

Rebase 2023 Year in review

Authors: Kenji K. Kojima, Weidong Bao, Noriko F. Kojima, and Oleksiy Kohany

Affiliation: Genetic Information Research Institute, 20380 Town Center Lane Suite 240, Cupertino, CA 95014, USA

Correspondence: Kenji K. Kojima, kojima@girinst.org

Published: 2024/02/13

Update of Rebase 2023

In the year 2023, Rebase increased the number of entries by 10,992, including 8,364 sequences published on Rebase Reports. The total number of entries reached 92,586 at the end of the year 2023. The majority (91,014: 98%) of these entries are transposable elements (TEs), either consensus or representative sequences. The remaining are satellite repeats and microsatellites (602 entries), multicopy genes (197 entries), integrated viruses (220 entries), multicopy genes (197 entries) and uncharacterized repeats.

Rebase transitioned to a subscription-based model on April 12, 2019. The last version of Rebase before the transition was Rebase Update 24.03 which included 50,356 entries. Since then, we have (1) increased the number of repeats published in each issue of Rebase Reports, and (2) covered more diverse organisms (Table 1). The choice of organisms is based on (1) the economic and scientific importance, (2) the quality of the genome sequence, and (3) the phylogenetic distance from well-analyzed organisms. We have continually updated, corrected, and refined existing repeats that contain sequence flaws or are ambiguously classified.

Rebase has been expanding to be a large dataset covering various eukaryotic repeats. The three major phylogenetic groups of Rebase entries are vertebrates, arthropods, and green plants (Figure 1). Vertebrate includes humans, lab animals, and almost all domestic or farmed animals for meat, labor/transportation, and pets (cow, sheep, goat, pig, camel, llama, horse, cat, dog, Guinea pig, mice, hamsters, rabbits, chicken, duck, goldfish, carp, salmon, tilapia, catfish, etc.). Major crops are classified into only a few orders in green plants: Poales (wheat, barley, rye, oat, maize, sorghum, millets, sugarcane, rice), Brassicales (cabbage/kale/broccoli, Chinese cabbage, radish), Fabales (pea, common bean, soybean, peanut, chickpea, alfalfa), Rosales (apple, pear, cherry, plum, strawberry), and Solanales (potato, tomato, sweet potato).

Arthropod is the largest group of animals and counts for >80% of all living animal species. Arthropod includes a vast majority of pests for humans and domestic animals (mosquitoes, flies, flea, bugs, lice, ticks, etc.), and for crops (caterpillars, flies, sawflies, leaf beetles, aphids, whiteflies, locusts, etc.). Other noticeable groups are mollusks (including oyster, mussel, clam, and abalone), fungi (including yeasts, edible mushrooms, and plant pathogens such as *Puccinia* and *Blumeria*), and stramenopiles (including plant pathogens such as *Phytophthora*).

At the end of the year 2023, Repbase contains repeat sequences from over 1,900 species. Four TE-rich species (yellow fever mosquito, Asian rice, zebrafish, and maize) each have over 2,000 entries (Table 2). Sometimes, multiple genomes in the same genus, such as *Drosophila* and *Arabidopsis*, have been sequenced and analyzed. Such analysis helps characterize very low-copy TE families or single-copy TE families in model organisms and reveals the genome evolution over a longer time scale. Table 3 shows the total number of Repbase entries for each genus. In some cases, such as *Chondrus* and *Locusta*, only one species has been analyzed, while in other cases, such as *Drosophila*, more than 2 species have been examined (Table 4). Table 5 shows the top 20 species whose entries increased in the year 2023. Even though the genomes of model organisms are very well analyzed, they still contain TE families to be discovered. It is challenging to assert that all TEs, regardless of their age, have been identified within a single species. However, it is our ongoing goal to consistently assemble an expanding, and near-complete set of TEs in some crucial genomes.

The characteristics of several species we focused on in the year 2023 are described below.

Triticum aestivum (bread wheat)

Wheat is one of the most important staple crops globally. It is a major ingredient in bread, pasta, cereals, and cakes. Alcoholic beverages such as beer (although the primary grain used in most traditional beer styles is barley) and vodka are also made from wheat. The bread wheat (common wheat) *Triticum aestivum* is believed to have originated from the hybridization of two wild grass species in the Fertile Crescent, a region in the Middle East. The two ancestral grass species are likely the emmer wheat *Triticum dicoccum* and the wild goatgrass *Aegilops tauschii*. In the year 2023, the entries from *T. aestivum* increased by 1,060, and the total number of entries became 1,144. In addition to the bread wheat, several other wheat species are also cultivated: durum wheat *T. durum*, spelt *T. spelta*, emmer wheat *T. dicoccum*, and einkorn

wheat *T. monococcum*. Rebase contains 8 sequences from *T. durum*, 1 from *T. spelta*, 67 from *T. monococcum*, as well as 8 sequences from *Aegilops tauschii*.

