



## Article

# Sequence Motif Analysis of PRDM9 and Short Inverted Repeats Suggests Their Contribution to Human Microdeletion and Microduplication Syndromes

Paris Ladias <sup>1,†</sup> , Georgios S. Markopoulos <sup>2,†</sup> , Charilaos Kostoulas <sup>1</sup> , Ioanna Bouba <sup>1</sup>, Agis Georgiou <sup>1</sup>, Sofia Markoula <sup>3</sup> and Ioannis Georgiou <sup>1,\*</sup>

<sup>1</sup> Laboratory of Medical Genetics in Clinical Practice, Faculty of Medicine, School of Health Sciences, University of Ioannina, 45 110 Ioannina, Greece

<sup>2</sup> Neurosurgical Institute, Faculty of Medicine, School of Health Sciences, University of Ioannina, 45 110 Ioannina, Greece

<sup>3</sup> Department of Neurology, Faculty of Medicine, School of Health Sciences, University of Ioannina, 45 110 Ioannina, Greece

\* Correspondence: igeorgio@uoi.gr; Tel.: +30-2651007822

† These authors contributed equally to this work.

**Abstract:** Holliday junctions are the first recognized templates of legitimate recombination. Their prime physiological role is meiotic homologous recombination, resulting in rearrangements of the genetic material. In humans, recombination hotspots follow a distinct epigenetic pattern designated by the presence of PR domain-containing protein 9 (PRDM9). Repetitive DNA elements can replicate in the genome and can pair with short inverted repeats (SIRs) that form Holliday junctions in a significantly high frequency in vitro. Remarkably, PRDM9 and SIR sequence motifs, which may have the potential to act as recombination primers associated with transposable elements (TEs) and their presence, may lead to gradual spreading of recombination events in human genomes. Microdeletion and microduplication syndromes (MMSs) constitute a significant entity of genetic abnormalities, almost equal in frequency to aneuploidies. Based on our custom database, which includes all MMSs shorter than 5 Mbs in length which is the cut-off point for the standard cytogenetic resolution, we found that the majority of MMSs were present in sequences shorter than 0.5 Mbs. A high probability of TE-associated and non-TE-associated PRDM9/SIR sequence motifs was found in short and long MMSs. Significantly, following the Reactome pathway analysis, a number of affected genes have been associated with the pathophysiological pathways linked to MMSs. In conclusion, PRDM9 or SIR sequence motifs in regions spanning MMSs hotspots underlie a potential functional mechanism for MMS occurrences during recombination.

**Keywords:** short inverted repeats (SIRs); retroelements; transposable element (TE); recombination; PRDM9; Holliday junction; DNA repair; microduplications; microdeletions; developmental disorders; sequence motifs



**Citation:** Ladias, P.; Markopoulos, G.S.; Kostoulas, C.; Bouba, I.; Georgiou, A.; Markoula, S.; Georgiou, I. Sequence Motif Analysis of PRDM9 and Short Inverted Repeats Suggests Their Contribution to Human Microdeletion and Microduplication Syndromes. *BioMedInformatics* **2023**, *3*, 267–279. <https://doi.org/10.3390/biomedinformatics3020018>

Academic Editors: Jörn Lötsch and Alexandre G. De Brevin

Received: 6 November 2022

Revised: 11 March 2023

Accepted: 22 March 2023

Published: 1 April 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

Over the past 10 years, the ability to analyze the sequence variations in entire genomes with high-throughput technologies (NGS, aCGH) in research and a clinical diagnostic framework has led to the identification of pathogenic genomic rearrangements causing genetic abnormalities with a wide range of syndromic or isolated phenotypic variations [1]. These variations are primarily related to ultrasound findings in embryos, congenital anomalies in neonates, or mental retardation and growth retardation in infants and children, commonly known as microdeletion and microduplication syndromes (MMSs) [2], and result in changes in the number of copies of one or more DNA segments (copy number variations, CNVs) leading to a gain or loss of gene content [2]. CNVs are essentially structural

genomic markers, which vary in size from 1 kb to a few megabases, and are characterized either as copy number polymorphisms (CNPs) or as rare polymorphisms responsible for many genetic syndromes or conditions [3]. CNVs can be pathogenic if the gain or loss involves one or more genes and present with a direct phenotype to genotype association, while other CNVs are benign variants within populations and no direct association between the sequence variation and a pathogenic phenotype can be established [4,5].

In order to determine whether a CNV is pathogenic, certain criteria can be applied. In general, the main criteria are heritability, size, and type (duplication or deletion) of the rearrangement, as well as the content of the underlying gene [6]. Large CNVs are more likely to cause genetic diseases than smaller CNVs, because of their large gene content (dose-dependent effect in transcription and translation) [7,8]. Chromosomal region losses can also lead to a deficiency; the consequence of which is known for involving specific genes and generating particular phenotypes [7]. On the other hand, duplications are more difficult to characterize and interpret because they often present with an unclear clinical phenotype that in most cases may alter gene dosage. Finally, gene content should be taken into consideration since CNVs that contain many genes or genes associated with known disorders are much more pathogenic than CNVs containing less genes or genes of unknown clinical significance or impact on the phenotype [9,10].

In most cases, CNVs are known to occur because of active non-allelic recombination-based mechanisms. There are two major categories of CNVs: (1) recurrent, and (2) non-recurrent. Recurrent CNVs are the result of non-allelic homologous recombination events during meiosis, with clustered breakpoints within defined regions, whereas non-recurrent CNVs arise mainly by non-homologous end joining, and fork stalling and template switching (FOSTES). In contrast to recurrent CNVs, non-recurrent CNVs do not share the same breakpoint regions [11–13].

It is important to note that the chromosomal rearrangements described above are a result of the genomic architecture of the human genome, which has been acquired over hundreds of thousands of years of evolution and countless meiotic recombinations. Many repeats (high copy or low copy) work as substrates for inter- or intrachromosomal recombination in meiosis and mitosis [14]. Large low copy repeats (LCRs) promote NAHR events with a frequency directly correlated with LCRs' homology length and sequence similarity. Previous and present studies aim to extend the evidence for the existence of a significantly high number of PRDM9 and short inverted repeats (SIR) sequence motifs related to genomic rearrangements found in almost all types of meiotic and mitotic recombinations. There is evidence that non-B DNA formed by SIRs plays a role in genetic instability, disease etiology, and evolution. Previous notable findings have explained the co-localization between SIRs and human genomic instability breakpoints [15]. Furthermore, the presence of PRDM9/SIR motifs is also associated with specific retroelements in proximity with genes, with transcription start sites (TSSs), and between genes, which can demonstrate the importance of controlled, mediated, or primed by specific mobile elements in recombination during the evolution of the genome [16]. The importance of PRDM9-associated meiotic DSBs has been thoroughly examined by Pratto et al. by generating high-quality recombination maps that has offered novel insights into the roles of specific PRDM9 alleles in meiotic recombination [17]. Subsequent studies revealed the importance of PRDM9 as a key molecule in the recognition of meiotic recombination hotspots in mammals in both physiological and pathological conditions, while it is dispensable in the megabase regulation of the phenomenon [18–21].

