

Communication

Composition and Structure of *Arabidopsis thaliana* Extrachromosomal Circular DNAs Revealed by Nanopore Sequencing

Pavel Merkulov^{1,2} , Ekaterina Egorova² and Ilya Kirov^{1,2,*} ¹ Moscow Institute of Physics and Technology, 141701 Dolgoprudny, Russia; paulmerkulov97@gmail.com² All-Russia Research Institute of Agricultural Biotechnology, 127550 Moscow, Russia;

egorova.ekaterina20021120@gmail.com

* Correspondence: kirovez@gmail.com

Abstract: Extrachromosomal circular DNAs (eccDNAs) are enigmatic DNA molecules that have been detected in a range of organisms. In plants, eccDNAs have various genomic origins and may be derived from transposable elements. The structures of individual eccDNA molecules and their dynamics in response to stress are poorly understood. In this study, we showed that nanopore sequencing is a useful tool for the detection and structural analysis of eccDNA molecules. Applying nanopore sequencing to the eccDNA molecules of epigenetically stressed *Arabidopsis* plants grown under various stress treatments (heat, abscisic acid, and flagellin), we showed that TE-derived eccDNA quantity and structure vary dramatically between individual TEs. Epigenetic stress alone did not cause eccDNA up-regulation, whereas its combination with heat stress triggered the generation of full-length and various truncated eccDNAs of the *ONSEN* element. We showed that the ratio between full-length and truncated eccDNAs is TE- and condition-dependent. Our work paves the way for further elucidation of the structural features of eccDNAs and their connections with various biological processes, such as eccDNA transcription and eccDNA-mediated TE silencing.

Keywords: extrachromosomal circular DNAs; *Arabidopsis*; long-read sequencing; transposable elements



Citation: Merkulov, P.; Egorova, E.; Kirov, I. Composition and Structure of *Arabidopsis thaliana*

Extrachromosomal Circular DNAs Revealed by Nanopore Sequencing. *Plants* **2023**, *12*, 2178. <https://doi.org/10.3390/plants12112178>

Academic Editors: Junjie Zou, Miaoyun Xu and Yong-Gu Cho

Received: 2 May 2023

Revised: 19 May 2023

Accepted: 29 May 2023

Published: 30 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Extrachromosomal circular DNAs (eccDNAs) are a type of double-stranded DNA that has been found in the cells of a variety of organisms, including humans, animals, and plants. The first report regarding DNA circularization was made in 1962, followed by Hotta and Basel's experiments, in which eccDNAs were discovered in boar sperm using electron microscopy [1]. Their results provided the first indication of the occurrence of eukaryotic DNA in a circular form. Circular DNAs were initially considered to originate from the mitochondria, as they were co-isolated with organellar DNA [2]. After examining mutant yeast lacking mitochondria, it was discovered that circular DNAs have nuclear origin [3]. Using electron microscopy, the first plant nuclei-derived small circular DNA forms were investigated in wheat and tobacco (Kinoshita et al., 1985). Two-dimensional (2D) gel electrophoresis was used to distinguish eccDNA structural forms in different families of higher plants (*Asteraceae*, *Brassicaceae*, *Fabaceae*, and *Poaceae*). Finally, the advent of high-throughput sequencing and bioinformatic methods facilitated insight into the diversity of circular molecules that comprise a plant circulome [4,5]. Genome-wide analysis of eccDNA-producing loci has been carried out for several plant species [5–11]. These studies indicated that different genomic loci can produce eccDNAs in plants, including genic and intergenic regions. A major source of eccDNAs in plant genomes is transposable elements (TEs) [4]. It was proposed that TE-derived eccDNAs are generated through homologous recombination and nonhomologous end joining of linear reverse-transcribed DNA of LTR retrotransposons [8,12]. Based on this, using eccDNAs as a marker for TE mobility was

proposed. An early method of plant eccDNA sequencing, known as Mobilome-seq, was published by Lanciano et al. [6]. In brief, this approach employs exonuclease to remove linear DNA, followed by Phi29 polymerase random rolling-circle amplification of intact circle molecules and Illumina short-read sequencing. After read mapping the genome, further data processing is often focused on the detection of split-reads, including eccDNA borders. Wang et al. [11] used Mobilome-seq to conduct an extensive investigation of the eccDNA landscape in several *A. thaliana* tissues. The results of Lanciano et al. in [6] and Wang et al. [11] demonstrated the diversity and high organ/tissue specificity of eccDNAs. Applying the Mobilome-seq approach facilitated the detection of potential transposition activity for a number of TEs, including *EVD* and *ONSEN* in *A. thaliana*; *Tos17*, *PopRice*, *Houba* in rice; *Nightshade* in potato; and *Alex1* and *Alex3* in carrot [5,6,8,10,13].

Rolling-circle amplification (RCA) products have a concatemer structure with several tandemly organized monomers that correspond to full eccDNA molecules. Long-read sequencing of RCA products is particularly appealing for eccDNA de novo identification due to this property. This led to the development of the long-read eccDNA sequencing method known as CIDER-seq [14,15]. This technique has been used to profile eccDNAs and full-length viral DNA in a variety of species, including plants [14,15]. Size-based selection of extrachromosomal DNA (rather than exonuclease-assisted linear DNA removal) and sequencing of obtained RCA amplicons using PacBio SMRT are the core components of the CIDER-seq technique. PacBio sequencing of *A. thaliana* eccDNAs revealed a significant percentage of TE-originated eccDNAs, primarily *Helitron*, *DNA/MuDr* DNA transposons, and *Gypsy/Copia* LTR-retrotransposons. This is an additional compelling proof of TEs' enormous influence on plant circulome development. Unfortunately, the long-read sequencing approach has not been extensively used for eccDNA sequencing, leaving structural features of plant eccDNAs unexplored.

The goal of this study was to learn more about the structure and composition of eccDNAs found in *Arabidopsis thaliana* plants grown under TE-relaxed circumstances (epigenetic stress) in combination with various treatments (ABA, heat stress, and flagellin). Using long-read nanopore sequencing, we dissected several TE- and gene-derived eccDNAs and showed that a single TE can produce eccDNAs that vary in structure and composition. We proved that under stress, wild-type (Col-0) and *ddm1* (decrease in DNA methylation 1) mutant plants' eccDNA structures could differ dramatically. Finally, our work demonstrates that nanopore sequencing is a useful tool for the rapid identification of active TEs and illuminating circulome changes in plants.

