




## Article

# Mutational Signature and Integrative Genomic Analysis of Human Papillomavirus-Associated Penile Squamous Cell Carcinomas from Latin American Patients

Luisa Matos do Canto <sup>1,†</sup> , Jenilson Mota da Silva <sup>2,3,†</sup>, Patrícia Valéria Castelo-Branco <sup>3</sup>, Ingrid Monteiro da Silva <sup>3</sup>, Jenilson Mota da Silva <sup>2,3,†</sup>, Leudivan Nogueira <sup>4</sup>, Carlos Eduardo Fonseca-Alves <sup>5</sup> , André Khayat <sup>6</sup>, Alexander Birbrair <sup>7,8,9</sup>  and Silma Regina Pereira <sup>3,\*</sup>

- <sup>1</sup> Clinical Genetics Department, University Hospital of Southern Denmark, 7100 Vejle, Denmark; luisaaamatos@gmail.com
- <sup>2</sup> Postgraduate Program in Health Science, Federal University of Maranhão, São Luís 65080-805, MA, Brazil; jenilson.motadasilva@gmail.com
- <sup>3</sup> Laboratory of Genetics and Molecular Biology, Department of Biology, Federal University of Maranhão, São Luís 65080-805, MA, Brazil; patricia.valeria@ufma.br (J.V.C.-S.); ingrid.monteiro@discente.ufma.br (I.M.d.S.)
- <sup>4</sup> Aldenora Bello Cancer Hospital, São Luís 65031-630, MA, Brazil; leudivan.nogueira@gmail.com
- <sup>5</sup> Health Science Institute, Paulista University, Bauru 186085-500, SP, Brazil; carloseduardofa@hotmail.com
- <sup>6</sup> Oncology Research Center, Federal University of Pará, Belém 66073-005, PA, Brazil; andresk@ufpa.br
- <sup>7</sup> Department of Dermatology, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53706, USA; birbrair@wisc.edu
- <sup>8</sup> Department of Pathology, Federal University of Minas Gerais, Belo Horizonte 31270-901, MG, Brazil
- <sup>9</sup> Department of Radiology, Columbia University Medical Center, New York, NY 10032, USA
- \* Correspondence: silma.pereira@ufma.br; Tel.: +55-98-32728543
- † These authors contributed equally to this paper.



**Citation:** Canto, L.M.; da Silva, J.M.; Castelo-Branco, P.V.; da Silva, I.M.; Nogueira, L.; Fonseca-Alves, C.E.; Khayat, A.; Birbrair, A.; Pereira, S.R. Mutational Signature and Integrative Genomic Analysis of Human Papillomavirus-Associated Penile Squamous Cell Carcinomas from Latin American Patients. *Cancers* **2022**, *14*, 3514. <https://doi.org/10.3390/cancers14143514>

Academic Editor: Kevin Gaston

Received: 1 June 2022

Accepted: 27 June 2022

Published: 20 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** DNA sequencing has been crucial to comprehending cancer mutational patterns, leading to the identification of driver genes and altered signaling pathways. Thus, identifying new pathogenic variants and their impact on tumor onset, progression, and treatment response has fueled tumor biology research. Here, we present novel findings addressing the first whole-exome sequencing (WES) of human papillomavirus (HPV)-associated penile squamous cell carcinoma (PSCC) from Latin Americans and its association with pathogenesis. We also compared the molecular profile of the tumors to that of three previous studies from populations with different genetic and socioeconomic backgrounds, the majority of which was HPV-negative. We describe the most altered genes and the main pathogenic variants found in the Latin Americans, ten of which are exclusive to our study sample. The data allowed us to identify molecular pathways and druggable targets with potential treatment value for this still-neglected HPV-associated carcinoma.

**Abstract:** High-throughput DNA sequencing has allowed for the identification of genomic alterations and their impact on tumor development, progression, and therapeutic responses. In PSCC, for which the incidence has progressively increased worldwide, there are still limited data on the molecular mechanisms involved in the disease pathogenesis. In this study, we characterized the mutational signature of 30 human papillomavirus (HPV)-associated PSCC cases from Latin Americans, using whole-exome sequencing. Copy number variations (CNVs) were also identified and compared to previous array-generated data. Enrichment analyses were performed to reveal disrupted pathways and to identify alterations mapped to HPV integration sites (HPV<sub>is</sub>) and miRNA–mRNA hybridization regions. Among the most frequently mutated genes were *NOTCH1*, *TERT*, *TTN*, *FAT1*, *TP53*, *CDKN2A*, *RYR2*, *CASP8*, *FBXW7*, *HMCN2*, and *ITGA8*. Of note, 92% of these altered genes were localized at HPV<sub>is</sub>. We also found mutations in ten novel genes (*KMT2C*, *SMARCA4*, *PTPRB*, *AJUBA*, *CRI1*, *KMT2D*, *NBEA*, *FAM135B*, *GTF2I*, and *CIC*), thus increasing our understanding of the potential HPV-disrupted pathways. Therefore, our study reveals innovative targets with potential therapeutic benefits for HPV-associated PSCCs. The CNV analysis by sequencing (CNV-seq) revealed five cancer-associated genes as the most frequent with gains (*NOTCH1*, *MYC*, *NUMA1*, *PLAG1*, and

*RAD21*), while 30% of the tumors showed *SMARCA4* with loss. Additionally, four cancer-associated genes (*CARD11*, *CSMD3*, *KDR*, and *TLX3*) carried untranslated regions (UTRs) variants, which may impact gene regulation by affecting the miRNAs hybridization regions. Altogether, these data contribute to the characterization of the mutational spectrum and its impact on cellular signaling pathways in PSCC, thus reinforcing the pivotal role of HPV infection in the molecular pathogenesis of these tumors.

**Keywords:** whole-exome sequencing; HPV; penis; carcinoma; biomarkers

## 1. Introduction

Several genes are frequently mutated in various cancers, yet the altered gene spectra vary greatly [1]. The advancement of high-throughput DNA sequencing has been crucial to comprehending cancer mutational patterns, leading to the identification of driver genes and altered signaling pathways. This knowledge enables the association of mutational profiles with specific tumors and their clinical–histopathological parameters. Thus, identifying new pathogenic variants and their impact on tumor onset, progression, and treatment response has fueled tumor biology research [2]. Here, our goal was to characterize the mutational profile of HPV-positive PSCC and its association with pathogenesis.

Penile cancer (PeCa) is one of the most neglected malignancies; it is described as a rare cancer in high-income nations, but it has higher incidences in developing countries. This can be partly attributed to social and economic issues, such as poor genital hygiene, lack of childhood circumcision, phimosis, and chronic inflammation. In addition, human papillomavirus (HPV) infection, often linked to unprotected sex with multiple partners, has been identified as a risk factor for penile carcinogenesis [3]. In some regions of Brazil, PeCa may account for up to 10% of male cancers [4–6]. Notably, the PeCa incidence is progressively increasing even in developed countries, along with HPV infections and oropharynx cancer [7–9].

Our previous studies corroborate the role of HPV in PeCa, showing a high frequency of cytobands with copy number alterations (CNAs) harboring HPVis, as well as the impact of miRNA/mRNA changes on HPV-related molecular pathways [10–12]. However, the molecular characterization of these tumors, especially upon HPV infection, is still limited. Furthermore, patients' treatments disregard HPV status, though infection contributes to PeCa classification [13]. Surgical resection and systemic chemotherapy are the primary treatment approaches. Nevertheless, they result in poor outcomes, low survival, and a high risk of comorbidities and psychosocial issues [3]. Information about the somatic mutation landscape of PeCa is limited to patients from Europe, Asia, and North America [14–16], who are mostly HPV-negative. Further investigation is needed to unravel HPV-associated molecular alterations and identify potential targets for more efficient personalized treatments.

In the present work, we report unique data on the mutational signature of HPV-associated PSCC from Latin American patients. We described novel genes and pathogenic variants that may contribute to a better understanding of the potential HPV-disrupted pathways. This study reveals innovative targets with potential therapeutic benefit for this carcinoma.

## 2. Materials and Methods

Supplementary Figure S1 shows the flowchart representing the study design.

### 2.1. Study Population

Thirty patients with PSCC underwent chemotherapy-naive surgical excision at the Aldenora Bello Cancer Hospital (São Luís, Maranhão, Brazil). All patients submitted written informed consent, which was approved by the Research Ethics Committee on Humans from the Federal University of Maranhão and by the National Research Ethics

Commission (CONEP, Brazil; CAAE: 46371515.5.0000.5087). Clinical and histopathological variables were extracted from the patients' medical records.

At diagnosis, the mean age of the patients was  $60.4 \pm 16.4$  years old, ranging from 31 to 85 years old. Most patients presented a low level of education (83.3% had no studies or incomplete primary education, and only one had an undergraduate degree), 43.3% self-declared themselves to be a current smoker or declared to have been a smoker in the past 5 years, and 36.7% declared that they drink alcohol regularly. Additionally, 62.5% of the patients presented phimosis. All tumors were SCC, presenting an average size of  $4.8 \pm 2.7$  cm. The tumors were localized in the glans, corpus cavernosum, and corpus spongiosum in 20% of the cases. The most frequent histological subtypes were condylomatous and usual, which were present in 46.7% and 40% of the cases, respectively. Tumor grades G2 and G3 were the most frequent, present in 66.7% and 26.7% of the patients, respectively. Partial and total penectomy were performed in 63% and 30% of the patients, respectively, while lymphatic and perineural invasion were performed in 18.5% and 33.3%, respectively. Table 1 and Supplementary Table S1 present the detailed patients' clinical–histopathological information.

**Table 1.** Clinical and histopathological profile of 30 PSCC patients analyzed by WES, qPCR (CNA and gene expression), and aCGH (n = 18).

Parameters	WES and qPCR (%)	aCGH (%)	Parameters	WES and qPCR (%)	aCGH (%)
Histologic site			Histologic type		
Glans	10 (33.3)	5 (27.8)	Usual	12 (40.0)	7 (38.8)
Glans + foreskin	8 (26.7)	8 (44.4)	Condylomatous	14 (46.6)	9 (50.0)
Foreskin	2 (6.7)	2 (11.1)	Mixed	2 (6.7)	1 (5.6)
Glans and other areas	10 (33.3)	3 (16.7)	Basaloid	2 (6.7)	1 (5.6)
Predominant lesion			Surgical procedure:		
Ulcerated	11 (36.7)	6 (33.3)	Preserved penis	1 (3.3)	1 (5.6)
Vegetative	10 (33.3)	6 (33.3)	Amputation (partial)	17 (56.7)	12 (66.6)
Ulcer/vegetative	3 (10.0)	2 (11.1)	Amputation (total)	9 (30.0)	5 (27.8)
Verrucous	2 (6.7)	2 (11.1)	Glandectomy (total)	2 (6.7)	0 (0.0)
Nodular ulcerated	3 (10.0)	2 (11.1)	No information	1 (3.3)	0 (0.0)
No information	1 (3.3)	0			
Lymphatic invasion			Perineural invasion		
Yes	5 (16.7)	3 (16.7)	Yes	9 (30.0)	5 (27.8)
No	22 (73.3)	13 (72.2)	No	18 (60.0)	11 (61.1)
No information	3 (10.0)	2 (11.1)	No information	3 (10.0)	2 (11.1)
Tumor Grade:			Tumor size (mm):		
G1	2 (6.7)	1 (5.6)	0.6–2.0	2 (6.7)	0
G2	20 (66.7)	12 (66.6)	2.1–5.0	17 (56.6)	13 (72.2)
G3	8 (26.6)	5 (27.8)	5.1–10.0	11 (36.7)	5 (27.8)
Primary tumor (T)			Infection HPV		
pT1	5 (16.7)	3 (16.7)	Positive	30 (100.0)	18 (100.0)
pT2	11 (36.7)	7 (38.9)	Negative	0	0
pT3	13 (43.3)	8 (44.4)			
No information	1 (3.33)	0			
Phimosis			Marital status		
Yes	15 (50.0)	7 (38.9)	Married	20 (66.7)	15 (83.3)
No	9 (30.0)	6 (33.3)	Single	7 (23.3)	3 (16.7)
No information	6 (20.0)	5 (27.8)	No information	3 (10.0)	0

PSCC, penile squamous cell carcinoma; CNA, copy number alteration; aCGH, array Comparative Genomic Hybridization; WES, whole-exome sequencing.

