	N	N	
438	*Eucalyptus camaldulensis*	N	N	
439	*Eucalyptus grandis*	N	N	
440	*Carica papaya*	N	N	
441	*Arabidopsis halleri subsp. gemmifera*	N	N	
442	*Arabidopsis lyrata subsp. lyrata*	N	N	
443	*Arabidopsis thaliana*	N	N	
444	*Camelina sativa*	N	N	
445	*Capsella rubella*	N	N	
446	*Brassica napus*	N	N	
447	*Brassica oleracea var. oleracea*	N	N	
448	*Brassica rapa*	N	N	
449	*Raphanus raphanistrum subsp. raphanistrum*	N	N	
450	*Raphanus sativus*	N	N	
451	*Aethionema arabicum*	N	N	
452	*Arabis alpina*	N	N	
453	*Eutrema parvulum*	N	N	
454	*Eutrema salsugineum*	N	N	
455	*Sisymbrium irio*	N	N	
456	*Leavenworthia alabamica*	N	N	
457	*Tarenaya hassleriana*	N	N	
458	*Gossypium arboreum*	N	N	
459	*Gossypium raimondii*	N	N	
460	*Theobroma cacao*	N	N	
461	*Aquilaria agallochum*	N	N	
462	*Azadirachta indica*	N	N	
463	*Citrus clementina*	N	N	
464	*Citrus sinensis*	N	N	
465	*Vitis vinifera*	N	N	

No	**Species**	**Verified BovBs in the genome (LASTZ)**	**Verified BovBs for the taxon (TBLASTN)**	**Status (blank=BovB-)**
466	*Amaranthus hypochondriacus*	N	N	
467	*Amaranthus tuberculatus*	N	N	
468	*Beta vulgaris subsp. vulgaris*	N	N	
469	*Spinacia oleracea*	N	N	
470	*Dianthus caryophyllus*	N	N	
471	*Actinidia chinensis*	N	N	
472	*Vaccinium macrocarpon*	N	N	
473	*Diospyros lotus*	N	N	
474	*Primula veris*	N	N	
475	*Solanum arcanum*	N	N	
476	*Solanum habrochaites*	N	N	
477	*Solanum lycopersicum*	N	N	
478	*Solanum melongena*	N	N	
479	*Solanum pennellii*	N	N	
480	*Solanum pimpinellifolium*	N	N	
481	*Solanum tuberosum*	N	N	
482	*Capsicum annuum*	N	N	
483	*Nicotiana sylvestris*	N	N	
484	*Nicotiana tomentosiformis*	N	N	
485	*Fraxinus excelsior*	N	N	
486	*Penstemon centranthifolius*	N	N	
487	*Penstemon grinnellii*	N	N	
488	*Sesamum indicum*	N	N	
489	*Genlisea aurea*	N	N	
490	*Mimulus guttatus*	N	N	
491	*Conyza canadensis*	N	N	
ECHINOIDEA				
492	*Lytechinus variegatus*	N	N	RTE+
493	*Strongylocentrotus purpuratus*	Y	Y	BovB+
ASTEROIDEA				
494	*Patiria miniata*	N	N	RTE+
ENTEROPNEUSTA				
495	*Saccoglossus kowalevskii*	N	N	RTE+
TUNICATA				
496	*Ciona intestinalis*	N	N	
497	*Ciona savignyi*	Y	N	BovB+
498	*Botryllus schlosseri*	N	N	RTE+
499	*Oikopleura dioica*	N	N	
LEPTOCARDII				
500	*Branchiostoma floridae*	N	N	RTE+

No	Species	Verified BovBs in the genome (LASTZ)	Verified BovBs for the taxon (TBLASTN)	Status (blank=BovB-)
CEPHALASPIDOMORPHI				
501	*Lethenteron camtschaticum*	N	N	RTE+
502	*Petromyzon marinus*	N	N	RTE+
SARCOPTERYGII				
503	*Latimeria chalumnae*	N	N	RTE+

Table 5: Genome coverage of L1 and BovB elements

Table C.5: **Genome covered by L1s/BovBs:** Shows the calculations used to generate the bargraph in Fig 1 of the manuscript. BovB and L1 base counts include both full-length elements and any fragments that were picked up during the genome screening.

