Comparing Roche Linear Array® to Next Generation Sequencing for HPV Genotype Identification

Luke Konur[1], Ashley Winters[1], Brock Williams[1], Meng Luo[1], Melody Baddoo[2], Jennifer Cameron PhD[1]

[1]Department of Microbiology, Immunology, and Parasitology, LSUHSC

[2]Tulane Cancer Center

Introduction

HPV in Cervical Cancer

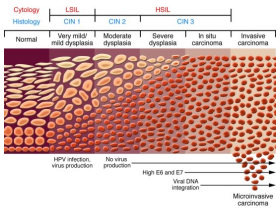

Figure 1: Progression of human cervical cancer after HPV infection[1]

- Human papillomavirus (HPV) is known to be a causative agent in most cervical neoplasia, thus identifying HPV in a patient's cervical swab is an important step in detecting risk for cervical cancer and preventing further development.[2]
- Over 200 genotypes have been found to date, of which 30-40 have a role in Cervical Cancer, with high-risk types being most prevalent in cervical dysplasia.[3]

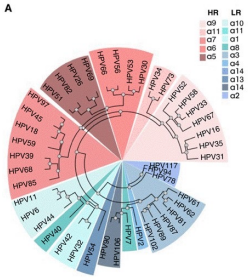

Figure 2: Phylogenetic tree of Alpha HPV family[4]. HR: High-oncogenic risk; LR: Low-oncogenic risk.

Linear Array

- The Linear Array Genotyping Test (LA) from Roche is a colorimetric reverse line blot hybridization array that detects 37 genotypes of HPV and has recently been discontinued.[5]
- The LA has been the gold standard assay used in research for HPV genotyping over the past two decades.
- Cross-hybridization between genotypes has been reported and could result in misidentification of HPV genotypes in a patient sample.[5]

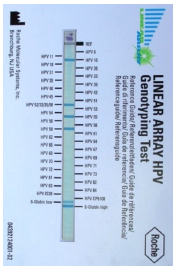

Figure 3: Roche LA genotyping kit reference guide and LA strip.[6]

Next Generation Sequencing

- Newer high-throughput assays like Next Generation Sequencing (NGS) have broadened the variety of HPV types that can be detected.
- NGS outputs sequence data that can be aligned with a continually updated library of reference genomes, allowing no lag time between new genotype discovery and its addition to the sequence library.
- The use of the MY09/11 degenerate primer set allow for nonspecific amplification of HPV late gene 1(L1) regions and sequencing of amplicons through NGS.[2]
- These primers target a 450 base pair region that is highly conserved in the L1 gene of a wide range of genital HPV types.[2]

Objective

The Objective of this work is to compare the performance of Next Generation Sequencing with MY09/11 Primers against the Roche Linear Array.

Genotype Results

Sample Number	Positive by LA only	Positive by NGS only	Positive by both
1	89	33	16
2	35, 52, 84	6, 53	71
3	55, 62, 73	6, 16	53
4	42, 61, 68		6, 31, 53
5	16, 51, 83, 84	81, 85	18
6	45, 58, 70	53	31
7			16, 54
8	62	54, 61, 83	
9	62, 89		61, 83
10	54, 62, 84	6, 31, 81, 83	
11	58	16	
12			16
13	42	16	62, 81
14		44, 85	16, 18, 42
15			18, 53, 70
16	61		
17			81
18			83

Table 1: Genotypes found by LA, NGS, or both for each individual sample. Red genotypes are undetectable by LA

HPV Genotype	Samples Positive by LA	Samples Positive by NGS	Samples Positive by Both LA and NGS
6	1	4	1
16	5	7	4
18	3	3	3
31	2	3	2
33	0	1	0
35	1	0	0
42	3	1	1
44	0	1	0
45	1	0	0
51	1	0	0
52	1	0	0
53	3	5	3
54	2	2	1
55	1	0	0
58	2	0	0
61	3	2	1
62	5	1	1
68	1	0	0
70	2	1	1
71	1	1	1
73	1	0	0
81	2	4	2
83	3	4	2
84	3	0	0
85	0	3	0
89	2	0	0

Table 2: Results of LA, NGS, and both for each genotype found in this work. Red genotypes are undetectable by LA

- A total of 23 HPV types were found by LA and 16 HPV types were found through NGS.
- There was an average concordance between LA and NGS of 92.01% (including positive and negative results).
- NGS found 2 HPV types not detected by LA, HPV 44, and 85, highlighted in red in Tables 1 and 2.
- LA found HPV 35, 45, 51, 52, 55, 58, 68, 73, 84, and 89 infections that were not detected by NGS.

