Coupling DNA Damage and Repair: an Essential Safeguard during Programmed DNA Double- Strand Breaks?

Mireille Bétermier, Valérie Borde, Jean-Pierre de Villartay

Coupling DNA damage and repair: an essential safeguard during programmed DNA double strand breaks?

Mireille Bétermier[1*], Valérie Borde[2*], and Jean-Pierre de Villartay[3*]

1. Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France
2. Institut Curie, CNRS UMR3244, Sorbonne Université, Paris, France
3. Laboratory of Genome Dynamics in the Immune System, INSERM UMR1163, Université Paris Descartes Sorbonne Paris Cité, Institut Imagine, Paris, France

***Correspondence:** mireille.betermier@i2bc.paris-saclay.fr, valerie.borde@curie.fr, and devillartay@gmail.com.

Keywords (two to six): Programmed DNA double strand break (prDSB), Genome stability, DNA recombination and Repair.

Abstract

DNA double strand breaks (DSBs) are the most toxic DNA lesions given their oncogenic potential. Nevertheless, programmed DSBs (prDSBs) contribute to several biological processes. Formation of prDSBs is the price to pay to achieve these essential biological functions. Generated by domesticated PiggyBac transposases, prDSBs have been integrated in the life cycle of ciliates. Created by Spo11 during meiotic recombination, they constitute a driving force of evolution and ensure balanced chromosome content for successful reproduction. Produced by the RAG1/2 recombinase, they are required for the development of the adaptive immune system in many species. The co-evolution of processes that couple introduction of prDSBs to their accurate repair may constitute an effective safeguard against genomic instability.

Introduction

Living organisms are constantly exposed to genotoxic assaults, which can be of endogenous origin such as cellular respiration or exogenous sources such as radiations or chemical exposures. Several highly conserved DNA repair mechanisms have been selected during evolution to cope with these various damages and maintain genomic integrity. Among DNA lesions, DNA double strand breaks (DSBs) are considered the most toxic and at least two DNA repair pathways (homologous recombination (HR) and non-homologous end joining, or NHEJ) have evolved to cope with DSBs. In addition to repairing pathologic DSBs, these DNA repair pathways are also important for the repair of physiological DSBs or programmed DSBs (prDSBs) created during programmed genome rearrangements (PGR) in ciliates, meiotic recombination for sexual reproduction, and V(D)J recombination. Defects in these processes result in death of progeny (PGR), sterility or aneuploidy (meiotic recombination), and severe immune deficiency (V(D)J recombination). Therefore, the introduction of prDSBs is "the price to pay" for some physiological processes. One can argue that efficient ways to control prDSBs have co-evolved to avoid the deleterious consequences of their mis-repair. Here, we discuss the view that the timely and physical coupling of DNA damage and repair may represent an efficient safeguard during prDSBs.

Coupling DNA damage and NHEJ-mediated repair of prDSBs?

NHEJ is one of the two main DSB repair mechanisms. It operates in all phases of the cell cycle, in contrast to HR, which is excluded from G0/G1. Its catalytic process can be schematically divided into three steps: (i) the heterodimer Ku70/80 identifies and is recruited to the break, prior to the recruitment of the DNA-dependent protein kinase catalytic subunit DNA-PKcs, forming the DNAPK holoenzyme; (ii) If needed, DNA ends are processed ("cleaned") by DNA polymerases, nucleases, and kinases. The processing step is important during V(D)J recombination for the opening of RAG1/2-generated DNA hairpins by the nuclease Artemis; (iii) the DSB is resealed by DNA ligaseIV assisted by the Xlf, MRI and PAXX accessory factors (see [1] for a recent review on the actors of NHEJ).

Programmed Genome Rearrangement in ciliates: The prototypical example of DNA coupling between DNA damage and repair.

In ciliates, which constitute a monophyletic group of unicellular eukaryotes, the somatic and germline functions of chromosomes are separated into two distinct types of

nuclei coexisting in the same cytoplasm [2-4]: (i) the diploid micronucleus (MIC), transcriptionally silent during vegetative growth, undergoes meiosis and transmits the parental germline genome to the zygotic nucleus of the following generation; (ii) the polyploid somatic macronucleus (MAC), responsible for gene expression, directs the cell phenotype but is destroyed at each sexual cycle (Figure 1A). Ciliates make their new MAC from a copy of the zygotic nucleus, through a process involving several rounds of whole-genome endo-duplication and massive **programmed genome rearrangements (PGR)** triggered by the introduction of **tens of thousands of prDSBs** at multiple loci in the genome of the developing new MAC.

During PGR, *Paramecium tetraurelia* eliminates 25 to 30% of germline DNA from its somatic genome [5, 6], including repeated sequences (transposable elements (TEs), minisatellites) and ~45,000 TE-related short and noncoding **Internal Eliminated Sequences (IESs)**, usually found as single-copy elements and scattered all along the ~100-Mbp germline genome. While repeated DNA is eliminated in a heterogeneous manner, IES excision in *Paramecium* is precise at the nucleotide level. Because IESs interrupt almost half of genes in the germline, their efficient and precise excision ensures that the somatic genome is correctly assembled, a prerequisite for accurate gene expression and progeny survival.