2. Results

2.1. Extrachromosomal Circular DNA (eccDNA) Sequencing of *ddm1* Plants

Nanopore sequencing has rarely been used to analyze plant eccDNAs. To verify this method, we first sequenced the eccDNAs of the *ddm1* *Arabidopsis thaliana* mutant, which has a high level of transposon activity and a well-known circulome composition [6,10]. Exonuclease-mediated linear DNA removal and rolling-circle amplification were used to enrich eccDNAs in a DNA sample [6]. The RCA products were then digested using a T7 endonuclease to debranch them (Figure 1A). Sequencing of *ddm1* eccDNA via MinION Nanopore produced 257,563 reads. Col-0 (332,457 reads, 20× genome coverage) and *ddm1* (194,354 reads, 25× genome coverage) whole-genome Oxford Nanopore Technologies (ONT) sequencing (WGS) data were utilized as a baseline to identify eccDNA-generating loci. We employed ONT reads containing two or more tandemly organized monomer units (sequence of one circle of the eccDNA molecule) and lengths > 500 bp to filter out reads that did not correspond to eccDNA RCA products. After this screening, 45,277 (17.6%) ONT reads remained for additional examination. Notably, for Col-0 (32,138 reads, or 9.7%) and *ddm1* (21,279 reads, or 10.9%) WGS samples, the proportion of ONT reads with two or more monomers was considerably lower (the chi-square statistic using the Yates correction *p*-value was 0.00001). The TAIR10 genome was used to map the WGS and

eccDNA concatemer reads, and the number of mapped reads for 500 bp genome windows was determined.

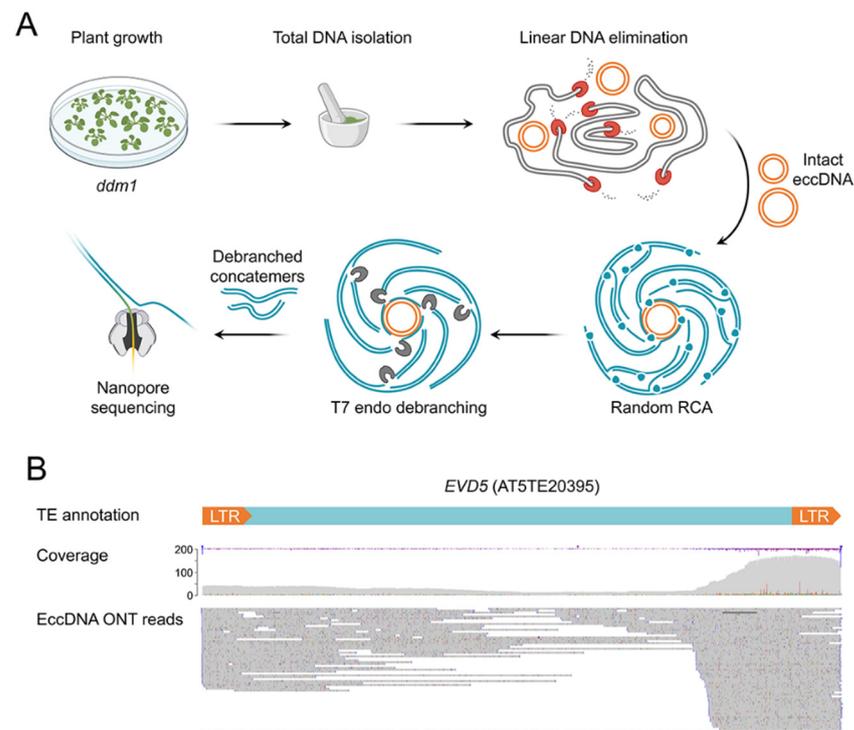


Figure 1. *A. thaliana ddm1* mutant eccDNA nanopore sequencing. **(A)** Schematic overview of the eccDNA nanopore sequencing experiment. **(B)** Coverage of EVD5 (AT5TE20395) by eccDNA *ddm1* reads.

For both WGS samples, peaks of high coverage associated with tandemly organized repeats were discovered, including centromeric repeats and rDNA loci (Supplementary Figure S1). These regions had significantly lower peak heights in the eccDNA *ddm1* sample, indicating that our protocol sufficiently removed linear DNA from the sample. We evaluated the \log_2 ratio between the read counts for the eccDNA *ddm1* sample relative to the WGS *ddm1* to identify loci with high coverage by eccDNA reads. Four eccDNA peaks with \log_2 ratios greater than three were found (Supplementary Figure S2); two of them, SRC2 (AT1G09070.1, Supplementary Figure S3) and MKK9 (AT1G73500, Supplementary Figure S4), were peaks on distal regions of chromosome 1.

Three copies of the LTR retrotransposons EVD1 (AT1TE41575), EVD2 (AT1TE41580), and EVD5 (AT5TE20395) from the ATCOPIA93 family made up the other two peaks (Figure 1B). In contrast to EVD1 and EVD2, which have not been observed to produce eccDNAs, EVD5 is a well-known characteristic of *ddm1* and epiRIL plants [6,16]. Additionally, EVD1 and EVD2 retrotransposons' transposition activity was not reported [16]. Indeed, when we manually verified the alignment of eccDNA readings for EVD1 and EVD2, we found many SNPs that unmistakably showed that the reads originated from EVD5. The concatemer reads of EVD5 were incorrectly mapped to EVD1 and EVD2 because they are tandemly organized and share many similarities with EVD5.

Thus, nanopore sequencing of eccDNA-enriched DNA samples from *A. thaliana* is an effective method for detecting eccDNA-produced loci such as TEs and genes. Our findings show that EVD5 is the most active eccDNAs-producing TE in the *ddm1* genetic background, which correlates with a prior discovery using short-read sequencing [6].

