4

PNN-1 and UAF-1 link RNA splicing to DNA repair by Non-Homologous End Joining

Lemmens BB, van Schendel R, Tijsterman M.

Abstract

Non-homologous End Joining (NHEJ) is the major DNA double strand break (DSB) repair route in somatic tissues and is vital to ensure genomic stability and proper animal development. Accordingly, NHEJ deficiency syndromes are characterized by severe developmental abnormalities and hypersensitivity to DSB-inducing agents such as ionizing radiation (IR). Given its highly effective but error-prone nature, NHEJ needs to be tightly regulated during development. In recent years RNA binding proteins (RBPs) have emerged as important regulator of genome stability, yet how and if they control DSB repair is unknown to date. Here, we constructed a transgenic reporter assay in *C. elegans* that allows detection of NHEJ activity *in vivo* and performed unbiased forward genetics screens to identify novel regulators of NHEJ. We found the THO ribonucleoprotein complex (*thoc-2*, *thoc-5* and *thoc-7*) and the RNA splicing regulator Pinin (*pnn-1*) to be required for efficient NHEJ and IR resistance in somatic tissues. In-depth transcriptome analysis revealed strikingly similar RNA expression alterations among the NHEJ mutants, including exon-specific splicing defects. Moreover, we found THO mutants to suffer from reduced expression of the essential splicing factor U2AF65 (*uaf-1*) of which depletion by RNA interference mimics the DSB repair defects in THO mutants. The identification of the splicing factors PNN-1 and UAF-1 in NHEJ regulation sets the stage for further dissection of RNA processing mechanisms during animal development and implicates potentially new roles of alternative RNA transcripts in DSB repair.

Introduction

During animal development cells are exposed to numerous DNA damaging agents that may hamper cellular function, ultimately leading to developmental defects, pathologies and reduced fitness. One of the most toxic DNA lesions a cell can encounter is a DNA double strand break (DSB); left unrepaired a DSB can cause chromosome segregation defects and cell death and its incorrect repair can lead to gross chromosomal aberrations, including deletions and translocations that promote oncogenic transformation (McKinnon and Caldecott 2007). To neutralize the toxic effects of DSBs an elaborate network of proteins has evolved that either can repair the damage or minimalize the consequences for animal development (Phillips and McKinnon 2007). Recent proteomic studies have revealed a vast amount of proteins being post-translationally modified upon DSB induction, including factors required for DSB repair, cell cycle arrest, chromatin modification and apoptosis (Aslanian *et al.* 2014; Beli *et al.* 2012; Jungmichel and Blasius 2013; Jungmichel *et al.* 2013; Matsuoka *et al.* 2007). Surprisingly, another major class of proteins modified during the DNA damage response consists of RNA binding proteins (RBPs), yet if and how these RBPs affect DSB repair is still an open question (Dutertre *et al.* 2014; Lenzken *et al.* 2013).

Eukaryotic cells posses a versatile toolbox of DNA repair factors that can act in different chromatin contexts, cell cycle stages and on a wide range of DNA substrates and repair templates. Accordingly, DSBs can be repaired via several different repair mechanisms, including non-homologous end joining (NHEJ), homologous recombination (HR) or single strand annealing (SSA), which all have unique abilities and limitations. For instance, NHEJ can seal DSBs independent of DNA sequence context, while both HR and SSA require a homologous DNA template to repair DSBs.

NHEJ is the major DSB repair route in human cells and also the pathway of choice in somatic tissues of the animal model *Caenorhabditis elegans (C. elegans) (Clejan et al. 2006; Lieber 2010)*. The core NHEJ machinery is well conserved and is based on DSB recognition by the Ku70/Ku80 heterodimer (CKU-70/CKU-80 in *C. elegans*) and subsequent recruitment of DNA ligase IV (LIG-4), which seals the break. Although this pathway is very efficient, there is no quality control and occasionally some nucleotides are lost or inserted, making it an error-prone repair route. Reduced *in vivo* NHEJ activity is associated with several human developmental disorders and is characterized by cellular radiosensitivity, microcephaly and severe immunodeficiency due to loss of NHEJ-dependent genetic variation among lymphocytes (McKinnon and Caldecott 2007). Although DSB repair via NHEJ is crucial to maintain genomic stability during development, it can also have detrimental toxic effects when left unrestrained. In fact, uncontrolled NHEJ activities can corrupt error-free repair by HR, promoting chromosomal abnormalities, tumorigenesis and infertility in animals (Adamo *et al.* 2010; Bunting *et al.* 2010; Lemmens *et al.* 2013). The activity of DSB repair pathways can

be regulated in many ways, for instance at the level of the DNA substrate chromatin status or protein modification, but also at the level of mRNA expression, localization and stability (Chapman *et al.* 2012; Hu and Gatti 2011; Lenzken *et al.* 2013).

In order to identify new regulators of NHEJ activity during animal development we performed an unbiased forward genetics screen in *C. elegans*. We constructed a dual reporter-based assay that allowed us to read-out NHEJ activity *in vivo* and identified both known NHEJ genes as well as factors not implicated in NHEJ before. Notably, we found three components of the THO ribonucleoprotein complex (*thoc-2*, *thoc-5* and *thoc-7*) and splicing-regulator Pinin (*pnn-1*) to be required for NHEJ in *C. elegans*, implying an important role for RBPs for NHEJ efficacy during development. Next to defective repair of endonuclease-induced DSBs, all mutants identified were hypersensitive to ionizing radiation (IR) during development, a characteristic also seen in NHEJ deficient human patients (Dvorak and Cowan 2010). Epistasis analysis with canonical NHEJ mutants confirmed a specific requirement for repair via NHEJ for these RBPs. To investigate possible effects on mRNA stability of canonical NHEJ factors, we performed genome-wide RNAseq in *thoc-5*, *thoc-7* and *pnn-1* mutants and found a small but very specific subset of transcripts to be affected in these mutant backgrounds. In-dept expression analysis revealed striking similarity between both THO and PNN-1 mutants including exon-specific splicing defects. Interestingly, we found the expression of the essential splicing factor UAF-1/U2AF to be reduced in THO mutants and UAF-1 depletion mimics the DSB repair defects in THO mutants, including reduced NHEJ and increased SSA activity. All together these results reveal a close link between mRNA splicing and *in vivo* NHEJ activity and set the stage for the analysis of potential novel mRNA isoforms that control DSB repair efficacy.

Results

Dual reporter system to measure NHEJ activity in *C. elegans*

In recent years the *C. elegans* model system has become a powerful tool to study DNA damage responses in a developmental context, leading to important insights in DSB repair in somatic as well as germline tissues (Lemmens and Tijsterman 2011). While classical DNA damage sensitivity assays and transgenic reporter systems uncovered robust and tissue-specific DSB repair mechanisms (Clejan *et al.* 2006; Pontier and Tijsterman 2009), cytological approaches revealed several regulators of DSB repair pathway choice, including suppressors of NHEJ (Adamo *et al.* 2010; Lemmens *et al.* 2013). We recently developed a variety of transgenic reporter systems based on site-specific DSB induction by the I-SceI meganuclease, which allowed us to measure both HR and SSA activities during worm development (Johnson *et al.* 2013; Pontier and Tijsterman 2009). Although these reporter systems can be used to examine HR and SSA activity directly, they can also be used to study NHEJ in an indirect manner, given that NHEJ defects result in a stark increase in compensatory pathways such as HR and SSA

(Johnson *et al.* 2013; Pontier and Tijsterman 2009). Yet such indirect measurements can be obscured by other repair defects that result in similar shifts in pathway usage, e.g. defective HR can also increase SSA activity.

We therefore developed a new I-SceI-based reporter that measures classical NHEJ activity directly (Figure 1A). This NHEJ reporter is based on restored expression of an out-of-frame GFP/LacZ sequence due to mutagenic repair of an upstream I-SceI target site: error-prone repair by NHEJ resulting in +1 or -2 frame shifts results in pharyngeal GFP/LacZ expression. We chose to measure DSB repair in pharyngeal muscle tissue because these cells are already terminally differentiated at the first larval stage and in contrast to cycling cells lack alternative end joining activities (D. Pontier and M. Tijsterman unpublished data). This provides us the unique possibility to measure mutagenic end joining *in vivo* that depends completely on classical NHEJ. Accordingly, DSB induction during larval development resulted in robust pharyngeal GFP/LacZ expression that was completely abolished by *lig-4* mutation (Figure 1B). To internally control for NHEJ efficacy within the same animal we combined the NHEJ reporter transgene with a well-characterized SSA reporter transgene, creating a dual reporter system that can detect *in vivo* NHEJ activity directly (by measuring NHEJ in non-replicating pharyngeal cells) and indirectly (by measuring SSA in various replicating somatic cells). We combined both reporter transgenes with a heat-shock inducible mCherry I-SceI fusion transgene allowing us to govern and monitor nuclear meganuclease expression (Figure 1A and 1C).

Inactivating NHEJ in dual reporter animals by *lig-4* mutation resulted in a dramatic shift in LacZ staining pattern: while LacZ staining disappeared in the pharynx, LacZ staining increased strongly in various other somatic tissues. Moreover, this pattern was extremely robust among reporter animals: while 0% of *lig-4* mutants showed GFP/LacZ positive pharynxes, nearly 100% of wild-type animals showed pharyngeal GFP/LacZ ORF correction within 24 hours post DSB induction (Figure 1D).

Forward genetics screen for regulators of NHEJ

Having such a distinct phenotype depending on classical NHEJ provided us with the opportunity to search for possible new regulators of NHEJ. To this end, we performed forward genetics screens in which we induced random mutations in dual reporter animals by ethyl methanesulfonate (EMS) and assessed NHEJ activity among mutant progeny (*i.e.* complex F3 populations were screened for animals with reduced pharyngeal GFP signal) (Figure 1E and S1). These GFP^{low} animals were selected and their clonal progeny was again tested for NHEJ activity to exclude stochastic events and identify heritable traits that affect pharyngeal GFP expression. To exclude false-positive mutants with reduced DSB induction or defective NHEJ transgenes we analyzed mCherry-I-SceI expression as well as SSA activity in all mutant candidates. Only NHEJ mutants that also showed increased SSA activity, indicative of repair defects post DSB induction, were selected for further analysis (Figure 1E).

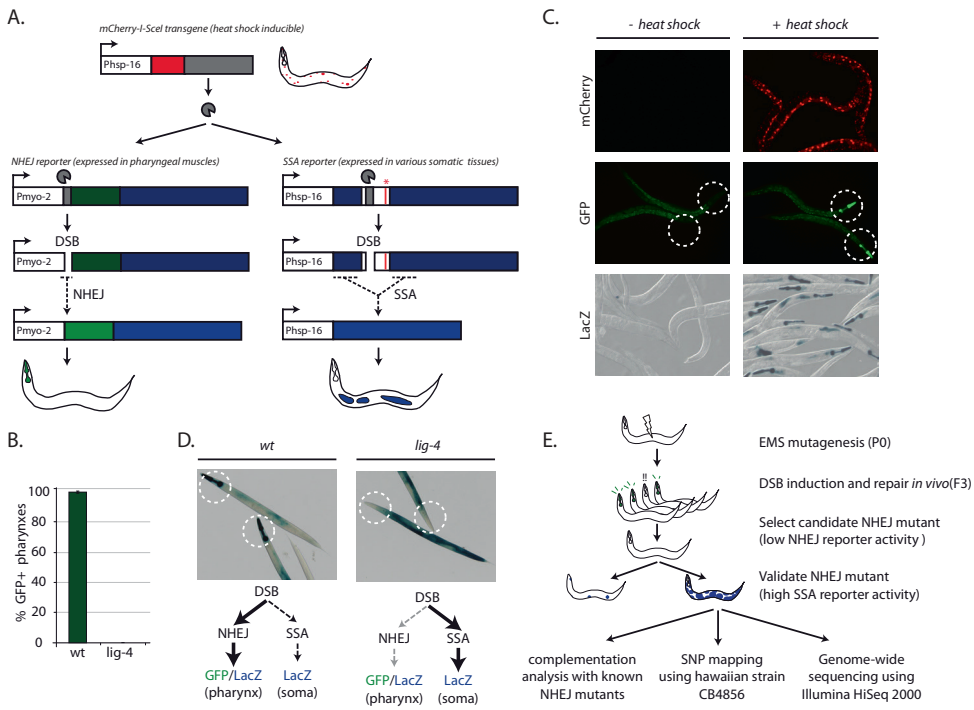

Figure 1. Dual reporter assay to measure NHEJ activity in developing *C. elegans*
A. Schematic diagram of dual reporter system based on heat-shock-inducible expression of mCherry::I-SceI and DSB repair-mediated ORF restoration at two multi-copy reporter transgenes. Mutagenic NHEJ is measured by GFP/LacZ ORF correction in non-dividing pharyngeal cells, while SSA is measured by LacZ ORF correction in dividing somatic cells. Asterisk indicates stop codon. **B.** Quantification of GFP-positive pharynxes in synchronized wild-type and *lig-4* deficient dual reporter animals, heat-shocked for 180 minutes at L1 stage and measured in adults. Average of three populations (n>200) is depicted; error bars represent S.E.M. **C.** Representative pictures of animals expressing nuclear mCherry::I-SceI (6 hours post adult heat-shock), pharyngeal GFP (3 days post L1 heat-shock, yet prior to adult heat-shock) and somatic LacZ (3 hours post adult heat-shock). Synchronized animals were heat-shocked for 180 minutes at L1 stage (to induce I-SceI expression /DSB formation) and at adult stage (to express the SSA reporter). **D.** LacZ expression patterns 3 hours post adult heat-shock of synchronized wild-type and *lig-4* deficient dual reporter animals, heat-shocked for 180 minutes at L1 and adult stage. **E.** Setup of forward genetics screen (more elaborate setup is depicted in Figure S1).

