

**From screening to function –
Evolutionary conservation of novel JAK/STAT
signal transduction pathway components**

Doctoral Thesis

In partial fulfillment of the requirements
for the degree “Doctor rerum naturalium (Dr. rer. nat.)”
in the Molecular Biology Program
at the Georg-August University Göttingen,
Faculty of Biology

submitted by

PATRICK MÜLLER

born in

Osnabrück

2007

AFFIDAVIT

Here I declare that my doctoral thesis entitled “From screening to function – Evolutionary conservation of novel JAK/STAT signal transduction pathway components” has been written independently with no other sources and aids than quoted.

Patrick Müller

Göttingen, January 2007

The majority of experiments in the present thesis was performed at the Max Planck Institute for Biophysical Chemistry in the Department of Molecular Developmental Biology (Göttingen, Germany) and at the German Cancer Research Center in the Boveri-Group Signaling and Functional Genomics (Heidelberg, Germany).

LIST OF PUBLICATIONS

Publication

Müller, P., D. Kutteneuler, V. Gesellchen, M. P. Zeidler, and M. Boutros. 2005. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature* 436: 871-5.

Selected Presentations

Pelte, N., **P. Müller**, M. P. Zeidler, and M. Boutros. 2007. Chifoumi is a novel negative regulator of JAK/STAT signalling. *48th Annual Drosophila Research Conference* (Philadelphia, USA).

Müller, P., and M. Boutros. 2006. CellScreen: A statistics package for the analysis of cell-based RNAi screens. *47th Annual Drosophila Research Conference* (Houston, USA).

Müller, P., M. Boutros, and M. P. Zeidler. 2005. dBRWD3 is a novel regulator of the JAK/STAT signal transduction pathway. *19th European Drosophila Research Conference* (Eger, Hungary, oral presentation).

Müller, P., and M. P. Zeidler. 2005. Identification and characterization of a phosphatase negatively regulating the JAK/STAT signal transduction pathway in *Drosophila*. *46th Annual Drosophila Research Conference* (San Diego, USA, oral presentation).

CONTENTS

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	1
ABSTRACT	2
LIST OF FIGURES	3
LIST OF TABLES.....	4
ABBREVIATIONS AND NOMENCLATURE	5
INTRODUCTION	6
The logic of developmental signaling control	6
The JAK/STAT signal transduction pathway.....	9
<i>Drosophila melanogaster</i> as a model organism to study JAK/STAT signaling	16
Dissection of cellular pathways by RNAi	21
MATERIALS AND METHODS	26
Analysis and manipulation of nucleic acids	26
Quantification of nucleic acid concentrations	26
Polymerase chain reaction (PCR)	26
Primer design	26
Preparation of plasmid DNA	26
Isolation of genomic DNA from single flies	28
Large-scale preparation of <i>Drosophila</i> genomic DNA.....	28
Sequencing.....	28
Restriction digest of DNA	28
DNA extraction from agarose gels.....	28
DNA ligation.....	28
Transformation of bacterial cells	30
Generation of DNA vectors	30
Generation of dsRNAs and siRNAs.....	32
QuantiGene assays	33
Analysis and manipulation of proteins.....	33
Quantification of protein concentrations	33
Polyacrylamide gel electrophoresis (PAGE)	34
Immunoblotting.....	34
Immunoprecipitation experiments	34
Protein sequence analysis of complex samples by LC-MS/MS	35
Cell culture	36
Cell lines	36
Transfections.....	36
Generation of stable cell lines.....	36
High-throughput RNAi screening.....	36

CONTENTS

Epistasis experiments in cells	37
Candidate phenotypes in cells.....	38
Microscopy	39
Preparation of samples from cell culture experiments	39
Epifluorescence microscopy	39
Confocal microscopy.....	39
Determination of cell growth rates.....	39
Computational analyses.....	39
Determination of candidate hits in the genome-wide RNAi screen.....	39
Sequence analysis of candidate hits in the genome-wide RNAi screen.....	40
Analysis of LC-MS/MS data	40
Protein interactions in FlyNet	41
Protein interactions in OsPrey	41
<i>Drosophila</i> genetics	41
Fly strains and maintenance.....	41
Generation of transgenic flies	41
Mobilization of P-elements.....	43
Genetic interaction assays	43
Collection and fixation of embryos.....	43
Whole mount RNA <i>in situ</i> hybridization.....	43
Germline clones	44
RESULTS.....	45
Dissection of signaling processes by RNAi	50
Design of a genome-wide RNAi screen.....	51
Genome-wide RNAi screening.....	52
Data analysis of the genome-wide RNAi screen	53
Normalization approaches	53
Validation of primary screening hits using an independent reporter	67
Chromosomal clustering of novel JAK/STAT modulators	67
Gene ontology classification.....	68
Comparison of the present data to published protein interaction studies.....	68
Epistasis analysis.....	70
Implementation of an interactive publicly accessible website	72
Functional conservation of orthologs.....	73
Selection of homologs	74
Establishment of human JAK/STAT assays.....	74
A small-scale screen for human JAK/STAT pathway modulators	78
Assessing the specificity of siRNA induced human JAK/STAT phenotypes	79
Ptp61F negatively regulates JAK/STAT signaling.....	82
BRWD3 is a novel component of the JAK/STAT pathway.....	86
<i>dBRWD3</i> mutants phenocopy a <i>stat92E</i> mutant wing phenotype	89
<i>dBRWD3</i> mutant germline clones have a defect in egg laying.....	90
<i>dBRWD3</i> acts downstream of the pathway kinase Hop	91
<i>dBRWD3</i> and Stat92E may physically interact.....	92
Analysis of Stat92E and <i>dBRWD3</i> associated subproteomes.....	94

CONTENTS

dBRWD3 is localized in the nucleus.....	99
DISCUSSION.....	102
A genome-wide RNAi screen to identify JAK/STAT regulators.....	102
Novel evolutionarily and functionally conserved pathway regulators.....	109
The role of BRWD3 in the JAK/STAT pathways of flies and humans.....	116
SUMMARY AND CONCLUSIONS	122
REFERENCES	123
APPENDIX	137
Supplementary Script 1. Source code for the analysis of the RNAi screen	137
Supplementary Tutorial. Computational analysis of the RNAi screen dataset	159
Supplementary Script 2. ImageJ macro for approximation of cell numbers.....	165
Supplementary Table 1. Sequence and cytological information for screen hits.....	166
Supplementary Table 2. Human homologs of <i>Drosophila</i> genes.....	167
Supplementary Table 3. Human disease homologs of <i>Drosophila</i> genes.....	168
Supplementary Figure 1. Interactions between random gene sets (Osprey).....	169
Supplementary Figure 2. Interactions between RNAi screen candidates (FlyNet)....	170
Supplementary Table 4. List of homologs for which siRNAs were generated.....	171
Supplementary Table 5. Results of growth analysis after siRNA treatment.....	172
Supplementary Table 6. Identity of siRNAs used in pools and individually.....	173
Supplementary Table 7. Common proteins in datasets identified by LC-MS/MS.....	174
Curriculum Vitae	175

ACKNOWLEDGEMENTS

I would like to thank Dr. Martin P. Zeidler and Prof. Dr. Herbert Jäckle for their support both scientifically and personally. I feel that the exposure to this unique combination of scientific backgrounds and ways of thinking has added value to my personal development. In the Department of Molecular Developmental Biology, I would like to thank all the people, whom I have worked with over the last years. These are in particular Cristiana Hentea, Kai Nolde, Anna Lena Ernst, Iris Plischke, Sabine Häder, Peter Karsten, Drs. Tina Mukherjee, Natasha Arbouzova, Ralf Pflanz, Ulrich Nauber and Alf Herzig.

The present work would not have been possible without the close collaboration with Dr. Michael Boutros. All the people in his lab have helped me a lot. In particular, I would like to thank Thomas Horn, Kerstin Bartscherer, Viola Gesellchen, David Kутtenkeuler, Mona Stricker, Barbara Mosterman, Dr. Nadège Pelte and Dr. Dierk Ingelfinger.

Proteomic analysis was performed in collaboration with the Bioanalytical Mass Spectrometry Group at the Max Planck Institute for Biophysical Chemistry. I would like to thank Dr. Henning Urlaub, Dr. Mads Grønborg, Monika Raabe and Uwe Pleßmann for their help in obtaining and analyzing the data.

A central feature of the graduate program ‘Molecular Biology’, the framework in which the present study was performed, is the thesis committee helping and promoting graduate students with additional advice. I would like to thank Prof. Dr. Reinhard Lührmann and in particular Dr. Martin P. Zeidler and Prof. Dr. Mary Osborn for the important advice and guidance I received within and also outside of the thesis committee meetings. I would also like to thank Dr. Steffen Burkhard, the graduate program coordinator, for his very helpful and enthusiastic dedication.

My time in the Max Planck Institute for Biophysical Chemistry has been wonderful. Extracurricularly responsible are in part my lunch mates Drs. Ralf Jauch and Ajay Pobbati, the MPI String Quartett (Miriam Schneider, Rebecca Medda, Dr. Oliver Lange), the Horizons Organization Committee and all the members of the PhD Committee for improving the situation for PhD students at the MPI.

Dr. Karin Schmidt has crystallized my interests by introducing me to the fascination of life sciences, and I enjoyed the many music sessions and discussions we had together. I would also like to thank Karin for proofreading the present work.

I would like to thank the Studienstiftung des deutschen Volkes (German National Academic Foundation) for the support of my undergraduate and graduate studies.

I am most grateful to my non-scientific advisory committee Anja Müller-Netter and Ella Margarethe Müller as well as Werner and Heidi Müller for the best non-scientific support and advice I could imagine.

ABSTRACT

Signal transduction pathways mediating the exchange of information between cells are essential for development, cellular differentiation and homeostasis. Their dysregulation is also frequently associated with human malignancies. The Janus tyrosine kinase/signal transducer and activator of transcription (JAK/STAT) cascade represents one such signaling pathway whose evolutionarily conserved roles include cell proliferation and hematopoiesis. In the present study, a systematic genome-wide screen for genes required for JAK/STAT pathway activity has been performed. By analyzing 20,026 RNA interference (RNAi)-induced phenotypes in cultured *Drosophila melanogaster* hemocyte-like cells, a total of four previously known and 86 novel and uncharacterized genes were discovered. Subsequently, cell-based epistasis experiments were used to classify these modulators based on their interaction with known components of the signaling cascade. To analyze the functional conservation of these novel components throughout evolution, putative human homologs of the candidates found in *Drosophila* were targeted in human cell culture systems to assess the activity of human STATs upon knockdown of candidate modulators. Interestingly, 30 of the human homologs display a similar JAK/STAT phenotype to their *Drosophila* counterparts. In addition to multiple human disease gene homologs, the protein tyrosine phosphatase Ptp61F and the *Drosophila* homolog of BRWD3, a bromo-domain-containing protein disrupted in leukemia, were found in the RNAi screen. *In vivo* analysis of these two novel pathway regulators demonstrates that disrupted *dBRWD3* and overexpressed *ptp61F* function as suppressors of leukemia-like blood cell tumors. Moreover, dBRWD3 is localized in the nucleus and can physically interact with *Drosophila* STAT, likely to induce target gene activity. The present study represents a comprehensive identification of novel loci required for JAK/STAT signaling and provides molecular insights into an important pathway relevant for human cancer. Human homologs of identified pathway modifiers may constitute targets for therapeutic interventions.

LIST OF FIGURES

Figure 1. Principles of signal transduction.....	8
Figure 2. The canonical JAK/STAT pathway	11
Figure 3. DNA binding specificities of different STAT proteins.....	12
Figure 4. Mechanism of gene silencing by RNA interference	23
Figure 5. Scheme of multimerized reporter.....	47
Figure 6. Development of an improved JAK/STAT reporter assay	48
Figure 7. Dissection of the JAK/STAT pathway by RNAi.....	51
Figure 8. Design of a genome-wide RNAi screen for JAK/STAT signaling factors	52
Figure 9. Pilot screen shows reproducibility and robustness	53
Figure 10. Screenshot of an example session in <i>CellScreen</i>	54
Figure 11. Median centering of screening plates.....	55
Figure 12. Properties of the data distribution from the genome-wide RNAi screen	56
Figure 13. Plots of an individual plate to assess data quality.....	58
Figure 14. Dynamic range of screening plates	59
Figure 15. Overview of primary RNAi screen data and normalization methods	60
Figure 16. Detection of spatial screening artefacts.....	62
Figure 17. Exemplary comparison of normalization approaches	64
Figure 18. Analysis of JAK/STAT activity modulators	67
Figure 19. Previously published interactions between RNAi screen candidates	70
Figure 20. Classification of JAK/STAT modulators by pathway position	71
Figure 21. Implementation of a web-version to access screening data.....	73
Figure 22. Simplified linear illustration of JAK/STAT signaling pathways.....	74
Figure 23. Analysis of cytokine induced posttranslational modifications	75
Figure 24. Transcriptional readout for human JAK/STAT pathways	77
Figure 25. Functionally conserved JAK/STAT signaling modulators.....	79
Figure 26. Effect of siRNAs on HeLa cell growth	81
Figure 27. Ptp61F is an evolutionarily conserved protein	82
Figure 28. Molecular function of Ptp61F.....	83
Figure 29. BRWD3 is an evolutionarily conserved protein	85
Figure 30. dBRWD3 functions as a JAK/STAT pathway component	86
Figure 31. Rescue of <i>hop^{Tum1}</i> -induced tumor formation	87
Figure 32. <i>dBRWD3</i> mutants phenocopy <i>stat92E</i> hypomorphic mutant effects	89
Figure 33. Flies with <i>dBRWD3</i> mutant germline clones are affected in egg laying.....	89
Figure 34. Human BRWD3 is specifically involved in STAT1 signal transduction.....	90
Figure 35. dBRWD3 acts downstream of the kinase Hop	91
Figure 36. Physical interaction of dBRWD3 and Stat92E.....	93
Figure 37. Identification of Stat92E and dBRWD3 associated subproteomes.....	94
Figure 38. Nuclear localization signals in BRWD3	98
Figure 39. Intracellular localization of dBRWD3 variants	99
Figure 40. A C-terminal dBRWD3 truncation acts dominant-negatively.....	101
Figure 41. Model of BRWD3 molecular function.....	120

LIST OF TABLES

Table 1. Evolutionary conservation of the JAK/STAT signal transduction pathway	10
Table 2. Mammalian factors activating the JAK/STAT pathway	14
Table 3. List of primers	27
Table 4. List of plasmids	29
Table 5. List of fly stocks	42
Table 6. Different output lists using different normalization approaches	63
Table 7. <i>Drosophila</i> JAK/STAT phenotypes (negative regulators)	65
Table 8. <i>Drosophila</i> JAK/STAT phenotypes (positive regulators)	66
Table 9. Expected and observed phenotype frequency	68
Table 10. Functional groups classified by InterPro prediction and GO	69
Table 11. Genetic interactions of <i>ptp61F</i> with <i>hop</i> ^{<i>Tuml</i>}	84
Table 12. Genetic interactions of <i>dBRWD3</i> with <i>hop</i> ^{<i>Tuml</i>}	88
Table 13. Proteins identified after α -Myc IP of Stat92E-Myc transfected cells	96
Table 14. Proteins identified after α -Flag IP of dBRWD3-Flag transfected cells	97
Table 15. Effect of <i>dBRWD3</i> ectopically expressed in <i>Drosophila</i> tissues	100
Table 16. Comparison of results from different JAK/STAT modulator screens	104
Table 17. Comparison of the experimental conditions for JAK/STAT RNAi screens...	107

ABBREVIATIONS AND NOMENCLATURE

Genes and their activities as well as alleles and transposon/transgene-construct names and symbols are italicized in the main text, whereas proteins are indicated in non-italic font with a capital initial letter. Wherever possible, the most recent gene names were used, and *x/y* indicates the *y* synonym for gene *x*. In some cases the nomenclature used in *Drosophila* and mammalian research is different (e.g. compare ‘Stat’ with ‘STAT’) and wherever possible, the most commonly found versions from the literature were used. A ‘d’ or an ‘h’ in front of a gene name is used to clarify that in this case the *Drosophila* or the human homolog, respectively, is indicated.

Below, the most important abbreviations used in the present work are listed.

1D	one-dimensional	miRNP	miRNA-protein complex
2D	two-dimensional	mRNA	messenger RNA
3D	three-dimensional	MYST	MOZ, YBF2/SAS 3, SAS 2, Tip60 family
aa	amino acid(s)	n	number
AngII	angiotensin II	na	not applicable / not available
APRE	Acute phase response element	NaCl	Sodium chloride
APS	Adaptor protein containing PH and SH2 domains	Na ₂ VO ₄	Sodium ortho vanadate
ATF6	Activating transcription factor 6	NCBI	National center for biotechnology information
ATM	Ataxia telangiectasia mutated protein	NH ₄ HCO ₃	Ammoniumhydrogencarbonat
B-CLL	B-cell chronic lymphocytic leukemia	NFAT	Nuclear factor of activated T cells
BLAST	Basic local alignment search tool	NF-κB	nuclear-factor κB
bp	basepair(s)	NHS	Normal horse serum
BS	Bloomington Stock Center	NICD	Notch intracellular domain
BSA	bovine serum albumine	NLS	Nuclear localization signal
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>	OSM	Oncostatin M
Ca ²⁺	Calcium ions	OTE(s)	Offtarget effect(s)
CaCl ₂	Calcium chloride	P	phosphoryl(-tyrosine or –serine)
Cdk	Cyclin dependent kinase	p...	plasmid
CG	Computed gene	p...(wt)	plasmid with a wildtype site
cg	collagen	p...Luc	plasmid with a luciferase gene
ChIP	chromatin immunoprecipitation	p...Mut	plasmid with a mutated site
CIS	Cytokine inducible SH2 containing protein	p...perf	plasmid with a perfect consensus site
CNS	Central nervous system	PAGE	Polyacrylamide gel electrophoresis
CNTF	Ciliary neurotrophic factor	PBS	Phosphate buffered saline
CREB	cAMP response element binding protein	PCR	Polymerase chain reaction
CTBP	C-terminal binding protein	PHIP	Pleckstrin homology domain interacting protein
Cub	Cubitus interruptus	PIAS	Protein inhibitor of activated STAT
d	day(s)	PKR	Protein kinase R
D1-SIE1	Cyclin D1 STAT5 induced element	PMSF	phenylmethanesulphonylfluoride
Da	Dalton	PP2A	Protein phosphatase 2A
DCAF	DDB1-Cul4-associated WD40 domain proteins	Prl	Prolactin
DFS	Dominant female sterile	PTP	Protein tyrosine phosphatase
DGRC	<i>Drosophila</i> genomics resource center	Rab5	Ras associated protein Rab5
DNA	Deoxyribonucleic acid	repeat	repeat associated siRNA
DNase	Deoxyribonuclease	Rel	Relish
Dome	Domeless	<i>Renilla</i>	<i>Renilla reniformis</i>
<i>Drosophila</i>	<i>Drosophila melanogaster</i>	RISC	RNA induced silencing complex
dsRNA	double-stranded RNA	RITS	RNA induced transcriptional silencing
DTT	Dithithreitol	RL	<i>Renilla</i> luciferase
ECM	extracellular matrix/milieu	RNA	Ribonucleic acid
EDTA	ethylenediaminetetraacetic acid	RNAi	RNA interference
EGTA	ethylene glycol tetraacetic acid	RNase	Ribonuclease
eIF-2	eukaryotic translation initiation factor 2	s	second(s)
ETS	erythroblast transformation specific	SDS	Sodium dodecyl sulfate
FBgn	Flybase gene	SH2	SRC homology 2
firefly	<i>Photinus pyralis</i>	Shh	Sonic hedgehog
FL	firefly luciferase	sIL6R	soluble IL6 receptor
FLP	Flipase	siRNA	small interfering RNA
Fos	FBJ murine osteosarcoma	SMAD	Sma and Mad related protein
FRT	Flipase recognition target sequence	SOCS	suppressors of cytokine signaling
Gal4	Gal4 transcription factor	SRC	Avian sarcoma virus protein
GAS	Interferon γ activated sequence	SREBP	Sterol regulatory element binding protein
GBP1	Guanylate binding protein 1	SRF	Serum response factor
GFP	Green fluorescence protein	STAT	signal transducer and activator of transcription
GI	gene identifier	STAT RE	STAT responsive element
GLI	Glioma associated oncogene	StIP	STAT interacting protein
GO	gene ontology	TAD	transactivation domain
Gp130	Membrane glycoprotein 130	TGF	Transforming growth factor
GPCR	G-Protein coupled receptor	TIF1	Transcriptional intermediary factor 1
h	hour(s)	TRIM33	Tripertite motif containing protein 33
HCl	Hydrochloric acid	Tris	Trishydroxymethylaminomethane
HeLa cells	Henrietta Lacks' cervical cancer cells, strain SS6	TSG	Tumor susceptibility gene
Hop	Hopscotch	UAS	Upstream activating sequence
IFN	Interferon	Upd	Unpaired
IL	Interleukin	Vps	Vacuolar protein sorting protein
IP	Immunoprecipitation	WB	Western Blot
JAK	Janus kinase	WD40	WD40 repeats
lacZ	β-galactosidase	WDR9	WD repeat domain 9
LB	Luria-Bertani broth	Wnt	Wingless and Int
LC-MS/MS	Liquid chromatography tandem mass spectrometry	ZnSO ₄	Zinc sulfate
LIF	Leukemia inhibitory factor		
M	mol/l		
MAD	median absolute deviation		
MADS box	domain in proteins Mcm1, Agamous, Deficiens, SRF		
Mef-2	Myocyte specific enhancer factor 2		
MgCl ₂	Magnesium chloride		
MGF	Mammary gland factor		
Mib	Mindbomb		
min	minute(s)		
miRNA	micro RNA		

INTRODUCTION

Multicellular organisms develop from a single cell, the fertilized egg, which divides to eventually produce all cells of the body. In order to differentiate into specialized organs, cells have to execute a correct transcriptional program, which is largely determined by signaling between cells via extracellular molecules. The recognition of molecules in the extracellular environment and the transduction of this information inside the cells are mediated by so-called signal transduction pathways. The discovery of these mechanisms can be traced back more than 100 years ago to experiments by Hans Spemann (Spemann 1901). These experiments eventually led to the identification of the ‘organizer’, which showed that cells in a host animal are able to follow a new direction of differentiation depending on the signals they received from a grafted tissue (Spemann and Mangold 1924).

The logic of developmental signaling control

During the last century, many regulatory mechanisms have been discovered which help to explain how cells receive the right amount of signal at the right time and how they interpret these signals. Generally, three types of signal interpretation can be distinguished (reviewed in Freeman and Gurdon 2002). In permissive ‘Threshold Inductions’, competent cells can only make a binary fate choice and the received signal determines when these cells should respond rather than determining the magnitude of response. In contrast, instructive ‘Concentration Dependent Inductions’ are more complex in that competent cells can make at least three fate choices depending on how much signal they receive. A concentration gradient of the signal, termed morphogen, in this process is created by the localized production of signal in one position, which tapers out with distance from the source. The ‘Community Effect’ represents a third principle of response, where cells have both inducing and responding capabilities so that the outcome depends on the number and vicinity of signaling cells.

Whatever the type of response may be, the response itself has to be regulated to achieve precision in the desired developmental outcome. Many regulatory mechanisms have evolved that are commonly used for different signal transduction pathways. The overall

simplified principle in signal transduction is that an extracellular signaling molecule, termed ligand, binds to the membrane bound receptor of a responding cell. The signal is then somehow transduced into the cell eventually leading to changes in gene activities. The most important regulatory step for the initiation of signaling appears to be the supply of active signaling molecule (Freeman and Gurdon 2002). The supply of such a ligand (e.g. morphogens may act *in vivo* in the range of 10^{-11} M of active concentration (Freeman and Gurdon 2002)) can be regulated in defined transcriptional patterns, which can be further shaped by the intracellular localization of the mRNA giving rise to the ligand after translation (e.g. Nilson and Schupbach 1999). A number of posttranslational mechanisms can add further regulatory control to the signal supply. This posttranslational control can be achieved, for example, by the regulated translocation through the secretory pathway (e.g. Tsruya et al. 2002), by the proteolytic activation of ligands (e.g. Peschon et al. 1998), by glycosylation (e.g. Harrison et al. 1998) and by attachment of lipids (e.g. Chuang and Kornberg 2000). Especially for morphogens, the supply of ligand in a concentration dependent manner is important. This concentration gradient can be established by modification, degradation and altering the stability of the ligand as well as by association with the extracellular matrix (e.g. Teleman et al. 2001).

Another level of regulatory control can be added by inhibiting the binding of the ligand signal to a receptor by other extracellular molecules (e.g. Holley et al. 1995), which may also be controlled by negative feedback loops (e.g. Chen and Struhl 1996), meaning that the signal itself induces the activity of an inhibitor thereby shutting down signaling activity. The harshest level of regulatory control in signal transduction is, of course, its complete termination. This can be achieved by the discontinuation of signal supply, through the removal of receptors from the membrane by endocytosis (e.g. Sorkin and Waters 1993) or by their transcriptional down-regulation (e.g. Sturtevant et al. 1994). Alternatively, antagonistic extra- or intracellular factors may be provided to terminate signaling, and cells can lose or change their 'competence' to respond to the signals without modulating receptor activity (e.g. Goldstein 1995, Xu et al. 2000).

Interestingly, only a small number of signal transduction pathways is sufficient for a wide variety of developmental decisions (Pires-daSilva and Sommer 2003). These pathways are repeatedly used in development and can lead to multiple different events depending

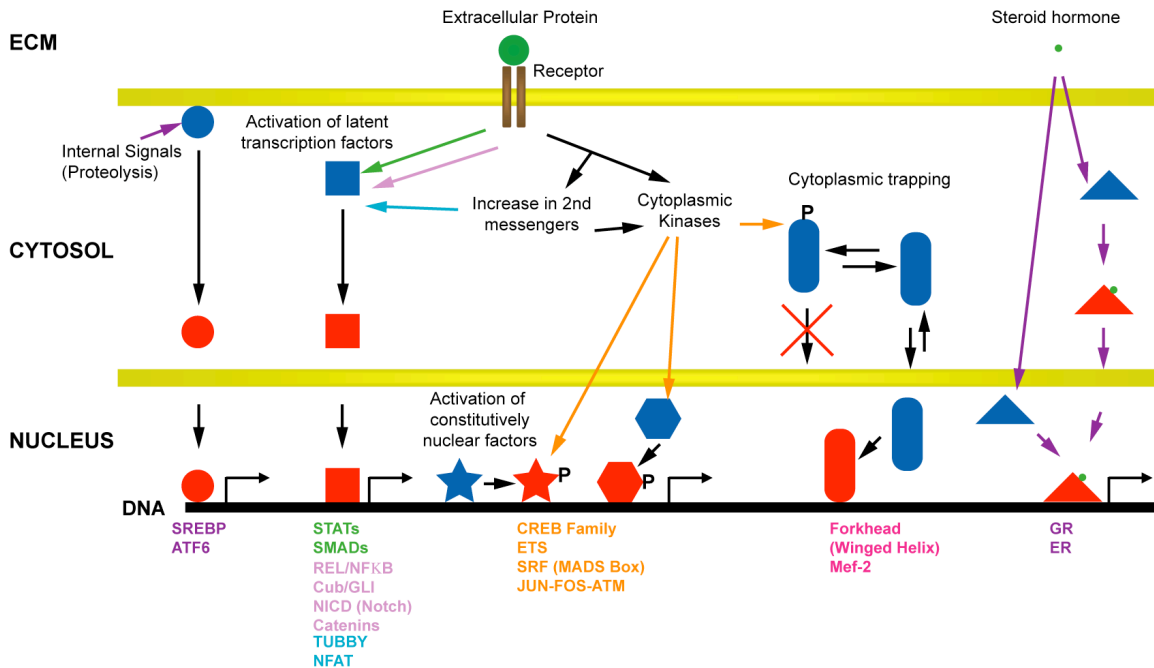


Figure 1. Principles of signal transduction from the extracellular environment to the nucleus. The extracellular protein signal is transduced through a receptor, which can lead to the activation of latent cytoplasmic transcription factors in a multitude of ways. These are from left to right (arrows indicate the mechanism; examples of the respective transcription factors are shown at the bottom of the figure): Activation by liberation through proteolysis (purple), directly by phosphorylation (green), through regulated phosphorylation or proteolysis (light purple) or through second messenger fluctuations (light blue). Kinases can phosphorylate resident nuclear proteins bound to DNA (star) or free in the nucleus (hexagon). Note that steroid hormones can diffuse through the membrane, and they bind and activate their receptors in the cytosol or in the nucleus (dark purple). Blue shapes indicate inactive transcription factors whereas red shapes represent the activated states. ECM indicates extracellular matrix, P stands for phosphorylated residues. Figure modified after Brivanlou and Darnell 2002.

on the time and place of activation, a phenomenon which can be subsumed under the term ‘pleiotropy’. Assuming a simplistic model, where signaling is transduced through only one receptor, spatial and temporal specificity can be achieved in principal by three different mechanisms. As a first example, a receptor may be activated by different ligands. Secondly, one ligand bound to the receptor may activate multiple intracellular signal transduction pathways (e.g. Pawson and Saxton 1999). Because signal transduction eventually ends in transcriptional events to achieve a particular cell fate, it is therefore conceivable, as a third example, that a single signal transduction pathway may activate several transcription factors. In addition to the overall ‘biography’ of a cell, the repertoire of transcription factors present in the receiving cell could thereby determine the outcome (e.g. Xu et al. 2000), and the combinatorial use of these proteins as a possibly unique

subset of constitutively active nuclear factors and regulatory transcription factors for every gene could ensure the right amount of the right protein at the right time for a given developmental program (Brivanlou and Darnell 2002). Figure 1 shows the diverse ways how transcription factors can be activated or regulated, e.g. by liberation through proteolysis (e.g. for the 'sterol response element binding protein' SREBP, which is regulated by internal sterol concentration), directly by phosphorylation (e.g. JAK/STAT and TGF- β pathways), through regulated phosphorylation or proteolysis (e.g. NF- κ B, Hedgehog, Notch and Wnt pathways) or through second messenger fluctuations (e.g. NFAT activation by Ca²⁺ increase and phosphoinositide dependent release of TUBBY). Constitutively nuclear factors, on the other hand, may be activated by serine kinases (e.g. c-Jun and c-Fos), whereas other proteins shuttle in and out of the nucleus and can be phosphorylated in the cytoplasm thereby preventing their reentry into the nucleus. Steroid hormones can bind and activate their receptors in the cytosol or in the nucleus, which then function as transcription factors (e.g. Glucocorticoid receptor (GR) and Oestrogen receptor (ER)). The interplay between these pathways and transcription factors is a key determinant to developmental decisions (Brivanlou and Darnell 2002).

The JAK/STAT signal transduction pathway

The Janus tyrosine kinase (JAK)/signal transducer and activator of transcription (STAT) cascade is one of the major signal transduction pathways in eukaryotes that allows the transfer of information among cells to choreograph the execution of transcriptional programs (Brivanlou and Darnell 2002). It has been conserved throughout evolution, and STATs are present in distantly related species such as the slime mold *Dictyostelium discoideum*, the nematode *Caenorhabditis elegans*, the fruitfly *Drosophila melanogaster*, the zebrafish *Danio rerio* as well as in the more complex mammals mouse, rat and human (Table 1, reviewed in Hou et al. 2002).

In the canonical model of JAK/STAT signaling, the non-receptor tyrosine kinase JAK is non-covalently associated with the intracellular carboxyterminal part of transmembrane receptors (Figure 2). Binding of an extracellular ligand to the extracellular aminoterminal domain of homo- or heterodimerized receptors possibly induces a conformational change

Table 1. Evolutionary conservation of the JAK/STAT signal transduction pathway as exemplified for the pathways present in mammals and *Drosophila melanogaster* (modified after Rawlings et al. 2004, Ungureanu et al. 2003).

	Mammals	<i>Drosophila melanogaster</i>
Ligands	many (see Table 2)	3 (Upd, Upd2, Upd3)
Receptors	many (see Table 2)	1-2 (Dome, CG14225*)
JAKs	4 (JAK1-3)	1 (Hop)
STATs	7 (STAT1-4, STAT5A, STAT5B, STAT6)	1 (Stat92E)
SOCS	8 (CIS, SOCS1-7)	3 (Socs16D*, Socs36E, Socs44A)
PIAS	5 (PIAS1/3, PIAS α , PIAS β , PIAS γ)	1 (dPIAS)

* Identification by sequence homology. Function in JAK/STAT signaling not yet established.

in the receptor, which brings two JAK molecules into juxtaposition. This allows the JAKs to tyrosine-phosphorylate each other as well as the receptor, thereby generating docking sites for the SRC homology 2 (SH2) domains of the latent cytoplasmic STATs, which are then themselves phosphorylated C-terminally by JAKs on an invariant tyrosine residue. Another protein StIP may serve as a bridging factor for the interaction between STATs and JAKs (Collum et al. 2000). The activity of JAKs can be positively and negatively regulated by the adaptor proteins SH2B β and APS, respectively (Kurzer et al. 2006), and JAKs may also have substrates other than STAT molecules (e.g. Shi et al. 2006, Takeshita et al. 1997). Other posttranslational modifications for the full activity of STATs, in addition to JAK-mediated tyrosine phosphorylation, may be serine phosphorylation (reviewed in Decker and Kovarik 2000), arginine methylation (Komyod et al. 2005, Meissner et al. 2004, Mowen et al. 2001) and lysine acetylation (Kramer et al. 2006, Wang et al. 2005, Yuan et al. 2005).

The STATs then dimerize by reciprocal phosphotyrosine-SH2 interactions to form an activated transcription factor (Chen et al. 1998). However, nonphosphorylated STATs have also been reported capable of forming dimers, although unable to bind DNA (Braunstein et al. 2003). STAT proteins are shuttled to the nucleus with the help of importins (Fagerlund et al. 2002) and may be retained in the nucleus after activation by

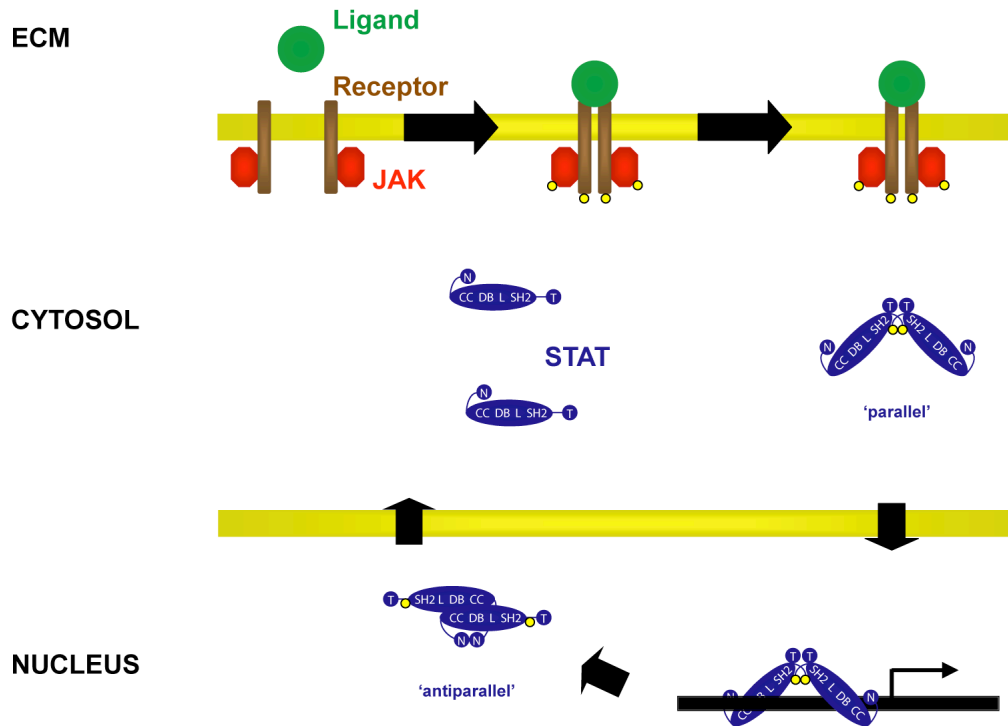


Figure 2. The canonical JAK/STAT pathway. Simplified schematic representation of the JAK/STAT signaling pathway. A ligand (green) in the extracellular milieu (ECM) binds to the receptor complex (brown) in the cell membrane associated with an intracellular Janus kinase (JAK, red). JAKs then auto-phosphorylate (yellow circles) and phosphorylate the dimerized receptor complex, thereby recruiting latent cytoplasmic STAT molecules (blue) to the receptor via an interaction between the phosphorylated receptor and the STAT SH2-domains. STATs then become phosphorylated by JAKs, dimerize, are translocated to the nucleus, bind in enhanceosomes and activate target genes. STATs are dephosphorylated following a structural rearrangement to leave the nucleus. ‘N’ indicates the N-terminal, ‘CC’ the coiled-coil, ‘DB’ the DNA binding, ‘L’ the linker, and ‘T’ the transactivation domain. Modified after Mertens et al. 2006.

cytokine stimulation (reviewed in Meyer and Vinkemeier 2004). There, they bind to consensus DNA sites within the promoters of genes as dimers or tetramers (John et al. 1999). They furthermore may interact with other transcription factors in so-called enhanceosomes, STATs often being the required factor for the increase in transcription (Darnell 1997, Lerner et al. 2003). Eventually, STATs need to be inactivated by dephosphorylation before they can leave the nucleus for another activation cycle (see below, McBride and Reich 2003, Figure 2).

The vast majority of STATs bind to specific palindromic motifs present in DNA (Ehret et al. 2001), the consensus binding site of which is conserved evolutionarily (Figure 3). This consensus binding site is represented by the DNA nucleotide sequence $TTC(N)_{2-4}GAA$, where $(N)_{2-4}$ indicates a spacer region of two to four nucleotides of any identity. Although

A

	Target gene	Species	Ligand(s)	Name of element	STAT binding sequence
STAT1	<i>α2-macroglobulin</i>	Rn	IFN γ , IL6	APRE	ATCC TTCTGGGAAT TCC
	<i>bcl-x</i> <i>GBP1</i>	Hs	LIF	GAS	GCAT TTCGGAGAA GACG GCAG TTCCAGGAAT CGG
STAT2	<i>IL20 Receptor α</i>	Hs	IFN-alfacon-1	GAS-like element	TTCCATGAA
STAT3	<i>gp130</i>	Hs	IL6, OSM	STAT binding element	CGCG TTACGGGAAT CGC
	<i>SOCS3</i>	Mm	LIF	STAT1/3 element	CATA TTACTCTAAAT CC
STAT4	<i>IFNγ</i>	Hs	IL12	STAT RE	TGCC TTCAAAGAA TCCC
STAT5	<i>β-casein</i>	Rn	Prl	prox. MGF site	GGAC TTCTTGGAAT TAA
	<i>CIS</i>	Hs	IL2	CIS1	GCGG TTCTAGGAAG ACG
	<i>cyclin D1</i>	Hs	IL3	D1-SIE1	GGCG TTCTTGGAAT G
STAT6	<i>eotaxin</i>	Hs	IL4	STAT6 element	AGGC TTCCCTGGAAT CTC
Stat92E	<i>eve 1</i>	Dmel	Upd		TTCCCCGAA
	<i>eve 2</i>	Dmel	Upd		TTCTAGGAA
	<i>dRaf</i>	Dmel	Upd		ATTCCGCGAAT
	Preferred	Dmel		consensus	TTCCCCGAAA

B

Figure 3. DNA binding specificities of different STAT proteins. (A) Shown are the preferred binding sequences for STATs from the species rat (Rn), human (Hs), mouse (Mm) and *Drosophila melanogaster* (Dmel). Red bases indicate the consensus binding nucleotides, green non-canonical ones (Ehret et al. 2001, Yan et al. 1996b, Brierley et al. 2006). Note that STAT6 preferably binds to sites with a spacer of four nucleotides. (B) Structure of the STAT3 protein bound to DNA (5'-TGCAT**TTCCCGTAAAT**CT-3', Becker et al. 1998) displayed in Cn3D (<http://130.14.29.110/Structure/CN3D/cn3d.shtml>). α -helical structures are shown as barrels, β -strands as sheets with arrows, domains are color-coded, and gold spheres indicate the backbone of the bound DNA with the bases pointing towards the interior. The N-terminus up to residue 319 is colored in purple, the blue (aa 320-347) and the brown domains include the DNA binding domain and the green domain includes the SH2 domain. For simplicity, only half of the homodimer is shown. The other dimer would be located approximately mirror-symmetrically to the first Stat3 monomer with their C-termini in close proximity and the overall homodimer forming a saddle-like structure on the DNA (Figure 2).

the sequence specificities for different STATs are very similar, they are not identical and the differences including the length and sequence of the interspersed spacer may govern specificity in the regulation of target genes (Ehret et al. 2001). Additional specificity of gene induction may come from the specific interaction with other transcription factors (e.g. c-Jun, c-Fos, Sp1, Glucocorticoid receptor (GR), p53, NF- κ B and others (Kramer et al. 2006, Lerner et al. 2003, Townsend et al. 2004, reviewed in Shuai 2000)), the interplay with other signaling pathways (e.g. TGF- β (Pitts et al. 2001), Notch (Kamakura et al. 2004) and MAPK pathways (David et al. 1995)) or by the state of the chromatin of a potential STAT target gene (Rusterholz et al. 1999). Transcription of the target gene after STAT binding to DNA can eventually be activated by the physical interaction with the basal transcription machinery through p300/CBP (reviewed in Shuai 2000).

Since the discovery of the first STAT in the early 1990s (Schindler et al. 1992), a total of seven different mammalian STATs (Table 1, Table 2) and many polypeptide ligands have been identified in mammals with a wide spectrum of function in development as well as in adult tissues including regulation of hematopoiesis, immune development, mammary gland development, lactation and adipogenesis (Table 2, reviewed in Levy and Darnell 2002, Rawlings et al. 2004). The 3D core structures (approximately residues 100 to 700 of their 750-850 aa sequence) of two phosphorylated STAT homodimers, STAT1 and STAT3, bound to DNA have been solved by crystallographic studies (Becker et al. 1998, Chen et al. 1998), confirming the modulatory structure of STATs suggested previously by sequence comparisons and mutagenesis studies (Darnell 1997). Both STAT1 and STAT3 protein dimers bound to DNA reveal a strikingly similar saddle like conformation (Figure 3B). The DNA binding domain lies in the center of the molecule surrounded by an N-terminal coiled-coil and the SH2 domain, which is followed by the phosphorylated tyrosine residue. The coiled-coil domain (beginning at residue 130) consists of a four-stranded helical coiled coil (Figure 3B (purple)). This represents an interface for a wide variety of protein-protein interactions. The β -pleated DNA binding domain (between residues 320 and 490, Figure 3B (blue)) shows a fold similar to the DNA binding domains of other transcription factors such as p53 or NF- κ B, and it is followed by a linker domain (Figure 3B (brown)), mutations in which affect the stability of DNA binding (Yang et al. 2002). Towards the C-terminus, the phosphorylated tyrosine residue

Table 2. Mammalian factors activating the JAK/STAT pathway (modified after Rawlings et al. 2004).

	Interferons			gp130 family			β c family	γ c family	Homo-dimeric Receptors	GPCRs
	α/β	γ	IL10	IL6, OSM*	IL12	Leptin				
JAK1	+	+	+	+	-	+	-	+	-	-
JAK2	-	+	-	+	+	+	+	+	+	+
JAK3	-	-	-	-	-	-	-	+	-	-
TYK2	+	-	+	+	+	-	-	+	+	+
STAT1	+	+	+	+	-	-	-	+	+	+
STAT2	+	-	-	-	-	-	-	-	-	+
STAT3	+	-	+	+	-	+	-	+	+	+
STAT4	-	-	-	-	+	-	-	+	-	-
STAT5A	+	+	-	+	-	-	+	+	+	+
STAT5B	+	+	-	+	-	-	+	+	+	+
STAT6	-	-	-	-	-	-	-	+	-	-

* e.g. IL6 (Interleukin 6), OSM (Oncostatin M), IL11, CNTF (Ciliary neurotrophic factor), LIF (leukemia inhibitory factor).

Data for β c family (e.g. IL3, IL5), γ c family (e.g. IL2, IL7, IL9, IL15, IL4, IL13), homodimeric receptors (growth hormone, erythropoietin, thrombopoietin) and the G-protein coupled receptors GPCRs (angiotensin, serotonin) is given cumulatively for each family due to space constraints. Note that STATs may also be directly phosphorylated by receptors with intrinsic tyrosine kinase activity, although this effect could also be mediated by other kinases such as Src (reviewed in Leaman et al. 1996).

+ indicates responsiveness, - indicates unresponsiveness.

(around aa 700) is preceded by a classical SH2 domain necessary for dimerization (between aa 580 and 680, Figure 3B (green)).

A domain not visible in the 3D structure is the C-terminal transactivation domain (TAD), which adds another 38-200 residues to the various proteins and probably governs specificity (Goenka et al. 2003). Another domain not included in the crystallographic analysis lies at the amino terminus (up to approximately residue 100). It is connected to the core via a flexible linker of approximately 24 aa (Chen et al. 1998) and appears to vary in structure among the different STATs (Xu et al. 1996). Nevertheless, the N-terminal domain plays a role in dimer-dimer formation, important for example in the STAT attachment to adjacent STAT binding tandem sites (Vinkemeier et al. 1996), phosphorylation and receptor recognition (Murphy et al. 2000), nuclear translocation as

well as dephosphorylation for inactivation (Strehlow and Schindler 1998). A recent report suggests that massive structural rearrangements facilitated by dimerization of the N-terminal domains from a 'parallel' (monomers of the dimer engage in reciprocal SH2-phosphotyrosine interactions with the phosphotyrosines buried in the dimer) to an 'antiparallel' form (with the phosphotyrosines exposed at the two ends of the dimer) inside the nucleus may be necessary to allow for dephosphorylation (Figure 2, Mertens et al. 2006).

STAT signaling can be negatively regulated in many ways such as direct deactivation through degradation (Kim and Maniatis 1996) or dephosphorylation by tyrosine phosphatases, e.g. in the cytosol by PTP1B (Aoki and Matsuda 2000) and in the nucleus by TC-PTP (ten Hoeve et al. 2002). Alternatively, STAT activity can be diminished via receptor or JAK dephosphorylation through cytoplasmic tyrosine phosphatases (Ali et al. 2003, reviewed in Wormald and Hilton 2004). 'Suppressors Of Cytokine Signaling' (SOCS) or 'Cytokine Inducible SRC homology 2-domain containing proteins' (CIS) suppress continued STAT activation by binding to the catalytic region of JAKs thereby preventing the phosphorylation of STATs (Yasukawa et al. 1999), by competing with STATs for binding to the receptor (Yamamoto et al. 2003) or by targeting JAKs for degradation (Ali et al. 2003). SOCS proteins themselves represent transcriptional targets of the JAK/STAT cascade and therefore provide a negative feedback mechanism. In the nucleus, the activation of gene expression may be further inhibited by 'Protein Inhibitors of Activated Stat' (PIAS) via sumoylation (Ungureanu et al. 2003), although PIAS proteins appear to be rather general repressors of transcription factors and regulators of chromatin structure (Hari et al. 2001, Long et al. 2003).

Experiments with STAT deficient organisms have shown the biological *in vivo* relevance for several processes, where the activities of STATs can be broadly classified as inhibiting and promoting proliferation. For example, STAT1 can be classified as inhibiting proliferation and promoting apoptosis. STAT1 knockout mice are more susceptible to viral infections and they lack an immune response (Durbin et al. 1996). STAT1 deficient mice also have high rates of spontaneous cancers due to the loss of 'immunosurveillance' by the immune system (Shankaran et al. 2001). On the other hand,

STAT3 has a strong proliferative effect. STAT3 is required for embryogenesis and important for cell growth (reviewed in Akira 2000). Also, STAT5a and STAT5b, which are more than 90% identical in their amino acid sequence, can suppress interferon induced and STAT1 mediated apoptosis (Jensen et al. 2005, Wellbrock et al. 2005) thereby exerting a stimulating effect on cell proliferation. The importance of the JAK/STAT pathway in proliferative processes is illustrated by the frequent identification of pathway overactivation in many human cancers leading to an anti-apoptotic state in the cancer cell (Bowman et al. 2000, Song and Grandis 2000). Given these important roles for the pathway, a comprehensive knowledge of the interactions of known pathway components with other proteins during signaling has the potential to provide targets for pharmacological inhibition of STAT activities and hence destruction of cancer cells (Darnell 2002). Unfortunately, a comprehensive search for modulators of mammalian *in vivo* pathway activity has been hampered by the complexity of the pathway in higher vertebrates, which consists of apparently redundant pathway components in multiple isoforms (Table 1, Table 2).

***Drosophila melanogaster* as a model organism to study JAK/STAT signaling**

As outlined above, the JAK/STAT pathway was originally discovered in mammalian cell culture systems. Subsequently, over the past years, similar biological roles for the pathway component homologs have been identified in organisms ranging from fruit flies to humans speaking for the evolutionary conservation of this important pathway (Pires-daSilva and Sommer 2003). Although systematic screens for modulators of vertebrate *in vivo* pathway activity have been hampered by the complexity of the pathway, they have been initiated in the low complexity system of *Drosophila melanogaster* (Bach et al. 2003, Mukherjee et al. 2006).

The fruit fly *Drosophila melanogaster*, hereafter referred to as *Drosophila*, is a well established model organism, which has been used for more than 100 years in biological research to study animal development and to discover novel cellular pathways by high-throughput approaches (reviewed in Beller and Oliver 2006). Due to the excellent genetic manipulation tools, many biological phenomena were first described in *Drosophila* and

could then also be characterized in higher organisms. The *Drosophila* genome was sequenced in the year 2000 (Adams et al. 2000). It comprises approximately 14,000 genes (release 4.2 of the Berkeley *Drosophila* Genome Project) distributed over only five chromosomes. More than 60% of the relevant genes in human diseases are also present as homologs in *Drosophila* (Chien et al. 2002, O'Kane 2003, Rubin et al. 2000) demonstrating the potential relevance of *Drosophila* research also for higher organisms.

Drosophila as an experimental system offers a variety of advantages compared to other higher model organisms such as the mouse. *Drosophila* is an undemanding small laboratory animal, which can be easily and inexpensively maintained in large numbers. The generation time of 10 d at 25°C is short, and fertilized flies can give rise to large numbers of offspring so that interesting phenotypes can be analyzed within a short period of time. After one day of embryogenesis, the first instar larva hatches, which one day later is followed by the second instar larval stage. After another day, the growing larva spends two to three days in the third instar larval stage, which finally ends with pupation. The metamorphosis is completed within 5 d, when the adult fly ecloses from the pupa (Weigmann et al. 2003). All of these developmental stages are large enough to observe the details of interesting phenotypes without big effort. Transgenic animals are easily generated (Spradling and Rubin 1982), and most processes during development are well characterized (Weigmann et al. 2003). Important genetic methods include the Gal4/UAS system (Brand and Perrimon 1993) for the time- and tissue-specific ectopic expression of genes during development, P-elements for induced mutagenesis (Spradling et al. 1999) as well as FLP/FRT-mediated mitotic recombination for the induction of time- and tissue-specific mutant clones (Xu and Rubin 1993). Since the early systematic screens for relevant genes in development and pattern formation of the larval cuticle in the 1980s by Christiane Nüsslein-Volhard and Eric Wieschaus (Nüsslein-Volhard et al. 1984), there are large collections of fly mutants available for many genes. Furthermore, there are established isolated cultured cells such as S2 (Schneider 1972), Kc₁₆₇ (Cherbas et al. 1977) and S2R⁺ cells (Yanagawa et al. 1998), which also facilitate the analysis of biological processes outside the whole organism. These methods, the known and conserved genome as well as the ease of handling therefore make *Drosophila* an ideal model organism to study conserved cellular pathways.

Developmental genetic screens in *Drosophila* have identified multiple JAK/STAT pathway components on the basis of their segmentation phenotype (Binari and Perrimon 1994, Harrison et al. 1998, Hou et al. 1996), and subsequent analysis of the pathway has characterized evolutionarily conserved roles during immune responses, hematopoiesis and cellular proliferation (Boutros et al. 2002, Lagueux et al. 2000, Meister and Lagueux 2003, Mukherjee et al. 2005). The JAK/STAT signaling cascade in *Drosophila* represents one of the most elementary complete pathways (Table 1). Whether this is really a simpler ancient pathway or rather the result of secondary simplification during evolution remains to be investigated (Raible et al. 2005). In contrast to vertebrates, where a multitude of proteins have been shown capable of stimulating the JAK/STAT pathway, only three *Drosophila* extracellular pathway ligands have been described so far (Table 1, Hombria et al. 2005), which are all located at the same chromosomal region. Mutational analysis indicates that these three ligands could account for all canonical JAK/STAT activity (Harrison et al. 1998, Agaisse et al. 2003, Hombria et al. 2005). The most extensively characterized ligand is Unpaired (Upd, Harrison et al. 1998). Although no clear homologs exist, Upd bears most similarity with vertebrate Leptin (Boulay et al. 2003). Upd is a secreted protein that is capable of stimulating pathway activity at a distance from its place of expression (Karsten et al. 2002, Zeidler et al. 1999) and has been visualized in the extracellular space (Zeidler et al. 1999). Upd is a glycoprotein which, in contrast to the diffusible ligand Upd2 (Hombria et al. 2005), is strongly associated with the extracellular matrix (ECM) in tissue culture, from which it can be removed by heparin treatment (Harrison et al. 1998). Upd can bind to a trans-membrane receptor termed Domeless (Dome, Brown et al. 2001) to activate the JAK/STAT pathway. Dome is the only currently identified invertebrate JAK/STAT pathway receptor (Table 1) and its extracellular domain shares most similarities with the vertebrate interleukin 6 receptor family (Table 2, Boulay et al. 2003, Brown et al. 2001). Prior to ligand binding, the receptor dimerizes, a process which has been illustrated *in vivo* and which does not appear to be ligand induced. The mechanisms by which this process is controlled are as yet unknown (Brown et al. 2003). Associated to the receptor Dome is a JAK called Hopscotch (Hop, Binari and Perrimon 1994), which phosphorylates a STAT transcription factor termed Stat92E (Hou et al. 1996, Yan et al. 1996b). Interestingly, this single STAT

appears to exert both the proliferative and antiproliferative functions that in vertebrates are distributed to different STATs (Mukherjee et al. 2005). Furthermore, an N-terminally truncated version of Stat92E derived from an alternative promoter site can negatively regulate Stat92E (Henriksen et al. 2002). Other known regulators of JAK/STAT signaling, including a family of *SOCS*- (Karsten et al. 2005) and *PIAS*-like genes (Betz et al. 2001), are also functionally conserved in *Drosophila* and were identified based on their homology to components originally characterized in mammalian cell culture studies (Castelli-Gair Hombria and Brown 2002).

A range of *Drosophila* developmental processes requires JAK/STAT pathway activity. During embryonic development, Upd is required for the initial process of sex selection (Sefton et al. 2000), before it activates pair-rule gene expression during embryonic segmentation (Beccari et al. 2002, Hou et al. 1996). The proliferation of primordial germ cells requires STAT activity stimulated by a receptor tyrosine kinase, and the active migration of these cells as well as the formation of the embryonic gonads later requires STAT activity again (Brown et al. 2006, Li et al. 2003b). STAT activity is further necessary for the proper masculinization of male gonads (Wawersik et al. 2005). The pathway is also required for the proper differentiation of the fore- and hind-gut (Josten et al. 2004, Lengyel and Iwaki 2002) and for the initial stages of tracheal morphogenesis and posterior spiracle development (Brown et al. 2001). During larval stages, the JAK/STAT pathway controls the proliferation of imaginal disc cells (Mukherjee et al. 2005) and the ommatidial rotation in the eye (Zeidler et al. 1999). In the adult fly, the pathway is once more required for the differentiation of veins in the wing (Yan et al. 1996a) and for the differentiation and proliferation of germ cells in both sexes (Baksa et al. 2002, Beccari et al. 2002, Ghiglione et al. 2002, Kiger et al. 2001, McGregor et al. 2002, Silver and Montell 2001, Tulina and Matunis 2001).

Amongst the roles of the *Drosophila* JAK/STAT pathway, those that have been conserved through the course of evolution are particularly interesting (reviewed in Arbouzova and Zeidler 2006). These conserved roles include the control of cellular proliferation (Bach et al. 2003, Mukherjee et al. 2005). Downstream targets of STAT that might mediate its proliferative effects include *c-myc*, cyclins, and *raf* (reviewed in Bromberg 2001, Kwon et al. 2000). The cyclin dependent kinase Cdk4, which is involved

in cell cycle control, has also been shown in *Drosophila* to activate Stat92E in a non-canonical manner (Chen et al. 2003). Other conserved roles of the JAK/STAT pathway include a function in hematopoietic development (Meister and Lagueux 2003) and for the innate immune system (Agaisse et al. 2003). Similarly to proliferative diseases in humans (James et al. 2005), a mutation in the *Drosophila* JAK homolog results in constitutive kinase activity causing the proliferation of hematopoietic cells and the formation of melanotic tumors (Harrison et al. 1995, Luo et al. 1997). Another conserved function of the JAK/STAT pathway appears to be the maintenance of stem cells. For example, the culturing of pluripotent murine embryonic stem cells to prevent the onset of differentiation requires the presence of leukemia inhibitory factor (LIF), a ligand of the JAK/STAT pathway activating STAT3 (Hao et al. 2006, Matsuda et al. 1999, Niwa et al. 1998). Similarly, the *in vivo* maintenance of germline stem cells in *Drosophila* adult male testes requires the activity of the JAK/STAT ligand Upd (Brawley and Matunis 2004, Kiger et al. 2001, Tulina and Matunis 2001), and a conceptually similar process may take place in female germline stem cells (Decotto and Spradling 2005).

Although successful in identifying the ‘core’ pathway members Upd, Dome, Hop and Stat92E, it is probable that forward genetic approaches have missed components, possibly due to non-saturating mutagenesis, genetic redundancy or phenotypic pleiotropy. However, several questions regarding the mechanism by which the pathway transduces information remain open. These include a comprehensive knowledge about the interactions and molecular links to other signaling pathways, the identity of co-activators and co-repressors of Stat92E, mechanisms to downregulate Stat92E (e.g. phosphatases, degradation), ligand processing, secretion and association with the ECM as well as ligand-receptor-coreceptor interactions. It is further unclear what the role of endocytosis of the activated ligand-receptor complexes into the receiving cell may be. For example, it was shown in mammalian systems that the endocytosis of some IFN receptors upon activation is required for STAT activity, whereas this may not be the case for others (Marchetti et al. 2006), and the internalization of an interleukin receptor may not even require JAK/STAT pathway activity (Thiel et al. 1998).

The identity and function of the gene products involved in these processes remains to be determined. The functional homology of the regulatory interactions throughout evolution

indicates that JAK/STAT signaling constitutes a general conserved mechanism, which can be explored using the genetically tractable model organism *Drosophila*. In particular, the identification of novel regulators of the pathway in *Drosophila* has the potential to provide homologous targets for pharmacological inhibition of pathway activity and hence destruction of cancer cells (Darnell 2002).

Dissection of cellular pathways by RNAi

The completion and availability of the *Drosophila* genome sequence in the year 2000 (Adams et al. 2000) has opened new avenues to discover genes in cellular pathways and to study the underlying genetic networks. Classically, unbiased genetic screens have been performed in *Drosophila* using random chemical or P-element induced mutations followed by detailed analysis of the screening candidates to elucidate the function by genetic interactions *in vivo* or *in vitro* by biochemistry. However, in modern approaches it is also possible to interfere with cellular pathways specifically at the level of gene activity (RNAi as outlined below) and at the level of the translated protein (e.g. by small molecules). Novel techniques are now often used to elucidate cellular networks including approaches such as metabolomics, proteomics, pharmacogenomics as well as techniques which can be summarized as functional genomics (reviewed in Beller and Oliver 2006). For example, the activity of genes following a certain environmental change can now be determined utilizing gene chips or microarrays, and the binding of transcription factors to specific DNA sites regulating the activity of genes can be analyzed by genome-wide ChIP-chip analysis (so-called tiling arrays (Biemar et al. 2006)). Until recently, systematic reverse genetic approaches to probe loss-of-function phenotypes have been difficult to conduct. However, the discovery of sequence-specific posttranscriptional silencing mechanisms by small interfering RNAs (siRNAs) has now allowed the development of tools to efficiently knockdown the expression of specific genes.

Interestingly, Britten and Davidson already proposed in 1969 that RNAs rather than proteins could specify which genes are turned on and off by simple Watson and Crick base pairing rules to DNA (Britten and Davidson 1969, reviewed in Zamore and Haley

2005). This theory of gene regulation by RNAs, although with a different degradative mechanism, was revived by pioneering plant studies in 1990 (Napoli et al. 1990, van der Krol et al. 1990) and later detailed studies in plants (e.g. Ingelbrecht et al. 1994) and the fungus *Neurospora crassa* (e.g. Cogoni et al. 1996), which led to the discovery of a process called ‘posttranscriptional gene silencing’. The idea of posttranscriptional gene silencing was then generalized by the dawn of the new millennium, when similar RNA-mediated mechanisms were discovered in organisms as diverse as nematodes (Fire et al. 1998), trypanosomes (Ngo et al. 1998), planaria (Sanchez Alvarado and Newmark 1999) and flies (Kennerdell and Carthew 1998). In this approach, long dsRNAs are processed inside the cell by the ribonuclease (RNase) III enzyme Dicer into siRNAs 21 to 23 bp in length, which subsequently direct the cleavage of homologous mRNAs via an RNA-induced silencing complex (RISC, Figure 4, Hannon 2002). In the *Drosophila* system, both embryos and cultured cells respond to dsRNA by strongly and specifically down-regulating the expression of targeted genes (Clemens et al. 2000). For use in cultured cells, 500-700 bp of dsRNA are synthesized from DNA templates containing terminal T7 promoters, the resulting molecules are added to the culture medium and are taken up by the cells in the absence of serum through scavenger receptor-mediated endocytosis (Ulvila et al. 2006, Saleh et al. 2006). Inside the cell, the dsRNAs are processed into shorter fragments of siRNA by Dicer-2 and R2D2 (Figure 4). The resulting duplicates are subsequently assembled as single strands into RISC, which eventually knocks down the corresponding gene products (Lee et al. 2004). Central for this process are the Argonaute family proteins (also known as ‘Slicer’), which are part of the RISC complex and which catalyze the cleavage of the target mRNA (Tolia and Joshua-Tor 2007). In *Drosophila*, the pathway of posttranscriptional gene silencing is diversified in that two different Dicer enzymes exist, which regulate the cleavage of dsRNAs or of endogenous regulators of gene activity - the micro RNAs (miRNAs). Additionally, endogenous repeat associated siRNA (rasiRNA) can be produced, and different small RNAs may lead to specific chromatin modification, RNA cleavage or translational repression.

Using RNA interference as a tool to study gene function in mammals, however, has been hampered by the antiviral interferon response induced by the use of long dsRNAs, which if longer than 30 bp can activate both protein kinase PKR (Manche et al. 1992) and 2',5'-

oligoadenylate synthetase (Minks et al. 1979; Reynolds et al. 2006 reported a potential induction of a cell-type specific IFN response even for dsRNAs between 23 bp and 30 bp in length). This leads to the unspecific inhibition of translation via phosphorylation of the initiation factor eIF-2 by activated PKR as well as the unspecific degradation of mRNA by 2',5'-oligoadenylate-activated ribonuclease L, respectively. Only by the use of synthetic small RNAs 21 bp in length has it been possible to circumvent the antiviral interferon response and to make use of this technique also in mammalian cells (Elbashir et al. 2001; the differences of RNAi in *Drosophila*, mammals and *C. elegans* are summarized in Figure 4).

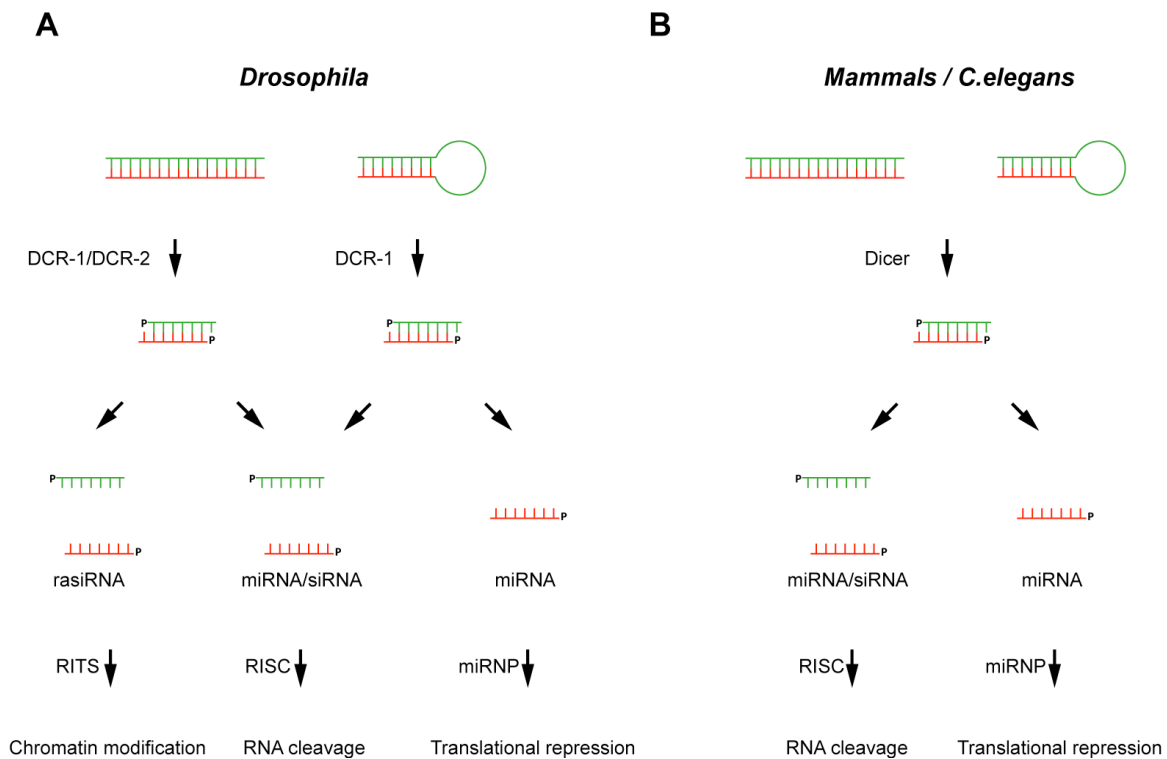


Figure 4. Mechanism of gene silencing by RNA interference. RNAi pathways are shown for *Drosophila* (A) and mammals / *C. elegans* (B). dsRNAs are shown as ladder-like structures including two strands in green and red. Precursors of miRNAs, pre-miRNAs, are depicted as a ladder with a loop attached. The RNase-III-like enzyme Dicer (DCR) cuts long dsRNA and miRNA precursors into siRNA or miRNA duplexes, which are then unwound and assembled into effector complexes. The effector complex RISC (RNA induced silencing complex) mediates target mRNA cleavage and subsequently degradation whereas the RITS (RNA induced transcriptional silencing) complex leads to heterochromatin condensation, and miRNPs (miRNA-protein complex) guide the translational repression of target mRNAs. P indicates 5' phosphate. Figure modified after Meister and Tuschl 2004.

RNAi has now been utilized to study individual gene functions in vertebrate and invertebrate cells (Elbashir et al. 2002), and several genome-wide collections of dsRNAs have already been applied to systematically screen for specific phenotypes in *Caenorhabditis elegans* (Fraser et al. 2000, Gonczy et al. 2000, Kamath et al. 2003). Recently, the first genome-wide RNAi library applicable to study a variety of biological processes in cultured *Drosophila* cells was synthesized (Boutros et al. 2004) using PCR fragments containing T7 promoter sequences on each end (Hild et al. 2003) as templates to generate approximately 20,000 dsRNAs by *in vitro* transcription. These are designed to target the activities of all known and predicted genes within the *Drosophila* genome. In a first screen, a systematic search for genes required for cell growth and cell survival was undertaken. This study demonstrated that RNAi screens using cultured *Drosophila* cells in high-density microtiter plates are feasible and can be rapidly and efficiently undertaken using semi-automated screening technologies.