## 2.2. HPV Genotyping

We performed HPV genotyping via nested-PCR according to BioGenetics Molecular Technologies (Uberlândia, Minas Gerais, Brazil, patent number BR102017004615.0), which allowed us to identify 40 viral subtypes classified as high risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 73, and 82), intermediate risk (MM7, MM8, 26, 30, 34, 53, 54, 55, 61,

62, 64, 67, 69, 70, 71, 72, 74, and 81), and low risk (6, 11, 42, 43, 44, and 57). In addition, we performed DNA sequencing according to our previous protocol [10].

### 2.3. Whole-Exome Sequencing

Genomic DNA was isolated from 30 surgically fresh tumor tissues that were macro-dissected by an experienced pathologist and presented at least 80% of tumor cells. Twenty of them were paired with nontumor tissue located roughly two centimeters away from the primary tumor.

DNA from tumor and nontumor tissues was extracted by using DNeasy<sup>®</sup> Tissue Kit (QIAGEN Group, Germantown, MD, USA), according to the manufacturer's protocol. DNA library was prepared by using 100 ng input gDNA of each tumor and their matched adjacent nontumor samples. DNA was fragmented by using a Covaris ME220 ultrasonicator and quantified on Qubit (Invitrogen, Waltham, MA, USA) according to the manufacturer's instructions. Library preparation was performed by using the Illumina<sup>®</sup> TruSeq<sup>®</sup> Exome Kit (Illumina, San Diego, CA, USA) and IDT custom probe enrichment (Integrated DNA Technologies, Coralville, IA, USA), according to the manufacturer's guidelines. The size, quantification, and quality control of the DNA library were evaluated on TapeStation (Agilent Technologies, Santa Clara, CA, USA). The captured DNA was subjected to high-performance sequencing, using Illumina NovaSeq 6000 platform to generate  $2 \times 150$ -bp paired-end reads.

### 2.4. Data Processing

To explore the spectrum of somatic variants, three pipelines were developed by VarStation Bioinformatics (<https://varsomics.com/varstation/> (last accessed on 10 May 2022)), following the steps described below. The quality control of the raw sequence reads and the presence of adapters were evaluated by FASTQC v0.11.9 (<https://qubeshub.org/resources/fastqc> (accessed on 22 March 2021)) and MultiQC v.1.10.1 [17]. Regions with low quality and containing adapters were trimmed, and the resulting sequences that were smaller than 50 base pairs were excluded by using Cutadapt v2.10 [18]. The FASTQ files were aligned to the reference human genome (hg38), using BWA-MEM v.0.7.13 [19]. The aligned BAM files were ordinated according to genomic annotation, PCR duplicates were removed, and the outputs were realigned by using the Indel Realignment tool of the Genome Analysis Toolkit (GATK) [20]. The recalibration was carried out by using the BaseRecalibrator tool from GATK. Moreover, the GATK Crosscheck Fingerprints tool was used to identify possible exchange between the tumor/nontumor sample pairs. The high-performance sequencing produced  $73.78 \pm 17.4$  million and  $75.35 \pm 18.2$  million of paired-reads for the tumors and nontumor samples, respectively, both with  $99.68\% \pm 0.1\%$  of DNA mapping on target. The mean coverage was  $294\times$  and  $300\times$  for tumor and nontumor tissues, with  $93\% \pm 6\%$  and  $94\% \pm 8\%$  of target DNA covered at least  $100\times$ , respectively (Supplementary Figure S2). All sequenced samples presented  $> 95\%$  of quality score  $\geq Q30$  (mean equals to  $97\% \pm 0.5\%$  for tumor and  $96.7\% \pm 0.7\%$  for nontumor) (Supplementary Table S2).

### 2.5. Variant Calling

The somatic and germline variant calling was performed by using the best practice of the GATK4 [21]. The HaplotypeCaller [22] was used for germline variants, and the Mutect2 tool [23] was used for somatic variants.

The normal panel was built with the 20 adjacent nontumor samples, which were paired to their respective tumor samples (somatic paired samples) to identify single-nucleotide variants (SNVs) and nucleotide(s) deletion coupled with a nucleotide(s) insertion (Indels). In addition, the 10 tumor samples without corresponding nontumor tissue (Somatic single samples) followed the tumor-only pipeline. For comparison purposes, the somatic paired samples were also analyzed as tumor-only, totaling 30 samples. The comparison of VCFs files obtained from these different analyses was performed by using vcfeval [24], which revealed 82% matching between somatic paired and somatic single samples, while there

was no matching between somatic paired and germline samples, thus confirming the usefulness of the pipeline for somatic variants through 20 paired samples.

### 2.6. Variant Filtering

Each sample was evaluated individually by using the VarStation platform (<https://app.varstation.com/#/analysis> (last accessed on 10 May 2022)), following predetermined parameters: read depth  $\geq 100$ , variant allele fraction (VAF)  $\geq 0.05$ , and populational frequency  $< 0.01$  (gnomAD, 1000 Genomes and ABRAoM, the Brazilian genomic variants database (<https://abraom.ib.usp.br> (accessed on 22 March 2021))). We also performed a visual analysis by using the Integrative Genomics Viewer (IGV, Broad Institute and UC San Diego) to filter out variants present in problematic regions of the genome [25]. The variant calling covered SNVs and Indels at exonic, intronic, or intergenic regions, such as splicing, frameshift, non-frameshift, non-coding RNA (ncRNA), UTR, stop-gain, stop-loss, start-gain, and start-loss.

Both synonymous and nonsynonymous variants that were predicted to be benign or probably benign were filtered out by the predictor's algorithms, including SIFT, LRT, Revel, MutationTaster, MutationAssessor, Fathmm-XF, Fathmm-MKL Vest, MetaSVM, Polyphen2, Polyphen2 HDIV, HVAR, CADD, DANN, MetaLR, MetaSVM, and Provean. Clinical information was extracted from ClinVar [26], and data on cancer-related mutations were obtained from the Catalogue of Somatic Mutations in Cancer (COSMIC) at <https://cancer.sanger.ac.uk/cosmic> (accessed on 22 March 2021) and OncoKB (<https://www.oncob.org> (accessed on 22 March 2021)). The variants obtained were annotated using by ANNOVAR [27].

Following the analysis of the 20 paired tumors, only the altered genes occurring in at least 15% of the cases were further analyzed in the 10 single tumors. In addition, we looked at the variants, especially the ones classified as pathogenic or likely pathogenic according to ClinVar (<http://www.ncbi.nlm.nih.gov/clinvar> (accessed on 22 March 2021)), following the somatic variant interpretation/reporting by the American College of Medical Genetics and Genomics (ACMG), Association for Molecular Pathology (AMP), and College of American Pathologists (CAP) recommendations [28,29]. All variants were searched in the COSMIC v9, released 28 May 2021 (<https://cancer.sanger.ac.uk/cosmic> (accessed on 1 June 2021)), and in dbSNP (<https://www.ncbi.nlm.nih.gov/snp/> (accessed on 1 June 2021)). Our data were compared to variants described for penile tumors from European [14], Asian [15], and North American [16] populations containing mostly HPV-negative.

### 2.7. Copy Number Variation by Sequencing (CNV-seq)

The 20 paired samples and the 30 single tumor samples (20 paired plus 10 unpaired) were also used for calling up somatic CNVs, using high-throughput sequencing. We followed the Best Practices guide GATK4, using the somatic pipeline for CNV-seq [21]. The normal panel was built with the same samples as for SNV/Indels. Variants were annotated with Functotator v1.6.20190124 s and AnnotSV v2.5 [30]. We considered cutoffs for calling gain and deletion as  $CN > 3$  and  $CN < 1$ , respectively. In addition, the specified regions should have over 10 Kb and should be present in at least 10% of samples [31].

### 2.8. UTR-Perturbing Variants

We used the miRabel platform [32] to identify miRNAs that regulate genes with variants in UTRs. Prediction tools available in miRabel, such as MiRanda, Pita, SVmicro, and TargetScan, were used. Only miRNA–mRNA interactions with a score  $\leq 0.05$  were considered. The miRNA–mRNA interactions identified in miRabel were then used in the STarMir program [33] to identify the seed regions. The construction of the design and nucleic acid fold (STarMir) is obtained by using the Mfold package [34] and Sfold containing the Srna module [35]. Sfold applies a two-step model for hybridization between mRNA and miRNA. In this model, the hybridization of the miRNA–target occurs at an accessible target site, and then the hybrid elongates to form the complete miRNA–target



duplex. The minimum free energy of hybridization was obtained from the RNAhybrid tool [36,37]. Only interactions with LogitProb values  $\geq 0.5$  were considered.

### 2.9. Integrated Genomic and Molecular Characterization of HPV-Associated PSCC

The thirty sequenced tumor samples were analyzed by qPCR for alterations in gene copy number and/or gene expression, while 18 tumors were evaluated for CNA by aCGH [10]. The data were cross-referenced to produce an overview of the main genes with SNV/Indels, CNV-seq, CNA, and gene expression changes (GECs) in HPV-associated PSCC in the Latin Americans.

### 2.10. Pathway Analysis

The Integrative Onco Genomics (IntOGen) platform [38] was used to identify variants of cancer-related genes in both paired and unpaired tumors, which were confirmed by using COSMIC [39]. The STRING v.11.5 database [40] was used to predict protein–protein interactions (PPI) based on direct (physical) and indirect (functional) associations (score  $> 0.9$ ). The KEGG pathways (Kyoto Encyclopedia of Genes and Genomes) were used for functional analysis through the enrichment of STRING clusters (FDR  $< 0.05$ , with  $p$ -values corrected by the Bonferroni method), which generate sets of functionally associated proteins at various levels of hierarchy, ranging from small groups of 5 proteins to large sets of up to 200 proteins [41].

## 3. Results

### 3.1. High Frequency of HPV Infection in PSCC

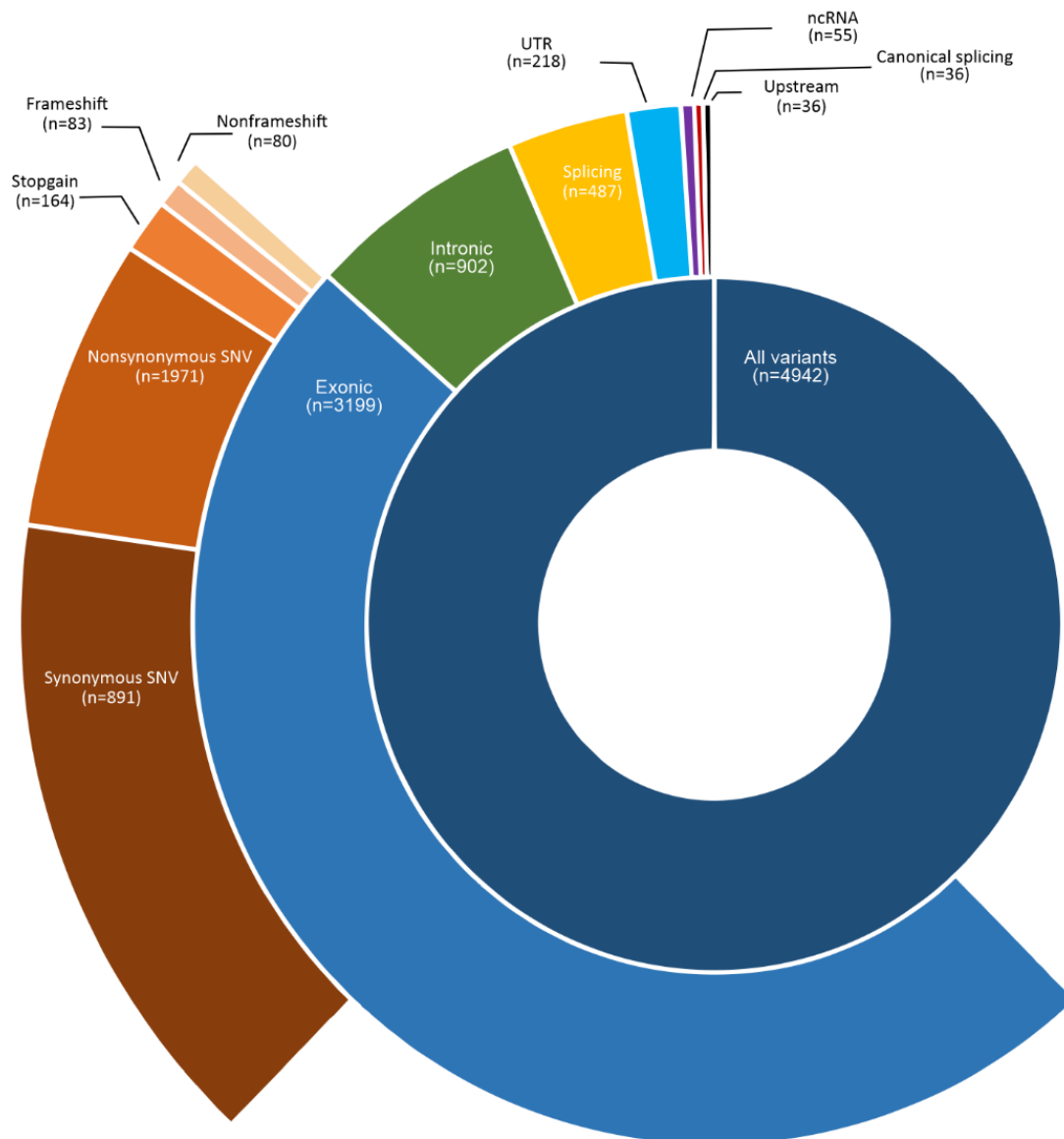
All patients ( $n = 30$ ) were tested for HPV infection by nested-PCR and direct DNA sequencing, revealing 100% of positivity for the virus (Table 1). Genotyping was possible in all but five samples, which yielded insufficient DNA. High-risk HPV subtypes were detected in 96% of the tumors, of which 29.2% presented multiple infections. The HPV16 was the most frequent (75%).

