

Functional analysis of genome polymorphisms in human papillomavirus 11 associated with recurrent respiratory papillomatosis with different clinical severity

Zsófia Nagy<sup>1</sup>, Gábor Kardos<sup>1</sup>, Mihály Szinai<sup>1</sup>, Attila Szűcs<sup>2</sup>, Zoltán Pethő<sup>3</sup>, Krisztina Szarka<sup>1</sup>

<sup>1</sup>Department of Medical Microbiology; <sup>2</sup>Clinic of Otorhinolaryngology and Head & Neck Surgery; <sup>3</sup>Department of Biophysics and Cell Biology, Faculty of Medicine, University of Debrecen, Nagyerdeikrt 98. Debrecen, Hungary

Polymorphisms of the long control region (LCR) of human papillomavirus (HPV) 11 may modify severity of respiratory papillomatosis (RRP). We also found unique amino acid polymorphisms in virus proteins. We analysed a HPV11 genome from a novel patient, and earlier reported and newly found polymorphisms *in silico* to assess their role in the clinical course of RRP.

The new patient has an adult-onset RRP with moderate dysplasia. All LCRs were examined for transcription factor binding sites using the tool <http://trap.molgen.mpg.de/cgi-bin/home.cgi>. *In silico* protein modelling (Molecular Operating Environment 2014.9) was performed on the following: i) The two polymorphisms (A476V and S486F) in L1 present in the virus from the papilloma but not from the healthy mucosa of Patient 5 with aggressive RRP. ii) E1 polymorphisms A72E and N100T in Patients 4 and 9, respectively, with moderately severe RRP. iii) E2 polymorphisms Q86K in Patient 4; S245F and N247T in Patient 9; K308R in multiple patients.

Complete genome of the newly enrolled Patient 10 contained 21 polymorphisms (16 silent, A45S in E7, S78L in E4, K308R in E2, I28F and V41L in E5a), a deletion and an insertion in the LCR; all known from other genomes. It also contains a large unique deletion in ORF E2-E4 (NT3381-3438) causing a frameshift with an early stop codon.

The identified seven LCR polymorphisms resulted in 28 new binding sites of 24 transcription factors, modified 66 binding sites of 52 transcription factors; no binding site was lost.

L1 polymorphisms are located in C-terminus playing a role in capsomer formation and possibly in heparan-sulfate binding of virions. In protein models both were involved in binding between monomers; S486F led to a new interaction with E382 and with A385 of the neighbouring monomer of the capsomer (free enthalpy -1.2 kcal/mol).

Both E1 polymorphisms are in the N-terminal regulatory region, but do not involve nuclear transport signal sequences. As this part of the protein is unstructured, structure modelling was not possible.

E2 polymorphism Q82K is located in the E1 binding domain, but due to lack of E1 model, this interaction could not be modelled. Polymorphisms S245F and N247T are in the hinge region of E2 containing phosphorylation sites involved in E2 stability and chromatin binding. As they lead to loss of one phosphorylation site and to gaining another, and alter the surface charge of E2, they may affect phosphorylation patterns. Polymorphism K308R is in the DNA binding domain, but is positioned distantly from the DNA during interaction.

The truncated E2 of the HPV11 from Patient 10 lost a large part of its DNA binding domain, and though the alpha-helix serving as DNA binding motif is preserved, its surface charge alters from basic to acidic, most probably impairing DNA binding.

In summary, protein modelling supports the role of some of these unique polymorphisms in pathogenesis of RRP.

K. Szarka was supported by János Bolyai Research Scholarship of the Hungarian Academy of Sciences.