Alliance of Genome Resources

A Lesson in Resource Integration and Sustainability

Kimberly M. Van Auken, PhD
Alliance of Genome Resources
ACKnowledge
Gene Ontology Consortium
WormBase

Thursday, May 15th, 2025

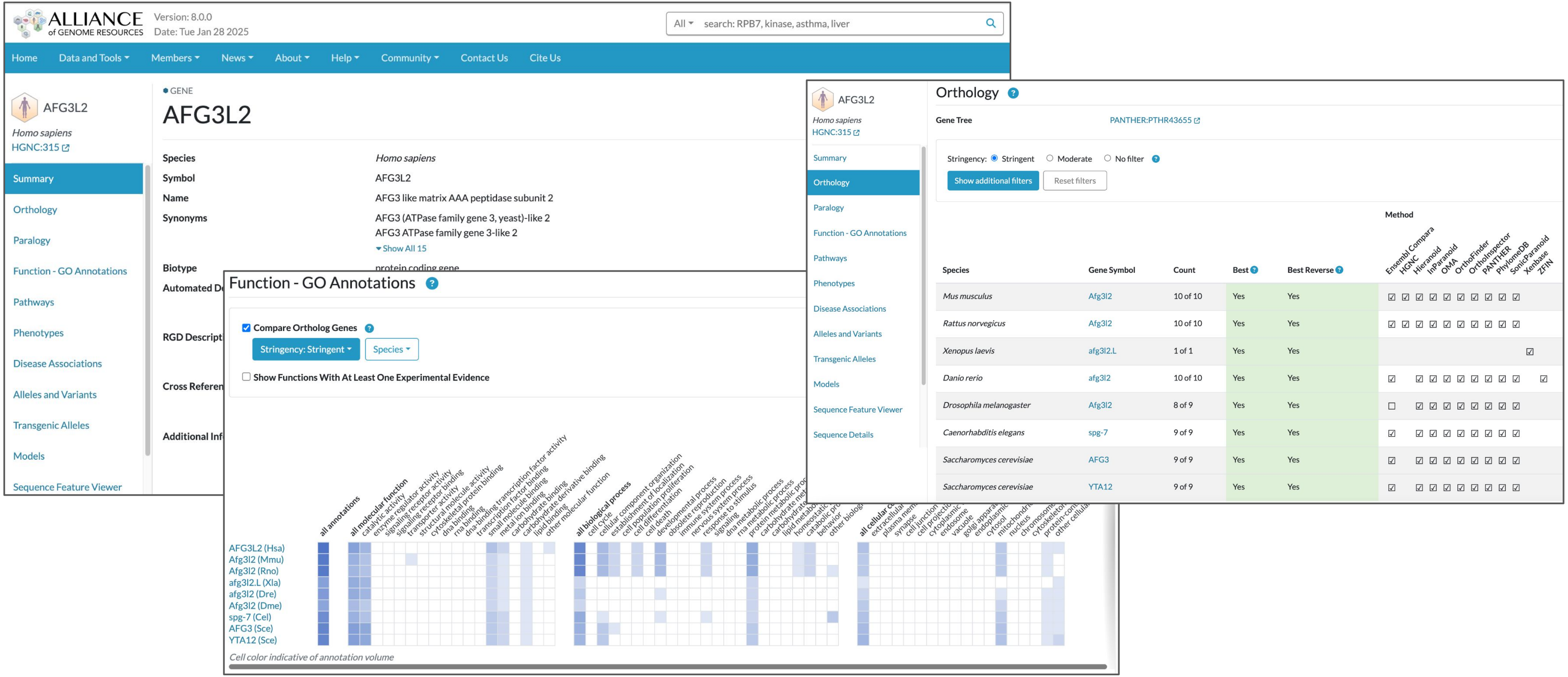
GBC
GLOBAL
CORE
BIODATA
RESOURCE

Alliance of Genome Resources:

a consortium of seven model organism knowledgebases and the Gene Ontology

GLOBAL
BIODATA
COALITION

Common, extensible platform for unified model organism and human data



AFG3L2
Homo sapiens
HGNC:315

GENE

AFG3L2

Species	*Homo sapiens*
Symbol	AFG3L2
Name	AFG3 like matrix AAA peptidase subunit 2
Synonyms	AFG3 (ATPase family gene 3, yeast)-like 2 AFG3 ATPase family gene 3-like 2 ▾Show All 15
Biotype	protein coding gene
Automated D	
RGD Descript	
Cross Referen	
Additional Inf	

Function - GO Annotations

☑ Compare Ortholog Genes

☐ Show Functions With At Least One Experimental Evidence

AFG3L2 (Hsa)
Afg3l2 (Mmu)
Afg3l2 (Rno)
afg3l2.L (Xla)
afg3l2 (Dre)
Afg3l2 (Dme)
spg-7 (Cel)
AFG3 (Sce)
YTA12 (Sce)

Cell color indicative of annotation volume

Orthology

AFG3L2
Homo sapiens
HGNC:315

Gene Tree PANTHER:PTHR43655

Stringency: ● Stringent ○ Moderate ○ No filter

Species	Gene Symbol	Count	Best	Best Reverse	Ensembl Compara	HGNC	Hieranoid	InParanoid	OMA	OrthoFinder	OrthoInspector	PANTHER	PhylomeDB	SonicParanoid	Xenbase	ZFIN
Mus musculus	Afg3l2	10 of 10	Yes	Yes	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑		
Rattus norvegicus	Afg3l2	10 of 10	Yes	Yes	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑		
Xenopus laevis	afg3l2.L	1 of 1	Yes	Yes											☑	
Danio rerio	afg3l2	10 of 10	Yes	Yes	☑		☑	☑	☑	☑	☑	☑	☑	☑		☑
Drosophila melanogaster	Afg3l2	8 of 9	Yes	Yes	☐		☑	☑	☑	☑	☑	☑	☑	☑		
Caenorhabditis elegans	spg-7	9 of 9	Yes	Yes	☑		☑	☑	☑	☑	☑	☑	☑	☑		
Saccharomyces cerevisiae	AFG3	9 of 9	Yes	Yes	☑		☑	☑	☑	☑	☑	☑	☑	☑		
Saccharomyces cerevisiae	YTA12	9 of 9	Yes	Yes	☑		☑	☑	☑	☑	☑	☑	☑	☑		

Why did the Alliance form?

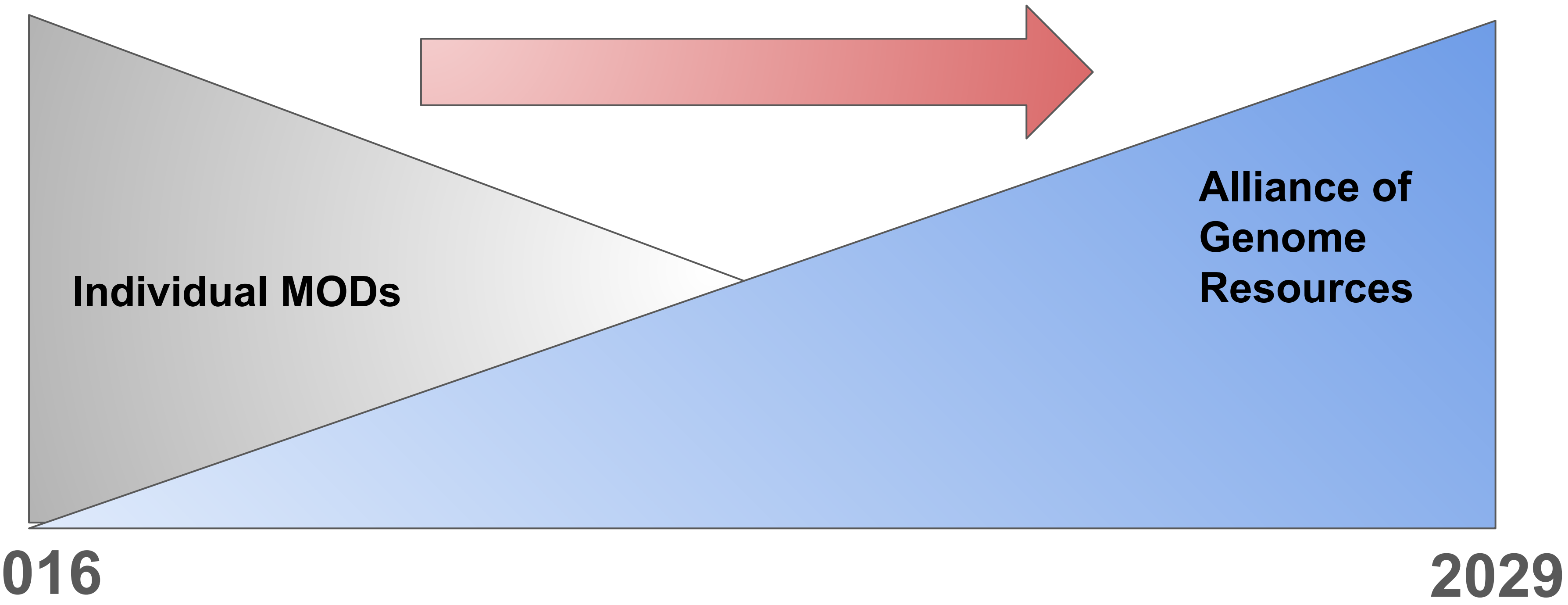
- Changes to funding priorities
- Desire for shared user experience (website, programmatic)
- Promote further data standardization
- Reduce redundant infrastructure, introduce new efficiencies

When and how did the Alliance form?

- NHGRI urged MODs and GOC to restructure organization, management, infrastructure - early 2016
- Initial concept from MODs and GOC presented to NHGRI - May 2016
- Formal proposal funded for first steps of integration work - September 2016

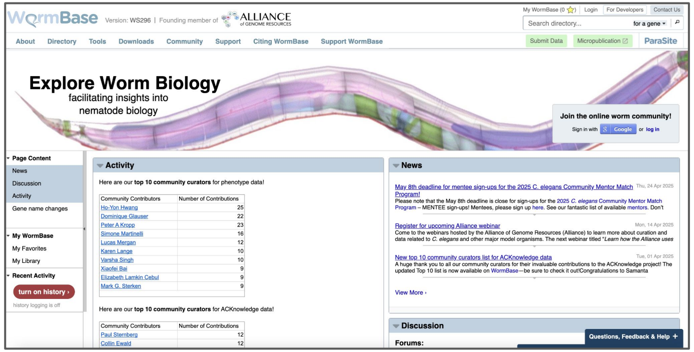
Alliance of Genome Resources Consortium. The Alliance of Genome Resources: Building a Modern Data Ecosystem for Model Organism Databases. Genetics. 2019

Approach: systematic harmonization and migration of data and software infrastructure



Core infrastructure of a model organism knowledgebase

- Data store
- API
- Website
    - data pages
    - analysis tools
    - file downloads
- Curation pipelines and tools



Activity

Here are our **top 10 community curators** for phenotype data!

Community Contributors	Number of Contributions
Ho-Yon Hwang	25
Dominique Glauser	22
Peter A Kropp	23
Simone Martinelli	16
Lucas Mergan	12
Karen Lange	10
Varsha Singh	10
Xiaofei Bai	9
Elizabeth Lamkin Cebul	9
Mark G. Sterken	9

Here are our **top 10 community curators** for ACKnowledge data!

Community Contributors	Number of Contributions
Paul Sternberg	12
Collin Ewald	12

News

May 8th deadline for mentee sign-ups for the 2025 C. elegans Community Mentor Match Program! Thu, 24 Apr 2025
Please note that the May 8th deadline is close for sign-ups for the 2025 *C. elegans* Community Mentor Match Program – MENTEE sign-ups! Mentees, please sign up here. See our fantastic list of available mentors. Don't

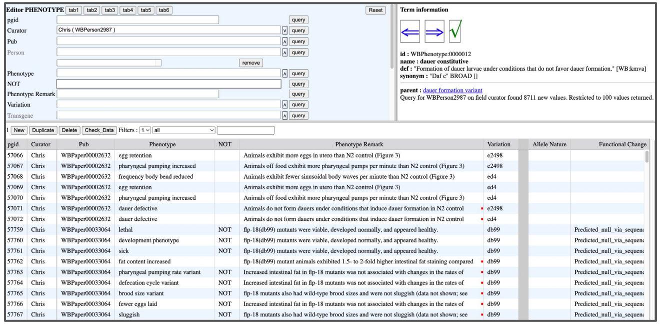
Register for upcoming Alliance webinar Mon, 14 Apr 2025
Come to the webinars hosted by the Alliance of Genome Resources (Alliance) to learn more about curation and data related to *C. elegans* and other major model organisms. The next webinar titled "*Learn how the Alliance uses*

New top 10 community curators list for ACKnowledge data Tue, 01 Apr 2025
A huge thank you to all our community curators for their invaluable contributions to the ACKnowledge project! The updated Top 10 list is now available on WormBase—be sure to check it out!Congratulations to Samanta

View More ›

Discussion

Forums:



pgid
Curator: Chris (WBPerson2987)
Pub
Person
Phenotype
NOT
Phenotype Remark
Variation
Transgene

