Research Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2024.59.009376

Investigation of SINE, LINE, and LTR as Potential Breakpoint Regions in Detected CNVs in a Single-Center Study

Ali Torabi[1]*, Ali Çiçekli[1], Sara Razzaghi[2], Ebru Marzioğlu Özdemir[3], Özkan Bağcı[3] and Nadir Koçak[4]

[1]*Medical Genetics Research Associate, University of Selcuk, Faculty of Medicine, Department of Medical Genetics, Turkey*

[2]*Medical Student, University of Selcuk, Faculty of Medicine, Turkey*

[3]*Assistant Professor, University of Selcuk, Faculty of Medicine, Department of Medical Genetics, Turkey*

[4]*Associate Professor, University of Selcuk, Faculty of Medicine, Department of Medical Genetics, Turkey*

***Corresponding author:** Ali Torabi, Medical Genetics Research Associate, University of Selcuk, Faculty of Medicine, Department of Medical Genetics, Konya, Turkey

ARTICLE INFO

Received: December 11, 2024

Published: December 16, 2024

Citation: Ali Torabi, Ali Çiçekli, Sara Razzaghi, Ebru Marzioğlu Özdemir, Özkan Bağcı and Nadir Koçak. Investigation of SINE, LINE, and LTR as Potential Breakpoint Regions in Detected CNVs in a Single-Center Study. Biomed J Sci & Tech Res 59(5)-2024. BJSTR. MS.ID.009376.

ABSTRACT

Copy Number Variations (CNVs) represent significant sources of genetic diversity and are associated with various genetic disorders, including developmental, neuropsychiatric, and oncological conditions. While Low-Copy Repeats (LCRs) are well-documented as mediators of CNV formation through Non-Allelic Homologous Recombination (NAHR), the contributions of other repetitive elements such as short interspersed nuclear elements (SINEs), long Interspersed Nuclear Elements (LINEs), and Long Terminal Repeats (LTRs) remain underexplored. This study investigates the prevalence and role of these repetitive elements as breakpoint regions in CNVs detected in a single-center cohort. A retrospective analysis of SNP-Microarray data from 116 postnatal patients between December 2022 and May 2023 identified 19 patients with 30 CNVs that met the study's inclusion criteria. Repetitive elements at CNV breakpoints were assessed using the UCSC Genome Browser and Repeat Masker track. Approximately 23.3% of CNVs were located in regions enriched with SINEs, LINEs, and LTRs, with LINE elements, particularly the L1 family, being most prominent. These findings underscore the significant role of these elements in CNV formation, likely mediated through mechanisms akin to NAHR. This study highlights the importance of considering diverse repetitive elements, beyond LCRs, in understanding CNV-mediated genomic instability. The results have implications for clinical genetics, emphasizing the need for comprehensive genomic analyses and advanced diagnostic tools to better characterize CNVs and their associated repetitive elements. Further research is warranted to elucidate the mechanisms by which these elements contribute to CNV formation and their potential therapeutic implications.

Keywords: CNV; Breakpoint; SINE; LINE; LTR

Abbreviations: CNVs: Copy Number Variations; LCRs: Low-Copy Repeats; NAHR: Non-Allelic Homologous Recombination; LINEs: Interspersed Nuclear Elements; UPS: Uniparental Disomy; LTRs: Long Terminal Repeats; TADs: Topologically Associated Domains NF 1: Neurofibromatosis Type 1

Introduction

Copy number variations (CNVs) are alterations in the genome that result in the duplication or deletion of large genomic regions, leading to a variation in the copy number of specific genes. CNVs are an important source of genetic diversity and have been associated with a variety of genetic disorders, including developmental and neuropsychiatric disorders, as well as cancer. Low-copy repeats (LCRs) have been extensively characterized for their role in CNV formation due to their high sequence homology which facilitates non-allelic homologous recombination (NAHR) [1]. However, other repetitive elements such as short interspersed nuclear elements (SINEs), long Interspersed Nuclear Elements (LINEs) and Long Terminal Repeats (LTRs) may also play critical roles in the formation of CNVs. SINEs, LINEs and LTRs are transposable elements that can insert themselves into

various locations within the genome, potentially leading to genomic instability and CNV formation [2]. Despite their potential importance, the role of these elements as CNV breakpoints has not been fully investigated. This study investigates the prevalence and significance of SINEs, LINEs and LTRs as breakpoint regions in CNVs detected in a single-centre cohort.

