

# Untersuchung verschiedener Virusspezies in neotropischen Nagetieren und den Einflüssen ökologischer Faktoren

## Dissertation

zur Erlangung des akademischen Grades  
Doctor rerum naturalium  
(Dr.rer.nat.)

vorgelegt dem Fachbereichs Biologie, Chemie, Pharmazie  
der Freien Universität Berlin

vorgelegt von

**Fabian Pirzer**

2022

Diese Arbeit wurde von 2017 bis 2021 am Institut für Virologie der  
Charité - Universitätsmedizin Berlin, unter der Leitung von Prof. Christian  
Drosten angefertigt.

1. Gutachter und Betreuer: Prof. Christian Drosten
2. Gutachter: Prof. Dino McMahon

Disputation am: 20.12.2023

## **Eidesstattliche Erklärung**

Hiermit versichere ich, die vorliegende Dissertation selbstständig und ohne unerlaubte Hilfe angefertigt zu haben.

Bei der Verfassung der Dissertation wurden keine anderen als die im Text aufgeführten Hilfsmittel verwendet.

Ein Promotionsverfahren wurde zu keinem früheren Zeitpunkt an einer anderen Hochschule oder bei einem anderen Fachbereich beantragt.

Fabian Pirzer

Berlin, 17.06.2022

## **Danksagungen:**

Mein erster Dank gilt Prof. Christian Drosten, sowie Dr. Victor Corman, die mir die Möglichkeit gegeben haben am Institut für Virologie an diesem spannenden und lehrreichen Thema zu arbeiten, sowie für ihre Betreuung, Führung und der hilfreichen Diskussionen

Als nächstes möchte ich Prof. Dino McMahon danken, dass er sich bereit erklärt hat als Gutachter meiner Dissertation zu fungieren.

Des Weiteren danke ich Tobias Bleicker, Sebastian Brünik und Monika-Eschbach Bludau für ihre tatkräftige Unterstützung im Laboralltag und für ein offenes Ohr bei unzähligen Fragen und Anliegen.

Auch danke ich all meinen anderen Kollegen des Instituts für Virologie für das professionelle und freundliche Arbeitsklima.

Besonderer Dank geht hierbei an das Büro der AG Jones, welches ein Sammelsurium fantastischer Menschen beinhaltet.

Zu guter Letzt danke ich meinen Freunden und meiner Familie für die Unterstützung während dieser Doktorarbeit.

# Inhaltsverzeichnis

<b>1. Einführung .....</b>	<b>1</b>
1.1. Zoonotische Pathogene und ihre Gefahr für den Menschen .....	1
1.2. Einfluss ökologischer Faktoren auf Pathogen .....	2
1.3. Bau des Panamakanals und dessen Einfluss auf den Urwald.....	3
1.4. Proechimys semispinosus .....	4
1.5. Einführung in die behandelten Viren .....	7
1.5.1. Picobirnavirus Grundlagen .....	7
1.5.2. Picornavirus Grundlagen .....	9
1.5.3. Deltavirus Grundlagen .....	10
1.5.4. Hepacivirus Grundlagen .....	12
1.5.5. Zielsetzung der vorliegenden Arbeit.....	14
<b>2. Material und Methoden .....</b>	<b>15</b>
2.1. Benutzte Materialien .....	15
2.1.1. Chemikalien.....	15
2.1.2. Kits .....	15
2.2. Software .....	16
2.3. Benutzte Geräte .....	16
2.4. Methoden .....	19
2.4.1. Fangprozedur und Probenahme .....	19
2.4.2. RNA Extraktion mittels MagNA Pure 96.....	21
2.4.3. Detektion viraler Erreger mittels PCR .....	22
2.4.3.1. Picobirnavirus.....	24
2.4.3.2. Picornavirus.....	25
2.4.3.3. Hepacivirus.....	28
2.4.3.4. Deltavirus .....	29
2.4.4. Deltavirus Zirkularisierungsassay .....	34
2.4.5. RT-PCR Amplikon-Quantifizierung .....	36
2.4.5.1. Herstellung einer in-vitro Transkript Standardreihe für qPCR.....	38
2.4.6. Aufreinigung von DNA aus PCR-Produkten.....	39
2.4.7. Aufreinigung der DNA aus Agarosegelen .....	39
2.4.8. Konzentrationsmessung von Nukleinsäuren.....	40
2.4.9. Sanger-Sequenzierung.....	40

2.4.10.	Next Generation Sequencing (NGS) .....	40
2.4.11.	Identifikation unbekannter Erreger mittels Pipeline .....	41
2.4.12.	Sequenzanalyse .....	42
2.4.13.	Phylogenetische Analyse .....	42
2.4.13.1.	Bayesianische Analyse zur Evolutionsrate per BEAST .....	42
2.4.13.2.	Modelauswahl .....	43
2.4.13.3.	Überprüfung temporaler Signale mit TempEst .....	44
2.4.14.	Statistische Auswertung .....	45
2.4.14.1.	Ökologische Parameter .....	45
<b>3.</b>	<b>Ergebnisse und Diskussion .....</b>	<b>47</b>
3.1.	Ergebnisse .....	47
3.1.1.	Spezies- und Probenübersicht .....	47
3.1.2.	Picornavirus Analysen .....	52
3.1.3.	Picobirnavirus Analysen .....	59
3.1.4.	Deltavirus Analysen .....	68
3.1.5.	Hepacivirus Analysen .....	74
3.1.6.	Detektionsraten aller Viren in <i>Proechimys semispinosus</i> .....	88
3.2.	Diskussion .....	90
3.2.1.	Picornavirus Auswertung .....	90
3.2.2.	Picobirnavirus Auswertung .....	91
3.2.3.	Deltavirus Auswertung .....	92
3.2.4.	Hepacivirus Auswertung .....	92
<b>4.</b>	<b>Zusammenfassung .....</b>	<b>95</b>
4.1.	Summary .....	96

<b>5. Abbildungsverzeichnis.....</b>	<b>97</b>
<b>6. Tabellenverzeichnis.....</b>	<b>99</b>
<b>7. Abkürzungsverzeichnis.....</b>	<b>101</b>
<b>8. Literaturverzeichnis.....</b>	<b>103</b>
<b>9. Erfolgte Publikationen.....</b>	<b>111</b>

## 1. Einführung

### 1.1. Zoonotische Pathogene und ihre Gefahr für den Menschen

Die Menschheit wächst zunehmend und beansprucht immer mehr Lebensraum und Ressourcen, was zu Lasten der natürlichen Flora und Fauna geschieht. Infolge dieser Urbanisierung werden Wälder und andere Ökosysteme zerstört, um Platz für Siedlungen, Landwirtschaft oder Industrie zu schaffen sowie um natürliche Rohstoffe abzubauen. Dieses Vordringen führt zu einer Verdrängung von Wildtierpopulationen aus ihren natürlichen Habitaten und kann auch mit einem Artensterben einhergehen. Sodass einerseits, durch diese Einflüsse der Artenreichtum (oder auch Biodiversität) in solchen Regionen abnehmen kann [1] und andererseits steigt die Anzahl menschlicher und tierischer Kontakte. Dies kann zum Auftreten von zoonotischen Infektionen führen.

Unter zoonotischen Erregern versteht man Pathogene, die die Artenbarriere überwinden und vom Tier auf den Menschen überspringen können. Zahlreiche Übertragungswege sind dabei beschrieben, darunter: direkter Kontakt, Einatmung von Aerosolen kontaminierter Ausscheidungen, sowie die Übertragung durch Vektoren.

Darüber hinaus ist auch der Verkauf und Verzehr von Wildtierfleisch ein entscheidender Faktor. In manchen Ländern Südostasiens, Afrikas oder Australiens wird sogenanntes Buschfleisch als essentielle Nahrungsquelle angesehen und ist teilweise sowohl kulturell als auch traditionell stark in den Lebensalltag dieser Regionen integriert [2]. Dies birgt ein großes Gefahrenpotential, da viele dieser Wildtierarten Wirte für noch unbekannte Krankheiten sein können. Unser Verständnis für das Auftreten dieser zoonotischen Pathogene ist momentan noch limitiert. Ein Beispiel dafür ist die aktuelle Pandemie, hervorgerufen durch SARS-CoV-2, welches seinen Ursprung höchstwahrscheinlich auf einen Wildtiermarkt in Wuhan, China hat [3].

Da die Anzahl zoonotischer Pathogene bislang nicht abgeschätzt werden kann ist es umso wichtiger grundlegende Mechanismen und Effekte zu erforschen. Dies ermöglicht eine bessere Abschätzung des Gefahrenpotentials und kann bei der Entwicklung effektiver Maßnahmen gegen Übertragungen mit solchen Erregern von Nutzen sein.

## 1.2. Einfluss ökologischer Faktoren auf Pathogen

Bei einer Virusübertragung spielen viele Parameter und Faktoren eine wichtige Rolle und bestimmen ob eine Infektion stattfindet. Neben viralen Faktoren, wie etwa Infektiosität oder Übertragungsweg, und wirtsspezifischen Faktoren wie z.B. Resistenz, sind auch Umweltparameter von großer Wichtigkeit. Bei diesen ökologischen Faktoren, kann man zwischen abiotischen und biotischen Faktoren unterscheiden. Abiotische Faktoren beschreiben unbelebten Faktoren wie Luftfeuchtigkeit, Jahreszeit, Temperatur oder Geländeart. Die biotischen Faktoren umfassen die Interaktionen und Wechselwirkungen zwischen Lebewesen, wie etwa Räuber-Beute Verhältnis, Speziesdichte oder Artenreichtum. Die wichtigsten ökologischen Faktoren umfassen die Anzahl der infizierten Individuen und die Kontakthäufigkeit mit den Pathogenen.

Bei Viren handelt es sich um Pathogene ohne eigenen Stoffwechsel. Aus diesem Grund sind diese nur mithilfe eines Wirtes und dessen zellulärer Ausstattung imstande sich zu replizieren [4].

Die Virusprävalenz, also der prozentuale Anteil einer Population, die mit einem Virus infiziert ist, hängt von verschiedenen direkten und indirekten Faktoren ab. Zu direkten Faktoren zählt die Wirtsspezifität oder Pathogenität eines Virus, während man unter indirekten Faktoren die Populationsdichte und Umweltparameter, wie unter anderem Klima und Jahreszeiten, versteht.

Viren, die eine hohe Wirtsspezifität besitzen, sind in dem Wirtsspektrum, dass sie infizieren können eingeschränkt. Durch diese Spezialisierung umgehen sie teilweise die Immunantwort und nutzen Zellmechanismen für ihre eigene Replikation. Diese starke Anpassung kann aber auch Nachteile mit sich bringen. Mit weniger potentielle Wirte, die für eine Infektion anfällig sind, stehen auch dauerhaft weniger Reservoir zur Verfügung in denen eine Replikation stattfinden kann.

In der Theorie bedeutet dies, dass je mehr unterschiedliche Arten in einem Habitat zu finden sind, je höher also die Diversität ist, desto geringer ist die Wahrscheinlichkeit einer Übertragung innerhalb einer bestimmten Art [5].

In einem Szenario in dem eine spezifische Art aus einem Habitat verdrängt wird kann dessen ökologische Nische durch andere Arten besetzt werden. Dabei handelt es sich in den meisten Fällen zunächst um Generalisten, die weniger auf ihren Lebensraum und dessen Konditionen spezialisiert sind. Durch die fehlende Spezialisierung sind sie in der

Lage unempfindlicher auf Schwankungen ökologischer Faktoren (Temperatur, Feuchtigkeit oder Lichteinfall) zu reagieren und sich an sie anzupassen. Betrachtet man nun die Diversität und Viruspopulationen in einem Habitat stellt man fest, dass bei einer Erhöhung der Diversität die Virusprävalenz sinkt, spricht man von dem Verdünnungseffekt [5-8].

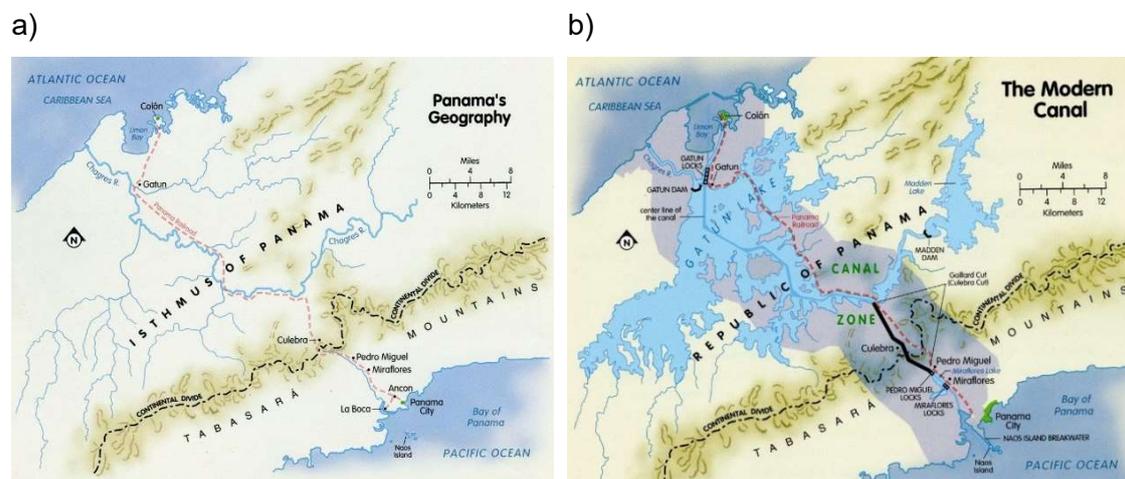
Der Verdünnungseffekt steht also im direkten Zusammenhang mit der Diversität und Populationsdichte in einem Habitat, welche ihrerseits durch menschliche Einflüsse wie Rodung oder Terraforming (darunter auch Bau künstlicher Strukturen) stark beeinflusst werden.

### **1.3. Bau des Panamakanals und dessen Einfluss auf den Urwald**

Bereits zu Beginn des 16. Jahrhunderts war eines der größten Probleme der frühen Schifffahrt die Verbindung zwischen dem Atlantischen und Pazifischen Ozean. Die kürzeste Verbindung zwischen den beiden Meeren stellte dabei eine Seestraße durch Mittelamerika dar. Viele Expeditionen wurden unternommen um die Möglichkeit eines solchen Unterfangens abzuschätzen [9]. Dabei bot sich der Isthmus (Landenge) von Panama an, da dort die geringste Distanz zwischen den beiden Ozeanen vorlag. Ein weiterer Vorteil den Panama bot, war die Tatsache, dass der Río Chagres, ein 193 Kilometer langer Fluss [10], bereits einen großen Teil des Isthmus durchzog und für ein solches Bauvorhaben genutzt werden konnte. Am Ende des 19. Jahrhunderts wurde der erste Versuch eines Bauvorhabens gestartet. Unter französischem Kommando begann Ferdinand de Lesseps 1880 mit den Bauarbeiten des Panama-Kanals. Ferdinand de Lesseps hatte sich bereits 1859 bis 1869 einen Namen, durch die Konstruktion des Suez-Kanals gemacht. Das Bauvorhaben wurde allerdings 1888 eingestellt – als Reaktion auf die großen Gefahren durch spontane Erdbeben nach heftigen Regenfällen sowie der Tatsache, dass es keine wirkungsvollen Schutzmaßnahmen gegen Gelbfieber und Malaria gab.

Unter Theodore Roosevelt startete 1904 das zweite Bauvorhaben. Trotz vieler Rückschläge und Führungswechsel der Hauptingenieure neigte sich 1911 der Bau des Panamakanals seiner Fertigstellung zu. In diesem Zeitraum begann die Überflutung des Isthmus von Panama, in der Nähe der Stadt Gatún (Abbildung 1). Durch den steigenden Wasserspiegel, veränderte sich die einst bewaldete Region stark. Als Folge entstanden

auf lokalen Erhöhungen kleinere Inseln, die nun durch den neu entstandenen Gatún See vom Festland isoliert waren.



**Abbildung 1: Darstellung des Isthmus von Panama vor und nach dem Bau des Kanals**

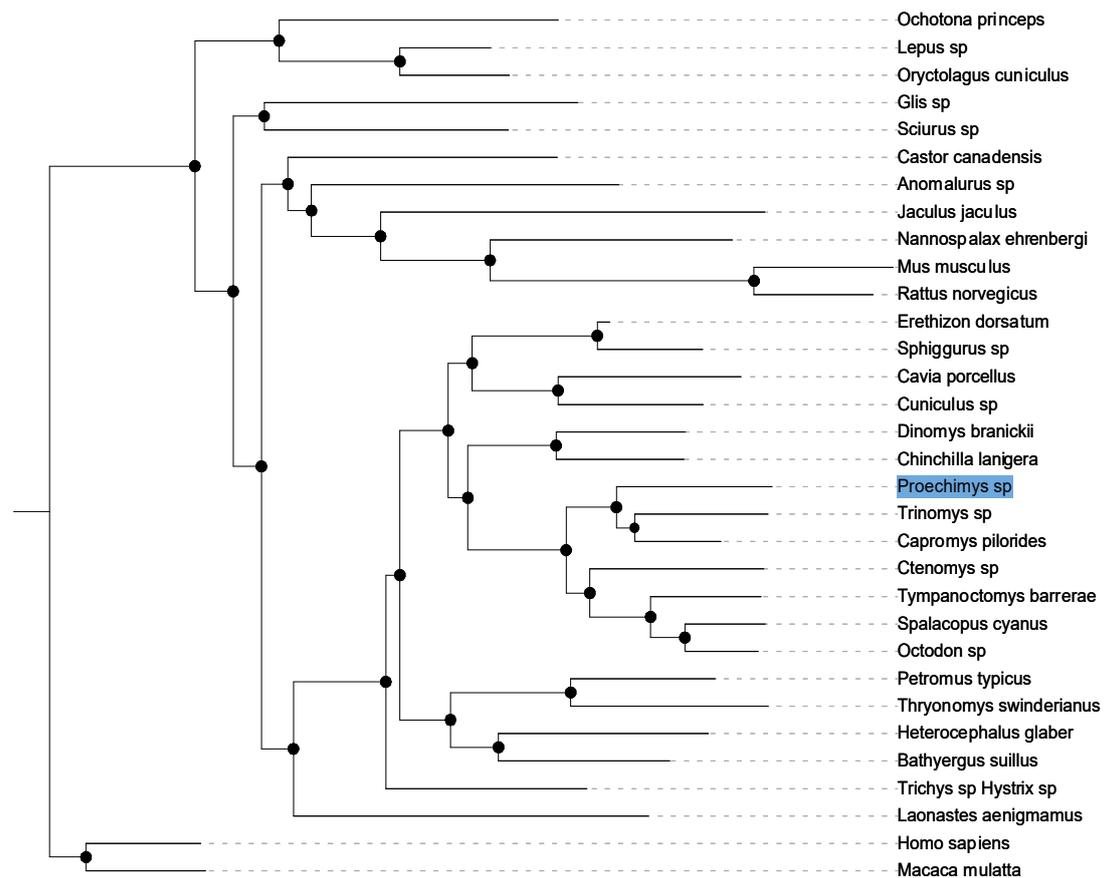
Schematische Übersichtskarten über die Panamakanal Region. Hierbei ist die Waldregion um den Chagres Fluss im nicht überfluteten Zustand gezeigt (a) sowie die Region um den Gatún See als Folge der Überflutung (b). Grafiken modifiziert von Travelark.org

## 1.4. *Proechimys semispinosus*

Durch eine solch drastische Veränderung des Lebensraumes, wie etwa durch den Bau des Panamakanals, müssen sich die einheimischen Arten sehr schnell anpassen, um zu überleben. Generalistenspezies haben meist eine höhere Toleranz gegenüber widrigen Umweltbedingungen und können aus diesem Grund in verschiedene ökologische Nischen überleben. Eine dieser weitverbreiteten Generalistenspezies in dem Gebiet um den Gatún See ist die Tomes-Stachelratte (*Proechimys semispinosus*; *P.sem*).

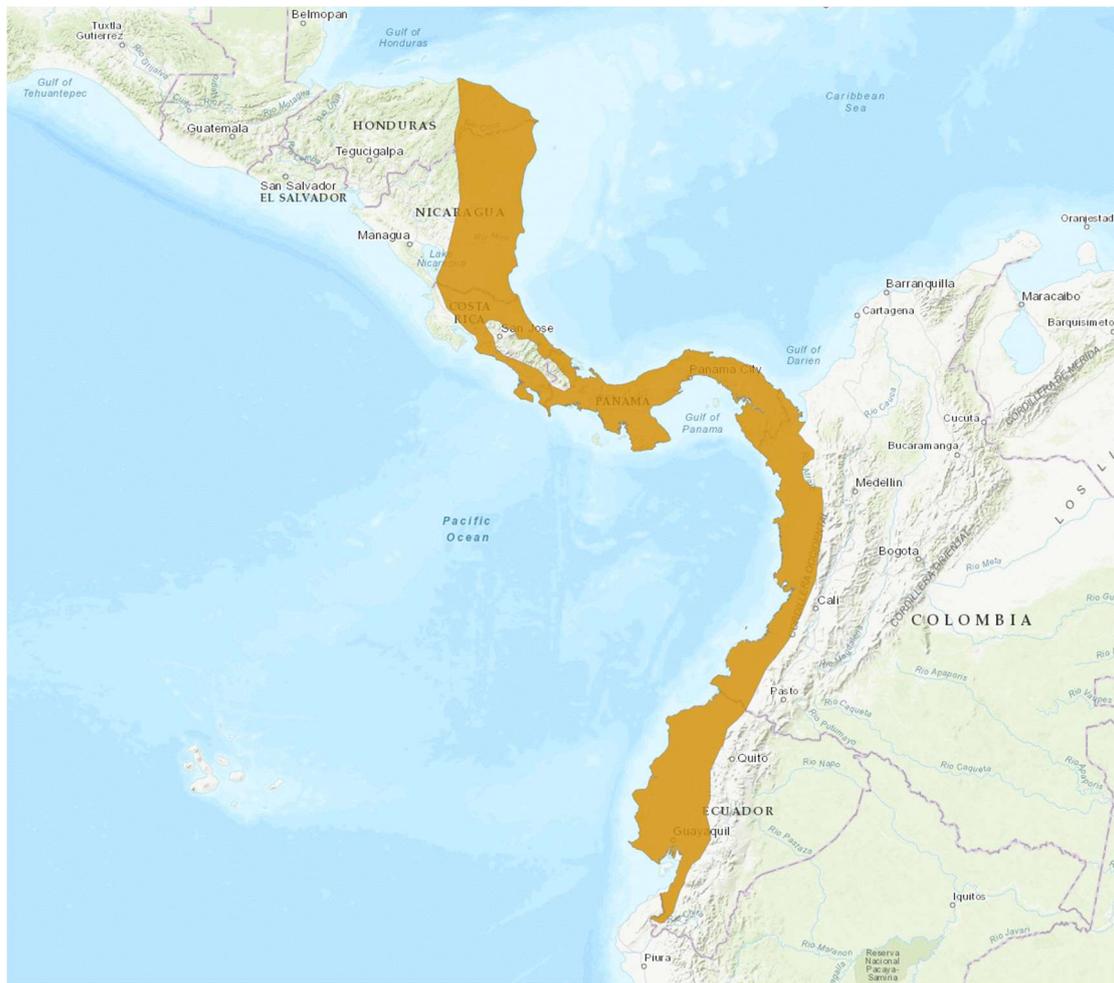
Diese wurde erstmals 1860 von Robert F. Tomes beschrieben, und dem Genus *Echimus* zugeordnet [11]. Nachdem Oldfield Thomas eine weitere Spezies, nämlich *Gymnurus*, 1897 innerhalb dieses Genus beschrieb, wurden die Unterschiede zwischen der neuen Spezies und *E. semispinosus* aufgezeigt und diskutiert [12]. Joel Allen fasste 1899 mehrere Spezies, einschließlich der Spezies *Semispinosus*, unter dem neuen Genus *Proechimys* zusammen [13]. In den folgenden Jahren wurden mehrere neue Arten beschrieben, die später als Subspezies innerhalb von *P. sem* kategorisiert wurden. Manche dieser Arten, wie zum Beispiel *P. gorgonae* (beschrieben von Outram Bangs, 1905 [14]) werden bis heute auch als Synonym für *P. sem* benutzt. *Proechimys* zählt zu den diversesten neotropischen Nagetiere [15], dessen Zuordnungen zu verschiedenen Spezies aufgrund neuer Erkenntnisse häufig geändert wurde [16]. Nach heutigem Stand

sind nur zwei Unterarten in der Gruppe *P.sem* zugeordnet: *P.sem* und *P.oconnelli* [17]. Sowohl *Proechimys*, als auch der naheverwandte Genus *Trinomys* (Abbildung 2), sind Nagetiere die primär auf dem zentral- bzw. süd-amerikanischen Kontinent anzutreffen sind (Abbildung 3 [18-20]).



**Abbildung 2: *Proechimys semispinosus* Stammbaum**

Maximum-Likelihood (ML) Baum (Substitution Model: General Time Reversible (GTR), 100 Bootstraps) eines konkatenierten Alignments aus 9 proteinkodierenden Genen von 32 Spezies. Bei den Genen handelt es sich um *ADRA2B*, *BRCA1*, *Cox1*, *Cyt-b*, *GHR*, *IRBP*, *RAG1*, *vWF* und dem mitochondrialen Genom. Das Alignment wurde in Geneious v9.1.7. [21] mit dem MAFFT Plugin Version 1.3.6 [22] generiert und der ML Baum mit dem PHYML Plugin Version 2.2.3 [23]. Schwarze Punkte markieren Knoten mit >70% Support. Der Genus *Proechimys* ist blau hervorgehoben. Als Außengruppe fungieren hier Sequenzen der Primatenspezies *Homo sapiens* und *Macaca mulatta* Alignment von Voloch 2013 [24], modifiziert.

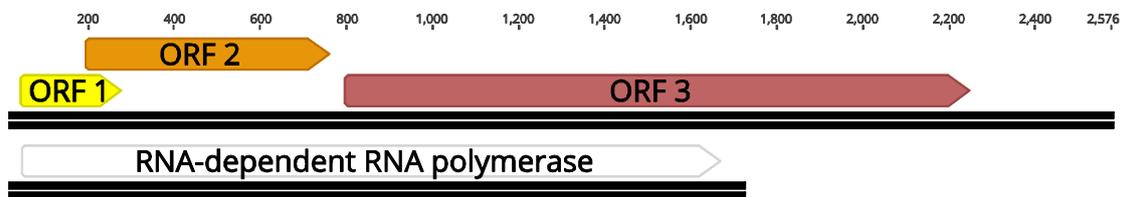


**Abbildung 3: Geografische Verbreitung von *Proechimys semispinosus***  
Ausschnitt Zentralamerikas mit Markierung über dem beobachteten Lebensraum von *P.sem.*. Das Gebiet erstreckt sich vom Osten Honduras bis zum Südwestens Ecuador (Stand Feb 2016).  
Die Informationen entstammen der „IUCN Red List of Threatened Species“  
Quelle: <https://www.iucnredlist.org/species/18297/22208264>

## 1.5. Einführung in die behandelten Viren

### 1.5.1. Picobirnavirus Grundlagen

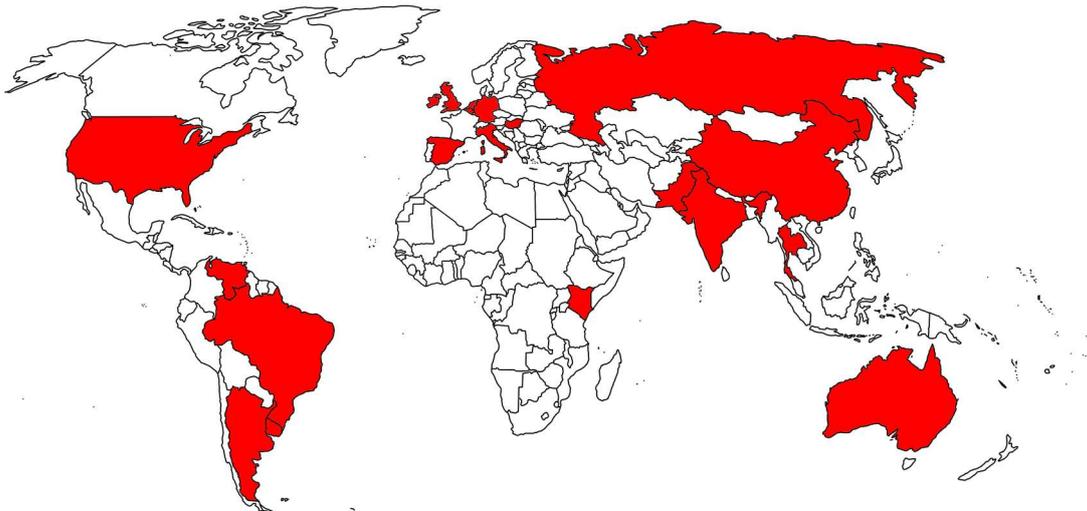
*Picobirnavirus* (PBV), das einzige Genus in der Familien *Picobirnaviridae*, umfasst nicht-umhüllte, doppelsträngige RNA Viren, mit einem zweigeteilten Genom. Der Name Picobirnavirus leitet sich aus dem Lateinischen ab und beschreibt die kleine („pico“) zweigeteilte („bi“) RNA Genomorganisation des Virus. Segment 1, das größere beider Genomsegmente, ist 2,3-2,6 Kilobasenpaare (kbp) groß [25] und kodiert drei offene Leserahmen (*open reading frames*; ORFs). Die Funktionen der Proteine, die von ORF1 und ORF2 kodiert werden, sind bislang unbekannt. ORF3 kodiert das Capsidprotein von Picobirnaviren. Segment 2, das eine Größe von 1,5-1,9 kbp [25] hat, kodiert einen ORF, der die RNA-abhängige RNA-Polymerase (RdRp) kodiert (siehe Abbildung 4).



**Abbildung 4: Vereinfachte Darstellung der Genomorganisation von Picobirnaviren**

Gezeigt ist das zweigeteilte Picobirnavirusgenom. Segment 1 (oben) kodiert ORF 1, ORF 2 und ORF 3, während Segment 2 (unten) ausschließlich die RdRp kodiert.

Innerhalb des Genus *Picobirnavirus* befinden sich drei Genotypen, von denen Genotyp 1 und 2 in Wirbeltieren beschrieben wurden [26] und Genotyp 3 in Wirbellosen [27-29]. Die Einteilung in die Genotypen, nach Angaben des internationalen Komitees für Virustaxonomie (ICTV), erfolgt anhand der Übereinstimmung in der Aminosäuresequenz der RdRp. Das Virus ist, seit der Entdeckung 1988 [30], in einer großen Bandbreite [31] von Vertebraten [32-39], Invertebraten [40] und Abwässern [41, 42] detektiert worden. Diese weite Verbreitung beschränkt sich allerdings nicht nur auf die Viruswirte, sondern auch auf die geografische Ausbreitung [43, 44] (siehe Abbildung 5).



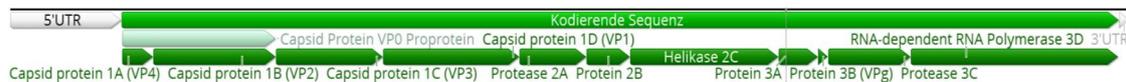
**Abbildung 5: Übersicht der Verbreitung von Picobirnaviren weltweit**

Weltkarte mit Ländern (rot markiert) in denen *Picobirnaviren* detektiert wurden. Daten entstammen Ganesh *et. al* [31]

Der Nachweis von PBV in humanen und tierischen Kotproben, deutet darauf hin, dass diese Viren mit dem Stuhl ausgeschieden werden und möglicherweise Gastroenteritis-Symptome auslösen können [45-47]. Dennoch konnte bislang kein Beweis, über die tatsächliche Pathogenese des Virus erbracht werden, da weder Tiermodelle noch Zellkultursysteme für die Anzucht existieren.

### 1.5.2. Picornavirus Grundlagen

Die Familie *Picornaviridae* umfasst 35 Genera, darunter zum Beispiel die Genera *Hunnivirus*, *Cardiovirus* und *Enterovirus* [48]. Alle Mitglieder dieser Virusfamilie haben gemein, dass es sich bei ihnen um nicht-umhüllte Viren mit einer einzelsträngigen RNA (*single-stranded* RNA, ssRNA) positiver Orientierung handelt. Der Name leitet sich, ähnlich wie bei den Picobirnaviren, vom lateinischen Wort „pico“ für „klein“ und RNA auf Basis des Genoms ab, mit einer Länge von etwa 7,5 kbp, ab [49, 50] (siehe Abbildung 6).



**Abbildung 6: Vereinfachte Darstellung der Genomorganisation von Picornaviren**

Genomorganisation mit 3' und 5' UTR, kodierende Regionen für die Capsidproteine 1A, 1B, 1C und 1D sowie Proteinen 2A (Protease), 2B, 2C (Helikase), 3A, 3B, 3C (Protease) und 3D (RdRp).

Das Virion der Picornaviren besitzt ein Capsid, welches aus vier Proteinen aufgebaut ist: 1A, 1B, 1C und 1D [49]. Dabei wird die Capsidoberfläche von 1B, 1C und 1D gebildet, indem sie ein Hetero-trimer formen. Das Protein 1A verbindet und stabilisiert fünf dieser Hetero-trimere innerhalb des Capsids [51, 52]. In manchen Viren verbleiben die Proteine 1A und 1B ungespalten als Protein 1AB. Dies übt dann dennoch die gleiche Funktion, die Bildung des Capsids, aus [48].

Bei den Nicht-strukturproteinen 2A und 3C handelt es sich um virale Proteinasen, die Vorläuferproteine proteolytisch spalten und in funktionale Untereinheiten unterteilen. Zusätzlich inhibieren diese viralen Proteine die Proteinbiosynthese der Wirtszelle [53-55].

Das Proprotein 2BC, bestehend aus den Untereinheiten 2B und 2C, inhibiert Sekretionswege der Wirtszelle, kann die Membranpermeabilität erhöhen und Golgi-Vesikel zu bilden [56-58]. Darüber hinaus besitzt 2C weitere Funktionen als Helikase, Autophagie-Inhibitor und Helfer in der Morphogenese des Virus [59].

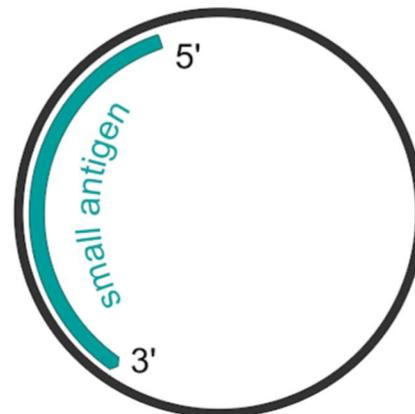
Die Replikation wird durch den Replikationskomplex gesteuert, welcher aus den Proteinen 3A, 3B und 3D besteht. Dabei sorgt ein Hetero-dimer aus 3A und 3B für die Bindung an das endoplasmatische Retikulum und 3D ist eine als RdRp und führt die Transkription durch [60, 61].

Die taxonomische Vielfalt der Familie *Picornaviridae* spiegelt sich auch im Spektrum der verursachten Erkrankungen wider. Von leichten Erkältungssymptomen, über Durchfall und Schleimhautentzündungen, können dabei auch schwere Leberentzündungen, Lähmungen und Infektionen des Zentralnervensystems auftreten. Da die Viren keine Lipidhülle besitzen, sind sie sehr resistent gegenüber organischen Lösungsmitteln. Außerdem überstehen viele Viren aus dieser Familie pH-Werte unter 3, so dass sie den Magen unbeschadet passieren und von der Darmschleimhaut aufgenommen werden können [62].

### 1.5.3. Deltavirus Grundlagen

Das Genus der *Deltaviren* beinhaltet bislang nur die Spezies *Hepatitis Delta Virus* (HDV) und ist keiner Familie zugeordnet. HDV besitzt ein zirkuläres, ssRNA Genom in negativer Orientierung mit einer Größe von etwa 1,7 kbp [63].

Im Virion wird das Genom von circa 70 Kopien des Hepatitis-Delta-Antigens (HDAg) [64] umgeben, welches auf dem Genom kodiert wird. Dabei existieren zwei unterschiedliche Formen dieses Antigens, L-HDAg (*large*) und S-HDAg (*small*). Auf dem Genom befindet sich ein einziger ORF, der das S-HDAg kodiert (Abbildung 7). Der ORF für das L-HDAg entsteht durch eine posttranskriptionale RNA Editierung am Stopcodon des ORFs für das S-HDAg, wodurch dieser verlängert wird und weitere 19 Aminosäuren (AA) zusätzlich kodiert [65, 66].

