

**Applications of Bioinformatics in Modern Molecular Biology**  
**PlasmoDB Workshop**  
**22<sup>nd</sup> October 2013, ICGEB, N. Delhi, India**

Workshop topics for discussion and hands on demonstration:

1. Introduction and highlights of the PlasmoDB website: What is available and how users can access and analyze genome datasets.
2. Searching for and retrieving gene(s) of interest: The concept of “Search Strategies”.  
[A] Text based searches.  
[B] Sequence based searches [blast / orthology].
3. My Basket: Saving searches, retrieving data for download, and graphing genomic views.
4. Genome browser (GBrowse): Viewing genes, DNA elements/motifs and functional genomics data mapping on a genomic context.
5. Identifying genes/proteins using functional genomics data: Gene/protein expression, phylogenetic profiling and orthology, genetic diversity, gene/protein function (including pathways), HTS datasets.
6. Hands-on exercise on the following:  
[A] Finding genes and exploring the gene page and running a blast.  
[B] Using the GBrowse.  
[C] HTS and population genetics data (Metadata and SNPs).  
[D] Functional genomics in EuPathDB: Transcriptomics and Proteomics.  
[E] Motif and domain searches.  
[F] Metabolic pathways and compounds.  
[G] Orthology and phyletic patterns.  
[H] Genomic colocation and other complex strategies
7. General discussion and Feedback on the workshop (Fill and submit the feedback form).
8. Live interaction *via* Skype with Dr. Omar Harb, EuPathDB Out Reach Manager, University of Pennsylvania, Philadelphia, USA.

## Introduction and highlights of the PlasmoDB website ([plasmodb.org](http://plasmodb.org)):

PlasmoDB is a functional genomics data resource for the various *Plasmodium* species. PlasmoDB belongs to a family of databases that fall under the EuPathDB (Eukaryotic Pathogen genome database) umbrella that includes diverse parasitic species belonging to *Amoeba*, *Cryptosporidium*, *Giardia*, *Microsporidia*, *Prioplasma*, *Toxoplasma*, *Trichomonas* and *Kinetoplastida*.

The home page of PlasmoDB has information about the DB itself, registration and login links, all search options, tutorial material, news regarding community events, publications and meetings, and contact information. The images below highlight these functions on the home page.

The screenshot shows the PlasmoDB home page. A red box highlights the PlasmoDB logo, which includes the text "Version 10.0 25 Sep 13". A yellow callout box points to the logo with the text: "The logo provides a quick link back to the home page. It also displays a database version number and data of release". The page features a navigation bar with links like Home, New Search, My Strategies, My Basket, Tools, Data Summary, Downloads, and Community. The main content area is divided into sections: News and Tweets, Data Summary, Identify Genes by, Identify Other Data Types, and Tools. The Tools section lists various analysis tools like BLAST, Sequence Retrieval, Pathogen Portal, PubMed and Entrez, and Genome Browser.

This screenshot shows the same PlasmoDB home page, but with a red box highlighting the search area. A yellow callout box points to the search boxes with the text: "Quick search boxes allow you to look for genes based on their IDs or a text term from anywhere in this database". The search area includes a "Gene ID:" field with the value "PF3D7\_1133400" and a "Gene Text Search:" field with the value "synth". The rest of the page layout is identical to the previous screenshot.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Login | Register | Contact Us

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

**Data Summary**

**News and Tweets**

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

All PlasmoDB News >>>

**Tweets** Follow

EuPathDB @eupathdb 9h

Gene ID search - get your list of genes in a strategy this example in ToxoDB retrieving a list from Walker et al. toxodb.org/toxo/im.do?s=b.

**Identify Genes by:**

- Expand All | Collapse All
- Text, IDs, Organism
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function
- Evolution
- Population Biology

**Identify Other Data Types:**

- Expand All | Collapse All
- Isolates
- Genomic Sequences
- Genomic Segments (DNA Motif)
- SNPs
- ESTs
- ORFs
- SAGE Tags
- Metabolic Pathways
- Compounds

**Tools:**

**BLAST**

- Identify Sequence Similarities

**Sequence Retrieval**

- Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**

- RNA sequence analysis, interactome maps and more

**PubMed and Entrez**

- View the Latest Pubmed and Entrez Results

**Genome Browser**

- View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above.....

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

Community Resources expand for 57 new items

Education and Tutorials

About PlasmoDB

You can also access helpful information, contact us, log into your account or register from the header section.

Registering and logging into your account allows you to access your previously run searches, add comments to gene or sequence pages, and use features like “My Basket” and “My Favorites” to save and retrieve sets of genes as shown below.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Login | Register | Contact Us

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

**Data Summary**

**News and Tweets**

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

All PlasmoDB News >>>

**Tweets** Follow

EuPathDB @eupathdb 10h

Gene ID search - get your list of genes in a strategy this example in ToxoDB retrieving a list from Walker et al. toxodb.org/toxo/im.do?s=b.

**Identify Genes by:**

- Expand All | Collapse All
- Text, IDs, Organism
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function
- Evolution
- Population Biology

**Identify Other Data Types:**

- Expand All | Collapse All
- Isolates
- Genomic Sequences
- Genomic Segments (DNA Motif)
- SNPs
- ESTs
- ORFs

**Tools:**

**BLAST**

- Identify Sequence Similarities

**Sequence Retrieval**

- Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**

- RNA sequence analysis, interactome maps and more

**PubMed and Entrez**

- View the Latest Pubmed and Entrez Results

**Genome Browser**

- View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above.....

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

Community Resources expand for 57 new items

Education and Tutorials

About PlasmoDB

**Account login**

Email:

Password:

☐ Remember me on this computer.

Login Cancel

Forgot Password? Register/Subscribe

Logging into your account allows you to access your previously run searches, add comments to gene or sequence pages, and use features like 'My Basket' and 'My Favorites'.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Login | Register | Contact Us

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

**Data Summary**

**News and Tweets**

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

All PlasmoDB News >>>

**Tweets** Follow

EuPathDB @eupathdb 10h  
Gene ID search - get your list of genes in a strategy this example in ToxoDB retrieving a list from Walker et al. toxodb.org/toxo/im.do?s=b.

**Tools:**

**BLAST**  
Identify Sequence Similarities

**Sequence Retrieval**  
Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**  
RNA sequence analysis, interactome maps and more

**PubMed and Entrez**  
View the Latest Pubmed and Entrez Results

**Genome Browser**  
View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above....

**Community Resources**  
expand for 57 new items

**Education and Tutorials**

**About PlasmoDB**

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

**EuPathDB**

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

**The tool bar provides access to new searches, the searches that you have already run, your basket, tools, data summary, file downloads, community resources and your favorites. Mousing over any of these items activates menus that allow you to navigate to your final destination.**

Navigating the TOOL BAR, and using pull down menus for “Text, IDs, Species” and “Transcript Expression” searches are shown below.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

**Data**

Search for Genes

Search for Isolates

Search for Genomic Sequences

Search for Genomic Segments (DNA Motif)

Search for SNPs

Search for ESTs

Search for ORFs

Search for SAGE Tags

Search for Metabolic Pathways

Search for Compounds

View all available searches

**Text, IDs, Organism**

Genomic Position

Gene Attributes

Protein Attributes

Protein Features

Similarity/Pattern

Transcript Expression

Protein Expression

Cellular Location

Putative Function

Evolution

Population Biology

**Text (product name, notes, etc.)**

Gene ID(s)

Organism

User Comments

Having Updated Annotation at GeneDB

Reagent Availability

**Data Types:**

ESTs

ORFs

SAGE Tags

Metabolic Pathways

Compounds

**Tools:**

**BLAST**  
Identify Sequence Similarities

**Sequence Retrieval**  
Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**  
RNA sequence analysis, interactome maps and more

**PubMed and Entrez**  
View the Latest Pubmed and Entrez Results

**Genome Browser**  
View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above....

**Community Resources**  
expand for 57 new items

**Education and Tutorials**

**About PlasmoDB**

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

**EuPathDB**

Please Contact Us with any questions or comments

POWERED BY Strategies WDK



The screenshot shows the PlasmoDB website interface. At the top, there's a header with the PlasmoDB logo, version information (10.0, 25 Sep 13), and a search bar for Gene ID (PF3D7\_1133400) and Gene Text Search (synth\*). Below the header is a navigation bar with links like Home, New Search, My Strategies, My Basket (0), Tools, Data Summary, Downloads, and Community. A side bar on the left contains links for Data, News, and Tweets. The main content area features a search menu with categories like Text, IDs, Organism, Genomic Position, Gene Attributes, Protein Attributes, Protein Features, Similarity/Pattern, Transcript Expression, Protein Expression, Cellular Location, Putative Function, Evolution, and Population Biology. A red arrow points to the 'EST Evidence' option under the 'Transcript Expression' category. Other sections include 'Identify Other Data Types' (Isolates, Genomic Sequences, Genomic Segments (DNA Motif), SNPs) and 'Tools' (BLAST, Sequence Retrieval, Pathogen Portal, PubMed and Entrez, Genome Browser). The footer includes 'Community Resources', 'Education and Tutorials', and 'About PlasmoDB'.

Data Summary, useful new items, tutorials, help menus and social media outreach can be accessed via the side bar highlighted below.

