

# G-QUADRUPLEX DNA STRUCTURES IN HUMAN PAPILLOMAVIRUS TYPES 16, 18 AND 58: PROMISING MOLECULAR TARGETS



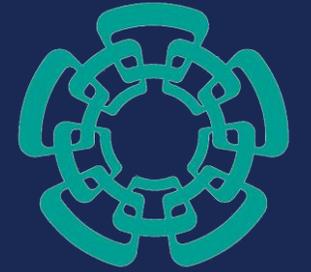
Kantún N.<sup>1</sup>, Ontiveros N.J.\*<sup>1,2</sup>, Montero J.<sup>3</sup>

<sup>1</sup> Centro de investigación Regional Dr. Hideyo Noguchi, Yucatán, México

<sup>2</sup> Universidad Autónoma de Yucatán, Yucatán, México

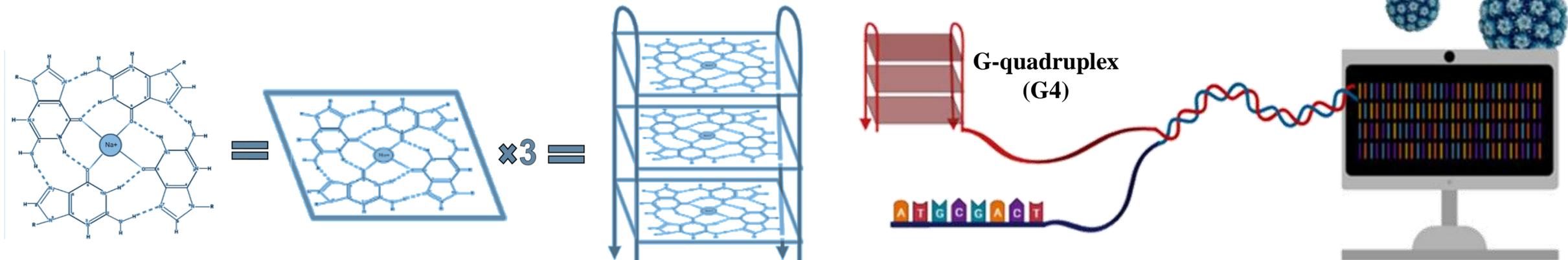
<sup>3</sup> Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Yucatán, México

\* [a15115717@alumnos.uady.com](mailto:a15115717@alumnos.uady.com) ; [josynan9@gmail.com](mailto:josynan9@gmail.com)



## INTRODUCTION

DNA is a structurally variable molecule, subject to folding into other conformations that disobey the Watson-Crick canon ( $\beta$ -DNA model); they are described as non-canonical forms and one of them are the G-quadruplexes (G4s). These are structures that form in DNA or RNA in guanine (G)-rich regions. G4s or Putative Quadruplex Sequences (PQSs) have been identified in eukaryotic and prokaryotic organisms and viruses with multiple biological roles. The structural diversity, folding topologies, and *in vitro* stability of G4s put them in the spotlight to be considered as a new pharmacological target, so it would be interesting to characterize them in DNA viruses, such as human papillomavirus (HPV), the causative agent of common reproductive tract infections and a variety of cancers.



# RESEARCH OBJECTIVE

Identify *in silico* DNA G-quadruplexes in the high-risk human papillomaviruses 16, 18 and 58 genomes to obtain possible molecular targets, as well as in some of their variants.

## METHODS

### STAGE 1

Selection and validation of bioinformatics predictors



Review of information sources.



Search and compilation of bioinformatics predictors for the elucidation of G4 in DNA genomes.



Validation and selection of at least 2 bioinformatics predictors

### STAGE 2

Identification of G4 structures in the genomes of HPV 16, 18 and 58, and the study of their variants.



*In silico* identification of G4 in HPV 16, 18 and 58 reference genomes, their diversity by gene region, G4 type and DNA strand.



Characterization of G4 in variants of the three HPV genotypes studied.



Determination of conserved and atypical G4 motifs among HPV genomes.

### STAGE 3

Molecular modelling of selected G4.



