Modeling of *SF3B1* cancer-related missense mutations in *C. elegans* uncovers targets for synthetic lethal interactions

Xènia Serrat[1], Dmytro Kukhtar[1], Eric Cornes[2], Anna Esteve-Codina[3], Helena Benlloch[1], Germano Cecere[2], Julián Cerón[1,*]

[1] Modeling human disease in *C. elegans* Group. Genes, Disease and Therapy Program, Institut d'Investigació Biomèdica de Bellvitge–IDIBELL. 08908, Barcelona, Spain.

[2] Mechanisms of Epigenetic Inheritance Group. Institut Pasteur. 75724, Paris, France.

[3] CNAG-CRG, Barcelona Institute of Science and Technology (BIST). 08028, Barcelona, Spain.

[*] Correspondence: jceron@idibell.cat

Abstract

SF3B1 is the most frequently mutated splicing factor in cancer. Mutations in SF3B1 confer growth advantages to cancer cells but they may also confer vulnerabilities that can be therapeutically targeted. In contrast to other animal models, SF3B1 cancer mutations can be maintained in homozygosis in *C. elegans*, allowing synthetic lethal screens with a homogeneous population of animals. These mutations cause alternative splicing (AS) defects in *C. elegans*, as it occurs in SF3B1-mutated human cells. In an RNAi screen, we identified U2 snRNP components as causing synthetic lethality with *sftb-1*/SF3B1 mutations. We also detected synthetic interactions between *sftb-1* mutants and cancer-related mutations in *uaf-2*/U2AF1 or *rsp-4*/SRSF2, demonstrating that this model can identify mutations that are mutually exclusive in human tumors. Thus, we have established a multicellular model amenable for high-throughput screening to uncover synthetic lethal interactions that could selectively kill cancer cells harboring SF3B1 mutations.

INTRODUCTION

Deregulated RNA splicing is emerging as a new hallmark of cancer following the discovery of several splicing factor genes harboring somatic mutations in different tumor types (Urbanski et al., 2018). *SF3B1*, a core component of the U2 snRNP, is the most frequently mutated splicing factor in human cancers. Somatic heterozygous mutations in *SF3B1* are particularly prevalent in myelodysplastic syndromes (MDS)—up to 80% in refractory anemia with ring sideroblasts (Papaemmanuil *et al.*, 2011; Yoshida *et al.*, 2011; Malcovati *et al.*, 2015); and in 15% of chronic lymphocytic leukemia (CLL) (Quesada *et al.*, 2011; Wang *et al.*, 2011; te Raa *et al.*, 2015). *SF3B1* mutations have also been reported in solid tumors, including 20% of uveal melanomas (UM) (Furney *et al.*, 2013; Harbour *et al.*, 2013; Martin *et al.*, 2013), 3% of pancreatic ductal adenocarcinomas (Biankin *et al.*, 2012), and 1.8% of breast cancers (Maguire *et al.*, 2015). RNA-sequencing (RNA-seq) analyses in different tumor types and cell lines with *SF3B1* missense mutations, including the most prevalent substitution K700E, identified distinct alternative splicing (AS) defects (Darman *et al.*, 2015; DeBoever *et al.*, 2015; Liberante *et al.*, 2019).

Recently, CRISPR/Cas9 genome editing has allowed researchers to faithfully reproduce human pathological mutations in animal models. Human *SF3B1* mutations have been introduced in cell lines (Alsafadi *et al.*, 2016; Shiozawa *et al.*, 2018; Gupta *et al.*, 2019; Liberante *et al.*, 2019), but mimicking these mutations in multicellular organisms has only been achieved by conditional alleles in murine models (Xu *et al.*, 2019). Taking advantage of the extraordinary conservation of splicing factors across evolution and the ease of

genetic manipulation of *C. elegans*, we established a multicellular model to study *SF3B1* cancer-related mutations. Since the spliceosome components, and particularly *SF3B1*, are being intensively studied as a target of antitumor drugs (Bonnal, Vigevani and Valcárcel, 2012; Effenberger, Urabe and Jurica, 2017; DeNicola and Tang, 2019), the multicellular model of *SF3B1* mutations described in this study would provide a convenient pre-clinical platform for identifying new targets and small molecules for use in cancer therapies.

RESULTS

sftb-1, the *C. elegans* ortholog of *SF3B1*, is an ubiquitously expressed gene essential for development

The *C. elegans* protein SFTB-1 is 66% identical to human SF3B1 in terms of amino acid composition. Homology is particularly high at the HEAT domain, reaching 89% identity, and the most frequently mutated amino acids in cancer are conserved. The SFTB-1 sequence also conserves some of the U2AF ligand motifs (ULMs) that bind U2AF homology motifs (UHMs) present in other splicing factors (**Figure 1A and S1**) (Loerch *et al.*, 2018).

As expected for a core splicing factor, the CRISPR-engineered endogenous fluorescent reporter *mCherry::SFTB-1* showed ubiquitous expression in somatic and germ cells, being absent in mature sperm only (**Figure 1b and 1c**). We also used CRISPR to generate the *sftb-1(cer6)* mutation, a deletion allele that produces a premature stop codon (**Figure S2a**) and causes an arrest at early larval stages, confirming that *sftb-1* is essential for development (**Figures 1d and 1e**). *sftb-1(cer6)* homozygous animals completed embryonic development but arrested as larvae that could undergo few germ cell divisions (**Figure S2b**), suggesting that maternal wild-type (WT) product can still be present in the early larva.

Different *sftb-1* missense mutations display additive effects when combined

We edited the *C. elegans* genome to mimic the K700E mutation, the most frequent *SF3B1* substitution in tumors. *sftb-1[K718E]* homozygous animals did not display any obvious phenotypes. Thus, we decided to reproduce two other *SF3B1* cancer-related mutations. R625C is a particularly prevalent mutation in uveal melanoma, and Q534P is less common but is predicted to have a strong impact on SF3B1 structure due to its location in an α-helix (Cretu *et al.*, 2016). Similar to *sftb-1[K718E]*, the two additional missense mutations, *sftb-1[R643C]* and *sftb-1[Q552P]* (**Figures S3a and S3b**), did not cause any evident phenotypes in *C. elegans*.

We were interested in producing animals with defective *sftb-1* function resulting in a viable and traceable phenotype for investigating modifiers of altered SFTB-1 activity. Since the effects of MDS-related missense mutations were reported to be additive in the yeast *SF3B1* ortholog Hsh155 (Carrocci *et al.*, 2017), we generated three strains with each of the possible pairs of mutations: *sftb-1[R643C, K718E], sftb-1[Q552P, K718E], and sftb-1[Q552P, R643C]*. None of the three double mutants showed an obvious phenotype. However, animals simultaneously carrying the three missense mutations *sftb-1[Q552P, R643C, K718E]* displayed several temperature-sensitive developmental defects including partial sterility, protruding vulva, and developmental delay (**Figure 1f–1h and S3c–e**). These results suggest that each of the three missense mutations provokes a functional impact that becomes unmasked only when the three mutations are combined.

Distinct *sftb-1* alleles produce alterations in alternative splicing and gene expression

The most commonly described AS defect caused by cancer-associated *SF3B1* mutations is the incorrect recognition of 3' splice sites, producing aberrant transcripts (Darman *et al.*, 2015; DeBoever *et al.*, 2015; Wang *et al.*, 2016; Pellagatti *et al.*, 2018; Shiozawa *et al.*, 2018) (**Figure 2a**). We investigated in *C. elegans* the transcriptional consequences of the K718E mutation, an equivalent of the most frequent *SF3B1* mutation K700E, by RNA-seq of two different strains with two independent K718E alleles (*cer3* and *cer7*) and their WT siblings at the L4 stage. Then, we used rMATS for robust detection of differential AS events between mutant and WT worms (Shen *et al.*, 2014). Using both splice junction and exon body reads, our analysis detected 21985 AS events, of which 134 were significantly altered (FDR < 0.05) (**Table S1**). A more restricted cut-off (inclusion level difference > 0.1), as used in a previous rMATS analysis of *SF3B1*-mutated samples (Pellagatti *et al.*, 2018), reduced the list of significant AS events to 78, with skipped exon (SE) being the most frequent event (**Figures 2b and 2c**).

Then, we investigated whether the two *sftb-1* alleles associated with a visible phenotype (the triple mutation allele *cer39*, and the deletion allele *cer6)* primarily induced cryptic 3' splice site selection. It is worth noting that, in this experiment, larvae were harvested and processed for RNA isolation at the L2 stage to avoid possible indirect effects due to developmental defects observed after this stage. Both alleles were associated with a high number of deregulated AS events as determined by rMATS (**Figure 2b**), and the predominant alteration was also exon skipping (**Figure 2c**). We then explored if there were

differentially expressed protein-coding (**Figure 2d**) and non-coding (**Figure 2e**) transcripts in the three RNA-seq datasets **(Table S2)**. As expected, the highest number of significantly altered transcripts was observed in the *cer6* allele.

In summary, we identified gene expression and alternative splicing alterations that are more frequent as *sftb-1* function becomes progressively compromised, with SE being the predominant AS defect.

A subset of U2 snRNP and U2-associated components synthetically interacts with *sftb-1* mutants

In the search for selective vulnerabilities in *sftb-1* defective backgrounds, we hypothesized that *sftb-1* mutant animals would be more sensitive to knockdown of other splicing factors than their WT counterparts. Based on a previous *C. elegans* RNA interference (RNAi) collection of splicing factors (Kerins *et al.*, 2010), we compiled and assayed a total of 104 RNAi clones targeting proteins that act in different steps of the splicing reaction. This led us to identify 27 splicing factors whose knockdown phenotype was enhanced in at least one of the three *sftb-1* single mutants, suggesting possible synthetic interactions (**Table S3**).