This screenshot highlights the side bar of the PlasmoDB website. A red box encloses the 'Data Summary', 'News and Tweets', 'Community Resources', 'Education and Tutorials', and 'About PlasmoDB' sections. A red text box with a yellow background is overlaid on the 'Data Summary' section, stating: 'The side bar section provides access to useful resources including a news section where you can find information about the current release or other important community related news.' The main content area shows the same search and navigation options as the previous screenshot, including the 'Identify Other Data Types' and 'Tools' sections.

All available tools and search options can be accessed from the “Identify Genes by”, “Identify Other Data Types” and “Tools” menu sections, which are highlighted below.

**PlasmoDB** Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

**Data Summary**

**News and Tweets**

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

**Tweets**

EuPathDB @eupathdb 10h

Gene ID search - get your list of genes in a strategy this example in ToxoDB retrieving a list from Walker et al. toxodb.org/toxo/im.do?s=b.

**Identify Genes by:**

- Expand All | Collapse All
- Text, IDs, Organism
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function
- Evolution
- Population Biology

**Identify Other Data Types:**

- Expand All | Collapse All
- Isolates
- Genomic Sequences
- Genomic Segments (DNA Motif)
- SNPs
- ESTs
- ORFs
- SAGE Tags
- Metabolic Pathways
- Compounds

**Tools:**

**BLAST**

- Identify Sequence Similarities

**Sequence Retrieval**

- Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**

- RNA sequence analysis, interactome maps and more

**PubMed and Entrez**

- View the Latest Pubmed and Entrez Results

**Genome Browser**

- View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above....

**Community Resources**

**Education and Tutorials**

**About PlasmoDB**

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

**EuPathDB**

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

## Searching for and retrieving gene(s) of interest:

Text and Gene ID based searches can be initiated from the tools and menus highlighted below. NOTE: Gene IDs can change between DB versions. Old IDs can be used for running searched but retrieved genes might list the new IDs. Mapping between the old and new IDs are available and can be downloaded.

**PlasmoDB** Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

**Data Summary**

**News and Tweets**

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

**Tweets**

EuPathDB @eupathdb 10h

Gene ID search - get your list of genes in a strategy this example in ToxoDB retrieving a list from Walker et al. toxodb.org/toxo/im.do?s=b.

**Identify Genes by:**

- Expand All | Collapse All
- Text, IDs, Organism
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function
- Evolution
- Population Biology

**Identify Other Data Types:**

- Expand All | Collapse All
- Isolates
- Genomic Sequences
- Genomic Segments (DNA Motif)
- SNPs
- ESTs
- ORFs
- SAGE Tags
- Metabolic Pathways
- Compounds

**Tools:**

**BLAST**

- Identify Sequence Similarities

**Sequence Retrieval**

- Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**

- RNA sequence analysis, interactome maps and more

**PubMed and Entrez**

- View the Latest Pubmed and Entrez Results

**Genome Browser**

- View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above....

**Community Resources**

**Education and Tutorials**

**About PlasmoDB**

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

**EuPathDB**

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

Consecutive steps in a text search are shown below.

STEP 1: Choose the text search menu highlighted by the red arrow below.

Version 10.0  
25 Sep 13

PlasmoDB Plasmodium Genomics Resource

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home | New Search | My Strategies | My Basket (0) | Tools | Community | My Favorites

**Data Summary**

**News and Tweets**

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

All PlasmoDB News >>>

**Tweets** Follow

EuPathDB @eupathdb 10h

Gene ID search - get your list of genes in a strategy this example in ToxoDB retrieving a list from Walker et al. toxodb.org/toxo/im.do?s=b.

**Community Resources**  
expand for 37 new items

**Education and Tutorials**

**About PlasmoDB**

**Identify Genes by:**

- Expand All | Collapse All
- Text (product name, notes, etc.)
- Gene ID(s)
- Organism
- User Comments
- Having Updated Annotation at GeneDB
- Reagent Availability
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function
- Evolution
- Population Biology
- Isolates
- Genomic Sequences
- Genomic Segments (DNA Motif)
- SNPs
- ESTs
- ORFs
- SAGE Tags
- Metabolic Pathways (BETA)
- Compounds (BETA)

**Tools:**

**BLAST**  
Identify Sequence Similarities

**Sequence Retrieval**  
Retrieve Specific Sequences using IDs and coordinates

**Pathogen Portal**  
RNA sequence analysis, interactome maps and more

**PubMed and Entrez**  
View the Latest Pubmed and Entrez Results

**Genome Browser**  
View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above.....

PlasmoDB 10.0 25 Sep 13  
©2013 The EuPathDB Project Team

EuPathDB

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

STEP 2: Choose the relevant organisms (1), enter the text search term (2), select the data type to display for each gene (3), run search (4), as shown below.

Version 10.0  
25 Sep 13

PlasmoDB Plasmodium Genomics Resource

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home | New Search | My Strategies | My Basket (0) | Tools | Data Summary | Downloads | Community | My Favorites

**Identify Genes based on Text (product name, notes, etc.)**

**Organism** select all | clear all | expand all | collapse all | reset to default

- ☐ Plasmodium berghei
- ☐ Plasmodium chabaudi
- ☐ Plasmodium cynomolgi
- ☒ Plasmodium falciparum
- ☒ Plasmodium falciparum 3D7
- ☒ Plasmodium falciparum IT
- ☐ Plasmodium gallinaceum
- ☒ Plasmodium knowlesi
- ☒ Plasmodium knowlesi strain H
- ☐ Plasmodium reichenowi
- ☒ Plasmodium vivax
- ☒ Plasmodium vivax Sal-1
- ☐ Plasmodium yoelii

select all | clear all | expand all | collapse all | reset to default

**Text term (use \* as wildcard)** Kinase

**Fields**

- ☒ Gene ID
- ☒ Alias
- ☒ Gene product
- ☒ Genes of previous release
- ☒ Rodent Malaria Phenotype
- ☒ GO terms and definitions
- ☒ Gene notes
- ☒ User comments
- ☒ Protein domain names and descriptions
- ☐ Similar proteins (BLAST hits v. NRDB/PDB)
- ☒ EC descriptions
- ☒ Metabolic pathway names and descriptions

select all | clear all

**Advanced Parameters**

Get Answer

Give this search a name

STEP 3: View results of the text search on the result page showing a list of genes fulfilling the search criteria. In this example, a search for the term 'Kinase' in any *Plasmodium falciparum*, *P. knowlesi*, and *P. vivax* returned **720 genes** as shown.

**720 Genes from Step 1**  
Strategy: Text

**Filter results by species** (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	<i>Plasmodium falciparum</i>		<i>Plasmodium vivax</i>	<i>Plasmodium yoelii</i>			<i>Plasmodium berghei</i>	<i>Plasmodium chabaudi</i>	<i>Plasmodium cynomolgi</i>	<i>Plasmodium gallinaceum</i>	<i>Plasmodium knowlesi</i>	<i>Plasmodium reichenowi</i>
		Distinct genes	3D7		IT	Distinct genes	yoelii 17XNL						
720	216	208	210	181	169	0	0	0	0	0	0	160	0

Gene Results | Genome View

First 1 2 3 4 5 Next Last | Advanced Paging | Add Columns

Gene ID	Organism	Genomic Location	Product Description	Found in	Score
PKH_131880	<i>P. knowlesi</i> strain H	Pk_strainH_chr13: 937,698 - 939,776 (+)	mitogen-activated protein kinase 1, MAP-kinase 1, putative	InterPro, Product, GoTerms, user comments	80
PKH_072870	<i>P. knowlesi</i> strain H	Pk_strainH_chr07: 1,305,282 - 1,309,597 (-)	diacylglycerol kinase, putative	InterPro, Product, GoTerms, EcNumber	72
PVX_099990	<i>P. vivax</i> Sal-1	Pv_Sal1_chr07: 1,168,881 - 1,173,406 (-)	diacylglycerol kinase, putative	InterPro, Product, GoTerms, EcNumber	72

The red circles highlight the number of genes from an individual species and clicking on this number will display only those genes, as shown below for *P. falciparum*.

**210 Genes from Step 1**  
Strategy: Text

**Filter results by species** (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	<i>Plasmodium falciparum</i>		<i>Plasmodium vivax</i>	<i>Plasmodium yoelii</i>			<i>Plasmodium berghei</i>	<i>Plasmodium chabaudi</i>	<i>Plasmodium cynomolgi</i>	<i>Plasmodium gallinaceum</i>	<i>Plasmodium knowlesi</i>	<i>Plasmodium reichenowi</i>
		Distinct genes	3D7		IT	Distinct genes	yoelii 17XNL						
720	216	208	210	181	169	0	0	0	0	0	0	160	0

Gene Results | Genome View

First 1 2 3 Next Last | Advanced Paging | Add Columns

Gene ID	Organism	Genomic Location	Product Description	Found in	Score
PF3D7_1106800	<i>P. falciparum</i> 3D7	PF3D7_11_v3: 283,869 - 289,224 (+)	protein kinase, putative	InterPro, Product, GoTerms, PreviousReleaseGenes	64
PF3D7_0213400	<i>P. falciparum</i> 3D7	PF3D7_02_v3: 541,811 - 543,807 (+)	protein kinase 7 (PK7)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber	56
PF3D7_0217500	<i>P. falciparum</i> 3D7	PF3D7_02_v3: 720,437 - 722,661 (+)	calcium-dependent protein kinase 1 (CDPK1)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber, Prev...	56
PF3D7_0515300	<i>P. falciparum</i> 3D7	PF3D7_05_v3: 628,981 - 635,382 (+)	phosphatidylinositol 3-kinase (PI3K)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber	56
PF3D7_0821100	<i>P. falciparum</i> 3D7	PF3D7_08_v3: 953,448 - 956,314 (+)	protein kinase 1 (PK1)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber, Note...	56



From this result, the gene of interest can be identified (as shown by the red star) and all available data for this gene can be viewed on the gene page, which can be reached by clicking the corresponding Gene ID (red circle). The gene page for this example is shown below (details of most datasets hidden to reduce image size).