By screening ~9000 haploid genomes we found seven *bona fide* NHEJ mutants: four showing reduced GFP expression and three showing no pharyngeal GFP at all. We next quantified both NHEJ activity (GFP expression) and SSA activity (somatic LacZ expression) of synchronized mutant populations and observed the expected inverse correlation between NHEJ defect severity and increased compensatory SSA activity: *lf151*, *lf152* and *lf153* having no detectable NHEJ activity and a strong (>4-fold) increase in SSA activity, and *lf158*, *lf160* and *lf161* having detectable but significantly reduced NHEJ activity and a milder (~3-fold) increase in SSA activity (Figure 2A). The seventh mutant *lf159* was also characterized in detail

but these results will be discussed later. All together these results indicate that our screening setup allowed us to find null alleles for essential NHEJ factors as well as modifier alleles that either partially impair essential NHEJ factors or fully block the function of non-essential NHEJ regulators.

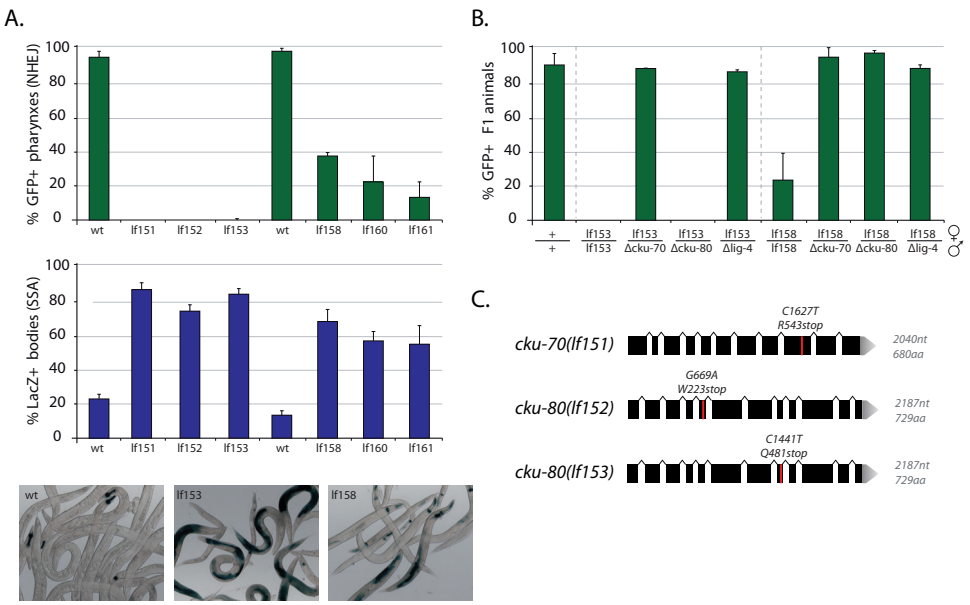

Figure 2. Characterization and complementation analysis of novel NHEJ mutants
A. Quantification of GFP-positive pharynxes and non-pharyngeal LacZ staining in clonal mutant populations of synchronized dual reporter animals, heat-shocked for 120 minutes at L1 and adult stage and measured in adults. Average of three populations (n>200) is depicted; error bars represent S.E.M. Lower panels depict LacZ expression patterns three hours post adult heat-shock. **B.** Quantification of GFP-positive pharynxes in trans-heterozygous F1 cross progeny, heat-shocked for 120min and measured in adults. Average percentage of GFP-positive pharynxes of three independent F1 populations is depicted. **C.** Gene models and newly identified alleles of *cku-70* and *cku-80*.

Screen validation

In order to identify the causal mutations and test if the new alleles affected the function of known NHEJ genes in *C. elegans*, we performed complementation analysis using null alleles of *lig-4*, *cku-80* and *cku-70*. All six mutants were crosses with *lig-4*, *cku-80* and *cku-70* mutant males and the trans-heterozygous F1 progeny was tested for NHEJ activity. As illustrated in Figure 2B, NHEJ activity of the *lf153* allele could be rescued by *lig-4* and *cku-70* mutant males, but not but *cku-80* mutants, implying that *lf153* affected *cku-80* function. Indeed, sequencing analysis of *lf153* mutants revealed a typical EMS induced C>T transversion in *cku-80*, leading to a premature stop codon that prevents expression of the well conserved C-terminus of CKU-80. Similarly, we found the NHEJ defect in *lf151* and *lf152* mutants to be caused by nonsense

mutations in *cku-70* and *cku-80*, respectively (Figure 2C). The identification of canonical NHEJ genes using this unbiased approach validated both our dual reporter system as well as our screening setup. Interestingly, although the NHEJ defect of all three null alleles *lf151, lf15*2 and *lf153* could be explained by mutations in known NHEJ genes, the NHEJ defect of the modifier alleles *lf158*, *lf160* and *lf161* was not caused by mutations in canonical NHEJ genes. For instance, paternally derived genomes containing *cku-70*, *cku-80* or *lig-4* null alleles could still restore NHEJ activity in *lf158* trans-heterozygotes (Figure 2B). These results suggest that we also picked up mutations in genes not described before to regulate NHEJ during *C. elegans* development.

THO complex is required for efficient NHEJ in *C. elegans*

To find the causal mutations in the NHEJ modifier strains we combined classical positional mapping and next generation sequencing techniques. We mapped the chromosomal regions responsible for the NHEJ defect using single nucleotide polymorphisms (SNPs) between a Hawaiian mapping strain CB4856 and Bristol dual reporter mutants, and found the causal genes of *lf158* and *lf161* to be located on the center of chromosome III and *lf161* on the right arm of chromosome IV (Figure 3A). Although located on the same chromosome, complementation analysis between *lf158* and *lf161* revealed that the causal mutations were in different genes (data not shown). Parallel to the mapping studies we used whole-genome sequencing to find non-synonymous SNPs that were unique for every mutant and were located in the mapped regions, culminating to a relatively small set of 4-7 candidate genes per mutant. Strikingly, we found that all three mutants carried a candidate mutation in a gene belonging to the THO complex, strongly suggesting that defective THO complex function was causing reduced NHEJ activity in *C. elegans* (Figure 3A).

The THO complex is a highly conserved ribonucleoprotein complex that has been implicated in transcription elongation, non-coding RNA metabolism, mRNA splicing and mRNA export (Luna *et al.* 2012). Intriguingly, defective THO complex function is shown to result in genome instability in various species, including yeast, worms and man, which in part could be explained by its role in preventing the formation of RNA:DNA hybrids throughout the genome (Castellano-Pozo *et al.* 2012b; Dominguez-Sanchez *et al.* 2011; Huertas and Aguilera 2003). The identification of several THO mutants in our NHEJ screen adds another possible link between these conserved RBPs and genome instability.

Here we identified a missense mutation in *thoc-2* (*lf158*) and *thoc-7* (*lf160*) and a nonsense mutation in *thoc-5* (*lf161*). While both the *thoc-5* and *thoc-7* alleles affected highly conserved residues and thus may reflect null alleles, the *thoc-2* allele affected a non-conserved amino acid, likely retaining some THOC-2 activity (Figure 3B). In fact, *lf158* animals are fertile while *thoc-2* null mutants are reported to be completely sterile, suggesting that we acquired either a hypomorfic or separation-of-function allele of *thoc-2* (Castellano-Pozo *et al.* 2012b).

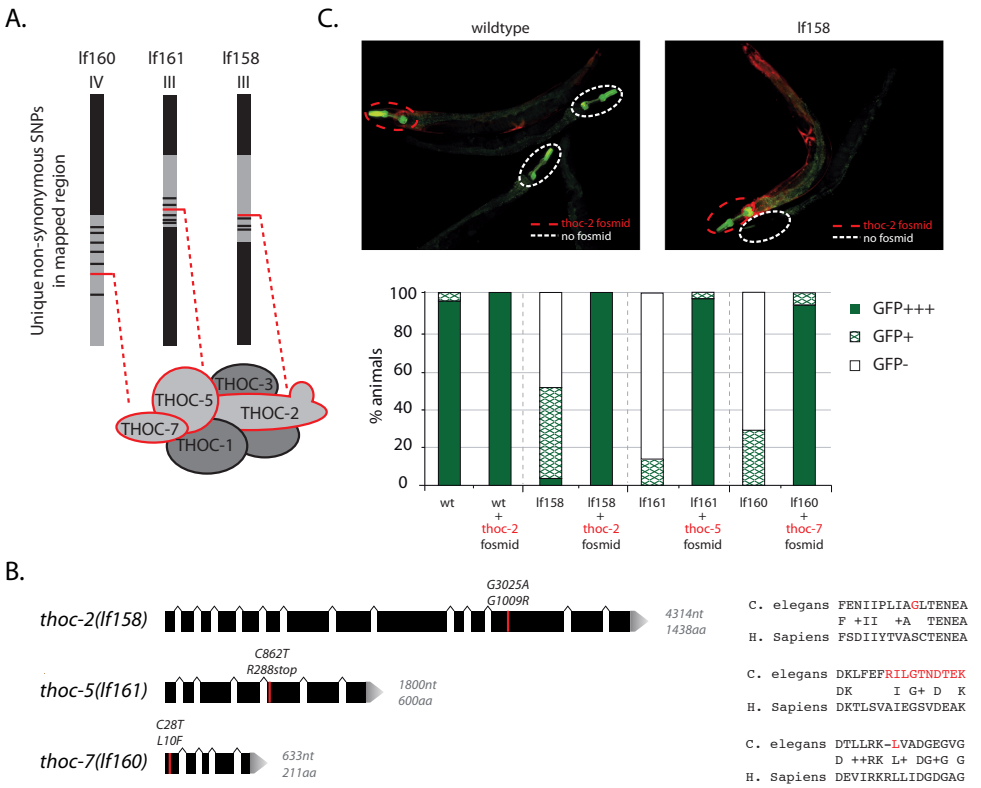

Figure 3. THO mutations cause reduced NHEJ activity in *C. elegans*
A. Schematic diagram of chromosomes III and IV in three NHEJ mutants; regions found to cause NHEJ defect based on Hawaiian SNP mapping in grey; black/red horizontal lines represent unique non-synonymous mutations. All three mutants bear mutations in THO genes **B.** Gene models and newly identified alleles of *thoc-2*, *thoc-5* and *thoc-7*. Right panel depicts amino acid context and conservation of THO mutations. **C.** Complementation analysis using fosmid arrays carrying wild-type THO genes and mCherry expression markers. Representative picture of synchronized populations of wildtype and lf158 animals with (red circle) or without (white circle) *thoc-2* fosmid array; animals were heat-shocked for 180 minutes at L1 stage. Histogram shows quantification of GFP-positive pharynxes in adults of the different genetic backgrounds. NHEJ activity was restored to wild-type levels by complementing the THO mutants with functional THO genes. Average percentage of GFP-positive pharynxes of three independent populations (n>150) is depicted.

To establish causality between the THO complex mutations and *in vivo* NHEJ efficacy we set out to complement the mutant dual reporter worms with functional THO genes. We performed fosmid microinjections to create animals carrying inheritable extra-chromosomal arrays harboring wild-type gene products of the mutated THO genes (Figure S2). These extra-chromosomal arrays also carried mCherry expression markers to be able to identify transgenic animals that express the functional THO genes. We next crossed the mCherry-marked fosmid arrays to the corresponding mutants: a *thoc-2* expression array to *lf158* animals, a *thoc-5*

expression array to *lf161* animals, and a *thoc-7* expression array to *lf161* animals. Both THO mutant and wild-type counterparts were isolated from each cross and NHEJ activity was measured in mCherry positive as well as mCherry negative progeny. Importantly, the presence of functional THO genes NHEJ activity in all three corresponding mutants, indicating that the NHEJ defect in these animals was indeed caused by impaired THO complex function (Figure 3C). Furthermore, we confirmed NHEJ restoration in complemented animals by analyzing SSA reporter activity and found somatic LacZ expression to be reduced in complemented animals ($mCherry^+$) compared to THO mutant controls ($mCherry^-$), further supporting a specific role for THO complex genes in NHEJ and DSB repair dynamics (Figure S3A). Moreover, we could increase SSA activity by reducing THOC-2 expression by RNA interference, suggesting that reduced THO gene expression, like THO gene mutations, affects NHEJ efficacy *in vivo* (Figure S3B).

