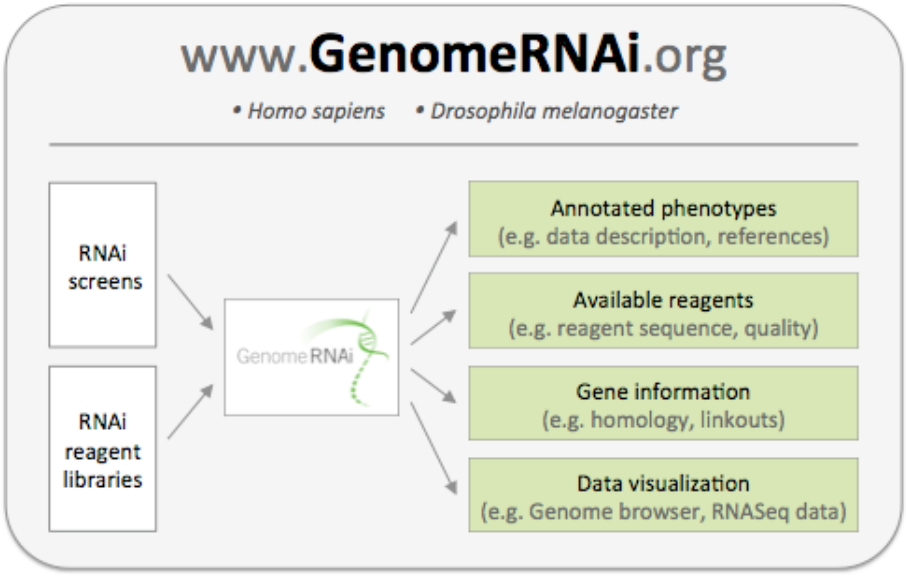
Home

GenomeRNAi database

The GenomeRNAi database currently contains phenotype information derived from RNA interference (RNAi) screens in *Drosophila melanogaster* and *Homo sapiens*. The database also provides an updated resource of RNAi reagents and predicted quality control information for libraries available for the *Drosophila* and the human genome. The database is accessible at www.genomernai.org.



Contact

Contact the GenomeRNAi team for suggestions and questions.

Reference

Gilsdorf M, Horn T, Arziman Z, Pelz O, Kiner E, Boutros M. GenomeRNAi: a database for cell-based RNAi phenotypes. 2009 Update. Nucleic Acids Res. 2010 Jan;38(Database issue):D448-52. Epub 2009 Nov 12.

Go to GenomeRNAi

or visit the following site to read more about it.

Database contents

Database contents

The GenomeRNAi database integrates information about RNAi reagents, their annotated targets and phenotypic information based on RNAi screens in *Drosophila* and human tissue culture as well as *in vivo Drosophila* screens.

Please read more about our latest release (September 2012, version 9.0) on the following page.

Reagents

The database contains information such as reagent sequence, reagent primers and primer parameters, predicted sequence specificity and predicted efficiency for:

- more than 140.000 constructs from eight RNAi libraries for *Drosophila*

- Ambion RNAi library
- BKN RNAi library (Heidelberg2 library)
- Drosophila RNAi Screening Center (DRSC) RNAi library (v2)
- HFA RNAi library (Heidelberg1 library, DRSC library v1)
- MRC RNAi library
- OpenBiosystems RNAi library
- Vienna Drosophila RNAi Center (VDRC) *in vivo* P-Element RNAi Library (GD)
- Vienna Drosophila RNAi Center (VDRC) *in vivo* phiC31 RNAi Library (KK)

- more than 300.000 RNAi reagents from six human RNAi libraries

- Ambion Silencer Select (siRNA)
- Dharmacon/ThermoFisher siGENOME (siRNA)
- The RNAi Consortium (TRC)/Sigma-Aldrich (shRNA)
- Qiagen druggable/whole genome supplement (siRNA)
- OpenBiosystems shRNAmir library
- Sigma Proligo library

Genes

The database provides information about gene names, location, homologs (to *Drosophila melanogaster* or *Homo sapiens*) and linkouts to other databases such as Entrez, Ensembl, FlyBase, NCBI RefSeq, UniProt, GeneCards, etc.

Latest annotations for the *Drosophila* and human genomes are obtained prior to each release from NCBI and FlyBase.

Phenotypes

GenomeRNAi contains phenotype data from:

- 127 screens performed in human cells,
- 117 screens performed in *Drosophila* cells and
- 53 *in vivo Drosophila* screens.

A list of all currently available screens can be accessed here.

In the process of updating pre-existing data and applying the new annotation guidelines, some screens have been removed from the GenomeRNAi database. Screen data that have been removed from the database have been archived and can be downloaded here.

Curation and annotation

Curation and annotation

We aim to provide comprehensive coverage of large-scale RNAi screening data in the literature. Priority is given to data of high scientific impact. Due to limited capacities, we currently focus on data from *Homo sapiens* and *Drosophila melanogaster*.

Annotation guidelines

Annotation guidelines are essential for conclusive comparisons across various datasets based on the application of different assay types. We defined controlled vocabularies and postulated vocabulary guidelines for each database field, in order to ensure a comparable description of experimental and phenotypic data within the GenomeRNAi database.

Update of pre-existing data

Some of the screening data in the database still need to be updated to the current annotation standard (pre-existing data, having GenomeRNAi IDs ending on '-A-0'). Application of the new annotation guidelines to these data is ongoing.

Data processing

Data processing

Release cycle

An update of the GenomeRNAi database is performed every 3 - 4 months. Typically, pure data releases alternate with releases that also include new features to the website. For every release the data in GenomeRNAi is mapped to the latest NCBI gene annotation to ensure currency of the data.

Mapping procedures

When collecting RNAi phenotype data from the literature, the diversity of gene and /or reagent identifiers used by the respective authors represents a challenge. In order to be able to display and compare data from different screening experiments, a common reference gene identifier is needed. Therefore we attempt to map the author-provided gene/reagent identifiers to the Entrez database and use the Entrez Gene ID as the GenomeRNAi reference identifier.

Update of gene information

Before every release we download the latest NCBI gene table as well as the HomoloGene mappings. UniProt entries are updated according to the “idmapping” data files provided by the UniProt ftp site.

Reagent-to-gene mapping

We map RNAi reagent sequences to the latest gene annotation using the NEXT-RNAi software tool, which additionally provides quality assessments with respect to specificity and efficiency of RNAi reagents.

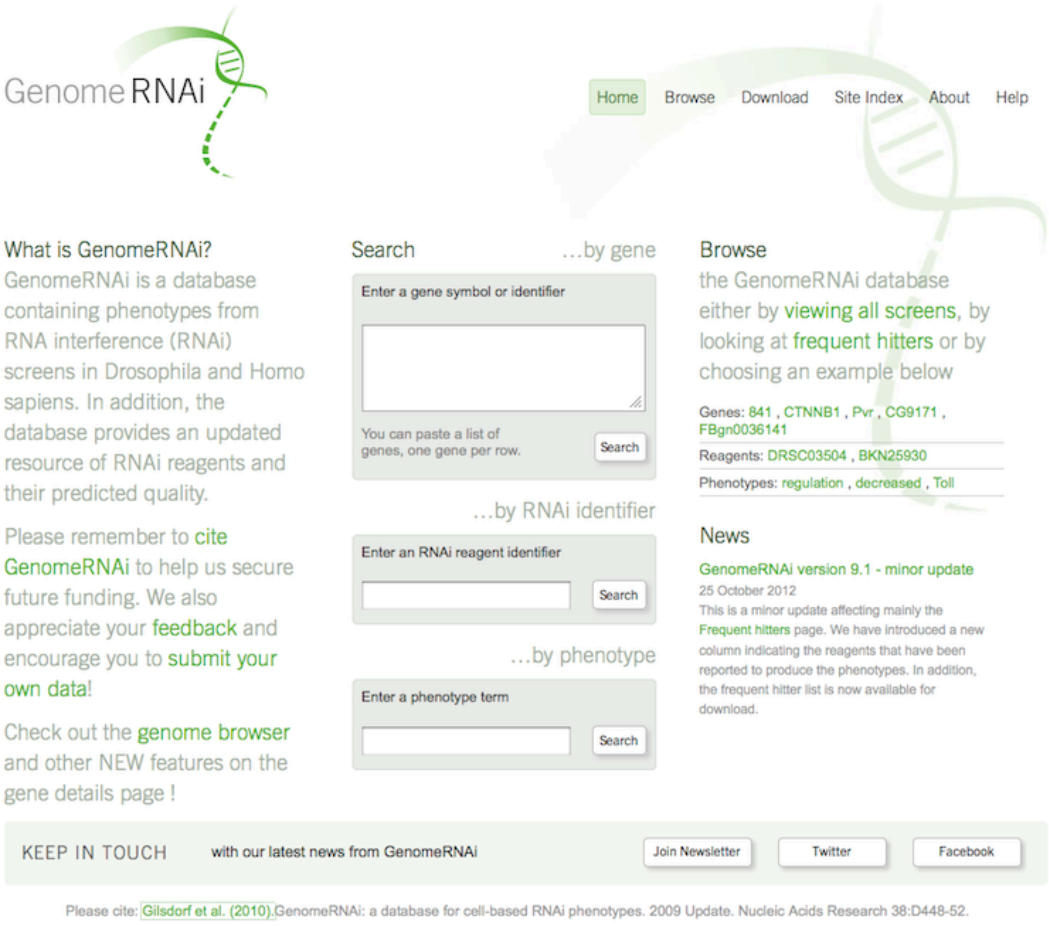
Phenotype mapping

Finally, we attempt to map all author-provided gene identifiers and / or gene symbols to Entrez Gene IDs and establish the gene – phenotype relationships. Reagent – phenotype relationships are directly extracted from the author-provided data.