Deficiencies in the coordination between mobile repeat elements and PRDM9/SIR motifs or the inability of the cells to maintain their potent cell cycle control properties may lead to unbalanced translocations and other somatic genomic defects such as mosaicism and cancer clonal defects [16]. While repeat elements have continuously and intensively been involved in mammalian and human evolution in order to create stability, diversity, and plasticity, they may also have the potential to destabilize the genome under a variety

of adverse conditions and result in clinically significant defects and disorders under certain circumstances [22].

This study examines associations of the common MMSs with PRDM9/SIR consensus sequences using statistical algorithms and tools, and presents evidence that the above-mentioned sequences are not randomly distributed throughout the genome and are present in most MMSs, which may indicate a functional mechanism for MMSs generation and occurrence with potential diagnostic and prognostic significance.

## 2. Materials and Methods

### 2.1. Bioinformatic Mining Tools for the Identification of PRDM9/SIRs Residing within the Boundaries of a Syndrome

To precisely identify the coordinates of MMS sequences in human genomes, we performed data mining for all microdeletion and microduplication syndromes [1,6]. In order to document all microdeletion and microduplication syndromes that have been reported in public databases and scientific publications, we developed two BED files. The overall syndrome hits were converted from HG19 to HG38 using the lift over tool in the UCSC browser. The algorithmic steps for the detection of sequence motifs within the boundaries of syndromes are listed below:

1. We used R 3.5.1 (<http://www.r-project.org/>) to perform the permutation analysis and the overlap analysis between Holliday junction hits and syndromes. Based on the permutation tests, we evaluated the statistical associations between our loci sets using the Regioner package, which was developed specifically for this purpose [21]. In order to extend the functionality of the permutation test framework, the user can provide custom functions. There are several predefined randomization and evaluation functions included in this framework, which are designed specifically for working with genomic regions. The exact algorithm and scripts are listed in the supplementary materials and methods files (Supplemental Dataset S1). With the Regioner package, we evaluated the associations between our region gene sequence sets based on the permutation tests. Here is a description of the methodology used to detect sequence variations of SIRs and PPRDM9 motifs that are statistically correlated with syndromes.

2. Data collection was implemented by utilizing a script written in C# that identified all different forms of PRDM9 sequence motifs (CCNCCNTNNCCNC) as well as all different forms of SIR motifs (CCnnnNNNGG where nnnNNN is an inverted repeat pattern). A total of 64 different SIR sequence variations and 1024 different PRDM9 sequence variations were identified.

3. R's matchPattern function was applied to all chromosomes of the genome. As a result, it produced a text file with all hits containing PRDM9 or SIR. Secondly, R's regioneR library was used to perform the overlapPermTest function. Next, R's regioneR library was used to perform R's local ZScore function. In the next step, overlap PermTest plots and local ZScore plots were saved as JPG images. Following that, the overlap Region's function in R was performed using R's regioneR library. In the last step, the output of the overlap Region's function was stored in a text file. The complete R code regarding the identification of PRDM9/SIR motifs within the boundaries of MMSs can be found here: [http://genescreening.gr/wp-content/uploads/PRDM9\\_code.r](http://genescreening.gr/wp-content/uploads/PRDM9_code.r). A *p*-value < 0.05 and a Z-score < −3 or >3 were considered significant.

Additionally, the UCSC genome browser provides bioinformatic tools (<http://genome.ucsc.edu/> last accessed on 14 October 2022) [23] to manually screen the human genome (version hg38) for PRDM9/SIR sequence motifs residing within the boundaries of syndromes and their surrounding sequences in each of the 24 human chromosomes, combined with ENCODE Project Transcription Factor Binding Site Data [24], as previously described [25]. By using Repeat Masker analysis (<http://www.repeatmasker.org> last accessed on 14 October 2022) [26], we defined whether and where there are repeated elements upstream and downstream of PRDM9/SIRs, based on a described framework [27]. The default settings were used as the gold standard in our analyses.

## 2.2. Genomic Annotation and Analysis of Genomic Distribution

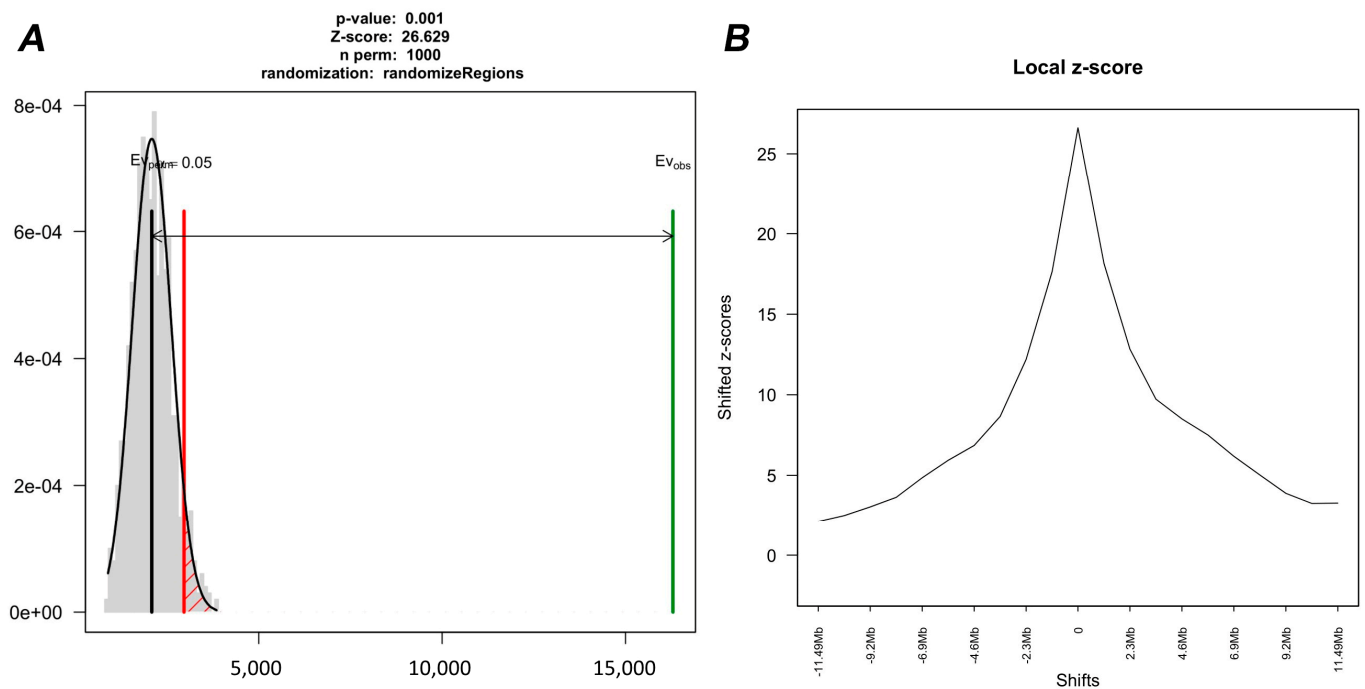
Genomic annotation and statistical enrichment analysis for genomic regions were performed using the PAVIS Annotation [28]. Gene-specific analysis was performed with Genomic Regions Enrichment of Annotations Tool (GREAT) ([great.stanford.edu](http://great.stanford.edu)). GREAT was used to analyze the relative distribution between MMS loci and transcription start sites (TSS) of human genes [29], using the default association rule settings for statistical analysis and the Human Genome Version hg38, following established protocols [30]. The MMS-associated genes were downloaded from the GREAT server and visualized in the UCSC genome browser. Reactome analysis ([www.reactome.org](http://www.reactome.org) last accessed on 14 October 2022) [31] was used to identify enriched pathways containing the gene set following the GREAT analysis, as described elsewhere [32].

