

Computational Exploration of Virus Diversity on Transcriptomic Datasets

Digitaler Anhang
der
Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Simon Käfer
aus Andernach

Bonn 2019

Table of Contents

1	Preliminary Work - Phylogenetic Tree Reconstruction	3
1.1	Non-segmented RNA Viruses	3
1.2	Segmented RNA Viruses	4
1.3	Flavivirus-like Superfamily	5
1.4	Picornavirus-like Viruses	6
1.5	Togavirus-like Superfamily	7
1.6	Nidovirales-like Viruses	8
2	TRAVIS - True Positive Details	9
2.1	INSnfrTABRAAPEI-14	9
2.2	INSnfrTADRAAPEI-16	10
2.3	INSnfrTAIRAAPEI-21	11
2.4	INSnfrTAORAAPEI-35	13
2.5	INSnfrTATRAAPEI-43	14
2.6	INSnfrTBERAAPEI-19	15
2.7	INSytvTABRAAPEI-11	16
2.8	INSytvTALRAAPEI-35	17
2.9	INSytvTBORAAPEI-47	18
2.10	INSswpTBBRAAPEI-21	19
2.11	INSeqtTAHRAAPEI-88	20
2.12	INShkeTCLRAAPEI-44	22
2.13	INSeqtTBNRAAPEI-11	23
2.14	INSeqtTCJRAAPEI-20	24
2.15	INSeqtTCZRAAPEI-47	25
2.16	INSeqtTDXRAAPEI-19	27
2.17	INSlupTBDRAAPEI-17	28
2.18	INSlupTBKRAAPEI-31	29
2.19	INSlupTBMRAAPEI-34	30
2.20	INSlupTBURAAPEI-45	31
2.21	INSlupTAFRAAPEI-44	32
2.22	INSntgTABRAAPEI-216	33
2.23	INSlupTASRAAPEI-89	34
2.24	INSqiqTBFRAAPEI-61	36
2.25	INSqiqTBLRAAPEI-83	37
2.26	INSqiqTBNRABPEI-90	38
2.27	INSqiqTCTRAAPEI-75	40

2.28	INSIupTATRAAPEI-90	41
2.29	INSqiqTCXRAAPEI-90	42
2.30	INSqiqTDLRAAPEI-72	43
2.31	INSobdTDTRAAPEI-18	44
2.32	INSobdTDYRAAPEI-30	45
2.33	INSerITCGRAAPEI-32	47
2.34	INSkzdTABRAAPEI-136	48
2.35	INSkzdTACRAAPEI-171	49
2.36	INSofmTBLRAAPEI-71	50
2.37	INSofmTCYRAAPEI-79	51
2.38	INSqiqTDDRABPEI-136	52
2.39	INSfkjTBIRAAPEI-202	53
2.40	INSerITAKRAAPEI-83	54
2.41	INSfkjTBMRAAPEI-206	55
2.42	INSofmTCERAAPEI-22	57
2.43	INSofmTCFRAAPEI-26	58
2.44	INSinITAARABPEI-43	59
2.45	INSinITAPRAAPEI-33	60
2.46	INSinITAWRAAPEI-44	61
2.47	RINSinITCARAAPEI-55	62
2.48	RINSinITCNRAAPEI-33	63
2.49	RINSymITABRAAPEI-202	64
2.50	RINSwvkTAURAAPEI-56	65
2.51	ANIsrmTAAWRAAPEI-225	66
2.52	WHANIsrmTMAFRAAPEI-14	67
2.53	WHANIsrmTMCHRAAPEI-56	68
2.54	INSeqtTBBRAAPEI-75	69
2.55	INSobdTDIRAAPEI-84	71

1 Preliminary Work - Phylogenetic Tree Reconstruction

1.1 Non-segmented RNA Viruses

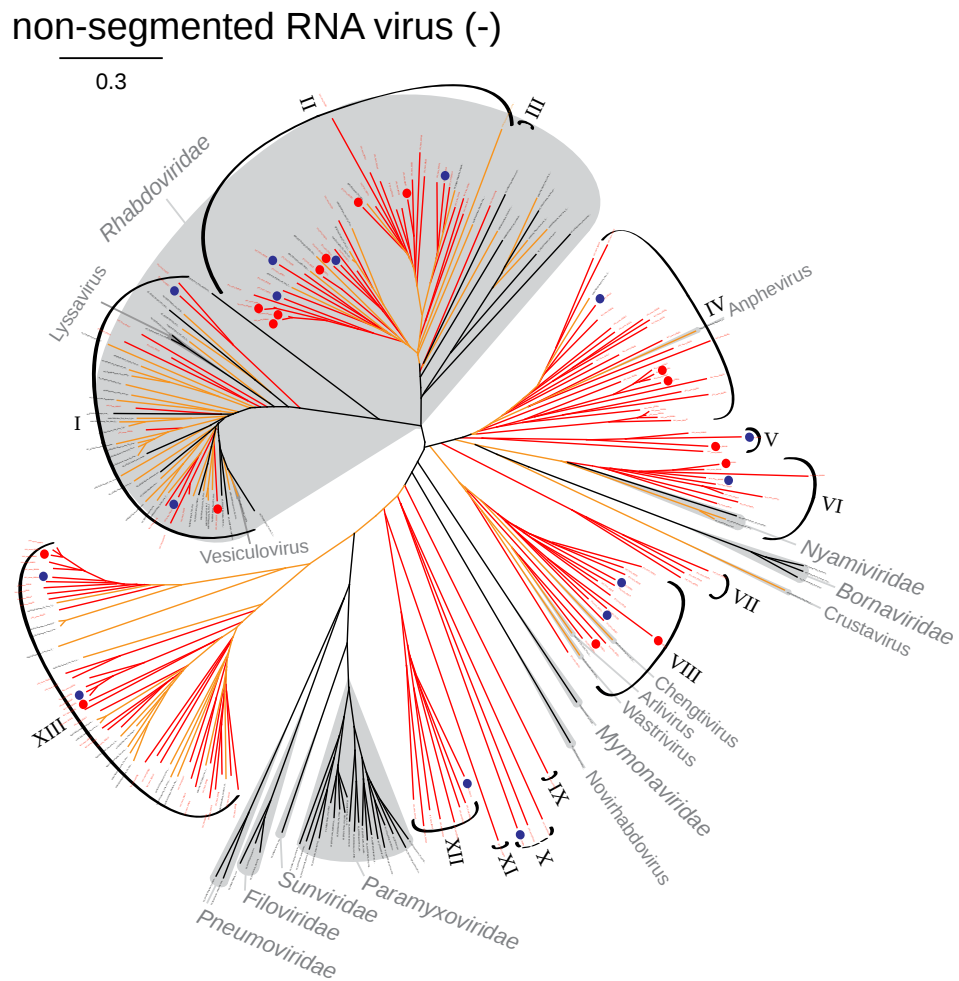


Figure 1: Non-segmented RNA Viruses

[h]

1.3 Flavivirus-like Superfamily

Flavivirus-like superfamily (+)

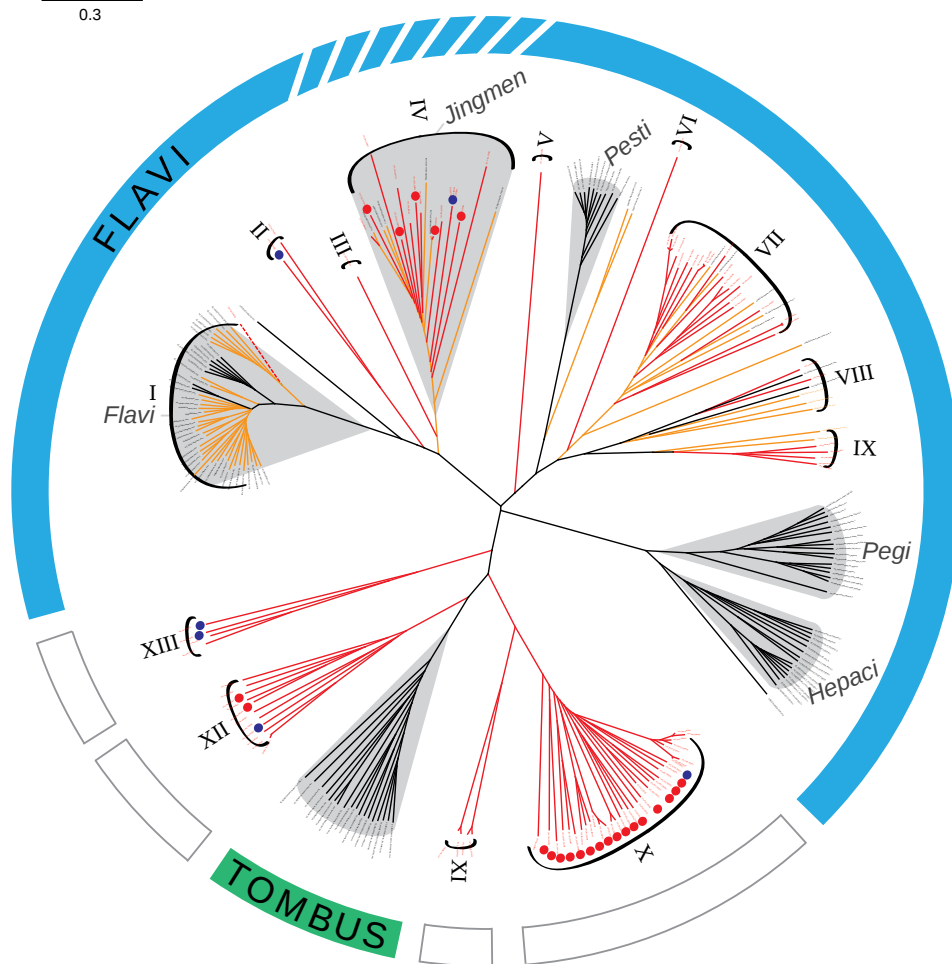


Figure 3: Flavivirus-like Superfamily

1.4 Picornavirus-like Viruses

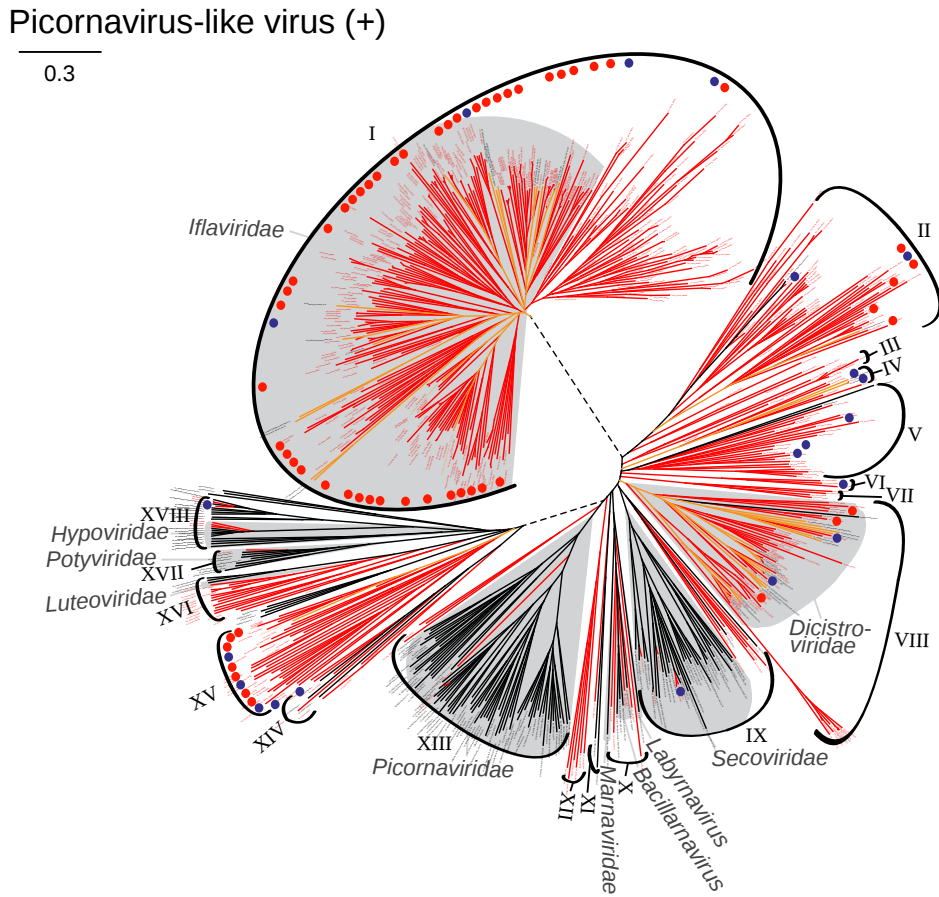


Figure 4: Picornavirus-like Viruses

1.5 Togavirus-like Superfamily

Togavirus-like superfamily (+)

