Submitter Name: J. Spencer Hatfield Submitter Email: jshatfi@clemson.edu

Genome-wide Association Analysis of Cocaine Preference in *Drosophila*

J. Spencer Hatfield[1], Vijay Shankar[1], Trudy F. C. Mackay[1], Robert R. H. Anholt[1]

[1]Department of Genetics and Biochemistry and Center for Human Genetics, Clemson University, Greenwood, S.C., U.S.A.

Genome-wide association studies (GWAS) of cocaine use disorder in human populations are challenging due to limited sample sizes, heterogeneity of genetic background, and environmental variability. *Drosophila melanogaster* presents a powerful model for GWAS of cocaine preference, a proxy for addiction behavior. Here we present a GWAS of 600 distinct genetic backgrounds of the wild-derived, inbred, sequenced *Drosophila* Genetic Reference Panel (DGRP). Using a high-throughput microplate feeder assay, we measured cocaine preference for more than 70,000 flies. We gave individual flies a choice between a control sucrose solution and one supplemented with cocaine and quantified cocaine preference for each fly after a 22-hour exposure. We found significant genetic variation for cocaine preference across the DGRP lines with significant sexual dimorphism. About 9% of DGRP lines exhibited cocaine preference with males on average exhibiting higher cocaine preference than females of the same line. GWAS identified 468 variants associated with cocaine preference at $p < 1x10^{-5}$ based on quantile-quantile plots, with multiple variants surpassing Bonferroni correction. Among the top associations are *Shaw*, which encodes a voltage-gated potassium channel, orthologous to human *KCNC2*; *CG33281*, which encodes a carbohydrate transporter orthologous to human *SLC2A6*; and *Btk29A*, which encodes Bruton tyrosine kinase, a target for ibrutinib which reduces startle-induced seizures in flies exposed to cocaine. Our observations that innate cocaine preference is dependent on sex and genetic background, and the identification of candidate genes not previously associated with cocaine use or addiction, is relevant for genetic risk susceptibility for cocaine use disorder in human populations.