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

Jennifer R. Smith

Scientific Project Manager
Rat Genome Database
Medical College of Wisconsin, Milwaukee, WI, USA

2024 Complex Trait Community and Rat Genomics Meeting

Introduction

- RGD has in the past lacked a substantial amount of data related to behavior, emotionality and addiction.
- Projects to increase this type of data that are either underway or in the *planning stages*:
  - Curate qualitative disease and phenotype data related to addiction for rat and human
  - Import human GWAS data from the GWAS Catalog, including data for behavioral traits, addiction, and diseases of mental health
  - Improve the representation of quantitative movement- and behavior-related phenotypes for rat
  - *Import GWAS data for rat, including data for behavioral traits and addiction*

Project: Curate qualitative disease and phenotype data related to addiction

- Curation priorities are set by
  - Maintenance of existing disease portals
  - The portal planned for next release
- The Addiction Portal is slated to be released in mid-2025.
- Terms related to addiction or behavior have been among the most used for annotation in the past year.

Term Acc	Term	# New Annotations
DOID:0050741	alcohol dependence	164
DOID:1574	alcohol use disorder	93
DOID:9976	heroin dependence	76
DOID:0050742	nicotine dependence	72
DOID:9975	cocaine dependence	69
DOID:2559	opiate dependence	54
DOID:0060001	withdrawal disorder	46
HP:0030955	Addictive alcohol use	41
DOID:9970	obesity	39
DOID:2560	morphine dependence	38
DOID:10591	pre-eclampsia	36
DOID:0050696	fetal alcohol spectrum disorder	34
DOID:9001131	stress-related disorder	32
DOID:9001234	prenatal exposure delayed effects	28
DOID:9006302	binge drinking	26
DOID:5844	myocardial infarction	26
DOID:670	amphetamine abuse	25
DOID:9006358	postoperative cognitive dysfunction	24
DOID:9002735	alcohol withdrawal syndrome	24
DOID:9004354	alcohol-related disorders	23

Project: Import human GWAS data from GWAS Catalog

- RGD recently began to import human GWAS data from the joint NHGRI-EBI GWAS Catalog database to provide a more substantial corpus of data for comparative studies.
- RGD curators are linking EFO terms to terms in the disease, phenotype, vertebrate trait and clinical measurement ontologies that RGD uses for curation. (Poster by Laulederkind, et al.)
- RGD has created and annotated both variant and quantitative trait locus records for the imported GWAS SNPs. (Poster by Smith, et al.)
- The "disease traits" studied cover a wide range of diseases, phenotypes and measurements, including a number related to addiction, behavior and mental health.

Term	# QTL annotated	# Variants annotated
alcohol consumption measurement	3212	2241
cigarettes per day measurement	734	596
substance abuse	591	591
alcohol use disorder measurement	364	315
smoking behaviour measurement	279	270
alcohol dependence	274	259
nicotine dependence symptom count	239	221
opioid dependence	184	182
cannabis dependence	146	144

Project: Improve quantitative movement and behavior phenotype data in RGD

- PhenoMiner is RGD's quantitative phenotype data store and mining tool.
- PhenoMiner is designed around the use of standardized vocabularies to specify
  - What trait was studied (Vertebrate Trait Ontology)
  - What animals were studied (Rat Strain Ontology)
  - What was measured (Clinical Measurement Ontology)
  - How it was measured (Measurement Method Ontology)
  - Under what conditions it was measured (Experimental Condition Ontology)
- Sometimes this is not as straightforward as it seems.

Project: Improve quantitative movement and behavior phenotype data in RGD

General

A genetic and multifactorial analysis of anxiety-related behaviours in Lewis and S

Authors:	Ramos, A Mellerin, Y Mormede, P Chaouloff, F
Citation:	Ramos A, etal., Behav Brain Res. 1998 Nov;96(1-2):195-205.
RGD ID:	4109649
Pubmed:	PMID:9821556 (View Abstract at PubMed)

- Selected one study to re-curate
- Two strains, control conditions
- Relatively few methods
  - Elevated plus maze (EPM)
  - Open field (OF)
- Relatively few measurements
  - Time in open/closed arms
  - # total/closed arm entries
  - % entries into open arms
  - Inner/outer locomotion in OF
  - Defecation

