VectorBase:

A Relational Bioinformatics Resource Center for Arthropod Vectors of Human Pathogens

VectorBase: Components

- EBI/EMBL (E. Birney): Ensembl
- EMBL Heidelberg (F. Kafatos): ESTs, Array Data, RNAi Data
- Harvard (W. Gelbart): Manual Annotation
- IMBB Crete (C. Louis): AnoBase
- U. Notre Dame (F. Collins): Coordination, Data Integration

Vector Genomes

- *Anopheles gambiae: 260 Mb*
- *Aedes aegypti: 900 Mb*
- *Culex pipiens quinquefasciatus: 450 Mb*
- *Ixodes scapularis: 2 Gb*
- *Glossina morsitans: 400 Mb*
- *Anopheles funestus: 260 Mb*

VectorBase: Species-Specific Genome Committees

- *Anopheles gambiae*: (VectorBase Investigators)
- *Aedes aegypti*: D. Severson, UND
- *Culex pipiens*: Peter Atkinson, UCR
- *Ixodes scapularis*: C. Hill, Purdue
- *Glossina morsitans*: S. Aksoy, Yale; M. Lehane, Liverpool
- *Anopheles funestus*: N. Besansky, UND

VectorBase Home

VectorBase.org

http://www.vectorbase.org/

VectorBase

All Organisms | A. gambiae | A. funestus | C. pipiens | I. scapularis | A. aegypti

Help: User: Not Logged In.

Genome | Gene Expression | EST | Images | More... | Get Data | Tools | Search

Welcome To VectorBase!

A Bioinformatics Resource Center for Invertebrate Vectors of Human Pathogens.

VectorBase Services:

- **Organisms »**
  A. gambiae, A. aegypti, A. funestus, C. pipiens, I. scapularis, more...
- **Data Sections »**
  Browse Genomes, EST Data, RNAi Data, Microarray Data, Images, more...
- **Tools »**
  BLAST, ClustalW, EST Analysis, more...
- **Retrieve Data »**
  Based on Gene Functionality (GO), Anatomy, BioMart, Terminology Search, more...
- **Submit Information »**
  Gene Names, Gene Functionality, Corrections, Images, more...
- **Documentation »**
  Getting Started, Help, Architecture, Data Submission , Naming Conventions, more...

Ixodes scapularis

Recent News:

09.23.04
Website Updated »

09.17.04
E-Mail Listserv Online »

09.01.04
First VectorBase meeting »

08.10.04
New Website »

Quick Links

- Getting started @ Vectorbase »
- Help Documentation »
- Ask a question »
- Downloads »
- Site Map »
- Contact Us »

Quick Search

Search:

Everything

for term or keyword:

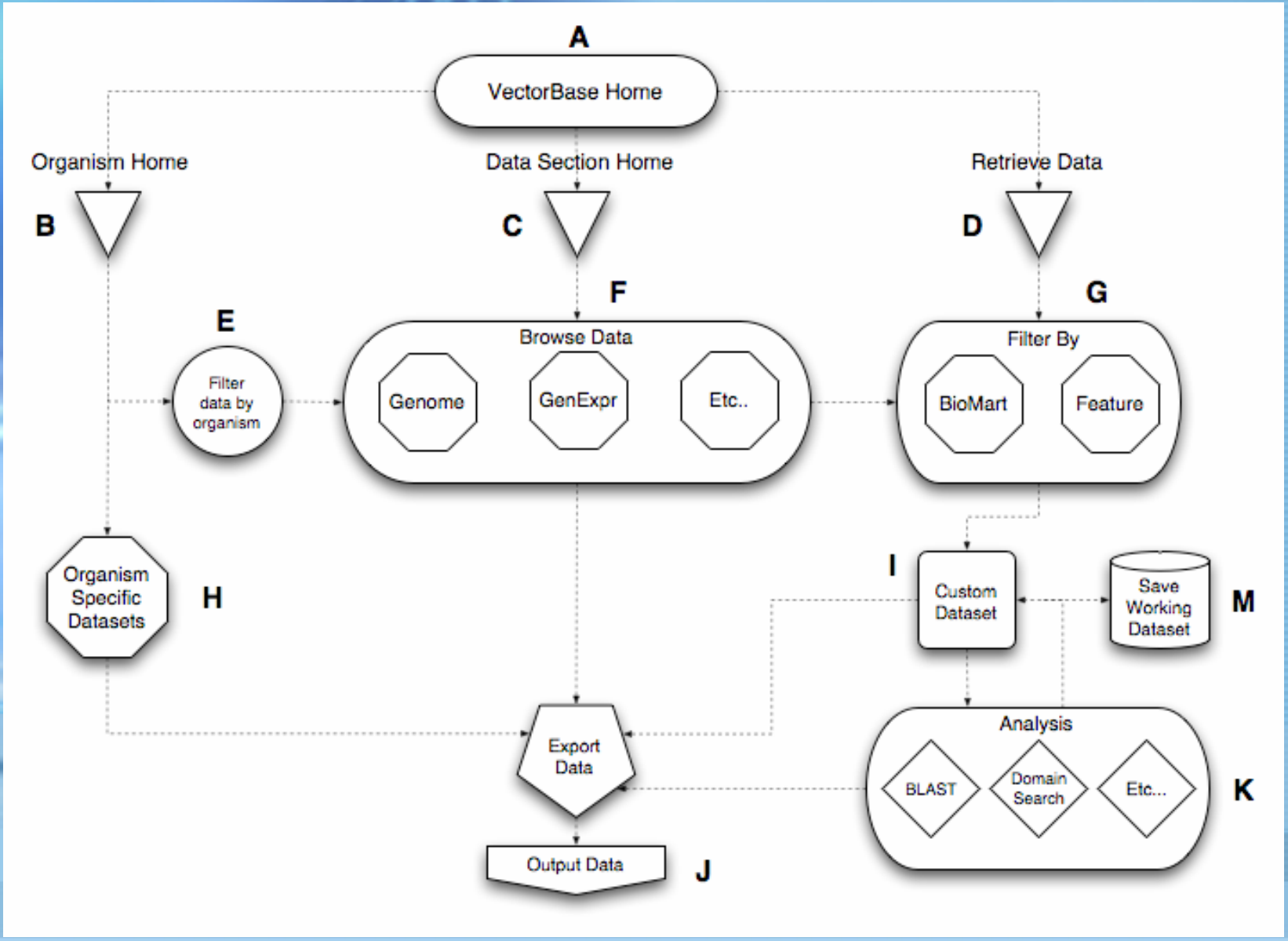
Go! • Advanced Search

Vectorbase Last Updated 09.29.04

Contact Webmaster

UNIVERSITY OF NOTRE DAME | Ensembl | IMBB | Harvard | EMBL

Workflow: Organism Home



Entry Point: Organism Home

- Entry point to all data available for a single organism.
- Links to custom organism datasets.
- Links to organism specific data sections.

Organism Home

A. gambiae @ VectorBase.org

http://agambiae.vectorbase.org/

Google

A. gambiae @ VectorBase.org

VectorBase: A. gambiae

All Organisms | A. gambiae | A. funestus | C. pipiens | I. scapularis | A. aegypti

Help: User: Not Logged In.

Genome | Gene Expression | EST | Images | More... | Get Data | Tools | Search

Welcome to the Anopheles gambiae section of VectorBase. Here you will find entrypoints to all gathered data related to the A. gambiae. Browse the genome by clicking on a chromosome, search for genes based on anatomy, or browse our catalog of images.

Anopheles gambiae @ VectorBase.org

Content Cordinator: TBA

Genomic Data

Browse A Chromosome:

- Chr 2R
- Chr 2L
- Chr 3R
- Chr 3L
- Chr X

Unmapped Genomic Data:

- Unmapped Scaffolds »
- Unplaced Chaff »
- Original Shotgun Data »

A. Gambiae Home last updated 10.11.04

Vectorbase Last Updated 09.29.04

Contact Webmaster

UNIVERSITY OF NOTRE DAME | Ensembl | IMBB | Harvard | EMBL

Organism Data Section

- Genome

A. gambiae @ VectorBase: Genome

http://agambiae.vectorbase.org/Genome/MapView/?chr=3R

VectorBase: A. gambiae

All Organisms | A. gambiae | A. funestus | C. pipiens | I. scapularis | A. aegypti

Help: User: Not Logged In.