PCR-based assay confirms role for the THO complex in mutagenic NHEJ

We next assayed DSB repair more directly and independent of reporter transcription via a PCR-based assay (Figure 4A). We heat-shocked synchronized L1 larva of two wild-type controls and four different NHEJ mutants (*i.e.* two Ku complex mutants and two THO complex mutants) to induce I-SceI expression and create DSBs at the NHEJ reporter locus. After allowing *in vivo* DSB repair for 24 hours, we isolated genomic DNA of all larva and PCR-amplified the genomic region surrounding the I-SceI target site. Subsequently, the PCR products were digested *in vitro* with recombinant I-SceI enzyme and restriction products were resolved on a polyacrylamide gel. As shown in Figure 4B, PCR products derived from non-heat-shocked animals were all susceptible to I-SceI cleavage, which indicated that the I-SceI site was intact in all the strains prior to DSB induction. However, after heat-shock induction different digestion patterns became apparent: while the majority of PCR products (~67%) derived from wild-type control animals became resistant to I-SceI cleavage (suggestive of mutagenic repair events at the DSB site), nearly all PCR products derived from both Ku mutants were susceptible to I-SceI cleavage (Figure 4C). This indicates that the resistant fraction in wild-type animals is dependent on Ku and likely reflects the error-pone activity of classical NHEJ. Notably, the majority of PCR products derived from both *thoc-5* and *thoc-7* mutants was still susceptible to I-SceI cleavage and only a minor fraction was I-SceI resistant (~36%), indicating that some mutagenic NHEJ activity is still present in these animals but that this activity is significantly reduced compared to wild-type animals (Figure 3C). Accordingly, both these THO mutants are able to restore GFP expression in an I-SceI dependent manner, but do so at a severely reduced level compared to wild-type animals (Figure 1 and 2).

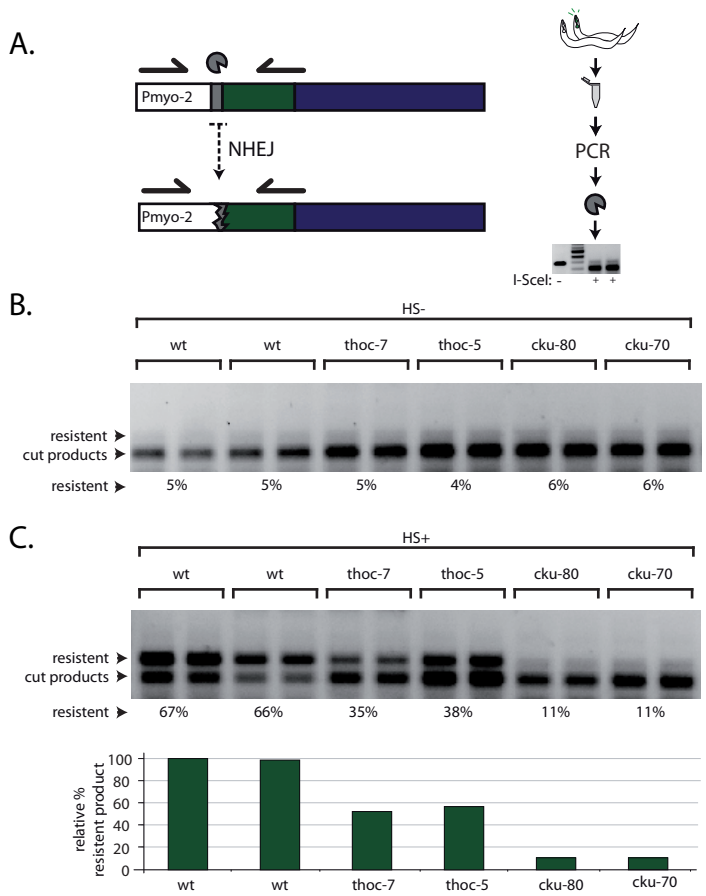

Figure 4. PCR-based assay confirms reduced mutagenic DSB repair in THO mutants
A. Setup of PCR-based assay to measure mutagenic DSB repair at NHEJ reporter transgene. Genomic I-SceI target sites of synchronized heat-shocked populations were PCR amplified and digested *in vitro* with recombinant I-SceI enzyme. Mutagenic repair of I-SceI-induced DSBs *in vivo* creates I-SceI resistant PCR products. **B.** Representative gel and band intensity quantification of I-SceI resistant fractions of synchronized untreated animals (non-heat shock controls). **C.** Representative gel and band intensity quantification of I-SceI resistant fractions of synchronized animals heat-shocked twice for 180 minutes. Histogram depicts I-SceI resistant fractions relative to wild-type controls (reflecting mutagenic DSB repair activity).

THO deficient somatic tissues are hypersensitive to IR

Given the NHEJ defect at endonuclease-induced DSBs in THO mutant animals, we wondered if the THO complex was also needed for efficient repair of DSBs inflicted throughout the genome by IR. IR sensitivity can be read-out by two different assays in *C. elegans*: a so-called 'L1 assay' in which somatic cells of L1 larva are irradiated and consequent developmental defects are analyzed, or a 'L4 assay' in which germ cells of L4 larva are irradiated and progeny

survival is analyzed (Figure 5A). While NHEJ mutants are hypersensitive in the L1 assay, most HR mutants are not, reflecting the key role of NHEJ in DSB repair in somatic cells (Clejan *et al.* 2006). In contrast, HR mutants are typically hypersensitive in the L4 assay, while NHEJ mutants are not, reflecting the strong HR bias in germ cells (Clejan *et al.* 2006; Johnson *et al.* 2013; Robert *et al.* 2008).

In line with previous studies, we found animals deficient for BRC-1 (the ortholog of well-established human HR factor BRCA1) to be hypersensive to IR in the L4 assay, while animals lacking the canonical NHEJ factor CKU-80 behaved like wild-types in this assay (Figure 5B) (Johnson *et al.* 2013; Lemmens *et al.* 2013). Similar to NHEJ mutants, germ cells defective for THOC-5 or THOC-7 were not hypersensitive to IR, indicating that lack of these THO factors does not sensitize genomes to IR *per se* (Figure 5B). However, when we subjected these animals to the L1 assay, we observed the inverse pattern. While *brc-1* mutants showed a very mild increase in IR-induced somatic defects, *cku-80, thoc-5* and *thoc-7* mutants were very sensitive in this assay, resulting into ~75%, 65% and 55% of animals having defective vulva development after 60Gy of IR, respectively (Figure 5C). Similar to canonical NHEJ mutants, *thoc-5* and *thoc-7* mutants showed various IR-dependent vulval defects, including protruding vulvas, ruptured vulvas and so-called “bag-of-worms” where progeny hatches within the mother because of an egg laying defect, all culminating to less progeny on the plate (Figure 5D). These data indicate that THOC-5 and THOC-7 are needed for IR resistance in somatic cells, like the vulval precursor cells, but not in germ cells, suggesting a specific need for the THO complex in DSB repair via NHEJ. This notion is also strongly supported by our previous data acquired via the transgenic DSB repair assays, where these THO mutants were specifically defective for NHEJ but not in other repair pathways such as SSA (Figure 2A). Moreover, both the IR sensitivity assays and the I-SceI-based reporter assays revealed a partial defect in NHEJ in THO mutants compared to CKU-80 null mutants, indicating that these THO alleles do not completely abolish THO function or that the THO complex promotes NHEJ in a non-essential manner.

While the read-out for *in vivo* NHEJ activity using the transgenic reporter system was completely saturated in NHEJ null mutants (showing no pharyngeal GFP at all), the L1 assay (measuring vulval defects) was not saturated, even when NHEJ was completely abolished (Figure 1 and 5). This attribute of the L1 assay allowed us to do epistasis analysis between the different NHEJ mutants and delineate the pathways involved. We created *thoc-5 thoc-7* and *cku-80; thoc-7* double mutants and analyzed IR sensitivity parallel to the single mutants described above. As expected, we found the *thoc-5; thoc-7* double mutants and *thoc-5* single mutants to be equally sensitive to IR, implying that THOC-5 and THOC-7 act in the same pathway/protein complex that confers IR resistance (Figure 5E). Notably, we also found *cku-80; thoc-7* double mutants and *cku-80* single mutants to be equally sensitive to IR, indicating that CKU-80 and THOC-7 also act in the same pathway that confers IR resistance in somatic cells (Figure 5F).

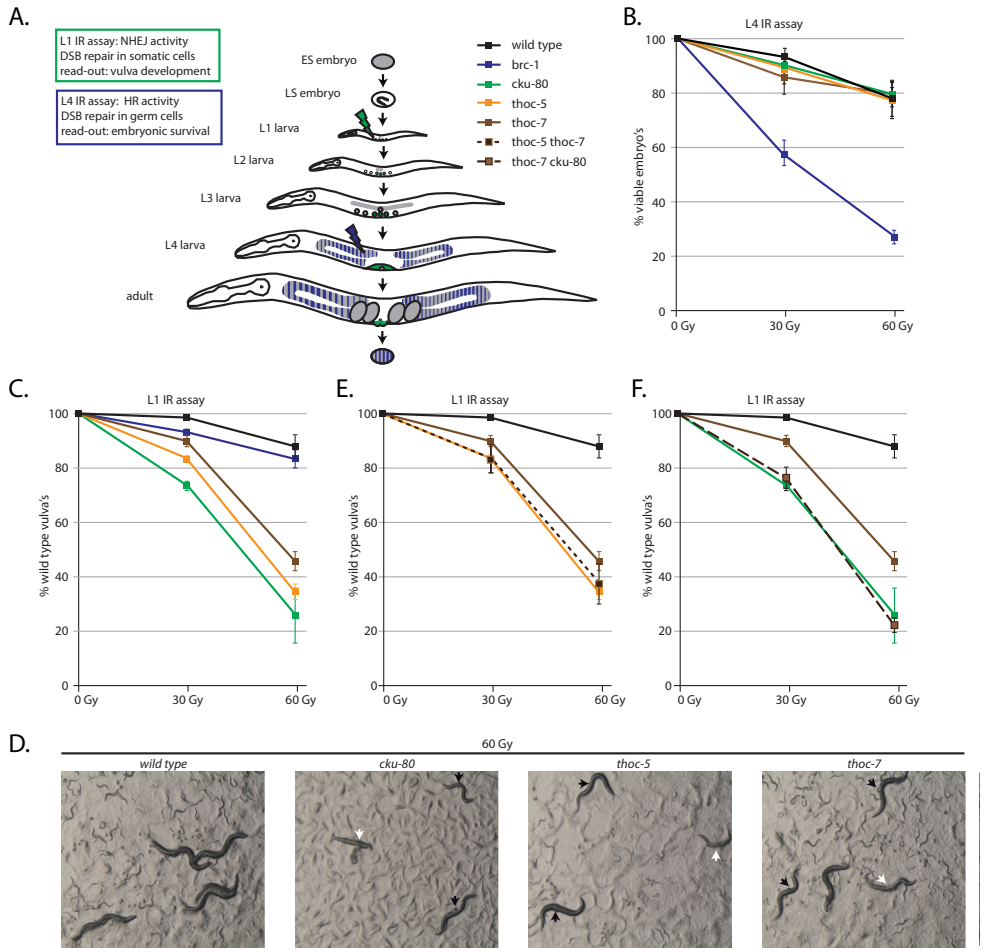

Figure 5. THO deficient somatic tissues are hypersensitive to IR

A. Schematic representation of two different IR assays in the context of the *C. elegans* life cycle and challenged tissues. While the L1 assay measures IR-resistance of arrested vulva precursor cells and mainly reflects NHEJ activity, the L4 assay measures IR-resistance of germ cells and typically reflects HR activity. **B.** L4 assay; L4 animals were challenged by the indicted dose of IR and percentage of viable progeny is plotted, see figure 5A for legends. Values depict the average of three independent experiments and error bars represent S.E.M. **C.** L1 assay; L1 animals were challenged by the indicted dose of IR and percentage of wild-type vulvas is plotted, see figure 5A for legends. Values depict the average of three independent experiments and error bars represent S.E.M. **D.** Representative pictures of synchronized animal populations 5 days post IR (60 Gy at L1 stage). Black arrows indicate protruding vulvas. White arrows indicate severe vulva defects resulting in germline extrusion or internal hatchlings/ bag-of-worms. **E. F.** L1 assay; Percentage of wild-type vulvas of double mutants treated as in Figure 5C; results of reference strains are depicted again, see figure 5A for legends. Values depict the average of three independent experiments and error bars represent S.E.M.

