

PosMed (Positional Medline)

User Manual

RIKEN Bioinformatics And Systems Engineering (BASE) division
Division Director Tetsuro Toyoda

Index

1. Introduction.....	2
2. Specify the search conditions.....	4
3. Specify the display conditions	6
4. Understand the relationship between the candidate genes and the keyword	10
4-1. Relationship displayed in the “Simple mode”	10
4-2. Relationship displayed in the “Tree mode”	11
4-3. Relationship displayed in the “Graph mode”	11
5. Browsing the details of hit results	12
5-1. Relationships between the gene and the keywords	12
5-2. Relationship between the gene and the gene	13
5-3. Relationship to gene cluster associated with the gene	19
5-4. Relationship between the gene and the keywords	20

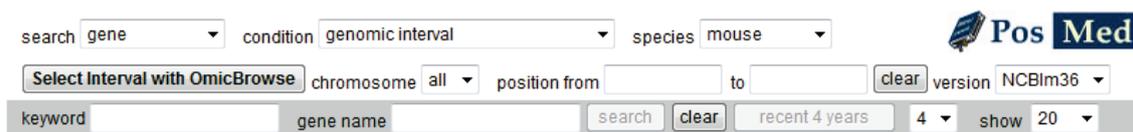
The URL of this website is <http://omicspace.riken.jp>

If you have any questions and comments, please feel free to send to
omicspace@gmail.com

1. Introduction

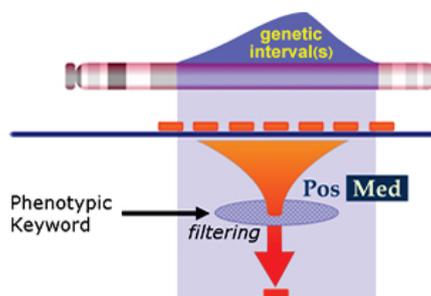
PosMed is RIKEN BASE's original system which possesses inference-type full text search functions. At first, PosMed performs a full-text search of the document database (containing MEDLINE abstracts etc) by using arbitrary keywords users have entered. And second, PosMed finds significant gene names or symbols that exist within the returned documents with Fisher's exact test and makes a ranked list (in this case, significant means that the P-value is as small as possible.) In the case where a user specifies the chromosomal interval, PosMed makes and displays a ranked list of genes restricted to the given interval. (That is to say, PosMed can search the connection of **Keyword -> Gene A -> Chromosomal interval.**)

Simultaneously, PosMed can also search for appropriate genes within the chromosomal interval that have an estimated gene-gene interaction, automatically. The selected genes will fulfill the conditions of network connections of **Keyword -> Gene B -> Gene C -> Chromosomal interval.** PosMed is a powerful tool which enables us to connect phenotypic functions and genes, because it uses relevant information of not only gene-gene interactions but also other biological objects' interactions (cell-gene, metabolite-gene, mutant mouse-gene, drug-gene, disease-gene, protein-protein interaction and ortholog data.) PosMed enables us to search human and mouse genes, drugs, diseases and document sets (*e.g.* MEDLINE and OMIM.)



Tutorial [[English](#)] [[Japanese](#)] [Sample trial \(search "diabetes" related genes on chr 1 in mouse\)](#)

PosMed: “Positional MEDLINE” assists your positional-cloning studies



Since whole genome sequences were first elucidated, knowledge-based ranking of candidate genes has become one of the most important bioinformatics tasks in the forward-genetics and positional-cloning approaches to identify phenotype-responsible gene mutations.

This task requires creating a form of artificial intelligence that can solve a genetic researcher's problems by learning computationally a vast amount of information accumulated in documents and published data.

We have developed a system named "PosMed," an artificial intelligence that guides you to the key information waiting to be discovered in the sea of data.

The search result page of PosMed

Search condition section

Display condition section

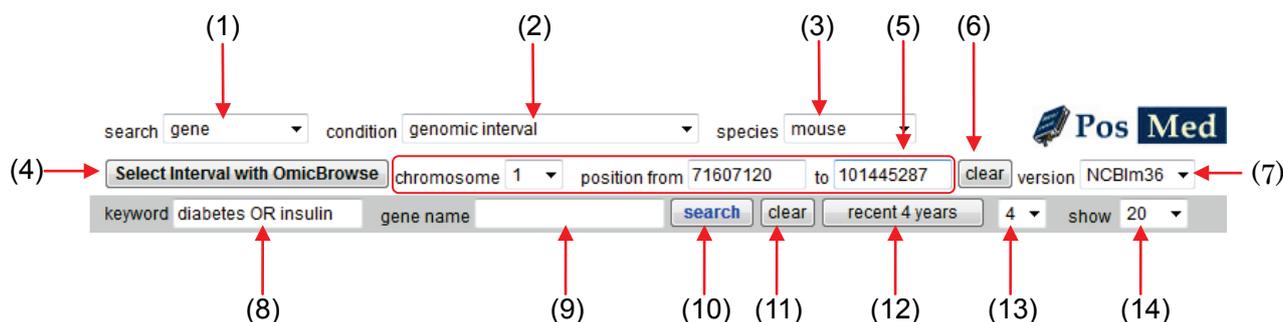
A list of the candidate genes

The hit list box

The maximum number of hit entries is 300 in the hit list box.

2. Specify the search conditions

The search condition section has some pull-down menus and entry boxes as below.



(1) Search target pull-down menu

The menu items are any, gene, metabolite, drug, cell, mutant, disease, document set, dictionary and database. According to the search target, each search tool shown previously will be displayed.

