

# GenomeRNAi: a database for cell-based RNAi phenotypes. 2009 update

Moritz Gilsdorf, Thomas Horn, Zeynep Arziman, Oliver Pelz, Evgeny Kiner and Michael Boutros\*

German Cancer Research Center and University of Heidelberg, Division Signaling and Functional Genomics, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

Received August 15, 2009; Revised October 15, 2009; Accepted October 16, 2009

## ABSTRACT

The GenomeRNAi database (<http://www.genomernai.org/>) contains phenotypes from published cell-based RNA interference (RNAi) screens in *Drosophila* and *Homo sapiens*. The database connects observed phenotypes with annotations of targeted genes and information about the RNAi reagent used for the perturbation experiment. The availability of phenotypes from *Drosophila* and human screens also allows for phenotype searches across species. Besides reporting quantitative data from genome-scale screens, the new release of GenomeRNAi also enables reporting of data from microscopy experiments and curated phenotypes from published screens. In addition, the database provides an updated resource of RNAi reagents and their predicted quality that are available for the *Drosophila* and the human genome. The new version also facilitates the integration with other genomic data sets and contains expression profiling (RNA-Seq) data for several cell lines commonly used in RNAi experiments.

## INTRODUCTION

RNA interference (RNAi) is a post-transcriptional gene silencing mechanism conserved from plants to humans and relies on the delivery of exogenous short double-stranded (ds) RNAs as a trigger for the degradation of homologous mRNA in cells (1,2). RNAi is now widely used as an experimental tool to silence the expression of genes in a broad spectrum of organisms (3). The availability of RNAi libraries targeting almost every transcript in an organism's genome has enabled researchers to query genomes for a broad spectrum of loss-of-function

phenotypes *in vitro* and *in vivo*. Such RNAi screens play an increasingly important role for the identification and characterization of gene function.

A crucial task is the integration and comparison of different RNAi screening datasets. For the 2007 version of GenomeRNAi (4) we collected and analyzed more than 91 000 long dsRNAs from different RNAi libraries targeting *Drosophila* transcripts and about 6100 published phenotype records from 29 large-scale studies in *Drosophila* cells.

Here we present an updated version of the GenomeRNAi database that has been significantly extended by adding new *Drosophila* RNAi screens and reagents curated from the literature. In addition, we have incorporated RNAi reagents and phenotypic screens performed in human cells and RNA-Seq data from selected cell lines. The database now contains more than 99 700 phenotypic classifications for *Drosophila* genes and about 97 600 for human genes including information about the RNAi libraries used in these studies, such as sequence information, predicted specificity (5,6) and predicted efficiency (7). Also, the availability of data from RNAi screens in human cells now allows for the evaluation of RNAi phenotypes across species. The new version of GenomeRNAi incorporates the possibility to present phenotypes from image-based screens. The database and user interface were completely re-implemented gaining significant improvements in performance and handling of queries. The GenomeRNAi database can be accessed at <http://www.genomernai.org/>.

## DATABASE CONTENT

The updated GenomeRNAi database integrates information about RNAi reagents, their annotated targets and phenotypic information based on large-scale RNAi screens in *Drosophila* and human tissue culture (Figure 1).

\*To whom correspondence should be addressed. Tel: +49 6221 421951; Fax: +49 6221 421959; Email: [m.boutros@dkfz.de](mailto:m.boutros@dkfz.de)