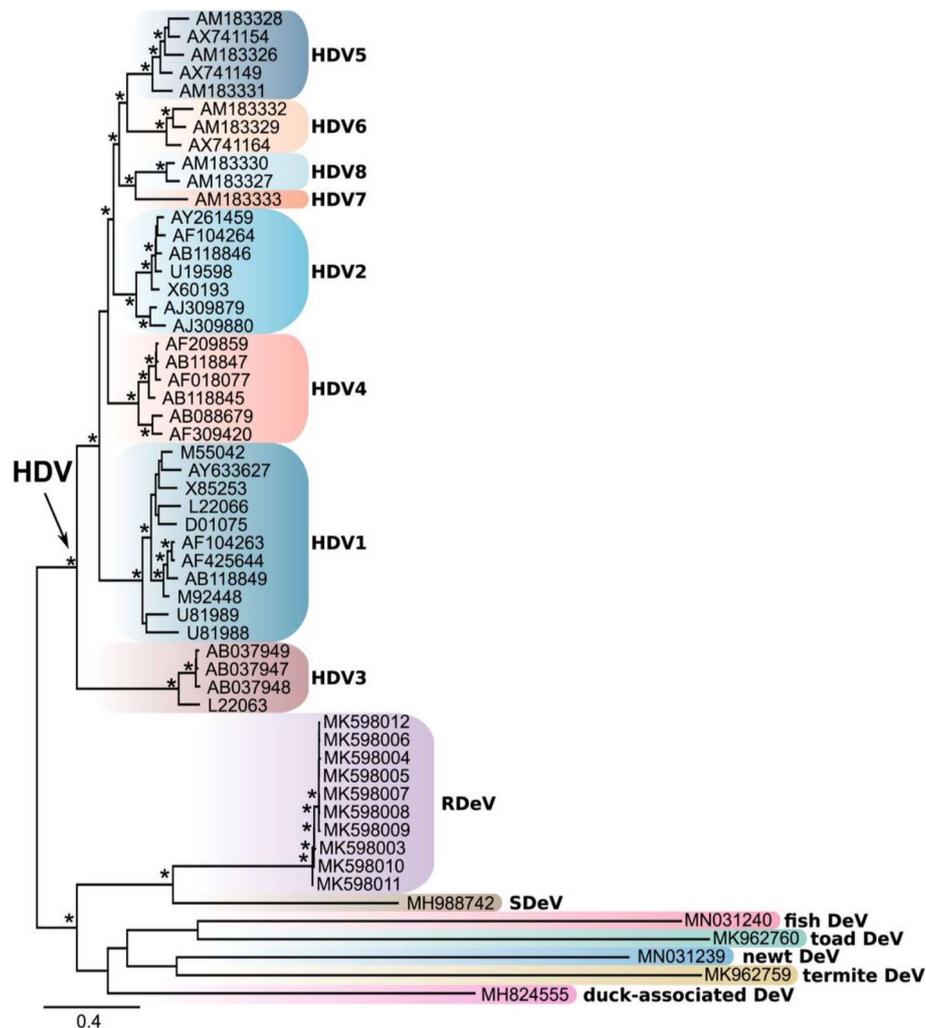


**Abbildung 7: Vereinfachte Darstellung der Genomorganisation von HDV**  
Zirkuläres Genom mit kodierender Region für das kleine Antigen (S-HDAg) und dessen Orientierung. Das große Antigen (L-HDAg) entsteht durch eine posttranskriptionale Modifikation der RNA von S-HDAg und ist hier nicht zu sehen.

Bei HDV handelt es sich um ein Virus, das auf eine Koinfektion eines Helfervirus angewiesen ist, um infektiöse Viruspartikel zu generieren. Solche Viren bezeichnet man auch als Satelliten-Viren [67]. Diese Viren sind nicht imstande eigenständig infektiöse Viruspartikel zu assemblieren. Im Falle von HDV wird das Hepatitis B Virus (HBV) als Helfervirus benötigt. Das Ribonukleinprotein, bestehend aus dem Genom und mehreren Kopien des HDAg, bindet innerhalb des endoplasmatischen Retikulums an dort vorliegende Hepatitis-B-Oberflächenantigene (*Hepatitis B surface antigenes*; HBsAg)

und bildet damit das Virion bevor es per exosomalen Transport aus der Zelle geleitet wird [68].

Bis 2018 war neben dem Menschen kein anderer Wirt für Deltaviren bekannt. Dies änderte sich jedoch mit dem Virusfund in Vögeln [69], Schlangen [70], Fischen, Amphibien und Invertebraten [71] sowie Nagetieren (diese vorliegende Arbeit) aus verschiedenen Transkriptomdaten (Abbildung 8)



**Abbildung 8: Phylogenetischer Stammbaum von Deltaviren in verschiedenen Wirten**

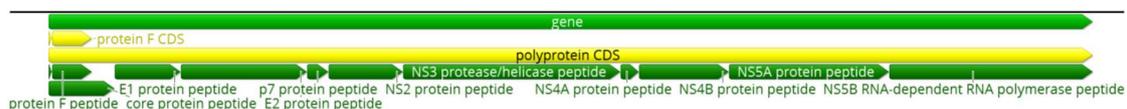
Maximum-Likelihood (ML) Baum eines Vollgenom-Alignments der acht verschiedenen humanen HDV Genotypen (HDV 1-8), sowie Vertretern nicht menschlicher Deltaviren (Nagetier Deltavirus [RDeV], Schlangen Deltavirus [SDeV], Enten-assoziiertes Deltavirus [duck-associated DeV], Molch Deltavirus [newt DeV], Kröten Deltavirus [toad DeV], Fisch Deltavirus [fish DeV], und Termiten Deltavirus [termite DeV]). Errechnet wurde der Baum mit RAXML-NG Version 0.7.0 BETA [72]). Sterne markieren Bootstrapsupportwerte >80% und die Äste zeigen die jeweiligen GenBank-Accession numbers. Die Wurzel wurde an dem Knoten der nicht menschlichen Deltaviren gesetzt.

#### 1.5.4. Hepacivirus Grundlagen

Das *Hepacivirus* ist, neben *Pestivirus*, *Flavivirus* und *Pegivirus*, eines von vier Genera der Familie *Flaviviridae*. Bei allen Mitgliedern dieser Familie handelt es sich um umhüllte Viren, mit einem einzelsträngigen RNA-Genom positiver Orientierung, welches ca. 9-13 kbp [73] groß ist. Es besteht aus einer 5' nicht-kodierenden-Region (NCR), einer kodierenden Region (CDS) für Strukturproteine, einer CDS für Nicht-Strukturproteine und einer 3'NCR [73] (siehe Abbildung 9).

Die 5'NCR, enthält eine interne ribosomale Eintrittsstelle (IRES) und ist 340 bp lang. IRES sind sekundäre Strukturen der Messenger-RNA (mRNA), die die Bindung an die Ribosomen vermitteln und dadurch die Translation ermöglichen. Bei den drei Strukturproteinen, die auf dem Genom kodiert sind handelt es sich um das Kernprotein C und zwei Hüllproteine E1 und E2. Das Kernprotein C formt, zusammen mit der viralen RNA, das Nukleocapsid [74], welches dann von einer Lipiddoppelschicht und den Hüllproteinen E1 und E2 umschlossen wird und das Virion bildet [75].

Die sechs Nicht-Strukturproteine (NS2, NS3, NS4a, NS4b, NS5a und NS5b) erfüllen wichtige Funktionen in der Generierung von Viruspartikeln und der Regulation der Translation beziehungsweise Replikation [76]. Zusammen mit NS2 hat NS3 proteolytische und autoproteolytische Funktionen [77]. NS3 kann, sowohl als Monomer als auch in einem Komplex mit NS4a, als Helikase fungieren [78]. NS4a und NS4b modifizieren gemeinsam den Vesikeltransport vom endoplasmatischen Retikulum zum Golgi-Apparat, wodurch sogenannte membranöse Netze („*membranous web*“) entstehen [79, 80]. NS5b ist eine RdRp und für die Replikation unablässig, weswegen die kodierende Sequenz hoch konserviert ist.



**Abbildung 9: Vereinfachte Darstellung der Genomorganisation von Hepatitis C Viren**

Genomorganisation mit codierenden Regionen für das Kernprotein C, Protein F, zwei Hüllproteine E1 und E2, sowie den sechs nicht-Strukturproteinen p7, NS2, NS3, NS4a, NS4b, NS5a und NS5b

Der Name Hepacivirus leitet sich vom griechischen Wort „*Hepar*“ ab, was Leber bedeutet. Die Viren dieses Genus infizieren die Leber und können chronische Entzündungen (Hepatitis) auslösen [81].

Laut WHO waren 2015 noch 71 Millionen Menschen mit dem Hepatitis-C-Virus (HCV) infiziert [82]. Vor 1990 bestand die größte Ansteckungsgefahr in der Gabe von Blutkonserven, da Hepatitis C noch unbekannt war und deswegen nicht darauf getestet wurde. Seit der Entdeckung eines neuen Virusstammes im Jahre 1989, der weder dem Hepatitis A- noch dem Hepatitis B Virus entsprach und deswegen die vorläufige Bezeichnung non-A-non-B Hepatitis (NANBH) erhielt [83, 84], wurden Blutkonserven routinemäßig auf das neue Virus getestet, worauf das Ansteckungsrisiko mit NANBH (später Hepatitis C Virus, HCV) durch Transfusionsblut gegen null sank.

Das Ausmaß, einer Infektion mit HCV, kann dabei von einer leichten nichtprogressiven Leberschädigung bis zu einer chronischen Hepatitis C reichen, wobei die Akutphase meist symptomarm verläuft und deswegen oft unerkannt bleibt [85]. Von einer chronischen Hepatitis C Infektion spricht man, wenn die akute Phase länger als 6 Monate andauert und die virale RNA über diesen Zeitraum nachweisbar ist. Die chronische Hepatitis C zeichnet sich dadurch aus, dass sich bei den Betroffenen eine Leberzirrhose bildet, die zu Leberversagen oder zum hepatozellulären Karzinom führen kann.

Neben HCV, welches Menschen infiziert, gibt es 13 weitere Spezies in dem Genus *Hepacivirus*, die Nagetiere, Fledermäuse, Rinder und andere Säugetiere infizieren [86-89]. Diese Virusspezies sind eingeteilt in Hepacivirus A bis Hepacivirus N auf Grundlage ihrer Diversität und phylogenetischen Verwandtschaft innerhalb des Genus.

#### 1.5.5. Zielsetzung der vorliegenden Arbeit

Vorangegangene Analysen, ließen vermuten, dass es im Naturschutzgebiete um den Gatún See, eine breites Spektrum an Viren gab. Das Ziel der vorliegenden Dissertation sollte es sein, diese, teils noch unbekanntes Virusspezies molekularbiologisch und phylogenetisch zu charakterisieren. Weiterhin sollte untersucht werden, in wie weit ökologische Faktoren Einfluss auf die Viruspopulationen haben. Logisch anknüpfend an diesen Punkt, wurde untersucht, ob die Überflutung, und die damit einhergehende Landschaftsänderung langfristige Folgen hatte, die sich bis heute in den ansässigen Virusspezies widerspiegelt. Um dies zu untersuchen wurden verschiedene Sequenzierungsmethoden genutzt und daraus die, daraus erhaltenen Informationen bioinformatisch aufbereitet. Die Ergebnisse dieser Arbeit sollen Hinweise darauf geben, in wie weit das Infektionspotential, von möglicherweise zoonotischen Erregern, durch äußere Faktoren beeinflusst werden kann.

## 2. Material und Methoden

### 2.1. Benutzte Materialien

#### 2.1.1. Chemikalien

<b>Chemikalie</b>	<b>Vertrieb</b>
1 kb plus DNA Ladder	Life Technologies
100 bp DNA Ladder	Life Technologies
Agarose	Sigma-Aldrich Chemie GmbH
Borsäure	VWR International GmbH
Bovines Serumalbumin (BSA)	Roche Diagnostics GmbH
Bromphenolblau	Carl Roth GmbH + Co. KG
Carrier-RNA	QIAGEN
dNTP Set, 100 mM	Life Technologies
Ethylendiamintetraessigsäure (EDTA)	AppliChem
Kanamycin	Sigma-Aldrich Chemie GmbH
LB-Agar (Lennox)	Carl Roth GmbH + Co. KG
LB-Medium (Lennox)	Carl Roth GmbH + Co. KG
Magnesiumchlorid (MgCl <sub>2</sub> )	Life Technologies
Magnesiumsulfat (MgSO <sub>4</sub> )	Sigma-Aldrich Chemie GmbH
Midori Green Gel staining	NIPPON Genetics Europe
Phosphatgepufferte Salzlösung (PBS)	Life Technologies
RNase freies Wasser (Rfa), AccuGENE	Lonza Cologne GmbH
RNAlater	Sigma-Aldrich Chemie GmbH
Saccharose	Carl Roth GmbH + Co. KG
Tris(hydroxymethyl)-aminomethan (Tris)	Merck KGaA

#### 2.1.2. Kits

<b>Kitname</b>	<b>Vertrieb</b>
ExoSAP-IT® For PCR Product Cleanup	Affymetrix USB®
MagNA Pure 96 DNA and Viral NA Small Volume Kit	Roche Diagnostics GmbH

MEGAscript® T7 Transcription Kit	Life Technologies
Platinum® Taq DNA Polymerase	Life Technologies
QIAGEN OneStep RT-PCR Kit	QIAGEN
RNeasy® Mini Kit	QIAGEN
SuperScript® III One-Step RT-PCR System with Platinum® Taq DNA Polymerase	Life Technologies
TOPO® TA Cloning® Kit for Sequencing, with One Shot® TOP10 Chemically Competent E.coli	Life Technologies

## 2.2. Software

Programm	Version	Entwickler
Geneious	9.1.7.	Biomatters Ltd.
MEGA	7.0.21	Kumar <i>et. al.</i> [90]
BEAUTI & BEAST	1.10.4	Suchard <i>et. al.</i> [91]
Tracer	1.7.1	Suchard <i>et. al.</i> [91]
Logcombiner	1.10.4	Suchard <i>et. al.</i> [91]
TreeAnnotator	1.10.4	Suchard <i>et. al.</i> [91]
FigTree	1.4.4	Rambaut <i>et. al.</i>
GraphPad Prism	5.03	GraphPad Software, Inc.
RDP4	4.95	Martin <i>et. al.</i> [92]
R	4.2.0	R Core Team (2022)
R Studio	2022.02.2 Build 485	RStudio Team (2022)
SpredD3	0.9.6	Bielejec <i>et. al.</i> [93]
TempEst	1.5.3.	Rambaut <i>et. al.</i> [94]
jModeltest	2.1.10	Darriba <i>et. al.</i> [95]
Office	Profession Plus 2016	Microsoft

## 2.3. Benutzte Geräte

Gerätetyp	Gerätename	Vertrieb
Automatische DNA/RNA-Extraktion	MagNA Pure 96	Roche Diagnostics GmbH

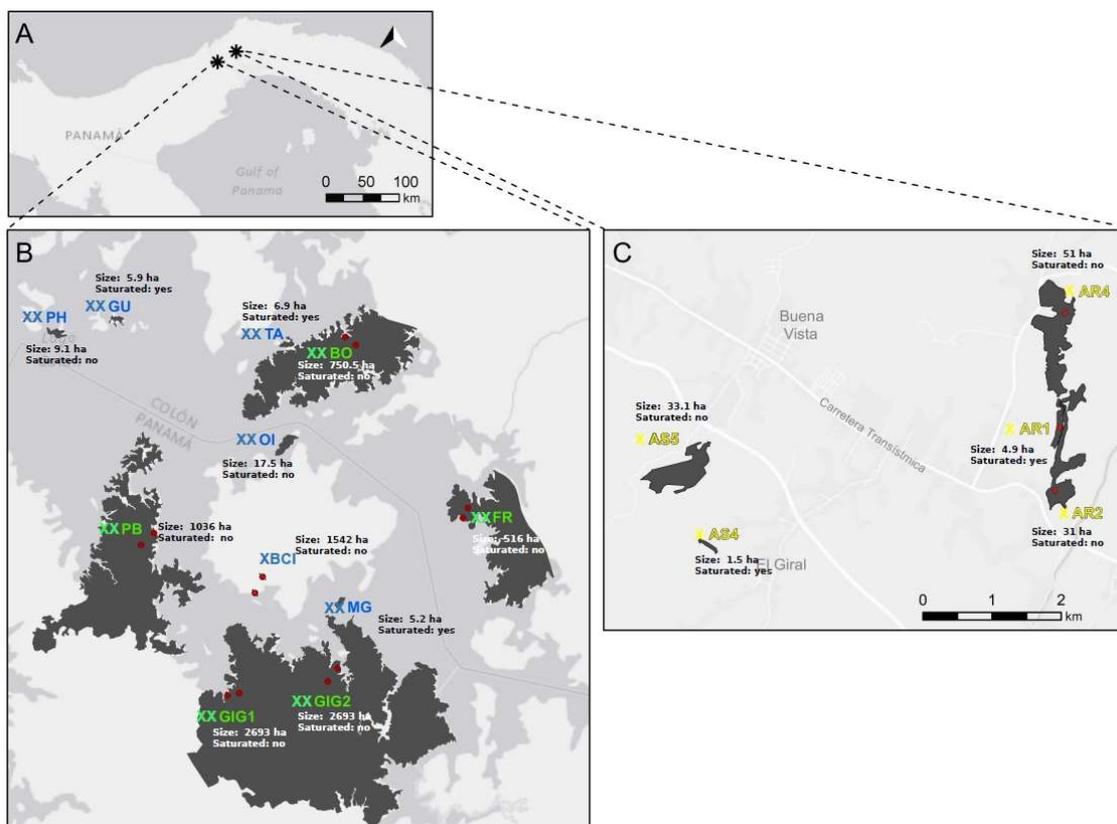
Feinwaage	LC2201P	Sartorius AG
Gelelektrophoresekammer	PerfectBlue Gelsystem Maxi S	PEQLAB Biotechnologie GmbH
Geldokumentation	Gel Doc XR+	Bio-Rad Laboratories
Geldokumentation, Drucker	P93DW	Mitsubishi Electric Europe B.V.
Heizblock	Thermomixer comfort	Eppendorf
Inkubator	Heraeus® B6126	Thermo Fisher Scientific GmbH
Stromverteiler für Gelelektrophorese-Kammern	Power Supply P25	Biometra GmbH
Pipetten	Research® plus variable 0,1-2,5 µL 0,5-10 µL 2-20 µL 10-100 µL 20-200 µL 100-1000 µL  Research® plus Multichannel Pipettes variable (8-channel) 0,5-10 µL 10-100 µL	Eppendorf
PCR-Cycler	Mastercycler ep Gradient S	Eppendorf
Real-Time PCR-Cycler	LightCycler® 480 II	Roche Diagnostics GmbH
Sicherheitswerkbank	Hera Safe KS12	Thermo Fisher Scientific GmbH
Spektralphotometer	NanoDrop 1000	PEQLAB Biotechnologie GmbH
Kühlschrank	FKUv 1610-22Q	Liebherr
Abzug	TA-1200	Hemling
Homogenisator	TissueLyser II	QIAGEN

Tiefkühlschrank (-20 °C)	GGU 1500	Liebherr
Tiefkühlschrank (-80 °C)	CryoCube F740	Eppendorf
Wasserbad	Typ 1002	GFL
Vortexer	VV3	VWR International GmbH
Tischzentrifuge	Centrifuge 5424	Eppendorf
Plattenzentrifuge	Centrifuge 5430	Eppendorf

## 2.4. Methoden

### 2.4.1. Fangprozedur und Probenahme

Die Probesammlung fand in Zentral Panama im Gebiet des Panamakanals statt. Dabei wurden verschiedenste Nagetiere in drei unterschiedlichen Habitattypen untersucht, wobei sich die Habitate in ihrem Grad an Einflüssen oder Veränderungen, sowohl menschlich als auch umweltbedingt, unterschieden. Bei den drei Habitattypen handelte es sich dabei um Flachland mit zusammenhängenden tropischen Wäldern um den Gatún See (Habitattyp C; continuous forest; 489 – 2694 ha), kleinere Waldabschnitte in langwirtschaftlich-genutzten Zonen (Habitattyp A; agricultural landscape; 1,5 – 51 ha) und isolierte, bewaldete Inseln (Habitattyp I; forested island; 5,2 – 17,5 ha). Der in dieser Arbeit genutzte 5-Buchstaben Code setzt sich zusammen aus dem Habitattyp (A,C,I,P), welcher am Anfang genannt wird, sowie einem anschließenden Habitatcode (Abbildung 10). Zwischen dem Habitattypcode und dem Habitatcode wird mit X aufgefüllt um sicherzustellen, dass jeder zusammengesetzte Code eine Länge von 5 Buchstaben hat.



**Abbildung 10: Geografische Lage der Fangorte und Habitate um den Gatún See**

Gezeigt sind Habitate innerhalb der drei großen Habitattypen: Wälder (grüne Schrift), Agrikulturgebiete (gelbe Schrift) und Inseln (blaue Schrift). Plantagen sind aufgrund ihrer geringen Anzahl nicht dargestellt. Die Grafik enthält außerdem Informationen über Größe der Habitate sowie die Information, ob eine Übersättigung der Fallen (viele Wiederfänge) stattgefunden hat.

Bei den kleinen, bewaldeten Inseln handelt es sich um ehemalige Berg- oder Hügelgipfel, die durch die Überflutung in Folge des Panamakanalbaues vor über 100 Jahren (1911), zu den heutigen Inseln wurden. Bei den momentan ansässigen terrestrischen Tierarten handelt es sich wahrscheinlich um Nachfahren von Tieren, die zum Überflutungszeitpunkt, bereits in diesem Areal beheimatet waren. Es ist nicht auszuschließen, dass es auch zu einer Wiederbesiedlung durch Tiere gekommen ist, die z.B. auf Pflanzenmaterial von benachbarten Inseln oder vom Festland schwammen. Alle Untersuchungsgebiete in den Habitattypen C und I stehen seit über 60 Jahren durch das Barro Colorado Nature Monument (BCNM) unter Naturschutz. Im Gegensatz dazu befinden sich die Waldfragmente im Habitattyp A in einem Gebiet, das in den letzten 60 Jahren, beginnend mit dem Bau der Transistmica-Autobahn in den 1950er Jahren, starken Landnutzungsänderungen und einer Zunahme der menschlichen Bevölkerung unterworfen war [96, 97]. Die Entfernung zwischen den Untersuchungsstandorten der Habitattypen C und I zum Habitattypen A beträgt ca. 25 km. Die klimatischen Bedingungen unterscheiden sich zwischen den Habitattypen nicht. Innerhalb jedes Habitattypen wurden fünf Studienorte untersucht, die als unabhängige Replikate dienen. Der Lebendfang von Kleinsäugetern wurde während drei Feldsaisonen durchgeführt (Saison 1: Okt 2013 - Mai 2014, Saison 2: Okt 2014 - April 2015, Saison 3: Juni 2016 – April 2017). Jedes Untersuchungsgebiet wurde einmal pro Feldsaison beprobt, wobei 100 gleichmäßig verteilte Fangstationen in Abständen von 20 m entlang paralleler Fanglinien in fünf aufeinander folgenden Nächten aufgestellt wurden. Jede der 100 Fangstationen bestand aus drei Fallen, die in einem Bereich von etwa 5 m aufgestellt wurden (d.h. insgesamt 300 Fallen/Studienstandorte). Eine Tomahawk-Falle (15,2 cm × 15,2 cm × 48,3 cm, livetraps.com) und eine Sherman-Falle (10,2 cm × 11,4 cm × 38,1 cm, shermantraps.com) wurden auf dem Boden aufgestellt. Weiterhin wurde sofern möglich eine Sherman-Falle in einer Höhe von 0,5-2,5 m an einem Lianen- oder Baumast angebracht. Drei der Untersuchungsgebiete (Habitattyp A: Untersuchungsgebiet AS4, Habitattyp I: Untersuchungsgebiete MG und GU) waren zu klein, um 100 Fangstationen unterzubringen, was zu einer geringeren Anzahl von Fallennächten führte. Die in den Habitattypen A und I beprobten Habitatfragmente waren so klein, dass selbst an den innersten Punkten die gleichen Umweltbedingungen wie an den Rändern herrschten und daher Randeffekte unwahrscheinlich waren. Im Gegensatz dazu waren in Habitattyp C Unterschiede in den Umweltbedingungen zwischen der Küstenlinie und dem Inneren großer Waldflächen theoretisch möglich. Um einen möglichen Randeffekt in den Untersuchungsgebieten des Habitattyp C zu testen, wurden 50 Fangstationen in Fanglinien entlang des Waldrandes (d.h. entlang der

geschützten Panamakanal-Uferlinie) und so weit wie möglich im Waldinneren (ca. 300 m) eingesetzt. Gefangene Tiere konnten anhand morphologischer Merkmale (z.B. Gewicht, Schienbeinlänge, Schwanzlänge) mit Ausnahme von *Marmosa sp.* leicht identifiziert werden. Zudem wurden für die Stachelratte *Proechimys semispinosus*, die der Fokus dieser Forschung war, auch das Alter (adulte oder jugendliche Ratte) anhand von Pelagemerkmalen [98] sowie das Geschlecht bestimmt. Für das Virus-Screening wurden Blutproben (10-75 µL) aus einem kleinen Ohrschnitt oder durch Venenpunktion der Schwanzvene mit EDTA-beschichteten Kapillaren entnommen (75 µL, Kabe, Deutschland). Etwaige Kotpellets, die die Tiere in der Falle hinterließen, wurden gesammelt. Zur Identifizierung von Wiederfängen während des Probennahmezeitraums an einem Studienort (d.h. fünf aufeinanderfolgende Nächten) wurden alle Tiere individuell markiert, indem ihnen Haare vom Rumpf abgeschnitten und die Haut mit Tiermarkierungsstäbchen gefärbt wurde (Raidex GmbH, Dettingen/Erms, Deutschland). Nach der Daten- und Probenerhebung wurden die Tiere am Fangort freigelassen. Individuen, die in mehreren Feldsaisons (n = 9) gefangen wurden, konnten aufgrund des Vorhandenseins des abgeschnittenen Ohres, von dem zuvor Blutproben entnommen wurde, als Wiederfänge identifiziert werden und wurden bei allen Analysen ausgelassen, um Pseudoreplikationen zu vermeiden. Alle Arbeiten wurden mit voller ethischer Genehmigung durchgeführt (Smithsonian IACUC-Protokoll 2013-0401-2016-A1-A7). Die Proben wurden mit Genehmigung der panamaischen Regierung nach Deutschland exportiert (SE/A-21-14, SE/A-69-14, SEX/A-22-15).

#### 2.4.2. RNA Extraktion mittels MagNA Pure 96

Bevor Nukleinsäuren untersucht werden, werden sie isolieren und aufgereinigt, um Substanzen zu entfernen, die nachfolgenden Analysen stören könnten. Dazu wurden für die gesammelten Kot- und Blutproben der MagNA Pure 96 von Roche benutzt. Bei dem MagNA Pure 96 handelt es sich um ein Instrument zur automatischen Extraktion von RNA und DNA mittels Silicabeads. Zunächst wird das Probematerial mit Lysispuffer behandelt, um Zellen bzw. Viruspartikel aufzuschließen. Im Anschluss können die freigelegten Nukleinsäuren an Silicabeads binden. Diese Bindung wird durch einen Bindungspuffer, der eine hohe Konzentration an chaotropen Salzen besitzt, begünstigt. Die Silicabeads besitzen einen magnetischen Kern, weswegen sie in anschließenden Waschvorgängen durch Magneten an der Reaktionsgefäßwandung gehalten werden können. Durch die Zugabe eines Elutionspuffers, der eine geringere Salzkonzentration

hat als die vorangegangenen Puffer, werden die Nukleinsäuren von den Beads gelöst und der Extraktionsvorgang abgeschlossen.

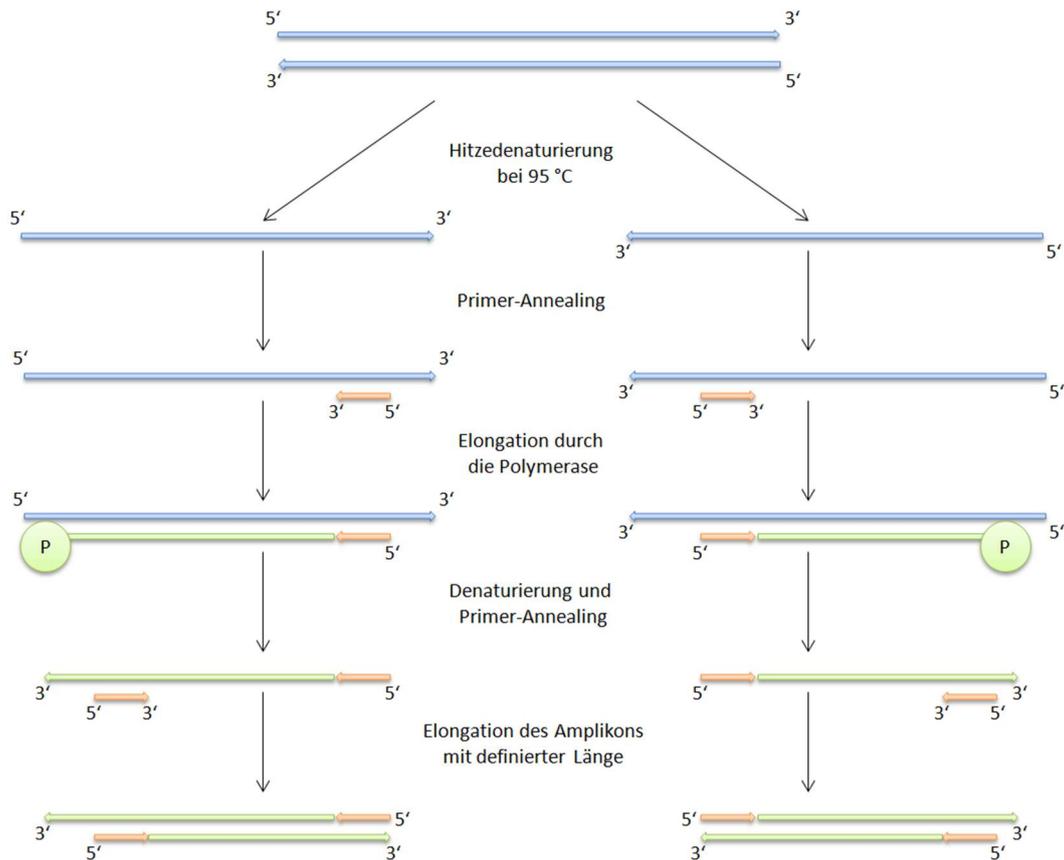
Für die Extraktion der totalen RNA aus Kot wurden die Proben (2-3 Kotpellets in RNAlater) im TissueLyser II homogenisiert. In einer MagNA Pure 96 Deep-well Platte wurden 180 µL PBS vorgelegt und mit 20 µL des Kot-Homogenisats vereint. Die Extraktion erfolgte unter Verwendung des DNA and Viral NA Small Volume Kits mit einem Elutionsvolumen von 100 µL.

Die Extraktion aus Blutproben erfolgte analog zur Extraktion der Kotproben mittels DNA and Viral NA Small Volume Kit. Blutproben wurden bei der Entnahme nicht in RNAlater aufgenommen. Bei manchen Probenentnahmen, bei denen eine Kapillare benutzt wurde, um ca. 5 µL Blut zu sammeln, verblieb die Kapillare im Reaktionsgefäß und wurde später mit 50 µL PBS ausgespült. 10 µL dieses verdünnten Blutes wurden mit 190 µL PBS in einer Deep-well Platte vereint und extrahiert. Wurde bei der Blutentnahme bereits die Kapillare entleert, wurden für die RNA Extraktion nur 2 µL des Vollblutes für die Extraktion benutzt und mit PBS auf 200 µL Extraktionsvolumen aufgefüllt.

Im Anschluss wurden jeweils Extrakte des gleichen Ausgangsmaterials in Pools vereinigt.

#### 2.4.3. Detektion viraler Erreger mittels PCR

Bei einer Polymerasekettenreaktion (PCR) handelt es sich um eine molekularbiologische Methode bei der, mithilfe von Oligonukleotiden, Puffern und Enzymen DNA-Abschnitte gezielt amplifiziert werden können. Die Primer (Oligonukleotide) werden hierbei so gewählt, dass sie die zu amplifizierende Sequenz flankieren. Primer dienen als Bindungsstellen für die Polymerase, die einen komplementären Strang zur gewünschten Sequenz synthetisiert. Durch zyklische Temperaturänderungen kommt es pro Runde zu einer Verdoppelung der Nukleinsäuresequenz, was in eine exponentielle Amplifikation am Ende der PCR resultiert (Abbildung 11). Somit können aus wenigen Kopien vom Ausgangsmaterial bis zu  $10^{12}$  Kopien entstehen und damit anschließende Analysen oder Experimente erst ermöglichen. Für die Amplifikation und Detektion von RNA wird diese zuvor durch eine Reverse Transkription in DNA umgeschrieben. Im Anschluss erfolgt die eigentliche Amplifikation per PCR.



**Abbildung 11: Schematischer Ablauf einer Polymerasekettenreaktion**

Blau: DNA-Strang; Orange: Primer; Grün: Synthetisierter DNA-Strang; P: Polymerase

Bei Screening PCRs handelt es sich um PCRs, die verwendet werden, um die Anwesenheit bestimmter Sequenzen in einer Probe zu prüfen. Dies wird genutzt um qualitative Aussagen zu treffen und zum Beispiel zu überprüfen, ob ein Organismus von einem Pathogen befallen ist. Je nach dem Design der Primer kann dabei sowohl eine ganze Virusfamilie als auch nur eine spezifische Unterklasse detektiert werden. Für das Design für Primerbindungsstellen eignen sich hochkonservierte Regionen am besten, da in diesen weniger Mutationen zu finden sind. Mutationen können für eine schlechtere Primerbindung sorgen und damit die Detektion erschweren.

Im Folgenden werden die einzelnen, in dieser Studie genutzten, Screening PCRs im Detail vorgestellt.

### 2.4.3.1. Picobirnavirus

Für die Suche nach Picobirnaviren wurden die Extrakte aus den Kotpellets von *Proechimys semispinosus* untersucht. Dafür wurde eine Picobirnavirus spezifische Screening PCR designt, um die RdRp Region zu detektieren. Dabei wurde ein 446 bp langes Fragment auf Segment 2 amplifiziert. Dafür wurde das SuperScript III OneStep RT-PCR kit (Life Technologies) in einem 12,5 µL Reaktionsansatz verwendet und die Reaktion entsprechend Tabelle 1 und Tabelle 2 angesetzt.

**Tabelle 1: Reaktionsansatz und Temperaturprotokoll der Picobirnavirus Screening PCR**

Reagenz	Menge pro Ansatz [µL]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	1,55	50 °C	20 min	x 1
2x Reaction Mix	6,25	94 °C	3 min	x 1
MgSO <sub>4</sub> (50 mM)	0,2			
BSA (1 mg/mL)	0,5	94 °C	15 s	x 10
PsemRdRpPBirna-F_M13 (10 µM)	0,5	60-50 °C *	30 s	
PsemRdRpPBirna-RnestM13 (10 µM)	0,5	72 °C	30 s	
One-Step SuperScript® III/Platinum® Taq	0,5	94 °C	15 s	x 40
		50 °C	30 s	
		72 °C	30 s	
Template RNA	2,5	72 °C	2 min	x 1

\*Touchdown: -1°C pro Zyklus

**Tabelle 2: Primerliste für die Picobirnavirus Screening PCR**

Name	Sequenz
PsemRdRpPBirna-F_M13	5'- GTAAAACGACGGCCAGTGAATCAGCGGGTGGTTTG GATGT -3'
PsemRdRpPBirna-RnestM13	5'- CACACAGGAAACAGCTATGACGTCATCGCCAAGAC ACTGTGA -3'

#### 2.4.3.2. Picornavirus

Für das Design von spezifischen Primern benötigt man Informationen über die Sequenz des zu untersuchenden Virusgenoms. Die nicht zielgerichtete Hochdurchsatzsequenzierung (Kapitel 2.4.10) ermöglicht die Generierung viraler Genominformation und Entdeckung neuer und neuartiger viraler Erreger. Durch einen solchen NGS-Lauf, der *P. sem* Kot-Extrakte untersuchte, wurden mehrere neue Picornaviren entdeckt. Anhand der erhaltenen Informationen aus dem NGS Lauf konnte ein vollständiges *P.sem* Picornavirusgenom sequenziert werden. Auf Basis dieses 6660 bp langem Genoms wurden eine Screening PCR entwickelt. Um die Detektionseffizienz zu erhöhen wurde hierbei ein zweistufiger Hemi-nested PCR Ansatz gewählt. Bei einer (Hemi-)nested PCR wird ein Anteil des PCR-Produktes aus der ersten Amplifikation eingesetzt und dient als Vorlage für die zweite PCR. In dieser wird durch ein zweites Primerpaar, das an Sequenzbereiche innerhalb dieser Vorlage bindet (nested Primer), ein kürzeres DNA-Fragment amplifiziert. Bei einer Hemi-nested PCR wird dabei nur einer der beiden Primer ausgetauscht. Die Initial-Screening PCR (Tabelle 3) amplifiziert hierbei ein 550 bp langes Stück der 3D Region des Virusgenoms in der ersten PCR Runde bzw. ein 380 bp langes Stück in der zweiten PCR Runde. Für die erste Runde der PCR wurde das SuperScript III OneStep RT-PCR Kit (Life Technologies) in einem 12,5 µL Reaktionsansatz verwendet. Um die Funktionalität der neuen Initial-Screening-Primer zu überprüfen, wurden diese in einer PCR mit Proben getestet bei denen bekannt war, dass sie die neuartigen Picornaviren enthalten.