## 3. Results

### 3.1. MMS Coincide with PRDM9/SIRs Motifs

We initially performed data mining for all microdeletion and microduplication syndromes [1,23] and we found that most syndromes (40% of overall syndromes present) are shorter than 0.5 Mb (Figure S1). As a further step in analyzing the relationship between syndromes and sequence data, we created two databases in BED format, one with 213 records associated with microdeletion syndromes across the human genome and one with 82 records associated with microduplication syndromes. The names and coordinates of the breakpoints are listed in each database as separate files (Supplemental Dataset S1). Correspondingly, we created databases also in BED format for all SIRs that form Holliday junctions according to the degenerated sequence motif (CCnnnNNNGG) across the human genome from our previous published data, and one more relating to the PRDM9-related sequences from the degenerated sequence motif (CCNCCNTNNCCNC) [16,33,34]. All databases were created based on the hg38 version of the human genome. There is no known association between PRDM9/SIR motifs, key elements of heterologous recombination, and any syndromes. A possible mechanism of MMS occurrence may be revealed by their presence within the breakpoints of a syndrome. For this reason, R statistical querying was used to define the overlap and statistical correlation between our databases. In particular, we demonstrated, using the Regioner package [35], which variations of the above-mentioned sequence motifs are statistically significant and not just random. The motifs were identified between the boundaries of syndromes. We did our search for gene identification in the regions between the breaks that promote the emergence of the syndromes. CpG and GC content are filtered by the Regioner package by programming in R language (<https://bioconductor.org/packages/devel/bioc/vignettes/regioner/inst/doc/regioner.html>).

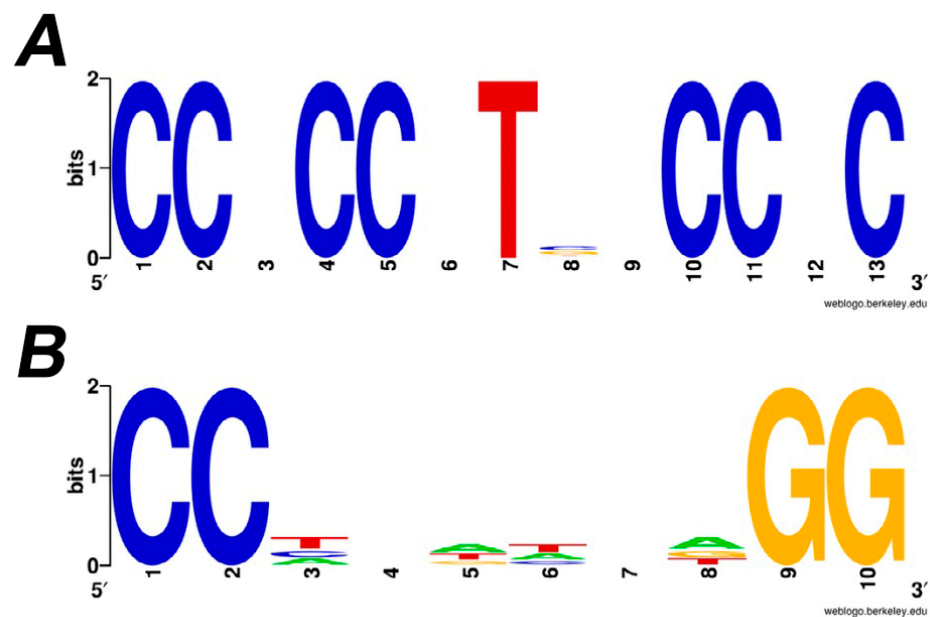
A number of 32 unique sequence variations of the total 64 SIR variations that form Holliday junctions and a result of 11,606 statistically significant genome hits were found present in 165 microdeletion and microduplication syndromes together with 363 PRDM9 sequence variations of the total 1024 resulting in 14,840 statistically significant genome hits which appeared more frequently in the syndromes with very low  $p$ -values (in all cases  $p < 0.05$ ). The results are listed in Supplementary Table S1. A plot of  $p$ -values/ $z$ -scores corresponding to PRDM9 hits located in microdeletion and microduplication syndromes is provided in Figure 1. Supplementary Dataset S1 also includes plots of  $p$ -values and  $z$ -scores corresponding to the most statistically associated SIRs hits. By displacing the PRDM9/SIRs outside the sequence boundaries of each individual syndrome, the  $z$ -scores drop, demonstrating that the association is dependent upon the exact position of the regions rather than being a non-specific regional effect. Detailed information regarding the number of SIR and PRDM9 motifs residing within the boundaries of a syndrome is provided in Supplementary Table S2.



**Figure 1.** Example plots extracted from R indicating the statistical significance of SIRs positions residing within the boundaries of microdeletion and microduplication syndromes based on permutational analysis. The resultant graph represents the results of the permutation test, based on  $p$ -value (**A**) and  $z$ -score (**B**). Subfigure (**A**) depicts a gray histogram representing the evaluation of the randomized region set with a fitted normal and a black bar representing the mean of the randomized evaluations. Accordingly, the green line represents the observed hits, while the red line represents the expected hits, and shows the limit at which the correlation can be statistically or by chance associated with MMSs. The  $p$ -value is equal to 0.05 when the green line coincides with the red line (observed hits = expected hits). In our analysis, the green line is statistically significant when it is on the right side of the red line. In addition, a red bar (and red shading) represents the significance limit (by default 0.05). Thus, if the green bar is in the red-shaded region, it means that the original evaluation is extremely unlikely and so the  $p$ -value will be significant. The X-axis represents the number of the overlaps between the region sets, while the Y-axis represents the density of probability. Subfigure (**B**) depicts plot of the local  $z$ -score of the permutation test in subfigure (**A**). The association is strongly related to the exact position of the SIRs peaks since the  $z$ -score drops sharply as soon as the regions are shifted a few hundreds of bases.

Different sequence variations show a different frequency of occurrence in MMSs. That was demonstrated by an association analysis of genomic regions based on permutation tests in R statistical environment. More specifically, the sequences with the degenerated motifs have been found to have a direct association with 165 MMS syndromes and indeed with a very low  $p$ -value ( $p < 0.001$ ). The consensus motifs of all significant PRDM9 and SIR motifs are depicted in Figure 2. In conclusion, we found a significant correlation between the boundaries of MMS with specific motifs of PRDM9/SIRs that may mediate recombination.



