Novel Chromosomal-abnormalities in biopsies of Pterygion found by RNA-seq

Beriozka Manzaneda-Murguia[1, 4], Jani Pacheco-Aranibar[2], Cinthia Cordova-Barrios[3], Rosemary Zapana Begazo[3]; Francis W. Jacobo-Valdivia[4]; Javier L. Valero Quispe[3], Ivo Palomino Valverde[3], Jhony R. Rodriguez Mamani[5], Jose Villanueva Salas[3, 4], Julio C. Bernabé-Ortiz[3, 4]
1Servicio de Oftamologia, Hospital III Juliaca EsSalud. Puno Peru
[2]Instituto de Biotecnologia del ADN Uchumayo, Arequipa - Peru.
3Universidad Católica de Santa María, Arequipa Peru
[4]Escuela de Postgrado, Universidad Católica de Santa María, Arequipa Peru
[5]Universidad Nacional del Altiplano. Puno - Peru.
* Correspondence: Email: jbernabe@ucsm.edu.pe

Abstract
Aims To identify the presence of fusion genes in pterygium extracted in a hospital in southern Peru. Methods In this perspective and exploratory study, we used next-generation sequencing (NGS) methodologies to process four pterygium samples obtained by surgical exeresis. The information obtained from RNA-seq was analyzed by Arriba (version 1.2.0), a novel computational tool with high sensitivity and a short processing time run to detect fusion genes. Results Fusion genes were found in all samples. Some of these genes did not show functional characteristics, unlike the PSMC3 - NXF1 and WNK1 - ZNF79 fusion genes, which have retained protein domains and in-frame fusions. This means that they synthesize proteins that can act as powerful drivers in the genesis of pterygium. Conclusion The discovery of fusion genes shows structural variations in the development of the pterygium. These findings suggest these genes should be new therapeutic targets in future research.
Keywords: pterygium, RNA-seq, genes fusion, UV-pterigium

1. Introduction

The discovery of fusion genes in several tumors has led to a deeper understanding of molecular mechanisms in proliferative tissues, revolutionizing the approach to diagnosis, prognosis, and therapy.[1] A fusion gene, also called a chimeric gene or hybrid gene, is the juxtaposition of two genes on the same chromosome or different chromosomes. They form because of structural variants, typically defined as significant variations in the structure of the human genome. Types of structural variants include insertions, deletions, inversions, and, in particular, translocations.[2]
Fusion genes have been identified initially in the pathogenesis of hematolymphoid malignancies. Nowadays, they are recognized as crucial driving events in the genesis of many benign and malignant neoplasms, including epithelial and soft tissue tumors.[1] Recurrent fusion genes are comparatively rare with other forms of somatic mutations, probably because of complicated generation mechanisms, and pathological fusion events are highly pervasive with prominent transformations. In ophthalmology, studies have reported fusion genes in lacrimal gland tumors, primary tumors of orbital soft tissues, and ocular adnexa.[3] However, in the literature, there are no reports of the presence of fusion genes in the pterygium tissues. Identifying these chimeric proteins will improve our understanding of the origins of pterygium and progression factors.
Pterygium is a pathological tissue of the ocular surface with a high degree of proliferative cells. It represents a chronic, fibrovascular, degenerative process in the corneo-conjunctival junction.[4] Pterygium is in the interpalpebral region and can be unipolar or bipolar, being more frequent on the nasal side. It can also be unilateral or bilateral, its most common shape being triangular, with its base in the periphery and the apex invading the cornea. According to the level of advancement, pterygium can be classified into four grades: the fibrovascular tissue reaches the limbus (Grade1); when it covers the cornea by 2 mm (Grade 2); when it reaches the margin of the pupil (Grade 3), and Grade 4 when it exceeds the pupillary margin.[5,6] There is ample evidence implicating anti-apoptotic and immunological mechanisms, cytokines, growth factors, modulators of the extracellular matrix, genetic factors, viral infections, exposure to ultraviolet radiation, and other factors in the pathogenesis of pterygium.[4-8] The pterygium produces astigmatism and loss of vision, thus, definitive treatment is surgical.[1] The identification of fusion genes will allow us to recognize new therapeutic targets for treatments against pterygium. In oncogenic processes, gene fusions are not automatically related since they can be found in many neoplasms, even non-tumor ones. These fusions have a vital role in the study of cancer development. The likelihood of fusion gene transcripts developing into tumors is a challenging

doi.org/10.31838/hiv22.02.722 Received: 16.07.21, Revised: 02.08.21, Accepted: 19.09.21

and not yet fully explored research problem.[7]
The main applications of NGS can be classified into whole-genome sequencing (WGS), whole-exome sequencing (WES), and whole transcriptome sequencing (RNA-Seq), enabling different levels of investigation.[2] In this sense, RNA-Seq is ideal for fusion gene discovery, allowing the detection of multiple alternative splice variants resulting from a fusion event. Besides its low cost and fast turnaround time, these distinctive features make RNA-Seq very popular in gene fusion studies.[2]

2. Material and Methods

Study design

The study was carried out from January 2020 to May 2021 under a prospective and exploratory approach. All the patients treated in the ophthalmology service of Hospital III Juliaca EsSalud were part of the work universe. The study population comprised four patients who underwent pterygium exeresis in February 2020 at Hospital III Juliaca.

Patient Selection

Inclusion criteria. All patients were operated on for pterygium excision with conjunctival autograft in the ophthalmology service of Hospital III Juliaca in February 2020. The patients whose preoperative diagnosis was clear and decisive for performing the surgery. For patients whose history and medical records were in the files of the institution and who filled out the required information, any lack of information would be observed for inclusion. Patients who voluntarily signed the informed consent.
Exclusion criteria. History of eye diseases: glaucoma, uveitis. Patients with systemic diseases, diabetes, or rheumatoid arthritis. Pterygium of the atrophic type. Patients voluntarily decided not to take part in the study.

Ethical aspects

The ethics committee of Hospital III Juliaca (Protocol Code: 1299-2020-494) approved the development of the study. The research was conducted in compliance with the primary and operational principles of internationally required ethical standards. Among the most important are respect for the individual, their right to self-determination, informed consent, and putting the health of patients and psychophysical integrity first.

Extraction of samples.

The samples were obtained in the operating room of Hospital III Juliaca. An expert surgeon performed the same technique of exeresis and free plasty in all patients. The procedure began with asepsis and antisepsis of the operative area. Topical anesthesia and 1 mL of 2% lidocaine with epinephrine at a concentration of 1:10,000 were applied to the subconjunctival tissue surrounding the pterygium. The apex of the pterygium was dissected from the cornea to the limbus with a number 15 scalpel, and the body of the pterygium was cut with Wescott scissors. The subconjunctival tissues were carefully removed. Mitomycin C 0.2% was then placed for 3 minutes and applied to cotton swabs in the subconjunctival tissue, and, subsequently, an eye wash was performed with 50 cc of 0.9% sodium chloride. The sclera was covered by an autograft of conjunctiva from the superior temporal region and sutured with separate 10/00 nylon stitches.
The fresh pterygium conjunctiva specimens were washed with 50 cc of distilled water and divided into two by making a longitudinal incision along the body and head of the pterygium. They were then placed in 500 µL of DMEM culture medium (Dulbecco Modified Eagle Medium) at 4 °C and another bottle with 500 µL of TRIzol reagent at -20 °C and transported to the laboratory.
The samples were processed in the Molecular Biology Laboratory of the Universidad Católica de Santa María in Arequipa, Peru. The sterile technique was used to handle the tissues, which were processed in the laminar flow cabinet.

RNA extraction

The sample was washed twice with sterile PBS, and the samples were cut with Vannas Rumex capsulotomy scissors and straight Katena forceps, getting cubes of approximately 1 x 1 mm. The sample was placed in an Eppendorf tube, and 500 µL of chloroform was added. Subsequently, it was centrifuged at 12,000 rpm for 10 minutes. Transfer of the aqueous phase to a new Eppendorf tube (approximately 300-400 pL of the sample) was performed without touching the interface. This process was repeated twice. Once finished, 500 µL of isopropanol was added and centrifuged at 12,000 rpm for 10 minutes. Isopropanol was removed by inversion on a paper towel once, isolating the pellet. The DNA pellet was resuspended with 500 µL of 70% ethanol and centrifuged at 12,000 rpm for 10 minutes.
Once the supernatant was removed by inversion on a paper towel, re-isolating the pellet. It was allowed to dry for approximately 30 seconds, then the lid of the Eppendorf was closed, and the pellet was resuspended with 50 pL of sterile water. The integrity and quantity of RNA were analyzed by electrophoresis and spectrophotometry, respectively. RNA was stored in isolation at -20°C.

Next-generation sequencing (NGS) technique

The RNA samples were sent to the Admera laboratory in New Jersey, United States, for sequencing using the NGS technology. The NGS platform used for sequencing was the NovaSeq 6000, which is the technology of Illumina high-throughput sequencing platform.[14] After applying quality control to the samples, RNA was extracted to construct the libraries.
The library is prepared from random fragmentation of the complementary DNA sample, followed by 5'

and 3′ ligation. Alternatively, "tagmentation" combines the fragmentation and ligation reactions in a single step, which increases the efficiency of the library preparation process. Adapter-ligated fragments are amplified by PCR and gel purified.

To get sequencing, the library is loaded into a flow cell where fragments are captured in a field of surface-bound oligos that are complementary to library adapters. Clusters then amplified each fragment. When cluster generation is complete, the templates are ready for actual sequencing on the machine.

The results obtained by NGS provide high-resolution data that allow the analysis of mutations. Another strength of NGS is its support for a wide range of applications to human genetic research.3 Finally, the software process the information and genera a FastQ type file.

Characteristic Chromosomal-abnormalities detection

The information obtained from RNA-seq was analyzed by the novel computational tool Arriba (version 1.2.0). Arriba has high sensitivity and a short execution time to detect and visualize gene fusions from RNA-Seq data using predetermined parameters.[15] Top performance provides the transcript sequence flanking the binding site, which helps with primer design for validation by Sanger sequencing. It also calculates the peptide sequence resulting from the chimeric transcript. This can serve as a basis for the prediction of fusion-derived neo-epitopes. In addition, it offers several useful features that go beyond mere fusion breakpoint prediction.[15] Supporting reads are minimal, the workflow above requires a single alignment step followed by a filtering process. The fraction of mismatches in supporting reads must be less than 0.8. Genes with more than 30% sequence identity are homologous and removed by the "homologous" filter. Once all the candidate alignments have been collected, a set of filters is applied to remove artifacts and enrich the results with a high level of confidence.[15]

3. Results

Sample 1

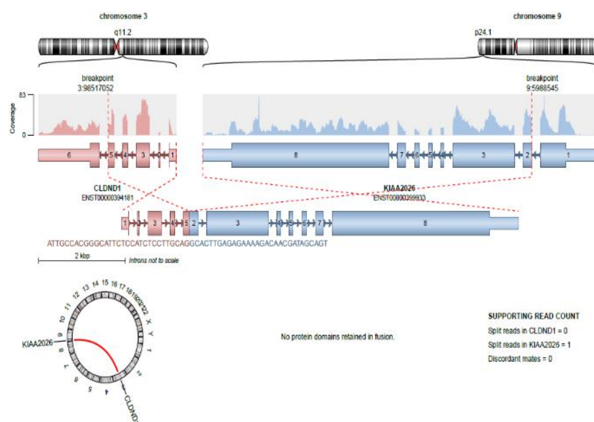

Figure 1. CLDND1 - KIAA2026 fusion gene.

Figure 1 shows a breakpoint at the level of exon 5 of the CLDND1 gene found on chromosome 3 and a breakpoint at exon 2 of the KIAA2026 gene located on chromosome 9. This results in the fusion gene that does not retain protein domains.

Sample 1 corresponds to a 47-year-old female patient who performs agricultural activities and underwent excision of a third-degree nasal pterygium in the left eye. She currently has pterygium in the temporal conjunctival region in the left eye and the nasal conjunctiva of the right eye.

Sample 2

In this sample, we found three fusion genes (Figures 2-4). The sample corresponds to a 33-year-old female patient who does housework and was operated on for a third-degree pterygium in the nasal conjunctiva of the right eye.

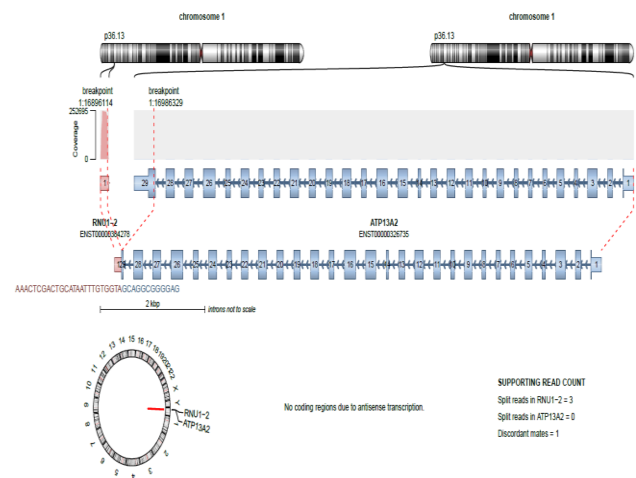

Figure 2. RNV1-2 - ATP13A2 fusion gene.

The RNU1-2 - ATP13A2 gene is formed on chromosome 1 (Figure 2). There is a breakpoint in part of exon 29 of the ATP13A2 protein that joins with exon 1 of the RNU1-2 gene. The RNU1-2 - ATP13A2 fusion gene has no coding regions because of antisense transcription.

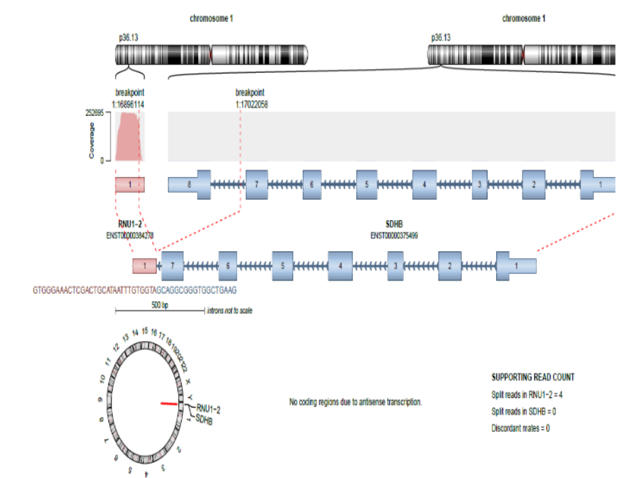

Figure 3. RNU1-2 - SDHB fusion gene.

The RNU1-2 - SDHB fusion gene is formed on chromosome 1 by breaking exon 7 of the SDHB protein that joins with exon 1 of the RNU1-2 gene (Figure 3). The RNU1-2 - SDHB gene does not present coding regions because of antisense transcription.

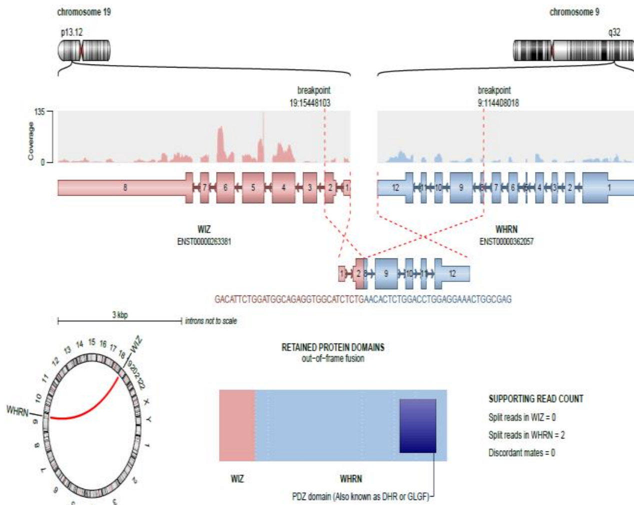

Figure 4. WIZ - WHRN fusion gene.

At the level of subband 2, subband 1, band 3, and region 1 on the short arm of chromosome 19, there is a break in exon 2 of the WIZ gene. This binds to the fragment produced because of the gap in exon 8 of the WHRN gene at the level of band 2, region 3 of the long arm of chromosome 9. Both fragments form the WIZ-WHRN fusion gene with retained protein domains, and out-of-frame fusion is obtained (Figure 4).

Sample 3

This sample reported five fusion genes (Figures 5-9). This tissue was removed by excision of a third-degree nasal pterygium in the left eye. The 66-year-old male patient performs agricultural activities and currently has a pterygium in his right eye.

In the case of the fusion gene 19:40114634 - SNORD3A (Figure5), this is formed by the union of exome 1 of the SNORD3A gene on chromosome 17, with the protein derived from the break at the level of subband 2, band 3, and region 1 from the long arm of chromosome 19. This gene does not code for proteins.

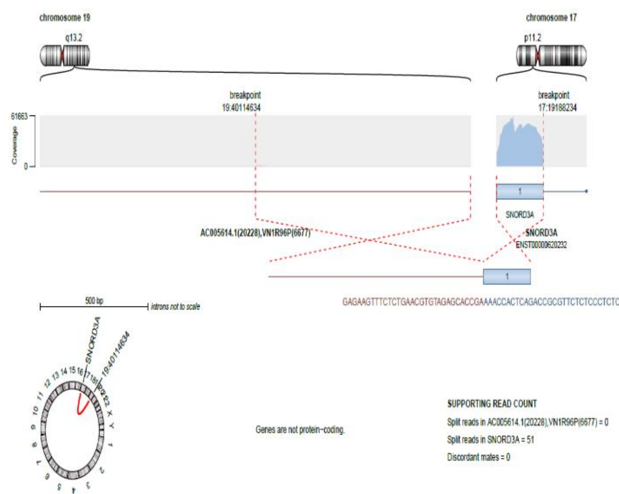

Figure 5. 19:40114634 - SNORD3A fusion gene.

The 6:164999470 - RNA5SP226 fusion gene is formed on chromosome 6 by the union of exon 1 of the RNA5SP226 gene with the proteins in band 7, region 2 of the long arm of chromosome 6. This gene does not code for proteins (Figure 6).

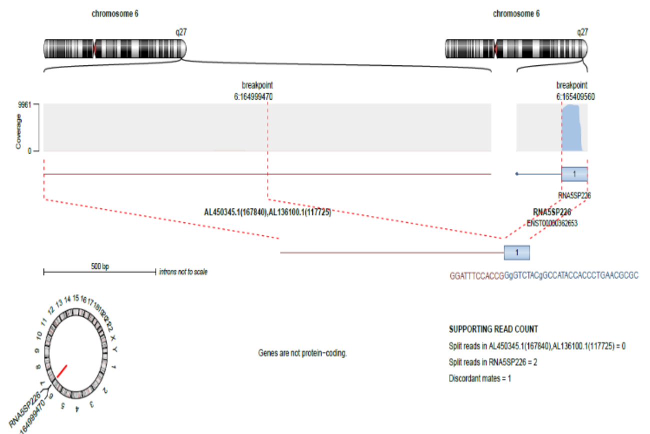

Figure 6. 6:164999470 - RNA5SP226 fusion gene.

The PSMC3 - NXF1 fusion gene is formed on chromosome 11 (Figure 7). There is a breakpoint in exon 7 of the NXF1 gene, which joins the segment formed by the break in exon 1 of the PSMC3 gene. This gene shows in-frame fusion and retained protein domains.

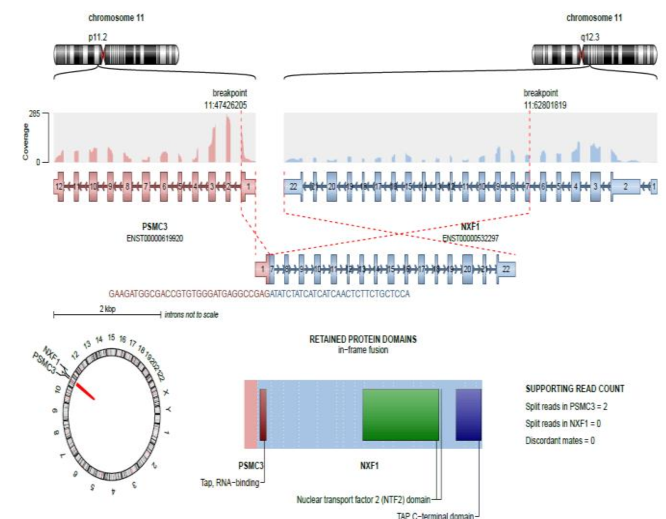

Figure 7. PSMC3 - NXF1 fusion gene.

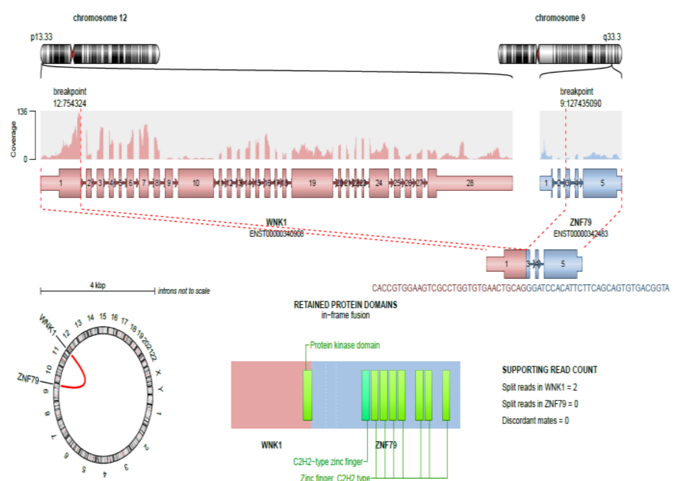

Figure 8. WNK1 - ZNF79 fusion gene.

The formation of the WNK1 - ZNF79 fusion gene occurs on chromosome 12 (Figure 8). The gene has a breakpoint in exon 1. This fragment joins the one formed by a breakpoint at the level of exon 3 of the ZNF79 gene on chromosome 9. The fusion gene formed has domains of protein retained and in-frame fusion.

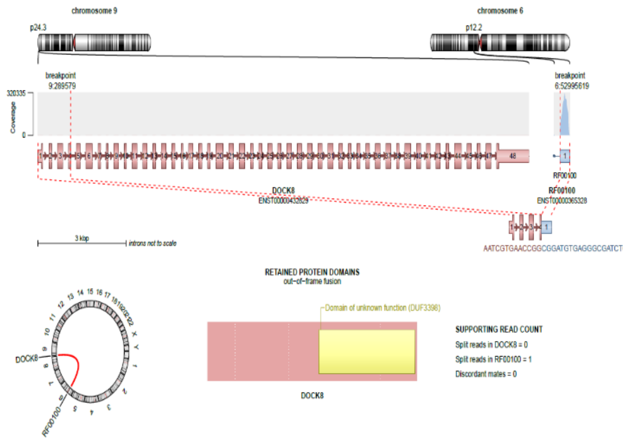

Figure 9. DOCK8 - RF00100 fusion gene.

The DOCK8 - RF00100 fusion gene is formed by the junction of exon 1 of the RF00100 gene and the segment derived by a break in exon 4 of the DOCK8 protein (Figure 9). This junction shows retained protein domains and out-of-frame fusion.

Sample 4

We reported four fusion genes in this sample. The sample was obtained by excision of a third-degree nasal pterygium in the left eye of a 39-year-old male patient who works as a teacher.

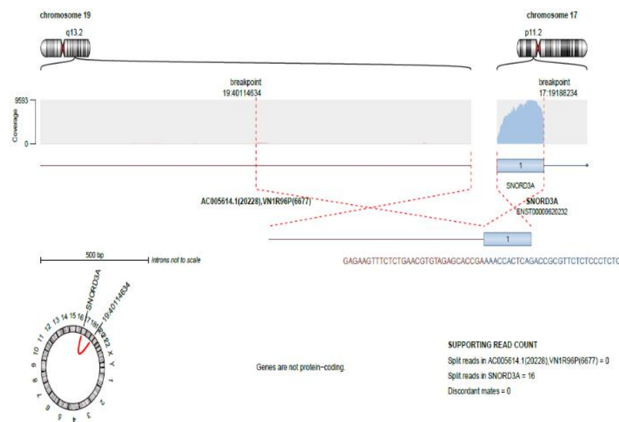

Figure 10. SNORD3A – 19:40114634 fusion gene.

The SNORD3A–19:40114634 fusion gene is formed by the union of exon 1 of SNORD3A on chromosome 17 and proteins located in sub-band 2, band 3, region 1 of the long arm of chromosome 19. The SNORD3A-19:40114634 gene does not code for proteins (Figure 10).

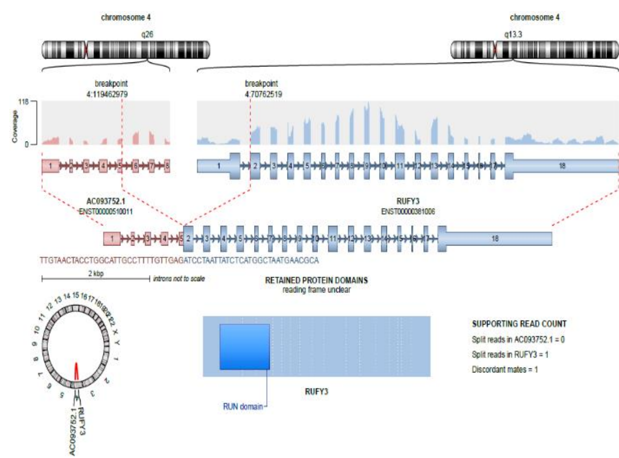

Figure 11. AC093752.1 - RUFY3 fusion gene.

The AC093752.1-RUFY3 fusion gene is formed on chromosome 4 in the AC093752.1 gene, which presented a breakpoint in exon 5. This fragment joined the fragment formed by the breakage in exon 2 of the RUFY3 gene. This structural variation features retained protein domains, and the reading frame is unclear (Figure 11).

The AC104151.1-CNTNAP4 fusion gene is formed on chromosome 16. In exon 3 of the AC104151.1 gene, there is a breakpoint that allows the joining of the fragment formed by the break in exon 18 of the CNTNAP4 gene (Figure 12). This fusion gene has retained protein domains, and the reading frame is unclear.

Finally, the AC104151.1-CNTNAP4 fusion gene is formed on chromosome 16. In the AC104151.1 gene, there is a breakpoint at the level of exon 2 that joins the fragment formed by the break in exon 18 of the CNTNAP4 gene (Figure 13). This fusion gene has retained protein domains, and the reading frame is unclear.

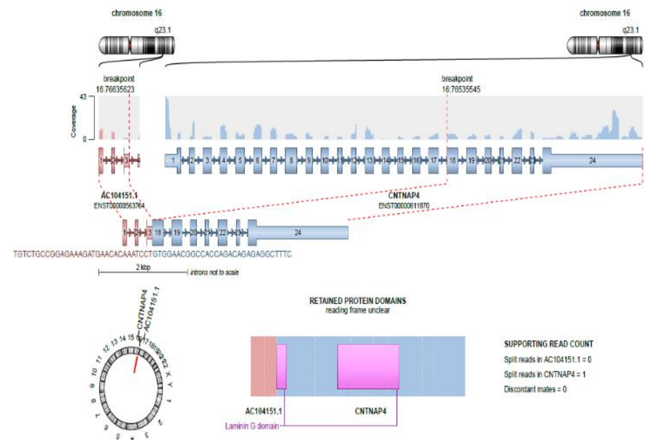

Figure 12. AC104151.1 - CNTNAP4 fusion gene.

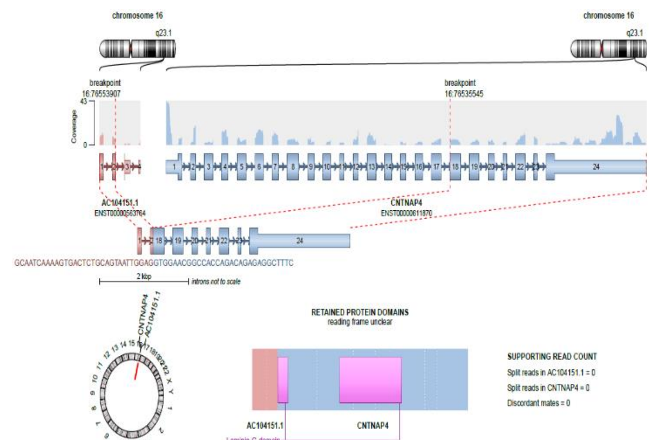

Figure 13. AC104151.1 - CNTNAP4 fusion gene.

4. Discussion

In the four processed samples, we found fusion genes, so there are structural variations in the development of the pterygium.2 The fusion genes reported in this study we do not find in the Mitelman database of chromosomal aberrations and fusion genes in cancer,8-14,11 means that we have discovered new fusion genes.

For sample 1, the fusion gene CLDND1 - KIAA2026

does not have protein domains. Instead, for sample 2, the RNV1-2 - ATP13A2 and RNU1-2 - SDHB fusion genes do not present coding regions for antisense transcription, and the WIZ - WHRN fusion gene has kept protein domains and fusion out of frame. In sample 3, the fusion genes 19:40114634 - SNORD3A, 6:164999470 - RNA5SP226 do not code for proteins, and DOCK8 - RF00100 have kept protein domains and out-of-frame fusion. In sample 4, the fusion gene 19:40114634 - SNORD3A does not code for proteins. All these results show that these fusion genes are not functional.9,16-22

In sample 3, the PSMC3-NXF1 and WNK1-ZNF79 fusion genes have retained protein domains and in-frame fusions. This means that they synthesize proteins that can act as powerful drivers in the genesis of tumors.13 In sample 4, the AC093752.1-RUFY3 and AC104151.1-CNTNAP4 fusion genes have retained protein domains, and the reading frame is unclear.

The described fusion genes have not been reported in leukemias, lymphomas, solid neoplasms, and benign tumors. This represents an essential finding of fusion genes in pterygium. These results are not conclusive; however, they represent a challenge for future research.1,23

Fusion genes have been linked to tumors for over three decades. Recent advances in sequencing technology have shown that gene fusions are much more common than previously assumed. This is corroborated by the presence of fusion genes in the pterygium samples. Over the last three years, over 9,000 novel fusion genes have been identified, and the cellular consequences remain unclear. Most fusion genes found in our samples are probably chance events, as they do not code for proteins. However, some fusion genes found did show protein-encoding. This allows us to conclude that they have the potential to activate, reduce, or eliminate the original functions of the proteins that manifest in the pathological growth of the pterygium.

The discovery of fusion genes in pterygium will provide further insight into molecular changes in proliferative tissue. This will allow us to improve the approach to diagnosis, prognosis, and therapy in patients with pterygium. The functional importance of these findings will promote the investigation of appropriate targeted therapies, paving the way toward individualized medicine.9

FUNDING: this article was financed by internal funds for Research Universidad Catolica Santa María, RESOLUTION 27144-R-2020. Project: "Isolation and culture of stem cells from the most prevalent cancers in Arequipa for their phenotypic characterization and gene expression profile by means of "next generation sequencing".

Acknowledgments: The study was conducted at the Laboratory of Research the Postgraduate school (Catholic University of Santa Maria, Arequipa, Peru).

Conflcts of Interest: The authors declare no conflct of interest.

References

1. Milman T, Ida C, Zhang PJL, Eagle RC. Gene fusions in ocular adnexal tumors. Am. J. Ophthalmol. 2021; 221: 211 -25.
2. Wang Q, Xia J, Jia P, Pao W, Zhao Z. Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. Brief. Bioinformatics. 2013; 14(4): 506-19.
3. Dai X, Theobard R, Cheng H, Xing M, Zhang J. Fusion genes: a promising tool combating against cancer. Reviews on Cancer. 2018; 1869(2): 149- 60.
4. Jiang B, Wang Y, Zhu Z, Wu Z. Expression of specificity protein 1 and collagen I in primary pterygial tissues. Arquivos Brasileiros de Oftalmologia. 2019; 82(5): p. 407-411.
5. García K, Romero M, Rodríguez M, Tenorio G. Correlación morfológica del pterigión y su evolución clínica. Revista Médica del Hospital General de México. 2006; 69(4): p. 205-211.
6. Wanzeler A, Barbosa I, Duarte B, Borges D, Barbosa E, Kamiji D, et al. Mechanisms and biomarker candidates in pterygium development. Arquivos Brasileiros de Oftalmologia. 2019; 82(6): p. 528-536.
7. Lovino M, Urgese G, Macii E, Di Cataldo S, Ficarra E. A deep learning approach to the screening of oncogenic gene fusions in humans. Int. J. Mol. Sci. 2019; 20(7): 1645.
8. Kim P, Zhou X. FusionGDB: fusion gene annotation DataBase. Nucleic Acids Res. 2019; 47(D1): D994-D1004.
9. Taniue K, Akimitsu N. Fusion genes and RNAs in cancer development. Noncoding RNA. 2021; 7(1): 10.
10. Brastianos PK, Ippen FM, Hafeez U, Gan HK. Emerging gene fusion drivers in primary and metastatic central nervous system malignancies: a review of available evidence for systemic targeted therapies. The Oncologist. 2018; 23(9): 1063-75.
11. Parker BC, Zhang W. Fusion genes in solid tumors: an emerging target for cancer diagnosis and treatment. Chin J Cancer. 2013; 32(11): 594-603.
12. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. Nat Commun. 2014; 5(4846).
13. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. Nat Rev Cancer. 2015; 15(6): 371-81.
14. Vega D. Secuenciamiento masivo (RNA-seq) y bioinformática del transcriptoma de tejido graso de Gallus gallus procedentes de centros de venta de Lima [Tesis Pregrado]. Lima: Facultad de Farmacia y Bioquímica, Universidad Nacional Mayor de San Marcos; 2019.
15. Uhrig S, Ellermann J, Walther T, Burkhardt P, Fröhlich M, Hutter B, Toprak UH, Neumann O, Stenzinger A, Scholl C, Fröhling S, Brors B. Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res. 2021 Mar;31(3):448-460.

16. Findley MK, Koval M. Regulation and roles for claudin-family tight junction proteins. IUBMB Life. 2009; 61(4): 431-7.
17. Matsuoka H, Shima A, Uda A, Ezaki H, Michihara A. The retinoic acid receptor-related orphan receptor α positively regulates tight junction protein claudin domain-containing 1 mRNA expression in human brain endothelial cells. J Biochem. 2017; 161(5): 441 -50.
18. Achari C, Winslow S, Larsson C. Down regulation of CLDND1 induces apoptosis in breast cancer cells. PLoS One. 2015; 10(6): e0130300.
19. Hofvander J, Tayebwa J, Nilsson J, Magnusson L, Brosjö O, Larsson O, et al. RNA sequencing of sarcomas with simple karyotypes: identification and enrichment of fusion transcripts. Lab Invest. 2015; 95(6): 603-9.
20. Humphray SJ, Oliver K, Hunt AR, Plumb RW, Loveland JE, Howe KL, et al. DNA sequence and analysis of human chromosome 9. Nature. 2004; 429(6990): 369-74.
21. Norppa AJ, Frilander MJ. The integrity of the U12 snRNA 3' stem-loop is necessary for its overall stability. Nucleic Acids Res. 2021; 49(5): 2835-47.
22. Kim P, Yiya K, Zhou X. FGviewer: an online visualization tool for functional features of human fusion genes. Nucleic Acids Res. 2020; 48(W1): W313- W320.
23. Hyrcza MD, Andreasen S, Melchior LC, Tucker T, Heegaard S, White VA. Primary secretory carcinoma of the lacrimal gland: report of a new entity. Am. J. Ophthalmol. 2018; 193: 178-83.