Home page overview

Home page overview

The image below shows the GenomeRNAi home page:



The user can query the database for genes, RNAi reagents and RNAi phenotypes (**search boxes** in the middle column). All searches are performed in an organism-independent manner, **examples** of search terms are shown in the right column next to the search boxes:

- A gene can be queried by its symbol / name or by an identifier such as Entrez, ENSEMBL or FlyBase identifier. An incomplete gene name can be used to search for gene families. Moreover, gene batch query is now possible - a list of genes (one gene per row) can be used for database search.
- A RNAi reagent can be queried by its reagent ID. This search will only identify exact matches in the database.
- Specific phenotypes can be queried by entering phenotype terms. As for the gene search, incomplete queries will also be matched to the database entries.

A list of all available screens / publications in the GenomeRNAi database can be obtained by the **'Browse' tab** in the navigation menu or the **'viewing all screens' link** in the right column.

The **'Site Index' tab** leads you to an overview of the GenomeRNAi pages.

The GenomeRNAi screens and phenotypes can be downloaded via the **'Download' tab** in the navigation menu.

Furthermore, we encourage the submission of RNAi phenotype data to the GenomeRNAi database (please follow the **'Submit your own RNAi phenotype data' link** in the left column).

The **'Frequent hitters' page** (link in the right column or via the Site Index button) provides lists of genes that frequently show a phenotype.

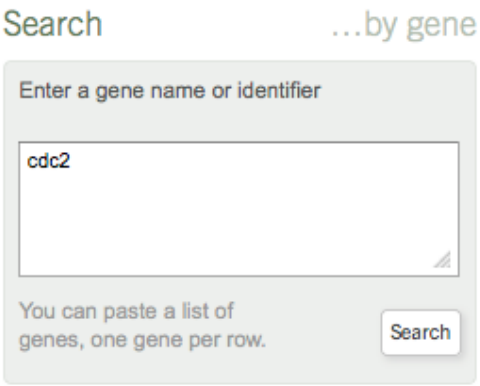
You can learn more about GenomeRNAi clicking on the **'About' tab** in the navigation menu - and to be always up-to-date, use the possibilities shown in the **'Keep in touch'-box** at the bottom of the page.

Search for genes

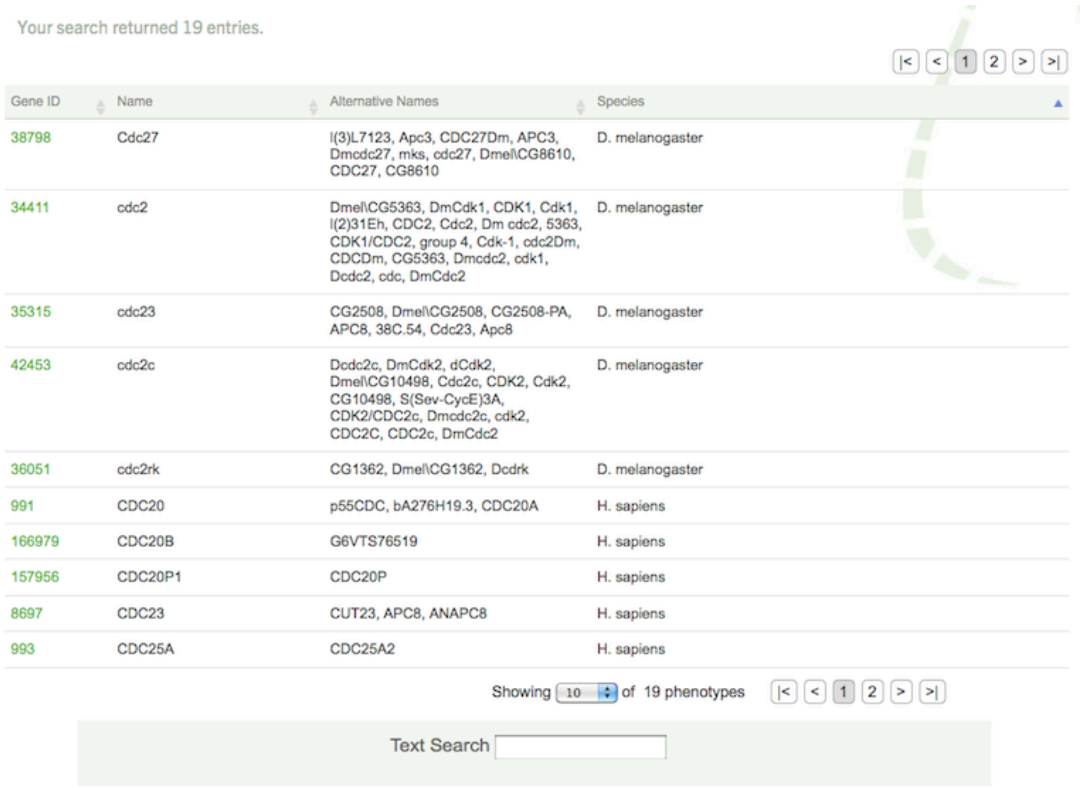
Search for genes

The search for genes is performed in an organism-independent manner. A gene can be queried in the gene search field by its official NCBI gene symbol (Entrez Gene ID), by synonym search or by search by other identifiers such as ENSEMBL or FlyBase identifiers. Gene batch query is possible - a list of genes (one gene per row) can be used for a database search. The result of a gene search will be all entries in the database matching the query. A search for e.g. cdc2 will identify *Drosophila* cdc2, cdc2rk, cdc2c, cdc23 and Cdc27 as well as human CDC genes and can be used to search for incomplete gene names or gene families.

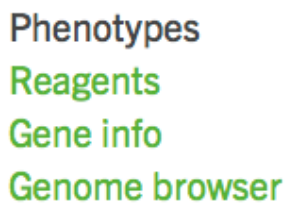
In the example shown below, GenomeRNAi was queried for the *Drosophila* gene *cdc2*.



In case the database identified more than one result matching the gene query, a list showing all identified hits will be provided. The database identifies 19 genes matching this query (5 *Drosophila* and 14 human). In addition the species, the gene IDs and the alternative names of the matched genes are displayed.



By clicking on the gene ID the user is directed to the gene details page. This page is subdivided in four tabs: 'Phenotypes', 'Reagents', 'Gene info' and 'Genome browser'.



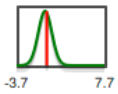
The following sections will explain these outputs in detail:

Phenotypes tab

The literature curated phenotypes for cdc2 (see below) include 55 entries from 45 screens.

For numerical scores, a graphical display of the score distribution and the position of the score in question is provided.

Phenotype information for gene 34411 (cdc2)

Screen Title	Gene ID [?]	Gene Symbol	Reagent ID [?]	Score	Phenotype [?]	Follow Up [?]
Cell cycle regulation (1)	CG5363	cdc2	cdc2	sp	Increased G2/M DNA content, increased cell size, decreased cytokinetic index, increased number of cells in prometaphase and metaphase, chromosome defects	yes
Cell cycle regulation (2)	CG5363	cdc2	cdc2	np	Increased G2/M DNA content, increased cell size	no
Cell growth and viability (1)			HFA03504	0,2  -3.7 7.7	none	yes
Cell growth and viability (2)			HFA03504	1,9  -2.9 7.3	none	no
Cell morphology	FBgn0004106	cdc2	np	np	Morphology	no
Cell size and cell-cycle regulation (1)	FBgn0004106	cdc2	np	sp	Decreased G1 DNA content, increased G2 DNA content	yes
Cell size and cell-cycle regulation (1)	FBgn0004106	cdc2	LD38718	sp	Decreased G1 DNA content, decreased viability	yes

cdc2 knock down was found to cause e.g. decreased G1 DNA content and decreased viability (Björklund et al. 2006). For literature curated phenotypes, the 'Screen Title' link provides information on the publication and the screen such as publication title, abstract, PubMed link as well as more details about the experiment. For example, this specific screen included kinases, phosphatases and selected genes, and monitored cell size, DNA content and viability in the *Drosophila* S2 cell line using dsRNA from custom-made libraries, named DGC1, DGC2 and PHOSPHO. Furthermore, links are provided to view or download the screen data.

