GGAA-Microsatellites of NR0B1 Promoter Region in Ewing's Sarcoma Patients and Healthy Individuals of a Southern Brazilian Population

doi: https://doi.org/10.32635/2176-9745.RBC.2022v68n2.2350

Microssatélites GGAA na Região Promotora de NR0B1 em Pacientes com Sarcoma de Ewing e Indivíduos Saudáveis de uma População no Sul do Brasil

Microsatélites GGAA en la Región Promotora de NR0B1 en Pacientes con Sarcoma de Ewing y Individuos Sanos de una Población del Sur de Brasil

Rodrigo Rosa de Stefani[1]; Elisa Cristina de Toni[2]; Caroline Brunetto de Farias[3]; André Tesainer Brunetto[4]; Algemir Lunardi Brunetto[5]; Rafael Roesler[6]; Clarice Sampaio Alho[7]; Deise Cristine Friedrich[8]

ABSTRACT

Introduction: The very aggressive soft tissue and bone pediatric tumor Ewing's sarcoma (ES) is caused in most cases by the chromosomal translocation t(11;22)(q24;q12), which encodes an aberrant chimeric transcription factor (EWS-FLI1) that regulates target genes, including the critical oncogene *NR0B1* (Xp21.2), via GGAA-microsatellites. **Objective:** Analyze the GGAA-microsatellites of *NR0B1* promoter region of ES patients and healthy subjects in the population investigated. **Method:** Ten male ES patients and 71 adult healthy males from Rio Grande do Sul state, Brazil, were included in this study. DNA from peripheral blood samples was extracted, amplified by PCR, sequenced by the Sanger method and analyzed by capillary electrophoresis. Total number of GGAA-motifs, length of microsatellite in base pairs, number of segments separated by "A" insertions, and the greatest number of consecutive GGAA-motifs were analyzed as well. Statistical analyses were performed in the SPSS® statistical software and p-value <0.05 was considered significant. **Results:** A total of 21 different alleles was identified in the 81 subjects, with 24.2 allele [$(GGAA)_7A(GGAA)_7A(GGAA)_{10}$] being the most frequent, but when comparing the data between the two groups, no significant difference was found. **Conclusion:** The sample investigated had a wide variation of microsatellite structure, including the presence of rare alleles, allowing the opportunity to describe this population as an essential step to identify genetic implications in ES tumorigenesis.

Key words: sarcoma, Ewing; DAX-1 orphan nuclear receptor; microsatellite repeats/genetics; genetic predisposition to disease; oncogenes.

RESUMO

Introdução: O sarcoma de Ewing (ES) é um tumor pediátrico de ossos e partes moles muito agressivo, causado, na maioria das vezes, pela translocação cromossômica t(11;22)(q24;q12), codificando um fator de transcrição quimérico aberrante (EWS-FLI1) que regula genes-alvo, incluindo o oncogene *NR0B1* (Xp21.2), via microssatélites GGAA. **Objetivo:** Analisar os microssatélites GGAA da região promotora de *NR0B1* em pacientes com ES e indivíduos saudáveis da população em investigação. **Método:** Foram incluídos dez pacientes do sexo masculino com diagnóstico de ES e 71 indivíduos adultos hígidos do sexo masculino do Estado do Rio Grande do Sul, Brasil. O DNA foi extraído de sangue periférico e amplificado por PCR, sequenciado pelo método de Sanger e analisado por eletroforese capilar. Foram analisados o número total de repetições GGAA, comprimento total do microssatélite em pares de bases, número de segmentos separados por inserções "A" e maior número de repetições GGAA consecutivas. As análises estatísticas foram realizadas no *software* estatístico SPSS® e o valor de p<0,05 foi considerado significativo. **Resultados:** Um total de 21 alelos diferentes foi identificado nos 81 indivíduos, com o alelo 24,2 [$(GGAA)_7A(GGAA)_7A(GGAA)_{10}$], sendo o mais frequente; mas, ao comparar os dados entre os dois grupos, nenhuma diferença significativa foi encontrada. **Conclusão:** A amostra estudada é altamente variável em termos de estrutura de microssatélites, incluindo a presença de alelos raros, dando a oportunidade de descrever essa população, o que é uma etapa fundamental na identificação de implicações genéticas na tumorigênese do ES.