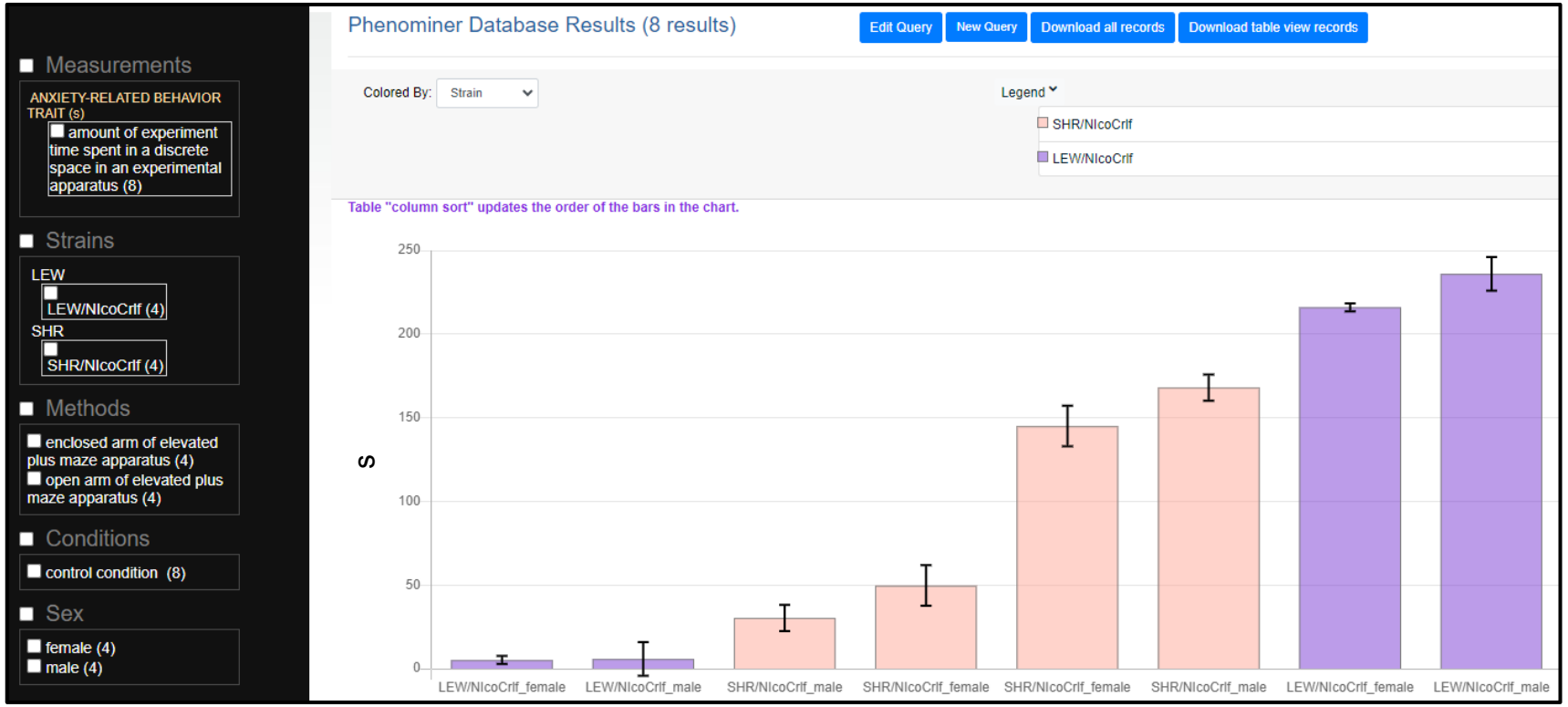
Elevated plus-maze, open field and systolic BP measures (X ± S.E.M.) of LEW and SHR rats grouped by sex

