Article

Somatic Mutational Landscape in Mexican Patients: *CDH1* Mutations and chr20q13.33 Amplifications Are Associated with Diffuse-Type Gastric Adenocarcinoma

Dennis Cerrato-Izaguirre [1,2], Yolanda I. Chirino [3], Diddier Prada [2], Ericka Marel Quezada-Maldonado [2], Luis A Herrera [4], Angélica Hernández-Guerrero [5], Juan Octavio Alonso-Larraga [5], Roberto Herrera-Goepfert [6], Luis F. Oñate-Ocaña [7], David Cantú-de-León [8], Abelardo Meneses-García [9], Patricia Basurto-Lozada [10], Carla Daniela Robles-Espinoza [10,11], Javier Camacho [1], Claudia M. García-Cuellar [2] and Yesennia Sánchez-Pérez [2,*]

1   Departamento de Farmacología, Centro de Investigación y de Estudios Avanzados del I.P.N. (CINVESTAV), Avenida Instituto Politécnico Nacional No. 2508, Ciudad de México CP. 07360, Mexico
2   Subdirección de Investigación Básica, Instituto Nacional de Cancerología (INCan), San Fernando No. 22, Tlalpan, Ciudad de México CP. 14080, Mexico
3   Unidad de Biomedicina, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Los Reyes Iztacala, Tlalnepantla de Baz, Estado de México CP. 54090, Mexico
4   Instituto Nacional de Medicina Genómica (INMEGEN), Periférico Sur No. 4809, Arenal Tepepan, Tlalpan, Ciudad de México CP. 14610, Mexico
5   Servicio de Endoscopía, Instituto Nacional de Cancerología (INCan), San Fernando No. 22, Tlalpan, Ciudad de México CP. 14080, Mexico
6   Departamento de Patología, Instituto Nacional de Cancerología (INCan), San Fernando No. 22, Tlalpan, Ciudad de México CP. 14080, Mexico
7   Subdirección de Investigación Clínica, Instituto Nacional de Cancerología (INCan), San Fernando No. 22, Tlalpan, Ciudad de México CP. 14080, Mexico
8   Dirección de Investigación, Instituto Nacional de Cancerología (INCan), San Fernando No. 22, Tlalpan, Ciudad de México CP. 14080, Mexico
9   Dirección General, Instituto Nacional de Cancerología (INCan), San Fernando No. 22, Tlalpan, Ciudad de México CP. 14080, Mexico
10 Laboratorio Internacional de Investigación Sobre el Genoma Humano, Universidad Nacional Autónoma de México, Santiago de Querétaro CP. 76010, Mexico
11 Wellcome Sanger Institute, Hinxton, Cambridgeshire CB10 1SA, UK
*   Correspondence: s_yesennia@yahoo.com.mx or yesennia.sanchez@salud.gob.mx

Abstract: The Hispanic population, compared with other ethnic groups, presents a more aggressive gastric cancer phenotype with higher frequency of diffuse-type gastric adenocarcinoma (GA); this could be related to the mutational landscape of GA in these patients. Using whole-exome sequencing, we sought to present the mutational landscape of GA from 50 Mexican patients who were treated at The Instituto Nacional de Cancerología from 2019 to 2020. We performed a comprehensive statistical analysis to explore the relationship of the genomic variants and clinical data such as tumor histology and presence of signet-ring cell, *H. pylori*, and EBV. We describe a potentially different mutational landscape between diffuse and intestinal GA in Mexican patients. Patients with intestinal-type GA tended to present a higher frequency of *NOTCH1* mutations, copy number gains in cytobands 13.14, 10q23.33, and 12q25.1, and copy number losses in cytobands 7p12, 14q24.2, and 11q13.1; whereas patients with diffuse-type GA tended to present a high frequency of *CDH1* mutations and CNV gains in cytobands 20q13.33 and 22q11.21. This is the first description of a mutational landscape of GA in Mexican patients to better understand tumorigenesis in Hispanic patients and lay the groundwork for discovering potential biomarkers and therapeutic targets.

Keywords: gastric cancer; Hispanic ethnicity; cancer genomics; precision medicine; copy number variation; mutations

1. Introduction

Worldwide, gastric cancer is the fifth cause of cancer-related deaths, with a 5-year survival rate of less than 30% in Western countries [1,2]. Asia has the highest incidence of gastric cancer, followed by Latin America (LATAM) and Europe [3]. Although LATAM ranks second in incidence, the knowledge regarding the molecular landscape of gastric cancer in this region is scarce [4]. Poor dietary habits, tobacco usage, and pathogens such as *Helicobacter pylori* and Epstein–Barr virus (EBV) are risk factors for gastric cancer [5]. Up to 95% of all patients with gastric cancer are diagnosed as having gastric adenocarcinoma (GA). According to the Lauren classification, the tumor histology of GA is classified as diffuse, intestinal, or mixed type [6]. The intestinal-type GA has been reported as the common type across Asian and European-descent ethnicities; however, a higher proportion of diffuse-type GA has been reported among the Hispanic population [7,8]; thus, a better comprehension of the molecular landscape of GA could help to understand carcinogenesis and develop new therapeutic approaches in this population.

A differential mutational pattern has been reported between diffuse- and intestinal-type GA. Somatic variations in *RHOA*, *CDH1*, *TAF1*, and *KMT2C* have been associated with diffuse-type GA [9,10]. Gene fusions, such as *CLDN18-ARHGAP26/6*, deletions in 9p, and gains in 20q are also some structural variations related to diffuse-type GA and signet-ring cell formations [9,11]. In contrast, intestinal-type GA is characterized by a high frequency of copy number variations (CNVs), commonly leading to the activation of tyrosine kinase-associated receptors, including ERBB2 and EGFR [12]. The study of GA mutational patterns has led to the identification of biomarkers and the use of targeted therapies, including those with trastuzumab and pembrolizumab, that benefit mainly patients with intestinal-type GA; on the other hand, diffuse-type GA has been associated with low benefit from chemotherapy [13].

The Cancer Genome Atlas (TCGA) consortium has divided GA into four molecular subtypes: EBV-positive (EBV), microsatellite-instable (MSI), chromosome-instable (CIN), and genome-stable (GS). Each GA subtype displays unique molecular and clinical characteristics [14]. The ethnicity of patients has an impact on the distribution of these subtypes; for instance, MSI and EBV subtypes are more common in South Korean patients, while GS is more abundant in Ukrainian patients [15]. In addition, the CIN subtype is more frequent in European-descent patients than in Asian patients [16]. When identifying the clinical significance of this classification, the EBV subtype is associated with the best prognosis, and the GS is associated with the worst prognosis [17].

The Hispanic population in the United States has been associated with an early age of onset, higher frequency of diffuse-type GA, and higher risk of peritoneal carcinomatosis compared with European-descent, Asian, and African American patients [18–21]. This characteristic has been associated with a distinctive genomic profile and a high frequency of *CDH1* germline mutations [22]. In Mexico, an estimated 8804 new cases and 6735 deaths related to gastric cancer were recorded in 2020 [3]. Differences in GA incidence are reported within the country, with the south and central regions especially affected [23]. The Mexican population has a rich genetic admixture with traces of Amerindian, European, African, and Asian ancestry, which varies across the country [24,25]. This genetic peculiarity may be related to the evolution of GA and the high frequency of diffuse-type GA. The GA mutational landscape has been described in populations of European-descent and Asian ethnicities; however, mutational data from the Hispanic population are elusive. Here, we sought to identify the somatic variants, tumor mutational burden (TMB), affected signaling pathways, and the CNVs that are present in Mexican patients with GA to describe a mutational landscape and explore its relationship with tumor histology and clinical characteristics.

2. Results

2.1. Clinicopathological Characteristics of the Patients

GA samples from 50 Mexican patients who were treated at The Instituto Nacional de Cancerología of Mexico (INCan) from January 2019 to January 2020 were analyzed.

The mean age of patients was 54.84 years (range, 22–89 years), with a predominant percentage of males (58%). In nearly half of the patients (42%), the tumor was found to have spread within two or more sites of the stomach. Tumors located in the fundus and body were the second most common (36%), followed by tumors located in the antrum and pylorus (22%). Diffuse-type GA was detected in 62% of the samples and intestinal-type GA was detected in 32%, whereas mixed-type GA was detected in the remaining 6% of the samples. Signet-ring cells were present in 68% of patients, and poorly differentiated tumor grade was found in 84% of patients. Patients with advanced clinical stages represented 62% of the cohort. The clinical characteristics of the patients included are presented in Table 1. After an early follow-up of 2 years, we observed that patients had a median overall survival of 11 months (Supplementary Figure S1). However, when comparing overall survival between patients with diffuse- and intestinal-type GA, no differences were observed (Supplementary Figure S2).

Table 1. The clinical and pathological characteristics of patients treated at The Instituto Nacional de Cancerología between January 2019 and January 2020 (N = 50).

Variable	Mean	SD
Age (years)	54.94	15.49
BMI (Kg/m^2)	23.60	5.18
Follow-up (months)	12.53	14.17

Variable	n	%
Sex		
Female	21	42.00
Clinical stage		
I–II	3	6.00
III–IV	47	94.00
Tumor location		
Antrum and pylorus	11	22.00
Cardia	0	0.00
Fundus and body	18	36.00
Two or more locations	21	42.00
Endoscopic impression [1]		
I	1	2.00
II	3	6.00
III	15	30.00
IV	22	44.00
V	6	12.00
Not reported	3	6.00
Tumor histology [2]		
Intestinal	16	32.00
Diffuse	31	62.00
Mixed	3	6.00
Tumor grade		
Well-differentiated	0	0.00
Moderately differentiated	8	16.00
Poorly differentiated	42	84.00
Signet-ring cell		
Positive	34	68.00
H. pylori		
Positive	21	42.00
EBV		
Positive	6	12.00

SD: standard deviation; BMI: body mass index, EBV: Epstein–Barr virus. [1] Endoscopic impression is according to Borrmann classification. [2] Tumor histology is grouped according to the Lauren classification.