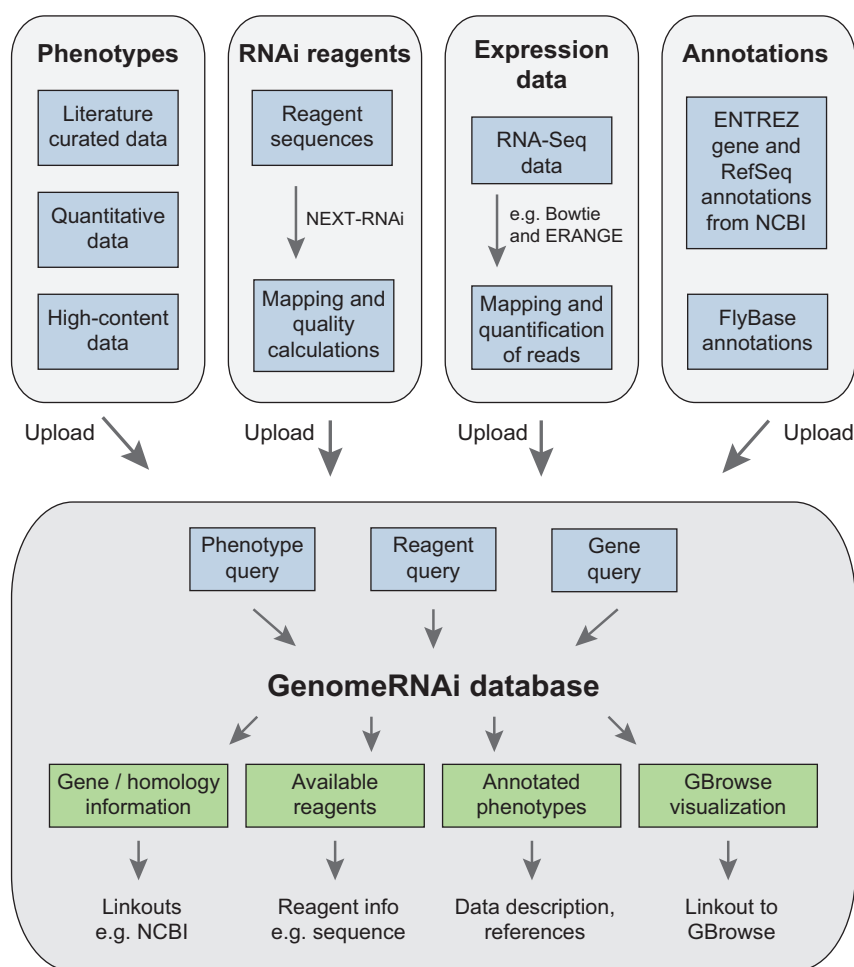
Present address:

Evgeny Kiner, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA02142, USA.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2009. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.5/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Overview of GenomeRNAi database content. Phenotypes (from literature, quantitative data and high-content data), RNAi reagents, expression data as well as gene annotations were collected in a database (upper panel). GenomeRNAi allows queries for reagents, genes and phenotypes and provides corresponding outputs (lower panel).

It contains 118 443 RNAi reagents from seven libraries available for *Drosophila* [DRSC library from Boston (8), HFA (9) and BKN libraries from Heidelberg, libraries from Ambion, OpenBiosystems (10) and the MRC, *in vivo* library from the VDRC (11)] and 302 786 RNAi reagents from four human siRNA- or shRNA-based libraries (Ambion Silencer Select, Dharmacon/ThermoFisher siGENOME, Sigma-Aldrich TRC and Qiagen druggable/whole genome supplement). Reagents were computationally mapped onto the latest genomic sequence using BLAT (12) and Bowtie (13). Annotations for targeted genes and transcripts were derived through the mapping on genome and transcriptome databases. In addition predicted specificities (5,6) and efficiencies (7) as well as potential regions of low complexity, such as simple nucleotide repeats and tandem repeats of the trinucleotide CAN (N indicated any base) (5,6), were calculated for each reagent. All calculations were performed using a new design/evaluation pipeline, NEXT-RNAi (T. Horn, manuscript in preparation). NEXT-RNAi also generates output files for the Generic Genome Browser (GBrowse) (14,15) that are used for the visualization of the mapped reagents in their genomic context (as GBrowse 'tracks').

The database contains RNAi phenotypes from tissue culture screens in *Drosophila* and human cells. Phenotypes were manually curated from published supplemental material. For some *Drosophila* screens the data was downloaded from FlyRNAi (16). For each entry we tried to assign the RNAi reagent used, the targeted gene, scoring methods and thresholds, the final score and the observed phenotype. In addition other data was extracted from publications including e.g. cell type, readout type, assay, assay length, reagent type and reagent amount. We uploaded phenotypic data from all to-date published screens and also included large-scale datasets generated in our lab. The new version of GenomeRNAi now contains data from 97 genome-scale screens performed in *Drosophila* (including nine *in vivo* screens) and 48 genome-scale screens in human cells. In total, more than 197 000 phenotypic classifications are currently stored in GenomeRNAi (99 700 *Drosophila*, 97 600 human). A list of all currently available screens can be accessed through the 'List all Screens' link on the GenomeRNAi webpage.

