Deep-Sequencing Analysis of E6 and E7 Single Nucleotide Variants and Genome Methylation of High-Risk Human Papillomaviruses

Giulia Masi, Enrico Lavezzo, Giorgio Palù, Luisa Barzon
Department of Molecular Medicine, University of Padova, Padova, Italy

Background

Persistent infection by high-risk HPV (HR-HPV) is a necessary cause, but not sufficient, for the development and progression of cervical cancer. Different putative molecular markers associated with the neoplastic progression of cervical lesions, such as methylation or presence of single nucleotide variants (SNVs) of gene sequences of HR-HPV, are currently under study.

Objective

Aim of this study was the development of high-throughput deep-sequencing techniques to study the presence of SNVs in the E6 and E7 genes of HR-HPV types and the methylation status of the E2, L1, and L2 genes of HPV16 and HPV18 in clinical specimens from patients with HR-HPV infection

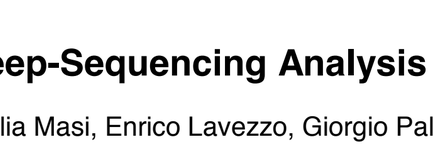

Figure 1. Map in HPV16 genome of the PCR primers which were designed for deep-sequencing analysis of E6/E7 SNVs and HPV genome methylation

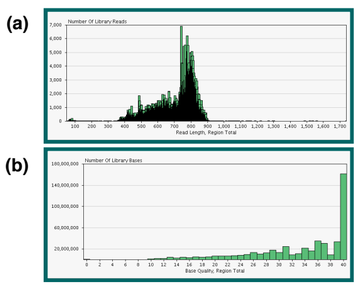

Figure 2. Representative QC results of HPV SNV and methylation analyses in 454 deep-sequencing run. Amplicon size distribution (a) and base quality score of reads (b) are reported.

Methods

PCR primer design

A dataset including all available HR-HPV genome sequences was retrieved from databases and used to design new sets of multiplex primers for SNV and methylation analysis on a 454 FLX+ platform (Roche).

For SNVs analysis, a set of 31 primers was designed to amplify a 800 bp region including the E6 and E7 oncogenes of all HR-HPV types in a single multiplex PCR reaction (Figure 1).

For methylation analysis, primers were designed to cover the E2, L1, and L2 region of HPV16 and HPV18 in four overlapping 800 bp amplicons (Figure 1). To reduce the risk of amplification biases, primer sequences were mapped in highly conserved sequences of HPV genome lacking putative CpG methylation islands.

Deep-sequencing

Both SNVs and methylation analysis of HR-HPV were performed by deep sequencing of amplicons in a 454 FLX+ NGS system (Roche) at 50,000X coverage. For methylation analysis, DNA was converted with sodium bisulfite before library preparation (Figure 2).

Analytical validation

The setup and analytical validation of the techniques was carried out on HeLa and CaSki cell lines, in WHO HPV genotyping proficiency testing samples, and in a group of clinical specimens positive for HR-HPV by Hybrid Capture 2 (Qiagen) testing and genotyped by INNO-LiPA HPV Genotyping Extra (Fujirebio Europe).

Results

Analysis of SNVs in the E6 and E7 genes of HR-HPVs

Analysis of HR-HPV E6 and E7 by 454 deep-sequencing in control and clinical samples demonstrated that the method could correctly identify all HR-HPV types both in single and in multiple infections (Table 1).

Table 1. Representative HPV typing results obtained by HR-HPV E6/E7 deep-sequencing in cervical swabs.

Sample ID	HPV types identified by LiPA	HPV types identified by E6/E7 deep sequencing	Reads count	Sample ID	HPV types identified by LiPA	HPV types identified by E6/E7 deep sequencing	Reads count
517	18, 31, (39, 52, 54)	31	50660	385	18, 31, (39, 52, 54)	31	45947
		18	3435			18	8915
507	18, 52, 66, 69/71, (39), HPVX	52	39860			52	22
		18	14913	877	18, 52, 66, 11, (39), HPVX	18	32879
		66	151			34	13895
665	16, 31, 6, 69/71, (52, 54)	31	69627			52	1012
		16	408			16	137
		18	4			66	68
		6	3			31	54
731	16, 18, (39)	18	42878			11	37
		52	40			91	12
		31	22	260	16, 39, 52, 6, 69/71, HPVX	39	59838
		16	19			34	15042
631	18, 56, (39, 74)	18	62182			16	6145
		56	1564			31	942
		31	47			18	3

SNV analysis of HR-HPV E6 and E7 in a group of cervical samples identified some SNVs that had not been previously reported. SNVs occurred in 100% of reads, confirming the high stability of HPV genome, which does not seem to accumulate mutations in the host (Table 2).

Table 2. Representative HPV E6/E7 SNV results obtained by E6/E7 deep-sequencing in cervical swabs.

Sample ID	HPV type	Gene	SNV	Amino acid variant	% of reads with SNV
517	HPV 18	E6	C287G	-	100
			T485C	-	100
			C549A	-	100
		E7	C751T	-	100
	HPV 31	E6	C285T	H>Y	100
		E6	A320T	-	100
		E6	G404A	-	100
		E6	A428G	-	100
665	HPV 16	E6	T350G	L>V	100
	HPV 18	E6	T485C	-	100
			C549A	-	100
		E7	C751T	-	100
	HPV 31	E6	A261C	I>L	100
			A297G	T>A	100
			A320T	-	100

HPV methylation analysis

For HPV16 and HPV18 methylation analysis, an adjusted preprocessing was carried out to avoid the elimination of bisulfite-treated sequences, rich in A and T homopolymers. In cervical samples, methylated CpG islands were correlated with pathological grade of the disease.

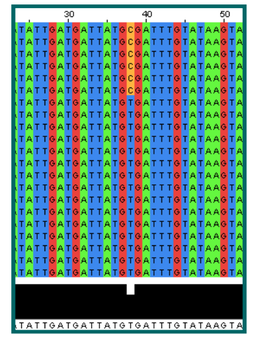

Figure 3. Representative analysis of HPV methylation. A detail of aligned reads of HPV18 L1 is represented, showing a CpG methylated island.

Conclusion

This study has enabled the development of deep-sequencing methods based of 454 FLX+ technology for methylation and SNV analysis in gene sequences of HR-HPVs. These methods are currently under clinical investigation as potential biomarkers of tumor progression in a longitudinal study of patients with HR-HPV-related cervical lesions.

Funded by MIUR grant no. 2010Z9FLM8

Università degli Studi di Padova

DMM Dipartimento di Medicina Molecolare Università di Padova