Variable	Males		Females	
	LEW	SHR	LEW	SHR
Plus maze				
Time in open arms (s)	5.8 ± 3.7	30.3 ± 7.8*	5.2 ± 2.4	49.7 ± 12.1*
Time in closed arms (s)	236 ± 10	168 ± 13*	216 ± 11	145 ± 13**
Open arm entries (%)	16.3 ± 7.6	28.8 ± 5.0	10.1 ± 6.8	35.3 ± 5.0*
Closed arm entries	3.8 ± 0.7	4.6 ± 0.5	7.0 ± 1.5	5.2 ± 0.4
Total arm entries	4.8 ± 1.1	6.8 ± 0.8	7.8 ± 1.6	8.5 ± 0.9
Open field				
Outer locomotion	47.8 ± 7.5	58.2 ± 4.7	74.3 ± 12.3	79.3 ± 4.4#
Inner locomotion	1.0 ± 0.0	4.7 ± 0.7**	1.3 ± 0.2	13.1 ± 2.1** # #
Defecation	0.6 ± 0.6	0.1 ± 0.1	0.0 ± 0.0	0.1 ± 0.1
BP (mmHg)	161 ± 4	229 ± 2**	157 ± 2	194 ± 2** # #

Project: Improve quantitative movement and behavior phenotype data in RGD

Current:

- What was measured:
  - amount of experiment time spent in a discrete space in an experimental apparatus
- How it was measured:
  - enclosed arm of elevated plus maze apparatus
  - open arm of elevated plus maze apparatus