### 3.2. Somatic Variants in HPV-Associated PSCC in Latin Americans

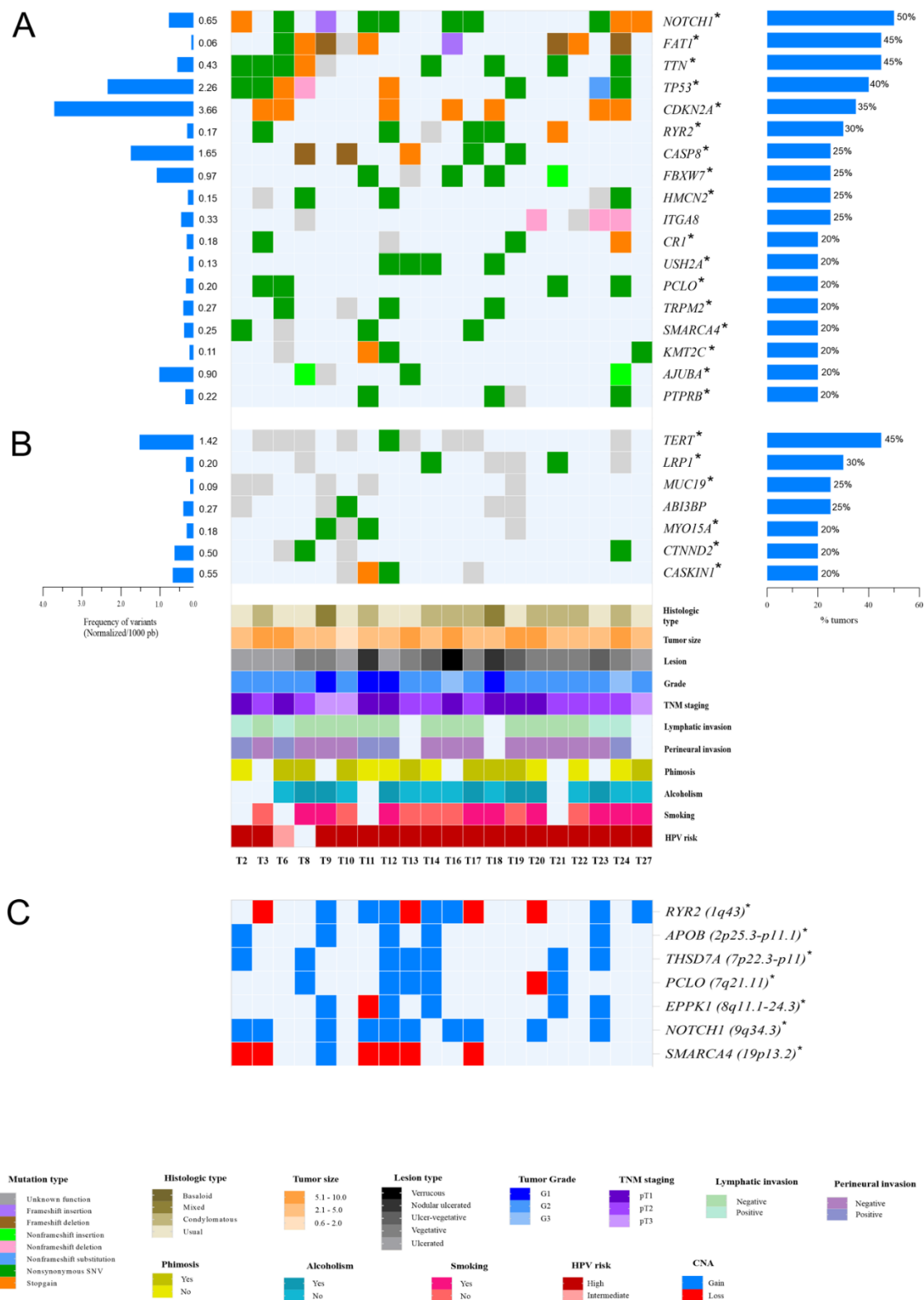
The WES of 30 PSCC and 20 matched adjacent normal penile tissues produced 149.14 million reads with high-quality control (Supplementary Figure S2). Firstly, the variant calling was performed in the 20 paired tumors. Considering the parameters described in the variant calling (see Materials and Methods), we identified 4942 variants affecting 3277 genes, most of them (64.73%) at exonic regions, of which 96% were SNV and 4% Indels (Figure 1). The most frequent alteration was nonsynonymous SNV (61.6%) due to transitions C  $>$  T (28.3%) and G  $>$  A (27.2%), while start loss ( $n = 8$ ) and stop-loss ( $n = 2$ ) ones were the less frequent. We also identified variants in intronic region (18.25%), splicing sites (9.85%), UTR (4.4%), and ncRNA (1%). Less frequent variants were in upstream ( $n = 36$ ), intergenic ( $n = 7$ ), and downstream ( $n = 2$ ) regions. As expected, the largest human chromosome (chr1) presented most of the variants (11%), followed by chromosomes 19 and 2 (7% each). Chromosome Y had only two variants.

Among 3277 genes presenting rare variants, 160 have been reported as cancer-associated genes according to Integrative Onco Genomic (<https://www.intogen.org/> (accessed on 14 October 2021) [38], as shown in Supplementary Table S3). The number of cancer-associated genes per tumor was  $12 \pm 7$ . Supplementary Table S4 shows the top mutated genes ( $n = 84$ ) found in at least 15% of the tumors, and their normalized score according to the gene size (bp), of which 18 (21.43%) are cancer-associated genes (*NOTCH*, *FAT1*, *TP53*, *CDKN2A*, *FBXW7*, *CASP8*, *SMARCA4*, *PTPRB*, *KMT2C*, *CR1*, *AJUBA*, *KMT2D*, *MTOR*, *NBEA*, *LRP1B*, *FAM135B*, *GTF2I*, and *CIC*). Figure 2A presents the most frequent ( $>20\%$  of the tumors) genes with exonic variants ( $n = 18$ ). *NOTCH1* was the most frequent (50% of the tumors), followed by *TTN* and *FAT1* (45% each); *TP53* (40%); *CDKN2A* (35%); *RYR2* (30%); *CASP8*, *FBXW7*, *HMCN2*, and *ITGA8* (25% each); *CR1*, *USH2A*, *PCLO*, *TRPM2*, *SAMARCA4*, *KMT2C*, *AJUBA*, and *PTPRB* (20% each). Considering the number of variants according to the gene size, *CDKN2A*, *TP53*, *CASP8*, *FBXW7*, and *AJUBA* are the hotspots. Among the

most frequent genes presenting non-exonic alterations ( $n = 7$ ), we observed notable somatic variants (45% of the tumors) in the *TERT* gene (Figure 2B). We highlight that 92% (23/25) of the most frequent mutant genes carry variants at HPVis.



**Figure 1.** Somatic variants identified in human papillomavirus-associated penile squamous cell carcinomas from Latin American patients. The full circle in darker blue represents the total number of variants identified in 3872 genes. The semicircle in the middle represents the number of variants identified according to the gene's location: exonic (2654 genes), intronic (850 genes), splicing (466 genes), UTR (214 genes), ncRNA (48 genes), canonical splicing (35 genes), and upstream (28 genes). The outer semicircle shows the types of variants identified in the exonic region: synonymous SNV (853 genes), non-synonymous SNV (1742 genes), stop-gain (149 genes), frameshift (73 genes), and non-frameshift (77 genes).



**Figure 2.** Top genes and variants identified in at least 20% of human papillomavirus-associated penile squamous cell carcinomas according to clinical and histopathological variables. The charts in (A) (exonic region), (B) (intronic region), and (C) (CNV-seq) present the mutational profile of the 20 paired tumors. The matrix at the bottom of (B) shows the clinical–histopathological profile and the HPV status of each case (identified by the samples ID). Horizontal bar plot on the right shows the frequency of each altered gene in the 20 paired tumors, and on the left is the number of variants normalized according to the gene size (1000 bp). \* Genes with variants localized at HPV integration sites; CNV-seq, copy number variation by sequencing; CNA, copy number alterations.



For the most frequently altered genes, the variants occur in a single tumor, such as *NOTCH1* and *TP53*. We highlight genes having repeated variants, such as *CDKN2A*. This gene presented only pathogenic variants, all of them stop-gain, of which c.238C > T (p.Arg80Ter) and c.172C > T (p.Arg58Ter) were observed in 15% and 10% of tumors, respectively, in addition to c.330G > A and c.329G > A (p.Trp110Ter). The *FBXW7* gene presented the variant c.1273A > G (p.Arg425Gly) in 10% of the tumors. Furthermore, some variants were found in both groups of tumors, paired and single, such as c.330G > A (*CDKN2A*), c.4389C > A (*NOTCH1*), and c.743G > A (*TP53*), each present in 6.7% of the tumors. Among the 229 variants detected in the main altered genes, including those found exclusively in Latin Americans, 123 (57.7%) are not described in COSMIC v9 or in the dbSNP (Supplementary Table S5).

We also highlight the occurrence of 10 (50%) variants in *TERT* (telomerase catalytic subunit reverse transcriptase), of which one occurs in exonic region (c.2713G > A; p.Val905Met) and nine (45% of tumors) in the upstream region of the gene. All *TERT* promoter (*TERT*p) variants were G > A transitions, with 66.7% of them being located at −146bp (rs1561215364). This variant was also found in 50% of the single tumors. The second most frequent (10% of paired tumors) was at −124bp (rs1242535815).