In der ersten Runde wurden dabei die Primer EnteroSPP\_Psem-F und EnteroSPP\_Psem-R verwendet, während EnteroSPP\_Psem-Fnest und EnteroSPP\_Psem-R in der zweiten Runde genutzt wurden (Primersequenzen siehe Tabelle 5).

**Tabelle 3: Reaktionsansatz und Temperaturprotokoll der Picornavirus Initial-Screening PCR****1. Runde der hemi-nested PCR**

Reagenz	Menge pro Ansatz [ $\mu\text{L}$ ]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	1,55	50 °C	20 min	x 1
2x Reaction Mix	6,25	94 °C	3 min	x 1
MgSO <sub>4</sub> (50 mM)	0,2			
BSA (1 mg/mL)	0,5	94 °C	15 s	x 10
EnteroSPP_Psem-F (10 $\mu\text{M}$ )	0,5	65-55 °C *	20 s	
EnteroSPP_Psem-R (10 $\mu\text{M}$ )	0,5	72 °C	30 s	
OneStep SuperScript® III/Platinum® Taq	0,5	94 °C	15 s	x 35
		56 °C	20 s	
		72 °C	30 s	
Template RNA	2,5	72 °C	2 min	x 1

\*Touchdown: -1°C pro Zyklus

**2. Runde der hemi-nested PCR**

Reagenz	Menge pro Ansatz [ $\mu\text{L}$ ]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	17,8	94 °C	2 min	x 1
10x Platinum Taq Puffer	2,5	94 °C	15 s	x 45
dNTP Mix (10 mM)	0,5	56 °C	20 s	
MgCl <sub>2</sub> (50 mM)	1	72 °C	30 s	
EnteroSPP_Psem-Fnest (10 $\mu\text{M}$ )	1	72 °C	2 min	x 1
EnteroSPP_Psem-R (10 $\mu\text{M}$ )	1			
Platinum® Taq	0,2			
PCR Produkt der 1. Runde	1			

Die PCR Produkte wurden mittels Agarosegelelektrophorese visualisiert und durch Sanger-Sequenzierung (SEQLAB Sequence Laboratories, Göttingen, Germany) sequenziert. Im Anschluss daran, wurden sämtliche Kot-Extrakt Pools (n=26) auf die neuen Viren getestet. Sechs Pools wurden positiv getestet und sequenziert. Aus den Ergebnissen der Sequenzierung konnten insgesamt 3 verschiedene Stämme identifiziert werden, für die neue, spezifische Hemi-nested Screening PCRs designt wurden (insgesamt 8 Primer). Die neue Hemi-nested Screening PCR (Tabelle 4 und Abbildung 12) zielt auf die gleiche 3D-Region ab wie die erste Screening PCR und amplifiziert ein 284 bzw. 189 bp langes Stück in der jeweils ersten und zweiten Runde.

Die Primer der ersten Runde EntSPP-F1, HunSPP-F1, EntSPP-R1, Ent2SPP-R1 und HunSPP-R1 wurden in equimolarer Konzentration als Forward-Primer Mix bzw. Reverse-Primer Mix in der PCR eingesetzt.

Für die zweiten Runde wurden die Primer EntSPP-Fnest1, Ent2SPP-Fnest1, HunSPP-Fnest1, EntSPP-R1, Ent2SPP-R1 und HunSPP-R1 in equimolarer Konzentration als Fnested-Primer Mix und Reverse-Primer Mix verwendet.

**Tabelle 4: Reaktionsansatz und Temperaturprotokoll der neuen Picornavirus Screening PCR**

**1. Runde der Hemi-nested PCR**

Reagenz	Menge pro Ansatz [ $\mu\text{L}$ ]	Temperatur	Dauer	Zyklusanzahl
2x Reaction Mix	6,25	50 °C	20 min	x 1
MgSO <sub>4</sub> (5 mM)	1,15	94 °C	3 min	x 1
BSA (1 mg/mL)	0,5			
Picorna Forward-Primer Mix (10 $\mu\text{M}$ )	0,8	94 °C	15 s	x 10
Picorna Reverse-Primer Mix (10 $\mu\text{M}$ )	0,8	60-50 °C *	20 s	
		72 °C	30 s	
One-Step SuperScript® III/Platinum® Taq	0,5	94 °C	15 s	x 35
		50 °C	20 s	
		72 °C	30 s	
Template RNA	2,5	72 °C	2 min	x 1

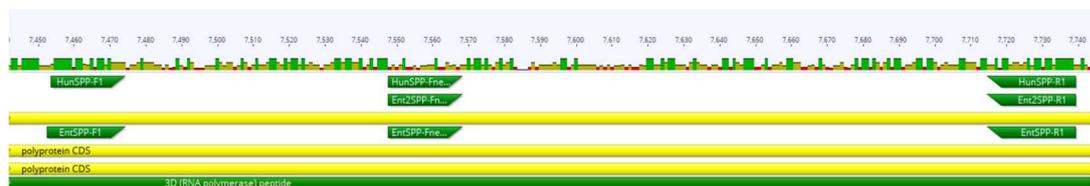
\*Touchdown: -1°C pro Zyklus

**2. Runde der Hemi-nested PCR**

Reagenz	Menge pro Ansatz [ $\mu\text{L}$ ]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	16,6	94 °C	2 min	x 1
10x Platinum Taq Puffer	2,5	94 °C	15 s	x 45
dNTP Mix (10 mM)	0,5	56 °C	20 s	
MgCl <sub>2</sub> (50 mM)	1	72 °C	30 s	
Picorna Fnested-Primer Mix (10 $\mu\text{M}$ )	1,6	72 °C	2 min	x 1
Picorna Reverse-Primer Mix (10 $\mu\text{M}$ )	1,6			
Platinum® Taq	0,2			
PCR Produkt der 1. Runde	1			

**Tabelle 5: Verwendete Oligonukleotide für die Picornavirus Screening PCR**

Name	Sequenz
EntSPP-F1	5'- GCAATCCTGATGTTGATTGGAC -3'
HunSPP-F1	5'- CAATCCTGATATGGACTGGCT -3'
EntSPP-R1	5'- GATGTTGTTTATTATAGTGTGAGA -3'
Ent2SPP-R1	5'- GATGTTATTTATGATGGTATTCAAG -3'
HunSPP-R1	5'- AATGTTGTTTCATTATGGTGTGAAG -3'
EntSPP-Fnest1	5'- GGACCGGTTATAATGGATGCT -3'
Ent2SPP-Fnest1	5'- GGACCAGTCATAATGGATGCT -3'
HunSPP-Fnest1	5'- ACTTCTTGGAGCTTCTTCTGC -3'

**Abbildung 12: Schematische Darstellung der Lage der Picornavirus Screening PCR Primer**

Gezeigt ist die schematische Lage der Primer für die Picornavirus Screening PCR. Alle 8 Primer liegen dabei auf der 3D Region

#### 2.4.3.3. Hepacivirus

Entnommene Blutproben wurden nach der Extraktion gepoolt und mit einer nested RT-PCR gescreent (s. [88, 99]). Dieser spezifische Assay, für die Detektion von Hepaciviren, wurde auf einem Alignment aus generierte Sequenzdaten aus *P.sem* zusammen mit verwandten Sequenzen aus GenBank (National Center for Biotechnology Information, USA) designt. Dabei wurde die konservierte Helikase Region, in Hepaciviren auch NS3 genannt, als Primerbindungsstelle ausgewählt (Tabelle 7). Da es sich bei Hepaciviren um RNA-Viren handelt, wurde das SuperScript III OneStep RT-PCR Kit (Thermo Fisher Scientific) für die PCR in einem 12,5 µL Reaktionsansatz (Tabelle 6) verwendet.

**Tabelle 6: Reaktionsansatz und Temperaturprotokoll der Hepacivirus Screening PCR**

Reagenz	Menge pro Ansatz [ $\mu\text{L}$ ]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	1,05	55 °C	20 min	x 1
2x Reaction Mix	6,25	94 °C	3 min	x 1
MgSO <sub>4</sub> (50 mM)	0,5			
BSA (1 mg/mL)	0,2	94 °C	15 s	x 10
HepaciPsem3276-F (10 $\mu\text{M}$ )	0,75	60-50 °C *	20 s	
HepaciPsem4263-R (10 $\mu\text{M}$ )	0,75	72 °C	1 min	
One-Step SuperScript® III/Platinum® Taq	0,5	94 °C	15 s	x 35
		50 °C	20 s	
		72 °C	1 min	
Template RNA	2,5	72 °C	2 min	x 1

\*Touchdown: -1°C pro Zyklus

**Tabelle 7: Verwendete Oligonukleotide für die Hepacivirus Screening PCR**

Name	Sequenz
HepaciPsem3276-F	5'- AGCCGCTGCTGATGAACAAGG -3'
HepaciPsem4263-R	5'- CRTTTGGRATGGTKGAGGCATC -3'

#### 2.4.3.4. Deltavirus

Die Hochdurchsatzsequenzierung von Blutproben identifizierte virale Nukleinsäure des Hepatitis Delta Virus. Für die PCR-Analyse wurden hier ebenfalls spezifische Primer etabliert (Tabelle 9), ähnlich wie in Kapitel 2.4.3.2 bereits beschrieben..

Alle RNA Extrakte wurden mit einer Echtzeit quantitativen PCR mit reverser Transkription (qRT-PCR; Abbildung 13) untersucht, unter Verwendung der spezifischen Primer für *P. sem* Deltaviren. Zielregion der Amplifikation war hier ein 94 bp großes Fragment des S-HDAg.

Dafür wurde erneut das SuperScript III OneStep RT-PCR Kit mit Platinum Taq DNA Polymerase Kit (Thermo Fisher Scientific) in einem 12,5  $\mu\text{L}$  Ansatz verwendet (Tabelle 8).

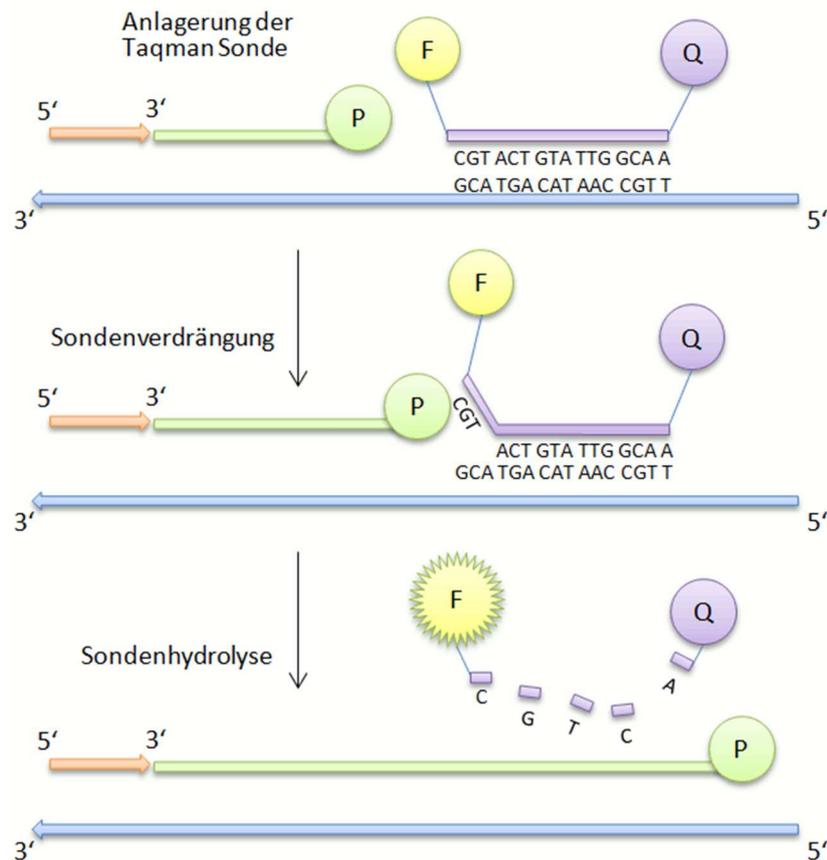
Tabelle 8: Reaktionsansatz und Temperaturprotokoll der Deltavirus Screening PCR

Reagenz	Menge pro Ansatz [ $\mu$ L]	Temperatur	Dauer	Zyklenanzahl
H <sub>2</sub> O	1,05	50 °C	20 min	x 1
2x Reaction Mix	6,25	94 °C	3 min	x 1
MgSO <sub>4</sub> (50 mM)	0,2			
BSA (1 mg/mL)	0,5	94 °C	15 s	x 10
rtHDVPsemAG-F1 (10 $\mu$ M)	0,5	60-50 °C *, **	30 s	
rtHDVPsemAG-R1 (10 $\mu$ M)	0,5	72 °C	30 s	
OneStep SuperScript® III/Platinum® Taq	0,5	94 °C	15 s	x 40
		50 °C **	30 s	
		72 °C	30 s	
TaqMan Sonde (5 $\mu$ M)	0,5	72 °C	2 min	x 1
Template RNA	2,5			

\*Touchdown: -1°C pro Zyklus

\*\* Fluoreszenzmessung

Das Echtzeit-RT-PCR-Protokoll wurde wie in Tabelle 8 beschrieben auf einem Roche LightCycler 480 durchgeführt.



**Abbildung 13: Schematische Darstellung der Detektion mittels Taq-Man Sonde**  
 Blau: DNA Strang; Orange: Primer; Grün: Amplikon; P: Polymerase; F: Fluorophor;  
 Q: Quencher

**Tabelle 9: Sequenzen und Verwendung der Oligonukleotide für Hepatitis Deltavirus PCR**  
 Ein Sternchen kennzeichnet identische Primer, die in mehreren Assays verwendet wurden.

Name	Sequenz (5'-3')	Primer Orientierung	Verwendeter Assay
rtHDVPsemAG-F1	AGG AAA GGG AGG ACC ATC GC	forward	Screening qRT-PCR Zielseq. HDag
rtHDVPsemAG-R1	GCC TCT TCC TCC TCG CTC A	reverse	Screening qRT-PCR Zielseq.HDag
rtHDVPsemAG-P (Sonde)	FAM-AGA AGC AGC-ZEN-TGG AGG AGC AAG GA-IBFQ	forward	Screening qRT-PCR Zielseq.HDag
Tbp_f	AAY CTT GGT TGT AAA CTT GAC CT	forward	TBP qRT-PCR

Tbp_r	GGG CTC CCT TAT TCT CAT GA	reverse	TBP qRT-PCR
Tbp_p (Sonde)	FAM-CCG AAA TGC-ZEN-TGA ATA TAA TCC CAA GC-IBFQ	forward	TBP qRT-PCR
HDV_Psem7-Fwd	CGA CGG CTC GCC GAG GA	forward	Zirkularisierungsassay 1 Erste Runde
HDV_Psem7-Rev	CTT CTT TCC TTG CTC CTC CAG C	reverse	Zirkularisierungsassay 1 Erste Runde
HDV_Psem7-Fnest	TCG CCG AGG AGG ACG AAC GTC	forward	Zirkularisierungsassay 1 Zweite Runde
HDV_Psem7-Rnest	TTT GTT CTC GAG GGC ACA CCT TCG	reverse	Zirkularisierungsassay 1 Zweite Runde
HDV_PsemCloneF1*	GAG GAA GAA GAA GAA GCT TGA GG	forward	Zirkularisierungsassay 2 Erste Runde
rtHDVPsem3-R1	TCC TGT CTG GGC TTG GGA GT	reverse	Zirkularisierungsassay 2 Erste Runde
HDV_Psem5-F	CAT GGC TGG GTA ACG TTC TTG GAA	forward	Zirkularisierungsassay 2 Zweite Runde
HDV_Psem8-Rev	ACC GGG AGT CTC CAT CCT GGA GT	reverse	Zirkularisierungsassay 2 Zweite Runde
rtHDVPsemRI-F2	CTT CGG GAC GAC AAG GAA ATC C	forward	Zirkularisierungsassay 3 Erste Runde
HDV_PsemCloneR1**	CTA GGG ATA TTC CCC TTC TCC TCT	reverse	Zirkularisierungsassay 3 Erste Runde
rtHDVPsemRI-F1	TCC TCT TCG GGA CGA CAA GG	forward	Zirkularisierungsassay 3 Zweite Runde
HDV_Psem5-R	CGT CTC CTC TTC TGG AGA ATG G	reverse	Zirkularisierungsassay 3 Zweite Runde
hHDV1_2-Fwd	TCC TCT TCG GGT CGG CAT G	forward	Humanes HDV Zirkularisierungsassay Erste Runde

## 2. Material und Methoden

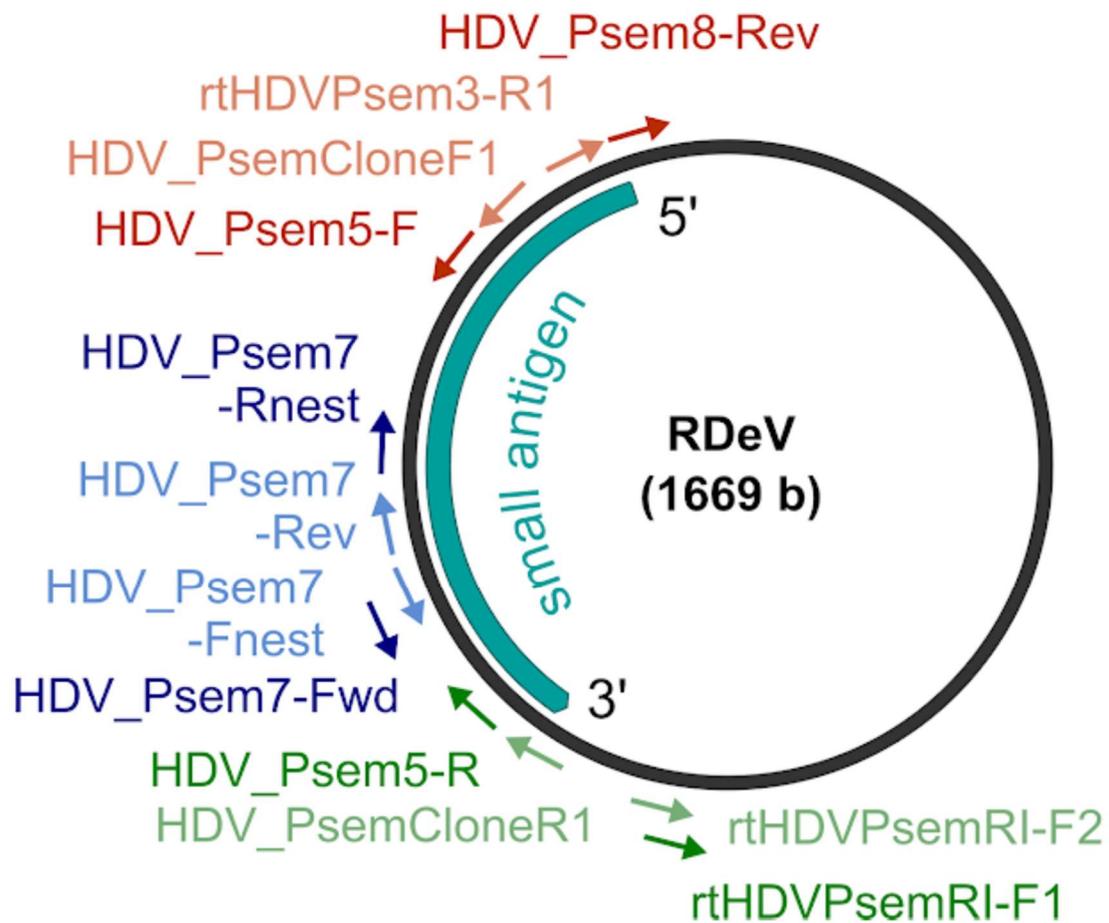
hHDV1_2-Rev	ATC GGA TGG AAA GAG TAT ATC C	reverse	Humanes HDV Zirkularisierungsassay Erste Runde
hHDV1_2-Fnest	GGT CGG CAT GGC ATC TCC	forward	Humanes HDV Zirkularisierungsassay Zweite Runde
hHDV1_2-Rnest	ACT CCG GAA CTC CTT GCA T	reverse	Humanes HDV Zirkularisierungsassay Zweite Runde
HDV_PsemCloneF1*	GAG GAA GAA GAA GAG GCT TGA GG	forward	Generierung eines Klonfragments
HDV_PsemCloneR1**	CTA GGG ATA TTC CCC TTC TCC TCT	reverse	Generierung eines Klonfragments
T7-fwd	TAA TAC GAC TCA CTA TAG GG	forward	Sequenzierung des Klonfragments
BGH-rev	TAG AAG GCA CAG TCG AGG	reverse	Sequenzierung des Klonfragments
Seq-Dimer-F	TCG CTC TCC GGA TGG	forward	Sequenzierung des Klonfragments
pcDNA3.1-Hind-rev	CAA GCT TAA GTT TAA ACG CTA GC	reverse	PsemD1 Amplifikation
rodHDV-for	ATG GAA ACA CCT CCT GGA GAA G	forward	PsemD1 Amplifikation
rodHDV-for-P	P-ATG GAA ACA CCT CCT GGA GAA G	forward	cDNA Genome Amplifikation
rodHDV-rev-P	P-CCG GAG AGC GAG GAC CG	reverse	cDNA Genom Amplifikation
pSEM-117F	GCT TGG GAG TTT TCT TCT TAC C	forward	Kolonie PCR, Sequenzierung des Klonfragments
pSEM-1540R	CTG CGT TTC CAG CAG CTA G	reverse	Kolonie PCR, Sequenzierung des Klonfragments

#### 2.4.4. Deltavirus Zirkularisierungsassay

Um zu bestätigen, dass das Genom des Nagetier-Deltavirus zirkulär ist, wurden alle Deltavirus-positiven Blut-RNA-Extrakte einer cDNA-Synthese unter Verwendung der SuperScript III Reverse Transcriptase (Thermo Fisher Scientific) mit spezifischen Forward-Primern unterzogen (s. Tabelle 9). Für die cDNA-Synthese wurden zwei Ansätze hergestellt (Tabelle 10).

cDNA Synthese-Mix 1		cDNA Synthese-Mix 2	
Reagenz	Menge pro Ansatz [ $\mu$ L]	Reagenz	Menge pro Ansatz [ $\mu$ L]
RNase-freies H <sub>2</sub> O	8	RNase-freies H <sub>2</sub> O	0,5
5x Erststrangpuffer	2	5x Erststrangpuffer	2
dNTP-Mix (je 10 mM)	1	Dithiothreitol (100 mM)	1
Forward-Primer (10 $\mu$ M)	1	BSA (1 mg/mL)	1
Template RNA	2,5	SuperScript III Reverse Transkriptase	1

Mix 1 wurde 5 min bei 65 °C inkubiert und 1 min auf Eis gelegt. Parallel zum Kühlschritt wurde Mix 2 auf 55 °C erhitzt. Nach dem Abkühlen auf Eis wurden 5,5  $\mu$ L von Mix 2 in Mix 1 gegeben, für 60 min auf 55 °C und anschließend für 15 min auf 70 °C erhitzt. Zum Abschluss wurde 1  $\mu$ L RNase H zugegeben und für 20 min bei 37 °C inkubiert. Diese cDNA wurde als Vorlage in drei verschiedenen Nested-PCRs (Tabelle 11) mit einem Reaktionsvolumen von 25  $\mu$ L bzw. 50  $\mu$ L für die erste und zweite Runde verwendet (Lage der Primer auf dem Genom kann Abbildung 14 entnommen werden).



**Abbildung 14: Vereinfachte Darstellung des Rodent-Deltavirus Genoms mit Zirkularisierungsassays**

Eingezeichnet sind drei verschiedene PCR Assays mit Primerbindungsstellen die zur Überprüfung der Zirkularität genutzt wurden. Primersequenzen können der Tabelle 7 entnommen werden. Abbildung wurde modifiziert aus Paraskevopoulou und Pirzer *et.al* [100]

**Tabelle 11: Reaktionsansatz und Temperaturprotokoll des Hepatitis Deltavirus Zirkularisierungsassays****1. Runde der Nested-PCR (25 µL)**

Reagenz	Menge pro Ansatz [µL]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	16,6	94 °C	3 min	x 1
10x Platinum Taq Puffer	2,5	94 °C 58 °C * 72 °C	15 s	x 45
MgCl <sub>2</sub> (50 mM)	1		20 s	
dNTP Mix (10 mM)	0,5		90 s	
Forward-Primer (10 µM)	1	72 °C	3 min	x 1
Reverse-Primer (10 µM)	1			
Platinum® Taq	0,2			
cDNA	2,5			

**2. Runde der Nested PCR (50 µL)**

Reagenz	Menge pro Ansatz [µL]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	33,2	94 °C	3 min	x 1
10x Platinum Taq Puffer	5	94 °C 58 °C 72 °C	15 s	x 45
MgCl <sub>2</sub> (50 mM)	1		20 s	
dNTP Mix (10 mM)	2		90 s	
Forward-Primer (10 µM)	3,2	72 °C	3 min	x 1
Reverse-Primer (10 µM)	3,2			
Platinum® Taq	0,4			
PCR Produkt der 1. Runde	2			

Die Amplikons wurden mittels Sanger-Sequenzierung sequenziert.

**2.4.5. RT-PCR Amplikon-Quantifizierung**

Zur Quantifizierung, der Konzentration an Deltaviren, wurde in allen quantitativen Reverse-Transkription PCR (qRT-PCR) Läufen ein seriell verdünntes und photometrisch quantifiziertes In-vitro-Transkript (IVT) verwendet. Dafür wurde ein Teil der HDAG-Kodierungsregion amplifiziert (Primer pSEM-117F und pSEM-1540R, Tabelle 9). Für die PCR wurde das SuperScript III OneStep RT-PCR System mit Platinum Taq DNA

Polymerase (Thermo Fisher Scientific) in einem Reaktionsvolumen von 25  $\mu\text{L}$  verwendet (Tabelle 12).

**Tabelle 12: Reaktionsansatz und Temperaturprotokoll der Deltavirus Klonierungsfragment PCR**

Reagenz	Menge pro Ansatz [ $\mu\text{L}$ ]	Temperatur	Dauer	Zyklusanzahl
H <sub>2</sub> O	2,1	55 °C	20 min	x 1
2x Reaction Mix	12,5	94 °C	3 min	x 1
MgSO <sub>4</sub> (50 mM)	0,4	94 °C	15 s	x 45
BSA (1 mg/mL)	1	58 °C	20 s	
pSEM-117F (10 $\mu\text{M}$ )	1,5	72 °C	40 s	
pSEM-1540R (10 $\mu\text{M}$ )	1,5	72 °C	2 min	x 1
OneStep SuperScript® III/ Platinum® Taq	1			
Template RNA	5			

Die Klonierung erfolgte gemäß den Herstellerangaben mit Hilfe des „TOPO TA cloning“-Systems mit OneShot TOP10 chemisch kompetenten *Escherichia coli* (*E. coli*) Zellen. Die Integration der DNA-Fragmente in den Vektor erfolgte über einen 3'-Oligo-A Überhang am PCR-Produkt. Dieser A-Überhang wurde durch eine PCR-Reaktion mit der Platinum Taq DNA-Polymerase an das gewünschte DNA-Fragment angehängen. Die Ligation der DNA in den Vektor erfolgte über eine TA-Überhangbindung mit Hilfe der an den linearisierten Vektor gekoppelten Topoisomerase.

Der Ligationsmix enthielt 0,5  $\mu\text{L}$  Salzlösung aus dem Klonierungskit, 2  $\mu\text{L}$  PCR-Produkt und 0,5  $\mu\text{L}$  pCR4 TOPO Vektor. Diese Mischung wurde 30 min bei Raumtemperatur und 2 min auf Eis inkubiert. 2  $\mu\text{L}$  der Ligationsmischung wurden in die kompetenten Zellen gegeben, 25 min auf Eis inkubiert und danach 30 s lang bei 42°C hitzegeschockt. Anschließend wurde die Mischung 1 min lang auf Eis inkubiert. 250  $\mu\text{L}$  SOC-Medium wurden zu der Transformationsmischung gegeben, 1 Stunde lang bei 37°C geschüttelt, auf LB-Platten mit Kanamycin ausplattiert und bei 37°C inkubiert. Die Klone wurden einer PCR und einer Sanger-Sequenzierung unterzogen. Die Klone mit der korrekten Orientierung wurden über Nacht in 2 mL SOC-Medium inkubiert. Die Plasmidkolonien wurden aus der Übernachtskultur mit dem QIAprep Spin Miniprep Kit nach dem Standardprotokoll extrahiert. In RNase-freiem Wasser wurde eine Verdünnungsreihe

hergestellt. Die Verdünnungsstufen  $10^{-6}$  bis  $10^{-10}$  wurden in einer PCR unter Verwendung der M13-Primer getestet (Tabelle 13).

**Tabelle 13: Reaktionsansatz und Temperaturprotokoll der Deltavirus Kolonie PCR**

Reagenz	Menge pro Ansatz [ $\mu$ L]	Temperatur	Dauer	Zyklenanzahl
H <sub>2</sub> O	38,6	94 °C	3 min	x 1
10x Platinum Taq Puffer	5	94 °C	15 s	x 45
dNTP Mix (10 mM)	1	58 °C	15 s	
MgCl <sub>2</sub> (50 mM)	2	72 °C	30 s	
M13-Forward (10 $\mu$ M)	1	72 °C	1 min	x 1
M13-Reverse (10 $\mu$ M)	1			
Platinum® Taq	0,4			
Plasmidverdünnung	1			

Um den Plasmidhintergrund zu reduzieren, wurde die Verdünnung mit der letzten deutlich sichtbaren Bande auf einem 2%igen Agarosegel verwendet (Verdünnungsfaktor  $10^{-6}$ ). Das PCR-Produkt wurde mit dem QiaQuick PCR Purification Kit (Qiagen) gemäß den Anweisungen des Herstellers aufgereinigt und anschließend mit dem MEGAscript T7 Kit (Thermo Fisher Scientific) transkribiert und erneut mit dem RNeasy Mini Kit (Qiagen) gereinigt, jeweils gemäß den Herstelleranweisungen. Das IVT wurde in RNase-freiem Wasser mit 10  $\mu$ g/mL carrier-RNA (Qiagen) verdünnt.

#### 2.4.5.1. Herstellung einer in-vitro Transkript Standardreihe für qPCR

Bei einer quantitativen PCR (qPCR) werden Reportersonden eingesetzt, die unter bestimmten Umständen Licht einer bestimmten Wellenlänge emittieren können. Die verwendete Taqman Sonde ist eine Hydrolysesonde aus einem einzelsträngigen Oligonukleotid. An den jeweiligen Enden sind sowohl ein Quencher als auch ein Fluorophor gekoppelt. Bei dem Quencher handelt es sich um ein Molekül, das die Fluoreszenz des Fluorophors unterdrückt, solange sich beide in räumlicher Nähe befinden. Sobald die Taqman Sonde an das Amplikon bindet, kann die Polymerase, durch ihre Exonukleaseaktivität das Fluorophor und den Quencher voneinander trennen. Da nun das Fluorophor vom Quencher räumlich getrennt ist, kann es, nach Anregung durch eine bestimmte Wellenlänge, Fluoreszenz emittieren. Da sich dieser Vorgang in

jedem Zyklus wiederholt und sich die Menge an freiem Fluorophor akkumuliert, ist die Stärke des nach der Anregung gemessenen Fluoreszenzsignals proportional zur Menge an amplifizierter DNA.

Um aus diesem Fluoreszenzsignal nun eine Genomkopienzahl zu bestimmen, muss parallel eine Probe oder Standardreihe mit bekannter Konzentration gemessen werden. Hierzu wird die Zielregion der PCR durch eine Klonierung mit anschließender Transformation amplifiziert. Nach der Messung am Nanodrop kann die Anzahl an Genomkopien durch folgende Formel bestimmt werden:

$$(1) \left( \frac{\beta}{\text{Fragmentlänge} * M} \right) * 6 * 10^{17} = \text{Kopienzahl}/\mu\text{L}$$

Wobei hierfür gilt:

$\beta$ : Massenkonzentration in g/L des IVTs (gemessen am Nanodrop)

$M$ : Durchschnitt der molaren Masse aller vier Nukleotide (340g/mol)

*Fragmentlänge*: Die Länge des IVTs in Basenpaaren [bp]

$6 * 10^{17}$ : Gekürzte Avogadrokonstante, um den Faktor von L auf  $\mu\text{L}$  auszugleichen und von der Einheit „Mol“ auf die Teilchenanzahl zu kommen.

#### 2.4.6. Aufreinigung von DNA aus PCR-Produkten

Um PCR-Produkte für weitere Analysen, z.B. die Sanger-Sequenzierung verwenden zu können, mussten diese zunächst aufgereinigt werden. Dazu wurde das PCR-Produkt über eine Affinitätssäule (QIAquick® Spin Kit) gemäß den Herstellerangaben aufgereinigt und mit destilliertem Wasser eluiert.

Alternativ dazu wurden PCR-Produkte auch per enzymatischem Verdau mit dem „ExoSAP-it for PCR Product Clean-Up“ Kit wie vom Hersteller beschrieben behandelt.

#### 2.4.7. Aufreinigung der DNA aus Agarosegelen

Bei der Sequenzierung eines PCR-Produktes können unspezifische Nebenbanden die Sequenzierung beeinträchtigen. Um dies zu verhindern kann man die Produkte aus der PCR auf einem Agarosegel elektrophoretisch auftrennen und die gewünschte Bande aus dem Gel schneiden. Im Anschluss muss jedoch die DNA aus dem Agarosegelstück

extrahiert werden. Hierfür wurde entweder das MinElute® Spin Kit oder alternativ das QIAex II DNA-Extraktionskit gemäß den Herstellerangaben verwendet.

#### 2.4.8. Konzentrationsmessung von Nukleinsäuren

Die Konzentration der Nukleinsäure (DNA sowie RNA) wurde mittels photometrischer Absorption bei einer Wellenlänge von 260 nm ( $OD_{260}$ ) am Nanodrop bestimmt. Dabei ist der Nukleinsäuregehalt der Probe direkt proportional zu der Menge an absorbierter Strahlung. Der Quotient aus verschiedenen Absorptionen gibt hierbei über die Reinheit der Probe Auskunft. Dieser Quotient liegt bei Verunreinigungen mit Proteinen bei  $OD_{260} / OD_{280} < 1,8$  und bei reiner DNA bei einem Wert von  $\geq 1,8$ . Kontaminationen mit organischen oder chaotropen Verbindungen (z.B. durch Phenol-Chloroform-Extraktion) haben ein Quotienten  $OD_{260} / OD_{230}$  von  $< 1,8$ .

#### 2.4.9. Sanger-Sequenzierung

Standardmäßig wurde die DNA aus PCR-Reaktionen mit der Kettenabbruchmethode nach Sanger sequenziert. Hierbei werden bei der Synthese des DNA Stranges zusätzlich fluoreszierende Didesoxyribonukleosid eingesetzt, die, durch die fehlende Hydroxy-Gruppe beim Einbau zu einem Kettenabbruch der Synthese führen. Dieser Vorgang findet willkürlich statt, wodurch eine große Bandbreite an Fragmenten unterschiedlicher Länge entstehen. Da somit jeweils die letzte Position des Stranges fluoreszenzmarkiert ist, kann durch eine Größenauftrennung per Chromatographie jede einzelne Position bestimmt werden. Die einzelnen Sequenzierungsreaktionen wurden kommerziell von der Firma Seqlab (Göttingen) durchgeführt.