In the present study, a systematic genome-wide screen for genes required for JAK/STAT signaling was performed. By analyzing 20,026 RNA interference (RNAi)-induced phenotypes in cultured *Drosophila* hemocyte-like cells with customized software tools, interacting genes were identified encoding four known and 86 previously uncharacterized proteins. To take these results further and to ‘place’ the novel components in the pathway, cell-based epistasis experiments were performed that enabled their classification based on their interaction with known components of the signaling cascade. Furthermore, the function of these novel candidates was analyzed with evolutionary aspects, and putative human homologs of the candidates found in *Drosophila* were targeted in a human cell culture system to assess the activity of human STATs upon knockdown of candidate modulators. Interestingly, 30 of the human homologs display a similar JAK/STAT phenotype to their *Drosophila* counterparts. In addition to multiple human disease gene homologs, the protein tyrosine phosphatase Ptp61F and the *Drosophila* homolog of BRWD3, a bromo-domain-containing protein disrupted in leukemia, were found in the RNAi screen and characterized further. *In vivo* analysis demonstrates that disrupted *dBWD3* and overexpressed *Ptp61F* function as suppressors of leukemia-like blood cell tumors. Moreover, *dBWD3* is localized in the nucleus, can physically interact with Stat92E and very likely induces target gene activity together with

Stat92E. In summary, the present study represents a comprehensive identification of novel evolutionarily and functionally conserved loci required for JAK/STAT signaling and provides molecular insights into an important pathway relevant for human cancer. Human homologs of identified pathway modifiers may constitute targets for therapeutic interventions.

MATERIALS AND METHODS

Analysis and manipulation of nucleic acids

Quantification of nucleic acid concentrations

For the determination of nucleic acid concentrations, a NanoDrop spectrophotometer (NanoDrop Technologies) was used. Complementary, nucleic acids were loaded onto TBE agarose gels (approximately 1% agarose) containing 0.5 $\mu\text{g/ml}$ ethidium bromide along with markers of known size and concentration (Gene Ruler, Fermentas) and subjected to electrophoresis according to standard protocols (Ausubel et al. 1999).

Polymerase chain reaction (PCR)

PCR to amplify or mutate DNA fragments was carried out according to standard protocols (Ausubel et al. 1999). Annealing temperatures and extension times were adjusted to fit the respective primer melting temperature and the length of the expected PCR amplicon. Most of the time, 35 cycles were used for the amplification of DNA except for site-directed mutagenic PCR using the QuikChange Site-Directed Mutagenesis Kit according to the manufacturer's instructions (Stratagene). Long PCR Enzyme Mix (Fermentas) was used for the amplification of long DNA pieces (> 500 bp), whereas conventional Taq polymerase (Roche) was used for all other purposes including PCR reactions for dsRNA production. Site-directed mutagenic PCR was carried out with Pfu polymerase (Stratagene).

Primer design

Primers were designed with the programs E-RNAi (Arziman et al. 2005), Primer3 (Rozen and Skaletsky 2000) or in Vector NTI 7 for Mac OS X (Invitrogen). Primers used in this study are shown in Table 3. Further information for primers used to generate the genome-wide RNAi library is available at <http://rnaï.dkfz.de>

Preparation of plasmid DNA

Plasmid DNA was purified using the QIAGEN Mini or Maxi Kits (QIAGEN) following the manufacturer's protocol.

Table 3. List of primers.

Name	Kind	Target	Tag	Sequence (5' -> 3')
Primers for dsRNA generation with T7 tag				
T7	F/R	T7 tag	T7	TAATACGACTCACTATAGGG
5-T7-DIAP	F	<i>diap</i>	T7	GAATTAATACGACTCACTATAGGGAGAGCCAAAGTG
3-T7-DIAP	R	<i>diap</i>	T7	GAATTAATACGACTCACTATAGGGAGATCCTCCAGC
T7DOME-F	F	<i>dome</i>	T7	GAATTAATACGACTCACTATAGGGAGACCAGCTGCC
DOMET7-R	R	<i>dome</i>	T7	GAATTAATACGACTCACTATAGGGAGACTGGACCCA
5T7gfp	F	<i>egfp</i>	T7	TAATACGACTCACTATAGGACGGCCGCCATTAACAAGCAAAAG
3T7gfp	R	<i>egfp</i>	T7	TAATACGACTCACTATAGGCTGGGCGGAGCGGATGATG
T7HOP-F	F	<i>hop</i>	T7	GAATTAATACGACTCACTATAGGGAGACTATTGCTT
T7HOP-R	R	<i>hop</i>	T7	GAATTAATACGACTCACTATAGGGAGATTGAGTGTG
5T7lacZ	F	<i>lacZ</i>	T7	GAATTAATACGACTCACTATAGGGAGACAGTGGCGTCTGGCGGAAAA
3T7lacZ	R	<i>lacZ</i>	T7	GAATTAATACGACTCACTATAGGGAGATCCGAGCCAGTTTACCCGCT
5'-T7pias	F	<i>pias</i>	T7	GAATTAATACGACTCACTATAGGGAGAACGCGACGCTTAATCAAAAGA
3'-T7pias	R	<i>pias</i>	T7	GAATTAATACGACTCACTATAGGGAGACTGTTTGACGTTGATGTGGG
relish_1	F	<i>rel</i>	T7	GAATTAATACGACTCACTATAGGGAGAAGCAGTGGCGCACTAA
relish_2	R	<i>rel</i>	T7	GAATTAATACGACTCACTATAGGGAGATGCTATAGCCACTGGT
T7RH5-F	F	<i>rh5</i>	T7	GAATTAATACGACTCACTATAGGGAGAAGGGCTTCC
RH5T7-R	R	<i>rh5</i>	T7	GAATTAATACGACTCACTATAGGGAGAGCTGACACT
STAT92E_r	R	<i>stat92E</i>	T7	TAATACGACTCACTATAGGGAGCAGCTGAGAACC
STAT92E_f	F	<i>stat92E</i>	T7	TAATACGACTCACTATAGGGAGAAGCTGCTTGCCCA
toll_1	F	<i>toll</i>	T7	GAATTAATACGACTCACTATAGGGAGATTTTTGATTTTCCAG
toll_2	R	<i>toll</i>	T7	GAATTAATACGACTCACTATAGGGAGAGCTTTTCTTAAGCTGC
T7upd-F	F	<i>upd</i>	T7	GAATTAATACGACTCACTATAGGGAGAGGTACC
T7upd-R	R	<i>upd</i>	T7	GAATTAATACGACTCACTATAGGGAGACTGCTTCTT
Primers for site-directed mutagenesis				
STATperfl-5'	F	Stat92E site	none	CCCGGGGATCCTAAATTTCCCGGAAAAGTAATAAAA
STATperfl-3'	R	Stat92E site	none	TTTTATTACTTTCCGGGAAATTTAGGATCCCCGGG
STATperf2-5'	F	Stat92E site	none	AGTAATAAAATTTCCCGGAAAAGTAAAGATCCCCGGT
STATperf2-3'	R	Stat92E site	none	AAACGGGGATCTTACTTTCCCGGAAAATTTACT
Primers for sequencing				
pAc5-S1	F	<i>actin 5c</i>	none	CCGTTTGAGTCTTGTGCTG
FFLuc-S1	R	<i>firefly luciferase</i>	none	CACTGCATACGACGATTCTGT
P-primer	F/R	<i>P-element</i>	none	CGACGGGACCCACTTATGTTATTTTCATCATG
pMT-5'	F	pMT vector	none	AATCATCTCAGTGCAACTAA
pMT-3'	R	pMT vector	none	TTAGGGATAGGCTTACCCTC
RenS1	R	<i>Renilla luciferase</i>	none	CTCAATATCAGGCCATTCATCCC
Primers for cloning				
NotI-dBrodl	F	<i>brwd3</i>	NotI	GCTTGCGGCCGCCACCATGGAACTAGACAACCCA
AscI-dBrodlOpen	R	<i>brwd3</i>	AscI	GCAAGGCGCGCCCTCCACTCCTTGAAGATACCGCG
NotI-dBrodl-bromo	F	<i>brwd3</i>	NotI	GCTTGCGGCCGCCACCATGAAGTCCGAGGAATGGC
AscI-dBrodl-bromoOpe	R	<i>brwd3</i>	AscI	GCAAGGCGCGCCCTTAAATGGCGCCTTCCAGC
AscI-dBrodl-WD40Open	R	<i>brwd3</i>	AscI	GCAAGGCGCGCCCTGTACTTTTCAATGCACACGCC
dPias-F	F	<i>pias</i>	BamHI	CATCGGATCCTGCAAAAAGGGGTCCAACGTACCGGAT
dPias-R	R	<i>pias</i>	Asp718	GGGTACC AAAAATGGTGATATGCTTCCGA
2xDrafsacII-5'	F	Stat92E site	SacII	AAAAAACC GCGGTGAGCTAACATAACCC
2xDrafsacII-3'	R	Stat92E site	SacI	GCGTAAGAGCTCCACCGCGGTGGC
Not2xSTAT-F	F	Stat92E site	NotI	TTGCGGCCGCTAAATTTCCCGGAAAAGTAAT
Not2xSTAT-R	R	Stat92E site	NotI	TGGCGGCCGCTCTAGAGTCGAGATC
SocsLuc-F	F	Stat92E site	Asp718	GTTAGGTACCGGGTCGCAGTATCGTTGGCG
SocsLuc-R	R	Stat92E site	BamHI	CGAAGGATCCCTGTCACCTCTCAGAAAATCGGTC

Kind describes directionality of primers to amplify the target. 'F' is the forward, 'R' the reverse primer.

The tags on the primers as derivatives of a T7 RNA polymerase recognition site are designated 'T7'. All other tags describe the attached restriction enzyme recognition site.

Isolation of genomic DNA from single flies

Single adult flies were homogenized on ice in 50 μ l squishing buffer (10 mM Tris-HCl pH 8.2, 1mM EDTA, 25 mM NaCl, 200 μ g/ml proteinase K). The mixture was incubated at 37°C for 30 min and proteinase K activity quenched by incubation at 95°C for 2 min. The supernatant was then directly used for PCR.

Large-scale preparation of Drosophila genomic DNA

Approximately 40 adult flies were subjected to the DNeasy Kit DNA preparation method (QIAGEN) following the protocol 'Isolation of total DNA from animal tissue'.

Sequencing

DNA sequencing was performed by Gordon Dowe at the MPI for Biophysical Chemistry on a ABI Prism 377 DNA sequencing machine (Applied Bioystems) or by the company MWG. DNA and primers were added following their instructions for preparing the sequencing reaction.

Restriction digest of DNA

All restriction endonucleases indicated below were obtained from NEB and used as suggested by the manufacturer for complete digestion, except for cases in which a partial digest was desired. In these cases, time courses of 0 min to 1 h were performed to determine the optimal time-point for the partial digest.

DNA extraction from agarose gels

Ethidium bromide stained DNA was excised from the agarose gel using a clean scalpel and extracted using the QIAquick Gel Extraction Kit (QIAGEN).

DNA ligation

Ligation using T4 DNA ligase (Fermentas) was performed with a total of 100 ng of DNA and a molar ratio of insert to host vector of 3:1 with 1 μ l T4 ligase in 10 μ l reaction volume overnight at 18°C, and otherwise according to the manufacturer's instructions.

Table 4. List of plasmids.

Name	Backbone	Promoter used	Insert	Reference / supplier
Basic published and unpublished vectors used for modification				
<i>pAc5.1A</i>	-	<i>actin 5c</i>	-	Invitrogen
<i>pAct-Gal4</i>	-	<i>actin 5c</i>	<i>Gal4</i>	M. P. Zeidler
<i>pAc5.1-Sid-1</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>Sid-1</i>	Feinberg and Hunter 2003
<i>pBS(KS+)</i>	-	-	-	Stratagene
<i>pBS-EGFP-B</i>	<i>pBS(KS+)</i>	-	<i>egfp</i>	M.P. Zeidler
<i>pCoBlast</i>	-	<i>copia</i>	blasticidin-resistance gene	Invitrogen
<i>pENTR-D</i>	-	-	-	A. Herzig
<i>pGL3</i>	-	-	<i>firefly luciferase</i>	Promega
<i>pMT A</i>	-	<i>metallothionein</i>	-	Invitrogen
<i>pBS-LD09022</i>	<i>pBS(SK-)</i>	-	<i>pias</i>	DGRC
<i>pOT2-LD40380</i>	<i>pOT2</i>	-	<i>brwd3</i>	DGRC
<i>pOT2-LP01280</i>	<i>pOT2</i>	-	<i>ptp61Fc</i>	DGRC
<i>pFlc-RE01370</i>	<i>pFlc-1</i>	-	<i>ptp61Fa</i>	DGRC
<i>pP_{Ac5c}-PL</i>	-	<i>actin 5c</i>	-	D. Curtis
<i>pRLSV40</i>	-	-	<i>Renilla luciferase</i>	Invitrogen
<i>pUAS-hop^{TumI}</i>	<i>pUAST</i>	<i>UAS</i>	<i>hop^{TumI}</i>	Harrison et al. 1995
<i>pUAS-STAT-GFP</i>	<i>pUAST</i>	<i>UAS</i>	<i>stat92E-egfp</i>	M.P. Zeidler
<i>pUAST</i>	-	<i>UAS</i>	-	Brand and Perrimon 1993
<i>pUbiP-rfa-3Flag</i>	<i>pBS(KS+)</i>	<i>ubiquitin</i>	-	A. Herzig
<i>pUbiP-rfa-EGFP</i>	<i>pBS(KS+)</i>	<i>ubiquitin</i>	-	A. Herzig
Vectors used for functional studies				
<i>p2xDrafSTAT(wt)</i>	TATA-PGVB	<i>draf</i>	<i>firefly luciferase</i>	Kwon et al. 2000
<i>p2xDrafSTAT(mut)</i>	TATA-PGVB	<i>draf</i> mutated	<i>firefly luciferase</i>	Kwon et al. 2000
<i>p2xDrafSTATperf</i>	<i>p2xDrafSTAT(wt)</i>	<i>draf</i> with Stat92E consensus binding site	<i>firefly luciferase</i>	this study
<i>p2x2xDrafLuc</i>	<i>p2xDrafSTAT(wt)</i>	multimerized <i>draf</i>	<i>firefly luciferase</i>	this study
<i>p2x2xDrafLuc(mut)</i>	<i>p2xDrafSTAT(mut)</i>	<i>draf</i> mutated	<i>firefly luciferase</i>	this study
<i>p3x2xDrafLuc</i>	<i>p2x2xDrafSTAT(wt)</i>	multimerized <i>draf</i>	<i>firefly luciferase</i>	this study
<i>p6x2xDrafLuc</i>	<i>p3x2xDrafSTAT(wt)</i>	multimerized <i>draf</i>	<i>firefly luciferase</i>	this study
<i>p4xSocsLuc</i>	<i>pGL3</i>	multimerized <i>socs36E</i>	<i>firefly luciferase</i>	this study
<i>pAct-RL</i>	<i>pAct</i>	<i>actin 5c</i>	<i>Renilla luciferase</i>	this study
<i>pAct-UpdGFP</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>upd-egfp</i>	this study
<i>pAc5.1-hop^{TumI}</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>hop^{TumI}</i>	this study
<i>pAc5.1-ptp61Fa</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>ptp61Fa</i>	this study
<i>pAc5.1-ptp61Fc</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>ptp61Fc</i>	this study
<i>pAct-STAT-GFP</i>	<i>pAct</i>	<i>actin 5c</i>	<i>stat92E-egfp</i>	this study
<i>pAc5.1-dBrodl</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>brwd3</i>	this study
<i>pAc5.1-dBrodlΔC</i>	<i>pAc5.1</i>	<i>actin 5c</i>	<i>brwd3ΔC</i>	this study
<i>pMT-UpdGFP</i>	<i>pMT A</i>	<i>metallothionein</i>	<i>upd-egfp</i>	this study
<i>pUAS-UpdGFP</i>	<i>pUAS</i>	<i>UAS</i>	<i>upd-egfp</i>	M.P. Zeidler
<i>pUAS-ptp61Fa</i>	<i>pUAST</i>	<i>UAS</i>	<i>ptp61Fa</i>	this study
<i>pUAS-ptp61Fc</i>	<i>pUAST</i>	<i>UAS</i>	<i>ptp61Fc</i>	this study
<i>pUAS-dPias-GFP</i>	<i>pUAST</i>	<i>UAS</i>	<i>pias</i>	this study
<i>pUAST-dBRWD3</i>	<i>pUAST</i>	<i>UAS</i>	<i>brwd3</i>	this study
<i>pBRWD3(full)-Flag</i>	<i>pUbi-Flag</i>	<i>ubiquitin</i>	<i>brwd3</i>	this study
<i>pBRWD3(WD40)-Flag</i>	<i>pUbi-Flag</i>	<i>ubiquitin</i>	<i>brwd3(WD40)</i>	this study
<i>pBRWD3(Bromo)-Flag</i>	<i>pUbi-Flag</i>	<i>ubiquitin</i>	<i>brwd3(Bromo)</i>	this study
<i>pBRWD3(ΔCterm)-Flag</i>	<i>pUbi-Flag</i>	<i>ubiquitin</i>	<i>brwd3(Bromo)</i>	this study
<i>pBRWD3(Bromo + Cterm)-Flag</i>	<i>pUbi-Flag</i>	<i>ubiquitin</i>	<i>brwd3(Bromo)</i>	this study
<i>pStat92E-10xMyc</i>	<i>pUbi-10xMyc</i>	<i>ubiquitin</i>	<i>stat92E</i>	A.Herzig / P. Karsten

DGRC is the *Drosophila* Genomics Resource Center. For the basic vectors, only the relevant information regarding backbone, promoters and inserts is given.

Transformation of bacterial cells

Escherichia coli DH5 cells were obtained from Invitrogen. Chemically competent cells were made as described in Inoue et al. 1990. For transformation, chemically competent cells were thawed on ice and approximately 30 ng DNA were added followed by an incubation time of 20 min. Cells were then heat shocked in a water bath at 42°C for 45 s. After 3 min on ice, Luria-Bertani (LB) medium was added, and the cells were allowed to recover for 40 min on a shaker. The cells were then plated on LB agar plates containing appropriate antibiotics according to standard protocols (Ausubel et al. 1999) and incubated overnight at 37°C.

Generation of DNA vectors

The primers used to generate DNA vectors are shown in Table 3. All constructs used in this study are shown in Table 4.

p6x2xDrafLuc: The JAK/STAT reporter *p6x2xDrafLuc* was constructed by multimerization of Stat92E binding sites. Specifically, a 165 bp blunted BamHI/XbaI fragment from the original *p2xDrafSTAT(wt)* (Kwon et al. 2000) was inserted into the SmaI cut *p2xDrafSTAT(wt)* to yield the *p2x2xDrafLuc* vector. The same fragment was amplified by PCR with NotI sites on both ends (primers Not2xSTAT-F and Not2xSTAT-R) and inserted into compatible sites to yield the *p3x2xDrafLuc* reporter containing six Stat92E binding sites. These fragments were amplified again using primers 2xDrafSacII-5' and 2xDrafSacI-3', and the resulting 540 bp fragment was inserted into the SacII cut *p3x2xDrafLuc* vector to generate the *p6x2xDrafLuc* reporter with an enhancer of approximately 1,000 bp containing a total of 12 Stat92E binding sites.

p2x2xDrafLuc(mut): A 165 bp blunted BamHI/XbaI fragment from the original *p2xDrafSTAT(mut)* vector (Kwon et al. 2000) was inserted into the SmaI cut *p2xDrafSTAT(mut)* vector to yield the *p2x2xDrafLuc(mut)* vector.

p4xsocsLuc: Another JAK/STAT pathway reporter, *p4xsocsLuc*, was generated by amplifying a 745 bp product from genomic DNA using the primers SocsLuc-F and SocsLuc-R. This was then cut with EcoRI/BamHI to give a 285 bp fragment, subcloned into *pBS(KS+)* and re-excised with Asp718/BamHI. This 340 bp fragment, containing four predicted Stat92E binding sites (Karsten et al. 2002) was cloned into the Asp718/BglII sites of the *pGL3* vector.

pAct-RL: The *pAct-RL* vector expressing *Renilla* luciferase from a constitutive reporter was generated by cloning a 974 bp fragment coding for *Renilla* luciferase from *pRLSV40* into the BamHI/XbaI cut *pP_{Ac5c}-PL* vector (a kind gift from Dan Curtis).

pAct-UpdGFP: To generate the *pAct-UpdGFP* vector, a cDNA coding for *upd* (Harrison et al. 1995) fused in frame to *EGFP* via a BamHI site from *pUAS-UpdGFP* was inserted into the BamHI/XbaI cut *pP_{Ac5c}-PL* vector.

pAc5.1-hop^{Tuml}: A vector expressing the dominant gain-of-function allele *hop^{Tuml}* was cloned by inserting the open reading frame obtained from *pUAS-hop^{Tuml}* into the NotI/XbaI cut *pAc5.1A* vector.

pAc5.1-ptp61F: To generate *ptp61F* expression constructs, cDNAs encoding *ptp61Fc* (LP01280) and *ptp61Fa* (RE01370) were obtained from the *Drosophila* Genomics Resource Center (DGRC, University of Indiana). cDNA clones were analyzed by restriction analysis and end sequencing to confirm their integrity before subcloning into *pAc5.1A* and *pUAST* (Brand and Perrimon 1993). For *ptp61Fc*, the coding region of LP01280 was excised as an EcoRI/XhoI (partial digest) fragment of 1.8 kb and cloned into *pUAST*. Subsequently, the insert was re-excised with EcoRI/XbaI and cloned into *pAc5.1A* (Invitrogen). For *ptp61Fa*, the coding region of RE01370 was cut out with EcoRI/Asp718(filled) and cloned into *pAc5.1A* cut EcoRI/XbaI(filled). To generate a *pUAST* construct, an EcoRI/Asp718 fragment was used.

pAct-STAT-GFP: The *stat92E-egfp* fusion was cut out of *pUAS-STAT-GFP* with Asp718(filled) and XbaI and inserted into BamHI(filled) and XbaI cut *pAct-RL* vector.

pAc5.1-dBrodl: DGRC clone LD40380 was partially digested with EcoRI and XhoI, and the resulting 6.6 kb fragment was inserted into EcoRI/XhoI digested *pAc5.1A*.

pAc5.1-dBrodlΔC: DGRC clone LD40380 was digested to completion with EcoRI and XhoI, and the resulting 5.4 kb fragment was inserted into EcoRI/XhoI digested *pAc5.1A*.

pMT-UpdGFP: The *upd-egfp* fusion was cut out of *pUAS-UpdGFP* with EcoRI and XbaI and inserted into the EcoRI/XbaI cut *pMT A* vector.

pUAS-dPIAS-GFP: To clone *pUAS-dPIAS-GFP*, the EST clone LD09022 was used as a template in conjunction with the oligos dPIAS-F and dPIAS-R to amplify a region coding for 522 amino acids. The resulting product was sequenced, cut with Asp718/BamHI and subcloned into *pBS-EGFP-B* to generate DNA coding for an in-frame C-terminal EGFP fusion protein. This gene was then subcloned as an Asp718/XbaI fragment into *pUAST*.

pUAST-dBRWD3: DGRC clone LD40380 was partially digested with BglIII and XhoI and inserted into BglIII/XhoI cut *pUAST*.

pBRWD3(full)-Flag: Full-length *dBWRD3* was amplified by PCR from DGRC clone LD40380 with primers NotI-dBrodl and AscI-dBrodlopen, cut with NotI/AscI and inserted into the NotI/AscI cut *pENTR-D* vector. The insert was then *in vitro* recombined into the destination vector *pUbiP-rfa-3Flag* according to the Gateway recombination protocol (Invitrogen).

pBRWD3(WD40)-Flag: A truncation covering the WD40 domain of *dBWRD3* was amplified by PCR from DGRC clone LD40380 with primers NotI-dBrodl and AscI-dBrodlWD40open, cut with NotI/AscI and inserted into the NotI/AscI cut *pENTR-D* vector. The insert was then *in vitro* recombined into the destination vector *pUbiP-rfa-3Flag* according to the Gateway recombination protocol (Invitrogen).

pBRWD3(Bromo)-Flag: A truncation covering the bromo-domains of *dBWRD3* was amplified by PCR from DGRC clone LD40380 with primers NotI-dBrodlbromo and AscI-dBrodlbromoopen, cut with NotI/AscI and inserted into the NotI/AscI cut *pENTR-D* vector. The insert was then *in vitro* recombined into the destination vector *pUbiP-rfa-3Flag* according to the Gateway recombination protocol (Invitrogen).

pBRWD3(Δ Cterm)-Flag: A truncation removing the C-terminus of *dBWRD3* was amplified by PCR from DGRC clone LD40380 with primers NotI-dBrodl and AscI-dBrodlbromoopen, cut with NotI/AscI and inserted into the NotI/AscI cut *pENTR-D* vector. The insert was then *in vitro* recombined into the destination vector *pUbiP-rfa-3Flag* according to the Gateway recombination protocol (Invitrogen).

pBRWD3(Bromo+Cterm)-Flag: A truncation covering the bromo-domains and the C-terminus of *dBWRD3* was amplified by PCR from DGRC clone LD40380 with primers NotI-dBrodlbromo and AscI-dBrodlopen, cut with NotI/AscI and inserted into the NotI/AscI cut *pENTR-D* vector. The insert was then *in vitro* recombined into the destination vector *pUbiP-rfa-3Flag* according to the Gateway recombination protocol (Invitrogen).

Generation of dsRNAs and siRNAs

For *Drosophila* experiments, PCR fragments containing T7 promoter sequences on each end (Hild et al. 2003) or as indicated in Table 3 were used as templates to generate dsRNAs by *in vitro* transcription (Boutros et al. 2004, Clemens et al. 2000). After DNase I treatment, dsRNAs were purified by ethanol precipitation (Ausubel et al. 1999) or using the RNeasy Kit (QIAGEN) following the manufacturer's instructions. dsRNAs were then individually quality controlled by gel electrophoresis as described above and diluted to a working stock concentration of approximately 100 ng/ μ l.

In case of the genome-wide library (prepared by David Kutenkeuler and Viola Gesellchen at the DKFZ, Heidelberg), dsRNAs were aliquoted in ready-to-screen 384-well tissue culture plates (Greiner). Computational mapping predicts that the 20,026 dsRNA fragments used target > 91% of all predicted genes in the *Drosophila* genome (Annotation 4.0, Misra et al. 2002). Complete primer and amplicon sequence information for double-stranded RNAs including calculation of predicted efficiency and off-target effects for the RNAi library is available at <http://rnaï.dkfz.de>.

For experiments in human cells, pooled siRNAs targeting human genes were ordered from Dharmacon. Catalog numbers for SMART pools (Dharmacon) are provided in Supplementary Table 4, and sequence information for individual siRNAs comprising the pools is available in Supplementary Table 6.

QuantiGene assays

Gene expression levels in human cells were quantitatively measured using a branched DNA assay (QuantiGene, Panomics) according to the manufacturer's instructions. Briefly, HeLa cells were grown to confluency in 96-well plates and lysed with 100 μ l per well of proprietary lysis buffer. For determination of all mRNA levels, 70 μ l of lysate were used except for the quantification of β -actin levels, where only 10 μ l were used. Mixed probe sets were added according to the manufacturer's instructions and hybridized in sealed capture plates along with the appropriate lysates at 56°C overnight. On the next day, plates were washed four times with proprietary washing buffer, incubated with 100 μ l per well label extender for 1 h at 56°C in sealed plates, washed three times, incubated with 100 μ l per well amplifier for 1 h at 56°C in sealed plates, washed three times and finally incubated with 100 μ l per well substrate for 30 min in sealed plates. Plates were allowed to cool down at room temperature for 10 min (compared to 5 and 15 min, values from these measurements were most stable within a 1 min measurement range), until plate seals were removed and luminescence detected for 0.2 s per well on a luminometer (Wallac Victor Light 1420 Luminescence Counter, PerkinElmer).

Analysis and manipulation of proteins

Quantification of protein concentrations

For protein quantification, the Bio-Rad Protein Assay (BioRad) was used essentially following the manufacturer's instructions. Briefly, scaled down to 96-well assay plates, 2 μ l of protein solution were mixed with 160 μ l of water and 40 μ l dye reagent and vortexed. Absorbance was measured after 5 min of incubation on a BioRad plate reader using a 595 nm filter setting. BSA

dilutions of known concentration were used for the generation of standard curves to deduce the concentration of the protein solutions.

Polyacrylamide gel electrophoresis (PAGE)

The method developed by Laemmli 1970 was used according to standard protocols (Ausubel et al. 1999), and most of the time 4% gels were used as stacking and 10% gels as separating gels. Where appropriate, gels were stained with colloidal PageBlue (Fermentas). Samples were prepared in 3x SDS sample buffer (187.5 mM Tris pH 6.8, 6% SDS, 30% glycerol, 150 mM DTT, 0.03% bromophenol blue), cooked for 5 min and loaded onto gels along with protein standard markers (Prestained Protein Molecular Weight Marker, Fermentas). For mass spectrometry, one-dimensional SDS-PAGE was performed using a precast NuPAGE Bis-Tris gradient polyacrylamide gel (4-12%, Invitrogen) run in MOPS buffer.

Immunoblotting

Proteins were transferred to nitrocellulose membranes (Schleicher & Schuell) in a sandwich setting (BioRad) for 1 h at 350 mA and otherwise according to standard protocols (Ausubel et al. 1999). Membranes were blocked with 5% BSA in TBST (50 mM Tris pH 7.4, 150 mM NaCl, 0.1% Tween-20), incubated with primary antibody at 4°C overnight, washed with TBST, incubated with secondary antibody at room temperature for 2 h, washed and then subjected to ECL detection.

Primary antibodies against human β -actin, STAT1, P-STAT1, STAT3, P-STAT3, STAT5 and P-STAT5 were supplied by Cell Signalling Technologies (Beverly, MA) and used at 1:1000 dilutions for Western Blots. α -Myc antibody (mAb 9e10, Evan et al. 1985) was used at a concentration of 1:10, and α -Flag antibody (Sigma) was used 1:1000 for immunoprecipitations and 1:1000 for Immunoblotting. α -GFP antibody (abcam) was used at a dilution of 1:1000.

The activity of HRP-conjugated secondary antibodies (Jackson labs) was determined using ECL Western Blotting Substrate (Pierce).

Immunoprecipitation experiments

Cells were grown to confluency in 10 cm dishes. The medium was removed, and cells were scraped off with 1 ml of Rinse Buffer (40 mM Tris pH 7.5, 1 mM EGTA, 150 mM NaCl). After centrifugation for 2 min at 2,000 rpm in a tabletop centrifuge, cells were once more washed with 1 ml Rinse Buffer and finally resuspended in 400 μ l RIPA buffer (50 mM Tris pH 7.5, 50 mM NaCl, 1mM DDT, 0.1 mM PMSF, 10% glycerol, 0.1% NP-40, 1 Tab/100 ml COMPLETE

protease inhibitor cocktail (Roche)). Cells were lysed with 3 cycles of freezing and thawing in liquid nitrogen and room temperature, respectively. The lysates were cleared by centrifugation at 13,000 rpm for 5 min at 4°C and the soluble fraction transferred to a fresh tube.

An appropriate amount of the supernatant was mixed with antibodies at concentrations ranging from 1:6.5 (α -Myc) to 1:100 (α -Flag) and rotated at 4°C for 3 h. Protein A agarose beads (Oncogene) were then added according to the manufacturer's recommendations, and complexes were pulled by rotating at 4°C overnight. After washing with RIPA buffer, 3x SDS sample buffer was added, the samples were vortexed and cooked for 5 min, centrifuged, and the supernatant was subjected to SDS-PAGE.

Protein sequence analysis of complex samples by LC-MS/MS

Immunoprecipitations were subjected to one-dimensional SDS-PAGE. After staining with colloidal coomassie blue, each lane of the gel was cut into 24 pieces each 2-3 mm thick to reduce the complexity of the individual samples. The gel pieces were further chopped into small cubes of approximately 1 mm³ using a clean scalpel. Sample preparation for mass spectrometry was performed similar to Shevchenko et al. 1996 and as follows. Briefly, after washing with water, gel pieces were shrunk and dehydrated in acetonitrile followed by drying in a vacuum centrifuge. Gel pieces were then rehydrated in 10 mM DTT at 56°C for 50 min to reduce the protein and dehydrated again in acetonitrile followed by rehydration with 55 mM iodoacetamide for 20 min at room temperature to modify the thiol groups of cysteine residues. After two more de- and rehydration cycles using acetonitrile and 10 mM NH₄HCO₃, respectively, the gel pieces were immersed in digestion buffer (42 mM NH₄HCO₃, 4.2 mM CaCl₂) containing 12.5 ng/ μ l sequence-grade trypsin for 45 min at 4°C. The pieces were then completely covered with digestion buffer and incubated at 37°C overnight. Peptides were extracted by dehydrating the gel pieces using acetonitrile and collecting the supernatant. The remainder peptides were eluted by rehydrating the gel pieces with 5% formic acid followed by treatment with acetonitrile and subsequent collection of the supernatant. The pooled supernatants were then dried in a vacuum centrifuge and re-dissolved in 15 μ l of 10% acetonitrile and 0.2% formic acid followed by sonication. The samples (tryptic peptides) were first separated on a C₁₈ reverse-phase column and subsequently analyzed on-line by tandem mass spectrometry (4000 Q-TRAP, Applied Biosystems).

Cell culture

Cell lines

Drosophila Kc₁₆₇ cells (Echalier and Ohanessian 1970) were maintained in Schneider's medium (Invitrogen) supplemented with 10% fetal bovine serum (PAA) and penicillin-streptomycin (Invitrogen). Cells were grown at 25°C at subconfluent densities. To test the capability of *Drosophila* cells to be stimulated by human cytokines, the following recombinant ligands were used in the indicated concentration ranges: IL3 (Sigma, 8 ng/ml – 8 µg/ml), IL6 (Sigma, 10 units/ml – 10,000 units/ml) and Leptin (Sigma, 1.25 ng/ml – 1.25 µg/ml), none of which exerted an effect in the *Drosophila* JAK/STAT reporter assay outlined below.

HeLa SS6 cells (a kind gift from Jens Gruber and Mary Osborn) were maintained in D-MEM + 4,500 mg/ml glucose + L-glutamine + pyruvate (Gibco) including 10% heat inactivated fetal bovine serum (PAA) and penicillin-streptomycin (Invitrogen). Cells were grown at 37°C in a humidified incubator with 5% CO₂ at subconfluent densities. HeLa cells were stimulated with the following recombinant human ligands: IL6 (Sigma, 100 units/ml), sIL6R (R&D Systems, 100 ng/ml), IL2 (R&D Systems, 10 ng/ml), IL3 (Sigma, 10 ng/ml), IFNγ (R&D systems, 40 ng/ml) and OSM (R&D systems, 20 ng/ml).

Transfections

For transfection of *Drosophila* cells, 5x10⁶ cells were transfected with Effectene (QIAGEN) in 6-well plates with a total of 2 µg of DNA, 20 µl Enhancer, 16 µl Effectene and otherwise according to the manufacturer's protocol.

siRNA transfections using HeLa SS6 cells were performed with Oligofectamine Reagent (Invitrogen) according to the manufacturer's instructions with 100 nM or 50 nM siRNA in 96-well plates and OptiMEM (Gibco).

Generation of stable cell lines

Drosophila cells were transfected with 2 µg of the construct of interest to be stably expressed along with 50 ng of the vector *pCoBlast* to confer resistance against the antibiotic blasticidin (InvivoGen). Positive cells were selected using 30 µg/ml blasticidin and maintained in 10 µg/ml blasticidin solution.

High-throughput RNAi screening

The genome-wide RNAi screening experiments were performed in white, polystyrene 384-well tissue culture plates (Greiner). Screening plates were loaded with an average concentration of 75

nM dsRNA in 5 μ l of 1 mM Tris pH 7. Kc₁₆₇ cells were transfected in batch in 6-well plates with 0.5 μ g of the *p6x2xDrafLuc* JAK/STAT signaling reporter, 0.6 μ g of *pAct-UpdGFP* expression vector, 0.25 μ g pAc5.1-Sid-1 (to facilitate RNA uptake, Feinberg and Hunter 2003) and 0.025 μ g of *pAct-RL* vector as a co-reporter. The total plasmid amount was normalized to 2 μ g with the *pAc5.1* plasmid. After 7 h incubation at 25°C, batch transfected cells were resuspended in serum-free medium. Subsequently 15,000 cells in 20 μ l were dispensed per dsRNA containing well using an automated liquid dispenser (MultiDrop, Thermo Labsystems). Cells were incubated for 45 min, and 30 μ l of serum-containing medium was added to each well. Cells were grown for 5 d to allow for protein depletion. Pathway activity was measured using a luminescence assay for firefly and *Renilla* luciferase on a Mithras LB940 plate reader (Berthold Technologies). Luminescence of the *Renilla* luciferase was measured using a 490 nm filter set. Screens were performed in duplicate. Each plate contained dsRNA targeting *stat92E*, *dome*, *hop* and *socs36E* in A01, A02, B01, B02, which were used as positive controls (see also Figure 13).

Retests were performed similarly except that the JAK/STAT pathway reporter *p4xSocsLuc* was used instead of *p6x2xDrafLuc*.

Epistasis experiments in cells

In order to map the putatively positively interacting candidates according to their position in the signaling cascade, the JAK/STAT pathway was activated in cells with different stimuli. In each case, cells were transfected with the appropriate vectors (see below) for 7 h, and 30,000 cells were seeded in 50 μ l of serum-free medium in wells of clear-bottom 96-well plates (Greiner), which contained 1.5 μ g of the dsRNAs to be tested (see also Figure 13). Following 1 h incubation, 75 μ l medium supplemented with 10% fetal bovine serum was added to the cells. The plates were then sealed and cells were lysed after 5 d to measure luciferase activities as described above on a luminometer.

dsRNA of the positive regulators was tested for their ability to suppress pathway activity under three conditions: (1) in *upd*-expressing cells (screening conditions), (2) in cells treated with Upd-conditioned medium (Upd-CM), and (3) in cells expressing the activated form of JAK, *hop^{Tuml}* (Harrison et al. 1995, Luo et al. 1995). Specifically, for overexpression of *upd*, 5x10⁶ Kc₁₆₇ cells were transfected with 0.6 μ g *pAct-UpdGFP*, 0.5 μ g *p6x2xDrafLuc* reporter, 0.25 μ g *pAc5.1-Sid-1*, 0.025 μ g *pAct-RL* and *pAc5.1* to a total of 2 μ g DNA. For *hop^{Tuml}* overexpression, 5x10⁶ Kc₁₆₇ cells were transfected with 0.2 μ g *pAc5.1-hop^{Tuml}*, 0.5 μ g *p6x2xDrafLuc* reporter, 0.25 μ g *pAc5.1-Sid-1*, 0.025 μ g *pAct-RL* and *pAc5.1* to a total of 2 μ g DNA. To analyze processes upstream of Upd, two batches of cells were transfected separately to generate ‘responder’ and ‘Upd-producer’

cells. The ‘responder’ cells were transfected with 0.5 μg *p6x2xDrafLuc* reporter, 0.25 μg *pAc5.1-Sid-1*, 0.025 μg *pAct-RL* and *pAc5.1* to a total of 2 μg plasmid DNA in batch and subsequently seeded into 96-well plates containing the respective dsRNAs as described above. The ‘Upd-producing’ cells were transfected with 2 μg *pAct-UpdGFP* and cultured in 10 cm dishes (Falcon). Three d after transfection, cells were treated with 50 $\mu\text{g/ml}$ Heparin (Sigma). After 24 h, the supernatant was harvested, cleared by centrifugation and passed through a 0.2 μm filter (Millipore). 50 μl of the Upd-conditioned medium were then used to stimulate pathway activity in the ‘responder’ cells for 24 h. Control medium from untransfected Heparin treated cells did not elicit pathway activity. Experiments were performed in eight replicates and repeated at least twice. Reporter activity in the firefly luciferase channel was divided by the *Renilla* luciferase channel to normalize for cell number. z-scores were calculated as the multiples of the standard deviation that a specific RNAi treatment differed from cells treated with *lacZ* dsRNA as negative controls. z-scores were subsequently transformed into a false-color representation as depicted in Figure 20. RNA controls as shown in Figure 20 were *in vitro* transcribed from PCR templates generated using the gene-specific primer sequences shown in Table 3. Note that the *gfp* dsRNA was used to target the *upd-gfp* transgene and leads to a loss of pathway activity. *lacZ* dsRNA was used as a negative control.

For epistasis analysis of the putative negative regulator *ptp61F*, cells were batch transfected with reporter and Upd inducer as described above. Subsequently, these cells were treated with 1.5 μg of dsRNA targeting the *ptp61F* transcript and 1.5 μg of dsRNA against *lacZ*, *dome*, *hop* or *stat92E*. In parallel, cells from the same transfection batch were treated with *lacZ*, *dome*, *hop* or *stat92E* dsRNAs alone. After normalization, the values of experiments with control dsRNA alone were set to a value of 1.

Candidate phenotypes in cells

To examine the JAK/STAT phenotype of *ptp61F* in cells, 5×10^6 Kc₁₆₇ cells were transfected with 0.6 μg *pAct-UpdGFP*, 0.5 μg *p6x2xDrafLuc* reporter, 0.25 μg *pAc5.1-Sid-1*, 0.025 μg *pAct-RL* and *pAc5.1* to a total of 2 μg DNA. To assess the effects of the different *ptp61F* splice forms, cells were transfected as described before with additional 0.5 μg of *pAc5.1-Ptp61Fa*, *pAc5.1-Ptp61Fc* or vector control, respectively. JAK/STAT pathway activation was expressed in relation to control cells.

To assess the *dBRWD3* phenotype in cells, a similar procedure was applied with overexpression of 0.5 μg *BRWD3* expression vectors. For all other experiments (immunoprecipitation,

immunofluorescence experiments), 0.5 μg of the appropriate expression construct were transfected following the protocols outlined above.

Microscopy

Preparation of samples from cell culture experiments

Cells on coverslips were freed from excess medium and then fixed for 10 min in 5% formaldehyde. The fixative was removed, cells were rinsed with TBS (50 mM Tris pH 7.4, 150 mM NaCl) and then incubated in 0.2% Triton X-100 for 5 min. Cells were washed with TBS + 0.1% Triton X-100 three times and rinsed with TBS followed by incubation in blocking solution (TBS + 10% NHS (normal horse serum)) for 1 h at room temperature. Cells were then incubated with the primary antibody (α -Flag 1:1000, α -Myc 1:10) in blocking solution overnight at 4°C. After washing twice with TBS + 0.1% Triton X-100 and once with TBS, secondary antibody (Cy2 α -mouse (Jackson), 1:70) was added in TBS + 10% NHS for 45 min at room temperature. Cells were washed three times with TBS and, if desired, treated with TRITC-phalloidin (Sigma, 1:1000) and DRAQ5 (Biostatus, 1:1000). After washing three times with TBS, the coverslips were then mounted on glass slides in ProLong Gold Antifade Reagent (Invitrogen).

Epifluorescence microscopy

Epifluorescence microscopy was performed on a Zeiss Axiovert 200M inverted microscope.

Confocal microscopy

Confocal microscopy pictures were taken on a Leica DM RXA2 confocal microscope.

Determination of cell growth rates

For the determination of HeLa growth curves after siRNA treatment, wells of a 96-well cell culture clear-bottom plate (Greiner) were imaged on a Zeiss Axiovert 200M inverted microscope using a 10x objective and a brightfield filter setting. Per time-point 5 frames were saved per well, which were then analyzed by a custom-made ImageJ macro to automatically count particles (Supplementary Script 2).

Computational analyses

Determination of candidate hits in the genome-wide RNAi screen

To identify candidate genes that significantly increase or decrease JAK/STAT signaling pathway activity, the raw luciferase results were normalized by median centering of each 384-well plate

(separately by channel). z-scores were calculated as the number of median absolute deviation (MAD, Gentleman et al. 2004) that a particular well differed from the median of the 384-well plate. To minimize false-negatives, a set of low-stringency criteria was applied to generate a list of candidate genes to be used in specific retests. First, dsRNA treatments were filtered with z-scores > 2 for negative regulators or < -2 for positive regulators, respectively. Treatments that showed a high variability between duplicates were excluded. Further, RNAi experiments that showed z-scores of > 2 or < -2 in the control channel were not selected for retesting. The results were also filtered against previously identified cell viability modifiers that show a phenotype in cultured *Drosophila* cells (Boutros et al. 2004), and genes were excluded that showed phenotypes in other screens (Michael Boutros, unpublished). These filtering steps led to a final list of approximately 107 candidates that were selected for retesting. Data analysis and representation were performed using R (R-Development-Core-Team 2004) and Bioconductor (Gentleman et al. 2004).

Sequence analysis of candidate hits in the genome-wide RNAi screen

The predicted genes targeted by 91 retested dsRNAs were classified according to InterPro (Mulder et al. 2005) and GO (Drysdale et al. 2005, Harris et al. 2004), and manual inspection was used to order genes into functional groups. Predicted proteins without InterPro domain or GO annotation were classified as 'Unknown', although these sequences might encode structurally conserved proteins. To determine whether *Drosophila* proteins have homologs in other species, BLASTP searches were performed using the command line utility 'blastall' (Altschul et al. 1990) against the protein predictions from *Homo sapiens* (NCBI build 35) with a cut-off of $E < 10^{-10}$. Databases were obtained from Ensembl (<http://www.ensembl.org>, Clamp et al. 2003) and Flybase (<http://www.flybase.org>, Drysdale et al. 2005). Reciprocal best BLASTP analysis was used to identify the human homolog of CG31132. CG31132 and human BRWD3 are classified as orthologous pairs by InParanoid (<http://inparanoid.cgb.ki.se/>). Domain locations for dBRWD3-A, hBRWD3-A and hWDR9-A were obtained from NCBI (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein&itool=toolbar>).

Analysis of LC-MS/MS data

The proteins were identified by searching the MS-data (14,057 queries for α -Flag IP from mock transfected cells, 13,462 queries for α -Flag IP from dBRWD3-Flag transfected cells, 11,529 queried for α -Myc IP from Stat92E-Myc transfected cells, and 13,066 queries for α -Myc IP from mock transfected cells) against the NCBI non-redundant database with a '*Drosophila* (fruit flies)'

taxonomy filter covering 47,203 protein sequences using the Mascot search engine (Matrix Science). For the searches, oxidation of methionine was selected as a variable protein modification. Peptide mass tolerance was set to +/- 1.4 Da and MS/MS tolerance set to +/- 0.4 Da with a maximum number of one missed cleavage. Only proteins from *Drosophila melanogaster* were considered for further validation. Peptides specific for either the dBRWD3-Flag IP or the Stat92E-Myc IP population were subjected to manual inspection to minimize the number of false-positives in the dataset. All proteins with at least one validated peptide were considered a confident hit since the data was manually validated.

Protein interactions in FlyNet

Interactions for RNAi screen candidates were downloaded from the webversion of Flynet (<http://www.jhubiomed.org/perl/flynet.pl>) with a confidence threshold of 0.5 and plotted in the automated graph layout software Graphviz 1.13 (v16).

Protein interactions in OsPrey

Gene names of *Drosophila* RNAi screen candidates were uploaded into Osprey V1.2.0 leading to 88 unique recognized genes, and only interactions within these nodes were searched in the Fly GRID database. Control datasets for interactions were obtained by randomly sampling 88 genes (FBgn) ten times from a list comprising a total of 11,795 genes. These ten datasets were uploaded into Osprey V1.2.0, and only interactions within these nodes were searched in the Fly GRID database. Note that Osprey did not recognize some FBgns, and therefore some of the datasets above may contain up to two genes less. The mean of these ten datasets (= expected interactions) is 1.3 for all interactions and 1.0 for only non-self interactions.

Drosophila genetics

Fly strains and maintenance

Flies were maintained on a complex cornflour-soyflour-molasse medium (supplemented with dry yeast) at 25°C with 50 – 70% humidity and an approximately 12 h/12 h light/dark cycle. If not noted differently, flies were handled according to standard protocols (Ashburner 1989). Flies used in this study are listed in Table 5.

Generation of transgenic flies

Multiple independent transgenic *Drosophila* stocks of each transformation vector construct were generated by microinjection of *w¹¹¹⁸* embryos using standard techniques (Rubin and Spradling 1982). Injection of the plasmid/P-Helper DNA mix was performed by Iris Plischke.

Table 5. List of fly stocks.

Name	Genotype	Reference
Wildtype lines		
OreR	wildtype	Lindsley and Zimm 1992
white	w ¹¹¹⁸	Lindsley and Zimm 1992
Gal4 driver and UAS lines		
<i>ey-Gal4</i>	y,w;p{ <i>ey-Gal4</i> }4-8 / <i>Cyo</i>	BS 5535
<i>GMR-Gal4</i>	w;p{w ⁺ , <i>GMR-Gal4</i> } / (<i>Cyo</i>)	Freeman 1996
<i>cg-Gal4</i>	w;p{w ⁺ , <i>cg-Gal4</i> }2	BS 7011
<i>MS1096-Gal4</i>	w;p{ <i>GawB</i> }Bx ^{MS1096}	BS 8860
<i>UAS-PIAS-GFP-1</i>	y,w;p{w ⁺ , <i>UAS-ZIMP-GFP</i> }26b.3 / (<i>TM3,Sb</i>)	this study
<i>UAS-PIAS-GFP-2</i>	w;p{w ⁺ , <i>UAS-ZIMP-GFP</i> }31a.2 / (<i>Cyo</i>)	this study
<i>UAS-BRWD3-1</i>	w;p{w ⁺ , <i>UAS-dBRWD3</i> }1.0.3 / (<i>TM3,Sb</i>)	this study
<i>UAS-BRWD3-2</i>	w;p{w ⁺ , <i>UAS-dBRWD3</i> }2.2.2 / (<i>CyO</i>)	this study
<i>UAS-BRWD3-3</i>	w;p{w ⁺ , <i>UAS-dBRWD3</i> }3.2.2 / (<i>CyO</i>)	this study
<i>UAS-EGFP</i>	w;p{w ⁺ , <i>UAS-EGFP</i> }5a.2 / (<i>Cyo</i>)	M.P. Zeidler
<i>UAS-PTP61Fa-1</i>	w;p{w ⁺ , <i>UAS-ptp61Fa</i> }1a.3 / (<i>TM3</i>)	this study
<i>UAS-PTP61Fa-2</i>	w;p{w ⁺ , <i>UAS-ptp61Fa</i> }1b.2 / (<i>CyO</i>)	this study
<i>UAS-PTP61Fa-3</i>	w;p{w ⁺ , <i>UAS-ptp61Fa</i> }3a.3 / (<i>TM3</i>)	this study
<i>UAS-PTP61Fa-4</i>	w;p{w ⁺ , <i>UAS-ptp61Fa</i> }7a.3 / (<i>TM3</i>)	this study
<i>UAS-PTP61Fc-1</i>	w;p{w ⁺ , <i>UAS-ptp61Fc</i> }1a.1 / (<i>FM7</i>)	this study
<i>UAS-PTP61Fc-2</i>	w;p{w ⁺ , <i>UAS-ptp61Fc</i> }2a.4	this study
<i>UAS-PTP61Fc-3</i>	w;p{w ⁺ , <i>UAS-ptp61Fc</i> }2b.3 / <i>TM3</i>	this study
<i>UAS-lacZ</i>	w;p{w ⁺ , <i>UAS-lacZ.B</i> }BG4-1-2	BS 1176
FLP and FRT lines		
<i>hs-FLP</i>	y,w;p{w ⁺ , <i>hs-FLP</i> }; <i>Dr</i> / <i>TM3,Sb</i> ,p{w ⁺ , <i>hs-hid</i> }	Xu and Rubin 1993
<i>FRT82B-STAT^{6C8}</i>	w;p{ <i>neo</i> ^r , <i>FRT</i> }82B, <i>STAT^{6C8}</i> / <i>TM3</i>	Tulina and Matunis 2001
<i>FRT-STAT³⁹⁷</i>	p{ <i>neo</i> ^r , <i>FRT</i> }82B, <i>STAT³⁹⁷</i> , <i>e</i> / <i>TM3,Sb</i>	Silver and Montell 2001
<i>FRT-STAT⁰⁶³⁴⁶</i>	w;p{ <i>ry</i> ⁺ , <i>neo</i> ^r , <i>FRT</i> }82B, <i>STAT92E⁰⁶³⁴⁶</i> / <i>TM6c,SB,Tb</i>	Hou et al. 1996
<i>FRT82B-BRWD3-1</i>	w;p{ <i>neo</i> ^r , <i>FRT</i> }82B,l(3)05842 / <i>TM3,Sb</i>	this study
<i>FRT82B-BRWD3-2</i>	w;p{ <i>neo</i> ^r , <i>FRT</i> }82B,l(3)05842 / <i>TM3,Sb</i>	this study
<i>FRT82B-BRWD3-3</i>	w;p{ <i>neo</i> ^r , <i>FRT</i> }82B,l(3)05842 / <i>TM3,Sb</i>	this study
Other mutant or transgenic lines		
<i>dBRWD3⁰⁵⁸⁴²</i>	<i>ry⁵⁰⁶</i> ,p{ <i>ry</i> ⁺ , <i>PZ</i> }l(3)05842 ⁰⁵⁸⁴² / <i>TM3,Sb¹</i>	Spradling et al. 1999
<i>GMR-Upd</i>	y,w;p{w ⁺ , <i>GMR-UpdΔ3</i> } / <i>FM7</i> ,p{w ⁺ , <i>Ubg-GFP</i> }	Bach et al. 2003
<i>hop^{Tuml}; cg-Gal4</i>	y,w, <i>hop^{Tuml}</i> / <i>FM7</i> ;p{w ⁺ , <i>cg-Gal4</i> } / (<i>CyO</i>)	this study
<i>OvoD</i>	w;p{ <i>neo</i> ^r , <i>FRT</i> }82B,p{w ⁺ , <i>ovoD1</i> }3R/ <i>st¹ βTub85D^D ss¹ e^r</i> / <i>Tm3,Sb¹</i>	BS 2149
<i>Df(BRWD3)-1</i>	<i>Df(3R)crb87-4, st¹ e¹</i> / <i>TM3,Ser¹</i>	BS 2362
<i>Df(BRWD3)-2</i>	<i>Df(3R)crb87-5, st¹ e¹</i> / <i>TM3,Ser¹</i>	BS 2363
<i>STAT^{HJ}</i>	<i>ry,e,STAT92E^{HJ}</i>	Yan et al. 1996a

BS indicates the stock number of the Bloomington Stock Center (<http://flystocks.bio.indiana.edu/>).

Mobilization of P-elements

Twenty-three independent stocks, in which the ry^+ marker present in the $p\{ry^+,PZ\}$ insertion of $l(3)05842$ had been lost following a cross to a transposase source, were established following a P-element mobilization scheme (Ashburner 1989). Of these, seven were viable revertants (30%) and include two stocks with a wing vein phenotype (Figure 32), two are semi-lethal with occasional escapers and the remainder were lethal.

Genetic interaction assays

GMR-Upd: Genetic interaction with $p\{w^+,GMR-upd\Delta3'\}$ was undertaken as described in Bach et al. 2003 using OreR and $Stat92E^{06346}$ as negative and positive controls, respectively. Suppression of $p\{w^+,GMR-upd\Delta3'\}$ induced eye overgrowth by $dBRWD3^{05842}$ was observed in multiple independent experiments in a majority of individuals of the expected genotype. Flies were photographed using a Zeiss STEMI 2000-C binocular microscope and Axiocam camera.

hop^{TumI}: For genetic interaction assays, females of the stock $y,w,hop^{TumI} / FM7; p\{w^+,cg-Gal4.A\}2$ (Drysdale et al. 2005, Harrison et al. 1995) were crossed to wild type controls (OreR and w^{118}), mutations in $stat92E$ (Hou et al. 1996, Silver and Montell 2001) and $dBRWD3$ (Spradling et al. 1999). The presence of the hemocyte specific Gal4 driver $p\{w^+,cg-Gal4.A\}2$ (Drysdale et al. 2005) also allowed specific UAS insertions to be tested for their potential influence on the tumor formation. Transgenes expressing *EGFP* or β -galactosidase were used as negative controls while misexpression of *Drosophila dPIAS-GFP* served as a positive control as previously described (Betz et al. 2001). Crosses were incubated at 25°C and adult females carrying the hop^{TumI} chromosome were scored within 24 h of eclosion for the presence of tumors classified as small (one or two small melanotic spots as shown in Figure 28 (right)) or large (large melanized growths or more than three small spots; Figure 28 (left)). Survival rates for hop^{TumI} females appeared to be independent of tumor frequency at the time-point counted. Assays were repeated at least twice for each genotype.

Collection and fixation of embryos

Embryos were collected on apple juice agar plates (Ashburner 1989), dechorionated in 50% commercial bleach, fixed in 2% formaldehyde and heptane, devitellinized by adding methanol and stored in methanol at -20°C until further use.

Whole mount RNA in situ hybridization

In situ hybridization was performed essentially following the protocols described in (Tautz and Pfeifle 1989). Specifically, antisense RNA probes were prepared using the DIG-labeling kit

(Roche) by transcription from T7 or T3 promoters followed by cleaning up using the RNeasy kit (QIAGEN). Fixed embryos were rehydrated in a methanol/PBS series, washed in PBT (PBS + 0.1% Triton X-100) and refixed in 4% formaldehyde. After additional washes, embryos were treated with proteinase K followed by washing and another refixation step. After washing in PBT, embryos were put in a dilution series of PBT and hybridization buffer (50% formamide, 5x SSC, 100 $\mu\text{g/ml}$ tRNA, 50 $\mu\text{g/ml}$ heparin) and then immersed in hybridization buffer for pre-hybridization at 62°C. Finally, embryos were hybridized with the labeled probe in hybridization buffer at 62°C overnight. After extensive washing, embryos were incubated with α -DIG AP-conjugated antibody in PBT and 5% NHS for 1 h, washed, and the color was developed using NBT/BCIP solution (Roche). Embryos were washed extensively and then mounted on glass slides for microscopy.

Germline clones

Germline mutant clones were produced using the dominant female sterile (DFS) technique essentially as described in Chou and Perrimon 1996 and in the RESULTS section of the present study. Mutant clones were induced in approximately the third larval instar stage by heatshocking larvae inside food-containing vials in a 37°C waterbath for 1 h.

RESULTS

Forward genetic screens in *Drosophila* (Bach et al. 2003, Mukherjee et al. 2006) have aimed to systematically identify novel components of the JAK/STAT pathway. However, forward genetic screens intrinsically fail to deliver exhaustive candidate lists, and it is very likely that these forward genetic approaches have missed components, possibly due to non-saturating mutagenesis, genetic redundancy or phenotypic pleiotropy. In contrast, reverse genetic approaches using genome-wide cell based RNAi to systematically knock down every annotated open reading frame in the *Drosophila* genome are more suitable to comprehensively dissect cellular pathways and signaling networks (Boutros et al. 2004, Fraser et al. 2000, Gonczy et al. 2000, Kittler et al. 2004).

In order to use this powerful technology for the identification of novel JAK/STAT signaling modulators in *Drosophila* cultured cells, first a specific phenotype and a robust reporter system had to be devised, with which pathway signaling levels can be assessed and accurately quantified. A previous RNAi screen to identify genes relevant for cell growth and viability identified 438 dsRNAs targeting essential genes (Boutros et al. 2004). Knocking down genes essential for general cellular processes could also lead a positive identification if only one reporter phenotype is used to monitor JAK/STAT signaling activity, thereby skewing the possible candidate list with false-positive interactors and unspecific phenotypes. Furthermore, technical rather than biological artefacts and variability, such as relative differences in growth between wells, transfection efficiency, pipetting variability and cell lysis efficiency may also lead to the identification of false-positives if only one reporter channel is used (Armknrecht et al. 2005). To separate these anticipated possible artefacts from the true JAK/STAT specific phenotype, a dual-reporter strategy reporting pathway specific and nonspecific effects was developed. Assuming a given ratio of pathway reporter to co-reporter, only the pathway reporter channel should be modulated (up or down) upon knockdown of a true pathway modulator whereas the co-reporter signal remains unchanged. Should a dsRNA knock down a transcript involved in positively or negatively regulating viability, both reporters are expected to be modulated (up or down) thus allowing this false-positive phenotype to be distinguished from a true JAK/STAT phenotype.

Generation of a JAK/STAT dependent reporter system

A Stat92E responsive firefly (*Photinus pyralis*) luciferase reporter termed *p2xDrafSTAT(wt)* for use in *Drosophila* cultured cells has been described before. Kwon et al. 2000 used the Stat92E binding site present in the promoter of the JAK/STAT target gene *raf* to drive the expression of the firefly luciferase gene. Compared to other reporter genes (such as chloramphenicol acetyltransferase (CAT), β -galactosidase (lacZ), alkaline phosphatase (AP), green fluorescence protein (GFP)), luciferase reporters offer the advantage of high sensitivity and broad dynamic ranges combined with convenient assays for dual-channel readout and scalability to high-throughput experiments using a luminometric plate reader (Arnone et al. 2004, Gould and Subramani 1988, Naylor 1999). The reporter construct *p2xDrafSTAT(wt)* shows a five-fold induction above basal levels in *Drosophila* S2 cells after co-transfection with vectors encoding Hop and Stat92E relative to total protein amounts. This activity was shown to be specific, as the induction fails if the Stat92E binding sites present in this reporter are mutated (Kwon et al. 2000).

Following the general scheme proposed above and using the *p2xDrafSTAT(wt)* construct as a starting point, a JAK/STAT specific reporter system was devised using the three-component system shown in Figure 5. Instead of stimulating the cells with Hop as in Kwon et al. 2000, the cells were transfected with a vector encoding the pathway ligand Upd to mimic the *in vivo* induction. Furthermore, in contrast to normalizing to total protein amount as in Kwon et al. 2000, a construct with a constitutively active *actin 5c* promoter driving the expression of *Renilla* (*Renilla reniformis*) luciferase was transfected to utilize the ease and power of a dual-luciferase system (Sherf et al. 1996). Once provided with substrates, these luciferases convert the substrates into light, which can be quantified on a luminometer linearly over at least five orders of magnitude (Gould and Subramani 1988).

However, the five-fold induction for the *p2xDrafSTAT(wt)* reporter system (Kwon et al. 2000) is not amenable to high-throughput screening experiments, since these experiments typically result in high levels of noise that have to be distinguished from the effect of 'real' modulators (Malo et al. 2006). This requires robust reporters with a broad dynamic range. Furthermore, the five-fold induction was only achieved after co-transfection with a

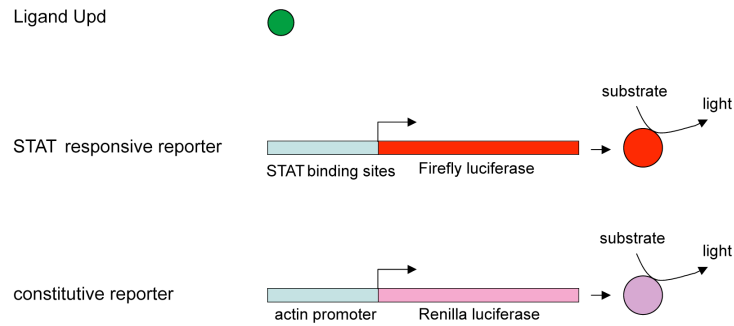


Figure 5. Scheme of multimerized reporter. The reporter system consists of a ligand (green) to stimulate the pathway and the phosphorylation of Stat92E, whose activity is reported by a STAT responsive reporter containing STAT binding sites. This promoter inducibly drives the expression of a firefly luciferase (red) and its protein product can convert a substrate into quantifiable light. A second coreporter bears the *actin* promoter, which constitutively drives the expression of the *Renilla* luciferase (purple) and whose protein product can convert a different substrate into light.

stat92E expressing construct. The *p2xDrafSTAT(wt)* reporter construct was thus taken as a lead to develop an improved more robust reporter assay. Four approaches were followed to optimize reporter activity:

- i) identification of a cell line with an intact pathway that does not require ectopic *stat92E* expression as in Kwon et al. 2000,
- ii) titration of the optimal amount and kind of ligand,
- iii) mutation of the ‘imperfect’ natural Stat92E binding sites present in the *p2xDrafSTAT(wt)* vector (Figure 3) towards the consensus binding site,
- iv) multimerization of Stat92E binding sites to allow for the binding of more STATs in the reporter region.

The *Drosophila* cell lines S2R+ (Yanagawa et al. 1998), S2 (Schneider 1972) and Kc₁₆₇ (Cherbas et al. 1977) were tested for their responsiveness to Upd ligand to induce JAK/STAT signaling. Basal levels of JAK/STAT pathway activity were all very low in these cells indicating only low-level expression of endogenous ligands. However upon addition of the ligand Upd, only S2 and Kc₁₆₇ cells showed significant induction whereas S2R+ appeared more unresponsive to Upd induction (Figure 6A and Figure 7). This is also consistent with microarray expression profiles available at <http://flight.licr.org> (Sims et al. 2006) showing low levels of *upd* in all cell lines but the highest expression values

for the most downstream pathway component *stat92E* in S2 and Kc₁₆₇ cells. Although endogenous *upd* expression levels are higher in S2R+ compared to S2 and Kc₁₆₇ cells, almost no *stat92E* is present in S2R+ cells, whereas *hop* and *dome* appear to be well present in S2 and Kc₁₆₇ cells with the highest relative expression levels in Kc₁₆₇ cells (Figure 6D). These results are further consistent given that Kc₁₆₇ cells are hemocyte-like *Drosophila* cells, a cell type in which JAK/STAT signaling is active *in vivo* (Meister and Lagueux 2003). Furthermore, the higher proliferation rates of Kc₁₆₇ cells compared to S2 and S2R+ cells are desirable traits that can be used to enhance the observed cellular phenotype due to faster protein depletion through more cell divisions after RNAi treatment.