Cell size and cell-cycle regulation (1)	FBgn0004106	cdc2	LD38718	sp	Decreased G1 DNA content, decreased viability	yes

REFERENCE

Identification of pathways regulating cell size and cell-cycle progression by RNAi. *Björklund et al., 2006*

View Phenotypes
Download
PubMed

Many high-throughput loss-of-function analyses of the eukaryotic cell cycle have relied on the unicellular yeast species Saccharomyces cerevisiae and Schizosaccharomyces pombe. In multicellular organisms, however, additional control mechanisms regulate the cell cycle to specify the size of the organism and its constituent organs. To identify such genes, here we analysed the effect of the loss of function of 70% of Drosophila genes (including 90% of genes conserved in human) on cell-cycle progression of S2 cells using flow cytometry. To address redundancy, we also targeted genes involved in protein phosphorylation simultaneously with their homologues. We identify genes that control cell size, cytokinesis, cell death and/or apoptosis, and the G1 and G2/M phases of the cell cycle. Classification of the genes into pathways by unsupervised hierarchical clustering on the basis of these phenotypes shows that, in addition to classical regulatory mechanisms such as Myc/Max, Cyclin/Cdk and E2F, cell-cycle progression in S2 cells is controlled by vesicular and nuclear transport proteins, COP9 signalosome activity and four extracellular-signal-regulated pathways (Wnt, p38betaMAPK, FRAP/TOR and JAK/STAT). In addition, by simultaneously analysing several phenotypes, we identify a translational regulator, eIF-3p66, that specifically affects the Cyclin/Cdk pathway activity.

SCREEN DETAILS

Stable Id: GR00048-A-1
Screen title: Cell size and cell-cycle regulation (1)
Assay [?]**:** Cell size, DNA content and viability
Method [?]**:** Flow cytometry
Scope: Kinases, phosphatases and selected genes
Screen type: Cell-based
Species: Drosophila melanogaster
Biosource [?]**:** Cell line
Biomodel: S2
Library: Custom-made, DGC1, DGC2 and PHOSPHO
Reagent type: dsRNA
Score Type [?]**:** Complex, sp
Cutoff [?]**:** Complex criteria
Notes: Additional information about the primary sccreen (pooled library) and a secondary screen (number of binucleate cells)

For phenotype entries reported in the screen "HeLa cell morphology" by Fuchs et al. there are images available for display. See for example:

Metaphase cells (81)

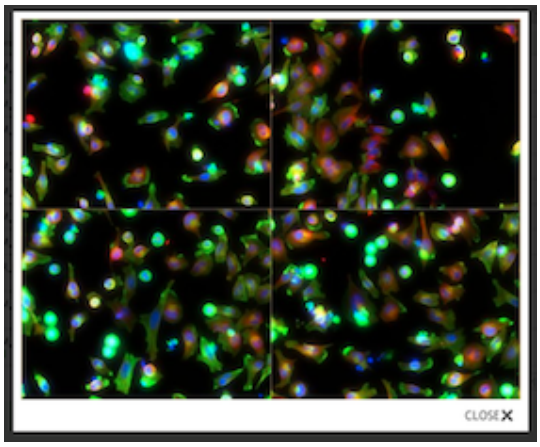
Mapped ID [?]	Gene ID [?]	Gene Symbol	Reagent ID [?]	Score Type [?]	Cutoff [?]	Score	Phenotype [?]	Follow Up [?]
1025	NM_001261	CDK9	M-003243-02	Complex, sp	np	np	Metaphase cells Click for image	no

Showing 10 of 1 entries (filtered from 81 total entries)

Text Search CDK9

"Click for image" opens the image in a new window. The colours are as follows:

Red: Actin, Green: Tubulin, Blue: DNA



Please refer to our Glossary for further information on the content of the columns.

Reagents tab

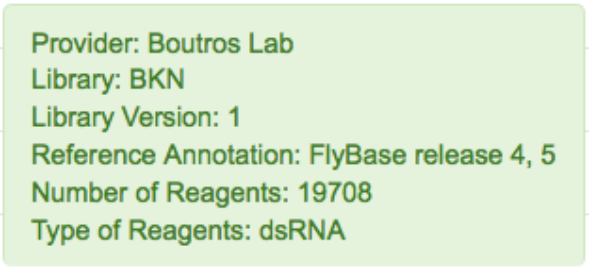
The 'Reagents' tab lists all RNAi reagents available targeting the queried gene (see below). For *cdc2* 13 reagents from 8 different RNAi libraries are available.

Reagent information for gene 34411 (cdc2)

Reagent ID [?]	Type	Library
106130	UAS-IR construct	KK [details]
41838	UAS-IR construct	GD [details]
41839	UAS-IR construct	GD [details]
62099	dsRNA	OpenBiosystems [details]
AMB33342	dsRNA	Ambion [details]
BKN22805	dsRNA	BKN [details]
BKN50710	dsRNA	BKN [details]
BKN51124	dsRNA	BKN [details]
DRSC03504	dsRNA	DRSC [details]
DRSC30705	dsRNA	DRSC [details]
DRSC30706	dsRNA	DRSC [details]
HFA03504	dsRNA	Heidelberg Fly Array (HFA) [details]
MRC022_F07	dsRNA	MRC [details]

The '[details]' link in the 'Library' column provides information about the RNAi library. For example, the BKN library (designed and synthesized in

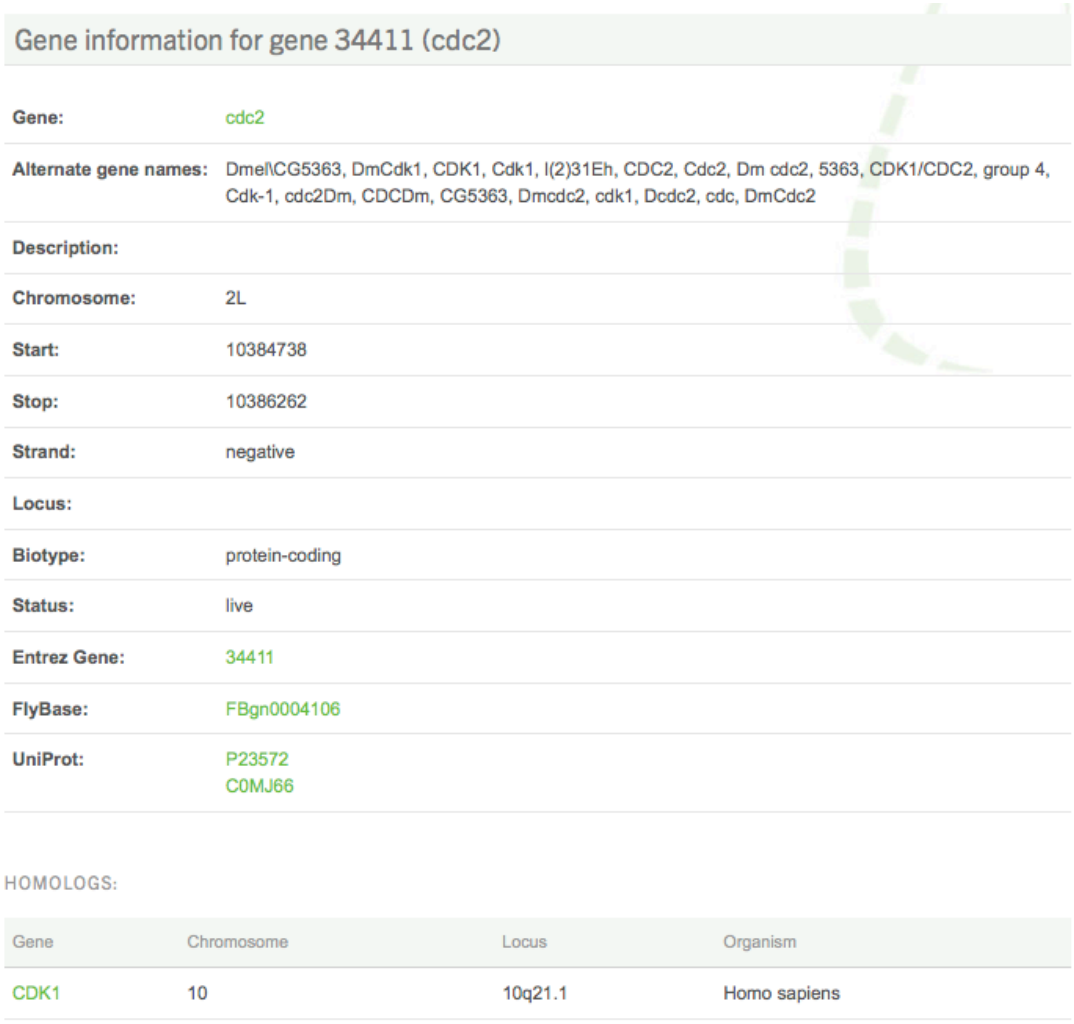
the Boutros lab) was designed for FlyBase releases 4 and 5 and contains 19.708 reagents (dsRNAs).