Genome | Gene Expression | EST | Images | More... | Get Data | Tools | Search

Home > Map View > Contig View > Gene View > Trans View > Exon View > Protein View >

Known Genes
Genes
% GC
Repeats
SNPs
Chromosome 3R

29A 29B 29C 29D 30A 30C 30D 30E 31B 31C 32A 32C 32D 33B 33C 33D 34B 34C 34D 35B 35C

Chromosome 3R

Length: 53,272,125 bps
Gene Count: 2670
Known Gene Count: 155
SNP Count: 250574

Change Chromosome

Chromosome: 3R Go

Jump to Contigview

Click anywhere on the chromosome ideogram or one of the feature distribution plots to jump to a contig-level view of features at that point. Alternatively, you can jump to contigview between any two markers on this chromosome:

Between:

and:

Go

Display contig-level view between any two features.

Organism Data Section

- Genome

A. gambiae @ VectorBase: Genome

http://agambiae.vectorbase.org/Genome/GeneView?species=Anopheles_gambiae&gene=ENS

VectorBase: A. gambiae

All Organisms | A. gambiae | A. funestus | C. pipiens | I. scapularis | A. aegypti

Help: User: Not Logged In.

Genome | Gene Expression | EST | Images | More... | Get Data | Tools | Search

Home > Map View > Contig View > Gene View > Trans View > Exon View > Protein View >

Ensembl Gene Report

Gene	GPRor46 (Anopheles_symbol ID)
Ensembl Gene ID	ENSANGG00000002851
Genomic Location	View gene in genomic location: 32489986 - 32494516 bp (32.5 Mb) on chromosome 3R This gene is located in sequence: AAAB01008835_35
Description	No description
Prediction Method	Ensembl genes were annotated by the Ensembl automatic analysis pipeline combining both homology matching methods with ab-initio gene prediction algorithms supported by functional domains and EST data.
Sequence Markup	View genomic sequence for this gene with exons highlighted
Export Data	Export gene data in EMBL, GenBank or FASTA
SNP information	View information about variations on this gene.

1: GPRor47 (ENSANGT00000015280) [Transcript information] [Exon information] [Protein information]
2: GPRor46 (ENSANGT00000025198) [Transcript information] [Exon information] [Protein information]

Transcript Structure

Features ▼

GPRor55 Ensembl known trans
GPRor13 Ensembl known trans
GPRor46 Ensembl known trans
GPRor17 Ensembl known trans
GPRor30 Ensembl known trans
GPRor16 Ensembl known trans
GPRor53 Ensembl known trans
GPRor47 Ensembl known trans
GPRor15 Ensembl kno

24.53 Kb

SwissProt

GeneDAS Sources

Manage Sources

Other Data Types

- Data currently in AnoBase
- Data Management Tools Under Development
- Data Tools Developed for Species-specific Communities
- Population Data (Population Genetics, Epidemiology and Ecology)