Paramecium IESs are flanked by conserved TA dinucleotides. IES excision is initiated by 4-base staggered double-strand DNA cleavages centered on each flanking TA [7]. The endonuclease responsible for prDSB introduction at IES ends is a domesticated transposase called **PiggyMac (Pgm)**, whose conserved DDD catalytic triad, characteristic of transposases from the PiggyBac family, is essential for its function [8, 9]. Pgm is expressed during MAC development and localizes specifically in the developing new MAC by the time DNA elimination takes place. Five groups of Pgm-like domesticated PiggyBac transposases (PgmL1 to PgmL5) assist Pgm in cleaving DNA [10]. Each PgmL can interact individually with Pgm and is essential for IES excision genome-wide. None of them harbors a fully conserved catalytic site, suggesting that they play an architectural role during assembly of the IES excision complex, with PgmL1 and PgmL3 fine-tuning the precise positioning of DNA cleavage at IES boundaries. Once introduced, chromosomal prDSBs are repaired by the Ligase IV- and Xrcc4-dependent **classical NHEJ pathway (C-NHEJ)** [11]. NHEJ-mediated DSB repair of excision sites must be efficient and precise in order to preserve the coding capacity of the rearranged somatic genome, especially for intragenic IESs. This precision is likely driven through the pairing of conserved TAs at each 5' overhang, removal of the

terminal 5' nucleotide, and addition of one nucleotide at recessed 3' ends [7], ensuring that open reading frames are faithfully reconstituted upon IES excision.

Functional studies of Ku70/Ku80, the earliest-acting NHEJ factors that bind broken DNA ends and protect them against extensive resection, suggest that different solutions to the problem have evolved among ciliate species.

P. tetraurelia harbors two almost identical *KU70* genes and three more divergent *KU80* paralogs, a single of which (*KU80c*) is specifically expressed during MAC development [12]. In a similar way to Pgm, the development-specific Ku70/Ku80c heterodimer localizes in the new MAC during PGR and expression of *KU70* or *KU80c* is essential for the recovery of a functional somatic genome. Ku80c interacts with Pgm when both proteins are co-expressed in a heterologous system, a property shared by PiggyBac transposases and Ku proteins from other organisms [13]. Strikingly, the depletion of Ku80c abolishes DNA cleavage at IES ends, resulting in retention of all 45,000 IESs genome-wide [12]. In *Paramecium*, therefore, Ku interaction with Pgm during MAC development is thought to license Pgm-dependent DNA cleavage through a mechanism that remains to be established. Such **tight coupling between DSB introduction and repair** would ensure that DSBs are introduced only if Ku proteins are present to channel broken ends to the NHEJ pathway (Figure 1B).

In *Tetrahymena thermophila*, most of the ~12,000 IESs identified in the germline genome lie in non-coding regions [14]. They are also excised by domesticated transposases including a Pgm ortholog (Tpb2, see [15]) and a Pgm-like protein (Lia5, see [16]). Likewise, the C-NHEJ pathway carries out the repair of intergenic IES excision sites but Tpb2 does not require the presence of Ku80 to cleave IES ends [17], suggesting that excision of *Tetrahymena* intergenic IESs has not imposed the same constraints on the system as compared with excision of *Paramecium* IESs, the majority of which are intragenic. *T. thermophila* also harbors 12 intragenic IESs that do not rely on Tpb2 for their elimination. Remarkably, their excision is extremely precise and carried out by two distinct domesticated PiggyBac transposases, Tpb1 and Tpb6 (Figure 1B), both of which are fused to a Ku80 domain at their N-terminal end [18, 19]. Even though the biological importance of the Ku80 domain still must be tested experimentally, it is tempting to speculate that Ku-transposase fusions have been selected in *Tetrahymena* to secure the precise excision of intragenic IESs.

In ciliates, Ku80, whether as a separate factor (*P. tetraurelia*) or linked to the transposase (*T. thermophila*), appears to play an essential role, independent of its *bona fide* DNA repair factor function, upstream of the prDSB during PGR.

V(D)J recombination: coupling DNA damage and repair to avoid genomic instability?

V(D)J recombination is the molecular process by which exons encoding the variable domain of immunoglobulins and T cell receptors are assembled prior to their expression, thus ensuring the generation of an almost infinite possibility of antigenic recognition specificities by the adaptive immune system B and T lymphocytes [20]. It is essentially a mechanism related to "cut and paste" transposition, in which previously scattered variable (V), diversity (D), and joining (J) segments are physically associated on the DNA by a combinatorial somatic rearrangement process. V(D)J recombination is initiated by a site specific prDSB introduced by the lymphoid specific, domesticated transposase **RAG1 and RAG2** on recombination signal sequences (RSS) that flank all the rearranging V, D, and J segments [21, 22], a catalytic mechanism that evolved from the ancient *Transib* transposon [23].

Although V(D)J recombination proceeds through the introduction of the most toxic DNA lesion, it is the "price to pay" for the development of an efficient adaptive immunity, and has been selected for this purpose since the jawless vertebrates [24]. Indeed, abortive V(D)J recombination caused by either the inability to introduce the prDSB by RAG1/2 or the inefficiency in processing/repairing these breaks results in the early arrest of B and T cell maturation and the ensuing Severe Combined Immune Deficiency (**SCID**) both in humans and mice [25]. Mouse models also revealed the substantial oncogenic power of V(D)J recombination with the early onset of aggressive pro-B cell lymphomas in mice harboring NHEJ deficiency coupled with TP53 targeted inactivation [26]. Likewise, RAG1/2 may have oncogenic mutator functions driving leukemias in humans, such as observed in the context of leukemias harboring the ETV6-RUNX1 chromosomal translocation [27].