1 Filters : 1 all

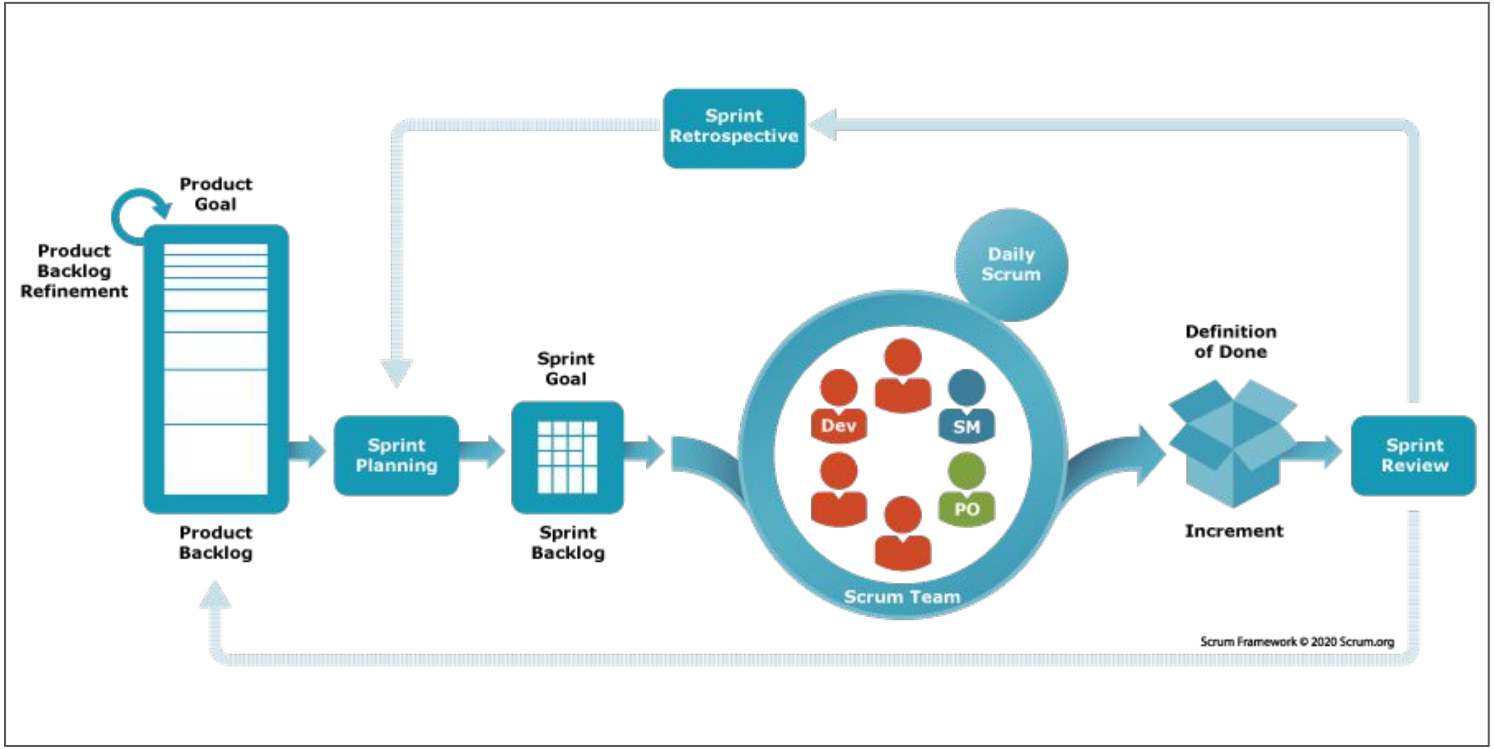
Term information

id : WBPhenotype:0000012
name : dauer constitutive
def : "Formation of dauer larvae under conditions that do not favor dauer formation." [WB:kmva]
synonym : "Daf c" BROAD []

parent : dauer formation variant
Query for WBPerson2987 on field curator found 8711 new values. Restricted to 100 values returned.

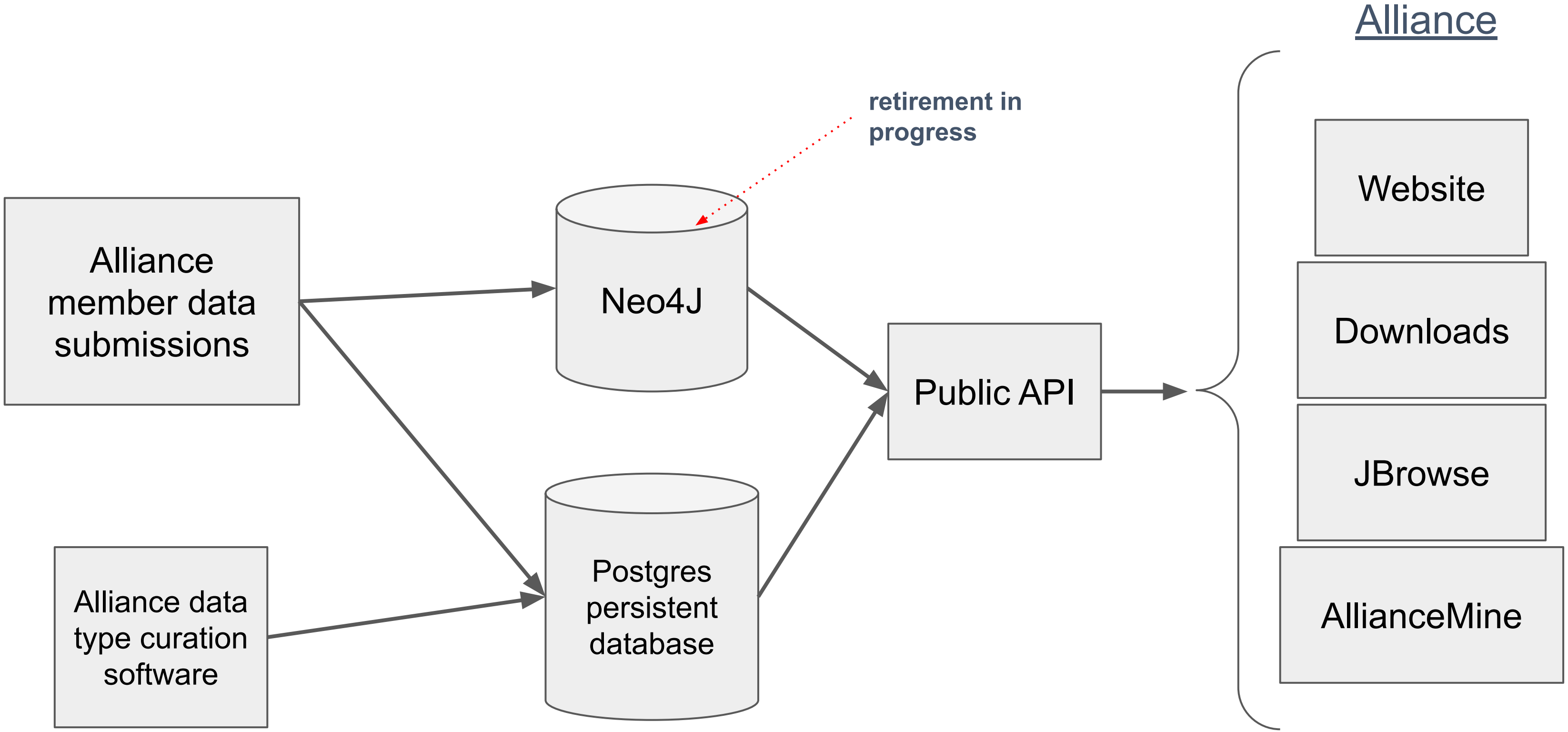
pgid	Curator	Pub	Phenotype	NOT	Phenotype Remark	Variation	Allele Nature	Functional Change
57066	Chris	WBPaper00002632	egg retention		Animals exhibit more eggs in utero than N2 control (Figure 3)	e2498		
57067	Chris	WBPaper00002632	pharyngeal pumping increased		Animals off food exhibit more pharyngeal pumps per minute than N2 control (Figure 3)	e2498		
57068	Chris	WBPaper00002632	frequency body bend reduced		Animals exhibit fewer sinusoidal body waves per minute than N2 control (Figure 3)	ed4		
57069	Chris	WBPaper00002632	egg retention		Animals exhibit more eggs in utero than N2 control (Figure 3)	ed4		
57070	Chris	WBPaper00002632	pharyngeal pumping increased		Animals off food exhibit more pharyngeal pumps per minute than N2 control (Figure 3)	ed4		
57071	Chris	WBPaper00002632	dauer defective		Animals do not form dauers under conditions that induce dauer formation in N2 control	e2498		
57072	Chris	WBPaper00002632	dauer defective		Animals do not form dauers under conditions that induce dauer formation in N2 control	ed4		
57759	Chris	WBPaper00033064	lethal	NOT	flp-18(db99) mutants were viable, developed normally, and appeared healthy.	db99		Predicted_null_via_sequenc
57760	Chris	WBPaper00033064	development phenotype	NOT	flp-18(db99) mutants were viable, developed normally, and appeared healthy.	db99		Predicted_null_via_sequenc
57761	Chris	WBPaper00033064	sick	NOT	flp-18(db99) mutants were viable, developed normally, and appeared healthy.	db99		Predicted_null_via_sequenc
57762	Chris	WBPaper00033064	fat content increased		flp-18(db99) mutant animals exhibited 1.5- to 2-fold higher intestinal fat staining compared	db99		Predicted_null_via_sequenc
57763	Chris	WBPaper00033064	pharyngeal pumping rate variant	NOT	Increased intestinal fat in flp-18 mutants was not associated with changes in the rates of	db99		Predicted_null_via_sequenc
57764	Chris	WBPaper00033064	defecation cycle variant	NOT	Increased intestinal fat in flp-18 mutants was not associated with changes in the rates of	db99		Predicted_null_via_sequenc
57765	Chris	WBPaper00033064	brood size variant	NOT	flp-18 mutants also had wild-type brood sizes and were not sluggish (data not shown; see	db99		Predicted_null_via_sequenc
57766	Chris	WBPaper00033064	fewer eggs laid	NOT	Increased intestinal fat in flp-18 mutants was not associated with changes in the rates of	db99		Predicted_null_via_sequenc
57767	Chris	WBPaper00033064	sluggish	NOT	flp-18 mutants also had wild-type brood sizes and were not sluggish (data not shown; see	db99		Predicted_null_via_sequenc

Approach: Agile and Scrum methodologies



https://www.scrum.org/resources/what-scrum-module

Database, Website, Data Curation: Scrum A Team



Alliance

retirement in progress

Alliance member data submissions

Alliance data type curation software

Neo4J

Postgres persistent database

Public API

Website

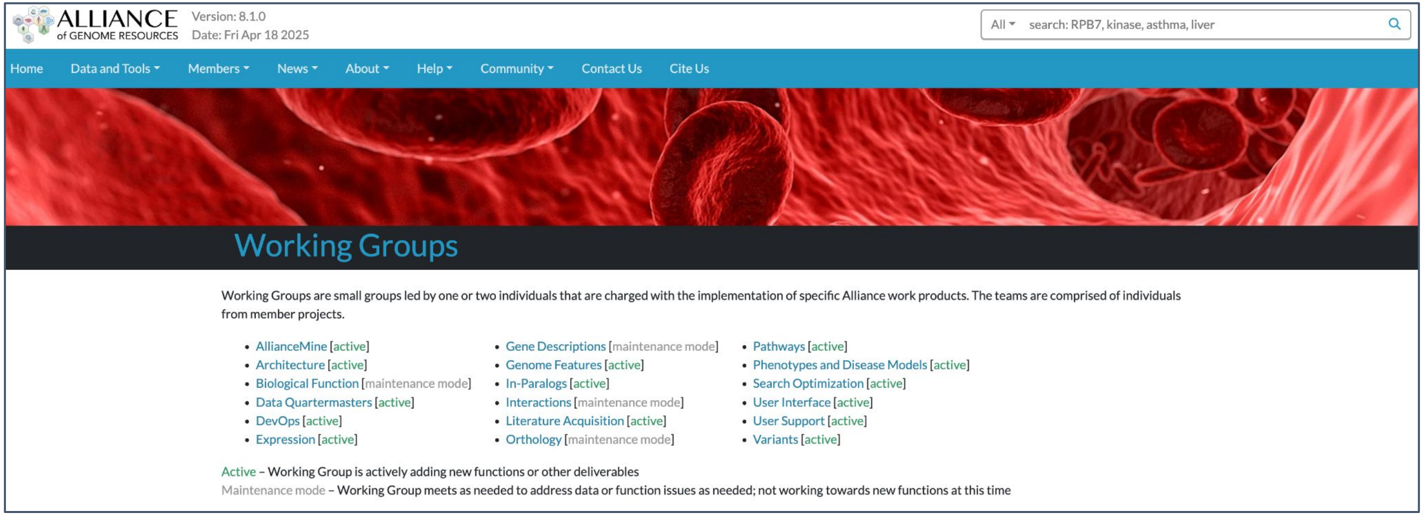
Downloads

JBrowse

AllianceMine

Alliance member data submissions

- Identify what types of data to include
- Understand how each member group curates that data
- Design and implement a harmonized data model
- Submit data to the central data store using that model



Working Groups

Working Groups are small groups led by one or two individuals that are charged with the implementation of specific Alliance work products. The teams are comprised of individuals from member projects.

- AllianceMine [active]
- Architecture [active]
- Biological Function [maintenance mode]
- Data Quartermasters [active]
- DevOps [active]
- Expression [active]
- Gene Descriptions [maintenance mode]
- Genome Features [active]
- In-Paralogs [active]
- Interactions [maintenance mode]
- Literature Acquisition [active]
- Orthology [maintenance mode]
- Pathways [active]
- Phenotypes and Disease Models [active]
- Search Optimization [active]
- User Interface [active]
- User Support [active]
- Variants [active]

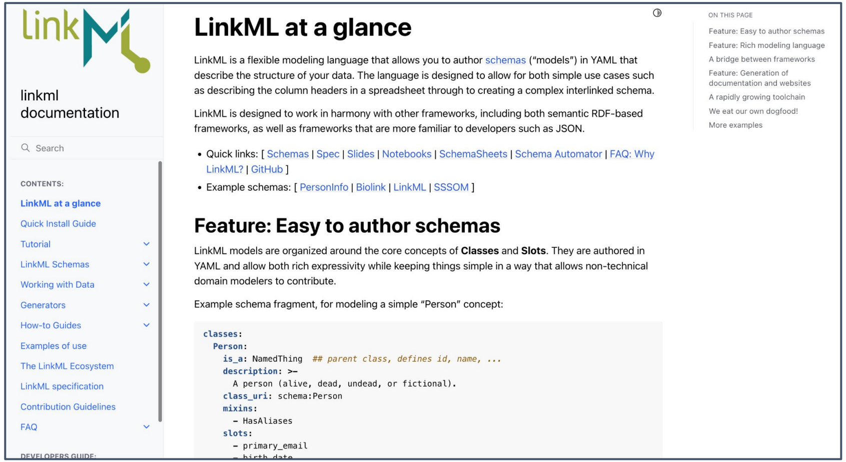
Active – Working Group is actively adding new functions or other deliverables

Maintenance mode – Working Group meets as needed to address data or function issues as needed; not working towards new functions at this time

https://www.alliancegenome.org/working-groups

Alliance data models: LinkML

- Linked Data Modeling Language (LinkML)
- Data submitted in JSON format by MOD data quarter masters (DQMs)
- Validation and consistency checks upon loading



linkml
documentation

Search

CONTENTS:

- **LinkML at a glance**
- Quick Install Guide
- Tutorial
- LinkML Schemas
- Working with Data
- Generators
- How-to Guides
- Examples of use
- The LinkML Ecosystem
- LinkML specification
- Contribution Guidelines
- FAQ

DEVELOPERS GUIDE:

LinkML at a glance

LinkML is a flexible modeling language that allows you to author schemas ("models") in YAML that describe the structure of your data. The language is designed to allow for both simple use cases such as describing the column headers in a spreadsheet through to creating a complex interlinked schema.