(2) Search condition pull-down menu

Available conditions are “gene name,” “genomic interval,” “multiple intervals,” “genomic interval (band),” “multiple intervals (band),” “accessions (*e.g.* MGI:107357, HGNC:6079),” “multiple accessions” and “genomic interval and accessions.”

(3) Species pull-down menu

The menus are any, mouse and human. (For this manual the “species” we use is “*M.musculus.*”)

(4) Select Interval with OmicBrowse button

At the stage of (2), when the menu condition is selected as “genomic interval” or “multiple intervals,” this button is displayed with (5). Click this to display OmicBrowse in the PosMed system.

(5) - [1]. Chromosome pull-down menu and base pair entry boxes

- [2]. Genomic band entry box
- [3]. Accessions entry box

At the stage of (2), when the menu condition is selected as “genomic interval” or “multiple intervals,” this part is displayed as [1] to specify genomic interval(s). In the case of “genomic interval (band)” or “multiple intervals (band),” this part is displayed as [2] to enter band name (*e.g.* 20p13). In the case of “accession” or “multiple accessions,” this part is displayed as [3] to enter accessions. And also in the case of “genomic interval and accessions,” both [1] and [2] are displayed at the

same time.

- (6) Clear button to reset chromosome pull-down menu and erase the base pair entry boxes
- (7) Pull-down menu specifying a genome version. (For this manual the “genome version” we use is “NCBI36.”)
- (8) Keyword entry box
PosMed enables us to specify more complex conditional equations using “OR (caps)”, “AND (caps)”, parentheses, double quotation marks for phrases, wild card (asterisk and question mark), “NOT (caps)” and “WHERE (caps).”

Keyword format in the Genome Browser & PosMed system

Keyword format	Genome Browser	PosMed
Union operator	OR (caps)	
Set intersection operator	AND (caps)	
Parenthesis expression	Available	
Phrase specification	group with “ ” <i>e.g.</i> “retinitis pigmentosa”	
Treatment of space	as “AND” operator	
Wild card	* (asterisk) <i>e.g.</i> diabet* matches <i>diabetes</i> , <i>diabetic</i> , etc.	
	? (question mark) <i>e.g.</i> ppar? matches <i>ppara</i> , <i>pparg</i> , <i>ppard</i> , etc	
Boolean operator	NOT (caps) <i>e.g.</i> diabetes NOT “type 1”	
Subquery	—	WHERE (caps)# <i>e.g.</i> plant WHERE synapse

#Note : PosMed searches mouse genes by “plant” and “synapse.” The search result of “plant” is displayed preferentially.

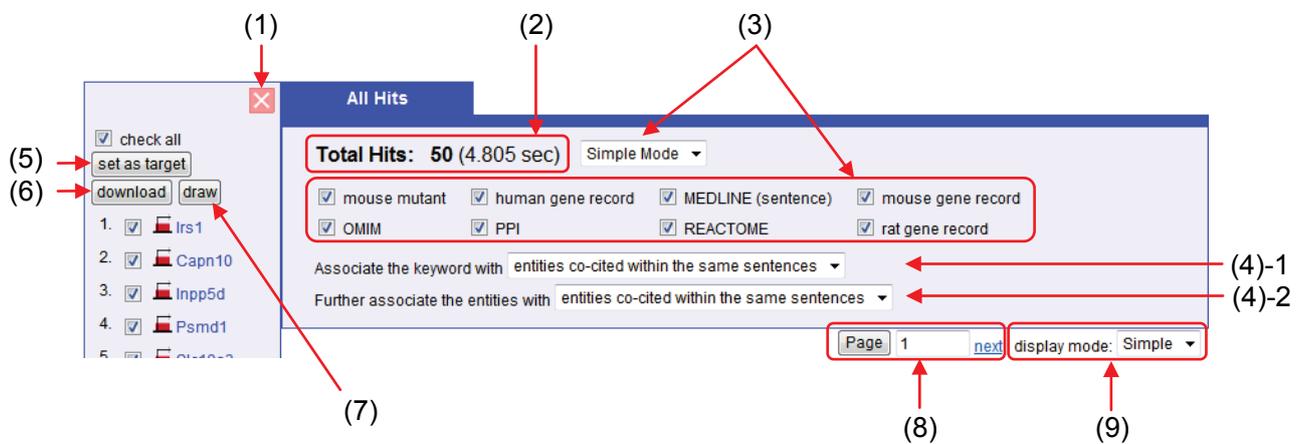
- (9) Gene name entry box to specify as determinate targets.
e.g. Irs1, “insulin receptor substrate 1”
- (10) Execute button to search
- (11) Clear button to erase the keyword entry box and the gene name entry box.

(12) Execute button to search with the specified number of years. The default number is set to 4. This button is linked with the “recent years pull-down menu” as below.

(13) Recent years pull-down menu (Available years are from 2 to 9)

