



Search for...

[expand all](#) | [collapse all](#)

Filter the searches below...

Genes

Organisms

Popset Isolate Sequences

Genomic Sequences

Genomic Segments

ESTs



News »

[VectorBase 57 Released](#)

THU APR 21 2022

[TrichDB 57 Released](#)

THU APR 21 2022

[TriTrypDB 57 Released](#)

THU APR 21 2022

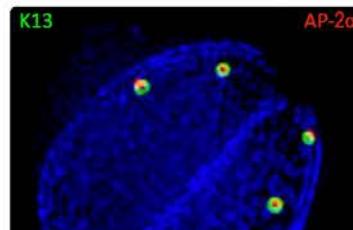
[ToxoDR 57 Released](#)[See all news](#)

Tweets by @veupathdb

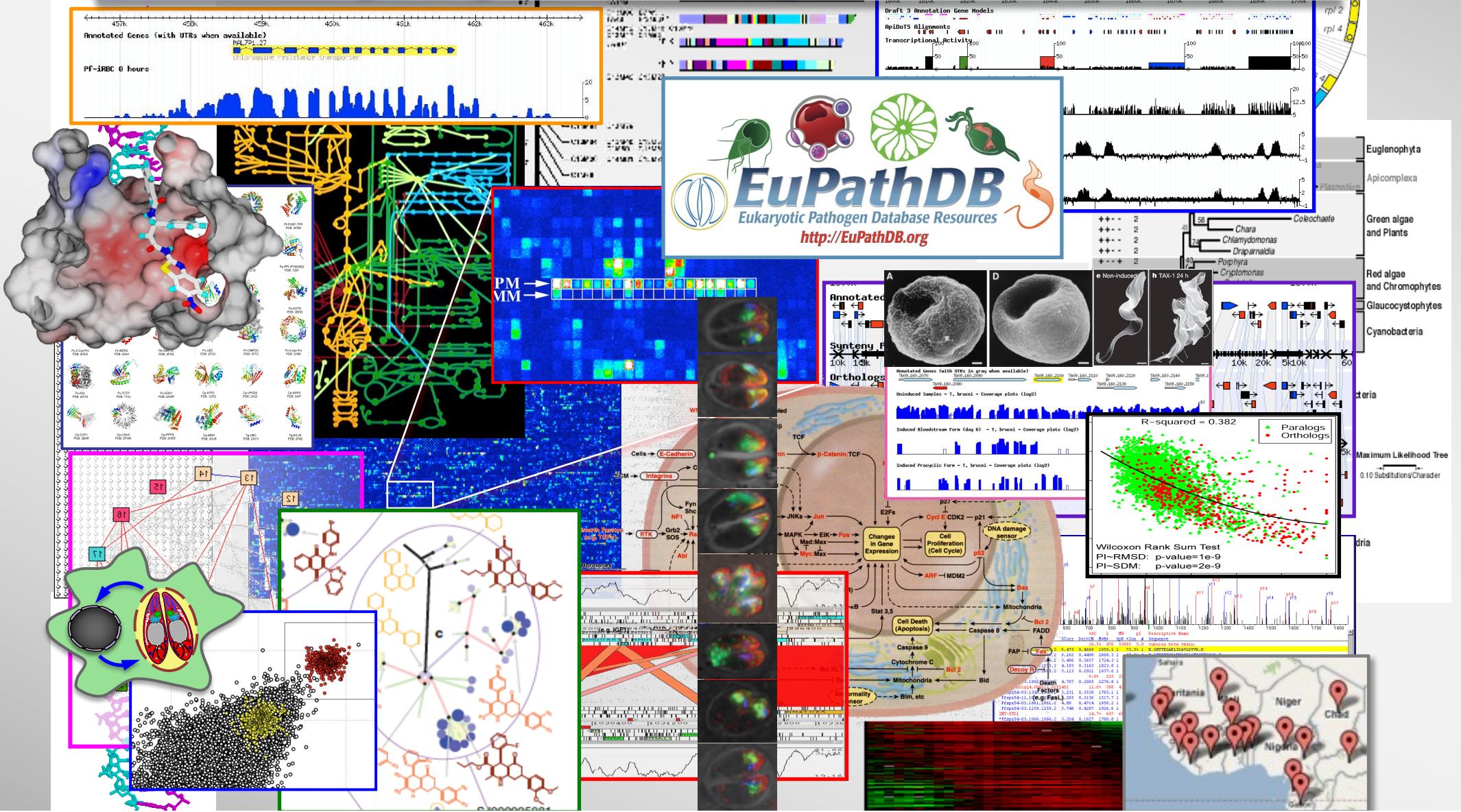
VEuPathDB Retweeted

 Ross Waller
@RossWaller3

Very pleased to present our recent study of #endocytosis in #apicomplexans that draws on #K13 and drug-resistance in #malaria to learn of a highly unusual endocytic machinery in #Toxoplasma with a deep evolutionary origin.
biorxiv.org/cgi/content/sh...



BIG DATA: Designing and Mining 'Omics Resources



Central Dogma: DNA —> RNA —> Protein —> Function

Genomes

Sequencing (*many technologies, changing*) – **Sanger, 454, Illumina, Nanopore, PacBio, etc**

Alignment & Assembly (*various algorithms*)

Variants & Population Biology (*within species; various algorithms*) – **SNPs, indels, genome rearrangement, microsatellites, CNVs, LoH, signatures of selection, geolocation mapping of isolates**

Comparative Genomics (*between species; various algorithms*) – **synteny, signatures of selection, gene duplication**

Structural Annotation

Gene Models (*various algorithms*) – **mRNAs, ncRNAs** (rRNA, tRNA, uRNA, lncRNA, etc)

Transcriptomics – **microarrays, ribosomal profiling, transcription factor binding sites**, etc (*many technologies: ESTs, SAGE, spotted arrays, Affymetrix, Illumina, RNA-seq, CAP-seq, FLAM-seq, scRNA-seq, etc*)

Epigenomics – **chromatin marks**, etc (*many technologies: ChIP-chip, ChIP-seq, FAIRE, MAINE, ATAC-seq, Hi-C, etc*)

Annotation – **standards, ontologies, manual curation, automated analysis** (*technologies: Artemis, Apollo, GO, EC, etc*)

Proteomics – **tandem mass-spectrometry, phosphoproteomics**, etc (*many technologies & analytical algoritms*)

Functional Assessment

Protein Domains (*various algorithms: Interpro domains: Pfam, etc*)

Metabolic Pathways (*automated analysis, manual curation*)

Metabolomics – **quadropole mass-spectrometry**, etc (*many technologies & analytical algoritms*)

Functional Assays – **knock-outs, knock-downs, high-content screening, functional enrichment** (*many technologies: insertional tagging, siRNA, morpholinos, CRISPr, protein tagging, protein destabilitation, etc*)

Orthology-based Inference – **in-paralogs, out-paralogs, phylogeny, phyletic patterns** (*various algorithms*)

A VEuPathDB Project

VectorBase

Bioinformatics Resources for
Invertebrate Vectors of Human Pathogens

Release 57
21 Apr 2022

Site search, e.g. AGAP004730 or *reductase or "binding protein"

My Strategies Searches Tools My Workspace Data About Help Contact Us

My Organism Preferences (58 of 58) enabled

Search for...

[expand all](#) | [collapse all](#)

Filter the searches below...

Genes

- ▶ Annotation, curation and identifiers
- ▶ Function prediction
- ▶ Gene models
- ▶ Genomic Location
- ▶ Immunology
- ▶ Orthology and synteny
- ▶ Pathways and interactions
- ▶ Protein features and properties
- ▶ Protein targeting and localization
- ▶ Proteomics
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Taxonomy
- ▶ Text
- ▶ Transcriptomics

Organisms

Genomic Sequences

Overview of Resources and Tools

Getting Started

VEuPathDB is packed with data, tools and visualizations that can help answer your research questions. We gather data from many sources, analyze according to standard workflows, and present the results for you to mine in a point and click interface. Here's how to get started:

SITE SEARCH: Explore the site; find what you need

Enter a term or ID in the site search box at the top of any page. The site search finds documents and records that contain your term and returns a summary of categorized matches. It's easy to find genes, pathways, searches, data sets and more with the site search.

