

## **Misclassification of a frequent loss of function variant from *PMS2CL* pseudogene as a *PMS2* variant in Brazilian patients**

Anthony Vladimir Campos Segura<sup>1</sup>, Sara Iolanda Oliveira da Silva<sup>1</sup>, Karina Miranda Santiago<sup>1</sup>, Rafael Canfield Brianese<sup>1</sup>, Dirce Maria Carraro<sup>1,2</sup>, Giovana Tardin Torrezan<sup>1,2\*</sup>

1- Clinical and Functional Genomics Group, A.C.Camargo Cancer Center, São Paulo, Brazil

2- National Institute of Science and Technology in Oncogenomics and Therapeutic Innovation, São Paulo, Brazil.

\*Corresponding author:

Giovana Tardin Torrezan, Clinical and Functional Genomics Group, International Research Center/CIPE, A.C.Camargo Cancer Center

Mailing address: 440 Taguá St., Sao Paulo, SP, 01508-010 Brazil.

[giovana.torrezan@accamargo.org.br](mailto:giovana.torrezan@accamargo.org.br)

<https://orcid.org/0000-0002-8659-5329>

### **Statements and Declarations:**

The authors declare no conflict of interests.

**Financial support:** This work was supported by the São Paulo Research Foundation (FAPESP 2014/50943-1 and 2022/05162-8); the National Council for Scientific and Technological Development (CNPq - 465682/2014-6) and the Coordination for the Improvement of Higher Education Personnel (CAPES - 88887.136405/2017-00).

### **Authors contribution:**

Conceptualization: G.T.T.; Formal analysis: A.V.C.S., S.I.O.S., K.M.S, R.C.B., D.M.C. and G.T.T; Methodology: A.V.C.S. and S.I.O.S., Investigation: A.V.C.S., S.I.O.S., K.M.S, R.C.B., and G.T.T; Funding acquisition: D.M.C. and G.T.T., Supervision: G.T.T.; Resources: D.M.C. and G.T.T., Writing—original draft preparation: A.V.C.S. and G.T.T.; Writing—review and editing: A.V.C.S., S.I.O.S., K.M.S, R.C.B., D.M.C. and G.T.T. All authors have read and agreed to the published version of the

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.  
manuscript.

## Abstract

*PMS2*, a Lynch Syndrome gene, presents challenges in genetic testing due to the existence of multiple pseudogenes. This study aims to describe a series of cases harboring a rare LoF variant in the *PMS2CL* pseudogene that has been incorrectly assigned to *PMS2* with different nomenclatures. We reviewed data from 647 Brazilian patients who underwent multigene genetic testing at a single center to identify those harboring the *PMS2* V1:c.2186\_2187delTC or V2:c.2182\_2184delACTinsG variants, allegedly located at *PMS2* exon 13. Gene-specific PCR and transcript sequencing was performed. Among the 647 individuals, 1.8% (12) carried the investigated variants, with variant allele frequencies ranging from 15 to 34%. By visually inspecting the alignments, we confirmed that both V1 and V2 represented the same variant and through gene-specific PCR and *PMS2* transcript analysis, we demonstrated that V1/V2 is actually located in the *PMS2CL* pseudogene. Genomic databases (ExAC and gnomAD) report an incidence of 2.5% - 5.3% of this variant in the African population. Currently, V1 is classified as "uncertain significance" and V2 as "conflicting" in ClinVar, with several laboratories classifying them as "pathogenic". We identified a frequent African *PMS2CL* LoF variant in the Brazilian population that is misclassified as a *PMS2* variant. It is likely that V1/V2 have been erroneously assigned to *PMS2* in several manuscripts and by clinical laboratories, underscoring a disparity-induced matter. Considering the limitations of short-read NGS differentiating between certain regions of *PMS2* and *PMS2CL*, using complementary methodologies is imperative to provide an accurate diagnosis.

**Keywords:** *PMS2*, *PMS2CL*, pseudogene, multigene panels, conflicting variants.

## Introduction

The use of Next-Generation Sequencing has significantly improved the accessibility and effectiveness of genetic diagnosis for patients with hereditary cancer predisposition syndromes. However, challenges arise in complex variants, such as errors found in classifying variants within genes that have homologous genomic regions or pseudogenes with sequences very similar to the original gene. These findings highlight the ongoing need for refinement in variant identification and classification methodologies.

*PMS2* is a gene involved in DNA mismatch repair. Deficiencies in this gene are associated with Lynch Syndrome (LS), a condition related to increased risks for developing colorectal, endometrial, ovarian and other cancers. Genetic testing for this gene is challenging due to the existence of multiple pseudogenes [1, 2]. Fourteen pseudogenes have been identified and described as  $\psi$ 1 to  $\psi$ 14, overlapping with some or all of *PMS2* exons 1 to 5 and varying in length. Additionally, there is the *PMS2CL* pseudogene (formerly known as  $\psi$ 0) with high homology to the 3' end of *PMS2* in exons 9 and 11 to 15 [3, 4].

Prior research has highlighted the misclassification of the loss-of-function (LoF) variant c.2182\_2184delACTinsG within the *PMS2CL* pseudogene, mistakenly ascribed to exon 13 of the *PMS2* gene [5]. In this study, we describe a series of cases harboring this frequent LoF variant in the *PMS2CL* that has been incorrectly identified as a pathogenic variant in *PMS2*, with varying nomenclatures throughout the years. We confirmed the variant location in *PMS2CL* using different molecular techniques in several patients. Additionally, we discuss the importance of performing alternative methods to circumvent the NGS limitations in distinguishing similar regions shared between *PMS2* and its pseudogenes and the relevance of increasing sequencing efforts across diverse populations. Finally, we empathize the pivotal role these endeavors play in ensuring precision in genetic diagnoses and mitigating disparities in genetic testing.

## Methods

### *Patient cohort and data collection*

We collected retrospective data from 647 patients who performed genetic testing with multigene panels harboring 26 to 126 cancer predisposing genes between 2018 and 2023 at the A.C. Camargo Cancer Center. A retrospective analysis was carried out using Sophia DDM platform to identify patients with *PMS2* (NM\_000535.5) c.2182\_2184delACTinsG (V1) or c.2186\_2187delTC (V2) variants. Clinical information (age of onset, tumor histology, familial history of cancer) was collected from hospital electronic records. All patients signed a written informed consent and were included in studies approved by the Institutional Review Board of A.C. Camargo Cancer Center (protocol numbers 2483/18 and 2497/18).

### *Gene-specific PCR*

Germline DNA from saliva or leucocytes from 12 patients with the presence of V1/V2 variant were obtained. A gene-specific PCR (GSP) was performed to examine exon 13 of *PMS2* and exon 4 of *PMS2CL*, like described by Hendricks [6]. Briefly, PCR primers were design to anneal preferentially in the desired gene by positioning the primer in a variable region between *PMS2* and *PMS2CL* that contains three mismatched bases between gene and pseudogene. Primers sequences used were PMS2\_E13GSP\_F: GAAGTTTTGTGACACTTAGCTGAGTAG and PMS2\_E13GSP\_R: TTGGCCTCCCAGAGTGCTG; PMS2CL\_E4GSP\_F: TTGTGACACTTAGCTGAATTATGTTGT and PMS2CL\_E4GSP\_R: TTATGTTAGCGAGGCTGGTCTCAAAC (underscored bases refer to those 3 discriminating bases). PCR products were analyzed by amplicon NGS using the Ion GeneStudio S5 system, followed by sequence analysis using the Integrative Genomics Viewer (IGV).

### *Transcript analysis*

As a complementary and confirmatory analysis, RNA was extracted from peripheral blood of 2 patients. RNA was then converted into cDNA and subjected to a nested PCR to analyze the transcription of *PMS2*. The first PCR targeted regions from exons 10 to 15, and the second PCR focused on exons 12 to 14. The PCR products were subjected to amplicon sequencing NGS, and the sequences were analyzed using the CLC Genomics workbench software.

## Results

The allegedly *PMS2* variant received two different nomenclatures on previous testing: V1: c.2186\_2187delTC; p.(Leu729Glnfs) - rs587779335, and V2: c.2182\_2184delinsG (p.Thr728Alafs) - rs1554294508. We observed that the nomenclature of the variant changed in 2020, after a software update. Upon visual inspection of the alignments from patients with these variants, we confirmed that both nomenclatures represent the same variant (Figure 1A). The difference between the variants is that for V1, a single nucleotide variant (SNV) *PMS2*:c.2182A>G; p.(Thr78Ala) is called as a separated variant, while in V2 this variant is considered part of the delins event, as recommended by HGVS nomenclature.

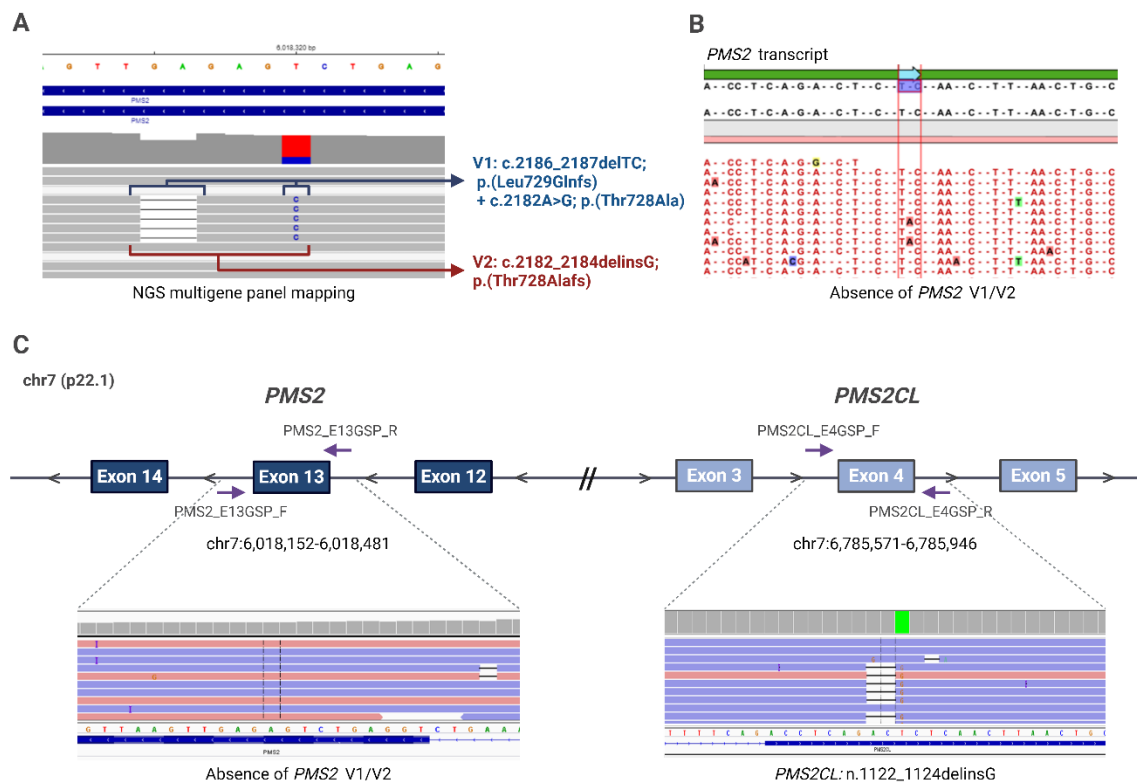
Among the 647 individuals evaluated, 12 (1.8%) patients carried the V1/V2 variant, exhibiting variant allele frequencies ranging from 15% to 34% (Table 1). These patients presented distinct cancer types, such as breast (8), colorectal (2) and gastric (1). Four patients had germline pathogenic variants detected in other cancer predisposing genes. The racial distribution of these patients was diverse, with three patients self-identified as Black (one having molecularly confirmed African ancestry), one as Brown and three as White.

We performed gene-specific PCR of all 12 cases and determined that V1/V2 is not present in the *PMS2* gene. Moreover, sequencing analysis of exon 4 of the *PMS2CL* pseudogene unequivocally demonstrated the presence of this variant (Figure 1B). The corresponding correct nomenclature for this variant in *PMS2CL* is n.1122\_1124delinsG. To further confirm that V1/V2 is not present in the *PMS2* gene, we conducted an analysis of *PMS2* transcripts by nested PCR in two patients, and in both cases no read containing V1/V2 variant was detected (Figure 1C).

**Table 1. Clinical and genetic features of patients with V1/V2 variant.**

ID	Tumor (age range)	Self-declared race	Genetic testing date	Prior nomenclature	Revised nomenclature	VAF	Additional tests/results
P_01	Breast (25-30)	Black	2018	PMS2: c.2186_2187delTC	PMS2CL: n.1122_1124delinsG	25.4%	Ancestry: AFR 66% <sup>a</sup>
P_02	Breast (40-45)	White	2018	PMS2: c.2186_2187delTC	PMS2CL: n.1122_1124delinsG	15.4%	<i>BRCA1</i> GPV
P_03	Breast (40-45)	White	2018	PMS2: c.2186_2187delTC	PMS2CL: n.1122_1124delinsG	17.2%	<i>BRCA1</i> GPV
P_04	Breast (35-40)	Black	2018	PMS2: c.2186_2187delTC	PMS2CL: n.1122_1124delinsG	34.2%	-
P_05	Gastric (45-0)	White	2019	PMS2: c.2186_2187delTC	PMS2CL: n.1122_1124delinsG	19.4%	-
P_06	CRC (25-30)	NA	2020	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	20.0%	dMMR (MLH1/PMS2); somatic <i>MLH1</i> LoF variant
P_07	Breast (50-55)	NA	2021	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	14.8%	-
P_08	Rectum (45-50)	Brown	2022	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	18.4%	<i>MSH2</i> GPV
P_09	Breast (50-55)	NA	2022	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	18.8%	-
P_10	Asymptomatic (40-45)	NA	2022	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	18.8%	-
P_11	Breast (60-65)	NA	2023	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	19.0%	<i>MSH2</i> GPV
P_12	Breast (60-65)	Black	2023	PMS2: c.2182_2184delinsG	PMS2CL: n.1122_1124delinsG	21.9%	-

Abbreviations: CRC, colorectal cancer; NA, not available; AFR, African; GPV, germline pathogenic variant; dMMR; deficient mismatch repair; LoF, loss of function. <sup>a</sup>Patient with confirmed African ancestry by Axiom Precision Medicine Diversity Array (PMDA).



**Figure 1.** Analysis of *PMS2* V1/V2 variants and reclassification as a *PMS2CL* variant. **A.** NGS reads alignment of multigene DNA panel sequencing, showing the *PMS2* exon 13 region in IGV. The image depicts the bases involved in V1 and V2 nomenclatures. **B.** RNA sequencing for the *PMS2* transcript and analysis by CLC Genomics software, showing absence of V1/V2. **C.** Depiction of gene-specific PCR (GSP) performed for *PMS2* exon 13 and *PMS2CL* exon 4. Genomic coordinates (Hg19) correspond to amplicons location. Left alignment shows mapping of *PMS2* GSP, demonstrating the absence of V1/V2, and right alignment shows mapping of *PMS2CL* GSP, demonstrating the presence of *PMS2CL*: n.1122\_1124delinsG.

## Discussion

Here, we present data on a frequent *PMS2CL* African variant found in the Brazilian population, which is incorrectly classified as a *PMS2* variant. We detected this variant with two distinct nomenclatures (*PMS2*:c.2182\_2184delACTinsG and c.2186\_2187delTC) in 1.8% of genetic tests performed in our center. Through gene-specific PCR and *PMS2* transcript analysis, we established that both variants represent the same *PMS2CL* variant (n.1122\_1124delinsG). Moreover, in our center 4 patients have received a genetic test report from external laboratories containing this

variant mistakenly described as a *PMS2* LoF variant, 3 of them describing the variant as VUS and one as pathogenic, highlighting the relevance of our investigation.

Currently, for the *PMS2* gene, V1: c.2186\_2187delTC is classified as "Uncertain significance reviewed by expert panel" and V2 as "Conflicting" in ClinVar, with several laboratories classifying both as "pathogenic." The conflicting data for classifying both variants is related to the need of confirming the variant as a *PMS2* variant. The annotation by the International Society for Gastrointestinal Hereditary Tumours (InSiGHT) expert group in Clinvar regarding this variant is "This variant is likely to come from pseudogene". When no confirmation is performed, most likely this variant should be assigned to the *PMS2CL* pseudogene, making it not disease-causing [7].

The V1:c.2186\_2187delTC has been reported in the literature in more than 12 articles, published between 1995 and 2019. Few articles described the variant in compound heterozygosity with another *PMS2* variant in patients affected with Turcot Syndrome [8] or Constitutional Mismatch Repair Deficiencies (CMMRD) [9], indicating a true occurrence of the variant in *PMS2*. However, V1 has also been reported in individuals with other cancer types, including colorectal [1, 10], breast [11, 12] and prostate [13], and pseudogene interference was not ruled out in most of these studies. A recent pediatric study reported the finding of V1 in 2 patients with pilocytic astrocytoma, but after applying long-range PCR, they determined that the variant belonged to the pseudogene [14].

The V2: c.2182\_2184delinsG is much less cited in the literature, with only 2 articles referring to the variant [1, 5]. The more recent report of this variant most likely reflects updates in variant calling algorithms that incorporated the *PMS2*:c.2182A>G SNV as part of the indel variant, as recommended by HGVS rules. While Guindalini [1] did not describe any sequential technique confirming the variant to *PMS2*, Chong et al showed that indeed the variant was located at *PMS2CL* in all 5 tested patients. In Clinvar, the variant is described as "Conflicting", with 4 clinical laboratories classifying the variant as Pathogenic/Likely pathogenic and one as VUS. Based on the evidence outlined above, we believe that both V1 and V2 have been incorrectly assigned to *PMS2* and mistakenly classified as pathogenic in several articles and that these variants should be classified as pathogenic only when unequivocal confirmed to be within the *PMS2* gene.



Global genomic databases (ExAC and gnomAD) report an incidence of these variants ranging from 2.5% to 5.3% in the African population, while in ABraOM (a Brazilian genomic database) [15] V1 appears at a frequency of 1.3%. Brazil has a significant percentage of Afro-descendant population with an average African genetic ancestry of 12.7% [5, 16]. In our study, 57% (4/7) of the patients tested with the presence of the *PMS2CL* variant declared themselves as Black or Brown, and one had a 66% African genetic ancestry molecularly confirmed.

Genomics disparities represent a current challenge in clinical genetics [17]. It is well-documented that minorities and underrepresented populations often have more VUS detected in clinical genetic testing. Our report emphasizes this issue by highlighting a common African variant frequently misclassified as a VUS or as pathogenic. Our results also underscore the importance of addressing these misclassification errors caused by NGS's inability to differentiate short regions that are very similar between *PMS2* and its pseudogenes. The misclassification of variants in the *PMS2* gene can lead to significant consequences in clinical genetics, with clinicians misinterpreting the genetic risk and potentially providing inappropriate management or surveillance recommendations. Moreover, the racial disparity in genomic analysis further exacerbates the challenges faced by minority populations in receiving accurate and timely genetic testing results, stressing the need for increasing the available genomic data from diverse populations.