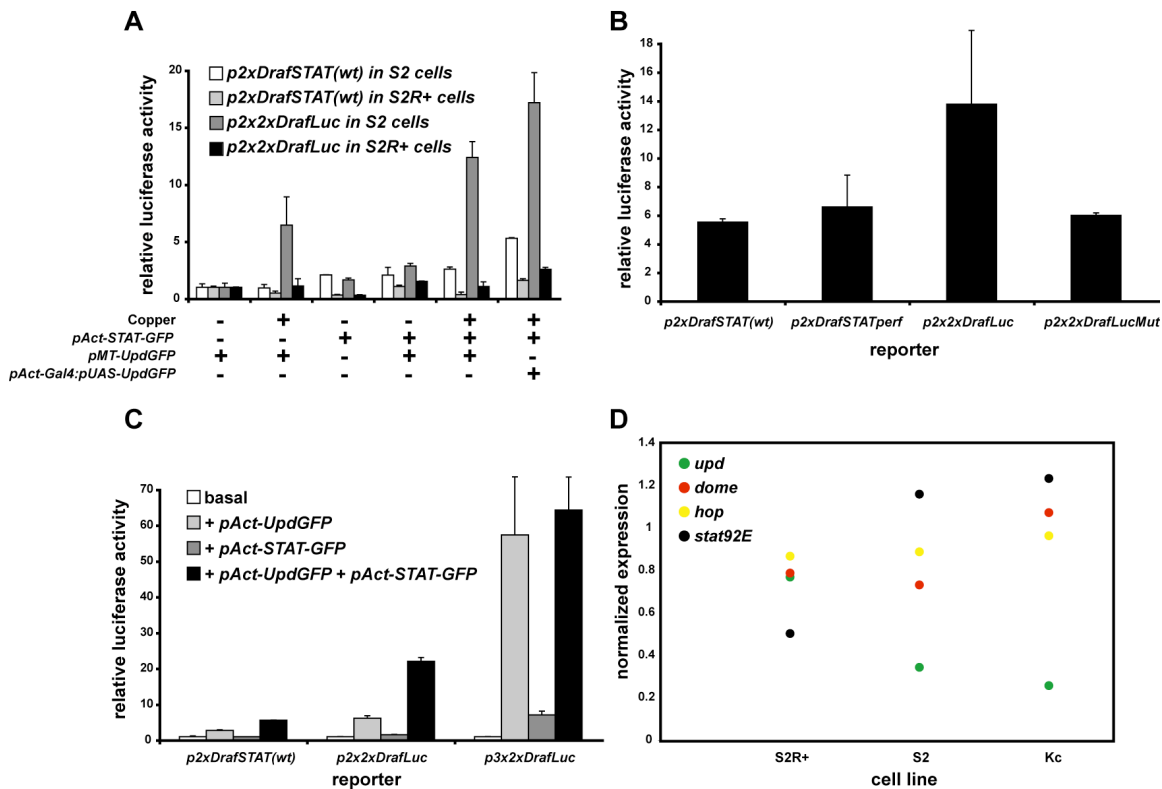


Figure 6. Development of an improved JAK/STAT reporter assay. (A) S2 or S2R+ cells were transfected with either *p2xDrafSTAT(wt)* or *p2x2xDrafLuc* reporters along with a *pAct-RL* co-reporter and the plasmids indicated with '+' below the figure. Cells were treated with CuSO₄ ('+' Copper) or not. Relative luciferase activity represents FL values divided by RL values, where all results are expressed to basal activities, which was set to a value of 1. Values are therefore unitless. Results for Kc₁₆₇ cells were similar in intensity as for S2 cells (Figure 7). (B) S2 cells were transfected with vectors encoding *upd-GFP*, *stat92E* and *RL* as well as the reporters indicated below the figure. Relative luciferase activity was calculated as in (A). (C) S2 cells were transfected with *pAct-RL* as well as the constructs indicated in the figure. Relative luciferase activity was calculated as in (A). (D) Normalized expression values from microarray experiments in untreated cell lines (S2R+, S2 and Kc₁₆₇) are shown. Data was obtained from <http://flight.licr.org> (Sims et al. 2006).

Next, the optimal amount of ligand to induce pathway activation was determined. In the planned high-throughput experiment, minimal manipulation of the cells would be most desirable. Therefore, cells stably expressing *pAct-UpdGFP* were generated. However, during the establishment of cell lines for the stable expression of *upd*, stable cell lines died within 3 weeks of *upd* expression suggesting that too much Upd is harmful for the cells in the longer term. On the other hand, cells inducibly expressing *upd* using a copper inducible reporter showed less reporter activation compared to *actin* promoter driven expression (Figure 8A, note that the expression of *upd-GFP* via Gal4/UAS system or directly by the *actin* promoter were similar). Additionally, the inducible cells would be more prone to copper-induced artefacts in a genome-wide RNAi screen. At the same time, inducing cells with copper after RNAi (i.e. re-opening the screening plates) increases the risk of contamination and is not feasible in high-throughput experiments. It is unlikely that recombinant Upd produced in bacterial cells, purified and consequently applied to *Drosophila* cells would exert an effect, as posttranslational modifications have been shown to be necessary for full Upd activity (Harrison et al. 1998). Stimulation with the related vertebrate cytokines IL6, IL3 and Leptin (Boulay et al. 2003) also did not elicit signaling activity. With this knowledge, a transient transfection approach using an *upd* expressing construct driven by an *actin* promoter was chosen to guarantee optimal pathway induction.

The Stat92E binding sites present in the *p2xDrafSTAT(wt)* reporter do not represent the preferred consensus binding sites that have been determined by *in vitro* studies (Figure 3A, Yan et al. 1996b). Therefore, the non-perfect Stat92E binding site TTCGCGGAA present in the *raf* promoter (Kwon et al. 2000) was mutated towards the perfect Stat92E binding site TTCCCGGAA. However, as shown in Figure 6B, the performance of the reporter *p2xDrafSTATperf* compared to *p2xDrafSTAT(wt)* did not show improved induction levels after *upd* transfection.

Another approach to get a better pathway reporter with more robust readout was to multimerize the Stat92E binding sites in the promoter region. This could potentially lead to the binding of more Stat92E proteins allowing multimerization (John et al. 1999) and thereby recruiting more transcription initiation complexes to the enhanceosome.

Ultimately, the 165 bp fragment from the original *p2xDrafSTAT(wt)* (Kwon et al. 2000) was inserted multiple times into the same vector to generate the *p6x2xDrafLuc* reporter with an enhancer of approximately 1,000 bp containing a total of 12 Stat92E binding sites. The intermediate constructs *p2x2xDrafLuc* and *p3x2xDrafLuc* with four and six Stat92E binding sites, respectively, showed greatly improved performance after Upd stimulation (Figure 6A-C). The higher activity in the *p2x2xDrafLuc(mut)* reporter, where the Stat92E binding sites are mutated, is probably due to residual Stat92E binding activity to less ideal binding sites (Figure 6B). Furthermore, already upon transfection of the *p3x2xDrafLuc* reporter and a plasmid to constitutively express the gene encoding the ligand Upd, a robust induction independent of ectopic *stat92E* expression (Figure 8C) was observed.

Dissection of signaling processes by RNAi

The results from these preliminary experiments to determine the optimal JAK/STAT reporter activity suggested the following setup for high-throughput screening: Transient transfection of Kc₁₆₇ cells with *pAct-UpdGFP*, *p6x2xDrafLuc* and *pAct-RL*, leading to a robust reporter induction (Figure 7). Next, it was examined whether depletion of known pathway components by RNAi (Clemens et al. 2000) could modify JAK/STAT signaling activity in Kc₁₆₇ cells. The effects of double-stranded RNAs (dsRNAs) targeting the mRNA of the genes *dome*, *stat92E* and *hop* was assessed as well as dsRNAs directed against the negative regulators *socs36E* and *dPIAS* after an incubation time of 5 d to allow for protein depletion. To enhance the efficiency of dsRNA uptake, a vector encoding the dsRNA transporter SID-1 from *C. elegans* was also used for the transfection of cells. As shown in Figure 7, knockdown of JAK/STAT components results in significant changes in reporter activity, whereas reporter activity in uninduced cells remains at low levels. For example, knockdown of the most downstream component *stat92E* leads to a decrease in signaling activity, and knockdown of the negative regulator *socs36E* leads to upregulated signaling levels, whereas depleting a component from a different unrelated signaling pathway (e.g. *relish/NFκB*) does not have an effect on reporter activity.

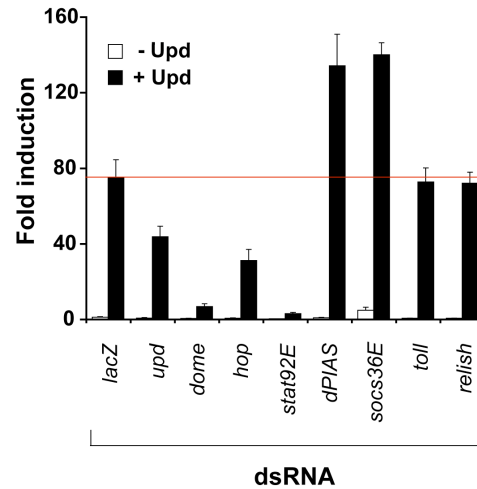


Figure 7. Dissection of the JAK/STAT pathway by RNAi. Knockdown of known JAK/STAT components leads to loss of pathway induction by Upd, whereas knockdown of *lacZ*, *toll* and *relish* shows no effect. The red line indicates an approximately 70-fold reporter induction relative to negative control dsRNA without Upd stimulation. Firefly luciferase (FL) values were divided by *Renilla* luciferase (RL) values and error bars represent standard deviations of six experiments.

Design of a genome-wide RNAi screen

To test whether this assay would be suitable for high-throughput conditions in 384-well plates, a pilot screen covering approximately 1,000 dsRNA probes was performed. The genome-wide RNAi screening library is designed such that PCR fragments containing T7 promoter sequences on each end (Hild et al. 2003) are used as templates to generate approximately 20,000 dsRNAs by *in vitro* transcription (Boutros et al. 2004). These quality controlled RNAs were diluted to a working stock concentration of approximately 100 ng/ μ l and aliquoted in ready-to-screen 384-well tissue culture plates. While dsRNAs from the library were present in all other wells, each plate also contained dsRNAs targeting *stat92E*, *dome*, *hop* and *socs36E* in positions A01, A02, B01, B02 which were used as positive controls with differently penetrant effects on the pathway readout. This way, the dynamic range could be determined to allow for an adjusted threshold in candidate selection. The pilot screen was performed in duplicates to decrease the width of data distribution by averaging between replicate data-points and to reduce the number of false-negatives while not increasing the number of false-positives (Malo et al. 2006).

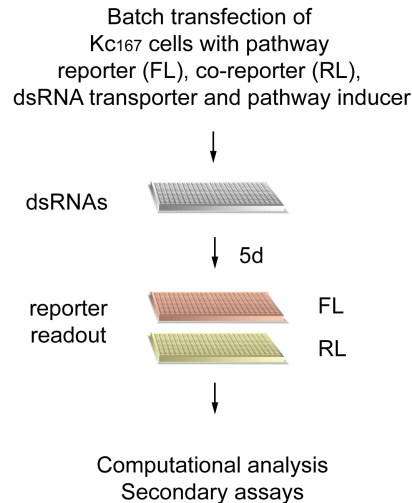


Figure 8. Design of a genome-wide RNAi screen for JAK/STAT signaling factors. Schematic diagram of screening approach using a multi-channel cell based setup. dsRNAs were screened in duplicate in 384-well plates. FL indicates firefly luciferase, RL indicates *Renilla* luciferase.

Figure 8 shows the overview of the screening strategy for the pilot screen and the whole-genome screen. Cells were batch transfected with the reporter system described above, seeded into wells containing individual dsRNAs in the absence of serum for 1 h to allow for dsRNA uptake, supplemented with serum and incubated for 5 d to allow for protein depletion. Finally, the reporter *p6x2xDrafLuc* and co-reporter *pAct-RL* activities were read out using a dual-luciferase assay. Figure 9 shows the results from the pilot screen (the scoring system is described in the section ‘Normalization approaches’ in more detail below) demonstrating that duplicate experiments are reproducible and robust enough to unambiguously identify dsRNAs from the screening set targeting components of the JAK/STAT pathway. For example, the ‘endogenous’ dsRNA targeting *hop* shows similar scores as the ‘spiked-in’ control targeting *hop*.

Genome-wide RNAi screening

As a next step, the complete library covering 20,026 dsRNAs targeting ~ 91% of the predicted transcripts in the *Drosophila* genome (Annotation 4.0, Misra et al. 2002) was screened in duplicate as outlined in Figure 8 using semi-automated logistics. These experiments resulted in large amounts of numerical tabular data to be computationally analyzed (i.e. a total of approximately 80,000 datapoints with ~ 40,000 datapoints from

the FL channel and $\sim 40,000$ datapoints from the RL channel). Normalization methods using robust estimators of center and spread for single channel experiments with only one assay read-out have only been described recently (Zhang et al. 2006, Boutros et al. 2006). Screens using more than one reporter readout have become more popular over the past years (Baeg et al. 2005, DasGupta et al. 2005), but a systematic benchmark study to set standards for normalization methods of multi-channel data as well as freely available software integrating algorithms to perform normalization steps have been lacking.

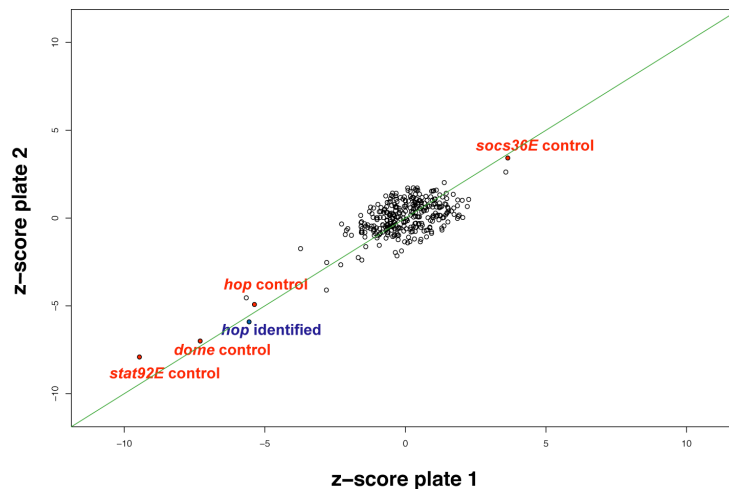
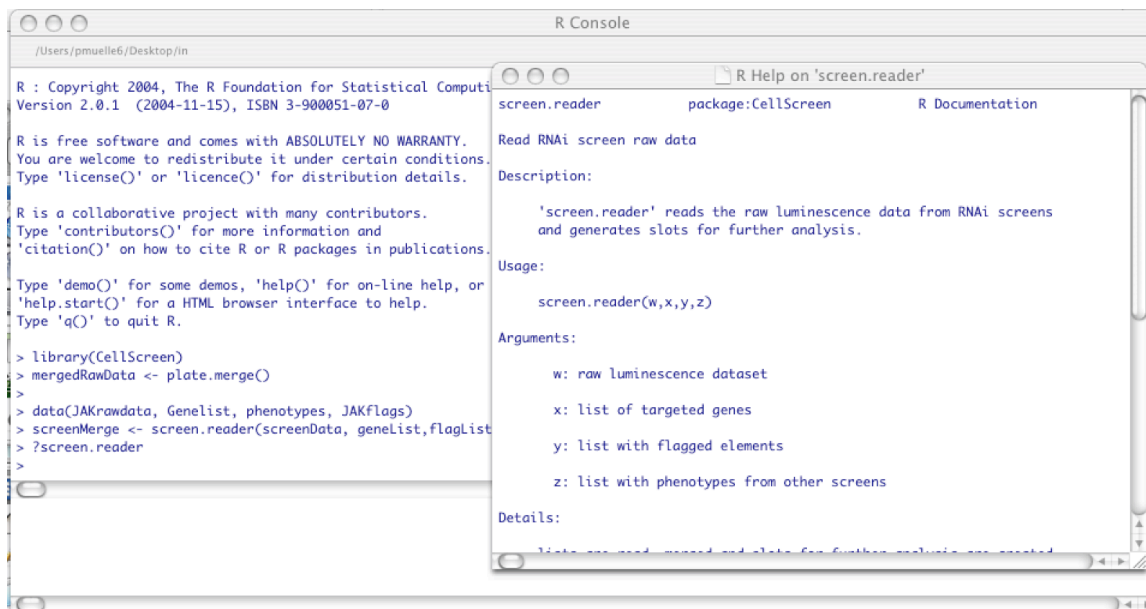


Figure 9. Pilot screen shows reproducibility and robustness. The FL z-scores (see section ‘Normalization approaches’) from two replicate plates covering 384 dsRNAs are plotted against each other. The green line represents a theoretical best fit with a correlation coefficient of $R=1$. The z-scores for ‘spiked-in’ controls are shown in red. Note that the data is well reproducible and that, as expected, most dsRNAs do not have a phenotype, i.e. the data-points are scattered around a z-score of 0. A dsRNA probe inside the plate targeting the pathway kinase *hop* (blue) can unambiguously be identified from the other dsRNAs.

Data analysis of the genome-wide RNAi screen

Normalization approaches

To provide a set of novel statistical approaches as well as tools facilitating the analysis of multi-channel datasets from reporter-based high-throughput RNAi screens, the software package *CellScreen* was therefore developed (Supplementary Script 1, Supplementary Tutorial). For the implementation of data analysis tools, the computational language and environment R (R-Development-Core-Team 2004) was chosen due to its ease in statistical computing, data handling, graphics and data distribution. A screenshot for the command line user interface and an example session in *CellScreen* is shown in Figure 10.



```
R : Copyright 2004, The R Foundation for Statistical Computing
Version 2.0.1 (2004-11-15), ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for a HTML browser interface to help.
Type 'q()' to quit R.

> library(CellScreen)
> mergedRawData <- plate.merge()
>
> data(JAKrawdata, Genelist, phenotypes, JAKflags)
> screenMerge <- screen.reader(screenData, geneList, flagList)
> ?screen.reader
>
```

```
screen.reader      package:CellScreen      R Documentation
Read RNAi screen raw data
Description:
'screen.reader' reads the raw luminescence data from RNAi screens
and generates slots for further analysis.
Usage:
screen.reader(w,x,y,z)
Arguments:
w: raw luminescence dataset
x: list of targeted genes
y: list with flagged elements
z: list with phenotypes from other screens
Details:
```

Figure 10. Screenshot of an example session in *CellScreen*. Shown is the command line interface of the R application and an open help file for the function *screen.reader*.

As shown in Figure 11A, there is considerable variability in the overall absolute values between individual plates of the genome-wide dataset, although the variability within plates is relatively minor (Armknecht et al. 2005). The causes for these inter-plate differences in absolute intensities could be, amongst others, varying environments inside the plates as well as the use of different reagent batches and/or screening of the plates on different days. In order to be able to compare the values from different plates, an independent measure has to be calculated for data normalization. Normalization methods can be used to remove systematic plate-to-plate variation making measurements comparable across plates. Boutros et al. 2006 and Zhang et al. 2006 have suggested z-score or quartile-based methods, respectively, using the robust estimators ‘median’ and ‘median absolute deviation’ (MAD) rather than ‘mean’ and ‘standard deviation’ (SD), which are more prone to the influence of outliers. Both z-score and quartile-based methods appear to function equally well in the robust identification of candidate modulators, even for non-symmetric datasets (Zhang et al. 2006). Log transformation of the raw data to stabilize the variance across the data range, a procedure frequently performed with microarray data (reviewed in Allison et al. 2006), does not appear to enhance the analysis of the data due to the smaller ranges of the raw values obtained in

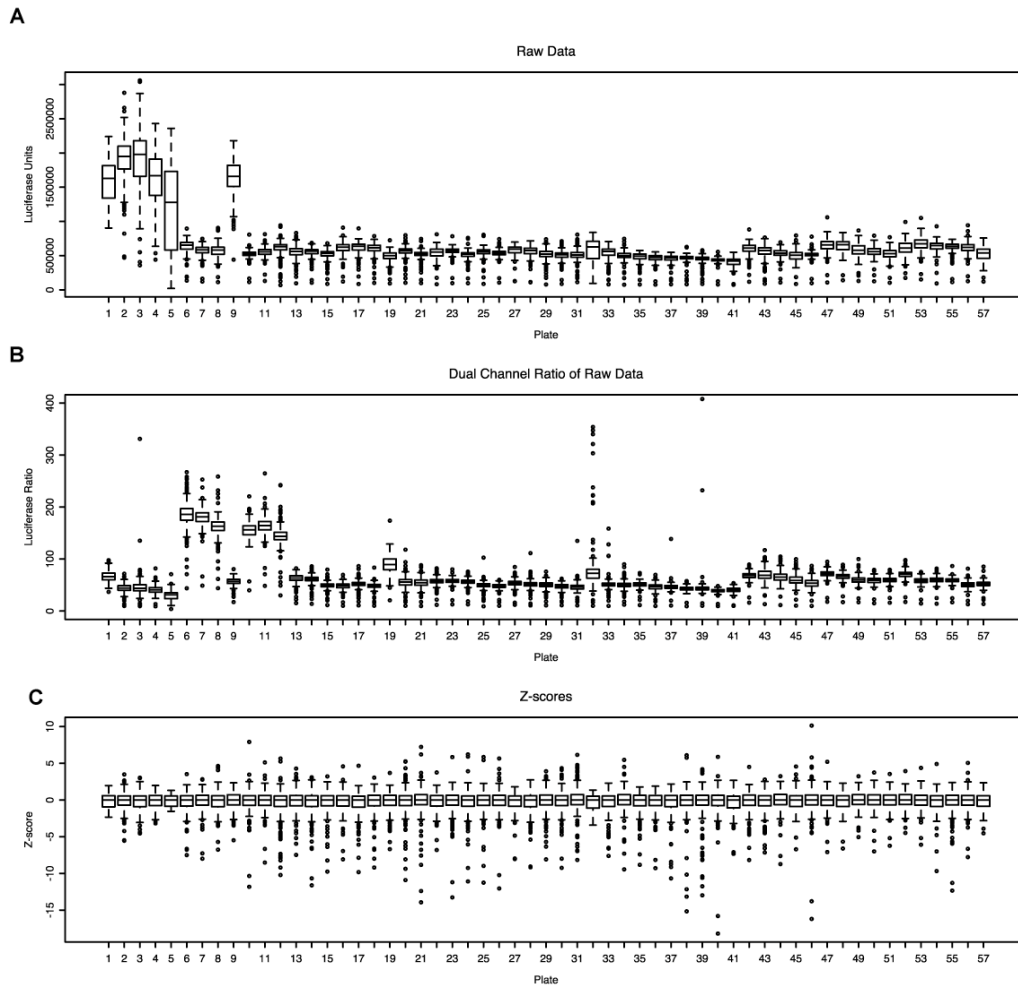


Figure 11. Median centering of screening plates. Boxplots are shown for individual plates from one replicate of a genome-wide RNAi screen. (A) Raw values for the pathway reporter channel. (B) Dual-channel normalized values expressed as the ratio of pathway reporter channel FL to co-reporter channel RL. (C) Values of the pathway reporter channel normalized by intra-plate median centering and calculation of z-scores.

RNAi screens (Boutros et al. 2006) compared to microarray approaches, where raw intensities of gene expression can be extremely high.

These normalization methods can likewise be used in multi-channel experiments. *A priori*, per-channel normalization methods appear to be more suitable for experiments where cells are ‘batch transfected’ before being transferred into individual dsRNA containing wells of screening plates due to the absence of differences in transfection efficiencies. In this scenario, each channel is treated separately and additional information can be obtained from the co-reporter channel, e.g. concerning non-specific

effects. In experiments where cells are transfected individually in each well of a screening plate, it may be necessary to derive a measure or ratio of the raw values obtained in the reporter channel relative to those from the co-reporter channel to normalize for transfection efficiency (Baeg et al. 2005). However, this dual channel normalization cannot compensate all variability in raw values between plates (Figure 11B). To identify candidate genes that significantly increase or decrease JAK/STAT signaling pathway activity, the raw luciferase results were therefore normalized by median centering of each 384-well plate separately per channel as a first step in data analysis (Figure 11C). z-scores were calculated as the number of median absolute deviation (MAD, Gentleman et al. 2004) that a particular well differed from the median of the 384-well plate making the values comparable between plates (Figure 11 C).

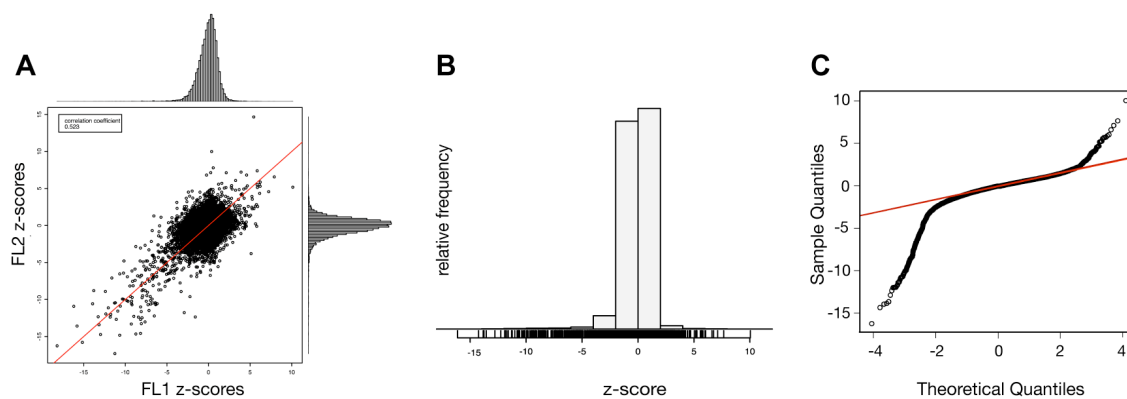


Figure 12. Properties of the data distribution from the genome-wide RNAi screen. (A) Reproducibility of screening data. Shown is a scatter plot of z-scores from replicate experiments for FL1 and FL2 channels as well as the histograms of the respective replicate datasets. The red line indicates a perfect fit of replicate experiments and a correlation coefficient of 1. (B) Narrow data distribution. Histogram plotting z-scores against relative frequency for the initial screen indicates that the majority of dsRNA molecules do not interact. Each band in the rug plot below indicates individual scores / phenotypes. (C) Non-symmetric data distribution. Q-Q plot of normally distributed quantiles against actual pathway screening result quantiles. The red line represents a fit to a normal distribution. Tails of positively and negatively interacting dsRNAs at each extreme with a z-score threshold of > 2 and < -2 represent RNAi experiments with significant phenotypes.

As a next step in data exploration, the normalized values from replicate experiments were compared (Figure 12A). As seen in the histograms for each replicate, the data distribution is very narrow indicating that most of the dsRNAs assayed did not have an effect. This is in agreement with the position-randomized library used for these experiments (Boutros et al. 2004). Secondly, the data appears to be reproducible with a Pearson's correlation

coefficient of 0.52 (a correlation coefficient of 1 represents a perfect fit). Similarly, the data-points averaged from normalized replicate experiments still show a very narrow distribution, indicating the similarity of the two replicate datasets (Figure 12B). A graph plotting the experimental quantiles against theoretical quantiles of a normal distribution (Q-Q-plot) further shows the non-normality as well as the asymmetry of the averaged dataset (Figure 12C). This indicates that the phenotypes obtained do not follow a normal distribution, which could be indicative of a random dataset. Furthermore, the ‘tails’ of the Q-Q-plot demonstrate that there are more positive regulators (with a negative z-score) than negative regulators (positive z-scores) identified in the screen.

Next, the data distribution and quality for individual plates was assessed. Figure 13A shows a narrow data distribution similar to the dataset as a whole, again in agreement with the hypothesis that most dsRNAs do not affect a specialized signaling cascade. Figure 13B demonstrates the plate setup that was used for the whole screen as well as the z-scores as phenotypes associated with the given dsRNAs in false-colors. Controls present in the top left corner of each plate were *hop* (A01), *dome* (A02), *stat92E* (B01) and *socs36E* (B02). dsRNAs from the library were present in all other wells including position B07 which targets *hopscotch* and L10 which targets *CG2033* and which was also previously identified in both Kc_{167} and S2R+ cells associated with a cell viability phenotype (Boutros et al. 2004). Similarly, I02 and G20 (which both target *sbr/CG17335*) were described in a previously published screen associated with a bi-nucleate phenotype (Kiger et al. 2003).

Dynamic range

The dynamic range can give insight into the robustness of the identification of hits in individual plates. The dynamic range can be defined as the spread between controls for all plates. Figure 14 shows such an analysis for the ‘spiked-in’ controls *stat92E* and *socs36E*, which among the controls had the strongest and most penetrant phenotypes. Averaged across the whole dataset, the dynamic range spans z-scores from approximately -8 to approximately +4 (Figure 14A), although plates with lower data quality exist (e.g. plate 32). This confirms the observation made in the Q-Q-plot from Figure 12C, where the tail of negative z-scores ranges lower than the tail for positive z-scores high in terms

of absolute z-score values. The same data is differently represented in the barplot of Figure 14B, where the z-scores for control dsRNAs targeting positive and negative regulators were subtracted making the visual identification of plates with possibly lower dynamic ranges, at least in the controls, identifiable. Figure 14C-F show a different representation for the assessment of data quality for individual plates. Here, the median of each plate is divided by the MAD for all channels separately (essentially the reverse of the ‘coefficient of variation’) revealing the inherent variability of the dynamic range in each plate. For example, a median/MAD of 10 would mean a variability of approximately 10% around the median, whereas a median/MAD of 5 would represent higher variability of approximately 20%. Given that the z-score is calculated as the fold MAD above or below the median, plates with a higher median/MAD in case of fixed thresholding for candidate selection would allow the facilitated identification of a modulator.

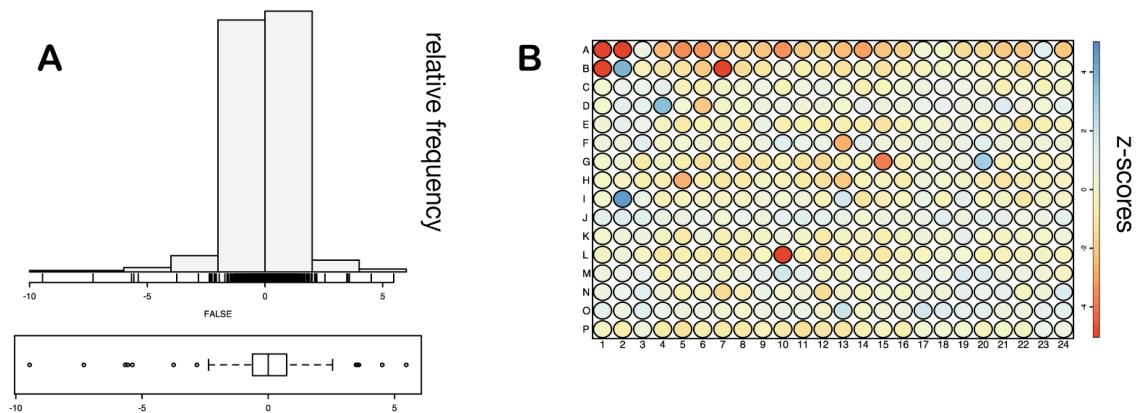


Figure 13. Plots of an individual plate to assess data quality. Shown are the distribution histogram, rug and barplot (A) as well as a false-color image (B) of averaged z-scores for the FL channel of plate 34. Color-coding for z-scores is shown in the key.

A way to take into account the dynamic range in the normalization of plates and to correct for plate-to-plate variability would be to express the data as percent of control per plate (Malo et al. 2006). However, the number of dsRNAs targeting controls is very limited in the setup of this genome-wide RNAi screen, whereas the density of non-interacting dsRNAs per plate to normalize against is much higher. The z-score excluding control measurements is therefore a more appealing approach for the normalization of this type of data to correct for plate-to-plate variability.

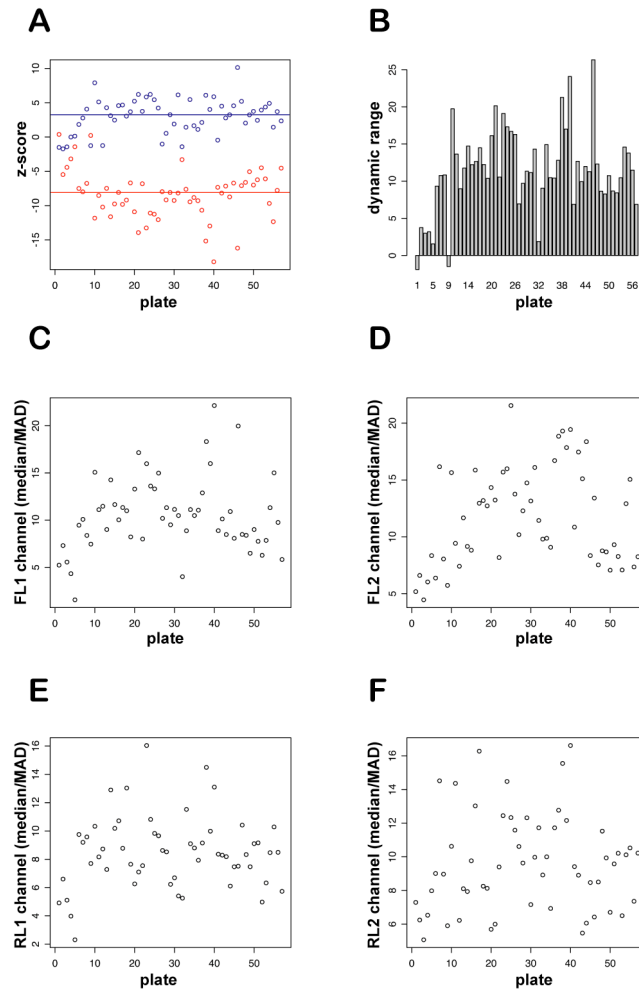


Figure 14. Dynamic range of screening plates. (A) Dynamic range for each plate shown for z-scores and the strong ‘spiked-in’ dsRNA controls targeting *stat92E* and *socs36E* in red and blue, respectively. The two red and blue straight lines represent the median for positive and negative regulator controls. In (B) the dynamic range is shown as the difference between the z-scores of the controls *stat92E* and *socs36E* on each plate. The ratios of median to MAD (labeled median/MAD) for all plates are shown for the channels FL1 (C), FL2 (D), RL1 (E) and RL2 (F). Plate 1 and 9 were initial trial plates in which the controls failed.

Detection of spatial artefacts

Similar to Figure 13B for a single plate, false-color graphs for the whole dataset were generated (Figure 15). Figure 15A shows the distribution of averaged z-scores for the firefly luciferase channel across the whole genome-wide RNAi screen for plates one to fifty-seven. The yellow background represents dsRNA treatments with insignificant scores confirming the observation from the histogram analysis in Figure 12B that only few gene activities are involved in JAK/STAT signaling, so that only few screening ‘hits’ are identified.

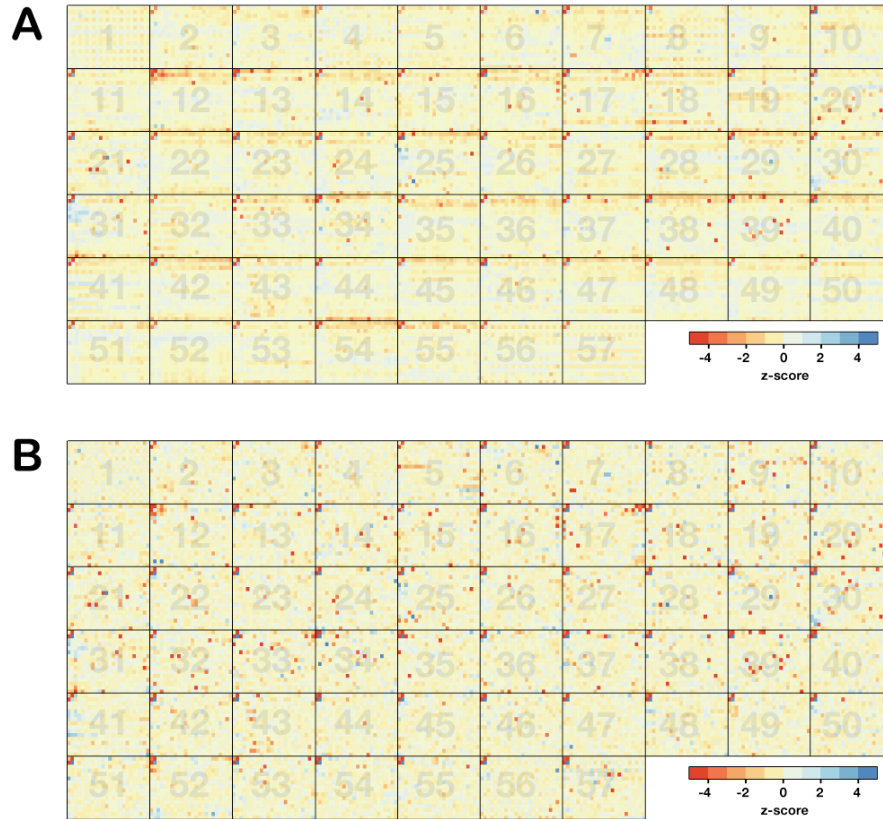


Figure 15. Overview of primary RNAi screen data and normalization methods. (A) False-color representation of z-scores for each well present in fifty-seven 384-well plates used for the initial screen. Key indicates the colors associated with the z-scores obtained: -4 (red) represents a decrease in reporter activity and hence a positive regulator. Conversely +4 (blue) represents an increase in activity and a negative regulator. Four controls were included in the top left corner of each plate. The interaction of these controls is visible in all plates except 1 and 9, which were initial trial plates in which these controls failed. (B) Heatmap showing a false-color representation of averaged B-Scores for all screening plates. Note the smoothing of edge-effects. Color-code for z-scores is the same as in (A).

However, technical or procedural factors can affect the outcome of screening experiments, e.g. poor pipette delivery, robotic failures, variable growth patterns, high versus low variability dsRNA probes, cell clumping or differences in dsRNA concentrations due to evaporation. These variations can increase the rate of false-positives and negatives (Malo et al. 2006). Formally this effect can be described such that

$$\text{Observed activity} = \text{true activity} + \text{effect of all errors}$$

Although these errors may be random, typically systematic errors also exist in high-throughput experiments (Malo et al. 2006). For example, it has been described for high-

throughput datasets that screening plates often show ‘edge-effects’ for uncertain reasons, maybe because of uneven evaporation. This means that wells located on the edges of screening plates may have a different activity in the assay channels compared to the rest of the plate (Brideau et al. 2003, Gunter et al. 2003). This could potentially cause the selection of false-positive candidates in further downstream applications and needs to be accounted for. Functions for plotting the processed data can therefore be useful for quality control to visually analyze systematic errors in the experimental data and to detect these spatial artefacts. The visual ‘heatmap’ of z-scores represented in the order how they appear can therefore also serve as another tool for the identification of these artefacts. A visual inspection of the plates shows a bias on the upper edges towards lower z-scores, and this bias seems to be systematic (Figure 15A). To refine the error assumption formalized above further, we can say that

$$\text{Observed activity} = \text{true activity} + \text{row-artefact} + \text{column-artefact} + \text{effect of all other errors}$$

where the row-artefacts and column-artefacts can be traced back to systematic errors. To further systematically analyze and reveal possible spatial across-plate and within-plate artefacts stemming from column or row edge-effects as well as position-related bias, another graphical analysis tool was developed. In Figure 16A, the screening scores are depicted such that the data is plotted by column of the 384-well plates. It is visually apparent that columns 1 and 2 have a striking pattern of significant positive and negative scores, whereas there is no apparent accumulation for the remainder of the dataset (the accumulation in columns 1 and 2, however, stems from the ‘spiked-in’ controls present in columns 1 and 2). On the other hand, the scores can also be plotted by row as shown in Figure 16B. This quality plot reveals what was already suspected from the graphical depiction in Figure 15A. There is a clear accumulation of significant scores in rows A and B as well as row P thereby revealing an edge-effect present in many plates of the dataset.

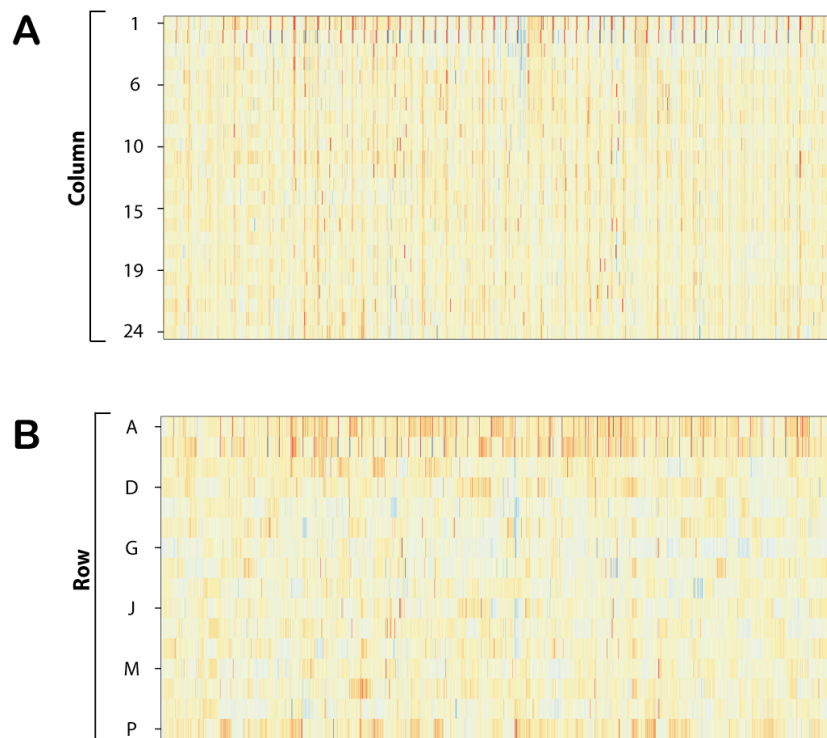


Figure 16. Detection of spatial screening artefacts. (A) Heatmap showing the presence of spiked-in controls in columns 1 and 2 for duplicate-averaged FL z-scores from fifty-seven 384-well plates. z-scores are plotted with their column index on the y-axis and their row index sequentially from plate 1 to plate 57 on the x-axis. (B) Heatmap showing positional artefacts in rows A, B and P for duplicate-averaged FL z-scores from fifty-seven 384-well plates. z-scores are plotted with their row index on the y-axis and their column index sequentially from plate 1 to plate 57 on the x-axis. Color-coding is as in Figure 15.

Given the systematic bias by well position detected using these plotting functions for the whole screening dataset, artefacts can be accounted for and smoothed using another approach. First, intra-plate normalization (2D) is performed, followed by normalization by well position through the whole dataset in the third dimension (3D). If the bias appears to be more randomly spread along rows and columns of the screening plates, these effects can be accounted for by the calculation of the so-called B-score using ‘Tukey’s two-way median polish procedure’ (Tukey 1977). The B-score is essentially similar to the z-score calculations but also takes into account plate-specific row and column errors masking the ‘true’ phenotype for a given well (Malo et al. 2006). Analyzing the same dataset shows a significant reduction of these edge-effects in Figure 15B as opposed to Figure 15A, and also the correlation coefficient between replicate experiments increases after two-way

median polishing (Pearson's correlation coefficient from 0.52 for z-scores to 0.60 for B-scores in 2D).

Table 6. Different output lists using different normalization approaches.

Reporter Channel	Normalization dimension	z-score	B-score	Common
FL	2D	1095	1401	628
RL	2D	829	1565	481
FL/RL	2D	1458	1541	783
FL	3D	871	1365	521
RL	3D	703	1523	454
FL/RL	3D	1314	1574	837

Shown is the number of candidates for each approach with scores > 2 or < -2 .

The column 'Common' indicates the overlap between the number of candidates from z-score and B-score calculations for the given normalization dimension and reporter channel.

Hit selection

To identify candidate genes that significantly increase or decrease JAK/STAT signaling pathway activity, z-scores were calculated as the number of MAD that a particular well differed from the median of a 384-well plate. To minimize false-negatives, a set of low-stringency criteria was applied to generate a list of candidate genes to be used in specific retests. dsRNA treatments with z-scores > 2 for negative regulators or < -2 for positive regulators were filtered, respectively (representing p-values of 0.05), theoretically leading to a selection of 5% of candidates for all treatments (Table 6). Interestingly, when this data is compared to other normalization methods like normalization by well position in 3D or by median polishing applying the B-score, different candidate lists are generated with varying overlap depending on the reporter channel analyzed and with the best overlap between z-score and B-score approaches after dual-channel normalization as well as normalization in 3D (Table 6).

Similarly, for individual plates, the chosen normalization approach can have a dramatic effect on candidate selection. As an example, plate 20 was analyzed using all the different normalization approaches described above (Figure 17). Both 3D-normalization and B-score are efficient in smoothening edge-effects (e.g. compare Figure 17A to Figure 17B as well as Figure 17G to Figure 17H). It is also apparent that the 3D normalization

approach cancels out the phenotypes for control dsRNA treatments in wells A01, A02, B01 and B02 (e.g. compare Figure 17A to Figure 17B). Furthermore, as the edge-effects decrease – i.e. those treatments with weak phenotypes that become normalized – the strength of putatively real interactors increases. Note for example the effect that 3D-normalization has on well G18 for a positive regulator and L19 for a negative regulator (Figure 17A and B). These effects are even more pronounced after calculation of B-scores, as can be seen for well D18 (e.g. Figure 17G and H). The dual-channel normalization approach, where the ratio of firefly to *Renilla* luciferase channel is calculated (FL/RL), has the potential to generate artefactual signals. Note the strong negative regulator identified in the RL channel in well A16 (Figure 17C) and which has no phenotype in the JAK/STAT pathway reporter channel FL (Figure 17A). In case of dual-channel normalization, this treatment would lead to the identification of a pseudo-positive regulatory dsRNA treatment (Figure 17E).

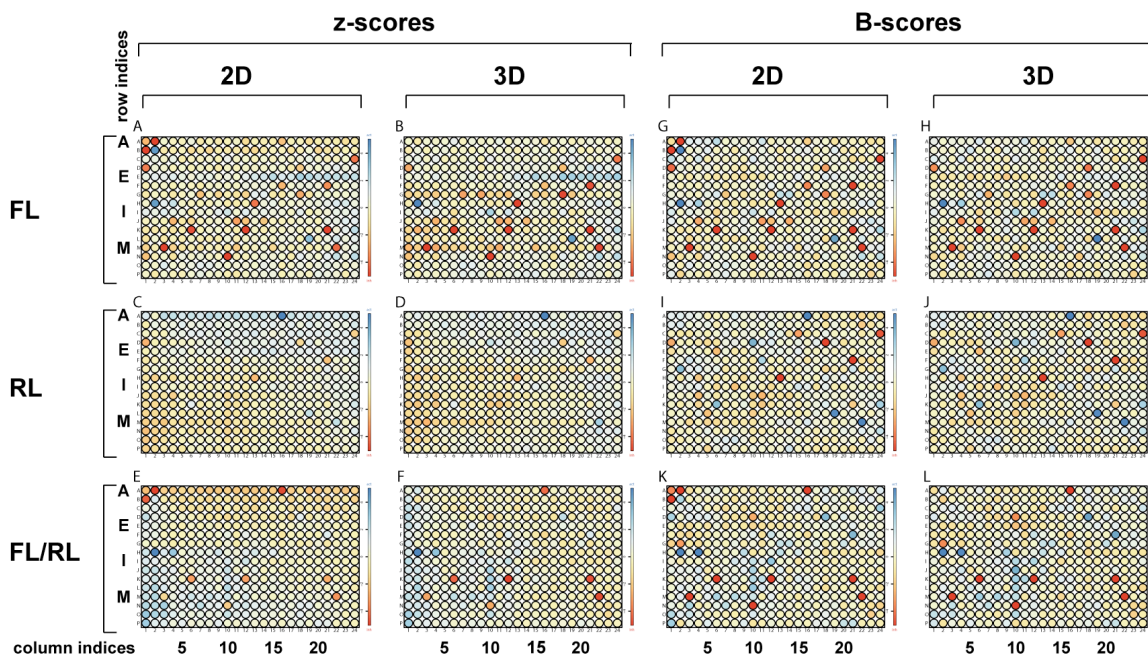


Figure 17. Exemplary comparison of normalization approaches. Shown are false-color representations for screening plate 20. z-scores are shown in (A)-(F), B-Scores in (G)-(L). (A), (B), (G) and (H) show the values for the FL channel. (C), (D), (I) and (J) show the values for the RL channel. (E), (F), (K) and (L) show the z-scores for the ratio of FL and RL channel. (A), (C), (E), (G), (I) and (K) show values normalized in ‘2D’ by plate centering. (B), (D), (F), (H), (J) and (L) show values normalized in ‘3D’ after plate and well centering. All values represent averages of duplicate experiments. To facilitate readability, column and row indices are indicated in bold once more on the bottom and left of the figure, respectively, in addition to their appearance below every plate depiction.

For candidate selection, the z-score in 2D for the firefly luciferase channel (FL) was eventually used to allow for the comparison with other published and unpublished datasets analyzed using the same approach. dsRNA treatments with z-scores > 2 or < -2 were selected and treatments that showed a high variability between duplicates were excluded. Further, RNAi experiments that showed z-scores of > 2 or < -2 in the control *Renilla* luciferase channel were not selected for retesting. Additionally, the data was filtered against previously identified cell viability modifiers that show a phenotype in cultured *Drosophila* cells (Boutros et al. 2004). Also, genes that showed phenotypes in other screens were excluded (Michael Boutros, unpublished) to identify those modulators that are specifically regulating the ‘core’ JAK/STAT pathway and to exclude the broader network influencing this cascade (Table 7, Table 8).

Table 7. *Drosophila* JAK/STAT phenotypes (negative regulators).

Gene	probe ID	z-score 1	z-score 2	Functional group	IPR	GO evidence
<i>bon</i>	HFA16914	5.6	4.8	Protein modifying enzymes / Metabolism	IPR001841	GO:0004842; ubiquitin-protein ligase activity
<i>Caf1</i>	HFA16596	3.0	2.6	Protein modifying enzymes / Metabolism	IPR001680	GO:0035035; histone acetyltransferase binding
<i>CG10077</i>	HFA09691	2.8	4.0	RNA processing and Translation	IPR001410	GO:0003724; RNA helicase activity
<i>CG11400</i>	HFA06070	2.6	2.2	Unknown	no IPR	na; na
<i>CG11501</i>	HFA14317	3.7	3.1	Unknown	no IPR	na; na
<i>CG13499</i>	HFA04144	2.5	3.1	Unknown	no IPR	na; na
<i>CG14247</i>	HFA14742	3.2	3.4	Unknown	IPR002557	na; na
<i>CG15706</i>	HFA06577	2.2	2.1	Unknown	IPR011701	na; na
<i>CG16975</i>	HFA02552	2.7	2.7	Transcription regulators	IPR001660	GO:0030528; transcription regulator activity
<i>CG17492</i>	HFA02623	2.5	2.1	Protein modifying enzymes / Metabolism	IPR001841	GO:0004842; ubiquitin-protein ligase activity
<i>CG18112</i>	HFA15304	2.1	2.1	Unknown	IPR001829	na; na
<i>CG30122</i>	HFA06935	3.3	2.8	Transcription regulators	IPR003034	GO:0003677; DNA binding
<i>CG4907</i>	HFA15673	3.3	3.5	Unknown	IPR007070	na; na
<i>dre4</i>	HFA08714	2.6	2.5	Transcription regulators	IPR000994	GO:0003712; transcription cofactor activity
<i>enok</i>	HFA04096	3.0	3.0	Transcription regulators	IPR001965	GO:0030528; transcription regulator activity
<i>lig</i>	HFA07247	2.2	2.1	Unknown	IPR009060	na; na
<i>Nup154</i>	HFA03384	2.9	2.9	Cytoskeleton and Transport	IPR011045	GO:0005487; nucleocytoplasmic transporter activity
<i>par-1</i>	HFA07660	4.4	4.2	Signal transduction	IPR000719	GO:0004674; protein serine/threonine kinase activity
<i>Pp1alpha-96A</i>	HFA16795	3.0	3.8	Signal transduction	IPR006186	GO:0004722; protein serine/threonine phosphatase activity
<i>PP2A-B'</i>	HFA16344	2.6	2.5	Signal transduction	IPR002554	GO:0008601; protein phosphatase type 2A regulator activity
<i>Ptp61F</i>	HFA08683	5.9	8.1	Signal transduction	IPR000863	GO:0004725; protein tyrosine phosphatase activity
<i>Rab5</i>	HFA00777	2.1	2.1	Signal transduction	IPR001806	GO:0005525; GTP binding
<i>Socs36E</i>	HFA02455	3.2	2.3	Signal transduction	IPR000980	GO:0007259; JAK-STAT cascade
<i>TSG101</i>	HFA11098	3.1	3.4	Protein modifying enzymes / Metabolism	IPR001440	GO:0004842; ubiquitin-protein ligase activity

The column ‘probe ID’ indicates the dsRNA used to obtain the JAK/STAT phenotype. Sequence information is available in Supplementary Table 1 and at <http://rna1.dkfz.de>

The column ‘z-score 1’ shows the averaged z-score phenotype for the FL channel obtained from replicate experiments in the initial genome-wide RNAi screen using the *p6x2xDrafluc* reporter. The column ‘z-score 2’ shows the phenotype obtained in the retest experiments using newly synthesized dsRNA and the *p4xSOCSLuc* reporter.

The column ‘IPR’ indicates InterPro evidence, which was taken from Mulder et al. 2005.

GO evidence was taken from Drysdale et al. 2005 (<http://flybase.org>).

All 384-well screening plates contained dsRNA against known JAK/STAT pathway components. Controls for the 57 screening plates were *stat92E* RNAi (identified 55 times), *hop* RNAi (identified 37 times), *dome* RNAi (identified 55 times) and *socs36E* RNAi (identified 45 times).

An interactive table with links to the InterPro records is available at <http://www.dkfz.de/signaling/jak-pathway/>

Table 8. *Drosophila* JAK/STAT phenotypes (positive regulators). Representation is the same as in Table 7.

Gene	probe ID	z-score 1	z-score 2	Functional group	IPR	GO evidence
<i>Arl2</i>	HFA00627	-2.9	-3.2	Protein modifying enzymes / Metabolism	IPR000051	GO:0016274; protein-arginine N-methyltransferase activity
<i>asf1</i>	HFA11324	-2.3	-2.5	Others	IPR008967	GO:0003682; chromatin binding
<i>bin3</i>	HFA04919	-3.1	-3.3	Unknown	IPR000051	na; na
<i>CG10007</i>	HFA14173	-3.2	-2.9	Unknown	no IPR	na; na
<i>CG10730</i>	HFA02102	-2.1	-2.3	Unknown	IPR004245	na; na
<i>CG10960</i>	HFA09807	-2.0	-2.1	Protein modifying enzymes / Metabolism	IPR005829	GO:0005355; glucose transporter activity
<i>CG11307</i>	HFA11648	-2.3	-2.4	Unknown	no IPR	GO:0016757; transferase activity
<i>CG11696</i>	HFA19417	-2.0	-2.3	Transcription regulators	IPR007087	GO:0003677; DNA binding
<i>CG12213</i>	HFA14478	-3.3	-3.2	Unknown	IPR009053	na; na
<i>CG12460</i>	HFA20970	-3.3	-3.4	Transcription regulators	IPR000504	GO:0030528; transcription regulator activity
<i>CG12479</i>	HFA19459	-2.3	-2.4	Unknown	IPR007512	na; na
<i>CG13243</i>	HFA01920	-2.7	-2.6	Unknown	IPR003117	na; na
<i>CG13473</i>	HFA10017	-2.4	-2.1	Cytoskeleton and Transport	IPR006662	GO:0005489; electron transporter activity
<i>CG14434</i>	HFA17927	-2.0	-2.3	Unknown	IPR008173	na; na
<i>CG15306</i>	HFA17993	-3.3	-3.1	Signal transduction	IPR001715	GO:0005102; receptor binding
<i>CG15418</i>	HFA00432	-2.1	-2.1	Protein modifying enzymes / Metabolism	IPR002223	GO:0004866; endopeptidase inhibitor activity
<i>CG15434</i>	HFA00449	-2.5	-2.9	Protein modifying enzymes / Metabolism	IPR007741	GO:0003954; NADH dehydrogenase activity
<i>CG15555</i>	HFA15093	-2.3	-2.6	Others	IPR001873	GO:0015268; alpha-type channel activity
<i>CG15784</i>	HFA18090	-2.4	-2.6	Unknown	IPR009072	na; na
<i>CG16903</i>	HFA18561	-2.8	-2.8	Transcription regulators	IPR011028	GO:0016251; general RNA polymerase II transcription factor activity
<i>CG17179</i>	HFA10258	-2.1	-2.8	Unknown	IPR001660	na; na
<i>CG18160</i>	HFA21006	-3.1	-2.4	Unknown	no IPR	na; na
<i>CG30069</i>	HFA06272	-2.9	-2.2	Protein modifying enzymes / Metabolism	no IPR	GO:0016491; oxidoreductase activity
<i>CG3058</i>	HFA00563	-3.4	-3.5	Cytoskeleton and Transport	IPR006663	GO:0005489; electron transporter activity
<i>CG31005</i>	HFA15507	-2.3	-3.0	Protein modifying enzymes / Metabolism	IPR000092	GO:0000010; trans-hexaprenyltransferase activity
<i>CG31132</i>	HFA16032	-2.8	-3.5	Unknown	IPR001487	na; na
<i>CG31132</i>	HFA15369	-2.3	-3.6	Unknown	IPR001487	na; na
<i>CG31358</i>	HFA15235	-2.0	-2.2	Cytoskeleton and Transport	IPR001972	GO:0005200; structural constituent of cytoskeleton
<i>CG31694</i>	HFA00415	-2.8	-2.7	Signal transduction	IPR006921	GO:0005102; receptor binding
<i>CG32406</i>	HFA09966	-2.1	-2.2	Signal transduction	IPR000980	na; na
<i>CG32573</i>	HFA19906	-3.1	-2.9	Unknown	IPR000719	na; na
<i>CG3281</i>	HFA15470	-3.1	-3.0	Transcription regulators	IPR007087	GO:0030528; transcription regulator activity
<i>CG3819</i>	HFA10378	-2.3	-2.3	Unknown	IPR001604	na; na
<i>CG4022</i>	HFA10395	-3.4	-3.7	Unknown	no IPR	na; na
<i>CG40351</i>	HFA20930	-2.6	-2.7	Transcription regulators	IPR001214	GO:0030528; transcription regulator activity
<i>CG4349</i>	HFA19892	-4.1	-2.1	Others	IPR009040	GO:0008199; ferric iron binding
<i>CG4446</i>	HFA10420	-2.7	-2.7	Protein modifying enzymes / Metabolism	IPR004625	GO:0008478; pyridoxal kinase activity
<i>CG4653</i>	HFA19909	-3.2	-3.0	Protein modifying enzymes / Metabolism	IPR001254	GO:0004263; chymotrypsin activity
<i>CG4781</i>	HFA04488	-2.5	-2.5	Unknown	IPR003591	na; na
<i>CG6422</i>	HFA16036	-3.3	-3.2	Unknown	IPR007275	na; na
<i>CG6434</i>	HFA10635	-2.8	-2.8	Unknown	IPR001680	na; na
<i>CG6946</i>	HFA16145	-2.3	-2.9	RNA processing and Translation	IPR000504	GO:0003723; RNA binding
<i>CG7635</i>	HFA20054	-2.9	-2.8	Cytoskeleton and Transport	IPR001972	GO:0005200; structural constituent of cytoskeleton
<i>CG8108</i>	HFA09675	-2.7	-2.7	Transcription regulators	IPR007087	GO:0003676; nucleic acid binding
<i>CG9086</i>	HFA20148	-2.8	-2.9	Signal transduction	IPR009030	GO:0005057; receptor signaling protein activity
<i>CKIIalpha</i>	HFA11946	-2.1	-2.5	Signal transduction	IPR000719	GO:0004702; receptor signaling protein serine/threonine kinase activity
<i>CKIIbeta</i>	HFA20230	-2.7	-2.6	Signal transduction	IPR000704	GO:0004702; receptor signaling protein serine/threonine kinase activity
<i>comm3</i>	HFA09995	-2.2	-2.2	Unknown	no IPR	na; na
<i>CtBP</i>	HFA16617	-2.9	-2.8	Transcription regulators	IPR006139	GO:0003714; transcription corepressor activity
<i>dome</i>	HFA19583	-6.2	-4.9	Signal transduction	IPR000194	GO:0004907; interleukin receptor activity
<i>eIF-4B</i>	HFA20983	-3.2	-3.0	RNA processing and Translation	IPR000504	GO:0003723; RNA binding
<i>HDC01676</i>	HFA01091	-2.3	-2.6	Unknown	IPR006202	na; na
<i>HDC11198</i>	HFA11427	-2.3	-2.2	Unknown	no IPR	na; na
<i>hop</i>	HFA20340	-5.7	-4.1	Signal transduction	IPR001245	GO:0004718; Janus kinase activity
<i>Ipk2</i>	HFA00357	-2.6	-4.0	Signal transduction	IPR005522	GO:0050516; inositol-polyphosphate multikinase activity
<i>jbug</i>	HFA04167	-2.7	-3.2	Cytoskeleton and Transport	IPR001298	GO:0005200; structural constituent of cytoskeleton
<i>kn</i>	HFA07637	-2.4	-2.4	Transcription regulators	IPR003523	GO:0030528; transcription regulator activity
<i>l(1)G0084</i>	HFA19450	-2.1	-2.1	Transcription regulators	IPR001965	GO:0003677; DNA binding
<i>larp</i>	HFA16984	-2.5	-2.4	Unknown	IPR006630	na; na
<i>mask</i>	HFA15370	-2.3	-2.7	Signal transduction	IPR002110	GO:0005102; receptor binding
<i>mst</i>	HFA20582	-2.2	-2.6	Unknown	no IPR	na; na
<i>nonA</i>	HFA20357	-3.0	-3.3	RNA processing and Translation	IPR000504	GO:0030528; transcription regulator activity
<i>Obp93a</i>	HFA15220	-2.4	-2.9	Cytoskeleton and Transport	IPR006170	GO:0005549; odorant binding
<i>Rrp1</i>	HFA00784	-4.3	-4.3	Others	IPR000097	GO:0004520; endodeoxyribonuclease activity
<i>sol</i>	HFA20587	-2.5	-3.0	Others	IPR001876	GO:0005516; calmodulin binding
<i>Stat92E</i>	HFA16870	-5.0	-5.2	Signal transduction	IPR001217	GO:0004871; signal transducer activity
<i>Taf2</i>	HFA11298	-2.7	-2.9	Transcription regulators	IPR002052	GO:0016251; general RNA polymerase II transcription factor activity

Validation of primary screening hits using an independent reporter

These filtering steps led to a final list of approximately 107 candidates that were selected for retesting. A new batch of dsRNA was re-synthesized and assayed with an independent reporter, derived from the promoter of the pathway target gene *socs36E* (Karsten et al. 2002) to exclude reporter-specific artefacts. This second independent JAK/STAT pathway reporter, *p4xsocsLuc* contains a 340 bp fragment with four predicted Stat92E binding sites (Karsten et al. 2002) driving the expression of firefly luciferase. Repeat assays were performed in quadruplicate and confirmed the identification of 24 dsRNAs that increase pathway activity (putative negative regulators, Table 7) and 67 dsRNAs that decrease pathway activity (putative positive regulators, Table 8) targeting a total of 90 predicted genes (see also Supplementary Table 1 for complete sequence and cytological information). This indicates a false-positive identification rate of approximately 15%.

Chromosomal clustering of novel JAK/STAT modulators

In order to obtain a first insight into the possible accumulation of these interacting candidates within the genome, the candidate loci were plotted on the *Drosophila* chromosomes (Figure 18, Supplementary Table 1 for cytological information). Although most modifiers are distributed throughout the genome, the X chromosome appears to be devoid of negative regulators (Table 9), a finding that may be linked to the role of the pathway in X:autosome ratio detection during *Drosophila* sex determination (Sefton et al. 2000).

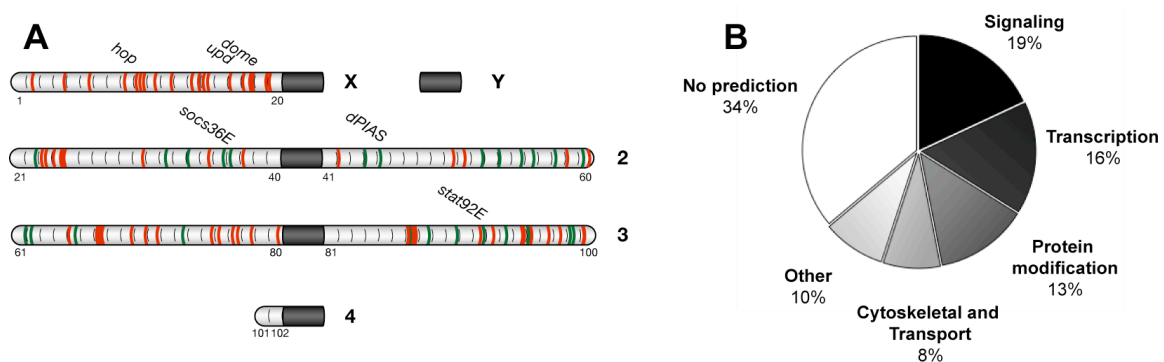


Figure 18. Analysis of JAK/STAT activity modulators. (A) Schematic representation of positive (red) and negative (green) regulator loci distributed within the *Drosophila* genome represented by schematic drawings of chromosomes. Cytology locations are indicated below the chromosomes. The chromosomal positions of canonical JAK/STAT signaling components are indicated. (B) Distribution of predicted gene functions. The identified JAK/STAT signaling modulators were grouped into the gene ontology groups indicated in the figure, and the percentage of these groups relative to all novel modulators was calculated.

Table 9. Expected and observed phenotype frequency.

Chromosome	No Genes*	%	Expected Phenotypes		Observed Phenotypes	
			Pos.	Neg.	Pos.	Neg.
X	2251	16.1	11	4	18	0
2L	2571	18.4	12	4	10	5
2R	2709	19.4	13	5	5	7
3L	2666	19.0	13	5	15	4
3R	3400	24.3	16	6	14	8
4	81	0.6	0	0	0	0
Unmapped	320	2.3	2	1	4	0

* Information from flybase.org according to release 4.2 of the Berkeley *Drosophila* Genome Project.

'Pos.' indicates positive and 'Neg.' negative regulators.

Gene ontology classification

The 90 predicted genes targeted by 91 dsRNAs were next classified according to their predicted functions by InterPro (Mulder et al. 2005) and gene ontology (GO, Drysdale et al. 2005, Harris et al. 2004). Manual inspection was used to order genes into functional groups (Table 10). Signaling factors, enzymes mediating posttranslational protein modifications and transcription factors cumulatively represent 48% of the genes identified (Figure 18B). Examples include *CG11501* encoding a putatively secreted, negative regulator of JAK/STAT signaling, *enok* encoding an acetyl-transferase, and the *tumour suppressor protein 101* gene, which encodes a ubiquitin conjugating enzyme. The molecular role of these genes in the regulation of JAK/STAT signaling remains to be determined. Seventy-four percent of the identified loci possess human homologs (E -values $<10^{-10}$, see Supplementary Table 2) compared to 62% for the whole proteome, of which 39% have been implicated in human disease (see Supplementary Table 3).

Comparison of the present data to previously published protein interaction studies

Another potentially interesting approach to look at the significance of the dataset is to look for an enrichment of previously described interactions between these genes or proteins. Parsing the dataset of this study against previously observed interactions gathered in Breitkreutz et al. 2003 reveals two previously published yeast two-hybrid interactions (CKIIalpha interacts with CKIIbeta and Par-1) and three genetic interactions

Table 10. Functional groups classified by InterPro prediction and GO.

Functional Group [†]	N*
Signaling factors	17
Transcription factors	14
Protein modification and metabolism	12
Cytoskeleton and transport	7
All others	9
Predicted proteins with identifiable functional group	59
Predicted proteins without identifiable functional group	31

Queries were performed with InterPro 8.0

[†] InterPro and GO results were classified into one of functionally related groups.

See Table 7 and Table 8 for complete list of genes, specific InterPro domains and GO assigned within each group.

* Number of proteins identified with InterPro domains and/or GO found in 90 translated gene sequences.

in the circuit of *stat92E*, *dome* and *hop* (Figure 19), which represents a four- to five-fold enrichment over the expected interactions (Supplementary Figure 1). The interaction of CKIIalpha with other proteins in the JAK/STAT interactor dataset, direct or indirect, is even more striking when a different interaction dataset is analyzed that includes in addition to the *Drosophila* protein-protein interactions those interactions that were observed in yeast and which were subsequently mapped to their *Drosophila* homologs (available at Flynet <http://www.jhubiomed.org/perl/flynet.pl>, Supplementary Figure 2). Although the knowledge of the genome represents a great challenge and screens produce many unproven interactions, this analysis shows the interaction of CKIIalpha with Dre4, CKIIbeta, CG10077 as well as CG3281 indirectly via CG8159, which had not been identified in the present JAK/STAT RNAi screen.

Overall, these comparisons reveal that the dataset obtained from the present genome-wide RNAi screen shows an accumulation of interactions also identified in other high-throughput datasets, thereby possibly revealing other regulators involved in the same cellular pathway, which may have been missed in the screen analysis as false-negatives.