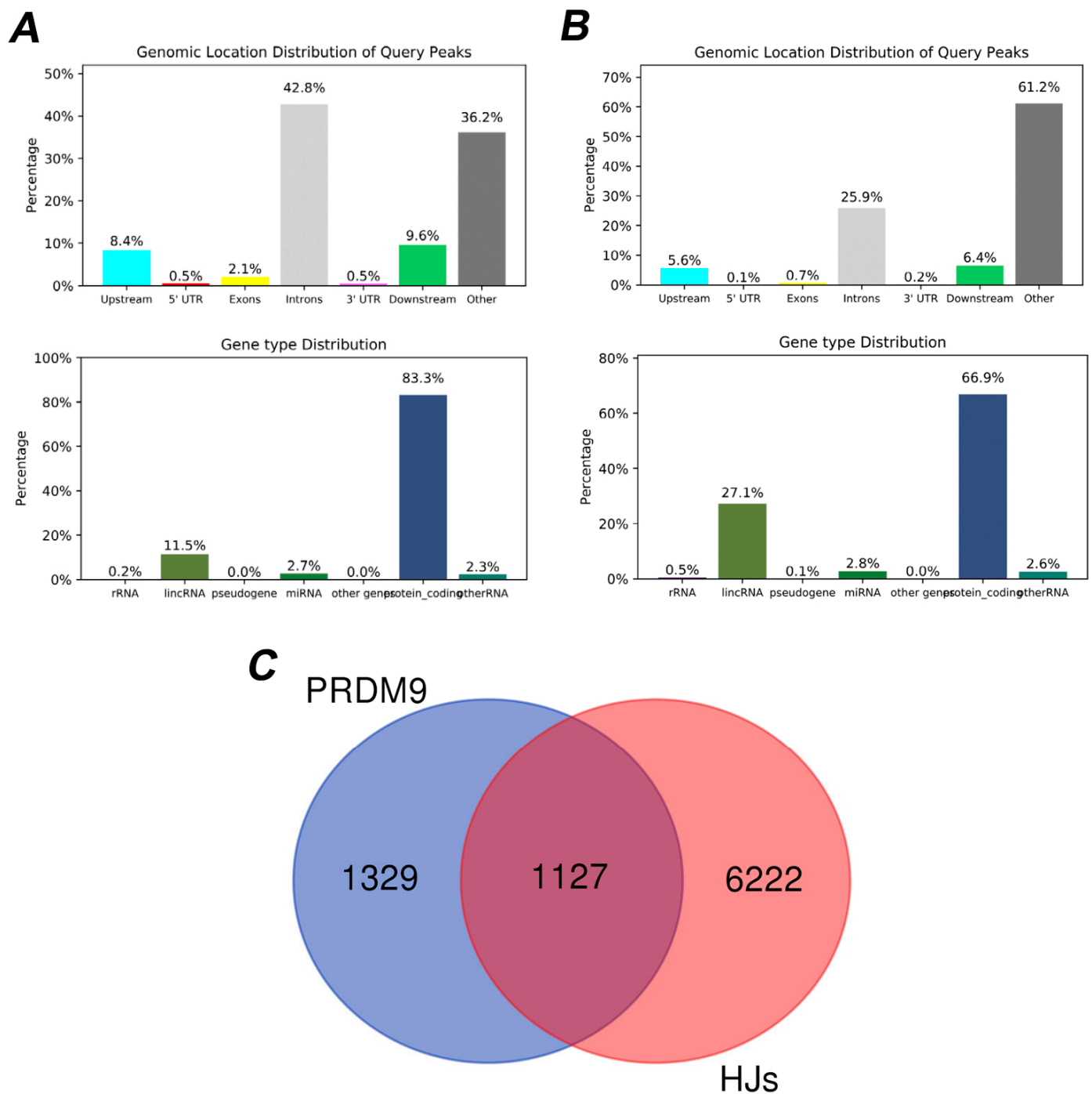


**Figure 2.** PRDM9 and SIR consensus sequences. The consensus motifs of all significant PRDM9 and SIR motifs are depicted in plot (A) (PRDM9) and plot (B) (SIRs).

### 3.2. Genomic Features of PRDM9/SIR Motifs in MMSs

Transposable elements (TEs) serve as substrates for illegitimate intra- or interchromosomal/interchromatid recombination during meiosis as well as in mitosis, as part of the repetitive genome. To this end, we analyzed PRDM9/SIR motifs in MMSs for the presence of TEs, based on the Repeat Masker analysis. In Supplementary Table S3, we present the distribution and percentage of TE- and non-TE-associated motifs residing in MMSs. The distribution of TE and non-TE PRDM9/SIR motifs in MMSs is specific, meaning that they have a particular imprint on individual retroelements. Our Repeat Masker analysis showed that the three major families of retroelements (ALUs, LINEs, HERVs) are co-existing with PRDM9/SIR motifs within the sequence boundaries of MMSs. More significantly, PRDM9 motifs appear to have evolved along with ALUs and LINEs. In particular, 6181 from the overall 14,840 hits identified as members of the ALUs repeat family (41.6%) while 1998 from the overall 14,840 hits identified as members of the LINEs repeat family. On the other hand, SIRs appeared to have evolved along with HERVs; especially 8342 hits from the overall 11,606 (71.9%) identified as members of the endogenous retroviruses (HERVs). The exact members of each repeat family are shown in Table S3. From our comparative analysis, we found that PRDM9 statistically significant sequence motifs reside in ALUs in a percentage of 41.6% when ALUs' overall percentage in human genomes is nearly 11%. On the other hand, SIR statistically significant sequence motifs reside in a percentage of nearly 72% in HERVs when HERVs' overall percentage in human genomes equals to 8%.

To gain further insight into the genomic distribution of MMS regions, we performed an enrichment analysis using the PAVIS Annotation Report from the [www.manticore.niehs.nih](http://www.manticore.niehs.nih) server [28]. This tool classifies each region according to whether it resides either in intergenic regions, promoters, exons, or introns, or it is partially contained within more than one category. We found that 42.8% of the PRDM9 hits regions reside in introns, 36.2% in intergenic regions, 9.6% downstream of gene regions, 8.4% in promoters, and 2.4% in exons. On the other hand, the genomic location distribution of SIRs was 61.2% in intergenic regions, 25.9% in introns, 6.4% downstream of gene regions, and 5.6% in promoters (Figure 3A,B). In general, MMS regions are more enriched and more likely to emerge in gene-associated regions. The emergence of extensive genomic data has stressed and established the idea that most genomic repetitive sequences and motifs may play a regulatory or functional role.



**Figure 3.** (A) Genomic distribution of PRDM9 sequence motifs. (**Upper Panel**) Genomic location distribution in relation to human genes, following PAVIS analysis. The relative abundance of sites upstream in 5' UTR, exons, introns, 3' UTR, downstream, and in intergenic regions are shown in distinct color bars. (**Lower Panel**) Gene type distribution. The relative gene type distribution is shown in distinct color bars. (B) Genomic distribution of SIRs sequence motifs. (**Upper Panel**) Genomic location distribution in relation to human genes, following PAVIS analysis. Various color bars indicate the relative abundance of sites in upstream, 5' UTR, exons, introns, 3' UTR, downstream, and intergenic regions. (**Lower Panel**) Gene type distribution. Different colored bars indicate the relative distribution of gene types. (C) Venn diagram for common PRDM9/SIR motifs that have a significant proximity to human genes.

Next, we performed a gene annotation analysis to find whether PRDM9/SIR motifs have a significant proximity to human genes, following the GREAT analysis [29]. We found that the loci of all our 26,446 significant genomic regions included in the analysis are associated with common 1127 genes in total, as shown in the Venn diagram presented in Figure 3C. The resultant genes are presented in Supplementary Table S4.