Species	Genome base count	BovB base count, percentage of genome	L1 base count, percentage of genome
MAMMALIA			
Tachyglossus aculeatus	2020007912	965926, 0.0478179%	0, 0%
Ornithorhynchus anatinus	1996811212	1157364, 0.0579606%	0, 0%
Monodelphis domestica	3605631728	346171, 0.00960084%	168269339, 4.66685%
Macropus eugenii	3075184024	588730, 0.0191445%	77395243, 2.51677%
Sarcophilus harrisii	3174693010	544524, 0.017152%	88996921, 2.80332%
Dasypus novemcinctus	3631505655	0, 0%	231152711, 6.3652%
Choloepus hoffmanni	2458927620	0, 0%	145205788, 5.90525%
Chrysochloris asiatica	4210093806	135792594, 3.22541%	49752325, 1.18174%
Echinops telfairi	2947007737	1811541, 0.0614705%	29871752, 1.01363%
Orycteropus afer afer	4444063711	256644763, 5.775%	117088389, 2.63471%
Elephantulus edwardii	3843982861	90987476, 2.36701%	47760267, 1.24247%
Trichechus manatus latirostris	3103791524	124360776, 4.00674%	152421150, 4.91081%
Procavia capensis	2985258999	28453440, 0.953131%	176734259, 5.92023%
Loxodonta africana	3271792967	175520711, 5.36466%	246282525, 7.52745%
Erinaceus europaeus	2715703478	0, 0%	38926214, 1.43337%
Sorex araneus	2423158183	0, 0%	49457765, 2.04105%
Condylura cristata	1769662895	0, 0%	16239492, 0.91766%
Pteropus alecto	1985958707	108073, 0.00544186%	79399739, 3.99806%
Pteropus vampyrus	1996076410	76025, 0.00380872%	66959412, 3.35455%
Eidolon helvum	1837754460	96463, 0.00524896%	50488801, 2.74731%
Megaderma lyra	1735931796	67208, 0.00387158%	43229597, 2.49028%
Rhinolophus ferrumequinum	1926439238	118497, 0.00615109%	49969516, 2.59388%
Pteronotus parnellii	1960317893	76126, 0.00388335%	55530997, 2.83275%
Eptesicus fuscus	2026629342	53474, 0.00263857%	37032955, 1.82732%
Myotis brandtii	2107242811	75647, 0.00358986%	53972449, 2.56128%
Myotis davidii	2059799708	67371, 0.00327075%	39997269, 1.9418%
Myotis lucifugus	2034575300	79572, 0.00391099%	65002870, 3.19491%
Ceratotherium simum simum	2464350348	284350, 0.0115385%	114395410, 4.64201%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Equus przewalskii	2395937679	280798, 0.0117198%	129687578, 5.41281%
Equus caballus (Thoroughbred)	2484532062	202155, 0.00813654%	133162075, 5.35964%
Equus caballus (Mongolian)	2377506909	267730, 0.011261%	122177903, 5.13891%
Manis pentadactyla	2204732179	0, 0%	109776390, 4.97913%
Felis catus	2455541136	0, 0%	107677871, 4.3851%
Panthera tigris altaica	2391065193	0, 0%	119470565, 4.99654%
Canis lupus familiaris	2410960148	0, 0%	120532167, 4.99934%
Ursus maritimus	2301362327	0, 0%	114457228, 4.97346%
Ailuropoda melanoleuca	2299509015	0, 0%	98315007, 4.27548%
Leptonychotes weddellii	3156886159	0, 0%	100033358, 3.16874%
Odobenus rosmarus divergens	2400133628	0, 0%	119232686, 4.96775%
Mustela putorius furo	2410863155	0, 0%	105028093, 4.35645%
Camelus dromedarius	2004047047	0, 0%	77193530, 3.85188%
Camelus ferus	2009177929	0, 0%	72311247, 3.59905%
Vicugna pacos	2172191320	0, 0%	83626855, 3.84988%
Sus scrofa (Duroc)	2808509378	0, 0%	110132425, 3.92138%
Sus scrofa (Tibetan)	2489155924	0, 0%	127504368, 5.12239%
Sus scrofa (Ellegaard Gottingen minipig)	2358017222	0, 0%	117308917, 4.9749%
Balaenoptera acutorostrata scammoni	2431671281	0, 0%	149364501, 6.14246%
Physeter catodon	2280711356	0, 0%	142493079, 6.24775%
Lipotes vexillifer	2429195737	0, 0%	252970985, 10.4138%
Tursiops truncatus	2551401796	0, 0%	118560068, 4.64686%
Orcinus orca	2372903489	0, 0%	163647834, 6.89652%
Pantholops hodgsonii	2696869832	228804957, 8.48409%	73461109, 2.72394%
Capra hircus	2635832257	236511890, 8.97295%	58230719, 2.2092%
Ovis aries (Texel)	2619054388	247270195, 9.4412%	67111961, 2.56245%
Ovis aries musimon	2589834840	236553802, 9.13393%	60594519, 2.33971%
Bubalus bubalis	2836150610	339483910, 11.9699%	94809063, 3.34288%
Bison bison bison	2953606000	360899526, 12.2189%	101439058, 3.43441%
Bos mutus	2645145588	258096922, 9.75738%	73612458, 2.78293%
Bos indicus	2673949103	304313141, 11.3807%	72928551, 2.72737%
Bos taurus	2670422299	316362787, 11.8469%	82022340, 3.07151%
Ochotona princeps	2229824103	0, 0%	17456997, 0.782887%
Oryctolagus cuniculus	2737490501	0, 0%	86512668, 3.16029%
Ictidomys tridecemlineatus	2478393770	0, 0%	73469520, 2.9644%
Heterocephalus glaber	2618188253	0, 0%	84855607, 3.241%
Fukomys damarensis	2333892479	0, 0%	106759044, 4.57429%
Cavia aperea	2716396567	0, 0%	80685572, 2.97032%
Cavia porcellus	2723219641	0, 0%	184907697, 6.79004%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Chinchilla lanigera	2390852391	0, 0%	80331690, 3.35996%
Octodon degus	2995872505	0, 0%	80354250, 2.68217%
Dipodomys ordii	2158502098	0, 0%	31840306, 1.47511%
Jaculus jaculus	2835233679	0, 0%	49057597, 1.73028%
Nannospalax galili	3061408210	0, 0%	97805179, 3.19478%
Mesocricetus auratus	2504908775	0, 0%	45068178, 1.79919%
Cricetulus griseus	2399770464	0, 0%	65170585, 2.7157%
Microtus ochrogaster	2287340943	0, 0%	23943280, 1.04677%
Peromyscus maniculatus bairdii	2630541020	0, 0%	39306831, 1.49425%
Rattus norvegicus	2909682625	0, 0%	152883889, 5.25431%
Mus musculus	2730855475	0, 0%	181250040, 6.63712%
Tupaia belangeri	3660774957	0, 0%	50529309, 1.38029%
Tupaia chinensis	2846580235	0, 0%	72668499, 2.55284%
Galeopterus variegatus	3187643824	0, 0%	132080752, 4.14352%
Otolemur garnettii	2519724550	0, 0%	101241758, 4.01797%
Microcebus murinus	2902270736	0, 0%	67564827, 2.328%
Tarsius syrichta	3453847770	0, 0%	229913092, 6.65672%
Callithrix jacchus	2914958544	0, 0%	188627606, 6.47102%
Saimiri boliviensis boliviensis	2608572064	0, 0%	146266740, 5.60716%
Rhinopithecus roxellana	2899535590	0, 0%	196631595, 6.78149%
Nasalis larvatus	3011966170	0, 0%	108305414, 3.59584%
Chlorocebus sabaeus	2789639778	0, 0%	153057893, 5.48665%
Macaca fascicularis	2946827162	0, 0%	160223837, 5.43716%
Macaca mulatta	2969988180	0, 0%	101505212, 3.4177%
Papio anubis	2948380710	0, 0%	158360113, 5.37109%
Nomascus leucogenys	2962077449	0, 0%	178826244, 6.03719%
Pongo abelii	3441227734	0, 0%	199557600, 5.79902%
Gorilla gorilla gorilla	3029537234	0, 0%	150808752, 4.97795%
Pan paniscus	2869173508	0, 0%	202665854, 7.06356%
Pan troglodytes	3309577922	0, 0%	183828361, 5.55444%
Homo sapiens	3101788170	0, 0%	185335660, 5.97512%
SAUROPSIDA			
Apalone spinifera	1931078847	0, 0%	1145101, 0.0592985%
Pelodiscus sinensis	2202466388	0, 0%	1223508, 0.0555517%
Chelonia mydas	2208393880	0, 0%	2364138, 0.107052%
Chrysemys picta bellii	2365749696	0, 0%	2283384, 0.0965184%
Struthio camelus australis	1225025301	0, 0%	18038, 0.00147246%
Tinamus guttatus	1047056493	0, 0%	6309, 0.000602546%
Anas platyrhynchos	1105035747	0, 0%	15012, 0.00135851%
Lyrurus tetrix tetrix	657025294	0, 0%	4332, 0.000659335%
Gallus gallus	1046932099	0, 0%	8421, 0.00080435%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Coturnix japonica	531959683	0, 0%	505, 9.4932e-05%
Meleagris gallopavo	1061817101	0, 0%	5265, 0.000495848%
Colinus virginianus	1171855925	0, 0%	61795, 0.00527326%
Acanthisitta chloris	1035876403	0, 0%	22950, 0.00221552%
Manacus vitellinus	1145871783	0, 0%	18613, 0.00162435%
Zonotrichia albicollis	1052600561	0, 0%	10446, 0.000992399%
Geospiza fortis	1065292181	0, 0%	14305, 0.00134282%
Serinus canaria	1152083301	0, 0%	57746, 0.00501231%
Taeniopygia guttata	1233186341	0, 0%	14036, 0.00113819%
Ficedula albicollis	1118326800	0, 0%	14642, 0.00130928%
Pseudopodoces humilis	1042980823	0, 0%	17094, 0.00163896%
Corvus brachyrhynchos	1091312783	0, 0%	20644, 0.00189167%
Corvus cornix cornix	1049947905	0, 0%	17351, 0.00165256%
Ara macao	1204683257	0, 0%	17568, 0.00145831%
Amazona vittata	1175404042	0, 0%	21900, 0.00186319%
Melopsittacus undulatus	1117373619	0, 0%	18355, 0.00164269%
Nestor notabilis	1053559886	0, 0%	24207, 0.00229764%
Falco cherrug	1174811715	0, 0%	23963, 0.00203973%
Falco peregrinus	1171955363	0, 0%	23041, 0.00196603%
Cariama cristata	1132245425	0, 0%	30441, 0.00268855%
Merops nubicus	1062961556	0, 0%	21529, 0.00202538%
Picoides pubescens	1167323935	0, 0%	6385, 0.000546978%
Buceros rhinoceros silvestris	1065782791	0, 0%	16397, 0.00153849%
Apaloderma vittatum	1070836417	0, 0%	22035, 0.00205774%
Leptosomus discolour	1136244952	0, 0%	22022, 0.00193814%
Haliaeetus albicilla	1133549865	0, 0%	30405, 0.00268228%
Haliaeetus leucocephalus	1178409481	0, 0%	31482, 0.00267157%
Aquila chrysaetos Canadensis	1192725744	0, 0%	30224, 0.00253403%
Cathartes aura	1152571117	0, 0%	32870, 0.00285188%
Tyto alba	1120143088	0, 0%	35233, 0.0031454%
Colius striatus	1075931597	0, 0%	21899, 0.00203535%
Charadrius vociferus	1219859583	0, 0%	28482, 0.00233486%
Balearica regulorum gibbericeps	1127605500	0, 0%	30641, 0.00271735%
Chlamydotis macqueenii	1086566339	0, 0%	28027, 0.00257941%
Cuculus canorus	1153894225	0, 0%	17368, 0.00150516%
Fulmarus glacialis	1141395646	0, 0%	31925, 0.00279701%
Aptenodytes forsteri	1254347440	0, 0%	39272, 0.00313087%
Pygoscelis adeliae	1216600033	0, 0%	37061, 0.00304628%
Phalacrocorax carbo	1138967842	0, 0%	26087, 0.00229041%
Pelecanus crispus	1160924693	0, 0%	30298, 0.00260982%
Nipponia nippon	1223846297	0, 0%	32043, 0.00261822%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Egretta garzetta	1206484573	0, 0%	30689, 0.00254367%
Phaethon lepturus	1152958507	0, 0%	27206, 0.00235967%
Gavia stellata	1129677294	0, 0%	33996, 0.00300935%
Tauraco erythrolophus	1155540733	0, 0%	25539, 0.00221013%
Opisthocomus hoazin	1203712246	0, 0%	39475, 0.00327944%
Columba livia	1107971856	0, 0%	24884, 0.00224591%
Pterocles gutturalis	1069324295	0, 0%	30850, 0.002885%
Calypte anna	1105676412	0, 0%	15355, 0.00138874%
Chaetura pelagica	1119188094	0, 0%	14308, 0.00127843%
Caprimulgus carolinensis	1119683066	0, 0%	29862, 0.002667%
Eurypyga helias	1088019637	0, 0%	17313, 0.00159124%
Mesitornis unicolor	1087290853	0, 0%	19855, 0.0018261%
Podiceps cristatus	1134922578	0, 0%	20835, 0.00183581%
Phoenicopterus ruber ruber	1132184511	0, 0%	33503, 0.00295915%
Alligator mississippiensis	2174259888	0, 0%	4385712, 0.201711%
Alligator sinensis	2270550999	0, 0%	4292056, 0.189031%
Crocodylus porosus	2120573303	0, 0%	3282608, 0.154798%
Gavialis gangeticus	2188353730	0, 0%	3613361, 0.165118%
Pogona vitticeps	1816116151	17015661, 0.936926%	1970781, 0.108516%
Anolis carolinensis	1799143587	581615, 0.0323273%	4939730, 0.27456%
Vipera berus berus	1532390814	4242323, 0.276843%	3433981, 0.224093%
Crotalus mitchellii pyrrhus	1126791715	291616, 0.0258802%	1618037, 0.143597%
Ophiophagus hannah	1594074654	844276, 0.0529634%	4928678, 0.309187%
Python bivittatus	1435034535	1454340, 0.101345%	2437029, 0.169824%
AMPHIBIA			
Nanorana parkeri	2053849526	0, 0%	2490814, 0.121275%
Xenopus tropicalis	1437513269	5977, 0.000415787%	3073965, 0.213839%
NEOPTERYGII			
Lepisosteus oculatus	945861706	22131, 0.00233977%	251058, 0.0265428%
Anguilla anguilla	1018701900	0, 0%	591743, 0.0580879%
Anguilla japonica	1151120721	0, 0%	799538, 0.0694574%
Danio rerio	1412464843	15763, 0.00111599%	3802559, 0.269214%
Astyanax mexicanus	1191242572	0, 0%	252459, 0.0211929%
Oryzias latipes	869801494	0, 0%	1614085, 0.185569%
Poecilia formosa	748923461	0, 0%	701799, 0.0937077%
Xiphophorus maculatus	729647787	0, 0%	555241, 0.0760971%
Fundulus heteroclitus	1021882034	0, 0%	1159036, 0.113422%
Takifugu flavidus	378032400	0, 0%	323072, 0.0854615%
Takifugu rubripes	391484715	0, 0%	300506, 0.0767606%
Tetraodon nigroviridis	358618246	0, 0%	100055, 0.0279001%
Cynoglossus semilaevis	470182763	5546, 0.00117954%	63219, 0.0134456%
Haplochromis burtoni	831411547	0, 0%	918561, 0.110482%
Pundamilia nyererei	830133247	0, 0%	922531, 0.11113%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Maylandia zebra	849595388	0, 0%	1066736, 0.125558%
Neolamprologus brichardi	847893845	0, 0%	886493, 0.104552%
Oreochromis niloticus	927696114	0, 0%	1592022, 0.17161%
Sebastes nigrocinctus	687549873	0, 0%	398069, 0.0578967%
Sebastes rubrivinctus	756296653	0, 0%	383118, 0.0506571%
Gasterosteus aculeatus	463354448	0, 0%	185171, 0.0399631%
Gadus morhua	824327835	0, 0%	575410, 0.0698035%
CHONDRICHTHYES			
Callorhinchus milii	974481817	0, 0%	37878, 0.00388699%
Carcharhinus brachyurus	2832541493	0, 0%	667676, 0.0235716%
ECDYSOZOA			
Cimex lectularius	650477627	28417, 0.00436864%	0, 0%
Aedes aegypti	1383957531	0, 0%	2702946, 0.195306%
Culex quinquefasciatus	579042118	0, 0%	764337, 0.132%
Anopheles albimanus	170508315	0, 0%	13633, 0.0079955%
Anopheles arabiensis	246567867	0, 0%	23925, 0.00970321%
Anopheles atroparvus	224290125	0, 0%	24163, 0.0107731%
Anopheles christyi	172658580	0, 0%	17976, 0.0104113%
Anopheles culicifacies	202998806	0, 0%	18159, 0.00894537%
Anopheles darlingi	136935538	0, 0%	11366, 0.00830026%
Anopheles dirus	216307690	0, 0%	38793, 0.0179342%
Anopheles epiroticus	223486714	0, 0%	30250, 0.0135355%
Anopheles farauti	183103254	0, 0%	44357, 0.0242251%
Anopheles funestus	225223604	0, 0%	33565, 0.014903%
Anopheles gambiae	287805703	0, 0%	67022, 0.0232872%
Anopheles maculatus	141894015	0, 0%	27906, 0.0196668%
Anopheles melas	224162116	0, 0%	19346, 0.00863036%
Anopheles merus	288048996	0, 0%	47015, 0.0163219%
Anopheles minimus	201793324	0, 0%	25619, 0.0126957%
Anopheles quadriannulatus	283828998	0, 0%	24501, 0.00863231%
Anopheles sinensis	220777669	0, 0%	35160, 0.0159255%
Anopheles stephensi	221324304	0, 0%	27033, 0.0122142%
Ceratitis capitata	484773492	0, 0%	52264, 0.0107811%
Drosophila albomicans	253560284	0, 0%	12982, 0.00511989%
Drosophila ananassae	230993012	0, 0%	5521, 0.00239012%
Drosophila biarmipes	169378599	0, 0%	9683, 0.00571678%
Drosophila bipectinata	167263958	0, 0%	9143, 0.00546621%
Drosophila elegans	171267669	0, 0%	7777, 0.00454085%
Drosophila erecta	152712140	0, 0%	6451, 0.00422429%
Drosophila eugracilis	156942009	0, 0%	8434, 0.00537396%
Drosophila ficusphila	152439475	0, 0%	62328, 0.040887%
Drosophila grimshawi	200467819	0, 0%	13335, 0.00665194%
Drosophila kikkawai	164292578	0, 0%	9541, 0.00580732%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Drosophila melanogaster	143706478	0, 0%	4915, 0.00342017%
Drosophila miranda	136728780	0, 0%	8889, 0.00650119%
Drosophila mojavensis	193826310	0, 0%	35669, 0.0184026%
Drosophila persimilis	188374079	0, 0%	7072, 0.00375423%
Drosophila pseudoobscura pseudoobscura	152696384	0, 0%	7046, 0.00461439%
Drosophila rhopaloa	197375704	0, 0%	4465, 0.00226218%
Drosophila sechellia	166577145	0, 0%	4293, 0.00257718%
Drosophila simulans	137828247	0, 0%	3689, 0.00267652%
Drosophila suzukii	232923092	0, 0%	17442, 0.00748831%
Drosophila takahashii	182106768	0, 0%	6426, 0.0035287%
Drosophila virilis	206026697	0, 0%	14242, 0.0069127%
Drosophila willistoni	235516348	0, 0%	8718, 0.00370165%
Drosophila yakuba	165693946	0, 0%	7898, 0.00476662%
Musca domestica	750403944	0, 0%	31670, 0.00422039%
Glossina austeni	370264922	0, 0%	37853, 0.0102232%
Glossina brevipalpis	315360362	0, 0%	55174, 0.0174955%
Glossina fuscipes fuscipes	374774708	0, 0%	29433, 0.00785352%
Glossina morsitans morsitans	363107242	0, 0%	28698, 0.00790345%
Glossina pallidipes	357332231	0, 0%	27822, 0.00778603%
Papilio glaucus	374815656	7956, 0.00212264%	0, 0%
Heliconius melpomene melpomene	273786188	5701, 0.00208228%	0, 0%
Danaus plexippus	272853388	5033, 0.00184458%	0, 0%
Bombyx mori	481803763	77499, 0.0160852%	0, 0%
Manduca sexta	419412261	7697, 0.00183519%	0, 0%
Plutella xylostella	393454548	18924, 0.0048097%	0, 0%
Orussus abietinus	201220334	0, 0%	5972, 0.00296789%
Microplitis demolitor	250525215	0, 0%	46654, 0.0186225%
Linepithema humile	219500750	2861, 0.00130341%	0, 0%
Solenopsis invicta	396009169	18170, 0.00458828%	0, 0%
Pogonomyrmex barbatus	235645958	2713, 0.0011513%	0, 0%
Blattella germanica	2037201033	0, 0%	409642, 0.0201081%
Zootermopsis nevadensis	485009472	0, 0%	13812, 0.00284778%
Daphnia pulex	197206209	0, 0%	33508, 0.0169914%
Eurytemora affinis	494890867	0, 0%	121545, 0.02456%
Hyalella azteca	1178848281	0, 0%	38545, 0.00326972%
Latrodectus hesperus	1137104656	0, 0%	41311, 0.003633%
Ixodes ricinus	391986416	0, 0%	507180, 0.129387%
Ixodes scapularis	1765382190	0, 0%	1217198, 0.0689481%
Rhipicephalus microplus	144692238	0, 0%	81227, 0.0561378%
Metaseiulus occidentalis	151699350	0, 0%	20528, 0.013532%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Centruroides exilicauda	926399729	104086, 0.0112355%	41597, 0.00449018%
Mesobuthus martensii	925546267	267724, 0.0289261%	0, 0%
Limulus polyphemus	1828256766	0, 0%	103261, 0.00564806%
Trichinella spiralis	63525422	0, 0%	6531, 0.0102809%
ROTIFERA			
Adineta vaga	217933776	0, 0%	9556, 0.00438482%
PLATYHELMINTHES			
Schistosoma curassoni	344202215	0, 0%	60507, 0.0175789%
Schistosoma haematobium	375894156	0, 0%	34542, 0.00918929%
Schistosoma japonicum	402743189	0, 0%	68372, 0.0169766%
Schistosoma mansoni	364518427	0, 0%	17400, 0.00477342%
Schistosoma margrebowiei	367396528	0, 0%	23316, 0.00634628%
Schistosoma mattheei	340818150	0, 0%	47277, 0.0138716%
Schistosoma rodhaini	343294136	0, 0%	29750, 0.00866604%
Clonorchis sinensis	547288241	0, 0%	45925, 0.00839137%
ANNELIDA			
Capitella teleta	333283208	0, 0%	2516, 0.000754914%
Helobdella robusta	235376169	86170, 0.0366095%	120582, 0.0512295%
MOLLUSCA			
Crassostrea gigas	557717710	0, 0%	767205, 0.137562%
Lottia gigantea	359505668	0, 0%	61570, 0.0171263%
Aplysia californica	927296314	0, 0%	445254, 0.0480164%
Biomphalaria glabrata	916374414	0, 0%	185937, 0.0202905%
CNIDARIA			
Nematostella vectensis	356613585	0, 0%	58238, 0.0163308%
Hydra vulgaris	852155112	0, 0%	482878, 0.0566655%
VIRIDIPLANTAE			
Chlamydomonas reinhardtii	120185366	0, 0%	103509, 0.0861245%
Volvox carteri f. nagariensis	137684403	0, 0%	193699, 0.140683%
Chlorella variabilis	46159512	0, 0%	4645, 0.0100629%
Coccomyxa subellipsoidea C-169	48826616	0, 0%	145747, 0.298499%
Physcomitrella patens	479985347	0, 0%	21805, 0.00454285%
Selaginella moellendorffii	212315224	0, 0%	221463, 0.104309%
Pinus taeda	265480119	0, 0%	2279230, 0.858531%
Amborella trichopoda	706332640	0, 0%	3951716, 0.55947%
Spirodela polyrhiza	132009443	0, 0%	99825, 0.0756196%
Phoenix dactylifera	555607186	0, 0%	3838482, 0.690863%
Elaeis oleifera	1402725009	0, 0%	2164632, 0.154316%
Ensete ventricosum	172241963	0, 0%	8973, 0.00520953%
Musa acuminata subsp. malaccensis	472235617	0, 0%	259105, 0.0548677%
Sorghum bicolor	738540932	0, 0%	3099868, 0.419729%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Zea mays	2066912289	0, 0%	5199039, 0.251537%
Setaria italica	405737341	0, 0%	2260412, 0.557112%
Brachypodium distachyon	271923306	0, 0%	1680270, 0.617921%
Leersia perrieri	266687832	0, 0%	1138563, 0.426927%
Oryza barthii	308272304	0, 0%	1322283, 0.428933%
Oryza brachyantha	259907595	0, 0%	335151, 0.12895%
Oryza glumipatula	372860283	0, 0%	1157663, 0.310482%
Oryza longistaminata	326442508	0, 0%	1291015, 0.39548%
Oryza meridionalis	335668232	0, 0%	981776, 0.292484%
Oryza nivara	337950324	0, 0%	1314956, 0.389097%
Oryza punctata	393816603	0, 0%	1224886, 0.31103%
Oryza sativa Japonica Group	382150945	0, 0%	1661866, 0.434872%
Zizania latifolia	603989347	0, 0%	1251284, 0.20717%
Aegilops tauschii	3313650219	0, 0%	18059360, 0.544999%
Triticum urartu	3747047519	0, 0%	18388049, 0.490734%
Nelumbo nucifera	804484421	0, 0%	7948224, 0.98799%
Lupinus angustifolius	523297923	0, 0%	2298681, 0.439268%
Phaseolus vulgaris	521076696	0, 0%	2796546, 0.536686%
Cajanus cajan	510809477	0, 0%	855794, 0.167537%
Vigna angularis var. angularis	291823841	0, 0%	53735, 0.0184135%
Vigna radiata var. radiata	463085359	0, 0%	79561, 0.0171806%
Glycine max	973224514	0, 0%	4085129, 0.419752%
Glycine soja	863568428	0, 0%	4282042, 0.495854%
Cicer arietinum	530768543	0, 0%	405892, 0.0764725%
Medicago truncatula	314353944	0, 0%	742744, 0.236276%
Trifolium pratense	304979311	0, 0%	1163753, 0.381584%
Lotus japonicus	147812252	0, 0%	967215, 0.654354%
Malus x domestica	526197889	0, 0%	1184710, 0.225145%
Pyrus x bretschneideri	508550595	0, 0%	1432644, 0.281711%
Prunus mume	233872527	0, 0%	560963, 0.239858%
Prunus persica	227251827	0, 0%	509252, 0.224091%
Fragaria iinumae	199627509	0, 0%	415411, 0.208093%
Fragaria nubicola	203686407	0, 0%	514953, 0.252817%
Fragaria orientalis	214184023	0, 0%	583311, 0.272341%
Fragaria vesca subsp. vesca	214217322	0, 0%	615360, 0.28726%
Fragaria x ananassa	173229572	0, 0%	343570, 0.198332%
Morus notabilis	320378613	0, 0%	329447, 0.102831%
Cannabis sativa	757438891	0, 0%	2576828, 0.340203%
Castanea mollissima	833240550	0, 0%	4792366, 0.575148%
Betula nana	564011153	0, 0%	1958322, 0.347213%
Cucumis melo	374772472	0, 0%	893429, 0.238392%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Cucumis sativus	242977439	0, 0%	518212, 0.213276%
Citrullus lanatus	321046825	0, 0%	612617, 0.190819%
Lagenaria siceraria	176727258	0, 0%	245807, 0.139088%
Populus euphratica	495875768	0, 0%	1123953, 0.22666%
Populus trichocarpa	417129638	0, 0%	1046935, 0.250986%
Jatropha curcas	318363250	0, 0%	723641, 0.2273%
Manihot esculenta subsp. flabellifolia	390835892	0, 0%	727795, 0.186215%
Ricinus communis	350458699	0, 0%	156776, 0.0447345%
Linum usitatissimum	282201865	0, 0%	325517, 0.115349%
Eucalyptus camaldulensis	654922307	0, 0%	4785524, 0.730701%
Eucalyptus grandis	691269672	0, 0%	4792995, 0.693361%
Carica papaya	369781828	0, 0%	1441176, 0.389737%
Arabidopsis halleri subsp. gemmifera	221139660	0, 0%	1987573, 0.898786%
Arabidopsis lyrata subsp. lyrata	206667935	0, 0%	2095939, 1.01416%
Arabidopsis thaliana	119146348	0, 0%	641751, 0.538624%
Camelina sativa	641356059	0, 0%	7721486, 1.20393%
Capsella rubella	133063876	0, 0%	823622, 0.618967%
Brassica napus	930121500	0, 0%	11114963, 1.195%
Brassica oleracea var. oleracea	488593889	0, 0%	5823329, 1.19185%
Brassica rapa	283975909	0, 0%	3312144, 1.16635%
Raphanus raphanistrum subsp. raphanistrum	253833977	0, 0%	3389208, 1.33521%
Raphanus sativus	341476849	0, 0%	4697370, 1.3756%
Aethionema arabicum	192487774	0, 0%	2130657, 1.10691%
Arabis alpina	308032609	0, 0%	4009565, 1.30167%
Eutrema parvulum	137073209	0, 0%	822202, 0.599827%
Eutrema salsugineum	243110105	0, 0%	2611512, 1.07421%
Sisymbrium irio	245550082	0, 0%	2436742, 0.99236%
Leavenworthia alabamica	173431618	0, 0%	665633, 0.383801%
Tarenaya hassleriana	249929577	0, 0%	413736, 0.165541%
Gossypium arboreum	1560831891	0, 0%	4269614, 0.273547%
Gossypium raimondii	761405269	0, 0%	2427290, 0.318791%
Theobroma cacao	345993675	0, 0%	919236, 0.26568%
Aquilaria agallochum	726710453	0, 0%	756168, 0.104054%
Azadirachta indica	261457759	0, 0%	259674, 0.0993178%
Citrus clementine	301364702	0, 0%	1271619, 0.421954%
Citrus sinensis	327669411	0, 0%	1358265, 0.414523%
Vitis vinifera	485326422	0, 0%	7065807, 1.45589%