Palavras-chave: sarcoma de Ewing; receptor nuclear órfão DAX-1; repetições de microssatélites/genética; predisposição genética para doença; oncogenes.

RESUMEN

Introducción: El sarcoma de Ewing (ES) es un tumor pediátrico de huesos y tejidos blandos muy agresivo, que se presenta con mayor frecuencia por translocación cromosómica t(11;22)(q24;q12), que codifica un factor de transcripción quimérico aberrante (EWS-FLI1) que regula los genes diana, incluido el oncogén *NR0B1* (Xp21.2), a través de microsatélites GGAA. **Objetivo:** Analizar los microsatélites GGAA de la región promotora de *NR0B1* en pacientes con ES y personas sanas de la población investigada. **Método:** Este estudio incluyó a diez pacientes varones con diagnóstico de ES y 71 varones adultos del estado de Rio Grande do Sul, Brasil. El ADN se extrajo de sangre periférica y se amplificó por PCR, secuenciado por el método de Sanger y analizado por electroforesis capilar. El número total de repeticiones GGAA, longitud total de microsatélites en pares de bases, número de segmentos separados por inserciones "A" y el mayor número de repeticiones GGAA consecutivas fueran analizados. Los análisis estadísticos se realizaron con el software estadístico SPSS® y se consideró significativo un valor de p<0,05. **Resultados:** Se identificaron un total de 21 alelos diferentes en los 81, siendo el alelo 24,2 [$(GGAA)_7A(GGAA)_7A(GGAA)_{10}$] el más frecuente, pero al comparar los datos entre los dos grupos, no hubo diferencia estadísticamente significativa. **Conclusión:** La muestra estudiada es muy variable en cuanto a estructura de microsatélites, incluyendo la presencia de alelos raros, lo que nos permite la oportunidad de describir la población estudiada, lo cual es un paso fundamental en la identificación de implicaciones genéticas en la tumorigénesis de ES.

Palabras clave: sarcoma de Ewing; receptor nuclear huérfano DAX-1; repeticiones de microsatélite/genética; predisposición genética a la enfermedad; oncogenes.

[1,2,7,8]Pontifícia Universidade Católica do Rio Grande do Sul, Faculdade de Medicina. Porto Alegre (RS), Brazil. E-mails: rodrigo.stefani@edu.pucrs.br; elisafar@gmail.com; csalho@pucrs.br; deise.friedrich@pucrs.br. Orcid iD: https://orcid.org/0000-0002-3395-7496; Orcid iD: https://orcid.org/0000-0002-2087-646X; Orcid iD: https://orcid.org/0000-0002-4819-9587; Orcid iD: https://orcid.org/0000-0002-4613-5301

[3,4,5]Universidade Federal do Rio Grande do Sul, Centro de Pesquisa Experimental, Hospital das Clínicas (CPE-HCPA) e Instituto de Câncer Infantil. Porto Alegre (RS), Brazil. E-mails: labpesquisa@ici.ong; adrebrunetto@ici.ong; institucional@ici.ong. Orcid iD: https://orcid.org/0000-0002-6435-6626; Orcid iD: https://orcid.org/0000-0002-7958-1279; Orcid iD: https://orcid.org/0000-0003-0668-6894

[6]Universidade Federal do Rio Grande do Sul. Instituto de Ciências Básicas da Saúde & CPE-HCPA, Departamento de Farmacologia. Porto Alegre (RS), Brazil. E-mail: rafaelroesler@hcpa.edu.br. Orcid iD: https://orcid.org/0000-0001-6016-2261

Corresponding author: Deise C. Friedrich. Avenida Ipiranga, 6681, Bloco 12C, Sala 132 - Jardim Botânico. Porto Alegre (RS), Brazil. CEP 90619-900. E-mail: deise.friedrich@pucrs.br

INTRODUCTION

Ewing's Sarcoma (ES) is a very aggressive soft tissue and bone tumor that occurs mainly from childhood to early adulthood, being the second most frequent primary bone cancer in children and adolescents[1]. Its prevalence is higher in European populations, with a slight preference for males. The 5-year survival rate is 70% in patients with localized tumors who undergo systemic chemotherapy and local control measures versus less than 30% that is recurrent or metastatic disease[1]. In Brazil, the age-adjusted incidence rate (0-14 years old) of bone tumors is 5,74 per million. ES is the second most frequent bone tumor with an age-adjusted incidence rate of 1,77 cases per million in children and adolescents (0-14 years old)[2].