By clicking on one of the reagents detailed information about the reagent is displayed (see Search for reagents).

Gene info tab

The 'Gene info' tab provides information about alternative gene names, the gene location and linkouts to other databases such as Entrez, Ensembl, FlyBase, NCBI RefSeq, UniProt, etc. In case homologs (to *Drosophila* or *Homo sapiens*) are available, they are listed below the gene description. In the example below the human gene CDK1 was found as homolog of the *Drosophila* gene cdc2. The link in the 'Gene' column under 'Homologs' forwards to the gene details page ('Phenotypes' tab) of the listed homologous gene. Homology mappings were obtained from NCBI HomoloGene.



Gene information for gene 34411 (cdc2)

Gene:	cdc2
Alternate gene names:	Dmel\CG5363, DmCdk1, CDK1, Cdk1, l(2)31Eh, CDC2, Cdc2, Dm cdc2, 5363, CDK1/CDC2, group 4, Cdk-1, cdc2Dm, CDCDm, CG5363, Dmcdc2, cdk1, Dcdc2, cdc, DmCdc2
Description:	
Chromosome:	2L
Start:	10384738
Stop:	10386262
Strand:	negative
Locus:	
Biotype:	protein-coding
Status:	live
Entrez Gene:	34411
FlyBase:	FBgn0004106
UniProt:	P23572 C0MJ66

HOMOLOGS:

Gene	Chromosome	Locus	Organism
CDK1	10	10q21.1	Homo sapiens

Cross-species search for phenotypes

GenomeRNAi allows for searches across species for homologous genes. E.g. for *Drosophila cdc2* a human homolog is annotated (CDK1) and can be accessed via the 'Gene info' tab (see above). The output structure for human queries is analogous to outputs for *Drosophila*.

Genome browser tab

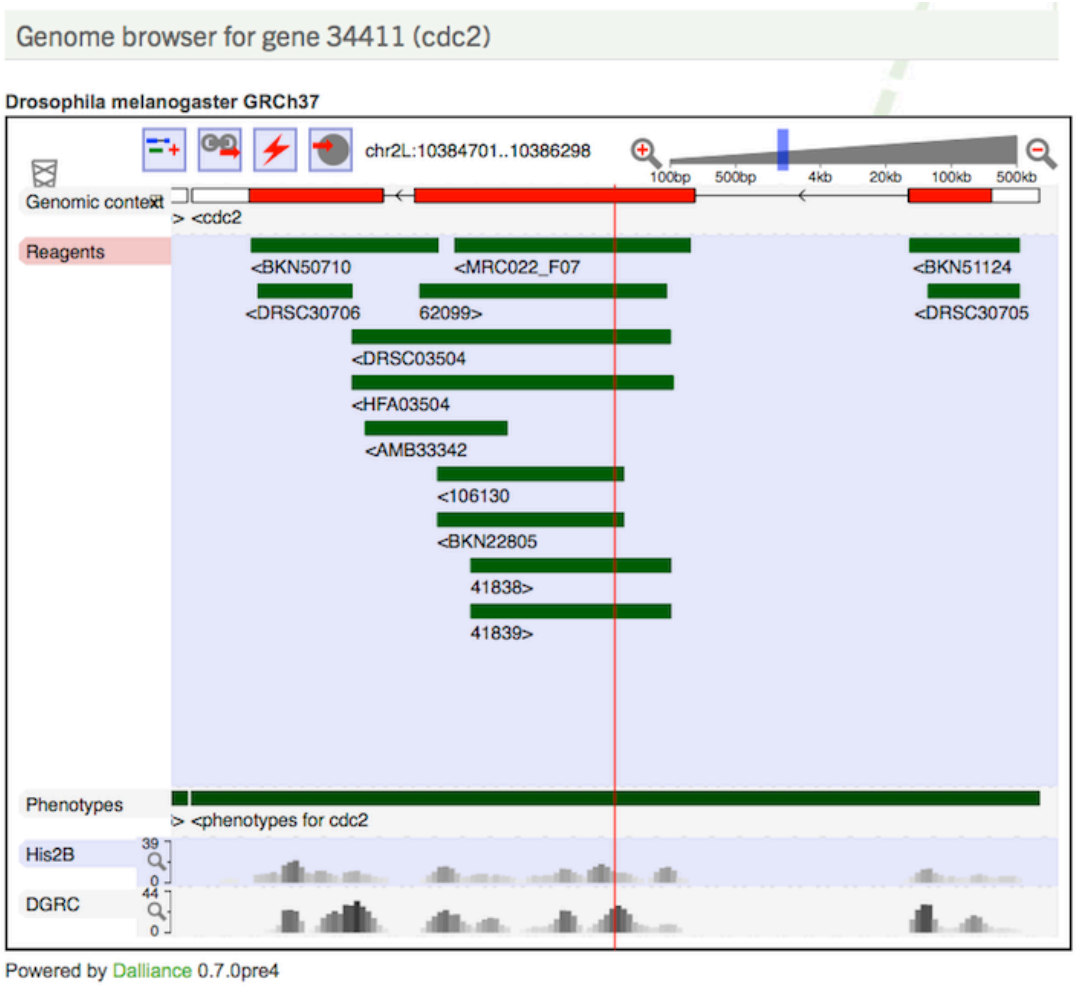
This tab displays the output of a dynamic genome browser implementation (Dalliance) to visualize different kinds of data in the genomic context of the targeted gene (see below for the output for *Drosophila* cdc2). The user can modify the display by zooming or scrolling, or by adding additional tracks for data sources available from the DAS registry. Further instructions are available from the Dalliance website.

For *Drosophila* the following 'tracks' are available by default:

- 'Genomic context': overall span and exon-intron structure of the gene
- 'Reagents': RNAi reagents available in GenomeRNAi – clicking on these tracks opens a window with further information and a link to the reagent page in GenomeRNAi
- 'Phenotypes': track spanning the entire length of the gene – clicking on this track opens a window with further information on the phenotypes and a link to the gene details page in GenomeRNAi
- 'RNASeq S2 cells (DGRC)' and 'RNASeq S2 cells (His2B)': Expression profiles for S2 cells (from DGRC) and S2-His2B-GFP cells (cells

kindly provided by the O'Farrell lab). The following table gives detailed information about the *Drosophila* RNASeq tracks:

Track Name	Platform	Mapping Algorithm	Uniquely Mapped Sequences
RNASeq S2 cells (DGRC)	Illumina	Tophat (default parameters)	5915932
RNASeq S2 cells (His2B)	Illumina	Tophat (default parameters)	5980329



The GBrowse implementation for *Homo sapiens* contains the following tracks:

- 'Genomic context': overall span and exon-intron structure of the gene
- 'Reagents': RNAi reagents available in GenomeRNAi – clicking on these tracks opens a window with further information and a link to the reagent page in GenomeRNAi
- 'Phenotypes': track spanning the entire length of the gene – clicking on this track opens a window with further information on the phenotypes and a link to the gene details page in GenomeRNAi
- 'RNASeq HEK293T cells', 'RNASeq DLD cells' and 'RNASeq HTC116 cells' : Expression profiles for the specified cell line. The following table gives detailed information about the human RNASeq tracks:

Track Name	Platform	Mapping Algorithm	Uniquely Mapped Sequences
RNASeq DLD cells	Illumina	Tophat (default parameters)	14840816
RNASeq HCT116 cells	SOLiD	Tophat (default parameters)	6494831
RNASeq HEK293 cells	Illumina	Tophat (default parameters)	3756389

Search for reagents

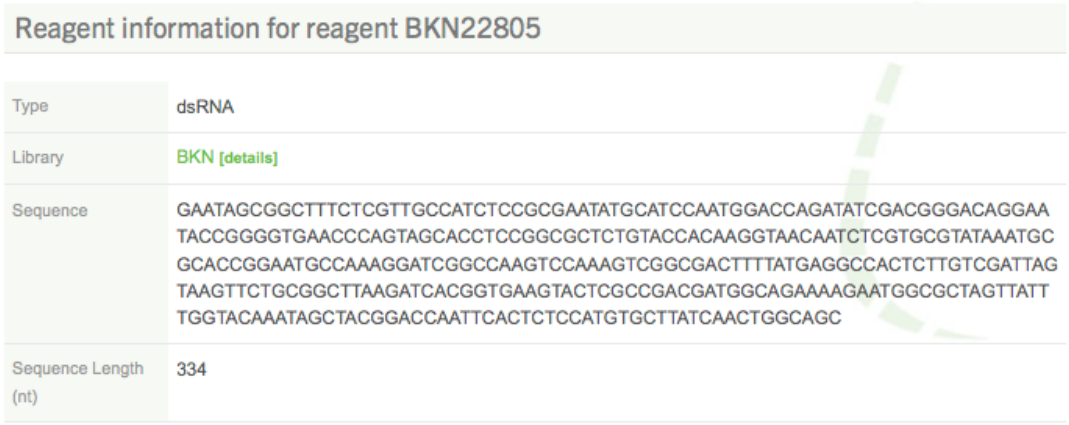
Search for reagents

The reagent identifier search on the GenomeRNAi home page will only identify exact matches in the database (no wild-card use). (Another option to show the reagent information is to search for a specific gene and choose a reagent in the 'Reagents' tab.)

