Université de Montréal

Identification et étude de mécanismes régulant l'expression de MAPK

par
Dariel Ashton-Beaucage

Institut de recherche en immunologie et cancérologie
Faculté de Médecine

Thèse présentée à la Faculté de Médecine
en vue de l'obtention du grade de Philosophiae Doctor
en biologie moléculaire

décembre 2014

Résumé

Les modèles classiques de signalisation cellulaire eucaryotes sont généralement organisés en voies linéaires et hiérarchiques, impliquant un ensemble de facteurs restreint. Ces facteurs forment un circuit isolé qui transmet une information externe vers sa destination, d'où une réponse cellulaire sera alors engendrée. Or, ces modèles sont justement le fruit d'approches expérimentales réductionnistes qui ne permettent pas d'intégrer aisément la contribution de facteurs multiples, ni de faire une évaluation quantitative de l'apport des composantes du système. Le développement de techniques d'investigation plus holistiques, telles la génomique fonctionnelle et la protéomique, permettent d'examiner de manière systématique et quantitative l'apport d'ensembles larges de facteurs et de les mettre en relation avec d'autres systèmes cellulaires. Il y aurait donc lieu de réévaluer le modèle de voie de signalisation linéaire au profit d'un modèle de réseau de signalisation multiparamétrique, comportant plusieurs branches d'entrée et sortie de signal interagissant avec d'autres systèmes cellulaires.

Cet ouvrage porte sur la voie RAS/MAPK, l'un des principaux axes de signalisation associé à la prolifération et la différenciation cellulaires. Le sujet y est d'abord abordé sous l'angle d'une perspective historique, en mettant l'emphase sur les contributions des études de génétique classique chez les organismes modèles *D. melanogaster* et *C. elegans*. Il fait ensuite état du développement du criblage par ARNi pan-génomique dans ces deux modèles en le comparant aux approches de criblage génétique classique. Le corps de l'ouvrage décrit ensuite les résultats expérimentaux d'une campagne de criblage par ARNi visant à dresser une carte globale des régulateurs de la voie chez la drosophile.

Trois groupes de régulateurs identifiés dans ce crible ont été caractérisés de manière plus détaillée. Dans un premier article, nous démontrons que les composantes du complexe EJC ont un impact sur l'épissage de *mapk*; une découverte doublement intéressante puisque l'EJC était jusqu'alors associé qu'à la régulation post-épissage des ARNm. Une seconde publication fait état de l'ensemble des résultats du crible ARNi, mettant l'emphase sur un ensemble de facteurs d'épissage qui modulent également *mapk*. Nous y montrons que l'impact de ces facteurs sur l'épissage alternatif est différent de celui de l'EJC, suggérant ainsi deux

modes de régulation distincts. Finalement, dans un troisième manuscrit, nous nous attardons au rôle d'*Usp47*, une déubiquitinase qui, contrairement aux autres facteurs identifiés dans le crible, régule l'expression de MAPK de manière post-traductionnelle. Nous y détaillons une stratégie de criblage d'interaction génétique par ARNi visant à identifier des facteurs reliés fonctionnellement à *Usp47*. Ce second crible a permis l'identification de trois facteurs reliés au « N-end rule », un mécanisme de dégradation des protéines caractérisé par la reconnaissance des résidus N-terminaux de protéines ou peptides.

Il existait jusqu'alors très peu de données quant à la régulation de l'expression des composantes de la voie MAPK, ce qui rend la description d'un large réseau de régulateurs agissant sur l'expression de MAPK d'autant plus insoupçonnée. L'absence d'un réseau équivalent rattaché aux autres composantes de la voie laisse supposer que MAPK serait un nœud servant de point d'entrée à ce type de régulation dans le système RAS/MAPK. De plus, nos travaux témoignent de la capacité de la génomique fonctionnelle à mettre en relation différents systèmes cellulaires de manière plus globale et à quantifier les liens établis entre eux.

Mots-clés : signalisation RAS/MAPK, *Drosophila melanogaster*, criblage par ARNi, complexe EJC (« exon junction complex »), épissage alternatif, spliceosome, définition d'exon/intron, système ubiquitine-protéasome, « N-end rule »

Abstract

The classical model of eukaryotic cellular signalling generally involves hierarchically organized linear pathways involving a restricted set of elements. These generally function together as an insulated circuit, transmitting information from the outside to the intracellular compartment involved in eliciting a response. These models, often the fruit of reductionist experimental approaches, do not allow for the integration of multiple inputs nor for a gradation of responses. The recent emergence of more holistic investigation techniques has brought about the re-evaluation of these classical models in favor of multiparametric signalling networks.

This thesis focuses on the RAS/MAPK pathway, one of the cell's main proliferation and differentiation signalling conduits, beginning with a historical perspective covering the contributions of model organism genetics to the current pathway model. This provides context for the description of a whole-genome RNAi screen experiment that we carried out to obtain a global view of regulators in *Drosophila*. Three groups of factors emerging from this screen were then examined in more detail.

A first article shows that the exon junction complex (EJC) plays a role in *mapk* alternative splicing, an observation that is unexpected given that this complex was not previously known to act on splicing. A second paper details the genome wide screening campaign and focuses on a large set of splicing factors that also regulate *mapk*, albeit in a distinct manner than the EJC's. Finally, a manuscript in a third segment examines *Usp47* function and finds it to control MAPK levels post-translationally. An RNAi-based genetic interaction screen is then used to identify factors functionally related to *Usp47* capable of counteracting its impact on MAPK levels. Three such factors identified through this technique are linked to the N-end rule protein degradation pathway. Regulation of core pathway component expression is a poorly described process, which makes the identification of a large set of factors regulating MAPK expression all the more unusual. Moreover, the absence of such regulation linked to other pathway components suggests that MAPK may act as a node incorporating inputs of this type into RAS/MAPK signaling dynamics.

Keywords: RAS/MAPK signalling, *Drosophila melanogaster*, RNAi screening, exon junction complex, alternative splicing, spliceosome, exon/intron definition, ubiquitin-proteasome system, N-end rule

Table des matières

Liste des tableaux

Liste des figures

Liste des sigles

Note sur la nomenclature des sigles de gènes :

Les sigles et les noms utilisés pour désigner les gènes, transcrits et protéines de *D. melanogaster* dans cet ouvrage respectent les règles de nomenclature établies par *Flybase* (flybase.org). Les noms et les sigles de gènes (« gene symbols ») apparaissent en italique alors que les noms et les sigles désignant les protéines sont en majuscules non-italique. Lorsque le sigle désigne plutôt le gène de manière générale ou dans une autre espèce, un sigle générique non-italique sera employé. Une liste des principaux facteurs et régulateurs de la voie RAS/MAPK et des principaux sigles employés chez *D. melanogaster*, *C. elegans* et *H. sapiens* figure au tableau 1.1.

ADN : acide désoxyribonucléique
ARN : acide ribonucléique
ARNdb : long ARN à double brin (voir aussi dsRNA)
ARNi : interférence ARN
ARNm : ARN messager
ARNt : ARN de transfert
btz : *barentsz*
Cdc37 : *Cell Division Cycle 37* S. cerevisiae *homolog*
Cdk12 : *Cyclin-dependent kinase 12*
CLIP : « cross-linking immunoprecipitation »
CRISPR : « clustered regularly interspaced short palindromic repeats » (série de répétitions courtes palindromiques espacées régulièrement)
dsRNA : « long double stranded RNA » (longs ARN à double brin; voir ARNdb)
Dsor1 : *Downstream of raf1*
eIF4AIII : *eukaryotic translation initiation factor 4AIII*
EJC : « exon junction complex » (complexe de jonction d'exon)

HCS : « high content screening » (crible à haut contenu)

hnRNP : « heterogeneous nuclear ribonucleoproteins »

HTS : « high throughput screening » (crible à haut débit ou crible à large échelle)

mago : *mago-nashi*

MAPK : mitogen-activated protein kinase

MEK : mitogen-activated protein kinase kinase

mRNP : « messenger ribonucleoprotein » (ribonucléoprotéine messagère)

NMD : « non-sense mediated decay » (dégradation des ARNm non-sens)

phl : *pole hole*

RAF : rapidly accelerated fibrosarcoma

Ras : rat sarcoma

Ras85D : *Ras oncogene at 85D*

RISC : « RNA-induced silencing complex » (complexe de silençage induit par l'ARN)

rl : *rolled*

RnpS1 : *RNA-binding protein S1*

RTK : récepteur tyrosine-kinase

S2 : Schneider 2

shRNA : « small hairpin RNA» (petit ARN en épingle à cheveux)

siRNA : « small interfering RNA» (petit ARN interférent)

SR : « serine/arginine rich » (riche en serine/arginine)

tsu : *tsunagi*

Liste des abréviations

C. elegans : *Caenorhabditis elegans*

D. melanogaster : *Drosophila melanogaster*

E. coli : *Escherichia coli*

H. sapiens : *Homo sapiens*

kb. : kilobase

nt. : nucléotide

S. cerevisiae : *Saccharomyces cerevisiae*

S. pombe : *Schizosaccharomyces pombe*

To my father who, early on, kindled the fire and nurtured the inquisitive mind.

Remerciements

J'aimerais d'abord exprimer ma profonde gratitude envers les collègues qui ont participé aux travaux décrits dans ces pages. Leurs précieuses contributions ont permis de mener nos projets de recherche – et avec eux mon parcours doctoral – à bon port. En particulier, je tiens à souligner les contribution de : Chris Udell, Hugo Lavoie, Caroline Lemieux et Caroline Baril pour leur apport expérimental significatif; Martin Lefrançois, pour son soutien sans égal tout au long de mon cheminement gradué; Malha Sahmi, pour m'avoir guidé lors de mes débuts en laboratoire et pour son aide au fil des projets; Sébastien Lemieux, pour ses conseils judicieux en statistiques et analyse de données; Jean Duchaine et les membres de la plateforme de criblage de l'IRIC pour leurs conseils et les nombreuses fois où j'ai sollicité leur aide; et Christian Charbonneau, pour son soutien technique en imagerie. Je suis aussi extrêmement reconnaissant envers mon directeur de thèse, Marc Therrien, qui aura grandement contribué à façonner mon identité de chercheur et qui, par sa curiosité passionnée, sa disponibilité et son ardeur au travail aura été d'un soutien inestimable tout au long de mon parcours.

Durant mon doctorat, j'ai eu le privilège exceptionnel de travailler à l'IRIC, où j'ai non-seulement pu bénéficier d'expertises et de ressources techniques hors pair, mais également d'un environnement scientifique fertile et propice à la créativité. Je souligne également le rôle important des organismes qui ont contribué financièrement à la réalisation de mes travaux de recherche, soit les Instituts de recherche en santé du Canada, la Fondation Cole, les programmes de biologie moléculaire et la Faculté des études supérieures. En dernier lieu, je remercie mes proches pour leur précieux soutien moral au long de mon cheminement et particulièrement ma femme, Jessica, qui a également aidé à la relecture du texte français de ce document.

[…] *il ne vivait guère que la nuit, pensant qu'on était mieux chez soi, plus seul, et que l'esprit ne s'excitait et ne crépitait réellement qu'au contact voisin de l'ombre ; il trouvait aussi une jouissance particulière à se tenir dans une chambre largement éclairée, seule éveillée et debout, au milieu des maisons enténébrées et endormies, une sorte de jouissance où il entrait peut-être une pointe de vanité, une satisfaction toute singulière, que connaissent les travailleurs attardés alors que, soulevant les rideaux des fenêtres, ils s'aperçoivent autour d'eux que tout est éteint, que tout est muet, que tout est mort.*

- *À rebours*, J.K. Huysmans

Yet his mind was preternaturally active. His thoughts went ever around and around in a circle. The centre of that circle was "work performed"; it ate at his brain like a deathless maggot. He awoke to it in the morning. It tormented his dreams at night. Every affair of life around him that penetrated through his senses immediately related itself to "work performed."

- *Martin Eden*, Jack London

1 Introduction

1.1 Importance de la voie RAS/MAPK

Une propriété fondamentale du vivant réside dans la capacité de percevoir des informations provenant du monde extérieur. Les voies de signalisation constituent la clé de voûte de la machinerie cellulaire, permettant la transmission des signaux externes vers le noyau et autres centres où l'information sera intégrée. La transmission des signaux externes est essentielle au bon déroulement des processus développementaux ainsi qu'au maintien de l'homéostasie. Par extension, bon nombre de pathologies et problèmes développementaux sont connus comme étant associés, voire même causés, par des processus de signalisation cellulaire déficients. À ce titre, la voie RAS/MAPK est sans doute l'une des voies de signalisation les plus célèbres, principalement de par son lien à l'oncogenèse et à la progression du cancer, à présent très bien documenté (Malumbres and Barbacid 2003, Roberts and Der 2007, Schubbert, Shannon et al. 2007, Zebisch, Czernilofsky et al. 2007).

Dans une cellule saine, la signalisation RAS/MAPK est centrale lors de la transmission de bon nombre de signaux différents. Cette signalisation joue un rôle essentiel à plusieurs étapes du développement et dans le maintien du fonctionnement cellulaire normal chez l'adulte. Selon le contexte, elle peut être associée à la prolifération, la différentiation, la motilité ou la sénescence (Pearson, Robinson et al. 2001). Plusieurs facteurs peuvent intervenir sur la réponse biologique ultimement suscitée; ceux-ci incluent l'interaction avec d'autres voies de signalisation, la durée et l'amplitude du signal ainsi que la régulation de facteurs agissant en aval tel que les facteurs de transcription (Qui and Green 1992, Kolch 2005, McKay and Morrison 2007, Hurlbut, Kankel et al. 2009). Une suractivation de la voie, typiquement causée par la mutation d'une de ses composantes centrales ou d'un régulateur périphérique, sera quant à elle associée à la réponse de prolifération excessive essentielle au développement du cancer ainsi qu'à d'autres pathologies (Schubbert, Shannon et al. 2007, Fernandez-Medarde and Santos 2011). En fait, plusieurs des composantes centrales RAS/MAPK sont des oncogènes connus, parmi les plus fréquemment observés dans les

cancers en général; un des meilleurs exemples est RAS qui, à lui seul, est muté dans 30 % des cancers (Roberts and Der 2007, Fernandez-Medarde and Santos 2011).

Connaissant les rôles centraux de la voie RAS/MAPK, à la fois dans un contexte cellulaire normal et dans le développement du cancer, il semble alors d'une évidence claire qu'il est important d'en approfondir notre compréhension. Jusqu'à présent, l'identification et la caractérisation des régulateurs de la voie a été l'un des principaux moteurs du développement de notre compréhension. L'ensemble du travail décrit dans ces pages vise la poursuite de cet effort en s'inspirant d'une lignée d'avancées successives remontant à la découverte d'un virus porteur de l'oncogène RAS en 1964 (Harvey 1964) en passant par les premiers liens tissés entre RAS et le récepteur tyrosine kinase (« RTK ») EGFR dans les années 1980 (Kamata and Feramisco 1984, Mulcahy, Smith et al. 1985, Smith, DeGudicibus et al. 1986).

1.2 La signalisation RTK/RAS/ERK élargie

Cette section résume les principales avancées dans l'identification de facteurs de la voie RAS/MAPK mettant l'emphase sur les avancées apportées par l'utilisation des organismes modèles *D. melanogaster* et *C. elegans*. Ce texte est paru dans revue Médecine Sciences en tant qu'article de synthèse avec le sous-titre : « La signalisation RTK/RAS/ERK élargie : Contributions de la génétique à l'assemblage d'un réseau de signalisation » (Ashton-Beaucage and Therrien 2010).

Contributions de l'auteur : Élaboration et rédaction du manuscrit. Montage des figures.

Contributions détaillées des co-auteurs : TM a supervisé la rédaction et apporté des corrections et changements au texte. L'éditeur de M/S a également apporté certains changements au niveau de la langue (grammaire/syntaxe et terminologie).

La signalisation RTK/RAS/ERK élargie

Contributions de la génétique à l'assemblage d'un réseau de signalisation

The Greater RTK/RAS/ERK signalling pathway

How Genetics has Helped Piece Together a Signalling Network

Dariel Ashton-Beaucage et Marc Therrien

D. Ashton-Beaucage : Institut de Recherche en Immunologie et Cancérologie, Université de Montréal, C.P. 6128 Succursale Centre-Ville, Montréal, Québec, Canada.
M. Therrien : Institut de Recherche en Immunologie et Cancérologie, Département de Pathologie et Biologie Cellulaire, Université de Montréal, C.P. 6128 Succursale Centre-Ville, Montréal, Québec, Canada.

1.2.1 Résumé

Les cellules possèdent des réseaux protéiques permettant de percevoir des stimulus externes et de réagir à ceux-ci. Ces réseaux sont décomposables en voies de signalisation qui transmettent les signaux du microenvironnement à divers effecteurs cellulaires. Les études sur la signalisation par les récepteurs à activité tyrosine kinase (RTK) ont été les premières à décrire le mécanisme par lequel un signal extracellulaire est transmis jusqu'au noyau pour induire une réponse transcriptionnelle. Quoique moins bien connues, les études génétiques utilisant la drosophile ou le nématode ont contribué de manière déterminante à façonner notre compréhension de la signalisation par les RTK. Nous décrivons ici brièvement l'apport de ces organismes à la mise en évidence des nombreuses protéines qui forment ou modulent l'axe de signalisation RTK/RAS/ERK.

1.2.2 Abstract

Cells respond to changes in their environment, to developmental cues and to pathogen aggression through the action of a complex network of proteins. These networks can be decomposed into a multitude of signalling pathways that relay signals from the microenvironment to the cellular components involved in eliciting a specific response. Perturbations in these signalling processes are at the root of multiple pathologies, the most notable of these being cancer. The study of receptor tyrosine kinase (RTK) signalling led to the first description of a mechanism whereby an extracellular signal is transmitted to the nucleus to induce a transcriptional response. Genetic studies conducted in *Drosophila* and nematodes have provided key elements to this puzzle. Here, we briefly discuss the somewhat lesser known contribution of these multicellular organisms to our understanding of what has become a prototype in cell signalling as well as to the more recent description of the complex network of regulators that is now known to govern RTK/RAS/ERK signalling.

1.2.3 Introduction

Les Récepteurs Tyrosine Kinase (RTKs) jouent un rôle prépondérant dans divers processus cellulaires de base, comme la régulation de la prolifération, de la différenciation et de la survie cellulaire. L'action des RTKs passe principalement par l'activation de voies de

signalisation effectrices, dont l'une des mieux caractérisées, et probablement des plus influentes, est communément appelée la voie RAS/ERK. La description des événements moléculaires reliant l'activation des RTKs à la modulation de facteurs de transcription par la kinase ERK (également connue sous le nom de MAPK) a constitué l'une des percées majeures en signalisation cellulaire. Au milieu des années 80, les premiers liens étaient déjà tissés par des données reliant l'activité de la petite GTPase RAS, l'oncoprotéine la plus fréquemment mutée dans les cancers humains (Roberts and Der 2007), à celle d'une autre oncoprotéine bien connue, soit le RTK EGFR (Malumbres and Barbacid 2003). Ces études initiales ont donné lieu à une série de découvertes qui ont permis de comprendre comment un signal extracellulaire est transmis à l'intérieur de la cellule. Ici, nous nous attardons principalement au recours à la drosophile (*D. melanogaster*) et au nématode (*C. elegans*) qui a contribué à l'élucidation de ce processus. Quoique les approches biochimiques aient résolu plusieurs aspects de cette question (Malumbres and Barbacid 2003), l'exploitation de la génétique de la drosophile et du nématode au cours des années 90 s'est avérée particulièrement utile pour non seulement confirmer l'identité et l'ordre respectif des composantes de base de la voie de signalisation, mais également pour distinguer de nouveaux régulateurs (Tableau 1.I).

Tableau 1.I Composantes de l'axe RTK/RAS/ERK communes à la drosophile, le nématode et l'humain.

Les symboles génériques représentés ici sont utilisés dans l'article pour désigner les homologues des protéines à travers les espèces. Les symboles officiels correspondants pour les protéines de D. melanogaster (FlyBase), C. elegans (WormBase) et les gènes de H. sapiens (HGNC) sont listés ici. Les symboles alternatifs communément employés sont entre parenthèses. Les homologues ne se trouvant pas dans la littérature ont été identifiées à l'aide de l'outil HCOP du HGCN (www.genenames.org), Ensembl (www.ensembl.org) ou par recherche BLAST. Dans un but de simplicité, les homologues des RTKs ne sont pas présentés ici. Les astérisques indiquent les protéines qui furent initialement associées à la signalisation RTK/RAS/ERK grâce à la drosophile ou au nématode. Un point d'interrogation est placé à la suite des homologues pour lesquels une fonction dans la voie MAPK n'a pas été décrite.

Nom générique	*D. melanogaster*	*C. elegans*	*H. Sapiens*	Fonction moléculaire
14-3-3	14-3-3ε, 14-3-3ζ	FTT-2 (?), PAR-5 (?)	YWHAE, YWHAB, YWHAZ, YWHAH, YWHAQ, SFN(?), YWHAG	Adaptateur / chaperonne
CBL	CBL	SLI-1*	CBL, CBLB, CBLC	Ubiquitine ligase E3
CDC37	CDC37	CDC-37 (?)	CDC37, CDC37L1 (?), CDC37P1 (?)	Chaperonne
CNK	CNK*	CNK-1	CNKSR1, CNKSR2, CNKSR3 (?), IPCEF1 (?)	Adaptateur
EIF4AIII	EIF4AIII*	Y65B4A.6 (?), F33D11.10 (?)	EIF4A3 (?)	Facteur d'épissage
ERK	RL	MPK-1	MAPK1 (ERK2), MAPK3 (ERK1)	S/T kinase
GAB	DOS	SOC-1	GAB1, GAB2, GAB3, GAB4 (?)	Adaptateur
GAP1	GAP1*	GAP-1	RASA2 (GAP1M), RASA3	GAP
GCKIII	GCKIII*	GCK-1 (?)	STK24	S/T kinase
GRB2	DRK	SEM-5	GRB2	Adaptateur
HSP90	HSP83	DAF-21 (?)	HSP90AB1, HSP90AA1 (?)	Chaperonne
HYP	AVE (HYP)*	AVE-1 (?)	SAMD10 (?), SAMD12 (?)	Adaptateur
KSR	KSR*	KSR-1*, KSR-2	KSR1, KSR2	Adaptateur
MAGO	MAGO*	MAG-1 (?)	MAGOH (?), MAGOHB (?)	Facteur d'épissage
MEK	DSOR1	MEK-2	MAP2K1, MAP2K2	S/T/Y kinase
MKP3	MKP3	LIP-1	DUSP6, DUSP7, DUSP9	S/T/Y phosphatase
NF1	NF1	-	NF1	GAP
PP1	PP1-13C (?), PP1-87B (?), PP1α-96A (?)	GSP-1 (?), GSP-2 (?), GSP-3 (?), PPH-1 (?)	PPP1CA	S/T phosphatase
PP2A C	MTS	LET-92	PPP2CA, PPP2CB	S/T phosphatase
PP2A B	TWS (?)	SUR-6	PPP2R2A, PPP2R2B, PPP2R2C, PPP2R2D	S/T phosphatase
PP2C	ALPH*	TAG-93(?)	PPM1A, PPM1B (?)	S/T phosphatase
PPM1	PPM1*	T23F11.1(?)	PPM1A, PPM1B (?)	S/T phosphatase
PTP-ER	PTP-ER	-	PTPRR (PTP-SL), PTPN5 (STEP)	Y phosphatase
PUM	PUM (?)	FBF-1*, FBF-2*	PUM1 (?), PUM2*	Facteur liant l'ARNm
RAF	PHL	LIN-45	RAF1, BRAF, ARAF	S/T kinase
RAS	RAS85D	LET-60	KRAS, NRAS, HRAS	GTPase
RASGAP	VAP	GAP-2	RASA1 (p120GAP)	GAP
RNPS1	RNPS1*	RNP-5 (?)	RNPS1 (?)	Facteur d'épissage
SHP-2	CSW*	PTP-2	PTPN11 (SHP2)	Y phosphatase
SOS	SOS*	SOS-1 (LET-341)	SOS1, SOS2	GEF
SPRY	STY*	-	SPRY1, SPRY2, SPRY3, SPRY4	Adaptateur
SRC	SRC42A	SRC-2 (?)	SRC, FRK	Y kinase
SUR-8	SUR-8 (?)	SOC-2 (SUR-8)*	SHOC2	Adaptateur
TSU	TSU*	RNP-4 (?)	RBM8A (?)	Facteur d'épissage

1.2.4 Premières descriptions de l'axe de signalisation RTK/RAS/ERK

Les premières études génétiques de la signalisation RTK/RAS parurent en 1990. Le duo Min Han et Paul Sternberg, ainsi que le groupe de Robert Horvitz, ont isolé indépendamment RAS (LET-60) comme facteur positif agissant en aval du RTK EGFR (LET-23). Les deux groupes ont employé des techniques de génétique fondées sur la perturbation du développement de la vulve de *C. elegans*, un processus dépendant de la signalisation RTK (Sternberg and Han 1998). Quelques mois plus tard, à l'aide d'une stratégie génétique analogue, mais exploitant le développement de l'œil de la drosophile, le groupe de Gerald Rubin décela RAS comme facteur requis en aval du RTK Sevenless (SEV) (Wassarman, Therrien et al. 1995). Ensemble, ces travaux confirmèrent *in vivo* la position relative de RAS par rapport aux RTKs, telle qu'on l'avait précédemment anticipée par des études biochimiques. Il devenait donc clair que les mécanismes de signalisation induit par les RTKs et RAS étaient apparentés et que l'application concertée d'approches génétiques et biochimiques permettrait vraisemblablement leur élucidation. En accord avec cette idée, plusieurs travaux effectués au cours de cette période permirent dans un premier temps de comprendre comment les RTKs modulent l'activité de RAS. Par exemple, une étude réalisée chez la drosophile distingua SOS (Son of Sevenless) comme facteur essentiel à la signalisation RTK et le positionna entre les RTKs et RAS (Wassarman, Therrien et al. 1995). Son homologie à CDC25 de levure, un facteur d'échange GDP/GTP pour RAS, suggéra que SOS joue probablement un rôle dans l'activation de RAS en favorisant sa liaison au GTP. Par ailleurs, le groupe de Robert Horvitz identifia chez *C. elegans* SEM-5, un multi-adaptateur apparenté à GRB2 agissant aussi entre les RTKs et RAS (Sternberg and Han 1998). À la suite de ces découvertes, huit articles (McCormick 1993) publiés en 1993 ont montré comment les RTKs activés recrutent SOS à leur portion cytoplasmique via GRB2, permettant ainsi l'activation de RAS. L'inactivation de RAS, quant à elle, avait été établie : elle est biochimiquement sous l'emprise des protéines GAPs (GTPase Activating Proteins) où deux protéines avaient été associées à cette activité, soit RASGAP et le suppresseur de tumeurs NF1 (Bernards 2003). Un crible génétique chez la drosophile en dégagea une troisième, soit GAP1 (Wassarman, Therrien et al. 1995).

Mais alors, qu'en était-il des trois kinases de base du module RAS/ERK, soit RAF, MEK et ERK? Dans les années 80, il avait été montré que la transformation cellulaire par une forme oncogénique de RAF ne dépendait pas de RAS ; en revanche, celle induite par RAS nécessitait RAF (Malumbres and Barbacid 2003). En accord avec ces résultats, Dickson *et coll*. positionnèrent génétiquement RAF de la drosophile en aval de SEV, SOS et RAS (Wassarman, Therrien et al. 1995), ce qui fut également confirmé chez *C. elegans* (Sternberg and Han 1998). Quelque temps après, le groupe de Rubin identifia un complexe de protéines chaperonnes, soit HSP90 (HSP83) et CDC37, comme facteurs agissant en aval des RTKs (Cutforth and Rubin 1994). Leur rôle demeura incertain jusqu'à ce qu'une étude génétique chez la drosophile démontre leur importance pour l'activité de RAF (van der Straten, Rommel et al. 1997). Par la suite, des études biochimiques confirmèrent ces résultats en montrant que ces protéines sont essentielles pour le repliement et la stabilité de RAF (Schulte, Blagosklonny et al. 1995; Grammatikakis, Lin et al. 1999). Le repérage de MEK par approche génétique fut réalisé chez la drosophile quelques mois après son identification biochimique. En effet, un crible génétique basé sur la réversion de la létalité causée par un allèle mutant de RAF, permit l'identification de MEK comme facteur agissant en aval de RAF et des RTKs (Wassarman, Therrien et al. 1995). Finalement, ERK fut positionnée par analyse génétique chez la drosophile et le nématode en aval de RAS et de RAF, mais en amont de facteurs de transcription RTK/RAS-dépendants (Wassarman, Therrien et al. 1995; Sternberg and Han 1998). En l'espace de quelques années, les composantes centrales de la voie avaient donc déjà été identifiées et disposées dans l'ordre où nous les voyons maintenant représentées dans tous les documents de référence (Figure 1.1).

Figure 1.1 La voie RTK/RAS/ERK simplifiée.

Représentation des composantes centrales de la voie MAPK.

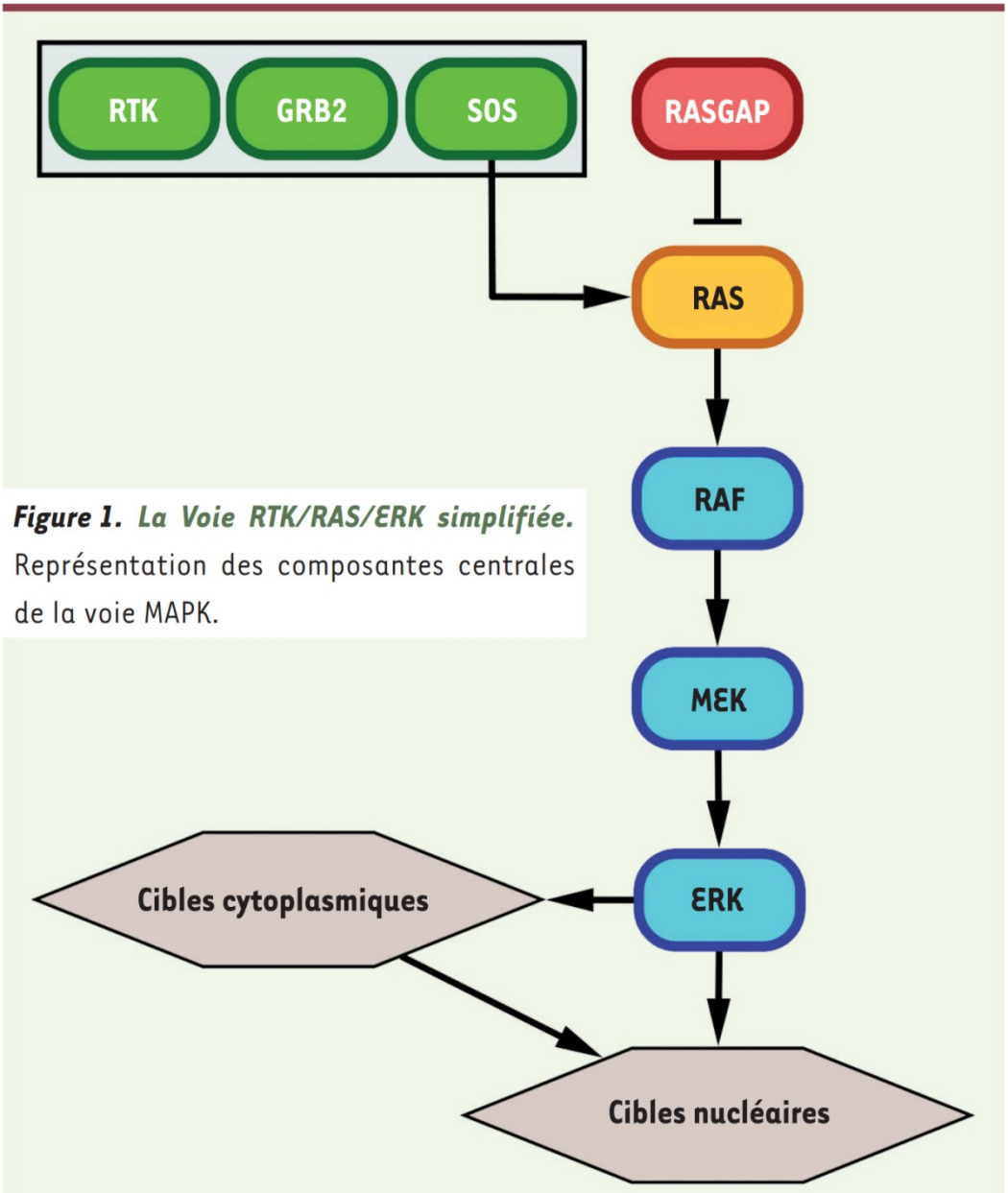


Figure 1. ***La Voie RTK/RAS/ERK simplifiée.***
Représentation des composantes centrales de la voie MAPK.

1.2.5 Percées de la génétique

L'utilité des approches génétiques se refléta aussi par leur capacité à repérer de nouveaux éléments encore insoupçonnés. Par exemple, le groupe de Norbert Perrimon distingua une composante à action positive en aval du RTK Torso (TOR), soit CSW (Corkscrew; aussi appelée SHP-2), une tyrosine phosphatase comprenant deux domaines SH2 (Perkins, Larsen et al. 1992). La fonction de SHP-2 dans la voie RAS/ERK demeura nébuleuse jusqu'à l'identification d'une autre protéine agissant entre les RTKs et RAS, soit DOS (Daugther of sevenless ; homologue des GABs de mammifères), une protéine à domaine PH (Raabe, Riesgo-Escovar et al. 1996). En effet, il fut démontré que cette GAB recrute SHP-2 dans la membrane plasmique après l'activation d'un RTK (Herbst, Zhang et al. 1999) et que cet événement permet une activation prolongée de RAS soit en déphosphorylant un site de liaison de RASGAP sur le RTK et probablement en favorisant le recrutement de GRB2 (Cleghon, Feldmann et al. 1998). Parallèlement aux protéines GAP, de nouvelles protéines ayant un rôle négatif furent aussi repérées. Par exemple, une étude cherchant à découvrir des facteurs participant au développement du système respiratoire de l'embryon de mouche identifia SPRY (Sprouty) comme antagoniste de la signalisation ERK déclenchée par le FGF. L'étude démontra également que l'expression de SPRY est dépendante de la signalisation FGF, décrivant ainsi une boucle de rétroaction négative (Hacohen, Kramer et al. 1998). L'activité de SPRY n'est pas restreinte à la signalisation par le FGF ; elle module aussi d'autres RTKs. Toutefois, son mécanisme d'action ne demeure que partiellement compris : différents scénarios ont été suggérés. Par exemple, selon le contexte cellulaire, SPRY semble capable d'inhiber l'activation de RAF, de séquestrer GRB2 ou bien de recruter GAP1 au RTK (Kim and Bar-Sagi 2004). De plus, des travaux récents ont montré que SHP-2 entre en jeu dans la déphosphorylation d'un résidu phospho-tyrosine de SPRY nécessaire pour ses effets inhibiteurs. Ceci suggère donc un autre mode par lequel le complexe GAB/SHP-2 modulerait la signalisation RAS/ERK (Jarvis, Toering et al. 2006). Une autre composante à action négative est le proto-oncogène CBL, soit une ubiquitine ligase de type E3 en action dans l'internalisation et la dégradation des RTKs (Dikic and Schmidt 2007). Cette dernière fut reliée pour la première fois à la signalisation RTK/RAS grâce à une étude génétique du développement de la vulve chez *C. elegans* (Yoon, Lee et al. 1995). De façon surprenante, il

semblerait que SPRY ait également la capacité de lier et de séquestrer CBL. Cette fonction positive de SPRY sur la signalisation RTK lui confère donc un rôle plus nuancé qui pourrait dépendre du contexte cellulaire (Kim and Bar-Sagi 2004).

Au milieu des années 90, il avait été démontré que l'activation de RAS entraine le recrutement et l'activation de RAF dans la membrane cellulaire, provoquant ensuite la phosphorylation et l'activation en cascade de MEK et d'ERK (Campbell, Khosravi-Far et al. 1998; Malumbres and Barbacid 2003). ERK activée phosphoryle à son tour un répertoire spécifique de substrats permettant ainsi une réponse cellulaire (Campbell, Khosravi-Far et al. 1998). Quoique cette description simple de la voie RAS/ERK corresponde toujours à notre compréhension de celle-ci, plusieurs observations laissaient présager l'existence d'autres protéines nécessaires à son fonctionnement. En 1995, les groupes de Rubin, Horvitz et Han effectuèrent des cribles génétiques dépendants de formes constitutivement actives de RAS chez la drosophile et le nématode dans le but de détecter de nouvelles protéines agissant en aval de RAS. Ces travaux permirent la découverte d'une nouvelle composante essentielle au module RAS/ERK, nommée KSR (Kinase Suppressor of RAS) (Kornfeld, Hom et al. 1995; Sundaram and Han 1995; Therrien, Chang et al. 1995). Étant donné son étroite homologie avec RAF, on lui attribua d'abord un rôle de kinase agissant sur RAF ou dans une voie parallèle. Par contre, des études biochimiques et fonctionnelles suggérèrent que KSR est dépourvue d'activité catalytique et lui attribuèrent plutôt un rôle de protéine d'échafaudage capable d'une part, de recruter MEK à RAF et, d'autre part, de participer au processus d'activation de RAF par RAS (Douziech, Sahmi et al. 2006). Des études sur des souris où le gène *KSR1* fut inactivé ont confirmé son rôle dans la signalisation RAS/MAPK et ont montré une perturbation du développement de tumeurs mammaires ou de papillomes RAS dépendants (Nguyen, Burack et al. 2002; Lozano, Xing et al. 2003). De plus, KSR1 et KSR2 semblent également avoir la capacité d'agir comme adaptateurs liant RAF, MEK et ERK (Nguyen, Burack et al. 2002; Claperon and Therrien 2007). Il convient de noter cependant qu'il existe très peu de données fonctionnelles sur KSR2. Finalement, l'inactivité catalytique de KSR demeure toutefois contestée par certains (Claperon and Therrien 2007).

À la suite de la découverte de KSR, un crible génétique basé sur un phénotype produit par l'expression d'une forme dominante-négative de KSR dans l'œil de mouche détecta CNK

(Connector eNhancer of KSR)(Therrien, Wong et al. 1998), une autre protéine multi-adaptatrice qui agit en reliant le complexe KSR/MEK et RAF. Des travaux récents de notre laboratoire ont montré que l'activité de CNK est régie par la tyrosine kinase Src42A (famille des kinases Src)(Laberge, Douziech et al. 2005). Quoique le recrutement de RAF à CNK semble direct, celui de KSR/MEK à CNK fait intervenir une autre protéine appelée HYP (*Hyphen* ou *ave*; *aveugle*) (Douziech, Sahmi et al. 2006). Cette dernière n'est constituée que d'un domaine SAM qui hétérodimérise avec le domaine SAM de CNK (Rajakulendran, Sahmi et al. 2008). La formation d'un complexe comprenant RAF, KSR, MEK, CNK et HYP s'avère donc critique pour la transmission de signaux au sein du module RAS/ERK chez la drosophile. Non seulement la formation d'un tel complexe permet le recrutement de MEK à RAF, mais il participe directement au mécanisme d'activation de RAF où KSR joue un rôle essentiel indépendamment de sa propriété de protéine d'échafaudage (Douziech, Sahmi et al. 2006). En effet, la formation du dimère KSR-RAF induit l'adoption d'une conformation active par RAF (Rajakulendran, Sahmi et al. 2009; Lavoie and Therrien 2010). Chez les mammifères, deux des trois homologues de CNK, soit CNK1 et CNK2, interagissent également avec RAF et jouent un rôle dans la signalisation MAPK. Toutefois, l'importance du rôle de CNK1 et CNK2 semble varier selon le RTK stimulant la voie (Claperon and Therrien 2007). Pour ce qui est des deux homologues de HYP chez les mammifères, SamD10 et Samd12, aucune étude n'est actuellement disponible à leur sujet.

Avec KSR et CNK, une autre protéine d'échafaudage, SUR-8, a été détectée chez *C. elegans* (Selfors, Schutzman et al. 1998; Sieburth, Sun et al. 1998). L'homologue humain de SUR-8 aurait la capacité d'associer RAS et RAF (Li, Han et al. 2000). En outre, une étude récente a montré que SUR-8 humain s'associerait plutôt à M-Ras, un membre distant de la famille Ras, de même qu'à la phosphatase PP1 et que cet événement provoquerait la déphosphorylation d'un site de liaison des protéines 14-3-3 sur la partie N-terminale de RAF (Rodriguez-Viciana, Oses-Prieto et al. 2006). Étant donné que l'association 14-3-3 / $RAF^{N\text{-terminal}}$ inhibe l'activité de RAF (voir ci-dessous), SUR-8 contribuerait à la signalisation RAS/ERK en défavorisant cette interaction.

Le crible RAS-dépendant qui révéla KSR chez la drosophile permit aussi l'identification de la sous-unité catalytique de la Ser/Thr-phosphatase PP2A, appelée MTS

(MicroTubule Star) (Wassarman, Solomon et al. 1996). Celle-ci semble jouer à la fois un rôle négatif entre RAS et RAF et un rôle positif en aval de RAF. Quoique le lien entre PP2A et le module RAS/ERK chez la Drosophile n'ait pas été étudié plus en détail, il est fort probable que son rôle positif soit relié à la régulation de l'association entre les protéines 14-3-3 et leurs sites de liaison phospho-dépendants situés dans la partie N-terminale de RAF et KSR, tel qu'on l'a observé chez les mammifères (Dougherty and Morrison 2004). En effet, cette interaction séquestrerait RAF et KSR dans le cytoplasme, les maintenant ainsi dans une conformation inactive. Cependant, les protéines 14-3-3 n'agissent pas uniquement comme antagonistes : leur liaison à l'extrémité C-terminale de RAF est également essentielle pour son activité catalytique (Tzivion, Luo et al. 1998). En accord avec ces rôles clés, deux études génétiques chez la mouche ont validé l'importance fonctionnelle des protéines 14-3-3 comme régulateurs de l'activité de RAF (Chang and Rubin 1997; Kockel, Vorbruggen et al. 1997). Par ailleurs, des analyses génétiques de la sous-unité catalytique (LET-92) et régulatrice B (SUR-6) de PP2A chez le nématode ont aussi confirmé l'importance de cette activité phosphatase pour réguler RAF et KSR (Sieburth, Sundaram et al. 1999; Yoder, Chong et al. 2004).

Une autre phosphatase découverte grâce au crible RAS-dépendant chez la drosophile est PTP-ER (Protein Tyrosine Phosphatase-ERK/enhancer of RAS) (Karim and Rubin 1999). Celle-ci est apparentée de façon éloignée aux Tyr-phosphatases PTP-SL et STEP de mammifères. Tout comme ses distants homologues, PTP-ER inhibe ERK en déphosphorylant le résidu tyrosine du motif TEY situé dans sa boucle activatrice. PTP-ER est elle-même phosphorylée et inactivée par ERK, décrivant ainsi une boucle de rétroaction. Par ailleurs, MKP3, une phosphatase à double spécificité (Ser/Thr et Tyr) agissant aussi sur ERK, a été mise en évidence par homologie de séquence chez la mouche et le nématode (Berset, Hoier et al. 2001; Kim, Kwon et al. 2002). Une étude chez la drosophile a montré que PTP-ER et MKP3 peuvent agir indépendamment l'une de l'autre, ou de manière concertée selon le contexte cellulaire (Rintelen, Hafen et al. 2003). Plus récemment, l'existence d'une boucle de rétroaction négative engageant l'expression ERK-dépendante de MKP3 a également été décrite (Gomez, Lopez-Varea et al. 2005). Finalement, PP2C (ALPH), un membre de la famille des Ser/Thr phosphatase de type PPM, a été identifiée par notre groupe à la suite du

crible KSR-dépendant mentionné ci-dessus. PP2C semble agir négativement à une étape en aval ou en parallèle de RAS (Baril and Therrien 2006). L'identité de son substrat demeure inconnue. Cependant, sa capacité à agir sur les kinases des voies JNK et p38 fonctionnellement apparentées à RAF, suggère qu'elle pourrait aussi agir sur RAF et/ou KSR (Baril, Sahmi et al. 2009).

Le contrôle de la signalisation n'opère pas uniquement au niveau post-traductionnel : plusieurs évidences indiquent un contrôle de l'expression de composantes centrales de la voie. Par exemple, il a été montré que le miARN *let-7* module l'expression de RAS chez *C. elegans* et chez l'homme (Johnson, Grosshans et al. 2005). Également chez *C. elegans*, FBF-1 et FBF-2, des facteurs de la famille Pumilio liant l'ARN, ont été montrés comme régulant la stabilité de l'ARN messager de ERK (Lee, Hook et al. 2007). Plus récemment, notre groupe ainsi que celui de Jessica Treisman ont aussi montré que l'épissage des pré-ARN messagers de ERK de drosophile est régulé par quatre composantes du complexe EJC (*exon junction complex*), soit MAGO, TSU, EIF4AIII et RNPS1 (Ashton-Beaucage, Udell et al. 2010; Roignant and Treisman 2010). Ainsi, le contrôle du niveau d'expression des composantes centrales de la voie, tel que RAS et ERK, ajoute un autre niveau de régulation au réseau RAS/ERK.

1.2.6 Perspectives

Nous avons présenté ici les principales protéines connues chez la drosophile et le nématode qui agissent entre les RTKs et ERK (Figure 1.2 et Tableau 1.I). Des homologues chez les mammifères ont été trouvés pour chaque nouvelle composante détectée, ce qui suggère que leur fonction s'est conservée au cours de l'évolution des métazoaires. Même si la majorité des protéines directement en cause dans cette voie de signalisation ont probablement été trouvées, il est aussi probable qu'il en reste encore de nouvelles à découvrir, de même que de nombreux modulateurs périphériques. Par exemple, un crible à grande échelle utilisant une banque d'ARNs interférants sur des cellules de drosophile en culture a récemment débusqué plusieurs nouvelles protéines dont la déplétion modulait la signalisation RTK/RAS/ERK (Friedman and Perrimon 2006). Quoique pour la plupart leur caractérisation reste à venir, deux d'entre elles, soit une kinase de la famille STE20 (GCKIII) et une Ser/Thr phosphatase de la famille PPM (PPM1), semblent respectivement capable de réguler l'activité

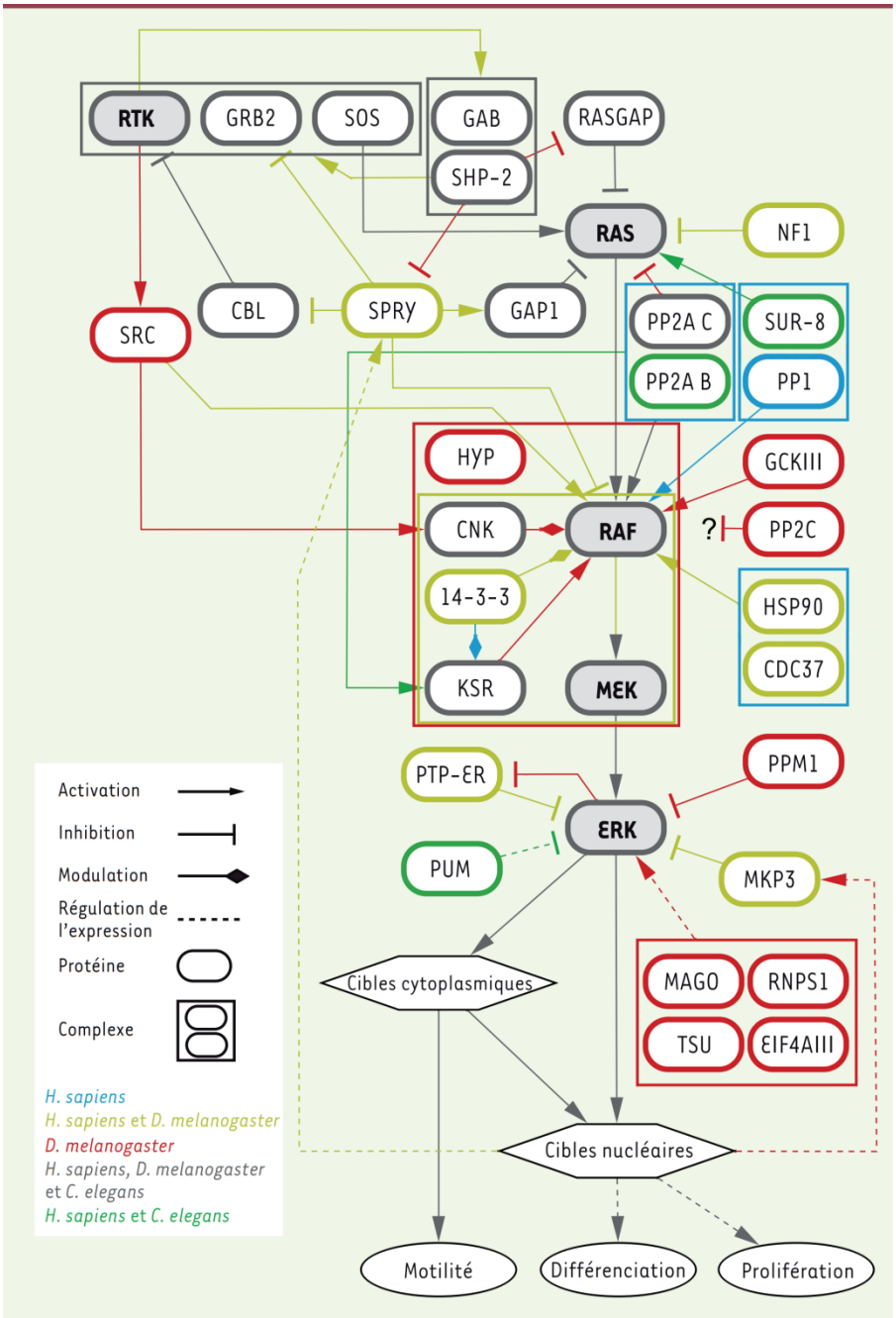


Figure 1.2 L'axe de signalisation RTK/RAS/ERK.

Représentation du module RAS/ERK et des principaux régulateurs identifiés au moyen de techniques génétiques. Les composantes, modulation et complexes sont colorés selon la présence d'évidences chez *D. melanogaster*, *C. elegans* et/ou *H. sapiens* (voir légende).

de RAF et d'ERK. Cette étude ne s'est pas limitée à répertorier de nouveaux régulateurs directs, elle a montré en outre l'importance de mécanismes de rétroaction, de même que la capacité de voies de signalisation parallèles, comme les voies JNK et Akt à influencer la signalisation RTK/RAS/ERK. De plus, cette étude suggère que diverses machineries cellulaires, notamment celle participant à la biogénèse des miARNs ou encore celle régulant le cycle cellulaire, ont pareillement la propriété de régir la signalisation RTK/RAS/ERK. Il semble donc que le niveau précis d'activation d'ERK dans une cellule donnée ne dépende pas uniquement de l'activation de la voie proprement dite, mais également de l'intégration de plusieurs signaux en provenance de diverses sources. Un défi pour les prochaines années sera d'identifier ces diverses sources et de déterminer à quelle étape chacune d'elle s'inscrit. Un autre défi de taille sera d'élucider au niveau moléculaire la nature des diverses boucles de rétroaction et comment ces dernières sont utilisées pour régler dynamiquement l'intensité et la durée du signal. Ainsi, dans la mesure où ces phénomènes sont conservés, la drosophile et le nématode devraient plusieurs années encore servir d'alliés puissants dans la quête d'une compréhension fine des mécanismes moléculaires régissant la signalisation RTK/RAS/ERK.

1.2.7 Remerciements

Nous remercions Caroline Baril, Gino Laberge et Martin Lefrançois pour leurs commentaires sur le manuscrit. DAB est récipiendaire d'une bourse d'études doctorales des IRSC. MT est récipiendaire d'une chaire de recherche du Canada en signalisation intracellulaire. Ce travail fut supporté par des fonds de recherche des IRSC et de la SCC octroyés à MT.

1.2.8 Références

Ashton-Beaucage, D., C. M. Udell, et al. (2010). "The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in *Drosophila*." Cell **143**(2): 251-262.

Baril, C., M. Sahmi, et al. (2009). "The PP2C Alphabet Is a Negative Regulator of Stress-Activated Protein Kinase Signaling in *Drosophila*." Genetics **181**(2): 567-579.

Baril, C. and M. Therrien (2006). "Alphabet, a Ser/Thr phosphatase of the protein phosphatase 2C family, negatively regulates RAS/MAPK signaling in *Drosophila*." Dev Biol **294**(1): 232-245.

Bernards, A. (2003). "GAPs galore! A survey of putative Ras superfamily GTPase activating proteins in man and *Drosophila*." Biochim Biophys Acta **1603**(2): 47-82.

Berset, T., E. F. Hoier, et al. (2001). "Notch inhibition of RAS signaling through MAP kinase phosphatase LIP-1 during C. elegans vulval development." Science **291**(5506): 1055-1058.

Campbell, S. L., R. Khosravi-Far, et al. (1998). "Increasing complexity of Ras signaling." Oncogene **17**(11 Reviews): 1395-1413.

Chang, H. C. and G. M. Rubin (1997). "14-3-3 epsilon positively regulates Ras-mediated signaling in *Drosophila*." Genes Dev **11**(9): 1132-1139.

Claperon, A. and M. Therrien (2007). "KSR and CNK: two scaffolds regulating RAS-mediated RAF activation." Oncogene **26**(22): 3143-3158.

Cleghon, V., P. Feldmann, et al. (1998). "Opposing actions of CSW and RasGAP modulate the strength of Torso RTK signaling in the *Drosophila* terminal pathway." Mol Cell **2**(6): 719-727.

Cutforth, T. and G. M. Rubin (1994). "Mutations in Hsp83 and cdc37 impair signaling by the sevenless receptor tyrosine kinase in *Drosophila*." Cell **77**(7): 1027-1036.

Dikic, I. and M. H. Schmidt (2007). "Malfunctions within the Cbl interactome uncouple receptor tyrosine kinases from destructive transport." Eur J Cell Biol **86**(9): 505-512.

Dougherty, M. K. and D. K. Morrison (2004). "Unlocking the code of 14-3-3." J Cell Sci **117**(Pt 10): 1875-1884.

Douziech, M., M. Sahmi, et al. (2006). "A KSR/CNK complex mediated by HYP, a novel SAM domain-containing protein, regulates RAS-dependent RAF activation in *Drosophila*." Genes Dev **20**(7): 807-819.

Friedman, A. and N. Perrimon (2006). "A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling." Nature **444**(7116): 230-234.

Gomez, A. R., A. Lopez-Varea, et al. (2005). "Conserved cross-interactions in *Drosophila* and Xenopus between Ras/MAPK signaling and the dual-specificity phosphatase MKP3." Dev Dyn **232**(3): 695-708.

Grammatikakis, N., J. H. Lin, et al. (1999). "p50(cdc37) acting in concert with Hsp90 is required for Raf-1 function." Mol Cell Biol **19**(3): 1661-1672.

Hacohen, N., S. Kramer, et al. (1998). "sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the *Drosophila* airways." Cell **92**(2): 253-263.

Herbst, R., X. Zhang, et al. (1999). "Recruitment of the protein tyrosine phosphatase CSW by DOS is an essential step during signaling by the sevenless receptor tyrosine kinase." EMBO J **18**(24): 6950-6961.

Jarvis, L. A., S. J. Toering, et al. (2006). "Sprouty proteins are in vivo targets of Corkscrew/SHP-2 tyrosine phosphatases." Development **133**(6): 1133-1142.

Johnson, S. M., H. Grosshans, et al. (2005). "RAS is regulated by the let-7 microRNA family." Cell **120**(5): 635-647.

Karim, F. D. and G. M. Rubin (1999). "PTP-ER, a novel tyrosine phosphatase, functions downstream of Ras1 to downregulate MAP kinase during *Drosophila* eye development." Mol Cell **3**(6): 741-750.

Kim, H. J. and D. Bar-Sagi (2004). "Modulation of signalling by Sprouty: a developing story." Nat Rev Mol Cell Biol **5**(6): 441-450.

Kim, S. H., H. B. Kwon, et al. (2002). "Isolation and characterization of a *Drosophila* homologue of mitogen-activated protein kinase phosphatase-3 which has a high substrate specificity towards extracellular-signal-regulated kinase." Biochem J **361**(Pt 1): 143-151.

Kockel, L., G. Vorbruggen, et al. (1997). "Requirement for *Drosophila* 14-3-3 zeta in Raf-dependent photoreceptor development." Genes Dev **11**(9): 1140-1147.

Kornfeld, K., D. B. Hom, et al. (1995). "The ksr-1 gene encodes a novel protein kinase involved in Ras-mediated signaling in C. elegans." Cell **83**(6): 903-913.

Laberge, G., M. Douziech, et al. (2005). "Src42 binding activity regulates *Drosophila* RAF by a novel CNK-dependent derepression mechanism." Embo J **24**(3): 487-498.

Lavoie, H. and M. Therrien (2010). "[It takes two RAFs to tango]." Med Sci (Paris) **26**(5): 459-460.

Lee, M. H., B. Hook, et al. (2007). "Conserved regulation of MAP kinase expression by PUF RNA-binding proteins." PLoS Genet **3**(12): e233.

Li, W., M. Han, et al. (2000). "The leucine-rich repeat protein SUR-8 enhances MAP kinase activation and forms a complex with Ras and Raf." Genes Dev **14**(8): 895-900.

Lozano, J., R. Xing, et al. (2003). "Deficiency of kinase suppressor of Ras1 prevents oncogenic ras signaling in mice." Cancer Res **63**(14): 4232-4238.

Malumbres, M. and M. Barbacid (2003). "RAS oncogenes: the first 30 years." Nat Rev Cancer **3**(6): 459-465.

McCormick, F. (1993). "Signal transduction. How receptors turn Ras on." Nature **363**(6424): 15-16.

Nguyen, A., W. R. Burack, et al. (2002). "Kinase suppressor of Ras (KSR) is a scaffold which facilitates mitogen-activated protein kinase activation in vivo." Mol Cell Biol **22**(9): 3035-3045.

Perkins, L. A., I. Larsen, et al. (1992). "corkscrew encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso." Cell **70**(2): 225-236.

Raabe, T., J. Riesgo-Escovar, et al. (1996). "DOS, a novel pleckstrin homology domain-containing protein required for signal transduction between sevenless and Ras1 in *Drosophila*." Cell **85**(6): 911-920.

Rajakulendran, T., M. Sahmi, et al. (2008). "CNK and HYP form a discrete dimer by their SAM domains to mediate RAF kinase signaling." Proc Natl Acad Sci U S A **105**(8): 2836-2841.

Rajakulendran, T., M. Sahmi, et al. (2009). "A dimerization-dependent mechanism drives RAF catalytic activation." Nature **461**(7263): 542-545.

Rintelen, F., E. Hafen, et al. (2003). "The *Drosophila* dual-specificity ERK phosphatase DMKP3 cooperates with the ERK tyrosine phosphatase PTP-ER." Development **130**(15): 3479-3490.

Roberts, P. J. and C. J. Der (2007). "Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer." Oncogene **26**(22): 3291-3310.

Rodriguez-Viciana, P., J. Oses-Prieto, et al. (2006). "A phosphatase holoenzyme comprised of Shoc2/Sur8 and the catalytic subunit of PP1 functions as an M-Ras effector to modulate Raf activity." Mol Cell **22**(2): 217-230.

Roignant, J. Y. and J. E. Treisman (2010). "Exon junction complex subunits are required to splice *Drosophila* MAP kinase, a large heterochromatic gene." Cell **143**(2): 238-250.

Schulte, T. W., M. V. Blagosklonny, et al. (1995). "Disruption of the Raf-1-Hsp90 molecular complex results in destabilization of Raf-1 and loss of Raf-1-Ras association." J Biol Chem **270**(41): 24585-24588.

Selfors, L. M., J. L. Schutzman, et al. (1998). "soc-2 encodes a leucine-rich repeat protein implicated in fibroblast growth factor receptor signaling." Proc Natl Acad Sci U S A **95**(12): 6903-6908.

Sieburth, D. S., Q. Sun, et al. (1998). "SUR-8, a conserved Ras-binding protein with leucine-rich repeats, positively regulates Ras-mediated signaling in C. elegans." Cell **94**(1): 119-130.

Sieburth, D. S., M. Sundaram, et al. (1999). "A PP2A regulatory subunit positively regulates Ras-mediated signaling during Caenorhabditis elegans vulval induction." Genes Dev **13**(19): 2562-2569.

Sternberg, P. W. and M. Han (1998). "Genetics of RAS signaling in C. elegans." Trends Genet **14**(11): 466-472.

Sundaram, M. and M. Han (1995). "The C. elegans ksr-1 gene encodes a novel Raf-related kinase involved in Ras-mediated signal transduction." Cell **83**(6): 889-901.

Therrien, M., H. C. Chang, et al. (1995). "KSR, a novel protein kinase required for RAS signal transduction." Cell **83**(6): 879-888.

Therrien, M., A. M. Wong, et al. (1998). "CNK, a RAF-binding multidomain protein required for RAS signaling." Cell **95**(3): 343-353.

Tzivion, G., Z. Luo, et al. (1998). "A dimeric 14-3-3 protein is an essential cofactor for Raf kinase activity." Nature **394**(6688): 88-92.

van der Straten, A., C. Rommel, et al. (1997). "The heat shock protein 83 (Hsp83) is required for Raf-mediated signalling in *Drosophila*." EMBO J **16**(8): 1961-1969.

Wassarman, D. A., N. M. Solomon, et al. (1996). "Protein phosphatase 2A positively and negatively regulates Ras1-mediated photoreceptor development in *Drosophila*." Genes Dev **10**(3): 272-278.

Wassarman, D. A., M. Therrien, et al. (1995). "The Ras signaling pathway in *Drosophila*." Curr Opin Genet Dev **5**(1): 44-50.

Yoder, J. H., H. Chong, et al. (2004). "Modulation of KSR activity in Caenorhabditis elegans by Zn ions, PAR-1 kinase and PP2A phosphatase." EMBO J **23**(1): 111-119.

Yoon, C. H., J. Lee, et al. (1995). "Similarity of sli-1, a regulator of vulval development in C. elegans, to the mammalian proto-oncogene c-cbl." Science **269**(5227): 1102-1105.

1.3 ARNi et génomique fonctionnelle

Le chapitre précédent fait état des avancées substantielles qu'aura apporté la génétique classique à la compréhension de la voie RAS/MAPK. De manière générale, ces techniques auront connu beaucoup de succès dès les années 1980, le premier crible à large échelle ayant permis l'identification de gènes relié à la létalité et à la formation de patrons embryonnaires (parfois aussi connu comme étant le « patterning ») chez la drosophile (Nusslein-Volhard and Wieschaus 1980). En plus de ses contributions au domaine de la signalisation RAS/MAPK, le crible génétique aura depuis permis de relier différents gènes à une gamme impressionnante de processus biologiques importants incluant, d'ailleurs, plusieurs autres voies de signalisation (St Johnston 2002, Nagy, Perrimon et al. 2003).

Cette technologie reste employée de nos jours pour l'identification et la caractérisation de nouveaux facteurs. Plus récemment, le séquençage du génome et le développement de l' interférence par ARN (« ARNi ») ont permis l'élaboration de techniques de criblage plus sophistiquées qui permettent de combler certaines lacunes du crible génétique et, très souvent, d'améliorer l'efficacité et la rapidité de détection. La présente section traitera de ces nouvelles technologies et de ses caractéristiques en les comparant aux méthodes de génétique classique.

1.3.1 Les techniques de génétique classique et leurs limitations

La plupart des techniques de crible génétique – dites de « forward genetics » – consistent à identifier la cause génétique responsable d'un phénotype. Typiquement, une grande quantité de mutants sont générés aléatoirement à l'aide d'agents mutagènes, tels que la radiation, l'utilisation d'un composé chimique ou l'insertion aléatoire de transposons. Dans certains cas, on utilisera aussi plutôt une collection de mutants portant des délétions connues. Les mutants sont alors criblés en se basant sur la présence d'un phénotype ou alors sur leur capacité à modifier un phénotype connu. Cette dernière technique de criblage, dite « second site modifier screen », est d'ailleurs celle qui aura permis de mapper plusieurs facteurs des voies de signalisation connues à ce jour, incluant ceux de la voie RAS/MAPK.

La deuxième étape du criblage génétique consiste à identifier le gène responsable du phénotype pour chaque candidat retenu. Outre le criblage lui-même, qui peut être un processus

assez laborieux en soi, l'identification des gènes mutés nécessite généralement un effort et un temps non négligeables. Ces considérations feront en sorte que seule une partie des mutants obtenus seront typiquement analysés. Par contre, avec le développement des techniques de séquençage à haut débit, il est devenu envisageable de contourner les techniques de génétique classique pour identifier plus facilement les sites de mutations (Schneeberger 2014). En somme, les techniques de criblage génétique permettent donc en principe une couverture complète et non-biaisée du génome, dans la mesure où une quantité suffisante de mutants sont criblés et associés avec succès à des gènes.

En plus de la lourdeur des étapes expérimentales requises qui a déjà été mentionnée, il existe cependant d'autres limitations à ces techniques qui méritent d'être abordées plus en détail. D'abord, le phénotype examiné est le plus souvent de nature qualitative; la subjectivité pourra donc venir influencer la qualité des données et limiter l'identification de phénotypes ou d'altérations plus subtiles. Ensuite, les différents agents mutagènes n'agissent pas de manière tout à fait aléatoire, puisqu'il est maintenant bien établi qu'ils concentrent leurs effets sur des régions ou des séquences d'ADN particulières, induisant ainsi un biais mutationnel envers certains gènes alors que d'autres demeureront que peu touchées (St Johnston 2002, Nagy, Perrimon et al. 2003). La modification d'un phénotype peut être relativement résistante aux perturbations mineures. Ce concept défini comme la robustesse (Stelling, Sauer et al. 2004), constitue une autre lacune des approches phénotypiques qualitatives.

En somme, le criblage génétique demeure un outil puissant pour identifier des gènes reliés à un processus biologique donné, bien qu'il montre des lacunes significatives faisant en sorte qu'elle peut difficilement être considérée comme offrant une couverture compréhensive du génome. À preuve, en ce qui concerne la signalisation RAS/MAPK, plusieurs cribles génomiques successifs auront été accomplis chez la drosophile et *C. elegans* et chacun d'entre eux aura rapporté l'identification de nouveaux gènes (Ferguson and Horvitz 1985, Beitel, Clark et al. 1990, Han and Sternberg 1990, Rogge, Karlovich et al. 1991, Simon, Bowtell et al. 1991, Stern and Horvitz 1991, Gaul, Mardon et al. 1992, Perkins, Larsen et al. 1992, Olivier, Raabe et al. 1993, Tsuda, Inoue et al. 1993, Kornfeld, Hom et al. 1995, Sundaram and Han 1995, Yoon, Lee et al. 1995, Dickson, van der Straten et al. 1996, Karim, Chang et al. 1996, Raabe, Riesgo-Escovar et al. 1996, van der Straten, Rommel et al. 1997, Hacohen, Kramer et

al. 1998, Selfors, Schutzman et al. 1998, Sieburth, Sun et al. 1998, Therrien, Morrison et al. 2000, Wang, Werz et al. 2008). Bref, l'efficacité de cette technique en tant qu'outil d'identification de gènes ne fait nul doute. Par contre, il est également clair qu'elle ne permet pas une couverture compréhensive du génome. Il est donc envisageable que plusieurs facteurs aient échappé à l'examen du criblage génétique, et ce malgré un nombre d'études et d'efforts considérables.

1.3.2 Le criblage par ARNi à large échelle

À l'instar du criblage génétique (une méthode de « forward genetics »), le criblage par ARNi est une technique dite de génétique inverse (« reverse genetics »); c'est-à-dire que cette approche utilise comme point de départ un gène dont la séquence est généralement connue pour ensuite l'associer à un phénotype. Les approches de génétique inverse impliquent généralement une perturbation de l'activité du gène; dans le cas de l'ARNi, il s'agira d'en réduire l'expression. Tout comme le criblage génétique, le criblage par ARNi à large échelle est rapidement apparu comme étant un puissant outil d'annotation fonctionnelle des gènes. D'ailleurs, grâce à sa relative aisance d'emploi chez la drosophile et le ver, il pourrait être considéré comme le successeur du criblage génétique. En réalité, dû à leurs différences et aux limitations qui leur sont propres, il serait probablement plus juste de les considérer comme deux techniques complémentaires.

1.3.2.1 Utilisation de l'ARNi chez *Drosophila melanogaster*

En 1998, Fire et Mello avaient établi les bases de la technique d'ARNi chez *C. elegans* (Fire, Xu et al. 1998). Peu de temps après, la technologie fut introduite chez la drosophile où il fut démontré que des longs ARN à double brin (« ARNdb ») pouvaient réduire l'expression de gènes dans des cellules S2 en culture (Clemens, Worby et al. 2000). Le modèle S2 s'est rapidement illustré comme étant un système particulièrement bien adapté à l'ARNi, principalement grâce à la capacité des cellules en culture d'absorber directement les ARNdb sans l'aide d'agents de transfection. De plus, l'ARNdb est considérablement plus facile et moins dispendieux à générer que les petits ARN interférents (« siRNA ») et les petits ARN en épingle à cheveux (« shRNA »), couramment utilisés dans d'autres modèles. Or, l'ARNdb n'est généralement pas utilisé chez les vertébrés, puisqu'il induit une réponse immune médiée

par les interferons (Karpala, Doran et al. 2005). Suite à l'absorption dans le milieu cytoplasmique, l'ARNdb est découpé en petits fragments siRNA de 20-25 nt. Un des brins sera alors chargé sur le complexe de silençage induit par l'ARN, dit le complexe RISC (« RNA-induced silencing complex ») et servira de matrice pour guider le complexe vers les ARN messagers (« ARNm ») portant une séquence correspondante. Les molécules d'ARNdb ont généralement une longueur entre 200 et 600 nt, permettant de générer un assortiment de siRNA et d'augmenter l'efficacité de l'effet de déplétion souhaité (Ramadan, Flockhart et al. 2007).

1.3.2.2 Criblage à large échelle par ARNi

La découverte de l'ARNi chez le ver et sa première utilisation chez la mouche a lieu en parallèle de la publication des génomes de ces deux organismes (Consortium 1998, Adams, Celniker et al. 2000). Il s'agit d'une coïncidence fortuite, rendant possible et très facile de cibler tout gène d'intérêt dans ces organismes. De plus, il devint désormais possible d'envisager la création de banques d'ARNi couvrant l'ensemble du génome. D'ailleurs, les premières collections d'ARNi servant au criblage à large échelle virent le jour peu de temps après. Chez *C. elegans*, des collections de bactéries *E. coli* portant des plasmides producteurs d'ARNdb ont été générées (Kamath, Fraser et al. 2003). Les collections d'ARNdb pan-génomiques servant à la déplétion en culture cellulaire suivirent rapidement chez la mouche (Boutros, Kiger et al. 2004). Très rapidement, on constate le grand potentiel de ces approches à identifier un très grand nombre de gènes candidats. Chez le ver par exemple, les premiers criblages ARNi relient plus de gènes à des nouveaux phénotypes que l'ensemble des études de génétique classique menées depuis les 40 années précédentes (Zipperlen, Fraser et al. 2001, Nagy, Perrimon et al. 2003). Le criblage par ARNi apparait donc comme un outil de choix pour l'annotation fonctionnelle.

Chez la mouche, on constate la même aptitude à identifier rapidement une grande quantité de candidats dans les cribles ARNi. De plus, l'utilisation des ARNdb dans les cellules S2 présente de nombreux avantages pour le criblage. En effet, la possibilité d'induire l'ARNi simplement en ajoutant les ARNdb au milieu cellulaire facilite grandement l'exécution des expériences à large échelle et est particulièrement bien adaptée au format des plaques de

culture à 96 et 384 puits. Il est donc aisé d'augmenter le débit des expériences et même d'automatiser plusieurs étapes. De plus, le génome de mouche présente un taux de paralogie relativement faible, ce qui a comme avantage de réduire la robustesse phénotypique reliée à la redondance génique. La voie RAS/MAPK en est d'ailleurs un bon exemple comme on peut le voir dans le tableau 1.1.

De manière plus générale, le criblage par ARNi se révèle être un outil complémentaire au criblage génétique. La couverture plus systématique du génome et l'identification rapide des gènes candidats sont des avantages considérables et permettent de couvrir plus de terrain et ce, plus rapidement qu'avec la génétique classique. Par contre, l'ARNi ne produit pas un effet équivalent à la perte de fonction complète d'un gène. Cette caractéristique peut être un avantage lorsque la perte de fonction d'un gène causerait la létalité, une déplétion partielle permettant la survie et donc, l'étude productive du gène.

1.3.2.3 Limitations du criblage par ARNi

S'il est vrai que le criblage par ARNi présente de nombreux avantages par rapport au criblage génétique, cette technologie a toutefois des limitations qui lui sont propres. Bien que l'effet de déplétion partielle induit par l'ARNi puisse être un avantage dans certains cas, il peut également s'avérer être un inconvénient dans des cas où une déplétion complète serait requise pour générer un phénotype. De plus, l'identification rapide d'une grande quantité de candidats est un avantage impressionnant, mais cet avantage peut poser problème lorsque vient le temps de sélectionner le candidat le plus pertinent pour en faire une étude plus approfondie.

Les premières études de criblage menées chez la mouche étaient marquées par la présence d'une quantité importante de faux positifs. On se rendit rapidement compte que la présence de séquences partagées par plusieurs gènes dans les ARNdb des collections de première génération causait des effets de déplétion non-spécifiques (dits « off-target effects ») était une source considérable de résultats faux positifs (Kulkarni, Booker et al. 2006, Ma, Creanga et al. 2006, Dasgupta, Nybakken et al. 2007, Moffat, Reiling et al. 2007). Une séquence partagée aussi petite que 20 nt présente dans un ARNdb est en mesure de dépléter un ARNm distinct de sa cible initiale. De plus, certaines séquences répétées (telles que les

répétitions de type CAN) sont présentes dans de nombreux gènes et augmentent considérablement le nombre de cibles potentielles d'un ARNdb qui en contiendrait.

Vu ces limitations, il est vite devenu clair que cette technologie ne pourrait pas remplacer entièrement le criblage génétique. De plus, la promesse d'une annotation fonctionnelle rapide et compréhensive du génome serait beaucoup plus difficile que ce qui avait été envisagé initialement (Mathey-Prevot and Perrimon 2006).

1.3.2.4 Évolution du criblage par ARNi

Depuis la réalisation des premiers cribles ARNi, plusieurs améliorations ont permis de faire progresser cette technologie. Afin d'adresser le problème majeur des faux positifs dus aux séquences non-spécifiques, il est maintenant monnaie courante de valider tous les résultats de criblage avec une, voire plusieurs, sondes ARNdb non-chevauchantes. De plus, les librairies ARNi de deuxième génération sont composées d'ARNdb optimisés *in silico* pour contenir le moins de séquences non-spécifiques possible (Arziman, Horn et al. 2005, Horn, Sandmann et al. 2010). Pour les rares gènes où il n'est pas possible de générer des sondes spécifiques, les effets hors cible prédits sont annotés et peuvent être intégrés dans l'analyse post-criblage.

Une autre avancée importante est le développement d'outils permettant la validation via l'expression d'ARNm résistants, qui peuvent être employés à relativement large échelle (Kondo, Booker et al. 2009, Kondo and Perrimon 2011). L'évolution de technologies d'édition du génome, tel le CRISPR (« clustered regularly interspaced short palindromic repeats »), offre davantage de possibilités lors de la validation de résultats de criblages ARNi (Mohr, Smith et al. 2014). Des collections de mouches ARNi sont maintenant disponibles (Dietzl, Chen et al. 2007, Ni, Markstein et al. 2008) et permettent non seulement de passer rapidement de la culture cellulaire à l'organisme entier, mais également de procéder au criblage à large échelle par ARNi ciblés dans des tissus spécifiques (Mummery-Widmer, Yamazaki et al. 2009).

Un problème important dans le domaine du criblage ARNi était l'absence d'outils d'analyse et de méthodes standardisées de traitement des données, un constat surtout très apparent lorsque qu'est faite la comparaison avec d'autres techniques expérimentales à large

échelle (Brazma, Hingamp et al. 2001, Knudsen, Daston et al. 2005, Taylor, Paton et al. 2007, Martinez-Bartolome, Binz et al. 2014). En 2006, malgré ses succès en tant qu'outil de génétique inverse, le criblage par ARNi n'avait pas encore atteint sa pleine maturité en tant que technique de biologie des systèmes due à ses lacunes (Mathey-Prevot and Perrimon 2006). Depuis, des approches combinées avec la protéomique et de criblage par ARNi ont montré que la combinaison de ces deux méthodes pouvait permettre une annotation fonctionnelle plus riche, ainsi qu'aider à la sélection des candidats (Miller, Lau et al. 2009, Friedman, Tucker et al. 2011, Vinayagam, Stelzl et al. 2011, Neumuller, Wirtz-Peitz et al. 2012, Vinayagam, Hu et al. 2013, Vinayagam, Zirin et al. 2014). D'autre part, la disponibilité d'outils de microscopie automatisée plus performants et d'outils d'analyse d'imagerie adaptés a permis l'émergence du criblage à haut contenu (« high content screening » ou « HCS »), où il est possible d'extraire des données multiparamétriques, permettant de cribler plusieurs phénotypes en parallèle et de mieux regrouper les facteurs similaires (Goshima, Wollman et al. 2007, Wang, Zhou et al. 2008, Fuchs, Pau et al. 2010). Finalement, on aura vu récemment l'émergence du criblage d'interaction génétique par ARNi chez la drosophile (Axelsson, Sandmann et al. 2011, Horn, Sandmann et al. 2011, Laufer, Fischer et al. 2014). Cette technique s'inspire des criblages d'interactions synthétiques à large échelle chez la levure et permet également d'identifier des groupes de gènes à fonction reliées. Nous avons employé une stratégie de criblage d'interaction génétique par ARNi pour identifier des facteurs reliés fonctionnellement à *Usp47*, une deubiquitinase régulant les niveaux protéiques de MAPK. Ces travaux ainsi que la technique de criblage sont décrits plus en détail au chapitre 4.

Des avancées importantes ont également eu lieu au plan des outils d'analyse et du traitement des données de criblage. Il existe maintenant plusieurs logiciels et outils disponibles pour chacune des étapes, allant de l'élaboration des ARNdb jusqu'à l'analyse des données (Boutros, Bras et al. 2006, Horn and Boutros 2010, Hu, Sopko et al. 2013). De plus, plusieurs ressources regroupent maintenant les données de cribles publiés, facilitant ainsi le minage de ces données (Tableau 1.II). Finalement, à défaut d'avoir des normes de standardisation bien définies, le fait que la majorité des criblages soient effectués dans quelques grands centres – principalement au *Drosophila RNAi Screening Center*, à Harvard – a permis une certaine standardisation des procédures et des réactifs employés.

Database	URL	Organisms	RNAi reagents
FLIGHT	http://flight.icr.ac.uk/	*Drosophila melanogaster*	Double-stranded RNA (dsRNA)
FlyRNAi.org	http://www.flyrnai.org/index.html	*D. melanogaster*	dsRNA
GenomeRNAi	http://genomernai.org/	Human and *D. melanogaster*	dsRNA, siRNA and short hairpin RNA (shRNA)
PubChem BioAssay	http://www.ncbi.nlm.nih.gov/pcassay	Human, rat and *D. melanogaster*	siRNA, shRNA and dsRNA
RNAiDB	http://www.rnai.org/	*Caenorhabditis elegans*	dsRNA
WormBase	http://www.wormbase.org	*C. elegans* and other nematodes	dsRNA

Tableau 1.II Bases de données regroupant les données de cribles ARNi publiés

(tiré de Mohr, Smith et al. 2014 avec la permission de *Nature Publishing Group*)

1.4 Facteurs identifiés dans le crible RAS^{V12}

La présente et dernière partie de l'introduction traitera des facteurs identifiés dans le crible ARNi pan-génomique ayant pour but d'identifier des facteurs agissant en aval de RAS^{V12} et modulant l'activation de MAPK. Le crible est décrit dans l'article qui figure au chapitre 3 (Ashton-Beaucage, Udell et al. 2014), alors que des articles portant sur la caractérisation de facteurs identifiés dans ce même crible figurent aux chapitres 2 et 4. Par souci de concision, la présente section ne vise pas à couvrir tous les facteurs identifiés dans le crible, ni à décrire de nouveau les éléments discourus dans ces trois articles. Il sera plutôt question ici de définir les concepts importants se rapportant aux principaux régulateurs identifiés. Dans certains cas, des informations complémentaires, parues depuis la publication de nos travaux, seront ajoutées ici pour mettre à jour l'état d'avancement des connaissances dans le domaine.

1.4.1 Les facteurs d'épissage

Les facteurs d'épissage constituent le premier groupe de candidats en terme du nombre de facteurs identifiés dans le crible RAS^{V12} (décrit dans l'article du chapitre 3). Il existait peu de données, préalablement à notre étude, sur rôle de l'épissage dans la régulation de la voie RAS/MAPK, à l'exception du gène H-RAS, dont l'épissage alternatif est bien décrit (Guil, de La Iglesia et al. 2003, Cheng, Yaffe et al. 2006, Barbier, Dutertre et al. 2007). Par contre, le phénomène inverse, c'est-à-dire la régulation de l'épissage alternatif par la voie MAPK, est amplement documenté (Weg-Remers, Ponta et al. 2001, Matter, Herrlich et al. 2002, Shin and Manley 2004, Carlson, Chouinard et al. 2011, Courcelles, Fremin et al. 2013) si bien qu'on aura donc plutôt tendance à considérer la voie MAPK comme régulant l'épissage plutôt que l'inverse. Notre crible identifie pourtant une grande quantité de facteurs d'épissage comme régulateurs de la signalisation RAS et les analyses subséquentes montrent un effet spécifique sur l'épissage de *mapk* pour une bonne part d'entre eux.

1.4.1.1 Épissage alternatif et épissage constitutif

La majorité des facteurs d'épissage identifiés dans le crible sont des composantes de la machinerie d'épissage constitutive, c'est-à-dire qu'ils font partie du spliceosome ou de ses facteurs associés. Cette deuxième constatation est également inattendue puisque l'on n'associe généralement pas une fonction spécifique à ces facteurs constitutifs. Le spliceosome et ses facteurs associés sont en effet impliqués dans l'épissage de l'ensemble des introns et il paraîtrait difficile de leur attribuer une fonction spécifique sur l'épissage d'un transcrit particulier. On aurait plutôt tendance à attribuer ce rôle aux facteurs d'épissage alternatifs, tels que les facteurs riches en sérine/arginine (les « facteurs SR ») et « heterogeneous nuclear ribonucleoproteins » (« hnRNP ») (Hertel 2008, House and Lynch 2008, Long and Caceres 2009).

De plus en plus, la distinction entre les facteurs d'épissage alternatifs et constitutifs devient difficile à établir clairement. En effet, le nombre grandissant de facteurs d'épissage constitutifs se voyant attribuer un rôle au sein de l'épissage alternatif contribue à brouiller la frontière entre ces deux ensembles de composantes (Park, Parisky et al. 2004, Hastings, Allemand et al. 2007, Saltzman, Pan et al. 2011, Papasaikas, Tejedor et al. 2014). D'autre part, la présence de plusieurs facteurs SR et hnRNP associés au spliceosome dans les études de caractérisation protéomique des différents états du spliceosome participe également à ce problème (Chen, Moore et al. 2007, Bessonov, Anokhina et al. 2008, Fabrizio, Dannenberg et al. 2009, Herold, Will et al. 2009).

La découverte révélant que certains facteurs constitutifs ont un impact spécifique sur l'épissage alternatif de *mapk* s'inscrit donc dans la lignée d'un ensemble grandissant de données remettant en question la distinction entre les facteurs constitutifs et alternatifs.

1.4.1.2 Exons constitutifs et alternatifs

De manière générale, les exons, tout comme les facteurs d'épissage, sont regroupés en deux catégories : constitutifs et alternatifs. Un exon alternatif est tout simplement un exon pour lequel la modulation par l'épissage alternatif a été démontrée. Il peut présenter des sites d'épissage à consensus plus faibles, ou alors des sites de liaison de facteurs d'épissage

alternatif qui lui confèrent une sensibilité à la modulation par épissage alternatif. À l'inverse, un exon constitutif serait un exon toujours inclus dans le transcrit. Ces exons présentent typiquement des sites d'épissage au consensus plus fort et leur épissage n'est pas régulé par l'épissage alternatif. Le même problème se pose ici pour la distinction entre constitutif et alternatif. D'abord, on peut concevoir qu'il est difficile de prouver qu'un exon constitutif n'est jamais épissé alternativement. D'autre part, l'émergence du séquençage à haut débit et l'augmentation de la profondeur de couverture des études de transcriptome a permis d'augmenter considérablement le nombre d'événements d'épissage alternatifs décrits, soit jusqu'à 61 % des gènes de drosophile et au-delà de 90 % des gènes humains (Graveley, Brooks et al. 2010, Keren, Lev-Maor et al. 2010, Chen 2014). Aussi, la définition d'un exon classé en tant que constitutif/alternatif peut également changer selon l'ensemble de données utilisé (Chen 2014).

1.4.2 Le complexe EJC

Le complexe EJC fait l'objet de l'article au chapitre 2 et est l'un des facteurs les plus intéressants identifiés dans notre crible ARNi RAS^{V12}. L'EJC est constitué d'un noyau de trois facteurs centraux : eIF4AIII, Mago et Tsu/Y14. On inclue parfois BTZ à la liste des facteurs centraux, malgré le fait que ce dernier est surtout associé à la forme cytoplasmique de l'EJC. Plusieurs facteurs, dits facteurs de l'EJC périphérique, s'associent aux facteurs centraux et sont associés à l'une ou l'autre des différentes fonctions du complexe. Ces facteurs périphériques incluent des facteurs d'épissage alternatif SR tels RnpS1, Acinus et SRm160 ainsi que d'autres facteurs associés à ses fonctions post-transcriptionnelles tels que SKAR et Upf3 (Gehring, Neu-Yilik et al. 2003, Ma, Yoon et al. 2008, Lee, Choe et al. 2009). Une liste plus exhaustive des composantes périphériques ainsi qu'une description plus complète des divers rôles post traductionnels de l'EJC sont comprises dans les articles de revue suivants (Tange, Nott et al. 2004, Le Hir and Andersen 2008, Le Hir and Seraphin 2008).

En outre, l'EJC est déposé légèrement en amont (20-24 nt) de la jonction formée entre deux exons épissés (Le Hir, Izaurralde et al. 2000, Le Hir, Moore et al. 2000). Cette déposition est dépendante de l'épissage et du recrutement d'eIF4AIII par le spliceosome (Gehring,

Lamprinaki et al. 2009, Alexandrov, Colognori et al. 2012, Barbosa, Haque et al. 2012, Steckelberg, Boehm et al. 2012). Les principaux rôles de l'EJC sont tous reliés à la régulation de différentes étapes de la biologie des ARNm épissés. Ces fonctions incluent un rôle dans la dégradation des ARNm non-sens (« non-sense mediated decay » ou « NMD »), un processus de contrôle de qualité des ARNm (Lykke-Andersen, Shu et al. 2001, Gatfield and Izaurralde 2002, Gehring, Neu-Yilik et al. 2003, Palacios, Gatfield et al. 2004, Gehring, Kunz et al. 2005), dans la régulation de l'export nucléaire des ARNm (Le Hir, Gatfield et al. 2001), dans la régulation de la localisation des ARNm (Hachet and Ephrussi 2001, Le Hir, Gatfield et al. 2001, Palacios, Gatfield et al. 2004, Zimyanin, Belaya et al. 2008) et dans le contrôle de la traduction des ARNm (Wiegand, Lu et al. 2003, Nott, Le Hir et al. 2004, Diem, Chan et al. 2007, Ma, Yoon et al. 2008, Lee, Choe et al. 2009).

Bien que l'association de l'EJC à divers facteurs d'épissage et au spliceosome soit bien répertoriée, il n'y avait pas de données publiées quant à l'implication des facteurs centraux de l'EJC dans l'épissage préalable à la parution de nos travaux ((Ashton-Beaucage, Udell et al. 2010); chapitre 2), publiés conjointement à ceux du groupe de Jessica Treissman (Roignant and Treisman 2010). À l'inverse, certains résultats publiés jusqu'alors excluaient même une telle fonction de l'EJC (Zhang and Krainer 2007, Crabb, Lam et al. 2010). Depuis, plusieurs nouvelles études sont parues et ont suggéré un rôle pour l'EJC dans l'épissage, lesquelles seront abordées en détail dans la discussion.

Deux études de génomique à large échelle parues plus récemment ont également apporté des éléments intéressants au modèle de fonctionnement de l'EJC. Une première étude de « cross-linking immunoprecipitation » (« CLIP ») combinée au séquençage à haut débit a permis de dresser une carte globale des sites de liaison de l'EJC sur le transcriptome (Sauliere, Murigneux et al. 2012). Cette étude confirme que l'EJC est associé principalement aux ARNm épissés et démontre, entre autres, que l'EJC est enrichi à proximité de certains motifs d'ARN et qu'il peut être présent sur des sites non-canoniques. L'EJC serait également plus fréquemment associé aux transcrits reliés aux catégories fonctionnelles de maturation des ARN et d'organisation des chromosomes. La deuxième étude combine le CLIP-seq à la protéomique pour établir un portrait d'ensemble de l'interactome ARN et protéique de l'EJC (Singh, Kucukural et al. 2012). Tout comme la première étude, elle confirme l'association aux

emplacements canoniques (c'est-à-dire aux jonctions exon-exon), y ajoute des emplacements non-canoniques exoniques et démontre un enrichissement sur les transcrits reliés au traitement des ARN (y joignant aussi un enrichissement sur les transcrits ciblés par le NMD). Le principal intérêt de cette seconde étude est cependant d'y avoir décrit des complexes de ribonucléoprotéines messagères (« mRNP ») d'ordre supérieur formés autour de l'EJC, incluant un bon nombre de facteurs SR. De plus, leurs données suggèrent que ces assemblages sont reliés à la compaction de l'ARN, évoquant un parallèle intéressant avec les nucléosomes.