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Amaranthus hypochondriacus	502147592	0, 0%	1253176, 0.249563%
Amaranthus tuberculatus	4347977	0, 0%	25948, 0.596783%
Beta vulgaris subsp. vulgaris	566198563	0, 0%	5319420, 0.939497%
Spinacia oleracea	474077292	0, 0%	2677940, 0.564874%
Dianthus caryophyllus	567661845	0, 0%	1992303, 0.350967%
Actinidia chinensis	604217145	0, 0%	1248455, 0.206624%
Vaccinium macrocarpon	414621889	0, 0%	925221, 0.223148%
Diospyros lotus	1104189	0, 0%	1950, 0.1766%
Primula veris	309692940	0, 0%	343837, 0.111025%
Solanum arcanum	665186956	0, 0%	5085725, 0.764556%
Solanum habrochaites	724284658	0, 0%	5697424, 0.786628%
Solanum lycopersicum	781353311	0, 0%	4363900, 0.558505%
Solanum melongena	833080760	0, 0%	4551351, 0.546328%
Solanum pennellii	720458090	0, 0%	4967222, 0.689453%
Solanum pimpinellifolium	688247059	0, 0%	5140272, 0.746864%
Solanum tuberosum	705779115	0, 0%	5349970, 0.758023%
Capsicum annuum	3063642317	0, 0%	15430743, 0.503673%
Nicotiana sylvestris	2221831232	0, 0%	11722446, 0.527603%
Nicotiana tomentosiformis	1688312294	0, 0%	8944938, 0.529815%
Fraxinus excelsior	875243685	0, 0%	2088618, 0.238633%
Penstemon centranthifolius	4471593	0, 0%	13235, 0.29598%
Penstemon grinnellii	3663524	0, 0%	10469, 0.285763%
Sesamum indicum	274906174	0, 0%	1733562, 0.630601%
Genlisea aurea	43357795	0, 0%	35770, 0.0824996%
Mimulus guttatus	321641293	0, 0%	851660, 0.264786%
Conyza canadensis	326165195	0, 0%	403314, 0.123653%
ECHINOIDEA			
Lytechinus variegatus	951759914	0, 0%	1010888, 0.106213%
Strongylocentrotus purpuratus	936564995	120641, 0.0128812%	1788637, 0.190978%
ASTEROIDEA			
Patiria miniata	811028858	0, 0%	108236, 0.0133455%
ENTEROPNEUSTA			
Saccoglossus kowalevskii	775840678	0, 0%	977220, 0.125956%
TUNICATA			
Ciona intestinalis	115212710	0, 0%	12183, 0.0105744%
Ciona savignyi	587352817	24890, 0.00423766%	1990976, 0.338974%
Botryllus schlosseri	579633380	0, 0%	10137, 0.00174886%
LEPTOCARDII			
Branchiostoma floridae	521895125	0, 0%	283235, 0.0542705%
CEPHALASPIDOMORPHI			

Species	**Genome base count**	**BovB base count, percentage of genome**	**L1 base count, percentage of genome**
Lethenteron camtschaticum	1030662718	0, 0%	358081, 0.0347428%
Petromyzon marinus	885550958	0, 0%	349957, 0.0395186%
SARCOPTERYGII			
Latimeria chalumnae	2860591921	0, 0%	6249848, 0.218481%