2.2. Epigenetic Stress Does Not Shift eccDNA Composition in *A. thaliana*

As TEs are the primary source of eccDNAs, we wondered if epigenetic stress could influence eccDNA formation. To test this, we cultured Col-0 plants in MS medium (ZA media) containing zebularine, a chemical demethylation agent, and α -amanitin, a PolII inhibitor, for 14 days. These toxins have the ability to reduce DNA methylation, hence allowing TE activation. Confirming the previous observations, Col-0 plants grown on toxin-supplemented media demonstrated a reduction in growth and root development (Supplementary Figure S5). We sequenced eccDNA samples from Col-0 plants growing on control (K, no toxins) and ZA media. Total ONT readings per duplicate ranged from 138,861 to 236,853. We examined the log₂ ratios of concatemer reads mapped to 500 bp windows between two biological replicates of ZA samples to determine a threshold. We discovered that 99.6% of the genomic windows had a $-\log_2$ ratio of two. Based on this, we established a log₂ threshold of three and a minimum length of 1000 bp for loci expressing eccDNAs. Additionally, only loci found in both biological replicates were chosen. We analyzed the number of reads per 100,000 reads in K and ZA variants using these criteria and determined the log₂ ratio of ZA to K ($\log_2(\text{ZA}/\text{K})$). Surprisingly, we found only one locus (Chr3: 13563500 . . . 13564500) that produces eccDNAs (Supplementary Figure S6). This region overlaps with the AT3TE55175 transposon. Although the data indicate that epigenetic stress alone is insufficient to increase eccDNA production, we anticipate that some loci may still produce a tiny amount of eccDNAs that are below the current detection limit.

2.3. Detection of eccDNAs under Abiotic, Hormone, and Flagellin Treatment of Epigenetically Stressed Plants

Previous investigations claimed that plants grown on ZA media and subjected to stress, such as heat stress, are stimulated to produce eccDNAs [10,17]. We combined epigenetic (ZA) stress with each of the following stimuli to detect stress-responsive eccDNA loci: heat stress (HS), flagellin treatment (Flg), and abscisic acid treatment (ABA). eccDNA sequencing of these three variations yielded 306,796, 399,064, and 681,418 reads, respectively, in two replicates. Concatemer reads with two or more monomer units were then chosen and aligned to the TAIR10 genome. We used Fisher's exact test with multiple corrections (Benjamini–Hochberg correction for multiple comparisons) to determine statistically significant peaks with increased read coverage in stressed samples compared with ZA samples. This test used only primary read alignments as input.

This analysis found no Flg- or ABA-responsive eccDNA peaks shared by two replicates (Supplementary Figures S6 and S7). This could imply that no eccDNA-produced genomic loci were activated by these conditions, or that eccDNA production was low and the ONT reads were insufficient to detect them. Following that, we examined ONT data for HS samples and discovered seven eccDNA-producing loci that had considerably higher coverage by concatemer reads in HS ZA samples than in ZA samples (Figure 2A). These peaks corresponded to seven copies of the ATCOPIA78 family's ONSEN LTR-retrotransposon: AT1TE59755 (ONSEN5), AT1TE71045 (ONSEN4), AT1TE12295 (ONSEN1), AT1TE24850 (ONSEN7), AT3TE89830 (ONSEN6), AT3TE92525 (ONSEN2), and AT5TE15240 (ONSEN3). ONSEN copies contributed differently to eccDNA contents; ONSEN3, ONSEN5, and ONSEN1 accounted for 92% of ONT reads, supporting previous results [18].

Thus, the data show that the eccDNA composition stabilized under ZA, ZA + ABA, and ZA + Flg treatments, whereas only heat stress significantly promoted eccDNA formation, but only from the ONSEN LTR-retrotransposon family.

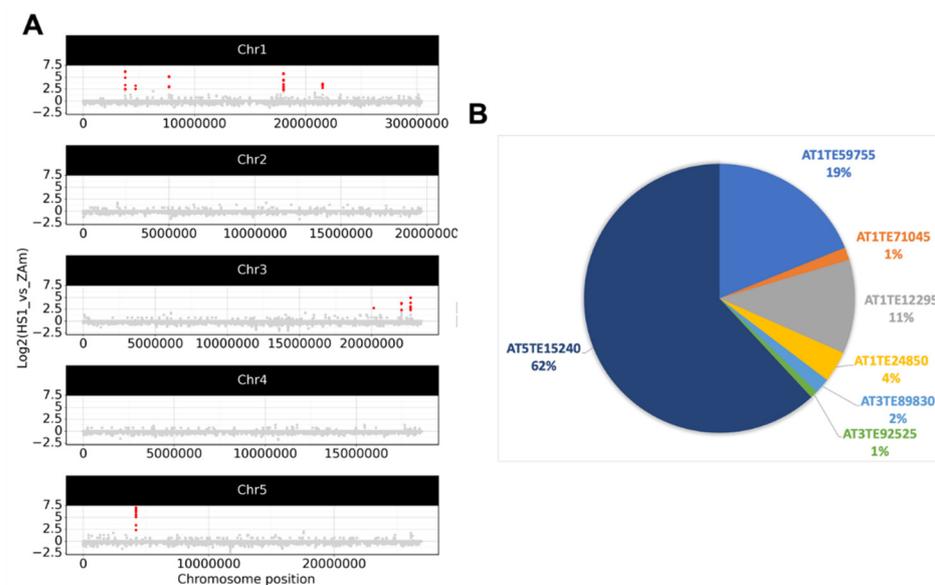


Figure 2. EccDNA loci that were responsive to heat stress. **(A)** Coverage of genomic loci by eccDNA reads in the HS sample. Dots represent the \log_2 ratio of the number of mapped concatemer reads of eccDNAs for the HS ZA vs. ZA samples (only primary alignments were utilized). In HS ZA samples compared with ZA samples, red dots indicate genomic areas with significantly increased read coverage (Fisher's exact test with Benjamini–Hochberg correction). **(B)** The ratio of eccDNA reads for seven ONSEN elements is presented as a pie chart.

2.4. ONT Reads Revealed Structure of eccDNA Molecules

We sought to shed light on the structure of eccDNA molecules generated from EVD and ONSEN TEs using long-read sequencing data. We created an original pipeline (eccStructONT) that made it possible to reconstruct eccDNA sequences. We performed ONT sequencing of the eccDNA molecules of *ddm1* plants grown in vitro and exposed to mild heat stress (30 °C) to estimate the structure of ONSEN and EVD eccDNAs in the same conditions. When combined with in vitro cultivation, mild heat stress could activate ONSEN while ensuring *ddm1* plant survival. To reduce noise, we only selected monomers that were repeated at least three times in the RCA concatemer read. We started by examining eccDNA molecules produced by EVD in *ddm1* mutants. For EVD (Supplementary Figure S9) and ONSEN elements (Figure 3A), the distribution of eccDNA lengths showed two peaks corresponding to full-length (fl_eccDNAs, near 5000 bp) and truncated (tr_eccDNAs, 1000 bp) eccDNAs. Next, we categorized various eccDNA molecules based on their region of origin in a TE frame. As seen in Figure 3B, the majority of EVD eccDNAs were fl_eccDNAs.