Subsequent gene ontology (GO) and pathway enrichment analysis revealed that a significant high number of these genes are taking part in many metabolic pathways and processes depicted in Figure S2. Among others, genes in GO biological processes with a significant high number of associated genes are involved in cellular metabolic processes (633 genes), regulation of biological processes (387 genes), metabolic processes (371 genes), response to stimulus (188 genes), localization (169 genes), signaling (157 genes), and multicellular organismal processes (111 genes). Collectively, PRDM9/SIR motifs reside adjacent to a large number of genes that are associated with significant processes in organismal physiology.

In the final step, in order to gain further insight into the regulatory footprint of genes with PRDM9/SIR motifs, we performed a pathway analysis on the Reactome server [31]. The enriched pathways are summarized in Table 1. Among the most enriched pathways, three contain interactions that include MECP2 protein, five interactions are associated with RUNX1 and RUNX3 (in mediated transcription regulation), and two are involved with FGFR2.

**Table 1.** Top 12 enriched pathways of genes with PRDM9/SIR motifs, based on Reactome analysis.

Pathway Name	Found	Total	<i>p</i> Value	FDR	Found	Total	Ratio
MECP2 regulates transcription of neuronal ligands	6	13	$8.73 \times 10^{-4}$	$7.76 \times 10^{-1}$	8	8	0.001
RUNX1 regulates transcription of genes involved in BCR signaling	4	7	$2.99 \times 10^{-3}$	$7.76 \times 10^{-1}$	4	4	0
TP53 regulates transcription of death receptors and ligands	6	18	$4.36 \times 10^{-3}$	$7.76 \times 10^{-1}$	4	7	0
RUNX3 regulates RUNX1-mediated transcription	3	4	$4.76 \times 10^{-3}$	$7.76 \times 10^{-1}$	2	2	0
RUNX1 regulates estrogen receptor-mediated transcription	4	10	$1.03 \times 10^{-2}$	$7.76 \times 10^{-1}$	7	8	0.001
Loss of MECP2 binding ability to 5hmC-DNA	2	2	$1.24 \times 10^{-2}$	$7.76 \times 10^{-1}$	1	1	0
Adrenoceptors	4	11	$1.41 \times 10^{-2}$	$7.76 \times 10^{-1}$	11	12	0.001
RUNX1 regulates transcription of genes involved in differentiation of keratinocytes	4	11	$1.41 \times 10^{-2}$	$7.76 \times 10^{-1}$	7	8	0.001
Class I peroxisomal membrane protein import	5	20	$2.71 \times 10^{-2}$	$7.76 \times 10^{-1}$	5	6	0
TRAIL signaling	3	8	$2.99 \times 10^{-2}$	$7.76 \times 10^{-1}$	6	6	0
Neurexins and neuroligins	10	60	$3.09 \times 10^{-2}$	$7.76 \times 10^{-1}$	16	19	0.001
L1CAM interactions	17	130	$4.81 \times 10^{-2}$	$7.76 \times 10^{-1}$	15	54	0.004

## 4. Discussion

### 4.1. What Does Statistical Significance Mean for the Biology of MMSs

This study examined whether PRDM9 and SIR motifs are statistically associated with MMSs in a non-random manner. Among the main findings of our study is the existence of a significant number of unique sequence motifs for PRDM9 and SIR in regions of specific chromosomes and loci where MMSs occur. In particular, 363 PRDM9 and 32 SIR variations are associated with a very low *p*-value and a very high *z*-score among MMSs. In conjunction with our previous report [16], our results suggest that mobile genetic elements may disperse recombination-sensitive sequences throughout the genome, suggesting that the observed relationship between PRDM9/SIR/MMS does not result by chance. To our knowledge, this is the first study to directly link the two phenomena.



The breakpoint analysis showed the presence of two of the major families of mobile elements (ALUs and HERVs) in the recombined loci, indicating that recombination between these PRDM9/SIRs and rich mobile elements may lead to syndromic genomic aberrations. Human mobile genetic elements provide a template for genetic and epigenetic innovation that can significantly impact physiology or pathology [36,37], as well as activation during early embryogenesis [38]; a notion that has been verified in other mammals such as the mouse [25,27]. Previous studies have shown that mobile elements are present in chromosomal rearrangements and early development [39,40] identified in routine cytogenetics; therefore, mobile elements are expected to play a role in short syndromes, between 0.5 and 5 Mb of genomic disorders [41], but also in clinically important rearrangements generating microdeletions and microduplications.

To further understand the importance of genomic loci containing joint MMSs with PRDM9 and SIR motifs, two types of analyses were performed. First, we conducted an enrichment analysis with PAVIS; a tool that performs general annotation and provides enrichment statistics for genomic regions. We found that MMS/PRDM9/SIR regions are highly enriched in gene-abundant genomic loci. More importantly, looking into the enrichment data we may also gain insight into possible mechanisms of MMSs occurrences. The enrichment in intergenic regions may probably mean that deletions in such regions may disturb gene regulatory mechanisms, while the enrichments in introns, respectively, may mean that in such cases MMSs are derived from conserved gene sequence, thus leading to developmental disruption through recombination of developmental and metabolic clusters.

Second, we performed an MMS-to-gene-specific enrichment using the GREAT tool and subsequently in gene ontology and PANTHER databases, leading to a major finding of the current study resulting from gene proximity findings. This analysis revealed that all our MMSs dataset regions are near or in proximity to genes and more importantly, near, adjacent, or within genes that control developmental and metabolic processes in Table S4). This finding comes to agreement with the fact that microdeletion and microduplication syndromes are interconnected with phenotypes involving developmental disorders. We believe that critical PRDM9-SIR-MMS motifs in developmental and metabolic gene regions may function as a template for recombination and may lead to unequal crossovers and structural variations within the affected breakpoints. Our claim is supported by previous studies that show that PRDM9-mediated recombination sites also influence genome evolution and the incidence of genetic diseases [17,19]. As regards metabolic gene regions, our finding is novel and may represent an additional link between PRDM9/SIRs and recombination-mediated pathology. The resulting microdeletions or microduplications may affect gene functionality, either by directly disrupting gene sequences or by affecting neighboring gene regulatory regions (locus control regions) and thus leading to developmental disorders. A novel finding that is extracted from this study is that MMSs-associated PRDM9 and SIRs are responsible for the generation of sequences near or within developmental and metabolic genes or gene clusters, and may control or participate in a potential mechanism that explains the underlying phenotype seen in each of these syndromes. More importantly, they may help to identify the boundaries containing disease-causing genes whose sequence or adjacent regulatory regions are altered. An interesting finding is the enrichment of pathways containing interactions of RUNX1, RUNX3, MECP2, and FGFR2. RUNX1 is an established epigenetic regulator in physiology and pathology with known roles in cancer development [42]. Importantly, microdeletions in RUNX1 locus have a clinical impact in myeloid malignancies [43]. On the other hand, MECP2 is a dynamic regulator of heterochromatin formation with significant roles in neurodevelopmental diseases [44]. A series of cases of MECP2 microduplications link the pathological expression of this gene with Rett syndrome [45–47]. FGFR2 genomic aberrations are a link between epigenetic dysregulation and cholangiocarcinoma [48]. Collectively, the pathway analysis reveals several genes associated with MMS as the most significantly affected among those with PRDM9/SIR motifs. Further studies in this direction may shed light on a unified regulatory network that links PRDM9/SIR motif-containing genes with the pathophysiology of MMS.