Capsicum annuum (bell pepper)

Capsicum annuum belongs to the Solanaceae family, which includes various crops like tomatoes, potatoes, and tobacco. *Capsicum annuum* produces a diverse range of fruits, including bell peppers (green, red, yellow, orange, or purple in color), as well as various chili peppers, such as jalapeños, serranos, cayenne, and others. The heat levels of these chili peppers can vary significantly, from mild to extremely spicy. The increase of Rebase entries in the year 2023 was 359. Rebase also contains 67 entries from its sibling species *Capsicum baccatum*.

Crassostrea gigas (Pacific oyster)

The Pacific oyster (*Crassostrea gigas*) is now the most widely farmed and the most commercially important oyster globally. It is native to the Pacific coast of Asia. It is also considered an invasive species where it has been introduced intentionally or accidentally. In the year 2023, the entries from *C. gigas* increased by 172, and the total number of entries became 1,818.

Corbicula fluminea (Asian clam)

The Asian clam *Corbicula fluminea* is a species of freshwater clam native to eastern Asia and it has successfully invaded throughout North America, South America, and Europe. At the end of the year 2023, the entries from *C. fluminea* was 764. The class Bivalvia are now classified into 4 major subclasses (Heterodonta, Palaeoheterodonta, Protobranchia, and Pteriomorphia). *Corbicula* belongs to Heterodonta, and is distant from Pteriomorphia, which includes oysters and mussels. Heterodonta includes around 5,600 species, over a half of the around 9,200 bivalve species.

Hydra vulgaris (*H. magnipapillata*)

Hydra vulgaris, commonly known as the freshwater hydra, is a small, freshwater, solitary cnidarian belonging to the family Hydraidae. *Hydra vulgaris* serves as a model organism in biological research, particularly in studies related to regeneration, development, and the functioning of simple nervous systems. The entries in Rebase in the year 2023 increased by 258, and now exceed 1,000.

Drosophila (fruit flies)

The genus *Drosophila* contains over 1,450 species, among which ~1,100 species belong to the subgenus *Drosophila*. *D. melanogaster* belongs to the subgenus *Sophophora*. In the year 2023, 449 TE families from the genus *Drosophila* were added to Rebase (Table 4). In total, Rebase contains 4,699 repeat entries from *Drosophila*.

Rhyzopertha dominica (lesser grain borer)

Rhyzopertha dominica, commonly known as the lesser grain borer or the American wheat weevil, is a significant pest that affects stored grains, particularly wheat and other cereal grains. This pest can cause losses ranging from 5% to 30% in stored grain quantities. Its origin is likely the Indian subcontinent, but now it is distributed worldwide. In 2023, 513 entries from *R. dominica* were added. The only other coleopteran species with >100 entries is the rice weevil *Sitophilus oryzae* (188) despite the fact that Coleoptera is the largest order with 400,000 described species.

Cydia splendana (chestnut tortrix)

The chestnut tortrix *Cydia splendana*, also known as the European oak processionary moth or oak processionary, is a moth species belonging to the family Tortricidae. It is native to central and southern Europe but has become an invasive species in various regions. The caterpillars primarily feed on the leaves of oak trees. Severe infestations can lead to defoliation and stress in oak trees. The entire 349 entries from *C. splendana* were added in the year 2023. Lepidoptera is one of the most diverse insect orders, with approximately 180,000 described species. Other lepidopteran species with >100 Rebase entries are: *Spodoptera exigua* (223, Noctuidae), *Papilio polytes* (181, Papilionidae), *Heliconius melpomene* (179, Nymphalidae), *Habrosyne pyritoides* (143, Drepanidae), *Trichoplusia ni* (137, Noctuidae), *Manduca sexta* (126, Sphingidae), *Lymantria dispar* (114, Lymantriidae or Erebidae), and *Bombyx mori* (106, Bombycidae).

Oxytricha trifallax

Oxytricha trifallax is a single-celled ciliate, with the remarkable ability to undergo programmed genome rearrangement during development. Like other ciliates, *O. trifallax* has two forms of nucleus: macronucleus (MAC) and micronucleus (MIC). MAC is generated after conjugation from MIC, through the fragmentation, rearrangement, and addition of telomeric

repeats. The nanochromosomes in MAC are short, linear DNA fragments that carry individual genes. Repetitive sequences are accumulated in MIC, but not in MAC. Now Repbase contains 194 *O. trifallax* repeat entries.

TE classification

The number of Repbase entries in each TE category is shown in Table 6, accompanied by the increase in the number of entries in the last three years. In the year 2023, we added *IS481EU* and *Replitron* under DNA transposon in our classification scheme. *IS481EU* is a small group of eukaryotic DNA transposons found in parabasallids. This group shows a close relationship with the bacterial *IS481* family of insertion sequences (Kojima and Bao 2023). *Replitron* encodes a protein containing a HUH nuclease, but unlike *Helitron*, it does not encode a helicase or a Y2 transposase (Craig 2023). Instead, *Replitron* encodes a Y1 transposase. Some *Crack* families were reclassified into the *Daphne* clade. Two new lineages of non-LTR retrotransposons (195 *Itch* families and 36 *Bite* families) from ticks were reported but not yet assigned to clades.