1.4.2.1 Mécanismes de définition d'exon et de définition d'intron

Un des concepts importants à notre interprétation du rôle de l'EJC dans l'épissage est celui de la définition d'exon. Dès la fin des années 1980, on avait déjà constaté que l'assemblage du spliceosome sur les sites d'épissage impliquait un effet de renforcement mutuel entre les sites d'épissage 5' et 3' d'un même intron. On proposait alors un assemblage du spliceosome de part et d'autre de l'intron puisque les modèles utilisés pour faire ces observations comportaient de courts introns et permettaient d'envisager la formation d'une association entre le complexe U1 (associé au site 5') et le complexe U2 (associé au site 3'). Ce modèle, appelé définition d'intron (« intron definition »), convenait d'ailleurs parfaitement aux eucaryotes inférieurs, tels la drosophile, où l'ensemble des introns sont relativement courts (Figure 1.3 A,B)(Sterner, Carlo et al. 1996, Fox-Walsh, Dou et al. 2005, McGuire, Pearson et al. 2008, Kandul and Noor 2009). Par contre, un problème se posait lorsque l'on appliquait ce modèle aux transcrits des vertébrés où la structure des transcrits présente généralement de courts exons flanqués de longs introns ((Figure 1.3 C,D) et (Sterner, Carlo et al. 1996, Fox-Walsh, Dou et al. 2005, McGuire, Pearson et al. 2008, Roy, Kim et al. 2008)). En effet, il est plus difficile d'envisager que le complexe U1 associé à un site 5' puisse facilement retrouver un complexe U2 situé plusieurs kilobases en aval et s'y lier afin de former une interaction stabilisatrice (Berget 1995, Black 1995, Reed 1996).

Le modèle de la définition d'exon fut d'abord proposé en réponse à ce problème par le groupe de Susan Berget (Robberson, Cote et al. 1990). Une série d'études démontrèrent que, dans le cas des exons courts bordés d'introns plus longs, le site d'épissage 5' de l'intron en

aval était important dans l'épissage de l'intron se trouvant en amont (Nasim, Spears et al. 1990, Robberson, Cote et al. 1990, Talerico and Berget 1990). Or, en réduisant la taille de l'intron ou en augmentant la taille de l'exon, le site 5' en aval perdait sa capacité à favoriser l'épissage en amont ((Robberson, Cote et al. 1990) et Figure 1.4 A,B). De même, la mutation du site 5' en aval avait un effet similaire ((Talerico and Berget 1990) et Figure 1.4 C) alors que renforcer un site 5' pauvre avait l'effet contraire, soit d'augmenter l'efficacité de l'épissage en amont (Grabowski, Nasim et al. 1991, Kuo, Nasim et al. 1991). De plus, l'examen de mutations de sites d'épissage présentes au sein de pathologies a révélé que le défaut d'épissage le plus fréquemment associé dans ce cas n'est pas l'utilisation d'un site alternatif, mais plutôt le saut d'exon en entier (Nakai and Sakamoto 1994). Ces évidences indiquaient donc déjà que les sites d'épissage pouvaient être reconnus en tant que paire exonique et, ensemble, fonctionner afin de promouvoir l'épissage des introns voisins. Ce phénomène suggérait donc une étape préalable d'association des composantes du spliceosome au-travers de l'exon qui permettrait la transition vers une configuration classique de définition d'intron lorsque deux exons définis seraient mis en proximité (Figure 1.3 E). Ce nouveau modèle avait également l'avantage d'expliquer pourquoi la taille moyenne des exons des eucaryotes supérieurs présentait une limite de taille se situant autour de 350-400 nt ((Figure 1.4 D) et (Robberson, Cote et al. 1990)); vraisemblablement, il s'agirait de la limite de taille d'exon permettant l'assemblage du spliceosome au travers de l'exon.

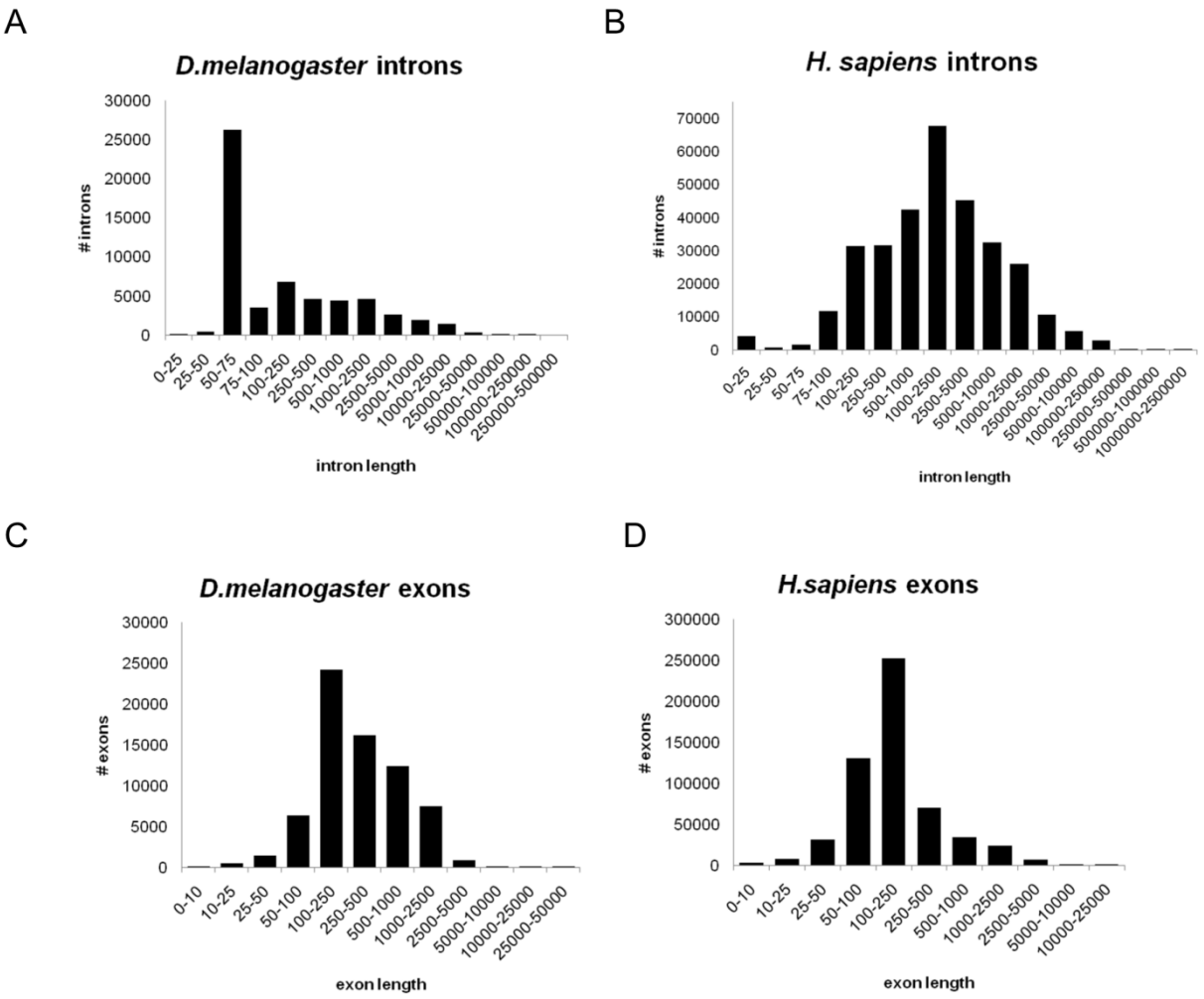


Figure 1.3 Les bases du modèle de définition d'exon

(A-D) Distribution des tailles d'exons et introns chez la drosophile et l'humain (*FlyMine* release 24.1; FB2010_03). Ces distributions sont montrées à titre d'exemple et sont généralement représentatives des distributions des eucaryotes inférieurs et supérieurs.

Figure 1.4 Modèle de définition d'exon

(A-C) Schématisation des premières évidences expérimentales à l'origine du modèle de définition d'exon. (A) Modèle de gène type eukaryote aux longs introns et courts exons. (B) La mutation d'un site d'épissage en aval aura l'effet de compromettre l'épissage de l'intron en amont et cause un saut d'exon plutôt qu'une rétention du second intron. (C) L'allongement de l'exon au-delà de 300 nt ne permet plus un épissage productif et mène au saut de l'exon en entier malgré que ses sites d'épissage restent inchangés. (C) Modèles de définition d'exon et d'intron. La définition d'exon serait le principale modèle de reconnaissance des sites d'épissage chez les vertébrés et implique un pré-assemblage sur les exons. Chez les eucaryotes inférieurs, la définition d'intron serait le modèle prédominant où le spliceosome reconnaît les paires de sites de part et d'autre de l'intron et s'assemble donc directement sur l'intron sans passer par l'étape préalable de la définition d'exon. (D) Étapes de la définition d'exon. Dans le cas d'une structure génique à multiple exons courts bordés de longs introns, le spliceosome reconnaît d'abord les paires de sites d'épissage exoniques et s'assemble sur ceux-ci. Par la suite, la juxtaposition des unités d'exons définis permettra la transition vers une structure classique de pre-spliceosome où les composantes du spliceosome lient des paires de sites introniques.

Les panneaux C et D sont tirés de Berget 1995 avec la permission de l'*American Society for Biochemistry and Molecular Biology*.

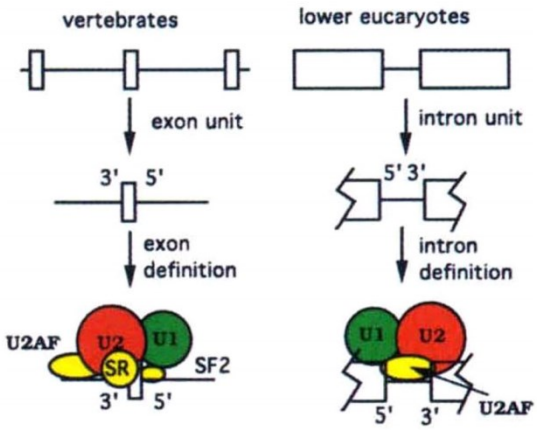
A


5'
3'

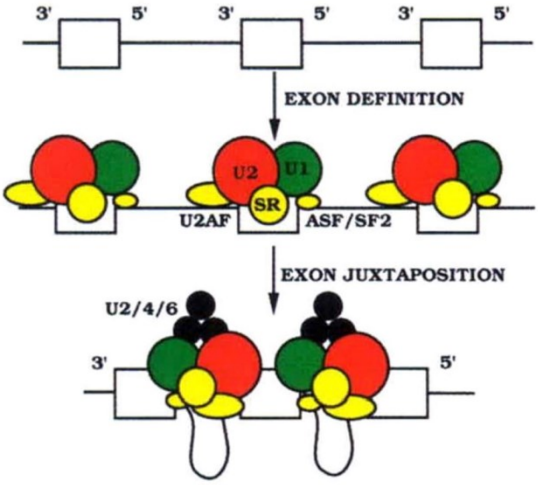
B

C

D


vertebrates
lower eucaryotes
exon unit
intron unit
3' 5'
5' 3'
exon
definition
intron
definition
U2AF
U2
U1
SR
SF2
3'
5'
U1
U2
5'
3'
U2AF

E


3'
5'
3'
5'
3'
5'
EXON DEFINITION
U2
U1
SR
U2AF
ASF/SF2
EXON JUXTAPOSITION
U2/4/6
3'
5'

Des travaux subséquents fournirent des évidences additionnelles à l'appui du modèle de définition d'exon, notamment en délimitant plus précisément la taille à laquelle la reconnaissance des sites d'épissage en paire devenait problématique, soit au-delà de 250 nt pour la définition d'intron et au-delà de 500 nt pour la définition d'exons (Sterner, Carlo et al. 1996, Fox-Walsh, Dou et al. 2005). De plus, chez la drosophile, il fut démontré que les exons correspondant aux critères de définition d'exon avaient 90 fois plus de chance d'être impliqués dans l'épissage alternatif (Fox-Walsh, Dou et al. 2005). D'autres études montrèrent également un impact de la longueur des introns dans des processus d'épissage alternatifs spécifiques (Bell, Cowper et al. 1998, Kandul and Noor 2009). Des facteurs d'épissage constitutifs, Prp5, SRp54, Rsd1 et SF3b, permettant le contact des complexes U1 et U2 au travers de l'intron furent également identifiés (Kennedy, Kramer et al. 1998, Xu, Newnham et al. 2004). D'autre part, la « polypyrimidine tract binding protein » (« PPT ») fut identifiée comme l'un des facteurs permettant la transition entre la définition d'exon et la définition d'intron (Sharma, Falick et al. 2005, Sharma, Kohlstaedt et al. 2008). Il fut également récemment démontré que le complexe de définition d'exon contenait les facteurs du tri-snRNP (U4, U5 et U6), des composantes du complexe spliceosomal B (dit le complexe spliceosomal pré-catalytique). Or, lors de l'association de ce complexe avec un site d'épissage 5' en amont, le spliceosome pouvait alors passer directement au complexe B canonique en sautant l'étape d'un complexe A intronique (Schneider, Will et al. 2010). La transition d'un complexe exonique apparenté au complexe B directement à un complexe B intronique canonique constitue un nouveau mode d'assemblage du spliceosome.

1.4.3 *Usp47* et les facteurs reliés à la dégradation de type « N-end rule »

Le troisième article présenté dans cet ouvrage porte sur *Usp47*, une déubiquitinase d'abord identifiée dans le crible RAS^{V12} décrit au chapitre 3. Tout comme une grande partie des facteurs identifiés dans ce crible, la déplétion d'*Usp47* était associée à une baisse de l'expression spécifique des niveaux de MAPK. Cependant, contrairement aux autres facteurs, *Usp47* n'avait aucun impact sur les niveaux ou l'épissage du transcrit *mapk*. Le chapitre 4 décrit la caractérisation de la régulation post-traductionnelle de MAPK par USP47.

Une seconde partie de l'article porte sur la stratégie de criblage d'interaction génétique par ARNi employée par la suite pour identifier des facteurs du système ubiquitine-protéasome reliés fonctionnellement à *Usp47*. Un facteur E2 (enzyme de conjugaison de l'ubiquitine) et deux E3 (ubiquitine ligase) sont identifiés à l'aide de cette stratégie. Ces trois facteurs sont reliés à un processus de dégradation protéasomique appelé « N-end rule » et qui se caractérise par la reconnaissance d'un dégron constitué par les résidus N-terminaux d'une protéine ou peptide. Les implications de ces résultats y sont également discutées.

2 The Exon Junction Complex Controls the Splicing of *mapk* and Other Long Intron-Containing Transcripts in *Drosophila*

Dariel Ashton-Beaucage[1†], Christian M. Udell[1†], Hugo Lavoie[1†], Caroline Baril[1], Martin Lefrançois[1], Pierre Chagnon[1], Patrick Gendron[1], Olivier Caron-Lizotte[1], Éric Bonneil[1], Pierre Thibault[1,3], and Marc Therrien[1,2,*]

[1] Institute for Research in Immunology and Cancer
Laboratory of Intracellular Signaling
Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montreal, Quebec, Canada, H3C 3J7

[2] Département de pathologie et de biologie cellulaire, Université de Montréal

[3] Département de chimie, Université de Montréal

† These authors contributed equally to this work

*** To whom correspondence should be addressed**

tel.: (514) 343-7837

fax.: (514) 343-6843

Contributions de l'auteur

Contributions expérimentales : Identification des facteurs de l'EJC en tant que régulateurs agissant en aval de Ras^{V12} (décrit au chapitre 3). Production de l'ensemble des données fonctionnelles en culture cellulaire (Fig.1 et Fig.S1) et de l'ensemble des données de génétique (Fig.2 et S2). Observation initiale de l'impact de l'EJC sur MAPK en culture cellulaire (Fig.3; version finale produite par UCM) et *in vivo* (Fig.4 et S4; versions finales produites par BC). Production des données de RT-PCR (Fig.5C), suite à une observation initiale de CMU. Élaboration d'un crible par RT-PCR (non-décrit dans l'article) qui a initialement révélé l'impact de l'EJC sur les gènes à long intron, dont *Dbp80*, *lt*, *sxc* et *Teq* (Fig. 7). Conception et analyse (avec HL et TM) de l'expérience de RNAseq (Fig.6, S6) qui a permis d'évaluer de manière plus compréhensive la corrélation avec la longueur des introns.

Contributions au manuscrit : Élaboration et rédaction du manuscrit. Montage des figures.

Contributions détaillées des co-auteurs

UCM a produit les données de la figure 3 suite à une observation initiale de ABD et une optimisation des conditions expérimentales. UCM a produit les données de qPCR (Fig. S1B-C et 5A), de la figure S3 et de la figure S5. UCM a également fait l'observation initiale pour la figure 5C. BC, avec l'assistance de LM, a fourni les données expérimentales de la figure 4B-F et S4, suite à une observation initiale de ABD et a ajouté les données pour les clones $mago^{3}$ (Fig. 4A). BC a également fourni les premières évidences expérimentales de l'impact de l'EJC sur l'épissage (Fig.4B). LH a élaboré les expériences de RNAseq (Fig.6 et S6) et de protéomique (Fig.S7) avec l'assistance de CP, TP, ABD et TM. CP et GP ont assisté LH et ABD à l'élaboration d'expériences de génomique et l'analyse de données bio-informatique. TP, CLO et BE ont assisté LH à l'élaboration et l'analyse d'expériences de protéomique. UCM, HL et TM ont donné leurs commentaires sur l'article et suggéré des modifications. TM a contribué de manière importante à l'élaboration et la supervision des travaux de recherche ainsi qu'à la rédaction de l'article. TM a modifié certains passages.

2.1 Abstract

Signaling pathways are controlled by a vast array of post-translational mechanisms. By contrast, little is known regarding the mechanisms that regulate the expression of their core components. We conducted an RNAi screen in *Drosophila* for factors modulating RAS/MAPK signaling and identified the Exon Junction Complex (EJC) as a key element of this pathway. The EJC binds the exon-exon junctions of mRNAs and thus far, has been linked exclusively to post-splicing events. Here, we report that the EJC is required for proper splicing of mapk transcripts by a mechanism that apparently controls exon definition. Moreover, whole transcriptome and RT-PCR analyses of EJC-depleted cells revealed that the splicing of long intron-containing genes, which includes mapk, is sensitive to EJC activity. These results identify a role for the EJC in the splicing of a subset of transcripts and suggest that RAS/MAPK signaling depends on the regulation of MAPK levels by the EJC.

2.2 Introduction

The RAS/MAPK signaling pathway is linked to a wide range of cellular processes that include proliferation, differentiation and survival (Kolch, 2005). Its critical role in oncogenesis and various developmental disorders has also been recognized early on and abundantly studied (Schubbert et al., 2007; Zebisch et al., 2007). The pathway is minimally defined by the small GTPase RAS and a core of three kinases (RAF, MEK and ERK/MAPK) that transmit signals incoming mostly from plasma membrane receptors. Upon its activation by a Guanine nucleotide Exchange Factor (GEF), GTP-loaded RAS triggers RAF activation, which in turn phosphorylates MEK. Activated MEK then phosphorylates and activates MAPK. MAPK then phosphorylates a specific set of substrates that will elicit distinct cell responses (McKay and Morrison, 2007; Turjanski et al., 2007).

Within recent years, it has become apparent that the RAS/MAPK pathway is not merely a linear route involving only four proteins as commonly referred to, but in fact corresponds to a larger protein network that comprises multiple regulatory characteristics such as feedback loops (Dougherty et al., 2005), compartmentalization (Ebisuya et al., 2005;

McKay and Morrison, 2007) and crosstalk with other pathways (Hurlbut et al., 2007). However, many of the features inherent to this network are still poorly understood and its protein composition remains incompletely defined. Despite their diversity, these regulatory mechanisms typically affect the enzymatic activity of the core components or the accessibility to their substrates (Kolch, 2005; McKay and Morrison, 2007). Although much less documented, mounting evidence suggests that the pathway is also influenced by other means that affect the steady-state levels of core components. For example, the expression of LET-60/RAS in *C. elegans* or mammalian RAS proteins has been shown to be modulated by the *let-7* family miRNAs (Johnson et al., 2005). More recently, Pumilio mRNA binding proteins were found to restrict MAPK activity by attenuating the expression of the *C. elegans* MAPK, Mpk-1, as well as ERK2/MAPK1 and p38α/MAPK14 in human ES cells, which occurred via the binding of specific sites in the 3'UTR of their respective transcripts (Lee et al., 2007). Moreover, the LARP-1 RNA-binding protein has also been shown to control the mRNA abundance of RAS/MAPK pathway components, including those of MAPK in the *C. elegans* germ line (Nykamp et al., 2008). Thus, it appears that the expression of the core components themselves can be the target of specific mechanisms that in turn impact signaling output.

In order to identify novel factors that modulate RAS/MAPK signaling, we conducted a genome-wide RNAi screen in *Drosophila* S2 cells. Interestingly, the screen led to the identification of the Exon Junction Complex (EJC) as a critical factor that specifically impacts MAPK protein levels in *Drosophila*. The EJC has recently emerged as an important determinant in many aspects of mRNA regulation (reviewed in (Tange et al., 2004)). This protein complex is deposited on mRNAs 20-24 nucleotides upstream of exon-exon junctions in a splicing-dependent manner and serves as a tether for other peripheral factors that are associated with various activities. Three main functions for the EJC have been described to date, namely, mRNA nuclear export (Le Hir et al., 2001b; Palacios et al., 2004), translational enhancement (Diem et al., 2007; Nott et al., 2004; Wiegand et al., 2003) and nonsense mediated mRNA decay (NMD) (Gatfield et al., 2003; Gehring et al., 2003; Lykke-Andersen et al., 2001; Palacios et al., 2004), which is an mRNA surveillance mechanism that eliminates faulty transcripts harboring premature stop codons (Chang et al., 2007). A few species-specific exceptions exist to these roles; for instance, the EJC does not mediate NMD in *Drosophila*, *C.*

elegans or *S. pombe* (Gatfield et al., 2003; Longman et al., 2007; Wen and Brogna, 2010). The EJC also seems to control mRNA subcellular localization in at least one case, where it has been shown to control localization of the *oskar* transcript to the posterior pole of the oocyte (Hachet and Ephrussi, 2004). The EJC is comprised of four core components: EIF4A3, MAGOH, RBM8A/Y14 and MLN51. These are respectively known in flies as EIF4AIII, Mago Nashi (MAGO), Tsunagi (TSU) and Barentsz (BTZ), which we will refer to herein. In contrast to the other EJC core factors, BTZ is predominantly cytoplasmic (Degot, 2004; Macchi et al., 2003; Palacios et al., 2004) and is thought to be added to the complex upon mRNA export from the nucleus, thus forming the cytoplasmic EJC. It is this configuration of the EJC that is involved in mRNA localization and NMD (Palacios et al., 2004). The nuclear EJC (EIF4AIII/MAGO/TSU) can modulate mRNA export, yet it is not required for export of all spliced transcripts (Gatfield and Izaurralde, 2002). Intriguingly, despite the fact that the EJC associates with many splicing co-factors such as RNPS1 and SRM160 and also interacts with the core spliceosome (Bessonov et al., 2008; Herold et al., 2009; Merz et al., 2007), it has not been implicated in pre-mRNA splicing *per se* and is therefore considered to be exclusively dedicated to post-splicing events.

Here, we investigated the function of the EJC in RAS/MAPK signaling. We show that the EJC acts downstream of MEK in the RAS/MAPK cascade and is required for MAPK expression. Unexpectedly, we found that the EJC controls the splicing of *mapk* pre-mRNAs as its disruption leads to exon skipping events. Interestingly, RNA-Seq and RT-PCR data revealed that the altered mRNA expression profile provoked by EJC disruption was not a general phenomenon, but instead correlated with intron length, whereby pre-mRNAs bearing very large introns, such as those of *mapk*, were more sensitive to EJC depletion. Given that exon exclusion was observed for several transcripts in addition to *mapk*, we propose that the EJC is involved in exon definition of large intron-containing genes. Together, our findings reveal a critical factor that controls RAS/MAPK signaling in *Drosophila* and unveils a role for the EJC in the splicing regulation of long intron-containing genes.

2.3 Results

2.3.1 EJC Components Modulate RAS1 Signaling Downstream of MEK

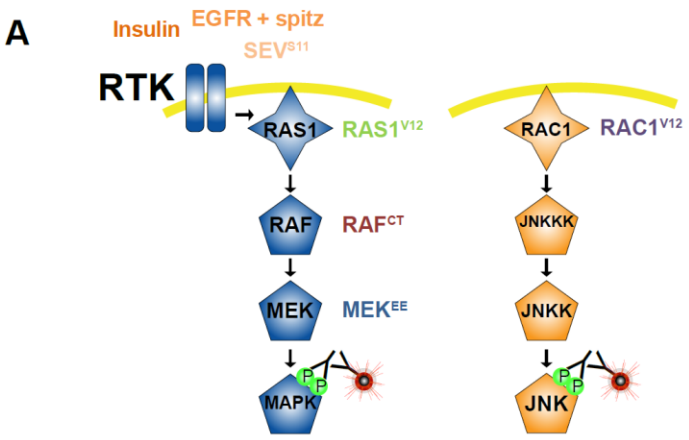
Expression of a constitutively activated form of *Drosophila* RAS1 ($RAS1^{V12}$) in S2 cells leads to sustained activation of endogenous MAPK that is quantitatively measurable by immunohistochemistry (Figure 2.1A and see Experimental Procedures). We used this assay to conduct a genome-wide RNAi screen to identify novel fly genes that modulate signaling through the RAS1/MAPK pathway. Approximately 100 validated hits were identified, of which two-thirds of the known components of the pathway were found including *raf/phl*, *mek*/*Dsor1* and *mapk/rl* (Ashton-Beaucage et al., unpublished data). Interestingly, among the candidate genes, the three components of the nuclear EJC (*eIF4AIII*, *mago* and *tsu*) were identified as suppressors considering that the average phospho-(p)MAPK signal per cell was found to decrease consistently upon knockdown of their respective transcripts (Figures 2.1B, S1A and S1B). This indicates that the EJC positively contributes to pathway activity at a step between RAS1 and MAPK. Intriguingly, although the peripheral EJC component *RnpS1* was also found as a suppressor (Figures S2.1A and S2.1C), *btz*, the fourth core component that is part of the cytoplasmic EJC, was not identified in the RNAi screen. Subsequent experiments confirmed that knockdown of *btz* did not reduce pMAPK levels (Figure S2.1A), even though the dsRNA targeting *btz* did efficiently deplete its transcript levels (Figure S2.1B) as did the dsRNAs targeting *mago*, *tsu*, *eIF4AIII* and *RnpS1* (Figures S2.1B and S2.1C).

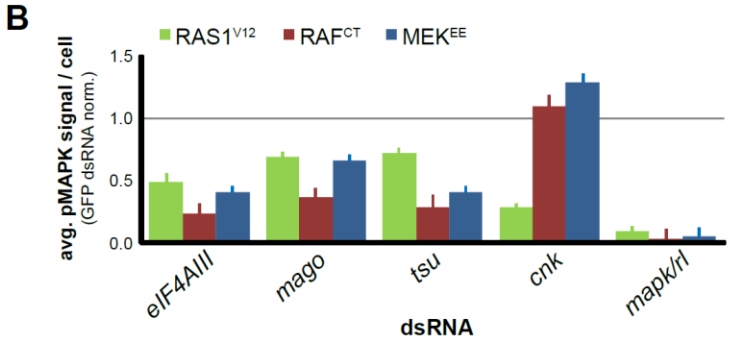
In addition to suppressing $RAS1^{V12}$, depletion of the nuclear EJC also reduced pMAPK induced by upstream RTKs including EGFR, SEV^{S11} and the insulin receptor (InR), but did not suppress $RAC1^{V12}$-induced JNK signaling (Figures 2.1A and 2.1C). This indicates that the EJC is required for RTK/RAS1/MAPK signaling, but is not required in the context of the RAC1/JNK pathway. Next, we assayed the position of the EJC's effect with regards to the kinase cascade downstream of RAS1. Epistasis experiments conducted using several cell lines expressing constitutively active forms of RAF and MEK also showed a consistent decrease in pMAPK levels upon EJC knockdown (Figures 2.1A, B and S2.1D). In contrast, *cnk*, a regulator of RAF activation (Douziech et al., 2003), suppressed $RAS1^{V12}$, but not active RAF

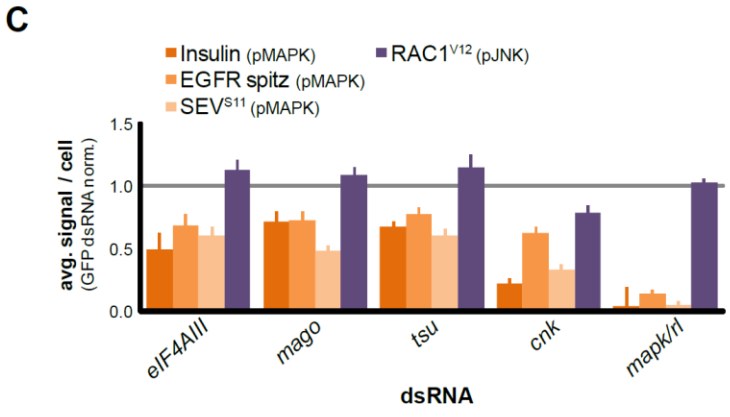
or MEK (Figure 2.1B). Thus, these results place the requirement for EJC activity at a step downstream of MEK.

To determine whether the ability of the EJC to modulate RAS1/MAPK signaling was not restricted to S2 cells, but also controls this pathway *in vivo*, we conducted genetic interaction experiments using *eIF4AIII*, *mago* and *tsu* loss-of-function alleles. RTK-induced RAS1/MAPK activity is required for neuronal photoreceptor and cone cell differentiation during *Drosophila* eye development (Wassarman et al., 1995). Expression of $Ras1^{V12}$ under the control of the eye specific *sev* promoter/enhancer regulatory sequences produces extra photoreceptor cells, which causes a characteristic rough eye phenotype (Fortini et al., 1992) that has been extensively used in genetic interaction assays (Karim et al., 1996; Therrien et al., 2000). As shown in Figure 2.2A, this rough eye phenotype was dominantly suppressed by heterozygous mutations in the *eIF4AIII*, *mago* or *tsu* genes. Four additional genetic interaction assays also tied EJC activity to RAS1/MAPK signaling *in vivo*. First, the lethality associated with raf/phl^{12} hemizygous mutant males was enhanced in *eIF4AIII*, *mago* or *tsu* heterozygous backgrounds (Figure S2.2A). Secondly, wing vein deletions and lethality caused by a hemizygous mutation in *csw*, which encodes a Shp-2 phosphatase homolog (Perkins et al., 1992), were enhanced by a $mago^{1}$ heterozygous allele (Figures 2.2B, S2.2B and S2.2C). Thirdly, extra wing vein material produced by a constitutively active *Egfr* allele was dominantly suppressed by *eIF4AIII*, *mago* or *tsu* mutant alleles (Figures 2.2C, S2.2D – S2.2F). Finally, the rough eye and wing vein deletion phenotypes of homozygous $mapk/rl^{1}$, which corresponds to a hypomorphic allele of the *mapk/rl* gene, strongly increased in severity in heterozygous *mago* mutant backgrounds (Figures 2.2D and 2.2E). Collectively, these data provide compelling evidence that the EJC is required for RAS1/MAPK signaling in *Drosophila*.

Figure 2.1 EJC components functionally impact RAS1/MAPK signaling downstream of MEK (A) Schematic representations of the MAPK and JNK pathways. The stimuli used in the different cell-based functional assays are positioned within these pathway models (colored text). (B) Epistasis analysis of the nuclear EJC components. Quantitative microscopy-based measurement of the average pMAPK intensity/cell for quadruplicate *pMet-Ras1*^{V12} S2 cell samples and duplicate *pMet-raf*^{CT} and *pMet-mek*^{EE} S2 cell line samples treated with the indicated dsRNAs. dsRNA probes targeting *mapk/rl* and *cnk* were added as controls of the epistasis strategy; *cnk* functions between RAS1 and RAF and MAPK functions downstream of MEK. (C) Values from RTK-based MAPK activation assays (orange) and RAC1^{V12}-based activation of JNK (purple) are shown for S2 cells treated with the indicated dsRNAs. The average pMAPK intensity/cell is shown for duplicate samples of insulin treated S2 cells, *pMet-Egfr* S2 cells treated with Spitz (EGFR + SPI) and heat-shock induced *pHS-Sev*^{S11} cells. The average pJNK intensity/cell is shown for duplicate samples of *pMet-Rac1*^{V12} S2 cells.


A
Insulin
EGFR + spitz
SEVS11
RTK
RAS1
RAS1V12
RAC1
RAC1V12
RAF
RAFCT
JNKKK
MEK
MEKEE
JNKK
MAPK
JNK
P
P


B
RAS1V12
RAFCT
MEKEE
avg. pMAPK signal / cell
(GFP dsRNA norm.)
1.5
1.0
0.5
0.0
eIF4AIII
mago
tsu
cnk
mapk/rl
dsRNA


C
Insulin (pMAPK)
RAC1V12 (pJNK)
EGFR spitz (pMAPK)
SEVS11 (pMAPK)
avg. signal / cell
(GFP dsRNA norm.)
1.5
1.0
0.5
0.0
eIF4AIII
mago
tsu
cnk
mapk/rl
dsRNA

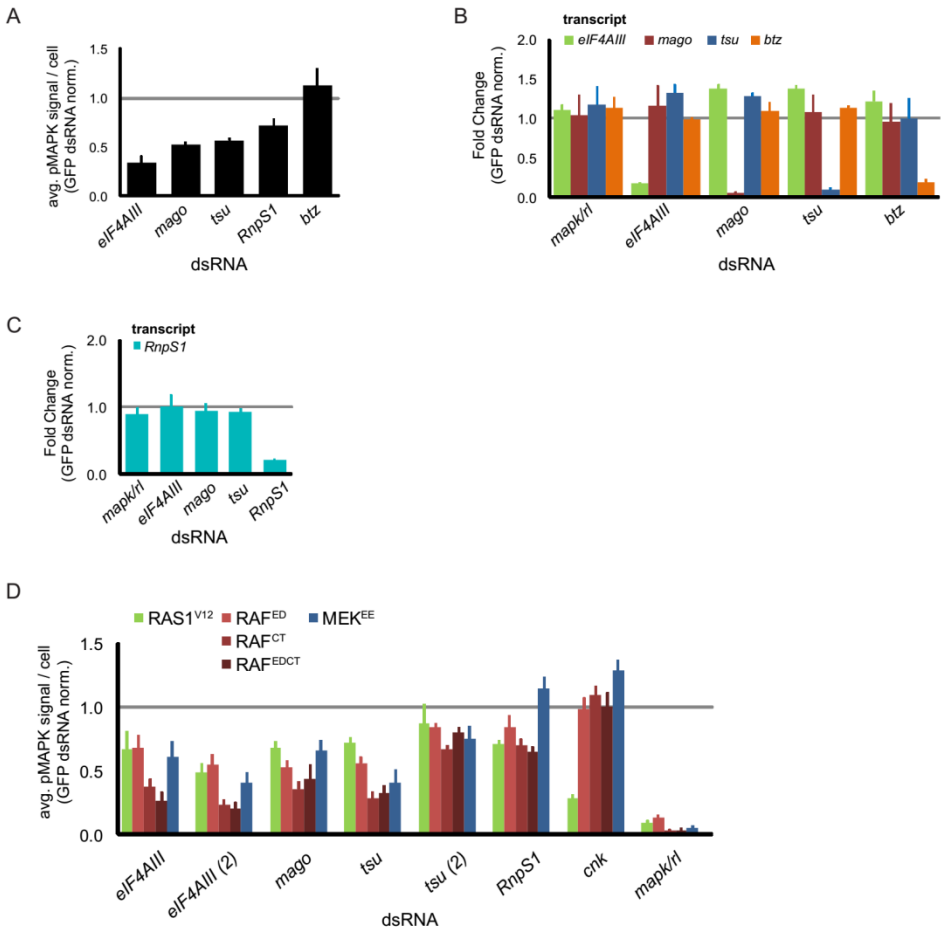


Figure S2.1 Nuclear EJC components functionally impact RTK/RAS1/MAPK signaling downstream of MEK.

(A) Average pMAPK intensity/cell of duplicate *pMet-Ras1*^{V12} cell samples treated with the indicated dsRNAs. (B) and (C) RNAi-mediated knockdown of EJC components was confirmed by RT-qPCR. The expression levels of the indicated transcripts were normalized to the amount present in cells treated with *GFP* dsRNA (D) Additional data for RAF and MEK-based MAPK activation assays. Average pMAPK intensity/cell values shown are for quadruplicate samples of *pMet-Ras1*^{V12} cells and duplicate samples of *pMet-raf*^{ED}, *pMet-raf*^{CT}, *pMet-raf*^{EDCT}, and *pMet-mek*^{EE} cells. Values for two distinct *tsu* and *eIF4AIII* dsRNA probes are shown in (D).

Figure 2.2 Components of the EJC genetically interact with RAS1/MAPK pathway components

(A) *Ras1*^{V12} rough eye phenotype is dominantly suppressed by heterozygous mutations in EJC components. Fly eyes of the indicated genotypes were scored for severity of the rough eye phenotype (n>30) and classified as strong, medium or weak. All of the EJC alleles scored as medium to strong suppressors. Representative ESEM microscopy images are presented here. The *mapk/rl*$^{S-352}$ allele is used as a positive control. Anterior is to the right. (B) Enhancement of *csw*^{lf} wing vein phenotypes by *mago*^{1} loss-of-function allele. A wild-type wing is shown as a reference (left panel). The characteristic deletion of the distal end of the L5 (and sometimes L4 and L2) wing veins typically observed in *csw*^{lf} hemizygous males is shown (central panel) as well as the enhanced phenotype observed in *mago*^{1} heterozygous background (right panel). Quantification of the enhancement is presented in Figure S2.2C. (C) Suppression of *Egfr* gain of function wing vein phenotype. Typical extra wing vein material produced by the *Egfr*^{Elp} gain of function allele near the distal end of the L2 vein is shown in the left panel (arrow). Suppression of the phenotype by the heterozygous *mapk/rl*^{1} allele (positive control; central panel) or by heterozygous *mago*^{1} (right panel) is also shown and quantification of the effect is presented in Figure S2.2E. (D - E) Enhancement of *mapk/rl*^{1} homozygous rough eye and wing vein phenotypes by two distinct *mago* alleles. Vein deletions (L4 mainly, arrow) and smaller/rougher eyes are observed in heterozygous mutant backgrounds for *mago*^{1} and *mago*^{3}.

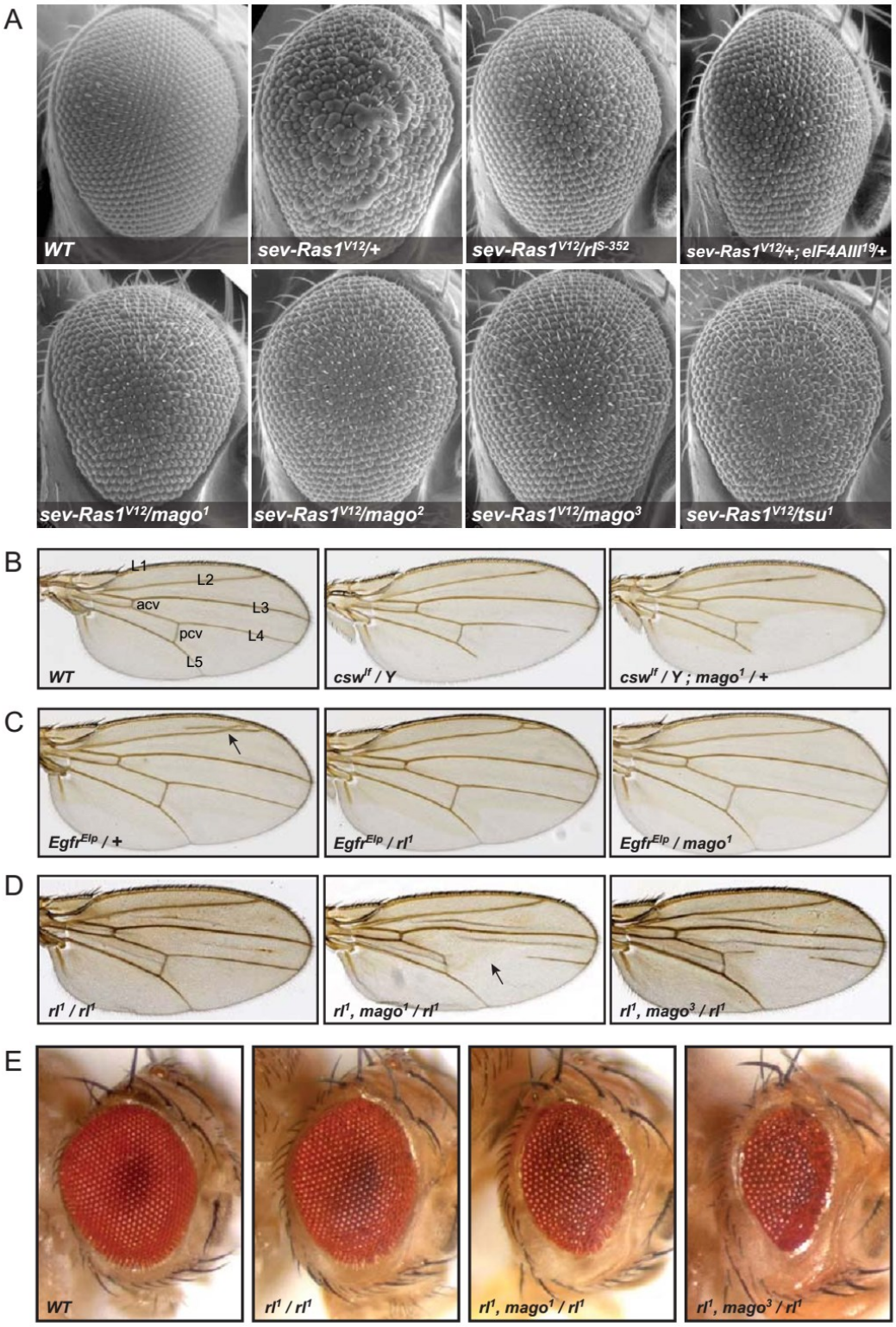
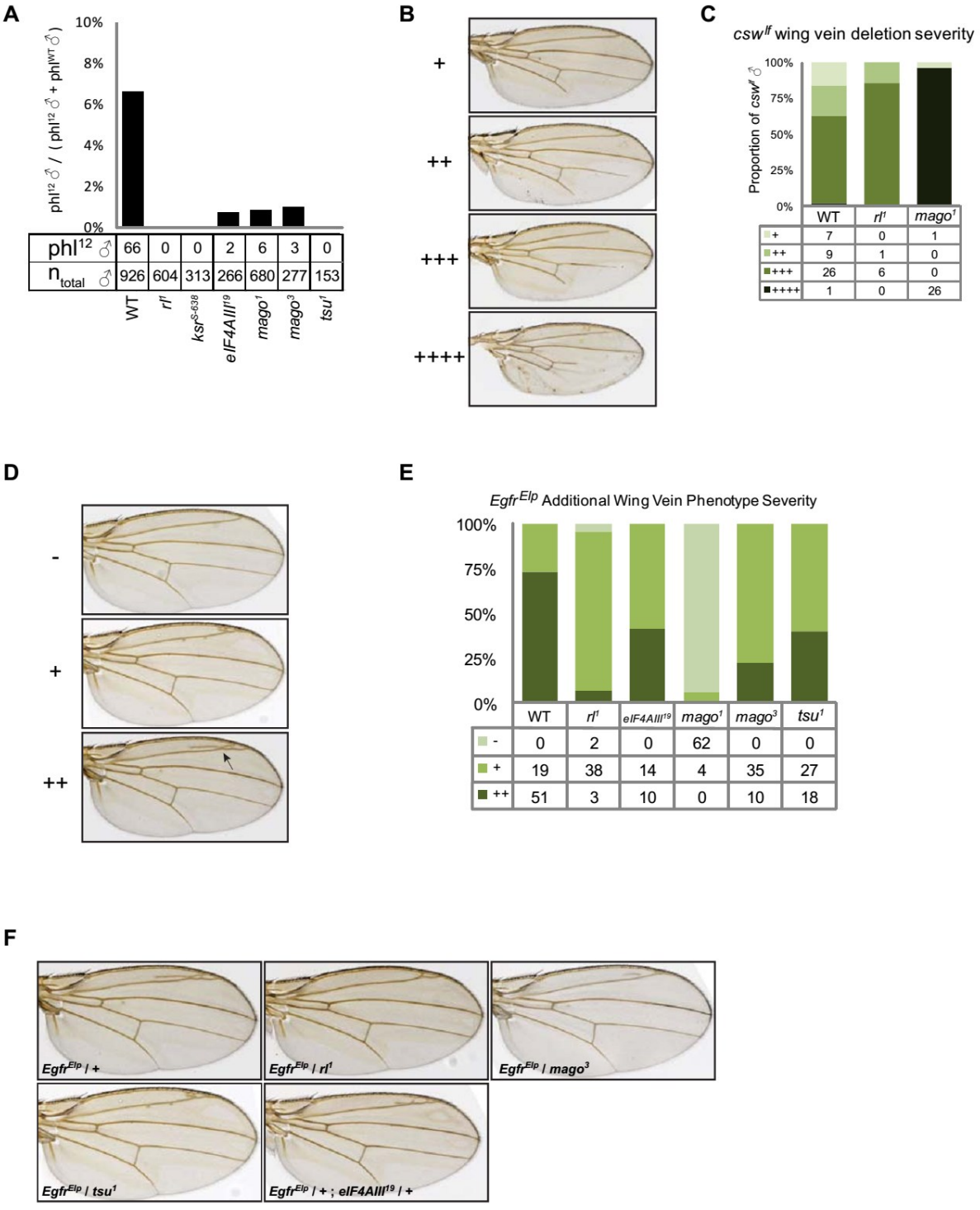

A
WT
sev-Ras1V12/+
sev-Ras1V12/rlS-352
sev-Ras1V12/+; eIF4AIII19/+
sev-Ras1V12/mago1
sev-Ras1V12/mago2
sev-Ras1V12/mago3
sev-Ras1V12/tsu1
B
L1
L2
acv
L3
pcv
L4
L5
WT
cswlf / Y
cswlf / Y ; mago1 / +
C
EgfrElp / +
EgfrElp / rl1
EgfrElp / mago1
D
rl1 / rl1
rl1, mago1 / rl1
rl1, mago3 / rl1
E
WT
rl1 / rl1
rl1, mago1 / rl1
rl1, mago3 / rl1

Figure S2.2. Additional genetic interaction data for EJC components.

Quantitative data for *raf/phl*^{12}, *csw*^{lf} and *Egfr*^{Elp} genetic interactions. (A) Enhancement of *raf/phl*^{12} viability. The proportion of *raf/phl*^{12} hemizygous males to the total male population is shown for flies of an otherwise wild-type background (WT) or flies heterozygous for the listed alleles. Loss of function alleles of *mapk/rl* (*rl*^{1}) or *ksr* (*ksr*$^{S\text{-}638}$) are shown as positive controls. (B - C) *csw*^{lf} genetic interactions. (B) Examples of phenotypic severity used for scoring the *csw*^{lf} wing vein deletion phenotypes. (C) Proportions of wings from *csw*^{lf} males with the indicated deletion severity for the *mapk/rl*^{1} and *mago*^{1} alleles. The *mago*^{3} allele significantly increased the lethality of *csw*^{lf} males and thus no wings could be scored. (D - F) *Egfr*^{Elp} genetic interactions. (D) Examples of *Egfr*^{Elp}-dependent additional wing vein severity used for phenotype scoring. Typical extra wing vein material produced by the *Egfr*^{Elp} gain of function allele near the distal end of the L2 vein is shown in the left panel (arrow). (E) Proportion of wings with the indicated phenotypic severity for *Egfr*^{Elp} flies with an otherwise wild-type background (WT) or heterozygous for the indicated alleles. The total number of wings scored is indicated below each allele. (F) Extra wing vein material produced by the *Egfr* gain of function allele and modulations of the wing vein phenotypes observed for the indicated genotypes.



phl12 ♂	66	0	0	2	6	3	0
n total ♂	926	604	313	266	680	277	153

	WT	rl1	mago1	mago3
+	7	0	1	
++	9	1	0	
+++	26	6	0	
++++	1	0	26	

	WT	rl1	eIF4AIII19	mago1	mago3	tsu1
-	0	2	0	62	0	0
+	19	38	14	4	35	27
++	51	3	10	0	10	18

2.3.2 The EJC Regulates *Drosophila* MAPK Protein Expression

In light of the EJC function in mRNA translation (Tange et al., 2004), we investigated the effects of its depletion on the protein levels of selected RAS1/MAPK pathway components. Strikingly, western blot analysis revealed that MAPK levels were significantly reduced upon the knockdown of EJC components (Figure 2.3). This effect appeared to be specific to MAPK as disruption of the EJC did not have an appreciable effect on the protein levels of other pathway components, such as RAS1, RAF, MEK, or CNK, on other kinases such as AKT and JNK, or on Actin (Figure 2.3). Moreover, silver-stained protein lysates from EJC-depleted S2 cells did not show any significant difference compared to control cells, thus providing additional evidence that EJC depletion does not lead to a global reduction of protein levels in S2 cells (Figure S2.3A and see below). Interestingly, we noticed that the levels of ectopically-produced MAPK from a *mapk* cDNA were insensitive to EJC depletion, which indicates that the effect of the EJC on MAPK does not occur following its translation (Figure S2.3B). Together, these findings provide an explanation as to why EJC depletion consistently reduced endogenous pMAPK levels in the various RAS1/MAPK signaling assays presented above and suggest that the EJC is specifically involved in MAPK protein expression at a step that precedes its translation or that occurs concomitantly.

To verify whether the EJC controls MAPK levels *in vivo* as it does in S2 cells, we generated *mago* homozygous mutant clones during eye development and stained third instar eye-antennal discs using an anti-MAPK antibody. In agreement with the cell culture data, both MAPK and pMAPK levels were strongly reduced in *mago* mutant clones, whereas CNK was unaffected (Figures 2.4A, D and S2.4A). Similar results were obtained when *mago*, *tsu* or *mapk* itself were clonally depleted by RNAi (Figures 2.4B, C, S2.4B). Consistent with reduced MAPK levels, EJC depletion during eye development significantly impeded photoreceptor and cone cell differentiation (Figures 2.4E and 2.4F). Together, these results indicate that the EJC plays a critical role in establishing proper MAPK protein levels *in vivo*.

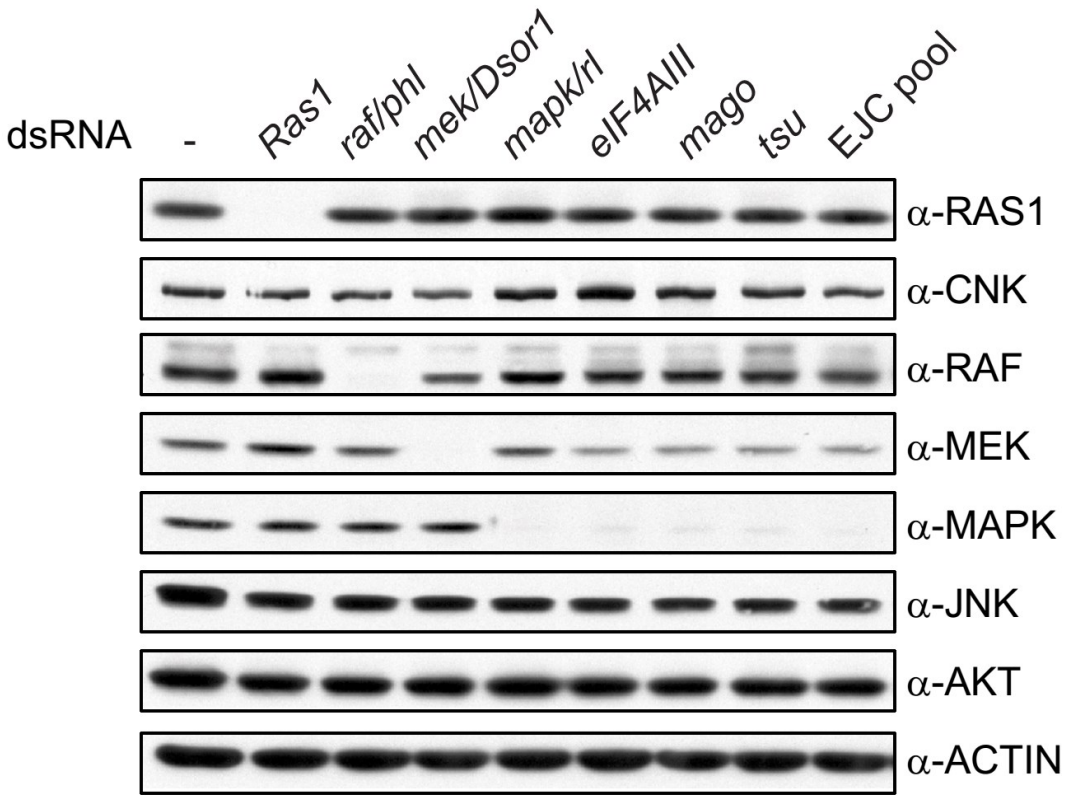


Figure 2.3 Depletion of EJC components specifically reduces MAPK levels in S2 cells

S2 cells were treated with the indicated dsRNA. EJC pool refers to a pool of three dsRNA targeting *eIF4AIII*, *mago*, and *tsu*. Protein lysates were subjected to SDS-PAGE and immunoblotted with antibodies indicated at the right of each panel.

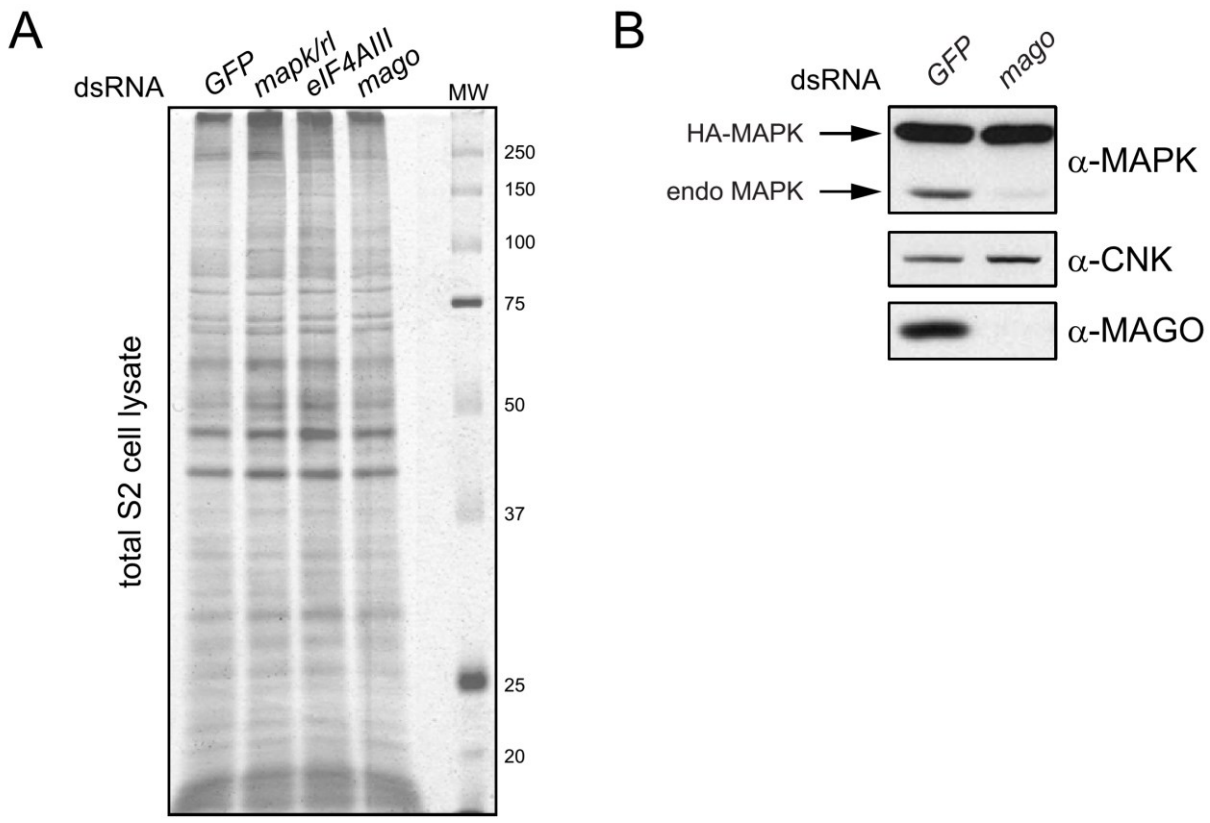
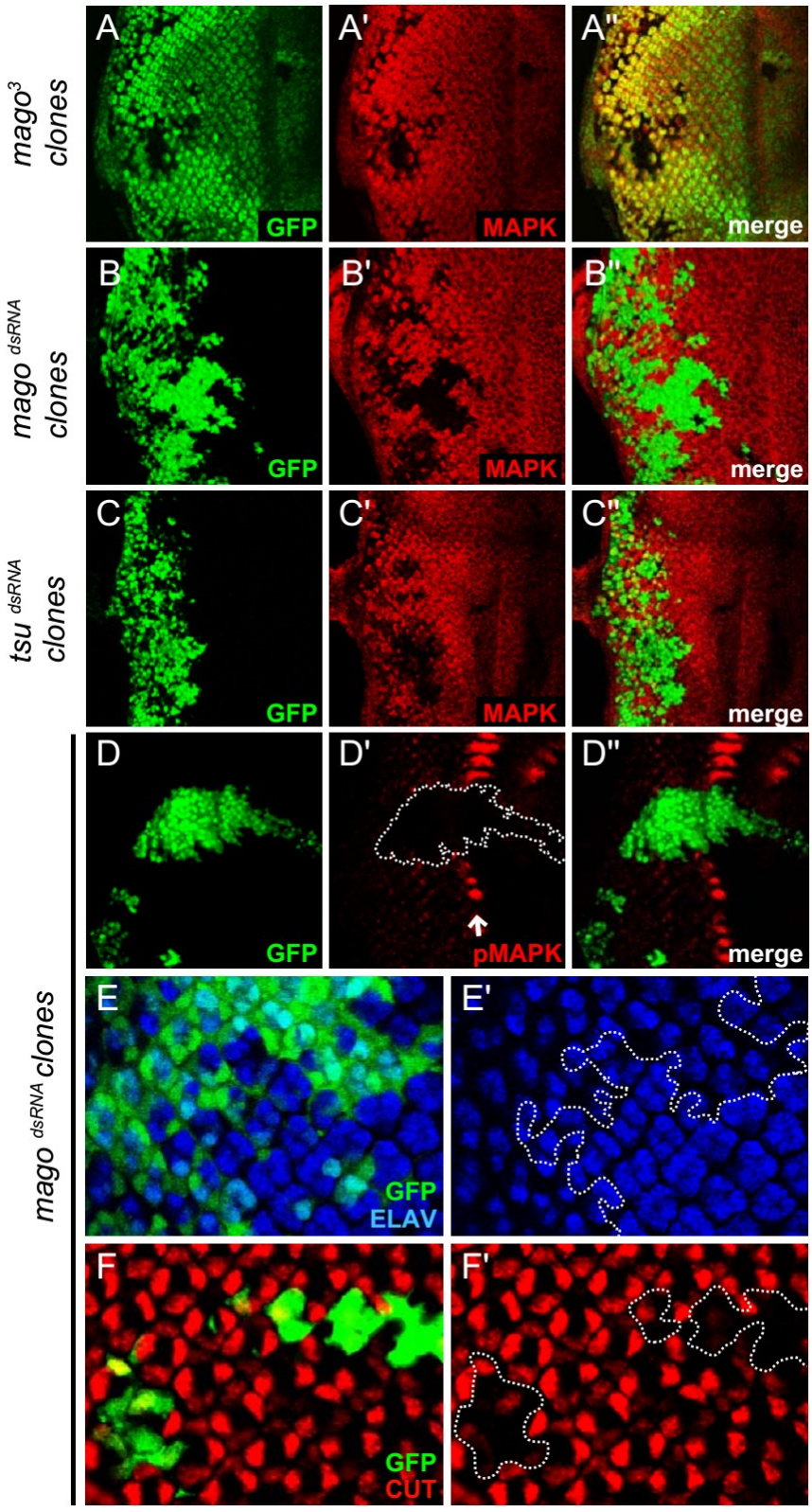


Figure S2.3 EJC knockdown does not have widespread effects on the S2 cell proteome

(A) Protein extracts were prepared from S2 cells treated with the indicated dsRNA, resolved by SDS-PAGE, and visualized by silver staining. (B) S2 cells stably expressing an HA epitope-tagged *mapk/rl* cDNA driven by the actin promoter were treated with *GFP* or *mago* dsRNA. Protein extracts were resolved by SDS-PAGE and immunoblotted with the antibodies indicated at the right of each panel. CNK served as a loading control.

Figure 2.4 Depletion of EJC components reduces MAPK levels *in vivo* and impacts photoreceptor cell differentiation

(A - C) MAPK staining of third instar larval eye discs. (A) MAPK levels are under detection levels in *mago*^{3} mutant clones, which are marked by the absence of GFP fluorescence. MAPK is also absent in GFP-positive clones expressing *mago* (B) and *tsu* (C) RNAi. (D - F) *mago* dsRNA-expressing clones marked by GFP fluorescence impede typical RTK-dependent signaling markers in third instar larval eye discs. (D) Phospho-MAPK (pMAPK) signal that is normally induced by EGFR signaling in cell clusters along the morphogenetic furrow (arrow) and in a limited set of cells posterior to it, is abrogated in *mago* dsRNA-expressing cells. (E - F) *mago* dsRNA-expressing cells are impaired in their ability to differentiate as neurons or cone cells as expression of the neuronal marker ELAV (E) or the cone marker CUT (F) is eliminated or significantly reduced in those cells. In all panels, anterior is to the right.


mago3 clones
mago dsRNA clones
tsu dsRNA clones
mago dsRNA clones
A
A'
A''
GFP
MAPK
merge
B
B'
B''
GFP
MAPK
merge
C
C'
C''
GFP
MAPK
merge
D
D'
D''
GFP
pMAPK
merge
E
E'
GFP
ELAV
F
F'
GFP
CUT

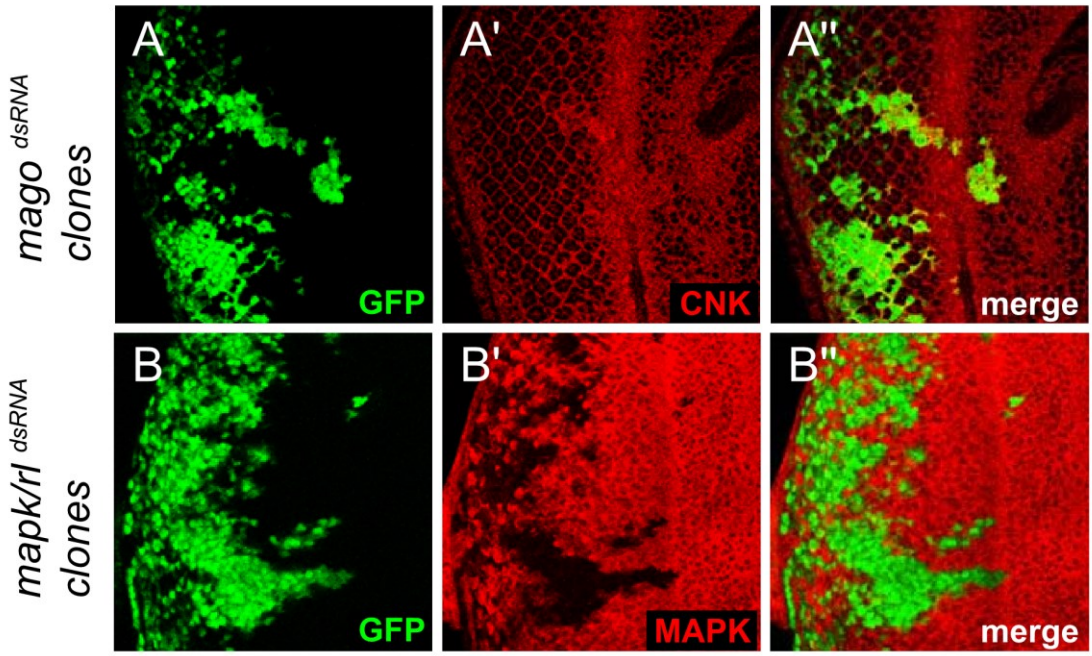


Figure S2.4 *mago* RNAi does not impact CNK levels *in vivo*

(A - B) CNK and MAPK staining of third instar larval eye discs. (A) CNK expression is not altered in GFP-positive clones expressing *mago* RNAi. (B) MAPK levels are effectively depleted in *mapk/rl* RNAi clones.

2.3.3 Loss of EJC Activity Induces Exon Skipping Events During *mapk* pre-mRNA Splicing

Given the role that the EJC plays in several aspects of mRNA maturation, we investigated the effect of EJC disruption on *mapk* transcript levels. An initial RT-qPCR evaluation of *mapk* transcript levels showed a reproducible 1.5 to 2-fold decrease upon EIF4AIII and MAGO depletion, whereas, consistent with the results from the functional assays, *mapk* transcript levels did not decrease upon knockdown of *btz* (Figure 2.5A). While the reduction in the amount of *mapk* mRNA was reproducible, it did not appear sufficient to account for the decrease in protein expression. Therefore, we examined whether *mapk* mRNA translation might be affected upon EJC depletion by comparing its incorporation into polysomes between control and MAGO-depleted cells. While the expected 1.5 to 2-fold decrease in *mapk* mRNA levels in MAGO-depleted cells was observed, no difference in polysome incorporation efficiency could be detected (Figures S2.5A-C). This suggests that the EJC does not directly impact *mapk* translation, nor significantly influence processes that would lead to a decrease in translation efficiency. Also, the fact that *btz* was not involved in this function further implied that the effect occurred at an earlier step in *mapk* expression involving the nuclear EJC heterotrimer.

In order to confirm the RT-qPCR results, we surveyed *mapk* mRNA by northern blot and RT-PCR analyses. In both cases, a reduction in the amount of the expected *mapk* transcripts was apparent upon EJC depletion, but surprisingly, truncated *mapk* transcripts were also visible in these samples (Figures 2.5B, C). Specifically, transcript size distribution, as assayed by northern blot, showed that the two previously described *mapk* transcripts of 1.7 and 2.7 kb (Berghella and Dimitri, 1996) were detected in control S2 cells (Figure 2.5B, lane 1). Both transcripts were significantly reduced upon EJC depletion and at least two novel transcripts of 1.1 and 2.1 kb could be detected (Figure 2.5B, lanes 3-4). Notably, the reduction in total *mapk* mRNA levels observed in EJC-depleted samples was comparable to those observed by RT-qPCR. Also, RT-PCR with primers in the 5' and 3' UTRs of *mapk* mRNA, which were expected to amplify the predicted RB, RD, RE and RF isoforms, revealed a significant decrease in the amount of the expected transcripts following EJC depletion (Figure 2.5C, compare lane 1 to lanes 3-6). Similar to the northern blot, several truncated species

could be observed in EJC-depleted samples by RT-PCR. In agreement with our previous observations, this effect on the *mapk* transcripts was observed for *eIF4AIII*, *mago*, *tsu* and *RnpS1* RNAi (Figure 2.5C, lanes 3-6), but not in *btz* dsRNA-treated cells (Figure 2.5C, lane 7). Finally, in support for the specificity of this effect, *jnk/bsk* full-length transcripts did not show any discernable change in abundance nor size when amplified by RT-PCR from EJC-depleted samples compared to control (Figure 2.5C, compare lane 8 to lanes 10-13).

In order to ascertain the nature of these truncated transcripts, we cloned and sequenced several RT-PCR products from both control cells and cells depleted of *eIF4AIII*. Remarkably, this analysis revealed that the shorter transcripts could be accounted for by a series of exon skipping events (Figure 2.5D). The vast majority of these events involved the precise skipping of multiple consecutive exons with the first exon being joined to different downstream exons at their expected 5' splice acceptor site. In agreement with the RT-PCR data, the most commonly observed transcripts in the *eIF4AIII*-depleted samples were those lacking exons II to III, II to IV and II to V (Figures 2.5C and D). A compilation of all skipping events revealed that exon II was the most frequently skipped exon (96% of the cases), followed by exon III (89%) (Figure S2.5E). As exons II and III include the start codon and a substantial portion of the kinase domain of MAPK, their absence in most *mapk* transcripts from EJC-depleted cells explains why EJC depletion leads to severe reduction in MAPK protein levels. Surprisingly, exon skipping was also observed, but to a limited extent, in control S2 cells (Figures 2.5C and D) and could also be observed in RNA samples prepared from flies at all developmental stages (Figure S2.5D). This indicates that exon skipping during the splicing of the *mapk* pre-mRNA occurs naturally at a certain rate. Taken as a whole, these findings suggest that the nuclear EJC plays a critical role in exon inclusion during *mapk* pre-mRNA splicing.

Figure 2.5 The EJC is required for faithful splicing of *mapk/rl* pre-mRNA

Total RNA was prepared from S2 cells treated with the indicated dsRNA. (A) *mapk/rl* transcript levels were assayed by RT-qPCR as described in the methods. Results are the average of three independent experiments (+/-SD). Statistical significance was evaluated using a two-tailed Student's t-test (* denotes *p*-value <0.01). (B) Poly(A)$^+$ mRNA was subjected to Northern blot with probes for *mapk/rl* or *jnk/bsk* mRNA (* indicate shifted *mapk/rl* species). (C) RT-PCR with primers in the 5' and 3' UTR of *mapk/rl* or *jnk/bsk* mRNA. The positions of species depicted in (D) are indicated at the right of the *mapk/rl* panel. (D) *mapk/rl* RT-PCR products were cloned and sequenced. The schematic represents the most abundant species observed (for example, ΔII-III indicates a transcript lacking exons II and III). Red bars indicate the position of primers used for RT-PCR. The coding sequence is coloured in blue and the exons are numbered I through VIII based on the RE transcript. The size of each species is shown in parentheses and lengths of the introns are indicated. Long introns (>250bp) are not drawn to scale and are represented by dashed lines. The predicted RF isoform was not observed and is therefore not represented.

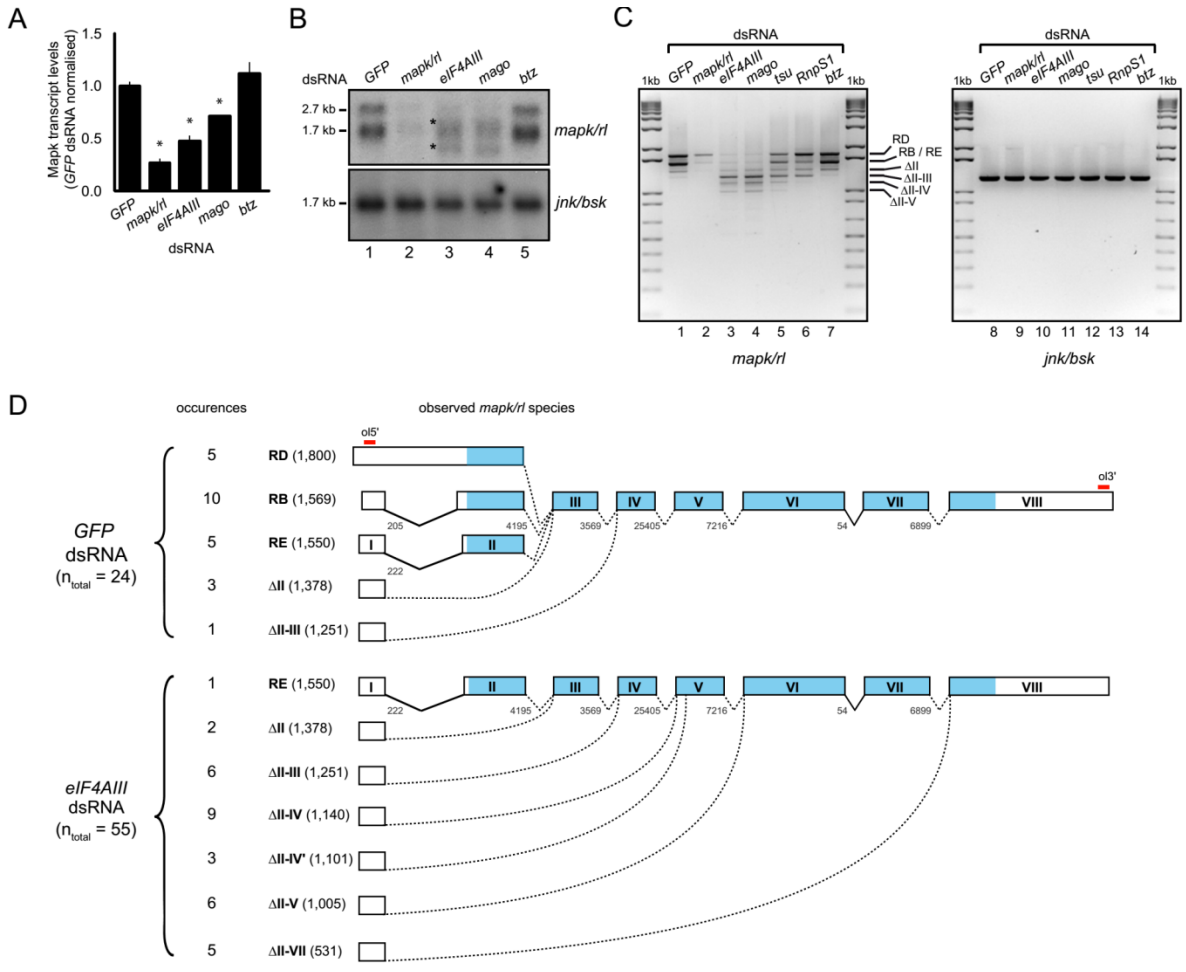
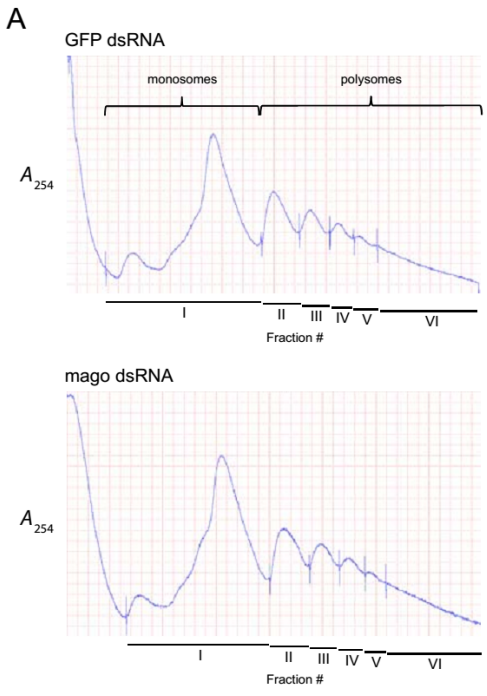
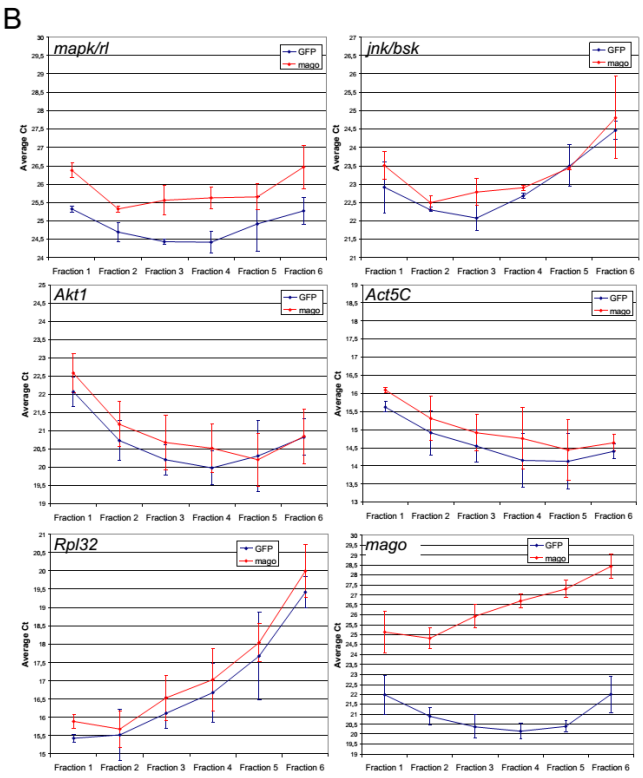
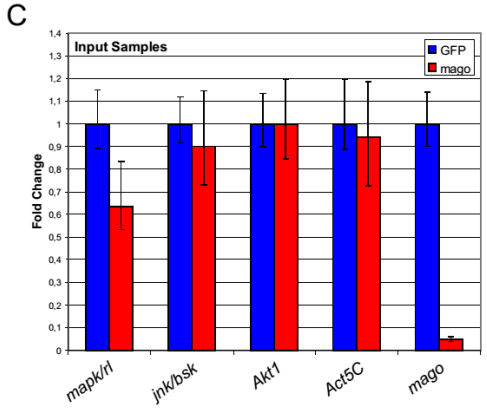

A
Mapk transcript levels (GFP dsRNA normalised)
1.5
1.0
0.5
0.0
GFP
mapk/rl
eIF4AIII
mago
btz
dsRNA
B
dsRNA
GFP
mapk/rl
eIF4AIII
mago
btz
2.7 kb
1.7 kb
mapk/rl
1.7 kb
jnk/bsk
1 2 3 4 5
C
dsRNA
1kb
GFP
mapk/rl
eIF4AIII
mago
tsu
RnpS1
btz
1kb
RD
RB / RE
ΔII
ΔII-III
ΔII-IV
ΔII-V
1 2 3 4 5 6 7
mapk/rl
dsRNA
1kb
GFP
mapk/rl
eIF4AIII
mago
tsu
RnpS1
btz
1kb
8 9 10 11 12 13 14
jnk/bsk
D
occurences
observed mapk/rl species
GFP dsRNA (n_total = 24)
5 RD (1,800)
10 RB (1,569)
5 RE (1,550)
3 ΔII (1,378)
1 ΔII-III (1,251)
ol5'
ol3'
I
II
III
IV
V
VI
VII
VIII
205
4195
222
3569
25405
7216
54
6899
eIF4AIII dsRNA (n_total = 55)
1 RE (1,550)
2 ΔII (1,378)
6 ΔII-III (1,251)
9 ΔII-IV (1,140)
3 ΔII-IV' (1,101)
6 ΔII-V (1,005)
5 ΔII-VII (531)

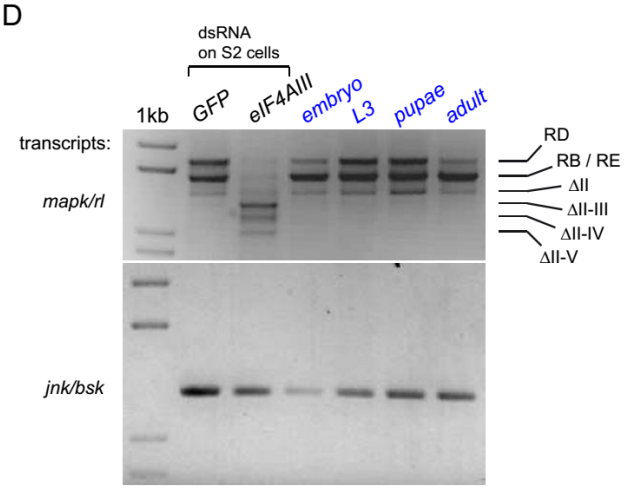
Figure S2.5 The EJC does not impact translation of *mapk/rl* observed in S2 cells

S2 cells were treated with *GFP* or *mago* dsRNA and post-nuclear supernatants were fractionated on sucrose density gradients as described in the supplemental methods. (A) Gradient profiles measuring the absorbance at 254 nm. The positions of monosomes (40S, 60S, and 80S) and polysomes are indicated. The fractions collected are indicated by vertical event marks on the profiles and labelled underneath. (B) Total RNA was prepared from each fraction and assayed by RT-qPCR as described in the methods. The graphs show C_t values plotted versus fraction number for the indicated mRNA. Results are the average of two independent experiments. The curves for *mapk/rl* are of the same shape, indicating no defect in polysome loading, but the *mapk/rl* curve for *mago* RNAi is upshifted relative to *GFP* RNAi due to there being less *mapk/rl* mRNA. (C) Total RNA from input samples was assayed by RT-qPCR. The histogram shows the fold change for each mRNA normalized to that present in the *GFP* dsRNA sample. The average fold change was calculated from the pooled data of two independent experiments (three technical replicates each). The error bars indicate the range of fold changes calculated from six technical replicates using *RpL32* as the reference gene. (D) RNA was prepared from S2 cells or fly tissue from the indicated developmental stage and *mapk/rl* or *jnk/bsk* transcripts were amplified by RT-PCR as described in the methods. PCR products were resolved on agarose/EtBr gels. The positions of *mapk/rl* transcripts are indicated at the right. (E) Percentage of *mapk/rl* transcripts lacking the indicated exon following depletion of eIF4AIII.









E

exon #	deleted (%)
II	96
III	89
IV	53
V	56
VI	25
VII	34

2.3.4 Intron Length Determines the EJC's Effect on Splicing

Given that the EJC is thought to generally bind to intron-containing mRNAs (Le Hir et al., 2000), we thought it unlikely that the splicing defect we observed was unique to the *mapk* mRNA. However, the observation that EJC depletion had no effect on the expression of several other proteins encoded by intron-containing genes indicated that splicing in general does not require the EJC. Thus, in order to evaluate the full spectrum of the EJC's effects on the transcriptome of S2 cells, we conducted a transcriptional profiling experiment using massively parallel RNA sequencing (RNA-Seq) technology on a SOLiD platform (Applied Biosystems) that yielded highly reproducible transcript coverage across samples (Figure S2.6A). The abundance of 6760 expressed transcripts (5X coverage threshold) assayed in this manner revealed that *mapk* was among the 50 genes whose expression was most reduced upon EJC depletion (Figure 2.6 and S2.6B and Table S2.I). Therefore, the *mapk* transcript is highly sensitive to perturbation of the EJC in comparison with most other genes expressed in S2 cells.

We next sought to determine what characteristics of *mapk* determined this sensitivity to the EJC. We noted three characteristics that distinguish *mapk* from other fly genes. First, *mapk* is localized in the pericentromeric constitutive heterochromatin. Second, the first exon which we found to be aberrantly joined to other downstream exons has a non-canonical splice donor site (the first three intronic residues, GUU, do not match the GUR consensus (Mount et al., 1992)). Finally, *mapk* has introns of larger than average size, including one 25kb intron, which is amongst the longest introns in the genome; less than 1% of *D. melanogaster* introns are larger than 25kb (Figure S2.6C). Of these three characteristics, intron length emerged as the determining factor of the EJC's effect. While localization in heterochromatin or the presence of atypical splice sites did somewhat correlate with sensitivity to EJC depletion, this effect was dependent on the presence of long introns. Indeed, genes with these characteristics, but with introns shorter than 250bp were not found to be sensitive to the EJC's effect (Table S2.II). On the other hand, transcripts with long introns were more likely to be sensitive to EJC depletion (Table S2.II and Figures 2.6A, B and S2.6B). Additionally, this sensitivity appeared to increase as a function of intron length in such a manner that a general trend could be observed for the entire population of transcripts after both *eIF4AIII* and *mago* knockdowns (Figure 2.6A, B and S2.6B). This relationship between EJC sensitivity and intron length is

highly significant since, of the 114 genes down-regulated by more than 2-fold after *eIF4AIII* depletion, 81 had an intron larger than 1000bp (p-value = 8.59 x 10^{-21} (calculated with a hypergeometric distribution) (Figure 2.6C)).

Following this observation, we sought to verify whether the fold-changes in transcript levels detected by RNA-Seq were due to splicing defects similar to those observed for the *mapk* transcript. Therefore, we performed RT-PCR on a subset of candidate transcripts which were affected by EJC depletion and also contained long introns. We found that a similar reduction in the levels of the full-length transcript could be observed and, in many cases, truncated products were also detected in EJC-depleted cells (Figure 2.7A), suggesting that splicing was also altered in these cases. Indeed, sequencing of the products for the *PMCA*, *lt*, and *Tequila* genes showed that exon skipping was, in large part, responsible for the appearance of the truncated products (Figure S2.7). In contrast, a set of control transcripts with multiple exons and whose introns were all smaller than 250bp did not show additional truncated isoforms or variation in abundance upon EJC depletion (Figure 2.7B). These results demonstrate a clear link between intron length, splicing and the EJC.

Finally, to complement our comprehensive study of transcript abundance after EJC depletion, we wanted to determine whether protein levels were accordingly modulated upon EJC knockdown in an intron length-dependent manner. For this, we used a non-biased, label-free quantitative proteomics approach to globally assess the abundance of proteins in S2 cells depleted for the EJC (see Supplementary Methods; Figure S2.6D). Of the 6760 expressed transcripts detected by RNA-Seq, we were able to correlate the abundance of 2410 corresponding proteins. Interestingly, no dramatic change was observed in the EJC-depleted proteome and no global trend could be found between protein fold-change and intron length (Figures S2.6E and F). This observation is consistent with western blot analysis of several proteins (Figure 2.3) and with the general aspect of silver-stained protein lysates from EJC-depleted S2 cells (Figure S2.3A). Nonetheless, MAPK, PMCA and SXC were amongst the most significantly down-regulated protein products in both *eIF4AIII* and *mago*-depleted samples compared to *GFP* dsRNA-treated cells (Figures S2.6E-G). This absence of strong correlation between transcript and protein abundance has been known for some times (Gygi et al., 1999) and is attributed to the contrasting stability of mRNAs and proteins, especially in the

context of short-term depletion of the EJC by RNAi. This global analysis of transcript and protein levels shows that most of the effect of EJC depletion is manifested at the level of transcript abundance and splicing and that *mapk* is one of the most sensitive targets of the EJC.

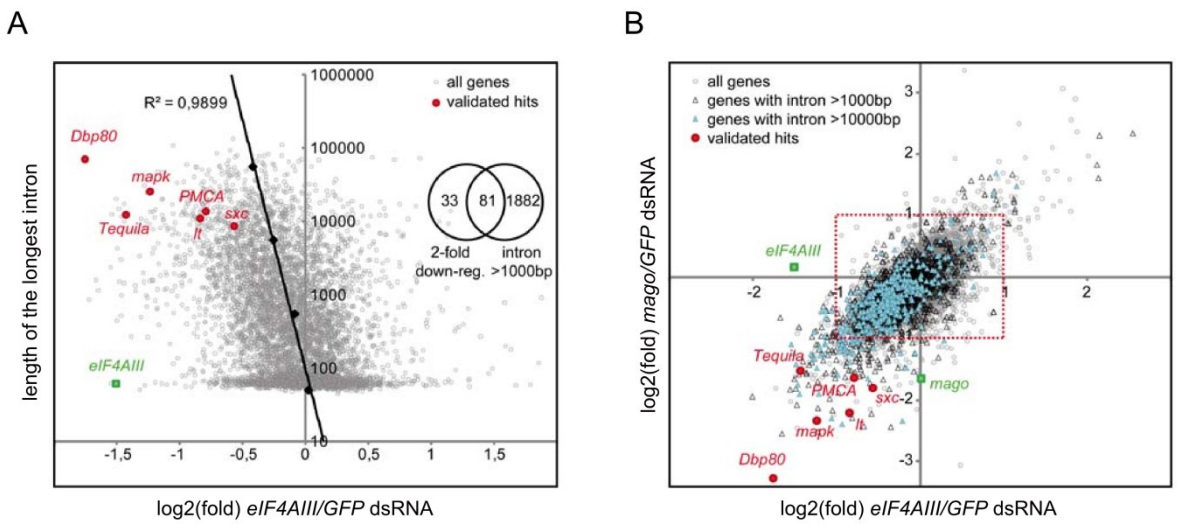


Figure 2.6 The impact of the EJC on transcript levels correlates with intron length

Poly(A)+, rRNA-depleted mRNAs from *eIF4AIII*, *mago* and *GFP* dsRNA-treated S2 cells were subjected to RNA-Seq on the SOLiD sequencing platform. (A) *eIF4AIII* knockdown causes a systematic and gradual down-regulation of transcripts $\log_2$(fold-change) in a manner proportional to the length of the longest intron of the gene. *mapk* is amongst the 50 most down-regulated gene detected by transcriptome analysis. The *lt*, *tequila*, *PMCA*, *sxc* and *Dbp80* transcripts that were subjected to further validation are labeled in red and *eIF4AIII*, in green. The Venn diagram depicts the overlap between the groups of 2-fold down-regulated transcripts and genes containing long introns (>1000 bp). (B) *mago* and *eIF4AIII* depletion yield a reproducible decrease of the *mapk* mRNA abundance together with other long intron-containing transcripts. Transcripts that were further validated are labeled in red. *eIF4AIII* and *mago* mRNAs are depicted in green and served as internal controls.

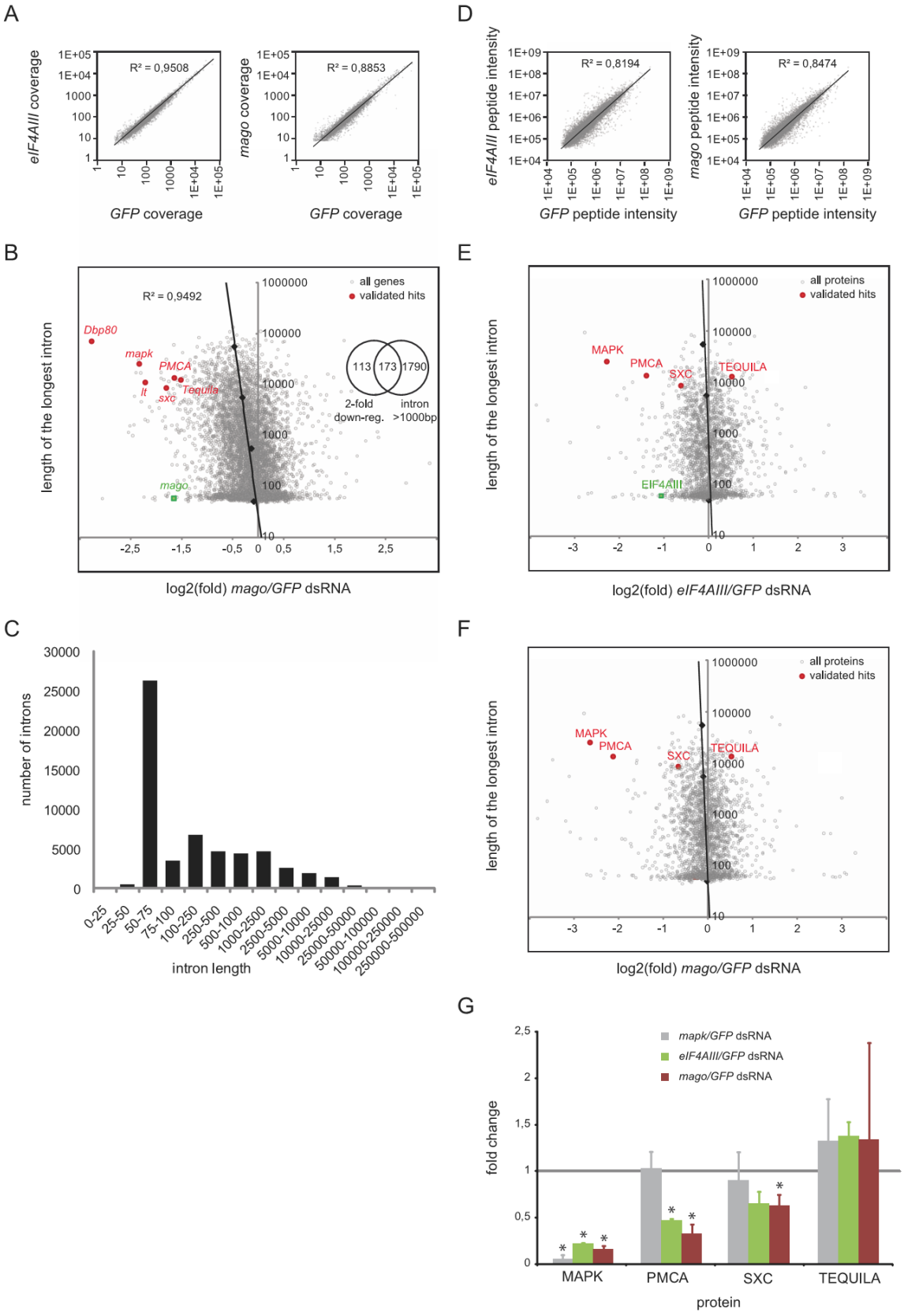

A
eIF4AIII coverage
R² = 0,9508
GFP coverage
mago coverage
R² = 0,8853
GFP coverage
D
eIF4AIII peptide intensity
R² = 0,8194
GFP peptide intensity
mago peptide intensity
R² = 0,8474
GFP peptide intensity
B
length of the longest intron
R² = 0,9492
all genes
validated hits
Dbp80
mapk
PMCA
It
sxc
Tequila
mago
113
173
1790
2-fold down-reg.
intron >1000bp
log2(fold) mago/GFP dsRNA
E
length of the longest intron
all proteins
validated hits
MAPK
PMCA
SXC
TEQUILA
EIF4AIII
log2(fold) eIF4AIII/GFP dsRNA
C
number of introns
intron length
0-25
25-50
50-75
75-100
100-250
250-500
500-1000
1000-2500
2500-5000
5000-10000
10000-25000
25000-50000
50000-100000
100000-250000
250000-500000
F
length of the longest intron
all proteins
validated hits
MAPK
PMCA
SXC
TEQUILA
log2(fold) mago/GFP dsRNA
G
mapk/GFP dsRNA
eIF4AIII/GFP dsRNA
mago/GFP dsRNA
fold change
MAPK
PMCA
SXC
TEQUILA
protein

Figure S2.6 Transcriptome and proteome analyses of EJC-depleted cells.