Given that *thoc-7* deficiency cannot increase IR sensitivity of NHEJ null mutants also suggests that THOC-7 loss does not result in additional (IR-induced) DSBs, which may be expected if RNA:DNA hybrids were to be abundant in the genome (Gomez-Gonzalez *et al.* 2011). In accordance with this notion, we found that increased SSA in *thoc-7* mutants depended on I-SceI-induced DSB formation and thus was not due to increased spontaneous DSBs (Figure S3C). All together, these data strongly argue for a role of the THO complex post DSB induction and at the level of DSB repair via classical NHEJ.

Splicing regulator PNN-1 is required for efficient NHEJ in *C. elegans*

The identification of the THO complex as an NHEJ regulator raised the intriguing question of how RBPs promote DNA repair? Interestingly, we identified one other NHEJ mutant (allele *lf159*), which based on epistasis and mapping studies was not affected in known NHEJ genes nor in THO complex genes. Subsequent genome-wide sequencing analysis revealed a candidate nonsense mutation in R186.7 (*pnn-1)*, the *C. elegans* homolog of Pinin/DRS/memA (Figure 6A and 6B). Pinin was first identified as a desmosome-associated protein involved in cell adhesion, but later was found to have nuclear functions as well, including regulation of mRNA splicing (Li *et al.* 2003; Wang *et al.* 2002). Notably, both the THO complex and Pinin associate with spliceosomes and have been implicated in pre-mRNA processing and mRNA export (Chi *et al.* 2013; Li *et al.* 2003). Given the potential functional overlap between PNN-1 and the THO complex, we investigated the putative role for PNN-1 in NHEJ regulation. To address if the NHEJ defect in *lf159* animals was caused by the early stop mutation in *pnn-1*, we crossed *lf159* hermaphrodites with wild-type males and males homozygous for a *pnn-1* deletion allele (*ok1872*). While NHEJ activity was restored in heterozygous cross progeny from wild-type males (*lf159/+)*, NHEJ activity in trans-heterozygous cross progeny (*lf159/ok1872)* was still defective, strongly arguing for a causal role for PNN-1 in NHEJ regulation (Figure 6C). Moreover, animals homozygous for the independently derived *pnn-1* null allele (*ok1872*) showed reduced NHEJ activity (pharyngeal LacZ) and increased SSA activity (somatic LacZ), revealing attenuated DSB repair and repair pathway specificity (Figure 6B and 6D).

To validate a role for PNN-1 in somatic NHEJ, we challenged *pnn-1* mutant larva with IR-induced DSBs during development. Like *thoc-5* mutants, *pnn-1* deletion mutants were hypersensitive to IR (Figure 6E and 6F). Interestingly, animals defective for both *thoc-5* and *pnn-1* were even more sensitive to IR than the respective single mutants, indicating that the THO complex and PNN-1 can act cooperatively to promote NHEJ (Figure 6F). This notion was confirmed using the dual reporter system: *thoc-5; pnn-1* double mutants showed reduced pharyngeal GFP (NHEJ) and increased somatic LacZ (SSA) compared to either single mutant (Figure 6D). Notably, *thoc-5; pnn-1* double mutants were nearly as sensitive to IR as *lig-4* deficient animals and showed severely reduced NHEJ reporter activity, illustrating the significance for these RBPs for efficient NHEJ *in vivo*.

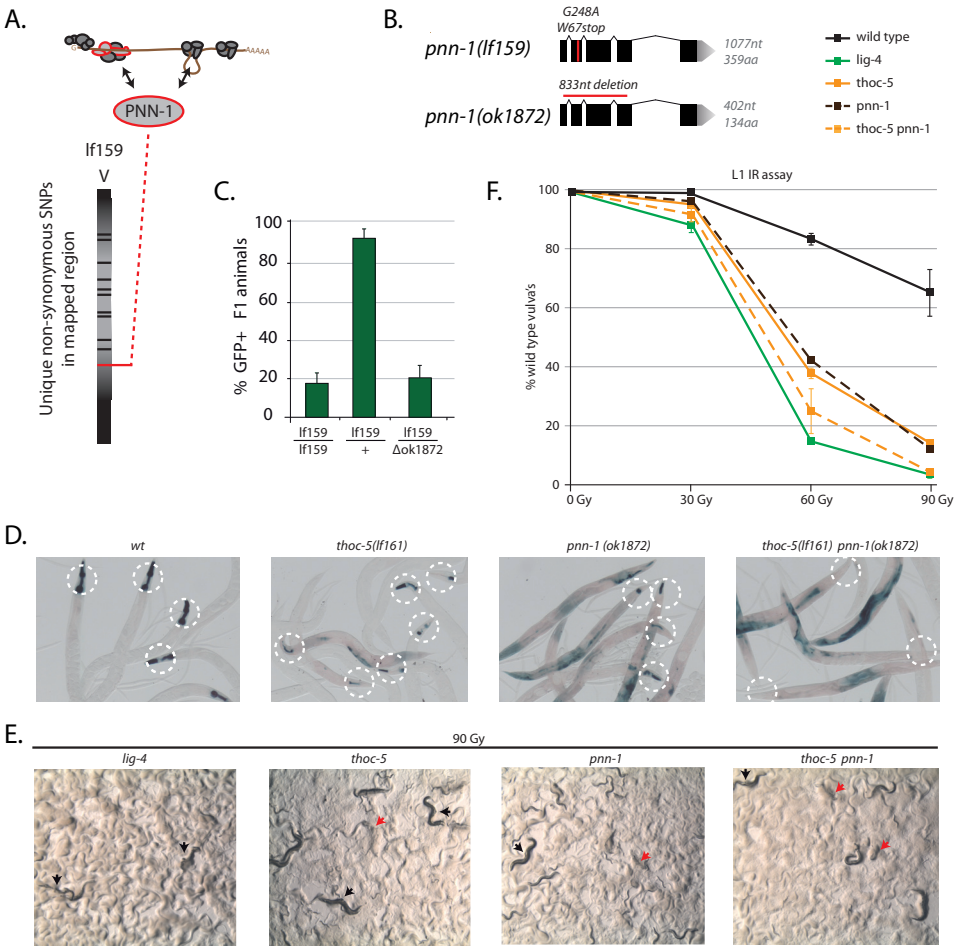

Figure 6. PNN-1 deficiency causes reduced NHEJ activity in *C. elegans*

A. Schematic diagram of chromosomes V in lf159 mutants; regions found to cause NHEJ defect based on Hawaiian SNP mapping in grey; black/red horizontal lines represent unique non-synonymous mutations. lf159 mutants have a nonsense mutation in *pnn-1*, a gene implicated in mRNA metabolism **B.** Gene model and alleles of pnn-1. **C.** Quantification of GFP-positive pharynxes in trans-heterozygous F1 cross progeny, heat-shocked for 120 minutes and measured in adults. **D.** LacZ expression patterns of synchronized dual reporter animals of the indicated genotype; animals were heat-shocked for 180 minutes at L1 and adult stage and stained three hours after the second heat-shock. **E.** Representative pictures of synchronized animal populations 6 days post IR (90 Gy at L1 stage). Black arrows indicate protruding vulvas. Red arrows indicate severe vulva defects resulting in germline extrusion or internal hatchlings/bag-of-worms. Wild-type control populations were starved under these conditions as they still had high fecundity. **F.** L1 assay; L1 animals were challenged by the indicted dose of IR and percentage of wild-type vulvas is plotted, see figure 5B for legends. Values depict the average of three independent experiments and error bars represent S.E.M.

PNN-1 and THO regulate mRNA stability of a specific set of transcripts

RBPs have been implicated in genome instability via the formation of RNA:DNA hybrids as well as their direct role in expression of specific DNA repair factors (Chan *et al.* 2014; Savage *et al.* 2014). Here we found Pinin/THO deficient L1 larva to be defective in NHEJ of both nuclease- and IR-induced DSBs. In order to search for gene expression alterations that may explain the NHEJ defect in these mutants, we extracted total RNA from unchallenged L1 larva from wild-type, *thoc-5*, *thoc-7* and *pnn-1* mutants and performed genome-wide RNAseq. We found total RNA levels not to be significantly different between the genotypes tested, implying that PNN-1 and these THO complex factors are not essential for transcription and stability of bulk RNA (Figure S4A). However, PNN-1 as well as THO deficiency resulted in altered expression of a specific set of genes, leading to both elevated and reduced expression patterns (Figure 7A). While THOC-5 loss resulted in slightly more down-regulated transcripts (535↓ versus 438↑), PNN-1 loss resulted in more up-regulated transcripts (195↓ versus 330↑), implying different roles for these RBPs in mRNA metabolism (Figure 7A). However, we also found a highly significant overlap in affected transcripts between the NHEJ mutants: *thoc-7* and *thoc-5* mutants shared 67% and 71% of up-regulated and down-regulated transcripts, and *pnn-1* and *thoc-5* shared 20% and 50% of up-regulated and down-regulated transcripts, respectively (Figure 7A). Thus PNN-1 and THOC-5/7 control mRNA stability in a similar but non-identical manner and both act highly selectively, affecting less than 5% of the transcriptome.

Since all three mutants were defective in NHEJ, we first focused on the genes of which the expression was significantly affected in all mutants, culminating to a common list of 18 up-regulated and 41 down-regulated transcripts (Figure 7A and S5). Interestingly, a quarter of the up-regulated genes was implicated in RNA metabolism and transcription, and 2 out of 18 up-regulated genes (11%) are known to interact with THO and promote mRNA splicing and export, potentially reflecting an *in vivo* response on a mRNA processing defect in these animals (Figure 7B) (Castellano-Pozo *et al.* 2012a; Chi *et al.* 2013). Surprisingly, none of the shared up-regulated and down-regulated genes were linked to DNA repair, cell cycle checkpoint or DNA damage signaling, suggesting that altered mRNA stability (by itself) was not the main cause for the NHEJ defect in these mutants (Figure S5).

Reduced LIG-4 levels in THO mutants do not limit NHEJ

Among the significantly down-regulated genes in *thoc-5* mutants, we found *lig-4* mRNA to be reduced to 36% of wild-type levels. Also *thoc-7* (68%) and *pnn-1* (86%) mutants seemed to suffer for reduced *lig-4* expression, yet this reduction was not statistically significant (Figure S4B). In contrast, *cku-80* and *cku-70* mRNA levels were not significantly affected in any of the RBP mutants, hinting towards a possible specific defect in *lig-4* mRNA metabolism (Figure S4C). To investigate if the NHEJ defect in THO mutants was due to low *lig-4* expression, we constructed several functional *lig-4* cDNA overexpression arrays and crossed these into the

different genetic backgrounds. While the *lig-4* cDNA arrays effectively restored NHEJ activity in *lig-4* mutants, they did not in *thoc-5* mutants, arguing that LIG-4 is not limiting in THO deficient animals (Figure S4D). Parallel approaches using fosmids or vectors expressing the full length *lig-4* ORF lead to similar results, further supporting the notion that reduced *lig-4* expression was not the cause of the NHEJ defect in *thoc-5* mutants. Besides, reducing functional *lig-4* transcripts by 50%, like in lig-4(ok719) heterozygotes (Figure 2B), also did not impede NHEJ efficacy, arguing that the residual *lig-4* mRNA levels in THO mutants (36-68%) are likely sufficient for efficient NHEJ.