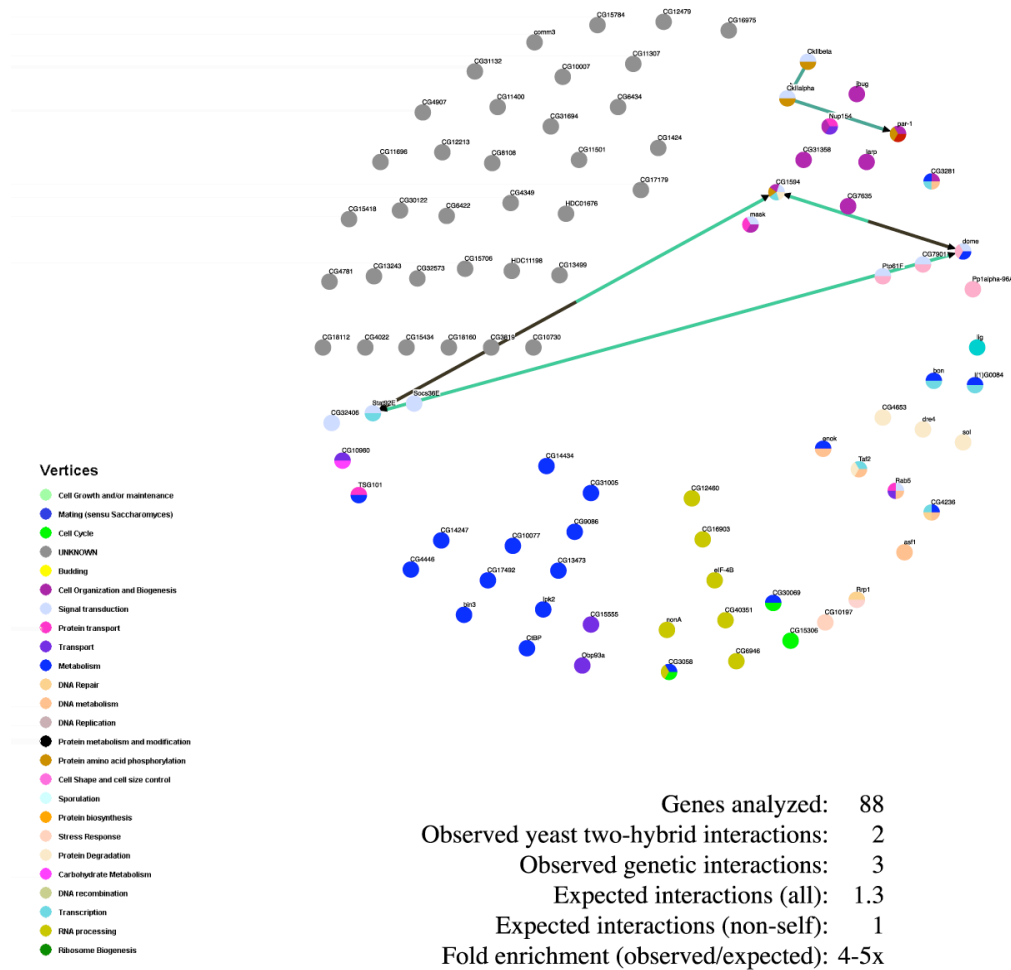


Figure 19. Previously published interactions between RNAi screen candidates. Fly genes identified in the genome-wide RNAi screen were uploaded into Osprey V1.2.0 and only interactions within these nodes were searched in the Fly GRID database. Circles represent identified genes, and their classification according to gene ontology is given in a pie chart with the colors explained in the key to the left.

Epistasis analysis

A genetic technique to characterize signaling molecules is the determination of their epistatic relationship with respect to defined pathway components. In order to map the putatively positively interacting candidate according to their position in the signaling cascade, the JAK/STAT pathway was activated in cells with different stimuli. dsRNAs of the positive regulators indicated in Figure 20 were then tested for their ability to suppress pathway activity under three conditions:

- (1) in *upd*-expressing cells ('Upd', screening conditions),
- (2) in cells treated with Upd-conditioned medium ('Upd-CM'), and

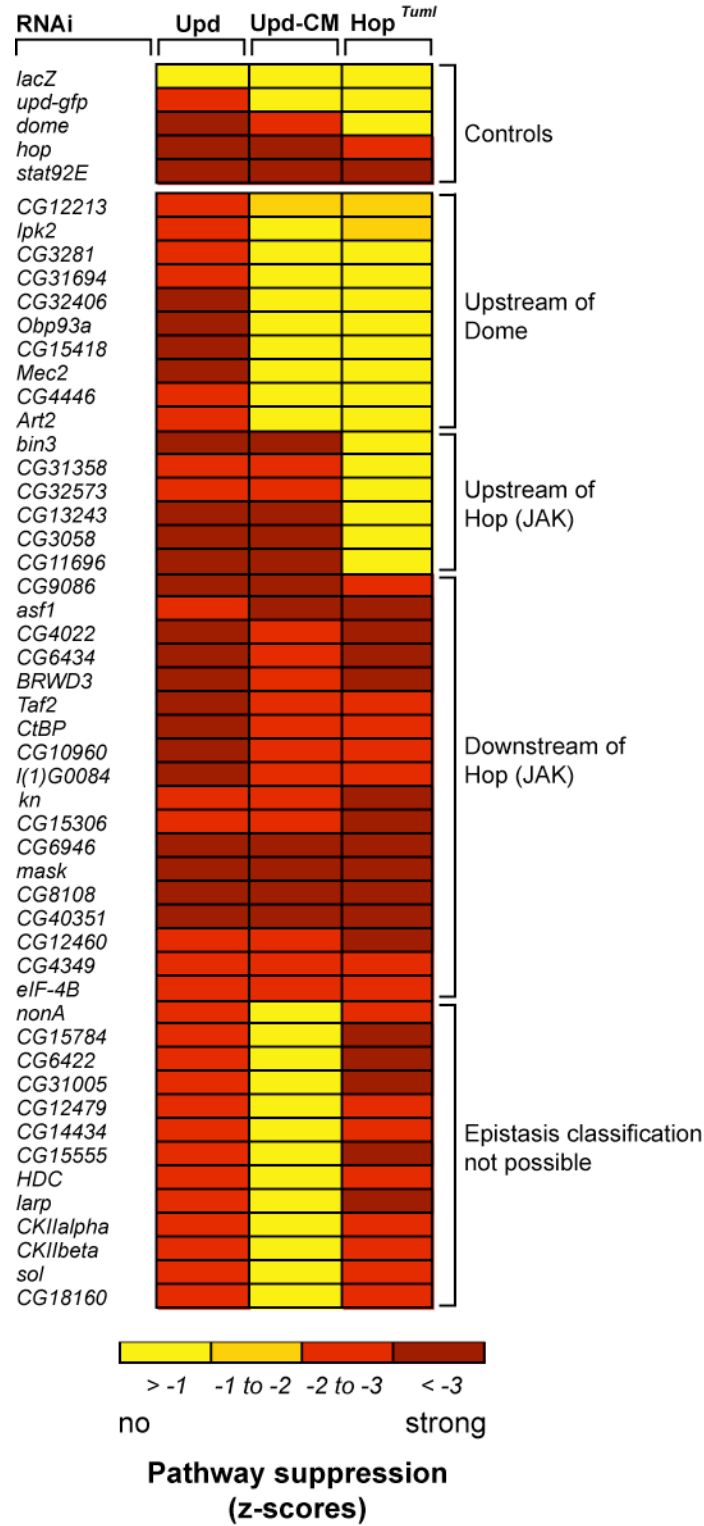


Figure 20. Classification of JAK/STAT modulators by pathway position. Epistasis analysis of the indicated positive pathway regulators showing interactions graded from none (yellow squares) to strong (red squares). Results shown have been obtained in independent experiments. Abbreviations used are: ectopic expression of *upd* (Upd), Upd conditioned medium (Upd-CM) and expression of a constitutively active JAK-allele (*Hop^{TumI}*, Harrison et al. 1995). Color-coding of z-scores is shown.

(3) in cells expressing the activated form of JAK, *hop^{Tuml}* (Hop^{Tuml}, Harrison et al. 1995, Luo et al. 1995).

In this way, dsRNA-silenced gene activities required upstream in the pathway can be characterized on the basis of their rescue by pathway activation further downstream. For example, although depletion of the IFN γ -related protein CG31694 results in downregulation of signaling stimulated by expression of *upd*, activation by Upd conditioned medium or *hop^{Tuml}* is unaffected (Figure 20). This suggests that CG31694 is required for the production and/or activity of the Upd ligand. Conversely, loss of pathway activity resulting from the knockdown of *kn* cannot be rescued by any form of pathway stimulus, implying a function downstream of JAK. Although this analysis suggests a role for multiple genes upstream of Dome, this classification is based on the lack of interaction observed under differing experimental conditions, and the molecular basis of these results remains to be established. Note that the *gfp* dsRNA was used to target the *upd-gfp* transgene and leads to a loss of pathway activity. *lacZ* dsRNA was used as a negative control.

Implementation of an interactive publicly accessible website

In order to facilitate the accessibility to the screening data presented in this study, a website was initiated and developed in collaboration with Thomas Horn and Michael Boutros at the German Cancer Research Center. The website can be publicly accessed at <http://www.dkfz.de/signaling/jak-pathway/>.

Figure 21 shows screenshots of examples how this website can be interactively used and how the data is linked to facilitate the analysis of interesting phenotypes. For example, an interactive map of the *Drosophila* chromosomes is available with the pathway interactors marked as red for positive and green for negative regulators linked to their respective chromosomal positions in the genome (Figure 21A). These chromosomal positions are linked with further information, which is accessible at a different page, to which the user is redirected. These further details include information about the dsRNA probe used as well as gene synonyms and information regarding potential off-target effects (OTEs).

Furthermore, the complete information regarding screening phenotypes as well as further annotation is available at this website (Figure 21B).

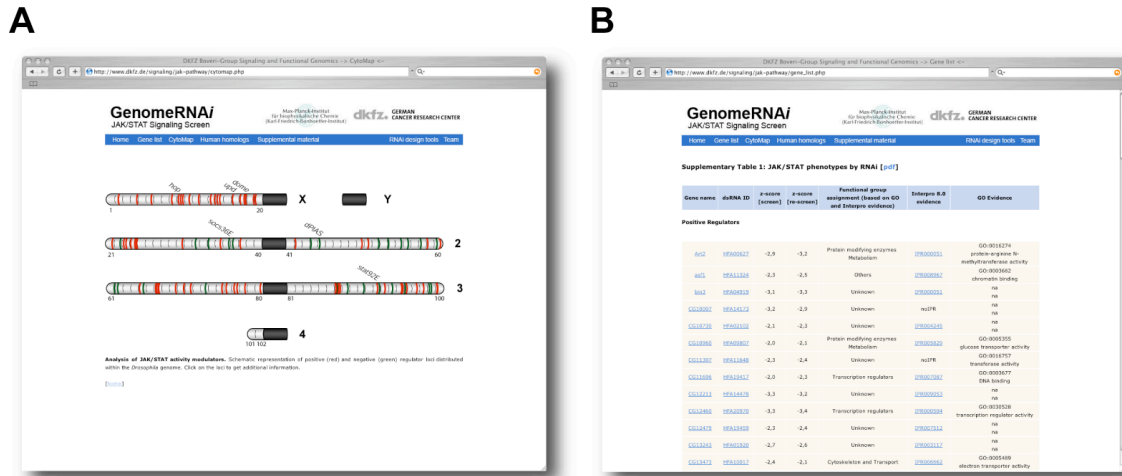


Figure 21. Implementation of a web-version to access screening data. An interactive version of the genome-wide RNAi screen data was developed with Thomas Horn at the German Cancer Research Center and is available at <http://www.dkfz.de/signaling/jak-pathway/>. (A) A screenshot of the interactive panel of JAK/STAT modulators mapped to the *Drosophila* chromosomes. Clicking on the red or green bands leads to more information regarding e.g. gene ontology information (B).

Functional conservation of orthologs

One of the fundamental tenets of the *Drosophila* genome-wide RNAi screen was the assumption that low levels of genetic redundancy within the fly genome will allow the identification of a greater range of factors with similar functions in higher organisms. At the same time, an interesting question is whether any of the factors identified in *Drosophila* would exert a specific effect on the combinatorial JAK/STAT pathways in mammalian systems (Table 1, Table 2). In other words: Are there factors important for Stat92E signal transduction in *Drosophila*, which are required specifically for the activity of one of the seven homologous mammalian STATs, but not for the others? To address this question and to validate the functional conservation of the regulators identified in the *Drosophila* RNAi screen, a comprehensive analysis of their homologs in the more complex human pathway (Figure 22, Table 1, Table 2) was therefore undertaken as a next step.

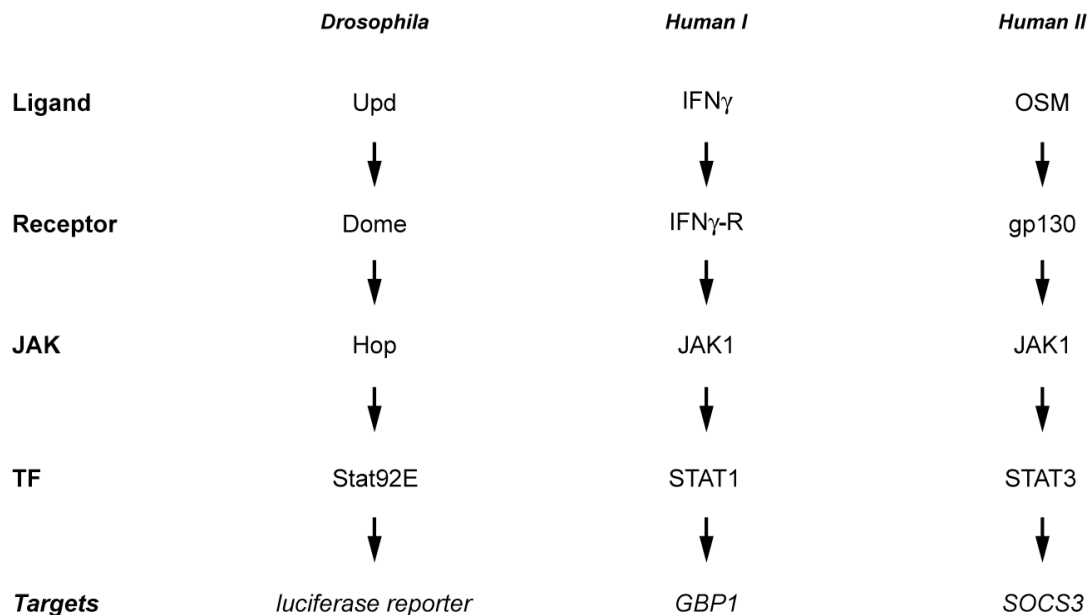


Figure 22. Simplified linear illustration of JAK/STAT signaling pathways. The models illustrate the comparison of the RNAi screening setup for *Drosophila* cells (*Drosophila*) as well as the assays to analyze the human JAK/STAT pathway for readouts of STAT1 (Human I) and STAT3 (Human II). TF indicates transcription factor.

Selection of homologs

The 90 previously identified *Drosophila* modulators were systematically screened for potential homologs using a variety of algorithms. HomoloGene, Inparanoid and best BLAST homologs were parsed from the Flight database (<http://flight.licr.org/>, Sims et al. 2006) for a total of 73 human candidate genes representing homologs of the interacting *Drosophila* genes (Supplementary Table 4). This list includes putative new regulators as well as controls such as the known pathway components STAT1, STAT3 and JAK1. Smart pool siRNAs (Dharmacon RNA Technologies) were ordered to target the transcript of each locus listed in Supplementary Table 4 with a pool of four independent dsRNAs (Supplementary Table 6), a technique designed to maximize the chance of effective knockdown while minimizing potential off-target effects (Birmingham et al. 2006).

Establishment of human JAK/STAT assays

Using HeLa cells as a representative human cancer cell line, assays for posttranslational modifications in human JAK/STAT pathways were first established. Following

stimulation with interferon γ (IFN γ), significant increases in the relative level of phospho-Y701 STAT1 (P-STAT1) are readily detected after 15 min while stimulation with the cytokine ligand Oncostatin M (OSM) is sufficient to cause the phosphorylation of Y705 of STAT3 (P-STAT3) (Figure 23A-D, Ehret et al. 2001). Knockdown of JAK1 reduces the proportion of phosphorylated STATs, and siRNA targeting the individual STAT transcripts specifically reduces both phospho- and non-phosphorylated forms (Figure 23C and D) showing the efficiency of protein depletion after 3 d.

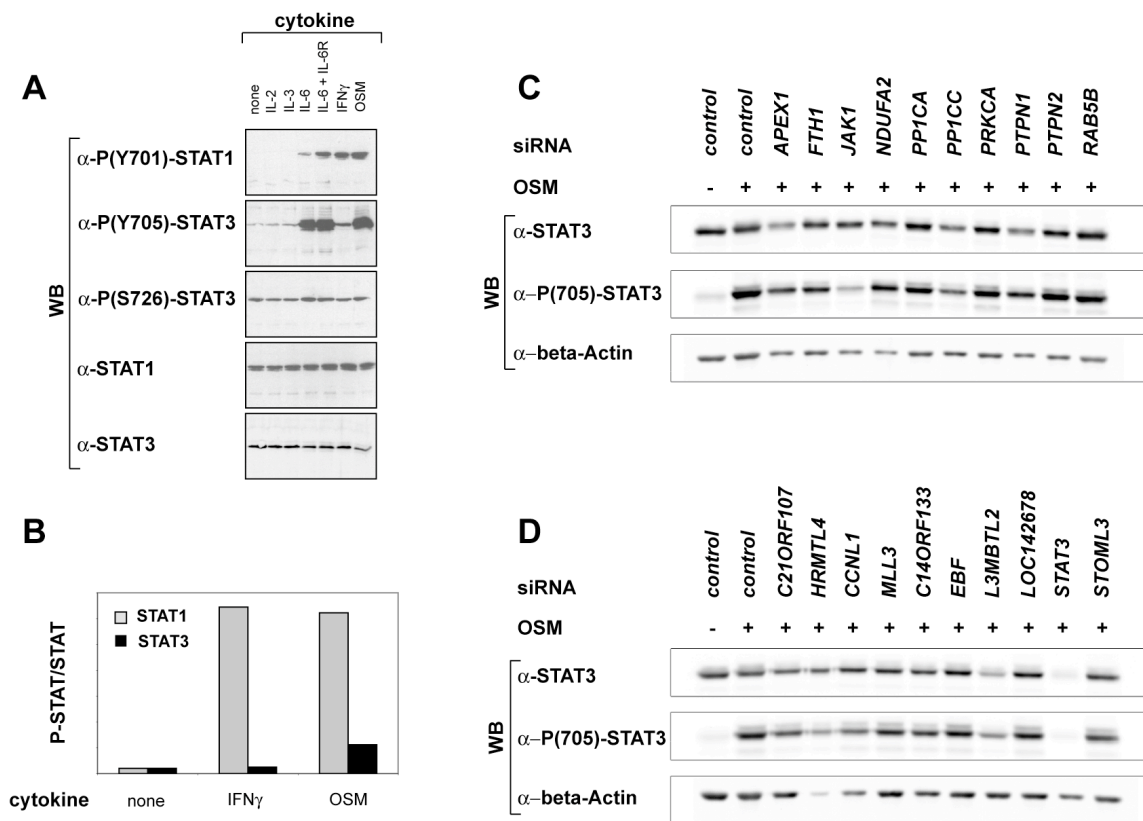


Figure 23. Analysis of cytokine induced posttranslational modifications on STAT proteins. (A) Initial trial to find cytokines capable of stimulating STAT activity in HeLa cells. Cytokine describes the kind of putative ligand applied to HeLa cells for 15 min before cells were lysed and subjected to Western blotting using the antibodies described on the left. (B) Quantification of the band intensity for the ratio of phosphorylated to unphosphorylated STAT proteins (P-STAT/STAT). (C) A sample from the Western blot based semi-quantitative screen. The membranes were first probed with α -STAT3 antibody, stripped and then incubated with α -phospho STAT3 along with α β -Actin antibody. Note that all lanes except for the first control lane were treated with OSM. siRNA against *JAK1* efficiently leads to the failure of STAT3 phosphorylation after cytokine stimulation. (D) Setup as in (C) but using different siRNAs. Note the efficiency of *STAT3* knockdown after 3 d.

Tyrosine phosphorylation of STATs is absolutely required for their transcriptional activity (Shuai et al. 1993). However, other posttranslational modifications have been identified that modulate the transcriptional potential of activated STAT molecules (Kramer et al. 2006, Wang et al. 2005, Yuan et al. 2005), and dominant-negative mutations of *Drosophila* Stat92E have been identified that are constitutively phosphorylated yet incapable of driving target gene transcription (Karsten et al. 2005). Therefore, as a direct measure of pathway activity, transcript levels of some endogenous STAT1 and STAT3 target genes following IFN γ and OSM stimulation were determined using a branched DNA assay in 96-well plates suitable for semi-high-throughput mRNA quantification (Collins et al. 1997, Gruber et al. 2005). In accordance with previous reports of STAT target genes (Ehret et al. 2001), multiple potential STAT1 and STAT3 target genes were tested (Figure 24A). This led to the identification of the robust induction of the STAT1 target gene *GBP1* and the STAT3 target gene *SOCS3* after 5 – 6 h of IFN γ or OSM stimulation, respectively (Figure 24A). In addition, up-regulation of *GBP1* is specific to IFN γ stimulation, while OSM stimulates high levels of *SOCS3*, with a minor up-regulation also being elicited by IFN γ (Figure 24A).

Knockdown of *JAK1* significantly reduces the expression of both target genes and, as would be expected for a *bona fide* target gene, knockdown of *STAT1* completely abolishes expression of *GBP1* and has no effect on *SOCS3* expression (Figure 24B+C). Similarly, knockdown of *STAT3* reduces the levels of OSM induced *SOCS3* expression to almost background levels while no change is detectable in IFN γ induced *GBP1* levels (Figure 24B+C). Intriguingly, knockdown of *STAT5A* and *STAT5B* appears to reveal the existence of compensatory mechanisms *in vivo*. While neither target gene is dependent on STAT5A/B, knockdown of these closely related STATs is sufficient to significantly increase the expression level of both target genes, with *GBP1* and *SOCS3* both being up-regulated in response to *STAT5A* or *STAT5B* knockdown. The complexity and potential redundancy of the human system must be considered when interpreting such results. For example, it has already been demonstrated that activated STAT5 can protect cells from IFN γ mediated apoptosis (Jensen et al. 2005) and that overexpression of STAT5 can counteract interferon signaling (Wellbrock et al. 2005).

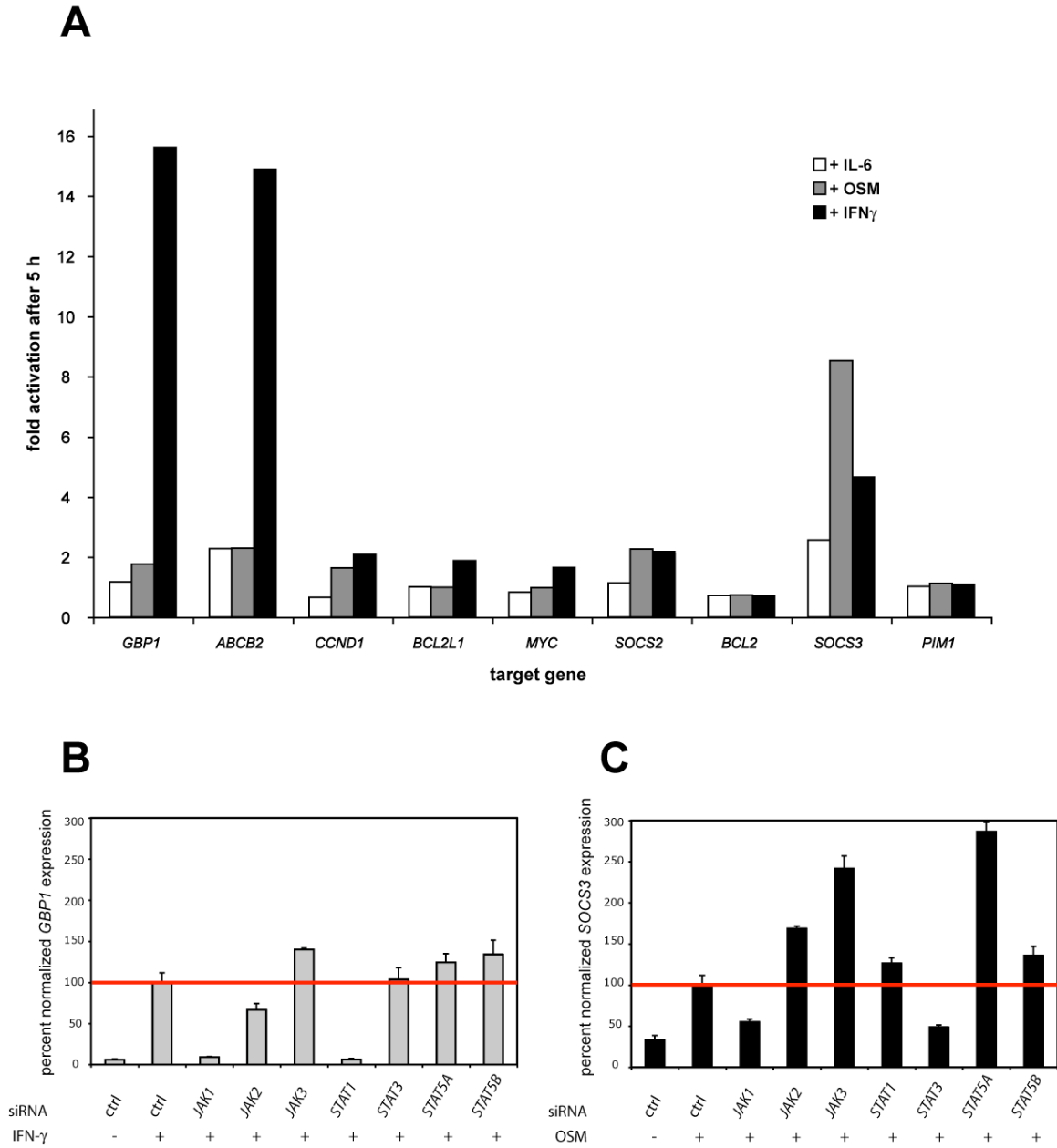


Figure 24. Transcriptional readout for human JAK/STAT pathways. Branched DNA assay technology was employed to measure relative mRNA levels in HeLa cells (Collins et al. 1997). (A) HeLa cells were treated with the indicated cytokines for 5 h followed by lysis and QuantiGene measurements. Fold activation is expressed as the level of the indicated target genes normalized to those treatments where cytokine was omitted. Note that for initial screening, each datapoint was only recorded once. Reproducibility is shown in the following panels. (B) Levels of *GBP1* relative to β -actin were assessed with QuantiGene measurements after siRNA treatments and IFN γ stimulation and normalized to non-targeting control siRNA (Dharmacon) treated cells, designated 100% (red line). (C) Levels of *SOCS3* relative to β -actin were assessed with QuantiGene measurements after siRNA treatments and OSM stimulation and normalized to non-targeting control siRNA treated cells, designated 100% (red line). Error bars in (B) and (C) represent standard deviations of 16 (ctrl), five (STAT1 and STAT3) or three (all other) biological replicates.

A small-scale screen for human JAK/STAT pathway modulators

Using these assays, all 73 siRNA pools were tested for their influence on the relative levels of phosphorylated STAT. In addition to changes in the overall level of phosphorylation of both STATs, specific changes in either STAT1 or STAT3 are also observed (Figure 23, Figure 25; given the difficulties in accurately quantifying Western blot data, the effect of each siRNA on the relative levels of STAT1 and STAT3 was classified simply as increase, decrease or no change) as well as siRNAs that result in the loss of all proteins (Supplementary Table 5).

Next, the effect of the 73 siRNA pools on cytokine induced target gene expression was assayed for up- or down-regulation compared to siRNA non-targeting controls. The heatmap in Figure 25 shows a cluster analysis of the changes in phosphorylation status and the quantitative target gene analysis for the total of 30 human *bona fide* JAK/STAT pathway interacting factors. These regulators include a number of pathway modulators not previously implicated in mammalian JAK/STAT signaling processes. For example, two homologs of *Drosophila mask* showed a specific phenotype in these assays. *mask* had been originally selected as a positive regulator of *Drosophila* Stat92E in this present genome-wide RNAi screen and encodes a large Ankyrin repeat and KH domain-containing protein, which has been implicated in receptor tyrosine kinase signaling and cellular proliferation (Smith et al. 2002, Tseng and Hariharan 2002). Two homologs of *dMask*, the Myeloid/lymphoid or mixed-lineage leukemia protein 3 (MLL3) and ANKHD1 both act as regulators of the human JAK/STAT pathway (Figure 25). MLL3 appears to be required as a positive regulator solely for STAT1 dependent transcription and has a negative regulatory effect on STAT3 signaling, whereas ANKHD1 acts as a negative regulator of STAT1 with levels of both P-STAT1 and *GBP1* increasing following *ANKHD1* knockdown. Conversely, levels of the STAT3 target gene *SOCS3* are strongly reduced, an effect that occurs despite the partial STAT1 dependency of *SOCS3* (Figure 24A). Although neither ANKHD1 nor MLL3 have been studied in detail, an involvement in JAK/STAT signaling is of potential significance for further analysis given the implication of MLL3 in the development of mixed-lineage leukemia.

Human gene, <i>Drosophila</i> gene	IFN γ stimulation									OSM stimulation					
	<i>Drosophila</i> z-score	fold normalized <i>GBP1</i> expression levels								fold normalized <i>SOCS3</i> expression levels					
		pooled siRNAs			individual siRNAs				P-STAT1	screen	individual siRNAs				P-STAT3
		screen 1	screen 2	retest	siRNA 1	siRNA 2	siRNA 3	siRNA 4			siRNA 1	siRNA 2	siRNA 3	siRNA 4	
<i>JAK1, hop</i>	-5.7	0.1	0.1	0.1	0.2	0.1	0.2	0.1	-	0.4	0.4	0.8	0.8	0.5	-
<i>STAT1, Stat92E</i>	-5.0	0.1	0.0	na	0.1	0.0	0.2	0.1	-	0.6	2.2	1.3	2.4	1.9	-
<i>STAT3, Stat92E</i>	-5.0	1.1	0.9	na	1.2	1.7	2.1	2.1	-	0.3	0.6	0.3	0.3	0.5	-
<i>MLL3, mask</i>	-2.3	0.7	0.8	0.4	0.4	0.8	0.6	0.9	-	1.2	1.4	1.0	3.4	2.8	-
<i>PHF10, I(1)G0084</i>	-2.1	0.9	1.0	0.6	0.9	0.7	0.7	0.8	-	0.4	1.8	0.9	1.6	0.7	-
<i>TPRT, CG31005</i>	-2.3	0.5	0.3	0.3	0.2	0.6	0.6	0.9	-	0.4	0.9	2.5	1.2	1.9	-
<i>BRODL, CG31132</i>	-2.8	1.4	0.5	0.5	1.2	0.5	0.6	0.4	-	0.9	2.7	0.6	1.7	2.1	-
<i>C21ORF107, CG31132</i>	-2.8	0.7	0.3	0.4	0.4	0.8	1.8	0.8	-	0.7	0.7	2.0	1.7	2.7	-
<i>CTBP2, CtBP</i>	-2.9	0.7	1.1	0.7	0.3	0.9	0.9	1.2	-	0.5	0.4	1.8	0.9	3.0	-
<i>MYST3, enok</i>	3.0	1.0	2.5	1.8	1.3	1.2	2.4	1.9	+	0.8	2.1	1.2	1.2	2.2	-
<i>TSG101, TSG101</i>	3.1	1.7	1.1	1.0	1.3	1.1	1.1	1.9	-	1.3	2.7	2.8	2.5	3.7	-
<i>PP1CA, Pp1alpha-96A</i>	3.0	5.6	3.2	2.6	1.9	0.9	1.2	0.3	-	0.6	2.6	1.3	3.0	1.0	-
<i>RAB5A, Rab5</i>	2.1	1.7	1.7	1.6	0.4	2.5	0.9	2.0	-	2.1	1.0	4.4	1.8	3.2	+
<i>PP1CC, Pp1alpha-96A</i>	3.0	3.0	1.6	1.7	0.4	1.1	0.9	1.6	-	0.4	1.3	1.1	1.4	1.7	-
<i>TRIM33, bon</i>	5.6	2.6	1.4	1.0	1.5	1.0	0.6	1.5	-	0.4	1.6	2.2	0.8	1.7	-
<i>DKFZP667B0120, kn</i>	-2.4	1.7	1.2	1.1	2.0	0.9	1.6	1.3	-	1.5	3.5	2.8	2.5	3.4	+
<i>FTH1, CG4349</i>	-4.1	2.9	1.0	1.0	2.5	1.7	0.9	1.3	-	0.3	2.0	3.2	2.7	2.9	-
<i>ANKHD1, mask</i>	-2.3	4.0	2.9	1.5	0.8	1.5	1.4	1.1	+	0.3	2.8	3.6	0.3	0.3	-
<i>HRMT1L4, Art2</i>	-2.9	6.2	2.2	2.4	0.8	1.2	2.3	1.0	-	0.8	2.0	1.7	1.1	2.9	-
<i>NDUFA2, CG15434</i>	-2.5	7.8	5.4	3.0	4.6	0.7	1.0	1.7	+	2.5	4.6	2.1	2.2	3.7	-
<i>IFRD2, CG31694</i>	-2.8	2.0	2.0	1.9	0.5	0.4	1.9	1.4	+	1.3	2.3	2.6	4.3	2.5	+
<i>DDX5, CG10077</i>	-3.2	5.0	5.1	3.1	2.0	1.1	0.7	0.8	-	0.6	0.6	2.9	1.0	1.7	-
<i>LOC142678, CG17492</i>	2.5	1.6	1.1	0.5	1.1	1.1	1.7	0.9	-	3.3	3.2	5.1	4.3	3.0	-
<i>C14ORF133, CG18112</i>	2.1	1.5	2.1	1.3	0.8	1.0	2.1	1.0	+	1.2	2.0	1.7	3.7	0.8	+
<i>NUP155, Nup154</i>	2.9	2.8	2.6	1.6	1.5	0.8	1.1	0.5	-	1.0	0.7	2.2	2.7	1.5	-
<i>RBBP4, Caf1</i>	3.0	1.2	1.3	1.6	0.8	2.0	0.9	0.8	-	1.4	1.7	1.8	1.1	2.6	-
<i>ENDOGL1, CG3819</i>	-2.3	1.8	1.3	1.5	0.7	1.4	0.6	1.0	-	0.3	0.6	2.2	1.7	1.9	-
<i>CAPN3, sol</i>	-2.5	3.3	2.7	1.6	0.9	1.7	0.9	0.8	+	1.6	3.4	2.1	1.8	2.3	-
<i>PPP2R5D, PP2A-B'</i>	2.6	2.3	2.4	1.3	0.7	1.5	0.7	0.6	-	1.6	2.7	2.5	1.9	1.7	-
<i>CHRNA7, HDC01676</i>	-2.3	2.6	1.9	0.8	1.2	0.7	1.1	1.1	-	0.6	2.7	2.7	2.2	1.2	-

Figure 25. Functionally conserved JAK/STAT signaling modulators. Heatmap of novel human JAK/STAT pathway regulating genes. The original *Drosophila* interactions (column 1) are expressed as z-scores. Further shown are the fold changes in the expression levels of human STAT1 and STAT3 target genes normalized to β -actin levels and the levels of phosphorylated human STAT1 and STAT3 proteins ('P-STAT1', 'P-STAT3'). In all columns, black represents a decrease (fold change < 0.9), white an increase (fold change > 1.1) and grey no change in activity (fold change = 1.0).

Assessing the specificity of siRNA induced human JAK/STAT phenotypes

It is possible that some of the siRNA-mediated effects observed are due to aberrant cellular proliferation or viability. For example, Fedorov and colleagues (Fedorov et al. 2006) have shown that 30% of the siRNAs designed for an exogenous target mRNA in their experimental system led to a non-specific viability phenotype, possibly due to off-target effects. In order to address this possibility, the growth rate of HeLa cells transfected with individual siRNAs was analyzed by quantifying both the levels of

endogenous β -actin mRNA levels and the amount of cells present after siRNA treatment (Figure 26, Supplementary Table 5). Although most treatments did not show major effects on the proliferation of HeLa cells, the differences in β -actin mRNA levels between treatments were not negligible (Supplementary Table 5). These differences necessitated the normalization approach chosen for the readout of target gene activities of *GBP1* and *SOCS3* as shown in Figure 24 and Figure 25. Note that this approach of normalizing target gene activity to β -actin levels is prone to artefacts as outlined in Figure 17. siRNAs knocking down e.g. *STOML3*, *EIF4B*, *POLR2A*, *CTBP1*, *AKT1* and *CAPN3* had severe effects on cellular proliferation, survival or morphology of the HeLa cells (Supplementary Table 5). These siRNAs, although reducing overall proliferation, were however not excluded from further analysis with the rationale that these could still specifically affect JAK/STAT signaling given the role of the pathway in cellular proliferation as well as the *STAT3* knockdown phenotype shown in Figure 26 and Supplementary Table 5. Although phosphorylation-activated STATs in the absence of cytokine could not be detected on Western Blots (Figure 23A), it is possible that even non-detectable low-level stimulation is necessary for cell survival.

As a more specific readout of potential off-target effects, the tested siRNA pools were next divided into their components and tested individually for their effect on the JAK/STAT reporter system. This way, a ‘poisoned’ siRNA pool can be identified which may contain mostly ineffective siRNAs with only one siRNA exerting an offtarget-effect leading to an artefactual apparent JAK/STAT pseudo-phenotype. The four siRNAs present in each pool were therefore tested individually for their effect on the transcriptional readout after IFN γ or OSM stimulation, respectively (Figure 25).

Three scenarios were observed, where the phenotype of the pooled siRNAs was repeated by

- i) all or three,
- ii) by two
- iii) or by only one

individual siRNA. In every case, only the JAK/STAT pathway reporter and not the efficiency of the siRNA to knock down target mRNA levels was analyzed. Therefore, a

lack of interaction may stem from experimental failure, lack of penetrance, inefficiency of knockdown or from a true false-positive phenotype for the pools. For simplicity, cases with at least three confirmed individual siRNAs (i.e. an individual phenotype similar to the pooled phenotype with a fold change either at least higher than 1.1 for siRNA targeting negative regulators or lower than 0.9 for siRNAs targeting positive regulators) are designated ‘confident’, cases with two confirmed individual siRNAs ‘less confident’ and cases with only one confirmed siRNA designated ‘potentially poisoned’. Using this classification, 53% (16 out of 30) of the pools are confident, 30% less confident and 17% potentially poisoned hits.

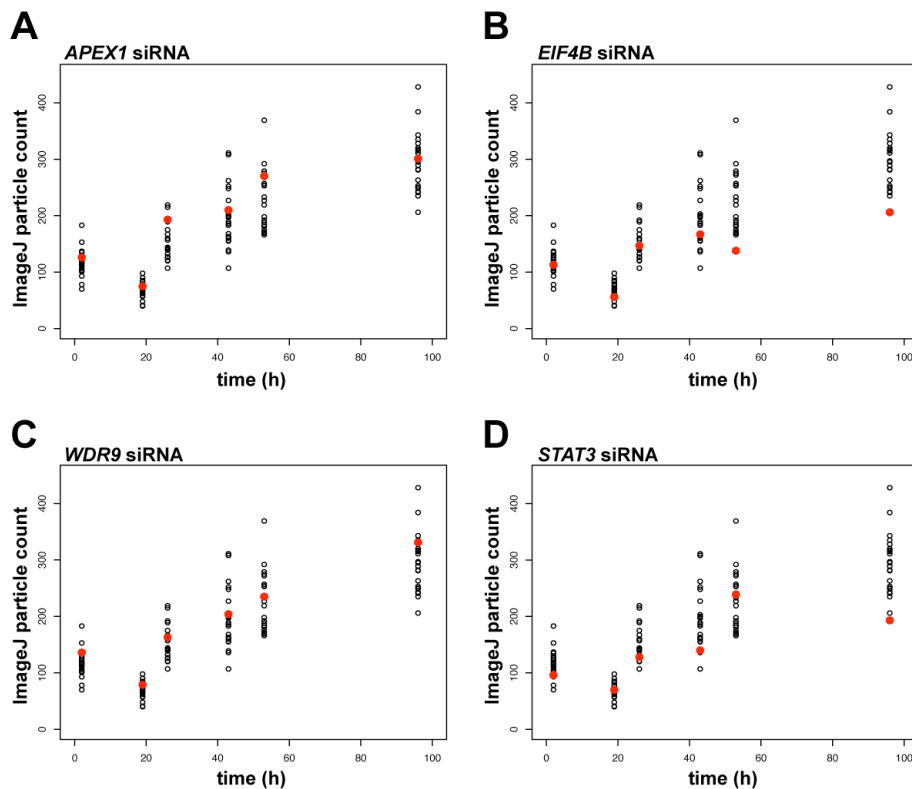


Figure 26. Effect of siRNAs on HeLa cell growth. Examples of the growth curve analysis after transfection of the siRNAs are shown. Cells were imaged at the time-points indicated, and the images were analyzed with an ImageJ macro to automatically count particles correlating to the number of cells. Cells were transfected with siRNAs approximately at time-point 18 h. Black dots represent wells with non-targeting control wells and are the same for all panels because all samples were present on a single 96-well plate. Red dots represent data-points for the following siRNA treatments: (A) No growth phenotype is observed for siRNA targeting *APEX1* (well B3), whereas in (B) a severe growth reduction is visible for knockdown of *EIF4B* (well A10). (C) Knockdown of *WDR9* (also known as *C21ORF107* and *BRWD1*, well G3) does not reduce the cell growth. (D) siRNA against *STAT3* (well G11) leads to less cell growth. Refer to Supplementary Table 5 for the complete dataset.

The validation of a significant proportion of the computationally determined hits from the genome-wide RNAi screen in *Drosophila* and human cell culture experiments indicates a low number of false-positives. However, this possibility cannot be entirely ruled out and analysis in immortalized cells growing in artificial medium along with a phenotypic readout of a transcriptional reporter following RNAi-mediated reduction in gene activity can only be a first step in the identification of possible interactions. These initial experiments cannot substitute the analysis and thorough characterization of gene function in the *in vivo* context of a developing organism using a genetic mutant. Therefore, as a next step towards revealing novel molecular mechanism in JAK/STAT signaling, examples of novel JAK/STAT signaling modulators were tested to confirm their regulatory role *in vivo* – from gene identification to gene analysis.

Ptp61F negatively regulates JAK/STAT signaling

As a first example Ptp61F was analyzed, a protein tyrosine phosphatase whose depletion led to an increase in JAK/STAT signaling activity and whose vertebrate homolog is the prototypic non-receptor protein tyrosine phosphatase PTP1B (also known as PTPN1, Aoki and Matsuda 2000).

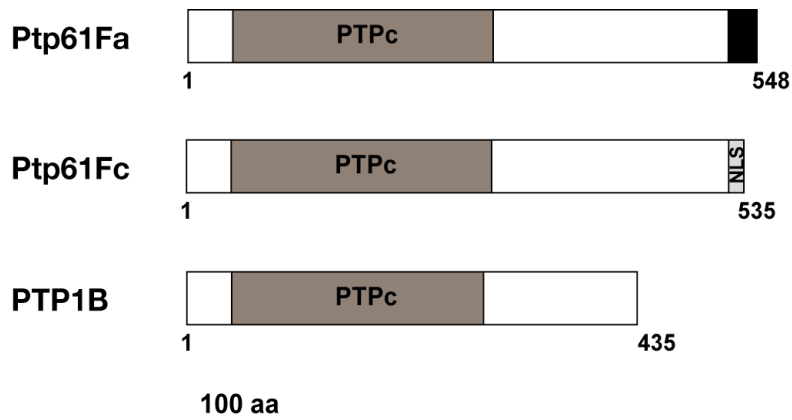


Figure 27. Ptp61F is an evolutionarily conserved protein. Shown are the comparisons between the two isoforms of *Drosophila* Ptp61F and the human homolog PTP1B. Note that Ptp61Fa and Ptp61Fc are identical except for the unique C-terminal parts marked in black and light grey, so that only Ptp61Fc contains a nuclear localization signal (NLS). PTPc is the catalytical domain of these protein tyrosine phosphatases (dark grey). Scale bar indicates the length of 100 amino acids. Information from NCBI and McLaughlin and Dixon 1993.

To perform epistasis analysis, known pathway components were removed and the effect of simultaneously targeting *ptp61F* was tested. Double RNAi against *ptp61F* together with *lacZ*, *dome* or *hop* results in pathway stimulation (Figure 28B). However, simultaneous removal of *ptp61F* and *stat92E* is sufficient to prevent signaling. Loss of this phosphatase therefore results in the stimulation of Stat92E activity even in the absence of upstream components, indicating that Ptp61F negatively regulates the pathway downstream of JAK – possibly via direct interaction with Stat92E.

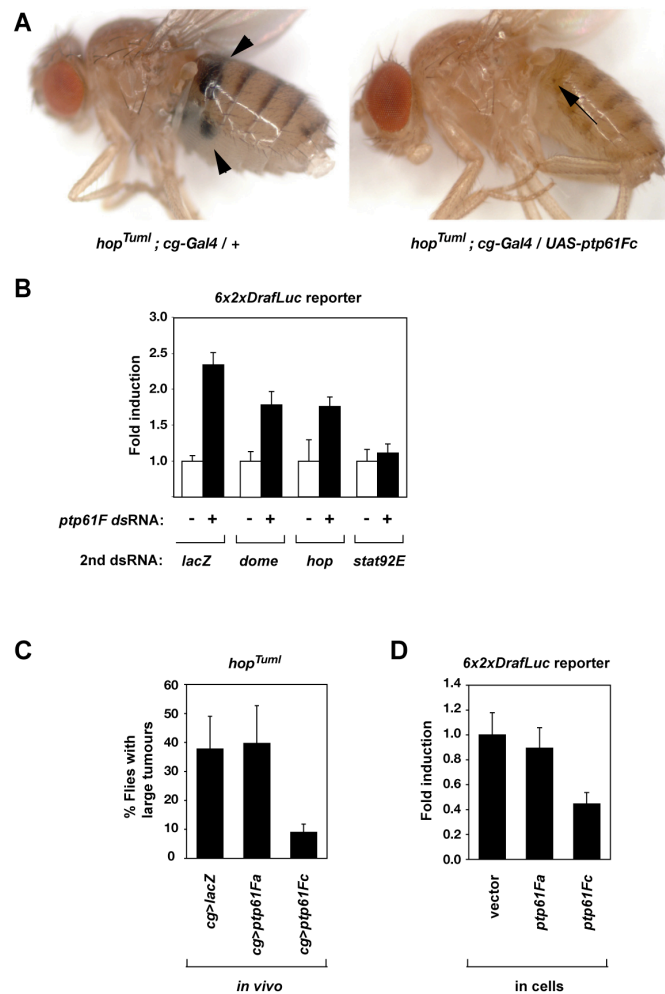


Figure 28. Molecular function of Ptp61F. (A) Ptp61F is a tumor suppressor *in vivo*. Hemocyte-specific misexpression of *ptp61F* can protect *hop^{Tum1}* mutants from melanotic tumor formation. Compare large black tumors in controls (arrowheads, left) with small tumors present in a *ptp61F*-expressing individual (right). (B) Epistasis analysis of *ptp61F* dsRNA in cell culture indicates that it acts downstream of Hop and upstream or parallel to Stat92E. (C) Quantitative analysis of large tumor formation in *hop^{Tum1}* mutants expressing cytoplasmic Ptp61Fa and nuclear Ptp61Fc showing specificity of rescue for the nuclear isoform. Error bars represent standard deviations of three or four independently tested transgenic lines. (D) Specificity of rescue by overexpression of the nuclear isoform in tissue culture based reporter assays. Error bars represent standard deviations of eight parallel cell culture experiments.

Table 11. Genetic interactions of *ptp61F* with *hop^{Tuml}*.

Exp	Genotype	Insert	Tumors (%)			n	z-score	
			None	Small	Large			
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-EGFP</i>	5a.2	19.9	35.1	45.0	151	1.2	(*)
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-EGFP</i>	6a.3	41.0	33.3	25.7	451	-1.7	(*)
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-lacZ</i>	BG4-1-2	25.8	26.4	47.8	341	0.4	(*)
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fa</i>	1b.2	46.5	27.7	25.7	101	-2.5	
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fa</i>	1b.2	46.5	29.1	24.3	230	-2.5	
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fa</i>	1a.3	22.6	28.8	48.6	177	0.8	
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fa</i>	1a.3	19.6	24.4	56.0	168	1.2	
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fa</i>	3a.3	35.8	28.5	35.8	165	-1.0	
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fa</i>	7a.3	16.4	36.1	47.5	61	1.6	
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fc</i>	1a.1	68.2	21.4	10.4	280	-5.4	
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fc</i>	2a.4	56.1	30.6	13.3	255	-3.8	
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fc</i>	2a.4	52.3	40.7	7.0	344	-3.2	
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fc</i>	2b.3	59.4	33.8	6.8	234	-4.2	
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-ptp61Fc</i>	2b.3	63.3	29.3	7.3	300	-4.7	
II	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-dPIAS-GFP</i>	26b.3	67.0	27.4	5.7	106	-5.2	(1)
I	<i>y,w,hop^{Tuml}/+;cg-Gal4/UAS-dPIAS-GFP</i>	26b.3	63.1	33.6	3.3	122	-4.7	(1)

Values shown represent percentage of 0-24 h old female flies containing no, small or large tumors visible in abdomen or thorax. Table shows results from two independent experiments (first column) undertaken under identical conditions. 'n' indicates the number of flies analyzed.

(*) 'wildtype' results used to calculate z-scores

Previously published interaction:

(1) Betz et al. 2001

The next question addressed was whether Ptp61F also interferes with JAK/STAT signaling *in vivo* by using the *cg-Gal4* transgene to misexpress *ptp61F* in blood cells of *hop^{Tuml}* mutant flies (Drysdale et al. 2005, Harrison et al. 1995). Transgenic flies expressing *EGFP* or β -galactosidase were used as negative controls and misexpression of *Drosophila dPIAS-GFP* served as a positive control as previously described (Table 11, Betz et al. 2001). Crosses were incubated at 25°C and adult females carrying the *hop^{Tuml}* chromosome were scored within 24 h of eclosion for the presence of tumors classified as small (one or two small melanotic spots as shown in Figure 28A (right)) or large (large melanized growths or more than three small spots as in Figure 28A (left)). Misexpression of *ptp61Fc* in a *hop^{Tuml}* mutant background resulted in suppression of melanotic tumor

formation, with the average frequency of large tumors reduced by approximately four-fold (Figure 28C), an effect also observed after the misexpression of *dPIAS* (Table 11, Betz et al. 2001).

Alternative splicing of *ptp61F* gives rise to nuclear and cytoplasmic protein forms that both contain the same phosphatase domain (Figure 27, McLaughlin and Dixon 1993). However, the tumor suppressor phenotype is only observed with nuclear Ptp61Fc (Figure 28C). This effect can be reproduced by overexpression of the nuclearly targeted isoform encoded by *ptp61Fc* in cell culture (Figure 28D). These results are consistent with the identification of Ptp61F as a negative regulator of pathway activity and suggest that it may function by targeting phosphorylated, nuclearly localized Stat92E for deactivation.

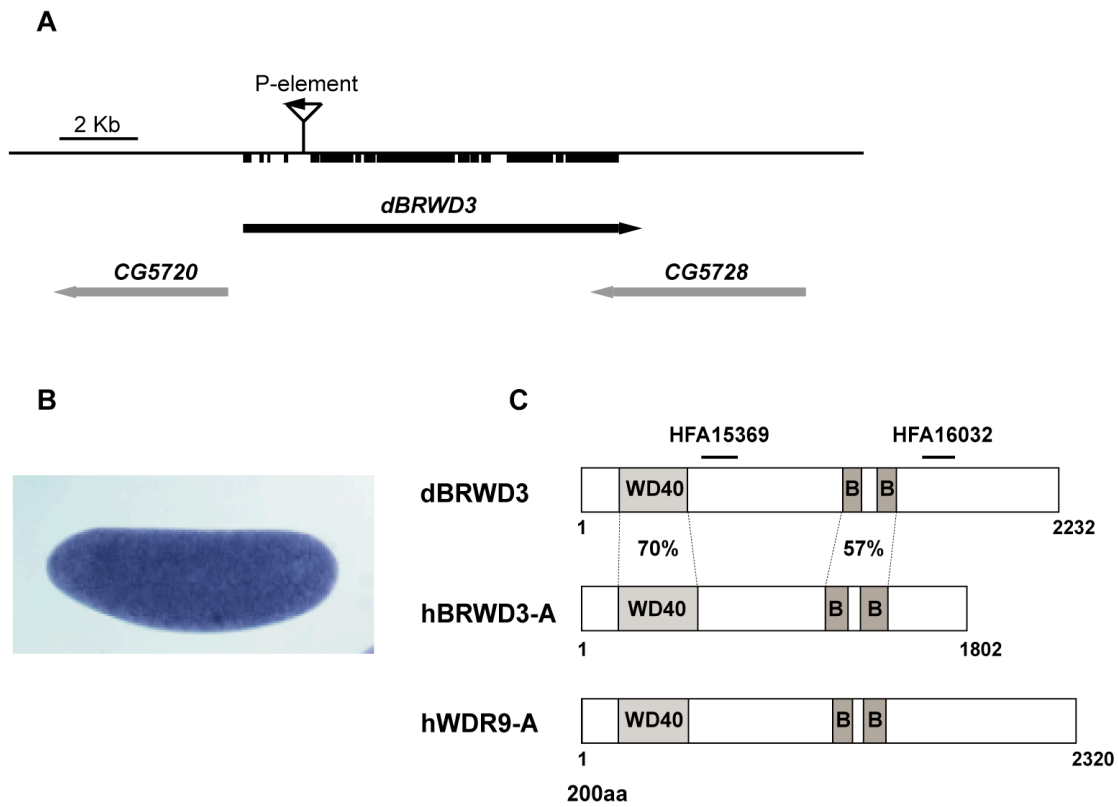


Figure 29. BRWD3 is an evolutionarily conserved protein. (A) Genomic organization at the *dBRWD3* locus. Exon and intron structure is shown, and arrowheads indicate the directionality of transcription. Scale bar indicates the length of 2 kb DNA. Information was obtained from <http://flybase.bio.indiana.edu/>. (B) Endogenous expression of *dBRWD3* in early *Drosophila* embryos (blue) indicates that *dBRWD3* transcript is provided maternally. (C) Domain structure and sequence similarity between *Drosophila* and human BRWD3 proteins as well as another homolog WDR9. Percentages show the similarity in the amino acid sequence and are similar for the comparison of dBRWD3 and hBRWD3 as well as dBRWD3 and hWDR9. Regions targeted by two independent dsRNAs (probe IDs HFA15369 and HFA16032) independently recovered in the *Drosophila* screen are shown. Scale bar indicates the length of 200 amino acids.

BRWD3 is a novel component of the JAK/STAT pathway

A strong positive regulator identified in the genome-wide RNAi screen is *CG31132*, which encodes a 2,232-amino-acid WD40- and bromo-domain-containing protein homologous to human BRWD3 and WDR9 (Figure 29). In the screen, a strong reduction of pathway activity was observed for two independent dsRNAs that target different regions of the transcript shown in Figure 29C (probe IDs HFA15369 and HFA16032). hBRWD3 and WDR9 show good homology to the *Drosophila* protein both between the domains themselves and overall. However, hBRWD3 is the best reciprocal BLAST hit (*E*-value of 0 compared to an *E*-value of 2E-173 for WDR9) and was therefore considered as the closest homolog to CG31132, hereafter referred to as dBRWD3. *hBRWD3* is a functionally uncharacterized locus recently identified at the breakpoint of a t(X;11)(q13;q23) translocation derived from a B-cell chronic lymphocytic leukemia (B-CLL) patient, and *hBRWD3* transcripts have been shown to be downregulated in the majority of B-CLL patients compared to sorted B cells of healthy donors (Kalla et al. 2005).

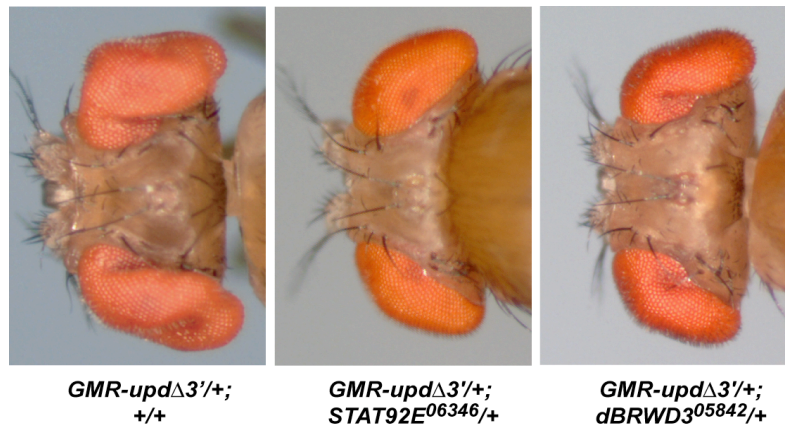


Figure 30. dBRWD3 functions as a JAK/STAT pathway component. Heads of adult *Drosophilae* heterozygous for the *GMR-updΔ3'* transgene crossed to wild type (left panel), *stat92E* (middle panel) and *dBRWD3* (right panel) mutants. Note the reduction in eye overgrowth after removal of pathway components.

In order to test whether *BRWD3* is also involved *in vivo* in JAK/STAT regulated process, flies mutant for *dBRWD3* were generated. A mutagenic P-element insertion termed *l(3)05842* and generated by the *Drosophila* genome project (Spradling et al. 1999) was identified in the fourth intron of *dBRWD3* (Figure 29A, Drysdale et al. 2005) and was

obtained from the Bloomington Stock Center (University of Indiana). The P-element insertion *l(3)05842* (hereafter termed *dBIRD3*⁰⁵⁸⁴²) is homozygous lethal (animals die in the larval stage) and fails to complement the *Df(3R)crb87-4* and *Df(3R)crb87-5* deficiencies spanning the *BRWD3* locus. Twenty-three independent stocks in which the *ry*⁺ marker present in the *p{ry⁺,PZ}* insertion had been lost following a cross to a transposase source were established. Of these, seven were viable revertants (30%), two are semi-lethal with occasional escapers and the remainder were lethal. Remobilization of the mutagenic P-element indicates that the transposon insertion is responsible for late embryonic lethality.

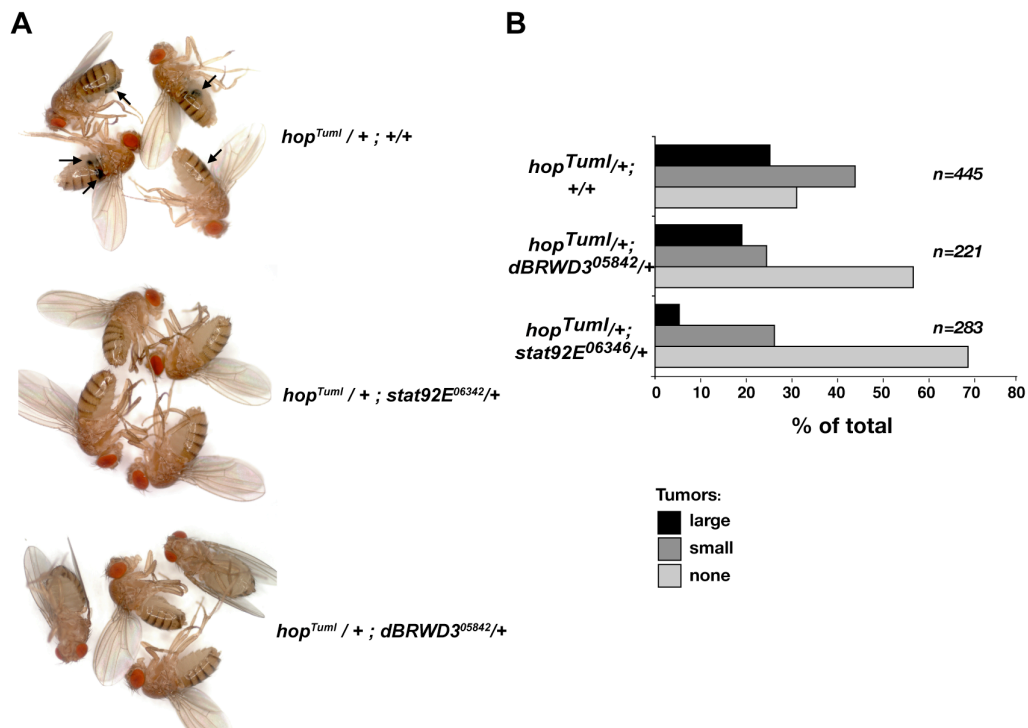


Figure 31. Rescue of *hop^{TumI}*-induced tumor formation by *dBIRD3* dosage reduction. (A) *hop^{TumI} / + ; +/+* females (top) frequently contain large black melanotic tumors (arrows). The size and frequency of *hop^{TumI}*-induced tumor formation is significantly decreased in *stat92E⁰⁶³⁴²* and *dBIRD3⁰⁵⁸⁴²* heterozygous backgrounds. Flies were grown in parallel independent experiments at 25°C and are representative examples of the individuals recovered. (B) Quantification of tumor formation. *hop^{TumI} / + ; stat92E⁰⁶³⁴² / +* heterozygotes which lack one copy of *stat92E* (bottom) contain fewer and smaller tumors. *hop^{TumI} / + ; dBIRD3⁰⁵⁸⁴² / +* (middle) also contain fewer and smaller tumors. ‘n’ indicates the number of flies counted.

Genetic interactions between *dBIRD3* and JAK/STAT signaling were tested by crossing flies bearing the *dBIRD3*⁰⁵⁸⁴² allele to *GMR-updΔ3* flies (Bach et al. 2003). The *GMR-updΔ3* transgene ectopically misexpresses *upd* during eye development, resulting in

cellular overproliferation and an enlarged adult eye. Genetic interaction with *GMR-updΔ3*' was undertaken as described in Bach et al. 2003 using OreR and *stat92E*⁰⁶³⁴⁶ as negative and positive controls, respectively. Removal of one copy of *stat92E* significantly suppresses eye overgrowth due to a reduction in the potency of JAK/STAT signaling (Bach et al. 2003). Removal of a single copy of *dBRWD3* was also able to suppress the *GMR-updΔ3*' phenotype in multiple independent experiments in a majority of individuals of the appropriate genotype, as expected for a positive regulator of JAK/STAT signaling (Figure 30). A chromosomal deficiency removing the region has also previously been independently identified as a suppressor of *GMR-updΔ3*' (Bach et al. 2003).

Table 12. Genetic interactions of *dBRWD3* with *hop*^{TumI}.

Exp	Genotype	Allele	Tumors (%)			n	z-score	
			None	Small	Large			
I	<i>y,w,hop</i> ^{TumI} /+ ; +/+	OreR	31.0	50.6	18.4	358	-0.4	(*)
II	<i>y,w,hop</i> ^{TumI} /+ ; +/+	OreR	31.0	43.8	25.2	445	-0.4	(*)
II	<i>y,w,hop</i> ^{TumI} /+ ; +/+	<i>w1118</i>	23.9	31.2	44.9	356	0.6	(*)
II	<i>y,w,hop</i> ^{TumI} /+ ; <i>STAT92E</i> /+	<i>397</i>	67.5	21.5	11.0	228	-5.3	(1)
I	<i>y,w,hop</i> ^{TumI} /+ ; <i>STAT92E</i> /+	<i>06346</i>	68.6	26.1	5.3	283	-5.4	(2)
II	<i>y,w,hop</i> ^{TumI} /+ ; <i>STAT92E</i> /+	<i>06346</i>	64.2	26.6	9.2	282	-4.9	(2)
II	<i>y,w,hop</i> ^{TumI} /+ ; <i>dBRWD3</i> /+	<i>05842</i>	56.6	24.4	19.0	221	-3.8	

Values shown represent percentage of 0-24 h old female flies containing no, small or large tumors visible in abdomen or thorax. Table shows results from two independent experiments (first column) undertaken under identical conditions.

(*) 'wildtype' results used to calculate z-scores

Previously published interactions:

(1) Silver and Montell 2001

(2) Hou et al. 1996

Another phenotypic consequence of constitutive JAK/STAT activation, which can be caused by the gain-of-function JAK allele *hop*^{TumI}, is the overproliferation of hemocytes and the frequent formation of melanotic tumors, a phenotype described as a *Drosophila* model for leukemia (Harrison et al. 1995, Luo et al. 1995). In these genetic interaction experiments, removal of one copy of *dBRWD3* is sufficient to reduce the size and the frequency of *hop*^{TumI}-induced melanotic tumors (Figure 31A and B, Table 12).

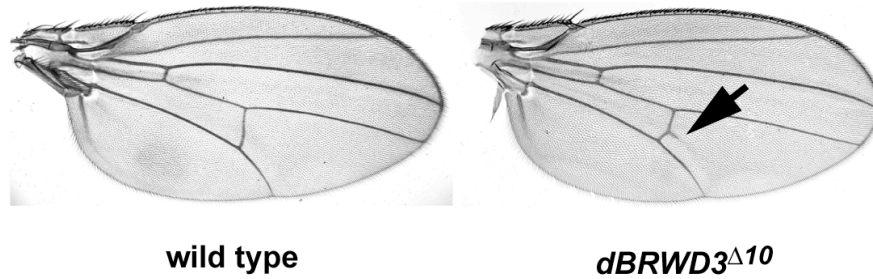


Figure 32. *dBRWD3* mutants phenocopy *stat92E* hypomorphic mutant effects in wing vein formation. By comparison to adult wild-type wings (left), ectopic wing vein material (arrow) is present in homozygous *dBRWD3^{Δ10}* mutant flies (a putative hypomorphic allele, right), a phenotype reminiscent of the *stat92E^{HJ}* mutant (Yan et al. 1996a).

dBRWD3 mutants phenocopy a *stat92E* mutant wing phenotype

Of the seven viable revertant mutants generated by remobilization and excision of the original mutagenic P-element, two include stocks with putative hypomorphic alleles. The homozygous escapers of these putative hypomorphic *dBRWD3*-alleles frequently develop ectopic wing vein material (Figure 32). This phenotype is reminiscent of the weak loss-of-function *stat92E^{HJ}* allele (Yan et al. 1996a), which contains a point mutation leading to a truncated protein version. The exact nature of the hypomorphic *dBRWD3* mutant and the relevance for JAK/STAT signaling remain to be investigated.