**PlasmoDB** Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shanmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (0) Tools Data Summary Downloads Community My Favorites

---

**Annotation** Protein Expression Sequence

**PF3D7\_0217500**  
calcium-dependent protein kinase 1 (CDPK1)

Previous ID(s): PF02\_0166, PFB0815w  
This gene has 2 user comments Add to Basket Add to Favorites  
[View updated annotation at GeneDB](#)

**Updated product name(s) from GeneDB: calcium-dependent protein kinase 1**

NOTE: These genomes are being resequenced. The sequence and annotation are being actively updated and curated, and are unpublished. Please consult with the Principal Investigator before undertaking large scale analyses of the annotation or underlying sequence.

**Overview**  
*P. falciparum* 3D7 protein coding gene on PF3D7\_02\_v3 from 720,437 to 722,661 (Chromosome: 2) [\[Data Sets\]](#)

**Genomic Context** [Hide](#)

Genomic context of the gene can be further explored by clicking this link [View in Genome Browser](#)  
(use right click or ctrl-click to open in a new window)

**SNP Overview**  
Total SNPs: 61  
Total Non-Synonymous SNPs: 6  
Total Synonymous SNPs: 55  
Non-Synonymous/Synonymous Ratio: 0.11 [Genetic diversity data](#)

**SNPs Alignment** [Show](#)

Regions/Spans associated by eQTL experiment on HB3 x DD2 progeny (LOD cut off = 1.5) [Hide](#)

Haplotype Block	Genomic Segment (Liberal)	Genomic Segment (Conservative)	LOD Score (opens a haplotype plot)	Search for Genes (Liberal by Default)	Search for Genes (Liberal by Default)
PF3D7_10_v3_80.2	PF3D7_10_v3:1336744-1398644	PF3D7_10_v3:1355962-1356107	1.94	<a href="#">Genes Contained in this Region</a>	<a href="#">Genes Associated to this Region</a>

Other genes that have similar associations based on eQTL experiments  
(use right click or ctrl-click to open in a new window)

**Comparison to previous release** [Hide](#) [\[Data Sets\]](#)

Previous Gene ID	Annotation type	Value
PFB0815w	product	calcium dependent protein kinase 1

**Multiple Sequence Alignment** [Show](#)

**Annotation** [Back to the Top](#)

**Add a comment on PF3D7\_0217500** [Show](#)

**User Comments** [Hide](#)

Comment ID	Headline	PubMed ID(s)	# Related Genes	# Uploaded Files	Made by	Date
138420	Calcium Dependent Protein Kinase 1	19307175	N/A	N/A	Ravikant Ranjan, National Instt of Immunology	2012-09-20 15:29:40.0
55893	Northern blot data reported. Transcript size of 3.2kb	23601558,8440720	N/A	N/A	Paul Horrocks, Keele University Medical School	2013-04-30 20:52:38.0

The Plasmodium falciparum genome is not finished. Please consult Plasmodium orthologs to support your conclusions.

**EC Number** [Show](#) [\[Data Sets\]](#)

**Metabolic Pathway Reactions** none [\[Data Sets\]](#)

**External Links** [Show](#) [\[Data Sets\]](#)

**Orthologs and Paralogs within PlasmoDB** [Show](#) [\[Data Sets\]](#)

**GO Terms** [Show](#) [\[Data Sets\]](#)

**Names, Previous Identifiers, and Aliases** [Show](#) [\[Data Sets\]](#)

**Notes** [Show](#) [\[Data Sets\]](#)

**Phenotype** [Show](#) [\[Data Sets\]](#)

**Curated Metabolic Pathways - Hagal Ginsburg** [Show](#)

**PlasmoCyc View**

**MR4 Reagents** none

Various gene annotation datasets and relevant external links. Full details can be viewed by clicking the 'Show' button highlighted by the arrow



[illegible]

The list of genes obtained as results (as shown above), can be saved and/or exported as depicted below.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

Home | New Search | My Strategies | My Basket (0) | Tools | Data Summary | Downloads | Community

My Strategies: New | Opened (1) | All (1) | Basket | Examples | Help

(Genes) Did You Know... 1 of 4  
...you can click Add Step to add a step to your strategy. Learn more...  
Add Step

Save search results by adding all the genes listed below to your 'Basket'

Download selected datasets for all genes listed below

210 Genes from Step 1  
Strategy: Text

Filter results by species (results removed by the filter will not be combined into the next step)

All Results	Ortholog Groups	Plasmodium falciparum	Plasmodium vivax	Plasmodium yoelii	Plasmodium berghei	Plasmodium chabaudi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium knowlesi	Plasmodium reichenowi
720	216	208	210	181	169	0	0	0	160	0

Click basket logo to select genes and save to gene basket

Gene ID	Genomic Location	Product Description	Found in	Score
PF3D7_1106800	P. falciparum 3D7 PI3D7_11_v3: 283,869 - 289,224 (+)	protein kinase, putative	InterPro, Product, GoTerms, PreviousReleaseGenes	64
PF3D7_0213400	P. falciparum 3D7 PI3D7_02_v3: 541,811 - 543,807 (+)	protein kinase 7 (PK7)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber	56
PF3D7_0217500	P. falciparum 3D7 PI3D7_02_v3: 720,437 - 722,661 (+)	calcium-dependent protein kinase 1 (CDPK1)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber, Prev...	56
PF3D7_0515300	P. falciparum 3D7 PI3D7_05_v3: 628,981 - 635,382 (+)	phosphatidylinositol 3-kinase (PI3K)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber	56
PF3D7_0821100	P. falciparum 3D7 PI3D7_08_v3: 953,448 - 956,314 (+)	protein kinase 1 (PK1)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber, Note...	56
PF3D7_0930500	P. falciparum 3D7 PI3D7_09_v3: 1,212,169 - 1,216,500 (-)	diacylglycerol kinase, putative (DGK1)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber	56
PF3D7_0934800	P. falciparum 3D7 PI3D7_09_v3: 1,361,935 - 1,363,948 (+)	cAMP-dependent protein kinase catalytic subunit (PKAc)	InterPro, Product, MetabolicPathways, GoTerms, EcNumber, user ...	56
PF3D7_1136500.1	P. falciparum 3D7 PI3D7_11_v3: 1,431,881 - 1,435,170 (-)	casein kinase 1 (CK1)	InterPro, Product, GoTerms, EcNumber, Notes	56

Add 210 Genes to Basket | Download 210 Genes

The download page options are shown below.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

Home | New Search | My Strategies | My Basket (210) | Tools | Data Summary | Downloads | Community

Download 210 Genes from the search:  
Text (product name, notes, etc.)

Please select a format from the dropdown list to download your search result. Select columns to include in the report. Optionally (see below) include a first line with column names.

Select a format:   
☒ Tab delimited (Excel): choose from columns  
☐ Configurable FASTA  
☐ GFF3: Gene models and optional sequences  
 Text: choose from columns and/or tables  
 XML: choose from columns and/or tables  
 json: choose from columns and/or tables

Data can be downloaded in different formats as chosen by user. In this case a spreadsheet format is chosen.

Columns

clear all | expand all | collapse all  
reset to current | reset to default

- ☒ Search-Specific
  - ☒ Found in
  - ☒ Score
- ☒ Text, IDs, Species
  - ☐ Genomic Sequence ID
  - ☒ Organism
  - ☐ Previous ID(s)
- ☒ Genomic Position
  - ☐ Chromosome
  - ☒ Genomic Location
  - ☐ Gene Strand
- ☐ Gene Attributes
- ☒ Protein Attributes
  - ☒ Product Description
  - ☐ Molecular Weight
  - ☐ Isoelectric Point
- ☐ Protein Features
  - ☒ Transcript Expression
  - ☒ Putative Function
- ☐ Evolution
  - ☒ Population Biology
  - ☐ Primary Seq
  - ☐ Primary Seq Start
  - ☐ Primary Seq End
  - ☐ Is Annotated
  - ☐ Release Policy
  - ☐ Predicted Protein Sequence
  - ☐ Predicted RNA/mRNA Sequence (introns spliced out)
  - ☒ No. of Pathways
  - ☐ Coding Sequence
  - ☐ Updated Sequence
  - ☐ Search Weight

clear all | expand all | collapse all  
reset to current | reset to default

Column names: ☒ include ☐ exclude

Download Type: ☐ Text File ☐ Excel File ☒ Show in Browser

Download options

\*\*Note: If you choose "Excel File" as Download Type, you can only download a maximum 10M (in bytes) of the results and the rest will be discarded. Opening a huge Excel file may crash your system. If you need to get the complete results, please choose "Text File" or "Show in Browser".

