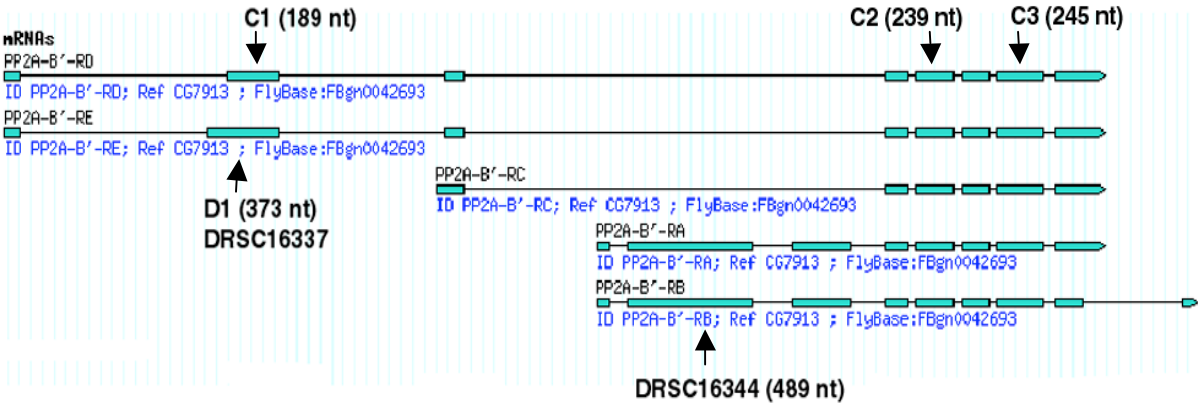
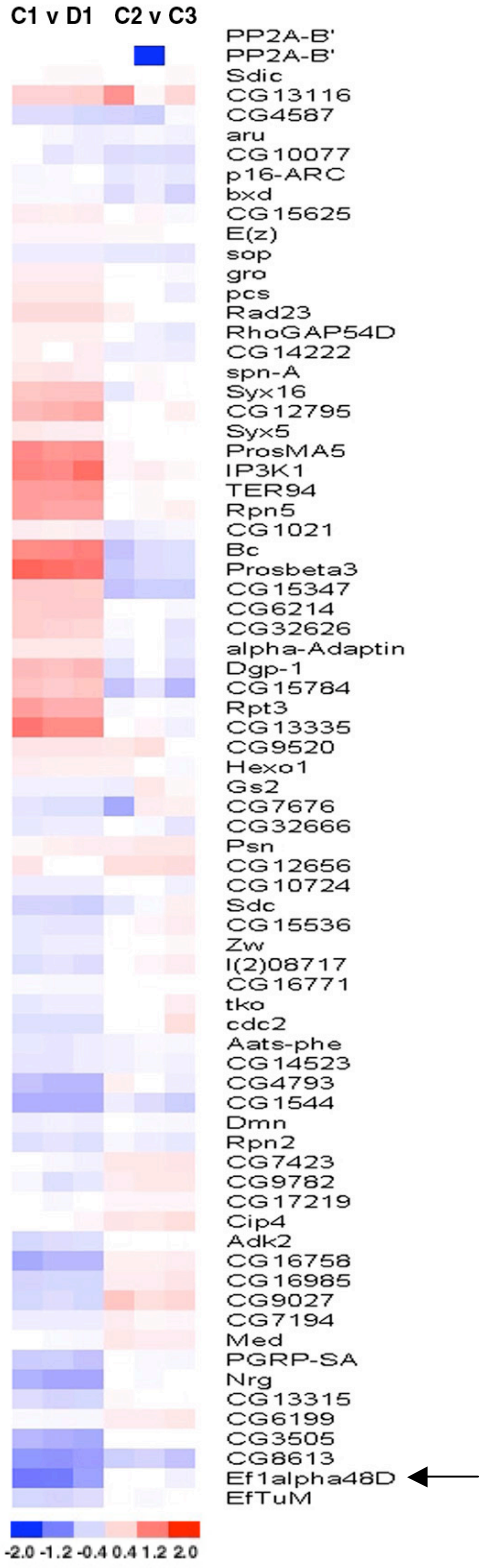
Supplementary Figure 1. Microarray analysis of off-target effects.