2.2. Helicobacter Pylori and EBV Identification

Helicobacter pylori infection was identified in 42% of patients, whereas EBV infection was present in 12% of patients. Concomitant infection between both pathogens was present in only 8% of patients. The presence of *H. pylori* (*p*-value = 0.69) or EBV (*p*-value = 0.673) was not associated with tumor histology. The presence of signet-ring cell was also not associated with the presence of *H. pylori* (*p*-value = 0.40) and EBV (*p*-value = 0.15).

2.3. Tumor Mutational Burden and Clinical Features

Whole-exome sequencing (WES) was performed successfully, with one sample failing the sequencing process. An average sequencing depth of 186.90× and a minimum coverage of 99.5% of the target region were achieved. The samples had a mean of 94.86% Q30 quality in the sequenced bases. FASTQ files were obtained, and further bioinformatic analyses were performed to detect somatic variations and CNVs (see Material and Methods section). A total of 119,744 somatic variants were identified across all samples. Histograms of the variant allele fraction (VAF) for each sample are depicted in Supplementary Figure S3. The total number of mutations, including driver and passenger somatic variants, present in the exome region were used to identify the tumor mutational burden (TMB) of each patient. The TMB values ranged from 0.16 to 158.35 mutations/Mb.

Patients were grouped according to their TMB status, and high TMB (TMB ≥ 10 mutations/Mb) was observed in 10% of the patients (Figure 1a). When exploring deeper into the patients with high TMB status, we observed that mutations in DNA mismatch repair were only found in patient CGTE_14 with a driver exonic mutation in *MSH6* (TV716-717X, R1024Q, T1085TX, and Q1314X), and the rest of the patients with high TMB had passenger intronic mutations in *MSH6* (CGTE_26, CGTE_18, and CGTE_40), *MSH2* (CGTE_40 and CGTE_28) and *MLH1* (CGTE_18). The potential relationship between the TMB values and clinical features was explored. Patients with intestinal-type GA (*p*-value = 0.026) and absence of signet-ring cell (*p*-value = 0.015) had greater TMB values (Figure 1b,c). The presence of *H. pylori* (*p*-value = 0.54) and EBV infection (*p*-value = 0.87) were not associated with TMB values (Supplementary Figure S4). We also evaluated if the TMB status of the patients affected their overall survival (Supplementary Figure S5); however, within this early follow-up time, no differences were identified between the groups (*p*-value = 0.78).
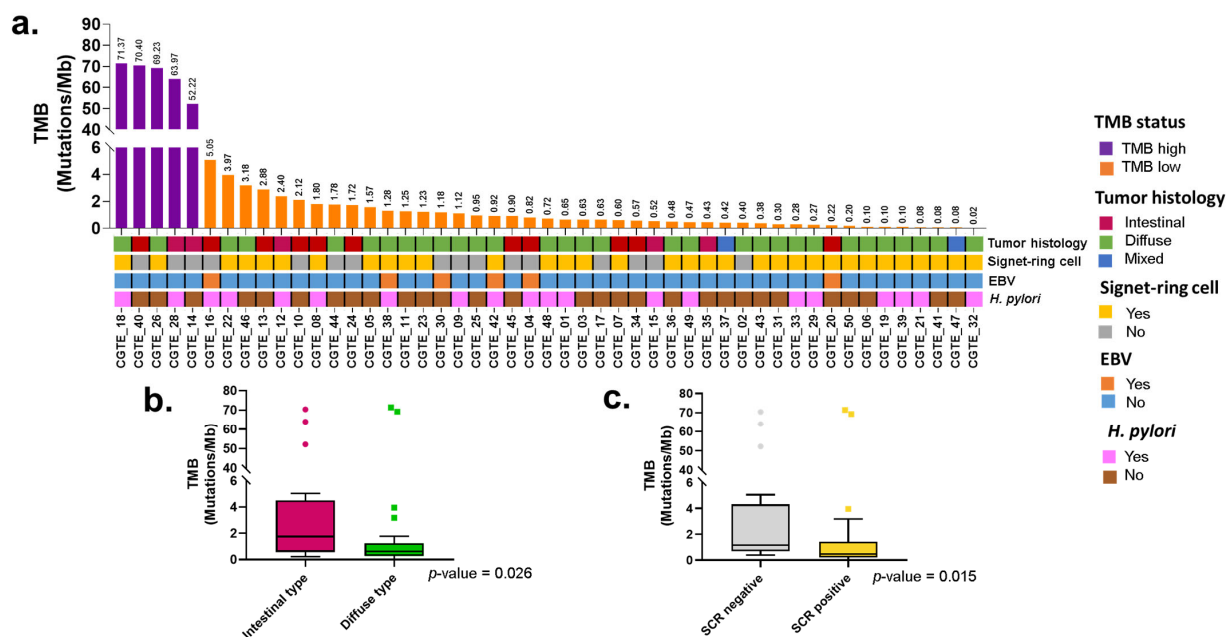
Figure 1. Tumor mutational burden of gastric adenocarcinoma in patients treated at The Instituto Nacional de Cancerología of Mexico between January 2019 and January 2020 (n = 49). (**a**) Tumoral mutational burden (TMB) across the sequenced samples. TMB was calculated with all the variants affecting the coding region. Purple bars indicate high TMB, and orange bars indicate low TMB. TMB ≥ 10 mutations/Mb was considered as high TMB. The lower panel represents the clinical features of the patients including tumor histology, signet-ring cell presence, Epstein–Barr virus (EBV) presence, and *Helicobacter pylori* (*H. pylori*) infection. (**b**) Comparison of TMB value according tumor histology. (**c**) Comparison of TMB value according to the presence of signet-ring cell. *p*-values were calculated using Kruskal–Wallis test.

2.4. Somatic Variants and Tumor Histology

Somatic variants were classified as driver and passenger using the BosstDM algorithm (*see* materials and methods). We found that eight samples had missing driver mutations. The tumor suppressor gene *TP53* was the most frequently mutated gene, altered in 41% of the patients. Missense variants were the most common mutation type observed across the genes (Figure 2).
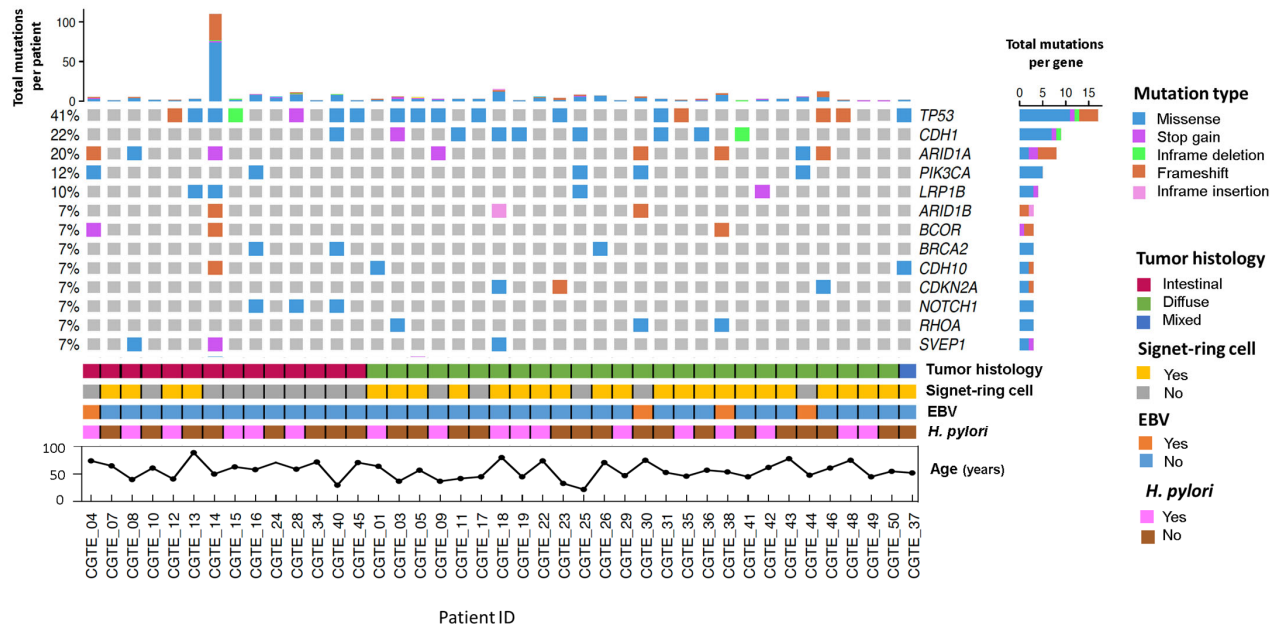
Figure 2. Driver somatic variants and clinical characteristics of patients with gastric adenocarcinoma treated at The Instituto Nacional de Cancerología between January 2019 and January 2020 (n = 41). Oncoprint, sorted by tumor histology, depicting the genes affected by driver somatic variants (single nucleotide variations, small insertions, and small deletions) in 7% or more of the samples. Less frequent driver somatic variants can be found in Supplementary Table S1. The variants are represented according to the mutation type, each described on the right color panel. The upper bar plot represents the number of driver somatic variants per patient, and the right bar plot represents the number of driver somatic variants per gene. The lower panner represents the tumor histology, the presence of signet-ring cells, Epstein–Barr virus (EBV), and *Helicobacter pylori* (*H. pylori*), and age per patient.

Patients were sorted by tumor histology, and the frequency of mutated genes between diffuse-type and intestinal-type GA was assessed. *NOTCH1* somatic variants were found exclusively in patients with intestinal-type GA histology (*p*-value = 0.03, false discovery rate (FDR) = 0.57). On the other hand, patients with diffuse histology tended to be associated with somatic variants in *CDH1* (*p*-value = 0.08, FDR = 0.57). Mutations in *RHOA* (*p*-value = 0.26, FDR = 0.57) and *CDKN2A* (*p*-value = 0.26, FDR = 0.57) were found to be present only in patients with diffuse-type GA. When comparing the presence of signet-ring cell and the frequency of somatic variants, mutations in *PIK3CA* (*p*-value < 0.01, FDR = 0.56) and *NOTCH1* (*p*-value = 0.03, FDR = 0.50) were exclusively present in patients without signet-ring cells. Somatic variants in *ARID1A* (*p*-value = 0.08, FDR = 0.56) also tended to be present in patients without signet-ring cells (Supplementary Figure S6).