#### 2.4.10. Next Generation Sequencing (NGS)

Für die Erregeridentifikation und zur Generierung von Genomsequenzen wurden Kot- und Blutproben mittels Hochdurchsatzsequenzierung untersucht. Dafür musste zunächst eine Sequenzierbibliothek hergestellt werden. Die RNA-Konzentration der zu untersuchenden Proben wurde mit dem Qubit RNA HS Assay-Kit (Thermo Fisher Scientific) in Kombination mit dem Qubit 2.0 Fluorometer-Instrument quantifiziert. Da alle Probenkonzentrationen zu niedrig waren, wurden 5  $\mu$ L jeder extrahierten RNA für die

Vorbereitung der Sequenzierbibliothek, unter Verwendung des KAPA RNA Hyper Prep Kits (Roche Molecular Diagnostics) mit halbem Reaktionsvolumen verwendet. Änderungen gegenüber dem Standardverfahren bestanden darin, dass die Fragmentierung bei 85 °C für 5 min durchgeführt und die Menge der Adapter in der Adapterligation auf 0,75 µM verringert wurde. Zur Aufreinigung der Bibliothek wurde ein 0,8-faches Bead-Clean-up (55 µL Bibliothek und 44 µL Beads) mit KAPA Pure Beads durchgeführt und mit EB-Puffer (Qiagen) eluiert. Die Amplifikation der Bibliothek begann mit einer Denaturierung bei 98 °C für 45 s, gefolgt von 12 Zyklen von 15 s bei 98 °C, 30 s bei 60 °C, 72 °C und einer abschließenden Elongation bei 72 °C für 1 Minute, gefolgt von einem 0,7-fachen Bead-Clean-up. Die indizierten DNA-Bibliotheken wurden dann mit RNase-freiem Wasser auf 1,5 nM normalisiert und equimolar gepoolt. Der Pool wurde mit dem Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) gemessen. Die gepoolten Bibliotheken wurden mit 0,2 M NaOH denaturiert und mit Tris-HCl 0,2 M/pH 7 neutralisiert. Die Pools wurden mit Hybridisierungspuffer (HT1, Illumina) auf eine Endkonzentration von 13 pM verdünnt. Die Sequenzierung wurde mit der 600 Zyklen umfassenden MiSeq-Reagenzienkartusche v3 (Illumina) gemäß den Angaben des Herstellers durchgeführt.

#### 2.4.11. Identifikation unbekannter Erreger mittels Pipeline

Um aus Rohdaten eines NGS Laufes ungezielt Pathogene zu bestimmen, wurden diese Reads durch eine Pipeline prozessiert. Dabei wurden mit dem Programm DIAMOND (Version 0.9.22.123 [101, 102]) Contigs gebildet. Diese Contigs wurden durch die RVDB-Protein-Datenbank (Version 13.0 [101]) gefiltert um Übereinstimmungen mit viralen Proteinen zu finden. Lagen ausreichend Reads vor, konnte durch In-Silco-Assembly ein Genom, oder größerer Genomabschnitt, erstellt werden. Mit Hilfe dieser Sequenzdaten wurden spezifische Screening PCRs designt (z.B Kapitel 1.5.2 Picornavirus Grundlagen). Falls keine ausreichende Anzahl an Virus Reads vorliegt, um eine Screening PCR zu erstellen, so ist die Pipeline Identifikation dennoch enorm hilfreich, da sie somit zur Probenbestimmung für neuartige Erreger benutzt werden kann. Auf diesen identifizierten Proben wurden dann, wie im Fall von Deltaviren (Kapitel 1.5.3 Deltavirus Grundlagen, weitere NGS Läufe angesetzt mit erhöhter Sequenziertiefe.

#### 2.4.12. Sequenzanalyse

Sequenzdaten wurden initial mit der Software Geneious (Version 9.1.7.[21]) analysiert. Hier erfolgte das Design von Primern, Annealing von Sequenzfragmenten, Alignment sowie phylogenetische Baumberechnungen. Außerdem wurden in Geneious sämtliche NGS Reads, durch iteratives Mapping an Referenzsequenzen, zu größeren Contigs assembliert. Die so entstandenen Konsensus-Sequenzen wurden mit weiteren Referenzsequenzen aligniert und für phylogenetische Analysen benutzt bzw. wurde in der Datenbank des „National Center for Biotechnology Information“ (NCBI) durch den BLAST (engl. „basic local alignment search tool“) Algorithmus nach verwandten Sequenzen gesucht.

#### 2.4.13. Phylogenetische Analyse

Für die Untersuchungen von Verwandtschaftsverhältnissen und Unterschieden in der Nukleotid- bzw. Aminosäuresequenz wurden funktioneller Genombereiche der identifizierten Viren mit repräsentativen Mitgliedern der gleichen Virusfamilie oder -ordnung über den MAFFT [22] Algorithmus aligniert. Die Erstellung der Stammbäume erfolgte, abhängig von der Fragestellung oder der Alignments, entweder mittels Neighbor-Joining oder Maximum-Likelihood (ML) Algorithmen in Geneious. Um die Qualität des Baumes an den Baumknoten zu berechnen, wurden 1000 Wiederholungen (bootstraps) durchgeführt.

##### 2.4.13.1. Bayesianische Analyse zur Evolutionsrate per BEAST

Bei BEAST (Bayesian Evolutionary Analysis Sampling Trees; [91]) handelt es sich um ein Software Paket, dass aus Nukleotid- oder Aminosäuresequenzalignments phylogenetische Rekonstruktionen erstellt. Dabei kann es sich, neben dem Stammbaum auch um eine Extrapolation eines Wurzel- oder Kladenalters handeln. BEAST liest hierbei eine XML-Datei ein, die man per Partnerprogramm BEAUTI (Bayesian Evolutionary Analysis Utility) aus einer Vorlage erstellen kann.

BEAST nutzt als mathematische Grundlage bayesianische Wahrscheinlichkeitstheorie sowie das Markovkette-Monte-Carlo Verfahren (MCMC). Die Markovkette nutzt eine Übergangswahrscheinlichkeitsmatrix um abzuschätzen in welchen zukünftigen Zustand ( $X_{t+1}$ ) ein aktueller Zustand ( $X_t$ ) übergeht. Dadurch kann durch eine Anreihung an Ereignissen ein beliebiger Zustand, in der Zukunft oder der Vergangenheit, abgeschätzt

werden. Ist dabei die Änderung zwischen den beiden Zuständen nur noch sehr gering, spricht man von einem stationären Zustand, bei dem das System die Konvergenz erreicht hat.

Die Wahrscheinlichkeitsebene wird aus mehreren Parametern, wie zum Beispiel der molekularen Clock-Rate oder die Populationsgröße, gebildet. Jeder dieser Parameter hat dabei zwei weitere Faktoren: Inklusion und Tuning. Der Inklusionsfaktor kann nur die Werte 0 und 1 annehmen und bestimmt damit, ob ein Parameter für das momentane Modell genutzt wird. Die Größe des Tuningfaktors bestimmt wie stark der Parameter Einfluss auf das Modell hat. In jedem Berechnungsschritt können dabei diese Faktoren geändert werden. Die Menge an Berechnungsschritten wird auch Kettenlänge der Markovkette bezeichnet. Handelt es sich bei den Parametern um nicht intrinsische Werte der Proben, also um Parameter die nicht bereits vor der Analyse bestimmbar sind, so müssen diese auch abgeschätzt werden. Durch diese steigende Menge an Faktoren und Parametern steigt die Anzahl an Möglichkeiten, die das System einnehmen kann (Freiheitsgrade).

Abhängig von der Komplexität der Alignments und der Fragestellung bzw. der Analyse muss die Kettenlänge deswegen erhöht werden, um sicher zu stellen, dass der Wahrscheinlichkeitsraum, den BEAST durchsucht, komplett abgedeckt und eine Konvergenz erreicht ist. Sollte man bereits konkretes Wissen über seine Daten haben, z.B. das Alter der Probe, so kann man diese Informationen als Prior in das Model hinzufügen. Durch die Einspeisung eines Priors werden dem System gewisse Grenzen gesetzt, was die Analyse beschleunigen kann. Dieser Prior wird im Verlauf der Analyse, durch die Verbindung mit den tatsächlichen Sequenzdaten, zum Posterior. Das Software-Paket von BEAST enthält eine hohe Anzahl an komplexen evolutionären Modellen zur Berechnung von Substitutionen, Insertions- und Deletionsverhalten, demografische Veränderungen, sowie Baummodellen und Clock-Raten.

#### 2.4.13.2. **Modelauswahl**

In einer bayesianischen Analyse werden die Übergänge von Zuständen in einen anderen Zustand durch Transitionmatrizen bestimmt. Aus diesem Grund, ist die Auswahl geeigneter Modelle, die die Transitionen errechnen, von größter Bedeutung. In der vorliegenden Arbeit wurde zur Bestimmung der Substitution-, Populations- und Clock-Modellen verschiedene Methoden genutzt. Für die Auswahl des Substitutionsmodells wurde jModeltest benutzt, welches durch statistische Analyse eines „best fit“-Modells und

verschiedener Strategien (Wahrscheinlichkeits-Verhältnis, Informationskriterium und Entscheidungs-Theorie) eine Rangliste aller verfügbaren Substitutionsmodelle ausgibt.

Bei der Auswahl des Clock-Modells kommt es darauf an, welche Fragestellung verfolgt wird und welche Anforderungen an den resultierenden, phylogenetischen Baum gestellt werden. Ein striktes Clock-Modell hat den Vorteil, dass es Rechenkapazität gespart wird, da für die Evolutionsrate nur eine Variable benötigt wird. Im Gegensatz dazu wird bei einem relaxierten Clock-Modell die Evolutionsrate für jeden Ast berechnet, was bei großen Alignments einen immensen Anstieg der Rechenleistung und damit auch der Laufzeit bedeutet. Dafür können jedoch die Evolutionsraten einzelner Kladen innerhalb eines Laufes miteinander verglichen werden. Allerdings ist das Risiko einer Überparametrisierung höher, was einen negativen Einfluss auf die Abschätzung der errechneten Werte haben kann.

Die Auswahl des Populationsmodells erfolgt ähnlich dem des Clock-Modells, da es auch hier eine Abwägung der höheren Rechenlaufzeit bedarf bzw. ob und welche Änderungen eine Population durchlebt hat. Ein Populationsmodell simuliert, an bestimmten Punkten, die Änderung der Populationsgröße. Die Populationsgröße hat direkten proportionalen Einfluss auf die Anzahl an Mutationsereignissen und damit der Evolutionsrate. Da im Fall der *P.sem* Hepaciviren keine Informationen über konkrete Änderungen der Population vorhanden sind, bietet es sich an diese berechnen bzw. abschätzen zu lassen. Dazu wurde ein Skygrid Model genutzt welches in vorher festgelegten Abschnitten die Änderung der Populationsgröße abschätzt.

#### 2.4.13.3. Überprüfung temporaler Signale mit TempEst

Bei TempEst handelt es sich um ein Programm zur Abschätzung der Stärke temporaler Signale in Sequenzdatensätzen. Dazu wird ein phylogenetischer Baum mit Wurzel in das Programm geladen und das Alter einer jeden Sequenz eingetragen. TempEst legt nun eine lineare Regressionsgrade durch alle Datenpunkte und gibt dabei an, wie weit dabei die Streuung aller Werte ist. Dies ist nützlich um einzelne Proben zu identifizieren, die aufgrund einer falschen Altersangabe oder Kontamination, älter bzw. jünger erscheinen als sie tatsächlich sind. Für diese Abschätzung ist jedoch ein robuster Baum, der auf stabilen temporalen Daten aufgebaut ist, nötig. Bei den vorliegenden Beprobungszeitraum (2013-2015) handelt es sich, im phylogenetischen Sinn, um einen sehr kurzen Zeitraum, der es erschwert einen Baum mit einem starken temporalen Signal zu generieren.

#### 2.4.14. Statistische Auswertung

Eine statistische Auswertung erfolgte in den Programmen GraphPad Prism, R und Excel. Für den Vergleich von zwei Gruppen wurde entweder der T-Test oder, falls die Voraussetzungen wie Normalverteilung der Werte nicht erfüllt waren, der Wilcoxon-Test. Für den Test auf Normalverteilung der Werte wurde der Shapiro-Wilk Test benutzt. Bei dem Vergleich von mehr als 2 Gruppen wurde entweder der Kruskal-Wallis-Test, für nicht-normalverteilte Werte, oder die Analyse der Varianz (ANOVA), für normalverteilte Werte, genutzt und anschließend mit einem Tukey Post-Hoc Test durchgeführt. Für die Korrelationsanalysen wurde zuerst die Stärke der Korrelation errechnet um im Anschluss daraus die Signifikanz, über eine lineare Regression, zu ermitteln. Bei sämtlichen Analysen wurde ein unteres Signifikanzniveau von  $p=0,05$  genutzt. Darüber hinaus wurden generalisierte, lineare Modelle (GLM) benutzt um den Einfluss unabhängiger Variablen (in dieser Arbeit sind es die ökologischen Faktoren) auf abhängige Variablen, wie etwa die Virusprävalenz zu testen. Dabei handelt es sich um eine lineare Regressionsfunktion, die mehrere Faktoren miteinander verknüpft und keine Normalverteilung der Variablen voraussetzt, weswegen es sich eignet, wenn man eine große Anzahl an Variablen betrachtet. Für den Vergleich von kategorischen Variablen, innerhalb des GLM, wird die erste Ausprägung innerhalb der Variable als Vergleichsregressor genutzt z.B. wird bei dem Geschlecht nur die Statistik für „Männlich“ widergegeben, da diese durch den Vergleich zu „Weiblich“ errechnet wurde. Die Güte des Modells wurde mit McFadden's Pseudo- $R^2$  berechnet (1 minus des Quotienten der Nullabweichung geteilt durch die Restabweichung). Dabei beschreibt die Güte, wie gut die Verteilung der Werte der abhängigen Variable durch die unabhängigen Variablen des Modells erklärt werden. McFadden's Pseudo- $R^2 >0,2$  steht dabei bereits für eine hohe Güte. Mit dem Chi Quadrat Test wird auf Unabhängigkeit zwischen Gruppen geprüft.

##### 2.4.14.1. Ökologische Parameter

Sämtliche Grundlegenden ökologischen Daten wurden bei der Probenahme aufgenommen. Die Proben wurden dabei von einem Team der AG Sommer (Universität Ulm) gesammelt.

Bei den ökologischen Parametern, die für Metaanalysen genutzt wurden handelte es sich um:

- Die Habitatgröße in [ha]
- Mittlere Blätterdachabdeckung
- Mittlere Blätterdachhöhe
- Mittlere Unterholzhöhe
- Anzahl an Attalea
- Anzahl an Astrocaryu
- Anzahl an Piper

Bei Attalea, Astrocaryu und Piper handelt es sich um Pflanzen die als Nahrungsquelle für Nagetiere dienen. Zusätzlich wurde mit Hilfe der Anzahl an Tieren die Speziesdiversität und die Populationsdichte berechnet. Die Speziesdiversität mit dem inversen Simpson Index berechnet. Die Formel dafür lautet:

$$(2) \quad D = \frac{N(N - 1)}{\sum_i n_i(n_i - 1)}$$

wobei  $n_i$  die Anzahl der Individuen einer Art  $i$  ist und  $N$  die Gesamtzahl aller Individuen. Der inverse Simpson Index hat als minimalen Wert 1 (keine Diversität) und steigt bei höher Diversität an. Die Populationsdichte wurde als Anzahl an Individuen einer Spezies pro Habitatfläche berechnet. Da zwischen manche dieser ökologischen Parameter eine Korrelation bzw. Autokorrelation besteht (z.B. der Habitatgröße und Populationsdichte) wurden für Regressionsanalysen vier Untergruppen aus Parametern gebildet, die gering miteinander korrelieren.

## 3. Ergebnisse und Diskussion

### 3.1. Ergebnisse

#### 3.1.1. Spezies- und Probenübersicht

Die Menge an untersuchten Blutproben (Tabelle 14) und Kotproben (Tabelle 15), aller gefundenen Spezies über alle Habitats, ist in den folgenden Tabellen gezeigt. Dabei handelt es sich um die vollständigen Datensätze aller untersuchten Proben. Dies beinhaltet auch Wiederfänge und Proben mit fehlenden Metadaten. Sämtliche der Analysen in den folgenden Kapiteln reduzieren diese Datenmengen auf unterschiedliche Weise, um den höchst möglichen Informationsgehalt zu gewährleisten.

Tabelle 14: Übersicht der Blutproben aus allen Spezies und Habitaten

Spezies	Habitats				Speziessumme
	Agrikulturgebiete	Wälder	Inseln	Plantagen	
<i>Liomys adpersus</i>				1	<b>1</b>
<i>Oecomys bicolor trinitatis</i>		1			<b>1</b>
<i>Transandinomys talamancae</i>			1		<b>1</b>
<i>Sigmodon hirsutus</i>				2	<b>2</b>
<i>Heteromys desmarestianus</i>		3			<b>3</b>
<i>Zygodontomys brevicauda</i>				3	<b>3</b>
<i>Metachirus nudicaudatus</i>	4				<b>4</b>
<i>Marmosa robinsoni isthmica</i>	5	16			<b>21</b>
<i>Hoplomys gymnurus</i>	22				<b>22</b>
<i>Philander opossum</i>	40	5		12	<b>57</b>
<i>Didelphis marsupialis</i>	73	15	8	9	<b>105</b>
<i>Proechimys semispinosus</i>	158	215	350	40	<b>763</b>
<b>Gesamt</b>	<b>302</b>	<b>255</b>	<b>359</b>	<b>67</b>	<b>983</b>

Da bei Wiederfängen nur identifiziert werden konnte ob sie bereits gefangen wurden, nicht aber welcher Probennummer dieser Fang entsprach, wurden diese Wiederfänge beispielsweise aus Metadatenanalysen entfernt. Bei Sequenzanalysen, sofern kein identischer Eintrag vorhanden war, wurden diese jedoch nicht ausgegliedert.

Tabelle 15: Übersicht der Kotproben aus allen Spezies und Habitaten

Spezies	Habitate				Speziessumme
	Agrikulturgebiete	Wälder	Inseln	Plantagen	
<b>Dasypus novemcinctus</b>	1				<b>1</b>
<b>Diplomys labilis</b>		1			<b>1</b>
<b>Transandinomys talamancae</b>			1		<b>1</b>
<b>Sigmodon hirsutus</b>				2	<b>2</b>
<b>Heteromys desmarestianus</b>		3			<b>3</b>
<b>Oecomys bicolor trinitatis</b>		1	2		<b>3</b>
<b>Zygodontomys brevicauda</b>				3	<b>3</b>
<b>Metachirus nudicaudatus</b>	4				<b>4</b>
<b>Hoplomys gymnurus</b>	17				<b>17</b>
<b>Marmosa robinsoni isthmica</b>	5	22	2	2	<b>31</b>
<b>Philander opossum</b>	47	9		8	<b>64</b>
<b>Didelphis marsupialis</b>	54	22	13	3	<b>92</b>
<b>Proechimys semispinosus</b>	165	236	392	27	<b>820</b>
<b>Gesamt</b>	293	294	410	45	<b>1042</b>

Für die ökologischen Faktoren wurde die Korrelationsstärke (Tabelle 16) und die Signifikanz dieser Stärke (Tabelle 17) in R und Excel errechnet.

**Tabelle 16: Korrelationsstärke aller ökologischen Faktoren untereinander**

Werte über Null sprechen für eine positive Korrelation, Werte unter Null für eine negative Korrelation. Die Höhe des jeweiligen Wertes bestimmt die Farbtiefe in der entsprechenden Richtung. Die Werte wurden sowohl mit R als auch Excel berechnet.

	Habitatgröße	Speziesdiversität	Mittlere Blätterdachabdeckung	Mittlere Blätterdachhöhe	Mittlere Unterholzdichte	Mittleres N Attalea	Mittleres N Astrocaryu	Mittleres N Piper	Populationsdichte
Habitatgröße	1,000								
Speziesdiversität	-0,130	1,000							
Mittlere Blätterdachabdeckung	0,308	0,302	1,000						
Mittlere Blätterdachhöhe	0,399	0,636	0,395	1,000					
Mittlere Unterholzdichte	-0,319	0,134	-0,490	-0,006	1,000				
Mittleres N Attalea	0,120	0,372	0,322	0,292	0,317	1,000			
Mittleres N Astrocaryu	0,289	-0,009	0,302	0,316	-0,185	0,495	1,000		
Mittleres N Piper	-0,075	0,406	0,210	0,034	-0,187	0,002	0,252	1,000	
Populationsdichte	-0,596	-0,348	-0,134	-0,553	-0,068	-0,311	-0,078	-0,181	1,000

Hierbei zeigte sich die größte Korrelation zwischen „Mittlere Unterholzdichte“ und „Mittlere Blätterdachabdeckung“, „Mittlere Blätterdachhöhe“ und „Populationsdichte“ sowie zwischen „Habitatgröße“ und „Populationsdichte“, wobei es sich bei allen drei um negative Korrelationen handelt. Da die Habitatgröße ein Faktor in der Berechnung der Populationsdichte ist, handelt es sich hierbei um eine Autokorrelation. Diese drei Korrelationen wurden durch die Studenten-t-Verteilung als signifikante Paarungen bestätigt.

**Tabelle 17: Signifikanz der errechneten Korrelationsstärke der ökologischen Faktoren**

Gezeigt ist die Student-t-Verteilung der Korrelationsstärken als Maß für die Signifikanz der Korrelation des Paares. Werte kleiner oder gleich 0,05 wurden grün markiert und gelten als signifikant.

	Habitatgröße	Speziesdiversität	Mittlere Blätterdachabdeckung	Mittlere Blätterdachhöhe	Mittlere Unterholzdichte	Mittleres N Attalea	Mittleres N Astrocaryu	Mittleres N Piper
Habitatgröße								
Speziesdiversität	0,673							
Mittlere Blätterdachabdeckung	0,306	0,316						
Mittlere Blätterdachhöhe	0,176	0,019	0,182					
Mittlere Unterholzdichte	0,288	0,663	0,089	0,985				
Mittleres N Attalea	0,697	0,210	0,284	0,332	0,291			
Mittleres N Astrocaryu	0,338	0,977	0,316	0,294	0,545	0,085		
Mittleres N Piper	0,807	0,169	0,492	0,912	0,541	0,994	0,407	
Populationsdichte	0,031	0,244	0,662	0,050	0,825	0,301	0,800	0,553

Obwohl nur drei Paarer als signifikant korrelierend identifiziert werden konnten, wurde zu Gunsten eines robusteren Modells, sämtliche Faktoren in vier Gruppen (Tabelle 18) aufgeteilt um in einem GLM benutzt zu werden.

**Tabelle 18: Einteilung der ökologischen Faktoren in vier Untergruppen**

	Model 1	Model 2	Model 3	Model 4
Habitatgröße	X			
Speziesdiversität	X			
Mittlere Blätterdachabdeckung	X			X
Mittlere Blätterdachhöhe				X
Mittlere Unterholzdichte		X		
Mittleres N Attalea	X			X
Mittleres N Astrocaryu		X	X	
Mittleres N Piper		X	X	X
Populationsdichte			X	

Darüber hinaus wurde in Model 5 die Abhängigkeit zu Geschlecht (m/f), Alter (Jungtier/Adult), Spezies und Habitattyp untersucht. Eine Auswahl an explorativen, grafische Darstellungen der ökologischen Faktoren ist in Abbildung 15 zu sehen.

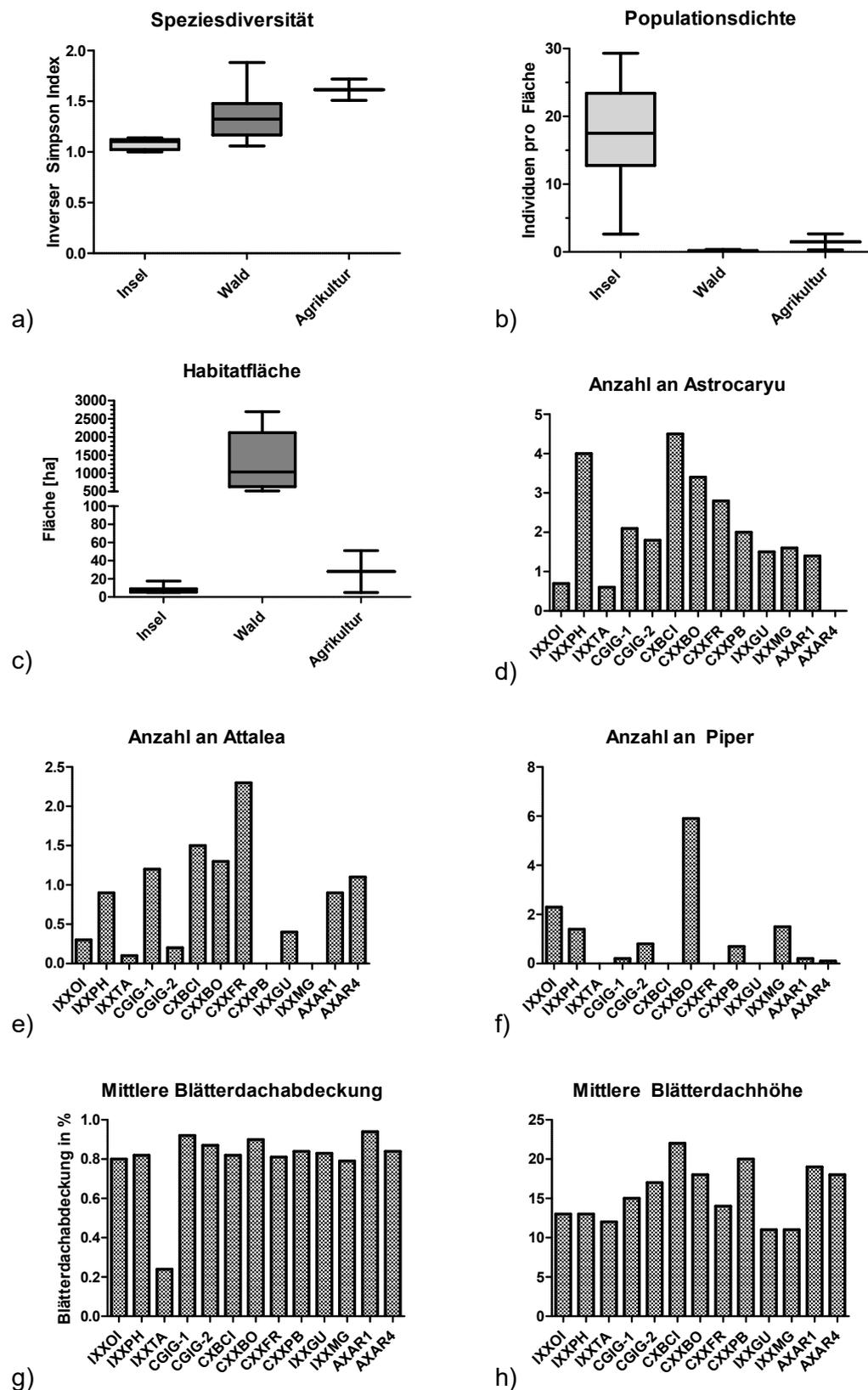
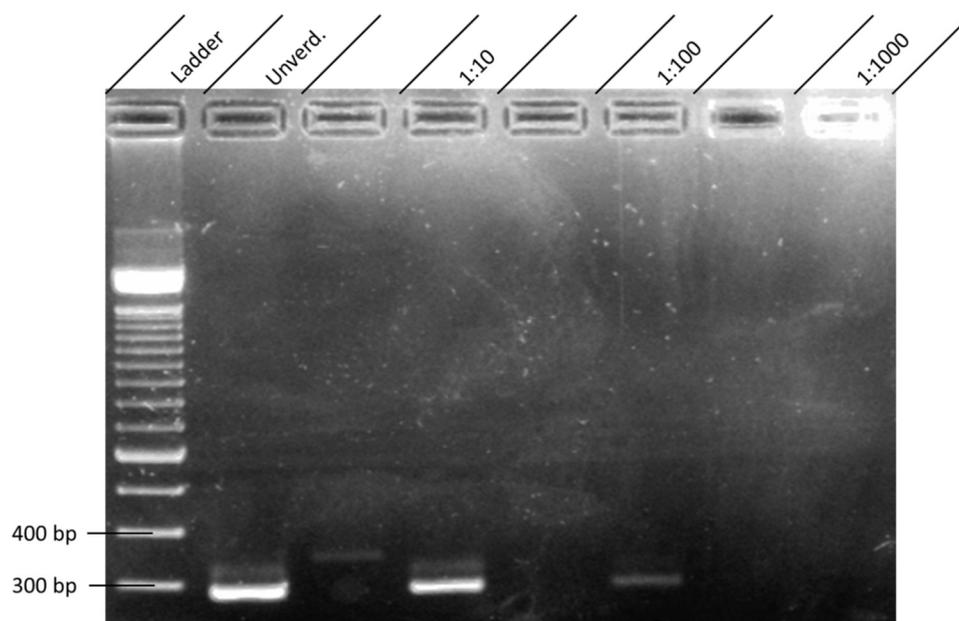


Abbildung 15: Grafische Darstellung mehrerer ökologischer Parameter

Gezeigt sind die a) Speziesdiversität berechnet über den inversen Simpson Index, b) Populationsdichte, c) Habitatfläche in [ha], Anzahl an d) Astrocaruyu, e) Attalea, f) Piper, sowie Mittlere g) Blätterdachabdeckung und h) Höhe

### 3.1.2. Picornavirus Analysen

Die Anwesenheit der Picornaviren innerhalb der *P.sem* Kot Proben wurde durch eine Pipeline Analyse eines NGS Lauf festgestellt (durchgeführt durch die AG Jones, Institut für Virologie der Charité Berlin). Mithilfe dieser Informationen konnte eine initiale Screening PCR designt werden (Tabelle 3). In insgesamt 6 Proben wurden Picornavirus Reads gefunden. Nach Sichtung der Reads konnte festgestellt werden, dass es sich um drei verschiedene Varianten handelt, weswegen für die Screening PCR eine Kombination aus mehreren Forward und Reverse Primern designt wurde (Tabelle 5). Dieser Screening Assay wurde auf bekannten positiven Proben (aus der Identifikation im NGS) getestet um die Sensitivität zu ermitteln. Hierfür wurde eine Verdünnungsreihe der Proben von 1:10 bis 1:10<sup>7</sup> angesetzt. Abbildung 16 zeigt hierbei das Agarosegel der PCR bis zu Verdünnung 1:1000, welches die ersten Verdünnung ohne Bande war (höhere Verdünnungen zeigten auch keine Banden und wurden deswegen nicht dargestellt).



**Abbildung 16: Ausschnitt des Sensivitätstest der Picornavirus Screening PCR**

Gezeigt ist ein 1,5%iges Agarosegel der ersten Runde der Picorna Screening PCR. Zu sehen sind die DNA-Leiter (Ladder) mit Markierung bei 400 und 300 bp. Gelbahnen sind mit der jeweiligen Probe beschriftet (unverdünnte Probe, 1:10 Verdünnung, 1:100 Verdünnung und 1:1000 Verdünnung). Die zu erwartende Höhe der ersten Runde des Assays liegt bei 284 bp.

Durch die Verdünnungsreihe konnte festgestellt werden, dass die PCR selbst nach einer 1:100 Verdünnung von extrahierten Kotproben noch Virus-RNA nachweisen kann. Im Anschluss wurden 112 Pools gescreent, die aus 587 Kotproben bestanden (die Größe der Pools variiert hier zwischen 5 und 10 Proben pro Pool). 26 der Pools wurden positiv getestet, woraufhin die einzelnen Proben der jeweiligen Pools (n=144) individuell nachgetestet wurden. Für 47/144 Pools konnte eine Proben im Screening Assay positiv getestet werden. Zur Bestätigung der Anwesenheit von Picornaviren, wurde die Sequenz jedes positiven PCR Produkts per Sanger-Sequenzierung ermittelt. Im Folgenden wurden 486 Proben, ohne vorherige Pooltestung, gescreent. Von diesen 486 gescreenten Proben, waren weitere 47 PCR positiv. Es wurden ausschließlich Picornaviren in *P.sem* gefunden (Tabelle 19). Eine Probe, der 94 PCR positiv getesteten Kotproben, gehörte ebenfalls zu den 6 Proben des initialen NGS Laufs, weswegen die Gesamtzahl an Picornavirus positiven Proben nur 99 statt 100 ist. Die sich daraus ergebende Picornavirusprävalenz innerhalb von *P.sem* ist in Tabelle 20 zu sehen.

**Tabelle 19: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von Picornavirus-RNA untersucht wurden**

Gezeigt sind die Ergebnisse der Suche nach Picornaviren in Kotproben. Die Ergebnisse umfassen Funde aus der Screening PCR sowie aus NGS Läufen.

Spezies	negativer Picornavirus Nachweis	positiver Picornavirus Nachweis	Speziessumme
<i>Dasyus novemcinctus</i>	1	0	1
<i>Diplomys labilis</i>	1	0	1
<i>Transandinomys talamancae</i>	1	0	1
<i>Sigmodon hirsutus</i>	2	0	2
<i>Zygodontomys brevicauda</i>	3	0	3
<i>Heteromys desmarestianus</i>	3	0	3
<i>Oecomys bicolor trinitatis</i>	3	0	3
<i>Metachirus nudicaudatus</i>	4	0	4
<i>Hoplomys gymnurus</i>	17	0	17
<i>Marmosa robinsoni isthmica</i>	31	0	31
<i>Philander opossum</i>	64	0	64
<i>Didelphis marsupialis</i>	92	0	92
<i>Proechimys semispinosus</i>	721	99	820
<b>Gesamtergebnis</b>	<b>943</b>	<b>99</b>	<b>1042</b>

**Tabelle 20: Picornavirusprävalenz Aufschlüsselung pro Habitattyp**

Gezeigt sind die Ergebnisse aus der Kombination von gerichtetem Virusscreening (PCR) und ungerichteter NGS Analyse. Hier sind die Anzahl der Funde pro Habitattyp aufgelistet, sowie der prozentuale Anteil an positiven Funden pro Habitattyp. Die Habitattypen sind Agrikulturgebiete (A), Wälder (C), Inseln (I) und Plantagen (P). Es handelt sich um die Prävalenz innerhalb von *P.sem* da nur in dieser Spezies Picornaviren gefunden wurden.

Habitattyp	negative Picornavirus Nachweis	positiver Picornavirus Nachweis	Summe pro Habitattyp	Picornavirusprävalenz (%)
<b>A</b>	145	20	<b>165</b>	<b>12,12</b>
<b>C</b>	207	29	<b>236</b>	<b>12,29</b>
<b>I</b>	344	48	<b>392</b>	<b>12,24</b>
<b>P</b>	25	2	<b>27</b>	<b>7,41</b>
<b>Gesamtergebnis</b>	721	99	<b>820</b>	<b>12,07</b>

Die Sequenzen, der insgesamt 94 PCR positiven Proben wurden per Sanger-Sequenzierung bestimmt. Diese kurzen Sequenzen (100 – 250 bp) wurden zusammen mit den sechs Sequenzen, auf denen die Screening PCRs designt wurden (~500bp bzw. ein Vollgenom), aligniert. In einer explorativen, phylogenetischen Baum konnten alle Sequenzen optische in 3 Gruppen eingeteilt werden (Tabelle 21).

**Tabelle 21: Numerische Verteilung der Picornavirussequenzen innerhalb der definierten Gruppen**  
Übersicht der Sequenzanzahl innerhalb der definierten Picornavirusgruppen

Bezeichnung	Anzahl an Sequenzen
<i>P.sem</i> Picornavirus Gruppe 1	29
<i>P.sem</i> Picornavirus Gruppe 2	38
<i>P.sem</i> Picornavirus Gruppe 3	30

Dieses Alignment aus 99 *P.sem* Picornavirussequenzen wurde zusammen mit dem ICTV Alignment der 3D Region aligniert und ein phylogenetischer Baum mit einem ML Algorithmus berechnet (Abbildung 17). In der Abbildung sind aus Gründen der Übersichtlichkeit nur der längste Vertreter einer Gruppe, sowie das einzelne Vollgenom dargestellt. Dabei entfiel der größte Teil der Sequenzen auf die näherverwandten Gruppen 2 und 3 (Tabelle 21 und Tabelle 22)

Um eine Aussage darüber zu Treffen in wie weit sich die drei Gruppen unterscheiden und wie nahe verwandt sie sind, wurde die genetische Distanz auf Nukleotid- und Aminosäureebene betrachtet (Tabelle 22 und Tabelle 23).

**Tabelle 22: Prozentuale Identität auf Nukleotidebene der *P.sem* Picornavirusgruppen untereinander und zu Referenzsequenzen**

Übersicht über die paarweise Identität auf Nukleotidebene zwischen den Gruppen. Die Distanzmatrix stammt aus dem Alignment dem der ML Baum aus Abbildung 17 zugrunde liegt. Die Identität zwischen den Gruppen und zur nächst-verwandten Referenz ist hierbei anhand der ~500bp Fragmenten der Gruppen errechnet worden.