Taking together both genomic- and gene-specific enrichment analyses revealed that MMSs may occur through disruption of regulatory or coding sequences by recombination. The affected genes in each case are disease-causing and may explain, directly or indirectly, the underlying phenotype.

#### 4.2. PRDM9/SIR *In Silico* Motifs and Independent Experimental Data

The current work is a systematic computational approach that combines the SIR motif identification from a previous study in our lab [16] with a number of computed sequence variants of PRDM9 motifs (1024 in total) in order to explore the nature and the possible implications of their co-occurrence. Our results are based on the effort to explain the non-random co-occurrence of these motifs in MM-associated DNA breaks, based on the structures derived in human genetic recombination sites. There has been significant progress made in experimental approaches to study recombination hotspots. A seminal study by Pratto et al. has linked PRDM9-associated meiotic DSBs with high-quality recombination maps and has offered novel insights into the roles of specific PRDM9 alleles in meiotic recombination [17]. Significantly, PRDM9 alleles regulate hotspot strength, contribute to the mutagenic effect of recombination, and play roles in genomic instability that is seen in MMSs [17]. The phenomenon has also been found to be associated with the epigenetic state of chromatin [49] and to be sex-specific [50]. Subsequent studies revealed on the one hand that PRDM9 is essential in meiotic recombination hotspot recognition, while on the other hand has been found not essential in the megabase-resolution regulation of the phenomenon [18–21].

In the present study, an independent approach offers an alternative view on possible additional genomic locations that take part in meiotic recombination, based on *in vitro* and *in silico* data. On the one hand, the lack of experimental verification of our computational model is a limitation of our study. However, we believe that the significance of our results is supported by the fact that SIR regions have been shown to take part in four-way junctions as intermediates of recombination [51], the impact of SIR cruciform structures in interaction with proteins in eukaryotes, prokaryotes, and viruses [52], and the impact of SIRs as hotspots for genetic instability in cancer [15]. The validity of our results remains to be verified by subsequent studies.

Thus, there is a need for future meta-analysis of present data, and high-throughput experimental data from the aforementioned studies would systematically verify the existence of concurrent data. Another important parameter that should be taken into consideration is the impact of copy number variations (CNVs) between individuals in the abundance of PRDM9-SIR motifs and how these would impact the recombination frequency. The lack of such dataset is an additional limitation of the current study. We believe that such studies might add to the resolution of genome-wide recombination maps.

#### 4.3. Conclusions and Future Perspectives

Bioinformatic analysis of NGS data permits the extraction of multiple results concerning either sequence substitutions or sequence rearrangements [53]. With the advent of high-throughput technology in genetic diagnosis and through the experience gained in prenatal genetic applications, we now know, in clinical practice, that specific recurrent MMSs are not rare as a group of genetic disorders and may have frequencies higher than aneuploidies [54]. Di-George recurrent microdeletion syndrome is more prevalent than aneuploidies 13 and 18, and is considered second only to Down syndrome [55]. A wide spectrum of MMSs exists, and even rare non-recurrences can have a high impact on prenatal investigation and diagnosis. Additionally, their phenotypic spectrum also demonstrates the clinical variation and significance of MMSs, whether they are detected by prenatal ultrasound studies, in neonates, or other genetic studies, suggesting a need for a more robust approach in molecular diagnosis. It has been found that a percentage of 30% or higher of MMS cases can be underdiagnosed by ultrasound, even when performed by an experienced fetomaternal specialist [6].

It would be helpful to understand the evolution and molecular biology of MMSs in order to develop and apply diagnostic approaches based on bioinformatic analyses of NGS data. This would address the issue of CNVs involved in the generation of MMSs as a defined clinical entity using strict criteria. In addition to chromosomal aneuploidies and microdeletions, WGS-NGS platforms can also monitor gene point mutations [56]. As a result, such technologies might be useful for cancer research as well as in future clinical practices. The reason for this is that both rare disorders and common disorders are likely to be caused by similar molecular processes that underlie the generation of multiple aberrations in embryos and somatic cells.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biomedinformatics3020018/s1>. Figure S1: Frequency distribution graphs containing all microdeletion and microduplication syndromes.; Figure S2: Gene Ontology (GO) pathway enrichment analysis; Table S1: MMS coordinates; Table S2: Total SIR/PRDM9 coordinates; Table S3: SIR/PRDM9 genomic Features; Table S4: SIR/PRDM9 associated Genes; Supplementary Dataset: Algorithm used for analysis and total results.

**Author Contributions:** Conceptualization, P.L. and I.G.; methodology, P.L. and G.S.M.; validation, I.B. and C.K.; investigation, A.G.; data curation, P.L. and G.S.M.; writing—original draft preparation, P.L. and G.S.M.; writing—review and editing, P.L., G.S.M., I.G., I.B. and S.M.; visualization, P.L., C.K. and G.S.M.; supervision, I.G.; All authors have read and agreed to the published version of the manuscript.

**Funding:** The project is co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation (EPAnEK), NSRF 2014–2020 (Project code MIS: 5047228).

**Data Availability Statement:** All data are available upon request.

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

## References

1. Vissers, L.E.; Stankiewicz, P. Microdeletion and microduplication syndromes. *Methods Mol. Biol.* **2012**, *838*, 29–75. [CrossRef]
2. Zhang, F.; Gu, W.; Hurles, M.E.; Lupski, J.R. Copy Number Variation in Human Health, Disease, and Evolution. *Annu. Rev. Genom. Hum. Genet.* **2009**, *10*, 451–481. [CrossRef]
3. Almal, S.H.; Padh, H. Implications of gene copy-number variation in health and diseases. *J. Hum. Genet.* **2012**, *57*, 6–13. [CrossRef]
4. Nowakowska, B. Clinical interpretation of copy number variants in the human genome. *J. Appl. Genet.* **2017**, *58*, 449–457. [CrossRef] [PubMed]
5. Redon, R.; Ishikawa, S.; Fitch, K.R.; Feuk, L.; Perry, G.H.; Andrews, T.D.; Fiegler, H.; Shapero, M.H.; Carson, A.R.; Chen, W.; et al. Global variation in copy number in the human genome. *Nature* **2006**, *444*, 444–454. [CrossRef]
6. Watson, C.T.; Marques-Bonet, T.; Sharp, A.J.; Mefford, H.C. The Genetics of Microdeletion and Microduplication Syndromes: An Update. *Annu. Rev. Genom. Hum. Genet.* **2014**, *15*, 215–244. [CrossRef]
7. Rice, A.M.; McLysaght, A. Dosage-sensitive genes in evolution and disease. *BMC Biol.* **2017**, *15*, 78. [CrossRef]
8. Henriksen, C.N.; Chaignat, E.; Reymond, A. Copy number variants, diseases and gene expression. *Hum. Mol. Genet.* **2009**, *18*, R1–R8. [CrossRef]
9. Mihaylova, M.; Staneva, R.; Toncheva, D.; Pancheva, M.; Hadjidekova, S. Benign, pathogenic and copy number variations of unknown clinical significance in patients with congenital malformations and developmental delay. *Balk. J. Med. Genet.* **2017**, *20*, 5. [CrossRef]
10. Zhang, L.; Wang, J.; Zhang, C.; Li, D.; Carvalho, C.M.B.; Ji, H.; Xiao, J.; Wu, Y.; Zhou, W.; Wang, H.; et al. Efficient CNV breakpoint analysis reveals unexpected structural complexity and correlation of dosage-sensitive genes with clinical severity in genomic disorders. *Hum. Mol. Genet.* **2017**, *26*, 1927–1941. [CrossRef]
11. Carvalho, C.M.B.; Pehlivan, D.; Ramocki, M.B.; Fang, P.; Alleva, B.; Franco, L.M.; Belmont, J.W.; Hastings, P.J.; Lupski, J.R. Replicative mechanisms for CNV formation are error prone. *Nat. Genet.* **2013**, *45*, 1319. [CrossRef]
12. Lindsay, S.J.; Khajavi, M.; Lupski, J.R.; Hurles, M.E. A chromosomal rearrangement hotspot can be identified from population genetic variation and is coincident with a hotspot for allelic recombination. *Am. J. Hum. Genet.* **2006**, *79*, 890–902. [CrossRef]
13. Hastings, P.J.; Lupski, J.R.; Rosenberg, S.M.; Ira, G. Mechanisms of change in gene copy number. *Nat. Rev. Genet.* **2009**, *10*, 551. [CrossRef]
14. Sen, S.K.; Han, K.; Wang, J.; Lee, J.; Wang, H.; Callinan, P.A.; Dyer, M.; Cordaux, R.; Liang, P.; Batzer, M.A. Human genomic deletions mediated by recombination between Alu elements. *Am. J. Hum. Genet.* **2006**, *79*, 41–53. [CrossRef] [PubMed]