Selection and validation of bioinformatics programs

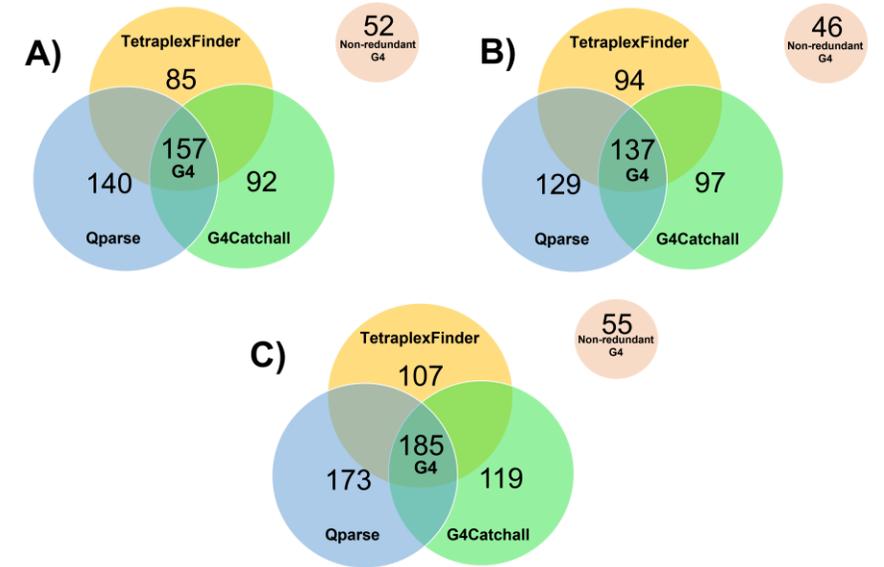


Models of folding-prone G-quadruplexes according to typical G4 topologies

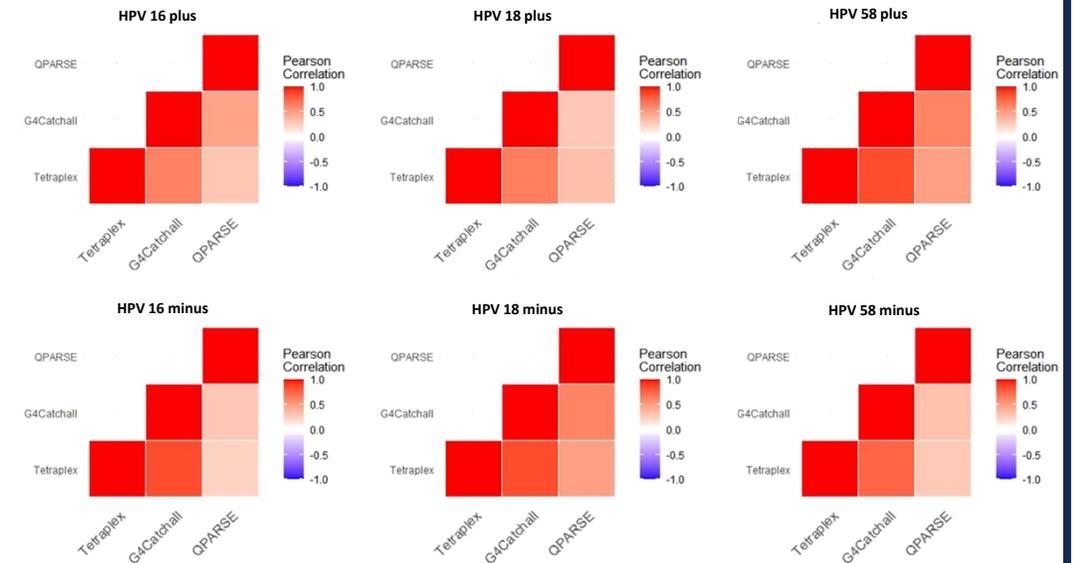
# RESULTS

**Table 1.** Comparative analysis of available software and its features for G4.

Software	DNA	G4	Total	# of	Loop	Island	Atypical	Stand	G4	Score
	position	position	motif	consecutive	size	size	motif	(+/-)	overlapping	
			size	islands	(1-12)	(2-5)	(max 2			
				(≥2)			B/M)			
G4CatchAll	✓	✓	✓	✓	✓	✓	✓	✓	✓	G4Hscore
AllQuads	✓	✓			✓			✓	✓	
ImGQfinder	✓	✓		✓	✓	✓	✓		✓	
Pqsfinder	✓	✓	✓	✓	✓	✓	✓	✓	✓	pqsfinder Score
G4P calculator	✓								✓	Percentage
G4PromFinder	✓			✓	✓	✓		✓		
G4 RNA screener		✓							✓	G4NN G4Hscore cGcC score
Quadron	✓	✓			✓			✓	✓	Quadron score
OPARSE	✓	✓	✓	✓	✓	✓	✓	✓	✓	OPARSE score
G4PREDICTORV.2	✓	✓	✓	✓	✓	✓		✓		cGcC Score
G4iM Grinder	✓	✓			✓			✓	✓	Total Score Frequency, Pqsfinder, G4Hunter y cGcC GScore
QGRS Mappers	✓	✓	✓	✓	✓	✓			✓	
nonB DNA Motif Search Tool (nBMST)	✓	✓	✓	✓						
IpiRid	✓							✓		
QuadBase	✓	✓	✓	✓	✓	✓	✓	✓	✓	
TetraPlexFinder										
G-quadruplex detector		✓	✓	✓		✓				