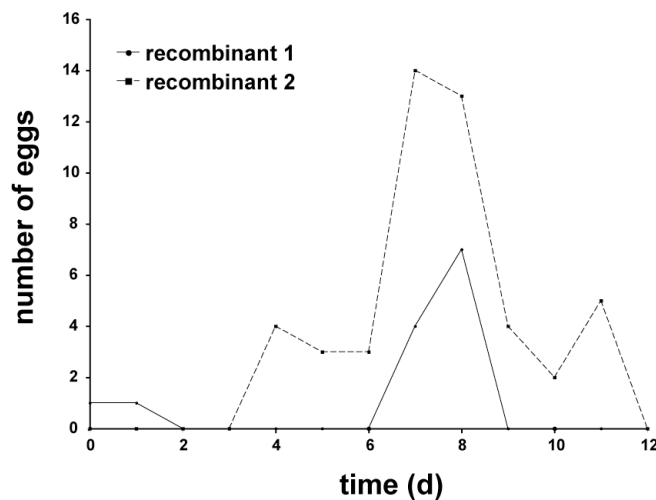


Figure 33. Flies with *dBRWD3* mutant germline clones are affected in egg laying. Germline clones were induced using the DFS technique (Chou and Perrimon 1996). The number of eggs per day laid by female flies mutant for *dBRWD3⁰⁵⁸⁴²* in the germline was counted for two different fly lines where the *dBRWD3⁰⁵⁸⁴²* mutation was recombined with FRT chromosomes. Day 0 is the time-point of crossing to *FRT-dBRWD3⁰⁵⁸⁴²* males.

dBRWD3 mutant germline clones have a defect in egg laying

As shown in Figure 29, *dBRWD3* appears to be maternally loaded into the embryo thereby obscuring the analysis of the early embryonic phenotype due to persistence of transcripts and possibly proteins. In order to circumvent the maternal provision of transcripts, germline mutant *dBRWD3*⁰⁵⁸⁴² clones were produced using the autosomal-FLP dominant female sterile technique as described by Chou and Perrimon 1996. *dBRWD3*⁰⁵⁸⁴² mutant clones were induced at late larval stages by heat shock, activating an FLP transgene and thereby allowing the recombination of FRT sites. Young virgins immediately after eclosion were mated with *dBRWD3*⁰⁵⁸⁴² mutant heterozygous males and their eggs were collected. The frequency of egg-laying in these females is shown in Figure 33. After an initial delay also observed in wildtype animals, eggs are deposited for a short period of time until the egg production ceases only few days later. This phenotype indicates a role for some aspect in ovary development, oogenesis, general cellular proliferation or germline stem cell maintenance. All of these phenotypes would be consistent with JAK/STAT regulatory roles of *dBRWD3*, although a direct examination of the ovary phenotypes present in the germline-clone females would be required to establish this definitively.

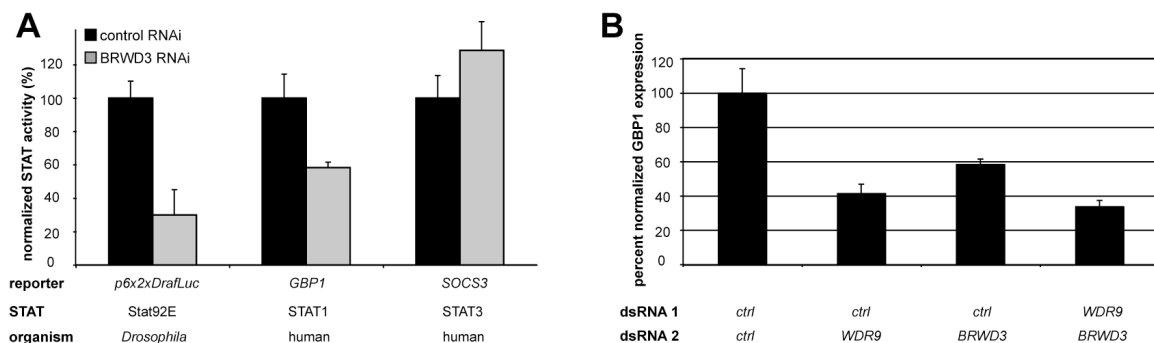


Figure 34. Human BRWD3 is specifically involved in STAT1 signal transduction. (A) Comparison of transcriptional readouts of Stat92E, STAT1 and STAT3 in *BRWD3* knockdown backgrounds compared to cells treated with a non-targeting dsRNA. Co-reporters used for normalization were *pAct-RL* in *Drosophila* and β -actin endogenous levels in human cells. (B) Double-knockdown experiment for *WDR9* (also known as *C21ORF107* and *BRWD1*) and *BRWD3*. *WDR9*, *BRWD3* and *WDR9/BRWD3* double-knockdowns show significantly less *GBP1* expression compared to the ctrl siRNA treatment (p-value < 0.0071 in a Mann-Whitney test), and the *WDR9/BRWD3* double-knockdown shows significantly less *GBP1* expression compared to the *WDR9* single knockdown (p-Value < 0.04 in a Mann-Whitney test).

dBRWD3 acts downstream of the pathway kinase Hop

The *in vivo* and *in vitro* data described above suggest a role for BRWD3 as a positive regulator of JAK/STAT signaling. Furthermore, the BRWD3 loss-of-function phenotype is functionally also conserved in both *Drosophila* and human cells (Figure 34). Knockdown of human *BRWD3* results in a reduction in human STAT1 induced target gene *GBP1* expression, whereas the levels of the human STAT3 reporter target gene *SOCS3* remains unchanged. Another potential homolog of *Drosophila* BRWD3 is BRWD1 (also known as WDR9, D'Costa et al. 2006). Knockdown of human *WDR9* has a similar phenotype as the *BRWD3* knockdown and double-knockdown of both *WDR9* and *BRWD3* leads to an enhanced JAK/STAT phenotype (Figure 34).

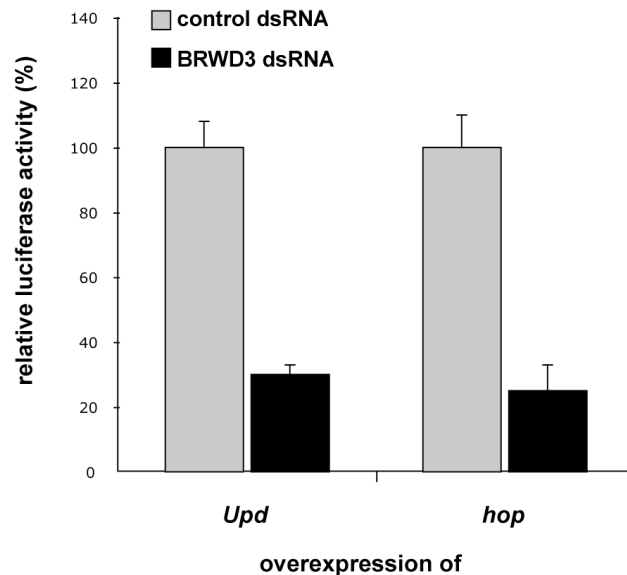


Figure 35. dBRWD3 acts downstream of the kinase Hop. Epistasis analysis of BRWD3. Results shown have been obtained in two independent replicate experiments. Abbreviations used are: ectopic expression of *upd* ('Upd') and expression of a constitutively active JAK-allele *hop^{Tumt}* ('hop', Harrison et al. 1995). Values were normalized to the full activity (100%) of Kc₁₆₇ cells treated with control dsRNA against *lacZ*.

The next question to be addressed was then, where BRWD3 would exert its effect in the signaling regime. The experiments *in vivo* in *Drosophila* with gain-of-function *hop* already indicated that BRWD3 likely acts downstream of JAK, since flies mutant for *dBRWD3* could counteract the JAK induced leukemia phenotype (Figure 31, Table 12). To confirm these findings, epistasis analysis was carried out in *Drosophila* cultured cells

with pathway stimulation either by the pathway ligand or the pathway kinase (Figure 35). Under both conditions, *dBRWD3* knockdown is effective indicating that BRWD3 acts downstream or in parallel to JAK.

dBRWD3 and Stat92E may physically interact

If BRWD3 acts downstream of the kinase Hop in the JAK/STAT pathway, the next question to be addressed was if BRWD3 could physically interact with the Stat92E transcription factor, which is also downstream of Hop in the pathway. To this end, a co-immunoprecipitation experiment was performed. *Drosophila* Kc₁₆₇ cells were transfected with vectors expressing tagged *stat92E* and/or tagged *dBRWD3*. Following incubation for 3 d, cells were lysed and proteins were immunoprecipitated with either α -Myc or α -Flag antibodies to pull down Stat92E-Myc or BRWD3-Flag and their associated binding partners, respectively. Subsequently, the complexes were assayed using a Western Blot with an α -Myc antibody to detect tagged Stat92E. The binding was specific as no proteins leading to a detectable signal were bound to the column material control (Figure 36B). Next, the individual tagged domains of dBRWD3 (Figure 36A) were tested for their interaction with Stat92E-Myc. The experiments shown in Figure 36B indicate that the WD40- and bromo-domains can independently bind Stat92E under the conditions tested and that the co-immunoprecipitation works in both directions. However, although immunoprecipitating the truncation comprising the bromo-domains and the C-terminus of BRWD3 (Bromo+Cterm) leads to a signal for the Stat92E-Myc protein, a signal cannot be detected for the BRWD3 truncation itself. Similarly, immunoprecipitating Stat92E-Myc does not lead to the detection of the full-length BRWD3 protein. Whether this is due to limits in the detectable range, degradation, artefactual binding or epitope-masking in the case of the Bromo+C-term truncation (since it can also not be detected in input) remains to be investigated. A further analysis of the individual WD40- and bromo-domains of dBRWD3 (Figure 36A) shows that both of these are capable of independently binding Stat92E under very stringent binding conditions ranging from 0 to 400 mM salt in the binding buffer (Figure 36C). Furthermore, the analysis shows that Stat92E under the conditions assayed is detectably associated with the BRWD3 domains independent of the presence of Upd ligand and independent of the phosphorylation status of Stat92E as

seen for the phosphorylation mutant Stat92E(Y704F), which is capable of binding BRWD3 domains (Figure 36D). The binding entity of Stat92E, at least to the WD40 domain of BRWD3, appears not to lie in the N-terminal region since an N-terminal truncation of Stat92E is still capable of binding the WD40-domain of BRWD3 (Figure 36E). Interestingly, the bromo-domains appear to bind the Stat92E-GFP protein less than the Stat92E-Myc protein. *Inter alia*, this could be due to the binding of the bromo-domains to the Myc-tag of Stat92E-Myc, a possibility that has to be investigated further in the future using negative controls of similarly tagged proteins unrelated to JAK/STAT signaling for co-immunoprecipitation with BRWD3-Flag.

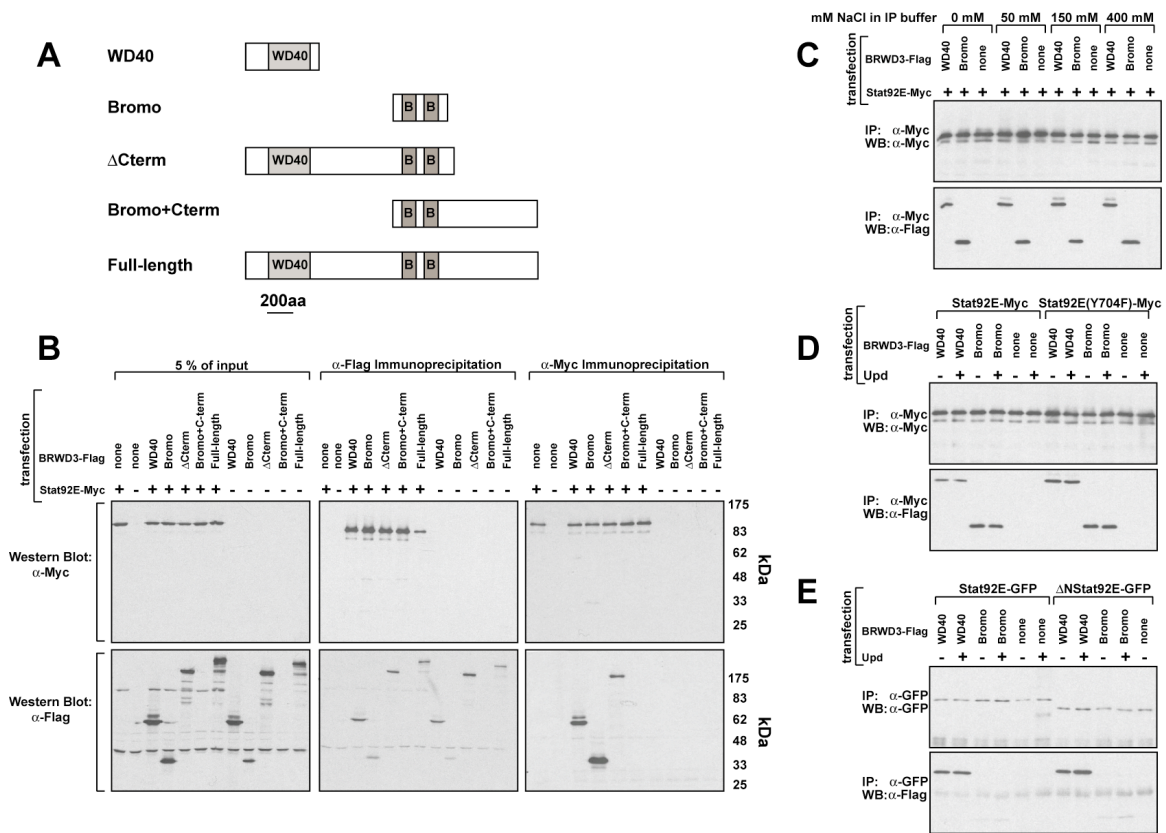


Figure 36. Physical interaction of dBRWD3 and Stat92E. In all experiments, Kc_{167} cells were transfected with plasmids expressing *upd*, *stat92E-10xMyc* or *dBRWD3-Flag* as indicated in the figures. Proteins were then immunoprecipitated (IP) from whole cell lysates with α -Myc or α -Flag antibodies. The immunoprecipitated complexes were subjected to Western Blot (WB) analysis and the blots were probed with the indicated antibodies. (A) Overview of the Flag-tagged BRWD3 truncation constructs used for transfection of *Drosophila* cells. (B) Screen to identify domains of BRWD3 capable of interacting with Stat92E. (C) Strength of BRWD3-Stat92E complex assayed using binding conditions ranging from 0 to 400 mM NaCl in the binding reaction. (D) Influence of phosphorylation on complex formation. (E) Influence of the Stat92E N-terminal domain in complex formation.

Analysis of Stat92E and dBRWD3 associated subproteomes

The interaction between Stat92E and BRWD3 was analyzed by ectopic expression of transfected constructs encoding tagged versions of these proteins in cell culture and was observed under stringent binding conditions ranging from 0 mM to 400 mM salt included in the binding buffer used for co-immunoprecipitation experiments. Experiments to verify that this interaction is not artefactual and can also take place between the endogenous proteins were hampered by the unavailability of specific antibodies. Antiserum generated from dBRWD3 protein comprising the WD40 domain injected into guinea pigs in the present study by Eurogentec detected multiple protein bands of the wrong size, and antibodies against Stat92E previously available from Cell Signaling Technologies were shown to be cross-reactive and fail to detect endogenous Stat92E (Melissa Henriksen, personal communication).

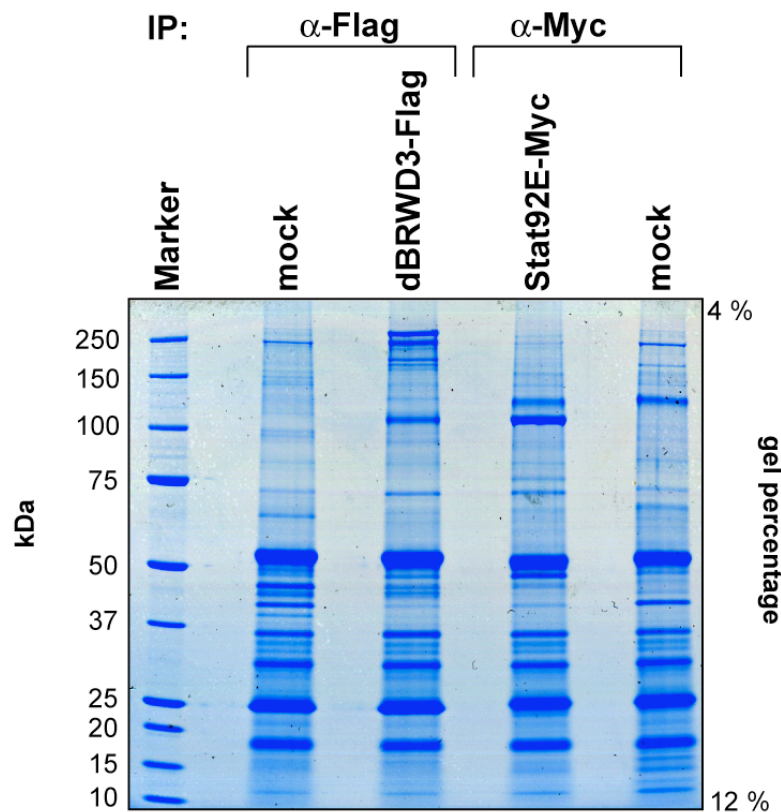


Figure 37. Identification of Stat92E and dBRWD3 associated subproteomes. Kc₁₆₇ cells were transfected with empty vector (mock), dBRWD3-Flag or Stat92E-Myc encoding constructs, grown for three days and lysed. Protein complexes were then immunoprecipitated using α -Myc or α -Flag antibodies as indicated in the figure. Immunoprecipitated proteins were subjected to SDS-PAGE using a NuPAGE gradient gel (4-12% polyacrylamide), and proteins were stained with colloidal coomassie blue. The marker on the left indicates the approximate size of bands in kDa.

Another sensitive technique to analyze protein complexes is mass spectrometry, which was therefore used to analyze the subproteomes associated with ectopic tagged BRWD3 and Stat92E proteins in *Drosophila* cells. Cells were transfected with empty vector (mock) or constructs encoding Flag-tagged BRWD3 or Myc-tagged Stat92E. After immunoprecipitation (IP) with the antibodies indicated in Figure 37, eluted proteins were separated on a 1D SDS-PAGE. The stained gel shows bands differentially present in the IPs of BRWD3-Flag and Stat92E-Myc positive cells, which correlate to the correct sizes of approximately 250 kDa and 100 kDa for the respective proteins (Figure 37). Notably, in addition to the 250 kDa band, there is a strong differentially present band in the BRWD3-Flag IP migrating at the same height as the 100 kDa band in the Stat92E-Myc IP, supporting the hypothesis that also the endogenous proteins may interact physically (Figure 37). To test this hypothesis systematically, each sample lane was cut into 24 pieces and subjected to trypsin in-gel digestion. The eluted peptides were then analyzed by LC-MS/MS (Liquid chromatography tandem mass spectrometry). In this technique, the peptides were first separated chromatographically before the sequentially eluting peptides were ionized. Doubly and/or triply charged peptides were automatically selected and fragmented in a collision chamber. The amino acid sequence could be deduced from the detection of the resulting fragments statistically differing by the mass of one amino acid. A database containing information about the *Drosophila melanogaster* proteome was then searched for proteins matching the LC-MS/MS dataset. Only proteins differentially present in the BRWD3-Flag or Stat92E-Myc IP were considered for further analysis and are listed in Table 13 and Table 14 (see Supplementary Table 7 for the common background in the IP samples). LC-MS/MS analysis of the Stat92E-Myc IP identified a total of 53 *Drosophila melanogaster* proteins with Mascot protein scores greater than zero, of which 29 are differentially present, four (14%, including CG10730 and the pathway receptor Dome) of which have a phenotype in the JAK/STAT RNAi screen (z -score > 2 or < -2) and nine of which can be considered as confident hits identified by LC-MS/MS. A similar analysis for the BRWD3-Flag IP identified 59 proteins of which 29 are differentially present, nine (31%, including CkII, eIF-4B and mask) of which have a JAK/STAT phenotype and 11 of which can be considered as confident hits (see MATERIALS AND METHODS). Since the selected z -score threshold

Table 13. Proteins identified after α -Myc immunoprecipitation of Stat92E-Myc transfected cells.

GI number	Protein name	Protein Score	Protein Mass	Peptides		J/S	FBgn
				Id.	Conf.		
gil45553427	Stat92E CG4257-PE, isoform E	1643	86	28	21	-5.0	FBgn0016917
gil17647519	Heat shock protein 26 CG4183-PA	69	23	5	3	1.2	FBgn0001225
gil66771263	IP07888p	56	32	2	1	na	FBgn0002611
gil24585375	CG10730-PA	46	76	1	1	-2.1	FBgn0032843
gil24646562	CG8863-PE, isoform E	31	45	5	1	-1.1	FBgn0038145
gil19922778	GDP-4-keto-6-deoxy-D-mannose 3,5-epimerase/4-reductase CG3495-PA	27	36	2	1	-0.8	FBgn0034794
gil19920556	CG15385-PA	25	65	1	1	1.2	FBgn0031397
gil24661872	nbs CG6754-PB	24	92	6	1	0.1	FBgn0026198
gil19921528	Asparaginyl-tRNA synthetase CG10687-PA	17	64	5	2	0.4	FBgn0027092
gil290262	prd-like homeobox protein	39	44	2	0	-1.5	FBgn0000061
gil20136450	pol protein	23	98	3	0	-1.4	FBgn0003277
gil19527521	RE22456p	20	101	5	0	-0.7	FBgn0011692
gil28572120	Diphthamide methyltransferase CG31289-PA	22	32	1	0	-0.2	FBgn0024558
gil24643078	lethal (1) G0003 CG6606-PA	25	92	1	0	0.5	FBgn0027335
gil45549338	CG1530-PA	16	86	2	0	0.5	FBgn0029983
gil24583065	CG13113-PA	18	24	2	0	-0.3	FBgn0032126
gil24583234	CG12441-PB, isoform B	30	25	2	0	-1.0	FBgn0032185
gil20129463	CG12264-PA	31	51	4	0	0.1	FBgn0032393
gil24654080	CG15701-PA	16	117	5	0	1.9	FBgn0034095
gil24657549	CG11474-PA, isoform A	19	49	2	0	-1.7	FBgn0034688
gil24655178	CG9168-PA	20	70	10	0	0.0	FBgn0035216
gil21358055	meso18E CG14233-PA	27	94	1	0	-0.3	FBgn0040089
gil21357231	CG12034-PA	22	50	3	0	-1.3	FBgn0035421
gil24662954	Ribosomal protein L10Ab CG7283-PC, isoform C	16	16	1	0	-3.9	FBgn0036213
gil21355769	CG18005-PA	15	61	3	0	0.6	FBgn0037660
gil45550777	Nepriylsin 4 CG4058-PA, isoform A	15	120	4	0	0.8	FBgn0038818
gil45552313	rhomboid-5 CG33304-PA	25	157	3	0	-0.5	FBgn0041723
gil17647357	domeless CG14226-PA	15	142	6	0	-6.2	FBgn0043903
gil24652768	CG30034-PA	29	39	2	0	na	FBgn0050034

Protein score was calculated by MudPIT Scoring using Mascot (Matrix Science), protein mass is given in kDa. Peptides refer to the number of unique identified peptides for the given protein (Id.) and the number of manually confirmed (Conf.) assigned peptides.

The JAK/STAT phenotype 'J/S' is expressed as the z-score calculated from two replicate FL channels obtained in the genome-wide RNAi screen of the present study. 'na' indicates that there was no matching probe found in the dataset.

Table 14. Proteins identified after α -Flag immunoprecipitation of dBRWD3-Flag transfected cells.

GI number	Protein name	Protein Score	Protein Mass	Peptides		J/S	FBgn
				Id.	Conf.		
gil24649631	BRWD3 CG31132-PA	2823	249	79	55	-2.8	FBgn0011785
gil21357503	DDB1 CG7769-PA	1421	126	41	30	-1.8	FBgn0027049
gil24668866	Casein kinase II subunit CG17520-PC, isoform C	77	40	8	6	-2.1	FBgn0000258
gil1359608	replication protein A	88	67	6	4	-0.9	FBgn0010173
gil51646256	TPA: eukaryotic initiation factor 4B	148	44	12	4	-3.2	FBgn0020660
gil8070	H3 histone	47	15	2	2	na	FBgn0001199
gil397852	ribosomal protein L27a	20	17	1	1	-5.3	FBgn0010410
gil24642434	cabeza CG3606-PA, isoform A	25	39	5	1	-0.6	FBgn0011571
gil4481810	EG:BACN32G11.5	21	70	4	1	-0.7	FBgn0028274
gil969093	ORF2	26	39	1	1	1.4	FBgn0032408
gil19528549	RH07841p	41	26	3	1	-0.6	FBgn0032906
gil17647515	Heat shock protein cognate 1 CG8937-PA, isoform A	56	71	8	0	0.0	FBgn0001216
gil157678	heat shock cognate 70 protein (partial) (at locus 88E)	38	11	3	0	-2.9	FBgn0001219
gil17945501	RE23308p	18	40	1	0	-0.5	FBgn0003345
gil21428640	LP10436p	31	51	3	0	2.4	FBgn0003890
gil289002	lysozyme precursor	23	16	1	0	1.1	FBgn0004425
gil510509	GCR 101	18	23	3	0	0.6	FBgn0011824
gil1321806	CG11538	38	74	2	0	0.2	FBgn0017424
gil17737907	Ribosomal protein L3 CG4863-PA, isoform A	31	47	5	0	-3.4	FBgn0020910
gil54650696	LP20978p	22	56	3	0	-0.2	FBgn0031497
gil19921698	CG3271-PB, isoform B	16	39	1	0	-0.6	FBgn0033088
gil24653937	CG8249-PA	24	58	2	0	-0.8	FBgn0034045
gil21406609	AT13486p	17	27	3	0	-0.6	FBgn0034601
gil33589306	RH51767p	27	31	4	0	-0.3	FBgn0037419
gil24647369	CG8927-PA, isoform A	18	40	2	0	-0.7	FBgn0038405
gil28571790	CG10825-PA, isoform A	28	79	5	0	1.0	FBgn0038860
gil18251232	multiple ankyrin repeat single KH domain protein	15	423	13	0	-2.3	FBgn0043884
gil45550446	CG30460-PC, isoform C	17	206	8	0	-4.0	FBgn0050460
gil24581260	CG3104-PB, isoform B	26	34	2	0	-0.7	FBgn0031473

Protein score was calculated by MudPIT Scoring using Mascot (Matrix Science), protein mass is given in kDa. Peptides refer to the number of unique identified peptides for the given protein (Id.) and the number of manually confirmed (Conf.) assigned peptides.

The JAK/STAT phenotype 'J/S' is expressed as the z-score calculated from two replicate FL channels obtained in the genome-wide RNAi screen of the present study. 'na' indicates that there was no matching probe found in the dataset.

should theoretically yield a total of 5% hits in a random dataset, the interacting proteins in the Stat92E-Myc IP are approximately three-fold enriched over a random selection, whereas the proteins present in the dBRWD3-Flag IP are approximately six-fold enriched over a random selection with respect to a JAK/STAT phenotype. Although the 1D SDS-PAGE analysis suggested a result consistent with the co-immunoprecipitation experiments (Figure 37), the initial LC-MS/MS data does not support an interaction between endogenous BRWD3 and Stat92E, and the band migrating at approximately 100 kDa in the BRWD3-Flag IP is very likely enriched in DDB1 (Table 14). The high abundance of DDB1 protein may have resulted in failure to identify the potentially extremely low abundant Stat92E in the BRWD3-Flag IP sample due to the limited detection time and automatic selection of the most abundant peptides in LC-MS/MS analysis. The same rationale may apply to the lack of BRWD3 in the Stat92E-Myc IP sample, thereby not excluding the possibility that the endogenous proteins truly physically interact.

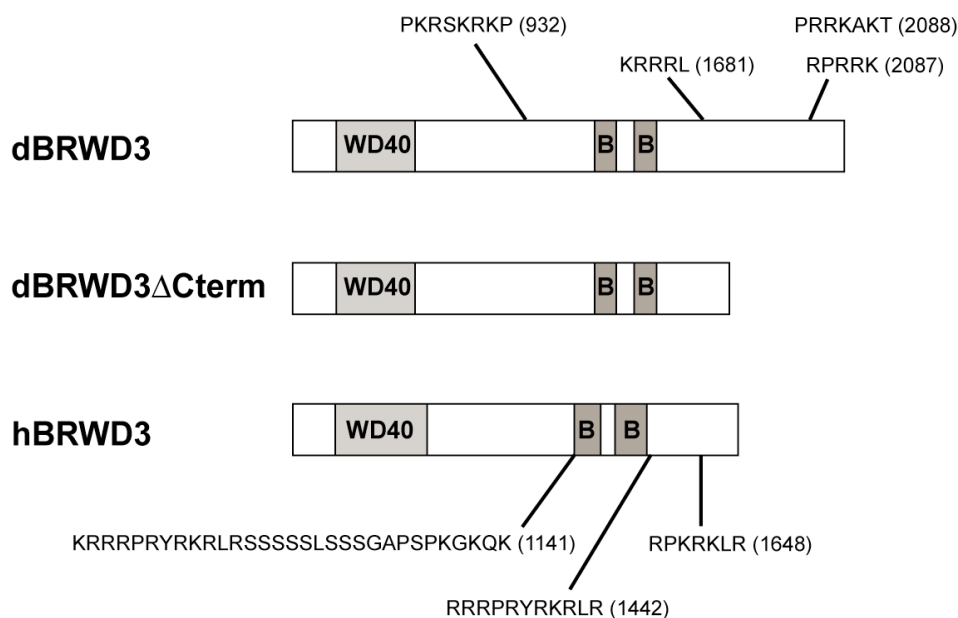


Figure 38. Nuclear localization signals in BRWD3. A similar distribution and spacing of three predicted nuclear localization signals (NLS) is present in the *Drosophila* and human BRWD3 proteins. The sequence in amino acids and the position is indicated in the figure. Note that the C-terminal truncations dBRWD3 Δ Cterm used in the following experiments lack the C-terminal most NLS. NLS were predicted by the program ‘PredictNLS’. For both *Drosophila* and human proteins, a high probability of nuclear localization (1.0 for dBRWD3 and 0.99 for hBRWD3) was predicted by the program ‘NucPred’ (<http://www.sbc.su.se/~maccallr/nucpred/>, Heddad et al. 2004).

dBRWD3 is localized in the nucleus

Given the finding from co-immunoprecipitation experiments that BRWD3 is able to physically interact with Stat92E, the next question to be addressed was where this interaction would take place, since the latent cytoplasmic transcription factor Stat92E can be shuttled to the nucleus after cytokine stimulation. A closer look at the BRWD3 protein amino acid composition revealed the likely existence of three putative nuclear localization signals (NLS). Additionally by computational prediction, this protein has a high probability of nuclear localization (Figure 38).

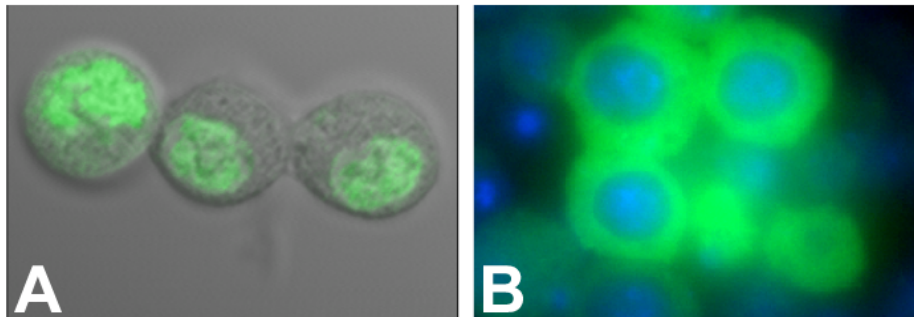


Figure 39. Intracellular localization of dBRWD3 variants. *Kc*₁₆₇ cells were transfected with Flag-tagged constructs encoding full-length (*pBRWD3(full)-Flag*) or truncated dBRWD3 (*pBRWD3(ΔCterm)-Flag*), and the tagged proteins were stained with α -Flag antibody. (A) Full-length dBRWD3 appears to be exclusively localized in the nucleus (grey: transmission, green: α -Flag). (B) dBRWD3 lacking the C-terminus (dBRWD3 Δ Cterm) is excluded from the nucleus (blue: DNA, green: α -Flag). Note that the overall number and morphology of the cells for the expression of both BRWD3 variants were comparable.

Strikingly, the number of NLS appears to be conserved between the proteins in *Drosophila* and humans. Furthermore, the spacing relative to the defined domains is very similar with two NLS located C-terminally of the bromo-domains and one NLS interspaced between the WD40- and bromo-domains. Interestingly, truncated versions of *BRWD3* have been found in human B-CLL patients, in which the region coding for the C-terminally located NLS is removed (Kalla et al. 2005; Claudia Kalla, personal communication). In order to test the relevance of the most C-terminal NLS for a potential disease-mechanism, a truncated version of dBRWD3, hereafter referred to as dBRWD3 Δ Cterm, was therefore generated and tested for its intracellular localization compared to its full-length counterpart. Figure 39 shows that the full-length *Drosophila*

BRWD3 is normally almost exclusively localized to the nucleus. This localization is independent of the presence of ectopic Upd ligand. In contrast, the C-terminally truncated version dBRWD3 Δ Cterm is almost exclusively localized to the cytosol and excluded from the nucleus. These data indicate that the most C-terminal NLS is necessary for normal nuclear localization of the protein.

Table 15. Effect of *dBRWD3* ectopically expressed in developing *Drosophila* tissues.

Temperature	Gal4 driver line	UAS line	Phenotype
25°C	<i>ey-Gal4</i>	<i>UAS-BRWD3-1</i>	lethal
25°C	<i>GMR-Gal4</i>	<i>UAS-BRWD3-1</i>	lethal
25°C	<i>cg-Gal4</i>	<i>UAS-BRWD3-1</i>	lethal
25°C	<i>MS1096-Gal4</i>	<i>UAS-BRWD3-1</i>	lethal
25°C	<i>ey-Gal4</i>	<i>UAS-BRWD3-2</i>	lethal
25°C	<i>GMR-Gal4</i>	<i>UAS-BRWD3-2</i>	lethal
25°C	<i>cg-Gal4</i>	<i>UAS-BRWD3-2</i>	lethal
25°C	<i>MS1096-Gal4</i>	<i>UAS-BRWD3-2</i>	lethal
25°C	<i>ey-Gal4</i>	<i>UAS-BRWD3-3</i>	mostly lethal, otherwise rough small eyes
25°C	<i>GMR-Gal4</i>	<i>UAS-BRWD3-3</i>	lethal
25°C	<i>cg-Gal4</i>	<i>UAS-BRWD3-3</i>	lethal
25°C	<i>MS1096-Gal4</i>	<i>UAS-BRWD3-3</i>	lethal
18°C	<i>ey-Gal4</i>	<i>UAS-BRWD3-1</i>	mostly lethal, otherwise no eyes
18°C	<i>GMR-Gal4</i>	<i>UAS-BRWD3-1</i>	lethal
18°C	<i>cg-Gal4</i>	<i>UAS-BRWD3-1</i>	lethal
18°C	<i>MS1096-Gal4</i>	<i>UAS-BRWD3-1</i>	unexpanded small wings
18°C	<i>ey-Gal4</i>	<i>UAS-BRWD3-2</i>	mostly wildtype with slightly reduced eye size
18°C	<i>GMR-Gal4</i>	<i>UAS-BRWD3-2</i>	mostly lethal, otherwise small eyes
18°C	<i>cg-Gal4</i>	<i>UAS-BRWD3-2</i>	lethal
18°C	<i>MS1096-Gal4</i>	<i>UAS-BRWD3-2</i>	unexpanded small wings
18°C	<i>ey-Gal4</i>	<i>UAS-BRWD3-3</i>	mostly wildtype with slightly reduced eye size
18°C	<i>GMR-Gal4</i>	<i>UAS-BRWD3-3</i>	small eyes
18°C	<i>cg-Gal4</i>	<i>UAS-BRWD3-3</i>	lethal
18°C	<i>MS1096-Gal4</i>	<i>UAS-BRWD3-3</i>	wrinkled wings

ey-Gal4 and *GMR-Gal4* were used to drive the expression in the developing eye, *cg-Gal4* for ectopic expression in hemocytes and *MS1096-Gal4* for ectopic expression in the wing.

Next, the effect of the C-terminally truncated protein without the NLS on JAK/STAT signal transduction was assessed with the same reporter assay used for the genome-wide RNAi screen. Ectopic full-length dBRWD3 does not seem to exert an effect, at least in cell culture (Figure 40), although overexpression *in vivo* in *Drosophila* tissues using the Gal4/UAS system (Brand and Perrimon 1993) leads to dramatic malformation in these tissues (Table 15). However, expression of ectopic dBRWD3 Δ Cterm leads to a

significant reduction in the cell culture based reporter assay for JAK/STAT signaling, showing a dominant-negative effect (Figure 40).

The dominant-negative effect on the transcriptional readout of the JAK/STAT pathway occurs for the C-terminal truncation dBRWD3 Δ Cterm, which is similar to the disease-relevant truncated versions of human BRWD3 found in B-CLL patients (Kalla et al. 2005; Claudia Kalla, personal communication). Whether the trapping of the truncated form in the cytosol and the corresponding effect on transcription is relevant for the onset of B-CLL has to be definitively established in future experiments.

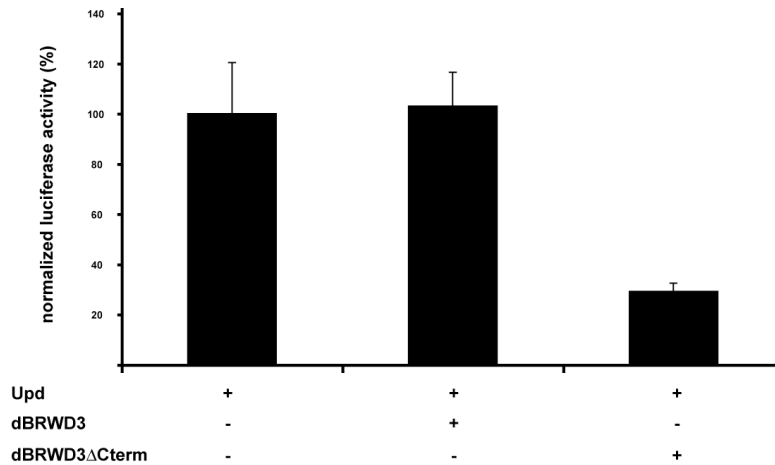


Figure 40. A C-terminal dBRWD3 truncation acts dominant-negatively. Kc₁₆₇ cells were transfected with the constructs indicated in the figure ('Upd' is the vector *pAct-UpdGFP*, 'dBRWD3' is *pAc5.1-dBrodl*, 'dBRWD3 Δ Cterm' is *pAc5.1-dBrodl Δ C*) along with the *p6x2xDrafLuc* and *pAct-RL* reporters. Normalized luciferase activity represents the FL/RL ratio normalized to the full activity without any dBRWD3 constructs. Error bars represent the standard deviations of eight biological replicate data-points.

DISCUSSION

The JAK/STAT signal transduction cascade is an important cellular pathway regulating many developmental processes in diverse multicellular organisms. Over the past 15 years, many of the canonical core pathway components have been identified and characterized in detail. However, much less is known about the *in vivo* regulation of these components that is required to control the right amount of signal at the right time and place. Considering the diversity of the complex developmental decisions regulated by the JAK/STAT pathway, it can be suspected *a priori* that complex regulatory networks exist to monitor and to control appropriate signaling levels. Dysregulation of proper pathway activity can result in severe developmental disorders and diseases (James et al. 2005, Rosenfeld et al. 2005), underlining the importance of systems-knowledge about dynamic pathway regulation. The present study extends the analysis of the JAK/STAT pathway to identify and characterize novel pathway regulators and shows that these have been functionally conserved throughout evolution.

A genome-wide RNAi screen to identify JAK/STAT regulators

A genome-wide RNAi screen was undertaken in *Drosophila* cultured cells to identify dsRNAs targeting gene activities involved in JAK/STAT signaling. To this end, a robust and sensitive dual reporter assay system was developed allowing the discrimination of pathway specific and unspecific responses upon dsRNA treatment. The stimulation of pathway activity relied on continuous ectopic *upd* expression thereby allowing the identification of gene activities involved in the production of fully functional pathway ligand. However, continuous pathway activation bears the danger that target genes involved in the regulation of JAK/STAT signaling by negative or positive feedback mechanisms may also be constantly induced thereby interfering with pathway induction, which is normally very transient *in vivo* (Lerner et al. 2003). Therefore, a screen in which cells would have been treated with RNAi first followed by specific pathway induction later, would have been more physiological with respect to transient pathway stimulation. However, this would have meant a much more technically challenging screening setup, where the inducing agent would have to be added after RNAi incubation making the data

itself also more prone to artefacts, e.g. by bacterial contamination after opening the screening plates. An inducible reporter system relying on copper induced expression of *upd* was developed but not used for screening because of the possibility of copper induced screening artefacts. The alternative, transfection after RNAi treatment, would have led to the necessity of normalizing the pathway reporter levels to those from the co-reporter due to differences in well-to-well transfection efficiencies. As outlined below, this dual reporter normalization can lead to the identification of normalization-artefact induced false-positives thereby skewing the final result hit list. Nevertheless, despite constitutive pathway activation, the effects of depleting downstream negative feedback targets were clearly detectable, as is the case for *socs36E* and *ptp61F*.

The genome-wide screen was performed in duplicate and every screening plate contained dsRNA targeting known components of the pathway. This experimental setup allowed for the calculation of statistics and quality metrics to assess overall data quality. When the screen in the present study was performed in the year of 2004, no publicly available data analysis tools had been developed to examine the data obtained. Therefore, new customized tools for data analysis had to be developed in the present study, which were implemented in the computational language R. This is open source code provided in Supplementary Script 1 and described in the Supplementary Tutorial. Data analysis of the present screen shows the reproducibility of replicate datasets as well as the inherent asymmetry and non-normality of the dataset. Plotting functions for calculated scores reveal systematic spatial bias effects, which can be corrected through normalization by well in 3D and by median polishing using ‘Tukey’s two-way median polish’ procedure (Tukey 1977).

Screens for novel JAK/STAT components have been performed before in *Drosophila*. Two forward genetic screens (Bach et al. 2003, Mukherjee et al. 2006) have made use of the *GMR-Upd* transgenic fly model to specifically drive the ectopic expression of the JAK/STAT pathway ligand Upd in the developing eye leading to increased cellular proliferation and a massively overgrown eye phenotype. One group used a deficiency screen in this sensitized background to first identify genetic regions whose loss modulates

Table 16. Comparison of results from different JAK/STAT modulator screens.

Gene	Bach	Mukherjee	Baeg	Müller
<i>stat92E</i>	Su	-	Decreased	-5.0
<i>ctBP</i>	En	-	na	-2.9
<i>ssdp</i>	na	-	Increased	2.1
<i>dome</i>	na	na	Decreased	-6.2
<i>hop</i>	na	na	Decreased	-5.7
<i>brm</i>	na	na	Increased	-5.7
<i>mor</i>	na	na	Increased	-5.2
<i>CG13235</i>	na	na	Increased	-5.2
<i>CG17836</i>	na	na	Decreased	-4.6
<i>CG11700</i>	na	na	Decreased	-4.3
<i>CG30460</i>	na	na	Increased	-4.0
<i>proct</i>	na	na	Increased	-3.2
<i>CG15563</i>	na	na	Increased	-3.0
<i>hsc70-4</i>	na	na	Increased	-2.9
<i>dBRWD3</i>	na	na	Decreased	-2.8
<i>cdc2</i>	na	na	Decreased	-2.8
<i>CG6434</i>	na	na	Increased	-2.8
<i>ken</i>	na	na	Increased	-2.3
<i>samuel</i>	na	na	Increased	-2.2
<i>mbl</i>	na	na	Increased	-2.0
<i>apt</i>	na	na	Increased	2.1
<i>enok</i>	na	na	Increased	3.0
<i>socs36E</i>	na	na	Increased	3.2
<i>ptp61F</i>	na	na	Increased	5.9

Data from Bach et al. 2003 ('Bach'), from Mukherjee et al. 2006 ('Mukherjee'), from Baeg et al. 2005 ('Baeg') and from the present study ('Müller').

The individual scoring systems for phenotypes of each study are shown. 'En' in enhancer, 'Su' is suppressor, '-' means suppression, and z-scores are given for the column 'Müller'. 'na' indicates that this gene was not identified in the respective screen.

the overgrown eye phenotype (Bach et al. 2003), whereas the other group used a library of flies with P-elements randomly integrated into their genes (Mukherjee et al. 2006).

Both groups then went on to characterize the interacting regions in more depth using candidate gene approaches. The comparison of the present dataset with these screens reveals a small overlap (Table 16). For example, only *ctBP* (encoding the C-terminal binding protein) was identified as an interactor in all three screens with varying roles as an enhancer or suppressor. Another gene identified in multiple screens is *ssdp* again with different roles as either an enhancer or suppressor (Table 16). This small overlap may first of all be explained by the use of different screening systems. Possibly there are

different requirements for pathway regulators in the developing eye tissue compared to the requirements in a cultured cell line. It also has to be noted, however, that the two genetic screens themselves only show a small overlap. This is likely an inherent feature of genetic approaches where the interaction can occur at many stages and levels, including processes from downstream of the pathway. Additionally, the genetic screens undertaken are inherently non-saturating. However, the interacting chromosomal deficiencies identified by Bach et al. 2003 overlap largely with the chromosomal locations of the interactors identified in the present screen, and it is likely that interacting genes lying in these deficiency regions share homology with the regions targeted by the dsRNAs.

Another reverse genetic screen using a genome-wide RNAi library to screen for novel JAK/STAT signaling modulators has also recently been published (Baeg et al. 2005). For the identification of novel regulators of JAK/STAT signaling, Baeg and colleagues generated a luciferase-based reporter that contains multimerized Stat92E binding sites taken from the *socs36E* enhancer region (Table 17). In a *Drosophila* Schneider cell line derivative, S2-NP, this reporter reflects signaling activity induced by endogenous levels of the ligand Upd2. For the genome-wide RNAi screen, cells were transfected with this reporter and a co-reporter along with individual dsRNAs per well. Four days later, reporter activities were determined. For candidate selection, the plate averages for the ratios of reporter to co-reporter (FL/RL) were calculated and phenotypes expressed as the fold standard deviation (SD) from the mean of each given plate. After retesting, Baeg and colleagues identified 116 novel genes that regulate JAK/STAT signaling in *Drosophila* and five previously known JAK/STAT pathway components (*upd2*, *dome*, *hop*, *stat92E*, *socs36E*). The comparison of the published lists shows an overlap between the present study and Baeg et al. 2005 for the eight genes *socs36E*, *CG6434*, *ptp61F*, *hop*, *stat92E*, *dBRWD3*, *enok* and *dome* (Table 16). Note, however, that this is a comparison between filtered datasets. For example, in the present study phenotypes also identified in other screens (e.g. for viability phenotypes (Boutros et al. 2004)) were filtered and excluded from the final list of 90 interactors (see MATERIALS AND METHODS, Table 17). Therefore, a closer look by parsing the Baeg dataset against the unfiltered data from the present study reveals 22 common phenotypes shown in Table 16 with an overlap of

approximately 20%, although some show opposite phenotypes – possible due to normalization artefacts as described in Figure 17. This overlap between the reverse genetic screens is larger than for the forward genetic screens. Furthermore, in both RNAi screens, the rate of identification of previously known canonical pathway components was very similar, suggesting that only few false-negatives were missed in the screening approaches. The presence of false-positives is possible, especially those related to so-called ‘offtarget effects’ (OTEs) associated with siRNAs binding to more than one target mRNA (Birmingham et al. 2006, Echeverri et al. 2006, Echeverri and Perrimon 2006, Fedorov et al. 2006, Kulkarni et al. 2006, Ma et al. 2006). However, essentially the same library was used for both screens suggesting that OTEs should be consistent between assays.

A more likely explanation for the differences in candidate lists are the differences in the experimental setup (summarized in Table 17). Discrepancies between microarray datasets generated in different laboratories analyzing the same biological process in yeast have been appreciated previously, and also the overlap identified in other genome-wide datasets obtained e.g. by proteomics and protein-protein interaction studies is very low (reviewed in Grünenfelder et al. 2002). These misleading results are likely due to experimental differences. Similarly, for the genome-wide RNAi screen for JAK/STAT modulators, different cell lines were used in the different studies. As demonstrated in Figure 6D, transcription profiles between cells derived from the same organism but from different tissue material can be very different overall. More specifically, the components of the JAK/STAT pathway are differentially present in these cell lines, which for example leads to more JAK/STAT unresponsive S2R+ cells in comparison to S2 and Kc₁₆₇ cells, likely due to reduced levels of the most downstream molecule Stat92E in these cells. Furthermore, in the assays performed in this study, basal levels of pathway induction were very low in the absence of ectopic ligand expression, also consistent with the expression levels of *upd* in these cells. In contrast, Baeg et al. 2005 used an alternative cell line in which the activity of a different ligand, Upd2, appeared to be sufficient for pathway induction. This approach may already determine the kind of regulators that can be identified as hits leading to a different data distribution. Both datasets appear to be non-symmetrical. However, while in the screen of this present study

Table 17. Comparison of the experimental conditions for two JAK/STAT RNAi screens.

Experimental condition	Baeg	Müller
Screening procedure		
Coverage of library	21,300 dsRNAs	20,026 dsRNAs
Pathway stimulation	endogenous Upd2	ectopically expressed Upd-GFP
Screening reporter	<i>10xStat92E-luciferase</i> (FL) containing five tandem repeats of a 441 bp fragment from the <i>socs36E</i> enhancer (each with 2 potential Stat92E binding sites)	<i>p6x2xDrafLuc</i> containing six repeats of a 165 bp fragment from the <i>raf</i> promoter (each with 2 Stat92E binding sites)
Co-reporter	<i>Act-RL</i>	<i>pAct-RL</i>
Cell line	S2-NP	Kc ₁₆₇
RNA concentration per well	80 ng/well	500 ng/well
Cells seeded per well of a 384-well plate	40,000	15,000
Transfection of reporter dsRNA uptake	per well transfection	in batch bathing + SID-1 dsRNA transporter
Time for RNAi	4 d	5 d
Replicate datasets	two	two
Data processing		
Data normalization	fold SD from the plate mean of FL/RL ratio for each plate	fold MAD from the plate median of FL channel for each plate
Selection of positive regulators	< 2 SD below plate mean	< 2 MAD below plate median
Selection of negative regulators	> 3 SD above plate mean	> 2 MAD above plate median
Exclusion of genes	genes not annotated by BDGP, ribosomal proteins, proteins involved in RNA processing and translation	previously published cell viability modifiers, treatments with high variability between replicates, treatments with z-scores > 2 or < 2 in the RL channel, genes with phenotypes in other screens
False-positive rate (determined by re-screens of primary hits)	29%	15%
Human homologs of hits	73%	74%

Data from Baeg et al. 2005 ('Baeg') and from the present study ('Müller').

with a high ectopic pathway induction more positive regulators of the pathway are identified, the study of Baeg et al. 2005 using endogenous basal levels of pathway induction shows a bias towards more negative regulators. Presumably a pathway stimulus in the medium range between these two extremes would have led to a more symmetrical data distribution.

The most likely reason for the difference in these datasets, however, probably lies in the different data analysis methods and approaches used for these two screens. Different normalization approaches can lead to quite different output lists (Table 6 and Figure 17). Surprisingly, two studies of the Wnt signal transduction pathway have recently shown that different co-reporters and different normalization procedures in dual-channel experiments can significantly affect the prediction of the regulatory role for a given candidate gene. For example, the endocytosis regulator Rab5 had initially been identified as a negative regulator of Wnt signaling in a genome-wide RNAi screen (DasGupta et al. 2005), but has subsequently been verified as a positive regulator of the pathway in a similar cell culture system and *in vivo* using a different co-reporter strategy (Seto and Bellen 2006). Furthermore, different sensitivity and signal-to-noise ratios in both screens could lead to different outcomes, and the selected cut-off threshold can significantly affect the final candidate lists. Reliable and robust statistical procedures are therefore important to identify interactors for downstream analysis. Moreover, post-processing filtering for pathway specificity appears to be another important step in generating candidate lists, as many of the overlapping hits found in Baeg et al. 2005 were excluded in the present study because they had been identified in other screens before. It would be very interesting to ascertain the overlap between the two datasets after the analysis using the same computational normalization methods. Surely, if a common analysis approach is not chosen, a public raw data repository to allow for independent comparative analysis in the RNAi screening field in general will be imperative. It would be interesting to see whether these data aggregations would then allow statistic inference on the choice of the best analysis methods, i.e. normalization in 2D, 3D, by z-score or by B-score.

Novel evolutionarily and functionally conserved pathway regulators

After post-processing and filtering of the genome-wide dataset, a total of 91 dsRNAs targeting 90 gene activities were identified that modulate JAK/STAT signaling levels. Interestingly, more positive (67) than negative (24) regulators were identified. However, this asymmetry is reversed in a different RNAi screen dataset (Baeg et al. 2005), arguing more for a matter of identification thresholds in the assay setup rather than for real biological significance. The present dataset contains many proteins previously associated with signal transduction, transcriptional regulation and protein modification. This finding is not necessarily surprising but rather confirms the functional gene ontology classes that would be expected from a screen for modulators of signaling and transcription. As judged from their epistatic relationship to known canonical pathway components (Figure 20), some of the novel modulators could be functionally involved in modification of the Upd ligand for full activity (e.g. CG12213, Ipk2, CG3281, CG31694), others may be involved in a process upstream of Hop but downstream of the receptor Dome. For example, CG31358 is annotated with the GO-term ‘plasma membrane (cellular component)’ and may act as a co-receptor of the pathway. The majority of positive pathway regulators identified, however, act downstream of the JAK Hop and could include proteins required for the full activation of Stat92E (e.g. by posttranslational modifications), proteins necessary for STAT translocation or proteins acting as transcriptional co-activators of Stat92E. The class, which exerts a phenotype under screening conditions as well as Hop^{TumI} stimulation but not under pathway induction by Upd-conditioned medium, cannot be explained in a linear pathway model and constitutes a group, for which further validation will be necessary.

Interestingly, the vast majority of *Drosophila* interactors have predicted homologs in *Homo sapiens*, many of which have been previously associated with human diseases. These putative human homologs were therefore tested for their functionality in human JAK/STAT signaling using a simplified system of STAT1 and STAT3 phosphorylation and transcriptional readouts. In this way, 30 interactors were identified which have a function for either STAT1 or STAT3 pathways in HeLa cells. Some of these had phenotypes opposite to what would have been expected from the phenotypes in *Drosophila* cells revealing possible compensatory mechanisms, which could be

explained, *inter alia*, by the escape of interferon response. The escape of interferon response has been described previously and has been linked to the activation of STAT5 (Jensen et al. 2005, Wellbrock et al. 2005). It is therefore possible that these compensatory effects are mediated by effectors acting on the activity of STAT5. It would also be very interesting to assess the effect of siRNAs with STAT5 dependent assays and in other cell lines to confirm this hypothesis. Unfortunately, a STAT5 reporter could not be established and the tested antibody against STAT5 detected multiple bands hampering the analysis on the level of posttranslational modifications. Moreover, it would be interesting to see to what extent the epistasis mapping of positive regulators in *Drosophila* would be in agreement with a similar categorization in human cells. This could, for example, be done by inducibly stimulating the pathway with a gain-of-function JAK allele, which has been described recently (James et al. 2005). Furthermore, epistasis analysis of the negative regulators could reveal additional valuable information. Although expression of both *GBP1* and *SOCS3* in HeLa cells is low in the absence of ligand and strongly upregulated following addition of IFN γ and OSM, it is possible that other JAK/STAT pathway independent mechanisms may also regulate the expression of these genes. Regulators of these independent mechanisms could be targeted for knockdown to produce a STAT independent, false-positive expression of *GBP1* or *SOCS3*. In order to exclude this possibility, it would be necessary, for example, to undertake double-knockdown experiments, in which the negative regulators of *GBP1* expression are targeted together with siRNAs designed to knockdown *STAT1*. Given that *STAT1* knockdown is sufficient to reduce IFN γ induced *GBP1* expression to basal levels, any remaining upregulation of *GBP1* in a *STAT1* knockdown background must occur via STAT1 independent mechanisms. Utilizing this approach, false-positive negative regulators could be excluded from the list of human JAK/STAT pathway modulators.

A pathway regulator found in three independent screens for JAK/STAT pathway modulators (Bach et al. 2003, Mukherjee et al. 2005, Müller et al. 2005) is CtBP. CtBP appears to have varying phenotypes in JAK/STAT signaling dependent on the alleles and methods used to score the phenotypes. For example, a deficiency removing *ctBP* was identified in the Bach screen as an enhancer, the Mukherjee screen describes enhancing and suppressing effects depending on the alleles used, and in the present study knock-

down of *ctBP* reveals an activity as a positive regulator epistatically downstream of Hop. A changing requirement for CtBP is not without precedent, as it has been previously characterized as a context-dependent transcription cofactor either activating or repressing transcription (Phippen et al. 2000). Furthermore, two mammalian homologs, CtBP1 and CtBP2, also exist (reviewed in Chinnadurai 2002), of which only CtBP2 appears to interact with JAK/STAT signaling (Figure 25). These CtBP proteins can be both nuclear and cytoplasmic, and a plant homolog has been shown to associate with the microtubule cytoskeleton (Kim et al. 2002) thereby opening the possibility of yet another mechanism by which CtBP could positively regulate JAK/STAT signaling. Another possibility could be an indirect action through Notch signaling, where CtBP represses the activation of Notch target genes (Barolo et al. 2002). Notch signaling in the eye disc can lead to the transcriptional activation of the JAK/STAT pathway ligand *upd* thereby indirectly inducing JAK/STAT signaling activity (Herz et al. 2006, Moberg et al. 2005, Vaccari and Bilder 2005), which could explain at least the forward genetic enhancer phenotype in JAK/STAT signaling.

One of the interesting questions unanswered before the present screen had been conducted was the involvement of endocytosis in JAK/STAT signaling and, if the pathway would be endocytically regulated, whether this would have a promoting or inhibiting role (Silver et al. 2005). The present RNAi screen reveals four novel *Drosophila* components probably involved in endocytosis, which also regulate JAK/STAT signaling (Rab5, Vps16B/CG18112, Mib2/CG17492 and TSG101). Interestingly, all of these have a negative regulatory role on JAK/STAT signal transduction, which is in contrast to other reports in vertebrate systems (Marchetti et al. 2006, Thiel et al. 1998). As expected, the homologs of these novel regulators also have a functional role in human JAK/STAT signaling, and all of them (RAB5A, LOC142678, C14ORF133 and TSG101) regulate STAT1 and STAT3 signaling negatively as do their *Drosophila* counterparts. Recently, a role for endocytosis in the regulation of Notch signaling has been described. For example, the ubiquitin ligase Mind bomb (Mib1) has recently been shown to promote Notch signaling by promoting the endocytosis of the pathway ligand. It is suspected that its paralog Mib2 also regulates these ligands (Lai et al. 2005), positively or negatively. Similarly, in cells with impaired TSG101 (Moberg et

al. 2005) or Vps25 function (Herz et al. 2006, Thompson et al. 2005, Vaccari and Bilder 2005), Notch is accumulated in intracellular compartments by endocytosis leading to increased signaling activity. Given that the JAK/STAT ligand Upd is a transcriptional target of the Notch pathway, and the effect that endocytosis has on Notch components, this could therefore also indirectly influence JAK/STAT signaling activity. It would be interesting to discriminate the effect of endocytosis to analyze whether the JAK/STAT pathway is directly endocytically regulated. This analysis would be especially interesting because endocytosis seems to negatively regulate JAK/STAT signaling and to positively regulate Notch signaling. Indeed, the option of an independent effect appears very likely since in the original *Drosophila* RNAi screen, endogenous levels of *upd* were very low and signaling was perhaps exclusively induced by ectopic *upd* expression. This experimental system, together with the ectopic induction of signaling by purified cytokines in HeLa cells, makes an indirect regulatory role of endocytosis via Notch signaling rather unlikely. If the role of endocytosis is directly acting on JAK/STAT signaling, it could be possible that this represents another mechanism to downregulate the pathway.

JAK/STAT signal transduction in *Drosophila* and vertebrate systems has previously been described to be regulated by negative feedback loops, where negative pathway regulators are direct transcriptional targets of the signaling cascade. This has been best demonstrated for the SOCS proteins (reviewed in Krebs and Hilton 2001). CG11501 was identified in the present study to regulate JAK/STAT signaling and could represent another member of a negative feedback loop mechanism. *CG11501* encodes a putatively secreted negative pathway regulator, which has previously been demonstrated to be a JAK/STAT pathway target gene (Boutros et al. 2002). A possible mechanism by which this extracellular negative regulator is acting could be at the level of the extracellular part of the pathway receptor or at the level of the pathway ligand itself. Both kinds of negative signal transduction regulation have been demonstrated before in Nodal signaling. For example, the negative regulator Lefty can antagonize Nodal signaling by binding to a coreceptor of Nodal (Cheng et al. 2004), and it can also attach to Nodal directly thereby blocking its binding to the receptor (Chen and Shen 2004). Studying *CG11501 in vivo*

and in greater detail holds great promise to potentially reveal a similar mechanism involved in JAK/STAT signaling.

Given the importance of posttranslational modifications and especially phosphorylation in JAK/STAT signal transduction, it seems likely that mechanisms to dephosphorylate pathway components have evolved as a regulatory mechanism. Another negative regulator identified in the screen is Ptp61F, a protein tyrosine phosphatase that has also been recovered from the Baeg et al. 2005 screen and which is expressed in both cytoplasmically and nuclearly localized splice forms. The closest mammalian homolog, PTP1B, has been implicated in JAK/STAT regulation before (Aoki and Matsuda 2000, Lund et al. 2005). Also the expression pattern of *ptp61F* (Ursuliak et al. 1997) is strikingly similar to that of *stat92E*, which could have led to the recognition of this functional link already many years before. Overexpression of the nuclearly localized *ptp61F* isoform both *in vivo* and in cultured cells specifically downregulates JAK/STAT signaling activity whereas the cytoplasmic isoform does not appear to be effective, indicating that Ptp61F antagonizes the pathway by dephosphorylating nuclearly localized Stat92E. This is in contrast to the observation of Baeg et al. 2005 that *ptp61F* loss-of-function leads to an increase in both phosphorylated Hop and Stat92E proteins. However, computational pathway modeling indicates that Ptp61F more likely acts on Stat92E directly, since pathway activity *in silico* can be modulated most by changing the levels of the nuclearly localized phosphatase downregulating STAT (Zi et al. 2005), and the increase in phosphorylated Hop protein may be due to reduced SOCS-mediated Hop degradation following a decrease in STAT activity. Further analysis of Ptp61F function shows that its gene activity is specifically induced after JAK/STAT pathway stimulation (Baeg et al. 2005) just like *socs36E*, demonstrating the presence of yet another feedback loop mechanism inside the cell to down-regulate signaling levels. Another phosphatase PP2A-B' also had a phenotype of a strong negative regulator upon knockdown of its gene activity in the genome-wide RNAi screen. Knockdown of its human homolog *PP2R5D* using pooled siRNAs leads to a similar phenotype for the expression levels of the JAK/STAT target genes *GBP1* and *SOCS3*, although the phenotype is less consistently reproducible with the individual siRNAs. *PP2R5D* is the delta isoform of the regulatory subunit B for the serine/threonine phosphatase PP2A. PP2A has been implicated

previously in the regulation of JAK/STAT signaling and it may act by dephosphorylating the serine-phosphorylated receptor protein Gp130 (Mitsuhashi et al. 2005). Inhibiting this phosphatase leads to the degradation of the receptor protein and in turn to higher transcription levels of *gp130* mRNA (Mitsuhashi et al. 2005). Furthermore, similar to the siRNA-mediated knockdown in the present study, longterm inhibition by a specific phosphatase inhibitor or an antisense construct of PP2A led to an increase in IL6-mediated transcriptional activity (Choi et al. 1998). Whether the effect of PP2R5D occurs through regulation of receptor mRNA levels, serine-phosphorylated STAT levels (Zhang et al. 1995) or via yet another mechanism remains to be investigated.

Many of the functional homologs modulating JAK/STAT pathway activity have been associated with human disease before and may constitute novel targets for therapeutic interventions. For example, the detailed knowledge of exact protein contact sites has already proven useful, when combinatorial chemical libraries were screened for inhibitors of dimerization of the leucine zipper proteins Myc/Max. The application of such an inhibitor eventually led to the successful blockage of transformation (Berg et al. 2002). One example of a human disease-associated novel negative regulator of JAK/STAT signaling is *bonus*, which encodes a single transcriptional cofactor and whose mammalian counterparts belong to the TIF1 (transcriptional intermediary factor 1) family of transcriptional coactivators and corepressors (Peng et al. 2002). Bonus is capable of binding nuclear hormone receptors (Beckstead et al. 2001) just like its mammalian counterpart TIF1 α (Le Douarin et al. 1995) acting as a coactivator. TIF1 β has been shown to act as a corepressor for some transcription factors (Zhong et al. 1999), whereas TIF1 γ (also known as TRIM33) may act by binding to and possibly ubiquitinating Smad4 thereby preventing or modulating TGF- β signaling (Dupont et al. 2005, He et al. 2006a). A zebrafish homolog of *bonus* and TIF1 γ has been identified and termed *moonshine*, which if mutated leads to disrupted embryonic and adult hematopoiesis (Ransom et al. 2004). Furthermore, human TIF1 γ is located on chromosome 1p13, a hotspot for chromosomal breakpoints associated with cancer (Johansson et al. 1994, Ng et al. 1999, Sawyer et al. 2002), and the fusion of TIF1 γ to a receptor tyrosine kinase leads to childhood papillary thyroid carcinomas (Klugbauer and Rabes 1999). Although Dupont et al. 2005 and He et al. 2006a show a role for TRIM33 in the control of differentiation

by TGF- β , these disease phenotypes would also be consistent with a role in JAK/STAT signaling. Recently, *bonus* has been also found to be involved in chromatin structure modulation (Beckstead et al. 2005). This finding is especially intriguing in the light of a recent study demonstrating the importance of the JAK/STAT pathway in regulating the cellular epigenetic status (Shi et al. 2006). As shown in the present study, human TRIM33 also regulates human JAK/STAT signaling, where it acts as a negative regulator of STAT1 signaling and maybe as a positive regulator of STAT3 signaling (Figure 25). This different regulatory role as a co-repressor or co-activator, respectively, has been described before and may be due to differences in hetero-oligomerization among the different members of the TIF1 family of transcriptional cofactors as, for example, the transcriptional coactivator TIF1 α has been shown to bind to the repressor TIF1 γ (Peng et al. 2002).

Another specific negative regulator of STAT transcriptional activity is *enok*, and its human homolog MOZ/MYST3 functions specifically by acting in the STAT1 and maybe in the STAT3 cascade. *MOZ* is a gene associated with acute myeloid leukemia (AML) resulting from chromosomal rearrangements that generate MOZ-CBP and MOZ-TIF2 fusion proteins (Borrow et al. 1996, Carapeti et al. 1999, Troke et al. 2006). Furthermore, recent analysis of the mouse loss-of-function *MOZ* allele showed that mutants die at day 15 of embryogenesis with severely reduced hematopoietic stem cells, lineage-committed progenitors and B-lineage cells as well as erythroid maturation defects and elevated myeloid lineage populations (Katsumoto et al. 2006). Given the siRNA-mediated knockdown findings in the present study, it seems possible that these loss-of-function developmental phenotypes are the consequence of a significant and specific increase in the level of STAT1 activity. It is tempting to hypothesize that the predicted enzymatic activity of MOZ as a MYST family histone acetyltransferase (Troke et al. 2006) may be of significance. Recent reports have shown that DNA binding and signaling of STAT3 is dependent on lysine acetylation (Yuan et al. 2005). Similar posttranslational modifications are present in STAT1 (Kramer et al. 2006), and it would be interesting to see if these are mediated by the MOZ acetyltransferase activity. Alternatively, a direct effect of MOZ on chromatin by changing the cellular epigenetic status could also be a

possible regulatory mechanism for JAK/STAT signaling (Hari et al. 2001, Shi et al. 2006).

The role of BRWD3 in the JAK/STAT pathways of flies and humans

Positive feedback loops regulating JAK/STAT pathway activity have also been discovered. For example, the expression of the pathway receptor (Brown et al. 2001) as well as of the expression of the STAT transcription factor (Ichiba et al. 1998) or its stability (Xi et al. 2003) may be upregulated following pathway stimulation. Although no such mechanisms have been identified in the present study, it remains likely that a subset of the novel identified positive pathway regulators are controlled in such a manner.