The visualisation of separate eccDNA structure groups for ONSEN components suggests that fl_eccDNAs and tr_eccDNAs were the two major eccDNA forms (Figure 3C). ONSEN tr_eccDNAs are primarily derived from one or two LTRs. Accordingly, EVD and ONSEN elements differed significantly in terms of their eccDNA set composition, with ONSEN having a significantly greater proportion of LTR-derived tr_eccDNAs than EVD did.

Next, we examined whether the ONSENS' fl_eccDNAs to tr_eccDNAs ratio remained the same in wild-type plants under ZA HS stress as in *ddm1* HS plants. Surprisingly, the structural study showed that all six ONSENS created primarily tr_eccDNAs, with very little fl_eccDNA production (Figure 4A). Again, LTR and LTR-adjusted internal ONSEN regions were the main sources of tr_eccDNAs (Figure 4A). Two primer pairs positioned on internal and LTR-adjusted parts of ONSENS were used in an RT-qPCR investigation to support these findings. The RT-qPCR results were corroborated using bioinformatic analysis (Figure 4B).

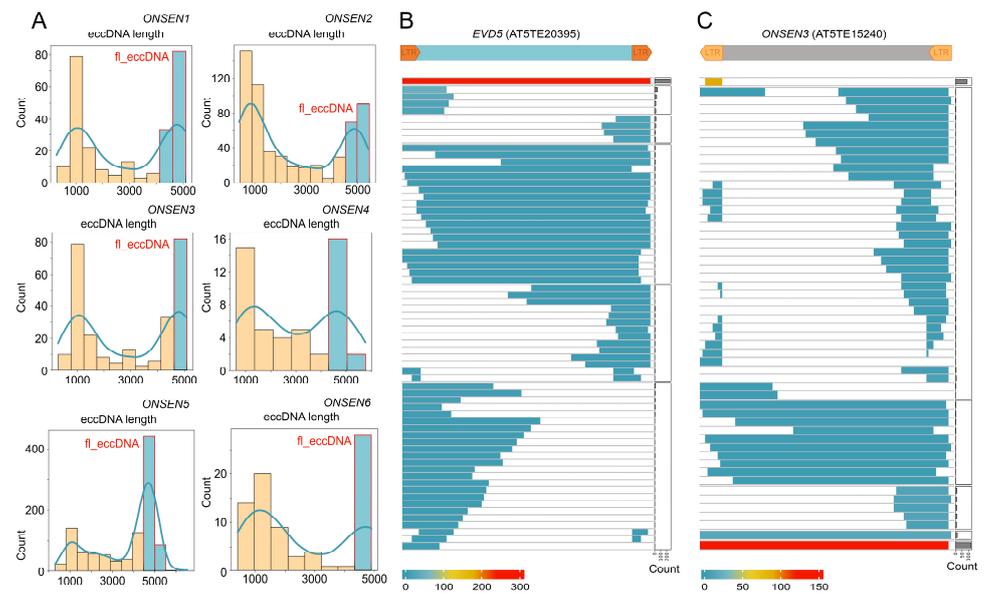


Figure 3. Structure of EVD and ONSEN eccDNAs generated in *ddm1* plants under mild heat stress conditions. **(A)** Histograms showing the number of ONSEN eccDNA molecules of different lengths. **(B,C)** Heatmaps showing structures and numbers of distinct eccDNAs from EVD and ONSEN3, respectively. The number of eccDNAs in individual groups is illustrated as a bar plot on the right side.

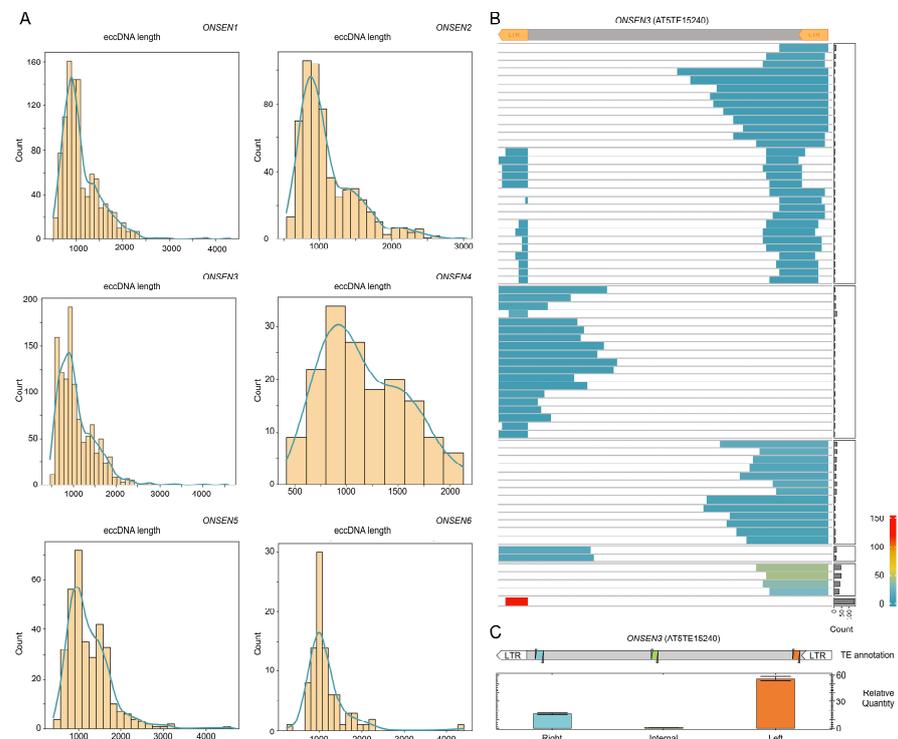


Figure 4. Structures of ONSEN3 eccDNAs detected in *Col-0* plants grown in ZA medium and subjected to heat stress conditions. **(A)** Histograms showing the number of ONSEN eccDNA molecules of different lengths. **(B)** Heat maps showing the structures and number of distinct eccDNAs from ONSEN3. The number of eccDNAs in individual groups is illustrated as a bar plot on the right side. **(C)** RT-qPCR results with eccDNA-enriched DNA of ZA HS *Col-0* and three primer pairs aligned to the left, right, and central parts of the TE.

