Identification of novel genomic structures and regulation patterns at HPV integration events in cervical cancer

Vanessa L. Porter, Kieran O'Neill, Signe A. MacLennan, Richard D. Corbett, Glenn Chang, Zeid Hamadeh, Luka Culibrk, Marissa Iden, Rachel Mutchler, Shirng-Wern Tsaih, Jeremy Fan, Vahid Akbari, Simon K. Chan, Richard Moore, Karen L. Mungall, Andrew J. Mungall, Inanc Birol, Steven J. M. Jones, Janet S. Rader, Marco A. Marra

Provincial Health Services Authority

Introduction

HPV and cervical cancer

Human papillomavirus (HPV) infection is necessary in the development of cervical cancer. Often, part of the transformation into cancer is the integration of the HPV genome into the host genome[1]. The integrated form of HPV changes the regulation of HPV genes, and transforms the genome and epigenome at the regions harbouring HPV[2].

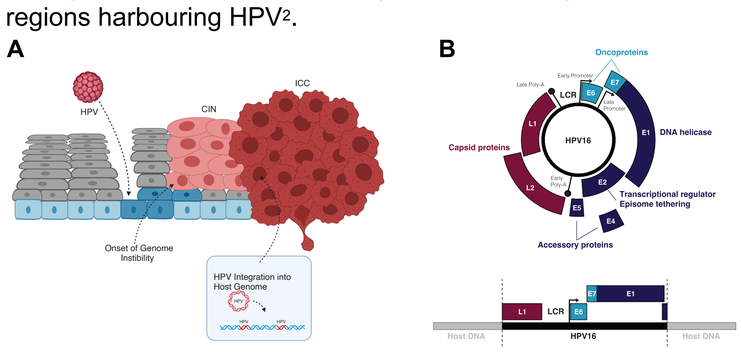


Figure 1: (A) The evolution of HPV infection, to cervical intraepithelial neoplasia (CIN), to invasive cervical cancer (ICC). **(B)** Examples of the episomal and integrated forms of the HPV genome including the genes and regulatory regions. Figure from Porter and Marra[3], 2022.

Objective and Aims

Objective: Investigate how integration of HPV affects the structure and regulation of cervical cancer genomes using Oxford Nanopore Technology (ONT) long read sequencing.

Aims:

1) Determine the structural categories of HPV integration using ONT long-read whole genome data (59 samples total).

2) Investigate and describe how each category affects the human genome.

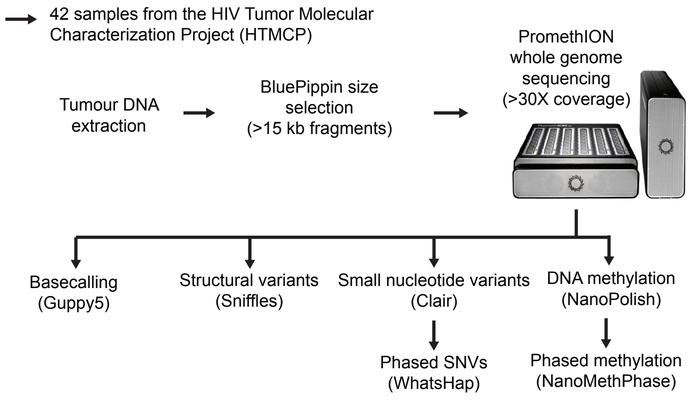
3) Assess the impact of HPV integration on local methylome

Methods

Whole genome sequencing

59 cervical cancer samples

- 17 samples from the Cancer Genome Atlas (TCGA)
- 42 samples from the HIV Tumor Molecular Characterization Project (HTMCP)



Detecting HPV integration

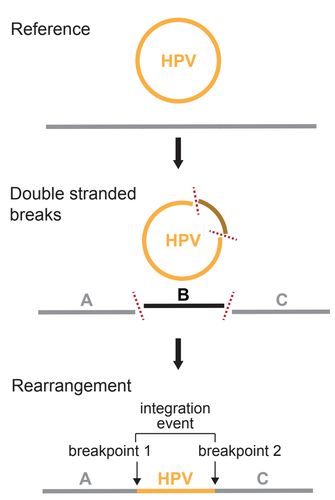
1. Detect HPV integration breakpoints as translocations between HPV and human chromosomes.

2. Group HPV breakpoints into integration events if they (1) co-occur on one or more of the same reads OR (2) map within 500 kb.

