

Review

Phylogenetic taxonomy in *Drosophila*

Problems and prospects

Patrick M. O'Grady^{1,*} and Therese A. Markow²

¹University of California, Berkeley; Department of Environmental Science; Policy and Management; Berkeley, California USA; ²University of California, San Diego; Division of Biological Sciences; San Diego, California USA

Key words: *Drosophila*, taxonomy, phylogenetics, evolution, classification

The genus *Drosophila* is one of the best-studied model systems in modern biology, with twelve fully sequenced genomes available. In spite of the large number of genetic and genomic resources, little is known concerning the phylogenetic relationships, ecology and evolutionary history of all but a few species. Recent molecular systematic studies have shown that this genus is comprised of at least three independent lineages and that several other genera are actually imbedded within *Drosophila*. This genus accounts for over 2,000 described, and many more undescribed, species. While some *Drosophila* researchers are advocating dividing this genus into three or more separate genera, others favor maintaining *Drosophila* as a single large genus. With the recent sequencing of the genomes of multiple *Drosophila* species and their expanding use in comparative biology, it is critical that the *Drosophila* research community understands the taxonomic framework underlying the naming and relationships of these species. The subdivision of this genus has significant biological implications, ranging from the accurate annotation of single genes to understanding how ecological adaptations have occurred over the history of the group.

Introduction

Over the course of the past 25 years, the tools of phylogenetic systematics have been applied to test existing taxonomies and to propose novel relationships among taxa. Phylogenetic systematics seeks to use explicit character information (e.g., morphology, DNA sequence) to infer evolutionary relationships among organisms. This approach is an improvement over traditional taxonomy because of the explicit, repeatable analytical methods used to infer evolutionary relationships. Powerful statistical methods can be applied to place taxonomy in an hypothesis-testing framework, an important consideration when making statements about morphological, developmental and ecological innovation, character evolution and phylogenetic relationships.

A given group of organisms can be classified as monophyletic, paraphyletic or polyphyletic (Fig. 1A–C). Monophyletic groups, or clades, consist of a common ancestor and all descendants of that ancestor (Fig. 1A). Basing taxonomic structure on clades is a powerful approach because it provides information about the composition and exclusivity of a group. Shared derived characters that delimit a group can also be used to tentatively place newly discovered species. Paraphyletic groups contain an ancestor and only some descendants of that ancestor (Fig. 1B). Because some descendants of an ancestor are not present in a paraphyletic group, these are less useful when trying to make an explicit statement about the evolutionary history of a lineage. Polyphyly is similar to paraphyly in that not all ancestors or descendants are members of the same group. Polyphyletic groups originate from multiple points on a phylogenetic tree (Fig. 1C).

Phylogenetics and Taxonomy of *Drosophila*

The debate between those who prefer to divide large genera into smaller units and others who propose to lump multiple genera into a single taxon is probably the oldest in taxonomy. The genus *Drosophila* provides an interesting modern example. This genus is polyphyletic and thus one will need to either divide it into smaller, monophyletic units or propose a single large clade that encompasses all currently described *Drosophila* species (Fig. 2A–C). The issue is not whether the genus represents a heterogeneous assemblage of lineages—it does. Rather, the question is what is the optimal way to utilize our current knowledge of phylogeny to improve the taxonomic structure within this genus?

Figure 2A shows a scenario where the genus *Drosophila* has been modified to include a number of other genera in order to create a large monophyletic group containing over 2,000 species. Under such a scheme, several genera (e.g., *Scaptomyza*, *Hirtodrosophila*, *Zaprionus*) are included as subgenera of *Drosophila* and their generic names must change. This action would create **homonyms** (see Box 1), species names that are identical yet refer to different species. For example, *Drosophila heedi* and *Scaptomyza heedi* are two separate taxonomic entities. However, if these two genera are at one point considered **synonyms** (see Box 1), one of those species names must change to avoid confusion about which morphological species the name *Drosophila heedi* is referring to. Creating a larger genus *Drosophila* would generate over 100 homonyms.¹

*Correspondence to: Patrick M. O'Grady; University of California, Berkeley; Department of Environmental Science; Policy and Management; 137 Mulford Hall; Berkeley, California 94720 USA; Email: ogrady@nature.berkeley.edu

Submitted: 01/05/09; Accepted: 01/06/09

Previously published online as a Fly E-publication:
<http://www.landesbioscience.com/journals/fly/article/7748>