The acquired data indicated that the eccDNA composition of a single TE can significantly differ between genetic backgrounds. As a result, it would be intriguing to investigate the genetic and environmental factors that influence eccDNA biogenesis.

3. Discussion

Genetic and environmental factors involved in eccDNA biogenesis in plants are not yet understood. However, it is clear that eccDNA composition and expression are dynamic processes. Previously, the results of a comparative analysis of eccDNAs in different *Arabidopsis* organs demonstrated that the eccDNA repertoire is influenced by unique routes and mechanisms involved in eccDNA production in different cell types [11]. Here, we did not find up-regulated eccDNAs under ZA Flg-stressed and ZA ABA-stressed conditions. It is possible that these stresses negatively regulated the expression of some TEs, as well as genes involved in the TE life cycle. The inhibition of TE-derived eccDNA production by stress was demonstrated in *Solanum tuberosum* and *S. commersonii* species [13]. Transposon *Nightshade* actively generates eccDNAs in control conditions, whereas cold stress inhibits eccDNA production [13]. This could explain why we did not detect any eccDNA peaks after applying stresses. Notably, the *Arabidopsis* genome has tens of TE families possessing transpositionally active elements [19,20]. However, the transposition of TEs from just a few families (e.g., *ATCOPIA93* (*EVD*), *ATCOPIA78* (*ONSENs*), *VANDAL*, etc.) has been observed in laboratory conditions. This implies that we still have very limited knowledge about conditions that can induce TE expression and transposition activity. New tools and techniques to trace TE transposition in different cell types and stress conditions are required.

Long-read sequencing followed by concatemer identification in raw ONT reads facilitated deciphering sequences of full-length eccDNAs, paving the way for the elucidation of eccDNA structure. Taking this into account, we investigated the structure of *EVD* and six *ONSEN* elements in the same genetic background (*ddm1*) and stress conditions (mild heat stress of in vitro growing plants). Our results indicated the presence of eccDNAs corresponding to full-length TE sequences (fl_eccDNAs) and certain TEs parts, including LTRs and LTR-adjusted regions (tr_eccDNAs). The ratios between fl_eccDNAs and tr_eccDNAs differed between *EVD* and *ONSEN* elements in *ddm1* under HS. In addition, fl_eccDNAs of *ONSEN* elements were almost not detectable in Col-0 plants under ZA HS stress; only tr_eccDNAs of different groups were identified. This further implies that the eccDNA repertoire may vary dramatically between individual TEs, stress conditions, and genetic backgrounds. The presence of different types of TE-derived tr_eccDNAs in a cell raised questions about the molecular mechanisms and possible functions of these molecules. A direct answer is that tr_eccDNAs are byproducts of intra- and intermolecular recombination events leading to fl_eccDNA conversion into tr_eccDNAs [21–23]. This scenario may have evolved as a defense mechanism against new TE integration. Most eccDNAs contain full-length LTR sequences that may function as promoters that trigger transcription from eccDNAs. It is well known that circular DNAs in the form of plasmids carrying target genes under promoter (e.g. 35S) are well transcribed, being transfected into the protoplast cells [24]. Moreover, RNA transcribed from eccDNAs was detected in humans [25] and plants [26]. Based on this, it could be proposed that RNA molecules transcribed from TE-derived eccDNAs may become substrates for RdDM-pathway-triggering TE silencing.

4. Materials and Methods

4.1. Plant Material and In Vitro Growth Conditions

Seeds of *ddm1* (*ddm1-2*, F7 generation) were kindly provided by Vincent Colot (Institut de Biologie de l'École Normale Supérieure (IBENS), Paris, France). *Arabidopsis* seeds were surface-sterilized with 75% ethanol (2 min) and washed with 5% sodium hypochlorite (5 min). After this, the seeds were rinsed with sterile distilled water 3 times. A total of 0.1% agarose was added to each tube and the seeds were resuspended and dripped on Petri dishes with Murashige and Skoog medium, supplemented with 3% of sucrose (PanReac

AppliChem, Darmstad, Germany) and 1% of agar (PanReac AppliChem, Darmstad, Germany). Plates with seeds were sealed with Parafilm (Pechiney Plastic Packaging Company, Chicago, IL, USA) and kept in the dark at 4 °C for 3 days for vernalization and synchronous germination. Afterward, dishes were transferred into a light chamber with 16 h day/8 h night photoperiods for further growth.

4.2. Stress Conditions

Surface-sterilized Col-0 seeds were resuspended in 0.1% agarose and transferred to solid 1/2 MS medium supplemented with sterile filtered 4.6 ug/mL α -amanitin (Sigma-Aldrich, CAS 23109-05-9) and 9 ug/mL zebularine (Sigma-Aldrich; CAS 3690-10-6). For flagellin treatment, 2-week-old plants grown on 1/2 MS media supplemented with α -amanitin and zebularine were transferred to the equal ZA medium but also supplemented with 0.5 ug/mL flagellin (Flagellin from *Salmonella typhimurium*, Sigma-Aldrich, SRP8029-10UG). For ABA treatment, 2-week-old plants grown on 1/2 MS media supplemented with α -amanitin and zebularine were transferred to the equal ZA medium but also supplemented with abscisic acid at final concentrations of 100 uM. For heat-stress treatment, 2-week-old plants grown on 1/2 MS media supplemented with α -amanitin and zebularine were subjected to elevated temperature (4 °C for 24 h followed by 37 °C for 24 h, HS). For mild heat stress treatment 2 weeks-old Col-0 *ddm1* plants were subjected to 30 °C for 24 h.

4.3. DNA Isolation

Total DNA was extracted by CTAB protocol [27] from plants immediately after stress (0 h recovery time). Briefly, 50 mg of young in vitro—grown *A. thaliana* plant tissues were ground with a pestle in a mortar with liquid nitrogen. A total of 0.5 mL of preheated to 75 °C CTAB1 buffer containing 6% of β -mercaptoethanol and 0.5% polyvinylpyrrolidone were immediately added to the frozen powder. A lysate was transferred to a 1.5 mL tube and incubated at 75 °C for 1 h. After cooling, an equal volume of chloroform was added to the sample. An upper water phase was transferred to a 1.5 mL tube containing 2 volumes of CTAB2 buffer. An obtained pellet was resuspended in 0.2 mL 1M NaCl, and DNA was precipitated with equal isopropanol volume followed by centrifugation. Pellet was washed with 70% ethanol and resuspended in nuclease-free water. An RNase treatment was performed followed by isopropanol reprecipitation and ethanol washing. The obtained DNA was resuspended in nuclease-free water again for downstream analysis.