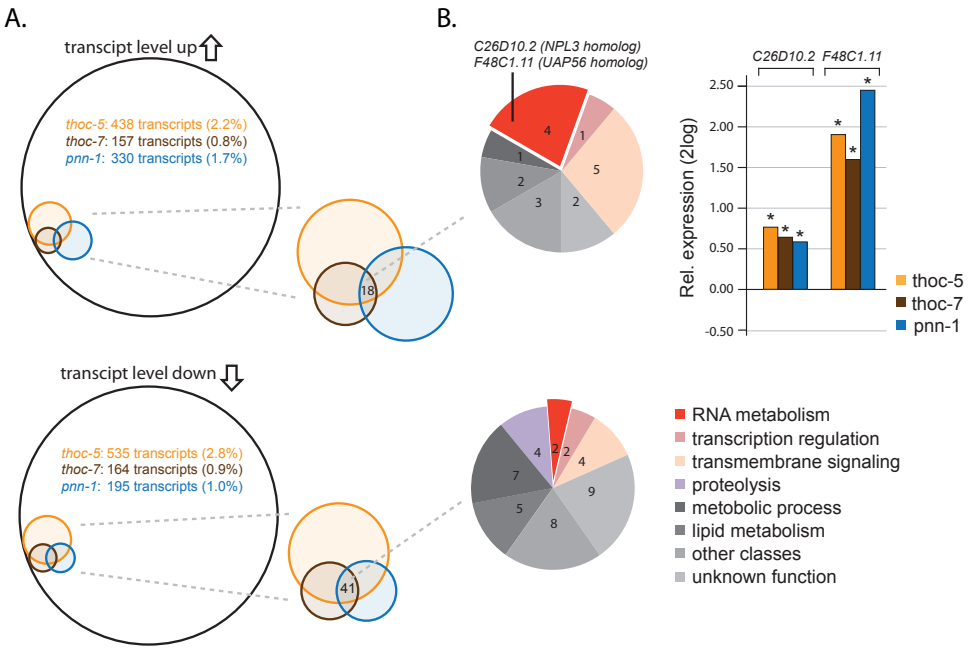

Figure 7. Transcriptome analysis by RNAseq of *thoc-5*, *thoc-7* and *pnn-1* mutants
A. Venn diagrams of significantly affected transcripts found by RNAseq in synchronized L1 animals of the indicated genotype. Upper Venn diagram depicts the number up-regulated transcripts; Lower Venn diagram depicts the number of down-regulated transcripts. Only a small fraction of the ~19.000 different transcripts identified was affected by *thoc-5*, *thoc-7* or *pnn-1* deficiency (% indicated between parentheses). A significant overlap was found between affected transcripts in *thoc-5*, *thoc-7* or *pnn-1* mutants, culminating in 18 and 41 common transcripts being up or down regulated, respectively. **B.** Pie-charts of the commonly affected transcripts in *thoc-5*, *thoc-7* or *pnn-1* mutants categorized by gene function, see bottom right for legends. Among the few up regulated transcripts, two encoded for genes that had orthologs implicated in RNA splicing and export: C26D10.2 and F48C1.11. Histogram depicts relative expression changes of C26D10.2 and F48C1.11 in the indicated genotypes compared to wild-type; asterisks indicates significance $q \leq 0.05$.

PNN-1 and THO deficiency results in alternative splicing at specific genes

Since Pinin and THO were also linked to mRNA splicing, we examined the genome-wide RNAseq data for splicing alterations that may explain the NHEJ defect. To identify putative splicing defects in these animals, we analyzed the relative expression of exons within the ~19.000 transcripts identified (Figure 8A). As expected, we identified altered expression patterns in the *pnn-1* gene itself, only in the strains carrying the *pnn-1* deletion allele, validated our bioinformatics analysis and substantiated the notion that we detected bona-fide exon-specific expression changes (Figure 8B). While the total number of transcripts with altered exon expression was relatively low in all three NHEJ mutants (*i.e.* 87, 17 and 19 transcripts in *thoc-5*, *thoc-7* and *pnn-1* mutants, respectively), the overlap in affected transcripts between the NHEJ mutants was strikingly large, being 100% between *thoc-7* and *thoc-5* and 42% between *pnn-1* and *thoc-5* (Figure 8A). Thus THOC5/7 and PNN-1 are highly selective regulators of mRNA metabolism, affecting exon expression of less than 0.5% of the transcriptome. This high selectivity was even seen at the level of individual exons: often the very same exons were included or excluded in THO and PNN-1 mutants, suggestive of highly specific splicing defects (Figure 8C). For example, all the NHEJ modifiers showed increased inclusion of the second exon of *oxi-1,* which encodes a well-conserved E3 ubiquitin ligase reported to be alternatively spliced in humans (Gong *et al.* 2003).

Although our RNAseq analysis revealed a common role for THOC5/7 and PNN-1 in RNA metabolism, they also revealed subtle differences. For instance, *thoc-5*, *thoc-7* and *pnn-1* mutants all showed alternative splicing of *tos-1*, resulting in the inclusion of the first intron; yet only the THO mutants showed additional expression defects of the latter three exons of *tos-1*, suggestive of a THO-specific alternative transcript (Figure 8D). The fact that THO and PNN-1 control mRNA metabolism in similar but non-identical ways may explain their cooperative functions in NHEJ regulation, as correct splicing of NHEJ genes may require both PNN-1 and the THO complex (Figure 6 and 8).

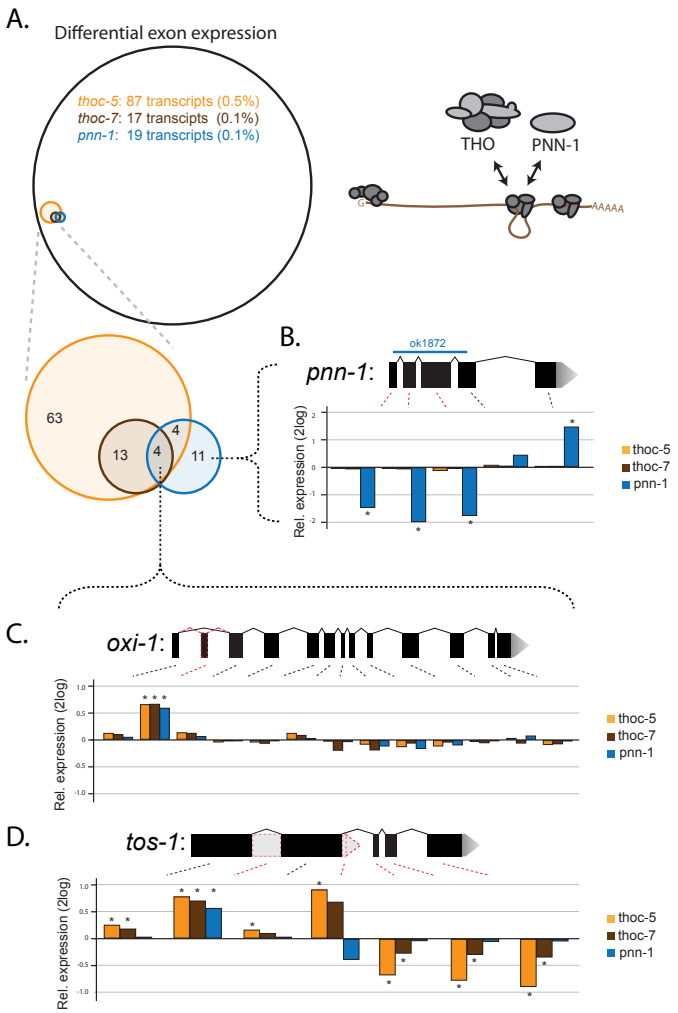

Figure 8. Differential exon expression in *thoc-5*, *thoc-7* and *pnn-1* mutants
A. Venn diagram of transcripts bearing significantly affected exons found by RNAseq in synchronized L1 animals of the indicated genotype. Only a small fraction of the ~19.000 transcripts identified had differentially expressed exons (% indicated between parentheses). Venn diagram blowup depicts unique and shared number of transcripts; a significant overlap was found between affected transcripts in *thoc-5*, *thoc-7* or *pnn-1* mutants, suggestive of rare but highly specific splicing defects. **B.** Gene model of *pnn-1* illustrates the position of the genomic deletion present in the *pnn-1* mutant (blue line). Histogram depicts relative expression changes of exons of the *pnn-1* transcript compared to wild-type as identified by DexSeq in the indicated genotypes; asterisks points out significance q ≤ 0.05. **C.** Gene model of *oxi-1*, a gene bearing alternatively expressed exons in *thoc-5*, *thoc-7* and *pnn-1* mutants. Histogram depicts relative expression changes of exons of *oxi-1* transcripts compared to wild-type as identified by DexSeq in the indicated genotypes; asterisks points out significance q ≤ 0.05. All three NHEJ mutants have increased inclusion of the second exon. **D** Gene model of *tos-1*, a gene bearing alternatively expressed exons in *thoc-5*, *thoc-7* and *pnn-1* mutants. Histogram depicts relative expression changes of exons of *tos-1* transcripts compared to wild-type as identified by DexSeq in the indicated genotypes; asterisks points out significance q ≤ 0.05. All three NHEJ mutants have increased inclusion of the first intron.

Reduced expression of splicing factor UAF-1 in THO mutants impedes NHEJ

Previous studies on alternative splicing have led to the identification of novel splicing regulators in *C. elegans*, including MFAP-1 and the ortholog of human U2AF large subunit UAF-1 (Ma *et al.* 2012; Ma and Horvitz 2009). Interestingly, *uaf-1* mutants have reported splicing defects reminisced to the ones occurring in the Pinin/THO mutants described here. In fact, *uaf-1* deficient animals also included the first intron of the endogenous splicing reporter gene *tos-1* (Ma *et al.* 2012). Our RNAseq analysis revealed that both *thoc-5* and *thoc-7* deficient L1 animals, but not *pnn-1* mutants, suffered from reduced *uaf-1* expression, implying that the splicing defects observed on the THO mutants may be caused (partially) by UAF-1 deficiency (Figure 9A and 9B). In order to establish a functional link between UAF-1 expression and *in vivo* DSB repair, we depleted *uaf-1* expression in dual DSB repair reporter animals by RNAi. Similar to THO mutation, *uaf-1* depletion by RNAi resulted in lower NHEJ activity and elevated SSA activity, strongly arguing for a causal link between altered UAF-1 levels and NHEJ efficacy (Figure 9C and 9D). All together these data suggest that both the splicing alterations and NHEJ defects in THO mutant animals can be readily explained by reduced expression of the essential splice factor UAF-1. The identification of selective splicing factors such as PNN-1 and UAF-1 in controlling *in vivo* NHEJ efficacy reveals another layer of regulation of error-prone DSB repair activities during animal development and sets the stage to discover alternative gene products that contribute to genome stability.

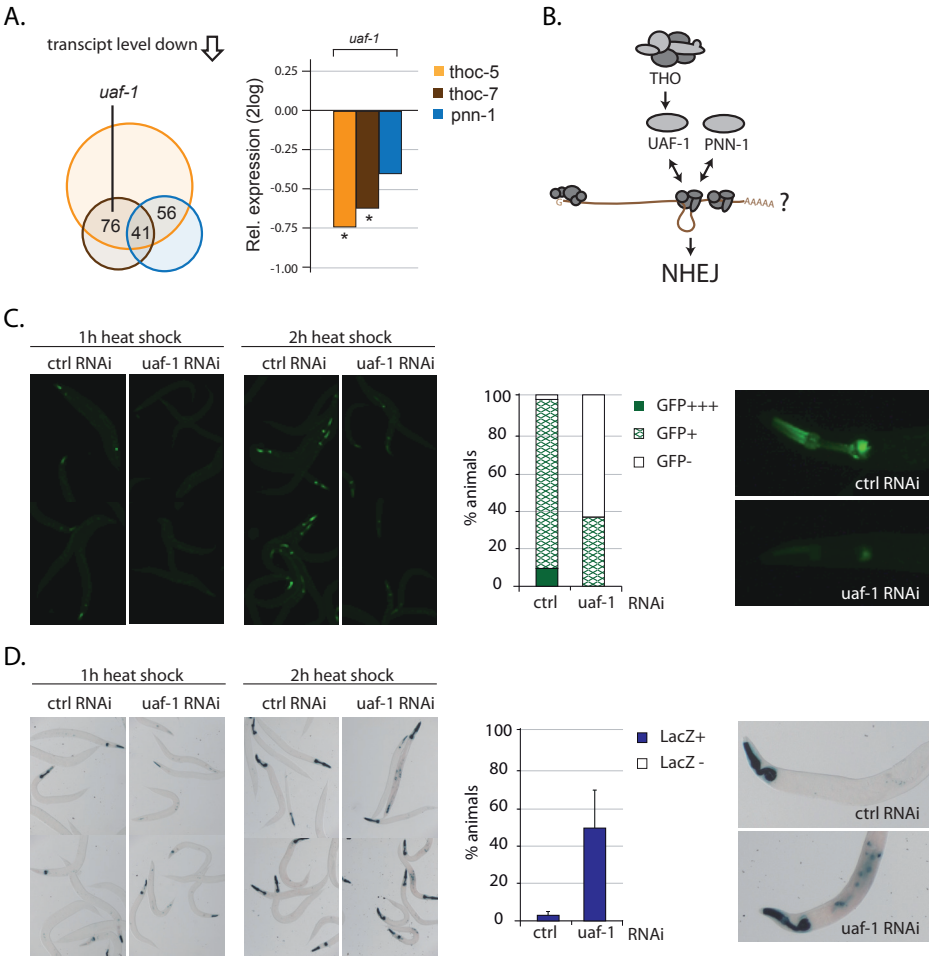

Figure 9. Reduced expression of splicing factor UAF-1 in THO mutants impedes NHEJ
A. Venn diagram of significantly down regulated transcripts compared to wild-type control as found by RNAseq in synchronized L1 animals deficient for *thoc-5* (orange), *thoc-7* (brown) and *pnn-1* (blue); numbers of commonly affected transcripts are indicated. Histogram depicts relative expression changes of *uaf-1* transcripts in the indicated genotypes compared to wild-type; asterisks indicates significance q ≤ 0.05. **B.** Putative model for the cooperative action of PNN-1 and the THO complex in NHEJ regulation. **C.** Representative pictures and quantifications of GFP-positive pharynxes in synchronized populations of dual reporter animals fed bacteria carrying empty vectors of *uaf-1* RNAi vectors. L4 animals were heat-shocked for 60 or 120 minutes and GFP was measured in adults. Histogram depicts average percentage of GFP-positive pharynxes of three independent populations (n>150) heat-shocked for 60 minutes. Right panels show representative pharyngeal GFP patterns in control and UAF-1 RNAi animals that were qualified as GFP+++ and GFP+, respectively. **D.** Representative pictures and quantifications of non-pharyngeal LacZ staining patterns in synchronized populations of dual reporter animals fed bacteria carrying empty vectors of *uaf-1* RNAi vectors. L4 animals were heat-shocked for 60 or 120min and stained 24 hours later. Histogram depicts average percentage of LacZ positive animals of three independent populations (n>150) heat-shocked for 120 minutes. Error bars represent S.E. Right panels show representative LacZ patterns in control and UAF-1 RNAi animals that were qualified as LacZ- and LacZ+, respectively.