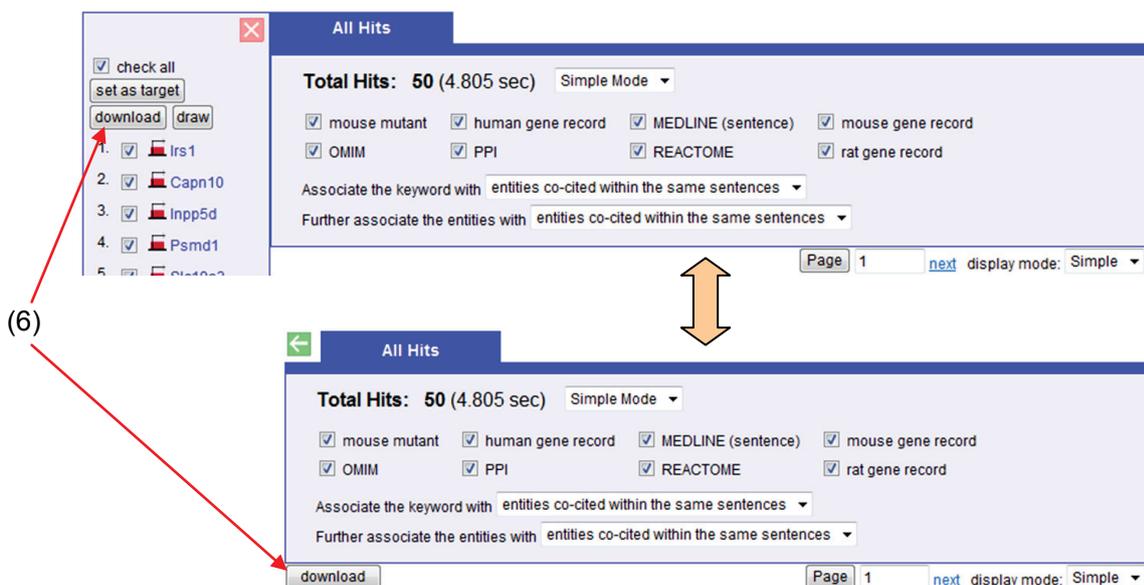
(14) Show pull-down menu (Available numbers are 5, 10, 20, 50, 100, 200 and 500)

3. Specify the display conditions



(1) Open all hits button

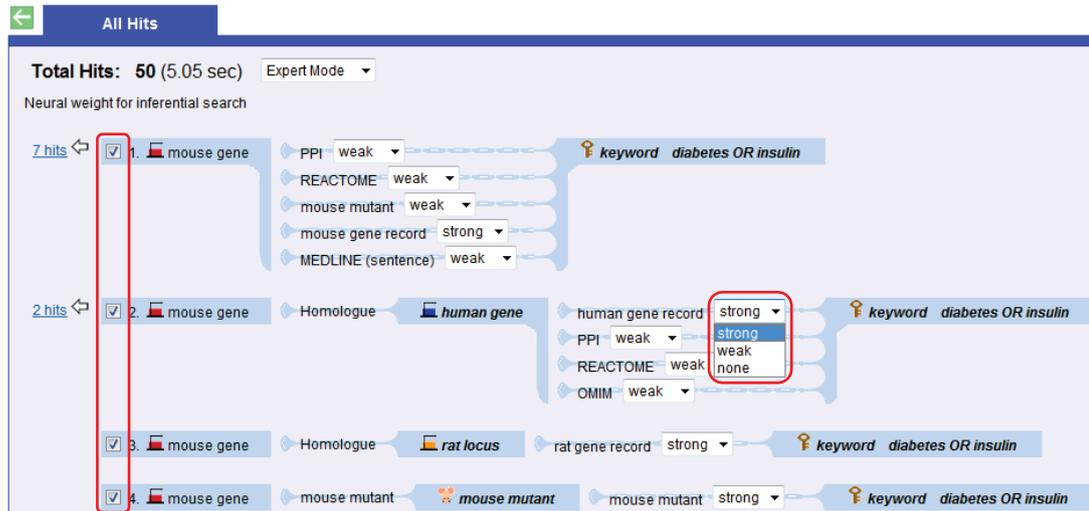
Click this button to show or close a list of genes. If you check some of the hits, you can set them as the next search targets by clicking “set as target” and download them as an excel file by clicking “download.”



(2) Number of hits and time taken for the search

(3) Search path display mode pull-down menu and path list

- Simple Mode : Users can check document sets for inference search process.
- Expert Mode : The display condition part expands with all search path with neural weight pull-down menu. Users can check path and select neural weight for inferential search process.



(4) Association pull-down menu

As above, PosMed can search the connection of **Keyword -> Gene A -> Gene B -> Chromosomal interval**. User can select the association status between the keyword and a single entity (gene A etc.) or two entities. The association status has can be set to either “co-cited within the same sentences” or “documents”.

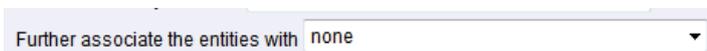
(4)-1 represents the association status between the keyword and Gene A.

The default is set to the association of “entities co-cited within the same sentences.”

