

## RESEARCH ARTICLE

# Human papillomavirus genotyping using next generation sequencing (NGS) in cervical lesions: Genotypes by histologic grade and their relative proportion in multiple infections

Jorge Alejandro Basiletti<sup>1\*</sup>, Joan Valls<sup>2,3</sup>, Tomás Poklépovich<sup>4</sup>, María Dolores Fellner<sup>1</sup>, Maryluz Rol<sup>2</sup>, Rafael Alonso<sup>5</sup>, Rita Mariel Correa<sup>1</sup>, María Celeste Colucci<sup>1</sup>, Mercedes Rodríguez de la Peña<sup>6</sup>, Paula Gabriela Falabella<sup>6</sup>, Agustina Saíno<sup>7</sup>, Josefina Campos<sup>4</sup>, Rolando Herrero<sup>8</sup>, Maribel Almonte<sup>2</sup>, María Alejandra Picconi<sup>1</sup>

**1** Servicio Virus Oncogénicos, Laboratorio Nacional y Regional de Referencia de HPV, Instituto Nacional de Enfermedades Infecciosas-ANLIS "Dr. Malbrán", Buenos Aires, Argentina, **2** Early Detection, Prevention & Infection Branch, International Agency for Research on Cancer, World Health Organization, Lyon, France, **3** Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain, **4** Unidad Operativa Centro de Genómica y Bioinformática, ANLIS "Dr. Malbrán", Buenos Aires, Argentina, **5** Departamento de Métodos Cuantitativos, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay, **6** Servicio de Ginecología, Hospital Nacional "Prof. Posadas", El Palomar, Provincia de Buenos Aires, Argentina, **7** Servicio de Anatomía Patológica, Hospital Nacional "Prof. Posadas", El Palomar, Provincia de Buenos Aires, Argentina, **8** Agencia Costarricense de Investigaciones Biomédicas, San José, Costa Rica

\* [jbasiletti@gmail.com](mailto:jbasiletti@gmail.com)



## OPEN ACCESS

**Citation:** Basiletti JA, Valls J, Poklépovich T, Fellner MD, Rol M, Alonso R, et al. (2022) Human papillomavirus genotyping using next generation sequencing (NGS) in cervical lesions: Genotypes by histologic grade and their relative proportion in multiple infections. PLoS ONE 17(11): e0278117. <https://doi.org/10.1371/journal.pone.0278117>

**Editor:** Cecilia Acuti Martellucci, University of Ferrara: Università degli Studi di Ferrara, ITALY

**Received:** May 31, 2022

**Accepted:** November 10, 2022

**Published:** November 23, 2022

**Copyright:** © 2022 Basiletti et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting Information](#) file.

**Funding:** MAP Grant CRA-PRI International Agency for Cancer Research <https://www.iarc.who.int/> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript MAP Grant 2016-2018 Instituto Nacional del Cancer (Argentina) <https://www.argentina.gob.ar/salud/inc> The funders had no role

## Abstract

Sensitive and specific genotyping of human papillomaviruses (HPVs) is critical for the surveillance and monitoring of the vaccine effectiveness. Here, HPV genotypes were identified in 137 cervical samples with different histology (79  $\leq$ CIN1 and 58 CIN3+) using Nested-PCR followed by Next-Generation sequencing (NGS) and relative proportions for each genotype in multiple infections were computed. All samples had been previously genotyped by PCR-Reverse Blotting Hybridization (PCR-RBH) thus allowing for a concordance analysis between both techniques. Multiple infections were present in 85% of  $\leq$ CIN1 cases compared to only 41% in CIN3+ cases ( $p < 0.001$ ). Among  $\leq$ CIN1 cases a towering genotypic diversity was observed, considering both low (LR-) and high risk (HR-) HPV genotypes; while among CIN3+, diversity was lower, HR-HPVs prevailing in most cases, especially HPV16. Furthermore, the predominance of HR-HPV genotypes in the proportions identified in each sample was higher in CIN3+ cases [(HPV16 (62.5%), followed by HPV31 and HPV58 (8.3% each)], than in  $\leq$ CIN1 cases [(HPV16 (17.7%), followed by HPV52 (14.7%) and HPV31 (10.3%)]. Agreement between PCR-RBH and NGS was higher than 90% for all genotypes (with an overall Kappa of 0.7), even though NGS identified eighty-nine positive results for HPV genotypes that had not been detected by PCR-RBH, evidencing its greater sensitivity. These results suggest that a reduction in genotypic diversity and/or an increase in the relative proportion of HR-HPVs in multiple infections can be considered as a bio-marker for the potential risk of malignant progression.

in study design, data collection and analysis, decision to publish, or preparation of the manuscript MAP Grant PICT 0364-2016 Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación <http://www.agencia.mincyt.gob.ar/> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

More than 200 genotypes of Human Papillomavirus (HPV) have been completely characterized, and new HPV types are still being found [1, 2]. Among them, 40 genotypes are known to infect the anogenital and aerodigestive tracts, and are grouped into high risk (HR)-HPV and low risk (LR)-HPV types based on their oncogenic potential [3, 4].

Classification of HPVs is based on the nucleotide sequence homology of the L1 gene, which is the most conserved region of the viral genome. Within the family, different genera share less than 60% nucleotide similarity. Within each genus, different species share similarity between 60% and 70%. Below the species level, a novel HPV type shares less than 90% similarity to any other type [1, 5]. The definition of a variant lineage is that the L1 open-reading frame differs by more than 1%, but less than the 10% that would make it another HPV type [6]. The International Agency for Research on Cancer (IARC) classifies twelve HPV genotypes as *carcinogenic* (HR-HPV) (Group 1: HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59), whereas HPV68 is *probably carcinogenic* (probable HR-HPV) (Group 2A) and HPVs 26, 53, 66, 73 and 82 are *possibly carcinogenic* (possible HR-HPV) (Group 2B); the so-called low-risk (LR) types (HPV6, 11, 40, 42, 43, 44, 55, 61, 81, 83) show strong evidence that they do not cause cancer (Group 4) [4, 7, 8].

Persistent HR-HPV infection is the main cause of cervical cancer (CC) [9], with HPV16 and HPV18 persistence accounting for over 70% of cases [10–15]. HPV has also been detected in other anatomic sites, causing oropharyngeal, vulvar, vaginal, penile and anal cancers [9, 16–18].

Viral genotyping for epidemiological and surveillance studies and vaccine trials needs highly sensitive, high-throughput and affordable HPV detection assays. Consensus or degenerate primers are widely used for broad-spectrum HPV PCR amplification. Their high sensitivity and robust format have allowed their wide use globally. However, they have some limitations, mainly due to the different performance of each genotype amplification, especially in the case of multiple infections, since the viruses present in a lower proportion could result in false negatives. The methods used for genotyping the amplicons often count on hybridization of individual type-specific probes (corresponding to the most relevant genotypes already characterized), allowing multiple types to be detected simultaneously despite a wide variation in the viral load and amplification bias [19–24].

DNA sequencing by Sanger has the potential to identify a wide range of HPV types in simple infections. However, in multiple infections, its reading accuracy decreases due to the overlapping of electropherograms, and requires prior cloning, which is laborious and time-consuming; in addition, minority genotypes that are part of mixed infections might not be detected.

More recently, the emergence of Next Generation Sequencing (NGS) has brought an innovative and attractive solution since it allows to detect a larger spectrum of HPV genotypes. This high-throughput methodology relies on massive and parallel sequencing of single molecules from small amounts of DNA. Therefore, it provides an enormous amount of information, allowing to identify all the genotypes present in a sample, regardless of the viral load and the characteristics of its sequence.

Although the prevalence and distribution of HPV genotypes in the cervix have been extensively studied, deep genomic analyses in cervical lesions with a different histological severity, regarding the diversity, abundance, and characteristics of the HPV genotypes are still scarce. NGS reports have shown that previous studies applying more limited techniques underestimated certain viral types and missed new genotypes, which could have an impact on pathogenesis, molecular epidemiology, virological surveillance and eventually on screening using HPV tests, particularly in vaccinated populations [25–27].