A new feature of GenomeRNAi is the presentation of data and phenotypes from image-based, high-content screens. To date, the database hosts images for a

genome-wide morphology screen performed in human HeLa cells (F. Fuchs, manuscript in preparation) and images of knock downs of all *Drosophila* kinases and phosphatases in *Drosophila* S2 cells (T. Horn, unpublished data). The *Drosophila* set will be expanded to the full genome in the near future.

Another new feature of the database is the cross-evaluation of RNAi phenotypes between *Drosophila* and *Homo sapiens*. Homology mappings were obtained from NCBI Homologene (17). This offers the opportunity to check whether a phenotype is conserved in *Drosophila* and human. As more comparable datasets become available (such as Wnt signaling pathway screens done in *Drosophila* and human cells) the value of interspecies comparisons is expected to increase further.

GBrowse offers a versatile tool to visualize gene models and mappings of RNAi reagents to the genome. We integrated RNA-Seq data (T. Sandmann, unpublished data) from *Drosophila* S2 cells and from human HEK293T cells as wiggle-plots in GBrowse. Absence of detectable gene expression may indicate that observed phenotypes resulted from cross silencing ('off-target' silencing) of other genes. The new version of GenomeRNAi also contains GBrowse 'tracks' for predicted specificities and efficiencies of the complete *Drosophila* and human genomes.

## DATA QUERY

The database can be queried by providing gene identifiers (NCBI, Entrez, Ensembl, FlyBase), RNAi reagent identifiers or phenotypes (Figure 2a). A list of all screens can be displayed via a direct link on the entry page that also allows accessing all phenotypes reported for a particular screen. Genes, RNAi reagents and phenotypes are linked to each other so that all kinds of queries allow accessing the other information available. Example-queries are also provided via links on the entry page. Advises how to query the database and further help can be obtained via the 'Help' link.

## DATA OUTPUT

For gene queries GenomeRNAi first provides detailed annotation information about the gene (from NCBI) including homology information (NCBI Homologene) in the 'Gene Info' tab (Figure 2b). To obtain more information about the gene, linkouts to other data source were implemented (including Entrez, NCBI RefSeq, HPRD, FlyBase). GenomeRNAi also summarizes all RNAi reagents available targeting the queried gene (tab 'Reagents', Figure 2c). The reagent links lead to more detailed information about the reagents (Figure 2d). The tab 'Phenotypes' (Figure 2e,f) lists three types of phenotypes: quantitative phenotypes, imaging phenotypes and literature-curated phenotypes. The 'GBrowse' tab (Supplementary Figure S1) shows the visualization of the gene model, the mapping of available RNAi reagents, plots for predicted specificities and efficiencies as well as available RNA-Seq data.

An example of a database session is shown in Figure 2. Here, we searched for RNAi reagents and annotated phenotypes available for the human gene COPB2, a factor required for vesicle trafficking (18). Figure 2b shows the results screen with detailed gene information, including a link to the *Drosophila* homologs (beta'Cop, CG6699). The 'Reagent' tab (Figure 2c) lists RNAi reagents from three different siRNA libraries available to target this gene. Following the Qiagen 'siRNA\_Pool' (SP00002881) link provides detailed information such as sequences, predicted specificity ('On-target' versus 'Off-target') and predicted efficiency as well as targeted transcripts ['Transcripts (Hits)'] can be obtained (Figure 2d). The tab 'Phenotypes' (Figure 2e) shows viability data from three different cell lines (HeLa, HEK293T and HepG2). The 'screen'-links provide more detailed information about the experiments. Knock down of COPB2 results in a viability defect in all three screens. The phenotype is quite severe in HeLa and HEK293T cells, where the 'Activity' (here viability) is below 10%. In HepG2 cells the phenotype is less strong with a remaining viability of about 70%. The images from the cell morphology screen in HeLa cells ('Imaging Phenotypes') support the phenotype as most cells died or show morphological defects. Additional support is provided by published experiments, e.g. COPB2 was found causing cell death in a 'Genome stability' assay. Going back to the 'Gene Info' tab and following the homology link to beta'Cop, similar phenotypes were found in *Drosophila* screens (Figure 2f). Both screens (in S2 cells from DGRC and S2-His2B-GFP cells) show a decreased viability after knock down of beta'Cop with a remaining viability of about 70% in both cell lines. In addition beta'Cop knock-down was found causing phenotypes in 12 other screens, such as a screens for viability and cell cycle as well as several 'infection' screens.