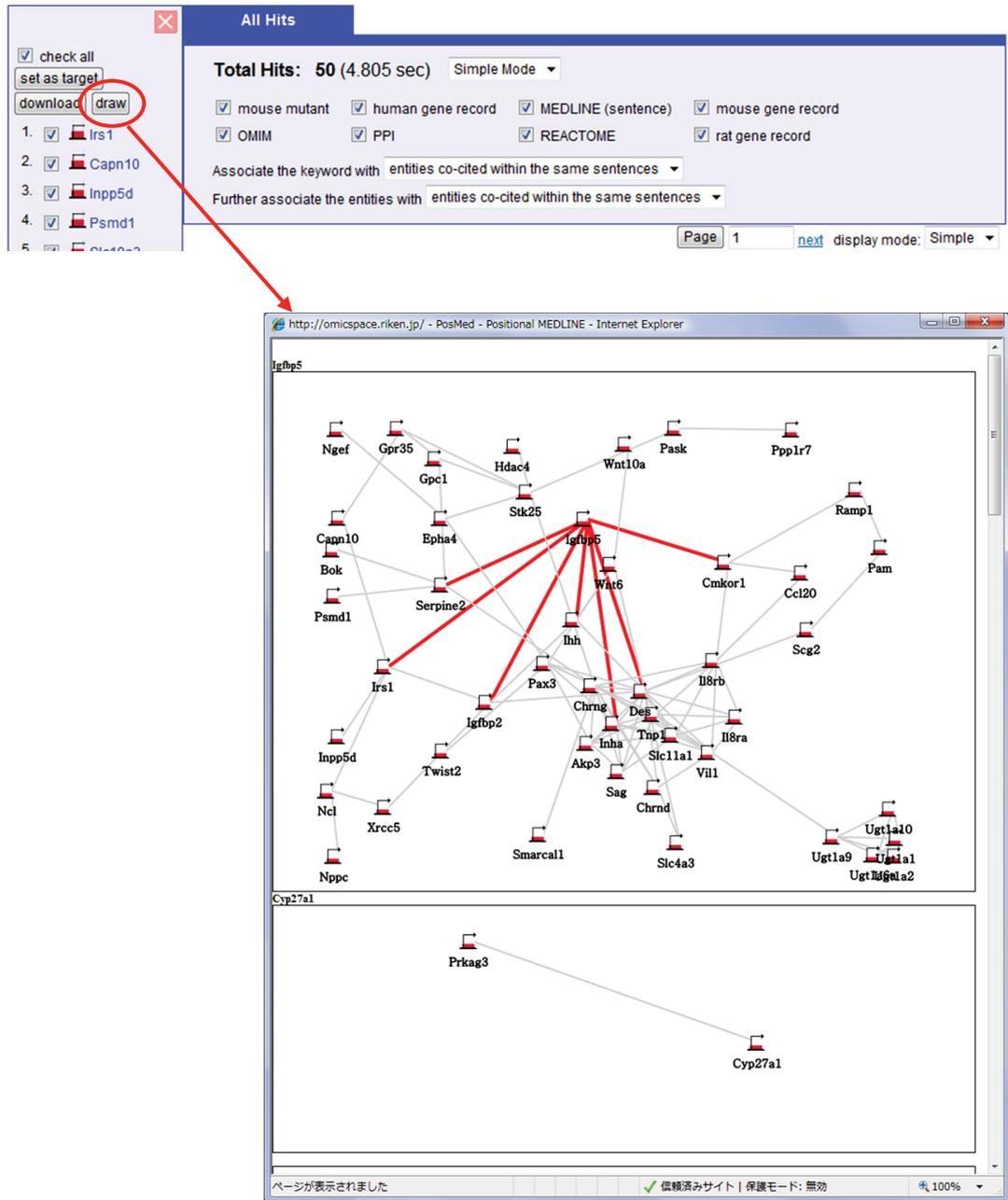


(4)-2 represents the association status between Gene A and Gene B.

The default is set to the association of “none.”



- (5) Set as target button to set them as the next search targets.
- (6) Download button to download all hits as an Excel file.
- (7) Draw button to draw the co-citation networks of hit entries in another window.



- (8) Page entry box and link to next (previous) page

(9) Search result display mode (mentioned in next section)

- “Simple” display mode shows a compact list of the search results with minimum information. Each result indicates the search path from keyword (bottom) to gene (top).

6. [Adipor1 , adiponectin receptor 1](#)

Adipor1 ← 172 docs P value: 1.82E-499 Position: [Mm:1:136231891-136248748](#) Link to: [MGI](#) [CAGE](#)
↑ ← 244 docs P value: 1.82E-499
Adipoq ← 3773 docs Position: [Mm:16:23061870-23073302](#) Link to: [MGI](#) [CAGE](#)
↑ ← 4378 hits P value: 2.49E-4632
♀ diabetes OR insulin

- “Tree” display mode shows a compact list of the search results in a hierarchical structure with minimum information similar to the “Simple mode.”

6. [Adipor1 , adiponectin receptor 1](#)

Adipor1 172 docs P value: 1.82E-499 Position: [Mm:1:136231891-136248748](#) Link to: [MGI](#) [CAGE](#)
└─ 244 docs P value: 1.82E-499
└─ Adipoq 3773 docs Position: [Mm:16:23061870-23073302](#) Link to: [MGI](#) [CAGE](#)
└─ 4378 hits P value: 2.49E-4632
└─ ♀ diabetes OR insulin

- “Graph” display mode shows a list of the search result with more information. Each result indicates the search path from keyword (right) to gene (left).

6. [Adipor1 , adiponectin receptor 1](#)

Interval ↔ Mouse Locus ↔ Co-citation ↔ Mouse Locus ↔ Keyword

Interval
DNA double helix icon

Mouse Locus
Gene track icon with [show graph](#) link
Symbol: Adipor1
Name: adiponectin receptor 1
P value: 1.82E-499
ID: MGI:1919924 [set as target](#)
172
Link to: [MGI](#) [CAGE](#)
Position: [Mm:1:136231891-136248748](#)

Co-citation
Book icon
P value: 1.82E-499
244

Mouse Locus
Gene track icon with [show graph](#) link
Symbol: Adipoq
Name: adiponectin, C1Q and collagen domain containing
Other aliases: adipoQ, adiponectin, GBP28, apM1, adipo, Acrp30, Acdc
ID: MGI:106675
3773
Link to: [MGI](#) [CAGE](#)
Position: [Mm:16:23061870-23073302](#)

Keyword
Key icon
P value: 2.49E-4632
Keyword: diabetes OR insulin

4. Understand the relationship between the candidate genes and the keyword

4-1. Relationship displayed in the “Simple mode”

The order from (1) to (5) indicates the search path from keywords (bottom) to genes (top).

(1) A keyword icon indicates the keyword(s) entered in the keyword entry box previously.

(2) An arrow icon indicates the relationship between the keyword(s) and a gene with a p value.

(3) A Mouse gene icon indicates an intermediate path has relationships with both the keyword(s) and gene.