Supplementary Figure 1a . Schematic of the genomic region of the *PP2A-B'* gene.



PP2A-B' exons are shown as filled rectangles. Regions targeted by the various dsRNAs are indicated with an arrow, together with the nomenclature of the dsRNAs used in the text and their length.

Supplementary Figure 1b. Comparison of microarray expression profiles after treatment with pairs of dsRNAs directed against *PP2A-B'*.

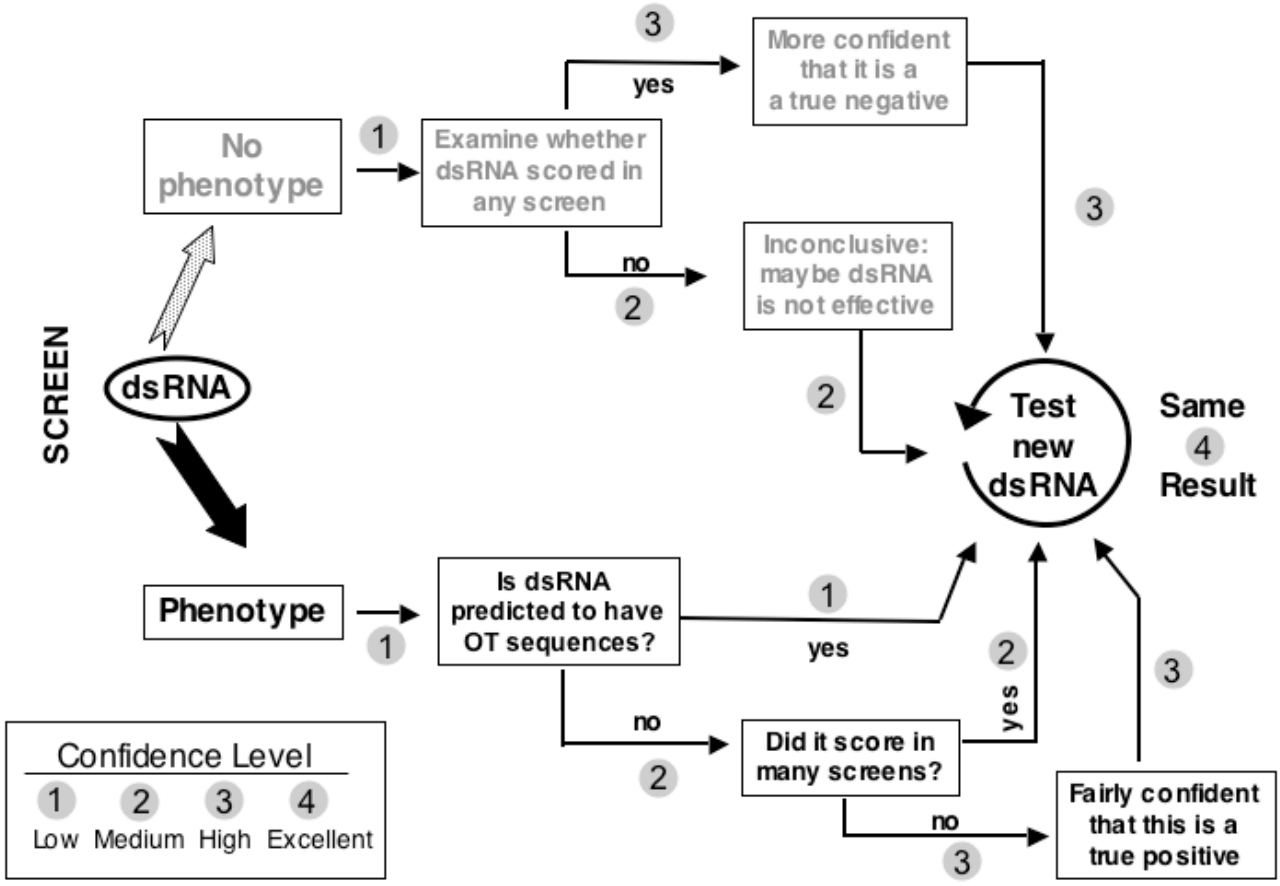


Drosophila melanogaster SL2 cells were treated by the bathing method with equal amounts (15 µg) of D1, C1, C2, and C3 long dsRNAs. The D1 and C1 dsRNAs target the PP2A-B' –RD and PP2A-B' –RE splice forms, while the C2 and C3 dsRNAs target all splice forms of PP2A-B'. For each long dsRNA, triplicate cultures were treated for 3 days before total RNA was isolated and reverse transcribed (see **Supplementary Methods** online). Labeled single strand cDNAs from D1 and C1 treated cells were mixed and competitively hybridized for a standard two-color (Cy3/Cy5) analysis to custom cDNA microarrays (C1 vs. D1 panel). For the comparison, three independent competitive hybridizations were carried out between C1 and D1 treated samples and between C2 and C3 treated samples. Virtually identical gene expression patterns were observed in the three independent experiments, demonstrating that changes in gene expression resulting from a particular dsRNA were reproducible and, thus, did not reflect either minor fluctuations in detection by microarray or inter-experiment fluctuations in transcript expression. The MIDAS V2.19 software was used to normalize the raw cDNA microarray signals according to the LOWESS method. A t-test was performed and p value of 0.001 for FDR was set to identify genes that were differentially expressed in each of the triplicates. The log2-transformed fold-changes calculated for the differentially expressed genes (DEGs) were used to perform a hierarchical clustering[1] based on the centroid-linkage method. A similar two color competitive hybridization was performed with cDNAs from C2 and C3 treated cultures (panel C2 vs. C3).

REFERENCES

1. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* **95**, 14863-8. (1998).