15. Lu, S.; Wang, G.; Bacolla, A.; Zhao, J.; Spitser, S.; Vasquez, K.M. Short Inverted Repeats Are Hotspots for Genetic Instability: Relevance to Cancer Genomes. *Cell Rep.* **2015**, *10*, 1674–1680. [[CrossRef](#)]
16. Ladias, P.; Markopoulos, G.; Lazaros, L.; Markoula, S.; Tzavaras, T.; Georgiou, I. Holliday Junctions Are Associated with Transposable Element Sequences in the Human Genome. *J. Mol. Biol.* **2016**, *428*, 658–667. [[CrossRef](#)]
17. Pratto, F.; Brick, K.; Khil, P.; Smagulova, F.; Petukhova, G.V.; Camerini-Otero, R.D. Recombination initiation maps of individual human genomes. *Science* **2014**, *346*, 1256442. [[CrossRef](#)]
18. Davies, B.; Hatton, E.; Altemose, N.; Hussin, J.G.; Pratto, F.; Zhang, G.; Hinch, A.G.; Moralli, D.; Biggs, D.; Diaz, R. Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature* **2016**, *530*, 171–176. [[CrossRef](#)]
19. Pratto, F.; Brick, K.; Cheng, G.; Lam, K.-W.G.; Cloutier, J.M.; Dahiya, D.; Wellard, S.R.; Jordan, P.W.; Camerini-Otero, R.D. Meiotic recombination mirrors patterns of germline replication in mice and humans. *Cell* **2021**, *184*, 4251–4267.e20. [[CrossRef](#)] [[PubMed](#)]
20. Mihola, O.; Pratto, F.; Brick, K.; Linhartova, E.; Kobets, T.; Flachs, P.; Baker, C.L.; Sedlacek, R.; Paigen, K.; Petkov, P.M. Histone methyltransferase PRDM9 is not essential for meiosis in male mice. *Genome Res.* **2019**, *29*, 1078–1086. [[CrossRef](#)]
21. Mihola, O.; Landa, V.; Pratto, F.; Brick, K.; Kobets, T.; Kusari, F.; Gasic, S.; Smagulova, F.; Grey, C.; Flachs, P. Rat PRDM9 shapes recombination landscapes, duration of meiosis, gametogenesis, and age of fertility. *BMC Biol.* **2021**, *19*, 86. [[CrossRef](#)] [[PubMed](#)]
22. La Spada, A.R.; Taylor, J.P. Repeat expansion disease: Progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* **2010**, *11*, 247–258. [[CrossRef](#)]
23. Kuhn, R.M.; Haussler, D.; Kent, W.J. The UCSC genome browser and associated tools. *Brief. Bioinform.* **2013**, *14*, 144–161. [[CrossRef](#)] [[PubMed](#)]
24. Funk, C.C.; Casella, A.M.; Jung, S.; Richards, M.A.; Rodriguez, A.; Shannon, P.; Donovan-Maiye, R.; Heavner, B.; Chard, K.; Xiao, Y.; et al. Atlas of Transcription Factor Binding Sites from ENCODE DNase Hypersensitivity Data across 27 Tissue Types. *Cell Rep.* **2020**, *32*, 108029. [[CrossRef](#)] [[PubMed](#)]
25. Markopoulos, G.; Noutsopoulos, D.; Mantziou, S.; Gerogiannis, D.; Thrasyvoulou, S.; Vartholomatos, G.; Kolettas, E.; Tzavaras, T. Genomic analysis of mouse VL30 retrotransposons. *Mob. DNA* **2016**, *7*, 10. [[CrossRef](#)] [[PubMed](#)]
26. Tarailo-Graovac, M.; Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **2009**, *25*, 4.10.1–4.10.14. [[CrossRef](#)] [[PubMed](#)]
27. Markopoulos, G.S. A Systems Biology Approach on the Regulatory Footprint of Human Endogenous Retroviruses (HERVs). *Diseases* **2022**, *10*, 98. [[CrossRef](#)]
28. Huang, W.; Loganantharaj, R.; Schroeder, B.; Fargo, D.; Li, L. Pavis: A tool for peak annotation and visualization. *Bioinformatics* **2013**, *29*, 3097–3099. [[CrossRef](#)] [[PubMed](#)]
29. McLean, C.Y.; Bristor, D.; Hiller, M.; Clarke, S.L.; Schaar, B.T.; Lowe, C.B.; Wenger, A.M.; Bejerano, G. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **2010**, *28*, 495–501. [[CrossRef](#)] [[PubMed](#)]
30. Mantziou, S.; Markopoulos, G.S. Origins and Function of VL30 lncRNA Packaging in Small Extracellular Vesicles: Implications for Cellular Physiology and Pathology. *Biomedicines* **2021**, *9*, 1742. [[CrossRef](#)] [[PubMed](#)]
31. Gillespie, M.; Jassal, B.; Stephan, R.; Milacic, M.; Rothfels, K.; Senff-Ribeiro, A.; Griss, J.; Sevilla, C.; Matthews, L.; Gong, C. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **2022**, *50*, D687–D692. [[CrossRef](#)] [[PubMed](#)]
32. Vartholomatos, E.; Mantziou, S.; Alexiou, G.A.; Lazari, D.; Sioka, C.; Kyritsis, A.; Markopoulos, G.S. An NF-kappaB- and Therapy-Related Regulatory Network in Glioma: A Potential Mechanism of Action for Natural Antiglioma Agents. *Biomedicines* **2022**, *10*, 935. [[CrossRef](#)] [[PubMed](#)]
33. Hays, F.A.; Schirf, V.; Ho, P.S.; Demeler, B. Solution formation of Holliday junctions in inverted-repeat DNA sequences. *Biochemistry* **2006**, *45*, 2467–2471. [[CrossRef](#)] [[PubMed](#)]
34. Myers, S.; Bowden, R.; Tumian, A.; Bontrop, R.E.; Freeman, C.; MacFie, T.S.; McVean, G.; Donnelly, P. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* **2010**, *327*, 876–879. [[CrossRef](#)] [[PubMed](#)]
35. Gel, B.; Díez-Villanueva, A.; Serra, E.; Buschbeck, M.; Peinado, M.A.; Malinverni, R. regioneR: An R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* **2016**, *32*, 289–291. [[CrossRef](#)]
36. Goodier, J.L.; Kazazian, H.H. Retrotransposons revisited: The restraint and rehabilitation of parasites. *Cell* **2008**, *135*, 23–35. [[CrossRef](#)]
37. Hancks, D.C.; Kazazian Jr, H.H. Active human retrotransposons: Variation and disease. *Curr. Opin. Genet. Dev.* **2012**, *22*, 191–203. [[CrossRef](#)] [[PubMed](#)]
38. Georgiou, I.; Noutsopoulos, D.; Dimitriadou, E.; Markopoulos, G.; Apergi, A.; Lazaros, L.; Vaxevanoglou, T.; Pantos, K.; Syrrou, M.; Tzavaras, T. Retrotransposon RNA expression and evidence for retrotransposition events in human oocytes. *Hum. Mol. Genet.* **2009**, *18*, 1221–1228. [[CrossRef](#)]
39. Dimitriadou, E.; Noutsopoulos, D.; Markopoulos, G.; Vlaikou, A.-M.; Mantziou, S.; Traeger-Synodinos, J.; Kanavakis, E.; Chrousos, G.P.; Tzavaras, T.; Syrrou, M. Abnormal DLK1/MEG3 imprinting correlates with decreased HERV-K methylation after assisted reproduction and preimplantation genetic diagnosis. *Stress* **2013**, *16*, 689–697. [[CrossRef](#)]
40. Vlaikou, A.-M.; Manolagos, E.; Noutsopoulos, D.; Markopoulos, G.; Liehr, T.; Vetro, A.; Ziegler, M.; Weise, A.; Kreskowski, K.; Papoulidis, I. An interstitial 4q31. 21q31. 22 microdeletion associated with developmental delay: Case report and literature review. *Cytogenet. Genome Res.* **2014**, *142*, 227–238. [[CrossRef](#)]
41. Riegel, M. Human molecular cytogenetics: From cells to nucleotides. *Genet. Mol. Biol.* **2014**, *37*, 194–209. [[CrossRef](#)] [[PubMed](#)]