[Read More](#)

Tutorials and Exercises

Grid view

[Apollo: Manual gene annotation](#)

[Gene Pages](#)

[Genetic Variation](#)

[Genome Annotation](#)

©2022 The VEuPathDB Project Team

COMMUNITY CHAT

News »

[VectorBase 57 Released](#)

THU APR 21 2022

[Ticks and Tick-Borne Pathogens Open Community Webinar Series](#)

WED FEB 16 2022

[VectorBase 56 Released](#)

TUE FEB 15 2022

[See all news](#)

Tweets by @VectorBase

VectorBase Retweeted

Pratima Chennuri
@PratimaChennuri

Research Associate position available in the Myles lab in Texas (Drosophila/Mosquito/small RNA/flaviviruses)
[tinyurl.com/TXresearchassoc...](http://tinyurl.com/TXresearchassoc)

@JobsVector @VectorBase

@FlyBaseDotOrg. RTs are much appreciated.

May 31, 2022

VectorBase Retweeted

Pratima Chennuri
@PratimaChennuri

PostDoc position available to study anti-

Export as a Search Strategy
to download or mine your results

All results matching odorant (filtered by organisms)

1 - 20 of 204

1

 Hide zero counts

Filter results

Genome	
Genes	199
Metabolism	
Compounds	4
Data access	
Data sets	1

Filter fields

Select a result filter above

Filter organisms

[Clear filter](#)[select only these](#) | [add these](#) | [clear these](#)

pest

- Arthropoda 8,841
- Insecta 8,749
- Diptera 8,359
 - Culicidae 6,381
 - Anopheles 4,126
 - gambiae species complex 1,688
 - Anopheles gambiae 200
 - Anopheles 200
 - gambiae PEST

Gene - AGAP003306 unspecified product

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains

Gene - AGAP001556 odorant-binding protein antennal 7

Gene name or symbol: OBP7

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions

Gene - AGAP009629 odorant-binding protein antennal 5

Gene name or symbol: OBP5

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions; PubMed

Gene - AGAP010489 odorant-binding protein antennal 4

Gene name or symbol: OBP4

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions

Gene - AGAP001409 odorant-binding protein antennal 3

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions

Gene - AGAP003307 unspecified product

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains

Gene - AGAP029062 odorant-binding protein 1

Gene name or symbol: OBP1

Organism: Anopheles gambiae PEST

▶ Fields matched: GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions

COMMUNITY CHAT

My Search Strategies

[Opened \(1\)](#) [All \(1\)](#) [Public \(21\)](#) [Help](#)

 Unnamed Search Strategy * 


text='Odorant' species=AgPEST
199 Genes

+ Add a step

Step 1

 199 Genes (116 ortholog groups) [Revise this search](#)
[Gene Results](#) [Genome View](#) [Analyze Results](#)
Organism Filter
[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)
 Hide zero counts

[Search organisms...](#)

[Arthropoda](#) 199

[Mollusca](#) 0

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)
 Hide zero counts

[Hide Organism Filter](#)
■ Genes on forward strand;

■ Genes on reversed strand;

[Search](#)


6 rows

 Show empty chromosomes

Rows per page: 20

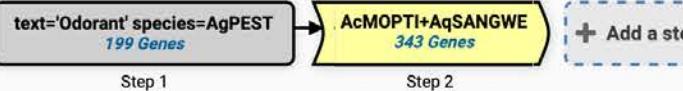
Sequence	Organism	Chromosome	#Genes	Length	Gene Locations
AgamP4_3R	Anopheles gambiae PEST	3R	55	53200684	
AgamP4_2R	Anopheles gambiae PEST	2R	48	61545105	
AgamP4_3L	Anopheles gambiae PEST	3L	36	41963435	
AgamP4_2L	Anopheles gambiae PEST	2L	35	49364325	
AgamP4_X	Anopheles gambiae PEST	X	16	24393108	
AgamP4_UNKN	Anopheles gambiae PEST	UNKN	9	42389979	

 COMMUNITY CHAT

My Search Strategies

Opened (1) All (1) Public (21) Help

Unnamed Search Strategy * [Edit](#)



343 Genes (113 ortholog groups) [Revise this search](#)

Gene Results [Genome View](#) [Analyze Results](#)

Show empty chromosomes

Rows per page: 20

Organism Filter

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Hide zero counts

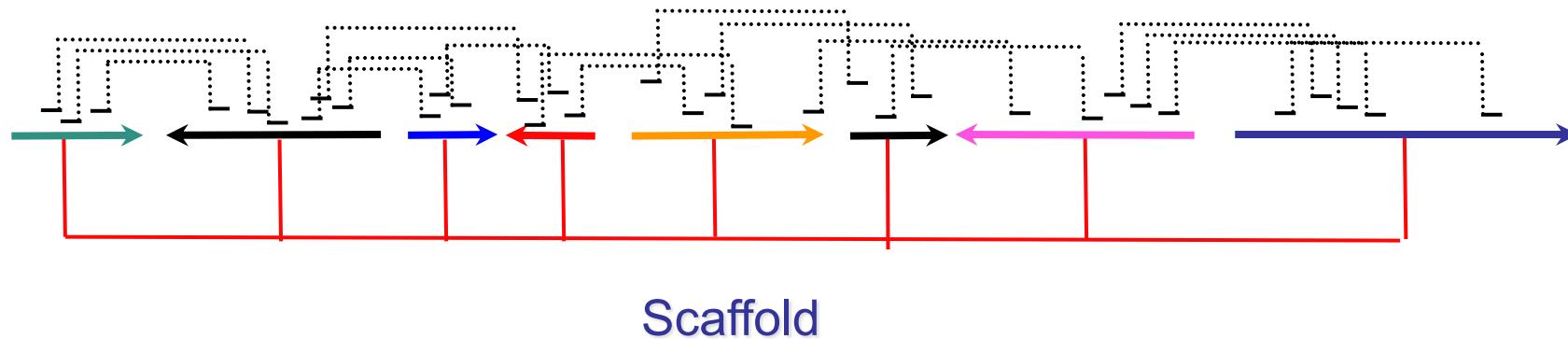
Search organisms... [?](#)

Sequence	Organism	Chromosome	#Genes	Length ?	Gene Locations
CM029348	Anopheles coluzzii MOPTI	2	69	114808232	
CM029349	Anopheles coluzzii MOPTI	3	77	99875506	
CM029350	Anopheles coluzzii MOPTI	X	14	27656523	
KB667777	Anopheles quadriannulatus SANGWE		7	4596278	
KB665621	Anopheles quadriannulatus SANGWE		4	3387546	
KB667999	Anopheles quadriannulatus		9	3255115	

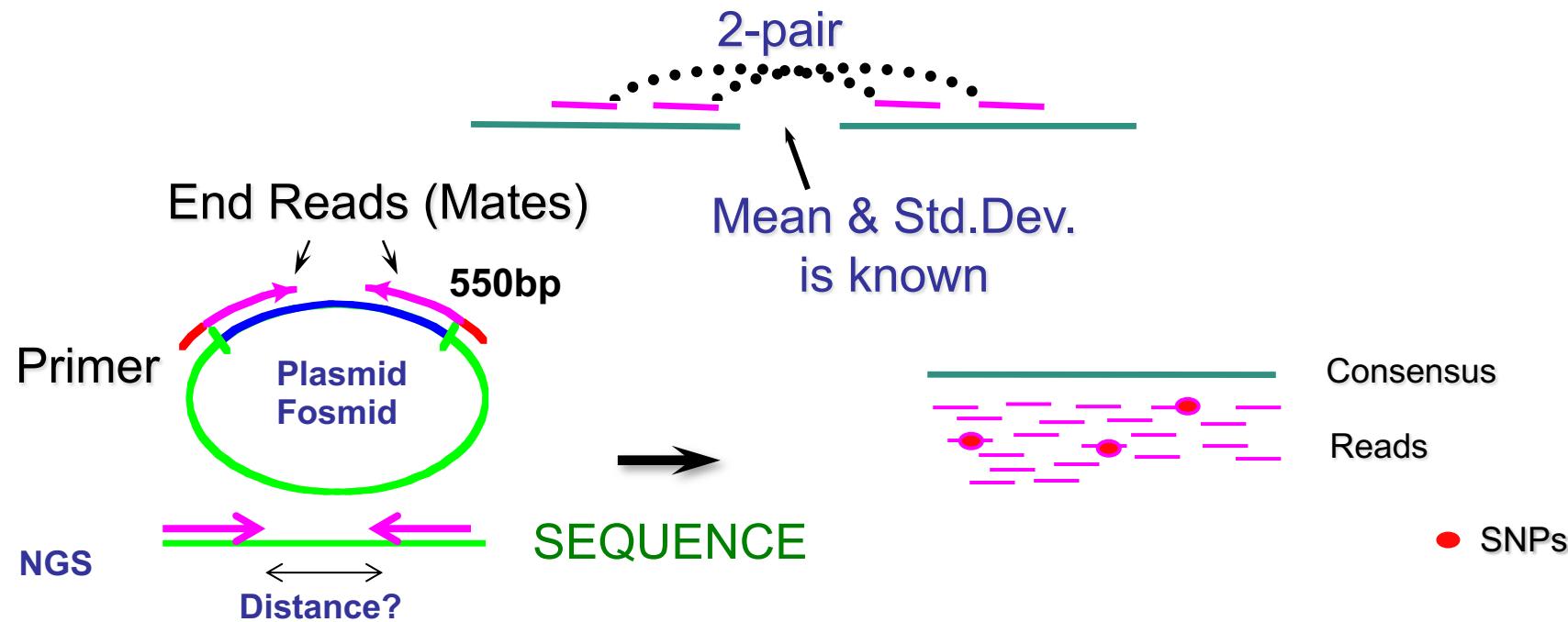
[Hide Organism Filter](#)