As seen in Figure 2, more than half of Repbase entries belong to LTR retrotransposon (if including endogenous retroviruses (ERVs) and *DIRS* (YR) retrotransposons), and a quarter to the *Gypsy* superfamily. In the current classification scheme of non-LTR retrotransposons, *LI* is the dominant, owing to the large number of *LI* families in mammals and land plants. The most dominant 6 superfamilies (*hAT*, *MuDR*, *Mariner/Tc1*, *Harbinger*, *Helitron*, and *EnSpm/CACTA*) sum up to three-quarters (18,269) of total entries of DNA transposons.

References

Bao W, Kojima KK, Kohany O.

Repbase Update, a database of repetitive elements in eukaryotic genomes.

Mob DNA, 2015; 6:11.

Kojima KK.

Structural and sequence diversity of eukaryotic transposable elements

Genes and Genetic Systems, 2020; 94(6):233-252.

Kojima KK, Bao W.

IS481EU shows a new connection between eukaryotic and prokaryotic DNA transposons.

Biology (Basel), 2023 Feb 25;12(3)

Craig RJ.

Replitrans: A major group of eukaryotic transposons encoding HUH endonuclease.

Proc Natl Acad Sci U S A. 2023 Jun 20;120(25):e2301424120.

Table 1. Organisms whose repeats were published in Rebase Reports in 2023.

Chordata	
- Vertebrata	
-- Aves	<i>Calidris pugnax</i> (ruff), <i>Calypte anna</i> (Anna's hummingbird), <i>Hirundo rustica</i> (barn swallow),
-- Amphibia	<i>Bufo bufo</i> (common toad), <i>Rana temporaria</i> (common frog), <i>Xenopus tropicalis</i> (western clawed frog)
-- Actinopterygii	<i>Ameiurus melas</i> (black bullhead), <i>Betta splendens</i> (Siamese fighting fish), <i>Danio rerio</i> (zebrafish), <i>Fundulus heteroclitus</i> (mummichog), <i>Gymnodraco acuticeps</i> (ploughfish), <i>Harpagifer antarcticus</i> (Antarctic spiny plunderfish), <i>Ilyophis</i> sp. 1 JC-2022 (arrowtooth eel), <i>Larimichthys crocea</i> (large yellow croaker), <i>Neogobius melanostomus</i> (round goby), <i>Pangasianodon hypophthalmus</i> (striped catfish), <i>Parambassis ranga</i> (Indian glass fish), <i>Pygocentrus nattereri</i> (red-bellied piranha)
- Tunicata	<i>Oikopleura dioica</i>
Arthropoda	
- Insecta/Diptera	<i>Drosophila albomicans</i> (fruit fly), <i>Drosophila grimshawi</i> (fruit fly), <i>Drosophila mojavensis</i> (fruit fly), <i>Drosophila takahashii</i> (fruit fly)
- Insecta/Lepidoptera	<i>Cydia splendana</i> (chestnut tortrix), <i>Habrosyne pyritoides</i> (buff arches)
- Insecta/Coleoptera	<i>Photinus pyralis</i> (common eastern firefly), <i>Rhyzopertha dominica</i> (lesser grain borer)
- Insecta/Hemiptera	<i>Empoasca (Matsumurasca) onukii</i> (tea green leafhopper)
- Malacostraca/Decapoda	<i>Litopenaeus vannamei</i> (whiteleg shrimp), <i>Macrobrachium nipponense</i> (oriental river prawn), <i>Penaeus monodon</i> (black tiger shrimp)
- Thecostraca/Balanomorpha	<i>Amphibalanus amphitrite</i> (barnacle)
- Arachnida/Ixodida	<i>Dermacentor silvarum</i> , <i>Haemaphysalis longicornis</i> (longhorned tick), <i>Hyalomma asiaticum</i> , <i>Ixodes scapularis</i> (black-legged tick), <i>Rhipicephalus sanguineus</i> (brown dog tick)
Nematoda	<i>Caenorhabditis remanei</i> , <i>Haemonchus contortus</i> (barber pole worm)
Mollusca	<i>Corbicula fluminea</i> (Asian clam), <i>Crassostrea gigas</i> (Pacific oyster), <i>Mercenaria mercenaria</i> (northern quahog), <i>Pomacea canaliculata</i> (golden apple snail), <i>Saccostrea glomerata</i> (Sydney rock oyster)
Annelida	<i>Owenia fusiformis</i>
Nemertea	<i>Lineus longissimus</i>
Platyhelminthes	<i>Schmidtea mediterranea</i>
Cnidaria	<i>Anthopleura sola</i> (starburst anemone), <i>Hydra vulgaris</i> (<i>H. magnipapillata</i>), <i>Metridium senile</i> (brown sea anemone)
Fungi	
Ascomycota	<i>Fusarium oxysporum</i> , <i>Geopyxis carbonaria</i> , <i>Mycosphaerella graminicola</i> , <i>Paracoccidioides brasiliensis</i> , <i>Phymatotrichopsis omnivora</i> (Texas root rot), <i>Pyrenophora teres</i> , <i>Pyrenophora tritici-repensis</i> , <i>Pyronema domesticum</i> , <i>Sphaerosporella brunnea</i> , <i>Talaromyces marneffei</i> , <i>Tirmania nivea</i> , <i>Tricharina praecox</i> , <i>Trichophaea hybrida</i>
Basidiomycota	<i>Heterobasidion irregulare</i> , <i>Leucocalocybe mongolica</i> , <i>Puccinia graminis</i> (wheat stem rust)
Angiospermae	
- Monocots	
-- Poales	<i>Ananas comosus</i> (pineapple), <i>Brachypodium distachyon</i> (purple false brome), <i>Coix lacryma-jobi</i> (Job's tears), <i>Echinochloa crus-galli</i> (Japanese millet), <i>Eragrostis curvula</i> (weeping love grass), <i>Oryza sativa</i> (Asian rice), <i>Panicum hallii</i> (Hall's panicgrass), <i>Setaria italica</i> (foxtail millet), <i>Sorghum bicolor</i> (sorghum), <i>Triticum aestivum</i> (bread wheat), <i>Zea mays</i> (maize)
-- Zingiberales	<i>Musa acuminata</i> (wild banana)
- Asterids	
-- Asterales	<i>Lactuca sativa</i> (lettuce), <i>Stevia rebaudiana</i> (sweetleaf)
-- Aquifoliales	<i>Ilex paraguariensis</i> (yerba mate)
-- Lamiales	<i>Avicennia marina</i> (grey mangrove), <i>Olea europaea</i> (common olive),
-- Solanales	<i>Capsicum annuum</i> (bell pepper), <i>Capsicum chinense</i> (habanero-type pepper)
- Rosids	
-- Fabales	<i>Arachis hypogaea</i> (peanut), <i>Cicer arietinum</i> (chickpea), <i>Glycine max</i> (soybean), <i>Medicago sativa</i> (alfalfa)
-- Malpighiales	<i>Manihot esculenta</i> (mandioca/cassava/yucca/manioc), <i>Salix purpurea</i> (purple willow)
-- Rosales	<i>Fragaria x ananassa</i> (garden strawberry), <i>Prunus persica</i> (peach)
-- Sapindales	<i>Citrus clementina</i> (clementine), <i>Citrus sinensis</i> (orange)
-- Myrtales	<i>Corymbia citriodora</i> (lemon-scented gum), <i>Eucalyptus grandis</i> (flooded gum/rose gum)
-- Brassicales	<i>Brassica rapa</i> (Chinese cabbage)
- Caryophyllales	<i>Chenopodium quinoa</i> (quinoa)
- Nymphaeales	Nymphaeaceae sp. (water lilies)
- Amborellales	<i>Amborella trichopoda</i>
Other eukaryotes	<i>Carpediemonas membranifera</i> , <i>Oxytricha trifallax</i> (ciliate), <i>Paulinella micropora</i> , <i>Tisochrysis lutea</i>

