THE OHIO STATE UNIVERSITY
WEXNER MEDICAL CENTER

Searle Lab

Exploring Biofluid Protein Expression with CATalog, an Interactive Proteomics Dashboard

Alex W. Joyce[1], Katelyn B. Brusach[2], Ariana Shannon[1,3], Jessica M. Quimby[2], and Brian C. Searle[1,3]

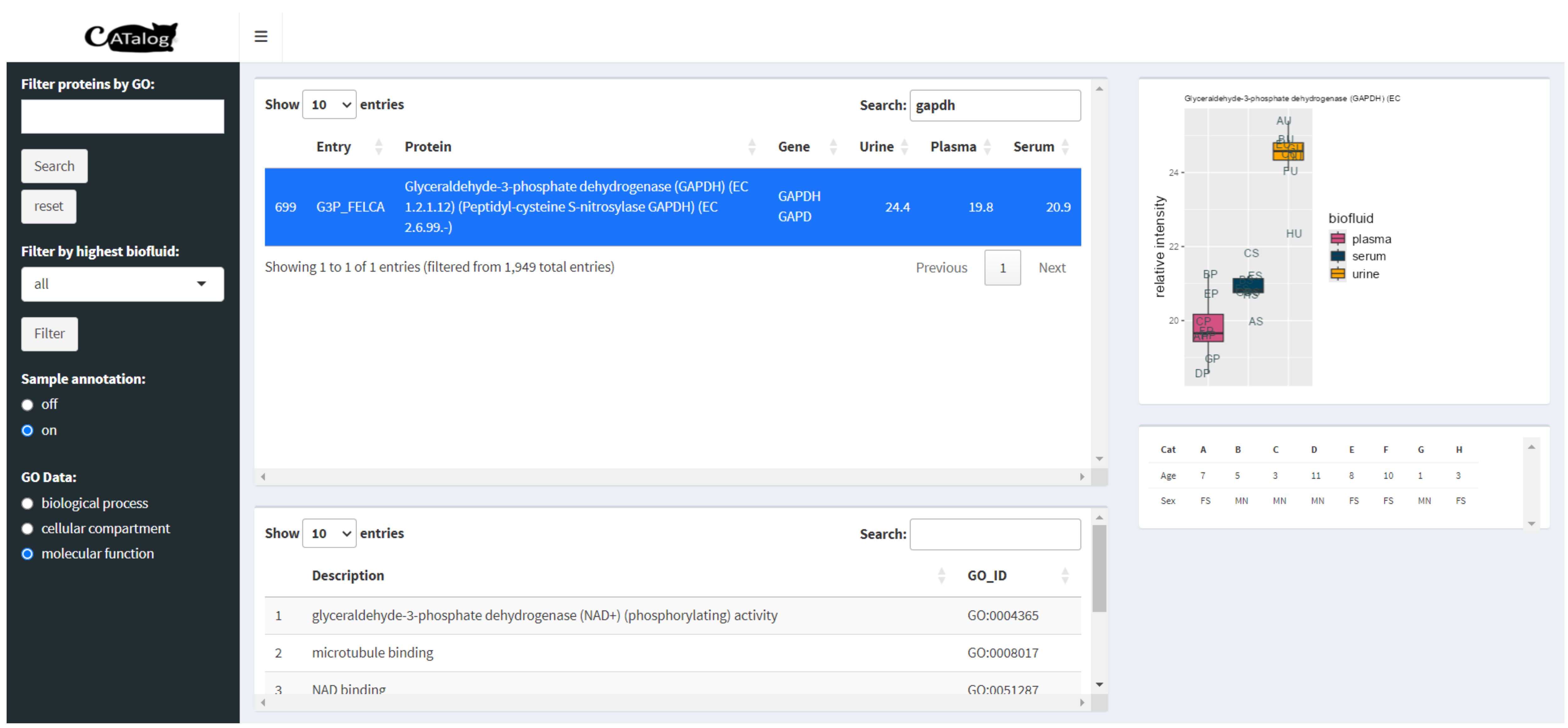
1-Department of Biomedical Informatics, 2-Veterinary Clinical Sciences, 3 – Pelatonia Institute of Immuno-Oncology