[Community Chat](#)

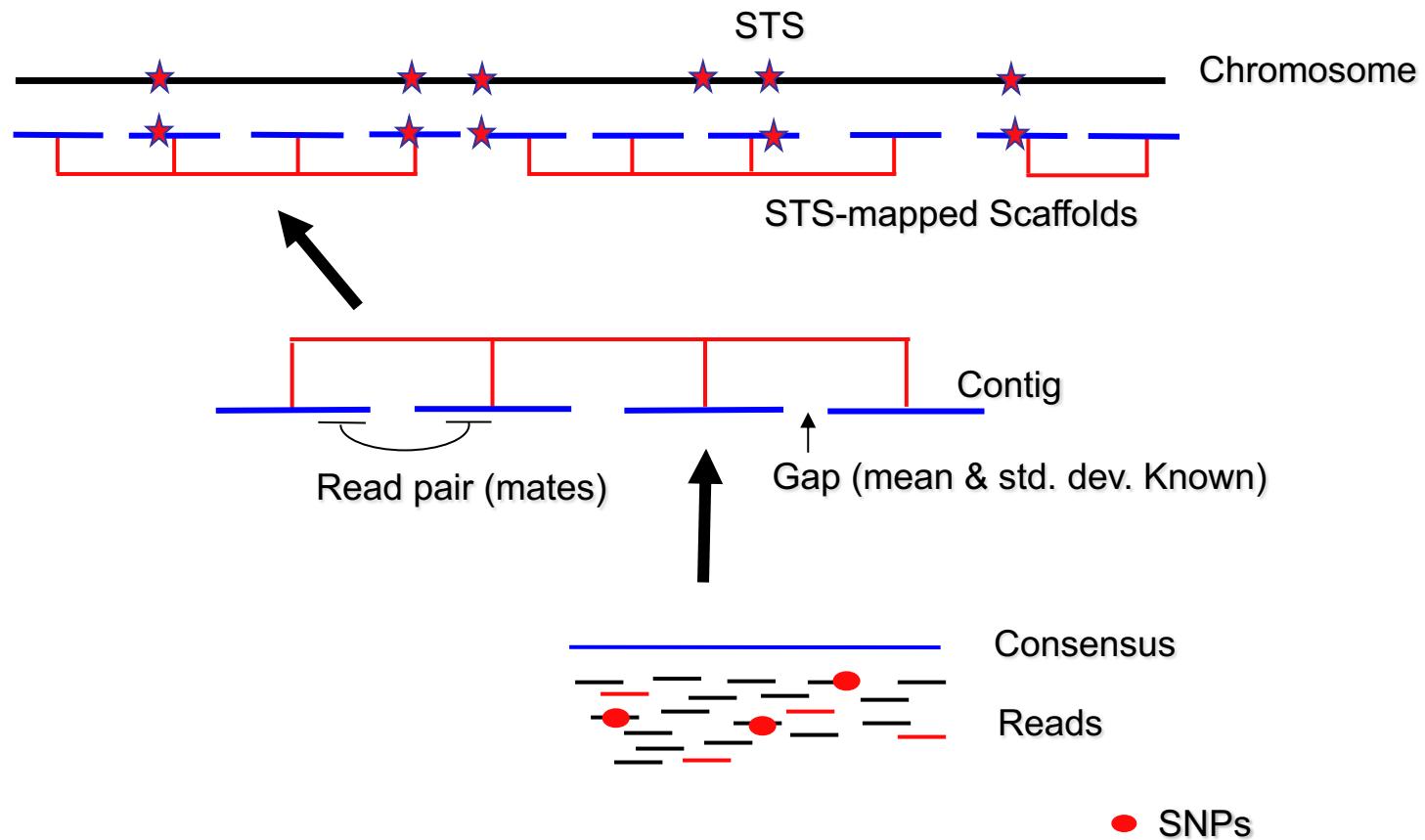
Pairs Give Order & Orientation



Gaps in scaffolds are traditionally indicated by 100 “N’ s”



Anatomy of a WGS Assembly





[Add to basket](#) [Add to favorites](#) [Download Gene](#)

AGAP012323 odorant-binding protein 27

Name: OBP27

Gene Type: protein coding gene

Biotype Classification: protein_coding

Chromosome: 3L

Location:

AgamP4_3L:40,218,226..40,218,764(+)

Species: *Anopheles gambiae*

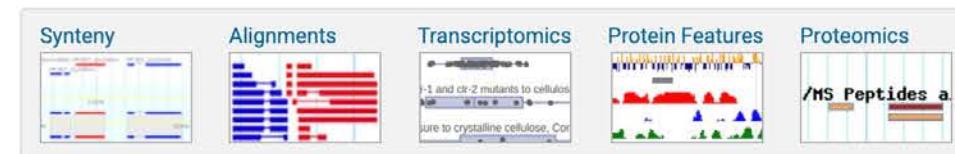
Strain: PEST

Status: Reference Strain

Add the first user comment

View and update community annotations in Apollo

Shortcuts



Also see AGAP012323 in the [Genome Browser](#) or [Protein Browser](#)

[Collapse all sections for better performance](#)

1 Gene models

AGAP012323



[expand all](#) | [collapse all](#)

Search section names...

- ▶ 1 Gene models
- ▶ 2 Annotation, curation and identifiers
- ▶ 3 Link outs
- ▶ 4 Genomic Location
- ▶ 5 Literature
- ▶ 6 Taxonomy
- ▶ 7 Orthology and synteny
- ▶ 8 Phenotype

Exons in Gene 5

Transcripts 2

▼ Gene Models

This gene is available in **Apollo** for community annotation. To find out more about Apollo, please visit this help page.

[View in JBrowse genome browser](#)

[Annotate in Apollo](#)

COMMUNITY CHAT

AGAP012323


[expand all](#) | [collapse all](#)

- Search section names... ?
- 1 Gene models
 - 2 Annotation, curation and identifiers
 - 3 Link outs
 - 4 Genomic Location
 - 5 Literature
 - 6 Taxonomy
 - 7 Orthology and synteny
 - 8 Phenotype
 - 9 Transcriptomics
 - 10 Sequence analysis
 - 11 Sequences
 - 12 Structure analysis
 - 13 Protein features and properties
 - 14 Function prediction
 - 15 Pathways and interactions
 - 16 Proteomics
 - 17 Immunology
- [expand all](#) | [collapse all](#)

7 Orthology and synteny

Ortholog Group ? OG6_315879

Orthologs and Paralogs within VectorBase ? Data sets

To run Clustal Omega, select genes from the table below. Then choose the sequence type and initiate the alignment with the 'Run Clustal Omega for selected genes' button.

 Search this table...


Clustal Omega ?	Gene	Product	Organism	Reference Strain?	Syntenic?	has comments ?
<input type="checkbox"/>	AARA21_001324	general odorant-binding protein 56d-like	Anopheles arabiensis DONGOLA 2021	no	no	no
<input type="checkbox"/>	AARA017061	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012323) VB Community Annotation]	Anopheles arabiensis Dongola	yes	yes	no
<input type="checkbox"/>	ACMO_011149	general odorant-binding protein 56d-like	Anopheles coluzzii MOPTI	no	no	no
<input type="checkbox"/>	ACOM026266	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012323) VB Community Annotation]	Anopheles coluzzii Mali-NIH	no	yes	no
<input type="checkbox"/>	ACON012323	unspecified product	Anopheles coluzzii Ngousso	yes	yes	no
<input type="checkbox"/>	AMAM014313	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012323) VB Community Annotation]	Anopheles maculatus maculatus3	yes	yes	no
<input type="checkbox"/>	AMEC006571	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012323) VB Community Annotation]	Anopheles melas CM1001059_A	yes	yes	no
<input type="checkbox"/>	AMEM010006	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012323) VB Community Annotation]	Anopheles merus MAF	yes	yes	no
<input type="checkbox"/>	AMEM21_002175	general odorant-binding protein 56d-like	Anopheles merus MAF 2021	no	yes	no
<input type="checkbox"/>	AQUA017091	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012323) VB Community Annotation]	Anopheles quadriannulatus SANGWE	yes	yes	no

[Check All](#) [Uncheck All](#)

Select sequence type for Clustal Omega multiple sequence alignment:

Please note: selecting a large flanking region or a large number of sequences will take several minutes to align.