(4) An arrow icon indicates relationships between two mouse genes.

(5) A Mouse gene icon indicates a gene within the interval.

4-2. Relationship displayed in the “Tree mode”

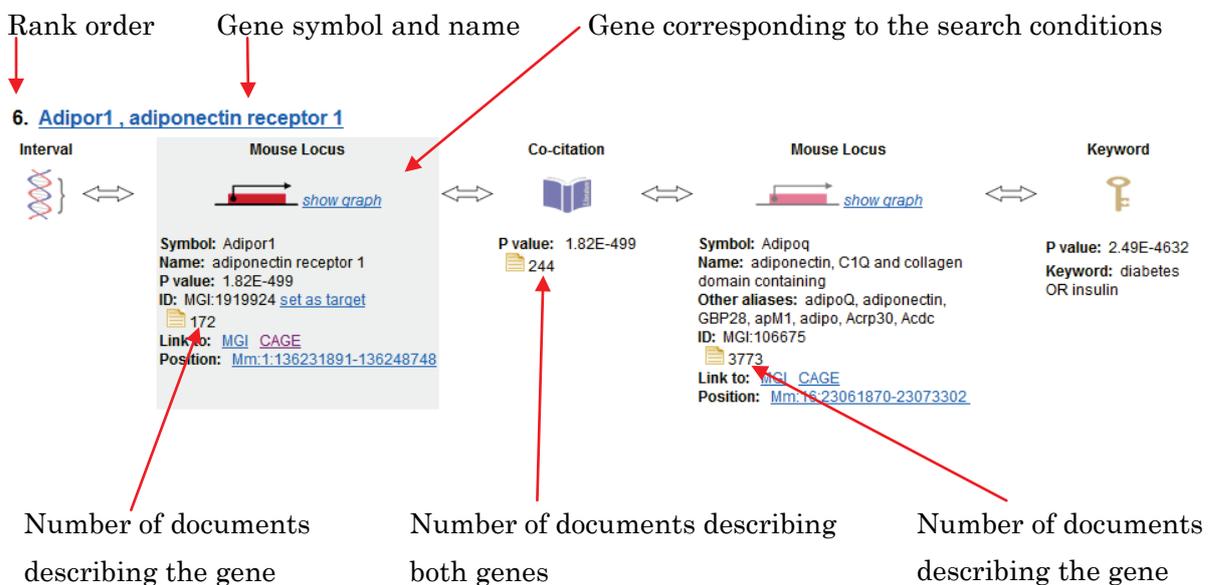
In the “Tree mode”, the hierarchy structure indicates the search path from keywords (bottom) to genes (top) in much the same way as the “Simple mode.”

6. [Adipor1 , adiponectin receptor 1](#)



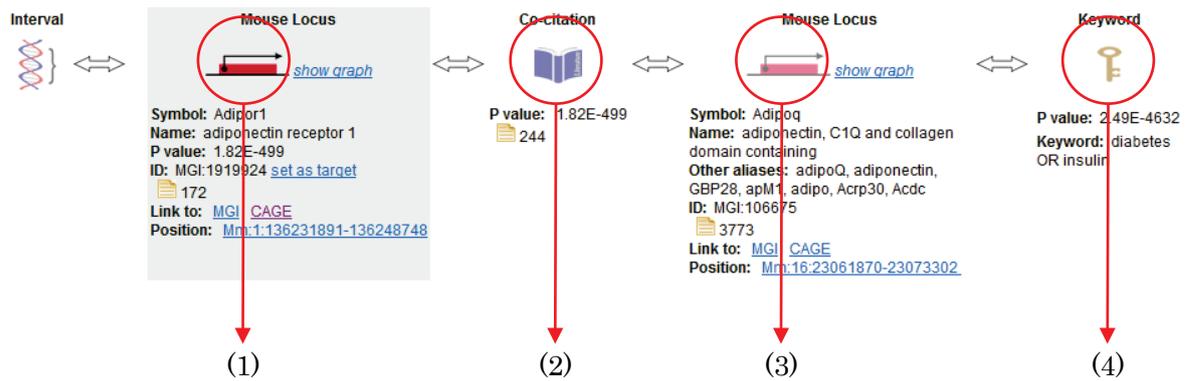
4-3. Relationship displayed in the “Graph mode”

In the “Graph mode”, each result indicates the search path from keywords (right) to genes (left).