Table 2. Top 40 species based on the number of Rebase entries.

Species	Total	2023 Increase
<i>Aedes aegypti</i> (yellow fever mosquito)	3648	0
<i>Oryza sativa</i> (Asian rice)	3439	199
<i>Danio rerio</i> (zebrafish)	2473	150
<i>Zea mays</i> (maize)	2108	71
<i>Crassostrea gigas</i> (Pacific oyster)	1818	172
<i>Arabidopsis lyrata</i> (lyrate rockcress)	1567	0
<i>Chondrus crispus</i> (Irish moss)	1185	2
<i>Triticum aestivum</i> (bread wheat)	1144	1060
<i>Arabidopsis thaliana</i> (thale cress)	1136	180
<i>Locusta migratoria</i> (migratory locust)	1129	2
<i>Hydra vulgaris</i> (hydra)	1003	258
<i>Medicago truncatula</i> (barrel medic)	1001	0
<i>Sorghum bicolor</i> (sorghum)	979	8
<i>Chrysemys picta bellii</i> (painted turtle)	843	0
<i>Drosophila takahashii</i> (fruit fly)	821	85
<i>Xenopus tropicalis</i> (western clawed frog)	766	69
<i>Corbicula fluminea</i> (Asian clam)	764	196
<i>Nematostella vectensis</i> (starlet sea anemone)	757	0
<i>Culex quinquefasciatus</i> (southern house mosquito)	660	0
<i>Lepeophtheirus salmonis</i> (salmon louse)	654	0
<i>Drosophila willistoni</i> (fruit fly)	637	1
<i>Carassius auratus</i> (goldfish)	632	1
<i>Homo sapiens</i> (human)	584	0
<i>Glycine max</i> (soybean)	578	20
<i>Raphanus sativus</i> (radish)	537	0
<i>Salmo salar</i> (Atlantic salmon)	535	0
<i>Malus domestica</i> (apple)	526	0
<i>Rhyzopertha dominica</i> (lesser grain borer)	513	513
<i>Alligator mississippiensis</i> (American alligator)	508	0
<i>Melampsora larici-populina</i> (poplar leaf rust fungus)	500	0
<i>Solanum tuberosum</i> (potato)	486	1
<i>Amblyraja radiata</i> (thorny skate)	467	0
<i>Mus musculus</i> (domestic mouse)	464	0
<i>Phytophthora infestans</i> ()	465	0
<i>Penaeus monodon</i> (Asian tiger shrimp)	452	132
<i>Litopenaeus (Penaeus) vannamei</i> (whiteleg shrimp)	441	3
<i>Monodelphis domestica</i> (gray short-tailed opossum)	434	0
<i>Brassica rapa</i> (Chinese cabbage)	433	52
<i>Amborella trichopoda</i>	412	22
<i>Ilyophis sp. 1 JC-2022</i> (eel)	406	301