Get Report

PlasmoDB 10.0 25 Sep 13  
©2013 The EuPathDB Project Team

EuPathDB

Please Contact Us with any questions or comments  
Strategies WDX

The list of genes saved to 'Basket' can be retrieved, modified or exported anytime in the future. NOTE: To be able to save your search results in 'Basket' you need to be logged in. An example 'Basket' page is shown below.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*  
 About PlasmoDB | Help | Dhanasekaran Shanmugam's Profile | Logout | Contact Us

Home | New Search | My Strategies | My Basket (0) | Tools | Data Summary | Downloads | Community

My Strategies: New | Opened (1) | All (1) | **Basket** | Examples | Help

Gene (210)  
 Refresh | Empty basket | Save basket to a strategy

**210 Genes**  
 Gene Results | Genome View | **Basket Page options**

First 1 2 3 Next Last Advanced Paging Add Columns

Gene ID	Genomic Location	Product Description
PF3D7_0102900	PI3D7_01_v3: 119,041 - 121,249 (-)	serine/threonine protein kinase, FIKK family (FIKK1)
PF3D7_0103700	PI3D7_01_v3: 166,513 - 168,887 (+)	L-seryl-tRNA(Sec) kinase, putative (PSTK)
PF3D7_0105900	PI3D7_01_v3: 249,231 - 252,010 (-)	conserved Plasmodium protein, unknown function
PF3D7_0107600	PI3D7_01_v3: 314,618 - 319,405 (+)	serine/threonine protein kinase, putative
PF3D7_0109000.1	PI3D7_01_v3: 383,284 - 387,940 (+)	phosphatidate cytidyltransferase, putative
PF3D7_0109000.2	PI3D7_01_v3: 383,284 - 387,940 (+)	phosphatidate cytidyltransferase, putative
PF3D7_0110600	PI3D7_01_v3: 405,739 - 412,169 (+)	phosphatidylinositol-4-phosphate 5-kinase (PIPSK)

Note on new releases: IDs sometimes change or are retired. Why? Old IDs are mapped to new IDs when possible. Retired IDs will not be in the basket. To keep a copy of your current basket please download your IDs now.

Download 210 Genes

**TIP:** Genes saved to 'Basket' can be graphed as chromosomal views based on their genomic locations, thus giving the user an instantaneous genomic view of their search results. These graphs can be useful as publication material. The genomic view of results from the above search example (210 genes from 'kinase' text search in *P. falciparum*) is shown below.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*  
 About PlasmoDB | Help | Dhanasekaran Shanmugam's Profile | Logout | Contact Us

Home | New Search | My Strategies | My Basket (0) | Tools | Data Summary | Downloads | Community

My Strategies: New | Opened (1) | All (1) | **Basket** | Examples | Help

Gene (210)  
 Refresh | Empty basket | Save basket to a strategy

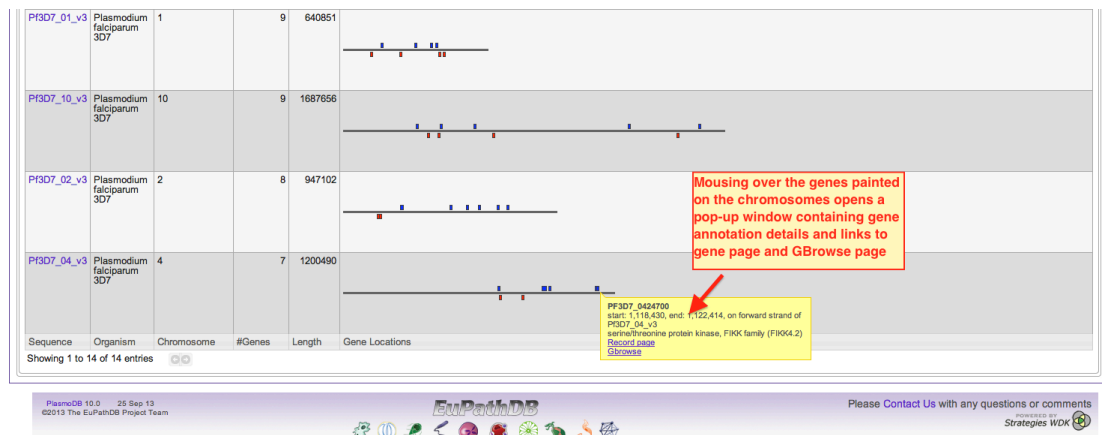
**210 Genes**  
 Gene Results | **Genome View** | Click for genomic view of genes present in Basket

Showing 1 to 14 of 14 entries Show 25 entries Search:

Sequence Organism Chromosome #Genes Length Gene Locations

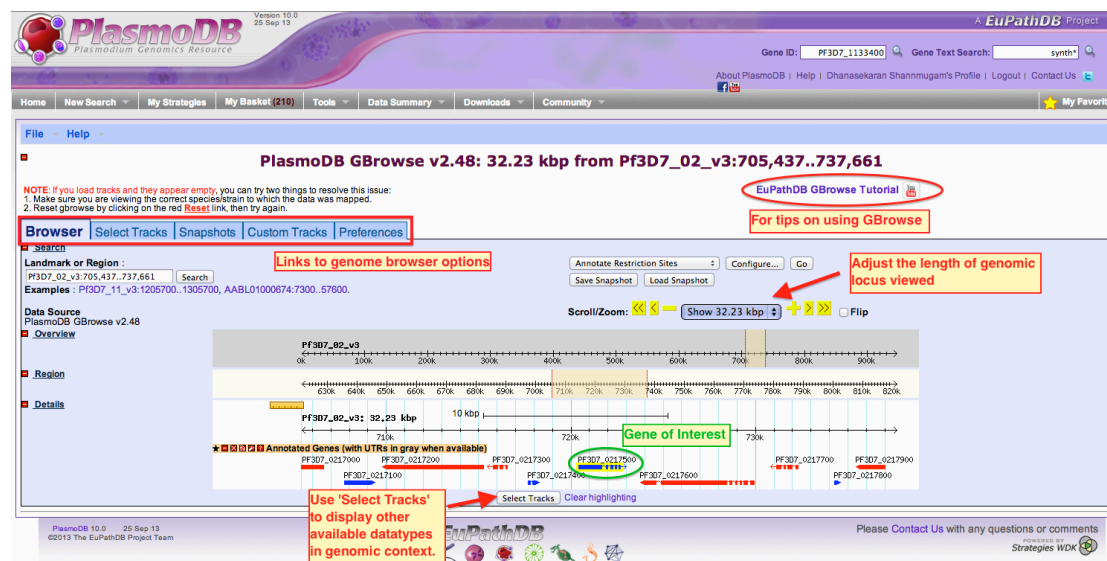
PF3D7_14_v3	Plasmodium falciparum 3D7	14	29	3291936	
PF3D7_11_v3	Plasmodium falciparum 3D7	11	24	2038340	
PF3D7_13_v3	Plasmodium falciparum 3D7	13	24	2925236	
PF3D7_09_v3	Plasmodium falciparum 3D7	9	21	1541735	
PF3D7_12_v3	Plasmodium falciparum 3D7	12	19	2271494	
PF3D7_07_v3	Plasmodium falciparum 3D7	7	13	1445207	
PF3D7_03_v3	Plasmodium falciparum 3D7	3	12	1067971	
PF3D7_05_v3	Plasmodium falciparum 3D7	5	12	1343557	
PF3D7_06_v3	Plasmodium falciparum 3D7	6	12	1418242	
PF3D7_08_v3	Plasmodium falciparum 3D7	8	11	1472805	

Legend:   
 Genes on forward strand: (blue bar)   
 Genes on reverse strand: (red bar)