Figure 2: Phylogeny showing two distinct BovB lineages in bats

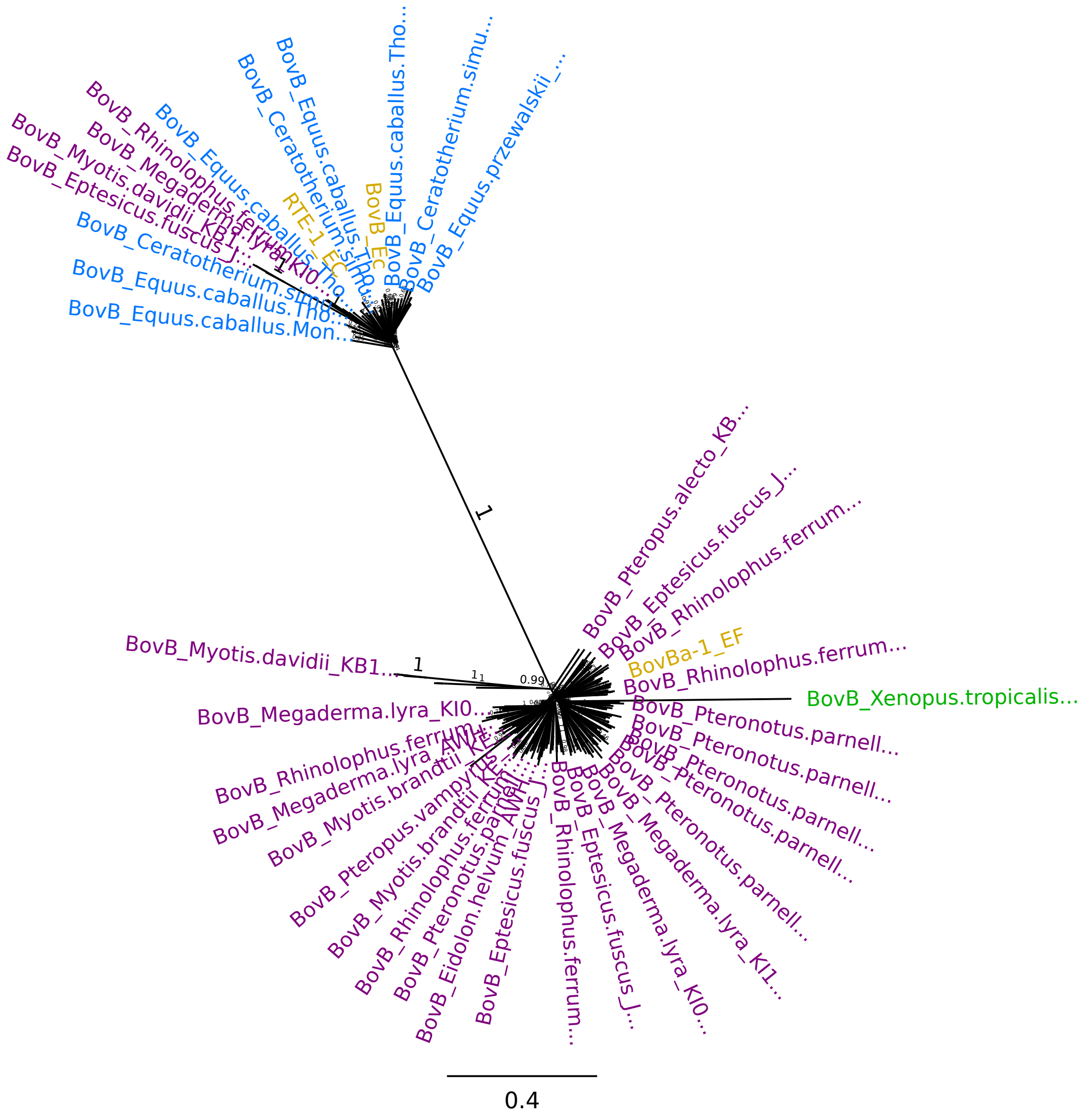


Figure C.2: **BovB lineages in bats.** Maximum likelihood tree inferred from 290 full-length nucleotide BovB sequences. Node labels are coloured to represent the species of origin: blue (perissodactyls), purple (bats), green (frog). RepBase sequences are coloured in light brown. MUSCLE was used to align sequences and FastTree was used to infer the phylogeny. This figure presents additional support for the two distinct BovB bat lineages, discussed in the manuscript.

Figure 3: Examples from the flanking region check

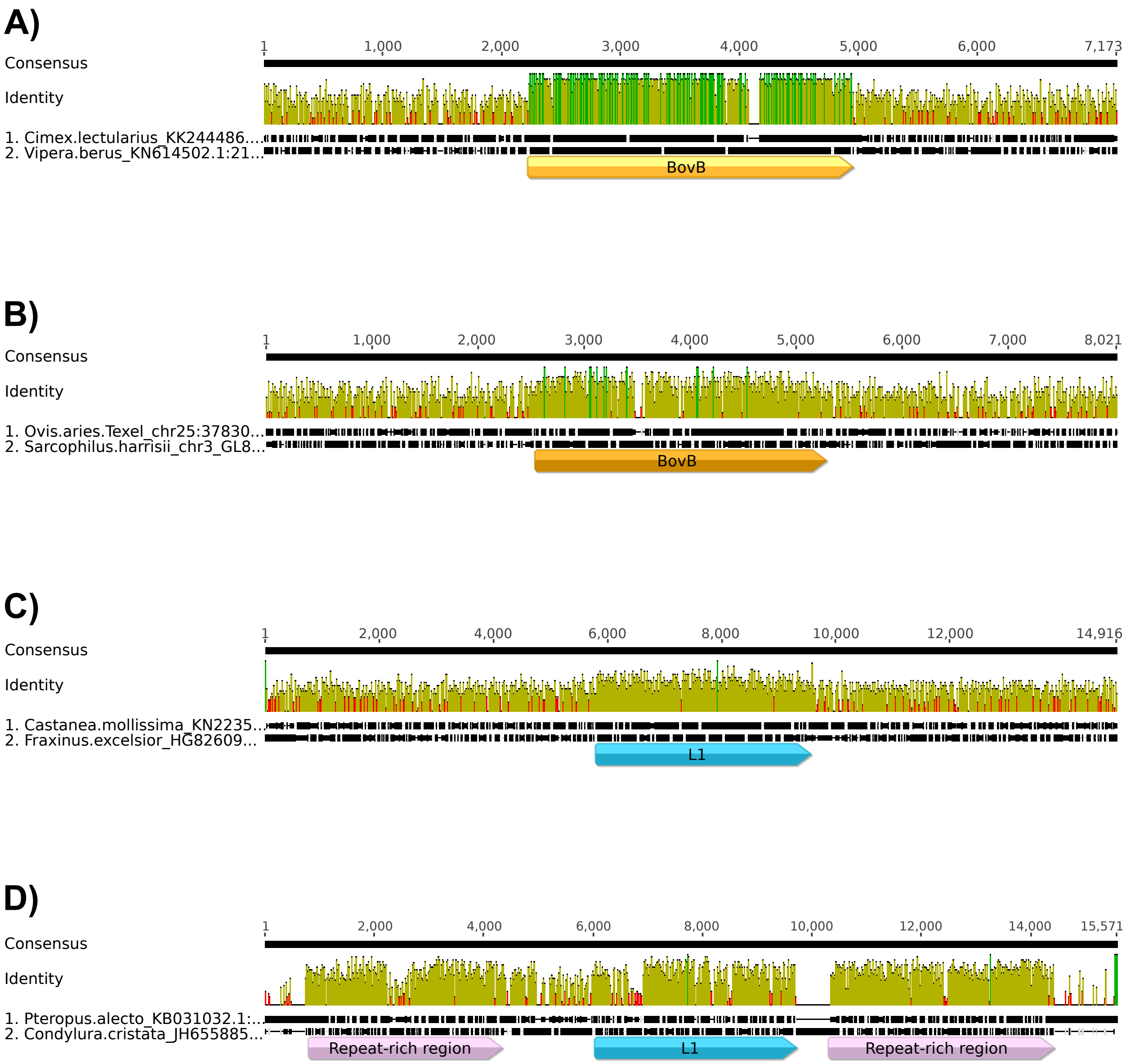


Figure C.3: **Flanking region checks.** (a) BovB FAM165: pairwise alignment of BovB HT candidates from bed bug *Cimex lectularius* and snake *Vipera berus*, including flanking regions. The high sequence similarity (>80%) is restricted to the transferred BovB. This is an example of a recent HT event. (b) BovB FAM14: pairwise alignment of BovB HT candidates from sheep *Ovis aries* and Tasmanian Devil *Sarcophilus harrisii*, including flanking regions. The sequence similarity (>50%) is still restricted to the transferred BovB, but much less obviously. This is an example of an ancient HT event. (c) L1 FAM802: pairwise alignment of L1 HT candidates from plants *Castanea mollissima* and *Fraxinus excelsior*, including flanking regions. As in (b), the identity is highest in the L1 region, but the difference is not obvious. This is an ancient L1 HT event. (d) L1 FAM1884: pairwise alignment of L1 HT candidates from bat *Pteropus alecto* and mole *Condyulra cristata*, including flanking regions. The L1 candidate pair is nestled in an orthologous, repeat-rich section of the genome. This is not a HT event.

Table 6: HT candidate families - BovB

Table C.6: **BovB HT families:** After the all-against-all BLAST and SiLiX procedure (at 50% identity), 215 BovB HT candidate clusters were identified. These clusters contain BovBs from at least two different species. The table below shows the 22 clusters which crossed between eularyotic Orders (e.g. *Anolis* lizard BovB with ruminant BovBs), and passed all the *in silico* validation tests.

Family	Number of BovBs from each species	Orders/clades
FAM1	2 *Anolis carolinensis*, 7364 *Bison bison*, 8752 *Bos indicus*, 4340 *Bos mutus*, 9561 *Bos taurus*, 6718 *Bubalus bubalis*, 4778 *Capra hircus*, 3 *Macropus eugenii*, 4 *Monodelphis domestica*, 2 *Ophiophagus hannah*, 4305 *Ovis aries musimon*, 5176 *Ovis aries Texel*, 2403 *Pantholops hodgsonii*, 39 *Pogona vitticeps*, 10 *Sarcophilus harrisi*i, 1 *Vipera berus*	Ruminants/ Squamates/ Marsupials
FAM2	208 *Bison bison*, 207 *Bos indicus*, 267 *Bos mutus*, 230 *Bos taurus*, 236 *Bubalus bubalis*, 261 *Capra hircus*, 1 *Macropus eugenii*, 232 *Ovis aries musimon*, 240 *Ovis aries Texel*, 255 *Pantholops hodgsonii*, 1 *Pogona vitticeps*, 3 *Vipera berus*	Ruminants/ Marsupial/ Squamates
FAM4	1 *Anolis carolinensis*, 1040 *Bison bison*, 1027 *Bos indicus*, 1123 *Bos mutus*, 1109 *Bos taurus*, 1130 *Bubalus bubalis*, 1062 *Capra hircus*, 16 *Macropus eugenii*, 3 *Monodelphis domestica*, 14 *Ophiophagus hannah*, 1003 *Ovis aries musimon*, 1085 *Ovis aries Texel*, 1028 *Pantholops hodgsonii*, 71 *Pogona vitticeps*, 7 *Sarcophilus harrisii*, 26 *Vipera berus*	Squamates/ Ruminants/ Marsupials
FAM5	2 *Anolis carolinensis*, 198 *Bison bison*, 189 *Bos indicus*, 235 *Bos mutus*, 210 *Bos taurus*, 216 *Bubalus bubalis*, 191 *Capra hircus*, 184 *Ovis aries musimon*, 201 *Ovis aries Texel*, 187 *Pantholops hodgsonii*, 3 *Pogona vitticeps*	Lizards/ Ruminants
FAM8	1486 *Bison bison*, 1238 *Bos indicus*, 478 *Bos mutus*, 1245 *Bos taurus*, 1101 *Bubalus bubalis*, 327 *Capra hircus*, 462 *Ovis aries musimon*, 366 *Ovis aries Texel*, 219 *Pantholops hodgsonii*, 1 *Pogona vitticeps*	Lizard/ Ruminants
FAM11	553 *Bison bison*, 270 *Bos indicus*, 425 *Bos mutus*, 289 *Bos taurus*, 566 *Bubalus bubalis*, 269 *Capra hircus*, 1 *Macropus eugenii*, 372 *Ovis aries musimon*, 213 *Ovis aries Texel*, 308 *Pantholops hodgsonii*, 4 *Pogona vitticeps*	Lizard/ Marsupial/ Ruminants
FAM14	22 *Bison bison*, 20 *Bos indicus*, 24 *Bos mutus*, 25 *Bos taurus*, 23 *Bubalus bubalis*, 24 *Capra hircus*, 20 *Ovis aries musimon*, 29 *Ovis aries Texel*, 32 *Pantholops hodgsonii*, 3 *Sarcophilus harrisii*	Marsupial/ Ruminants
FAM18	67 *Bison bison*, 61 *Bos indicus*, 65 *Bos mutus*, 68 *Bos taurus*, 69 *Bubalus bubalis*, 75 *Capra hircus*, 1 *Macropus eugenii*, 49 *Ovis aries musimon*, 61 *Ovis aries Texel*, 54 *Pantholops hodgsonii*, 1 *Sarcophilus harrisii*	Marsupials/ Ruminants
FAM20	2357 *Bison bison*, 2362 *Bos indicus*, 815 *Bos mutus*, 2448 *Bos taurus*, 1929 *Bubalus bubalis*, 647 *Capra hircus*, 926 *Ovis aries musimon*, 1044 *Ovis aries Texel*, 404 *Pantholops hodgsonii*, 8 *Pogona vitticeps*, 1 *Vipera berus*	Ruminants/ Squamates
FAM24	49 *Bison bison*, 36 *Bos indicus*, 55 *Bos mutus*, 40 *Bos taurus*, 52 *Bubalus bubalis*, 35 *Capra hircus*, 1 *Monodelphis domestica*, 37 *Ovis aries musimon*, 41 *Ovis aries Texel*, 42 *Pantholops hodgsonii*, 1 *Sarcophilus harrisii*, 2 *Vipera berus*	Snake/ Ruminants/ Marsupials
FAM27	32 *Bison bison*, 32 *Bos indicus*, 49 *Bos mutus*, 36 *Bos taurus*, 38 *Bubalus bubalis*, 33 *Capra hircus*, 1 *Monodelphis domestica*, 29 *Ovis aries musimon*, 29 *Ovis aries Texel*, 54 *Pantholops hodgsonii*, 1 *Python bivittatus*	Snake/ Marsupial/ Ruminants

