FlyBase has started to collect and display this information from primary research articles. Many of these papers describe Drosophila models of human diseases or investigate functional conservation between human-fly orthology. To help researchers find genes and alleles that model human diseases in Drosophila FlyBase has started to collect and display this information from primary research articles.

The number and proportion of Drosophila articles published each year that mention disease has risen steadily. Disease models we have applied this to are at the allele level i.e. at-locus-mutations and transgenes.

What alleles are associated with standardized human disease terms from the Disease Ontology (disease-ontology.org). An allele’s relationship with a disease can be described as:
- a modifier: ameliorates or exacerbates a disease modeled by another allele
- When >1 allele models the disease, “in combination with” is used to describe their relationship.
- *DOES NOT* is used for cases when a particular allele fails to manifest the expected phenotype.
- Therefore using simple expressions there are multiple ways in which the relationship between an allele and disease can be described.

Finding disease-related papers
- To identify potential papers for disease curation we have:
  - Searched the archives for papers with “disease” in the title or abstract,
  - Examined papers known to use human transgenes,
- From this we have been able to associate ≥1000 alleles of over 500 genes, spanning ≥100 diseases (DO terms).

Drosophila has proven an incredibly tractable model for neurodegenerative disease. The fly has a complex brain and a short life-span enables the study of these diseases which often have an age-dependent pathology. It is therefore not surprising that the majority of Drosophila human-disease models relate to neurodegenerative disease. This is illustrated in the word cloud above, generated from the Disease Ontology terms we have associated with alleles in FlyBase.

FlyBase is supported by a grant from the National Human Genome Research Institute at the U.S. National Institutes of Health (HG000873). Support is also provided by the British Medical Research Council (this work was supported by grant number G8300948), the Indiana Genomics Initiative, and the National Science Foundation through XSEDE resources provided by Indiana University.