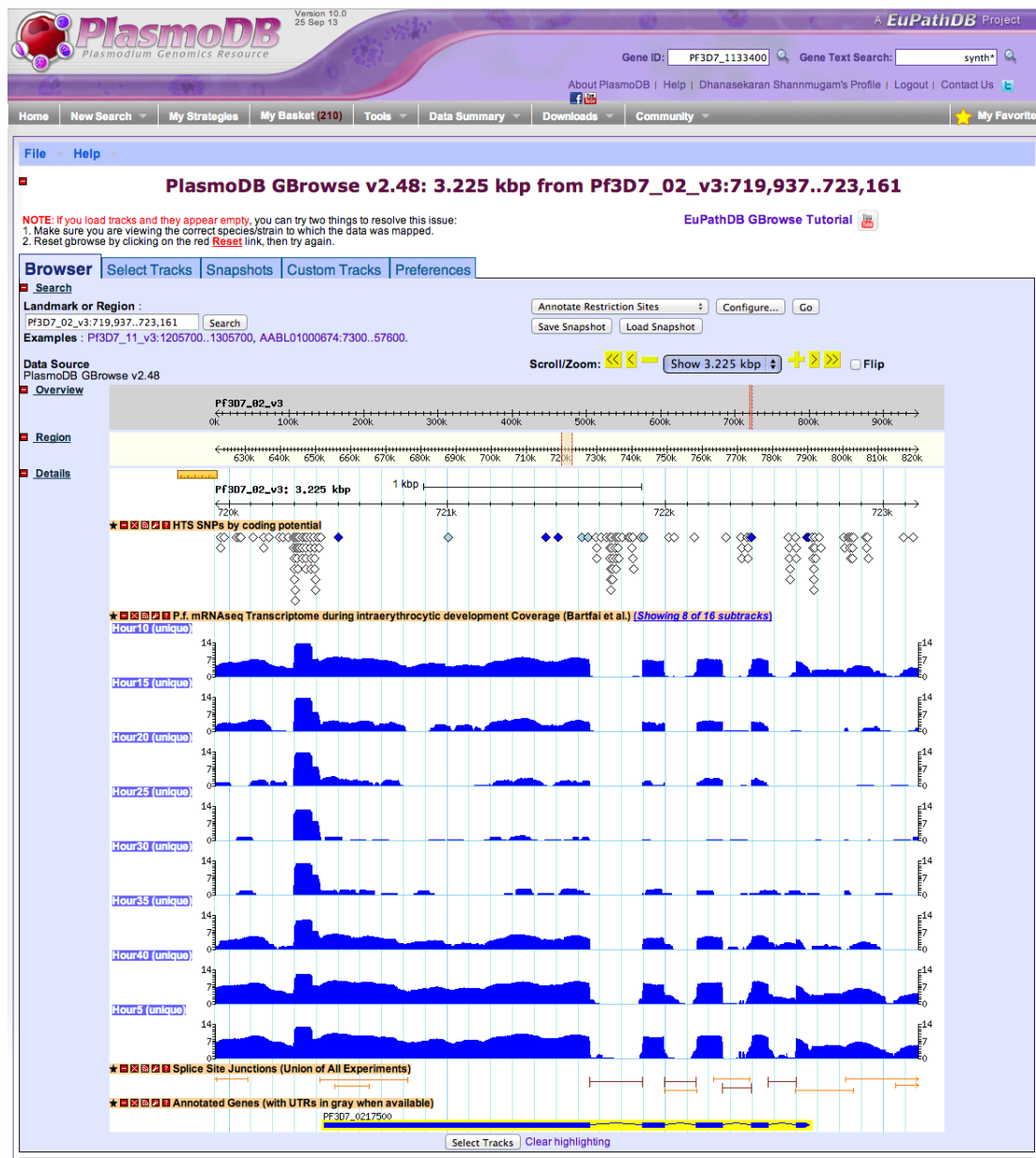
A magnified view of the genomic location for the gene of interest can be viewed along with other available genome wide data types using the GBrowse (Genome Browser) tool. The GBrowse page can be accessed by either clicking on the GBrowse link in the pop-up window from the genomic view graphs or from the individual gene pages.

An example GBrowse view is shown below.



Use 'Select Tracks' option to select and display other available datatypes in genomic context. Some of this data can be gene specific (e.g., microarray probe sequence, splice site junctions from RNA Seq data, peptide data from mass spectrometry) while others are genome wide datasets (e.g., position of restriction sites, SNP data, non-coding RNA mapping).

The size of the genomic fragment viewed can be adjusted using the scaling tool highlighted above. More easily, the genomic context for only the gene of interest can be displayed simply by clicking on the GBrowse link in the pop-up window that comes up when mousing over the gene. The magnified GBrowse view for the gene highlighted in green above along with data for splice site junctions, temporal RNA Seq data during intra-erythrocytic development, and SNP data is depicted below.



## Blast searches using DNA and protein sequences:

Thus far we have seen how gene associated text terms can be used for searching and retrieving genes. Now we will see how to run sequence based searches using the blast tool to find genes as well as non-coding regions of interest.

The blast tool is easily accessed from the home page of PlasmoDB by clicking on the 'Blast' link under the 'Tools' menu bar. The blast page allows the users to select the data type that has to be searched (transcripts, proteins, genomic sequence, etc.), the program to use, the organism on which the search must be performed, enter query sequence in a text box, and modify search parameters.



In the search example depicted, a *P. falciparum* FIKK1 kinase protein sequence is used to find out how many members of this family are present and are there any related kinases in any other *Plasmodium* species.

The screenshot shows the PlasmoDB search interface with several red annotations highlighting key steps:

- Choose the datatype to run the search:** Points to the 'Target Data Type' section where 'Proteins' is selected.
- Choose correct blast program:** Points to the 'BLAST Program' section where 'blastp' is selected.
- Choose species of interest:** Points to the 'Target Organism' section where multiple *Plasmodium* species are checked.
- Enter the query sequence in the text box:** Points to the 'Input Sequence' text area containing the FIKK1 protein sequence.
- Modify parameters if required:** Points to the 'Expectation value' and 'Maximum descriptions/alignments (V=B)' fields.
- Run the search:** Points to the 'Get Answer' button.

The 'Input Sequence' text area contains the following protein sequence:

```
MLNMIQKKKNLFLRSILIKFIHISFLGILYFFYNVEVLE
DKETNNLELIVNESRYLAES
RKDLKKGKDKPNINVDNFCVRCGSKNKCINSEER
RCNNCEGISKLFVCGCTKGLKFFWNK
ISCDPLCIEKNDSDNNYNNNNIQQNTCSVND
LLKNEFEIKDNNVKEVDDHISRD
NKNVCYNDNNMKKEDSKITFTDILGRLGKLLDS
SINFSINGVKYDDWETTPITCCA
SRVMEKQCKMYKVKPKKEDDKVDNKKMGNCN
DCDKIKGKDNNEEGDEIKLFMKKVP
```

The 'Expectation value' is set to 10, and 'Maximum descriptions/alignments (V=B)' is set to 50. The 'Low complexity filter' is set to 'no'.

The 'Advanced Parameters' section is collapsed. Below the search form, there is a 'Description' section with the following text:

Find genes, genomic sequences, ESTs, Isolates, or ORFs that have BLAST similarity to your input sequence. The BLAST input (nucleotide) sequence can be up to 2MB long. The search uses WU BLAST.

The blast result shows that the *P. falciparum* FIKK1 gene has duplicated and expanded in this organism while the other *Plasmodium* species have only one copy of the gene.

The blast result page (shown below) contains the list of genes with blast hit values and the alignments for the hits. As already discussed, from this page, genes can be viewed (by clicking on gene ID, exported out and/or saved to 'Basket'. The result tab allows the users to either view the blast result itself, or just the list of genes, or the graphical genomic view of the results. From the graphical genomic view it appears that 19 out of 22 FIKK kinases are present at chromosomal ends and each chromosome has at least one copy of the gene while chromosome 9 has 7 copies of the gene.

Version 10.0  
25 Sep 13

PlasmoDB  
Plasmodium Genomics Resource

A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New Opened (1) All (3) Basket Examples Help

(Genes) Strategy: BLAST(2) \* Rename Duplicate Save As Share Delete

Did You Know... 3 of 4  
...you can get a menu of step actions by clicking on the plus icon, step name or the Venn diagram. Learn more...  
<< Previous Next >> ☐ Never show me this again Close

BLAST 22 Genes Step 1 Add Step

22 Genes from Step 1 Strategy: BLAST(2) Expanded in *P. falciparum*..... Add 22 Genes to Basket | Download 22 Genes

Filter results by species (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	Plasmodium falciparum		Plasmodium vivax	Plasmodium yoelii		Plasmodium berghei	Plasmodium chabaudi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium knowlesi	Plasmodium reichenowi
		Distinct genes	3D7 MT		Distinct genes	yoelii 17XNL						
50	17	22	21	1	1	1	1	1	1	0	1	0

BLAST Gene Results Genome View

BLASTP 2.OMP-WashU [Linux26-i786-ILP32F6] [linux26-i786-ILP32F6]..... but single copy in other Plasmodium species

Copyright (C) 1996-2006 Washington University, Saint Louis, Missouri USA.  
All Rights Reserved  
Click for genomic view  
Reference: Gish, W. (1996-2006) http://blast.wustl.edu

Query= MySeq1 (630 letters)

Database: /var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pb/ergheiANKA/blast/Pb/ergheiANKAAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pc/habaudichabaudi/blast/Pc/habaudichabaudiAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pc/ynomolgiB/blast/Pc/ynomolgiBAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pf/alciiparum3D7/blast/Pf/alciiparum3D7AnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pf/alciiparum17/blast/Pf/alciiparum17AnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pg/allinaceum8A/blast/Pg/allinaceum8AAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pk/novlesiiB/blast/Pk/novlesiiBAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pr/eichenowiDennis/blast/Pr/eichenowiDennisAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Pv/ivaxSall/blast/Pv/ivaxSallAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Py/oelliiyoelii17XNL/blast/Py/oelliiyoelii17XNLAnnotatedProteins;  
/var/www/Common/apiSiteFilesMirror/webServices//PlasmoDB/build-19/Py/oelliiyoeliiYM/blast/Py/oelliiyoeliiYMANnotatedProteins  
51,195 sequences; 33,248,649 total letters.

