Final Report

Frog and toad herpesvirus detection 2018-2019: Differences in Bufonid herpesvirus 1 and in toads and Ranid herpesvirus 3 in common frogs.

Project financed by FOEN (contract n. R414-1866)

submitted by Francesco C. Origgi,

March 31st, 2021

Table of content

Investigation into the genetic features of Ranid herpesvirus 3 (RaHV3) and Bufonid herpesvirus 1 (BfHV1). Project summary

Two new herpesviruses have been recently discovered. *Ranid herpesvirus* 3 (RHV3) was found associated with a proliferative dermatitis in common frogs (*Rana temporaria*) (Origgi et al., 2017), whereas *Bufonid herpesvirus* 1 (BfHV1) has been found associated with a proliferative dermatitis in common toads (*Bufo bufo*) (Origgi et al., 2018). Very little is currently known about these viruses and it is not clear if there is one or distinct strains of RHV3 and BfHV1 circulating in Switzerland and we do not know if their putative differences might be associated with distinct virulence and associated pathology phenotype in their respective hosts.

In order to clarify these aspects and better characterize these newly discovered infectious agents we proposed to sequence the genomes of all the available strains of RHV3 and BfHV1 collected in Switzerland and in other European countries. The sequencing of these viruses will reveal if one or more viral species/variants are present on the Swiss territory and in Europe, and if any of them is particularly associated with decline of the associated local population of common frog or common toad, respectively. The sequencing will be carried out with state of the art technology (Next generation sequencing) and the obtained information will be critical in understanding the role of these viruses in amphibian disease ecology.

Achievements:

We have carried out an extensive genetic characterization of the largest repository of Batrachoviruses (amphibian herpesviruses) in the world. This investigation was critical in unraveling several critical aspects of these viruses. Of paramount importance has been the discovery of at least two distinct lineages within both RaHV3 and BfHV1. This discovery provides critical genotyping tools, which will open the way to future investigations clarifying possible associations between the distinct lineages and RaHV3/BfHV1 outbreaks, distinct pathology phenotypes, population dynamics including decline. The observation that the distinct lineages appear to be geographically linked, provide a powerful tool to monitor the distribution and diffusion of these agents and if these movements are associated with disease outbreak (also associated with other pathogens) and population decline. Thanks to this research project, new tools to understand RaHV3 and BfHV1-associated disease ecology in Switzerland and Europe are finally available.

Furthermore, the higher genetic conservation of BfHV1 is suggesting a more ancient origin of BfHV1 with a more recent evolution of RaHV3. Both viral species have revealed that a gene encoding for a putative immunomodulatory protein similar to the Tumor necrosis factor receptor (TNFr) is highly conserved, which is consistent with the results of parallel investigation, making this gene a critical target to understand the RaHV3 and BfHV1 associated pathogenesis.

Ranid herpesvirus 3

Isolates

A total of 7 DNA samples obtained from common frog (Rana temporaria; n=5) and agile frog (Rana dalmatina; n=2) were available (Table 1). Of these 4 were originally from Switzerland, 2 from Hungary and 1 from Italy. The availability from isolates from close countries to Switzerland (Italy) and from relatively more distant (Hungary) was considered important to assess if genomic variabilities was associated with a geographical component. All the isolates were obtained from frogs showing the classic clinical signs, consisting in proliferative gray skin patches. All samples were obtained from fresh tissue, with the exception of BS_1994, which was obtained from formalin-fixed, paraffin-embedded samples.

Table 1. RaHV3 isolates

RaHV3 isolate ID	Host	Origin	Year	Genome length (nt)
BS_1994	*R. dalmatina*	Italy	1994	207,902
FO_2015	*R. temporaria*	Switzerland_ZH	2015	207,914
W17_4184	*R. temporaria*	Switzerland_Ne	2017	207,900
W18_2160	*R. dalmatina*	Switzerland_Tg	2018	207,961
RaHV3_D	*R. temporaria*	Switzerland_Gr	2019	207,833
RaHV3_HU_Sk	*R. temporaria*	Hungary	2019	207,919
RaHV3_HU_Sw	*R. temporaria*	Hungary	2019	207,926

Isolate sequencing

Next generation sequencing was carried out by a private Biotech Company (Fasteris SA, Geneva) using the Illumina platform (2x100 bp - 100 million clusters). All the sequences obtained were then mapped to the reference strain FO_2015. The coverage obtained for the different isolates ranged from 89.73% (BS_1994) to 99.98% (RaHV3_D). A consensus sequence was obtained from each isolate by patching the missing regions using the reference strain FO_2015. The genome sequence length of all the isolates were very close, ranging from 207,833 (RaHV3_D) to 207,961 (W18_2160) (Table 1).

Isolate analysis

The obtained consensus sequences of each isolate were compared among each other using the software package Geneious prime (Geneious Prime® 2020.2.4; www.geneious.com). The parameters examined included pairwise nucleotide identity, identification of highly conserved, moderately conserved and poorly conserved genes across the available isolates, identification of all the nucleotide changes resulting in amino acid substitutions, open reading frame (ORF) modifications (insertion of stop codon, suppression of stop codon, insertion of start codon, suppression of start codon); large insertions and deletions. Figure 1 shows the aligned genomes with predicted ORFs.

ORF conservation was assessed according two distinct parameters. The first parameter was the pairwise identity. An arbitrary separation of three groups of genes (highly, moderately, and poorly conserved) was carried out using as cutoffs values up to 94.9% for poorly conserved, up to 97.9% for moderately conserved and up to 100% for highly conserved genes (Table 2.). The second parameter considered was the ORF conservation across isolates (same reading frame). A highly conserved gene was that showing a fully conserved frame across all the isolates, whereas a moderately conserved gene showed a conserved ORF across all isolates, but the frame size might have included variations, such as frame fragmentation, shortening or extension. Finally, a poorly conserved gene showed the absence of the specific ORF in one or more of the isolates (Table 2). According to these parameters, the RaHV3 isolates contained 16 poorly conserved genes, 104 moderately conserved and 66 highly conserved ones. Overall, the RaHV3 ORFs accounted for 170 out of 186 with a pairwise conservation equal or higher than 95%. Poorly ORF conservation was higher (43.7%) among the group of genes with lower pairwise identities, and low (1.5%) among those with high pairwise identity, as expected. On the contrary, ORF conservation was highest (47%) in the group with high pairwise identity and lowest

(31.3%) in the group with low pairwise identity. Moderately conserved ORFs were predominant in the group of genes with intermediate pairwise identity (Table 2).

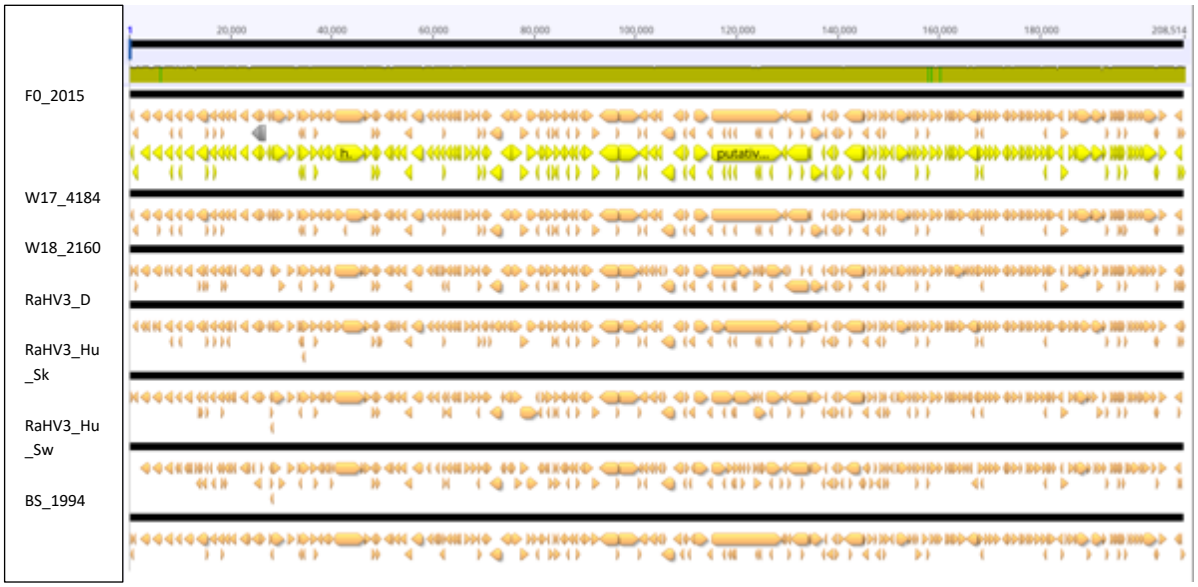

Figure 1. Alignment of the RaHV3 isolates with their ID (left), genome length ruler (top) and ORFs prediction. The ORFs of the reference strain (FO_2015) are shown in yellow.

Table 2 below, shows the color-coded genes, identified as highly conserved (yellow), moderately conserved (green) and poorly conserved (blue). This grouping was carried out on the basis of the pairwise identity of each of the specific gene (shown as percentage value in parenthesis; higher percentage=higher conservation) and on the level of conservation of the specific ORFs. Specifically, for the ORF conservation, a highly conserved gene showed a fully conserved frames across all the isolates, whereas a moderately conserved gene showed a conserved ORF across all isolates including frame size variation. Differently a poorly conserved gene showed the absence of that specific ORF in one or more of the isolates.

Table 2. Pairwise identity RaHV3 genes and conserved open reading frames (ORFs)

ORF	Pairwise identity (%)	ORF	Pairwise identity (%)	ORF	Pairwise identity (%)
153a	86.6	8	95	52	98
96	91	118	95.2	62a	98
86d	91.1	149	95.3	84	98
36	91.2	13	95.4	87	98
86e	92.4	42	95.4	88a	98
2	92.8	145a	95.4	90	98
7a	92.8	156	95.4	99	98
157	93.1	15	95.6	104	98
86b	93.4	46	95.6	129	98
35	93.9	152	95.6	136	98
1	94	86	95.7	142	98
86a	94.2	139	95.7	38	98.1

3	94.5	156a	95.7	50	98.1
141	94.5	9	95.8	55	98.1
153	94.6	20	95.8	60	98.1
44	94.8	86c	95.8	105	98.1
		45	95.9	121	98.1
		151	95.9	135	98.1
		123	96	21	98.2
		155	96	34a	98.2
		5	96.2	53	98.2
		106	96.2	57	98.2
		10	96.3	95	98.2
		7	96.4	102	98.2
		40	96.4	137	98.2
		86f	96.4	51	98.3
		131	96.4	58	98.3
		6	96.5	68	98.3
		11a	96.5	93	98.3
		18	96.5	69	98.4
		27	96.5	71	98.4
		30	96.5	80	98.4
		122	96.5	107	98.4
		132	96.5	126	98.4
		145	96.5	127	98.4
		41	96.6	49	98.5
		117	96.6	63	98.5
		31	96.7	72	98.5
		33	96.7	73	98.5
		43	96.7	81	98.5
		43a	96.7	83	98.5
		77	96.7	74	98.6
		133	96.7	88b	98.6
		147	96.7	119	98.6
		148	96.7	23	98.7
		32	96.8	66	98.7
		91	96.8	75	98.7
		98a	96.8	89	98.7
		19	96.9	92	98.7
		26	96.9	113	98.7
		11	97	154	98.7
		12	97	125	98.8
		29	97.1	67	98.9
		120	97.1	74a	98.9
		131a	97.1	97a	98.9
		140	97.1	110	98.9
		17	97.2	111	98.9
		25	97.2	70	99
		28	97.2	144	99.1
		64	97.2	4	99.3
		78	97.2	109	99.4
		124	97.2	112	99.5
		16	97.3	128	99.5

		24a	97.3	109a	99.6
		39	97.3	108	99.9
		52a	97.3	110a	100
		65	97.3		
		85	97.3		
		119a	97.3		
		130	97.3		
		76	97.4		
		118a	97.4		
		24	97.5		
		115	97.5		
		54	97.6		
		97	97.6		
		134	97.6		
		143	97.6		
		22	97.7		
		34	97.7		
		48	97.7		
		88	97.7		
		98	97.7		
		100	97.7		
		103	97.7		
		150	97.7		
		14	97.8		
		37	97.8		
		59	97.8		
		64a	97.8		
		79	97.8		
		82	97.8		
		94	97.8		
		101	97.8		
		116	97.8		
		146	97.8		
		47	97.9		
		56	97.9		
		61	97.9		
		62	97.9		
		76a	97.9		
		76b	97.9		
		114	97.9		
		138	97.9		
547		346010		31341	
16		104		66	
31.3; 25;43.7%		32.7;57.7;9,6%		47;51.5;1.5%	

The moderately conserved genes, with conserved ORFs across all isolates, but showing variability in its length and integrity were further examined to determine which were the nucleotide changes, which were determining the observed variations, which are all summarized in Table 3 below.

Table 3. Nucleotide variations within moderately conserved RaHV3 ORFs.

Color coding shows fully conserved genes (yellow), moderately conserved (green) and poorly (blue).

Strain 1= FO_2015; Strain 2= W17_4184; Strain 3=W18_2160; Strain 4= RaHV3_D; Strain 5= RaHV3_HU_Sk; Strain 6= RaHV3_HU_Sw; Strain 7= BS_1994

ORF	Conserved	Notes
1	No	In 3, 4, 5 and 6 not present; 3 and 6; (575); 4(600-3, 609-12); 3,5,6(T/C, 635)
2	No	Not present in 6; Truncated in 7; 7 (1288); 6 (1282)
3	Partially	4 has a split frame with two deletions (3174 and 3142), which determine a stop codon. A second frame restart at 3002. All the other strains have only one of these two deletion in that portion of the ORF (do not consider the del in strain1)
4	Partially	4 has a T deletion at 6014 resulting in a stop codon in frame at nt 5058. There is a new start in frame with the ORFs of the other strains at nt 4919. 3 has an A to G substitution at 4589 suppressing the stop codon in frame, with the ORF being 15 nt longer until the following stop codon is reached.
5	Partially	3 has a mutation (T to A) at nt 7489 introducing a stop codon. A new ORF restart at nt 7354 in frame with the other strains. Strain 3, 5 and 6 have a mutation (T to A) at nt 6872 suppressing the stop codon and extending the ORFs of 10 nt.
6	No	3,5,7 not present
7	Partially	In 5 there is an insertion (G) at nt 9567 creating a stop codon at nt 9522; anew ORF starts at 9527 in frame with the other strains; A C to A mutation in 3, 5 and 6 at 8986 suppress the stop codon at nt 8987 and prolongs the frame of 15nt; in 3, 5 and 6 there is a 27 nt insertion between nt 9975 and 10001.
7a	No	Not present in 5,6,7
8	Partially	In strains 3, 5 and 6 there is a deletion (C) at nt 12437 with start codon position 120nt downstream the start of the other ORFs; A deletion (TT) is present in strains 1,2,4 and 7 at nt 12422 and 23, but no effect on the ORF; Strain 6 has a deletion (T) at nt 11843 introducing a stop codon in frame at nt 11825 and a new restart at 11829; A 4nt deletion (CTTT) is present at nt 11556-53 of the same strain resulting in a stop codon in frame at 11516 and a new restart at 11364 with the stop codon in frame with the ORFs of the other strains.
9	No	In 3 there is an insertion (A) at nt 14680 resulting in a stop codon creation at the same position. There is a restart of a new ORF in frame with the other strains at nt 14691; In 4 there is a deletion (A) at nt 14709 resulting in a stop codon in frame at nt 14703 and a restart of a new ORF in frame with the other strains at nt 14701; 5 and 6 have an insertion (C) at nt 14250 risulting in a stop codon at 14238 and a new start at nt 14105 ending in frame with the other strains; 6 has an insertion (T) at nt 15135 resulting in a stop codon at 15107 and a restart of the frame at 15169.
10	No	Not present in 6