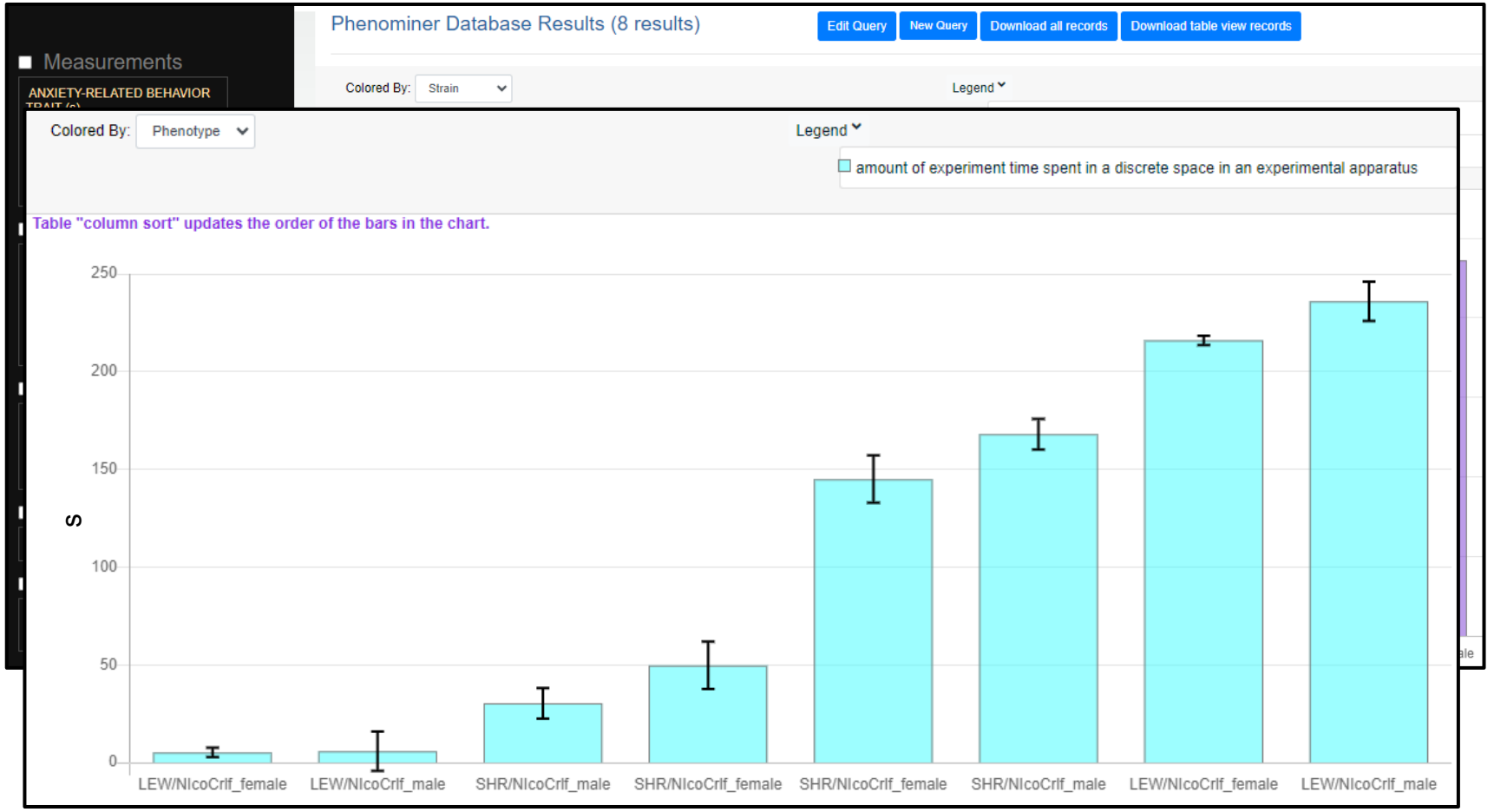


https://rgd.mcw.edu

Project: Improve quantitative movement and behavior phenotype data in RGD

Current:

- What was measured:
  - amount of experiment time spent in a discrete space in an experimental apparatus
- How it was measured:
  - enclosed arm of elevated plus maze apparatus
  - open arm of elevated plus maze apparatus