3. Subset sequencing reads from the integration event and create an assembly of the region.

4. Categorize integration event using the assembly and read alignment patterns.

5. Use the subsetted reads to investigate integrant lengths and methylation patterns.



Samples

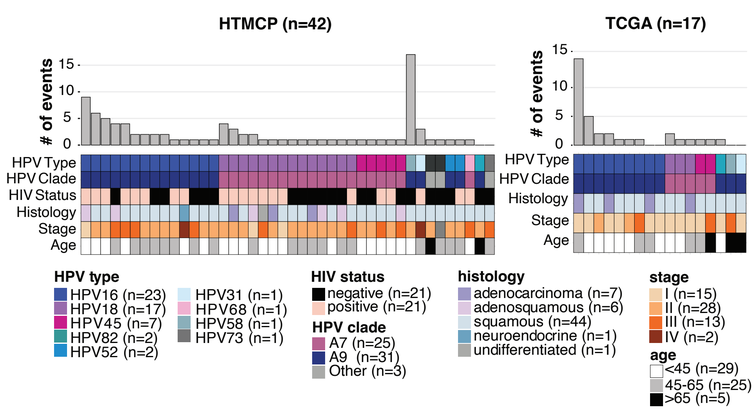


Figure 2: (A) The number of integration events across the HTMCP and TCGA samples along with the listed clinical and molecular characteristics of the patient samples.

Results

The frequency of HPV integration structures in cervical cancer

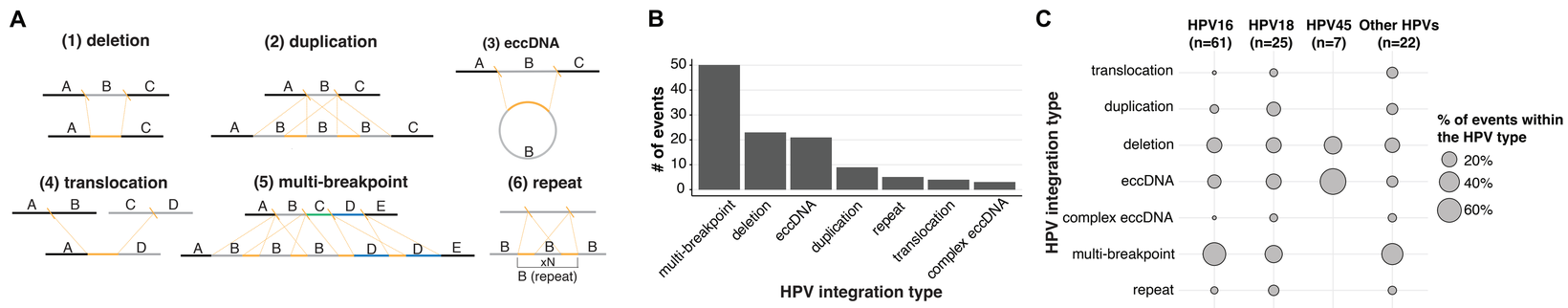


Figure 3: (A) Graphical representations of the six categories of HPV integration event structures (eccDNA = extrachromosomal circular DNA). **(B)** The number of each HPV integration event type across the cohorts. **(C)** The frequency of the integration event types between HPV types.