11	Partially	Strains 5 and 6 have a deletion (A) at nt 16903 resulting the starting of the ORF at nt 16869m 92 nt downstream the start in the other strains; 6 has an insertion (A) at 16248 with the formation of a stop codon in frame at nt 16227 and a restart at 16262 ending in frame with the other strains.
11a	No	Not present in 3, 5, 6, 7
12	Partially	Strain 6 has an insertion (T) at nt 18490 resulting in the positioning of the start codon 124nt downstream the start in the other frames at nt 18467; An insertion (C) at nucleotide 17995 creates a stop codon in frame at nt 17978 and a restart at 17938 ending in frame with the other strains; Potential presence of a previously undetected ORF between 18093 and 18362 present in all strains. This frame is longer in 3, 5, 6 and 7 (18093-18861) because the mutation of a T to a C, suppressing the stop codon at that position.
13	Yes	
14	Partially	Strains 3 and 5 have a deletion (T) at nt 20526 and another deletion (A) at 20528 resulting in a stop codon in frame at 20503; strains 4 and 6 have an insertion (A) at nt 20527 with a stop codon in frame at the same position. Frames restarts at nt 20106 for strains 3 and 4, at nt 20454 for strain 5 and at nt 20200 for strain 6.
15	Partially	ORF conserved in all strains, however, it is 369nt longer in strain 3 because of an upstream codon, created by a triple deletion (T at 23679 and T and G at 23736-7). An overlapping frame, but not fused with the main one is also present in strain 6, but because of the lack of the double deletion in 23736-7 has a stop codon in frame at nt 23427, restarting at 23417 with the main frame of all the other strain.
16	Partially	3 and 5 are longer and 6 have a longer ORF starting at nt 25817, 66 nt upstream the common ORF to all the strains. A triple deletion (TAT/G) is present in all the strain with longer ORFs at position 25724-6.
17	Partially	3 and 6 have a double deletion (TT) at nt 26372-3 creating a stop codon in frame at nt 26398
18	No	Not present in 3,5,6,7
19	Yes	
20	Partially	Strains 3, 5 and 6 start at nt 28241; strains 1 and 4 starts at 28406 and strain 7 starts at 28153. This because strains 1, 2 and 4 have a deletion (T) at nt 28357. Furthermore, strains 1,2,3,4,5,6 have an insertion (G/A) at position 28353. All these mutations cause stop codons in frame variably shortening or extending the ORFs, accordingly; Strain 2 has a mutation (G to T) at position 29195 creating a stop codon there. The frame restarts at 29258 and ends at 30489 because of a deletion (A) at 30468 creating a stop codon in frame at 30487. The same occurs in strains 5 and 7; Strain 3 and 6 have an insertion (A) at nt 29685 creating a stop codon in frame at 29789. It restarts at 29682 and stops at 30565 because of a deletion (T) at 30524 causing a stop codon in frame at 30563. Same stop and mutation also in strain 6. Strain 4 has an insertion (A) at nt 30523 causing the formation of a stop codon in frame at 30523.
21	Partially	Strain 2 has a mutation (G to T) at nt 32400 introducing a stop codon there and truncating the ORF at that level.
22	Yes	

23	Partially	Strains 3,5,6,7 have a longer frame starting at 33924, differently from the others, which start at 33804. This is because a mutation (T to G) at 33863 in strains 3, 5, 6 and 7.
24	Partially	3,5,6 have an insertion (T) at 35593 resulting in a stop codon in frame at 35615. 7 has a deletion at 35590. Strain 4 has a deletion (A) at nt 34166 leading to a stop codon in frame at 34303. Restarts at 34196
24a	No	Not present in 3, 5 and 6
25	Yes	
26	Yes	
27	Partially	Strain 4 is 89 bp shorter than the others because of a deletion (C) at 39732 nt.
28	Partially	Strain 4 has a shorter frame because of a double mutation (CC to TT) introducing a stop codon at nt 38330 and 301; Strains 5 and 6 have a deletion (G) at nt 38403 resulting in a stop codon at nt 38361. A ORF starts for both strains at nt 38321, which end together with the main frames of strains 1,2,3, 7 .
29	Partially	Strains 3 and 6 have a double insertion (TC) at nt 39642 and 3 resulting in a stop codon in frame at nt 39651; A new ORF starts for both strains at nt 39623; However, in strain 3 the new ORF ends together with that of strains 1,2, 4, 5, 7; Differently, strain 6 new ORF has a deletion (A) at 39911 leading to a stop codon in frame at nt 39936. An additional frame starts at nt 39962 and ends together with the main frame of strains 1,2,4,5,7.
30	Partially	Strain 2 has a deletion (C) at nt 42572, creating a stop codon in that position; A new frame start at 42655 and ends together with the main frame; Strain 4 has an insertion (C) at nt 41998 leading to a stop codon in frame at position 4201 and a new frame starts at 42124; Strain 3 has a deletion (T) at 45668 leading to a stop codon in frame at nt 45710; A new frame starts at 45739 with an end together with that of the main frame.
31	Yes	
32	Yes	
33	Partially	Strains 3 and 5 have a double mutation (CG to GA) at nt 48513 and 14 creating a stop codon in that position, shortening the ORF of 28bp.
34	Partially	Strain 4 has a deletion (C) at 4878 leading to a starting codon 121 nt downstream the main frame; An insertion (C) at 49217 results in a stop codon in frame at nt at 49356. A new frame starts at nt 49397 and ends at 49681 as strain 1. Strains 3, 5, 6, 7 ends at 49675; strains 2 has a deletion (A) at 49674 and an insertion (T) at 49680 with a 11nt longer frame 3'; Strains 3 and 5 have an insertion at 49675 with a 5nt shorter frame 3'; strains 6 and 7 have an insertion (AAAAAT) at nt49675-80 and a 6nt shorter frame 3'.
34a	Yes	
35	Yes	
36	Yes	
37	Partially	Strain 4 has a double insertion (CT) at nt 52878 and 9 resulting in a stop codon in frame at 52831; a new frame starts at 52643 and ends with the main frames; Strain 6 has a double deletion (AA) at nt 52391 and 2.

38	Partially	Strain 2 has a deletion (A) at nt 52500 resulting in a stop codon in frame at 53457; Strain 4 has a deletion (T) at 53602 resulting in a stop codon in frame at nt 53544.
39	Yes	
40	Partially	Strain 3 has a mutation (G to A) at nt 55964 leading to a stop codon in frame at the same site; Strain 5 and 6 have a deletion (A) at nt 55992 (present also in strain 3), which creates a stop codon in frame at 55908.
41	Partially	Strain 6 has a deletion (A) at 59697, shortening the frame of 353 nt (start codon downstream the main frame)
42	Partially	Strains 3 and 7 have a mutation (G to A) at nt 60674 creating a stop codon in that site. A new frame starts at nt 60682 and ends with the main frame. Strain 6 has a double deletion (TA) at 60538-9 creating a stop codon in frame at nt 60523 with shortening of the frame of 439 nt.
43	Partially	Strain 2 has a mutation (G to A) at 62291 leading to an early start codon at 62292 and a 21nt longer frame 5'; Strain 3 has a deletion (T) at 61886 causing a stop codon at 61883 shortening the frame of 522 nt; Strain 5 has a deletion (T) at nt 62275 leading to an early start codon 141bp upstream the main frame, but also causing a stop codon in frame at 61713 and a shortening of the frame of 352nt. Strain 6 has the same shorter frame with the same early stop codon because of two insertions (G at 61787 and A/T at 61785), in comparison with strains 1,4 and 7, which have the full frame.
43a	Yes	
44	Yes	
45	Partially	Strain 6 has an insertion (C) at nt 63748 creating a stop codon at 63727 and a shortening of the frame of 217nt.
46	Partially	Strains 2 and 5 have a deletion (A) at nt64348 leading to a 34nt longer frame 3'; Strains 3 and 6 have two deletion (G) at 64350 and another one (A) at 64348, leading to a 29nt longer frame 3'.
47	Yes	
48	Yes	
49	Yes	
50	Partially	Strains 5 and 6 have a mutation (A to G) at nt 68774, suppressing the stop codon at the same site and extending the frame of 108 nt with a new stop codon at 68666 nt.
51	Yes	
52	Yes	
52a	No	Missing in 5,6,7
53	Partially	Strain 4 has a deletion (A) at nt 70995 creating a stop codon in frame at 71009. A new frame starts at 70989 and ends with the main frame. Strain 6 has a deletion (A) at nt 70605 with a shortening of the frame of 221 nt at the 5 prime site, starting at nt 70699.
54	Partially	Strain 4 has a deletion (A) at nt 71769 causing a stop codon at 71764. A new frame starts at 71633 and ends with main frame.
55	Partially	Strain 4 and 6 have a deletion (T) at 74500 causing a stop codon at 74461. A new frame starts at 74288 ending with main frame, while strain 6 has an insertion (G) as strain 5, at nt 73787 causing a stop codon at 73745. A new frame starts at nt 73732 (as in strain 5) ending with main frame. Strain 5 has the first frame starting with the

		main one, but ending as the second of strain 6. The second frame of strain 5 overlaps with the third of strain 6.
56	Partially	Strain 6 has an insertion (C) at nt 76237 leading to a stop codon in frame at nt 76240. A new start at 76208 end with the main frame.
57	Partially	Strain 5 has a deletion (T) at nt 78742 suppressing the stop codon in frame and fusion with ORF 58.
58	Partially	Strain 5 has ORF 57 and 58 fused together. Strain 7 has a deletion (T) at nt 79495 causing a stop codon at 79551 and a restart at 79631, which then ends with main frame.
59	No	Missing in 4 and 6
60	Yes	
61	Partially	Strains 3, 5, 6 and 7 have a mutation (A to G) at nt 81407 resulting in an early start codon with a 9nt longer ORF 5'.
62	Partially	Strain 4 has a deletion (A) at 82369 causing a later start with start codon at 82841. A mutation (C to G) in strain 6 and 7 at nt 82807 creates a start codon at 82805. The same mutation is also present in strain 3 but it is ininfluencial.
62a	Yes	
63	Partially	Strains 3, 5, 6 and 7 have a later start with a starting codon at 83686 resulting in a 75nt shorter frame. This because of a mutation (G to T) at nt83860 in strains 3, 5, 6, 7 resulting in the formation of a stop codon in that position.
64	Yes	
64a	Yes	
65	Yes	
66	Partially	Strain 2 has a deletion (T) at nt 86098 with a downstream start codon at 86167, corresponding to a 146nt shorter main frame; Strain 6 has a deletion (T) at nt 86352 leading to a stop codon in frame at 86363 and a restart at 86371 and ending as the frames of the other strains.
67	Yes	
68	Yes	
69	Partially	Strain 4 has a deletion (A) at nt 87856 with a stop in frame at 87835 resulting in a 206nt shorter frame.
70	Yes	
71	Yes	
72	Partially	Strain 7 has a deletion (T) at nt 90990 with a stop codon at 91058 and a shortening of the frame of 507nt.
73	Partially	Strain 7 has a deletion (A) at 91508 leading to an early start with a start codon at nt 91315 with a 181bp longer frame.
74	Yes	
74a	Yes	
75	Partially	Strain 4 has a deletion (T) at 97446 with stop codon at 97450 in frame, with a new start at 97450 and end of the frame together with those of the other strains
76	Partially	Strain 3 has a deletion (A) at nt 101338 with a stop codon in frame at 101290 and a new start at 101165 with ending as the other main frames; Strain 6 has a deletion (A) at 101338 creating a stop codon there. It then starts a new frame at 101165 and ends with the other main frames. Strains 3, 5, 6 and 7 have a mutation (T to A) at 102361 determining an early start codon with a 42nt longer frame 5'.
76a	Yes	

76b	No	Missing in 5 and 7
77	Partially	Strain 3 has a deletion (CTGG) at nt 103593 and an insertion (CGAATA) at 103577 resulting in a stop codon at 103560; Restart at 103463 and ending with main frame; Strain 6 has an insertion (CGAATAATAATTT) at nt 103589 resulting in a stop codon at 103520; restart at 103463 and ending in frame with the other strains.
78	Partially	Strain 3 and 5 have a deletion (A) at nt 104246 resulting in a stop at 104220; Strain 6 has an additional deletion (G/A) at 104249 resulting in a stop codon at 104247nt.
79	Yes	
80	Partially	Strain 7 has a deletion (A) at nt 108872 with a stop at 108855; new start at 108929 and ending with the other main frames.
81	Yes	
82	Partially	Strain 6 has a mutation (A to T) at nt 109891 resulting in a new stop codon shortening the frame of 36nt 5'.
83	Partially	Strain 6 has an insertion (T) at nt 111010 leading to a stop codon at 111008. There is a new start at 110923 and the frame ends together with the main of the other strains; Strain 7 has a deletion (C) at nt 110845 and a resulting stop codon at 110823. A new start at 110824 and end in frame together with the main of the other strains.
84	Yes	
85	Yes	
86	Partially	A single ORF is present only in strains 1, 2 and 7. Strains 3,4,5 and 6 shows multiple ORFs within the main frame. More specifically, strains 3 has 6 sub ORFs; strain 4 has 2 sub ORFs, strain 5 has 5 and strain 6 has 12. This is all likely associated with frameshifting.
86a	No	Missing in 4
86b	No	Missing in 2, 3,5,6
86c	Partially	Strains 3, 5 and 6 have a mutation (A to C) at 119864 suppressing the start codon and shortening the frame 5' of 36 nt.
86d	Partially	Strain 2 has a mutation (C to T) creating a start codon at 124235 extending the ORF 36 nt at 5'. Strain 5, 6 and 7 have a mutation (G to T) leading to a stop codon at 123817, shortening the frame of 12 nt.
86e	Partially	Strain 5 has an insertion (T), which change the frame and result in a fusion with one of the sub-ORFs present and becomes 300nt longer; Strains 3, 5 and 6 have a mutation (G to C) at nt124687 leading to an early start codon there and an extra 3 nt long frame 5'.
86f	Yes	
87	Partially	Strain 7 has a deletion (T) at nt 129471 leading to a stop codon in frame at 129447 and a restart at 129361 to end together with the other main frames
88	Partially	Strain 3 has a deletion (T) at nt 129743, suppressing the stop codon in frame at nt 129737 and extending the frame up to 129570; Strain 7 has a deletion (T) at 130798 leading to a stop codon at 130686 and a new start at 130509 ending with the main frames.
88a	Partially	Strains 3, 5 and 6 have a mutation (C to G) at nt 130231, leading to an early start codon in that position, expanding the ORF 5' of 105 nt; Strain 7 has a deletion (A) at nt 130532 leading to a stop codon in frame at 130566, shortening the ORF at 3' of 38nt.
88b	Yes	

89	Partially	Strain 7 has a deletion (T) at nt 134646 leading to a truncation of the ORF on the 5' end of 202 nt, having a start codon at nt 134497. Strains 4, 5 and 6 have an insertion (G) at nt 134687 with a change of frame and a shortening of 7 nt 5' of the frame.
90	Partially	Strain 4, 5, 6 have an insertion (G) at nt 134687 leading to a shortening of the ORF 5' of 134nt
91	Yes	
92	Yes	
93	Yes	
94	Yes	
95	Yes	
96	Yes	
97	Partially	Strain 6 has an insertion (C) at nt 144394 leading to a stop in frame at 144298. A new ORF starts at 144321 and thanks to an insertion (C) at 141474, stops at 141405, 238 nt shorter 3'.
97a	Yes	
98	Yes	
98a	Yes	
99	Partially	Strain 6 has a deletion (A) at nt 146357 leading to a stop codon in frame at 146429. Restarts at 146343 and ends with the main frames.
100	Yes	
101	Yes	
102	Yes	
103	Partially	Strains 3, 5, 6 and 7 have a mutation (T to A) at nt 149216 that insert a stop codon at 149215, shortening the frame of 18 nt.
104	Partially	Strain 5 and 6 have an insertion (C) at 149582 resulting in a stop codon in frame at 149642. They both restart at 149553, however, while 5 runs to the end of the main frame, 6 has a deletion (C) at 149824 leading to a stop codon in frame at 149842. It then restart at 149925 and ends with the main frames.
105	Yes	
106	Partially	Strains 3, 5 and 6 have a mutation (C to T) creating a stop codon at 151919. They all have a new frame starting at 152049 finishing with the main frames of all but strain 6, which together with strain 7 has a double deletion (GT) at 153947-8 creating a stop codon in frame at 153965 and shortening the frame of 7 nt.
107	Yes	
108	Partially	Strain 7 has a deletion (T) at 155239 leading to a stop codon in frame at 155283 shortening the frame of 35nt.
109	Partially	Strain 7 has a deletion (G) at 155707 leading to a stop codon at 155713. A second frame starts at 155555, the same start codon of ORF109a of the other strains and ends with the main frame of ORF109 of all the other strains. This could be a frameshift.
109a	Partially	In strain 7, this ORF is embedded in the second ORF of 109. Likely a frameshift.
110	Partially	Strain 6 has a double deletion (AA) at 157020-1 leading to a stop codon at 157031. A new frame starts at 157545 and ends with the main frames.
110a	Yes	
111	Partially	Strain 7 has a deletion (T) at 159127, leading to a stop codon in frame at 159131. A new frame starts at 159238 and ends together with the main frame of the other strains.