Supplementary Figure 2. Flow-chart and recommendations for assessing the specificity of dsRNAs in genome-wide-screens.



Diagrammed here are general considerations to be used for all dsRNAs reported in publications or RNAi databases. The main steps involve an analysis of predicted off-targets (threshold placed at 19 nt) for any given dsRNA which scores in an assay and how often that dsRNA scores in other assays. Although we indicate relative confidence levels, it is important to stress that off-target effects can arise from silencing transcripts via the miRNA pathway which only requires a 7-8 nt homology and can be hard to predict computationally[1]. Therefore, we strongly recommend to test a second and preferably a third dsRNA, targeting the same gene as the initial dsRNA associated with a phenotype, to insure that the measured activity results from the specific knockdown of

the intended target. In order to facilitate the implementation of these guidelines, the DRSC has replaced all dsRNAs predicted to have off-targets (due to sequence homology to other targets) in our collection, and has put in place a mechanism by which screeners will be provided with the means of confirming their primary results using a set of distinct dsRNAs. Lastly, as more screen results become publicly available, individual dsRNAs that have been associated with a phenotype in a screen can be checked against the list of dsRNAs reported as hits in published screens. Such knowledge can be used to gauge the specificity of dsRNA and may help refine ultimately the choice of candidate genes to be actively pursued.

REFERENCES

1. Birmingham, A. et al. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat Methods.* **3**, 199-204. (2006).

Supplementary Table 3. Amplicon ID for dsRNAs used in Figure 3

Gene	DRSC	Amplicon
CG12155	DRSC17826	Amp1
CG12155	DRSC31035	Amp2
CG12155	DRSC31036	Amp3
CG30421	DRSC04532	Amp1
CG30421	DRSC30735	Amp2
CG30421	DRSC30736	Amp3
CG32791	DRSC17892	Amp1
CG32791	DRSC31039	Amp2
CG32791	DRSC31040	Amp3
CG3563	DRSC13053	Amp1
CG3563	DRSC30936	Amp2
CG3563	DRSC30937	Amp3
shot	DRSC05459	Amp1
shot	DRSC30757	Amp2
shot	DRSC30758	Amp3
Smox	DRSC18716	Amp1
Smox	DRSC31057	Amp2
Smox	DRSC31058	Amp3
trio	DRSC08527	Amp1
trio	DRSC30819	Amp2
trio	DRSC30820	Amp3