V(D)J recombination occurs during the G0/G1 phase of the cell cycle, resulting in prDSBs that are repaired through **NHEJ** [28]. The NHEJ factor Cernunnos was identified through the analysis of immune compromised patients with clinical presentation resembling that of patients with Lig4 syndromes but lacking DNA Ligase IV mutations [29]. Cernunnos shares sequence homology and interacts with XRCC4 (X4), two features at the basis of its concomitant identification as XRCC4-like factor (Xlf) [30]. XRCC4 and Cernunnos/Xlf interact through their globular head domains, forming long filaments readily visible by electronic microscopy and live imaging [31-35]. Given this structure, it was proposed that the

X4/Xlf filament could form a synapse that would facilitate DNA end tethering for their subsequent ligation by NHEJ.

It was anticipated that because Xlf represents a *bona fide* NHEJ factor, its deficiency would result in impaired V(D)J recombination. However, V(D)J recombination did not seem to be significantly affected *in vivo* in the lymphoid lineages either in Xlf-deficient human patients or mouse models [36-38]. First, B cell maturation in the bone marrow of Cernunnos patients was not arrested at the pro-B cell stage as expected and found in case of RAG1/2, Artemis, or DNA ligaseIV deficiencies. Second, Xlf KO mouse did not experience severe immune deficiency, only a slight decrease in lymphocyte numbers. Third, Xlf deficiency did not result in V(D)J recombination-driven chromosomal translocation or development of pro-B cell lymphoma when introduced on a $TP53^{-/-}$ background as observed with all the other models of NHEJxTP53 combined inactivations. Together, it appears that Cernunnos/Xlf is dispensable for V(D)J recombination although its absence clearly results in a major DNA repair defect when it comes to random or accidental DNA lesions such as the ones inflicted by ionizing radiations (IR).

One striking difference between IR- and V(D)J-driven DSBs is the presence of RAG1/2 in the latter. The RAG1/2 complex is known to remain on the DSB it has initiated as the post cleavage complex (PCC) [21], providing a possible means to tether DNA ends, which would be redundant to the expected function of the X4/Xlf filament (Figure 2A). Under this hypothesis, the sole presence of the PCC would provide a DNA repair synapse complementing the absence of the X4/Xlf filament during V(D)J recombination while such synapse would be missing at genotoxic-driven DSBs. The stability of the PCC relies on the C terminus region of RAG2, a region outside of the core and not essential for V(D)J recombination, as shown *in vitro* and *in vivo* in the $RAG2^{cc}$ mouse model specifically engineered to restrict RAG2 to its core domain [39, 40]. V(D)J recombination is not grossly affected in $RAG2^{cc}$ or $Xlf^{-/-}$ single mutant conditions, but is fully abrogated in $RAG2^{cc}xXlf^{-/-}$ mice, resulting in SCID animals devoid of mature B and T lymphocytes [41]. Based on these observations a 2-synapse model (Fig. 2A) was proposed in which the PCC complex on one hand and the X4/Xlf filament on the other hand help maintain genome integrity during V(D)J recombination. A similar functional redundancy operating specifically during V(D)J recombination was observed between Xlf and ATM or H2AX [42], and 53BP1 [43, 44] as well as between Xlf and the recently described NHEJ factors PAXX [45-50] or MRI [51] thus establishing a "synthetic dysfunction" among these factors (Fig. 2B). $RAG2^{cc}$ and ATM or PAXX combined deficiencies do not result in impaired V(D)J recombination, suggesting that

these three factors are epistatic, opposite to that of Xlf (Figure 2B) [41, 49]. It will be interesting to better understand the status of XRCC4 in this two-tier mechanism given its physical interaction with Xlf. However, XRCC4 KO mice are embryonic lethal and demonstrate impaired V(D)J recombination in fetuses because XRCC4 is required for DNA-ligase IV stability [52, 53]. Hence, XRCC4 KO phenocopies DNA-Lig4 KO condition. In contrast, the absence of immune deficiency in human primordial dwarfism with microcephaly syndrome caused by hypomorphic (yet severe) XRCC4 mutations suggests that, like Xlf, XRCC4 is not required for V(D)J recombination [54].

The demonstration of these functional interplays between RAG2, Xlf, and other DNA repair factors during V(D)J recombination suggests the possible existence of a "coupling" of DNA damage and repair during V(D)J recombination like the one described during PGR in ciliates. The challenge is now to fully understand how functional links between the RAG1/2 complex and the DNA repair apparatus translate into physical interaction of key components. Interestingly with regard to the analogy with PGR in ciliates, Raval et al. reported on the interaction of RAG1 with the Ku70/Ku80 complex [55]. Moreover, an interaction between RAG1 and another critical DNA repair factor, MDC1, was also reported [56]. The existence of a physical link between RAG1/2 and the DNA repair machinery certainly accredits the hypothesis of a DNA damage-repair coupling during V(D)J recombination.

Meiotic recombination: Homologous Recombination is also concerned

As opposed to NHEJ, homologous recombination uses DNA sequence homology on an intact DNA template to repair the broken DNA molecule after a DSB. The repair template can be located on the sister chromatid, on a homologous chromosome, or elsewhere in the genome. The first step of homologous recombination is the resection of the 5' ends of the DSB, first by the MRE11 complex, then by EXO1 and BLM/DNA2, which generates protruding 3' ends that invade the repair DNA template, through the action of a RecA-related recombinase, such as Rad51 [57]. In somatic cells, the preferred DNA repair template is the sister chromatid and therefore, homologous recombination is restricted to the G2/M phases of the cell cycle. During the meiotic prophase of sexually reproducing organisms, the topoisomerase-like protein **Spo11** initiates **meiotic recombination** by introducing hundreds of prDSBs along chromosomes. These prDSBs are important for recognition and pairing of homologs and a few of them will be repaired by crossovers, generating a physical link between homologs essential for their accurate segregation into the future gametes [58].