Poly(A)+, rRNA-depleted mRNAs from *eIF4AIII*, *mago* and *GFP* dsRNA treated S2 cells were subjected to RNA-Seq on the SOLiD sequencing platform. (A) Scatter plots representing the high correlation in transcript coverage between *eIF4AIII* (left panel) and *mago* (right panel) dsRNA-treated samples and *GFP dsRNA*-treated control. (B) *mago* knockdown causes a systematic and gradual down-regulation of transcripts that is proportional to the length of their longest intron. *mapk/rl* is amongst the 50 most down-regulated transcripts detected by transcriptome analysis. The *mago* transcript is labeled in green. The Venn diagram depicts the overlap between the group of 2-fold down-regulated transcripts and genes coding for introns larger than 1000 bp. (C) *Drosophila* intron length distribution extracted using FlyMine (FMv24.1) (Lyne et al., 2007). (D) Total protein extracts from EJC-depleted S2 cells were subjected to label-free proteome evaluation. Scatter plots show the high level of correlation obtained between samples in terms of peptide intensity for *eIF4AIII* (left panel) and *mago* (right panel) compared to the *GFP dsRNA*-treated control. *eIF4AIII* (E) and *mago* (F) knockdowns do not affect global protein levels in a intron length-dependent manner. *mapk* is amongst the 15 most down-regulated proteins detected by quantitative proteomics analysis of EJC-depleted S2 cells. The MAPK, PMCA, SXC and TEQUILA proteins are highlighted in red. EIF4AIII is labeled in green and served as an internal control. (G) Fold-change values for the candidate proteins MAPK, PMCA, SXC and TEQUILA detected in the proteomics dataset. Statistical analysis was conducted using a two-tailed Student's t test to compare the protein abundance obtained in the *GFP* versus the *mapk*, *eIF4AIII* or *mago*-depleted triplicates; asterisks denotes a *p*-value smaller than 0,05.

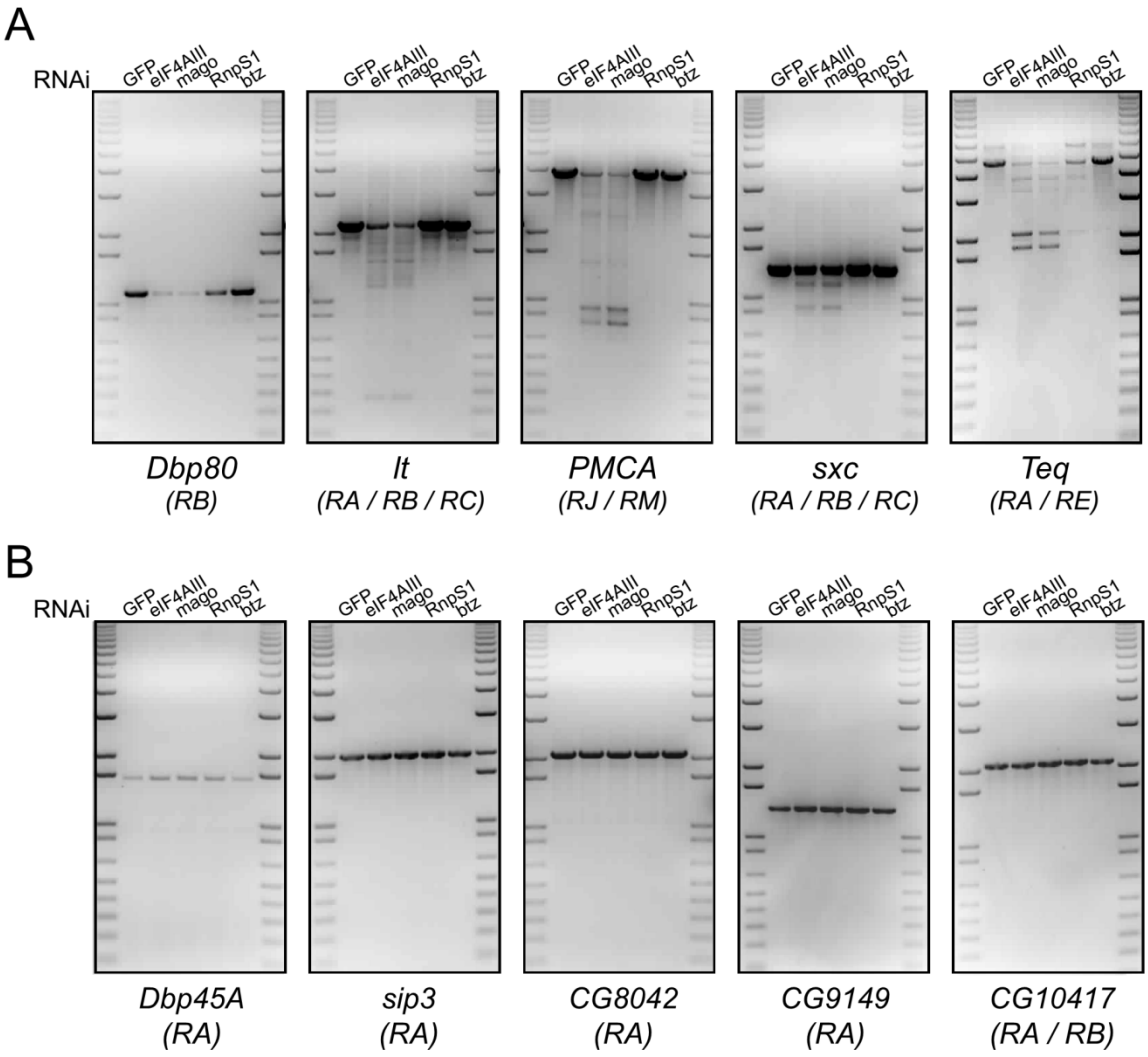


Figure 2.7 Truncated transcripts are observed for other EJC targets

Total RNA was prepared from S2 cells treated with the indicated dsRNA. RT-PCR products for EJC-sensitive transcripts (A) and EJC-insensitive transcripts (B) were resolved on agarose/EtBr gels. Specific transcript isoforms targeted by RT-PCR are indicated in parentheses. Each of the transcripts assayed in A has at least one large intron (>1000bp), while the transcripts in B have at least five exons, but no introns longer than 200 bp.

Figure S2.7 Exon skipping events are also observed in other EJC target genes

RT-PCR products for *lt-RA* (A), *PMCA-RM* (B) and *Tequila-RA* (C) were cloned and sequenced as described in the methods, and are shown schematically. In each panel, the gene models for wild-type transcripts are shown at the top (exons and introns are not drawn to scale). Long introns (>250bp) are represented by dashed lines. The coding sequence is colored in blue and red bars indicate the position of primers used for PCR. Red shaded boxes between exons indicate an intron inclusion event. The identity, size (bp), and number of occurrences of each transcript is indicated on the left.

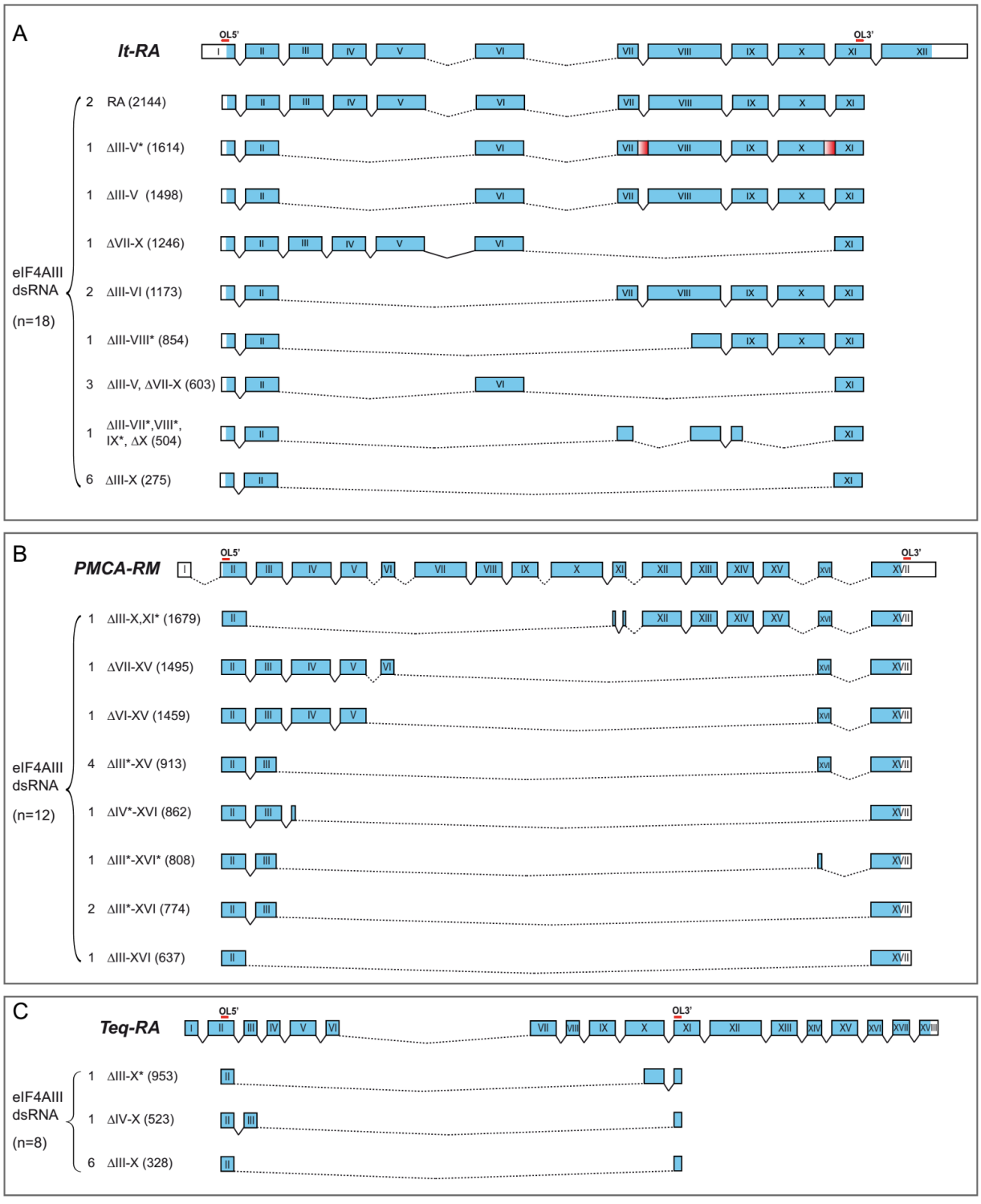

A
It-RA
OL5'
OL3'
eIF4AIII dsRNA (n=18)
2 RA (2144)
1 ΔIII-V* (1614)
1 ΔIII-V (1498)
1 ΔVII-X (1246)
2 ΔIII-VI (1173)
1 ΔIII-VIII* (854)
3 ΔIII-V, ΔVII-X (603)
1 ΔIII-VII*,VIII*, IX*, ΔX (504)
6 ΔIII-X (275)
B
PMCA-RM
eIF4AIII dsRNA (n=12)
1 ΔIII-X,XI* (1679)
1 ΔVII-XV (1495)
1 ΔVI-XV (1459)
4 ΔIII*-XV (913)
1 ΔIV*-XVI (862)
1 ΔIII*-XVI* (808)
2 ΔIII*-XVI (774)
1 ΔIII-XVI (637)
C
Teq-RA
eIF4AIII dsRNA (n=8)
1 ΔIII-X* (953)
1 ΔIV-X (523)
6 ΔIII-X (328)

Table S2.I List of genes whose mRNA expression is down-regulated by at least 2-fold after *eIF4AIII* or *mago* knockdown compared to *GFP* dsRNA treatment

The normalized coverage and $\log_2$(fold-change) obtained by RNA-Seq on the SOLiD platform is indicated for all transcripts that match a 5X coverage threshold. The $\log_2$(fold-change) is displayed for all protein products detected by quantitative proteomics after *eIF4AIII* and *mago* knockdown. The *mapk*, *lt*, *tequila*, *PMCA*, *sxc* and *Dbp80* transcripts are labeled in red and *eIF4AIII* and *mago* are highlighted in green. All transcripts containing at least one intron larger than 1000bp are shaded in gray.

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

Criteria	# of genes	average fold-change upon EJC knockdown	*p*-value
largest intron >1000bp	1962	-0.32	1.72E-124
atypical 5'SS	1654	-0.24	1.15E-37
heterochromatin	116	-0.47	2.82E-11
atypical 5'SS and largest intron <250bp	555	-0.12	0.836
heterochromatin and largest intron <250bp	44	-0.27	0.027
all	6760	-0.12	

Table S2.II Summary of Gene Expression Changes Induced by EJC Knockdown

The fold-change upon EJC knockdown is the average of the gene expression values for *mago* and *eIF4AIII* dsRNA treated cells normalized to GFP dsRNA control cells. The values shown are for the indicated gene sets. Statistical analysis was conducted using a two-tailed Student's t test to compare the fold-changes of the indicated groups of genes to the remaining genes in the transcriptome dataset. Atypical 5' splice sites (5'SS) were defined as sites where the first three intronic residues did not match the consensus (GTR) sequence. Genes, introns and intronic sequences were extracted using flymine (FMv24.1) and genes containing at least one GTR mismatch were considered as atypical 5'SS genes for purpose of this analysis.

2.4 Discussion

Previous work has chiefly associated the EJC to the control of post-splicing events. Here, we describe a novel function for the nuclear EJC in the regulation of splicing. This function is not universal, because in the context of RAS1/MAPK signaling it was limited to *mapk*. Our observation that intron length determines sensitivity to EJC depletion provides an explanation for this. Importantly, this also provides an important insight into the splicing of transcripts with long introns, which suggests that the EJC is required for exon definition in this context.

Splicing of short introns (<200bp) occurs via the recognition of 5' and 3' splice sites across the intron (intron definition). The process differs in long introns (>250bp), where bordering exons require *a priori* recognition of their respective splice sites across the exon before splicing can occur (exon definition) (Fox-Walsh et al., 2005; Sterner et al., 1996). Exon definition is less robust than intron definition, and is thus thought to be more permissive to regulation (Fox-Walsh et al., 2005; Sterner et al., 1996). Consistent with this, long genes with multiple exons tend to present more alternative splice variants (Budagyan and Loraine, 2004) and exons bordered by large introns are much more likely to be excluded (Fox-Walsh et al., 2005; Kim et al., 2007; McGuire et al., 2008; Roy et al., 2008). This correlation is more pronounced in lower eukaryotes, where large introns are comparatively rare and seem to act as major determinants of alternative splicing (Fox-Walsh et al., 2005). In vertebrates, the presence of additional modes of regulation is thought to explain the less predominant impact of intron size. Still, it remains unclear whether the effect of long introns on alternative splicing involves specific protein factors, although the binding of multiple hnRNPs to long introns has been proposed as one explanation for this effect (Fox-Walsh et al., 2005).

Our finding that the EJC is associated with exon definition in transcripts containing long introns suggests that the nuclear EJC is one of the factors determining alternative splicing in this context. Importantly, amongst the different types of alternative splicing events, exon exclusion – often of multiple consecutive exons – was the salient feature of EJC disruption. This is consistent with previous reports that exon exclusion is predominantly associated to long introns (Fox-Walsh et al., 2005; Kim et al., 2007; McGuire et al., 2008). However, we

also observed some intron retention and alternative donor and acceptor site usage in the transcripts we sequenced (Figure 2.5 and S2.7). In addition to these splicing changes, EJC disruption also caused a reduction in transcript levels of genes with long introns (Figures 2.6 and 7). This may be due to a lower stability of the new transcript variants or to other splicing defects culminating in the degradation of the RNA product.

As introns tend to be larger in higher eukaryotes compared to *Drosophila* and other lower eukaryotes (Fox-Walsh et al., 2005; Kim et al., 2007; McGuire et al., 2008), it will be interesting to verify the extent to which the EJC is involved in splicing in higher eukaryotes. With respect to the *MAPK* genes, it is intriguing to note that human *MAPK1*/*ERK2* bears an intron just under 60kb, which is significantly above the average human intron length (Lander et al., 2001). In contrast, *MAPK3*/*ERK1* has significantly smaller introns, raising the possibility that these two *MAPK* genes bear different sensitivities to regulation by the EJC.

Despite the fact that the core components of the nuclear EJC have not been previously linked to splicing, they have been shown to associate with the spliceosome during splicing and before actual deposition of the EJC on the mRNA (Gehring et al., 2009; Merz et al., 2007). Furthermore, deposition of EIF4AIII upstream of the exon junction site has been found to occur during the second step of the splicing reaction, before cleavage of the 3' splice site (Gehring et al., 2009). Thus, these previous observations do not preclude an involvement of the nuclear EJC in the splicing process.

The cytoplasmic EJC, which includes BTZ, is required for *oskar* mRNA localization and translation in *Drosophila* (Hachet and Ephrussi, 2004; Mohr et al., 2001; van Eeden et al., 2001). However, *btz* depletion did not influence MAPK activity, nor did it cause reduction in *mapk* mRNA or protein levels. Though BTZ is sometimes described as a core EJC component, EJC complexes devoid of BTZ have been found to associate with the spliceosome (Bessonov et al., 2008; Herold et al., 2009; Merz et al., 2007), and human EIF4A3, RBM8A/Y14 and MAGOH have been shown to form a trimeric complex in the absence of MLN51/BTZ (Ballut et al., 2005; Gehring et al., 2009). Furthermore, BTZ deposition on the mRNA has been found to occur after completion of the splicing reaction (Gehring et al., 2009). Also, whereas MAGO, TSU and EIF4AIII have been described as being mainly located in the nucleus and in nuclear speckles (Le Hir et al., 2001a; Palacios et al., 2004), BTZ bears a nuclear export signal

and is predominantly cytoplasmic at steady state (Degot, 2004; Macchi et al., 2003; Palacios et al., 2004). Thus, BTZ is likely not involved in nuclear events regulated by the EJC, which is consistent with our observation that *btz* depletion does not impact the splicing of *mapk* or of other pre-mRNAs (Figures 2.5C and 2.7A).

While the EJC has not been previously implicated in splicing, this is not the case for some EJC-associated factors such as the SR (serine/arginine-rich) factors RNPS1 and SRM160. SR factors are important determinants of constitutive and alternative splicing (Long and Caceres, 2009). RNPS1 was initially characterized as a splicing factor (Mayeda et al., 1999), and has since been shown to regulate alternative splicing through alternate exon usage (Sakashita et al., 2004) and to enhance spliceosomal activity (Trembley et al., 2005). The fact that we also identified RNPS1 as a factor linked to the EJC's splicing function is consistent with this and further suggests that the recruitment of RNPS1 or other SR factors could provide an additional level of specificity to the EJC's effect. In support of this idea, our RT-PCR data showed that *RnpS1* did not impact the splicing of all EJC target genes (Figure 2.7). Importantly, RNPS1 takes part in more than one aspect of the EJC's functions as it is also linked to NMD (Lykke-Andersen et al., 2001) and translational enhancement (Nott et al., 2004; Wiegand et al., 2003). SRM160, another EJC-associated SR factor, has also been shown to function as a splicing co-activator which binds to exonic splicing enhancers via other SR factors (Blencowe et al., 1998; Eldridge et al., 1999). However, *SRm160* knockdown had modest effects on MAPK expression and no splicing defects in *mapk* could be observed by RT-PCR (data not shown). Interestingly, SRM160 and RNPS1 also promote pre-mRNA 3' end cleavage, but only SRM160 can function independently of the EJC in this context (McCracken et al., 2002; McCracken et al., 2003), indicating that its activity can be uncoupled from that of the EJC. Still, it is possible that this component acts in concert with the EJC in the splicing of other transcripts.

One important issue raised by our findings is whether the EJC's effect on splicing is regulated, or if it is part of a constitutive process involved in exon definition. EJC activity is known to be regulated in the context of translational enhancement where mTOR signaling modulates this activity via the EJC co-factor, SKAR (Ma et al., 2008). A similar mechanism involving EJC co-factors could also be responsible for bridging different inputs with the

regulation of splicing. Accordingly, the CK2 kinase has been found to phosphorylate RNPS1 and regulate its splicing activity (Trembley et al., 2005). More generally, the control of alternative splicing through SR and hnRNP factors is regulated by different signaling events (reviewed in (House and Lynch, 2008; Stamm, 2008)). For example, the RAS/MAPK dependent regulation of CD44 splicing occurs via the SRM160 and SAM68 SR factors (Cheng and Sharp, 2006). Thus, SR factors could provide further specificity and signal-integration properties to the EJC's function in splicing. Further investigation of the splicing changes brought about by co-factors such as RNPS1 will help to understand their contribution to EJC-regulated splicing. Indeed, one possibility is that the EJC acts as an adaptor platform for SR factors which are required for exon definition. In this model, the individual SR factors would be the effectors involved in providing specificity to splice site selection. Alternatively, it is possible that the EJC also directly impacts splice site selection (and exon definition) by masking either binding sites for other splicing factors or RNA motifs directly involved in splice site selection (Yu et al., 2008). A third possibility is that the EJC stabilizes the interaction of spliceosome complexes with splice sites and that this is of particular importance for exon definition. Investigation of these possibilities will be important in order to further understand the EJC's role in splicing at the mechanistic level.

The example of CD44 is of particular interest in the context of this study as the modulation of CD44 also constitutes a feedback mechanism that regulates RAS activity (Cheng and Sharp, 2006; Cheng et al., 2006). Since we initially identified the EJC as specifically regulating *mapk* expression in the context of the RAS1/MAPK pathway, and as *mapk* was amongst the most highly modulated EJC targets, both at the mRNA and protein level, the control of *mapk* splicing, like CD44, may also be a key element in modulating signal flow. The identification of the inputs regulating the splicing of *mapk* via the EJC will provide additional insights into how signal modulation is achieved through this route. Also, although many signaling genes are known to encode alternatively spliced transcripts with different functions, the extent to which regulation of signaling processes is achieved by splicing is still largely underappreciated.

2.5 Experimental Procedures

2.5.1 Quantitative Immunofluorescence Microscopy

S2 cell lines were distributed in 96 well clear plates (Corning) containing a final concentration of 200ng/μL dsRNA. Cells were then fixed, blocked and incubated overnight with a primary antibody (anti-pMAPK 1/2000, Sigma #M8159 or an anti-pJNK 1/500, NEB #9251S), then stained with a secondary antibody (anti-mouse Alexa Fluor 555-conjugated 1/1000, Invitrogen #A-21424) and with DAPI. Mowiol (9.6% PVA, Fluka). An automated fluorescence microscopy system (Zeiss Axiovert) was employed for plate imaging. Autofocus, image acquisition and analysis were conducted using MetaMorph (Molecular Devices) software. The cell-scoring application in MetaMorph was used for quantitative image analysis. Information on cell lines and cell culture conditions is included in the Extended Experimental Procedures.

2.5.2 Fly Genetics, Immunohistochemistry and ESEM

Fly husbandry was conducted according to standard procedures. Crosses to *raf/phl*^{12} flies were performed at 18°C. All other crosses were performed at 25°C. The *sev-Ras*^{V12} and *raf/phl*^{12} (formerly referred to as *raf*^{HM7}) lines have been described previously (Karim et al., 1996; Melnick et al., 1993). *Egfr*^{Elp} was described in (Baker and Rubin, 1989). The *csw*^{lf} (Perkins et al., 1996) was kindly provided by L. Perkins. The *mago* alleles were described in (Boswell et al., 1991). RNAi fly lines were obtained from the VDRC (Dietzl et al., 2007). All other mutant lines described herein were obtained from the Bloomington stock center.

Homozygous *mago* mutant clones were generated using the *flp-FRT* technique (Xu and Rubin, 1993). Third instar eye discs were fixed and stained with anti-MAPK (1/1000, Cell Signaling #4695), anti-CUT (1/200, DSHB) and anti-ELAV (1/20, DSHB) antibodies. Eye disc images were acquired on a Zeiss LSM 510 laser scanning confocal microscope. Fly eyes were imaged using an environmental scanning electron microscope (Quanta 200 FEG) or stereomicroscope (Leica MZ FL III). Permount-mounted wings were imaged using a Nanozoomer (Hamamatsu).

2.5.3 RT-qPCR, RT-PCR and Northern Blot

S2 cells were cultured in dsRNA (15 μg/ml) for seven days and total RNA was prepared using TRIzol reagent (Invitrogen).

For RT-qPCR, 2 μg of total RNA was reverse transcribed using the High Capacity cDNA Archive Kit with random primers (Applied Biosystems). PCR reactions for 384-well plate formats were performed using 2 μl of cDNA, 5 μl of the TaqMan fast Universal PCR Master Mix (Applied Biosystems), 2 μM of each primer and 1 μM of the Universal TaqMan probe in a total volume of 10 μl. The ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) was used to detect the amplification level.

For RT-PCR, 1 μg of total RNA was primed with oligo$(dT)_{18}$ followed by reverse transcription (RT) with SuperScript II Reverse Transcriptase (Invitrogen). 1/20 of the RT reaction was used as template for PCR.

For northern blot analyses, poly$(A)^+$ mRNA was isolated from total S2 cell RNA using oligotex resin (Qiagen). mRNA samples (1.5μg) were separated on a 5% formaldehyde-1% agarose gel and transferred to a nylon membrane (Hybond-N+; GE Healthcare). Hybridizations were conducted in 0.125M Na_2HPO_4 pH 7.4, 4mM EDTA, 7% SDS. ^{32}P-labeled probes were synthesized by random priming using *mapk/rl* or *jnk/bsk* cDNA fragments. Membranes were washed three times in 1X SSC/0.1% SDS and once in 0.1X SSC/0.1% SDS for 20 minutes at 65°C, and then exposed for 3 days at -80°C using an intensifying screen.

For additional information regarding primer sequences see the Extended Experimental Procedures and Table S2.III.

Table S2.III RT-qPCR Primers and TaqMan Hydrolysis Probes

Gene	Left Primer	Right Primer	Probe#
eIF4AIII	ccgattcgcatcttggtc	cgacgaagaactgcttgatg	47
mago	agatcatgcaggaggacga	tccgatgacgatctccagt	37
tsu	attgacaatgcggaggagtt	gtgcttcgccttttccttc	37
btz	aagcgctactccagcttgc	ctgctcagccaagtgctgt	18
RnpS1	tcaacaatcgtggaggaggt	tcctcatgggtgactgacg	70
jnk/bsk	catagtttaaaatgctcgccact	tgagagcaaggagcatatcg	104
Akt1	ctttgcgagtattaactggacaga	ggatgtcacctgaggcttg	70
RpL32	cggatcgatatgctaagctgt	gcgcttgttcgatccgta	105
Act5C	accgagcgcggttactct	cttgatgtcacggacgatttc	77

2.5.4 Whole Transcriptome Sequencing and Analysis

Total RNA from S2 cells subjected to treatment with *eIF4AIII, mago* and *GFP* dsRNA in duplicate and was prepared using TRIzol reagent (Invitrogen) and Poly(A)+ mRNAs were enriched using oligo(dT) selection with the Oligotex mRNA Midi kit (Qiagen). The resulting mRNA was then depleted of rRNA molecules using the Ribominus Eukaryote kit (Invitrogen). High-throughput sequencing libraries were prepared according to the SOLiD whole transcriptome library preparation protocol (Applied Biosystems). Whole transcriptome reads were aligned using the Bioscope software package (Applied Biosystems) using the UCSC *Drosophila* genome release 4.2 as a reference (**http://genome.ucsc.edu/**). 953, 837 and 524 million reads mapped uniquely to the reference genome and yielded an average exon coverage of 154X, 134X and 77X for the *GFP*, *eIF4AIII* and *mago* knocked down samples respectively.

2.6 Acknowledgements

We are grateful to R. Boswell, E. Perkins, N. Perrimon and D. St-Johnston as well as the Bloomington and VDRC stock centers for fly stocks and cell lines. We thank Christian Charbonneau and Monica Nelea for assistance with microscopy, Philippe Roux, Pierre Zindy and Katherine Borden for their help with the polysome fractionation experiments. We also extend our gratitude to the IRIC HTS platform for use of the automated fluorescence microscope, to Michael Kubal (ABI) for help with analysis of the RNA-Seq data and to Raphaëlle Lambert for assistance with RT-qPCR and sequencing. DAB is a recipient of Frederick Banting and Charles Best Canada Doctoral Scholarship. HL is a recipient of a Cancer Research Society postdoctoral fellowship. MT is recipient of a Tier II Canada Research Chair in Intracellular Signaling. This work was supported by the Canadian Cancer Society and by the Canadian Institutes for Health Research.

2.7 References

Baker, N.E., and Rubin, G.M. (1989). Effect on eye development of dominant mutations in *Drosophila* homologue of the EGF receptor. Nature *340*, 150-153.

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B., and Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat Struct Mol Biol *12*, 861-869.

Berghella, L., and Dimitri, P. (1996). The heterochromatic rolled gene of *Drosophila* melanogaster is extensively polytenized and transcriptionally active in the salivary gland chromocenter. Genetics *144*, 117-125.

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., and Luhrmann, R. (2008). Isolation of an active step I spliceosome and composition of its RNP core. Nature *452*, 846-850.

Blencowe, B.J., Issner, R., Nickerson, J.A., and Sharp, P.A. (1998). A coactivator of pre-mRNA splicing. Genes & Development *12*, 996-1009.

Boswell, R.E., Prout, M.E., and Steichen, J.C. (1991). Mutations in a newly identified *Drosophila* melanogaster gene, mago nashi, disrupt germ cell formation and result in the formation of mirror-image symmetrical double abdomen embryos. Development *113*, 373-384.

Budagyan, B., and Loraine, A. (2004). Gene length and alternative transcription in fruit fly. Computational Systems Bioinformatics Conference, 2004: 2004, Stanford, CA, USA: Proceedings, 515 - 516.

Chang, Y.F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem *76*, 51-74.

Cheng, C., and Sharp, P.A. (2006). Regulation of CD44 alternative splicing by SRm160 and its potential role in tumor cell invasion. Mol Cell Biol *26*, 362-370.

Cheng, C., Yaffe, M.B., and Sharp, P.A. (2006). A positive feedback loop couples Ras activation and CD44 alternative splicing. Genes Dev *20*, 1715-1720.

Degot, S. (2004). Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module. J Biol Chem *279*, 33702-33715.

Diem, M.D., Chan, C.C., Younis, I., and Dreyfuss, G. (2007). PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs. Nat Struct Mol Biol *14*, 1173-1179.

Dietzl, G., Chen, D., Schnorrer, F., Su, K.C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S.*, et al.* (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. Nature *448*, 151-156.

Dougherty, M.K., Muller, J., Ritt, D.A., Zhou, M., Zhou, X.Z., Copeland, T.D., Conrads, T.P., Veenstra, T.D., Lu, K.P., and Morrison, D.K. (2005). Regulation of Raf-1 by direct feedback phosphorylation. Mol Cell *17*, 215-224.

Douziech, M., Roy, F., Laberge, G., Lefrancois, M., Armengod, A.V., and Therrien, M. (2003). Bimodal regulation of RAF by CNK in *Drosophila*. Embo J *22*, 5068-5078.

Ebisuya, M., Kondoh, K., and Nishida, E. (2005). The duration, magnitude and compartmentalization of ERK MAP kinase activity: mechanisms for providing signaling specificity. J Cell Sci *118*, 2997-3002.

Eldridge, A.G., Li, Y., Sharp, P.A., and Blencowe, B.J. (1999). The SRm160/300 splicing coactivator is required for exon-enhancer function. Proceedings of the National Academy of Sciences of the United States of America *96*, 6125-6130.

Fortini, M.E., Simon, M.A., and Rubin, G.M. (1992). Signalling by the sevenless protein tyrosine kinase is mimicked by Ras1 activation. Nature *355*, 559-561.

Fox-Walsh, K.L., Dou, Y., Lam, B.J., Hung, S.P., Baldi, P.F., and Hertel, K.J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proc Natl Acad Sci U S A *102*, 16176-16181.

Gatfield, D., and Izaurralde, E. (2002). REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. J Cell Biol *159*, 579-588.

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P., and Izaurralde, E. (2003). Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. Embo J *22*, 3960-3970.

Gehring, N.H., Lamprinaki, S., Hentze, M.W., and Kulozik, A.E. (2009). The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. PLoS Biol *7*, e1000120.

Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W., and Kulozik, A.E. (2003). Y14 and hUpf3b Form an NMD-Activating Complex. Molecular Cell *11*, 939-949.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Mol Cell Biol *19*, 1720-1730.

Hachet, O., and Ephrussi, A. (2004). Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. Nature *428*, 959-963.

Herold, N., Will, C.L., Wolf, E., Kastner, B., Urlaub, H., and Luhrmann, R. (2009). Conservation of the protein composition and electron microscopy structure of *Drosophila* melanogaster and human spliceosomal complexes. Mol Cell Biol *29*, 281-301.

House, A.E., and Lynch, K.W. (2008). Regulation of alternative splicing: more than just the ABCs. J Biol Chem *283*, 1217-1221.

Hurlbut, G.D., Kankel, M.W., Lake, R.J., and Artavanis-Tsakonas, S. (2007). Crossing paths with Notch in the hyper-network. Curr Opin Cell Biol *19*, 166-175.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the let-7 microRNA family. Cell *120*, 635-647.

Karim, F.D., Chang, H.C., Therrien, M., Wassarman, D.A., Laverty, T., and Rubin, G.M. (1996). A screen for genes that function downstream of Ras1 during *Drosophila* eye development. Genetics *143*, 315-329.

Kim, E., Magen, A., and Ast, G. (2007). Different levels of alternative splicing among eukaryotes. Nucl Acids Res *35*, 125-131.

Kolch, W. (2005). Coordinating ERK/MAPK signalling through scaffolds and inhibitors. Nat Rev Mol Cell Biol *6*, 827-837.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Le Hir, H., Gatfield, D., Braun, I.C., Forler, D., and Izaurralde, E. (2001a). The protein Mago provides a link between splicing and mRNA localization. EMBO Rep *2*, 1119-1124.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001b). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J *20*, 4987-4997.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20[ndash]24 nucleotides upstream of mRNA exon-exon junctions. EMBO J *19*, 6860-6869.

Lee, M.H., Hook, B., Pan, G., Kershner, A.M., Merritt, C., Seydoux, G., Thomson, J.A., Wickens, M., and Kimble, J. (2007). Conserved regulation of MAP kinase expression by PUF RNA-binding proteins. PLoS Genet *3*, e233.

Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochem J *417*, 15-27.

Longman, D., Plasterk, R.H.A., Johnstone, I.L., and Cáceres, J.F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes & Development *21*, 1075-1085.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science *293*, 1836-1839.

Ma, X.M., Yoon, S.O., Richardson, C.J., Julich, K., and Blenis, J. (2008). SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. Cell *133*, 303-313.

Macchi, P., Kroening, S., Palacios, I.M., Baldassa, S., Grunewald, B., Ambrosino, C., Goetze, B., Lupas, A., St Johnston, D., and Kiebler, M. (2003). Barentsz, a new component of the Staufen-containing ribonucleoprotein particles in mammalian cells, interacts with Staufen in an RNA-dependent manner. J Neurosci *23*, 5778-5788.

Mayeda, A., Badolato, J., Kobayashi, R., Zhang, M.Q., Gardiner, E.M., and Krainer, A.R. (1999). Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing. EMBO J *18*, 4560-4570.

McCracken, S., Lambermon, M., and Blencowe, B.J. (2002). SRm160 splicing coactivator promotes transcript 3'-end cleavage. Mol Cell Biol *22*, 148-160.

McCracken, S., Longman, D., Johnstone, I.L., Cáceres, J.F., and Blencowe, B.J. (2003). An Evolutionarily Conserved Role for SRm160 in 3′-End Processing That Functions

Independently of Exon Junction Complex Formation. Journal of Biological Chemistry *278*, 44153-44160.

McGuire, A., Pearson, M., Neafsey, D., and Galagan, J. (2008). Cross-kingdom patterns of alternative splicing and splice recognition. Genome Biology *9*, R50.

McKay, M.M., and Morrison, D.K. (2007). Integrating signals from RTKs to ERK/MAPK. Oncogene *26*, 3113-3121.

Melnick, M.B., Perkins, L.A., Lee, M., Ambrosio, L., and Perrimon, N. (1993). Developmental and molecular characterization of mutations in the *Drosophila*-raf serine/threonine protein kinase. Development *118*, 127-138.

Merz, C., Urlaub, H., Will, C.L., and Luhrmann, R. (2007). Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. RNA *13*, 116-128.

Mohr, S.E., Dillon, S.T., and Boswell, R.E. (2001). The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during *Drosophila* oogenesis. Genes Dev *15*, 2886-2899.

Mount, S.M., Burks, C., Hertz, G., Stormo, G.D., White, O., and Fields, C. (1992). Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. Nucleic Acids Res *20*, 4255-4262.

Nott, A., Le Hir, H., and Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev *18*, 210-222.

Nykamp, K., Lee, M.H., and Kimble, J. (2008). C. elegans La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. RNA *14*, 1378-1389.

Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature *427*, 753-757.

Perkins, L.A., Johnson, M.R., Melnick, M.B., and Perrimon, N. (1996). The nonreceptor protein tyrosine phosphatase corkscrew functions in multiple receptor tyrosine kinase pathways in *Drosophila*. Dev Biol *180*, 63-81.

Perkins, L.A., Larsen, I., and Perrimon, N. (1992). corkscrew encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso. Cell *70*, 225-236.

Roy, M., Kim, N., Xing, Y., and Lee, C. (2008). The effect of intron length on exon creation ratios during the evolution of mammalian genomes. RNA *14*, 1 - 13.

Sakashita, E., Tatsumi, S., Werner, D., Endo, H., and Mayeda, A. (2004). Human RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator in vivo. Mol Cell Biol *24*, 1174-1187.

Schubbert, S., Shannon, K., and Bollag, G. (2007). Hyperactive Ras in developmental disorders and cancer. Nat Rev Cancer *7*, 295-308.

Stamm, S. (2008). Regulation of alternative splicing by reversible protein phosphorylation. J Biol Chem *283*, 1223-1227.

Sterner, D.A., Carlo, T., and Berget, S.M. (1996). Architectural limits on split genes. Proc Natl Acad Sci U S A *93*, 15081-15085.

Tange, T.O., Nott, A., and Moore, M.J. (2004). The ever-increasing complexities of the exon junction complex. Curr Opin Cell Biol *16*, 279-284.

Therrien, M., Morrison, D.K., Wong, A.M., and Rubin, G.M. (2000). A genetic screen for modifiers of a kinase suppressor of Ras-dependent rough eye phenotype in *Drosophila*. Genetics *156*, 1231-1242.

Trembley, J.H., Tatsumi, S., Sakashita, E., Loyer, P., Slaughter, C.A., Suzuki, H., Endo, H., Kidd, V.J., and Mayeda, A. (2005). Activation of pre-mRNA splicing by human RNPS1 is regulated by CK2 phosphorylation. Mol Cell Biol *25*, 1446-1457.

Turjanski, A.G., Vaque, J.P., and Gutkind, J.S. (2007). MAP kinases and the control of nuclear events. Oncogene *26*, 3240-3253.

van Eeden, F.J., Palacios, I.M., Petronczki, M., Weston, M.J., and St Johnston, D. (2001). Barentsz is essential for the posterior localization of oskar mRNA and colocalizes with it to the posterior pole. J Cell Biol *154*, 511-523.

Wassarman, D.A., Therrien, M., and Rubin, G.M. (1995). The Ras signaling pathway in *Drosophila*. Curr Opin Genet Dev *5*, 44-50.

Wen, J., and Brogna, S. (2010). Splicing-dependent NMD does not require the EJC in Schizosaccharomyces pombe. Embo J.

Wiegand, H.L., Lu, S., and Cullen, B.R. (2003). Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc Natl Acad Sci USA *100*, 11327-11332.

Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult *Drosophila* tissues. Development *117*, 1223-1237.

Yu, Y., Maroney, P.A., Denker, J.A., Zhang, X.H., Dybkov, O., Luhrmann, R., Jankowsky, E., Chasin, L.A., and Nilsen, T.W. (2008). Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition. Cell *135*, 1224-1236.

Zebisch, A., Czernilofsky, A.P., Keri, G., Smigelskaite, J., Sill, H., and Troppmair, J. (2007). Signaling through RAS-RAF-MEK-ERK: from basics to bedside. Curr Med Chem *14*, 601-623.

2.8 Supplemental experimental procedures

2.8.1 Cell Lines and Plasmids

Drosophila S2 cells were cultured at 27°C in Schneider medium (Invitrogen) supplemented with 10% fetal bovine serum. S2 cell lines stably transfected with the following constructs were used to conduct RNAi screening assays: *pMet-Ras^{V12}* (Therrien et al., 1999), *pMet-raf^{ED}* (Douziech et al., 2006), *pMet-mek^{EE}* (Douziech et al., 2003), *pHS-Sev^{S11}* (Therrien et al., 1998), *pMet-EGFR* (a gift from N. Perrimon), *pMet-Rac1^{V12}* (Baril et al., 2009). *pMet-raf^{CT}* and *pMet-raf^{EDCT}* respectively encode N-terminal deleted (2-371) RAF made from *pMet-RAF* and *pMet-RAF^{ED}* (Douziech et al., 2006). *pMet-GFP* was made by inserting an *EGFP* cDNA in the *pMet* vector (Therrien et al., 1998).

2.8.2 Cell Culture Functional Assays

S2 cells were diluted in Schneider medium (to a concentration of $1x10^6$ cells/mL) before distribution in 96 well plates. Plates were placed in plastic containers to reduce evaporation and incubated at 27°C for four or six days. pMet driven expression of *Ras1^{V12}*, *raf^{CT}*, *raf^{ED}*, *raf^{EDCT}*, *mek^{EE}* and *Rac1^{V12}* was induced by adding 0.7mM $CuSO_4$ 24h prior to fixation. *pHS-Sev^{S11}* cells were induced for 30' at 37°C and incubated 2.5h at 27°C before fixation. Cells were resuspended and transferred to concanavalin A-coated plates and allowed to settle and adhere for 1h before being fixed and stained.

2.8.3 Western Blots

For western blot analyses of S2 cell lysates, cells were lysed in cold lysis buffer (20 mM Tris, pH 8.0, 137 mM NaCl, 10% glycerol, 1% Igepal CA-630, 1 mM EDTA) supplemented with 1X phosphatase inhibitor cocktail (Sigma #P2850), 10 μg/mL each aprotinin and leupeptin, and 1 mM PMSF. Lysates were then clarified by centrifugation at 10,000 x g. Protein samples were resolved by electrophoresis on 12% SDS-polyacrylamide gels and transferred to nitrocellulose membranes. Specific *Drosophila* proteins were detected using the following antibodies: anti-CNK (1:5; Douziech et al., 2003); anti-RAS1 (1:10; Douziech et al., 2003); anti-RAF (1:5000; Douziech et al., 2003); anti-MEK (1:1000; Cell Signaling #9122); anti-MAPK (1:1000; Cell Signaling #4695); anti-JNK (1:2000; Santa Cruz #SC-571); anti-AKT (1:2000; Cell Signaling #9272); anti-Actin (1:2000; Chemicon #MAB1501).

2.8.4 RT-qPCR

Gene-specific assays (primer sets and TaqMan hydrolysis probes) were designed using the Universal ProbeLibrary assay design center (Roche Applied Science), with the exception of the *mapk/rl* assay. The assay for *mapk/rl* was a custom assay from Integrated DNA Technologies and the primer and probe sequences are as follows; ATAAACGGATTCCTGTCGAGG (left primer), ATCAACGACTTCAGGGCATC (right primer), and TCCGAAATGGCACTTCAGCGACA (probe, labeled with 5' FAM fluorophore and 3' Iowa Black quencher). For the other assays, the primer sequences and the Universal ProbeLibrary probe number are listed in the Table S3. Assays were designed such that the amplified regions did not overlap with sequences targeted by dsRNA. The quantification of target genes was determined by using the Ct method. Briefly, the Ct (threshold cycle) values of target genes were normalized to a reference gene (Act5C unless indicated otherwise) where $\Delta Ct = Ct_{target} - Ct_{Act5C}$, and then compared with a calibrator sample (GFP RNAi) where $\Delta\Delta Ct = \Delta Ct_{Sample} - \Delta Ct_{Calibrator}$. Relative expression (RQ) or fold change was calculated using the

Sequence Detection System (SDS) 2.2.2 software (Applied Biosystems) and the formula RQ = $2^{-\Delta\Delta CT}$.

2.8.5 Additional Information for Northern Blot, RT-PCR and Sequencing

For Northern Blots, the ^{32}P-labeled probes were synthesized by PCR using the following primers: *mapk*-OL5' TTAATTCGAGCGGATCAGTAG and *mapk*-OL3' GATCCAGGTAATGTTTTCCAG, or *jnk*-OL5' CAGGATGTCTACCTGGTCATG and *jnk*-OL3' GATTAGCTCCTTCCACTGCTC).

Mapk transcripts were amplified with the following forward and reverse primers; 5'–GTCGGATCCTACGCCGTCGATTTTGATAAATC-3' and 5'-GTCGAGCTCTTCTTTACTTTCTTTAATCGATCTTTAATATTC-3'. Transcripts for *lt* were amplified with 5'- GTCGAATTCTCCGTTTGCTTTAAATGGCTA-3' and 5'-GTCGCGGCCGCGGTCACAATTTCAGGGTGCT-3'. *PMCA* transcripts were amplified with 5'- GCTCTAGAGTGTTTGCACCAAATATGGTTAAGCC-3' and 5'-GGGGTACCCAAATGGCCACTATAGATGGAAGACC-3'. *Tequila* transcripts were amplified with 5'-CTAGCTAGCGGATACACAACGAAGCCAAGA-3' and 5'-ATAAGAATGCGGCCGCATCATCCCAGCAATGTGGTT-3'. Restriction sites in each primer are underlined.

PCR products were digested with restriction enzymes and inserted into the *pMet* vector (Therrien et al., 1998) or *pBluescript II KS+* (Stratagene). Mapk PCR products were inserted into the BamHI and SacI sites of *pMet*, while lt PCR products were inserted into the EcoRI and NotI sites of *pMet*. Clones were sequenced with the following flanking primers; 5'-GCCAATGTGCATCAGTTGTG-3' and 5'- GTGAGTGTGCATCGAATAGC-3'. PMCA PCR products were inserted into the XbaI and KpnI sites of *pBluescript II KS+*, Tequila PCR products were inserted into the NheI and NotI sites of *pBluescript II KS+*, and these clones were sequenced with M13 forward and reverse primers.

2.8.6 Polysome Fractionation

S2 cells were cultured in GFP or mago-specific dsRNA (15 μg/ml) for four days and then treated with cycloheximide (CHX) at 0.1 mg/ml for 30 minutes. For each RNAi, 10^8 cells were pelleted and washed once with cold PBS containing 0.1 mg/ml CHX. The cells were then lysed in polysome lysis buffer (PLB; 20 mM Tris-Cl, pH 7.5, 250 mM KCl, 10 mM $MgCl_2$, 1% Triton X-100, 1 mM DTT, 0.1 mg/ml CHX, 200 units/ml SUPERase-In (Ambion), 10 μg/ml each leupeptin and aprotinin, and 1 mM PMSF). The lysates were centrifuged at 12,000 x g for 10 minutes at 4°C and 1/10 of the supernatant was removed as input sample. The remainder of each supernatant was loaded on a 20-50% sucrose gradient prepared in PLB (without Triton X-100) and resolved by centrifugation in an SW41 rotor (Beckman) at 38,000 rpm for 3 hours at 4°C. Fractions were collected from the top of the gradient using a Gradient Station (Biocomp Instruments) while monitoring the absorbance at 254 nm. RNA was prepared from each fraction and assayed by RT-qPCR as described in the methods. RpL32 was used as the reference gene to calculate relative mRNA levels in the input samples.

2.8.7 Label-free Quantitative Proteomics of EJC-depleted S2 cells

Total proteins from S2 cells subjected to *eIF4AIII*, *mago*, *mapk* or *GFP* dsRNA treatment (see Material and Methods) were obtained in triplicate in lysis buffer (20 mM Tris, pH 8.0, 137 mM NaCl, 10% glycerol, 1% Igepal CA-630, 1 mM EDTA) supplemented with 1X phosphatase inhibitor cocktail (Sigma #P2850), 10 μg/mL each aprotinin and leupeptin, and 1 mM PMSF. 80 μg of proteins were separated on a 15-well 4%–12% precast NuPAGE gel (Invitrogen) at 100V for 3h30min at 4°C. The gel was stained with coomassie G250 (0,2% coomassie G250, 20% methanol, 0,5% acetic acid), and the lanes were cut into 12 equal-sized pieces. The gel pieces were digested by trypsin, and peptides were extracted three times with 90% acetonitrile (ACN) in 0.5 M urea. Combined extracts were dried and resuspended in 5% ACN, 0.1% formic acid (FA) prior to mass-spectrometry analyses.

Peptides were separated on a 150 μm ID, 10 cm reversed-phase nano-LC column (Jupiter C18, 3 μm, 300 Å, Phenomex) with a loading buffer of 0.2% formic acid (FA).

Peptide elution was achieved using a gradient of 5%–40% ACN in 56 min. The nano-LC column was coupled to an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific), and samples were injected in an interleaved manner. The mass spectrometer was operated in a data-dependent acquisition mode with a 1 s survey scan at 60,000 resolution, followed by three product ion scans (MS/MS) in the ion trap for the most abundant precursors above a threshold of 10,000 counts.

The centroided MS/MS data were merged into single peak-list files and searched with the Mascot search engine v2.10 (Matrix Science) against the combined forward and reversed Uniprot *Drosophila* database to obtain a false discovery rate < 1%. The mass tolerance on precursor and fragment ions was set to ±0.03 and ±0.5 Daltons, respectively.

Orbitrap raw LC-MS data files were transformed into peptide maps using the in-house peptide detection and clustering software (Marcantonio et al., 2008). Peptide maps belonging to one experiment were clustered and aligned using clustering parameters of $\Delta m/z = 0.02$ and +/- 2 min (wide), +/- 0.5 min (narrow). Peptide clusters were aligned with mascot identification files to assign sequence identity. Expression analyses were performed on proteins identified by at least 2 different peptide sequences across triplicate analyses. Expression values and relative standard deviation (RSD) were obtained by averaging the intensity differences and standard deviation of the corresponding peptide triplets, respectively.

2.9 Supplementary references

Baril, C., Sahmi, M., Ashton-Beaucage, D., Stronach, B., and Therrien, M. (2009). The PP2C Alphabet Is a Negative Regulator of Stress-Activated Protein Kinase Signaling in *Drosophila*. Genetics *181*, 567-579.

Douziech, M., Roy, F., Laberge, G., Lefrancois, M., Armengod, A.V., and Therrien, M. (2003). Bimodal regulation of RAF by CNK in *Drosophila*. Embo J *22*, 5068-5078.

Douziech, M., Sahmi, M., Laberge, G., and Therrien, M. (2006). A KSR/CNK complex mediated by HYP, a novel SAM domain-containing protein, regulates RAS-dependent RAF activation in *Drosophila*. Genes Dev *20*, 807-819.

Lyne, R., Smith, R., Rutherford, K., Wakeling, M., Varley, A., Guillier, F., Janssens, H., Ji, W., McLaren, P., North, P.*, et al.* (2007). FlyMine: an integrated database for *Drosophila* and Anopheles genomics. Genome Biol *8*, R129.

Marcantonio, M., Trost, M., Courcelles, M., Desjardins, M., and Thibault, P. (2008). Combined enzymatic and data mining approaches for comprehensive phosphoproteome analyses: application to cell signaling events of interferon-gamma-stimulated macrophages. Mol Cell Proteomics *7*, 645-660.

Therrien, M., Wong, A.M., Kwan, E., and Rubin, G.M. (1999). Functional analysis of CNK in RAS signaling. Proc Natl Acad Sci U S A *96*, 13259-13263.

Therrien, M., Wong, A.M., and Rubin, G.M. (1998). CNK, a RAF-binding multidomain protein required for RAS signaling. Cell *95*, 343-353.

3 A Functional Screen Reveals an Extensive Layer of Transcriptional and Splicing Control Underlying RAS/MAPK Signaling in *Drosophila*

Running head: Multiple Regulators Converge on MAPK Expression

Dariel Ashton-Beaucage[1], Christian M. Udell[1], Patrick Gendron[1], Malha Sahmi[1], Martin Lefrançois[1], Caroline Baril[1], Anne-Sophie Guenier[1], Jean Duchaine[1], Daniel Lamarre[1,2], Sébastien Lemieux[1], Marc Therrien[1,3,*]

[1] Institute for Research in Immunology and Cancer
Laboratory of Intracellular Signaling
Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montreal, Quebec, Canada, H3C 3J7

[2] Département de médecine, Université de Montréal

[3] Département de pathologie et de biologie cellulaire, Université de Montréal

***To whom correspondence should be addressed**
tel.: (514) 343-7837
fax.: (514) 343-6843

Contributions détaillées de l'auteur

Contributions expérimentales : Production de l'ensemble des données expérimentales à l'exception de la figure 4C,D et S7 (fournies par SM), de la figure S6 (fournie par UCM) et de la figure S9 (fournie par BC). Développement et mise au point de l'essai de criblage pMAPK. Développement de la procédure de microscopie automatisée. Production et analyse des données de criblage primaire avec DJ et GAS, assisté de GP. Production et analyse des données de criblage de validation et criblage secondaire. Élaboration d'une formule (avec l'assistance de LS) permettant d'établir les groupes d'épistasie et attribuer des scores de spécificité à partir des données de criblage secondaire. Conception, avec GP et LS, de la base de données *IRIC RNAi*. Conception de l'ensemble des expériences avec TM.

Contributions au manuscrit : Élaboration et rédaction du manuscrit. Montage des figures.

Contributions détaillées des co-auteurs

UCM a contribué les données de la figure S6 qui constituent la première observation de l'impact sur les niveaux de MAPK de plusieurs facteurs d'épissage. UCM a également mis au point la méthode qui a servi de base à l'essai d'export des ARN décrit en figure S4B,C. SM et BC ont contribué des données expérimentales (Fig.4C,D; Fig.S7 et Fig. S6) avec l'assistance de LM. LS et GP ont assisté ABD à l'analyse de données et GP a élaboré la base de données *IRIC RNAi*. DJ, GAS et LD ont assisté ABD au criblage ARNi primaire. TM a contribué de manière importante à l'élaboration et la supervision des travaux de recherche ainsi qu'à la rédaction de l'article.

Abbreviations

AS (alternative splicing), EJC (exon junction complex), *hyp* (hyphen; designates the *ave*/*hyp* gene). The following *Drosophila* genes are named based on their human counterparts: *Fnta* (farnesyl transferase alpha; *CG2976*), *Fgop2* (fibroblast growth factor receptor 1 oncogene partner 2; *CG10158*), *Slmap* (sarcolemma associated protein; *CG17494*), *Strip* (Striatin interacting protein; *CG11526*). *Ras85D* (refers to the *Drosophila* gene encoding RAS), *ras* (gene encoding an IMP dehydrogenase) is referred to by its full name, "raspberry", to avoid confusion. Following the nomenclature recommended by Flybase, gene symbols are in lower-cased italics and protein symbols are in upper-case without italics.

3.1 Abstract

The small GTPase RAS is among the most prevalent oncogenes. The evolutionarily conserved RAF-MEK-MAPK module that lies downstream of RAS is one of the main conduits through which RAS transmits proliferative signals in normal and cancer cells. Genetic and biochemical studies conducted over the last two decades uncovered a small set of factors regulating RAS/MAPK signaling. Interestingly, most of these were found to control RAF activation, thus suggesting a central regulatory role for this event. Whether additional factors are required at this level or further downstream remains an open question. To obtain a comprehensive view of the elements functionally linked to the RAS/MAPK cascade, we used a quantitative assay in *Drosophila* S2 cells to conduct a genome-wide RNAi screen for factors impacting RAS-mediated MAPK activation. The screen led to the identification of 101 validated hits, including most of the previously known factors associated to this pathway. Epistasis experiments were then carried out on individual candidates to determine their position relative to core pathway components. While this revealed several new factors acting at different steps along the pathway - including a new protein complex modulating RAF activation - we found that most hits unexpectedly work downstream of MEK and specifically influence MAPK expression. These hits mainly consist of constitutive splicing factors which thereby suggests that splicing plays a specific role in establishing MAPK levels. We further characterized two representative members of this group and surprisingly found that they act by regulating mapk alternative splicing. This study provides an unprecedented assessment of the factors modulating RAS/MAPK signaling in *Drosophila*. In addition, it suggests that pathway output does not solely rely on classical signaling events, such as those controlling RAF activation, but also on the regulation of MAPK levels. Finally, it indicates that core splicing components can also specifically impact alternative splicing.

3.2 Author summary

The RAS/MAPK pathway is a cornerstone of the cell proliferation signaling apparatus. It plays key roles in both normal development and diseased contexts, most notably in cancer where it is associated to aberrant proliferation. Previous work has focused predominantly on post-translational regulation of RAS/MAPK signaling such that a large and intricate network of factors is now known to act on core pathway components. However, regulation at the pre-translational level has not been examined nearly as extensively and is comparatively poorly understood. In this study, we used an unbiased and global screening approach to survey the *Drosophila* genome for novel regulators of this pathway. Surprisingly, a majority of our hits were associated to either transcription or mRNA splicing. We used a series of secondary screening assays to position the regulatory impact of the new candidates. We found that these factors were not equally distributed along the pathway, but rather converged predominantly on *mapk* mRNA expression and processing. Our findings raise the intriguing possibility that regulation of *mapk* transcripts production is a key focal point for a diverse set of regulatory inputs, and may play an important part in RAS/MAPK signaling dynamics.

3.3 Introduction

The RAS/MAPK pathway consists of a core module of three kinases (RAF, MEK and ERK/MAPK) that transmit signals downstream of the small GTPase RAS. Upstream factors such as Receptor Tyrosine Kinases (RTK), which respond to extracellular signals, lead to RAS activation by a Guanine nucleotide Exchange Factor (GEF). GTP-loaded RAS then triggers the sequential activation of RAF, MEK and MAPK; active RAF phosphorylates and activates MEK, which in turn phosphorylates and activates MAPK[1]. Unlike RAF and MEK, MAPK has a variety of cytoplasmic and nuclear substrates that include transcription factors such as c-Jun, c-Fos, p53, ELK1, c-Myc, c-Myb, STAT1/3, SRF and SMAD1/2/3/4[2-4]. Phosphorylation of these targets, and others, by MAPK induces a wide range of cellular responses that include proliferation, differentiation and survival[5]. Also, RAS/MAPK signaling's important role in oncogenesis and various developmental disorders has been recognized early on and abundantly studied[6,7].

Over the last two decades, genetic screens in metazoan models such as *Drosophila* and *C. elegans* have been instrumental in identifying a growing list of key regulators of the RAS/MAPK pathway such as *sos*[8], *csw*[9], *ksr*[10-12], *Cbl*[13], *dos*[14], *mts*/*PP2A*[15], *sur-8*/*soc-2*[16,17], *cnk*[18], *spry*[19], *sur-6*[20], *PTP-ER*[21], *let-7*[22], *alph*/*PP2C*[23] and *hyp*/*ave*[24,25]. Thus, these studies and research conducted in other systems have revealed a large network of factors whose regulatory activity converges on the core MAPK module[1,5,6,26,27]. This regulatory network includes complex features such as feedback loops[28-30], compartmentalization[1,31], crosstalk with other signaling pathways[32], allosteric modulation via dimerization[27,33] and the formation of larger order complexes called nanoclusters[34]. While the function of the core module is well characterized, many aspects of the network that surround it are still poorly understood, including its protein composition. Also, much of the identified regulators influence RAS-mediated RAF activation, which is in agreement with the fact that this particular step is subjected to a tight and complex regulation[27]. In comparison, fewer positively-acting components have been found to act downstream of RAF, suggesting that MEK and MAPK activation depends on more simple regulatory mechanisms. Alternatively, such modulators might have eluded detection. Finally,

most of the regulatory input that has been described so far acts at the post-translational level. Comparatively little is known on how RAS/MAPK component expression is controlled.

The success of the aforementioned genetic screens typically relied on the qualitative modification of a visible phenotype. This consideration, together with the technical limitations associated with genetic screening procedures, usually limit results to a handful of confirmed hits. RNA interference (RNAi) used as a functional genomics tool provides the possibility of a more comprehensive type of analysis providing a systematic means to functionally annotate the genome[35,36]. Moreover, the possibility of using quantitative assays, in particular, allows for the identification of a much wider range of regulators[37]. However, the considerable amount of candidates often identified by this methodology has made the perspective of rapid functional annotation a daunting task.

Here, we present the results of a genome-wide RNAi screen in *Drosophila* S2 cells that specifically focused on signal regulation between RAS and MAPK. Validated hits were submitted to a series of secondary assays aimed at positioning their regulatory input with respect to the three core kinases. In addition to identifying and correctly positioning most of the components previously known to mediate RAS-induced MAPK activation, the screen led to the discovery of several new factors that act at different steps along the pathway. Notably, we identified five novel components that act upstream of RAF. The homologs of these five proteins are part of a complex named Striatin-Interacting Phosphatase and Kinase Complex (STRIPAK)[38] that also includes PP2A, which is known to regulate RAF activation[39,40]. Unexpectedly, the majority of our candidates did not map to the interval between RAS and RAF, but were instead positioned further downstream. These included some transcription factors which we find regulate the transcript abundance of *mek*, *mapk* or of the MAPK phosphatase, *PTP-ER*. However, most of the novel factors were associated to mRNA processing and were found to act downstream of MEK and to regulate *mapk* splicing. Among these were components of the Exon Junction Complex (EJC), which we and others have previously reported to be involved in regulating the splicing of the *mapk* pre-mRNA[41,42]. In particular, depletion of the EJC was found to alter the splicing of *mapk*'s long introns and cause a reduction in the amount of functional protein product. In this study, we focus on the function of a larger group of canonical splicing factors that also regulate *mapk* splicing. We

show that the impact of these factors on alternative splicing (AS) of *mapk* differs from what we previously described for the EJC, indicating that two different types of regulatory input act on this step in *mapk* expression.

Thus, in addition to providing a comprehensive view of regulatory factors influencing signal transmission between RAS and MAPK, this work suggests that pathway output does not solely rely on post translational regulatory events, such as those controlling RAF activation, but is also tightly governed by the regulation of the expression of core components. In particular, the expression of MAPK emerges as a focal point for multiple different regulatory inputs.

3.4 Results

3.4.1 Identification of new RAS/MAPK pathway regulators

To systematically search for and categorize new factors that specifically modulate signaling between RAS and MAPK, we employed a screening strategy that involved three distinct steps: 1) a primary genome-wide RNAi screen, 2) a validation screening step aimed at eliminating false positives and 3) validated candidates were submitted to a series of a secondary screens to establish the position of their regulatory input relative to known pathway components (epistasis) and assess their specificity to RAS/MAPK signaling (Figure S3.1A).

We employed an automated immunofluorescence-based microscopy assay that quantitatively detected variations of dually-phosphorylated MAPK (pMAPK) in *Drosophila* S2 cells. This assay was used to screen a genome-wide dsRNA (long double-stranded RNA) library for modulation of pMAPK levels induced by RAS^{V12} expression (Figure S3.1B and C). The results from this primary screen and all subsequent screens are made available online at the IRIC RNAi database (http://www.bioinfo.iric.ca/iricrnai). 309 hit genes, which reproducibly altered pMAPK signal, were identified in the primary screen (Table S3.I). Importantly, core RAS/MAPK pathway components (e.g. *raf/phl*, *mek/Dsor1*, *mapk/rl*, *cnk* and *ksr*) were amongst the strongest hits that decreased the pMAPK signal (Figure 3.1A and Table S3.I). Other known positively-acting genes were also identified, such as *14-3-3ζ* and the RAF chaperone *Cdc37*[43]. Another expected hit was *βggt-I* which encodes a factor involved in RAS prenylation[10,44,45]. Also expected, the PP2C phosphatase *alph* was identified as a negative regulator[23].

We next conducted two successive validation steps to address readily identifiable sources of false positives, namely effects on the *pMet-*RAS^{V12} expression system and dsRNA off-target effects. 101 genes of the initial 309 primary hits passed both validation criteria (Figure S3.2 and Table S3.I). Validated genes were then assigned to broad functional categories based on their associated Gene Ontology (GO) terms and on the functions of predicted homologs. Interestingly, transcription and mRNA processing factors comprised, together, roughly half of our candidates (Figure 3.1B) and mRNA processing was the most

highly enriched GO term of our hit set (Table S3.II). Despite the fact that mRNA splicing factors are often enriched in RNAi screen hit lists[46], we chose not to apply a selection bias against any group of genes at this stage. Therefore, all of the candidate genes passing both primary and validation criteria were evaluated in secondary screens without distinction.

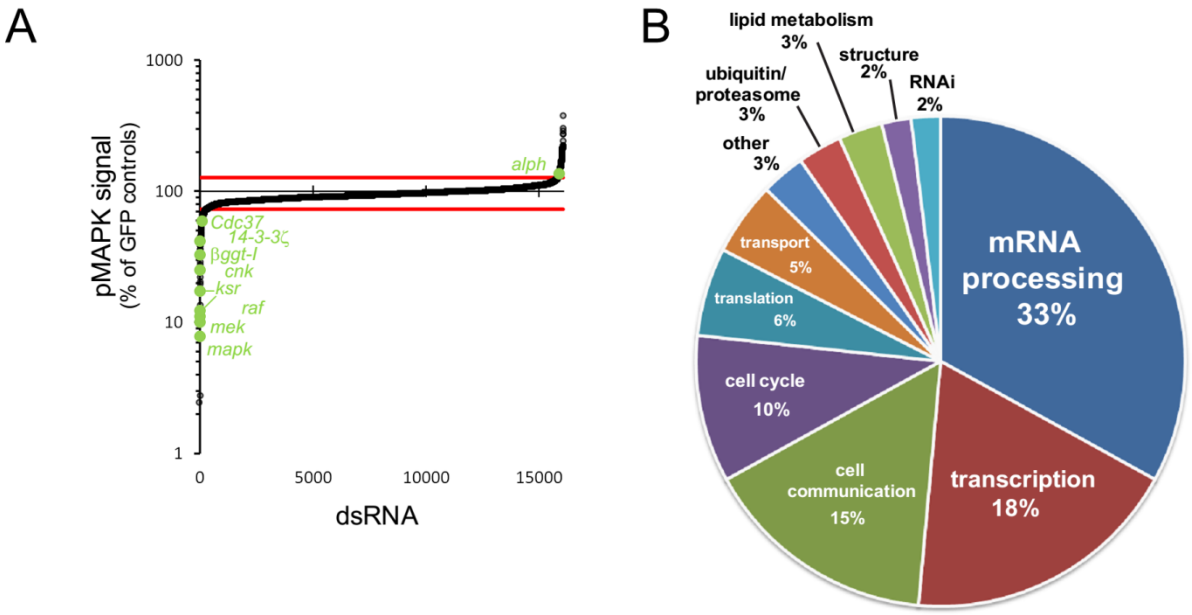


Figure 3.1 Primary Screen Results

(A) pMAPK signal of primary screen dsRNA probes normalized to plate-specific *GFP* dsRNA-treated controls. A majority of expected pathway regulators were identified as hits (labeled in green) outside of the cutoff margins (red lines). (B) Functional distribution of validated hits, based on GO term annotation of *Drosophila* genes or their predicted homologs.

Figure S3.1 Primary Screen and Screening Strategy

(A) Primary screen hits are submitted to validation screening steps to eliminate false positives. Remaining candidates were submitted to secondary screens to assess the position of the regulatory input relative to known pathway components (epistasis) and specificity to the RAS/MAPK signaling context. In addition to this, candidates were screened for their impact on core RAS/MAPK component expression, both at the transcript (qPCR and RT-PCR) and protein level (quantitative immunofluorescence and Western blot). (B) The robustness of the primary screen assay was evaluated by monitoring changes in RAS^{V12}-induced pathway activity following knockdown of *mek* and *PTP-ER*. The levels of pMAPK are then measured by quantitative microscopy. Results shown for each dsRNA are the mean of 43 sample wells in 3 separately prepared plates. The calculated Z'-factor was 0.643 for *mek* depletion and 0.175 for *PTP-ER* depletion. (C) Distribution of primary screen probe data organized in plate screening order (grey data points). *GFP* dsRNA (blue) was used as a negative control and reference to normalize screening plate data. *mek* and *PTP-ER* dsRNAs (green and red, respectively) were used as positive controls to verify dsRNA knockdown efficiency.

A

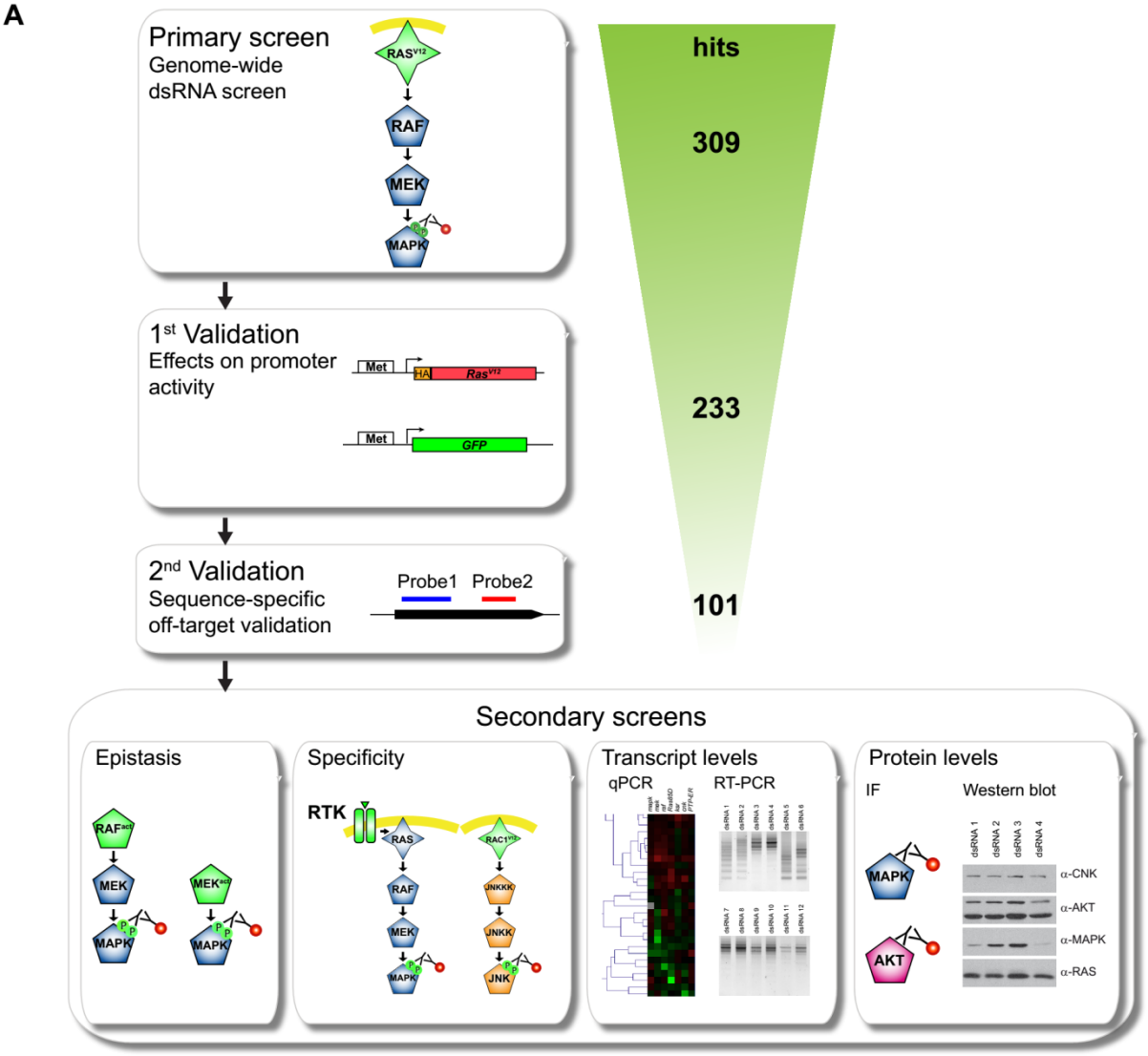

Primary screen
Genome-wide
dsRNA screen
RAS
RAF
MEK
MAPK
hits
309
1st Validation
Effects on promoter
activity
Met
HA
Met
GFP
233
2nd Validation
Sequence-specific
off-target validation
Probe1
Probe2
101
Secondary screens
Epistasis
Specificity
RTK
RAS
RAF
MEK
MAPK
JNKKK
JNKK
JNK
Transcript levels
qPCR
RT-PCR
Protein levels
IF
Western blot
AKT
α-CNK
α-AKT
α-MAPK
α-RAS

Figure S3.2 Promoter Validation Screens

(A - B) *pMet* promoter validation screen assays and control experiments. (A) Ras^{V12} expression in *pMet-HA-*Ras^{V12} stably transfected S2 cells is monitored by quantitative immunofluorescence through the use of anti-HA antibody. Values shown are the average HA signals of duplicate samples normalized to $CuSO_4$ induced controls with no dsRNA treatment. (B) GFP expression in *pMet-GFP* stably transfected S2 cells is monitored by quantitative immunofluorescence. Values shown are the average GFP signals of duplicate samples normalized to $CuS0_4$ induced controls with no dsRNA treatment. (C) Variation of pMAPK, GFP and HA signal in response to increasing amounts of *MTF-1* dsRNA. Values shown are average of triplicate samples normalized to $CuSO_4$ induced controls with no dsRNA treatment. (D - E) Promoter validation screen results (x axis) plotted against pMAPK primary screen values (y axis). The cutoff (dashed line) to identify false positive candidates with effects on promoter activity (data points in red areas) is a function of the pMAPK signal observed in the primary screen (see suppl. methods). MAPK regulators (red), STRIPAK (blue), EJC (green) as well as *Sur-8*, *Pp1-87B*, *Fnta* and *Prp19* (black) are shown. Factors known to be involved in *pMet*-driven expression, such as *MTF-1*, TBP-associated factors (Tafs) and *Ctr1B* are also shown (grey). (D) *pMet-HA-*Ras^{V12} validation screen results. The HA signal from the validation screen (x axis) plotted against the pMAPK signal from the primary screen (y axis). The HA and pMAPK values shown are normalized to plate-specific *GFP* dsRNA controls. (E) *pMet-GFP* validation screen results. The GFP signal from the validation screen (x axis) plotted against the pMAPK signal from the primary screen (y axis). The GFP values shown are normalized to plate-specific *mek* dsRNA controls. The pMAPK values shown are normalized to plate-specific *GFP* dsRNA controls.

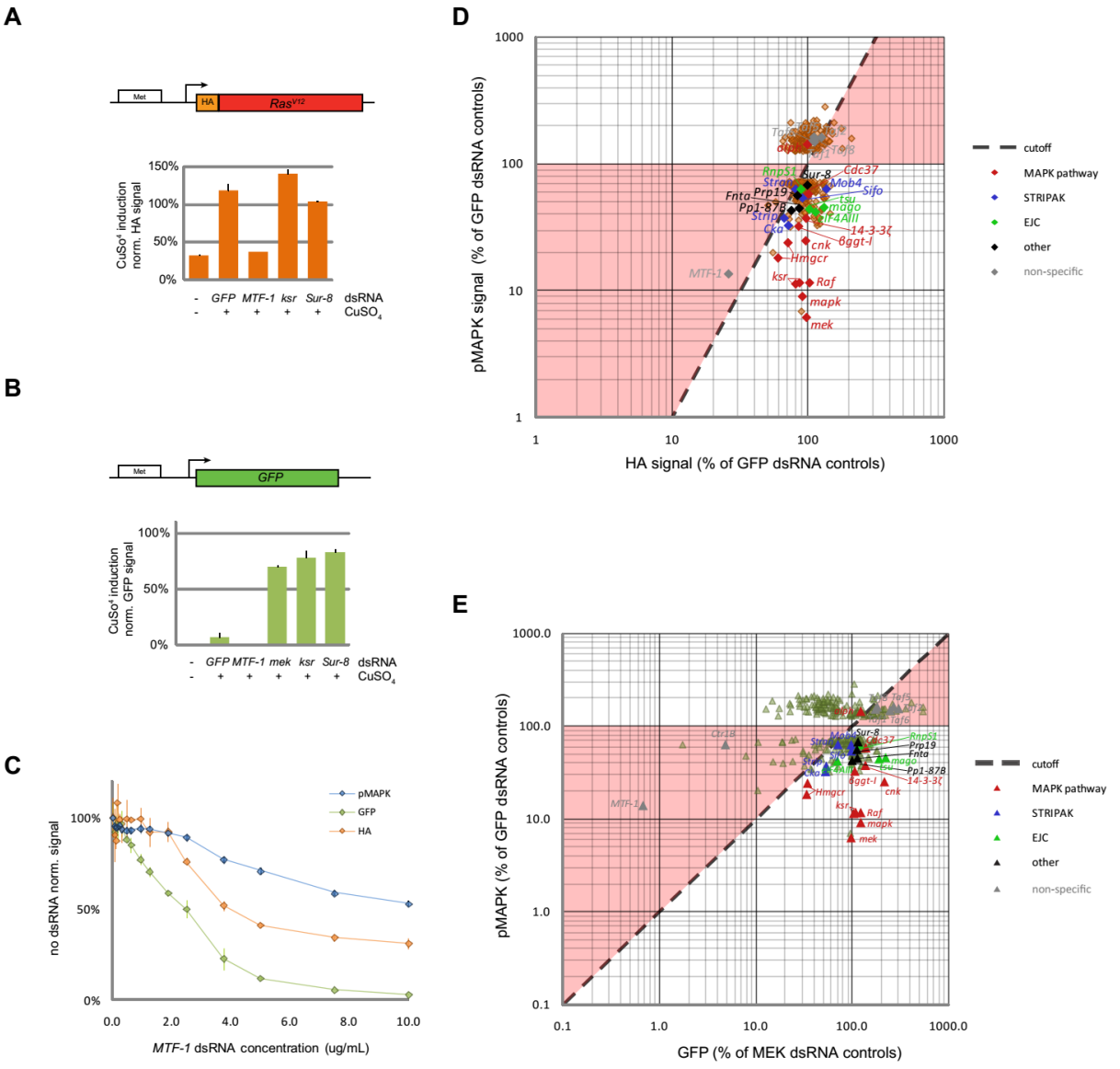

A
Met
HA
Ras^V12
CuSo4 induction norm. HA signal
150%
100%
50%
0%
- GFP MTF-1 ksr Sur-8 dsRNA
- + + + + CuSO4
B
Met
GFP
CuSo4 induction norm. GFP signal
100%
50%
0%
- GFP MTF-1 mek ksr Sur-8 dsRNA
- + + + + CuSO4
C
no dsRNA norm. signal
pMAPK
GFP
HA
100%
50%
0%
0.0 2.0 4.0 6.0 8.0 10.0
MTF-1 dsRNA concentration (ug/mL)
D
pMAPK signal (% of GFP dsRNA controls)
1000
100
10
1
1 10 100 1000
HA signal (% of GFP dsRNA controls)
Rnp55
Sur-8
Cdc37
Mob4
Sifo
tsu
mago
14-3-3ζ
cnk
Hmgcr
ksr
Raf
mapk
mek
MTF-1
cutoff
MAPK pathway
STRIPAK
EJC
other
non-specific
E
pMAPK (% of GFP dsRNA controls)
1000.0
100.0
10.0
1.0
0.1
0.1 1.0 10.0 100.0 1000.0
GFP (% of MEK dsRNA controls)
Sur-8
Cdc37
Rnp55
Prp19
Fnta
tsu
mago
Pp1-87B
14-3-3ζ
cnk
Hmgcr
ksr
Raf
mapk
mek
MTF-1
cutoff
MAPK pathway
STRIPAK
EJC
other
non-specific

Table S3.I Primary and Promoter Validation Screen Data (MS Excel file)

Primary screen and validation screen data. The average pMAPK observed in the *pMet-Ras^{V12}* primary screen (from the primary and confirmation steps) are listed for the 309 initial primary screen hits. The associated values and results for the first of two validation screening steps (elimination of candidates impacting the *pMet* expression system using *pMet-GFP* and *pMet-HA-Ras^{V12}*) are also shown.

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

GO accession	GO term	genes	p-value
GO:0006397	**mRNA processing**	35	1.1E-24
GO:0008380	**RNA splicing**	33	1.9E-24
GO:0006396	**RNA processing**	37	1.6E-23
GO:0022008	**neurogenesis**	51	4.2E-20
GO:0030154	**cell differentiation**	59	3.9E-18
GO:0007049	**cell cycle**	38	2.8E-16
GO:0010467	**gene expression**	56	8.7E-14
GO:0000226	**microtubule cytoskeleton organization**	26	5.6E-13
GO:0006996	**organelle organization**	35	3.2E-07
GO:0009790	**embryo development**	19	1.1E-04
GO:0051716	**cellular response to stimulus**	32	8.2E-04
GO:0051298	**centrosome duplication**	7	1.8E-03
GO:0006974	**response to DNA damage stimulus**	11	3.1E-03

Table S3.II GO Term Enrichment

Gene ontology terms found to be enriched in our set of validated hits using the Flymine[129] GO enrichment function. P values are calculated using Benjamini and Hochberg multiple hypothesis test.

3.4.2 Secondary screens: epistasis

To further characterize the 101 candidate genes, we conducted a series of secondary screens that can be subdivided into three groups: 1) MAPK activation induced by stimuli upstream of RAS, 2) epistasis screens involving MAPK activation at the level of or downstream of RAS, and 3) JNK activation screens aimed at addressing specificity to the MAPK pathway context (Figure S3.3; detailed results in Table S3.III, Table S3.IV and Table S3.V).

Four distinct MAPK activation assays using RTKs or *GAP* RNAi (Figure S3.3) were conducted to assess the degree to which pathway activity could be perturbed by depletion of the candidate genes in different activation contexts occurring upstream of RAS. Although a few exceptions were found, in most cases, we observed signal modulation which was generally consistent with our RAS^{V12} results.

Next, epistasis experiments were carried out using S2 cell lines expressing either constitutively activated forms of RAF or MEK (Figure 3.2 and Figure S3.3). The aim of these experiments was to position the identified genes in relation to the core kinases of the pathway by comparing the values from the RAS, RAF and MEK activation assays. To do this, we calculated the correlation of our screening data with theoretical profiles of hypothetical components acting within three possible epistasis intervals (RAS-RAF, RAF-MEK and MEK-MAPK) using an *Uncentered Pearson's Correlation* metric (Figure 3.2C and Table S3.III). All known pathway components were positioned correctly by this approach. For example, *Ras85D*, *ksr*, *cnk*, *hyp/ave*, *14-3-3ζ*, *14-3-3ε,* and *βggt-I* are part of a group of genes that suppressed RAS^{V12}, but not activated RAF or activated MEK; these components were thus correctly positioned in the RAS-RAF interval (Figure 3.2B and C). Nine additional genes also fell into this category (Table S3.III) and therefore represent potentially novel pathway regulators acting at this level. Strikingly, while only eight hits (including the RAF chaperone Cdc37) mapped between RAF and MEK, most of the candidates (69) were assigned to the MEK-MAPK interval (Table S3.III), with the majority of these being factors not previously linked to RAS/MAPK signaling. This unexpected finding suggests that additional regulatory events that escaped prior detection are lying downstream of MEK.

RAC1^{V12} and peptidoglycan (PGN) were then used as stimuli in two JNK activation assays as a proxy to evaluate specificity to the RAS/MAPK signaling context (Figure S3.3). Very few of the candidates modulated pJNK to a similar extent as they did pMAPK (Table S3.IV). One of these was the ALPH PP2C phosphatase, whose depletion increased both pMAPK and pJNK signals. This is consistent with our recent findings demonstrating that ALPH negatively regulates both MAPK and JNK signaling[23,47]. Remarkably, the vast majority of the RNA processing factors identified in the primary screen did not modulate pJNK levels and thereby argued for their specific role in RAS/MAPK signaling (Table S3.IV).

3.4.3 Predicted protein complexes have similar functional profiles

We next submitted our secondary screen data to unsupervised hierarchical clustering to group candidates with similar profiles together (Figure 3.3A). *Bona fide* pathway components with similar functions are clearly grouped together by this analysis. For example, *Ras85D*, *ksr*, *cnk* and *hyp* all act at the level of RAF activation and all show very similar profiles. Both 14-3-3 isoforms, which also act at this level, are grouped together and are also close to the first group of genes involved in RAF activation, as is *βggt-I*, a component involved in RAS prenylation. Based on this, we can expect that candidates who have a similar profile to *bona fide* RAS/MAPK pathway components might in fact share the same function as these components.

Next, we sought to identify putative protein complexes as well as related factors in our set of candidates by constructing an interaction network (PIN) based on publicly available protein and genetic interaction data (Figure 3.3B). The canonical RAS/MAPK components are clearly grouped together in this network and at least two other complexes can be clearly distinguished. The first consists of components of the STRIPAK complex and the second is composed primarily of mRNA processing factors. Remarkably, the components of both complexes also group together in similar functional profiles in our clustering analysis (groups "1" and "2" in both panels of Figure 3.3).

Figure 3.3 Predicted protein complexes have similar secondary screen functional profiles

(A) Unsupervised hierarchical clustering of secondary screen results. The 13 secondary assays are listed at the top of the clustering diagram. *Bona fide* MAPK pathway components are identified with an asterisk. (B) Protein interaction network (PIN) assembled using interaction data from predicted homologs of medium and high specificity hits. Edge color represents the source of the interaction data and edge width denotes the number of distinct experimental evidences for a given interaction. The coloring of the node border represents RAS^{V12} screen results while the coloring of the node center reflects MAPK protein levels. Node shapes reflect the functional category of the hits and specificity results are represented by the size of nodes. In both panels, epistasis results are represented by the coloring of gene symbols in orange (RAS-RAF), magenta (RAF-MEK), dark blue (MEK-MAPK) or grey (ambiguous). MAPK pathway components, STRIPAK (shaded area 1) and splicing factors (shaded area 2) group together in both the clustering analysis and PIN. Note that the full name is used for the gene raspberry, instead of *ras*, to avoid confusion with *Ras85D*.