The three resolutions of two-breakpoint HPV integration events

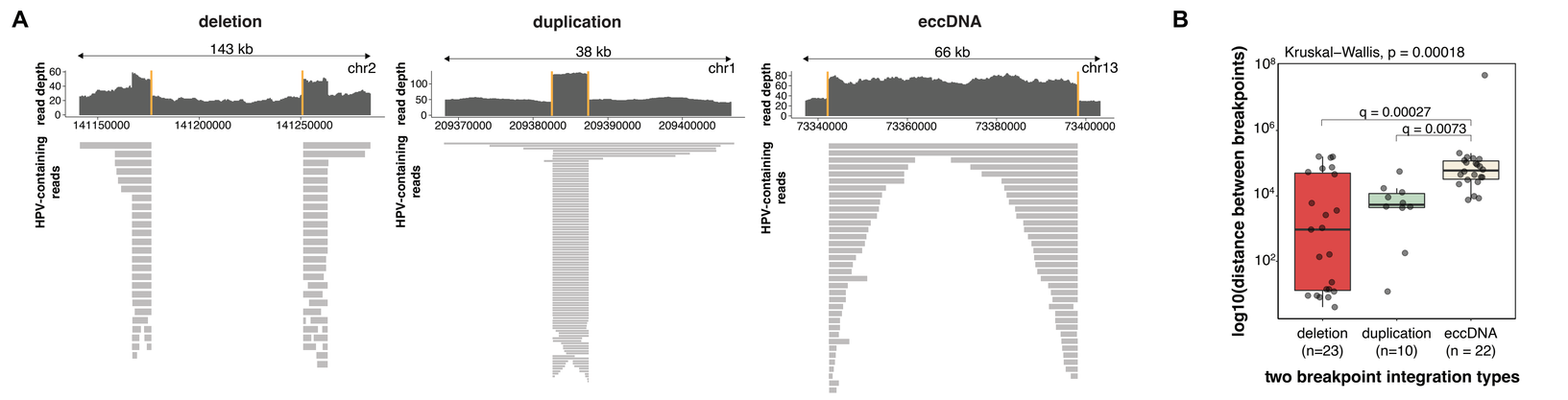


Figure 4: (A) An example of sequencing read coverage at each two-breakpoint integration type. Orange lines denote the HPV integration breakpoints. **(B)** The distance between breakpoints in the two-breakpoint integration types.

HPV integrants can have heterologous structures of the HPV genome

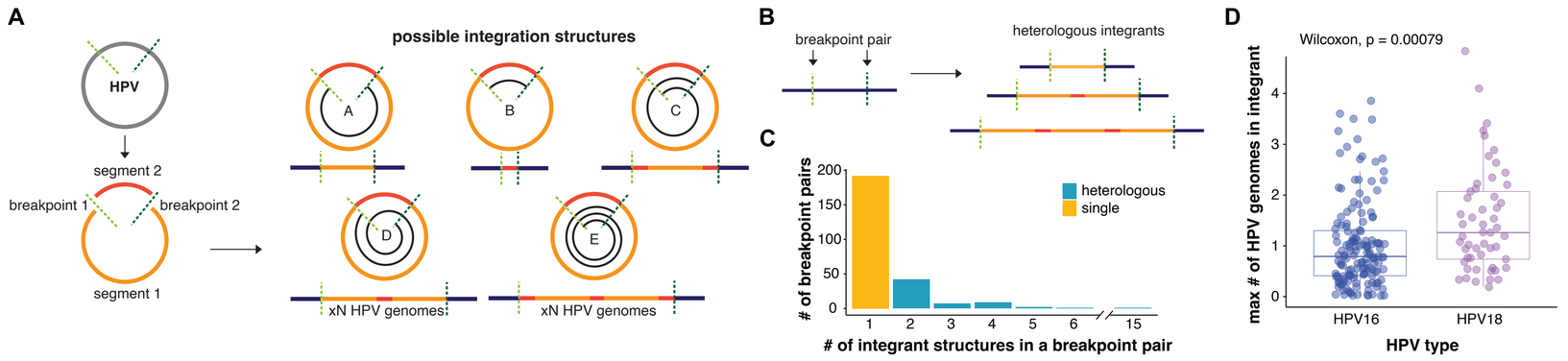


Figure 5: (A) Schematic of possible integrant structures in an HPV integrant with two breakpoints. **(B)** Schematic of how several heterologous integrants can exist between a single breakpoint pair, with the size of the integrant increasing by n HPV copies. **(C)** The number of integrant structures between all identified breakpoint pairs within the cohort. Integrants with 2+ identified structures were classified as heterologous. **(D)** The maximum size of the integrant structure in a breakpoint pair as shown by the number of HPV genomes in HPV16 and HPV18 integrants.