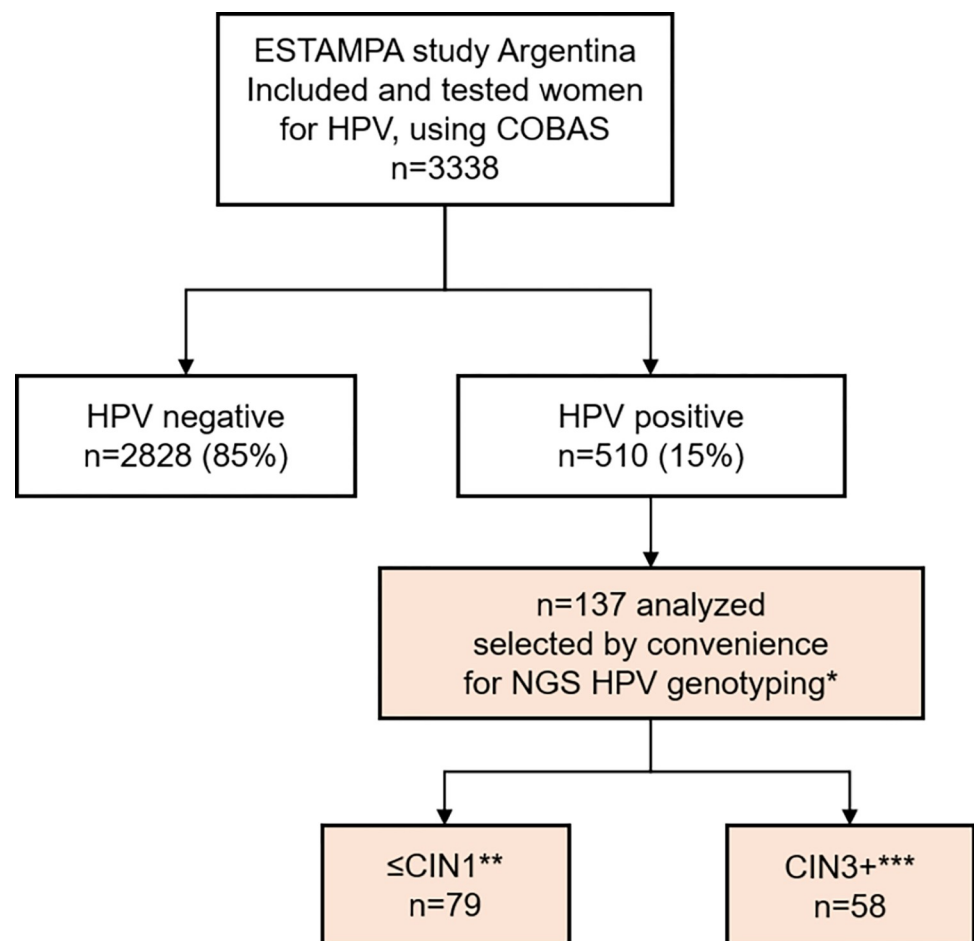
Moreover, the benefit is even greater when analysing multiple infections, whose implications in cervical carcinogenesis are still controversial [28–31]. In samples with coinfections, where genotype identification is complex and occasionally imprecise, NGS provides a powerful lens through which to look at viral communities in detail that would be unnoticeable with conventional methods, allowing to evaluate the proportion of each genotype in the viral mixture. Some authors have determined a greater average abundance of HR-HPVs in HSIL [25], but data are still limited in histology-confirmed cervical lesions.

This study aimed to identify HPV genotypes with NGS in cervical samples with different histologies, and to determine the relative proportion of each genotype within each sample in multiple infections.

## Materials and methods

### Study design

This is a cross-sectional, observational study, ancillary of the multicentric ESTAMPA study (Spanish acronym for “ESTudio multicéntrico de TAMizaje y triaje del cáncer cervicouterino con pruebas de virus PApiloma humano”) [32] to assess the HPV-genotype distribution using NGS including a series of 137 women selected from Argentina (Fig 1).



**Fig 1. Selection of the study population.** \* All 137 samples had previously been genotyped for PCR/RBH. \*\* ≤CIN1: Normal Cytology (Negative for Intraepithelial lesion or Malignancy, NILM) and CIN1. \*\*\* CIN3+: CIN3 and CC.

<https://doi.org/10.1371/journal.pone.0278117.g001>

## Clinical samples

Our series for the present study included 137 cervical samples collected at the initial visit of ESTAMPA Argentina, classified into two groups according to histology results: (1)  $\leq$ CIN1 ( $n = 79$ ), which included 39 normal cytology (NILM) or negative histology and 40 CIN1, selected by convenience sampling and (2) CIN3+ ( $n = 58$ ), which included 54 CIN3 and 4 CC (all the CIN3+ diagnosed were considered in this group).

All samples had been previously genotyped separately by two PCR strategies combined with Reverse Blotting Hybridization (RBH) using type-specific biotinylated probe sets [33]: i) Broad-Spectrum General Primers 5+/6+ PCR (BSGP) [24, 34] and ii) PGMY9/11 (CHUV) [23, 35], which allow to amplify highly conserved 140 bp and 450 bp fragments of the HPV-L1 gene, respectively [45]. These “in house” techniques coupled PCR amplification to hybridization to specific probes for 37 anogenital HPV genotypes (LR and HR) immobilized on a nylon membrane. These assays have been analytically validated by their developers [23, 24, 34, 35] and verified by us upon implementation in our laboratory; they are usually used in many laboratories for HPV genotyping due to their highly valuable sensitive for HPV epidemiological studies and surveillance [33, 36–40]. The Regional HPV Reference Laboratory from Argentina annually participates in the external control of HPV genotyping (HPV DNA typing proficiency panel), provided by the International Papillomavirus Reference Centre, which adds reliability to the results obtained [41].

## HPV genotyping by NGS

**Analytical characterization.** To determine the sensitivity and specificity of the NGS genotyping protocol applied in this study, 20 ng of genomic DNA + 5, 50 and 500 copies per 5ul of HPV16, HPV18, and 50 and 500 copies per 5ul of HPV31, HPV33, HPV35, HPV39, HPV45, HPV51, HPV58, HPV59, HPV68, HPV6 and HPV11 plasmid and HPV negative cell lines (C33A cells) were used as reference samples. The different HPV types were combined in various amounts to mimic multiple infections.

**PCR amplification.** Was performed by a nested PGMY/GP primer set, using 5ul of each plasmid sample and 20 ng of DNA previously extracted from cervical cells collected in Thin-Prep PreservCyt Solution (Hologic, Bedford, MA, USA) with commercial columns on a robotic system (QIAcube system, Qiagen). Briefly, two successive rounds of PCR were carried out to obtain a final target sequence of about 200 bp length of the L1 gen. The first round of PCR with a mixture of primers for PGMY11/09 [20] and the second round with nested HPV PCR primers GP5+/GP6+ [24, 34, 42].

**HPV DNA sample library preparation.** Amplicons were purified using DNA Clean & Concentrator (Zymo Research Corp.), according to manufacturer’s instructions. Purified DNA concentration was determined using a Qubit fluorometer (Life Technologies Corporation, Carlsbad, CA, USA). DNA libraries were prepared using an Illumina-compatible kit (NEBNext® Ultra™ II DNA Library Prep Kit, NEB, Frankfurt am Main, Germany), following the manufacturer’s recommendations (New England BioLabs, Herts, UK). Briefly, purified amplicons were end-repaired, adaptor-ligated, and cleaned up. Subsequently, DNA libraries obtained from each sample were amplified by nested-PCR with different index primers and then purified using Am-Pure XP beads (Beckman Coulter, Porterville, CA, USA). DNA libraries with different index primers were pooled in equal amounts.

**Sequencing using the MiSeq platform.** Paired-end (150×2) deep sequencing was performed using the MiSeq v2 reagent kit (Illumina, Inc., San Diego, CA, USA) on the MiSeq platform (Illumina, Inc.) with a standard protocol.