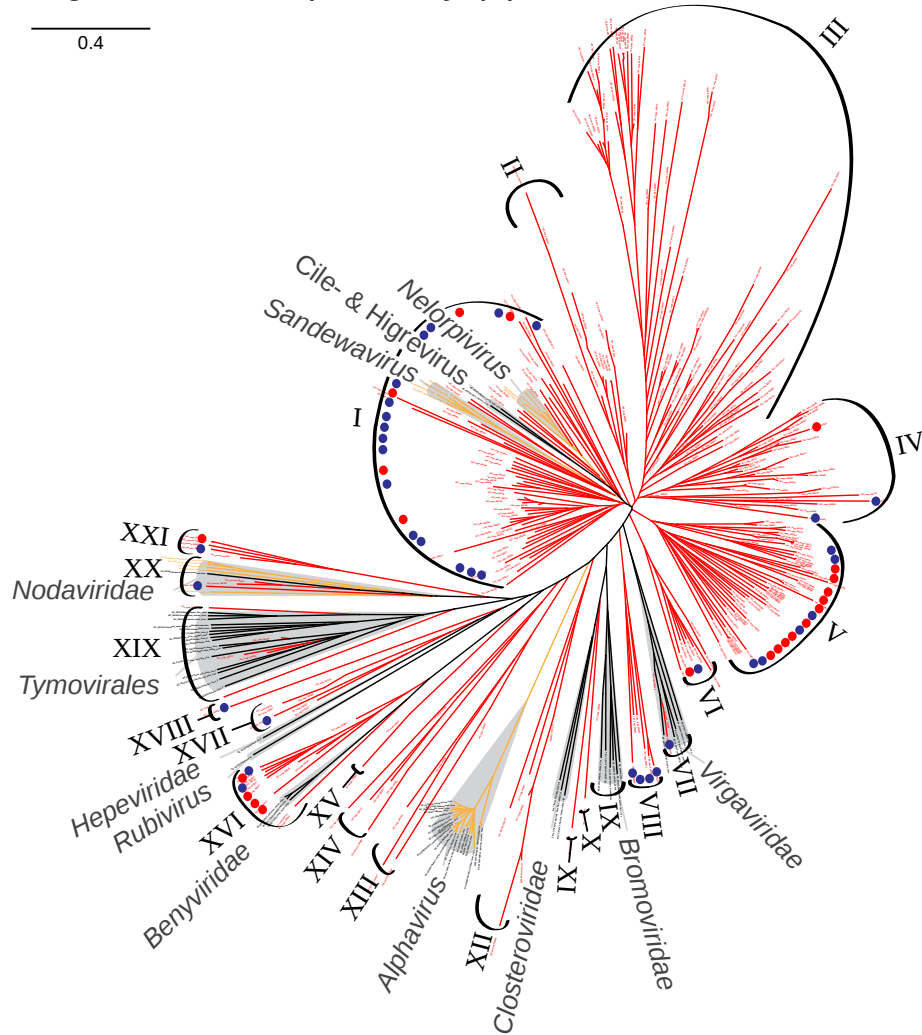


Figure 5: Togavirus-like Superfamily

1.6 Nidovirales-like Viruses

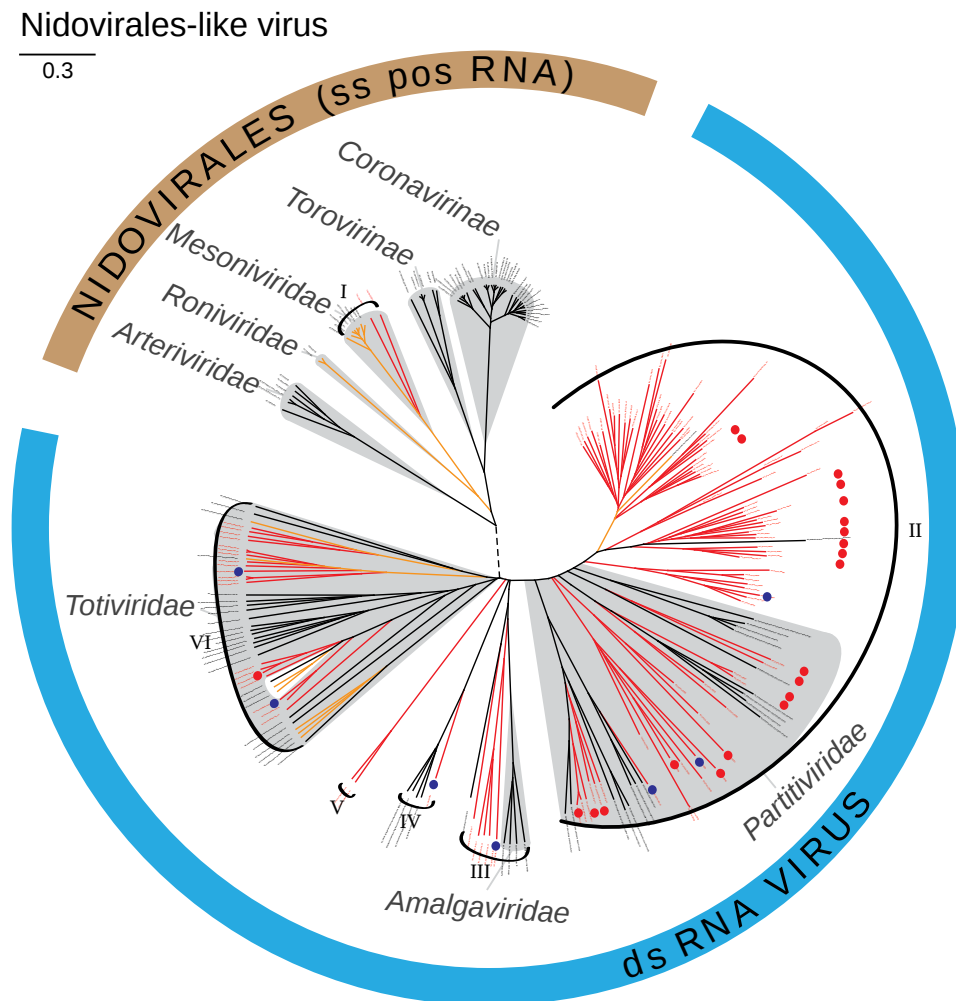


Figure 6: Nidovirales-like Viruses

2 TRAVIS - True Positive Details

2.1 INSnrTABRAAPEI-14

This transcriptome contained a starting region of the RdRp segment and the end region of a capsid protein. It also contained a sequence similar to a virus cell attachment protein. However this is questionable because it belongs to a group of sequences that caused a lot of the false positives. chapter 3.2.2.

Table 1: Sample Information of INSnrTABRAAPEI-14.

Filename	120107_I247_FCD0KMHACXX_L8_INSnrTABRAAPEI-14.free.fas
Assembly ID	INSnrTABRAAPEI-14
Order	Hemiptera
Order details	Heteroptera
Family	Pleidae
Family details	NA
Species	<i>Plea minutissima</i>
Number of specimen	60
Stage	adult
Sample location	Germany, Lower Saxony, Lüchow-Dannenberg, Hühbeck, Pevestorf
Sample date	Aug-2011
Blood-feeding	no
Suspicious sequences	18

Table 2: Suspicious Sequences in INSnrTABRAAPEI-14.

2 of 18 sequences were true positives, 2 are questionable and 14 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C129227_a_11_0_l_843	ORF_005	1. sigma 1, Mammalian Orthoreovirus (JQ412761) 2. segment 1, Nelson bay reovirus (AF218360)	18% 26%	partial (start) partial (start)
C133938_a_8_0_l_1034	ORF_003	1. RdRp, Rice gall dwarf virus (DQ494209) 2. RdRp, Homalodisca vitripennis reovirus (GU362064)	36% 34%	partial (start) partial (start)
C138664_a_11_0_l_1299	ORF_006	1. minor outer capsid, Rice gall dwarf virus (AY556484) 2. segment 2, Homalodisca vitripennis reovirus (GU369683, GU384984)	27% 27%	partial (end) partial (end)
(?) C143329_a_10_0_l_1751	ORF_005	1. segment 10, Dendrolimus punctatus cypovirus (YP_009111319) 2. poly ADP-ribose glycohydrolases (GBGD01000886, XM_012431649)	30% 40-50%	partial partial

INSnrTABRAAPEI-14 Sequence Organization

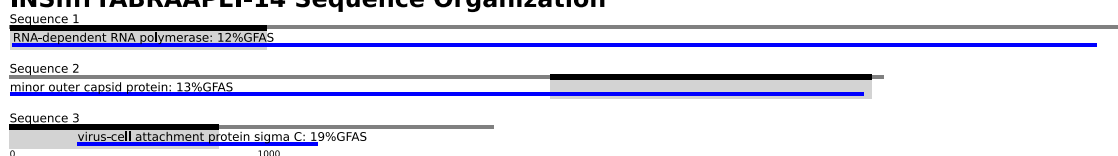


Figure 7: Sequence Organization of INSnrTABRAAPEI-14.

2.2 INSnrTADRAAPEI-16

This transcriptome did not contain a valid true positive RdRp. However, a sequence similar to *Hubei Chuvirus-like virus* and *Lishi Spider Virus* has been found with a similar sequence organization.

Table 3: Sample Information of INSnrTADRAAPEI-16.

Filename	120107_I247_FCD0KMHACXX_L8_INSnrTADRAAPEI-16.free.fas
Assembly ID	INSnrTADRAAPEI-16
Order	Coleoptera
Order details	NA
Family	Dytiscidae
Family details	NA
Species	<i>Cybister lateralimarginalis</i>
Number of specimen	1
Stage	adult
Sample location	Germany, Lower Saxony, Lüchow-Dannenberg, Hühbeck, Pevestorf
Sample date	27-Aug-2011
Blood-feeding	no
Suspicious sequences	16

Table 4: Suspicious Sequences in INSnrTADRAAPEI-16.

1 of 16 sequences was true positive and 15 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C127169_a_60_0_l_4158	ORF_001	1. glycoprotein, Hubei chuvirus-like virus 3 (NC_033015)	47%	full
		2. glycoprotein, Lishi Spider Virus 1 (KM817596)	36%	full
C127169_a_60_0_l_4158	ORF_002	1. hypothetical protein, Hubei chuvirus-like virus 3 (NC_033015)	41%	full
		2. nucleoprotein, Lishi Spider Virus 1 (KM817596)	39%	full

INSnrTADRAAPEI-16 Sequence Organization



Figure 8: Sequence Organization of INSnrTADRAAPEI-16.

2.3 INSnrTAIRAAPEI-21

In this transcriptome, a full RdRp has been detected as well as seven other fragmentary segments. One of the other segments (VP2) might also be nearly complete because it consisted of two contigs where a central connecting part is missing.

Table 5: Sample Information of INSnrTAIRAAPEI-21.

Filename	120107_I247_FCD0KMHACXX_L8_INSnrTAIRAAPEI-21.free.fas
Assembly ID	INSnrTAIRAAPEI-21
Order	Collembola
Order details	NA
Family	Neanuridae
Family details	NA
Species	<i>Anurida maritima</i>
Number of specimen	ca 100
Stage	adult
Sample location	Netherlands, North Holland, Texel, Ferry Bay
Sample date	01-Sep 2011
Blood-feeding	no
Suspicious sequences	20

Table 6: Suspicious Sequences in INSnrTAIRAAPEI-21.

8 of 20 sequences were true positives, one is questionable and 11 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
C87378_a_10_0_l_1411	ORF_013	VP3, Eyach virus (KU163323, NC_003698)	26%	partial (start-mid)
(?) C88412_a_5_0_l_1508	ORF_003	1. Dendrolimus punctatus cypovirus (NC_025838) 2. hypothetical poly ADP-ribose glycohydrolases (XM_012814485, XM_014007017, XM_018466349)	36% 33%	partial (end) partial (end)
C91152_a_9_0_l_1820	ORF_003	segment 10, Eyach virus (NC_003705, KU163330)	25%	full
C91298_a_4_0_l_1840	ORF_005	VP2, Eyach virus (EU789380)	22%	partial (start)
C91402_a_6_0_l_1853	ORF_003	VP2, Eyach virus (EU789378, EU789378, KU163322, NC_003697)	30%	partial (end)
C92366_a_11_0_l_1992	ORF_004	segment 3, Eyach virus (EU789381, KU163323, NC_003698)	25%	partial (start-mid)
C94214_a_4_0_l_2312	ORF_005	segment 4, Eyach virus (KU163324, NC_003699)	30%	full
C98610_a_10_0_l_4248	ORF_016	RdRp , Eyach virus (NC_003696)	34%	full
s368_L_360_0_a_7_6_l_809	ORF_003	sigma-c like virus cell attachment protein Cangyuan orthoreovirus (NC_025806)	26%	partial (start)
s6362_L_12907_0_a_8_4_l_1869	ORF_004	segment 6, Dendrolimus punctatus cypovirus (NC_025850)	21%	full

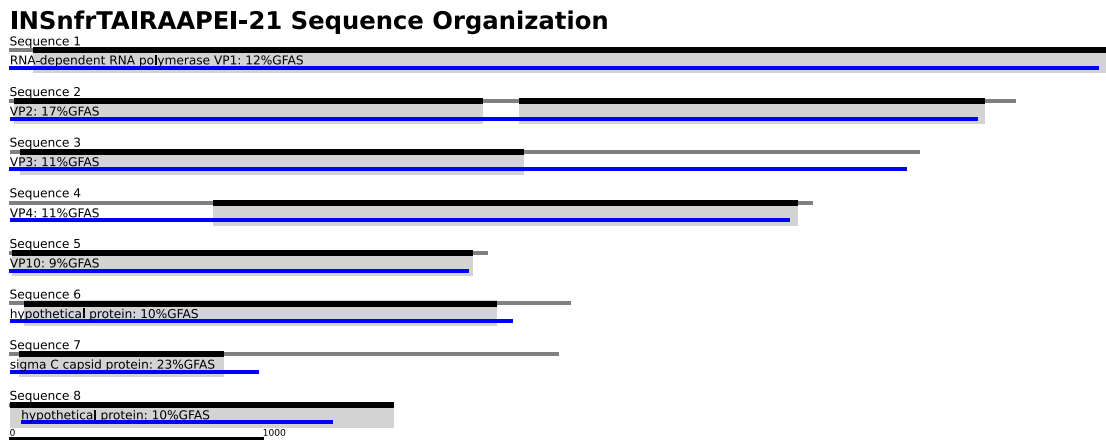


Figure 9: Sequence Organization of INSnfrTAIRAAPEI-21.

2.4 INSnrTAORAAPEI-35

In this transcriptome four small fragments of different segments have been detected. Three of those are potentially related to *Orthoptera Brumata Reovirus*.

Table 7: Sample Information of INSnrTAORAAPEI-35.

Filename	120126_I283_FCD0L80ACXX_L1_INSnrTAORAAPEI-35.free.fas
Assembly ID	INSnrTAORAAPEI-35
Order	Hemiptera
Order details	Heteroptera
Family	Veliidae
Family details	NA
Species	<i>Velia caprai</i>
Number of specimen	20
Stage	adult
Sample location	Germany, Lower Saxony, H/"ohbeck, Pevestorf
Sample date	12-Aug-2011
Blood-feeding	no
Suspicious sequences	13

Table 8: Suspicious Sequences in INSnrTAORAAPEI-35.

4 of 13 sequences were true positives and 9 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C46324_a_5_0_l_227	ORF_001	segment 3, Operophtera brumata reovirus (NC_007561)	33%	partial
C50775_a_3_0_l_247	ORF_001	1. segment 6, Dendrolimus punctatus cypovirus (NC_025850) 2. glycoprotein, Changping Tick Virus 2 (NC_028260)	33% 38%	partial partial
C67213_a_3_0_l_384	ORF_002	RdRp , Operophtera brumata reovirus (NC_007559)	42%	partial
C94877_a_63_0_l_1501	ORF_005	1. Hubei diptera virus 21 (KX884697) 2. segment 2, Operophtera brumata reovirus (NC_007560)	29% 28%	partial partial

INSnrTAORAAPEI-35 Sequence Organization

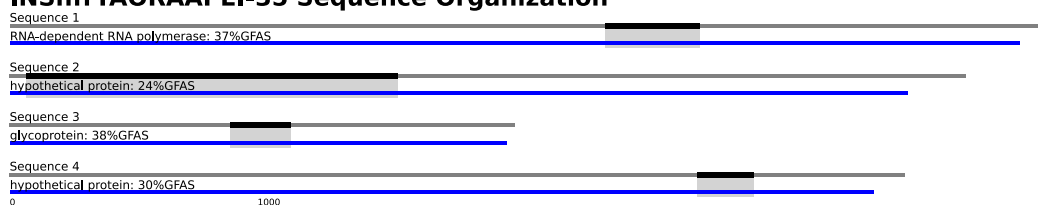


Figure 10: Sequence Organization of INSnrTAORAAPEI-35.

2.5 INSnfrTATRAAPEI-43

In this transcriptome only a small fragmentary RdRp-matching sequence has been indentified. It is matching several other RdRps at a similar location with about 50-55% identity on amino acid level and thus is considered a true positive.

Table 9: Sample Information of INSnfrTATRAAPEI-43.

Filename	120126_I283_FCD0L80ACXX_L1_INSnfrTATRAAPEI-43.free.fas
Assembly ID	INSnfrTATRAAPEI-43
Order	Neuroptera
Order details	NA
Family	Hemerobiidae
Family details	NA
Species	<i>Micromus variegatus</i>
Number of specimen	12
Stage	adult
Sample location	Germany, North Rhine-Westphalia, Bonn, Zoological Research Museum A Koenig ZFMK
Sample date	Jun-2011
Blood-feeding	no
Suspicious sequences	19

Table 10: Suspicious Sequences in INSnfrTATRAAPEI-43.

1 of 19 sequences was true positive and 18 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C94914_a_4_0_l_237	ORF_001	1. RdRp, Morris orbivirus (KX907618)	55%	partial (mid)

INSnfrTATRAAPEI-43 Sequence Organization



Figure 11: Sequence Organization of INSnfrTATRAAPEI-43.

2.6 INSnfrTBERRAPEI-19

In this transcriptome, two small fragments probably related to *Rice dwarf virus* have been detected.

Table 11: Sample Information of INSnfrTBERRAPEI-19.

Filename	120126_I283_FCD0L80ACXX_L2_INSnfrTBERRAPEI-19.free.fas
Assembly ID	INSnfrTBERRAPEI-19
Order	Coleoptera
Order details	NA
Family	Gyrinidae
Family details	NA
Species	<i>Gyrinus marinus</i>
Number of specimen	12
Stage	adult
Sample location	Germany, Lower Saxony, H6hbeck, Pevestorf
Sample date	11-Aug 2011
Blood-feeding	no
Suspicious sequences	12

Table 12: Suspicious Sequences in INSnfrTBERRAPEI-19.

2 of 12 sequences were true positives and 10 sequences were false positives similar to the false positives listed in chapter 3.2.2. Both true positive sequences were matching two distinct areas of reference RdRps suggesting they belong together.

Sequence ID	ORF	Match	Identity	Completeness
C92930_a_4_0_l_564	ORF_002	RdRp , Rice dwarf virus (NC_009248)	39%	partial (start)
C98326_a_8_0_l_737	ORF_001	RdRp , Rice dwarf virus (NC_009248)	33%	partial (mid)

INSnfrTBERRAPEI-19 Sequence Organization

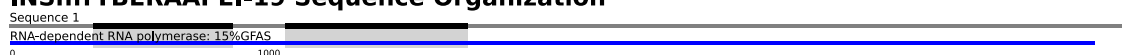


Figure 12: Sequence Organization of INSnfrTBERRAPEI-19.

2.7 INSyTvTABRAAPEI-11

Table 13: Sample Information of INSyTvTABRAAPEI-11.

Filename	120429_I266_FCC0HG0ACXX_L7_INSyTvTABRAAPEI-11.free.fas
Assembly ID	INSyTvTABRAAPEI-11
Order	Hymenoptera
Order details	NA
Family	Mutillidae
Family details	NA
Species	<i>Smicromyrme rufipes</i>
Number of specimen	21
Stage	adult
Sample location	Germany, Rhineland-Palatinate, Birkenheide
Sample date	22-May-2011
Blood-feeding	no
Suspicious sequences	23

Table 14: Suspicious Sequences in INSyTvTABRAAPEI-11.