112	Yes	
113	Partially	Strain 7 has a deletion (A) at nt 161615 leading to a stood codon in frame at 161763 shortening the frame of 35 nt
114	Yes	Strain 4 and 7 have a deletion (G at 161891, leading to an early start codon at nt 161822 and extending the frame of 61nt, still in frame with the other strains.
115	Yes	
116	Partially	Strain 3 has an insertion (C) leading to a frame shift and suppression of stop codon at 167793 and the fusion of the frame with ORF117. Strain 7 has only one of the two deletions present around a polyA motif at nt 163780 with suppression of the stop at 163793 allowing an extension of the frame of 35nt.
117	Partially	The fused 166 and 117 ORF of strain3 and the frames of strain 5 and 6 have a deletion (A/T) at nt 164830 resulting in a stop at 164881 and a restart at 165018 with the end with the main frames of the other strains.
118	Partially	Strain 2 has a deletion (166915-31) leading to a stop at 166911. Restarts at 166820 and ends with the other main frames. Strain 3 has a deletion (166802-9) resulting in a stop at 166796. A second ORF starts from here and an insertion (T) at 165921 cause a stop at 165916. A new frame starts at 165857 and ends with the other main frames. Strain 5 has a deletion (166913-4 and 917-9) causing a stop at 166911. A new ORF starts at 166933 and it has two deletions, one (TT) at 166912-3 and the other (TAG) at 166916-9. A new ORF starts at 166444 and has an insertion (G) at 166427 and a deletion (A) at 166189, it stops at 166085. Last ORF starts at 166062 and ends with the other main frames. Strain 6 has 3 nt deletion at 166916-and an insertion (G) at 166427 resulting in a stop codon at 166395. A new frame starts at 166444 and has an insertion (G) at 166427 and a deletion (A) at 166189 and resulting in a stop at 166085. The last frame starts at 165857 and ends with the other main frames.
118a	No	In 2,3,5,6,7 is missing
119	Partially	Strain 7 has a deletion (T) at 169945 leading to a stop in frame at 168952, shortening the frame of 122 nt. A T to A mutation in strain 2 at 169063 cause a stop codon in that position and a shortening of the ORF of 12 nt.
119a	Partially	One C to G mutation in strains 3, 5 6 and 7 at nt 168725 and a G to A in strain 2 and 4 at nt 168695 determines the suppression of a start codon at 168727 and a shortening of the frame of 75nt.
120	Partially	Strains 3, 5, 6 and 7 have a mutation (G to A) at nt 169102 with a shortening of 9 nt 5' end.
121	Yes	
122	Yes	
123	Partially	Strain 5 has a deletion (C) at 172665 resulting in a stop codon at 172648; Strain 6 has a deletion (A) at 173211 causing a stop at 173176. A new ORF starts at 173165 and ends at 172648. Strain 3 has a deletion (172571-6) extending the ORF of 24nt.
124	Partially	Strains 2,5 and 6 have a deletion (174317-20) associated with a stop codon at 174325; Strain 3 has an insertion (G) at 174279 leading to a stop at 174308.
125	Yes	

126	Partially	Strain 7 has a deletion (T) at 176082 resulting in a late start codon at 176063 and a shortening of the frame of 125nt.
127	No	In 5 and 6 is missing
128	Yes	
129	Partially	Strains 3 and 6 have an insertion (T) at 178836 suppressing the stop codon at 178864 and extending the ORF of 35nt.
130	Yes	
131	Partially	Strains 3 and 6 have a deletion (G) at 180538 determining a late start codon (180538) shortening the ORF 5' of 260nt.
131a	Yes	
132	Partially	Strain 3 has a frame 1 nt shorter because of an insertion (AA) at 1818847-8 causing an early stop codon. Similarly, strains 5 and 6 have an insertion (A) at 181848 causing as well an early stop.
133	Partially	Strains 5 and 6 have an insertion (TA) at 182260-1 associated with a stop codon in frame at 182276. A new frame starts at 182428 and ends with the main frames.
134	Partially	Strain 3 and 6 have an insertion (C) at 183943 leading to a stop codon at 183931, shortening the frame of 568 nt 3'. Strain 5 has an insertion at (T) 184316 resulting in a late start codon at 184234, resulting in a shortening of 233nt, 5'.
135	Partially	Strain 4 has a deletion (T) at 184316 leading to a late start at 184467 with shortening of the frame of 233nt 5'. Strain 7 has a deletion (T) at 184571 leading to a stop at 184693 and a restart at 184953 and ending with the main frames.
136	Yes	
137	Yes	
138	Partially	Strains 3, 5 and 6 have a deletion (GGAA) at nt 189322-5 resulting in a stop codon at 189342. A new frame starts at nt 189344 and ends with main frame. Strain 4 has an insertion (A) at 187893 with a stop codon at 187915. A new frame starts at 187994 and stops with main frame. Strain 7 has a deletion (A) at 189458 and stops at 189531. A new frame starts at 189451 and ends with the main frame.
139	Partially	Strain 3 has a deletion (A) at 190458 causing a late start 5' at 190610 shortening the frame of 284 nt. The frame has a deletion (A) at 191358 causing a stop at 191437 with a restart at 191412 and ending with main frame. Strain 5 has a deletion (AA) at 191358-9 causing a stop at 191384; it restarts at 191305 to end with main frames. Strain 6 has a late start at 190610 because of a deletion (A) at 190458; The main frame has an insertion (T) at 191170 resulting in a stop at 191183; It restarts at 191266 ending in frame with mains.
140	Yes	Strain 3 has a deletion (A) at 192881 resulting in a stop at 192892 with 25 nt shortening of the frame 3'.
141	Yes	
142	Yes	Strains 1,2,4 and 7 have a 6nt shorter frame 5' because of a mutation (A to G) at nt 194002 in strains 1, 4 and 7 or a mutation (T to A) at nt 194003 in strain 2
143	Partially	Strain 6 has an insertion (A) resulting in a stop codon at 194837, shortening the frame 3' of 178nt.
144	Yes	
145	Yes	
145a	Yes	

146	Yes	
147	Partially	Strain 6 has two deletions, one (ATATAT) at nt 196285 and another (T/A) at nt 196318 resulting in an early start at 196225 and a 93nt longer frame 5'.
148	Partially	Strain 6 has a mutation (T to G) at 196660 suppressing a start codon at 196659 and resulting in a 237nt shorter frame 5'.
149	Partially	Strain 3 has an insertion (T) at 197308 causing a late start at 197329 and a 67bp shorter frame 5'; Strain 4 has a deletion (T/G) at 197541 causing a stop at 197657 and a 362nt shorter frame 3'; Strain 7 has a deletion (G) at 197507 with a stop at 197657 and a 362nt shorter frame 3'.
150	Partially	Strain 3 and 6 have an insertion (A) at nt 198856 resulting in the fusion between ORF150 and 151
151	Partially	Strain 4 has a deletion (CT) at nt 198957-6 associated with a late start at 198896 with a frame 31bp shorter 5'. Strain 5 has a mutation (A to T) creating a stop codon at 199213. A new start is at 199504 with an end together with the main frames; Strains 3 and 6 are fused with ORF150.
152	Yes	
153	Partially	Strain 3 has a mutation (G to T) at 200888 creating a stop codon there; a new frame starts at 200975 and contains 5 deletions located at 201419-21, 201710 and 202114 resulting in a stop at 202192 and a new start at 200191 ending in frame; Strain 4 has 2 deletions at 201710-1 with a stop at 202192 and a restart at 2020191 and end in frame; Strain 5 has 5 deletions at 201419-21, 201710 and 202114 resulting in a stop at 202192 and a new start at 200191 ending in frame; Strain 6 has 4 deletions at 201419-21 and 202114 resulting in a stop at 201728, restart at 2020191 and end in frame
153A	No	Not present in strain 6
154	Yes	
155	Yes	
156	Partially	Strain 3 and 6 have a deletion (T) at 207884 leading to a late start codon at 207760 resulting in a 281nt shortened ORF 5'. Strain 7 has a deletion (G) at 207472 leading to a stop at 207465 with a restart at 207425 and ending in frame.
156a	No	Not present in 7; Extraframe in 3 and 6
157	No	Not present in 5 and 6.

By extracting all the nucleotide changes resulting in significant ORFs changes, we observed that the highest number of nucleotide changes in comparison with the type strain FO_2015 was observed in strain W_18_2160, within Switzerland with 63 relevant nucleotide changes, which was lower only than one of the two Hungarian strains (RaHV3_HU_Sw). Differently, the isolate showing the least number of significant nucleotide changes in comparison with the reference strain FO_2015, was strain W17_4184, another Swiss strain. Interestingly, isolates W17_4184, RaHV3_D, BS_1994 are those showing the least variations in comparison with the reference strain (FO_2015). Curiously, among the most variable isolate, is the Swiss strain W18/2160, which shows a level of variability consistent with the Hungarian strains. Overall, the most common nucleotide changes were deletions, whereas insertion and mutations were comparable. All this information are summarized in Table 4 below.

Table 4. Nucleotide changes resulting in frame variation across RaHV3 isolates

Strain	Deletions	Insertions	Mutations	Fully conserved genes (main ORF)	Total changes
FO_2015	1	1		96	2
W17_4184	11	2	6	80	19
W18_2160	28	15	20	46	63
RaHV3_D	21	9	2	67	32
RaHV3_HU_Sk	26	14	17	52	57
RaHV3_HU_Sw	36	31	18	29	85
BS_1994	26	1	9	64	36
Total	151	73	72		

By examining the moderately conserved ORFs, we observed that the majority of the changes were present in only one isolate per examined frame (36). Curiously, the changes occurring in the frames of 2, 3 or 4 isolates per examined frame were relatively consistent ranging from 18 to 21. ORF changes occurring in almost all the isolates, per specific ORV examined. Differently, ORFs account for frames variations in most of the isolates (affecting 5 or 6 of the seven strains) were substantially rarer (1 and 3, respectively). This information is summarized in Table 5 below.

Table 5. Frame variations across isolates in moderately conserved genes.

N.	ORF	Not conserved frames	Remarks*		
1	3	1/7			
2	4	1/7			
3	5	1/7			
4	7	3/7			
5	8	3/7			
6	9	4/7	*		
7	11	2/7			
8	12	1/7			
9	14	4/7	*		
10	15	1/7			
11	16	3/7			
12	17	2/7			
13	20	6/7	****		
14	21	1/7			
15	23	4/7	*		
16	24	3/7			
17	27	1/7			
18	28	3/7			
19	29	2/7			
20	30	4/7	*		
21	33	1/7			
22	34	6/7	****		
23	37	2/7			
24	38	2/7			
25	40	3/7			
26	41	1/7			

27	42	3/7			
28	43	3/7			
29	45	1/7			
30	46	4/7	*length		
31	50	2/7			
32	53	2/7			
33	54	1/7			
34	55	2/7			
35	56	1/7			
36	57	1/7			
37	58	1/7			
38	61	4/7	*		
39	62	3/7			
40	63	4/7	*		
41	66	2/7			
42	69	1/7			
43	72	1/7			
44	73	1/7			
45	75	1/7			
46	76	4/7	*		
47	77	2/7			
48	78	3/7			
49	80	1/7			
50	82	1/7			
51	83	2/7			
52	86	4/7	***		
53	86c	3/7			
54	86d	4/7	*		
55	86e	2/7			
56	87	1/7			
57	88	2/7			
58	88a	4/7	*		
59	89	4/7	*length		
60	90	3/7			
61	97	1/7			
62	99	1/7			
63	103	4/7	*length		
64	104	2/7			
65	106	3/7			
66	108	1/7			
67	109	1/7			
68	109a	1/7			
69	110	1/7			
70	111	1/7			
71	113	1/7			
72	114	2/7			
73	116	1/7			
74	117	3/7			
75	118	4/7	**		
76	119	1/7			
77	119a	6/7	****length		

78	120	4/7	*length		
79	123	3/7			
80	124	4/7	*length		
81	126	1/7			
82	129	2/7			
83	131	2/7			
84	132	3/7			
85	133	2/7			
86	134	3/7			
87	135	2/7			
88	138	5/7	****		
9	139	3/7			
90	140	1/7			
91	142	3/7			
92	143	1/7			
93	147	1/7			
94	148	1/7			
95	149	3/7			
96	151	4/7	**		
97	153	4/7	**		
98	156	3/7			
Total					
1/7	2/7	3/7	4/7	5/7	6/7
36	19	21	18	1	3

*The number of asterisks is proportional to the proportion of strains affected

In tables 6a-c are summarized all the identities of the variably or fully conserved ORFs on the basis of their pairwise identity and frame conservation.

Table 6a. *Ranid herpesvirus 3* Open reading frames (ORF) features. Most conserved genes (pairwise identity). In yellow the ORFs with fully conserved frame.

ORF	Viral Homologue	Protein	Predicted function/ description	Similarity or confidence	Position	Length (nt)	Frame
4	None	Unknown	Cap-Gly domain	53.2%[b] (7% coverage)	4,504-6,099	1,596	-
23	None	Unknown	TNF receptor superfamily	99.5%[b] (57% coverage)	33,212-33,550	339	-
34A	None	Unknown	AMMERC1-like superfamily	35.6%[b] (39% coverage)	48,589-49,248	660	+
49	None	Unknown	SAM domain-like	53.2%[b] (9% coverage)	67,175-68,371	1,197	+
51	None	Unknown	Unknown	N/A	68,449-69,126	678	-
52	RHV1	ORF85	ORF85	50%[a] 44/88	69,120-70,136	1,017	-
60	None	Unknown	Mss4-like superfamily motif	48.8%[b] (11% coverage)	80,337-81,062	726	-
62A	None	Unknown	Alpha-alpha superhelix	44%[b] (36% coverage)	82,217-82,489	273	-

66	RHV1	ORF83	Multiple membrane spanning protein	50%[a] 111/219	85,667-86,419	753	+
67	None	Unknown	Membrane protein	49.3%[b] (20% coverage)	86,104-86,397	294	-
68	RHV1	ORF82 (not conf)	Multiple membrane spanning protein	46%[a] 54/116	86,513-87,247	735	-
70	None	Unknown	Putative isomerase	29.7%[b] (19% coverage)	88,028-88,555	528	+
71	RHV1	ORF76 Not conf.	Capsid triplex protein	40%[a] 46/113	88,552-89,514	963	-
72	RHV1	ORF75	Sim. to IcHV1 ORF54	40%[a] 230/571	89,522-91,216	1,695	+
73	None	Unknown	Putative DNA binding protein	27.1%[b] (8% coverage)	91,143-92,411	1,269	+
74	RHV1	ORF73	Sim. to IcHV1 ORF56	47%[a] 581/1,218	92,405-96,094	3,690	-
74A	None	Unknown	Unknown	N/A	95,880-96,149	270	+
75	RHV1	ORF72	DNA polymerase	47%[a] 677/1,431	96,082-100,182	4,101	+
80	RHV1	ORF63	Capsid protease and scaffolding protein	47%[a] 221/464	107,208-109,142	1,935	-
81	RHV2	ORF87	ORF87	40%[a] 57/141	108,976-109,509	534	-
83	RHV1	ORF59	ORF59	51%[a] 187/366	110,091-111,263	1,173	-
84	None	Unknown	Putative isomerase	53.7%[b] (10% coverage)	111,246-113,828	2,583	+
88B	None	Unknown	Unknown	N/A	132,392-132,676	285	+
89	RHV1	ORF53	ORF53	45%[a] 61/135	133,625-134,290	666	-
92	None	Unknown	Thioredoxin-like	42.7%[b] (23% coverage)	136,809-137,375	567	-
93	RHV1	ORF49	Sim to IcHV1 ORF34	56%[a] 253/446	137,290-138,633	1,344	-
95	None	Unknown	Centromere-like motif	19.3%[b] (5% coverage)	139,057-139,677	621	+
97A	None	Unknown	Similar to capsid protein VP1 of food and mouth disease virus	10%[b] (42% coverage)	141,894-142,220	327	+
102	RHV1	ORF42	Exon of terminase	52%[a] 138/264	147,960-148,739	780	+
105	None	Unknown	DNA binding protein	47%[b] (7% coverage)	149,884-150,423	540	-
107	None	Unknown	C-lectin type protein	97.1%[b] (44% coverage)	153,558-154,190	633	+
108	None	Unknown	C-lectin type protein	96.6%[b] (50% coverage)	154,285-154,839	555	+
109	None	Unknown	C-lectin type protein	74.5%[b] (24% coverage)	154,962-156,122	1,161	+
109A	None	Unknown	Unknown	N/A	155,072-155,347	276	+
110	None	Unknown	C-type lectin protein	74.1%[b] (Coverage 25%)	156,195-157,379	1,185	+
110A	None	Unknown	SacY-like RNA binding domain	39.7%[b] (27% coverage)	157,246-157,518	273	+
111	None	Unknown	C-lectin type protein	76.4%[b] (12% coverage)	157,791-159,236	1,446	+

112	None	Unknown	C-lectin type protein	82.4%[b] (32% coverage)	159,313-160,191	879	+
113	None	Unknown	C-lectin protein (c-type mannose receptor-like)	100%[b] (98% coverage)	160,770-161,303	534	+
115	None	Unknown	C-lectin type protein related to NK receptor ly49l4	99.5%[b] (52% coverage)	161,813-162,268	456	+
119	None	P450 superfamily	P450 superfamily	43%[a] 199/455	167,211-168,551	1,341	+
121	None	Unknown	C-lectin like protein (c-type mannose receptor-like)	100%[b] (71% coverage)	169,501-170,475	975	+
125	None	Unknown	Protein binding motif	33.4%[b] (2% coverage)	173,863-175,302	1,440	+
126	None	Unknown	C-type lectin protein	97.7%[b] (64% coverage)	175,381-175,878	498	+
127	None	Unknown	C-type lectin protein	99.9%[b] (56% coverage)	175,924-176,490	567	+
128	None	Unknown	Unknown	N/A	176,736-177,272	537	+
131A	None	Unknown	Toxin-like motif	45.8%[b] (12% coverage)	180,373-180,675	303	-
136	None	Unknown	ROP-like motif	41.6%[b] (5% coverage)	185,121-186,212	1,092	+
137	None	Unknown	Putative glycoprotein	80%[b] (54% coverage)	186,256-186,705	450	+
144	None	Unknown	Beta-beta-alpha zinc fingers superfamily	58.3%[b] (17% coverage)	194,412-194,795	384	+
154	None	Unknown	Toxin-like protein	22%[b] (15% coverage)	202,339-202,857	519	+
157	None	Unknown	Immunoglobulin-like motif	23.3%[b] (10% coverage)	207,314-207,874	561	+

Table 6b. *Ranid herpesvirus 3* Open reading frames (ORF) features. Moderately conserved genes-(pairwise identity). In yellow the ORFs with fully conserved frame.