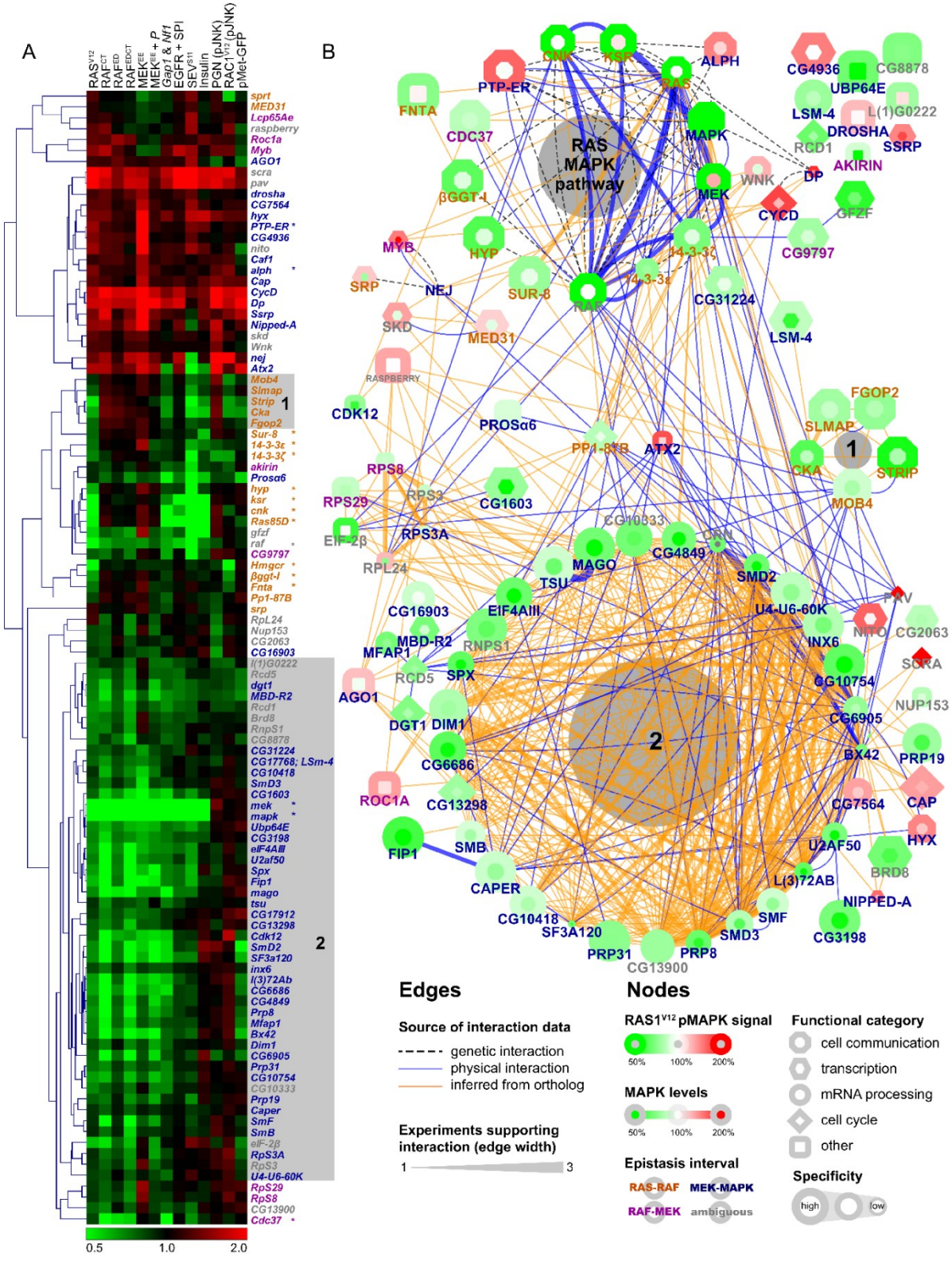

A
B
RAS MAPK pathway
Edges
Source of interaction data
genetic interaction
physical interaction
inferred from ortholog
Experiments supporting interaction (edge width)
Nodes
RAS1V12 pMAPK signal
MAPK levels
Epistasis interval
RAS-RAF
MEK-MAPK
RAF-MEK
ambiguous
Functional category
cell communication
transcription
mRNA processing
cell cycle
other
Specificity
high
low

Table S3.III Epistasis Screen Results and Calculated Pearson Correlations (MS Excel file)

Results from the secondary screens used in the epistasis analysis are shown (screens # 1-6). Values shown are $\log_{10}$ transformed normalized pMAPK signals. Calculated *Pearson Uncentered Correlation* (*r*) are shown for the three epistasis profiles (**RAS-RAF** =[1 0 0 0 0 0], **RAF-MEK** =[1 1 1 1 0 0], **MEK-MAPK** =[1 1 1 1 1 1]). The final epistasis interval is also shown as well as the confidence score for the epistasis result. Candidates that were not positioned are marked with "?" while those with ambiguous positioning (confidence score < 0.2) results are marked with "(A)".

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

Symbol	Functional group	dsRNA ID	times scored as hit in DRSC screens	WB result	WB depletion strength	MAPK AKT score	WB score	Nuc Export score	cell count score	pMet promoter score	JNK score	epistasis score	Combined 2ry screen specificity score	Total specificity score	Specificity group
14-3-3ε	cell communication	RNATH0478	8	N		0	0		0	0	1	0	1	9	L
14-3-3ζ	cell communication	RNATH0031	1	NS	W	0	1		0	0	1	0	1	3	H
AGO1	RNAi	RNATH0224	5	N		0	0		0	0	0	0	0	5	H
akirin	other	RNATH0489	7	MAPK	S	1	-1	0	0	0	1	0	1	8	M
alph	cell communication	RNATH0221	3	N		0	0		0	1	1	0	2	5	H
Atx2	structure	RNATH0227	7	NS	M	0	1		0	2	0.5	0	2.5	10.5	L
Brd8	transcription	RNATH0232	1	MAPK	W	-1	-1	0	0	1	0	0	1	1	H
Bx42	mRNA processing	RNATH0234	10	PS	S	0	1	1	0	2	0	0	2	14	L
Caf1	transcription	RNATH0236	3	N		1	0		0	1	1	0	2	6	M
Cap	cell cycle	RNATH0238	0	N		1	0		0	1	1	0	2	3	H
Caper	mRNA processing	RNATH0251	2	MAPK	S	-1	-1	0	0	0	0	0	0	1	H
Cdc37	cell communication	RNATH0240	0	N		1	0		0	0	0.5	0	0.5	1.5	H
Cdk12	transport	RNATH0325	7	MAPK	S	1	-1	0	1	1	1	0	2	10	L
CG10333	mRNA processing	RNATH0243	1	PS	S	-1	1	1	1	0	0	0	0	3	H
CG10418	mRNA processing	RNATH0245	3	MAPK	S	0	-1	1	0	0	0	0	0	3	H
CG10754	mRNA processing	RNATH0247	1	MAPK	S	-1	-1	1	1	0	0	0	0	1	H
CG13298	cell cycle	RNATH0261	3	NS	M	0	1		1	0	0.5	0	0.5	5.5	M
CG13900	mRNA processing	RNATH0265	5	MAPK	S	-1	-1	1	0	0	0	0	0	4	H
CG1603	transcription	RNATH0275	0	MAPK	S	-1	-1	0	0	0	0	0	0	1	H
CG16903	mRNA processing	RNATH0278	3	MAPK	W	0	-1	1	1	0	0.5	0	0.5	4.5	H
CG17768/ LSm-4	mRNA processing	RNATH0514	0	PS	W	0	1	1	0	0	0.5	0	0.5	2.5	H
CG17912	transcription	RNATH0283	1	MAPK	M	-1	-1	0	0	1	0	0	1	1	H
CG2063	transcription	RNATH0285	5	MAPK	S	0	-1	1	0	0	0.5	0	0.5	5.5	M
CG31224	transcription	RNATH0296	0	N		0	0		0	2	0.5	0	2.5	2.5	H
CG3198	mRNA processing	RNATH0299	0	MAPK	W	-1	-1	0	0	0	0	0	0	1	H
CG4849	mRNA processing	RNATH0309	4	MAPK	S	-1	-1	1	1	0	0	0	0	4	H
CG4936	transcription	RNATH0310	1	N		0	0		0	0	0	0	0	1	H
CG6686	mRNA processing	RNATH0315	3	MAPK	S	-1	-1	1	1	0	0	0	0	3	H
CG6905	mRNA processing	RNATH0320	4	MAPK	S	1	-1	1	1	2	0.5	0	2.5	8.5	L
CG7564	mRNA processing	RNATH0324	5	N		0	0	0	0	1	0.5	0	1.5	6.5	M
CG8878	other	RNATH0331	3	MAPK	M	-1	-1	0	0	1	0	0	1	2	H
CG9797	transcription	RNATH0335	2	N		0	0		0	0	0	0	0	2	H
Cka	cell communication	RNATH0184	5	RAS	W	0	-1		0	0	0.5	0	0.5	4.5	H
cnk	cell communication	RNATH0128	3	CNK	S	0	1		0	0	0	0	0	4	H
crn	mRNA processing	RNATH0525	10	?			1	1	0			0	0	12	L
CycD	cell cycle	RNATH0346	4	N		0	0		0	1	1	0	2	6	M
dgt1	cell cycle	RNATH0284	3	NS	W	0	1		0	1	1	0	2	6	M
Dim1	mRNA processing	RNATH0294	2	MAPK	W	0	-1	1	0	0	0	0	0	2	H
Dp	transcription	RNATH0351	9	NS	W	1	1		0	2	1	0	3	14	L
drosha	RNAi	RNATH0352	3	N		0	0		0	1	0	0	1	4	H

eIF-2β	translation	RNATH0508	7	NS	W	0	1		0	0	0	0	0	8	**M**
eIF4AIII	mRNA processing	RNATH0509	5	MAPK	S	-1	-1	0	0	0	0	0	0	3	**H**
Fgop2	cell communication	RNATH0241	1	N		0	0		0	0	0	0	0	1	**H**
Fip1	mRNA processing	RNATH0248	3	MAPK	S	-1	-1	0	0	0	0	0	0	1	**H**
Fnta	lipid metabolism	RNATH0286	0	RAS shift		0	-1		0	0	0.5	0	0.5	1	**H**
gfzf	transcription	RNATH0361	5	MAPK	M	-1	-1	0	0	0	0	0	0	3	**H**
Hmgcr	lipid metabolism	RNATH0378	9	RAS shift		0	-1		0	0	0.5	0	0.5	8.5	**L**
hyp	cell communication	RNATH0142	1	N		0	0		0	0	0	0	0	1	**H**
hyx	cell communication	RNATH0381	4	N		0	0		1	1	1	0	2	7	**M**
inx6	mRNA processing	RNATH0318	2	MAPK	M	-1	-1	1	0	0	0	0	0	1	**H**
ksr	cell communication	RNATH0129	3	N		0	0		0	0	0	0	0	3	**H**
l(1)G0222	other	RNATH0388	0	NS	S	1	1		0	2	1	0	3	5	**H**
l(3)72Ab	mRNA processing	RNATH0312	11	MAPK	S	-1	-1	1	1	0	0	0	0	11	**L**
Lcp65Ae	structure	RNATH0390	0	N		0	0		0	0	0.5	0	0.5	1	**H**
mago	mRNA processing	RNATH0185	1	MAPK	S	-1	-1	0	0	0	0	0	0	1	**H**
mapk	cell communication	RNATH0025	5	MAPK	S	-1	-1	0	0	0	0	0	0	3	**H**
MBD-R2	transcription	RNATH0400	4	NS	M	0	1		0	1	1	0	2	7	**M**
MED31	transcription	RNATH0402	4	N		0	0		0	0	1	0	1	5	**H**
mek	cell communication	RNATH0024	3	N		1	0		0	0	0	0	0	4	**H**
Mfap1	mRNA processing	RNATH0242	10	MAPK	S	-1	-1	1	0	0	0	0	0	9	**L**
Mob4	cell communication	RNATH0306	3	RAS	W	0	-1		0	0	0.5	0	0.5	2.5	**H**
Myb	transcription	RNATH0409	7	NS	W	1	1		1	0	1	0	1	11	**L**
nej	transcription	RNATH0516	12	NS	S	1	1		1	2	1	0	3	18	**L**
Nipped-A	transcription	RNATH0412	10	NS	S	0	1		0	2	0.5	0	2.5	13.5	**L**
nito	transcription	RNATH0413	5	N		0	0		0	0	0	0	0	5	**H**
Nup153	transport	RNATH0417	9	NS	S	0	1		0	0	0.5	0	0.5	10.5	**L**
pav	cell cycle	RNATH0421	10	N		1	0		1	1	1	0	2	14	**L**
Pp1-87B	cell cycle	RNATH0182	5	N		0	0		1	0	0.5	0	0.5	6.5	**M**
Prosα6	ubiquitin/proteasome	RNATH0425	4	NS	M	0	1		0	1	1	0	2	7	**M**
Prp19	mRNA processing	RNATH0360	2	MAPK	S	-1	-1	1	1	0	0	0	0	2	**H**
Prp31	mRNA processing	RNATH0319	0	MAPK	M	-1	-1	1	0	0	0	0	0	1	**H**
Prp8	mRNA processing	RNATH0426	8	MAPK	S	-1	-1	1	1	0	0	0	0	8	**M**
PTP-ER	cell communication	RNATH0048	1	N		0	0		0	0	0	0	0	1	**H**
raf	cell communication	RNATH0023	3	N		0	0		0	0	0	0	0	3	**H**
ras	transport	RNATH0428	3	N		0	0		0	0	0.5	0	0.5	3.5	**H**
Ras85D	cell communication	RNATH0022	7	RAS	S	0	-1		0	0	0	0	0	6	**M**
Rcd1	cell cycle	RNATH0329	4	N		0	0		0	2	1	0	3	7	**M**
Rcd5	cell cycle	RNATH0252	6	N		0	0		0	2	1	0	3	9	**L**
RnpS1	mRNA processing	RNATH0430	1	MAPK	W	-1	-1	0	0	0	0	0	0	1	**H**
Roc1a	ubiquitin/proteasome	RNATH0431	4	N		0	0		0	0	1	0	1	5	**H**
RpL24	translation	RNATH0433	10	NS	M	-1	1		0	0	0.5	1	1.5	11.5	**L**
RpS29	translation	RNATH0436	6	NS	M	0	1		0	0	0	0	0	7	**M**
RpS3	translation	RNATH0437	11	NS	S	0	1		0	1	1	0	2	14	**L**
RpS3A	translation	RNATH0438	12	NS	S	0	1		0	1	1	0	2	15	**L**
RpS8	translation	RNATH0439	9	NS	S	0	1		0	0	1	0	1	11	**L**
scra	cell cycle	RNATH0441	8	N		1	0		1	1	1	0	2	12	**L**
SF3a120	mRNA processing	RNATH0279	12	PS	S	-1	1	1	1	2	0	0	2	16	**L**

skd	transcription	RNATH0444	3	CNK	W	0	1		0	2	1	0	3	7	**M**
Slmap	cell communication	RNATH0282	1	RAS	W	0	-1		0	0	0.5	0	0.5	1	**H**
SmB	mRNA processing	RNATH0445	3	PS	M	0	1	1	0	0	1	0	1	6	**M**
SmD2	mRNA processing	RNATH0259	6	MAPK	S	-1	-1	1	1	2	0	0	2	8	**M**
SmD3	mRNA processing	RNATH0446	6	PS	M	-1	1	1	0	0	0	0	0	7	**M**
SmF	mRNA processing	RNATH0348	4	NS	M	-1	1	1	0	0	0	0	0	5	**H**
sprt	transport	RNATH0449	2	CNK	M	0	1		0	0	1	0	1	4	**H**
Spx	mRNA processing	RNATH0450	5	PS	S	-1	1	1	1	0	0.5	0	0.5	7.5	**M**
srp	transcription	RNATH0451	10	RAS	W	0	-1		0	0	0	0	0	9	**L**
Ssrp	transcription	RNATH0454	4	N		1	0		0	2	1	0	3	8	**M**
Strip	cell communication	RNATH0255	2	RAS	W	0	-1		0	0	0.5	0	0.5	1.5	**H**
Sur-8	cell communication	RNATH0213	0	N		0	0		0	0	0	0	0	1	**H**
tsu	mRNA processing	RNATH0460	0	MAPK	S	-1	-1	0	0	1	0	0	1	1	**H**
U2af50	mRNA processing	RNATH0463	7	PS	S	-1	1	0	1	0	0	0	0	8	**M**
U4-U6-60K	mRNA processing	RNATH0314	0	MAPK	M	-1	-1	1	0	0	0	0	0	1	**H**
Ubp64E	ubiquitin/proteasome	RNATH0464	0	MAPK	S	-1	-1	0	0	0	0	0	0	1	**H**
Wnk	cell communication	RNATH0321	6	N		1	0		0	0	0.5	0	0.5	7.5	**M**
βggt-I	cell communication	RNATH0228	0	RAS shift		0	-1		0	0	0.5	0	0.5	1	**H**

Table S3.IV Specificity Score Results

The data used to calculate the specificity score is presented in the following order: number of times this gene has been identified in previous *Drosophila* RNAi screens; result observed in the western blot screen (NS: non-specific effect, MAPK or RAS: specific effect on this component, - no effect) and associated specificity score; specificity score derived from the total mRNA nuclear export screen; score associated to a decreased cell count; effect on *pMet* driven expression derived from the *pMet-GFP* screen; JNK specificity result derived from the two pJNK secondary assays; combined specificity score derived from all secondary screen results; final specificity score and specificity group (high, medium, low). Specificity score calculation is described in the Text S1.

Table S3.V Summary of Secondary Screen Results (MS Excel file)

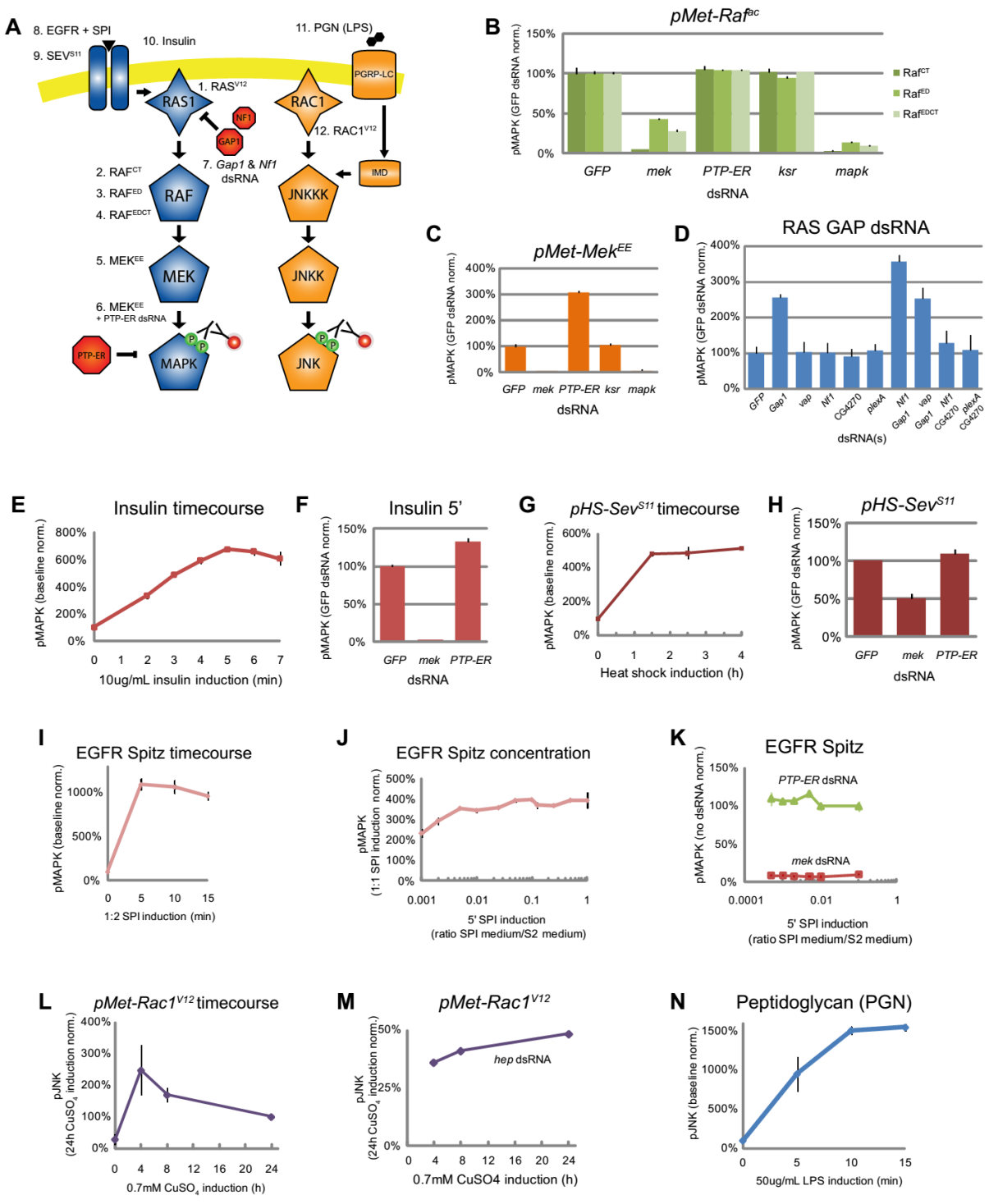
List of validated screen hits and secondary screen results. Values are expressed as $\log_2$ transformed ratios normalized to *GFP* dsRNA negative controls. Also presented in this table is: the calculated epistasis interval, specificity score, functional group (based on GO annotation) and predicted homologs in *H.sapiens* and *S.cerevisiae*. *Bona fide* RAS/MAPK components are marked with "*". Genes to which we have associated new gene symbols are marked with an "†".

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

Figure S3.3 Secondary Screens

(A) MAPK and JNK pathway models depicting the secondary screen assays used in this study. See Suppl. Methods for screen descriptions. (B - N) Secondary screen control assays. Values shown are duplicate sample averages normalized to *GFP* dsRNA treated controls unless otherwise indicated. (B) to (K) are MAPK activation assays with a pMAPK readout. (L) to (N) are JNK pathway assays with a pJNK readout. (B) RAF-based MAPK activation: Values shown are for *pMet-raf*^{ED}, *pMet-raf*^{CT} and *pMet-raf*^{EDCT} stably transfected cell lines treated with the indicated dsRNAs and induced with $CuSO_4$ for 24h. (C) MEK-based MAPK activation: dsRNA treated *pMet-mek*^{EE} stably transfected cells were induced with $CuSO_4$ for 24h. (D) RAS GAP RNAi-based MAPK activation: S2 cells were treated with dsRNAs targeting the indicated RAS GAPs or predicted RAS GAPs. Values are normalized to untreated control samples. Combined knockdown of *Gap1* and *Nf1* produced the highest pMAPK activation and was used in the secondary screen assay. (E & F) Insulin-based MAPK activation. (E) Insulin induction time course: The values shown are baseline normalized signal averages from single well samples of insulin treated S2 cells induced for the indicated times. (F) The signals shown are the average from dsRNA-treated single well samples of insulin induced S2 cells induced for 5'. (G & H) SEV^{S11}-based MAPK activation. (G) Heat shock induction time course of *pHS-sev*^{S11} stably transfected S2 cells. Samples are induced for 30' at 37°C and incubated at 27°C for the indicated times before sample preparation. Values shown are normalized to a non-induced control. (H) dsRNA-treated *pHS-sev*^{S11} cells were induced with a 30' heat shock at 37°C followed by a 2.5h incubation at 27°C. (I - K) EGFR-based MAPK activation assays. (I) Induction time course was performed on *pMet-Egfr* stably transfected cells induced with 1:2 supernatant from *pMet-Spi* cells. Values are normalized to non-induced controls. (J) Calibration of Spitz concentration: *pMet-Egfr* cells were induced for 5' with the indicated dilution of Spitz supernatant and normalized to 1:1 Spitz induced controls. (K) dsRNA-treated *pMet-Egfr* cells were induced for 5' with the indicated concentrations of Spitz supernatant. Values shown are normalized to samples not treated with dsRNA. (L & M) RAC1^{V12}-based JNK activation time courses. (L) Induction time course: *pMet-Rac1*^{V12} stably transfected cells were induced with $CuSO_4$ for the indicated times. Values shown are from triplicate samples of two experiments. (M) *pMet-Rac1*^{V12} cells were

treated with *hep* dsRNA and induced with $CuSO_4$ for the indicated times. Values shown are from single well samples normalized to samples not treated with dsRNA. (N) Peptidoglycan (PGN) based JNK activation time course: S2 cells were induced with 50μg/mL LPS for the indicated times. The values shown are averages of four samples from two separate experiments.


A
8. EGFR + SPI
9. SEV^{S11}
10. Insulin
11. PGN (LPS)
PGRP-LC
1. RAS^{V12}
RAS1
RAC1
NF1
GAP1
12. RAC1^{V12}
7. Gap1 & Nf1 dsRNA
IMD
2. RAF^{CT}
3. RAF^{ED}
4. RAF^{EDCT}
RAF
JNKKK
5. MEK^{EE}
MEK
JNKK
6. MEK^{EE} + PTP-ER dsRNA
PTP-ER
MAPK
JNK
B
pMet-Raf^{ac}
pMAPK (GFP dsRNA norm.)
150%
100%
50%
0%
Raf^{CT}
Raf^{ED}
Raf^{EDCT}
GFP
mek
PTP-ER
ksr
mapk
dsRNA
C
pMet-Mek^{EE}
pMAPK (GFP dsRNA norm.)
400%
300%
200%
100%
0%
GFP mek PTP-ER ksr mapk
dsRNA
D
RAS GAP dsRNA
pMAPK (GFP dsRNA norm.)
400%
300%
200%
100%
0%
GFP Gap1 vap Nf1 CG4270 plexA Nf1 Gap1 vap Gap1 Nf1 CG4270 plexA CG4270
dsRNA(s)
E
Insulin timecourse
pMAPK (baseline norm.)
800%
600%
400%
200%
0%
0 1 2 3 4 5 6 7
10ug/mL insulin induction (min)
F
Insulin 5'
pMAPK (GFP dsRNA norm.)
150%
100%
50%
0%
GFP mek PTP-ER
dsRNA
G
pHS-Sev^{S11} timecourse
pMAPK (baseline norm.)
600%
400%
200%
0%
0 1 2 3 4
Heat shock induction (h)
H
pHS-Sev^{S11}
pMAPK (GFP dsRNA norm.)
150%
100%
50%
0%
GFP mek PTP-ER
dsRNA
I
EGFR Spitz timecourse
pMAPK (baseline norm.)
1000%
500%
0%
0 5 10 15
1:2 SPI induction (min)
J
EGFR Spitz concentration
pMAPK (1:1 SPI induction norm.)
500%
400%
300%
200%
100%
0%
0.001 0.01 0.1 1
5' SPI induction (ratio SPI medium/S2 medium)
K
EGFR Spitz
pMAPK (no dsRNA norm.)
150%
100%
50%
0%
PTP-ER dsRNA
mek dsRNA
0.0001 0.01 1
5' SPI induction (ratio SPI medium/S2 medium)
L
pMet-Rac1^{V12} timecourse
pJNK (24h CuSO$_4$ induction norm.)
400%
300%
200%
100%
0%
0 4 8 12 16 20 24
0.7mM CuSO$_4$ induction (h)
M
pMet-Rac1^{V12}
pJNK (24h CuSO$_4$ induction norm.)
50%
25%
0%
hep dsRNA
0 4 8 12 16 20 24
0.7mM CuSO4 induction (h)
N
Peptidoglycan (PGN)
pJNK (baseline norm.)
1500%
1000%
500%
0%
0 5 10 15
50ug/mL LPS induction (min)

3.4.4 Regulation of RAS/MAPK pathway gene expression and specificity

Given that several of our candidate genes are linked to RNA processing and transcription, we hypothesized that these factors might be acting on the expression of one or multiple RAS/MAPK pathway components. We first investigated the impact of our candidates on the expression of core RAS/MAPK pathway component transcripts (*Ras85D, raf, mek, mapk, ksr, cnk,* and *PTP-ER*) by qPCR (Figure S3.4A and Table S3.V). A specific effect on the *mapk* transcript was observed upon depletion of *mago* and *eIF4AIII* (Figure 3.4A, Table S3.V and Table S3.VI) as we have previously reported[41]. At least three other factors (*Cdk12*, *Fip1* and *CG1603*) also seemed to modulate the transcript levels of *mapk*. Moreover, two factors, *gfzf* and *CG4936*, were found to modulate the levels of *mek* and *PTP-ER*, respectively (Figure 3.4A and Table S3.VI). *CG1603*, *gfzf* and *CG4936* were subsequently tested in larval eye disc tissue where similar results were obtained (Figure 3.4B and Table S3.VII).

Surprisingly, aside from the candidates mentioned above, most hits did not appear to cause a significant change in the transcript levels of pathway components. We had previously observed that qPCR assays targeting *mapk* are not always strongly affected by the splicing changes induced by EJC depletion. On the other hand, an RT-PCR assay spanning the whole *mapk* transcript is a more sensitive tool allowing for detection of small splicing changes[41]. Based on this, and on the fact that almost all the splicing factors we identified mapped downstream of MEK, we decided to systematically examine the impact of these factors on *mapk* splicing. To do this, we used the RT-PCR assay that had been used with the EJC to examine *mapk* splicing. Interestingly, not only did this experiment reveal that nearly all the splicing factors in our set caused shifts in the *mapk* RT-PCR profile, but these RT-PCR profiles were clearly different from those produced by EJC depletion (Figure S3.5). Thus, while the impact of most of our candidates on *mapk* expression may not be apparent when measuring total transcript abundance, a clear impact on the different *mapk* isoforms can be observed by RT-PCR, indicating that these factors regulate AS of *mapk*.

In the case of the EJC, the splicing changes were accompanied by a corresponding decrease in MAPK protein levels. Because of this, we decided to also measure the impact of our candidates on MAPK protein levels using quantitative immunofluorescence. This analysis

confirmed that most of the factors positioned downstream of MEK (including most of the RNA processing factors) also caused a reduction of MAPK protein levels. Conversely, AKT protein levels, which were used as a control, were not generally sensitive to depletion of these same factors (Figure 3.5A). We also verified the impact on MAPK and two other pathway components (RAS and CNK) by Western blot. Most candidates that caused a reduction in MAPK levels did not impact the levels of RAS, CNK or AKT (Figure S3.6), which mirrored the results from the immunofluorescence experiment. Thus, both evaluations of MAPK protein levels agreed with the RT-PCR experiments suggesting that the changes in splicing results in a reduction of MAPK protein abundance.

Most of the other hits, including those positioned at the RAS-RAF and RAF-MEK intervals, did not appear to cause a change in the expression of RAS/MAPK pathway components. However, we observed that a minority of candidates seemed to cause fluctuations in multiple proteins or transcripts. Likewise, these candidates also tended to impact some or all of the non-RAS/MAPK related assays we tested them in (Table S3.V). Moreover, these hits were also more frequently present in hit lists of other published RNAi screens (Table S3.IV). Consequently, in order to discriminate such non-specific hits from higher quality candidates, we derived a scoring system that factored in results from the pJNK assays, pMet-GFP expression, hit occurrence in previously published RNAi screens and the impact on measured protein and transcript levels (see supplemental text and Table S3.IV). As expected, a few of the splicing factors in our list had low specificity scores. On the other hand, most of the factors that selectively affected MAPK levels, including the majority of the splicing factors, were not present in previous screen hit lists more frequently than *bona fide* RAS/MAPK factors and generally displayed a good specificity to the RAS/MAPK context (Table S3.IV and Table S3.V). This suggests that discriminating against an entire category of genes based on the enrichment of that category in previous screen hit sets is a strategy that can lead to elimination of meaningful candidates.

Figure S3.4. qPCR Screen

(A) Unsupervised hierarchical clustering of qPCR screen results. SYBR green qPCR assays were used to assess the levels of RAS/MAPK pathway transcripts (top labels) following depletion of the indicated Ras^{V12} screen candidates (labels, left). The transcript levels are presented as $\log_2$ transformed ratios of *GFP* dsRNA treated controls. RAS/MAPK control dsRNAs are indicated by asterisks. The *gfzf* dsRNA has the closest profile to the *mek* control dsRNA, with both reagents causing a reduction in *mek* levels. *PTP-ER* and *CG4936* also have very similar profiles and cause a reduction in *PTP-ER* transcript levels. As we previously reported, the EJC components (blue) *mago* and *eIF4AIII* cause an observable decrease in *mapk* transcript levels. However, the knockdown of *tsu* and *RnpS1*, which have a weaker impact on pMAPK and MAPK levels, does not cause a readily observable change in *mapk*. Only *CG1603* and *Cdk12* had similar profiles to the two stronger EJC components, causing a decrease in *mapk* transcript levels. The *Fip1* dsRNA results in the qPCR screen were not valid, but *Fip1* dsRNA was re-tested in the confirmation qPCR experiment presented in Figure 3.2 and found to have a similar profile to the two other factors. Most of the splicing factors – other than the two EJC components – were not found to cause changes in the levels of *mapk* or any of the other RAS/MAPK transcripts. (B) Total mRNA FISH performed using an oligo-dT Cy3-labeled probe. The proportion of cells displaying nuclear retention was evaluated by performing segmentation and scoring the overlap of the FISH signal with a DAPI nuclear stain. dsRNA targeting *sbr*, the homolog of the mammalian nuclear RNA export factor 1 (NXF1), and *CG2063*, a factor previously linked to mRNA export, both caused an increase in nuclear retention. In comparison *mago* dsRNA did not visibly alter total mRNA retention. (C) Total mRNA export screen results. A subset of 44 candidates were tested to evaluate their impact on total mRNA export: these included all factors with functions linked to mRNA splicing (red) as well as other factors found to alter MAPK protein or *mapk* mRNA levels. A nuclear retention index was calculated by normalizing the rate of nuclear retention to the average results of *GFP* dsRNA treated controls (green). *sbr* dsRNA was used as a positive control (red arrowhead) for nuclear retention. In addition to *sbr*, 5 candidates in our set (orange circles) were also identified in a mRNA nuclear export RNAi screen[90].

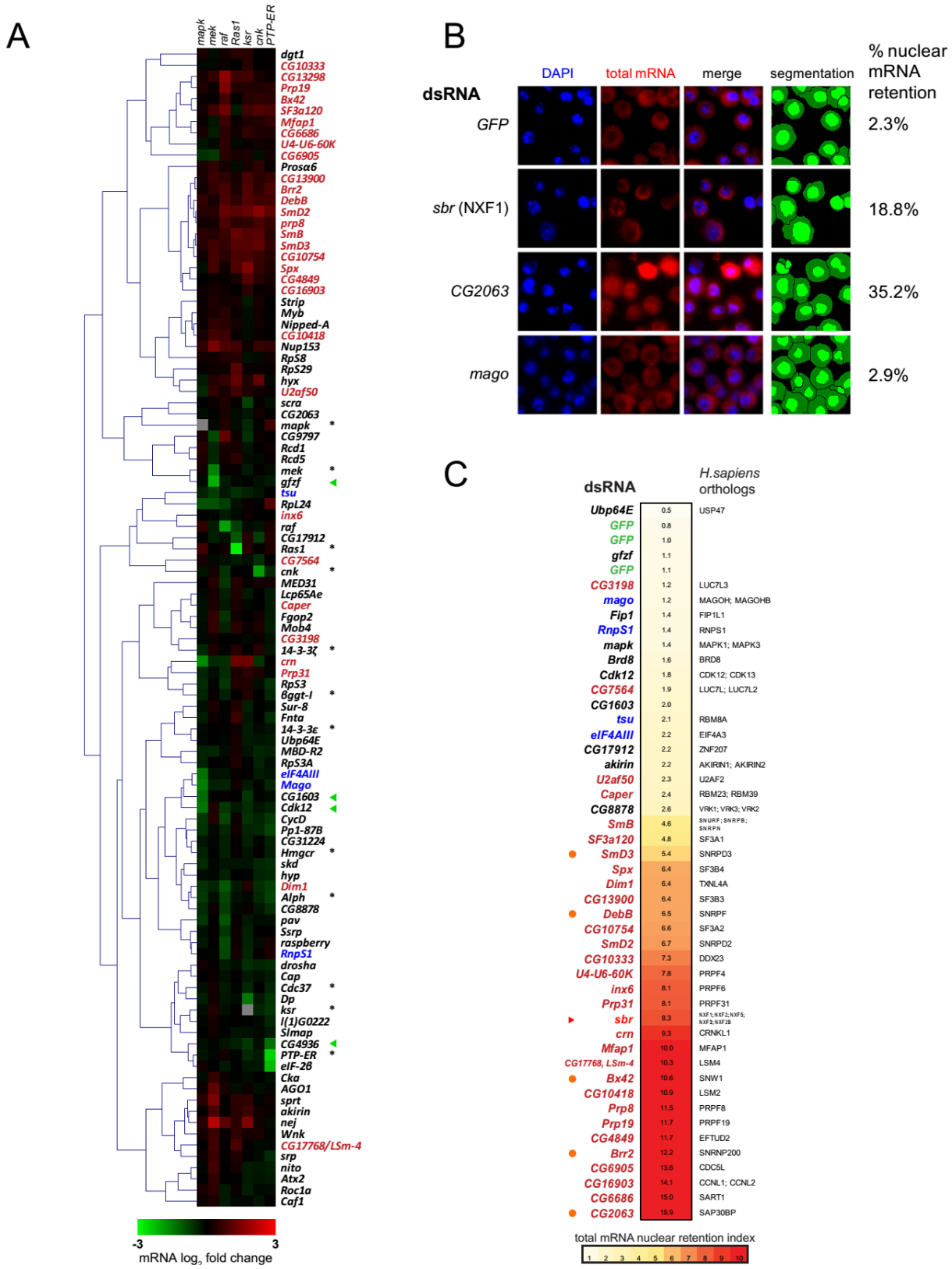
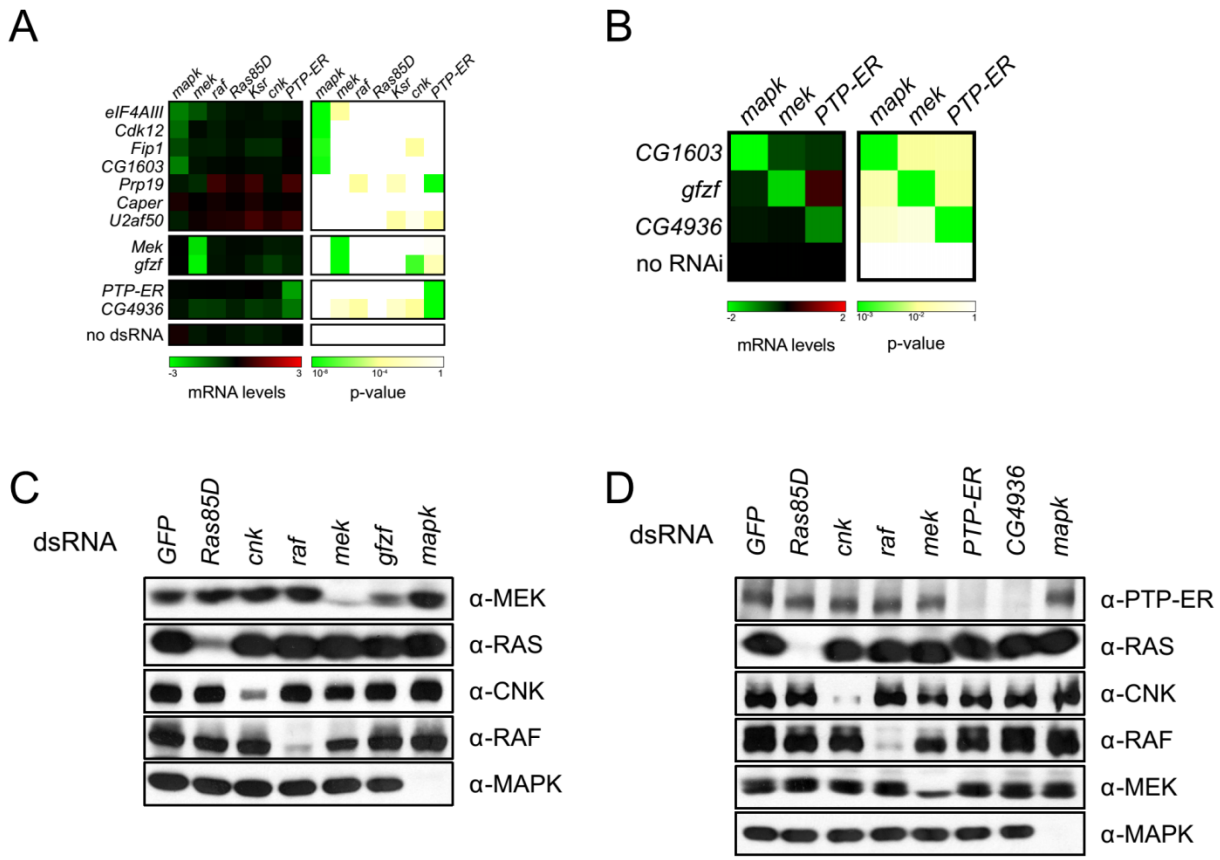

A
mapk
mek
raf
Ras1
ksr
cnk
PTP-ER
dgt1
CG10333
CG13298
Prp19
Bx42
SF3a120
Mfap1
CG6686
U4-U6-60K
CG6905
Prosα6
CG13900
Brr2
DebB
SmD2
prp8
SmB
SmD3
CG10754
Spx
CG4849
CG16903
Strip
Myb
Nipped-A
CG10418
Nup153
RpS8
RpS29
hyx
U2af50
scra
CG2063
mapk
CG9797
Rcd1
Rcd5
mek
gfzf
tsu
RpL24
inx6
raf
CG17912
Ras1
CG7564
cnk
MED31
Lcp65Ae
Caper
Fgop2
Mob4
CG3198
14-3-3ζ
crn
Prp31
RpS3
6ggt-I
Sur-8
Fnta
14-3-3ε
Ubp64E
MBD-R2
RpS3A
eIF4AIII
Mago
CG1603
Cdk12
CycD
Pp1-87B
CG31224
Hmgcr
skd
hyp
Dim1
Alph
CG8878
pav
Ssrp
raspberry
RnpS1
drosha
Cap
Cdc37
Dp
ksr
l(1)G0222
Slmap
CG4936
PTP-ER
eIF-2B
Cka
AGO1
sprt
akirin
nej
Wnk
CG17768/LSm-4
srp
nito
Atx2
Roc1a
Caf1
-3
3
mRNA log2 fold change
B
DAPI
total mRNA
merge
segmentation
% nuclear mRNA retention
dsRNA
GFP
2.3%
sbr (NXF1)
18.8%
CG2063
35.2%
mago
2.9%
C
dsRNA
H.sapiens orthologs
Ubp64E 0.5 USP47
GFP 0.8
GFP 1.0
gfzf 1.1
GFP 1.1
CG3198 1.2 LUC7L3
mago 1.2 MAGOH; MAGOHB
Fip1 1.4 FIP1L1
RnpS1 1.4 RNPS1
mapk 1.4 MAPK1; MAPK3
Brd8 1.6 BRD8
Cdk12 1.8 CDK12; CDK13
CG7564 1.9 LUC7L; LUC7L2
CG1603 2.0
tsu 2.1 RBM8A
eIF4AIII 2.2 EIF4A3
CG17912 2.2 ZNF207
akirin 2.2 AKIRIN1; AKIRIN2
U2af50 2.3 U2AF2
Caper 2.4 RBM23; RBM39
CG8878 2.6 VRK1; VRK3; VRK2
SmB 4.6 SNURF; SNRPB; SNRPN
SF3a120 4.8 SF3A1
SmD3 5.4 SNRPD3
Spx 6.4 SF3B4
Dim1 6.4 TXNL4A
CG13900 6.4 SF3B3
DebB 6.5 SNRPF
CG10754 6.6 SF3A2
SmD2 6.7 SNRPD2
CG10333 7.3 DDX23
U4-U6-60K 7.8 PRPF4
inx6 8.1 PRPF6
Prp31 8.1 PRPF31
sbr 8.3 NXF1; NXF2; NXF5; NXF3; NXF2B
crn 9.3 CRNKL1
Mfap1 10.0 MFAP1
CG17768, LSm-4 10.3 LSM4
Bx42 10.6 SNW1
CG10418 10.9 LSM2
Prp8 11.5 PRPF8
Prp19 11.7 PRPF19
CG4849 11.7 EFTUD2
Brr2 12.2 SNRNP200
CG6905 13.6 CDC5L
CG16903 14.1 CCNL1; CCNL2
CG6686 15.0 SART1
CG2063 15.9 SAP30BP
total mRNA nuclear retention index
1 2 3 4 5 6 7 8 9 10

Figure 3.4 Screen candidates modify the expression of RAS/MAPK components

(A) The transcript levels of RAS/MAPK components are altered by the depletion of some candidates. All candidates were tested in an initial qPCR secondary screen; the results shown here are from a separate qPCR confirmation experiment (see Text S1). The left panel shows transcript levels for the RAS/MAPK components listed on the top following treatment with the indicated dsRNAs (labels to the left). mRNA levels are expressed as $\log_2$ ratios of *GFP* dsRNA treated controls. The right panel shows the associated p-values (unpaired two-tailed Student's *t*-test). dsRNA targeting *gfzf* had a similar effect to the *mek* dsRNA with a -2.86 reduction in *mek* transcript levels (p-val. 4.2×10^{-8}). *CG4936* dsRNA caused a -1.37 reduction in *PTP-ER* transcript levels (p-val. 1.2×10^{-8}), which was slightly weaker than the -1.88 reduction (p-val. 3.0×10^{-9}) measured for the *PTP-ER* dsRNA . *Cdk12*, *Fip1* and *CG1603* dsRNAs behaved similarly to *eIF4AIII* dsRNA, used as a control for *mapk* transcript depletion, with *mapk* transcript levels < -0.75 and a p-values < 1×10^{-4}. (B) *In vivo* RNAi experiments confirm cell culture qPCR results. Hairpin RNAi constructs were expressed in larvae under the control of a heat shock-inducible flip-out actin promoter. qPCR experiments were performed on L3 eye disc lysates. mRNA levels are expressed as $\log_2$ ratios of a no RNAi control (flies carrying the flip-out promoter without a RNAi construct). Results were similar to those in cell culture qPCR experiments in (A): *CG1603* RNAi caused a reduction in *mapk* levels (-2.06; p-val. 5.3×10^{-4}), *gfzf* RNAi reduced *mek* levels (-1.66; p-val. 5.4×10^{-4}) and *CG4936* reduced *PTP-ER* levels (-1.06; p-val. 5.7×10^{-4}). (C and D) The levels of RAS/MAPK pathway components were evaluated by Western blot with the indicated antibodies (labels to right of panel) following treatment with the indicated dsRNA reagents (top labels). (B) Mirroring the qPCR data, a specific depletion of MEK levels was observed in *gfzf* dsRNA treated cells. (C) Also in agreement with the qPCR data, a specific decrease in PTP-ER levels was observed upon depletion of *CG4936*.


A
mapk
mek
raf
Ras85D
Ksr
cnk
PTP-ER
eIF4AIII
Cdk12
Fip1
CG1603
Prp19
Caper
U2af50
Mek
gfzf
PTP-ER
CG4936
no dsRNA
mRNA levels
p-value
B
mapk
mek
PTP-ER
CG1603
gfzf
CG4936
no RNAi
mRNA levels
p-value
C
dsRNA
GFP
Ras85D
cnk
raf
mek
gfzf
mapk
α-MEK
α-RAS
α-CNK
α-RAF
α-MAPK
D
dsRNA
GFP
Ras85D
cnk
raf
mek
PTP-ER
CG4936
mapk
α-PTP-ER
α-RAS
α-CNK
α-RAF
α-MEK
α-MAPK

RNAi	*mapk* (mRNA)	*mek* (mRNA)	*raf* (mRNA)	*Ras85D* (mRNA)	*ksr* (mRNA)	*cnk* (mRNA)	*PTP-ER* (mRNA)	*Act5C* (mRNA)	*RpL32* (mRNA)	p-val *mapk* (mRNA)	p-val *mek* (mRNA)	p-val *raf* (mRNA)	p-val *Ras85D* (mRNA)	p-val *ksr* (mRNA)	p-val *cnk* (mRNA)	p-val *PTP-ER* (mRNA)	p-val *Act5C* (mRNA)	p-val *RpL32* (mRNA)
no dsRNA	0.26	-0.26	-0.09	-0.23	-0.37	-0.25	-0.06	-0.12	0.12	3.2E-02	2.5E-01	5.8E-01	1.9E-01	8.5E-02	5.0E-02	6.0E-01	4.3E-01	3.6E-01
GFP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00
mapk		0.19	0.16	-0.02	0.36	-0.02	0.10	0.15	-0.15		3.4E-01	3.0E-01	9.0E-01	7.9E-02	8.3E-01	2.8E-01	3.4E-01	1.7E-01
mek	0.01	-2.60	-0.26	-0.10	-0.13	-0.31	-0.27	0.09	-0.09	9.2E-01	8.2E-08	1.1E-01	5.5E-01	5.6E-01	1.6E-02	9.0E-03	5.8E-01	4.1E-01
raf	0.04	-0.27	-2.10	-0.28	-0.29	-0.32	-0.01	0.06	-0.06	7.6E-01	1.8E-01	5.1E-08	1.3E-01	1.4E-01	1.4E-02	8.9E-01	6.9E-01	5.5E-01
Ras85D	-0.08	0.05	-0.32	-3.09	-0.43	-0.47	-0.15	0.31	-0.31	4.5E-01	7.9E-01	4.5E-02	4.2E-09	4.4E-02	5.9E-04	1.5E-01	8.5E-02	1.4E-02
Ksr	0.35	0.02	-0.14	-0.12		-0.13	-0.01	0.01	-0.01	7.8E-03	9.0E-01	3.7E-01	4.5E-01		2.0E-01	9.2E-01	9.2E-01	8.9E-01
cnk	-0.05	-0.15	-0.39	-0.21	-0.35	-2.53	-0.31	0.22	-0.22	6.2E-01	4.4E-01	1.8E-02	2.1E-01	8.1E-02	1.1E-10	1.3E-02	1.4E-01	5.1E-02
PTP-ER	-0.06	-0.03	-0.01	-0.05	0.03	-0.18	-1.88	0.13	-0.13	6.2E-01	8.7E-01	9.4E-01	7.9E-01	8.8E-01	1.1E-01	3.0E-09	4.3E-01	2.3E-01
eIF4AIII	-1.49	-0.98	-0.41	-0.19	-0.15	-0.29	-0.23	0.12	-0.12	6.0E-08	3.9E-04	1.6E-02	2.6E-01	4.3E-01	1.0E-02	2.0E-02	4.6E-01	2.7E-01
Prp19	-0.26	-0.49	0.71	0.28	0.70	-0.15	0.77	-0.07	0.07	3.1E-02	6.9E-02	4.7E-04	1.1E-01	3.2E-03	1.3E-01	2.7E-06	6.3E-01	5.2E-01
Caper	0.29	0.02	0.22	0.05	0.32	-0.04	0.08	0.03	-0.03	1.9E-02	9.2E-01	1.6E-01	7.4E-01	1.1E-01	6.9E-01	4.2E-01	8.5E-01	7.9E-01
U2af50	-0.33	0.16	0.33	0.38	0.80	0.42	0.73	-0.02	0.02	1.1E-02	4.0E-01	4.2E-02	4.2E-02	1.3E-03	8.2E-03	1.9E-04	8.7E-01	8.2E-01
Cdk12	-1.24	-0.11	-0.31	-0.12	-0.17	-0.18	-0.04	0.30	-0.30	3.5E-07	5.9E-01	5.2E-02	4.7E-01	3.7E-01	8.9E-02	6.9E-01	7.3E-02	1.3E-02
Fip1	-0.84	-0.33	-0.35	-0.16	-0.49	-0.50	0.05	0.20	-0.20	1.2E-05	1.3E-01	3.5E-02	3.3E-01	2.2E-02	6.6E-04	5.5E-01	1.9E-01	1.6E-01
CG1603	-1.48	-0.20	-0.10	-0.05	-0.15	-0.16	-0.03	0.10	-0.10	6.6E-06	3.4E-01	5.0E-01	7.7E-01	4.4E-01	1.2E-01	7.7E-01	5.1E-01	3.5E-01
gfzf	0.00	-2.86	-0.23	-0.12	-0.38	-0.70	-0.35	0.23	-0.23	9.8E-01	4.2E-08	1.3E-01	4.7E-01	6.1E-02	2.5E-05	2.8E-03	1.3E-01	5.5E-02
CG4936	-0.20	-0.69	-0.64	-0.36	-0.66	-0.55	-1.37	0.29	-0.29	8.1E-02	5.0E-03	1.1E-03	4.3E-02	4.5E-03	1.5E-04	1.2E-08	7.2E-02	2.0E-02
AGO1	0.74	0.86	0.37	-0.04	0.37	-0.46	0.00	-0.16	0.16	6.3E-05	1.3E-03	2.4E-02	8.3E-01	7.0E-02	1.4E-03	9.7E-01	3.0E-01	1.8E-01
sprt	0.36	0.41	0.01	0.24	-0.15	-0.68	0.13	0.05	-0.05	6.9E-03	5.5E-02	9.4E-01	1.6E-01	4.2E-01	1.9E-05	1.7E-01	7.0E-01	6.1E-01

Table S3.VI Detailed confirmation qPCR results

Values for the confirmation qPCR experiment presented in Figure 3.5A. mRNA levels are expressed as $\log_2$ ratios of negative control (*GFP* dsRNA treated S2 cells or no RNAi flies). The right panel shows the associated p-values (unpaired two-tailed Student's *t*-test). *AGO1* and *sprt* were not followed-up as the observed increase in *mek* mRNA levels in the initial qPCR secondary screen was not confirmed in this experiment. The dsRNA probes from our set targeting *mapk* and *ksr* overlap with the SYBR green qPCR probes invalidating those results. Both qPCR assays were validated with separate dsRNAs (data not shown).

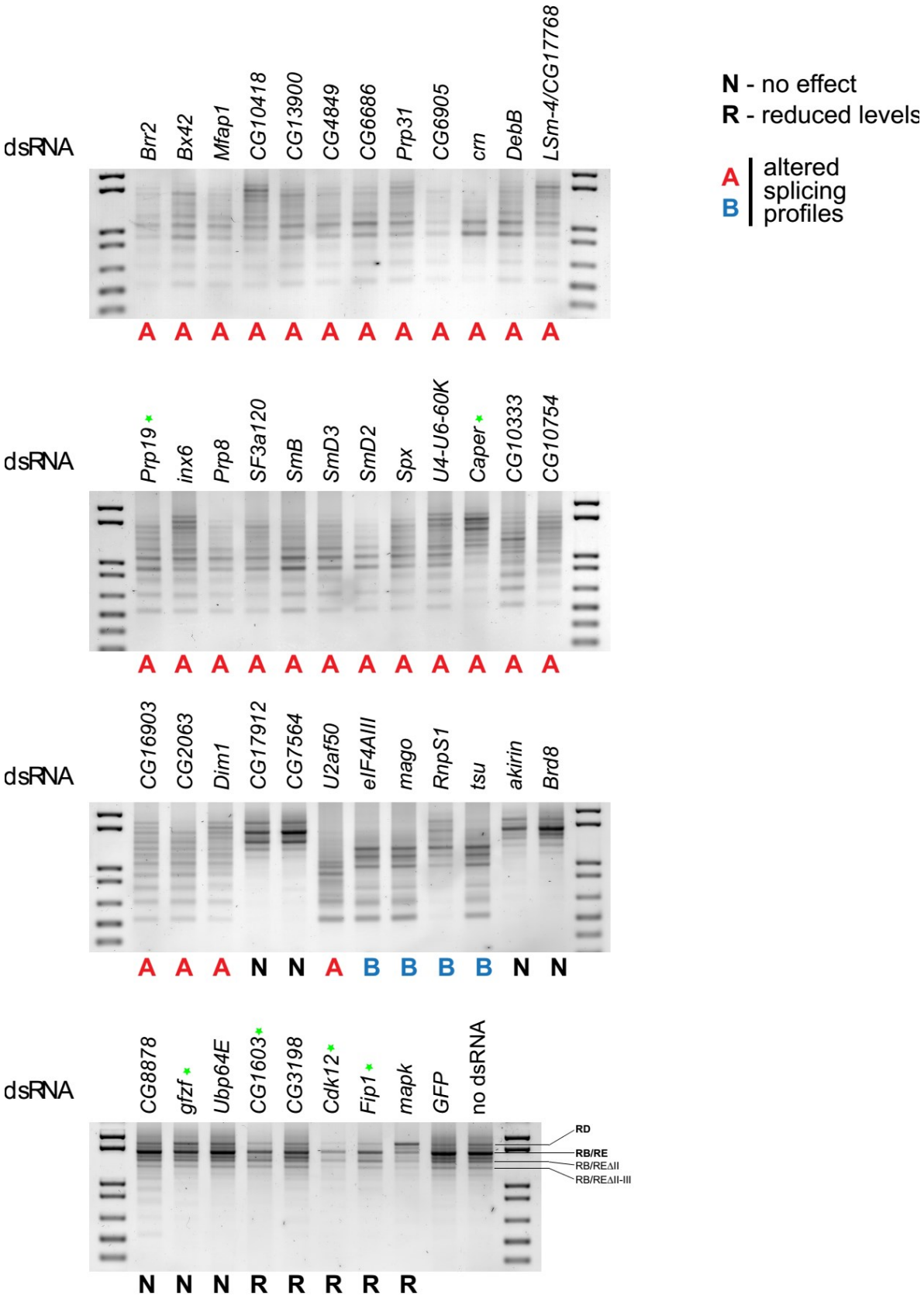
RNAi	*mapk* (mRNA)	*mek* (mRNA)	*PTP-ER* (mRNA)	*Act5C* (mRNA)	*RpL32* (mRNA)	p-val *mapk* (mRNA)	p-val *mek* (mRNA)	p-val *PTP-ER* (mRNA)	*p-val Act5C* (mRNA)	*p-val RpL32* (mRNA)
CG1603	-2.07	-0.55	-0.40	-0.37	0.41	5.3E-04	7.5E-02	1.2E-01	2.3E-02	1.9E-02
gfzf	-0.29	-1.66	0.47	-0.52	0.51	2.4E-01	5.4E-04	3.8E-02	2.2E-06	1.9E-04
CG4936	-0.18	-0.11	-1.06	-0.33	0.30	5.1E-01	6.8E-01	5.7E-04	1.4E-02	5.7E-02
hs-flipout (no RNAi)	0.00	0.00	0.00	0.00	0.00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00

Table S3.VII Detailed *in vivo* qPCR results

Values for the *in vivo* qPCR presented in Figure 3.5 B. mRNA levels are expressed as $\log_2$ ratios of negative control (*GFP* dsRNA treated S2 cells or no RNAi flies). The right panel shows the associated p-values (unpaired two-tailed Student's *t*-test).

Figure S3.5 RT-PCR Screen

An RT-PCR assay encompassing the entire *mapk* RB, RD, RE and RF transcripts was used to evaluate the impact on the distribution of *mapk* transcript isoforms. The same subset of 44 hits selected for the mRNA export screen were also tested here. A majority of factors tested caused a change in the *mapk* RT-PCR profile when compared to *GFP* dsRNA treated or untreated controls. The altered profiles were grouped into 2 categories (A and B) based on similarities in the sizes of the bands observed. Most splicing factors in the set produced an "A" type profile while EJC factors produced a different "B" type profile. Other candidates causing no obvious change in product size, but causing a reduction in band intensity are labeled "R". Those causing no observable change are labeled "N". The labels on the right of the bottom panel refer to *mapk* transcript isoforms that correspond to the bands visible on the gel. Candidates featured in the manuscript are labeled with a green star.


dsRNA
Brr2
Bx42
Mfap1
CG10418
CG13900
CG4849
CG6686
Prp31
CG6905
crn
DebB
LSm-4/CG17768
A A A A A A A A A A A A
dsRNA
Prp19 *
inx6
Prp8
SF3a120
SmB
SmD3
SmD2
Spx
U4-U6-60K
Caper *
CG10333
CG10754
A A A A A A A A A A A A
dsRNA
CG16903
CG2063
Dim1
CG17912
CG7564
U2af50
eIF4AIII
mago
RnpS1
tsu
akirin
Brd8
A A A N N A B B B B N N
dsRNA
CG8878
gfzf *
Ubp64E
CG1603 *
CG3198
Cdk12 *
Fip1 *
mapk
GFP
no dsRNA
RD
RB/RE
RB/REΔII
RB/REΔII-III
N N N R R R R R
N - no effect
R - reduced levels
A altered splicing profiles
B

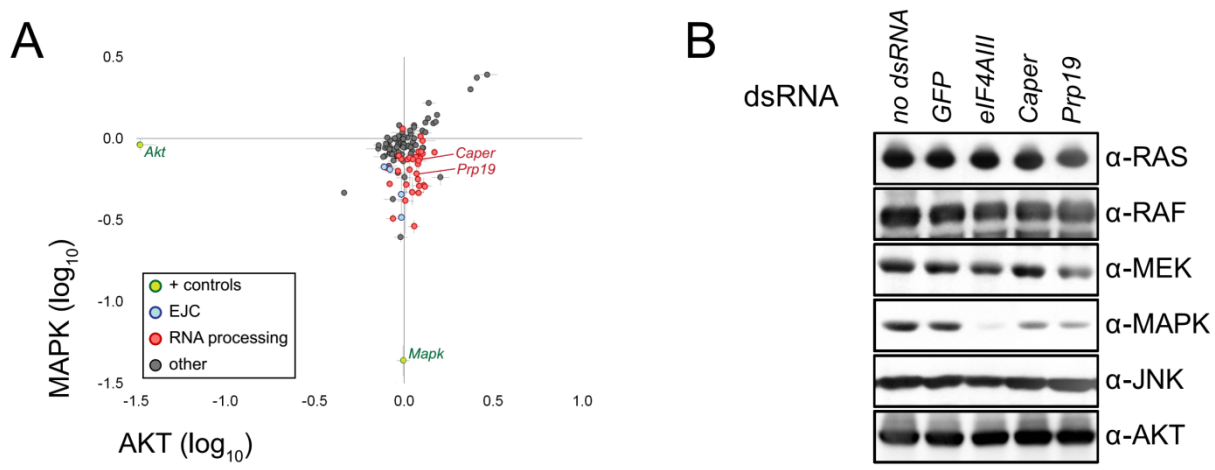
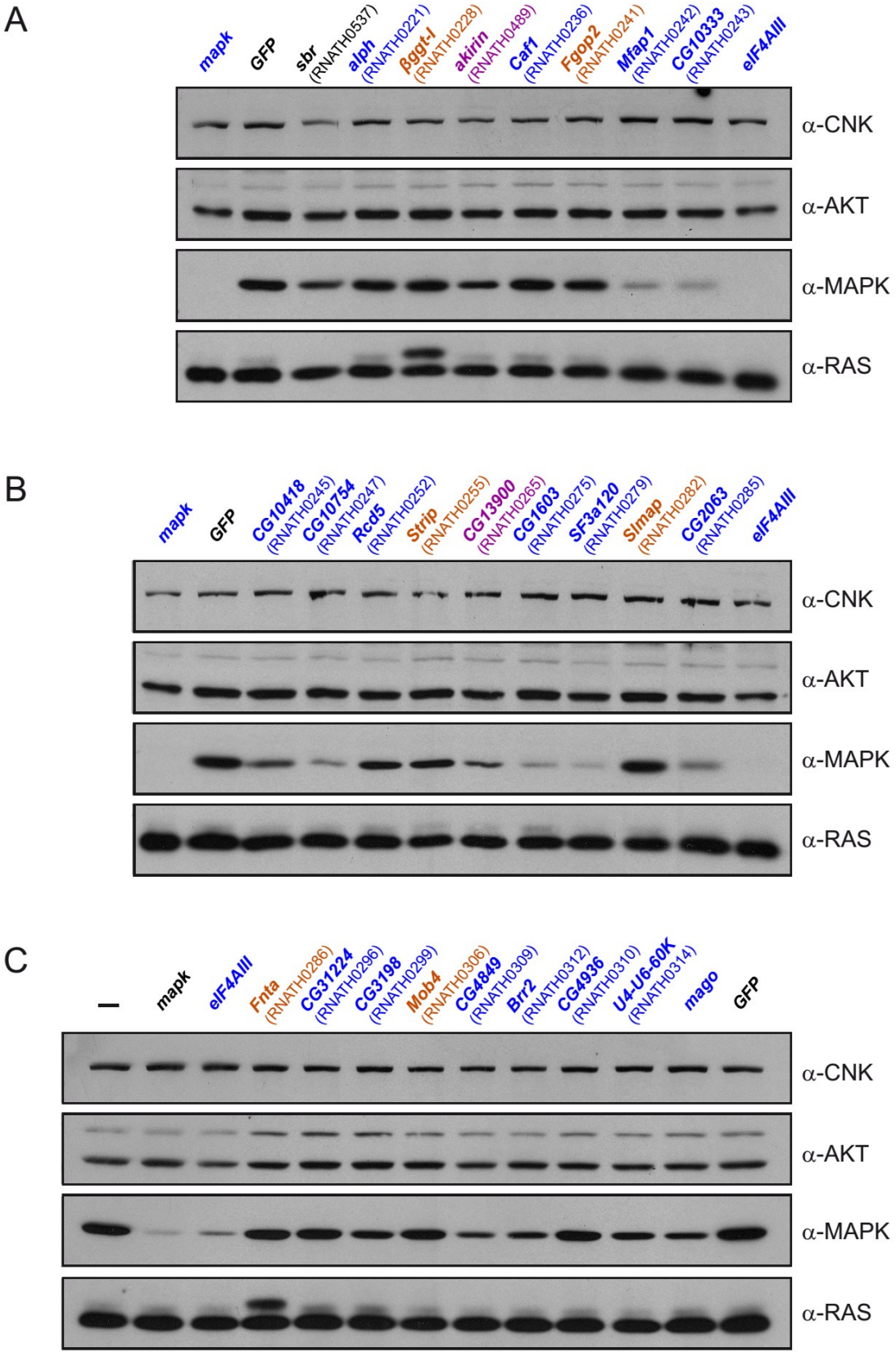


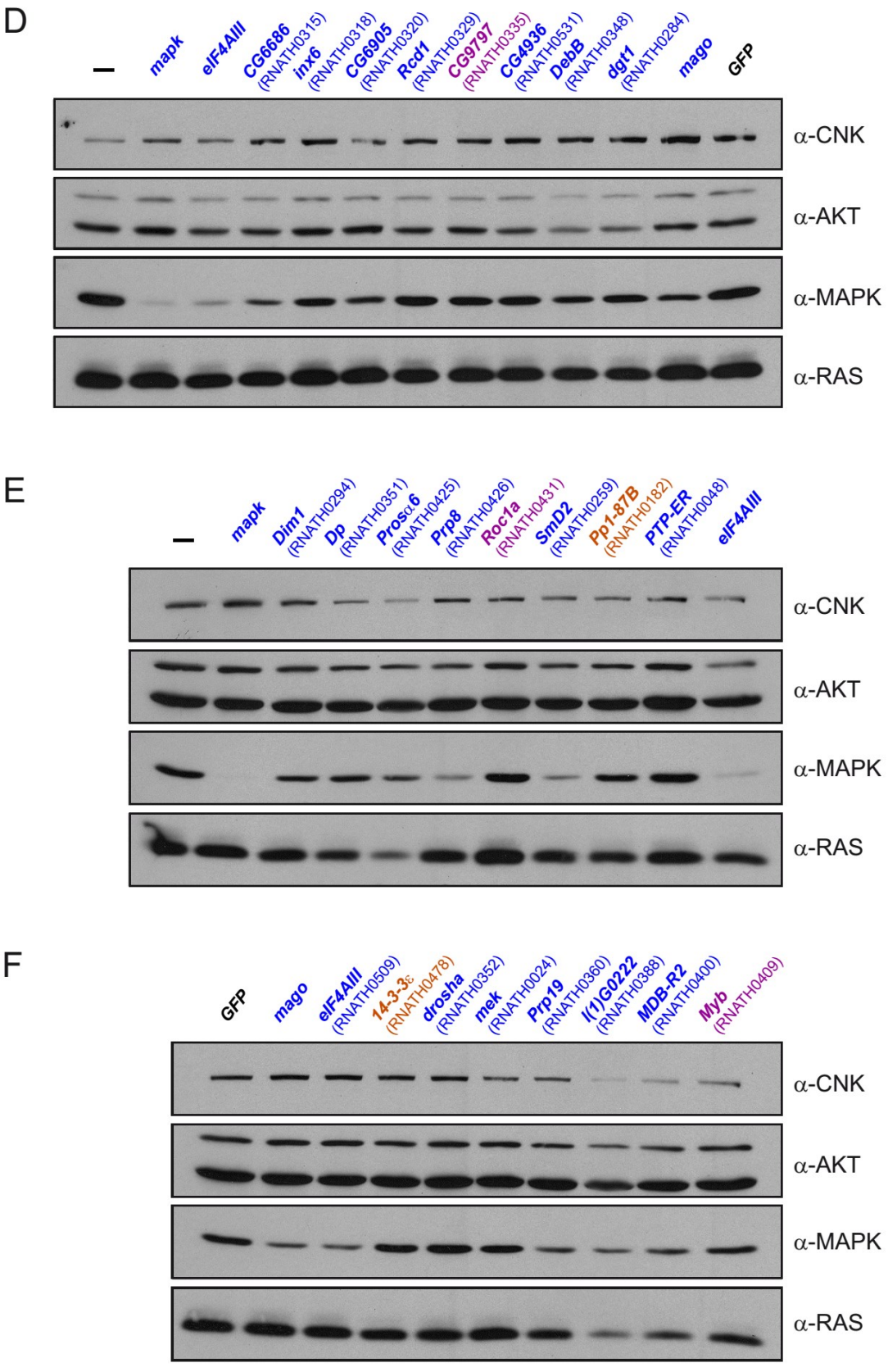
Figure 3.5. Splicing factors cause a decrease in MAPK protein levels

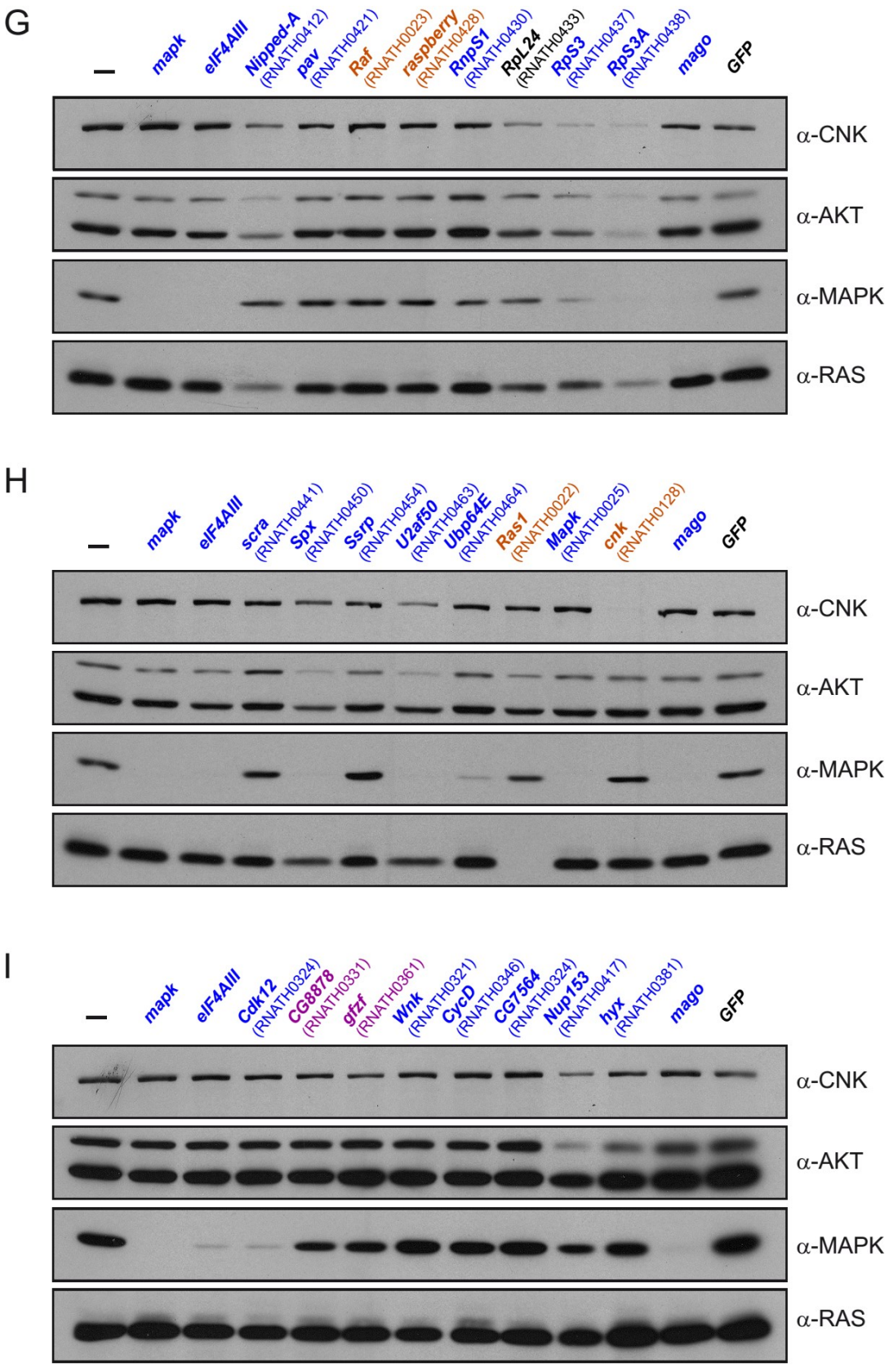
(A) Candidates were tested in an immunofluorescence-based secondary screen to evaluate their impact on MAPK protein levels. The results from this experiment show that RNA processing factors (red) cause a reduction in MAPK levels without impacting AKT (used as a negative control). This impact is similar to that which was observed for EJC components (blue). *mapk* and *Akt* dsRNA positive controls are also shown (green). (B) The specific effect on MAPK levels is confirmed by Western blot for both *Prp19* and *Caper* dsRNA treated samples. The impact on MAPK is similar, albeit slightly weaker, to that observed following *eIF4AIII* depletion. The levels of other MAPK pathway components (RAS, RAF and MEK) as well as other signaling pathway components (AKT and JNK) did not appear to vary significantly following depletion of these factors.

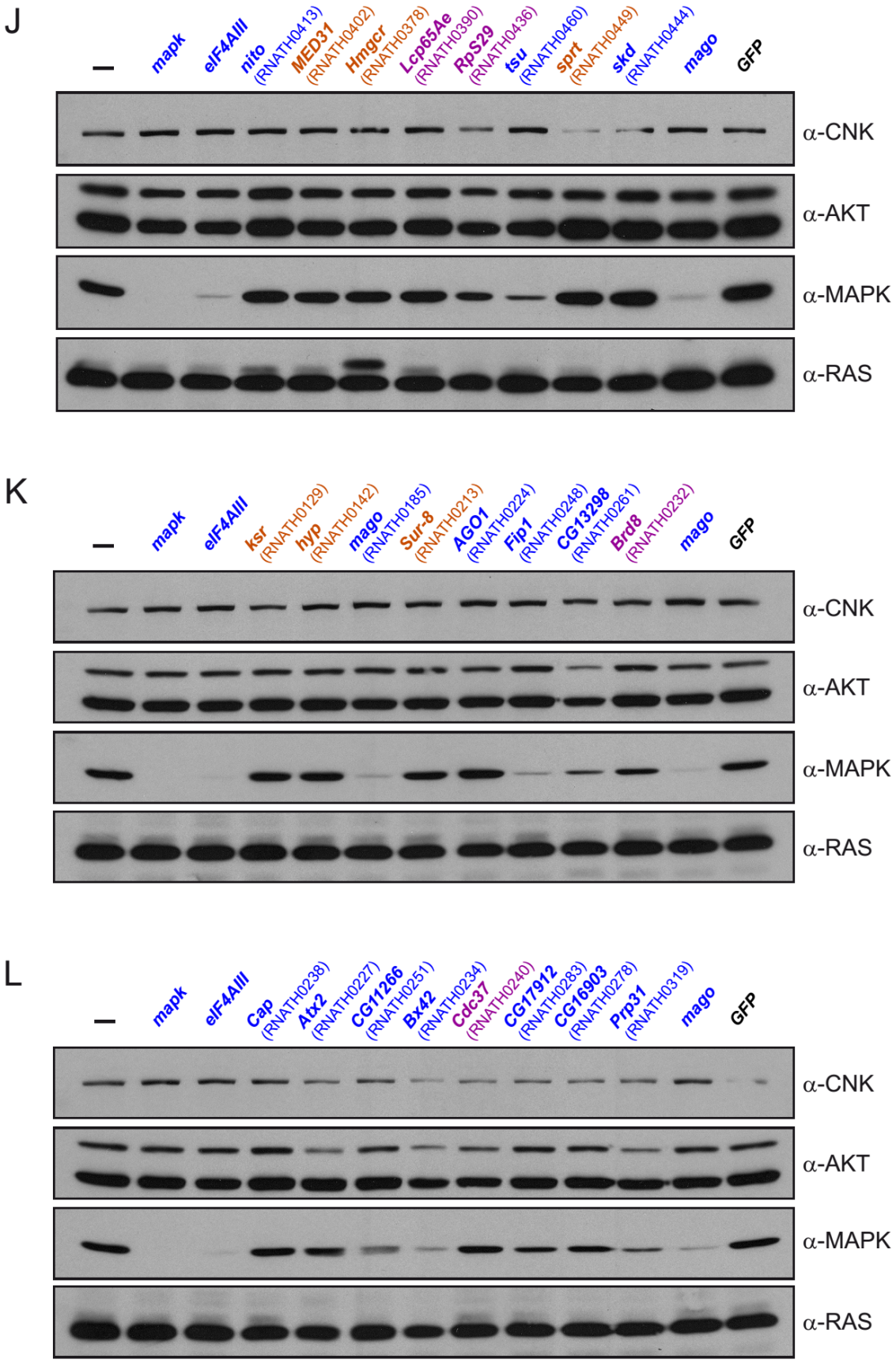
Figure S3.6 Immunoblot Screen

(A to M). Western blot analyses of S2 cells treated with the indicated dsRNAs. Endogenous levels of CNK, AKT, MAPK and RAS were monitored using specific antibodies. A negative (*GFP* dsRNA) and two positive (*mago* and *eiF4AIII dsRNAs*) controls were included in this experiment.

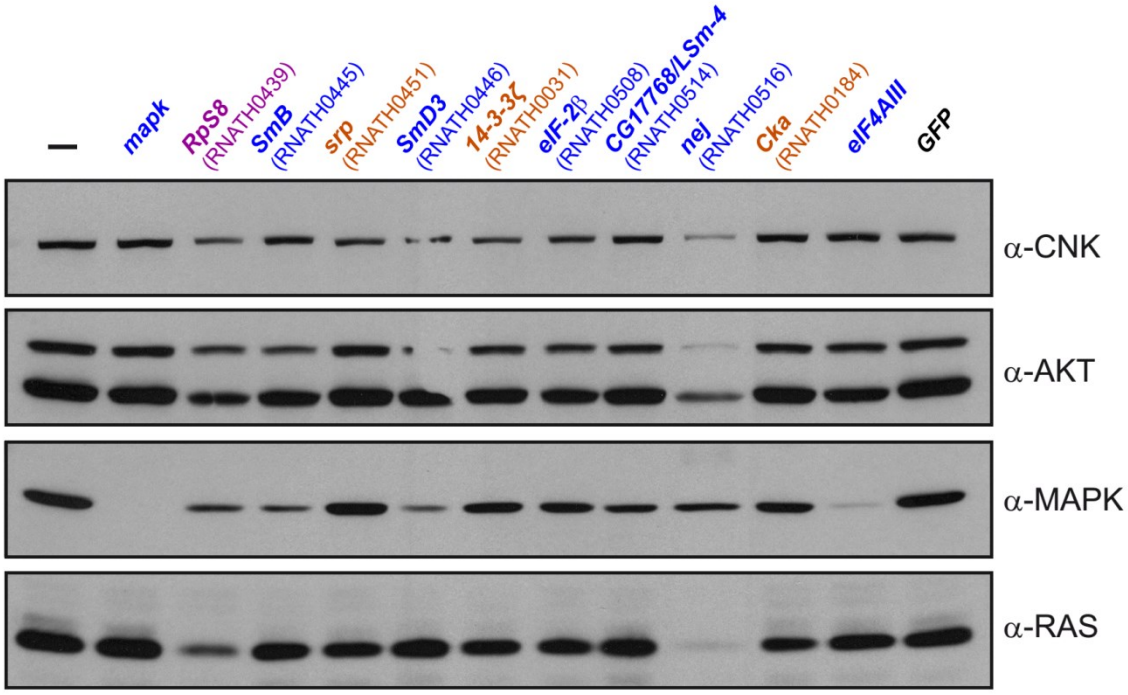

A
mapk
GFP
sbr (RNATH0537)
alph (RNATH0221)
βggt-I (RNATH0228)
akirin (RNATH0489)
Caf1 (RNATH0236)
Fgop2 (RNATH0241)
Mfap1 (RNATH0242)
CG10333 (RNATH0243)
eIF4AIII
α-CNK
α-AKT
α-MAPK
α-RAS
B
mapk
GFP
CG10418 (RNATH0245)
CG10754 (RNATH0247)
Rcd5 (RNATH0252)
Strip (RNATH0255)
CG13900 (RNATH0265)
CG1603 (RNATH0275)
SF3a120 (RNATH0279)
Slmap (RNATH0282)
CG2063 (RNATH0285)
eIF4AIII
α-CNK
α-AKT
α-MAPK
α-RAS
C
–
mapk
eIF4AIII
Fnta (RNATH0286)
CG31224 (RNATH0296)
CG3198 (RNATH0299)
Mob4 (RNATH0306)
CG4849 (RNATH0309)
Brr2 (RNATH0312)
CG4936 (RNATH0310)
U4-U6-60K (RNATH0314)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS


D
–
mapk
eIF4AIII
CG6686 (RNATH0315)
inx6 (RNATH0318)
CG6905 (RNATH0320)
Rcd1 (RNATH0329)
CG9797 (RNATH0335)
CG4936 (RNATH0531)
DebB (RNATH0348)
dgt1 (RNATH0284)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS
E
–
mapk
Dim1 (RNATH0294)
Dp (RNATH0351)
Prosα6 (RNATH0425)
Prp8 (RNATH0426)
Roc1a (RNATH0431)
SmD2 (RNATH0259)
Pp1-87B (RNATH0182)
PTP-ER (RNATH0048)
eIF4AIII
α-CNK
α-AKT
α-MAPK
α-RAS
F
GFP
mago
eIF4AIII (RNATH0509)
14-3-3ε (RNATH0478)
drosha (RNATH0352)
mek (RNATH0024)
Prp19 (RNATH0360)
l(1)G0222 (RNATH0388)
MDB-R2 (RNATH0400)
Myb (RNATH0409)
α-CNK
α-AKT
α-MAPK
α-RAS


G
–
mapk
eIF4AIII
Nipped-A (RNATH0412)
pav (RNATH0421)
Raf (RNATH0023)
raspberry (RNATH0428)
RnpS1 (RNATH0430)
RpL24 (RNATH0433)
RpS3 (RNATH0437)
RpS3A (RNATH0438)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS
H
–
mapk
eIF4AIII
scra (RNATH0441)
Spx (RNATH0450)
Ssrp (RNATH0454)
U2af50 (RNATH0463)
Ubp64E (RNATH0464)
Ras1 (RNATH0022)
Mapk (RNATH0025)
cnk (RNATH0128)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS
I
–
mapk
eIF4AIII
Cdk12 (RNATH0324)
CG8878 (RNATH0331)
gfzf (RNATH0361)
Wnk (RNATH0321)
CycD (RNATH0346)
CG7564 (RNATH0324)
Nup153 (RNATH0417)
hyx (RNATH0381)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS


J
–
mapk
eIF4AIII
nito (RNATH0413)
MED31 (RNATH0402)
Hmgcr (RNATH0378)
Lcp65Ae (RNATH0390)
RpS29 (RNATH0436)
tsu (RNATH0460)
sprt (RNATH0449)
skd (RNATH0444)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS
K
–
mapk
eIF4AIII
ksr (RNATH0129)
hyp (RNATH0142)
mago (RNATH0185)
Sur-8 (RNATH0213)
AGO1 (RNATH0224)
Fip1 (RNATH0248)
CG13298 (RNATH0261)
Brd8 (RNATH0232)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS
L
–
mapk
eIF4AIII
Cap (RNATH0238)
Atx2 (RNATH0227)
CG11266 (RNATH0251)
Bx42 (RNATH0234)
Cdc37 (RNATH0240)
CG17912 (RNATH0283)
CG16903 (RNATH0278)
Prp31 (RNATH0319)
mago
GFP
α-CNK
α-AKT
α-MAPK
α-RAS

M


–
mapk
RpS8 (RNATH0439)
SmB (RNATH0445)
srp (RNATH0451)
SmD3 (RNATH0446)
14-3-3ζ (RNATH0031)
eIF-2β (RNATH0508)
CG17768/LSm-4 (RNATH0514)
nej (RNATH0516)
Cka (RNATH0184)
eIF4AIII
GFP
α-CNK
α-AKT
α-MAPK
α-RAS

3.4.5 RAS-RAF candidates: the usual suspects and STRIPAK

Most of the positive regulators positioned in the RAS-RAF interval were factors that had previously been linked either to RAS prenylation or regulation of RAF activation. As previously mentioned, one of the RAS prenylation factors we identified was *βggt-I*, which was originally identified in *Drosophila*[10,45]. In addition to *βggt-I*, two other factors that are known to function in RAS prenylation in other organisms were also identified: *Hmgcr* and *Fnta* (CG2976). FNTA is the alpha farnesyltransferase subunit for mammalian RAS proteins[48]. The hydroxymethylglutaryl-CoA reductase HMGCR functions in the cholesterol biosynthesis pathway and is required for farnesylation of RAS and other membrane-associated proteins[49,50]. In our western blot experiments, all three of these factors were observed to cause a mobility shift in RAS suggesting that RAS geranylgeranylation is impaired[51], and thus that these factors act on RAS in *Drosophila* (Figure S3.6A,C,J).

RAF activation is arguably the most tightly regulated step of the MAPK module[1,27,52]. Multiple components are involved in a series of events that link up RAF to RAS, anchor RAF at the plasma membrane, allow RAF to adopt and maintain an active conformation and finally enable efficient substrate targeting[52]. Phosphorylation and dephosphorylation events control progression throughout these steps[52]. Of these, the removal of the phosphate moiety on the S346 residue of *Drosophila* RAF (equivalent to S259 of human RAF1) is one of the pivotal regulatory events as it is thought to trigger the release of 14-3-3, which otherwise sequesters RAF in the cytoplasm[52]. Of the factors involved in RAF activation, all of the expected factors (*ksr*, *cnk*, *hyp*, *14-3-3ε* and *14-3-3ζ*) were correctly positioned at the RAS-RAF interval (Table S3.II). Two other candidates, *Pp1-87B* and *Sur-8*, also clustered together with the set of known factors acting in the RAS-RAF interval. These two factors had not previously been shown to act on MAPK signaling in *Drosophila*, but evidence from other organisms indicates that they might act at this level[16,17,40] which is consistent with our results. Of particular interest, one study has found a mammalian complex composed of PP1, SUR-8 and MRAS and linked it to dephosphorylation of the S259 residue on C-RAF[53].

The only other positive regulators in the RAS-RAF epistasis group that were not previously linked to RAS/MAPK signaling were five components that form the smallest of

two complexes in our network (CKA, STRIP, SLMAP, FGOP2 and MOB4) (Figure 3.3B). Three of these components are distantly related to budding yeast alpha factor arrest (FAR) complex components, which are involved in signaling G1 arrest upon alpha factor stimulation[54]. More recently, a protein complex comprising Striatin, the catalytic subunit of PP2A, the STE20 family kinase STK24, and four additional core proteins was identified in human cells and named the Striatin-interacting phosphatase and kinase (STRIPAK) complex[38]. The core of this complex was suggested to serve as a protein platform that specifies PP2A and/or STK24 action. Remarkably, the five RAS-RAF proteins identified are homologs of the non-catalytic members that make up the core STRIPAK complex. CKA, which is the fly Striatin homolog, has previously been demonstrated genetically to act as a positive regulator of JNK signaling[55]. MOB4 has also been previously studied genetically in *Drosophila*, where it appears to participate in mitotic spindle assembly[56]. The three other members, STRIP, FGOP2 and SLMAP, have not been extensively studied in flies and are named based on their mammalian counterparts.

Consistent with their ability to work together as a complex, the five STRIPAK homologs had similar effects in all the secondary screens and epitope-tagged variants co-immunoprecipitated in binary co-expression experiments (Figure 3.3A and Figure S3.7). Notably, their depletion also suppressed JNK activation induced by $RAC1^{V12}$, suggesting that CKA/Striatin modulates signaling through this pathway and that the other STRIPAK members act in conjunction with Striatin in this context. This is also consistent with the findings that T3JAM (TRAF3-interacting JNK-activating modulator), one of the SLMAP homologs, is linked to JNK signaling[57] and with a recent report that identified *Cka* as a suppressor of JNK signaling[58]. Furthermore, depletion of STRIPAK components reduced pMAPK signal induced by insulin, activated Sevenless RTK (SEV^{S11}) and GAP RNAi, but only marginally affected EGFR signaling (Figure S3.5). This suggests that the role of STRIPAK differs depending on the MAPK and JNK activation contexts.

To confirm the involvement of STRIPAK complex components in RAS/MAPK signaling *in vivo*, we conducted genetic interaction experiments using *Cka*/Striatin mutant alleles[55]. RAS/MAPK activity is required for neuronal photoreceptor and cone cell differentiation during *Drosophila* eye development[59,60]. Expression of Ras^{V12} under the

control of the eye specific *sev* promoter/enhancer regulatory sequences produces extra photoreceptor cells, which causes a characteristic rough-eye phenotype ([61] and Figure 3.6B). This rough eye phenotype was dominantly suppressed in a *Cka* heterozygous mutant background (Figure 3.6E and Figure S3.8A,). Extra wing vein material produced by a constitutively active *Egfr* allele, $Egfr^{Elp}$, was also dominantly suppressed by *Cka* mutant alleles to a degree comparable to a weak loss-of-function allele of *rl*/*mapk* (Figure S3.8E,J). In agreement with these results, wing vein deletions were significantly enhanced in a *shp-2*/*csw* hemyzygous mutant background when a *Cka* mutation was introduced in this context (Figure S3.8D,I,K). Moreover, consistent with the role of STRIPAK in RAS/MAPK signaling, loss-of-function of *Cka* activity impaired R7 photoreceptor cell differentiation, which is a classical RAS/MAPK-dependent developmental event (Figure S3.9).

In addition to PP2A/*mts*, we noted that some of the other STRIPAK components described in Goudreault et al.[38], were not identified in our primary screen. However, one of these, *GckIII*, was one of the validated regulators in the InR-driven MAPK screen reported by Adam Friedman and Norbert Perrimon[62]. This raised the possibility that *GckIII* might have an impact on pathway activity in alternate activation contexts. To address this, we examined the effects of depleting *GckIII* in RAS^{V12}, insulin and *GAP* RNAi assays. We found that while this had little impact on RAS^{V12}-induced MAPK activation, *GckIII* depletion had an effect comparable to *Fgop2* depletion in the insulin and *GAP* RNAi contexts (data not shown). This indicates that *GckIII* may function upstream or in parallel to RAS and raises the intriguing possibility that the STRIPAK complex regulates multiple aspects of the larger RTK/RAS/MAPK pathway.

Since Striatins are defined as PP2A regulatory (B) subunits[63] and STRIPAK was initially described as a PP2A associated complex, it is possible that STRIPAK assumes this role in the context of RAF activation. We observed a modulation of RAS^{V12} signaling upon depletion of the catalytic subunit of PP2A (*mts*), but not of the regulatory B subunit, *tws*, which support the notion that STRIPAK components are functioning as PP2A regulatory (B) subunits in this context (data not shown). A similar function for STRIPAK has recently been described in the context of Hippo signaling, where it was found to associate with PP2A and HPO[64]. Finally, in agreement with our findings, another recent report has linked the CKA

subunit to RAS/MAPK signaling[65] and another study conducted in *Neurospora* has indicated that MAPK regulates STRIPAK function, suggesting the possibility of regulatory feedback phosphorylation[66].

Figure S3.7 Co-immunoprecipitation of *Drosophila* STRIPAK complex components

S2 cells were transfected with expression plasmids carrying cDNAs encoding the fusion proteins indicated at the top of A, B, and C. Cell lysates were immunoprecipitated with the indicated antibodies (bottom of upper panels in A and B, and left of upper panels in C. The immunoprecipitates and equal amounts of cell lysates (normalized for total protein content) were fractionated by SDS-PAGE and immunoblotted with the antibodies indicated at the right (A – C). The tagged STRIPAK complex protein shown in each panel is indicated in parentheses. None of the STag fusion proteins nor GFP-FGOP2 are present in α-HA or α-HSV immunoprecipitates when these fusion proteins are expressed alone, and GFP does not co-immunoprecipitate with HA-CKA, HSV-MOB4, HSV-STRIP, or HSV-SLMAP (data not shown).