**Figure 1.** Number of G4 predicted in HPV 16 (A), HPV 18 (B) and HPV 58 (C) genomes.

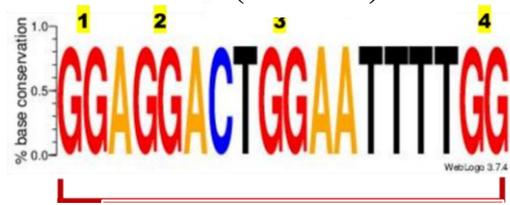


**Figure 2.** Pearson correlation between the 3 softwares, on each strand by HPV

**Figure 5 .** Distribution of G4 in HPV 16, 18 and 58 genomes by size and gene position.

**A) Type of G4**

(10-16%)

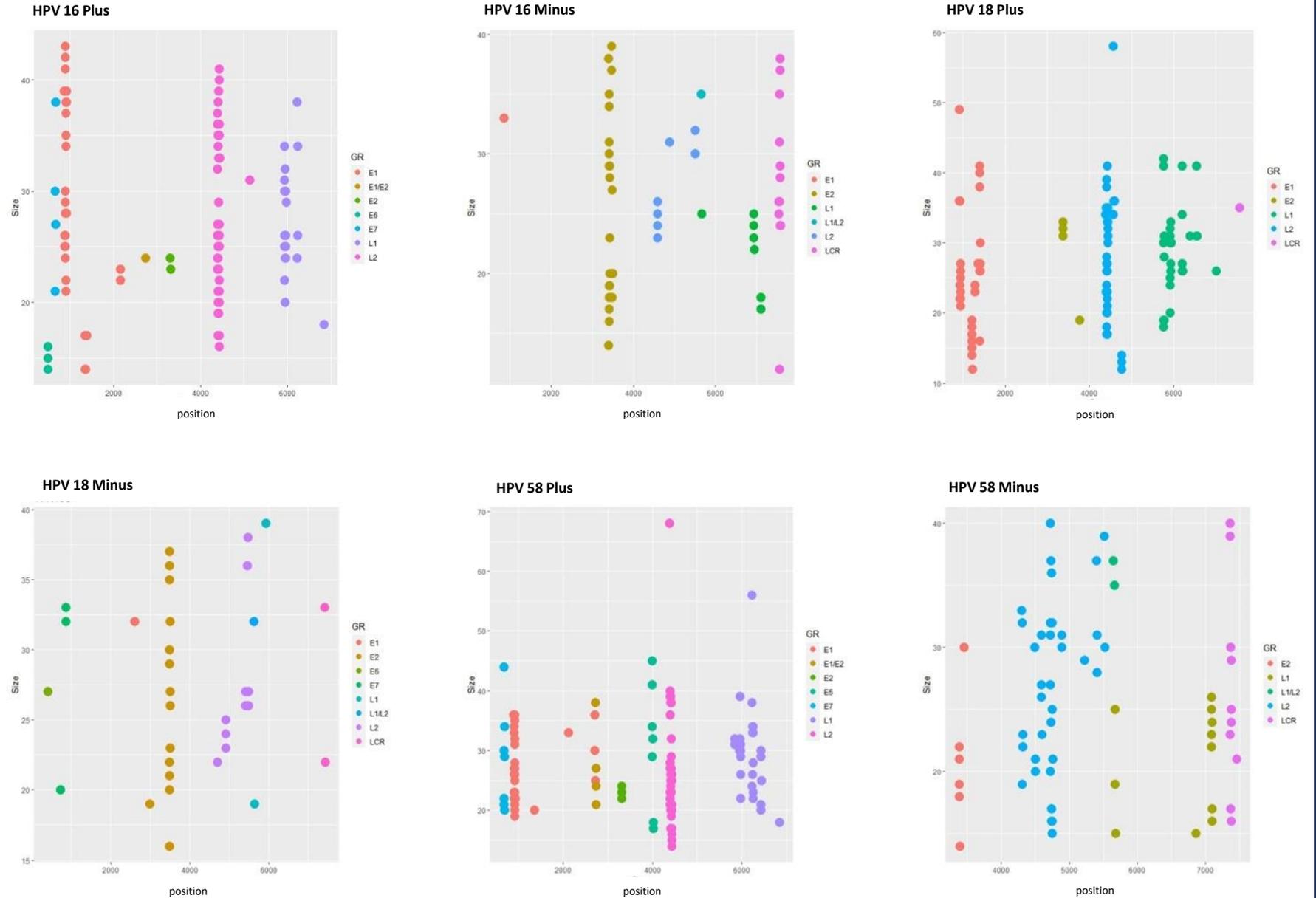


**Figure 3.** Typical G4

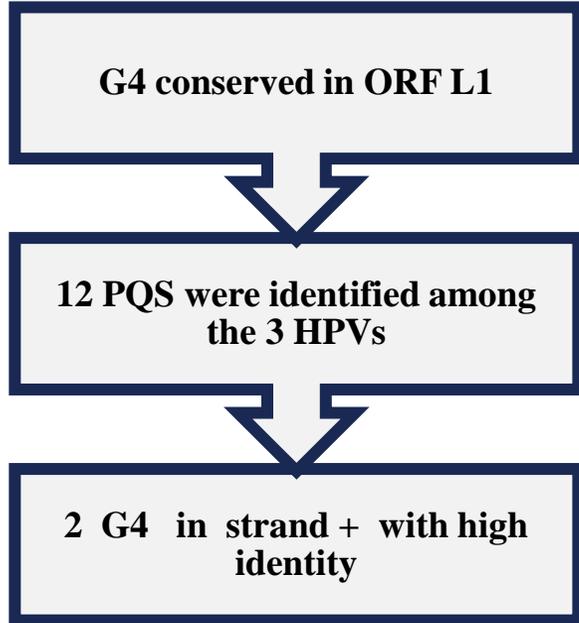
(83-89%)



**Figure 4.** Atypical G4



**A) G4 in ORF L1**

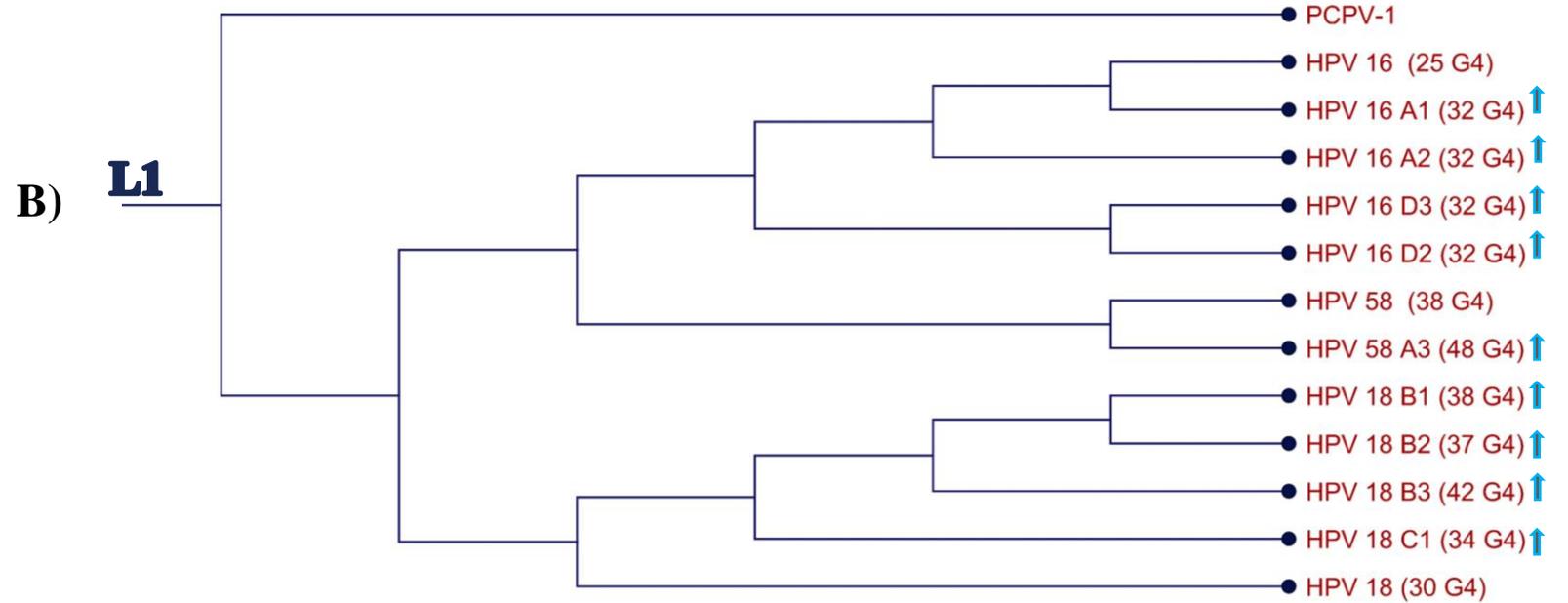


The density of G4 in the L1 gene was :