**Bioinformatics.** The FASTQ data went through a first processing to exclude the low quality reads (Q-score <30) and the adapter sequences, by cutting regions with the FastQC [43]

and Trimgalore software [44]. The presence of HPV and potential contaminants reads, such as human DNA, were identified with Kraken 2 software [45]. Trimmed sequences were mapped into human genome hg38 using Bowtie [46]. Reads with at least 95% identity and 75% coverage of human DNA were removed. Only unmapped paired-end reads were assembled to generate about 180 bp L1 contigs using Trinity software [47] and then, consensus sequences were generated from contigs with high similarity (pairwise distance  $\leq 3\%$ ). Consensus L1 sequences were aligned to L1 region of HPV reference sequences obtained from the NCBI and Papillomavirus Episteme PaVE databases (<http://pave.niaid.nih.gov>) using Bowtie [46]. Only the consensus sequences that, when compared to the PaVE base, showed at least 90% identity and  $>70\%$  coverage, were considered positive for one HPV genotype.

HPV genotypes were identified for each sample (as present or absent) and their relative presence further quantified based on the number of reads matched to each consensus genotype in relation to the total number of reads, allowing to evaluate the relative proportions of each genotype within a sample. The weeSAM software [48] was used to count the mapped reads on each fragment.

Therefore, absolute frequencies for each genotype were used first to compute their prevalence within histological grade. Second, the proportion of each genotype present in the samples with multiple infections was determined as the ratio of the number of reads corresponding to each genotype by the total number of reads in each sample. Third, the prevalence of predominant genotypes, i.e. that with the highest observed proportion, among samples where the genotype was present, was computed within histologic grade. In addition, the proportions were also used to compute a Gini index as a measure of genotypic diversity for each sample; a heatmap was produced to visualize these proportions within samples by histologic grade and sorting them out with the Gini index.

NGS genotyping results were compared with those previously obtained by PCR-RBH in the same samples. This analysis was restricted to the 37 HPV genotypes detectable by all RBH probes.

## Ethical considerations

The ESTAMPA protocol was approved by the Ethics Committee of the International Agency for Research on Cancer (IEC Project 12–27-A7), the Pan American Health Organization (PAHO) Ethical Committee and Ethical Committees of the Hospital Nacional “Prof. Posadas” (Argentina). The participants signed an informed consent that included details on the background, study procedures, risks and benefits, statement of confidentiality, specimen use and study staff to contact.

## Statistical analyses

A chi-square test was used to compare the distribution of single and multiple infections by lesion severity. A Student t-test was used to assess whether genotypic diversity, as measured with the Gini index, was associated to histologic grade. Cohen’s Kappa was used to globally assess the concordance between NGS and PCR-RBH-genotyping in the positivity for any genotype. To further investigate the concordance at the genotype level, an overall agreement was calculated for each HPV type as the sum of the number of specimens positive by both assays plus the number of negative ones by both assays, divided by the total number of specimens tested multiplied by 100, using McNemar test to evaluate the differences. In addition, positive agreement was also presented, i.e. excluding those that were negative for both tests. Threshold for statistical significance was set at  $p < 0.05$ .

## Results

One hundred thirty-seven cervical samples were sequenced by NGS and bioinformatically analysed for HPV detection.

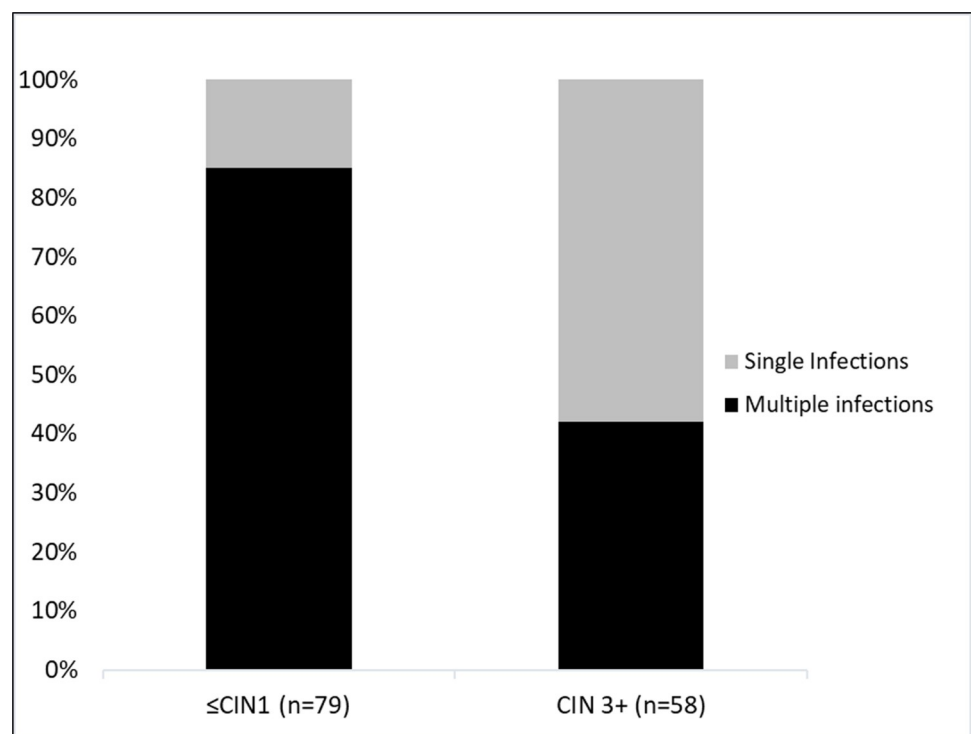
When using a cut-off value of 50 reads, HPV types in plasmid samples and HPV negative cell lines were correctly detected and typed with NGS at different DNA amounts (5, 50 and 500 copies/5 $\mu$ L for HPV16 and HPV18; 50 and 500 copies/5 $\mu$ L for the other HPV types). Therefore, a threshold of 50 reads was set for positive samples. Furthermore, all samples showed complete agreement for all types, even when containing multiple infections.

The HPV aligned reads in HPV positive samples showed a median of 1267 (minimum: 50; maximum: 31983) in the  $\leq$ CIN1 group, while in the CIN3+ group the median was 5507 (minimum: 51; maximum: 26783) (S1 Table).

Fig 2 illustrates the distribution of single and multiple infections by lesion severity. In the  $\leq$ CIN1 group, multiple infections stand for 85% of cases, whereas in the CIN3+ cases they encompass 41% ( $p < 0.001$ ); therefore, fewer multiple infections were detected in the higher severity cervical lesions.