Family	Number of BovBs from each species	Orders/clades
FAM28	97 *Bison bison*, 84 *Bos indicus*, 107 *Bos mutus*, 89 *Bos taurus*, 94 *Bubalus bubalis*, 97 *Capra hircus*, 1 *Macropus eugenii*, 87 *Ovis aries musimon*, 83 *Ovis aries Texel*, 94 *Pantholops hodgsonii*	Marsupial/ Ruminants
FAM39	9012 *Chrysochloris asiatica*, 5 *Echinops telfairi*, 1 *Lepisosteus oculatus*, 1 *Ornithorhynchus anatinus*	Fish/ Monotreme/ Afrotherians
FAM53	42 *Bison bison*, 45 *Bos indicus*, 51 *Bos mutus*, 47 *Bos taurus*, 52 *Bubalus bubalis*, 45 *Capra hircus*, 1 *Ophiophagus hannah*, 47 *Ovis aries musimon*, 47 *Ovis aries Texel*, 56 *Pantholops hodgsonii*, 1 *Python bivittatus*, 1 *Sarcophilus harrisii*	Marsupial/ Snakes/ Ruminants
FAM54	2 *Bison bison*, 5 *Bos indicus*, 4 *Bos mutus*, 3 *Bos taurus*, 3 *Bubalus bubalis*, 5 *Capra hircus*, 7 *Ovis aries musimon*, 7 *Pantholops hodgsonii*, 1 *Pogona vitticeps*	Lizard/ Ruminants
FAM83	4 *Bison bison*, 2 *Bos indicus*, 1 *Bos mutus*, 2 *Bos taurus*, 1 *Bubalus bubalis*, 3 *Capra hircus*, 1 *Ovis aries musimon*, 1 *Ovis aries Texel*, 1 *Pantholops hodgsonii*, 1 *Pogona vitticeps*	Lizard/ Ruminants
FAM107	41 *Ceratotherium simum*, 26 *Equus caballus Mongolian*, 31 *Equus caballus Thoroughbred*, 31 *Equus przewalskii*, 1 *Megaderma lyra*, 7 *Pogona vitticeps*, 1 *Pteronotus parnellii*, 2 *Rhinolophus ferrumequinum*	Lizard/ Bats/ Perissodactyls
FAM131	1 *Bison bison*, 3 *Bos indicus*, 2 *Bos mutus*, 5 *Bos taurus*, 3 *Bubalus bubalis*, 5 *Capra hircus*, 5 *Ovis aries musimon*, 7 *Ovis aries Texel*, 1 *Pantholops hodgsonii*, 1 *Sarcophilus harrisii*	Marsupial/ Ruminants
FAM165	1 *Cimex lectularius*, 9 *Ophiophagus hannah*, 1 *Python bivittatus*, 54 *Vipera berus*	Arthropod/ Snakes
FAM174	3 *Crotalus mitchellii*, 1 *Helobdella robusta*, 4 *Ophiophagus hannah*, 1 *Pogona vitticeps*, 26 *Python bivittatus*	Annelid/ Squamates
FAM199	4 *Bison bison*, 3 *Bos indicus*, 1 *Bos mutus*, 3 *Bos taurus*, 2 *Bubalus bubalis*, 1 *Capra hircus*, 1 *Ovis aries musimon*, 2 *Ovis aries Texel*, 3 *Pantholops hodgsonii*, 2 *Pogona vitticeps*	Lizard/ Ruminants
FAM233	1 *Danio rerio*, 1 *Helobdella robusta*, 3 *Mesobuthus martensii*	Annelid/ Fish/ Arthropod

Table 7: Orthologous families - L1

Table C.7: **L1 ortholog families:** The all-against-all clustering method, at 50% identity, produced 2815 L1 HT candidate families (i.e. families that contained L1s from at least two different species). The vast majority of these were from mammals (and a few non-mammal). All of the animal L1 candidates failed the flanking region check. The table below shows some examples of what these families looked like. Each family only contains 1 L1 from each species. In contrast, the BovB HT families and L1 plant HT families all contain multiple elements from at least one species. Accordingly, the number of elements from each species in a family should be taken into account when screening for HT events.

Family	Number of L1s from each species	Orders/clades

Family	Number of L1s from each species	Orders/clades
FAM3	1 *Acanthisitta chloris*, 1 *Apaloderma vittatum*, 1 *Aptenodytes forsteri*, 1 *Aquila chrysaetos*, 1 *Calypte anna*, 1 *Cariama cristata*, 1 *Cathartes aura*, 1 *Chaetura pelagica*, 1 *Chlamydotis macqueenii*, 1 *Colius striatus*, 1 *Columba livia*, 1 *Cuculus canorus*, 1 *Egretta garzetta*, 1 *Falco cherrug*, 1 *Falco peregrinus*, 1 *Fulmarus glacialis*, 1 *Gavia stellata*, 1 *Haliaeetus albicilla*, 1 *Haliaeetus leucocephalus*, 1 *Leptosomus discolour*, 1 *Manacus vitellinus*, 1 *Merops nubicus*, 1 *Mesitornis unicolor*, 1 *Nestor notabilis*, 1 *Nipponia nippon*, 1 *Opisthocomus hoazin*, 1 *Pelecanus crispus*, 1 *Phaethon lepturus*, 1 *Phalacrocorax carbo*, 1 *Phoenicopterus ruber*, 1 *Pterocles gutturalis*, 1 *Pygoscelis adeliae*, 1 *Tauraco erythrolophus*, 1 *Tyto alba*	Various bird Orders
FAM1174	1 *Canis lupus*, 1 *Felis catus*, 1 *Lipotes vexillifer*, 1 *Mustela putorius*, 1 *Orcinus orca*, 1 *Panthera tigris*, 1 *Tursiops truncatus*, 1 *Ursus maritimus*	Carnivores/ Artiodactyls
FAM1735	1 *Chlorocebus sabaeus*, 1 *Homo sapiens*, 1 *Myotis brandtii*, 1 *Myotis davidii*, 1 *Myotis lucifugus*, 1 *Nomascus leucogenys*, 1 *Pteropus alecto*, 1 *Rhinopithecus roxellana*	Bats/ Primates
FAM2296	1 *Equus caballus Mongolian*, 1 *Lipotes vexillifer*, 1 *Nomascus leucogenys*, 1 *Orcinus orca*, 1 *Pan paniscus*, 1 *Tursiops truncatus*	Equid/ Primates/ Artiodactyls

Table 8: HT candidate families - L1

Table C.8: **L1 HT candidate families:** Only 4 plant families showed evidence of L1 horizontal transfer across eukaryotic Orders. Each of these passed the flanking region check and *in silico* validation.

Family	Number of L1s from each species	Orders/clades
FAM802	89 *Betula nana*, 142 *Castanea mollissima*, 113 *Fraxinus excelsior*	Fagales/ Lamiales
FAM922	2 *Brassica napus*, 206 *Capsicum annuum*, 263 *Nicotiana sylvestris*, 367 *Nicotiana tomentosiformis*, 17 *Solanum arcanum*, 16 *Solanum habrochaites*, 13 *Solanum lycopersicum*, 3 *Solanum melongena*, 20 *Solanum pennellii*, 12 *Solanum pimpinellifolium*, 56 *Solanum tuberosum*	Brassicales/ Solanales
FAM925	2 *Brassica napus*, 3 *Capsicum annuum*, 152 *Nicotiana sylvestris*, 97 *Nicotiana tomentosiformis*, 7 *Solanum arcanum*, 7 *Solanum habrochaites*, 1 *Solanum lycopersicum*, 27 *Solanum melongena*, 7 *Solanum pennellii*, 7 *Solanum pimpinellifolium*, 23 *Solanum tuberosum*	Brassicales/ Solanales
FAM3016	1 *Medicago truncatula*, 7 *Spinacia oleracea*	Fabales/ Caryophyllales

Appendix D

Supplementary for Chapter 4

Figure 1: Phylogeny of elephants based on SNP data

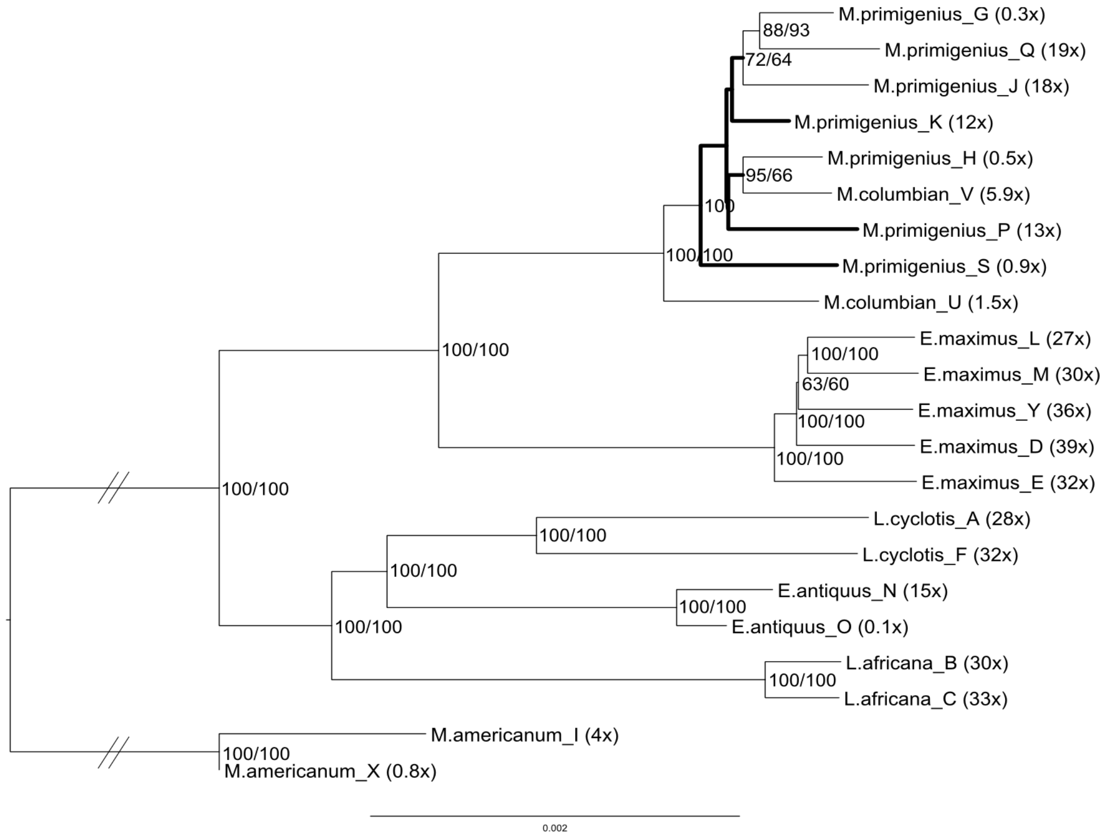


Figure D.1: **Evolutionary relationships between elephants based on SNP data:** Pairwise distance neighbour-joining tree, provided by Elle Palkopoulou (David Reich lab).