Details on specific sequences for each dsRNAs can be found using the tool Gene and Amplicon Lookup available at http://flyrnai.org/cgi-bin/RNAi_gene_lookup_public.pl

Supplementary Table 4. Phenotype of multiple amplicons targeting seven genes in ERK activation assay.

	smox		CG32791		CG3563		trio		CG12155		CG30421		shot	
Amp1 19nt OT	138		121		111		78		76		58		48	
	%Luc	***p***	**%Luc**	***p***	**%Luc**	***p***	**%Luc**	***p***	**%Luc**	***p***	**%Luc**	***p***	**%Luc**	***p***
Amp1 (+)	121	$2x10^{-8}$	81	$4x10^{-7}$	79	$7x10^{-11}$	65	$1x10^{-15}$	80	$3x10^{-7}$	88	$6x10^{-3}$	78	$1x10^{-9}$
Amp2 (-)	99	0.7	98	0.4	102	0.5	98	0.3	101	0.8	97	0.3	98	0.2
Amp3 (-)	104	0.1	107	0.3	96	0.6	101	0.7	102	0.2	92	0.1	98	0.8

The table describes the phenotype of each amplicon in the ERK activation assay, relative to a luciferase dsRNA negative control, and the *p* value for each comparison. Only the first amplicon for each gene (with the number of predicted 19nt off-targets given below the gene name) significantly affected ERK activation denoted by +. Note there was no correlation between phenotype and extent of mRNA knockdown (see **Fig. 3**), indicating the phenotype of the first amplicon is likely due to an off-target effect.

Supplementary Table 5. Assay reproducibility and discovery rate for known components of the Wg, Hh and JAK/STAT pathways measured in 3 genome-wide RNAi screens.

Pathway	Reference	Genes passing primary	Genes tested in secondary	Genes verified in secondary	Canonical genes identified
Wg	DasGupta *et al.* [1]	238	238	213/238 89.5%	15/17 88%
Hh	Nybakken *et al.* [2]	509	255	204/255 80%	12/15 80%
JAK/STAT	Baeg & Zhou *et al.* [3]	474	286	202/286 70.6%	5/6 83%

Data reported in the three published screens were used to compare the respective assay reproducibility by assessing the number of genes that satisfied the criteria for causing a phenotype in the primary screen and the subset of these that satisfied the criteria of secondary screens (71-90% reproducibility). In addition, the number of genes identified in each screen and corresponding to canonical pathway components was compared to the expected number of genes known to belong to this category.

REFERENCES

1. DasGupta, R., Kaykas, A., Moon, R. T. & Perrimon, N. Functional Genomic Analysis of the Wnt-Wingless Signaling Pathway. *Science* **308**, 826-833 (2005).

2. Nybakken, K., Vokes, S. A., Lin, T. Y., McMahon, A. P. & Perrimon, N. A genome-wide RNA interference screen in Drosophila melanogaster cells for new components of the Hh signaling pathway. *Nat Genet.* **37**, 1323-32. Epub 2005 Nov 20. (2005).

3. Baeg, G. H., Zhou, R. & Perrimon, N. Genome-wide RNAi analysis of JAK/STAT signaling components in Drosophila. *Genes Dev.* **19**, 1861-70. Epub 2005 Jul 29. (2005).