Select Tracks

My Tracks Currently Active Recently Used

Category

- 1 Comparative Genomics
- 6 Epigenomics
- 3 Gene Models
- 22 Genetic Variation
- 6 Proteomics
- 15 Sequence Analysis**
- 895 Transcriptomics

Subcategory

- 1 (no data)
- 6 Array Probes
- 2 BLAT and Blast Alignments
- 1 Sequence assembly
- 3 Sequence composition, complexity and repeats
- 2 Sequence sites, features and

? About this track

- Pin to top
- Edit config
- Delete track
- Save track data

Display mode Show labels Select Subtracks

Agilent 13k v1

Track Type

- 1 Reference Sequence
- 13 Segments
- 1 XYPlot

aalbSTECLA2020 gene

aalbSTECLA2020 span

aaraDONGOLA2021 gene

aaraDONGOLA2021 span

aaraDongola gene

aaraDongola span

aatrEBRO gene

aatrEBRO span

achrACHKN1017 gene

achrACHKN1017 span

acolMOPTI gene

acolMOPTI span

acolNgouso gene

acolNgouso span

adarCoari gene

adarCoari span

adirWRAIR2 gene

adirWRAIR2 span

Back to browser Contains text

15 matching tracks

Name	Category	Subcategory	Dataset	Track Type	RNA-Seq Alignment
Centromere	Sequence Analysis	Sequence sites, features and motifs	...	Segments	...
<input checked="" type="checkbox"/> GC Content	Sequence Analysis	Sequence composition, complexity and repeats	...	XYPlot	...
<input checked="" type="checkbox"/> Low Complexity Regions	Sequence Analysis	Sequence composition, complexity and repeats	...	Segments	...
Microarray Probes A-GEOD-13157 - Agilent A. gambiae 020449 44k v2	Sequence Analysis	Array Probes	A-GEOD-13157 - Agilent A. gambiae 020449 44k v2	Segments	...
Microarray Probes A-MEXP-2196 - LSTM_Angambiae_s.s._AGAM15K_V 1.0	Sequence Analysis	Array Probes	A-MEXP-2196 - LSTM_Angambiae_s.s._AGAM15K_V 1.0	Segments	...
Microarray Probes A-MEXP-225 - EMBL A. gambiae MMC1 20k v1.0	Sequence Analysis	Array Probes	A-MEXP-225 - EMBL A. gambiae MMC1 20k v1.0	Segments	...
Microarray Probes A-MEXP-863 - LIV A. gambiae DETOX	Sequence Analysis	Array Probes	A-MEXP-863 - LIV A. gambiae DETOX	Segments	...

0,000 30,000,000 32,000,000 34,000,000 36,000,000 38,000,000 40,000,000

40,218,750 40,220,000 40,221,250

AGAP012324-RA
AGAP012324
odorant-binding protein 68

AGAP012325-RA
AGAP012325
odorant-binding protein 28

Stein 27

COMMUNITY CHAT

Central Dogma: DNA —> RNA —> Protein —> Function

Genomes

Sequencing (*many technologies, changing*) – Sanger, 454, Illumina, Nanopore, PacBio, etc

Alignment & Assembly (*various algorithms*)

Variants & Population Biology (*within species; various algorithms*) – SNPs, indels, genome rearrangement, microsatellites, CNVs, LoH, signatures of selection, geolocation mapping of isolates

Comparative Genomics (*between species; various algorithms*) – synteny, signatures of selection, gene duplication

Structural Annotation

Gene Models (*various algorithms*) – mRNAs, ncRNAs (rRNA, tRNA, uRNA, lncRNA, etc)

Transcriptomics – microarrays, ribosomal profiling, transcription factor binding sites, etc (*many technologies: ESTs, SAGE, spotted arrays, Affymetrix, Illumina, RNA-seq, CAP-seq, FLAM-seq, scRNA-seq, etc*)

Epigenomics – chromatin marks, etc (*many technologies: ChIP-chip, ChIP-seq, FAIRE, MAINE, ATAC-seq, Hi-C, etc*)

Annotation – standards, ontologies, manual curation, automated analysis (*technologies: Artemis, Apollo, GO, EC, etc*)

Proteomics – tandem mass-spectrometry, phosphoproteomics, etc (*many technologies & analytical algoritms*)

Functional Assessment

Protein Domains (*various algorithms: Interpro domains: Pfam, etc*)

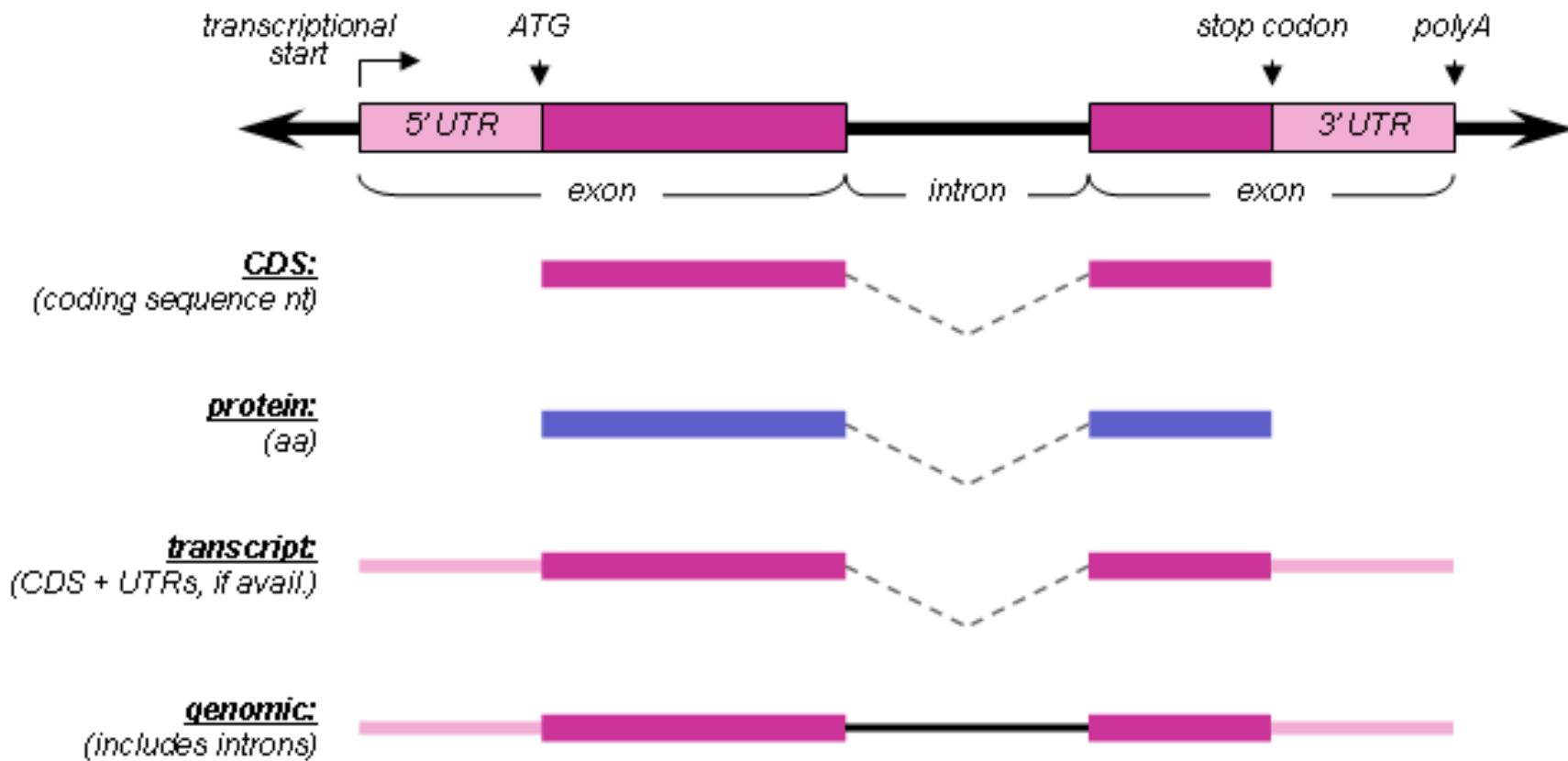
Metabolic Pathways (*automated analysis, manual curation*)

Metabolomics – quadropole mass-spectrometry, etc (*many technologies & analytical algoritms*)

