

Improved representation of behavior data in the Rat Genome Database's PhenoMiner tool

Jennifer R Smith¹, Stanley JF Laulederkind¹, G Thomas Hayman¹, Shur-Jen Wang¹, Monika Tutaj¹, Mary L Kaldunski¹, Mahima Vedi¹, Wendy M Demos¹, Marek A Tutaj¹, Jyothi Thota¹, Logan Lamers¹, Adam C Gibson¹, Akhilanand Kundurthi¹, Varun Reddy Gollapally¹, Kent C Brodie², Stacy Zacher³, Jeffrey L De Pons¹, Melinda R Dwinell¹, Anne E Kwitek¹

¹Rat Genome Database, Department of Physiology, Medical College of Wisconsin, Milwaukee, WI, USA. ²Clinical and Translational Science Institute, Medical College of Wisconsin, Milwaukee, WI, USA. ³Finance and Administration, Medical College of Wisconsin, Milwaukee, WI, USA.

The Rat Genome Database (RGD, <https://rgd.mcw.edu>) is the premiere online resource for genetic, genomic, disease and phenotype data for rat, as well as a substantial corpus of integrated data for human, mouse and seven other mammalian species used as models for human disease.

Rodents in general, and rats in particular, have traditionally been the models of choice for studies of behavior, emotionality, and addiction. However, historically the extent of such data at RGD has been relatively limited. To rectify this, RGD is embarking on several projects to expand the representation of addiction, movement, and behavior data in the database. Work is underway to curate qualitative disease and phenotype data related to addiction for rat and human. A project to import addiction- and behavior-related GWAS data for rat is in the planning stages. RGD already regularly imports human GWAS data from the GWAS Catalog, including a significant number of addiction-related results. In addition, the representation of quantitative movement and behavior phenotypes in the RGD PhenoMiner database and tool are currently being overhauled.

Early work with quantitative behavior measurements used relatively general terms to group measurements wherever possible. However, user feedback clearly indicated that this approach was not optimal. The first step toward improving the representation of behavior in PhenoMiner has been redevelopment of the vocabularies used to specify measurements made and methods used, including work with subject matter experts to ensure that the data and metadata are clear and correctly represented. Second, the experimental results currently in the database are being re-curated to make the data more usable, findable, and understandable. Finally, focused curation of published literature and imports of larger bulk datasets directly from researchers will substantially improve and expand behavior and addiction measurement data in PhenoMiner.