2 of 23 sequences were questionable and 21 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C64493_a_4_0_l_495	ORF_001	RdRp, Hubei reo-like virus 14 (KX884702)	49%	partial (mid)
(?) C84625_a_5_0_l_1012	ORF_005	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	30%	partial (start)

INSyTvTABRAAPEI-11 Sequence Organization

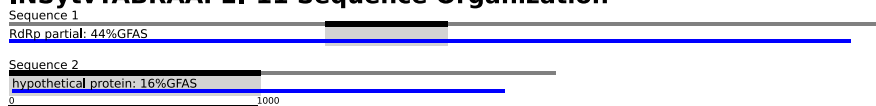


Figure 13: Sequence Organization of INSyTvTABRAAPEI-11.

2.8 INSyTvTALRAAPEI-35

In this transcriptome four full segments similar to *Hubei reo-like virus 6* have been identified.

Table 15: Sample Information of INSyTvTALRAAPEI-35.

Filename	120429_I266_FCC0HG0ACXX_L8_INSyTvTALRAAPEI-35.free.fas
Assembly ID	INSyTvTALRAAPEI-35
Order	Hemiptera
Order details	Sternorrhyncha
Family	Triozidae
Family details	NA
Species	<i>Acanthocasuarina muellerianae</i>
Number of specimen	few
Stage	not determined
Sample location	Australia, South Australia, Kangaroo Island, Sedden Conservation Park
Sample date	09-Feb-2012
Blood-feeding	no
Suspicious sequences	10

Table 16: Suspicious Sequences in INSyTvTALRAAPEI-35.

4 of 10 sequences were true positives, 1 questionable and 5 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
C203002_a_61_0_l_2753	ORF_002	hypothetical protein, Hubei reo-like virus 6 (KX884719)	26%	full
C203574_a_61_0_l_3410	ORF_003	hypothetical protein, Hubei reo-like virus 6 (KX884718)	26%	full
C203812_a_56_0_l_4203	ORF_001	RdRp , Hubei reo-like virus 6 (KX884716)	30%	full
C203826_a_57_0_l_4320	ORF_003	hypothetical protein, Hubei reo-like virus 6 (KX884717)	25%	full
(?) s10555_l_43618_0_a_47_0_l_7342	ORF_003	hypothetical protein, Dendrolimus punctatus cypovirus 22 (NC_025850)	25%	full

INSyTvTALRAAPEI-35 Sequence Organization



Figure 14: Sequence Organization of INSyTvTALRAAPEI-35.

2.9 INSyTvTBORAAPEI-47

In this transcriptome two fragments similar to *Bloomfield virus* have been identified. Also two small fragments similar to segment 6 of *Dendrolimus punctatus cypovirus 22* gave two questionable hits.

Table 17: Sample Information of INSyTvTBORAAPEI-47.

Filename	120521_I249_FCC0U4RACXX_L7_INSyTvTBORAAPEI-47.free.fas
Assembly ID	INSyTvTBORAAPEI-47
Order	Hymenoptera
Order details	NA
Family	Melittidae
Family details	NA
Species	<i>Macropis fulvipes</i>
Number of specimen	8
Stage	adult
Sample location	Germany, Rhineland-Palatinate, Albersweiler
Sample date	2011
Blood-feeding	no
Suspicious sequences	15

Table 18: Suspicious Sequences in INSyTvTBORAAPEI-47.

2 of 15 sequences were true positives, 2 were questionable and 11 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C31726_a_6_0_l_364	ORF_002	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	36%	partial (end)
(?) C39651_a_7_0_l_539	ORF_005	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	35%	partial (mid)
C46583_a_3_0_l_801	ORF_002	RdRp , <i>Bloomfield virus</i> (KP714090)	31%	partial (end)
C62529_a_11_0_l_2839	ORF_007	RdRp , <i>Bloomfield virus</i> (KP714090)	35%	partial (start-mid)

INSyTvTBORAAPEI-47 Sequence Organization

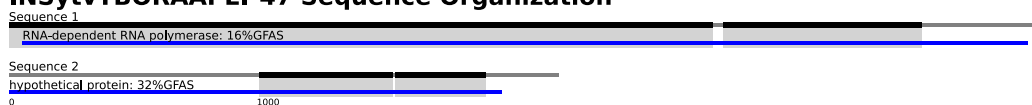


Figure 15: Sequence Organization of INSyTvTBORAAPEI-47.

2.10 INSswpTBBRAAPEI-21

This transcriptome contained a large part of the RdRp segment similar to *Hubei reo-like virus 14*.

Table 19: Sample Information of INSswpTBBRAAPEI-21.

Filename	120707_I249_FCD111GACXX_L3_INSswpTBBRAAPEI-21.free.fas
Assembly ID	INSswpTBBRAAPEI-21
Order	Hymenoptera
Order details	NA
Family	Apidae
Family details	NA
Species	<i>Epeolus variegatus</i>
Number of specimen	3
Stage	adult
Sample location	Italy, Sardinia, SW Santa Teresa Gallura
Sample date	06-Sep-2011
Blood-feeding	no
Suspicious sequences	14

Table 20: Suspicious Sequences in INSswpTBBRAAPEI-21.

1 of 14 sequences was true positive and 13 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C95104_a_11_0_l_3069	ORF_012	RdRp, Hubei reo-like virus 14 (KX884607)	37%	partial (start-mid)

INSswpTBBRAAPEI-21 Sequence Organization



Figure 16: Sequence Organization of INSswpTBBRAAPEI-21.

2.11 INSeqtTAHRAAPEI-88

This transcriptome contained many small ORFs similar to segment 6 of *Dendrolimus punctatus cypovirus 22* but no true positive segment similar to an RdRp.

Table 21: Sample Information of INSeqtTAHRAAPEI-88.

Filename	121010_I249_FCD1C4BACXX_L6_INSeqtTAHRAAPEI-88.free.fas
Assembly ID	INSeqtTAHRAAPEI-88
Order	Hymenoptera
Order details	NA
Family	Chalcididae
Family details	NA
Species	<i>Brachymeria minuta</i>
Number of specimen	4
Stage	adult
Sample location	Germany, Hessen, Osthofen
Sample date	28-Jun-2012
Blood-feeding	no
Suspicious sequences	42

Table 22: Suspicious Sequences in INSeqtTAHRAAPEI-88.

16 of 42 sequences were questionable and 26 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C106753_a_3_0_l_390	ORF_002	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	30%	partial (start)
(?) C113935_a_11_0_l_454	ORF_002	segment 8, Kadipiro virus chromosome (NC_004208)	22%	partial (start)
(?) C117709_a_23_0_l_494	ORF_003	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	31%	partial (mid)
(?) C127403_a_4_0_l_631	ORF_002	hypothetical protein, Hubei diptera virus 21 (KX884697)	38%	partial (mid)
(?) C127695_a_3_0_l_636	ORF_003	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	29%	partial (end)
(?) C133384_a_8_0_l_761	ORF_005	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	26%	partial (start)
(?) C136012_a_8_0_l_834	ORF_001	hypothetical protein, Hubei diptera virus 21 (KX884697)	34%	partial (mid)
(?) C142742_a_10_0_l_1115	ORF_003	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	26%	partial (end)
(?) C144414_a_9_0_l_1213	ORF_006	hypothetical protein, Hubei diptera virus 21 (KX884697)	33%	partial (mid)
(?) C68927_a_4_0_l_205	ORF_001	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	27%	partial (mid)
(?) C76527_a_6_0_l_233	ORF_001	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	27%	partial (start)
(?) C77495_a_5_0_l_235	ORF_001	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	38%	partial (start)
(?) C82049_a_3_0_l_246	ORF_001	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	40%	partial (mid)
(?) C90211_a_17_0_l_280	ORF_001	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	43%	partial (mid)
(?) C91643_a_21_0_l_288	ORF_001	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	34%	partial (end)
(?) C95097_a_3_0_l_310	ORF_003	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	37%	partial (start)

INSeqtTAHRAAPEI-88 Sequence Organization

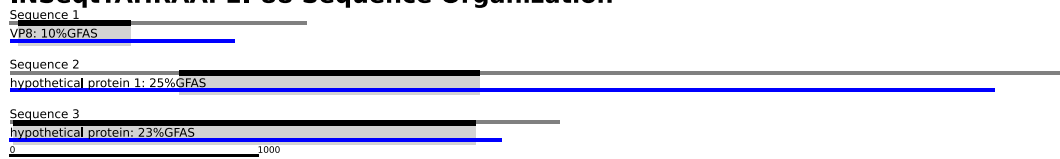


Figure 17: Sequence Organization of INSeqtTAHRAAPEI-88.

2.12 INShkeTCLRAAPEI-44

Table 23: Sample Information of INShkeTCLRAAPEI-44.

Filename	121014_I189_FCC173YACXX_L1_INShkeTCLRAAPEI-44.free.fas
Assembly ID	INShkeTCLRAAPEI-44
Order	Blattodea
Order details	NA
Family	Blattidae
Family details	NA
Species	<i>Deropeltis erythrocephala</i>
Number of specimen	1
Stage	adult
Sample location	Germany, Lab culture with Samples originating from Germany, private breeder
Sample date	13-Mar-2011
Blood-feeding	no
Suspicious sequences	6

this sequence is questionable because it's not an ongoing ORF

Table 24: Suspicious Sequences in INShkeTCLRAAPEI-44.

1 of 6 sequences was questionable and 5 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) s2137_L_2837_0_a_4_2_1_945	ORF_003	RdRp, Yunnan orbivirus (NC_007656)	31%	partial (mid)

INShkeTCLRAAPEI-44 Sequence Organization

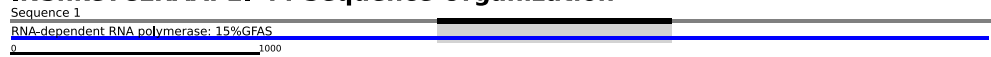


Figure 18: Sequence Organization of INShkeTCLRAAPEI-44.

2.13 INSeqtTBNRAAPEI-11

This transcriptome contained several near full segments similar to *Rice dwarf virus*.

Table 25: Sample Information of INSeqtTBNRAAPEI-11.

Filename	121030_I251_FCC19KWACXX_L1_INSeqtTBNRAAPEI-11.free.fas
Assembly ID	INSeqtTBNRAAPEI-11
Order	Odonata
Order details	Zygoptera
Family	Lestidae
Family details	NA
Species	<i>Indolestes peregrinus</i>
Number of specimen	2
Stage	adult
Sample location	Japan, Nagano, Ueda
Sample date	27-May-2012
Blood-feeding	no
Suspicious sequences	13

Table 26: Suspicious Sequences in INSeqtTBNRAAPEI-11.

6 of 13 sequences were true positives and 7 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C106661_a_58_0_l_2619	ORF_001	minor core structural protein, Rice dwarf virus (NC_009247)	30%	full
C107435_a_53_0_l_3182	ORF_015	major core structural protein, Rice dwarf virus (NC_009243)	21%	full
C107553_a_48_0_l_3308	ORF_007	RdRp , Rice dwarf virus (NC_003773)	31%	partial (mid-end)
C107715_a_60_0_l_3560	ORF_001	minor core structural protein, Rice dwarf virus (NC_003774)	25%	full
C91752_a_32_0_l_905	ORF_002	RdRp , Rice dwarf virus (NC_003773)	29%	partial (start)
C96118_a_13_0_l_1114	ORF_001	nonstructural protein, Rice dwarf virus (NC_003766)	24%	full

INSeqtTBNRAAPEI-11 Sequence Organization

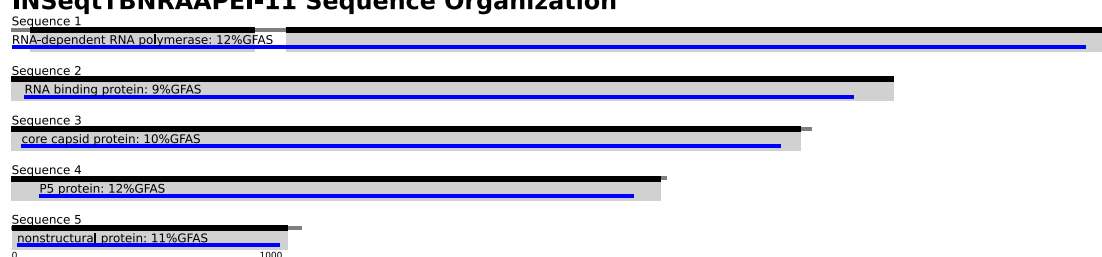


Figure 19: Sequence Organization of INSeqtTBNRAAPEI-11.

2.14 INSeqtTCJRAAPEI-20

This transcriptome contained several partial fragments similar to *Bloomfield virus*.

Table 27: Sample Information of INSeqtTCJRAAPEI-20.

Filename	121030_I251_FCC19KWACXX_L1_INSeqtTCJRAAPEI-20.free.fas
Assembly ID	INSeqtTCJRAAPEI-20
Order	Raphidioptera
Order details	NA
Family	Raphidiidae
Family details	NA
Species	<i>Ornatoraphidia flavilabris</i>
Number of specimen	3
Stage	larva
Sample location	Austria, Lab culture with Samples originating from Greece, Arcadia Mainalon, Mountains Levidi, Kardaras
Sample date	20-May-2012
Blood-feeding	no
Suspicious sequences	33

Table 28: Suspicious Sequences in INSeqtTCJRAAPEI-20.

5 of 33 sequences were true positives, 1 questionable and 27 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
C234593_a_7_0_l_724	ORF_003	RdRp , Bloomfield virus (KP714090)	34%	partial (end)
C251492_a_47_0_l_2048	ORF_005	minor core structural protein, Bloomfield virus (KP714094)	24%	partial (mid-end)
(?) C251660_a_4_0_l_2101	ORF_003	1. glycoprotein, Lishi Spider Virus 1 (KM817596) 2. hypothetical protein, Dendrolimus punctatus cypovirus 22 (NC_025850)	34% 28%	partial (start) partial (start)
C252938_a_26_0_l_2905	ORF_007	RdRp , Bloomfield virus (KP714090)	36%	partial (start-mid)
C253318_a_53_0_l_3951	ORF_009	major core protein, Bloomfield virus (KP714091)	26%	full
s15123_L_40031_0_a_13_9_l_1642	ORF_001	glycoprotein, Wuchang Cockroach Virus 3 (KM817605)	34%	partial (mid-end)

INSeqtTCJRAAPEI-20 Sequence Organization

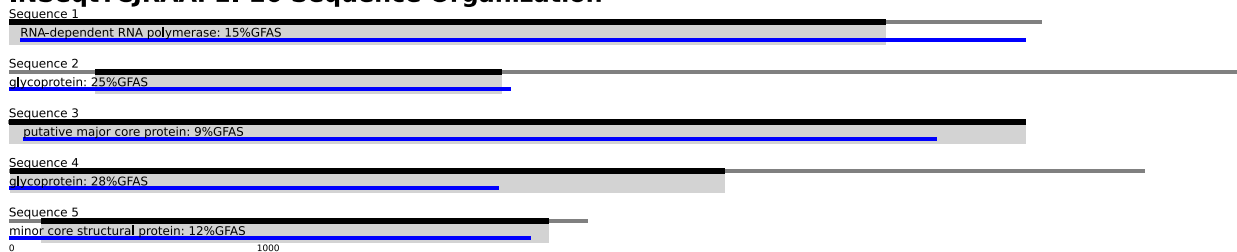


Figure 20: Sequence Organization of INSeqtTCJRAAPEI-20.

2.15 INSeqTCZRAAPEI-47

This transcriptome contained several small fragments similar to *Liao ning virus*. However, a near full RdRp-segment could be assembled by three fragments.

Table 29: Sample Information of INSeqTCZRAAPEI-47.

Filename	121030_I251_FCC19KWACXX_L1_INSeqTCZRAAPEI-47.free.fas
Assembly ID	INSeqTCZRAAPEI-47
Order	Plecoptera
Order details	NA
Family	Nemouridae
Family details	Nemourinae
Species	<i>Protonemura ausonia</i>
Number of specimen	ca 7
Stage	nymph
Sample location	Italy, Viterbo, Arlena River
Sample date	01-Mar-2007
Blood-feeding	no
Suspicious sequences	26

Table 30: Suspicious Sequences in INSeqTCZRAAPEI-47.