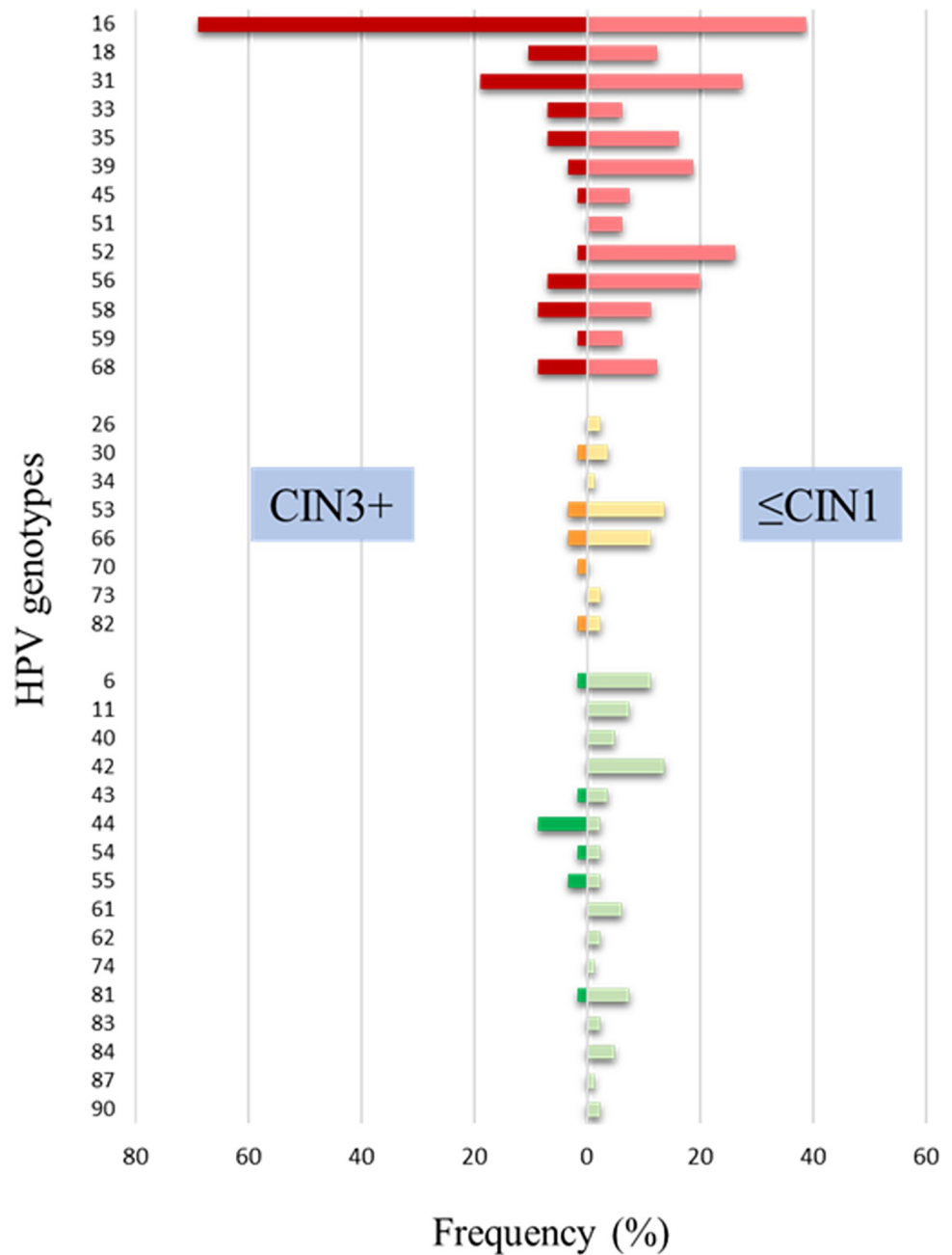
Fig 3 details the frequencies of the 37 genotypes identified by NGS for each histological category. In case of multiple infections, each genotype was counted independently. HPV16 was the most prevalent genotype in both groups. The ten more frequent genotypes, in descending order were: for  $\leq$ CIN1: HPV16 (38.8%), HPV31 (27.5%), HPV52 (26.3%), HPV56 (20%), HPV39 (18.8%), HPV35 (16.3%), HPV18 (12.5%), HPV68 (12.5%), HPV58 (11.3%), HPV45 (7.5%), and for CIN3+: HPV16 (69%), HPV31 (19%), HPV53 (13.8%), HPV66 (11.3%), HPV18 (10.3%), HPV58 (8.6%), HPV68 (8.6%), HPV33 (6.9%), HPV35 (6.9%), and HPV56 (6.9%).

Our results showed a higher genotype diversity in the  $\leq$ CIN1 group, both HR and LR-HPV, while in CIN3+; the viral spectrum was more restricted, detecting almost exclusively HR-HPVs, particularly HPV16.



**Fig 2. Distribution of single and multiple infections, according to histology.**

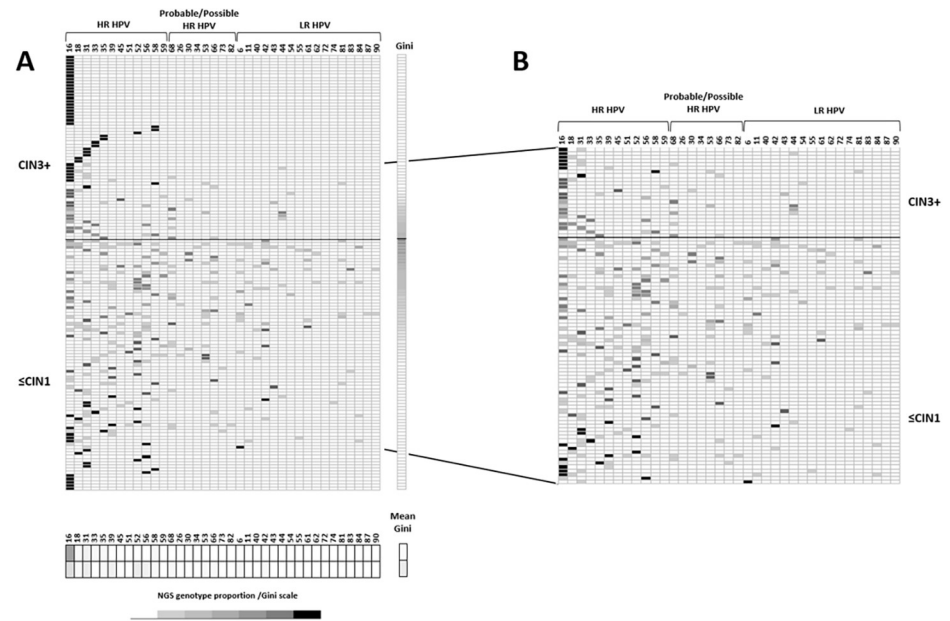
<https://doi.org/10.1371/journal.pone.0278117.g002>



**Fig 3. Distribution of HPV genotypes by severity of the lesion.** HR-HPVs (Group 1 IARC) are in red, possible or probable HR-HPVs (Group 2A/2B IARC) are in orange and low-risk HPV (Group 4 IARC) are in green.

<https://doi.org/10.1371/journal.pone.0278117.g003>

The analysis of the relative proportion of the genotypes identified in each sample is detailed in the heatmap (Fig 4A). According to the severity of the lesion different patterns were observed: in  $\leq$ CIN1, a wide diversity of HPV genotypes, both LR and HR, was observed in different relative proportions in each sample, without a marked predominance of any of them; while the CIN3+ group exhibits a sharp drop in the diversity of genotypes with a notable predominance of HR-HPVs, especially HPV 16. Notably, mean Gini index was significantly higher in  $<$ CIN1 samples, showing a 2.2-fold increase in genotypic diversity when compared



**Fig 4. HPV genotypic diversity and relative proportion of HPV genotypes in cervical samples by NGS, according to histology.** HPV genotypes were classified according IARC criteria from carcinogenic (Group 1), probably/possibly carcinogenic (Group 2A/B) to probably not carcinogenic (Group 4). Lines correspond to each of the studied samples, the columns to each genotype identified and the different shades of gray in the cells represent the relative proportion, in percentage, of each genotype detected within a sample (the lightest corresponding to the lowest proportion and the darkest to the highest proportion). A: Individual and mean Gini coefficient to measure genotypic diversity are shown for each sample and histological group, respectively (79  $\leq$ CIN1 and 58 CIN3+). Samples grade are ordered by Gini coefficient, from minimal (white) to maximal (black) genotypic diversity. B: Zoom in showing the HPV genotypic relative proportions in 92 cervical samples (68  $\leq$ CIN1 and 24 CIN3+) with multiple infections.

<https://doi.org/10.1371/journal.pone.0278117.g004>

to CIN3+ samples ( $0.20 \pm 0.22$  vs  $0.09 \pm 0.17$  respectively,  $p = 0.001$ ). Genotypic diversity Fig 4B shows in greater detail the situation in the 92 multiple infections samples. In the  $\leq$ CIN1 group, the relative proportions of genotypes found in each sample do not show a marked predominance of one or some of them, the highest relative proportion corresponding to HPV16 (17.7%), followed by HPV52 (14.7%) and HPV31 (10.3%); while in the CIN3+ group, a worthy superiority in the relative proportions of HR-HPV stands out, the predominance being HPV16 (62.5%), followed by HPV31 and HPV58 (8.3% each).

Results of the HPV genotypes identified by RBH and NGS in the 137 cervical samples are shown in Table 1, reporting a good concordance both when genotypes were considered individually and as a whole group (Kappa 0.70). The overall agreement was higher than 90% for all genotypes although the positive agreement showed higher variations suggesting less concordance between the two assays when positivity was present. However, NGS identified eighty-nine new positive results for genotypes included in the RBH probe pool which had not been detected by PCR-RBH. Among them, forty-eight corresponded to ten HR-HPVs (HPVs 16, 18, 31, 33, 35, 39, 45, 51, 52 and 56), eight were probable HR-HPV (HPV 68), nine were possible HR-HPVs (HPVs 26, 53 and 66), and twenty-four corresponded to nine LR-HPVs (HPVs 6, 11, 40, 42, 43, 55, 61, 70 and 81). Differences in agreement were statistically significant for HPVs 16, 31, 35, 39 and 68 (McNemar's  $p < 0.05$ ).