### 3.3. Novel Genes Altered Exclusively in PSCC from Latin Americans

We compared the top altered genes in Latin Americans (all of them HPV-positive) with those described in three other populations (European, Asian, and North American) (in which most of the tumors are HPV-negative) (Supplementary Table S6). These populations present some commonly altered genes (Figure 3), of which only two, *NOTCH1* and *CASP8*, were found in the four populations. Latin Americans, North Americans, and Europeans share an alteration in *CDKN2A* and *FBXW7*, while Latin Americans, North Americans, and Asians have in common changes in *TP53*, *FAT1*, and *TTN* genes. The highest frequency of exclusive altered genes occurred in the Latin American population (88%), followed by North American (63%), European (60%), and Asian (18%). Among the exclusive genes in the Latin American population, we identified 10 genes related to cancer as novel altered genes potentially associated with HPV-associated PSCC: *KMT2C*, *SMARCA4*, *PTPRB*, *AJUBA*, *CR1* (in 20% of the tumors, each), *KMT2D*, *CIC*, *NBEA*, *FAM135B*, and *GTF2I* (in 15% of the tumors, each). Table 2 shows the variants detected in cancer-associated genes found exclusively in Latin Americans. Interestingly, all loci are in HPVis, according to HPVbase (<http://crdd.osdd.net/servers/hpvbase/index.html> (accessed on 17 March 2022)) [42].

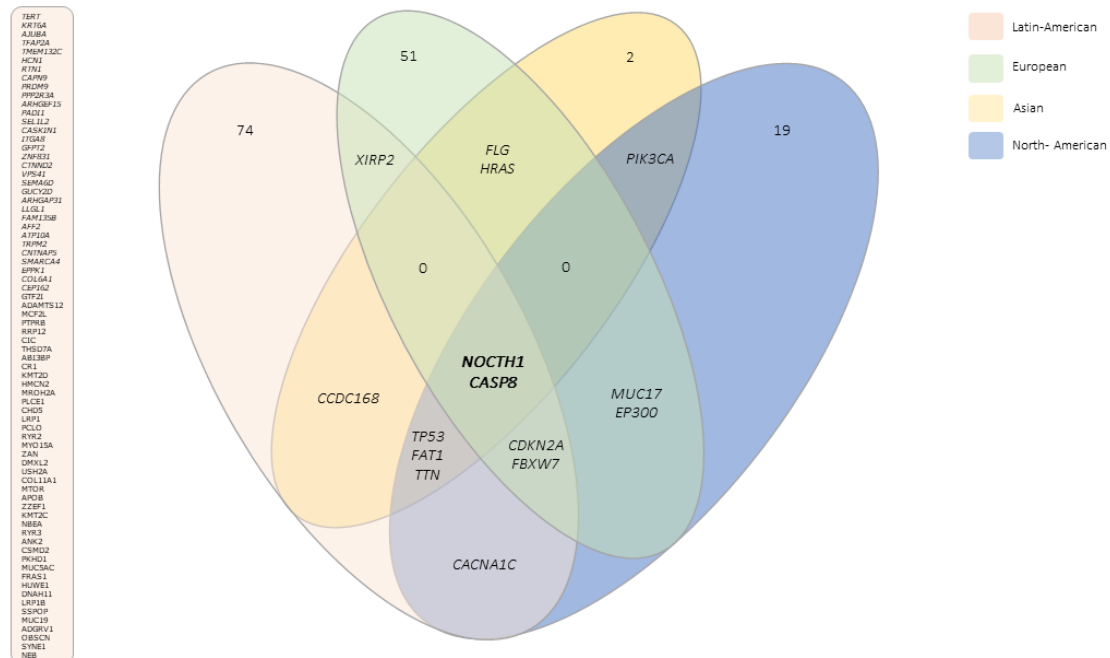
### 3.4. Copy Number Variation in HPV-Associated PSCC

The CNV-seq analyses revealed 558 genes presenting  $\geq 3$  copies, of which 42 are associated with cancers. *NOTCH1*, *PTCH1*, *TSC1*, *NUP214*, *SET*, *ABL1*, *TNC*, *FANCC*, *PLAG1*, *MYC*, *RAD21*, and *NUMA1* were detected in 10% of the tumors. When comparing the CNV-seq data with our previous data on CNAs detected by aCGH [10], we found 80 genes in common presenting gene copy gain. Among them, five cancer-associated genes were confirmed by both analyses as the most frequent genes with gains: *NOTCH1*, *MYC*, *NUMA1*, *PLAG1*, and *RAD21*. Deletions were less frequent in both aCGH and CNV-seq analysis; however, we highlight that 30% of the tumors showed *SMARCA4* loss (Figure 2C). Table 3 presents the main altered cytobands and the most frequent genes with gain, all of them localized at HPVis for subtype 16.

### 3.5. UTRs Variants and Their Biological Consequences in miRNA-mRNA Interactions

Among all altered genes, 5.5% have variants in their UTRs (Figure 1). To investigate the biological consequences of these variants, we identified miRNAs that may regulate these genes, as well as their seed regions (miRNA  $\times$  mRNA interacting regions) that may be impacted by the variants. As a result, 14 genes were found to be regulated by miRNAs (miRabel score  $\leq 0.05$ ), four of which are cancer-associated genes (*CARD11*, *CSMD3*, *KDR*, and *TLX3*). The variants are located within the gene hybridization sites (Supplementary

Table S7). The *CSMD3* gene (regulated by miR-1237-3p, miR-224-5p, and miR-132-3p) presented the highest number of hybridization regions affected by the c.-19T > C variant (sites for miR-1237-5p). In addition, the ‘c.-19T > C’ variant has not yet been described in the dbSNP and COSMIC databases.



**Figure 3.** Common genes described in four studies in which whole-exome sequencing was performed in penile carcinomas, including the present study. Data from Europeans, Asians, and North Americans were adapted from [14–16], respectively. Left side of the figure: genes observed exclusively in penile squamous cell carcinomas of Latin American patients.

**Table 2.** Cancer-associated genes presenting variants detected by whole-exome sequencing in human-papillomavirus—associated penile squamous cell carcinomas exclusively altered in Latin American patients.

Gene	Cytoband	Integration Sites (HPV Subtypes) *	Variants			
<i>KMT2C</i>	7q36.1	yes (16, 18)	c.11261C > T	c.4844G > A	c.162-17T > G	c.9059C > G
<i>KMT2D</i>	12q13.12	yes (16, 18)	c.511C > T c.9614G > C	c.9232G > C c.9186G > C	c.8914G > A c.12133C > T	
<i>SMARCA4</i>	19p13.2	yes (16, 18)	c.3575G > A	c.1249C > T	c.3694G > A	chr19:g.11003471C > T
<i>CIC</i>	19q13.2	yes (16, 18)	c.1264C > T	chr19:g.42274618G > A	chr19:g.42287248CC > TT;p.?	
<i>PTPRB</i>	12q15	yes (18)	c.2886G > C	c.979G > C	c.2357C > G	chr12:g.70609388C > T
<i>NBEA</i>	13q13.3	yes (16, 26) **	c.6226C > T c.1572T > G	c.4644 + 5G > A	c.6060G > A	c.5463C > A
<i>FAM135B</i>	8q24.23	yes (16, 18, 45)	c.2974T > A	chr8:g.138132825A > C		
<i>GTF2I</i>	7q11.23	yes (16)	c.178G > C	c.20C > G	c.233A > G	chr7:g.74589997C > A
<i>AJUBA</i>	14q11.2	yes (16, 18)	c.1006G > C	chr14:g.22978448GAGAA > -p.?	c.762_763insGA	c.445dupG
<i>CR1</i>	1q32.2	yes (16, 18)	c.4067G > C	c.68G > A	c.5383G > T	chr1:g.207588649A > T

\* Data were obtained from HPV base (<http://crdd.osdd.net/servers/hpvbase/index.html>, last accessed on 17 March 2022) [42]; \*\* Data were obtained from Virus Integration Site DataBase (VISDB) (<https://bioinfo.uth.edu/VISDB/index.php/homepage>, last accessed on 17 March 2022) [43].

**Table 3.** Cancer-related genes with gain ( $\geq 3$  copies) detected in 10% of human-papillomavirus-associated penile squamous cell carcinomas from Latin American patients. The specified cytobands have over 10 Kb.

Gene	Start	End	Cytoband	HPV Integration Site (Subtypes) *
NOTCH1 †	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	136407694	136518542	9q34.3–9q34.3	yes (16, 18)
PTCH1	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	84866993	98578322	9q21.33–9q22.31	yes (16, 18)
TSC1	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	131060785	133209449	9q34.12–9q34.2	yes (16, 18)
NUP214	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	131060785	133209449	9q34.12–9q34.2	yes (16, 18)
SET	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	128220134	128822687	9q34.11–9q34.11	yes (16, 18)
ABL1	129873232	131057397	9q34.11–9q34.12	yes (16, 18)
	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
TNC †	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	114792157	120990030	9q32–9q33.1	yes (16, 18)
FANCC †	89311043	138219684	9q22.2–9q34.2	yes (16, 18)
	84866993	98578322	9q21.33–9q22.31	yes (16, 18)
PLAG1 †	54136261	89971540	8q21.12–8q13.2	yes (16, 18, 68)
	56072840	143934540	8q21.12–8q24.22	yes (16, 18, 68)
MYC †	117528401	133257113	8q24.12–8q24.22	yes (16, 18)
	56072840	143934540	8q21.12–8q24.22	yes (16, 18, 68)
RAD21	109399826	117521303	8q23.2–8q24.11	yes (16, 18)
	56072840	143934540	8q21.12–8q24.22	yes (16, 18, 68)
NUMA1	70078357	73083990	11q13.3–11q13.4	yes (16)
	71463803	72023397	11q13.4–11q13.4	yes (16)

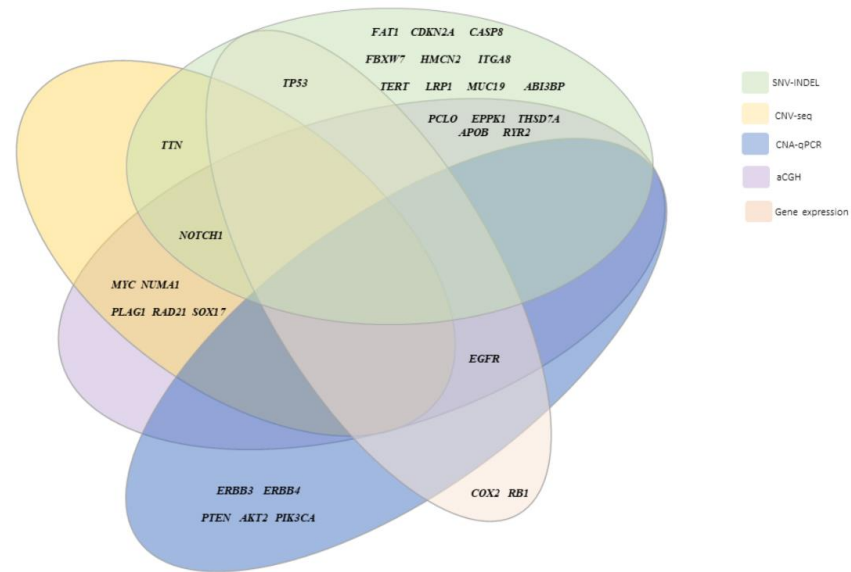
† HPV target gene; \* data obtained from HPVbase (<http://crdd.osdd.net/servers/hpvbase/index.html> (accessed on 17 March 2022)) [42] and Virus Integration Site DataBase [43].