Searching.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100% done

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Sum P(N)	Probability N		
				organism=Plasmodium_falciparum_3D7	prod...
PF3D7_0102600	organism=Plasmodium_falciparum_3D7	prod...	933	4.9e-114	3
PF3D7_0301200	organism=Plasmodium_falciparum_3D7	prod...	907	2.6e-113	2
PF3D7_0500000	organism=Plasmodium_falciparum_3D7	prod...	1021	3.1e-112	2
PF3D7_0805700	organism=Plasmodium_falciparum_3D7	prod...	867	1.1e-106	2
PF3D7_1200800	organism=Plasmodium_falciparum_3D7	prod...	859	1.2e-104	2
PF3D7_0424500	organism=Plasmodium_falciparum_3D7	prod...	866	1.0e-103	2
PF3D7_1149300	organism=Plasmodium_falciparum_3D7	prod...	851	1.0e-103	2
PF3D7_1476400	organism=Plasmodium_falciparum_3D7	prod...	892	2.6e-103	3
PF3D7_0902500	organism=Plasmodium_falciparum_3D7	prod...	850	2.7e-103	2
PF3D7_0902300	organism=Plasmodium_falciparum_3D7	prod...	697	1.9e-100	2
PF3D7_1039000	organism=Plasmodium_falciparum_3D7	prod...	968	2.7e-98	1
PF3D7_0902400	organism=Plasmodium_falciparum_3D7	prod...	952	1.3e-96	1
PF3D7_1371700	organism=Plasmodium_falciparum_3D7	prod...	584	9.6e-94	3
PF3D7_0902600	organism=Plasmodium_falciparum_3D7	prod...	588	4.7e-93	3

Result for *P. falciparum* FIKK1 protein blast

## Identifying gene/proteins using functional genomics data and the 'Search Strategy' concept:

As already discussed, a variety of functional genomics datasets, such as expression, putative function, subcellular localization, phylogeny and population biology, are available for querying in PlasmoDB. These search attributes can be used independently of each other or in combination using the search strategy. For example, a combination search for all metabolic enzymes expressed during intra-erythrocytic development is depicted below as consecutive steps.

STEP 1: Find all metabolic enzymes from *P. falciparum* using EC number data under the putative function search option.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

### Data Summary

#### News and Tweets

- 25 September 2013 PlasmoDB 10.0 Released
- 10 March 2013 PlasmoDB 9.3 Released
- 20 January 2013 EuPathDB Annual Workshop

All PlasmoDB News >>>

#### Tweets

Follow @eupathdb

Gene ID search - get your list

Find genes that have been annotated as enzymes, i.e. that have been assigned an Enzyme Commission (EC) number.

toxodb.org/toxo/rm.do?rs=0.

#### Identify Genes by:

- Expand All | Collapse All
- Text, IDs, Organism
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function description
- GO Term
- EC Number
- Metabolic Pathway
- Y2H Protein Interaction
- Predicted Functional Interaction
- Evolution
- Population Biology

#### Identify Other Data Types:

- Expand All | Collapse All
- Isolates
- Genomic Sequences
- Genomic Segments (DNA Motif)
- SNPs
- ESTs
- ORFs
- SAGE Tags
- Metabolic Pathways BETA
- Compounds BETA

#### Tools:

- BLAST**
  - Identify Sequence Similarities
- Sequence Retrieval**
  - Retrieve Specific Sequences using IDs and coordinates
- Pathogen Portal**
  - RNA sequence analysis, interactome maps and more
- PubMed and Entrez**
  - View the Latest Pubmed and Entrez Results
- Genome Browser**
  - View Sequences and Features in the genome browser

For additional tools, use the Tools menu in the gray toolbar above.....

#### Community Resources

expand for 36 new items

#### Education and Tutorials

#### About PlasmoDB

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

EuPathDB

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

### Identify Genes based on EC Number

Organism select all | clear all | expand all | collapse all | reset to default

- ☐ Plasmodium berghei
- ☐ Plasmodium chabaudi
- ☒ Plasmodium falciparum
- ☒ Plasmodium falciparum 3D7
- ☐ Plasmodium falciparum IT
- ☐ Plasmodium knowlesi
- ☐ Plasmodium vivax
- ☐ Plasmodium yoelii

EC Number or Term \*

Type three characters to see suggestions. Or use \* as a wildcard, like this: "you-term"

Advanced Parameters

Get Answer

Pfal Enz

#### Description

Find *P. falciparum*, *P. yoelii* and/or *P. knowlesi* genes by their Enzyme Commission (EC) number. Use '-' in place of numbers to denote partial EC Numbers, for example, '2.7.7.-'. You may use '\*' as a wild card character. For help with obtaining EC numbers, try the [EXPASY-ENZYME](#) browser.

E.C. number annotation for *P. falciparum* are from Andy Berry, Matt Berriman (Sanger Institute) and Hagai Ginsburg (Hebrew University).

#### Data Sets used by this search

- Enzyme Nomenclature Database**
  - EC Numbers from ENZYME
    - The ENZYME database in 2000. Bairoch et al. Nucleic Acids Res. 2000 Jan 1;28(1):304-5

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

EuPathDB

Please Contact Us with any questions or comments

Strategies WDK

The results show that there are 1174 genes in *P. falciparum* mapped with at least one EC number. Now, the next step is to run a query for gene expression and combine the two results. This can be done using the "My Strategies" section located at the top of the search result page (highlighted below).

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New Opened (1) All (4) Basket Examples Help

(Genes) Strategy: Pfal Enz \* 1 of 4

Did You Know... 1 of 4

...you can click Add Step to add a step to your strategy. [Learn more...](#)

<< Previous Next >> ☐ Never show me this again Close

Rename Duplicate Save As Share Delete

Pfal Enz 1174 Genes Step 1 Add Step

1174 Genes from Step 1 Strategy: Pfal Enz Add 1174 Genes to Basket | Download 1174 Genes

Filter results by species (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	Plasmodium falciparum		Plasmodium vivax	Plasmodium yoelii			Plasmodium berghei	Plasmodium chabaudi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium knowlesi	Plasmodium reichenowi
		Distinct genes	3D7		IT	Distinct genes	yoelii 17XNL						
1174	1063	1148	1174	0	0	0	0	0	0	0	0	0	0

Gene Results Genome View

First 1 2 3 4 5 Next Last

Number of genes identified as enzymes based on EC number

Gene ID	Organism	Genomic Location	Product Description	EC Numbers
mal_mito_1	P. falciparum 3D7	M76611: 734 - 1,486 (-)	cytochrome c oxidase subunit 3 (cox3)	1.9.3.1 (Cytochrome-c oxidase)
mal_mito_2	P. falciparum 3D7	M76611: 2,035 - 3,471 (+)	cytochrome c oxidase subunit I (cox1)	1.9.3.1 (Cytochrome-c oxidase)
mal_mito_3	P. falciparum 3D7	M76611: 3,492 - 4,622 (+)	apocytochrome b (cyb)	1.10.2.2 (Ubiquinol-cytochrome-c reductase)
PF3D7_0102400	P. falciparum 3D7	PF3D7_01_v3: 107,197 - 108,348 (-)	lysophospholipase, putative, pseudogene	3.1.1.5 (Lysophospholipase)
PF3D7_0102600	P. falciparum 3D7	PF3D7_01_v3: 119,041 - 121,249 (-)	serine/threonine protein kinase, FIKK family (FIKK1)	2.7.11.1 (Non-specific serine/threonine protein kinase)
PF3D7_0102900	P. falciparum 3D7	PF3D7_01_v3: 128,960 - 130,840 (-)	aspartate-tRNA ligase	6.1.1.12 (Aspartate-tRNA ligase)

STEP 2: In the 'My Strategies' section, use the 'Add Step' button to further filter the results using a new search.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New Opened (1) All (4) Basket Examples Help

(Genes) Strategy: Pfal Enz \* 1 of 4

Click here for 'Add Step' menu

Add Step

Select transcript expression search by microarray evidence

Run a new Search for Transform by Orthology Add contents of Basket Add existing Strategy Filter by assigned Weight Transform to Pathways Transform to Compounds

Genes Genomic Segments SNPs ORFs

Text, IDs, Organism Genomic Position Gene Attributes Protein Attributes Protein Features Similarity/Pattern Transcript Expression Protein Expression Cellular Location Putative Function Evolution Population Biology

EST Evidence SAGE Tag Evidence Microarray Evidence RNA Seq Evidence ChIP on Chip Evidence TF Binding Site Evidence

Close

1174 Genes from Step 1 Strategy: Pfal Enz

Filter results by species (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	Plasmodium falciparum		Plasmodium vivax	Plasmodium yoelii			Plasmodium berghei	Plasmodium chabaudi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium knowlesi	Plasmodium reichenowi
		Distinct genes	3D7		IT	Distinct genes	yoelii 17XNL						
1174	1063	1148	1174	0	0	0	0	0	0	0	0	0	

Gene Results Genome View

First 1 2 3 4 5 Next Last

Advanced Paging

Gene ID	Organism	Genomic Location	Product Description	EC Numbers
mal_mito_1	P. falciparum 3D7	M76611: 734 - 1,486 (-)	cytochrome c oxidase subunit 3 (cox3)	1.9.3.1 (Cytochrome-c oxidase)
mal_mito_2	P. falciparum 3D7	M76611: 2,035 - 3,471 (+)	cytochrome c oxidase subunit I (cox1)	1.9.3.1 (Cytochrome-c oxidase)
mal_mito_3	P. falciparum 3D7	M76611: 3,492 - 4,622 (+)	apocytochrome b (cyb)	1.10.2.2 (Ubiquinol-cytochrome-c reductase)
PF3D7_0102400	P. falciparum 3D7	PF3D7_01_v3: 107,197 - 108,348 (-)	lysophospholipase, putative, pseudogene	3.1.1.5 (Lysophospholipase)
PF3D7_0102600	P. falciparum 3D7	PF3D7_01_v3: 119,041 - 121,249 (-)	serine/threonine protein kinase, FIKK family (FIKK1)	2.7.11.1 (Non-specific serine/threonine protein kinase)
PF3D7_0102900	P. falciparum 3D7	PF3D7_01_v3: 128,960 - 130,840 (-)	aspartate-tRNA ligase	6.1.1.12 (Aspartate-tRNA ligase)
PF3D7_0103400	P. falciparum 3D7	PF3D7_01_v3: 147,915 - 152,777 (-)	zinc-carboxypeptidase, putative	3.4.17.10 (Carboxypeptidase E)
PF3D7_0103600	P. falciparum 3D7	PF3D7_01_v3: 161,131 - 166,229 (+)	ATP-dependent RNA helicase, putative	3.6.5.2 (Small monomeric GTPase)
PF3D7_0103700	P. falciparum 3D7	PF3D7_01_v3: 168,513 - 168,687 (+)	L-seryl-tRNA(Sec) kinase, putative (PSTK)	2.7.1.164 (O-phosphoserine-tRNA(Sec) kinase)
PF3D7_0104300	P. falciparum 3D7	PF3D7_01_v3: 190,269 - 201,230 (+)	ubiquitin carboxyl-terminal hydrolase 1, putative (UBP1)	3.4.19.12 (Ubiquitinyl hydrolase 1)