On the other hand, RBH yielded eight positive results for HPVs 58 and 59 (HR-HPVs) and HPVs 73 and 82 (probable HR-HPV), which were not obtained by NGS, but they lacked statistical significance.

**Table 1. Comparison of identified HPV genotypes by RBH and NGS on the 137 cervical samples.** The analysis was restricted to the 37 genotypes which can be detected by both methods.

HPV Genotypes	No. of specimens						% Agreement	% Positive Agreement	McNemar P Value
	RBH positive	NGS positive	RBH-NGS-	RBH+ NGS-	RBH-NGS+	RBH+ NGS+			
6	5	10	127	0	5	5	96	50	0,07
11	2	6	131	0	4	2	97	33	0,13
16	64	71	66	0	7	64	95	90	<b>0,02</b>
18	13	16	121	0	3	13	98	81	0,25
26	1	2	135	0	1	1	99	50	1,00
31	21	33	104	0	12	21	91	64	<b>0,001</b>
33	6	9	128	0	3	6	98	67	0,25
34	1	1	135	1	1	0	99	0	0,48
35	11	17	120	0	6	11	96	65	<b>0,04</b>
39	9	17	120	0	8	9	94	53	<b>0,01</b>
40	3	4	133	0	1	3	99	75	1,00
42	9	11	126	0	2	9	99	82	0,48
43	2	4	133	0	2	2	99	50	0,48
44	7	7	126	4	4	3	94	27	0,72
45	6	7	130	0	1	6	99	86	1,00
51	4	5	132	0	1	4	99	80	1,00
52	18	22	115	0	4	18	97	82	0,88
53	9	13	123	1	5	8	96	57	0,22
54	8	3	128	6	1	2	95	22	0,13
55	3	4	132	1	2	2	98	40	1,00
56	17	20	114	3	6	14	93	61	0,50
58	16	14	121	2	0	14	99	88	0,48
59	9	6	127	4	1	5	96	50	0,37
61	1	5	132	0	4	1	97	20	0,13
66	7	11	126	0	4	7	97	64	0,13
57	0	0	137	0	0	0	100	NA	N/A
68	7	15	122	0	8	7	94	47	<b>0,01</b>
69	0	0	137	0	0	0	100	NA	N/A
70	0	1	136	0	1	0	99	0	1,00
71	0	0	137	0	0	0	100	NA	N/A
72	0	0	137	0	0	0	100	NA	N/A
73	4	2	132	3	1	1	97	20	0,62
81	3	7	130	0	4	3	97	43	0,13
82	4	3	132	2	1	2	98	40	1,00
83	2	2	135	0	0	2	100	100	N/A
84	4	4	129	4	4	0	94	0	0,72
89	1	0	136	1	0	0	99	0	1,00

<https://doi.org/10.1371/journal.pone.0278117.t001>

Additionally, NGS allowed to identify five genotypes (HPVs 30, 62, 74, 87 and 90), beyond the 37 probes included in the RBH assay.

## Discussion

For optimal epidemiological and surveillance studies and vaccine trials, it is essential to as far as possible identify all genotypes present in the sample, individually, and avoid false-negative results. This work takes advantage of the deep analysis of the nucleic acid sequences using

NGS to expand the information about the HPV genotypes' distribution among the histologically characterized cervical samples, and to explore the proportion of each genotype in coinfections.

In agreement with previous publications, HPV16 was by far the most frequently detected type in all histological grades [11, 49, 50]; moreover, our data also add evidence on the strong increase of its relative contribution compared to other HR-HPVs with disease severity from  $\leq$ CIN1 (38.8%) to CIN3+ (69.0%). Our series, albeit small, reproduces the broad spread of HPV16 at global level that was already well established in numerous and large studies [10, 14, 50, 51], indicating a substantial advantage over all other mucosal HPV types in terms of transmissibility and/or persistence.

In our series, the  $\leq$ CIN1 group, exhibited the greatest genotype diversity, both HR and LR-HPV (36 genotypes identified). A recent study from our group [33] assessed the genotypic diversity with PCR-RHB genotyping, using an extension of the series analysed here and also showing higher genotypic diversity in  $\leq$ CIN1 than in CIN3 both for HR and LR HPV's genotypes. Our findings, in the Latin American region, are consistent with what has been already described in the literature [14, 52] in other regions of the world. It confirms the broad spread of these viruses in the sexually active population. On the other hand, in CIN3+, the viral spectrum was much more restricted (23 genotypes identified), detecting almost exclusively HR-HPVs, as previously reported [10–15]. The inverse correlation between HPV diversity and progressive disease is also consistent with the findings of 1,518 cervical biopsies ranging from negative to CIN3 in the ATHENA (Addressing The Need for Advanced HPV diagnostics) trial [53].

The detection of multiple HR-HPV infections has become a key issue in the development of cervical lesions and the epidemiological status of the population, which is why more efficient genotyping strategies are needed. Different studies have reported that multiple HR-HPV acted synergistically in cervical carcinogenesis [54, 55], and cancers with multiple HPV infections could be more resistant to therapy than those with a single infection [56]. On the other hand, Schmitt et al found that the occurrence of multiple HPV infections did not affect the risk of a lesion being high- or low-grade [57], and Wentzensen et al reported no association between disease status and the number of genotypes detected in a woman [31]. Previous works showing an increased risk of CIN with multiple HPV infections had few CIN2+ cases, and restricted to younger women [58], a subgroup known to harbour a larger number of HPV infections [59].

In our study, a significant larger percentage of coinfections was observed in  $\leq$ CIN1 (85%) compared to CIN3+ (41%). Such a high amount of multiple infections has important implications to estimate the fraction of low and high-grade cervical abnormalities that are causally attributable to anyone, or a group, of HPV types. Moreover, as clearly shown by the zoom inside the heat map (Fig 4), when CIN3+ histology is compared to  $\leq$ CIN1 group in multiple infections cases, there is a strong dominance of HR- genotypes (higher relative proportions in the viral mix). Conversely, in the  $\leq$ CIN1 group a marked diversity of genotypes was observed, with none prevailing. Deep amplicon sequencing generated abundant mapped reads and deciphered both the genotype composition within each sample and the proportion of each one in the viral mixture. Our results agree with those of Shen-Gunther et al, who showed that the viral community differed between LSIL and HSIL with a loss of genotypic diversity and domination by carcinogenic HPVs, in particular, HPV-16 in HSIL [25]. HR- HPVs, in particular HPV- 16 and -18, have been shown to be indicators and predictors of CIN3 development [53, 60], being carcinogenic HPV dominance ( $\geq$ 50%) a probable indicator of underlying high-grade disease. These findings are also in line with Depuydt et al, who described in women with multiple HPV types that developed CIN3+, that there was always one of the detected HPV- types present representing a clonal population; the infected basal cell divides with a fixed

number of HPV DNA copies per clonal cell, corresponding to an exponential linear growth over time [28]. These metagenomic characteristics are consistent with the ecological principles of competitive exclusion and carcinogenesis hallmarked by clonal expansion and evolution of transformed cells [28, 61, 62]. Distinctive features of altered diversity and dominance among HPV communities within the histological spectrum could be used as a biomarker of disease severity.

The comparison of HPV genotypes identified by PCR-RBH and NGS showed good concordance both when genotypes were analysed individually and when compared as a whole group (genotypes detectable by all 37 RBH probes). However, NGS was able to identify eighty-nine new positive results for HPV genotypes not detected by PCR-RBH, evidencing NGS' greater sensitivity. The lesser detection capability of PCR-RBH may be partially due to the presence of a viral variant not covered by the probes, or to cross reactivity with another type present in the same sample [63].

In agreement with previous reports, we noticed that NGS was more sensitive than PCR-RBH, particularly in detecting HPV16, HPV31, HPV35, HPV39 and HPV68 [42, 64, 65]. Among them, HPV16 and HPV31, which are phylogenetically related (gender alpha papillomavirus, species 9), have a high epidemiological and pathogenic relevance, given their leading role in the development of cervical precancer and cancer [60]. NGS gives as advantage a massive and parallel sequencing of single molecules from small amounts of DNA, whose benefit increases when the reaction contains multiple HPV types and primers [66]. The high rate of multiple infections may partly explain the discrepant results observed in the present work.