### 3.6. Spectrum of Somatic Variants in HPV-Associated PSCC

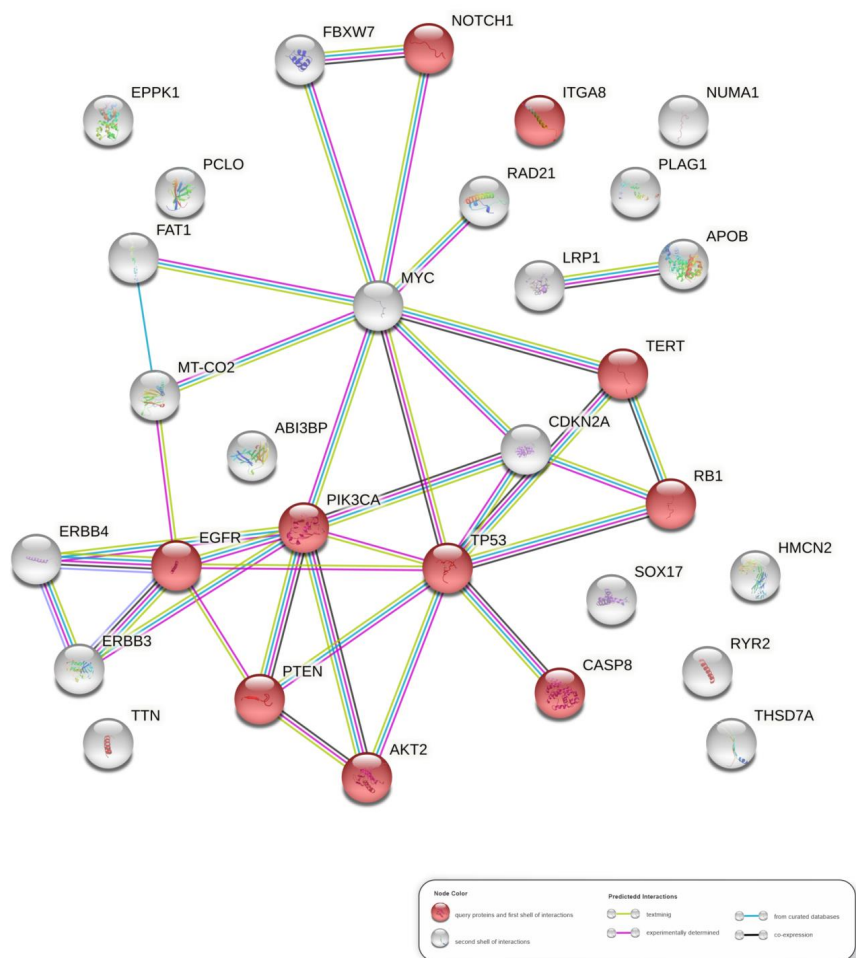
Finally, to generate a spectrum of somatic variants in HPV-associated PSCC from Latin Americans, we crossed WES data with genomic alterations identified in our previous study [10], in which tumors were characterized for GEC and CNA. The profile of the most frequently altered genes in both studies is represented in Figure 4.

### 3.7. Molecular Pathways in HPV-Associated PSCC

We predicted the main pathways (highest confidence) by using the most frequently altered genes in the HPV-associated PSCC (Figure 4). Supplementary Table S8 lists the top 20 KEGG pathways identified and ranked by FDR value. The HPV infection pathway (hsa05165) is one of the top five, with *PIK3CA*, *RBI*, *TP53*, *EGFR*, *NOTCH1*, *TERT*, *CASP8*, *PTEN*, *ITGA8*, and *AKT2* supporting it. The network of gene interactions supporting the top pathways for HPV-associated PSCC in Latin Americans is visualized in Figure 5.



**Figure 4.** The spectrum of main genes exhibiting somatic variants in human papillomavirus-associated penile squamous cell carcinomas in Latin American patients.



**Figure 5.** Protein-protein interaction network of the main genes presenting somatic variants in human papillomavirus-associated penile squamous cell carcinomas from Latin Americans. The red nodes represent genes supporting the human papillomavirus infection pathway.

#### 4. Discussion

Our study revealed the first mutational PSCC signature of Latin Americans, for whom we performed a high-throughput WES. Our study population is from the state of Maranhão in Northeast Brazil, where HPV prevalence has been reported to be among the highest in the world [5]. We also compared our findings with previous WES studies of penile tumors, in which mutational landscapes were described in Europeans [14], Asians [15], and North Americans [16], all of which had a similar sample size ( $30 \pm 3.5$ ). This comparison is particularly interesting because, in addition to the differences in genetic backgrounds and HPV status (mostly HPV-negative in the previous studies), the clinical and social characteristics of patients from developed countries contrast with those from Brazil.

HPV is the second most prevalent cancer-causing viral agent worldwide and is related to several types of tumors, such as cervical, oropharyngeal, and anogenital, including penile [44]. However, while HPV DNA is identified in over 90% of cervical carcinoma cells [7], there is substantial variation in viral positivity among penile tumors, mainly due to differences in molecular testing. Our study population is a subset from 55 previously investigated PSCC cases [10], with the highest HPV positivity among other PeCa populations [45]. This higher HPV positivity is likely a result of two more sensitive detection/genotyping methods (DNA sequencing in addition to nested-PCR able to detect 40 HPV subtypes), improving the accuracy compared to the use of a single method. As result, in this present study, 100% of the tumors were HPV positive, of which 96% were high risk.

The high-coverage WES enabled us to detect variants in exonic and intronic regions, splicing sites, UTR, and ncRNA, in addition to the major CNV-seq. Additionally, these alterations were integrated with CNAs and gene expression variations detected in 18/30 cases [10,11]. Among the altered genes found in at least 15% of tumors, 18 (*NOTCH*, *FAT1*, *TP53*, *CDKN2A*, *FBXW7*, *CASP8*, *SMARCA4*, *PTPRB*, *KMT2C*, *CR1*, *AJUBA*, *KMT2D*, *MTOR*, *NBEA*, *LRP1B*, *FAM135B*, *GTF2I*, and *CIC*) had already been linked to different types of cancer [38]. Considering the number of variants per gene size, *CDKN2A*, *TP53*, *CASP8*, and *FBXW7* were the hotspots. Schwaederle et al. [46] proposed *CDKN2A*, *NOTCH1*, *TP53*, and *FBXW7* as “squamousness” genes in head and neck, lung, cutaneous, gastrointestinal, and gynecologic cancers. Interestingly, all of those genes were altered at a high frequency in our cases and were classified as PSCC. Therefore, our data strengthen the importance of translational research that focuses on potential molecular targets that are common to squamous cell cancers, thereby mitigating the difficulties associated with conducting effective clinical trials due to the low prevalence of PeCa worldwide.

*CDKN2A* (p16) expression is a predictive marker for HPV infection [47], and as such, practically all HPV-associated cancers overexpress p16. HPV-unrelated tumors are often more aggressive and p16-negative, which is usually inactivated by homozygous deletion and promoter hypermethylation [48,49]. In the present study, *CDKN2A* had the highest frequency (35% of cases) of tumor-shared pathogenic variants, with all being stop-gain (p.R80X, p.R58X, and p.W110X). Previously, the nonsense variant was also reported, but in only two p16-negative penile tumors [16]. Similarly, other authors attributed loss-of-function variants to explain the negativity of p16 in 5% of HPV-associated cervical cancer [50]. Hence, considering the high frequency of *CDKN2A* nonsense variants in high-risk HPV advanced PSCC, we reinforce that p16 immunoreexpression may be ineffective as an HPV marker in a significant subset of PeCa. Our results corroborate those of Zito Marino et al. [51], who highlighted the importance of using at least two methods to avoid the misclassification of HPV penile tumors in clinical practice.

By comparing our mutational signature to those reported before [14–16], we identified a common list of altered genes. Latin Americans’ signature shares gene alterations with North Americans and Europeans (*CDKN2A* and *FBXW7*) and with North Americans and Asians (*TP53*, *FAT1*, and *TTN*). These genes have been described as being involved in the establishment and progression of other tumors [52–54] and may also be potential drivers of this rare carcinoma. *NOTCH1* and *CASP8* were found to be frequently altered in all



PeCa WES studies (LA = 33% and 17%; NA = 41% and 24%, respectively; EU = 12.5% each; AS = 13.3% each). Feber's group [14] proposed two mutational signatures based on the HPV status: an HPV-associated APOBEC mutation signature and an NpCpG mutation signature in HPV-negative disease, whereas Wang and collaborators [15] identified the Notch, RTK–RAS, and *hippo* pathways as the most disrupted pathways. Chahoud et al. [16] highlighted a molecular similarity between PSCC and head and neck squamous cell carcinoma (HNSC), specifically involving the Notch pathway, corroborating our findings of *NOTCH1* variants in 50% of cases, as well as those of Wang et al. [15].

It is worth noting that whole-genome sequencing was performed on a panel of HPV-negative and epithelial-like PSCC cell lines, in which the exomic variants were highly consistent with their corresponding cancer tissues [55]. Similar to our study, *TP53*, *CDKN2A*, and *NOTCH1* were among the most commonly altered genes. Furthermore, the authors observed amplification of *MYC*, *PLAG1*, and *EGFR*, revealing MAPK, Jak-STAT, TGF- $\beta$ , Notch, and apoptosis signaling pathways as major players in the PeCa. Altogether, these results point out that SNV/Indels in these genes may be related to penile carcinogenesis regardless of HPV infection. On the other hand, we have previously suggested that downregulation of *TP53* and *RB1* mRNAs in PSCC [10] occurs through the upregulation of 13 miRNAs mapped at HPVis [12]. We also reported the impact of HPVis-localized CNAs on miRNA/mRNA interactions, which may alter critical pathways, such as P53 signaling, *hippo* and TGF- $\beta$  signaling, proteoglycans in cancer, and viral carcinogenesis [11].

Although studies have described potential targets for PeCa treatment (see [45] for a review), most of them are frequently altered in other cancers; only a few clinical trials are available, with discouraging endpoints. Furthermore, despite the association between prognosis and HPV status [47], treatment is not chosen based on HPV infection. Thus, cisplatin-based chemotherapy remains the most used treatment, although there is no consensus on a standard regimen and, unfortunately, the majority of patients has a poor response to treatment [56].

We describe ten novel PSCC-associated genes, all of which have variants found exclusively in Latin Americans (*KMT2C*, *SMARCA4*, *PTPRB*, *AJUBA*, *CR1*, *KMT2D*, *NBEA*, *FAM135B*, *GTF2I*, and *CIC*). Notably, all of these genes are at HPVis, mainly for subtype 16, which is the most prevalent in our samples. HPV integration can cause cancer by activating oncogenes or inactivating tumor suppressors [42,43], and, although the process is still unclear, HPV infection has been related to penile carcinogenesis [7,57]. Four of the novel genes were described as predictive markers and potential pharmacological targets (<https://www.dgidb.org> (accessed on 3 May 2022) [58]); *SMARCA4* (Tazemetostat, Abemaciclib, Ribociclib, Vinorelbine, Palbociclib and Cisplatin), *CIC* (Vemurafenib, Selumetinib and Trametinib), *NBEA* (Metformin), and *CR1* (Eculizumab).

In terms of targeted therapies, the mitogen-activated protein kinase (MAPK) pathway has received considerable attention; however, most MAPK inhibitors induce chemoresistance [59]. Furthermore, some of these inhibitors are highly selective, such as Vemurafenib, a drug with specificity for BRAF-V600E tumors [31], which were never reported in PeCa. On the other hand, although the *SMARCA4* loss is not directly targetable, Januario and collaborators [60] demonstrated that inhibiting histone methyltransferase EZH2 was effective in *SMARCA4*-mutated cancer cell lines derived from various tumor types. Clinical trials have also demonstrated that Tazemetostat, an oral EZH2-specific inhibitor, has a favorable safety profile and antitumor activity [61–64]. Studies have also indicated that combining tumor immunotherapy with epigenetic modifiers, particularly EZH2, could enhance treatment efficacy (see Reference [65] for a review). Additionally, given the efficacy of CDK4/6 inhibitors in certain types of cancers harboring *SMARCA4* loss [66], we highlight that the deletion and nonsense variants in *CDKN2A* (p16) were found in 30% and 35% of our tumor samples, respectively. Hence, considering that p16 normally inhibits CDK4 and CDK6 [67], the use of tyrosine kinase inhibitors in tumors with *SMARCA4* and *CDKN2A* deficiency is a rational potential therapeutic option for this set of PSCCs.

We identified noncoding variants in the *TERT* promoter (*TERT*<sub>p</sub>) in 45% of tumors, with G250A (rs1561215364) and G228A (rs1242535815) accounting for 66.7% and 22.2%, respectively. Both G > A transitions were described as the most frequent in a large core of various cancers, 73% of which overexpressed *TERT* [68]. In contrast to a recent study [69] that linked *TERT*<sub>p</sub> alterations to non-HPV-related PSCC, we emphasize that all tumors with these alterations are HPV positive in the current study. *TERT* is normally silenced in nontumor cells but is reactivated via regulatory mechanisms that maintain the telomere length in human cancer cells, allowing them to survive. Interestingly, GABP (GA-binding protein) was identified as a critical transcription factor specifically recruited for *TERT* activation via the G > A mutant promoter [70]. Hence, the high frequency of these variants in PSCC provides opportunities for the development of anticancer strategies and applications of *TERT* as a potential biomarker for the diagnosis and prognosis of neoplasms [71]. Additionally, we identified an exonic p. (Val905Met) variant and a novel upstream (*TERT*:GG > AA at 1295127-1295128) variant, making *TERT* and *NOTCH1* the most frequently altered genes in HPV-associated PSCC (50% each).