Meiotic recombination is a risky business for genome integrity of germ cells. Indeed, the burden of prDSBs that are introduced by Spo11 during meiotic recombination is at high risk of generating unwanted translocations or chromosome rearrangements, and their formation is therefore highly controlled by the use of several processes specific to meiotic cells. First, the DSB formation and repair steps both take place in a specific chromosome compartment, **the chromosome axis**, from which chromatin loops emanate, and which will be the place where the homologs become fully aligned within the synaptonemal complex (Figure 3). The DNA sequences that are cleaved by **Spo11** are preferentially located on the chromatin loops, although these sequences interact with proteins present along the chromosome axis, implying a spatial folding of the loop towards the chromosome axis during recombination. Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination [59]. This allows both to easily control the number of prDSBs generated, and also to physically "isolate" the cleaved DNA sequence from non-allelic/non-homologous sequences. This **physical tethering of the sequence to be cleaved** onto the chromosome axis is facilitated by a specific histone modification, H3K4me3, which bridges the DSB sequences to the DSB proteins located on the chromosome axis [60] thus ensuring that **meiotic prDSBs are formed within the correct spatial context**. Another regulation is exerted at the level of the "DSB forming complex". Indeed, in all organisms studied, the catalytic subunit Spo11 alone is not sufficient for DSB formation to take place, and multiple other proteins (9 in budding yeast, at least 5 in mammals) are required [61-63]. Mainly discovered in budding yeast, these proteins now appear conserved among many organisms, although they lack clear sequence homology, which renders them difficult to identify. These proteins form several proposed subcomplexes that interact together to promote DSB formation: a "core complex" composed of Spo11 (homolog of the catalytic TopoVI-A subunit) and a homolog of the TopoVI-B subunit [64, 65], a "RMM" complex proposed to interact directly with the chromosome axis, and other components that can vary depending on the species [61]. Among the Spo11 protein partners required for DSB formation or localization, several are also important for DSB repair: the **Mre11 complex** (in budding yeast, the worm *C. elegans,* and maybe in mammals), Narya in the fruit fly *Drosophila*, and PRDM9, which, is essential for their targeting to specific DNA sequences in humans and mice, despite not being essential for DSB formation.

The MRE11 complex

The MRE11 complex is well known for the signaling and processing of DSBs [66]. Its endonuclease activity is required to process meiotic DSBs, by removing the Spo11 protein together with a short oligonucleotide from DSB ends [67]. In addition, at least in the budding yeast *S. cerevisiae* and in the nematode *C. elegans*, this complex is integrated in the process that is **required for prDSB formation by Spo11** [68, 69]. The functions of the MRE11 complex in DSB formation and repair are genetically separable, and its function for DSB formation seems to involve mainly the Mre11 and Rad50 subunits [70, 71]. In mammals, whether the MRE11 complex is needed for Spo11-induced DSB formation is not clear at the moment, because members of the MRE11 complex are essential for viability, so only hypomorphic mutants of the MRE11 complex could be tested in the mouse for their effect on fertility, and they showed only defects in meiotic DSB repair [72]. Conditional knock out of the Mre11 complex specifically in meiotic cells would be required to know if it is also required for meiotic DSB formation in mammals. In budding yeast, components of the MRE11 complex appear to directly interact with other DSB formation proteins [73], implying a specialization of the MRE11 complex for the immediate signaling and processing of these programmed meiotic DSBs. Likewise, in the plant *Arabidopsis thaliana,* although not strictly required for meiotic DSB formation, a member of the MRE11 complex directly interacts with a protein required for DSB formation ([74] and Mathilde Grelon, personal communication). It is attractive to propose that incorporating the signaling and repair Mre11 complex in the step of meiotic DSB formation allows the immediate processing of meiotic DSBs into the homologous recombination pathway.

Narya

Another example of coupling between meiotic DSB formation and repair comes from the fruit fly *Drosophila*, where a protein, **Narya**, fulfills functions both for formation and repair of meiotic DSBs with a crossing over, to ensure proper homolog segregation and successful meiosis [75]. Narya is a RING finger protein that is redundant with another related protein, Nanya, for meiotic DSB formation and repair. It also interacts with Vilya, a third protein of the family, required for DSB formation and interacting with MEI22, one of the *Drosophila* DSB proteins [76]. Interestingly, a separation-of-function allele of Narya, in its RING finger domain, shows that Narya is not only involved in DSB formation, but also required for their repair as a crossover. In addition, the three proteins show two sequential localization patterns; first, early with DSB sites and then, to crossover sites [75, 76]. This illustrates again a double function for meiotic DSB formation and their repair into crossovers within a single protein, therefore directly coupling these two steps of meiotic recombination.

PRDM9

The histone methyltransferase **PRDM9**, responsible for targeting meiotic prDSBs to its consensus DNA binding sequence in many organisms including mice and humans, is also critical for meiotic prDSB repair. In its absence, Spo11 forms meiotic prDSBs at "default" chromatin accessible locations within functional genomic elements, which are not well repaired for unclear reasons [77, 78]. In addition, if PRDM9 is present on only one homolog owing to a polymorphism affecting its consensus binding sequence, this also creates problems in DSB repair [79]. This suggests that the symmetric binding of PRDM9 to both homologs, thanks to its sequence specificity, facilitates the repair, perhaps by bringing close together into the chromosome axis the two chromatid sequences that will experience the recombination event (Figure 3). In favor of this hypothesis, PRDM9 physically interacts with several components of the chromosome axis [80]. PRDM9 therefore represents yet another example of coupling prDSB formation and repair within a single protein during programmed meiotic recombination.