Methodology

A total of 116 postnatal patients who underwent SNP-Microarray analysis in our centre between December 2022 and May 2023 were included in this study. The inclusion criteria were based on the presence of CNVs and excluded normal results, mosaic results, Uniparental Disomy (UPD), results with aneuploidy and results with segmental aneuploidy to focus on relevant CNVs. This resulted in a final study population of 19 patients. CNVs reported in selected patients were analysed using SNP-Microarray technology (Illumina Beadchip Microarray, Infinium HTS). The genomic coordinates of CNVs were obtained and cross-referenced with UCSC Genome Browser to identify the presence of repetitive elements (SINE, LINE, LTR) using the RepeatMasker track (Figure 1). A total of 30 CNVs were analysed from the selected patients. Repeats of the same class and family located at the start and end points of CNVs were considered as potential breakpoints. The analysis focused on identifying the presence of SINEs, LINEs and LTRs in these breakpoint regions.

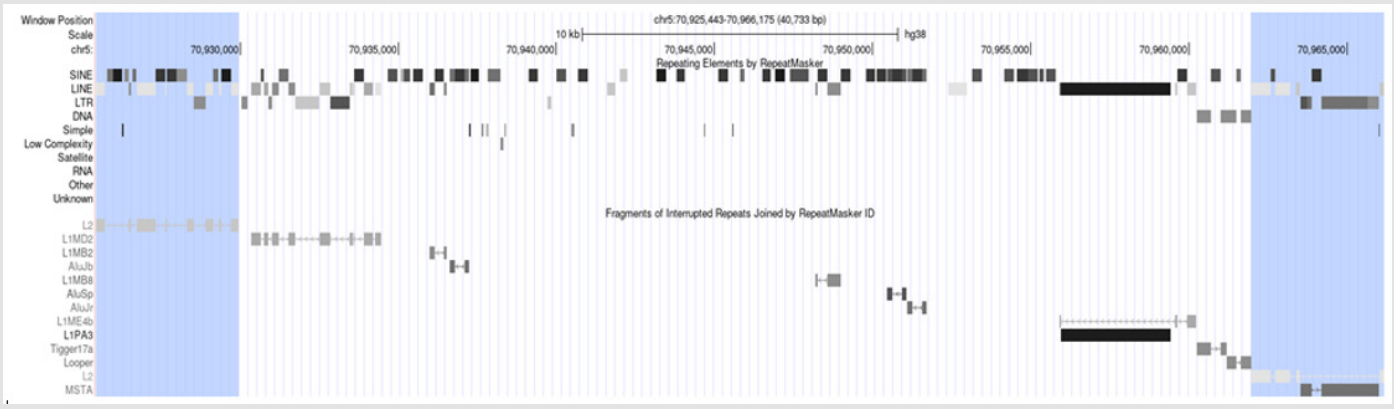

Figure 1: Visualization of repetitive elements analysis associated with CNV breakpoints on chromosome 5 using the UCSC Genome Browser. Genomic coordinates spanning positions 70,925,443 to 70,966,175 (hg38) were analyzed with RepeatMasker to identify SINEs, LINEs, LTRs, and other repeat families. The diagram highlights the distribution and density of various repetitive elements, which were cross-referenced to determine their potential role as CNV breakpoint regions.

Results

This study revealed that approximately %23.3 of CNVs were located in regions enriched with SINEs, LINEs and LTRs. The majority of these copy number variations (CNVs) were found to be associated with LINE families, particularly L1 elements. The presence of these elements at CNV breakpoints indicates their potential role in mediating CNV formation. Notably, CNVs located between these repetitive elements exhibited a higher recurrence probability, aligning with the propensity of these regions to undergo non-allelic homologous recombination. This finding highlights the importance of considering various types of repetitive elements, not only LCRs, in studies of CNV formation and genomic instability.