12 of 26 sequences were true positives, 3 were questionable and 11 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C145258_a_4_0_l_204	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	51%	partial (mid)
C148370_a_3_0_l_209	ORF_001	segment 2, Liao ning virus (AY317100)	43%	partial (mid)
C194142_a_3_0_l_243	ORF_001	segment 2, Liao ning virus (NC_007737)	29%	partial (mid)
C266489_a_15_0_l_406	ORF_003	segment 5, Liao ning virus (NC_007740)	33%	partial (end)
C267864_a_3_0_l_412	ORF_001	segment 2, Liao ning virus (NC_007737)	33%	partial (mid)
C269215_a_3_0_l_418	ORF_001	segment 4, Liao ning virus (NC_007739)	35%	partial (start)
C269977_a_6_0_l_422	ORF_002	RdRp , Liao ning virus (NC_007736)	40%	partial (end)
(?) C280721_a_11_0_l_481	ORF_002	segment 8, Kadipiro virus (NC_004208)	28%	partial (start)
(?) C305375_a_8_0_l_784	ORF_001	1. segment 11, Liao ning virus (NC_007746) 2. R2D2, Nilaparvata lugens (KC316044)	33% 44%	partial (start) partial(start)
C307017_a_6_0_l_823	ORF_001	segment 3, Liao ning virus (NC_007738)	37%	partial (end)
C311087_a_10_0_l_946	ORF_003	segment 5, Liao ning virus (NC_007740)	27%	partial (start)
C315239_a_6_0_l_1127	ORF_002	segment 3, Liao ning virus (NC_007738)	46%	partial (start)
C315997_a_5_0_l_1176	ORF_001	RdRp , Liao ning virus (NC_007736)	45%	partial (mid)
C319123_a_7_0_l_1437	ORF_004	segment 6, Liao ning virus (NC_007741)	25%	partial (start-mid)
C321585_a_7_0_l_1930	ORF_010	RdRp , Liao ning virus (NC_007736)	34%	partial (start-mid)

INSeqTCZRAAPEI-47 Sequence Organization

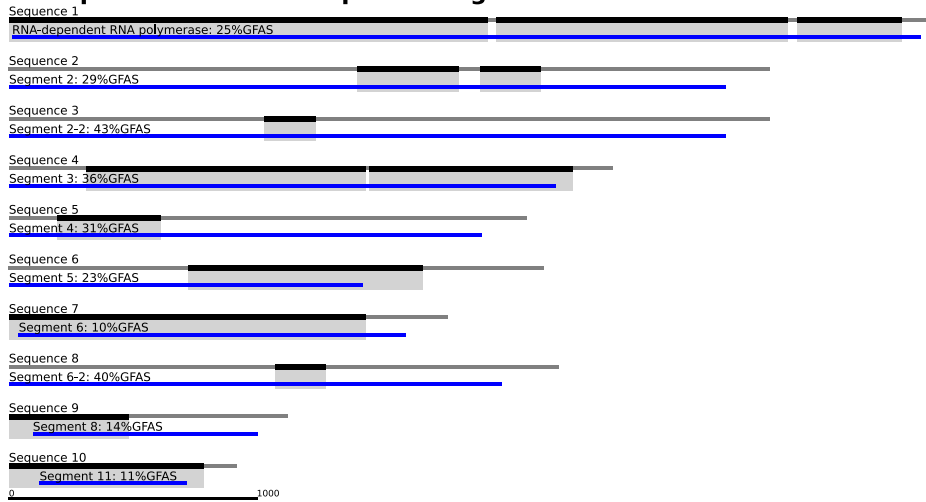


Figure 21: Sequence Organization of INSeqTCZRAAPEI-47.

2.16 INSeqTDXRAAPEI-19

This transcriptome contained a partial segment similar to *Hubei diptera virus 21* and a very small fragment of an RdRp similar to the same virus.

Table 31: Sample Information of INSeqTDXRAAPEI-19.

Filename	121030_I251_FCC19KWACXX_L3_INSeqTDXRAAPEI-19.free.fas
Assembly ID	INSeqTDXRAAPEI-19
Order	Neuroptera
Order details	NA
Family	Hemerobiidae
Family details	NA
Species	<i>Hemerobius marginatus</i>
Number of specimen	3
Stage	adult
Sample location	Austria, Lower Austria, Vienna-Surroundings, Klosterneuburg
Sample date	May-2012
Blood-feeding	no
Suspicious sequences	14

Table 32: Suspicious Sequences in INSeqTDXRAAPEI-19.

3 of 14 sequences were true positives, 1 questionable and 10 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
C102680_a_3_0_l_570	ORF_001	hypothetical protein 1, Hubei diptera virus 21 (KX884697)	46%	partial (mid)
C109650_a_7_0_l_734	ORF_001	hypothetical protein 1, Hubei diptera virus 21 (KX884697)	45%	partial (mid)
(?) C83898_a_5_0_l_334	ORF_001	RdRp, Hubei diptera virus 21 (KX884696)	35%	partial (end)
C89115_a_3_0_l_381	ORF_002	hypothetical protein 1, Hubei diptera virus 21 (KX884697)	43%	partial (start)

INSeqTDXRAAPEI-19 Sequence Organization

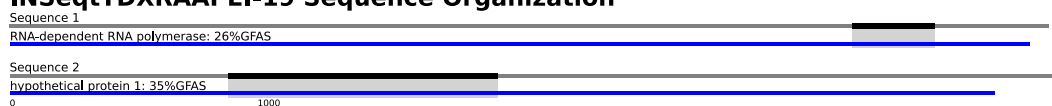


Figure 22: Sequence Organization of INSeqTDXRAAPEI-19.

2.17 INSlupTBDRAAPEI-17

This transcriptome contained a near full RdRp-segment similar to *Hubei reo-like virus 14*.

Table 33: Sample Information of INSlupTBDRAAPEI-17.

Filename	121221_I260_FCC1GFFACXX_L3_INSlupTBDRAAPEI-17.free.fas
Assembly ID	INSlupTBDRAAPEI-17
Order	Hymenoptera
Order details	NA
Family	Crabronidae
Family details	NA
Species	<i>Mellinus arvensis</i>
Number of specimen	2
Stage	adult
Sample location	Germany, Rhineland-Palatinate, Osthofen
Sample date	21-Jun-2012
Blood-feeding	no
Suspicious sequences	19

Table 34: Suspicious Sequences in INSlupTBDRAAPEI-17.

1 of 19 sequences was true positive and 18 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
s5948_L_19400_0_a_33_1_1_3436	ORF_022	RdRp, Hubei reo-like virus 14 (KX884702)	39%	full

INSlupTBDRAAPEI-17 Sequence Organization



Figure 23: Sequence Organization of INSlupTBDRAAPEI-17.

2.18 INSlupTBKRAAPEI-31

This transcriptome contained several near full segments similar to *Rice black streaked dwarf virus*.

Table 35: Sample Information of INSlupTBKRAAPEI-31.

Filename	121221_I260_FCC1GFFACXX_L3_INSlupTBKRAAPEI-31.free.fas
Assembly ID	INSlupTBKRAAPEI-31
Order	Hemiptera
Order details	Auchenorrhyncha Fulgoromorpha
Family	Cixiidae
Family details	NA
Species	<i>Tachycixius pilosus</i>
Number of specimen	6
Stage	adult
Sample location	Germany, Thuringia, Jena
Sample date	May-2012
Blood-feeding	no
Suspicious sequences	14

Table 36: Suspicious Sequences in INSlupTBKRAAPEI-31.

6 of 14 sequences were true positives and 8 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C136255_a_5_0_l_1616	ORF_001	segment 10, Rice black streaked dwarf virus (NC_003733)	23%	full
C136503_a_9_0_l_1696	ORF_008	segment 7, Rice black streaked dwarf virus (NC_003730)	25%	partial (start-mid)
C137679_a_54_0_l_3342	ORF_005	core capsid protein, Rice black streaked dwarf virus (NC_003728)	23%	full
C137715_a_25_0_l_3809	ORF_003	major core protein, Rice black streaked dwarf virus (NC_003734)	21%	full
s6698_L_26721_0_a_19_3_l_3499	ORF_001	segment 4, Rice black streaked dwarf virus (NC_003735)	18%	full
s7010_L_30471_0_a_63_6_l_4223	ORF_011	RdRp , Rice black streaked dwarf virus (NC_003729)	34%	full

INSlupTBKRAAPEI-31 Sequence Organization

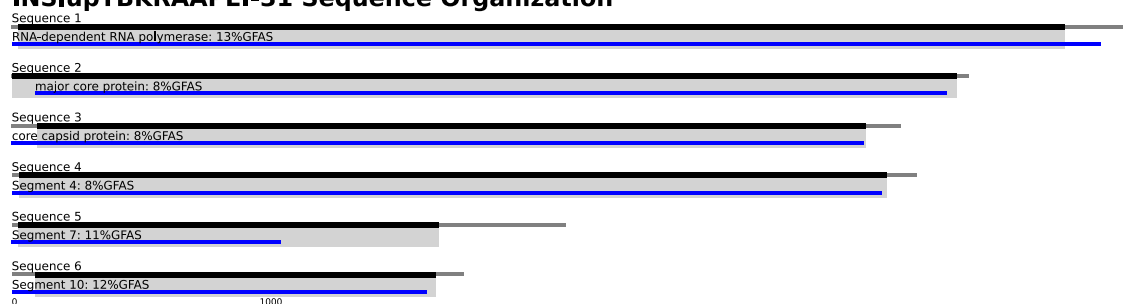


Figure 24: Sequence Organization of INSlupTBKRAAPEI-31.

2.19 INSlupTBMRAAPEI-34

This transcriptome contained three near full segments similar to *Equine encephalitis virus*.

Table 37: Sample Information of INSlupTBMRAAPEI-34.

Filename	121221_I260_FCC1GFFACXX_L3_INSlupTBMRAAPEI-34.free.fas
Assembly ID	INSlupTBMRAAPEI-34
Order	Hemiptera
Order details	Heteroptera
Family	Aphelocheiridae
Family details	NA
Species	<i>Aphelocheirus aestivalis</i>
Number of specimen	3
Stage	adult
Sample location	Germany, Thuringia, Maua
Sample date	01-Aug-2012
Blood-feeding	no
Suspicious sequences	14

Table 38: Suspicious Sequences in INSlupTBMRAAPEI-34.

3 of 14 sequences were true positives and 11 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C148506_a_38_0_l_2007	ORF_003	segment 4, Equine encephalosis virus (AB811638)	29%	full
C149424_a_61_0_l_3917	ORF_003	RdRp, Equine encephalosis virus (AB811635)	26%	full
s7155_L_42109_0_a_61_0_l_3092	ORF_006	segment 3, Equine encephalosis virus (AB811637)	21%	full

INSlupTBMRAAPEI-34 Sequence Organization

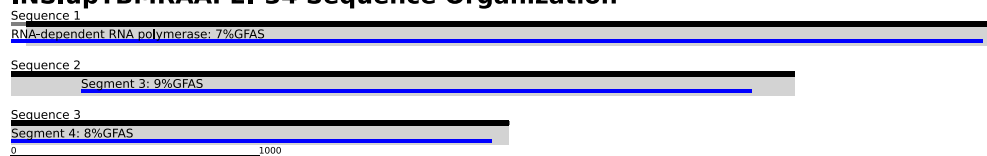


Figure 25: Sequence Organization of INSlupTBMRAAPEI-34.

2.20 INSlupTBURAAPEI-45

This transcriptome contained a small fragment of an RdRp similar to *Reptilian orthoreovirus*.

Table 39: Sample Information of INSlupTBURAAPEI-45.

Filename	121221_I260_FCC1GFFACXX_L8_INSlupTBURAAPEI-45.free.fas
Assembly ID	INSlupTBURAAPEI-45
Order	Mantodea
Order details	NA
Family	Hymenopodidae
Family details	NA
Species	<i>Acromantis sp</i>
Number of specimen	1
Stage	adult
Sample location	Germany, Lab culture with Samples originating from Hong Kong
Sample date	Aug-2012
Blood-feeding	no
Suspicious sequences	8

Table 40: Suspicious Sequences in INSlupTBURAAPEI-45.

1 of 8 sequences was questionable and 7 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C142720_a_3_0_l_484	ORF_001	RdRp, Reptilian orthoreovirus (KT696549)	31%	partial (mid)

INSlupTBURAAPEI-45 Sequence Organization



Figure 26: Sequence Organization of INSlupTBURAAPEI-45.

2.21 INSlupTAFRAAPEI-44

This transcriptome contained a small fragment that might be similar to an RdRp of *Rice black streaked dwarf virus* and two other segments that are in relation to *Liao ning virus* and *Umatilla virus*.

Table 41: Sample Information of INSlupTAFRAAPEI-44.

Filename	130112_I269_FCC1M19ACXX_L8_INSlupTAFRAAPEI-44.free.fas
Assembly ID	INSlupTAFRAAPEI-44
Order	Hemiptera
Order details	Auchenorrhyncha Cicadomorpha
Family	Aphrophoridae
Family details	NA
Species	<i>Aphrophora alni</i>
Number of specimen	ca 7
Stage	adult
Sample location	Germany, Thuringia, Jena
Sample date	30-Jul-2012
Blood-feeding	no
Suspicious sequences	12

Table 42: Suspicious Sequences in INSlupTAFRAAPEI-44.

1 of 12 sequences was true positive, 3 questionable and 8 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C116451_a_8_0_l_385	ORF_001	RdRp , Rice black streaked dwarf virus (NC_003729)	42%	partial (mid)
(?) C127494_a_7_0_l_494	ORF_003	1. segment 11, Liao ning virus (NC_007746) 2. transcribed RNA sequence, Homalodisca liturata (GECU01029485)	29% 54%	partial (start-mid) partial (start-mid)
C137751_a_3_0_l_666	ORF_002	segment 5, Umatilla virus (NC_024507)	28%	partial (end)
(?) C139507_a_7_0_l_712	ORF_004	1. segment 11, Liao ning virus (NC_007746) 2. transcribed RNA sequence, Homalodisca liturata (GECU01023362)	34% 85%	partial (start) partial (start-mid)

INSlupTAFRAAPEI-44 Sequence Organization

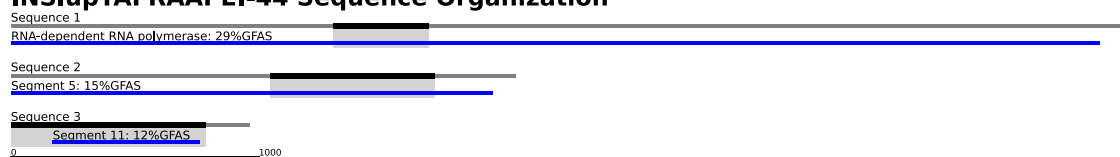


Figure 27: Sequence Organization of INSlupTAFRAAPEI-44.

2.22 INSntgTABRAAPEI-216

This transcriptome contained a fragment similar to the RdRp of *Hubei reo-like virus 6*.

Table 43: Sample Information of INSntgTABRAAPEI-216.

Filename	130125_I266_FCC1MY6ACXX_L3_INSntgTABRAAPEI-216.free.fas
Assembly ID	INSntgTABRAAPEI-216
Order	Neuroptera
Order details	NA
Family	Coniopterygidae
Family details	NA
Species	<i>Coniopteryx sp</i>
Number of specimen	10
Stage	adult
Sample location	Austria, Lower Austria, Krems-Land, Duernstein
Sample date	27-May-2012
Blood-feeding	no
Suspicious sequences	12

Table 44: Suspicious Sequences in INSntgTABRAAPEI-216.

1 of 12 sequences was questionable and 1 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C119964_a_3_0_l_381	ORF_002	RdRp , Hubei reo-like virus 6 (KX884716)	47%	partial (mid)

INSntgTABRAAPEI-216 Sequence Organization



Figure 28: Sequence Organization of INSntgTABRAAPEI-216.

2.23 INSlupTASRAAPEI-89

This transcriptome contained several near full segments of a virus similar to *Fji disease virus*.

Table 45: Sample Information of INSlupTASRAAPEI-89.

Filename	130206_I238_FCC1LVUACXX_L1_INSlupTASRAAPEI-89.free.fas
Assembly ID	INSlupTASRAAPEI-89
Order	Hemiptera
Order details	Auchenorrhyncha Fulgoromorpha
Family	Dictyopharidae
Family details	NA
Species	<i>Dictyophara europaea</i>
Number of specimen	ca 10
Stage	adult
Sample location	Germany, Thuringia, Jena
Sample date	31-Jul-2012
Blood-feeding	no
Suspicious sequences	12

Table 46: Suspicious Sequences in INSlupTASRAAPEI-89.