Concluding remarks

Besides meiosis, PGR, and V(D)J recombination, prDSBs have been identified during signal-induced transcription in several experimental settings (see [81] for a recent review). These activity-induced prDSBs occur primarily in early response genes and are introduced by the topoisomerase IIβ. This is in particular the case in the response of MCF-7 cells to estradiol [82] or activation through glucocorticoid receptors [83]. prDSBs also occur *in vivo* and *in vitro* upon neuronal activity [84, 85]. Interestingly, in the case of the glucocorticoid receptor-induced transcriptional activation, recruitment of Top2b and Ku70/86 via the BRG1 transcription activator-containing chromatin-remodeling complex is required at GR-responsive promoters [83]. This suggests that a subset of DNA repair factors may be in place before prDSBs, thus accrediting the hypothesis of a possible coupling of DNA damage and repair during signal-induced transcription. Finally, beside Pgm in ciliates and RAG1/2, two other domesticated transposases have been identified in human cells, the piggyBac transposable element-derived 5 (PGBD5) [86] and THAP9, related to the *Drosophila* P-element transposase [87], the exact functions of which are presently unknown. The deregulated expression of PGBD5 in rhabdoid tumors in children participates in the oncogenic transformation by promoting site-specific DNA rearrangements within tumor suppressor genes [88]. Whether this deleterious activity of PGBD5 is counteracted in its

physiological context by a mechanism related to DNA damage-repair coupling is of course speculative but represents an interesting issue to follow. If this were to be true, it would reinforce the idea that DNA damage-repair coupling may represent an essential step in the transposase domestication process.

Altogether coupling DNA damage and repair may have co-evolved with prDSBs to ensure their efficient repair and thus avoid any associated genomic instability. Several questions remain to be addressed: (i) what are the exact mechanisms governing this coupling, (ii) given the oncogenic power of DSBs, what would be the consequences of losing this coupling (see Outstanding Questions)?

Acknowledgments

We thank Mathilde Grelon (INRA, Versailles) for personal communication. Work in our respective laboratories is supported by institutional grants from INSERM, CNRS, ANR ("Investissements d'avenir" program ANR-10-IAHU-01; ANR-13-PRTS-0004; ANR-18-CE12-0018; ANR-14-CE10-0005-01; and ANR-18-CE12-0005-02), INCa (PLBIO16-280), and grants from Ligue Nationale Contre le Cancer (Equipe Labellisée), Fondation pour la Recherche Médicale (Equipe FRM EQU201903007785) and AT-Europe Fundation.

References

1 Hnizda, A. and Blundell, T.L. (2019) Multicomponent assemblies in DNA-double-strand break repair by NHEJ. *Current opinion in structural biology* 55, 154-160

2 Betermier, M. and Duharcourt, S. (2014) Programmed Rearrangement in Ciliates: Paramecium. *Microbiology spectrum* 2

3 Yao, M.C., *et al.* (2014) Programmed Genome Rearrangements in Tetrahymena. *Microbiology spectrum* 2

4 Yerlici, V.T. and Landweber, L.F. (2014) Programmed Genome Rearrangements in the Ciliate Oxytricha. *Microbiology spectrum* 2

5 Arnaiz, O., *et al.* (2012) The *Paramecium* germline genome provides a niche for intragenic parasitic DNA: evolutionary dynamics of internal eliminated sequences. *PLoS genetics* 8, e1002984

6 Guerin, F., *et al.* (2017) Flow cytometry sorting of nuclei enables the first global characterization of Paramecium germline DNA and transposable elements. *BMC genomics* 18, 327

7 Gratias, A. and Bétermier, M. (2003) Processing of double-strand breaks is involved in the precise excision of *Paramecium* IESs. *Mol. Cell. Biol.* 23, 7152-7162

8 Baudry, C.*, et al.* (2009) PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate Paramecium tetraurelia. *Genes & development* 23, 2478-2483

9 Dubois, E.*, et al.* (2017) Multimerization properties of PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements. *Nucleic Acids Res* 45, 3204-3216

10 Bischerour, J.*, et al.* (2018) Six domesticated PiggyBac transposases together carry out programmed DNA elimination in Paramecium. *Elife* 7

11 Kapusta, A.*, et al.* (2011) Highly precise and developmentally programmed genome assembly in *Paramecium* requires ligase IV-dependent end joining. *PLoS Genet* 7, e1002049

12 Marmignon, A.*, et al.* (2014) Ku-mediated coupling of DNA cleavage and repair during programmed genome rearrangements in the ciliate Paramecium tetraurelia. *PLoS genetics* 10, e1004552

13 Jin, Y.*, et al.* (2017) DNA-PK facilitates piggyBac transposition by promoting paired-end complex formation. *Proc Natl Acad Sci U S A* 114, 7408-7413

14 Hamilton, E.P.*, et al.* (2016) Structure of the germline genome of *Tetrahymena thermophila* and relationship to the massively rearranged somatic genome. *Elife* 5

15 Cheng, C.Y.*, et al.* (2010) A domesticated piggyBac transposase plays key roles in heterochromatin dynamics and DNA cleavage during programmed DNA deletion in *Tetrahymena thermophila*. *Molecular biology of the cell* 21, 1753-1762

16 Shieh, A.W. and Chalker, D.L. (2013) LIA5 is required for nuclear reorganization and programmed DNA rearrangements occurring during *Tetrahymena* macronuclear differentiation. *PLoS One* 8, e75337

17 Lin, I.T.*, et al.* (2012) An essential role for the DNA breakage-repair protein Ku80 in programmed DNA rearrangements in Tetrahymena thermophila. *Molecular biology of the cell* 23, 2213-2225