LinkML is designed to work in harmony with other frameworks, including both semantic RDF-based frameworks, as well as frameworks that are more familiar to developers such as JSON.

- Quick links: [Schemas | Spec | Slides | Notebooks | SchemaSheets | Schema Automator | FAQ: Why LinkML? | GitHub]
- Example schemas: [PersonInfo | Biolink | LinkML | SSSOM]

Feature: Easy to author schemas

LinkML models are organized around the core concepts of **Classes** and **Slots**. They are authored in YAML and allow both rich expressivity while keeping things simple in a way that allows non-technical domain modelers to contribute.

Example schema fragment, for modeling a simple "Person" concept:

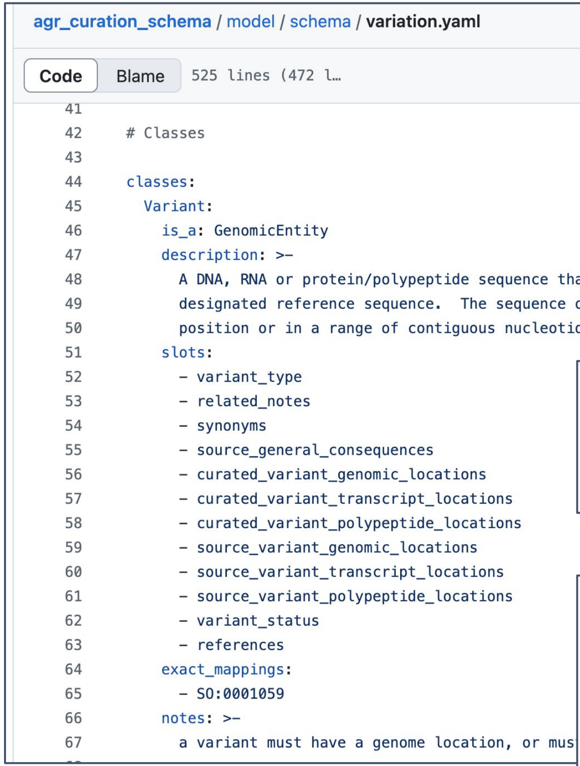
```
classes:
  Person:
    is_a: NamedThing  ## parent class, defines id, name, ...
    description: >-
      A person (alive, dead, undead, or fictional).
    class_uri: schema:Person
    mixins:
      - HasAliases
    slots:
      - primary_email
      - birth_date
```

ON THIS PAGE

- Feature: Easy to author schemas
- Feature: Rich modeling language
- A bridge between frameworks
- Feature: Generation of documentation and websites
- A rapidly growing toolchain
- We eat our own dogfood!
- More examples

https://linkml.io/linkml/intro/overview.html

Alliance LinkML model for variants



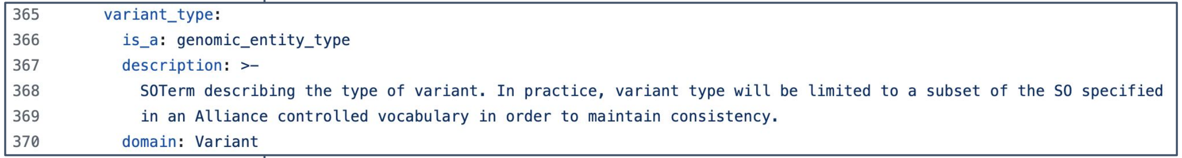
agr_curation_schema / model / schema / variation.yaml

525 lines (472 l…

```
# Classes

classes:
  Variant:
    is_a: GenomicEntity
    description: >-
      A DNA, RNA or protein/polypeptide sequence that differs relative to a
      designated reference sequence.  The sequence occurs at a single
      position or in a range of contiguous nucleotides or amino acids.
    slots:
      - variant_type
      - related_notes
      - synonyms
      - source_general_consequences
      - curated_variant_genomic_locations
      - curated_variant_transcript_locations
      - curated_variant_polypeptide_locations
      - source_variant_genomic_locations
      - source_variant_transcript_locations
      - source_variant_polypeptide_locations
      - variant_status
      - references
    exact_mappings:
      - SO:0001059
    notes: >-
      a variant must have a genome location, or mus
```

LinkML models describe attributes of an entity or type of data; use ontologies and CVs wherever possible

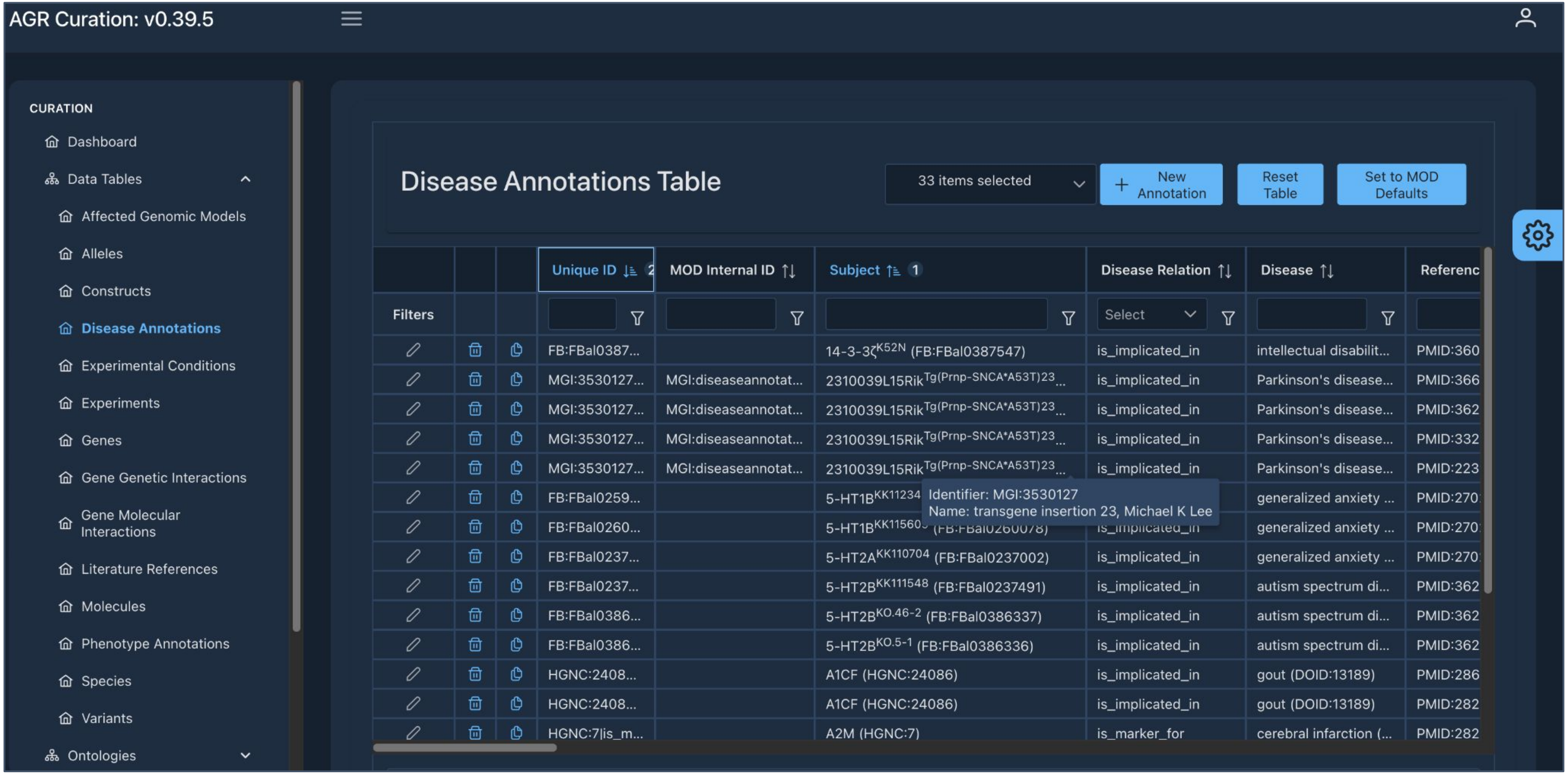


```
variant_type:
  is_a: genomic_entity_type
  description: >-
    SOTerm describing the type of variant. In practice, variant type will be limited to a subset of the SO specified
    in an Alliance controlled vocabulary in order to maintain consistency.
  domain: Variant
```

Allele/Variant Symbol	Allele Synonyms	Category	Variant	Variant Type	Molecular Consequence
ox299		allele with one associated variant	NC_003282.8:g.12781517T>A	point mutation	missense variant
m68		allele with one associated variant	NC_003282.8:g.12776551C>T	point mutation	stop gained
tg65		allele with one associated variant	NC_003282.8:g.12777572C>T	point mutation	stop gained

https://github.com/alliance-genome/agr_curation_schema

Alliance Data Curation Tools

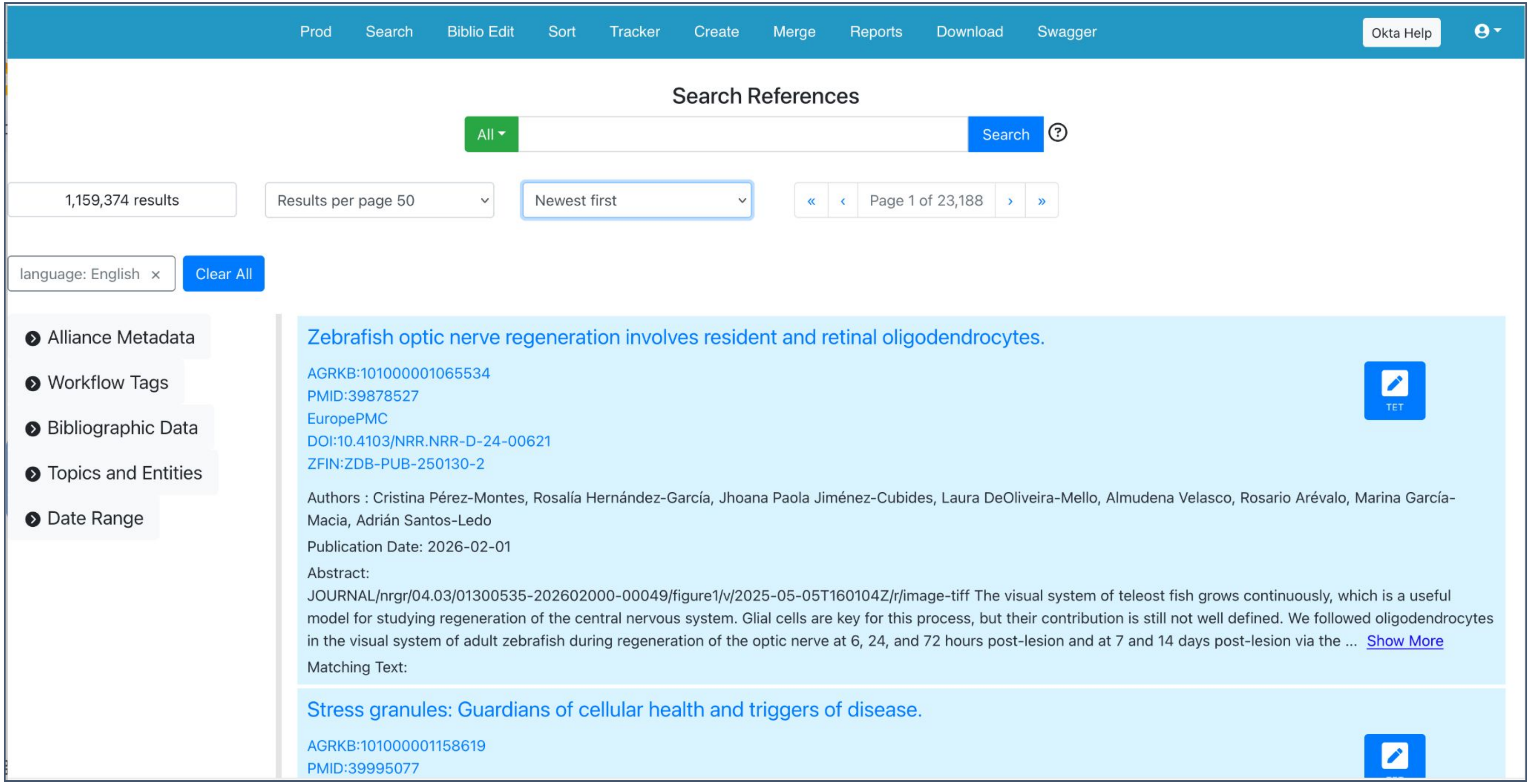


Disease Annotations Table

	Unique ID 2	MOD Internal ID	Subject 1	Disease Relation	Disease	Referenc
Filters				Select		
	FB:FBal0387...		14-3-3ζ^{K52N} (FB:FBal0387547)	is_implicated_in	intellectual disabilit...	PMID:360
	MGI:3530127...	MGI:diseaseannotat...	2310039L15Rik$^{Tg(Prnp-SNCA*A53T)23}$...	is_implicated_in	Parkinson's disease...	PMID:366
	MGI:3530127...	MGI:diseaseannotat...	2310039L15Rik$^{Tg(Prnp-SNCA*A53T)23}$...	is_implicated_in	Parkinson's disease...	PMID:362
	MGI:3530127...	MGI:diseaseannotat...	2310039L15Rik$^{Tg(Prnp-SNCA*A53T)23}$...	is_implicated_in	Parkinson's disease...	PMID:332
	MGI:3530127...	MGI:diseaseannotat...	2310039L15Rik$^{Tg(Prnp-SNCA*A53T)23}$...	is_implicated_in	Parkinson's disease...	PMID:223
	FB:FBal0259...		5-HT1B^{KK11234}		generalized anxiety ...	PMID:270
	FB:FBal0260...		5-HT1B^{KK11560} (FB:FBal0260078)	is_implicated_in	generalized anxiety ...	PMID:270
	FB:FBal0237...		5-HT2A^{KK110704} (FB:FBal0237002)	is_implicated_in	generalized anxiety ...	PMID:270
	FB:FBal0237...		5-HT2B^{KK111548} (FB:FBal0237491)	is_implicated_in	autism spectrum di...	PMID:362
	FB:FBal0386...		5-HT2B$^{KO.46-2}$ (FB:FBal0386337)	is_implicated_in	autism spectrum di...	PMID:362
	FB:FBal0386...		5-HT2B$^{KO.5-1}$ (FB:FBal0386336)	is_implicated_in	autism spectrum di...	PMID:362
	HGNC:2408...		A1CF (HGNC:24086)	is_implicated_in	gout (DOID:13189)	PMID:286
	HGNC:2408...		A1CF (HGNC:24086)	is_implicated_in	gout (DOID:13189)	PMID:282
	HGNC:7|is_m...		A2M (HGNC:7)	is_marker_for	cerebral infarction (...	PMID:282

Identifier: MGI:3530127
Name: transgene insertion 23, Michael K Lee

Alliance Bibliography Central: Scrum Blue Team



Search References

1,159,374 results

Results per page 50

Newest first

Page 1 of 23,188

language: English ×

Zebrafish optic nerve regeneration involves resident and retinal oligodendrocytes.