It is believed that the neoplasm originates from the malignant transformation of progenitor cells from the mesoderm and the neural crest with its driver mutation being the chromosomal translocation t(11;22)(q24;q12) in 85-90% of cases. It encodes the EWS/ETS family fusion proteins[3]. EWS-FLI, the most prevalent aberrant chimeric transcription factor generated, regulates many genes, including the critical oncogenic target *NR0B1*, via GGAA-microsatellites on its promoter[4,5].

Before developing new therapeutic approaches to such a highly malignant disease it is necessary to understand the molecular structure and the emergent properties that regulate the promoter of *NR0B1*[6]. This region contains microsatellites consisting of multiple copies of GGAA core motif, with an occasional single "A" base insertion, which were identified as EWS-FLI binding sites, being necessary and sufficient for gene regulation[6]. Not only the absolute number of repeats present in a construct that determines the microsatellite's ability to function as a response element, but also the number of consecutive (in tandem) repeats is important[7]. Experimental analysis of DNA probes and promoter fragments with a variable number of repeats indicated that at least three consecutive GGAA-motifs are required for DNA binding and gene activation[8]. The binding and subsequent target gene activation is highly dependent on the number of motifs in microsatellites near promoter regions and the length of 18–26 was shown to confer maximal EWS-FLI responsiveness to target genes, but the mechanistic basis for this remains unknown[8,9]. Additionally, it was demonstrated that the position of the "A" insertions influences the binding of EWS-FLI oncoprotein to DNA[6].

The objective of this study was to identify and describe the molecular structure of GGAA-microsatellites on the *NR0B1* promoter region of unrelated ES patients and healthy subjects from a southern Brazilian population. Since the length and the microsatellite structure are important for DNA binding and gene activation, it was hypothesized that a longer and more complex microsatellite could be a risk factor for ES.

METHOD

In a cross-sectional study, 24 patients (age ranged from 1 to 21 years old; 10 males and 14 females) were diagnosed with Ewing's Sarcoma between 1991 and 2004 in the Pediatric Oncology Department of Porto Alegre Clinic Hospital (HCPA), Rio Grande do Sul State, Brazil. All the clinical diagnoses were confirmed by imaging and biopsies studies. Phenotypic information can be found in a previous study[10]. The non-ES group consisted of 71 adult males, geographically matched to the patients.

Ten ES male patients were included in the study. Only male subjects were enrolled because *NR0B1* is located in the X chromosome, which could lead to DNA sequencing errors in female samples (particularly in heterozygotes).

This study was approved by the Institutional Review Board, number 2.409.676, in compliance with the Declaration of Helsinki. All the subjects and/or parents/legal representatives, on behalf of the minors enrolled in the study, signed the Informed Consent Form (ICF).

DNA was isolated from peripheral blood samples according to Lahiri and Nurnberger[11] protocol. DNA was amplified in a final volume of 25µl reaction containing 1-10ng of DNA, 12.5µL of Master Mix Multiplex PCR Kit (QIAGEN GmbH, Hilden, Germany), 10 pmol/µL of upstream and downstream primers, and 5.0µL of Q Solution (QIAGEN GmbH, Hilden, Germany). PCR was performed on a Veriti Thermal Cycler (Applied Biosystems; Life Technologies, USA) using predenaturation, 38 cycles of 94°C-60s, 57°C-90s, and 72°C-30s, and a final extension. Sequencing was conducted under BigDyeTM terminator cycling conditions and performed using Automatic Sequencer ABI 3730XL (Applied Biosystems; Life Technologies, USA). The primers used to amplify the targeted regions (design using the Primer3 algorithm)[12] were 5'-TCTTATGCTGAGAATTCCAGGTC-3' and 5'- AAGAAGAGGGAGGATGGGA-3'. The sequencing had a better result with the forward primer, which was used for the proceeding analysis. In all sequences involved in this study, it was clearly discerned the forward primer followed by a preliminary sequence, GGAA-repeats, a posterior sequence, and finishing with primer reverse (Figure 1). Statistical analyses were performed in the SPSS® statistical software and p-value <0.05 were considered significant.