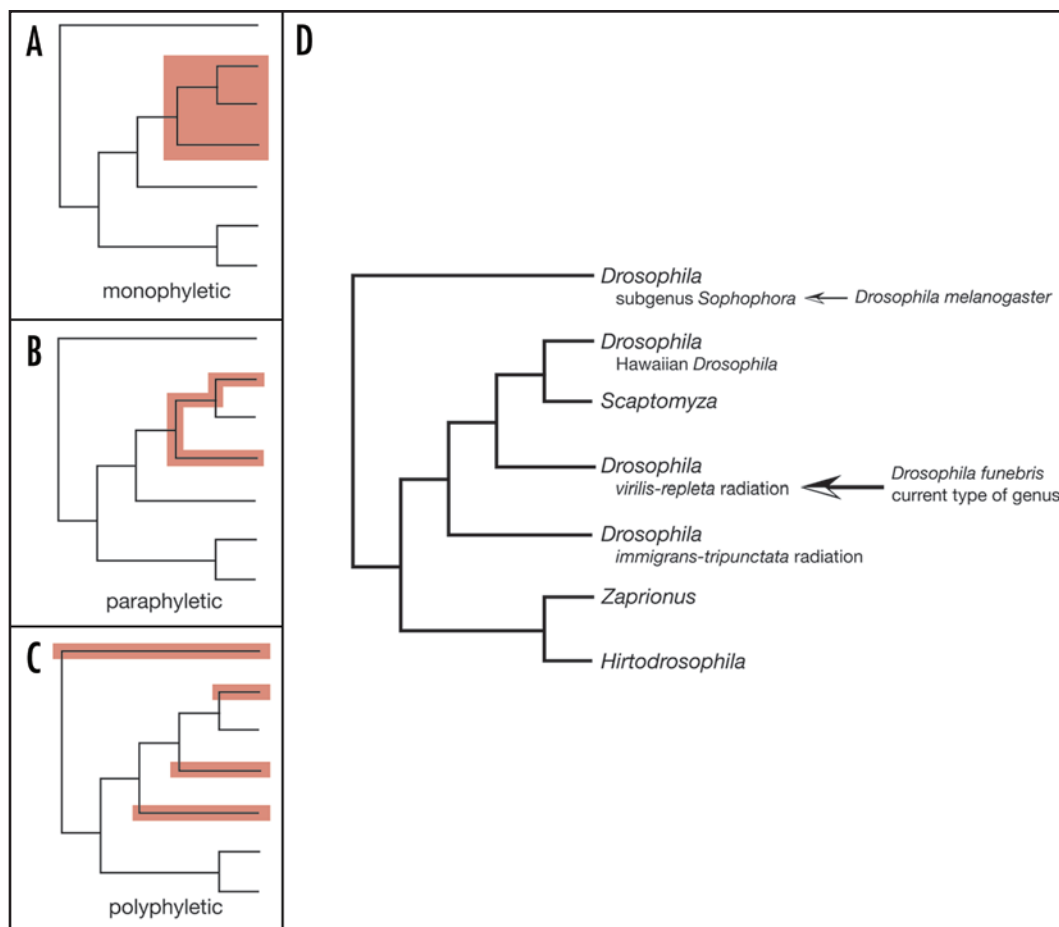


Figure 1. Phylogenetic trees showing (A) monophyletic, (B) paraphyletic and (C) polyphyletic groups outlined in red. (D) Simplified version of phylogenetic relationships supported by O'Grady and DeSalle²⁴ to illustrate the polyphyly of the genus *Drosophila*. The type of the genus, *D. funebris* is indicated, as is the genetic model system *D. melanogaster*.

Figure 2B illustrates two scenarios where the genus *Drosophila* is split into smaller monophyletic groups. While these would require some change in generic names, there would be no issue with homonyms. However, the first scenario would require a change in some generic names. This is due to the rules of nomenclature and how types are designated. When new species are described, a single specimen is designated as the **holotype** (see Box 1), the specimen to which all others are compared when determining species identity. The rule of priority states that the earliest described species in a genus is designated the type of that genus. This provides a link between the name and the morphological definition of a species and/or genus. All newly described species are placed into genera based on their similarity to the type species of that genus. *Drosophila funebris* Fallen 1823 is the type of the genus *Drosophila*.² The rule of priority also holds for genera. The earliest valid genus name should be used whenever possible. When genera are synonymized, or equated with one another, the oldest type species determines which genus name is used. All subsequent generic names are considered junior synonyms and are not used, although they may be resurrected in the future if the taxonomy changes. *Drosophila funebris*, the type of the genus *Drosophila*, is not in the same clade as *Drosophila melanogaster*. Thus, if the genus *Drosophila* were to be split into smaller monophyletic groups and *Drosophila funebris*

maintained as the type of the genus, according to the rules of nomenclature the name *Drosophila melanogaster* would change to *Sophophora melanogaster* (Fig. 2B).