Among the 27 candidates, we selected 3 U2 snRNP components, including *sftb-1*, and 2 U2-associated proteins for further validation **(Figure 3A)**. The identification of *sftb-1* as a candidate certified our approach since it demonstrated that any RNAi affecting *sftb-1*-related functions would have a stronger effect on *sftb-1* mutants. Since a partial loss-of-function mutant for the candidate *teg-4* was available, we proceeded to validate the interaction between *sftb-1* mutations and *teg-4*, the *C. elegans* ortholog of *SF3B3*, which

physically interacts with SF3B1 HEAT repeats (Cretu *et al.*, 2016). We found that the K718E mutation exacerbated the phenotype of *teg-4(oz210)* at 20ºC (Mantina *et al.*, 2009), whereas the Q552P mutation did so to a lesser extent (**Figure 3B**).

Then, we validated the rest of the candidates using RNAi. As we had observed that *sftb-1* mutations had an additive effect (**Figure 1f–1h**), we reasoned that *sftb-1* double mutants might be more sensitive to splicing perturbations than single mutants. Consistently, *sftb-1(RNAi)* led to an additive loss of function in double mutants, which arrested at earlier stages compared to WT worms (**Figure 3c**). A similar result was obtained with knockdown of *uaf-2*/*U2AF1* (**Figure 3d**), another hit from the RNAi screen.

Among the remaining U2 and U2-associated candidates, we selected *mog-2*/*SNRPA1* and *smr-1*/*SMNDC1* for further validation in the double-mutation backgrounds. Strikingly, whereas RNAi of *mog-2* resulted in very low penetrant sterility in WT worms, the incidence of the phenotype incremented to over 70% in the three *sftb-1* double-mutation strains (**Figure 3e**). A similar synthetic phenotype was observed upon knockdown of *smr-1* (**Figure 3f**). Finally, we further validated the synthetic interaction between *teg-4* and *sftb-1,* as *sftb-1* double mutants were more sensitive to *teg-4(RNAi)* (**Figure 3g**).

Together, our results suggest that targeting other splicing factors such as *SF3B3* to compromise their function can be an opportunity to selectively impair cancer cells harboring *SF3B1* mutations.

The *C. elegans* mutations equivalent to human *U2AF1*^{S34P} and *SRSF2*^{P95H} genetically interact with *sftb-1[Q552P, R643C, K718E]* mutants

Along with *SF3B1* mutations, recurrent missense mutations in U2 small nuclear RNA auxiliary factor 1 (*U2AF1*) and serine/arginine-rich splicing factor 2 (*SRSF2*) were discovered in hematological malignancies (Yoshida *et al.*, 2011), and later reported to be present in some solid tumors (Imielinski *et al.*, 2012; Seiler, Peng, *et al.*, 2018). Somatic hotspot mutations in these three splicing factors are only found in heterozygosis and are mutually exclusive, suggesting redundant functional impact or limited tolerance to splicing perturbations.

Missense mutations in *U2AF1* are commonly found at position S34, which corresponds to a residue located within the two zinc finger domains through which U2AF1 is thought to interact with the pre-mRNA (Ilagan *et al.*, 2015). Similarly, *SRSF2* is typically mutated at position P95, mainly bearing missense mutations but also in-frame insertions or deletions (Yoshida *et al.*, 2011).

Based on a recently reported synthetic interaction between *Srsf2*$^{P95H/+}$ and *Sf3b1*$^{K700E/+}$ in mice (Lee *et al.*, 2018), and on the fact that *sftb-1* mutants were more sensitive to inactivation of some splicing factors (**Figure 3**), we wondered if the combination of *sftb-1[K718E]* with mutations equivalent to *U2AF1*^{S34F} or *SRSF2*^{P95H} would be tolerated in *C. elegans*.

Thus, we edited the worm orthologs of these genes to generate the *uaf-2[S42F]* and *rsp-4[P100H]* mutant strains. Neither of these two mutations caused any obvious phenotypes. We did not detect any interaction when combining each of these mutants with *sftb-1[K718E]* or the three *sftb-1* double-mutation strains. In contrast, *rsp-4[P100H]* and *uaf-2[S42F]* significantly increased the low percentage of sterile worms observed in the *sftb-1[Q552P, R643C, K718E]*

background at 20ºC (**Figure 4a**). The most dramatic effect was observed by the addition of the *uaf-2[S42F]* mutation to the *sftb-1* triple-mutation strain. Moreover, this combination of *uaf-2* and *sftb-1* mutations induced developmental delay at 20ºC (**Figure 4b**). These results show for the first time in an animal model, a synthetic interaction between *sftb-1* and *uaf-2* missense mutations.

DISCUSSION

SF3B1 cancer-related mutations have been intensely studied in the past few years. They are particularly recurrent in MDS or UM, which commonly lack effective therapeutic approaches. In this study, we establish the first multicellular model for *SF3B1* mutations expressed non-conditionally from the endogenous locus. The rapid *C. elegans* life cycle (three days at 25ºC) and the around 300 self-progeny obtained from a single hermaphrodite, among other features, make this model organism ideal for large-scale genetic and drug screens. The task of modeling human pathological mutations has been facilitated by the advent of the CRISPR/Cas9 genome editing technique, which has been efficiently implemented in this nematode.

As expected for a core splicing factor, and as previously reported in zebrafish (An and Henion, 2012) and mice (Isono *et al.*, 2005), the *C. elegans SF3B1* ortholog is essential for development. Cancer cells bear *SF3B1* hotspot mutations in heterozygosis and depend on *SF3B1* WT function for viability. However, homozygous *sftb-1* missense mutations are viable in *C. elegans,* although they generate defects that are unmasked when additional alterations in the spliceosome are introduced. Such a feature is advantageous for investigating vulnerabilities of *SF3B1*-mutated tumors, which are not only conferred by missense mutations, but also by the loss of one *SF3B1* WT copy (Alsafadi *et al.*, 2016; Paolella *et al.*, 2017).

Differently from MDS with *SF3B1* mutations, where aberrant transcripts have been directly related to phenotypic traits (Dolatshad *et al.*, 2015; Pellagatti *et al.*, 2018), transcripts affected by *sftb-1[K718E]* do not produce any obvious phenotypes. Still, *Sf3b1*$^{K700E/+}$ mice models only partially recapitulate the hematopoietic phenotype characteristic of MDS (Obeng *et al.*, 2016; Mupo *et al.*, 2017) and CLL (Yin *et al.*, 2019). In these models, the splicing alterations caused by *Sf3b1*$^{K700E/+}$ were similar to that reported in human tumors, but most of the affected transcripts were different, probably due to low conservation in intronic sequences.

RNA-seq of *sftb-1[K718E]* animals did not identify the aberrant use of alternative 3' splice sites (A3'SS) as the predominant splicing defect caused by this mutation, as has been observed in human cells (Darman *et al.*, 2015; DeBoever *et al.*, 2015; Alsafadi *et al.*, 2016; Pellagatti *et al.*, 2018; Shiozawa *et al.*, 2018). Indeed, in three distinct *sftb-1* mutants, we found a mild effect on AS, with SE being the most abundant splicing defect. Such disparity in the consequences of the mutation K700E/K718E between humans and nematodes may be explained by different factors. First, diverse computational pipelines have been used to assess differences in alternative splicing between WT and *SF3B1*-mutated samples, resulting in distinct conclusions. For instance, *Liberante et al.* reanalyzed previously published data and found that *SF3B1* mutation was mostly associated with differential skipped exon events, rather than A3'SS selection (Liberante *et al.*, 2019). Second, intronic elements directing splicing have evolved differently in *C. elegans* compared to other organisms. *C. elegans* introns are typically shorter, do not have a consensus

branch site (BS) sequence, and the vertebrate polypyrimidine (Py) tract is replaced by the consensus UUUUCAG/R sequence (Morton and Blumenthal, 2011). Third, only ~35% of genes are alternatively spliced in *C. elegans* (Tourasse, Millet and Dupuy, 2017) compared to 95% in humans (Pan *et al.*, 2008). Fourth, cell type-specific alternative isoforms generated by *sftb-1* mutations could have been underrepresented in our analyses as we sequenced the transcriptome of whole animals, and A3'SS selection can be regulated in a tissue-specific manner in *C. elegans* (Ragle *et al.*, 2015). And fifth, the nonsense-mediated decay (NMD) system could efficiently eliminate transcripts with A3'SS resulting from *sftb-1* mutations, as occurs in human cells (Darman *et al.*, 2015). However, we did not detect predominance of downregulation among the differentially expressed transcripts (**Figures 2D and 2E**).

Beyond its action on A3'SS usage, *SF3B1* missense mutations could have an impact on the structure, the interaction with other proteins in the spliceosome, or the robustness of the protein. In fact, the small molecule H3B-8800, which induces lethality in spliceosome-mutant cancers by modulating the SF3b complex, affects AS activity similarly in WT and in *SF3B1*^{K700E} human cells (Seiler, Yoshimi, *et al.*, 2018). Therefore, *SF3B1*-mutated cells present vulnerabilities that can be exploited in synthetic lethal screens. We found that *sftb-1* mutants are sensitive to RNAi of different splicing factors, including itself, and we further validated the synthetic interaction with the inactivation of other components of the U2 snRNP. These results point to other components of the spliceosome as targets to preferentially kill cells with deficiencies in SF3B1 activity. Knowing the suitability of *sftb-1* mutants for RNAi screens, the

collection of RNAi clones to be tested should be expanded, with particular interest placed on deubiquitinases and chromatin factors that have been related with SF3B1 functions (Kfir *et al.*, 2015; Paolella *et al.*, 2017).

