Novel polyomavirus and papillomavirus detected in gilthead seabream infected by lymphocystis disease virus

Alejandro Labella[1], David Carrera[1], Alberto López-Bueno[2], Alí Alejo[3], Juan J. Borrego[1], Dolores Castro[1*]

[1]Universidad de Málaga, Departamento de Microbiología, Facultad de Ciencias, Málaga, Spain; [2]Centro de Biología Molecular Severo Ochoa (Consejo Superior de Investigaciones Científicas and Universidad Autónoma de Madrid), Cantoblanco, Madrid, Spain; [3]Centro de Investigación en Sanidad Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Valdeolmos, Madrid, Spain

Lymphocystis disease virus (LCDV), a member of the genus *Lymphocystivirus*, family *Iridoviridae*, is the etiological agent of the lymphocystis disease (LCD), a common pathology that has been described in more than 150 different fish species worldwide. Direct sequencing of the virome of lymphocystis lesions from affected gilthead seabream (*Sparus aurata*) was used to obtain the complete genome sequence of a new LCDV species, named LCDV-Sa, that is the largest vertebrate iridovirus sequenced to date. This approach allowed us to assemble the full-length circular genomes of two previously unknown viruses, tentatively identified as members of the *Polyomaviridae* and *Papillomaviridae* families, and named Sparus aurata polyomavirus 1 (SaPyV1) and Sparus aurata papillomavirus 1 (SaPV1), respectively. SaPyV1 genome is a circular 7,299-bp-long DNA with a 52.1% GC content, and contains five non-overlapping ORFs carried on opposite DNA strands in an organization reminiscent to that found in polyomaviruses. Phylogenetic analyses based on the amino acid sequence of the large T antigen and the VP1 proteins revealed that SaPyV1 clustered with the other currently available polyomavirus-like full-length genomes obtained from fish. SaPV1 genome corresponds to a circular 5,748-bp-long DNA and has a GC content of 39.5%. Although this genome is smaller than that of previously described papillomaviruses (about 8 kbp), it presents a typical papillomavirus organization, with seven ORFs carried on the same strand. Similarity searches identified distant orthologues of the early E1 and E2 proteins involved in replication and transcription, and late structural proteins L1 and L2. Furthermore, one of the others ORFs encodes a small protein (52 amino acids) that contains both a pRB binding domain (LXCXE) and a C-terminal PDZ class 2 binding motif, with are elements typically present in the longer E7 and E6 proteins found in most known papillomaviruses. The conserved genomic organization, the similarity of the main proteins, and the phylogenetical analysis, based on the amino acid sequence of the L1 protein, support that SaPV1 is the first member of the *Papillomaviridae* family described in fish. Epizootic surveys carried out in gilthead seabream farms in the Mediterranean area showed that LCD is frequently associated with the concurrent appearance of one or both of the new viruses, 98.3% of diseased fish being positive for SaPyV1 and/or SaPV1. In LCDV-infected asymptomatic fish, SaPyV1 and/or SaPV1 were detected in 32.1% of the animals analysed.

This study was funded by project P12-RNM-2261 from Junta de Andalucía.

Contact e-mail: dcastro@uma.es