The same database output can be obtained by direct searches for reagents or phenotypes.

## CONCLUSION AND OUTLOOK

The GenomeRNAi database hosts RNAi phenotype information from large-scale studies in *Drosophila* and human cells connected with the underlying perturbation reagents and annotated target genes. Since genome annotations are in flux and RNAi reagents could exert unspecific effects, it is important to provide a regularly updated match of RNAi reagents to intended target genes. GenomeRNAi facilitates the evaluation of phenotypes at multiple levels. It provides the latest reagent annotations and internally calculates quality information, such as predicted specificity and efficiency. The information whether an observed phenotype was 'validated' (by retests with independent RNAi designs or other secondary assays) also contributes to the phenotype assessment. Furthermore the visualization of RNAi reagents in GBrowse uncovers design limitations such as incomplete coverage of annotated splice variants or designs biases (e.g. towards UTR regions) and provides the user with information about the expression of the

(a)

Please select organism: **Homo sapiens**

Search for Gene:

Options: ☐ Alternative gene names  
☐ External IDs (ENSEMBL, FLYBASE)

Search by RNAi identifier:

Search by Phenotype:

[List all Screens](#)

(b)

## Gene Description

Gene:	<a href="#">COPB2</a>
Alternate gene names:	beta'-COP
Description:	coatomer protein complex, subunit beta 2 (beta prime)
Chromosome:	3
Locus tag:	3q23
Biotype:	protein-coding
Status:	live
Entrez Gene ID:	<a href="#">9276</a>
Homologs:	1
ENSEMBL	<a href="#">ENSG00000184432</a>
HPRD	<a href="#">06103</a>
RefSeq ID:	<a href="#">NM_004766</a>

## Homologs

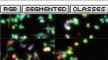
Gene	Chromosome	Locus	Organism
<a href="#">beta'Cop</a>	2L	3489-3489	<i>Drosophila melanogaster</i>

(c)

## Phenotypes from internal screens

id	Title	Score	Locus Id	Activity Ch1	Reag	Access	Image
13	HeLa cell viability	3.94	<a href="#">Dha_1_B11</a>	7.4 % (σ: 1.0%)	M-019847-01		
14	HEK293T cell viability	16.6	<a href="#">Dha_1_B11</a>	5.1 % (σ: 0.4%)	M-019847-01		
17	HepG2 cell viability	2.6	<a href="#">Dha_1_B11</a>	71.5 % (σ: 10.3%)	M-019847-01		

## Imaging Phenotypes

Experiment	Cell Type	Image
<a href="#">Cell morphology</a>	HeLa (cervical cancer)	

## Curated phenotypes from publications

Screen Id	Experiment	Reagent Id	Score Type	Score Cutoff	Score	Phenotype	Validated
71	<a href="#">Genome stability</a>	NM_004766	% gammaH2AX	Significance >= 1	18.37% (significance 4)	Increased number of gammaH2AX foci	NA
71	<a href="#">Genome stability</a>	NM_004766	ND	ND	ND	Cell death	NA
75	<a href="#">Wnt signaling</a>	M-019847-01	Z-score	<= -2	-2.55	Decreased Wnt reporter activity	Yes
119	<a href="#">Cell cycle</a>	COPB2	Effect on cell cycle	ND	ND	G0/1 arrest	Yes
128	<a href="#">Cell viability</a>	COPB2	Essentiality	>= 8/12 cancer cell lines	11	Decreased viability	Yes