2.5. Somatic Variants, H. pylori, and EBV

Patients who were EBV-positive were associated with a higher frequency of *PIK3CA* (*p*-value < 0.01, FDR = 0.04), *ARID1A* (*p*-value < 0.01, FDR = 0.31), *SMAD4* (*p*-value = 0.01, FDR = 0.82), *BCOR* (*p*-value = 0.03, FDR = 0.87), and *RHOA* (*p*-value = 0.03, FDR = 0.87) somatic variants. No associations between the presence of *H. Pylori* infection and the frequency of somatic variants were found (Supplementary Figure S7).

2.6. Signaling Pathways Affected by Somatic Mutations According to Tumor Histology

Patients were grouped according to their tumor histology to observe the differences between diffuse- and intestinal-type GA in the levels of enrichment ratio for the signaling pathways affected by driver mutations. Using the PANTHER dataset as reference [26], the hypoxia response via HIF activation was the most enriched signaling pathway in intestinal-type GA (enrichment ratio (ER) = 33%, FDR < 0.01) and diffuse-type GA (ER = 29.89, FDR < 0.01). Enrichment of the p53-related pathways, as well as the insulin/IGF and VEGF signaling pathways, was shared between diffuse- and intestinal-type GA. However, the JAK/STAT signaling pathway and Axon guidance mediated by netrin tended to be more enriched in diffuse-type GA (Figure 3a).
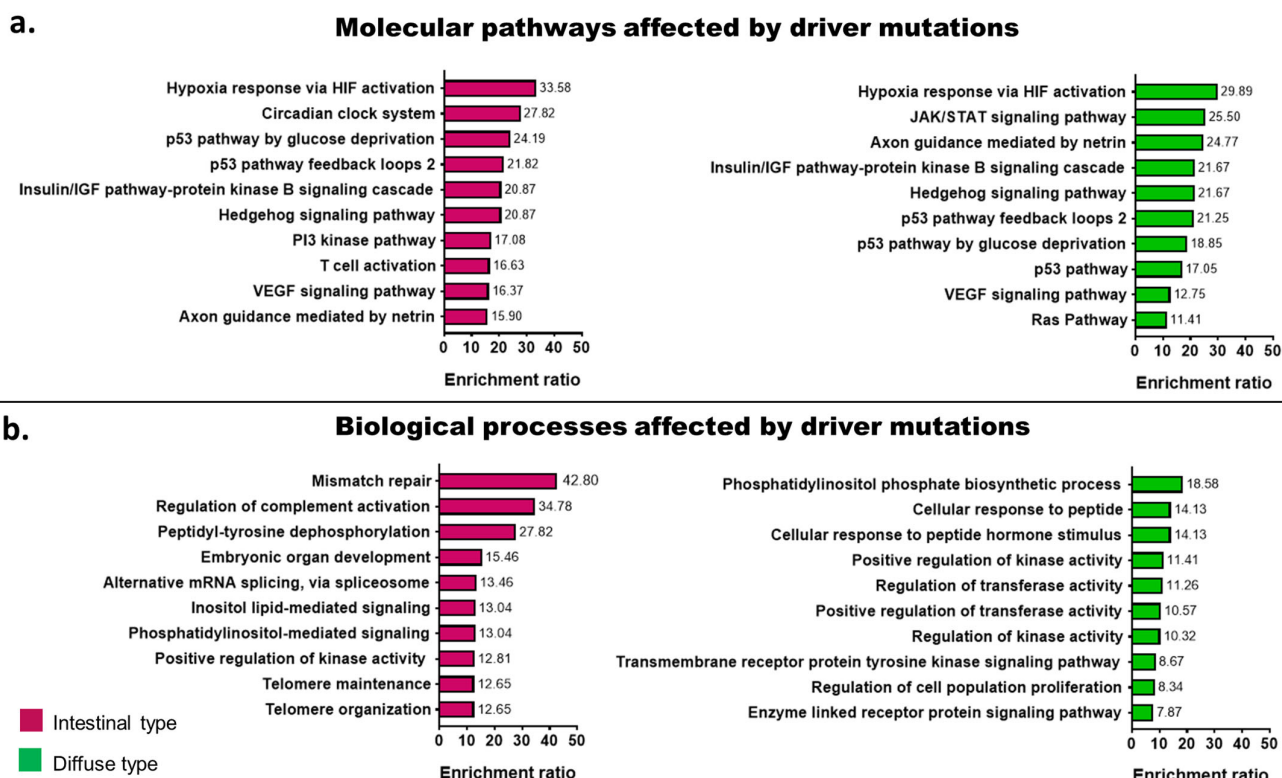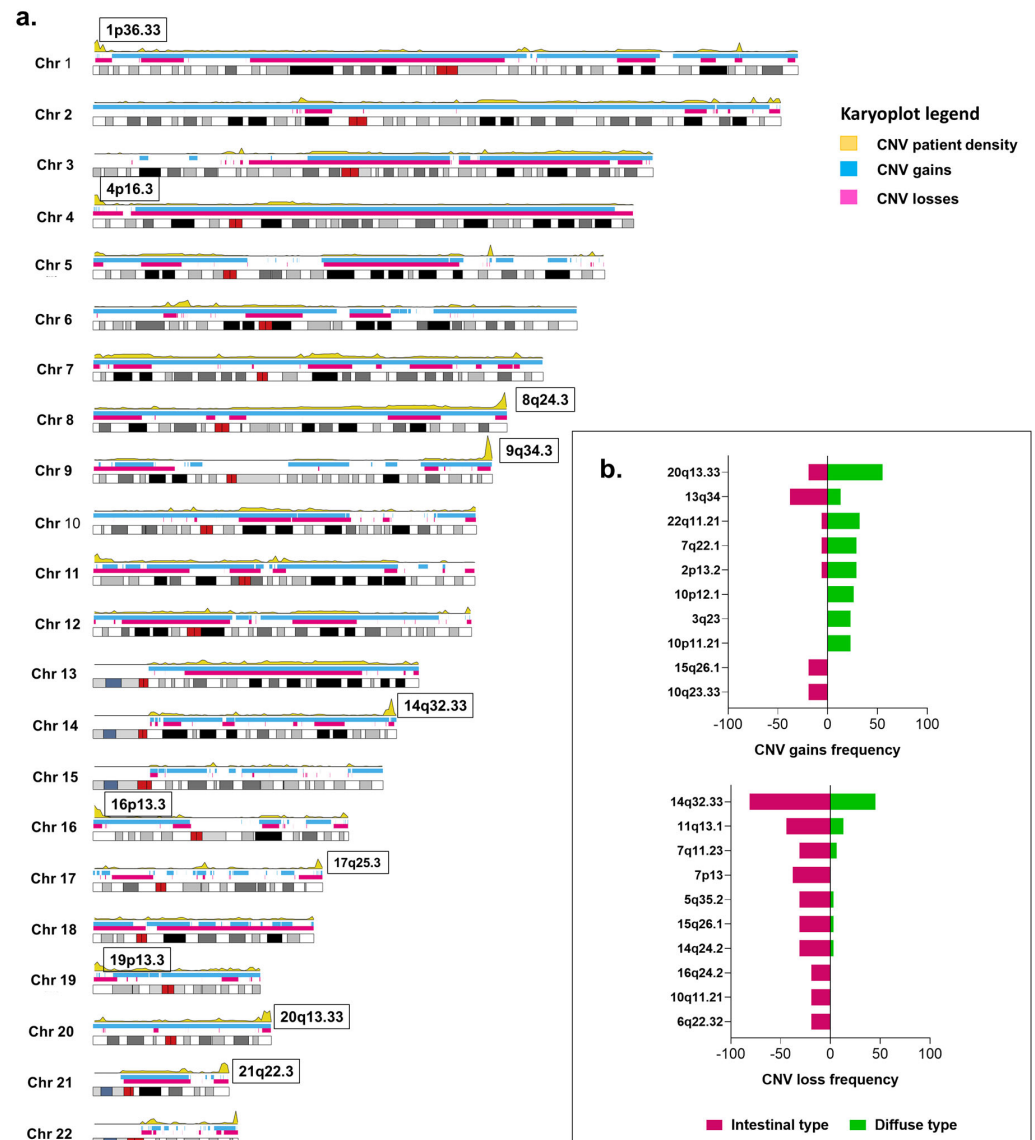

Figure 3. Molecular pathways affected by driver somatic variants from patients with gastric adenocarcinoma treated at The Instituto Nacional de Cancerología of Mexico between January 2019 and January 2020, grouped according to the tumor histology. (**a**) Bar plot representing the molecular pathways affected by driver somatic variants using PANTHER as reference dataset. (**b**) Bar plot representing depicting the biological processes affected by driver somatic variants, using as reference the Gene Ontology dataset.

The biological processes affected by somatic variants were different in patients with diffuse- and intestinal-type GA, and Figure 3b shows the biological processes affected by somatic variants according to the Gene Ontology database. The mismatch repair mechanism (ER = 42.80, FDR < 0.01) was the most affected biological process in patients with intestinal-type GA, whereas the biosynthetic process of phosphatidylinositol (ER = 18.58, FDR < 0.01) was affected in patients with diffuse-type GA.

2.7. Copy Number Variations and Tumor Histology

CNVs affecting patients with GA were identified for different genomic regions. CNV gains were the most commonly found within the patients (Supplementary Figure S8). Chromosomes 2, 7, 8, and 20 were affected almost entirely by CNV gains, whereas chromosomes 4 and 17 were affected by CNV losses. Even though large genomic regions seem to be affected by CNVs, our cohort showed high-density areas where CNV tends to occur more

frequently (Figure 4a). Cytoband 20q13.33 was the most frequently affected, altered in 73.46% of the patients. This cytoband harbors different genes including *LAMA5*, which encodes for the alpha subunit of laminin 5. *LAMA5* was the most frequently affected gene by CNVs, CNV losses occurring in 61.11% of the cases. *SYCP2*, also present in 20q13.33, encodes for a protein from the synaptonemal complex and was affected by CNV gains in 93.75% of the cases. *NOTCH1*, harbored in cytoband 9q34.4, presented CNV losses in 71% of the cases, whereas *JAG2*, which is a *NOTCH1* ligand and is encoded within cytoband 14q32.33, was affected by CNV losses in 78.57% of the cases. A list of all the CNVs found in this study and the genes affected within the cytobands can be found in Supplementary Table S2.