Pladienolide and derivatives have been proven to be effective in selectively killing cells with *SF3B1* mutations (Obeng *et al.*, 2016; Seiler, Yoshimi, *et al.*, 2018; Wu *et al.*, 2018). We tested Pladienolide B and Sudemycin D6 in WT and *sftb-1* mutant worms and we did not observe any effects (**Table S4**). The inhibitory activity of these drugs is likely dependent on protein structure, which may differ from nematodes to humans. In fact, human VAL1078 corresponds to ILE1096 in *C. elegans,* and a mutation to isoleucine or alanine at this position is sufficient to render *SF3B1*-mutant cells insensitive to Pladienolide B (Teng *et al.*, 2017). Since editing of this residue in the *SF3B1* yeast ortholog Hsh155 sensitizes *S. cerevisiae* to splicing inhibitors (Hansen *et al.*, 2019), these drugs could also inhibit SF3B1 chimeric proteins in *C. elegans*.

SF3B1 mutations are mutually exclusive with other splicing factor mutations present in tumors. *U2AF1* and *SRSF2* missense mutations are reported in 11% and 12-15% of MDS, respectively (Xu *et al.*, 2019), and very rarely overlap with *SF3B1* mutations, which are present in 17-28% of MDS patients (Lee *et al.*, 2018; Seiler, Peng, *et al.*, 2018). Hence cells may not handle all the AS defects produced by the combination of mutations in any of these three proteins. To further validate our model as a tool for identifying vulnerabilities in cells with compromised SF3B1 function, we mimicked *U2AF1*^{S34F} and *SRSF2*^{P95H} mutations in *C. elegans* and detected a functional interaction between both mutants and our *sftb-1* mutant carrying the three missense mutations.

Additionally, *C. elegans rsp-4* and *uaf-2* missense mutations could be used for RNA-seq analyses and RNAi screens to study the consequences and vulnerabilities of these mutations, respectively.

We have demonstrated that *C. elegans* can be a valuable model to explore vulnerabilities of cells with mutations in *SF3B1*, although the molecular consequences of *sftb-1* mutations were different from the ones observed in other organisms. Replacement of several HEAT repeats in yeast Hsh155 with their human counterparts is functional and supports splicing (Carrocci, Paulson and Hoskins, 2018). Similarly, partial or total replacement of *sftb-1* for its human ortholog could result in a more accurate model. To date, only one human gene has been functionally transplanted to *C. elegans* (McDiarmid *et al.*, 2018), in part due to low efficiencies of inserting long DNA fragments by CRISPR. Such technical difficulties can now be bypassed by the use of Nested CRISPR (Vicencio *et al.*, 2019). Thus, humanized *sftb-1* nematodes obtained by this method would be very efficient as *in vivo* models to further understand the impact of *SF3B1* mutations in cancer and as a platform for large-scale drug screens.

Author Contributions

Conceptualization, X.S. and J.C.; Methodology, X.S., G.C., and J.C.; Validation, X.S.; Formal analysis, X.S. and A.E.C.; Investigation, X.S., D.K., H.B., and E.C.; Writing – original draft, X.S. and J.C., with help from the other authors; Writing – Review & Editing, X.S. and J.C.; Supervision, J.C.; Funding acquisition J.C.

Acknowledgments

We acknowledge the members of the Cerón Laboratory for helpful discussions and comments on the manuscript. We also thank Alan Zahler and Sol Katzman for their advice in transcriptomic analyses. The *teg-4(oz210)* strain was kindly provided by Dave Hansen. Pladienolide B and Sudemycin D6 were gifts from Juan Valcárcel. Thanks to @shookstudio for their assistance in making figures the graphical abstract.This work has been supported by a grant from the Instituto de Salud Carlos III (ISCIII) to JC (PI15-00895), co-funded by FEDER funds/European Regional Development Fund (ERDF) — a way to build Europe. We thank CERCA Program / Generalitat de Catalunya for their institutional support. XS has an FPU PhD fellowship from MINECO and DK has an FI AGAUR fellowship from Generalitat de Catalunya.

Declaration of Interests

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1. Characterization of *sftb-1,* the *C. elegans* ortholog of *SF3B1*, and different *sftb-1* mutations.

(**a**) Top: Number of mutations annotated in COSMIC database (Forbes *et al.*, 2017) at each amino acid position. Middle: schematic structure of the SF3B1 protein. Purple boxes—U2AF ligand motifs (ULMs); yellow box—p14-interacting region; blue boxes—20 HEAT repeats. Labels indicate the three residues that are relevant to this study. Bottom: bars indicate residues that are conserved in the *C. elegans* ortholog *sftb-1.* (**b**) Representative image of an embryo, L2, and L4 larvae showing *mCherry::SFTB-1* endogenous expression. The fluorescence image (top panel) is merged with the corresponding differential interference contrast (DIC) image (bottom panel). (**c**) Fluorescence images of dissected gonads from adult worms expressing *mCherry::SFTB-1*. Gonads were stained with DAPI to visualize nuclei. Insets show differentiated sperm nuclei, where *mCherry::SFTB-1* is not expressed. (**d**) Left: diagram depicting how the *cer6* deletion allele is maintained in heterozygosis. The WT *sftb-1* copy is genetically linked to a red fluorescent marker (*oxTi615*). Right: worm body length quantification of WT and heterozygotes (red), and *cer6* homozygotes (gray) grown for 48h at 20°C (n=475, 134; N=3). (**e**) Representative images of synchronized red and non-red siblings after growing for 24h and 48h at 20°C. (**f**) Incidence of different *sftb-1[Q552P, R643C, K718E]* phenotypes at 25°C. Dots represent percentages observed in each replicate (n=119, 116; N=2). (**g**) Total number of progeny laid by WT or *sftb-1[Q552P, R643C, K718E]* fertile worms at 25°C (n=22, 35; N=2). (H) *sftb-1[Q552P, R643C, K718E]* mutant displays

developmental delay at 25ºC. The proportion of the population at each stage is represented by the size of the dots (n=119, 116; N=2).

In (**d**) and (**g**), dots represent measures in individual worms, and overlaid Tukey-style boxplots represent the median and interquartile range (IQR). Whiskers extend to the minimum and maximum values or at a maximum distance of 1.5xIQR from the box limits. Statistics: (**d**) and (**g**), Mann-Whitney's test; (**f**), Fisher's exact test. n.s., no significant difference, **** $p < 0.0001$. Scale bars: 100 µm (**b** and **e**), 10 µm (**c**).

See also **Figures S1, S2 and S3**.

Figure 2. Different *sftb-1* mutations induce alternative splicing and gene expression alterations.

(**a**) Types of AS events. Gray boxes represent constitutive exons, while colored boxes represent alternatively spliced exons. (**b**) Total number of significant (FDR < 0.05) AS alterations detected by rMATS with an inclusion level difference > 0.1. Event types are color coded as in panel A. Light colors represent AS events that were more significantly included in WT samples, while dark colors symbolize splicing alterations associated to mutant samples. (**c**) Doughnut charts showing the distribution of significant events represented in panel B by event type. In the three mutants, skipped exon (SE) was the most predominant alteration. (**d**) MA-plots displaying changes in protein-coding transcript expression levels (corrected $\log_2$fold change) over the mean of normalized counts, analyzed by DESeq2. Significant transcripts (padj > 0.05) are colored in blue. (**e**) MA-plots of expression changes in non-coding

transcripts, represented as in panel D. Transcripts that are significantly deregulated are colored by transcript type.

See also **Tables S1 and S2**.

Figure 3. An RNAi screen targeting splicing genes leads to the validation of synthetic interactions between *sftb-1* mutants and five U2 snRNP and U2-associated components.

(**a**) Schematics of the RNAi screen. In brief, WT or mutant worms bearing any of the three *sftb-1* single mutations were fed RNAi clones against a total of 104 splicing factors. 77 of these clones had the same effect in all genetic backgrounds. In contrast, we identified 27 clones that caused a stronger phenotype in at least one of the three single mutant strains compared to WT, which was indicative of a synthetic interaction. Listed are the 3 U2 snRNP and 2 U2 snRNP-related candidates from the screen that were selected for further validation. (**b**) *sftb-1[K718E]* and *sftb-1[Q552P]* synthetically interact with *teg-4(oz210).* The number of hatched larvae (orange) and dead embryos (gray) laid by worms of the indicated genotype at 20°C is represented (n≥28; N=3). *m+ z-*, homozygous mutant progeny of heterozygous mothers. No significant differences were observed in the number of hatched larvae laid by *sftb-1[K718E], sftb-1[R643C],* and *sftb-1[Q552P]* compared to WT. (**c**) Worm length after 48h of RNAi treatment against *gfp* (black) or *sftb-1* (blue) at 25°C (n≥150; N=2). (**d**) Worm length after 48h of RNAi treatment against *gfp* (black) or *uaf-2* (blue) at 25°C (n≥150; N=2). In both (c) and (d), worm length was not significantly different between *sftb-1* mutants and WT upon *gfp(RNAi)*. (e-g)

Mean percentage of sterile worms observed upon RNAi of *mog-2* (E; n≥30), *smr-1* (F; n≥27) and *teg-4* (G; n≥34) at 25ºC (N=2). Red dots indicate percent sterility observed in each replicate. Gray bars, P_0-treated worms giving rise to <5 F_1 larvae and some dead embryos ('dead progeny' category); black bars, P_0-treated worms that laid neither larvae nor dead embryos ('no progeny' category).