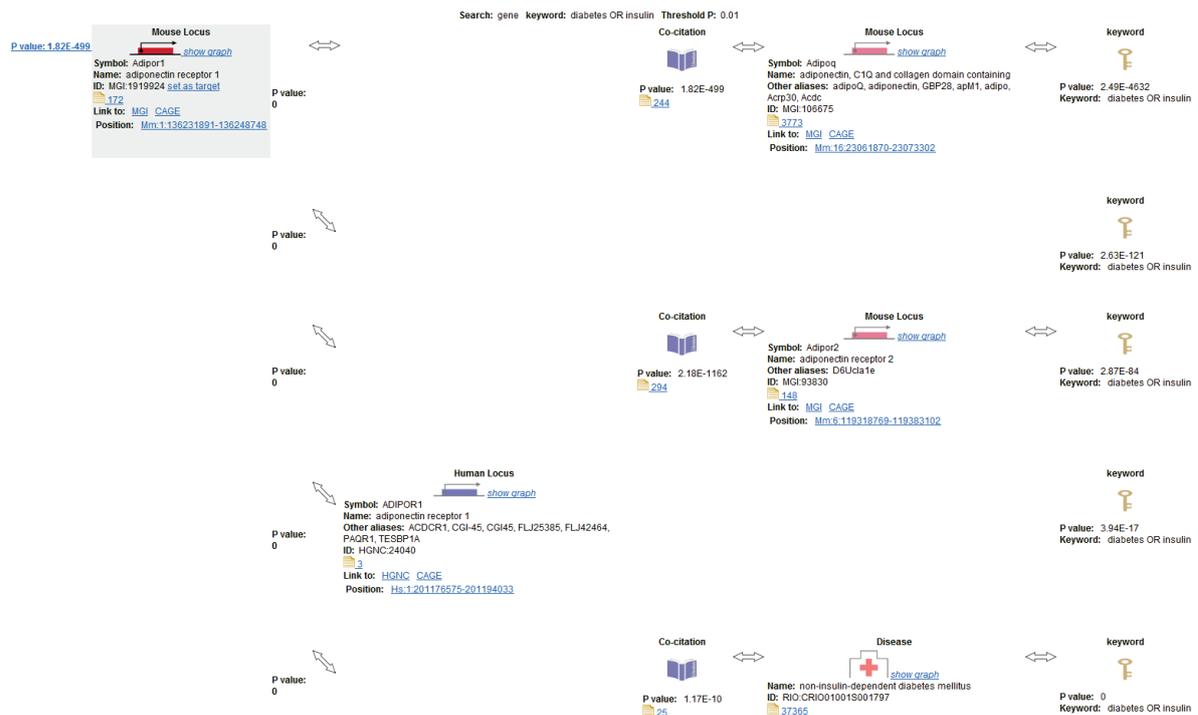
5. Browsing the details of hit results

6. Adipor1 , adiponectin receptor 1



5-1. Relationships between the gene and the keywords

When you click (1), the gene, various inference relationships from the gene to the corresponding biological objects are displayed.



5-2. Relationship between the gene and another gene

When you click (2), the relationship, document digests that include both the genes are displayed.

In the document digest display, the gene names and the keywords are highlighted with light blue (*Adipor1* etc.), light green (*Adipoq* etc.) and red (*diabetes* or *insulin*) respectively. Additional biological objects are displayed in gray. In the default display sequence, the documents including the keyword(s) are ranked higher. Click a document number in order to transfer to [PubMed](#) at NCBI.

The hit list box The hit and related molecules tabs

The hit list box

The hit and related molecules tabs

Related molecules and diseases box

Documents section

5-2-1. Hit and related molecules tabs

These tabs display the descriptions of hit molecules (Adipor1 etc.), related molecules (Adipoq etc) and their relationships. Click a tab to display each description and change information of the “related molecules and disease box” and the “documents section” respectively.

The figure illustrates the navigation between different molecule tabs in a database interface. It shows three sequential views:

- Adipor1 Tab:** The 'Adipor1' tab is selected. It displays the molecule's symbol (Adipor1), name (adiponectin receptor 1), and various identifiers (HGNC, CAGE, MGI). A bar chart shows the number of documents published from 1999 to 2008, with a legend for 'with keyword' (red) and 'without keyword' (blue).
- Relation Tab:** The 'Relation' tab is selected, showing the relationship between Adipor1 and Adipoq. It includes statistics for both molecules, such as P value, hits, and mouse mutants.
- Adipoq Tab:** The 'Adipoq' tab is selected. It displays the molecule's symbol (Adipoq), name (adiponectin, C1Q and collagen domain containing), and various identifiers. A bar chart shows the number of documents published from 1999 to 2008.

5-2-2. Hit document tabs

These tabs indicate the number of hit documents in each [document set](#). In the case of mouse, the document sets for the inference search process are mutant information, [MEDLINE](#), gene description of [MGI](#), PPI description of the [HPRD](#) and REACTOME description of [BioPAX](#). The number of hit documents which contain the keywords are highlighted in red. The default setting is all.

all (3038/3773)	mouse mutant (3/4)	HsPPI (0/2)	MEDLINE (3035/3766)	mouse gene record (0/1)	REACTOME (0/0)
-----------------	--------------------	-------------	---------------------	-------------------------	----------------

Click a tab to display the corresponding document digests as described above.
(e.g. The transition from “all” to “mouse mutant.”)

The screenshot shows the search interface with the 'all' tab selected. The search results include a bar chart showing document counts from 1999 to 2008, and a list of related entities for Adipor1.

publication year	with keyword	without keyword
unknown	0	0
1999	0	0
2000	0	0
2001	0	0
2002	0	0
2003	0	0
2004	0	0
2005	0	0
2006	0	0
2007	0	0
2008	0	0

The screenshot shows the search interface with the 'mouse mutant' tab selected. The search results include a bar chart showing document counts from 1999 to 2008, and a list of related entities for Adipor1 via mouse mutant.

publication year	with keyword	without keyword
unknown	0	0
1999	0	0
2000	0	0
2001	0	0
2002	0	0
2003	0	0
2004	0	0
2005	0	0
2006	0	0
2007	0	0
2008	0	0

5-2-3. Related molecules and diseases box

Other related molecules or diseases are listed as images in the left yellow box. If the documents contain the keywords, the related objects will be displayed with hit images . You can click a name (e.g. Adipor2) to display the relationships, document digests that include the keywords (**diabetes** or **insulin**), the gene name (**Adipor1** etc.) and the biological object (e.g. **Adipor2** etc.) in the documents section. This function is the same as the “Hit and related molecules tabs” described previously.