8 of 12 sequences were true positives and 4 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C233648_a_60_0_l_1298	ORF_002	segment 7, Fiji disease virus (NC_007163)	25%	full
C234606_a_26_0_l_1429	ORF_006	segment 9, Fiji disease virus (NC_007156)	27%	full
C236042_a_45_0_l_1734	ORF_004	major outer capsid protein, Fiji disease virus (NC_007162)	22%	full
C236682_a_38_0_l_1998	ORF_004	segment 6, Fiji disease virus (NC_007161)	24%	full
C237738_a_43_0_l_4257	ORF_001	RdRp , Fiji disease virus (NC_007159)	33%	full
s11280_l_56553_0_a_56_1_l_3431	ORF_005	segment 4, Fiji disease virus (NC_007155)	23%	full
s11335_l_57448_0_a_64_5_l_3564	ORF_005	segment 3, Fiji disease virus (NC_007158)	25%	full
s11389_l_58703_0_a_46_0_l_3754	ORF_001	segment 2, Fiji disease virus (NC_007154)	20%	full

INSlupTASRAAPEI-89 Sequence Organization

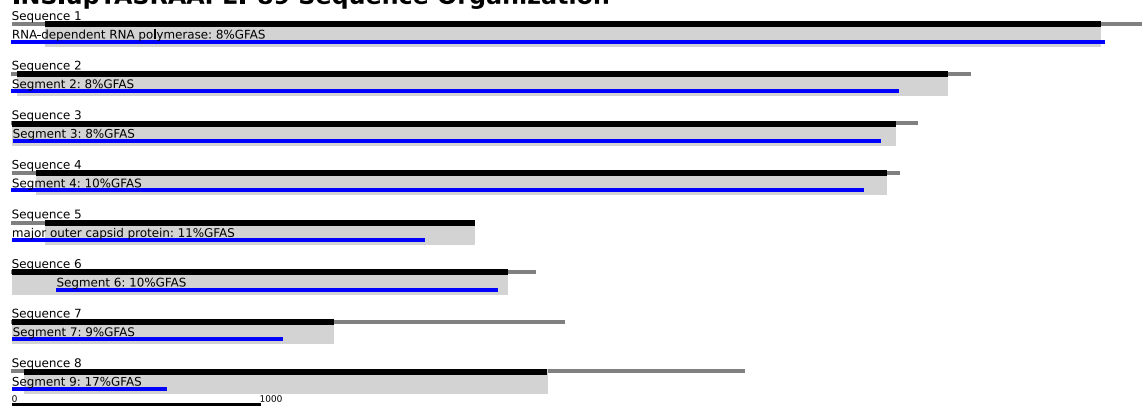


Figure 29: Sequence Organization of INSlupTASRAAPEI-89.

2.24 INSqiqTBFRAAPEI-61

This transcriptome contained a partial RdRp similar to the one of *Hubei tetragrathia maxillosa virus*.

Table 47: Sample Information of INSqiqTBFRAAPEI-61.

Filename	130206_I238_FCC1LVUACXX_L1_INSqiqTBFRAAPEI-61.free.fas
Assembly ID	INSqiqTBFRAAPEI-61
Order	Hymenoptera
Order details	NA
Family	Chrysididae
Family details	Chrysidinae
Species	<i>Praestochrysis megerlei</i>
Number of specimen	2
Stage	adult
Sample location	Italy, Emilia-Romagna, Parma Oriano
Sample date	29-Aug-2012
Blood-feeding	no
Suspicious sequences	11

Table 48: Suspicious Sequences in INSqiqTBFRAAPEI-61.

4 of 13 sequences were true positives and 9 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
s7034_L_44580_0_a_6_7_I_2092	ORF_005	RdRp, Hubei tetragrathia maxillosa virus 9 (KX884675)	30%	partial (mid-end)

INSqiqTBFRAAPEI-61 Sequence Organization



Figure 30: Sequence Organization of INSqiqTBFRAAPEI-61.

2.25 INSqiqTBLRAAPEI-83

This transcriptome contained a near full RdRp similar to the one of *Hubei tetragratha maxillosa virus* and two other segments similar to *Dendrolimus punctatus cypovirus 22* and *Wuchang Cockroach virus*.

Table 49: Sample Information of INSqiqTBLRAAPEI-83.

Filename	130206_I238_FCC1LVUACXX_L2_INSqiqTBLRAAPEI-83.free.fas
Assembly ID	INSqiqTBLRAAPEI-83
Order	Neuroptera
Order details	NA
Family	Myrmeleontidae
Family details	NA
Species	<i>Myrmeleon formicarius</i>
Number of specimen	1
Stage	adult
Sample location	Japan, Ibaraki, Tsukuba
Sample date	25-Jun-2012
Blood-feeding	no
Suspicious sequences	20

Table 50: Suspicious Sequences in INSqiqTBLRAAPEI-83.

3 of 20 sequences were true positives and 17 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C101877_a_3_0_I_1449	ORF_008	segment 6, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025850)	32%	full
C105769_a_8_0_I_2999	ORF_010	glycoprotein, <i>Wuchang Cockroach Virus 3</i> (KM817605)	31%	full
C106019_a_29_0_I_3543	ORF_011	RdRp , <i>Hubei tetragratha maxillosa virus 9</i> (KX884675)	35%	full

INSqiqTBLRAAPEI-83 Sequence Organization

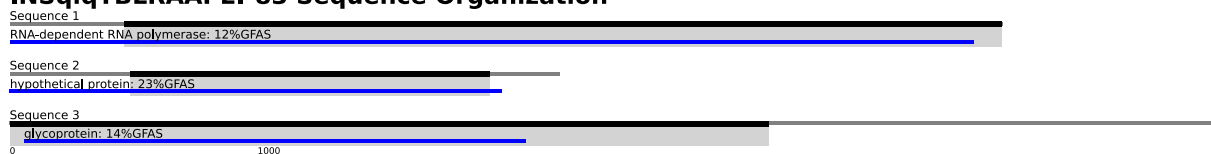


Figure 31: Sequence Organization of INSqiqTBLRAAPEI-83.

2.26 INSqiqTBNRABPEI-90

This transcriptome contained three fragments similar to the RdRp of *Hubei diptera virus 21* as well as several fragments similar to other segments of the same virus. Additionally, some fragments similar to different segments of *Dendrolimus punctatus cypovirus* could have been identified.

Table 51: Sample Information of INSqiqTBNRABPEI-90.

Filename	130206_I238_FCC1LVUACXX_L3_INSqiqTBNRABPEI-90.free.fas
Assembly ID	INSqiqTBNRABPEI-90
Order	Hymenoptera
Order details	NA
Family	Torymidae
Family details	NA
Species	<i>Bootanomyia dorsalis</i>
Number of specimen	ca 20
Stage	adult
Sample location	Germany, Baden-Wuerttemberg, Stuttgart, Rosensteinpark
Sample date	19-Jul-2012
Blood-feeding	no
Suspicious sequences	42

INSqiqTBNRABPEI-90 Sequence Organization



Figure 32: Sequence Organization of INSqiqTBNRABPEI-90.

Table 52: Suspicious Sequences in INSqiqTBNRABPEI-90.

6 of 42 sequences were true positives, 21 questionable and 15 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C32191_a_3_0_l_222	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	28%	partial (mid)
(?) C32667_a_10_0_l_226	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	25%	partial (mid)
(?) C35078_a_3_0_l_243	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	36%	partial (mid)
(?) C38974_a_5_0_l_2633	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (start)
(?) C39945_a_4_0_l_267	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (end)
(?) C40125_a_3_0_l_269	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (start)
(?) C43181_a_6_0_l_289	ORF_001	hypothetical protein 1, Hubei diptera virus 21 (KX884697)	42%	partial (mid)
C46563_a_38_0_l_319	ORF_001	RdRp , Hubei diptera virus 21 (KX884696)	35%	partial (start)
(?) C46813_a_3_0_l_321	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	29%	partial (end)
(?) C46971_a_6_0_l_323	ORF_002	hypothetical protein 2, Hubei diptera virus 21 (KX884698)	51%	partial (mid)
(?) C48774_a_7_0_l_339	ORF_001	hypothetical protein 1, Hubei diptera virus 21 (KX884697)	49%	partial (mid)
C51752_a_4_0_l_370	ORF_003	RdRp , Hubei diptera virus 21 (KX884696)	53%	partial (mid)
(?) C54619_a_3_0_l_401	ORF_003	segment 10, Dendrolimus punctatus cypovirus 22 (NC_025838)	35%	partial (mid-end)
(?) C55252_a_9_0_l_410	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	23%	partial (end)
(?) C57419_a_8_0_l_440	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	21%	partial (start)
C59361_a_3_0_l_469	ORF_001	RdRp , Hubei diptera virus 21 (KX884696)	53%	partial (mid)
(?) C60811_a_3_0_l_493	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	44%	partial (mid)
(?) C61091_a_5_0_l_496	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	26%	partial (start)
(?) C61837_a_4_0_l_509	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	26%	partial (mid)
(?) C63597_a_6_0_l_539	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (end)
(?) C69161_a_3_0_l_666	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	28%	partial (mid)
(?) C69279_a_4_0_l_670	ORF_005	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (mid)
(?) C70559_a_5_0_l_703	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	33%	partial (mid)
C74861_a_6_0_l_852	ORF_001	hypothetical protein 1, Hubei diptera virus 21 (KX884697)	55%	partial (mid)
C84762_a_5_0_l_1517	ORF_006	hypothetical protein 5, Hubei diptera virus 21 (KX884701)	47%	full
s3488_l_10393_0_a_6_4_l_545	ORF_001	hypothetical protein 2, Hubei diptera virus 21 (KX884698)	66%	partial (end)
(?) s4274_l_16530_0_a_3_6_l_882	ORF_001	segment 10, Dendrolimus punctatus cypovirus 22 (NC_025838)	32%	partial (mid-end)

2.27 INSqiqTCTRAAPEI-75

Table 53: Sample Information of INSqiqTCTRAAPEI-75.

Filename	130206_I238_FCC1LVUACXX_L3_INSqiqTCTRAAPEI-75.free.fas
Assembly ID	INSqiqTCTRAAPEI-75
Order	Mantodea
Order details	NA
Family	Mantidae
Family details	NA
Species	<i>Omomantis zebrata</i>
Number of specimen	1
Stage	juvenile
Sample location	Germany, Lab culture with Samples originating from South Africa
Sample date	Dec-2012
Blood-feeding	no
Suspicious sequences	6

Table 54: Suspicious Sequences in INSqiqTCTRAAPEI-75.

1 of 6 sequences was true positive and 5 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C137185_a_7_0_l_1730	ORF_004	RdRp, Fiji disease virus (NC_007159)	25%	partial (start)

INSqiqTCTRAAPEI-75 Sequence Organization



Figure 33: Sequence Organization of INSqiqTCTRAAPEI-75.

2.28 INSlupTATRAAPEI-90

This transcriptome contained a near full RdRp similar to the one of *Hubei diptera virus 21* and three other segments similar to the same virus.

Table 55: Sample Information of INSlupTATRAAPEI-90.

Filename	130206_I238_FCC1LVUACXX_L1_INSlupTATRAAPEI-90.free.fas
Assembly ID	INSlupTATRAAPEI-90
Order	Hymenoptera
Order details	NA
Family	Torymidae
Family details	NA
Species	<i>Podagrion pachymerum</i>
Number of specimen	52
Stage	adult
Sample location	Slovakia, Rudno nad Hronom
Sample date	Jan-2012
Blood-feeding	no
Suspicious sequences	128

Table 56: Suspicious Sequences in INSlupTATRAAPEI-90.

6 of 128 sequences were true positives and 122 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C92943_a_62_0_l_3472	ORF_003	minor capsid protein, Hubei diptera virus 21 (KX884698)	41%	full
C93031_a_37_0_l_3587	ORF_001	structural protein, Hubei diptera virus 21 (KX884699)	26%	full
C93187_a_43_0_l_3841	ORF_016	hypothetical protein, Hubei diptera virus 21 (KX884697)	31%	full
C93315_a_47_0_l_4221	ORF_007	RdRp , Hubei diptera virus 21 (KX884693)	39%	full
s6000_l_27906_0_a_30_4_l_3801	ORF_006	RdRp , Hubei rhabdo-like virus 6 (KX884421)	53%	partial (start)
s6000_l_27906_0_a_30_4_l_3801	ORF_005	hypothetical protein, Hubei rhabdo-like virus 6 (KX884421)	42%	partial(start)

INSlupTATRAAPEI-90 Sequence Organization

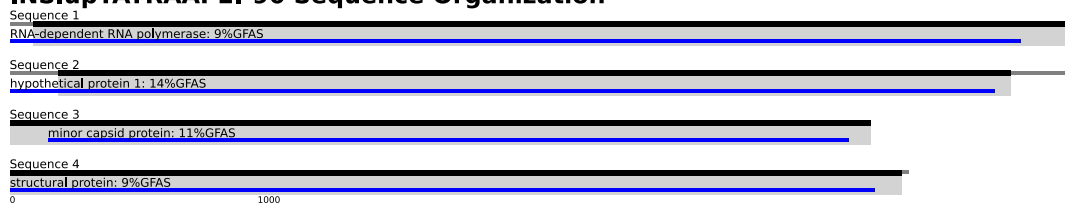


Figure 34: Sequence Organization of INSlupTATRAAPEI-90.

2.29 INSqiqTCXRAAPEI-90

This transcriptome contained few questionable fragments but noch true positive RdRp-like sequence.

Table 57: Sample Information of INSqiqTCXRAAPEI-90.

Filename	130206_I238_FCC1LVUACXX_L4_INSqiqTCXRAAPEI-90.free.fas
Assembly ID	INSqiqTCXRAAPEI-90
Order	Neuroptera
Order details	NA
Family	Hemerobiidae
Family details	NA
Species	<i>Hemerobius humulinus</i>
Number of specimen	10
Stage	NA
Sample location	Italy, Pordenone, Barcis Prescudin
Sample date	02-Aug-2012
Blood-feeding	no
Suspicious sequences	17

Table 58: Suspicious Sequences in INSqiqTCXRAAPEI-90.

4 of 17 sequences were questionable and 13 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C103846_a_3_0_l_322	ORF_002	hypothetical protein1, Hubei diptera virus 21 (KX884697)	44%	partial (mid)
(?) C130385_a_6_0_l_631	ORF_001	hypothetical protein1, Hubei diptera virus 21 (KX884697)	26%	partial (end)
(?) C136315_a_3_0_l_794	ORF_002	hypothetical protein1, Hubei diptera virus 21 (KX884697)	50%	partial (mid)
(?) C145349_a_13_0_l_1352	ORF_006	glycoprotein, Dendrolimus punctatus cypovirus 22 (NC_025838)	34%	full
(?) C74377_a_3_0_l_201	ORF_001	hypothetical protein1, Hubei diptera virus 21 (KX884697)	50%	partial (mid)

INSqiqTCXRAAPEI-90 Sequence Organization

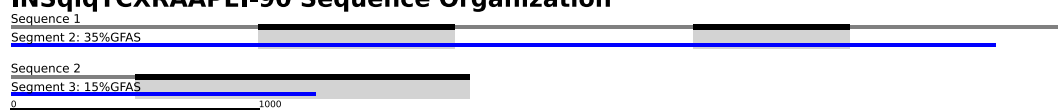


Figure 35: Sequence Organization of INSqiqTCXRAAPEI-90.

2.30 INSqiqTDLRAAPEI-72

This transcriptome contained a potential glzcoprotein but no true positive RdRp-like sequence. howeeever...i do not trust those viruses

Table 59: Sample Information of INSqiqTDLRAAPEI-72.

Filename	130206_I238_FCC1LVUACXX_L4_INSqiqTDLRAAPEI-72.free.fas
Assembly ID	INSqiqTDLRAAPEI-72
Order	Phasmatodea
Order details	NA
Family	Pseudophasmatidae
Family details	Xerosomatinae Hesperophasmatini
Species	<i>Creoxylus spinosus</i>
Number of specimen	1
Stage	adult
Sample location	Germany, lab culture
Sample date	16-Oct-2012
Blood-feeding	no
Suspicious sequences	8

Table 60: Suspicious Sequences in INSqiqTDLRAAPEI-72.