Organism Custom Dataset

- In-Situ Images @ AnoBase

Search Results

http://www.anobase.org/cgi-bin/insitu.pl

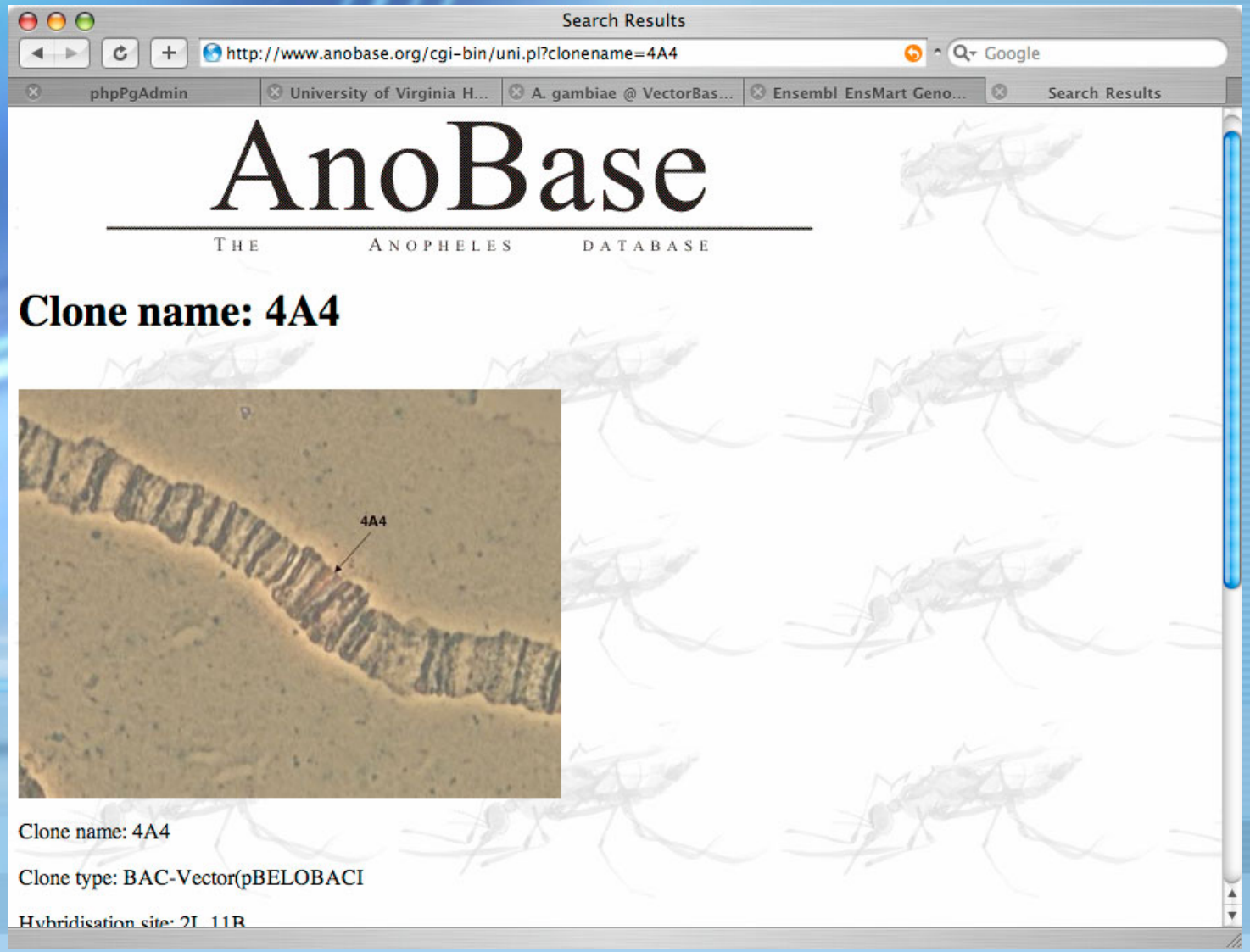
phpPgAdmin | University of Virginia H... | A. gambiae @ VectorBas... | Ensembl EnsMart Geno... | Search Results