18 Cheng, C.Y.*, et al.* (2016) The piggyBac transposon-derived genes TPB1 and TPB6 mediate essential transposon-like excision during the developmental rearrangement of key genes in Tetrahymena thermophila. *Genes & development* 30, 2724-2736

19 Feng, L., *et al.* (2017) A germline-limited piggyBac transposase gene is required for precise excision in *Tetrahymena* genome rearrangement. *Nucleic Acids Res* 45, 9481-9502

20 Jung, D., *et al.* (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annual review of immunology* 24, 541-570

21 Schatz, D.G. and Swanson, P.C. (2011) V(D)J recombination: mechanisms of initiation. *Annual review of genetics* 45, 167-202

22 Zhang, Y., *et al.* (2019) Transposon molecular domestication and the evolution of the RAG recombinase. *Nature* 569, 79-84

23 Liu, C., *et al.* (2019) Structures of a RAG-like transposase during cut-and-paste transposition. *Nature* Nov 13. doi: 10.1038/s41586-019-1753-7

24 Boehm, T. and Swann, J.B. (2014) Origin and evolution of adaptive immunity. *Annual review of animal biosciences* 2, 259-283

25 de Villartay, J.P., *et al.* (2003) The mechanisms of immune diversification and their disorders. *Nature reviews. Immunology* 3, 962-972

26 Ferguson, D.O. and Alt, F.W. (2001) DNA double strand break repair and chromosomal translocation: lessons from animal models. *Oncogene* 20, 5572-5579

27 Papaemmanuil, E., *et al.* (2014) RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nature genetics* 46, 116-125

28 Lieber, M.R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry* 79, 181-211

29 Buck, D., *et al.* (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* 124, 287-299

30 Ahnesorg, P., *et al.* (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124, 301-313

31 Andres, S.N., *et al.* (2012) A human XRCC4-XLF complex bridges DNA. *Nucleic acids research* 40, 1868-1878

32 Hammel, M., *et al.* (2010) XLF regulates filament architecture of the XRCC4.ligase IV complex. *Structure* 18, 1431-1442

33 Reid, D.A., *et al.* (2015) Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proceedings of the National Academy of Sciences of the United States of America* 112, E2575-2584

34 Ropars, V., *et al.* (2011) Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. *Proceedings of the National Academy of Sciences of the United States of America* 108, 12663-12668

35 Wu, Q., *et al.* (2011) Non-homologous end-joining partners in a helical dance: structural studies of XLF-XRCC4 interactions. *Biochemical Society transactions* 39, 1387-1392, suppl 1382 p following 1392

36 Li, G., *et al.* (2008) Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in V(D)J recombination. *Molecular cell* 31, 631-640

37 van der Burg, M. and Gennery, A.R. (2011) Educational paper. The expanding clinical and immunological spectrum of severe combined immunodeficiency. *European journal of pediatrics* 170, 561-571

38 Vera, G., *et al.* (2013) Cernunnos deficiency reduces thymocyte life span and alters the T cell repertoire in mice and humans. *Molecular and cellular biology* 33, 701-711

39 Deriano, L., *et al.* (2011) The RAG2 C terminus suppresses genomic instability and lymphomagenesis. *Nature* 471, 119-123

40 Liang, H.E., *et al.* (2002) The "dispensable" portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B and T cell development. *Immunity* 17, 639-651

41 Lescale, C., *et al.* (2016) RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG complex in DNA repair. *Nature communications* 7, 10529

42 Zha, S., *et al.* (2011) ATM damage response and XLF repair factor are functionally redundant in joining DNA breaks. *Nature* 469, 250-254

43 Oksenych, V., *et al.* (2012) Functional redundancy between repair factor XLF and damage response mediator 53BP1 in V(D)J recombination and DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* 109, 2455-2460

44 Liu, X., *et al.* (2012) Overlapping functions between XLF repair protein and 53BP1 DNA damage response factor in end joining and lymphocyte development.

Proceedings of the National Academy of Sciences of the United States of America 109, 3903-3908

45 Abramowski, V.*, et al.* (2018) PAXX and Xlf interplay revealed by impaired CNS development and immunodeficiency of double KO mice. *Cell death and differentiation* 25, 444-452

46 Balmus, G.*, et al.* (2016) Synthetic lethality between PAXX and XLF in mammalian development. *Genes & development* 30, 2152-2157

47 Hung, P.J.*, et al.* (2017) Deficiency of XLF and PAXX prevents DNA double-strand break repair by non-homologous end joining in lymphocytes. *Cell cycle* 16, 286-295

48 Kumar, V.*, et al.* (2016) PAXX and XLF DNA repair factors are functionally redundant in joining DNA breaks in a G1-arrested progenitor B-cell line. *Proceedings of the National Academy of Sciences of the United States of America* 113, 10619-10624

49 Lescale, C.*, et al.* (2016) Specific Roles of XRCC4 Paralogs PAXX and XLF during V(D)J Recombination. *Cell reports* 16, 2967-2979

50 Liu, X.*, et al.* (2017) PAXX promotes KU accumulation at DNA breaks and is essential for end-joining in XLF-deficient mice. *Nature communications* 8, 13816

51 Hung, P.J.*, et al.* (2018) MRI Is a DNA Damage Response Adaptor during Classical Non-homologous End Joining. *Molecular cell* 71, 332-342 e338

52 Frank, K.M.*, et al.* (2000) DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Molecular cell* 5, 993-1002

53 Gao, Y.*, et al.* (2000) Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature* 404, 897-900

54 de Villartay, J.P. (2015) When natural mutants do not fit our expectations: the intriguing case of patients with XRCC4 mutations revealed by whole-exome sequencing. *EMBO molecular medicine* 7, 862-864