Discussion

Our findings contribute to the growing evidence that repetitive elements such as SINEs, LINEs and LTRs are crucial in CNV formation. While the role of LCRs in genomic rearrangements has been extensively studied, the involvement of other repetitive elements has received less attention. The identification of SINEs, LINEs and LTRs at CNV breakpoints in our study cohort suggests that these elements also contribute to genomic instability and the formation of CNVs. LINEs, particularly the L1 family, are autonomous retrotransposons capable of self-replication and insertion into new positions within the genome. This activity can disrupt genes and regulatory elements, leading to genomic instability. LINEs significantly contribute to shaping the genome's structure and function by mediating recombination events that lead to deletions, duplications, and other structural variations. Additionally, the retrotransposition of LINEs can influence gene expression and contribute to the evolution of gene families by creating new exons and promoters [3]. SINEs, such as Alu elements, are non-autonomous retrotransposons that rely on the enzymatic machinery of LINEs for their mobilisation. Alu elements are the most abundant transposon elements in the human genome and have been implicated in various genomic rearrangements. They can mediate non-allelic homologous recombination, leading to deletions, duplications and inversions. Furthermore, SINEs can affect gene regulation by inserting into promoter regions or creating alternative splice sites, thus altering gene expression patterns [4].

LTRs are sequences derived from endogenous retroviruses integrated into the host genome. They contain regulatory elements that can influence the expression of nearby genes by acting as promoters, enhancers or silencers. LTRs may also contribute to the formation of CNVs and other structural variations by mediating recombination events. Additionally, LTRs play a role in the regulation of gene networks involved in immune responses and cellular differentiation [5]. In addition to their role in genomic instability and CNV formation, LINEs, SINEs and LTRs contribute to the overall plasticity and evolution of the genome. These elements can actively generate genetic diversity by creating new genetic variants and influencing the architecture of regulatory networks. They have been implicated in the evolution of species-specific traits and adaptations by enabling the acquisition of novel functions and regulatory mechanisms [6]. For example, the amylase gene family has evolved through duplications mediated by repetitive elements, leading to differences in starch digestion abilities among different populations [7]. Furthermore, repetitive elements may impact the three-dimensional organisation of the genome. They contribute to the formation of Topologically Associated Domains (TADs), regions of the genome that interact more frequently with each other than with adjacent regions. This organisation is crucial for the regulation of gene expression and the maintenance of genomic integrity. Disruptions in TADs caused by repetitive elements can lead to abnormal gene expression and disease [8]. The propensity of SINEs, LINEs and LTRs to mediate non-allelic homologous recombination is supported by their ability to generate regions of sequence homology that can misalign during DNA replication or repair processes [9]. This mechanism is similar to how LCRs facilitate CNV formation and indicates that multiple types of repetitive elements can act as substrates for recombination events leading to CNVs [10]. Our study aligns with previous research indicating that repeated sequences are enriched in the vicinity of CNV breakpoints [11].

This enrichment indicates that repetitive elements form hotspots for genomic instability, making them key players in the generation of structural variations within the genome. The involvement of LINEs, particularly L1, in our study is consistent with their known ability to retrotranspose and integrate into new genomic locations and further contribute to genomic instability [12]. The interplay between different types of repetitive elements and genomic instability highlights the complexity of structural variations of the genome. The presence of repetitive elements such as SINEs, LINEs and LTRs at CNV breakpoints suggests that these elements contribute to the genomic architecture's dynamic nature and potentially facilitate both deleterious and beneficial variations [13]. For example, SINEs, particularly Alu elements, have been implicated in disease-associated CNVs, such as those involved in Neurofibromatosis Type 1 (NF1) and Peutz-Jeghers syndrome [14,15].

LINEs, particularly the L1 family, are known to promote genomic instability through their ability to retrotranspose and mediate non-allelic homologous recombination events. These events can lead to the formation of pathogenic CNVs, such as those associated with intellectual disability and other neurodevelopmental disorders [16,17]. LTRs, derived from endogenous retroviruses, also play a significant role in genomic rearrangements by providing sequences that can recombine and result in CNVs. LTR-mediated CNVs have been observed in various genetic disorders, including those affecting the immune response and cellular differentiation [18] Furthermore, the involvement of multiple types of repetitive elements in CNV formation underscores the need for comprehensive genomic analyses in clinical settings. Diagnostic tools that can accurately detect and characterize CNVs, including breakpoint regions, are essential for understanding the genetic basis of various disorders and developing effective treatment strategies [19,20].

The integration of high-throughput sequencing technologies and advanced bioinformatics approaches can enhance our ability to identify and interpret CNVs and their associated recurrent elements, thereby improving patient care and outcomes [21]. Future research should continue to explore the specific mechanisms by which SINEs, LINEs and LTRs contribute to CNV formation. In addition, studies should investigate the potential therapeutic interventions that may reduce the effects of CNVs mediated by these elements. Understanding the broader landscape of repetitive elements involved in CNV formation will enhance our ability to address genetic disorders associated with genomic instability.