Functional Assays – knock-outs, knock-downs, high-content screening, functional enrichment (*many technologies: insertional tagging, siRNA, morpholinos, CRISPr, protein tagging, protein destabilitation, etc*)

Orthology-based Inference – in-paralogs, out-paralogs, phylogeny, phyletic patterns (*various algorithms*)

Terminology





[Add to basket](#) [Add to favorites](#) [Download Gene](#)

AGAP012323 odorant-binding protein 27

Name: OBP27

Gene Type: protein coding gene

Biotype Classification: protein_coding

Chromosome: 3L

Location:

AgamP4_3L:40,218,226..40,218,764(+)

Species: *Anopheles gambiae*

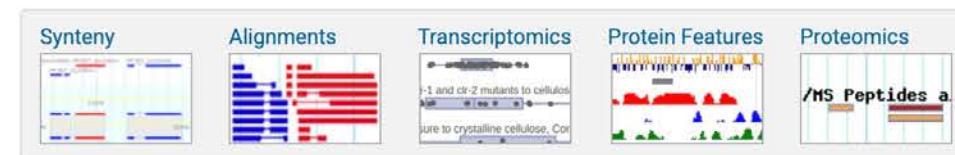
Strain: PEST

Status: Reference Strain

Add the first user comment

View and update community annotations in Apollo

Shortcuts



Also see AGAP012323 in the [Genome Browser](#) or [Protein Browser](#)

[Collapse all sections for better performance](#)

1 Gene models

AGAP012323



[expand all](#) | [collapse all](#)

Search section names...

- ▶ 1 Gene models
- ▶ 2 Annotation, curation and identifiers
- ▶ 3 Link outs
- ▶ 4 Genomic Location
- ▶ 5 Literature
- ▶ 6 Taxonomy
- ▶ 7 Orthology and synteny
- ▶ 8 Phenotype

Exons in Gene 5

Transcripts 2

Gene Models

This gene is available in **Apollo** for community annotation. To find out more about Apollo, please visit this help page.

[View in JBrowse genome browser](#)

[Annotate in Apollo](#)

COMMUNITY CHAT

1 Gene models

AGAP012323



[expand all](#) | [collapse all](#)

- Search section names...
- [1 Gene models](#)
 - [2 Annotation, curation and identifiers](#)
 - [3 Link outs](#)
 - [4 Genomic Location](#)
 - [5 Literature](#)
 - [6 Taxonomy](#)
 - [7 Orthology and synteny](#)
 - [8 Phenotype](#)
 - [9 Transcriptomics](#)
 - [10 Sequence analysis](#)
 - [11 Sequences](#)
 - [12 Structure analysis](#)
 - [13 Protein features and properties](#)
 - [14 Function prediction](#)
 - [15 Pathways and interactions](#)
 - [16 Proteomics](#)
 - [17 Immunology](#)

[expand all](#) | [collapse all](#)

Exons in Gene 5

Transcripts 2

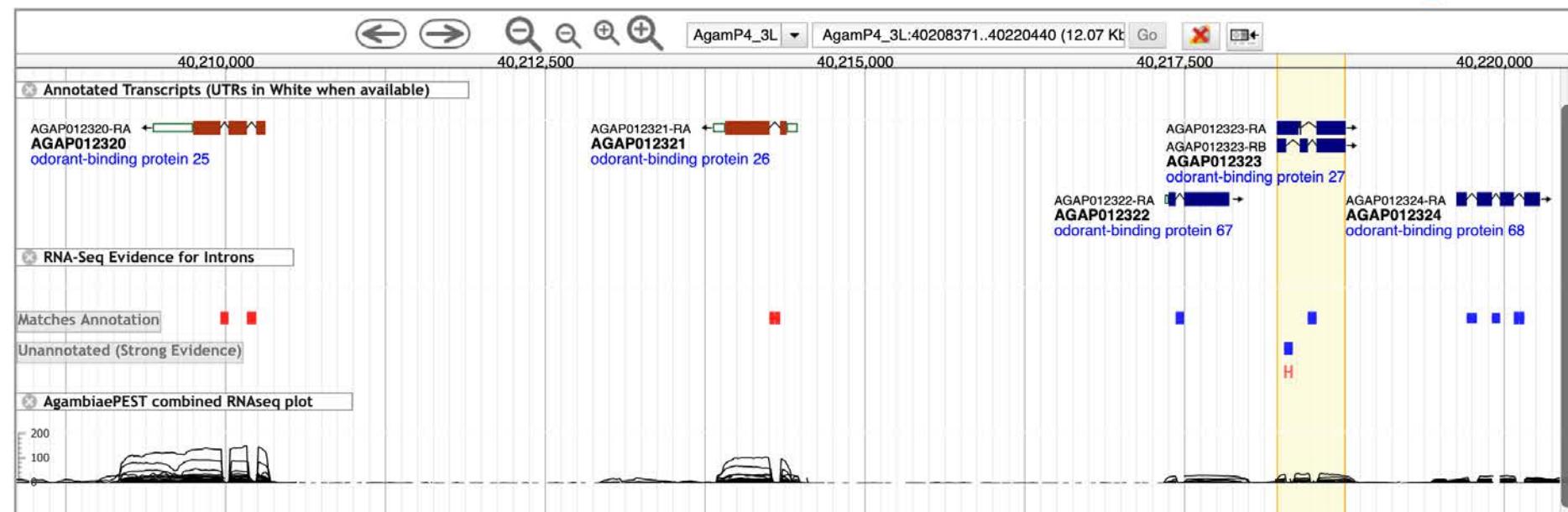
Gene Models

This gene is available in **Apollo** for community annotation. To find out more about Apollo, please visit [this help page](#).

[View in JBrowse genome browser](#)

[Annotate in Apollo](#)

[Scroll and zoom](#)



[View in JBrowse genome browser](#)

[Annotate in Apollo](#)

GFF format of gene and transcript features Click to open GFF in a new tab

[Transcripts](#)

COMMUNITY CHAT

AGAP012323



[expand all](#) | [collapse all](#)

- Search section names...
- ▶ 1 Gene models
 - ▶ 2 Annotation, curation and identifiers
 - ▶ 3 Link outs
 - ▶ 4 Genomic Location
 - ▶ 5 Literature
 - ▶ 6 Taxonomy
 - ▶ 7 Orthology and synteny
 - ▶ 8 Phenotype
 - ▶ 9 Transcriptomics
 - 10 Sequence analysis
 - ▶ 11 Sequences
 - ▶ 12 Structure analysis
 - ▶ 13 Protein features and properties
 - ▶ 14 Function prediction
 - ▶ 15 Pathways and interactions
 - ▶ 16 Proteomics
 - ▶ 17 Immunology
- [expand all](#) | [collapse all](#)

9 Transcriptomics

▼ RNA-Seq Transcription Summary NEW

► *Summary of expression values. Each row represents a distinct RNA-Seq experiment. [Click to read more...](#)*

Gene: AGAP012323

0 2 4 6 8 10

Anopheles gambiae small non-coding RNAs by tissue (aligned to A. gambiae PEST) - agamPEST, Bryant et al 2020

Antennae vs maxillary palps (An. gambiae), Hodges et al 2014

Antennal transcriptome expression profiles following a blood meal, Rinker et al 2013

Chemosensory appendages, male and female, Pitts et al 2011

Comparative transcriptome analyses of deltamethrin resistant and susceptible mosquitoes, Bonizzoni et al 2012

Effect of bloodmeal on mosquito whole body transcriptome, Vannini et al 2014

Female antennal transcriptome of A. gambiae sensu stricto aligned to A. gambiae PEST - agamPEST, Rinker et al 2013

Gene expression levels in Anopheles male reproductive organs (An. gambiae, aligned to PEST strain), Izquierdo et al 2019

HMBPP-supplemented blood meal, Emami et al 2017

Male and female sex-tissues vs carcass transcriptomes in adults and larvae, Papa et al 2017

Midgut and salivary glands transcription, Gómez-Díaz et al 2014

Mosquito gene expression after ivermectin ingestion, Seale et al 2015

Mosquito midgut and salivary glands (SGs) infected with P. falciparum (An. coluzzii), Izquierdo et al 2021

Plasmodium falciparum infection (polysomal and non-polysomal fractions), Meadow et al 2010