Figure 4. Copy number variations (CNVs) present in patients with gastric adenocarcinoma treated at The Instituto Nacional de Cancerología between January 2019 and January 2020 (n = 49). (**a**) Karyoplot depicting the genomic regions affected by CNVs in all the patients (n = 49). Blue represents the CNV gains, pink, the CNV losses. The yellow density plot represents the frequency of patients affected by CNVs across the genome. Top-ten frequently affected cytobands across all samples are shown in a box in bold. (**b**) Divergent bar plot showing the comparison between the affected cytobands in patients with diffuse- and intestinal-type gastric cancer.

When grouping the patients according to their tumor histology (Figure 4b), patients with diffuse-type GA tended to have gains in 20q13.33 (*p*-value = 0.01, FDR = 1) and 22q11.21 (*p*-value = 0.04, FDR = 1), whereas patients with intestinal-type GA presented a higher frequency of CNV losses in 7p12 (*p*-value < 0.01, FDR = 0.27), 14q24.2 (*p*-value = 0.01, FDR = 0.66), and 11q13.1 (*p*-value = 0.02, FDR = 0.66). Patients with intestinal-type GA also showed a high frequency of CNV gains in cytoband 13.q14 (*p*-value = 0.04, FDR = 1), 10q23.33 (*p*-value = 0.03, FDR = 1), and 15q25.1 (*p*-value = 0.03, FDR = 1). In addition, the presence of signet-ring cells was associated with CNV gains in cytobands 6p22.2 (*p*-value = 0.02, FDR = 0.67) and 1q21.3 (*p*-value = 0.07, FDR = 0.67). CNV losses in 4p16.33 were associated with absence of signet-ring cells (*p*-value = 0.03, FDR = 0.67) (Supplementary Figure S9).

### 2.8. Copy Number Variations, H. pylori, and EBV

When grouping patients according to the presence of *H. pylori*, patients with *H. Pylori* presented a higher frequency of CNV gains in 17q21.2 (*p*-value = 0.02, FDR = 0.57), 10q26.3 (*p*-value = 0.02, FDR = 0.57), and 10q25.3 (*p*-value = 0.02, FDR = 0.57). CNV losses in 12q12 (*p*-value = 0.01, FDR = 0.57) were also common in patients with *H. pylori* infection. Compared to when grouping according to the EBV presence, we observed that patients with EBV infection tended to present a high frequency of CNV gains in cytobands 19q13.42 (*p*-value = 0.01, FDR = 087), 18q22.2 (*p*-value = 0.01, FDR = 0.67), and 2p25.1 (*p*-value = 0.01, FDR = 0.67). Differences in CNV losses were not found between the patients with and without EBV infection (Supplementary Figure S10).

2.9. Mutational Processes and Their Associations with Clinical–Pathological Characteristics

The mutational processes involved in GA carcinogenesis of the patients were quantified through a mutational signature analysis to identify the contribution of single-base signatures (SBSs). The SBS16 related to liver cancer [27] and high mutational rate due to an inefficient nucleotide excision repair pathway [28] was the most common signature, present in 33% of the patients that showed a contribution higher than 0.2 (Supplementary Figure S11), followed by the SBS1 related to the aging process, which was found to be overrepresented in 28%.

Patients were categorized according to similarities in the pattern of the SBS contributions; however, no defined clusters were found. When comparing the contribution of SBS with the clinical features of the patients, we identified that tumors with high TMB had higher contribution of SBS16 (Figure 5a), patients with signet-ring cells had a lower contribution of SBS01 (Figure 5b), and patients with intestinal-type GA displayed a higher contribution of SBS17 (Figure 5c). Supplementary Figure S12 shows a complete panorama of the associations between the most frequent mutational signatures (SBS01, SBS03, SBS05, SBS09, SBS16, and SBS17) found in our cohort and the clinical–pathological characteristics of the patients.
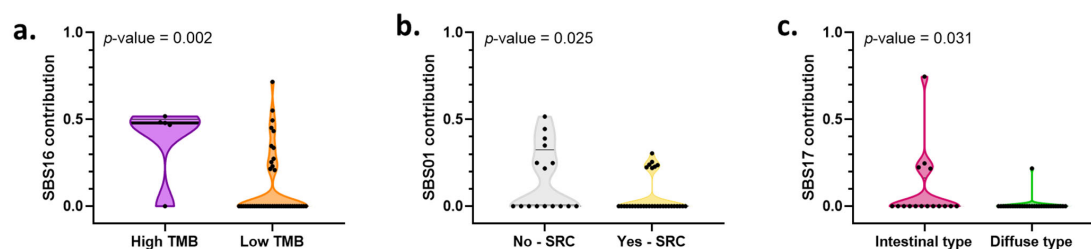
Figure 5. Mutational processes associated with clinical–pathological characteristics of patients with gastric adenocarcinoma treated at The Instituto Nacional de Cancerología between January 2019 and January 2020 (n = 49). (**a**) Violin plot depicting the association between the contribution of single-base substitution signature (SBS) 16 and the tumoral mutational burden (TMB) status of the patients. (**b**) Violin plot depicting the association between the contribution of SBS01 and the presence of signet-ring cells (SRSs). (**c**) Violin plot depicting the association between the contribution of SBS17 and the tumor histology.

2.10. Comparison of Genes Affected by Somatic Variants and Copy Number Variations between Mexican Patients and Other Cohorts

To identify the similitudes and differences in the mutational landscape of Mexican patients with the mutated genes reported by European-descent and Asian patients, we compared this cohort with two other cohorts. We compared the patients included in this study (the INCan cohort) with the TCGA stomach adenocarcinoma firehouse legacy study cohort [14] and the OncoSG cohort [29]. The TCGA cohort includes European-descent and Asian patients, the OncoSG cohort includes Asian patients, and the INCan cohort includes Mexican patients. A total of 14,296 genes affected by somatic variants and 10,521 genes affected by CNVs were shared between the three cohorts (Figure 6a). However, about 5000 mutated genes were exclusive to the INCan cohort. *TP53* was the gene most frequently mutated across all cohorts. CDH1 seemed to be more frequently mutated in the INCan cohort (Figure 6b). A list showing the frequency of all the genes with driver mutations in the three cohorts can be found in Supplementary Table S3. Additionally, the genes encoded in cytoband 1p36.33: *AGRN*, *MEGF6*, *TNFRSF18*, and *TNFRSF4* seemed to be more affected by CNVs in the INCan cohort; they were affected by CNV losses. *WDR90* and *ABCA2*, which are encoded within cytobands 16p13.3 and 9q34.4, respectively, were also affected by CNV losses.
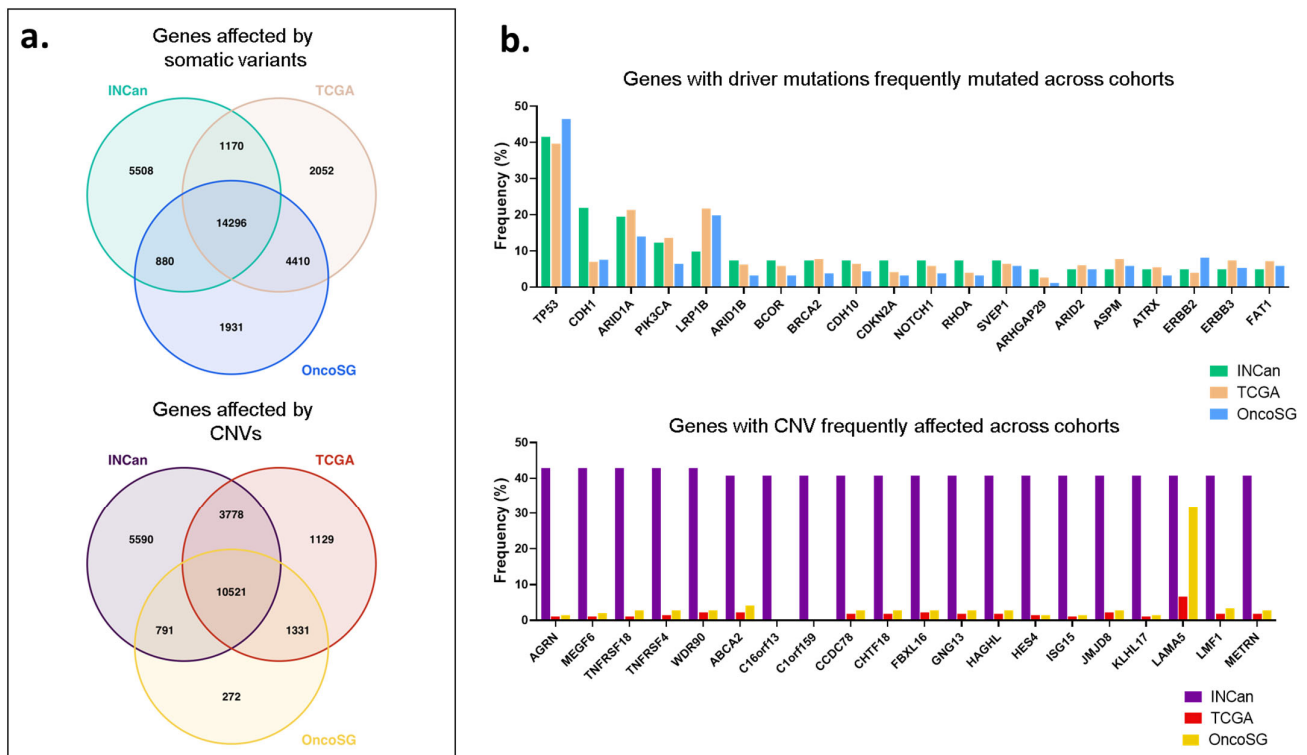
Figure 6. Genes affected by somatic variants and copy number variations in patients with gastric adenocarcinoma among three different cohorts. (**a**) Venn diagram depicting the number of somatic and structural variants shared between the patients from The Instituto Nacional de Cancerología (INCan), The Cancer Genome Atlas (TCGA) stomach adenocarcinoma firehouse legacy, and the OncoSG cohorts. (**b**) Bar plot presenting the 20 most frequently affected genes by driver mutations and copy number variations (CNVs) found in our analysis (INCan cohort) and their frequency found in the TCGA and OncoSG cohorts.