In the example shown below, GenomeRNAi was queried for the reagent BKN22805:

The resulting page is subdivided in four tabs: 'Reagent info', 'Gene(s)', 'Phenotypes' and 'Genome browser'.

Detailed information about the reagent is displayed in the 'Reagent info' tab. Besides the full sequence and sequence length the primers and primer parameters (such as length, melting temperature and percent GC content) are listed as well as predicted sequence specificity ('Reagent quality' and 'Intended targets') and predicted efficiency. Information marked with an asterix is generated by NEXT-RNAi.



PRIMER FORWARD

Sequence	GAATAGCGGCTTTCTCGTTG
Length	20
Position	1
Tm (°C) *	59.982
GC (%) *	50.000

PRIMER REVERSE

Sequence	GCTGCCAGTTGATAAGCACA
Length	20
Position	334
Tm (°C) *	60.019
GC (%) *	50.000

REAGENT QUALITY *

siRNAs (20nt)	On-target	Off-target	No-target	Efficiency (%)
316	316	0	0	51.52

INTENDED TARGETS *

Gene	Transcripts (Hits)
FBgn0004106	FBtr0080051

'siRNAs [19nt]' is the number of all possible 19 nt siRNAs obtained from the dsRNA sequence (with an off-set of 1).

'On-target' is the number of siRNAs mapping only to the transcripts of the intended target gene.

'Off-target' is the number of siRNAs with mappings to transcripts of multiple target genes.

'No target' is the number of siRNAs with no transcript-target at all (e.g. for intron-spanning designs). BKN22805 has no predicted off-target to unintended transcripts.

The predicted efficiency ('Efficiency[%]') is calculated for each 19 nt siRNA according to the rules defined in Shah et al. and averaged over all siRNAs. It is on a scale from 0% to 100% and is basically a weighted scoring of overall 12 parameters (such as GC content, asymmetry between 3' and 5' end certain base preferences).

The 'Intended targets' table lists the primary target gene, all annotated splice variants and the number of siRNAs targeting each transcript. BKN22805 targets the transcript FBtr0080051 with all 316 siRNAs.

In case transcripts of other genes are targeted an additional table ('Off-targets') is displayed. An example is shown below for BKN21751 targeting *Drosophila* wls (CG6210, FBgn0036141, Entrez Gene ID 39259). This dsRNA has one perfect 19 nt match to the two splice forms of CG13366 (FBgn0025633, Entrez Gene ID 31004).

▾ OFF-TARGETS *

Gene	Transcripts (Hits)
FBgn0025633	FBtr0070161 [...]

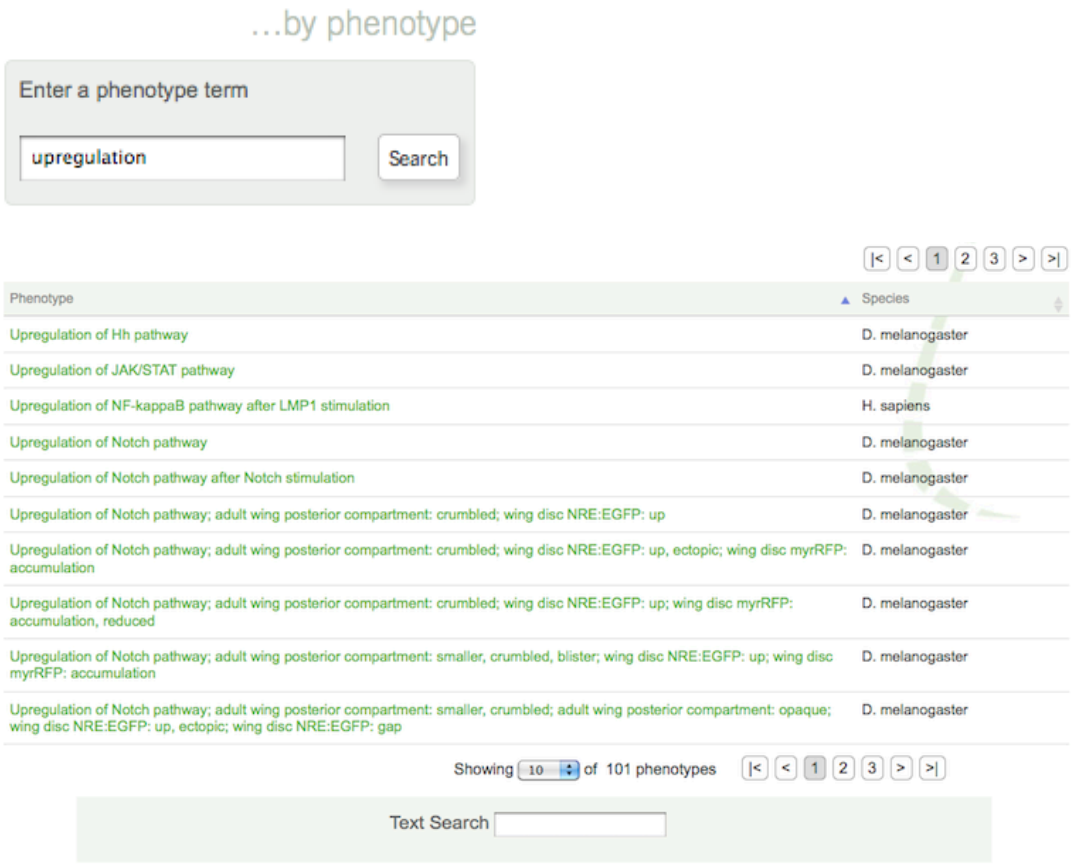
The other outputs ('Gene(s)', 'Phenotypes' and 'Genome browser' tabs) are analogous to the outputs described for gene searches.

Search for phenotypes

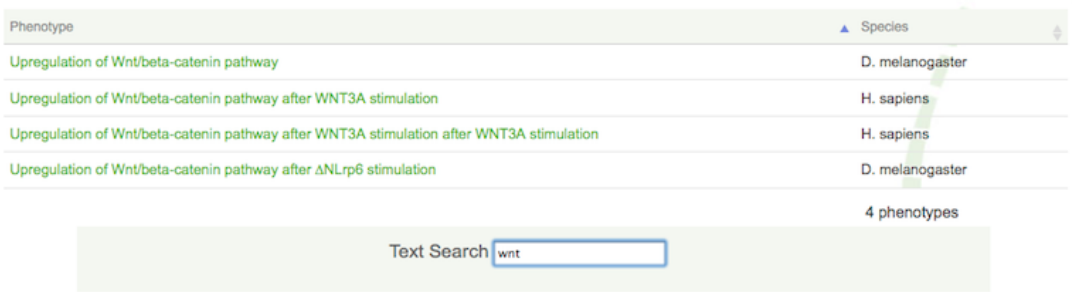
Search for phenotypes

The phenotype search field allows the user to query for specific phenotypes in an organism-independent manner. Incomplete queries will also be matched to the database entries, e.g. a search for 'regulation' can be used to search for both 'upregulation' and 'downregulation'.

A search for 'upregulation' phenotypes identifies all phenotypes related to upregulation such as 'Upregulation of Hh pathway' and 'Upregulation of Notch pathway after Notch stimulation' (see example below). For each phenotype description the species in which the screen was performed is displayed.

