Title: Machine learning analysis identifies *Drosophila Grunge/Atrophin* as an important learning and memory gene required for memory retention and social learning.

Authors: Balint Z Kacsoh[†], Casey S. Greene[§] and Giovanni Bosco[†*]

Affiliations:

[†] Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, NH, 03755, USA.

[§] Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, PA, 19104, USA.

*Corresponding author: (GB) giovanni.bosco@dartmouth.edu

ABSTRACT

High throughput experiments are becoming increasingly common, and scientists must balance hypothesis driven experiments with genome wide data acquisition. We sought to predict novel genes involved in Drosophila learning and long-term memory from existing public high-throughput data. We performed an analysis using PILGRM, which analyzes public gene expression compendia using machine learning. We evaluated the top prediction alongside genes involved in learning and memory in IMP, an interface for functional relationship networks. We identified *Grunge/Atrophin* (*Gug/Atro*), a transcriptional repressor, histone deacetylase, as our top candidate. We find, through multiple, distinct assays, that *Gug* has an active role as a modulator of memory retention in the fly and its

function is required in the adult mushroom body. Depletion of *Gug* specifically in neurons of the adult mushroom body, after cell division and neuronal development is complete, suggests that *Gug* function is important for memory retention by regulating neuronal activity, and not simply by altering neurodevelopment. Our study provides a previously uncharacterized role for *Gug* as a possible regulator of neuronal plasticity at the interface of memory retention and memory extinction.