AGRKB:101000001065534
PMID:39878527
EuropePMC
DOI:10.4103/NRR.NRR-D-24-00621
ZFIN:ZDB-PUB-250130-2

Authors : Cristina Pérez-Montes, Rosalía Hernández-García, Jhoana Paola Jiménez-Cubides, Laura DeOliveira-Mello, Almudena Velasco, Rosario Arévalo, Marina García-Macia, Adrián Santos-Ledo

Publication Date: 2026-02-01

Abstract:
JOURNAL/nrgr/04.03/01300535-202602000-00049/figure1/v/2025-05-05T160104Z/r/image-tiff The visual system of teleost fish grows continuously, which is a useful model for studying regeneration of the central nervous system. Glial cells are key for this process, but their contribution is still not well defined. We followed oligodendrocytes in the visual system of adult zebrafish during regeneration of the optic nerve at 6, 24, and 72 hours post-lesion and at 7 and 14 days post-lesion via the ... Show More

Matching Text:

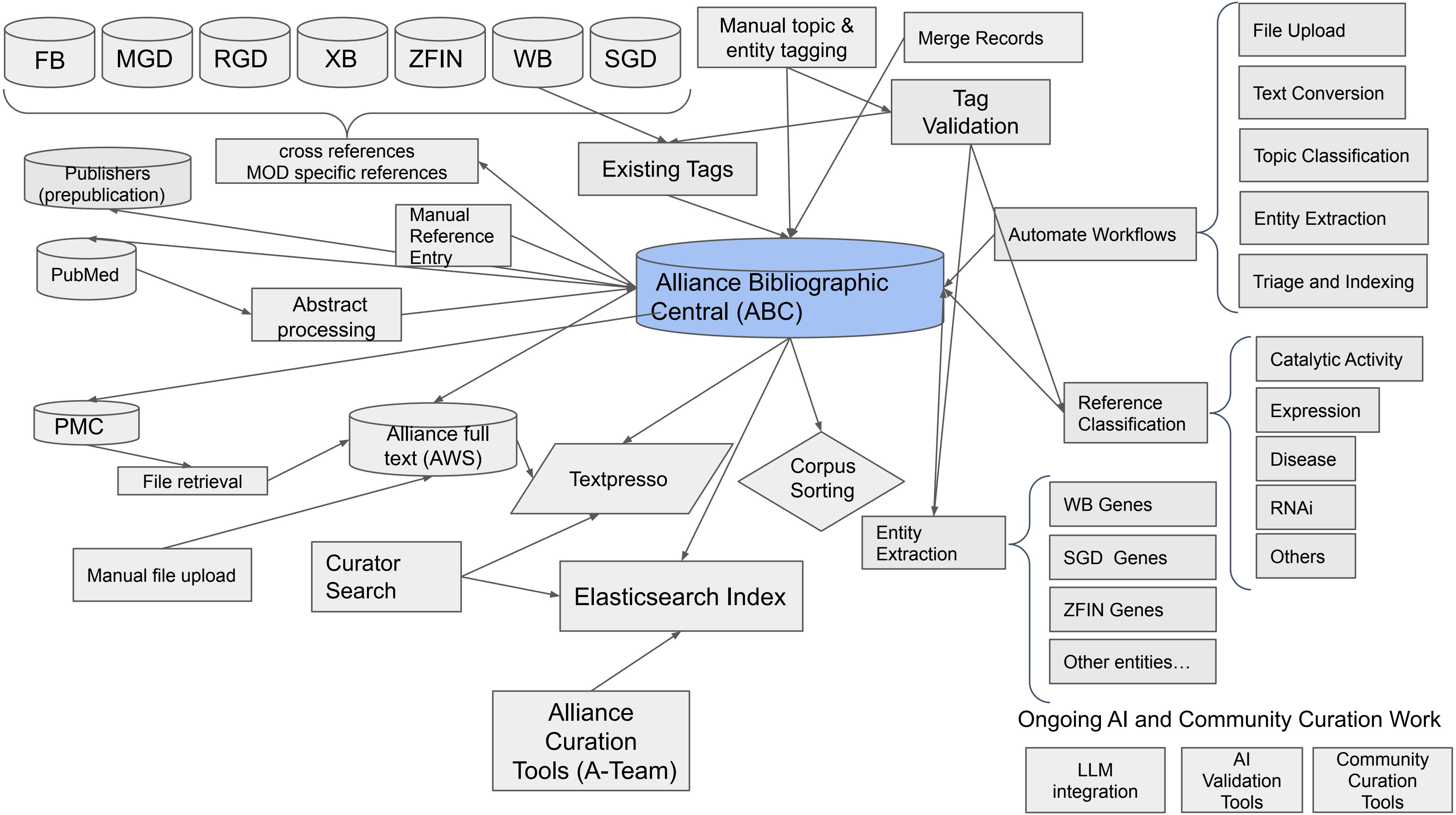
Stress granules: Guardians of cellular health and triggers of disease.

AGRKB:101000001158619
PMID:39995077

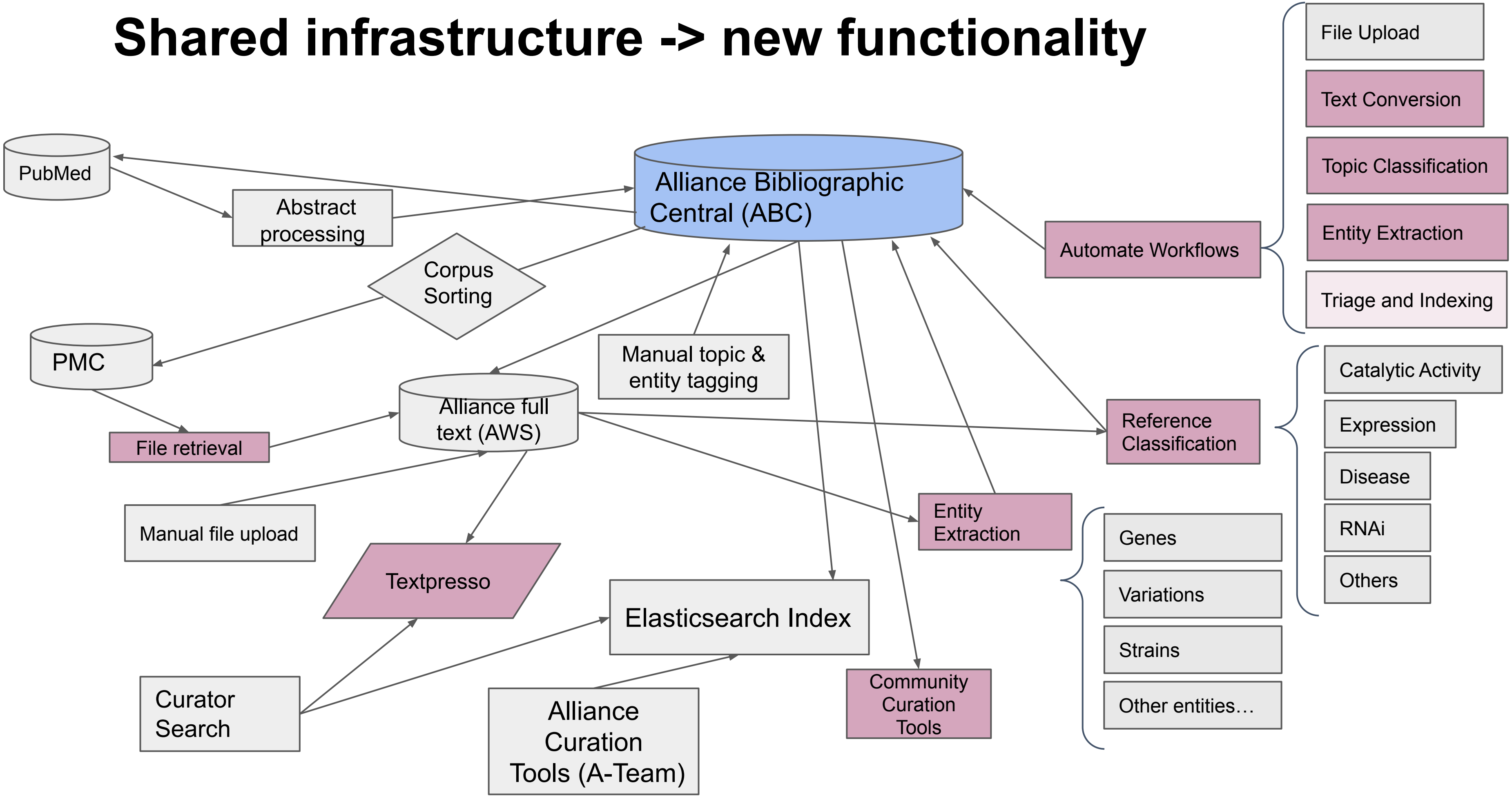
Alliance Bibliography Central: multiple working groups contribute to development

Requirements Gathering:

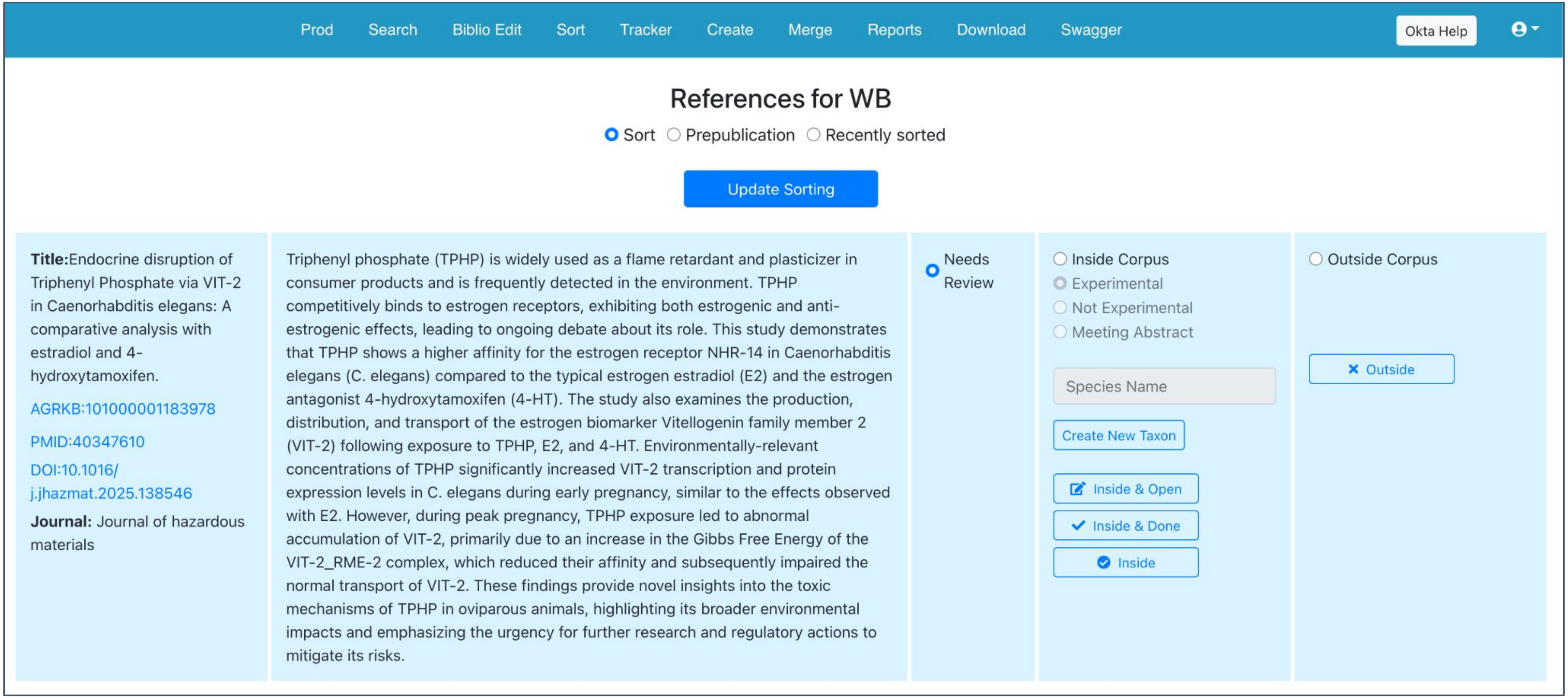
- Literature Acquisition Working Group (LAWG) - curators and developers representing MODs, meets every other week
- Metadata Modeling Group - subset of LAWG, meets weekly
- Search Working Group - subset of LAWG, meets every other week
- AI Working Group - curators and developers who are interested in LLM/ML/NLP, meets weekly


FB
MGD
RGD
XB
ZFIN
WB
SGD
cross references
MOD specific references
Manual topic & entity tagging
Merge Records
Tag Validation
Existing Tags
Publishers (prepublication)
Manual Reference Entry
PubMed
Abstract processing
Alliance Bibliographic Central (ABC)
Automate Workflows
File Upload
Text Conversion
Topic Classification
Entity Extraction
Triage and Indexing
PMC
Alliance full text (AWS)
File retrieval
Textpresso
Corpus Sorting
Reference Classification
Catalytic Activity
Expression
Disease
RNAi
Others
Entity Extraction
WB Genes
SGD Genes
ZFIN Genes
Other entities…
Manual file upload
Curator Search
Elasticsearch Index
Alliance Curation Tools (A-Team)
Ongoing AI and Community Curation Work
LLM integration
AI Validation Tools
Community Curation Tools

Shared infrastructure -> new functionality



Sorting inside corpus: automated PubMed searches and metadata addition



References for WB

● Sort ○ Prepublication ○ Recently sorted

Title: Endocrine disruption of Triphenyl Phosphate via VIT-2 in Caenorhabditis elegans: A comparative analysis with estradiol and 4-hydroxytamoxifen.