The WES also allowed us to identify CNV-seq in 42 cancer-associated genes with *cn* > 3, including *NOTCH1*, *MYC*, *NUMA1*, *PLAG1*, and *RAD21* amplified in 10% of tumors. These findings confirm our previous aCGH data, showing that all alterations are mapped at high-risk HPVis. Moreover, we were able to identify genes presenting variants at miRNA binding sites (seed bindings in UTR regions). MiRNAs normally interact with the target gene/mRNA via their seed sequence (5' end) or the untranslated region (3'UTR) of the gene [72]. Base pairing has also been reported in the 5' UTR region and in the coding region [73]. Herein, we found 14 genes with UTR variants, of which four cancer-associated genes harbor variants in 3' UTR (*KDR* and *CARD11*) and 5' UTR (*CSMD3* and *TLX3*). These genes have been described as oncogenes, coding for vascular endothelial growth factor receptor 2 [57,74,75], promoting cell survival by activating NF-kb [76,77], activating downstream MAPK/ERK signaling, inhibiting p53 and activating cMyc [78,79], and promoting cell growth in hematological neoplasms by activating LINC00478/miR-125b [80]. The UTR variants can alter the binding site of miRNAs, impairing their regulatory capacity; they may also generate a new binding site for other miRNAs, resulting in an imbalance of cellular homeostasis. A limited number of studies have described variants in UTRs in cancer [81–83], and our findings form the basis for future studies into the biological and functional impact of somatic variants in these regions on penile carcinogenesis.

Finally, the spectrum of genomic alterations obtained across an integrative analysis of the copy number and gene expression alterations in the analyzed penile tumors allowed us to predict the main pathways involved in penile carcinogenesis. As previously discussed, although other authors have proposed a link between HPV infection and penile carcinogenesis, the variety of tissue collection methods and techniques used to detect the virus allowed false-negative results. Thus, the association of specific markers or cellular pathways with viral infection still remains controversial. We used two highly sensitive molecular methods for virus detection/genotyping, which allowed us to identify the virus precisely in the tumor samples. Nevertheless, our methods did not detect active infection, which was a limitation of our study, as well as the absence of HPV-negative cases.

The enrichment analysis revealed the HPV infection pathway (hsa05165) among the top five, supported by the interaction network of the genes *PIK3CA*, *RB1*, *TP53*, *EGFR*, *NOTCH1*, *TERT*, *CASP8*, *PTEN*, *ITGA8*, and *AKT2*. This corroborates with our previous study in which HPV-mediated viral carcinogenesis was also supported for both gene and miRNAs CNAs [11]. Furthermore, most of the altered genes identified by WES, presenting SNV/Indels or CNV-seq, as well as the majority of altered cytobands identified by aCGH, lie within HPVis, reinforcing the pivotal role of HPV infection in genomic alterations through integration of the virus into the host genome.

## 5. Conclusions

The data from the first whole-exome sequencing of penile tumors from Latin Americans that are characterized by HPV infection and an advanced disease revealed that most of the alterations are located at HPVis. *NOTCH1*, *TERT*, *TTN*, *FAT1*, *TP53*, *CDKN2A*, *RYR2*, *CASP8*, *FBXW7*, *HMCN2*, and *ITGA8* were the most frequently altered genes with SNV/Indels. In addition, ten novel genes (*KMT2C*, *SMARCA4*, *PTPRB*, *AJUBA*, *CR1*, *KMT2D*, *NBEA*, *FAM135B*, *GTF2I*, and *CIC*) were found in the Latin American population only, suggesting innovative targets with a potential therapeutic benefit for HPV-associated PSCC. We highlighted potential tyrosine kinase inhibitors since a significant proportion of tumors have *SMARCA4* and *CDKN2A* deficiency. In addition, *NOTCH1*, *MYC*, *NUMA1*, *PLAG1*, and *RAD21* were found to be amplified in 10% of tumors, thus supporting the aCGH findings. We also described, for the first time, UTRs variants (*KDR*, *CARD11*, *CSMD3*, and *TLX3*), impacting the miRNAs' binding sites in penile tumors. All of these genomic alterations allowed us to predict important molecular pathways and biological markers involved in penile carcinogenesis, thereby improving our knowledge and confirming previous studies. The HPV molecular pathway was highly supported by our findings. Additionally, our data strengthen the importance of translational research that focuses on potential molecular targets common to squamous cell cancers, thereby mitigating the difficulties associated with conducting effective clinical trials due to the low prevalence of penile tumors worldwide. We further highlight the importance of collaborative research involving the integration of data from different populations, whose genetic, socioeconomic, and cultural characteristics, in addition to HPV status, must have a substantial impact on the initiation, development, and treatment response of penile cancer.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14143514/s1>. Figure S1: Flowchart representing the study design. Fresh penile squamous cell carcinomas were collected from 30 Latin American patients. Tumors were screened for human papillomavirus (HPV) infection and were genotyped in most cases (5 cases yielded insufficient DNA). Whole-exome sequencing (WES) was performed in 20 paired tumors and adjacent normal tissues and 10 single tumors. Sequencing results were used for identifying somatic alterations. Furthermore, enrichment for miRNA–mRNA interactive regions (miR–mRNAir) and HPV integration sites (HPVis) was performed. The copy number variants (CNVs) were contrasted to previous CNA results of the same cases based on aCGH, with validation of selected genes by qPCR and RT-qPCR<sup>10</sup>. The mutated genes identified in this study were compared to those described by three previous studies reporting WES data from PeCa patients from Europe<sup>14</sup>, Asia<sup>15</sup>, and North America<sup>16</sup>. Figure S2: Depth and coverage of (A) HPV-associated penile squamous cell carcinoma samples (n = 30) and (B) nontumor adjacent penile tissues (n = 20). Table S1: Clinical–histopathological characterization of the Latin Americans' patients with HPV-associated penile squamous cell carcinomas (n = 30). Table S2: Quality control of the whole-exome sequencing performed in 30 HPV-associated penile squamous cell carcinomas. Table S3: Cancer-associated genes presenting alterations in HPV-associated penile squamous cell carcinomas from Latin Americans. Table S4: Top 84 genes with somatic variants identified by WES in HPV-associated penile squamous cell carcinomas from Latin Americans, and the normalized number of variants based on gene size (bp). Cancer-associated genes are highlighted in bold. Table S5: Variants detected in the most altered genes in Latin Americans with HPV-associated penile squamous cell carcinomas. Table S6: Main altered genes described in three previous studies in which whole-exome sequencing was performed in penile squamous cell carcinomas from Europeans [14], Asians [15], and North Americans [16], compared to Latin Americans (this study). Table S7: Genes with variants in untranslated regions (UTRs), and the miRNAs that binding by seed region. Table S8: The top twenty KEGG pathways identified from a spectrum of major genes with somatic variants in HPV-associated penile squamous cell carcinomas from Latin Americans, ranked by FDR value.

**Author Contributions:** Conceptualization, S.R.P.; methodology, L.M.C., C.E.F.-A., L.N., A.K., and S.R.P.; formal analysis, S.R.P. and L.M.C.; investigation, S.R.P., L.M.C., P.V.C.-B.; I.M.S. and J.M.S.; resources, S.R.P.; writing—original draft preparation, S.R.P.; writing—review and editing, S.R.P., L.M.C., A.B., and J.M.S.; visualization, L.M.C., A.K., and A.B.; supervision, S.R.P.; project administration,

S.R.P.; funding acquisition, S.R.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA; Grant #COOPI-07895/17 and IECT-05551/18) and by Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES- code 001). “The APC was partially funded by Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) POSGRAD 02598/21”.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Research Ethics Committee on Humans from the Federal University of Maranhão and by the National Research Ethics Commission (CONEP, Brazil; CAAE: 46371515.5.0000.5087).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to patient confidentiality.

**Acknowledgments:** We are grateful to the Vejle Hospital, Institute of Regional Health Research, University of Southern Denmark, for the technical assistance with Whole-exome sequencing. The authors are also thankful to the Hospital Albert Einstein/VarStation for the Bioinformatics service. We are grateful to Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) for funding this research and to Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES-code 001) for providing scholarship to J.M.S. (Master Postgraduate Program in Health Science, Federal University of Maranhão, Brazil).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Sanchez-Vega, F.; Mina, M.; Armenia, J.; Chatila, W.K.; Luna, A.; La, K.C.; Dimitriadoy, S.; Liu, D.L.; Kantheti, H.S.; Saghafeinia, S.; et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **2018**, *173*, 321–337. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Alexandrov, L.B.; Kim, J.; Haradhvala, N.J.; Huang, M.N.; Tian Ng, A.W.; Wu, Y.; Boot, A.; Covington, K.R.; Gordenin, D.A.; Bergstrom, E.N.; et al. The Repertoire of Mutational Signatures in Human Cancer. *Nature* **2020**, *578*, 94–101. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Thomas, A.; Necchi, A.; Muneer, A.; Tobias-Machado, M.; Tran, A.T.H.; Van Rompuy, A.-S.; Spiess, P.E.; Albersen, M. Penile Cancer. *Nat. Rev. Dis. Primers* **2021**, *7*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Koifman, L.; Vides, A.J.; Koifman, N.; Carvalho, J.P.; Ornellas, A.A. Epidemiological Aspects of Penile Cancer in Rio de Janeiro: Evaluation of 230 Cases. *Int. Braz. J. Urol.* **2011**, *37*, 231–243. [\[CrossRef\]](#)
5. Coelho, R.W.P.; Pinho, J.D.; Moreno, J.S.; Garbis, D.V.E.O.; do Nascimento, A.M.T.; Larges, J.S.; Calixto, J.R.R.; Ramalho, L.N.Z.; da Silva, A.A.M.; Nogueira, L.R.; et al. Penile Cancer in Maranhão, Northeast Brazil: The Highest Incidence Globally? *BMC Urol.* **2018**, *18*, 50. [\[CrossRef\]](#)
6. Douglawi, A.; Masterson, T.A. Penile Cancer Epidemiology and Risk Factors: A Contemporary Review. *Curr. Opin. Urol.* **2019**, *29*, 145–149. [\[CrossRef\]](#)
7. Kidd, L.C.; Chaing, S.; Chipollini, J.; Giuliano, A.R.; Spiess, P.E.; Sharma, P. Relationship between Human Papillomavirus and Penile Cancer—Implications for Prevention and Treatment. *Transl. Androl. Urol.* **2017**, *6*, 791–802. [\[CrossRef\]](#)
8. Hansen, B.T.; Orumaa, M.; Lie, A.K.; Brennhovd, B.; Nygård, M. Trends in Incidence, Mortality and Survival of Penile Squamous Cell Carcinoma in Norway 1956–2015: Trends in Penile Squamous Cell Carcinoma. *Int. J. Cancer* **2018**, *142*, 1586–1593. [\[CrossRef\]](#)
9. D’Hauwers, K.W.M.; Depuydt, C.E.; Bogers, J.J.; Noel, J.C.; Delvenne, P.; Marbaix, E.; Donders, A.R.T.; Tjalma, W.A.A. Human Papillomavirus, Lichen Sclerosus and Penile Cancer: A Study in Belgium. *Vaccine* **2012**, *30*, 6573–6577. [\[CrossRef\]](#)
10. Macedo, J.; Silva, E.; Nogueira, L.; Coelho, R.; Silva, J.; Santos, A.; Teixeira-Júnior, A.A.; Belfort, M.; Silva, G.; Khayat, A.; et al. Genomic Profiling Reveals the Pivotal Role of HrHPV Driving Copy Number and Gene Expression Alterations, Including MRNA Downregulation of *TP53* and *RB1* in Penile Cancer. *Mol. Carcinog.* **2020**, *59*, 604–617. [\[CrossRef\]](#)
11. Silva, J.D.; Nogueira, L.; Coelho, R.; Deus, A.; Khayat, A.; Marchi, R.; Oliveira, E.D.; Santos, A.P.D.; Cavalli, L.; Pereira, S. HPV-Associated Penile Cancer: Impact of Copy Number Alterations in MiRNA/MRNA Interactions and Potential Druggable Targets. *CBM* **2021**, *32*, 147–160. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Silva, J.; Costa, C.C.; Farias Ramos, I.; Laus, A.C.; Sussuchi, L.; Reis, R.M.; Khayat, A.S.; Cavalli, L.R.; Pereira, S.R. Upregulated MiRNAs on the TP53 and RB1 Binding Seedless Regions in High-Risk HPV-Associated Penile Cancer. *Front. Genet.* **2022**. [\[CrossRef\]](#)