ORF	Viral Homologue	Protein	Predicted function/ description	Similarity or confidence	Position	Length (nt)	Frame
13	None	Unknown	Pili subunit/adhesion motif	88.4%[b] (23% coverage)	18,460-19,341	882	-
19	None	Unknown	Ferrodoxin-like motif	21.4%[b] (25% coverage)	27,727-28,008	282	-
22	None	Unknown	BCL-2 superfamily inhibitors of programmed cell death	71%[b] (56% coverage)	32,899-33,243	345	+
25	None	Unknown	Suppressor of cytokine signaling 5 motif	51.1%[b] (2% coverage)	35,550-36,818	1,269	+

26	None	Unknown	probable ubiquitin fold modifier conjugating enzyme motif	39.5%[b] (9% coverage)	36,146-36,883	738	+
31	None	Unknown	Hydrolase-like motif	48%[b] (5% coverage)	46,175-47,413	1,239	+
32	None	Unknown	SipA N-terminal domain-like superfamily	35.1%[b] (27% coverage)	47,445-48,224	780	-
39	None	Unknown	Propionate kinase-like	93.4[b] (43% coverage)	54,188-55,285	1,098	-
43A	None	Unknown	Surfactant-like protein	43.9%[b] (43% coverage)	61,651-61,929	279	+
47	None	Unknown	Lyase-like motif	39.7%[b] (11% coverage)	64,545-65,102	558	+
48	None	Unknown	Nucleoside triphosphate hydrolases/Sulfotransfease family	100%[b] (coverage 61%)	65,965-67,017	1,053	+
64	RHV1	ORF95	Capsid triplex protein sub.2	44%[a] 151/336	83,721-84,728	1,008	+
64A	None	Unknown	Spectrin-like repeat	52.8%[b] (28% coverage)	84,037-84,318	282	-
65	RHV2 (RHV1)	ORF118 (ORF84)	ORF118	46%[a] 124/266	84,736-85,668	933	+
76A	None	Unknown	Putative histidine-rich glycoprotein	18.8%[b] (18% coverage)	100,281-100,592	312	+
79	None	Unknown	Restriction endonuclease-like motif	50.2%[b] (4% coverage)	104,701-107,124	2,424	-
85	RHV1	ORF57	ORF57	40%[a] 113/280	113,802-114,656	855	-
86F	None	Unknown	Snake toxin-like	17.4%[b] (25% coverage)	126,243-126,551	309	-
91	RHV2	ORF77	ORF77	48%[a] 12/25	136,393-136,842	450	-
94	RHV1	ORF48	ORF48	41%[a] 54/131	138,459-139,193	735	-
98	RHV1	ORF45	Putative membrane glycoprotein (or signal peptide?)	48%[a] 164/340	144,399-145,607	1,209	-
98A	None	Unknown	Similar to e62 human papillomavirus oncoprotein	32.2%[b] (14% coverage)	144,638-145,024	387	+
100	RHV1	ORF42	ATPase subunit of terminase	63%[a] 106/168	146,778-147,302	525	+
101	RHV1	ORF43	ORF43	44%[a] 118/265	146,975-147,934	960	-
115	None	Unknown	C-lectin type protein related to NK receptor ly49l4	99.5%[b] (52% coverage)	161,813-162,268	456	+
122	None	Unknown	C-lectin like protein	83.2%[a] (14% coverage)	170,561-171,487	927	+
130	None	Unknown	(c-type mannose receptor-like (C-type lectin protein)	97.6%[b] (49% coverage)	178,385-179,647	1,263	+
131A	None	Unknown	Toxin-like motif	45.8%[b] (12% coverage)	180,373-180,675	303	-
145	None	Unknown	Transcription associated protein	30.9%[b] (7% coverage)	194,684-195,142	459	+

145A	None	Unknown	SAM domain-like	26.1%[b] (24% coverage)	194,806-195,087	282	-
146	None	Unknown	Putative metal binding protein	25.8%[b] (coverage 31%)	195,185-195,637	453	+
152	None	Unknown	Lyase-like	38.6%[b] (39% coverage)	199,193-199,762	570	+
155	None	Unknown	Heme-binding protein-like	30.1%[b] (11% coverage)	202,911-204,221	1,311	+

Table 6c. *Ranid herpesvirus 3* Open reading frames (ORF) features. Less conserved genes (pairwise identity). In yellow the ORFs with fully conserved frame.

ORF	Viral Homologue	Protein	Predicted function/ description	Similarity or confidence	Position	Length (nt)	Frame
1	None	Unknown	Predicted TNF receptor associated factor	11.3% [b] (27% coverage)	432-734	303	-
2	None	Unknown	Predictive putative 4-hydroxythreonine-4-phosphate2 dehydrogenase motifs from salmonella typhimurium	53.3%[b] (14% coverage)	628-1,491	288	-
3	None	Unknown	Isomerase motif	48.3%[b] (9% coverage)	2,020-3,774	1,755	-
7A	None	Unknown	PMP inhibitors motif	51.6%[b] (4% coverage)	9,699-9,977	278	-
35	None	Unknown	DNA biding protein motif	40.8%b (2% coverage)	49,954-51,192	1,239	-
36	None	Unknown	Forkhead box protein m1 motif/transcription/DNA	59.8%[b] (42% coverage)	51,286-51,750	465	-
44	None	Unknown	Unknown	N/A	62,055-63,065	1,011	-
86A	None	Unknown	Cyclophillin-like	40.3%[b] (21% coverage)	116,973-117,314	342	-
86B	None	Unknown	Putative metal transport protein	26.1%[b] (20% coverage)	118,077-118,451	375	-
86D	None	Unknown	Antimicrobial peptide motif	15%[b] (4% coverage)	123,434-123,787	354	-
86E	None	Unknown	Putative structural protein	21.7%[b] (9% coverage)	123,909-124,271	362	-
96	RHV1	ORF47	Membrane glycoprotein	37%[a] 98/264	139,656-140,699	1,044	+
141	None	Unknown	DNA/RNA-binding 3-helical bundle motif	52.5%[b] (35% coverage)	192,353-192,787	435	+
153	None	Unknown	Putative hydrolase	100%[b] (38% coverage)	199,787-202,414	2,628	+
153A	None	Unknown	Putative transferase	19.2%[b] (11% coverage)	201,892-202,302	411	-

157	None	Unknown	Immunoglobulin -like motif	23.3%[b] (10% coverage)	207,314-207,874	561	+

Among the most conserved genes are those encoding for structural elements (capsid), involved in DNA metabolism (terminase-ORF102) and in immunomodulation (TNFr-ORF23, but with not fully conserved frame across isolates).

Additional insertions and deletions

A series of insertion and deletions was observed either within ORFs or in intergenic regions in one or more isolates. Below is a list of these changes:

- ORF7, insertion (27nt) in isolates 3, 5, 6

-ORF intergene 8-9, insertion (24nt) in isolate 3

-ORF15, deletion (15nt) in isolate 3, 5, 6

-ORF intergene 24-25, insertion (9nt) in isolate 3 and partial in 5 (2) and 6 (5)

-ORF34, insertion (5nt) in isolate 6 and 7

-ORF intergene 44-45, insertion (8nt) in isolate 5 and partial (3) in 3

-ORF77 insertion (13nt) in isolate 6 and partial (6) in 3

-ORF82 insertion (6nt) in isolate 6 and partial (3) in 3 and 5

-ORF86 insertion (31nt) in isolate 2

-ORF96 insertion (48nt) in isolates 3, 5 and 6

-ORF intergene 130-131, deletion (10nt) in isolates 3, 4, 5, 6

-ORF intergene 133-134, deletion (49nt) in isolates 2 and 4, partial (2 to 6) in all the others

-ORF134 insertion (6nt) in isolates 2 and 4

-ORF139 insertion (40nt) in isolate 3

-ORF153 insertion (24nt) in isolates 3 and 6

Repeated motifs

Additionally to the analyses described above, we investigated the presence of short repeated motifs across the 7 isolates examined. The search settings included decamers repeated at least 5 times. The analysis was carried out using the online software Gramene SSrtool (https://archive.gramene.org/db/markers/ssrtool). Here below in figures 2a-g are reported the results of this analysis.

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-1	ta	5	20490	20499	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-2	ct	6	52818	52829	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-3	tg	5	80439	80448	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-4	ta	5	135116	135125	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-5	at	5	153928	153937	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-6	at	5	175820	175829	208550

Figure 2a. Short repeated motifs identified in strain FO1_2015

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-1	ta	5	20493	20502	208553
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-2	ct	6	52821	52832	208553
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-3	tg	5	80442	80451	208553
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-4	ta	5	135119	135128	208553
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-5	at	6	153929	153940	208553
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-6	at	7	175819	175832	208553
GCCTTCATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-7	tg	5	191964	191973	208553

Figure 2b. Short repeated motifs identified in strain W17_4184

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-1	ct	5	694	703	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-2	ta	5	20490	20499	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-3	cg	5	37239	37248	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-4	ct	6	52818	52829	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-5	ta	5	60781	60790	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-6	tg	5	80439	80448	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-7	gt	6	153885	153896	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-8	at	5	153928	153937	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-9	tg	5	191961	191970	208550

Figure 2c. Short repeated motifs identified in strain W18_2160

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-1	ta	5	20493	20502	208553
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-2	ct	7	52819	52832	208553
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-3	at	5	153931	153940	208553
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-4	at	5	175823	175832	208553
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAA-5	at	6	196226	196237	208553

Figure 2d. Short repeated motifs identified in strain RaHV3_D

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-1	ct	5	694	703	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-2	ta	5	20490	20499	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-3	cg	5	37239	37248	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-4	ct	6	52818	52829	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-5	ta	5	60781	60790	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-6	tg	5	80439	80448	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-7	gt	6	153885	153896	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-8	at	6	153926	153937	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-9	tg	5	191961	191970	208550

Figure 2e. Short repeated motifs identified in strain RaHV3_HU_Sk

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-1	ct	5	694	703	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-2	ta	5	20490	20499	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-3	cg	5	37239	37248	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-4	ta	5	50124	50133	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-5	ct	6	52818	52829	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-6	ta	5	60781	60790	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-7	gt	5	153887	153896	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-8	at	5	153928	153937	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-9	tg	5	191961	191970	208550

Figure 2f. Short repeated motifs identified in strain RaHV3_HU_Sw

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-1	ta	5	20490	20499	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-2	cg	5	37239	37248	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-3	ct	6	52818	52829	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-4	tg	5	80439	80448	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-5	gt	5	153887	153896	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-6	at	5	153928	153937	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-7	at	5	175820	175829	208550
GCCTTGATACTAGCTTACGCAACACATTTACTCACATATGATTAAACCGGTTTAATAATTTT-8	tg	5	191961	191970	208550

Figure 2g. Short repeated motifs identified in strain BS_1994

Phylogenetic analysis

Phylogenetic analysis was carried out using either the whole genome of the isolates using the specific Genious software package application or specific genes, which would provide a similar tree topology to that obtained with the full genome and accordingly, ideal to be used as proxy of the whole genome itself using the software package Mega 7 (https://www.megasoftware.net/). Phylogenetic analysis unambiguously identified at least two separate clusters of Swiss RaHV3 strains. Additionally, the Italian isolate appears to cluster separately, whereas the two Hungarian isolates appear to be more closely related to one of the Swiss strains as shown in Figure 3 below.

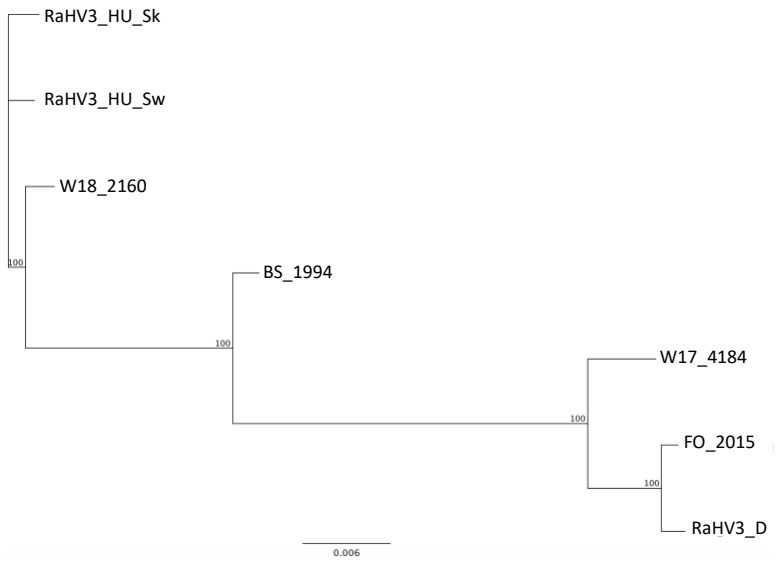

Figure 3. Whole genome-Neighbor-Joining consensus tree Nucleotide alignment of the investigated RaHV3 isolates. Bootstraps value are indicated at the branching of the nodes.

The most represented lineage is arbitrarily defined as lineage 1, whereas, the less common one, is arbitrarily defined as lineage 2.
Similar tree topology has been obtained using the entire or partial sequences of several genes. Below is the tree obtained with the sequences of the gene encoded by ORF96, a poorly conserved gene, ideal to evaluate fine strain variability.

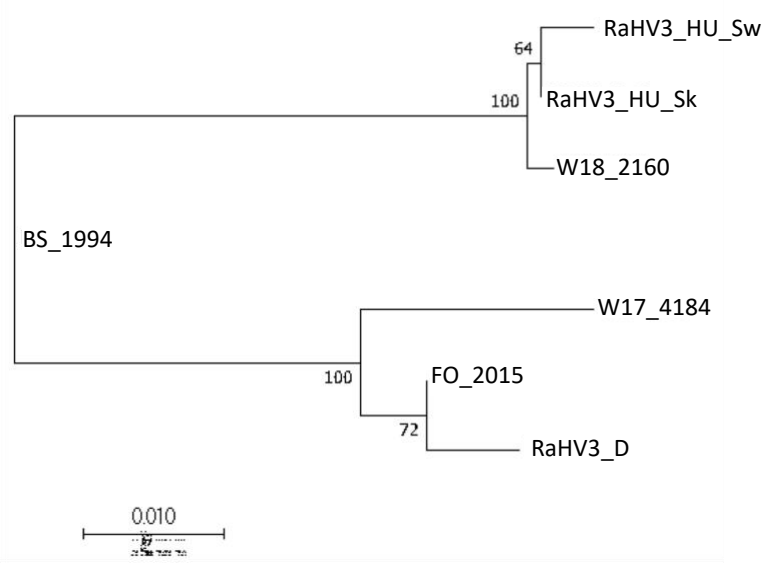

Figure 4. Maximum Likelyhood phylogentic tree based on the whole nucleotide sequence of RaHV3 ORF96. Bootstraps value are indicated at the branching of the nodes.

Finally, using the partial sequence of the same ORF 99 (266nt) it was possible to obtain a similar topology (Figure 5) making this region ideal for genotyping the circulating RaHV3 strains.

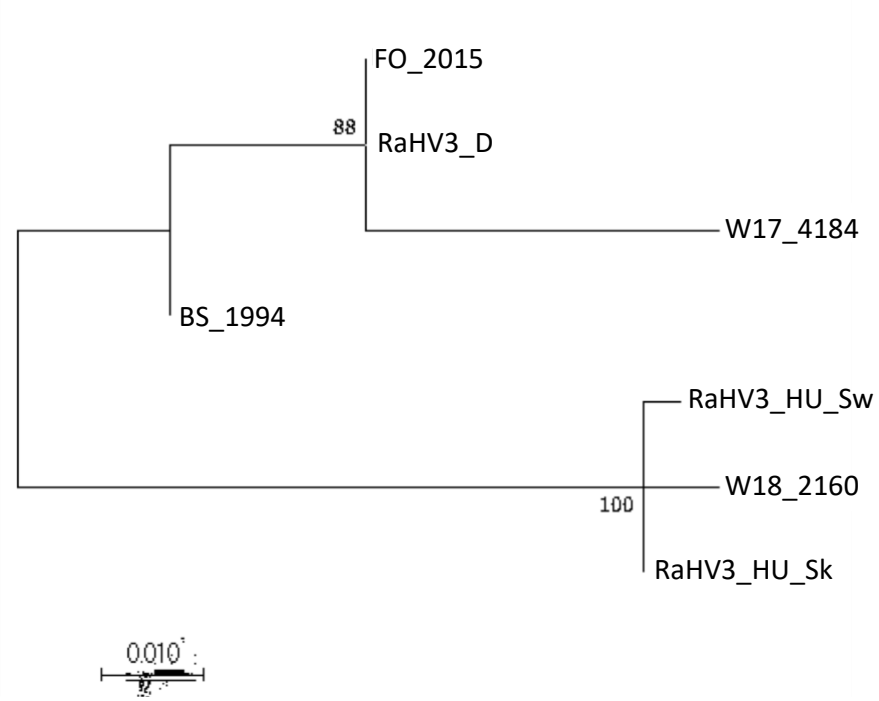

Figure 5. Maximum Likelyhood phylogentic tree based on the partial nucleotide sequence of RaHV3 ORF96. Bootstraps value are indicated at the branching of the nodes.