AGRKB:101000001183978

PMID:40347610

DOI:10.1016/j.jhazmat.2025.138546

Journal: Journal of hazardous materials

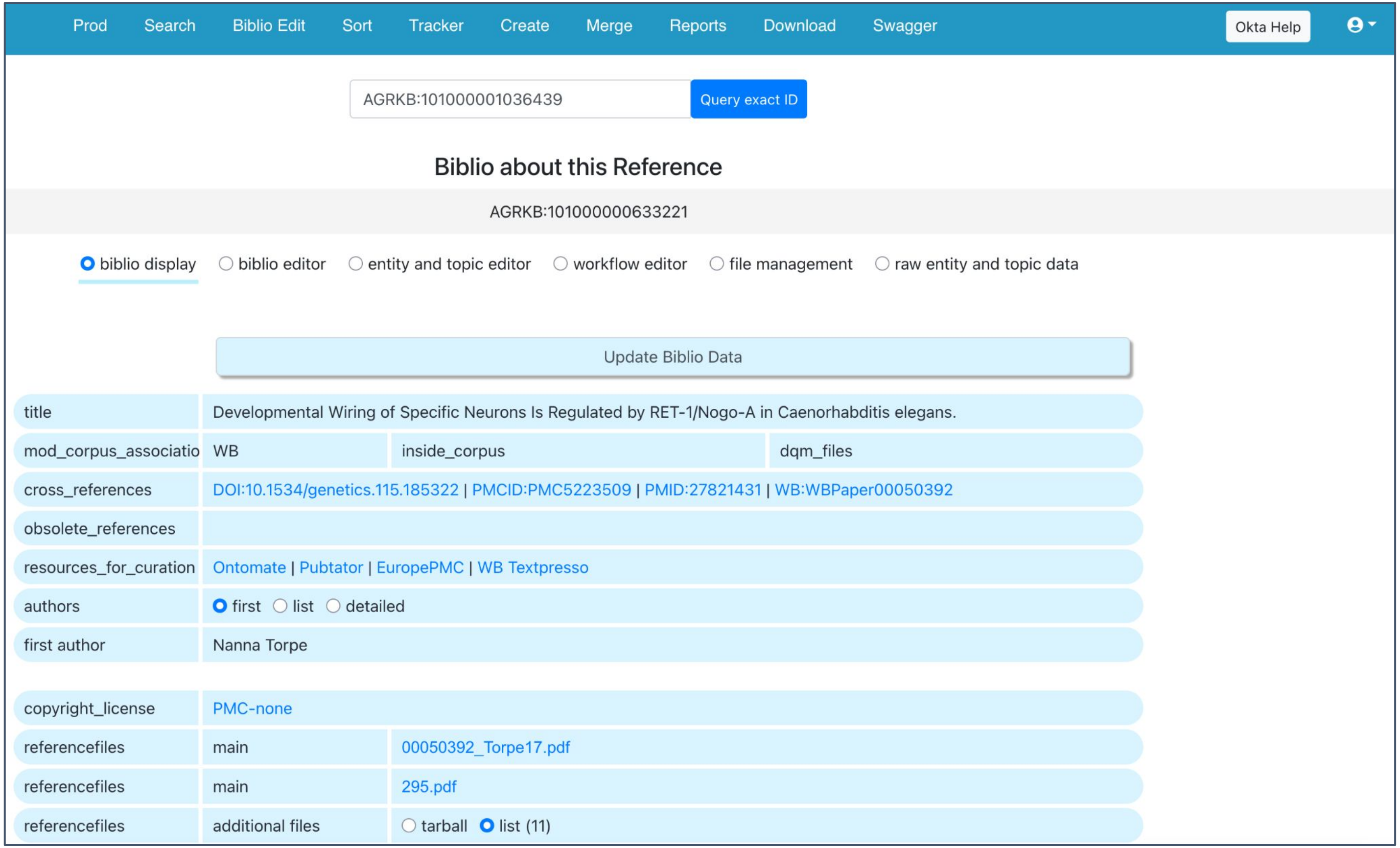
Triphenyl phosphate (TPHP) is widely used as a flame retardant and plasticizer in consumer products and is frequently detected in the environment. TPHP competitively binds to estrogen receptors, exhibiting both estrogenic and anti-estrogenic effects, leading to ongoing debate about its role. This study demonstrates that TPHP shows a higher affinity for the estrogen receptor NHR-14 in Caenorhabditis elegans (C. elegans) compared to the typical estrogen estradiol (E2) and the estrogen antagonist 4-hydroxytamoxifen (4-HT). The study also examines the production, distribution, and transport of the estrogen biomarker Vitellogenin family member 2 (VIT-2) following exposure to TPHP, E2, and 4-HT. Environmentally-relevant concentrations of TPHP significantly increased VIT-2 transcription and protein expression levels in C. elegans during early pregnancy, similar to the effects observed with E2. However, during peak pregnancy, TPHP exposure led to abnormal accumulation of VIT-2, primarily due to an increase in the Gibbs Free Energy of the VIT-2_RME-2 complex, which reduced their affinity and subsequently impaired the normal transport of VIT-2. These findings provide novel insights into the toxic mechanisms of TPHP in oviparous animals, highlighting its broader environmental impacts and emphasizing the urgency for further research and regulatory actions to mitigate its risks.

● Needs Review

○ Inside Corpus
○ Experimental
○ Not Experimental
○ Meeting Abstract

○ Outside Corpus

Each paper has links to external resources and access to full text when available



AGRKB:101000001036439

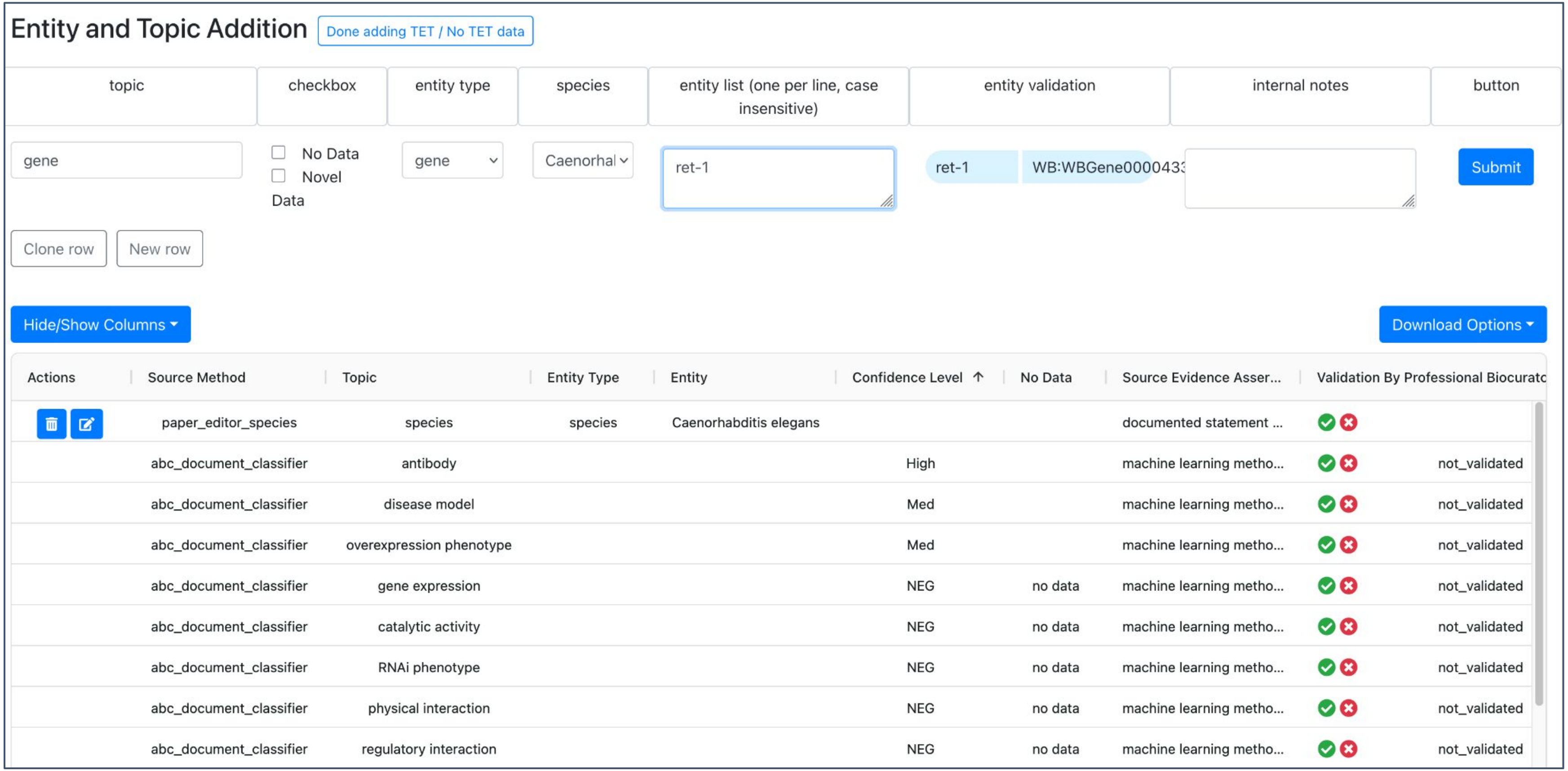
Biblio about this Reference

AGRKB:101000000633221

● biblio display ○ biblio editor ○ entity and topic editor ○ workflow editor ○ file management ○ raw entity and topic data

title	Developmental Wiring of Specific Neurons Is Regulated by RET-1/Nogo-A in Caenorhabditis elegans.		
mod_corpus_associatio	WB	inside_corpus	dqm_files
cross_references	DOI:10.1534/genetics.115.185322 | PMCID:PMC5223509 | PMID:27821431 | WB:WBPaper00050392		
obsolete_references			
resources_for_curation	Ontomate | Pubtator | EuropePMC | WB Textpresso		
authors	● first ○ list ○ detailed		
first author	Nanna Torpe		
copyright_license	PMC-none		
referencefiles	main	00050392_Torpe17.pdf	
referencefiles	main	295.pdf	
referencefiles	additional files	○ tarball ● list (11)	

Manual and automated flagging methods are collected in the TET (topic and entity table)

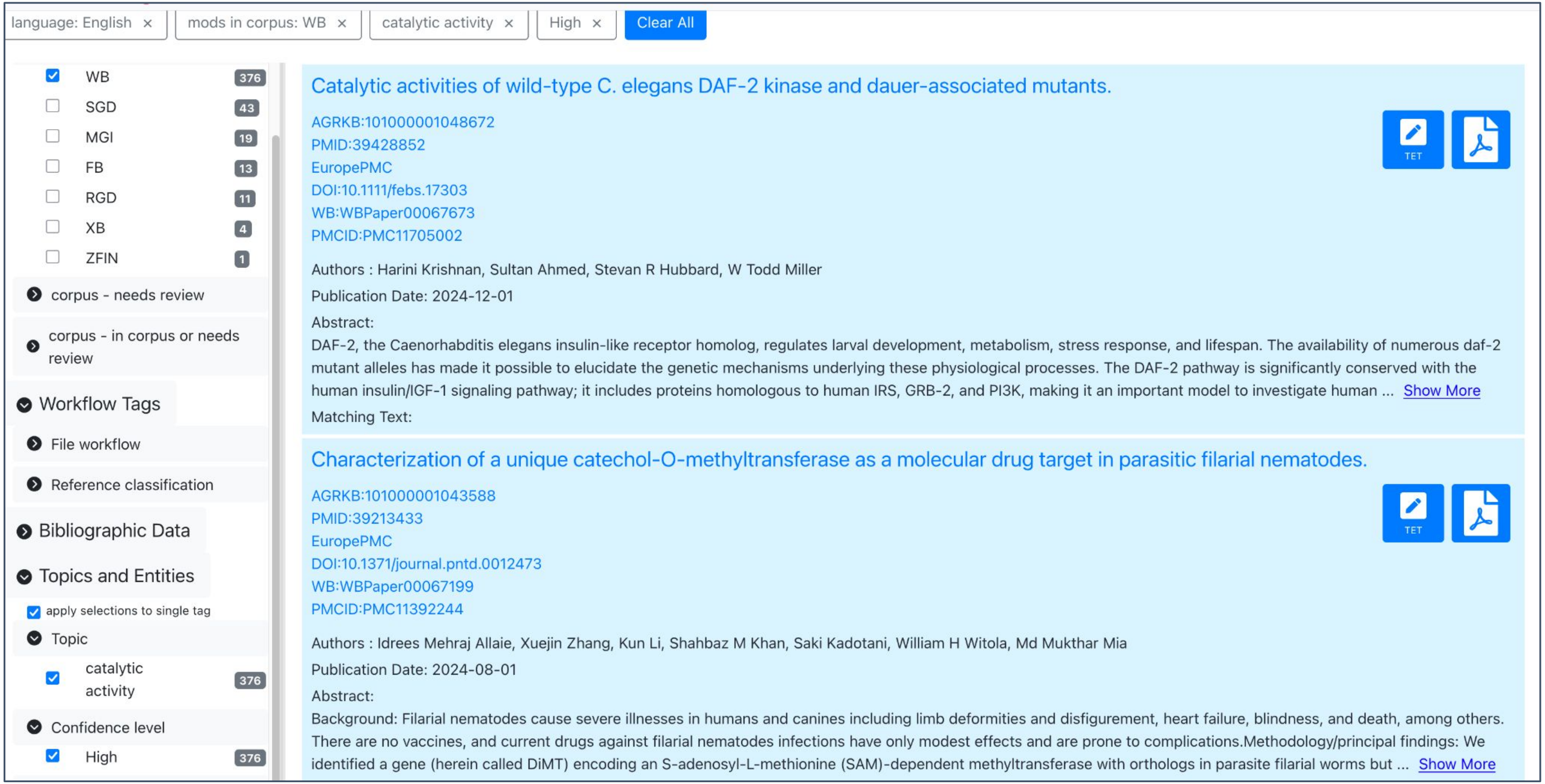


Entity and Topic Addition

topic	checkbox	entity type	species	entity list (one per line, case insensitive)	entity validation	internal notes	button
gene	☐ No Data ☐ Novel Data	gene	Caenorhal	ret-1	ret-1 WB:WBGene000043		Submit