An alternate approach to splitting the genus *Drosophila* is also shown in Figure 2B. This involves proposing to redesignate *Drosophila melanogaster* as the type of the genus *Drosophila*. Such an action, however, requires making an exception to the rules of nomenclature and an application to do this has been made to the International Committee of Zoological Nomenclature.³ The authors base their application on two main arguments; (1) that changing “*Drosophila*” (i.e., *Drosophila melanogaster*) to *Sophophora* would cause widespread confusion in the literature and should be avoided and (2) phylogenetic analyses support their splitting approach. While this seems like an acceptable course of action, there are some drawbacks presented in a series of comments written in opposition.⁴⁻¹⁰ Most importantly, the rules of nomenclature clearly state that the Commission may only render decisions on issues of nomenclature and should not endorse any single classification scheme—taxonomic, phylogenetic or otherwise.¹¹ By placing this issue before the Commission, the petitioners are asking for an endorsement of their taxonomic scheme (Fig. 2B, splitting the genus) over alternate scenarios (Fig. 2A and C) that are equally well supported by the data and involve less drastic taxonomic change.

Box 1 Glossary of terms used in this paper

Nomenclature	The science of naming new species. All animal species are governed by rules set forth in the International Code of Zoological Nomenclature. A Commission of experts accepts and rules on petitions to modify the code and alter species names. http://www.iczn.org/iczn/index.jsp
Binomen	The full genus and species name of an organism.
Homonym	Two separate biological entities with the same genus and species combination. The rules of nomenclature state that the oldest name is valid and subsequent names be changed.
Synonym	A single biological entity with two different genus and species combinations. The rules of nomenclature state that the oldest species name is valid and subsequent names are considered junior synonyms. Higher level taxonomic groups (e.g., genera) can also be considered synonymous with one another.
Holotype (or type)	The individual specimen that characterizes a species and to which all other specimens are compared when assessing species identity. The type of a genus is the first described species in that genus.
Taxonomy	Organizing living organisms into hierarchical groups based on some measure of morphological similarity (kingdom, phylum, class, order, family, genus, species).
Phylogenetic Systematics (or phylogenetics)	The field concerned with inferring the evolutionary history of all life using explicit computational analysis of morphological and/or molecular characters and rigorous statistical analyses. Groups present in phylogenetic trees can either be monophyletic, paraphyletic or polyphyletic (Fig. 1).

If the Commission votes in favor of the van der Linde et al.³ proposal, they would make *D. melanogaster* the type of the genus *Drosophila* and preserve the **binomen** (see Box 1), *Drosophila melanogaster*. Should the Commission ultimately oppose the petition, there is, in fact, still no change to the name *D. melanogaster*. Only someone revising the genus *Drosophila* into smaller units can change the name. If such a revision were published, the community could either accept or reject that revision through usage.

Figure 2C involves a fourth option, to base the relationships in the genus *Drosophila* on evolutionary history rather than taxonomy. Such an approach would result in no change of generic names, create no homonyms, and require no redesignation of the type of the genus. While this requires that researchers be aware of the phylogenetic relationships among drosophilid taxa, this option offers the greatest flexibility as it can be modified easily as more data are obtained. Furthermore, because the generic names of the Hawaiian *Drosophila*, cactophilic *Drosophila* and *Drosophila virilis* would remain unchanged, this option also preserves the greatest degree of continuity in the literature.