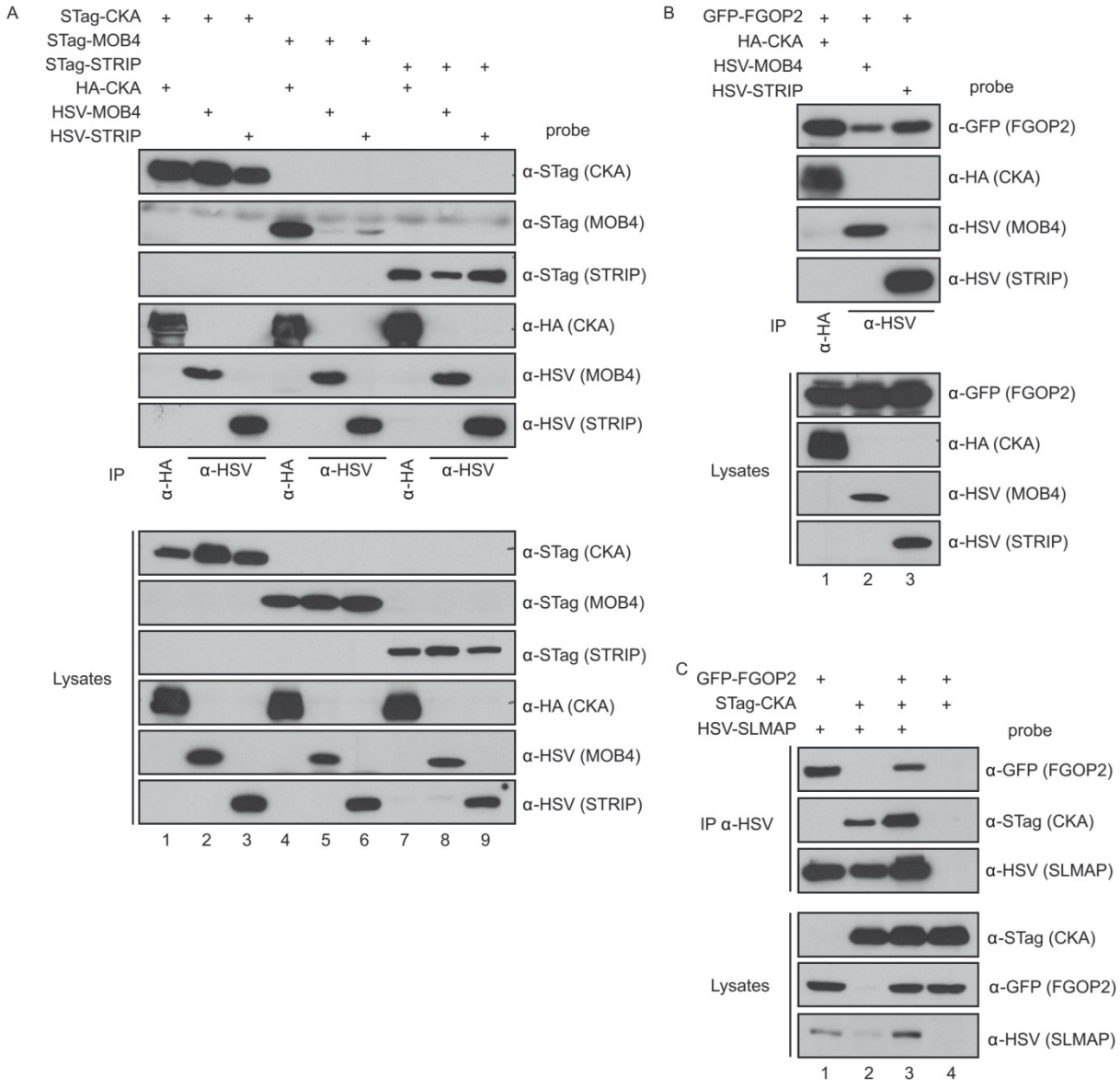

A
STag-CKA + + +
STag-MOB4 + + +
STag-STRIP + + +
HA-CKA + + +
HSV-MOB4 + + +
HSV-STRIP + + +
probe
α-STag (CKA)
α-STag (MOB4)
α-STag (STRIP)
α-HA (CKA)
α-HSV (MOB4)
α-HSV (STRIP)
IP
α-HA
α-HSV
α-HA
α-HSV
α-HA
α-HSV
Lysates
α-STag (CKA)
α-STag (MOB4)
α-STag (STRIP)
α-HA (CKA)
α-HSV (MOB4)
α-HSV (STRIP)
1 2 3 4 5 6 7 8 9
B
GFP-FGOP2 + + +
HA-CKA +
HSV-MOB4 +
HSV-STRIP +
probe
α-GFP (FGOP2)
α-HA (CKA)
α-HSV (MOB4)
α-HSV (STRIP)
IP
α-HA
α-HSV
Lysates
α-GFP (FGOP2)
α-HA (CKA)
α-HSV (MOB4)
α-HSV (STRIP)
1 2 3
C
GFP-FGOP2 + + +
STag-CKA + + +
HSV-SLMAP + + +
probe
IP α-HSV
α-GFP (FGOP2)
α-STag (CKA)
α-HSV (SLMAP)
Lysates
α-STag (CKA)
α-GFP (FGOP2)
α-HSV (SLMAP)
1 2 3 4

Figure 3.6 RNAi screen candidates interact genetically with RAS/MAPK pathway components

(A-J) The Ras^{V12} rough eye phenotype is dominantly suppressed by heterozygous mutations in *Cka*, *gfzf*, *CG1603*, *Fip1*, *Prp19*, *Caper* and a trans-heterozygous mutation in *CG4936*. Fly eyes of the indicated genotypes were imaged by stereomicroscopy. The *mapk* alleles $mapk^{E1171}$ and rl^{1} are used as positive controls. All fly eye images are from female flies except $CG4936^{DG10305}/CG4936^{EY10172}$ which is from a male fly; the rough eye phenotype was observed to be similar in males and females except in this case where males displayed a stronger genetic interaction. (K-N) Genetic interactions with rl^{1} wing vein deletion phenotypes. rl^{1}/ rl^{1} flies display a slight deletion of the mid-section of the L4 wing vein that is not fully penetrant. The L4 deletion is enhanced, sometimes extending to the posterior cross vein (pcv) in $Prp19^{CE162}$ and $Caper^{f07714}$ heterozygous backgrounds (pictures shown served to illustrate detailed scoring results in Figure S3.8H). (O-R) Genetic interactions with rl^{1} rough-eye phenotypes. The weak rough eye phenotype observed in rl^{1} homozygotes is shown. The severity of this phenotype is increased in heterozygous mutant backgrounds for *Prp19* and *Caper*; these flies display a further decrease in eye size and an increased eye roughness.

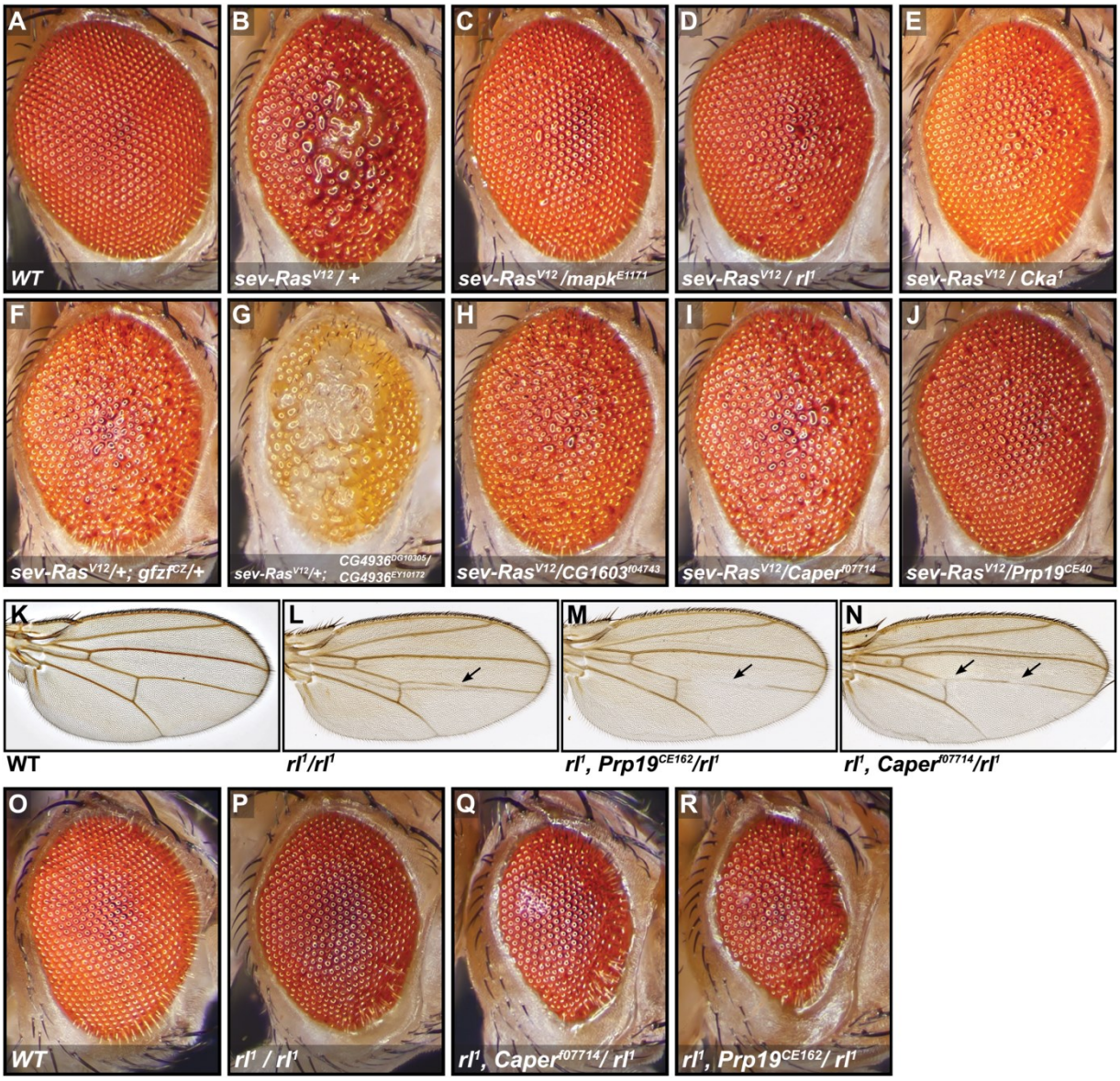
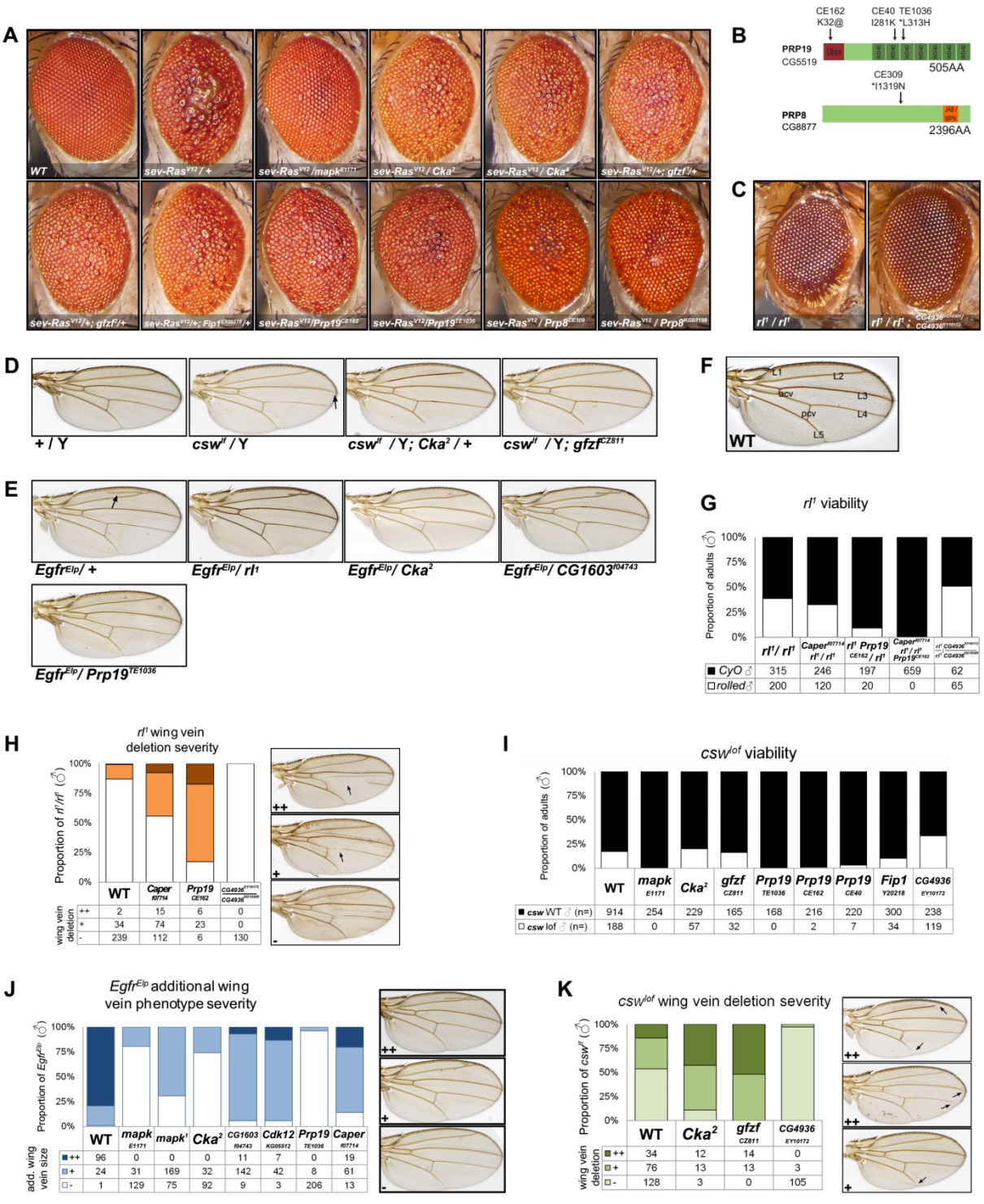

A
WT
B
sev-Ras^{V12} / +
C
sev-Ras^{V12} /mapk^{E1171}
D
sev-Ras^{V12} / rl^{1}
E
sev-Ras^{V12} / Cka^{1}
F
sev-Ras^{V12}/+; gfzf^{CZ}/+
G
sev-Ras^{V12}/+; CG4936^{DG10305}/ CG4936^{EY10172}
H
sev-Ras^{V12}/CG1603^{f04743}
I
sev-Ras^{V12}/Caper^{f07714}
J
sev-Ras^{V12}/Prp19^{CE40}
K
WT
L
rl^{1}/rl^{1}
M
rl^{1}, Prp19^{CE162}/rl^{1}
N
rl^{1}, Caper^{f07714}/rl^{1}
O
WT
P
rl^{1} / rl^{1}
Q
rl^{1}, Caper^{f07714}/ rl^{1}
R
rl^{1}, Prp19^{CE162}/ rl^{1}

Figure S3.8. Additional genetic interaction data

(A) Ras^{V12} genetic interaction data for additional alleles of *Cka*, *gfzf* and *Prp19* as well as alleles of *Fip1* and *Prp8*. (B) Details of molecular lesions present in the three alleles of *Prp19* and the $Prp8^{CE309}$ allele identified in a *cnk* dominant negative genetic screen (Baril, Lefrançois et al., manuscript in preparation) and used in our genetic validation experiments. A protein map of PRP19 and PRP8 showing the amino acid changes found in the *Prp19* and *Prp8* alleles is presented with "*" indicating residues conserved in humans. The mutation in $Prp8^{CE309}$, in addition to causing an amino acid change, is located on the second residue of a 3'SS (T4088A) and may impact splicing of *Prp8*. (C) Additional rl^1 rough-eye phenotype genetic interactions for *CG4936*. Trans-heterozygous $CG4936^{DG10305}/CG4936^{EY10172}$ suppresses the weak rough eye phenotype of rl^1 homozygotes. Eye size is also slightly restored. (H) Representative wings for csw^{lf} hemizygous males scored in (K). The deletion of the distal end of the L2, L3 and/or L5 vein is a frequently observed phenotype in a csw^{lf} background. The occurrence, number and severity of this deletion are more pronounced in Cka^2 and $gfzf^{CZ811}$ heterozygous backgrounds. (E) Representative wing images for the $Egfr^{Elp}$ males scored in (J). The additional wing vein near the extremity of the L2 vein (arrow) is characteristic in $Egfr^{Elp}$ flies. Wings displaying a suppressed phenotype (reduced frequency and length of the extra vein material) are shown for Cka^2, $CG1603^{f04743}$, $Prp19^{TE1036}$ and $mapk/rl^1$ (positive control) heterozygotes. (F) Image of a wild-type fly wing with labels indicating the location of the 5 wing veins (L1 – L5) and of the anterior (acv) and posterior (pcv) cross-veins. (G) Proportion of $rl^1/\ rl^1$ males to $rl^1/\ CyO$ males observed following a $rl^1/\ CyO$ X $rl^1/\ CyO$ cross. The wings of $rl^1/\ rl^1$ flies have a "rolled" phenotype; they are slightly curved downwards along the anterior-posterior axis. $rl^1/\ CyO$ flies display a regular *CyO* phenotype; pronounced upward curve along a lateral axis. Both $rl^1\ Prp19^{CE162}/\ CyO$ and $Caper^{f07714}\ rl^1/\ CyO$ were crossed to $rl^1/\ CyO$ flies. The frequency of $rl^1\ Prp19^{CE162}/\ rl^1$ and $Caper^{f07714}\ rl^1/\ rl^1$ was lower than that of $rl^1/\ rl^1$. The $rl^1\ Prp19^{CE162}/\ Caper^{f07714}\ rl^1$ combination was not viable at 25°C. Trans-heterozygous *CG4936* slightly enhanced the proportion of $rl^1/\ rl^1$ flies recovered. (H) $rl^1/\ rl^1$ flies display a deletion of the mid-section of the L4 wing vein that is not fully penetrant. Examples are Shown for wings with weak "+" (single small deletion) and strong "++" (large or multiple deletions) wing vein deletions. Wing deletions were scored for the indicated

genotypes. The total amount of individuals scored is indicated below the allele labels. (I) Proportion of csw^{lf} males to csw^{+} males observed for the indicated genotypes with the total amount of individuals scored indicated below the allele labels. The trans-heterozygous allelic combination of *CG4936* behaved similarly to heterozygous $CG4936^{EY10172}$ in csw^{lf} genetic interaction experiments (wing vein deletions and viability) and is not shown here. (J) Proportions of wings from $Egfr^{Elp}$ flies with the indicated extra wing vein material phenotype severity for the indicated alleles. Examples of strong and weak additional wing vein phenotypes are shown next to the graph. The total number of wings scored is indicated below the allele labels. The *Fip1* and *gfzf* alleles displayed only slight suppression of the wing vein phenotype and are not shown here. (K) Proportions of wings from csw^{lf} males with the indicated deletion severity for the indicated alleles. Examples of strong (large or multiple small) and weak (single, small) wing vein deletions are shown. The total number of wings scored is indicated below the alleles. The *Caper* and *CG1603* alleles only displayed very weak genetic interaction with csw^{lf} and are not shown here. The *Prp19* flies produced too few escapers for accurate scoring of the wing phenotypes.


A
WT
sev-Ras^{V12} / +
sev-Ras^{V12} /mapk^{E1171}
sev-Ras^{V12} / Cka^{2}
sev-Ras^{V12} / Cka^{4}
sev-Ras^{V12}/+; gfzf^{1}/+
sev-Ras^{V12}/+; gfzf^{2}/+
sev-Ras^{V12}/+; Fip1^{EY20218}/+
sev-Ras^{V12}/Prp19^{CE162}
sev-Ras^{V12}/Prp19^{TE1036}
sev-Ras^{V12} / Prp8^{CE309}
sev-Ras^{V12} / Prp8^{KG03188}
B
CE162 K32@
CE40 TE1036 I281K *L313H
PRP19 CG5519
505AA
CE309 *I1319N
PRP8 CG8877
2396AA
C
rl^{1} / rl^{1}
rl^{1} / rl^{1} ; CG4936^{EY10172}/ CG4936^{EY10172}
D
+ / Y
csw^{ll} / Y
csw^{ll} / Y; Cka^{2} / +
csw^{ll} / Y; gfzf^{CZ811}
E
Egfr^{Elp}/ +
Egfr^{Elp}/ rl^{1}
Egfr^{Elp}/ Cka^{2}
Egfr^{Elp}/ CG1603^{f04743}
Egfr^{Elp}/ Prp19^{TE1036}
F
L1
L2
L3
L4
L5
acv
pcv
WT
G
rl^{1} viability
Proportion of adults ♂
100%
75%
50%
25%
0%
rl^{1}/ rl^{1} | Caper^{f07714} rl^{1} / rl^{1} | rl^{1} Prp19^{CE162} / rl^{1} | Caper^{f07714} rl^{1} / rl^{1} Prp19^{CE162} | rl^{1} CG4936^{EY10172}/ rl^{1} CG4936^{EY10172}
CyO ♂ | 315 | 246 | 197 | 659 | 62
rolled ♂ | 200 | 120 | 20 | 0 | 65
H
rl^{1} wing vein deletion severity
Proportion of rl^{1}/rl^{1} ♂
100%
75%
50%
25%
0%
WT | Caper^{f07714} | Prp19^{CE162} | CG4936^{EY10172} CG4936^{EY10359}
wing vein deletion
++ | 2 | 15 | 6 | 0
+ | 34 | 74 | 23 | 0
- | 239 | 112 | 6 | 130
++
+
-
I
csw^{lof} viability
Proportion of adults ♂
100%
75%
50%
25%
0%
WT | mapk^{E1171} | Cka^{2} | gfzf^{CZ811} | Prp19^{TE1036} | Prp19^{CE162} | Prp19^{CE40} | Fip1^{Y20218} | CG4936^{EY10172}
csw WT ♂ (n=) | 914 | 254 | 229 | 165 | 168 | 216 | 220 | 300 | 238
csw lof ♂ (n=) | 188 | 0 | 57 | 32 | 0 | 2 | 7 | 34 | 119
J
Egfr^{Elp} additional wing vein phenotype severity
Proportion of Egfr^{Elp} ♂
100%
75%
50%
25%
0%
WT | mapk^{E1171} | mapk^{1} | Cka^{2} | CG1603^{f04743} | Cdk12^{KG05512} | Prp19^{TE1036} | Caper^{f07714}
add. wing vein size
++ | 96 | 0 | 0 | 0 | 11 | 7 | 0 | 19
+ | 24 | 31 | 169 | 32 | 142 | 42 | 8 | 61
- | 1 | 129 | 75 | 92 | 9 | 3 | 206 | 13
++
+
-
K
csw^{lof} wing vein deletion severity
Proportion of csw^{lof} ♂
100%
75%
50%
25%
0%
WT | Cka^{2} | gfzf^{CZ811} | CG4936^{EY10172}
wing vein deletion
++ | 34 | 12 | 14 | 0
+ | 76 | 13 | 13 | 3
- | 128 | 3 | 0 | 105
++
++
+

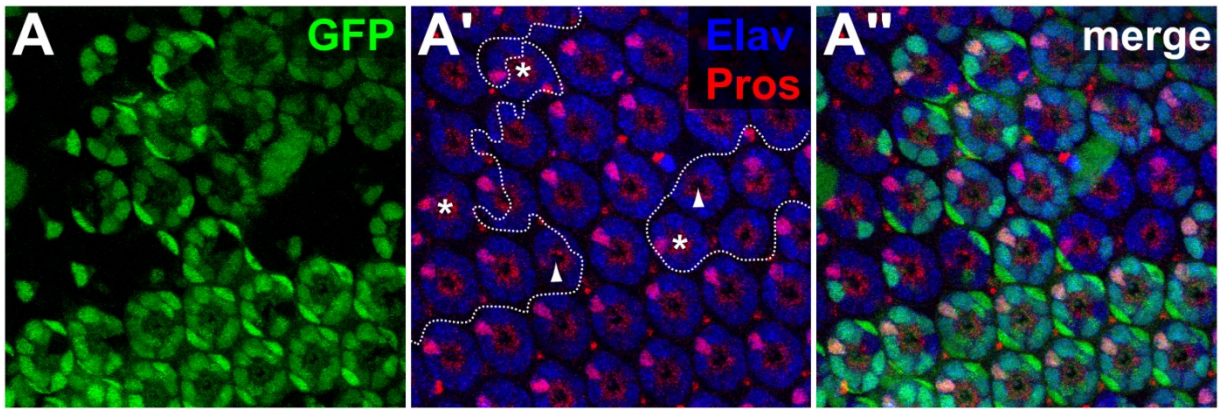


Figure S3.9. *Cka* is important for photoreceptor development

(A) *Cka*^{2} mutant clones in pupal eye discs dissected 42h APF are marked by the absence of GFP (A). The differentiated R7 photoreceptor is marked by the overlapping Elav and Pros stainings (in purple, A' and A''). In *Cka*^{2} clones, ommatidia lacking the R7 photoreceptor are marked with a white arrow. We also noted some displaced R7 photoreceptors (R7s outside the focal plane are marked with white arrowheads), and rotated ommatidia (white asterisks) in the *Cka*^{2} clones.

3.4.6 RAF-MEK candidates: gfzf

Only 13 of our candidates were found to act downstream of RAF and upstream of MEK. Of these, the RAF chaperone *Cdc37* was the only *bona fide* pathway regulator. Among the others, *CG8878* was of interest as it is the homolog of the mammalian Vaccinia-Related Kinase (VRK) genes of which two (VRK1 and VRK3) have been recently identified in a recent screen for KRAS synthetic-lethal factors[67]. Also, another study has shown that VRK2 interacts with MEK and KSR1, potentially acting on MEK activation[68]. However, in this context, VRK2 acts as a negative regulator whereas *CG8878* appears to act positively on RAS signaling in our experiments.

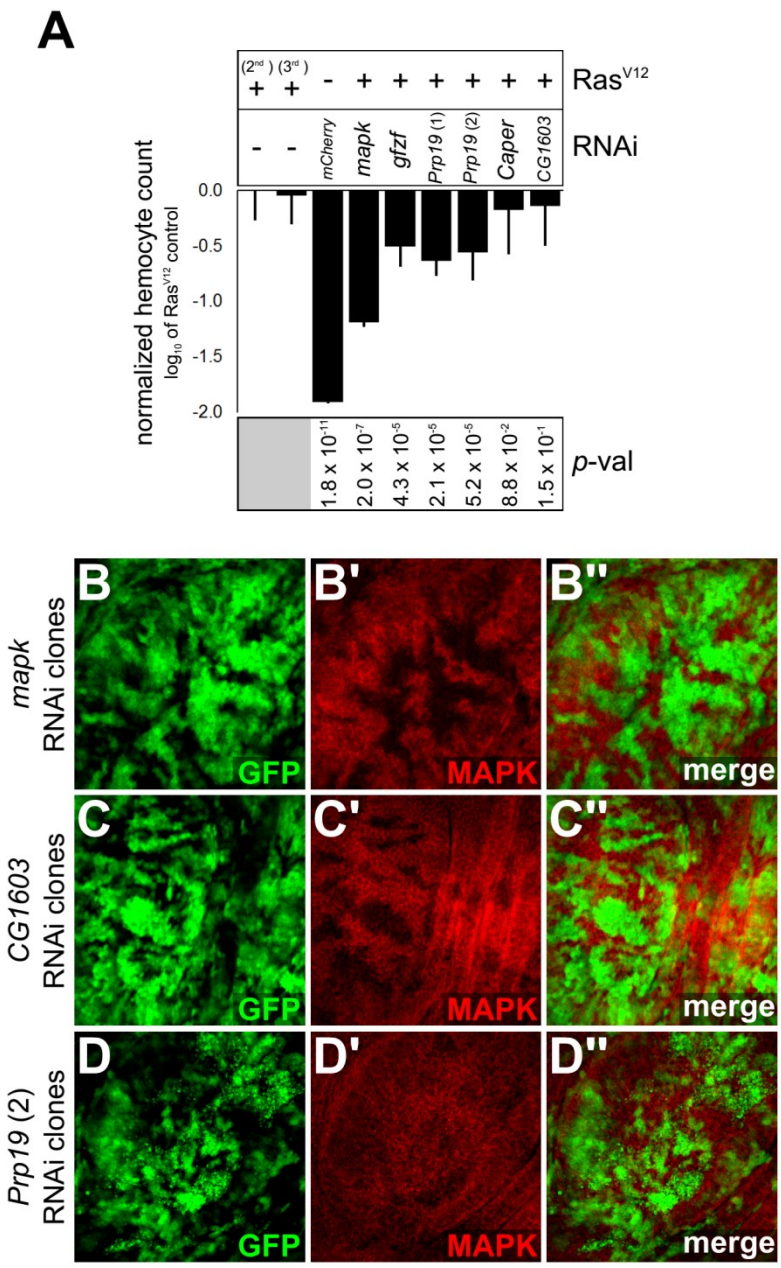
The most interesting candidate that fell within the RAF-MEK interval was GST-containing FLYWCH zinc-finger protein (*gfzf*). GFZF was initially found because of the property of its GST domain to bind to glutathione sepharose beads[69]. It does not, however, have any clearly identified function assigned to it, but there are indications that it may act as a co-factor for the E2F transcription factor[70]. Also, recent RNAi screens have suggested that *gfzf* may be acting downstream of PDGF to control cell size[71] and as a factor functioning in the G2/M DNA damage checkpoint[72]. In our qPCR experiments, *gfzf* clearly stood out from other candidates as it was the only factor that caused a strong reduction in *mek* transcript levels and had the closest profile to the *mek* RNAi itself (Figure 3.4A,B and Figure S3.4A). This observation also extended to the protein product as we observed a clear reduction in MEK levels upon *gfzf* knockdown (Figure 3.4C). In flies, two different *gfzf* loss of function alleles suppressed the Ras^{V12} rough eye phenotype (Figure 3.6F and Figure S3.8A). $gfzf^{ez811}$also increaded the severity of wing vein deletions in hemizygous csw^{lf} males (Figure S3.8D,K). Finally, knocking down *gfzf* reduced RAS^{V12}-induced hemocyte proliferation (Figure 3.7A). Together, these observations suggest *gfzf* is regulating *mek*, possibly by acting as a positive transcription factor. Since FLYWCH domain proteins have been found to negatively control miRNA expression in *C.elegans*[73], one interesting alternative is that *gfzf* represses the production of a miRNA that targets the *mek* transcript.

Another candidate, *CG9797*, was clustered close to *gfzf* through the functional screen data and the qPCR results (Figure 3.3A and Figure S3.4A) and its protein product is predicted to interact with GFZF (Figure 3.3B). In the qPCR screen, *CG9797* knockdown caused a weak

reduction in *mek* transcript levels; it is the 4th strongest hit in terms of reducing *mek* levels, after the dsRNAs targetting *gfzf*, *mek* and *RpL24* (Table S3.V). From this, and since *CG9797* encodes another zinc finger protein, it is possible that this factor works in conjunction with GFZF as a transcriptional regulator. Recently, the chromatin remodeling factors *Geminin* and *Brahma* have been found to modulate MEK protein expression in *Drosophila* wing discs[74] raising the possibility that *gfzf* might be acting in conjunction with these factors.

Figure 3.7. *In vivo* evidence of impact of RAS/MAPK signalling

(A) Impact of candidates on Ras^{V12}-induced hemocyte proliferation. A *Hemolectin-Gal4* driver was used to co-express RAS^{V12} (on either chromosome 2 or 3) with the RNAi constructs or a *UAS-lacZ* control. GFP positive hemocytes were counted by automated microscopy. The total hemocyte count is expressed as $\log_{10}$ ratio of the Ras^{V12} control (2^{nd} chromosome *UAS-*Ras^{V12} fly line). Expression of Ras^{V12} increased hemocyte count by approximately 100 fold compared to a *UAS-mcherry* RNAi negative control without Ras^{V12}. Co-expression of a *mapk* RNAi with Ras^{V12} was used as a control for reduced proliferation. A Student's *t*-test (unpaired, two tailed) was performed comparing candidates to the appropriate Ras^{V12} control. (B-D) 3^{rd} instar larval imaginal discs showing reduction in MAPK protein levels in GFP positive clones expressing RNAi hairpin constructs. Two different *Prp19* RNAi constructs were tested in wing imaginal discs (one is shown here) and found to produce a slight, but consistent reduction in MAPK levels in the clonal tissue. This was more visible in clones with a stronger GFP signal and was sometimes accompanied by signs of apoptosis (small GFP positive fragments) which are visible in (D).


A
(2nd) (3rd)
+ +
- + + + + + +
RasV12
- - mCherry mapk gfzf Prp19 (1) Prp19 (2) Caper CG1603
RNAi
normalized hemocyte count
log10 of RasV12 control
0.0
-0.5
-1.0
-1.5
-2.0
1.8 x 10-11
2.0 x 10-7
4.3 x 10-5
2.1 x 10-5
5.2 x 10-5
8.8 x 10-2
1.5 x 10-1
p-val
mapk RNAi clones
CG1603 RNAi clones
Prp19 (2) RNAi clones
B
B'
B''
C
C'
C''
D
D'
D''
GFP
MAPK
merge

3.4.7 MEK-MAPK candidates: multiple inputs converge on mapk splicing

mRNA processing and transcription factors formed the largest group of hits in our study and also the largest complex in our network analysis (Figure S3.1B and Figure 3.3B) and almost all these factors mapped to the MEK-MAPK epistasis interval (Table S3.V). While the majority of the candidates in this category cause changes in *mapk* expression, one clear exception was *CG4936*. This gene encodes a zinc finger protein of unknown function that is distantly related to human ZBTB20, a BCL-6 like transcription factor that is expressed in hematopoietic tissues and lymphoid neoplasms[75]. *CG4936* also mapped downstream of MEK, but was not found to influence *mapk* expression. A first clue as to the function of this factor was provided by the observation that *CG4936* behaved very similarly to *PTP-ER* in our functional screens; of all the candidates, it has the closest profile to *PTP-ER* (Figure 3.3A). One of the two MEK^{EE}-based MAPK activation assays involved the use of *PTP-ER* RNAi to increase the pMAPK signal. Surprisingly, while both *PTP-ER* and *CG4936* RNAi significantly increase MEK^{EE}-induced pMAPK levels, combining *CG4936* and *PTP-ER* RNAi did not have an additive effect on pMAPK levels (Table S3.II and Table S3.V). This suggests that these factors work together in the same regulatory pathway. *CG4936* also clustered close to PTP-ER in our qPCR screen (Figure S3.4A). Consistent with this, both qPCR and protein analysis revealed that *CG4936* knockdown respectively caused a specific reduction in *PTP-ER* transcript and protein levels (Figure 3.4A,B and D). Furthermore, in our genetic interaction experiments, flies trans-heterozygous for two P-element insertion alleles of *CG4936* ($CG4936^{EY10172}/CG4936^{DG10305}$) showed an enhancement of the Ras^{V12} rough eye phenotype (Figure 3.6G) and suppressed lethality and phenotypes caused by a homozygous $mapk/rl^{1}$ hypomorphic allele (Figure S3.8C,G,H). $CG4936^{EY10172}$ alone suppressed the lethality and wing vein deletions of csw^{lf} hemizygous males (Figure S3.8I,K), as did the double *CG4936* mutant (not shown). Moreover, csw^{lf} homozygous females were observed in a background heterozygous for $CG4936^{EY10172}$, which is indicative of suppression as csw^{lf} is not normally homozygous viable. Altogether, this data suggests that *CG4936* acts on *PTP-ER* transcription and that this action has bearing on RAS/MAPK dependent developmental processes.

Excepting *CG4936*, most of the other candidates positioned downstream of MEK (39) caused a decrease in MAPK protein levels (< -0.25 $\log_2$ fold and $p < 1 \times 10^{-5}$)(Table S3.V). A

few of these factors caused a significant and reproducible reduction in *mapk* transcript levels as well, without significantly impacting the levels of other RAS/MAPK pathway components (Figure S3.4A and Table S3.V). These factors include two EJC components (*eIF4AIII* and *mago*) as well as three other candidates: *Cdk12*, *Fip1* and *CG1603*. Depletion of these latter candidates led to a significant drop in *mapk* transcript levels as measured by qPCR (Figure 3.4A and Table S3.VI). However, these three factors could be distinguished from the EJC components by the fact that they did not cause a shift in the *mapk* RT-PCR profile (Figure S3.5), but only a decrease in the overall transcript abundance, suggesting that they do act at a different step of *mapk* expression. Finally, lethal alleles of *CG1603* and *Fip1* dominantly suppressed a Ras^{V12} induced rough eye phenotype (Figure 3.6H and Figure S3.8A). The *CG1603* allele also suppressed the extra wing vein phenotype caused by $Egfr^{Elp}$, as did $Cdk12^{KG05512}$ (Figure S3.8E,J). However, none of the alleles had a readily observable impact on csw^{lf} phenotypes (not shown); although *Fip1* did cause a slight enhancement of csw^{lf} lethality (Figure S3.8I). Finally, consistent with our cell culture data, knockdown of *CG1603* in larval imaginal disc clones was found to cause a pronounced decrease in MAPK protein levels (Figure 3.7C).

Cdk12 is the main RNA polymerase II C-terminal domain kinase. Phosphorylation of the C-terminal domain of RNAP II is required for transcription elongation, RNA processing and splicing[76,77]. Thus it is possible that *Cdk12* influences either of these steps in the case of *mapk* expression althought it does not seem to induce AS. CDK12 has previously been shown to function in conjunction with Cyclin L in regulating AS[78]. Consistent with this data, the *Drosophila* CycL ortholog, *CG16903*, was also a hit in our screen that mapped to the MEK-MAPK interval (Table S3.III) and caused a shift in the *mapk* RT-PCR profile (Figure S3.5). However, *CG16903* did not significantly change *mapk* transcript levels measured by qPCR (Table S3.V).

Fip1 is also involved in transcript processing; its best studied ortholog in yeast is part of the pre-mRNA cleavage and polyadenylation complex[79]. mRNA polyadenylation plays an important part in stabilizing spliceosome assembly on the 3' most exon of the transcript and is also an important pre-requisite for effective mRNA export and also influences mRNA

stability[80-82]. Thus, like *Cdk12*, *Fip1* may be regulating *mapk* expression by controlling transcript abundance or by influencing splicing efficiency.

CG1603 encodes a protein of unknown function that contains MADF type zinc finger domains which are related to Myb DNA binding domains[83]. *CG1603* is poorly conserved in humans, bearing distant homology to ZNF664 and ZNF322. Interestingly, ZNF322 was found to act as a transcriptional co-activator of SRF and AP-1 in humans, which would position it downstream of MAPK signaling[84], thus representing a potential feedback mechanism.

The remainder of the factors that were observed to lower MAPK levels were either spliceosome components or factors associated to the splicing machinery (Figure 3.5A and Table S3.V). Intriguingly, multiple lines of evidence suggested that these splicing components can also play a specific role in modulating RAS/MAPK signaling: 1) most splicing factors did not have any detectable impact on CNK, AKT and RAS levels, even though these proteins are all derived from intron-containing genes (Figure S3.6), 2) depletion of most splicing factors by RNAi did not significantly modulate PGN and $RAC1^{V12}$-induced JNK activation (Figure 3.3A and Table S3.IV) and 3) these splicing factors scored as hits in previous screens less often than *bona fide* RAS/MAPK pathway components (Table S3.IV). Accordingly, the majority of these splicing factors were categorized in the high specificity group. Importantly, many of these had impacts on MAPK levels that were comparable to or greater in strength than those splicing factors of the lower specificity group. This suggests that the higher specificity score is not simply attributable to lower knockdown efficiency or a weaker impact on constitutive splicing.

Interestingly, one indication that the canonical splicing factors might be acting differently on MAPK signaling than the EJC came from our qPCR expression data. Namely, some high specificity splicing factors such as the *CG10754* (the counterpart of human SF3A2, a U2 snRNP associated factor involved in branch point binding[85]), *CG3198* (the ortholog of LUC7L3, a predicted splicing factor[86]), *Prp19* (the central component of the PRP19 spliceosomal complex involved in C complex assembly[87]), as well as *CG4849* and *CG6686* (two predicted tri-snRNP components[88,89]) all caused a reduction in MAPK protein levels comparable to the EJC factors *mago* and *eIF4AIII*. However, no reduction of *mapk* mRNA was observed by qPCR (Figure 3.4A and Table S3.V). Also, another important difference between the EJC and the canonical splicing factors was that the depletion of many candidates

of the latter group caused an increase in nuclear pre-mRNA retention as measured by fluorescence in-situ hybridization using a poly-A probe (Figure S3.4B,C)[90]. This indicates – as might be expected – that most of these splicing factors also play a more general role in splicing.

The key difference between the EJC and canonical splicing factors suggested by the qPCR and pre-mRNA export data was readily observable using the whole-transcript *mapk* RT-PCR assay; both groups caused an alteration in the *mapk* RT-PCR product, but the canonical splicing factors produced a clearly different pattern (Figure S3.5). Since most of the candidates in the canonical splicing factors group produced similar shifts in the *mapk* RT-PCR profile, we selected two representative factors for more detailed follow-up experiments. The first, *Prp19*, was selected because it is a core spliceosome component that caused a strong reduction in MAPK levels and had a clear impact on the *mapk* RT-PCR profile (Figure 3.5A,B and Figure 3.8B). The second, *Caper*, was selected because it is a serine rich (SR) AS factor[91] that did not perturb global pre-mRNA export, caused a weaker reduction in MAPK protein levels and a less severe change in the *mapk* RT-PCR profile (Figure S3.5, Figure 3.5A,B and Figure 3.8B). Consistent with the protein expression data, while both factors clearly altered the *mapk* RT-PCR products, similar RT-PCR assays targeting *Ras85D*, *raf*, *mek, cnk* or *ksr* did not display any obvious change (data not shown). Also, Western blot experiments showed a clear drop in MAPK protein levels with no effect on RAS, RAF, MEK or AKT (Figure 3.5B). This suggests that splicing of *mapk*, and not of other pathway components, is affected.

Also lending strength to the idea that splicing factors may have a specific role in the RAS/MAPK context was the fact that three *Prp19* alleles and one *Prp8* allele, were isolated by our group in an independent genetic screen for modifiers of a dominant-negative form of CNK (Figure S3.8B and Baril, Lefrançois, et al., manuscript in preparation). In addition to this, *Prp8* has been previously found to enhance the small wing phenotype induced by expression of the *Egfr* inhibitor *aos*[92]. Consistent with this, the *Prp19* and *Prp8* alleles dominantly suppressed the RAS^{V12} rough eye phenotype, as did an allele of *Caper* (Figure 3.6I,J and Figure S3.8A). The *Prp19* alleles also dramatically enhanced csw^{lf} lethality and suppressed the $Egfr^{Elp}$ wing phenotype (Figure S3.8E,I,J). Moreover, in $mapk/rl^{1}$ homozygous flies, *Prp19*

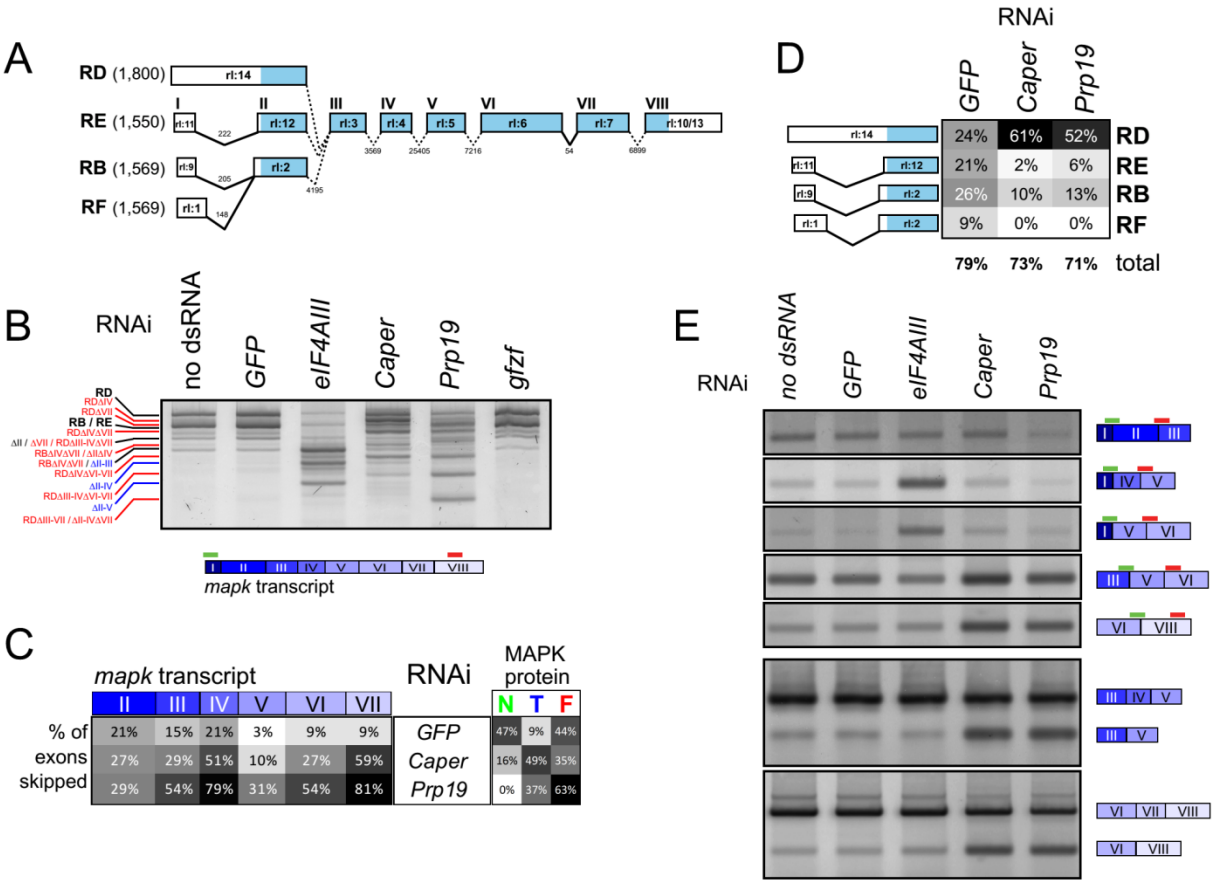
and *Caper* alleles enhanced the severity of wing vein deletions and rough eye phenotypes (Figure 3.6K-R and Figure S3.8G,H). Importantly, although flies carrying one copy of both $Caper^{f07714}$ and $Prp19^{CE162}$ were perfectly viable, this allelic combination was entirely lethal in a $mapk/rl^{1}$ homozygous background (Figure S3.8G). This constitutes another indication that these splicing factors are acting in concert on *mapk* expression. Finally, splicing of *mapk* was found to be altered in $Prp19^{CE162}$ flies also homozygous for rl^{1} (Figure S3.10A,B). The rl^{1} mutant alone reduces *mapk* transcript levels without altering the RT-PCR splicing profile[93; and data not shown].

Complementing the genetic interaction experiments, expression of *Prp19* RNAi reduced the Ras^{V12}-induced proliferation of larval hemocytes (Figure 3.7A). Also, clonal tissue expressing *Prp19* RNAi in wing imaginal discs consistently caused a reduction in MAPK levels (Figure 3.7D), although the reduction was not as pronounced as that of *mapk* RNAi. Clonal regions sometimes showed signs of apoptosis (in one of the two RNAi constructs tested and in wing discs in particular) suggesting that these tissues may be more sensitive to knockdown of *Prp19* than S2 cells. Finally, splicing of *mapk* was found to be altered in wing disc segments where *Caper* had been knocked down (Figure S3.10C).

Altogether, these experiments suggest that these transcription and splicing factors are important in regulating MAPK levels, and thus are important for MAPK signaling. Furthermore, our genetic interaction data suggests that they can act in a number of different *in vivo* contexts. In particular, the *Prp19* alleles had an impact in all our RAS/MAPK genetic interaction experiments and displayed some of the strongest phenotype modifications. Thus, of the different groups of candidates, it is quite possible that splicing factors are relevant to the broadest range of RAS/MAPK regulatory and developmental contexts.

Figure 3.8 *Prp19* and *Caper* regulate *mapk* AS

(A) Schematic representation of the four annotated *mapk* splice isoforms observed in S2 cells. Exons are numbered from I to VIII based on the rl-RE transcript for simplicity in addition to the official Flybase exon names (e.g. "rl:14"). Intron lengths are also indicated. (B) An RT-PCR assay encompassing the entire *mapk* transcript (primers bind in exons I and VIII) for the four principal *mapk* isoforms is used to evaluate changes in the *mapk* transcript. In the untreated and *GFP* dsRNA controls, the two most abundant bands on the gel correspond to the RD isoform (topmost) and the RB/RE isoform (immediately below RD). Both *Caper* and *Prp19* knockdown are found to cause important shifts in the abundance and length of the *mapk* transcript (red labels) which differ from those produced by the depletion of the EJC component, *eIF4AIII* (blue labels). By default, all labels refer to the RB/RE isoform unless otherwise indicated. (C) Sequencing of the RT-PCR products from (B) reveals that the shorter products can be attributed to exon skipping events. In particular, exons IV and VII are the most frequently skipped following *Caper* and *Prp19* knockdown. This contrasts with *eIF4AIII* knockdown where we previously observed skipping of multiple consecutive exons[41]. The proportion of normal "N", truncated "T" and frameshifted "F" protein products is also indicated for the sequenced transcripts. (D) *Caper* and *Prp19* cause a shift from the RB/RE/RF forms towards the RD form, which is characterized by retention of the first intron. (E) Exon-exon junction spanning primers are used to detect specific exon skipping events (top panels). Skipping of exons II-III (2^{nd} row) and exons II-IV (3^{rd} row) is more abundant in *eIF4AIII* depleted cells. Exon IV (4^{th} row) and exon VII (5^{th} row) skipping is more prevalent following *Caper* and *Prp19* knockdown. Exon IV and VII skipping could also be detected using assays in which both primers lie within an exon region (bottom panels).


A
RD (1,800)
RE (1,550)
RB (1,569)
RF (1,569)
rl:14
rl:11
rl:12
rl:3
rl:4
rl:5
rl:6
rl:7
rl:10/13
rl:9
rl:2
rl:1
I
II
III
IV
V
VI
VII
VIII
222
205
149
4195
3569
25405
7216
54
6899
B
RNAi
no dsRNA
GFP
eIF4AIII
Caper
Prp19
gfzf
RD
RD∆IV
RD∆VII
RB / RE
RD∆IV∆VII
∆II / ∆VII / RD∆II-IV∆VII
RB∆IV∆VII / ∆II∆IV
RB∆IV∆VII / ∆II-III
RD∆IV∆VI-VII
∆II-IV
RD∆III-IV∆VI-VII
∆II-V
RD∆III-VII / ∆II-IV∆VII
mapk transcript
C
mapk transcript
RNAi
MAPK protein
II III IV V VI VII
N T F
% of exons skipped
21% 15% 21% 3% 9% 9% GFP 47% 9% 44%
27% 29% 51% 10% 27% 59% Caper 16% 49% 35%
29% 54% 79% 31% 54% 81% Prp19 0% 37% 63%
D
RNAi
GFP
Caper
Prp19
24% 61% 52% RD
21% 2% 6% RE
26% 10% 13% RB
9% 0% 0% RF
79% 73% 71% total
E
RNAi
no dsRNA
GFP
eIF4AIII
Caper
Prp19

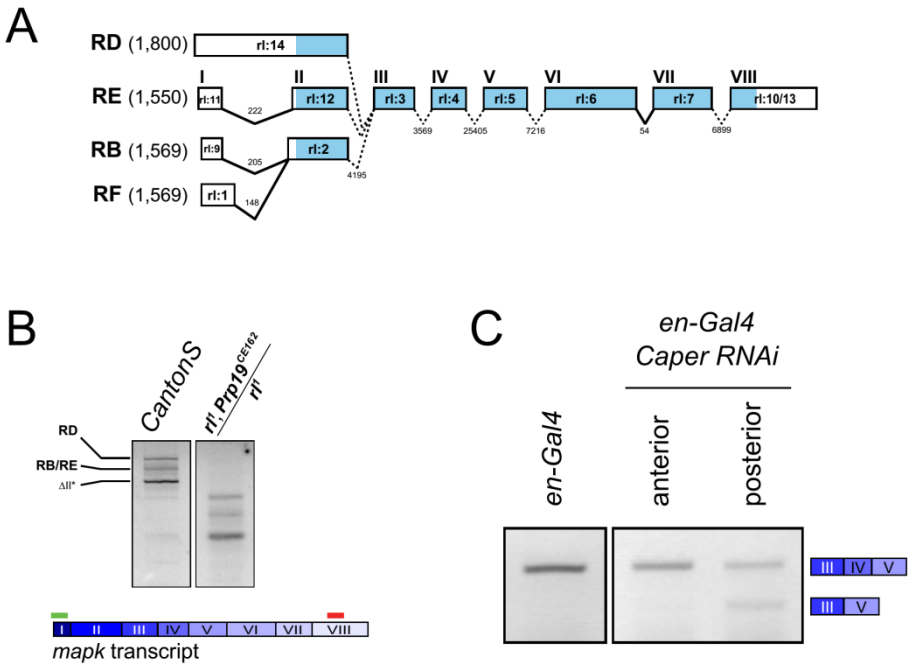


Figure S3.10 *In vivo* evidence for alternative splicing of *mapk*

(A) Schematic representation of the *mapk* splice isoforms from Figure 3.8. (B) The RT-PCR assay spanning the whole *mapk* transcript was used to detect splicing changes in adult flies. Samples were prepared from 5 adult flies of the indicated genotypes. Escapers homozygous for rl^1/rl^1 and also carrying the $Prp19^{CE162}$ mutant displayed splicing changes compared to a CantonS wild type control strain. The products from this experiment were not verified by sequencing though they were similar to those observed in S2 cells upon *Prp19* knockdown (the lower sized band in the WT control is assumed to be due to skipping of exon II, based on our S2 cell data). rl^1 homozygous flies have a reduced amount of *mapk* though the size of the RT-PCR product is unchanged (not shown). (C) *mapk* splicing is altered in larval wing imaginal discs following *Caper* knockdown. A RNAi construct targeting *Caper* was expressed in the posterior segment of wing imaginal discs using an *engrailed*-GAL4 driver. Imaginal discs from 3rd instar larvae were microdissected to separate the posterior (control) and anterior (*Caper* knockdown) segments. The RT-PCR assay spanning exons III-V was used on extracts from both samples to evaluate inclusion of exon IV. All *Prp19* RNAi tested with the *Engrailed-Gal4* driver caused a high rate of lethality and escapers that could be recovered were of reduced size and were not found to contain wing discs.

3.4.8 Analysis of mapk AS induced by Caper and Prp19 depletion

Prp19 and *Caper* RNAi display similar *mapk* RT-PCR profiles, but *Caper* produces a more subtle shift in product size with less lower size bands observable (Figure 3.8B). To verify that these changes in the RT-PCR profile of *mapk* were due to altered splicing, we cloned and sequenced the RT-PCR products. We found that the lower size *mapk* transcripts produced by both *Caper* and *Prp19* RNAi were generated by a series of exon skipping events and, to a lesser extent, to retention of the 5' most intron of the RB/RE isoforms (Figure 3.8A-D and Table S3.VIII). *Prp19* and *Caper* RNAi produced many single exon skipping events with exons IV and VII being the most frequently skipped. These AS events differed from those we had previously observed in EJC depleted samples, where skipping of multiple consecutive exons was observed immediately to the 3' end of exon I[41]. Interestingly, exon skipping events associated with the EJC mostly resulted in frameshifting due to the loss of the start site in exon II and/or skipping of exon III. On the other hand, the skipping of exons IV and VII associated with *Prp19* and *Caper* produce an in-frame deletion potentially giving rise to a truncated protein product. That we did not observe any smaller size products may be due to the epitope being removed or to the smaller products being unstable and degraded (MAPK is mostly comprised of a kinase domain and it is likely that these deletions would disrupt proper folding).

In order to confirm that depletion of *Prp19* and *Caper* produced AS changes in *mapk* that were different from those produced by EJC depletion, we designed RT-PCR assays aimed at detecting specific AS events. Using primer pairs in which the 5' primer overlapped the exon junction between exons I and III or exons I and IV, we were able to detect an enrichment in exon II-III and II-IV skipping in *eIF4AIII* depleted samples, consistent with our previous results. However, these exon skipping events were not enriched in *Caper* and *Prp19* RNAi samples. Conversely, using the same strategy, increased skipping of exons IV and VII was observed for *Caper* and *Prp19*, but not for *eIF4AIII* (Figure 3.8E). Also, using primers targeting the exons on either side of the skipped exons, we were able to detect both the canonical and the alternative transcripts produced by either *Prp19* or *Caper* depletion in both S2 cells and wing disc tissue (Figure 3.8E and Figure S3.10C). In sum, multiple lines of evidence indicate that *Prp19* and *Caper* cause specific changes in *mapk* splicing that differ

from those associated to the EJC. This suggests that at least two different types of regulation act on *mapk* AS. By extension, it is likely that the other splicing factors identified in our screen can be grouped with *Caper* and *Prp19* – and not the EJC – since they produced *mapk* RT-PCR profiles similar to these factors.

Table S3.VIII Detailed Sequencing Data from the mapk RT-PCR Experiment (MS Excel file)

Summary of sequencing data from cloned *mapk* RT-PCR products. 142 clones were succesfully sequenced and aligned in order to investigate the nature of the splicing defects induced by *Caper* and *Prp19* knockdown. *mapk* exons are labeled according to the numbering scheme we used in this study (I-VIII) as well as Flybase exon IDs. “rl:8/10/13alt” refers to an alternate exon located downstream of exon VIII and has been described previously[130]. Presence or absence of exons in a given clone is indicated by “1” or “0”. Exons not annotated in Flybase are indicated in the individual clone entries (ex.: rl:2-B in clone CaperKD_41 is a rl:2-like exon with an alternate splice acceptor site). All exon sequences (annotated and novel) are listed in the “mapk exons” tab.

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

3.5 Discussion

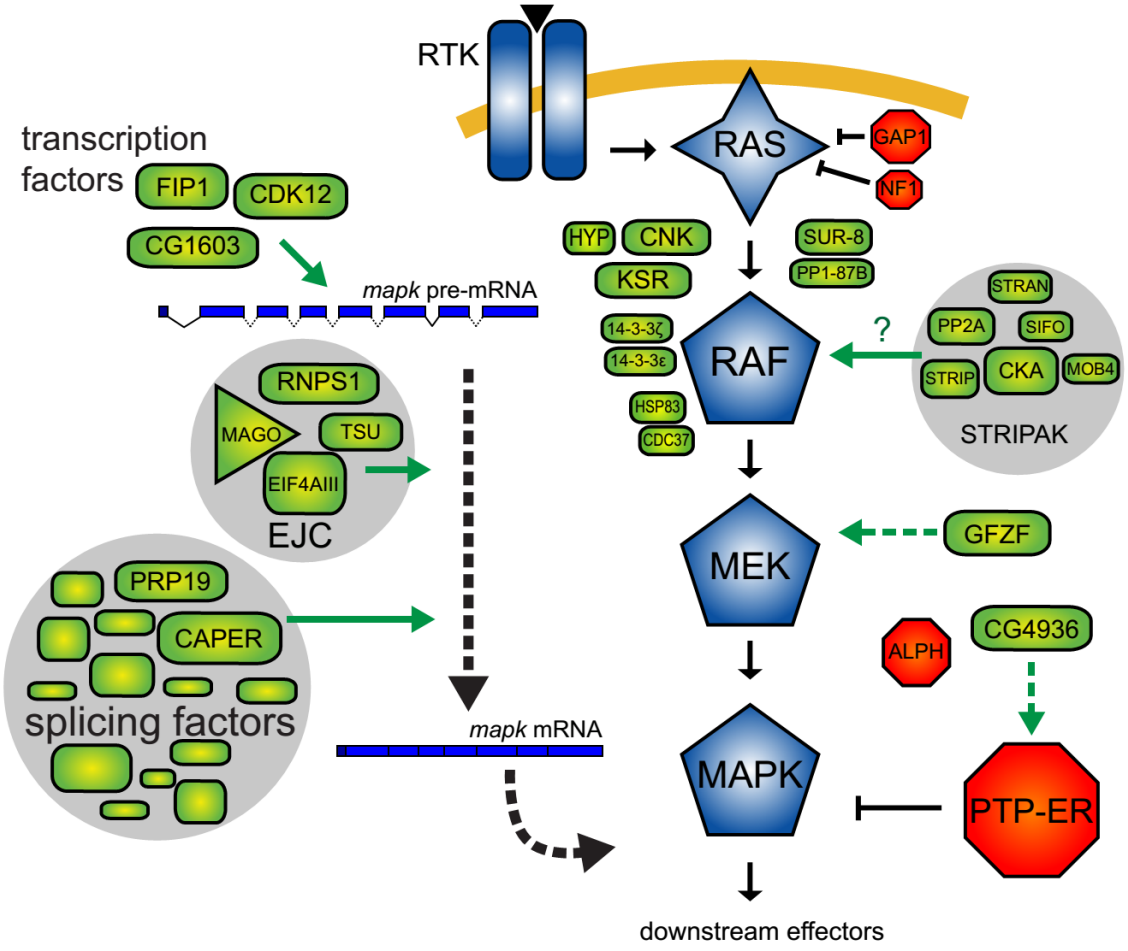
In this report, we presented the results of an RNAi screen for factors influencing signaling between RAS and MAPK in *Drosophila*. Most previously known pathway components have been identified in our screen, including a few that had not yet been found in flies (e.g. *Sur-8, Pp1-87B, Hmgcr* and *Fnta*). Based on analyses of our secondary screen results and publicly available data, we have assessed the specificity of our candidates, grouped them into protein complexes and positioned their effect relative to the core RAS/MAPK pathway components (Table S3.II). We have also evaluated the impact of our candidates on the expression of core components of the RAS/MAPK pathway. From this data, we identified distinct groups of factors with different roles in modulating RAS/MAPK signaling. Furthermore, we show that the largest group of candidates – which is composed of splicing factors – acts specifically on *mapk* expression. This discovery is surprising since most of the previously described regulators of MAPK signaling act at the level of RAF activation. Thus, our results uncover an unappreciated point of control governing *Drosophila* RAS/MAPK signaling which takes place through the control of *mapk* expression (Figure 3.9). Moreover, the identification of two other factors (*gfzf* and *CG4936*) that act to control *mek* and *PTP-ER* expression adds another layer to the gene expression control network surrounding the RAS/MAPK pathway.

When comparing our results to those of three previous RNAi screens examining insulin receptor (InR) and EGFR induced MAPK signaling in *Drosophila* cells in culture[37,62], we found that our validated hit set had a relatively limited overlap with those studies; 44 of our 78 positive regulators were present in their list of 986 positive regulators and 18 of our 28 negative regulators were also present in their group of 1266 negative regulators (Figure S3.11A,B). 331 hits reported in the InR screen were tested for their ability to modulate RAS^{V12} in an assay similar to the one used in our screen. However, very few of the candidates found to modulate RAS^{V12} signaling in their secondary screen were also present in our validated hit set; 14 of our 106 validated genes were found to alter RAS^{V12} induced pMAPK beyond 5% of controls in their study (Figure S3.11C,D) . Furthermore, 9 of these 14 genes were *bona fide* pathway components. This limited overlap between our studies can be

explained by two things: First, their S2R+ InR assay involved the use of an exogenous source of YFP-tagged MAPK where the YFP signal was used to normalize pMAPK signal. This strategy makes detection of factors that modify endogenous MAPK expression impractical (a large proportion of our candidates are involved in exactly this type of regulation). Second, when selecting which hits to follow up in secondary screening, the authors elected to exclude certain genes linked to large molecular complexes, most likely excluding many of the splicing factors that were retained in our study.

Figure 3.9 Regulatory input at the level of MEK, PTP-ER and MAPK expression adds another layer to the network of factors that control RAS/MAPK signaling

Schematic model of proteins associated with RAS/MAPK signal transmission discussed in this work. Components used in secondary screens (GAP1, NF1) are also depicted. *Sur-8*, *PP1-87B* and the five STRIPAK complex components were positioned between RAS and RAF in our epistasis assays. Their position would be consistent with a role in the RAF activation process. As it has been previously shown in mammalian models, SUR-8 and PP1 may be acting on RAF activation by dephosphorylating the N-terminal 14-3-3 binding site. Because PP2A is also known to dephosphorylate this site and because STRIPAK has been characterized as a PP2A-associated complex, STRIPAK may be involved in facilitating PP2A binding to RAF. GFZF was positioned at the level of MEK and was found to impact MEK expression, presumably by regulating *mek* transcription. CG4936 was found to impact expression of the MAPK phosphatase, PTP-ER, and also probably acts at the level of transcriptional regulation. CG1603, FIP1 and CDK12 were found to act on MAPK expression, most likely acting as transcriptional regulators (CG1603) or involved in transcript maturation/processing (FIP1 and CDK12). Finally, components of the spliceosome, splicing factors and the EJC were found to modulate MAPK expression by altering the splicing of the *mapk* transcript. The particular sensitivity of MAPK to disruption of these spliceosome components may be due to their involvement in recruiting specific mRNA processing factors such as the EJC. Alternatively, the reason why *mapk* displays an increased requirement for this set of spliceosome components may be due to a feature in *mapk*'s gene structure. For example, intron length is correlated with sensitivity of transcripts to EJC depletion[41].


RTK
RAS
GAP1
NF1
transcription
factors
FIP1
CDK12
CG1603
mapk pre-mRNA
HYP
CNK
KSR
SUR-8
PP1-87B
14-3-3ζ
14-3-3ε
HSP83
CDC37
RAF
?
STRAN
PP2A
SIFO
STRIP
CKA
MOB4
STRIPAK
RNPS1
MAGO
TSU
EIF4AIII
EJC
MEK
GFZF
PRP19
CAPER
splicing factors
ALPH
CG4936
mapk mRNA
MAPK
PTP-ER
downstream effectors

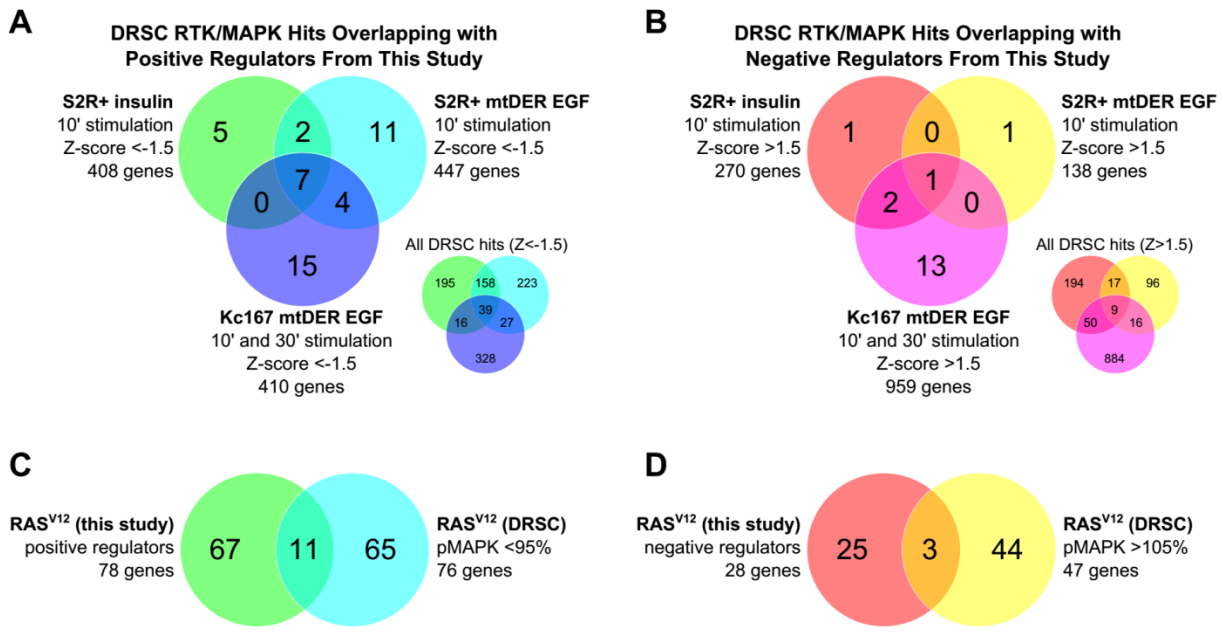


Figure S3.11 Comparison with Previous RTK/MAPK RNAi Screens

(A and B) Overlap of validated positive (A) and negative (B) regulators of RAS^{V12} identified in this study with positive and negative regulators reported in three *Drosophila* RNAi Screening Center (DRSC) RTK/MAPK screens[37,62]. The diagrams include all genes reported as Insulin and EGF modulators (both in S2R+ and Kc cells) with a Z-score +/- 1.5 (986 positive and 1266 negative regulators in total). The large Venn diagrams display hits in the 3 DRSC screens overlapping with hits reported here. The total hits reported in the DRSC screens are displayed in the small diagrams. Reported gene identifiers were first updated to the current annotation before comparison was performed. (C and D) Overlap of validated positive (C) and negative (D) regulators of RAS^{V12} identified in this study with 331 validated hits from the DRSC S2R+ insulin screen that were tested in a RAS^{V12} secondary screen[62]. 123 genes were reported to modulate pMAPK signal by over +/- 5% and were considered as hits for the purpose of this comparison.

3.5.1 Control of mapk expression

With the exception of the well described regulation of RAS by *let-7* family miRNAs[22], surprisingly little is known on the regulation of MAPK module component expression. It was therefore surprising to find that our largest group of hits specifically decreased MAPK levels. Yet, control of MAPK expression is not unprecedented. For example, in yeast, both the FUS3 and HOG1 MAPKs are transcriptional targets of their respective pathways[94]. Pumilio mRNA binding proteins are also known to reduced MAPK activity lowering the expression of the *C. elegans* MAPK, Mpk-1, as well as ERK2/MAPK1 and p38α/MAPK14 in human ES cells, through the binding of specific sites on the 3'UTR of their respective transcripts[95]. Finally, LARP-1 RNA-binding proteins have been found to control the abundance the transcripts of Mpk-1 and other pathway members in the *C. elegans* germ line[96].

In this study, we found that multiple splicing factors act downstream of MEK to specifically control MAPK levels. While some of these factors were associated to the regulation of AS, most were components of the spliceosome or factors that co-purify with the spliceosome. Interestingly, many lines of evidence suggest that AS can be modified by spliceosomal factors[97-106]. Global analyses of splicing events in yeast using a series of temperature sensitive alleles and deletions of splicing factors, found that spliceosome components differed in terms of which splicing events they altered[98,99,105]. Among these, the yeast homologs of *Prp19, Prp8*, *U4-U6-60K*, *l(3)72Ab*, *Prp6*, *SmD2*, *SF3a120*, *CG4849* and *CG10333* – factors included in our set of hits – were found to differentially modulate the splicing of specific sets of transcripts[98]. Moreover, studies conducted in *Drosophila* using an RNAi-based strategy have demonstrated that knocking down many so-called "core" spliceosome components caused specific changes in specific AS contexts[100,101]. In particular, some of the "core" components identified in our screen have been shown to have selective effects. For example, knocking down *Prp6*, *l(3)72Ab, crn*, *Caper*, *SF3a120* and *CG10418*, was found to differentially influence AS of *Dscam*, *para*, *TAF1* and/or *dAdar*[100,101]. In mammals, SmB was recently shown to be involved in regulating inclusion of alternative exons[103]. Finally, our observation that a *Prp19* knockdown had specific effects on MAPK protein levels and the identification of *Prp19* (as well as *Prp8*) in a separate

CNK-based genetic screen supports the notion that these splicing factors are important for *mapk* expression *in vivo*. Thus, the fact that we identified specific spliceosome components in our screen and not others may reflect the importance of these particular components in regulating *mapk* splicing.

The specific elements of the *mapk* gene structure that dictate requirement of particular splicing factors have yet to be determined. Likewise, many – if not most – other genes may have structures that preclude AS regulation by components of the spliceosome. Still, AS processes in at least four other *Drosophila* genes seem to involve some of the same components of the splicing machinery that we have linked to *mapk*, indicating that a common characteristic may dictate the involvement of these spliceosome components in AS. On the other hand, the fact that some of the spliceosome factors are involved in specific AS contexts and not others, suggests that key differences exist. This observation is important as it implies that specific AS events can be modulated by controlling the activity of these spliceosomal factors. In support of this idea, one of the previously mentioned studies indicates that the control of *TAF1* AS by a series of splicing factors – which include *Caper* and other spliceosome components – is downstream of a Camptothecin-induced ATR pathway[101]. Another example is the regulation of CD44 AS by RAS/MAPK-induced phosphorylation of Sam68, which has been found to function by regulating the activity of the U2AF65 spliceosome component[107]. In addition to Sam68, other splicing factors and spliceosome components – including some that were identified in our screen – were recently found to be the targets of ERK phosphorylation[108-110]. Finally, a general splicing repressor, SRp38, has also been shown to function as a sequence specific splicing activator upon phosphorylation in response to cellular stress[111].

One explanation for our results, as well as some of the previous observations, would be that some of the spliceosomal factors in our set may be interchangeable or function as non-essential co-regulators. Thus, removing these components would not greatly disrupt general spliceosome function, but rather lead to an altered spliceosome activity that selectively impacts sensitive AS contexts. An example of this is the interplay between PUF60 and U2AF65 which both function in 3' splice site recognition; the two factors can either work cooperatively or independently, producing different splicing outcomes based on the presence

or absence of either protein[112]. A further example is the stress-induced relocalization of certain spliceosome and splicing factors which leads to changes in spliceosome configuration and AS[113]. Cellular stress has also been associated to the production of "non-productive" AS variants or silent messengers (transcripts that are either degraded by the quality control machinery or that do not encode functional proteins)[114,115]. In fact, the AS we observed in *mapk* is reminiscent of the AS observed for the E3 ligase MDM2 following camptothecin-induced genotoxic stress[116]. In this study, the authors found that stress induced a number of non-productive MDM2 transcript isoforms that resulted in lower MDM2 protein levels and a stabilization of the MDM2 target, p53. It will be important to determine whether stress – or another signal – acts to control MAPK protein levels by inducing the AS we observe in our experiments.

The discovery that MAPK expression is specifically modulated by mRNA processing factors raises multiple questions. Not only will it be vital to define the upstream signals that dictate this activity but it will also be important to assess which other genes are similarly regulated and to identify the common characteristic that renders them sensitive to this type of regulation. Also, the time frame within which these changes occur will have implications as to the role that this type of regulation can play in MAPK signaling. Typically, RAS/MAPK signal modulation has been observed to occur through either rapid post-translational mechanisms or through slightly slower mechanisms involving control of protein stability and transcriptional control. However, the reduction in protein levels we observed as a consequence of AS are only apparent over a period of days, probably because the MAPK protein is relatively stable. This indicates that this regulation will not be relevant over the shorter timeframes of previously characterized regulatory events and also implies that a prolonged stimulus will be necessary to produce the effects we observe. Therefore, control of *mapk* splicing may be more important in the context of certain tissues and organs, in development or in disease. Interestingly, the abundance of core spliceosome components has been shown to be regulated and vary both temporally and across different tissues[117-121]. What is more, disruption in core spliceosome components has also been found to cause changes in AS in diseases such as spinal muscular atrophy and retinitis pigmentosa[122]. Mutations in splicing factors have also been found to occur in a large proportion of myelodysplasia

patients[104,123] as well as in melanoma[106]. Another study has shown that, in glioblastomas and astrocytomas, splicing factors controlled by c-Myc play a role in controlling the expression of pyruvate kinase, a factor that is important for aerobic glycolysis[124]. It will be interesting to explore whether, in contexts such as these, the altered activity of splicing factors may regulate MAPK levels with important functional consequences for either normal or diseased cellular function.

3.6 Methods

3.6.1 Genome-Wide RNAi Screen

pMet-Ras^{V12} S2 cells diluted in Schneider medium (to a concentration of 1x10^6 cells/mL) were distributed in 96 well clear plates (Corning) containing 5µL dsRNA aliquots at a concentration of ~ 200ng/µL. Plates were placed in plastic containers to reduce evaporation and incubated at 27°C for four days. *Ras^{V12}* expression was induced by adding 0.7mM $CuSO_4$ to medium 24h prior to fixation. Cells were resuspended and transferred to concanavalin A coated plates and allowed to settle and adhere for 1h. Cells were then fixed in 4% paraformaldehyde / PBS, washed and blocked in 0.2% Triton X-100 / 0.2% BSA / PBS (PBT/BSA) and incubated overnight with an anti-pMAPK antibody (1/2000; Sigma #M8159), washed in PBT/BSA, and revealed using an anti-mouse Alexa Fluor 555-conjugated secondary antibody (1/1000; Invitrogen #A-21424). DAPI (0.04 µg/ml) was used to stain nuclei. Mowiol (9.6% PVA, Fluka) was added to wells prior to imaging. An automated fluorescence microscopy system (Zeiss Axiovert) was employed for plate imaging. Autofocus, image acquisition and analysis were conducted using MetaMorph (Molecular Devices) software. The cell-scoring application in MetaMorph was used for cell segmentation and quantification of fluorescent signal.

3.6.2 Epistasis

Candidates were assigned to one of three possible epistasis intervals (RAS-RAF, RAF-MEK or MEK-MAPK) based on the data from the following secondary screens: 1. RAS^{V12}, 2. RAF^{ED}, 3. RAF^{CT}, 4. RAF^{EDCT}, 5. MEK^{EE}, 6. MEK^{EE} + PTP-ER dsRNA. The correlations between normalized secondary screen log transformed values and three predetermined epistasis profiles were calculated using a modified *Uncentered Pearson's Correlation*:

$$r(x,y) = \frac{\sum_{i=1}^{n} w_i(x_i y_i)}{\sqrt{\sum_{i=1}^{n} w_i(x_i^2)}\ \sqrt{\sum_{i=1}^{n} w_i(y_i^2)}}$$

Where *r* is the correlation value[-1,1], *x* is the secondary screen value (for screens #1, 2, 3, 4, 5 and 6), *y* is the predetermined profile value and ***w*** is the weight applied to a given screening experiment (where ***w*** =[3 1 1 1 1 2]; see suppl. methods). The following predetermined epistasis profiles ***y*** were used:

RAS-RAF =[1 0 0 0 0 0]
RAF-MEK =[1 1 1 1 0 0]
MEK-MAPK =[1 1 1 1 1 1]

A negative *r* indicates reverse correlation and is observed for positive pathway regulators.

3.6.3 Fly Genetics and Microscopy

Fly husbandry was conducted according to standard procedures. All crosses were performed at 25°C. The *sev-Ras*^{V12} line has been described previously[45]. *Egfr*^{Elp} was described in[125]. The *Cka* alleles[55] were kindly provided by S. Hou. The *csw*^{lf}[126] flies were originally obtained from L. Perkins. The *mapk*^{E1171} allele was identified in a genetic screen as a dominant suppressor of a dominant negative form of KSR[127]. The *Prp19*^{CE162}, *Prp19*^{CE40}, *Prp19*^{TE1036} and *Prp8*^{CE309} alleles were recovered in a genetic screen for modifiers of a dominant negative form of CNK (Baril, Lefrançois et al., manuscript in preparation). RNAi fly lines were obtained from the VDRC[128]. All other fly lines described herein were obtained from the Bloomington stock center.

Adult fly eyes were imaged using a stereomicroscope (Leica MZ FL III) and CombineZP, a freely available software package (http://www.hadleyweb.pwp.blueyonder.co.uk/CZP/News.htm), was used for focus stacking. Adobe Lightroom® and GNU Image Manipulation Program (GIMP) were used for image processing. Wings were mounted in Permount (Fisher) on glass slides and imaged using a Nanozoomer (Hamamatsu).

RNAi clones were generated using a line carrying a heat shock inducible flip-out actin promoter driving the expression of GAL4 and GFP in clonal tissues (*hs-flp;; Act5C>CD2>Gal4, UAS-GFP*). L1 larvae were heat shocked for 15 minutes at 37°C and later collected for dissection upon reaching late L3 (wandering) stage. Third instar eye-antennal and

wing discs were dissected in Schneider medium, fixed and stained with DAPI and an anti-MAPK antibody (1/1000, Cell Signaling #4695) following the same procedure described above for S2 cells.

3.7 Acknowledgements

We thank the Bloomington and VDRC stock centers for fly stocks and cell lines and Therrien lab members for critical reading of the manuscript. We thank Marie-Noëlle Séguin-Grignon for assistance with the RTK and JNK assays. We are grateful to Steve Hou, Liz Perkins, Adam Friedman, Norbert Perrimon as well as the Bloomington and VDRC stock centers for fly stocks and cell lines. We thank Patrick Lapointe, and Christian Charbonneau for a assistance with microscopy, Raphaëlle Lambert for assistance with qPCR and sequencing and Thomas Coudrat for help with the RNA export screen data analysis.

3.8 References

1. McKay MM, Morrison DK (2007) Integrating signals from RTKs to ERK/MAPK. Oncogene 26: 3113-3121.

2. Turjanski AG, Vaque JP, Gutkind JS (2007) MAP kinases and the control of nuclear events. Oncogene 26: 3240-3253.

3. Gineitis D, Treisman R (2001) Differential usage of signal transduction pathways defines two types of serum response factor target gene. J Biol Chem 276: 24531-24539.

4. Pulverer BJ, Kyriakis JM, Avruch J, Nikolakaki E, Woodgett JR (1991) Phosphorylation of c-jun mediated by MAP kinases. Nature 353: 670-674.

5. Kolch W (2005) Coordinating ERK/MAPK signalling through scaffolds and inhibitors. Nat Rev Mol Cell Biol 6: 827-837.

6. Zebisch A, Czernilofsky AP, Keri G, Smigelskaite J, Sill H, et al. (2007) Signaling through RAS-RAF-MEK-ERK: from basics to bedside. Curr Med Chem 14: 601-623.

7. Schubbert S, Shannon K, Bollag G (2007) Hyperactive Ras in developmental disorders and cancer. Nat Rev Cancer 7: 295-308.

8. Rogge RD, Karlovich CA, Banerjee U (1991) Genetic dissection of a neurodevelopmental pathway: Son of sevenless functions downstream of the sevenless and EGF receptor tyrosine kinases. Cell 64: 39-48.

9. Perkins LA, Larsen I, Perrimon N (1992) corkscrew encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso. Cell 70: 225-236.

10. Therrien M, Chang HC, Solomon NM, Karim FD, Wassarman DA, et al. (1995) KSR, a novel protein kinase required for RAS signal transduction. Cell 83: 879-888.

11. Kornfeld K, Hom DB, Horvitz HR (1995) The ksr-1 gene encodes a novel protein kinase involved in Ras-mediated signaling in C. elegans. Cell 83: 903-913.

12. Sundaram M, Han M (1995) The C. elegans ksr-1 gene encodes a novel Raf-related kinase involved in Ras-mediated signal transduction. Cell 83: 889-901.

13. Yoon CH, Lee J, Jongeward GD, Sternberg PW (1995) Similarity of sli-1, a regulator of vulval development in C. elegans, to the mammalian proto-oncogene c-cbl. Science 269: 1102-1105.

14. Raabe T, Riesgo-Escovar J, Liu X, Bausenwein BS, Deak P, et al. (1996) DOS, a novel pleckstrin homology domain-containing protein required for signal transduction between sevenless and Ras1 in *Drosophila*. Cell 85: 911-920.

15. Wassarman DA, Solomon NM, Chang HC, Karim FD, Therrien M, et al. (1996) Protein phosphatase 2A positively and negatively regulates Ras1-mediated photoreceptor development in *Drosophila*. Genes Dev 10: 272-278.

16. Selfors LM, Schutzman JL, Borland CZ, Stern MJ (1998) soc-2 encodes a leucine-rich repeat protein implicated in fibroblast growth factor receptor signaling. Proc Natl Acad Sci U S A 95: 6903-6908.

17. Sieburth DS, Sun Q, Han M (1998) SUR-8, a conserved Ras-binding protein with leucine-rich repeats, positively regulates Ras-mediated signaling in C. elegans. Cell 94: 119-130.

18. Therrien M, Wong AM, Rubin GM (1998) CNK, a RAF-binding multidomain protein required for RAS signaling. Cell 95: 343-353.

19. Hacohen N, Kramer S, Sutherland D, Hiromi Y, Krasnow MA (1998) sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the *Drosophila* airways. Cell 92: 253-263.

20. Sieburth DS, Sundaram M, Howard RM, Han M (1999) A PP2A regulatory subunit positively regulates Ras-mediated signaling during Caenorhabditis elegans vulval induction. Genes Dev 13: 2562-2569.

21. Karim FD, Rubin GM (1999) PTP-ER, a novel tyrosine phosphatase, functions downstream of Ras1 to downregulate MAP kinase during *Drosophila* eye development. Mol Cell 3: 741-750.

22. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, et al. (2005) RAS is regulated by the let-7 microRNA family. Cell 120: 635-647.

23. Baril C, Therrien M (2006) Alphabet, a Ser/Thr phosphatase of the protein phosphatase 2C family, negatively regulates RAS/MAPK signaling in *Drosophila*. Dev Biol 294: 232-245.

24. Roignant JY, Hamel S, Janody F, Treisman JE (2006) The novel SAM domain protein Aveugle is required for Raf activation in the *Drosophila* EGF receptor signaling pathway. Genes Dev 20: 795-806.

25. Douziech M, Sahmi M, Laberge G, Therrien M (2006) A KSR/CNK complex mediated by HYP, a novel SAM domain-containing protein, regulates RAS-dependent RAF activation in *Drosophila*. Genes Dev 20: 807-819.

26. Malumbres M, Barbacid M (2003) RAS oncogenes: the first 30 years. Nat Rev Cancer 3: 459-465.

27. Udell CM, Rajakulendran T, Sicheri F, Therrien M (2011) Mechanistic principles of RAF kinase signaling. Cell Mol Life Sci 68: 553-565.

28. Dougherty MK, Muller J, Ritt DA, Zhou M, Zhou XZ, et al. (2005) Regulation of Raf-1 by direct feedback phosphorylation. Mol Cell 17: 215-224.

29. Fritsche-Guenther R, Witzel F, Sieber A, Herr R, Schmidt N, et al. (2011) Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. Mol Syst Biol 7: 489.

30. Gomez AR, Lopez-Varea A, Molnar C, de la Calle-Mustienes E, Ruiz-Gomez M, et al. (2005) Conserved cross-interactions in *Drosophila* and Xenopus between Ras/MAPK signaling and the dual-specificity phosphatase MKP3. Dev Dyn 232: 695-708.

31. Ebisuya M, Kondoh K, Nishida E (2005) The duration, magnitude and compartmentalization of ERK MAP kinase activity: mechanisms for providing signaling specificity. J Cell Sci 118: 2997-3002.

32. Hurlbut GD, Kankel MW, Lake RJ, Artavanis-Tsakonas S (2007) Crossing paths with Notch in the hyper-network. Curr Opin Cell Biol 19: 166-175.

33. Rajakulendran T, Sahmi M, Lefrancois M, Sicheri F, Therrien M (2009) A dimerization-dependent mechanism drives RAF catalytic activation. Nature 461: 542-545.

34. Harding A, Hancock JF (2008) Ras nanoclusters: combining digital and analog signaling. Cell Cycle 7: 127-134.

35. Mathey-Prevot B, Perrimon N (2006) *Drosophila* genome-wide RNAi screens: are they delivering the promise? Cold Spring Harb Symp Quant Biol 71: 141-148.

36. Boutros M, Ahringer J (2008) The art and design of genetic screens: RNA interference. Nat Rev Genet.

37. Friedman AA, Tucker G, Singh R, Yan D, Vinayagam A, et al. (2011) Proteomic and functional genomic landscape of receptor tyrosine kinase and ras to extracellular signal-regulated kinase signaling. Sci Signal 4: rs10.

38. Goudreault M, D'Ambrosio LM, Kean MJ, Mullin MJ, Larsen BG, et al. (2009) A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. Mol Cell Proteomics 8: 157-171.

39. Abraham D, Podar K, Pacher M, Kubicek M, Welzel N, et al. (2000) Raf-1-associated protein phosphatase 2A as a positive regulator of kinase activation. J Biol Chem 275: 22300-22304.

40. Jaumot M, Hancock JF (2001) Protein phosphatases 1 and 2A promote Raf-1 activation by regulating 14-3-3 interactions. Oncogene 20: 3949-3958.

41. Ashton-Beaucage D, Udell CM, Lavoie H, Baril C, Lefrancois M, et al. (2010) The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in *Drosophila*. Cell 143: 251-262.

42. Roignant JY, Treisman JE (2010) Exon junction complex subunits are required to splice *Drosophila* MAP kinase, a large heterochromatic gene. Cell 143: 238-250.

43. Grammatikakis N, Lin JH, Grammatikakis A, Tsichlis PN, Cochran BH (1999) p50(cdc37) acting in concert with Hsp90 is required for Raf-1 function. Mol Cell Biol 19: 1661-1672.

44. Kato K, Cox AD, Hisaka MM, Graham SM, Buss JE, et al. (1992) Isoprenoid addition to Ras protein is the critical modification for its membrane association and transforming activity. Proc Natl Acad Sci U S A 89: 6403-6407.

45. Karim FD, Chang HC, Therrien M, Wassarman DA, Laverty T, et al. (1996) A screen for genes that function downstream of Ras1 during *Drosophila* eye development. Genetics 143: 315-329.

46. Ramadan N, Flockhart I, Booker M, Perrimon N, Mathey-Prevot B (2007) Design and implementation of high-throughput RNAi screens in cultured *Drosophila* cells. Nat Protoc 2: 2245-2264.

47. Baril C, Sahmi M, Ashton-Beaucage D, Stronach B, Therrien M (2009) The PP2C Alphabet Is a Negative Regulator of Stress-Activated Protein Kinase Signaling in *Drosophila*. Genetics 181: 567-579.

48. Reiss Y, Goldstein JL, Seabra MC, Casey PJ, Brown MS (1990) Inhibition of purified p21ras farnesyl:protein transferase by Cys-AAX tetrapeptides. Cell 62: 81-88.

49. Schafer WR, Kim R, Sterne R, Thorner J, Kim SH, et al. (1989) Genetic and pharmacological suppression of oncogenic mutations in ras genes of yeast and humans. Science 245: 379-385.

50. Kim R, Rine J, Kim SH (1990) Prenylation of mammalian Ras protein in Xenopus oocytes. Mol Cell Biol 10: 5945-5949.

51. Lerner EC, Zhang TT, Knowles DB, Qian Y, Hamilton AD, et al. (1997) Inhibition of the prenylation of K-Ras, but not H- or N-Ras, is highly resistant to CAAX peptidomimetics and requires both a farnesyltransferase and a geranylgeranyltransferase I inhibitor in human tumor cell lines. Oncogene 15: 1283-1288.