AnoBase
the *Anopheles* database

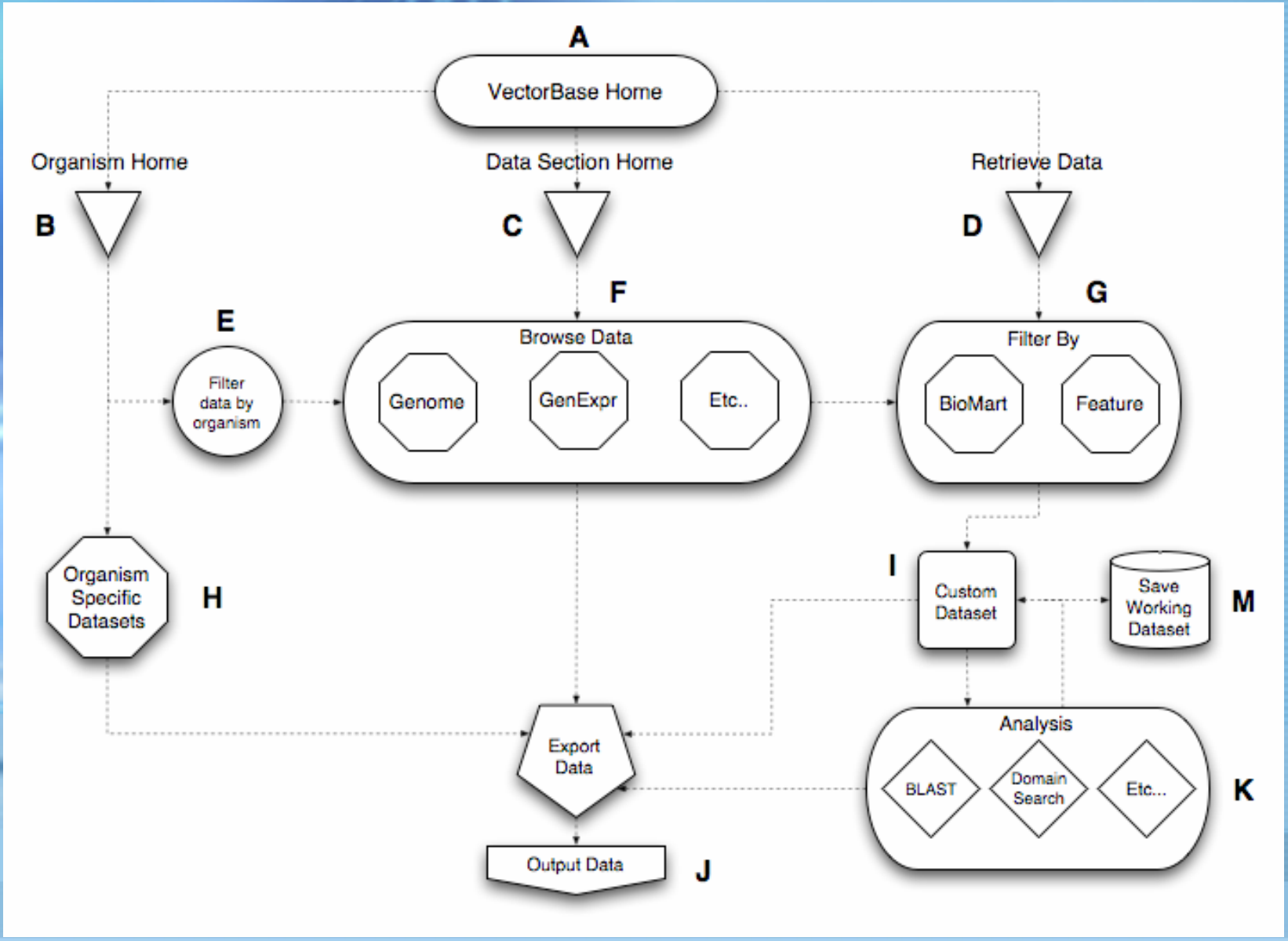
Clone name	Primary site	Secondary site	Clone type
4A4	2L 11B		BAC-Vector(pBELOBACI
4A7	2L 23B		BAC-Vector(pBELOBACI
4C12	2L 25C		BAC-Vector(pBELOBACI
6D13	2L 28B		BAC-Vector(pBELOBACI
6E5	2L 23B		BAC-Vector(pBELOBACI
6H13	2L 25A		BAC-Vector(pBELOBACI
6M13	2L 28C		BAC-Vector(pBELOBACI
6M3	2L 24C		BAC-Vector(pBELOBACI
101D16	2L24B	35B,perice	BAC
123C09	2L22A		BAC
131K01	2L20D		BAC
150K24	2L21F/22A	pericentro	BAC

Organism Custom Dataset

- In-Situ Images @ AnoBase



Workflow: Retrieve Data



Retrieving & Filtering Data

- Users want to gather data based on certain characteristics for further analysis.

Filtering Data: BioMart

- A simple, distributed data integration system with powerful query capabilities.
- Originally developed as a data retrieval tool specifically for EnsEMBL, now expanded to target other databases.
- Collects data based on genomic location and gene structural characteristics.

BioMart

Ensembl EnsMart Genome Browser (MartView)

http://www.ensembl.org/Multi/martview/DFzpwyUynX.mart

phpPgAdmin | University of Virginia H... | A. gambiae @ VectorBas... | Ensembl EnsMart Geno...

project Ensembl MartView

The Wellcome Trust Sanger Institute

EBI

Home ▸ EnsMart ▲ TextSearch ▲ BlastSearch ▲ MartSearch ▲ Download ▲

new | START | FILTER | OUTPUT | export

refresh | Online Help | Help Desk

FILTER

◂ back | next ▸

Further refine your search or click 'next':

REGION:

☑ Limit to (uncheck for entire genome):