Phylogeography

The results obtained from the phylogenetic analysis were then used to determine if there was any association between specific isolates and geographic regions. Interestingly, the cluster determined by phylogenetic analysis suggest that in Switzerland are circulating at least two distinct lineages of RaHV3. One present in Canton Zurich, Neuchatel and Graubunden, whereas the second lineage has been detected only in Thurgau and interestingly, this lineage share similarities with isolates from Hungary. Thurgau is located at the Northwestern rim of the Swiss borders in close range with German and Austrian border, however, it is almost entirely surrounded by water on the border side. It is not clear if this location might have favored the spread into Switzerland, of a lineage, which appears to be present in Eastern Europe. Further investigation with larger sample sizes are needed to investigate the phylogeography of RaHV3 and draw more definitive conclusions concerning the distribution and movement of RaHV3 isolates across Europe.

Figure 6. Map of Switzerland showing the origin of the RaHV3 isolates. Yellow dots show the origin of RaHV3 lineage 1, whereas, the single red dot shows the origin of the isolate sharing similarities with the Hungarian strains (https://www.freeworldmaps.net/europe/switzerland/switzerland-physical-map.jpg).

Conclusions (RaHV3)

The analysis of the 7 RaHV3 isolates available have revealed that RaHV3 contains genes sharing between 86.6 to 100% pairwise identity. The genes can be grouped in fully conserved, partially conserved and not conserved. Fully conserved share overlapping ORFs and reading frames in all the strains; partially conserved genes share at least a portion of the specific frame in all the strains; not conserved genes lack the homologous ORF in one or more of the strains. Overall RaHV3 genome

contains 70 fully conserved genes; 98 partially conserved genes and 18 not conserved genes. Predominant conserved genes are within the group of genes with higher pairwise identity.

The majority of the partially conserved genes are characterized by changes in a single strain (36). Changes in up to two strains are present in 19 genes, in up to three strains are present in 21 genes and in up to 4 strains are detected in 18 genes. Changes in up to 5 strains are present in one gene only and two genes shows changes in the frame of up to 6 strains. The partially conserved genes are characterized by the accumulation of mutations (insertions, deletions, substitutions) responsible of frame changes (frameshift or/n and truncation). Of these, the majority were deletions (151) followed by almost equivalent insertions (73) and mutations (72). Strains 3 and 6 are those with more mutations detected. Among the most conserved genes are those encoding for structural elements (capsid), involved in DNA metabolism (terminase) and interestingy in immunomodulation (TNFr). Phylogenetic analysis revealed the presence of two distinct lineages, with a more diffuse one (lineage 1) including strains from Zurich, Neuchatel and Graubunden and a less diffuse one (lineage 2) including a strain from Thurgau. A reliable phylogenetic marker has been identified in the partial sequence of ORF96. This marker will allow the fine characterization of RaHV3 phylogeography in Switzerland and the potential association between the different lineages, population dynamics and disease occurrence.

Bufonid herpesvirus 1

Isolates

A total of 6 DNA samples obtained from common toad (*Bufo bufo*) were available (Table 7). Of these 5 were originally from Switzerland and 1 from Hungary. Similarly to what discusses concerning the RaHV3 isolates, the availability of an isolate from relatively distant country (Hungary) was considered important to assess if genomic variabilities was associated with a geographical component. All the isolates were obtained from common toads showing the classic clinical signs, consisting in proliferative brown skin patches. All samples were obtained from fresh tissue. Of them the reference strain, FO_2015T, was the pool of three samples collected from three individuals collected in from the same pond. Similarly, another isolate (ZH_2017) was obtained pooling samples from 8 individuals collected in the same pond.

Table 7. RaHV3 isolates

RaHV3 isolate ID	Host	Origin	Year	Genome length (nt)
FO_2015T	*Bufo bufo*	Switzerland_BL	2015	158,250
W14_1083	*Bufo bufo*	Switzerland_BL	2014	158,218
W16_9699	*Bufo bufo*	Switzerland_ZH	2016	158,246
ZH_2017	*Bufo bufo*	Switzerland_ZH	2017	158,256
W17_4268	*Bufo bufo*	Switzerland_Ne	2017	158,201
BfHV1_HU_Sw	*Bufo bufo*	Hungary	2018	158,216

Isolate sequencing

Next generation sequencing was carried out by a private Biotech Company (Fasteris SA, Geneva) using the Illumina platform (2x100 bp - 100 million clusters). All the sequences obtained were then mapped to the reference strain FO_2015T. The coverage obtained for the different isolates ranged from 96.68% (BfHV1_HU_Sw) to 100% (W14_1083 and W17_4268). A consensus sequence was obtained from each isolate by patching the missing regions using the reference strain FO_2015T. The genome sequence length of all the isolates were very close, ranging from 158,201 (W17_4268) to 158,256 (ZH_2017) (Table 1).

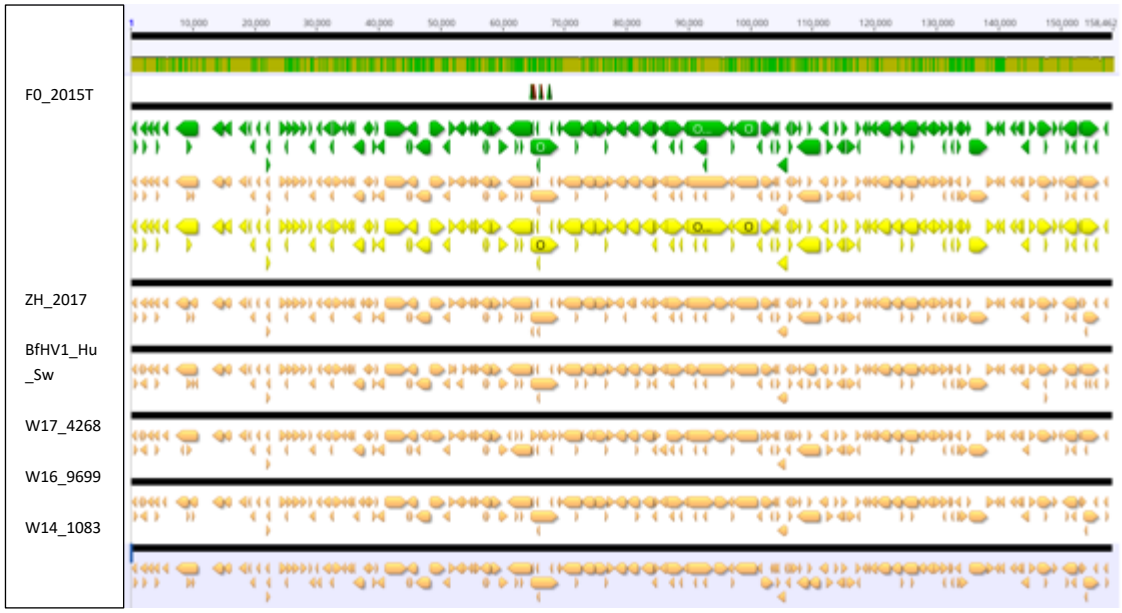

Figure 7. Alignment of the BfHV1 isolates with their ID (left), genome length ruler (top) and ORFs prediction. The ORFs of the reference strain (FO_2015T) are shown in yellow.

Isolate analysis

Similarly to what described for the RaHV3 isolates above, an overlapping analysis approach was undertaken also for the BfHV1 isolates. Accordingly, the obtained consensus sequences of each isolate were compared among each other using the software package Geneious prime (Geneious Prime® 2020.2.4; www.geneious.com). The parameters examined included pairwise nucleotide identity, identification of highly conserved, moderately conserved and poorly conserved genes across the available isolates, identification of all the nucleotide changes resulting in amino acid substitutions, open reading frame (ORF) modifications (insertion of stop codon, suppression of stop codon, insertion of start codon, suppression of start codon); large insertions and deletions. Figure 7 shows the aligned genomes with predicted ORFs.

Similarly, to what reported above concerning RaHV3 isolates analysis, ORF conservation was assessed according two distinct parameters. The first parameter was the pairwise identity. An arbitrary separation of three groups of genes (highly, moderately, and poorly conserved) was carried out. However, differently from RaHV3 and with the exception of 6 out of 152 predicted ORFs with a pairwise identity lower than 98%, all the other ORFs showed a relatively high pairwise identity (overall higher than for the RaHV3 isolates) and accordingly, the separation between highly, moderately and poorly conserved genes was arbitrarily determined using smaller intervals than those used with RaHV3. Accordingly, for BfHV1, poorly conserved ORFs were defined as those with pairwise identity ranging from 85.1 to 98.9%, moderately conserved genes had a pairwise identity between 99 and 99.8% and highly conserved ORFs were defined as those showing a pairwise identity ranging from 99.9 to 100% (Table 8). The second parameter considered was the ORF conservation across isolates (same reading frame). A highly conserved gene was considered that showing a fully conserved frame across all the isolates, whereas a moderately conserved gene showed a conserved ORF across all isolates, but the frame size might have included variations, such as frame fragmentation, shortening or extension. Finally, a poorly conserved gene showed the absence of the specific ORF in one or more of the isolates (Table 8). According to these parameters, the BfHV1 isolates contained 18 poorly conserved genes, 93 moderately conserved and 41 highly conserved ones. Overall, the RaHV3 ORFs accounted for 134 out of 152 with a pairwise conservation equal or higher than 99%. Poorly ORF conservation was higher (5.6%) among the group of genes with lower pairwise identities, and lower (2.2-2.4%) among those with moderate and high pairwise identity, as expected. On the contrary, ORF conservation was highest (85.4%) in the group with high pairwise identity and lowest (22.2%) in the group with low pairwise identity. Moderately conserved ORFs were predominant in the group of genes with lower pairwise identity (Table 8).

Table 8 below, shows the color-coded genes, identified as highly conserved (yellow), moderately conserved (green) and poorly conserved (blue). This grouping was carried out on the basis of the pairwise identity of each of the specific gene (shown as percentage value in parenthesis; higher percentage=higher conservation) and on the level of conservation of the specific ORFs. Specifically, for the ORF conservation, a highly conserved gene showed a fully conserved frames across all the isolates, whereas a moderately conserved gene showed a conserved ORF across all isolates including frame size variation. Differently a poorly conserved gene showed the absence of that specific ORF in one or more of the isolates.

Table 8. Pairwise identity RaHV3 genes and conserved open reading frames (ORFs)

ORF	Pairwise identity (%)	ORF	Pairwise identity (%)	ORF	Pairwise identity (%)
114	85.1	2	99	3a	99.9
110a	92.7	10	99	15a	99.9
113b	96.1	115a	99	20	99.9
6	97.1	18	99.1	35	99.9
46a	97.1	21	99.1	39	99.9
88	97.5	53	99.1	42	99.9
110	98	84a	99.1	43	99.9
115	98.3	85	99.1	46b	99.9
11	98.4	116	99.1	49	99.9
1	98.4	3	99.2	50	99.9
115b	98.4	27	99.2	50d	99.9
89	98.5	105	99.2	51	99.9
90a	98.5	46	99.3	60	99.9
7a	98.6	65	99.3	62	99.9
113	98.6	76	99.3	64	99.9
4	98.7	77a	99.3	70	99.9
107	98.7	91	99.3	73	99.9
112	98.9	23	99.4	74	99.9
		53a	99.4	80	99.9
		106	99.4	81	99.9
		108	99.4	90b	99.9
		109	99.4	99	99.9
		7	99.5	26	99.96
		14	99.5	16	99.97
		14b	99.5	44	100
		22	99.5	49a	100
		28	99.5	49b	100
		31	99.5	50a	100
		94	99.5	66	100
		5	99.6	67	100
		5a	99.6	70a	100
		14a	99.6	71	100
		30	99.6	82	100
		34	99.6	82a	100
		41	99.6	95	100
		59	99.6	96	100
		63	99.6	97	100
		63a	99.6	98	100
		79	99.6	98a	100
		84	99.6	103	100
		90	99.6	104	100
		111	99.6		
		113a	99.6		
		72	99.7		
		9	99.7		
		12	99.7		
		13	99.7		

		15	99.7		
		24	99.7		
		29	99.7		
		50b	99.7		
		50c	99.7		
		54	99.7		
		56	99.7		
		57	99.7		
		58	99.7		
		63b	99.7		
		65a	99.7		
		65b	99.7		
		68	99.7		
		72a	99.7		
		75	99.7		
		77b	99.7		
		78	99.7		
		86	99.7		
		92	99.7		
		101	99.7		
		8	99.8		
		17	99.8		
		19	99.8		
		25	99.8		
		32	99.8		
		33	99.8		
		36	99.8		
		36a	99.8		
		37	99.8		
		38	99.8		
		40	99.8		
		45	99.8		
		47	99.8		
		48	99.8		
		52	99.8		
		55	99.8		
		55a	99.8		
		61	99.8		
		69	99.8		
		77	99.8		
		83	99.8		
		87	99.8		
		93	99.8		
		94a	99.8		
		100	99.8		
		102	99.8		
4;13;1		62;29;2		35;5;1	
18		93		41	
22.2;72.2;5.6%		66.7;31.2;2.2 %		85.4; 12.2; 2.4%	

The moderately conserved genes, with conserved ORFs across all isolates, but showing variability in its length and integrity were further examined to determine which were the nucleotide changes, which were determining the observed variations, which are all summarized in Table 3 below.

Table 9. Nucleotide variations within moderately conserved RaHV3 ORFs.

Color coding shows fully conserved genes (yellow), moderately conserved (green) and poorly (blue).

Strain1= FO_2015T; Strain 2= 2017_ZH; Strain 3= BfHV1_HU_Sw; Strain 4= W17_4268; Strain 5= W16_9699; Strain 6= W14_1083

ORF	Conserved	Notes
1	yes	
2	Yes	
3	Partially	The frame is conserved, however, there is a 9nt deletion between 2008-16, with loss of a repeat element. Possible extra ORF in strains 3, 4 and 5
3a	Yes	
4	Partially	The frame is conserved, however there are alternative insertions/deletions between 3204-14 in all strains (poly_D area). Insertion in 4 and deletion in all the others
5	Yes	
5a	Partially	At 4001 there is a mutation G to A in strains 2,3,4,5 and 6 determining the occurrence of a start codon in this position and an extension of the frame of 81nt 5'.
6	Partially	Strain 6 has an insertion of 12nt from 5808 to 5819. It has also a 3 nt deletion at 5786-9, which is also present in strain 2 and 5. A 4nt insertion is present in strain 4 at nt 5841 and a 40nt deletion is present at nt 5801 of the same strain. Strain 5 has a mutation (C to A) at 5893 introducing a stop codon in frame and shortening the frame of 270nt 5' with a start codon at nt5721
7	Partially	Strain 2 has a deletion (T) at 9448, leading to a stop codon in frame at 9439, a new frame starts at 9320 and ends in frame with the others. Strain 5 has a deletion (T) at 10505 leading to a stop codon in frame at nt 10436. A new frame starts at 10500 and ends with the others. Strain 5 has a mutation (T to A) at 9345 resulting in a stop codon in that region. A new frame starts at 9320 and ends in frame with the others.
7a	Partially	The deletion (T) at 9448n of strain 2, leads to a stop codon in frame at nt9552. An insertion (T) at 9447 lead to a stop codon in frame at nt 9499 of strain 5.
8	Yes	
9	Yes	
10	Yes	At position 17873 there is a triplet G (GGG) insertion in strain 5 and 2 leading to no frame disruption. (change not added to the mutation table)
11	Yes	
12	Yes	
13	Partially	Strain 2 has a deletion (T) at nt 20903 leading to a stop codon in frame at 20900 and shortening the main frame of 164 nt 5'.