(c)

Reagent Id	Type	Library
AS00005X	siRNA	Ambion
AS00005Y	siRNA	Ambion
AS00005Z	siRNA	Ambion
D-019847-02	siRNA	Dharmacon
D-019847-03	siRNA	Dharmacon
D-019847-04	siRNA	Dharmacon
D-019847-05	siRNA	Dharmacon
M-019847-01	siRNA_Pool	Dharmacon
SI00068096	siRNA	Qiagen
SI02664809	siRNA	Qiagen
SI02664816	siRNA	Qiagen
SI03095792	siRNA	Qiagen
SP00002881	siRNA_Pool	Qiagen
TRCN0000065113	siRNA	TBC
TRCN0000065114	siRNA	TBC
TRCN0000065115	siRNA	TBC
TRCN0000065116	siRNA	TBC
TRCN0000065117	siRNA	TBC

(d)

## Reagent Information

Name:	SP00002881
Type:	siRNA_Pool
Library:	<a href="#">Qiagen</a>
siRNA Id	Sequence
<a href="#">SI03095792</a>	CTGCTAGATCTGATCGAGT
<a href="#">SI00068096</a>	CAGTACGTATTGGCATTC
<a href="#">SI02664809</a>	ACGATCTTCAGAGTATGC
<a href="#">SI02664816</a>	CAGGTTTCAAGGOTAGTGA
Pool quality	
siRNAs	On-target Off-target No-target Efficiency[%]
4	4 0 0 57.555
Intended targets	
Gene	Transcripts (Hits)
9276	NM_004766.2 (4), NR_023350.1 (4)

(f)

## Phenotypes from internal screens

id	Title	Score	Locus Id	Activity Ch1	Reag	Access	Image
16	S2-His2B GFP cell viability	2.52	<a href="#">BKN_2_I20</a>	74.0 % (σ: 2.8%)	BKN20346		
24	S2-DGRC viability	2.61	<a href="#">BKN_2_I20</a>	69.1 % (σ: 1.4%)	BKN20346		

## Curated phenotypes from publications

Screen Id	Experiment	Reagent Id	Score Type	Score Cutoff	Score	Phenotype	Validated
2	<a href="#">Lipid storage</a>	beta'Cop	B-score	>= 2.0 / <= -1.7	ND	overstorage	NA
19	<a href="#">Chlamydia infection</a>	DRSC03492	Visual inspection	ND	ND	Less inclusions	NA
27	<a href="#">ERK signaling</a>	DRSC03492	weak, moderate, strong	ND	strong	ERK activation increased	Yes
35	<a href="#">Drosophila C virus infection</a>	DRSC03492	Visual inspection	ND	ND	Drosophila C virus replication down	NA
41	<a href="#">E. coli / S. aureus infection</a>	CG6699	Visual inspection / FACS	20% decreased phagocytosis	ND	Decreased phagocytosis	NA
48	<a href="#">Cell viability</a>	HFA03492	Z-score	3	0.4 (Kc167)	ND	NA
48	<a href="#">Cell viability</a>	HFA03492	Z-score	3	2 (S2R+)	ND	NA
50	<a href="#">Cytokinesis</a>	DRSC03492	Visual inspection		medium	Binucleate, low cell count	Yes
53	<a href="#">L. monocytogenes infection</a>	DRSC03492	Visual inspection	ND	Strong	Infection down	Yes
63	<a href="#">Hedgehog signaling</a>	DRSC03492	Z-score	>= 3 / <= -2	-2.34	Decreased Hedgehog reporter activity	Yes
66	<a href="#">Cell cycle</a>	beta'Cop	Z-score	>= 5 / <= -5	ND	G1 decreased	NA
111	<a href="#">Lipid storage</a>	CG6699	Visual inspection	ND	Strong	Normal/larger sized lipid droplets, more dispersed	NA
157	<a href="#">Mitochondrial Ca2+ and H+ regulation</a>	DRSC26650	Percent positive control	>= 25% (1), >= 50% (2), >= 75% (3)	3		Yes
157	<a href="#">Mitochondrial Ca2+ and H+ regulation</a>	DRSC03492	Percent positive control	>= 25% (1), >= 50% (2), >= 75% (3)	2		Yes