1 of 8 sequences was true positive and 7 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C176836_a_14_0_l_2093	ORF_017	glycoprotein,Wuchang Cockroach Virus 3 (KM817605)	32%	full

INSqiqTDLRAAPEI-72 Sequence Organization

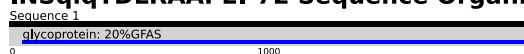


Figure 36: Sequence Organization of INSqiqTDLRAAPEI-72.

2.31 INSobdTDTRAAPEI-18

This transcriptome contained five near full segments similar do different other viruses.

Table 61: Sample Information of INSobdTDTRAAPEI-18.

Filename	130314_I269_FCC1KFEACXX_L7_INSobdTDTRAAPEI-18.free.fas
Assembly ID	INSobdTDTRAAPEI-18
Order	Phasmatodea
Order details	NA
Family	Phylliidae
Family details	Phylliini tribe
Species	<i>Phyllium philippicum</i>
Number of specimen	1
Stage	adult
Sample location	Germany, lab culture
Sample date	13-Nov-2012
Blood-feeding	no
Suspicious sequences	11

Table 62: Suspicious Sequences in INSobdTDTRAAPEI-18.

5 of 10 sequences were true positives and 5 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C143807_a_61_0_l_1898	ORF_002	minor structural protein, Fiji disease virus (NC_007161)	21%	full
C147136_a_60_0_l_3324	ORF_006	segment 4, Mal de Rio Cuarto virus (NC_008729)	21%	full
C147356_a_62_0_l_3744	ORF_006	hypothetical protein, Hubei diptera virus 20 (KX884685)	28%	full
C147412_a_62_0_l_3919	ORF_006	major core protein, Wuhan heteroptera virus 3 (NC_032510)	20%	full
C147496_a_62_0_l_4272	ORF_021	RdRp , Mal de Rio Cuarto virus (NC_008733)	26%	full

INSobdTDTRAAPEI-18 Sequence Organization

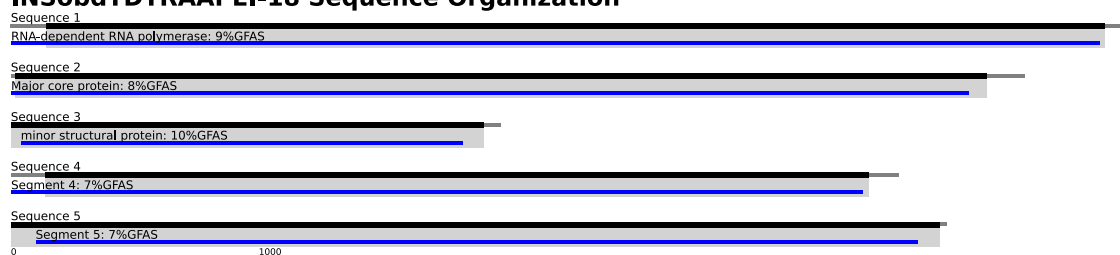


Figure 37: Sequence Organization of INSobdTDTRAAPEI-18.

2.32 INSobdTDYRAAPEI-30

This transcriptome contained a near full RdRp similar to *Hubei reo-like virus 12*. There are several questionable fragments similar to segment 6 of *Dendrolimus punctatus cypovirus 22* and a near full segment similar to segment 10 of *Dendrolimus punctatus cypovirus 22*.

Table 63: Sample Information of INSobdTDYRAAPEI-30.

Filename	130501_I249_FCC1UW6ACXX_L1_INSobdTDYRAAPEI-30.free.fas
Assembly ID	INSobdTDYRAAPEI-30
Order	Diptera
Order details	Brachycera
Family	Rhagionidae
Family details	NA
Species	<i>Chrysopilus thoracicus</i>
Number of specimen	1
Stage	adult
Sample location	USA, North Carolina, Wake County, Raleigh
Sample date	28-May-2012
Blood-feeding	no
Suspicious sequences	45

INSobdTDYRAAPEI-30 Sequence Organization

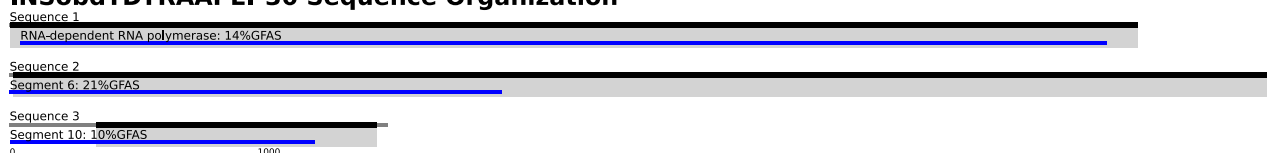


Figure 38: Sequence Organization of INSobdTDYRAAPEI-30.

Table 64: Suspicious Sequences in INSobdTDYRAAPEI-30.

1 of 45 sequences was true positive, 16 were questionable and 28 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C101978_a_3_0_l_243	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	39%	partial (mid)
(?) C111852_a_8_0_l_277	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (mid)
(?) C112940_a_6_0_l_282	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	38%	partial (mid)
(?) C126394_a_3_0_l_364	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	25%	partial (end)
(?) C129516_a_28_0_l_389	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	37%	partial (start)
(?) C132062_a_5_0_l_413	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	30%	partial (end)
(?) C134838_a_7_0_l_442	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	36%	partial (mid)
(?) C137952_a_20_0_l_479	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	30%	partial (end)
(?) C148614_a_6_0_l_673	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (start)
(?) C152290_a_5_0_l_780	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	35%	partial (mid)
(?) C156110_a_5_0_l_944	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	32%	partial (mid-end)
(?) C159321_a_4_0_l_1131	ORF_006	segment 10, Dendrolimus punctatus cypovirus 22 (NC_025838)	24%	full
C169853_a_60_0_l_4548	ORF_010	RdRp , Hubei reo-like virus 12 (KX884634)	40%	full
(?) s10802_l_38510_0_a_38_6_l_3377	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	24%	full
(?) s383_l_329_0_a_3_0_l_375	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (start)
(?) s3885_l_4850_0_a_23_1_l_1267	ORF_006	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	36%	partial (end)
(?) s4657_l_6430_2_a_8_3_l_598	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	30%	partial (end)

2.33 INSerITCGRAAPEI-32

This transcriptome contained two near full sequences similar to the glycoprotein of *Dendrolimus punctatus cypovirus 22* but no RdRp-like true positive sequence.

Table 65: Sample Information of INSerITCGRAAPEI-32.

Filename	130608_I189_FCD20KDACCXX_L3_INSerITCGRAAPEI-32.free.fas
Assembly ID	INSerITCGRAAPEI-32
Order	Raphidioptera
Order details	NA
Family	Inocelliidae
Family details	NA
Species	<i>Fibla maclachlani</i>
Number of specimen	1
Stage	larva
Sample location	Austria, Labstock culture originating from Italy, Sardinia, Sassari
Sample date	22-Feb-2013
Blood-feeding	no
Suspicious sequences	27

Table 66: Suspicious Sequences in INSerITCGRAAPEI-32.

2 of 27 sequences were questionable and 9 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) s5144_L_8051_0_a_48_6_I_2295	ORF_003	glycoprotein, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025838)	30%	full
(?) s5883_L_10241_0_a_25_9_I_3513_a	ORF_011	glycoprotein, <i>Dendrolimus punctatus cypovirus 22</i> (NC_025838)	32%	full

INSerITCGRAAPEI-32 Sequence Organization



Figure 39: Sequence Organization of INSerITCGRAAPEI-32.

2.34 INSkzdTABRAAPEI-136

This transcriptome contained a fragment of an RdRp-like sequence similar to *Dendrolimus punctatus cypovirus*.

Table 67: Sample Information of INSkzdTABRAAPEI-136.

Filename	130720_I246_FCD23MRACXX_L4_INSkzdTABRAAPEI-136.free.fas
Assembly ID	INSkzdTABRAAPEI-136
Order	Psocodea
Order details	Psocomorpha
Family	Lachesillidae
Family details	NA
Species	<i>Lachesilla abiesicola</i>
Number of specimen	NA
Stage	adult
Sample location	Mexico, Cumbres del Ajusco National Park
Sample date	27-Feb-2013
Blood-feeding	no
Suspicious sequences	8

Table 68: Suspicious Sequences in INSkzdTABRAAPEI-136.

1 of 8 sequences was questionable and 7 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C60185_a_5_0_l_422	ORF_001	RdRp, <i>Dendrolimus punctatus cypovirus</i> 22 (NC_025847)	24%	partial (mid)

INSkzdTABRAAPEI-136 Sequence Organization

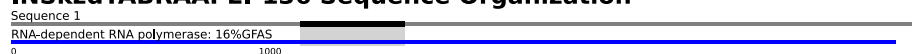


Figure 40: Sequence Organization of INSkzdTABRAAPEI-136.

2.35 INSkzdTACRAAPEI-171

This transcriptome contained two small fragments similar to the RdRp of *Mal de Rio Cuarto virus*.

Table 69: Sample Information of INSkzdTACRAAPEI-171.

Filename	130720_I246_FCD23MRACXX_L4_INSkzdTACRAAPEI-171.free.fas
Assembly ID	INSkzdTACRAAPEI-171
Order	Odonata
Order details	Zygoptera
Family	Pseudostigmatidae
Family details	NA
Species	<i>Megaloprepus caerulatus</i>
Number of specimen	1
Stage	adult
Sample location	Panama, Barro Colorado Island
Sample date	2011
Blood-feeding	no
Suspicious sequences	4

Table 70: Suspicious Sequences in INSkzdTACRAAPEI-171.

3 of 4 sequences were true positives, 1 questionable and 9 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
C143832_a_26_0_l_483	ORF_002	RdRp, Mal de Rio Cuarto virus (NC_008733)	29%	partial (mid)
C157093_a_4_0_l_696	ORF_001	RdRp, Mal de Rio Cuarto virus (NC_008733)	36%	partial (mid)
(?) C168499_a_19_0_l_1118	ORF_005	segment 10, Dendrolimus punctatus cypovirus 22 (NC_025838)	32%	full

INSkzdTACRAAPEI-171 Sequence Organization

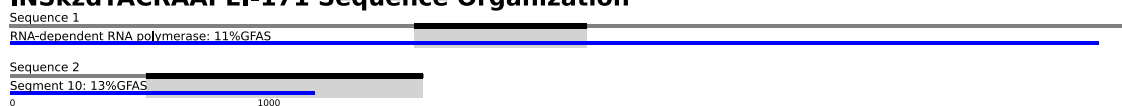


Figure 41: Sequence Organization of INSkzdTACRAAPEI-171.

2.36 INSoFmTBLRAAPEI-71

This transcriptome contained five near full segments similar to *Colorado tick fever virus*, including an RdRp segment.

Table 71: Sample Information of INSoFmTBLRAAPEI-71.

Filename	130728_I263_FCD23HKACXX_L3_INSoFmTBLRAAPEI-71.free.fas
Assembly ID	INSoFmTBLRAAPEI-71
Order	Odonata
Order details	Anisoptera
Family	Aeshnidae
Family details	NA
Species	<i>Anax imperator</i>
Number of specimen	1
Stage	adult
Sample location	Germany, Rhineland-Palatinate, Steinfeld
Sample date	22-Jun-2012
Blood-feeding	no
Suspicious sequences	15

Table 72: Suspicious Sequences in INSoFmTBLRAAPEI-71.

5 of 15 sequences were true positives and 10 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C83693_a_58_0_l_1857	ORF_003	segment 10, Colorado tick fever virus (NC_004189)	27%	full
C86843_a_61_0_l_2465	ORF_006	segment 4, Colorado tick fever virus (NC_004184)	26%	full
C89155_a_61_0_l_3382	ORF_010	segment 3, Colorado tick fever virus (NC_004183)	32%	full
C89824_a_61_0_l_3972	ORF_006	segment 2, Colorado tick fever virus (NC_004182)	26%	full
C89996_a_62_0_l_4241	ORF_012	RdRp , Colorado tick fever virus (NC_004181)	32%	full

INSoFmTBLRAAPEI-71 Sequence Organization

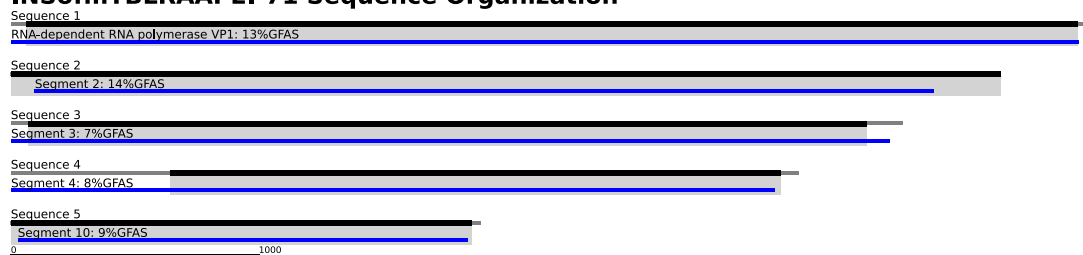


Figure 42: Sequence Organization of INSoFmTBLRAAPEI-71.

2.37 INSofmTCYRAAPEI-79

This transcriptome contained a near full RdRp segment similar to the one of *Rice gall dwarf virus* and a partial capsid protein similar to the same virus.

Table 73: Sample Information of INSofmTCYRAAPEI-79.

Filename	130728_I263_FCD23HKACXX_L4_INSofmTCYRAAPEI-79.free.fas
Assembly ID	INSofmTCYRAAPEI-79
Order	Neuroptera
Order details	NA
Family	Coniopterygidae
Family details	NA
Species	<i>Coniopteryx pygmaea</i>
Number of specimen	18
Stage	not determined
Sample location	Austria, Lower Austria, Wien-Surroundings, Klosterneuburg
Sample date	28-Apr-2013
Blood-feeding	no
Suspicious sequences	17

Table 74: Suspicious Sequences in INSofmTCYRAAPEI-79.

4 of 17 sequences were true positives and 13 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C107296_a_6_0_I_729	ORF_001	minor outer capsid protein, Rice gall dwarf virus (NC_009244)	27%	partial end
C123576_a_32_0_I_3800	ORF_007	RdRp , Rice gall dwarf virus (NC_003773)	32%	full
C80679_a_4_0_I_263	ORF_001	RdRp , Rice gall dwarf virus (NC_009244)	32%	partial (end)

INSofmTCYRAAPEI-79 Sequence Organization

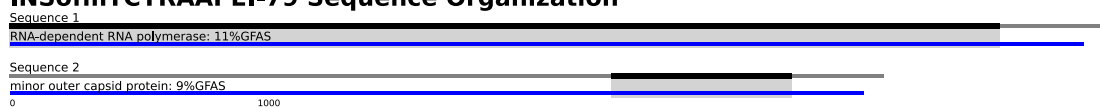


Figure 43: Sequence Organization of INSofmTCYRAAPEI-79.

2.38 INSqiqTDDRABPEI-136

This transcriptome contained two small fragments of an RdRp similar to the one of *Colorado tick fever virus*.

Table 75: Sample Information of INSqiqTDDRABPEI-136.

Filename	130728_I263_FCD23HKACXX_L4_INSqiqTDDRABPEI-136.free.fas
Assembly ID	INSqiqTDDRABPEI-136
Order	Archaeognatha
Order details	NA
Family	Machilidae
Family details	NA
Species	<i>Petridiobius arcticus</i>
Number of specimen	8
Stage	NA
Sample location	USA, labstock Alaska, US Fish and Wildlife service, Kenai National Wildlife Refuge, Soldotna
Sample date	Mar-2012
Blood-feeding	no
Suspicious sequences	15

Table 76: Suspicious Sequences in INSqiqTDDRABPEI-136.

2 of 15 sequences were true positives, 1 questionable and 12 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
C188009_a_5_0_l_439	ORF_002	RdRp, Colorado tick fever virus (NC_004181)	42%	partial (mid)
(?) C225936_a_6_0_l_1366	ORF_003	segment 10, Dendrolimus punctatus cypovirus 22 (NC_025838)	30%	full
s11273_L_59364_0_a_3_0_l_1125	ORF_001	RdRp, Colorado tick fever virus (NC_004181)	37%	partial (mid)

INSqiqTDDRABPEI-136 Sequence Organization

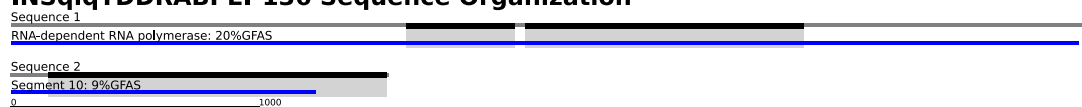


Figure 44: Sequence Organization of INSqiqTDDRABPEI-136.