One strong positive regulator found with two independent dsRNAs in the genome-wide screen of the present study was *dBRWD3*, named after its human homolog, which has been found at the breakpoint of a translocation implicated in the development of B-cell chronic lymphocytic leukemia (B-CLL (Kalla et al. 2005)). *dBRWD3* has further been identified independently in an RNAi screen for JAK/STAT modulators (Baeg et al. 2005) and a chromosomal deficiency removing the region has also previously been recovered as a suppressor of *GMR-updΔ3'* (Bach et al. 2003). *Drosophila* BRWD3 possesses two homologs in humans – BRWD3 and WDR9 (also known as C21ORF107 and BRWD1), both of which have a similar protein domain structure and which only differ slightly in length (D'Costa et al. 2006). Predictions by Inparanoid to identify eukaryotic ortholog groups identify both human BRWD3 and WDR9 as orthologs of *Drosophila* BRWD3 with bootstrap values of 100% indicating that BRWD3 and WDR9 could be paralogs, although hBRWD3 is the best reciprocal BLAST hit for *dBRWD3*. Both human BRWD3 and WDR9 show a similar functional interaction in JAK/STAT signaling as their *Drosophila* counterpart, with both human BRWD3 and WDR9 specifically acting on STAT1 signaling as positive regulators and with an additive phenotype after double-knockdown of both. Furthermore, a number of genetic interaction experiments in *Drosophila* show that *dBRWD3* is also involved in regulating JAK/STAT activity *in vivo*. *dBRWD3* encodes a protein with a WD40-domain and two bromo-domains and is an essential gene in *Drosophila* as homozygous mutants die during larval life. Interestingly, overexpression of the full-length *dBRWD3* in cultured cells is not detrimental to neither

JAK/STAT signaling nor overall viability, whereas the ectopic expression in *Drosophila* tissues *in vivo* leads to malformations, indicating that a well balanced dosage of BRWD3 *in vivo* is necessary for the organism's survival. Epistasis analysis places dBRWD3 downstream of the JAK Hop, where it may physically interact with Stat92E in the nucleus. The lack of identification of an endogenous interaction between dBRWD3 and Stat92E analyzed by mass spectrometric analysis using immunoprecipitated complexes associated with epitope-tagged version of either BRWD3 or Stat92E does not exclude the possibility of a true physical interaction between the endogenous versions. The high abundance of DDB1 (damage-specific DNA-binding protein 1), co-migrating with the Stat92E band on a 1D SDS-PAGE in the BRWD3-Flag IP sample may have hampered identification of potentially extremely low abundant Stat92E in the BRWD3-Flag IP sample due to the limited detection time and automatic selection of the most abundant peptides in LC-MS/MS analysis. Moreover, the lack of identification of BRWD3 in the Stat92E-Myc IP sample may be due to the intrinsic sensitivity limits of LC-MS/MS for detection of a protein eluted from a polyacrylamide gel (0.5 – 1 fmol; Henning Urlaub, personal communication). To facilitate the identification of a less abundant protein in a pool of significantly more abundant proteins, peptides masses specific to Stat92E (from an *in silico* digest of Stat92E) can be put into an inclusion list of the mass spectrometer, which will increase the chance of sequencing these specific peptides during the automated LC-MS/MS run. Nevertheless, co-immunoprecipitation experiments indicate that two independent domains of dBRWD3, WD40- and the bromo-domains, are capable of binding Stat92E. Although further experiments are required to exclude epitope-tag mediated artefacts, the binding of two independent domains to another protein is not without precedent. STAT proteins themselves have been shown to bind to cofactors with more than one of their domains (reviewed in Bromberg 2001). For example, STAT1 can bind CBP/p300 with both its N-terminal and transactivation domain (Zhang et al. 1996). These distinct regions are distant in the primary sequence of the protein but may be in close proximity in the 3D structure.

A recent publication has further characterized BRWD3 as a chromatin-associated protein required for cell morphology in the developing *Drosophila* eye (D'Costa et al. 2006). Like Stat92E (Li et al. 2003a), dBRWD3 appears to be enriched in the developing

nervous system (D'Costa et al. 2006), and also the human and murine homologs *WDR9* show detectable expression in neural tissues in addition to expression in other tissues with a high proliferation rate (Ramos et al. 2002, Huang et al. 2003). dBRWD3 is not essential for general cell growth, viability or proliferation because mutant clones maintain the same size as their twospot after clonal induction in the larval eye disc. Furthermore, dBRWD3 is localized with active chromatin, although it does not appear to be a general transcription factor but only required for the transcription of some genes, as some RNA polymerase II occupied loci show no detectable dBRWD3 levels (D'Costa et al. 2006). The dBRWD3-associated subproteome shows that it can interact with the protein DDB1, the knockdown of which has a z-score of -1.8 in the genome-wide JAK/STAT RNAi screen. Therefore, knockdown of *DDB1* could be an unconfirmed positive regulatory phenotype missed by the threshold criteria for hit selection in the present study. DDB1 can function as a DNA repair protein in conjunction with the replication protein A (Wakasugi et al. 2001), which was also identified in the dBRWD3-associated complex (Table 14). However, more diverse roles other than in DNA repair have been suggested for DDB1 (Nichols et al. 2000) and described *in vivo*, e.g. in cellular proliferation and in the development of *Drosophila* and mouse (Takata et al. 2002, Cang et al. 2006). For example, knockdown of *DDB1* can induce melanotic tumors in *Drosophila* (Takata et al. 2004). Furthermore, DDB1 together with the transcription factor E2F may be required for the activity of some genes (Hayes et al. 1998) that are also controlled by STATs, such as *c-myc* and *cyclin D1* which contain STAT binding sites (Ehret et al. 2001) as well as potential E2F binding sites (Motokura and Arnold 1993, Hiebert et al. 1989). DDB1 can also reside in complexes with histone methyltransferases (Higa et al. 2006) and histone acetyltransferases and could play a role in chromatin remodeling (Martinez et al. 2001, Raptic-Otrin et al. 2002). For example, DDB1 is able to bind to the histone acetyltransferase p300 (Raptic-Otrin et al. 2002) just like STATs, which can interact with p300 via their N-terminal and their transactivation domain (reviewed in Bromberg et al. 2001). Interestingly, DDB1 has also been shown in recent studies to associate with the potential human homologs of dBRWD3, WDR9 (Angers et al. 2006) and PHIP (Angers et al. 2006, Jin et al. 2006), which were subsequently renamed to DCAF19 and DCAF14 for 'DDB1-Cul4-associated WD40 domain proteins'. The DDB1 and Cul4 proteins are

part of a cullin-RING ubiquitin ligase complex regulating DNA repair (Wang et al. 2006), DNA replication (Hu et al. 2004) and transcription (Wertz et al. 2004), and substrate specificity for the ubiquitin ligase may be governed by the use of particular WD40 proteins as molecular adaptors in substrate recognition (Angers et al. 2006, He et al. 2006b, Higa et al. 2006). These complexes can be hijacked by viruses to escape the interferon response by degrading STAT1 (reviewed in Horvath 2004) without DDB1 directly bound to STAT1 (Angers et al. 2006). Furthermore, murine WDR9 is also localized in the nucleus, excluded from pericentric heterochromatin and possesses a functional transcriptional activation domain between the WD40- and bromo-domains consisting of a polyglutamine-containing region (Huang et al. 2003). Taken these findings together, it appears likely that dBRWD3, DDB1 and Stat92E may cooperatively modulate transcriptional activity.

Analysis of BRWD3 in the *Drosophila* eye suggests that *BRWD3* is genetically upstream of Hedgehog and that a mutation can act as a dominant suppressor of Hedgehog loss-of-function (D'Costa et al. 2006). It could therefore be possible that this effect on Hedgehog signaling is due to upstream activation of hedgehog by the JAK/STAT pathway and BRWD3. For example, it has been shown in mice with temporal control of IFN γ expression that the mRNAs of *shh* and *gli-1* were increased along with known IFN γ regulated genes following induction of IFN γ transgene expression in the cerebellum, eventually leading to the ectopic expression of *shh* by granule neurons (Wang et al. 2004). Further, Wang et al. 2003 have demonstrated that *shh* and *gli-1* are specifically induced by IFN γ and not by IFN α in cultured granule neurons implying a role for the regulation of *shh* signaling in the CNS by IFN γ . It is therefore tempting to speculate that a similar role for the JAK/STAT cascade exists in *Drosophila*, maybe specifically in the eye. A role for BRWD3 also in other signaling systems other than the JAK/STAT cascade appears more unlikely as it has so far only been identified in RNAi screens for modulators of JAK/STAT signaling (for a comparison to other RNAi screen datasets see the FLIGHT database at <http://flight.licr.org>, Sims et al. 2006).

Late in embryogenesis, dBRWD3 is also localized in the germ cells (D'Costa et al. 2006). This is consistent with the phenotype seen when clones were induced in the *Drosophila*

germline leading to a loss of egg production possibly due to loss of germline stem cell maintenance. Additionally, the BRWD3 associated protein DDB1 is also localized in adult testis and ovaries (Takata et al. 2002). Whether a transfer of DDB1 from the nucleus to the cytosol shown during spermatogenesis (Takata et al. 2002) is of significance for the loss of pluripotency remains to be investigated. It could be possible that there is a general role for JAK/STAT signaling and BRWD3 in stem cell maintenance, for example in the neural tissue for neuroblast stem cell maintenance, where Stat92E, dBRWD3 and DDB1 are also expressed (Li et al. 2003a, D'Costa et al. 2006, Takata et al. 2002).

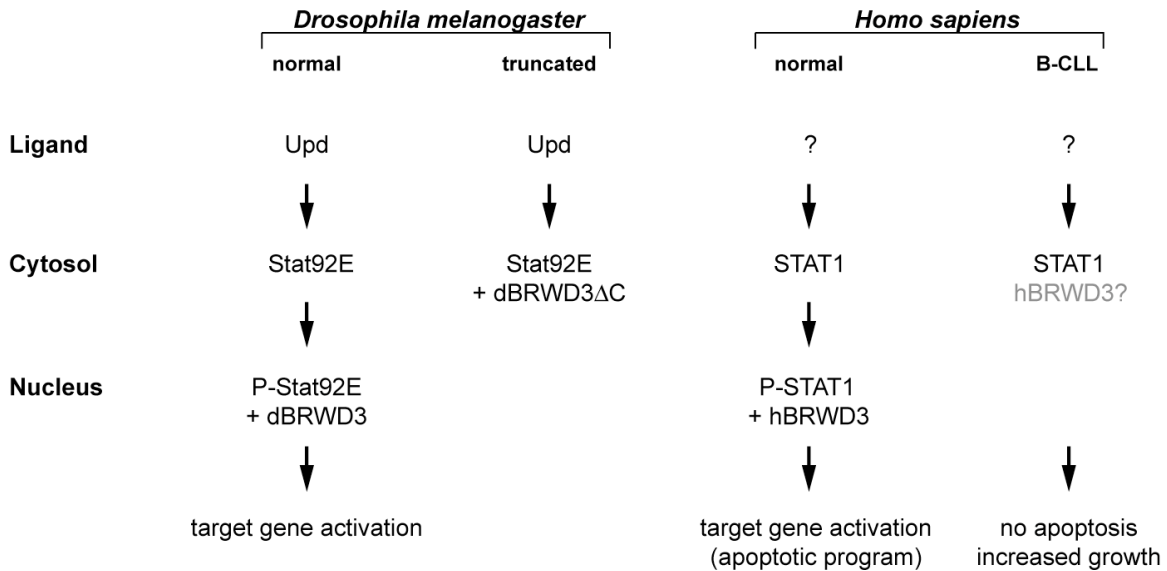


Figure 41. Model of BRWD3 molecular function. The pathway ligand normally induces the phosphorylation of *Drosophila* Stat92E. After translocation to the nucleus, P-Stat92E may interact with the nuclear localized dBRWD3 to activate the transcription of target genes ('normal'). If the C-terminus of BRWD3 including its NLS is truncated, BRWD3 Δ C is retained in the cytosol and JAK/STAT target genes cannot be activated ('truncated'). In *Homo sapiens*, a pathway ligand of unknown identity may normally induce STAT1 phosphorylation to activate target genes including apoptosis regulators. If hBRWD3 activity is compromised, the apoptosis program cannot be initiated leading to increased growth and tumor formation ('B-CLL'). hBRWD3 is shown in grey with a question mark to indicate that it could be present in the nucleus in reduced levels or trapped in the cytosol as a truncated version. Note that this is a speculative model.

Taken the data on BRWD3 and JAK/STAT signaling together, a speculative model can be derived of how BRWD3 acts to promote signaling in normal development as well as disease (Figure 41). In *Drosophila*, signaling is initiated by the binding of the Upd ligand to the Dome receptor. This leads to the phosphorylation of Stat92E and eventually to its

translocation to the nucleus. There, BRWD3 is associated to chromatin, maybe priming the chromatin or the promoters directly to become transcriptionally active with the binding of Stat92E leading to the activation of JAK/STAT target genes. In case the C-terminally located NLS is deleted in BRWD3 (BRWD3 Δ C), this protein is retained in the cytosol thereby preventing the successful activation of target genes even after stimulation of the JAK/STAT signaling pathway by Upd. It would be very interesting to analyze in further experiments whether the apparent dominant-negative effect of C-terminally truncated BRWD3 is caused by cytoplasmic trapping of Stat92E directly bound to the BRWD3 truncation, or whether this happens indirectly through a different factor or even via a different mechanism. A tempting speculation is that the onset of B-CLL in these patients is caused by the lack or reduced activity of BRWD3. While stimulation by a postulated STAT1 activating ligand is normally sufficient to turn on an apoptotic transcriptional program, it is not sufficient to prevent increased growth in these cells due to the lacking interaction of STAT1 with reduced hBRWD3 or due to cytoplasmic trapping of STAT1 by a C-terminal truncation of hBRWD3. These truncated versions of hBRWD3, similar to *Drosophila* BRWD3 Δ C, have been found in patients with B-CLL (Kalla et al. 2005; Claudia Kalla, personal communication). Similarly an escape of interferon response caused by overactive STAT5 has previously been reported to lead to the development of tumors (Wellbrock et al. 2005). Interestingly, the dosage reduction of *dBRWD3* in flies bearing a gain-of-function JAK leads to a reduction of JAK/STAT-induced tumor formation (Figure 31), whereas the reduced activity of *hBRWD3* in B-CLL patients correlates with the development of leukemia (Kalla et al. 2005). A possible explanation for these controversial phenotypes of reduced *BRWD3* levels is the functional diversity of *Drosophila* Stat92E (Mukherjee et al. 2005), which can exert both proliferative and anti-proliferative roles that are distributed to distinct STATs in mammals (e.g. STAT3 and STAT1). It would be very interesting to perform transcriptional profiling experiments in the future to identify the BRWD3 JAK/STAT pathway dependent target genes, e.g. in HeLa cells. This would help to clarify the mechanism of JAK/STAT pathway modulation by BRWD3 and to eventually definitively establish the proposed model, in which lack of apoptosis leads to tumorigenesis and leukemia.

SUMMARY AND CONCLUSIONS

From screening to function – this is how the present study can be best summarized. The aim of this study was to comprehensively identify novel modulators of JAK/STAT signaling using genome-wide RNAi screening. Novel methodologies were developed for high-throughput data analysis, novel selected candidates were mapped into the pathway and a significant proportion of the *Drosophila* candidates could be shown to have a functionally conserved role in mammalian systems. Two examples were analyzed *in vivo* to demonstrate their roles in regulating the pathway in *Drosophila*. One of these is the homolog of human *BRWD3*, a gene recently implicated in the development of B-cell chronic lymphocytic leukemia (B-CLL). Given the functional analysis of *dBRWD3* and the known roles for JAK/STAT signaling during normal hematopoiesis, it seems likely that a breakdown in BRWD3-mediated STAT regulation may represent a molecular mechanism leading to the development of B-CLL. The present study has generated a wealth of data and candidate lists, which further remain to be validated *in vivo*. From the basic research point of view, many interesting potential novel mechanisms have been discovered of how JAK/STAT signaling could be regulated positively and negatively, inside and outside the cell and at many levels of the signaling cascade. Aberrant JAK/STAT signaling has been implicated in multiple human malignancies and its components have been previously proposed as molecular targets for the development of therapeutic compounds. Therefore, also from the view of applied research, this study could well have implications for potential novel therapeutic approaches in the treatment of human disease. Thus, comprehensive genetic surveys by RNAi using *Drosophila* as a model organism represent a powerful approach for identifying targets relevant to human diseases.

REFERENCES

- Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen, R. C. Brandon, Y. H. Rogers, R. G. Blazej, M. Champe, B. D. Pfeiffer, K. H. Wan, C. Doyle, E. G. Baxter, G. Helt, C. R. Nelson, G. L. Gabor, J. F. Abril, A. Agbayani, H. J. An, C. Andrews-Pfannkoch, D. Baldwin, R. M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E. M. Beasley, K. Y. Beeson, P. V. Benos, B. P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M. R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K. C. Burtis, D. A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J. M. Cherry, S. Cawley, C. Dahlke, L. B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A. D. Mays, I. Dew, S. M. Dietz, K. Dodson, L. E. Doup, M. Downes, S. Dugan-Rocha, B. C. Dunkov, P. Dunn, K. J. Durbin, C. C. Evangelista, C. Ferraz, S. Ferreira, W. Fleischmann, C. Fosler, A. E. Gabrielian, N. S. Garg, W. M. Gelbart, K. Glasser, A. Glodek, F. Gong, J. H. Gorrell, Z. Gu, P. Guan, M. Harris, N. L. Harris, D. Harvey, T. J. Heiman, J. R. Hernandez, J. Houck, D. Hostin, K. A. Houston, T. J. Howland, M. H. Wei, C. Ibegwam, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287: 2185-95.
- Agaisse, H., U. M. Petersen, M. Boutros, B. Mathey-Prevot, and N. Perrimon. 2003. Signaling role of hemocytes in *Drosophila* JAK/STAT-dependent response to septic injury. *Dev Cell* 5: 441-50.
- Akira, S. 2000. Roles of STAT3 defined by tissue-specific gene targeting. *Oncogene* 19: 2607-11.
- Ali, S., Z. Nouhi, N. Chughtai, and S. Ali. 2003. SHP-2 regulates SOCS-1-mediated Janus kinase-2 ubiquitination/degradation downstream of the prolactin receptor. *J Biol Chem* 278: 52021-31.
- Allison, D. B., X. Cui, G. P. Page, and M. Sabripour. 2006. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7: 55-65.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J Mol Biol* 215: 403-10.
- Angers, S., T. Li, X. Yi, M. J. MacCoss, R. T. Moon, and N. Zheng. 2006. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* 443: 590-3.
- Aoki, N., and T. Matsuda. 2000. A cytosolic protein-tyrosine phosphatase PTP1B specifically dephosphorylates and deactivates prolactin-activated STAT5a and STAT5b. *J Biol Chem* 275: 39718-26.
- Arbouzova, N. I., and M. P. Zeidler MP. 2006. JAK/STAT signalling in *Drosophila*: insights into conserved regulatory and cellular functions. *Development* 133: 2605-16.
- Armknecht, S., M. Boutros, A. Kiger, K. Nybakken, B. Mathey-Prevot, and N. Perrimon. 2005. High-throughput RNA interference screens in *Drosophila* tissue culture cells. *Methods Enzymol* 392: 55-73.
- Arnone, M. I., I. J. Dmochowski, and C. Gache. 2004. Using reporter genes to study cis-regulatory elements. *Methods Cell Biol* 74: 621-52.
- Arziman, Z., T. Horn, and M. Boutros. 2005. E-RNAi: a web application to design optimized RNAi constructs. *Nucleic Acids Res* 33: W582-8.
- Ashburner, M. 1989. *Drosophila - A Laboratory Manual*. Cold Spring Harbour Laboratory Press.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, eds. 1999. *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc.
- Bach, E. A., S. Vincent, M. P. Zeidler, and N. Perrimon. 2003. A sensitized genetic screen to identify novel regulators and components of the *Drosophila* janus kinase/signal transducer and activator of transcription pathway. *Genetics* 165: 1149-66.
- Baeg, G. H., R. Zhou, and N. Perrimon. 2005. Genome-wide RNAi analysis of JAK/STAT signaling components in *Drosophila*. *Genes Dev* 19: 1861-70.
- Baksa, K., T. Parke, L. L. Dobens, and C. R. Dearolf. 2002. The *Drosophila* STAT protein, Stat92E, regulates follicle cell differentiation during oogenesis. *Dev Biol* 243: 166-75.
- Barolo, S., T. Stone, A. G. Bang, and J. W. Posakony. 2002. Default repression and Notch signaling: Hairless acts as an adaptor to recruit the corepressors Groucho and dCtBP to Suppressor of Hairless. *Genes Dev* 16: 1964-76.
- Beccari, S., L. Teixeira, and P. Rorth. 2002. The JAK/STAT pathway is required for border cell migration during *Drosophila* oogenesis. *Mech Dev* 111: 115-23.

- Becker, S., B. Groner, and C. W. Müller. 1998. Three-dimensional structure of the Stat3 β homodimer bound to DNA. *Nature* 394: 145-51.
- Beckstead, R., J. A. Ortiz, C. Sanchez, S. N. Prokopenko, P. Chambon, R. Losson, and H. J. Bellen. 2001. Bonus, a *Drosophila* homolog of TIF1 proteins, interacts with nuclear receptors and can inhibit betaFTZ-F1-dependent transcription. *Mol Cell* 7: 753-65.
- Beckstead, R. B., S. S. Ner, K. G. Hales, T. A. Grigliatti, B. S. Baker, and H. J. Bellen. 2005. Bonus, a *Drosophila* TIF1 homolog, is a chromatin-associated protein that acts as a modifier of position-effect variegation. *Genetics* 169: 783-94.
- Beller, M., and B. Oliver. 2006. One hundred years of high-throughput *Drosophila* research. *Chromosome Res* 14: 349-62.
- Berg, T., S. B. Cohen, J. Desharnais, C. Sonderegger, D. J. Maslyar, J. Goldberg, D. L. Boger, and P. K. Vogt. 2002. Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts. *Proc Natl Acad Sci U S A* 99: 3830-5.
- Betz, A., N. Lampen, S. Martinek, M. W. Young, and J. E. Darnell, Jr. 2001. A *Drosophila* PIAS homologue negatively regulates *stat92E*. *Proc Natl Acad Sci U S A* 98: 9563-9568.
- Biemar, F., D. A. Nix, J. Piel, B. Peterson, M. Ronshaugen, V. Sementchenko, I. Bell, J. R. Manak, and M. S. Levine. 2006. Comprehensive identification of *Drosophila* dorsal-ventral patterning genes using a whole-genome tiling array. *Proc Natl Acad Sci U S A* 103: 12763-8.
- Binari, R., and N. Perrimon. 1994. Stripe-specific regulation of pair-rule genes by hopscotch, a putative Jak family tyrosine kinase in *Drosophila*. *Genes Dev* 8: 300-12.
- Birmingham, A., E. M. Anderson, A. Reynolds, D. Ilsley-Tyree, D. Leake, Y. Fedorov, S. Baskerville, E. Maksimova, K. Robinson, J. Karpilow, W. S. Marshall, and A. Khvorova. 2006. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat Methods* 3: 199-204.
- Borrow, J., V. P. Stanton, Jr., J. M. Andresen, R. Becher, F. G. Behm, R. S. Chaganti, C. I. Civin, C. Distche, I. Dube, A. M. Frischauf, D. Horsman, F. Mitelman, S. Volinia, A. E. Watmore, and D. E. Housman. 1996. The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nat Genet* 14: 33-41.
- Boulay, J. L., J. J. O'Shea, and W. E. Paul. 2003. Molecular phylogeny within type I cytokines and their cognate receptors. *Immunity* 19: 159-63.
- Boutros, M., H. Agaisse, and N. Perrimon. 2002. Sequential activation of signaling pathways during innate immune responses in *Drosophila*. *Dev Cell* 3: 711-22.
- Boutros, M., L. Bras, and W. Huber. 2006. Analysis of cell-based RNAi screens. *Genome Biol* 7: R66.
- Boutros, M., A. A. Kiger, S. Armknecht, K. Kerr, M. Hild, B. Koch, S. A. Haas, R. Paro, and N. Perrimon. 2004. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 303: 832-5.
- Bowman, T., R. Garcia, J. Turkson, and R. Jove. 2000. STATs in oncogenesis. *Oncogene* 19: 2474-88.
- Brand, A. H., and N. Perrimon. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118: 401-15.
- Braunstein, J., S. Brutsaert, R. Olson, and C. Schindler. 2003. STATs dimerize in the absence of phosphorylation. *J Biol Chem* 278: 34133-40.
- Brawley, C., and E. Matunis. 2004. Regeneration of male germline stem cells by spermatogonial dedifferentiation *in vivo*. *Science* 304: 1331-4.
- Brideau, C., B. Gunter, B. Pikounis, and A. Liaw. 2003. Improved statistical methods for hit selection in high-throughput screening. *J Biomol Screen* 8: 634-47.
- Brierley, M. M., K. L. Marchington, I. Jurisica, and E. N. Fish. 2006. Identification of GAS-dependent interferon-sensitive target genes whose transcription is STAT2-dependent but ISGF3-independent. *FEBS J* 273: 1569-81.
- Britten, R. J., and E. H. Davidson. 1969. Gene regulation for higher cells: a theory. *Science* 165: 349-57.
- Brivanlou, A. H., and J. E. Darnell, Jr. 2002. Signal transduction and the control of gene expression. *Science* 295: 813-8.
- Bromberg, J. F. 2001. Activation of STAT proteins and growth control. *Bioessays* 23: 161-9.
- Brown, S., N. Hu, and J. C. Hombria. 2001. Identification of the first invertebrate interleukin JAK/STAT receptor, the *Drosophila* gene *domeless*. *Curr Biol* 11: 1700-5.

- Brown, S., N. Hu, and J. C. Hombria. 2003. Novel level of signalling control in the JAK/STAT pathway revealed by *in situ* visualisation of protein-protein interaction during *Drosophila* development. *Development* 130: 3077-84.
- Brown, S., M. P. Zeidler, and J. E. Hombria. 2006. JAK/STAT signalling in *Drosophila* controls cell motility during germ cell migration. *Dev Dyn* 235: 958-66.
- Cang, Y., J. Zhang, S. A. Nicholas, J. Bastien, B. Li, P. Zhou, and S.P. Goff. 2006. Deletion of DDB1 in mouse brain and lens leads to p53-dependent elimination of proliferating cells. *Cell* 127: 929-40.
- Carapeti, M., R. C. Aguiar, A. E. Watmore, J. M. Goldman, and N. C. Cross. 1999. Consistent fusion of MOZ and TIF2 in AML with inv(8)(p11q13). *Cancer Genet Cytogenet* 113: 70-2.
- Castelli-Gair Hombria, J., and S. Brown. 2002. The fertile field of *Drosophila* Jak/STAT signalling. *Curr Biol* 12: R569-75.
- Chen, C., and M. M. Shen. 2004. Two modes by which Lefty proteins inhibit nodal signaling. *Curr Biol* 14: 618-24.
- Chen, X., S. W. Oh, Z. Zheng, H. W. Chen, H. H. Shin, and S. X. Hou. 2003. Cyclin D-Cdk4 and cyclin E-Cdk2 regulate the Jak/STAT signal transduction pathway in *Drosophila*. *Dev Cell* 4: 179-90.
- Chen, X., U. Vinkemeier, Y. Zhao, D. Jeruzalmi, J. E. Darnell, Jr., and J. Kuriyan. 1998. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* 93: 827-39.
- Chen, Y., and G. Struhl. 1996. Dual roles for patched in sequestering and transducing Hedgehog. *Cell* 87: 553-63.
- Cheng, S. K., F. Olale, A. H. Brivanlou, and A. F. Schier. 2004. Lefty blocks a subset of TGFbeta signals by antagonizing EGF-CFC coreceptors. *PLoS Biol* 2: E30.
- Cherbas, P., L. Cherbas, and C. M. Williams. 1977. Induction of acetylcholinesterase activity by beta-ecdysone in a *Drosophila* cell line. *Science* 197: 275-77.
- Chien, S., L. T. Reiter, E. Bier, and M. Gribskov. 2002. Homophila: human disease gene cognates in *Drosophila*. *Nucleic Acids Res* 30: 149-51.
- Chinnadurai, G. 2002. CtBP, an unconventional transcriptional corepressor in development and oncogenesis. *Mol Cell* 9: 213-24.
- Choi, I., M. J. Lee, E. J. Kim, H. S. Kang, and K. H. Pyun. 1998. Roles of protein phosphatase 2A in IL-6 signal transduction in Hep3B cells. *Immunol Lett* 61: 103-7.
- Chou, T. B., and N. Perrimon. 1996. The autosomal FLP-DFS technique for generating germline mosaics in *Drosophila melanogaster*. *Genetics* 144: 1673-9.
- Chuang, P. T., and T. B. Kornberg. 2000. On the range of hedgehog signaling. *Curr Opin Genet Dev* 10: 515-22.
- Clamp, M., D. Andrews, D. Barker, P. Bevan, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, R. Durbin, E. Eyra, J. Gilbert, M. Hammond, T. Hubbard, A. Kasprzyk, D. Keefe, H. Lehvaslaiho, V. Iyer, C. Melsopp, E. Mongin, R. Pettett, S. Potter, A. Rust, E. Schmidt, S. Searle, G. Slater, J. Smith, W. Spooner, A. Stabenau, J. Stalker, E. Stupka, A. Ureta-Vidal, I. Vastrik, and E. Birney. 2003. Ensembl 2002: accommodating comparative genomics. *Nucleic Acids Res* 31: 38-42.
- Clemens, J. C., C. A. Worby, N. Simonson-Leff, M. Muda, T. Maehama, B. A. Hemmings, and J. E. Dixon. 2000. Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc Natl Acad Sci U S A* 97: 6499-6503.
- Cogoni, C., J. T. Irelan, M. Schumacher, T. J. Schmidhauser, E. U. Selker, and G. Macino. 1996. Transgene silencing of the al-1 gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *EMBO J* 15: 3153-63.
- Collins, M. L., B. Irvine, D. Tyner, E. Fine, C. Zayati, C. Chang, T. Horn, D. Ahle, J. Detmer, L. Shen, J. Kolberg, S. Bushnell, M. S. Urdea, and D. D. Ho. 1997. A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. *Nucleic Acids Res* 25: 2989-84.
- Collum, R. G., S. Brutsaert, G. Lee, and C. Schindler. 2000. A Stat3-interacting protein (StIP1) regulates cytokine signal transduction. *Proc Natl Acad Sci U S A* 97: 10120-5.
- D'Costa, A., R. Reifegerste, S. Sierra, and K. Moses. 2006. The *Drosophila ramshackle* gene encodes a chromatin-associated protein required for cell morphology in the developing eye. *Mech Dev* 123: 591-604.

- Darnell, J. E., Jr. 1997. STATs and gene regulation. *Science* 277: 1630-5.
- Darnell, J. E., Jr. 2002. Transcription factors as targets for cancer therapy. *Nat Rev Cancer* 2: 740-9.
- DasGupta, R., A. Kaykas, R. T. Moon, and N. Perrimon. 2005. Functional genomic analysis of the Wnt-wingless signaling pathway. *Science* 308: 826-33.
- David, M., E. Petricoin, 3rd, C. Benjamin, R. Pine, M. J. Weber, and A. C. Larner. 1995. Requirement for MAP kinase (ERK2) activity in interferon alpha- and interferon beta-stimulated gene expression through STAT proteins. *Science* 269: 1721-3.
- Decker, T., and P. Kovarik. 2000. Serine phosphorylation of STATs. *Oncogene* 19: 2628-37.
- Decotto, E., and A. C. Spradling. 2005. The *Drosophila* ovarian and testis stem cell niches: similar somatic stem cells and signals. *Dev Cell* 9: 501-10.
- Drysdale, R. A., M. A. Crosby, W. Gelbart, K. Campbell, D. Emmert, B. Matthews, S. Russo, A. Schroeder, F. Smutniak, P. Zhang, P. Zhou, M. Zytovicz, M. Ashburner, A. de Grey, R. Foulger, G. Millburn, D. Sutherland, C. Yamada, T. Kaufman, K. Matthews, A. DeAngelo, R. K. Cook, D. Gilbert, J. Goodman, G. Grumblin, H. Sheth, V. Strelets, G. Rubin, M. Gibson, N. Harris, S. Lewis, S. Misra, and S. Q. Shu. 2005. FlyBase: genes and gene models. *Nucleic Acids Res* 33: D390-5.
- Dupont, S., L. Zacchigna, M. Cordenonsi, S. Soligo, M. Adorno, M. Rugge, and S. Piccolo. 2005. Germ-layer specification and control of cell growth by Ectodermin, a Smad4 ubiquitin ligase. *Cell* 121: 87-99.
- Durbin, J. E., R. Hackenmiller, M. C. Simon, and D. E. Levy. 1996. Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. *Cell* 84: 443-50.
- Echalier, G., and A. Ohanessian. 1970. *In vitro* culture of *Drosophila melanogaster* embryonic cells. *In Vivo* 6: 162-72.
- Echeverri, C. J., P. A. Beachy, B. Baum, M. Boutros, F. Buchholz, S. K. Chanda, J. Downward, J. Ellenberg, A. G. Fraser, N. Hacohen, W. C. Hahn, A. L. Jackson, A. Kiger, P. S. Linsley, L. Lum, Y. Ma, B. Mathey-Prevot, D. E. Root, D. M. Sabatini, J. Taipale, N. Perrimon, and R. Bernards. 2006. Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat Methods* 3: 777-9.
- Echeverri, C. J., and N. Perrimon. 2006. High-throughput RNAi screening in cultured cells: a user's guide. *Nat Rev Genet* 7: 373-84.
- Ehret, G. B., P. Reichenbach, U. Schindler, C. M. Horvath, S. Fritz, M. Nabholz, and P. Bucher. 2001. DNA binding specificity of different STAT proteins. Comparison of *in vitro* specificity with natural target sites. *J Biol Chem* 276: 6675-88.
- Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411: 494-8.
- Elbashir, S. M., J. Harborth, K. Weber, and T. Tuschl. 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26: 199-213.
- Evan, G. I., G. K. Lewis, G. Ramsay, and J. M. Bishop. 1985. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol Cell Biol* 5: 3610-6.
- Fagerlund, R., K. Melen, L. Kinnunen, and I. Julkunen. 2002. Arginine/lysine-rich nuclear localization signals mediate interactions between dimeric STATs and importin alpha 5. *J Biol Chem* 277: 30072-8.
- Fedorov, Y., E. M. Anderson, A. Birmingham, A. Reynolds, J. Karpilow, K. Robinson, D. Leake, W. S. Marshall, and A. Khvorova. 2006. Off-target effects by siRNA can induce toxic phenotype. *RNA* 12: 1188-96.
- Feinberg, E. H., and C. P. Hunter. 2003. Transport of dsRNA into cells by the transmembrane protein SID-1. *Science* 301: 1545-7.
- Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391: 806-11.
- Fraser, A. G., R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann, and J. Ahringer. 2000. Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* 408: 325-30.
- Freeman, M. 1996. Reiterative use of the EGF receptor triggers differentiation of all cell types in the *Drosophila* eye. *Cell* 87: 651-60.

- Freeman, M., and J. B. Gurdon. 2002. Regulatory principles of developmental signaling. *Annu Rev Cell Dev Biol* 18: 515-39.
- Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
- Ghigliione, C., O. Devergne, E. Georgenthum, F. Carballes, C. Medioni, D. Cerezo, and S. Noselli. 2002. The *Drosophila* cytokine receptor Domeless controls border cell migration and epithelial polarization during oogenesis. *Development* 129: 5437-47.
- Goenka, S., C. Marlar, U. Schindler, and M. Boothby. 2003. Differential roles of C-terminal activation motifs in the establishment of Stat6 transcriptional specificity. *J Biol Chem* 278: 50362-70.
- Gonczy, P., C. Echeverri, K. Oegema, A. Coulson, S. J. Jones, R. R. Copley, J. Duperon, J. Oegema, M. Brehm, E. Cassin, E. Hannak, M. Kirkham, S. Pichler, K. Flohrs, A. Goessen, S. Leidel, A. M. Alleaume, C. Martin, N. Ozlu, P. Bork, and A. A. Hyman. 2000. Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* 408: 331-36.
- Gould, S. J., and S. Subramani. 1988. Firefly luciferase as a tool in molecular and cell biology. *Anal Biochem* 175: 5-13.
- Gruber, J., T. Lampe, M. Osborn, and K. Weber. 2005. RNAi of FACE1 protease results in growth inhibition of human cells expressing lamin A: implications for Hutchinson-Gilford progeria syndrome. *J Cell Sci* 118: 689-96.
- Grünenfelder, B. and E. A. Winzeler. 2002. Treasures and traps in genome-wide data sets: case examples from yeast. *Nat Rev Genet* 3: 653-61.
- Gunter, B., C. Brideau, B. Pikounis, and A. Liaw. 2003. Statistical and graphical methods for quality control determination of high-throughput screening data. *J Biomol Screen* 8: 624-33.
- Hannon, G. J. 2002. RNA interference. *Nature* 418: 244-51.
- Hao, J., T. G. Li, X. Qi, D. F. Zhao, and G. Q. Zhao. 2006. WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev Biol* 290: 81-91.
- Hari, K. L., K. R. Cook, and G. H. Karpen. 2001. The *Drosophila* *Su(var)2-10* locus regulates chromosome structure and function and encodes a member of the PIAS protein family. *Genes Dev* 15: 1334-48.
- Harris, M. A., J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, K. Eilbeck, S. Lewis, B. Marshall, C. Mungall, J. Richter, G. M. Rubin, J. A. Blake, C. Bult, M. Dolan, H. Drabkin, J. T. Eppig, D. P. Hill, L. Ni, M. Ringwald, R. Balakrishnan, J. M. Cherry, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. Engel, D. G. Fisk, J. E. Hirschman, E. L. Hong, R. S. Nash, A. Sethuraman, C. L. Theesfeld, D. Botstein, K. Dolinski, B. Feierbach, T. Berardini, S. Mundodi, S. Y. Rhee, R. Apweiler, D. Barrell, E. Camon, E. Dimmer, V. Lee, R. Chisholm, P. Gaudet, W. Kibbe, R. Kishore, E. M. Schwarz, P. Sternberg, M. Gwinn, L. Hannick, J. Wortman, M. Berriman, V. Wood, N. de la Cruz, P. Tonellato, P. Jaiswal, T. Seigfried, and R. White. 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32: D258-61.
- Harrison, D. A., R. Binari, T. S. Nahreini, M. Gilman, and N. Perrimon. 1995. Activation of a *Drosophila* Janus kinase (JAK) causes hematopoietic neoplasia and developmental defects. *Embo J* 14: 2857-65.
- Harrison, D. A., P. E. McCoon, R. Binari, M. Gilman, and N. Perrimon. 1998. *Drosophila unpaired* encodes a secreted protein that activates the JAK signaling pathway. *Genes Dev* 12: 3252-63.
- Hayes S., P. Shiyonov, X. Chen, and P. Raychaudhuri. 1998. DDB, a putative DNA repair protein, can function as a transcriptional partner of E2F1. *Mol Cell Biol* 18: 240-9.
- He, W., D. C. Dorn, H. Erdjument-Bromage, P. Tempst, M. A. Moore, and J. Massague. 2006a. Hematopoiesis controlled by distinct TIF1gamma and Smad4 branches of the TGFbeta pathway. *Cell* 125: 929-41.
- He, Y. J., C. M. McCall, J. Hu, Y. Zeng, and Y. Xiong. 2006b. DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4-ROC1 ubiquitin ligases. *Genes Dev* 20: 2949-54.
- Heddad, A., M. Brameier, and R. M. MacCallum. 2004. Evolving Regular Expression-Based Sequence Classifiers for Protein Nuclear Localisation. Pages 31-40. *Lecture Notes in Computer Science*.

- Henriksen, M. A., A. Betz, M. V. Fuccillo, and J. E. Darnell. 2002. Negative regulation of STAT92E by an N-terminally truncated STAT protein derived from an alternative promoter site. *Genes Dev* 16: 2379-2389.
- Herz, H. M., Z. Chen, H. Scherr, M. Lackey, C. Bolduc, and A. Bergmann. 2006. vps25 mosaics display non-autonomous cell survival and overgrowth, and autonomous apoptosis. *Development* 133: 1871-80.
- Hiebert, S. W., M. Lipp, and J. R. Nevins. 1989. E1A-dependent transactivation of the human MYC promoter is mediated by the E2F factor. *Proc Natl Acad Sci USA* 86: 3594-8.
- Higa, L. A., M. Wu, T. Ye, R. Kobayashi, H. Sun, and H. Zhang. 2006. CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat Cell Biol* 8: 1277-83.
- Hild, M., B. Beckmann, S. A. Haas, B. Koch, V. Solovyev, C. Busold, K. Fellenberg, M. Boutros, M. Vingron, F. Sauer, J. D. Hoheisel, and R. Paro. 2003. An integrated gene annotation and transcriptional profiling approach towards the full gene content of the *Drosophila* genome. *Genome Biol* 5: R3.
- Holley, S. A., P. D. Jackson, Y. Sasai, B. Lu, E. M. De Robertis, F. M. Hoffmann, and E. L. Ferguson. 1995. A conserved system for dorsal-ventral patterning in insects and vertebrates involving sog and chordin. *Nature* 376: 249-53.
- Hombria, J. C., S. Brown, S. Häder, and M. P. Zeidler. 2005. Characterisation of Upd2, a *Drosophila* JAK/STAT pathway ligand. *Dev Biol* 288: 420-33.
- Horvath, C. M. 2004. Weapons of STAT destruction. Interferon evasion by paramyxovirus V protein. *Eur J Biochem* 271: 4621-8.
- Hou, X. S., M. B. Melnick, and N. Perrimon. 1996. Marelle acts downstream of the *Drosophila* HOP/JAK kinase and encodes a protein similar to the mammalian STATs. *Cell* 84: 411-9.
- Hou, X. S., Z. Zheng, X. Chen, and N. Perrimon. 2002. The Jak/STAT pathway in model organisms: emerging roles in cell movement. *Dev Cell* 3: 765-78.
- Hu, J., C. M. McCall, T. Ohta, and Y. Xiong. 2004. Targeted ubiquitination of CDT1 by the DDB1-Cul4-ROC1 ligase in response to DNA damage. *Nature Cell Biol* 6: 1003-9.
- Huang, H., I. Rambaldi, E. Daniels, and M. Featherstone. 2003. Expression of the *Wdr9* gene and protein products during mouse development. *Dev Dyn* 227: 608-14.
- Ichiba, M., K. Nakajima, Y. Yamanaka, N. Kiuchi, and T. Hirano. 1998. Autoregulation of the Stat3 gene through cooperation with a cAMP-responsive element-binding protein. *J Biol Chem* 273: 6132-8.
- Ingelbrecht, I., H. Van Houdt, M. Van Montagu, and A. Depicker. 1994. Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. *Proc Natl Acad Sci USA* 91: 10502-6.
- Inoue, H., H. Nojima, and H. Okayama. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96: 23-8.
- James, C., V. Ugo, J. P. Le Couedic, J. Staerk, F. Delhommeau, C. Lacout, L. Garcon, H. Raslova, R. Berger, A. Bennaceur-Griscelli, J. L. Villeval, S. N. Constantinescu, N. Casadevall, and W. Vainchenker. 2005. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* 434: 1144-8.
- Jensen, J., E. D. Galsgaard, A. E. Karlsen, Y. C. Lee, and J. H. Nielsen. 2005. STAT5 activation by human GH protects insulin-producing cells against interleukin-1beta, interferon-gamma and tumour necrosis factor-alpha-induced apoptosis independent of nitric oxide production. *J Endocrinol* 187: 25-36.
- Jin, J., E. E. Arias, J. Chen, J. W. Harper, and J. C. Walter. 2006. A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol Cell* 23: 709-21.
- Johansson, M., C. Dietrich, N. Mandahl, G. Hambreus, L. Johansson, P. P. Clausen, F. Mitelman, and S. Heim. 1994. Karyotypic characterization of bronchial large cell carcinomas. *Int J Cancer* 57: 463-7.
- John, S., U. Vinkemeier, E. Soldaini, J. E. Darnell, Jr., and W. J. Leonard. 1999. The significance of tetramerization in promoter recruitment by Stat5. *Mol Cell Biol* 19: 1910-8.
- Josten, F., B. Fuss, M. Feix, T. Meissner, and M. Hoch. 2004. Cooperation of JAK/STAT and Notch signaling in the *Drosophila* foregut. *Dev Biol* 257: 181-9.

- Kalla, C., H. Nentwich, M. Schlotter, D. Mertens, K. Wildenberger, H. Dohner, S. Stilgenbauer, and P. Lichter. 2005. Translocation t(X;11)(q13;q23) in B-cell chronic lymphocytic leukemia disrupts two novel genes. *Genes Chromosomes Cancer* 42: 128-43.
- Kamakura, S., K. Oishi, T. Yoshimatsu, M. Nakafuku, N. Masuyama, and Y. Gotoh. 2004. Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signalling. *Nat Cell Biol* 6: 547-54.
- Kamath, R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, M. Gotta, A. Kanapin, N. Le Bot, S. Moreno, M. Sohrmann, D. P. Welchman, P. Zipperlen, and J. Ahringer. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421: 231-7.
- Karsten, P., S. Häder, and M. P. Zeidler. 2002. Cloning and expression of *Drosophila* SOCS36E and its potential regulation by the JAK/STAT pathway. *Mech Dev* 117: 343-6.
- Karsten, P., I. Plischke, N. Perrimon, and M. P. Zeidler. 2005. Mutational analysis reveals separable DNA binding and trans-activation of *Drosophila* STAT92E. *Cell Signal* 18: 819-29.
- Katsumoto, T., Y. Aikawa, A. Iwama, S. Ueda, H. Ichikawa, T. Ochiya, and I. Kitabayashi. 2006. MOZ is essential for maintenance of hematopoietic stem cells. *Genes Dev* 20: 1321-30.
- Kennerdell, J. R., and R. W. Carthew. 1998. Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* 95: 1017-26.
- Kiger, A. A., B. Baum, S. Jones, M. R. Jones, A. Coulson, C. Echeverri, and N. Perrimon. 2003. A functional genomic analysis of cell morphology using RNA interference. *J Biol* 2: 27.
- Kiger, A. A., D. L. Jones, C. Schulz, M. B. Rogers, and M. T. Fuller. 2001. Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue. *Science* 294: 2542-2545.
- Kim, G. T., K. Shoda, T. Tsuge, K. H. Cho, H. Uchimiya, R. Yokoyama, K. Nishitani, and H. Tsukaya. 2002. The ANGUSTIFOLIA gene of Arabidopsis, a plant CtBP gene, regulates leaf-cell expansion, the arrangement of cortical microtubules in leaf cells and expression of a gene involved in cell-wall formation. *Embo J* 21: 1267-79.
- Kim, T. K., and T. Maniatis. 1996. Regulation of interferon-gamma-activated STAT1 by the ubiquitin-proteasome pathway. *Science* 273: 1717-9.
- Kittler, R., G. Putz, L. Pelletier, I. Poser, A. K. Heninger, D. Drechsel, S. Fischer, I. Konstantinova, B. Habermann, H. Grabner, M. L. Yaspo, H. Himmelbauer, B. Korn, K. Neugebauer, M. T. Pisabarro, and F. Buchholz. 2004. An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* 432: 1036-40.
- Klugbauer, S., and H. M. Rabes. 1999. The transcription coactivator HTIF1 and a related protein are fused to the RET receptor tyrosine kinase in childhood papillary thyroid carcinomas. *Oncogene* 18: 4388-93.
- Komyod, W., U. M. Bauer, P. C. Heinrich, S. Haan, and I. Behrmann. 2005. Are STATs arginine-methylated? *J Biol Chem* 280: 21700-5.
- Kramer, O. H., D. Baus, S. K. Knauer, S. Stein, E. Jager, R. H. Stauber, M. Grez, E. Pfitzner, and T. Heinzel. 2006. Acetylation of Stat1 modulates NF-kappaB activity. *Genes Dev* 20: 473-85.
- Krebs, D. L., and D. J. Hilton. 2001. SOCS proteins: negative regulators of cytokine signaling. *Stem Cells* 19: 378-387.
- Kulkarni, M. M., M. Booker, S. J. Silver, A. Friedman, P. Hong, N. Perrimon, and B. Mathey-Prevot. 2006. Evidence of off-target effects associated with long dsRNAs in *Drosophila melanogaster* cell-based assays. *Nat Methods* 3: 833-8.
- Kurzer, J. H., P. Saharinen, O. Silvennoinen, and C. Carter-Su. 2006. Binding of SH2-B family members within a potential negative regulatory region maintains JAK2 in an active state. *Mol Cell Biol* 26: 6381-94.
- Kwon, E. J., H. S. Park, Y. S. Kim, E. J. Oh, Y. Nishida, A. Matsukage, M. A. Yoo, and M. Yamaguchi. 2000. Transcriptional regulation of the *Drosophila* raf proto-oncogene by *Drosophila* STAT during development and in immune response. *J Biol Chem* 275: 19824-30.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-5.
- Lagueux, M., E. Perrodou, E. A. Levashina, M. Capovilla, and J. A. Hoffmann. 2000. Constitutive expression of a complement-like protein in *toll* and JAK gain-of-function mutants of *Drosophila*. *Proc Natl Acad Sci U S A* 97: 11427-32.

- Lai, E. C., F. Roegiers, X. Qin, Y. N. Jan, and G. M. Rubin. 2005. The ubiquitin ligase *Drosophila* Mind bomb promotes Notch signaling by regulating the localization and activity of Serrate and Delta. *Development* 132: 2319-32.
- Le Douarin, B., C. Zechel, J. M. Garnier, Y. Lutz, L. Tora, P. Pierrat, D. Heery, H. Gronemeyer, P. Chambon, and R. Losson. 1995. The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *Embo J* 14: 2020-33.
- Leaman, D. W., S. Leung, X. Li, and G. R. Stark. 1996. Regulation of STAT-dependent pathways by growth factors and cytokines. *Faseb J* 10: 1578-88.
- Lee, Y. S., K. Nakahara, J. W. Pham, K. Kim, Z. He, E. J. Sontheimer, and R. W. Carthew. 2004. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* 117: 69-81.
- Lengyel, J. A., and D. D. Iwaki. 2002. It takes guts: the *Drosophila* hindgut as a model system for organogenesis. *Dev Biol* 243: 1-19.
- Lerner, L., M. A. Henriksen, X. Zhang, and J. E. Darnell, Jr. 2003. STAT3-dependent enhanceosome assembly and disassembly: synergy with GR for full transcriptional increase of the alpha 2-macroglobulin gene. *Genes Dev* 17: 2564-77.
- Levy, D. E., and J. E. Darnell, Jr. 2002. Stats: transcriptional control and biological impact. *Nat Rev Mol Cell Biol* 3: 651-62.
- Li, J., W. Li, H. C. Calhoun, F. Xia, F. B. Gao, and W. X. Li. 2003a. Patterns and functions of STAT activation during *Drosophila* embryogenesis. *Mech Dev* 120: 1455-68.
- Li, J., F. Xia, and W. X. Li. 2003b. Coactivation of STAT and Ras is required for germ cell proliferation and invasive migration in *Drosophila*. *Dev Cell* 5: 787-98.
- Lindsley, D. L., and G. G. Zimm. 1992. The Genome of *Drosophila melanogaster*. Academic Press.
- Long, J., I. Matsuura, D. He, G. Wang, K. Shuai, and F. Liu. 2003. Repression of Smad transcriptional activity by PIASy, an inhibitor of activated STAT. *Proc Natl Acad Sci U S A* 100: 9791-6.
- Lund, I. K., J. A. Hansen, H. S. Andersen, N. P. Moller, and N. Billestrup. 2005. Mechanism of protein tyrosine phosphatase 1B-mediated inhibition of leptin signalling. *J Mol Endocrinol* 34: 339-51.
- Luo, H., W. P. Hanratty, and C. R. Dearolf. 1995. An amino acid substitution in the *Drosophila* hop^{Tum-1} Jak kinase causes leukemia-like hematopoietic defects. *Embo J* 14: 1412-20.
- Luo, H., P. Rose, D. Barber, W. P. Hanratty, S. Lee, T. M. Roberts, A. D. D'Andrea, and C. R. Dearolf. 1997. Mutation in the Jak kinase JH2 domain hyperactivates *Drosophila* and mammalian Jak-Stat pathways. *Mol Cell Biol* 17: 1562-71.
- Ma, Y., A. Creanga, L. Lum, and P. A. Beachy. 2006. Prevalence of off-target effects in *Drosophila* RNA interference screens. *Nature* 443: 359-63.
- Malo, N., J. A. Hanley, S. Cerquozzi, J. Pelletier, and R. Nadon. 2006. Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* 24: 167-75.
- Manche, L., S. R. Green, C. Schmedt, and M. B. Mathews. 1992. Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol Cell Biol* 12: 5238-48.
- Marchetti, M., M. N. Monier, A. Fradagrada, K. Mitchell, F. Baychelier, P. Eid, L. Johannes, and C. Lamaze. 2006. Stat-mediated signaling induced by type I and type II interferons (IFNs) is differentially controlled through lipid microdomain association and clathrin-dependent endocytosis of IFN receptors. *Mol Biol Cell* 17: 2896-909.
- Martinez E., V.B. Palhan, A. Tjernberg, E.S. Lyman, A. M. Gamper, T.K. Kundu, B. T. Chait, and R. G. Roeder. 2001. Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors *in vivo*. *Mol Cell Biol* 21: 6782-95.
- Matsuda, T., T. Nakamura, K. Nakao, T. Arai, M. Katsuki, T. Heike, and T. Yokota. 1999. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *Embo J* 18: 4261-9.
- McBride, K. M., and N. C. Reich. 2003. Nuclear trafficking of STAT proteins. Pages 269-283 in P. B. Sehgal, D. E. Levy, and T. Hirano, eds. *Signal Transducers and Activators of Transcription (STATs); Activation and Biology*. Kluwer Academic Publishers.
- McGregor, J. R., R. Xi, and D. A. Harrison. 2002. JAK signaling is somatically required for follicle cell differentiation in *Drosophila*. *Development* 129: 705-17.

- McLaughlin, S., and J. E. Dixon. 1993. Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in *Drosophila*. *J Biol Chem* 268: 6839-42.
- Meissner, T., E. Krause, I. Lödige, and U. Vinkemeier. 2004. Arginine Methylation of STAT1: a reassessment. *Cell* 119: 587-9.
- Meister, G., and T. Tuschl. 2004. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431: 343-9.
- Meister, M., and M. Lagueux. 2003. *Drosophila* blood cells. *Cell Microbiol* 5: 573-80.
- Mertens, C., M. Zhong, R. Krishnaraj, W. Zou, X. Chen, and J. E. Darnell, Jr. 2006. Dephosphorylation of phosphotyrosine on STAT1 dimers requires extensive spatial reorientation of the monomers facilitated by the N-terminal domain. *Genes Dev* 20: 3372-81.
- Meyer, T., and U. Vinkemeier. 2004. Nucleocytoplasmic shuttling of STAT transcription factors. *Eur J Biochem* 271: 4606-12.
- Minks, M. A., D. K. West, S. Benvin, and C. Baglioni. 1979. Structural requirements of double-stranded RNA for the activation of 2',5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. *J Biol Chem* 254: 10180-3.
- Misra, S., M. A. Crosby, C. J. Mungall, B. B. Matthews, K. S. Campbell, P. Hradecky, Y. Huang, J. S. Kaminker, G. H. Millburn, S. E. Prochnik, C. D. Smith, J. L. Tupy, E. J. Whitfield, L. Bayraktaroglu, B. P. Berman, B. R. Bettencourt, S. E. Celniker, A. D. de Grey, R. A. Drysdale, N. L. Harris, J. Richter, S. Russo, A. J. Schroeder, S. Q. Shu, M. Stapleton, C. Yamada, M. Ashburner, W. M. Gelbart, G. M. Rubin, and S. E. Lewis. 2002. Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol* 3: RESEARCH0083.
- Mitsuhashi, S, H. Shima, N. Tanuma, S. Sasa, K. Onoe, M. Ubukata, and K. Kikuchi. 2005. Protein phosphatase type 2A, PP2A, is involved in degradation of gp130. *Mol Cell Biochem* 269: 183-7.
- Moberg, K. H., S. Schelble, S. K. Burdick, and I. K. Hariharan. 2005. Mutations in erupted, the *Drosophila* ortholog of mammalian tumor susceptibility gene 101, elicit non-cell-autonomous overgrowth. *Dev Cell* 9: 699-710.
- Motokura, T., and A. Arnold. 1993. PRAD1/cyclin D1 proto-oncogene: genomic organization, 5'DNA sequence, and sequence of a tumor-specific rearrangement breakpoint. *Genes Chromosomes Cancer* 7: 89-95.
- Mowen, K. A., J. Tang, W. Zhu, B. T. Schurter, K. Shuai, H. R. Herschman, and M. David. 2001. Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription. *Cell* 104: 731-41.
- Mukherjee, T., J. C. Hombria, and M. P. Zeidler. 2005. Opposing roles for *Drosophila* JAK/STAT signalling during cellular proliferation. *Oncogene* 24: 2503-11.
- Mukherjee, T., U. Schäfer, and M. P. Zeidler. 2006. Identification of *Drosophila* genes modulating janus kinase/signal transducer and activator of transcription signal transduction. *Genetics* 172: 1683-97.
- Mulder, N. J., R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bradley, P. Bork, P. Bucher, L. Cerutti, R. Copley, E. Courcelle, U. Das, R. Durbin, W. Fleischmann, J. Gough, D. Haft, N. Harte, N. Hulo, D. Kahn, A. Kanapin, M. Krestyaninova, D. Lonsdale, R. Lopez, I. Letunic, M. Madera, J. Maslen, J. McDowall, A. Mitchell, A. N. Nikolskaya, S. Orchard, M. Pagni, C. P. Ponting, E. Quevillon, J. Selengut, C. J. Sigrist, V. Silventoinen, D. J. Studholme, R. Vaughan, and C. H. Wu. 2005. InterPro, progress and status in 2005. *Nucleic Acids Res* 33 Database Issue: D201-5.
- Müller, P., D. Kutenkeuler, V. Gesellchen, M. P. Zeidler, and M. Boutros. 2005. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature* 436: 871-5.
- Murphy, T. L., E. D. Geissal, J. D. Farrar, and K. M. Murphy. 2000. Role of the Stat4 N domain in receptor proximal tyrosine phosphorylation. *Mol Cell Biol* 20: 7121-31.
- Napoli, C., C. Lemieux, and R. Jorgensen. 1990. Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. *Plant Cell* 2: 279-89.
- Naylor, L. H. 1999. Reporter gene technology: the future looks bright. *Biochem Pharmacol* 58: 749-57.
- Ng, K. C., A. M. Tan, Y. Y. Chong, L. C. Lau, and J. Lou. 1999. Congenital acute megakaryoblastic leukemia (M7) with chromosomal t(1;22)(p13;q13) translocation in a set of identical twins. *J Pediatr Hematol Oncol* 21: 428-30.
- Ngo, H., C. Tschudi, K. Gull, and E. Ullu. 1998. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 95: 14687-92.

- Nichols, A. F., T. Itoh, J. A. Graham, W. Liu, M. Yamaizumi, and S. Linn. 2000. Human damage-specific DNA-binding protein p48. Characterization of XPE mutations and regulation following UV irradiation. *J Biol Chem* 275: 21422-28.
- Nilson, L. A., and T. Schupbach. 1999. EGF receptor signaling in *Drosophila* oogenesis. *Curr Top Dev Biol* 44: 203-43.
- Niwa, H., T. Burdon, I. Chambers, and A. Smith. 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 12: 2048-60.
- Nüsslein-Volhard, C., E. Wieschaus, and K. H. 1984. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* I. Zygotic loci on the second chromosome. *Roux's Arch Dev Biol* 193: 267-82.
- O'Kane, C. J. 2003. Modelling human diseases in *Drosophila* and *Caenorhabditis*. *Semin Cell Dev Biol* 14: 3-10.
- Pawson, T., and T. M. Saxton. 1999. Signaling networks--do all roads lead to the same genes? *Cell* 97: 675-8.
- Peng, H., I. Feldman, and F. J. Rauscher, 3rd. 2002. Hetero-oligomerization among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: a potential mechanism for regulating the switch between coactivation and corepression. *J Mol Biol* 320: 629-44.
- Peschon, J. J., J. L. Slack, P. Reddy, K. L. Stocking, S. W. Sunnarborg, D. C. Lee, W. E. Russell, B. J. Castner, R. S. Johnson, J. N. Fitzner, R. W. Boyce, N. Nelson, C. J. Kozlosky, M. F. Wolfson, C. T. Rauch, D. P. Cerretti, R. J. Paxton, C. J. March, and R. A. Black. 1998. An essential role for ectodomain shedding in mammalian development. *Science* 282: 1281-4.
- Phippen, T. M., A. L. Sweigart, M. Moniwa, A. Krumm, J. R. Davie, and S. M. Parkhurst. 2000. *Drosophila* C-terminal binding protein functions as a context-dependent transcriptional co-factor and interferes with both mad and groucho transcriptional repression. *J Biol Chem* 275: 37628-37.
- Pires-daSilva, A., and R. J. Sommer. 2003. The evolution of signalling pathways in animal development. *Nat Rev Genet* 4: 39-49.
- R-Development-Core-Team. 2004. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.
- Raible, F., K. Tessmar-Raible, K. Osoegawa, P. Wincker, C. Jubin, G. Balavoine, D. Ferrier, V. Benes, P. de Jong, J. Weissenbach, P. Bork, and D. Arendt. 2005. Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* 310: 1325-6.
- Ramos, V. C., J. M. Vidal-Taboada, S. Bergonon, A. Egeo, E. M. C. Fischer, P. Scartezzini, and R. Oliva. 2002. Characterisation and expression analysis of the *WDR9* gene, located in the Down critical region-2 of the human chromosome 21. *Biochim Biophys Acta* 1577: 377-83.
- Ransom, D. G., N. Bahary, K. Niss, D. Traver, C. Burns, N. S. Trede, N. Paffett-Lugassy, W. J. Saganic, C. A. Lim, C. Hersey, Y. Zhou, B. A. Barut, S. Lin, P. D. Kingsley, J. Palis, S. H. Orkin, and L. I. Zon. 2004. The zebrafish *moonshine* gene encodes transcriptional intermediary factor 1gamma, an essential regulator of hematopoiesis. *PLoS Biol* 2: E237.
- Rapic-Otrin, V., M. P. Mclenigan, D. C. Bisi, M. Gonzalez, and A. S. Levine. 2002. Sequential binding of UV DNA damage binding factor and degradation of the p48 subunit as early events after UV irradiation. *Nucleic Acids Res* 30: 2588-98.
- Rawlings, J. S., K. M. Rosler, and D. A. Harrison. 2004. The JAK/STAT signaling pathway. *J Cell Sci* 117: 1281-3.
- Reynolds, A., E. M. Anderson, A. Vermeulen, Y. Fedorov, K. Robinson, D. Leake, J. Karpilow, W. S. Marshall, and A. Khorova. 2006. Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA* 12: 988-93.
- Rosenfeld, R. G., E. Kofoed, C. Buckway, B. Little, K. A. Woods, J. Tsubaki, K. A. Pratt, L. Bezrodnik, H. Jasper, A. Tepper, J. J. Heinrich, and V. Hwa. 2005. Identification of the first patient with a confirmed mutation of the JAK-STAT system. *Pediatr Nephrol* 20: 303-5.
- Rozen, S., and H. Skaletsky. 2000. Primer 3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132: 365-86.
- Rubin, G. M., and A. C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218: 348-53.
- Rubin, G. M., M. D. Yandell, J. R. Wortman, G. L. Gabor Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, P. W. Li, R. Apweiler, W. Fleischmann, J. M. Cherry, S. Henikoff, M. P. Skupski, S.

- Misra, M. Ashburner, E. Birney, M. S. Boguski, T. Brody, P. Brokstein, S. E. Celniker, S. A. Chervitz, D. Coates, A. Cravchik, A. Gabrielian, R. F. Galle, W. M. Gelbart, R. A. George, L. S. Goldstein, F. Gong, P. Guan, N. L. Harris, B. A. Hay, R. A. Hoskins, J. Li, Z. Li, R. O. Hynes, S. J. Jones, P. M. Kuehl, B. Lemaitre, J. T. Littleton, D. K. Morrison, C. Mungall, P. H. O'Farrell, O. K. Pickeral, C. Shue, L. B. Vosshall, J. Zhang, Q. Zhao, X. H. Zheng, and S. Lewis. 2000. Comparative genomics of the eukaryotes. *Science* 287: 2204-15.
- Rusterholz, C., P. C. Henrioud, and M. Nabholz. 1999. Interleukin-2 (IL-2) regulates the accessibility of the IL-2-responsive enhancer in the IL-2 receptor alpha gene to transcription factors. *Mol Cell Biol* 19: 2681-9.
- Saleh, M. C., R. P. van Rij, A. Hekele, A. Gillis, E. Foley, P. H. O'Farrell, and R. Andino. 2006. The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. *Nat Cell Biol* 8: 793-802.
- Sanchez Alvarado, A., and P. A. Newmark. 1999. Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc Natl Acad Sci U S A* 96: 5049-54.
- Sawyer, J. R., E. L. Thomas, J. L. Lukacs, C. M. Swanson, Y. Ding, D. M. Parham, J. R. Thomas, and R. W. Nicholas. 2002. Recurring breakpoints of 1p13 approximately p22 in osteochondroma. *Cancer Genet Cytogenet* 138: 102-6.
- Shankaran, V., H. Ikeda, A. T. Bruce, J. M. White, P. E. Swanson, L. J. Old, and R. D. Schreiber. 2001. IFN γ and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* 410: 1107-11.
- Schindler, C., K. Shuai, V. R. Prezioso, and J. E. Darnell, Jr. 1992. Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor. *Science* 257: 809-13.
- Schneider, I. 1972. Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *J Embryol Exp Morphol* 27: 353-65.
- Sefton, L., J. R. Timmer, Y. Zhang, F. Beranger, and T. W. Cline. 2000. An extracellular activator of the *Drosophila* JAK/STAT pathway is a sex-determination signal element. *Nature* 405: 970-3.
- Seto, E. S., and H. J. Bellen. 2006. Internalization is required for proper Wingless signaling in *Drosophila melanogaster*. *J Cell Biol* 173: 95-106.
- Sherf, B. A., S. L. Navarro, R. R. Hannah, and K. V. Wood. 1996. Dual-Luciferase Reporter Assay: An Advanced Co-Reporter Technology Integrating Firefly and *Renilla* Luciferase Assays. *Promega Notes*: 2-9.
- Shevchenko, A., M. Wilm, O. Vorm, and M. Mann. 1996. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68: 850-8.
- Shi, S., H. C. Calhoun, F. Xia, J. Li, L. Le, and W. X. Li. 2006. JAK signaling globally counteracts heterochromatic gene silencing. *Nat Genet* 38: 1071-6.
- Shuai, K. 2000. Modulation of STAT signaling by STAT-interacting proteins. *Oncogene* 19: 2638-44.
- Shuai, K., G. R. Stark, I. M. Kerr, and J. E. Darnell, Jr. 1993. A single phosphotyrosine residue of Stat91 required for gene activation by interferon-gamma. *Science* 261: 1744-6.
- Silver, D. L., E. R. Geisbrecht, and D. J. Montell. 2005. Requirement for JAK/STAT signaling throughout border cell migration in *Drosophila*. *Development* 132: 3483-92.
- Silver, D. L., and D. J. Montell. 2001. Paracrine signaling through the JAK/STAT pathway activates invasive behavior of ovarian epithelial cells in *Drosophila*. *Cell* 107: 831-41.
- Sims, D., B. Bursteinas, Q. Gao, M. Zvelebil, and B. Baum. 2006. FLIGHT: database and tools for the integration and cross-correlation of large-scale RNAi phenotypic datasets. *Nucleic Acids Res* 34: D479-83.
- Smith, R. K., P. M. Carroll, J. D. Allard, and M. A. Simon. 2002. MASK, a large ankyrin repeat and KH domain-containing protein involved in *Drosophila* receptor tyrosine kinase signaling. *Development* 129: 71-82.
- Song, J. I., and J. R. Grandis. 2000. STAT signaling in head and neck cancer. *Oncogene* 19: 2489-95.
- Sorkin, A., and C. M. Waters. 1993. Endocytosis of growth factor receptors. *Bioessays* 15: 375-82.
- Spemann, H. 1901. Über Correlationen in der Entwicklung des Auges. *Verh anat Ges* 15: 61-79.
- Spemann, H., and H. Mangold. 1924. Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Arch mikrok Anat u Entw mechan* 100: 599-638.
- Spradling, A. C., and G. M. Rubin. 1982. Transposition of cloned P-elements into *Drosophila* germ line chromosomes. *Science* 218: 341-7.