4.4. Extrachromosomal Circular DNA (eccDNA) Enrichment

For eccDNA enrichment, the protocol of Lanciano et al. [6] was used with some modifications. Briefly, linear DNA elimination was performed with PlasmidSafe DNase (LGC Biosearch Technologies, UK). For this, 750 ng of total DNA was combined with 2 μ L ATP 25 mM, 5 μ L 10 \times PlasmidSafe buffer, and 1 μ L PlasmidSafe DNase (10 units) in 50 μ L reaction volume. Incubation mode was 72 h at 37 °C with the addition of an extra amount of reagents (2 μ L 25 mM ATP, 0.3 μ L PlasmidSafe buffer, 1 μ L PlasmidSafe DNase (10 units) after the first 24 h. Enzyme inactivation was performed at 70 °C for 30 min. The remaining DNA was precipitated by overnight incubation with 1/10V 3M sodium acetate (pH 5.2) and 2.5V absolute ethanol followed by centrifugation at 12,000 \times g for 30 min. The DNA pellet was subjected to random RCA reaction using illustra TempliPhi 100 Amplification Kit (GE Healthcare, catalog #25-6400-10). The reaction was carried out by adding 5 μ L of TempliPhi sample buffer (the buffer was preliminary preheated at 95 °C for 3 min and then cooled to RT) to the DNA pellet. After adding 5 μ L of the premix solution (5 μ L TempliPhi Reaction buffer, 0.2 TempliPhi Enzyme mix) to the sample, the reaction was incubated at 28 °C for 72 h. After enzyme inactivation at 65 °C for 10 min, the sample volume was adjusted to 50 μ L and purified with AmpureXP at 0.5 \times ratios according to the manufacturer's instructions. A purified RCA product was debranched to remove the hyperbranched structures generated. The following reaction mix was prepared for debranching: 500 ng of RCA product, 5 μ L of 10 \times reaction buffer, and 1 μ L of T7 endonuclease I (New England

Biolabs, M0302S) in 50 µL reaction volume. After the incubation at 37 °C for 15 min, the reaction was stopped and purified by adding equal chloroform volume. Debranched RCA product was precipitated by adding 1/10V 3M sodium acetate (pH 5.2) and 2.5V absolute ethanol followed by incubation at −80 °C for 30 min and centrifugation at 12,000× *g* for 30 min. The obtained pellet was dissolved in nuclease-free water and used for nanopore sequencing.

4.5. RT-qPCR Analysis of *fl_eccDNA* and *tr_eccDNAs* of *ONSEN*

To estimate the *ONSEN* eccDNA enrichment profile after RCA, the following primers were designed for the three distinct regions of *ONSEN* sequences: left part—cONS_Lq (TGAAGATCCTAAAGATGGCGAG; TGCTCCTAGGATAGCCTTCA); internal part—cONS_Iq (CTCATGCTCATGTACCGGATG; AGCTTGTAGCCTTTGGAGTTG); right part—cONS_Rq (AAGTCGGCAATAGCTTTGGC; CATACTCCAATTGCACGTCCT). The real-time PCR reaction mix was prepared in 25 µL with the use of 12.5 µL BioMaster HS-qPCR SYBR Blue (2×) (Biolabmix, Russia), 1 µL of 10 pmol of each primer, 9.5 µL of nuclease-free water (Biolabmix, Russia), and 1 µL of RCA product. PCR amplification was run on CFX96 (BioRad, USA) with the following program: 95 °C for 2 min, then 95 °C for 15 s, 58 °C for 15 s, and 40 cycles.

4.6. Nanopore Library Preparation and Sequencing

Library preparation was carried out with 1 µg of eccDNA or genomic DNA using the Native Barcoding Expansion 1–12 (Oxford Nanopore Technologies (Oxford, UK), catalog no. EXP-NBD104) and the Ligation Sequencing Kit SQK-LSK109 (Oxford Nanopore Technologies). Sequencing was performed by MinION equipped with an R9.4.1 flow cell.

4.7. Bioinformatic Analysis of eccDNA Sequencing and Data Visualization

The TAIR10 genome was downloaded from the NCBI database. For the analysis of eccDNA, concatemer reads were selected by running TideHunter [28] with the following additional parameters: ‘-f 2 -c 2’ for eccDNA peak determination and ‘-f 2 -c 3’ for eccDNA structure analysis. The selected reads or monomers (for structural analysis) were aligned to the reference genomes TAIR10 using minimap2 software [29] with the following parameters: -ax map-ont -t 100. The obtained sam file was converted to bam format, sorted, and indexed using SAMtools [30]. Then only primary alignments with a minimum MQ value of 30 were left for further analysis using the following command: ‘samtools view -F 3840 -q 30’. The sorted bam files were used for alignment inspection using locally installed JBrowse2 [31]. To detect eccDNA peaks, the genome was split into 500 bp windows using the bedtools [32] command ‘bedtools makewindows -w 500’. The number of reads at each window was determined by ‘bedtools intersect’. To find peaks, the number of reads in each window was compared between control (ZA) and stress variants. The statistical analysis (significance level = 0.01) with two-sided Fisher’s exact test was performed using the Python `scipy.stats.fisher_exact` function [33] followed by Benjamini–Hochberg correction for multiple comparisons carried out by the function `statsmodels.stats.multitest.multipletests` (with the argument “method = ‘fdr_bh’”). The obtained results were visualized using the plotnine Python module (<https://plotnine.readthedocs.io/>, last accessed on 22 February 2023). The code used for this analysis (Jupyter Notebook `eccDraw_HS1_vs_mergedZA.ipynb`) can be found in the GitHub repository: <https://github.com/Kirovez/eccStructONT> (last accessed on 27 April 2023).

For the structural analysis of eccDNAs, the Python package ‘eccDNA_struct.py’ (<https://github.com/Kirovez/eccStructONT>) was written to obtain the matrix that was used for the heatmap visualization performed by ComplexHeatmap R package [34]. This script also generated eccDNA length distribution histograms using seaborn (<https://seaborn.pydata.org/>, last accessed on 21 April 2023) and pandas (<https://pandas.pydata.org/>, last accessed on 9 February 2023) modules.