42. Hong, D.; Fritz, A.J.; Gordon, J.A.; Tye, C.E.; Boyd, J.R.; Tracy, K.M.; Fietze, S.E.; Carr, F.E.; Nickerson, J.A.; Van Wijnen, A.J. RUNX1-dependent mechanisms in biological control and dysregulation in cancer. *J. Cell. Physiol.* **2019**, *234*, 8597–8609. [[CrossRef](#)]
43. Baer, C.; Walter, W.; Hutter, S.; Nadarajah, N.; Kern, W.; Haferlach, T.; Haferlach, C.; Meggendorfer, M. Whole Genome Sequencing Identifies Microdeletions Affecting TET2 and RUNX1 with Clinical Impact in Myeloid Malignancies. *Blood* **2021**, *138*, 3672. [[CrossRef](#)]
44. Li, C.H.; Coffey, E.L.; Dall'Agnese, A.; Hannett, N.M.; Tang, X.; Henninger, J.E.; Platt, J.M.; Oksuz, O.; Zamudio, A.V.; Afeyan, L.K. MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature* **2020**, *586*, 440–444. [[CrossRef](#)]
45. Pitzianti, M.B.; Santamaria Palombo, A.; Esposito, S.; Pasini, A. Rett syndrome in males: The different clinical course in two brothers with the same microduplication MECP2 Xq28. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3075. [[CrossRef](#)]
46. Candelo, E.; Ramirez-Montaña, D.; Pachajoa, H. Microduplication of Xp22. 31 and MECP2 pathogenic variant in a girl with Rett syndrome: A case report. *Iran. J. Med. Sci.* **2019**, *44*, 347. [[PubMed](#)]
47. Tekendo-Ngongang, C.; Dahoun, S.; Nguefack, S.; Moix, I.; Gimelli, S.; Zambo, H.; Morris, M.A.; Sloan-Béna, F.; Wonkam, A. MECP2 duplication syndrome in a patient from Cameroon. *Am. J. Med. Genet. Part A* **2020**, *182*, 619–622. [[CrossRef](#)]
48. Mahipal, A.; Tella, S.H.; Kommalapati, A.; Anaya, D.; Kim, R. FGFR2 genomic aberrations: Achilles heel in the management of advanced cholangiocarcinoma. *Cancer Treat. Rev.* **2019**, *78*, 1–7. [[CrossRef](#)]
49. Lam, K.-W.G.; Brick, K.; Cheng, G.; Pratto, F.; Camerini-Otero, R.D. Cell-type-specific genomics reveals histone modification dynamics in mammalian meiosis. *Nat. Commun.* **2019**, *10*, 3821. [[CrossRef](#)]
50. Brick, K.; Thibault-Sennett, S.; Smagulova, F.; Lam, K.-W.G.; Pu, Y.; Pratto, F.; Camerini-Otero, R.D.; Petukhova, G.V. Extensive sex differences at the initiation of genetic recombination. *Nature* **2018**, *561*, 338–342. [[CrossRef](#)]
51. Eichman, B.F.; Vargason, J.M.; Mooers, B.H.; Ho, P.S. The Holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3971–3976. [[CrossRef](#)] [[PubMed](#)]
52. Bowater, R.P.; Bohálová, N.; Brázda, V. Interaction of Proteins with Inverted Repeats and Cruciform Structures in Nucleic Acids. *Int. J. Mol. Sci.* **2022**, *23*, 6171. [[CrossRef](#)] [[PubMed](#)]
53. Pereira, R.; Oliveira, J.; Sousa, M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J. Clin. Med.* **2020**, *9*, 132. [[CrossRef](#)] [[PubMed](#)]
54. Liang, D.; Cram, D.S.; Tan, H.; Linpeng, S.; Liu, Y.; Sun, H.; Zhang, Y.; Tian, F.; Zhu, H.; Xu, M.; et al. Clinical utility of noninvasive prenatal screening for expanded chromosome disease syndromes. *Genet. Med.* **2019**, *21*, 1998–2006. [[CrossRef](#)] [[PubMed](#)]
55. Poirsier, C.; Besseau-Ayasse, J.; Schluth-Bolard, C.; Toutain, J.; Missirian, C.; Le Caignec, C.; Bazin, A.; de Blois, M.C.; Kuentz, P.; Catty, M.; et al. A French multicenter study of over 700 patients with 22q11 deletions diagnosed using FISH or aCGH. *Eur. J. Hum. Genet.* **2016**, *24*, 844–851. [[CrossRef](#)]
56. Boon, E.M.J.; Faas, B.H.W. Benefits and limitations of whole genome versus targeted approaches for noninvasive prenatal testing for fetal aneuploidies. *Prenat. Diagn.* **2013**, *33*, 563–568. [[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.