- Spradling, A. C., D. Stern, A. Beaton, E. J. Rhem, T. Laverty, N. Mozden, S. Misra, and G. M. Rubin. 1999. The Berkeley *Drosophila* genome project gene disruption project. Single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* 153: 135-77.
- Strehlow, I., and C. Schindler. 1998. Amino-terminal signal transducer and activator of transcription (STAT) domains regulate nuclear translocation and STAT deactivation. *J Biol Chem* 273: 28049-56.
- Sturtevant, M. A., J. W. O'Neill, and E. Bier. 1994. Down-regulation of *Drosophila Egf-r* mRNA levels following hyperactivated receptor signaling. *Development* 120: 2593-600.
- Takata, K., G. Ishikawa, F. Hirose, and K. Sakaguchi. 2002. *Drosophila* damage-specific DNA-binding protein 1 (D-DDB1) is controlled by the DRE/DREF system. *Nucleic Acids Res* 30:3795-808.
- Takata, K., H. Yoshida, M. Yamaguchi, and K. Sakaguchi. 2004. *Drosophila* damaged DNA-binding protein 1 is an essential factor for development. *Genetics* 168: 855-65.
- Takeshita, T., T. Arita, M. Higuchi, H. Asao, K. Endo, H. Kuroda, N. Tanaka, K. Murata, N. Ishii, and K. Sugamura. 1997. STAM, signal transducing adaptor molecule, is associated with Janus kinases and involved in signaling for cell growth and c-myc induction. *Immunity* 6: 449-57.
- Tautz, D., and C. Pfeifle. 1989. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* 98: 81-5.
- Teleman, A. A., M. Strigini, and S. M. Cohen. 2001. Shaping morphogen gradients. *Cell* 105: 559-62.
- ten Hoeve, J., M. de Jesus Ibarra-Sanchez, Y. Fu, W. Zhu, M. Tremblay, M. David, and K. Shuai. 2002. Identification of a nuclear Stat1 protein tyrosine phosphatase. *Mol Cell Biol* 22: 5662-8.
- Thiel, S., I. Behrmann, E. Dittrich, L. Muys, J. Tavernier, J. Wijdenes, P. C. Heinrich, and L. Graeve. 1998. Internalization of the interleukin 6 signal transducer gp130 does not require activation of the Jak/STAT pathway. *Biochem J* 330: 47-54.
- Thompson, B. J., J. Mathieu, H. H. Sung, E. Loeser, P. Rorth, and S. M. Cohen. 2005. Tumor suppressor properties of the ESCRT-II complex component Vps25 in *Drosophila*. *Dev Cell* 9: 711-20.
- Tolia, N. H., and L. Joshua-Tor. 2007. Slicer and the Argonautes. *Nat Chem Biol* 3: 36-43.
- Townsend, P. A., T. M. Scarabelli, S. M. Davidson, R. A. Knight, D. S. Latchman, and A. Stephanou. 2004. STAT-1 interacts with p53 to enhance DNA damage-induced apoptosis. *J Biol Chem* 279: 5811-20.
- Troke, P. J., K. B. Kindle, H. M. Collins, and D. M. Heery. 2006. MOZ fusion proteins in acute myeloid leukaemia. *Biochem Soc Symp*: 23-39.
- Tseng, A. S., and I. K. Hariharan. 2002. An overexpression screen in *Drosophila* for genes that restrict growth or cell-cycle progression in the developing eye. *Genetics* 162: 229-43.
- Tsruya, R., A. Schlesinger, A. Reich, L. Gabay, A. Sapir, and B. Z. Shilo. 2002. Intracellular trafficking by Star regulates cleavage of the *Drosophila* EGF receptor ligand Spitz. *Genes Dev* 16: 222-34.
- Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading Massachusetts: Addison-Wesley.
- Tulina, N., and E. Matunis. 2001. Control of stem cell self-renewal in *Drosophila* spermatogenesis by JAK-STAT signaling. *Science* 294: 2546-9.
- Ulvila, J., M. Parikka, A. Kleino, R. Sormunen, R. A. Ezekowitz, C. Kocks, and M. Ramet. 2006. Double-stranded RNA is internalized by scavenger receptor-mediated endocytosis in *Drosophila* S2 cells. *J Biol Chem* 281: 14370-5.
- Ungureanu, D., S. Vanhatupa, N. Kotaja, J. Yang, S. Aittomaki, O. A. Janne, J. J. Palvimo, and O. Silvennoinen. 2003. PIAS proteins promote SUMO-1 conjugation to STAT1. *Blood* 102: 3311-3.
- Ursuliak, Z., J. C. Clemens, J. E. Dixon, and J. V. Price. 1997. Differential accumulation of DPTP61F alternative transcripts: regulation of a protein tyrosine phosphatase by segmentation genes. *Mech Dev* 65: 19-30.
- Vaccari, T., and D. Bilder. 2005. The *Drosophila* tumor suppressor vps25 prevents nonautonomous overproliferation by regulating notch trafficking. *Dev Cell* 9: 687-98.
- van der Krol, A. R., L. A. Mur, M. Beld, J. N. Mol, and A. R. Stuitje. 1990. Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2: 291-9.
- Vinkemeier, U., S. L. Cohen, I. Moarefi, B. T. Chait, J. Kuriyan, and J. E. Darnell, Jr. 1996. DNA binding of *in vitro* activated Stat1 alpha, Stat1 beta and truncated Stat1: interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites. *Embo J* 15: 5616-26.

- Wakasugi, M., M. Shimizu, H. Morioka, S. Linn, O. Nikaido, and T. Matsunaga. 2001. Damaged DNA-binding Protein DDB stimulates the excision of cyclobutane pyrimidine dimers *in vitro* in concert with XPA and replication protein A. *J Biol Chem* 18: 15434-40.
- Wang, H., L. Zhai, J. Xu, H. Y. Joo, S. Jackson, H. Erdjument-Bromage, P. Tempst, Y. Xiong, and Y. Zhang. 2006. Histone H3 and H4 ubiquitylation by the CUL4–DDB–ROC1 ubiquitin ligase facilitates cellular response to DNA damage. *Mol Cell* 22: 383–94.
- Wang, J., W. Lin, B. Popko, and I. L. Campbell. 2004. Inducible production of interferon-gamma in the developing brain causes cerebellar dysplasia with activation of the Sonic hedgehog pathway. *Mol Cell Neurosci* 27: 489-96.
- Wang, R., P. Cherukuri, and J. Luo. 2005. Activation of Stat3 sequence-specific DNA binding and transcription by p300/CREB-binding protein-mediated acetylation. *J Biol Chem* 280: 11528-34.
- Wawersik, M., A. Milutinovich, A. L. Casper, E. Matunis, B. Williams, and M. Van Doren. 2005. Somatic control of germline sexual development is mediated by the JAK/STAT pathway. *Nature* 436: 563-7.
- Weigmann, K., R. Klapper, T. Strasser, C. Rickert, G. Technau, H. Jäckle, W. Janning, and C. Klämbt. 2003. FlyMove—a new way to look at development of *Drosophila*. *Trends Genet* 19: 310-1.
- Wellbrock, C., C. Weisser, J. C. Hassel, P. Fischer, J. Becker, C. S. Vetter, I. Behrmann, M. Kortylewski, P. C. Heinrich, and M. Schartl. 2005. STAT5 contributes to interferon resistance of melanoma cells. *Curr Biol* 15: 1629-39.
- Wertz, I. E., K. M. O'Rourke, Z. Zhang, D. Dornan, D. Arnott, R. J. Deshaies, and V. M. Dixit. 2004. Human-De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase. *Science* 303: 1371-4.
- Wormald, S., and D. J. Hilton. 2004. Inhibitors of cytokine signal transduction. *J Biol Chem* 279: 821-4.
- Xi, R., J. R. McGregor, and D. A. Harrison. 2003. A gradient of JAK pathway activity patterns the anterior-posterior axis of the follicular epithelium. *Dev Cell* 4: 167-177.
- Xu, C., R. C. Kauffmann, J. Zhang, S. Kladny, and R. W. Carthew. 2000. Overlapping activators and repressors delimit transcriptional response to receptor tyrosine kinase signals in the *Drosophila* eye. *Cell* 103: 87-97.
- Xu, T., and G. M. Rubin. 1993. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* 117: 1223-37.
- Xu, X., Y. L. Sun, and T. Hoey. 1996. Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain. *Science* 273: 794-7.
- Yamamoto, K., M. Yamaguchi, N. Miyasaka, and O. Miura. 2003. SOCS-3 inhibits IL-12-induced STAT4 activation by binding through its SH2 domain to the STAT4 docking site in the IL-12 receptor beta2 subunit. *Biochem Biophys Res Commun* 310: 1188-93.
- Yan, R., H. Luo, J. E. Darnell, Jr., and C. R. Dearolf. 1996a. A JAK-STAT pathway regulates wing vein formation in *Drosophila*. *Proc Natl Acad Sci U S A* 93: 5842-7.
- Yan, R., S. Small, C. Desplan, C. R. Dearolf, and J. E. Darnell, Jr. 1996b. Identification of a Stat gene that functions in *Drosophila* development. *Cell* 84: 421-30.
- Yanagawa, S., J. S. Lee, and A. Ishimoto. 1998. Identification and characterization of a novel line of *Drosophila* Schneider S2 cells that respond to *wingless* signaling. *J Biol Chem* 273: 32353-9.
- Yang, E., M. A. Henriksen, O. Schaefer, N. Zakharova, and J. E. Darnell, Jr. 2002. Dissociation time from DNA determines transcriptional function in a STAT1 linker mutant. *J Biol Chem* 277: 13455-62.
- Yasukawa, H., H. Misawa, H. Sakamoto, M. Masuhara, A. Sasaki, T. Wakioka, S. Ohtsuka, T. Imaizumi, T. Matsuda, J. N. Ihle, and A. Yoshimura. 1999. The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop. *Embo J* 18: 1309-20.
- Yuan, Z. L., Y. J. Guan, D. Chatterjee, and Y. E. Chin. 2005. Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science* 307: 269-73.
- Zamore, P. D., and B. Haley. 2005. Ribo-gnome: the big world of small RNAs. *Science* 309: 1519-24.
- Zeidler, M. P., N. Perrimon, and D. I. Strutt. 1999. Polarity determination in the *Drosophila* eye: a novel role for *unpaired* and JAK/STAT signaling. *Genes Dev* 13: 1342-53.
- Zhang, X., J. Blenis, H. C. Li, C. Schindler, and S. Chen-Kiang. 1995. Requirement of serine phosphorylation for formation of STAT-promoter complexes. *Science* 267: 1990-4.

- Zhang, J. J., U. Vinkemeier, W. Gu, D. Chakravarti, C. M. Horvath, and J. E. Darnell, Jr. 1996. Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling. *Proc Natl Acad Sci* 93: 15092-6.
- Zhang, X. D., X. C. Yang, N. Chung, A. Gates, E. Stec, P. Kunapuli, D. J. Holder, M. Ferrer, and A. S. Espeseth. 2006. Robust statistical methods for hit selection in RNA interference high-throughput screening experiments. *Pharmacogenomics* 7: 299-309.
- Zhong, S., L. Delva, C. Rachez, C. Cenciarelli, D. Gandini, H. Zhang, S. Kalantry, L. P. Freedman, and P. P. Pandolfi. 1999. A RA-dependent, tumour-growth suppressive transcription complex is the target of the PML-RARalpha and T18 oncoproteins. *Nat Genet* 23: 287-95.
- Zi, Z., K. H. Cho, M. H. Sung, X. Xia, J. Zheng, and Z. Sun. 2005. *In silico* identification of the key components and steps in IFN-gamma induced JAK-STAT signaling pathway. *FEBS Lett* 579: 1101-8.

APPENDIX

Supplementary Script 1. Source code for the analysis of the genome-wide RNAi screen including help pages (from the *CellScreen* package).

Note: analysis functions start with the pattern “**xyz** <- function(...) {...}”, whereas corresponding help pages start with the pattern “% --- Source file: man/xyz.Rd ---”.

```
center.byplate <- function(x)
{
  screenData <- x

  ## split by factor (Plate)
  s <- split(screenData, screenData$Plate384)

  ## apply median and mad to all four channels
  for(i in 1:length(s))
  {
    spos <- (i-1)*length(s[[i]]$Element384)+s[[i]]$Element384
    screenData$f11med[spos] <- median(s[[i]]$f11, na.rm=TRUE)
    screenData$f12med[spos] <- median(s[[i]]$f12, na.rm=TRUE)
    screenData$f11mad[spos] <- mad(s[[i]]$f11, na.rm=TRUE)
    screenData$f12mad[spos] <- mad(s[[i]]$f12, na.rm=TRUE)
    screenData$r11med[spos] <- median(s[[i]]$r11, na.rm=TRUE)
    screenData$r12med[spos] <- median(s[[i]]$r12, na.rm=TRUE)
    screenData$r11mad[spos] <- mad(s[[i]]$r11, na.rm=TRUE)
    screenData$r12mad[spos] <- mad(s[[i]]$r12, na.rm=TRUE)
    screenData$FLRL1med[spos] <- median(s[[i]]$FLRL1, na.rm=TRUE)
    screenData$FLRL2med[spos] <- median(s[[i]]$FLRL2, na.rm=TRUE)
    screenData$FLRL1mad[spos] <- mad(s[[i]]$FLRL1, na.rm=TRUE)
    screenData$FLRL2mad[spos] <- mad(s[[i]]$FLRL2, na.rm=TRUE)
  }
  return(screenData)
}
```

```
center.bywell <- function(x)
{
  screenData <- x

  ## split by factor (well)
  t <- split(screenData, screenData$Element384)

  ##create slots for processed data
  screenData$f11med3D <- numeric(nrow(screenData))
  screenData$f12med3D <- numeric(nrow(screenData))
  screenData$f11mad3D <- numeric(nrow(screenData))
  screenData$f12mad3D <- numeric(nrow(screenData))
  screenData$r11med3D <- numeric(nrow(screenData))
  screenData$r12med3D <- numeric(nrow(screenData))
  screenData$r11mad3D <- numeric(nrow(screenData))
  screenData$r12mad3D <- numeric(nrow(screenData))
  screenData$FLRL1med3D <- numeric(nrow(screenData))
  screenData$FLRL2med3D <- numeric(nrow(screenData))
}
```

```

screenData$FLRL1mad3D <- numeric(nrow(screenData))
screenData$FLRL2mad3D <- numeric(nrow(screenData))
screenData$fl1Bmed3D <- numeric(nrow(screenData))
screenData$fl2Bmed3D <- numeric(nrow(screenData))
screenData$fl1Bmad3D <- numeric(nrow(screenData))
screenData$fl2Bmad3D <- numeric(nrow(screenData))
screenData$rl1Bmed3D <- numeric(nrow(screenData))
screenData$rl2Bmed3D <- numeric(nrow(screenData))
screenData$rl1Bmad3D <- numeric(nrow(screenData))
screenData$rl2Bmad3D <- numeric(nrow(screenData))
screenData$FLRL1Bmed3D <- numeric(nrow(screenData))
screenData$FLRL2Bmed3D <- numeric(nrow(screenData))
screenData$FLRL1Bmad3D <- numeric(nrow(screenData))
screenData$FLRL2Bmad3D <- numeric(nrow(screenData))

## apply median and mad to z-scores
for(i in 1:length(t))
  {
    for (i in 1:length(t)) {
      spos <- seq(i, length(screenData$Plate384), length(t))
      screenData$fl1med3D[spos] <- median(t[[i]]$fl1score, na.rm = TRUE)
      screenData$fl2med3D[spos] <- median(t[[i]]$fl2score, na.rm = TRUE)
      screenData$fl1mad3D[spos] <- mad(t[[i]]$fl1score, na.rm = TRUE)
      screenData$fl2mad3D[spos] <- mad(t[[i]]$fl2score, na.rm = TRUE)
      screenData$rl1med3D[spos] <- median(t[[i]]$rl1score, na.rm = TRUE)
      screenData$rl2med3D[spos] <- median(t[[i]]$rl2score, na.rm = TRUE)
      screenData$rl1mad3D[spos] <- mad(t[[i]]$rl1score, na.rm = TRUE)
      screenData$rl2mad3D[spos] <- mad(t[[i]]$rl2score, na.rm = TRUE)
      screenData$FLRL1med3D[spos] <- median(t[[i]]$FLRL1score, na.rm = TRUE)
      screenData$FLRL2med3D[spos] <- median(t[[i]]$FLRL2score, na.rm = TRUE)
      screenData$FLRL1mad3D[spos] <- mad(t[[i]]$FLRL1score, na.rm = TRUE)
      screenData$FLRL2mad3D[spos] <- mad(t[[i]]$FLRL2score, na.rm = TRUE)
      screenData$fl1Bmed3D[spos] <- median(t[[i]]$fl1Bscore, na.rm = TRUE)
      screenData$fl2Bmed3D[spos] <- median(t[[i]]$fl2Bscore, na.rm = TRUE)
      screenData$fl1Bmad3D[spos] <- mad(t[[i]]$fl1Bscore, na.rm = TRUE)
      screenData$fl2Bmad3D[spos] <- mad(t[[i]]$fl2Bscore, na.rm = TRUE)
      screenData$rl1Bmed3D[spos] <- median(t[[i]]$rl1Bscore, na.rm = TRUE)
      screenData$rl2Bmed3D[spos] <- median(t[[i]]$rl2Bscore, na.rm = TRUE)
      screenData$rl1Bmad3D[spos] <- mad(t[[i]]$rl1Bscore, na.rm = TRUE)
      screenData$rl2Bmad3D[spos] <- mad(t[[i]]$rl2Bscore, na.rm = TRUE)
      screenData$FLRL1Bmed3D[spos] <- median(t[[i]]$FLRL1Bscore, na.rm = TRUE)
      screenData$FLRL2Bmed3D[spos] <- median(t[[i]]$FLRL2Bscore, na.rm = TRUE)
      screenData$FLRL1Bmad3D[spos] <- mad(t[[i]]$FLRL1Bscore, na.rm = TRUE)
      screenData$FLRL2Bmad3D[spos] <- mad(t[[i]]$FLRL2Bscore, na.rm = TRUE)
    }
  }
return(screenData)
}

channel.norma <- function(x)
{
  x$FLRL1 <- x$fl1/x$rl1
  x$FLRL2 <- x$fl2/x$rl2
  return(x)
}

```



```

histo.byplate <- function(x, plate = 34, data.col = x$fl1score)
{
  screenData <- x
  library(Simple)
  simple.hist.and.boxplot(data.col[screenData$Plate384==plate], main="Plate histogram")
}

image.byplate <- function(x, plate = 34, data.col = x$fl1score)
{
  screenData <- x
  library(prada)
  library(RColorBrewer)
  library(geneplotter)

  f <- colorRampPalette(brewer.pal(9,"RdYlBu"))
  ## replace Na with 0
  x <- is.na(data.col)
  data.col[x] <- 0

  plotPlate(data.col[screenData$Plate384==plate], nrow=16, ncol=24, col=f(20), xrange=c(-5,5),
width=7, desc=c("act", "inh"))
}

medmad.byplate <- function(x)
{
  screenData <- x
  quartz()
  plot(screenData$Plate384, screenData$fl1med/screenData$fl1mad,xlab= "plate",
ylab="FL(median/mad)", main="Screen med/mad by plate, FL1 channel")
  quartz()
  plot(screenData$Plate384, screenData$rl1med/screenData$rl1mad,xlab= "plate",
ylab="RL(median/mad)", main="Screen med/mad by plate, RL1 channel")
  quartz()
  plot(screenData$Plate384, screenData$fl2med/screenData$fl2mad,xlab= "plate",
ylab="FL(median/mad)", main="Screen med/mad by plate, FL2 channel")
  quartz()
  plot(screenData$Plate384, screenData$rl2med/screenData$rl2mad,xlab= "plate",
ylab="RL(median/mad)", main="Screen med/mad by plate, RL2 channel")
}

plate.merge <- function()
{
  d <- list.files("./FL")
  dataFL= NULL
  for(i in d) dataFL <- rbind(dataFL, cbind(file=i,read.table(paste("./FL/",i,sep=""))))
  e <- list.files("./RL")
  dataRL= NULL
  for(i in e) dataRL <- rbind(dataRL, cbind(file=i,read.table(paste("./RL/",i,sep=""))))
  mergedRawData <- data.frame(dataFL$V3, dataRL$V3)
  return(mergedRawData)
}

screen.histo <- function(x, data.col = x$flavg)
{
  library(Simple)

```

```

    simple.hist.and.boxplot(data.col, main="Screen histogram")
  }

screen.image <- function(x, data.col = x$flavg, threshold1 = -7, threshold2 = 7)
  {
    screenData <- x
    library(RColorBrewer)
    library(geneplotter)
    f <- colorRampPalette(brewer.pal(9,"RdYlBu"))
    data_col <- data.col

    ## reformat data
    im_data <- rep(as.numeric(0), 23040)
    dim(im_data) <- c(240, 96)
    for(x in 0:9)
      {
        for(y in 0:5)
          {
            for(n in 1:24)
              {
                for(m in 1:16)
                  {
                    im_data[x*24+n,96-m-y*16+1] <- data_col[y*3840+x*384+(m-
1)*24+n]
                  }
                }
            }
          }
        }

    ## cap extreme values for image
    x <- im_data < threshold1
    im_data[x] <- threshold1
    x <- im_data > threshold2
    im_data[x] <- threshold2

    ## replace NA with 0
    x <- is.na(im_data)
    im_data[x] <- 0

    image(matrix(im_data,ncol=96, nrow=240), col=f(20), main="Screen overview")
  }

screen.qqPlot <- function(x, data.col = x$flavg)
  {
    qqnorm(data.col, main="Screen normal Q-Q Plot")
    qqline(data.col, col="red")
  }

screen.reader <- function(w, x, y, z)
  {
    ## read raw luminescence values
    screenData <- w

    ## read gene annotation file
    geneList <- x
  }

```

```
## read flag file
flagList <- y

## read phenotype file
phenoData <- z

## create slots for analysis
screenData$fl1med <- numeric(nrow(screenData))
screenData$fl2med <- numeric(nrow(screenData))
screenData$fl1mad <- numeric(nrow(screenData))
screenData$fl2mad <- numeric(nrow(screenData))
screenData$fl1score <- numeric(nrow(screenData))
screenData$fl2score <- numeric(nrow(screenData))
screenData$flavg <- numeric(nrow(screenData))
screenData$fldiff <- numeric(nrow(screenData))
screenData$rl1med <- numeric(nrow(screenData))
screenData$rl2med <- numeric(nrow(screenData))
screenData$rl1mad <- numeric(nrow(screenData))
screenData$rl2mad <- numeric(nrow(screenData))
screenData$rl1score <- numeric(nrow(screenData))
screenData$rl2score <- numeric(nrow(screenData))
screenData$rlavg <- numeric(nrow(screenData))
screenData$rldiff <- numeric(nrow(screenData))
screenData$FLRL1 <- numeric(nrow(screenData))
screenData$FLRL2 <- numeric(nrow(screenData))
screenData$FLRL1med <- numeric(nrow(screenData))
screenData$FLRL1mad <- numeric(nrow(screenData))
screenData$FLRL2med <- numeric(nrow(screenData))
screenData$FLRL2mad <- numeric(nrow(screenData))
screenData$FLRL1score <- numeric(nrow(screenData))
screenData$FLRL2score <- numeric(nrow(screenData))
screenData$FLRLavg <- numeric(nrow(screenData))
screenData$FLRLdiff <- numeric(nrow(screenData))

## create slots for annotation
screenData$locid <- character(nrow(screenData))
screenData$hfa <- character(nrow(screenData))
screenData$scg <- character(nrow(screenData))
screenData$genes <- character(nrow(screenData))
screenData$affx <- character(nrow(screenData))

## create slots for flags
screenData$flag <- numeric(nrow(screenData))

## create slots for phenotypes from other screens
screenData$viakc <- numeric(nrow(screenData))
screenData$vias2r <- numeric(nrow(screenData))
screenData$sother <- character(nrow(screenData))
screenData$comment <- numeric(nrow(screenData))

## add annotation information
screenData$locid <- geneList$LocationID
screenData$hfa <- geneList$HFA
screenData$scg <- geneList$CG
```

```

screenData$genes <- geneList$Gene
screenData$affx <- geneList$Affx
screenData$flag <- flagList$Flag

## add phenotype data
screenData$viakc <- phenoData$Via_Kc
screenData$vias2r <- phenoData$Via_S2R
screenData$other <- phenoData$Other
return(screenData)
}

write.out <- function(x)
{
  screenData <- x
  write.table(screenData, file="ScreenDataOut.txt", sep="\t", row.names=FALSE,
col.names=TRUE, quote=FALSE)
}

z.score.2D <- function(x)
{
  ## calculate z-score and difference for each well
  screenData <- x
  screenData$fl1score <- (screenData$fl1-screenData$fl1med)/screenData$fl1mad
  screenData$fl2score <- (screenData$fl2-screenData$fl2med)/screenData$fl2mad
  screenData$rl1score <- (screenData$rl1-screenData$rl1med)/screenData$rl1mad
  screenData$rl2score <- (screenData$rl2-screenData$rl2med)/screenData$rl2mad
  screenData$flavg <- (screenData$fl1score+screenData$fl2score)/2
  screenData$rlavg <- (screenData$rl1score+screenData$rl2score)/2
  screenData$flldiff <- abs(screenData$fl1score-screenData$fl2score)
  screenData$rlldiff <- abs(screenData$rl1score-screenData$rl2score)
  x <- screenData$flag == "C" | screenData$flag == "A"
  x <- screenData$flag == "D" | screenData$flag == "B"
  screenData$flavg[x] <- screenData$fl1score[x]
  screenData$rlavg[x] <- screenData$rl1score[x]
  screenData$flavg[x] <- screenData$fl2score[x]
  screenData$rlavg[x] <- screenData$rl2score[x]
  x <- is.na(screenData$flavg)
  screenData$flavg[x] <- screenData$fl1score[x]
  x <- is.na(screenData$flavg)
  screenData$flavg[x] <- screenData$fl2score[x]
  x <- is.na(screenData$rlavg)
  screenData$rlavg[x] <- screenData$rl1score[x]
  x <- is.na(screenData$rlavg)
  screenData$rlavg[x] <- screenData$rl2score[x]
  screenData$FLRL1score <- (screenData$FLRL1-
screenData$FLRL1med)/screenData$FLRL1mad
  screenData$FLRL2score<- (screenData$FLRL2-screenData$FLRL2med)/screenData$FLRL2mad
  screenData$FLRLavg <- (screenData$FLRL1score+screenData$FLRL2score)/2
  screenData$FLRLdiff <- abs(screenData$FLRL1score-screenData$FLRL2score)
  x <- screenData$flag == "C" | screenData$flag == "A"
  screenData$FLRLavg[x] <- screenData$FLRL2score[x]
  x <- screenData$flag == "D" | screenData$flag == "B"
  screenData$FLRLavg[x] <- screenData$FLRL1score[x]
  return(screenData)
}

```

```

z.score.3D <- function(x)
{
  screenData <- x

  ##create slots for processed data
  screenData$fl1score3D <- numeric(nrow(screenData))
  screenData$fl2score3D <- numeric(nrow(screenData))
  screenData$flavg3D <- numeric(nrow(screenData))
  screenData$fl1Bscore3D <- numeric(nrow(screenData))
  screenData$fl2Bscore3D <- numeric(nrow(screenData))
  screenData$flBavg3D <- numeric(nrow(screenData))
  screenData$flBdiff3D <- numeric(nrow(screenData))
  screenData$rl1score3D <- numeric(nrow(screenData))
  screenData$rl2score3D <- numeric(nrow(screenData))
  screenData$rlavg3D <- numeric(nrow(screenData))
  screenData$rl1Bscore3D <- numeric(nrow(screenData))
  screenData$rl2Bscore3D <- numeric(nrow(screenData))
  screenData$rlBavg3D <- numeric(nrow(screenData))
  screenData$rlBdiff3D <- numeric(nrow(screenData))
  screenData$FLRL1score3D <- numeric(nrow(screenData))
  screenData$FLRL2score3D <- numeric(nrow(screenData))
  screenData$FLRLavg3D <- numeric(nrow(screenData))
  screenData$FLRLdiff3D <- numeric(nrow(screenData))
  screenData$FLRL1Bscore3D <- numeric(nrow(screenData))
  screenData$FLRL2Bscore3D <- numeric(nrow(screenData))
  screenData$FLRLBavg3D <- numeric(nrow(screenData))
  screenData$FLRLBdiff3D <- numeric(nrow(screenData))

  ## calculate z-score and difference for each well
  screenData$fl1score3D <- (screenData$fl1score - screenData$fl1med3D)/screenData$fl1mad3D
  screenData$fl2score3D <- (screenData$fl2score - screenData$fl2med3D)/screenData$fl2mad3D
  screenData$rl1score3D <- (screenData$rl1score - screenData$rl1med3D)/screenData$rl1mad3D
  screenData$rl2score3D <- (screenData$rl2score - screenData$rl2med3D)/screenData$rl2mad3D
  screenData$flavg3D <- (screenData$fl1score3D + screenData$fl2score3D)/2
  screenData$rlavg3D <- (screenData$rl1score3D + screenData$rl2score3D)/2
  screenData$fldiff3D <- abs(screenData$fl1score3D - screenData$fl2score3D)
  screenData$rlldiff3D <- abs(screenData$rl1score3D - screenData$rl2score3D)
  screenData$fl1Bscore3D <- (screenData$fl1Bscore -
screenData$fl1Bmed3D)/screenData$fl1Bmad3D
  screenData$fl2Bscore3D <- (screenData$fl2Bscore -
screenData$fl2Bmed3D)/screenData$fl2Bmad3D
  screenData$rl1Bscore3D <- (screenData$rl1Bscore -
screenData$rl1Bmed3D)/screenData$rl1Bmad3D
  screenData$rl2Bscore3D <- (screenData$rl2Bscore -
screenData$rl2Bmed3D)/screenData$rl2Bmad3D
  screenData$flBavg3D <- (screenData$fl1Bscore3D + screenData$fl2Bscore3D)/2
  screenData$rlBavg3D <- (screenData$rl1Bscore3D + screenData$rl2Bscore3D)/2
  screenData$flBdiff3D <- abs(screenData$fl1Bscore3D - screenData$fl2Bscore3D)
  screenData$rlBdiff3D <- abs(screenData$rl1Bscore3D - screenData$rl2Bscore3D)
  x <- screenData$flag == "C" | screenData$flag == "A"
  x <- screenData$flag == "D" | screenData$flag == "B"
  screenData$flavg3D[x] <- screenData$fl1score3D[x]
  screenData$rlavg3D[x] <- screenData$rl1score3D[x]
  screenData$flavg3D[x] <- screenData$fl2score3D[x]
  screenData$rlavg3D[x] <- screenData$rl2score3D[x]
}

```

```

screenData$flBavg3D[x] <- screenData$fl1Bscore3D[x]
screenData$rlBavg3D[x] <- screenData$rl1Bscore3D[x]
screenData$flBavg3D[x] <- screenData$fl2Bscore3D[x]
screenData$rlBavg3D[x] <- screenData$rl2Bscore3D[x]
x <- is.na(screenData$flavg3D)
screenData$flavg3D[x] <- screenData$fl1score3D[x]
x <- is.na(screenData$flavg3D)
screenData$flavg3D[x] <- screenData$fl2score3D[x]
x <- is.na(screenData$rlavg3D)
screenData$rlavg3D[x] <- screenData$rl1score3D[x]
x <- is.na(screenData$rlavg3D)
screenData$rlavg3D[x] <- screenData$rl2score3D[x]
x <- is.na(screenData$flBavg3D)
screenData$flBavg3D[x] <- screenData$fl1Bscore3D[x]
x <- is.na(screenData$flBavg3D)
screenData$flBavg3D[x] <- screenData$fl2Bscore3D[x]
x <- is.na(screenData$rlBavg3D)
screenData$rlavg3D[x] <- screenData$rl1Bscore3D[x]
x <- is.na(screenData$rlBavg3D)
screenData$rlBavg3D[x] <- screenData$rl2Bscore3D[x]
screenData$FLRL1score3D <- (screenData$FLRL1score -
screenData$FLRL1med3D)/screenData$FLRL1mad3D
screenData$FLRL2score3D <- (screenData$FLRL2score -
screenData$FLRL2med3D)/screenData$FLRL2mad3D
screenData$FLRLavg3D <- (screenData$FLRL1score3D + screenData$FLRL2score3D)/2
screenData$FLRLdiff3D <- abs(screenData$FLRL1score3D - screenData$FLRL2score3D)
x <- screenData$flag == "C" | screenData$flag == "A"
screenData$FLRLavg3D[x] <- screenData$FLRL2score3D[x]
x <- screenData$flag == "D" | screenData$flag == "B"
screenData$FLRLavg3D[x] <- screenData$FLRL1score3D[x]
screenData$FLRL1Bscore3D <- (screenData$FLRL1Bscore -
screenData$FLRL1Bmed3D)/screenData$FLRL1Bmad3D
screenData$FLRL2Bscore3D <- (screenData$FLRL2Bscore -
screenData$FLRL2Bmed3D)/screenData$FLRL2Bmad3D
screenData$FLRLBavg3D <- (screenData$FLRL1Bscore3D + screenData$FLRL2Bscore3D)/2
screenData$FLRLBdiff3D <- abs(screenData$FLRL1Bscore3D - screenData$FLRL2Bscore3D)
x <- screenData$flag == "C" | screenData$flag == "A"
screenData$FLRLBavg3D[x] <- screenData$FLRL2Bscore3D[x]
x <- screenData$flag == "D" | screenData$flag == "B"
screenData$FLRLBavg3D[x] <- screenData$FLRL1Bscore3D[x]
return(screenData)
}

```

```

b.score <- function (x)
{
screenData <- x
s <- split(screenData, screenData$Plate384)
for (i in 1:length(s))
{
polish=matrix(c(1:384), ncol=24, nrow=16)
spos <- (i - 1) * length(s[[i]]$Element384) + s[[i]]$Element384
polish<- medpolish(matrix(s[[i]]$fl1, ncol=24, nrow=16, byrow=T), na.rm = TRUE)
screenData$fl1Bscore[spos] <- c(matrix((polish$residuals/mad(polish$residuals,
na.rm=T)), ncol=16, byrow=T))
polish<- medpolish(matrix(s[[i]]$fl2, ncol=24, nrow=16, byrow=T), na.rm = TRUE)

```

```

        screenData$fl2Bscore[spos] <- c(matrix((polish$residuals/mad(polish$residuals,
na.rm=T)), ncol=16, byrow=T))
        polish<- medpolish(matrix(s[[i]]$rl1, ncol=24, nrow=16, byrow=T), na.rm = TRUE)
        screenData$rl1Bscore[spos] <- c(matrix((polish$residuals/mad(polish$residuals,
na.rm=T)), ncol=16, byrow=T))
        polish<- medpolish(matrix(s[[i]]$rl2, ncol=24, nrow=16, byrow=T), na.rm = TRUE)
        screenData$rl2Bscore[spos] <- c(matrix((polish$residuals/mad(polish$residuals,
na.rm=T)), ncol=16, byrow=T))
        polish<- medpolish(matrix(s[[i]]$FLRL1, ncol=24, nrow=16, byrow=T), na.rm = TRUE)
        screenData$FLRL1Bscore[spos] <- c(matrix((polish$residuals/mad(polish$residuals,
na.rm=T)), ncol=16, byrow=T))
        polish<- medpolish(matrix(s[[i]]$FLRL2, ncol=24, nrow=16, byrow=T), na.rm = TRUE)
        screenData$FLRL2Bscore[spos] <- c(matrix((polish$residuals/mad(polish$residuals,
na.rm=T)), ncol=16, byrow=T))
    }
    screenData$flBavg=(screenData$fl1Bscore+ screenData$fl2Bscore)/2
    screenData$rlBavg=(screenData$rl1Bscore+ screenData$rl2Bscore)/2
    screenData$FLRLBavg=(screenData$FLRL1Bscore+ screenData$FLRL2Bscore)/2
    screenData$flBdiff=abs(screenData$fl1Bscore- screenData$fl2Bscore)
    screenData$rlBdiff=abs(screenData$rl1Bscore- screenData$rl2Bscore)
    screenData$FLRLBdiff=abs(screenData$FLRL1Bscore- screenData$FLRL2Bscore)
    return(screenData)
}

```

```
HeatMapRows <- function (x, threshold1 = -7, threshold2 = 7)
```

```

{
s <- split(x, x$Plate384)
store=matrix(ncol=length(x$Plate384)/16, nrow=16)
i=1
j=1
k=24
while(i<=length(s))
{
store[c(1:16),c(j:k)]=matrix(s[[i]]$fl1score, ncol=24, nrow=16, byrow=T)
i=i+1
j=j+24
k=k+24
}

```

```
## NOTE: taken from package image.R (http://www.maths.lth.se/help/R/image/)
```

```
mirror.matrix=function(y)
```

```

{
xx=as.data.frame(y);
xx=rev(xx);
xx=as.matrix(xx);
xx;
}

```

```
## NOTE: taken from package image.R (http://www.maths.lth.se/help/R/image/)
```

```
rotate270.matrix=function(z)
```

```

{
mirror.matrix(t(z))
}

```

```
store2= rotate270.matrix(store)
```

```

library(RColorBrewer)
library(geneplotter)
f <- colorRampPalette(brewer.pal(9, "RdYlBu"))
x <- store2 < threshold1
store2[x] <- threshold1
x <- store2 > threshold2
store2[x] <- threshold2
x <- is.na(store2)
store2[x] <- 0
image(store2, col = f(20))
}

```

```

hits.perplate <- function(x, data.col = x$flavg)
{
s <- split(x, x$Plate384)
i=1
a=c(1:length(s))
while(i<=length(s))
{
a[i]=sum(data.col[x$Plate384==i]< -2 | data.col[x$Plate384==i]>2, na.rm=T)
i=i+1
}
barplot(a, names.arg=c(1:length(s)))
}

```

```

HeatMapCols <- function(x, threshold1 = -7, threshold2 = 7)
{
s <- split(x, x$Plate384)
store=matrix(ncol=length(x$Plate384)/24, nrow=24)
i=1
j=1
k=16
while(i<=length(s))
{
store[c(1:24),c(j:k)]=matrix(s[[i]]$fl1score, ncol=16, nrow=24)
i=i+1
j=j+16
k=k+16
}

## NOTE: taken from package image.R (http://www.maths.lth.se/help/R/image/)
mirror.matrix=function(x)
{
xx=as.data.frame(x);
xx=rev(xx);
xx=as.matrix(xx);
xx;
}

## NOTE: taken from package image.R (http://www.maths.lth.se/help/R/image/)
rotate270.matrix=function(x)
{
mirror.matrix(t(x))
}

```



```

store2= rotate270.matrix(store)
library(RColorBrewer)
library(geneplotter)
f <- colorRampPalette(brewer.pal(9, "RdYlBu"))
x <- store2 < threshold1
store2[x] <- threshold1
x <- store2 > threshold2
store2[x] <- threshold2
x <- is.na(store2)
store2[x] <- 0
image(store2, col = f(20))
}

dynamicRange <- function(x, data.col = x$fl1score, ctrl1 = 2, ctrl2 = 26)
{
a=data.col[x$Element384==ctrl1]
b=data.col[x$Element384==ctrl2]
z=c(1:length(a))
plot(z,a, ylim=c(min(a),max(b)), col="red", xlab="plate", ylab="score")
points(z,b, col="blue")
lines(c(-1:(length(a)+2)), rep(median(a), (length(a)+4)), col="red")
lines(c(-1:(length(a)+2)), rep(median(b), (length(a)+4)), col="blue")
quartz()
barplot((b-a), names.arg=z, ylab="dynamic range", xlab="plate")
}

boxplot.byPlate <- function (x)
{
par(mfrow=c(3,1))
boxplot(x$fl1 ~ x$Plate384, xlab="Plate", ylab="Luciferase Units")
title("Raw Data")
boxplot(x$fl1/x$rl1 ~ x$Plate384, ylim=c(0,400), xlab="Plate", ylab="Luciferase Ratio")
title("Dual Channel Ratio of Raw Data")
boxplot(x$fl1score ~ x$Plate384, xlab="Plate", ylab="Z-score")
title("Z-scores")
par(mfrow=c(1,1))
}

simple.scatterplot <- function (x, y, ...)
{
## function modified from John Verzani's package Simple
def.par <- par(no.readonly = TRUE)
n <- length(x)
xhist <- hist(x, sqrt(n), plot = FALSE)
yhist <- hist(y, sqrt(n), plot = FALSE)
top <- max(c(xhist$counts, yhist$counts))
xrange <- c(min(x), max(x))
yrange <- c(min(y), max(y))
nf <- layout(matrix(c(2, 0, 1, 3), 2, 2, TRUE), c(3, 1),
c(1, 3), TRUE)
layout.show(nf)
par(mar = c(3, 3, 1, 1))
plot(x, y, xlab = "x", ylab = "y", ...)
lines(c(-40:40), c(-40:40), col="red")
legend(-18,15, c("correlation coefficient", round(cor(x, y, use="complete.obs"), digits=3)))
}

```

```

    par(mar = c(0, 3, 1, 1))
    barplot(xhist$counts, axes = FALSE, ylim = c(0, top), space = 0,col = gray(0.95))
    par(mar = c(3, 0, 1, 1))
    barplot(yhist$counts, axes = FALSE, xlim = c(0, top), space = 0,col = gray(0.95), horiz = TRUE)
    par(def.par)
  }

% --- Source file: man/center.byplate.Rd ---
\name{center.byplate}
\title{Median centering of RNAi screening data in 2D}
\description{'center.byplate' calculates the plate statistics for RNAi screens
}
\usage{
center.byplate(x)
}
\arguments{
  \item{x}{merged luminescence dataset}
}
\details{
  For each plate the median and mad are calculated individually,
  output is a list with gene annotations and plate statistics
}
\seealso{
\code{\link[CellScreen::z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
}
\
\eof

% --- Source file: man/center.bywell.Rd ---
\name{center.bywell}
\title{Median centering of RNAi screening data in 3D}
\description{'center.bywell' calculates the plate statistics for RNAi screens
}
\usage{
center.bywell(x)
}
\arguments{
  \item{x}{processed dataset in 2D}
}
\details{
  To normalize 'edge-effects', median and mad are calculated based on well positions of the screening
  plates.
}
\seealso{
\code{\link[CellScreen::center.byplate]{center.byplate}}
\code{\link[CellScreen::z.score.3D]{z.score.3D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{

```

```

data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
screenZscoresBscores <- b.score(screenZscores)
screenCentered3D <- center.bywell(screenZscoresBscores)
}
\
\eof

% --- Source file: man/channel.norma.Rd ---
\name{channel.norma}
\title{Reporter channel normalization}
\description{'channel.norma' calculates the ratio of the reporter and co-reporter channels}
}
\usage{
channel.norma(x)
}
\arguments{
\item{x}{merged dataset}
}
\details{
}
\seealso{
\code{\link[CellScreen::screen.reader]{screen.reader}}
}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenNorma <- channel.norma(screenMerge)
}
\
\eof

% --- Source file: man/histo.byplate.Rd ---
\name{histo.byplate}
\title{Analyze data in histogram per plate}
\description{'histo.byplate' shows the distribution of scores in a histogram for a selected plate.}
}
\usage{
histo.byplate(x, plate = 34, data.col = x$f11 score)
}
\arguments{
\item{x}{normalized dataset}
\item{plate}{plate to be analyzed}
\item{data.col}{data column to be plotted}
}
\seealso{
\code{\link[CellScreen::z.score.2D]{z.score.2D}}
}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)

```

```
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
histo.byplate(screenZscores)
}
\
\eof

% --- Source file: man/image.byplate.Rd ---
\name{image.byplate}
\title{Plot RNAi screen phenotypes per plate}
\description{'image.byplate' generates a false-color representation of scores for a selected plate.}
}
\usage{
image.byplate(x, plate = 34, data.col = x$f11score)
}
\arguments{
  \item{x}{normalized dataset}
  \item{plate}{plate to be analyzed}
  \item{data.col}{data column to be plotted}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
image.byplate(screenZscores)
}
\
\eof

% --- Source file: man/medmad.byplate.Rd ---
\name{medmad.byplate}
\title{Analyze signal to noise ratio}
\description{'medmad.byplate' shows the distribution of med/mad values for all plates.}
}
\usage{
medmad.byplate(x)
}
\arguments{
  \item{x}{normalized dataset from 'z.score.2D'}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
medmad.byplate(screenZscores)
}
```

```
}
\  

\eof

% --- Source file: man/plate.merge.Rd ---
\name{plate.merge}
\title{Merging Raw Data Files}
\description{'plate.merge' reads files from one folder and merges them into one
}
\usage{
plate.merge(x)
}
\arguments{
\item{x}{folder with raw data}
}
\seealso{
\code{\link[CellScreen::screen.reader]{screen.reader}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\  

\eof

% --- Source file: man/screen.histo.Rd ---
\name{screen.histo}
\title{Analyze data in histogram}
\description{'screen.histo' shows the distribution of scores in a histogram for the whole dataset.
}
\usage{
screen.histo(x, data.col = x$flavg)
}
\arguments{
\item{x}{normalized dataset}
\item{data.col}{data column to be plotted}
}
\seealso{
\code{\link[CellScreen::z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
screen.histo(screenZscores)
}
\  

\eof

% --- Source file: man/screen.image.Rd ---
\name{screen.image}
\title{Plot RNAi screen phenotypes }
\description{'screen.image' generates a false-color representation of scores for all plates.}
\usage{
screen.image(x, data.col = x$flavg, threshold1 = -7, threshold2 = 7)
}
}
```

```

\arguments{
  \item{x}{normalized dataset}
  \item{threshold1}{the threshold above which z-scores are cut}
  \item{threshold2}{the threshold below which z-scores are cut}
  \item{data.col}{data column to be plotted}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
screen.image(screenZscores)
}
\
\eof

```

```

% --- Source file: man/screen.qqPlot.Rd ---
\name{screen.qqPlot}
\title{Analyze data in qq-plot}
\description{'screen.qqPlot' shows the distribution of scores in a Q-Q-plot.}

```

```

\usage{
screen.qqPlot(x, data.col = x$flavg)
}

```

```

\arguments{
  \item{x}{normalized dataset}
  \item{data.col}{data column to be plotted}
}

```

```

\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{

```

```

data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
screen.qqPlot(screenZscores)
}
\
\eof

```

```

% --- Source file: man/screen.reader.Rd ---
\name{screen.reader}
\title{Read RNAi screen raw data}
\description{'screen.reader' reads the raw luminescence data from RNAi screens and generates slots for further analysis.}

```

```

\usage{
screen.reader(w,x,y,z)
}

```

```

\arguments{
  \item{w}{raw luminescence dataset}
  \item{x}{list of targeted genes}
  \item{y}{list with flagged elements}
  \item{z}{list with phenotypes from other screens}
}
\details{
  lists are read, merged and slots for further analysis are created
}
\seealso{
\code{\link[CellScreen::center.byplate]{center.byplate}}
\code{\link[CellScreen::z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
}
\
\eof

% --- Source file: man/write.out.Rd ---
\name{write.out}
\title{write out screening data}
\description{'write.out' saves processed data as text-file.}
\usage{
write.out(x)
}
\arguments{
  \item{x}{dataset}
}
\seealso{
\code{\link[CellScreen::z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
write.out(screenZscores)
}
\
\eof

% --- Source file: man/z.score.2D.Rd ---
\name{z.score.2D}
\title{Calculate RNAi screen phenotypes}
\description{
  'z.score.2D' calculates z-scores for individual wells based on plate median and mad and automatically
  annotates the phenotypes.
}
\usage{
z.score.2D(x)
}

```

```

\arguments{
  \item{x}{plate centered luminescence dataset}
}
\details{
  z-scores are calculated individually per reporter channel, output is a list with phenotypes and annotations.
}
\seealso{
\code{\link[CellScreen::center.byplate]{center.byplate}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
}
\
\eof

% --- Source file: man/z.score.3D.Rd ---
\name{z.score.3D}
\title{Calculate RNAi screen phenotypes in 3D}
\description{'z.score.3D' calculates z-scores for individual wells based on well median and mad and
automatically annotates the phenotypes.
}
\usage{
z.score.3D(x)
}
\arguments{
  \item{x}{plate centered luminescence dataset}
}
\details{
  z-scores are calculated individually per reporter channel, output is a list with phenotypes and annotations.
  Requires dataframe with already calculated z-scores in 2D.
}
\seealso{
\code{\link[CellScreen::center.bywell]{center.bywell}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
screenZscoresBscores <- b.score(screenZscores)
screenCentered3D <- center.bywell(screenZscoresBscores)
screenZscores3D <- z.score.3D(screenCentered3D)
}
\
\eof

% --- Source file: man/b.score.Rd ---
\name{b.score}
\title{Calculate B-scores for RNAi screen phenotypes}
\description{'b.score' calculates B-scores based on Tukey's two-way median polish
}

```



```

\usage{
b.score(x)
}
\arguments{
  \item{x}{plate centered luminescence dataset}
}
\seealso{
\code{\link[CellScreen::center.bywell]{center.bywell}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
screenZscoresBscores <- b.score(screenZscores)
}
\
\eof

```

```

% --- Source file: man/HeatMapRows.Rd ---
\name{HeatMapRows}
\title{Analyze row artefacts}
\description{'HeatMapRows' generates a false-color representation of z-scores for all plates.}
\usage{
HeatMapRows(x, threshold1 = -7, threshold2 = 7)
}
\arguments{
  \item{x}{normalized dataset}
  \item{threshold1}{the threshold above which z-scores are cut}
  \item{threshold2}{the threshold below which z-scores are cut}
}
\seealso{
\code{\link[CellScreen::z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
HeatMapRows(screenZscores)
}
\
\eof

```

```

% --- Source file: man/hits.perplate.Rd ---
\name{hits.perplate}
\title{Analyze hits per plate}
\description{'hits.perplate' generates a barplot for the number of hits with scores > 2 or < -2 per plate}
\usage{
hits.perplate(x, data.col = x$flavg)
}
\arguments{
  \item{x}{normalized dataset}
}

```

```

\item{data.col}{data column to be plotted}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
hits.perplate(screenZscores)
}
\
\eof

```

```

% --- Source file: man/HeatMapCols.Rd ---
\name{HeatMapCols}
\title{Analyze column artefacts}
\description{'HeatMapCols ' generates a false-color representation of z-scores for all plates.}
\usage{
HeatMapCols(x, threshold1 = -7, threshold2 = 7)
}
\arguments{
\item{x}{normalized dataset}
\item{threshold1}{the threshold above which z-scores are cut}
\item{threshold2}{the threshold below which z-scores are cut}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
HeatMapCols(screenZscores)
}
\
\eof

```

```

% --- Source file: man/dynamicRange.Rd ---
\name{dynamicRange}
\title{Analyze the dynamic range of screening plates}
\description{'dynamicRange ' generates two graphical representations of the dynamic range as determined by the spread between control dsRNA treatments targeting a positive regulator (Stat92E) and control dsRNA treatments targeting a negative regulator (Socs36E). In the first plot, the z-scores for these regulators for every plate are shown as well as their median as a straight line. The second barplot depicts the difference between these scores for every plate.}
\usage{
dynamicRange(x, data.col = x$f1 score, ctrl1 = 2, ctrl2 = 26)
}
\arguments{
\item{x}{normalized dataset}

```

```

\item{data.col}{data column to be plotted}
\item{ctrl1}{first control to be plotted (positive regulator), position by 'Element384' number}
\item{ctrl2}{second control to be plotted (negative regulator), position by 'Element384' number}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
dynamicRange(screenZscores)
}
\
\eof

```

```

% --- Source file: man/boxplot.byPlate.Rd ---
\name{boxplot.byPlate}
\title{Analyze the data distribution of screening plates in boxplots}
\description{'boxplot.byPlate' generates three graphical representations. The first boxplot shows the raw luciferase values, the second the dual-channel normalized values and the third boxplot shows the z-scores for one luciferase channel.}
\usage{
boxplot.byPlate(x)
}
\arguments{
\item{x}{normalized dataset}
}
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
boxplot.byPlate(screenZscores)
}
\
\eof

```

```

% --- Source file: man/simple.scatterplot.Rd ---
\name{simple.scatterplot}
\title{Analyze the reproducibility between two screening replicates channels}
\description{'simple.scatterplot' plots the scores from one channel against the other replicate channel and calculates Pearson's correlation coefficient.}
\usage{
simple.scatterplot(x)
}
\arguments{
\item{x}{normalized dataset}
}

```

```
\seealso{
\code{\link[CellScreen:z.score.2D]{z.score.2D}}
\references{http://www.dkfz.de/signaling/jak-pathway/}
\author{Patrick Mueller <pmuller@mpi-bpc.mpg.de>, Michael Boutros <m.boutros@dkfz-heidelberg.de>}
\examples{
data(JAKrawdata, Genelist, phenotypes, JAKflags)
screenMerge <- screen.reader(screenData, geneList, flagList, phenoData)
screenCentered <- center.byplate(screenMerge)
screenZscores <- z.score.2D(screenCentered)
simple.scatterplot(screenZscores)
}
\
\eof
```

Supplementary Tutorial. Computational analysis of the genome-wide RNAi screen dataset.

The software package provided in Supplementary Script 1 was developed to provide a set of novel statistical approaches as well as tools facilitating the analysis of multi-channel datasets from reporter-based high-throughput RNAi screens. For the implementation of data analysis tools, the computational language and environment R was chosen (R-Development-Core-Team 2004) due to its ease in statistical computing, data handling, graphics and data distribution. The package enables the user with the initial steps of reading raw data from plate-based RNAi screenings, normalization of screening results in ‘2D’ and ‘3D’ as well as calculation of scores for candidate selection. Graphical representations of a multitude of statistical parameters serve to validate the quality of the screen. The operation of the package is described in the following.

MATERIALS

EQUIPMENT

Standard computer

Computer operating systems: Mac OS X, Windows or Linux

R software (Version 2.0.1 or higher)

Additional R packages (*Simple*, *RColorBrewer*, *geneplotter*, *prada*)

EQUIPMENT SETUP

The R software is available at <http://www.r-project.org/>. Download and install the version appropriate for your operating system. Additional R packages can be uploaded directly via the R application menu ‘Packages’ using the package manager. Recommended packages for use with are *Simple*, *RColorBrewer*, *geneplotter* and *prada*. R runs as a command line program. The following commands in the section PROCEDURE are described in the format

R> command

The ‘*R>*’ in the PROCEDURE section is written solely for readability and only the command needs to be entered for execution.

PROCEDURE

Setting up the application

1 Launch R by double-clicking the application icon. Check that all necessary additional packages are installed and loaded by clicking ‘Packages’ in the application menu. Also install the package described in Supplementary Script 1 (e.g. by copying and pasting the functions).

Preprocessing the dataset

2 Datasets can be assembled *de novo* from plate reader files (A). Alternatively, already merged datasets can be used (B).

A) Merging raw data from plate reader files

- i) Create a directory (e.g. “.../in” and move all the relevant raw data files (data from one channel in a subdirectory FL, the other in a subdirectory RL) to that directory)
- ii) Make this the working directory in R (i.e. use “File -> Change Dir -> .../in)
- iii) merge individual files by

```
R> mergedRawData <- plate.merge()
```

B) Usage of preprocessed datasets

- i) The following steps assume a sample dataset from a genome-wide *Drosophila* RNAi screen.

```
R> data(JAKrawdata, Genelist, phenotypes, JAKflags)
```

can be used to load the relevant data files all at once. For example, *JAKrawdata* includes the values and plate positions of FL and RL channels in duplicate and yields an object *screenData*. *Genelist* is the gene annotation list for the plate positions and yields an object *geneList*.

phenotypes contains information about phenotypes from other published and unpublished screens for the respective genes and yields an object *phenoData*. *JAKflags* is an annotation file with flagged data (detailed information is contained in the header of this file) and yields an object *flagList*.

ii) These files then need to be merged by:

```
R> screenMerge <- screen.reader(screenData, geneList,flagList, phenoData)
```

which puts the data from the appropriate files into annotation slots.

Data normalization

3 Single (A) or multi-channel (B) datasets can be analyzed

A) Single channel normalization

i) For intra-plate normalization, omit step (B) i) and directly proceed with step (B) ii)

B) Dual channel normalization

i) get the FL/RL channel ratio

```
R> screenNorma <- channel.norma(screenMerge)
```

ii) For intra-plate normalization, calculate the median and the median absolute deviation (MAD) for all plates

```
R> screenCentered <- center.byplate(screenNorma)
```

iii) Calculate z-scores for all wells

```
R> screenZscores <- z.score.2D(screenCentered)
```

- iv) Output of the calculated datasets can be generated at any time to write a tab delimited text-file named “ScreenDataOut.txt” into the working directory

```
R> write.out(screenZscores)
```

Graphical data exploration for quality control and detection of spatial artefacts

4 After initial data normalization, it is recommended to visualize the complete dataset for the identification of possible artefactual data distribution. The data distribution of the values before and after normalization can be best analyzed by boxplots. The command

```
R> boxplot.byPlate(screenZscores)
```

will draw boxplots for the raw as well as the plate median centered values. Plotting a histogram reveals the width of the data distribution.

```
R> screen.histo(screenZscores)
```

Ideally, the width of data distribution would be narrow with the rationale that most dsRNA probes do not affect the reporter readout and will not yield a phenotype.

Plotting the experimental against the theoretical quantiles in a normal quantile-quantile plot (Q-Q-Plot) furthermore shows the symmetry or asymmetry of the dataset.

```
R> screen.qqPlot(screenZscores)
```

Assess the reproducibility between replicate datasets in simple scatterplots, where Pearson’s correlation coefficient will also be calculated. For example, plot the z-scores for the FL1 channel (fl1score) against the FL2 channel (fl2score):

```
R> simple.scatterplot(screenZscores$fl1score, screenZscores$fl2score)
```

Next, plot a false-color heatmap image of the scores to visually inspect the dataset for the distribution of potential hits.

```
R> screen.image(screenZscores)
```


5 Looking at the data distribution within individual selected normalized plates can also be very informative. As a first step, plot a histogram for a selected plate, e.g. plot the distribution of z-scores for the F11 channel (`f11score`) for plate 34:

```
R> histo.byplate(screenZscores, plate=34, data.col=screenZscores$f11score)
```

As described above for the whole dataset, plot false-color heatmap images for single plates.

```
R> image.byplate(screenZscores, plate=34, data.col=screenZscores$f11score)
```

6 To further assess the data quality, calculate the dynamic range, i.e. the spread between the positive and negative regulator spiked-in controls for every plate. Plotting the dynamic range will help visually to identify outlier plates with lower dynamic range spreads

```
R> dynamicRange(screenZscores)
```

This command and function yields a plot for the z-scores of a negative and a positive regulator control as well as the difference between the z-scores shown in a barplot.

It is also very helpful to analyze the ratio of plate median to plate MAD for all screening plates and channels, which can help to identify plates with higher and lower data quality.

```
R> medmad.byplate(screenZscores)
```

7 The heatmap image of the whole dataset from step 3 already gives an idea about spatial artefacts hidden in the data. Use

```
R> HeatMapCols(screenZscores)
```

to further systematically analyze potential column artefacts. The next step allows the same analysis to look for row artefacts:

```
R> HeatMapRows(screenZscores)
```

Correction of spatial artefacts

8 Steps 3-7 are suitable to detect spatial artefacts. To smoothen the data and to account for systematic as well as non-systematic errors, Tukey's two-way median polishing procedure can be applied to calculate the so-called B-score. Calculate the B-score by median polishing the row as well as the column effects:

```
R> screenZscoresBscores <- b.score(screenZscores)
```

9 To correct for potential systematic errors stemming from a certain well position bias within the screening plates, calculate z-scores by centering the values by well position through the dataset in a third dimension. First calculate median and MAD by well position in '3D':

```
R> screenCentered3D <- center.bywell(screenZscoresBscores)
```

Note: This is a computationally time-consuming process, which – depending on the computer memory – can take several hours to be completed.

Next, derive the z-scores by well position:

```
R> screenZscores3D <- z.score.3D(screenCentered3D)
```

Post-analysis processing

10 After data normalization and smoothening, extract the number of dsRNA treatments with significant scores by

```
R< hits.perplate(screenZscoresBscore$f1score)
```

Supplementary Script 2. ImageJ macro for approximation of cell numbers used for growth curve analysis of HeLa cells after siRNA treatment.

```
requires("1.33n");
dir = getDirectory("Choose a Directory ");
list = getFileList(dir);
start = getTime();
setBatchMode(true);
for (i=0; i<list.length; i++)
    {
    path = dir+list[i];
    showProgress(i, list.length);
    open(path);
    run("8-bit");
    run("Find Edges");
    setThreshold(19, 42);
    run("Analyze Particles...", "minimum=100 maximum=999999 bins=20
show=Nothing clear summarize");
    run("Close");
    print("Count: "+nResults);
    }
```

Supplementary Table 1. Sequence and cytological information for RNAi screen hits.

dsRNA ID	Amplicon primer 1 ^a	Amplicon primer 2	No of efficient siRNAs ^b	Target gene (Symbol)	Cytological location
HFA00627	TGC CTG TTT TCT GGA AAT ATG	CTC GCT GGG TTT CAT GGT	51/496	Art2	24E1
HFA11324	TCG AAC TCA CGT TCG AGT	ATC ATC TTC GGG ATG GAT AAC	61/489	asf1	76B9
HFA04919	GAG ATA CCC CGT GAT GAC A	CTT GGG AAT ACG CAC AAA GA	87/487	bin3	42A13--14
HFA16914	AGG TGC TGG TGG AAA AGA A	ACC CGT CAC CCG GAA AG	60/496	bon	92F2--3
HFA16596	TAT TTG CTG TCA GCC TCT G	TGG TCC GTC CTC AGC ATC	81/496	Caf1	88E3
HFA14173	CGC CCT GAT CTT TGT GGG	GGA CGA GTA CAT CGC AAT G	139/494	CG10007	87A4
HFA09691	GCA CCA CCT CGT TGA AGA	GGG CAG CCA CAT CGG T	72/484	CG10077	65D3--4
HFA02102	GAA CTT CAT TTG GAA GCG TTT	CTT GCG CCG GAA CCA G	103/497	CG10730	38B2
HFA09807	GCC GCC GGT ACC GTC	AAG TAG GTG GGC GAT TCC	63/495	CG10960	69E5--6
HFA11648	CCG TGG CCA CAG GAA CA	CAG TCC TGT TCA TGT GGA AAT	51/242	CG11307	78E1
HFA06070	TTG TCT GGC TGT GTC TGT C	GAC AAT CCT TGG CCC AAT AAC	85/459	CG11400	54A1
HFA14317	ATG GCA TCC CCA GTA GTC A	GTG ACT TTG ATG ATC TGG ATT C	28/312	CG11501	99B1
HFA19417	GCC GAC GAA CAG CCA AA	TCG CAC ACC TCG GGA C	64/486	CG11696	10C7
HFA14478	TAA CGG TGA CGG AAC CCA	CCG AAT CCT CGA TGG GTT	78/498	CG12213	87A3
HFA20970	GCC AAA ATC AAG CGA ATC AG	CTT AAT TGC CTG CAC CTC C	50/114	CG12460*	heterochromatin
HFA19459	ATC GGC TGC GTG AGA AC	TTT GTT GGC CAA ACT TTA CA	19/181	CG12479	12E2
HFA01920	GAT TGG ACG CTT CTG TTT GA	GTT GAA ACA TTG CTG GGT GA	112/494	CG13243	35D4--5
HFA10017	TGG CTG CCA TGC AGA AG	CCA ATT TCG GCA CCG TAG	73/391	CG13473	70F3
HFA04144	AGT GGC AGC GGA GGT G	CGC CTC GCA GTG GGT T	19/256	CG13499	58B1
HFA14742	AAA ATA AAT GGA GTA ACT TCC CC	TAC GCC TCG CAC TCC A	34/497	CG14247	97D1
HFA17927	CGC AAT GTG GAG GTG AAG	ACT GAA ATA CGA GCC GAT C	52/490	CG14434	6D7
HFA17993	TTT GAG GGC CCA CAA TGT	TGG CAA GTC GCA ACT TTA C	122/475	CG15306	9B7
HFA00432	CAA AGG CAC CTG GTT TGT G	CAG TAG CGC AGA CGT TG	19/143	CG15418	24A2
HFA00449	GGT ATT ACT CTG TTC CGA TTG	CTT CCA GGT TTT TGT GTA TGT C	30/217	CG15434	24F3
HFA15093	GGC AAA GAT CCC AAG CAG	GTT GAA GGT GCA GCA GAA G	58/283	CG15555	100B9
HFA06577	CAG CCA TCG ATT GGA ACA G	CTC CAA GTG CCA GAA CAT AAA	77/477	CG15706	52F11
HFA18090	GGC CAC AAG CAT GGT CG	CTT TGC CCT TGC ACT TCT	51/500	CG15784	4F10
HFA18561	TCG CCC ATG GTG CTA GA	CGA TCC ACG GTG ATT ACA G	72/477	CG16903	2C10
HFA02552	CAG ACT CCT ACC TCG TTT TG	AAC ATG CGC TCC AGA TAG T	54/495	CG16975	34A7--8
HFA10258	GCC AAG AGA CGG AGA AGA	TAC GGA TGC TGG TTG ATG T	3/155	CG17179*	3L
HFA02623	CCC AGG GCC ATT TGG ATT T	TCC TTT AAG CGC TGC ATG	71/486	CG17492	37B10--11
HFA15304	GGG CAT GGC CTC ATT ACA	CGG CGA TAT TTG CTG GTC	78/475	CG18112	99C2
HFA21006	GTG GCG CAC CGG AAA G	GAT GAA CTT CAT TGT TGT TGA AA	50/114	CG18160*	U
HFA06272	TGA CGA AGC ATA TAC AAG GAT A	TGG GTT TTT CTG GTG AAA CAA	111/489	CG30069	50E2--3
HFA06935	GTT TGC ATC GGC CAA ACC	GTG TCA GAG AAA TTC ACT AAG TA	62/463	CG30122	55E3
HFA00563	AAT ACG TTT CGG TCA CGA TT	GTA TCT GTA CTT GGT AGA GTA GT	69/326	CG3058	24F1
HFA15507	CCC CGA GCT GAA TCC CA	CTT CAT GCG GTT GAT GAC TA	8/197	CG31005	100B8
HFA15369	CGT AAG TGC TAG TTC CTC TG	TCC CGA GCG TCC CTT T	34/488	CG31132	95F12--13
HFA16032	CCC ACG GAG CTG TTC TTT	AAA CGA CTA CCC AGG ACA TT	63/495	CG31132	95F12--13
HFA15235	AGG CAT CTG CAG ATT CTC T	GAG GAA TGG GAA TGG ATG AAG	112/488	CG31358	87A5
HFA00415	GTG ATG GGT CCC GGG ATG	CGT CTT GTC ACG ATT CTT T	28/159	CG31694	23B7--8
HFA09966	CCG CCA CAA TGA TAA CCA AC	CGG GTG CGT GAA GAG T	68/477	CG32406	65A2--3
HFA19906	ATC TGT TGA ACG CCG AGG	GGT ATC GGT GAA GTT CTT CTC	39/495	CG32573	14F5
HFA15470	TTG TCG CGA CCT TCC CA	ACT TCT TGG AGC AGA TCT TG	66/500	CG3281	87A3
HFA10378	CGG ACA CCG GCT ATG TG	ATG TTC TTG GCC GAG TCA A	70/482	CG3819	75E6
HFA10395	TAC TCA AGG ATC GCG ATA TC	GGC TGG GTG TGG GAG TG	53/484	CG4022	67B4--5
HFA20930	GCA GGA CGT TCG GAA TAT C	TCC CAT TAC AGA CTT TTG ATT G	199/540	CG40351*	U
HFA19892	GGC GCC ACA TGT GCA TG	GCC GCT GCC CAT ATA CTT	97/480	CG4349	11D11
HFA10420	IGT GGC TGT CGC TTA TCT T	AAA AAT ATA CAG CCG TTT CCT T	56/481	CG4446	67B2
HFA19909	ACC CAG CTA AAT CCT ACA ATG	ACT CCA GAT GCT GGG TCA	55/496	CG4653	15A3
HFA04488	TTG ACG GAT TGC CAC ATC T	GCC TCC CCG TCC AAG T	75/482	CG4781	60D10
HFA15673	TGG GCT CCG CAG AGA TA	CAA GTA GAG GAG CCC GAT	105/492	CG4907	94C2
HFA16036	TCT TTG TCA TCA AAT GCT ACT C	CAT CCG GCC CAT GCA TT	102/487	CG6422	96B17
HFA10635	TTG AAC ATC GTG GCT TCT TT	CCT CGC AAA CTC GAT GC	39/148	CG6434	77B4
HFA16145	CAA CAA CAT GCT GGG CTT C	CGA AGT TCG AGC CGA CA	118/468	CG6946	86F8--9
HFA20054	GAG CCG GCG ATC ATC TT	CTC GGC GGC GAT CAC	33/452	CG7635	18A6
HFA09675	GAT GAG AAG GAC GAG AAG AG	CTT GAT GCG GCA ATG GAC	56/481	CG8108	67C11--D1
HFA20148	ATA GGT TCA ACA CGA TCC CC	GAA GGC TGG TGT TAG TTT TG	93/492	CG9086	15C5--6
HFA11946	ACT TGC GTG GAG GAA CTA A	ATG CTT AGA GTT CTT CGG T	144/490	CkHalp	80D1
HFA20230	AGC TCG AGG ACA ATC CAC	GGC TGA CTT TCA CAG TAG AC	27/152	CkHbeta	10E3
HFA09995	CGT ACG ATG ATG CAC TGG	GAA CCG GCA GAA TGG TTG	37/499	comm3	71E3--4
HFA16617	GGC AGT GGC AGC TCT GA	CTC GGG TCC GGT GAA CT	30/254	CiBP	87D8--9
HFA19583	CGT CTG CGC AGT GAT CC	TGG GCT CCG ATG GAT AGA	39/480	dome	18D13--E1
HFA08714	AGC GAC GAG GAA GAT GTG	TGA CAA ATG TGG CCT CTG G	65/476	dre4	62B7
HFA20983	TTG GAA AAT CGA GAG GAT TTA A	CAC ATT TTT CGA ATT CAA TTG TC	159/488	eIF-4B*	U
HFA04096	CGT CTA ATG AGG CAA AGA AAC	CCG TTT TTG CCA CTT TAA CC	84/487	enok	60B10
HFA01091	TCG TGA TGG TGT TGG TGA C	TCC ACT GAA AGT GCT TTG GT	48/220	HDC01676	30D1
HFA11427	GGG CGA ATG CAC GGA AT	TGG CAT ACC TCG AAT AAC TG	105/477	HDC11198	77D4
HFA20340	TAA TCG A CG ATC AGG A AC AG	GTG TGG CCT CCG AGG TG	52/493	hop	10B5--6
HFA00357	CGT CCC CCG GTT TTA CG	ATC AGC CAG TCT TGA ATA GTC	75/488	lpk2	21E2
HFA04167	ATA AAA GGC GCC AAG GTG A	TCA CCT GCA TTC CCG TTT C	18/202	jbug	59A3
HFA07637	GAC GGC CTT CAA TTC CTA TG	GCG ACG AGG AGA GTG TG	10/228	kn	51C2--3
HFA19450	TGC TGC GCA AGC GAC	CAT TTG GCT GGA AGA TGA CA	38/496	l(1)G0084	18D8--11
HFA16984	CAC AAA GCC GCT GAA CAG	TTC GTG GTT ACA CAC ACA GT	88/496	larp	98C3--4
HFA07247	CCG CCG GAA CGA CTT	TGA TCG CTT ATC ATC GTA TAT TA	35/377	lig	44A4
HFA15370	ACT AGT AGC AGT CAG TCC TC	GCG CCA CCG TTG CTA T	30/486	mask	95F3--5
HFA20582	ACA GCA TTC GGG TGG TAA A	GCC ATC CGA AGT TGA TCG	58/473	mst	20A1
HFA20357	AAC CAG AAC CAG AAT CAA AAT G	GTT TCC AGC GCG ATT ATT G	27/118	nonA	14B18--C1
HFA03384	GCC TGG ATG GAG TTG TTT G	GGA CTT ATG GGC TGA TTG AAC	87/500	Nup154	32C5
HFA15220	AGC GGC TGC AGG AGT TC	TTC TTA TTA CTG GCC ACA TCA T	36/167	Obp93a	93C1
HFA07660	CAC GTT CTG CCG TAG CC	GCT TGG GAT CCG CTA AAT C	60/324	par-1	56D9--11
HFA16795	TTG TGG GTA AAT TTT TAC AGA AG	GCA ATT CCC CCG AGT AGT	12/118	Pp1alpha-96A	96A5
HFA16344	CGG ATC CCG AGC ACC C	GCG ATG GAG CTG CTG G	32/469	PP2A-B'	90F4--5
HFA08683	CTT GAC GCT GAA GAA CCC	CCT GGA ATT GGA TCG ATG C	72/495	Ptp61F	61F7--62A1
HFA00777	GGC AAC CAC TCC ACG CA	TCC TGG CCA GCC GTG T	28/244	Rab5	22E1
HFA00784	AGA GCC GCC GAA ACA AC	GCG TTG GTT TCA GTA GAG G	98/487	Rrp1	23C3--4
HFA02455	CAG CAG TAA AGC ACT TTC AA	CGG ATT CCG GCA TGG C	38/490	Soes36E	36E6
HFA20587	GAG TAC AAG CAT GTG TAC AAG	CCT CTT GGT GGA GGT AGT G	31/359	sol	19F5
HFA16870	CTT GCC CAA AAC TAC AGT TAC	CGA CTG TGG GTG GAT TGT T	64/479	Stat92E	92F1
HFA11298	AAG GAA AGC GCA TTT CGT	AAA TCC ATA TCC ACT TCC TCA C	114/481	Taf2	67D1
HFA11098	ATC CCT CAA ATC CCA GTT CC	AAA GTG GCG CTG TGG TG	53/319	TSG101	73D1

^aComplete amplicon information can be obtained at <http://mai.dkfz.de>^bEfficiency calculated based on Reynolds et al., (2004). All siRNAs with score of 6 or higher were counted as efficient.