5. Conclusions

In this study, we showed that nanopore sequencing is a useful tool in the detection and structural analysis of eccDNA molecules in plants. Using nanopore sequencing of eccDNA molecules of epigenetically stressed *Arabidopsis* plants grown under various stress treatments (heat, abscisic acid, and flagellin), we showed that TE-derived eccDNA quantities and structures significantly vary between individual TEs. Epigenetic stress alone does not cause eccDNA up-regulation; its combination with heat stress triggers a generation of full-length and various truncated eccDNAs of *ONSEN* elements. We showed that the ratio between full-length and truncated eccDNAs is TE- and condition-dependent. Our work paves the way for further elucidation of eccDNA structural features and their connections with various biological processes, such as eccDNA transcription and eccDNA-mediated TE silencing.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants12112178/s1>, Figure S1: Mapping of selected concatemer reads with 2 or more monomers was obtained from nanopore sequencing eccDNA-enriched *ddm1* DNA as well as whole-genome sequencing of Col-0 and *ddm1* plants; Figure S2: EccDNA read coverage of genome loci; Figure S3: Coverage of AT1G09070 gene (Chr1:2,927,502 . . . 2,929,107) by nanopore concatemer reads from eccDNA-enriched DNA of *ddm1*; Figure S4: Coverage of AT1G73500 gene (Chr1: 27,637,748 . . . 27,641,755) by nanopore concatemer reads from eccDNA-enriched DNA of *ddm1*; Figure S5: Inhibition of *A.thaliana* wild-type growth and development under toxin treatment; Figure S6: EccDNA read coverage of genomic loci; Figure S7: EccDNA read coverage of genomic loci from Flg samples (replicate 1); Figure S8: EccDNA read coverage of genomic loci from Flg samples (replicate 2); Figure S9.: The histogram showing a number of EVD eccDNAs of different lengths.

Author Contributions: Conceptualization, I.K.; methodology, P.M. and I.K.; software, I.K.; validation, P.M., E.E. and I.K.; formal analysis, P.M.; investigation, P.M. and E.E.; resources, IK; writing—original draft preparation, P.M. and I.K.; writing—review and editing, I.K.; visualization, P.M. and I.K.; supervision, I.K; funding acquisition, I.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation (grant No 22-74-10055).

Data Availability Statement: The nanopore data produced for this study are available in the Sequence Read Archive (SRA), NCBI, under Bioproject Accession PRJNA975338.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hotta, Y.; Bassel, A. Molecular Size And Circularity Of DNA In Cells Of Mammals And Higher Plants*. *Proc. Natl. Acad. Sci. USA* **1965**, *53*, 356–362. [[CrossRef](#)] [[PubMed](#)]
2. Billheimer, F.E.; Avers, C.J. Nuclear And Mitochondrial DNA From Wild-Type And Petite Yeast: Circularity, Length, And Buoyant Density*. *Proc. Natl. Acad. Sci. USA* **1969**, *64*, 739–746. [[CrossRef](#)] [[PubMed](#)]
3. Clark-Walker, G.D. Size Distribution of Circular DNA from Petite-Mutant Yeast Lacking ρ DNA. *Eur. J. Biochem.* **1973**, *32*, 263–267. [[CrossRef](#)] [[PubMed](#)]
4. Peng, H.; Mirouze, M.; Bucher, E. Extrachromosomal Circular DNA: A Neglected Nucleic Acid Molecule in Plants. *Curr. Opin. Plant Biol.* **2022**, *69*, 102263. [[CrossRef](#)] [[PubMed](#)]
5. Zhang, P.; Peng, H.; Llauro, C.; Bucher, E.; Mirouze, M. Ecc_finder: A Robust and Accurate Tool for Detecting Extrachromosomal Circular DNA From Sequencing Data. *Front. Plant Sci.* **2021**, *12*, 743742. [[CrossRef](#)]
6. Lanciano, S.; Carpentier, M.-C.; Llauro, C.; Jobet, E.; Robakowska-Hyzorek, D.; Lasserre, E.; Ghesquière, A.; Panaud, O.; Mirouze, M. Sequencing the Extrachromosomal Circular Mobilome Reveals Retrotransposon Activity in Plants. *PLoS Genet.* **2017**, *13*, e1006630. [[CrossRef](#)]
7. Vanderschuren, H.; Zaidi, S.S.-A.; Shakir, S.; Mehta, D.; Nguyen, V.; Gutzat, R. Profiling of EccDNAs in Arabidopsis Indicates Shift in Population of TE-Derived EccDNAs in Response to Stress, Cellular State, and Epigenetic Processes. 2022. Available online: https://assets.researchsquare.com/files/rs-2256837/v3_covered.pdf?c=1680543096 (accessed on 1 March 2023).
8. Kwolek, K.; Kędzierska, P.; Hankiewicz, M.; Mirouze, M.; Panaud, O.; Grzebelus, D.; Macko-Podgórn, A. Diverse and Mobile: EccDNA-based Identification of Carrot Low-copy-number LTR Retrotransposons Active in Callus Cultures. *Plant J.* **2022**, *110*, 1811–1828. [[CrossRef](#)]