Conclusion

In conclusion, our study emphasizes the important role of SINEs, LINEs, and LTRs as breakpoint regions in the formation of CNVs. These repetitive elements, similar to LCRs, contribute to genomic instability and facilitate the formation of CNVs through mechanisms such as non-allelic homologous recombination. Understanding these mechanisms provides valuable insights into genetic disorders and offers potential avenues for improving diagnostic and therapeutic strategies. Future research should continue to explore the diverse roles of repetitive elements in genomic instability and their implications in clinical genetics.

References

1. Ondrej Pös, Jan Radvanszky, Gergely Buglyó, Zuzana Pös, Diana Rusnakova, et al. (2021) DNA copy number variation: Main characteristics, evolutionary significance, and pathological aspects. biomedical journal 44(5): 548-559.
2. Cardoso AR, Oliveira M, Amorim A, Azevedo L (2016) Major influence of repetitive elements on disease-associated copy number variants (CNVs). Human Genomics 10(1): 30.
3. Cordaux R, Batzer MA (2009) The impact of retrotransposons on human genome evolution. Nature reviews genetics 10(10): 691-703.
4. Hedges D, Deininger P (2007) Inviting instability: transposable elements, double-strand breaks, and the maintenance of genome integrity. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 616(1-2): 46-59.

5. Medstrand P, Van De Lagemaat LN, Mager DL (2002) Retroelement distributions in the human genome: variations associated with age and proximity to genes. Genome research 12(10): 1483-1495.
6. Kazazian Jr HH (2004) Mobile elements: drivers of genome evolution. Science 303(5664): 1626-1632.
7. GH P, Nathaniel J Dominy, Katrina G Claw, Arthur S Lee, Heike Fiegler, et al. (2007) Diet and the evolution of human amylase gene copy number variation. Nat Genet 39: 1188-1190.
8. Dixon JR, Selvaraj S, Yue F, Audrey Kim, Yan Li, et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485(7398): 376-380.
9. Beck CR, Garcia-Perez JL, Badge RM (2011) Moran JV LINE-1 elements in structural variation and disease. Annual review of genomics and human genetics 12(1): 187-215.
10. Bailey JA, Eichler EE (2006) Primate segmental duplications: crucibles of evolution, diversity and disease. Nature Reviews Genetics 7(7): 552-564.
11. Carvalho CM, Lupski JR (2016) Mechanisms underlying structural variant formation in genomic disorders. Nature Reviews Genetics 17(4): 224-238.
12. Kazazian Jr HH (1998) Mobile elements and disease. Current opinion in genetics & development 8(3): 343-350.
13. Belancio VP, Hedges DJ (2008) Deininger P Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome research 218(3): 343-358.
14. MR W, D A Marchuk, L B Andersen, R Letcher, H M Odeh, et al. (1990) Type 1 neurofibromatosis gene: Identification of a large transcript disrupted in three patients. Science 249(4965): 182-186.
15. Boardman LA, Couch FJ, Burgart LJ, D Schwartz, R Berry, et al. (2000) Genetic heterogeneity in Peutz-Jeghers syndrome. Human mutation 16(1): 23-30.
16. Weening R, De Boer M, Kuijpers T, Neefjes V, Hack W, et al. (2000) Point mutations in the promoter region of the CYBB gene leading to mild chronic granulomatous disease. Clinical & Experimental Immunology 122(3): 410-417.
17. Van de Lagemaat LN, Landry J-R, Mager DL, Medstrand P (2003) Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. TRENDS in Genetics 19(10): 530-536.
18. Chen JM, Chuzhanova N, Stenson PD, Férec C, Cooper DN (2005) Meta-Analysis of gross insertions causing human genetic disease: Novel mutational mechanisms and the role of replication slippage. Human mutation 25(2): 207-221.
19. Pös O, Jan Radvanszky, Jakub Styk, Zuzana Pös, Gergely Buglyó, et al. (2021) Copy number variation: methods and clinical applications. Applied Sciences 11(2): 819.
20. Zhao M, Wang Q, Wang Q, Jia P, Zhao Z (2013) Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. BMC bioinformatics 14(Suppl 11): S1.
21. Korbel JO, Urban AE, Affourtit JP, Brian Godwin, Fabian Grubert, et al. (2007) Paired-end mapping reveals extensive structural variation in the human genome. Science 318(5849): 420-426.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2024.59.009376

Ali Torabi. Biomed J Sci & Tech Res