14	Partially	There is a mutation (T to A) in strains 3,4,5 and 6 suppressing the start codon in that position and a shortening of the main frame of 21 nt 5'.
14a	Yes	
14b	Yes	
15	Yes	
15a	Yes	
16	Yes	
17	Partially	Strain 6 has a deletion (T) at 27327, leading to a stop codon in frame at nt27342.
18	Yes	
19	Yes	
20	Yes	
21	Partially	In strain 3 there is a mutation (C to T) at nt 30367, leading to a stop codon in that site and shortening the frame of 150bp
22	Yes	
23	Partially	Strain 4 has a C to A mutation at 32764 creating a stop codon in frame in that position and shortening then the main frame of 171nt 5'.
24	Yes	
25	Yes	
26	Partially	Strain 3 has a deletion (G) at nt 35468 resulting in a stop codon in frame at 35441. A new frame starts at 35448, which ends with the other main ones.
27	Partially	Strain 3 has a deletion (A) at nt 35979, resulting in a stop codon in frame at nt 35952.
28	Partially	Strain 2 has a deletion (G) at nt 37251, resulting in a stop codon in frame at nt 37223. A new frame starts at 37173, which ends in frame with the others. Strain 5 has a deletion (C) at nt 37058 resulting in a stop codon at nt36884. A new frame starts at nt36978, ending in frame with the others.
29	Yes	
30	Yes	
30	Yes	
31	Yes	
32	Yes	
33	Yes	
34	Yes	
35	Yes	
36	Yes	
36a	Yes	
37	Partially	Strain 4 has a deletion (T) at nucleotide 46954 resulting in a stop codon in frame at nt46931. A new frame starts at nt 46818.
38	Yes	
39	Yes	
40	Yes	
41	Partially	Strain 3 has a double deletion at nt 52663 (GC) leading to a stop codon in frame (also thanks to a G to A mutation) at nt 52657.
42	Yes	
43	Yes	
44	Yes	
45	Yes	
46	Yes	Frame are conserved, however strain 2 has a 9nt insertion between 57203 and 57211. Strain 5 has a 6nt insertion between 57206-211, and strain 6

		has a 3 nt insertion at 57209-11. Finally, strain 4 has a 3 nt insertion at 57212-4. This region spans across a poly D.
46a	Partially	Frame are conserved, however strain 2 has a 9nt insertion between 57203 and 57211. Strain 5 has a 6nt insertion between 57206-211, and strain 6 has a 3 nt insertion at 57209-11. Finally, strain 4 has a 3 nt insertion at 57212-4. This region spans across a poly M.
46b	Yes	
47	Yes	
48	Partially	Strain 2 has a double deletion (TG) at nt 60168 leading to a stop codon in frame at nt60170. A new frame starts at 60285, which ends in frame with the other strains.
49	Partially	Strain 4 has a deletion (G) leading to a stop codon in frame at nt61413. A new frame starts at nt61478, ending in frame with the other strains. In strain 2 there is a deletion (C) at nt 64680 leading to an earlier start codon in frame at nt 64805, resulting in a 103nt longer frame 5'.
49a	Yes	
49b	Yes	
50	Partially	Strain 2 has a deletion (C) at nt 64680 resulting in a stop codon in frame at nt64697 and resulting in a later start codon at nt 64714 with a 56nt shorter frame 5' (this is the same deletion impacting the longer frame of strain 2 in ORF49). Strain 4 has a double insertion (CA) at nt 65900, resulting in a stop codon in frame at nt 65953. A new frame starts at nt65982 and it has a G to T mutation at 66516, introducing a stop codon there and truncating the frame. A new frame starts at nt66555 which also has a G to T mutation at 68019 introducing a stop codon there and truncating the frame. A new frame starts at 68163 and ends together with the other strains.
50a	Yes	
50b	Yes	
50c	No	Missing in strain 4
50d	Yes	
51	No	Missing in strain 4
52	Yes	
53	Partially	Frames are conserved, however, strain 3, 5 and 6 have a 3 nt deletion at 71167; Strain 4 has 3 nt insertion at 71764. This is a poly-D site. Strain 4 has a 9nt insertion at 72587-95, which is a poly-D area again.
53a	Yes	
54	Partially	Strain 4 has a deletion (C) at nt 73438, leading to a stop codon in frame at nt73421. A new frame starts at 73185 and it finished together with the other strains
55	Yes	
55a	Yes	
56	Partially	The frames are conserved, however, Strain 4 has a 3nt deletion at 77215
57	Partially	Strain 2 has a deletion (T) at nt 79788, leading to a stop codon in frame at nt 79750. A new frame starts at nt 79779, which ends together with the other frames. Strain 3 has a 9 nt insertion at 79880-79888.
58	Partially	Strain 2 has a double insertion (TG) at nt 82023 together with a deletion (A) at 81875, all leading to a delayed start of the frame at nt81693, with a 556nt shorter frame 5'. The TG insertion at 82023 is present also in strains 4 and 6, both resulting in a delayed start codon at nt 81945, with a shortening of 305nt, 5'.
59		Strain 2 has a deletion (C) at nt 83412, resulting in a stop codon in frame at nt83381. A new frame starts at 83289 and end together with the other

		strains; Strain 5 has a deletion (T) at nt 82761 leading to a stop codon in frame at nt82721; Strains 4, 5 and 6 have a triple deletion at nt 82836.
60	Yes	
61	Yes	
62	Yes	
63	Partially	Strain 3 has a 6nt deletion at nt 89174
63a	Yes	
63b	Yes	
64	Yes	
65	Partially	Strains 2, 4 and 5 have an insertion (A) at nt 94124, leading to a stop codon in frame at nt94165. A new frame starts at nt 94364 ending in frame with the other strains; Strain 6 has a mutation (C to A) at nt 94088, creating a stop codon in that site and truncating the frame; strain 3 has a deletion (A) at nt 91627, resulting in a stop codon in frame at nt91643 and a new frame starting at nt91639; Strains 2, 5 and 6 have an 18 nt insertion at 95588-605; Strain 4 has a 3nt deletion at 95585-7 and a 9nt insertion at nt 95588-96.
65a	Yes	
65b	Yes	
66	Yes	
67	Yes	
68	Partially	Strain 6 has a deletion (C) at nt 101231 leading to a stop codon in frame, which shortens the frame of 143nt 5' and starting at nt 101148. The main frame of strains 6 has a deletion (G) at nt 97495 leading to a stop codon in frame at nt97400, shortening the frame of 23nt 3'.
69	Yes	
70	Partially	Strain 4 has a deletion (T) at nt 102667 leading to a stop codon in frame at nt102679. A new frame starts at nt102792 and ends in frame with the other strains.
70a	Yes	
71	Yes	
72	Yes	
72a	Yes	
73	Partially	Strains 4 and 6 have a deletion (C) leading to a stop codon in frame at nt105684. A new frame starts at nt 105772 and ends in frame with the other strains.
74	Yes	
75	Yes	
76	Partially	Frame is conserved, however, Strain 4 has a 6nt insertion at nt106907.
77	Partially	Strain 4 has a deletion (C) at nt 110598, leading at a stop codon in frame at nt110463. A new frame starts at nt 110596 and has a deletion (T) at nt 108475, leading to a stop codon in frame at nt108459. A new frame starts at nt108442 and ends in frame with the other strains; Strain 6 has a deletion (G) at nt 109544 leading to a stop codon in frame at nt109521. A new frame starts at nt109498 and ends in frame with the other strains.
77a	Yes	
77b	Partially	Strain 6 has a deletion (G) at nt 109544 leading to an early start codon at nt109497 with the frame extended 7nt 5'.
78	Partially	Strain 4 has a deletion (T) at nt 111491, resulting in a stop codon in frame at nt111421.
79	Yes	
80	Yes	

81	Yes	
82	Yes	
82a	Yes	
83	Yes	
84	Yes	
84a	Yes	
85	Yes	
86	Yes	
87	Yes	
88	Partially	Strain 4 has a deletion (G) at nt 120836 resulting in a stop codon in frame at nt120798. A new frame starts at nt 128811 and ends in frame with the other strains; Strains 4 and 5 have a 31nt insertion at 122175-205. The same strains have a 4nt deletion at nt 122206. Strains 4 and 5 have a 4nt insertion at nt122109. Finally, strains 4 and 5 have a 31nt deletion at nt 122078-108.
89	Partially	Strains 4 and 5 have a 14nt insertion at 124098-111 and a 14nt deletion at nt124112-125.
90	Partially	Strain 5 has a 3nt insertion at nt 124743; Strains 2, 3 4 and 6 have a 3nt deletion at nt124746
90a	Partially	Strain 5 has a 3nt insertion at nt 124743; Strains 2, 3 4 and 6 have a 3nt deletion at nt124746
90b	Yes	
91	Yes	
92	Yes	
93	Yes	
94	Yes	
94a	Yes	
95	Yes	
96	Yes	
97	Yes	
98	Yes	
98a	Yes	
99	Partially	
100	No	
101	Yes	
102	Yes	
103	Yes	
104	Yes	
105	Yes	
106	Yes	
107	Yes	
108	Partially	Strain 4 has a mutation (C to A) resulting in a stop codon at nt144122.
109	Yes	
110	Partially	Strain 4 has a 24nt deletion at nt 147885-908. Strain 6 has a 12nt deletion at nt 147885-896.
110a	Partially	Strain 4 has a 24nt deletion at nt 147885-908. Strain 6 has a 12nt deletion at nt 147885-896.
111	Yes	
112	No	
113	Partially	Strain 4 has a 27nt deletion at nt 152392-418 and a 3nt insertion at nt 152419.

113a	Yes	
113b	Partially	Strain 4 has a 27nt deletion at nt 152392-418 and a 3nt insertion at nt 152419.
114	Partially	The frame is conserved however strains 4 and 5 have a 37nt insertion at nt153226-261 and a 38nt deletion at nt153262-299; Strain 3 has an insertion (A) at nt 153297 leading to a frame shift suppressing the stop codon at nt153201 and prolonging the main frame of 41nt 3', with a stop codon at nt 153160.
115	Partially	Strain 2 has a G to T mutation at nt 153626, suppressing the start codon at that site and shortening the frame of 234nt. An insertion in the first frame (not specific location) leads to a stop codon in frame at nt154269. A new frame starts at nt154087 and ends with the main frames. Strains 5 and 6 have also an insertion (not specific region) leading to a stop codon in frame at nt 154269. A new frame starts for both of them at nt 154087 ending in frame with the other strains. Strain 3 has a 19nt insertion at nt 155151-169.
115a	Yes	
115b	Partially	Strain 3 has a A to C mutation at 156176, originating a start codon in that position and prolonging the frame of 39nt 5'.
116	Yes	

By extracting all the nucleotide changes resulting in significant ORFs changes, we observed that the highest number of nucleotide changes in comparison with the type strain FO_2015T was observed in strain W17_4268, within Switzerland with 47 relevant nucleotide changes, which was even higher than the two Hungarian strain (BfHV1_HU_Sw). Remarkably, the isolate showing the least number of significant nucleotide changes in comparison with the reference strain FO_2015T, was the Hungarian strain (BfHV1_HU_Sw). Overall, the most common nucleotide changes were deletions, followed by insertion and mutations. All this information is summarized in Table 4 below.

Table 10. Nucleotide changes resulting in frame variation across RaHV3 isolates

Strain	Deletions	Insertions	Mutations	Fully conserved genes (main ORF)*,**,****,*****	Total changes
FO_2015T	1***			47	1
2017_ZH	14	6	2	30	22
BfHV1_HU_Sw	9	3	4	31	16
W17_4268	24	17	6	16	47
W16_9699	10	12	4	30	26
W14_1083	14	6	3	28	23
Total	72	44	19		

*Changes in 46/46a were counted twice
**The deletion of strain 2 in ORF49 and 50 are the same and were both counted.
***ORF4
****Changes in 90 and 90a were counted twice
***** Changes in 110 and 110a were counted twice

Similary to what we have described concerning the RaHV3 isolates, by examining the moderately conserved ORFs, we observed that the majority of the changes were present in only one isolate per

examined frame (22). The second most frequent occurrence was the presence of changes in two out of the 6 isolates across all those examined (n=11). Changes occurring simultaneously in 3, 4 or 5 in the same ORF across all the strains examined was observed in 3, 7 and 3 genes, respectively. This information is summarized in Table 11 below.

Table 11. Frame variations across isolates in moderately conserved genes.

N.	ORF*	Not conserved frames (out of 6)	Remarks	
1	5a	5		
2	6	1		
3	7	3		
4	7a	2		
5	13	1		
6	14	4		
7	17	1		
8	21	1		
9	23	1		
10	26	1		
11	27	1		
12	28	2		
13	37	1		
14	41	1		
15	46	4		
16	46a	4		
17	48	1		
18	49	2		
19	50	2		
20	53	4		
21	54	1		
22	56	1		
23	57	2		
24	58	3		
25	59	4		
26	63	1		
27	65	4		
28	68	1		
29	70	1		
30	73	2		
31	76	1		
32	77	2		
33	77b	1		
34	78	1		
35	88	2		
36	89	2		
37	90	5		
38	90a	5		
39	108	1		
40	110	2		
41	110a	2		
42	113	1		

43	113b	1		
44	114	3		
45	115	4		
46	115b	1		
1/6=22	2/6=11	3/6=3	4/6=7	5/6=3

*ORF3 and 4 are not counted (yet)

In tables 12 a-c are summarized all the identities of the variably or fully conserved ORFs on the basis of their pairwise identity and frame conservation.

Table 12a. *Bufonid herpesvirus 1* Open reading frames (ORF) features. Most conserved genes (pairwise identity). In yellow the ORFs with fully conserved frame.

ORF	Herpes or other virus Homologue	Protein	Predicted function/ description	Similarity or confidence	Position	Length (nt)	Frame
3	None	Unknown	C-type Mannose receptor-2	100%[b] (40% coverage)	1,531-2,442	912	-
16	CyHV-2	Ribonucleotide reductase-Subunit 2	De novo DNA synthesis	217/312[a] (69%)	25,176-26,138	963	+
20	CyHV-1	ORF136a	Signal peptide (immune-regulation)-CD27-like	41/93[a] (44%)	28,752-29,753	822	-
26	RHV1	ORF95	Capsid triplex subunit 2	169/317[a] (53%)	34,599-35,525	927	-
35	CyHV3	ORF96	ORF96 (transferase inhibitor)	24/55[a] (43%)	44,292-44,627	336	-
39	RHV1	ORF42	putative ATPase subunit of terminase	191/264 [a] (72%)	50,273-51,247	951	-
42	RHV1	ORF83	multiple membrane-spanning protein	104/173[a] (60%)	53,485-54,207	723	+
43	RHV1	ORF82	multiple membrane-spanning protein	72/165[a] (43%)	54,324-54,932	609	-
44	RHV2	ORF115	multiple membrane-	38/87[a]	54,942-55,322	381	-

			spanning protein	(43%)			
46b	Unknown	Unknown	Unknown	N/A	57,382-57,657	276	+
49	RHV1	ORF73	similar to RaHV-2 ORF111 and IcHV-1 ORF56	675/1234[a] (54%)	60,655-64,401	3,747	-
49a	None	Unknown	Unknown	N/A	61,932-62,210	279	+
49b	None	Unknown	Unknown	N/A	62,738-63,112	375	+
50	RHV1	ORF72	DNA polymerase	843/1480[a] (56%)	64,356-68,672	4,317	+
50a	None	Unknown	Transcription motif	24%[b] (28% coverage)	64,425-64,769	345	-
50d	None	Unknown	DNA repair motif	58%[b] (34% coverage)	67,562-67,843	282	-
51	None	Unknown	Immunoglobulin -like beta-sandwich motif	15.1%[b] (40% coverage)	68,673-68,954	282	-
60	None	Unknown	Unknown	NA	84,008-84,595	588	-
62	RHV1	ORF59	Hypothetical Protein	215/363[a] (59%)	85,182-86,285	1,104	-
64	RHV1	ORF57	Hypothetical protein	110/277[a] (39%)	88,881-89,738	858	-
66	RHV1	ORF55	similar to RaHV-2 ORF81	131/226[a] (57%)	96,201-97,007	807	-
67	None	Unknown	Hydrolase	33.3%[b] (22% coverage)	96,804-97,121	318	+
70	RHV1	ORF52	Similar to RaHV-2 ORF78 and IcHV-1 ORF37	324/622[a] (52%)	101,479-103,404	1,926	+
70a	None	Unknown	Transcription/transcription activator	16%[b] (31% coverage)	103,095-103,397	303	-

71	RHV1	ORF50	similar to RaHV-2 ORF76; similar to IcHV-1 ORF35	75/175[a] (42%)	103,401-103,838	438	-
73	RHV1	ORF49	similar to RaHV-2 ORF75 and IcHV-1 ORF34	274/450[a] (60%)	104,262-105,617	1,356	-
74	None	Unknown	Delta-endotoxin-like	52.2%[b] (23% coverage)	105,461-106,126	666	-
80	RHV1	ORF42	Terminase	123/174[a] (70%)	113,375-113,902	528	+
81	RHV1	ORF43	similar to RaHV-2 ORF69 and IcHV-1 ORF70	101/198[a] (51%)	113,808-114,665	858	-
82	RHV1	ORF42	Terminase	156/269[a] (57%)	114,569-115,366	798	+
82a	None	Unknown	Hydrolase-like motif	60.9%[b] (31% coverage)	114,778-115,101	324	-
90b	None	Unknown	Unknown	N/A	125,746-126,030	285	+
95	None	Unknown	Putative short chain cytokine motif	30.7%[b] (16% coverage)	131,847-132,170	324	+
96	None	Unknown	Putative calcium-binding protein	21.3%[b] (16% coverage)	132,167-132,445	279	+
97	None	Unknown	tumor necrosis factor receptor superfamily member 10a	99.6%[b] (77% coverage)	132,174-132,479	306	-
98	RHV1	ORF26	similar to RaHV-2 ORF95; similar to IcHV-1 ORF75	159/317[a] (50%)	132,739-133,665	927	-
98a	None	Unknown	Isomerase	31.4%[b] (34% coverage)	133,133-133,576	444	+
99	RHV1	ORF25	Predicted Zn-binding protein	167/337[a] (49%)	133,742-134,734	993	+

103	None	Unknown	Unknown	N/A	139,184-139,807	624	-
104	TeHV3	UL17	Tegument protein	35/75[a] (46%)	139,862-140,737	876	-

Table 12b. *Bufonid herpesvirus 1* Open reading frames (ORF) features. Moderately conserved genes- (pairwise identity). In yellow the ORFs with fully conserved frame.