2.39 INSfjktBIRAAPEI-202

This transcriptome contained a fragment of an RdRp-like sequence similar to *Rice dwarf virus*.

Table 77: Sample Information of INSfjktBIRAAPEI-202.

Filename	130815_I162_FCD2DLYACXX_L2_INSfjktBIRAAPEI-202.free.fas
Assembly ID	INSfjktBIRAAPEI-202
Order	Hymenoptera
Order details	NA
Family	Platygastridae
Family details	NA
Species	<i>Inostemma sp</i>
Number of specimen	51
Stage	adult
Sample location	Germany, North Rhine-Westphalia, Bonn, garden of the ZFMK
Sample date	Jun-2013
Blood-feeding	no
Suspicious sequences	21

Table 78: Suspicious Sequences in INSfjktBIRAAPEI-202.

2 of 21 sequences were questionable and 19 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C171416_a_6_0_l_469	ORF_001	glycoprotein, Lishi Spider Virus 1 (KM817596)	33%	partial (start)
(?) C191441_a_36_0_l_928	ORF_003	RdRp , Rice dwarf virus (NC_003773)	32%	partial (start)

INSfjktBIRAAPEI-202 Sequence Organization

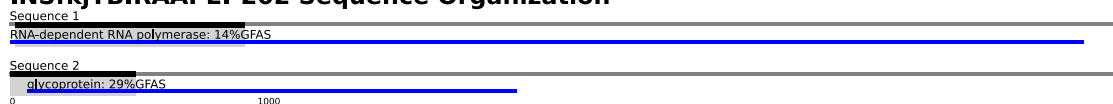


Figure 45: Sequence Organization of INSfjktBIRAAPEI-202.

2.40 INSerITAKRAAPEI-83

This transcriptome contained a glycoprotein similar to the one of *Wuchang Cockroach Virus 3* but no RdRp-like sequence.

Table 79: Sample Information of INSerITAKRAAPEI-83.

Filename	130608_I189_FCD20KDACCXX_L1_INSerITAKRAAPEI-83.free.fas
Assembly ID	INSerITAKRAAPEI-83
Order	Zygentoma
Order details	NA
Family	Lepismatidae
Family details	NA
Species	<i>Ctenolepisma lineata</i>
Number of specimen	4
Stage	not determined
Sample location	Portugal, Faro, near Salir
Sample date	04-Jan-2013
Blood-feeding	no
Suspicious sequences	27

Table 80: Suspicious Sequences in INSerITAKRAAPEI-83.

1 of 27 sequences was true positive and 26 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C184308_a_31_0_l_2127	ORF_003	glycoprotein, Wuchang Cockroach Virus 3 (KM817605)	28%	full

INSerITAKRAAPEI-83 Sequence Organization

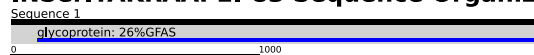


Figure 46: Sequence Organization of INSerITAKRAAPEI-83.

2.41 INSfkjTBMRAAPEI-206

This transcriptome contained several near full segments similar to *Hubei diptera virus 21* and one segment similar to segment 5 of *Operophtera brumata reovirus*.

Table 81: Sample Information of INSfkjTBMRAAPEI-206.

Filename	130831_I260_FCC2BWAACXX_L1_INSfkjTBMRAAPEI-206.free.fas
Assembly ID	INSfkjTBMRAAPEI-206
Order	Hymenoptera
Order details	NA
Family	Encyrtidae
Family details	NA
Species	<i>Metaphycus flavus</i>
Number of specimen	ca 40
Stage	adult
Sample location	Lab culture of unknown geographical origin
Sample date	03-May-2013
Blood-feeding	no
Suspicious sequences	19

Table 82: Suspicious Sequences in INSfkjTBMRAAPEI-206.

8 of 19 sequences were true positives and 11 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C63072_a_16_0_l_1537	ORF_006	hypothetical protein 5, Hubei diptera virus 21 (KX884701)	32%	full
C63130_a_26_0_l_1543	ORF_002	RdRp , Hubei diptera virus 21 (KX884696)	37%	partial (start)
C65354_a_18_0_l_1884	ORF_005	segment 5, Operophtera brumata reovirus (NC_007563)	19%	full
C65544_a_25_0_l_1922	ORF_006	hypothetical protein 4, Hubei diptera virus 21 (KX884700)	24%	full
C67533_a_11_0_l_2759	ORF_012	RdRp , Hubei diptera virus 21 (KX884696)	38%	partial (mid-end)
C67888_a_13_0_l_3447	ORF_001	hypothetical protein 2, Hubei diptera virus 21 (KX884698)	41%	full
C67916_a_19_0_l_3593	ORF_004	hypothetical protein 3, Hubei diptera virus 21 (KX884699)	26%	full
C67944_a_11_0_l_3775	ORF_003	major core capsid protein, Hubei diptera virus 21 (KX884697)	30%	full

INSfkjTBMRAAPEI-206 Sequence Organization

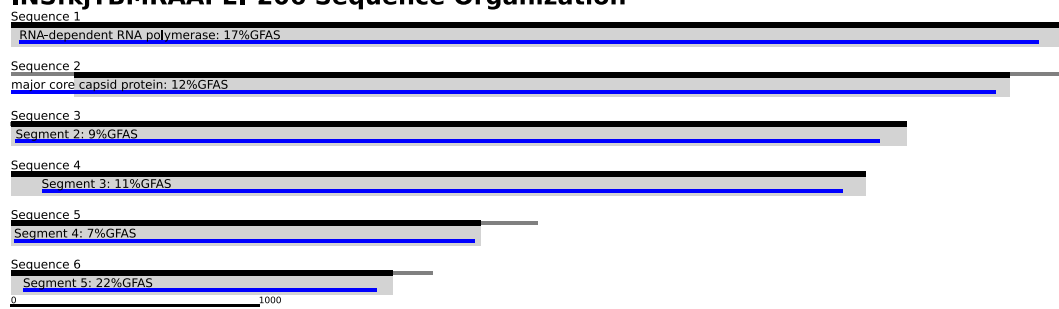


Figure 47: Sequence Organization of INSfkjTBMRAAPEI-206.

2.42 INSofmTCERAAPEI-22

This transcriptome contained a questionably fragment of an RdRp similar to the one of *Rice ragged stund virus* and two other questionable sequences form viruses that are related to many false positives.

Table 83: Sample Information of INSofmTCERAAPEI-22.

Filename	130919_I247_FCC2V7VACXX_L1_INSofmTCERAAPEI-22.free.fas
Assembly ID	INSofmTCERAAPEI-22
Order	Mantodea
Order details	NA
Family	Hymenopodidae
Family details	NA
Species	<i>Harpagomantis tricolor</i>
Number of specimen	1
Stage	NA
Sample location	Germany, Lab culture with Samples originating from South Africa, Johannesburg
Sample date	2012
Blood-feeding	no
Suspicious sequences	10

Table 84: Suspicious Sequences in INSofmTCERAAPEI-22.

4 of 10 sequences were questionable and 6 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C129377_a_3_0_l_312	ORF_001	RdRp, Rice ragged stunt virus (NC_003771)	45%	partial (start)
(?) C142385_a_6_0_l_382	ORF_003	segment 11, Liao ning virus (NC_007746)	34%	partial (start)
(?) C186124_a_8_0_l_1235	ORF_002	sigma 1, Mammalian Orthoreovirus (JQ412761)	20%	full

INSofmTCERAAPEI-22 Sequence Organization

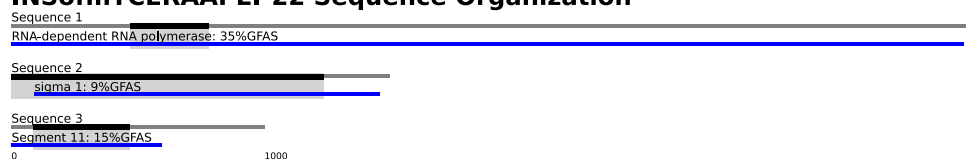


Figure 48: Sequence Organization of INSofmTCERAAPEI-22.

2.43 INSoftmTCFRAAPEI-26

This transcriptome contained three questionable fragments similar to the RdRp of *Equine encephalosis virus*.

Table 85: Sample Information of INSoftmTCFRAAPEI-26.

Filename	130919_I247_FCC2V7VACXX_L1_INSoftmTCFRAAPEI-26.free.fas
Assembly ID	INSoftmTCFRAAPEI-26
Order	Blattodea
Order details	NA
Family	Corydiidae
Family details	Tiviinae
Species	<i>Tivia sp</i>
Number of specimen	1
Stage	adult
Sample location	Namibia, Otjozondjupa, Waterberg
Sample date	05-Apr-2013
Blood-feeding	no
Suspicious sequences	3

Table 86: Suspicious Sequences in INSoftmTCFRAAPEI-26.

2 of 3 sequences were questionable and 1 sequence was false positive similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C429509_a_3_0_l_558	ORF_002	RdRp, Equine encephalosis virus (AB811635)	32%	partial (mid)
(?) s11468_L_28221_0_a_5_6_l_982	ORF_001	RdRp, Equine encephalosis virus (AB811635)	39%	partial (mid)
(?) s11468_L_28221_0_a_5_6_l_982	ORF_005	RdRp, Equine encephalosis virus (AB811635)	39%	partial (mid)

INSoftmTCFRAAPEI-26 Sequence Organization

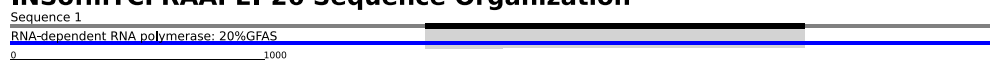


Figure 49: Sequence Organization of INSoftmTCFRAAPEI-26.

2.44 INSinTAARABPEI-43

This transcriptome contained a small fragment of an RdRp similar to the one of *Hubei diptera virus 10*.

Table 87: Sample Information of INSinTAARABPEI-43.

Filename	130919_I247_FCC2V7VACXX_L3_INSinTAARABPEI-43.free.fas
Assembly ID	INSinTAARABPEI-43
Order	Orthoptera
Order details	Caelifera
Family	Pyrgomorphidae
Family details	NA
Species	<i>Psedna nana</i>
Number of specimen	1
Stage	adult
Sample location	Australia, Western Australia, Toodyay
Sample date	Dec-2011
Blood-feeding	no
Suspicious sequences	1

Table 88: Suspicious Sequences in INSinTAARABPEI-43.

1 of 1 sequences was questionable.

Sequence ID	ORF	Match	Identity	Completeness
(?) s4225_L_7567_0_a_3_0_I_271	ORF_001	RdRp, Hubei diptera virus 20 (KX884693)	56%	partial (mid)

INSinTAARABPEI-43 Sequence Organization

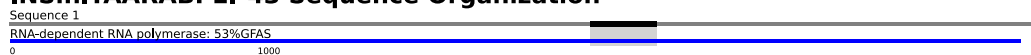


Figure 50: Sequence Organization of INSinTAARABPEI-43.

2.45 INSinITAPRAAPEI-33

This transcriptome contained a fragmentary RdRp similar to the one of *Hubei rhabdo-like virus*. Additionally, another hypothetical ORF showed similarity to the same virus.

Table 89: Sample Information of INSinITAPRAAPEI-33.

Filename	130919_I247_FCC2V7VACXX_L6_INSinITAPRAAPEI-33.free.fas
Assembly ID	INSinITAPRAAPEI-33
Order	Hymenoptera
Order details	NA
Family	Agaonidae
Family details	NA
Species	<i>Courtella sp</i>
Number of specimen	50
Stage	adult
Sample location	South Africa, Western Cape, Kirstenhof, Waterford Circle
Sample date	21-Apr-2013
Blood-feeding	no
Suspicious sequences	8

Table 90: Suspicious Sequences in INSinITAPRAAPEI-33.

1 of 8 sequences was true positive and 7 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C111078_a_9_0_I_661	ORF_005	RdRp, Hubei rhabdo-like virus 6 (KX884421)	43%	partial (mid)

INSinITAPRAAPEI-33 Sequence Organization

Sequence 1
RNA-dependent RNA polymerase: 616GFAS

hypothetical protein: 474GFAS

Figure 51: Sequence Organization of INSinITAPRAAPEI-33.

2.46 INSinTAWRAAPEI-44

This transcriptome contained a fragmentary RdRp similar to the one of *Diaphorina citri* associated *C virus*.

Table 91: Sample Information of INSinTAWRAAPEI-44.

Filename	130919_I247_FCC2V7VACXX_L6_INSinTAWRAAPEI-44.free.fas
Assembly ID	INSinTAWRAAPEI-44
Order	Diptera
Order details	Brachycera
Family	Lonchopteridae
Family details	NA
Species	<i>Lonchoptera bifurcata</i>
Number of specimen	10
Stage	adult
Sample location	USA, North Carolina, Wake County, Raleigh, Schenck forest
Sample date	25-May-2013
Blood-feeding	no
Suspicious sequences	7

Table 92: Suspicious Sequences in INSinTAWRAAPEI-44.

2 of 7 sequences were true positives and 5 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
s2259_L_4066_0_a_5_0_I_449	ORF_001	RdRp, <i>Diaphorina citri</i> associated C virus (KX235518)	46%	partial (mid)
s2260_L_4066_1_a_3_3_I_878	ORF_001	RdRp, <i>Diaphorina citri</i> associated C virus (KX235518)	45%	partial (mid)

INSinTAWRAAPEI-44 Sequence Organization



Figure 52: Sequence Organization of INSinTAWRAAPEI-44.

2.47 RINSinITCARAAPEI-55

This transcriptome contained two fragments of an RdRp similar to the one of *Hubei odonate virus 15*.

Table 93: Sample Information of RINSinITCARAAPEI-55.

Filename	130919_I247_FCC2V7VACXX_L7_RINSinITCARAAPEI-55.free.fas
Assembly ID	RINSinITCARAAPEI-55
Order	Diptera
Order details	Brachycera
Family	Tachinidae
Family details	NA
Species	<i>Euthera bicolor</i>
Number of specimen	1
Stage	adult
Sample location	USA, Mississippi, Noxubee County, Sam D Hamilton, Noxubee National Wildlife Refuge
Sample date	19-May-2013
Blood-feeding	no
Suspicious sequences	4

Table 94: Suspicious Sequences in RINSinITCARAAPEI-55.

2 of 4 sequences were true positives and 2 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C61403_a_15_0_l_1491	ORF_004	RdRp, Hubei odonate virus 15 (KX884664)	46%	partial (mid)
s3320_l_16467_0_a_24_6_l_1796	ORF_008	RdRp, Hubei odonate virus 15 (KX884664)	49%	partial (start)

RINSinITCARAAPEI-55 Sequence Organization



Figure 53: Sequence Organization of RINSinITCARAAPEI-55.

2.48 RINSinITCNRAAPEI-33

This transcriptome contained several full segments similar to *Cimodo virus*. The contig s10827_L_38271_0_a_47_1_l_7816 seemed to be a missassembly and has been cut down to the matching region.

Table 95: Sample Information of RINSinITCNRAAPEI-33.

Filename	130928_I232_FCC2UV4ACXX_L1_RINSinITCNRAAPEI-33.free.fas
Assembly ID	RINSinITCNRAAPEI-33
Order	Hymenoptera
Order details	NA
Family	Eulophidae
Family details	NA
Species	<i>Tamarixia radiata</i>
Number of specimen	ca 131
Stage	adult
Sample location	Pakistan, Lab culture with Samples originating from Pakistan, Punjab
Sample date	May-2013
Blood-feeding	no
Suspicious sequences	44

Table 96: Suspicious Sequences in RINSinITCNRAAPEI-33

4 of 44 sequences were true positives and 40 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C135862_a_52_0_l_1820	ORF_008	segment 6, Cimodo virus (NC_024920)	27%	full
C143103_a_33_0_l_3239	ORF_003	segment 3, Cimodo virus (NC_024917)	32%	full
C144637_a_50_0_l_4075	ORF_004	RdRp, Cimodo virus (NC_023420)	41%	full
s10827_L_38271_0_a_47_1_l_7816	ORF_013	segment 2, Cimodo virus (NC_024916)	32%	full

RINSinITCNRAAPEI-33 Sequence Organization

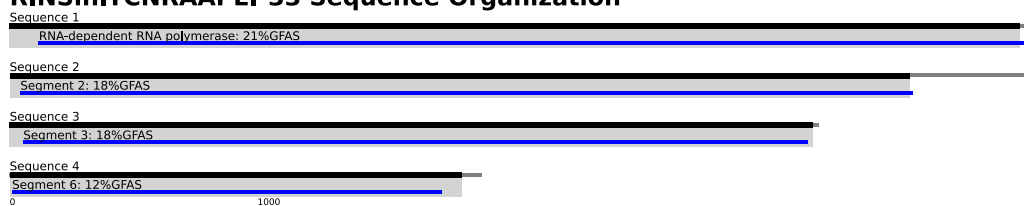


Figure 54: Sequence Organization of RINSinITCNRAAPEI-33.

