How to use KAIKObase Version 3.1.0

http://sgp2010.dna.affrc.go.jp/KAIKObase/

Copyright © National Institute of Agrobiological Sciences. All rights reserved.
Outline

1. System overview

2. Top page

3. Browsers, viewer, and external databases
   3.1 Pgmap
   3.2 UnifiedMap
   3.3 Gbrowse
   3.4 GeneViewer
   3.5 *Bombyx* Trap Database
   3.6 Proteome database (KAIKO2DDB)
   3.7 KAIKOGAAS (An integrated database and automated annotation system)
   3.8 Full-length cDNA Database
   3.9 EST Database

4. Keyword and position search, and sequence search
   4.1 Example of keyword and position search
   4.2 Example of sequence search
1. System overview

KAIKObase is an integrated silkworm genome browser with 3 map browsers, 1 gene viewer, and 5 independent databases.

* PGmap : physical map and genetic linkage map for each chromosome.
* UnifiedMap : PGmap containing genome contigs, BAC-ends, fosmid-ends, and Fingerprint Contigs (FPC).
* GBrowse : genome annotation viewer based on GBrowse developed by GMOD.

* GeneViewer : gene information viewer based on the gene viewer of KAIKOGAAS.

* Bombyx Trap DataBase : Bombyx gene function database by transgenic works using GAL4 system.
* KAIKO 2DDB : proteomic data of various developmental stages and tissues of silkworm.
* KAIKOGAAS : An integrated database and automated annotation system.
* Full-length cDNA (FLcDNA) Database
* EST Database

KAIKObase provides three data mining approaches:

(1) Narrowing the range of data sets using PGmap or UnifiedMap in Chromosome Overview
(2) Keyword and position search
(3) Sequence search using BLAST.
2. Top page

The top page provides access to silkworm genome information.

1) Chromosome select: Two choices are available for the genome maps – the PGmap can be accessed by clicking the chromosome number after clicking PGmap (colored pink) and the UnifiedMap by clicking the chromosome number after clicking UnifiedMap (colored green).

2) Keyword and Position Search: Input a keyword and choose data set from “Search Field”, “Search Chromosome and/or Position” or “Scaffold name and/or Position”, and “Search Data Set” delimit the area and the region of the search. The data sets include genes, scaffolds, contig within scaffolds, markers, fingerprint contigs (FPC), BAC ends, fosmid ends, BAC clones, Locus, FLcDNAs, complete mRNAs, gene models, and ESTs. As long as “Graphic View ON”, search results are shown in chromosome image as well as in Excel table.

3) Sequence Search: Nucleotide or amino acid sequence can be used as a query. Search results are shown in chromosome images.
3. Browsers, viewer, and external database

3.1 PGmap
3.2 UnifiedMap
3.3 GBrowse
3.4 GeneViewer
3.5 *Bombyx* Trap Database
3.6 Proteome Database (KAIKO2DDB)
3.7 KAIKOGAAS (An integrated database and automated annotation system)
3.8 Full-length cDNA (FLcDNA) database
3.9 EST database
3.1 PGmap

[Features]
- Position of markers in the linkage map and physical map at the chromosome level.

[Link]
- Choose designated region and link to GBrowse.
- Choose designated region and link to view sequence.

1. Select a chromosome
2. Specify the desired region in the chromosome by moving the blue frame.
3. Click display on/off to highlight the area for GBrowse link.
4. Specify the region for GBrowse viewing or sequence viewing by moving the red frame.
5. Click “Go to GBrowse” button to view link.
6. Click “View sequence” button to view link.
7. Alternatively, enter the exact position within the chromosome and click “Go to GBrowse” button to view link or click “View sequence” button to view link.

As the size of gap between adjacent scaffolds is unknown, we set an uniform space of 500 kbps between scaffolds. The direction of a scaffold containing plural markers is defined, while the direction of a scaffold having only one marker cannot be determined and the original direction is used.
3.2 UnifiedMap

[Features]
- Pgmap is integrated with the information of scaffolds and contigs of Build2, BAC ends, Fosmid ends, FPC, SNP markers and Trait markers.

[Link]
- Choose designated region on the physical map and link to GBrowse or get to sequence.
- Choose designated region on the linkage map. A red arrow appears on right side if marker information is available.

1. Select a chromosome
2. Specify the zooming level.
3. Check the desired items for display.
3.3 GBrowse

[Features]
- Genome annotation viewer with a graphical representation of a section of a chromosome.