Chromosome name: 2L

From Base pair 1 ---

To Base pair 10M ---

GENE:

☐ Limit to genes whose annotation includes Known status — ◉ Only ○ Excluded

☐ Limit to genes with the following Anopheles Symbol ID(s): — Choose File no file selected

☐ Transcripts per gene: ◉ Single ○ Multiple

☐ Entries with a 5' UTR ◉ Only ○ Excluded

Summary

▸ start

- Focus: Ensembl Genes
- Species: Anopheles gambiae

ⓘ *14364 Entries Total*

▸ filter

ⓘ *Not yet initialised*

▸ output

ⓘ *Not yet initialised*

BioMart

Ensembl EnsMart Genome Browser (MartView)

http://www.ensembl.org/Multi/martview/DFzpwyUynX.mart

phpPgAdmin | University of Virginia H... | A. gambiae @ VectorBas... | Ensembl EnsMart Geno...

project Ensembl MartView

The Wellcome Trust Sanger Institute

Home ▸ EnsMart ▲ TextSearch ▲ BlastSearch ▲ MartSearch ▲ Download ▲

new | START ◂▸ FILTER ◂▸ OUTPUT | export

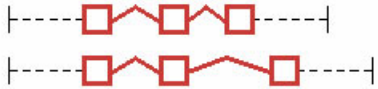
OUTPUT

◂ back | export

Features | SNPs | Structures | Sequences

Type of Sequence to Export (all in 5'-3' direction):

- (•) Transcripts/proteins
- ( ) Genes - transcript information ignored (one output per gene)



- (•) Gene sequence only
- ( ) Gene plus 5' flank
- ( ) 5' upstream only
- ( ) 5' UTR only
- ( ) 3' UTR only
- ( ) Exon sequences only
- ( ) Exons plus 5' flanks
- ( ) cDNA sequence only
- ( ) Peptide
- ( ) Gene plus 5' and 3' flanks
- ( ) Gene plus 3' flank
- ( ) 3' downstream only
- ( ) 5' upstream and UTR
- ( ) 3' UTR and downstream
- ( ) Exons plus 5' and 3' flanks
- ( ) Exons plus 3' flanks
- ( ) Coding sequence only