Searle

Introduction

Biomarkers are a vital tool for identifying aberrant states in a biological system, either for early detection of a disease or for assessing symptoms of an existing condition. One of the abundant sources of biomarkers are biofluids which can be easily sampled from an organism and stored for later analysis. However, due to the diversity of potential biomarkers, they oftentimes appear in different concentrations among various types of biofluids, making it difficult to identify the optimal sample to analyze. CATalog can assist in both biofluid and biomarker selection by displaying the relative abundances of selected biomarkers from healthy cats across three paired samples (urine, serum, and plasma) [1] collected using mass spectrometry. Additionally, relevant gene ontology (GO) [2] information is provided for each biomarker, showing involved biological processes, subcellular localization, and molecular function, annotating each biomarker and assisting with selection.

CATalog Presents Data in an Interactive Format



CATalog

Filter proteins by GO:

Search

reset

Filter by highest biofluid:

all

Filter

Sample annotation:

○ off

● on

GO Data:

○ biological process

○ cellular compartment

● molecular function

Show 10 entries Search: gapdh

	Entry	Protein	Gene	Urine	Plasma	Serum
699	G3P_FELCA	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (EC 1.2.1.12) (Peptidyl-cysteine S-nitrosylase GAPDH) (EC 2.6.99.-)	GAPDH GAPD	24.4	19.8	20.9

Showing 1 to 1 of 1 entries (filtered from 1,949 total entries) Previous 1 Next

Show 10 entries Search:

	Description	GO_ID
1	glyceraldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity	GO:0004365
2	microtubule binding	GO:0008017
3	NAD binding	GO:0051287