3. Discussion

To the best of our knowledge, this is the first study able to describe a mutational landscape of GA in a cohort of Mexican patients using fresh samples. Here, we analyzed the somatic variants and CNVs using WES and identified a mutational landscape for Mexican

patients with GA. We compared a potentially different mutational landscape between diffuse- and intestinal-type GA in Mexican patients. Patients with intestinal-type GA tended to present a higher frequency of *NOTCH1* mutations, CNV gains in cytobands 13.14, 10q23.33, and 12q25.1, and CNV losses in cytobands 7p12, 14q24.2, and 11q13.1; whereas patients with diffuse-type GA tended to present a higher frequency of *CDH1* mutations and CNV gains in cytobands 20q13.33 and 22q11.21.

This cohort showed similar clinicopathological characteristics reported for other Mexican and Hispanic patients with GA. The patients in this study had a higher proportion of diffuse-type GA, clinical state III–IV, and a low proportion of EBV infection. These characteristics align with those reported in other studies using larger cohorts of Mexican patients, including a previous study by our team [30], which also mentioned a predominance of diffuse-type GA [31,32].

Similar to Asian cohorts where TMB has been proposed as a biomarker for survival and use of immunotherapy, we identified that 10% of patients with GA had high TMB [33,34]. High TMB has been associated with intestinal-type GA [35], and we found that Mexican patients with greater TMB raw values were also associated with intestinal-type GA. It seems that the Mexican patients behave similar to other cohorts in terms of TMB. We found an association between a greater contribution of SBS16 and high TMB values. However, no evidence of somatic mutations in genes associated with the mismatch repair pathway, a common explanation for high TMB of the patients with the microsatellite-instable (MSI) phenotype, was found. In African American patients with esophageal squamous cell carcinoma, a prominent association between SBS16 and high TMB was attributed to defects in the nucleotide excision repair (NER) pathway of the patients with high TMB (27,28). Other studies are necessary to expose the implications of defects on the NER pathway in Mexican patients with high TMB.

In gastric cancer, TMB does not correlate with the MSI phenotype, but both are promising predictive biomarkers for the use immunotherapy [36]. Currently, the Precision Medicine Working Group from the European Society for Medical Oncology recommends TMB as a biomarker to select PD-1/PD-L1 therapy for well- and moderately differentiated neuroendocrine tumors and cervical, salivary, thyroid, and vulvar cancers [37]; however, no recommendations have been issued for gastric cancer. The US Food and Drug Administration (FDA) has approved the use of TMB as a biomarker for selecting pembrolizumab for solid tumors with high levels of TMB (TMB > 10 mutations/mb) based on the results of the KEYNOTE-158 study [38]. In Mexico, TMB is not used as a biomarker for gastric cancer treatment. Here, we found no correlation between TMB and Lauren classification; however, other studies should be performed to investigate the use of TMB as a biomarker in other aspects of the disease, such as its utility in selecting targeted therapies.

Whole-exome sequencing studies performed in Asian and European-descent populations have identified a set of mutations and genes with the potential to drive personalized medicine in gastric cancer [35,39,40]. Mexican patients exhibited somatic variants in known cancer-related genes, such as *TP53*, *CDH1*, and *ARID1A*. In concordance with reports in the Hispanic population [22], we observed a high frequency of *CDH1* mutations. The high frequency of *CDH1* could be due to the greater number of patients with diffuse-type GA included in our cohort. *CDH1* encodes the adhesion protein E-cadherin. Somatic mutations in *CDH1* have been associated with the epithelial–mesenchymal transition in tumor cells, resulting in a more aggressive tumor phenotype [41,42]. *CDH1* mutations are commonly associated with diffuse-type GA, and here, we observed the same tendency.

Due to the high frequency of diffuse-type GA reported for Hispanic patients, we decided to focus on comparing the mutational pattern of patients with diffuse- and intestinal-type GA. Patients with diffuse-type GA display mutations in *CDH1*, *RHOA,* and *CDKN2A*. *RHOA* encodes for the small GTPase family member Rho. *RHOA* mutations have been associated with cell survival and cell migration through the inactivation of ROCK [43]. In patients with lymphoma, *RHOA* mutations were assessed as a potential biomarker for poor prognosis [44], and gain-of-function mutations were studied as a possible therapeutic

target for gastric cancer [45]. All *RHOA* mutations identified in this study were missense mutations; however, due to the high frequency of diffuse-type GA in the Mexican population, further studies are required to identify their utility both as a prognosis biomarker as well as a therapeutic target in Mexican patients.

Amplifications of cytoband 20q13.33 were associated with diffuse-type gastric cancer. In colorectal cancer, amplifications in cytoband 20q13.33 were associated with early tumoral stages and mutations in *APC* and *KRAS* [46]. *LAMA5* was the most frequently affected gene within cytoband 20q13.33 in our cohort. In colorectal cancer, *LAMA5* overexpression is associated with hepatic metastasis, angiogenesis, and NOTCH1 pathway inhibition [47]. In ovarian cancer, *LAMA5* overexpression was associated with a good prognosis [48]. No information about the potential use of *LAMA5* as a biomarker for prognosis in gastric cancer has been reported. *PSMA7* and *GID8* are two genes encoded within 20q13.33 that were described as possible gene targets in gastric cancer [49]. The overexpression of *PSMA7* was associated with poor prognosis in patients with gastric cancer [50]. Identifying whether 20q13.33 is affected in the early stages of GA in Mexican patients, or which genes are affected by CNVs in this cytoband, requires further studies.

Patients with intestinal-type GA were associated with mutations in *NOTCH1.* NOTCH signaling has been widely dissected in gastric cancer tumorigenesis [51–53]. In colorectal cancer, increased NOTCH1 expression is associated with greater immune cell infiltration, tumor volume, and depth invasion [51,54]. In cervical cancer, the loss of nuclear NOTCH1 was reported to be an independent predictor of malignancy [55]. We identified that *NOTCH1* and *JAG2* were affected exclusively by deleterious mutations and mainly by CNV losses. We suspect that the NOTCH signaling pathway is underactive in intestinal-type GA; however, the clinical implications of these genomic variants are still to be elucidated.

A limitation of this study is the number of patients included in this cohort, which could restrict the statistical power of the analysis and hinder the frequency of genomic variants identified. We are aware that a small number of patients would restrain the representation of the true molecular landscape of gastric cancer of Mexican patients and could limit the statistical power to compare the molecular behavior of our cohort with other larger cohorts. However, in this first study using fresh samples from a prospective Mexican cohort, we reported the mutational landscape of GA and managed to identify associations between genomic variants and clinicopathological characteristics, such as tumor histology, signet-ring cell morphology, and *H. pylori* and EBV presence, which could guide us in the design of further studies that feature a larger cohort to test some of the identified variants as biomarkers of prognosis for patients treated with immune checkpoint inhibitors or other therapies.

The ethnicity of patients may influence the mutational landscape of GA. For example, it is reported that African American patients tend to present more *TP53* somatic variants than Asian, European-descent, and Hispanic patients [56]; European-descent patients tend to present deletions of the phosphatase gene *PTPRD* more frequently than Asian patients (16); and a high frequency of somatic variants of *APC*, *ARID1A*, *KMT2A*, *PIK3CA*, and *PTEN* has been found in Asian patients compared to European-descent patients [57]. Hispanics are one of the most genetically diverse ethnic groups [58,59]; however, GA somatic and structural variants in this ethnic group have not been thoroughly explored. A recent review published by our team presented the mutational landscape of Hispanic GA patients, identifying a differential distribution of the variants according to the countries [4]. Thus, Hispanics should not be considered a genetically homogeneous ethnic group in GA research. Here, we present a mutational landscape for GA in Mexican patients to better understand tumorigenesis in patients of Hispanic ethnicity and lay the groundwork for the development of potential biomarkers as well as therapeutic targets.

4. Materials and Methods

4.1. Population and Sample Collection

Patients included in this study were older than 18 years old, had a positive diagnosis of gastric adenocarcinoma, did not receive any type of treatment, and signed an informed consent prior to sample collection. Both tumor and non-tumor gastric tissue samples were collected for each patient by two trained endoscopists. The presence of tumoral cells in the tumor samples were confirmed by a trained pathologist. Samples were immediately stabilized after collection in RNAlater (Invitrogen, Waltham, MA, USA, AM7021) and stored for 24 h at 4 °C. Thereafter, RNAlater was removed, and the tissue samples were stored at −80 °C until DNA extraction.

4.2. Clinical Data Collection

Clinical data, including age at diagnosis, body mass index (BMI), sex, clinical stage, tumor location, endoscopic impression (Borrmann classification), tumor histology (Lauren classification), tumor grade, and the presence of signet-ring cells, were collected from the electronic clinical records of each patient.

4.3. DNA Extraction and Quality Control

Tumor and non-tumor samples were homogenized using QIAshredder columns (Qiagen, Germantown, MD, USA, 79654), and DNA was extracted using the AllPrep® DNA/RNA Mini Kit (Qiagen, Germantown, MD, USA, 80204), following the protocol recommendations [60]. DNA purity was evaluated using NanoDrop 2000. The quantity and DNA fragmentation status were evaluated with the Agilent 2200 TapeStation System using the genomic DNA ScreenTape assay (Agilent, Santa Clara, CA, USA, 5067-5365) to obtain the DNA integrity number (DIN). Samples with a DIN between 6 and 10 proceeded to whole-exome sequencing.