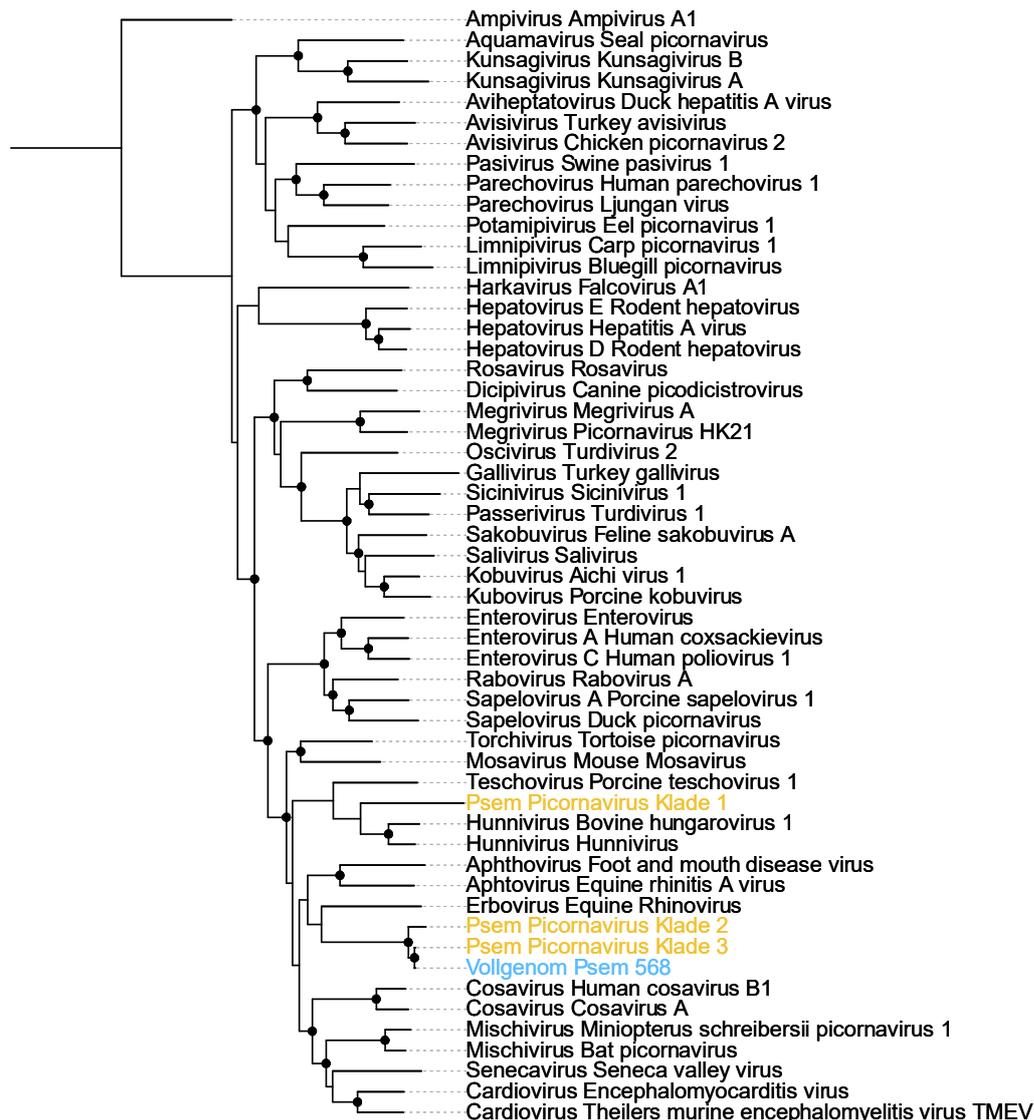
	Gruppe 1	Gruppe 2	Gruppe 3	Nächste verwandte Referenz
Gruppe 1	87,76-100%			Hunnivirus (46,31%)
Gruppe 2	44,64%	84,67-100%		Cardiovirus –Theilers murine encephalomyelitis-virus [TMEV] (45,85%)
Gruppe 3	48,51%	84,66%	77,26-100%	Cardiovirus –Theilers murine encephalomyelitis-virus [TMEV] (45,57%)

**Tabelle 23: Prozentuale Identität auf Aminosäureebene der *P.sem* Picornavirusgruppen untereinander und zu Referenzsequenzen**

Übersicht über die paarweise Identität auf Aminosäureebene zwischen den Gruppen. Die Distanzmatrix stammt aus dem Alignment dem der ML Baum aus Abbildung 17 zugrunde liegt, nachdem es translatiert wurde. Die Identität zwischen den Gruppen und zur nächst-verwandten Referenz ist hierbei anhand der ~500bp Fragmenten der Gruppen errechnet worden.

	Gruppe 1	Gruppe 2	Gruppe 3	Nächste verwandte Referenz
Gruppe 1	87,43-100%			Hunnivirus (76,58%)
Gruppe 2	44,64%	84,81-100%		Erbovirus (46,15%)
Gruppe 3	48,51%	98,23%	85,55-100%	Erbovirus (44,73%)

Hier ist zu sehen, dass trotz der geringen Fragmentlänge (~100-250 bp) eine relativ hohe Grundvariabilität innerhalb aller Gruppen zu sehen ist. Ebenso sind die Distanzen zu bekannten Referenzsequenzen höher als die Grundvariabilität.



**Abbildung 17: Phylogenetischer Baum der 3D Region von Picornaviren**

Der phylogenetische Baum beinhaltet eine 3D Region des Vollgenoms der Picornaviren (blau) und drei Picornavirus Screening PCR Fragmente (~500bp) (gelb) aus *P.sem*. Die Referenzsequenzen stammen aus dem ICTV Alignment der 3D Region von *Picornaviridae*. Das Gesamtalignment wurde mit dem dem MAFFT Plugin Version 1.3.6 [22] in Geneious generiert und der ML Baum mit dem PHYML Plugin Version 2.2.3 [23], mit 1000 Bootstraps und GTR Substitutionsmatrix, berechnet. Schwarze Kreise markieren alle Knoten mit einem Support >75%. Außengruppe hier ist das Ampivirus.

Für eine genauere taxonomische Einordnung, wurde versucht die Menge an Sequenzinformationen zu erhöhen. Hierfür wurden Genomprimer mithilfe der vorhandenen NGS Datensätze designt um Vollgenome der *P.sem* Picornaviren zu erstellen. Dabei handelte es sich um sechs, sich jeweils überlappende Hemi-nested PCRs mit einer Fragmentgröße über 1100 bp, die das gesamte Genom abdecken. Insgesamt konnte ein Vollgenom komplett sequenziert werden, welches sich im Vergleich zur ersten NGS Sequenz nur um 6 Basenpaare unterscheidet.

Dieses Vollgenom fügt sich in Gruppe 3 ein, aus dem auch das analoge, kürzere Fragment derselben Probe stammt.

Im Folgenden wurde die Genbank Accession numbers aus dem ICTV 3D Regio Alignment genutzt um alle vollständigen Sequenzen daraus zu erhalten. Mit 50 dieser Referenzsequenzen und dem *P.sem* Picornavirus Vollgenom, wurde über den MAFFT Algorithmus ein Nukleotidalignment erstellt. Dieses wurde in ein Aminosäurealignment übersetzt. Aus diesem Alignment wurden jeweils die P1, 2C, 3C und 3D Regionen extrahiert und realigniert. Diese vier Alignments wurden benutzt um jeweils einen ML Baum über das PHYML Plugin in Geneious zu berechnen (1000 Bootstraps mit GTR Substitutionsmatrix). Einzig in dem Baum der 3C Region clusterte das Vollgenom monophyletisch mit einer Referenz, dem Rosavirus. Die maximale Identität der jeweiligen AA Alignments ist in Tabelle 24 zu sehen.

**Tabelle 24: AA Identität des Picornavirus Vollgenomes zu nahverwandten Referenzen**

Gezeigt sind die maximalen Identitäten pro AA Alignment der Regionen P1, 2C, 3C und 3D von Picornaviren. Die Alignments wurden in Geneious mit dem MAFFT Plugin erstellt. Dabei wurde aus 50 Referenzen und dem *P.sem* Picornavirusvollgenom

Region	Maximale Identität nahverwandtesten Referenz	Name der Referenz
P1	29,74%	Kubovirus; Porcine Kobuvirus
2C	42,61%	Erbovirus; Equine Rhinovirus
3C	27,26%	Rosavirus; Rosavirus
3D	41,03%	Aphtovirus; Equine rhinitis A virus

Abschließend wurden GLM genutzt um den Einfluss der ökologischen Faktoren auf die Virusverteilung zu zeigen (Tabelle 25). Dabei wurden die fünf Modelle benutzt wie in Kapitel 3.1.1 beschrieben. Die 820 Proben wurden bereinigt, so das keine fehlenden Metadaten mehr vorhanden waren. Dies reduzierte den Datensatz auf 649 Proben.

**Tabelle 25: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von Picornaviren**

Gezeigt sind die Ergebnisse des, in R berechneten, GLM der detektierten Picornavirus Infektionen in Abhängigkeit zu ökologischen Faktoren. Als Vergleichsregressor für kategoriale Variablen diente der erste Wert der Variable. Als Güte wurde McFadden's Pseudo-R<sup>2</sup> berechnet. Für die Regression wurden jeweils 649 Proben benutzt. In Model 5 werden, ausser *P.sem*, nur signifikante Spezies gezeigt. Signifikanzbeschriftung wurde wie folgt festgesetzt:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

Model 1 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.095)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Habitatgröße	-7.18E-05	1.74E-05	4.30E-05	***
Biodiversität	-1.42E-01	1.98E-02	1.67E-12	***
Blätterdachabdeckung	-9.27E-03	2.34E-03	8.38E-05	***
Anzahl Attalea	1.03E-02	1.72E-02	0.553	

Model 2 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.042)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Unterholzdichte	-0.006903	0.002002	0.000602	***
Anzahl Astrocaryu	0.006747	0.011308	0.550945	
Anzahl Piper	-0.042333	0.009057	3.60E-06	***

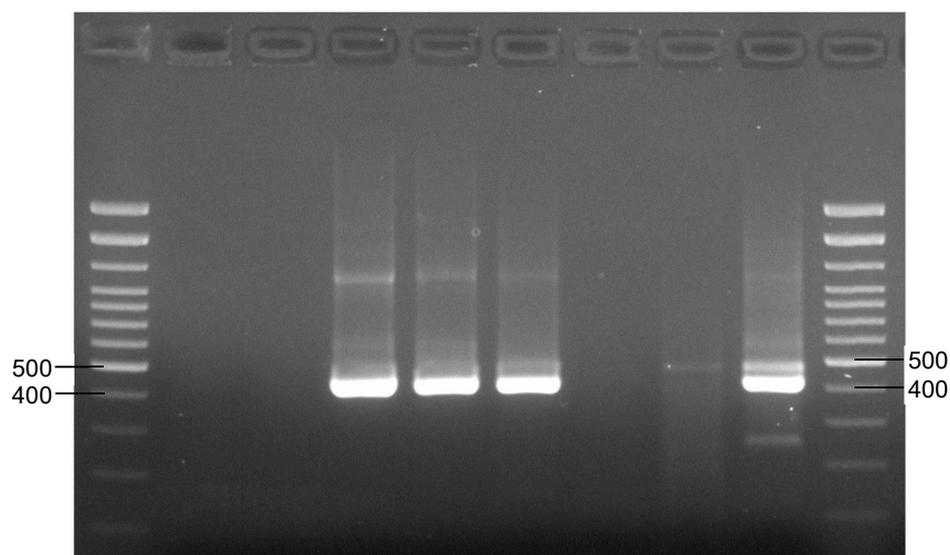
Model 3 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.057)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Anzahl Astrocaryu	0.006787	0.011188	0.544338	
Anzahl Piper	-0.027072	0.008796	0.002174	**
Populationsdichte	0.005806	0.001233	3.03E-06	***

Model 4 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.050)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Blätterdachabdeckung	-0.002644	0.00265	0.31873	
Blätterdachhöhe	-0.014059	0.004384	0.00141	**
Anzahl Attalea	0.004941	0.017774	0.78109	
Anzahl Piper	-0.025849	0.008485	0.00241	**

Model 5 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.077)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Sex	-0.0363317	0.0253793	0.152765	
Alter	-0.0028863	0.0454958	0.949435	
Spezies <i>P.sem</i>	0.1290895	0.0428424	0.002689	**
Habitattyp C	-0.0020093	0.0295551	0.945819	
Habitattyp I	0.1477182	0.0391211	0.000174	***

### 3.1.3. Picobirnavirus Analysen

Durch die Hochdurchsatzsequenzierung von *P.sem* Kotproben wurden virale Reads detektiert, die der Familie der *Picobirnaviridae* zugeordnet werden konnten. Darüber hinaus gelang es eine Konsensussequenz zu generieren, die das gesamte kleinere Fragment abdeckte. Dieses Fragment kodiert die RdRp, die ein essentieller Teil der Virusreplikation darstellt und somit hoch konserviert ist. Aus diesem Grund wurde diese Region für die Screening PCR ausgewählt. Zur Überprüfung der Spezifität der Screening PCR wurde zunächst eine Visualisierung per Agarosegel durchgeführt (Abbildung 18). Amplifikate mit der zu erwarteten Größe (446 bp) wurden Sanger sequenziert. Nach der erfolgreichen Sequenzierung von 434 Proben, wurde auf die Generierung weiterer Sequenzen, zur Assaybestätigung verzichtet und Proben mit einer entsprechenden Bandengröße als positiv bewertet.



**Abbildung 18: Beispielhafte Visualisierung der Amplifikate der Screening PCR für PBV**

Zu sehen sind die Amplifikate der ersten Runde der Screening PCR für PBV, aufgetrennt in einem 1,5%iges Agarosegel. Die 400 und 500 bp Banden auf der DNA Leiter sind beschriftet. Erwartete Bandenhöhe für das PCR Produkt ist 446 bp.

Für 198 der 434 sequenzierten Proben konnten Konsensussequenzen generiert werden, die eine ausreichende Qualität sowie eine Mindestlänge von 360 bp erreichten. Dabei entstand jedoch kein Konsensussequenz aus Proben, die aus dem Habitattyp „Plantage“ stammen, weswegen bei den Sequenzanalysen nur die drei Habitattypen „Wald“, „Insel“ und „Agrikulturgebiet“ vorkommen. Diese Vortestung ergab eine hohe Abundanz an PBV in den Kotproben, weswegen auf eine Pooltestung verzichtet und Einzelproben direkt

analysiert wurden. In 626/1042 (60,08 %) der Kotproben konnte PBV RNA detektiert werden (siehe Tabelle 26 und Tabelle 27).

**Tabelle 26: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von PBV-RNA untersucht wurden**

Gezeigt sind die Ergebnisse der Suche nach PBV in allen Kotproben.

Spezies	negativer PBV Nachweis	positiver PBV Nachweis	Speziessumme
<i>Dasypus novemcinctus</i>	0	1	1
<i>Diplomys labilis</i>	1	0	1
<i>Transandinomys talamancae</i>	0	1	1
<i>Sigmodon hirsutus</i>	0	2	2
<i>Zygodontomys brevicauda</i>	0	3	3
<i>Heteromys desmarestianus</i>	1	2	3
<i>Oecomys bicolor trinitatis</i>	0	3	3
<i>Metachirus nudicaudatus</i>	1	3	4
<i>Hoplomys gymnurus</i>	11	6	17
<i>Marmosa robinsoni isthmica</i>	28	3	31
<i>Philander opossum</i>	40	24	64
<i>Didelphis marsupialis</i>	50	42	92
<i>Proechimys semispinosus</i>	284	536	820
<b>Gesamtergebnis</b>	416	626	1042

**Tabelle 27: Picobirnavirusprävalenz aufgeschlüsselt nach Habitattyp**

Gezeigt sind die Ergebnisse aus der Kombination von gerichtetem Virusscreening (PCR). Hier sind die Anzahl der Funde pro Habitattyp aufgelistet, sowie der prozentuale Anteil an positiven Funden pro Habitattyp. Die Habitattypen sind Agrikulturgebiete (A), Wälder (C), Inseln (I) und Plantagen (P)

Habitattyp	negativer PBV Nachweis	positiver PBV Nachweis	Summe pro Habitattyp	PBV-prävalenz (%)
<b>A</b>	131	162	<b>293</b>	<b>55,29</b>
<b>C</b>	140	154	<b>294</b>	<b>55,38</b>
<b>I</b>	111	299	<b>410</b>	<b>72,93</b>
<b>P</b>	34	11	<b>45</b>	<b>24,44</b>
<b>Gesamtergebnis</b>	416	626	<b>1042</b>	<b>60,08</b>

In Tabelle 28 ist eine Übersicht über die Spezies und Fangorte der Tiere dargestellt, aus denen diese Sequenzen generiert werden konnten.

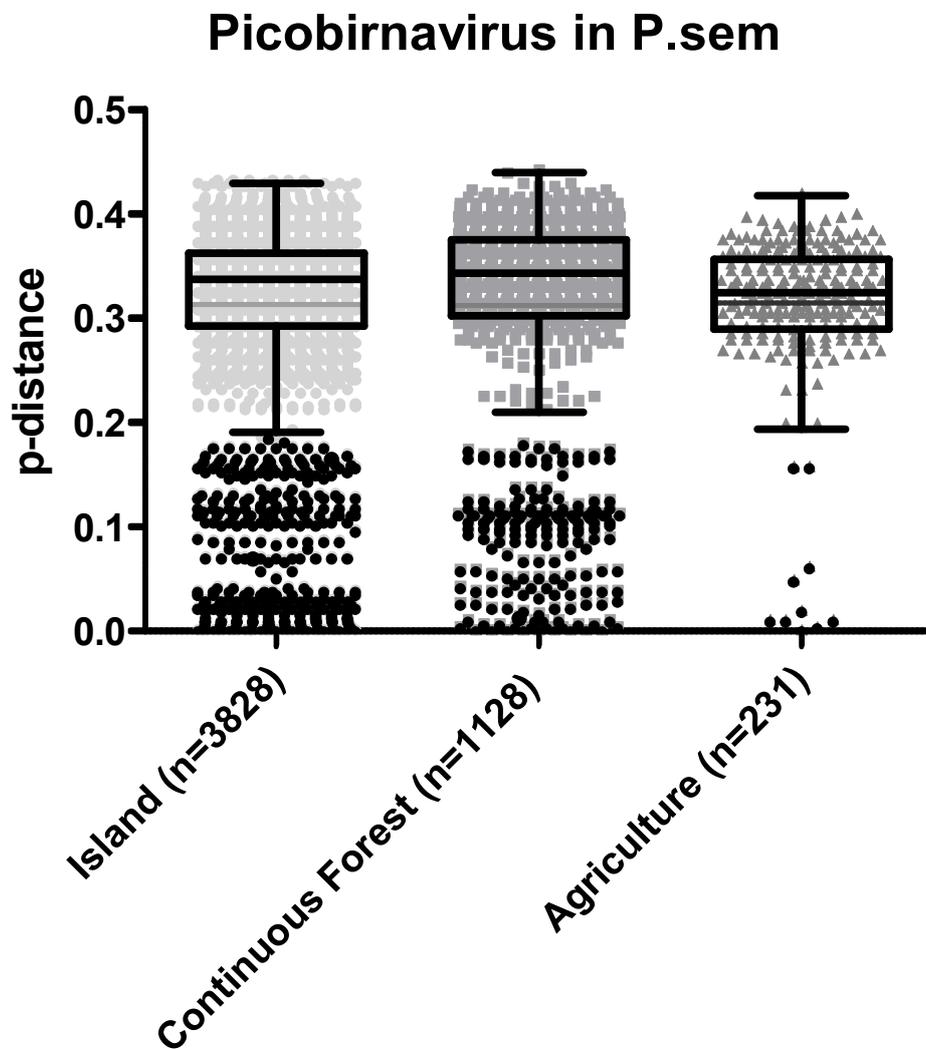
**Tabelle 28: Speziesverteilung innerhalb der PBV-Sequenzcontigs**

Übersicht der Speziesverteilung der Sequenzcontigs die eine ausreichende Qualität und Mindestlänge von 360 bp verfügten.

Spezies	Habitattypen			Anzahl
	Agrikulturgebiete	Wälder	Inseln	
<i>Marmosa robinsoni isthmica</i>	0	1	0	1
<i>Transandinomys talamancae</i>	0	0	1	1
<i>Dasyopus novemcinctus</i>	1	0	0	1
<i>Metachirus nudicaudatus</i>	3	0	0	3
<i>Oecomys bicolor trinitatis</i>	0	1	2	3
<i>Hoplomys gymnurus</i>	4	0	0	4
<i>Philander opossum</i>	12	0	0	12
<i>Didelphis marsupialis</i>	9	3	3	15
<i>Proechimys semispinosus</i>	22	48	88	158
<b>Gesamtergebnis</b>	<b>51</b>	<b>53</b>	<b>94</b>	<b>198</b>

Da die Grunddiversität, laut ICTV in allen bislang bekannten PBV sehr hoch ist und keine Einordnung anhand mehrere Genomregion, wie im Falle von Picornaviren, möglich ist, wurde als erstes die Diversität der gefunden PBV ermittelt. Um die Diversität dieser Sequenzen zu untersuchen, wurde ein Nukleotid-Alignment erstellt. Die genetische Distanz reichte von 0% (identische Sequenzen) bis zu 50,4%. Die höchste PBV-Diversität wurde dabei zwischen Sequenzen aus *Didelphis marsupialis* und *P.sem* beobachtet. Im nächsten Schritt wurden die Sequenzen, der Generalistenspezies *P.sem*, welche das höchste Aufkommen hatten, allein betrachtet. Die genetische Distanz innerhalb des *P.sem* Nukleotid-Alignments lag zwischen 0% und 45,5%. In MEGA wurde mit dem Distanzmatrix-Tool die paarweise p-Distanzen, des Nukleotidalignments berechnet. Dabei entstanden aus 22 Sequenzen des Habitattyp „A“ (Agrikulturgebiete), 231 Distanzpaare. Für Habitattyp „C“ (Wald) wurden aus 48 Sequenzen 1128 Distanzpaare und für Habitattyp „I“ (Insel) entstanden aus 88 Sequenzen 3828 Distanzpaare. Dabei zeigt sich, dass die mittlere Nukleotiddistanz innerhalb der einzelnen Habitattypen bei ca. 31% lag (Abbildung 19). Erwartungsgemäß ist dabei die p-Distanzen dabei nicht normalverteilt. Dies wurde per Shapiro-Wilk-Test in GraphPad Prism überprüft und ergab, dass keine der drei p-Distanz Datensets laut des Tests nicht normalverteilt vorlagen ( $p < 0.0001$  für alle drei Datensets). Aus diesem Grund wurde, zur Errechnung des Vorhandenseins eines signifikanten Unterschieds zwischen den Gruppen, ein Kurskal-Wallis-Test mit Dunn's Post-Hoc-Test durchgeführt. Dieser errechnete einen signifikanten Unterschied zwischen den Habitaten Wald/Insel, sowie

Wald/Agrikulturgebiet. Da dies der grafischen Auswertung des Boxplots widersprach wurde die Effektstärke dieser Paare errechnet.



**Abbildung 19: Nukleotiddistanz von PBV in *P.sem* innerhalb der drei Habitattypen.**

Kombination aus Boxplot mit unterliegendem Scatterplot. Der Boxplot zeigt Median (dicke schwarze Linie), Interquartilsabstand und Whiskers nach Tukey (1,5-Fache des Interquartilsabstands) und der Scatterplot den Mittelwert (dünne, graue Linie) der Nukleotiddistanzen. Werte außerhalb der Whiskers sind schwarz gefärbt.

Die Berechnung der Effektstärke, oder auch Cohen's D [103], erfolgte hierbei anhand der Formeln (3) und (4). Die Ergebnisse dieser Berechnung sind in Tabelle 29 zu sehen.

$$(3) \quad \text{Effektstärke } (d) = \left| \frac{\text{Mittelwert A} - \text{Mittelwert B}}{\text{Standardabweichung } (s)} \right|$$

$$(4) \quad \text{Standardabweichung } (s) = \sqrt{\frac{s_A^2 * n_A + s_B^2 * n_B}{n_A + n_B}}$$

Dabei gilt:

Mittelwert A und B entsprechen den Mittelwerten der p-Distanzen der jeweiligen Gruppen

Standardabweichung Gruppe A ( $s_A$ ) und Gruppe B ( $s_B$ )

Jeweiligen Probenanzahl von Gruppe A oder B ( $n_A$  bzw.  $n_B$ )

**Tabelle 29: Berechnete Effektstärken der p-Distanzen von Picobirnaviren**

Gezeigt sind die Interpretationen der Effektstärken nach Cohen [103], sowie die errechneten Effektstärken zwischen den drei Habitattypen. Jede Effektstärke der Habitatpaare ist dabei kleiner als das untere Niveau der Cohen Effektstärke, was bedeutet, dass es sich hierbei um nicht signifikante Effekte handeln.

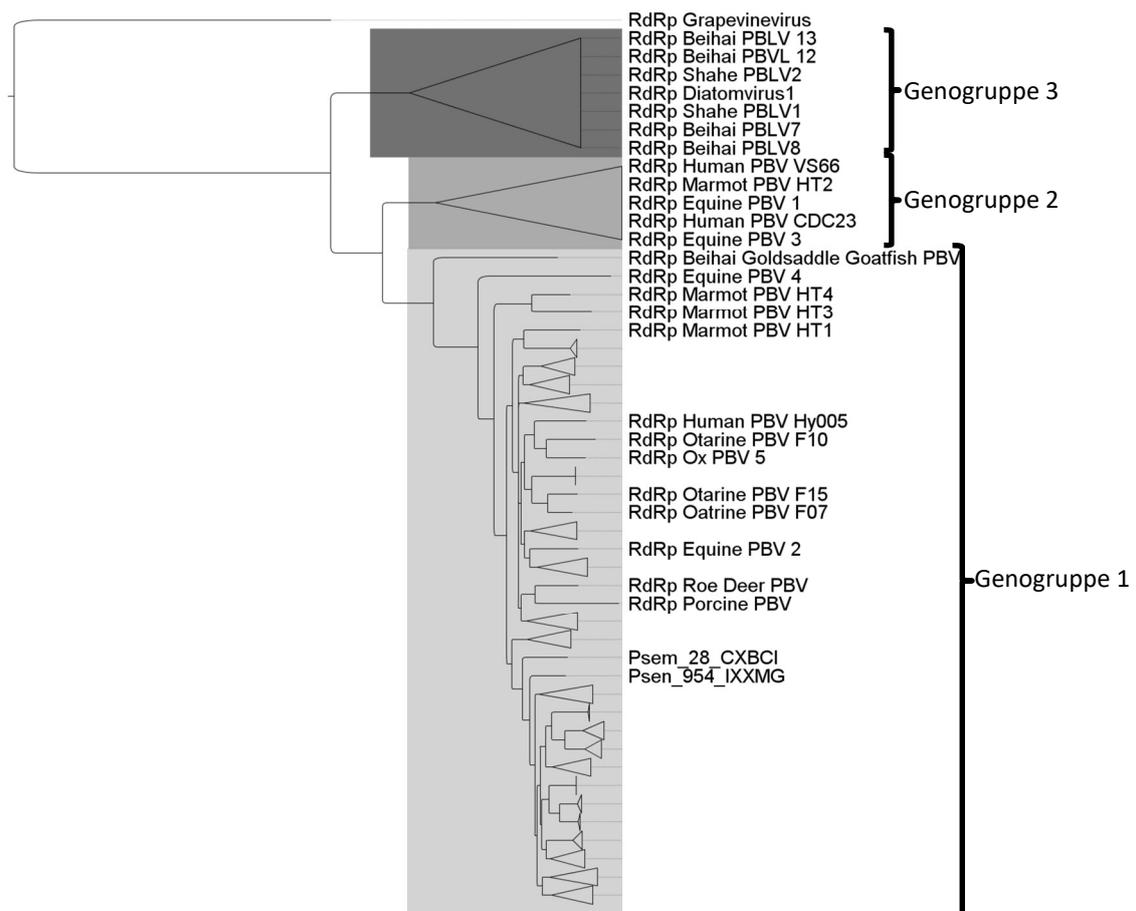
	Effektstärke (Cohen's d)
Schwache Effektstärke	0.2
Mittlere Effektstärke	0.5
Große Effektstärke	0.8
Effektstärke Insel gegen Wald	<b>0.0008</b>
Effektstärke Insel gegen Agrikultur	<b>0.0011</b>
Effektstärke Wald gegen Agrikultur	<b>0.0014</b>

Durch die fehlende Effektstärke, ist der Kurskal-Wallis-Test nicht länger aussagekräftig und es wird angenommen, dass es zwischen den p-Distanzen der drei verglichenen Gruppen keinen signifikanten Unterschied gibt.

Zum jetzigen Zeitpunkt umfasst, laut ICTV, der Genus Picobirnavirus drei Genogruppen. Wobei Gruppe 1 und 2 Wirbeltiere und Genogruppe 3 Wirbellose infiziert. Hierbei stellt die Grunddiversität ein großes Problem dar für eine Speziesordnung. Das ICTV

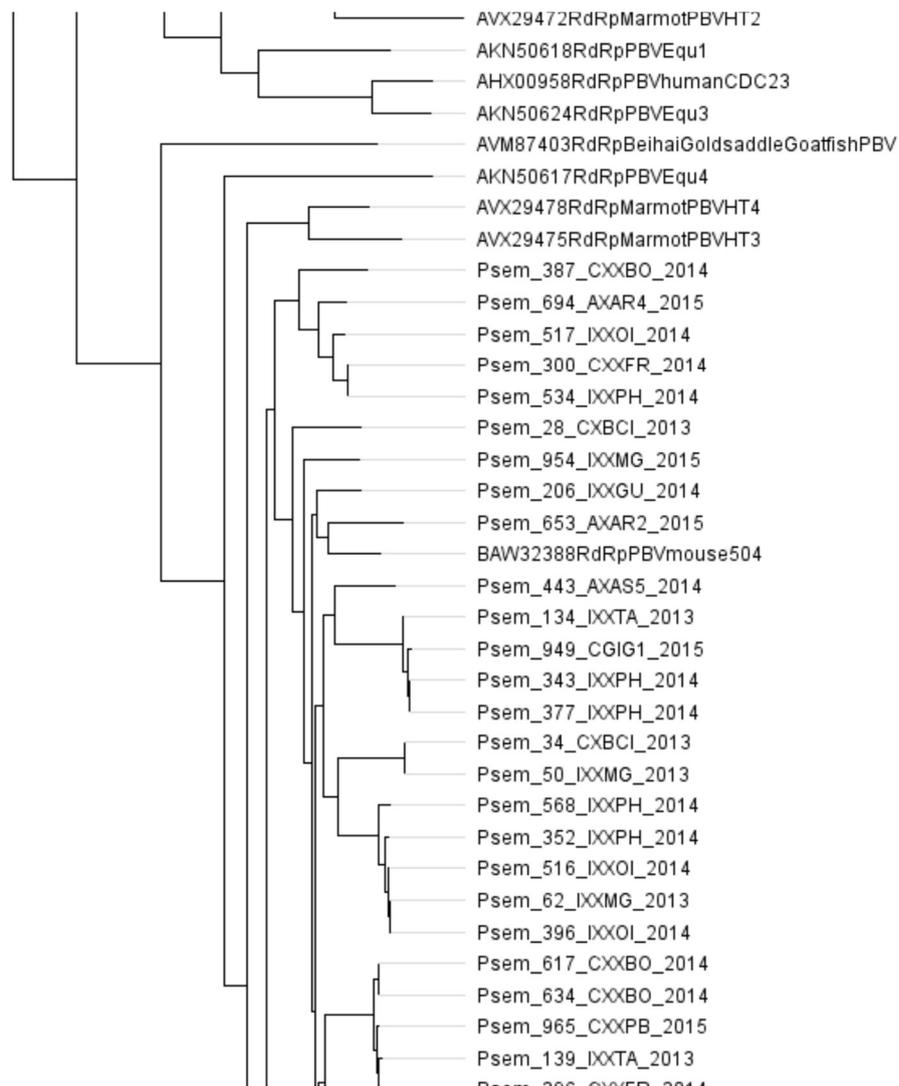
schreibt hierzu: "In Anbetracht des hohen Maßes an Divergenz der RdRP- und Kapsidsequenzen unter den vollständig sequenzierten Picobirnaviren (20 infizieren Wirbeltiere und 7 wurden in Wirbellosen gefunden), könnte jedes der Picobirnaviren, die durch ein vollständig sequenziertes Genom definiert sind, letztendlich einer eigenen Virusart zugeordnet werden. Paarweise Alignments der RdRP-Aminosäuresequenzen vollständig sequenzierter Picobirnaviren zeigen die Existenz von drei Genogruppen mit einem hohen Maß an Sequenzdivergenz in jeder dieser Genogruppen., (Quelle: ICTV Homepage, Unterseite *Picobirnaviridae* )

Aus diesem Grund wurden nur die Einordnung in Genogruppen durchgeführt, jedoch aber auf eine taxonomische Aufteilung nach Subgruppen verzichtet, da es bis jetzt aufgrund der hohen Diversität an einer strukturellen Unterkategorie mangelt. Hierfür wurden die Nukleotidsequenzen in Aminosäuresequenzen übersetzt und zusammen mit Referenzen, der jeweiligen Genogruppen, zusammen aligniert. Aus diesem Alignment wurde mit dem Neighbor-Joining Algorithmus ein phylogenetischer Baum errechnet (Abbildung 20). In diesem NJ-Baum ist zu sehen, dass es sowohl Murmeltier (Marmot) als auch Humane PBV Sequenzen gibt, die sowohl Genogruppe 1 als auch Genogruppe 2 angehören. Gruppen an Sequenzen, die durch kurze Äste verbunden waren, wurden in Kladen kollabiert (hier durch entstanden 21 Untergruppen). Da diese Untergruppen jedoch nicht monophyletisch für ein Habitat und die Parameter nachdem sie kollabiert wurde, willkürlich gewählt waren, wurde diese Gruppeneinteilung nicht für Analysen herangezogen.



**Abbildung 20: Phylogenetischer Baum der Genogruppen von PBV inkl. Referenzen**

NJ-Baum mit 1000 Bootstraps und Jukes-Cantor Substitutionsmodell mit Aminosäuresequenzen. Grapevinevirus wurde als Außengruppe definiert. Hervorgehoben sind die drei Genogruppen: 1, 2 und 3. Aufgrund der Übersichtlichkeit sind die meisten *P.sem* PBV Kladen zusammengeklappt. Alle PBV aus *P.sem* befinden sich in Genogruppe 1. Die Referenzsequenzen inkl. der Außengruppe wurden ICTV entnommen.



**Abbildung 21: Vergrößerter Ausschnitt des phylogenetischer Baum der Genogruppen von PBV inkl. Referenzen**

Vergrößerter Abschnitt des NJ-Baumes aus Abbildung 20. Zu sehen ist der obere Abschnitt der Genogruppe 1 ohne kollabierte Kladen.

In diesem Baum konnte kein phylogeographischer Zusammenhang erkannt werden (Abbildung 21), da die Kladen zum Großteil aus Sequenzen unterschiedlicher Habitate bestanden (welche durch eine größere geographische Distanz getrennt waren). Abschließend wurden GLM genutzt um den Einfluss der ökologischen Faktoren auf die Virusverteilung zu zeigen (Tabelle 30). Dabei wurden die fünf Modelle benutzt wie in Kapitel 3.1.1 beschrieben. Die 820 Proben wurden bereinigt, so das keine fehlenden Metadaten mehr vorhanden waren. Dies reduzierte den Datensatz auf 649 Proben.