ORF	Herpes or other virus Homologue	Protein	Predicted function/ description	Similarity or confidence	Position	Length (nt)	Frame
2	None	Unknown	Unknown	N/A	560-952	393	+
3	None	Unknown	C-type Mannose receptor-2	100%[b] (40% coverage)	1,531-2,442	912	-
5	None	Unknown	Unknown	N/A	3,529-4,434	906	-
5a	None	Unknown	Unknown	N/A	4,021-4,347	327	+
7	HHV1	US11	DNA binding protein	52%[a] 25/48	7,280-10,684	3,405	-
8	None	Unknown	Unknown	N/A	13,193-14,710	1,518	-
9	CeHV9	UL20	Envelope protein	46%[a] 24/52	14,832-16,133	1,302	-
10	None	Unknown	Allene oxide cyclase-like	29.2%[b] (20% coverage)	17,430-18,482	1,053	-
12	Hantavirus	Unknown	Nucleoprotein	24.4%[b] (4% coverage)	19,163-19,837	675	-
13	EBV	Unknown	gp42	99.8%[b] (56% coverage)	20,175-20,747	573	-
14	None	Unknown	Unknown	N/A	21,714-22,190	477	-

14a	None	Unknown	De novo protein	52%[b] (15% coverage)	21,767-22,282	516	-
14b	None	Unknown	Transport protein	47.6%[b] (60% coverage)	21,835-22,248	414	+
15	None	Unknown	Hydrolase	100%[b] (79% coverage)	23,993-25,192	1,200	+
17	None	Unknown	Isomerase	100%[b] (86% coverage)	26,165-27,211	1,047	+
18	TeHV3	TE11	C-lectin type protein	70/118[a] (59%)	27,370-28,266	897	+
19	None	Unknown	Lectin-like	31.9%[b] (22% coverage)	28,735-29,043	309	+
21	None	Unknown	lectin-related NK cell receptor ly49l1	100%[b] (54% coverage)	29,934-30,593	660	-
22	None	Unknown	Unknown	N/A	30,862-32,136	1,275	-
23	None	Unknown	Unknown	N/A	32,031-32,510	546	-
24	None	Unknown	Unknown	N/A	32,465-34,015	1,551	+
25	RHV1	ORF99	Putative double stranded RNA binding motif	31/78[a] (39%)	34,076-34,564	489	-
27	RHV1	ORF94	Similar to IcHV-1 ORF26 (RNA bind-prot.)	68/146[a] (46%)	35,506-35,994	489	-
28	RHV1	ORF93	Helicase	312/531[a] (58%)	35,793-37,382	1,590	-
29	RHV1	ORF92	similar to IcHV-1 ORF24	117/245[a] (47%)	37,382-38,338	957	-
30	None	Unknown	Unknown	N/A	38,332-39,150	819	+
31	None	Unknown	Ribosome-like protein	61.4%[b] (26% coverage)	39,000-39,353	354	+

32	None	Unknown	Oxidoreductase -like protein	100%[b] (coverage 61%)	39,390-39,662	273	+
33	None	Unknown	Putative Ras association domain-containing family protein 5	29.4%[b] (9% coverage)	39,631-40,776	1,146	-
34	RHV1	ORF89	chromosome segregation protein SMC	202/524[a] (38%)	40,877-44,368	3,492	+
36	RHV1	ORF88	Similar to RaHV-2 ORF122 and IcHV-1 ORF64	260/510[a] (50%)	44,446-45,963	1,518	-
36a	Unknown	Unknown	Lyase	28[b] (8% coverage)	45,089-45,433	345	+
37	RHV1	ORF87	Helicase-primase subunit	309/683[a] (45%)	45,927-48,074	2,184	-
38	RHV1	ORF86	DNA (cytosine-5-)-methyltransferase	393/740[a] (53%)	48,129-50,426	2,298	+
40	RHV1	ORF85	Similar to IcHV-1 ORF61	126/284[a] (44%)	51,246-52,196	951	+
41	RHV1	ORF84	similar to RaHV-2 ORF118 and IcHV-1 ORF60	222/428[a] (51%)	52,254-53,483	1,230	-
45	Unknown	Unknown	Unknown	N/A	55,630-56,160	531	+
46	RHV1	ORF76	capsid triplex protein 1	116/273[a] (42%)	56,157-57,806	1,650	-
47	RHV1	ORF75	similar to RaHV-2 ORF113 and IcHV-1 ORF54	292/570[a] (51%)	57,817-59,454	1,638	+
48	RHV1	ORF74	probable similar to RaHV-2 ORF112 and IcHV-1 ORF55	176/383[a] (45%)	59,369-60,601	1,233	+
50b	None	Unknown	Ferrodoxin-like	39%[b] (37% coverage)	65,280-65,576	297	-
50c	None	Unknown	Unknown	NA	65,501-65,806	306	-

52	None	Unknown	Putative transcription factor	49.6%[b] (29% coverage)	68,953-69,414	462	+
53	RHV1	ORF69	Possible member of PK gene family	48/115[a] (41%)	69,415-72,456	3,042	-
53a	Unknown	Unknown	Unknown	NA	71,626-71,895	270	+
54	RHV1	ORF68	Protein Kinases, catalytic domain	229/528[a] (43%)	72,456-74,651	2,196	-
55	RHV1	ORF67	Predicted signal peptide	239/482[a] (49%)	74,641-76,752	2,112	+
55a	None	Unknown	Unknown	NA	76,368-76,691	324	+
56	RHV1	ORF66	member of ORF24 gene family-predicted signal peptide	188/327[a] (57%)	76,771-77,880	1,110	+
57	RHV1	ORF65	Zinc-RING-finger protein	137/354[a] (38%)	77,916-79,727	1,812	-
58	RHV1	ORF65	Zinc-RING-finger protein	182/438[a] 41.5%	79,858-81,909	2,052	-
59	RHV1	ORF63	putative capsid maturational protease	282/511[a] (55%)	82,207-84,054	1,848	-
61	HHV1	UL31	Putative virion egress molecule	50.6%[b] (22% coverage)	84,351-85,193	843	+
63	RHV1	ORF58	similar to IcHV-1 ORF43 (putative isomerase)	404/828[a] (48%)	86,284-88,884	2,601	+
63a	None	Unknown	Unknown	N/A	86,709-87,386	678	-
63b	None	Unknown	Unknown	N/A	88,391-88,663	273	-
65	Unknown	None	Unknown	43%[a] 280/641	89,802-96,137	6,336	+
65a	None	Unknown	Unknown	N/A	90,650-91,045	396	-

65b	None	Unknown	Unknown	NA	92,186-92,707	522	-
68	RHV1	ORF54	Major capsid protein	738/1330[a] (55%)	97,013-100,930	3,918	-
69	None	Unknown	Apopotosis-associated molecule (putative)	28.1%[b] (21% coverage)	100,896-101,564	669	-
72	None	Unknown	Unknown	N/A	103,835-104,413	579	-
72a	None	Unknown	Oxydoreductase-like	24.8%[b] (29% coverage)	104,023-104,466	444	+
75	None	Unknown	Unknown	N/A	105,975-106,514	540	+
76	RHV1	ORF47	membrane glycoprotein	103/203[a] (50%)	106,514-107,506	993	+
77	RHV1	ORF46	Membrane glycoprotein	667/1161[a] (57%)	107,352-110,897	3,546	-
77a	None	Unknown	Putative glycoprotein	9.5%[b] (44% coverage)	107,584-107,877	294	+
77b	None	Unknown	Unknown	N/A	109,136-109,564	429	+
78	RHV1	ORF45	membrane glycoprotein	216/394[a] (54%)	110,953-112,164	1,212	-
79	RHV1	ORF44	similar to RaHV-2 ORF70; similar to IcHV-1 ORF68	155/340[a] (45%)	112,052-113,128	1,077	+
83	None	Unknown	GIY-YIG endonuclease	97.6%[b] (18% coverage)	115,333-116,721	1,389	+
84	RHV1	ORF40	Thymidine Kinase-like	106/221[a] (47%)	116,822-117,472	651	+
84a	None	Unknown	Actin-binding protein	29.9%[b] (51% coverage)	117,052-117,333	282	-
85	None	Unknown	Unknown	N/A	117,884-118,525	642	-

86	Rhesus rotavirus	VP4	Sialic acid binding domain	33.1%[b] (17% coverage)	118,600-118,977	378	-
87	RHV1	ORF35	member of ORF35 gene family	150/393[a] (38%)	118,995-120,158	1,164	-
90	RHV1	ORF67	Predicted signal peptide	187/467[a] (40%)	124,140-126,656	2,517	-
91	RHV1	ORF33	Predicted protein (putative transcription regulator)	58/145[a] (40%)	126,681-128,069	1,389	-
92	None	Unknown	Unknown	N/A	128,087-129,313	1,227	-
93	None	Unknown	Unknown	NA	129,312-130,496	1,185	+
94	RHV1	ORF30	Member of ORF30 gene family	159/385[a] (41%)	130,507-131,646	1,185	+
94a	None	Unknown	Unknown	N/A	130,866-131,135	270	-
100	None	Unknown	Chelatase-like	38.6%[b] (29% coverage)	134,805-135,317	513	+
101	None	Unknown	Unknown	N/A	135,115-137,901	2,787	+
102	RHV1	ORF24	Predicted signal peptide	156/358[a] (43%)	137,888-139,120	1,233	+
105	None	Unknown	Unknown	N/A	141,709-142,488	780	-
106	None	Unknown	Unknown	N/A	142,593-143,438	846	-
108	None	Unknown	Unknown	N/A	143,556-144,599	1,044	-
109	None	Unknown	Unknown	N/A	144,766-145,857	1,092	+
111	Murine RV	ORF67	Capsid maturation protease (complement c3-like)	39/71[a] (54%)	148,521-149,216	696	+
113a	None	Unknown	Unknown	N/A	150,792-151,088	297	+
115a	None	Unknown	Unknown	N/A	153,737-154,063	327	-

116	None	Unknown	SipA N-terminal domain-like	71.2%[b] (21% coverage)	156,778-157,230	453	-

Table 12c. *Bufonid herpesvirus 1* Open reading frames (ORF) features. Poorly conserved genes- (pairwise identity). In yellow the ORFs with fully conserved frame.

ORF	Herpes or other virus Homologue	Protein	Predicted function/ description	Similarity or confidence	Position	Length (nt)	Frame
							-
1	None	Unknown	Unknown	N/A	232-777	546	
4	None	Unknown	Putative C-type lectin protein	100%[b] (40% coverage)	2,550-3,398	849	-
6	None	Unknown	Unknown	N/A	5,157-5,933	777	-
7a	None	Unknown	Transferase	44.1% [b] (59% coverage)	9,030-9,611	582	+
11	None	Unknown	DNA/RNA binding protein	68.9%[b] (42% coverage)	18,835-19,212	378	-
46a	Unknown	Unknown	Unknown	N/A	56,726-57,034	309	-
88	None	Unknown	Unknown	N/A	120,159-122,309	2,151	-
89	RHV1	ORF34	Major outer envelope glycoprotein	135/362[a] (37%)	122,477-124,132	1,656	-
90a	None	Unknown	Unknown	NA	124,159-124,443	285	+
107	None	Unknown	Unknown	N/A	143,489-143,776	288	-
110	None	Unknown	Hydrolase	95.2%[b] (14% coverage)	146,154-148,283	2,130	+
110a	AngHV1	ORF34	FAM75 family	54/97[a] (55%)	147,211-147,657	447	+

112	None	Unknown	Transferase-like	18.8%[b] (52% coverage)	149,580-149,897	318	-
113	PaHV-2	Unknown	Putative transcriptional regulator motif (ICP4)	19/33[a] (57%)	149,984-152,317	2,334	-
113b	None	Unknown	Unknown	N/A	151,543-152,139	597	-
114	None	Unknown	Unknown	N/A	152,744-153,028	285	-
115	LCV Macaca	BPLF1	Predicted tegument protein (Ph2 transc. Regulator c.97.1%-Cov.13%)	87/195[a] (44%)	153,125-155,725	2,601	+
115b	None	Unknown	Unknown	N/A	155,111-155,620	510	-

Among the most conserved genes are several involved in DNA metabolism and immunomodulatory genes such as ORF97, encoding for a Tumor necrosis factor receptor-like molecule.

Additional significant deletions/insertions

A series of insertion and deletions was observed either within ORFs or in intergenic regions in one or more isolates. Below is a list of these changes:
-43nt insertion between 1087-1129 in strain 6, not coding region
-4nt insertion between 1251-1254 in strain 6, not coding region
-9nt deletion between 2008-2016 in strain 4, ORF3, loss of a repeat element
-Insertion/deletion area between 3204-14 in strains in a poly-D stretch
-a 40nt deletion is present at nt 5801 of strain 4. Variable deletions present also in other strains. This is a repeat region (poly E-K motif)
-There is a 20 and 22 nt insertion in strains 2 and 5 respectively between 11152-73, a non coding region
-Strain 4 has an 82 nucleotide deletion between 11175 and 11256, not coding region as well. Also, it has an 80nt insertion at 11385-11464, not coding region. Finally, it has a 40nt insertion at 11647-88, not coding.
-Strain 5 has a 22nt insertion at 11362-83, not coding, and a 40nt deletion at 11606-45, not coding.
-ORF46/46a, strain 2 has a 9nt insertion between 57203 and 57211. Strain 5 has a 6nt insertion between 57206-211, and strain 6 has a 3 nt insertion at 57209-11. Finally, strain 4 has a 3 nt insertion at 57212-4. This region spans across a poly M.
-ORF53, Strain 3, 5 and 6 have a 3 nt deletion at 71167; Strain 4 has 3 nt insertion at 71764
-ORF56a, Strain 4 has a 3nt deletion at 77215

-ORF57, Strain 3 has a 9 nt insertion at 79880-79888.
-ORF59, Strains 4, 5 and 6 have a triple deletion at nt 82836.
-ORF63, Strain 3 has a 6nt deletion at nt 89174
-ORF65, Strains 2, 5 and 6 have an 18 nt insertion at 95588-605; Strain 4 has a 3nt deletion at 95585-7 and a 9nt insertion at nt 95588-96.
-ORF76, Strain 4 has a 6nt insertion at nt106907.
-ORF88, Strains 4 and 5 have a 31nt insertion at 122175-205. The same strains have a 4nt deletion at nt 122206. Strains 4 and 5 have a 4nt insertion at nt122109. Finally, strains 4 and 5 have a 31nt deletion at nt 122078-108.
-ORF89, Strains 4 and 5 have a 14nt insertion at 124098-111 and a 14nt deletion at nt124112-125.
-ORF90, Strain 5 has a 3nt insertion at nt 124743; Strains 2, 3 4 and 6 have a 3nt deletion at nt124746.
-Intergenic region ORF104-105, strain 4, 26 nt insertion and 4 nt deletion at nt141216-247.
-ORF110 and 110a, Strain 4 has a 24nt deletion at nt 147885-908. Strain 6 has a 12nt deletion at nt 147885-896.
-ORF113 and 113b, Strain 4 has a 27nt deletion at nt 152392-418 and a 3nt insertion at nt 152419.
-ORF114, strains 4 and 5 have a 37nt insertion at nt153226-261 and a 38nt deletion at nt153262-299.
-ORF115, Strain 3 has a 19nt insertion at nt 155151-169.
-Intergening region ORF115-116, there is a 40nt insertion in strain 1 at nt 156481-520 and an 11nt insertion in strain 2 at nt 156510-520.

Repeated motifs

Additionally to the analyses described above, we investigated the presence of short repeated motifs across the 7 isolates examined. The search settings included decamers repeated at least 5 times. The analysis was carried out using the online software Gramene SSrtool (https://archive.gramene.org/db/markers/ssrtool). Here below in figures 8a-g are reported the results of this analysis.