In (b), (c), and (d), data are shown as Tukey-style boxplots and overlaid dots representing measures in individual animals. Statistics: (b), (c), and (d), Kruskal-Wallis test with Dunn's multiple comparisons test. Data were compared to *teg-4(oz210)* (B), or WT animals fed the same RNAi (c) and (d). (e), (f), and (g), Fisher's exact test with Bonferroni correction for multiple comparisons. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

See also **Table S3**.

Figure 4. Genetic interaction between *rsp-4/SRSF2*, *uaf-2/U2AF1*, and *sftb-1* mutants.

(A) Mean percentage of sterile worms of the indicated genotype at 20ºC (N=2). Black dots denote observed values in each replicate. (B) *rsp-4[P100H]* and *uaf-2[S42F]* mutations gradually exacerbate the mild developmental delay displayed by the *sftb-1[Q552P, R643C, K718E]* mutant at 20ºC. Dot sizes symbolize the percentage of worms at each developmental stage on any given day. In both panels, n≥110, N=2.

Statistics: (A) Fisher's exact test with Bonferroni correction for multiple comparisons. Lines indicate compared groups. ** $p < 0.01$, **** $p < 0.0001$.

METHODS

Caenorhabditis elegans

C. elegans strains were maintained using standard methods (Stiernagle, 2006; Porta-de-la-Riva *et al.*, 2012) at the specified temperature. Before conducting the experiments, strains were grown for at least two generations at the experimental temperature. Unless otherwise specified, we used Bristol N2 as the WT strain. Mutant strains with single *sftb-1* missense mutations (CER218, CER220, CER236, and CER238) were outcrossed x2. WT siblings resulting from the second outcross were given specific strain names (CER217, CER221, CER237, and CER239, respectively) and kept for closer comparison. CER351, CER442 and CER450 strains were outcrossed x2. In order to maintain the *cer6* deletion, heterozygous mutants were crossed with the EG7893: *oxTi615 unc-119(ed3) III* strain so that a red fluorescent marker was inserted 0.15 cM away from the *sftb-1* WT locus.

Protein alignments

H. sapiens SF3B1 (Uniprot: O75533) and *C. elegans* SFTB-1 (Uniprot: G5EEQ8) protein sequences were aligned using the T-Coffee algorithm (Notredame, Higgins and Heringa, 2000) with default settings. **Figure S1** was generated with Jalview (Waterhouse *et al.*, 2009).

CRISPR/Cas9 mutant and reporter strains

Specific guide RNAs were designed using both *Benchling* (www.benchling.com) and *CCTop* (Stemmer *et al.*, 2015) online tools. All CRISPR/Cas9 mutant and reporter strains were obtained following a co-CRISPR strategy (Kim *et al.*,

2014) using *dpy-10* as a marker to enrich for genome-editing events. In all cases, mixes were injected into gonads of young adult P_0 hermaphrodites using the XenoWorks Microinjection System and following standard *C. elegans* microinjection techniques. F_1 progeny was screened by PCR using specific primers and F_2 homozygotes were confirmed by Sanger sequencing. As we optimized the protocol for CRISPR/Cas9 genome editing in our laboratory while preparing this manuscript, different reagents and conditions were used to produce the distinct mutants. The *sftb-1(cer6)* deletion mutant allele was obtained by nonspecific double-strand break repair on our first attempt to generate the K718E mutation. Only in this experiment, guide RNAs and Cas9 were expressed from plasmids that were purified with the NucleoBond® Xtra Midi Kit (Macherey-Nagel). In the remaining mixes, ribonucleoprotein complexes containing crRNA, tracrRNA and Cas9 were annealed at 37°C for 15 minutes prior to injection. All the reagents and injection mixes used in this study are listed in **Table S5**.

C. elegans microscopy

Worms were mounted on 2% agar pads with 10 mM tetramisole hydrochloride and imaged using a ZEISS Axio Observer Z1 inverted fluorescence microscope. Images were processed using ZEISS ZEN 2012 (blue edition) software and Fiji.

Gonad dissection and DAPI staining

Worms expressing *mCherry::SFTB-1* were anesthetized with 0.33 mM tetramisole hydrochloride and their heads cut using a pair of 25-gauge needles to extrude gonad arms. Released gonads were fixed in 4% paraformaldehyde in PBS for 20–30 minutes and washed with PBS + 0.1% Tween 20 for 10 minutes.

After performing 3 washes, gonads were transferred to a microscope slide and mounted using 10 µl DAPI Fluoromount-G® (SouthernBiotech). Images were acquired the same day to avoid mCherry signal loss.

sftb-1 **deletion mutant larval arrest**

Synchronized L1-arrested F_2 worms from heterozygous *sftb-1(cer6)/oxTi615* P_0s were seeded onto NGM plates with OP50 and imaged after 48 hours of growth at 20ºC using a ZEISS Axio Observer Z1 inverted fluorescence microscope at 10X magnification. Body length was measured manually by drawing the midline of the worm using the freehand line tool in Fiji.

Phenotypic characterization

Similar experiments were conducted to characterize *sftb-1[Q552P, R643C, K718E], rsp-4[P100H]* and *uaf-2[S42F]* mutant phenotypes. Briefly, synchronized L1-arrested worms of the specified genotype were first seeded onto NGM plates with OP50 and, after a few hours to allow recovery from the L1 arrest, 60 animals per genotype were separated individually to 12-well NGM plates seeded with OP50. Plates were incubated at 20ºC (**Figures 4 and S3c–E**) or 25ºC (**Figure 1f–h**). The developmental stage of individual worms was assessed visually every 24 hours under a stereomicroscope based on animal size and germline morphology. Phenotypes were scored at day 4. In order to determine brood sizes of *sftb-1[Q552P, R643C, K718E]* and WT animals, a proportion of worms were transferred into new 35-mm NGM plates at day 4 and subsequently moved to fresh plates every 12–24 h until egg laying ceased. Progeny were counted 36–48 h after passages.

RNA isolation

In order to study the transcriptional consequences of the K718E mutation, we used two homozygous mutant lines (CER218 and CER220) derived from two independent CRISPR/Cas9 events (*cer3* and *cer7* alleles, respectively), and the corresponding WT siblings (CER217 and CER221). Synchronized L4 animals were collected in M9 after growing for 41 hours at 20ºC and total RNA was isolated using TRI Reagent (MRC, Inc.) and the PureLink® RNA Mini Kit (Ambion), following the manufacturer's instructions.

sftb-1(cer6) homozygous larvae were obtained by automatically sorting non-red animals from a synchronized CER191 population grown at 20ºC for 23 hours, using the COPAS Biosorter system (Union Biometrica). We conducted two biological replicates with an average of 1.6% non-red recombinants and 5.25% red contaminants in the samples. The same size and fluorescence window was used to sort two N2 WT populations collected under the same experimental conditions. Two CER276 biological replicates were collected independently without sorting, after growing for 23 hours at 20ºC. Total RNA from the six samples was extracted using TRI Reagent, followed by aqueous and organic phase separation using 1-Bromo-3-chloropropane (Sigma-Aldrich), and RNA was precipitated with 2-propanol. After treatment with TURBO DNase (Thermo Fisher Scientific), RNA was purified following a phenol/chloroform extraction protocol.

In all cases, total RNA was quantified by Qubit® RNA BR Assay kit (Thermo Fisher Scientific) and the integrity was checked by using the Agilent RNA 6000 Nano Kit (Agilent).

RNA sequencing

Four RNA-seq libraries (from CER217, CER218, CER220, and CER221 strains) were prepared with the KAPA Stranded mRNA-Seq Illumina® Platforms Kit (Roche-Kapa Biosystems) following the manufacturer's recommendations. Briefly, 500 ng of total RNA was used as the input material, the poly-A fraction was enriched with oligo-dT magnetic beads and the mRNA was fragmented. Strand specificity was achieved during the second strand synthesis performed in the presence of dUTP instead of dTTP. The blunt-ended double stranded cDNA was 3´adenylated and Illumina indexed adapters from TruSeq™ Stranded Total RNA Sample Preparation Kit with Ribo-Zero Gold (Illumina) were ligated.

The remaining six samples (two biological replicates from N2, CER276, and homozygous *sftb-1(cer6)* larvae) were prepared using the TruSeq™ Stranded Total RNA kit protocol (Illumina). Briefly, rRNA was depleted from 500 ng of total RNA using the Ribo-Zero Gold rRNA Removal Kit and fragmented by divalent cations, with a major peak at 160 nt. Following the fragmentation, first and second strand synthesis was performed, also ensuring the library directionality. The cDNA was adenylated and ligated to xGen® Dual Index UMI Adapters (IDT) for paired-end sequencing.

Both types of libraries ligation products were enriched with 15 PCR cycles and the final library was validated on an Agilent 2100 Bioanalyzer with the Agilent DNA 1000 Kit (Agilent).

The libraries were sequenced on HiSeq 2000 (Illumina) in paired-end mode with a read length of 2x76bp using the HiSeq SBS Kit V4 50 cycle kit in a fraction of a sequencing v4 flow cell lane (HiSeq PE Cluster Kit V4 - cBot). Image analysis,

base calling, and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 1.18.66.3) and followed by generation of FASTQ sequence files.

Bioinformatic analysis

RNA-seq reads were mapped against *C. elegans* reference assembly (WBcel235.87) with STAR (Dobin *et al.*, 2013). Genes and transcripts were quantified with RSEM (Li and Dewey, 2011). Differential expression analysis was performed with DESeq2 R package (Love, Huber and Anders, 2014), and detection of differential alternative splicing events was done with rMATS (Shen *et al.*, 2014).