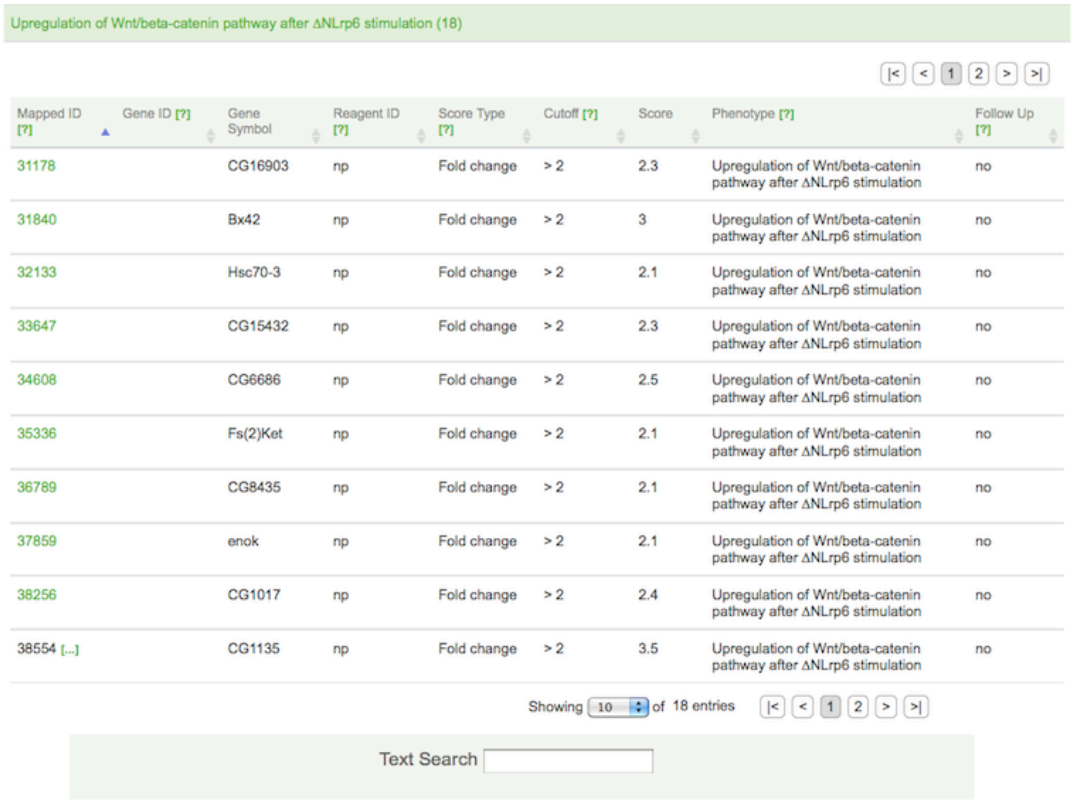


The 'Text Search' box allows searches within the results, e.g. a search for 'wnt' in the 'Text Search' box after searching for 'Upregulation' phenotypes will display only entries with the search term (see below).



By following the 'Phenotype' link, a list of the screens in which the particular phenotypic classification was observed is shown as well as the other phenotypic descriptions in those screens.

Click on the 'Title' link expands the detailed information about the screen and the publication as described above. Click on the phenotype heading expands the detailed information about the phenotypes; the 'Mapped ID' and 'Reagent ID' (if available) links query the database for the corresponding gene and reagent.



The 'Text Search' box allows filtering within the phenotype entries, e.g. a search for 'CG1' will display only entries with the search term (see below).

Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation (18)

Mapped ID [?]	Gene ID [?]	Gene Symbol	Reagent ID [?]	Score Type [?]	Cutoff [?]	Score	Phenotype [?]	Follow Up [?]
31178		CG16903	np	Fold change	> 2	2.3	Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation	no
33647		CG15432	np	Fold change	> 2	2.3	Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation	no
38256		CG1017	np	Fold change	> 2	2.4	Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation	no
38554 [...]		CG1135	np	Fold change	> 2	3.5	Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation	no
43019		CG11848	np	Fold change	> 2	2.3	Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation	yes
43529		CG18041	np	Fold change	> 2	3.2	Upregulation of Wnt/beta-catenin pathway after ΔNLrp6 stimulation	no

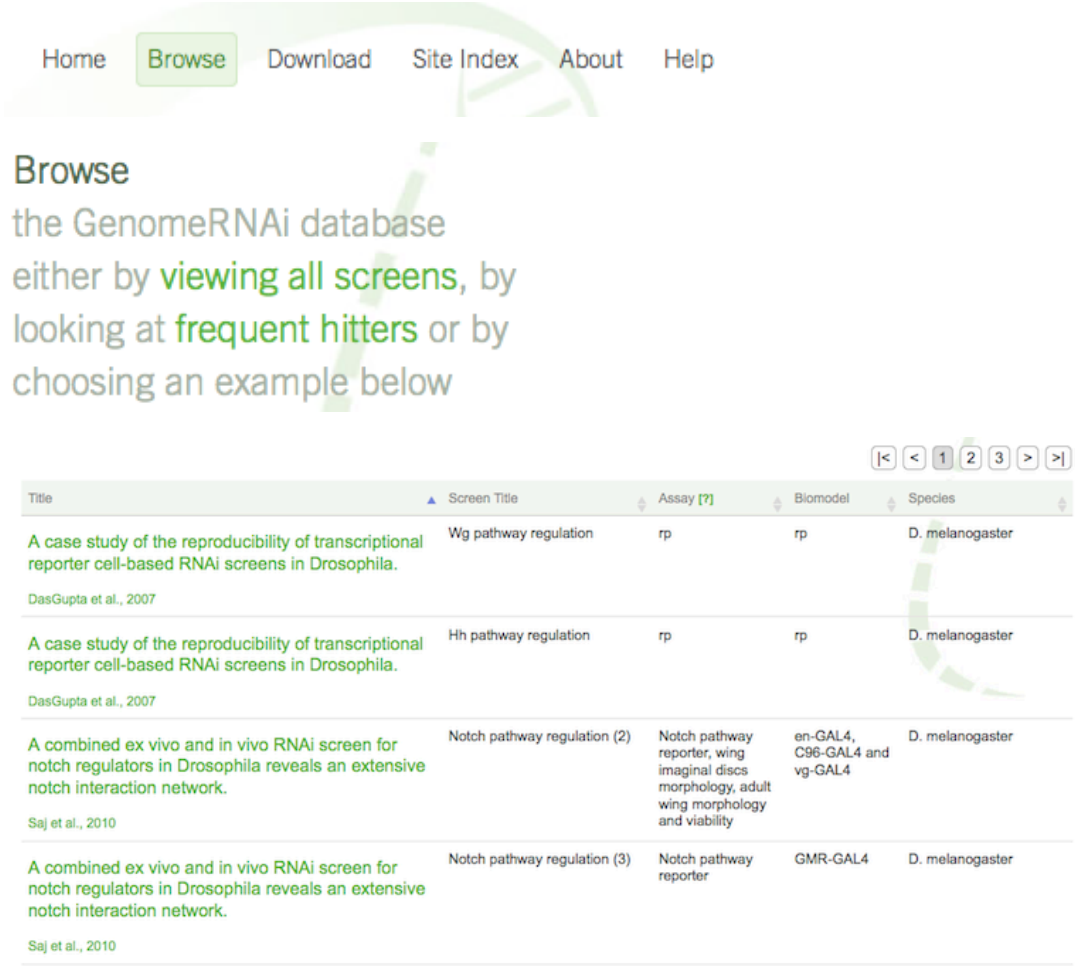
Showing 10 of 6 entries (filtered from 18 total entries)

Text Search CG1

Browse all screens

Browse all screens

A list of all available screens / publications in the GenomeRNAi database can be obtained by the 'Browse' tab or the 'viewing all screens' link on the GenomeRNAi home page (see below).



The 'Title' link provides information on the publication and the screen such as abstract, PubMed link as well as more details about the experiment. Furthermore, links are provided to view or download the screen data.

A combined ex vivo and in vivo RNAi screen for notch regulators in Drosophila reveals an extensive notch interaction network. Saj et al., 2010	Notch pathway regulation (3)	Notch pathway reporter	GMR-GAL4	D. melanogaster

Abstract

Notch signaling plays a fundamental role in cellular differentiation and has been linked to human diseases, including cancer. We report the use of comprehensive RNAi analyses to dissect Notch regulation and its connections to cellular pathways. A cell-based RNAi screen identified 900 candidate Notch regulators on a genome-wide scale. The subsequent use of a library of transgenic Drosophila expressing RNAi constructs enabled large-scale in vivo validation and confirmed 333 of 501 tested genes as Notch regulators. Mapping the phenotypic attributes of our data on an interaction network identified another 68 relevant genes and revealed several modules of unexpected Notch regulatory activity. In particular, we note an intriguing relationship to pyruvate metabolism, which may be relevant to cancer. Our study reveals a hitherto unappreciated diversity of tissue-specific modulators impinging on Notch and opens new avenues for studying Notch regulation and function in development and disease.

View Phenotypes
Download
PubMed

Screen details

Stable ID: GR00189-A-3
Screen title: Notch pathway regulation (3)
Assay [?]**:** Notch pathway reporter
Method [?]**:** Fluorescence
Scope: Selected genes
Screen type: in vivo
Species: D. melanogaster
Biosource [?]**:** Tissue
Biomodel: GMR-GAL4
Library: np, Notch pathway component enriched RNAi library
Reagent type: UAS-IR construct
Score Type [?]**:** Average fluorescent intensity
Cutoff [?]**:** Complex criteria
Notes:

The 'Text Search' box at the bottom of the page allows searches within all screen entries, e.g. a search for 'Hh' will display only entries with the search term (see below).