The screenshot shows a web interface with tabs for 'All Hits', 'Adipor1', 'Relation', and 'Adipor2'. The 'Relation' tab is active, displaying details for Adipor1 and Adipor2. Adipor1 details include Symbol: Adipor1, Name: adiponectin receptor 1, ID: MGI:1919924, and Position: Mm:1:136231891-136248748. Adipor2 details include Symbol: Adipor2, Name: adiponectin receptor 2, Other aliases: D6Ucla1e, ID: MGI:93830, and Position: Mm:6:119318769-119383102. A bar chart shows document counts from 1999 to 2008, with a legend for 'with keyword' (red) and 'without keyword' (blue). A list of 'Adipor1 related entities' includes Adipoq, Adipor2 (circled in red), non-insulin-dependent dia..., ADIPOR1, Adipor1, Deltagen and Lexicon Knoc..., Edn1, Amh, Mapk8, and Agrp. A search box contains 'diabetes OR insulin' and 'year all'. A document digest for PMID:17607322 is shown, with the 'Find' button circled in red.

The diagram illustrates the process of selecting a condition from a pull-down menu. The top part shows the 'Adipor1 related entities' list with the 'Metabolite' condition selected. An arrow points to the 'Find' button, which is circled in red. The bottom part shows the 'Keyword' search box with 'Glycerol' entered and the 'go' button clicked.

Top fifty related molecules and diseases are displayed as biological objects initially. You can select a particular condition from the pull down menu. (e.g. Metabolite)

Click the “Find” image to display the keyword search box. Click the go button to find a biological object of interest in this box. (e.g. Glycerol)

5-2-4. Documents section

The documents section has three convenient functions to find desired documents. The functions are the “show” pull-down menu, “order by” pull-down menu, “year” pull-down menu and “Find.”

keyword year

Year	with keyword	without keyword
unknown	0	0
1999	0	0
2000	0	0
2001	0	0
2002	0	0
2003	2	0
2004	8	2
2005	15	5
2006	25	10
2007	32	10
2008	12	5

show order by

1-20 of 128 Page 1

1: Pattern of expression of adiponectin receptors in human adipose tissue depots and its relation to the metabolic state.
[PMID:17607322](#)
 International journal of obesity (2005) 2007 Dec
OBJECTIVE: To investigate whether [adiponectin](#) receptor genes ([AdipoR1](#) and [AdipoR2](#)) expression in human subcutaneous (SAT) and visceral (VAT) adipose tissue in severely obese patients with or without **diabetes** is related to [adiponectin](#) gene ([APM1](#)) expression and in vivo metabolic parameters.

The “keyword” pull-down menu enables us to change the display sequence of the hit documents by keyword(s) or biological terms. The default is the previously entered keyword(s) in this case “diabetes OR insulin.” Select “physical” to superordinate documents including physical terms (*e.g.* **binds**) and highlight these terms.

keyword year

Year	with keyword	without keyword
unknown	0	0
1999	0	0
2000	0	0
2001	0	0
2002	0	0
2003	0	0
2004	0	10
2005	0	20
2006	0	35
2007	0	45
2008	0	15

show order by

1-20 of 128 Page 1

1: Adiponectin and its role in cardiovascular diseases.
[PMID:18336252](#)
 Cardiovascular & hematological disorders drug targets 2008 Mar
[Adiponectin](#) **binds** two receptors, [AdipoR1](#) and [AdipoR2](#).

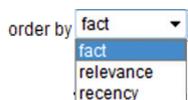
year

- all
- unknown
- 1999
- 2000
- 2001

The published year of documents is narrowed down by the “year” pull-down menu. You can select years from 1999 to the current year or any with unknown published years. The default setting is “all.”



The “show” pull-down menu has two modes: the default is the digest mode to display the minimum sentences including keyword(s) and biological objects, and the other one is the all mode to display whole text.



The “order by” pull-down menu enables us to change the display sequence of the hit documents by biological objects, keyword(s) or their recentness. The default is the “fact” mode to display sentences including previously selected biological objects, in this case [AdipoR1 \(Adipor1\)](#) and the biological object [AdipoR2 \(Adipor2\)](#). Select “relevance” to superordinate documents including keywords ([diabetes](#) or [insulin](#)) and highlight these terms. The “recency” mode superordinates the latest sentences.

You can download the list of all documents by clicking the “download” button. The list has information about the PMID (PubMed identifier), title, name of the journal, published year and date, abstract and links to NCBI.

Click the “Find” image to display the keyword search box. Click the go button to superordinate documents including the highlighted keyword of interest. (e.g. [lipid](#))

The screenshot shows a search interface. On the left, a 'Keyword' search box contains the word 'lipid' and has 'go' and 'close' buttons. An arrow points to the right, where a search results page is displayed. At the top, there are dropdown menus for 'keyword' (set to 'physical') and 'year' (set to 'all'), and a 'Find' button. Below this is a bar chart showing the number of documents published from 1999 to 2008, categorized by 'with keyword' (red) and 'without keyword' (blue). The chart shows a significant increase in documents with the keyword starting in 2006. Below the chart are controls for 'show' (set to 'digest'), 'order by' (set to 'fact'), and a 'download' button. At the bottom right, it says '1-20 of 128' and 'Page 1'. The search results list two main entries:

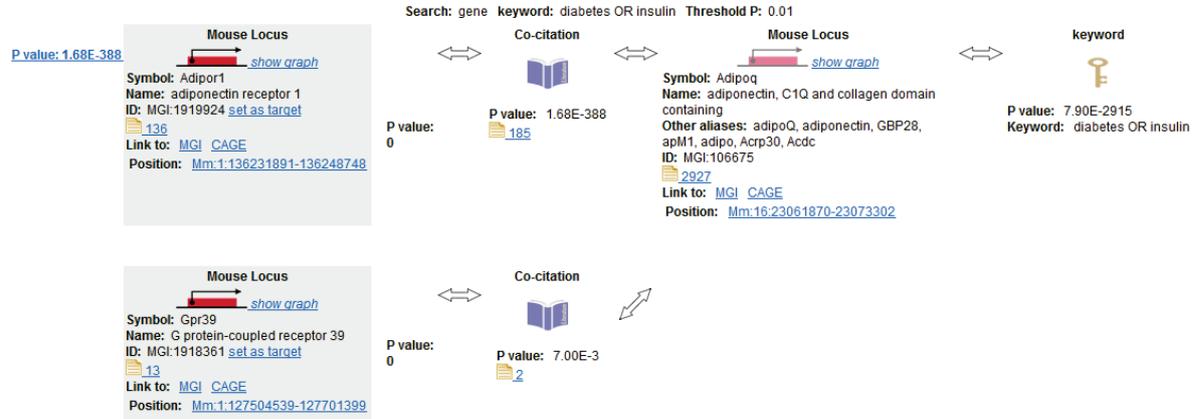
- 1:** Tissue-specific effect of dietary cysteamine on expression of adiponectin receptors in rats.
 PMID:17696487
 Journal of agricultural and food chemistry 2007 Sep 19
 Adiponectin is synthesized by adipocytes and affects glucose and lipid metabolism by binding to its receptors, [AdipoR1](#) and [AdipoR2](#).

? [Adipor1](#), adiponectin receptor 1 ([references](#), [MGI:1919924](#), [CAGE](#), [Mm:1:136231891-136248748](#))
 ? [Adipog](#), adiponectin, C1Q and collagen domain containing ([references](#), [MGI:106675](#), [CAGE](#), [Mm:16:23061870-23073302](#))
 ? [Adipor2](#), adiponectin receptor 2 ([references](#), [MGI:93830](#), [CAGE](#), [Mm:6:119318769-119383102](#))
- 2:** Gene expression of adiponectin receptors in human visceral and subcutaneous adipose tissue is related to insulin resistance and metabolic parameters and is altered in response to physical training.
 PMID:17878241
 Diabetes care 2007 Dec
 OBJECTIVE: Adiponectin receptors 1 and 2 ([AdipoR1](#) and [AdipoR2](#), respectively) mediate the effects of adiponectin on glucose and lipid metabolism in vivo.

? [Adipor1](#), adiponectin receptor 1 ([references](#), [MGI:1919924](#), [CAGE](#), [Mm:1:136231891-136248748](#))
 ? [Adipog](#), adiponectin, C1Q and collagen domain containing ([references](#), [MGI:106675](#), [CAGE](#), [Mm:16:23061870-23073302](#))
 ? [Adipor2](#), adiponectin receptor 2 ([references](#), [MGI:93830](#), [CAGE](#), [Mm:6:119318769-119383102](#))

5-3. Relationship to gene cluster associated with the gene

When you click (3), gene (Adipoq), a list of relationships to genes associated with the gene is displayed.



5-4. Relationship between the gene and the keywords

When you click (4), the keyword icon, document digests describing both the keywords and the gene (Adipoq, adiponectin etc.) are displayed. In the document digest display, the gene name and the keywords are highlighted with light green (**adiponectin**) and red (**diabetes or insulin**) respectively. In the default display sequence, documents including the keyword(s) are ranked higher. Click a document number in order to transfer to PubMed at NCBI.

The screenshot displays a web interface for the gene **Adipoq**. At the top, there are tabs for "All Hits" and "Adipoq". Below the tabs, a "show graph" button is visible. The main content area provides details for the gene: Symbol: Adipoq, Name: adiponectin, C1Q and collagen domain containing, Other aliases: adipoQ, adiponectin, GBP28, apM1, adipo, Acrp30, Acdc, ID: MGI:106675, and counts of 3773 and 3038. It lists the Human, Mouse, and Rat Locus information with corresponding database links.

Below the locus information, it states: "Reportedly, **Adipoq** is associated with the keyword as follows." This is followed by a table of database counts:

all (3038/3773)	mouse mutant (3/4)	HsPPI (0/2)	MEDLINE (3035/3766)	mouse gene record (0/1)	REACTOME (0/0)	other databases
-----------------	--------------------	-------------	---------------------	-------------------------	----------------	-----------------

The interface also features a search section for "Adipoq related entities" with a "Find" button and a list of related genes including Lep, Ghrl, Pzp, a, Cyp2e1, Drd2, Vamp2, Gpr120, Insig2, Tlr2, Csf3, Csf2, and Cxd11.

A bar chart shows the number of documents published from 1999 to 2008, categorized by "with keyword" (red) and "without keyword" (blue). The x-axis is "publication year" and the y-axis is "documents".

Year	With keyword	Without keyword
unknown	0	0
1999	0	0
2000	0	0
2001	0	0
2002	~10	0
2003	~100	~100
2004	~200	~200
2005	~400	~200
2006	~600	~200
2007	~800	~200
2008	~500	~100

Below the chart, there are controls for "show digest", "order by fact", and "download". A pagination bar shows "1-20 of 300" and "Page 1".

The first document listed is: **1: Adiponectin and risk of coronary heart disease in older men and women.** PMID:18593765. The abstract mentions: "RESULTS: **Adiponectin** exhibited significant negative correlations with baseline adiposity, **insulin** resistance, dyslipidemia, inflammatory markers, and **leptin**." Below this, related terms like "Rouqnon-Heberden disease" and "Cholesterol" are listed with their respective reference counts.

Additional biological objects are highlighted in gray. (e.g. leptin) Other related molecules are listed in the left yellow box. Click the images to display the relationships and to find further related biological objects.