Figure 2: Example of a variant site

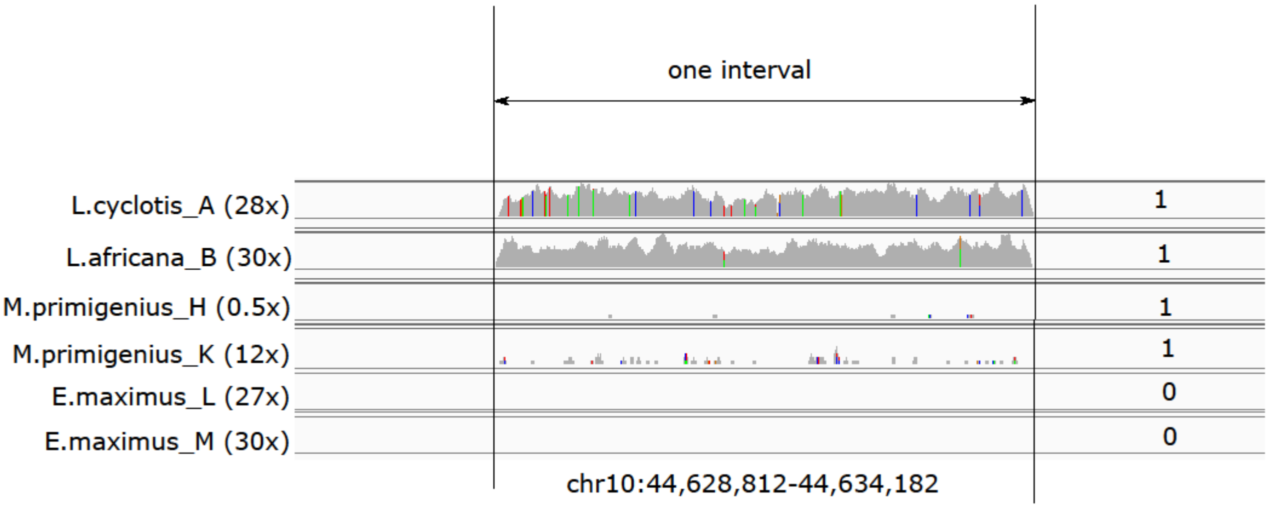


Figure D.2: **Variant site example:** Shows an interval which would be classified 'variant' because it is present in some elephants (labelled '1') but absent in others (labelled '0'). In this trivial example, an interval is 'present' if there is at least 1 bp in the specified region.

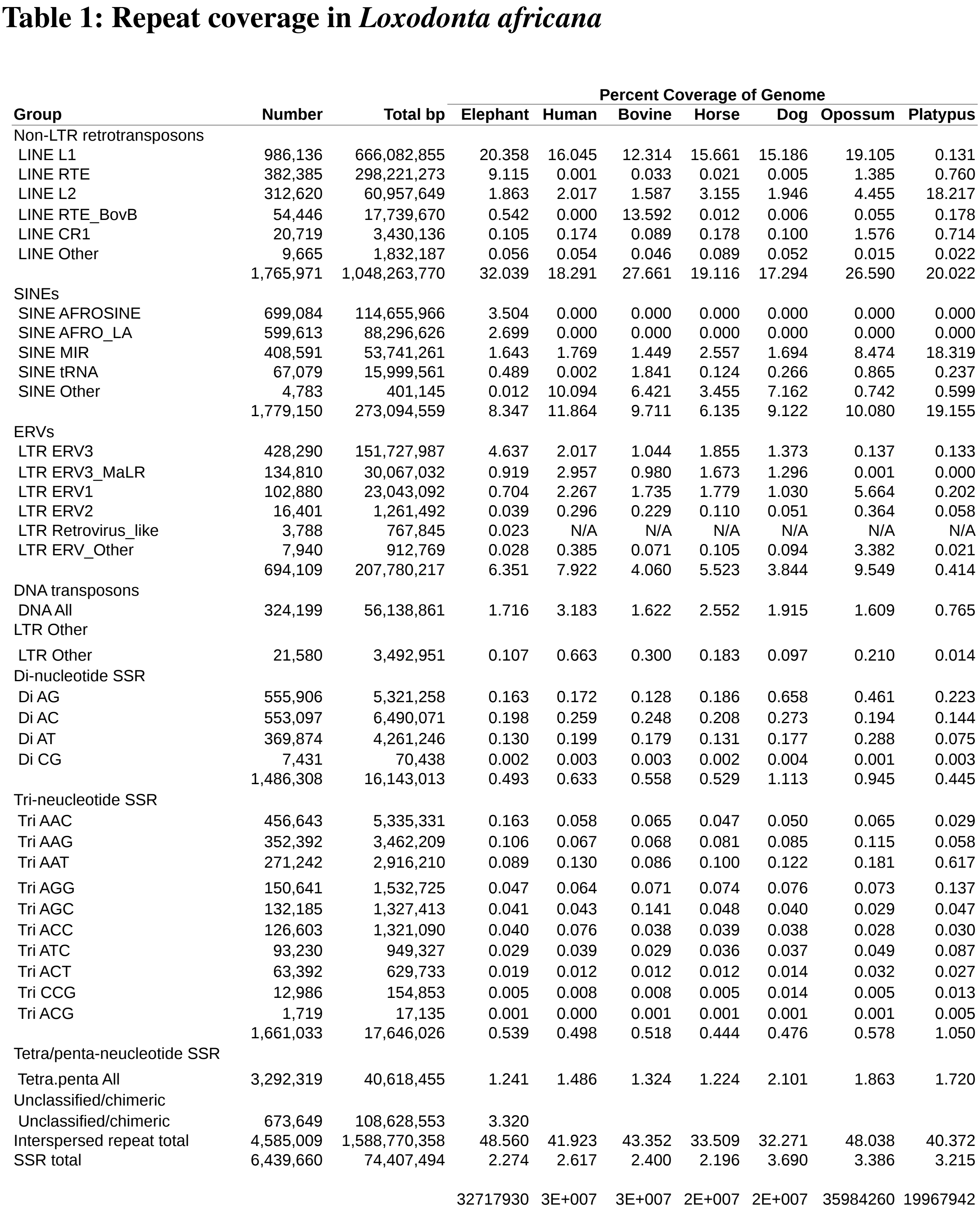


Table 1: Repeat coverage in *Loxodonta africana*

			Percent Coverage of Genome						
Group	**Number**	**Total bp**	**Elephant**	**Human**	**Bovine**	**Horse**	**Dog**	**Opossum**	**Platypus**
Non-LTR retrotransposons									
LINE L1	986,136	666,082,855	20.358	16.045	12.314	15.661	15.186	19.105	0.131
LINE RTE	382,385	298,221,273	9.115	0.001	0.033	0.021	0.005	1.385	0.760
LINE L2	312,620	60,957,649	1.863	2.017	1.587	3.155	1.946	4.455	18.217
LINE RTE_BovB	54,446	17,739,670	0.542	0.000	13.592	0.012	0.006	0.055	0.178
LINE CR1	20,719	3,430,136	0.105	0.174	0.089	0.178	0.100	1.576	0.714
LINE Other	9,665	1,832,187	0.056	0.054	0.046	0.089	0.052	0.015	0.022
	1,765,971	1,048,263,770	32.039	18.291	27.661	19.116	17.294	26.590	20.022
SINEs									
SINE AFROSINE	699,084	114,655,966	3.504	0.000	0.000	0.000	0.000	0.000	0.000
SINE AFRO_LA	599,613	88,296,626	2.699	0.000	0.000	0.000	0.000	0.000	0.000
SINE MIR	408,591	53,741,261	1.643	1.769	1.449	2.557	1.694	8.474	18.319
SINE tRNA	67,079	15,999,561	0.489	0.002	1.841	0.124	0.266	0.865	0.237
SINE Other	4,783	401,145	0.012	10.094	6.421	3.455	7.162	0.742	0.599
	1,779,150	273,094,559	8.347	11.864	9.711	6.135	9.122	10.080	19.155
ERVs									
LTR ERV3	428,290	151,727,987	4.637	2.017	1.044	1.855	1.373	0.137	0.133
LTR ERV3_MaLR	134,810	30,067,032	0.919	2.957	0.980	1.673	1.296	0.001	0.000
LTR ERV1	102,880	23,043,092	0.704	2.267	1.735	1.779	1.030	5.664	0.202
LTR ERV2	16,401	1,261,492	0.039	0.296	0.229	0.110	0.051	0.364	0.058
LTR Retrovirus_like	3,788	767,845	0.023	N/A	N/A	N/A	N/A	N/A	N/A
LTR ERV_Other	7,940	912,769	0.028	0.385	0.071	0.105	0.094	3.382	0.021
	694,109	207,780,217	6.351	7.922	4.060	5.523	3.844	9.549	0.414
DNA transposons									
DNA All	324,199	56,138,861	1.716	3.183	1.622	2.552	1.915	1.609	0.765
LTR Other									
LTR Other	21,580	3,492,951	0.107	0.663	0.300	0.183	0.097	0.210	0.014
Di-nucleotide SSR									
Di AG	555,906	5,321,258	0.163	0.172	0.128	0.186	0.658	0.461	0.223
Di AC	553,097	6,490,071	0.198	0.259	0.248	0.208	0.273	0.194	0.144
Di AT	369,874	4,261,246	0.130	0.199	0.179	0.131	0.177	0.288	0.075
Di CG	7,431	70,438	0.002	0.003	0.003	0.002	0.004	0.001	0.003
	1,486,308	16,143,013	0.493	0.633	0.558	0.529	1.113	0.945	0.445
Tri-neucleotide SSR									
Tri AAC	456,643	5,335,331	0.163	0.058	0.065	0.047	0.050	0.065	0.029
Tri AAG	352,392	3,462,209	0.106	0.067	0.068	0.081	0.085	0.115	0.058
Tri AAT	271,242	2,916,210	0.089	0.130	0.086	0.100	0.122	0.181	0.617
Tri AGG	150,641	1,532,725	0.047	0.064	0.071	0.074	0.076	0.073	0.137
Tri AGC	132,185	1,327,413	0.041	0.043	0.141	0.048	0.040	0.029	0.047
Tri ACC	126,603	1,321,090	0.040	0.076	0.038	0.039	0.038	0.028	0.030
Tri ATC	93,230	949,327	0.029	0.039	0.029	0.036	0.037	0.049	0.087
Tri ACT	63,392	629,733	0.019	0.012	0.012	0.012	0.014	0.032	0.027
Tri CCG	12,986	154,853	0.005	0.008	0.008	0.005	0.014	0.005	0.013
Tri ACG	1,719	17,135	0.001	0.000	0.001	0.001	0.001	0.001	0.005
	1,661,033	17,646,026	0.539	0.498	0.518	0.444	0.476	0.578	1.050
Tetra/penta-neucleotide SSR									
Tetra.penta All	3,292,319	40,618,455	1.241	1.486	1.324	1.224	2.101	1.863	1.720
Unclassified/chimeric									
Unclassified/chimeric	673,649	108,628,553	3.320						
Interspersed repeat total	4,585,009	1,588,770,358	48.560	41.923	43.352	33.509	32.271	48.038	40.372
SSR total	6,439,660	74,407,494	2.274	2.617	2.400	2.196	3.690	3.386	3.215
			32717930	3E+007	3E+007	2E+007	2E+007	35984260	19967942

Figure D.3: **Repeat coverage:** Simple and interspersed repeats in the reference elephant genome (LA4v2), as annotated by CENSOR.

Figure 3: Correlations among repeat groups in *Loxodonta africana*

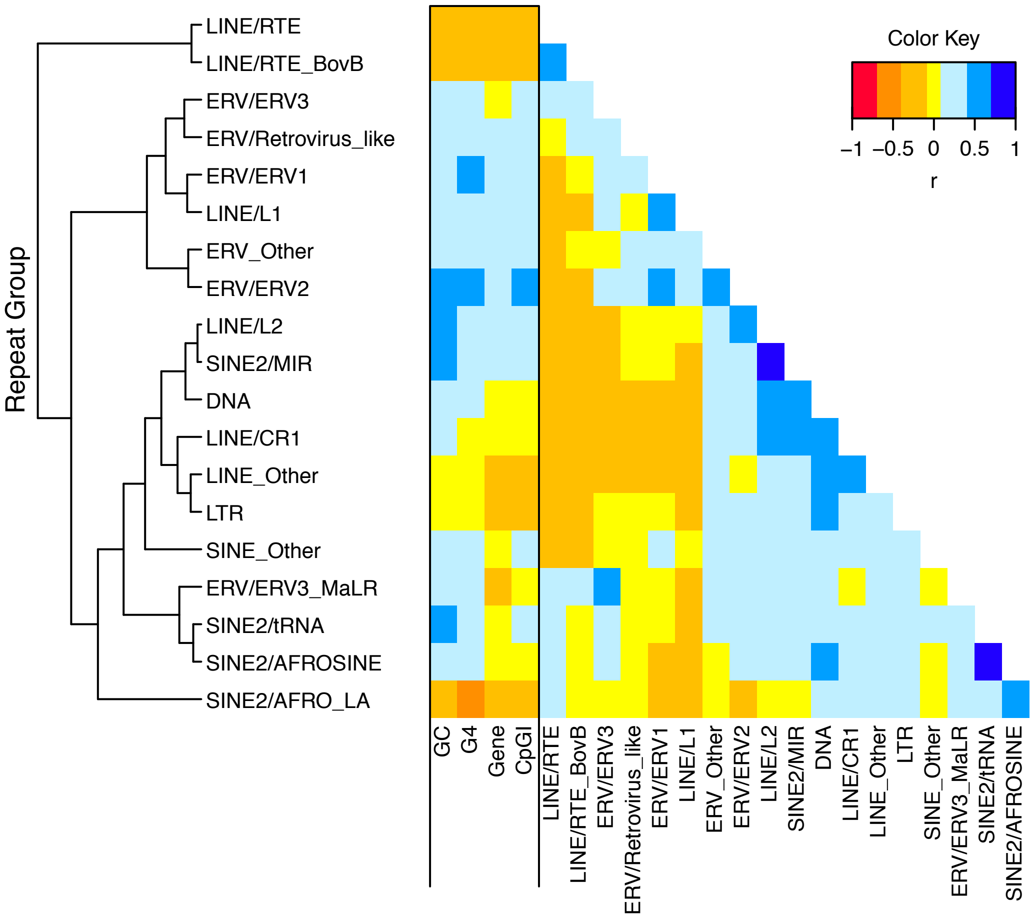


Figure D.4: **Correlations among repeat groups:** Heatmap showing correlations between repeats groups in the reference elephant genome (LA4v2).

