

## **A probabilistic method to infer the role of INDEL-generating DNA repair mechanisms on genome diversity across species**

Guy Tadmor, Avraham A. Levy, Fabrizio Mafessoni

Weizmann Institute of Science, Department of Plant and Environmental Sciences

Genome integrity is maintained by molecular mechanisms which repair DNA whenever molecular, physical, or chemical processes inflict damage upon it. Double Strand Breaks (DSBs) can have genetic outcomes ranging from large-scale chromosomal rearrangements to small indels. DSBs are frequently repaired through different mechanisms, such as non-homologous-end joining (NHEJ), microhomology-mediated end joining (MMEJ), and Synthesis-dependent MMEJ (SD-MMEJ), which leave specific genomic signatures in the genome in the form of short insertions and deletions (INDELs). These mechanisms have been characterized in experimental systems, where DSBs can be induced using molecular scissors like CRISPR-Cas9, but their expected footprints have not been analyzed thoroughly in whole genome datasets due to limitations in the currently available methods to survey indels and to infer the DNA repair mechanisms underlying natural variation among individual genomes. Here we report a probabilistic method to infer INDEL-generating DNA repair mechanisms through a comparison of patterns of DNA repair outcomes in experimentally-induced and naturally occurring INDELs. Using this method, we estimate the proportion of different DNA repair mechanisms occurring in plant and animal genomes under different biological scenarios and genomic contexts, providing a better understanding of their role in evolution and in the origin of new genetic diversity.