RNAi screen

Information about the selected splicing genes was collected from SpliceosomeDB (Cvitkovic and Jurica, 2013) or from the literature. RNAi clones were obtained from the ORFeome library (Rual *et al.*, 2004) or the Ahringer library (Kamath *et al.*, 2003) and clone insert size was validated by PCR. For screening, we used 24-well plates containing NGM supplemented with 50 µg/mL ampicillin, 12.5 µg/mL tetracycline and 3 mM IPTG. Each well was seeded with 80 µl of bacterial culture and dsRNA expression was induced overnight at room temperature. WT and mutant worms were tested in duplicates. Between 10 and 20 worms synchronized at the L1 stage were seeded in each well, and phenotypes were scored visually at the adult stage (72–96 h post-seeding at 25ºC). Bacteria expressing dsRNA against *gfp* were used as negative control.

Interaction between *teg-4* and *sftb-1* mutations

Worms of the indicated genotype were singled onto individual 35-mm NGM plates with OP50 at the L4 stage and transferred to new plates every 12 hours until all eggs were laid. Numbers of dead embryos and hatched larvae were scored 24 and 48 hours after passages, respectively. Since the double mutant *sftb-1(K718E); teg-4(oz210)* could not be maintained in homozygosis, F_1 worms from *sftb-1(K718E); teg-4(oz210/+)* P_0 worms, including animals homozygous and heterozygous for *teg-4(oz210)*, were picked at the L4 stage and their progeny scored. These F_1s were genotyped after conducting the experiment to include only *sftb-1(K718E)* and *sftb-1(K718E); teg-4(oz210)* data.

Validation of candidates from the RNAi screen

In all cases, we used 55-mm or 35-mm NGM plates supplemented with 50 µg/mL ampicillin, 12.5 µg/mL tetracycline and 3 mM IPTG, seeded with specific RNAi bacterial clones. RNAi targeting *gfp* was used as negative control.

In **Figures 3c and 3d**, synchronized L1-arrested worms of the indicated genotypes were seeded onto 55-mm plates, and imaged directly on the plates after 48 hours of growth at 25°C using a stereomicroscope. Body length was determined automatically with WormSizer (Moore, Jordan and Baugh, 2013).

In **Figures 3e–g,** synchronized L1-arrested animals were seeded onto 55-mm plates and separated into individual 35-mm plates at the L4 stage. Worms were transferred every 24 hours and the presence of dead embryos was determined 24 hours after passages.

Number of germ cells assay

WT or *sftb-1(cer6)* homozygous worms were collected at the indicated time points, fixed using Carnoy's solution (60% absolute ethanol, 30% chloroform, 10% acetic acid), and washed three times in PBS + 0.1% Tween 20. Whole worms were transferred to a microscope slide and stained with DAPI Fluoromount-G® in order to visualize and count germ cell nuclei.

Spliceosome inhibitors experiments

Experiments in liquid media were conducted in 96-well plates. 25 µl of dead OP50 resuspended in a mix of freshly prepared S Medium (Stiernagle, 2006), 4 µg/ml cholesterol, 250 µg/ml streptomycin, and 62.5 µg/ml tetracycline were added to each well. Subsequently, 25 µl 2.4X drug and 10 µl L1-arrested worms at a density of 1–2 worms/µl were added to each well. Each condition was tested in duplicate in the same 96-well plate. Plates were sealed with parafilm and incubated in a humid chamber with gentle agitation at room temperature for the indicated time period. 25 µl water (experiments 1–3) or 0.5% DMSO (Sigma-Aldrich) (experiments 4 and 6) were used as negative controls. In experiment 5, 50 µM Sudemycin D6 or 0.5% DMSO were injected into the gonads or intestine of young adult hermaphrodites using standard microinjection techniques, and injected animals were recovered in individual 35-mm NGM plates with OP50 and incubated at 20ºC.

Statistical analyses

In all figure legends, 'N' denotes the number of independent replicate experiments performed, while 'n' indicates the total number of animals analyzed

in each condition (different genotypes are separated by commas). Statistical analyses were performed in GraphPad Prism 6 and R. Statistical tests used are reported in the figure legends. **Figures 1a**, **1d**, **1f–h**, **2b–e**, **3b–g**, **4**, **S2b**, and **S3c–e** were generated by using the ggplot2 R package.

Data availability

The RNA sequencing data generated in this study have been deposited in the Gene Expression Omnibus repository and are available under the accession number GEO: GSE129642.

Supplementary information

Figure S1. Most of the SF3B1 mutated residues are conserved in the *C. elegans* ortholog *sftb-1*.

Figure S2. *sftb-1(cer6)* deletion mutant.

Figure S3. *sftb-1* missense mutations design and *sftb-1[Q552P, R643C, K718E]* phenotypes at 20ºC.

Table S1. Alternative splicing analysis of *sftb-1* mutants by rMATS.

Table S2. Differential expression analysis of *sftb-1* mutants by DeSeq2.

Table S3. Information and phenotypes of the RNAi splicing screen.

Table S4. Experiments with spliceosome inhibitors.

Table S5. CRISPR reagents and injection mixes.

REFERENCES

Alsafadi, S. *et al.* (2016) 'Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage', *Nature Communications*. Nature Publishing Group, 7:10615. doi: 10.1038/ncomms10615.

An, M. and Henion, P. D. (2012) 'The zebrafish *sf3b1*^{b460} mutant reveals differential requirements for the *sf3b1* pre-mRNA processing gene during neural crest development', *International Journal of Developmental Biology*, 56(4), pp. 223–237. doi: 10.1387/ijdb.113383ma.

Arribere, J. A. *et al.* (2014) 'Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*.', *Genetics*, 198(3), pp. 837–846. doi: 10.1534/genetics.114.169730.

Biankin, A. V. *et al.* (2012) 'Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes', *Nature*, 491(7424), pp. 399–405. doi: 10.1038/nature11547.

Bonnal, S., Vigevani, L. and Valcárcel, J. (2012) 'The spliceosome as a target of novel antitumour drugs.', *Nature reviews. Drug discovery*, 11(11), pp. 847–59. doi: 10.1038/nrd3823.

Carrocci, T. J. *et al.* (2017) 'SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast', *Nucleic Acids Research*, 45(8), pp. 4837–4852. doi: 10.1093/nar/gkw1349.

Carrocci, T. J., Paulson, J. C. and Hoskins, A. A. (2018) 'Functional analysis of Hsh155/SF3b1 interactions with the U2 snRNA/branch site duplex', *RNA*, 24(8), pp. 1028–1040. doi: 10.1261/rna.065664.118.

Cretu, C. *et al.* (2016) 'Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations', *Molecular cell*, 64(2), pp. 307–319. doi: 10.1016/j.molcel.2016.08.036.

Cvitkovic, I. and Jurica, M. S. (2013) 'Spliceosome database: A tool for tracking

components of the spliceosome', *Nucleic Acids Research*, 41((Database issue)), pp. D132–D141. doi: 10.1093/nar/gks999.

Darman, R. B. *et al.* (2015) 'Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point.', *Cell reports*, 13(5), pp. 1033–1045. doi: 10.1016/j.celrep.2015.09.053.

DeBoever, C. *et al.* (2015) 'Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in *SF3B1*-mutated cancers.', *PLoS Computational Biology*, 11(3), p. e1004105. doi: 10.1371/journal.pcbi.1004105.

DeNicola, A. B. and Tang, Y. (2019) 'Therapeutic approaches to treat human spliceosomal diseases', *Current Opinion in Biotechnology*, 60, pp. 72–81. doi: 10.1016/j.copbio.2019.01.003.

Dobin, A. *et al.* (2013) 'STAR: ultrafast universal RNA-seq aligner', *Bioinformatics*, 29(1), pp. 15–21. doi: 10.1093/bioinformatics/bts635.

Dolatshad, H. *et al.* (2015) 'Disruption of *SF3B1* results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells', *Leukemia*, 29, pp. 1092–1103. doi: 10.1038/leu.2014.331.

Effenberger, K. A., Urabe, V. K. and Jurica, M. S. (2017) 'Modulating splicing with small molecular inhibitors of the spliceosome', *WIREs RNA*, 8(2), p. e1381. doi: 10.1002/wrna.1381.

Forbes, S. A. *et al.* (2017) 'COSMIC: somatic cancer genetics at high-resolution', *Nucleic Acids Research*, 45, pp. D777–D783. doi: 10.1093/nar/gkw1121.

Friedland, A. E. *et al.* (2013) 'Heritable genome editing in C. elegans via a CRISPR-Cas9 system', *Nature Methods*, 10(8), pp. 741–743. doi: 10.1038/nmeth.2532.

Furney, S. J. *et al.* (2013) '*SF3B1* mutations are associated with alternative

splicing in uveal melanoma', *Cancer Discovery*, 3(10), pp. 1122–1129. doi: 10.1158/2159-8290.CD-13-0330.

Gupta, A. K. *et al.* (2019) 'Degenerate minigene library analysis enables identification of altered branch point utilization by mutant splicing factor 3B1 (SF3B1)', *Nucleic Acids Research*, 47(2), pp. 970–980. doi: 10.1093/nar/gky1161.

Hansen, S. R. *et al.* (2019) 'Chemical Inhibition of Pre-mRNA Splicing in Living *Saccharomyces cerevisiae*', *Cell Chemical Biology*, 26(3), pp. 443–448. doi: 10.1016/j.chembiol.2018.11.008.

Harbour, J. *et al.* (2013) 'Recurrent mutations at codon 625 of the splicing factor *SF3B1* in uveal melanoma', *Nature Genetics*, 45(2), pp. 133–135. doi: 10.1038/ng.2523.