Figure 4: Ancient-ness classification in *Loxodonta africana*

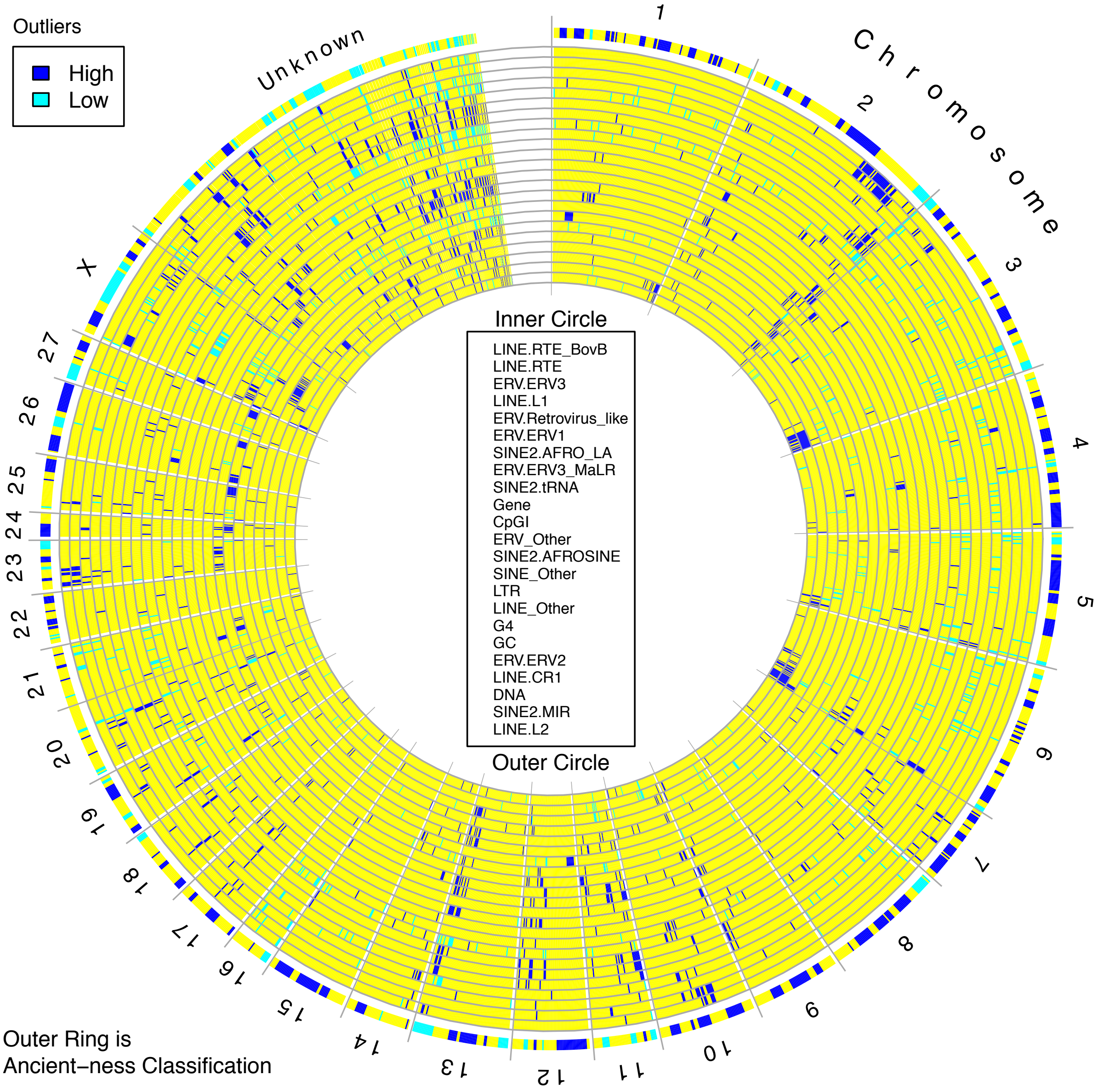
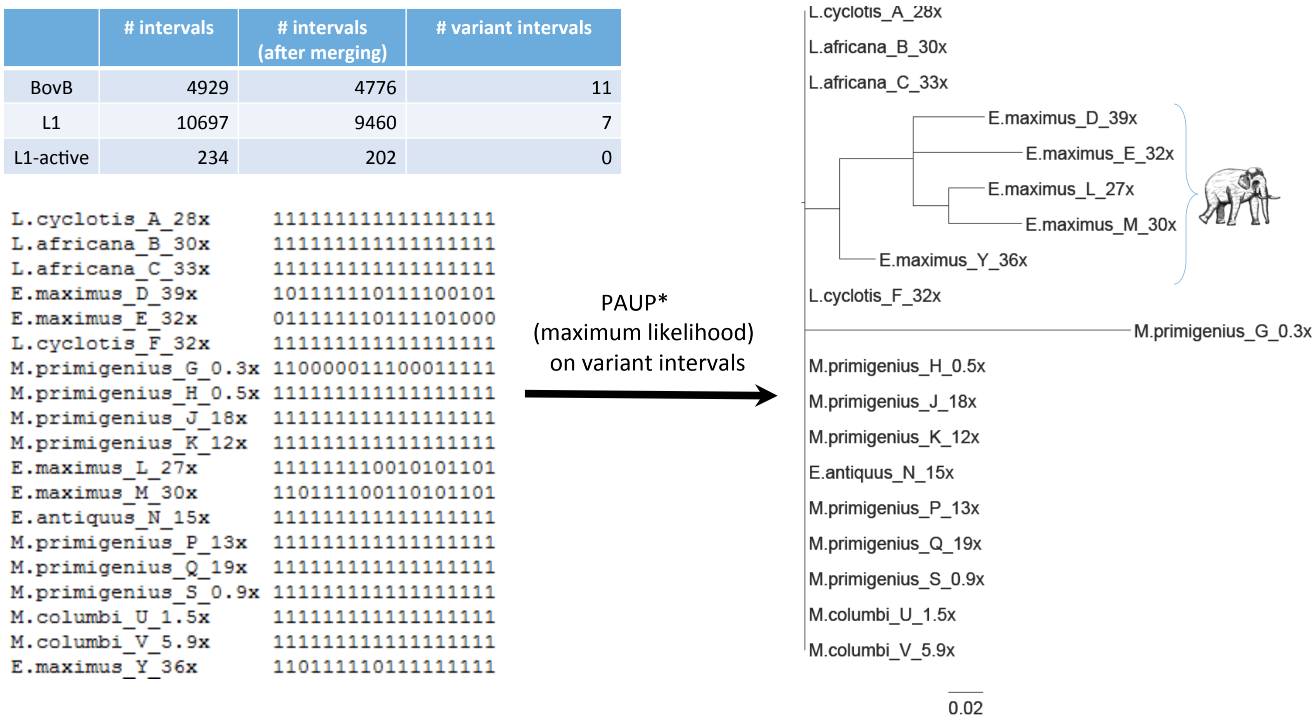


Figure D.5: **Ancient-ness classification and feature outlier bins:** Classification of repeats into high and low density regions in the reference elephant genome (LA4v2).

Figure 5: Initial test set of full-length LINEs



	# intervals	# intervals (after merging)	# variant intervals
BovB	4929	4776	11
L1	10697	9460	7
L1-active	234	202	0

```
L.cyclotis_A_28x       111111111111111111
L.africana_B_30x       111111111111111111
L.africana_C_33x       111111111111111111
E.maximus_D_39x        101111110111100101
E.maximus_E_32x        011111110111101000
L.cyclotis_F_32x       111111111111111111
M.primigenius_G_0.3x 110000011100011111
M.primigenius_H_0.5x 111111111111111111
M.primigenius_J_18x  111111111111111111
M.primigenius_K_12x  111111111111111111
E.maximus_L_27x        111111110010101101
E.maximus_M_30x        110111100110101101
E.antiquus_N_15x       111111111111111111
M.primigenius_P_13x  111111111111111111
M.primigenius_Q_19x  111111111111111111
M.primigenius_S_0.9x 111111111111111111
M.columbi_U_1.5x       111111111111111111
M.columbi_V_5.9x       111111111111111111
E.maximus_Y_36x        110111110111111111
```

Figure D.6: **Results from the initial subset of full-length LINEs:** Only 18 variant sites were found using this subset. The subsequent phylogeny (inferred using PAUP, maximum likelihood) is only useful in regards to the *E. maximus* elephants.

Bibliography

[1] S. Schaack, C. Gilbert, and C. Feschotte, “Promiscuous dna: horizontal transfer of transposable elements and why it matters for eukaryotic evolution,” *Trends Ecol Evol*, vol. 25, no. 9, pp. 537–46, 2010.

[2] E. L. Loreto, C. M. Carareto, and P. Capy, “Revisiting horizontal transfer of transposable elements in drosophila,” *Heredity (Edinb)*, vol. 100, no. 6, pp. 545–54, 2008.

[3] S. B. Daniels, L. D. Strausbaugh, L. Ehrman, and R. Armstrong, “Sequences homologous to p elements occur in drosophila paulistorum,” *Proc Natl Acad Sci U S A*, vol. 81, no. 21, pp. 6794–7, 1984.

[4] S. B. Daniels, K. R. Peterson, L. D. Strausbaugh, M. G. Kidwell, and A. Chovnick, “Evidence for horizontal transmission of the p transposable element between drosophila species,” *Genetics*, vol. 124, no. 2, pp. 339–55, 1990.

[5] J. P. de Castro and C. M. Carareto, “Canonical p elements are transcriptionally active in the saltans group of drosophila,” *J Mol Evol*, vol. 59, no. 1, pp. 31–40, 2004.

[6] M. A. Houck, J. B. Clark, K. R. Peterson, and M. G. Kidwell, “Possible horizontal transfer of drosophila genes by the mite proctolaelaps regalis,” *Science*, vol. 253, no. 5024, pp. 1125–8, 1991.

[7] E. L. Loreto, V. L. Valente, A. Zaha, J. C. Silva, and M. G. Kidwell, “Drosophila mediopunctata p elements: a new example of horizontal transfer,” *J Hered*, vol. 92, no. 5, pp. 375–81, 2001.

[8] J. C. Silva and M. G. Kidwell, “Horizontal transfer and selection in the evolution of p elements,” *Mol Biol Evol*, vol. 17, no. 10, pp. 1542–57, 2000.

[9] F. Brunet, F. Godin, C. Bazin, and P. Capy, “Phylogenetic analysis of mos1-like transposable elements in the drosophilidae,” *J Mol Evol*, vol. 49, no. 6, pp. 760–8, 1999.

[10] A. R. Lohe, E. N. Moriyama, D. A. Lidholm, and D. L. Hartl, “Horizontal transmission, vertical inactivation, and stochastic loss of mariner-like transposable elements,” *Mol Biol Evol*, vol. 12, no. 1, pp. 62–72, 1995.

[11] K. Maruyama and D. L. Hartl, “Evidence for interspecific transfer of the transposable element mariner between drosophila and zaprionus,” *J Mol Evol*, vol. 33, no. 6, pp. 514–24, 1991.

[12] H. M. Robertson and D. J. Lampe, “Recent horizontal transfer of a mariner transposable element among and between diptera and neuroptera,” *Mol Biol Evol*, vol. 12, no. 5, pp. 850–62, 1995.

[13] M. Turnbull and B. Webb, “Perspectives on polydnavirus origins and evolution,” *Adv Virus Res*, vol. 58, pp. 203–54, 2002.