Distinct methylation patterns within the integrated HPV genome

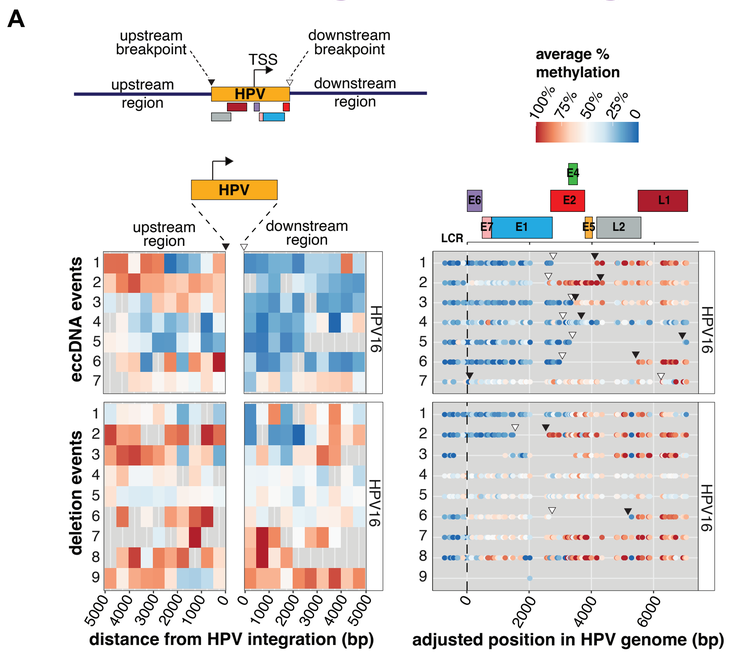


Figure 6: (A) The methylation frequency within and adjacent to HPV integrants in deletion and eccDNA integration events. The regions 5kb upstream and downstream (relative to the direction of HPV transcription) are divided into 500bp bins. Within HPV, the methylation of each CpG group is shown as a point. Open and closed arrowheads indicate where that breakpoint on HPV connects to the matching arrowhead on the human breakpoints shown on the left.

Haplotype-specific DNA methylation changes around HPV-integration

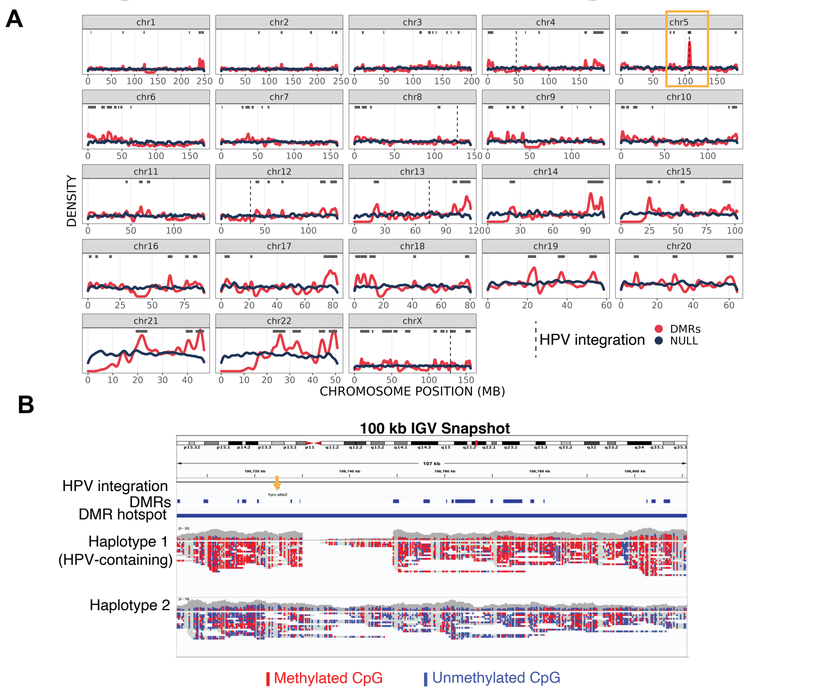


Figure 7: (A) The density of differentially methylated regions (DMRs) across the chromosomes in an example sample. Regions harbouring HPV integration are noted by a dotted line. HPV integrated region shown in (B) is marked by an orange box. **(B)** Integrative Genome Browser (IGV) snapshot of the DMR hotspot at the HPV-integrated region on chr5 from (A).

Conclusions

1: HPV integration events often involve structural variants that can be delineated and categorized using Oxford Nanopore Sequencing.

2: The HPV integrant structure can be heterologous between two breakpoint pairs, resulting in read evidance of multiple different integrant lengths.

3: The DNA methylation within and immediately adjacent to the HPV integrant has distinct patterns, with eccDNA integration events having the most consistant methylation.

4: HPV integration corresponds with long-range differential methylation between allelic haplotypes in the tumour.

***** Oxford Nanopore Technologies products are not intended for use for health assessment or to diagnose, treat, mitigate, cure, or prevent any disease or condition**

References

(1) de Sanjosé, Silvia, et al. "Burden of human papillomavirus (HPV)-related cancers attributable to HPVs 6/11/16/18/31/33/45/52 and 58." JNCI cancer spectrum 2.4 (2018): pky045.

(2) Akagi, Keiko, et al. "Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability." Genome research 24.2 (2014): 185-199.

(3) Porter, Vanessa L., and Marco A. Marra. "The Drivers, Mechanisms, and Consequences of Genome Instability in HPV-Driven Cancers." Cancers 14.19 (2022): 4623.

Acknowledgements

CANADA'S MICHAEL SMITH GENOME SCIENCES CENTRE

THE UNIVERSITY OF BRITISH COLUMBIA

CIHR IRSC

Canadian Institutes of Health Research | Instituts de recherche en santé du Canada