Ilagan, J. O. *et al.* (2015) 'U2AF1 mutations cause allele-specific alterations in 3' splice site recognition in myeloid malignancies', *Genome Research*, 25(1), pp. 14–26. doi: 10.1101/gr.181016.114.

Imielinski, M. *et al.* (2012) 'Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing', *Cell*, 150(6), pp. 1107–1120. doi: 10.1016/j.cell.2012.08.029.

Isono, K. *et al.* (2005) 'Mammalian Polycomb-mediated repression of *Hox* genes requires the essential spliceosomal protein Sf3b1', *Genes and Development*, 19(5), pp. 536–541. doi: 10.1101/gad.1284605.

Kamath, R. S. *et al.* (2003) 'Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi.', *Nature*, 421(6920), pp. 231–237. doi: 10.1038/nature01278.

Kerins, J. A. *et al.* (2010) 'PRP-17 and the Pre-mRNA Splicing Pathway Are Preferentially Required for the Proliferation Versus Meiotic Development Decision and Germline Sex Determination in *Caenorhabditis elegans*', *Developmental Dynamics*, 239(5), pp. 1555–1572. doi: 10.1002/dvdy.22274.

Kfir, N. *et al.* (2015) ‘SF3B1 association with chromatin determines splicing outcomes.’, *Cell reports*, 11(4), pp. 618–29. doi: 10.1016/j.celrep.2015.03.048.

Kim, H. *et al.* (2014) ‘A co-CRISPR strategy for efficient genome editing in *Caenorhabditis elegans*.’, *Genetics*. Genetics Society of America, 197(4), pp. 1069–80. doi: 10.1534/genetics.114.166389.

Lee, S. C.-W. *et al.* (2018) ‘Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations’, *Cancer Cell*, 34(2), p. 225–241.e8. doi: 10.1016/j.ccell.2018.07.003.

Li, B. and Dewey, C. N. (2011) ‘RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome’, *BMC Bioinformatics*, 12, p. 323. doi: 10.1186/1471-2105-12-323.

Liberante, F. G. *et al.* (2019) ‘Altered splicing and cytoplasmic levels of tRNA synthetases in SF3B1-mutant myelodysplastic syndromes as a therapeutic vulnerability’, *Scientific Reports*, 9(1), p. 2678. doi: 10.1038/s41598-019-39591-7.

Loerch, S. *et al.* (2018) ‘The pre-mRNA splicing and transcription factor Tat-SF1 is a functional partner of the spliceosome SF3b1 subunit via a U2AF homology motif interface.’, *The Journal of biological chemistry*, 294(8), pp. 2892–2902. doi: 10.1074/jbc.RA118.006764.

Love, M. I., Huber, W. and Anders, S. (2014) ‘Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2’, *Genome Biology*, 15(12), p. 550. doi: 10.1186/s13059-014-0550-8.

Maguire, S. L. *et al.* (2015) ‘*SF3B1* mutations constitute a novel therapeutic target in breast cancer’, *Journal of Pathology*, 235(4), pp. 571–580. doi: 10.1002/path.4483.

Malcovati, L. *et al.* (2015) ‘*SF3B1* mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts’, *Blood*, 126(2), pp. 233–241. doi: 10.1182/blood-2015-03-633537.

Mantina, P. *et al.* (2009) ‘A mutation in *teg-4*, which encodes a protein homologous to the SAP130 pre-mRNA splicing factor, disrupts the balance between proliferation and differentiation in the *C. elegans* germ line’, *Mechanisms of Development*. Elsevier Ireland Ltd, 126(5–6), pp. 417–429. doi: 10.1016/j.mod.2009.01.006.

Martin, M. *et al.* (2013) ‘Exome sequencing identifies recurrent somatic mutations in *EIF1AX* and *SF3B1* in uveal melanoma with disomy 3’, *Nature Genetics*, 45(8), pp. 933–936. doi: 10.1038/ng.2674.

McDiarmid, T. A. *et al.* (2018) ‘CRISPR-Cas9 human gene replacement and phenomic characterization in *Caenorhabditis elegans* to understand the functional conservation of human genes and decipher variants of uncertain significance’, *Disease Models & Mechanisms*, 11(12), p. dmm036517. doi: 10.1242/dmm.036517.

Moore, B. T., Jordan, J. M. and Baugh, L. R. (2013) ‘WormSizer: High-throughput Analysis of Nematode Size and Shape’, *PLoS ONE*, 8(2), p. e57142. doi: 10.1371/journal.pone.0057142.

Morton, J. J. and Blumenthal, T. (2011) ‘RNA Processing in C. elegans’, in *Methods in Cell Biology*. Second Edi. Elsevier Inc., pp. 187–217. doi: 10.1016/B978-0-12-544172-8.00007-4.

Mupo, A. *et al.* (2017) ‘Hemopoietic-specific *Sf3b1*-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts’, *Leukemia*. Nature Publishing Group, 31(3), pp. 720–727. doi: 10.1038/leu.2016.251.

Notredame, C., Higgins, D. G. and Heringa, J. (2000) ‘T-coffee: A novel method for fast and accurate multiple sequence alignment’, *Journal of Molecular Biology*, 302(1), pp. 205–217. doi: 10.1006/jmbi.2000.4042.

Obeng, E. A. *et al.* (2016) ‘Physiologic Expression of *Sf3b1*^{K700E} Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation’, *Cancer Cell*. Elsevier Inc., 30(3), pp. 404–417. doi:

10.1016/j.ccell.2016.08.006.

Pan, Q. *et al.* (2008) 'Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing', *Nature Genetics*, 40(12), pp. 1413–1415. doi: 10.1038/ng.259.

Paolella, B. R. *et al.* (2017) 'Copy-number and gene dependency analysis reveals partial copy loss of wild-type SF3B1 as a novel cancer vulnerability', *eLife*, 6, p. e23268. doi: 10.7554/eLife.23268.

Papaemmanuil, E. *et al.* (2011) 'Somatic *SF3B1* Mutation in Myelodysplasia with Ring Sideroblasts', *New England Journal of Medicine*, 365(15), pp. 1384–1395. doi: 10.1056/NEJMoa1103283.

Pellagatti, A. *et al.* (2018) 'Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations', *Blood*, 132(12), pp. 1225–1240. doi: 10.1182/blood-2018-04-843771.

Porta-de-la-Riva, M. *et al.* (2012) 'Basic *Caenorhabditis elegans Methods*: Synchronization and Observation', *Journal of Visualized Experiments*, 64, p. e4019. doi: 10.3791/4019.

Quesada, V. *et al.* (2011) 'Exome sequencing identifies recurrent mutations of the splicing factor *SF3B1* gene in chronic lymphocytic leukemia', *Nature Genetics*. Nature Publishing Group, 44(1), pp. 47–52. doi: 10.1038/ng.1032.

R Core Team (2018) 'R: A Language and Environment for Statistical Computing'. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.r-project.org/.

te Raa, G. *et al.* (2015) 'The impact of *SF3B1* mutations in CLL on the DNA-damage response', *Leukemia*, 29(5), pp. 1133–1142. doi: 10.1038/leu.2014.318.

Ragle, J. M. *et al.* (2015) 'Coordinated tissue-specific regulation of adjacent alternative 3' splice sites in *C. elegans*.', *Genome research*, 25(7), pp. 982–94.

doi: 10.1101/gr.186783.114.

Rual, J.-F. *et al.* (2004) 'Toward Improving Caenorhabditis elegans Phenome Mapping With an ORFeome-Based RNAi Library', *Genome Research*, 14, pp. 2162–2168. doi: 10.1101/gr.2505604.7.

Schindelin, J. *et al.* (2012) 'Fiji: an open-source platform for biological-image analysis.', *Nature methods*, 9(7), pp. 676–82. doi: 10.1038/nmeth.2019.

Seiler, M., Yoshimi, A., *et al.* (2018) 'H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers', *Nature Medicine*, 24(4), pp. 497–504. doi: 10.1038/nm.4493.

Seiler, M., Peng, S., *et al.* (2018) 'Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types', *Cell Reports*, 23(1), p. 282–296.e4. doi: 10.1016/j.celrep.2018.01.088.

Shen, S. *et al.* (2014) 'rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(51), pp. E5593-601. doi: 10.1073/pnas.1419161111.

Shiozawa, Y. *et al.* (2018) 'Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia', *Nature Communications*, 9(1), p. 3649. doi: 10.1038/s41467-018-06063-x.

Stemmer, M. *et al.* (2015) 'CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool.', *PloS one*, 10(4), p. e0124633. doi: 10.1371/journal.pone.0124633.

Stiernagle, T. (2006) 'Maintenance of *C. elegans*.', in *WormBook*, pp. 1–11. doi: 10.1895/wormbook.1.101.1.

Teng, T. *et al.* (2017) 'Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A-SF3b complex', *Nature Communications*. Nature Publishing Group, 8, p. 15522. doi: 10.1038/ncomms15522.

Tourasse, N. J., Millet, J. R. M. and Dupuy, D. (2017) ‘Quantitative RNA-seq meta-analysis of alternative exon usage in *C. elegans*’, *Genome Research*, 27(12), pp. 2120–2128. doi: 10.1101/gr.224626.117.

Urbanski, L. M., Leclair, N. and Anczuków, O. (2018) ‘Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics’, *WIREs RNA*, 9(4), p. e1476. doi: 10.1002/wrna.1476.

Vicencio, J. *et al.* (2019) ‘Efficient Generation of Endogenous Fluorescent Reporters by Nested CRISPR in *Caenorhabditis elegans*’, *Genetics*, 211(4), pp. 1143–1154. doi: 10.1534/genetics.119.301965.