Table 3. Top 30 genera based on the number of Repbase entries.

Genus (Species with >10 entries)	Total	2023 Increase
<i>Drosophila</i> ¹	4699	449
<i>Aedes</i> (<i>A. aegypti</i>)	3649	0
<i>Oryza</i> (<i>O. sativa</i>)	3456	199
<i>Arabidopsis</i> (<i>A. thaliana</i> , <i>A. lyrata</i>)	2703	0
<i>Danio</i> (<i>D. rerio</i>)	2487	150
<i>Zea</i> (<i>Z. mays</i>)	2110	69
<i>Crassostrea</i> (<i>C. gigas</i> , <i>C. virginica</i>)	2004	172
<i>Puccinia</i> (<i>P. striiformis</i> , <i>P. graminis</i> , <i>P. hordei</i> , <i>P. triticina</i> , <i>P. coronata</i> , <i>P. horiana</i>)	1360	64
<i>Triticum</i> (<i>T. aestivum</i>)	1226	1060
<i>Chondrus</i> (<i>C. crispus</i>)	1185	2
<i>Locusta</i> (<i>L. migratoria</i>)	1129	2
<i>Xenopus</i> (<i>X. tropicalis</i> , <i>X. laevis</i>)	1102	69
<i>Medicago</i> (<i>M. truncatula</i>)	1010	3
<i>Hydra</i> (<i>H. vulgaris</i>)	1003	257
<i>Sorghum</i> (<i>S. bicolor</i>)	979	8
<i>Panaeus</i> (<i>P. monodon</i> , <i>Litopanaeus</i> (<i>P.</i>) <i>vannamei</i>)	893	135
<i>Anopheles</i> (<i>A. gambiae</i> , <i>A. funestus</i> , <i>A. merus</i> , <i>A. stephensi</i>)	886	0
<i>Chrysemys</i> (<i>C. picta</i>)	843	0
<i>Phytophthora</i> (<i>P. infestans</i> , <i>P. sojae</i> , <i>P. ramorum</i> , <i>P.</i> <i>parasitica</i>)	791	5
<i>Corbicula</i> (<i>C. fluminea</i>)	764	196
<i>Nematostella</i> (<i>N. vectensis</i>)	758	0
<i>Culex</i> (<i>C. quinquefasciatus</i>)	669	0
<i>Lepeophtheirus</i> (<i>L. salmonis</i>)	654	0
<i>Mus</i> (<i>M. musculus</i> , <i>M. pahari</i> , <i>M. caroli</i> , <i>M. spretus</i>)	637	0
<i>Carassius</i> (<i>C. auratus</i>)	632	1
<i>Homo</i> (<i>H. sapiens</i>)	584	0
<i>Solanum</i> (<i>S. tuberosum</i> , <i>S. lycopersicum</i> , <i>S. demissum</i>)	582	5
<i>Glycine</i> (<i>G. max</i>)	578	20
<i>Brassica</i> (<i>B. rapa</i> , <i>B. oleracea</i>)	571	52
<i>Caenorhabditis</i> (<i>C. briggsae</i> , <i>C. elegans</i> , <i>C. inopinata</i> , <i>C.</i> <i>remanei</i> , <i>C. japonica</i> , <i>C. brenneri</i>)	565	0

¹ Species are shown in Table 4.

Table 4. Repbase entries from the genus *Drosophila*.