Title	Screen Title	Assay [?]	Biomodel	Species
A case study of the reproducibility of transcriptional reporter cell-based RNAi screens in Drosophila. DasGupta et al., 2007	Hh pathway regulation	rp	rp	D. melanogaster
A genome-wide RNA interference screen in Drosophila melanogaster cells for new components of the Hh signaling pathway. Nybakken et al., 2005	Hedgehog signaling	Hedgehog signaling	Clone 8	D. melanogaster
A genome-wide RNA interference screen uncovers two p24 proteins as regulators of Wingless secretion. Port et al., 2011	Wg pathway regulation (4)	Viability and wing morphology	hhGAL4	D. melanogaster

Showing 10 of 3 screens

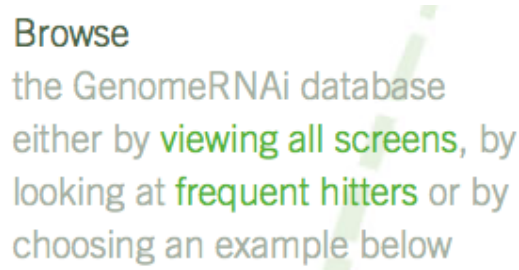
Text Search Hh

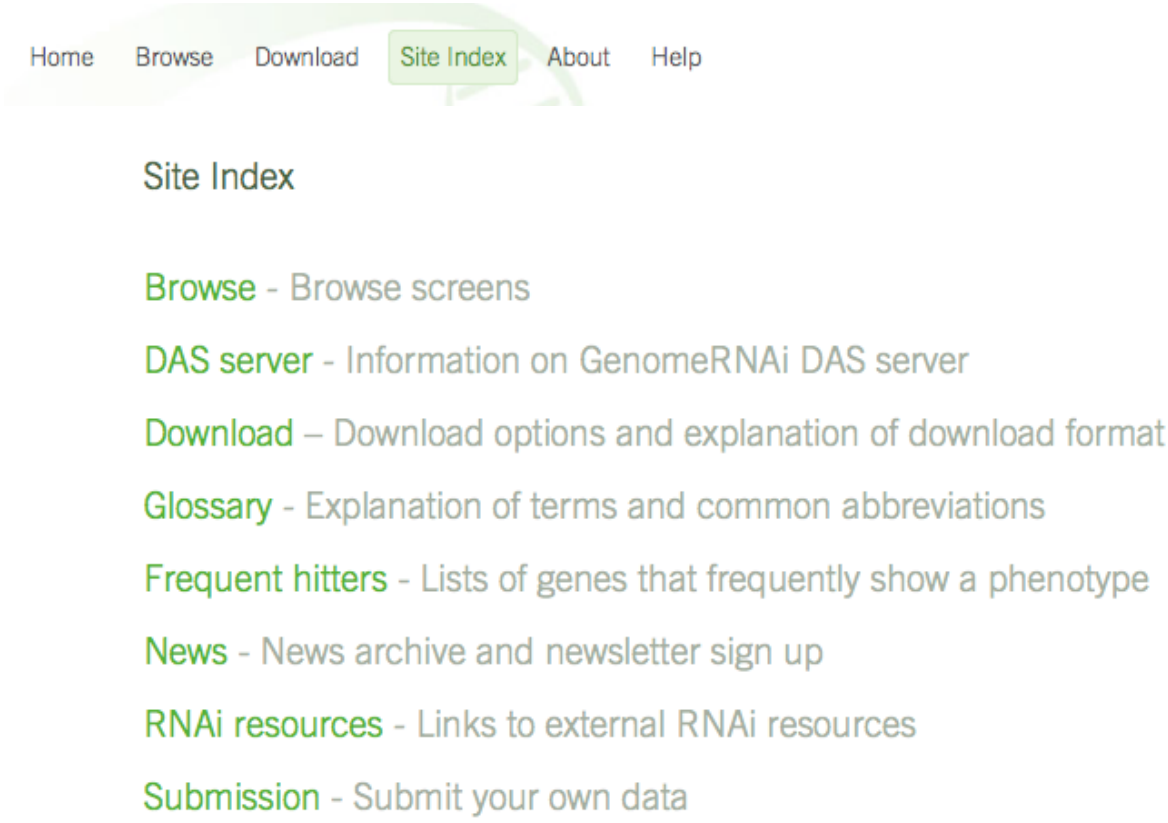
Frequent hitters

Frequent hitters

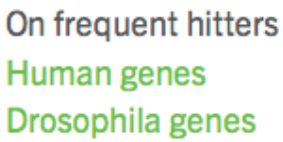
The 'Frequent hitters' page provides lists of genes that frequently show a phenotype according to the data included in GenomeRNAi.

In order to access the page, follow the link in the right column or go to the 'Site Index' tab and then choose the 'Frequent hitters' link:





The 'Frequent hitters' page is subdivided in three tabs: 'On frequent hitters', 'Human genes' and 'Drosophila genes':



The **'On frequent hitters' tab** provides information about frequent hitters, e.g. how the calculation of the number of phenotypes per gene was done. It also provides a link for downloading the frequent hitters list.

The **'Human genes' and 'Drosophila genes' tabs** provide gene lists sorted in descending order as to the number of times they have shown a phenotype. The lists can be sorted by other columns as well by clicking on the sorting arrows in the title bar. Clicking on the number of reagents link opens a window with a list of the reagents used. The output for human genes is shown below as an example:

Number of Hits	Total Entries	Percentage Hits	Reagents	Entrez Gene ID	Gene	Author-Provided Name(s) and ID(s)
20	31	64.5	11	5347	PLK1	ENSG00000166851, NM_005030, PLK, PLK1
18	27	66.7	14	10594	PRPF8	ENSG00000174231, NM_006445, PRPF8
17	26	65.4	10	10291	SF3A1	ENSG00000099995, NM_001005409, NM_005877, SF3A1
16	22	72.7	10	6651	SON	ENSG00000159140, NM_003103, NM_032195, SON
15	22	68.2	13	9276	COPB2	COPB2, ENSG00000184432, NM_004766
15	22	68.2	9			ENSG00000181789, 128
15	22	68.2	10			IF3S10, ENSG00000107581, NM_003750
15	21	71.4	12	1314	COPA	COPA, ENSG00000122218, NM_004371
14	27	51.9	7	4609	MYC	731404, C-MYC, ENSG00000136997, MYC, NM_002467
14	24	58.3	7	22938	SNW1	ENSG00000100603, NM_012245, SKIIP, SNW1

D-019138-01, D-019138-03, D-019138-04, ENSG00000181789, Hs_COPG_5, Hs_COPG_7, M-019138-00, PL-50054, np

Showing 10 of 4,940 genes |< < 1 2 3 > >|

Text Search

The 'Text Search' box allows searches within the frequent hitters, e.g. a search for 'COP1' in the 'Text Search' box will display only entries with the search term (see below).

Number of Hits	Total Entries	Percentage Hits	Reagents	Entrez Gene ID	Gene	Author-Provided Name(s) and ID(s)
3	13	23.1	6	64326	RFWD2	COP1, ENSG00000143207, NM_001001740, NM_022457, RFWD2
2	8	25	2	114769	CARD16	CARD16, COP, COP1, NM_052889

Showing 10 of 2 genes (filtered from 4,940 total entries)

Text Search COP1

Download information

Download data from an individual screen

Data for individual screens can be downloaded directly from the website. For example, go to 'Browse' and click on your screen of interest. Then click on the 'Download' button to the right.

Download format

The download files consist of two sections per screen: The first section provides general information about the screen, such as authors, Pubmed ID, biomodel, assay, and the second section provides the actual gene-phenotype associations in tab delimited format. An example is given below.

NOTE: Entrez ID vs Gene ID - the Gene ID column contains the gene identifier as provided by the authors of the screen and can therefore come from various sources when comparing different screens. We have attempted to map all Gene IDs to Entrez Gene IDs in order to provide one data column with uniform identifiers.

NOTE: Stable IDs are assigned according to the following convention:
GR00123-A = core ID for a screen curated from the literature
GR00123-C = core ID for a screen imported via cellHTS2
GR00123-A-0 = indicates that the data from this screen still need to be updated to the current annotation standard (pre-existing data)
GR00123-A-1, GR00123-A-2, etc. = indicates that this is one screen out of several, published within the same publication

For more information about data fields see also our Glossary.

Download Example

#Screen ID=GR00184-A-1
#Screen Title=Self-renewal and pluripotency in human embryonic stem cells (1)
#Publication Title=A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity
#Authors=Chia et al.
#Publication Year=2010
#Pubmed ID=20953172
#Organism=Homo sapiens
#Screen Type=Cell-based
#Biosource=Cell line
#Biomodel=hESC H1
#Assay=POU5F1-GFP protein expression
#Method=Fluorescence
#Library Manufacturer=Dharmacon
#Library=SMARTpool siRNA library
#Scope=Genome-wide
#Reagent Type=siRNA
#Score Type=Z-score
#Cutoff=< -2
#Notes=
#Screen ID Entrez ID Gene ID Gene Symbol Reagent ID Score Phenotype Conditions Follow Up Comment
GR00184-A-1 390992 XM_372757 HES3 M-027394-00 -0.011 none no
GR00184-A-1 3704 NM_033453 ITPA M-013681-00 -1.153 none no
GR00184-A-1 10673 NM_006573 TNFSF13B M-017586-00 -0.013 none no
GR00184-A-1 284837 NM_194310 LOC284837 M-019296-00 0.562 none no
GR00184-A-1 56999 NM_020249 ADAMTS9 M-005779-00 -0.67 none no
GR00184-A-1 1212 NM_001834 CLTB M-004003-00 -2.263 decreased POU5F1-GFP protein expression no

Download all screens

A file containing data for all screens in GenomeRNAi is available via e-mail. Please use this form to contact us and we'll send you a download link.