Discussion

Recent large-scale genetic and molecular studies identified RBPs as important players in the maintenance of genome stability, either because they directly affect DNA repair or because they control proper expression of crucial DNA repair factors (Dutertre *et al.* 2014; Lenzken *et al.* 2013). At the same time several studies have shown that uncontrolled DSB repair activities can be detrimental for animal development and human health, ultimately leading to severe diseases such as cancer (Bunting *et al.* 2010; McKinnon and Caldecott 2007). How the activity of error-prone repair routes like NHEJ are controlled during development and if RBPs play a role in this regulation is unknown to date. We performed an unbiased forward genetics screen to find factors required for efficient DSB repair in *C. elegans* and identified the well-conserved RBPs PNN-1 and THOC-2/5/7 to promote NHEJ (Figure 1, 2, 3 and 6). Multiple transgenic DSB repair reporter assays and independent IR-sensitivity assays confirmed a specific requirement of these RBPs for efficient NHEJ but not DSB repair via HR/SSA (Figure 2, 5 and 6). Subsequent genome-wide transcriptome analysis revealed that the novel NHEJ mutants suffered from highly specific splicing defects, implying a functional link between NHEJ regulation and mRNA splicing. This notion was substantiated by the fact that the essential splicing factor UAF-1/U2AF was mis-expressed in THO deficient animals and the NHEJ defects seen in THO mutants could be mimicked by reduced expression of UAF-1. We propose that altered processing of specific mRNAs could provide means to affect the activity of essential gene products and may compromise DSB repair efficacy in somatic tissues.

Notably, many cancer-associated genes are regulated through alternative splicing, arguing for a significant role of this post-transcriptional regulatory mechanism in the production of oncogenes and tumor suppressors. Recently, more intimate links between splicing and DNA repair complexes became apparent, as for instance the tumor suppressor BRCA1 was found to be part of a DNA damage-induced splicing complex that controlled the activity of multiple DNA repair genes (Savage *et al.* 2014). Moreover, reduced expression of U2AF was shown to result in IR hypersensitivity in human cells, which could be partially restored by overexpression of specific DNA repair genes, suggesting a conserved but indirect role for U2AF in regulating DSB repair (Savage *et al.* 2014). Interestingly, Pinin was recently found to be part of a novel RNA-processing complex that recruits U2AF to specific mRNAs (Bracken *et al.* 2008) and both Pinin and U2AF are phosphorylated upon DNA damage (Beli *et al.* 2012). In support of a post-transcriptional role for such splicing factors in DNA repair, Pinin is redistributed upon DNA damage but is excluded from DNA damage sites (Beli *et al.* 2012; Lenzken *et al.* 2013). Alternative splicing factors are known to respond to external and internal stimuli and recent high-throughput screenings in human cells and budding yeast revealed many RNA processing factors to be targets of DNA damage checkpoint kinases. In fact, human THOC5 is also directly phosphorylated by ataxia-telangiectasia-mutated (ATM) kinase upon DNA damage

and this modification impaired THOC5/mRNA complex formation, drastically decreasing the cytoplasmic pool of THOC5-dependent mRNAs (Ramachandran *et al.* 2011). We show here that deficiencies of several well-conserved RNA processing factors, including PNN-1 and THOC-5 can result in highly selective splicing defects that are paralleled by specific defects in NHEJ, including somatic IR hypersensitivity. Given that RNA processing factors are often deregulated in response to genotoxic treatments and many chemotherapeutic therapies are based on DNA-damaging agents, alternative splicing can be an important determinant of how tumor cells respond to therapy (Ladomery 2013; Lenzken *et al.* 2013). The identification of the abovementioned RBPs in NHEJ sets the stage for further dissection of mRNA processing mechanisms during development and potentially will reveal important roles of alternative RNA transcripts in DSB repair.

Materials and Methods

Genetics

All strains were cultured according to standard *C. elegans* procedures (Brenner 1974). Alleles used in this study include: *LGIII: cku-70(tm1524), cku-80(ok861), lig-4(ok716), brc-1(tm1145), cku-70(lf151), cku-80(lf152), cku-80(lf153), thoc-2(lf158), thoc-5/Y32H12A.2(lf161), LGIV: thoc-7/B0513.2(lf160), LGV: pnn-1/R186.7(lf159), pnn-1/R186.7(ok1872), LGX: pkIs2379 [Phsp-16.41::I-SceI-ORF; rol-6(su1006)], pkIs2170 [SSA reporter Phsp-16.41::ATG::LacZ::I-SceI-site::stops::LacZ-ORF); unc-119(+)], lfIs104 [NHEJ reporter Pmyo-2::ATG::I-SceI-site:GFP-ORF::LacZ-ORF, Phsp-16.41::mCherry::I-SceI-ORF; rol-6(su1006)], lfEx164 [thoc-2 fosmid WRM0614bD12; Pmyo-3::mCherry; Prab-3::mCherry], lfEx166 [thoc-7 fosmid WRM0640bD11; Pmyo-3::mCherry; Prab-3::mCherry], lfEx164 [thoc-5 fosmid WRM0617bE04; Pmyo-3::mCherry; Prab-3::mCherry], lfEx190 [lig-4 fosmid WRM0634bF07; Pmyo-3::mCherry; Prab-3::mCherry], lfEx195 [Prpl-28::lig-4-cDNA; Pmyo-3::mCherry; Prab-3::mCherry], lfEx196 [Prpl-28::lig-4-ORF; Pmyo-3::mCherry; Prab-3::mCherry].* All transgenic strains were obtained by microinjection of plasmid/fosmid DNA into the germ line and data presented are from a single representative transgenic line unless noted otherwise. The parental NHEJ reporter transgene *lfIs104* was obtained via IR-mediated genomic integration and combined with *pkIs2379* and *pkIs2170* to create the dual reporter strain XF540, which served as the starting strain for the forward genetics screen.

DSB repair reporter assays

Synchronized L1 animals were obtained by harvesting eggs from hypochlorite-treated gravid adults and overnight starvation in M9 solution (Lewis and Fleming 1995). Hatched L1 larvae were transferred on NGM plates seeded with either *E. coli* OP50 or HT115 bacteria (Kamath and Ahringer 2003). In order to insure complete RNAi before DSB induction, L1 worms were

cultured at 20°C for at least 20hrs. Heat-shock driven I-SceI expression was induced by putting the worms at 34°C for 60-180 minutes, as indicated. After the heat-shock procedure, worms were cultured at 20°C to allow DSB formation, DSB repair and worm development. NHEJ activity was measured by scoring pharyngeal GFP expression using a Leica M165FC fluorescence dissecting-microscope. Experiments were performed in triplicate with 50-200 animals tested for each condition. After GFP quantification, ~25 adult animals were transferring onto microscope slides with 3% agarose pads and representative pictures were acquired using a Leica DM6000 microscope with 10X objective. SSA activity was measured by scoring animals showing LacZ positive cells in non-pharyngeal somatic tissues (Pontier and Tijsterman 2009). One hour prior fixation/LacZ staining, young adults were heat-shocked at 34°C for 120 minutes to induce SSA reporter expression.

Forward genetics screen

Dual reporter larvae (XF540) were mutagenized with ethyl methanesulfonate (EMS) using standard procedures (Brenner 1974). Complex F2 populations, each derived from 50 mutagenized P0s, were bleached and synchronized L1 larvae (F3) were seeded on NGM/OP50 plates. On two consecutive days larvae were heat-shocked at 34°C for 180 minutes in order to maximize GFP ORF correction. GFP^{low} F3 animals were selected using a Leica M165FC fluorescence dissecting-microscope and clonal F4 populations were tested again for NHEJ activity. Populations showing reduced GFP expression were fixed and stained with X-gal as described previously (Pontier and Tijsterman 2009).

Positional cloning, genome-wide sequencing and transgenesis

Causal mutations in *thoc-2*, *thoc-5* and *thoc-7* were mapped by crossing the respective mutants (Bristol) to the related Hawaiian strain CB4856 and performing single-nucleotide-polymorphism mapping on NHEJ proficient versus NHEJ deficient F2 lines (Davis *et al.* 2005). Unique EMS-induced genetic alterations in the mapped regions were identified by comparing genome-wide paired-end sequencing data of the parental mutant strains using the Illumina Hiseq 2000 platform, the *C. elegans* reference genome (Wormbase version 225) and MaqGene software (Bigelow *et al.* 2009). Causality was established by complementation analysis using wild-type fosmid arrays. Complemented regions spanned by the fosmids contained only one non-synonymous SNP (Figure S2). To create transgenic animals carrying fosmid arrays an injection mix containing 100 ng/µl pBluescript, 10 ng/µl pGH8 (Prab-3::mCherry::unc-54-3′UTR), 5 ng/µl pCFJ104 (Pmyo-3::mCherry::unc-54-3′UTR) and 10-50 ng/µl fosmid DNA (lig-4 WRM0634bF07, thoc-2 WRM0614bD12, thoc-7 WRM0640bD11, thoc-5 WRM0617bE04, pnn-1 WRM0637aA06) was injected into the gonads of young adults. For lig-4 cDNA and lig-4 ORF containing vectors, 5 ng/µl plamsid DNA was added instead of the fosmid DNA.

PCR-based DSB repair assay

Synchronized L1 larvae were seeded on NGM/OP50 plates and heat-shocked at 34°C for 180 minutes at L1 and L2 stage. Genomic DNA of L3 larvae was extracted using DNeasy Blood & Tissue kit (Qiagen) and 100ng of genomic DNA was digested for 2 hours at 37°C with 7.5 units I-SceI (New England Biolabs) to enrich for cleavage-resistant sites. PCR on pre-digested DNA was performed in duplicate (2x 50ng) using primers flanking the I-SceI target site of the NHEJ reporter: CTCGCGCATCCCACCGAGCGG and CAGGTAGTTTTCCAGTAGTGC. PCR products were purified (QIAquick) and digested for 2 hours at 37°C with 7.5 units I-SceI (New England Biolabs) and analyzed on a 2% agarose gel. Band intensities were quantified using ImageJ software.

IR sensitivity assays

All IR experiments were performed with a dose rate of 10Gy/minute using an electronic X-ray generator (XYLON International). Figures provide mean values of three independent experiments. L4 assay: three L4 animals per plate were treated with various doses of IR (three plates per condition) and progeny survival was scored 2 days post IR. L1 assay: ~200 L1 larvae per plate were treated with various doses of IR (three plates per condition) and vulval phenotypes were scored 5 days post IR. Representative pictures of irradiated populations were acquired using a Leica DFC295 camera/M165FC microscope.

Transcriptome sequencing

To obtain clean L1 populations and remove dead corpses, o/n starved L1 progeny from hypochlorite-treated gravid adults were filtered using 10µm nylon filters (Millipore). Total RNA of >3000 L1s per sample was extracted as follows: L1 animas were collected in 100µl M9, 400µl Trizol was added, vortexed 2 minutes, followed by 4 snap freeze/thaw cycles at -196°C/37°C, 200µl Trizol was added, incubated 5 minute RT, 120ul chloroform was added, incubated 2 minute at RT, centrifuged 15 minutes at 4°C (16000 rcf), 350µl supernatant was mixed with 70% ethanol (1:1) and total RNA was purified using Purelink® RNA columns (Ambion) and stored at -80°C. Total RNA was DNAse treated (Turbo DNA-free, Ambion) and RNA quality was verified using RNA 6000 Pico kit (Agilent).