Subgenus - Group (subgroup)	Species	Abbr.	Repbase entries	2023 Increase
<i>Sophophora - melanogaster</i> (<i>melanogaster</i> subgroup)	<i>D. melanogaster</i>	DM	296	0
	<i>D. simulans</i>	DSim	43	0
	<i>D. sechellia</i>	DSe	33	0
	<i>D. yakuba</i>	DY	139	0
	<i>D. erecta</i>	DEre	18	0
(<i>suzukii</i> subgroup)	<i>D. suzukii</i>	DSuz	207	0
	<i>D. biarmipes</i>	DBi	80	0
(<i>ananassae</i> subgroup)	<i>D. ananassae</i>	DAn	240	0
	<i>D. bipectinata</i>	DBp	158	0
(<i>elegans</i> subgroup)	<i>D. elegans</i>	DEl	314	0
(<i>takahashii</i> subgroup)	<i>D. takahashii</i>	DTa	821	85
(<i>rhopaloa</i> subgroup)	<i>D. rhopaloa</i>	DRh	116	0
(<i>montium</i> subgroup)	<i>D. kikkawai</i>	DKi	97	0
	<i>D. serrata</i>	DSer	71	0
<i>Sophophora - ficusphila</i>	<i>D. ficusphila</i>	DF	129	0
<i>Sophophora - eugracilis</i>	<i>D. eugracilis</i>	DEu	200	0
<i>Sophophora - obscura</i>	<i>D. azteca</i>	DAzt	102	0
	<i>D. pseudoobscura</i>	DPse	80	0
	<i>D. persimilis</i>	DPer	106	0
<i>Sophophora - willistoni</i>	<i>D. willistoni</i>	DWil	637	0
<i>Drosophila - immigrans</i>	<i>D. albomicans</i>	DAlb	254	254
<i>Siphlodora - repleta</i>	<i>D. mojavensis</i>	DMoj	235	81
	<i>D. hydei</i>	DHyd	39	0
<i>Siphlodora - virilis</i>	<i>D. virilis</i>	DVi	148	0
<i>Idiomyia</i>	<i>D. grimshawi</i>	DGri	65	24

¹ Only species with >10 entries are shown.

² Subgroups are shown only in the *melanogaster* group.

Table 5. Top 20 species whose entries increased in 2023.

Species	2023 Increase	Total
<i>Triticum aestivum</i> (bread wheat)	1060	1144
<i>Rhyzopertha dominica</i> (lesser grain borer)	513	513
<i>Cydia splendana</i> (chestnut tortrix)	349	349
<i>Capsicum annuum</i> (chili pepper)	324	359
<i>Ilyophis</i> sp. (eel)	301	406
<i>Haemaphysalis longicornis</i> (Asian longhorned tick)	266	266
<i>Hydra vulgaris</i> (hydra)	257	1003
<i>Drosophila albomicans</i> (fruit fly)	254	254
<i>Rana temporaria</i> (common frog)	241	242
<i>Amphibalanus amphitrite</i> (barnacle)	241	241
<i>Fundulus heteroclitus</i> (mummichog)	233	235
<i>Mercenaria mercenaria</i> (quahog)	202	210
<i>Corbicula fluminea</i> (Asian clam)	196	764
<i>Oryza sativa</i> (Asian rice)	191	3439
<i>Murina aurata</i> (little tube-nosed bat)	185	185
<i>Neogobius melanostomus</i> (round goby)	176	177
<i>Saccostrea glomerata</i> (Sydney rock oyster)	173	184
<i>Crassostrea gigas</i> (Pacific oyster)	172	1818
<i>Pygocentrus nattereri</i> (red-bellied piranha)	172	173
<i>Oxytricha trifallax</i> (ciliate)	168	194

Table 6. TE classification and the numbers of entries in Repbase.

A. LTR retrotransposon

Superfamily	Total	2023 increase	2022 increase	2021 increase
<i>Gypsy</i>	26081	3972	4508	4139
<i>Copia</i>	9983	514	605	908
<i>BEL</i>	4603	581	974	432
<i>DIRS</i>	1137	201	104	187
Endogenous Retrovirus (ERV)				
<i>ERV1</i>	4436	551	1578	45
<i>ERV2</i>	2745	305	877	0
<i>ERV3</i>	1338	146	375	0
<i>ERV4</i>	202	0	9	0
<i>Lentivirus</i>	4	0	0	0
<i>Lokiretrovirus</i>	55	8	4	-
<i>Spumaretrovirus</i>	7	1	1	-
Unclassified ERV	322	0	4	-10
Unclassified LTR retrotransposon	752	-22	38	9
Total	51605	6188	9077	5767

B. Penelope-like retrotransposon

Superfamily	Total	2023 increase	2022 increase	2021 increase
<i>Penelope/Poseidon</i>	344	3	52	-
<i>Neptune</i>	236	11	5	-
<i>Nematis</i>	5	0	0	-
<i>Athena</i>	42	0	0	-
<i>Coprina</i>	15	0	3	-
<i>Hydra</i>	28	0	6	-
<i>Naiad/Chlamys</i>	119	0	110	-
Unclassified <i>Penelope</i>	23	1	5	-
Total	812	15	181	77