Testes transcriptome of males vs mated, Pitts et al 2014

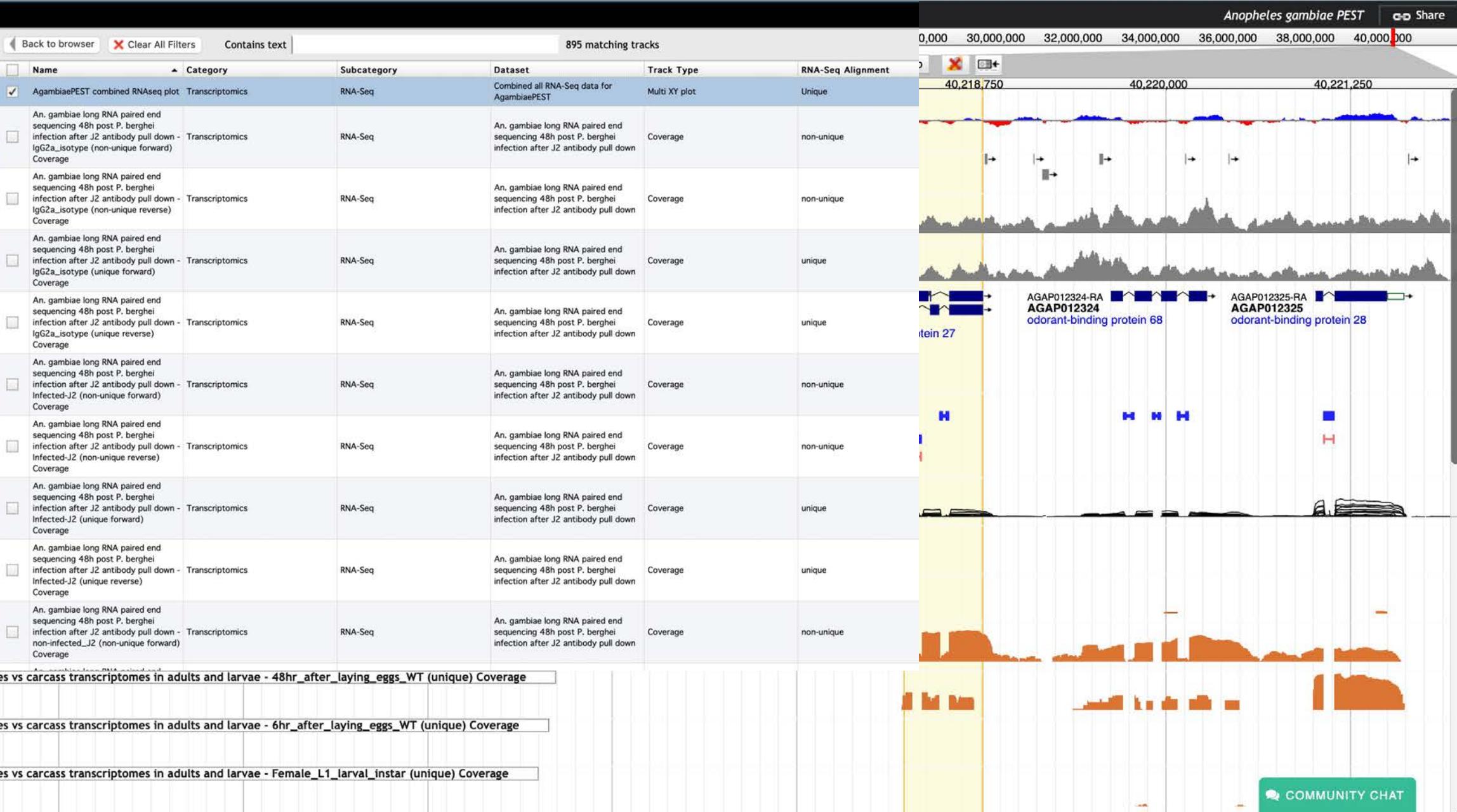
Testis vs carcass gene expression (An. coluzzii), Cassone et al 2017

The male mosquito contribution towards malaria transmission mating influences the Anopheles female midgut transcriptome and increases female susceptibility to human malaria parasites, [Cassone et al 2017](#)



Select Tracks

My Tracks	
Currently Active	
Recently Used	
Category	
1 Comparative Genomics	
6 Epigenomics	
3 Gene Models	
22 Genetic Variation	
6 Proteomics	
15 Sequence Analysis	
895 Transcriptomics	
Subcategory	
1 BLAT and Blast Alignments	
894 RNA-Seq	
Dataset	
2 (no data)	
16 An. gambiae long RNA paired end sequencing 48h post P. berghei infection after J2 antibody pull down	
12 An. gambiae small RNA 24h post P. berghei infection	
12 An. gambiae small RNA 48h post P. berghei infection	
20 Anopheles gambiae Small RNA Transcriptome	
20 Anopheles gambiae small non-coding RNAs by tissue (aligned to A. gambiae PEST)	
12 Anopheles gambiae strain:G3 Transcriptome or Gene expression	
6 Anopheles_sex_determination	
8 Antennae vs maxillary palps (An. gambiae)	
24 Antennal transcriptome expression profiles following a blood meal	
16 Chemosensory appendages, male and female	
1 Combined all RNA-Seq data for AgambiaePEST	
8 Comparative transcriptome analyses of deltamethrin resistant and susceptible mosquitoes	
16 Deep-RNA-Seq of Anopheles gambiae life stages (L1, L3, Adult Male and Adult Female)	
18 Developmental transcriptome of sexed Anopheles gambiae embryos	
24 Developmental transcriptome of sexed Anopheles gambiae larvae and adult mosquitoes	



Central Dogma: DNA —> RNA —> Protein —> Function

Genomes

Sequencing (*many technologies, changing*) – **Sanger, 454, Illumina, Nanopore, PacBio, etc**

Alignment & Assembly (*various algorithms*)

Variants & Population Biology (*within species; various algorithms*) – **SNPs, indels, genome rearrangement, microsatellites, CNVs, LoH, signatures of selection, geolocation mapping of isolates**

Comparative Genomics (*between species; various algorithms*) – **synteny, signatures of selection, gene duplication**

Structural Annotation

Gene Models (*various algorithms*) – **mRNAs, ncRNAs** (rRNA, tRNA, uRNA, lncRNA, etc)

Transcriptomics – **microarrays, ribosomal profiling, transcription factor binding sites**, etc (*many technologies: ESTs, SAGE, spotted arrays, Affymetrix, Illumina, RNA-seq, CAP-seq, FLAM-seq, scRNA-seq, etc*)

Epigenomics – **chromatin marks**, etc (*many technologies: ChIP-chip, ChIP-seq, FAIRE, MAINE, ATAC-seq, Hi-C, etc*)

Annotation – **standards, ontologies, manual curation, automated analysis** (*technologies: Artemis, Apollo, GO, EC, etc*)

Proteomics – **tandem mass-spectrometry, phosphoproteomics**, etc (*many technologies & analytical algoritms*)

Functional Assessment

Protein Domains (*various algorithms: Interpro domains: Pfam, etc*)

Metabolic Pathways (*automated analysis, manual curation*)

Metabolomics – **quadropole mass-spectrometry**, etc (*many technologies & analytical algoritms*)

Functional Assays – **knock-outs, knock-downs, high-content screening, functional enrichment** (*many technologies: insertional tagging, siRNA, morpholinos, CRISPr, protein tagging, protein destabilitation, etc*)

Orthology-based Inference – **in-paralogs, out-paralogs, phylogeny, phyletic patterns** (*various algorithms*)

AGAP012323

[expand all](#) | [collapse all](#)

- [Search section names...](#)
- ▶ 1 Gene models
- ▶ 2 Annotation, curation and identifiers
- ▶ 3 Link outs
- ▶ 4 Genomic Location
- ▶ 5 Literature
- ▶ 6 Taxonomy
- ▶ 7 Orthology and synteny
- ▶ 8 Phenotype
- ▶ 9 Transcriptomics
- 10 Sequence analysis
- 11 Sequences
- 12 Structure analysis
- 13 Protein features and properties
- 14 Function prediction
- 15 Pathways and interactions
- 16 Proteomics
- 17 Immunology

[expand all](#) | [collapse all](#)