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-1	gt	5	1053	1062	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-2	ac	6	6098	6109	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-3	ca	61	6181	6302	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-4	at	5	16152	16161	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-5	ca	5	17020	17029	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-6	ca	5	65537	65546	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-7	tg	5	139521	139530	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-8	gtc	5	3079	3093	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-9	tct	5	5663	5677	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-10	tga	5	8199	8213	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-11	tca	9	56852	56878	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-12	cat	6	71398	71415	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-13	ttg	5	82432	82446	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-14	aga	5	95190	95204	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-15	tgt	5	124264	124278	158188
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-17	tttct	8	5696	5743	158188

Figure 8a. Short repeated motifs identified in strain FO_2015T

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-1	gt	5	1057	1066	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-2	ac	6	6099	6110	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-3	ca	61	6182	6303	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-4	at	5	16147	16156	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-5	ca	6	17015	17026	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-6	ca	5	65546	65555	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-7	tg	5	139555	139564	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-8	gtc	5	3083	3097	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-9	tga	5	8199	8213	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-10	tca	13	56853	56891	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-11	cat	6	71407	71424	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-12	ttg	5	82450	82464	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-13	aga	5	95190	95204	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-14	aga	14	95208	95249	158194
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-16	tttct	8	5697	5744	158194

Figure 8b. Short repeated motifs identified in strain 2017_ZH

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-1	gt	5	1057	1066	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-2	ac	5	6105	6114	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-3	ca	61	6186	6307	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-4	at	6	16152	16163	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-5	ca	7	17022	17035	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-6	ca	5	65533	65542	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-7	tg	5	139523	139532	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-8	tct	5	5670	5684	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-9	tga	5	8204	8218	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-10	tca	9	56849	56875	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-11	cat	5	71394	71408	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-12	ttg	5	82445	82459	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-13	gaa	8	95189	95212	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-15	ttttct	8	5703	5750	158154
CAAAATTAATAGTCTGGAGACAGGACATTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-18	taggcaaagg	7	5020	5089	158154

Figure 8d. Short repeated motifs identified in strain BfHV1_HU_Sw

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-1	ac	7	6077	6090	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-2	ca	61	6162	6283	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-3	at	6	16130	16141	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-4	ca	5	17000	17009	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-5	ca	6	65522	65533	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-6	tg	5	139528	139537	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-7	ctt	13	5673	5711	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-8	tca	10	56837	56866	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-9	cat	7	71386	71406	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-10	aga	11	95195	95227	158139
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-17	taggcaaagg	7	5016	5085	158139

Figure 8d. Short repeated motifs identified in strain W17_4268

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-1	ac	6	6097	6108	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-2	ca	61	6180	6301	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-3	at	5	16151	16160	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-4	ca	6	17019	17030	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-5	ca	5	65554	65563	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-6	tg	5	139561	139570	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-7	gtc	5	3081	3095	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-8	tga	5	8198	8212	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-9	tca	12	56861	56896	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-10	cat	5	71415	71429	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-11	aga	14	95209	95250	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-12	tgt	6	124301	124318	158184
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-14	ttttct	8	5695	5742	158184

Figure 8e. Short repeated motifs identified in strain W16_9699

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-1	gt	5	1057	1066	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-2	gt	5	1097	1106	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-3	ac	5	6110	6119	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-4	ca	61	6191	6312	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-5	at	5	16160	16169	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-6	ca	5	17028	17037	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-7	ca	5	65554	65563	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-8	tg	5	139543	139552	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-9	tga	5	8209	8223	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-10	tca	11	56864	56896	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-11	cat	5	71415	71429	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-12	aga	11	95202	95234	158156
CAAAATTAATAGTCTGGAGACAGGACACTGTATTTTAAATCTTTTACAAATGCCATTGTGTG-14	ttttct	10	5696	5755	158156

Figure 8f. Short repeated motifs identified in strain W14_1083

Phylogenetic analysis

Similary to the RaHV3 isolates, phylogenetic analysis was carried out using either the whole genome of the isolates using the specific Genious software package application or specific genes, which would provide a similar tree topology to that obtained with the full genome and accordingly, ideal to be used as proxy of the whole genome itself using the software package Mega 7 (https://www.megasoftware.net/). Phylogenetic analysis unambiguously identified at least two separate clusters of Swiss BfHV1 strains, whereas the two Hungarian isolates appear to be cluster separately from the Swiss strains as shown in Figure 9 below. The cluster gathering 4 of the 5 Swiss strains investigated appears to be the most common and widespread in Switzerland and is arbitrarily named as lineage 1, whereas the second cluster composed of a single isolate, putative representing a less widespread BfHV1 strain in Switzerland is arbitrarily named lineage 2.

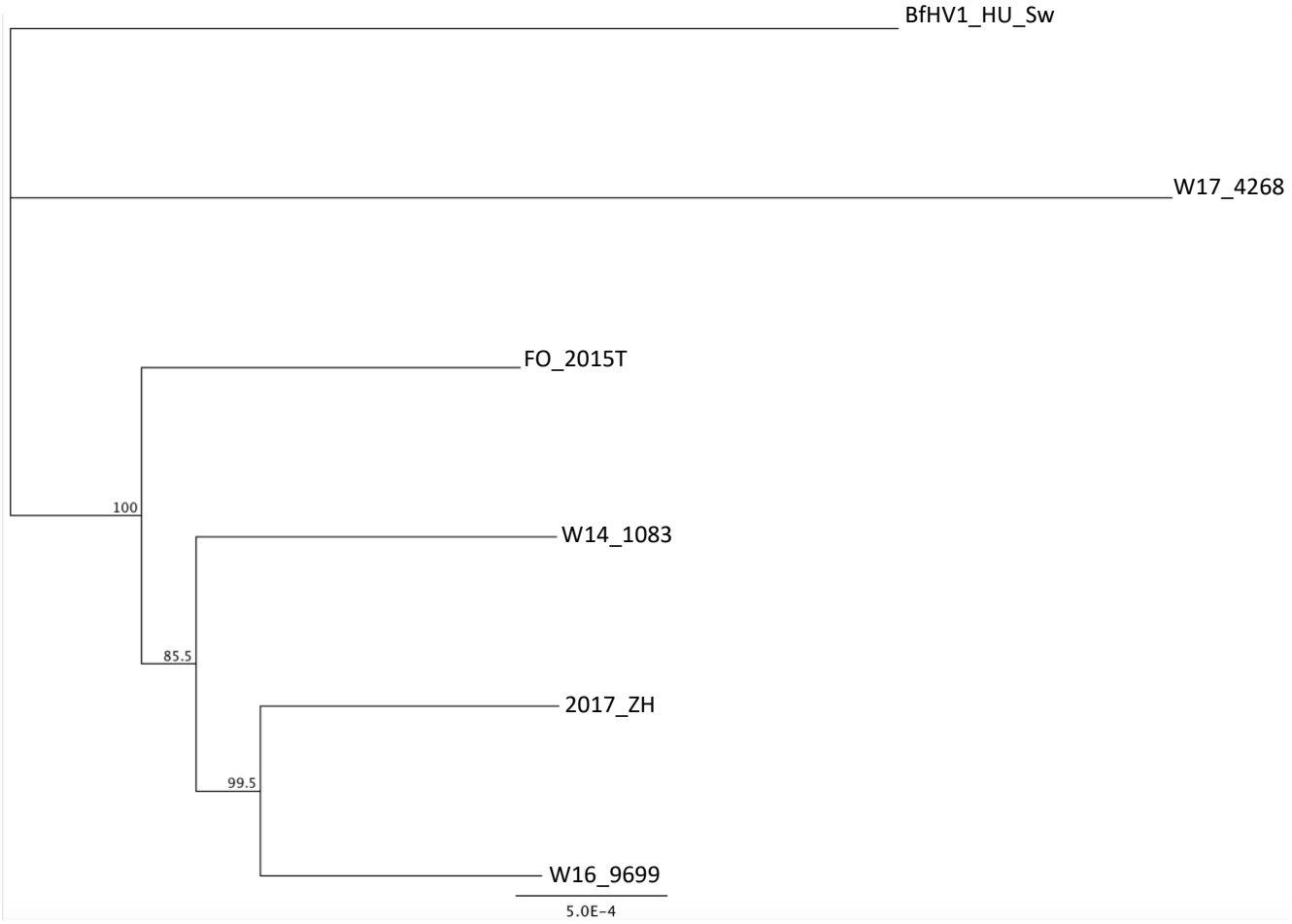

Figure 9. Whole genome-Neighbor-Joining consensus tree Nucleotide alignment of the investigated RaHV3 isolates. Bootstraps value are indicated at the branching of the nodes.

Similar tree topology has been obtained using the entire or partial sequences of the ORF46a, a poorly conserved gene, ideal for fine genetic resolution. Below is the correspondent tree (Figure 10).

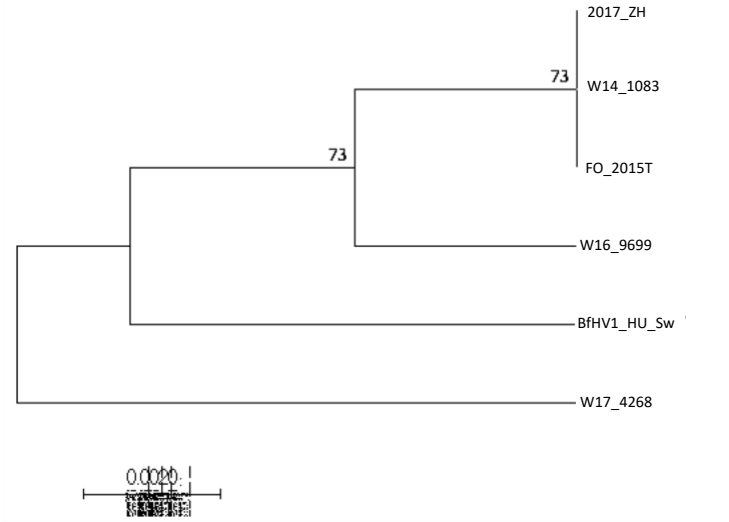

Figure 10. Maximum likelihood phylogenetic tree based on the sequence of the ORF46a. Bootstraps value are indicated at the branching of the nodes. The tree shows a similar topology to that obtained with the whole genome (Figure 9)

Phylogeography

The results of the analysis on the available BfHV1 isolates revealed the existence of at least two distinct lineages present in Switzerland. Lineage 1, has been detected in different ponds in Canton Zurich and in Basel land. Differently, the only strain of the putative lineage 2 has been detected in Canton Neuchatel. The Hungarian strain appears to cluster separately from the Swiss strains, differently from what observed with the RaHV3 strains, which showed closer relatedness with one of the Hungarian isolates.

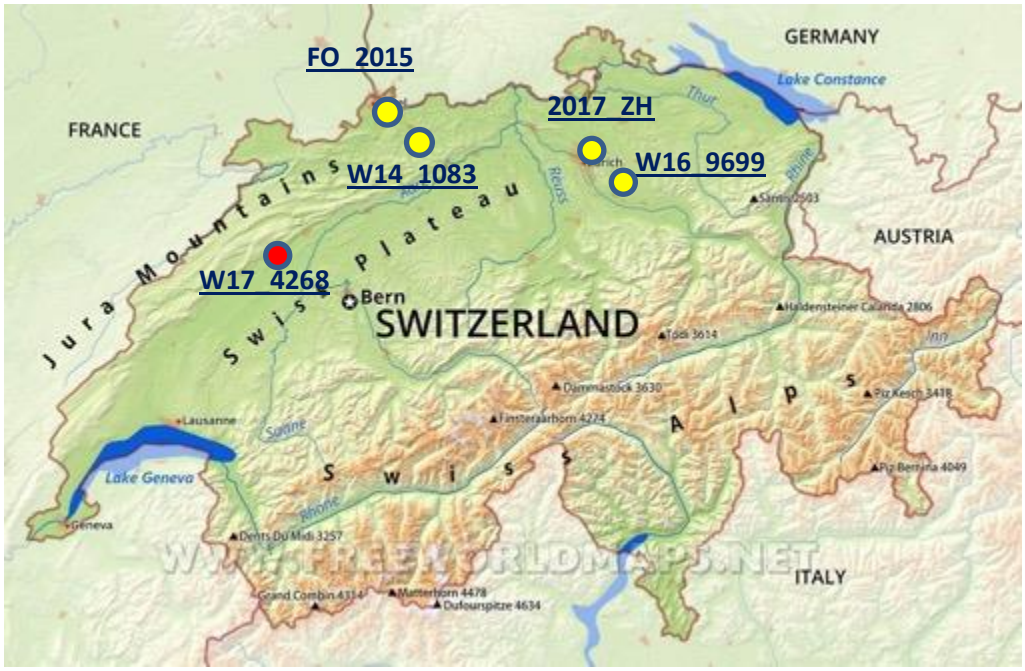

Figure 11. Map of Switzerland showing the origin of the BfHV1 isolates. Yellow dots show the origin of BfHV1 lineage 1, whereas, the single red dot shows the origin of the isolate belonging to lineage 2 (https://www.freeworldmaps.net/europe/switzerland/switzerland-physical-map.jpg).

Conclusions (BfHV1)

The analysis of the 6 BfHV1 isolates available have revealed that BfHV1 contains genes sharing between 85.1 to 100% pairwise identity. The genes can be grouped in fully conserved, partially conserved and not conserved. Fully conserved share overlapping ORFs and reading frames in all the strains; partially conserved genes share at least a portion of the specific frame in all the strains; not conserved genes lack the homologous ORF in one or more of the strains. Overall RaHV3 genome contains 101 fully conserved genes; 47 partially conserved genes and 4 not conserved genes. Predominant conserved genes are within the group of genes with higher pairwise identity.

The majority of the partially conserved genes are characterized by changes in a single strain (2). Changes in up to two strains are present in 11 genes, in up to three strains are present in 3 genes and in up to 4 strains are detected in 7 genes. Changes in up to 5 strains are present in 3. The partially conserved genes are characterized by the accumulation of mutations (insertions, deletions, substitutions) responsible of frame changes (frameshift or/n and truncation). Of these, the majority were deletions (72) followed by almost equivalent insertions (44) and mutations (19). Strains 3 and 6 are those with more mutations detected. Among the most conserved genes are those encoding for proteins involved in DNA metabolism and interestingy in immunomodulation (TNFr). Phylogenetic analysis revealed the presence of two distinct lineages, with a more diffuse one (lineage 1) including strains from Zurich and Basel land and a less diffuse one (lineage 2) including a strain from Neuchatel. A reliable phylogenetic marker has been identified in the partial sequence of 46a. This marker will allow the fine characterization of RaHV3 phylogeography in Switzerland and the potential association between the different lineages, population dynamics and disease occurrence.

Overall project conclusions

This investigation allowed for the first time a through characterization of 13 Batrachovirus isolates, including 7 RaHV3 and 6 BfHV1 isolates. A very important achievement of this investigation is the identification within both RaHV3 and BfHV1 of the existence of at least two distinct viral lineages, apparently associated with distinct geographic areas. Despite the number of strains available is relatively low and does not allow to draw more definitive conclusions, the observations of the existence of more than one, and probably more lineages within these viral species will allow in the future to clarify the hypothesis which were at the basis of this investigation. More specifically, the mid and long-term monitoring of free ranging frogs and toads in Switzerland will allow to verify if any of the current discovered lineages, and eventually of others still to be detected, is associated with a distinct and/or more variable pathology and if any of these lineages is associated with population declines in any of the regions where it is observed. The identification of reliable markers, which thanks to a simple PCR and sequencing will allow to genotype and correlate the presence and circulations of different strains with variation of population dynamics will be a major diagnostic and conservation tools for amphibians in Switzerland and in Europe. Ongoing collaborations with colleagues from other European countries will allow to draw a more complete and more comprehensive epidemiological picture of the variability of RaHV3 and BfHV1, contributing to understand not only within Switzerland, but at a European scale the meaning of RaHV3 and BfHV1 in the disease ecology of free-ranging amphibian populations.

Additionally, the analysis revealed that BfHV1 strains contain ORF with an overall higher pairwise identity across strains than with RaHV3, suggesting that BfHV1 might be an “older” Batrachovirus species. In both RaHV3 and BfHV1, it appears that the majority of the genes are well conserved across strains and only a minority show a relatively high variability. Of great interest is that among the relatively highly conserved genes, we observed also the presence of putative immunomodulatory genes, such as those predicted to encode a TNFr-like molecule. In a recent parallel investigation, we have shown how this gene is highly transcribed both in RaHV3 infected frogs and BfHV1 infected toads. The observation of its relatively high stability across strains and homologous viral species (RaHV3 and BfHV1) is highly suggestive of a relevant role in the pathogenesis of the associated disease opening the venue for future investigation for better understanding the pathogenesis of these diseases with consequential benefits to amphibian conservation.

Acknowledgements

I would like to thank Drs Frances Cordillot and Beatrice Werffeli for supporting this project. I am in debt with Benedikt Schmidt, Petra Lohmann, Ursula Sattler, Patricia Otten and Helmut Segner for their help and support.

References

Origgi FC, Schmidt BR, Lohmann P, Otten P, Akdesir E, Gaschen V, Aguilar-Bultet L, Wahli T, Sattler U, Stoffel MH. Ranid Herpesvirus 3 and Proliferative Dermatitis in Free-Ranging Wild Common Frogs (Rana Temporaria). Vet Pathol. 2017 Jul;54(4):686-694. doi: 10.1177/0300985817705176.

Origgi FC, Schmidt BR, Lohmann P, Otten P, Meier RK, Pisano SRR, Moore-Jones G, Tecilla M, Sattler U, Wahli T, Gaschen V, Stoffel MH. Bufonid herpesvirus 1 (BfHV1) associated dermatitis and mortality in free ranging common toads (Bufo bufo) in Switzerland. Sci Rep. 2018 Oct 3;8(1):14737. doi: 10.1038/s41598-018-32841-0.