4.4. PCR

The presence of *H. pylori* was detected in samples with 100 ng of DNA from both the tumor tissue and adjacent tissue using the Step One Plus Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA). RealQ Plus 2× Master Mix (AMPLIQON, A323402) was used according to the manufacturer's instructions. The HP-16S primer set (5′-TCGGAATCACTGGGCGTAA-3′ and 5′-TTCTATGGTTAAGCCATAGGATTT CACAC-3′) was used to detect the presence of *H. pylori*. β-Globin was used as an internal control; the specific primers used were 5′-GACAGGTACGGCTGTCATCA-3′ and 5′-TAGATGGCTCTGCCCTGACT-3′.

4.5. In Situ Hybridization

The presence of EBV in tumoral tissue was assessed by in situ hybridization for the EBV-encoded RNA (EBER1), following the procedures previously described [61]. A tumor was considered EBV-negative if EBER1 staining was undetected or only expressed in lymphoid cells with a benign appearance. A tumor was considered positive if EBER1 staining was observed in the nucleus of malignant cells.

4.6. Library Preparation, Hybridization Capture, and WES

Library preparation, hybridization capture, and WES procedures were performed by Novogene Corporation Inc. (Sacramento, California, USA). Sequencing libraries from the extracted DNA were generated using the Agilent SureSelect Human All Exon Kit V6 (Agilent, Santa Clara, California, USA, 5190-8865) following the manufacturer's recommendations. Products were purified using the AMPure XP system (Beckman Coulter, Brea, California, USA, A63882) and quantified using an Agilent high sensitivity DNA assay on the Agilent Bioanalyzer 2100 system. The post-capture libraries were sequenced on the Illumina NovaSeq 6000 platform.

4.7. Bioinformatic Analysis Pipeline

Raw sequencing data were filtered to eliminate adapter-contaminated and low-quality reads. A read pair was discarded if either read contained adapter contamination, if more than 10% of bases were uncertain, or if the proportion of low-quality bases was below 50%. Burrows-Wheeler Aligner (BWA) (https://github.com/lh3/bwa accessed on 12 February 2021) (version [v] 0.7.8-r455) [62] was used to map the paired-end clean reads to the GRCh37 human reference genome. SAMtools (http://www.htslib.org accessed on 12 February 2021) (v1.0) [63] was used to sort the BAM file, and Picard (v1.111) [64] was used to mark and remove duplicated reads. Somatic variants (SNVs) were identified using three variant callers: MuTect (https://github.com/broadinstitute/mutect accessed on 12 February 2021) (v1.1.4), VarScan2 (https://github.com/Jeltje/varscan2 accessed on 12 February 2021) (v4.1) [65], and MuTect2 (https://gatk.broadinstitute.org/hc/en-us/articles/360037593851-Mutect2 accessed on 12 February 2021) (v4.2) [66]. Only the variants that passed all the filtering sets to the variant calling programs and labeled as "PASS" were selected. Only SNVs identified by two or more variant callers were selected for further analysis. Small insertions and deletions (InDels) were identified with Strelka (https://github.com/Illumina/strelka, accessed on 12 February 2021) (v1.0.13) [67], VarScan2, and MuTect2. InDels identified by two or more variant callers were selected. All the SNVs and InDels that had a minor allele frequency of 1% or higher in gnomAD (v2.0.1) were filtered out to eliminate potential germline SNP. In addition, to prevent any contamination with germline variants, we performed a paired analysis for each patient, filtering out all the mutations shared by the adjacent and the tumoral tissues and leaving only the mutations found in the tumoral tissue with a genotype 0/1 in the final VCF files; also, we filtered out variants with a variant allele frequency smaller than 0.05 and greater than 0.7. The biological and clinical relevance of SNVs was obtained using The Cancer Genome Interpreter (CGI) web interphase [68]. For predicting driver mutations, a tissue-specific model (stomach adenocarcinoma) was selected in the CGI. This model uses a machine learning algorithm named BoostDM to annotate the clinical and biological relevance of the somatic variants with a BoostDM score of 0.5 and an accuracy for predicting driver mutations (F_{50}-score) above 0.9 [69]. Control-FREEC (https://github.com/BoevaLab/FREEC accessed on 12 February 2021) (v11.4) [70] was used for the identification of CNVs. Segments with log2 greater than 0.2 were classified as CNV gains, and segments with log2 lower than -0.235 were CNV losses.

4.8. Tumor Mutational Burden

TMB was calculated using the total variants affecting the coding region. Samples with TMB equal or greater than 10 mutations/Mb were considered as having high TMB [71].

4.9. Enrichment Analysis

Enrichment analysis was performed using the Gene Ontology web interphase (http://geneontology.org, accessed on 12 March 2022) with the PANTHER classification system to identify the biological processes and the molecular pathways affected by somatic variants [26]. Fisher's exact test with the Benjamini–Hochberg procedure for *p*-value correction was used. The top-ten enriched pathways were selected.

4.10. Mutational Signatures

Mutational signatures' contribution to the mutational process was calculated for each sample using the R package deconstuctSigs (v1.9.0) [72]. COSMIC signatures were set as reference signatures, and the count method was set as default. Mutational signature weight ≥0.2 was considered as a positive contribution to the mutational process of the sample.

4.11. TCGA and OncoSG Data Acquisition

The open-access web resource, c-Bioportal, was used to access TCGA genomic data [73,74]. Samples from the stomach adenocarcinoma (TCGA, Firehouse Legacy) study (14) and from the Gastric Cancer (OncoSG, 2018) study (29) were selected. Somatic and structural variant

data were retrieved. A Venn diagram was used to observe similitudes and differences the number of genes affected by somatic variants and CNVs shared by TCGA, OncoSG, and our study. Then, the frequency of the ten most affected genes by somatic variants and CNVs in our cohort was presented in a bar plot to depict differences and similarities with the other cohorts.

4.12. Statistical Analysis

Associations between clinical characteristics such as: tumor histology and the presence of signet-ring cells, H. pylori, and EBV were assessed with the presence of somatic variations and CNVs. Patients were grouped according to their clinical characteristics, and then the frequency of the variants between the groups was compared using Fisher's exact test. *p*-values were adjusted using the Benjamini–Hochberg procedure. TMB raw values were also evaluated for association with the clinical characteristics of the patients using the Kruskal–Wallis test. The positive contribution of mutational signatures was categorized and compared with the clinical characteristics of the patients. Contingency tables were analyzed for independence using the chi-squared test with Yates' correction. Statistical significance was set at $p < 0.05$. All statistical analyses were performed using R software (https://cran.r-project.org, accessed on 28 July 2021).

5. Conclusions

We described a mutational landscape for Mexican patients with GA, exploring the differential mutational pattern between clinical features such as tumor histology and the presence of signet-ring cells, *H. Pylori*, and EBV infection. We focused our discussion on the potentially different mutational pattern observed between diffuse- and intestinal-type GA. Mexican patients with intestinal-type GA tended to present a high frequency of *NOTCH1* mutations, CNV gains in cytobands 13.14, 10q23.33, and 12q25.1, and CNV losses in cytobands 7p12, 14q24.2, and 11q13.1; whereas patients with diffuse-type GA tended to present a high frequency of *CDH1* mutations and CNV gains in cytobands 20q13.33 and 22q11.21. Whether these variants occur in other populations and their role in GA development and treatment require further investigation with a larger cohort.