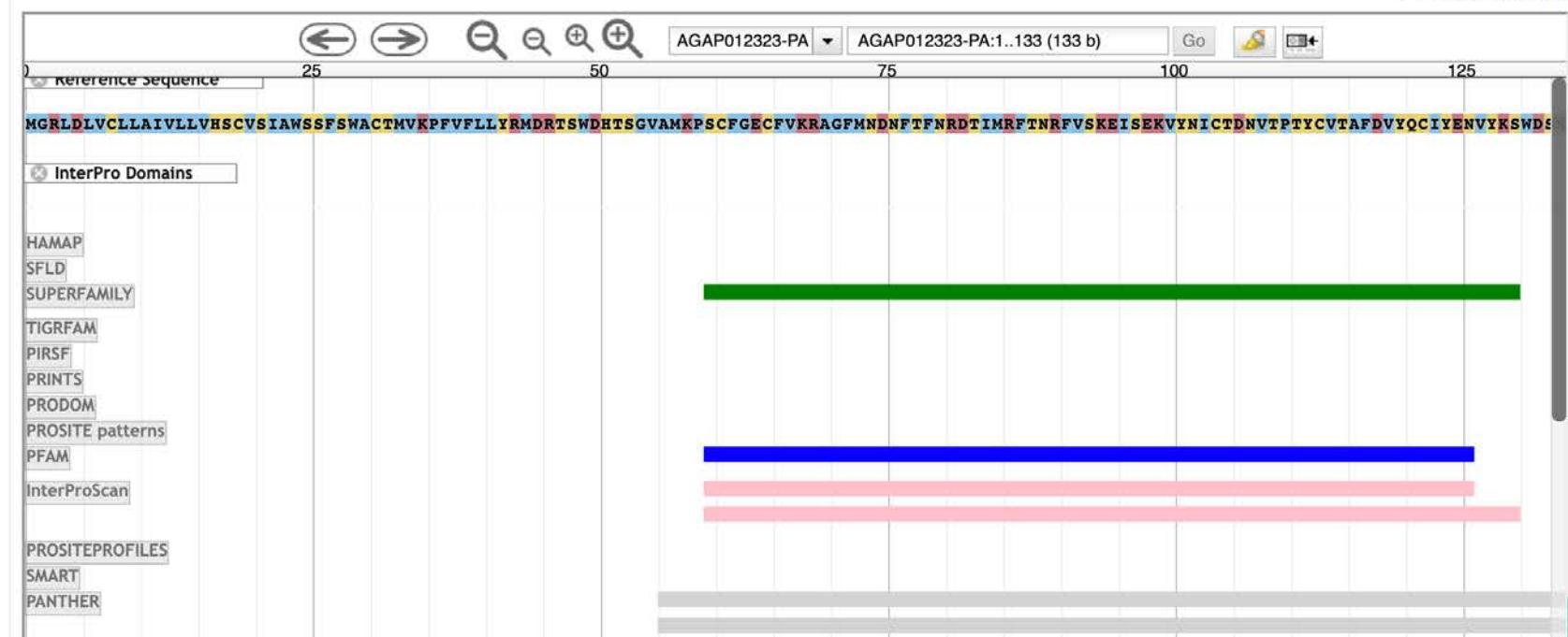
13 Protein features and properties

▼ Attributes and Protein Browser

This section contains features of each protein isoform: isoelectric point, molecular weight, Signal peptide (SignalP) domain, length, sequence, transmembrane domains (TMHMM), InterPro domains, predicted export domains, hydropathy plot, low complexity regions, PSIPRED Helix, Strand and Coil predictions.

[Search this table...](#)

	↑↓ Transcript ID	↑↓ Isoelectric Point	↑↓ Molecular Weight	↑↓ Has SignalP	↑↓ Has TMHMM	↑↓ Protein Length	↑↓ Protein Browser
▼	AGAP012323-RA	8.08	15645	yes	yes	134	Interactive

[View in protein browser](#) Scroll and zoom 

COMMUNITY CHAT

AGAP012323


[expand all](#) | [collapse all](#)

Search section names...

- ▶ 1 Gene models
- ▶ 2 Annotation, curation and identifiers
- ▶ 3 Link outs
- ▶ 4 Genomic Location
- ▶ 5 Literature
- ▶ 6 Taxonomy
- ▶ 7 Orthology and synteny
- ▶ 8 Phenotype
- ▶ 9 Transcriptomics
- ▶ 10 Sequence analysis
- ▶ 11 Sequences
- ▶ 12 Structure analysis
- ▶ 13 Protein features and properties
- ▶ 14 Function prediction
- ▶ 15 Pathways and interactions
- ▶ 16 Proteomics
- ▶ 17 Immunology

[expand all](#) | [collapse all](#)

14 Function prediction

▶ EC Numbers Data sets

▶ EC Numbers (scores for computational inference)

▶ GO Slim Download Data sets

Search this table...



⬆ Ontology	⬇ GO ID	⬇ Is Not	⬇ GO Term Name	⬇ GO Slim ID	⬇ GO Slim Term Name
Biological Process	GO:0007608	N/A	sensory perception of smell	GO:0050877	nervous system process
Biological Process	GO:0050896	N/A	response to stimulus	GO:0008150	biological_process
Molecular Function	GO:0005549	N/A	odorant binding	GO:0003674	molecular_function

▶ GO Terms Download Data sets

Search this table...



⬆ Ontology	⬇ GO ID	⬇ Is Not	⬇ GO Term Name	⬇ Source	⬇ Evidence Code	⬇ Reference	⬇ Support for Evidence Code Assignment	⬇ Transcript ID(s)
Biological Process	GO:0007608	N/A	sensory perception of smell	goa_import	IBA	N/A	N/A	AGAP012323-RA
Biological Process	GO:0007608	N/A	sensory perception of smell	gouniprot	IBA	N/A	N/A	AGAP012323-RA
Biological Process	GO:0007608	N/A	sensory perception of smell	gouniprot	IEA	N/A	N/A	AGAP012323-RB
Biological Process	GO:0050896	N/A	response to stimulus	gouniprot	IEA	N/A	N/A	AGAP012323-RA, AGAP012323-RB
Molecular Function	GO:0005549	N/A	odorant binding	goa_import	IEA	N/A	N/A	AGAP012323-RA, AGAP012323-RB
Molecular Function	GO:0005549	N/A	odorant binding	gouniprot	IEA	N/A	N/A	AGAP012323-RA, AGAP012323-RB

COMMUNITY CHAT



Central Dogma: DNA —> RNA —> Protein —> Function

Genomes

Sequencing (*many technologies, changing*) – **Sanger, 454, Illumina, Nanopore, PacBio, etc**

Alignment & Assembly (*various algorithms*)

Variants & Population Biology (*within species; various algorithms*) – **SNPs, indels, genome rearrangement, microsatellites, CNVs, LoH, signatures of selection, geolocation mapping of isolates**

Comparative Genomics (*between species; various algorithms*) – **synteny, signatures of selection, gene duplication**

Structural Annotation

Gene Models (*various algorithms*) – **mRNAs, ncRNAs** (rRNA, tRNA, uRNA, lncRNA, etc)

Transcriptomics – **microarrays, ribosomal profiling, transcription factor binding sites**, etc (*many technologies: ESTs, SAGE, spotted arrays, Affymetrix, Illumina, RNA-seq, CAP-seq, FLAM-seq, scRNA-seq, etc*)

Epigenomics – **chromatin marks**, etc (*many technologies: ChIP-chip, ChIP-seq, FAIRE, MAINE, ATAC-seq, Hi-C, etc*)

Annotation – **standards, ontologies, manual curation, automated analysis** (*technologies: Artemis, Apollo, GO, EC, etc*)

Proteomics – **tandem mass-spectrometry, phosphoproteomics**, etc (*many technologies & analytical algoritms*)

Functional Assessment

Protein Domains (*various algorithms: Interpro domains: Pfam, etc*)

Metabolic Pathways (*automated analysis, manual curation*)

Metabolomics – **quadropole mass-spectrometry**, etc (*many technologies & analytical algoritms*)

Functional Assays – **knock-outs, knock-downs, high-content screening, functional enrichment** (*many technologies: insertional tagging, siRNA, morpholinos, CRISPr, protein tagging, protein destabilitation, etc*)

Orthology-based Inference – **in-paralogs, out-paralogs, phylogeny, phyletic patterns** (*various algorithms*)

DNA —> RNA —> protein —> function

Query Strategies
(*in silico* experiments)

Genomes

Sequencing (many technologies, changing) – Sanger, 454, Illumina, Nanopore, PacBio, etc

Alignment & Assembly (various algorithms)

Variants & Population Biology (within species; various algorithms) – SNPs, indels, genome rearrangement, microsatellites, CNVs, LoH, signatures of selection, geolocation mapping of isolates

Comparative Genomics (between species; various algorithms) – synteny, signatures of selection, gene duplication

Structural Annotation

Gene Models (various algorithms) – mRNAs, ncRNAs (rRNA, tRNA, uRNA, lncRNA, etc)

Transcriptomics – microarrays, ribosomal profiling, transcription factor binding sites, etc (many technologies: ESTs, SAGE, spotted arrays, Affymetrix, Illumina, RNA-seq, CAP-seq, FLAM-seq, scRNA-seq, etc)

Epigenomics – chromatin marks, etc (many technologies: ChIP-chip, ChIP-seq, FAIRE, MAINE, ATAC-seq, Hi-C, etc)