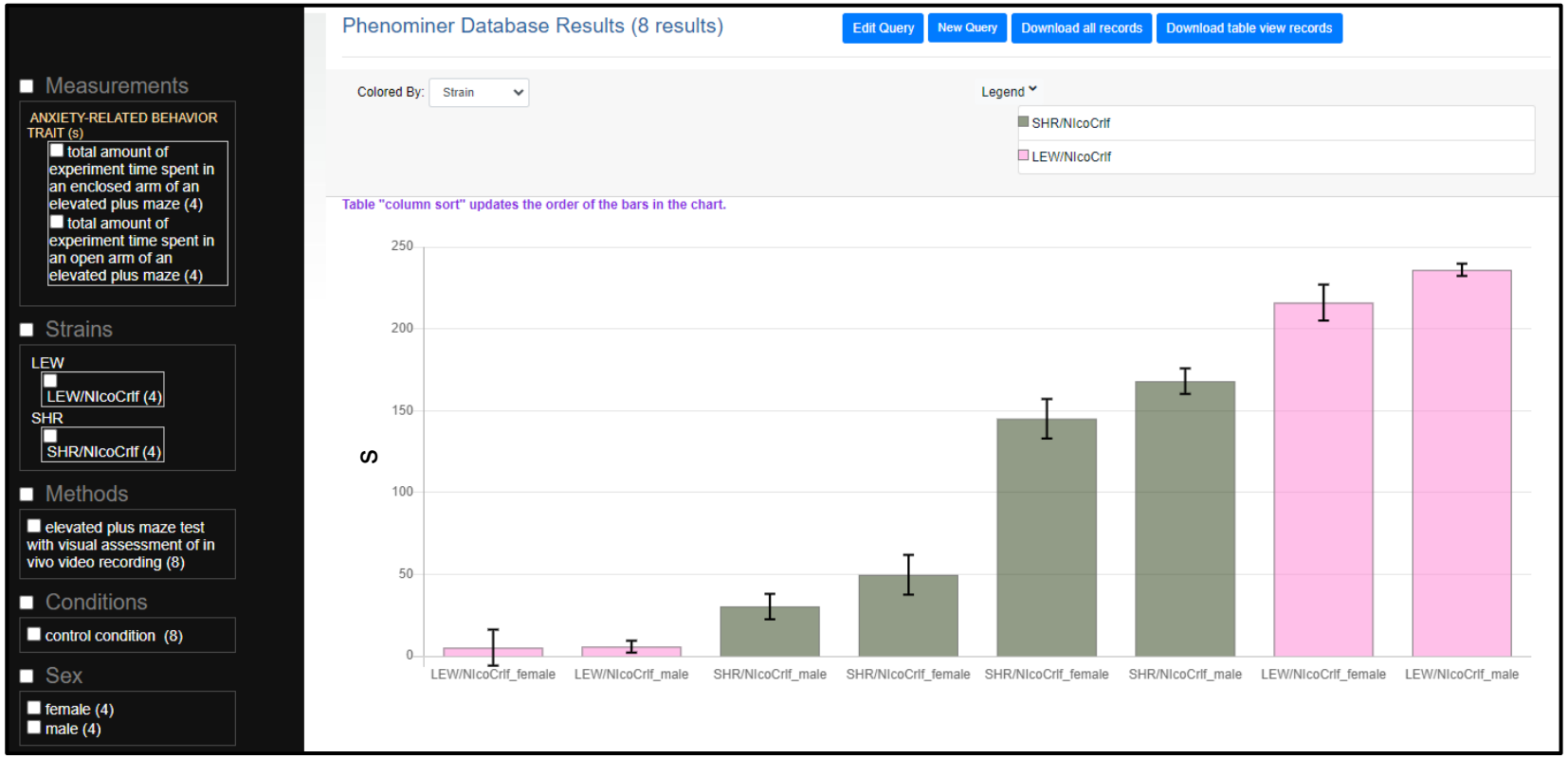


https://rgd.mcw.edu

Project: Improve quantitative movement and behavior phenotype data in RGD

Proposed changes:

- What was measured:
  - ➢ total amount of experiment time spent in an enclosed arm of an elevated plus maze
  - ➢ total amount of experiment time spent in an open arm of an elevated plus maze
- How it was measured:
  - ➢ elevated plus maze test with visual assessment of in vivo video recording