Actions	Source Method	Topic	Entity Type	Entity	Confidence Level ↑	No Data	Source Evidence Asser...	Validation By Professional Biocurato
🗑 ✎	paper_editor_species	species	species	Caenorhabditis elegans			documented statement ...	✓ ✗
	abc_document_classifier	antibody			High		machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	disease model			Med		machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	overexpression phenotype			Med		machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	gene expression			NEG	no data	machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	catalytic activity			NEG	no data	machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	RNAi phenotype			NEG	no data	machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	physical interaction			NEG	no data	machine learning metho...	✓ ✗ not_validated
	abc_document_classifier	regulatory interaction			NEG	no data	machine learning metho...	✓ ✗ not_validated

Faceted search allows for curation prioritization



☑ WB 376
☐ SGD 43
☐ MGI 19
☐ FB 13
☐ RGD 11
☐ XB 4
☐ ZFIN 1

☑ apply selections to single tag

☑ catalytic activity 376

☑ High 376

Catalytic activities of wild-type C. elegans DAF-2 kinase and dauer-associated mutants.

AGRKB:101000001048672
PMID:39428852
EuropePMC
DOI:10.1111/febs.17303
WB:WBPaper00067673
PMCID:PMC11705002

Authors : Harini Krishnan, Sultan Ahmed, Stevan R Hubbard, W Todd Miller

Publication Date: 2024-12-01

Abstract:
DAF-2, the Caenorhabditis elegans insulin-like receptor homolog, regulates larval development, metabolism, stress response, and lifespan. The availability of numerous daf-2 mutant alleles has made it possible to elucidate the genetic mechanisms underlying these physiological processes. The DAF-2 pathway is significantly conserved with the human insulin/IGF-1 signaling pathway; it includes proteins homologous to human IRS, GRB-2, and PI3K, making it an important model to investigate human ... Show More

Matching Text:

Characterization of a unique catechol-O-methyltransferase as a molecular drug target in parasitic filarial nematodes.

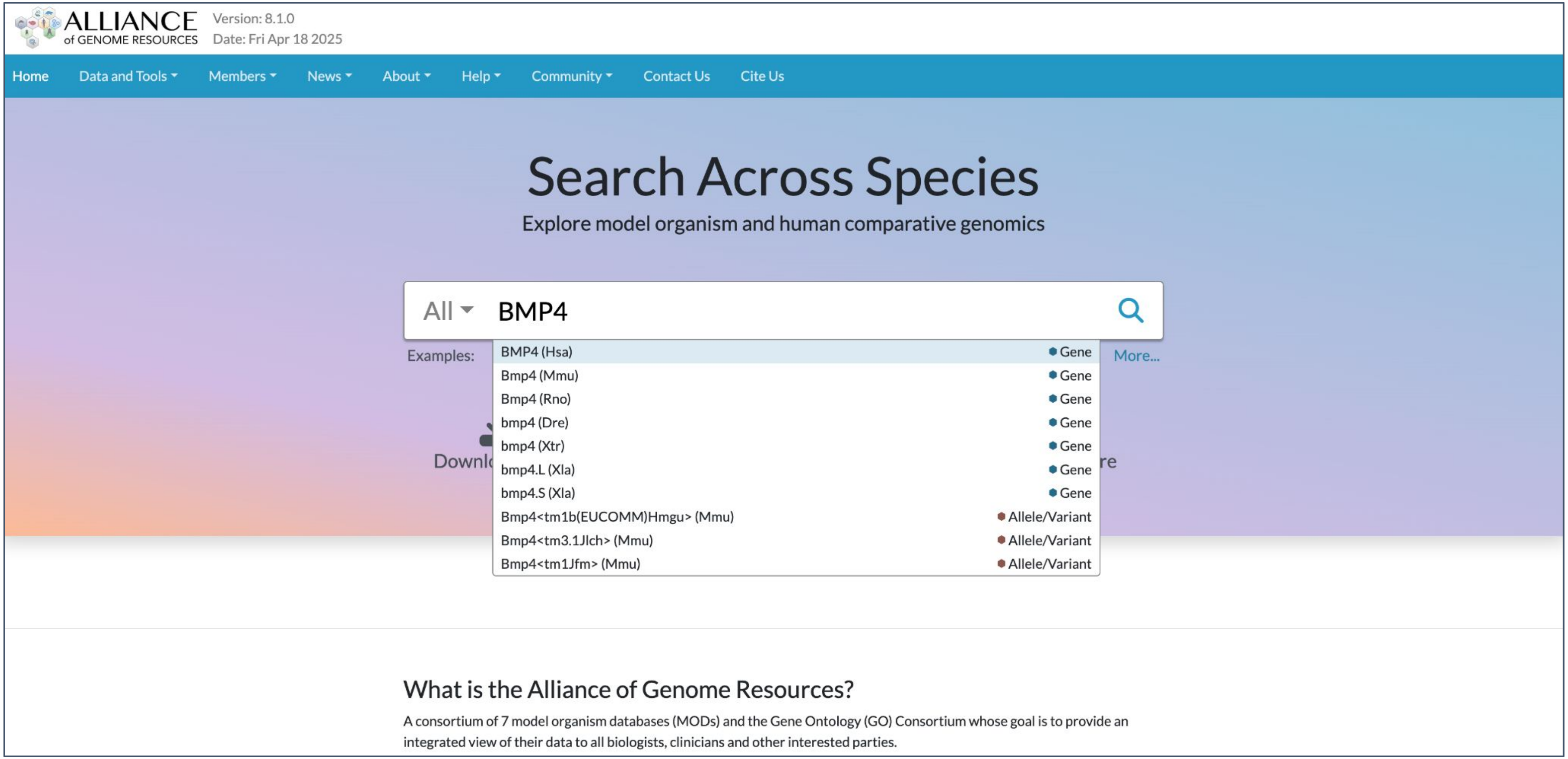
AGRKB:101000001043588
PMID:39213433
EuropePMC
DOI:10.1371/journal.pntd.0012473
WB:WBPaper00067199
PMCID:PMC11392244

Authors : Idrees Mehraj Allaie, Xuejin Zhang, Kun Li, Shahbaz M Khan, Saki Kadotani, William H Witola, Md Mukthar Mia

Publication Date: 2024-08-01

Abstract:
Background: Filarial nematodes cause severe illnesses in humans and canines including limb deformities and disfigurement, heart failure, blindness, and death, among others. There are no vaccines, and current drugs against filarial nematodes infections have only modest effects and are prone to complications.Methodology/principal findings: We identified a gene (herein called DiMT) encoding an S-adenosyl-L-methionine (SAM)-dependent methyltransferase with orthologs in parasite filarial worms but ... Show More

Putting it all back together



Search Across Species

Explore model organism and human comparative genomics

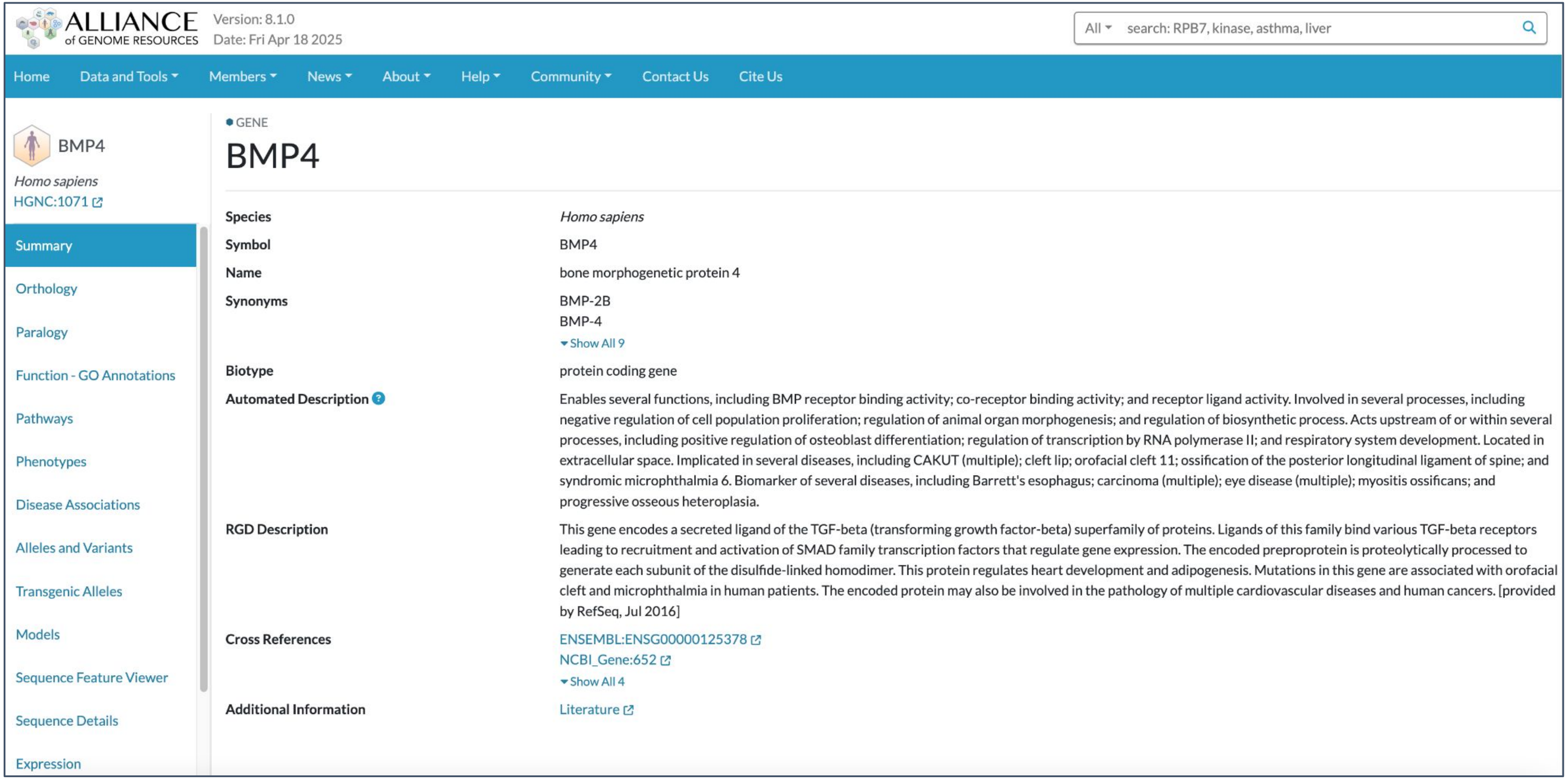
BMP4

BMP4 (Hsa) ● Gene
Bmp4 (Mmu) ● Gene
Bmp4 (Rno) ● Gene
bmp4 (Dre) ● Gene
bmp4 (Xtr) ● Gene
bmp4.L (Xla) ● Gene
bmp4.S (Xla) ● Gene
Bmp4<tm1b(EUCOMM)Hmgu> (Mmu) ● Allele/Variant
Bmp4<tm3.1Jlch> (Mmu) ● Allele/Variant
Bmp4<tm1Jfm> (Mmu) ● Allele/Variant

What is the Alliance of Genome Resources?

A consortium of 7 model organism databases (MODs) and the Gene Ontology (GO) Consortium whose goal is to provide an integrated view of their data to all biologists, clinicians and other interested parties.