Figure 2C offers a compelling argument for both historical and methodological reasons. The genus *Drosophila* has always been what is referred to as a “catch all” or “trash can” genus. When new species didn’t clearly belong to any other genus, they were placed in *Drosophila*. In an attempt to ameliorate this problem and organize some of the divergent lineages in *Drosophila*, Sturtevant^{12,13} proposed a series of subgenera and species groups. Subsequent workers have expanded upon this framework as new taxa were discovered. Throckmorton¹⁴ recognized a new taxonomic rank, the radiation, below the level of subgenus and used it to associate related species groups. This historical artifact means that, in essence, *Drosophila* workers have been following the scheme proposed in Figure 2C for well over 40 years. Interestingly, because *Drosophila* biologists have always had a large amount of data (chromosomal, behavioral, morphological, genetic, molecular) available to infer evolutionary relationships, the species groups and radiations proposed by Sturtevant,^{12,13} Throckmorton¹⁴ and others are highly congruent with the recent molecular phylogenetic work. Therefore, the classification scheme presented in Figure

2C is still based on monophyletic groups; the only difference is that these clades are not necessarily genera. Instead, species groups, subgenera and radiations are the main units of comparison, a custom long employed in the *Drosophila* literature.

General Considerations

The recent completion of 12 whole *Drosophila* genome sequences,¹⁵ and the promise of several more in the near future, has stimulated a number of comparative studies.^{16–20} Comparative approaches rely on clear, stable taxonomy and well-resolved phylogenetic hypotheses or evolutionary relationships. However, many *Drosophila* groups are either unplaced or have not been examined in sufficient detail to provide a strong test of monophyly. Others, such as the Hawaiian *Drosophila*, have over 100 species that are known, yet remain to be described. Proposing drastic taxonomic changes or making all but the broadest of generalizations from comparative studies given the current state of phylogenetic information within *Drosophilidae* may be ill advised at this time.

Elucidating the phylogenetic relationships within and among the major lineages that comprise the genus *Drosophila* is a major goal for the next ten years of *Drosophila* systematics research. Specifically, the monophyly and relationships among the basal lineages within the genus *Drosophila*, including the *immigrans-tripunctata* and *virilis-repleta* radiations and the subgenus *Sophophora*, should be resolved. Relationships among species groups within each of the two main radiations also require investigation. One of the most important considerations for future work is the level of statistical support for given evolutionary relationships. There is actually very little consensus or statistical support for many nodes on the phylogeny of *Drosophilidae*.^{21,22} For example, it is unclear whether the *immigrans-tripunctata* or *virilis-repleta* radiations, as currently defined, are monophyletic. This is because these groups diversified rapidly 30 million years ago as their ancestors adapted to fungi and necrotic plant tissues as host substrates. Many other nodes are resolved in phylogenetic analyses, but only with modest support. Support in phylogenetic analysis is often assessed using the bootstrap, a procedure that resamples the data multiple times (with replacement)