9. Mann, L.; Seibt, K.M.; Weber, B.; Heitkam, T. ECCsplorer: A Pipeline to Detect Extrachromosomal Circular DNA (EccDNA) from next-Generation Sequencing Data. *BMC Bioinform.* **2022**, *23*, 40. [[CrossRef](#)]
10. Thieme, M.; Lanciano, S.; Balzergue, S.; Daccord, N.; Mirouze, M.; Bucher, E. Inhibition of RNA Polymerase II Allows Controlled Mobilisation of Retrotransposons for Plant Breeding. *Genome Biol.* **2017**, *18*, 134. [[CrossRef](#)]
11. Wang, K.; Tian, H.; Wang, L.; Wang, L.; Tan, Y.; Zhang, Z.; Sun, K.; Yin, M.; Wei, Q.; Guo, B.; et al. Deciphering Extrachromosomal Circular DNA in Arabidopsis. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1176–1183. [[CrossRef](#)]
12. Hirochika, H.; Otsuki, H. Extrachromosomal Circular Forms of the Tobacco Retrotransposon Ttol. *Gene* **1995**, *165*, 229–232. [[CrossRef](#)] [[PubMed](#)]
13. Esposito, S.; Barteri, F.; Casacuberta, J.; Mirouze, M.; Carputo, D.; Aversano, R. LTR-TEs Abundance, Timing and Mobility in *Solanum commersonii* and *S. tuberosum* Genomes Following Cold-Stress Conditions. *Planta* **2019**, *250*, 1781–1787. [[CrossRef](#)]
14. Mehta, D.; Hirsch-Hoffmann, M.; Were, M.; Patrignani, A.; Zaidi, S.S.-A.; Were, H.; Gruissem, W.; Vanderschuren, H. A New Full-Length Circular DNA Sequencing Method for Viral-Sized Genomes Reveals That RNAi Transgenic Plants Provoke a Shift in *Geminivirus* Populations in the Field. *Nucleic Acids Res.* **2018**, *47*, e9. [[CrossRef](#)] [[PubMed](#)]
15. Mehta, D.; Cornet, L.; Hirsch-Hoffmann, M.; Zaidi, S.S.-A.; Vanderschuren, H. Full-Length Sequencing of Circular DNA Viruses and Extrachromosomal Circular DNA Using CIDER-Seq. *Nat. Protoc.* **2020**, *15*, 1673–1689. [[CrossRef](#)] [[PubMed](#)]
16. Kirov, I.; Merkulov, P.; Dudnikov, M.; Polkhovskaya, E.; Komakhin, R.A.; Konstantinov, Z.; Gvaramiya, S.; Ermolaev, A.; Kudryavtseva, N.; Gilyok, M.; et al. Transposons Hidden in *Arabidopsis thaliana* Genome Assembly Gaps and Mobilization of Non-Autonomous LTR Retrotransposons Unravelling by Nanotei Pipeline. *Plants* **2021**, *10*, 2681. [[CrossRef](#)] [[PubMed](#)]
17. Kirov, I.; Merkulov, P.; Polkhovskaya, E.; Konstantinov, Z.; Kazancev, M.; Saenko, K.; Polkhovskiy, A.; Dudnikov, M.; Garibyan, T.; Demurin, Y.; et al. Epigenetic Stress and Long-Read cDNA Sequencing of Sunflower (*Helianthus annuus* L.) Revealed the Origin of the Plant Retrotranscriptome. *Plants* **2022**, *11*, 3579. [[CrossRef](#)] [[PubMed](#)]
18. Cavrak, V.V.; Lettner, N.; Jamge, S.; Kosarewicz, A.; Bayer, L.M.; Scheid, O.M. How a Retrotransposon Exploits the Plant's Heat Stress Response for Its Activation. *PLoS Genet.* **2014**, *10*, e1004115. [[CrossRef](#)]
19. Baduel, P.; Leduque, B.; Ignace, A.; Gy, I.; Gil, J.; Loudet, O.; Colot, V.; Quadrana, L. Genetic and Environmental Modulation of Transposition Shapes the Evolutionary Potential of *Arabidopsis thaliana*. *Genome Biol.* **2021**, *22*, 138. [[CrossRef](#)]
20. Quadrana, L.; Silveira, A.B.; Mayhew, G.F.; LeBlanc, C.; Martienssen, R.A.; Jeddeloh, J.A.; Colot, V. The *Arabidopsis thaliana* Mobilome and Its Impact at the Species Level. *eLife* **2016**, *5*, e15716. [[CrossRef](#)]
21. Arrey, G.; Keating, S.T.; Regenberg, B. A Unifying Model for Extrachromosomal Circular DNA Load in Eukaryotic Cells. *Semin. Cell. Dev. Biol.* **2022**, *128*, 40–50. [[CrossRef](#)]
22. Møller, H.D.; Larsen, C.E.; Parsons, L.; Hansen, A.J.; Regenberg, B.; Mourier, T. Formation of Extrachromosomal Circular DNA from Long Terminal Repeats of Retrotransposons in *Saccharomyces cerevisiae*. *G3 Genes Genomes Genet.* **2016**, *6*, 453–462. [[CrossRef](#)]
23. Møller, H.D.; Parsons, L.; Jørgensen, T.S.; Botstein, D.; Regenberg, B. Extrachromosomal Circular DNA Is Common in Yeast. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3114–E3122. [[CrossRef](#)] [[PubMed](#)]
24. Deshayes, A.; Herrera-Estrella, L.; Caboche, M. Liposome-mediated Transformation of Tobacco Mesophyll Protoplasts by an *Escherichia coli* Plasmid. *EMBO J.* **1985**, *4*, 2731–2737. [[CrossRef](#)] [[PubMed](#)]
25. Ling, X.; Han, Y.; Meng, J.; Zhong, B.; Chen, J.; Zhang, H.; Qin, J.; Pang, J.; Liu, L. Small Extrachromosomal Circular DNA (EccDNA): Major Functions in Evolution and Cancer. *Mol. Cancer* **2021**, *20*, 113. [[CrossRef](#)] [[PubMed](#)]
26. Molin, W.T.; Yaguchi, A.; Blenner, M.; Sasaki, C.A. The EccDNA Replicon: A Heritable, Extranuclear Vehicle That Enables Gene Amplification and Glyphosate Resistance in *Amaranthus palmeri*. *Plant Cell* **2020**, *32*, 2132–2140. [[CrossRef](#)]
27. Pucker, B.; Kleinbölting, N.; Weisshaar, B. Large Scale Genomic Rearrangements in Selected *Arabidopsis thaliana* T-DNA Lines Are Caused by T-DNA Insertion Mutagenesis. *BMC Genom.* **2021**, *22*, 599. [[CrossRef](#)]
28. Gao, Y.; Liu, B.; Wang, Y.; Xing, Y. TideHunter: Efficient and Sensitive Tandem Repeat Detection from Noisy Long-Reads Using Seed-and-Chain. *Bioinformatics* **2019**, *35*, i200–i207. [[CrossRef](#)]
29. Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [[CrossRef](#)]
30. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. Subgroup, 1000 Genome Project Data Processing The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
31. Diesh, C.; Stevens, G.J.; Xie, P.; Martinez, T.D.J.; Hershberg, E.A.; Leung, A.; Guo, E.; Dider, S.; Zhang, J.; Bridge, C.; et al. JBrowse 2: A Modular Genome Browser with Views of Synteny and Structural Variation. *Genome Biol.* **2023**, *24*, 74. [[CrossRef](#)]
32. Quinlan, A.R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr. Protoc. Bioinform.* **2014**, *47*, 11.12.1–11.12.34. [[CrossRef](#)] [[PubMed](#)]
33. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)] [[PubMed](#)]
34. Gu, Z. Complex Heatmap Visualization. *Imeta* **2022**, *1*, e43. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.