References

1. Asplund, J.; Kauppila, J.H.; Mattsson, F.; Lagergren, J. Survival Trends in Gastric Adenocarcinoma: A Population-Based Study in Sweden. *Ann Surg Oncol.* **2018**, *25*, 2693–2702. [CrossRef]
2. Jim, M.A.; Pinheiro, P.S.; Carreira, H.; Espey, D.K.; Wiggins, C.L.; Weir, H.K. Stomach cancer survival in the United States by race and stage (2001–2009): Findings from the CONCORD-2 study. *Cancer* **2017**, *123*, 4994–5013. [CrossRef]
3. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef]
4. Cerrato-Izaguirre, D.; Chirino, Y.I.; García-Cuellar, C.M.; Santibáñez-Andrade, M.; Prada, D.; Hernández-Guerrero, A.; Larraga, O.A.; Camacho, J.; Sánchez-Pérez, Y. Mutational landscape of gastric adenocarcinoma in Latin America: A genetic approach for precision medicine. *Genes Dis.* **2022**, *9*, 928–940. [CrossRef]
5. Bonequi, P.; Meneses-González, F.; Correa, P.; Rabkin, C.S.; Camargo, M.C. Risk factors for gastric cancer in Latin America: A meta-analysis. *Cancer Causes Control* **2013**, *24*, 217–231. [CrossRef]
6. Lauren, P. The two histological main types of gastric carcinoma: Diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. *Acta Pathol. Microbiol. Scand.* **1965**, *64*, 31–49. [CrossRef]
7. Chang, E.T.; Gomez, S.L.; Fish, K.; Schupp, C.W.; Parsonnet, J.; DeRouen, M.C.; Keegan, T.H.; Clarke, C.A.; Glaser, S.L. Gastric cancer incidence among Hispanics in California: Patterns by time, nativity, and neighborhood characteristics. *Cancer Epidemiol. Biomarkers Prev.* **2012**, *21*, 709–719. [CrossRef]
8. Corral, J.E.; Delgado Hurtado, J.J.; Domínguez, R.L.; Valdez de Cuéllar, M.; Balmore Cruz, C.; Morgan, D.R. The descriptive epidemiology of gastric cancer in Central America and comparison with United States Hispanic populations. *J. Gastrointest. Cancer* **2015**, *46*, 21–28. [CrossRef]
9. Wang, R.; Song, S.; Harada, K.; Ghazanfari Amlashi, F.; Badgwell, B.; Pizzi, M.P.; Xu, Y.; Zhao, W.; Dong, X.; Jin, J.; et al. Multiplex profiling of peritoneal metastases from gastric adenocarcinoma identified novel targets and molecular subtypes that predict treatment response. *Gut* **2020**, *69*, 18–31. [CrossRef]
10. Wang, K.; Yuen, S.T.; Xu, J.; Lee, S.P.; Yan, H.H.; Shi, S.T.; Siu, H.C.; Deng, S.; Chu, K.M.; Law, S.; et al. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat. Genet.* **2014**, *46*, 573–582. [CrossRef]
11. Shu, Y.; Zhang, W.; Hou, Q.; Zhao, L.; Zhang, S.; Zhou, J.; Song, X.; Zhang, Y.; Jiang, D.; Chen, X.; et al. Prognostic significance of frequent CLDN18-ARHGAP26/6 fusion in gastric signet-ring cell cancer. *Nat. Commun.* **2018**, *9*, 2447. [CrossRef]
12. Tan, P.; Yeoh, K.G. Genetics and Molecular Pathogenesis of Gastric Adenocarcinoma. *Gastroenterology* **2015**, *149*, 1153–1162. [CrossRef]
13. Ma, J.; Shen, H.; Kapesa, L.; Zeng, S. Lauren classification and individualized chemotherapy in gastric cancer. *Oncol. Lett.* **2016**, *11*, 2959–2964. [CrossRef]
14. Network CGAR. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **2014**, *513*, 202–209. [CrossRef]
15. Liao, P.; Jia, F.; Teer, J.K.; Knepper, T.C.; Zhou, H.H.; He, Y.J.; McLeod, H.L. Geographic variation in molecular subtype for gastric adenocarcinoma. *Gut* **2019**, *68*, 1340–1341. [CrossRef]
16. Schumacher, S.E.; Shim, B.Y.; Corso, G.; Ryu, M.H.; Kang, Y.K.; Roviello, F.; Saksena, G.; Peng, S.; Shivdasani, R.A.; Bass, A.J.; et al. Somatic copy number alterations in gastric adenocarcinomas among Asian and Western patients. *PLoS ONE* **2017**, *12*, e0176045. [CrossRef]
17. Sohn, B.H.; Hwang, J.E.; Jang, H.J.; Lee, H.S.; Oh, S.C.; Shim, J.J.; Lee, K.W.; Kim, E.H.; Yim, S.Y.; Lee, S.H.; et al. Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. *Clin. Cancer Res.* **2017**, *23*, 4441–4449. [CrossRef]
18. Sanjeevaiah, A.; Cheedella, N.; Hester, C.; Porembka, M.R. Gastric Cancer: Recent Molecular Classification Advances, Racial Disparity, and Management Implications. *J. Oncol. Pract.* **2018**, *14*, 217–224. [CrossRef]
19. Holowatyj, A.N.; Ulrich, C.M.; Lewis, M.A. Racial/Ethnic Patterns of Young-Onset Noncardia Gastric Cancer. *Cancer Prev. Res.* **2019**, *12*, 771–780. [CrossRef]
20. Choi, A.H.; Ji, L.; Babcock, B.; Ramos, V.; Kwong, M.L.M.; Morgan, J.W.; Selleck, M.J.; Langridge, W.H.R.; DeLeon, M.; Wall, N.R.; et al. Peritoneal carcinomatosis in gastric cancer: Are Hispanics at higher risk? *J. Surg. Oncol.* **2020**, *122*, 1624–1629. [CrossRef]
21. Merchant, S.J.; Kim, J.; Choi, A.H.; Sun, V.; Chao, J.; Nelson, R. A rising trend in the incidence of advanced gastric cancer in young Hispanic men. *Gastric. Cancer* **2017**, *20*, 226–234. [CrossRef]

22. Wang, S.C.; Yeu, Y.; Hammer, S.T.G.; Xiao, S.; Zhu, M.; Hong, C.; Clemenceau, J.R.; Yoon, L.Y.; Nassour, I.; Shen, J.; et al. Hispanic/Latino Patients with Gastric Adenocarcinoma Have Distinct Molecular Profiles Including a High Rate of Germline. *Cancer Res.* **2020**, *80*, 2114–2124. [CrossRef]

23. Sánchez-Barriga, J.J. Tendencias de mortalidad y años potenciales de vida perdidos por cáncer gástrico en México, 2000–2012. *Rev. Gastroenterol. México* **2016**, *81*, 65–73. [CrossRef]

24. Moreno-Estrada, A.; Gignoux, C.R.; Fernández-López, J.C.; Zakharia, F.; Sikora, M.; Contreras, A.V.; Acuña-Alonzo, V.; Sandoval, K.; Eng, C.; Romero-Hidalgo, S.; et al. Human genetics. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* **2014**, *344*, 1280–1285. [CrossRef]

25. Romero-Hidalgo, S.; Ochoa-Leyva, A.; Garcíarrubio, A.; Acuña-Alonzo, V.; Antúnez-Argüelles, E.; Balcazar-Quintero, M.; Carnevale, A.; Cornejo-Granados, F.; Fernández-López, J.C. Demographic history and biologically relevant genetic variation of Native Mexicans inferred from whole-genome sequencing. *Nat. Commun.* **2017**, *8*, 1005. [CrossRef]

26. Mi, H.; Ebert, D.; Muruganujan, A.; Mills, C.; Albou, L.P.; Mushayamaha, T.; Thomas, P.D. PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **2021**, *49*, D394–D403. [CrossRef]

27. Alexandrov, L.B.; Nik-Zainal, S.; Wedge, D.C.; Aparicio, S.A.; Behjati, S.; Biankin, A.V.; Bignelli, G.R.; Bolli, N.; Borg, A.; Børresen-Dale, L.A.; et al. Signatures of mutational processes in human cancer. *Nature* **2013**, *500*, 415–421. [CrossRef]

28. Erkizan, H.V.; Sukhadia, S.; Natarajan, T.G.; Marino, G.; Notario, V.; Lichy, J.H.; Wadleigh, R.G. Exome sequencing identifies novel somatic variants in African American esophageal squamous cell carcinoma. *Sci. Rep.* **2021**, *11*, 14814. [CrossRef]

29. Guo, Y.A.; Chang, M.M.; Huang, W.; Ooi, W.F.; Xing, M.; Tan, P.; Skanderup, A.J. Mutation hotspots at CTCF binding sites coupled to chromosomal instability in gastrointestinal cancers. *Nat. Commun.* **2018**, *9*, 1520. [CrossRef]

30. Sánchez, Y.; Vaca-Paniagua, F.; Herrera, L.; Oñate, L.; Herrera-Goepfert, R.; Navarro-Martínez, G.; Cerrato, D.; Díaz-Velázquez, C.; Quezada, E.M.; García-Cuellar, C.; et al. Nutritional Indexes as Predictors of Survival and Their Genomic Implications in Gastric Cancer Patients. *Nutr. Cancer* **2021**, *73*, 1429–1439. [CrossRef]

31. Martínez-Galindo, M.G.; Zamarripa-Dorsey, F.; Carmona-Castañeda, A.; Angeles-Labra, A.; Peñavera-Hernández, R.; Ugarte-Briones, C.; Blanco-Vela, C.I. Histopathologic characteristics of gastric adenocarcinoma in Mexican patients: A 10-year experience at the Hospital Juárez de México. *Rev. Gastroenterol. México* **2015**, *80*, 21–26. [CrossRef] [PubMed]

32. Trujillo Rivera, A.; Sampieri, C.L.; Morales Romero, J.; Montero, H.; Acosta Mesa, H.G.; Cruz Ramírez, N.; Novoa Del Toro, E.M.; León Córdoba, K. Risk factors associated with gastric cancer in Mexico: Education, breakfast and chili. *Rev. Esp. Enferm. Dig.* **2018**, *110*, 372–379. [CrossRef]

33. Folprecht, G. Tumor mutational burden as a new biomarker for PD-1 antibody treatment in gastric cancer. *Cancer Commun.* **2019**, *39*, 74. [CrossRef] [PubMed]

34. Kim, J.; Kim, B.; Kang, S.Y.; Heo, Y.J.; Park, S.H.; Kim, S.T.; Kang, W.K.; Lee, J.; Kim, K.M. Tumor Mutational Burden Determined by Panel Sequencing Predicts Survival After Immunotherapy in Patients With Advanced Gastric Cancer. *Front. Oncol.* **2020**, *10*, 314. [CrossRef]

35. Cai, H.; Jing, C.; Chang, X.; Ding, D.; Han, T.; Yang, J.; Lu, Z.; Hu, X.; Liu, Z.; Wang, J.; et al. Mutational landscape of gastric cancer and clinical application of genomic profiling based on target next-generation sequencing. *J. Transl. Med.* **2019**, *17*, 189. [CrossRef] [PubMed]

36. Zhao, Z.; Li, W.; Zhang, X.; Ge, M.; Song, C. Correlation between TMB and MSI in patients with solid tumors. *J. Clin. Oncol.* **2020**, *38*, e15169. [CrossRef]

37. Mosele, F.; Remon, J.; Mateo, J.; Westphalen, C.B.; Barlesi, F.; Lolkema, M.P.; Normanno, N.; Scarpa, A.; Robson, M.; Meric-Bernstam, F.; et al. Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: A report from the ESMO Precision Medicine Working Group. *Ann. Oncol.* **2020**, *31*, 1491–1505. [CrossRef]

38. Chung, H.C.; Ros, W.; Delord, J.P.; Perets, R.; Italiano, A.; Shapira-Frommer, R.; Manzuk, L.; Piha-Paul, S.A.; Xu, L.; Zeigenfuss, S.; et al. Efficacy and Safety of Pembrolizumab in Previously Treated Advanced Cervical Cancer: Results From the Phase II KEYNOTE-158 Study. *J. Clin. Oncol.* **2019**, *37*, 1470–1478. [CrossRef]

39. Nemtsova, M.V.; Kalinkin, A.I.; Kuznetsova, E.B.; Bure, I.V.; Alekseeva, E.A.; Bykov, I.I.; Khorobrykh, T.V.; Mikhaylenko, D.S.; Tanas, A.S.; Kutsev, S.I.; et al. Clinical relevance of somatic mutations in main driver genes detected in gastric cancer patients by next-generation DNA sequencing. *Sci. Rep.* **2020**, *10*, 504. [CrossRef]