**Tabelle 30: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von PBV**

Gezeigt sind die Ergebnisse des, in R berechneten, GLM der detektierten PBV Infektionen in Abhängigkeit zu ökologischen Faktoren. Als Vergleichsregressor für kategoriale Variablen diente der erste Wert der Variable. Als Güte wurde McFadden's Pseudo-R<sup>2</sup> berechnet. Für die Regression wurden jeweils 649 Proben benutzt. In Model 5 werden, ausser *P.sem*, nur signifikante Spezies gezeigt. Signifikanzbeschriftung wurde wie folgt festgesetzt:

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

Model 1 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.079)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Habitatgröße	8.93E-06	2.65E-05	0.7362	
Biodiversität	-4.25E-02	3.00E-02	0.15751	
Blätterdachabdeckung	5.16E-06	3.56E-03	0.99884	
Anzahl Attalea	-1.87E-01	2.62E-02	2.80E-12	***

Model 2 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.041)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Unterholzdichte	-0.014736	0.003018	1.32E-06	***
Anzahl Astrocaryu	-0.009409	0.017047	0.581	
Anzahl Piper	0.006748	0.013654	0.621	

Model 3 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.065)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Anzahl Astrocaryu	-0.009747	0.016784	0.56161	
Anzahl Piper	0.03866	0.013195	0.00351	**
Populationsdichte	0.011843	0.001849	2.89E-10	***

Model 4 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.079)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Blätterdachabdeckung	0.001977	0.003932	0.6153	
Blätterdachhöhe	-0.00585	0.006504	0.3687	
Anzahl Attalea	-0.184705	0.02637	6.27E-12	***
Anzahl Piper	0.015511	0.012589	0.2183	

Model 5 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.143)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Sex	-0.007431	0.036863	0.840303	
Alter	-0.154512	0.066081	0.019685	*
Spezies Marmosa robinsoni isthmica	-0.271414	0.10852	0.012633	*
Spezies <i>P.sem</i>	0.187112	0.062227	0.002743	**
Habitattyp C	-0.083589	0.042928	0.051952	
Habitattyp I	0.219615	0.056822	0.000123	***

### 3.1.4. Deltavirus Analysen

Um ein Verständnis über das Vorkommen viraler Erreger in *P.sem* zu generieren, wurden 120 Blutproben mittels Hochdurchsatzsequenzierung (Illumina MiSeq) untersucht. Die Probeauswahl hierfür erfolgte für eine Hepacivirussequenzierung und wird im Kapitel 3.1.5 genauer erläutert. Der Vergleich von de novo-assemblierten Sequenzen mit einer viralen Referenzbibliothek offenbarte eine hohe Übereinstimmung zu humanem HDV. Durch iteratives Alignieren der Rohdaten gegen Referenzsequenzen aller bekannten HDV-Genotypen, konnten HDV-ähnliche Sequenzen zusammengestellt werden. Insgesamt konnten so HDV-Vollgenome von zwei *P.sem* generiert werden. Nach diesen anfänglichen Genom-Assemblies mit HDV-Referenzgenotypen, wurde ein RT-PCR-Assay entworfen, der ein Fragment in der kodierenden Region des Nager-Deltavirus-Antigens (rodent delta antigen; RDeAg) amplifiziert. Dieser Assay wurde auf gepoolte Blutproben von 763 *P. sem* Individuen angewendet. Die Auflösung der positiv getesteten Pools führte schließlich zur Identifikation von 30 positiven Individuen. Neben *P.sem* wurden Blutproben von Individuen (n = 220) aus 11 anderen Nagetier- und Beuteltierarten untersucht. In keiner anderen Spezies, außer *P. semispinosus*, wurden Hinweise für die Präsenz von Nager-Deltavirus (rodent deltavirus; RDeV) gefunden (Tabelle 31).

**Tabelle 31: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von Deltavirus-RNA untersucht wurden**

Gezeigt sind die Ergebnisse der Suche nach Deltaviren in Blutproben. Die Ergebnisse umfassen Funde aus der Screening PCR sowie aus NGS Läufen.

Spezies	negativer Picornavirus Nachweis	positiver Picornavirus Nachweis	Speziessumme
<b>Liomys adspersus</b>	1	0	<b>1</b>
<b>Oecomys bicolor trinitatis</b>	1	0	<b>1</b>
<b>Transandinomys talamanceae</b>	1	0	<b>1</b>
<b>Sigmodon hirsutus</b>	2	0	<b>2</b>
<b>Heteromys desmarestianus</b>	3	0	<b>3</b>
<b>Zygodontomys brevicauda</b>	3	0	<b>3</b>
<b>Metachirus nudicaudatus</b>	4	0	<b>4</b>
<b>Marmosa robinsoni isthmica</b>	21	0	<b>21</b>
<b>Hoplomys gymnurus</b>	22	0	<b>22</b>
<b>Philander opossum</b>	57	0	<b>57</b>
<b>Didelphis marsupialis</b>	105	0	<b>105</b>
<b>Proechimys semispinosus</b>	733	30	<b>763</b>
<b>Gesamtergebnis</b>	953	30	<b>983</b>

Da HDV ein zirkuläres Genom besitzt und die NGS Daten ebenfalls auf ein zirkuläres Genom hinwiesen, wurden mehrere Primerpaare entwickelt (Abbildung 14), um zu bestätigen, dass es dies bei dem gefundenen RDeV vorliegt. Die Amplikons, die bei der RT-PCR erzeugt wurden, decken nahezu das vollständige Deltavirus-Genom ab, wenn das Template-RNA-Molekül tatsächlich zirkulär ist. Dies gelang in drei Proben mit hoher Viruslast. Die 30 positiv getesteten Proben wurden per NGS sequenziert und aus den erhaltenen Reads konnten anschließend 10 vollständige, zirkuläre RDeV-Genome und 10 partielle Genome assembliert werden. Die Nukleotidsequenzenidentität der zehn vollständigen Genome lag dabei zwischen 97,5 und 99,6%.

Betrachtet man die habitatspezifische Prävalenz, so fällt auf, dass ein erheblicher Teil der RDeV Funde in Wald Habitaten zu finden ist (Tabelle 33). In einem GLM wird geprüft ob der Habitattyp einen signifikanten Einfluss auf die Wahrscheinlichkeit hat, mit RDeV infiziert zu werden (Tabelle 32). Da ein GLM nicht mit fehlenden Werten umgehen kann, wurde der Datensatz so reduziert, dass keine fehlenden Werte mehr vorhanden waren. Von den ursprünglich 983 Proben standen für diese Analyse nach der Reduktion noch 887 Proben zur Verfügung.

**Tabelle 32: Logistische Regression zur Untersuchung des Effekts des Habitattyps auf das Virusaufkommen von RDeV anhand eines GLM**

Gezeigt sind die Ergebnisse des GLM der detektierten ReDV Infektionen in Abhängigkeit zu den nominalen Variablen des Habitattyps. Als Vergleichsregressor diente in diesem Fall Habitattyp A. Als Güte wurde McFadden's Pseudo-R<sup>2</sup> berechnet. Für die Regression wurden jeweils 887 Proben benutzt. Signifikanzbeschriftung wurde wie folgt festgesetzt: \*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05. Das GLM wurde in R gerechnet.

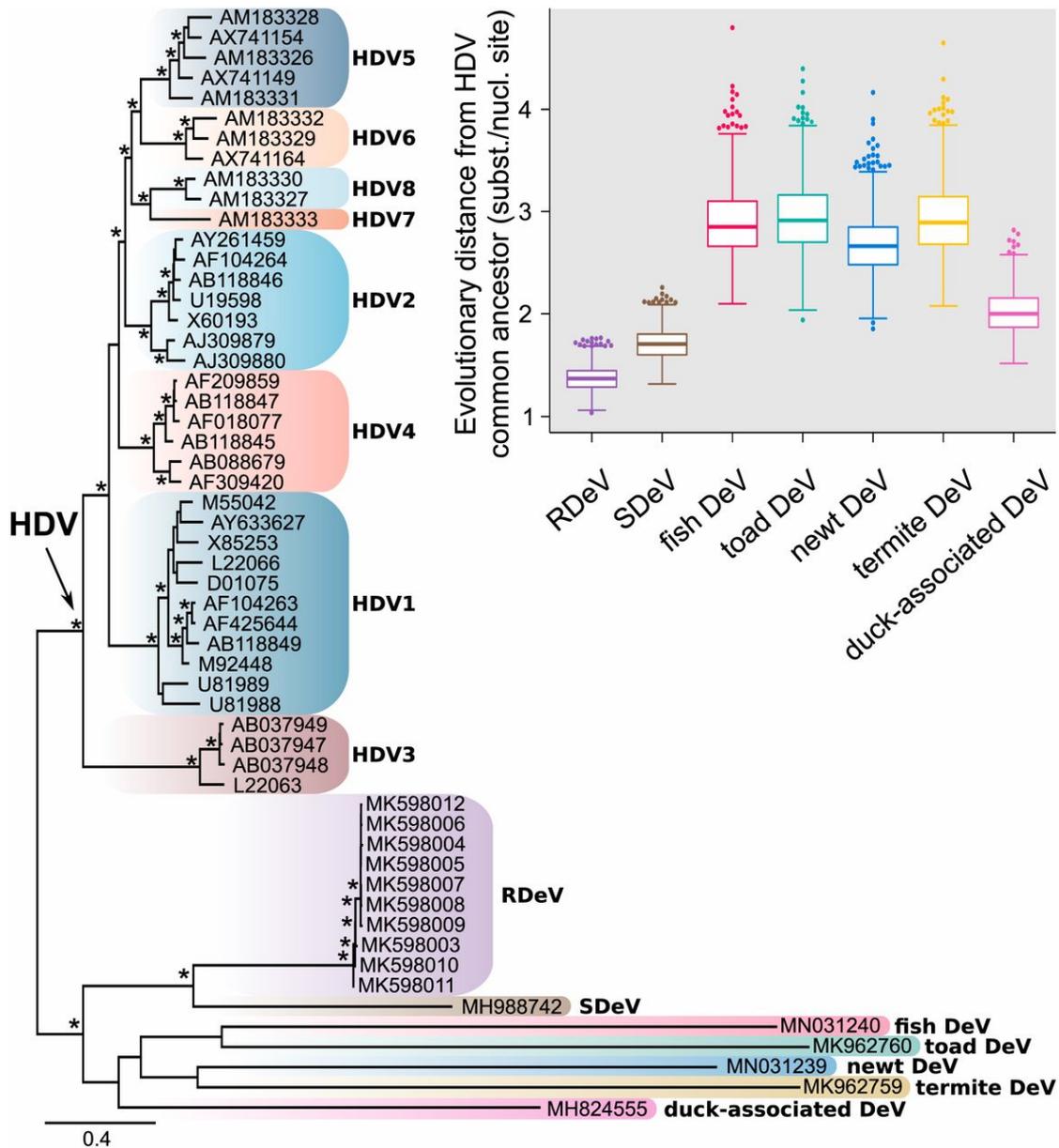
Habitattyp	Estimate	Std. Error	P-Value	Signifikanz
C	0.064226	0.014404	9.74E-06	***
I	-0.003774	0.018027	0.834	
Güte nach McFadden's Pseudo-R <sup>2</sup> : 0.037				

Das GLM errechnete einen eindeutigen Effekt für den Habitattyp Wald und einen vernachlässigbaren Effekt für den Habitattyp Insel, da dieser das Signifikanzniveau von p<0,05 überschritt.

**Tabelle 33: Deltavirusprävalenz Aufschlüsselung pro Habitattyp**

Gezeigt sind die Ergebnisse aus der Kombination von gerichtetem Virusscreening (PCR) und ungerichteter NGS Analyse. Hier sind die Anzahl der Funde pro Habitattyp aufgelistet, sowie der prozentuale Anteil an positiven Funden pro Habitattyp. Die Habitattypen sind Agrikulturgebiete (A), Wälder (C), Inseln (I) und Plantagen (P)

Habitat	negativer Deltavirus Nachweis	positiver Deltavirus Nachweis	Summe pro Habitattyp	Deltavirusprävalenz (%)
AXAR1	56	0	56	0
AXAR2	64	1	65	1,54
AXAR4	83	0	83	0
AXAS1	9	0	9	0
AXAS2	9	0	9	0
AXAS3	5	0	5	0
AXAS4	15	0	15	0
AXAS5	50	0	50	0
AXLP1	9	1	10	10,00
CGIG1	22	3	25	12,00
CGIG2	16	4	20	20,00
CXBCI	24	2	26	15,38
CXXBO	42	0	42	0
CXXFR	106	1	107	0,93
CXXPB	28	7	35	20,00
IXXGU	120	0	120	0
IXXMG	30	0	30	0
IXXOI	23	1	24	4,17
IXXPH	108	9	117	7,69
IXXTA	68	0	68	0
PXTP1	16	0	16	0
PXTP2	16	1	17	5,88
PXTP3	13	0	13	0
PXTP5	9	0	9	0
PXTP8	12	0	12	0
<b>Gesamtergebnis</b>	953	30	983	



**Abbildung 22: Phylogenetischer Baum der Deltaviren mit Vergleich der Nt-Distanzen**

ML Baum aus einem Alignment von Repräsentanten der HDV Vollgenomen aus allen acht Genogruppen, sowie RDeV, SDeV, Enten-assoziiertes DeV, Molch DeV, Kröten DeV, Fisch DeV und Termiten DeV. Support >80% ist per Sternchen markiert. Baumspitzen geben die GenBank Accession numbers an. Die Wurzel des Baumes wurde auf den Ast der nicht-menschlichen Sequenzen gelegt. Die Verteilung der patristischen Nt-Distanzen im Boxplot (1000 Bootstraps) aller nicht-menschlichen Sequenzen wurde mit dem gemeinsamen Vorfahren aller HDV Sequenzen (hier markiert durch einen Pfeil) errechnet. Der Boxplot zeigt Median, Interquartilsabstand und Minimum/Maximum Whiskers nach Tukey. Modifiziert aus Paraskevopoulou *et al.* [100]

Diese Sequenzinformationen wurden phylogenetisch ausgewertet und, zusammen mit anderen nicht-humanen Deltaviren, mit HDV verglichen (Abbildung 22).

Wichtige strukturelle und funktionelle HDV-Domänen sind in allen wiedergewonnenen RDeV-Sequenzen vorhanden [100], einschließlich eines einzelnen offenen Leserahmens (ORF), der das Delta-Antigen kodiert. Ein 19-Aminosäuren-Region, die das S-HDAg vom L-HDAg in humanem HDV unterscheidet, liegt auch in RDeV vor,

allerdings mit einer anderen Aminosäurezusammensetzung. Da bei HDV ein Organotropismus der Leber vorliegt wurden als nächstes die Infektionsmuster des Nager-Deltavirus in den Organen infizierter Tiere untersucht. Dazu wurden Organen von 18 Tieren, die während der Probenahme tot aufgefunden wurden, auf virale RNA getestet. Hierbei ist zu erwähnen, dass aufgrund des nicht-invasiven Probenahmeverfahrens Tiere nicht eingeschläfert wurden. Eines dieser Tiere wurde RDeV RNA-positiv in allen verfügbaren Organen (Leber, Niere, Lunge, Herz, und Dünndarm) getestet. Die organspezifischen RNA-Konzentrationen wiesen jedoch nicht auf eine spezifische Virusreplikation in der Leber hin (Tabelle 34). Für diese quantitative Bestimmung ist es jedoch nötig, ein positives Material mit bekannter Konzentration zu besitzen. Aus diesem Grund wurde ein in-vitro Transkript hergestellt und verwendet wie unter Kapitel 2.4.5.1 beschrieben.

<b>Tabelle 34: Anzahl der RDeV-RNA Kopien in Organen von <i>P.sem</i></b>	
Kalkulation erfolgte über die Steigungsfunktion der Verdünnungsreihe eines in-vitro Transkripts mit bekannter Konzentration	
<b>Organ</b>	<b>RNA Kopien von RDeV pro [g] Gewebe</b>
Herz	1.045.688
Niere	1.855.127
Leber	955.878
Lunge	311.926
Dünndarm	1.589.207

Um die potenzielle Virusausscheidung zu überprüfen, wurden Kotproben von 820 Individuen (alle verfügbaren Kotproben aus *P.sem*) per RT-PCR getestet. Zehn von diesen Proben waren positiv für RDeV-RNA. Über eine qRT-PCR wurde anschließend die Konzentrationen an Virus in den verschiedenen Proben bestimmt.

Die Konzentration in Stuhlproben betrug  $2,9 \times 10^9$  RNA-Kopien pro Gramm, mit einer maximalen Konzentration von  $2,2 \times 10^{10}$  RNA-Kopien pro Gramm aus 10 Proben. Durchschnittliche und maximale Konzentrationen im Blut betrugen  $1,8 \times 10^8$  und  $2,1 \times 10^9$  Kopien pro mL aus 30 Proben.

Abschließend wurden GLM genutzt um den Einfluss der ökologischen Faktoren auf die Virusverteilung zu zeigen (Tabelle 35). Dabei wurden die fünf Modelle benutzt wie in Kapitel 3.1.1 beschrieben. Die 983 Proben wurden bereinigt, so dass keine fehlenden Metadaten mehr vorhanden waren. Dies reduzierte den Datensatz auf 634 Proben.

**Tabelle 35: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von RDeV**

Gezeigt sind die Ergebnisse des, in R berechneten, GLM der detektierten RDeV Infektionen in Abhängigkeit zu ökologischen Faktoren. Als Vergleichsregressor für kategoriale Variablen diente der erste Wert der Variable. Als Güte wurde McFadden's Pseudo-R<sup>2</sup> berechnet. Für die Regression wurden jeweils 634 Proben benutzt. In Model 5 werden, ausser *P.sem*, nur signifikante Spezies gezeigt. Signifikanzbeschriftung wurde wie folgt festgesetzt:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

Model 1 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.095)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Habitatgröße	6.63E-05	9.19E-06	1.52E-12	***
Biodiversität	4.71E-03	9.71E-03	0.62798	
Blätterdachabdeckung	-3.43E-04	1.24E-03	0.7825	
Anzahl Attalea	-3.00E-02	9.16E-03	0.00113	**

Model 2 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.047)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Unterholzdichte	-0.005115	0.001032	9.26E-07	***
Anzahl Astrocaryu	0.016595	0.005851	0.00471	**
Anzahl Piper	-0.010154	0.004845	0.03651	*

Model 3 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.022)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Anzahl Astrocaryu	0.0144859	0.0059206	1.47E-02	*
Anzahl Piper	-0.0084529	0.0049378	0.08741	
Populationsdichte	-0.0016309	0.0005949	0.00629	**

Model 4 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.040)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Blätterdachabdeckung	-0.000811	0.001372	0.554559	
Blätterdachhöhe	0.008602	0.002194	9.77E-05	***
Anzahl Attalea	-0.033854	0.009449	0.000366	***
Anzahl Piper	-0.005861	0.004683	0.211175	

Model 5 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.049)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Sex	0.006326	0.013273	0.633811	
Alter	-0.031213	0.018784	0.097089	
Spezies <i>P.sem</i>	0.026001	0.021116	0.21867	
Habitattyp C	0.057673	0.015904	0.000311	***
Habitattyp I	-0.017718	0.019908	0.373824	

### 3.1.5. Hepacivirus Analysen

Die Etablierung einer Hepacivirus Screening PCR verlief analog zu den anderen spezifischen *P.sem* Viren. Durch einen NGS Lauf wurden Hepacivirus-Reads detektiert, die im Anschluss zu einem Vollgenom assembliert wurden. Auf Basis dieses Vollgenomes konnten nun Primer für eine Screening PCR designt werden [88, 99]. In einer Vortestung fiel auf, dass die Hepacivirusprävalenz in Spezies außerhalb von *P.sem* (4/60 positive PCR Detektion; 6,67%), im Vergleich zu *P.sem* (95/110 positive PCR Detektion, 86,3%) geringer ausfiel. Da wenige Einzeldetektionen in Randspezies kaum einen Mehrwert für Analysezwecke brachen, wurde entschieden, gepoolte Blutproben, die kein *P.sem* enthielten nicht mit der Screening PCR zu testen.

**Tabelle 36: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von PBV-RNA untersucht wurden**  
Gezeigt sind die Ergebnisse der Suche nach PBV in allen Kotproben.

Spezies	negativer Hepacivirus Nachweis	positiver Hepacivirus Nachweis	Ungetestete Proben	Speziessumme
<i>Liomys adspersus</i>	0	1	0	1
<i>Oecomys bicolor trinitatis</i>	1		0	1
<i>Transandinomys talamancae</i>	1	0	0	1
<i>Sigmodon hirsutus</i>	2	0	0	2
<i>Heteromys desmarestianus</i>	0	0	3	3
<i>Zygodontomys brevicauda</i>	3	0		3
<i>Metachirus nudicaudatus</i>	3	0	1	4
<i>Marmosa robinsoni isthmica</i>	14	0	7	21
<i>Hoplomys gymnurus</i>	11	0	11	22
<i>Philander opossum</i>	6	17	34	57
<i>Didelphis marsupialis</i>	26	0	79	105
<i>Proechimys semispinosus</i>	169	594	0	763
<b>Gesamtergebnis</b>	<b>236</b>	<b>612</b>	<b>135</b>	<b>983</b>

Mit dieser Screening Strategie wurden aus den 763 extrahierten und gepoolten *P.sem* Blutproben insgesamt 594 Proben in der Screening PCR positive getestet (Tabelle 36). Aufgrund von der geringen Konzentration an Hepaciviren in manchen Proben (erkennbar durch schwache Banden im Agarosegel der Screening PCR) konnte nur von 349 dieser Proben ein Screening-Fragmente (~1 kbp Länge) sequenziert werden. Durch die hohen

Kosten einer NGS Laufes und der Tatsache, dass für nachfolgende Analysen die begrenzten Mengen an Probematerial geschont werden mussten, wurde entschieden nur einen Bruchteil der positiv getesteten *P.sem* Blutproben (Tabelle 37) per Hochdurchsatzsequenzierung zu sequenzieren. Dafür wurde, soweit möglich, aus sämtlichen Habitaten maximal 10 positive Proben ausgewählt. Die Proben wurden hierbei nach der Stärke, der Bande im Agarosegel der Screening PCR ausgewählt. Waren mehrere Kandidaten vorhanden wurde versucht ein möglichst breites Spektrum an Wirten abzudecken (Männlich/Weiblich, Jugendliche/Adult sowie das Vorhandensein von Kotproben aus demselben Tier).

**Tabelle 37: Übersicht der Anzahl an Proben aus *P.sem* von Hepaciviren über alle Habitats und Datensets**

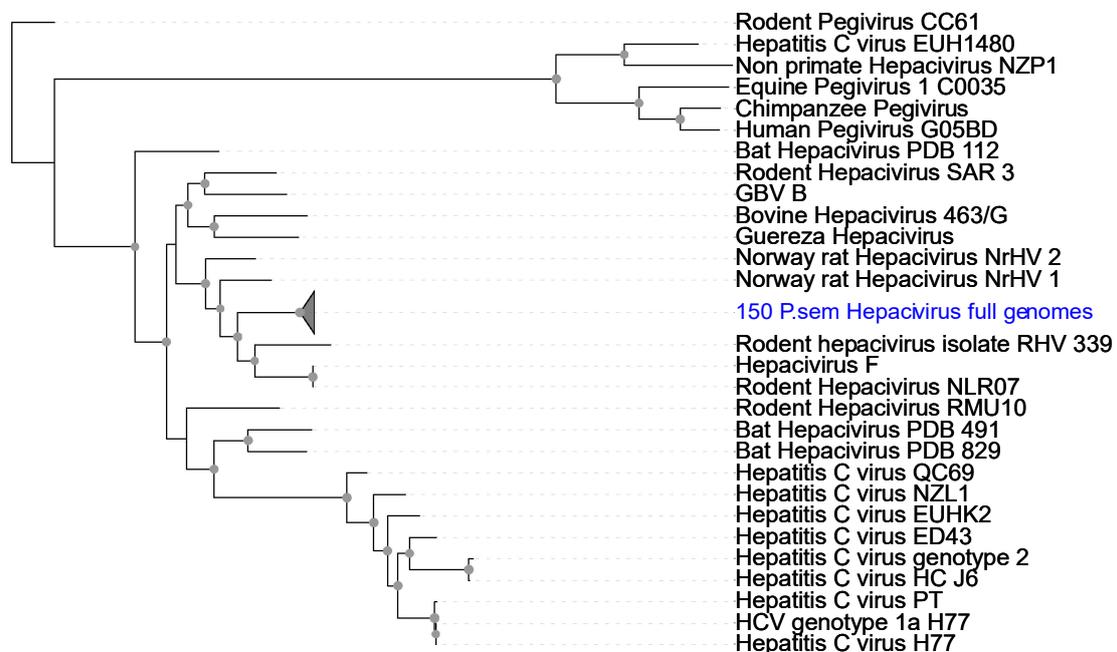
Gezeigt ist hier die Anzahl an Proben die in der Hepaci-Screening PCR positiv detektiert wurde, die Anzahl an sequenzierten Proben sowie die Anzahl an Vollgenomen unterteilt in jeweils alle Habitats, sofern diese einen Fund beinhalteten.

Habitat	Detektionen per Hepacivirus Screening PCR	Anzahl an Sequenzfragmenten	Anzahl an Vollgenomen
AXAR1	32	25	10
AXAR2	9	5	5
AXAR4	27	27	9
AXAS4	3	3	3
AXAS5	9	8	8
AXLP1	9	0	0
CGIG1	20	17	10
CGIG2	15	13	10
CXBCI	16	14	10
CXXBO	26	24	10
CXXFR	84	31	10
CXXPB	24	24	10
IXXGU	93	41	10
IXXMG	26	22	10
IXXOI	21	18	10
IXXPH	108	40	10
IXXTA	61	26	10
PXTP1	2	1	1
PXTP2	0	0	0
PXTP3	2	0	0
PXTP5	3	0	0
PXTP8	4	4	4
<b>Gesamtergebnis</b>	<b>594</b>	<b>343</b>	<b>150</b>

Diese wurden mit Hilfe von neu designten, spezifischen Genom-Primern und NGS sequenziert. Die daraus resultierten Vollgenome (n=150) besaßen eine Länge von 8289 bp. Aus den gesammelten Sequenzdaten wurde per MAFFT Plugin in Geneious

ein Alignment erstellt. Das Alignment aus 343 Sequenzen, bestand dabei aus 193 Screening Fragmenten (~1000 bp), sowie 150 Vollgenomen (8289 bp). Aufgrund der geringen Anzahl an Sequenzen aus dem Habitattyp „Plantage“ (PXTP1 und PXTP8), sowie der geographischen Nähe, wurden dieser mit dem Habitattyp „Agrikultur“ in weiteren Analysen gruppiert und zusammen betrachtet.

Zur Genuseinordnung wurde ein Alignment erstellt, dass neben den 150 Hepacivirusvollgenomen zusätzlich 28 Referenzsequenzen von Hepaci- und Pegiviren enthält (Abbildung 23). Dazu wurde von ICTV ein Alignment an *Flaviviridae* Referenzsequenzen heruntergeladen und die NS5b Region extrahiert (dies ist die RdRp auf der die Screening Primer designt wurden).

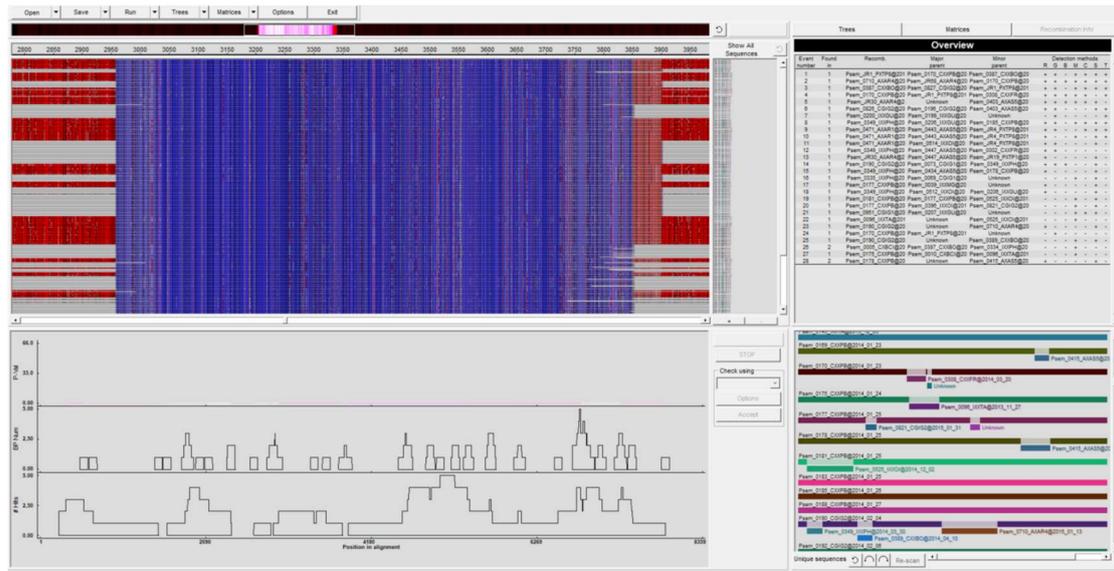


**Abbildung 23: Phylogenetischer Analyse der NS5b Region von Hepaci- und Pegiviren**

Der ML Baum wurde in Geneious mit 1000 Bootstraps berechnet und beinhaltet die NS5b Region der 150 *P.sem* Hepacivirusvollgenome (blaue Klade; kollabiert), sowie 28 Hepaci- und Pegivirusreferenzsequenzen. Graue Kreise markieren alle Knoten mit einem Support >75%. Die Wurzel wurde auf „Rodent Pegivirus CC61“ gelegt.

Durch die große Menge an Sequenzdaten für *P.sem* Hepaciviren konnten tieferegehende Analysen durchgeführt werden, die anhand der Datensätze der anderen Viren nicht möglich waren. In explorative Analysen von ML Bäumen konnte gezeigt werden, dass trotz einer großen Menge an Sequenzinformationen, der Support an vielen Knoten in reinen *P.sem* Hepaciviren Bäumen sehr niedrig war (<60%). Dieses Verhalten lässt sich z.B. durch eine hohe Grundvariabilität oder Rekombination erklären. Zur Prüfung auf

Rekombination wurde daraufhin das Alignment, aus 343 Sequenzen, per RDP4 (Abbildung 24) einer Rekombinationsanalyse unterzogen. Unter Einbeziehung sämtlicher Methoden im „Full Exploration“ Modus wurden mehrere Regionen in verschiedenen Sequenzen identifiziert, die mit einer hohen Wahrscheinlichkeit (identifiziert durch mindestens 3 der 7 Algorithmen) aus einem Rekombinationsereignis hervorgegangen sind.



**Abbildung 24: Rekombinationsanalyse von Hepaciviren aus *P.sem***

Zu sehen sind die vier Hauptfenster des Rekombinationsanalyseprogramms RDP4. Oben links ist die Alignmentübersicht mit Sequenzdichte (wieviel Sequenzinformationen pro Position vorhanden sind), unten links ist die Übersicht an Funden von Rekombinationsereignissen pro Position und Algorithmus. Oben rechts ist die Rangliste der gefundenen Rekombinationsereignissen inkl. Stammsequenz, Elternsequenz (1. und 2. Grades) und Anzahl an Treffern pro Algorithmus. Unten rechts sieht man die grafische Übersicht an möglichen Rekombinationsereignissen inkl. der Lage auf dem Genom und hypothetischer Elternsequenz.

Durch mehrere Iterationsrunden im „Full Exploration“ Modus, wurden diese Sequenzen entfernt und das Alignment auf 245 Sequenzen reduziert. Da jedoch noch immer einzelne Rekombinationsereignisse gefunden wurden (identifiziert durch 1-2 Algorithmen) und der Support in ML Bäumen an manchen Knoten immer noch niedrig war, wurde das Alignment in mehrere Fragmente unterteilt. Da dies ist ein Indiz dafür, dass Teilabschnitt im Baum weiterhin nicht eindeutig zugeordnet werden können, was für verbleibende, rekombinierte Regionen spricht. Durch systematische Aufteilung des Alignments und weitere Iterationen in RDP4 konnte in 231 Sequenzen eine Region identifiziert werden, die ML Bäume mit hohen Supportwerten erzeugte. Von diesen 231 Sequenzen stammen 69 aus den Vollgenomen und die restlichen 162 aus den 1kb Screening Fragmenten. Dabei ist die längste Sequenz 3511 bp lang. Die Verteilung über die Habitate in diesem Alignment ist in Tabelle 38 zu sehen.

**Tabelle 38: Verteilung der Proben innerhalb des reduzierten *P.sem* Hepacivirus Alignments aus 231 Sequenzen über die Habitate**

Gezeigt ist sind hierbei die Anzahl an Proben die sich in dem Alignment befinden, sowie die Anzahl an Vollgenomen und Sequenzfragmenten

Habitat	Anzahl an Screeningfragmente	Anzahl an Vollgenomen
AXAR1	11	1
AXAR4	15	1
CGIG1	5	6
CGIG2	3	3
CXBCI	1	5
CXXBO	11	4
CXXFR	18	9
CXXPB	9	3
IXXGU	30	9
IXXMG	12	10
IXXOI	8	5
IXXPH	26	6
IXXTA	13	7
<b>Gesamtergebnis</b>	<b>162</b>	<b>69</b>

Für eine tiefgehende bayesianische Analyse mussten sichergestellt werden, dass die temporalen Daten, die aus den verwendeten Sequenzdaten und den Probenahmezeitpunkten entstehen, in sich konsistent und keine Widerspruch in sich aufweisen. Hierfür eignet sich eine Analyse des temporalen Signals mit dem Programm TempEST [94]. Dazu wird ein ML-Baum (berechnet in Geneious mit dem PHYML Plugin, 100 Bootstraps und GTR Substitutionsmatrix), des Alignments aus 231 Sequenzen (mit maximaler Länge von 3511 bp), in das Programm geladen und eine Abschätzung der Astlängen im Verhältnis zu deren Probenahmedaten erstellt. In diesem Fall, konnte keine Sequenz gefunden werden bei der die Astlänge in Relation zu anderen Sequenzen heraussticht.

Nun musste für diesen Datensatz die geeignetsten Substitutions-, Populations- und Clock-Modelle bestimmt werden.

Für das Substitutionsmodell wurde das Programm jModeltest (Abbildung 25) benutzt.

AICc Selection Results				BIC Selection Results			
Model selected				Model selected			
Model	GTR+I+G			Model	GTR+I+G		
partition	012345			partition	012345		
-lnL	27957.2350			-lnL	27957.2350		
K	414			K	414		
freqA	0.2559	R(a)	1.4300	freqA	0.2559	R(a)	1.4300
freqC	0.2665	R(b)	12.2241	freqC	0.2665	R(b)	12.2241
freqG	0.2344	R(c)	2.8264	freqG	0.2344	R(c)	2.8264
freqT	0.2431	R(d)	0.3066	freqT	0.2431	R(d)	0.3066
ti/tv	-	R(e)	14.4006	ti/tv	-	R(e)	14.4006
		R(f)	1.0000			R(f)	1.0000
p-inv	0.6710	gamma	0.7440	p-inv	0.6710	gamma	0.7440

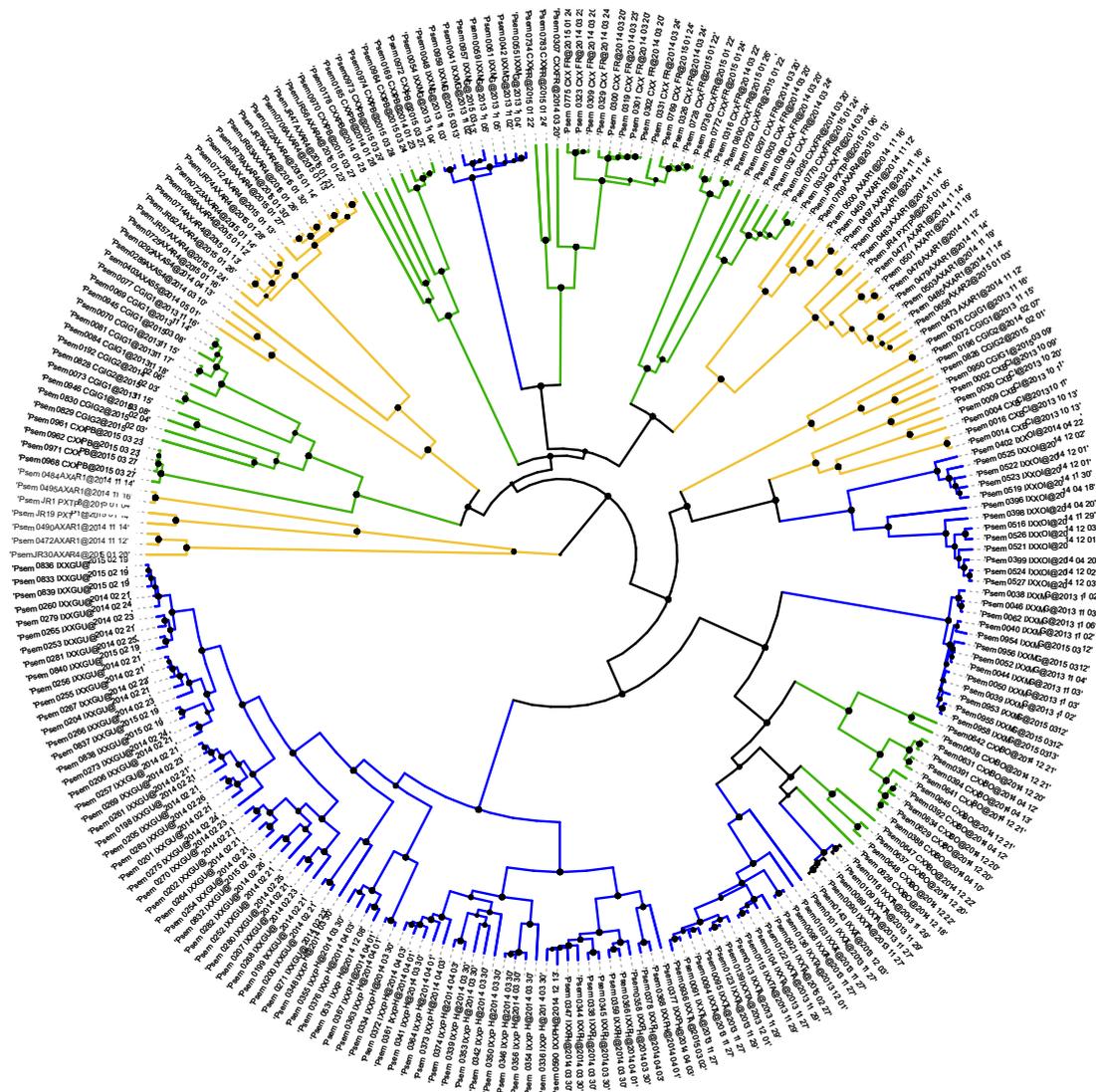
**Abbildung 25: Bestimmung des Substitutionsmodells mit jModeltest**

Gezeigt sind die beide höchsten Treffer der Ranglisten für Akaike Information Criterion corrected (AICc) und Bayesian Information Criterion (BIC). Die Berechnung erfolgte ML optimiert, mit 11 Substitutionsmodellen und unter Einbeziehung aller Ratenvariationen (+I, +G) sowie unterschiedlicher ungleicher Basenverteilung (+F). Beiden Ranglisten bestimmten GTR+I+G (general time reversible) als passendstes Substitutionsmodell.