Waaijers, S. *et al.* (2013) ‘CRISPR/Cas9-targeted mutagenesis in *Caenorhabditis elegans*’, *Genetics*, 195(3), pp. 1187–1191. doi: 10.1534/genetics.113.156299.

Wang, L. *et al.* (2011) ‘*SF3B1* and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia’, *New England Journal of Medicine*, 365, pp. 2497–2506. doi: 10.1056/nejmoa1109016.

Wang, L. *et al.* (2016) ‘Transcriptomic Characterization of *SF3B1* Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia’, *Cancer Cell*. Elsevier Inc., 30(5), pp. 750–763. doi: 10.1016/j.ccell.2016.10.005.

Waterhouse, A. M. *et al.* (2009) ‘Jalview Version 2-A multiple sequence alignment editor and analysis workbench’, *Bioinformatics*, 25(9), pp. 1189–1191. doi: 10.1093/bioinformatics/btp033.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. doi: 10.1007/978-3-319-24277-4.

Wu, G. *et al.* (2018) ‘Inhibition of SF3B1 by molecules targeting the spliceosome results in massive aberrant exon skipping’, *RNA*, 24(8), pp. 1056–1066. doi: 10.1261/rna.065383.117.

Xu, J. J. *et al.* (2019) ‘Modeling human RNA spliceosome mutations in the mouse: not all mice were created equal’, *Experimental Hematology*, 70, pp. 10–23. doi: 10.1016/j.exphem.2018.11.001.

Yin, S. *et al.* (2019) ‘A Murine Model of Chronic Lymphocytic Leukemia Based on B Cell-Restricted Expression of *Sf3b1* Mutation and *Atm* Deletion’, *Cancer Cell*, 35(2), p. 283–296.e5. doi: 10.1016/j.ccell.2018.12.013.

Yoshida, K. *et al.* (2011) ‘Frequent pathway mutations of splicing machinery in myelodysplasia’, *Nature*, 478(7367), pp. 64–69. doi: 10.1038/nature10496.

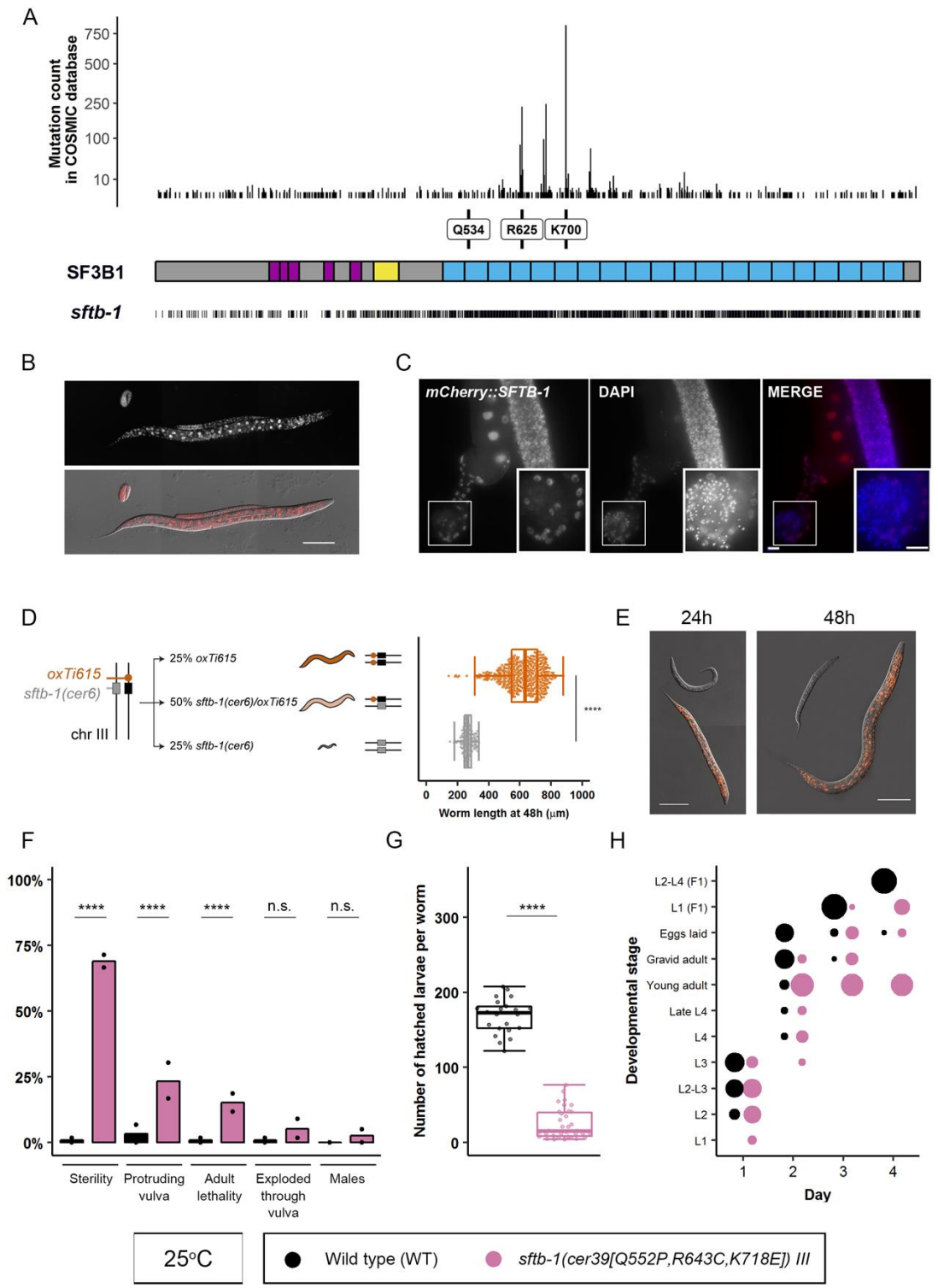
A
Mutation count in COSMIC database
750
500
250
100
10
Q534
R625
K700
SF3B1
sftb-1
B
C
mCherry::SFTB-1
DAPI
MERGE
D
oxTi615
sftb-1(cer6)
chr III
25% oxTi615
50% sftb-1(cer6)/oxTi615
25% sftb-1(cer6)
****
0
200
400
600
800
1000
Worm length at 48h (μm)
E
24h
48h
F
100%
75%
50%
25%
0%
****
****
****
n.s.
n.s.
Sterility
Protruding vulva
Adult lethality
Exploded through vulva
Males
G
Number of hatched larvae per worm
300
200
100
0
****
H
Developmental stage
L2-L4 (F1)
L1 (F1)
Eggs laid
Gravid adult
Young adult
Late L4
L4
L3
L2-L3
L2
L1
1
2
3
4
Day
25°C
Wild type (WT)
sftb-1(cer39[Q552P,R643C,K718E]) III

Figure 1

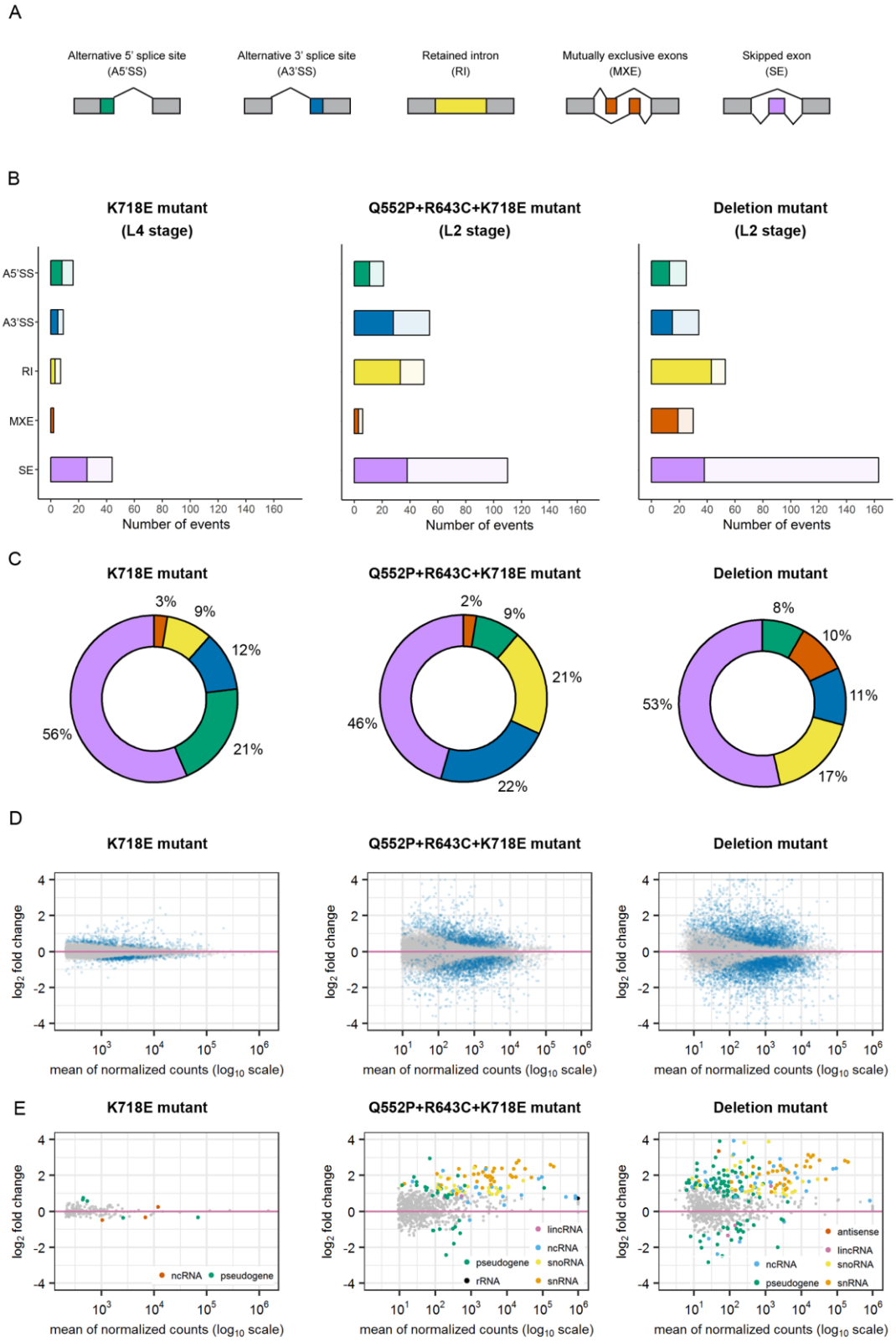
A
Alternative 5' splice site (A5'SS)
Alternative 3' splice site (A3'SS)
Retained intron (RI)
Mutually exclusive exons (MXE)
Skipped exon (SE)
B
K718E mutant (L4 stage)
Q552P+R643C+K718E mutant (L2 stage)
Deletion mutant (L2 stage)
A5'SS
A3'SS
RI
MXE
SE
Number of events
C
K718E mutant
3%
9%
12%
21%
56%
Q552P+R643C+K718E mutant
2%
9%
21%
22%
46%
Deletion mutant
8%
10%
11%
17%
53%
D
K718E mutant
Q552P+R643C+K718E mutant
Deletion mutant
log2 fold change
mean of normalized counts (log10 scale)
E
K718E mutant
ncRNA
pseudogene
Q552P+R643C+K718E mutant
lincRNA
ncRNA
pseudogene
snoRNA
rRNA
snRNA
Deletion mutant
antisense
lincRNA
ncRNA
snoRNA
pseudogene
snRNA