55 Raval, P.*, et al.* (2008) Evidence for Ku70/Ku80 association with full-length RAG1. *Nucleic acids research* 36, 2060-2072

56 Coster, G.*, et al.* (2012) A dual interaction between the DNA damage response protein MDC1 and the RAG1 subunit of the V(D)J recombinase. *The Journal of biological chemistry* 287, 36488-36498

57 Kowalczykowski, S.C. (2015) An Overview of the Molecular Mechanisms of Recombinational DNA Repair. *Cold Spring Harbor perspectives in biology* 7

58 Hunter, N. (2015) Meiotic Recombination: The Essence of Heredity. In *Cold Spring Harbor perspectives in biology*

59 Panizza, S.*, et al.* (2011) Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. *Cell* 146, 372-383

60 Borde, V. and de Massy, B. (2013) Programmed induction of DNA double strand breaks during meiosis: setting up communication between DNA and the chromosome structure. *Current opinion in genetics & development* 23, 147-155

61 Lam, I. and Keeney, S. (2014) Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harbor perspectives in biology* 7, a016634

62 Robert, T.*, et al.* (2016) A new light on the meiotic DSB catalytic complex. *Semin Cell Dev Biol* 54, 165-176

63 Stanzione, M.*, et al.* (2016) Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. *Nat Cell Biol* 18, 1208-1220

64 Robert, T.*, et al.* (2016) The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* 351, 943-949

65 Vrielynck, N.*, et al.* (2016) A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science* 351, 939-943

66 Oh, J. and Symington, L.S. (2018) Role of the Mre11 Complex in Preserving Genome Integrity. *Genes (Basel)* 9

67 Neale, M.J.*, et al.* (2005) Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* 436, 1053-1057

68 Chin, G.M. and Villeneuve, A.M. (2001) C. elegans mre-11 is required for meiotic recombination and DNA repair but is dispensable for the meiotic G(2) DNA damage checkpoint. *Genes Dev* 15, 522-534

69 Usui, T.*, et al.* (2001) A DNA damage response pathway controlled by Tel1 and the Mre11 complex. *Mol Cell* 7, 1255-1266

70 Girard, C.*, et al.* (2018) Interdependent and separable functions of Caenorhabditis elegans MRN-C complex members couple formation and repair of meiotic DSBs. *Proc Natl Acad Sci U S A* 115, E4443-E4452

71 Oh, J., *et al.* (2016) Xrs2 Dependent and Independent Functions of the Mre11-Rad50 Complex. *Mol Cell* 64, 405-415

72 Cherry, S.M., *et al.* (2007) The Mre11 complex influences DNA repair, synapsis, and crossing over in murine meiosis. *Curr Biol* 17, 373-378

73 Arora, C., *et al.* (2004) Antiviral protein Ski8 is a direct partner of Spo11 in meiotic DNA break formation, independent of its cytoplasmic role in RNA metabolism. *Mol. Cell* 13, 549-559

74 Waterworth, W.M., *et al.* (2015) Arabidopsis TAF1 is an MRE11-interacting protein required for resistance to genotoxic stress and viability of the male gametophyte. *Plant J* 84, 545-557

75 Lake, C.M., *et al.* (2019) Narya, a RING finger domain-containing protein, is required for meiotic DNA double-strand break formation and crossover maturation in Drosophila melanogaster. *PLoS Genet* 15, e1007886

76 Lake, C.M., *et al.* (2015) Vilya, a component of the recombination nodule, is required for meiotic double-strand break formation in Drosophila. *Elife* 4, e08287

77 Brick, K., *et al.* (2012) Genetic recombination is directed away from functional genomic elements in mice. *Nature* 485, 642-645

78 Grey, C., *et al.* (2018) PRDM9, a driver of the genetic map. *PLoS Genet* 14, e1007479

79 Davies, B., *et al.* (2016) Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature* 530, 171-176

80 Parvanov, E.D., *et al.* (2017) PRDM9 interactions with other proteins provide a link between recombination hotspots and the chromosomal axis in meiosis. *Molecular biology of the cell* 28, 488-499

81 Puc, J., *et al.* (2017) Physiological functions of programmed DNA breaks in signal-induced transcription. *Nature reviews. Molecular cell biology* 18, 471-476

82 Ju, B.G., *et al.* (2006) A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science* 312, 1798-1802

83 Trotter, K.W., *et al.* (2015) Glucocorticoid Receptor Transcriptional Activation via the BRG1-Dependent Recruitment of TOP2beta and Ku70/86. *Molecular and cellular biology* 35, 2799-2817

84 Madabhushi, R., *et al.* (2015) Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell* 161, 1592-1605

85 Suberbielle, E., *et al.* (2013) Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-beta. *Nature neuroscience* 16, 613-621

86 Henssen, A.G., *et al.* (2015) Genomic DNA transposition induced by human PGBD5. *eLife* 4

87 Majumdar, S., *et al.* (2013) The human THAP9 gene encodes an active P-element DNA transposase. *Science* 339, 446-448

88 Henssen, A.G., *et al.* (2017) PGBD5 promotes site-specific oncogenic mutations in human tumors. *Nature genetics* 49, 1005-1014

Highlights

- Several biological processes (meiosis, V(D)J recombination, PGR in ciliates, signal-induced transcription) proceed via introduction of programmed DSBs (prDSBs).
- DSBs being the most toxic DNA lesions, as potentially oncogenic, prDSBs are likely associated with very efficient, multi-layered DNA repair mechanisms. Coupling DNA damage and repair is one critical layer.
- Ku80 is a critical factor to link DNA damage and repair during PGR in ciliates
- The C terminus of RAG2 may be responsible for the DNA damage-repair coupling during V(D)J recombination as a safeguard against genome instability.
- During meiotic recombination, a specific pathway ensures that meiotic DSBs are formed within the correct spatial chromosomal context.
- The MRE11 complex is required for the formation of prDSBs by Spo11 during meiotic recombination.
- DNA damage-repair coupling may represent an essential step in the domestication process of PiggyMac, RAG1/2 and other transposases.