13. Moch, H.; Cubilla, A.L.; Humphrey, P.A.; Reuter, V.E.; Ulbright, T.M. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours. *Eur. Urol.* **2016**, *70*, 93–105. [[CrossRef](#)]
14. Feber, A.; Worth, D.C.; Chakravarthy, A.; de Winter, P.; Shah, K.; Arya, M.; Saqib, M.; Nigam, R.; Malone, P.R.; Tan, W.S.; et al. CSN1 Somatic Mutations in Penile Squamous Cell Carcinoma. *Cancer Res.* **2016**, *76*, 4720–4727. [[CrossRef](#)] [[PubMed](#)]
15. Wang, Y.; Wang, K.; Chen, Y.; Zhou, J.; Liang, Y.; Yang, X.; Li, X.; Cao, Y.; Wang, D.; Luo, L.; et al. Mutational Landscape of Penile Squamous Cell Carcinoma in a Chinese Population. *Int. J. Cancer* **2019**, *145*, 1280–1289. [[CrossRef](#)]
16. Chahoud, J.; Gleber-Netto, F.O.; McCormick, B.Z.; Rao, P.; Lu, X.; Guo, M.; Morgan, M.B.; Chu, R.A.; Martinez-Ferrer, M.; Eterovic, A.K.; et al. Whole-Exome Sequencing in Penile Squamous Cell Carcinoma Uncovers Novel Prognostic Categorization and Drug Targets Similar to Head and Neck Squamous Cell Carcinoma. *Clin. Cancer Res.* **2021**, *27*, 2560–2570. [[CrossRef](#)]
17. Ewels, P.; Magnusson, M.; Lundin, S.; Källner, M. MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report. *Bioinformatics* **2016**, *32*, 3047–3048. [[CrossRef](#)]
18. Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet J.* **2011**, *17*, 10. [[CrossRef](#)]
19. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
20. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernysky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)]
21. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data. *Nat. Genet.* **2011**, *43*, 491–498. [[CrossRef](#)] [[PubMed](#)]
22. Poplin, R.; Ruano-Rubio, V.; DePristo, M.A.; Fennell, T.J.; Carneiro, M.O.; Van der Auwera, G.A.; Kling, D.E.; Gauthier, L.D.; Levy-Moonshine, A.; Roazen, D.; et al. Scaling Accurate Genetic Variant Discovery to Tens of Thousands of Samples. *bioRxiv* **2017**. [[CrossRef](#)]
23. Benjamin, D.; Sato, T.; Cibulskis, K.; Getz, G.; Stewart, C.; Lichtenstein, L. Calling Somatic SNVs and Indels with Mutect2. *bioRxiv* **2019**. [[CrossRef](#)]
24. Cleary, J.G.; Braithwaite, R.; Gaastra, K.; Hilbush, B.S.; Inglis, S.; Irvine, S.A.; Jackson, A.; Littin, R.; Rathod, M.; Ware, D.; et al. Comparing Variant Call Files for Performance Benchmarking of Next-Generation Sequencing Variant Calling Pipelines. *bioRxiv* **2015**. [[CrossRef](#)]
25. Amemiya, H.M.; Kundaje, A.; Boyle, A.P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci. Rep.* **2019**, *9*, 9354. [[CrossRef](#)]
26. Landrum, M.J.; Lee, J.M.; Riley, G.R.; Jang, W.; Rubinstein, W.S.; Church, D.M.; Maglott, D.R. ClinVar: Public Archive of Relationships among Sequence Variation and Human Phenotype. *Nucleic Acids Res.* **2014**, *42*, D980–D985. [[CrossRef](#)]
27. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)]
28. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. [[CrossRef](#)]
29. Li, M.M.; Datto, M.; Duncavage, E.J.; Kulkarni, S.; Lindeman, N.I.; Roy, S.; Tsimberidou, A.M.; Vnencak-Jones, C.L.; Wolff, D.J.; Younes, A.; et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer. *J. Mol. Diagn.* **2017**, *19*, 4–23. [[CrossRef](#)]
30. Geoffroy, V.; Herenger, Y.; Kress, A.; Stoetzel, C.; Piton, A.; Dollfus, H.; Muller, J. AnnotSV: An Integrated Tool for Structural Variations Annotation. *Bioinformatics* **2018**, *34*, 3572–3574. [[CrossRef](#)]
31. Liu, J.; Yan, J.; Mao, R.; Ren, G.; Liu, X.; Zhang, Y.; Wang, J.; Wang, Y.; Li, M.; Qiu, Q.; et al. Exome Sequencing Identified Six Copy Number Variations as a Prediction Model for Recurrence of Primary Prostate Cancers with Distinctive Prognosis. *Transl. Cancer Res.* **2020**, *9*, 2231–2242. [[CrossRef](#)] [[PubMed](#)]
32. Quillet, A.; Saad, C.; Ferry, G.; Anouar, Y.; Vergne, N.; Lecroq, T.; Dubessy, C. Improving Bioinformatics Prediction of MicroRNA Targets by Ranks Aggregation. *Front. Genet.* **2020**, *10*, 1330. [[CrossRef](#)] [[PubMed](#)]
33. Rennie, W.; Liu, C.; Carmack, C.S.; Wolenc, A.; Kanoria, S.; Lu, J.; Long, D.; Ding, Y. STarMir: A Web Server for Prediction of MicroRNA Binding Sites. *Nucleic Acids Res.* **2014**, *42*, W114–W118. [[CrossRef](#)]
34. Zuker, M. Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res.* **2003**, *31*, 3406–3415. [[CrossRef](#)] [[PubMed](#)]
35. Ding, Y.; Chan, C.Y.; Lawrence, C.E. Sfold Web Server for Statistical Folding and Rational Design of Nucleic Acids. *Nucleic Acids Res.* **2004**, *32*, W135–W141. [[CrossRef](#)]
36. Rehmsmeier, M.; Steffen, P.; Höchsmann, M.; Giegerich, R. Fast and Effective Prediction of MicroRNA/Target Duplexes. *RNA* **2004**, *10*, 1507–1517. [[CrossRef](#)]
37. Long, D.; Lee, R.; Williams, P.; Chan, C.Y.; Ambros, V.; Ding, Y. Potent Effect of Target Structure on MicroRNA Function. *Nat. Struct. Mol. Biol.* **2007**, *14*, 287–294. [[CrossRef](#)]
38. Martínez-Jiménez, F.; Muiños, F.; Sentís, I.; Deu-Pons, J.; Reyes-Salazar, I.; Arnedo-Pac, C.; Mularoni, L.; Pich, O.; Bonet, J.; Kranas, H.; et al. A Compendium of Mutational Cancer Driver Genes. *Nat. Rev. Cancer* **2020**, *20*, 555–572. [[CrossRef](#)]



39. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **2019**, *47*, D941–D947. [[CrossRef](#)]
40. Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; et al. The STRING Database in 2021: Customizable Protein–Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res.* **2021**, *49*, D605–D612. [[CrossRef](#)]
41. Franceschini, A.; Szklarczyk, D.; Frankild, S.; Kuhn, M.; Simonovic, M.; Roth, A.; Lin, J.; Minguez, P.; Bork, P.; von Mering, C.; et al. STRING v9.1: Protein–Protein Interaction Networks, with Increased Coverage and Integration. *Nucleic Acids Res.* **2012**, *41*, D808–D815. [[CrossRef](#)] [[PubMed](#)]
42. Kumar Gupta, A.; Kumar, M. HPVbase—A Knowledgebase of Viral Integrations, Methylation Patterns and MicroRNAs Aberrant Expression: As Potential Biomarkers for Human Papillomaviruses Mediated Carcinomas. *Sci. Rep.* **2015**, *5*, 12522. [[CrossRef](#)] [[PubMed](#)]
43. Tang, D.; Li, B.; Xu, T.; Hu, R.; Tan, D.; Song, X.; Jia, P.; Zhao, Z. VISDB: A Manually Curated Database of Viral Integration Sites in the Human Genome. *Nucleic Acids Res.* **2020**, *48*, D633–D641. [[CrossRef](#)] [[PubMed](#)]
44. Anna Szymonowicz, K.; Chen, J. Biological and Clinical Aspects of HPV-Related Cancers. *Cancer Biol. Med.* **2020**, *17*, 864–878. [[CrossRef](#)]
45. Ribera-Cortada, I.; Guerrero-Pineda, J.; Trias, I.; Veloza, L.; Garcia, A.; Marimon, L.; Diaz-Mercedes, S.; Alamo, J.R.; Rodrigo-Calvo, M.T.; Vega, N.; et al. Pathogenesis of Penile Squamous Cell Carcinoma: Molecular Update and Systematic Review. *Int. J. Mol. Sci.* **2021**, *23*, 251. [[CrossRef](#)]
46. Schwaederle, M.; Elkin, S.K.; Tomson, B.N.; Carter, J.L.; Kurzrock, R. Squamousness: Next-Generation Sequencing Reveals Shared Molecular Features across Squamous Tumor Types. *Cell Cycle* **2015**, *14*, 2355–2361. [[CrossRef](#)]
47. Cubilla, A.L.; Lloveras, B.; Alejo, M.; Clavero, O.; Chaux, A.; Kasamatsu, E.; Monfuleda, N.; Tous, S.; Alemany, L.; Klaustermeier, J.; et al. Value of P16INK4a in the Pathology of Invasive Penile Squamous Cell Carcinomas: A Report of 202 Cases. *Am. J. Surg. Pathol.* **2011**, *35*, 253–261. [[CrossRef](#)]
48. Gadhikar, M.A.; Zhang, J.; Shen, L.; Rao, X.; Wang, J.; Zhao, M.; Kalu, N.N.; Johnson, F.M.; Byers, L.A.; Heymach, J.; et al. *CDKN2A/P16* Deletion in Head and Neck Cancer Cells Is Associated with CDK2 Activation, Replication Stress, and Vulnerability to CHK1 Inhibition. *Cancer Res.* **2018**, *78*, 781–797. [[CrossRef](#)]
49. Sievers, P.; Hielscher, T.; Schrimpf, D.; Stichel, D.; Reuss, D.E.; Berghoff, A.S.; Neidert, M.C.; Wirsching, H.-G.; Mawrin, C.; Ketter, R.; et al. *CDKN2A/B* Homozygous Deletion Is Associated with Early Recurrence in Meningiomas. *Acta Neuropathol.* **2020**, *140*, 409–413. [[CrossRef](#)]
50. da Mata, S.; Ferreira, J.; Nicolás, I.; Esteves, S.; Esteves, G.; Lérias, S.; Silva, F.; Saco, A.; Cochicho, D.; Cunha, M.; et al. P16 and HPV Genotype Significance in HPV-Associated Cervical Cancer—A Large Cohort of Two Tertiary Referral Centers. *Int. J. Mol. Sci.* **2021**, *22*, 2294. [[CrossRef](#)]
51. Zito Marino, F.; Sabetta, R.; Pagliuca, F.; Brunelli, M.; Aquino, G.; Perdonà, S.; Botti, G.; Facchini, G.; Fiorentino, F.; Di Lauro, G.; et al. Discrepancy of P16 Immunohistochemical Expression and HPV RNA in Penile Cancer. A Multiplex in Situ Hybridization/Immunohistochemistry Approach Study. *Infect. Agents Cancer* **2021**, *16*, 22. [[CrossRef](#)] [[PubMed](#)]
52. Zhu, Q.; Zhang, J.; Chen, Y.; Hu, Q.; Shen, H.; Huang, R.; Liu, Q.; Kaur, J.; Long, M.; Battaglia, S.; et al. Whole-exome Sequencing of Ovarian Cancer Families Uncovers Putative Predisposition Genes. *Int. J. Cancer* **2020**, *146*, 2147–2155. [[CrossRef](#)] [[PubMed](#)]
53. Bailey, M.H.; Meyerson, W.U.; Dursi, L.J.; Wang, L.-B.; Dong, G.; Liang, W.-W.; Weerasinghe, A.; Li, S.; Li, Y.; Kelso, S.; et al. Retrospective Evaluation of Whole Exome and Genome Mutation Calls in 746 Cancer Samples. *Nat. Commun.* **2020**, *11*, 4748. [[CrossRef](#)] [[PubMed](#)]
54. Tan, F.H.; Bai, Y.; Saintigny, P.; Darido, C. mTOR Signalling in Head and Neck Cancer: Heads Up. *Cells* **2019**, *8*, 333. [[CrossRef](#)] [[PubMed](#)]
55. Zhou, Q.; Deng, C.; Li, Z.; Chen, J.; Yao, K.; Huang, K.; Liu, T.; Liu, Z.; Qin, Z.; Zhou, F.; et al. Molecular Characterization and Integrative Genomic Analysis of a Panel of Newly Established Penile Cancer Cell Lines. *Cell Death Dis.* **2018**, *9*, 684. [[CrossRef](#)]
56. Stecca, C.E.; Alt, M.; Jiang, D.M.; Chung, P.; Crook, J.M.; Kulkarni, G.S.; Sridhar, S.S. Recent Advances in the Management of Penile Cancer: A Contemporary Review of the Literature. *Oncol. Ther.* **2021**, *9*, 21–39. [[CrossRef](#)]
57. Huang, K.-B.; Guo, S.-J.; Li, Y.-H.; Zhang, X.-K.; Chen, D.; Spiess, P.E.; Li, Z.-S.; Deng, C.-Z.; Chen, J.-P.; Zhou, Q.-H.; et al. Genome-Wide Profiling Reveals HPV Integration Pattern and Activated Carcinogenic Pathways in Penile Squamous Cell Carcinoma. *Cancers* **2021**, *13*, 6104. [[CrossRef](#)]
58. Freshour, S.L.; Kiwala, S.; Cotto, K.C.; Coffman, A.C.; McMichael, J.F.; Song, J.J.; Griffith, M.; Griffith, O.L.; Wagner, A.H. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with Open Crowdsourcing Efforts. *Nucleic Acids Res.* **2021**, *49*, D1144–D1151. [[CrossRef](#)]
59. Braicu, C.; Buse, M.; Busuioc, C.; Drula, R.; Gulei, D.; Raduly, L.; Rusu, A.; Irimie, A.; Atanasov, A.G.; Slaby, O.; et al. A Comprehensive Review on MAPK: A Promising Therapeutic Target in Cancer. *Cancers* **2019**, *11*, 1618. [[CrossRef](#)]
60. Januario, T.; Ye, X.; Bainer, R.; Alicke, B.; Smith, T.; Haley, B.; Modrusan, Z.; Gould, S.; Yauch, R.L. PRC2-Mediated Repression of SMARCA2 Predicts EZH2 Inhibitor Activity in SWI/SNF Mutant Tumors. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12249–12254. [[CrossRef](#)]