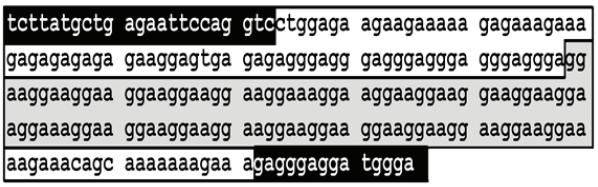

Figure 1. Standard *NR0B1* promoter allele with 25 GGAA-motifs and two single insertions of "A" base to show the reference patterns used to know the other alleles. In the upper black background is the forward primer; in up white background is the preliminary sequence; in gray background is the GGAA-microsatellite; in down white background is the posterior sequence; and in down black background is the reverse primer

RESULTS

The same primers, preliminary and posterior sequences identification ensured the exact starting location of the GGAA-microsatellite. With this accurate information, it was possible to identify the arrangement of this region on the *NR0B1* promoter. A total of 21 different alleles were identified in the 81 subjects, which were classified according to the amount and position of GGAA-motifs and single 'A' base insertions (Table 1).

Comparisons between patients and controls were made to verify if there were differences related to microsatellite length or structure that could be a risk factor for ES.

The allele frequencies between patients and controls were compared but no significant difference was found. Additionally, the alleles were separated into three groups according to the number of GGAA-motifs (<18, 18-26, and >26) since 18-26 would confer maximal EWS/FLI responsiveness[8,9], although none of the comparisons was significant (data not shown).

The 24.2 allele [$(GGAA)_7A(GGAA)_7A(GGAA)_{10}$ sequence] was the most frequent, being present in 33% of all subjects (50% in patients and 31% in controls). The highest frequency in patients was also not significant (p=0.4, Chi-Square Test with Yates correction).

The population was also described in terms of the total number of GGAA-motifs, the length of microsatellite in base pairs, the number of segments separated by "A"

Table 1. Different alleles by GGAA-motifs in *NR0B1* promoter

Allele's name[a]	Total of base pairs[b]	Number of patients/ controls	Allele sequence (between the preliminary and the posterior sequences)
16.1	65	1/0	$(GGAA)_5A(GGAA)_{11}$
17.1	69	0/7	$(GGAA)_5A(GGAA)_{12}$
18.1	73	0/2	$(GGAA)_5A(GGAA)_{13}$
19.1	77	0/3	$(GGAA)_5A(GGAA)_{14}$
21.2	86	0/1	$(GGAA)_7A(GGAA)_6A(GGAA)_8$
22.1	90	0/1	$(GGAA)_5A(GGAA)_{17}$
23.2-A	94	0/1	$(GGAA)_7A(GGAA)_6A(GGAA)_{10}$
23.2-B	94	0/2	$(GGAA)_7A(GGAA)_7A(GGAA)_9$
24.2	98	5/22	$(GGAA)_7A(GGAA)_7A(GGAA)_{10}$
25.2-A	102	2/14	$(GGAA)_7A(GGAA)_7A(GGAA)_{11}$
25.2-B	102	0/1	$(GGAA)_8A(GGAA)_7A(GGAA)_{10}$
26.2-A	106	1/4	$(GGAA)_7A(GGAA)_7A(GGAA)_{12}$
26.2-B	106	0/4	$(GGAA)_7A(GGAA)_8A(GGAA)_{11}$
26.2-C	106	0/3	$(GGAA)_8A(GGAA)_7A(GGAA)_{11}$
27.2	110	0/1	$(GGAA)_7A(GGAA)_7A(GGAA)_{13}$
28.2	114	1/0	$(GGAA)_7A(GGAA)_7A(GGAA)_{14}$
57.6-A	234	0/1	$(GGAA)_5A(GGAA)_{10}A(GGAA)_8A(GGAA)_4A(GGAA)_{11}A(GGAA)_9A(GGAA)_{10}$
57.6-B	234	0/1	$(GGAA)_5A(GGAA)_{10}A(GGAA)_8A(GGAA)_4A(GGAA)_{12}A(GGAA)_9A(GGAA)_9$
59.6	242	0/1	$(GGAA)_5A(GGAA)_{10}A(GGAA)_8A(GGAA)_4A(GGAA)_{14}A(GGAA)_9A(GGAA)_9$
60.6	246	0/1	$(GGAA)_5A(GGAA)_{10}A(GGAA)_8A(GGAA)_4A(GGAA)_{15}A(GGAA)_9A(GGAA)_9$
75.8	308	0/1	$(GGAA)_5A(GGAA)_{11}A(GGAA)_8A(GGAA)_4A(GGAA)_{13}A(GGAA)_4A(GGAA)_{12}A(GGAA)_9A(GGAA)_9$