52. Wellbrock C, Karasarides M, Marais R (2004) The RAF proteins take centre stage. Nat Rev Mol Cell Biol 5: 875-885.

53. Rodriguez-Viciana P, Oses-Prieto J, Burlingame A, Fried M, McCormick F (2006) A phosphatase holoenzyme comprised of Shoc2/Sur8 and the catalytic subunit of PP1 functions as an M-Ras effector to modulate Raf activity. Mol Cell 22: 217-230.

54. Kemp HA, Sprague GF, Jr. (2003) Far3 and five interacting proteins prevent premature recovery from pheromone arrest in the budding yeast Saccharomyces cerevisiae. Mol Cell Biol 23: 1750-1763.

55. Chen HW, Marinissen MJ, Oh SW, Chen X, Melnick M, et al. (2002) CKA, a novel multidomain protein, regulates the JUN N-terminal kinase signal transduction pathway in *Drosophila*. Mol Cell Biol 22: 1792-1803.

56. Trammell MA, Mahoney NM, Agard DA, Vale RD (2008) Mob4 plays a role in spindle focusing in *Drosophila* S2 cells. J Cell Sci 121: 1284-1292.

57. Dadgostar H, Doyle SE, Shahangian A, Garcia DE, Cheng G (2003) T3JAM, a novel protein that specifically interacts with TRAF3 and promotes the activation of JNK(1). FEBS Lett 553: 403-407.

58. Bond D, Foley E (2009) A Quantitative RNAi Screen for JNK Modifiers Identifies Pvr as a Novel Regulator of *Drosophila* Immune Signaling. PLoS Pathog 5: e1000655.

59. Dickson B, Sprenger F, Morrison D, Hafen E (1992) Raf functions downstream of Ras1 in the Sevenless signal transduction pathway. Nature 360: 600-603.

60. Simon MA, Bowtell DD, Dodson GS, Laverty TR, Rubin GM (1991) Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase. Cell 67: 701-716.

61. Fortini ME, Simon MA, Rubin GM (1992) Signalling by the sevenless protein tyrosine kinase is mimicked by Ras1 activation. Nature 355: 559-561.

62. Friedman A, Perrimon N (2006) A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling. Nature 444: 230-234.

63. Moreno CS, Park S, Nelson K, Ashby D, Hubalek F, et al. (2000) WD40 repeat proteins striatin and S/G(2) nuclear autoantigen are members of a novel family of calmodulin-binding proteins that associate with protein phosphatase 2A. J Biol Chem 275: 5257-5263.

64. Ribeiro PS, Josué F, Wepf A, Wehr MC, Rinner O, et al. (2010) Combined Functional Genomic and Proteomic Approaches Identify a PP2A Complex as a Negative Regulator of Hippo Signaling. 39: 521-534.

65. Horn T, Sandmann T, Fischer B, Axelsson E, Huber W, et al. (2011) Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. Nat Methods 8: 341-346.

66. Dettmann A, Heilig Y, Ludwig S, Schmitt K, Illgen J, et al. (2013) HAM-2 and HAM-3 are central for the assembly of the Neurospora STRIPAK complex at the nuclear envelope and regulate nuclear accumulation of the MAP kinase MAK-1 in a MAK-2-dependent manner. Mol Microbiol 90: 796-812.

67. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, et al. (2009) Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462: 108-112.

68. Fernandez IF, Blanco S, Lozano J, Lazo PA (2010) VRK2 inhibits MAPK signaling and inversely correlates with ErbB2 in human breast cancer. Mol Cell Biol: MCB.01581-01509.

69. Dai MS, Sun XX, Qin J, Smolik SM, Lu H (2004) Identification and characterization of a novel *Drosophila* melanogaster glutathione S-transferase-containing FLYWCH zinc finger protein. Gene 342: 49-56.

70. Ambrus AM, Rasheva VI, Nicolay BN, Frolov MV (2009) Mosaic genetic screen for suppressors of the de2f1 mutant phenotype in *Drosophila*. Genetics 183: 79-92.

71. Sims D, Duchek P, Baum B (2009) PDGF/VEGF signaling controls cell size in *Drosophila*. Genome Biol 10: R20.

72. Kondo S, Perrimon N (2011) A genome-wide RNAi screen identifies core components of the G(2)-M DNA damage checkpoint. Sci Signal 4: rs1.

73. Ow MC, Martinez NJ, Olsen PH, Silverman HS, Barrasa MI, et al. (2008) The FLYWCH transcription factors FLH-1, FLH-2, and FLH-3 repress embryonic expression of microRNA genes in C. elegans. Genes Dev 22: 2520-2534.

74. Herr A, McKenzie L, Suryadinata R, Sadowski M, Parsons LM, et al. (2010) Geminin and Brahma act antagonistically to regulate EGFR-Ras-MAPK signaling in *Drosophila*. Dev Biol 344: 36-51.

75. Zhang W, Mi J, Li N, Sui L, Wan T, et al. (2001) Identification and characterization of DPZF, a novel human BTB/POZ zinc finger protein sharing homology to BCL-6. Biochem Biophys Res Commun 282: 1067-1073.

76. Phatnani HP, Greenleaf AL (2006) Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev 20: 2922-2936.

77. Kornblihtt AR (2006) Chromatin, transcript elongation and alternative splicing. Nat Struct Mol Biol 13: 5-7.

78. Chen HH, Wang YC, Fann MJ (2006) Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. Mol Cell Biol 26: 2736-2745.

79. Preker PJ, Lingner J, Minvielle-Sebastia L, Keller W (1995) The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. Cell 81: 379-389.

80. Hagiwara M, Nojima T (2007) Cross-talks between transcription and post-transcriptional events within a 'mRNA factory'. J Biochem 142: 11-15.

81. Houseley J, Tollervey D (2009) The many pathways of RNA degradation. Cell 136: 763-776.

82. Rigo F, Martinson HG (2008) Functional coupling of last-intron splicing and 3'-end processing to transcription in vitro: the poly(A) signal couples to splicing before committing to cleavage. Mol Cell Biol 28: 849-862.

83. Bhaskar V, Courey AJ (2002) The MADF-BESS domain factor Dip3 potentiates synergistic activation by Dorsal and Twist. Gene 299: 173-184.

84. Li Y, Wang Y, Zhang C, Yuan W, Wang J, et al. (2004) ZNF322, a novel human C2H2 Kruppel-like zinc-finger protein, regulates transcriptional activation in MAPK signaling pathways. Biochem Biophys Res Commun 325: 1383-1392.

85. Staknis D, Reed R (1994) Direct interactions between pre-mRNA and six U2 small nuclear ribonucleoproteins during spliceosome assembly. Mol Cell Biol 14: 2994-3005.

86. Shipman KL, Robinson PJ, King BR, Smith R, Nicholson RC (2006) Identification of a family of DNA-binding proteins with homology to RNA splicing factors. Biochem Cell Biol 84: 9-19.

87. Chan SP, Kao DI, Tsai WY, Cheng SC (2003) The Prp19p-associated complex in spliceosome activation. Science 302: 279-282.

88. Achsel T, Ahrens K, Brahms H, Teigelkamp S, Luhrmann R (1998) The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. Mol Cell Biol 18: 6756-6766.

89. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ (2002) Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA 8: 426-439.

90. Farny NG, Hurt JA, Silver PA (2008) Definition of global and transcript-specific mRNA export pathways in metazoans. Genes Dev 22: 66-78.

91. Dowhan DH, Hong EP, Auboeuf D, Dennis AP, Wilson MM, et al. (2005) Steroid hormone receptor coactivation and alternative RNA splicing by U2AF65-related proteins CAPERalpha and CAPERbeta. Mol Cell 17: 429-439.

92. Coelho CM, Kolevski B, Walker CD, Lavagi I, Shaw T, et al. (2005) A genetic screen for dominant modifiers of a small-wing phenotype in *Drosophila* melanogaster identifies proteins involved in splicing and translation. Genetics 171: 597-614.

93. Lu BY, Emtage PC, Duyf BJ, Hilliker AJ, Eissenberg JC (2000) Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in *Drosophila*. Genetics 155: 699-708.

94. Roberts CJ, Nelson B, Marton MJ, Stoughton R, Meyer MR, et al. (2000) Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. Science 287: 873-880.

95. Lee MH, Hook B, Pan G, Kershner AM, Merritt C, et al. (2007) Conserved regulation of MAP kinase expression by PUF RNA-binding proteins. PLoS Genet 3: e233.

96. Nykamp K, Lee MH, Kimble J (2008) C. elegans La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. RNA 14: 1378-1389.

97. Adams DJ, van der Weyden L, Mayeda A, Stamm S, Morris BJ, et al. (2001) ZNF265--a novel spliceosomal protein able to induce alternative splicing. J Cell Biol 154: 25-32.

98. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C (2007) Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS Biol 5: e90.

99. Clark TA, Sugnet CW, Ares M, Jr. (2002) Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science 296: 907-910.

100. Park JW, Parisky K, Celotto AM, Reenan RA, Graveley BR (2004) Identification of alternative splicing regulators by RNA interference in *Drosophila*. Proc Natl Acad Sci U S A 101: 15974-15979.

101. Katzenberger RJ, Marengo MS, Wassarman DA (2009) Control of alternative splicing by signal-dependent degradation of splicing-regulatory proteins. J Biol Chem 284: 10737-10746.

102. Olson S, Blanchette M, Park J, Savva Y, Yeo GW, et al. (2007) A regulator of Dscam mutually exclusive splicing fidelity. Nat Struct Mol Biol 14: 1134-1140.

103. Saltzman AL, Pan Q, Blencowe BJ (2011) Regulation of alternative splicing by the core spliceosomal machinery. Genes Dev 25: 373-384.

104. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, et al. (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 478: 64-69.

105. Albulescu LO, Sabet N, Gudipati M, Stepankiw N, Bergman ZJ, et al. (2012) A Quantitative, High-Throughput Reverse Genetic Screen Reveals Novel Connections between Pre-mRNA Splicing and 5' and 3' End Transcript Determinants. PLoS Genet 8: e1002530.

106. Furney SJ, Pedersen M, Gentien D, Dumont AG, Rapinat A, et al. (2013) SF3B1 Mutations Are Associated with Alternative Splicing in Uveal Melanoma. Cancer Discov 3: 1122-1129.

107. Tisserant A, Konig H (2008) Signal-regulated Pre-mRNA occupancy by the general splicing factor U2AF. PLoS ONE 3: e1418.

108. Courcelles M, Fremin C, Voisin L, Lemieux S, Meloche S, et al. (2013) Phosphoproteome dynamics reveal novel ERK1/2 MAP kinase substrates with broad spectrum of functions. Mol Syst Biol 9: 669.

109. Carlson SM, Chouinard CR, Labadorf A, Lam CJ, Schmelzle K, et al. (2011) Large-scale discovery of ERK2 substrates identifies ERK-mediated transcriptional regulation by ETV3. Sci Signal 4: rs11.

110. Pan C, Olsen JV, Daub H, Mann M (2009) Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics. Mol Cell Proteomics 8: 2796-2808.

111. Feng Y, Chen M, Manley JL (2008) Phosphorylation switches the general splicing repressor SRp38 to a sequence-specific activator. Nat Struct Mol Biol 15: 1040-1048.

112. Hastings ML, Allemand E, Duelli DM, Myers MP, Krainer AR (2007) Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF(65). PLoS ONE 2: e538.

113. Biamonti G, Caceres JF (2009) Cellular stress and RNA splicing. Trends Biochem Sci 34: 146-153.

114. Dutertre M, Sanchez G, Barbier J, Corcos L, Auboeuf D (2011) The emerging role of pre-messenger RNA splicing in stress responses: Sending alternative messages and silent messengers. RNA Biol 8.

115. Ip JY, Schmidt D, Pan Q, Ramani AK, Fraser AG, et al. (2011) Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. Genome Res 21: 390-401.

116. Dutertre M, Sanchez G, De Cian MC, Barbier J, Dardenne E, et al. (2010) Cotranscriptional exon skipping in the genotoxic stress response. Nat Struct Mol Biol 17: 1358-1366.

117. Ruiz-Lozano P, Doevendans P, Brown A, Gruber PJ, Chien KR (1997) Developmental expression of the murine spliceosome-associated protein mSAP49. Dev Dyn 208: 482-490.

118. Luo HR, Moreau GA, Levin N, Moore MJ (1999) The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes. RNA 5: 893-908.

119. Grosso AR, Gomes AQ, Barbosa-Morais NL, Caldeira S, Thorne NP, et al. (2008) Tissue-specific splicing factor gene expression signatures. Nucleic Acids Res 36: 4823-4832.

120. Castle JC, Armour CD, Lower M, Haynor D, Biery M, et al. (2010) Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification. PLoS ONE 5: e11779.

121. Xiao R, Tang P, Yang B, Huang J, Zhou Y, et al. (2012) Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. Mol Cell 45: 656-668.

122. Wang GS, Cooper TA (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet 8: 749-761.

123. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, et al. (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med 365: 1384-1395.

124. David CJ, Chen M, Assanah M, Canoll P, Manley JL (2009) HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. Nature.

125. Baker NE, Rubin GM (1989) Effect on eye development of dominant mutations in *Drosophila* homologue of the EGF receptor. Nature 340: 150-153.

126. Perkins LA, Johnson MR, Melnick MB, Perrimon N (1996) The nonreceptor protein tyrosine phosphatase corkscrew functions in multiple receptor tyrosine kinase pathways in *Drosophila*. Dev Biol 180: 63-81.

127. Therrien M, Morrison DK, Wong AM, Rubin GM (2000) A genetic screen for modifiers of a kinase suppressor of Ras-dependent rough eye phenotype in *Drosophila*. Genetics 156: 1231-1242.

128. Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, et al. (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. Nature 448: 151-156.

129. Lyne R, Smith R, Rutherford K, Wakeling M, Varley A, et al. (2007) FlyMine: an integrated database for *Drosophila* and Anopheles genomics. Genome Biol 8: R129.

130. Ashton-Beaucage D, Therrien M (2011) The exon junction complex: A splicing factor for long intron containing transcripts? Fly (Austin) 5.

3.9 Additional Notes on the EJC

The exon junction complex was originally identified as a regulator of RAS/MAPK signalling in the RNAi screen described in this manuscript. The characterization of the EJC's function in the regulation of *mapk* splicing is described in detail in a previous paper from our group[1] as well as in another study[2]. In this study, we chose to focus on other hits from the screen and on the splicing factor group in particular. Experiments involving EJC components are used to highlight the differences between the previously descrived EJC function and the novel regulation of *mapk* splicing described here.

Recently, a new EJC associated factor, CWC22, was identified[3-5]. CWC22 is proposed to act as a link between EIF4A3 to spliceosome. Interestingly, CWC22 is also a splicing factor[5] which fits in well with our initial observation that the EJC components can also act on splicing. This also ties in nicely with a recent paper reporting a comprehensive EJC-RNA and EJC-protein interactome[6]. Namely, binding of the EJC to non-canonical sites and assembly into higher order structures that include SR factors is observed, which would be consistent with a function in alternative splicing.

The *Drosophila* homologue of CWC22, *ncm*, while not a hit in our screen, was very close to the cutoff threshold. Moreover, *ncm*, like *Prp8*, was found to enhance the *aos* induced small wing phenotype[7], suggesting that it also acts on RTK/MAPK signalling.

3.10 Supplemental Methods

3.10.1 Genome-wide dsRNA Library

The dsRNA library was generated from dsDNA templates purchased from Open Biosystems (Huntsville, AL; http://www.openbiosystems.com). The Expression Arrest *Drosophila* RNAi Library (versions 1 and 2) covers most of the protein-coding genes in the annotated *Drosophila* genome and is described in detail in[8]. The 15880 individual dsRNAs were generated at the IRIC high throughput screening facility by *in vitro* transcription using T7 RNA polymerase. dsRNA probe concentration was assessed using RiboGreen RNA reagent (Invitrogen). Transcription was repeated when necessary. Products were diluted to an average concentration of 0.2μg/μl and purified on filter plates (Millipore). Additionally, 25% of the dsRNAs were verified on gel (E-Gel, Invitrogen). The library is stored in sealed 384-well plate aliquots at -30°C.

3.10.2 Primary Screen Assay

The Z'-factor[9] was used to calculate assay suitability for high throughput screening considerations. This calculation was done using *GFP* dsRNA negative controls and *mek* and *PTP-ER* dsRNA positive controls for suppressors and enhancers, respectively. The calculated Z' was 0.643 for *mek* and 0.175 for *PTP-ER*. This meant that the assay was highly suitable for detection of suppressors, but was not as robust in the case of enhancers. This lower Z' reflects the slightly larger variation in the calculated signals of *PTP-ER* controls as well as the lower difference between *GFP* and *PTP-ER* control means (Figure S3.1B). The lower reliability in the case of enhancers may be caused by the strong pathway activation state due to the overexpression of RAS^{V12}. The difference between the negative control and enhancer populations would therefore not be as pronounced because pathway activity is already at a heightened state.

For the primary and promoter validation screening steps, the library dsRNAs as well as control dsRNAs were arrayed in bar-coded 96 well clear flat bottom cell culture microplates (Corning) by a Biomek FX (Beckman) in a sterile culture hood. The cells were then plated at a concentration of 50,000 cells per well using a multi-channel pipettor (BioHit). After a three

day 25°C incubation period, 0.7mM $CuSO_4$ (diluted Schneider medium) was added using a multi-channel pipettor and cells were incubated for another 24h, for a total of four days of dsRNA incubation. Following incubation and $CuSO_4$ induction, the subsequent staining steps were performed using an automated procedure developed on an integrated roboticized platform. Briefly, a Biomek was used to resuspend and transfer cells to a second set of identical ConA coated plates. After a 1h incubation step, cells were fixed in a 4% paraformaldehyde solution for 15'. A PBT (PBS Triton 0.2%) BSA 0.2% solution was used to wash, permeabilize and block cells using a washing station (BioTek ELx405). PBT BSA 2% was used as a solution for incubation in both anti-pMAPK, as well as a mixed anti-mouse Alexa555 antibody and DAPI solution. These reagents were distributed using a Multidrop (Thermo). Following the staining procedure, cells were mounted in Mowiol (9.6% PVA, 25% Glycerol) and sealed using adhesive plate covers. The plates were read using a wide field inverted microscope (Axiovert, Zeiss) equipped with a motorized XY stage as well as an objective turret mounted on a mobile Z axis. Metamorph (Molecular Devices) software was used to automate autofocus and image capture procedures. Image were acquired at a 20X magnification at a rate of two fields per well (primary screen) or five fields per well (primary screen confirmation step and validation step).

To control for problems in cell resuspension and transfer, the initial plates were stained with Methylene Blue in order to detect cases of increased cell adhesion. Stain was then dissolved in DMSO and OD 600nm was measured to approximate cell number. The Methylene Blue images and values are available on the IRIC *RNAi* database.

Quantification was carried out using the Cell Scoring application in Metamorph. Plate-specific background values were subtracted to all images. The integrated intensity from Alexa555 positive cells was divided by the total DAPI cell nucleus count to obtain an average measure of pMAPK signal per cell. The values of the *GFP* negative control wells were used for plate data normalization.

Cutoffs for our primary dataset were set at -0.45 (73.2%; suppressors) and 0.34 (126.8%; enhancers) $\log_2$ ratio of the *GFP* dsRNA control average (Figure 3.1A). By comparison, the standard deviation of the dsRNA set was 0.21 and that of the *GFP* negative controls was of 0.09. Out of the 1322 *GFP* control well values, one is outside the cutoff

margins (Figure S3.1C), allowing to estimate the rate of false positives due to assay noise at 0.0756%. Out of 1598 positive controls tested, only 6 positive controls, 1 *mek* and 5 *PTP-ER*, resulted in false negatives (Figure S3.1C). Accordingly, the global false negative rate for the assay could therefore be estimated at 0.375%, or more specifically, at 0.125% for suppressors and at 0.626% for enhancers. The selected cutoff value for suppressors is over 2x the standard deviation of the dsRNA set and 4.5x the GFP standard deviation. The cutoff for enhancers was more permissive (1.6x and 3.9x) in order to include more enhancers in our set.

Following the genome-wide screen, probes that scored as preliminary hits were re-tested to confirm the primary screen data. Only those probes were retained whose average signal (of the primary and primary confirmation screen values) was outside the cutoffs. Of the confirmed hits, additional candidates were also eliminated based on cell clustering phenotypes (visually scored) or low cell count values, which interfere with proper quantification.

3.10.3 Validation Screens

Because MTF-1 (Metal response element-binding Transcription Factor-1), a key transcription factor involved in Metallothionein promoter (*pMet*) activity[10], was also amongst the strongest suppressors identified (Table S3.I), we conducted a first series of validation assays using *pMet-GFP* and *pMet-HA-Ras^{V12}* cell lines to eliminate false positives that influenced *pMet* activity. This was done by measuring the GFP and HA signals and comparing them to the primary screen pMAPK signal. Because these assays relied on the production of a protein signal, they also allowed for the elimination of hits that affected basic cellular machineries involved in protein synthesis such as mRNA export and translation. Depletion of *MTF-1*, but not of RAS/MAPK pathway components, caused reduction in GFP and HA levels, allowing to identify false positive hits of this type (Figure S3.2). Both validation screens were carried out using a similar protocol to the one used in the primary screen (see above).

We observed that pMAPK levels (induced by *pMet-Ras^{V12}*) was less sensitive to disruption of promoter activity than GFP or HA levels (induced by *pMet-GFP* or *pMet-HA-Ras^{V12}*) (Figure S3.2). Therefore, we used a cutoff that was a function of the primary screen signal for the GFP and HA screens:

$$\text{GFP screen cutoff} = x$$

$$\text{HA screen cutoff} = 10^{2\log_{10} x - 2}$$

Where *x* is the normalized signal from the GFP or HA screen. The calculated cutoff is applied to the primary screen result. Candidates with primary screen pMAPK value below the cutoff (for suppressors) or above the cutoff (for enhancers) are retained (Figure S3.2). This result indicates that there is no effect on promoter activity or that it is not significant enough to account for the observed variation in pMAPK. 233 hits passed these *pMet* validation steps (Table S3.I).

In the second validation step (Figure S3.1A), we addressed potential dsRNA off-target effects. We synthesized new, non-overlapping, dsRNA probes targeting these 233 genes and retested them in the RAS^{V12} assay in four replicates, but in conditions otherwise similar to those described for the primary assay. Importantly, these probes were designed to be devoid of off-predicted target sequences using either the E-RNAi (http://www.dkfz.de/signaling/e-rnai3/)[11] or SnapDragon (http://www.flyrnai.org/cgi-bin/RNAi_find_primers.pl) dsRNA design tools. 101 probes from this validation step recapitulated the effect of the primary probes on pMAPK levels.

3.10.4 Secondary Screens: Functional Assays

The procedure for conducting the secondary screens was similar to the one described for the primary and validation steps, except that the protocol was not automated. In total, 13 secondary screens were performed using the secondary probe set that had passed the validation step criteria. 12 signaling screens (Figure S3.3A) and a second *pMet-GFP* promoter validation step were performed. The same dsRNA probe set was used for the protein and transcript level screens as well as the bulk mRNA export screen. For the secondary RAS^{V12} assay, the data from validated probes (in the second validation step) was used here and considered as secondary screen data). Validation of the 11 other secondary assays is shown in Figure S3.3 along with a description of the induction procedures. All dsRNA incubations were performed for four days in duplicate or larger replicate sizes.

As in the primary screen assay, automated microscopy and image analysis were used for quantification and secondary screen and functional assay data was also normalized to plate-specific dsRNA controls (*GFP* dsRNA, or *mek* dsRNA in the case of the *pMet-GFP* assay).

3.10.5 qPCR and RT-PCR

For the qPCR secondary screen, triplicate S2 cell culture replicates were lysed using a Cells-to-cDNA™ (Life technologies) lysis buffer following manufacturer guidelines. The qPCR was performed in two steps using a High Capacity cDNA Reverse Transcription Kit (Life technologies) for reverse transcription and SYBR® Green PCR kit (Life technologies) for DNA quantification. When selecting hits for follow-up analysis, we considered significant changes in transcript abundance greater than +/- 1 $\log_2$ fold change with a p-value < 0.005 (unpaired two-tailed Student's t-test). We eliminated dsRNAs that significantly altered the levels of more than one transcript. We performed confirmation experiments only on the candidates for which the modulation of transcript levels was consistent with the functional effect (ex.: *Cdk12* dsRNA decreases *mapk* transcript levels and also supresses RAS^{V12}-induced pMAPK). For the qPCR confirmation experiment, RNA extracts were prepared using TRIzol® reagent (Life technologies) following manufacturer instructions. qPCR was then performed in technical triplicate as in the qPCR screen.

For qPCR performed on larval eye-antennal imaginal discs, RNAi constructs were expressed under the control of the same heat shock flip-out system used to generate clones for microscopy. To maximize RNAi expression, L1 larvae were heat shocked 40 minutes at 37°C. In these conditions, GFP positive clones covered the majority of dissected imaginal discs. For each sample, 3 eye-antennal imaginal discs were lysed in 5uL of Cells-to-cDNA™ lysis buffer (Life Technologies), treated with DNAse and immediately used to conduct RT using SuperScript™ II. Three samples were prepared for each condition tested (biological triplicates) and qPCR was performed in technical triplicate.

For S2 cell culture RT-PCR experiments, RT was carried out on RNA samples prepared using TRIzol® reagent (Life technologies). For the RT-PCR screen, we selected a subset of hits that were positioned downstream of MEK as well as predicted splicing factors

positioned at other intervals. S2 cells were seeded in triplicate in 96 well plates with dsRNA. The samples were then pooled for TRIzol lysis and RNA extraction. Following this, RT-PCR was performed using a SuperScript™ II (Life Technologies) as per manufacturer guidelines. PCR was performed using Taq DNA polymerase (Life Technologies) and samples were loaded on a 2% agarose gel. All other RT-PCR experiments were performed using a High Capacity cDNA Reverse Transcription Kit (Life technologies). Cloning and sequencing of *mapk* RT-PCR products was performed as previously described in Ashton-Beaucage et al.[1]. The second exon of the RE and RB/RF (rl:12 and rl:2) as well as the first exon of RD (rl:14) were considered as exon II for purposes of calculating the exon skipping rates.

Adult fly RT-PCR experiments were performed on RNA extracted from 5 adult male flies using TRIzol. Flies were first frozen at -80°C and homogenized in TRI reagent. Following this, RT-PCR was carried out using the conditions described above.

For the RT-PCR on larval tissue, wing imaginal discs were first extracted from third instar larvae. A *UAS-GFP* under the control of the *engrailed-GAL4* promoter was used to mark the anterior segment of the wing discs. A UV lamp system (NIGHTSEA™ Fluorescence Viewing Systems) was used to visualize GFP and guide manual micro-dissection with sharpened tungsten needles. A single wing disc fragment was homogenized in 5uL of Cells-to-cDNA^{TM} lysis buffer (Life Technologies), treated with DNAse and immediately used to conduct RT using SuperScript™ II.

All primer sequences used for the qPCR and RT-PCR assays are listed in Table S3.IX.

assay type	assay name	gene symbol	OI 5'	OI 3'
qPCR	qPCR Ras85D	*Ras85D*	AGCGGAAAAACAAGGAGGA	GTTTGTGTGGGTGACTGTGC
qPCR	qPCR raf	*raf*	TTTAAATTGAACACGCTCATCG	AGATGTTGTTTGACTTGAGGTCTCT
qPCR	qPCR mek	*mek*	CGGTAAATACGGAGGCAACT	TCCCCAGCAAACTGTGTGT
qPCR	qPCR rl	*rl*	ACGGAAGTTCCTCAATCTAATGCTGAAGT	GTTTGTTAGCGTGTCATCCGCAGA
qPCR	qPCR ksr	*ksr*	CGCAGCAGGAGATTCGTT	GCAGCAGCTCGGAGAAGTAT
qPCR	qPCR cnk	*cnk*	AGGCGGTGGAGAACTTAAGAAACTTCCATT	TGGCCAGTTCGCGATGCAGAT
qPCR	qPCR PTP-ER	*PTP-ER*	AGCTGACTAGCTTGCTGGGGGAA	TATTGGGCAGTGGACCCTGCGT
RT-PCR	RT-PCR screen mapk	*mapk*	TACGCCGTCGATTTTGATAAATC	TTGAGTCCACGCAATATCTGA
RT-PCR	RT-PCR mapk full transcript	*mapk*	CGCCGTCGATTTTGATAAATCATATTTACGC	AGGCGCATTGTCTGGTTGTCGT
RT-PCR	RT-PCR ksr	*ksr*	GCGCGTAGTGGGCCTTAGCC	GAGCTGCTGCCCGCCTGATC
RT-PCR	RT-PCR raf	*raf*	CGGGCCGCAGGAGAAGGA	GCGGCCTAAAGAGCGGTCGA
RT-PCR	RT-PCR Mek	*mek*	CGCCGCCATTCAAAACCCCCT	CGCAGAGGCGTCTAGTTGGGC
RT-PCR	RT-PCR cnk	*cnk*	CCGAGTGGACGCCCGACCAG	TGCCAGCCCACAACGGTCTG
RT-PCR	RT-PCR Ras85D	*Ras85D*	CCGCGCGCGAGAAGAGAGAG	TAAAAGGGCGCCTCGCAGCC
RT-PCR	RT-PCR mapk I-II II-III	*mapk*	CATTTGTCTAAAGAGTAAAAAAAATATTTCAA	GTGTCATCCGCAGACACAAC
RT-PCR	RT-PCR mapk I-IV IV-V	*mapk*	GCATTTGTCTAAAGAGTAAAAAAAAATTATTGA	TGTGATCATTACTTAGCCTCTGTG
RT-PCR	RT-PCR mapk I-V V-VI	*mapk*	ATTTGTCTAAAGAGTAAAAAAAAAGGCTA	AATCCAAAGTCGCAAATTTTTAAG
RT-PCR	RT-PCR mapk III-V VI	*mapk*	TTGACCAGATTTAAACATGAAAACAG	AGTATGATCGTGCTCGGGAT
RT-PCR	RT-PCR mapk VI-VIII VIII	*mapk*	GATTTAGAGTGTATTATTAATGAAAAGCCTGT	TAAGGCGCATTGTCTGGTTG
RT-PCR	RT-PCR mapk III V	*mapk*	ACCATATTGACCAGATTTAAACATGAA	ACTTGGCTTAAGGTCCCGAT
RT-PCR	RT-PCR mapk VI VIII	*mapk*	TCCCGGGACGATTTAGAGTG	AAATGGCACTTCAGCGACAG

Table S3.IX qPCR and RT-PCR Primers

Sequence of RT-PCR and qPCR primers used in this study. dsRNA primer sequences are available at the IRIC *RNAi* database (http://www.bioinfo.iric.ca/iricrnai).

3.10.6 FISH Total mRNA Export Screen

The fluorescence *in situ* hybridization (FISH) was adapted from Herold et al.[12]. An oligo-dT FISH probe was used to visualize polyadenylated total mRNA. Cells were transferred to 96 well concanavalin coated A glass-bottomed plates and fixed in 4% paraformaldehyde (4% PFA in PBS) for 10 minutes. The PFA was then quenched in PBS-Glycine 1.25M for 5 minutes and PBS-triton 0.5% (10 min) was used for permeabilization. PBS wash steps preceded and followed each of these steps. Next, samples were incubated in pre-hybridization buffer (2X saline-sodium citrate[SSC; 0.3M sodium chloride, 30mM trisodium citrate, pH 7.0], 20% formamide, 0.2% BSA, 1mg/mL yeast tRNA) for 15 min at 37°C, then hybridized with the oligo-dT probe for 2h in hybridization buffer (2X SCC, 20% formamide, 0.2% BDS, 1mg/mL yeast tRNA, 10% dextran sulfate, 0.1 pmol/uL Cy3-oligo-dT) at 37°C. Following hybridization, samples were incubated twice in 2X SCC 20% formamide and twice in 2X SCC for 5 minutes at 42°, once in 1X SCC for 5 min, once in DAPI (DAPI 0.1 µg/mL in PBS) and washed once in PBS before being coated in Mowiol.

Cells were imaged by automated microscopy using an Operetta high content imaging system (PerkinElmer). Segmentation and signal quantification was performed using Harmony® software (PerkinElmer).

3.10.7 The IRIC RNAi Database

Images and quantitative data were imported into the IRIC RNAi database (www.bioinfo.iric.ca/iricrnai/). IRIC RNAi is a custom SQL database intended for use in storing and referencing RNAi screen data conducted at the Institute for Research in Immunology and Cancer (IRIC). The screening data described in this paper is made available at *IRIC RNAi* as a web resource accompanying this paper in the "Ras^{V12} pMAPK" project section. Information on the screening assays described in this study is also made available. The database allows access to specific experimental results, probe sequence and genomic information. It is also possible to view the automated microscopy images used for signal quantification. To find information relating to a specific gene or RNAi probe, a gene and probe identifier search can be performed. A genome browser allows the user to visualize dsRNA probe localization relative to gene sequence.

Part of the image analysis was also conducted through the IRIC RNAi database. Out-of focus images from the primary screen were filtered out using a custom algorithm. The algorithm calculates the average frequency of the set of DAPI images using a two-dimensional Fourier transform. A threshold was set to sort out blurred images with lower average frequencies. A manual quality control and visual inspection was also performed using the plate overview and image viewing functions. All negative controls and primary hits were inspected visually to sort out errors due to obvious staining artefacts.

3.10.8 Additional information for epistasis analysis

We used a modified *Pearson's Uncentered Correlation* with weight (***w***) values in order to correct for the fact that data from one RAS-based screen, three RAF-based screens and two MEK-based screens were being used to position our candidates. w_1 (for 1. RAS^{V12}) was set to 3, w_2, w_3, and w_4 (2. RAF^{CT}, 3. RAF^{ED}, 4. RAF^{EDCT}) to 1, w_5 (5. MEK^{EE}) to 1 and w_6 (6. MEK^{EE} + PTP-ER dsRNA) to 2. Thus data from RAS, RAF and MEK contribute equally to the calculated correlation *r*. Screen #6 (MEK^{EE} + *PTP-ER* dsRNA) was given a higher weight than screen #5 (MEK^{EE}) based on the our observation that the addition of *PTP-ER* dsRNA produced a higher signal induction (Figure S3.3C) and that modulation of this signal was more robust than for MEK^{EE} alone (data not shown).

Candidates with poor correlation ($-0.5 < r < 0.5$) in all three profiles were not assigned an epistasis interval. Candidates with similar *r* values for all three profiles were marked as ambiguous based on a confidence score:

$$c = \left| r_1 - \frac{r_2 r_3}{2} \right|$$

Where *c* is the epistasis confidence score and r_1, r_2 and r_3 the first, second and third highest correlation values. Thus, candidates for which the highest correlation value (r_1) was close to the average of the other two calculated correlation values ($c < 0.2$) were labeled as ambiguous. See Table S3.III for the full set of correlation values and confidence scores.

Both unsupervised clustering analyses presented in Figure 3.3A were performed on $\log_{10}$ transformed secondary screen values using a hierarchical clustering function and *Pearson's Uncentered* distance metric in the MeV application (http://www.tm4.org/mev/)[13]. The 3D

graphical representation shown in Figure 3.2C was assembled using OriginLab (http://www.originlab.com/).

3.10.9 Specificity score

Four datasets were used to calculate the specificity score: 1) secondary functional screen data, 2) mRNA export screen data, 3) hit occurrence in previously published RNAi screens and 4) the western blot secondary screen. Scores assessed from these three sources were added to obtain the final specificity score (with a higher specificity score corresponding to a less specific hit).

Secondary screen data was analyzed as follows. Specificity with regards to JNK signaling and effects on *pMet* promoter activity were both assessed. As in the epistasis analysis, *Pearson's Uncentered Correlation* was used to assess correlation with predetermined specificity profiles in both cases. The correlations were calculated on normalized secondary screen Log_{10} transformed values:

$$r(x,y) = \frac{\sum_{i=1}^{n}(x_i y_i)}{\sqrt{\sum_{i=1}^{n} x_i^2}\ \sqrt{\sum_{i=1}^{n} y_i^2}}$$

Where *r* is the correlation value[-1,1]. The secondary screen values *x* and specificity profiles ***y*** are specified below.

For JNK specificity assessment *x* corresponds to data from screens # 1, 11 and 12 (1. RAS^{V12}, 11. PGN and 12. $RAC1^{V12}$). Correlation with 4 possible profiles was assessed:

PGN-$RAC1^{V12}$ =[1 1 1] (score: 2; similar variation in RAS^{V12}, PGN and $RAC1^{V12}$)
PGN =[1 1 0] (score: 1; similar variation in RAS^{V12} and PGN only)
$RAC1^{V12}$ =[1 1 0] (score: 1; similar variation in RAS^{V12} and $RAC1^{V12}$ only)
NE =[1 0 0] (score: 0; no effect: PGN and $RAC1^{V12}$ not similar to RAS^{V12})

Specificity scores noted above were assigned to candidates closest to the indicated profiles. Candidates with low correlation values ($-0.5 < r < 0.5$) were assigned a score of 1.

Total mRNA nuclear export scores exceeding 4 times that of the *GFP* dsRNA controls were assigned a score of 1.

For assessment of effects on *pMet* promoter activity, *x* corresponds to data from screens # 1, 7, 10, 8, 12 and pMet-GFP (1. RAS^{V12}, 7. *Gap1/ Nf1* dsRNA, 10. Insulin, 8. EGFR SPI, 12.

$RAC1^{V12}$ and the pMet-GFP screen conducted using secondary probes). The pMet-$RAC1^{V12}$ and pMet-GFP screens were included as indicators of non-specific modulation of pMet (in processes not related to MAPK signaling). The Insulin and *GAP* RNAi screens were included as indicators of specific modulation of MAPK (in a context not dependent on pMet-induced stimulation of pathway activity). Correlation with 4 possible profiles was assessed:

pMet =[1 0 0 1 1 1] (score: 2; modulates pMet, but not pMet-independent assays)
All =[1 1 1 1 1 1] (score: 1; similar variation in all assays)
none =[1 0 0 0 0 0] (score: 0; no modulation of the other assays)
non-pMet =[1 1 1 0 0 0] (score: 0; modulates pMet-independent, but not pMet assays)

As for the JNK test, specificity scores noted above were assigned to candidates closest to the indicated profiles. Candidates with low correlation values ($-0.5 < r < 0.5$) were assigned a score of 1. Notably, the epistasis positioning in either the RAS-RAF or RAF-MEK interval is also an indication of the lack of non-specific effect on *pMet* expression (because positioning in these intervals implies lack of effect on *pMet* expression of either RAF and/or MEK). To reflect this, specificity scores were corrected as follows: -2 for RAS-RAF candidates and -1 for RAF-MEK, while candidates. The final *pMet* specificity score could not be lower than 0.

Hit occurrence in previously published *Drosophila* RNAi screens was evaluated using data from FlyMine. The number of screens in which the candidate was also identified was directly added to the specificity score. As a reference, RAS/MAPK components were identified in an average of 4 previous RNAi screens. On the other hand, some splicing factors and ribosome components were hits in over 10 previous screens.

Finally, the effects on MAPK, RAS, CNK and AKT levels were also used to assess specificity. The effects on protein levels in the western blots shown in Figure S3.4 were scored visually by three separate evaluators. Specificity scores were attributed as follows: No effect on protein levels: 0; depletion of two or more proteins: 1 (weak), 2 (medium), 3 (strong); and depletion observed for a single MAPK pathway protein: -1.

Global specificity scores were defined as follows: high specificity (score ≤ 5) medium specificity hits ($5 <$ score ≤ 8) and low specificity (score > 8). The final specificity score as well as a summary of the different specificity assessments is presented in Table S3.V.

3.10.10 Protein Interaction Network

We used the DroID[14] Cytoscape[15] plugin to access *Drosophila* genetic interaction data. The combined network was created by combining the different homolog networks using the Advanced Network Merge Cytoscape plugin.

3.10.11 Additional Fly Genetics and Immunohistochemistry Information

Homozygous *Cka* mutant clones were generated using the *flp-FRT* technique[16]. Pupal eye discs were prepared as described in[17] and stained with DAPI, mouse anti-prospero (1/100, DSHB) and rat anti-Elav (1/1000; DSHB).

3.10.12 Larval Hemocyte Assay

Ras^{V12} is known to induce overproliferation of larval hemocytes[18]. We used *Hemolectin*-GAL4 (*Hml*-GAL4) to drive expression of Ras^{V12} in larval hemocytes. RNAi hairpin and GFP constructs were also expressed under the control of the UAS promoter to knockdown candidate genes and mark hemocytes with GFP. Two different fly lines containing *UAS-*Ras^{V12} (on either chromosome 2 or 3) and *UAS-lacZ* (used in lieu of an RNAi contruct) were used as positive controls for increased larval hemocyte proliferation. These two controls were needed as the UAS-RNAi constructs were located on either chromosome 2 or 3, and *UAS-*Ras^{V12} was added to the chromosome not containing an RNAi construct.

Extraction and scoring of hemocytes was conducted as follows: the hemolymph from 3rd instar larvae was extracted and transferred to Schneider medium in 384-well cell culture plates (Greiner). Each well contained the extracted hemolymph of one larva. Fixed and DAPI-stained cells were then counted using automated microscopy on an Operetta high content imaging system (PerkinElmer). Segmentation and signal quantification was performed using Harmony® software (PerkinElmer). Two separate experiments were conducted, each with at least three biological replicates per condition tested. RNAi were also tested alone (without *RasV12*) to verify if they impacted cell proliferation. None of the RNAi tested (*gfzf, CG4936, Prp19, Caper, CG1603, mapk*) appeared to alter hemocyte counts (not shown) although

hemocyte counts were low in larvae not expressing Ras^{V12} making these results less robust than our experiment with Ras^{V12} expressing hemocytes.

3.10.13 Cell Lines and Plasmids

Drosophila S2 cells were cultured at 27°C in Schneider medium (Invitrogen) supplemented with 10% fetal bovine serum. S2 cell lines stably transfected with the following constructs were used to conduct RNAi screening assays: *pMet-*Ras^{V12}[19], *pMet-*raf^{ED}[20], *pMet-*mek^{EE}[21], *pHS-*sev^{S11}[22], *pMet-EGFR* (a gift from N. Perrimon), *pMet-*$Rac1^{V12}$[23]. *pMet-*raf^{CT} and *pMet-*raf^{EDCT} respectively encode N-terminal deleted (2-371) RAF made from *pMet-RAF* and *pMet-*RAF^{ED}[20]. *pMet-GFP* was made by inserting an *EGFP* cDNA in the *pMet* vector[22].

To construct tagged forms of CKA, STRIP, SLMAP, FGOP2 and MOB4, the respective cDNAs were amplified by PCR with a forward primer containing the desired epitope tag and cloned into the *pMet* vector. For the GFP fusion to FGOP2, the tag was removed from the STag-SIFO construct and replaced with a cDNA encoding GFP. In the case of CKA, MOB4, and FGOP2, the tags were fused to the first amino acid following the initiator methionine of the only predicted ORF. For SLMAP, the HSV epitope was fused to the valine at position two of the 897 amino acid ORF (CG17494-PA). For STRIP, the tags were fused to the leucine at position three of the 882 amino acid ORF (CG11526-PA). All constructs were verified by sequencing.

3.10.14 Protein analysis

For western blot analyses of S2 cell lysates, cells were lysed in cold lysis buffer (20 mM Tris, pH 8.0, 137 mM NaCl, 10% glycerol, 1% Igepal CA-630, 1 mM EDTA) supplemented with 1X phosphatase inhibitor cocktail (Sigma #P2850), 10 μg/mL each aprotinin and leupeptin, and 1 mM PMSF. Lysates were then clarified by centrifugation at 10,000 x g. Protein samples were resolved by electrophoresis on 8 – 12% SDS-polyacrylamide gels and transferred onto nitrocellulose membranes. Specific *Drosophila* proteins were immunodetected using the following antibodies: anti-CNK (1:5; Douziech et al., 2003); anti-RAS (1:10; Douziech et al., 2003); anti-RAF (1:5000; Douziech et al., 2003); anti-MEK (1:1000; Cell Signaling #9122); anti-MAPK (1:2000; Chemicon #AB3053); anti-JNK (1:2000; Santa Cruz #SC-571); anti-

AKT (1:2000; Cell Signaling #9272); anti-Actin (1:2000; Chemicon #MAB1501); anti-PTP-ER (1:10) is a mouse monoclonal antibody that was kindly provided by G.M. Rubin.

For STRIPAK protein expression, S2 cells were transfected with plasmids using Effectene reagent (QIAGEN)., $CuSO_4$ (0.7 mM) was added to the media 24 hours after transfection To induce protein expression. Cells were harvested 36 hours post-induction and lysed as indicated above. For immunoprecipitation of epitope-tagged proteins, lysates were incubated with either 0.1 mL anti-HA 12CA5 hybridoma supernatant or 1.5 μg goat polyclonal anti-HSV (Abcam #ab19354) for 4 hours at 4°C in the presence of 20 μL Protein A/G PLUS-Agarose (Santa Cruz). The immunoprecipitates were collected by centrifugation (1000 × g for 2 minutes at 4°C), washed three times with 1 mL cold lysis buffer, and solubilized in Laemmli Buffer. Immunoprecipitates and cell lysates (150 μg total protein) were fractionated on 8% SDS-PAGE gels and transferred to nitrocellulose membranes. The membranes were immunoblotted with anti-STag (GenScript #A00625) at 1:5000 in TBST/5% milk, anti-HSV (GenScript #A00624) at 1:5000 in TBST/5% milk, anti-GFP (Santa Cruz #sc-8334) at 1:1000 in TBST/5% milk, or anti-HA 12CA5 hybridoma supernatant at 1:10 in TBST/5% milk. Antigens were revealed by ECL with HRP-conjugated secondary antibodies.

3.11 Supplementary references

1. Ashton-Beaucage D, Udell CM, Lavoie H, Baril C, Lefrancois M, et al. (2010) The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in *Drosophila*. Cell 143: 251-262.
2. Roignant JY, Treisman JE (2010) Exon junction complex subunits are required to splice *Drosophila* MAP kinase, a large heterochromatic gene. Cell 143: 238-250.
3. Barbosa I, Haque N, Fiorini F, Barrandon C, Tomasetto C, et al. (2012) Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. Nat Struct Mol Biol 19: 983-990.
4. Alexandrov A, Colognori D, Shu MD, Steitz JA (2012) Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. Proc Natl Acad Sci U S A 109: 21313-21318.
5. Steckelberg AL, Boehm V, Gromadzka AM, Gehring NH (2012) CWC22 connects pre-mRNA splicing and exon junction complex assembly. Cell Rep 2: 454-461.
6. Singh G, Kucukural A, Cenik C, Leszyk JD, Shaffer SA, et al. (2012) The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell 151: 750-764.
7. Coelho CM, Kolevski B, Walker CD, Lavagi I, Shaw T, et al. (2005) A genetic screen for dominant modifiers of a small-wing phenotype in *Drosophila* melanogaster identifies proteins involved in splicing and translation. Genetics 171: 597-614.
8. Goshima G, Wollman R, Goodwin SS, Zhang N, Scholey JM, et al. (2007) Genes required for mitotic spindle assembly in *Drosophila* S2 cells. Science 316: 417-421.
9. Zhang JH, Chung TD, Oldenburg KR (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen 4: 67-73.
10. Marr MT, 2nd, Isogai Y, Wright KJ, Tjian R (2006) Coactivator cross-talk specifies transcriptional output. Genes Dev 20: 1458-1469.
11. Horn T, Boutros M (2010) E-RNAi: a web application for the multi-species design of RNAi reagents--2010 update. Nucleic Acids Res 38: W332-339.

12. Herold A, Klymenko T, Izaurralde E (2001) NXF1/p15 heterodimers are essential for mRNA nuclear export in *Drosophila*. RNA 7: 1768-1780.

13. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, et al. (2006) TM4 microarray software suite. Methods Enzymol 411: 134-193.

14. Murali T, Pacifico S, Yu J, Guest S, Roberts GG, 3rd, et al. (2011) DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for *Drosophila*. Nucleic Acids Res 39: D736-743.

15. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27: 431-432.

16. Xu T, Rubin GM (1993) Analysis of genetic mosaics in developing and adult *Drosophila* tissues. Development 117: 1223-1237.

17. Baril C, Therrien M (2006) Alphabet, a Ser/Thr phosphatase of the protein phosphatase 2C family, negatively regulates RAS/MAPK signaling in *Drosophila*. Dev Biol 294: 232-245.

18. Asha H, Nagy I, Kovacs G, Stetson D, Ando I, et al. (2003) Analysis of Ras-induced overproliferation in *Drosophila* hemocytes. Genetics 163: 203-215.

19. Therrien M, Wong AM, Kwan E, Rubin GM (1999) Functional analysis of CNK in RAS signaling. Proc Natl Acad Sci U S A 96: 13259-13263.

20. Douziech M, Sahmi M, Laberge G, Therrien M (2006) A KSR/CNK complex mediated by HYP, a novel SAM domain-containing protein, regulates RAS-dependent RAF activation in *Drosophila*. Genes Dev 20: 807-819.

21. Douziech M, Roy F, Laberge G, Lefrancois M, Armengod AV, et al. (2003) Bimodal regulation of RAF by CNK in *Drosophila*. Embo J 22: 5068-5078.

22. Therrien M, Wong AM, Rubin GM (1998) CNK, a RAF-binding multidomain protein required for RAS signaling. Cell 95: 343-353.

23. Baril C, Sahmi M, Ashton-Beaucage D, Stronach B, Therrien M (2009) The PP2C Alphabet Is a Negative Regulator of Stress-Activated Protein Kinase Signaling in *Drosophila*. Genetics 181: 567-579.

4 Post Translational Control of MAPK Expression by the Deubiquitinase USP47 and N-end Rule Ubiquitin Ligases

Dariel Ashton-Beaucage[1], Caroline Lemieux[1], Chris Udell[1], Malha Sahmi[1] and Marc Therrien[1,3,*]

[1] Institute for Research in Immunology and Cancer
Laboratory of Intracellular Signaling
Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montreal, Quebec, Canada, H3C 3J7

[3] Département de pathologie et de biologie cellulaire, Université de Montréal

***To whom correspondence should be addressed**
tel.: (514) 343-7837
fax.: (514) 343-6843

Contributions détaillées de l'auteur

Contributions expérimentales : Identification initiale de *Ubp64E* en tant que régulateur agissant en aval de Ras^{V12}. Production de l'ensemble des données de génétique (Fig.1C,D et Fig.S1). Démonstration de l'impact de *Usp47* sur MAPK *in vivo* (Fig.2B,C) et de l'absence d'impact sur le transcrit *mapk* par qPCR (Fig.2E). Conception et élaboration du criblage d'intéraction génétique par ARNi (Fig.3C-F, S3 et S4) et du mini-crible DUB (Fig.S5). Conception de la méthode d'analyse permettant d'évaluer l'interaction génétique entre deux ARNdb. Identification de *UbcD6*, *poe* et *CG5604* comme facteurs opposant l'effet de *Ubp64E* en culture cellulaire et *in-vivo* (Fig.4A-B).

Contributions au manuscrit : Élaboration et rédaction du manuscrit. Montage des figures.

Contributions détaillées des co-auteurs

UCM a fourni la première observation de l'impact de *Ubp64E* sur MAPK (Fig.2A; version finale produite par LC). UCM a produit les données de RT-PCR (Fig.2D) et de chargement des polysomes (Fig.2F). UCM a fourni la première observation que *Uba1* et le protéasome contrecarraient *Ubp64E* (Fig.3A,B), ce qui a servi de base au développement de l'essai de criblage par interaction génétique. LC a produit les données de la figure 2G-I et de la figure S2. LC et UCM ont confirmé les effets de « rescue » des E2/3 en Western (Fig.4C-D). TM a contribué de manière importante à l'élaboration et à la supervision des travaux de recherche.

4.1 Summary

RAS/MAPK signaling is a central component of the cell's proliferation signaling apparatus. Disruptions in this pathway's activity are widely found in cancer and other pathologies. Consequently, much effort has been devoted to understanding the mechanistic aspects of signal transmission and regulation. Importantly, while much attention has been devoted to characterizing the steps leading up to the activation/inactivation of pathway components, comparatively little is known on the regulation of component expression and turnover. Here, we describe the regulation of MAPK protein levels by the *Usp47* deubiquitinase. We also detail an RNAi based genetic-interaction screening strategy that allowed us to identify ubiquitin ligases that counteract *Usp47*'s effect on MAPK.

4.2 Introduction

The RAS/MAPK pathway is one of the principal conduits of proliferation and differentiation signaling in metazoans and perturbations in pathway activity are closely associated to oncogenesis (Schubbert, Shannon et al. 2007; Zebisch, Czernilofsky et al. 2007). Signaling through the small GTPase RAS is typically initiated by signal-receiving transmembrane receptors. Active RAS triggers the successive phosphorylation and activation of the three pathway kinases: RAF, MEK and ERK/MAPK. A complex and intricate network of factors are now known to regulate the steps leading up to MAPK activation (Kolch 2005; McKay and Morrison 2007). RAF activation, in particular, involves numerous post-translational regulatory events leading up to its activation (Udell, Rajakulendran et al. 2011). To date, much effort has been devoted to identifying and understanding the post-translational mechanisms that govern pathway signaling dynamics (Malumbres and Barbacid 2003; Kolch 2005; McKay and Morrison 2007; Udell, Rajakulendran et al. 2011). By comparison very little attention has been allocated to the study of the mechanisms that regulate RAS/MAPK pathway component expression and turnover.

Recently, we reported the identification of a series of factors in an RNAi screen that impacted RAS/MAPK signaling at a pre-translational level (Ashton-Beaucage, Udell et al. 2010; Ashton-Beaucage, Udell et al. 2014). Unexpectedly, these factors were in large part associated to either transcription or splicing of *mapk*, suggesting that this might be another focal point for regulatory input into this pathway. Here, we focus on *Usp47* (also commonly referred to as *Ubp64E*), a factor initially identified in the same RNAi screen that differs from the other candidates in that it is associated to post-translational control of MAPK expression.

Usp47 is a deubiquitinase (DUB) of the ubiquitin specific protease (USP) family that has previously been associated to the regulation of the transcriptional factors *tramtrack* and *Slbo* in *Drosophila* (Rorth, Szabo et al. 2000; Bajpe, van der Knaap et al. 2008). Its human ortholog, USP47, has been found to interact with the beta-TRCP E3 ligase complex (Peschiaroli, Skaar et al. 2010) and has been found to regulate base-excision repair by controlling the levels of BER Polymerase β (Parsons, Dianova et al. 2011; Weinstock, Wu et al. 2012). USP47 has also been linked to axonal growth by working antagonistically to the E3

ligase CHIP in regulating the microtubule severing protein katanin-p60 (Yang, Oh et al. 2013). Additionally, it has been linked to cell-cell adhesion through its ability to inhibit E-cadherin cleavage, which stabilizes adherens junction formation (Sako-Kubota, Tanaka et al. 2014).

In this study, we show that *Usp47* acts to stabilize MAPK protein levels. In order to identify the putative ubiquitin-ligases involved in destabilizing MAPK, we conducted a targeted RNAi-based genetic interaction screen. This approach led to the identification of three factors that acted antagonistically to *Usp47* in controlling MAPK levels. Remarkably, the three ubiquitin ligases are linked to a protein degradation mechanism called N-end rule. This finding has interesting implications as it suggests that MAPK signaling is regulated by an N-end rule mechanism. Moreover, it suggests a new role for the *Usp47* DUB in the N-end rule process. Taken together, these findings could have important repercussions in both the RAS signaling and N-end rule fields.

4.3 Results

We initially identified *Usp47* in a genome-wide RNAi screen for factors modulating RAS^{V12}-induced pMAPK activity in *Drosophila* S2 cells (Ashton-Beaucage, Udell et al. 2014). Like most of the factors identified in this screen, *Usp47* was found to act downstream of MEK. Indeed, signaling by constitutively active RAS, RAF and MEK were all equally suppressed by a dsRNA targeting *Usp47* (Figure 4.1A). *Usp47* also demonstrated a broad capacity to modulate MAPK signaling in different contexts; signal induced by the insulin-like receptor (InR), sevenless and EGFR was suppressed to the same degree by *Usp47* RNAi (Figure 4.1B). JNK activity induced by $RAC1^{V12}$ was, however, not modified by Usp47 (Figure 4.1B).

We next conducted genetic interaction experiments to validate Usp47's newfound function *in vivo*. In flies, RAS/MAPK signaling is important in many aspects of development, including photoreceptor and cone cell differentiation in the eye. Genetic lesions that lower pathway activity cause a readily observable eye roughness phenotype in adults that is due to missing photoreceptor cells (Fortini, Simon et al. 1992) and (Figure 4.1C). Two alleles of *Usp47* that cause a reduction in USP47 protein expression (Bajpe, van der Knaap et al. 2008) were found to increase the severity of the rough eye phenotype induced by a hypomorphic allele of *mapk* called rl^{1} (Figure 4.1C). Similar observations were also made in a $shp\text{-}2/csw^{lf}$ hemyzygous and KSR dominant negative backgrounds (Figure 4.1D and Figure S4.1A). These experiments support our cell culture-based observations pointing towards a positive regulatory role for *Usp47* in MAPK signaling.

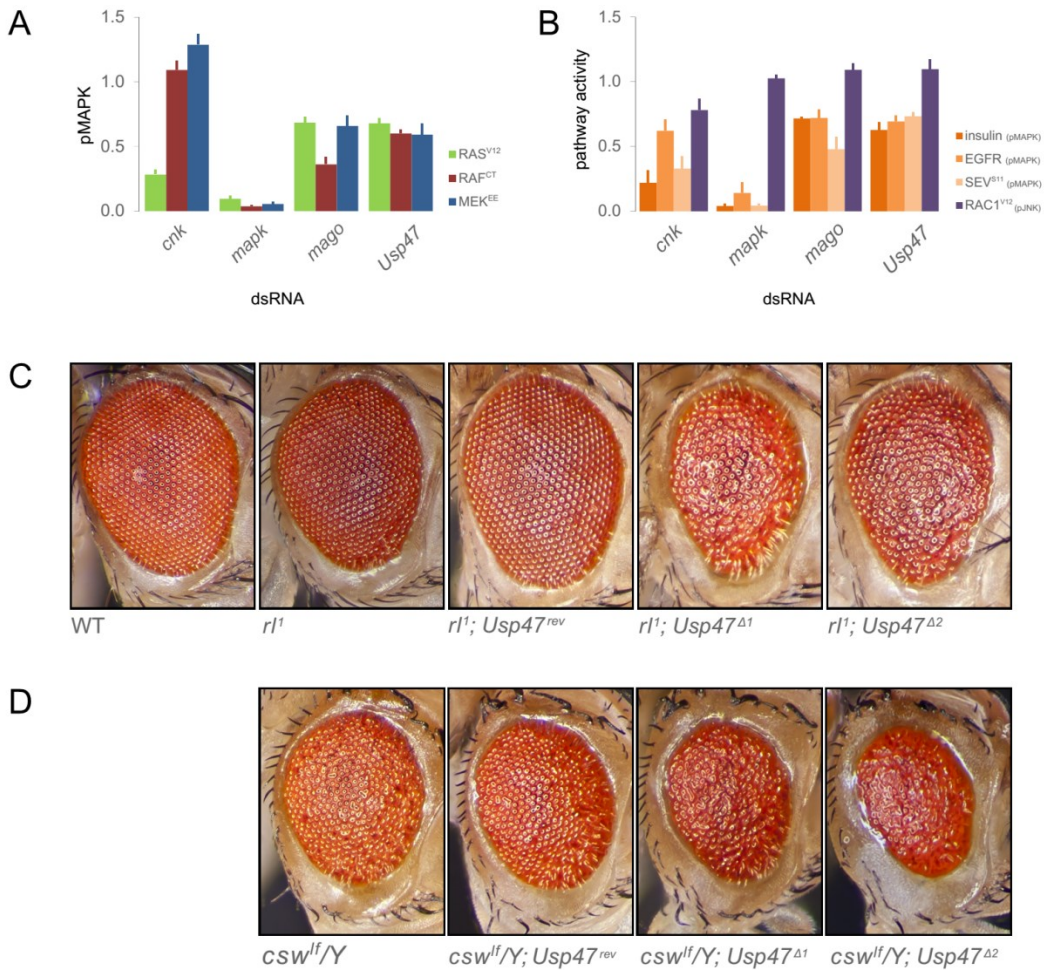


Figure 4.1 Usp47 RNAi impacts RTK/RAS signaling downstream of MEK

(A) Epistasis analysis of *Usp47* depletion on MAPK activation induced by constitutively active forms of RAS, RAF and MEK in *Drosophila* S2 cells. dsRNA targeting *cnk* was used as a control for a factor acting at the level of RAF. The EJC component, *mago*, was used as a control for a factor acting downstream of MEK. Phosphorylated MAPK levels were measured using quantitative microscopy. (B) MAPK activation is induced by three different RTKs acting upstream: insulin stimulation, EGFR expressing cells stimulated with the Spitz ligand and a heat-shock inducible sevenless (sev). JNK pathway activation induced by $RAC1^{V12}$ is used as a negative control. (C) *Usp47* alleles enhance the severity of the rough eye phenotype of mapk/rl^1 homozygous flies. However, a precise p-element excision revertant control line ($Usp47^{rev}$) did not increase eye roughness. (D) *Usp47* alleles also enhance eye roughness in *csw* loss of function (csw^{lf}) hemizygotes.

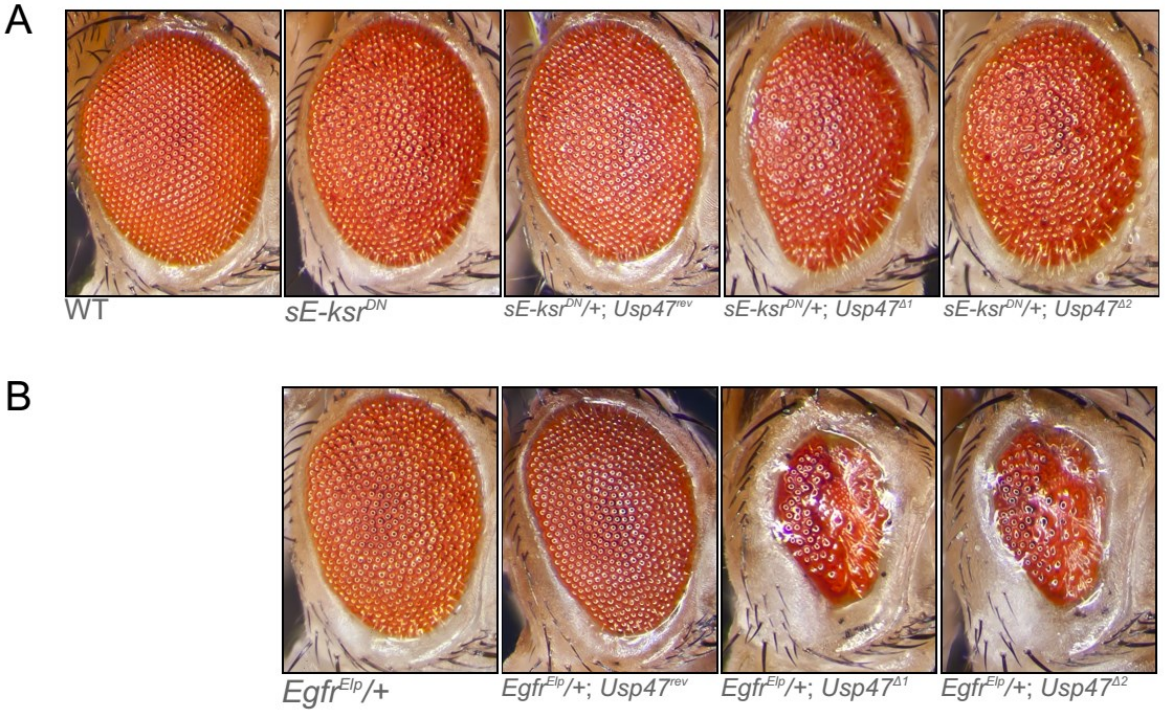


Figure S4.1 CKDN and EGFR genetic interaction experiments

(A) *Usp47* alleles enhance the severity of the rough eye phenotype in flies expressing a dominant-negative form of *ksr* (ksr^{DN}). This result is in agreement with our *csw* and *mapk* genetic interaction experiments. (B) In contrast, *Usp47* does not seem to suppress the rough eye phenotype induced by an activated form of *Egfr* gain of function ($Egfr^{Elp}$). Rather, eye size is strongly reduced and there appears to be ommatidial fusion suggesting that *Usp47* strongly enhances the phenotype. This would be consistent with *Usp47*'s previously identified function where it was shown to act on TTK, a transcriptional repressor functioning downstream of MAPK.

Like many of the candidates identified in our initial genome-wide screen, *Usp47* RNAi was found to cause a decrease in MAPK protein levels (Figure 4.2A) and (Ashton-Beaucage, Udell et al. 2014). This effect was also observed *in vivo* where knocking down *Usp47* in fly larval imaginal discs caused a drop in MAPK protein levels (Figure 4.2B and C). A clear distinction between *Usp47* and other candidates from our initial screen arose when examining their impact on *mapk* mRNA. Indeed, while the other factors caused changes in *mapk* mRNA levels or splicing, *Usp47* stood out from the lot as it had no impact on *mapk* mRNA (Figure 4.2D and E). It did not either cause a change in *mapk* translation as measured by polysome loading (Figure 4.2F) and radioactive isotope labelling assays (Figure 4.2G) and (Figure S4.2). Conversely, radioactive isotope pulse-chase experiments did reveal a more rapid decay of MAPK in cells treated with *Usp47* RNAi (Figure 4.2H and I). Together, these experiments strongly suggest that USP47 functions post-translationally to regulate MAPK protein levels.

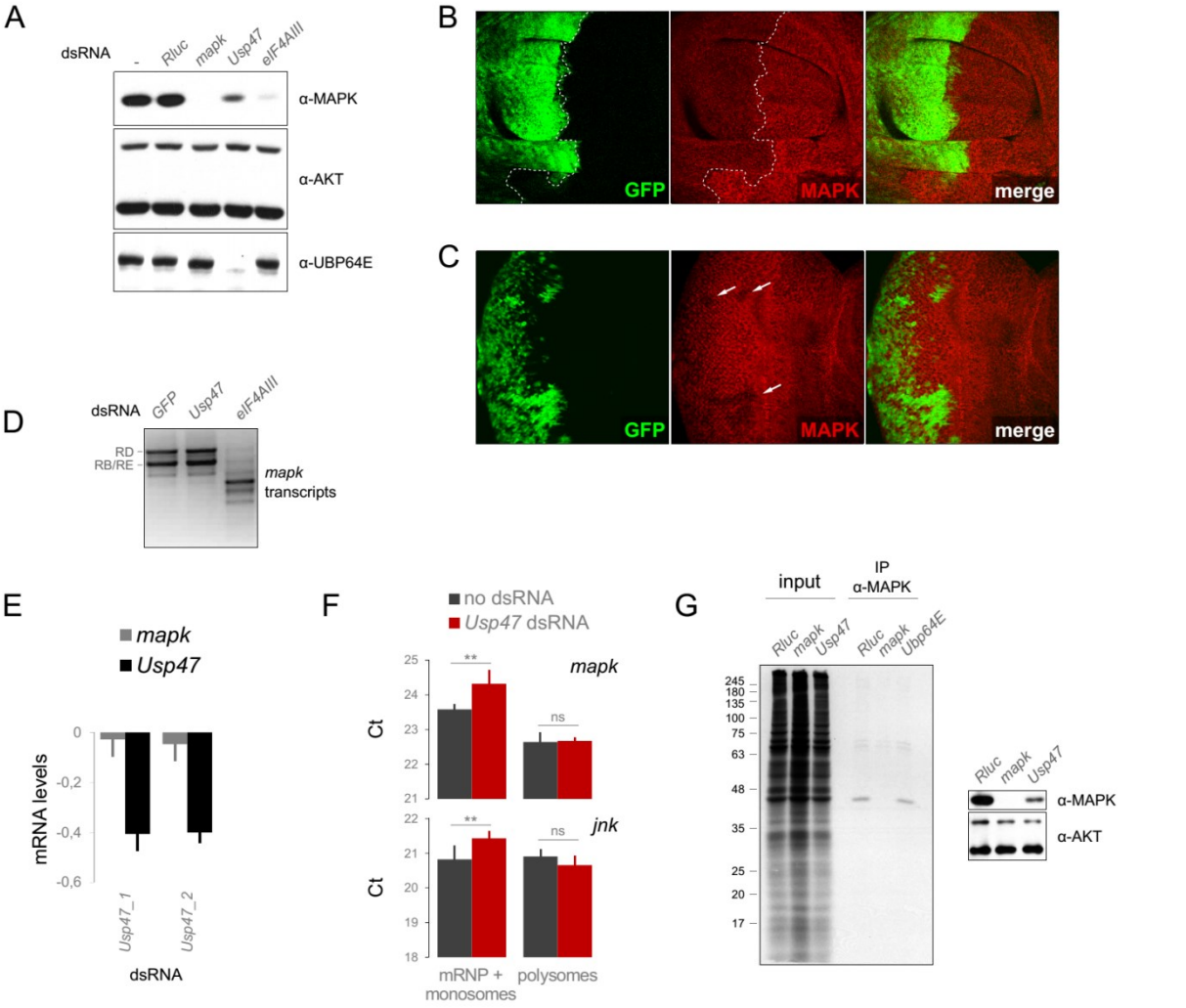
As MAPK has no known associated ubiquitin ligases and as no such factor had been identified any in our genome-wide RNAi screen, we next sought to examine if USP47 activity could be counteracted by the ubiquitin-proteasome system. UBA1 is the sole ubiquitin-activating enzyme in *Drosophila*. We found that co-depletion of *Uba1* and *Usp47* by RNAi restored MAPK levels to near baseline (Figure 4.3A). Proteasome inhibition using epoxomicin had a similar effect (Figure 4.3B). Thus, as might be expected, the ubiquitin-proteasome system appears to act in opposition to USP47.

Following this observation, we next set out to identify potential specific regulators involved in counteracting USP47 activity, including presumptive E2 and E3 ligases. A few previous studies have shown that RNAi co-depletion can be used to effectively identify genetic interactions between related factors (Axelsson, Sandmann et al. 2011; Horn, Sandmann et al. 2011; Laufer, Fischer et al. 2014). In a somewhat similar manner, we sought to identify candidates that, like *Uba1*, displayed a specific genetic interaction with *Usp47*; genetic interaction being defined here as a co-depletion outcome that differs significantly from the sum of the individual depletion outcomes (Mani, St Onge et al. 2008; Horn, Sandmann et al. 2011). Using a quantitative immunofluorescence assay, it was possible to accurately quantify MAPK level variations induced by *Usp47* RNAi and the rescue effect induced by co-depletion of *Uba1* (Figure 4.3C). We employed this assay to conduct a targeted RNAi screen

with a subset of dsRNAs linked to ubiquitin-proteasome function. RNAi reagents were tested both in combination with a *GFP* control dsRNA and in co-depletion with *Usp47* dsRNA (Figure 4.3D-E, Figure S4.3A and Table S4.I) and a genetic interaction score, Δ*m*, was derived from both these values in order to select candidates from our targeted screen ((Figure 4.3F, Figure S4.3B) and (Supplemental Text)). Using *Uba1* as a reference, we selected a Δ*m* cutoff to guide our choice of candidates for follow-up validation ((Figure S4.3C-E) and (Supplemental Text)).

Figure 4.2 *Usp47* regulates MAPK levels post-translationally

(A) *Usp47* depletion causes a reduction in MAPK levels. S2 cells were treated with the indicated dsRNAs. The EJC component *eIF4AIII* was used as a control for a factor that reduces MAPK expression. AKT levels were not impacted by *Usp47* depletion. (B-C) MAPK levels are also reduced by *Usp47* knockdown in vivo. Tissue expressing *Usp47* RNAi is marked with GFP. (B) The expression of *Usp47* RNAi was induced using an engrailed-Gal4 promoter that drives expression in the posterior segment of wing discs. (C) Clones expressing *Usp47* RNAi in eye imaginal discs also show reduced MAPK levels. (D-E) *Usp47* depletion does not impact the *mapk* transcript levels. (D) *mapk* transcript levels were measured following dsRNA treatment using an RT-PCR assay targeting the 5' and 3' UTR. *eIF4AIII* dsRNA is used as a control that alters *mapk* splicing. (E) *mapk* transcript levels measured by qPCR are not impacted by two different dsRNA reagents targeting *Usp47*. (F) Polysome loading on the *mapk* transcript is not altered by *Usp47* depletion. The polysome and mRNP fractions were separated on a sucrose gradient. qPCR was used to assay *mapk* and *jnk* levels in both fractions. A slight increase (**: $p > 0.01$) was observed in the mRNP/monosome fractions. However, this observed for both transcripts and was not accompanied by a concomitant reduction in the polysome fraction. (G) S2 cells were treated with the indicated dsRNA for five days. Radio isotope S35 methionine labelling followed by immunoprecipitation of MAPK shows that *Usp47* depletion does not cause an appreciable reduction in newly-translated MAPK levels (left panel). A control showing the impact on total MAPK and AKT levels is shown as a control (right panel). (H-I) Pulse-chase labelling of MAPK shows an impact on MAPK half-life. S2 cells were treated with the indicated dsRNAs for five days followed by 6h of S35 methionine labelling. Samples were prepared at 0-36h chase times. MAPK is found to decrease more rapidly in *Usp47* depleted cells (H) and in quantification of this experiment performed in triplicate (I).


A
dsRNA
-
Rluc
mapk
Usp47
eIF4AIII
α-MAPK
α-AKT
α-UBP64E
B
GFP
MAPK
merge
C
GFP
MAPK
merge
D
dsRNA
GFP
Usp47
eIF4AIII
RD
RB/RE
mapk
transcripts
E
mapk
Usp47
mRNA levels
0
-0,2
-0,4
-0,6
Usp47_1
Usp47_2
dsRNA
F
no dsRNA
Usp47 dsRNA
mapk
jnk
Ct
**
ns
mRNP + monosomes
polysomes
G
input
IP α-MAPK
Rluc
mapk
Usp47
Ubp64E
245
180
135
100
75
63
48
35
25
20
17
α-MAPK
α-AKT
H
input
IP α-MAPK
Rluc
Usp47
dsRNA
h
0 6 12 24 36
I
Chase Time (h)
0
10
20
30
MAPK Levels (log10)
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
-1,2
Rluc dsRNA
Usp47 dsRNA
R² = 0,9614
R² = 0.8989

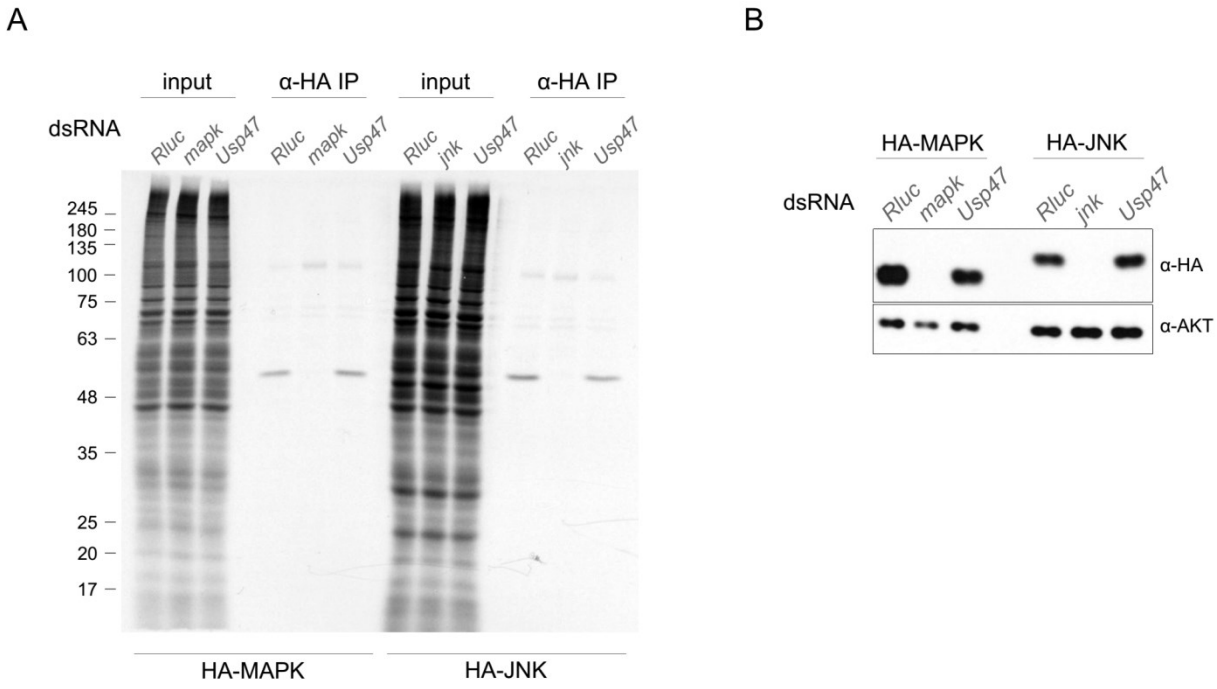


Figure S4.2 Labelling exogenous MAPK and JNK

S2 cell lines expressing HA-MAPK and HA-JNK were treated with the indicated dsRNA for five days. Radio isotope S35 methionine labelling followed by anti-HA immunoprecipitation shows that *Usp47* depletion does not cause an appreciable reduction in newly-translated exogenous MAPK or JNK levels (left panel). A control showing the impact on total MAPK and AKT levels is shown as a control (right panel).

Figure 4.3 Elaboration of an RNAi screening strategy to identify factors that modify USP47's impact on MAPK levels

(A) Knocking down the *Drosophila* E1 ligase, *Uba1*, rescues the impact of *Usp47* RNAi on MAPK levels. S2 cells were treated with the indicated dsRNAs for five days. Note that while MAPK levels are restored by *Uba1* co-depletion, the levels of a CNK control were slightly reduced. As might be expected, we observed an appreciable impact on cell viability when knocking down *Uba1*. (B) Inhibiting proteasome function by treating S2 cells with epoxomyxcin also restored MAPK levels in a *Usp47* depleted context. (C) *Uba1*'s rescue effect can be robustly measured in a plate-based quantitative microscopy assay suitable for large scale screening. *Usp47* and *Uba1* co-depletion significantly restored MAPK levels (***: $p < 0.001$) compared to a *Usp47* single depletion (**: $p < 0.01$). *Uba1* single depletion did not significantly alter MAPK levels compared to *GFP* dsRNA treated control cells (used for normalization. In this experiment, S2 cells were pre-incubated with either *GFP* dsRNA (1st and 3rd sample) or with *Usp47* dsRNA (2nd and 4th sample) for 3 days. The cells were then distributed in 384 well plates containing the indicated dsRNAs and incubated for another three days. Following dsRNA treatment, anti-MAPK stained cells were imaged and analyzed by high content microscopy. (D) Screening strategy for a large-scale RNAi screen focused on ubiquitin-proteasome associated factors. As in (C), S2 cells were pre-incubated with either *GFP* or *Usp47* dsRNA for three days and then distributed in screening plates containing the ubiquitin-proteasome dsRNA set and incubated for another three days followed by quantitative microscopy quantification of MAPK levels. Each condition was tested in quadruplicate. (E) Histogram depicting RNAi screen results. MAPK levels were normalized to GFP dsRNA treated cells. The *Usp47* co-depleted cells show a clear shift towards a reduction in MAPK levels. The *Uba1* dsRNA control stands out from the lot as it completely counteracts the *Usp47* dsRNA. (F) Distribution of *Usp47* genetic interaction scores (Δm) for the screening set. In order to distinguish specific modifiers of *Usp47* from other candidates that might simply alter MAPK levels in an unrelated process (such as transcriptional or splicing regulation), we derived a genetic interaction score to select candidates for validation. The Δm indicates the difference between the expected outcome of a co-depletion effect (neutral phenotype, $\Delta m = 0$) and outcomes that result from synergistic (positive Δm) or alleviating

(negative Δm) genetic interactions. This is further detailed in the supplemental text. The negative Δm obtained for Uba1 is consistent with the observed alleviation of the MAPK level reduction.

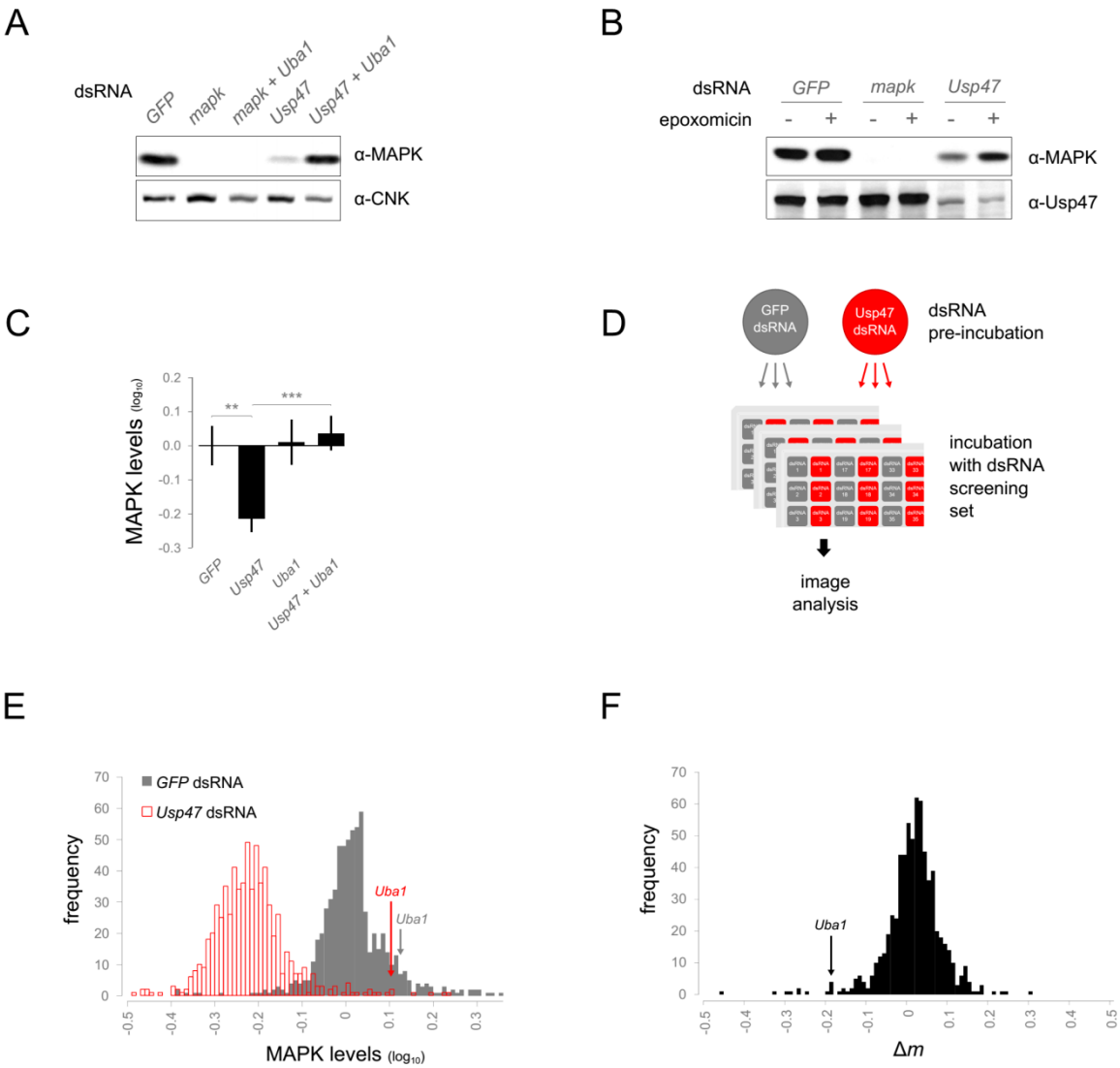

A
dsRNA
GFP
mapk
mapk + Uba1
Usp47
Usp47 + Uba1
α-MAPK
α-CNK
B
dsRNA
GFP
mapk
Usp47
epoxomicin
- + - + - +
α-MAPK
α-Usp47
C
MAPK levels (log10)
0.2
0.1
0.0
-0.1
-0.2
-0.3
**
***
GFP
Usp47
Uba1
Usp47 + Uba1
D
GFP dsRNA
Usp47 dsRNA
dsRNA pre-incubation
incubation with dsRNA screening set
image analysis
E
GFP dsRNA
Usp47 dsRNA
frequency
Uba1
Uba1
MAPK levels (log10)
F
frequency
Uba1
Δm

Figure S4.3 Additional data and hit selection strategy for the *Usp47* genetic interaction RNAi screen

(A) *Usp47* RNAi screen MAPK levels are plotted for *GFP* co-depletion samples against the corresponding *Usp47* co-depletion samples. X-axis values are normalized to *GFP* dsRNA controls while y-axis values are normalized to *Usp47* dsRNA controls. *Usp47* co-depletion MAPK levels generally mirror those of the *GFP* co-depletion controls. This suggests that most candidates do not differ from the expected neutral phenotype of a purely additive co-depletion effect. (B) Volcano plot of Δ*m* scores plotted against Δ*m* p-values. The two values are generally well correlated prompting us to use the p-value to guide candidate selection. (C) Average cells per well imaged for both *GFP* and *Usp47* co-depletion samples throughout the screen. The absence of a significant difference between the two suggests that there was not a widespread synthetic lethality effect caused by co-depletion of *Usp47* and other ubiquitin-proteasome factors. (D) A reduction in cell count is correlated with a non-neutral Δ*m* (***: $p < 0.001$). This led to the use of cell count as a secondary criteria to increase the stringency of selection criteria for hits that caused cell lethality. (E) Graphical representation of the hit selection criteria. The false discovery rate (FDR) of Δ*m* is plotted against the FDR of the cell count values. A global Δ*m* cutoff (blue line) is used to select hits irrespective of their impact on viability. A second less stringent Δ*m* cutoff (red line) is used to select candidates that had little or no impact on cell count.

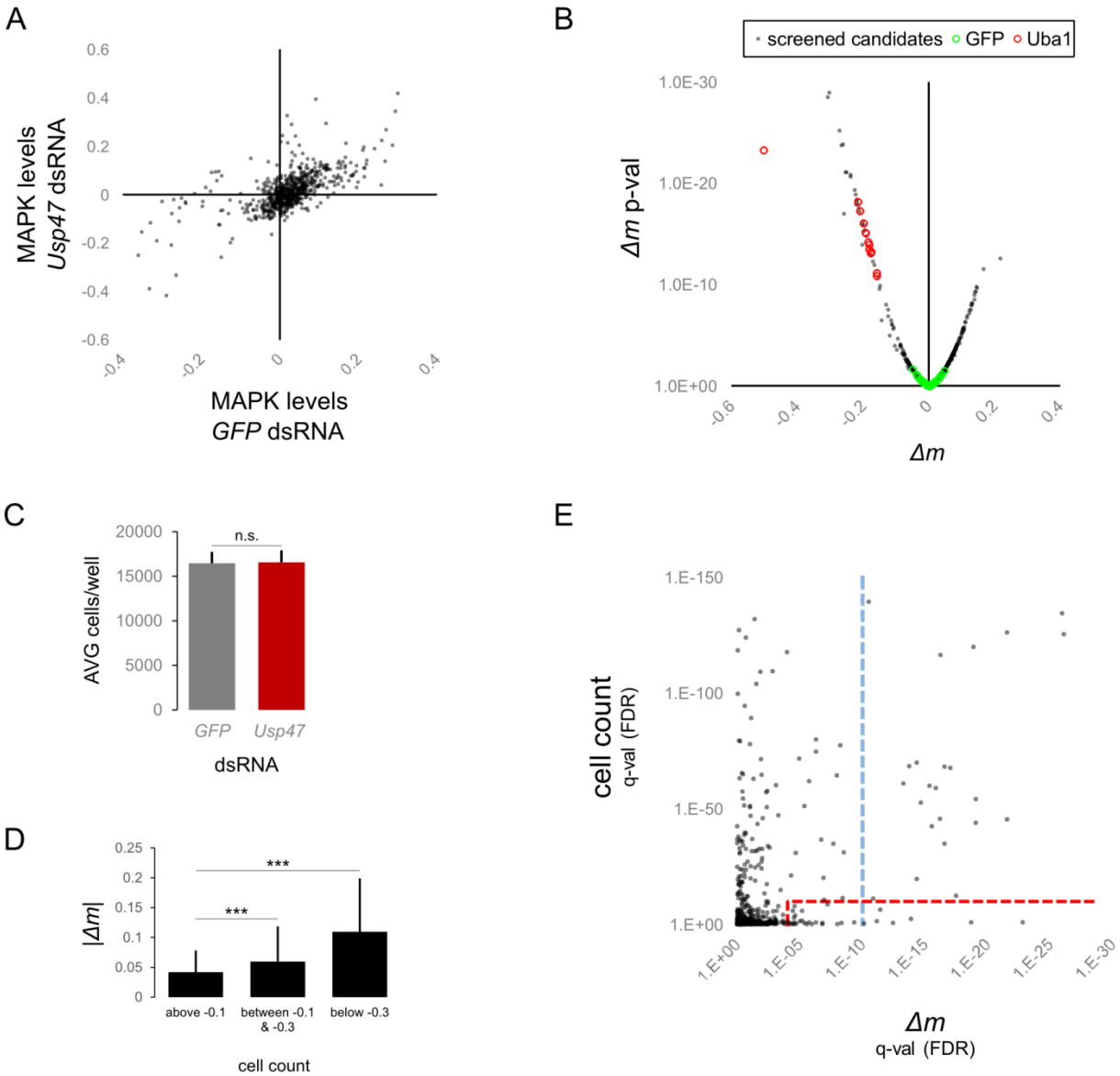

A
MAPK levels
Usp47 dsRNA
MAPK levels
GFP dsRNA
B
screened candidates
GFP
Uba1
Δm p-val
Δm
C
AVG cells/well
n.s.
GFP
Usp47
dsRNA
D
|Δm|
***
***
above -0.1
between -0.1 & -0.3
below -0.3
cell count
(log10 fold of GFP dsRNA)
E
cell count
q-val (FDR)
Δm
q-val (FDR)

Table S4.I Primary screen results (MS Excel file)

Results from the *Usp47* co-depletion RNAi screen. MAPK levels, Δ*m*, actin levels and cell count are $\log_{10}$ normalized values. *GFP* dsRNA treated controls were used for normalization of *GFP* co-depleted samples whereas *Usp47* co-treated samples were used as normalization controls for their *Usp47* co-depletion counterparts. The false discovery rate (FDR) is also included for both the Δ*m* and cell count.