* information according to release 2 of the Berkeley Drosophila Genome Project (not available in later releases)

Supplementary Table 2. Human homologs of *Drosophila* genes with JAK/STAT phenotypes.

<i>Drosophila</i> Gene	BLASTP	Identity [%]	Human Gene	RefSeq
Art2	1.60E-77	44.2	protein arginine N-methyltransferase 4	NP_062828.2
asf1	3.20E-68	61.7	ASF1 anti-silencing function 1 homolog A	NP_054753.1
bin3	8.80E-49	34.3	hypothetical protein FLJ20257	NP_062552.2
bon	3.60E-45	30.5	tripartite motif-containing 33 protein	NP_056990.2
Caf1	0	91.9	retinoblastoma binding protein 4	NP_005601.1
CG10007	8.80E-50	34.5	chromosome 2 open reading frame 18	NP_060347.2
CG10077	7.00E-171	67.7	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	NP_004387.1
CG10960	3.40E-78	36.9	solute carrier family 2, (facilitated glucose transporter) member 8	NP_055395.2
CG11696	5.10E-33	29.9	zinc finger protein 502	NP_149987.2
CG12460	1.80E-17	54.0	splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)	NP_005057.1
CG13473	3.90E-17	34.9	thioredoxin 2 precursor	NP_036605.2
CG15306	3.30E-27	45.1	microtubule-associated protein, RP/EB family, member 1	NP_036457.1
CG15418	1.40E-10	41.1	tissue factor pathway inhibitor 2	NP_006519.1
CG15434	1.60E-17	50.6	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8kDa	NP_002479.1
CG15706	6.60E-20	20.0	FLJ20160 protein	NP_060164.2
CG16903	2.00E-100	63.1	cyclin L1	NP_064703.1
CG16975	4.00E-123	49.4	l(3)mbt-like 2 isoform a	NP_113676.2
CG17492	0	48.3	zinc finger, ZZ type with ankyrin repeat domain 1	NP_543151.1
CG18112	1.80E-20	27.8	chromosome 14 open reading frame 133	NP_071350.2
CG30122	7.20E-42	40.9	E1B-55kDa-associated protein 5 isoform a	NP_008971.2
CG3058	2.60E-80	95.8	thioredoxin-like 4	XP_499552.1
CG31005	9.00E-100	52.5	trans-prenyltransferase	NP_055132.2
CG31132	0	49.9	bromo domain-containing protein disrupted in leukemia	NP_694984.2
CG31358	2.10E-51	44.2	stomatin-like 3	NP_660329.1
CG31694	2.00E-66	36.3	interferon-related developmental regulator 2	NP_006755.3
CG32406	5.50E-15	37.8	C1 domain-containing phosphatase and tensin-like protein isoform 3	NP_938072.1
CG3281	4.00E-41	31.7	zinc finger protein 91	NP_003421.1
CG40351	2.20E-94	56.6	PREDICTED: KIAA1076 protein	XP_037523.9
CG4349	4.60E-35	45.2	ferritin, heavy polypeptide 1	NP_002023.2
CG4446	1.90E-66	47.2	pyridoxal kinase	NP_003672.1
CG4653	8.80E-23	30.7	protease, serine, 2 preproprotein	NP_002761.1
CG4781	2.00E-17	33.8	PREDICTED: similar to KIAA0644 protein	XP_379800.1
CG4907	1.10E-47	28.9	phosphatidylinositol glycan, class N	NP_036459.1
CG6422	1.30E-71	53.8	YTH domain family, member 1	NP_060268.2
CG6434	6.20E-69	71.2	retinoblastoma binding protein 5	NP_005048.2
CG6946	4.00E-39	46.2	heterogeneous nuclear ribonucleoprotein F	NP_004957
CG7635	2.90E-76	62.1	stomatin isoform a	NP_004090.4
CG9086	0	32.0	ubiquitin protein ligase E3 component n-recognin 1	NP_777576.1
Ckl1alpha	4.00E-163	88.7	casein kinase II alpha 1 subunit isoform a	NP_001886.1
Ckl1beta	5.00E-107	89.2	casein kinase 2, beta polypeptide	NP_001311.3
CtBP	8.00E-152	72.4	C-terminal binding protein 2 isoform 1	NP_001320.1
dome	7.60E-15	28.2	sidekick 2	NP_061937.2
dre4	0	59.9	chromatin-specific transcription elongation factor large subunit	NP_009123.1
eIF-4B	4.70E-32	27.2	eukaryotic translation initiation factor 4B	NP_001408.1
enok	1.80E-94	33.4	MYST histone acetyltransferase (monocytic leukemia) 3	NP_006757.1
HDC01676	2.80E-15	61.0	cholinergic receptor, nicotinic, alpha polypeptide 7 precursor	NP_000737.1
hop	1.60E-59	26.7	Janus kinase 2	NP_004963.1
Ipk2	1.80E-26	33.6	inositol polyphosphate multikinase	NP_689416.1
jbug	5.30E-45	27.6	filamin B, beta (actin binding protein 278)	NP_001448.1
kn	0	69.7	early B-cell factor	NP_076870.1
l(1)G0084	7.40E-28	31.5	PHD finger protein 10 isoform a	NP_060758.1
larp	2.00E-103	48.1	KIAA0731 protein	NP_056130.2
lig	5.50E-25	32.8	ubiquitin associated protein 2 isoform 2	NP_065918.1
mask	0	74.0	multiple ankyrin repeats, single KH-domain protein isoform 1	NP_060217.1
mst	7.00E-52	29.7	misato	NP_060586.2
nonA	1.80E-60	40.6	splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)	NP_005057.1
Nup154	0	32.6	nucleoporin 155kDa isoform 1	NP_705618.1
par-1	0	54.1	MAP/microtubule affinity-regulating kinase 3	NP_002367.4
Pp1alpha-96A	8.00E-169	88.9	protein phosphatase 1, catalytic subunit, alpha isoform 1	NP_002699.1
PP2A-B'	0	78.9	delta isoform of regulatory subunit B56, protein phosphatase 2A isoform 1	NP_006236.1
Ptp61F	5.00E-32	37.9	hypothetical protein LOC9671	NP_055468
Rab5	4.90E-85	75.0	RAB5A, member RAS oncogene family	NP_004153.2
Rrp1	6.10E-82	55.2	APEX nuclease	NP_542380.1
Soes36E	4.80E-65	68.0	suppressor of cytokine signaling 5	NP_054730.1
Stat92E	6.40E-86	41.6	signal transducer and activator of transcription 5B	NP_036580.2
Taf2	0	52.5	TBP-associated factor 2	NP_003175.1
TSG101	4.30E-98	48.7	tumor susceptibility gene 101	NP_006283.1

Shown are human homologs of *Drosophila* genes with a BlastP E value of 10^{10} or less.

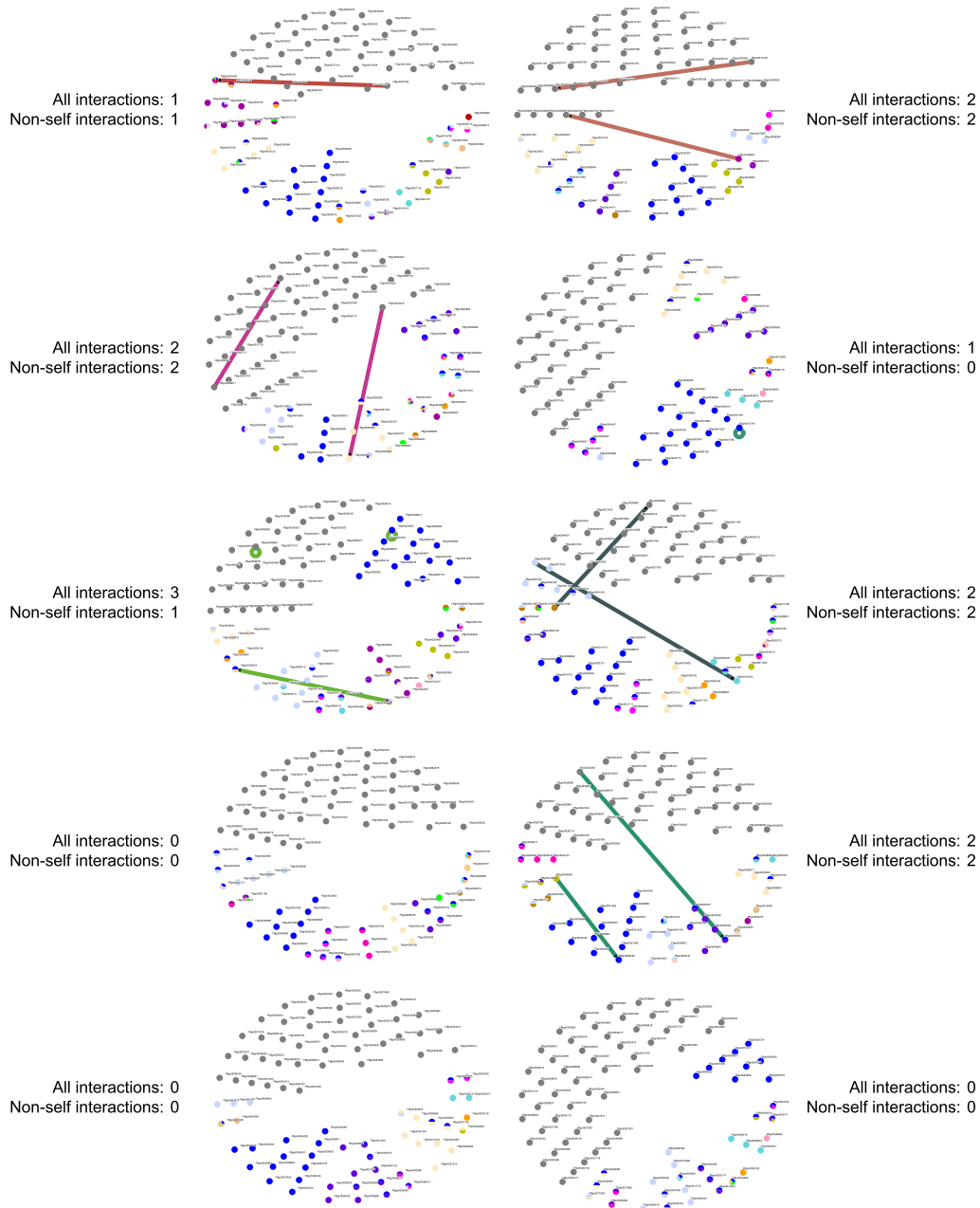
Supplementary Table 3. Human disease homologs of *Drosophila* genes with JAK/STAT phenotypes.

<i>Drosophila</i> Gene	BLASTP	Human Gene	RefSeq	Disease
bon	3.60E-45	tripartite motif-containing 33 protein	NP_056990.2	Thyroid carcinoma, papillary (2)
Cad1	9.90E-17	peroxin 7	NP_000279	Refsum disease (1)
CG10960	1.40E-39	erythrocyte/hepatoma glucose transporter	NP_006507	Glucose transport defect, blood-brain barrier (1)
CG11696	2.60E-25	zinc finger protein 41	NP_006051	Mental Retardation, X-linked nonsyndromic (1)
CG17492	4.80E-23	ankyrin, brain	NP_001139	Long QT syndrome 4 (1)
CG31132	2.70E-17	Lissencephaly-1 Gene	NP_000421	Subcortical laminar heterotopia (1)
CG31132	0	bromo domain-containing protein disrupted in leukemia	NP_694984	Leukemia (3)
CG31358	7.40E-38	Podocin	NP_055440	Nephrotic syndrome, steroid-resistant (1)
CG32573	7.80E-47	Protein Kinase C, alpha	NP_002728	Pituitary Tumor, invasive (1)
CG3281	2.00E-33	zinc finger protein 41	NP_009061	Mental Retardation, X-linked nonsyndromic (1)
CG40351	2.10E-14	Androgen Receptor-Associated Coregulator 267	NP_071900	Stons Syndrome, sporadic (1)
CG4349	4.60E-35	ferritin, heavy polypeptide 1	NP_002023.2	Iron overload, autosomal dominant (2)
CG4349	3.50E-35	ftih	NP_002023	Pancreatitis, hereditary (1)
CG4653	6.90E-22	Protease, Serine, 1	NP_002760	Stomatocytosis 1 (2)
CG7635	2.90E-76	stomatin isoform a	NP_004090.4	Nephrotic syndrome, steroid-resistant (1)
CG7635	5.70E-62	Podocin	NP_055440	Ret Syndrome, atypical (1)
CK1alpha	1.20E-23	serine/threonine protein kinase 9	NP_003150	Phosphoglycerate dehydrogenase deficiency (1)
CKBP	4.20E-24	3-phosphoglycerate dehydrogenase; 3pgdh	NP_006614	Wolman disease (1)
dros4	4.60E-46	Lipase A precursor	NP_000226	Schizophrenia, neuropsychologic defect in (2)
HDC01676	2.80E-15	cholinergic receptor, nicotinic, alpha polypeptide 7 precursor	NP_000737.1	SCID, autosomal recessive, T-negative/B-positive type (1)
hop	8.40E-52	Janus kinase 3	NP_000206	Atelosteogenesis, type I (2)
jbug	5.30E-45	filamin B, beta (actin binding protein 278)	NP_001448.1	Larsen syndrome (2)
jbug	5.30E-45	filamin B, beta (actin binding protein 278)	NP_001448.1	Spondylocarpotarsal synostosis syndrome (2)
jbug	5.30E-45	filamin B, beta (actin binding protein 278)	NP_001448.1	Frontometaphyseal dysplasia (1)
jbug	6.00E-102	actin-binding protein 280; abp280	NP_001447	Long QT syndrome 4 (1)
mask	5.30E-59	ankyrin, brain	NP_001139	Diabetes mellitus, type II (1)
par-1	5.30E-39	Oncogene Akl2	NP_001617	Insulin resistance, susceptibility to (1)
Pip61F	7.80E-79	Protein phosphotyrosylphosphatase 1B	NP_002818	Griscelli Syndrome (1)
Rab5	8.90E-23	ras-associated protein RAB27A	NP_004571	Muscular dystrophy, limb-girdle, type 2A (1)
sol	2.80E-33	calcium-activated neutral protease 3	NP_000061	Leukemia, acute promyelocytic, STAT5B/RARA type (2)
Stat92E	6.40E-86	signal transducer and activator of transcription 5B	NP_036580.2	Growth hormone insensitivity with immunodeficiency (2)
Stat92E	6.40E-86	signal transducer and activator of transcription 5B	NP_036580.2	Breast cancer (2)
TSG101	4.30E-98	tumor susceptibility gene 101	NP_006283.1	

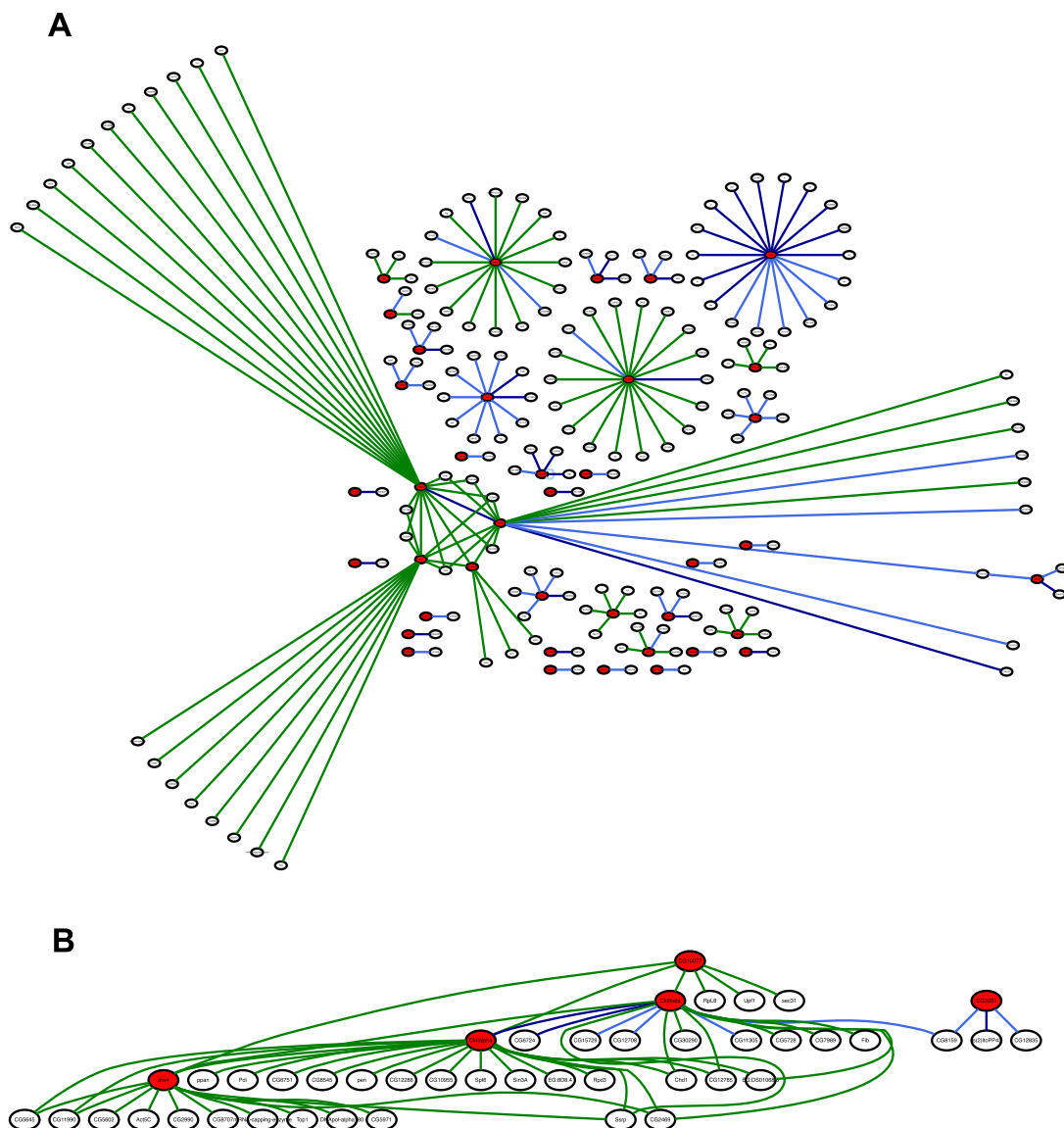
Shown are the relevant diseases of the human homologues with a BlastP E value of 10⁻¹⁰ or less.

References:

- (1) data sets were downloaded from Homophila Version 2.1 (update 13 Apr 2005). Originally published: Chien, S., Reiter, L. T., Bier, E. and Gribskov M. Homophila: human disease gene cognates in *Drosophila*. *Nucleic Acids Research* 30: 149-151 (2002)
- (2) data sets from: Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), 2000. World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>
- (3) information from: Kallia C., Newkitch H., Schlotter M., Mertens D., Wildenberger K., Doerner H., Stiggenbauer S. and Lichter P. Translocation (X:11)(q13;q23) in B-cell chronic lymphocytic leukemia disrupts two novel genes. *Genes Chromosomes Cancer* 42 (2): 128-143 (2005)



Supplementary Figure 1. Interactions between random gene sets (Osprey). Control datasets for interactions were obtained by randomly sampling 88 genes (FBgn) ten times from a list comprising 11,795 genes. These ten datasets were uploaded into Osprey V1.2.0 and only interactions within these nodes were searched in the Fly GRID database. Note that some FBgns were not recognized by Osprey and therefore some of the datasets above may contain up to two genes less. The mean of these ten data sets (= expected interactions) is 1.3 for all interactions and 1.0 for only non-self interactions. Graphical representation is as in Figure 19 of the main text.



Supplementary Figure 2. Interactions between RNAi screen candidates (FlyNet). (A) Interactions for RNAi screen candidates were downloaded from Flynet with a confidence threshold of 0.5 and plotted in the automated graph layout software Graphviz 1.13 (v16). (B) Blowup of (A) to show the various interactions linking RNAi screen candidates (CG10077, CkIibeta, CKIIalpha, dre4, CG3281) through interactions with proteins not identified in the RNAi screen. Red circles represent nodes identified in the RNAi screen, whereas white circles indicate interactors not identified in the RNAi screen. Green lines indicate physical interactions identified in yeast and consequently mapped to the fly ortholog. Light blue lines indicate physical interactions identified for fly proteins. Dark blue lines represent interactions identified under more than one condition, including genetic interactions in yeast mapped to the corresponding fly ortholog. Data was downloaded from <http://www.jhubiomed.org/perl/flynet.pl>

Supplementary Table 4. Candidate human homologs chosen for functional assays using siRNAs.

Gene name <i>Drosophila melanogaster</i>	Gene Name <i>Homo sapiens</i>	Accession Number	Dharmoon siRNA catalog number	Molecular function (1)	Localization (1)	Alternative Localizations (1)	Homolog identification by			
							BestBlast (2)	Inparanoid (2)	Homologene (2)	ScreenBlast (3)
Positive regulators										
Art2	HRMT1L4	NM_019854	M-009763-00	Methyltransferase activity	-	-	+	-	-	+
asf1	ASPIA	NM_014034	M-020222-00	chaperone activity	nucleus	-	-	-	+	+
bin3	BIN3	NM_018688	M-016113-00	receptor signaling complex scaffold activity	mitochondrion	-	-	-	+	+
CG11696	ZNFEN1A1	NM_006060	M-019092-00	Transcription factor activity	nucleus	-	-	-	+	+
CG12460	SFPQ	NM_005066	M-006455-01	RNA binding	nucleus	-	-	-	+	+
CG13473	TXN2	NM_012473	M-017448-00	Oxidoreductase activity	mitochondrion	-	-	-	+	+
CG15306	MAPRE1	NM_012225	M-015287-00	Regulation of cell cycle	cytoplasm	-	-	-	+	+
CG15434	NDUFA2	NM_002488	M-018899-00	Oxidoreductase activity	cytoplasm	-	-	-	+	+
CG15555	ACCSN	NM_017419	M-009278-00	Ion channel activity	membrane	-	-	-	+	+
CG16003	CCNL1	NM_020307	M-008682-00	RNA binding	nucleus	-	-	-	+	+
CG30069	POLR2A	NM_000937	M-011186-01	DNA-directed RNA polymerase	nucleus	-	-	-	+	+
CG30358	TXNL4	NM_006701	M-008361-00	Ribonucleoprotein	nucleus	-	-	-	+	+
CG31005	TPT1	NM_014317	M-008464-00	Transferase Enzyme	ER	-	-	-	+	+
CG31132	WDR9	NM_018963	M-010963-00	unknown	nucleus	-	-	-	+	+
CG31132	BRODL	NM_153252	M-016759-00	unknown	-	-	-	-	+	+
CG31132	PHIP	NM_017934	M-019291-00	Receptor binding	Plasma membrane	-	-	-	+	+
CG31358	NPHS2	NM_014625	M-021200-00	Receptor binding	cytoplasm	-	-	-	+	+
CG31694	IFRD1	NM_001550	M-019615-00	unknown	nucleus	-	-	-	+	+
CG31694	IFRD2	NM_006764	M-016430-00	unknown	nucleus	-	-	-	+	+
CG32573	PRKCA	NM_002737	M-003523-03	Protein Serine/Threonine kinase activity	nucleus	-	-	-	+	+
CG3819	ENDODGL1	NM_005107	M-008572-00	unknown	-	-	-	-	+	+
CG4349	FTH1	NM_002032	M-019634-01	Storage protein	cytoplasm	-	-	-	+	+
CG4781	KIAA0644	NM_037980	M-021245-00	Storage protein	cytoplasm	-	-	-	+	+
CG6422	YTHDF1	NM_017798	M-018095-00	unknown	-	-	-	-	+	+
CG6434	RBBP5	NM_005057	M-012008-00	Transcription regulator activity	nucleus	-	-	-	+	+
CG7635	STOM	NM_004099	M-016971-00	-	secretory granule	-	-	-	+	+
CG7635	STOML3	NM_145286	M-016122-00	-	Plasma membrane	-	-	-	+	+
CG9086	UBR1	NM_174916	M-010691-01	Ubiquitin-specific protease activity	Plasma membrane	-	-	-	+	+
CIBP	CTBP1	NM_001328	M-008609-01	Transcription regulator activity	nucleus	-	-	-	+	+
CIBP	CTBP2	NM_001329	M-008692-01	Transcription regulator activity	nucleus	-	-	-	+	+
eIF-4B	EIF4B	NM_001417	M-020179-00	Translation regulator activity	cytoplasm	-	-	-	+	+
HDCU1676	CHRNA7	NM_000746	M-003443-01	Voltage-gated ion channel activity	Plasma membrane	-	-	-	+	+
hop	JAK3	NM_000215	M-003147-01	Protein-tyrosine kinase activity	cytoplasm	-	-	-	+	+
hop	JAK2	NM_004972	M-003146-02	Protein-tyrosine kinase activity	cytoplasm	-	-	-	+	+
hop	JAK1	NM_002227	M-003145-01	Protein-tyrosine kinase activity	cytoplasm	-	-	-	+	+
kn	EBF	NM_024007	M-028500-01	Transcription factor activity	nucleus	-	-	-	+	+
kn	EBF3	NM_00105463	M-011848-00	Transcription factor activity	nucleus	-	-	-	+	+
l(1)G0084	PHF10	NM_018288	M-013349-00	Transcription regulator activity	-	-	-	-	+	+
larp	LARP	NM_015315	M-027187-00	unknown	nucleus	-	-	-	+	+
mask	MLL3	NM_021230	M-007039-01	Transcription regulator activity	nucleus	-	-	-	+	+
mask	ANKHD1	NM_017747	M-014405-00	DNA binding	nucleus	-	-	-	+	+
nomA	SFPQ	NM_005066	M-006455-01	RNA binding	nucleus	-	-	-	+	+
Rpl1	APEX1	NM_001641	M-010237-01	DNA repair protein	nucleus	-	-	-	+	+
sol	SOLH	NM_005632	M-006037-00	Transcription factor activity	cytoplasm	-	-	-	+	+
sol	CAPN3	NM_000070	M-005805-01	Cysteine-type peptidase activity	nucleus	-	-	-	+	+
Stat92E	STAT5B	NM_012448	M-010539-01	Transcription factor activity	cytoplasm	-	-	-	+	+
Stat92E	STAT1	NM_007315	M-003543-01	Transcription factor activity	cytoplasm	-	-	-	+	+
Stat92E	STAT3	NM_003150	M-003544-00	Transcription factor activity	nucleus	-	-	-	+	+
Stat92E	STAT5A	NM_003152	M-003169-02	Transcription factor activity	cytoplasm	-	-	-	+	+
Taf2	TAF2	NM_003184	M-009341-01	Transcription factor activity	nucleus	-	-	-	+	+
Negative regulators										
bon	TRIM33	NM_015906	M-005392-02	Transcription regulator activity	nucleus	-	-	-	+	+
Car1	RBBP4	NM_003610	M-012137-00	Transcription regulator activity	nucleus	-	-	-	+	+
CG10077	DID3	NM_004396	M-003774-00	RNA binding	nucleus	-	-	-	+	+
CG15706	FLJ20160	NM_017694	M-020665-00	-	Plasma membrane	-	-	-	+	+
CG16975	L3MBTL2	NM_031488	M-010678-00	Transcription regulator activity	nucleus	-	-	-	+	+
CG17492	MIB2	NM_008075	M-004259-02	Cytoskeletal protein binding	cytoplasm	-	-	-	+	+
CG18112	C14ORF133	NM_022067	M-016131-00	unknown	-	-	-	-	+	+
CG4907	PIGN	NM_012327	M-012463-00	Transferase Enzyme	ER	-	-	-	+	+
dros4	SUPT16H	NM_007192	M-009517-00	Transcription factor activity	nucleus	-	-	-	+	+
enok	MYST3	NM_006766	M-019849-00	Transcription regulator activity	nucleus	-	-	-	+	+
lig	UBAP2	NM_018449	M-013168-00	unknown	-	-	-	-	+	+
Nup154	NUP155	NM_004298	M-011967-00	Transporter activity	nucleus	-	-	-	+	+
par-1	AKT2	NM_001626	M-003801-01	Protein Serine/Threonine kinase activity	cytoplasm	-	-	-	+	+
par-1	MARK1	NM_018650	M-006824-00	Protein Serine/Threonine kinase activity	cytoplasm	-	-	-	+	+
Pp1alpha-96A	PPP1CA	NM_002708	M-008927-00	Protein Serine/Threonine phosphatase activity	cytoskeleton	-	-	-	+	+
Pp1alpha-96A	PPP1CC	NM_002710	M-006827-00	Protein Serine/Threonine phosphatase activity	nucleus	-	-	-	+	+
PP2A-B'	PPP2R5D	NM_006245	M-009799-01	Protein Serine/Threonine phosphatase activity	nucleus	-	-	-	+	+
Pp61F	PTPN1	NM_002827	M-003529-04	Protein Tyrosine phosphatase activity	ER	-	-	-	+	+
Pp61F	PTPN2	NM_002828	M-008969-00	Protein Tyrosine phosphatase activity	ER	-	-	-	+	+
Rab5	RAB5B	NM_002868	M-004010-01	GTPase activity	Plasma membrane	-	-	-	+	+
Rab5	RAB5A	NM_004162	M-004009-00	GTPase activity	Plasma membrane	-	-	-	+	+
Rab5	RAB5C	NM_004583	M-004011-01	GTPase activity	endosome	-	-	-	+	+
Socs35E	SOCSS	NM_014011	M-017374-00	receptor signaling complex scaffold activity	cytoplasm	-	-	-	+	+
TSG101	TSG101	NM_006292	M-003549-01	Ubiquitin-specific protease activity	cytoplasm	-	-	-	+	+

References:

- (1) data sets were downloaded from HRPD (05/09/06). Originally published: Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, Ibarrola N, Deshpande N, Shanker K, Shivashankar HN, Rashmi BP, Ramya MA, Zhao Z, Chandrika KN, Padma N, Harsha HC, Yatish AJ, Kavitha MP, Menezes M, Choudhury DR, Suresh S, Ghosh N, Saravana K, Chandran S, Krishna S, Joy M, Anand SK, Madavan V, Joseph A, Wong GW, Schiemann WP, Constantinescu SN, Huang L, Khosravi-Far R, Steen H, Tewari M, Ghaffari S, Blobel GA, Dang CV, Garcia JG, Pevsner J, Jensen ON, Roepstorff P, Deshpande KS, Chinmayan AM, Hamosh A, Chakravarti A, Pandey A. Development of human protein reference database as an initial platform for uniochase systems biology in humans. *Genome Research*. 13:2363-2371. (2003) World Wide Web URL: <http://www.hprd.org/>
- (2) data sets were downloaded from FLIGHT (update 01 January 2006). Originally published: Sims D, Bursteinas B, Gao Q, Zvelebil M, and Baum B. FLIGHT: database and tools for the integration and cross-correlation of large-scale RNAi phenotypic datasets. *Nucleic Acids Res*;34(Database issue):D479-83. (2006) World Wide Web URL: <http://flight.flier.org/>
- (3) ScreenBlast refers to the identification by automated blast searches described in Supplementary Table 3.

Supplementary Table 5. Results of HeLa cell growth curve analysis after siRNA treatments.

Well	Growth Plate1	Growth Plate 2	Growth Phenotype	Human gene	β -actin mRNA 1	Fold change of β -actin compared to controls 1	β -actin mRNA phenotype 1	β -actin mRNA 2	Fold change of β -actin compared to controls 2	β -actin mRNA phenotype 2	β -actin mRNA 3	Fold change of β -actin compared to controls 3	β -actin mRNA phenotype 3	Summary of growth phenotypes	JAK/STAT phenotype
A01				none	8827840	1.06		223200	0.93		5209720	0.84		0	
A02				none	5295965	0.64		223541	0.93		5702250	0.92		0	
A03		reduced	yes	CAPN3	3520090	0.42	yes	112930	0.47	yes	2069735	0.33	yes	4	+
A04				JAK3	2618270	0.31	yes	195563	0.81		2453680	0.40	yes	2	
A05				CHRNA7	3933665	0.47	yes	208163	0.86		5042270	0.81		1	+
A06		reduced	yes	POLR2A	438660	0.05	yes	31744	0.13	yes	141125	0.02	yes	4	+
A07				DKFZP667B0210	4350405	0.52	yes	225156	0.93		2743675	0.44	yes	2	+
A08		mildly reduced	yes	CTBP1	2087860	0.25	yes	113427	0.47	yes	545215	0.09	yes	4	
A09					4859955	0.58	yes	224359	0.93		5574005	0.90	yes	1	-
A10		strongly reduced	yes	EIF4B	2532305	0.30	yes	138032	0.57	yes	2025390	0.33	yes	4	
A11				IFRD1	3222655	0.39	yes	140244	0.58	yes	2740290	0.44	yes	3	
A12		strongly reduced	yes	AKT2	4581985	0.55	yes	128934	0.53	yes	2719280	0.44	yes	4	
B01				none	9686020	1.16		230840	0.96		5895455	0.95		0	
B02				none	8326435	1.00		292629	1.21		6125605	0.99		0	
B03				APEX1	3570855	0.43	yes	206886	0.86		2298825	0.37	yes	2	
B04				FTH1	4413515	0.53	yes	259706	1.08		3377250	0.54	yes	2	+
B05				JAK1	2565310	0.31	yes	271081	1.12		2509420	0.40	yes	2	-
B06				NDUFA2	1970775	0.24	yes	220621	0.91		1878655	0.30	yes	2	+
B07		reduced	yes	PPP1CA	2186705	0.26	yes	206662	1.08		2502965	0.40	yes	3	+
B08	mildly reduced		yes	PPP1CC	2851280	0.34	yes	241894	1.00		3723760	0.60	yes	2	+
B09				PRKCA	4786725	0.57	yes	307865	1.28		4020995	0.65	yes	1	+
B10	strongly reduced	mildly reduced	yes	PTPN1	2597835	0.31	yes	217393	0.90		2411280	0.39	yes	3	
B11	mildly reduced	reduced	yes	PTPN2	4033500	0.48	yes	181715	0.75		2736270	0.44	yes	3	
B12		reduced	yes	RAB5B	5568865	0.67		217343	0.90		3579315	0.58	yes	2	
C01				none	9169670	1.10		283692	1.18		6198605	1.00		0	
C02				none	7554430	0.91		256855	1.06		6207990	1.00		0	
C03				STAT5A	5930675	0.71		225185	0.93		3231525	0.52	yes	1	+
C04				TAF2	2921845	0.35	yes	181378	0.75		2248260	0.36	yes	2	
C05	mildly reduced		yes	STOM	3433900	0.41	yes	288480	1.20		3164390	0.51	yes	3	
C06				RAB5A	3238535	0.39	yes	156168	0.65		2243880	0.36	yes	2	+
C07				NUP155	1917675	0.23	yes	179649	0.74		1990125	0.32	yes	2	+
C08				DDX5	2307645	0.28	yes	201405	0.83		2328755	0.38	yes	2	+
C09				JAK2	4593230	0.55	yes	184062	0.76		4808870	0.77	yes	1	+
C10	strongly reduced		yes	RBBP5	3831515	0.46	yes	196195	0.81		2014265	0.32	yes	3	
C11				SFPQ	2035570	0.24	yes	122653	0.51	yes	808745	0.13	yes	3	
C12	mildly reduced	mildly reduced	yes	ENDOGL1	4390620	0.53	yes	94878	0.39	yes	3960910	0.64	yes	3	+
D01				none	11206700	1.35		155868	0.65		6323915	1.02		0	
D02				none	7655155	0.92		249042	1.03		6392760	1.03		0	
D03				RBBP4	8228815	0.99		213583	0.89		3756705	0.61	yes	0	+
D04				SOLH	2444805	0.29	yes	24730	0.10	yes	1077830	0.17	yes	2	
D05	mildly reduced		yes	ZNFN1A1	3628035	0.44	yes	272860	1.13		4428830	0.71	yes	2	
D06				PPP2R5D	3854305	0.46	yes	181284	0.75		2951410	0.48	yes	2	+
D07				TSG101	1785335	0.21	yes	187340	0.78		1840440	0.30	yes	2	+
D08				TXNL4	4939550	0.59	yes	210441	0.87		2689780	0.43	yes	2	
D09				IFRD2	5073755	0.61		262397	1.09		3951150	0.64	yes	0	+
D10				MYST3	4209035	0.51	yes	166207	0.69		2968805	0.48	yes	2	+
D11		reduced	yes	SUPT16H	4653585	0.56	yes	183277	0.76		2464440	0.40	yes	3	
D12		reduced	yes	STAT1	7007265	0.84		87247	0.36	yes	4189080	0.67	yes	2	-
E01				none	10475950	1.26		374556	1.55		7209975	1.16		0	
E02				none	8603765	1.03		283357	1.17		5912140	0.95		0	
E03				MAPRE1	7765585	0.93		198348	0.82		3129880	0.50	yes	1	
E04				PIGN	3714340	0.45	yes	155341	0.64		2178945	0.35	yes	2	
E05				STAT5B	5853805	0.70		190638	0.79		3269030	0.53	yes	1	+
E06				TXN2	2743815	0.33	yes	151248	0.63		1215585	0.20	yes	2	
E07				SOC35	3793695	0.46	yes	189378	0.78		2541580	0.41	yes	2	
E08		reduced	yes	ASF1A	4224090	0.51	yes	243607	1.01		3508900	0.57	yes	3	
E09				TPST1	8887090	1.07		278616	1.15		5352020	0.86	yes	0	-
E10				NPHS2	4422440	0.53	yes	280785	1.16		2508755	0.40	yes	2	
E11		reduced	yes	KIAA0644	4571890	0.55	yes	170644	0.71		2470120	0.40	yes	3	
E12				LARP	3114805	0.37	yes	121025	0.50	yes	1879205	0.30	yes	3	
F01				none	10646640	1.28		288103	1.19		7651040	1.23		0	
F02				none	7603055	0.91		234650	0.97		5580260	0.90		0	
F03	reduced	reduced	yes	TRIM33	5393665	0.65		182348	0.76		2326865	0.37	yes	2	+
F04				ACCN5	3889420	0.47	yes	166533	0.69		1295915	0.21	yes	2	
F05		strongly reduced	yes	FLJ20160	3687155	0.44	yes	214276	0.89		3004820	0.48	yes	3	
F06				ANKHD1	3654185	0.44	yes	191146	0.79		3391610	0.55	yes	2	+
F07				YTHDF1	3091235	0.37	yes	219841	0.91		2330985	0.38	yes	2	
F08	reduced	strongly reduced	yes	PHF	3468000	0.42	yes	217023	0.90		2663170	0.43	yes	3	
F09				PHF10	4915925	0.59	yes	229046	0.95		3428450	0.55	yes	2	-
F10				UBAP2	8534220	1.02		331412	1.37		9174895	1.48	yes	0	
F11	mildly reduced		yes	MARK1	3219480	0.39	yes	212654	0.88		2940145	0.47	yes	3	
F12	strongly reduced	reduced	yes	BIN3	5860125	0.70		114538	0.47	yes	4503250	0.73	yes	2	
G01				none	10204000	1.23		357452	1.48		4692485	0.76	yes	0	
G02				none	6834975	0.82		343503	1.42		4119885	0.66	yes	0	
G03				C21ORF107	6357090	0.76		221939	0.92		3776215	0.61	yes	0	-
G04				HRMT1L4	3135535	0.38	yes	136208	0.56	yes	1247580	0.20	yes	3	+
G05				CCNL1	2627355	0.32	yes	180189	0.75		1899215	0.31	yes	2	-
G06				MLL3	4117445	0.49	yes	257373	1.07		4282880	0.69	yes	1	-
G07				C14ORF133	3915900	0.47	yes	232528	0.96		2959015	0.48	yes	2	+
G08		reduced	yes	EBF	3069895	0.37	yes	252031	1.04		1865770	0.30	yes	3	
G09				L3MBTL2	4404745	0.53	yes	158230	0.66		4063960	0.65	yes	1	
G10	reduced		yes	LOC142678	3966495	0.48	yes	84983	0.35	yes	2471345	0.40	yes	4	+
G11	reduced	strongly reduced	yes	STAT3	2092555	0.25	yes	104096	0.44	yes	2086740	0.34	yes	4	
G12	strongly reduced		yes	STOML3	6748235	0.81		102113	0.42	yes	3493035	0.56	yes	3	
H01				none	10406590	1.25		241265	1.00		4528860	0.73	yes	0	
H02				none	8010795	0.96		191303	0.79		4797770	0.77	yes	0	
H03				BRODL	6338075	0.76		183443	0.76		3980880	0.64	yes	1	-
H04		mildly reduced	yes	UBR1	3205015	0.38	yes	98960	0.41	yes	2134660	0.34	yes	3	
H05				RAB5C	5423535	0.65		231270	0.96		5844185	0.94	yes	0	
H06				none	7335465	0.88		189082	0.78		8972020	1.45	yes	0	
H07				none	7439250	0.89		250661	1.04		8593460	1.38	yes	0	
H08				none	8604355	1.03		203989	0.85		7190390	1.16	yes	0	
H09				none	7132960	0.86		244869	1.01		7228035	1.16	yes	0	
H10				none	6825915	0.82		181028	0.75		6260630	1.01	yes	0	
H11				none	6978920	0.84		192113	0.80		6978180	1.12	yes	0	
H12				none	9975040	1.20		98825	0.41		7265430	1.17	yes	0	

Growth curves were determined over a course of 3 d after siRNA transfection by counting particles in recorded images using ImageJ.

β -actin mRNA levels were determined using a branched DNA assay during the initial siRNA transfection screen with a stimulation of IFN- γ for the first two experiments or OSM for the last experiment, respectively. A phenotype for β -actin levels was determined if the fold change of β -actin levels was <0.6 compared to controls.

The column 'Summary of growth phenotypes' represents the cumulative score in all comparisons (e.g. "1" indicates a phenotype in one experiment, "4" indicates phenotypes in all 4 conditions). The column 'JAK/STAT phenotype' is the phenotype after siRNA transfection for the IFN γ experiment. "+" represents an increase in target gene activity, whereas a "-"

Supplementary Table 6. Identity of siRNAs (Dharmacon) used in pools and individually for retests.

Gene Symbol	Catalog Nr.	Locus ID	Accession Nr.	Sequence
ANKHD1	D-014405-01	54882	NM_017747	GUAUAUUGCUAGAUGAAGG
ANKHD1	D-014405-02	54882	NM_017747	UGGCAGCUCUACUUAUUGA
ANKHD1	D-014405-03	54882	NM_017747	GCGCUAUGUGCAUGCUUC
ANKHD1	D-014405-04	54882	NM_017747	ACACUGCGCUAACUUAUGC
BRODL	D-016759-01	254065	NM_153252	GACCUAGAGCUAAUUAAG
BRODL	D-016759-02	254065	NM_153252	GGUGGUUAUGUCAUACAA
BRODL	D-016759-03	254065	NM_153252	GCAAGAAUUGGCAGUUAUC
BRODL	D-016759-04	254065	NM_153252	GAUCCGAAAGCCGAGCUUAU
C14ORF133	D-016131-01	63894	NM_022067	GAACCGCAGUACUUAUUG
C14ORF133	D-016131-02	63894	NM_022067	CAGAAGAGCUUGCGCUAUC
C14ORF133	D-016131-03	63894	NM_022067	UAAGGUAGAUAAAGGAUCA
C14ORF133	D-016131-04	63894	NM_022067	GCUCCAAGUUAAGGCUUU
C21ORF107	D-010963-01	54014	NM_018963	GAAAUUGAGUGAUUGUGAA
C21ORF107	D-010963-02	54014	NM_018963	CGAAAGAGAGUUAUUAUA
C21ORF107	D-010963-03	54014	NM_018963	GAGCUGUUGUUAUGUCUUA
C21ORF107	D-010963-04	54014	NM_018963	GACAGAUUCCGCUUAUUA
CAPN3	D-005805-01	825	NM_000070	CAACAUCUCCGGAACAUAU
CAPN3	D-005805-03	825	NM_000070	CAGUAGAGCUACAAGAAGU
CAPN3	D-005805-04	825	NM_000070	UCAACAACCAGCUCUUAUGA
CAPN3	D-005805-05	825	NM_000070	UGAUUCCGCUACUAGAUAC
CHRNA7	D-004143-01	1139	NM_000746	ACAAGGAGCUGUCAAGAA
CHRNA7	D-004143-02	1139	NM_000746	GAUGAGAAGAACCAAGUUA
CHRNA7	D-004143-03	1139	NM_000746	GAGAAGAACAAGUUUUA
CHRNA7	D-004143-05	1139	NM_000746	GGGUAAGACUUGUCUUUA
CTBP2	D-008962-03	1488	NM_001329	UCGCAUCCAGAAAGCUIUA
CTBP2	D-008962-04	1488	NM_001329	CAAUUGGUGCCACUACAGA
CTBP2	D-008962-05	1488	NM_001329	GCUCUUAUGGCGCCAUUAC
CTBP2	D-008962-06	1488	NM_001329	UGAGAGUACUGUGCGGAU
DDX5	D-003774-01	1655	NM_004396	GCAAGUAGCUGCUUAUUA
DDX5	D-003774-02	1655	NM_004396	CACAAGAGGUGGAACAUAU
DDX5	D-003774-03	1655	NM_004396	GUGUAGGCUUACCAGAAA
DDX5	D-003774-04	1655	NM_004396	GCAAAUGUUAUGGAGUGUA
DKFZP667B0210	D-026500-01	253738	NM_001005463	GCACUAGCAUAUGGUAUA
DKFZP667B0210	D-026500-02	253738	NM_001005463	AACCAUAGAUUCCGCUUA
DKFZP667B0210	D-026500-03	253738	NM_001005463	CUACACCGCCUUAUAUGAA
DKFZP667B0210	D-026500-04	253738	NM_001005463	GGGUGAUCCCGAAAGGUUA
ENDOGL1	D-008572-01	9941	NM_005107	UAGAAGAACAUAAGCUIUA
ENDOGL1	D-008572-02	9941	NM_005107	GGACGAGAAUAGAAUUGUA
ENDOGL1	D-008572-03	9941	NM_005107	CGUACUUGUUAUUAUGUA
ENDOGL1	D-008572-04	9941	NM_005107	GACCUUACCCAGACUAGA
FTH1	D-019634-01	2495	NM_002032	GGUGAAAGCAUACAAGAA
FTH1	D-019634-02	2495	NM_002032	ACUCAGAGGCGCCAUCAA
FTH1	D-019634-03	2495	NM_002032	CCAUACAAGAAUUGGUGUA
FTH1	D-019634-04	2495	NM_002032	GAGGAGGCGGCUUAUGCAA
HRMT1L4	D-009763-01	56341	NM_019854	GCUCUAGUUAUUAUGUA
HRMT1L4	D-009763-02	56341	NM_019854	GACCAUUGCCUUAUUGUA
HRMT1L4	D-009763-03	56341	NM_019854	GAAGAGCUAUCGUUACUAU
HRMT1L4	D-009763-04	56341	NM_019854	CAACAUCUACCAUUAUUA
IFRD2	D-016430-01	7866	NM_006764	UGAAUGCGAAGAAGAGUAU
IFRD2	D-016430-02	7866	NM_006764	UCAAGGAAGUGUCUGGUAU
IFRD2	D-016430-03	7866	NM_006764	AGAAGGACCCUGUAACAUG
IFRD2	D-016430-04	7866	NM_006764	UCGCAAGGCUUCCAGCUIUA
JAK1	D-003145-05	3716	NM_002227	CCCAUAGGCUUUAUGGUAU
JAK1	D-003145-06	3716	NM_002227	UGAAUACACUACAUAUUA
JAK1	D-003145-07	3716	NM_002227	UAAGGAACCCUUAUUAUUA
JAK1	D-003145-08	3716	NM_002227	GCAGGUGGCUUUAUUAUUA
LOC142678	D-015287-01	142678	NM_080875	GCAAGAAACUCCGCCAGAA
LOC142678	D-015287-02	142678	NM_080875	CAACCCGAGGCGCCAAACAU
LOC142678	D-015287-03	142678	NM_080875	CCUCAAGGCUACCGGCUUA
LOC142678	D-015287-04	142678	NM_080875	UGGCGGAGUUUAUGGCAU
MLL3	D-007039-01	58508	NM_021230	GAGAAGAGCUCUGUAUUA
MLL3	D-007039-02	58508	NM_021230	GAACAAGGAUCCCGGAAU
MLL3	D-007039-03	58508	NM_021230	GCAUUGGUCUUAUGGUAU
MLL3	D-007039-05	58508	NM_021230	CAAGGUCUACAACAGUUA
MYST3	D-019849-01	7994	NM_006766	GAACAGCUAUGAUAUGACA
MYST3	D-019849-02	7994	NM_006766	GGAGUUGAGUUAUUAUUA
MYST3	D-019849-03	7994	NM_006766	GCGCUUAUUAUUAUUAUUA
MYST3	D-019849-04	7994	NM_006766	AGAUGUAGAACCAGAAUUA
NDUFA2	D-018869-01	4695	NM_002488	GAAGGCGAAUCCCGACCUA
NDUFA2	D-018869-02	4695	NM_002488	GAGAUUCCUCCACUUAUUA
NDUFA2	D-018869-03	4695	NM_002488	UCUUAUGUUAUUAUUAUUA
NDUFA2	D-018869-04	4695	NM_002488	CCGCUACGCUUUAUGGCAA
NUP155	D-011967-01	9631	NM_004298	GGACUCAGCUAUCUUAUUA
NUP155	D-011967-02	9631	NM_004298	CAACUCAGGCGCAAAUUAU
NUP155	D-011967-03	9631	NM_004298	ACAUAGAGCUCUUAUUAUUA
NUP155	D-011967-04	9631	NM_004298	CAUUAUGGUAUUAUUAUUA
PHF10	D-013349-01	55274	NM_018288	GAUAUUGCUGAAGGUGUA
PHF10	D-013349-02	55274	NM_018288	UCAGUGCACUUAUGGCUUA
PHF10	D-013349-03	55274	NM_018288	GCGCAGUUAUGAAGUGUAU
PHF10	D-013349-04	55274	NM_018288	GAACGCUAACGAUUAUUAU
PPP1CA	D-008927-01	5499	NM_002708	CAAGAUUCGCGUUAUUAU
PPP1CA	D-008927-02	5499	NM_002708	CAAGAGACCCUACAACUUA
PPP1CA	D-008927-03	5499	NM_002708	GAACGACCCUUGGCGUUAU
PPP1CA	D-008927-04	5499	NM_002708	CCAGUUAUUAUUAUUAUUA
PPP1CC	D-006827-01	5501	NM_002710	GGAGUGUUAUUAUUAUUAU
PPP1CC	D-006827-02	5501	NM_002710	GUAAAGUUAUUAUUAUUAU
PPP1CC	D-006827-03	5501	NM_002710	CUUAAGUUAUUAUUAUUAU
PPP1CC	D-006827-04	5501	NM_002710	GCGGAGAGUUAUUAUUAUUA
PPP2R5D	D-009799-01	5528	NM_006245	GAGCGGACUUAUUAUUAUUA
PPP2R5D	D-009799-03	5528	NM_006245	GUACAUCGACGAAUUAUUA
PPP2R5D	D-009799-04	5528	NM_006245	UCCAUGGAGUUAUUAUUAU
PPP2R5D	D-009799-05	5528	NM_006245	GCUUAUUAUUAUUAUUAUUA
RAB5A	D-004009-01	5868	NM_004162	GCAAGCAAGCUUAUUAUUAU
RAB5A	D-004009-02	5868	NM_004162	GGAGAGGAGUUAUUAUUAU
RAB5A	D-004009-03	5868	NM_004162	AGGAUAACGCUUAUUAUUAU
RAB5A	D-004009-04	5868	NM_004162	GAAGAGGAGUUAUUAUUAU
RBBP4	D-012137-01	5928	NM_005610	GAUACUGUUAUUAUUAUUAU
RBBP4	D-012137-02	5928	NM_005610	GAUCGCUUAUUAUUAUUAU
RBBP4	D-012137-03	5928	NM_005610	GGUAUCUGUUAUUAUUAUUA
RBBP4	D-012137-04	5928	NM_005610	AACAUAUUAUUAUUAUUAU
STAT1	D-003543-01	6772	NM_007315	AGAAAGAGCUUAUUAUUAU
STAT1	D-003543-03	6772	NM_007315	UAAGGAAACUGGUAUUAUUA
STAT1	D-003543-04	6772	NM_007315	GAGCUUAUUAUUAUUAUUAU
STAT1	D-003543-05	6772	NM_007315	GAACCGACUUAUUAUUAUUA
STAT3	D-003544-01	6774	NM_003150	CCAAGCAGCAGCAUUAUUAU
STAT3	D-003544-02	6774	NM_003150	GGAGAAGCAUUAUUAUUAUUA
STAT3	D-003544-03	6774	NM_003150	CCACUUAUUAUUAUUAUUAU
STAT3	D-003544-04	6774	NM_003150	NM_003150
TPRT	D-008464-01	23590	NM_014317	AGACAGAGCUGCAAGUUAU
TPRT	D-008464-02	23590	NM_014317	GAAGACCCGAGCUGUAUUA
TPRT	D-008464-03	23590	NM_014317	GCAUAUUAUUAUUAUUAUUA
TPRT	D-008464-04	23590	NM_014317	UUAAGGAAUUAUUAUUAUUA
TRIM33	D-005392-01	51592	NM_015906	GGACAACCCAUUAUUAUUAU
TRIM33	D-005392-02	51592	NM_015906	CAACGCGAGUUAUUAUUAU
TRIM33	D-005392-06	51592	NM_015906	UGAAACAUGUUAUUAUUAU
TRIM33	D-005392-07	51592	NM_015906	GUGUAUUAUUAUUAUUAUUA
TSG101	D-003549-01	7251	NM_006292	AAACUGGAGUUAUUAUUAU
TSG101	D-003549-02	7251	NM_006292	GAACCUACUGGAAACAUAU
TSG101	D-003549-04	7251	NM_006292	CCGUUAUUAUUAUUAUUAU
TSG101	D-003549-05	7251	NM_006292	UCCACAGCUCUUAUUAUUA

Supplementary Table 7. Common proteins between immunoprecipitated complexes identified by LC-MS/MS.

GI number	Protein name	identified in
gil51092169	LD24035p	A + B
gil41617226	TPA: TPA_inf: HDC06756	A + B
gil415634	cloned by ability to arrest the cell cycle when expressed in the fission yeast Schizosaccharomyces	A + B
gil3915856	40S ribosomal protein S3a (C3 protein)	A + B
gil24653084	CG13151-PA	A + B
gil21429784	AT24407p	A + B
gil20976820	GH04183p	A + B
gil17946379	RE65203p	A + B
gil17737731	Ribosomal protein LP0 CG7490-PA	A + B
gil17647925	Septin-1 CG1403-PA, isoform A	A + B
gil17136832	Ribosomal protein S13 CG13389-PA, isoform A	A + B
gil17136734	string of pearls CG5920-PA	A + B
gil6911895	ribosomal protein S17	A + C
gil40882499	RE57333p	A + C
gil24655737	beta-Tubulin at 56D CG9277-PB, isoform B	A + C
gil18860087	eIF-2 CG9946-PA	A + C
gil17945758	RE33114p	A + C
gil17647883	Ribosomal protein L23 CG3661-PA	A + C
gil8127	hsp 82	A + D
gil8040	unnamed protein product	A + D
gil7739653	rasputin	A + D
gil558485	ribosomal protein DL11	A + D
gil48958421	RH21963p	A + D
gil24642576	Ribosomal protein S19a CG4464-PC, isoform C	A + D
gil21357739	Glycoprotein 93 CG5520-PA	A + D
gil19922662	CG10540-PA	A + D
gil17137626	Inos CG11143-PA	A + D
gil1653979	60kDa heat shock protein	A + D
gil158489	alpha-spectrin	A + D
gil157667	heat shock protein cognate 71	A + D
gil156759	actin	A + D
gil1141790	nonmuscle myosin-II heavy chain	A + D
gil396531	ribosomal protein S19	B + C
gil4378006	ribosomal protein L23a	B + D
gil296094	ribosomal protein S3	B + D
gil21355973	CG7185-PA	B + D
gil6016262	Heat shock protein 83 (HSP 82)	C + D
gil158483	small optic lobes protein	C + D
gil6249321	EG:BACR19J1.4	A + B + C
gil46409240	AT25469p	A + B + C
gil39752635	LP10071p	A + B + C
gil21392200	RH09938p	A + B + C
gil156750	actin	A + B + C
gil17738071	Septin-2 CG4173-PA	A + B + D
gil508229	Peanut	A + B + D
gil495594	poly(A)-binding protein	A + C + D
gil24641739	thioredoxin peroxidase 1 CG1633-PB, isoform B	A + C + D
gil24581558	CG31957-PA	A + C + D
gil6094051	60S ribosomal protein L31	B + C + D
gil77403899	Heat shock protein cognate 4	A + B + C + D
gil7919	Elongation factor 2b	A + B + C + D
gil733532	unknown	A + B + C + D
gil2274922	14-3-3zeta	A + B + C + D
gil17136226	Ecdysone-inducible gene L3 CG10160-PA	A + B + C + D
gil157658	heat shock protein cognate 72	A + B + C + D

Proteins identified were selected if from *Drosophila melanogaster* and if the protein score was higher than 0.

A indicates the mock immunoprecipitation (IP) using α -Flag antibody

B indicates the IP from BRWD3-Flag transfected cells using α -Flag antibody

C indicates the IP from Stat92E-Myc transfected cells using α -Myc antibody

D indicates the mock IP using α -Myc antibody

Curriculum Vitae

Personal Data

Full name Patrick Müller
Date of birth April 27, 1979
Place of birth Osnabrück, Germany

Education

since 2004 Max Planck Institute for Biophysical Chemistry Göttingen, Germany
PhD Thesis in Dr. Martin P. Zeidler's / Prof. Dr. Herbert Jäckle's laboratories

2003 – 2004 Rockefeller University New York, USA
MSc Thesis in Prof. Dr. James E. Darnell's laboratory

2002 – 2003 Georg-August University Göttingen, Germany
International MSc/PhD Program Molecular Biology

2001 – 2002 University of California at Berkeley, USA
General biology

1999 – 2001 Georg-August University Göttingen, Germany
Biology

Awards

since 2006 stipend of the Max Planck Society

2005 winner of esiWay contest with best research proposal (Roche and RZPD)
companies supplied free biochemicals for RNAi experiments

2005 DAAD fellowship for RISE program
organized scholarship for a summer student (University of Illinois, USA)

2005 travel award for conference participation (International Max Planck Research School)

since 2004 member / mentee of the Mentoring Program at Roche Diagnostics

since 2003 fellow of the German National Academic Foundation (Studienstiftung des Deutschen Volkes) for undergraduate and graduate studies

since 2003 fellowship from e-fellows.net

2003 Irmgard-Ulderup-Scholarship for a language course in Rome, Italy

2002-2003 KWS/PLANTA stipend
undergraduate stipend from a plant biotechnology company

- 2002-2003** stipend from the International Max Planck Research School
2002 Money for Musicians Scholarship (ASUC Senate, Berkeley)
2001-2002 Education Abroad Program (University of California)

Publications

Dietz, G. P. H., P. C. Valbuena, B. Dietz, K. Meuer, **P. Müller**, J. H. Weishaupt, and M. Bähr. 2006. Application of a blood-brain barrier-penetrating form of GDNF in a mouse model for Parkinson's Disease. *Brain Research* 1082: 61-6.

Müller, P., D. Kutteneuler, V. Gesellchen, M. P. Zeidler, and M. Boutros. 2005. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature* 436: 871-5.

Patent

P. Müller, M. P. Zeidler, and M. Boutros. Identification of JAK/STAT pathway modulating genes by genome wide RNAi screening. *European Patent Application Nr. 05 012 934.5*.

Selected Presentations

Pelte, N., **P. Müller**, M. P. Zeidler, and M. Boutros. 2007. Chifoumi is a novel negative regulator of JAK/STAT signalling. *48th Annual Drosophila Research Conference* (Philadelphia, USA).

Müller, P., and M. Boutros. 2006. CellScreen: A statistics package for the analysis of cell-based RNAi screens. *47th Annual Drosophila Research Conference* (Houston, USA).

Müller, P., M. Boutros, and M. P. Zeidler. 2005. dBRWD3 is a novel regulator of the JAK/STAT signal transduction pathway. *19th European Drosophila Research Conference* (Eger, Hungary, oral presentation).

Müller, P., and M. P. Zeidler. 2005. Identification and characterization of a phosphatase negatively regulating the JAK/STAT signal transduction pathway in *Drosophila*. *46th Annual Drosophila Research Conference* (San Diego, USA, oral presentation).

Sudame, A., S. Lee, H. Lee, **P. Mueller**, K. Hida, H. Ng, T. Lundquist, P. Strom, and T. Leighton. 2002. Selenite reducing bacteria of Panoche Algal Bacterial Selenium Removal (ABSR) Facility, CA. *34th Mid-Atlantic Industrial & Hazardous Waste Conf* (New Jersey, USA).