Annotation – standards, ontologies, manual curation, automated analysis (technologies: Artemis, Apollo, GO, EC, etc)

Proteomics – tandem mass-spectrometry, phosphoproteomics, etc (many technologies & analytical algorithms)

Protein Domains (various algorithms: Interpro domains: Pfam, etc)

Metabolic Pathways (automated analysis, manual curation)

Metabolomics – quadropole mass-spectrometry, etc (many technologies & analytical algorithms)

Functional Assays – knock-outs, knock-downs, high-content screening, functional enrichment (many technologies: insertional tagging, siRNA, morpholinos, CRISPR, protein tagging, protein destabilization, etc)

Orthology-based Inference – *in-paralogs, out-paralogs, phylogeny, phyletic patterns* (various algorithms)

Functional Assessment

My Search Strategies

[Opened \(2\)](#) All (2) Public (21) Help

Mid-Atlantic Tick Summit 11 example *



16 Genes (16 ortholog groups)

Organism Filter		Gene Results				Genome View		Analyze Results		
select all clear all expand all collapse all <input type="checkbox"/> Hide zero counts		Rows per page: 20						Download	Add to Basket	Add Columns
<input type="checkbox"/> Arthropoda 16 <input type="checkbox"/> Mollusca 0 select all clear all expand all collapse all <input type="checkbox"/> Hide zero counts		Gene ID Transcript ID Genomic Location (Gene) Product Description # Transcripts								
		ISCW001655 ISCW015883 ISCW006008 ISCW018541 ISCW007220 ISCW009102 ISCW009204 ISCW010226 ISCW010371 ISCW021550	ISCW001655-RA ISCW015883-RA ISCW006008-RA ISCW018541-RA ISCW007220-RA ISCW009102-RA ISCW009204-RA ISCW010226-RA ISCW010371-RA ISCW021550-RA	DS636787:157,403..162,989(-) DS642473:468,151..480,607(-) DS750534:11,016..18,937(-) DS762074:249,972..314,299(+) DS802612:66,797..70,832(-) DS825751:61,274..72,460(+) DS826205:770..7,607(-) DS833764:323,376..324,165(+) DS866430:77,615..100,122(+) DS868232:34,862..40,399(+)	Pupal cuticle protein G1A, putative [Source:UniProtKB/TrEMBL;Acc:B7P4Q6] monocarboxylate transporter, putative Secreted protein, putative [Source:UniProtKB/TrEMBL;Acc:B7PN67] Secreted salivary gland peptide, putative [Source:UniProtKB/TrEMBL;Acc:B7PPZ7] Cell adhesion molecule, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PVV1] Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PZE9] Cathepsin-like protease, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PZK7] uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7Q0S8] serine carboxypeptidase, putative uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7Q6I9]	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1				
		Hide Organism Filter								

COMMUNITY CHAT

▼ Schedule

Sunday, 5th June 2022

Time	Topic (and handout)	Instructor
6:00 - 9:00 pm	Registration Welcome dinner (provided) General introduction to omics Attendees introductions Group photo	GIGC, OH, MAM, DSR, SR

Monday, 6th June

2022

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	VectorBase general introduction: remit, data tools and resources	GIGC
9:30 am	Search: Finding genes (with Site Search and BLAST), building search strategies and visiting gene pages	GIGC
10:00 am	Coffee/Tea Break (provided)	
10:20 am	Search – Continued	GIG
11:00 am	Genome Browser I: Alignments and Comparative Genomics	DSR
12:00 m	Lunch (provided)	
1:00 pm	RNA sequence data analysis via Galaxy, Part I Uploading data (Group Exercise)	OSH
1:45 pm	Genome Browser II: Interpreting RNAseq data (Group Exercise)	DSR
2:30 pm	Coffee/Tea Break (provided)	
2:50 pm	Homology gene relationships via OrthoMCL DB	GIGC
4:00 pm - 5:00 pm	RNA sequence data analysis via Galaxy, Part II starting the workflow (Group Exercise)	OSH



**Monday, 6th June
2022**

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	VectorBase general introduction: remit, data tools and resources	GIGC
9:30 am	Search: Finding genes (with Site Search and BLAST), building search strategies and visiting gene pages	GIGC
10:00 am	Coffee/Tea Break (provided)	
10:20 am	Search – Continued	GIG
11:00 am	Genome Browser I: Alignments and Comparative Genomics	DSR
12:00 m	Lunch (provided)	
1:00 pm	RNA sequence data analysis via Galaxy, Part I Uploading data (Group Exercise)	OSH
1:45 pm	Genome Browser II: Interpreting RNAseq data (Group Exercise)	DSR
2:30 pm	Coffee/Tea Break (provided)	
2:50 pm	Homology gene relationships via OrthoMCL DB	GIGC
4:00 pm - 5:00 pm	RNA sequence data analysis via Galaxy, Part II starting the workflow (Group Exercise)	OSH

**Tuesday, 7th June
2022**

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	Sequence Exercises: Motifs, domains and colocation	GIGC
9:30 am	Data retrieval and download	GICG
10:00 am	Coffee/Tea Break (provided)	



**Tuesday, 7th June
2022**

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	Sequence Exercises: Motifs, domains and colocation	GIGC
9:30 am	Data retrieval and download	GIGC
10:00 am	Coffee/Tea Break (provided)	
10:50 am	RNA sequence data analysis via Galaxy, Part III: DESeq2 and Exporting your data to VectorBase FASTQ Groomer Trimmomatic FASTQC HISAT2 bamCoverage htseq-count HTSeqCountToTPM	OSH
12:00 pm	Lunch (provided)	
1:00 pm	RNA sequence data analysis via Galaxy, Part IV: viewing and analyzing your results, Bam to BigWig (Group Exercise)	OSH
2:30 pm	Coffee/Tea Break (provided)	
2:50 pm	Functional Genomics 1: Transcriptomics and Proteomics	GIGC
4:00 - 5:00 pm	Workshop group project, Part 1: Define your own questions and develop a method to answer them with VectorBase (Group exercise)	

**Wednesday, 8th June
2022**

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	Population data with MapVEu: population abundance, pathogen infection status, insecticide resistance genotypes and phenotypes, blood meal host, microsatellites and SNP data ... etc.	SR



**Wednesday, 8th June
2022**

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	Population data with MapVEu: population abundance, pathogen infection status, insecticide resistance genotypes and phenotypes, blood meal host, microsatellites and SNP data ... etc.	SR
10:00 am	Coffee/Tea Break (provided)	
10:20 am	Population data with the genome browser: SNPs	SR
11:00 am	RNA sequence data analysis Part V: finish data analyses	OSH
12:00 m	Lunch (provided)	
1:00 pm	Functional Genomics 2: GO Enrichment, Metabolic Pathways and Host response data	GIGC
2:00 pm	Workshop group project, Part 2: Data analyses (Group exercise) Start outline of presentation, you can do a live demo and/or a few slides	
3:00 pm	Coffee/Tea Break (provided)	
3:20 - 5:00 pm	Workshop group project, Part 3: Finalize project (Group exercise)	

**Thursday, 9th June
2022**

Time	Topic (and handout)	Instructor
8:00 am	Breakfast (provided)	
9:00 am	Manual gene annotation with Apollo	GIGC, RW
10:30 am	Coffee/Tea Break (provided)	
10:50 am	Workshop group project, Part 4: Present your findings	
12:00 - 12:15 pm	Workshop Feedback and wrap-up. Digital certificates will be provided.	



Arthropod Genomics 2022

IN PERSON AND VIRTUAL



VEuPathDB
Eukaryotic Pathogen, Vector & Host
Informatics Resources



National Institute of
Allergy and
Infectious Diseases



VectorBase
Bioinformatics Resource for
Invertebrate Vectors of Human Pathogens



 UNIVERSITY OF
NOTRE DAME



ECK INSTITUTE FOR
**GLOBAL
HEALTH**



 ILLINOIS
Carl R. Woese Institute
for Genomic Biology





Search for...

[expand all](#) | [collapse all](#)

Filter the searches below...

Genes

Organisms

Popset Isolate Sequences

Genomic Sequences

Genomic Segments

ESTs



News »

[VectorBase 57 Released](#)

THU APR 21 2022

[TrichDB 57 Released](#)

THU APR 21 2022

[TriTrypDB 57 Released](#)

THU APR 21 2022

[ToxoDR 57 Released](#)[See all news](#)

Tweets by @veupathdb

VEuPathDB Retweeted

 Ross Waller
@RossWaller3

Very pleased to present our recent study of #endocytosis in #apicomplexans that draws on #K13 and drug-resistance in #malaria to learn of a highly unusual endocytic machinery in #Toxoplasma with a deep evolutionary origin.
biorxiv.org/cgi/content/sh...