61. Morschhauser, F.; Tilly, H.; Chaidos, A.; McKay, P.; Phillips, T.; Assouline, S.; Batlevi, C.L.; Campbell, P.; Ribrag, V.; Damaj, G.L.; et al. Tazemetostat for Patients with Relapsed or Refractory Follicular Lymphoma: An Open-Label, Single-Arm, Multicentre, Phase 2 Trial. *Lancet Oncol.* **2020**, *21*, 1433–1442. [[CrossRef](#)]
62. Izutsu, K.; Ando, K.; Nishikori, M.; Shibayama, H.; Teshima, T.; Kuroda, J.; Kato, K.; Imaizumi, Y.; Nosaka, K.; Sakai, R.; et al. Phase II Study of Tazemetostat for Relapsed or Refractory B-cell Non-Hodgkin Lymphoma with *EZH2* Mutation in Japan. *Cancer Sci.* **2021**, *112*, 3627–3635. [[CrossRef](#)] [[PubMed](#)]
63. Italiano, A.; Soria, J.-C.; Toulmonde, M.; Michot, J.-M.; Lucchesi, C.; Varga, A.; Coindre, J.-M.; Blakemore, S.J.; Clawson, A.; Suttle, B.; et al. Tazemetostat, an *EZH2* Inhibitor, in Relapsed or Refractory B-Cell Non-Hodgkin Lymphoma and Advanced Solid Tumours: A First-in-Human, Open-Label, Phase 1 Study. *Lancet Oncol.* **2018**, *19*, 649–659. [[CrossRef](#)]
64. Gounder, M.; Schöffski, P.; Jones, R.L.; Agulnik, M.; Cote, G.M.; Villalobos, V.M.; Attia, S.; Chugh, R.; Chen, T.W.-W.; Jahan, T.; et al. Tazemetostat in Advanced Epithelioid Sarcoma with Loss of *INI1/SMARCB1*: An International, Open-Label, Phase 2 Basket Study. *Lancet Oncol.* **2020**, *21*, 1423–1432. [[CrossRef](#)]
65. Pyziak, K.; Sroka-Porada, A.; Rzymiski, T.; Dulak, J.; Łoboda, A. Potential of Enhancer of Zeste Homolog 2 Inhibitors for the Treatment of SWI/SNF Mutant Cancers and Tumor Microenvironment Modulation. *Drug Dev. Res.* **2021**, *82*, 730–753. [[CrossRef](#)]
66. Xue, Y.; Meehan, B.; Macdonald, E.; Venneti, S.; Wang, X.Q.D.; Witkowski, L.; Jelinic, P.; Kong, T.; Martinez, D.; Morin, G.; et al. *CDK4/6* Inhibitors Target *SMARCA4*-Determined Cyclin D1 Deficiency in Hypercalcemic Small Cell Carcinoma of the Ovary. *Nat. Commun.* **2019**, *10*, 558. [[CrossRef](#)]
67. Braal, C.L.; Jongbloed, E.M.; Wilting, S.M.; Mathijssen, R.H.J.; Koolen, S.L.W.; Jager, A. Inhibiting *CDK4/6* in Breast Cancer with Palbociclib, Ribociclib, and Abemaciclib: Similarities and Differences. *Drugs* **2021**, *81*, 317–331. [[CrossRef](#)]
68. Barthel, F.P.; Wei, W.; Tang, M.; Martinez-Ledesma, E.; Hu, X.; Amin, S.B.; Akdemir, K.C.; Seth, S.; Song, X.; Wang, Q.; et al. Systematic Analysis of Telomere Length and Somatic Alterations in 31 Cancer Types. *Nat. Genet.* **2017**, *49*, 349–357. [[CrossRef](#)]
69. Kim, S.K.; Kim, J.-H.; Han, J.H.; Cho, N.H.; Kim, S.J.; Kim, S.I.; Choo, S.H.; Kim, J.S.; Park, B.; Kwon, J.E. TERT Promoter Mutations in Penile Squamous Cell Carcinoma: High Frequency in Non-HPV-Related Type and Association with Favorable Clinicopathologic Features. *J. Cancer Res. Clin. Oncol.* **2021**, *147*, 1125–1135. [[CrossRef](#)]
70. Bell, R.J.A.; Rube, H.T.; Kreig, A.; Mancini, A.; Fouse, S.D.; Nagarajan, R.P.; Choi, S.; Hong, C.; He, D.; Pekmezci, M.; et al. The Transcription Factor GABP Selectively Binds and Activates the Mutant TERT Promoter in Cancer. *Science* **2015**, *348*, 1036–1039. [[CrossRef](#)]
71. Dratwa, M.; Wysoczańska, B.; Łacina, P.; Kubik, T.; Bogunia-Kubik, K. TERT—Regulation and Roles in Cancer Formation. *Front. Immunol.* **2020**, *11*, 589929. [[CrossRef](#)] [[PubMed](#)]
72. Liu, J.; Zhang, C.; Zhao, Y.; Feng, Z. MicroRNA Control of P53: MicroRNAs and P53. *J. Cell Biochem.* **2017**, *118*, 7–14. [[CrossRef](#)] [[PubMed](#)]
73. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front. Endocrinol.* **2018**, *9*, 402. [[CrossRef](#)] [[PubMed](#)]
74. Secker, G.A.; Harvey, N.L. Regulation of VEGFR Signalling in Lymphatic Vascular Development and Disease: An Update. *Int. J. Mol. Sci.* **2021**, *22*, 7760. [[CrossRef](#)]
75. Bai, M.; Li, Z.-G.; Ba, Y. Influence of KDR Genetic Variation on the Efficacy and Safety of Patients with Chemotherapy Refractory Metastatic CRC Who Received Apatinib Treatment. *Int. J. Mol. Sci.* **2021**, *14*, 1041–1055. [[CrossRef](#)]
76. Phelan, J.D.; Young, R.M.; Webster, D.E.; Roulland, S.; Wright, G.W.; Kasbekar, M.; Shaffer, A.L.; Ceribelli, M.; Wang, J.Q.; Schmitz, R.; et al. A Multiprotein Supercomplex Controlling Oncogenic Signalling in Lymphoma. *Nature* **2018**, *560*, 387–391. [[CrossRef](#)]
77. Seeholzer, T.; Kurz, S.; Schlauderer, F.; Woods, S.; Gehring, T.; Widmann, S.; Lammens, K.; Krappmann, D. BCL10-CARD11 Fusion Mimics an Active CARD11 Seed That Triggers Constitutive BCL10 Oligomerization and Lymphocyte Activation. *Front. Immunol.* **2018**, *9*, 2695. [[CrossRef](#)]
78. La Fleur, L.; Falk-Sörqvist, E.; Smeds, P.; Berglund, A.; Sundström, M.; Mattsson, J.S.; Brandén, E.; Koyi, H.; Isaksson, J.; Brunnström, H.; et al. Mutation Patterns in a Population-Based Non-Small Cell Lung Cancer Cohort and Prognostic Impact of Concomitant Mutations in *KRAS* and *TP53* or *STK11*. *Lung Cancer* **2019**, *130*, 50–58. [[CrossRef](#)]
79. Qiu, Z.; Lin, A.; Li, K.; Lin, W.; Wang, Q.; Wei, T.; Zhu, W.; Luo, P.; Zhang, J. A Novel Mutation Panel for Predicting Etoposide Resistance in Small-Cell Lung Cancer. *Drug Des. Dev. Ther.* **2019**, *13*, 2021–2041. [[CrossRef](#)]
80. Renou, L.; Boelle, P.-Y.; Deswarte, C.; Spicuglia, S.; Benyoucef, A.; Calvo, J.; Uzan, B.; Belhocine, M.; Cieslak, A.; Landman-Parker, J.; et al. Homeobox Protein *TLX3* Activates *MiR-125b* Expression to Promote T-Cell Acute Lymphoblastic Leukemia. *Blood Adv.* **2017**, *1*, 733–747. [[CrossRef](#)]
81. Kawahara, Y. Human Diseases Caused by Germline and Somatic Abnormalities in MicroRNA and MicroRNA-Related Genes: MicroRNAs and Human Genetic Diseases. *Congenit. Anom.* **2014**, *54*, 12–21. [[CrossRef](#)] [[PubMed](#)]
82. Chhichholiya, Y.; Suryan, A.K.; Suman, P.; Munshi, A.; Singh, S. SNPs in MiRNAs and Target Sequences: Role in Cancer and Diabetes. *Front. Genet.* **2021**, *12*, 793523. [[CrossRef](#)] [[PubMed](#)]
83. Zhao, D. Single Nucleotide Alterations in MicroRNAs and Human Cancer—A Not Fully Explored Field. *Non-Coding RNA Res.* **2020**, *5*, 27–31. [[CrossRef](#)] [[PubMed](#)]