40. Pan, X.; Ji, X.; Zhang, R.; Zhou, Z.; Zhong, Y.; Peng, W.; Sun, N.; Xu, X.; Xia, L.; Li, P.; et al. Landscape of somatic mutations in gastric cancer assessed using next-generation sequencing analysis. *Oncol. Lett.* **2018**, *16*, 4863–4870. [CrossRef]

41. Zhao, H.; Hu, H.; Chen, B.; Xu, W.; Zhao, J.; Huang, C.; Xing, Y.; Lv, H.; Nie, C.; Wang, J.; et al. Overview on the Role of E-Cadherin in Gastric Cancer: Dysregulation and Clinical Implications. *Front. Mol. Biosci.* **2021**, *8*, 689139. [CrossRef] [PubMed]

42. Shenoy, S. CDH1 (E-Cadherin) Mutation and Gastric Cancer: Genetics, Molecular Mechanisms and Guidelines for Management. *Cancer Manag. Res.* **2019**, *11*, 10477–10486. [CrossRef] [PubMed]

43. Nam, S.; Kim, J.H.; Lee, D.H. RHOA in Gastric Cancer: Functional Roles and Therapeutic Potential. *Front. Genet.* **2019**, *10*, 438. [CrossRef] [PubMed]

44.   Lage, L.A.d.P.C.; Barreto, G.C.; Culler, H.F.; Cavalcanti, J.B.; Alves, L.d.O.; Nardinelli, L.; Bendit, I.; Rocha, V.; Pereira, J. Rhoa Mutation Is a Potential Biomarker Associated with Adverse Prognosis and High- Tumor Burden in Patients with Nodal Peripheral Lymphomas with T-Helper Follicular Phenotype (nPTCL-Thf): Data from a Brazilian Retrospective Cohort of Nodal PTCL. *Blood* **2021**, *138*, 4482. [CrossRef]

45.   Kim, J.H.; Park, S.; Lim, S.M.; Eom, H.J.; Balch, C.; Lee, J.; Jeong, J.H.; Nam, S.; Kim, Y.H. Rational design of small molecule RHOA inhibitors for gastric cancer. *Pharm. J.* **2020**, *20*, 601–612. [CrossRef]

46.   Bui, V.M.; Mettling, C.; Jou, J.; Sun, H.S. Genomic amplification of chromosome 20q13.33 is the early biomarker for the development of sporadic colorectal carcinoma. *BMC Med. Genom.* **2020**, *13*, 149. [CrossRef]

47.   Gordon-Weeks, A.; Lim, S.Y.; Yuzhalin, A.; Lucotti, S.; Vermeer, J.A.; Jones, K.; Chen, J.; Muschel, R.J. Tumour-Derived Laminin α5 (LAMA5) Promotes Colorectal Liver Metastasis Growth, Branching Angiogenesis and Notch Pathway Inhibition. *Cancers* **2019**, *11*, 630. [CrossRef]

48.   Diao, B.; Yang, P. Comprehensive Analysis of the Expression and Prognosis for Laminin Genes in Ovarian Cancer. *Pathol. Oncol. Res.* **2021**, *27*, 1609855. [CrossRef]

49.   Cheng, L.; Wang, P.; Yang, S.; Yang, Y.; Zhang, Q.; Zhang, W.; Zhang, W.; Xiao, H.; Gao, H. Identification of genes with a correlation between copy number and expression in gastric cancer. *BMC Med. Genom.* **2012**, *5*, 14. [CrossRef]

50.   Xia, S.; Tang, Q.; Wang, X.; Zhang, L.; Jia, L.; Wu, D.; Xu, P.; Zhang, X.; Tang, G.; Yang, T.; et al. Overexpression of PSMA7 predicts poor prognosis in patients with gastric cancer. *Oncol. Lett.* **2019**, *18*, 5341–5349. [CrossRef]

51.   Yao, Y.; Ni, Y.; Zhang, J.; Wang, H.; Shao, S. The role of Notch signaling in gastric carcinoma: Molecular pathogenesis and novel therapeutic targets. *Oncotarget* **2017**, *8*, 53839–53853. [CrossRef] [PubMed]

52.   Katoh, M. Precision medicine for human cancers with Notch signaling dysregulation (Review). *Int. J. Mol. Med.* **2020**, *45*, 279–297. [CrossRef] [PubMed]

53.   Du, X.; Cheng, Z.; Wang, Y.H.; Guo, Z.H.; Zhang, S.Q.; Hu, J.K.; Zhou, Z.G. Role of Notch signaling pathway in gastric cancer: A meta-analysis of the literature. *World J. Gastroenterol.* **2014**, *20*, 9191–9199. [CrossRef] [PubMed]

54.   Hu, J.; Yu, J.; Gan, J.; Song, N.; Shi, L.; Liu, J.; Zhang, Z.; Du, J. Notch1/2/3/4 are prognostic biomarker and correlated with immune infiltrates in gastric cancer. *Aging* **2020**, *12*, 2595–2609. [CrossRef] [PubMed]

55.   Vázquez-Ulloa, E.; Ramos-Cruz, A.C.; Prada, D.; Avilés-Salas, A.; Chávez-Blanco, A.D.; Herrera, L.A.; Lizano, M.; Contreras-Paredes, A. Loss of nuclear NOTCH1, but not its negative regulator NUMB, is an independent predictor of cervical malignancy. *Oncotarget* **2018**, *9*, 27. [CrossRef]

56.   Van Beek, E.J.A.H.; Hernandez, J.M.; Goldman, D.A.; Davis, J.L.; McLoughlin, K.C.; Ripley, R.T.; Kim, T.S.; Tang, L.H.; Hechtman, J.F.; Zheng, J.; et al. Correction to: Rates of TP53 Mutation are Significantly Elevated in African American Patients with Gastric Cancer. *Ann. Surg. Oncol.* **2020**, *27*, 963. [CrossRef]

57.   Jia, F.; Teer, J.K.; Knepper, T.C.; Lee, J.K.; Zhou, H.H.; He, Y.J.; McLeod, H.L. Discordance of Somatic Mutations Between Asian and Caucasian Patient Populations with Gastric Cancer. *Mol. Diagn. Ther.* **2017**, *21*, 179–185. [CrossRef]

58.   Conomos, M.P.; Laurie, C.A.; Stilp, A.M.; Gogarten, S.M.; McHugh, C.P.; Nelson, S.C.; Sofer, T.; Fernández-Rhodes, L.; Justice, A.E.; Graff, M.; et al. Genetic Diversity and Association Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos. *Am. J. Hum. Genet.* **2016**, *98*, 165–184. [CrossRef]

59.   González Burchard, E.; Borrell, L.N.; Choudhry, S.; Naqvi, M.; Tsai, H.-J.; Rodriguez-Santana, J.R.; Chapela, R.; Rogers, S.D.; Mei, R.; Rodriguez-Cintron, W.; et al. Latino Populations: A Unique Opportunity for the Study of Race, Genetics, and Social Environment in Epidemiological Research. *Am. J. Public Health* **2005**, *95*, 2161–2168. [CrossRef]

60.   Peña-Llopis, S.; Brugarolas, J. Simultaneous isolation of high-quality DNA, RNA, miRNA and proteins from tissues for genomic applications. *Nat. Protoc.* **2013**, *8*, 2240–2255. [CrossRef]

61.   Chang, K.L.; Chen, Y.-Y.; Shibata, D.; Weiss, L.M. Description of an In Situ Hybridization Methodology for Detection of Epstein-Barr Virus RNA in Paraffin-Embedded Tissues, with a Survey of Normal and Neoplastic Tissues. *Diagn. Mol. Pathol.* **1992**, *1*, 246–255. [CrossRef] [PubMed]

62.   Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **2010**, *26*, 589–595. [CrossRef] [PubMed]

63.   Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Subgroup, G.P.D.P. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef] [PubMed]

64.   McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, 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] [PubMed]

65.   Koboldt, D.C.; Zhang, Q.; Larson, D.E.; Shen, D.; McLellan, M.D.; Lin, L.; Miller, C.A.; Mardis, E.R.; Ding, L.; Wilson, R.K. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **2012**, *22*, 568–576. [CrossRef]

66.   Benjamin, D.; Sato, T.; Cibulskis, K.; Getz, G.; Stewart, C.; Lichtenstein, L. Calling Somatic SNVs and Indels with Mutect2. *bioRxiv* **2019**, 861054. [CrossRef]

67.   Saunders, C.T.; Wong, W.S.; Swamy, S.; Becq, J.; Murray, L.J.; Cheetham, R.K. Strelka: Accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **2012**, *28*, 1811–1817. [CrossRef]

68. Tamborero, D.; Rubio-Perez, C.; Deu-Pons, J.; Schroeder, M.P.; Vivancos, A.; Rovira, A.; Tusquets, I.; Albanell, J.; Rodon, J.; Tabernero, J.; et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* **2018**, *10*, 25. [CrossRef]

69. Muiños, F.; Martínez-Jiménez, F.; Pich, O.; Gonzalez-Perez, A.; Lopez-Bigas, N. In silico saturation mutagenesis of cancer genes. *Nature* **2021**, *596*, 428–432. [CrossRef]

70. Boeva, V.; Popova, T.; Bleakley, K.; Chiche, P.; Cappo, J.; Schleiermacher, G.; Janoueix-Lerosey, I.; Delattre, O.; Barillot, E. Control-FREEC: A tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* **2012**, *28*, 423–425. [CrossRef]

71. Chalmers, Z.R.; Connelly, C.F.; Fabrizio, D.; Gay, L.; Ali, S.M.; Ennis, R.; Schrock, A.; Campbell, B.; Shlien, A.; Chmielecki, J.; et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **2017**, *9*, 34. [CrossRef] [PubMed]

72. Rosenthal, R.; McGranahan, N.; Herrero, J.; Taylor, B.S.; Swanton, C. DeconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **2016**, *17*, 31. [CrossRef] [PubMed]

73. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Schultz, N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [CrossRef] [PubMed]

74. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2012**, *2*, 401–404. [CrossRef] [PubMed]