([a]) *NR0B1* promoter alleles were named using the total number of GGAA-motifs followed by the number of single "A" insertions; Letters -A, -B, and -C were used to discriminate the alleles when the "A" insertions were in different positions.
([b]) Number of base pairs was calculated considering the first nucleotide after the preliminary sequence until the last nucleotide before the posterior sequence.

insertions, and the greatest number of consecutive GGAA-motifs in patients and control groups (Table 2).

DISCUSSION

The results demonstrate that the population investigated is highly variable since only 81 chromosomes were sequenced allowing the identification of 21 alleles. Table 1 also shows that no allele had a segment with less than four consecutive GGAA-motifs, which is a susceptibility factor for ES since previous findings showed that EWS/FLI only binds regions with at least three consecutive motifs[8].

In different Ewing cell lines, the level of expression of the oncogenic *NR0B1* had a positive correlation to the number of GGAA-repeats in its promoter[13]. Long and complex GGAA-microsatellites were found not only in the patient group but in the control group as well. So, the length of the microsatellite does not seem to be a risk factor for ES.

The data analyzed in Table 2 was based on a previous study of the same promoter region of *NR0B1* in European and African populations. In that study, it was shown that the GGAA-microsatellite is highly polymorphic in both populations, with longer and more variant sequences in the African population but they did not evaluate those characteristics in ES patients[14]. Since the length and the microsatellite structure are important for gene activation[6,7], the samples were described and analyzed according to these criteria. As there is no study with ES patients to compare with and there were no significant differences between patients and controls, the result may not be relevant.

The most frequent allele in European and African populations[14] was the allele 24.2, which corresponds to the findings hereof. Interestingly, 17-26 GGAA-motifs alleles are also the most common found in ES cell lines according to the same study.

It is known that the small number of subjects is the major limitation of the study, however, that was the number of males diagnosed in the period in Rio Grande do Sul, and this is a given scenario.

CONCLUSION

It was possible to identify many rare alleles in the *NR0B1* promoter region and to describe the population, which is a fundamental step in identifying genetic implications in ES tumorigenesis. As many authors hypothesized that microsatellites are involved in tumorigenesis and there are scarce studies on this theme, the current article can contribute to the knowledge that long and complex structured GGAA-repeats are characteristics of the healthy population as well. Finally, more studies are necessary to elucidate a possible relationship between the GGAA-microsatellite structure of the *NR0B1* promoter region and the susceptibility to ES in the population investigated.

CONTRIBUTIONS

Rodrigo Rosa de Stefani contributed with data interpretation and wording; Elisa Cristina de Toni with data collection analysis and wording; Caroline Brunetto de Farias with study design and critical review; André Tesainer Brunetto with study design and critical review;

Table 2. Comparison between patients (n=10) and controls (n=71) concerning *NR0B1* promoter allele data

	Mean	Minimum	Maximum	Standard deviation	P-value[a]
Total number of GGAA-motifs					
Patients	24	16	28	3.1	0.9
Controls	24	17	75	10.4	
Length of microsatellite in base pairs					
Patients	97.9	65	114	12.6	0.9
Controls	106.2	69	308	42.9	
Number of segments separated by "A" insertions					
Patients	2.9	2	3	0.3	0.9
Controls	3.1	2	9	1.2	
Greatest number of GGAA-motifs in a segment					
Patients	10.9	10	14	1.3	0.6
Controls	11.1	8	17	1.5	

([a]) P-value calculated using Mann-Whitney test comparing the distribution between patients and controls.