2.49 RINSymITABRAAPEI-202

This transcriptome contained a full RdRp similar to the one of *Hubei odonate virus 15*.

Table 97: Sample Information of RINSymITABRAAPEI-202.

Filename	131212_I249_FCC39K4ACXX_L6_RINSymITABRAAPEI-202.free.fas
Assembly ID	RINSymITABRAAPEI-202
Order	Diptera
Order details	Brachycera
Family	Oestridae
Family details	NA
Species	<i>Cuterebra austeni</i>
Number of specimen	1
Stage	adult
Sample location	USA, New Mexico, Grant County, Silver City, Gomez Park
Sample date	27-May-2013
Blood-feeding	yes-larvae
Suspicious sequences	3

Table 98: Suspicious Sequences in RINSymITABRAAPEI-202.

1 of 3 sequences were true positives and 2 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C40647_a_62_0_l_4009	ORF_011	RdRp, Hubei odonate virus 15 (KX884664)	45%	full

RINSymITABRAAPEI-202 Sequence Organization



Figure 55: Sequence Organization of RINSymITABRAAPEI-202.

2.50 RINSwvkTAURAAPEI-56

This transcriptome contained a glycoprotein similar to the one of *Wuchang Cockroach Virus* 3 but no RdRp-like true positive.

Table 99: Sample Information of RINSwvkTAURAAPEI-56.

Filename	140430_I162_FCC4EFCACXX_L1_RINSwvkTAURAAPEI-56.free.fas
Assembly ID	RINSwvkTAURAAPEI-56
Order	Neuroptera
Order details	NA
Family	Chrysopidae
Family details	NA
Species	<i>Chrysopa formosa</i>
Number of specimen	3
Stage	adult
Sample location	Austria, Lower Austria, Krems-Land, Duernstein
Sample date	27-Jul-2013
Blood-feeding	no
Suspicious sequences	13

Table 100: Suspicious Sequences in RINSwvkTAURAAPEI-56.

1 of 13 sequences was true positive and 12 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
s3755_L_6696_1_a_79_3_I_2180	ORF_001	glycoprotein, Wuchang Cockroach Virus 3 (KM817605)	39%	partial

RINSwvkTAURAAPEI-56 Sequence Organization

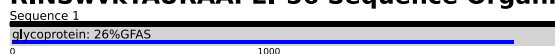


Figure 56: Sequence Organization of RINSwvkTAURAAPEI-56.

2.51 ANIsrmTAAWRAAPEI-225

This transcriptome contained a fragmentary RdRp similar to the one of *Morris orbivirus*.

Table 101: Sample Information of ANIsrmTAAWRAAPEI-225.

Filename	140710_I812_FCH8K85ADXX_L2_ANIsrmTAAWRAAPEI-225.free.fas
Assembly ID	ANIsrmTAAWRAAPEI-225
Order	Lepidoptera
Order details	NA
Family	Nepticulidae
Family details	NA
Species	<i>Stigmella atricapitella</i>
Number of specimen	1
Stage	adult
Sample location	Germany, Rhineland-Palatinate, Oberhausen an der Nahe, Rabenfelsen
Sample date	28-Jun-2011
Blood-feeding	no
Suspicious sequences	14

Table 102: Suspicious Sequences in ANIsrmTAAWRAAPEI-225.

1 of 14 sequences was true positive and 13 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C128088_a_7_0_I_1169	ORF_004	RdRp, <i>Morris orbivirus</i> (KX907618)	42%	partial (mid)

ANIsrmTAAWRAAPEI-225 Sequence Organization



Figure 57: Sequence Organization of ANIsrmTAAWRAAPEI-225.

2.52 WHANIsrmTMAFRAAPEI-14

This transcriptome contained a small fragment of an RdRp similar to the one of *Equine encephalosis virus*.

Table 103: Sample Information of WHANIsrmTMAFRAAPEI-14.

Filename	140813_I652_FCC4L86ACXX_L1_WHANIsrmTMAFRAAPEI-14.free.fas
Assembly ID	WHANIsrmTMAFRAAPEI-14
Order	Coleoptera
Order details	NA
Family	Coccinellidae
Family details	Scymninae
Species	<i>Cryptolaemus montrouzieri</i>
Number of specimen	1
Stage	adult
Sample location	Australia, Australian Capital Territory, Australian National Insect Collection, Acton 2601, Black Mountain
Sample date	13-Oct-2013
Blood-feeding	no
Suspicious sequences	18

Table 104: Suspicious Sequences in WHANIsrmTMAFRAAPEI-14.

1 of 18 sequences was questionable and 17 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C100756_a_7_0_1_408	ORF_004	RdRp, Equine encephalosis virus (AB811635)	37%	partial (mid)

WHANIsrmTMAFRAAPEI-14 Sequence Organization

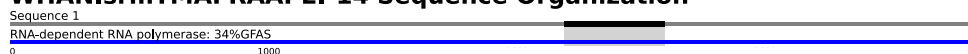


Figure 58: Sequence Organization of WHANIsrmTMAFRAAPEI-14.

2.53 WHANIsrmTMCHRAAPEI-56

This transcriptome contained a very small fragment of an RdRp similar to the one of *Hubei insect virus 2* and another segment with a putative major core capsid protein similar to the one of the same virus.

Table 105: Sample Information of WHANIsrmTMCHRAAPEI-56.

Filename	140813_I652_FCC4L86ACXX_L4_WHANIsrmTMCHRAAPEI-56.free.fas
Assembly ID	WHANIsrmTMCHRAAPEI-56
Order	Lepidoptera
Order details	NA
Family	Epiptropidae
Family details	NA
Species	<i>Epipomponia nawai</i>
Number of specimen	1
Stage	NA
Sample location	South Korea, Ulsan City, Uiju, Mount Ganweolsan
Sample date	28-Aug-2012
Blood-feeding	no
Suspicious sequences	18

Table 106: Suspicious Sequences in WHANIsrmTMCHRAAPEI-56.

2 of 18 sequences were true positives and 16 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C149031_a_5_0_l_342	ORF_001	RdRp, Hubei insect virus 2 (NC_032167)	45%	partial (mid)
C200126_a_3_0_l_1375	ORF_001	major core capsid protein, Hubei insect virus 2 (NC_032161)	25%	partial (mid)

WHANIsrmTMCHRAAPEI-56 Sequence Organization

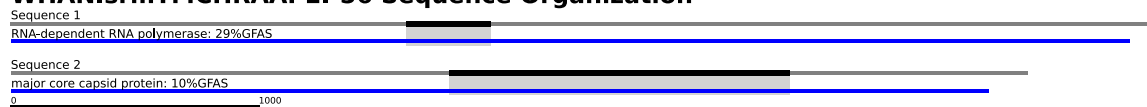


Figure 59: Sequence Organization of WHANIsrmTMCHRAAPEI-56.

2.54 INSeqtTBBRAAPEI-75

This transcriptome contained several near full segments of a virus that has similarities with *Operophtera brumata reovirus*. Additionally, there were as well many small ORFs yielding hits for segment 6, *Dendrolimus punctatus cypovirus 22*.

Table 107: Sample Information of INSeqtTBBRAAPEI-75.

Filename	121010_I249_FCD1C4BACXX_L7_INSeqtTBBRAAPEI-75.free.fas
Assembly ID	INSeqtTBBRAAPEI-75
Order	Hymenoptera
Order details	NA
Family	Diapriidae
Family details	NA
Species	<i>Trichopria drosophilae</i>
Number of specimen	ca 80
Stage	adult
Sample location	Lab culture with Samples originating from France, 60 km south of Lyon Sablons, Is re district
Sample date	2012
Blood-feeding	no
Suspicious sequences	48

INSeqtTBBRAAPEI-75 Sequence Organization

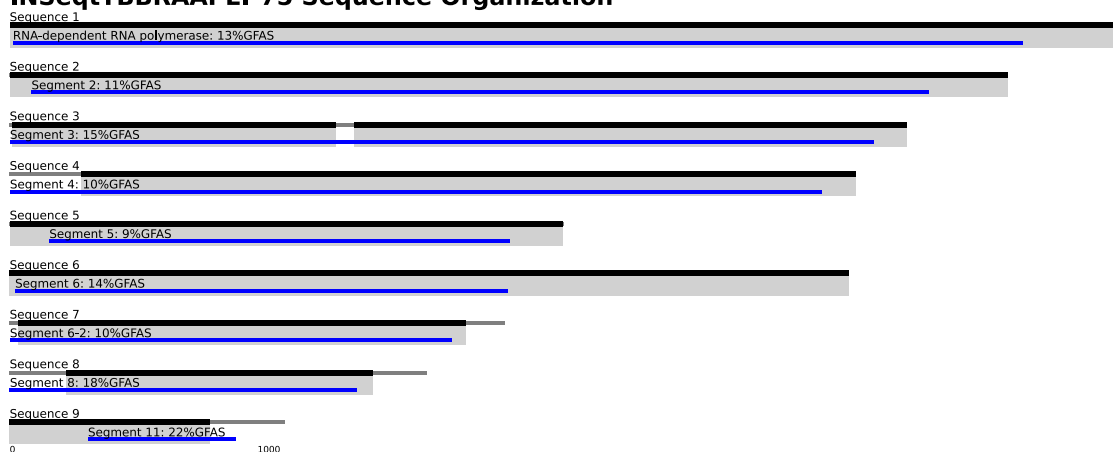


Figure 60: Sequence Organization of INSeqtTBBRAAPEI-75.

Table 108: Suspicious Sequences in INSeqTBBRAAPEI-75.

11 of 48 sequences were true positives, 21 questionable and 16 sequences were false positives similar to the false positives listed in chapter 3.2.2. Questionable sequences are marked with (?).

Sequence ID	ORF	Match	Identity	Completeness
(?) C17507_a_3_0_l_204	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	32%	partial (start)
(?) C18879_a_3_0_l_223	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (start)
(?) C21267_a_3_0_l_244	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	38%	partial (start)
(?) C23307_a_3_0_l_257	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	58%	partial (mid)
(?) C24999_a_4_0_l_273	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	35%	partial (end)
(?) C27963_a_3_0_l_324	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	37%	partial (mid)
C36977_a_47_0_l_528	ORF_002	RdRp , Operophtera brumata reovirus (NC_007559)	38%	partial (end)
(?) C37285_a_4_0_l_537	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	43%	partial (end)
(?) C37453_a_6_0_l_542	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	38%	partial (mid)
(?) C40235_a_5_0_l_643	ORF_005	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	37%	partial (end)
C41097_a_12_0_l_677	ORF_002	segment 6, Operophtera brumata reovirus (NC_007564)	37%	partial (end)
(?) C42503_a_6_0_l_744	ORF_003	segment 11, Liao ning virus (NC_007746)	21%	full
(?) C43717_a_4_0_l_802	ORF_005	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	40%	partial (end)
(?) C44833_a_3_0_l_866	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (mid)
C47520_a_10_0_l_1047	ORF_001	segment 3, Operophtera brumata reovirus (NC_007561)	40%	partial (mid)
C49320_a_9_0_l_1189	ORF_001	segment 6, Operophtera brumata reovirus (NC_007564)	24%	partial (mid-end)
C49430_a_19_0_l_1203	ORF_005	segment 3, Operophtera brumata reovirus (NC_007561)	29%	partial (end)
C49840_a_20_0_l_1238	ORF_003	segment 8, Operophtera brumata reovirus (NC_007566)	27%	partial (mid-end)
C50472_a_8_0_l_1305	ORF_002	segment 3, Operophtera brumata reovirus (NC_007561)	25%	partial (start)
(?) C52230_a_7_0_l_1503	ORF_002	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	32%	partial (mid-end)
(?) C52824_a_3_0_l_1571	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (start-mid)
C56798_a_53_0_l_2229	ORF_001	segment 5, Operophtera brumata reovirus (NC_007563)	19%	full
C59458_a_42_0_l_3125	ORF_006	segment 4, Operophtera brumata reovirus (NC_007562)	32%	full
C60500_a_30_0_l_3935	ORF_012	segment 2, Operophtera brumata reovirus (NC_007560)	27%	full
C60856_a_44_0_l_4409	ORF_001	RdRp , Operophtera brumata reovirus (NC_007559)	35%	full
(?) s1322_l_1611_0_a_3_0_l_526	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (mid)
(?) s1726_l_2289_0_a_15_5_l_3299	ORF_001	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	30%	full
(?) s1831_l_2457_0_a_5_7_l_646	ORF_004	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (mid)
(?) s2132_l_3130_2_a_31_5_l_832	ORF_005	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (start)
(?) s406_l_427_0_a_6_1_l_767	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	45%	partial (end)
(?) s407_l_427_1_a_5_9_l_767	ORF_003	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	34%	partial (end)
(?) s924_l_1022_0_a_3_8_l_1718	ORF_010	segment 6, Dendrolimus punctatus cypovirus 22 (NC_025850)	31%	partial (mid-end)

2.55 INSobdDIRAAPEI-84

This transcriptome contained several full fragments similar to *Operopthera brumata reovirus*.

Table 109: Sample Information of INSobdDIRAAPEI-84.

Filename	130314_I269_FCC1KFEACXX_L8_INSobdDIRAAPEI-84.free.fas
Assembly ID	INSobdDIRAAPEI-84
Order	Hymenoptera
Order details	NA
Family	Pteromalidae
Family details	NA
Species	<i>Lariophagus distinguendus</i>
Number of specimen	70
Stage	adult
Sample location	Germany, Lab culture with Samples originating from Germany, Berlin
Sample date	Nov-2012
Blood-feeding	no
Suspicious sequences	29

Table 110: Suspicious Sequences in INSobdDIRAAPEI-84.

6 of 29 sequences were true positives and 23 sequences were false positives similar to the false positives listed in chapter 3.2.2.

Sequence ID	ORF	Match	Identity	Completeness
C74954_a_59_0_l_1852	ORF_001	segment 6, <i>Operopthera brumata reovirus</i> (NC_007564)	25%	full
C75590_a_61_0_l_1947	ORF_003	segment 5, <i>Operopthera brumata reovirus</i> (NC_007563)	20%	full
C79547_a_47_0_l_3588	ORF_003	segment 3, <i>Operopthera brumata reovirus</i> (NC_007561)	25%	full
C79689_a_62_0_l_3813	ORF_001	segment 2, <i>Operopthera brumata reovirus</i> (NC_007560)	30%	full
C79803_a_56_0_l_4127	ORF_001	RdRp , <i>Operopthera brumata reovirus</i> (NC_007559)	34%	full
s4658_L_21876_0_a_45_2_l_3350	ORF_007	segment 4, <i>Operopthera brumata reovirus</i> (NC_007562)	36%	full

INSobdDIRAAPEI-84 Sequence Organization

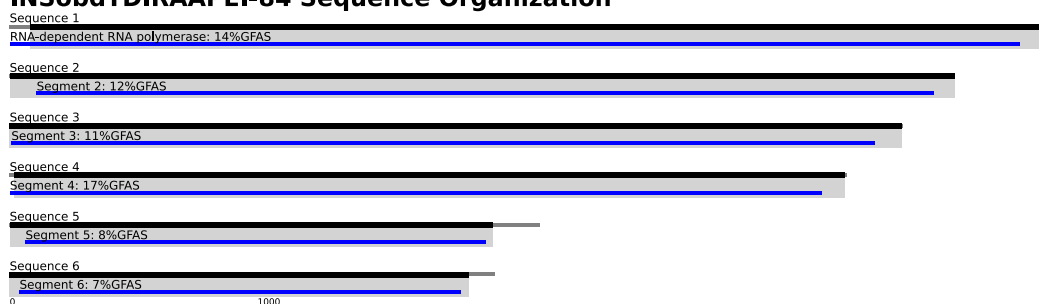


Figure 61: Sequence Organization of INSobdDIRAAPEI-84.