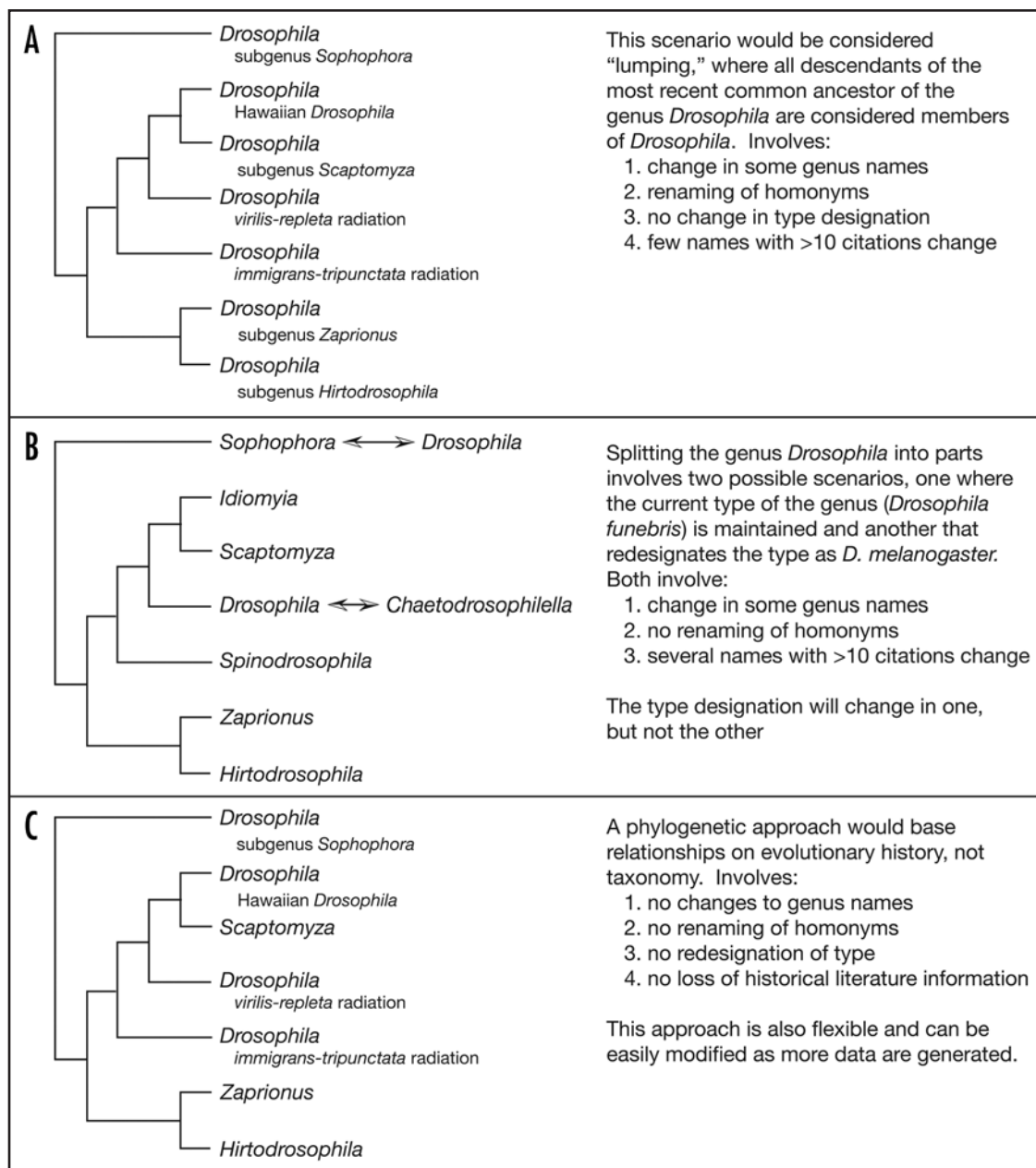


Figure 2. Four possible resolutions to a polyphyletic genus *Drosophila* and their implications to generic names, homonyms, type designations and links to historical literature. (A) A broadly defined genus *Drosophila* with many subgenera. (B) Splitting the genus *Drosophila* with and without the redesignation of the type. If *D. funebris* is maintained as the type, then the name *Drosophila melanogaster* changes to *Sophophora melanogaster*. Species in Hawaiian *Drosophila* become *Idiomyia* and members of the *immigrans-tripunctata* radiation would belong to the genus *Spinodrosophila*. Only species placed in the *virilis-repleta* group would be referred to as *Drosophila*. If *D. melanogaster* is redesignated the type of *Drosophila*, then *Drosophila melanogaster* is preserved as a name and only taxa in the subgenus *Sophophora* are considered *Drosophila*. As above, Hawaiian *Drosophila* becomes *Idiomyia* and the *immigrans-tripunctata* radiation belongs to the genus *Spinodrosophila*. Members of the *virilis-repleta* radiation would belong to the genus *Chaetodrosophilella*. (C) A possible phylogenetic scenario where clades are named and referred to in the literature and genera are not equivalent with monophyletic groups. This is essentially how taxonomy in *Drosophila* has been addressed in the past, from Sturtevant’s species groups and subgenera to Throckmorton’s radiations.

and constructs a tree based on each pseudoreplicate matrix. In the case of the relationships above, only 60–80% of resampled matrices contain support for that node. Strong bootstrap support is generally considered to be above 90%. Furthermore, the subgenus *Sophophora* has been shown to be paraphyletic with respect to the genus *Lordiphosa*,¹ a situation that would recreate the same problem currently facing the genus *Drosophila*. Clarification of the systematic relationships of these species becomes more critical as researchers

seek the most informative *Drosophila* species for additional whole genome sequencing efforts.

In his book on *Drosophila* evolution, Powell²³ states that:

“One could question whether the great diversity and numbers of species really should be a single genus; a genus with 1,600 named species (and probably at least 2,000 total) is seldom found in any other group. For example, genera could be raised to subfamilies, subgenera to genera, and so forth. *However, such radical taxonomic*

revision is not advisable at this time, as the literature and traditions are so well established that any such formal reassessment would not be worth the confusion engendered. Drosophila workers, by and large, are quite comfortable with the informal subdivisions traditionally used, and it would seem advisable to keep them.”