**Figure 2.** Example of a database search for human COPB2. (a) The entry page allows for gene, reagent and phenotype queries or to 'List all Screens'. Here COPB2 was queried. (b) The 'Gene Info' tab provides detailed information about the queried gene and linkouts to other sources. One *Drosophila* homolog, beta'Cop, was found (f). (c) List of all RNAi reagents available that target COPB2. The 'Library' link leads to more information about the RNAi library, the 'Reagent Id link' provides more information about the RNAi reagent. (d) Detailed information for the Qiagen siRNA pool SP00002881 containing four siRNA sequences. Sequence information, information about 'On-' and 'Off-target' hits defining the predicted specificity as well as the predicted efficiency are presented. All transcripts of the targeted gene covered by the siRNAs are listed, with the number of siRNA hits in braces. (e) All reported phenotypes are listed in the 'Phenotype' tab. Data from three quantitative viability assays is available for COPB2. 'Score' (z-score) and 'Activity' (activity normalized to negative controls) columns provide a measure for the phenotype strength and reproducibility (given by the standard deviation). Also images (raw, segmented and phenotypic classified) from one high-content screen performed in HeLa cells are available. Clicking on the thumbnail enlarges the images. COPB2 was also found causing phenotypes in four published screens. The column 'Experiment' assigns a short name to the screen. The link can be followed up to obtain detailed information about the experiment. The other columns show the utilized reagent, the 'Score type' and 'Score Cutoff' used for the analysis, the actual 'Score' and 'Phenotype'. The column 'Validated' states whether a phenotype was retested (e.g. by a second RNAi reagent or by secondary assays). (f) Phenotypes for the *Drosophila* homolog beta'Cop. Data from two quantitative viability assays is presented. In addition data from 12 published screens is available.



targeted genes in several cell lines. The large number of screens and phenotypes hosted by the database reveals possible pleiotropic phenotypes of a candidate. Finally, the availability of data from two organisms enables cross-species comparisons of phenotypes, which will be extended by other organisms when screening becomes available.

The focus of GenomeRNAi is to present phenotypes in the context of genomic information. The database is not limited to a single organism and reports quantitative, literature-curated and imaging data. Together with the underlying pipeline for mapping of RNAi reagents from different libraries and phenotypes it is unique compared to other RNAi databases (16,19).

The integration and validation of data from primary publications is still a major issue due to the lack of standards on minimal information that need to be provided from large-scale screening approaches. There is also no general ontology to uniquely describe cellular phenotypes and literature reports vary in how phenotypes are described and what level of detail is provided. With GenomeRNAi we try to provide common terms for the 'type' of screen and the applied 'assay', which helps to identify and compare similar screens. To facilitate the cross-correlation of numerical phenotypic data and to document the screen-analysis-route we implemented the upload of data analyzed using the R/Bioconductor package cellHTS2 (20) for internal screens and plan to make this upload function available to users in the future. A screening dataset could then be analyzed and documented online using the implementation of cellHTS2 as web-tool (web-cellHTS; <http://web-cellhts2.dkfz.de>) before uploading the analyzed data to GenomeRNAi. The availability of *in vivo* RNAi libraries for *Drosophila* now enables genome-scale studies in the whole fly. Although the database already contains *Drosophila in vivo* phenotypes, more datasets will be added when available. In addition, the integration of RNAi phenotypes with other genomic data sets, such as RNA-Seq data, will be implemented in the future.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENT

We are grateful to Thomas Sandmann for critical comments on the manuscript.

## FUNDING

Fellowship of the German National Academic Foundation (to T.H.); fellowship by the MISTI's MIT Germany Program (to E.K.); grants from the Helmholtz Alliance for Systems Biology and the Emmy Noether Program of the German Research Council (partial). Funding for open access charge: German Cancer Research Center.

*Conflict of interest statement.* None declared.