RNA-Seq was performed on an Illumina HiSeq 2000 platform using standard reagents. Raw reads were aligned by TopHat on Wormbase assembly 238, which allows for reads to be split over splice junctions. Next, we used Cuffdiff, part of the Cufflinks package, to identify differentially expressed transcripts ($q \leq 0.05$) (Trapnell *et al.* 2012). For differential exon expression, DEXSeq was used using standard settings ($p_{adjust} \leq 0.05$) (Anders *et al.* 2012). For each mutant two or three samples were sequenced to account for variation between isogenic L1 populations.

Supporting Information

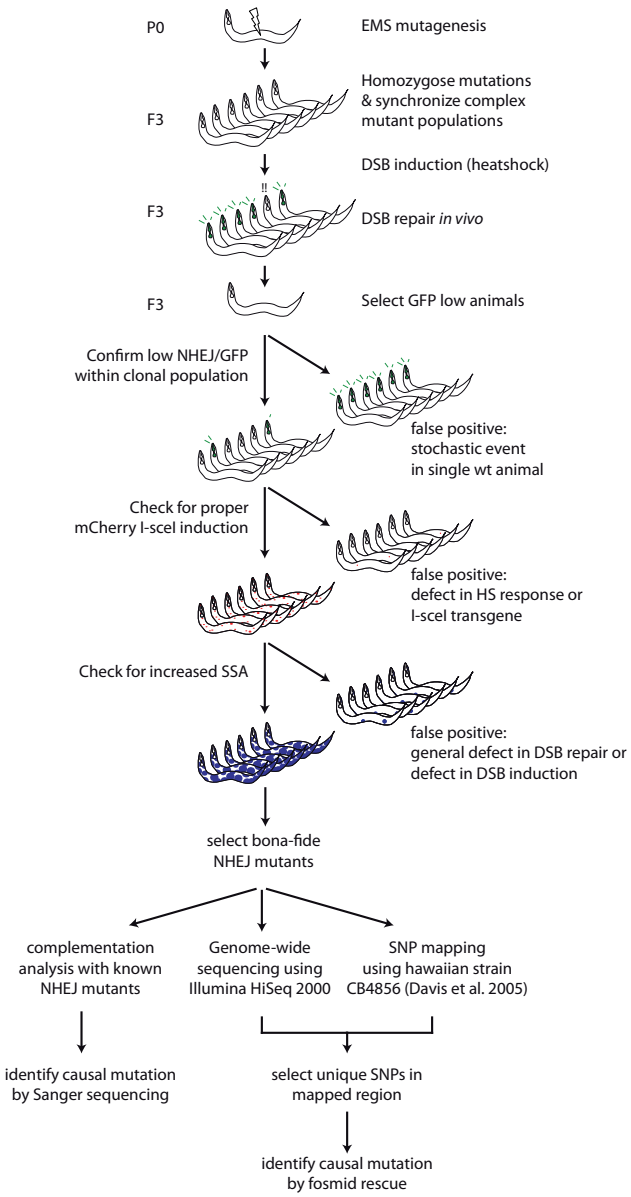

Figure S1. Schematic setup of forward genetics screens to identify NHEJ mutants

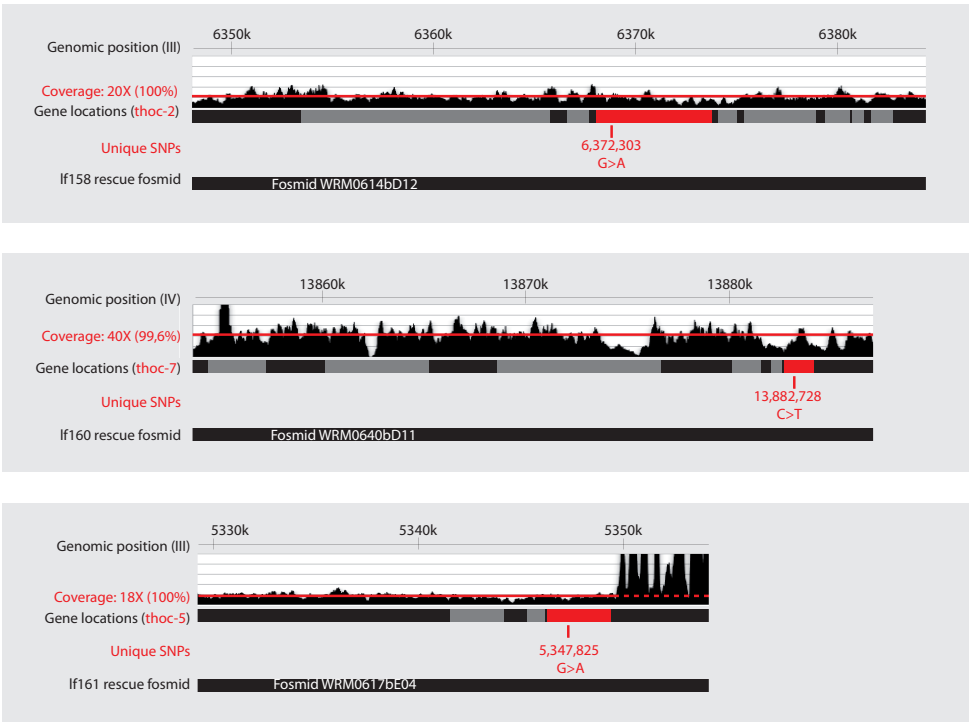

Figure S2. Sequencing coverage and genomic context of regions spanned by rescue fosmids Unique non-synonymous SNPs in the genomic regions spanned by the rescue fosmids (highlighted in red). Despite high sequence coverage, no other unique variants than those in the THO genes were found in the respective NHEJ mutants, strongly suggesting that we specifically complemented THO deficiency despite the use of 2-4kb fosmids.

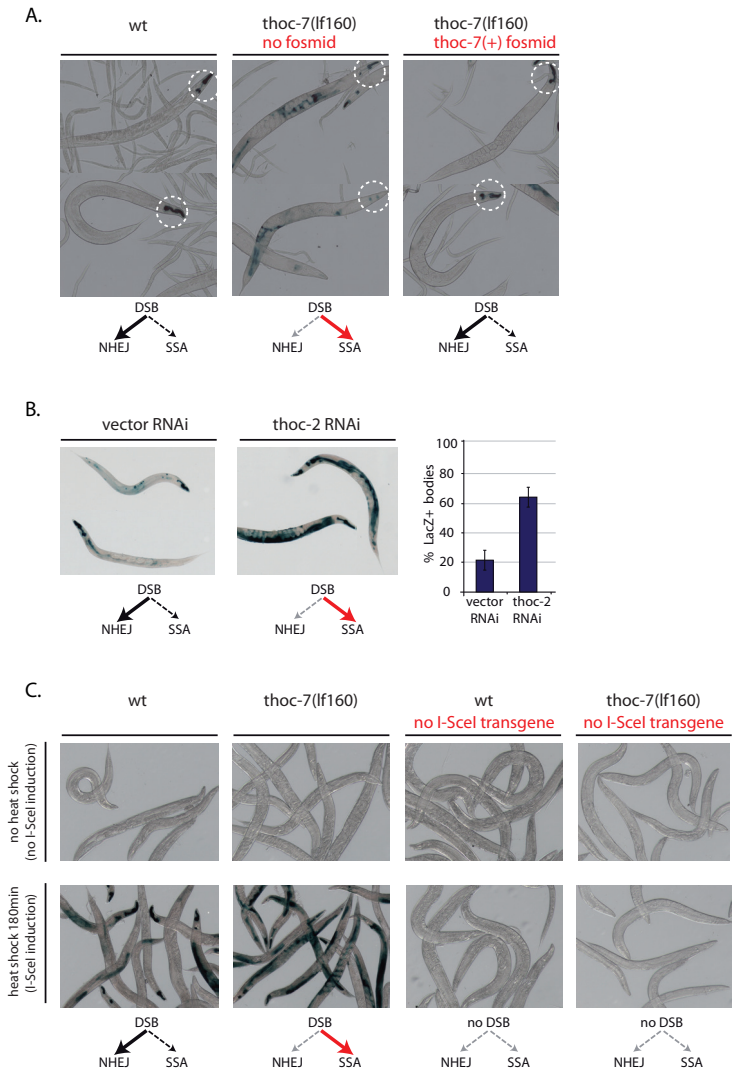

Figure S3 THO deficiency results in increased DSB repair via SSA

A. Representative pictures of synchronized and LacZ-stained dual-reporter animals of the indicated genotype. For the *thoc-7* mutant lines carrying the mCherry-marked *thoc-7* fosmid arrays, mCherry negative and mCherry positive clonal populations were grown and synchronized by bleaching. Animals were heat-shocked for 90 min at L1 stage and stained for LacZ expression as adults. While mCherry negative populations typically showed low pharyngeal LacZ (NHEJ) and increased levels of somatic LacZ (SSA), mCherry positive populations often showed animals having high levels of pharyngeal LacZ (NHEJ). These animals typically also had low levels of somatic LacZ (SSA), indicative of systemic restoration of functional NHEJ **B.** Representative pictures and quantification of LacZ-stained dual-reporter animals treated by feeding RNAi. Synchronized L1 animals were fed on *thoc-2*/empty vector RNAi plates, heat-shocked for 120 minutes at L4 stage and stained as adults. **C.** Representative pictures of transgenic populations of the indicated genotype. Synchronized animals were heat-shocked 0 or 120 minutes at L1 stage and adults were stained for LacZ expression. Increased somatic LacZ (SSA) in *thoc-7* mutants required heat-shock-driven I-SceI expression and the presence of the I-sceI transgene.

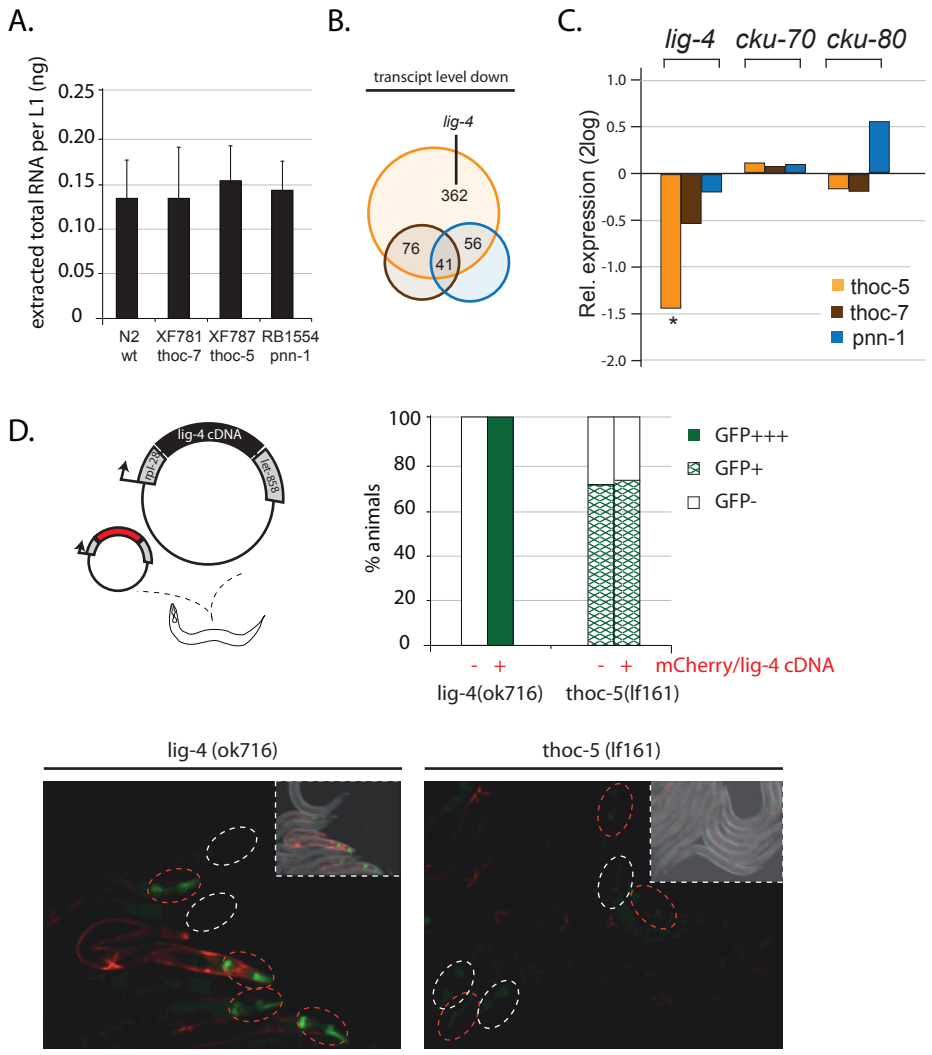

Figure S4 Reduced LIG-4 levels in THO mutants do not limit NHEJ

A. Average total RNA yield per L1 animal determined from four independent L1 populations of the indicated genotype. **B.** Venn diagram of significantly down-regulated transcripts compared to wild-type control as found by RNAseq in synchronized L1 animals deficient for *thoc-5* (orange), *thoc-7* (brown) and *pnn-1* (blue); numbers of commonly affected transcripts are indicated. **C.** Histogram depicts relative expression changes of *lig-4, cku-70* and *cku-80* transcripts in the indicated genotypes compared to wild-type; asterisks indicates significance q ≤ 0.05. **D.** Extra-chromosomal multi-copy arrays carrying P*rpl-28*::*lig-4* cDNA expression constructs and mCherry-based markers were made by micro-injection and crossed to *lig-4(ok716)* and *thoc-5(lf161)* mutants. Histogram shows quantification of GFP-positive pharynxes in adults of the different genetic backgrounds. Average percentage of GFP-positive pharynxes of three independent populations (n>100) is depicted. The presence of the *lig-4* overexpression array (mCherry^{+}) fully restored NHEJ activity in *lig-4* null mutants but did not improve NHEJ in *thoc-5* mutants. Lower panels are representative pictures of synchronized populations of *lig-4* and *thoc-5* deficient animals with (red circle) or without (white circle) *lig-4* overexpression constructs; animals were heat-shocked for 180min at L1 stage.