Cat	A	B	C	D	E	F	G	H
Age	7	5	3	11	8	10	1	3
Sex	FS	MN	MN	MN	FS	FS	MN	FS

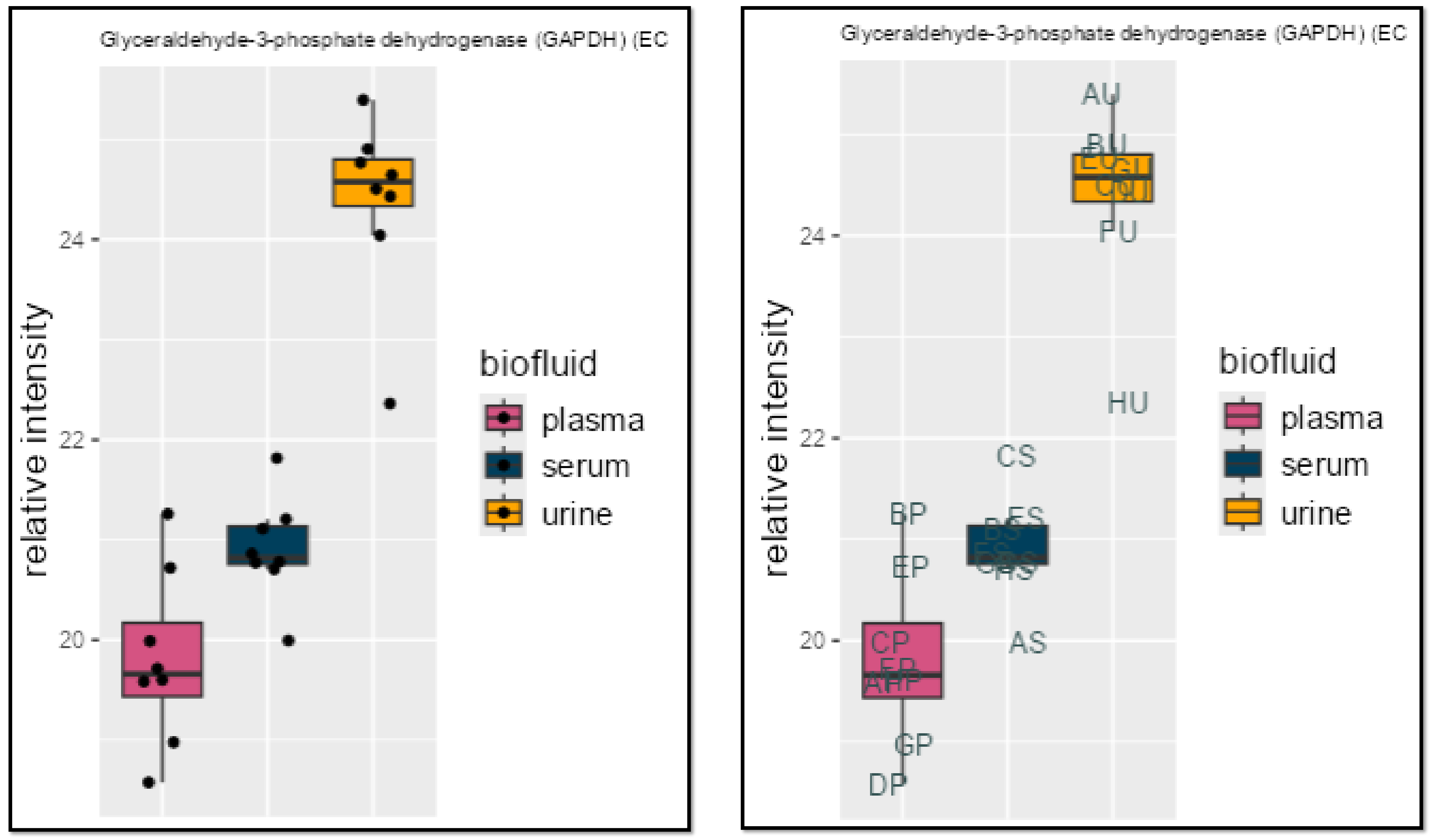
GO Information Facilitates Biomarker Selection