[Link]
- Click designated region on “Details” and sequence and ID are displayed in another page.
- Click gene model on “Details” for link to GeneViewer and KAIKO2DDB (proteome database).
- Click trait marker on “Details” for link to Bombyx Trap DataBase.

Click desired position of Overview, then a display will change to information of the region.

After clicking gene model, jump to URLs is shown.

Check display items that you want to see in the display, then, press “Redraw” to present the desired items on the display.
3.4 GeneViewer

[Features]
- Gene information viewer, showing the detailed information of the gene model.

[Link]
- “Search Results” links to results of homology search by BLASTn, BLASTp, HMMER and ProfileScan with alignments of the sequence.
- “Analysis of amino acid sequences” links to PSORT, SOSUI, MOTIF and Gene ontology (InterProScan).
- “Sequence” provides the nucleotide sequence, spliced nucleotide sequence and translated protein sequence of the predicted gene.

Result of domain search in Pfam database.

Detailed information of the predicted gene including chromosome number, positions of exons and GC content.

Results of homology search using BLASTn (top 3 ESTs), BLASTp (top 10 proteins), HMMER and ProfileScan with the alignments of the sequence.

Display the results of amino acid analysis in PSORT, SOSUI, MOTIF and Gene ontology (InterProScan).

Links to the nucleotide sequence, spliced nucleotide sequence and translated protein sequence of the predicted gene.

Graphical representation of InterProScan results

Character representation of InterProScan results
3.5 *Bombyx* Trap DataBase

[Features]
- Information on reporter expressions and positions of transposon vectors (mutators) in transposon insertion lines (enhancer trap strains, or gene trap strains) is presented with this database.
For more details, please refer to the document, http://sgp.dna.affrc.go.jp/ETDB/howto/how_to_use_Bombyx_Trap_DB.pdf

[Link]
This database is linked from Keyword search result page of KAIKObase.

*Bombyx* Trap DataBase top page
3.6 Proteome database (KAIKO2DDB)

[Features]
- Database for silkworm proteome analysis using two-dimensional polyacrylamide gel electrophoresis images of various developmental stages and tissues.
- For more details, please refer to the document, http://au.expasy.org/ch2d/manch2d.html

[Link]
- This database is linked from Keyword search result page or GBrowse of KAIKObase.

Proteome database (KAIKO2DDB) top page
3.7 KAIKOGAAS

[Features]
- Database of detailed annotation of a total of 192 scaffolds anchored to the 28 silkworm chromosomes.
- For more details, please refer to the KAIKOGAAS document, http://kaikogaas.dna.affrc.go.jp/howtokaikogaas.pdf

[Link]
- This database is linked from GBrowse of KAIKObase.

KAIKOGAAS top page
3.8 Full-length cDNA (FLcDNA) Database

[Features]
- 10,165 full-length cDNAs (FLcDNAs) and 2,627 complete mRNAs in public database are included in the Full-length cDNA (FLcDNA) Database.
- Keyword search provides a list of BLAST top hit annotation using the NCBI NR database.

[Link]
- 9,146 FLcDNAs of 10,165 FLcDNAs and 2,555 complete mRNAs of 2,627 complete mRNA are mapped on scaffold and linked from Gbrowse of KAIKObase.
- 10,165 FLcDNAs and 2,627 complete mRNAs are linked from keyword search result page of KAIKObase

Full-length cDNA (FLcDNA) top page

Search Entries of full cDNA-length clone Database

[notes] If you search without search option, all entries are displayed.
3.9 EST database

[Features]
- A total of 463,607 ESTs correspond to unreleased EST sequences, Bombyx EST sequences in public database as of March 2010, and Bombyx partial cDNA sequences in public database as of June 2010.
- Keyword search provides a list of Top hit annotation of InterProScan and BLAST using database of NCBI NR, Drosophila, C.elegans, Anopheles, Apis, and Tribolium.

[Link]
- 399,510 ESTs of 463,607 ESTs are mapped on scaffold and linked from Gbrowse of KAIKObase.
- 463,607 ESTs are linked from keyword search result page of KAIKObase.

EST database top page
4. Keyword and Position search, and Sequence search

4.1 Example of keyword and position search
4.2 Example of sequence search
4.1 Example of keyword and position search

Search for “catalase” in gene data set on whole chromosome.

Enter “catalase” in query box

Do not click if all scaffolds would be scanned.

Click “Gene model”

Choose “on” from pull-down menu.

Click “Search” button.

Search result is downloaded by clicking “download hit text data” button.

Choose “description” from pull-down menu

Click any button to view link.

Result of keyword search
4.2 Example of sequence search

Example of search for location of a specified amino acid sequence on scaffold build2.