[14] B. Arca and C. Savakis, “Distribution of the transposable element minos in the genus drosophila,” *Genetica*, vol. 108, no. 3, pp. 263–7, 2000.

[15] L. M. de Almeida and C. M. Carareto, “Multiple events of horizontal transfer of the minos transposable element between drosophila species,” *Mol Phylogenet Evol*, vol. 35, no. 3, pp. 583–94, 2005.

[16] C. Bartolome, X. Bello, and X. Maside, “Widespread evidence for horizontal transfer of transposable elements across drosophila genomes,” *Genome Biol*, vol. 10, no. 2, p. R22, 2009.

[17] A. Sanchez-Gracia, X. Maside, and B. Charlesworth, “High rate of horizontal transfer of transposable elements in drosophila,” *Trends Genet*, vol. 21, no. 4, pp. 200–3, 2005.

[18] A. F. Smit and A. D. Riggs, “Tiggers and dna transposon fossils in the human genome,” *Proc Natl Acad Sci U S A*, vol. 93, no. 4, pp. 1443–8, 1996.

[19] L. M. Gomulski, C. Torti, M. Bonizzoni, D. Moralli, E. Raimondi, P. Capy, G. Gasperi, and A. R. Malacrida, “A new basal subfamily of mariner elements in ceratitis rosa and other tephritid flies,” *J Mol Evol*, vol. 53, no. 6, pp. 597–606, 2001.

[20] Z. Ivics, Z. Izsvak, A. Minter, and P. B. Hackett, “Identification of functional domains and evolution of tc1-like transposable elements,” *Proc Natl Acad Sci U S A*, vol. 93, no. 10, pp. 5008–13, 1996.

[21] H. M. Robertson, “Multiple mariner transposons in flatworms and hydras are related to those of insects,” *J Hered*, vol. 88, no. 3, pp. 195–201, 1997.

[22] M. J. Leaver, “A family of tc1-like transposons from the genomes of fishes and frogs: evidence for horizontal transmission,” *Gene*, vol. 271, no. 2, pp. 203–14, 2001.

[23] J. K. Biedler, H. Shao, and Z. Tu, “Evolution and horizontal transfer of a dd37e dna transposon in mosquitoes,” *Genetics*, vol. 177, no. 4, pp. 2553–8, 2007.

[24] G. M. Simmons, “Horizontal transfer of hobo transposable elements within the drosophila melanogaster species complex: evidence from dna sequencing,” *Mol Biol Evol*, vol. 9, no. 6, pp. 1050–60, 1992.

[25] C. Torti, L. M. Gomulski, M. Bonizzoni, V. Murelli, D. Moralli, C. R. Guglielmino, E. Raimondi, D. Crisafulli, P. Capy, G. Gasperi, and A. R. Malacrida, “Cchobo, a hobo-related sequence in ceratitis capitata,” *Genetica*, vol. 123, no. 3, pp. 313–25, 2005.

[26] C. Gilbert, J. K. Pace, and C. Feschotte, “Horizontal spinning of transposons,” *Commun Integr Biol*, vol. 2, no. 2, pp. 117–9, 2009.

[27] C. Gilbert, S. Schaack, n. Pace, J. K., P. J. Brindley, and C. Feschotte, “A role for host-parasite interactions in the horizontal transfer of transposons across phyla,” *Nature*, vol. 464, no. 7293, pp. 1347–50, 2010.

[28] C. Gilbert, S. S. Hernandez, J. Flores-Benabib, E. N. Smith, and C. Feschotte, “Rampant horizontal transfer of spin transposons in squamate reptiles,” *Mol Biol Evol*, vol. 29, no. 2, pp. 503–15, 2012.

[29] n. Pace, J. K., C. Gilbert, M. S. Clark, and C. Feschotte, “Repeated horizontal transfer of a dna transposon in mammals and other tetrapods,” *Proc Natl Acad Sci U S A*, vol. 105, no. 44, pp. 17023–8, 2008.

[30] P. Novick, J. Smith, D. Ray, and S. Boissinot, “Independent and parallel lateral transfer of dna transposons in tetrapod genomes,” *Gene*, vol. 449, no. 1-2, pp. 85–94, 2010.

[31] C. Gilbert, P. Waters, C. Feschotte, and S. Schaack, “Horizontal transfer of oc1 transposons in the tasmanian devil,” *BMC Genomics*, vol. 14, no. 1, p. 134, 2013.

[32] A. Koga, A. Shimada, A. Shima, M. Sakaizumi, H. Tachida, and H. Hori, “Evidence for recent invasion of the medaka fish genome by the tol2 transposable element,” *Genetics*, vol. 155, no. 1, pp. 273–81, 2000.

[33] N. R. Mota, A. Ludwig, V. L. Valente, and E. L. Loreto, “Harrow: new drosophila hat transposons involved in horizontal transfer,” *Insect Mol Biol*, vol. 19, no. 2, pp. 217–28, 2010.

[34] M. Deprá, Y. Panzera, A. Ludwig, V. L. Valente, and E. L. Loreto, “hosimary: a new hat transposon group involved in horizontal transfer,” *Molecular Genetics and Genomics*, vol. 283, no. 5, pp. 451–459, 2010.

[35] E. A. Gladyshev and I. R. Arkhipova, “A single-copy is5-like transposon in the genome of a bdelloid rotifer,” *Molecular biology and evolution*, vol. 26, no. 8, pp. 1921–1929, 2009.

[36] H. J. Pagan, J. D. Smith, R. M. Hubley, and D. A. Ray, “Piggybac-ing on a primate genome: novel elements, recent activity and horizontal transfer,” *Genome biology and evolution*, vol. 2, pp. 293–303, 2010.

[37] J. Thomas, S. Schaack, and E. J. Pritham, “Pervasive horizontal transfer of rolling-circle transposons among animals,” *Genome biology and evolution*, vol. 2, pp. 656–664, 2010.

[38] N. de Setta, M.-A. Van Sluys, P. Capy, and C. M. Carareto, “Multiple invasions of gypsy and micropia retroelements in genus zaprionus and melanogaster subgroup of the genus drosophila,” *BMC evolutionary biology*, vol. 9, no. 1, p. 1, 2009.

[39] A. Ludwig and E. Loreto, “Evolutionary pattern of the gtwin retrotransposon in the drosophila melanogaster subgroup,” *Genetica*, vol. 130, no. 2, pp. 161–168, 2007.

[40] P. Gonzalez and H. Lessios, “Evolution of sea urchin retroviral-like (surl) elements: evidence from 40 echinoid species.,” *Molecular biology and evolution*, vol. 16, no. 7, pp. 938–952, 1999.

[41] L. M. De Almeida and C. M. Carareto, “Sequence heterogeneity and phylogenetic relationships between the copia retrotransposon in drosophila species of the repleta and melanogaster groups,” *Genetics Selection Evolution*, vol. 38, no. 5, p. 1, 2006.

[42] I. K. Jordan and J. F. McDonald, “Evolution of the copia retrotransposon in the drosophila melanogaster species subgroup.,” *Molecular biology and evolution*, vol. 15, no. 9, pp. 1160–1171, 1998.

[43] I. K. Jordan, L. V. Matyunina, and J. F. McDonald, “Evidence for the recent horizontal transfer of long terminal repeat retrotransposon,” *Proceedings of the National Academy of Sciences*, vol. 96, no. 22, pp. 12621–12625, 1999.

[44] M. Evgen’ev, H. Zelentsova, L. Mnjoian, H. Poluectova, and M. G. Kidwell, “Invasion of drosophila virilis by the penelope transposable element,” *Chromosoma*, vol. 109, no. 5, pp. 350–357, 2000.

[45] G. T. Lyozin, K. S. Makarova, V. V. Velikodvorskaja, H. S. Zelentsova, R. R. Khechumian, M. G. Kidwell, E. V. Koonin, and M. B. Evgen’ev, “The structure and evolution of penelope in the virilis species group of drosophila: an ancient lineage of retroelements,” *Journal of Molecular Evolution*, vol. 52, no. 5, pp. 445–456, 2001.

[46] R. Morales-Hojas, C. P. Vieira, and J. Vieira, “The evolutionary history of the transposable element penelope in the drosophila virilis group of species,” *Journal of molecular evolution*, vol. 63, no. 2, pp. 262–273, 2006.

[47] K. P. Gogolevsky, N. S. Vassetzky, and D. A. Kramerov, “Bov-b-mobilized sines in vertebrate genomes,” *Gene*, vol. 407, no. 1, pp. 75–85, 2008.

[48] V. Zupunski, F. Gubensek, and D. Kordis, “Evolutionary dynamics and evolutionary history in the rte clade of non-ltr retrotransposons,” *Mol Biol Evol*, vol. 18, no. 10, pp. 1849–63, 2001.

[49] M. Hamada, Y. Kido, M. Himberg, J. D. Reist, C. Ying, M. Hasegawas, and N. Okada, “A newly isolated family of short interspersed repetitive elements (sines) in coregonid fishes (whitefish)

with sequences that are almost identical to those of the smai family of repeats: possible evidence for the horizontal transfer of sines," *Genetics*, vol. 146, no. 1, pp. 355–367, 1997.

[50] J.-N. Volff, C. Körting, and M. Schartl, "Multiple lineages of the non-ltr retrotransposon rex1 with varying success in invading fish genomes," *Molecular Biology and Evolution*, vol. 17, no. 11, pp. 1673–1684, 2000.

[51] O. Novikova, E. Sliwinska, V. Fet, J. Settele, A. Blinov, and M. Woyciechowski, "Cr1 clade of non-ltr retrotransposons from maculinea butterflies (lepidoptera: Lycaenidae): evidence for recent horizontal transmission," *BMC Evol Biol*, vol. 7, p. 93, 2007.

[52] I. Sormacheva, G. Smyshlyaev, V. Mayorov, A. Blinov, A. Novikov, and O. Novikova, "Vertical evolution and horizontal transfer of cr1 non-ltr retrotransposons and tc1/mariner dna transposons in lepidoptera species," *Mol Biol Evol*, vol. 29, no. 12, pp. 3685–702, 2012.

[53] A. M. Walsh, R. D. Kortschak, M. G. Gardner, T. Bertozzi, and D. L. Adelson, "Widespread horizontal transfer of retrotransposons," *Proc Natl Acad Sci U S A*, vol. 110, no. 3, pp. 1012–6, 2013.

[54] A. M. Ivancevic, R. D. Kortschak, T. Bertozzi, and D. L. Adelson, "Lines between species: Evolutionary dynamics of line-1 retrotransposons across the eukaryotic tree of life," *Genome Biology and Evolution*, p. Advance Access, 2016.

[55] R. S. Harris, *Improved pairwise alignment of genomic DNA.*
ProQuest, 2007.

[56] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of molecular biology*, vol. 215, no. 3, pp. 403–410, 1990.

[57] O. Kohany, A. J. Gentles, L. Hankus, and J. Jurka, "Annotation, submission and screening of repetitive elements in repbase: Repbasesubmitter and censor," *BMC bioinformatics*, vol. 7, no. 1, p. 1, 2006.

[58] J. Jurka, V. V. Kapitonov, A. Pavlicek, P. Klonowski, O. Kohany, and J. Walichiewicz, "Repbase update, a database of eukaryotic repetitive elements," *Cytogenetic and genome research*, vol. 110, no. 1-4, pp. 462–467, 2005.

[59] R. C. Edgar, "Search and clustering orders of magnitude faster than blast," *Bioinformatics*, vol. 26, no. 19, pp. 2460–2461, 2010.

[60] R. D. Finn, J. Clements, and S. R. Eddy, "Hmmer web server: interactive sequence similarity searching," *Nucleic acids research*, p. gkr367, 2011.

[61] R. C. Edgar, "Muscle: multiple sequence alignment with high accuracy and high throughput," *Nucleic acids research*, vol. 32, no. 5, pp. 1792–1797, 2004.

[62] J. Castresana, “Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis,” *Molecular biology and evolution*, vol. 17, no. 4, pp. 540–552, 2000.

[63] M. N. Price, P. S. Dehal, and A. P. Arkin, “Fasttree 2–approximately maximum-likelihood trees for large alignments,” *PloS one*, vol. 5, no. 3, p. e9490, 2010.

[64] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, *et al.*, “Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data,” *Bioinformatics*, vol. 28, no. 12, pp. 1647–1649, 2012.

[65] M. El Baidouri, M.-C. Carpentier, R. Cooke, D. Gao, E. Lasserre, C. Llauro, M. Mirouze, N. Picault, S. A. Jackson, and O. Panaud, “Widespread and frequent horizontal transfers of transposable elements in plants,” *Genome research*, vol. 24, no. 5, pp. 831–838, 2014.

[66] V. Miele, S. Penel, and L. Duret, “Ultra-fast sequence clustering from similarity networks with silix,” *BMC bioinformatics*, vol. 12, no. 1, p. 1, 2011.