Methods

- DNA extracts from cervical swabs, collected previously from patients in an HPV observational study.
- The samples were tested by the Linear Array Genotyping Test (Roche Diagnostics), according to manufacturer specifications.
- Next, HPV libraries were prepared using MY09/11 primers, the samples were sequenced using the Illumina MiSeq platform and sequence reads aligned to HPV reference genomes using custom Perl scripts.
- Best concordance between LA and NGS was found when a read count threshold of ≥ 20 reads was applied to NGS output.
- Data from NGS was then compared to Linear Array results using Excel and Graphpad Prism.

This work was supported by the Center of Biomedical Research Excellence (CoBRE) in Translational Viral Oncology (CTVO) funded by the NIH NIGMS, award # P20GM121288.

Concordance vs. Mismatch Count

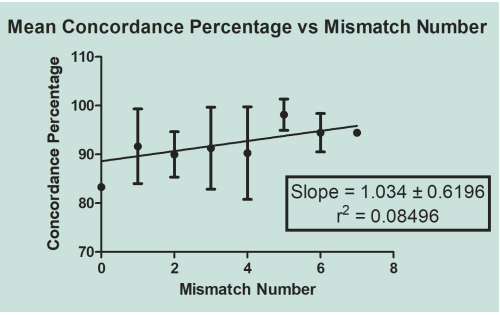

Figure 1: Comparison of concordance for each HPV genotype against number of mismatches between HPV reference consensus sequence and MY09/11 primer sequence.

- First the sequence of the primer binding region for each HPV genotype was compared to the primer set, and the number of nucleotide mismatches summed(see example).
- Then, the concordance of each HPV genotype was plotted against this mismatch number, with concordance of types with same mismatch number averaged.
- Simple linear regression yielded no significant trend in this comparison.

Example Alignment for HPV 31

MY11

MY11 - 5'-GCM CAG GGW CAT AAY AAT GG-3'

HPV 31 - 5'-GCT CAG GGA CAC AAT AAT GG-3'

MY09

MY09 - 5'-GAT CAG TWT CCY YTK GGA CG-3'

HPV 31 - 5'-GAT CAG TTT CCA CTG GGT CG-3'

This sample alignment shows the primer sequence first with degenerate primers, then the reference genome consensus sequence. Mismatches are highlighted in yellow. Type 31 had a total of 4 mismatches, and it had a concordance of 94.44%. Degenerate primers: M = A/C ; W = A/T ; Y = C/T ; K = G/T .

Conclusion

- NGS with MY09/11 primers were able to identify two genotypes not detected by linear array, HPV 44 and 85.
- HPV 85 is classified as High-risk, and HPV 44 is classified as low risk, so these are important infections that are not detected by LA.
- Cross-hybridization has been reported possibly explaining why LA detected genotypes not found by NGS in individual samples.
- The lack of correlation between mismatch number and concordance could be due to low sample number in this work.
- Perhaps, in a larger sample size this correlation could be seen as a better match between primer sequence and genotype sequence should lead to better amplification, and therefore better concordance.
- Cross-hybridization could also be confounding this correlation.
- Overall, the NGS platform circumvents the cross-hybridization of LA as well as allows for the detection of a broader range of HPV types.

References

1. J Clin Invest. 2006;116(5):1167-1173.
2. Depuydt CE, Boulet GA, Horvath CA, et al. Comparison of MY09/11 consensus PCR and type-specific PCRs in the detection of oncogenic HPV types. Journal of Cellular and Molecular Medicine. 2007 Jul-Aug;11(4):881-891.
3. Bansal D, Elmi AA, Skariah S, Hoddad P, Abu-Raddad LJ, Al Hamadi AH, Mohamed-Nady N, Affifi NM, Chedira R, Hassen E, Al-Thani AA, Al-Ansari AA, Sultan AA. Molecular epidemiology and genotype distribution of Human Papillomavirus (HPV) among Arab women in the State of Qatar. J Transl Med. 2014 Nov 26;12:300.
4. Auslander N, Wolf YI, Shabalina SA and Koonin EV. A unique insert in the genomes of high-risk human papillomaviruses with a predicted dual role in conferring oncogenic risk [version 1; peer review: 2 approved]. F1000Research 2019, 8:1000
5. Artaza-Irigaray, C., Flores-Miramontes, M.G., Olszewski, D. et al. Cross-hybridization between HPV genotypes in the Linear Array Genotyping Test confirmed by Next-Generation Sequencing. Diagn Pathol 14, 31 (2019).
6. Vaiphei K, Kochhar R, Bhardawaj S, Dutta U, Singh K. High prevalence of human papillomavirus in esophageal squamous cell carcinoma: a study in paired samples. Dis Esophagus. 2013 Apr;26(3):282-7. doi: 10.1111/j.1442-2050.2012.01365.x. Epub 2012 Jun 7. PMID: 22676445.