## REFERENCES

- Chapman,E.J. and Carrington,J.C. (2007) Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.*, **8**, 884–896.
- Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.
- Boutros,M. and Ahringer,J. (2008) The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.*, **9**, 554–566.
- Horn,T., Arziman,Z., Berger,J. and Boutros,M. (2007) GenomeRNAi: a database for cell-based RNAi phenotypes. *Nucleic Acids Res.*, **35**, D492–D497.
- Kulkarni,M.M., Booker,M., Silver,S.J., Friedman,A., Hong,P., Perrimon,N. and Mathey-Prevot,B. (2006) Evidence of off-target effects associated with long dsRNAs in *Drosophila melanogaster* cell-based assays. *Nat. Methods*, **3**, 833–838.
- Ma,Y., Creanga,A., Lum,L. and Beachy,P.A. (2006) Prevalence of off-target effects in *Drosophila* RNA interference screens. *Nature*, **443**, 359–363.
- Shah,J.K., Garner,H.R., White,M.A., Shames,D.S. and Minna,J.D. (2007) siR: siRNA Information Resource, a web-based tool for siRNA sequence design and analysis and an open access siRNA database. *BMC Bioinformatics*, **8**, 178.
- Ramadan,N., Flockhart,I., Booker,M., Perrimon,N. and Mathey-Prevot,B. (2007) Design and implementation of high-throughput RNAi screens in cultured *Drosophila* cells. *Nat. Protoc.*, **2**, 2245–2264.
- Boutros,M., Kiger,A.A., Armknecht,S., Kerr,K., Hild,M., Koch,B., Haas,S.A., Paro,R. and Perrimon,N. (2004) Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science*, **303**, 832–835.
- Foley,E. and O'Farrell,P.H. (2004) Functional dissection of an innate immune response by a genome-wide RNAi screen. *PLoS Biol.*, **2**, E203.
- Dietzl,G., Chen,D., Schnorrrer,F., Su,K.C., Barinova,Y., Fellner,M., Gasser,B., Kinsey,K., Oppel,S., Scheiblaue,S. *et al.* (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature*, **448**, 151–156.
- Kent,W.J. (2002) BLAT—the BLAST-like alignment tool. *Genome Res.*, **12**, 656–664.
- Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
- Stein,L.D., Mungall,C., Shu,S., Caudy,M., Mangone,M., Day,A., Nickerson,E., Stajich,J.E., Harris,T.W., Arva,A. *et al.* (2002) The generic genome browser: a building block for a model organism system database. *Genome Res.*, **12**, 1599–1610.
- Stajich,J.E., Block,D., Boulez,K., Brenner,S.E., Chervitz,S.A., Dagdigan,C., Fuellen,G., Gilbert,J.G., Korf,I., Lapp,H. *et al.* (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res.*, **12**, 1611–1618.
- Flockhart,I., Booker,M., Kiger,A., Boutros,M., Armknecht,S., Ramadan,N., Richardson,K., Xu,A., Perrimon,N. and Mathey-Prevot,B. (2006) FlyRNAi: the *Drosophila* RNAi screening center database. *Nucleic Acids Res.*, **34**, D489–D494.
- Sayers,E.W., Barrett,T., Benson,D.A., Bryant,S.H., Canese,K., Chetvernin,V., Church,D.M., DiCuccio,M., Edgar,R., Federhen,S. *et al.* (2009) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **37**, D5–D15.
- Coutinho,P., Parsons,M.J., Thomas,K.A., Hirst,E.M., Saude,L., Campos,I., Williams,P.H. and Stemple,D.L. (2004) Differential requirements for COPI transport during vertebrate early development. *Dev. Cell*, **7**, 547–558.
- Sims,D., Bursteinas,B., Gao,Q., Zvelebil,M. and Baum,B. (2006) FLIGHT: database and tools for the integration and cross-correlation of large-scale RNAi phenotypic datasets. *Nucleic Acids Res.*, **34**, D479–D483.
- Boutros,M., Bras,L.P. and Huber,W. (2006) Analysis of cell-based RNAi screens. *Genome Biol.*, **7**, R66.