C. Non-LTR retrotransposon

Superfamily (clade)	Total	2023 increase	2022 increase	2021 increase
SINE				
<i>SINE1/7SL</i>	223	31	92	0
<i>SINE2/tRNA</i>	2335	547	977	133
<i>SINE3/5S</i>	33	1	6	-6
<i>SINEU/snRNA</i>	17	0	0	0
Unclassified SINE	74	-1	2	-39
<i>CRE</i>	87	32	2	5
<i>R4</i>	58	2	3	0
<i>Hero</i>	32	3	1	0
<i>NeSL</i>	119	0	0	0
<i>R2</i>	184	3	21	1
<i>RandI/Dualen</i>	13	0	0	0
<i>Proto1</i>	10	0	0	4
<i>L1</i>	4184	783	1237	195
<i>Tx1</i>	407	36	66	1
<i>RTETP</i>	1	0	0	0
<i>Proto2</i>	68	15	1	0
<i>RTEX</i>	298	4	141	5
<i>RTE</i>	761	86	108	14
<i>Outcast</i>	41	3	7	8
<i>Ingi</i>	36	2	8	4
<i>Vingi</i>	154	1	0	0
<i>I</i>	251	15	28	4
<i>Nimb</i>	183	57	9	0
<i>Tad1</i>	548	29	32	177
<i>Loa</i>	98	8	2	1
<i>R1</i>	312	18	15	13
<i>Jockey</i>	369	26	44	19
<i>Rex1</i>	141	3	5	7
<i>CR1</i>	1049	73	59	44
<i>Kiri</i>	130	22	3	13
<i>L2</i>	318	7	12	2
<i>L2A</i>	7	0	2	0
<i>L2B</i>	29	0	0	0
<i>Crack</i>	134	-14	5	1
<i>Daphne</i>	337	92	6	1
<i>Ambal</i>	8	0	0	0
Unclassified non-LTR retrotransposon	446	221	23	23
Total	13495	2105	2917	630

D. DNA transposon

Superfamily	Total	2023 increase	2022 increase	2021 increase
<i>EnSpm/CACTA</i>	1278	133	37	195
<i>Transib</i>	258	33	43	39
<i>hAT</i>	5847	572	557	643
<i>MuDR</i>	2301	188	68	313
<i>P</i>	302	29	27	32
<i>Kolobok</i>	852	50	411	40
<i>Dada</i>	171	0	120	0
<i>Mariner/Tc1</i>	4755	731	865	136
<i>Zator</i>	101	7	5	6
<i>piggyBac</i>	606	57	118	10
<i>Merlin</i>	148	35	20	2
<i>Harbinger</i>	2156	224	146	221
<i>ISL2EU</i>	203	11	28	22
<i>Ginger1</i>	22	2	-19	0
<i>Ginger2/TDD</i>	66	2	20	5
<i>IS3EU</i>	78	33	0	0
<i>IS481EU</i>	30	30	-	-
<i>Sola</i>				
<i>Sola1</i>	162	23	30	2
<i>Sola2</i>	159	35	21	5
<i>Sola3</i>	45	1	0	2
<i>Academ</i>	410	23	24	34
<i>Novosib</i>	9	0	0	0
<i>Zisupton</i>	44	1	1	1
<i>Helitron</i>	1932	242	250	218
<i>Replitron</i>	36	36	-	-
<i>Polinton</i>	234	39	41	12
<i>Crypton</i>				
<i>CryptonF</i>	26	2	0	0
<i>CryptonA</i>	31	13	0	0
<i>CryptonI</i>	30	20	0	0
<i>CryptonS</i>	59	0	0	0
<i>CryptonV</i>	59	6	0	0
Unclassified <i>Crypton</i>	91	2	0	0
Unclassified DNA transposon	2356	23	22	35
Total	24857	2603	2816	1973