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New Opened (1) All (4) Basket Examples Help

(Genes) Strategy: Pfal Enz

1174 Genes from Step 1 Strategy: Pfal Enz

Filter results by species (results removed)

All Results	Ortholog Groups	Plasmodium falciparum	Plasmodium reichenowi
1174	1063	1148	1174

Gene Results Genome View

First 1 2 3 4 5 Next Last

Gene ID	Organism
mal_mito_1	P. falciparum 3D7
mal_mito_2	P. falciparum 3D7
mal_mito_3	P. falciparum 3D7
PF3D7_0102400	P. falciparum 3D7
PF3D7_0102600	P. falciparum 3D7
PF3D7_0102900	P. falciparum 3D7
PF3D7_0103400	P. falciparum 3D7
PF3D7_0103600	P. falciparum 3D7
PF3D7_0103700	P. falciparum 3D7
PF3D7_0104300	P. falciparum 3D7
PF3D7_0104400	P. falciparum 3D7
PF3D7_0105900	P. falciparum 3D7
PF3D7_0106100	P. falciparum 3D7

### Add Step 2 : Microarray Evidence

Filter Data Sets: Type keyword(s) to filter Legend: AGS Association to G FC Fold Change FD Fold Difference P Percentile S Similarity SA Similarity of Ass

Organism Data Set Choose a search

P. berghei ANKA	Transcript Profiling of DOZI Mutant Strain (Mair et al.)	FC	P
P. berghei ANKA	Transcript Profiling of Developmental Stages - High Producer (Andrew Waters)	FC	P
P. falciparum 3D7	eQTL for HB3, Dd2 and 34 progeny (Gonzales et al.)	AGS	S
P. falciparum 3D7	Life cycle expression data (Le Roch et al.)	FC	P
P. falciparum 3D7	Erythrocytic expression time series (3D7, DC, Linas et al.)	FC	P
P. falciparum 3D7	Asexual blood stage transcriptomes of clonal strains (Rovira-Graells et al.)	FC	FD
P. falciparum 3D7	Gametocyte stages I-V transcriptomes (Young et al.)	FC	P
P. falciparum 3D7	Invasion pathway knockouts (Stubbs et al.)	FC	P
P. falciparum 3D7	Two Sir2 KO lines (Tonkin et al.)	FC	P
P. falciparum 3D7	Three Isogenic Lines w/ CQ Treatment (Jiang et al.)	FC	P
P. vivax Sal-1	Intraerythrocytic developmental cycle of three isolates (Bozdech et al.)	FC	P
P. yoelii yoelii 17XNL	Liver, mosquito and blood stage expression profiles (Tarun et al.)	FC	P

Description

Find genes based on transcript expression data from microarray experiments.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New

(Genes) Strategy: Pfal Enz

1174 Genes from Step 1 Strategy: Pfal Enz

Filter results by species (results removed)

All Results	Ortholog Groups	Plasmodium falciparum	Plasmodium reichenowi
1174	1063	1148	1174

Gene Results Genome View

First 1 2 3 4 5 Next Last

Gene ID	Organism
mal_mito_1	P. falciparum 3D7
mal_mito_2	P. falciparum 3D7
mal_mito_3	P. falciparum 3D7
PF3D7_0102400	P. falciparum 3D7
PF3D7_0102600	P. falciparum 3D7
PF3D7_0102900	P. falciparum 3D7
PF3D7_0103400	P. falciparum 3D7
PF3D7_0103600	P. falciparum 3D7
PF3D7_0103700	P. falciparum 3D7
PF3D7_0104300	P. falciparum 3D7
PF3D7_0104400	P. falciparum 3D7
PF3D7_0105900	P. falciparum 3D7

### Add Step 2 : P.f. Asexual Cycle (percentile)

Experiment Sorbitol

Samples

- ☒ Early Ring
- ☒ Late Ring
- ☒ Early Trophozoite
- ☒ Late Trophozoite
- ☒ Early Schizogony
- ☒ Late Schizogony
- ☒ Merozoite
- ☐ select all | clear all

Minimum expression percentile 20

Maximum expression percentile 100

Matches Any or All Selected Samples? any

Protein Coding Only: protein coding

Advanced Parameters

#### Combine Genes in Step 1 with Genes in Step 2:

1 Intersect 2 1 Minus 2

1 Union 2 2 Minus 1

1 Relative to 2, using genomic colocation

Run Step Name and run the search

Description

Asexual cycle expression on Affy chips, from the Winzeler lab

Data Sets used by this search

- P. falciparum 3D7 Life cycle expression data

Nine different stages during the development of P. falciparum 3D7 were assayed by oligonucleotide microarray analysis.

- Discovery of gene function by expression profiling of the malaria parasite life cycle. Le Roch et al. Science 2003 Sep 12;301(5639):1503-8

Close

PF3D7_01_v3: 202,536 - 204,143 (+)	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (Lyb)	1.17.1.2 (4-hydroxy-3-methylbut-2-enyl diphosphate reductase)
PF3D7_01_v3: 253,584 - 254,930 (+)	DNA binding protein, putative	3.1.11.1 (Exodeoxyribonuclease I); 3.1.11.2 (Exodeoxyribonuclease III)



plasmodb.org/plasmo/showApplication.do

Version 10.0  
25 Sep 13

EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New Opened (1) All (4) Basket Examples Help

(Genes)

Strategy: Pfal Enz

Results of intersection combination.

1130 Genes from Step 2  
Strategy: Pfal Enz

Add 1130 Genes to Basket | Download 1130 Genes

Filter results by species (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	Plasmodium falciparum		Plasmodium vivax	Plasmodium yoelii			Plasmodium berghei	Plasmodium chabaudi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium knowlesi	Plasmodium reichenowi
		Distinct genes	3D7		Distinct genes	yoelii 17XNL	yoelii VM						
1130	1024	1106	1130	0	0	0	0	0	0	0	0	0	0

Gene Results Genome View

First 1 2 3 4 5 Next Last Advanced Paging Add Columns

Gene ID	Genomic Location	Product Description
mal_mito_1	M76611: 734 - 1,486 (-)	cytochrome c oxidase subunit 3 (cox3)
mal_mito_2	M76611: 2,035 - 3,471 (+)	cytochrome c oxidase subunit 1 (cox1)
mal_mito_3	M76611: 3,492 - 4,622 (+)	apocytochrome b (cyb)
PF3D7_0102400	PF3D7_01_v3: 107,197 - 108,348 (-)	lysophospholipase, putative, pseudogene
PF3D7_0102600	PF3D7_01_v3: 119,041 - 121,249 (-)	serine/threonine protein kinase, FIKK family (FIKK1)
PF3D7_0102900	PF3D7_01_v3: 128,960 - 130,840 (-)	aspartate-tRNA ligase
PF3D7_0103400	PF3D7_01_v3: 147,915 - 152,777 (-)	zinc-carboxypeptidase, putative
PF3D7_0103600	PF3D7_01_v3: 161,131 - 166,229 (+)	ATP-dependent RNA helicase, putative
PF3D7_0103700	PF3D7_01_v3: 166,513 - 168,687 (+)	L-seryl-tRNA(Sec) kinase, putative (PSTK)
PF3D7_0104300	PF3D7_01_v3: 190,269 - 201,230 (+)	ubiquitin carboxyl-terminal hydrolase 1, putative (UBP1)
PF3D7_0104400	PF3D7_01_v3: 202,536 - 204,143 (+)	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB)
PF3D7_0105900	PF3D7_01_v3: 253,584 - 254,930 (+)	DNA binding protein, putative
PF3D7_0106100	PF3D7_01_v3: 260,658 - 261,809 (-)	vacuolar ATP synthase subunit c, putative
PF3D7_0106300	PF3D7_01_v3: 265,208 - 269,173 (-)	calcium-transporting ATPase (ATP6)
PF3D7_0106800	PF3D7_01_v3: 288,061 - 288,018 (-)	DNA GTPase 5c (DAB5c)

This search result shows that most of the enzymes 1130 out of 1174 genes are expressed in at least one intra-erythrocytic stage *P. falciparum* parasite with a minimum expression percentile of 20%. Now, we can modify the search to look for only those genes whose expression value is less than 20% during the entire intra-erythrocytic development.