E-mail address:

Name:

Institution:

Submit

Download frequent hitter list

Frequent hitters are gene lists sorted in descending order according to the number of times a gene has shown a phenotype. Genes showing a phenotype only once, or not at all, are not included.

Download frequent hitter list.

Download archived screens

In the process of updating pre-existing data and applying the new annotation guidelines, some screens have been removed from the GenomeRNAi database.

Typically the reason is a very low number of specific data that can be extracted from the publication, e.g. only one hit described from a genome-wide screen without providing information on the genes actually tested.

Screen data that have been removed from the database have been archived and can be downloaded here.

Submission of RNAi phenotype and/or reagent data

Have you published or are you about to publish RNAi screening data? Would you like to increase the visibility of your data in the RNAi community and beyond? We welcome direct data submissions from authors into GenomeRNAi, the most up-to-date database for RNAi phenotypes and reagents in human and drosophila.

Why submit data to GenomeRNAi?

GenomeRNAi provides high-quality RNAi phenotype data in a structured format, allowing searches and comparisons across multiple screens.

GenomeRNAi is well integrated with other publically available databases, providing links with Entrez Gene identifiers, Ensembl, UniProt and Flybase.

Direct author submissions are much preferred over curation by a curator team - the authors know the data best! Indeed, we encourage authors to get in touch with us even before publication - we are happy to provide advice in respect to good practice of reporting RNAi phenotype data.

How does it work?

We provide an Excel data template for download. This file contains detailed instructions with regard to our annotation guidelines, with dropdown menus where appropriate, to facilitate straightforward and consistent annotation. The author sends the completed file to us.

Our curator team performs a quality check to make sure the data are complete and consistent with our annotation guidelines.

The curator team gets in touch with the author to confirm the status of the submission and sort out any open issues.

At the next GenomeRNAi release, the author receives a confirmation with a direct link to the submitted data.

Which species can be included in GenomeRNAi?

Currently, GenomeRNAi holds RNAi phenotype data for *Homo sapiens* and *Drosophila melanogaster*. Our curator team focusses their efforts into manual curation from the literature on these two species.

However, we are happy to discuss inclusion of other species into GenomeRNAi, if there is interest to submit such data. Do let us know and we can make the necessary arrangements.

Any other questions?

Don't hesitate to contact us with any further questions!

Download data template

An Excel template for submission of RNAi data into the GenomeRNAi database is available via e-mail. Please use this form to contact us and we'll send you a download link.

E-mail address:

Name:

Institution:

Submit

You can link to GenomeRNAi by using the following URLs:

Genes:
http://www.genomernai.de/GenomeRNAi/genedetails/<ENTREZ GENE ID>
http://www.genomernai.de/GenomeRNAi/ExternalLink/ensembl/<ENSEMBL ID>
http://www.genomernai.de/GenomeRNAi/ExternalLink/uniprot/<UNIPROT ID>
http://www.genomernai.de/GenomeRNAi/ExternalLink/cg/<CG ID>
http://www.genomernai.de/GenomeRNAi/ExternalLink/flybase/<FLYBASE ID>

Reagents:
http://www.genomernai.de/GenomeRNAi/reagentdetails/<REAGENT NAME>

Phenotypes:
http://www.genomernai.de/GenomeRNAi/Index.thesubmit3/<PHENOTYPE TERM>

Screens:
http://www.genomernai.de/GenomeRNAi/ExternalLink/stableid/<GenomeRNAi STABLE ID>

German Cancer Research Center (DKFZ) - Division of Signaling and Functional Genomics
Heidelberg University, Department of Cell and Molecular Biology

GenomeRNAi: a database for cell-based RNAi phenotypes. 2009 update.
Gilsdorf M, Horn T, Arziman Z, Pelz O, Kiner E, Boutros M.
Nucleic Acids Res. 2010 Jan;38(Database issue):D448-52. Epub 2009 Nov 12.

GenomeRNAi: a database for cell-based RNAi phenotypes.
Horn T, Arziman Z, Berger J, Boutros M.
Nucleic Acids Res. 2007 Jan;35(Database issue):D492-7. Epub 2006 Nov 28.

When you have used GenomeRNAi data and cited our publication, we'd very much appreciate if you let us know. Any feedback is also welcome!

German Cancer Research Center (DKFZ) - Division of Signaling and Functional Genomics
Heidelberg University, Department of Cell and Molecular Biology

Glossary

This glossary gives short definitions for the major terms used in this website.

Assay – Short description of the assay method from a biological point of view, e.g. ‘viability’, ‘muscle cell morphology’, ‘hTERT mRNA expression’.

Method – Short description of the assay method from a technical point of view (readout), e.g. ‘luminescence’, ‘real-time qPCR’, ‘flow cytometry’.

Biomodel – The concrete biomodel used, e.g. ‘HeLa’, ‘primary embryonic cells’.

Biosource – Type of biomodel used in RNAi screen; controlled vocabulary (‘Cell line’, ‘Primary cells’, ‘Tissue’, ‘Organism’). For *in vivo* screens: ‘tissue’, if tissue-specific expression of UAS-IR, ‘organism’, if ubiquitous expression of UAS-IR.

Conditions – Only used when the same assay was done under different conditions (e.g. with or without an additional agent, different incubation times).

Follow Up – Indicates whether any further experimental analysis has been performed with the gene (rather than the reagent). The abbreviation 'NA' can be found in the previously annotated data and indicates that the phenotype was either not validated by secondary test or this information was not available from the publication.

Gene ID – Gene identifiers as provided by the authors. Gene IDs separated by ‘, ’ indicate that two or more genes have been knocked down.

Mapped ID – Entrez IDs, obtained by multi-step mapping of author-provided gene identifiers, gene symbols or reagent identifiers to the Entrez database.

Phenotype – Description of the actual observation upon knockdown of a gene, rather than an interpretation of the result, e.g. ‘increased activity…’ rather than ‘suppressor…’. ‘none’ = non-hit, or not provided.

Reagent ID - Reagent IDs separated by ‘, ’ indicate that two or more reagents have been used for RNAi.

Scope – Extent of the screen, e.g. ‘genome-wide’, ‘phosphatases’, ‘selected genes’.

Score – The score actually used for defining hits, consistent with score type and score cutoff.

Cutoff – Score threshold or other criteria used by the authors to define hits and separate them from non-hits.

Score Type – Type of score used in the screen for defining hits, e.g. ‘Z-score’, ‘percentage’, ‘frequency’.

Stable ID - Stable Ids are assigned according to the following convention:
GR00123-A= core ID for a screen curated from the literature
GR00123-C = core ID for a screen imported via cellHTS2
GR00123-A-0 = indicates that the data from this screen still need to be updated to the current annotation standard (pre-existing data)
GR00123-A-1, GR00123-A-2, etc = indicates that this is one screen out of several, published within the same publication

Common Abbreviations

np - (None Provided) might appear in several data fields. The abbreviation means that no information for this field was available from the publication.

rp - (Reference Provided) might appear in several data fields. The abbreviation is used when the authors did not describe the information but provided a reference instead.

sp - (See Publication) might appear in several data fields. For example, it is used in the „Conditions“ field to describe complex conditions used for the assay, e.g. summarized phenotypes from several lines in *in vivo* screens. If the „Score Type“ it indicates that several score types were taken for selecting hits, or when the score type was not described precisely. In the „Score“ field „SP“ is used when two or more scores were needed for defining hits.

DAS server

DAS server

We provide RNAi phenotype and reagent data from GenomeRNAi as a DAS source (format according to the DAS/1.53 specification). Please use the following URL:
http://genomernai.de/DASGenomeRNAi/das/

If you do not have a genome browser installed, we suggest the use of the Ensembl Genome Browser, e.g. through the URL upload panel ('Manage Your Data' -> 'Attach DAS', please refer to the "Adding Custom Tracks to Ensembl" help page for details and other options).

We provide the following tracks:

Drosophila melanogaster tracks:

- "GenomeRNAi Phenotypes": RNAi knockdown phenotypes from the GenomeRNAi database for *Drosophila melanogaster*
- "GenomeRNAi Reagents": RNAi reagents from the GenomeRNAi database for *Drosophila melanogaster*

Homo sapiens tracks:

- "GenomeRNAi Phenotypes": RNAi knockdown phenotypes from the GenomeRNAi database for *Homo sapiens*
- "GenomeRNAi Reagents": RNAi reagents from the GenomeRNAi database for *Homo sapiens*

See for example the output of the GenomeRNAi data for the *Drosophila melanogaster* cdc2 gene in the Ensembl Genome Browser: