Targeted next-generation sequencing of glandular odontogenic cyst: a preliminary study

Elisa Carvalho de Siqueira, DDS, MSc, a Silvia Ferreira de Sousa, DDS, MSc, PhD, b Josiane Alves França, BMsc, MSc, c Marina Gonçalves Diniz, BS, MSc, PhD, c Thaís dos Santos Fontes Pereira, DDS, MSc, PhD, d Rennan Garcia Moreira, BS, MSc, d Pablo Agustin Vargas, DDS, PhD, f Ricardo Santiago Gomez, DDS, MSc, PhD, f and Carolina Cavaliere Gomes, DDS, MSc, PhD f

Objective. Glandular odontogenic cyst (GOC) is an uncommon developmental cyst. Its molecular pathogenesis is unclear, and deep sequencing may help identify causative low-frequency variants in tumors. We investigated in GOC mutations in 50 genes commonly altered in human cancers.

Study Design. Targeted next-generation sequencing was used to interrogate a panel of approximately 2800 mutations in GOC.

Results. Six missense single nucleotide variations (SNVs) were reported. Three SNVs (TP53 rs1042522, KDR rs1870377, and KIT rs38222214) are listed as “common single-nucleotide polymorphisms” at the UCSC Genome Browser. The other SNVs (PIK3CA p.Glu689Lys, PIK3CA p.Ala708Thr, and TP53 p.Leu289Phe) are predicted to have deleterious or damaging effects on proteins, but they showed very low frequency in our samples and could not be further validated by orthogonal methods.

Conclusions. No pathogenic SNV was detected in this cohort of GOCs. Further studies with larger gene panels or whole exome sequencing are needed to find the genetic basis of GOC. (Oral Surg Oral Med Oral Pathol Oral Radiol 2017;124:490–494)

Glandular odontogenic cyst (GOC) is an uncommon developmental cyst, which accounts for less than 0.5% of all odontogenic cysts. 1 It was first reported in 1987, and in 1988, the term glandular odontogenic cyst was suggested for the first time. 2,3 GOC occurs more frequently in the mandible, and it has a marked propensity to recur. 4,5 GOC shows epithelial features with glandular differentiation and is thought to arise from the remnants of the dental lamina. 6 Specific microscopic criteria for the diagnosis of GOC have been reported. 6 GOC shares microscopic similarities with intraosseous mucoepidermoid carcinoma, making the diagnosis by incisional biopsy challenging. Currently, these 2 entities are regarded as different lesions, considering the differences in cytokeratin expression profiles, as well as the lack of the MAML2 gene rearrangements in GOC. 7,8 Other molecular markers, such as ki67 and p53 immunohistochemical positivity, have been tested as diagnostic tools to distinguish low-grade mucoepidermoid carcinoma and GOC, but this approach did not prove to be very effective. 9

Molecular alterations in GOC are rarely investigated, and its molecular pathogenesis remains unclear. No PTCH gene mutation has been found in GOC, in contrast to odontogenic keratocyst. 10 In this context, next-generation sequencing (NGS) is a valuable tool to simultaneously interrogate several genetic mutations in rare samples, and we used this technology to investigate hundreds of cancer hotspot mutations in GOC. We used a panel to target cancer hotspot mutations, including 2856 previously described somatic mutations in 50 oncogenes and tumor suppressor genes.

MATERIALS AND METHODS

Samples selection and DNA isolation

A convenience sample of 7 formalin-fixed, paraffin-embedded (FFPE) specimens of GOC, with sufficient lesional material, were retrieved from the files of the oral pathology service of the authors’ institutions. All the samples fulfilled the criteria for GOC diagnosis. 6 The study was approved by the university ethics committee.

Statement of Clinical Relevance

A panel of 50 oncogenes and tumor suppressor genes in glandular odontogenic cysts were subjected to deep sequencing, revealing six missense single nucleotide variants. None of these variants is a known pathogenic variant.
Table I. Target genes included in the Ion AmpliSeq Cancer Hotspot Panel v2

<table>
<thead>
<tr>
<th>Gene</th>
<th>Target</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL1</td>
<td>EGRF</td>
<td></td>
</tr>
<tr>
<td>AKT1</td>
<td>ERBB2</td>
<td></td>
</tr>
<tr>
<td>ALK</td>
<td>ERBB4</td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>EZH2</td>
<td></td>
</tr>
<tr>
<td>ATM</td>
<td>FBXW7</td>
<td></td>
</tr>
<tr>
<td>BRAF</td>
<td>FGF1R</td>
<td></td>
</tr>
<tr>
<td>CDH1</td>
<td>FGF2R</td>
<td></td>
</tr>
<tr>
<td>CDKN2A</td>
<td>FGF3R</td>
<td></td>
</tr>
<tr>
<td>CSF1R</td>
<td>FLT3</td>
<td></td>
</tr>
<tr>
<td>CTNNB1</td>
<td>GNA11</td>
<td></td>
</tr>
</tbody>
</table>

DNA was extracted by using QIAamp DNA FFPE Tissue Kit (Qiagen Inc., Valencia, CA), according to the manufacturer instructions. Measurement of nucleic acid purity was performed by using NanoDrop 2000 (ThermoFisher Scientific Inc., Waltham, MA), and quantification of DNA was carried out by using Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA). DNA from one sample did not reach the quality needed for DNA library preparation and was discarded.

Next-generation sequencing

DNA libraries from 6 GOC were constructed. Library preparation was carried out by using the Ion AmpliSeq Cancer Hotspot Panel v2 (Life Technologies, Carlsbad, CA) (Table I) on the Ion Torrent PGM system (Life Technologies, Carlsbad, CA). To prepare barcode libraries, the Ion AmpliSeq Library kit 2.0 (Life Technologies, Carlsbad, CA) and Ion Xpress Barcode Adapters 1-16 Kit (Life Technologies, Carlsbad, CA) were used. The barcoded libraries were purified with Agencourt AMPure XP Kit (Beckman Coulter, Indianapolis, IN) and quantified with Ion Library TaqMan Quantitation Kit (Life Technologies, Carlsbad, CA) on the StepOne Plus system (Applied Biosystem, Foster City, CA). After library quantification, 2 samples did not pass the quality control and were excluded. The remaining samples underwent template preparation with Ion OneTouch 2 System and Ion One Touch ES (Life Technologies, Carlsbad, CA). The sequencing of 4 good-quality samples libraries was performed on Ion 316 chip (Life Technologies, Carlsbad, CA).

The sequences were aligned to human genome reference (hg19) and filtered by type of variant (single nucleotide variant [SNV], multiple nucleotide variant, indel, and longdel) and variant effects, which include non–frame shift insertion, frame shift insertion, non–frame shift deletion, frame shift deletion, stoploss, missense, non–frame shift block substitution, frame shift block substitution, nonsense, and unknown by using Ion Reporter Software 5.2 (Thermo Fisher Scientific, Waltham, MA). Reported variants are those in high-depth regions (greater than ×100) and with frequency greater than 5%. The software Integrative Genomics Viewer version 2.3 (Broad Institute, Cambridge, MA) was used to exclude false variants.

In silico analysis was carried out to predict the likely pathogenicity of missense variants, investigating the degree of conservation at the affected residue, by using both SIFT and PolyPhen-2 (accessed at http://sift.jcvi.org and http://genetics.bwh.harvard.edu/pph2/), respectively. Variant elected to be candidates for Sanger validation were established as those reported at ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) as pathogenic variants or those classified as damaging and nontolerated by SIFT and PolyPhen-2 scores.

RESULTS

Three samples sequenced were from male patients and 1 from a female patient (25%), with a mean age of 54.25 ± 9.54 years (range 42-65 years). Although 2 lesions occurred in the mandible, 1 was located in the maxilla, and location information was not available for the other case. Representative photomicrographs of a case included in the study are shown in Figure 1. Sequencing of all 4 samples resulted in a high number of sequenced bases (ranging from 28,700,000 to 70,850,000), in most part being sequenced at Q20. Samples were sequenced at a hight depth, ranging from 265,000 to 650,000 reads per sample.

Filtered variants are shown in Table II. After excluding false variants by the inspection of each variant with IVG, we identified 6 missense SNVs, which occurred at a depth greater than ×100 in the samples: TP53 p.Pro72Arg, KDR p.Gln472His, PIK3CA p.Glu689Lys, PIK3CA p.Ala708Thr, KIT p.Met541Leu, and TP53 p.Leu289Phe. Three of these SNVs—PIK3CA p.Glu689Lys, PIK3CA p.Ala708Thr, and TP53 p.Leu289Phe—occurred at a frequency less than 5%, below the detection limit of this assay. None of the variants filtered at frequency greater than 5% were reported as pathogenic or predicted to be deleterious or damaging. The 3 SNVs that occurred at depth greater than ×100 and a frequency greater than 5% (TP53 p.Pro72Arg, KDR p.Gln472His, and KIT p.Met541Leu) are reported in The UCSC Genome Browser (https://genome.ucsc.edu) as “common single nucleotide polymorphisms (SNPs).”

DISCUSSION

The molecular basis of GOC pathogenesis remains unknown, and molecular studies of this entity are scarce. PTCH1 mutation was absent in 1 GOC sample investigated, and contrarily to mucoepidermoid carcinomas, GOC lacks the MAML2 gene rearrangements. We used NGS to interrogate in GOC mutations in 50 different tumor suppressor and oncogenes that are...
commonly mutated in several types of human cancers. We started our study with 7 samples. However, only 4 cases were investigated, as 3 failed in quality-control steps. NGS allows for simultaneous sequencing of multiple targets, which is an enormous advantage when studying rare lesional types with small tissue samples, even working with FFPE material. Another advantage of this technology is that deep sequencing allows for identification of low-frequency mutations, which occur either in heterogeneous tumor cells subpopulations or when there is background contamination by adjacent normal tissue.

We detected missense SNVs in all 4 sequenced samples. However, the 3 SNVs that are not listed as common SNPs in the UCSC database, including 2 PIK3CA variants predicted to be damaging or probably damaging by Polyphen2, did not reach the 5% frequency set as the sensitivity detection limit of the assay. Therefore, we could not validate the presence of this mutation by orthogonal methods, such as Sanger sequencing, which have detection sensitivity that is estimated to be 10% to 20%. None of the missense SNVs read at high depth in a frequency greater than 5% is pathogenic, according to ClinVar, all of these being tolerated at the protein level according to in silico predictions. Although NGS is a highly sensitive method, and we confirmed the presence of a lesion in each sample included in the study, the possibility of normal cell contamination or the issue of intra-tumor heterogeneity cannot be excluded. Furthermore, the SNVs with frequency less than 5% that showed either G>A or C>T substitutions, which are nucleotide changes considered signatures of DNA damage resulting from formalin fixation, supporting the hypothesis of these alterations being a sequencing artefacts.  

The other 3 SNVs we detected (TP53 p.Pro72Arg, KDR p.Gln472His, and KIT p.Met541Leu) are reported in the UCSC Genome Browser as “common SNPs.” As we could not evaluate these SNVs in matched normal tissue, we could not confirm if they were somatic or germline mutations. However, all of them were detected at a frequency around either 50% or 100%, consistent with heterozygous or homozygous SNPs. Although our study design did not allow for inferring any participation of these SNVs in the context of GOC tumorigenesis, we could exclude such a possibility. Recently, large-scale SNP analyses have identified genetic loci associated with cancer susceptibility as well as with disease prognosis and drug response. Interestingly, TP53 p.Pro72Arg is one of the most well-studied SNPs, as it is implicated in differentiated apoptosis capacity as well as in drug response.

The NGS cancer gene targeted panel used in the present study includes 2856 hotspot mutations in 50 genes. Targeted panels offer a cheap, reliable, fast, and easy approach to detection of cancer. Although this is the largest gene panel investigated in GOCs, the absence of pathogenic mutations does not exclude genetic mutations in other genes or other genetic alterations, such as gene fusions, copy number alteration, and so on. The rate of newly discovered gene characterizations has been increasing rapidly in the last few years. In this context, whole exome sequencing offers wide coding genome coverage. However, the limitations of whole exome sequencing are its high cost and difficulties in interpretation and reporting, making this approach highly complex and less frequently used.