Additionally, NGS allowed to identify five genotypes (HPVs 30, 62, 74, 87 and 90), beyond the 37 probes included in the RBH assay. Although these genotypes have no proven association with cervical carcinogenesis, it is important to detect them to expand their characterization and add information related to HPV epidemiology and virological surveillance.

On the other hand, RBH yielded eight positive results for HPVs 58 and 59 (HR-HPVs) and HPVs 73 and 82 (probable HR-HPV), which were not obtained by NGS. They may represent occasional false-positive results generated by a cross-reaction that has been described for hybridization-based assays [26].

Our study has several strengths, including the implementation of NGS, an advanced technology for HPV genotyping, to expand knowledge on the distribution and relative proportions of each HPV genotype within a sample, in a series of histologically diagnosed cervical samples, in the context of a screening and triage trial in Latin American women. The study has limitations as well; it was a cross-sectional study in a sub-cohort selected by convenience sampling and cannot be claimed to represent the wider screening population.

In conclusion, NGS is a very sensitive and high-throughput tool that may be useful for HPV epidemiological surveillance and for clinical purposes. It may expand the range of genotypes identified in a sample aliquot in a single run. Furthermore, its high detection capability of an expanded range of HPV genotypes and its ability to determine the proportion of each genotype in coinfections allows for an accurate measure rather than a nominal "positive vs negative" measure for a given HPV genotype. In severe lesions, the drop of genotypic diversity and/or the increase of the relative proportion of some of the HR-HPVs in multiple infections could be applied to the follow-up of patients as a biomarker of potential risk for malignant progression.

## Supporting information

**S1 Table. HPV genotyping by NGS: Reads data.**  
(XLSX)

## Acknowledgments

The team is thankful to all women who participated in this study.

The authors are particularly grateful to medical staff of Hospital Nacional “Prof. Posadas” (El Palomar, Provincia de Buenos Aires, Argentina): Juan Mural (Head of Gynecology Service at the launching of the study), Maria Alejandra Borelli (Current Head of Gynecology Service) and Alicia Ines Figueredo and Natalia Gabriela Kaznowiecki (Cervical Pathology specialists) for their for their commitment in recruitment of patients, colposcopies and treatments; Sandra García (Head of Pathological Anatomy Service, at the launching of the study), Adrián Moreno and Julieta Guyot (Cytology Section), Melisa Alejandra Lencioni (Current Head of Pathological Anatomy Service), for their great dedication in coordinating the movement and archiving of samples and morphological studies, and Nerina Cattarin and Mariano Quinteiro for helping in the HPV testing.

We are also indebted to Valeria Padin for her technical cooperation throughout HPV genotyping.

Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

## Author Contributions

**Conceptualization:** Jorge Alejandro Basiletti, Rolando Herrero, Maribel Almonte, María Alejandra Picconi.

**Data curation:** Jorge Alejandro Basiletti, Joan Valls, María Dolores Fellner, Rafael Alonso.

**Formal analysis:** Jorge Alejandro Basiletti, Joan Valls, Rolando Herrero, Maribel Almonte, María Alejandra Picconi.

**Funding acquisition:** Rolando Herrero, Maribel Almonte, María Alejandra Picconi.

**Investigation:** Jorge Alejandro Basiletti, Tomás Poklépovich, María Dolores Fellner, Maryluz Rol, Rita Mariel Correa, María Celeste Colucci, Mercedes Rodríguez de la Peña, Paula Gabriela Falabella, Agustina Saíno, Rolando Herrero, María Alejandra Picconi.

**Methodology:** Jorge Alejandro Basiletti, Joan Valls, Maribel Almonte.

**Project administration:** María Alejandra Picconi.

**Resources:** Mercedes Rodríguez de la Peña, Paula Gabriela Falabella, Agustina Saíno.

**Software:** Joan Valls, Tomás Poklépovich, Rafael Alonso, Josefina Campos.

**Supervision:** Josefina Campos, María Alejandra Picconi.

**Visualization:** María Alejandra Picconi.

**Writing – original draft:** Jorge Alejandro Basiletti, María Alejandra Picconi.

**Writing – review & editing:** Jorge Alejandro Basiletti, Joan Valls, María Dolores Fellner, Maryluz Rol, Rita Mariel Correa, María Celeste Colucci, Rolando Herrero, Maribel Almonte, María Alejandra Picconi.

## References

1. Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H zur, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology*. 2010 May; 401(1):70–9. <https://doi.org/10.1016/j.virol.2010.02.002> PMID: 20206957

2. Karolinska Institute. International Human Papillomavirus Reference Center [Internet]. [cited 2021 Nov 2]. Available from: <https://ki.se/en/labmed/international-hpv-reference-center>.
3. Muñoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, et al. Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. *N Engl J Med*. 2003 Feb 6; 348(6):518–27. <https://doi.org/10.1056/NEJMoa021641> PMID: 12571259
4. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Biological agents. Volume 100 B. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012; 100(Pt B):1–441. PMID: 23189750
5. de Villiers EM. Cross-roads in the classification of papillomaviruses. *Virology*. 2013 Oct; 445(1–2):2–10. <https://doi.org/10.1016/j.virol.2013.04.023> PMID: 23683837
6. Burk RD, Harari A, Chen Z. Human papillomavirus genome variants. *Virology*. 2013 Oct; 445(1–2):232–43. <https://doi.org/10.1016/j.virol.2013.07.018> PMID: 23998342
7. Arbyn M, Tommasino M, Depuydt C, Dillner J. Are 20 human papillomavirus types causing cervical cancer? *J Pathol*. 2014 Dec; 234(4):431–5. <https://doi.org/10.1002/path.4424> PMID: 25124771
8. Burd EM. Human Papillomavirus Laboratory Testing: the Changing Paradigm. *Clin Microbiol Rev*. 2016 Apr; 29(2):291–319. <https://doi.org/10.1128/CMR.00013-15> PMID: 26912568
9. Walboomers JMM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol*. 1999 Sep; 189(1):12–9. [https://doi.org/10.1002/\(SICI\)1096-9896\(199909\)189:1<12::AID-PATH431>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F) PMID: 10451482
10. de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *The Lancet Oncology*. 2010 Nov; 11(11):1048–56. [https://doi.org/10.1016/S1470-2045\(10\)70230-8](https://doi.org/10.1016/S1470-2045(10)70230-8) PMID: 20952254
11. Guan P, Howell-Jones R, Li N, Bruni L, de Sanjosé S, Franceschi S, et al. Human papillomavirus types in 115,789 HPV-positive women: A meta-analysis from cervical infection to cancer. *Int J Cancer*. 2012 Nov 15; 131(10):2349–59. <https://doi.org/10.1002/ijc.27485> PMID: 22323075
12. Herrero R, Castle PE, Schiffman M, Bratti MC, Hildesheim A, Morales J, et al. Epidemiologic Profile of Type-Specific Human Papillomavirus Infection and Cervical Neoplasia in Guanacaste, Costa Rica. *J INFECT DIS*. 2005 Jun; 191(11):1796–807. <https://doi.org/10.1086/428850> PMID: 15871111
13. Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J Cancer*. 2011 Feb 15; 128(4):927–35. <https://doi.org/10.1002/ijc.25396> PMID: 20473886
14. Bruni L, Diaz M, Castellsagué X, Ferrer E, Bosch FX, de Sanjosé S. Cervical Human Papillomavirus Prevalence in 5 Continents: Meta-Analysis of 1 Million Women with Normal Cytological Findings. *J INFECT DIS*. 2010 Dec 15; 202(12):1789–99. <https://doi.org/10.1086/657321> PMID: 21067372
15. Bosch FX, Castellsagué X, de Sanjosé S. HPV and cervical cancer: screening or vaccination? *Br J Cancer*. 2008 Jan; 98(1):15–21. <https://doi.org/10.1038/sj.bjc.6604146> PMID: 18182975
16. Schiffman M, Doorbar J, Wentzensen N, de Sanjosé S, Fakhry C, Monk BJ, et al. Carcinogenic human papillomavirus infection. *Nat Rev Dis Primers*. 2016 Dec 22; 2(1):16086. <https://doi.org/10.1038/nrdp.2016.86> PMID: 27905473
17. IARC-WHO. Globocan [Internet]. 2020 [cited 2021 Sep 10]. Available from: <https://gco.iarc.fr/>.
18. ICO-IARC. HPV Information centre [Internet]. [cited 2021 Oct 21]. Available from: <https://hpvcentre.net/>.
19. Jacobs MV, Snijders PJ, van den Brule AJ, Helmerhorst TJ, Meijer CJ, Walboomers JM. A general primer GP5+/GP6(+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *J Clin Microbiol*. 1997 Mar; 35(3):791–5. <https://doi.org/10.1128/jcm.35.3.791-795.1997> PMID: 9041439
20. Gravitt PE, Peyton CL, Apple RJ, Wheeler CM. Genotyping of 27 Human Papillomavirus Types by Using L1 Consensus PCR Products by a Single-Hybridization, Reverse Line Blot Detection Method. *J Clin Microbiol*. 1998 Oct; 36(10):3020–7. <https://doi.org/10.1128/JCM.36.10.3020-3027.1998> PMID: 9738060
21. Söderlund-Strand A, Carlson J, Dillner J. Modified General Primer PCR System for Sensitive Detection of Multiple Types of Oncogenic Human Papillomavirus. *J Clin Microbiol*. 2009 Mar; 47(3):541–6. <https://doi.org/10.1128/JCM.02007-08> PMID: 19144817
22. Kleter B, van Doorn LJ, Schrauwen L, Molijn A, Sastrowijoto S, ter Schegget J, et al. Development and Clinical Evaluation of a Highly Sensitive PCR-Reverse Hybridization Line Probe Assay for Detection