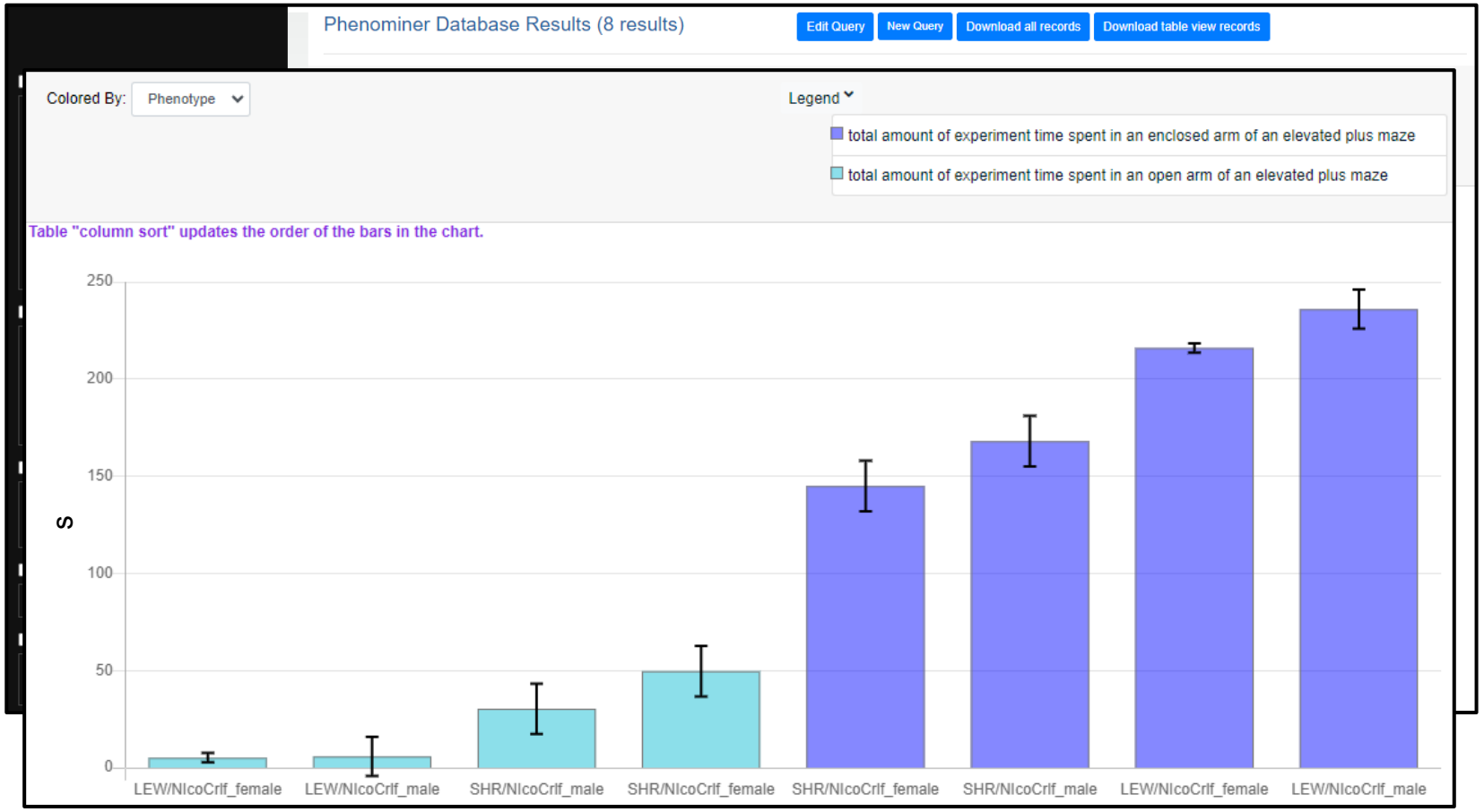


https://rgd.mcw.edu

Project: Improve quantitative movement and behavior phenotype data in RGD

Proposed changes:

- What was measured:
  - total amount of experiment time spent in an enclosed arm of an elevated plus maze
  - total amount of experiment time spent in an open arm of an elevated plus maze
- How it was measured:
  - elevated plus maze test with visual assessment of in vivo video recording



https://rgd.mcw.edu

Project: Improve quantitative movement and behavior phenotype data in RGD

- Proposed changes:
  - Multiple traits for a single experiment
    - Distance traveled can be a measurement of either locomotor behavior or fear/anxiety (or both)
  - Move information like "open arm of an elevated plus maze" out of the method and into the clinical measurement terms
    - The open arm of an elevated plus maze is not a method
    - Ensure that clinical measurement terms faithfully represent the measurements that were made
    - Ensure that measurements that are qualitatively different are not lumped together
  - Work more closely with subject matter experts for both ontology development and establishing curation practices for movement, behavior and addiction data

Project: Improve quantitative movement and behavior phenotype data in RGD

- Future work:
  - Continue work to redevelop the PhenoMiner vocabularies for behavior and addiction measurements
  - Work with subject matter experts to ensure that the vocabularies are usable for researchers and the data and metadata are clear and correctly represented
  - Establish best practices for curation of quantitative movement, behavior and addiction data
  - Review and re-curate the data that has already been entered into PhenoMiner for behavior
  - Focused curation of literature for movement, behavior and addiction data
  - Find and import bulk datasets directly from researchers

Project: Improve quantitative movement and behavior phenotype data in RGD

If you study rat behavior, we need your help!

We are looking for Subject Matter Experts to provide input on any and all aspects of quantitative behavior data curation and visualization.

https://rgd.mcw.edu

Thank you!

jrsmith@mcw.edu
https://rgd.mcw.edu

The RGD Team:

Principal Investigator:
Anne Kwitek, PhD

Co-Investigator:
Mindy Dwinell, PhD

Curation Team:
Jennifer Smith, MSc
Stan Laulederkind, PhD
Tom Hayman, PhD
Shur-Jen Wang, PhD
Monika Tutaj, PhD
Mary Kaldunski, MSc
Mahima Vedi, PhD
Wendy Demos, MSc

Development Team:
Jeff De Pons, BSc
Marek Tutaj, MSc
Jyothi Thota, MSc
Logan Lamers, BSc
Adam Gibson, BSc
Akhilanand Kundurthi, MSc
Varun Reddy Gollapally, MSc

Systems Administrator:
Kent Brodie, MSMI

Database Administrator:
Stacy Zacher, MSc

We gratefully acknowledge our funders:
RGD is funded by the National Heart, Lung, and Blood Institute (NHLBI; R01HL064541), and the National Human Genome Research Institute (NHGRI) as part of the Alliance of Genome Resources (U24HG010859).

We also acknowledge our collaborators and the researchers who contribute data and who faithfully use our website and data!

physiology

MEDICAL COLLEGE OF WISCONSIN

RGD

ALLIANCE of GENOME RESOURCES FOUNDING MEMBER

GBC GLOBAL CORE BIODATA RESOURCE