DOWN	
Transcript	*Function*
Y39B6A.1	*Ca ion binding*
F23H12.4	*cuticle formation*
F53E10.4	*infection response*
T08A9.12	*infection response*
T01C3.4	*lipid metabolism*
W02A2.2	*lipid metabolism*
W03G1.7	*lipid metabolism*
Y51H4A.5	*lipid metabolism*
Y65B4BR.1	*lipid metabolism*
B0310.5	*metabolic process*
F55F3.2	*metabolic process*
F58A6.1	*metabolic process*
K05B2.4	*metabolic process*
T01G1.4	*metabolic process*
T02B5.1	*metabolic process*
ZK550.6	*metabolic process*
M04B2.4	*oxidation respons*
R10E8.6	*protein binding*
T12B5.6	*protein binding*
F36D3.9	*proteolysis*
F44B9.1	*proteolysis*
F59D6.3	*proteolysis*
Y75B8A.4	*proteolysis*
M04B2.1	*RNA binding*
M04B2.2	*RNA binding*
M04B2.3	*transcription regulation*
R07B7.3	*transcription regulation*
C03G6.3	*transmembrane signalling/transport*
E03H4.10	*transmembrane signalling/transport*
R07E3.4	*transmembrane signalling/transport*
ZK666.7	*transmembrane signalling/transport*
C48D1.1	*ubiquitin binding*
C17F4.7	*unknown*
C26E1.2	*unknown*
F46C5.10	*unknown*
K07A9.4	*unknown*
K09F6.6	*unknown*
M04B2.6	*unknown*
Y38C1AA.12	*unknown*
Y51F10.7	*unknown*
ZK1290.13	*unknown*

UP	
Transcript	*Function*
Y48G1C.8	development
ZC168.5	*intercellular signalling*
B0285.9	*lipid metabolism*
Y38E10A.7	*lipid metabolism*
Y50D7A.3	*metabolic process*
WBGene00022951	*ncRNA*
C26D10.2	*RNA binding*
D2030.6	*RNA binding*
F48C1.11	*RNA binding*
B0213.2	*RNA binding?*
C13C4.3	*transcription regulation*
C10F3.3	*transmembrane signalling/transport*
F44F4.4	*transmembrane signalling/transport*
F54F12.1	*transmembrane signalling/transport*
T11G6.3	*transmembrane signalling/transport*
Y80D3A.7	*transmembrane signalling/transport*
F48C1.9	*unknown*
T02G5.14	*unknown*

Figure S5. Common transcripts affected significantly in *thoc-5*, *thoc-7* and *pnn-1* mutants
List of significantly affected transcipts (q ≤ 0.05) that are shared by *thoc-5*, *thoc-7* and *pnn-1* mutants. Putative protein functions based on orthologs or structural domains are indicated.

Acknowledgements
The authors thank Shohei Mitani (National Bioresource Project, Japan) and the Caenorhabditis Genetics Center for strains; Jennemiek van Arkel and Evelina Papaioannou for technical support.

References

Adamo, A., S. J. Collis, C. A. Adelman, N. Silva, Z. Horejsi *et al.*, 2010 Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. Mol Cell **39:** 25-35.

Anders, S., A. Reyes and W. Huber, 2012 Detecting differential usage of exons from RNA-seq data. Genome Res **22:** 2008-2017.

Aslanian, A., J. R. Yates, 3rd and T. Hunter, 2014 Mass spectrometry-based quantification of the cellular response to methyl methanesulfonate treatment in human cells. DNA Repair (Amst) **15:** 29-38.

Beli, P., N. Lukashchuk, S. A. Wagner, B. T. Weinert, J. V. Olsen *et al.*, 2012 Proteomic investigations reveal a role for RNA processing factor THRAP3 in the DNA damage response. Mol Cell **46:** 212-225.

Bigelow, H., M. Doitsidou, S. Sarin and O. Hobert, 2009 MAQGene: software to facilitate C. elegans mutant genome sequence analysis. Nat Methods **6:** 549.

Bracken, C. P., S. J. Wall, B. Barre, K. I. Panov, P. M. Ajuh *et al.*, 2008 Regulation of cyclin D1 RNA stability by SNIP1. Cancer Res **68:** 7621-7628.

Brenner, S., 1974 The genetics of Caenorhabditis elegans. Genetics **77:** 71-94.

Bunting, S. F., E. Callen, N. Wong, H. T. Chen, F. Polato *et al.*, 2010 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell **141:** 243-254.

Castellano-Pozo, M., T. Garcia-Muse and A. Aguilera, 2012a The Caenorhabditis elegans THO complex is required for the mitotic cell cycle and development. PLoS One **7:** e52447.

Castellano-Pozo, M., T. Garcia-Muse and A. Aguilera, 2012b R-loops cause replication impairment and genome instability during meiosis. EMBO Rep **13:** 923-929.

Chan, Y. A., P. Hieter and P. C. Stirling, 2014 Mechanisms of genome instability induced by RNA-processing defects. Trends Genet **30:** 245-253.

Chapman, J. R., M. R. Taylor and S. J. Boulton, 2012 Playing the end game: DNA double-strand break repair pathway choice. Mol Cell **47:** 497-510.

Chi, B., Q. Wang, G. Wu, M. Tan, L. Wang *et al.*, 2013 Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. Nucleic Acids Res **41:** 1294-1306.

Clejan, I., J. Boerckel and S. Ahmed, 2006 Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics **173:** 1301-1317.

Davis, M. W., M. Hammarlund, T. Harrach, P. Hullett, S. Olsen *et al.*, 2005 Rapid single nucleotide polymorphism mapping in C. elegans. BMC Genomics **6:** 118.

Dominguez-Sanchez, M. S., S. Barroso, B. Gomez-Gonzalez, R. Luna and A. Aguilera, 2011 Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. PLoS Genet **7:** e1002386.

Dutertre, M., S. Lambert, A. Carreira, M. Amor-Gueret and S. Vagner, 2014 DNA damage: RNA-binding proteins protect from near and far. Trends Biochem Sci **39:** 141-149.

Dvorak, C. C., and M. J. Cowan, 2010 Radiosensitive severe combined immunodeficiency disease. Immunol Allergy Clin North Am **30:** 125-142.

Gomez-Gonzalez, B., M. Garcia-Rubio, R. Bermejo, H. Gaillard, K. Shirahige *et al.*, 2011 Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. EMBO J **30:** 3106-3119.

Gong, T. W., L. Huang, S. J. Warner and M. I. Lomax, 2003 Characterization of the human UBE3B gene: structure, expression, evolution, and alternative splicing. Genomics **82:** 143-152.
Hu, H., and R. A. Gatti, 2011 MicroRNAs: new players in the DNA damage response. J Mol Cell Biol **3:** 151-158.

Huertas, P., and A. Aguilera, 2003 Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol Cell **12:** 711-721.

Johnson, N. M., B. B. Lemmens and M. Tijsterman, 2013 A role for the malignant brain tumour (MBT) domain protein LIN-61 in DNA double-strand break repair by homologous recombination. PLoS Genet **9:** e1003339.

Jungmichel, S., and M. Blasius, 2013 Rapid and transient protein acetylation changes in response to DNA damage. Cell Cycle **12:** 1993.

Jungmichel, S., F. Rosenthal, M. Altmeyer, J. Lukas, M. O. Hottiger *et al.*, 2013 Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses. Mol Cell **52:** 272-285.

Kamath, R. S., and J. Ahringer, 2003 Genome-wide RNAi screening in Caenorhabditis elegans. Methods **30:** 313-321.

Ladomery, M., 2013 Aberrant Alternative Splicing Is Another Hallmark of Cancer. Int J Cell Biol **2013:** 463786.

Lemmens, B. B., N. M. Johnson and M. Tijsterman, 2013 COM-1 promotes homologous recombination during Caenorhabditis elegans meiosis by antagonizing Ku-mediated non-homologous end joining. PLoS Genet **9:** e1003276.

Lemmens, B. B., and M. Tijsterman, 2011 DNA double-strand break repair in Caenorhabditis elegans. Chromosoma **120:** 1-21.

Lenzken, S. C., A. Loffreda and S. M. Barabino, 2013 RNA Splicing: A New Player in the DNA Damage Response. Int J Cell Biol **2013:** 153634.

Lewis, J. A., and J. T. Fleming, 1995 Basic culture methods. Methods Cell Biol **48:** 3-29.

Li, C., R. I. Lin, M. C. Lai, P. Ouyang and W. Y. Tarn, 2003 Nuclear Pnn/DRS protein binds to spliced mRNPs and participates in mRNA processing and export via interaction with RNPS1. Mol Cell Biol **23:** 7363-7376.

Lieber, M. R., 2010 The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem **79:** 181-211.

Luna, R., A. G. Rondon and A. Aguilera, 2012 New clues to understand the role of THO and other functionally related factors in mRNP biogenesis. Biochim Biophys Acta **1819:** 514-520.

Ma, L., X. Gao, J. Luo, L. Huang, Y. Teng *et al.*, 2012 The Caenorhabditis elegans gene mfap-1 encodes a nuclear protein that affects alternative splicing. PLoS Genet **8:** e1002827.

Ma, L., and H. R. Horvitz, 2009 Mutations in the Caenorhabditis elegans U2AF large subunit UAF-1 alter the choice of a 3' splice site *in vivo*. PLoS Genet **5:** e1000708.

Matsuoka, S., B. A. Ballif, A. Smogorzewska, E. R. McDonald, 3rd, K. E. Hurov *et al.*, 2007 ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science **316:** 1160-1166.

McKinnon, P. J., and K. W. Caldecott, 2007 DNA strand break repair and human genetic disease. Annu Rev Genomics Hum Genet **8:** 37-55.

Phillips, E. R., and P. J. McKinnon, 2007 DNA double-strand break repair and development. Oncogene **26:** 7799-7808.

Pontier, D. B., and M. Tijsterman, 2009 A robust network of double-strand break repair pathways governs genome integrity during C. elegans development. Curr Biol **19:** 1384-1388.

Ramachandran, S., D. D. Tran, S. Klebba-Faerber, C. Kardinal, A. D. Whetton *et al.*, 2011 An ataxia-telangiectasia-mutated (ATM) kinase mediated response to DNA damage down-regulates the mRNA-binding potential of THOC5. RNA **17:** 1957-1966.

Robert, V. J., M. W. Davis, E. M. Jorgensen and J. L. Bessereau, 2008 Gene conversion and end-joining-repair double-strand breaks in the Caenorhabditis elegans germline. Genetics **180:** 673-679.

Savage, K. I., J. J. Gorski, E. M. Barros, G. W. Irwin, L. Manti *et al.*, 2014 Identification of a BRCA1-mRNA Splicing Complex Required for Efficient DNA Repair and Maintenance of Genomic Stability. Mol Cell **54:** 445-459.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim *et al.*, 2012 Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc **7:** 562-578.

Wang, P., P. J. Lou, S. Leu and P. Ouyang, 2002 Modulation of alternative pre-mRNA splicing *in vivo* by pinin. Biochem Biophys Res Commun **294:** 448-455.