## References

1. Guindalini RSC, Win AK, Gulden C, et al (2015) Mutation Spectrum and Risk of Colorectal Cancer in African American Families with Lynch Syndrome. *Gastroenterology* 149:1446–1453. <https://doi.org/10.1053/J.GASTRO.2015.07.052>
2. Soto JL, Blanco I, Díez O, et al (2018) Documento de consenso sobre la implementación de la secuenciación masiva de nueva generación en el diagnóstico genético de la predisposición hereditaria al cáncer. *Med Clin (Barc)* 151:80.e1-80.e10. <https://doi.org/10.1016/J.MEDCLI.2017.12.010>
3. Blount J, Prakash A (2018) The changing landscape of Lynch syndrome due to PMS2 mutations. *Clin Genet* 94:61–69. <https://doi.org/10.1111/CGE.13205>
4. Bouras A, Lefol C, Ruano E, et al (2023) PMS2 or PMS2CL? Characterization of variants detected in the 3' of the PMS2 gene. *Genes, Chromosom Cancer*. <https://doi.org/10.1002/GCC.23193>
5. Chong AS, Chong G, Foulkes WD, Saskin A (2020) Reclassification of a frequent African-origin variant from PMS2 to the pseudogene PMS2CL. *Hum Mutat* 41:749–752. <https://doi.org/10.1002/HUMU.23978>
6. Hendriks YMC, Jagmohan-Changur S, Van Der Klift HM, et al (2006) Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome). *Gastroenterology* 130:312–322. <https://doi.org/10.1053/J.GASTRO.2005.10.052>
7. Thompson BA, Goldgar DE, Paterson C, et al (2013) A multifactorial likelihood model for MMR gene variant classification incorporating probabilities based on sequence bioinformatics and tumor characteristics: a report from the Colon Cancer Family Registry. *Hum Mutat* 34:200–209. <https://doi.org/10.1002/HUMU.22213>
8. De Vos M, Hayward BE, Picton S, et al (2004) Novel PMS2 Pseudogenes Can Conceal Recessive Mutations Causing a Distinctive Childhood Cancer Syndrome. *Am J Hum Genet* 74:954. <https://doi.org/10.1086/420796>
9. Bakry D, Aronson M, Durno C, et al (2014) Genetic and clinical determinants of constitutional mismatch repair deficiency syndrome: Report from the

- constitutional mismatch repair deficiency consortium. *Eur J Cancer* 50:987–996.  
<https://doi.org/10.1016/J.EJCA.2013.12.005>
10. Rodríguez-Salgado LE, Silva-Aldana CT, Medina-Méndez E, et al (2022) Frequency of actionable Exomic secondary findings in 160 Colombian patients: Impact in the healthcare system. <https://doi.org/10.1016/j.gene.2022.146699>
  11. Eliade M, Skrzypski J, Baurand A, et al (2017) The transfer of multigene panel testing for hereditary breast and ovarian cancer to healthcare: What are the implications for the management of patients and families? *Oncotarget* 8:1957–1971. <https://doi.org/10.18632/ONCOTARGET.12699>
  12. Gardner SA, Weymouth KS, Kelly WS, et al (2018) Evaluation of a 27-gene inherited cancer panel across 630 consecutive patients referred for testing in a clinical diagnostic laboratory. *Hered Cancer Clin Pract* 16:.. <https://doi.org/10.1186/S13053-017-0083-8>
  13. Leongamornlert D, Saunders E, Dadaev T, et al (2014) Frequent germline deleterious mutations in DNA repair genes in familial prostate cancer cases are associated with advanced disease. *Br J Cancer* 110:1663–1672. <https://doi.org/10.1038/BJC.2014.30>
  14. Stoltze UK, Foss-Skiftesvik J, van Overeem Hansen T, et al (2023) Genetic predisposition and evolutionary traces of pediatric cancer risk: a prospective 5-year population-based genome sequencing study of children with CNS tumors. *Neuro Oncol* 25:761–773. <https://doi.org/10.1093/NEUONC/NOAC187>
  15. Naslavsky MS, Scliar MO, Yamamoto GL, et al (2022) Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. *Nat Commun* 13:26. <https://doi.org/10.1038/S41467-022-28648-3>
  16. Durães RO, Berardinelli GN, da Costa AM, et al (2020) Role of Genetic Ancestry in 1,002 Brazilian Colorectal Cancer Patients From Barretos Cancer Hospital. *Front Oncol* 10:494547. <https://doi.org/10.3389/FONC.2020.00145/BIBTEX>
  17. Manrai AK, Funke BH, Rehm HL, et al (2016) Genetic Misdiagnoses and the Potential for Health Disparities. *N Engl J Med* 375:655–665. <https://doi.org/10.1056/NEJMSA1507092>