- and Identification of Anogenital Human Papillomavirus. *J Clin Microbiol.* 1999 Aug; 37(8):2508–17. <https://doi.org/10.1128/JCM.37.8.2508-2517.1999> PMID: 10405393
23. Estrade C, Menoud PA, Nardelli-Haeffliger D, Sahli R. Validation of a Low-Cost Human Papillomavirus Genotyping Assay Based on PGMY PCR and Reverse Blotting Hybridization with Reusable Membranes. *J Clin Microbiol.* 2011 Oct; 49(10):3474–81. <https://doi.org/10.1128/JCM.05039-11> PMID: 21832011
  24. Schmitt M, Dondog B, Waterboer T, Pawlita M. Homogeneous Amplification of Genital Human Alpha Papillomaviruses by PCR Using Novel Broad-Spectrum GP5+ and GP6+ Primers. *J Clin Microbiol.* 2008 Mar; 46(3):1050–9. <https://doi.org/10.1128/JCM.02227-07> PMID: 18199790
  25. Shen-Gunther J, Wang Y, Lai Z, Poage GM, Perez L, Huang THM. Deep sequencing of HPV E6/E7 genes reveals loss of genotypic diversity and gain of clonal dominance in high-grade intraepithelial lesions of the cervix. *BMC Genomics.* 2017 Dec; 18(1):231. <https://doi.org/10.1186/s12864-017-3612-y> PMID: 28288568
  26. Artaza-Irigaray C, Flores-Miramontes MG, Olszewski D, Vallejo-Ruiz V, Limón-Toledo LP, Sánchez-Roque C, et al. Cross-hybridization between HPV genotypes in the Linear Array Genotyping Test confirmed by Next-Generation Sequencing. *Diagn Pathol.* 2019 Dec; 14(1):31. <https://doi.org/10.1186/s13000-019-0808-2> PMID: 31010421
  27. Flores-Miramontes MG, Olszewski D, Artaza-Irigaray C, Willemsen A, Bravo IG, Vallejo-Ruiz V, et al. Detection of Alpha, Beta, Gamma, and Unclassified Human Papillomaviruses in Cervical Cancer Samples From Mexican Women. *Front Cell Infect Microbiol.* 2020 Jun 9; 10:234. <https://doi.org/10.3389/fcimb.2020.00234> PMID: 32582561
  28. Depuydt CE, Thys S, Beert J, Jonckheere J, Salembier G, Bogers JJ. Linear viral load increase of a single HPV-type in women with multiple HPV infections predicts progression to cervical cancer: HPV-type specific viral load slope predicts cervical cancer. *Int J Cancer.* 2016 Nov 1; 139(9):2021–32.
  29. Dickson EL, Vogel RI, Geller MA, Downs LS. Cervical cytology and multiple type HPV infection: A study of 8182 women ages 31–65. *Gynecologic Oncology.* 2014 Jun; 133(3):405–8. <https://doi.org/10.1016/j.ygyno.2014.03.552> PMID: 24657488
  30. Senapati R, Nayak B, Kar SK, Dwibedi B. HPV genotypes co-infections associated with cervical carcinoma: Special focus on phylogenetically related and non-vaccine targeted genotypes. *Consolaro MEL*, editor. *PLoS ONE.* 2017 Nov 21; 12(11):e0187844. <https://doi.org/10.1371/journal.pone.0187844> PMID: 29161285
  31. Wentzensen N, Schiffman M, Dunn T, Zuna RE, Gold MA, Allen RA, et al. Multiple human papillomavirus genotype infections in cervical cancer progression in the study to understand cervical cancer early endpoints and determinants. *Int J Cancer.* 2009 Nov; 125(9):2151–8. <https://doi.org/10.1002/ijc.24528> PMID: 19585494
  32. Almonte M, Murillo R, Sánchez GI, González P, Ferrera A, Picconi MA, et al. Multicentric study of cervical cancer screening with human papillomavirus testing and assessment of triage methods in Latin America: the ESTAMPA screening study protocol. *BMJ Open.* 2020 May; 10(5):e035796. <https://doi.org/10.1136/bmjopen-2019-035796> PMID: 32448795
  33. Correa RM, Baena A, Valls J, Colucci MC, Mendoza L, Rol M, et al. Distribution of human papillomavirus genotypes by severity of cervical lesions in HPV screened positive women from the ESTAMPA study in Latin America. *Tornesello ML*, editor. *PLoS ONE.* 2022 Jul 29; 17(7):e0272205. <https://doi.org/10.1371/journal.pone.0272205> PMID: 35905130
  34. van den Brule AJC, Pol R, Franssen-Daalmeijer N, Schouls LM, Meijer CJLM, Snijders PJF. GP5+/6+ PCR followed by Reverse Line Blot Analysis Enables Rapid and High-Throughput Identification of Human Papillomavirus Genotypes. *J Clin Microbiol.* 2002 Mar; 40(3):779–87.
  35. Estrade C, Sahli R. Updating the PGMY Primers and Probes for Improved Detection of HPV68a: Validation of Version 2 of the PGMY-CHUV Assay. *Caliendo AM*, editor. *J Clin Microbiol.* 2014 Nov; 52(11):4033–5. <https://doi.org/10.1128/JCM.01950-14> PMID: 25143570
  36. Mendoza LP, Arbiza J, Paez M, Kasamatsu E, Castro A, Giménez G, et al. Distribution of human papillomavirus genotypes in Paraguayan women according to the severity of the cervical lesion. *J Med Virol.* 2011 Aug; 83(8):1351–7. <https://doi.org/10.1002/jmv.22112> PMID: 21618554
  37. Jacot-Guillarmod M, Pasquier J, Greub G, Bongiovanni M, Achdari C, Sahli R. Impact of HPV vaccination with Gardasil® in Switzerland. *BMC Infect Dis.* 2017 Dec; 17(1):790. <https://doi.org/10.1186/s12879-017-2867-x> PMID: 29273004
  38. González JV, Deluca GD, Correa RM, Liotta DJ, Basiletti JA, Fellner MD, et al. Strong reduction in prevalence of HPV16/18 and closely related HPV types in sexually active adolescent women following the introduction of HPV vaccination in Argentina. *Papillomavirus Research.* 2020 Dec; 10:100208. <https://doi.org/10.1016/j.pvr.2020.100208> PMID: 33161174