Gene pages are the information hub of the Alliance



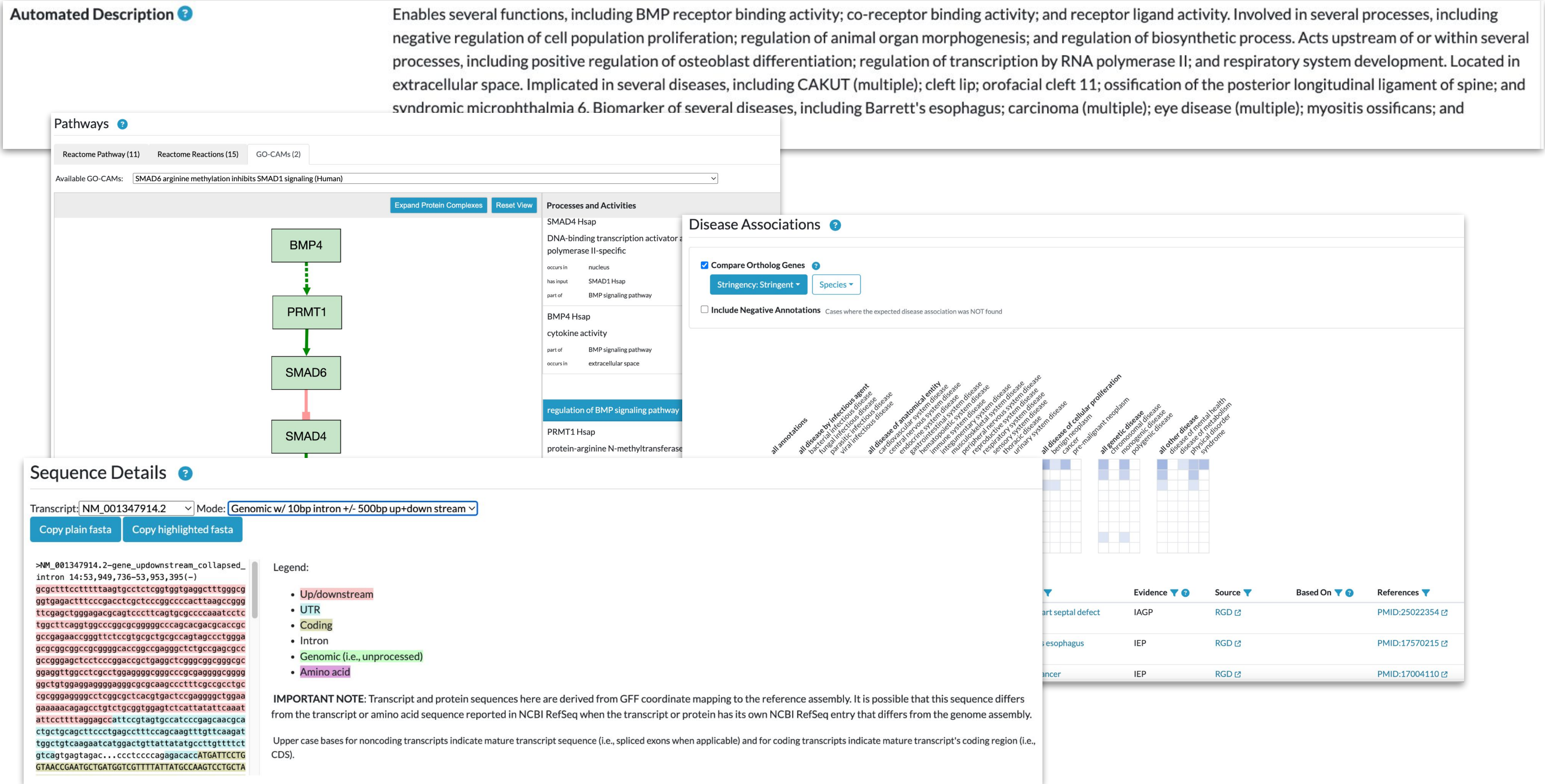
Version: 8.1.0
Date: Fri Apr 18 2025

Homo sapiens
HGNC:1071

GENE

BMP4

Species	*Homo sapiens*
Symbol	BMP4
Name	bone morphogenetic protein 4
Synonyms	BMP-2B BMP-4 Show All 9
Biotype	protein coding gene
Automated Description	Enables several functions, including BMP receptor binding activity; co-receptor binding activity; and receptor ligand activity. Involved in several processes, including negative regulation of cell population proliferation; regulation of animal organ morphogenesis; and regulation of biosynthetic process. Acts upstream of or within several processes, including positive regulation of osteoblast differentiation; regulation of transcription by RNA polymerase II; and respiratory system development. Located in extracellular space. Implicated in several diseases, including CAKUT (multiple); cleft lip; orofacial cleft 11; ossification of the posterior longitudinal ligament of spine; and syndromic microphthalmia 6. Biomarker of several diseases, including Barrett's esophagus; carcinoma (multiple); eye disease (multiple); myositis ossificans; and progressive osseous heteroplasia.
RGD Description	This gene encodes a secreted ligand of the TGF-beta (transforming growth factor-beta) superfamily of proteins. Ligands of this family bind various TGF-beta receptors leading to recruitment and activation of SMAD family transcription factors that regulate gene expression. The encoded preproprotein is proteolytically processed to generate each subunit of the disulfide-linked homodimer. This protein regulates heart development and adipogenesis. Mutations in this gene are associated with orofacial cleft and microphthalmia in human patients. The encoded protein may also be involved in the pathology of multiple cardiovascular diseases and human cancers. [provided by RefSeq, Jul 2016]
Cross References	ENSEMBL:ENSG00000125378 NCBI_Gene:652 Show All 4
Additional Information	Literature



Automated Description

Enables several functions, including BMP receptor binding activity; co-receptor binding activity; and receptor ligand activity. Involved in several processes, including negative regulation of cell population proliferation; regulation of animal organ morphogenesis; and regulation of biosynthetic process. Acts upstream of or within several processes, including positive regulation of osteoblast differentiation; regulation of transcription by RNA polymerase II; and respiratory system development. Located in extracellular space. Implicated in several diseases, including CAKUT (multiple); cleft lip; orofacial cleft 11; ossification of the posterior longitudinal ligament of spine; and syndromic microphthalmia 6. Biomarker of several diseases, including Barrett's esophagus; carcinoma (multiple); eye disease (multiple); myositis ossificans; and

Pathways

Available GO-CAMs: SMAD6 arginine methylation inhibits SMAD1 signaling (Human)

Processes and Activities

SMAD4 Hsap

DNA-binding transcription activator
polymerase II-specific

occurs in — nucleus
has input — SMAD1 Hsap
part of — BMP signaling pathway

BMP4 Hsap

cytokine activity

part of — BMP signaling pathway
occurs in — extracellular space

regulation of BMP signaling pathway

PRMT1 Hsap

protein-arginine N-methyltransferase

Disease Associations

☑ **Compare Ortholog Genes**

☐ **Include Negative Annotations** Cases where the expected disease association was NOT found

	Evidence	Source	Based On	References
art septal defect	IAGP	RGD		PMID:25022354
s esophagus	IEP	RGD		PMID:17570215
ancer	IEP	RGD		PMID:17004110

Sequence Details

Transcript: NM_001347914.2 Mode: Genomic w/ 10bp intron +/- 500bp up+down stream

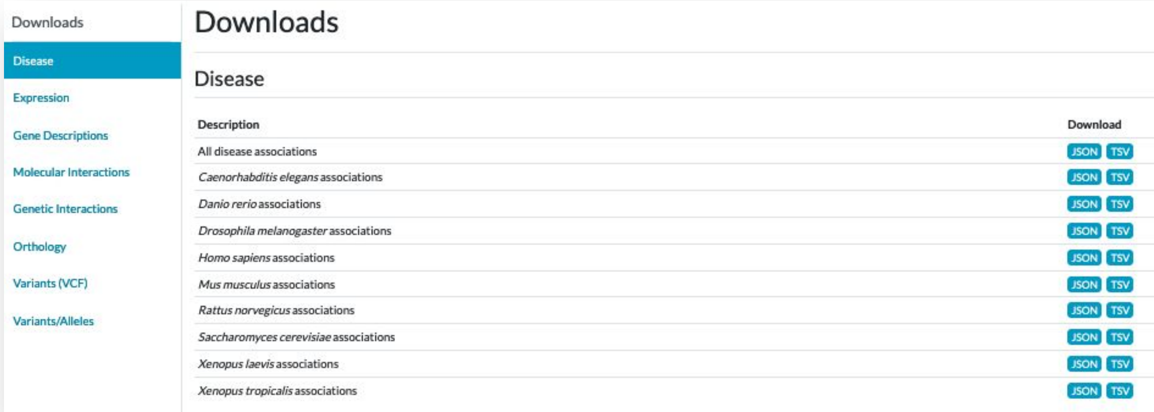
```
>NM_001347914.2-gene_updownstream_collapsed_
intron 14:53,949,736-53,953,395(-)
gcgctttccttttttaagtgcctctcggtggtgaggctttgggcg
ggtgagactttcccgacctcgctcccggccccacttaagccggg
ttcgagctgggagacgcagtcccttcagtgcgccccaaatcctc
tggcttcaggtggcccggcgcgggggcccagcacgacgcaccgc
gccgagaaccgggttctccgtgcgctgcgccagtagccctggga
gcgcggcggccgcggggcaccggccgagggctctgccgagcgcc
gccgggagctcctcccggaccgctgaggctcgggcggcgggcgc
ggaggttggcctcgcctggagggggcgggcccgcgagggggcgggg
ggctgtggaggaggggagggcgcgcaagccctttcgccgcctgc
cgcgggaggggcctcggcgctcacgtgactccgaggggctggaa
gaaaaacagagcctgtctgcggtggagtctcattatattcaaat
attccttttaggagccattccgtagtgccatcccgagcaacgca
ctgctgcagcttccctgagcctttccagcaagtttgttcaagat
tggctgtcaagaatcatggactgttattatatgccttgttttct
gtcagtgagtagac...ccctccccagagacaccATGATTCCTG
GTAACCGAATGCTGATGGTCGTTTTATTATGCCAAGTCCTGCTA
```

Legend:

- Up/downstream
- UTR
- Coding
- Intron
- Genomic (i.e., unprocessed)
- Amino acid

IMPORTANT NOTE: Transcript and protein sequences here are derived from GFF coordinate mapping to the reference assembly. It is possible that this sequence differs from the transcript or amino acid sequence reported in NCBI RefSeq when the transcript or protein has its own NCBI RefSeq entry that differs from the genome assembly.

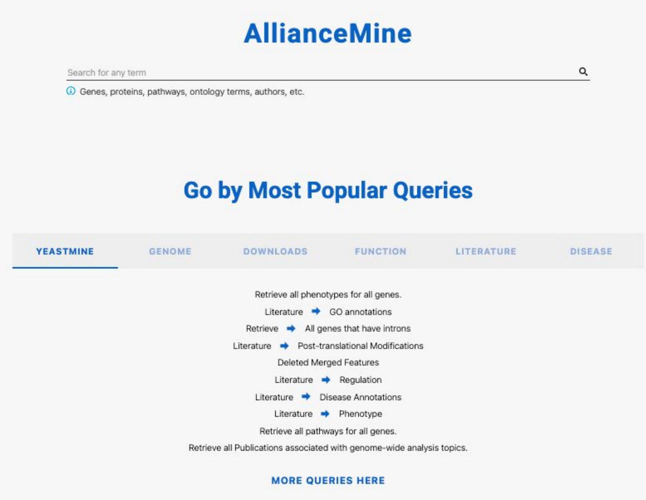
Upper case bases for noncoding transcripts indicate mature transcript sequence (i.e., spliced exons when applicable) and for coding transcripts indicate mature transcript's coding region (i.e., CDS).



Downloads

Disease

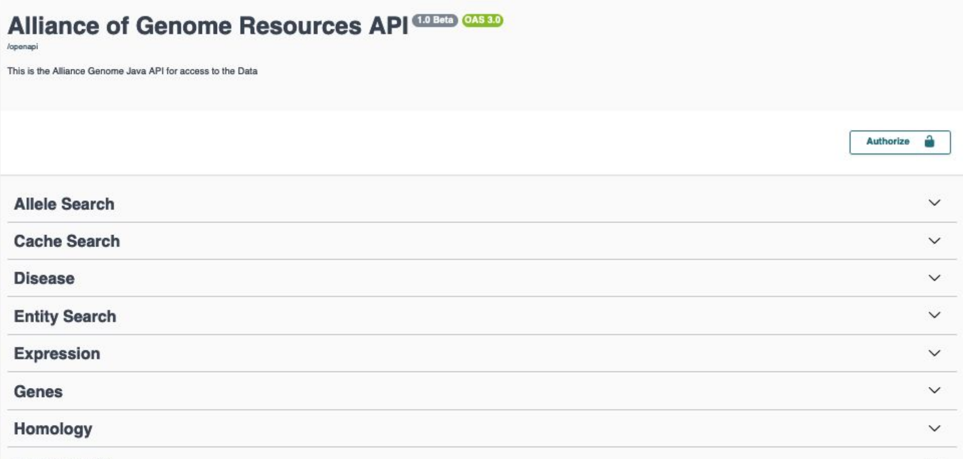
Description	Download
All disease associations	JSON TSV
Caenorhabditis elegans associations	JSON TSV
Danio rerio associations	JSON TSV
Drosophila melanogaster associations	JSON TSV
Homo sapiens associations	JSON TSV
Mus musculus associations	JSON TSV
Rattus norvegicus associations	JSON TSV
Saccharomyces cerevisiae associations	JSON TSV
Xenopus laevis associations	JSON TSV
Xenopus tropicalis associations	JSON TSV



AllianceMine

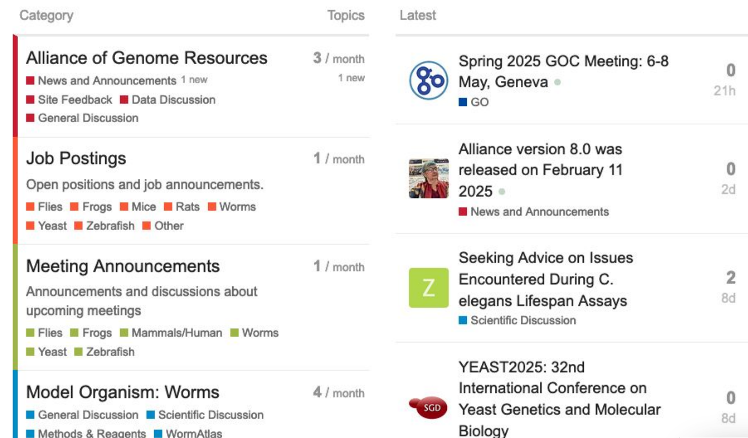
Go by Most Popular Queries

Retrieve all phenotypes for all genes.
Literature → GO annotations
Retrieve → All genes that have introns
Literature → Post-translational Modifications
Deleted Merged Features
Literature → Regulation
Literature → Disease Annotations
Literature → Phenotype
Retrieve all pathways for all genes.
Retrieve all Publications associated with genome-wide analysis topics.