- HPV 16: 3 PQS/Kb
- HPV 18: 4 PQS/Kb
- HPV 58: 5 PQS/Kb

**L1C5: 83.3% IDENTITY**

**L1C10: 80.76% IDENTITY**

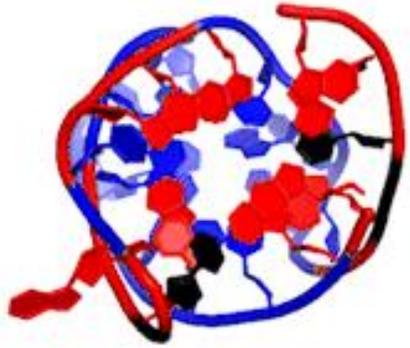


**C) Table 2.** Position of polymorphisms between the reference genome and its variants per PQS L1C5

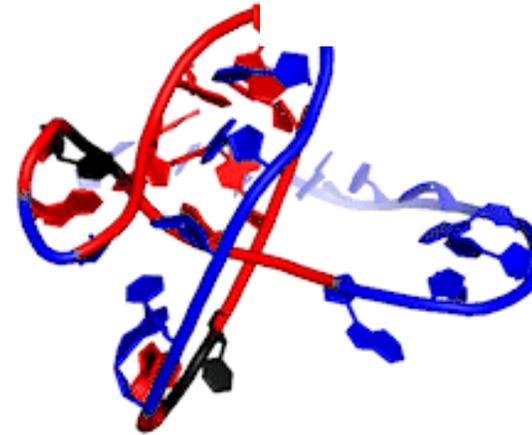
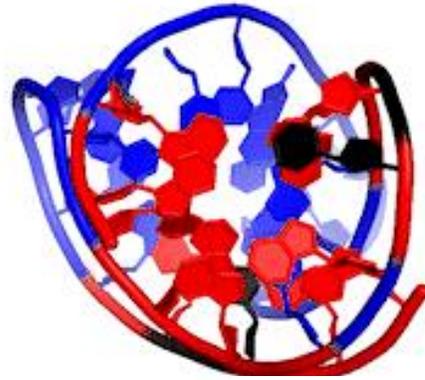
Genotype	Position	SNPs	Variant with SNPs
HPV 18	5940-5969	G-5929	C1

L1C5 was 100% conserved in HPV 16 and 58 among its variants; it only presented one mutation in HPV18, in the C1 variant.

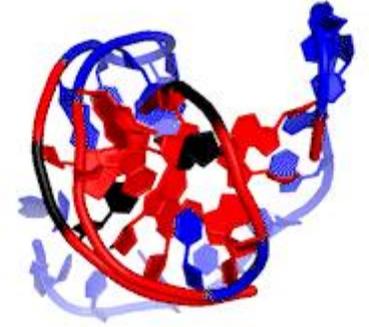
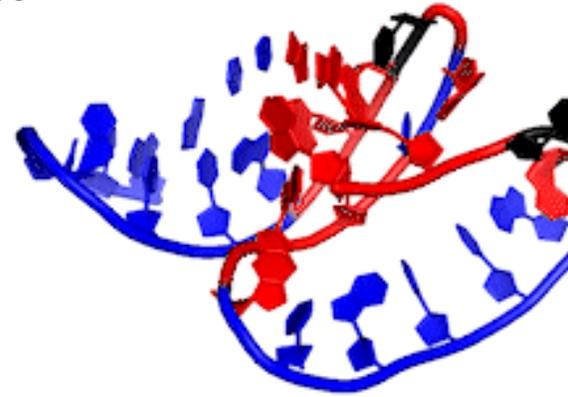
L1C5



Antiparallel G4



Parallel G4



Mixed G4

## CONCLUSION

More than 100 PQS were identified in the genomes of HPV 16, 18 and 58. In the L1 region encoding the main structural protein, two conserved G4 motifs (L1C5 and L1C10) were identified for all three HPVs. L1C5 is an atypical G4 that is conserved in cervical cancer-related variants, so it could be a target for further study. The implementation of three predictors helped in the elucidation of G4 diversity, with TetraplexFinder and G4Catchall giving very similar results.

*This work form part of UADY project (CIRB-2021-0005). Funding: none.*