The continuity of the literature is an important consideration—many databases and literature resources (e.g., Flybase, NCBI) employ the current classification scheme. If the Hawaiian *Drosophila* and parts of the genus *Drosophila* were to suddenly have new names, then there would be a great deal of confusion. It would become difficult to associate the past literature dealing with, for example, ecological adaptation in *Drosophila mojavensis* with the newly coined name *Chaetodrosophilella mojavensis*. Eventually, this would be resolved, but not without massive reorganization of data management systems. Perhaps a more advisable course of action is to increase sampling of poorly known drosophilids, use newly completed genome sequences to generate additional characters for molecular systematics, and attempt to increase statistical support for certain key nodes in the phylogenetic tree of *Drosophila*. Once we have a more comprehensive and certain picture of the evolutionary relationships within this group, we can address taxonomic hierarchy within this important group of model organisms. Is now the time to adopt a new nomenclature for *Drosophila*? It is difficult to imagine that a change with such broad impact should or could successfully occur without first obtaining a consensus of the *Drosophila* research community.

References

- Hu YG, Toda MJ. Polyphyly of Lordiphosa and its relationships in Drosophilinae (Diptera: Drosophilidae). *Syst Entomol* 2001; 26:15-31.
- Fallen CF. Diptera Sueciae. Descriptio a Carolo Frederico Fallen II. Dipterorum antennis paramarticulatis instructorum. Sectionen posteriorem continen 1823.
- van der Linde K, Bächli G, Toda M, Zhang W-X, Katoh T, Hu Y-G, et al. Case 3407: *Drosophila* Fallen, 1832 (Insecta, Diptera): proposed conservation of usage. *Bull Zoolog Nomenclature* 2007; 64:238.
- Gaimari SD. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:146.
- McEvey SF, Schiffer M, DaLage J-L, David JR, Lemunier F, Joly D, et al. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:147.
- O'Grady PM, Ashburner M, Castrezana S, DeSalle R, Kaneshiro KY, Lapoint RT, et al. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:141.
- Prigent SR. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:137.
- Sidorenko V. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:98.
- Stys P. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:144.
- Yassin A. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:96.
- Thompson FC, Evenhuis NL, Pape T, Pont AC. Comments on the proposed conservation of usage of *Drosophila* (Insecta, Diptera). *Bull Zoolog Nomenclature* 2008; 65:140.
- Sturtevant AH. On the subdivision of the genus *Drosophila*. *Proc Natl Acad Sci USA* 1939; 25:137-41.
- Sturtevant AH. The classification of the genus *Drosophila*, with descriptions of nine new species. *Univ Texas Pubs* 1942; 4213:5-51.
- Throckmorton LH. The phylogeny, ecology and geography of *Drosophila*. In: King RC, ed. *A Manual of Invertebrates of Genetic Interest*. New York, NY, Plenum Press 1975; 421-69.
- Drosophila 12 Genomes Consortium. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 2007; 350:203-18.
- Bosco G, Campbell P, Leiva-Neto JT, Markow TA. Analysis of *Drosophila* species genome size and satellite DNA content reveals significant differences among strains as well as between species. *Genetics* 2007; 177:1277-90.
- Calvi BR, Byrnes BA, Kolpakas AJ. Conservation of epigenetic regulation, ORC binding and developmental timing of DNA replication origins in the genus *Drosophila*. *Genetics* 2007; 177:1291-301.
- Haerty W, Jagadeeshan S, Kulathinal RJ, Wong A, Ram KR, Sirot LK, et al. Evolution in the fast lane: Rapidly evolving sex-related genes in *Drosophila*. *Genetics* 2007; 177:1321-35.
- Koerich LB, Wang Z, Clark AG, Carvalho AB. Low conservation of gene content in the *Drosophila* Y chromosome. *Nature* 2008; 456:949-51.
- McBride CS, Arguello JR, O'Meara BC. Five *Drosophila* genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics* 2007; 177:1395-416.
- Markow T, O'Grady PM. *Drosophila: A Guide to Species Identification and Use*. Academic Press, London 2006.
- O'Grady PM, Lapoint RT, Bennett GM. The potential and peril of the supertree approach: A response to van der Linde and Houle. *Insect Syst Evol* 2008; 39:269-80.
- Powell JR. *Progress and prospects in evolutionary biology: The Drosophila model*. Oxford University Press, Oxford 1997.
- O'Grady PM, DeSalle R. Out of Hawaii: The biogeographic history of the genus *Scaptomyza* (Diptera: Drosophilidae). *Biol Lett* 2008; 4:195-9.