Figure 2

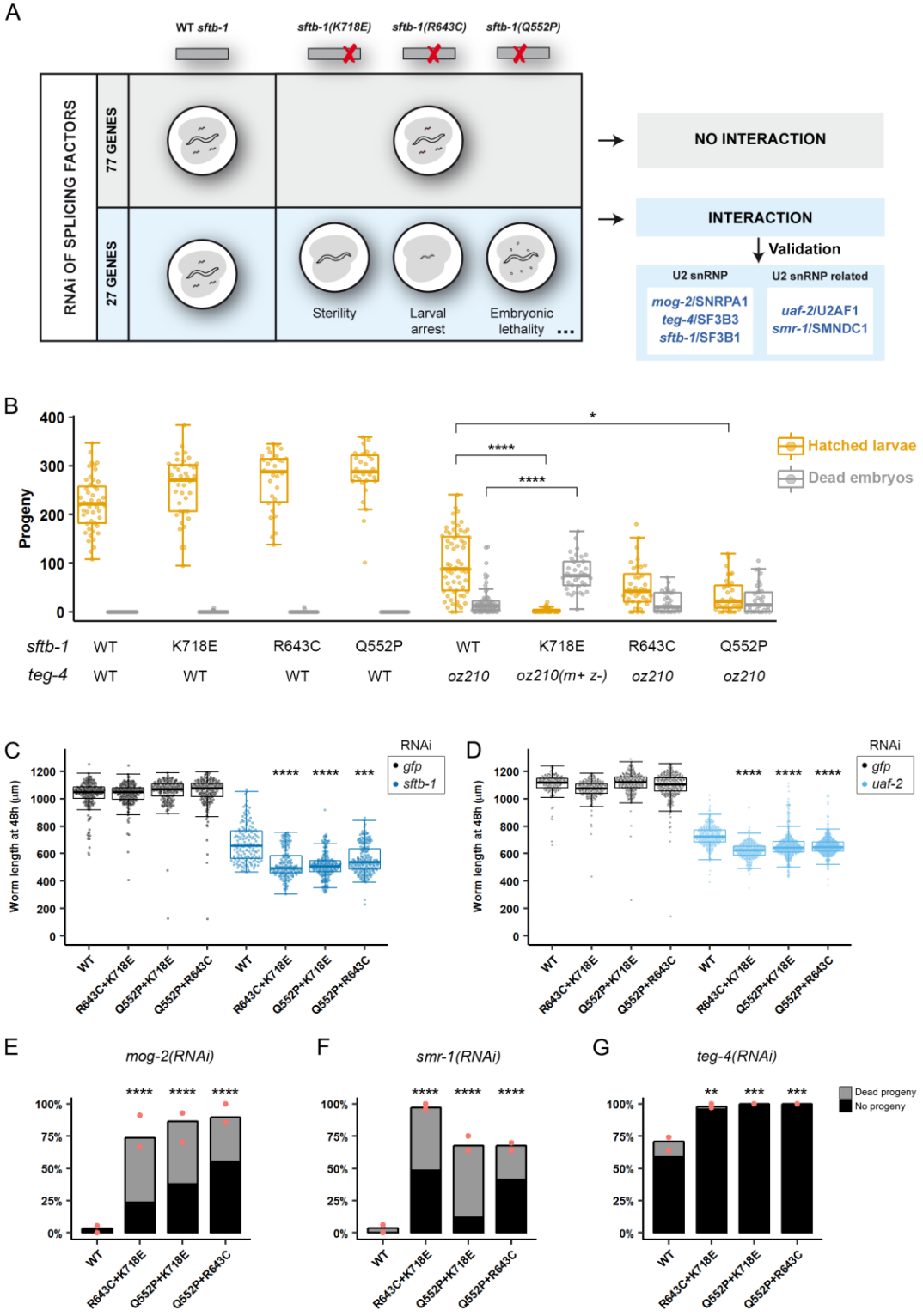
A
WT sftb-1
sftb-1(K718E)
sftb-1(R643C)
sftb-1(Q552P)
RNAi OF SPLICING FACTORS
77 GENES
27 GENES
Sterility
Larval arrest
Embryonic lethality ...
NO INTERACTION
INTERACTION
Validation
U2 snRNP
U2 snRNP related
mog-2/SNRPA1
teg-4/SF3B3
sftb-1/SF3B1
uaf-2/U2AF1
smr-1/SMNDC1
B
Progeny
Hatched larvae
Dead embryos
sftb-1 WT K718E R643C Q552P WT K718E R643C Q552P
teg-4 WT WT WT WT oz210 oz210(m+ z-) oz210 oz210
C
Worm length at 48h (µm)
RNAi
gfp
sftb-1
D
uaf-2
WT
R643C+K718E
Q552P+K718E
Q552P+R643C
E
mog-2(RNAi)
F
smr-1(RNAi)
G
teg-4(RNAi)
Dead progeny
No progeny

Figure 3

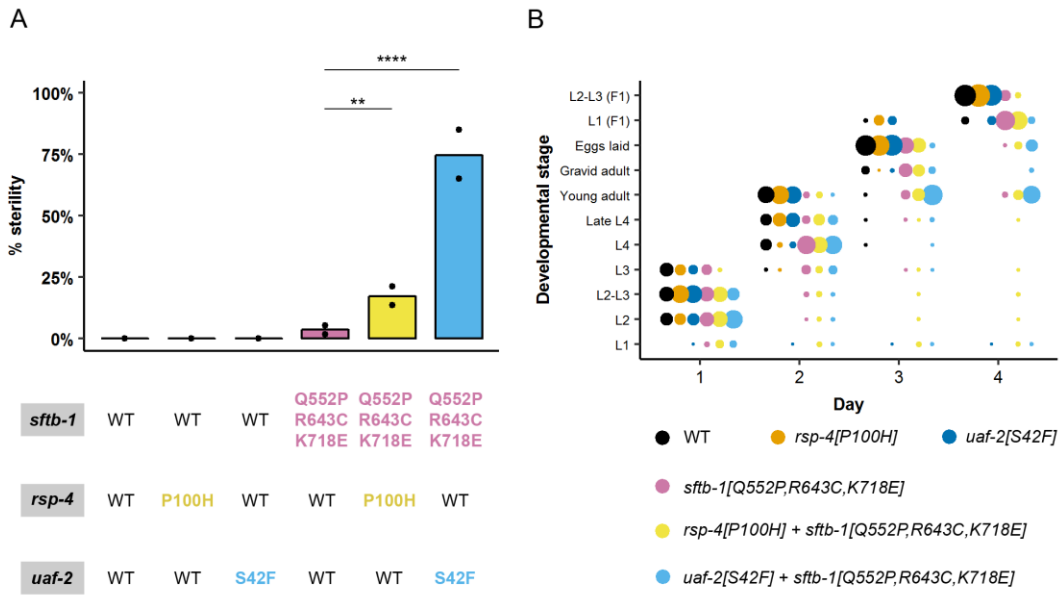
A
B
100%
75%
50%
25%
0%
% sterility
****
**
sftb-1
WT
WT
WT
Q552P R643C K718E
Q552P R643C K718E
Q552P R643C K718E
rsp-4
WT
P100H
WT
WT
P100H
WT
uaf-2
WT
WT
S42F
WT
WT
S42F
Developmental stage
L2-L3 (F1)
L1 (F1)
Eggs laid
Gravid adult
Young adult
Late L4
L4
L3
L2-L3
L2
L1
1
2
3
4
Day
WT
rsp-4[P100H]
uaf-2[S42F]
sftb-1[Q552P,R643C,K718E]
rsp-4[P100H] + sftb-1[Q552P,R643C,K718E]
uaf-2[S42F] + sftb-1[Q552P,R643C,K718E]

Figure 4