Die Analyse zeigte, dass das GTR+I+G Substitutionsmodell den besten Fit für die vorliegenden Sequenzdaten hatte.

Zur Analyse ob sich die Isolation der Inseln auf die Evolutionsrate der Hepaciviren ausgewirkt hat, wurde initial eine relaxed Clock gewählt. Dies ermöglicht die unterschiedlichen Evolutionsraten an verschiedenen Ästen zu betrachten und vergleichen.

Für die Berechnung der Populationsgröße wurde ein bayesianisches Skygrid-Model benutzt, welches in gleichmäßigen Abständen eine mögliche Änderung der Population abschätzt. Da durch vorangegangene, explorative Läufe ein ungefährender Bereich für das Wurzelalter von ca. 50-150 Jahren abgeschätzt werden konnte, wurde das Skygrid entsprechend definiert. Im Fall von *P.sem* Hepaciviren, wurden die Werte so gewählt, dass das Modell einen Zeitraum von 180 Jahren betrachtet und diesen in 60 Teilabschnitte unterteilt. Anhand dieser Parameter wird die Populationsgröße, innerhalb des Baumes, alle 3 Jahre abgeschätzt.

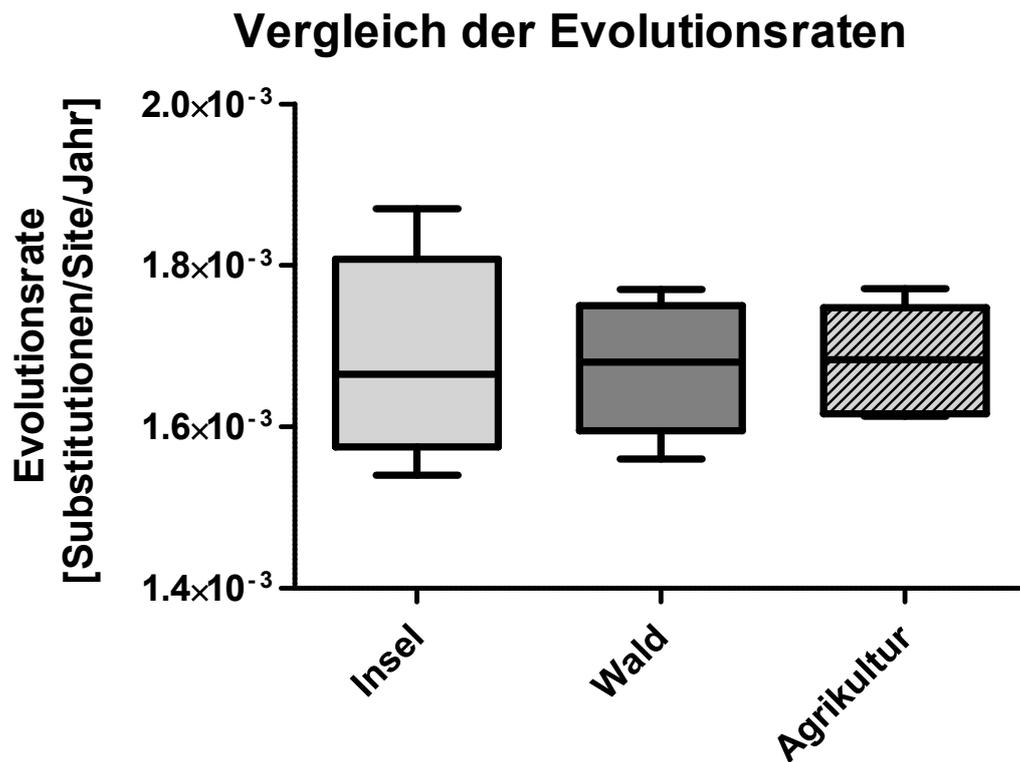


**Abbildung 26: Phylogenetischer bayesianischer Baum der *P.sem* Hepaciviren in radialer Ansicht**  
Gezeigt ist ein kombinierter Baum aus fünf BEAST Läufen, des Alignments aus 231 *P.sem* Hepacivirussequenzen, mit einer Kettenlänge von jeweils 150 Millionen (Mio.), gerechnet mit einem relaxed Clock-Modell, Skygrid Populationsmodell und GTR+I+G Substitutionsmodell. Gefärbte Äste zeigen Sequenzen aus den jeweiligen Habitats Agrikultur/Plantage (gelb), Wälder (grün) und Inseln (blau) an. Schwarze Punkte markieren Knoten mit einem Support >75%.

Mit den gewählten Modellen wurde mit der bayesianische Analyse, mit BEAST, begonnen und ein phylogenetischer Baum berechnet (Abbildung 26). Bei dem gezeigten Baum handelte es sich um eine Kombination aus fünf verschiedenen Läufen mit einer Kettenlänge (MCMC) von jeweils 150 Mio. Alle Läufe wurden vor der Kombination, mit dem Programm „Logcombiner“ auf vollständige Konvergenz (ESS aller Parameter >200) geprüft.

In diesem phylogenetischen Baum (Abbildung 26) zeigte sich, dass die unterschiedlichen Habitat zu großen Teilen monophyletisch zusammen liegen. Dabei bilden besonders Habitate vom Typ Insel (im Baum blau markiert) große Cluster.

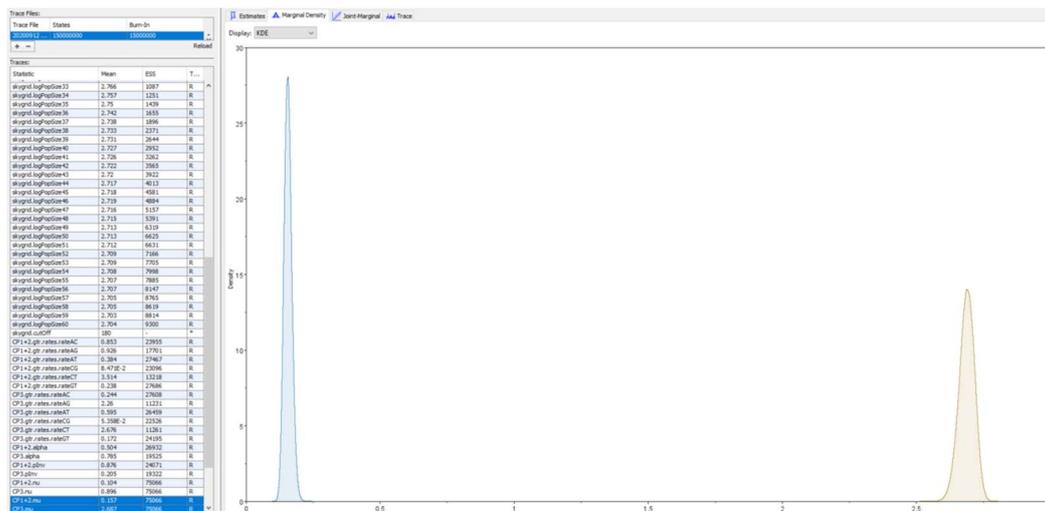




**Abbildung 28: Vergleich der Evolutionsraten von *P.sem* Hepaciviren**

Boxplot, mit Whiskers nach Tukey, erstellt in GraphPad Prism. Es handelt sich um 23 Evolutionsraten, aus dem bayesianischen Baum (Abbildung 27) aus fünf kombinierten BEAST Analysen. Evolutionsraten sind dabei nach Habitattypen zusammengefasst. Es wurde ein Kruskal-Wallis Test mit Dunns Post-Hoc Test durchgeführt ( $p = 0,8852$ ), der keinen signifikanten Unterschied zwischen den drei Gruppen zeigen konnte.

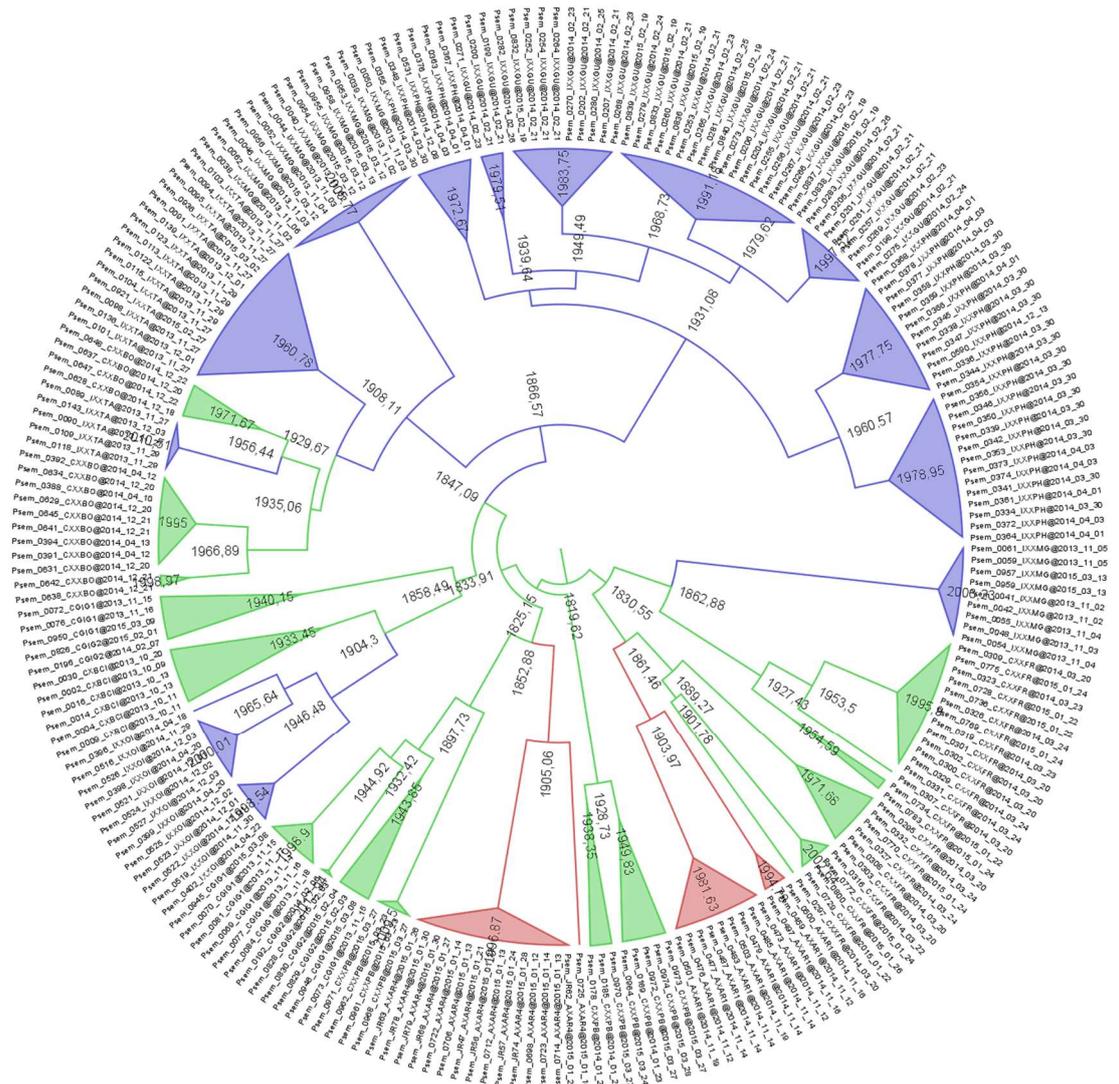
Es wurde außerdem festgestellt, dass eine Mutation in einem hohen Maß häufiger auf der dritten Codonposition stattfindet (Abbildung 29). Durch den Vergleich der Marginalverteilungen der ersten beiden Codonpositionen (0,157) und der dritten Codonposition (2,687) errechnet sich eine 17fach höhere Mutationsrate auf der letzten Position des Codons.



**Abbildung 29: Traceranalyse der Mutationsrate zwischen den Codonpositionen**

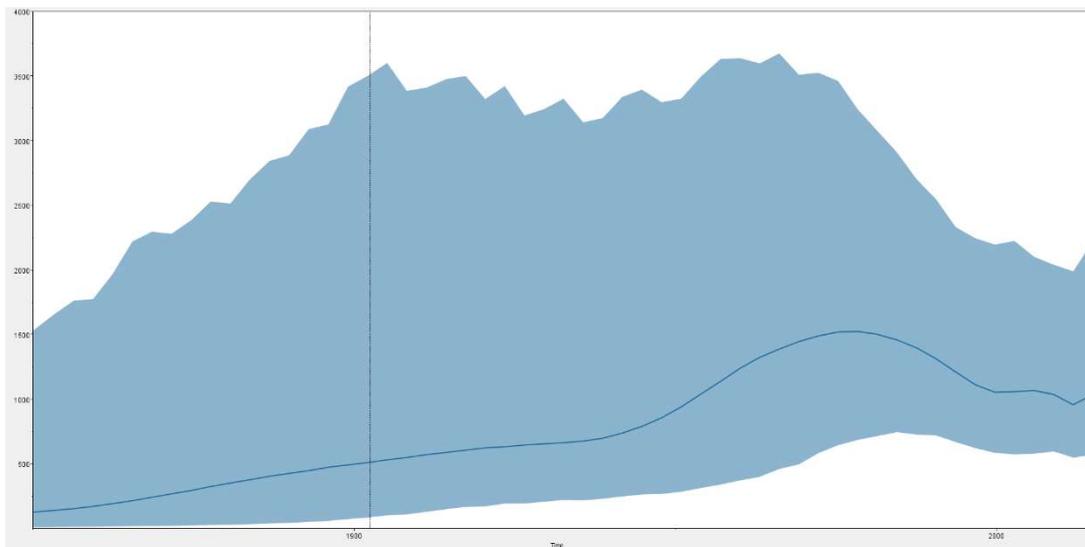
Gezeigt ist die Marginalverteilung der Mutation zwischen den ersten beiden Codonpositionen (blau; 0,157) und der dritten Codonposition (orange; 2,687) in Tracer. Die zugrundeliegenden Daten stammen dabei aus fünf kombinierten BEAST Läufen mit einer Kettenlänge von 150 Mio.

Um eine Verfälschung der Resultate durch Überparametrisierung zu vermeiden, wurden die BEAST Läufe mit einer Strict-Clock wiederholt. Da durch ein strict Clock-Model die Markovkette schneller konvergiert wurde die Kettenlänge auf 50 Mio. reduziert. Zusätzlich wurde auf die Habitate ein Verfahren der „Bayes'schen stochastischen Suche zur Variablenauswahl“ (BSSVS) angewandt. Mit diesem Verfahren wird versucht, die Anzahl der Transitionsraten (mindestens  $k-1$ , wobei  $k$  die Anzahl der Zustände ist) auf diejenigen zu beschränken, die den phylogenetischen Diffusionsprozess angemessen erklären. Es wurden fünf Wiederholungsläufe mit jeweils einer Kettenlänge von 50 Mio. durchgeführt. Analog zur vorherigen Analyse, des relaxed Clock Laufes, wurden die Log Dateien mit „Logcombiner“ zusammengeführt. Durch die veränderte Modelauswahl wurde die Mutationsrate niedriger abgeschätzt als im vorherigen Lauf (vergleich Abbildung 27 und Abbildung 30). Sie lag nach dieser Modelauswahl bei  $6,04 \times 10^{-4}$  Mutationen/Basenpaar/Jahr. Mit dieser Mutationsrate datiert BEAST das Wurzelalter des Baums, oder das Alter des jüngsten, gemeinsamen Vorfahrens (MRCA) auf ca. 1819. Die Auswertung des Verlaufs der Populationsgröße durch das Skygrid Model (Abbildung 31).



**Abbildung 30: Phylogenetischer bayesianischer Baum der *P.sem* Hepaciviren in radialer Ansicht des BEAST Laufs mit einem strict Clock Model**

Gezeigt ist ein kombinierter Baum aus fünf BEAST Läufen, des Alignments aus 231 *P.sem* Hepacivirussequenzen, mit einer Kettenlänge von jeweils 50 Millionen (Mio.), gerechnet mit einem strict Clock Model, Skygrid Populationsmodell und GTR+I+G Substitutionsmodell. Gefärbte Äste zeigen Sequenzen aus den jeweiligen Habitaten Agrikultur/Plantage (gelb), Wälder (grün) und Inseln (blau) an. Schwarze Punkte markieren Knoten mit einem Support >75%.

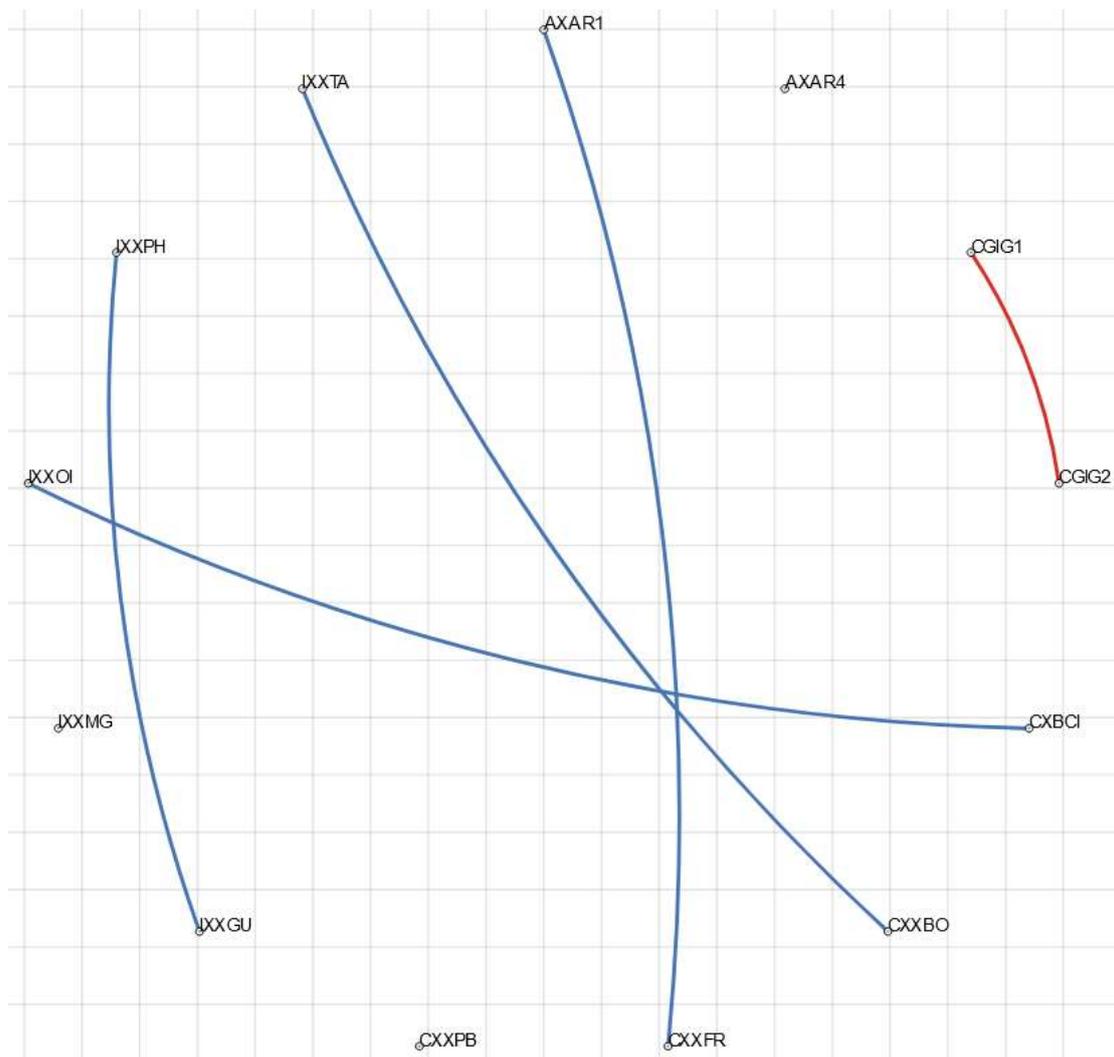


**Abbildung 31: Plot der Skygrid Rekonstruktion des Verlaufs der Populationsgröße von *P.sem* Hepaciviren**

Gezeigt ist die abgeschätzte Änderung der Populationsgröße über die Jahre anhand des Plot der Skygridanalyse. Der dazugehörigen phylogenetische Baum des BEAST Laufs ist in Abbildung 27 gezeigt. Der Plot zeigt den graduellen Anstieg der Populationsgröße mit zwei Einbrüchen (erkennbar an den beiden Spitzen der maximalen Werte) um ca. 1900 sowie 1980.

Bei der BSSVS Analyse wurde die Übergangswahrscheinlichkeit, oder Transition zwischen den Habitaten errechnet. Zusätzlich kann auch das Maß der Effektstärke der Transitionen, als Bayes Faktor, beschrieben werden (Abbildung 32 und Tabelle 39).

Im Anschluss wurden GLM genutzt um den Einfluss der ökologischen Faktoren auf die Virusverteilung zu zeigen (Tabelle 40). Dabei wurden die fünf Modelle benutzt wie in Kapitel 3.1.1 beschrieben. Der Datensatz aus 983 Proben wurden bereinigt, so das keine fehlenden Metadaten mehr vorhanden waren. Dies reduzierte den Datensatz auf 523 Proben.



**Abbildung 32: Plot der BSSV Analyse mit einem 50% Cutoff der Posterior-Wahrscheinlichkeit**  
Gezeigt ist ein Plot der Transitionswahrscheinlichkeit zwischen Habitaten. Der Farbgradient verläuft dabei anhand des Bayes Faktors von 0,5 (blau) bis 130000 (rot). Exakte Werte sind in Tabelle 39

**Tabelle 39: Resultate der BSSVS Analyse als Übergangswahrscheinlichkeiten zwischen Habitatpaaren und die Effektstärke des Übergangs**

Gezeigt sind die Übergangswahrscheinlichkeiten (Posterior Wahrscheinlichkeit) und der Bayes Faktor als Effektstärke zu Habitatpaaren.

Von	Zu	Bayes Faktor	Posterior Wahrscheinlichkeit
CGIG1	CGIG2	128569.6	1.000
CXXBO	IXXTA	746.8	0.993
IXXGU	IXXPH	194.5	0.974
CXBCI	IXXOI	16.7	0.764
AXAR1	CXXFR	9.9	0.658

**Tabelle 40: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von Hepaciviren in *P.sem***

Gezeigt sind die Ergebnisse des, in R berechneten, GLM der detektierten Hepacivirus Infektionen in Abhängigkeit zu ökologischen Faktoren. Als Vergleichsregressor für kategoriale Variablen diente der erste Wert der Variable. Als Güte wurde McFadden's Pseudo-R<sup>2</sup> berechnet. Für die Regression wurden jeweils 523 Proben benutzt. In Model 5 werden, außer *P.sem*, nur signifikante Spezies gezeigt. Signifikanzbeschriftung wurde wie folgt festgesetzt:

\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

Model 1 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.100)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Habitatgröße	2.18E-05	2.67E-05	0.41472	
Biodiversität	-2.21E-01	3.19E-02	1.27E-11	***
Blätterdachabdeckung	4.06E-04	4.04E-03	0.91997	
Anzahl Attalea	4.18E-02	2.66E-02	0.11607	

Model 2 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.091)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Unterholzdichte	-0.010412	0.003083	0.000788	***
Anzahl Astrocaryu	0.115327	0.017798	2.14E-10	***
Anzahl Piper	-0.011915	0.015314	0.436908	

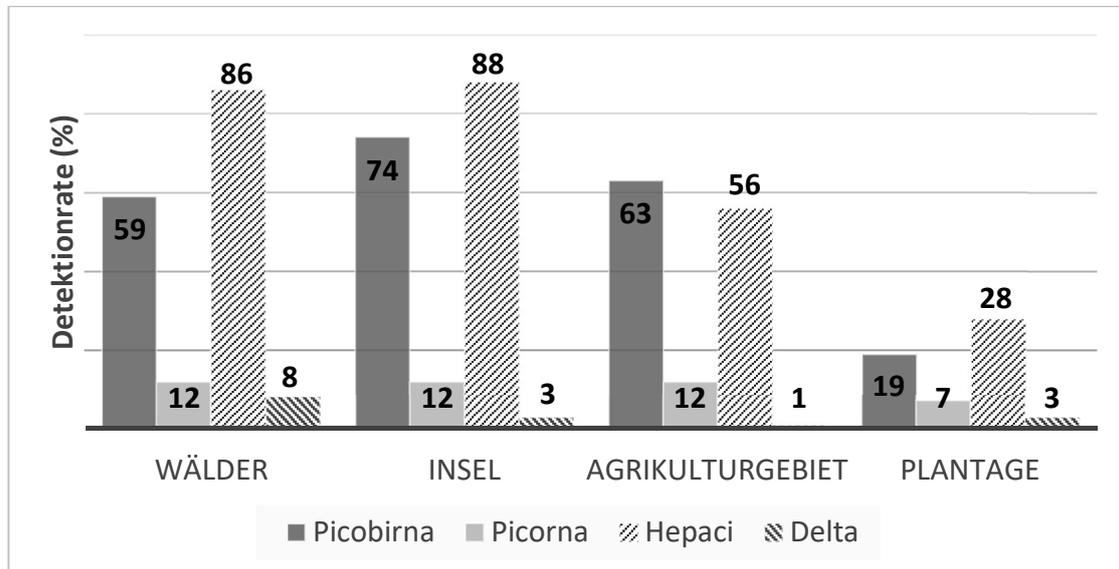
Model 3 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.100)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Anzahl Astrocaryu	0.11859	0.017763	6.31E-11	***
Anzahl Piper	0.011244	0.015282	0.462	
Populationsdichte	0.006821	0.001662	4.73E-05	***

Model 4 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.013)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Blätterdachabdeckung	0.006157	0.004546	0.1762	
Blätterdachhöhe	-0.012435	0.006584	0.0595	
Anzahl Attalea	0.026473	0.027834	0.342	
Anzahl Piper	0.026385	0.015692	0.0933	

Model 5 (Güte als McFadden's Pseudo-R <sup>2</sup> = 0.143)				
Faktor	Estimate	Std. Error	P-Value	Signifikanz
Sex	-0.007431	0.036863	0.840303	
Alter	-0.154512	0.066081	0.019685	*
Spezies Philander opossum	1.04179	0.17733	7.62E-09	***
Spezies <i>P.sem</i>	0.65119	0.09614	3.46E-11	***
Habitattyp C	0.29521	0.04158	4.18E-12	***
Habitattyp I	0.23645	0.04924	2.07E-06	***

### 3.1.6. Detektionsraten aller Viren in *Proechimys semispinosus*

Im Folgenden sind alle vorangegangener Detektionsraten der vier untersuchten Viren in allen Habitaten graphisch (Abbildung 33) und tabellarisch (Tabelle 41) aufgelistet.



**Abbildung 33: Detektionsraten aller *P.sem* Virusfunde für jeden Habitattyp**

Angegeben sind die Detektionsraten basierend auf den vorangegangenen Virusdetektionen per PCR und NGS pro Gesamtzahl an getesteten Proben in den vier unterschiedlichen Habitattypen.

**Tabelle 41: Gesammelte Resultate aller Virusdetektionen innerhalb der Generalistenspezies *Proechimys semispinosus***

Angegeben ist die absolute Anzahl an detektierten Virusfunden für Picobirna-, Picorna-, Hepaci- und Rodentdeltaviren aus *P.sem*, mit prozentualen Anteil in Klammern, in den vier Habitattypen Wald, Insel, Agrikulturgebiet und Plantagen.

Habitattyp	Detektion PBV	Detektion Picornavirus	Detektion Hepaciviren	Detektion RDeV
Wald	139/236 (58,9%)	29/236 (12,3%)	185/215 (86,0%)	17/215 (7,9%)
Insel	288/392 (73,5%)	48/392 (12,2%)	309/350 (88,3%)	10/350 (2,9%)
Agrikulturgebiet	104/165 (63,0%)	20/165 (12,1%)	89/158 (56,3%)	2/158 (1,3%)
Plantagen	5/27 (18,5%)	2/27 (7,4%)	11/40 (27,5%)	1/40 (2,5%)
<b>Gesamtergebnis</b>	<b>536/820 (65,4%)</b>	<b>99/820 (12,1%)</b>	<b>594/763 (77,9%)</b>	<b>30/763 (3,9%)</b>

Mit dem Chi Quadrat Test wurde nun geprüft ob sich die Detektionsraten der jeweiligen Viren zwischen allen Habitattypen unterscheiden bzw. ob es Unterschiede in den Detektionsraten aller Viren in jeweiligen Habitattypen gab (Tabelle 42).

**Tabelle 42: Resultate des Chi Quadrat-Tests auf Unabhängigkeit zwischen der Detektionsrate und den Habitattyp**

Angegeben ist der p-Wert der Chi-Quadrat-Tests. Zur Berechnung wurden pro untersuchter Unabhängigkeit (Habitattyp oder Detektionsrate), jeweils vier 2\*4 Matrizen gebildet und mit den Werten für positive oder negative Detektion gefüllt. Aus diesen Matrizen wurden die erwarteten Werte errechnet und in Excel der CHIQU-Test benutzt.

Habitattyp	p-Wert	Virus	p-Wert
Wald	$1.94 \cdot 10^{-64}$	Picornavirus	0.298
Insel	$8.14 \cdot 10^{-148}$	Picobirnavirus	$3.46 \cdot 10^{-9}$
Agrikulturgebiet	$5.49 \cdot 10^{-37}$	Deltavirus	0.004
Plantagen	0.009	Hepacivirus	$2.30 \cdot 10^{-28}$

Damit unterschieden sich die Detektionsraten aller Viren innerhalb eines Habitattyps. Die Detektionsrate der jeweiligen Viren, betrachtet über alle verschiedenen Habitattypen unterscheidet sich auch signifikant, mit Ausnahme von Picornaviren, bei denen kein Unterschied zwischen den Detektionsraten über die Habitattypen festgestellt werden konnte.

## 3.2. Diskussion

### 3.2.1. Picornavirus Auswertung

Es wurden mehrere PCRs mit Vollgenom-Primern, auf bekannt positiven Proben, angesetzt um die Menge an Sequenzinformationen zu vergrößern und eventuell vollständige Virusgenome zu erhalten. Bei vielen der Proben stellte sich heraus, dass die vorhandene Viruskonzentration nicht ausreichend war um mehr Sequenzinformationen zu erhalten, so dass es nur möglich war ein komplettes Vollgenom und ein zu 80% vollständiges Vollgenom zu generieren. Durch die, im Vergleich zu den anderen untersuchten Viren, niedrige Prävalenz (~12%, s. Abbildung 33) in Kombination mit den Schwierigkeiten der Vollgenomgenerierung, durch niedrige Konzentrationen, ist es schwer große Mengen an Sequenzinformation zu sammeln.

Die vorhandenen Daten lassen darauf schließen, dass drei distinkte Kladen identifiziert wurden. Eine genaue Einordnung der drei Kladen ist zum jetzigen Zeitpunkt nicht möglich, da nicht ausreichend Sequenzdaten zur Verfügung standen. Für das Vollgenom aus einer Klade (Gruppe 3, Abbildung 17) ist dies jedoch möglich. In Tabelle 24, in der die Aminosäure Identität zu naheverwandten Referenzsequenzen gezeigt wird, sieht man, dass sich das neuentdeckte Picornavirus in allen vier Regionen stark von anderen Viren der Familie *Picornaviridae* unterscheidet. Da somit die formalen Voraussetzungen erfüllt sind, handelt es sich hierbei um ein neues Genus innerhalb der Familie. Die Distanz auf Aminosäureebene (Tabelle 23) lässt darüber hinaus vermuten, dass es sich bei Gruppe 2 um das selbe, oder zumindest nahverwandtes Genus handelt. Bei Gruppe 1 hingegen, liegt die Vermutung nahe, dass es sich um ein weiteres Genus handelt. Dieser interessante Aspekt sollte aus diesem Grund durch weitere Hochdurchsatzsequenzierungen weiter erforscht werden.

Die logistische Regression fand multiple Faktoren, die Einfluss auf die Picornavirusprävalenz nehmen (Tabelle 25). Betrachtet man Model 5, so wird klar, dass es sich bei den gefundenen ökologischen Faktoren um Variablen handelt, die die Inseln beschreiben. Laut aller Model wird die Picornavirusprävalenz durch geringere Biodiversität, kleinere Habitate mit einer höheren Populationsdichte und weniger dichte Flora, besonders in *P.sem*, erhöht. All dies spricht dafür, dass die Umweltbedingungen auf kleinen, isolierten Inseln die Übertragung mit diesem Virus begünstigen. Da, aller Wahrscheinlichkeit das *P.sem* Picornavirus auch über eine oral-fäkale Route übertragen wird spielen ein geringeres Nahrungsaufkommen natürlich eine große Rolle. Bilden

dabei die Bäume der Gattung Piper, die Hauptnahrungsquelle, so sind diese starke Anlaufpunkte der Stachelratten, wodurch der Kontakt mit Kot umso wahrscheinlicher wird. Alleine dieser Fund ist ein Paradebeispiel, wie durch den menschlichen Eingriff in das Ökosystem, isolierte Virusverbunde geschaffen werden, die sich, durch die höhere Prävalenz in den spezifischen Habitaten schneller akkumulieren können.

### 3.2.2. Picobirnavirus Auswertung

Die Picobirnaviren zeichnen sich nicht nur durch ein enorm breites Wirtsspektrum, sondern auch auf eine ubiquitäre, globale Verbreitung aus. Trotz dieser Tatsachen ist über die Pathogenese nicht viel bekannt. In den hier gezeigten Daten spiegelt sich diese Tatsache wider, da neben *P.sem* auch ein hoher Teil an anderen Tieren infiziert wurde (Tabelle 28). Über die Hälfte alle getesteten Proben enthielten hierbei das PBV. Zusätzlich bestand, eine sehr hohe Grunddiversität in den Panamaproben. So lag die maximale Intrahost-Nukleotiddistanz (in *P.sem*) bei 45,5% und im Vergleich die Interhost-Nukleotiddistanz (zwischen *P.sem* und *D.mars*) 50,4%. Diese Diversität lässt am wahrscheinlichsten durch eine hohe Mutationsrate erklären. Durch die geringe Größe des Sequenzfragments (bzw. die allgemein niedrige Größe des Segments 2 auf denen die RdRp liegt) in Kombination mit dem natürlichen Sequenzspektrum von PBV sind Aussagen zur konkreten Genuseinordnung nur schwer möglich. Die Tatsache, dass die Wirte von PBV auch nur sehr grob zusammengefasst werden können (z.B. Säugetiere für die ersten beiden Genogruppen, Abbildung 20) erschwert die taxonomische Klassifizierung.

Zusammenfassend gesagt, ermöglichte der PBV Datensatz keine tiefergehenden phylogenetischen Analysen. Wie auch schon bei Picornaviren, zeigte die logistische Regressionsanalyse, einen starken Zusammenhang zwischen Nahrungsquellenmangel und Populationsdichte der sich positiv auf die Picobirnavirusprävalenz ausübt. Außerdem sollte erwähnt werden, dass *Marmosa robinsoni isthmica* eine interessante Rolle bei der Übertragung von PBV zu spielen scheint: Die Anwesenheit wirkt sich negativ auf die Virusprävalenz aus. Dies gibt Hinweis darauf, dass *Marmosa robinsoni isthmica* einen natürlichen Resistenz für die Infektion von PBV zu haben scheint.