Algemir Lunardi Brunetto with study design and critical review; Rafael Roesler with study design and critical review; Clarice Sampaio Alho with study design, data interpretation and critical review; Deise Cristine Friedrich with data interpretation and critical review. All the authors approved the final version to be published.

CONFLICTS OF INTEREST

There is no conflict of interest to declare.

FUNDING

Grant from "*Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul*" (FAPERGS) and the National Council for Scientific and Technological Development (CNPq).

REFERENCES

1. Grünewald TGP, Cidre-Aranaz F, Surdez D, et al. Ewing sarcoma. Nat Rev Dis Primers. 2018;4(1):5. doi: https://doi.org/10.1038/s41572-018-0003-x
2. Balmant NV, Reis RS, Santos MO, et al. Incidence and mortality of bone cancer among children, adolescents and young adults of Brazil. Clinics. 2019;74:e858. doi: https://doi.org/10.6061/clinics/2019/e858
3. Lawlor ER, Sorensen PH. Twenty years on: What do we really know about Ewing sarcoma and what is the path forward? Crit Rev Oncog. 2015;20(3-4):155-71. doi: https://doi.org/10.1615/critrevoncog.2015013553
4. Guillon N, Tirode F, Boeva V, et al. The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. PLoS One. 2009;4(3):e4932. doi: https://doi.org/10.1371/journal.pone.0004932
5. Gangwal K, Sankar S, Hollenhorst PC, et al. Microsatellites as EWS/FLI response elements in Ewing's sarcoma. Proc Natl Acad Sci U S A. 2008;105(29):10149-54. doi: https://doi.org/10.1073/pnas.0801073105
6. Gangwal K, Close D, Enriquez CA, et al. Emergent properties of EWS/FLI regulation via GGAA microsatellites in Ewing's sarcoma. Genes Cancer. 2010;1(2):177-87. doi: https://doi.org/10.1177/1947601910361495
7. Gangwal K, Lessnick SL. Microsatellites are EWS/FLI response elements: genomic "junk" is EWS/FLI's treasure. Cell Cycle. 2008;7(20):3127-32. https://doi.org/10.4161/cc.7.20.6892
8. Johnson KM, Taslim C, Saund RS, et al. Identification of two types of GGAA-microsatellites and their roles in EWS/FLI binding and gene regulation in Ewing sarcoma. PLoS One. 2017;12(11):e0186275. doi: https://doi.org/10.1371/journal.pone.0186275
9. Johnson KM, Mahler NR, Saund RS, et al. Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. Proc Natl Acad Sci U S A. 2017;114(37):9870-5. doi: https://doi.org/10.1073/pnas.1701872114
10. Silva DSB, Sawitzki FR, Toni EC, et al. Ewing's sarcoma: analysis of single nucleotide polymorphism in the EWS gene. Gene. 2012;509(2):263-6. doi: https://doi.org/10.1016/j.gene.2012.08.012
11. Lahiri DK, Nurnberger JI Jr. A rapid non-enzymatic method for the preparation of HMV DNA from blood for RFLP studies. Nucleic Acids Res. 1991;19(19):5444. doi: https://doi.org/10.1093/nar/19.19.5444
12. Primer3 [Internet]. v. 0.4.0. Cambridge (MA): Whitehead Institute for Biomedical Research. c2007 - [cited 2021 Apr 22]. Available from: http://frodo.wi.mit.edu/primer3
13. García-Aragoncillo E, Carrillo J, Lalli E, et al. DAX1, a direct target of EWS/FLI1 oncoprotein, is a principal regulator of cell-cycle progression in Ewing's tumor cells. Oncogene. 2008;27:6034-43. doi: https://doi.org/10.1038/onc.2008.203
14. Beck R, Monument MJ, Watkins WS, et al. EWS/FLI-responsive GGAA microsatellites exhibit polymorphic differences between European and African populations. Cancer Genet. 2012;205(6):304-12. doi: https://doi.org/10.1016/j.cancergen.2012.04.004

Recebido em 30/9/2021
Aprovado em 18/2/2022

Associate Editor: Claudio Gustavo Stefanoff. Orcid iD: https://orcid.org/0000-0001-7050-3269
Scientific Editor: Anke Bergmann. Orcid iD: https://orcid.org/0000-0002-1972-8777