Supplementary Methods

Custom cDNA-based Microarray: DNA probes to a set of approximately 3,000 strategically selected genes were generated by PCR. These probes were obtained from the PCR amplicons used to generate the dsRNA library[1]. The 3,000 genes represented on the array include genes for all the Kinases (350) and Phosphatases (161) encoded in the *Drosophila* genome, known/canonical components and targets of the major signaling pathways such as Wg/Wnt, Hh, TGF-β, JAK/STAT, NFκB, Notch, RTK, JNK, and INR pathways[2, 3]; http://genome.med.yale.edu/Lifecycle). Approximately 1,000 of the 3,000 genes were randomly selected to facilitate normalization of signal intensity across the array. Only PCR products showing a clear and strong band were recovered using Montage PCR plates (Millipore). The DNA concentration was determined using PicoGreen reagent (Molecular Probes). The final concentration of DNA averaged 400 ng/μl. Samples were transferred to 384 well plates (Genetix) containing DMSO to give a final concentration of 200 ng/μl in 50% DMSO. The DNA probes were directly spotted onto Corning UltraGAPS coated glass slides using a Genetix QArray robot in the Biopolymer Facility at Harvard Medical School. Spots were arrayed in a 14 x 14 arrangement using 24 (4 x 6) Genetix 150μm diameter tip solid microarray pins with a center-to-center spacing of 315 μm. A number of controls were prepared and printed on the glass slides with the probe DNAs. These controls include Negative controls (water, 50% DMSO, sonicated salmon sperm DNA, human Cot1, mouse Cot1, GFP, CFP, YFP, LacZ, and Luciferase) to assess the degree of nonspecific hybridization, and Spiked controls (Lucidea Universal Scorecard, Amersham Biosciences).

Labeling and Hybridization: 25 μg of total RNA from the RNAi treated samples (D1, C1, C2, and C3) were reverse transcribed and labeled using Superscript RT II (Invitrogen) and Genisphere's 3DNA Array 350 labeling kit. First, the total RNA is reverse transcribed using the included deoxynucleotide triphosphate mix and special Cy3/Cy5 RTdT primers. Thus, total RNA from the D1- and C1- treated samples were reverse transcribed using complementary Cy3 and Cy5 RTdT primers. Similarly total RNA from C2- and C3- treated samples were reverse transcribed using complementary Cy3/Cy5 RTdT primers. Then, the complementary cDNAs are mixed (D1 and C1; C2 and C3) and hybridized to the cDNA array. Cy3/Cy5 Fluorescent 3DNA reagents are then hybridized to the microarray in a second hybridization step. The Cy3 and Cy5 fluorescent 3DNA reagents will hybridize to the cDNA because they include a "capture

sequence" that is complementary to a sequence on the 5' end of the Cy3/Cy5 RT primer respectively. Each competitive hybridization was independently performed in triplicate including a color-flip to preclude dye bias. Post hybridization, the microarrays were washed in SSC and scanned using PMT-adjusted Axon's GenePix 4000B microarray scanner.

Computational analyses: GenePix Pro 5.0 was used for raw intensity value calculations. GenePix Results files (.GPR) were used as input for MIDAS V2.19 software (Saeed et al. 2003) to normalize the raw cDNA microarray signals by LOWESS method[4]. We used t-test followed by the false discovery rate (FDR) adjustment[5] to detect differentially expressed genes. The FDR adjusted p-value cutoff was 0.01. We calculated the base 2 log of the fold-changes. The clustering was performed by a hierarchical clustering algorithm[6] with the centroid-linkage method.

REFERENCES

1. Boutros, M. et al. Genome-wide RNAi analysis of growth and viability in Drosophila cells. *Science* **303**, 832-5. (2004).

2. Boutros, M., Agaisse, H. & Perrimon, N. Sequential activation of signaling pathways during innate immune responses in Drosophila. *Dev Cell.* **3**, 711-22. (2002).

3. Neal, S. J., Gibson, M. L., So, A. K. & Westwood, J. T. Construction of a cDNA-based microarray for Drosophila melanogaster: a comparison of gene transcription profiles from SL2 and Kc167 cells. *Genome.* **46**, 879-92. (2003).

4. Quackenbush, J. Microarray data normalization and transformation. *Nat Genet.* **32**, 496-501. (2002).

5. Benjamini, Y. & Yekutieli, D. Quantitative trait Loci analysis using the false discovery rate. *Genetics.* **171**, 783-90. Epub 2005 Jun 14. (2005).

6. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* **95**, 14863-8. (1998).