39. Kukimoto I, Matsumoto K, Takahashi F, Iwata T, Tanaka K, Yamaguchi-Naka M, et al. Human Papillomavirus (HPV) Genotyping Assay Suitable for Monitoring the Impact of the 9-Valent HPV Vaccine. *Tohoku J Exp Med*. 2020; 251(4):287–94. <https://doi.org/10.1620/tjem.251.287> PMID: 32759554
40. Bel Haj Rhouma R, Ardhaoui M, El Fehri E, Marzougui A, Laassili T, Guizani I, et al. Distribution of human papillomavirus in precancerous and cancerous cervical neoplasia in Tunisian women. *Infect Agents Cancer*. 2021 Dec; 16(1):52. <https://doi.org/10.1186/s13027-021-00392-1> PMID: 34271960
41. Eklund C, Mühr LSA, Lagheden C, Forslund O, Robertsson KD, Dillner J. The 2019 HPV Labnet international proficiency study: Need of global Human Papillomavirus Proficiency Testing. *Journal of Clinical Virology*. 2021 Aug; 141:104902. <https://doi.org/10.1016/j.jcv.2021.104902> PMID: 34186414
42. Ardhaoui M, Ennaifer E, De Matos Salim AC, Gomez FM, Laassili T, Boubaker MS, et al. Nested PCR followed by NGS: Validation and application for HPV genotyping of Tunisian cervical samples. Aguayo F, editor. *PLoS ONE*. 2021 Aug 11; 16(8):e0255914. <https://doi.org/10.1371/journal.pone.0255914> PMID: 34379683
43. Andrews S. FastQC. A quality control tool for high throughput sequence data. [Internet]. Babraham Bioinformatics; 2010. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
44. Krueger F. Trim Galore [Internet]. Babraham Bioinformatics; 2019. Available from: [https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).
45. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*. 2014; 15(3):R46. <https://doi.org/10.1186/gb-2014-15-3-r46> PMID: 24580807
46. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009; 10(3):R25. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: 19261174
47. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011 May 15; 29(7):644–52. <https://doi.org/10.1038/nbt.1883> PMID: 21572440
48. Hughes J. WeeSAM [Internet]. CVR Bioinformatics; Available from: <https://bioinformatics.cvr.ac.uk/weesam-version-1-5/>.
49. De Vuyst H, Clifford GM, Nascimento MC, Madeleine MM, Franceschi S. Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: A meta-analysis. *Int J Cancer*. 2009 Apr 1; 124(7):1626–36. <https://doi.org/10.1002/ijc.24116> PMID: 19115209
50. Ciapponi A, Bardach A, Glujovsky D, Gibbons L, Picconi MA. Type-Specific HPV Prevalence in Cervical Cancer and High-Grade Lesions in Latin America and the Caribbean: Systematic Review and Meta-Analysis. Borrow R, editor. *PLoS ONE*. 2011 Oct 4; 6(10):e25493. <https://doi.org/10.1371/journal.pone.0025493> PMID: 21991313
51. Bzhalava D, Guan P, Franceschi S, Dillner J, Clifford G. A systematic review of the prevalence of mucosal and cutaneous human papillomavirus types. *Virology*. 2013 Oct; 445(1–2):224–31. <https://doi.org/10.1016/j.virol.2013.07.015> PMID: 23928291
52. Clifford GM. Human Papillomavirus Genotype Distribution in Low-Grade Cervical Lesions: Comparison by Geographic Region and with Cervical Cancer. *Cancer Epidemiology Biomarkers & Prevention*. 2005 May 1; 14(5):1157–64. <https://doi.org/10.1158/1055-9965.EPI-04-0812> PMID: 15894666
53. Stoler MH, Wright TC, Sharma A, Apple R, Gutekunst K, Wright TL. High-Risk Human Papillomavirus Testing in Women With ASC-US Cytology. *American Journal of Clinical Pathology*. 2011 Mar 1; 135(3):468–75.
54. Trottier H. Human Papillomavirus Infections with Multiple Types and Risk of Cervical Neoplasia. *Cancer Epidemiology Biomarkers & Prevention*. 2006 Jul 1; 15(7):1274–80. <https://doi.org/10.1158/1055-9965.EPI-06-0129> PMID: 16835323
55. Wang L, Wang P, Ren Y, Du J, Jiang J, Jia X, et al. Prevalence of High-Risk Human Papillomavirus (HR-HPV) Genotypes and Multiple Infections in Cervical Abnormalities from Northern Xinjiang, China. Tornesello ML, editor. *PLoS ONE*. 2016 Aug 5; 11(8):e0160698. <https://doi.org/10.1371/journal.pone.0160698> PMID: 27494179
56. Munagala R, Donà MG, Rai SN, Jenson AB, Bala N, Ghim SJ, et al. Significance of multiple HPV infection in cervical cancer patients and its impact on treatment response. *Int J Oncol*. 2009 Jan; 34(1):263–71. PMID: 19082497
57. Schmitt M, Depuydt C, Benoy I, Bogers J, Antoine J, Arbyn M, et al. Multiple Human Papillomavirus Infections with High Viral Loads Are Associated with Cervical Lesions but Do Not Differentiate Grades of Cervical Abnormalities. *J Clin Microbiol*. 2013 May; 51(5):1458–64. <https://doi.org/10.1128/JCM.00087-13> PMID: 23447632

58. Chaturvedi AK, Katki HA, Hildesheim A, Rodríguez AC, Quint W, Schiffman M, et al. Human Papillomavirus Infection with Multiple Types: Pattern of Coinfection and Risk of Cervical Disease. *The Journal of Infectious Diseases*. 2011 Apr 1; 203(7):910–20. <https://doi.org/10.1093/infdis/jiq139> PMID: 21402543
59. Cuschieri KS. Multiple high risk HPV infections are common in cervical neoplasia and young women in a cervical screening population. *Journal of Clinical Pathology*. 2004 Jan 1; 57(1):68–72. <https://doi.org/10.1136/jcp.57.1.68> PMID: 14693839
60. Adcock R, Cuzick J, Hunt WC, McDonald RM, Wheeler CM. Role of HPV Genotype, Multiple Infections, and Viral Load on the Risk of High-Grade Cervical Neoplasia. *Cancer Epidemiol Biomarkers Prev*. 2019 Nov; 28(11):1816–24. <https://doi.org/10.1158/1055-9965.EPI-19-0239> PMID: 31488417
61. Bravo IG, Félez-Sánchez M. Papillomaviruses. *Evolution, Medicine, and Public Health*. 2015; 2015(1):32–51.
62. Quint W, Jenkins D, Molijn A, Struijk L, van de Sandt M, Doorbar J, et al. One virus, one lesion-individual components of CIN lesions contain a specific HPV type. *J Pathol*. 2012 May; 227(1):62–71. <https://doi.org/10.1002/path.3970> PMID: 22127961
63. Meisal R, Rounge TB, Christiansen IK, Eieland AK, Worren MM, Molden TF, et al. HPV Genotyping of Modified General Primer-Amplicons Is More Analytically Sensitive and Specific by Sequencing than by Hybridization. Aguayo FR, editor. *PLoS ONE*. 2017 Jan 3; 12(1):e0169074. <https://doi.org/10.1371/journal.pone.0169074> PMID: 28045981
64. Nilyanimit P, Chansaenroj J, Poomipak W, Praianantathavorn K, Payungporn S, Poovorawan Y. Comparison of Four Human Papillomavirus Genotyping Methods: Next-generation Sequencing, INNO-LiPA, Electrochemical DNA Chip, and Nested-PCR. *Ann Lab Med*. 2018 Mar 28; 38(2):139–46. <https://doi.org/10.3343/alm.2018.38.2.139> PMID: 29214758
65. Arroyo Mühr LS, Lagheden C, Lei J, Eklund C, Nordqvist Kleppe S, Sparén P, et al. Deep sequencing detects human papillomavirus (HPV) in cervical cancers negative for HPV by PCR. *Br J Cancer*. 2020 Dec 8; 123(12):1790–5. <https://doi.org/10.1038/s41416-020-01111-0> PMID: 33020595
66. Mori S, Nakao S, Kukimoto I, Kusumoto-Matsuo R, Kondo K, Kanda T. Biased amplification of human papillomavirus DNA in specimens containing multiple human papillomavirus types by PCR with consensus primers. *Cancer Science*. 2011 Jun; 102(6):1223–7. <https://doi.org/10.1111/j.1349-7006.2011.01922.x> PMID: 21388488