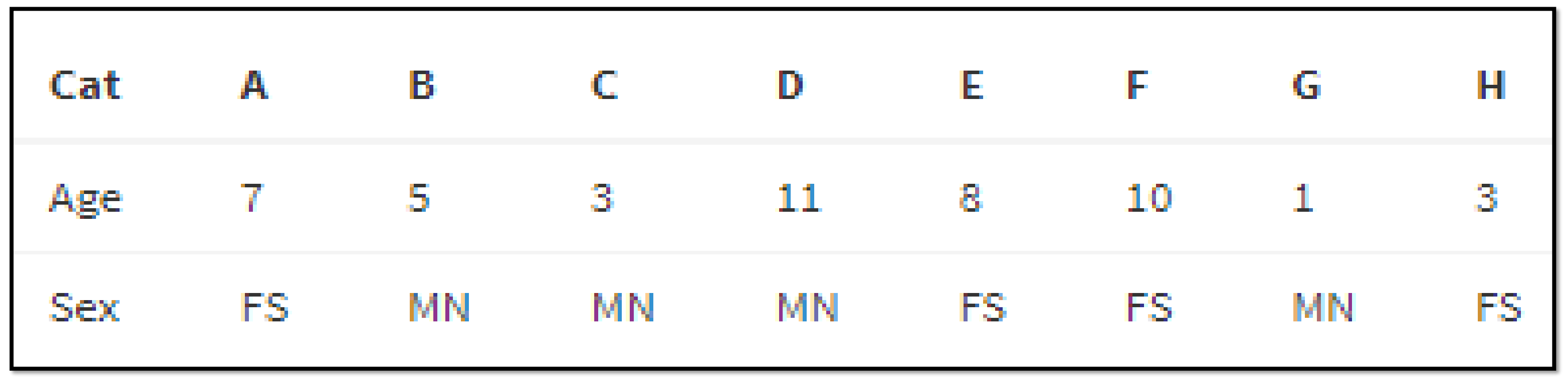
Figure 2: Table of GO information for the glyceralderhyde-3-phosphatase dehydrogenase, concatenated from the three individual tables in CATalog.