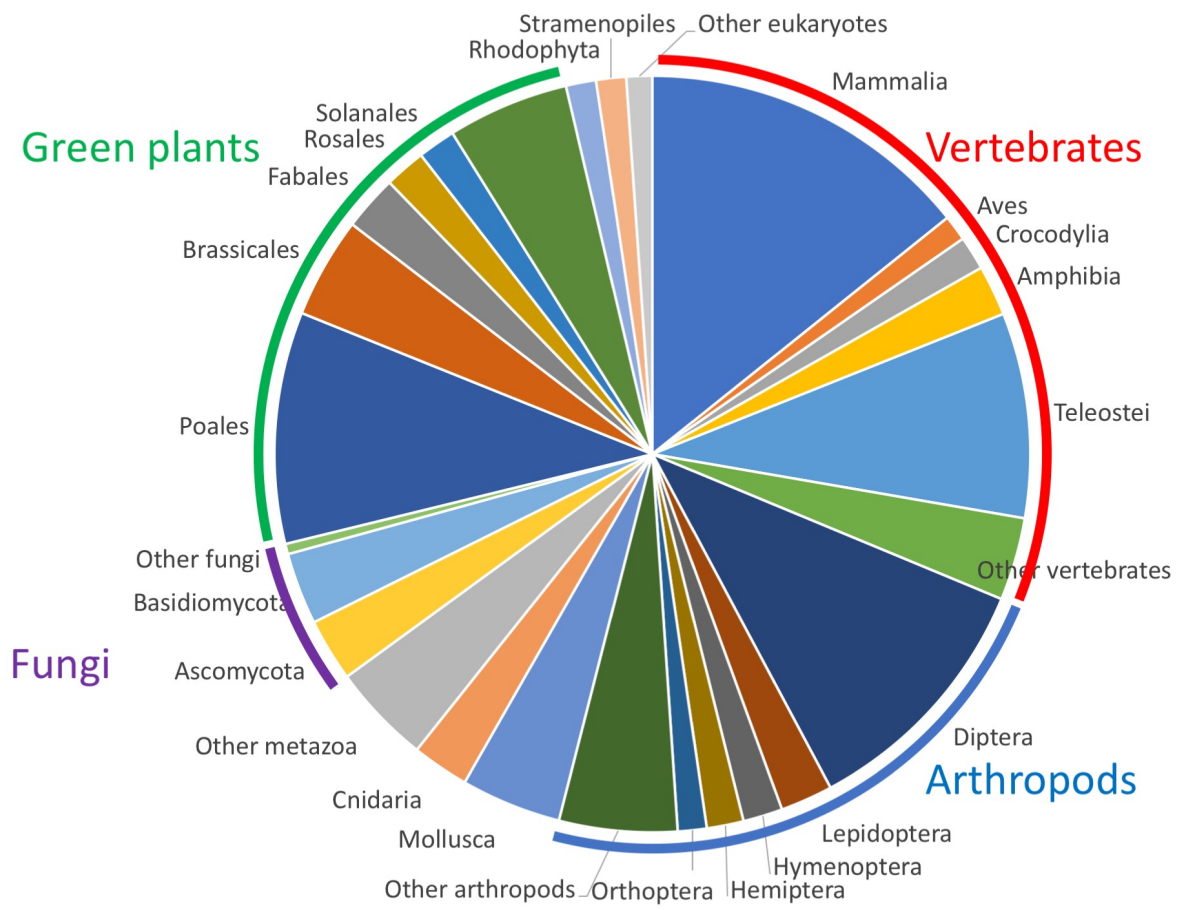


Figure 1. Proportions of Rebase entries in phylogenetic lineages. Lineages with over 1000 entries are shown.

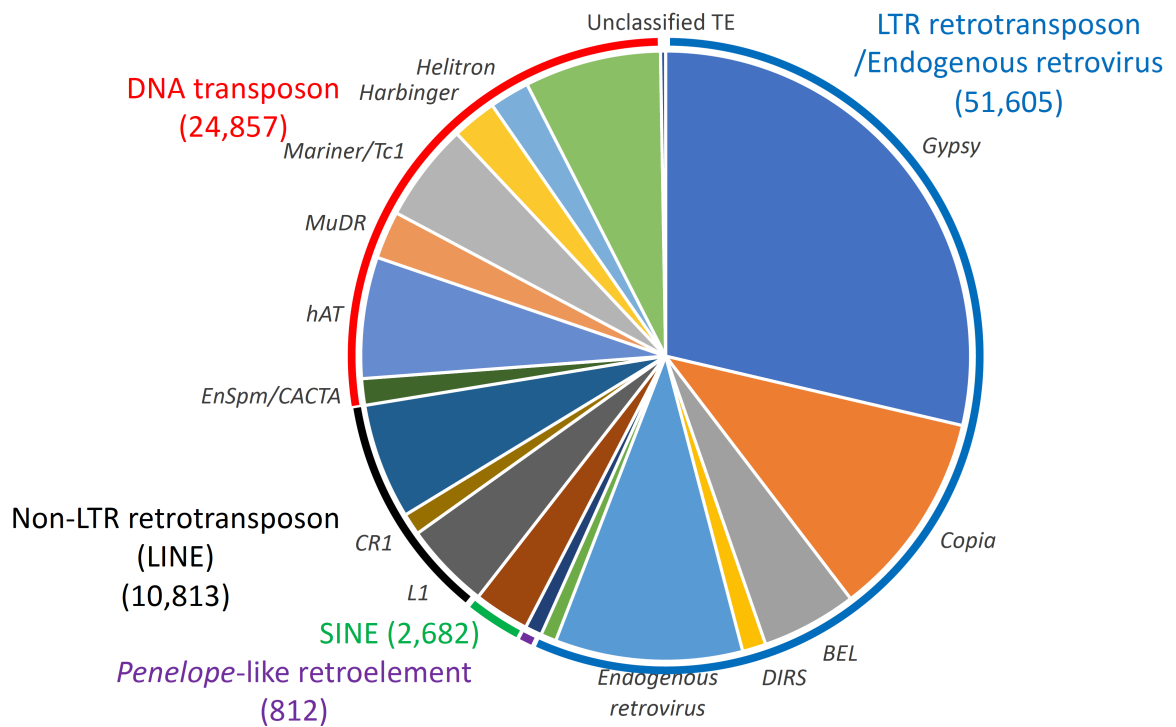


Figure 2. Proportions of Rebase entries in transposon groupings. Groups with over 1000 entries are shown. Numbers in parentheses show the numbers of Rebase entries.