In line with our negative results for pathogenic mutations in GOC, which is a benign lesion, it is known that even aggressive cancers may show a stable genome, with low mutation rates. For example, pediatric rhabdoid tumors, which are highly aggressive types of cancer, were shown to be driven by mutations in genes related to the chromatin remodeling complex, showing a very low mutation rate. The notion that mutations in the chromatin remodeling complex can drive carcinogenesis has now gained force, and we cannot exclude this possibility for benign lesions, such as GOC.

Fig. 1. Photomicrograph of a glandular odontogenic cyst (GOC) sample. A, Histologically, the cyst is characterized by an epithelial lining of variable thickness with cuboidal or columnar superficial cells, and clear cells located in the basal and parabasal layers. B, The epithelium revealed surface eosinophilic cells exhibiting apocrine snouting, intraepithelial microcyst formation and mucous cells. H&E staining. (A, ×200 magnification; B, ×400 magnification).
Table II. Characterization of the missense single-nucleotide variant (SNV) detected in the glandular odontogenic cyst (GOC) samples

<table>
<thead>
<tr>
<th>Sample/Locus</th>
<th>Genes</th>
<th>Transcript</th>
<th>Coding</th>
<th>Amino acid change</th>
<th>SIFT</th>
<th>PolyPhen-2</th>
<th>Minor allele frequency (MAF)*</th>
<th>UCSC common single-nucleotide polymorphisms (SNPs)</th>
<th>ClinVar</th>
<th>Coverage (%)</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOC1 chr17:7579472</td>
<td>TP53</td>
<td>NM_000546.5</td>
<td>c.215 C &gt; G</td>
<td>p.Pro72Arg</td>
<td>Tolerated</td>
<td>Benign</td>
<td>rs1042522</td>
<td>0.4571</td>
<td>Yes</td>
<td>Benign/Uncertain significance</td>
<td>639</td>
</tr>
<tr>
<td>GOC2 chr4:55972974, chr17:7579472</td>
<td>KDR</td>
<td>NM_002253.2</td>
<td>c.1416 A &gt; T</td>
<td>p.Gln472His</td>
<td>Tolerated</td>
<td>Benign</td>
<td>rs1870377</td>
<td>0.2119</td>
<td>Yes</td>
<td>Not provided</td>
<td>1506</td>
</tr>
<tr>
<td>GOC3 chr3:178938823, chr3:178938880, chr17:7579472</td>
<td>PIK3CA</td>
<td>NM_006218.2</td>
<td>c.2122 G &gt; A</td>
<td>p.Ala708Thr</td>
<td>Tolerated</td>
<td>Benign</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
<td>Not provided</td>
<td>1992</td>
</tr>
<tr>
<td>GOC4 chr4:55593464, chr17:7577073</td>
<td>KIT</td>
<td>NM_000222.2</td>
<td>c.1621 A &gt; C</td>
<td>p.Met541Leu</td>
<td>Tolerated</td>
<td>Benign</td>
<td>rs3822214</td>
<td>0.0645</td>
<td>Yes</td>
<td>Benign/Likely benign</td>
<td>1817</td>
</tr>
<tr>
<td>GOC4 chr4:55972974, chr17:7577073</td>
<td>KDR</td>
<td>NM_002253.2</td>
<td>c.865 C &gt; T</td>
<td>p.Leu289Phe</td>
<td>Tolerated</td>
<td>Benign</td>
<td>NA</td>
<td>NA</td>
<td>No</td>
<td>Not provided</td>
<td>482</td>
</tr>
</tbody>
</table>

NA, not available.

†Frequency of this single-nucleotide variant (SNV) was <5%.
CONCLUSIONS

Our study has shown the absence of genetic variants with pathogenic effect in GOC. Further studies, including a larger cohort of samples and interrogating mutations in a larger gene panel or even in the whole exome, may help clarify the molecular pathogenesis of such lesions.

We thank the Genomics Multi-user Laboratory (Centro de Laboratórios Multiusuários, ICB/Universidade Federal de Minas Gerais) for providing support on the next-generation sequencing.

REFERENCES