

Computational & Comparative Genomics  
Cold Spring Harbor, New York, United States  
9-15 November 2011

## BioMart

Bert Overduin, Ph.D.  
Vertebrate Genomics Team  
EMBL – European Bioinformatics Institute  
Wellcome Trust Genome Campus  
Hinxton, Cambridge, CB10 1SD, UK



## BioMart

- Data retrieval tool
- Originally developed for Ensembl (EnsMart)
- Now used by many large data resources
- Integrated with several widely used software packages
- Joint project between the European Bioinformatics Institute (EBI) and the Ontario Institute for Cancer Research (OICR)
- Central portal: <http://www.biomart.org>



bio::mart

HOME MARTVIEW MARTSERVICE DOCS CONTACT NEWS CREDITS

Try BioMart 0.8 release candidate 6

### BioMart Project

BioMart is a query-oriented data management system developed jointly by the [Ontario Institute for Cancer Research \(OICR\)](#) and the [European Bioinformatics Institute \(EBI\)](#).

The system can be used with any type of data and is particularly suited for providing 'data mining' like searches of complex descriptive data. BioMart comes with an 'out of the box' website that can be installed, configured and customised according to user requirements. Further access is provided by graphical and text based applications or programmatically using web services or API written in Perl and Java. BioMart has built-in support for query optimisation and data federation and in addition can be configured to work as a DAS 1.5 Annotation server. The process of converting a data source into BioMart format is fully automated by the tools included in the package. Currently supported RDBMS platforms are MySQL, Oracle and Postgres.



BioMart is completely Open Source, licensed under the LGPL, and freely available to anyone without restrictions.

**Powered by BioMart software:**

- [BioMart Central Portal](#)
- [ICGC Data Portal](#)
- [Ensembl](#)
- [Ensembl Bacteria](#)
- [Ensembl Metazoa](#)
- [Ensembl Protists](#)
- [Ensembl Plants](#)
- [Ensembl Fungi](#)
- [SalmonDB](#)
- [Phytozome](#)
- [Gramene](#)
- [Europhenome](#)
- [UniProt](#)
- [InterPro](#)
- [HGNC](#)
- [Rice-Map](#)
- [IKMC](#)
- [Wormbase](#)
- [DroSpeGe](#)
- [ArrayExpress DW](#)
- [Eurexpress](#)
- [HapMap](#)
- [Dictybase](#)
- [COSMIC](#)
- [IntOGen](#)
- [GermOnLine](#)
- [PRIDE](#)
- [PepSeeker](#)
- [VectorBase](#)
- [HTGT](#)
- [Cildb](#)
- [European Mouse Mutant Archive](#)
- [sigReannot](#)
- [Pancreatic Expression Database](#)
- [Reactome](#)
- [EU Rat Mart](#)
- [Paramecium DB](#)
- [International Potato Center \(CIP\)](#)
- [Mouse Genome Informatics \(MGI\)](#)
- [CyanoBase](#)
- [EMAGE](#)

**Third party software with BioMart Plugin:**

[Bioclipse](#) [BioExtract](#) [biomaRT-BioConductor](#) [Cytoscape](#) [Galaxy](#) [Gitools](#) [Ruby API](#) [Taverna](#) [WebLab](#)

EMBL-EBI   Ontario Institute for Cancer Research

Page last updated: 06/07/2011 16:09:59

## BioMart - Principle

- Step 1 – Dataset  
Choose your dataset and species
- Step 2 – Filters  
Limit your dataset
- Step 3 – Attributes  
Specify what information you want to output
- Step 4 – Results  
Preview and output your results

# Q&A

QUESTIONS

ANSWERS

## Worked example

Query:

Williams syndrome is a rare neurodevelopmental disorder caused by the deletion of genetic material from the region q11.23 of human chromosome 7

([http://en.wikipedia.org/wiki/Williams\\_syndrome](http://en.wikipedia.org/wiki/Williams_syndrome)).

Retrieve all protein-coding Ensembl genes that are located in this region.

Export the Ensembl Gene ID, name and description to an Excel spreadsheet.

# Exercises

## BIOMART

In Exercise 2 and 4, copy the list of IDs to be filtered for from the Exercises and answers pdf, which can be found on the course website