Outstanding Questions Box

- What are the exact mechanisms and critical players of DNA damage-repair coupling during prDSBs?
- What are the downstream consequences of uncoupling DNA damage-repair during prDSBs driven processes (for example: genetic instability and tumor development, cellular degeneracy)?

- Is DNA damage-repair coupling a general rule that applies to all biological processes that proceed through prDSBs?

Glossary

- **DSBs:** DNA double strand breaks can be "accidental" as a result of environment insult or "programmed" (prDSBs) as part of essential physiological processes (meiosis, V(D)J recombination, PGR in ciliates).
- **HR:** The Homologous Recombination DNA repair pathway is one of the two main mechanisms, with NHEJ, to repair DSBs. It operates exclusively in S phase of the cell cycle when a sister chromatid is available as template.
- **Meiosis:** Meiosis is the process of chromosome segregation during the formation of gametes. prDSBs are introduced by Spo11 during meiosis
- **MRE11 complex:** Composed of meiotic recombination 11 (MRE11), RAD50 and Nijmegen breakage syndrome 1 (NBS1 or Nibrin, Xrs2 in budding yeast). Acts in the sensing and signaling of DSBs. The endonuclease activity of MRE11 is essential for the processing of protein-linked meiotic DSBs.
- **NHEJ:** The Non Homologous End Joining DNA repair pathway is one of the two main mechanisms, with HR, to repair DSBs. It operates in all phases of the cell cycle.
- **PGR:** Programmed Genome Rearrangement that reproducibly eliminates large fractions of germline DNA (25 to 95% according to species) during formation of the Macronucleus (MAC) during the ciliate sexual cycle
- **PiggyMac (Pgm):** Domesticated transposase responsible for PGR in the ciliate *Paramecium*.
- **RAG1 & RAG2:** The Recombination Activating Gene 1 & 2 constitute the domesticated transposase initiating V(D)J recombination in immature B and T lymphocytes.
- **SCID:** Severe Combined Immune Deficiency is a rare condition in humans (and engineered mouse models) characterized by a profound defect in the development/function of the adaptive immune system. Impaired V(D)J recombination results in SCID.
- **Spo11:** Catalytic subunit of a topoisomerase-like complex that introduces prDSBs during meiosis

- **V(D)J Recombination:** Lymphoid-specific somatic DNA rearrangement process of immunoglobulin and T cell receptor (TCR) genes initiated by the RAG1/2 factors aimed at generating the antigenic diversity (repertoire) of the adaptive immune system.

Figure Legends

Figure 1: Coupling between DSB formation and repair during programmed genome rearrangements in ciliates

A) In the ciliate *Paramecium tetraurelia,* vegetative cells harbor two MICs (black) and one MAC (gray). During sexual reproduction, MICs undergo meiosis, a single meiotic product divides and yields two gametic nuclei (black), while all others are degraded (light gray). During conjugation, following reciprocal exchange of gametic nuclei between mating partners, the resident and incoming nuclei fuse to give a diploid zygotic nucleus (black); during autogamy (a self-fertilization process), the zygotic nucleus results from the fusion of both gametic nuclei from the same cell. The zygotic nucleus then divides twice: two of the resulting nuclei become the new MICs (black) and the other two differentiate into new MACs (hatched purple and blue). Programmed DNA elimination takes place during MAC development. Following mitosis of the new MICs, the new MACs segregate into each daughter cell. Throughout the whole process, the old MAC is fragmented and is eventually lost after a few cell divisions. **B)** In *Paramecium*, the presence of Ku is required for Pgm to cleave DNA at IES ends, indicative of tight coupling between DSB formation and repair. *Tetrahymena* intragenic IESs are excised by domesticated transposases (Tpb1 and Tpb6) fused to a Ku80-like domain, the functional importance of which has not been established.

Figure 2: Coupling between DSB formation and repair during V(D)J Recombination?

A) Two complementary synapses (the XLF/XRCC4 filament and the RAG1/2 post cleavage complex) ensure DNA end tethering during V(D)J recombination. The loss of both synapses results in major genomic instability with the development of lymphomas. **B)** Synthetic dysfunction of V(D)J recombination highlights two complementary axes during the repair phase of V(D)J recombination. RAG2 participates in the ATM-PAXX-MRI axis, complementary to the XLF-XRCC4 axis.

Figure 3: Coupling between DSB formation and repair during programmed meiotic recombination.

During meiosis, homologous chromosomes experience homologous recombination, which promotes their pairing, crossover and physical attachment through the chiasmata. Chromosomes are organized around a protein axis (blue and red lines for maternal and paternal chromosomes), from which chromatin loops emanate (gray). DSBs (yellow lightening) are formed to initiate recombination, which culminates into at least one crossover pair of homologs. Programmed DSBs occur at hotspot sequences, which become transiently tethered to the chromosome axis where DSB proteins (including the catalytic subunit Spo11) are located. Among these, the Mre11 complex is necessary both for the formation of DSBs by Spo11 and for their repair, highlighting the specialized coupling between DSB formation and repair during meiosis. Other examples of such coupling are discussed in the text.