5' Flank (bp) 1000 | 3' Flank (bp) 1000

refresh | Online Help | Help Desk

Summary

▸ start
- Focus: Ensembl Genes
- Species: Anopheles gambiae

ⓘ *14364 Entries Total*

▸ filter
- Chromosome: 2L
- From base: 1
- To base: 10000000

ⓘ *477 Entries pass Filters*

▸ output
- Sequence

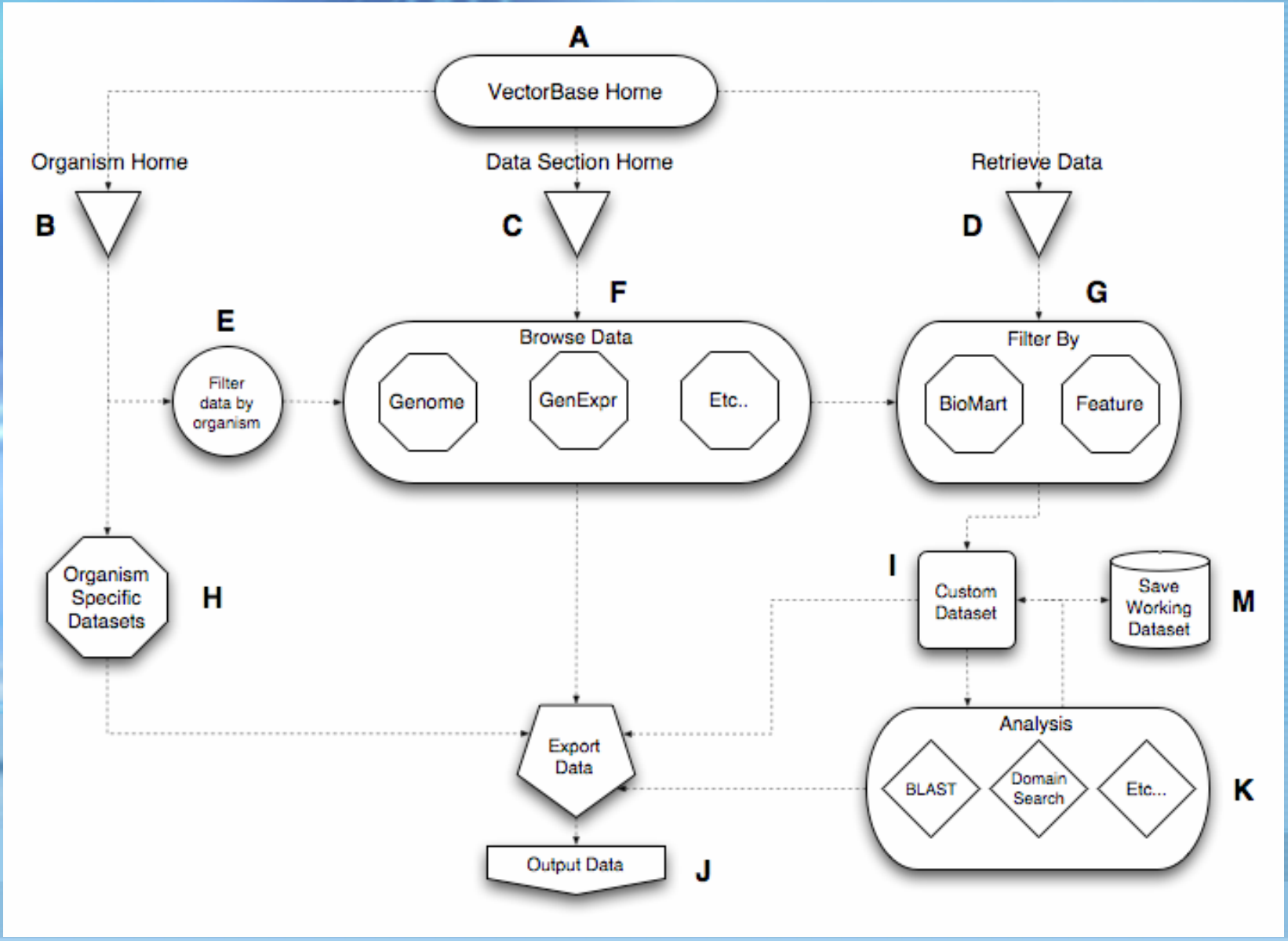
Filtering Data: Sequence Features

- Retrieve data based on controlled vocabulary.
  - GO Associations
  - Anatomical Characteristics
  - Experimental Conditions

Filtering Data: Completion

- Retrieved data can be downloaded in a variety of formats for further personalized use.
- Data can be saved as a custom dataset for further manipulation. **Restricted**

Workflow: Data Analysis



Data Analysis

- Computed on VectorBase Cluster
- Restricted to privileged users
- Analysis available to all users after completion

Data Analysis

- Grouped data run through analysis programs.
- BLAST / InterPro / Exonerate
- Web Interface
- Results can be exported or saved as gene features if approved by content coordinator.

Data Analysis: EST Processing

VBEST

http://vbest.vectorbase.org/www/interface/index.php?content=dataset_info&dataset

phpPgAdmin | University of Virginia H... | A. gambiae @ VectorBas... | Ensembl EnsMart Geno... | VBEST

VBEST

rbruggne logged in

- home
- public datasets
- search
- user home
- user datasets
- user libraries
- administrative tools
- log out

Dataset DTSTID_00000000010 : Ap

General Info

Creator:	Robert Bru
Owner:	Robert Bru
Associated Library:	LIBRID_00
Child Datasets:	DTSTID_00
Active Sequences:	28
Dataset Description:	BAC Cover

Basic Dataset Functionality

- Edit dataset general information
- Delete dataset
- Add sequences to dataset
- Copy dataset for personal use
- Add dataset to your basket
- View dataset active sequences
- View dataset quality values
- View dataset inactive sequences
- Export sequences, quality values, ass
- Run a FASTA using this sequence as

Sequence Manipulation and Annotation

- Run Sequence Cleanup Assistant
- Assemble dataset sequences

VBEST Assistant: Sequence Assembly

- Sequence Assembly Assistant »

Welcome!

This assistant will assemble your sequences for you. A new dataset will be created with the newly assembled sequences.

Please click the right arrow to get started.

Page 1