Alliance of Genome Resources API

This is the Alliance Genome Java API for access to the Data



Category Topics Latest

Alliance of Genome Resources 3 / month 1 new
News and Announcements 1 new
Site Feedback Data Discussion
General Discussion

Job Postings 1 / month
Open positions and job announcements.
Flies Frogs Mice Rats Worms
Yeast Zebrafish Other

Meeting Announcements 1 / month
Announcements and discussions about upcoming meetings
Flies Frogs Mammals/Human Worms
Yeast Zebrafish

Model Organism: Worms 4 / month
General Discussion Scientific Discussion
Methods & Reagents WormAtlas
WormBase WormBook Facts & Figures

Model Organism: Yeast 3 / month
Information about all things yeast-related, populated and monitored by your friendly neighborhood SGD biocurators

Spring 2025 GOC Meeting: 6-8 May, Geneva
GO
0 21h

Alliance version 8.0 was released on February 11 2025
News and Announcements
0 2d

Seeking Advice on Issues Encountered During C. elegans Lifespan Assays
Scientific Discussion
2 8d

YEAST2025: 32nd International Conference on Yeast Genetics and Molecular Biology
Conferences
0 8d

The Jarriault Lab is looking for a postdoc on cellular reprogramming
Worms

Alliance of Genome Resources

Follows you @alliancegenome.bsky.social
1.6K followers 102 following 141 posts

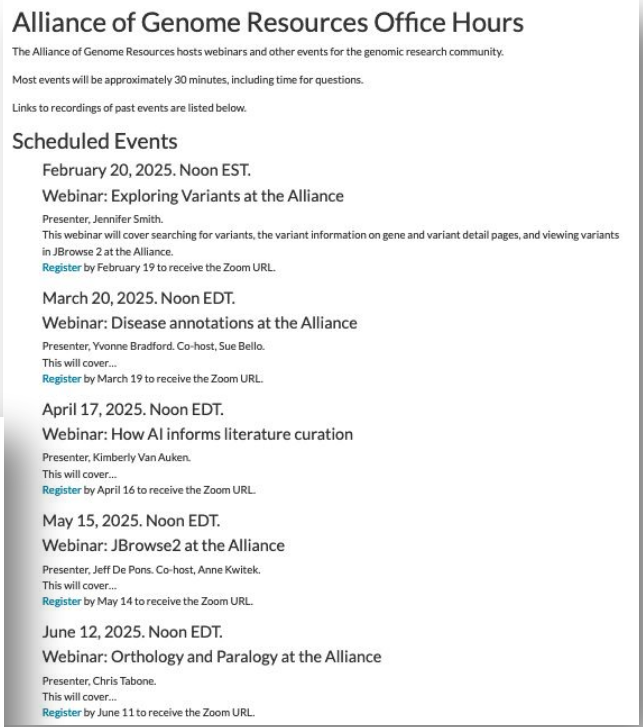
The Alliance of Genome Resources supports biological science via data integration, standardization of data models and interfaces, and unified outreach.

Find us at www.alliancegenome.org.

Followed by FlyBase and Rat Genome Database

Alliance of Genome Resources @alliancegenome.bsky.social · 6m
The next Alliance webinar is scheduled for next Thursday, February 20, at noon ET. This month's webinar is on Exploring Variants at the Alliance, and will be presented by Jennifer Smith. Register by Wednesday Feb 19 (before midnight ET please) to receive the Zoom link forms.gle/GzMnmwK23SzP...

Alliance Of Genome Resources Event Registration



Alliance of Genome Resources Office Hours

The Alliance of Genome Resources hosts webinars and other events for the genomic research community.

Most events will be approximately 30 minutes, including time for questions.

Links to recordings of past events are listed below.

Scheduled Events

February 20, 2025. Noon EST.

Webinar: Exploring Variants at the Alliance

Presenter, Jennifer Smith.
This webinar will cover searching for variants, the variant information on gene and variant detail pages, and viewing variants in JBrowse 2 at the Alliance.
Register by February 19 to receive the Zoom URL.

March 20, 2025. Noon EDT.

Webinar: Disease annotations at the Alliance

Presenter, Yvonne Bradford. Co-host, Sue Bello.
This will cover...
Register by March 19 to receive the Zoom URL.

April 17, 2025. Noon EDT.

Webinar: How AI informs literature curation

Presenter, Kimberly Van Auken.
This will cover...
Register by April 16 to receive the Zoom URL.

May 15, 2025. Noon EDT.

Webinar: JBrowse2 at the Alliance

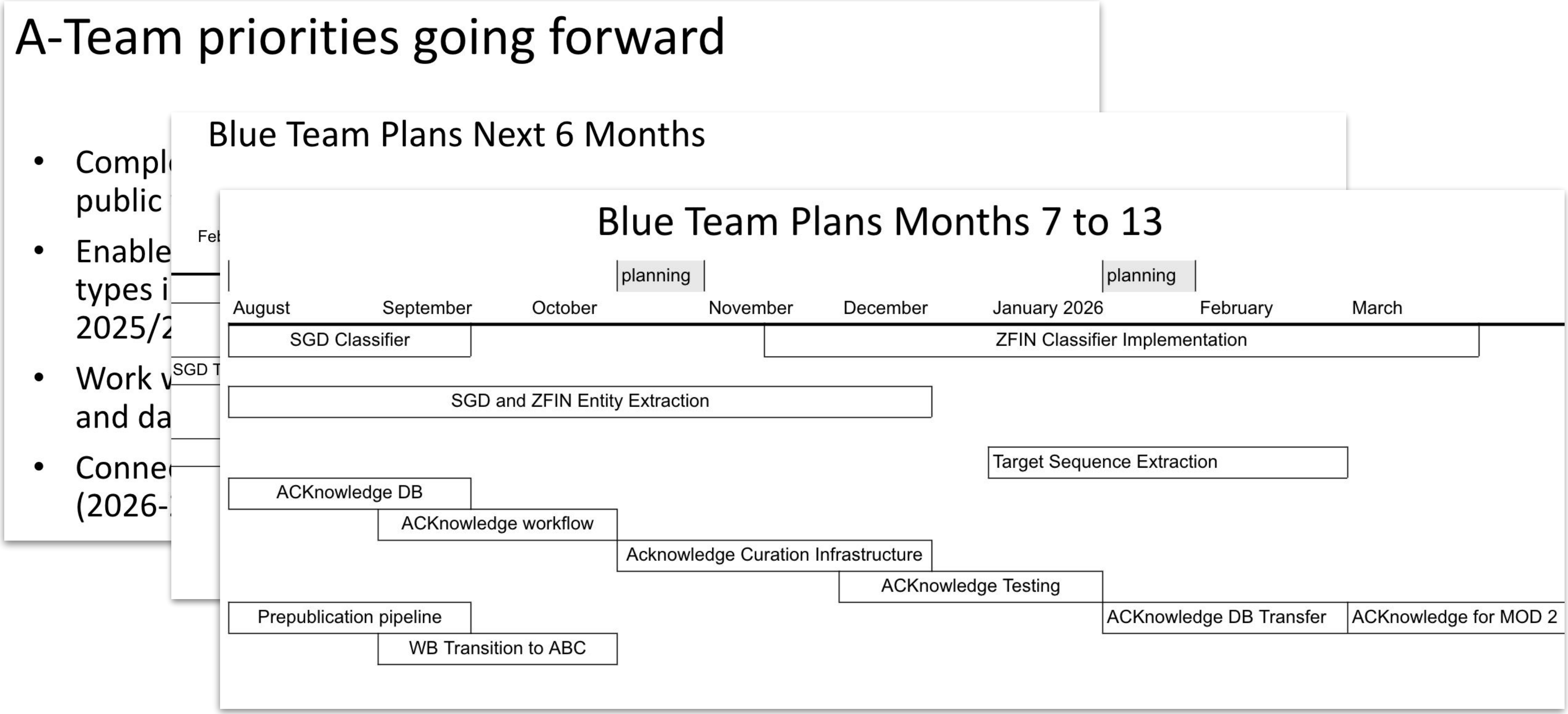
Presenter, Jeff De Pons. Co-host, Anne Kwitek.
This will cover...
Register by May 14 to receive the Zoom URL.

June 12, 2025. Noon EDT.

Webinar: Orthology and Paralogy at the Alliance

Presenter, Chris Tabone.
This will cover...
Register by June 11 to receive the Zoom URL.

Where are we on the roadmap?



A-Team priorities going forward

- Compl
  public
- Enable
  types i
  2025/2
- Work v
  and da
- Conne
  (2026-

Blue Team Plans Next 6 Months

Feb

SGD T

Blue Team Plans Months 7 to 13

planning

planning

August

September

October

November

December

January 2026

February

March

SGD Classifier

ZFIN Classifier Implementation

SGD and ZFIN Entity Extraction

Target Sequence Extraction

ACKnowledge DB

ACKnowledge workflow

Acknowledge Curation Infrastructure

ACKnowledge Testing

ACKnowledge DB Transfer

ACKnowledge for MOD 2

Prepublication pipeline

WB Transition to ABC

A-Team members

- Carol Bult (PI)
- Chris Grove (Product Owner)
- Jeff De Pons (Scrum Master)
- Olin Blodgett
- Mark Quinton-Tulloch
- Andrés Becerra Sandoval
- Adam Gibson
- Christian Pich
- Varun Reddy Gollapally

Subject Matter Experts (SMEs)

- Sue Bello
- Yvonne Bradford

Blue Team members

- Paul Sternberg (PI)
- Ceri Van Slyke (Product Owner)
- Valerio Arnaboldi
- Juancarlos Chan
- Paul Hale
- Ian Longden
- Shuai Weng
- Pinglei Zhou

Subject Matter Experts (SMEs)

- Kimberly Van Auken
- Gillian Millburn

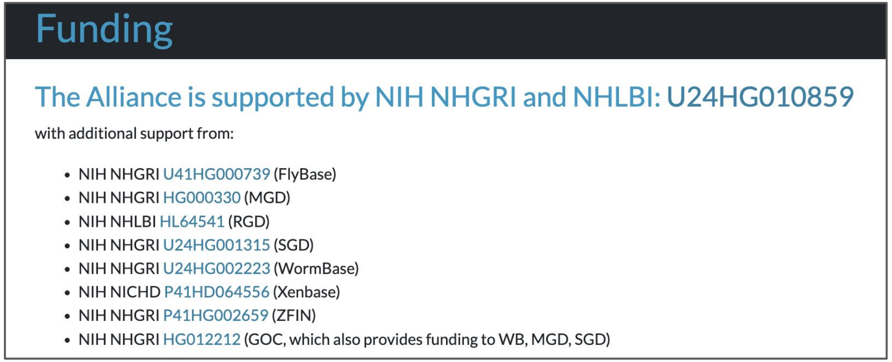
Alliance Software Specialist Team

David Shaw (now retired)	User Support
Sian Gramates	User Support / Social Outreach
Sierra Moxon	LinkML/Modelling
Dustin Ebert	Pathway Viewer
Hans-Michael Mueller	Textpresso
Seth Carbon	GO Software Coordinator
Paulo Nuin	AllianceMine, BLAST
Stuart Miyasato	AWS, DevOps
Paul Hale	JBrowse
Manuel Luypaert	Proteins, Annotations, and Variants Inspector
Tremayne Mushayahama	Pathway Viewer, GO Software
Adam Wright	BLAST, DevOps
Christopher Tabone	Coordinator, DevOps, DIOPT

Thanks to all!

The Alliance of Genome Resources Consortium:

Suzanne A Aleksander, Anna V Anagnostopoulos, Giulia Antonazzo, Valerio Arnaboldi, Helen Attrill, Andrés Becerra, Susan M Bello, Olin Blodgett, Yvonne M Bradford, Carol J Bult, Scott Cain, Brian R Calvi, Seth Carbon, Juancarlos Chan, Wen J Chen, J Michael Cherry, Jaehyoung Cho, Madeline A Crosby, Jeffrey L De Pons, Peter D'Eustachio, Stavros Diamantakis, Mary E Dolan, Gilberto dos Santos, Sarah Dyer, Dustin Ebert, Stacia R Engel, David Fashena, Malcolm Fisher, Saoirse Foley, Adam C Gibson, Varun R Gollapally, L Sian Gramates, Christian A Grove, Paul Hale, Todd Harris, G Thomas Hayman, Yanhui Hu, Christina James-Zorn, Kamran Karimi, Kalpana Karra, Ranjana Kishore, Anne E Kwitek, Stanley J F Laulederkind, Raymond Lee, Ian Longden, Manuel Luypaert, Nicholas Markarian, Steven J Marygold, Beverley Matthews, Monica S McAndrews, Gillian Millburn, Stuart Miyasato, Howie Motenko, Sierra Moxon, Hans-Michael Muller, Christopher J Mungall, Anushya Muruganujan, Tremayne Mushayahama, Robert S Nash, Paulo Nuin, Holly Paddock, Troy Pells, Norbert Perrimon, Christian Pich, Mark Quinton-Tulloch, Daniela Raciti, Sridhar Ramachandran, Joel E Richardson, Susan Russo Gelbart, Leyla Ruzicka, Gary Schindelman, David R Shaw, Gavin Sherlock, Ajay Shrivatsav, Amy Singer, Constance M Smith, Cynthia L Smith, Jennifer R Smith, Lincoln Stein, Paul W Sternberg, Christopher J Tabone, Paul D Thomas, Ketaki Thorat, Jyothi Thota, Monika Tomczuk, Vitor Trovisco, Marek A Tutaj, Jose-Maria Urbano, Kimberly Van Auken, Ceri E Van Slyke, Peter D Vize, Qinghua Wang, Shuai Weng, Monte Westerfield, Laurens G Wilming, Edith D Wong, Adam Wright, Karen Yook, Pinglei Zhou, Aaron Zorn, and Mark Zytkovicz



Funding

The Alliance is supported by NIH NHGRI and NHLBI: U24HG010859

with additional support from:

- NIH NHGRI U41HG000739 (FlyBase)
- NIH NHGRI HG000330 (MGD)
- NIH NHLBI HL64541 (RGD)
- NIH NHGRI U24HG001315 (SGD)
- NIH NHGRI U24HG002223 (WormBase)
- NIH NICHD P41HD064556 (Xenbase)
- NIH NHGRI P41HG002659 (ZFIN)
- NIH NHGRI HG012212 (GOC, which also provides funding to WB, MGD, SGD)