STEP 3: From the 'Strategy' section, edit the desired search parameter.

plasmodb.org/plasmo/showApplication.do

Version 10.0  
25 Sep 13

EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

My Strategies: New Opened (1) All (4) Basket Examples Help

(Genes)

Strategy: Pfal Enz

Results of intersection combination.

1130 Genes from Step 2  
Strategy: Pfal Enz

Add 1130 Genes to Basket | Download 1130 Genes

Filter results by species (results removed by the filter will not be combined into the next step.)

All Results	Ortholog Groups	Plasmodium falciparum		Plasmodium vivax	Plasmodium yoelii			Plasmodium berghei	Plasmodium chabaudi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium knowlesi	Plasmodium reichenowi
		Distinct genes	3D7		Distinct genes	yoelii 17XNL	yoelii VM						
1130	1024	1106	1130	0	0	0	0	0	0	0	0	0	

Gene Results Genome View

First 1 2 3 4 5 Next Last Advanced Paging Add Columns

STEP 2 : Step 1 Pf Asexual %ile

Revise Operation

1 INTERSECT 2 1 UNION 2 1 MINUS 2 2 MINUS 1

Results: 1130 Genes

Gene ID	Genomic Location	Product Description
mal_mito_1	M76611: 734 - 1,486 (-)	cytochrome c oxidase subunit 3 (cox3)
mal_mito_2	M76611: 2,035 - 3,471 (+)	cytochrome c oxidase subunit 1 (cox1)
mal_mito_3	M76611: 3,492 - 4,622 (+)	apocytochrome b (cyb)
PF3D7_0102400	PF3D7_01_v3: 107,197 - 108,348 (-)	lysophospholipase, putative, pseudogene
PF3D7_0102600	PF3D7_01_v3: 119,041 - 121,249 (-)	serine/threonine protein kinase, FIKK family (FIKK1)
PF3D7_0102900	PF3D7_01_v3: 128,960 - 130,840 (-)	aspartate-tRNA ligase
PF3D7_0103400	PF3D7_01_v3: 147,915 - 152,777 (-)	zinc-carboxypeptidase, putative
PF3D7_0103600	PF3D7_01_v3: 161,131 - 166,229 (+)	ATP-dependent RNA helicase, putative
PF3D7_0103700	PF3D7_01_v3: 166,513 - 168,687 (+)	L-seryl-tRNA(Sec) kinase, putative (PSTK)
PF3D7_0104300	PF3D7_01_v3: 190,269 - 201,230 (+)	ubiquitin carboxyl-terminal hydrolase 1, putative (UBP1)
PF3D7_0104400	PF3D7_01_v3: 202,536 - 204,143 (+)	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB)

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

**My Strategies:** New Opened (1) All (4) Basket Examples Help

(Genes)

Strategy: Pfal Enz

Rename Duplicate Save As Share Delete

Pfal Enz 1174 Genes Step 1

44 Genes Step 2

Add Step

After modifying the search combination we find that there are only 44 genes that do not express during intra-erythrocytic development.

44 Genes from Step 2 Strategy: Pfal Enz

Add 44 Genes to Basket | Download 44 Genes

Filter results by species (results removed by the filter will not be shown)

All Results	Ortholog Groups	Plasmodium falciparum	Plasmodium vivax	Plasmodium knowlesi	Plasmodium reichenowi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium berghei	Plasmodium chabaudi	Plasmodium yoelii	Plasmodium yoelii
44	43	43	44	0	0	0	0	0	0	0	0

Gene Results Genome View

Advanced Paging Add Columns

Gene ID	Genomic Location	Product Description
PF3D7_0107800	PF3D7_01_v3: 314,618 - 319,405 (+)	serine/threonine protein kinase, putative
PF3D7_0203200	PF3D7_02_v3: 158,137 - 159,660 (-)	conserved Plasmodium protein, unknown function
PF3D7_0216000	PF3D7_02_v3: 658,636 - 664,629 (+)	DEAD/DEAH box helicase, putative
PF3D7_0305800	PF3D7_03_v3: 281,916 - 285,260 (+)	P-loop containing nucleoside triphosphate hydrolase, putative
PF3D7_0311400	PF3D7_03_v3: 488,804 - 496,510 (+)	protein kinase, putative
PF3D7_0420200	PF3D7_04_v3: 910,187 - 911,971 (+)	holo-(acyl-carrier protein) synthase, putative
PF3D7_0605600	PF3D7_06_v3: 228,132 - 234,835 (-)	nucleoside diphosphate kinase, putative
PF3D7_0607500	PF3D7_06_v3: 311,550 - 313,115 (+)	para-hydroxybenzoate-polyphenyltransferase, putative
PF3D7_0609800	PF3D7_06_v3: 421,157 - 422,622 (-)	zinc finger protein, putative
PF3D7_0618900.1	PF3D7_06_v3: 784,354 - 788,027 (+)	N-acetylglucosamine transferase (GPI1)
PF3D7_0618900.2	PF3D7_06_v3: 784,354 - 788,027 (+)	N-acetylglucosamine transferase (GPI1)

STEP 4: To validate this result, we can further extend the search strategy using expression evidence from RNA sequence data.

PlasmoDB Plasmodium Genomics Resource Version 10.0 25 Sep 13 A EuPathDB Project

Gene ID: PF3D7\_1133400 Gene Text Search: synth\*

About PlasmoDB | Help | Dhanasekaran Shannmugam's Profile | Logout | Contact Us

Home New Search My Strategies My Basket (210) Tools Data Summary Downloads Community My Favorites

**My Strategies:** New Opened (1) All (4) Basket Examples Help

(Genes)

Strategy: Pfal Enz

Rename Duplicate Save As Share Delete

Pfal Enz 1174 Genes Step 1

44 Genes Step 2

8 Genes Step 3

Add Step

Genes with RNA seq evidence for greater than 20 percentile expression during the entire intra-erythrocytic development

Only 8 out of the 44 enzymes selected by microarray data are validated by RNA seq data as having no expression during intra-erythrocytic development.

8 Genes from Step 3 Strategy: Pfal Enz

Add 8 Genes to Basket | Download 8 Genes

Filter results by species (results removed by the filter will not be shown)

All Results	Ortholog Groups	Plasmodium falciparum	Plasmodium vivax	Plasmodium knowlesi	Plasmodium reichenowi	Plasmodium cynomolgi	Plasmodium gallinaceum	Plasmodium berghei	Plasmodium chabaudi	Plasmodium yoelii	Plasmodium yoelii
8	8	8	8	0	0	0	0	0	0	0	0

Gene Results Genome View

Advanced Paging Add Columns

Gene ID	Genomic Location	Product Description
PF3D7_0818900	PF3D7_08_v3: 770,946 - 771,773 (+)	adenylate kinase 2 (AK2)
PF3D7_0910000	PF3D7_09_v3: 452,399 - 455,743 (-)	SET domain protein, putative (SET4)
PF3D7_1121800	PF3D7_11_v3: 823,239 - 829,313 (-)	petidase, M16 family
PF3D7_1139500.2	PF3D7_11_v3: 1,576,132 - 1,578,978 (-)	AAA family ATPase, putative
PF3D7_1468600	PF3D7_14_v3: 2,816,100 - 2,822,286 (+)	aminophospholipid transporter, putative
PF3D7_1475500	PF3D7_14_v3: 3,106,949 - 3,111,811 (+)	LCCL domain-containing protein (CCP1)
PFC10_API0016	PFC10_API_IRAB: 10,108 - 11,835 (+)	RNA polymerase beta' subunit, putative (rpoC)
PFC10_API0017	PFC10_API_IRAB: 11,844 - 14,725 (+)	RNA polymerase D (rpoD)

PlasmoDB 10.0 25 Sep 13 ©2013 The EuPathDB Project Team

EuPathDB

Please Contact Us with any questions or comments

POWERED BY Strategies WDK

As previously discussed, the list of genes in the result can be saved and exported.

## **Exercises:**

1. Register and create your PlasmoDB account.
2. Using text search, find all possible kinases in *Plasmodium falciparum*.  
Hint: use the keyword “kinase” (without quotations) in the “Gene Text Search” box.
3. From the results obtained above, select only those with a protein kinases domain.  
  
Hint: Use the Pfam domain ID PF00069 to filter the genes.
4. Find all the proteins kinases that are likely to be secreted from the cell.  
  
Hint: Secreted proteins will have a signal peptide.
5. Which of these secreted kinases have their gene expression up-regulated during the late stages of intra-erythrocytic development?  
  
Hint: Use transcript expression data for this.
6. How many of these secreted kinases belong to expanded gene families?  
  
Hint: Use the search options under the “Evolution” section.
7. Select only the expanded kinases and place them in your Basket. Now use the gene in the basket to generate a genomic view of these kinases.
8. Download the gene ID list for secreted kinases along with all available datasets as a spread sheet.
9. Select a gene from the genomic view and view the GBrowse page.
10. On the GBrowse page, choose the appropriate tracks to display the synteny among all available *Plasmodium* species.
11. Find all *P. falciparum* genes absent in mammalian species.
12. Find all *P. falciparum* gene present exclusively in the Apicomplexan phylum.
13. Find all gene present exclusively in *Plasmodium* species.
14. Identify all metabolic enzymes under diversifying selection.
15. Identify the number restriction sites for the enzyme EcoRI (GAATTC) on chromosome-1 of *P. falciparum*.