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

55 genes were selected for a validation screen using independent dsRNA reagents (Table S4.II). These included a number of proteasome components and the ubiquitin genes that caused an important reduction in cell count. Importantly, *Usp47* co-depletion was not found to rescue this lethality (Figure S4.4), indicating that it could be separated from the effect on MAPK levels. This guided our choice not to exclude general proteasome components and ubiquitin genes from our validation experiments; many such factors were indeed confirmed to rescue *Usp47*'s impact on MAPK levels (Table S4.III), further indicating a functional interaction with the ubiquitin-proteasome machinery.

In our validation experiments, only candidates with negative genetic interaction scores were confirmed (alleviating genetic interactions) and these all produced *Usp47* rescue effects (Table S4.II). The absence of synthetic and *Usp47* epistatic genetic interactions could signify that this type of regulation is not operating on USP47/MAPK or, if such factors exist, that they were not present in our screening set. Still, as *Usp47* depletion had only a partial impact on MAPK levels compared to the effect of *mapk* dsRNA itself, we reasoned that it might be acting redundantly with another DUB to protect MAPK from degradation. This observation was verified for multiple *Usp47* dsRNAs, even when tested at prolonged depletion times (data not shown). In order to further investigate the possibility of a redundant DUB, we also re-screened the remaining 41 predicted DUBs in the fly genome. However, we found no other DUB that further impacted MAPK levels, either alone or when co-depleted with *Usp47* (Figure S4.5 and Table S4.IV). Taken together with our primary screen data, this suggests that *Usp47* is not acting redundantly with another DUB and that its partial impact on MAPK levels must be explained by another feature of this regulatory mechanism. For instance, the degradation machinery may not be able to access the entire pool of MAPK.

We next further narrowed down our list to 13 candidates that displayed a reproducible rescue effect that was also consistent with what we had observed in our primary screen (Table S4.III) and (Supplemental Text). Of particular interest, an E2 ubiquitin-conjugating enzyme, *UbcD6*, and two E3 ligases, *poe* and *CG5604* were retained in this set (*CG31053*, a potential E3 also present in this list had a weak rescue impact that we were not able to reproduce using other techniques). All three factors displayed a capacity to restore *Usp47* RNAi-induced MAPK levels (Figure 4.4A and Table S4.V). This effect was confirmed by western blot

(Figure 4.4B) as well as *in vivo*, in *Drosophila* wing imaginal discs (Figure 4.4C). Moreover, exogenously expressed MAPK was also found to be sensitive to these three factors (Figure 4.4D), indicating that the observed effect was not due to an unrelated transcriptional (or pre-translational) regulation of MAPK expression.

Table S4.II RNAi validation screen results (MS Excel file)

Results from the validation experiments using independent dsRNAs from those used in the primary screen. MAPK levels, Δ*m*, actin levels and cell count are $\log_{10}$ normalized values. *GFP* dsRNA treated controls were used for normalization of *GFP* co-depleted samples whereas *Usp47* co-treated samples were used as normalization controls for their *Usp47* co-depletion counterparts. In addition to the Δ*m* p-value, two other selection criteria are shown: consistency with the primary screen results and MAPK level variation above 0.1 (absolute value, $\log_{10}$ normalized). P-values shown are for a Student's t-test comparing the experimental result to those of *GFP* dsRNA control samples.

Fichier Microsoft Excel. Disponible en format électronique associé à ce document.

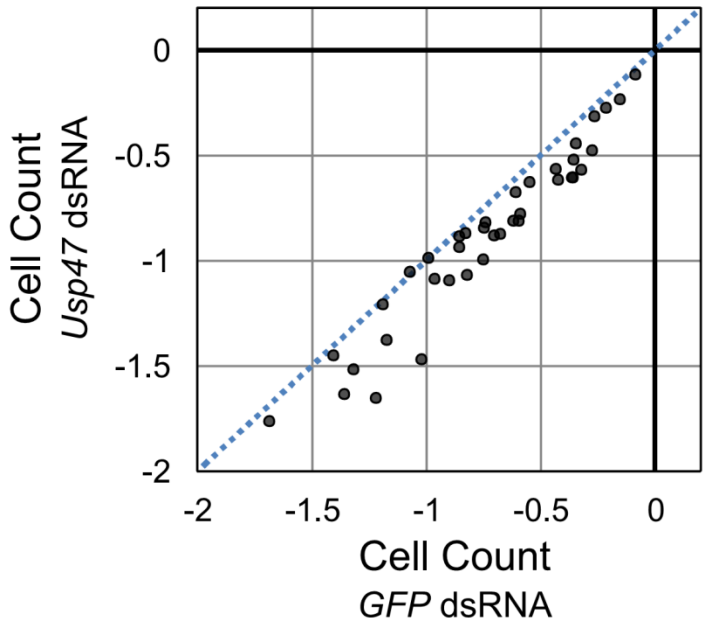


Figure S4.4 Absence of alleviating or synthetic lethality between *Usp47* and ubiquitin-proteasome factors

Data for the ubiquitin-proteasome factors from the RNAi validation experiments (using independent dsRNAs from the primary screen). Cell count values are shown for *GFP* co-depletion (x-axis) plotted against *Usp47* co-depletion (y-axis). Values are $\log_{10}$ normalized to the respective control dsRNAs (*GFP* and *Usp47*). Ubiquitin-proteasome component depletion conceivably has a broad impact on cell function causing cell lethality and the MAPK level increase may be a by-product of this non-specific effect. However, *Usp47* co-depletion did not appreciably alter the cell count reduction associated with these factors; their impact was generally equivalent to the *GFP* co-depletion condition (blue dotted line). This suggests that Usp47 is not simply countering some general function of the ubiquitin-proteasome machinery.

gene	FBgn	# dsRNA probes confirming	reduced cell count	function
CG31053	FBgn0051053	1		encodes RING domain protein
CG5604	FBgn0032208	1		UFD4 ortholog; HECT domain E3 ligase
fzy	FBgn0001086	1	+	CDC20 homolog; member of the APC complex
poe	FBgn0011230	2		UBR4 ortholog; UBR box domain E3 ligase
Prosalpha5	FBgn0016697	2	++	proteasome subunit
Prosbeta1	FBgn0010590	1	++	proteasome subunit
Prosbeta3	FBgn0026380	2	++	proteasome subunit
Rpn2	FBgn0028692	2	++	proteasome regulatory subunit
RpS27A	FBgn0003942	2	++	ubiquitin and ribosomal protein
UbcD6	FBgn0004436	1	+	RAD6 ortholog; E2 ubiquitin ligase
Ubi-p5E	FBgn0086558	2	++	ubiquitin
Ubi-p63E	FBgn0003943	2	++	ubiquitin
Ufd1-like	FBgn0036136	2	+	UFD1 ortholog; involved in endoplasmic reticulum associated degradation (ERAD)

Table S4.III Validated hits list

Summary table of validated hits. RNAi reagents causing a reduction in cell count are marked "+" (moderate) and "++" (strong).

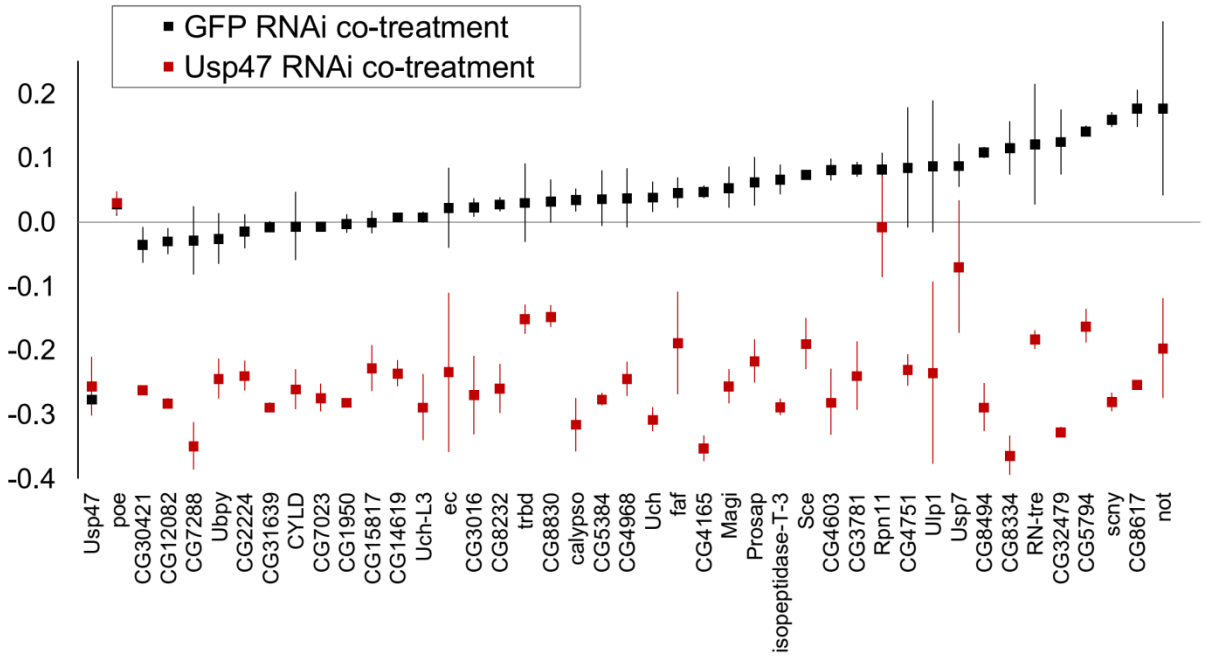


Figure S4.5 DUB redundancy screen

Results from the DUB co-depletion screen. Usp47 RNAi depletion causes a limited impact on MAPK levels, even at prolonged depletion times. An independent set of dsRNAs targeting the predicted DUBs in Drosophila was designed to specifically re-screen these genes to look for any potential redundant factors that might act synergistically with Usp47. However, as was the case in the primary screen, no other DUB was found to cause an appreciable MAPK reduction, either on its own or in conjunction with Usp47 depletion.

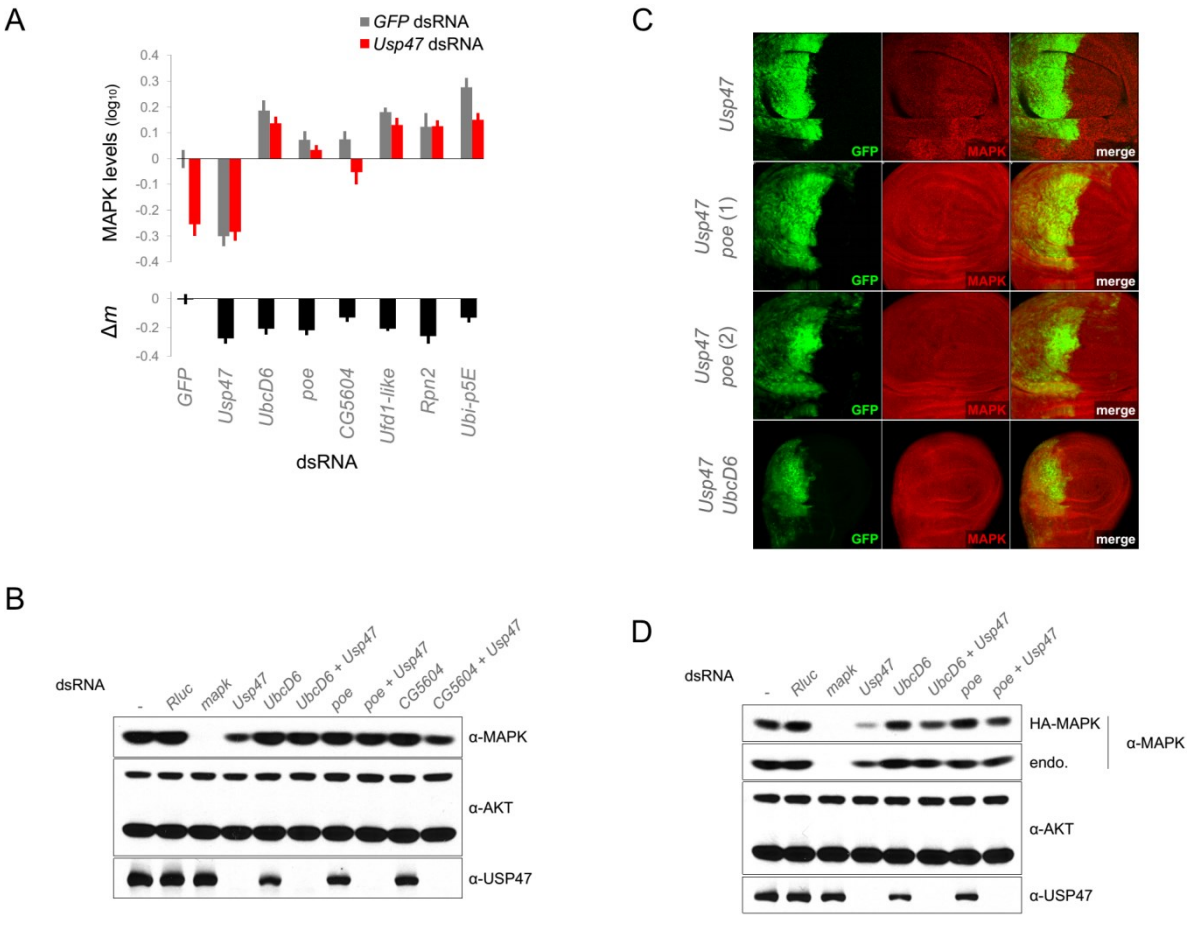
RNAi probe	gene symbol	closest Hsap ortholog	closest Scer ortholog	*GFP* dsRNA (MAPK)	*Usp47* dsRNA (MAPK)	*poe* dsRNA (MAPK)	*Usp47;poe* dsRNA (MAPK)	Δm	p-val. GFP dsRNA (MAPK)	p-val. Usp47 dsRNA (MAPK)	p-val. poe dsRNA (MAPK)	p-val. Usp47+poe dsRNA (MAPK)	p-val Δm	*GFP* dsRNA (cell count)	*Usp47* dsRNA (cell count)	*poe* dsRNA (cell count)	*Usp47;poe* dsRNA (cell count)	GFP dsRNA (actin)	Usp47 dsRNA (actin)	poe dsRNA (actin)	Usp47;poe dsRNA (actin)
Usp47_3	*Usp47*	USP47	UBP6	-0.28	-0.26	-0.05	-0.11	**-0.25**	4.1E-12	5.7E-11	6.2E-02	4.1E-05	2.0E-05	0.02	0.00	0.01	-0.03	0.00	0.02	0.02	0.01
poe_1	*poe*	UBR4	0	0.03	0.03	0.13	0.01	**-0.23**	2.1E-01	1.7E-01	7.6E-06	7.0E-01	3.3E-05	0.02	-0.04	-0.02	-0.07	0.00	0.04	0.04	0.04
CG30421	*CG30421*	USP31	UBP12	-0.04	-0.26	0.04	-0.07	**0.00**	1.4E-01	1.0E-11	5.2E-02	5.6E-03	8.8E-01	0.08	0.00	0.00	-0.09	-0.04	0.01	0.02	0.03
CG12082	*CG12082*	USP5	UBP14	-0.03	-0.28	0.01	-0.04	**0.02**	2.0E-01	2.1E-12	5.8E-01	1.1E-01	6.9E-01	0.07	0.01	0.00	-0.13	-0.03	-0.01	0.00	0.05
CG7288_3	*CG7288*	USP39	SAD1	-0.03	-0.35	-0.01	-0.05	**0.09**	2.5E-01	4.2E-14	7.3E-01	5.5E-02	9.0E-02	-0.14	-0.21	-0.12	-0.15	0.08	0.12	0.07	0.08
Ubpy	*Ubpy*	USP8	DOA4	-0.03	-0.24	0.02	-0.07	**-0.01**	2.8E-01	8.1E-11	3.4E-01	5.0E-03	7.5E-01	0.06	-0.08	-0.05	-0.05	-0.02	0.05	0.03	0.03
CG2224	*CG2224*	STAMBP	RFU1	-0.02	-0.24	0.04	-0.07	**-0.01**	5.4E-01	8.9E-11	5.7E-02	7.9E-03	8.6E-01	0.03	0.01	-0.01	-0.10	-0.03	0.01	0.02	0.04
CG31639	*CG31639*	ODF1	0	-0.01	-0.29	0.03	-0.06	**0.05**	7.6E-01	1.2E-12	1.8E-01	1.5E-02	3.1E-01	0.06	0.02	0.03	-0.04	-0.03	0.00	0.00	0.02
CYLD	*CYLD*	CYLD	0	-0.01	-0.26	0.06	-0.03	**0.02**	8.0E-01	2.0E-11	1.3E-02	1.7E-01	6.7E-01	0.04	-0.05	-0.04	-0.15	-0.01	0.02	0.03	0.05
CG7023	*CG7023*	USP46	UBP13	-0.01	-0.27	0.06	-0.02	**0.04**	8.1E-01	5.0E-12	9.8E-03	3.1E-01	4.7E-01	0.05	0.02	-0.03	-0.12	-0.02	0.01	0.02	0.06
CG1950	*CG1950*	UCHL5	0	0.00	-0.28	0.03	-0.05	**0.05**	9.3E-01	2.2E-12	1.6E-01	4.3E-02	3.4E-01	0.04	0.02	0.03	-0.10	-0.01	0.00	0.00	0.05
CG15817	*CG15817*	USP1	UBP13	0.00	-0.23	0.00	-0.05	**0.00**	1.0E+00	3.8E-10	8.9E-01	6.5E-02	9.0E-01	0.06	0.02	0.04	-0.04	-0.01	0.00	-0.01	0.03
CG14619	*CG14619*	USP2	DOA4	0.01	-0.24	0.01	-0.08	**0.01**	7.3E-01	1.1E-10	7.6E-01	2.1E-03	8.3E-01	0.07	0.07	0.06	-0.02	-0.03	-0.02	-0.01	0.02
Uch-L3	*Uch-L3*	UCHL5	0	0.01	-0.29	0.05	-0.03	**0.07**	7.0E-01	5.9E-12	3.1E-02	2.3E-01	1.9E-01	0.06	0.00	0.02	-0.04	-0.02	0.01	0.00	0.03
ec	*ec*	USP53	0	0.02	-0.24	0.00	-0.06	**0.03**	3.5E-01	6.3E-08	9.9E-01	2.2E-02	6.7E-01	0.02	0.01	0.03	-0.02	0.00	0.01	-0.01	0.02
CG3016	*CG3016*	USP30	UBP1	0.02	-0.27	0.04	-0.03	**0.06**	3.0E-01	4.8E-11	1.2E-01	1.8E-01	2.3E-01	0.01	0.03	0.00	-0.07	0.00	0.00	0.02	0.04
CG8232	*CG8232*	PAN2	PAN2	0.03	-0.26	0.03	-0.03	**0.06**	2.1E-01	3.0E-11	1.7E-01	1.5E-01	2.6E-01	0.04	0.02	0.01	-0.08	-0.01	0.00	0.01	0.05
trbd	*trbd*	ZRANB1	0	0.03	-0.15	0.08	0.00	**-0.05**	2.2E-01	5.1E-07	1.3E-03	9.6E-01	2.9E-01	0.02	0.00	-0.05	-0.14	0.01	0.02	0.05	0.08
CG8830	*CG8830*	USP35	UBP11	0.03	-0.15	0.04	-0.03	**-0.05**	1.6E-01	6.9E-07	6.2E-02	2.6E-01	2.6E-01	0.03	0.01	0.01	-0.10	0.00	-0.01	0.00	0.05
calypso	*calypso*	0	0	0.03	-0.32	-0.03	-0.06	**0.12**	1.3E-01	5.0E-13	2.7E-01	1.3E-02	2.1E-02	0.02	0.04	0.03	-0.08	0.00	-0.01	-0.01	0.05
CG5384	*CG5384*	USP14	UBP6	0.04	-0.28	-0.01	-0.02	**0.08**	1.2E-01	3.4E-12	5.7E-01	3.5E-01	1.0E-01	0.02	0.03	0.00	-0.10	0.01	0.00	-0.01	0.07
CG4968	*CG4968*	OTUB1	0	0.04	-0.24	0.04	-0.04	**0.05**	1.2E-01	6.9E-11	1.0E-01	8.1E-02	3.3E-01	0.02	0.03	0.03	-0.10	-0.01	0.00	0.00	0.04
Uch	*Uch*	UCHL3	YUH1	0.04	-0.31	0.01	-0.06	**0.12**	9.0E-02	3.6E-13	5.3E-01	2.2E-02	2.2E-02	0.03	0.02	0.02	-0.05	-0.01	0.00	0.00	0.03
faf	*faf*	USP9X	UBP15	0.04	-0.19	0.01	-0.07	**0.00**	4.9E-02	1.4E-07	5.9E-01	1.1E-02	9.7E-01	0.01	-0.02	0.01	-0.05	0.01	0.03	0.00	0.03
CG4165	*CG4165*	USP45	UBP8	0.05	-0.35	-0.02	-0.05	**0.17**	4.0E-02	1.8E-14	4.4E-01	5.8E-02	1.5E-03	-0.01	0.03	0.03	-0.09	0.02	-0.01	-0.02	0.06
Magi	*Magi*	MAGI1	0	0.05	-0.26	0.02	-0.01	**0.08**	2.4E-02	2.6E-11	3.9E-01	6.8E-01	1.2E-01	0.00	0.01	0.00	-0.15	0.01	0.01	0.01	0.07
Prosap	*Prosap*	SHANK1	MGA2	0.06	-0.22	0.02	-0.04	**0.05**	1.0E-02	1.0E-09	3.9E-01	1.4E-01	3.4E-01	0.03	0.01	0.01	-0.10	0.00	0.01	0.00	0.04
isopeptidase-T-3	*isopeptidase-T-3*	UBAC1	0	0.07	-0.29	0.01	0.00	**0.12**	6.5E-03	1.4E-12	5.8E-01	9.0E-01	1.5E-02	-0.05	0.02	0.01	-0.17	0.04	0.00	0.00	0.09
Sce	*Sce*	RNF2	0	0.07	-0.19	0.01	-0.02	**0.03**	2.4E-03	1.6E-08	5.0E-01	5.5E-01	5.3E-01	0.02	0.02	-0.01	-0.14	0.00	0.01	0.01	0.07
CG4603	*CG4603*	YOD1	OTU1	0.08	-0.28	0.06	0.00	**0.13**	1.1E-03	1.1E-11	1.6E-02	9.1E-01	1.2E-02	-0.06	0.01	0.02	-0.20	0.05	0.00	0.00	0.10
CG3781	*CG3781*	JOSD2	0	0.08	-0.24	0.06	-0.01	**0.09**	9.6E-04	3.2E-10	1.2E-02	6.7E-01	7.2E-02	0.03	0.03	0.01	-0.13	0.00	0.00	0.01	0.07
Rpn11	*Rpn11*	PSMD14	RPN11	0.08	-0.01	0.11	0.12	**-0.14**	1.0E-03	7.9E-01	2.2E-05	2.5E-05	7.0E-03	-1.17	-1.20	-0.82	-1.00	0.05	0.06	0.14	0.14
CG4751	*CG4751*	MPND	0	0.08	-0.23	0.03	-0.03	**0.08**	4.7E-03	2.2E-10	2.8E-01	2.5E-01	1.2E-01	-0.02	0.01	0.02	-0.19	0.03	0.00	0.01	0.08
Ulp1	*Ulp1*	SENP2	ULP1	0.09	-0.24	-0.06	-0.07	**0.13**	4.9E-03	4.9E-07	1.4E-02	5.9E-03	8.0E-02	-0.26	-0.06	0.03	-0.15	0.10	0.03	-0.02	0.07
Usp7	*Usp7*	USP7	UBP15	0.09	-0.07	0.11	0.03	**-0.07**	6.5E-04	2.2E-02	6.4E-05	2.4E-01	1.6E-01	-0.09	-0.09	-0.04	-0.17	0.07	0.06	0.03	0.07
CG8494	*CG8494*	USP20	UBP10	0.11	-0.29	0.00	0.03	**0.17**	4.5E-05	2.9E-12	9.2E-01	1.6E-01	1.7E-03	0.00	0.02	0.02	-0.13	0.03	0.01	0.00	0.08
CG8334	*CG8334*	USP32	UBP12	0.11	-0.36	-0.03	0.03	**0.25**	3.5E-05	1.3E-14	2.7E-01	2.5E-01	2.2E-05	0.00	0.03	0.05	-0.13	0.02	0.00	-0.02	0.08
RN-tre	*RN-tre*	USP6NL	MSB3	0.12	-0.18	0.05	0.04	**0.07**	1.6E-04	1.5E-08	3.8E-02	1.2E-01	1.7E-01	-0.04	-0.04	0.00	-0.16	0.04	0.04	0.01	0.07
CG32479_2	*CG32479*	USP10	UBP3	0.12	-0.33	-0.03	-0.02	**0.22**	1.7E-05	7.8E-14	2.2E-01	3.4E-01	1.0E-04	-0.11	-0.04	-0.04	-0.20	0.07	0.01	0.00	0.09
CG5794	*CG5794*	USP34	UBP15	0.14	-0.16	0.11	0.03	**0.07**	9.8E-07	1.7E-07	2.0E-05	2.2E-01	1.5E-01	-0.04	-0.06	-0.07	-0.13	0.05	0.06	0.06	0.08
scny	*scny*	USP36	UBP10	0.16	-0.28	0.04	0.05	**0.21**	1.4E-07	2.5E-12	8.4E-02	2.3E-02	1.5E-04	-0.28	-0.10	-0.06	-0.32	0.13	0.03	0.02	0.14
CG8617	*CG8617*	STAMBP	0	0.18	-0.25	0.05	0.06	**0.20**	3.4E-08	2.1E-11	4.8E-02	1.6E-02	3.0E-04	-0.14	-0.11	-0.05	-0.25	0.10	0.04	0.02	0.14
not	*not*	USP22	UBP8	0.18	-0.20	0.01	0.02	**0.14**	9.9E-06	6.4E-08	6.7E-01	3.9E-01	2.3E-02	-0.11	0.00	0.02	-0.14	0.06	0.03	0.01	0.10

Table S4.IV DUB screen results

Results from the DUB RNAi mini-screen using independent dsRNAs from those used in the primary screen. MAPK levels, Δm, actin levels and cell count are $\log_{10}$ normalized values. All values shown here are normalized to *GFP* dsRNA treated controls. In addition to co-treatment with *Usp47*, all DUBs were also tested in co-depletion with *poe* and in triple-depletion with *Usp47* and *poe*. P-values shown are for a Student's t-test comparing the experimental result to those of *GFP* dsRNA control samples.

Figure 4.4 *UbcD6*, *poe*/*UBR4* and *CG5604*/*UFD4* alleviate *Usp47*'s impact on MAPK levels

(A) Selected hits from our *Usp47* co-depletion screen. After validation with independent dsRNA reagents, our set of candidates included one E2 conjugating enzyme, *UbcD6*, and two E3 ligases, *poe* and *CG5604*. The set also included proteasome components such as *Rpn2* and ubiquitin genes such as *Ubi-p5E*. *Ufd1-like*, a proteasome-associated factor also linked to ER-associated protein degradation (ERAD) was also present in our set. Of the dsRNA shown here, only *poe* and *CG5604* did not cause an accompanying reduction in cell count (see Table S2 and Table S3 for full validation screen data). Data shown here is from RNAi validation experiments using a second set of dsRNAs. All MAPK levels are normalized to *GFP* dsRNA controls. (B) Western blot experiments confirm that *UbcD6, poe* and *CG5604* also rescue endogenous MAPK levels in *Usp47* depleted cells. The rescue for CG5604 is weaker, possibly due to a comparatively moderate depletion efficiency (see Table S5). AKT levels are shown as a control. (C) Co-depletion experiments conducted in third instar *Drosophila* wing discs show that *poe* and *UbcD6* can rescue *Usp47* depletion in vivo. The *en-gal4* driver was used to drive expression of RNAi in the posterior segment (GFP-positive) of the disc. *UbcD6* RNAi expression caused extensive larval lethality. Those wing discs that could be recovered were of reduced size. (D) Exogenous MAPK levels are rescued by co-depletion of *Usp47* with *poe* or *UbcD6*. An *HA-mapk* stable cell line was treated with the indicated dsRNA for four days. A rescue effect was observed on both endogenous and exogenous MAPK.


A
GFP dsRNA
Usp47 dsRNA
MAPK levels (log10)
Δm
GFP
Usp47
UbcD6
poe
CG5604
Ufd1-like
Rpn2
Ubi-p5E
dsRNA
C
Usp47
Usp47 poe (1)
Usp47 poe (2)
Usp47 UbcD6
GFP
MAPK
merge
B
dsRNA
-
Rluc
mapk
Usp47
UbcD6
UbcD6 + Usp47
poe
poe + Usp47
CG5604
CG5604 + Usp47
α-MAPK
α-AKT
α-USP47
D
dsRNA
-
Rluc
mapk
Usp47
UbcD6
UbcD6 + Usp47
poe
poe + Usp47
HA-MAPK
endo.
α-MAPK
α-AKT
α-USP47

qPCR assay	*GFP* dsRNA	dsRNA1	dsRNA2	p-value 1	p-value 2
CG31053	0.00	0.35	-0.06	3.7E-02	7.1E-01
CG5604	0.00	-0.57	-0.61	3.8E-03	2.3E-03
fzy	0.00	-3.58	-2.69	1.3E-08	2.0E-07
poe	0.00	-0.92	-1.59	1.7E-04	3.0E-06
UbcD6	0.00	-4.12	-2.95	3.2E-09	1.6E-07
Ufd1-like	0.00	-2.56	-2.12	3.8E-09	1.7E-08

Table S4.V qPCR validation of selected candidate dsRNAs

RNAi reagents targeting selected hits were validated for depletion efficacy by qPCR. Results are $\log_2$ normalized to *GFP* dsRNA control samples.

4.4 Discussion

One striking commonality shared by the E2 (*UbcD6*) and both of the E3s (*poe* and *CG5604*) we identified is that they are all linked to a particular ubiquitin-dependent protein degradation process called N-end rule. The N-end rule refers to the regulation of protein stability by a degron located at its N-terminal extremity (N-degron) and whose principal determinant is the identity of the N-terminal amino acid of the protein itself (Varshavsky 2011; Kim, Kim et al. 2014). *UbcD6* is the fly ortholog of RAD6, the main E2 enzyme linked to N-end rule dependent protein degradation (Dohmen, Madura et al. 1991; Varshavsky 2011). *Poe* (UBR4) is a member of the UBR family of E3 ligases called N-recognins, that recognize the N-degron through their UBR box domain. In contrast to the canonical N-recognins, UBR1 and UBR2, non-canonical N-recognins such as UBR4 remain relatively poorly characterized. Moreover, UBR4 appears to lack a ubiquitylation domain(Tasaki, Mulder et al. 2005) suggesting that it may play a different role than its canonical counterparts, possibly by acting in conjunction with another E3. *CG5604* is HECT domain E3 ligase whose closest ortholog is a member of the Ubiquitin Fusion Degradation (UFD) factor UFD4. UFD4 has been shown to recognize the N-terminal Ub moiety of UFD substrates (Ju, Wang et al. 2007). Interestingly, UFD4 (and other UFD pathway factors) have been shown to work synergistically with N-end rule pathway (Hwang, Shemorry et al. 2009) and to interact physically and functionally with UBR1 in a manner that contributes to the ubiquitylation and degradation of N-end rule substrates (Hwang, Shemorry et al. 2010).

The identification of these three factors together strongly suggests that MAPK is the target of an N-end rule type regulation. When the different *Drosophila* UBR family orthologs were re-tested, we found that the UBR1/2 ortholog also showed a capacity to restore MAPK levels (Figure S4.6 and Table S4.VI) though, unlike poe/UBR4, it caused a significant reduction in cell count (Table S4.VI), perhaps explaining why it was not retained in our screen. It is possible that UBR4, like UBR1, has the capacity to form a complex and work together with UFD4 to recognize an N-terminal or internal degron on MAPK. Intriguingly, UBC4 (the main E2 associated to UFD4) was not found to act on MAPK levels indicating that if there is such an UBR4-UFD4 complex, it most likely would operate differently from the UBR1/RAD6-UFD4/UBC4 complex described in yeast. Moreover, poe/UBR4's apparent lack

of an ubiquitylation domain suggests that it relies on other E3s such as CG5604/UFD4 or even UBR1/2.

In contrast to the previously mentioned factors, *Usp47* has never been linked to the N-end rule pathway to data. Our current understanding of the N-end rule process does not include any major players in terms of deubiquitinase activity. Thus, our data linking *Usp47* to N-end rule factors presents the intriguing possibility that this DUB may act on a subset of N-end rule targets. Alternatively, as N-end rule substrates are not generally known to be targeted by DUBs, this may indicate that *UbcD6*, *poe* and *CG5604* act on MAPK through an N-end rule independent mechanism.

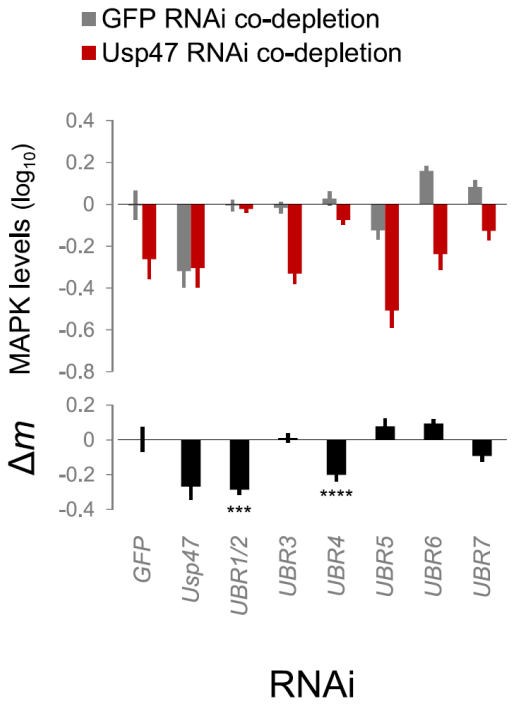
The screening technique we employed in this study differs from classical techniques in that it was based on a genetic interaction approach through co-depletion by RNAi. Our approach is reminiscent of techniques commonly used in yeast genetic screens but has only recently been introduced to the RNAi screening field (Horn, Sandmann et al. 2011; Laufer, Fischer et al. 2014). Our results further support this approach as a viable technique in identifying functionally related factors, such as factors counteracting *Usp47* in this case. The advantage of such co-depletion based approaches is highlighted by the fact that the ubiquitin-proteasome components and ligases identified here were not identified in the initial RAS-MAPK pathway screen that uncovered *Usp47*. This can be readily explained by our observation that these factors only have a sizeable impact on MAPK levels in a context where *Usp47* is depleted. Thus, while the regulatory landscape of the RAS/MAPK pathway has now been extensively explored through single-depletion RNAi screening (Friedman, Tucker et al. 2011; Ashton-Beaucage, Udell et al. 2014), it is likely that other factors would be uncovered by further co-depletion/genetic interaction RNAi screens.

Our findings indicate that *Usp47*, together with *UbcD6*, *poe* and *CG5604*, is part of a regulatory process that acts to control MAPK levels post-translationally. This finding adds another layer of regulation to the set of transcriptional and splicing regulators we have described previously (Ashton-Beaucage, Udell et al. 2010; Ashton-Beaucage, Udell et al. 2014) and lends more weight to the notion that MAPK is subjected to a significant amount of regulatory input compared to other pathway components. *Usp47*'s activity is, however, not specific to MAPK and has even been previously described as a negative regulator acting

downstream of MAPK in the context of eye development where it stabilizes the transcriptional repressor TTK (*Tramtrack*) (Bajpe, van der Knaap et al. 2008). Importantly, a different E3 ligase, SINA, was implicated in the context of TTK, indicating a difference in the way these two components are regulated. It is possible that *Usp47* plays a dual function in the context of eye development. This would explain the different outcomes we obtained in our genetic interaction experiments (Figure 4.1 C, D and Figure S4.1 A, B) and (Bajpe, van der Knaap et al. 2008); specific genetic lesions in the MAPK pathway may be more sensitive to the positive or negative impact of *Usp47* regulation. One plausible model is that *Usp47* inactivation leads to a rapid degradation of TTK followed by a slower degradation of MAPK. This would constitute a positive input at first that then transitions into a negative feedback loop as MAPK levels decrease.

Figure S4.6 Co-depletion of other UBR family factors with *Usp47*

We re-tested all of the Drosophila UBR family factors to verify whether other factors of this family might also suppress the impact of *Usp47* dsRNA on MAPK levels. Independent dsRNAs were designed for *UBR1/2* (*CG9086*), *UBR3* (*CG42593*), *UBR4* (*poe*), *UBR5* (*hyd*), *UBR6* (*FBX011*) and *UBR7* (*CG15141*). Other than *UBR4/poe*, only *UBR1/2* had a robust alleviating impact on MAPK levels when co-depleted with *Usp47* (***: p<0.001; ****: p<0.0001). However, unlike *UBR4*, *UBR1/2* RNAi caused a sharp reduction in cell count (see detailed experimental data Table S4.VI).



dsRNA	*GFP* dsRNA (MAPK)	*Usp47* dsRNA (MAPK)	Δ*m*	p-val. GFP dsRNA	p-val. Usp47 dsRNA	p-val Δm	*GFP* dsRNA (actin)	*Usp47* dsRNA (actin)	*GFP* dsRNA (cell count)	*Usp47* dsRNA (cell count)
GFP	-0.01	-0.26	**0.00**	1.0E+00	2.1E-18	1.0E+00	0.00	0.02	0.00	-0.01
Usp47	-0.32	-0.31	**-0.27**	1.5E-13	1.8E-12	7.0E-08	-0.02	0.00	0.04	0.03
UBR1/2	-0.01	-0.02	**-0.29**	9.9E-01	6.3E-01	9.8E-06	0.12	0.14	-0.72	-0.73
UBR3	-0.02	-0.33	**0.01**	7.7E-01	1.5E-10	8.8E-01	-0.02	-0.02	0.04	0.04
poe/UBR4	0.03	-0.07	**-0.20**	3.7E-01	5.9E-02	8.5E-04	-0.02	-0.02	0.03	0.06
UBR5	-0.13	-0.51	**0.08**	2.0E-03	4.3E-15	1.8E-01	0.01	0.01	-0.11	-0.11
UBR6	0.16	-0.24	**0.09**	4.5E-05	3.7E-07	1.1E-01	0.16	0.07	-0.97	-1.24
UBR7	0.08	-0.13	**-0.09**	1.8E-02	1.8E-03	9.7E-02	0.08	0.04	-0.13	-0.06

Table S4.VI UBR co-depletion experiment data

Results from the UBR family factor RNAi co-depletion experiment. dsRNAs used were different from those used in the primary screen. MAPK levels, Δ*m*, actin levels and cell count are $\log_{10}$ normalized values. All values shown here are normalized to *GFP* dsRNA treated controls. P-values shown are for a Student's t-test comparing the experimental result to those of *GFP* dsRNA control samples.

4.5 Experimental Procedures

4.5.1 RNAi

RNAi in S2 cells was conducted using dsRNA following previously described procedures (Ashton-Beaucage, Udell et al. 2014). For the *Usp47* co-depletion screen and other quantitative microscopy-based experiments, cells were pre-incubated with *Usp47* or GFP control dsRNA for three days and then distributed in clear bottom 384 well plates (Greiner) containing the dsRNA sets for a second three day depletion. This was done as *Usp47*'s impact on MAPK levels was strongest at incubation times greater than three days whereas other dsRNAs can cause lethality at longer depletion times.

4.5.2 High Content Microscopy

RNAi-treated cells S2 cells in 384 well plates were fixed in 4% paraformaldehyde and stained overnight with an anti-MAPK antibody (1/1000, Cell Signaling #4695), phalloïdin and DAPI. Plates were imaged on an Operetta (Perkin Elmer) imaging platform. Acquired images were segmented and analyzed using Harmony software (Perkin Elmer) to derive the average MAPK and actin levels per cell as well as the average cell count per well.

4.5.3 Fly Genetics and Microcopy

Fly husbandry was conducted according to standard procedures. The $Usp47^{\Delta 1}$, $Usp47^{\Delta 2}$ and $Usp47^{rev}$ lines were initially described in (Bajpe, van der Knaap et al. 2008). RNAi clones were generated as previously described (Ashton-Beaucage, Udell et al. 2014). RNAi expression in wing discs was carried out using an *engrailed-Gal4* driver line also carrying *UAS-dcr2* to enhance RNAi depletion effectiveness.

4.6 Acknowledgements

We thank the Bloomington, NIG-fly and VDRC stock centers for fly stocks. We also thank C.P. Verrijzer for providing us with the *Usp47* mutant flies.

4.7 References

Ashton-Beaucage, D., C. M. Udell, et al. (2014). "A functional screen reveals an extensive layer of transcriptional and splicing control underlying RAS/MAPK signaling in *Drosophila*." PLoS Biol **12**(3): e1001809.

Ashton-Beaucage, D., C. M. Udell, et al. (2010). "The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in *Drosophila*." Cell **143**(2): 251-262.

Axelsson, E., T. Sandmann, et al. (2011). "Extracting quantitative genetic interaction phenotypes from matrix combinatorial RNAi." BMC Bioinformatics **12**: 342.

Bajpe, P. K., J. A. van der Knaap, et al. (2008). "Deubiquitylating enzyme UBP64 controls cell fate through stabilization of the transcriptional repressor tramtrack." Mol Cell Biol **28**(5): 1606-1615.

Dohmen, R. J., K. Madura, et al. (1991). "The N-end rule is mediated by the UBC2(RAD6) ubiquitin-conjugating enzyme." Proc Natl Acad Sci U S A **88**(16): 7351-7355.

Fortini, M. E., M. A. Simon, et al. (1992). "Signalling by the sevenless protein tyrosine kinase is mimicked by Ras1 activation." Nature **355**(6360): 559-561.

Friedman, A. A., G. Tucker, et al. (2011). "Proteomic and functional genomic landscape of receptor tyrosine kinase and ras to extracellular signal-regulated kinase signaling." Sci Signal **4**(196): rs10.

Horn, T., T. Sandmann, et al. (2011). "Mapping of signaling networks through synthetic genetic interaction analysis by RNAi." Nat Methods **8**(4): 341-346.

Hwang, C. S., A. Shemorry, et al. (2010). "The N-end rule pathway is mediated by a complex of the RING-type Ubr1 and HECT-type Ufd4 ubiquitin ligases." Nat Cell Biol **12**(12): 1177-1185.

Hwang, C. S., A. Shemorry, et al. (2009). "Two proteolytic pathways regulate DNA repair by cotargeting the Mgt1 alkylguanine transferase." Proc Natl Acad Sci U S A **106**(7): 2142-2147.

Ju, D., X. Wang, et al. (2007). "The armadillo repeats of the Ufd4 ubiquitin ligase recognize ubiquitin-fusion proteins." FEBS Lett **581**(2): 265-270.

Kim, H. K., R. R. Kim, et al. (2014). "The N-terminal methionine of cellular proteins as a degradation signal." Cell **156**(1-2): 158-169.

Kolch, W. (2005). "Coordinating ERK/MAPK signalling through scaffolds and inhibitors." Nat Rev Mol Cell Biol **6**(11): 827-837.

Laufer, C., B. Fischer, et al. (2014). "Measuring genetic interactions in human cells by RNAi and imaging." Nat Protoc **9**(10): 2341-2353.

Malumbres, M. and M. Barbacid (2003). "RAS oncogenes: the first 30 years." Nat Rev Cancer **3**(6): 459-465.

Mani, R., R. P. St Onge, et al. (2008). "Defining genetic interaction." Proc Natl Acad Sci U S A **105**(9): 3461-3466.

McKay, M. M. and D. K. Morrison (2007). "Integrating signals from RTKs to ERK/MAPK." Oncogene **26**(22): 3113-3121.

Parsons, J. L., Dianova, II, et al. (2011). "USP47 is a deubiquitylating enzyme that regulates base excision repair by controlling steady-state levels of DNA polymerase beta." Mol Cell **41**(5): 609-615.

Peschiaroli, A., J. R. Skaar, et al. (2010). "The ubiquitin-specific protease USP47 is a novel beta-TRCP interactor regulating cell survival." Oncogene **29**(9): 1384-1393.

Rorth, P., K. Szabo, et al. (2000). "The level of C/EBP protein is critical for cell migration during *Drosophila* oogenesis and is tightly controlled by regulated degradation." Mol Cell **6**(1): 23-30.

Sako-Kubota, K., N. Tanaka, et al. (2014). "Minus end-directed motor KIFC3 suppresses E-cadherin degradation by recruiting USP47 to adherens junctions." Mol Biol Cell.

Schubbert, S., K. Shannon, et al. (2007). "Hyperactive Ras in developmental disorders and cancer." Nat Rev Cancer **7**(4): 295-308.

Tasaki, T., L. C. Mulder, et al. (2005). "A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons." Mol Cell Biol **25**(16): 7120-7136.

Udell, C. M., T. Rajakulendran, et al. (2011). "Mechanistic principles of RAF kinase signaling." Cell Mol Life Sci **68**(4): 553-565.

Varshavsky, A. (2011). "The N-end rule pathway and regulation by proteolysis." Protein Sci.

Weinstock, J., J. Wu, et al. (2012). "Selective Dual Inhibitors of the Cancer-Related Deubiquitylating Proteases USP7 and USP47." ACS Med Chem Lett **3**(10): 789-792.

Yang, S. W., K. H. Oh, et al. (2013). "USP47 and C terminus of Hsp70-interacting protein (CHIP) antagonistically regulate katanin-p60-mediated axonal growth." J Neurosci **33**(31): 12728-12738.

Zebisch, A., A. P. Czernilofsky, et al. (2007). "Signaling through RAS-RAF-MEK-ERK: from basics to bedside." Curr Med Chem **14**(5): 601-623.

4.8 Supplemental experimental procedures

4.8.1 RNAi Reagents

The dsRNA library was generated from dsDNA templates purchased from Open Biosystems (Huntsville, AL; http://www.openbiosystems.com) and is described in detail in (Goshima, Wollman et al. 2007). In-house dsRNA reagents were also added to supplement the collection. The individual dsRNAs were generated by *in vitro* transcription using T7 RNA polymerase. Following NaOAc ethanol precipitation, dsRNA concentration was assessed by gel dosage. Individual dsRNA probes were added to cells to a final concentration of 10 ug/mL.

4.8.2 Genetic Interaction Score

For the purposes of the RNAi screen, genetic interaction was defined as deviation from the expected result (neutral phenotype), which is a common definition of genetic interaction when working with a quantitative phenotype (MAPK levels in our case) (Mani, St Onge et al. 2008). The expected result is obtained by adding the single RNAi depletion effects together. The combined depletion effect can then be compared to the expected result. The genetic interaction score, Δm, is thus derived as follows:

$$\Delta m = \log\left(\frac{m_x}{m_g}\right) - \log\left(\frac{m_{xu}}{m_{gu}}\right)$$

Where m_x is the measured MAPK signal upon knocking down a given gene x and m_g is the MAPK signal for *GFP* RNAi controls. m_{xu} is the MAPK signal for co-depletion of gene x with *Usp47*. m_{gu} is the MAPK signal of *GFP* and *Usp47* co-depletion controls. Thus, a dsRNA whose impact on MAPK levels is equal in single knockdown or in co-depletion with *Usp47* will have a Δm of 0 (no genetic interaction). Non-equivalent depletion effects will differ from 0 and indicate genetic interaction with *Usp47*.

A negative Δm would be obtained in the case of factors that alleviate *Usp47*'s effect on MAPK (alleviating genetic interaction). These would include factors, such as *Uba1*, that restore MAPK levels by partially or completely negating *Usp47*'s impact (alleviating rescue effect; the factor is epistatic to *Usp47*) and also potential redundant factors, such as DUB, whose co-depletion with *Usp47* would be less than the sum of the individual RNAi (alleviating redundant).

On the other hand, a positive Δ*m* would occur in cases where the impact on MAPK levels in *Usp47* co-depletion is synergistic or greater than the sum of the individual depletion effects (synthetic genetic interaction). For instance, a potential factor acting redundantly with *Usp47* to stabilize MAPK might have little impact when depleted on its own. However, in the absence of *Usp47*, depleting this factor might now cause a further drop in MAPK levels (synthetic redundant). Another case would be a negative regulator of *Usp47* whose impact on MAPK levels is negated by *Usp47* co-depletion (*Usp47* would be epistatic to this factor in this case).

4.8.3 Candidate Selection Criteria

For the *Usp47* co-depletion screen, we used a dual cutoff based hit selection strategy. The hits with a Δ*m* false discovery rate (FDR) below $1x10^{-10}$ were retained irrespective of other parameters. Hits with a Δ*m* FDR between $1x10^{-10}$ and $1x10^{-3}$ were only retained if their cell count FDR was also above $1x10^{-10}$ (Figure S3E). This second cutoff critera was introduced to bias hit selection towards factors that did not cause a significant change in cell count but had a weaker confidence Δ*m*.

An RNAi validation experiment was then conducted on selected hits. For this, we synthesized two separate dsRNA reagents per candidate that differed from the primary screen dsRNA and, when possible, did not overlap with its sequence. The confirmation criteria used for validation were a Δ*m* p-value < 0.05 and an absolute change in MAPK levels above 0.1 (log_{10} normalized). Validate hits also were required to have an identical genetic interact type to that of the primary screen (same sign Δ*m*). A number of proteasome components failed on this last criteria, possibly due to their broader impact that includes cell lethality; it is conceivable that an alleviating impact may be observed following depletion followed by a drop in MAPK levels due to the general impact of cell death on protein stability due to caspases, for example. Notably, none of the positive Δ*m* candidates (synergistic interactions) were confirmed in our validation experiments.

4.9 Supplementary references

Goshima, G., R. Wollman, et al. (2007). "Genes required for mitotic spindle assembly in *Drosophila* S2 cells." Science **316**(5823): 417-421.

Mani, R., R. P. St Onge, et al. (2008). "Defining genetic interaction." Proc Natl Acad Sci U S A **105**(9): 3461-3466.

5 Discussion

Les éléments de discussion ayant été approfondis dans les chapitres précédents ne seront pas présentés de nouveau dans cette section. On visera plutôt ici à amener des éléments additionnels qui découlent, soit d'avancées plus récentes, soit d'idées qui n'ont pas pu être inclues dû aux restrictions de format. Cette section inclut également un article de commentaire publié dans la revue *Fly* en 2011 en complémentaire à l'article figurant au chapitre 2. Ce commentaire présente quelques données additionnelles sur l'EJC et, surtout, une discussion plus approfondie des résultats et des implications de ceux-ci en regard du modèle d'action de l'EJC sur l'épissage alternatif.

5.1 Les facteurs d'épissage constitutifs

L'article sur le crible RNAi RAS^{V12} du chapitre 3 met en évidence plusieurs facteurs d'épissage constitutifs et leur impact sur l'épissage alternatif de *mapk*. La discussion y fait mention de différentes évidences liant les facteurs constitutifs à l'épissage alternatif. À celles-ci, il serait sans doute important de mentionner la parution récente d'une étude à large échelle de l'impact fonctionnel des facteurs d'épissage constitutifs sur l'épissage alternatif d'un ensemble d'événements d'épissage alternatifs liés à l'apoptose et la prolifération (Papasaikas, Tejedor et al. 2014). Les auteurs y constatent que les facteurs centraux du spliceosome ont des effets plus prononcés et plus large sur l'épissage alternatif alors que les facteurs périphériques ou associés au spliceosome de manière plus transitoire ont des effets plus restreints et moins prononcés. Ils observent aussi que la déplétion de différents facteurs constitutifs est associée à des profils d'épissage alternatifs distincts et que ces profils sont plus ou moins semblables selon l'ordre et la durée de recrutement de ces composantes au spliceosome. De plus, ils montrent aussi que la perturbation des facteurs d'épissage constitutifs n'est pas associée à une baisse globale de l'efficacité d'épissage. En effet, l'épissage de plusieurs introns demeure inchangé et des événements d'exclusion d'intron sont observés dans d'autres cas. En dernier lieu, comme nous, ils observent que plusieurs composantes associés aux formes plus avancées du spliceosome ont

également un impact sur l'épissage alternatif, ce qui indique que la modulation de l'épissage alternatif peut se produire même lors des étapes tardive, allant jusqu'à l'étape de l'activation catalytique. Cette observation entre en opposition avec le modèle classique de l'épissage alternatif où celui-ci régule l'association du pré-spliceosome aux sites d'épissage.

Dans l'étude précédente, les événements d'épissage examinés étaient tous des événements d'épissage alternatif caractérisés. Dans le cas de notre étude portant sur *mapk*, comme il s'agit de la description de nouveaux événements d'épissage, on aurait pu penser qu'il s'agissait de défauts d'épissage impliquant des exons constitutifs, d'autant plus que les facteurs d'épissage impliqués sont surtout des facteurs dits constitutifs. Par contre, certains des transcrits résultant étaient également observables dans des cellules de type sauvage ainsi que dans des bases de données publiques répertoriant des données de séquençage d'ARN de différents tissus de mouche (Graveley, Brooks et al. 2010). Il est donc envisageable qu'il s'agisse d'événements d'épissage modulables in vivo. Des études plus approfondies devront cependant établir le contexte où de telles modulations sont produites et l'étendue réelle de leur impact sur la signalisation MAPK. Or, la régulation des facteurs d'épissage constitutifs dans un contexte d'épissage alternatif est un domaine encore très peu exploré. Ce sera donc doublement intéressant d'explorer cet aspect de l'expression de *mapk*.

5.2 L'EJC : un facteur d'épissage pour les transcrits à longs introns?

Cette section a été publiée en 2011 dans la revue *Fly* sous le format « Extra view », à titre de complément à l'article figurant ici au chapitre 2 (Ashton-Beaucage, Udell et al. 2010). Le texte est paru sous le titre original « The Exon Junction Complex: A splicing factor for long intron containing transcripts? ». On y retrouve quelques données et analyses additionnelles ainsi qu'une discussion plus approfondie du rôle de l'EJC dans l'épissage.

Contributions détaillées de l'auteur

Contributions expérimentales : Production de l'ensemble des données expérimentales décrites dans l'article. Analyses supplémentaires des données de transcriptome (Fig.2). Élaboration d'un outil informatique permettant d'évaluer la force des sites d'épissage et utilisé pour calculer le « information weight matrix » des sites d'épissage 5' de *mapk* (tableau II).

Contributions au manuscrit : Élaboration et rédaction du manuscrit. Montage des figures. Conception des modèles reliant l'EJC au mécanisme de définition par exon (« exon definition »).

Contributions détaillées des co-auteurs

TM a supervisé la rédaction et apporté des corrections mineures au texte.

5.2 The Exon Junction Complex: A Splicing Factor for Long Intron Containing Transcripts?

Running Title: The EJC and Splicing of Long Introns

Dariel Ashton-Beaucage[1] and Marc Therrien[1,2,*]

[1] Institute for Research in Immunology and Cancer
Laboratory of Intracellular Signaling
Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montreal, Quebec, Canada, H3C 3J7

[2] Département de pathologie et de biologie cellulaire, Université de Montréal

*** To whom correspondence should be addressed**
tel.: (514) 343-7837
fax.: (514) 343-6843

Fly

Volume 5 • Issue 3 • July/August/September 2011

LANDES
BIOSCIENCE

5.2.1 Summary

Alternative splicing provides one of the major mechanisms by which eukaryotic genomes can increase the diversity of transcripts and protein products encoded by a limited amount of genes. The spliceosome is involved in recognizing and removing intronic sequences and it is guided in this activity by splicing regulatory factors that determine the particular splice sites that are to be recognized and used for intron removal. Thus, by modulating the spliceosome's activity, these factors can cause different mRNA transcripts to be generated from the same precursor mRNA. In a recent RNAi screen for factors modulating RAS/MAPK signaling, we identified the *Drosophila* exon junction complex (EJC) as one of the components regulating the splicing of *mapk* transcripts. We showed that removal of EJC components caused multiple consecutive *mapk* exons to be skipped. Moreover, we determined that other fly genes that, like *mapk*, had particularly large introns, were also sensitive to disruption of the EJC. The importance of intron length in determining sensitivity to EJC disruption suggests that the EJC is functioning in exon definition, a splicing mechanism that occurs for long introns in particular.

5.2.2 Introduction

Splicing of precursor mRNA (pre-mRNA), whereby intronic sequences are removed and exons are ligated to form a mature mRNA, is an important step in gene expression. Introns are removed by the spliceosome which first binds to a 5' donor splice site (5'SS) and a 3' acceptor splice site (3'SS) on either side of the intron, then proceeds to excision of the intron and ligation of the two exon ends. The spliceosome is a large complex composed of many proteins and snRNAs [1-3]. In addition to the constitutive core spliceosomal components, other factors act to guide the spliceosome's choice of splice sites. These include serine/arginine rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNP) which bind to the pre-mRNA and act to promote or prohibit the use of nearby splice sites [4, 5]. Thus, alternative splicing occurs when different combinations of exons are ligated, introns retained or different 5'SS and 3'SS used. Together, these changes are responsible for a large part of transcriptome and proteome diversity in many eukaryotic genomes [6]. Illustrating this principle, recent high throughput sequencing (RNA-seq) analyses of the *D. melanogaster* transcriptome by the modENCODE consortium has reported 14,862 annotated genes that give rise to 36,139 different transcripts [7].

The splicing process also leaves traces on the mRNA; different splicing factors have been shown to remain bound to the mRNA after splicing occurs and other factors are known to be recruited as a consequence of splicing. These marks left on the mRNA are involved in the regulation of many post-splicing events including mRNA export, stability and translation [8]. The exon junction complex (EJC) is perhaps one of the best examples of a factor linking splicing to post-splicing events. The EJC is recruited by the spliceosome during the splicing process [9] and is deposited on the mRNA slightly upstream of exon junction sites after the splicing reaction is completed [10]. Once bound to the spliced mRNA, the EJC has been shown to play multiple parts, namely it regulates the export, subcellular localization and translation of certain transcripts [11-16]. It also plays a central role in nonsense mediated decay (NMD) – a quality control mechanism which is responsible for the degradation of transcripts with premature stop codons – by recruiting proteins like the Upf factors [13, 17-20]. Thus, the EJC is chiefly thought of as a central regulator of many aspects of the mRNA's post-splicing journey.

Recently, a study published by Jean-Yves Roignant and Jessica Treisman and another study by our group have demonstrated that – in addition to its post-splicing functions – the EJC

also influences the splicing process *per se* [21, 22]. Treisman's group initially identified the EJC factor *mago-nashi* (*mago*) in a genetic screen for factors affecting photoreceptor differentiation in the *Drosophila* eye [23] and they subsequently showed that disruption of *mago* specifically impedes signaling through the EGFR/RAS1/MAPK pathway [21]. In an RNAi screen conducted in S2 cells for factors modulating $RAS1^{V12}$-dependent activation of MAPK, we identified *mago* as well as three other EJC components, *eIF4AIII, tsunagi* (*tsu*) and the SR factor, *RnpS1* [22]. Both studies show that disruption of the EJC inhibits RAS1/MAPK signaling and that this effect is due to a specific reduction of MAPK protein levels. This decrease is shown to be attributable to changes in the splicing of the *mapk* transcripts (Figure 5.1). Finally, the EJC is found to produce similar splicing defects in other fly transcripts which, like *mapk*, are multi exon genes bearing large introns.

In this commentary, we provide further details and results relative to the EJC's effect on splicing of the *mapk* transcripts and expand on its general splicing function. We also present some data on other SR factors which, like RNPS1, may be involved in regulating *mapk* expression. Finally, we cover some of the differences between our findings and those of Roignant and Treisman. In particular, we discuss the importance of heterochromatic location, which we do not find to be a common characteristic of all EJC regulated genes, and provide data indicating that SRM160 does not have an EJC-like effect on the *mapk* transcripts.

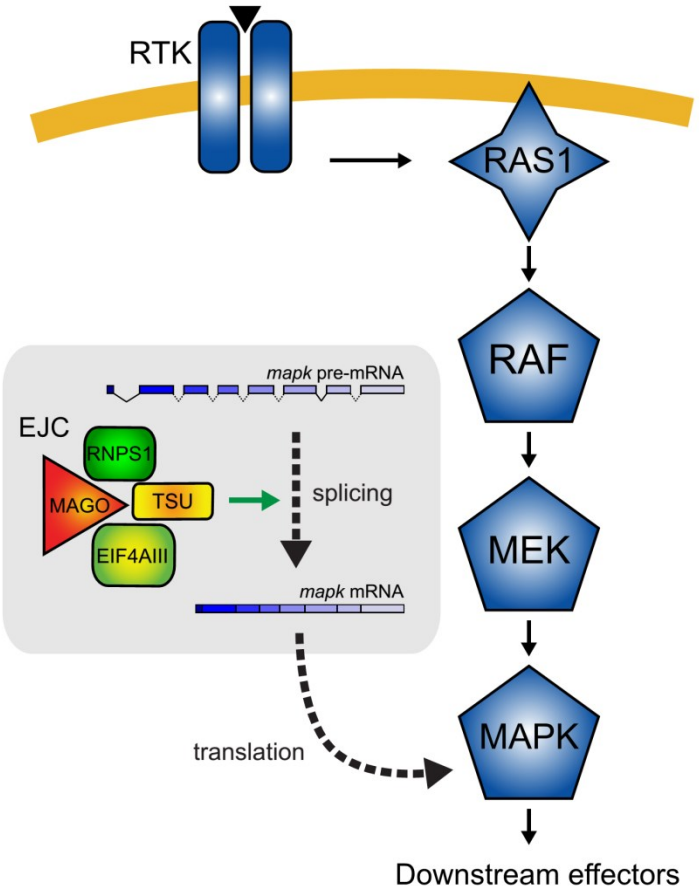


Figure 5.1 Schematic Representation of the EJC's Impact on RAS1/MAPK Signaling

The three nuclear EJC components (MAGO, TSU and EIF4AIII) as well as RNPS1, an SR factor and peripheral EJC component, are involved in the expression of MAPK. Presence of the EJC is specifically required for productive splicing of the *mapk* transcript.

5.2.3 The EJC's Involvement in the Splicing of Transcripts with Long Introns

One of the most striking observations in our study was that the EJC – a protein complex traditionally associated to post-splicing functions – does in fact impact the splicing process *per se*. Moreover, the EJC's effect on splicing was limited to certain genes, which was surprising given that it was generally thought to constitutively bind exon junctions in a sequence-independent manner. Interestingly, one recent paper suggests that this assumption is not accurate as the EJC was found to preferentially bind to certain exon junctions in *Drosophila* [20]. In this paper, the authors also show that the EJC's ability to induce NMD on one of the transcripts to which it binds depends on the sequence upstream of the exon junction site. Previous observations that the EJC is not required for NMD in *Drosophila* [17] can thus be explained by its restricted binding specificity. In light of this, it also seems entirely plausible that the EJC would bind only a subset of pre-mRNAs to modulate their splicing. However, the EJC's function in NMD requires it to remain bound to transcripts after they are spliced. This is not necessarily the case for splicing, implying that the binding specificity observed for NMD is not necessarily relevant to the EJC's splicing function.

Our study as well as the one of Roignant and Treisman show that the EJC's splicing function was specific to *mapk* and a limited set of other genes. In our case, both RT-PCR and transcriptional profiling using high throughput mRNA sequencing (RNA-seq) clearly demonstrate that intron length determines sensitivity to disruption of the EJC. Moreover, we found the intensity of the decrease in transcript abundance to correlate with intron length. Thus, the transcripts with the longest introns showed the highest sensitivity to EJC disruption. *mapk*'s largest intron is over 25kb long and is amongst the 1% largest *Drosophila* introns, explaining its high sensitivity to EJC disruption.

In addition to bearing unusually long introns, the *mapk* gene also has the particularity of being located in the pericentromeric heterochromatin. Roignant and Treisman propose that this characteristic, in addition to intron length, also determines sensitivity to disruption of the EJC [21]. However, their conclusion that euchromatic genes with long introns were not significantly affected differs from our own. Firstly, in our paper, we noted that EJC knockdown caused the levels of all transcripts with introns >1000nt to decrease reproducibly in our RNA-seq analysis. Second, when we specifically examined the transcripts of *Tequila* and *PMCA* – two euchromatic

genes with long introns – by RT-PCR, they showed a decrease in their abundance as well as exon skipping. This strongly suggests that a gene need not be localized in a heterochromatic region to be sensitive to EJC depletion. The summary of the transcriptome data we presented in our paper does not directly address the difference between long intron-bearing heterochromatic and euchromatic genes. We have therefore added this data here. Table 5.I shows that the 1901 euchromatic transcripts with introns >1000nt tend to consistently have lower expression levels upon EJC depletion. Still, the extent of the decrease in transcript abundance is more pronounced in the 61 heterochromatic genes with large introns. However, the associated p-value is much weaker in this case due to the lower amount of transcripts and variability of transcript abundance. Heterochromatic transcripts with short introns also decreased, although this result may not be significant due to similar reasons. In short, we believe that intron length alone can be sufficient to determine sensitivity to EJC depletion. Nevertheless, chromatin structure is known to impact splicing [24-26]. Also, heterochromatic genes like *mapk* have been previously found to require the presence of heterochromatin in their vicinity to be efficiently expressed [27]. Therefore, it is likely that chromatin structure also has an important impact on the splicing of genes located in heterochromatic regions. As a final note on the topic of heterochromatin, it is important to mention that data released by the modENCODE consortium and published recently [28] has greatly expanded our knowledge of chromatin states across the *Drosophila* genome. In light of this, it may be that many of the EJC target genes that we have defined as euchromatic – based on the heterochromatin regions released by the DHGP [29] – are in fact located in heterochromatin or heterochromatin-like pockets. Thus we cannot rule out that chromatin modification plays a role in determining the EJC's splicing function.

Finally, transcripts bearing short introns (<250nt) were not generally sensitive to disruption of the EJC by RNAi (Table 5.I), nor were genes without introns (Figure 5.2D). The EJC has been found to bind to some transcripts with short introns and regulate their sensitivity to NMD [17]. However, the observation that the expression level of luciferase reporters containing the short introns of CG7956 and *Eaat* increases upon disruption of the EJC and *Upf1* also suggests that splicing is not significantly altered in these cases [17]. Moreover, increasing the length of short introns has been shown to lower the efficiency of the splicing reaction [30]. In fact, a slightly different splicing mechanism, called exon definition, takes place when long introns (>250nt) are involved. In exon definition, the spliceosome initially assembles across the 3'SS and 5'SS of an

exon which is bordered by large introns. Exons thus "defined" can then join together and allow for the transition to intron definition – where the spliceosome is assembled across the intron – to occur and the splicing reaction to take place [30-33]. Thus it is possible that the EJC acts to bolster the less efficient exon definition splicing process that occurs when larger introns are present. Alternatively, mechanistic differences between the intron and exon definition processes may preclude the EJC's involvement in the former or allow its involvement in the latter. Since most introns in *Drosophila* are relatively small, this provides an explanation for the relatively specific impact on splicing we observed upon EJC disruption.

largest intron <250nt		X		X	X		
largest intron >1000nt			X			X	X
heterochromatin				X		X	
euchromatin					X		X
all	X						
genes	6760	3519	1962	44	3475	61	1901
average fold-change upon EJC knockdown	-0.12	0.00	-0.32	-0.27	0.00	-0.65	-0.31
p-value*		6.1E-137	1.7E-124	2.7E-02	3.7E-142	5.2E-11	3.4E-109

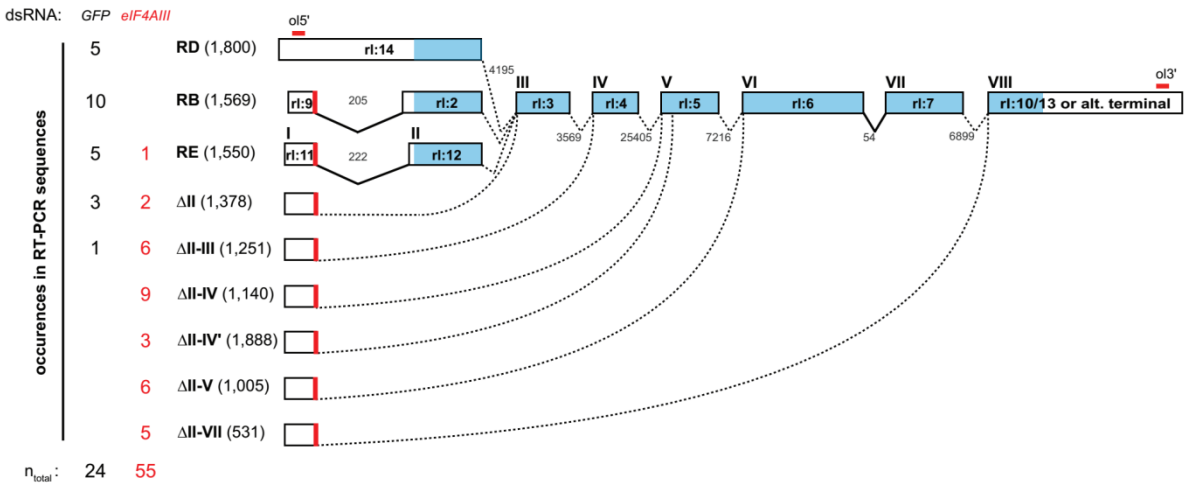
Table 5.I Summary of Heterochromatic and Euchromatic Gene Expression Changes Induced by EJC Knockdown

The number of genes for which transcripts could be detected in our transcriptional profiling experiments is indicated for each of the listed categories. Log2 fold-change is calculated from the average of the gene expression values for *mago* and *eIF4AIII* dsRNA treated cells normalized to GFP dsRNA control cells. p-values were calculated using a two-tailed Student's t test to compare the fold-changes of the indicated groups of genes to the remaining genes in the transcriptome dataset.

Figure 5.2 EJC Depletion Induces Skipping of Multiple Consecutive Exons Predominantly in the 5' of the *mapk* Transcript

Panels A and B are reproduced from [22] with some modifications and are provided as a support for some the additional data and discussion elements presented in this commentary. (A) Schematic representation of the alternative splicing isoforms observed in *GFP* and *eIF4AIII* dsRNA treated S2 cells. Each occurrence represents an RT-PCR product that was cloned and sequenced. The flybase exon names are represented (e.g. "rl:14") and exons are also numbered from I to VIII based on the rl-RE transcript in order to simplify the description of the transcripts resulting from exon skipping events. Introns lengths lengths are indicated next to the corresponding introns (not drawn to scale). The non-consensus 5'SS present on the first exon of RB and RE (rl:9 and rl:11) is shown in red. The splice isoforms that include the alternate unannotated terminal exon (see text) are not depicted separately. Instead, the different terminal exons (rl:10, rl:13 and the alternate terminal exon corresponding to a duplication of the rl:10 and rl:13 sequence) are lumped together for simplicity. The position of the primers used for RT-PCR is also shown. (B) The frequency at which individual *mapk* exons were found to be deleted in the sequenced RT-PCR products. This is represented as a percentage of the total set of sequenced RT-PCR products listed in (A). (C) The average coverage (top panels) and fold change (bottom panels) of *mapk* exons (left) or exon junctions (right) measured in our RNA-seq experiment. Exons are numbered according to the RE transcript (see (A)). The I-II exon junction is not presented because of the low abundance of sequencing reads at this site. (D) Transcriptional profiling of RNA extracts from S2 cells treated with *eIF4AIII* and *GFP* dsRNA showing only intronless genes.

A

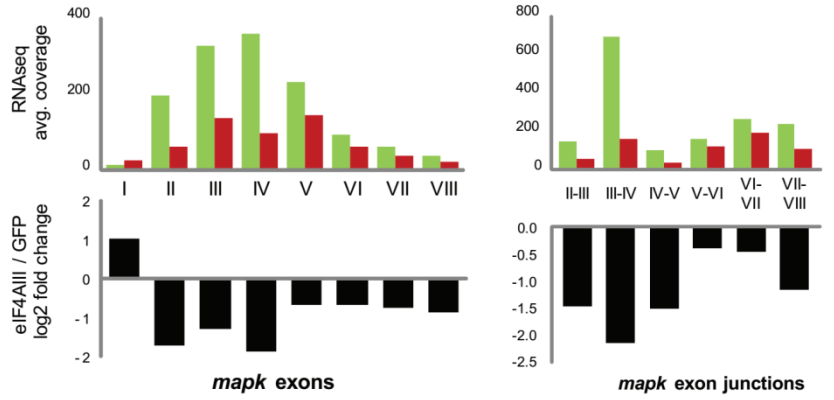


B

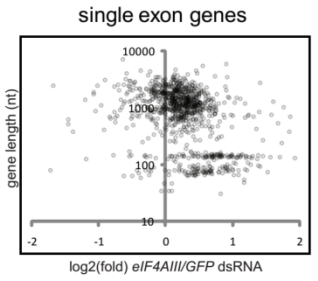
RT-PCR sequences

exon	II	III	IV	V	VI	VII
deleted *GFP* dsRNA (%)	16	5	0	0	0	0
deleted *eIF4AIII* dsRNA (%)	96	89	53	56	25	34

C



D



5.2.4 Altered Splicing of *mapk* Caused by Disruption of the EJC

In our case, the first line of evidence indicating that the EJC might be implicated in splicing came from a Northern blot experiment where we observed a decrease in the abundance of the 2.7 and 1.7kb *mapk* transcripts as well as downwards shift in their size in EJC dsRNA treated samples [22]. Further investigation of the splicing defects was conducted by RT-PCR targeting the first and last exons at sites common for the RB, RD, RE and RF *mapk* isoforms (Figure 5.2A; RC and RA could not be amplified by RT-PCR and were not detected in our RNA-seq transcriptional profiling of S2 cells). Multiple lower sized RT-PCR products were observed to be enriched in EJC depleted samples. Cloning and sequencing of these products revealed that this shift in size was due to exon skipping. Some exon skipping was also observed in GFP dsRNA treated control, indicating that these isoforms of *mapk* are also present at steady state (Figure 5.2A). Moreover, an RT-PCR product corresponding to the *mapk* RB or RE transcript with its second exon missing could clearly be observed, not only in untreated S2 cells, but in whole fly extracts from all developmental stages. Additionally, while not present in the annotated Flybase *mapk* gene model, an RB/RE-like transcript lacking the second exon is predicted based on data from the modENCODE project [7] and (http://modencode.org). This indicates that these additional alternatively spliced *mapk* isoforms are present at steady state and that disruption of the EJC may favor their production over the annotated full-sized forms. In addition to this, we also observed an alternate terminal exon in some of the RT-PCR isoforms we sequenced. This alternate exon is identical to the annotated terminal *mapk* exon except for a 1nt deletion and a 6nt insertion. It is situated approximately 4kb downstream of the annotated *mapk* gene and probably is the result of a duplication of the terminal exon. Moreover, evidence for this alternate exon is also supported by data from the modENCODE project. The alternate terminal exon was only observed in 7 of transcripts we sequenced in eIF4III depleted samples and additional work would be needed to determine if the EJC is also involved in this alternative splicing event. Finally, the length of most of the internal exons of *mapk* is divisible by 3 – rl:12 and rl:3 being the exceptions to this – which is a characteristic of alternatively spliced exons. Thus, there seems to be more alternative splicing than previously anticipated occurring in the *mapk* gene and the EJC may be one of the factors involved in modulating some part of this.

The exon skipping occurring in EJC depleted cells tended to involve the skipping of multiple consecutive exons (Figure 5.2A). Exon skipping was most frequent for exons near the 5' end and least frequent in the exons at the 3' end (Figure 5.2B). This finding was also mirrored in Roignant and Treisman's RT-qPCR results and by their RNA-seq data. Additionally, our RNA-seq results also support this observation (Figure 5.2C). Furthermore most exon skipping involved the 5' splice donor site (5'SS) of the first exon of the RB (rl:9) or RE (rl:11) transcript being joined to an alternate downstream 3' splice acceptor site (3'SS) (Figure 5.2). Interestingly, the 5'SS of this exon presents the poorest match to the donor site consensus (-1.99 bits) when compared to the other *mapk* exon 5'SSs (Table 5.II). Thus, in EJC depleted cells, exon skipping events in *mapk* predominantly involve the joining of this poor 5'SS to different downstream exons. However, while other genes containing similar poor consensus 5'SS did tend to be downregulated in EJC depleted cells, those that contained short introns were not significantly affected. Thus, intron length still remains the principal determinant of sensitivity to EJC depletion. Like heterochromatin, the presence of atypical splice sites may also shape the nature of the EJC's splicing function.

In their paper, Roignant and Treisman report a splicing defect involving the skipping of the rl:6 and rl:7 exons and the use of a cryptic splice site in rl:5 that is directly joined to the terminal exon of *mapk*. Although we did observe similar splicing changes in the RT-PCR products we sequenced, they were much less abundant than those involving the first exon (rl:9/11). Interestingly, the intron dividing rl:6 and rl:7 is the only small intron (54nt) present in the *mapk* transcript (excluding the first intron of RB and RE which is an insert intron) and those two exons were nearly always included or excluded together. Accordingly, Roignant and Treisman found that, unlike the genomic DNA *mapk* construct, the MAPK cDNA construct containing only this short intron was not sensitive to EJC depletion. This fits in nicely with the notion that introns spliced via intron definition are not subjected to the EJC's activity. Moreover, short introns are not generally permissive to exon skipping [30, 34], which explains why we did not observe a transcript with rl:9/11 joined to rl:7 (Figure 5.2A).

exon intron

	-3	-2	-1	1	2	3	4	5	6	7	
	M	A	G	G	U	R	A	G	U	W	
A	0.37	0.51	0.16	0.00	0.00	0.59	0.74	0.08	0.13	0.36	$f(b, l)$
C	0.27	0.16	0.07	0.00	0.01	0.01	0.06	0.03	0.12	0.14	
G	0.19	0.16	0.64	1.00	0.00	0.34	0.10	0.83	0.07	0.20	
U	0.18	0.16	0.13	0.00	0.99	0.05	0.09	0.06	0.68	0.30	
A	0.53	1.02	-0.67	-8.65	-9.44	1.23	1.55	-1.61	-0.92	0.49	$R_{iw}(b, l)$
C	0.07	-0.68	-1.78	-11.29	-4.74	-4.13	-1.99	-3.26	-1.10	-0.88	
G	-0.40	-0.64	1.34	1.98	-9.91	0.43	-1.28	1.71	-1.94	-0.31	
U	-0.53	-0.62	-0.97	-10.61	1.97	-2.27	-1.45	-2.06	1.43	0.26	
rl:16	A	U	G	G	U	G	G	G	A	A	5.64
rl:14	U	G	U	G	U	A	A	G	U	A	8.22
rl:1	A	A	A	G	U	A	A	A	U	A	7.92
rl:3	A	A	C	G	U	A	A	A	U	A	6.80
rl:4	C	A	G	G	U	A	A	U	G	G	4.85
rl:5	A	A	A	G	U	A	U	G	U	A	8.25
rl:6	A	A	G	G	C	A	A	G	U	A	6.54
rl:7	G	A	G	G	U	A	C	U	A	C	1.29
rl:11 / rl:9	A	A	A	G	U	U	G	A	G	U	-1.99

R_i (bits per site)

Table 5.II Comparison of the 5'SS Sequences of the *mapk* Exons

The 5'SS sequences of the annotated *mapk* exons are presented here alongside the consensus 5'SS proposed in [51]. The frequency of occurrence of each base in all the 5'SS present in FlyBase (R5.34) is also shown. The conservation of each splice donor site was assessed by using the individual information R_i technique described in [49]. The information weight matrix $R_{iw}(b,l)$ calculated from the frequency of occurrence of the bases *f(b,l)* is displayed as well as the calculated R_i for each of the *mapk* 5'SSs.

5.2.5 Additional Factors Involved in *mapk* Splicing

In addition to the three core EJC factors, we also reported the identification of *RnpS1* as another factor regulating the splicing of *mapk*. RNPS1 is a protein of the SR family, which is typically associated to the regulation of splicing and, accordingly, was initially characterized as a splicing factor [35, 36]. Subsequent studies also associated *RnpS1* to the EJC via a role for this component in RNA export and NMD [15, 16, 19]. Thus, RNPS1 is both known as a splicing factor and as a so-called peripheral EJC component. The fact that *RnpS1* depletion has a similar impact on *mapk* splicing to that produced by the depletion of *mago*, *tsu* or *eIF4AIII* indicates that these two aspects of RNPS1 function (splicing factor and EJC component) can effectively be reconciled.

It is important to note that *RnpS1* depletion had a weaker effect on *mapk* transcript levels and did not seem to impact some of the other EJC targets. This suggests that RNPS1 may not be involved in all the splicing functions that the EJC is involved in. In this perspective, the core EJC factors might be acting as a common platform to recruit different splicing factors like RNPS1. This would explain why knocking down core EJC components has a much more profound impact on the levels of *mapk* and other transcripts. Alternatively, RNPS1 may be redundant with other EJC-associated splicing factors, explaining why its depletion only produces a fraction of the EJC's effect.

5.2.5.1 SRm160

Because *RnpS1* had a less drastic effect on *mapk* splicing than the other EJC components, we thought that this might be due to functional redundancy with another SR factor. *SRm160*, for example, is another peripheral EJC component [10] that has also been previously linked to modulation of splicing [37, 38]. Indeed, Roignant and Treisman mention that they observe a drop in *mapk* levels upon depletion of *SRm160* and that this effect is additive when both *RnpS1* and *SRm160* are depleted [21]. In our RNAi screen, *SRm160* caused a modest decrease of pMAPK levels that was not sufficient for it to be retained as a hit (data not shown). Still, we also observed a decrease of MAPK protein levels upon depletion of *SRm160* (Figure 5.3). Moreover, when combining *SRm160* and *RnpS1* dsRNA, we could also observe an additive effect (Figure 5.3).

However, *SRm160* RNAi did not produce an increase in the abundance of the shorter *mapk* transcripts that were a characteristic of the exon skipping induced by the depletion of the other EJC components (Figure 5.4).

In order to identify other SR factors involved in *mapk* splicing that might be redundant with *RnpS1*, we screened a set of candidate SR factors by RNAi and measured their effect on MAPK protein levels. The different dsRNA probes were tested alone and in combination with *RnpS1* to look for any possible additive effects. Some of the candidates we tested caused a reduction in MAPK levels (Figure 5.3) which tended to be modest. However, other than *SRm160*, only *U2af38* – a SR factor involved in the basal splicing machinery – had a clear additive effect when combined with *RnpS1* dsRNA (Figure 5.3). Because perturbing the basal splicing machinery also seems to cause a synergistic effect with *RnpS1*, this suggests that measuring *mapk* expression levels alone might not be an accurate method to identify additional components that have EJC-like role on splicing; factors that have distinct splicing functions might also show such additive properties.

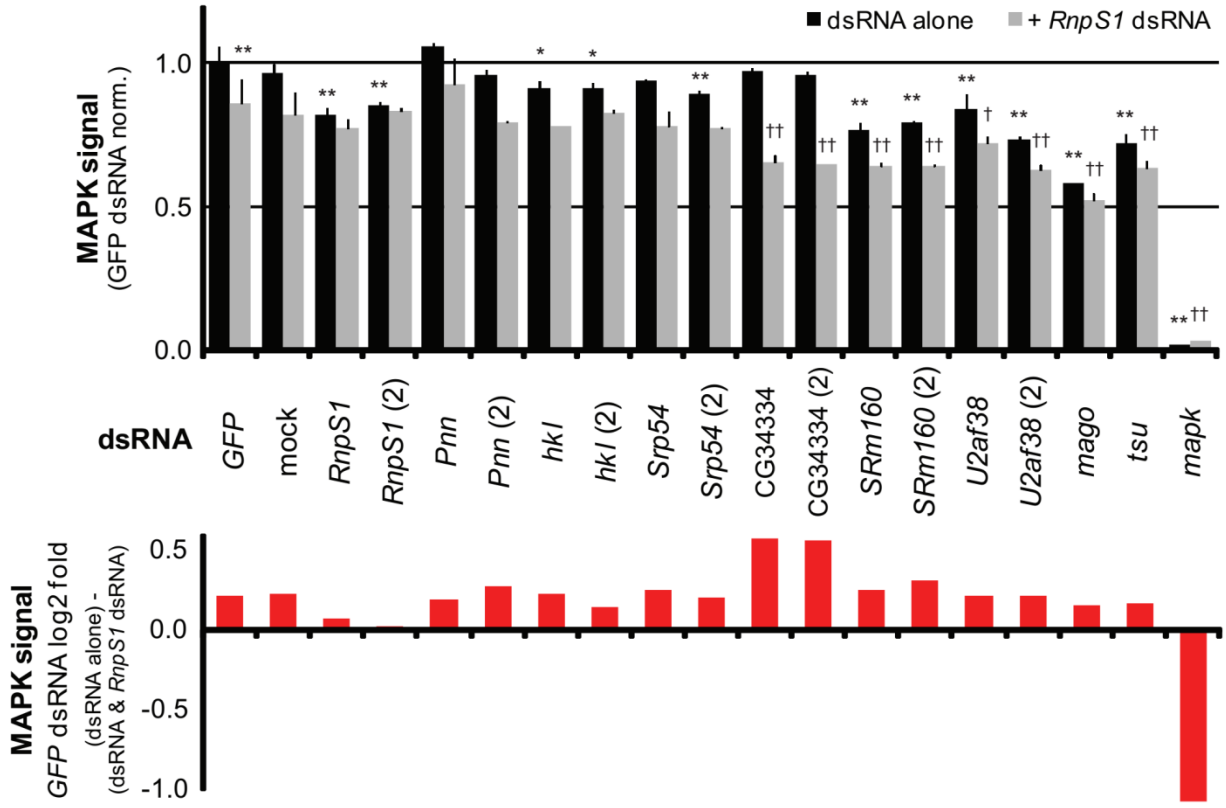


Figure 5.3 Additional SR Factors Impact MAPK Protein Levels

MAPK protein levels measured using quantitative microscopy in S2 cells treated with the indicated dsRNAs. MAPK levels are represented as a ratio of *GFP* dsRNA treated controls (top panel). The indicated dsRNAs were used alone (black) or in combination with RnpS1 dsRNA (grey). A Students T-test was used to compare the tested experimental dsRNAs to the control sample treated with *GFP* dsRNA alone ("*": p-value <0.05 and "**": p-value <0.01). To evaluate additive effects with *RnpS1* dsRNA, experimental dsRNAs combined with *RnpS1* dsRNA were compared to the control sample teated with *GFP* and *RnpS1* dsRNA ("†": p-value <0.05 and "††": p-value <0.01). The difference between the MAPK levels measured for the tested dsRNAs (+/- RnpS1 dsRNA) is represented as log2 fold change of the *GFP* dsRNA controls (bottom panel).

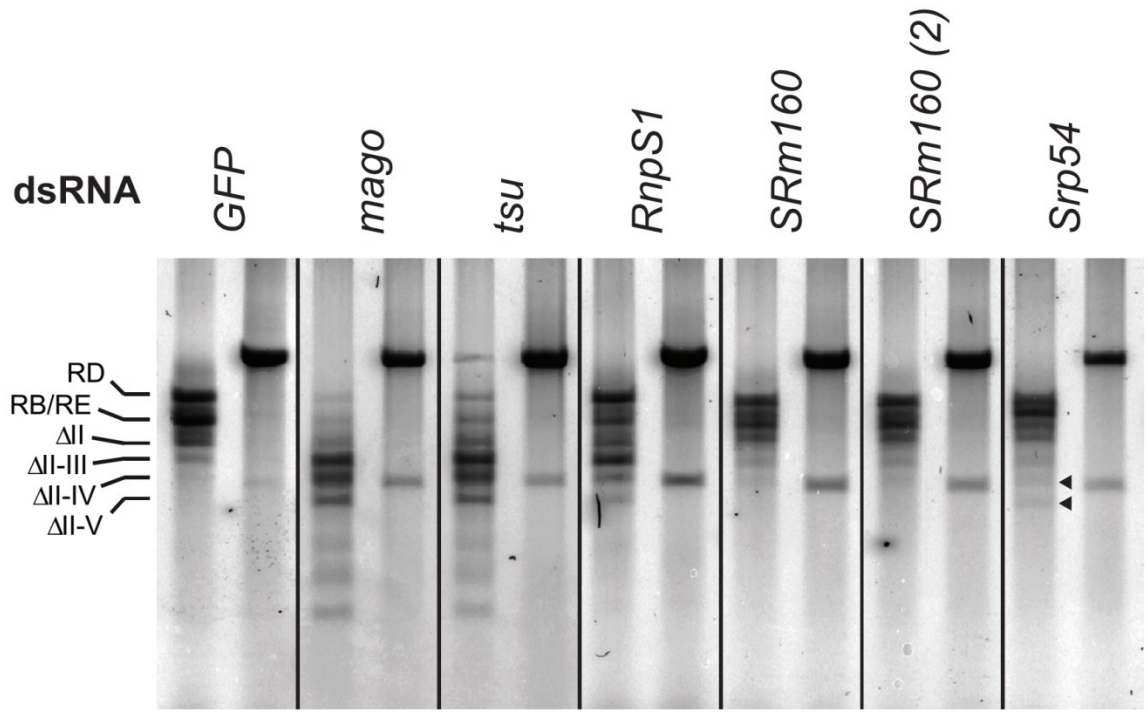


Figure 5.4 *SRm160* dsRNA Does Not Produce Clear EJC-like Changes in the *mapk* Transcripts

mapk (left columns) and *Src42A* (right columns) transcripts visualized by RT-PCR. Expected sizes of the annotated *mapk* transcript isoforms as well as the observed alternate isoforms resulting from exon skipping (see Figure 5.2) are indicated on the left-hand side. S2 cells treated with *mago*, *tsu* and *RnpS1* dsRNAs show a clear increase in the ΔII-III, ΔII-IV and ΔII-V isoforms. *SRm160* dsRNA does not produce an increase in these forms. *SRp54* dsRNA, which has a less potent impact on MAPK protein levels compared to *SRm160* (see Figure 5.3), produces a slight increase in the ΔII-IV and ΔII-V *mapk* isoforms (arrowheads). None of the tested dsRNAs seem to impact the splicing of *Src42A*. The faint lower sized band observable in some of the *Src42A* lanes was not consistently observed to vary significantly across different experiments we conducted and may possibly be a PCR contaminant.

5.2.5.2 Srp54 and hkl

Thus, to address if any of these SR factors might have a function in *mapk* splicing similar to the EJC's, we looked at their effect on the abundance of the different *mapk* transcript isoforms by RT-PCR. Of the candidates we tested, only *Srp54* produced a slight increase in the ΔII-IV and ΔII-V *mapk* splice isoforms (Figure 5.4 and data not shown). *Srp54* is an SR protein which, like *RnpS1* and *SRm160*, has also been previously characterized as a regulator of splicing [39]. Moreover, *Srp54* has been found to interact with and function in conjunction to *RnpS1* in modulating alternative splicing [36]. Pnn (pinin), another *RnpS1* interacting factor identified in the same study, did not seem to have a significant impact on MAPK levels (Figure 5.3). However, hkl (a homolog of Acinus), another factor known to function with *RnpS1* as part of the apoptosis- and splicing-associated protein complex [40], caused a slight reduction in MAPK levels (Figure 5.3). Importantly, combining either *Srp54* or *hkl* dsRNA with *RnpS1* dsRNA did not produce a synergistic effect on MAPK levels (Figure 5.3). Further investigation would be needed to validate the involvement of either of these factors in the EJC's splicing function.