GO Term	GO ID
Biological Process	
glucose metabolic process	GO:0006006
glycolytic process	GO:0006096
innate immune response	GO:0045087
…	…
Cellular Compartment	
cytoplasm	GO:0005737
cytosol	GO:0005829
GAIT complex	GO:0097452
,,,	,,,
Molecular Function	
glyceraldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity	GO:0004365
microtubule binding	GO:0008017
NAD binding	GO:0051287
…	…

Boxplots Provide Demographic Context

Figure 3: Each biomarker in the database has an associated boxplot which displays sample-level information of relative abundances. Annotations provide information about the corresponding cat's age and sex, connecting demographics to potential outliers in the biomarker's relative abundance.

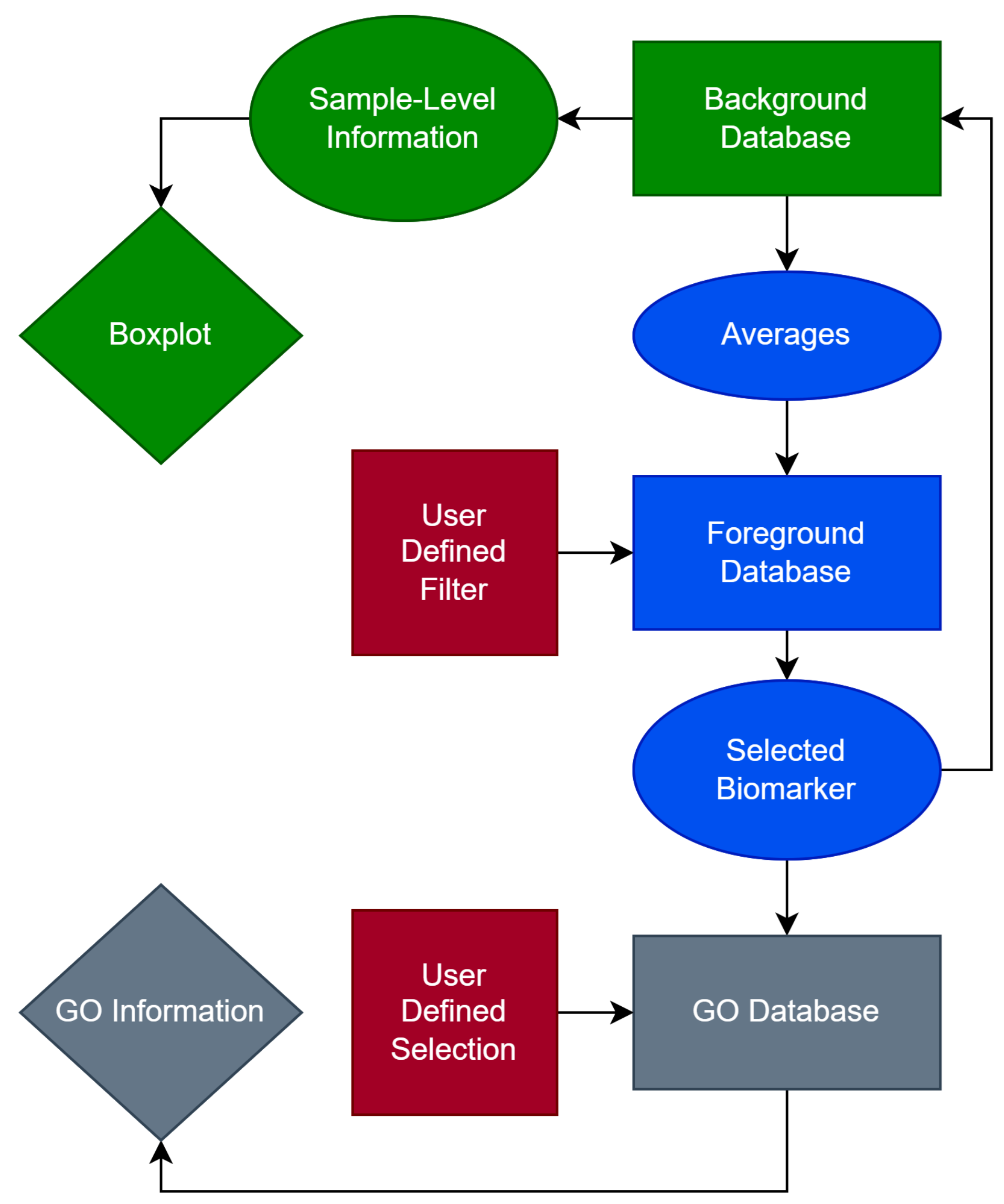




Cat	A	B	C	D	E	F	G	H
Age	7	5	3	11	8	10	1	3
Sex	FS	MN	MN	MN	FS	FS	MN	FS

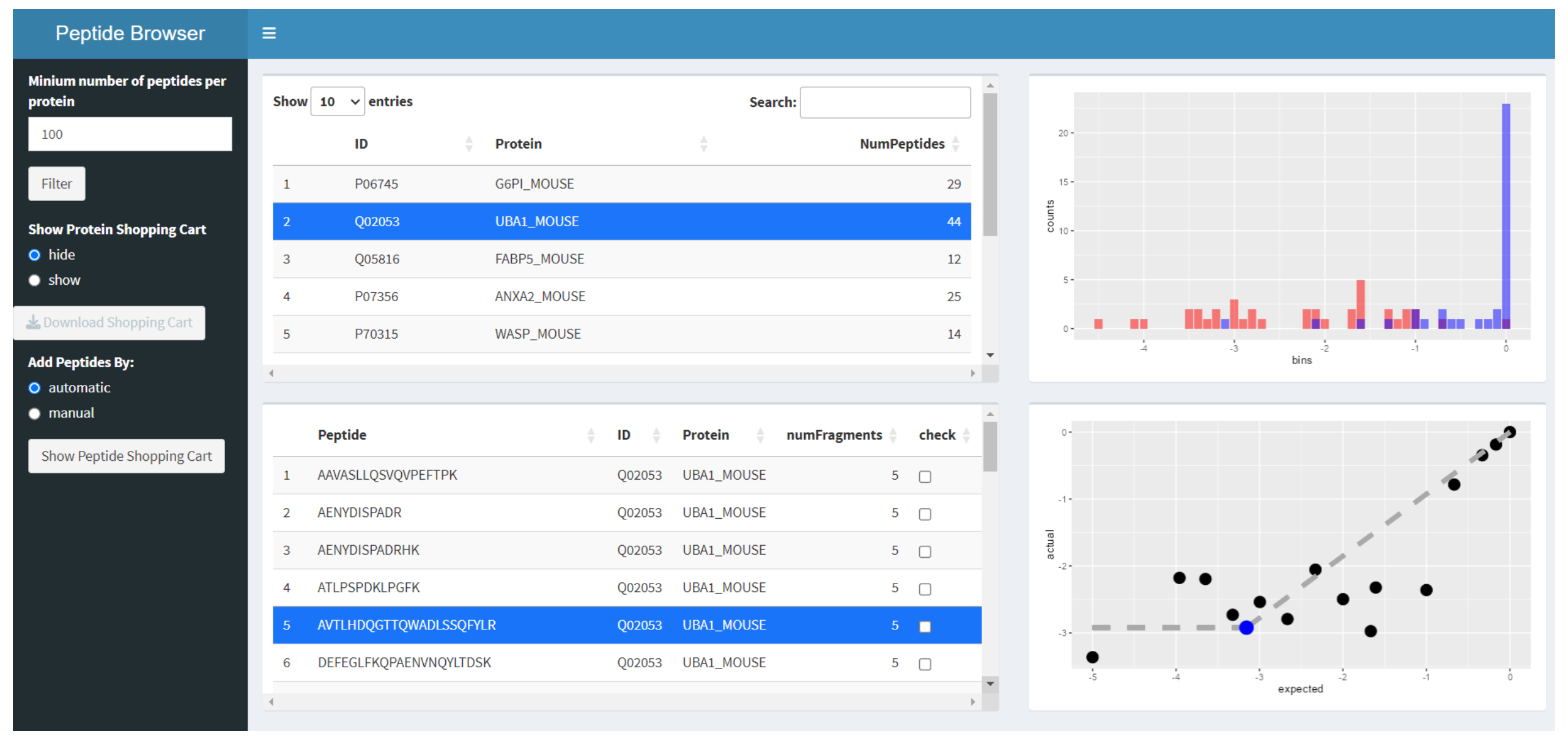
Data Structure

- CATalog was written in R using the Shiny framework which provides a modular structure.
- Only the foreground dataset is shown to the user, greatly reducing the amount of information at the forefront.
- Complete sample-level information is held in the background which is displayed to the user in the form of a boxplot.
- The gene ontology (GO) database was obtained from UniProt for the organism felus catus (taxon id:). Which portion of the information is displayed can be selected by the user.



Future Directions

- Adding a feature to download selected peptides from a "shopping cart"
- Comparison of user-uploaded data to enable direct comparison with CATalog's control data.
- Expansion of the central CATalog design to peptide-level quantitative analysis and/or with other species. This is currently in development as 'PeptideBrowser'.



Peptide Browser

Minium number of peptides per protein

100

Filter

Show Protein Shopping Cart

● hide

○ show

Download Shopping Cart

Add Peptides By:

● automatic

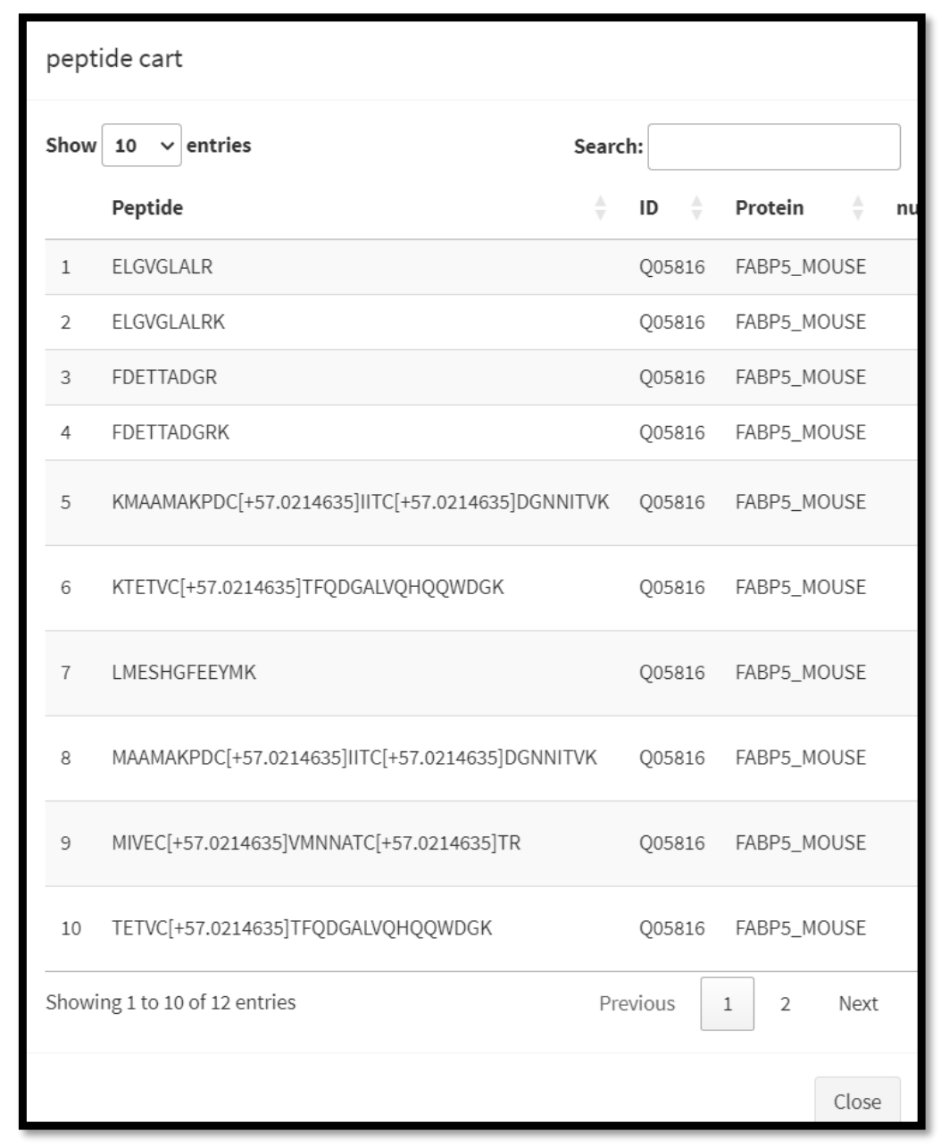
○ manual

Show Peptide Shopping Cart

Show 10 entries Search:

	ID	Protein	NumPeptides
1	P06745	G6PI_MOUSE	29
2	Q02053	UBA1_MOUSE	44
3	Q05816	FABP5_MOUSE	12
4	P07356	ANXA2_MOUSE	25
5	P70315	WASP_MOUSE	14

	Peptide	ID	Protein	numFragments	check
1	AAVASLLQSVQVPEFTPK	Q02053	UBA1_MOUSE	5	☐
2	AENYDISPADR	Q02053	UBA1_MOUSE	5	☐
3	AENYDISPADRHK	Q02053	UBA1_MOUSE	5	☐
4	ATLPSPDKLPGFK	Q02053	UBA1_MOUSE	5	☐
5	AVTLHDQGTTQWADLSSQFYLR	Q02053	UBA1_MOUSE	5	☐
6	DEFEGLFKQPAENVNQYLTDSK	Q02053	UBA1_MOUSE	5	☐



peptide cart

Show 10 entries Search:

	Peptide	ID	Protein
1	ELGVGLALR	Q05816	FABP5_MOUSE
2	ELGVGLALRK	Q05816	FABP5_MOUSE
3	FDETTADGR	Q05816	FABP5_MOUSE
4	FDETTADGRK	Q05816	FABP5_MOUSE
5	KMAAMAKPDC[+57.0214635]IITC[+57.0214635]DGNNITVK	Q05816	FABP5_MOUSE
6	KTETVC[+57.0214635]TFQDGALVQHQQWDGK	Q05816	FABP5_MOUSE
7	LMESHGFEEYMK	Q05816	FABP5_MOUSE
8	MAAMAKPDC[+57.0214635]IITC[+57.0214635]DGNNITVK	Q05816	FABP5_MOUSE
9	MIVEC[+57.0214635]VMNNATC[+57.0214635]TR	Q05816	FABP5_MOUSE
10	TETVC[+57.0214635]TFQDGALVQHQQWDGK	Q05816	FABP5_MOUSE

Showing 1 to 10 of 12 entries Previous 1 2 Next

Close