5.2.5.3 CG34334 (*RnpS1*-like)

Interestingly, CG34334, an *RnpS1*-like gene, had no effect on MAPK levels on its own, but when combined to *RnpS1* dsRNA, potentiated the depletion of MAPK (Figure 5.3). Of all the SR factors we examined by RNAi, it was the only candidate to display this property. CG34334 is a potential paralog of *RnpS1* – mainly based on its RRM domain being most similar to RNPS1's – which suggests that it may be redundant with RNPS1 with regards to regulating *mapk* splicing. The fact that RNAi targeting CG34334 produces a reduction in MAPK levels only when combined to *RnpS1* indicates that CG34334 can compensate for the lack of *RnpS1* but that, in the presence of *RnpS1*, it is not otherwise required. Interestingly, there seems to be no human equivalent of CG34334; *Drosophila RnpS1* and CG34334 are both predicted orthologs of human RNPS1 (as per the HGNC's Comparison of Orthology Prediction tool [41]). Clear CG34334 equivalents are limited to the other drosophilid species, indicating that the presence of such an *RnpS1* paralog may be a particularity of this family.

5.2.6 Speculation on the EJC's Function in Splicing

Past research widely characterized the EJC as a factor involved in post-splicing functions. However, experiments where the EJC was shown not to be required for splicing were based on assays where transcripts with short introns were used [42, 43]. This presumably precluded the detection of any splicing defects similar to those we observed in transcripts with long introns. Moreover, components of the EJC have been shown to bind to the spliceosome during the early stages of spliceosome assembly which would allow for a function in splicing [2, 9]. The question still remains as to how the EJC exerts its effect on splicing and how this relates to the splicing of long introns in particular. One model would be that the EJC associates with spliceosomal components and that it is part of the components that are involved in exon bridging that occurs in exon definition. In the absence of the EJC, exons would be more poorly defined leading to more exon skipping (Figure 5.5). This would explain why a large number of splice isoforms involving different exon skipping events and sometimes combinations of exons skipping events, were observed in EJC knockdown samples.

Another possibility is that the EJC, once deposited on exon junction sites, is involved in stabilizing the exon bridging interactions of larger segments formed when two or more exons have been spliced together in intermediate splice forms. Exon definition requires exons to be relatively short, presumably in order for the spliceosomal components binding the 3'SS and 5'SS to contact each other; the EJC would thus permit the extension of the interaction of the 3'SS to the 5'SS across the larger exon formed when multiple exons are joined together (Figure 5.5). In the absence of the EJC, such intermediates would be poorly defined and would thus be more likely to be skipped. In the case of the *mapk* splicing in EJC depleted cells, this would provide an elegant explanation as to why the frequency of transcript isoforms lacking multiple consecutive exons was increased whereas the frequency of the isoform lacking only the second exon remained relatively unchanged compared to the control sample (Figure 5.5). Recent work suggests that exon definition in multi-exon splice intermediates is maintained by spliceosomal components remaining bound to the splice junction sites rather than by the EJC [43]. However, the splicing assay used in this study was a "hybrid" transcript with one short intron and a larger one slightly over the 250nt exon definition limit. While transcript levels were unchanged by EJC depletion, the fact that one of the two introns is spliced by intron definition may preclude any

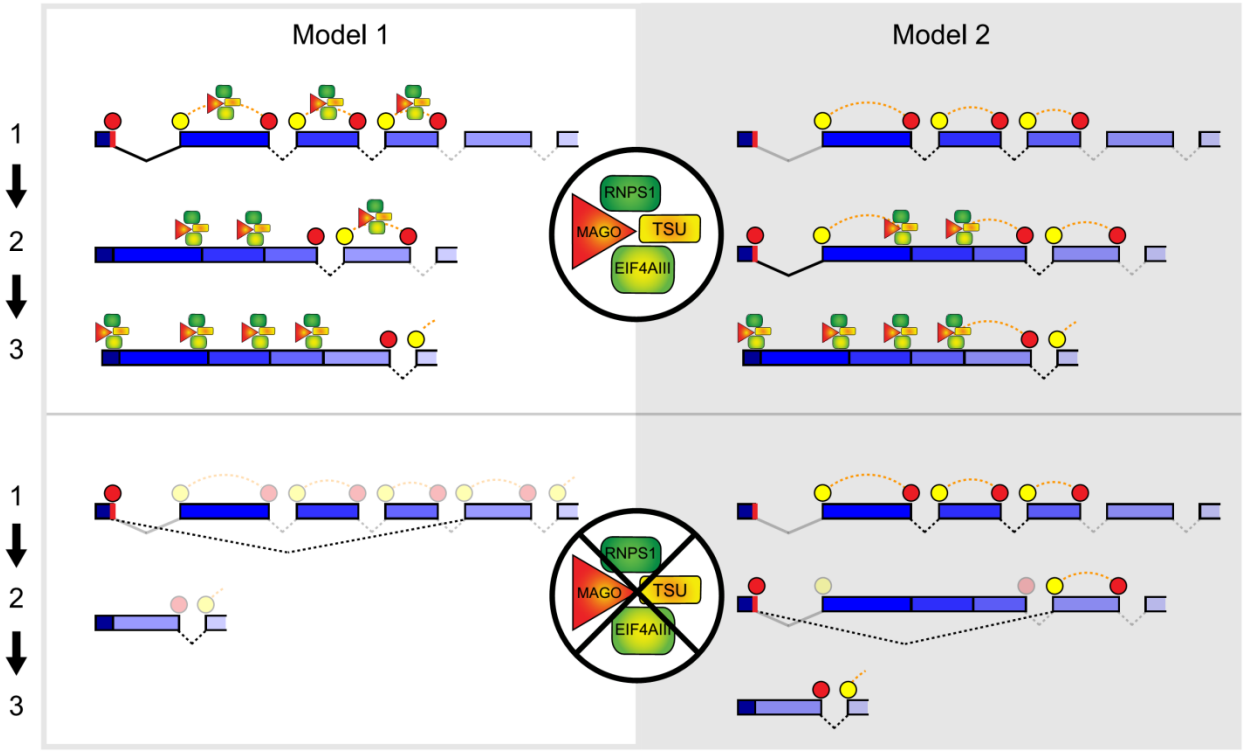
exon skipping in such a case. In light of this, it would be interesting to verify the importance of the EJC in assays involving transcripts with at least one internal exon bordered by two long introns. Also, increasing intron length may be required in order to effectively detect any splicing defects.

Both Treisman's and our studies show that the EJC is involved in regulating splicing in *Drosophila*. It remains to be demonstrated that this function is conserved in other species. In particular, if the EJC is involved in splicing in humans – where the majority of introns are long [30, 44] – it will be interesting to see if intron size remains a determining factor. If so, the EJC may have a more widespread function in modulating splicing in mammals than in *Drosophila* where long introns are more uncommon. However, given the larger complement of SR and hnRNP factors present in vertebrates [45], it may be that the EJC plays a less predominant role in regulation of splicing simply because these additional factors are also involved. Finally, while we demonstrate that absence of the EJC causes splicing changes to occur, it still remains to be established if the EJC's splicing function is open to modulation. It has been shown that mTOR signaling can modulate the EJC's translation function via the EJC co-factor SKAR [46]. SR and hnRNP factors in general are known to be modulated by different cellular stimuli, causing changes in alternative splicing [4, 47]. CK2 is known to phosphorylate RNPS1 and thus produce changes in alternative splicing [48]. Thus, EJC cofactors may be the sites of regulatory input that influence EJC function. Alternatively, the EJC may function as a basal component required for the splicing of long introns and its function may not be regulated in normal physiological contexts, but may be altered in certain diseased states and have important consequences in these cases. Further investigation will be needed to explore these possibilities.

Figure 5.5 Model of the EJC's Function In the Splicing of Long Introns

Two different models are presented to explain how the EJC might specifically impact the splicing of long introns in particular. The spliceosomal U1 (red circles) and U2 (yellow circles) complexes are represented binding to the 5'SS and 3'SS of the *mapk* transcript as splicing progresses. Exon definition interactions between U1 and U2 complexes are represented by orange dotted lines. Introns are represented by grey lines and those not drawn to scale are represented by dotted lines. Introns primed to be spliced out are darker.

In model 1, the EJC physically associates to the spliceosomal complex that forms during exon definition. The EJC serves to bolster the interactions necessary for exon definition and thus allows for exons bordered by large introns to be efficiently recognized and spliced out. This model is supported by the observation that EJC components bind to the spliceosome during the early stages of the splicing process [2, 9]. Without the EJC, some exons definition does not proceed as efficiently resulting in random exon skipping. In model 2, the EJC, deposited on splice junctions after splicing occurs, serves as an extension to allow exon definition to reach across the larger exon formed by multiple joined exons. Exon definition requires that exon lengths be short, presumably so that U1 and U2 complexes can come into contact with each other. In the larger exons formed in splice intermediates where multiple consecutive exons are joined together, the EJC thus would serve as a platform to maintain exon definition and loss of exon definition. In the absence of the EJC, there is a loss of exon definition of the larger exons formed in splice intermediates resulting in skipping of two or more exons. This model explains the increased abundance of the *mapk* ΔII-III, ΔII-IV and ΔII-V isoforms (where multiple consecutive exons are skipped) we observed. The weaker 5'SS on the first *mapk* exon may cause the use of this site to be delayed – it is not used at all in some *mapk* isoforms – allowing for downstream introns to be spliced out beforehand.


Model 1
Model 2
1
2
3
RNPS1
MAGO
TSU
EIF4AIII
1
2
3
RNPS1
MAGO
TSU
EIF4AIII

5.2.7 Additional Experimental Procedures

The additional transcriptome data presented here was based on the RNA-seq experiment initially described in [22].

5.2.7.1 Quantitative Microscopy

Quantification of MAPK levels described in Figure 5.3 was carried out according to a similar procedure to the one described in [22] with the following changes to the procedure: Cells were plated in 384 well plates at a concentration of $2x10^4$ cells/well in 25uL Schneider S2 medium and incubated for six days with dsRNA (10µg/ml or 20µg/ml for the combined dsRNA treatments). Anti-MAPK (1/2000, Cell Signaling #4695) was used to detect MAPK levels.

5.2.7.2 RT-PCR

RNA extraction and RT-PCR was carried as described in [22]. The RT-PCR primers used to detect the *Src42A* control were: 5'-GCAGAGAGTCAGTGAAGAGG and 3'-AGGGTCTCGGCATATGAAGG.

5.2.7.3 Splice Site Evaluation

The information content of the *mapk* 5'SS was assessed by comparing them to the consensus derived from the 5'SS splice sites in the *Drosophila* genome (R5.27) using the methodology described in [49]. An individual information weight matrix $R_{iw}(b,l)$ was first derived for the entire set of aligned splice sites:

$$R_{iw}(b,l) = H_{nb} + log_2 f(b,l) \quad (1)$$

where $f(b,l)$ is the frequency of each base b at position l in the aligned splice site sequence. The $R_{iw}(b,l)$ matrix is shown in Table 2. H_{nb} is the uncertainty calculated as in [50] and is obtained as follows:

$$H_{nb} = -\sum_{b}^{B}\left(\frac{nb}{n}\right) log_2\left(\frac{nb}{n}\right) \quad (2)$$

where n is the number of nucleotides in the genome and b is the specific base residue within B = { A,C,G,T } such that nA, nC, cG, nT are the number of A, C, G and T residues. The calculated ratios were (nA/n = nT/n = 0.289) and (nC/n = nG/n = 0.211). Thus the calculated H_{nb} for the *Drosophila* genome is 1.98 bits, or slightly lower than the maximal uncertainty of 2 bits that corresponds to equal probability of the occurrence of all four bases.

The individual information (R_i) of splice sites was then calculated as a dot product of the weight matrix derived in (1) and of the matrix formed by the individual splice site sequences:

$$R_i(j) = \sum_{l} \sum_{b}^{B} s(b,l,j) R_{iw}(b,l) \quad (3)$$

where $s(b,l,j)$ is the individual sequence matrix for each base b at position l in the aligned splice site sequence such that $s(b,l,j)$ is 1 when position l in splice site j is b and 0 otherwise. For example, the rl:16 5'SS (AUGGUGGGAA) is assigned 0.53 + (-0.62) + 1.34 + 1.98 + 1.97 + 0.43 + (-1.28) + 1.71 + (-0.92) + 0.49 = 5.64 bits per site (Table 2).

5.2.8 Acknowledgements

We would like to thank Hugo Lavoie and Chris Udell for critical reading of the manuscript as well as for many invaluable discussions. We are also greatful to Sebastien Lemieux for providing advice on the calculation of the information content of individual splice sites. D.A.B. is the recipient of a Frederick Banting and Charles Best Canada Doctoral Scholarship. M.T. is the recipient of a Tier II Canada Research Chair in Intracellular Signaling. This work was supported by the Canadian Cancer Society and by the Canadian Institutes for Health Research.

5.2.9 References

1. Herold N, Will CL, Wolf E, Kastner B, Urlaub H, Luhrmann R. Conservation of the protein composition and electron microscopy structure of *Drosophila* melanogaster and human spliceosomal complexes. Mol Cell Biol 2009; 29:281-301,18981222.
2. Bessonov S, Anokhina M, Will CL, Urlaub H, Luhrmann R. Isolation of an active step I spliceosome and composition of its RNP core. Nature 2008; 452:846-50,18322460.
3. Chen YI, Moore RE, Ge HY, Young MK, Lee TD, Stevens SW. Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. Nucleic Acids Res 2007; 35:3928-44,17537823.
4. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. Biochem J 2009; 417:15-27
5. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol 2009,19773805.
6. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature 2010; 463:457-63,20110989.
7. Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, et al. The developmental transcriptome of *Drosophila* melanogaster. Nature 2010,21179090.
8. Moore MJ, Proudfoot NJ. Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 2009; 136:688-700,19239889.
9. Gehring NH, Lamprinaki S, Hentze MW, Kulozik AE. The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. PLoS Biol 2009; 7:e1000120,19478851.
10. Le Hir H, Izaurralde E, Maquat LE, Moore MJ. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 2000; 19:6860-9
11. Hachet O, Ephrussi A. Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. Nature 2004; 428:959-63,15118729.
12. Le Hir H, Gatfield D, Braun IC, Forler D, Izaurralde E. The protein Mago provides a link between splicing and mRNA localization. EMBO Rep 2001; 2:1119-24

13. Palacios IM, Gatfield D, St Johnston D, Izaurralde E. An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature 2004; 427:753-7,14973490.
14. Diem MD, Chan CC, Younis I, Dreyfuss G. PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs. Nat Struct Mol Biol 2007; 14:1173-9
15. Nott A, Le Hir H, Moore MJ. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev 2004; 18:210-22
16. Wiegand HL, Lu S, Cullen BR. Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc Natl Acad Sci USA 2003; 100:11327-32
17. Gatfield D, Unterholzner L, Ciccarelli F, Bork P, Izaurralde E. Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. Embo J 2003; 22:3960 - 70,doi:10.1093/emboj/cdg371.
18. Gehring NH, Neu-Yilik G, Schell T, Hentze MW, Kulozik AE. Y14 and hUpf3b Form an NMD-Activating Complex. Molecular Cell 2003; 11:939-49
19. Lykke-Andersen J, Shu MD, Steitz JA. Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science 2001; 293:1836-9,11546874.
20. Sauliere J, Haque N, Harms S, Barbosa I, Blanchette M, Le Hir H. The exon junction complex differentially marks spliced junctions. Nat Struct Mol Biol 2010; advance online publication
21. Roignant JY, Treisman JE. Exon junction complex subunits are required to splice *Drosophila* MAP kinase, a large heterochromatic gene. Cell 2010; 143:238-50,20946982.
22. Ashton-Beaucage D, Udell CM, Lavoie H, Baril C, Lefrancois M, Chagnon P, et al. The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in *Drosophila*. Cell 2010; 143:251-62,20946983.
23. Janody F, Lee JD, Jahren N, Hazelett DJ, Benlali A, Miura GI, et al. A mosaic genetic screen reveals distinct roles for trithorax and polycomb group genes in *Drosophila* eye development. Genetics 2004; 166:187-200,15020417.

24. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. Nat Struct Mol Biol 2009; 16:990-5,19684600.
25. Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcarcel J, et al. Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol 2009; 16:996-1001,19684599.
26. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res 2009; 19:1732-41,19687145.
27. Eberl DF, Duyf BJ, Hilliker AJ. The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of *Drosophila* melanogaster. Genetics 1993; 134:277-92,8514136.
28. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, et al. Comprehensive analysis of the chromatin landscape in *Drosophila* melanogaster. Nature 2010,21179089.
29. Smith CD, Shu S, Mungall CJ, Karpen GH. The Release 5.1 Annotation of *Drosophila* melanogaster Heterochromatin. Science 2007; 316:1586-91
30. Fox-Walsh KL, Dou Y, Lam BJ, Hung SP, Baldi PF, Hertel KJ. The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proc Natl Acad Sci U S A 2005; 102:16176-81,16260721.
31. Sterner DA, Carlo T, Berget SM. Architectural limits on split genes. Proc Natl Acad Sci U S A 1996; 93:15081-5,8986767.
32. Nasim FH, Spears PA, Hoffmann HM, Kuo HC, Grabowski PJ. A Sequential splicing mechanism promotes selection of an optimal exon by repositioning a downstream 5' splice site in preprotachykinin pre-mRNA. Genes Dev 1990; 4:1172-84,2210374.
33. Robberson BL, Cote GJ, Berget SM. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol Cell Biol 1990; 10:84-94
34. Kim E, Magen A, Ast G. Different levels of alternative splicing among eukaryotes. Nucl Acids Res 2007; 35:125-31
35. Mayeda A, Badolato J, Kobayashi R, Zhang MQ, Gardiner EM, Krainer AR. Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing. EMBO J 1999; 18:4560-70,10449421.

36. Sakashita E, Tatsumi S, Werner D, Endo H, Mayeda A. Human RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator in vivo. Mol Cell Biol 2004; 24:1174-87,14729963.
37. Blencowe BJ, Issner R, Nickerson JA, Sharp PA. A coactivator of pre-mRNA splicing. Genes & Development 1998; 12:996-1009
38. Blencowe BJ, Baurén G, Eldridge AG, Issner R, Nickerson JA, Rosonina E, et al. The SRm160/300 splicing coactivator subunits. RNA 2000; 6:111-20
39. Zhang WJ, Wu JY. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. Mol Cell Biol 1996; 16:5400-8,8816452.
40. Schwerk C, Prasad J, Degenhardt K, Erdjument-Bromage H, White E, Tempst P, et al. ASAP, a novel protein complex involved in RNA processing and apoptosis. Mol Cell Biol 2003; 23:2981-90,12665594.
41. Eyre TA, Wright MW, Lush MJ, Bruford EA. HCOP: a searchable database of human orthology predictions. Brief Bioinform 2007; 8:2-5,16951416.
42. Zhang Z, Krainer AR. Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. Proc Natl Acad Sci U S A 2007; 104:11574-9,17606899.
43. Crabb TL, Lam BJ, Hertel KJ. Retention of spliceosomal components along ligated exons ensures efficient removal of multiple introns. RNA 2010:-
44. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409:860-921,11237011.
45. Barbosa-Morais NL, Carmo-Fonseca M, Aparicio S. Systematic genome-wide annotation of spliceosomal proteins reveals differential gene family expansion. Genome Res 2006; 16:66-77,16344558.
46. Ma XM, Yoon SO, Richardson CJ, Julich K, Blenis J. SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. Cell 2008; 133:303-13,18423201.
47. Stamm S. Regulation of alternative splicing by reversible protein phosphorylation. J Biol Chem 2008; 283:1223-7,18024427.

48. Trembley JH, Tatsumi S, Sakashita E, Loyer P, Slaughter CA, Suzuki H, et al. Activation of pre-mRNA splicing by human RNPS1 is regulated by CK2 phosphorylation. Mol Cell Biol 2005; 25:1446-57,15684395.

49. Schneider TD. Information content of individual genetic sequences. J Theor Biol 1997; 189:427-41,9446751.

50. Schneider TD, Stormo GD, Gold L, Ehrenfeucht A. Information content of binding sites on nucleotide sequences. J Mol Biol 1986; 188:415-31,3525846.

51. Mount SM, Burks C, Hertz G, Stormo GD, White O, Fields C. Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. Nucleic Acids Res 1992; 20:4255-62,1508718.

5.3 Avancées récentes et conséquences sur le modèle de l'EJC en tant que facteur d'épissage alternatif

Depuis la parution de notre article portant sur le complexe EJC en 2010 et du commentaire présenté dans la section précédente, plusieurs travaux ont confirmé et approfondi nos connaissances sur le rôle de l'EJC dans l'épissage alternatif. Cette section discute des implications de ces découvertes plus récentes.

5.3.1 Données additionnelles confirmant le rôle de l'EJC dans l'épissage alternatif

Une des questions qui demeuraient ouvertes suite à l'identification du rôle de l'EJC dans l'épissage chez la mouche était à savoir si cette fonction était conservée dans d'autres espèces. Une réponse à cette question fut apportée lorsqu'il fut documenté que des composantes de l'EJC peuvent influencer l'épissage alternatif de *Bcl-x* ainsi que d'autres facteurs liés à l'apoptose (Michelle, Cloutier et al. 2012). D'autres données de cette étude allaient de pair avec les nôtres et celles du groupe de Jessica Treissman. Il y est en effet décrit l'implication du facteur EJC périphérique, *RnpS1*, dans la modulation de l'épissage et, à l'inverse, l'absence d'implication de certains facteurs connus comme étant associés aux rôles post-épissage de l'EJC. En plus de l'EJC et de *RnpS1* deux autres facteurs de l'EJC périphérique, Acinus et SAP18, y sont également rapportés comme altérant l'épissage de *Bcl-x*. Pour ce qui est de RnpS1, les données obtenues par Michelle et al. indiquent cependant que ce facteur se lie à des segments d'ARN distincts de ceux reconnus par les facteurs centraux de l'EJC. Ces données suggèrent donc que RnpS1 pourrait fonctionner de manière indépendante au complexe principal ou qu'il existe différents sous-complexes.

Si l'EJC fonctionnait bel et bien comme régulateur de l'épissage dans d'autres espèces, il restait désormais à savoir si la longueur des introns est un facteur déterminant dans ce type de régulation. Déjà lors, l'étude de Michelle et al. semblait indiquer que la longueur des introns n'était pas un facteur déterminant chez les mammifères dans le cas de l'épissage de *Bcl-x* (Michelle, Cloutier et al. 2012). Suite à cette étude, deux publications subséquentes sont venues approfondir notre compréhension du fonctionnement de l'EJC dans l'épissage en

utilisant le gène de mouche *piwi* comme point de départ. La première étude démontre que l'excision du quatrième intron du gène *piwi* est dépendante de l'EJC et de son co-facteur RNPS1 (Malone, Mestdagh et al. 2014). Par contre, l'intron en question est petit et il semble plutôt que ce soit la présence d'une séquence poly-pyrimidine faible qui détermine la sensibilité de *piwi* à l'EJC. Tout comme l'étude de Michelle et al., ce groupe démontre que le facteur de l'EJC périphérique, *Acinus*, est également impliqué. La deuxième étude récente menée chez la mouche porte également sur l'épissage de l'intron court de *piwi*, mais confirme en même temps le lien entre la longueur des introns et la modulation EJC-dépendante de l'épissage alternatif (Hayashi, Handler et al. 2014). En effet, elle réconcilie ces deux observations en faisant la distinction entre la régulation de la rétention d'intron et celle du saut d'exon. Leurs données de séquençage à haut débit démontrent que le saut d'exon serait bel et bien associé à la longueur des introns alors que la rétention EJC-dépendante d'intron serait plutôt associée à un site d'épissage faible et à la présence d'un intron long en aval. Les auteurs du deuxième article observent également l'implication d'*Acinus* dans la rétention d'intron et y ajoutent la distinction qu'*Acinus* n'est peu, voire nullement, associé au saut d'exon associé aux longs introns. Ce constat est en accord avec nos propres données qui montrent que la rétention du premier intron de *mapk*, un intron court, est modulée par l'EJC (Figure 5.6). Dans le article de Hayashi et al., bien que les auteurs ne l'aient pas indiqué expressément, on peut voir qu'*Acinus* semble causer la rétention du premier intron du transcrit *mapk*-RB, ce qui mène à un enrichissement de la forme *mapk*-RD dans les cellules souches somatiques ovariennes de drosophile. En somme, il semblerait donc que la longueur des introns soit associée à au moins une partie des événements d'épissage modulés par l'EJC.

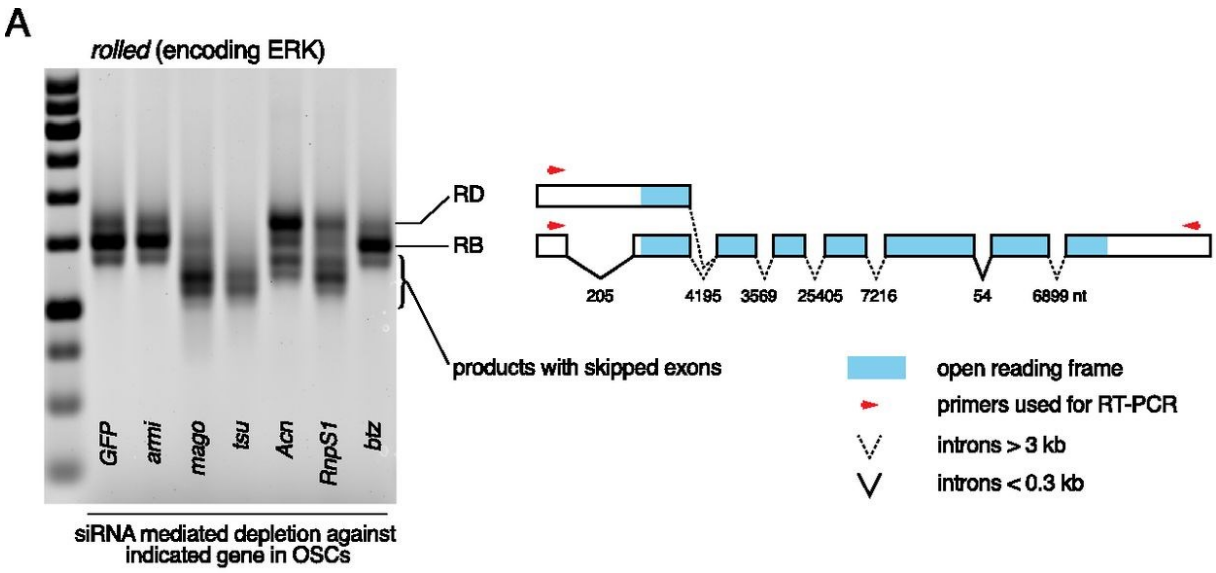


Figure 5.6 Modulation de la rétention du premier intron de *mapk*

Confirmation des effets de l'EJC sur l'épissage de *mapk* dans les cellules souches somatiques ovariennes de drosophile. Le facteur de l'EJC périphérique, *Acinus*, qui semble avoir peu d'impact sur le saut d'exon, semble par contre causer une rétention accrue du premier intron. Figure tirée de Hayashi et al. 2014 avec la permission de *Cold Spring Harbor Laboratory Press*.

Plus récemment, la parution d'une étude globale de l'épissage dépendant de l'EJC est venue ajouter des informations importantes quant au rôle de l'EJC chez les mammifères. Cette étude fait usage du séquençage à haut débit couplé à des donnés de « cross-linking immunoprecipitation » (CLIP) et des analyses d'épissage alternatif pour offrir un portrait global de l'impact sur l'épissage EJC-dépendant dans des cellules HeLa (Wang, Murigneux et al. 2014). Les auteurs ne constatent toutefois pas le même impact de l'EJC que nous sur l'abondance des transcrits à longs introns. Par contre, en examinant l'épissage alternatif des exons cassette, ils observent que l'épissage de ceux-ci est plus fortement altéré lorsqu'ils sont flanqués de longs introns. De plus, ils décrivent également plusieurs événements de saut d'exons constitutifs, qui ne sont pas habituellement connus comme étant associés à un épissage alternatif, ce qui n'est pas sans rappeler nos propres observations. Pour ce qui est de la rétention d'intron modulée par l'EJC, comme chez la drosophile, celle-ci aurait plutôt lieu dans le cas des introns courts possédant des sites d'épissage à consensus plus faibles.

Toujours chez les mammifères, le groupe de Benjamin Blencowe a récemment rapporté qu'un facteur d'épissage alternatif, MBNL, joue un rôle important à la différentiation des cellules souches embryonnaires (Han, Irimia et al. 2013). Un criblage subséquent visant à l'identification de facteurs d'épissage additionnels impliqués dans ce tissu a récemment identifié des composantes de l'EJC (Benjamin Blencowe, communication personnelle). De plus, l'EJC est rapporté comme induisant des événements d'épissage alternatif dans des gènes reliés à l'apoptose et la prolifération au même titre que des facteurs associés au spliceosome dans les cellules HeLa (Papasaikas, Tejedor et al. 2014). Le rôle de l'EJC dans l'épissage alternatif semble donc conservé à travers l'évolution. De plus, tout indique que ce rôle inclurait au moins deux modes de régulation distincts, dont l'un serait associé à la longueur des introns et probablement responsable de la modulation de *mapk*.

5.3.2 Régulation de MAPK par l'EJC

Depuis la parution de notre article rapportant la régulation de *mapk* par l'EJC, deux autres groupes ont publié des données décrivant spécifiquement l'impact de l'EJC sur MAPK. L'article de Hayashi et al. confirme les observations en faisant état de la modulation de

l'épissage de *mapk* dans des cellules ovariennes de drosophile ((Hayashi, Handler et al. 2014) et Figure 5.6), ce qui assoit bien le rôle de l'EJC en tant que facteur d'épissage alternatif chez la mouche. La question de la modulation de MAPK chez d'autres espèces a été examinée chez *Xenopus laevis* suite à une déplétion de EIF4AIII (Haremaki and Weinstein 2012). Cette étude rapporte un effet sur l'épissage du transcrit du gène *ryr* et un impact associé sur la fonction musculaire de l'embryon. Par contre, aucun effet n'est observé sur l'épissage du plus long intron de MAPK (19 kb). Les auteurs ont subséquemment utilisé un essai RT-PCR couvrant le transcrit entier et n'ont pas observé de modulation de MAPK suite à la déplétion de Eif4a3 (Daniel Weinstein, communication personnelle). Si l'impact de l'EJC sur MAPK ne semble pas être conservé chez *Xenopus*, la question demeure toutefois ouverte chez d'autres espèces de vertébrés.

5.3.3 Mécanismes d'action potentiels de l'EJC sur l'épissage alternatif

La présente section traitera des avancées récentes portant sur notre compréhension du mécanisme d'action de l'EJC en tant que facteur d'épissage. L'article figurant à la section 6.4 décrit deux hypothèses de mécanismes d'action potentiels qui étaient envisagés lors de la publication initiale de nos travaux. Or, il se trouve que les études parues depuis permettent de raffiner les deux modèles proposés et d'y ajouter un troisième modèle.

5.3.3.1 L'EJC en tant que co-facteur associé au spliceosome

L'association de l'EJC au pré-ARNm suite à l'épissage est très bien documentée. Par contre, tel qu'indiqué précédemment, l'association de l'EJC au spliceosome débute avant cette étape, alors que le processus d'épissage n'a pas encore pris fin (Gehring, Lamprinaki et al. 2009). Il serait donc aisé de penser que cette association lors du processus d'épissage influencerait le choix du site d'épissage. L'EJC jouerait ainsi un rôle clé dans la définition des exons, particulièrement importante en présence de longs introns (Figure 5.5, modèle 1). Récemment, deux nouvelles publications ont détaillé le rôle essentiel du facteur d'épissage CWC22 dans l'association entre l'EJC et le spliceosome (Barbosa, Haque et al. 2012, Steckelberg, Boehm et al. 2012). Ces articles montrent qu'EIF4AIII entre tout d'abord en

contact avec le spliceosome et permet ensuite l'association des deux autres composantes de l'EJC avec le spliceosome. S'il n'y a toujours pas de démonstration que l'EJC agit sur le spliceosome lors de son recrutement, ces études clarifient toutefois ce processus et indiquent que CWC22 serait sans doute une cible d'intérêt pour de futures études de l'épissage EJC-dépendant. D'ailleurs, l'homologue de CWC22 chez la drosophile, *ncm*, causait une réduction des niveaux de pMAPK dans notre crible RAS^{V12}, mais n'a pas été retenu, n'étant que légèrement au-dessus du seuil de sélection.

5.3.3.2 L'EJC en tant que facteur d'épissage classique

Des études de génomique à large échelle basées sur le CLIP ont démontré que l'EJC n'était pas distribué sur les ARNm épissés de manière uniforme tel qu'envisagé précédemment (Sauliere, Murigneux et al. 2012, Singh, Kucukural et al. 2012). En fait, ces travaux démontrent plutôt une distribution non-uniformes avec une prépondérance dans certaines familles de transcrits et un enrichissement sur certains sites non-canoniques présentant des motifs caractéristiques de la liaison des facteurs d'épissage riches en serine/arginine (SR). Ces observations suggèrent donc que l'EJC pourrait lier l'ARN, soit directement, soit à travers un de ses facteurs associés et ainsi agir tel un facteur d'épissage alternatif classique en aidant au recrutement du spliceosome. Ce modèle est également privilégié par Michelle et al. dans leur étude de l'épissage de *Bcl-x* où il est d'ailleurs démontré que la régulation de l'épissage par l'EJC passe par la liaison à des sites non-canoniques (Michelle, Cloutier et al. 2012). Par contre, les données de CLIP tendent, de manière générale, à montrer que l'EJC est surtout enrichi dans des positions canoniques sur les exons et que sa fonction plus large dans l'épissage serait donc généralement associée aux sites de liaison canonique typiquement reliés à ses autres rôles (Sauliere, Murigneux et al. 2012, Wang, Murigneux et al. 2014). L'existence de complexes mRNP d'ordre supérieur, incluant de multiples EJCs et facteurs SR, rapportée dans Singh et al., suggèrent que plusieurs facteurs pourraient agir de pair pour lier l'ARN et/ou l'EJC et influencer l'épissage ensemble (Singh, Kucukural et al. 2012). Cette étude indique également que les présences d'EJC sur des sites canoniques et non-canoniques sont fortement associées. Il semble donc difficile de découpler l'apport de ces multiples facteurs et

points de liaison dans la modulation de l'épissage. Le recours à des techniques d'interaction in vitro sera sans doute nécessaire pour découpler ces différents apports et à l'identification d'un site de liaison non-canonique propre à EIF4AIII.

5.3.3.3 Modulation des événements d'épissage adjacents par l'EJC déposé sur les jonctions exon-exon

Suite à l'épissage d'un intron, les deux exons sont joints pour former un exon plus long. Ce nouvel exon doit ensuite continuer à être défini correctement pour permettre l'épissage et la jonction avec l'exon suivant. Puisque l'EJC est déposé aux jonctions exon-exon, il est tentant d'envisager qu'il agisse pour cimenter la définition des exons ainsi joints et promouvoir l'utilisation des sites d'épissage de part et d'autre (Figure 5.5, modèle 2).

L'idée que la déposition de l'EJC sur son site canonique, c'est à dire aux jonctions exon-exon, influencerait l'épissage en amont ou en aval a été explorée précédemment et a été rejetée dans deux études initiales. D'abord, EIF4AIII ne semblait pas être requis pour l'épissage de pre-ARNm in vitro (Zhang and Krainer 2007). Une autre étude plus détaillée avait également démontré que la déposition de l'EJC sur une jonction exon-exon n'a pas d'effet sur l'épissage des introns adjacents et cet effet serait plutôt attribuable à la rétention de composantes du spliceosome (Crabb, Lam et al. 2010).

Suite à notre publication, cette idée a été explorée davantage dans le domaine. D'abord, les deux études basées sur le gène *piwi* (voir section 5.3.1) montrent un rôle clair de l'EJC dans l'épissage alternatif de son quatrième intron. Dans les deux cas, on évoque un mécanisme nécessitant la déposition préalable de l'EJC sur un site canonique (jonction exon-exon) pour expliquer la modulation de l'épissage de *piwi*. La première étude montre en effet que l'épissage du quatrième intron dépend de l'épissage préalable des introns voisins et donc probablement de la déposition de l'EJC sur ces jonctions épissées (Malone, Mestdagh et al. 2014). La deuxième étude rapporte que la déposition du complexe EJC sur une jonction exon-exon a une incidence sur l'épissage alternatif se produisant en amont et en aval dans le cas de *piwi* ainsi que d'autres gènes examinés (Hayashi, Handler et al. 2014). Dans leur discussion, les auteurs citent notre modèle de mécanisme de prolongation de la définition des exons et

évoquent l'idée que ce mécanisme opère dans le cas des sauts d'exons multiples observés dans leur étude ainsi que dans la nôtre. Par contre, ils ajoutent une distinction en ce qui concerne la rétention d'intron; dans ce cas, la déposition de l'EJC sur un site à proximité (à moins de 250 nt) permettrait de promouvoir l'épissage. Finalement, ils évoquent également un mécanisme de temporalité où l'ordre des événements d'épissage – et de la déposition de l'EJC associée aux jonctions exoniques – aurait un impact sur l'épissage alternatif.

Si l'influence de la déposition de l'EJC aux jonctions exon-exon sur les événements d'épissage voisins a pu être vérifiée chez la mouche, elle ne l'a toujours pas été chez les mammifères. Les travaux récents de Wang et al. observent une augmentation des sauts d'exon et une modulation de la rétention d'intron, mais n'ont toutefois pas pu observer que ceux-ci dépendaient d'événements d'épissage à proximité (Wang, Murigneux et al. 2014). Il n'est donc pas clair que l'EJC assemblé et déposé aux jonctions exon-exon soit responsable de l'impact plus large sur l'épissage alternatif observé au cours de cette étude.

5.3.3.4 Impact de la transcription sur l'épissage EJC-dépendant

L'article de Wang et al. privilégie un autre mécanisme d'action de l'EJC. En ralentissant la transcription, les auteurs parviennent à contrecarrer les effets d'épissage alternatifs induits par la déplétion de l'EJC. De plus, ils constatent un impact de la déplétion de l'EJC sur le taux d'élongation de l'ARN (Wang, Murigneux et al. 2014). Les données récentes montrent que l'EJC est au cœur de la formation des structures mRNP complexes (Singh, Kucukural et al. 2012). Or, il est possible que l'empaquetage des ARN ait un impact sur la transcription et qu'il influence alors la disponibilité des sites d'épissage du pre-ARNm naissant.

Nous avons nous-même identifié deux régulateurs de l'ARN polymérase II dans notre crible RAS^{V12} : *Cdk12*, une kinase régulant l'ARN polymérase II par phosphorylation de sa queue C-terminale, et *Fip1*, un membre du complexe de clivage et polyadénylation des pré-ARNm. Cependant, ces deux facteurs produisaient un profil d'épissage de *mapk* différent de celui associé à l'EJC. Il n'est donc pas clair que le même type de relation entre l'EJC et la transcription soit présent chez la drosophile.

De manière plus générale, il est bien établi que l'ARN polymérase II et la transcription jouent un rôle important dans la modulation des facteurs d'épissage alternatif (Kornblihtt 2006, Phatnani and Greenleaf 2006, Rigo and Martinson 2008, Ip, Schmidt et al. 2011, Luco, Allo et al. 2011). Le complexe EJC ne fait probablement pas exception à la règle. Par contre, l'impact de l'EJC sur la transcription doit encore être exploré plus en détail. D'ailleurs, avec l'ajout d'un rôle transcriptionnel, l'EJC se retrouvera donc impliqué dans chacune des étapes de la vie des ARNm, allant de leur production jusqu'à leur dégradation.

5.3.3.5 Importance des facteurs associés à l'EJC

Un autre élément important apporté par les études récentes est la notion de la composition du complexe EJC. L'idée que la composition du complexe EJC varie et permet de préciser son rôle moléculaire est déjà bien présente en ce qui se rapporte à ses fonctions post-épissage. Par exemple, BTZ est connue comme étant centrale aux rôles post-épissage du complexe (Mohr, Dillon et al. 2001, van Eeden, Palacios et al. 2001, Hachet and Ephrussi 2004). Le co-facteur SKAR, quant à lui, est plutôt associé à sa fonction de modulateur de la traduction (Ma, Yoon et al. 2008). Dans nos propres travaux, nous avions noté une absence d'impact de BTZ sur la fonction d'épissage de l'EJC ((Ashton-Beaucage, Udell et al. 2010); chapitre 2). Nous avons également fait état de l'association de l'EJC à plusieurs facteurs d'épissage alternatifs connus (voir chapitre 2.4, discussion). Il est donc raisonnable de penser que les fonctions d'épissage alternatif de l'EJC seraient déterminées par des conformations particulières du complexe associées à ce rôle et distinctes des conformations rattachées à ses autres fonctions.

La composition du complexe EJC pourrait même expliquer les différents rôles de l'EJC dans le processus d'épissage. En effet, tel que mentionné précédemment, l'importance du facteur de l'EJC périphérique *Acinus* dans l'exclusion d'intron et, à l'inverse, l'absence d'effet sur le saut d'exon, indique probablement deux configuration distinctes de l'EJC (Hayashi, Handler et al. 2014). Ces configurations distinctes seraient d'ailleurs en mesure d'expliquer les différences observées chez les mammifères, *Acinus* régulant le choix du site d'épissage de *Bcl-x* (Michelle, Cloutier et al. 2012), mais n'ayant pas les mêmes effets larges sur l'épissage

alternatif associés aux facteurs centraux de l'EJC (Wang, Murigneux et al. 2014). Il se pourrait donc que l'EJC central agisse comme une plateforme générique d'épissage alternatif où peuvent s'associer différents régulateurs spécifiques de l'épissage alternatif.

5.4 MAPK un nœud au sein d'un ensemble de régulateurs à paliers multiples

Au chapitre 2, nous avons fait état de l'impact de l'EJC sur l'épissage de *mapk*. Le chapitre 3 y ajoute une autre série de régulateurs de l'épissage produisant des transcrits différents. De plus, nous y décrivons des facteurs agissant probablement au niveau de la transcription de *mapk*. Le chapitre 4 décrit la modulation post-traductionnelle des niveaux de MAPK par USP47. Il semblerait donc que l'expression de *mapk* soit la cible d'un ensemble de régulateurs agissant à plusieurs étapes de son expression. Le nombre et la diversité de ces régulateurs est impressionnant d'autant plus que, à l'exception des facteurs de transcription *gfzf* et *CG4936*, nous n'avons pas identifié de facteurs comparables agissant sur d'autres composantes de la voie.

L'activation de RAF est une autre étape de la voie RAS/MAPK qui implique un ensemble complexe de régulateurs post-traductionnels incluant, entre autres, des boucles de rétroaction et l'apport de signaux externes (« crosstalk ») (Roskoski 2010, Udell, Rajakulendran et al. 2011). Serait-il possible que *mapk* soit un autre nœud central de la voie RAS/MAPK, servant à intégrer des signaux de nature moins immédiats? Pour élucider cette question importante, il faudra d'abord mieux comprendre le rôle des nouveaux régulateurs de *mapk*. En effet, les mécanismes contrôlant l'activité de ces facteurs reste entièrement à définir. S'agit-il de processus ayant cours dans le cadre du développement? Ou s'agit-il simplement de facteurs basaux servant tout au plus à veiller à l'intégrité de l'expression de *mapk*? Et s'il s'agit de facteurs basaux, pourquoi *mapk* y serait-il plus sensible que d'autres gènes? La réponse à ces questions réside peut-être dans les caractéristiques qui distinguent *mapk* des autres gènes de drosophile. D'abord, il s'agit d'un gène à longs introns se situant dans

l'extrême par rapport aux autres gènes de mouche (voir chapitre 2). Ensuite, il est situé dans l'hétérochromatine péricentromérique, caractéristique qu'il partage avec très peu d'autres gènes. Dernièrement, nous avons également noté que *mapk* présentait une distribution de codons atypique pour un gène de drosophile (Figure 6.3), indiquant que sa traduction serait également plus lente ou alors sensible à l'abondance d'ARNt plus rares.

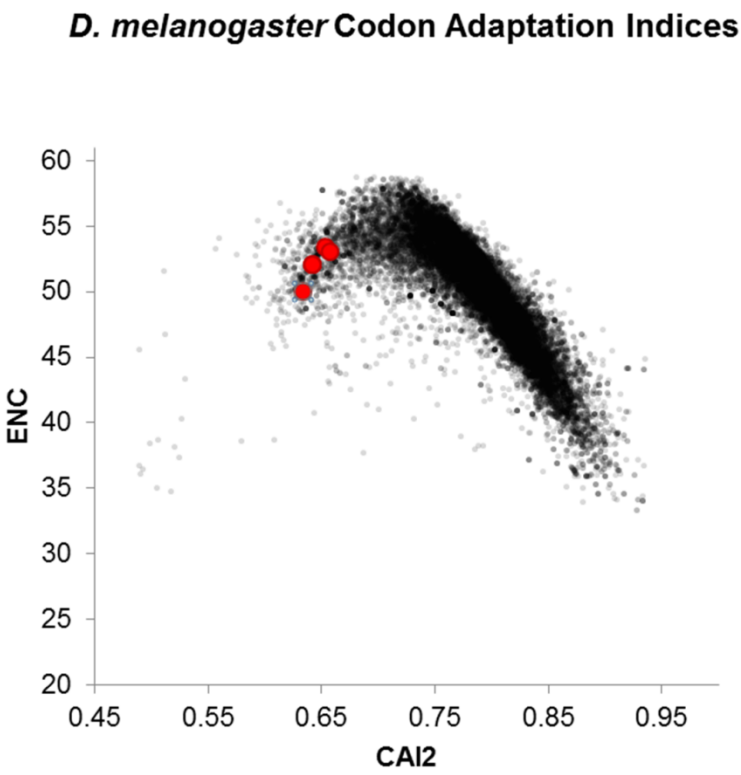


Figure 5.7 Distribution de l'indice d'adaptation des codons chez la drosophile

L'indice d'adaptation des codons (CAI; en abscisse) est comparé au nombre effectif de codons (ENC; en ordonnée) pour l'ensemble des transcrits chez *D. melanogaster*. Les transcrits *mapk* sont indiqués par les points rouges et sont parmi les indices d'adaptation les plus faibles du transcriptome. Ceci signifie qu'ils emploient un nombre élevé de codons rares chez la drosophile. Les scores CAI et ENC ont été dérivés à l'aide du logiciel DAMBE (Xia and Xie 2001)

La présence de longs introns est associée à une permissivité accrue aux changements évolutifs, prenant souvent la forme d'élimination d'un exon ou de l'intégration d'un pseudo-exon (Ben-Dov, Hartmann et al. 2008, Birzele, Csaba et al. 2008, Kandul and Noor 2009). Il est possible que *mapk* soit donc plus susceptible à de tels changements et qu'il y ait un avantage évolutif à ce que ce gène ait plus de plasticité. Par contre, les transcrits alternatifs de *mapk* détectés dans les cellules sauvages produiraient soient une protéine équivalente ou un changement de cadre de lecture qui mènerait à une dégradation via NMD due à la présence d'un site de terminaison précoce. Il en va d'ailleurs de même pour plusieurs des transcrits observés suite à la déplétion de l'EJC. Il ne semble donc pas y avoir d'émergence de formes alternatives de *mapk* productives chez la drosophile et l'épissage alternatif observé mènerait simplement à une réduction de l'expression protéique.

Une autre possibilité est que l'un ou l'autre de ces facteurs fasse partie d'un mécanisme de réponse au stress. Cette avenue est discutée au chapitre 3 où il est mentionné que les facteurs d'épissage constitutifs répondent à certaines formes de stress menant à des changements d'épissage (Biamonti and Caceres 2009). Il est possible qu'il s'agisse d'une modulation active passant par un des régulateurs de *mapk* pour moduler son expression. À l'inverse, *mapk* pourrait être un gène particulièrement sensible aux perturbations induites par le stress sur des processus tels que l'épissage ou le N-end rule et, conséquemment, parmi l'un des premiers à voir son expression diminuer. Une diminution des niveaux de MAPK pourrait ainsi aider l'organisme à survivre au stress en ralentissant son taux de prolifération par exemple.

5.5 Incidence du criblage à haut débit par ARNi sur le paradigme des voies de signalisation

Nous avons fait état des avancées dans le domaine du criblage par RNAi en faisant mention également que cette technologie n'avait probablement toujours pas atteint son plein potentiel en tant qu'outil de biologie des systèmes. Or, les différents cribles RNAi pan-génomiques menés jusqu'à présent auront tout de même permis d'accomplir des avancées importantes, notamment dans l'identification de nombre de nouveaux facteurs de voies de signalisation (Mathey-Prevot and Perrimon 2006, Neumuller and Perrimon 2011). Au-delà de son utilité en tant qu'outil de découverte de gènes, la criblage RNAi a tout de même su s'imposer comme outil de biologie des systèmes et ce, malgré le manque de standardisation toujours problématique en 2014; les quelques initiatives tentant d'établir des standards n'ayant toujours pas porté fruit ((Haney 2007) et miare.sourceforge.net).

Une des conséquences principales des données de criblage portant sur les voies de signalisation a été de favoriser la transition du modèle de voies de signalisation linéaires vers un modèle de réseau (Friedman and Perrimon 2007, Neumuller and Perrimon 2011). En effet, les données cumulées d'un répertoire grandissant de cribles RNAi remettent en question le modèle classique de voie de signalisation impliquant un assemblage hiérarchique de facteurs centraux qui fonctionnent en vase clos et transmettent un signal tel les composantes d'un circuit électrique ((Noselli and Perrimon 2000) et (Figure 5.8 A)). Ces données suggéreraient donc plutôt un modèle de réseau dense où la transmission de signaux est assurée par la collaboration de facteurs multiples à contribution inégale ((Figure 5.8 B), (Friedman and Perrimon 2007) et section 1.2 de cet ouvrage).

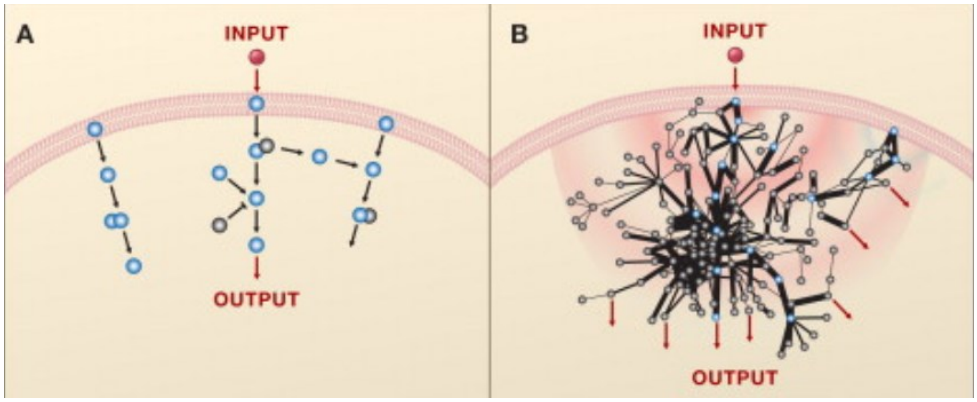


Figure 5.8 Modèles classique de voie et de réseau de signalisation

(A) Modèle classique des voies de signalisation en tant que circuits fermés. Ces modèles sont généralement dérivés d'études développementales et n'incorporent pas d'informations quantitatives. (B) Nouveaux modèle de réseau de signalisation suggéré, entre autre, par les études de criblage RNAi à large échelle. Un signal donné se propage à travers un réseau complexe de facteurs n'ayant pas tous un apport égal à la propagation du signal. De plus, le signal peut se disperser à travers un ensemble de branches qui, ensemble, auront un impact sur la réponse biologique finale. Figure tirée de Friedmand et Perrimon 2007 et reproduite avec la permission de *Elsevier*.

Notre évaluation de la signalisation RAS/MAPK tend à soutenir le modèle d'un large réseau aux points de régulation multiples tout en confirmant l'importance centrale des composantes de la voie classique. S'il est vrai que les données de cribles ARNi peuvent indiquer des modèles de réseaux de signalisation, ils ne sont toutefois pas encore en mesure de reconstituer des modèles de réseaux fonctionnels. À ce titre, des études combinées intégrant le criblage par ARNi à d'autres technologies seront certainement une voie d'avenir prometteuse (Vinayagam, Zirin et al. 2014).

D'autre part, la contribution d'autres techniques expérimentales au modèle de réseau complexe de signalisation n'est pas négligeable comme le témoigne ce catalogage des différentes branches de la voie mTOR (Caron, Ghosh et al. 2010). Les études de protéomique à large échelle contribuent également de manière importante en élargissant le répertoire des interactions et cibles enzymatiques des facteurs de signalisation connus (Kosako, Yamaguchi et al. 2009, Old, Shabb et al. 2009, Carlson, Chouinard et al. 2011, Courcelles, Fremin et al. 2013). Il semblerait donc plus juste d'avancer que les cribles ARNi participent à la refonte des modèles de signalisation de pair avec un ensemble de technologies complémentaires.

6 Bibliographie

Adams, M., S. Celniker, R. Holt, C. Evans, J. Gocayne, P. Amanatides, S. Scherer, P. Li, R. Hoskins and R. Galle (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185 - 2195.

Alexandrov, A., D. Colognori, M. D. Shu and J. A. Steitz (2012). "Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay." Proc Natl Acad Sci U S A **109**(52): 21313-21318.

Arziman, Z., T. Horn and M. Boutros (2005). "E-RNAi: a web application to design optimized RNAi constructs." Nucleic Acids Res **33**(Web Server issue): W582-588.

Ashton-Beaucage, D. and M. Therrien (2010). "[The greater RTK/RAS/ERK signalling pathway: how genetics has helped piece together a signalling network.]." Med Sci (Paris) **26**(12): 1067-1073.

Ashton-Beaucage, D., C. M. Udell, P. Gendron, M. Sahmi, M. Lefrancois, C. Baril, A. S. Guenier, J. Duchaine, D. Lamarre, S. Lemieux and M. Therrien (2014). "A functional screen reveals an extensive layer of transcriptional and splicing control underlying RAS/MAPK signaling in Drosophila." PLoS Biol **12**(3): e1001809.

Ashton-Beaucage, D., C. M. Udell, H. Lavoie, C. Baril, M. Lefrancois, P. Chagnon, P. Gendron, O. Caron-Lizotte, E. Bonneil, P. Thibault and M. Therrien (2010). "The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila." Cell **143**(2): 251-262.

Axelsson, E., T. Sandmann, T. Horn, M. Boutros, W. Huber and B. Fischer (2011). "Extracting quantitative genetic interaction phenotypes from matrix combinatorial RNAi." BMC Bioinformatics **12**: 342.

Barbier, J., M. Dutertre, D. Bittencourt, G. Sanchez, L. Gratadou, P. de la Grange and D. Auboeuf (2007). "Regulation of H-ras splice variant expression by cross talk between the p53 and nonsense-mediated mRNA decay pathways." Mol Cell Biol **27**(20): 7315-7333.

Barbosa, I., N. Haque, F. Fiorini, C. Barrandon, C. Tomasetto, M. Blanchette and H. Le Hir (2012). "Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly." Nat Struct Mol Biol **19**(10): 983-990.

Beitel, G. J., S. G. Clark and H. R. Horvitz (1990). "Caenorhabditis elegans ras gene let-60 acts as a switch in the pathway of vulval induction." Nature **348**(6301): 503-509.

Bell, M. V., A. E. Cowper, M.-P. Lefranc, J. I. Bell and G. R. Screaton (1998). "Influence of Intron Length on Alternative Splicing of CD44." Mol. Cell. Biol. **18**(10): 5930-5941.

Ben-Dov, C., B. Hartmann, J. Lundgren and J. Valcarcel (2008). "Genome-wide analysis of alternative pre-mRNA splicing." J Biol Chem **283**(3): 1229-1233.

Berget, S. (1995). "Exon Recognition in Vertebrate Splicing." J Biol Chem **270**(6): 2411 - 2414.

Bessonov, S., M. Anokhina, C. L. Will, H. Urlaub and R. Luhrmann (2008). "Isolation of an active step I spliceosome and composition of its RNP core." Nature **452**(7189): 846-850.

Biamonti, G. and J. F. Caceres (2009). "Cellular stress and RNA splicing." Trends Biochem Sci **34**(3): 146-153.

Birzele, F., G. Csaba and R. Zimmer (2008). "Alternative splicing and protein structure evolution." Nucleic Acids Res **36**(2): 550-558.

Black, D. L. (1995). "Finding splice sites within a wilderness of RNA." RNA **1**(8): 763-771.

Boutros, M., L. P. Bras and W. Huber (2006). "Analysis of cell-based RNAi screens." Genome Biol **7**(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**(5659): 832-835.

Brazma, A., P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, J. Aach, W. Ansorge, C. A. Ball, H. C. Causton, T. Gaasterland, P. Glenisson, F. C. Holstege, I. F. Kim, V. Markowitz, J. C. Matese, H. Parkinson, A. Robinson, U. Sarkans, S. Schulze-Kremer, J. Stewart, R. Taylor, J. Vilo and M. Vingron (2001). "Minimum information about a microarray experiment (MIAME)-toward standards for microarray data." Nat Genet **29**(4): 365-371.

Carlson, S. M., C. R. Chouinard, A. Labadorf, C. J. Lam, K. Schmelzle, E. Fraenkel and F. M. White (2011). "Large-scale discovery of ERK2 substrates identifies ERK-mediated transcriptional regulation by ETV3." Sci Signal **4**(196): rs11.

Caron, E., S. Ghosh, Y. Matsuoka, D. Ashton-Beaucage, M. Therrien, S. Lemieux, C. Perreault, P. P. Roux and H. Kitano (2010). "A comprehensive map of the mTOR signaling network." Mol Syst Biol **6**: 453.

Chen, F. C. (2014). "Are all of the human exons alternatively spliced?" Brief Bioinform **15**(4): 542-551.

Chen, Y. I., R. E. Moore, H. Y. Ge, M. K. Young, T. D. Lee and S. W. Stevens (2007). "Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors." Nucleic Acids Res **35**(12): 3928-3944.

Cheng, C., M. B. Yaffe and P. A. Sharp (2006). "A positive feedback loop couples Ras activation and CD44 alternative splicing." Genes Dev **20**(13): 1715-1720.

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**(12): 6499-6503.

Consortium, C. e. S. (1998). "Genome sequence of the nematode C. elegans: a platform for investigating biology." Science **282**(5396): 2012-2018.

Courcelles, M., C. Fremin, L. Voisin, S. Lemieux, S. Meloche and P. Thibault (2013). "Phosphoproteome dynamics reveal novel ERK1/2 MAP kinase substrates with broad spectrum of functions." Mol Syst Biol **9**: 669.

Crabb, T. L., B. J. Lam and K. J. Hertel (2010). "Retention of spliceosomal components along ligated exons ensures efficient removal of multiple introns." RNA: -.

Dasgupta, R., K. Nybakken, M. Booker, B. Mathey-Prevot, F. Gonsalves, B. Changkakoty and N. Perrimon (2007). "A case study of the reproducibility of transcriptional reporter cell-based RNAi screens in Drosophila." Genome Biol **8**(9): R203.

Dickson, B. J., A. van der Straten, M. Dominguez and E. Hafen (1996). "Mutations Modulating Raf signaling in Drosophila eye development." Genetics **142**(1): 163-171.

Diem, M. D., C. C. Chan, I. Younis and G. Dreyfuss (2007). "PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs." Nat Struct Mol Biol **14**(12): 1173-1179.

Dietzl, G., D. Chen, F. Schnorrer, K. C. Su, Y. Barinova, M. Fellner, B. Gasser, K. Kinsey, S. Oppel, S. Scheiblauer, A. Couto, V. Marra, K. Keleman and B. J. Dickson (2007). "A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila." Nature **448**(7150): 151-156.

Fabrizio, P., J. Dannenberg, P. Dube, B. Kastner, H. Stark, H. Urlaub and R. Luhrmann (2009). "The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome." Mol Cell **36**(4): 593-608.

Ferguson, E. L. and H. R. Horvitz (1985). "Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode Caenorhabditis elegans." Genetics **110**(1): 17-72.

Fernandez-Medarde, A. and E. Santos (2011). "Ras in cancer and developmental diseases." Genes Cancer **2**(3): 344-358.

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**(6669): 806-811.

Fox-Walsh, K. L., Y. Dou, B. J. Lam, S. P. Hung, P. F. Baldi and K. J. Hertel (2005). "The architecture of pre-mRNAs affects mechanisms of splice-site pairing." Proc Natl Acad Sci U S A **102**(45): 16176-16181.

Friedman, A. and N. Perrimon (2007). "Genetic screening for signal transduction in the era of network biology." Cell **128**(2): 225-231.

Friedman, A. A., G. Tucker, R. Singh, D. Yan, A. Vinayagam, Y. Hu, R. Binari, P. Hong, X. Sun, M. Porto, S. Pacifico, T. Murali, R. L. Finley, Jr., J. M. Asara, B. Berger and N. Perrimon (2011). "Proteomic and functional genomic landscape of receptor tyrosine kinase and ras to extracellular signal-regulated kinase signaling." Sci Signal **4**(196): rs10.

Fuchs, F., G. Pau, D. Kranz, O. Sklyar, C. Budjan, S. Steinbrink, T. Horn, A. Pedal, W. Huber and M. Boutros (2010). "Clustering phenotype populations by genome-wide RNAi and multiparametric imaging." Mol Syst Biol **6**: 370.

Gatfield, D. and E. Izaurralde (2002). "REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export." J Cell Biol **159**(4): 579-588.

Gaul, U., G. Mardon and G. M. Rubin (1992). "A putative Ras GTPase activating protein acts as a negative regulator of signaling by the Sevenless receptor tyrosine kinase." Cell **68**(6): 1007-1019.

Gehring, N. H., J. B. Kunz, G. Neu-Yilik, S. Breit, M. H. Viegas, M. W. Hentze and A. E. Kulozik (2005). "Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements." Mol Cell **20**(1): 65-75.

Gehring, N. H., S. Lamprinaki, M. W. Hentze and A. E. Kulozik (2009). "The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay." PLoS Biol **7**(5): e1000120.

Gehring, N. H., G. Neu-Yilik, T. Schell, M. W. Hentze and A. E. Kulozik (2003). "Y14 and hUpf3b Form an NMD-Activating Complex." Molecular Cell **11**(4): 939-949.

Goshima, G., R. Wollman, S. S. Goodwin, N. Zhang, J. M. Scholey, R. D. Vale and N. Stuurman (2007). "Genes required for mitotic spindle assembly in Drosophila S2 cells." Science **316**(5823): 417-421.

Grabowski, P. J., F. U. Nasim, H. C. Kuo and R. Burch (1991). "Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle." Mol Cell Biol **11**(12): 5919-5928.

Graveley, B. R., A. N. Brooks, J. W. Carlson, M. O. Duff, J. M. Landolin, L. Yang, C. G. Artieri, M. J. van Baren, N. Boley, B. W. Booth, J. B. Brown, L. Cherbas, C. A. Davis, A. Dobin, R. Li, W. Lin, J. H. Malone, N. R. Mattiuzzo, D. Miller, D. Sturgill, B. B. Tuch, C. Zaleski, D. Zhang, M. Blanchette, S. Dudoit, B. Eads, R. E. Green, A. Hammonds, L. Jiang, P. Kapranov, L. Langton, N. Perrimon, J. E. Sandler, K. H. Wan, A. Willingham, Y. Zhang, Y. Zou, J. Andrews, P. J. Bickel, S. E. Brenner, M. R. Brent, P. Cherbas, T. R. Gingeras, R. A. Hoskins, T. C. Kaufman, B. Oliver and S. E. Celniker (2010). "The developmental transcriptome of Drosophila melanogaster." Nature.

Guil, S., N. de La Iglesia, J. Fernandez-Larrea, D. Cifuentes, J. C. Ferrer, J. J. Guinovart and M. Bach-Elias (2003). "Alternative splicing of the human proto-oncogene c-H-ras renders a new Ras family protein that trafficks to cytoplasm and nucleus." Cancer Res **63**(17): 5178-5187.

Hachet, O. and A. Ephrussi (2001). "Drosophila Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport." Curr. Biol. **11**: 1666-1674.

Hachet, O. and A. Ephrussi (2004). "Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization." Nature **428**(6986): 959-963.

Hacohen, N., S. Kramer, D. Sutherland, Y. Hiromi and M. A. Krasnow (1998). "sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the Drosophila airways." Cell **92**(2): 253-263.

Han, H., M. Irimia, P. J. Ross, H. K. Sung, B. Alipanahi, L. David, A. Golipour, M. Gabut, I. P. Michael, E. N. Nachman, E. Wang, D. Trcka, T. Thompson, D. O'Hanlon, V. Slobodeniuc, N. L. Barbosa-Morais, C. B. Burge, J. Moffat, B. J. Frey, A. Nagy, J. Ellis, J. L. Wrana and B. J. Blencowe (2013). "MBNL proteins repress ES-cell-specific alternative splicing and reprogramming." Nature **498**(7453): 241-245.

Han, M. and P. W. Sternberg (1990). "let-60, a gene that specifies cell fates during C. elegans vulval induction, encodes a ras protein." Cell **63**(5): 921-931.

Haney, S. A. (2007). "Increasing the robustness and validity of RNAi screens." Pharmacogenomics **8**(8): 1037-1049.

Haremaki, T. and D. C. Weinstein (2012). "Eif4a3 is required for accurate splicing of the Xenopus laevis ryanodine receptor pre-mRNA." Dev Biol **372**(1): 103-110.

Harvey, J. J. (1964). "An Unidentified Virus Which Causes the Rapid Production of Tumours in Mice." Nature **204**: 1104-1105.

Hastings, M. L., E. Allemand, D. M. Duelli, M. P. Myers and A. R. Krainer (2007). "Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF(65)." PLoS ONE **2**(6): e538.

Hayashi, R., D. Handler, D. Ish-Horowicz and J. Brennecke (2014). "The exon junction complex is required for definition and excision of neighboring introns in Drosophila." Genes Dev **28**(16): 1772-1785.

Herold, N., C. L. Will, E. Wolf, B. Kastner, H. Urlaub and R. Luhrmann (2009). "Conservation of the protein composition and electron microscopy structure of Drosophila melanogaster and human spliceosomal complexes." Mol Cell Biol **29**(1): 281-301.

Hertel, K. J. (2008). "Combinatorial control of exon recognition." J Biol Chem **283**(3): 1211-1215.

Horn, T. and M. Boutros (2010). "E-RNAi: a web application for the multi-species design of RNAi reagents--2010 update." Nucleic Acids Res **38**(Web Server issue): W332-339.

Horn, T., T. Sandmann and M. Boutros (2010). "Design and evaluation of genome-wide libraries for RNA interference screens." Genome Biology **11**(6): R61.

Horn, T., T. Sandmann, B. Fischer, E. Axelsson, W. Huber and M. Boutros (2011). "Mapping of signaling networks through synthetic genetic interaction analysis by RNAi." Nat Methods **8**(4): 341-346.

House, A. E. and K. W. Lynch (2008). "Regulation of alternative splicing: more than just the ABCs." J Biol Chem **283**(3): 1217-1221.

Hu, Y., R. Sopko, M. Foos, C. Kelley, I. Flockhart, N. Ammeux, X. Wang, L. Perkins, N. Perrimon and S. E. Mohr (2013). "FlyPrimerBank: an online database for Drosophila melanogaster gene expression analysis and knockdown evaluation of RNAi reagents." G3 (Bethesda) **3**(9): 1607-1616.

Hurlbut, G. D., M. W. Kankel and S. Artavanis-Tsakonas (2009). "Nodal points and complexity of Notch-Ras signal integration." Proc Natl Acad Sci U S A **106**(7): 2218-2223.

Ip, J. Y., D. Schmidt, Q. Pan, A. K. Ramani, A. G. Fraser, D. T. Odom and B. J. Blencowe (2011). "Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation." Genome Res **21**(3): 390-401.

Kamata, T. and J. R. Feramisco (1984). "Epidermal growth factor stimulates guanine nucleotide binding activity and phosphorylation of ras oncogene proteins." Nature **310**(5973): 147-150.

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**(6920): 231-237.

Kandul, N. P. and M. A. Noor (2009). "Large introns in relation to alternative splicing and gene evolution: a case study of Drosophila bruno-3." BMC Genet **10**(1): 67.

Karim, F. D., H. C. Chang, M. Therrien, D. A. Wassarman, T. Laverty and G. M. Rubin (1996). "A screen for genes that function downstream of Ras1 during Drosophila eye development." Genetics **143**(1): 315-329.

Karpala, A. J., T. J. Doran and A. G. Bean (2005). "Immune responses to dsRNA: implications for gene silencing technologies." Immunol Cell Biol **83**(3): 211-216.

Kennedy, C. F., A. Kramer and S. M. Berget (1998). "A role for SRp54 during intron bridging of small introns with pyrimidine tracts upstream of the branch point." Mol Cell Biol **18**(9): 5425-5434.

Keren, H., G. Lev-Maor and G. Ast (2010). "Alternative splicing and evolution: diversification, exon definition and function." Nat Rev Genet **11**(5): 345-355.

Knudsen, T. B., G. P. Daston and S. Teratology (2005). "MIAME guidelines." Reprod Toxicol **19**(3): 263.

Kolch, W. (2005). "Coordinating ERK/MAPK signalling through scaffolds and inhibitors." Nat Rev Mol Cell Biol **6**(11): 827-837.

Kondo, S., M. Booker and N. Perrimon (2009). "Cross-species RNAi Rescue Platform in Drosophila melanogaster." Genetics.

Kondo, S. and N. Perrimon (2011). "A genome-wide RNAi screen identifies core components of the G-M DNA damage checkpoint." Sci Signal **4**(154): rs1.

Kornblihtt, A. R. (2006). "Chromatin, transcript elongation and alternative splicing." Nat Struct Mol Biol **13**(1): 5-7.

Kornfeld, K., D. B. Hom and H. R. Horvitz (1995). "The ksr-1 gene encodes a novel protein kinase involved in Ras-mediated signaling in C. elegans." Cell **83**(6): 903-913.

Kosako, H., N. Yamaguchi, C. Aranami, M. Ushiyama, S. Kose, N. Imamoto, H. Taniguchi, E. Nishida and S. Hattori (2009). "Phosphoproteomics reveals new ERK MAP kinase targets and links ERK to nucleoporin-mediated nuclear transport." Nat Struct Mol Biol **16**(10): 1026-1035.

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**(10): 833-838.

Kuo, H. C., F. H. Nasim and P. J. Grabowski (1991). "Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle." Science **251**(4997): 1045-1050.

Laufer, C., B. Fischer, W. Huber and M. Boutros (2014). "Measuring genetic interactions in human cells by RNAi and imaging." Nat Protoc **9**(10): 2341-2353.

Le Hir, H. and G. R. Andersen (2008). "Structural insights into the exon junction complex." Curr Opin Struct Biol **18**(1): 112-119.

Le Hir, H., D. Gatfield, I. C. Braun, D. Forler and E. Izaurralde (2001). "The protein Mago provides a link between splicing and mRNA localization." EMBO Rep. **2**: 1119-1124.

Le Hir, H., D. Gatfield, E. Izaurralde and M. J. Moore (2001). "The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay." EMBO J. **20**: 4987-4997.

Le Hir, H., E. Izaurralde, L. E. Maquat and M. J. Moore (2000). "The spliceosome deposits multiple proteins 20[ndash]24 nucleotides upstream of mRNA exon-exon junctions." EMBO J. **19**: 6860-6869.

Le Hir, H., M. J. Moore and L. E. Maquat (2000). "Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions." Genes Dev **14**(9): 1098-1108.

Le Hir, H. and B. Seraphin (2008). "EJCs at the heart of translational control." Cell **133**(2): 213-216.

Lee, H. C., J. Choe, S. G. Chi and Y. K. Kim (2009). "Exon junction complex enhances translation of spliced mRNAs at multiple steps." Biochem Biophys Res Commun.

Long, J. C. and J. F. Caceres (2009). "The SR protein family of splicing factors: master regulators of gene expression." Biochem J **417**(1): 15-27.

Luco, R. F., M. Allo, I. E. Schor, A. R. Kornblihtt and T. Misteli (2011). "Epigenetics in alternative pre-mRNA splicing." Cell **144**(1): 16-26.

Lykke-Andersen, J., M. D. Shu and J. A. Steitz (2001). "Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1." Science **293**(5536): 1836-1839.

Ma, X. M., S. O. Yoon, C. J. Richardson, K. Julich and J. Blenis (2008). "SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs." Cell **133**(2): 303-313.

Ma, Y., A. Creanga, L. Lum and P. A. Beachy (2006). "Prevalence of off-target effects in Drosophila RNA interference screens." Nature **443**(7109): 359-363.

Malone, C. D., C. Mestdagh, J. Akhtar, N. Kreim, P. Deinhard, R. Sachidanandam, J. Treisman and J. Y. Roignant (2014). "The exon junction complex controls transposable element activity by ensuring faithful splicing of the piwi transcript." Genes Dev **28**(16): 1786-1799.

Malumbres, M. and M. Barbacid (2003). "RAS oncogenes: the first 30 years." Nat Rev Cancer **3**(6): 459-465.

Martinez-Bartolome, S., P. A. Binz and J. P. Albar (2014). "The Minimal Information about a Proteomics Experiment (MIAPE) from the Proteomics Standards Initiative." Methods Mol Biol **1072**: 765-780.

Mathey-Prevot, B. and N. Perrimon (2006). "Drosophila genome-wide RNAi screens: are they delivering the promise?" Cold Spring Harb Symp Quant Biol **71**: 141-148.

Matter, N., P. Herrlich and H. Konig (2002). "Signal-dependent regulation of splicing via phosphorylation of Sam68." Nature **420**(6916): 691-695.

McGuire, A., M. Pearson, D. Neafsey and J. Galagan (2008). "Cross-kingdom patterns of alternative splicing and splice recognition." Genome Biology **9**(3): R50.

McKay, M. M. and D. K. Morrison (2007). "Integrating signals from RTKs to ERK/MAPK." Oncogene **26**(22): 3113-3121.

Michelle, L., A. Cloutier, J. Toutant, L. Shkreta, P. Thibault, M. Durand, D. Garneau, D. Gendron, E. Lapointe, S. Couture, H. Le Hir, R. Klinck, S. A. Elela, P. Prinos and B. Chabot (2012). "Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators." Mol Cell Biol **32**(5): 954-967.

Miller, B. W., G. Lau, C. Grouios, E. Mollica, M. Barrios-Rodiles, Y. Liu, A. Datti, Q. Morris, J. L. Wrana and L. Attisano (2009). "Application of an integrated physical and functional screening approach to identify inhibitors of the Wnt pathway." Mol Syst Biol **5**: 315.

Moffat, J., J. H. Reiling and D. M. Sabatini (2007). "Off-target effects associated with long dsRNAs in Drosophila RNAi screens." Trends Pharmacol Sci **28**(4): 149-151.

Mohr, S. E., S. T. Dillon and R. E. Boswell (2001). "The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during Drosophila oogenesis." Genes Dev **15**(21): 2886-2899.

Mohr, S. E., J. A. Smith, C. E. Shamu, R. A. Neumuller and N. Perrimon (2014). "RNAi screening comes of age: improved techniques and complementary approaches." Nat Rev Mol Cell Biol **15**(9): 591-600.

Mulcahy, L. S., M. R. Smith and D. W. Stacey (1985). "Requirement for ras proto-oncogene function during serum-stimulated growth of NIH 3T3 cells." Nature **313**(5999): 241-243.

Mummery-Widmer, J. L., M. Yamazaki, T. Stoeger, M. Novatchkova, S. Bhalerao, D. Chen, G. Dietzl, B. J. Dickson and J. A. Knoblich (2009). "Genome-wide analysis of Notch signalling in Drosophila by transgenic RNAi." Nature **458**(7241): 987-992.

Nagy, A., N. Perrimon, S. Sandmeyer and R. Plasterk (2003). "Tailoring the genome: the power of genetic approaches." Nat Genet **33 Suppl**: 276-284.

Nakai, K. and H. Sakamoto (1994). "Construction of a novel database containing aberrant splicing mutations of mammalian genes." Gene **141**(2): 171-177.

Nasim, F. H., P. A. Spears, H. M. Hoffmann, H. C. Kuo and P. J. Grabowski (1990). "A Sequential splicing mechanism promotes selection of an optimal exon by repositioning a downstream 5' splice site in preprotachykinin pre-mRNA." Genes Dev **4**(7): 1172-1184.

Neumuller, R. A. and N. Perrimon (2011). "Where gene discovery turns into systems biology: genome-scale RNAi screens in Drosophila." Wiley Interdiscip Rev Syst Biol Med **3**(4): 471-478.

Neumuller, R. A., F. Wirtz-Peitz, S. Lee, Y. Kwon, M. Buckner, R. A. Hoskins, K. J. Venken, H. J. Bellen, S. E. Mohr and N. Perrimon (2012). "Stringent analysis of gene function and protein-protein interactions using fluorescently tagged genes." Genetics **190**(3): 931-940.

Ni, J. Q., M. Markstein, R. Binari, B. Pfeiffer, L. P. Liu, C. Villalta, M. Booker, L. Perkins and N. Perrimon (2008). "Vector and parameters for targeted transgenic RNA interference in Drosophila melanogaster." Nat Methods **5**(1): 49-51.

Noselli, S. and N. Perrimon (2000). "Signal transduction. Are there close encounters between signaling pathways?" Science **290**(5489): 68-69.

Nott, A., H. Le Hir and M. J. Moore (2004). "Splicing enhances translation in mammalian cells: an additional function of the exon junction complex." Genes Dev. **18**: 210-222.

Nusslein-Volhard, C. and E. Wieschaus (1980). "Mutations affecting segment number and polarity in Drosophila." Nature **287**(5785): 795-801.

Old, W. M., J. B. Shabb, S. Houel, H. Wang, K. L. Couts, C. Y. Yen, E. S. Litman, C. H. Croy, K. Meyer-Arendt, J. G. Miranda, R. A. Brown, E. S. Witze, R. E. Schweppe, K. A. Resing and N. G. Ahn (2009). "Functional proteomics identifies targets of phosphorylation by B-Raf signaling in melanoma." Mol Cell **34**(1): 115-131.

Olivier, J. P., T. Raabe, M. Henkemeyer, B. Dickson, G. Mbamalu, B. Margolis, J. Schlessinger, E. Hafen and T. Pawson (1993). "A Drosophila SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of Ras guanine nucleotide exchange, Sos." Cell **73**(1): 179-191.

Palacios, I. M., D. Gatfield, D. St Johnston and E. Izaurralde (2004). "An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay." Nature **427**(6976): 753-757.

Papasaikas, P., J. R. Tejedor, L. Vigevani and J. Valcarcel (2014). "Functional Splicing Network Reveals Extensive Regulatory Potential of the Core Spliceosomal Machinery." Mol Cell.

Park, J. W., K. Parisky, A. M. Celotto, R. A. Reenan and B. R. Graveley (2004). "Identification of alternative splicing regulators by RNA interference in Drosophila." Proc Natl Acad Sci U S A **101**(45): 15974-15979.

Pearson, G., F. Robinson, T. Beers Gibson, B. E. Xu, M. Karandikar, K. Berman and M. H. Cobb (2001). "Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions." Endocr Rev **22**(2): 153-183.

Perkins, L. A., I. Larsen and N. Perrimon (1992). "corkscrew encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso." Cell **70**(2): 225-236.

Phatnani, H. P. and A. L. Greenleaf (2006). "Phosphorylation and functions of the RNA polymerase II CTD." Genes Dev **20**(21): 2922-2936.

Qui, M. S. and S. H. Green (1992). "PC12 cell neuronal differentiation is associated with prolonged p21ras activity and consequent prolonged ERK activity." Neuron **9**(4): 705-717.

Raabe, T., J. Riesgo-Escovar, X. Liu, B. S. Bausenwein, P. Deak, P. Maroy and E. Hafen (1996). "DOS, a novel pleckstrin homology domain-containing protein required for signal transduction between sevenless and Ras1 in Drosophila." Cell **85**(6): 911-920.

Ramadan, N., I. Flockhart, M. Booker, N. Perrimon and B. Mathey-Prevot (2007). "Design and implementation of high-throughput RNAi screens in cultured Drosophila cells." Nat Protoc **2**(9): 2245-2264.

Reed, R. (1996). "Initial splice-site recognition and pairing during pre-mRNA splicing." Curr Opin Genet Dev **6**(2): 215-220.

Rigo, F. and H. G. Martinson (2008). "Functional coupling of last-intron splicing and 3'-end processing to transcription in vitro: the poly(A) signal couples to splicing before committing to cleavage." Mol Cell Biol **28**(2): 849-862.

Robberson, B. L., G. J. Cote and S. M. Berget (1990). "Exon definition may facilitate splice site selection in RNAs with multiple exons." Mol. Cell. Biol. **10**(1): 84-94.

Roberts, P. J. and C. J. Der (2007). "Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer." Oncogene **26**(22): 3291-3310.

Rogge, R. D., C. A. Karlovich and U. Banerjee (1991). "Genetic dissection of a neurodevelopmental pathway: Son of sevenless functions downstream of the sevenless and EGF receptor tyrosine kinases." Cell **64**(1): 39-48.

Roignant, J. Y. and J. E. Treisman (2010). "Exon junction complex subunits are required to splice Drosophila MAP kinase, a large heterochromatic gene." Cell **143**(2): 238-250.

Roskoski, R., Jr. (2010). "RAF protein-serine/threonine kinases: structure and regulation." Biochem Biophys Res Commun **399**(3): 313-317.

Roy, M., N. Kim, Y. Xing and C. Lee (2008). "The effect of intron length on exon creation ratios during the evolution of mammalian genomes." RNA **14**(11): 2261-2273.

Saltzman, A. L., Q. Pan and B. J. Blencowe (2011). "Regulation of alternative splicing by the core spliceosomal machinery." Genes Dev **25**(4): 373-384.

Sauliere, J., V. Murigneux, Z. Wang, E. Marquenet, I. Barbosa, O. Le Tonqueze, Y. Audic, L. Paillard, H. Roest Crollius and H. Le Hir (2012). "CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex." Nat Struct Mol Biol **19**(11): 1124-1131.

Schneeberger, K. (2014). "Using next-generation sequencing to isolate mutant genes from forward genetic screens." Nat Rev Genet **15**(10): 662-676.

Schneider, M., C. L. Will, M. Anokhina, J. Tazi, H. Urlaub and R. Luhrmann (2010). "Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes." Mol Cell **38**(2): 223-235.

Schubbert, S., K. Shannon and G. Bollag (2007). "Hyperactive Ras in developmental disorders and cancer." Nat Rev Cancer **7**(4): 295-308.

Selfors, L. M., J. L. Schutzman, C. Z. Borland and M. J. Stern (1998). "soc-2 encodes a leucine-rich repeat protein implicated in fibroblast growth factor receptor signaling." Proc Natl Acad Sci U S A **95**(12): 6903-6908.

Sharma, S., A. M. Falick and D. L. Black (2005). "Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex." Mol Cell **19**(4): 485-496.

Sharma, S., L. A. Kohlstaedt, A. Damianov, D. C. Rio and D. L. Black (2008). "Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome." Nat Struct Mol Biol **15**(2): 183-191.

Shin, C. and J. L. Manley (2004). "Cell signalling and the control of pre-mRNA splicing." Nat Rev Mol Cell Biol **5**(9): 727-738.

Sieburth, D. S., Q. Sun and M. Han (1998). "SUR-8, a conserved Ras-binding protein with leucine-rich repeats, positively regulates Ras-mediated signaling in C. elegans." Cell **94**(1): 119-130.

Simon, M. A., D. D. Bowtell, G. S. Dodson, T. R. Laverty and G. M. Rubin (1991). "Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase." Cell **67**(4): 701-716.

Singh, G., A. Kucukural, C. Cenik, J. D. Leszyk, S. A. Shaffer, Z. Weng and M. J. Moore (2012). "The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus." Cell **151**(4): 750-764.

Smith, M. R., S. J. DeGudicibus and D. W. Stacey (1986). "Requirement for c-ras proteins during viral oncogene transformation." Nature **320**(6062): 540-543.

St Johnston, D. (2002). "The art and design of genetic screens: Drosophila melanogaster." Nat Rev Genet **3**(3): 176-188.

Steckelberg, A. L., V. Boehm, A. M. Gromadzka and N. H. Gehring (2012). "CWC22 connects pre-mRNA splicing and exon junction complex assembly." Cell Rep **2**(3): 454-461.

Stelling, J., U. Sauer, Z. Szallasi, F. J. Doyle, 3rd and J. Doyle (2004). "Robustness of cellular functions." Cell **118**(6): 675-685.

Stern, M. J. and H. R. Horvitz (1991). "A normally attractive cell interaction is repulsive in two C. elegans mesodermal cell migration mutants." Development **113**(3): 797-803.

Sterner, D. A., T. Carlo and S. M. Berget (1996). "Architectural limits on split genes." Proc Natl Acad Sci U S A **93**(26): 15081-15085.

Sundaram, M. and M. Han (1995). "The C. elegans ksr-1 gene encodes a novel Raf-related kinase involved in Ras-mediated signal transduction." Cell **83**(6): 889-901.

Talerico, M. and S. M. Berget (1990). "Effect of 5' splice site mutations on splicing of the preceding intron." Mol. Cell. Biol. **10**(12): 6299-6305.

Tange, T. O., A. Nott and M. J. Moore (2004). "The ever-increasing complexities of the exon junction complex." Curr Opin Cell Biol **16**(3): 279-284.

Taylor, C. F., N. W. Paton, K. S. Lilley, P. A. Binz, R. K. Julian, Jr., A. R. Jones, W. Zhu, R. Apweiler, R. Aebersold, E. W. Deutsch, M. J. Dunn, A. J. Heck, A. Leitner, M. Macht, M. Mann, L. Martens, T. A. Neubert, S. D. Patterson, P. Ping, S. L. Seymour, P. Souda, A. Tsugita, J. Vandekerckhove, T. M. Vondriska, J. P. Whitelegge, M. R. Wilkins, I. Xenarios, J. R. Yates, 3rd and H. Hermjakob (2007). "The minimum information about a proteomics experiment (MIAPE)." Nat Biotechnol **25**(8): 887-893.

Therrien, M., D. K. Morrison, A. M. Wong and G. M. Rubin (2000). "A genetic screen for modifiers of a kinase suppressor of Ras-dependent rough eye phenotype in Drosophila." Genetics **156**(3): 1231-1242.

Tsuda, L., Y. H. Inoue, M. A. Yoo, M. Mizuno, M. Hata, Y. M. Lim, T. Adachi-Yamada, H. Ryo, Y. Masamune and Y. Nishida (1993). "A protein kinase similar to MAP kinase activator acts downstream of the raf kinase in Drosophila." Cell **72**(3): 407-414.

Udell, C. M., T. Rajakulendran, F. Sicheri and M. Therrien (2011). "Mechanistic principles of RAF kinase signaling." Cell Mol Life Sci **68**(4): 553-565.

van der Straten, A., C. Rommel, B. Dickson and E. Hafen (1997). "The heat shock protein 83 (Hsp83) is required for Raf-mediated signalling in Drosophila." EMBO J **16**(8): 1961-1969.

van Eeden, F. J., I. M. Palacios, M. Petronczki, M. J. Weston and D. St Johnston (2001). "Barentsz is essential for the posterior localization of oskar mRNA and colocalizes with it to the posterior pole." J Cell Biol **154**(3): 511-523.

Vinayagam, A., Y. Hu, M. Kulkarni, C. Roesel, R. Sopko, S. E. Mohr and N. Perrimon (2013). "Protein complex-based analysis framework for high-throughput data sets." Sci Signal **6**(264): rs5.

Vinayagam, A., U. Stelzl, R. Foulle, S. Plassmann, M. Zenkner, J. Timm, H. E. Assmus, M. A. Andrade-Navarro and E. E. Wanker (2011). "A directed protein interaction network for investigating intracellular signal transduction." Sci Signal **4**(189): rs8.

Vinayagam, A., J. Zirin, C. Roesel, Y. Hu, B. Yilmazel, A. A. Samsonova, R. A. Neumuller, S. E. Mohr and N. Perrimon (2014). "Integrating protein-protein interaction networks with phenotypes reveals signs of interactions." Nat Methods **11**(1): 94-99.

Wang, J., X. Zhou, P. L. Bradley, S. F. Chang, N. Perrimon and S. T. Wong (2008). "Cellular phenotype recognition for high-content RNA interference genome-wide screening." J Biomol Screen **13**(1): 29-39.

Wang, Y., C. Werz, D. Xu, Z. Chen, Y. Li, E. Hafen and A. Bergmann (2008). "Drosophila cbl Is Essential for Control of Cell Death and Cell Differentiation during Eye Development." PLoS ONE **3**(1): e1447.

Wang, Z., V. Murigneux and H. Le Hir (2014). "Transcriptome-wide modulation of splicing by the Exon Junction Complex." Genome Biol **15**(12): 551.

Weg-Remers, S., H. Ponta, P. Herrlich and H. Konig (2001). "Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway." EMBO J **20**(15): 4194-4203.

Wiegand, H. L., S. Lu and B. R. Cullen (2003). "Exon junction complexes mediate the enhancing effect of splicing on mRNA expression." Proc. Natl. Acad. Sci. USA **100**: 11327-11332.

Xia, X. and Z. Xie (2001). "DAMBE: software package for data analysis in molecular biology and evolution." J Hered **92**(4): 371-373.

Xu, Y. Z., C. M. Newnham, S. Kameoka, T. Huang, M. M. Konarska and C. C. Query (2004). "Prp5 bridges U1 and U2 snRNPs and enables stable U2 snRNP association with intron RNA." Embo J **23**(2): 376-385.

Yoon, C. H., J. Lee, G. D. Jongeward and P. W. Sternberg (1995). "Similarity of sli-1, a regulator of vulval development in C. elegans, to the mammalian proto-oncogene c-cbl." Science **269**(5227): 1102-1105.

Zebisch, A., A. P. Czernilofsky, G. Keri, J. Smigelskaite, H. Sill and J. Troppmair (2007). "Signaling through RAS-RAF-MEK-ERK: from basics to bedside." Curr Med Chem **14**(5): 601-623.

Zhang, Z. and A. R. Krainer (2007). "Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components." Proc Natl Acad Sci U S A **104**(28): 11574-11579.

Zimyanin, V. L., K. Belaya, J. Pecreaux, M. J. Gilchrist, A. Clark, I. Davis and D. St Johnston (2008). "In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization." Cell **134**(5): 843-853.

Zipperlen, P., A. G. Fraser, R. S. Kamath, M. Martinez-Campos and J. Ahringer (2001). "Roles for 147 embryonic lethal genes on C.elegans chromosome I identified by RNA interference and video microscopy." Embo J **20**(15): 3984-3992.