### 3.2.3. Deltavirus Auswertung

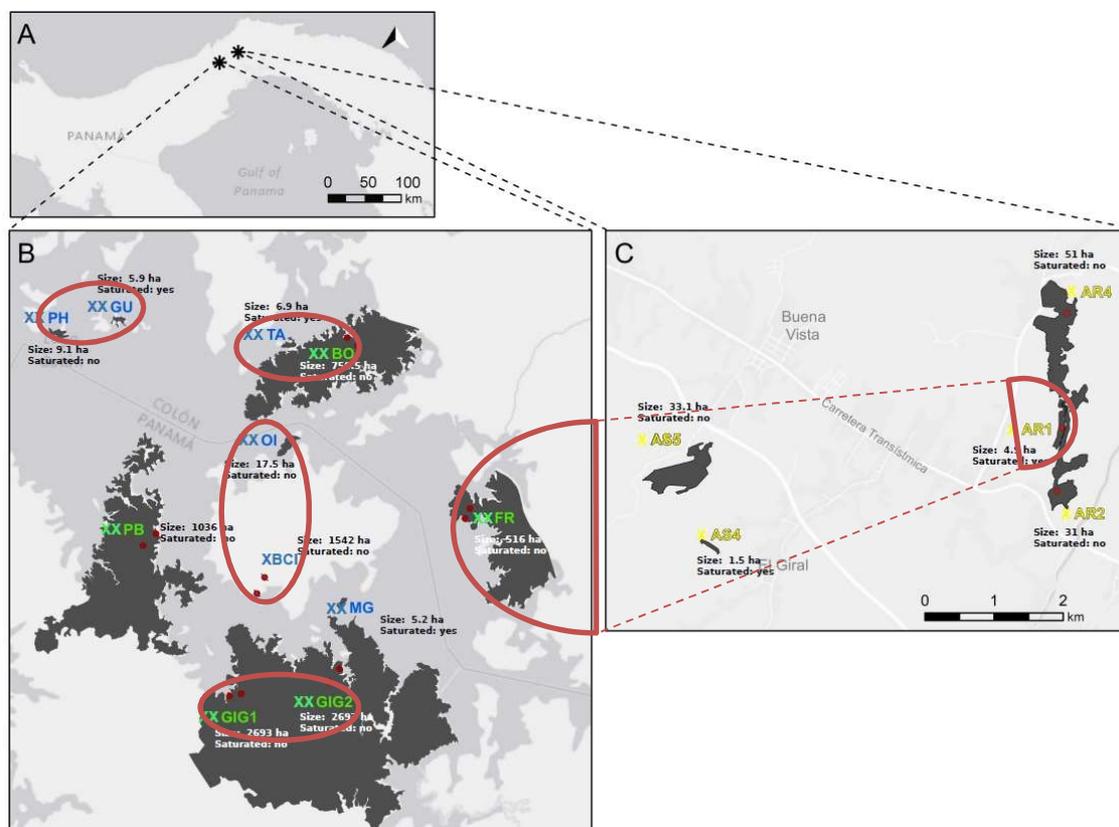
Deltaviren wurden bis 2018 [63] ausschließlich in Menschen gefunden. Umso überraschender war der Fund in den Stachelratten als zweiter bekannter Säugetierwirt. Durch die Anwendung verschiedener PCR Techniken zeigten sich offensichtliche Parallelen aber auch Unterschiede, zum humanen Hepatitis Deltavirus. So liegt RDeV, wie auch HDV, zirkulär vor (Abbildung 14). Im Menschen ist HDV nur als sekundär Infektion einer Hepatitis B Virusinfektion möglich, da HDV die Hüllproteine von HBV benötigt um infektiöse Partikel zu bilden. In den *P.sem* Proben konnte jedoch kein Anzeichen für ein Hepadna-Helfervirus gefunden werden. Ein weiterer Unterschied konnte in der Konzentrationsbestimmung in den Organen gezeigt werden, da RDeV, im Gegensatz zu HDV, keinen Lebertropismus aufwies (Tabelle 34). Daneben zeigte die Sequenzanalyse, dass RDeV die geringste patristische Distanz, im Vergleich zu anderen Deltaviren, zu HDV besaß (Abbildung 22). Die Abwesenheit eines Helfervirus konnte, bis jetzt nicht geklärt werden. Allerdings zeigt die logistische Regressionsanalyse, dass ein starker Zusammenhang zwischen der RDeV Prävalenz und größeren Habitaten besteht. Da die untersuchten Wälder, durch ihre Größe, keinen vermehrten Wirtkontakt begünstigen, stellt sich die Frage was der Grund hierfür ist. In Kombination mit der Tatsache, dass kein Helfervirus identifiziert werden konnte, lässt die Wald-Assoziation vermuten, dass diese am wahrscheinlichsten dort zu finden ist. Eine weitere Möglichkeit wäre, dass sich eine andere primäre Infektionsquelle in den Wäldern befindet. Diese Überlegung basiert auf der Tatsache, dass es viele Vermutungen über den tatsächlichen Ursprung von Deltaviren gibt. Eine Theorie besagt, dass Deltaviren aufgrund ihrer Ähnlichkeit zu Viroiden, eine Ursprung in Pflanzen haben könnten. Mit dem primären Auftreten in Wald Habitaten ist dies ein spannender Ansatz, der weitere Untersuchungen bedarf.

Weitere experimentelle Daten, zur RDeV Bestimmung (wie Zellkultur, Replikationsexperimente und Proteindetektion) sind im assoziierten Paper Paraskevopoulou, Pirzer [100] 2020 zu finden.

### 3.2.4. Hepacivirus Auswertung

Von allen vier betrachteten Viren war die Prävalenz der Hepaciviren am höchsten. Durch die hohe Anzahl an positiven Proben, mit hohen Konzentrationen, konnten per NGS mehrere Vollgenome generiert werden. Diese hohe Anzahl an Sequenzinformationen

(343 Sequenzen) wurde durch die Detektion von wahrscheinlichen Rekombinationsereignissen wieder reduziert. Bei dieser Reduktion wurde die letzten Vertreter des Habitattyps „Plantage“ entfernt, weswegen in folgenden Analysen nur noch der Vergleich zwischen „Wald“, „Insel, und „Agrikulturgebiet“ gezogen wurde. Eine initiale Analyse per BEAST mit einem relaxed Clock Model, bestätigte die Vermutung, dass die Mutationsrate über alle Habitatstypen in etwa gleich groß war (Abbildung 28) und dennoch mit Literaturwerten vergleichbar blieb [104]. Aus diesem Grund wurde für weitere BEAST Analysen ein strict Clock Model gewählt, da es nun nicht mehr nötig war, mögliche Ratenvariationen an verschiedenen Ästen des phylogenetischen Baums zuzulassen. Durch diese Reduktion der Parameter wurde eine zusätzliche BSSVS Analyse in die BEAST Läufe mitaufgenommen.



**Abbildung 34: Visualisierung der BSSVS Analyse aus BEAST Läufen von Hepaciviren**

Durch die BSSVS Analyse gefundenen signifikanten Übergänge (50% Posterior cutoff) (Abbildung 32 und assoziierte Daten) sind hier rot eingetragen.

Die Ergebnisse der BSSVS Analyse zeigen einen eindeutigen Zusammenhang zwischen den phylogenetischen und den geografischen Daten (Abbildung 34). Das Skygrid Resultat, aus demselben Lauf zeigt zwei Populationseinbrüche. Bei dem ersten Einbruch, um ca. 1900, könnte es sich um den Zeitpunkt der Überflutung des Waldgebietes handeln. Der Bau des Panamakanals wurde zwar erst 1911 fertiggestellt

aber bereits Jahre zuvor wurde damit begonnen, die umgebenden Flüsse aufzustauen. Zu beachten ist weiterhin, dass der Probenahmezeitraum nur 3 Jahre umfasste, was zu einem sehr schwachen, temporalen Signal führt. Durch diese beiden Tatsachen, fällt der Schwankungsbereich für das erste Ereignis grob mit der Überflutung des Urwaldes zusammen. Der stärkere, zweite Einbruch konnte leider nicht durch historische Ereignisse oder andere Faktoren definitiv benannt werden. Jedoch finden sich in der Literatur andere Langzeitstudien, die über einen konstanten Artenrückgang, anderer Tierarten in Panama, berichten [105, 106].

Die Theorie des „Verdünnungseffekts“ besagt, dass durch eine höhere Biodiversität, die Virusprävalenz sinkt. Die logistische Regression (Tabelle 40) bestätigt diesen Zusammenhang. Für die Hepacivirusprävalenz ist ein positiver Effekt, bei sinkender Biodiversität zu beobachten. Auch spielt hier die Populationsdichte eine große Rolle. Dies kann damit erklärt werden, dass es bei *P.sem* in kleinen Habitaten zu Revierkämpfen kommen kann. Neben dem, offensichtlichen Habitataffekt, ist die der starke Effekt von *Philander opossum* sehr interessant. Opossums sind Allesfresser und ernähren sich von Insekten, Pflanzen aber auch Kleinstnagetieren. Der hier beobachtete Effekt könnte aus der Tatsache entstehen, dass *P.sem* in *Philander opossum* einen Fressfeind hat.

## 4. Zusammenfassung

Im Rahmen dieser Dissertation wurde ein breites Spektrum an Nage- und Beuteltierproben analysiert. Dies führte zu vielen interessanten Entdeckungen, wie etwa die große Prävalenz an Picobirnaviren oder die Detektion von mindestens zwei, bis dato unbekanntem, neuen Genera aus der Familie *Picornaviridae*. Zudem gelang erstmaligen Nachweis eines Deltavirus aus einem Säugetier abseits vom Menschen, sowie dessen molekularbiologische Charakterisierung und Unterschied im Replikationsverhalten gegenüber Hepatitis Deltaviren. Darüber hinaus brachte die Detektion und bioinformatische Analyse der Hepaciviren einen enorm großen Wissensgewinn. Darüber hinaus konnte die Theorie, die hinter dem ökologischen „Verdünnungseffekt“, nämlich die Abhängigkeit von Speziesdiversität zu Prävalenz, klar bestätigt werden. All diese Untersuchungen verschiedener Proben um den Gatún See in Panama zeigten viele, menschengemachte Einflüsse auf die ansässigen Viruspopulationen auf. Durch den Bau des Panama Kanals wurden weite Teile des Urwalds überflutet, woraufhin der Gatún See entstand. Zusätzlich zu diesem initialen Terraforming, kam es durch den wirtschaftlichen Aufschwung zu einer Ausbreitung urbaner Flächen. All dies führte dazu, dass der Mensch mehr Kontakt zu diversifizierenden, potenziell zoonotischen, sowie bislang noch unbekanntem Viruspopulationen hat. Und wie uns die jüngste Zeit zeigt, ist die Gefahr einer Pandemie mit zoonotischen Ursprung allgegenwärtig und sollte nicht unterschätzt werden. Aus diesem Grund ist es wichtiger denn je die Zusammenhänge zwischen verschiedenen Einflussfaktoren zu verstehen und eine Ausbreitung beziehungsweise das Erscheinen neuer zoonotischer Erreger vorhersagen zu können. Die aktuelle Pandemie zeigt wie relevant das Wissen um die Auswirkungen dieser Faktoren auf Pathogenität und Infektiosität (bzw. Infektionsrisiko) ist. Die Ergebnisse dieser Arbeit sollen dazu dienen, wie im Fall der Deltaviren, den bislang unbekanntem Ursprung dieser Familie Deltaviren zu ergründen. All diese Grundlagen können in weiterführenden Studien nicht nur helfen, Pathogen-Wirts Interaktionen besser zu verstehen, sondern auch mögliche medizinische und ökologische Relevanzen für den Menschen abzuschätzen.

## 4.1. Summary

In this dissertation, a wide range of rodent and marsupial samples were analyzed. This led to many interesting discoveries, such as the high prevalence of picobirnaviruses or the detection of at least two, previously unknown, new genera from the family *Picornaviridae*. In addition, the first detection of deltavirus outside of humans was achieved, as well as its molecular characterization and difference in replication behavior compared to hepatitis deltaviruses. Furthermore, the detection and bioinformatic analysis of hepaciviruses provided a tremendous increase in knowledge. Moreover, the theory behind the ecological "dilution effect", namely the dependence of species diversity on prevalence, was clearly confirmed. All these studies of different samples around Lake Gatún in Panama revealed many, man-made influences on the resident virus populations. The construction of the Panama Canal flooded large areas of virgin forest, resulting in the creation of Lake Gatún. In addition to this initial terraforming, an expansion of urban areas occurred as a result of the economic boom. All of this led to humans having more contact with diversifying, potentially zoonotic, as well as yet unknown viral populations. And as recent times have shown us, the threat of a pandemic of zoonotic origin is ever-present and should not be underestimated. For this reason, it is more important than ever to understand the interrelationships between different influencing factors and to be able to predict the spread or emergence of new zoonotic pathogens. The current pandemic shows how relevant the knowledge of the impact of these factors on pathogenicity and infectivity (or risk of infection) is. The results of this work will be used, as in the case of deltaviruses, to elucidate the hitherto unknown origin of this family of deltaviruses. All these fundamentals can help in further studies not only to better understand pathogen-host interactions, but also to assess possible medical and ecological relevance for humans.

## 5. Abbildungsverzeichnis

Abbildung 1: Darstellung des Isthmus von Panama vor und nach dem Bau des Kanals	
Abbildung 2: <i>Proechimys semispinosus</i> Stammbaum.....	5
Abbildung 3: Geografische Verbreitung von <i>Proechimys semispinosus</i> .....	6
Abbildung 4: Vereinfachte Darstellung der Genomorganisation von Picobirnaviren .....	7
Abbildung 5: Übersicht der Verbreitung von Picobirnaviren weltweit.....	8
Abbildung 6: Vereinfachte Darstellung der Genomorganisation von Picornaviren .....	9
Abbildung 7: Vereinfachte Darstellung der Genomorganisation von HDV .....	10
Abbildung 8: Phylogenetischer Stammbaum von Deltaviren in verschiedenen Wirten	11
Abbildung 9: Vereinfachte Darstellung der Genomorganisation von Hepatitis C Viren	12
Abbildung 10: Geografische Lage der Fangorte und Habitate um den Gatún See .....	19
Abbildung 11: Schematischer Ablauf einer Polymerasekettenreaktion.....	23
Abbildung 12: Schematische Darstellung der Lage der Picornavirus Screening PCR Primer.....	28
Abbildung 13: Schematische Darstellung der Detektion mittels Taq-Man Sonde .....	31
Abbildung 14: Vereinfachte Darstellung des Rodent-Deltavirus Genoms mit Zirkularisierungsassays .....	35
Abbildung 15: Grafische Darstellung mehrerer ökologischer Parameter .....	51
Abbildung 16: Ausschnitt des Sensivitätstest der Picornavirus Screening PCR .....	52
Abbildung 17: Phylogenetischer Baum der 3D Region von Picornaviren.....	56
Abbildung 18: Beispielhafte Visualisierung der Amplifikate der Screening PCR für PBV .....	59
Abbildung 19: Nukleotiddistanz von PBV in <i>P.sem</i> innerhalb der drei Habitattypen. ....	62
Abbildung 20: Phylogenetischer Baum der Genogruppen von PBV inkl. Referenzen..	65
Abbildung 21: Vergrößerter Ausschnitt des phylogenetischer Baum der Genogruppen von PBV inkl. Referenzen .....	66
Abbildung 22: Phylogenetischer Baum der Deltaviren mit Vergleich der Nt-Distanzen	71
Abbildung 23: Phylogenetischer Analyse der NS5b Region von Hepaci- und Pegiviren .....	76
Abbildung 24: Rekombinationsanalyse von Hepaciviren aus <i>P.sem</i> .....	77
Abbildung 25: Bestimmung des Substitutionsmodells mit jModeltest .....	79
Abbildung 26: Phylogenetischer bayesianischer Baum der <i>P.sem</i> Hepaciviren in radialer Ansicht.....	80
Abbildung 27: Phylogenetischer bayesianischer Baum der <i>P.sem</i> Hepaciviren mit Metadaten errechnet durch ein relaxed Clock-Model .....	81
Abbildung 28: Vergleich der Evolutionsraten von <i>P.sem</i> Hepaciviren .....	82
Abbildung 29: Traceranalyse der Mutationsrate zwischen den Codonpositionen .....	83
Abbildung 30: Phylogenetischer bayesianischer Baum der <i>P.sem</i> Hepaciviren in radialer Ansicht des BEAST Laufs mit einem strict Clock Model.....	84
Abbildung 31: Plot der Skygrid Rekonstruktion des Verlaufs der Populationsgröße von <i>P.sem</i> Hepaciviren.....	85
Abbildung 32: Plot der BSSV Analyse mit einem 50% Cutoff der Posterior-Wahrscheinlichkeit.....	86
Abbildung 33: Detektionsraten aller <i>P.sem</i> Virusfunde für jeden Habitattyp .....	88

Abbildung 10: Visualisierung der BSSVS Analyse aus BEAST Läufen von Hepaciviren  
..... 93

## 6. Tabellenverzeichnis

Tabelle 1: Reaktionsansatz und Temperaturprotokoll der Picobirnavirus Screening PCR .....	24
Tabelle 2: Primerliste für die Picobirnavirus Screening PCR .....	24
Tabelle 3: Reaktionsansatz und Temperaturprotokoll der Picornavirus Initial-Screening PCR .....	26
Tabelle 4: Reaktionsansatz und Temperaturprotokoll der neuen Picornavirus Screening PCR .....	27
Tabelle 5: Verwendete Oligonukleotide für die Picornavirus Screening PCR .....	28
Tabelle 6: Reaktionsansatz und Temperaturprotokoll der Hepacivirus Screening PCR .....	29
Tabelle 7: Verwendete Oligonukleotide für die Hepacivirus Screening PCR .....	29
Tabelle 8: Reaktionsansatz und Temperaturprotokoll der Deltavirus Screening PCR .....	30
Tabelle 9: Sequenzen und Verwendung der Oligonukleotide für Hepatitis Deltavirus PCR .....	31
Tabelle 10: Reaktionsansatz der cDNA Synthese des Hepatitis Deltavirus .....	34
Tabelle 11: Reaktionsansatz und Temperaturprotokoll des Hepatitis Deltavirus Zirkularisierungsassays .....	36
Tabelle 12: Reaktionsansatz und Temperaturprotokoll der Deltavirus Klonierungsfragment PCR .....	37
Tabelle 13: Reaktionsansatz und Temperaturprotokoll der Deltavirus Kolonie PCR .....	38
Tabelle 14: Übersicht der Blutproben aus allen Spezies und Habitaten .....	47
Tabelle 15: Übersicht der Kotproben aus allen Spezies und Habitaten .....	48
Tabelle 16: Korrelationsstärke aller ökologischen Faktoren untereinander .....	49
Tabelle 17: Signifikanz der errechneten Korrelationsstärke der ökologischen Faktoren .....	50
Tabelle 18: Einteilung der ökologischen Faktoren in vier Untergruppen .....	50
Tabelle 19: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von Picornavirus-RNA untersucht wurden .....	53
Tabelle 20: Picornavirusprävalenz Aufschlüsselung pro Habitattyp .....	54
Tabelle 21: Numerische Verteilung der Picornavirussequenzen innerhalb der definierten Gruppen .....	54
Tabelle 22: Prozentuale Identität auf Nukleotidebene der <i>P.sem</i> Picornavirusgruppen untereinander und zu Referenzsequenzen .....	55
Tabelle 23: Prozentuale Identität auf Aminosäureebene der <i>P.sem</i> Picornavirusgruppen untereinander und zu Referenzsequenzen .....	55
Tabelle 24: AA Identität des Picornavirus Vollgenomes zu nahverwandten Referenzen .....	57
Tabelle 25: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von Picornaviren .....	58
Tabelle 26: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von PBV-RNA untersucht wurden .....	60
Tabelle 27: Picobirnavirusprävalenz aufgeschlüsselt nach Habitattyp .....	60
Tabelle 28: Speziesverteilung innerhalb der PBV-Sequenzcontigs .....	61
Tabelle 29: Berechnete Effektstärken der p-Distanzen von Picobirnaviren .....	63

Tabelle 30: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von PBV.....	67
Tabelle 31: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von Deltavirus-RNA untersucht wurden.....	68
Tabelle 32: Logistische Regression zur Untersuchung des Effekts des Habitattyps auf das Virusaufkommen von RDeV anhand eines GLM .....	69
Tabelle 33: Deltavirusprävalenz Aufschlüsselung pro Habitattyp .....	70
Tabelle 34: Anzahl der RDeV-RNA Kopien in Organen von <i>P.sem</i> .....	72
Tabelle 35: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von RDeV .....	73
Tabelle 36: Überblick über die untersuchten Nagetierspezies per Habitattyp, die auf das Vorhandensein von PBV-RNA untersucht wurden .....	74
Tabelle 37: Übersicht der Anzahl an Proben aus <i>P.sem</i> von Hepaciviren über alle Habitate und Datensets .....	75
Tabelle 38: Verteilung der Proben innerhalb des reduzierten <i>P.sem</i> Hepacivirus Alignments aus 231 Sequenzen über die Habitate.....	78
Tabelle 39: Resultate der BSSVS Analyse als Übergangswahrscheinlichkeiten zwischen Habitatpaaren und die Effektstärke des Übergangs .....	86
Tabelle 40: Logistische Regression zur Untersuchung des Effekts der ökologischen Faktoren auf das Virusaufkommen von Hepaciviren in <i>P.sem</i> .....	87
Tabelle 41: Gesammelte Resultate aller Virusdetektionen innerhalb der Generalistenspezies <i>Proechimys semispinosus</i> .....	88
Tabelle 41: Resultate des Chi Quadrat-Tests auf Unabhängigkeit zwischen der Detektionsrate und den Habitattyp .....	89

## 7. Abkürzungsverzeichnis

µL	Mikroliter
AA	Aminosäure (aminoacid)
ANOVA	Varianzanalyse (analysis of variance)
BEAST	Bayesian Evolutionary Analysis Sampling Trees
BEAUTI	Bayesian Evolutionary Analysis Utility
bp	Basenpaare
CDS	Kodierende Sequenz (coding sequence)
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleosid-Triphosphat
GLM	Generalized linear model
GTR	General Time Reversible
ha	Hektar
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HDAg	Hepatitis Delta Antigen
HDV	Hepatitis Delta Virus
IVT	In-Vitro Transkript
L-HDAg	large Hepatitis Delta Antigen
MCMC	Markov chain Monte Carlo
mL	Milliliter
ML	Maximum-Likelihood
NANBH	non-A-non-B Hepatitis

---

NGS	Next generation sequencing
NJ	Neighbor-Joining
OD	Optische Dichte
ORF	Open Reading Frame
P.sem	Proechimys semispinosus
PBS	Phosphatgepufferte Salzlösung (phosphate buffered saline)
PBV	Picobirnavirus
PCR	Polymerasekettenreaktion (polymerase chain reaction)
qPCR	Quantitative PCR
RDeV	Rodent Delta Virus
RdRp	RNA abhängige RNA-Polymerase
RNA	Ribonukleinsäure
RT-PCR	Reverse-Transkription PCR
S-HDAg	small Hepatitis Delta Antigen

## 8. Literaturverzeichnis

1. Forester, D.J. and G.E. Machlis, *Modeling human factors that affect the loss of biodiversity*. Conservation Biology, 1996. **10**(4): p. 1253-1263.
2. Wolfe, N.D., et al., *Bushmeat hunting, deforestation, and prediction of zoonoses emergence*. Emerg Infect Dis, 2005. **11**(12): p. 1822-7.
3. Holmes, E.C., et al., *The origins of SARS-CoV-2: A critical review*. Cell, 2021. **184**(19): p. 4848-4856.
4. Sanchez, E.L. and M. Lagunoff, *Viral activation of cellular metabolism*. Virology, 2015. **479-480**: p. 609-618.
5. Civitello, D.J., et al., *Biodiversity inhibits parasites: Broad evidence for the dilution effect*. Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(28): p. 8667-8671.
6. Johnson, P.T.J. and D.W. Thieltges, *Diversity, decoys and the dilution effect: how ecological communities affect disease risk*. Journal of Experimental Biology, 2010. **213**(6): p. 961-970.
7. Ostfeld, R.S., F. Keesing, and V.T. Eviner, *Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems*. 2010: Princeton University Press.
8. Turney, S., A. Gonzalez, and V. Millien, *The negative relationship between mammal host diversity and Lyme disease incidence strengthens through time*. Ecology, 2014. **95**(12): p. 3244-3250.
9. Bennett, I.E., *History of the Panama Canal; its construction and builders*. 1915, Washington, D.C.: Historical Publishing Company. xi, 5-543 p. incl. col. front.
10. Weaver, P.L., G.P. Bauer, and I.I.o.T. Forestry, *The San Lorenzo Protected Area: A Summary of Cultural and Natural Resources*. 2004: U.S. Department of Agriculture, Forest Service, International Institute of Tropical Forestry.
11. *Proceedings of the Zoological Society of London*. Vol. 28 (1860). 1860, London: Zoological Society of London. p. 265.
12. Thomas, O., *The Annals and magazine of natural history; zoology, botany, and geology*. Vol. 20 (1897). 1897, London: Taylor and Francis, Ltd. 550.
13. Allen, J.A., *The generic names Echimys and Loncheres*. Vol. 12. 1899: Bulletin of the American Museum of Natural History. p. 257.

14. *Bulletin of the Museum of Comparative Zoology at Harvard College*. Vol. 46 (1904-1906). 1905, Cambridge: Museum of Comparative Zoology at Harvard College. p. 89.
15. Patton, J.L., *Species groups of spiny rats, genus Proechimys (Rodentia: Echimyidae)*. Studies in neotropical mammalogy : essays in honor of Philip Hershkovitz /, 1987. **39**: p. 305-345.
16. Gardner, A.L., *Proechimys semispinosus (Rodentia: Echimyidae): Distribution, type locality, and taxonomic history*. Proceedings of the Biological Society of Washington., 1983. **96**: p. 134-144.
17. Patton, J.L., U.F.J. Pardiñas, and G. D'Elía, *Mammals of South America, Volume 2: Rodents*. 2015: University of Chicago Press.
18. Adler, G.H. and T.D. Lambert, *Ecological correlates of trap response of a Neotropical forest rodent, Proechimys semispinosus*. Journal of Tropical Ecology, 1997. **13**: p. 59-68.
19. Adler, G.H. and J.O. Seamon, *Distribution and Abundance of a Tropical Rodent, the Spiny Rat, on Islands in Panama*. Journal of Tropical Ecology, 1991. **7**: p. 349-360.
20. Wright, S.J., *How Isolation Affects Rates of Turnover of Species on Islands*. Oikos, 1985. **44**(2): p. 331-340.
21. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data*. Bioinformatics, 2012. **28**(12): p. 1647-9.
22. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7: improvements in performance and usability*. Mol Biol Evol, 2013. **30**(4): p. 772-80.
23. Guindon, S., et al., *New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0*. Syst Biol, 2010. **59**(3): p. 307-21.
24. Voloch, C.M., et al., *Phylogeny and chronology of the major lineages of New World hystricognath rodents: insights on the biogeography of the Eocene/Oligocene arrival of mammals in South America*. BMC Research Notes, 2013. **6**(1): p. 160.
25. Rosen, B.I., et al., *Cloning of human picobirnavirus genomic segments and development of an RT-PCR detection assay*. Virology, 2000. **277**(2): p. 316-29.
26. Fregolente, M.C., et al., *Molecular characterization of picobirnaviruses from new hosts*. Virus Res, 2009. **143**(1): p. 134-6.

27. Delmas, B., et al., *ICTV virus taxonomy profile: Picobirnaviridae*. J Gen Virol, 2019. **100**(2): p. 133-134.
28. Ganesh, B., et al., *Picobirnavirus infections: viral persistence and zoonotic potential*. Rev Med Virol, 2012. **22**(4): p. 245-56.
29. Fregolente, M.C. and M.S. Gatti, *Nomenclature proposal for picobirnavirus*. Arch Virol, 2009. **154**(12): p. 1953-4.
30. Pereira, H.G., et al., *A virus with a bisegmented double-stranded RNA genome in rat (*Oryzomys nigripes*) intestines*. J Gen Virol, 1988. **69** ( Pt 11): p. 2749-54.
31. Ganesh, B., G. Masachessi, and Z. Mladenova, *Animal picobirnavirus*. Virusdisease, 2014. **25**(2): p. 223-38.
32. Day, J.M., et al., *Metagenomic analysis of the turkey gut RNA virus community*. Virol J, 2010. **7**: p. 313.
33. Finkbeiner, S.R., et al., *Metagenomic analysis of human diarrhea: viral detection and discovery*. PLoS Pathog, 2008. **4**(2): p. e1000011.
34. Ghosh, S., et al., *Molecular characterization of full-length genomic segment 2 of a bovine picobirnavirus (PBV) strain: evidence for high genetic diversity with genogroup I PBVs*. J Gen Virol, 2009. **90**(Pt 10): p. 2519-24.
35. Alfieri, A.A., et al., *Ocurrence of Escherichia Coli, rotavirus, picobirnavirus and Cryptosporidium Parvum in a postweaning diarrhoea focus in swine*. Semina: Ciências Agrárias, 1994. **15**(1): p. 5-7.
36. Haga, I.R., et al., *Identification of a bisegmented double-stranded RNA virus (Picobirnavirus) in faeces of giant anteaters (*Myrmecophaga tridactyla*)*. Vet J, 1999. **158**(3): p. 234-6.
37. Masachessi, G., et al., *Picobirnavirus (PBV) natural hosts in captivity and virus excretion pattern in infected animals*. Arch Virol, 2007. **152**(5): p. 989-98.
38. Navarro, J.O., et al., *Genetic diversity of bovine Picobirnavirus, Brazil*. Virus Genes, 2018.
39. Takiuchi, E., et al., *Electrophoretic RNA genomic profiles of Brazilian Picobirnavirus (PBV) strains and molecular characterization of a PBV isolated from diarrheic calf*. Virus Res, 2016. **211**: p. 58-63.
40. Shi, M., et al., *Redefining the invertebrate RNA virosphere*. Nature, 2016.
41. Adriaenssens, E.M., et al., *Viromic Analysis of Wastewater Input to a River Catchment Reveals a Diverse Assemblage of RNA Viruses*. mSystems, 2018. **3**(3).

42. Zhang, S., et al., *Detection and evolutionary analysis of picobirnaviruses in treated wastewater*. Microb Biotechnol, 2015. **8**(3): p. 474-82.
43. Malik, Y.S., et al., *Epidemiology, phylogeny, and evolution of emerging enteric Picobirnaviruses of animal origin and their relationship to human strains*. Biomed Res Int, 2014. **2014**: p. 780752.
44. Yinda, C.K., et al., *Cameroonian fruit bats harbor divergent viruses, including rotavirus H, bastroviruses, and picobirnaviruses using an alternative genetic code*. Virus Evol, 2018. **4**(1): p. vey008.
45. Gallimore, C.I., et al., *Detection and characterisation of bisegmented double-stranded RNA viruses (picobirnaviruses) in human fecal specimens*. Journal of Medical Virology, 1995. **45**(2): p. 135-140.
46. van Leeuwen, M., et al., *Human picobirnaviruses identified by molecular screening of diarrhea samples*. J Clin Microbiol, 2010. **48**(5): p. 1787-94.
47. Sun, G., et al., *Viral metagenomics analysis of picobirnavirus-positive feces from children with sporadic diarrhea in China*. Arch Virol, 2016. **161**(4): p. 971-5.
48. Zell, R., et al., *ICTV Virus Taxonomy Profile: Picornaviridae*. J Gen Virol, 2017. **98**(10): p. 2421-2422.
49. Lin, J.Y., et al., *Viral and host proteins involved in picornavirus life cycle*. J Biomed Sci, 2009. **16**: p. 103.
50. Ehrenfeld, E., E. Domingo, and R.P. Roos, *The picornaviruses*. 2010: American Society for Microbiology Press.
51. Arnold, E., et al., *Implications of the picornavirus capsid structure for polyprotein processing*. Proc Natl Acad Sci U S A, 1987. **84**(1): p. 21-5.
52. Jiang, P., et al., *Picornavirus morphogenesis*. Microbiol Mol Biol Rev, 2014. **78**(3): p. 418-37.
53. Lu, H.H., et al., *Analysis of picornavirus 2A(pro) proteins: separation of proteinase from translation and replication functions*. J Virol, 1995. **69**(12): p. 7445-52.
54. Hughes, P.J. and G. Stanway, *The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation*. J Gen Virol, 2000. **81**(Pt 1): p. 201-7.
55. Sun, D., et al., *Roles of the Picornaviral 3C Proteinase in the Viral Life Cycle and Host Cells*. Viruses, 2016. **8**(3): p. 82.

56. de Jong, A.S., et al., *Functional Analysis of Picornavirus 2B Proteins: Effects on Calcium Homeostasis and Intracellular Protein Trafficking*. Journal of Virology, 2008. **82**(7): p. 3782-3790.
57. Moffat, K., et al., *Inhibition of the secretory pathway by foot-and-mouth disease virus 2BC protein is reproduced by coexpression of 2B with 2C, and the site of inhibition is determined by the subcellular location of 2C*. J Virol, 2007. **81**(3): p. 1129-39.
58. Nieva, J.L., et al., *Mechanisms of membrane permeabilization by picornavirus 2B viroporin*. FEBS Lett, 2003. **552**(1): p. 68-73.
59. Cheng, Z., et al., *The nonstructural protein 2C of a Picorna-like virus displays nucleic acid helix destabilizing activity that can be functionally separated from its ATPase activity*. J Virol, 2013. **87**(9): p. 5205-18.
60. Hope, D.A., S.E. Diamond, and K. Kirkegaard, *Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB*. J Virol, 1997. **71**(12): p. 9490-8.
61. Wessels, E., et al., *Effects of picornavirus 3A Proteins on Protein Transport and GBF1-dependent COP-I recruitment*. J Virol, 2006. **80**(23): p. 11852-60.
62. Came, P.E., et al., *Chemotherapy of Viral Infections*. 2012: Springer Berlin Heidelberg.
63. Magnius, L., et al., *ICTV Virus Taxonomy Profile: Deltavirus*. Journal of General Virology, 2018. **99**(12): p. 1565-1566.
64. Ryu, W.S., et al., *Ribonucleoprotein complexes of hepatitis delta virus*. Journal of virology, 1993. **67**(6): p. 3281-3287.
65. Sureau, C. and F. Negro, *The hepatitis delta virus: Replication and pathogenesis*. J Hepatol, 2016. **64**(1 Suppl): p. S102-S116.
66. Taylor, J.M., *Hepatitis D Virus Replication*. Cold Spring Harb Perspect Med, 2015. **5**(11).
67. Tobler, K., M. Ackermann, and C. Fraefel, *Allgemeine Virologie*. 2016: UTB GmbH.
68. Selzer, L. and A. Zlotnick, *Assembly and Release of Hepatitis B Virus*. Cold Spring Harb Perspect Med, 2015. **5**(12).
69. Wille, M., et al., *A Divergent Hepatitis D-Like Agent in Birds*. Viruses, 2018. **10**(12).
70. Hetzel, U., et al., *Identification of a Novel Deltavirus in Boa Constrictors*. mBio, 2019. **10**(2).

71. Chang, W.S., et al., *Novel hepatitis D-like agents in vertebrates and invertebrates*. *Virus Evol*, 2019. **5**(2): p. vez021.
72. Kozlov, A.M., et al., *RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference*. *Bioinformatics*, 2019. **35**(21): p. 4453-4455.
73. Simmonds, P., et al., *ICTV Virus Taxonomy Profile: Flaviviridae*. *J Gen Virol*, 2017. **98**(1): p. 2-3.
74. Gawlik, K. and P.A. Galloway, *HCV core protein and virus assembly: what we know without structures*. *Immunol Res*, 2014. **60**(1): p. 1-10.
75. Moradpour, D., F. Penin, and C.M. Rice, *Replication of hepatitis C virus*. *Nat Rev Microbiol*, 2007. **5**(6): p. 453-63.
76. Spaeth, G.B., C.L. Murray, and C.M. Rice, *Hepacivirus*, in *The Springer Index of Viruses*, C. Tidona and G. Darai, Editors. 2011, Springer New York: New York, NY. p. 473-481.
77. Boukadida, C., et al., *NS2 proteases from hepatitis C virus and related hepaciviruses share composite active sites and previously unrecognized intrinsic proteolytic activities*. *PLoS Pathog*, 2018. **14**(2): p. e1006863.
78. Jennings, T.A., et al., *NS3 helicase from the hepatitis C virus can function as a monomer or oligomer depending on enzyme and substrate concentrations*. *J Biol Chem*, 2009. **284**(8): p. 4806-14.
79. Konan, K.V., et al., *Nonstructural protein precursor NS4A/B from hepatitis C virus alters function and ultrastructure of host secretory apparatus*. *J Virol*, 2003. **77**(14): p. 7843-55.
80. Blanchard, E. and P. Roingeard, *The Hepatitis C Virus-Induced Membranous Web in Liver Tissue*. *Cells*, 2018. **7**(11).
81. King, A.M., et al., *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*. 2011: Elsevier.
82. Organization, W.H., *Global hepatitis report 2017*. 2017, World Health Organization.
83. Feinstone, S.M., et al., *Transfusion-associated hepatitis not due to viral hepatitis type A or B*. *N Engl J Med*, 1975. **292**(15): p. 767-70.
84. Choo, Q.L., et al., *Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome*. *Science*, 1989. **244**(4902): p. 359-62.
85. Freeman, A.J., et al., *Estimating progression to cirrhosis in chronic hepatitis C virus infection*. *Hepatology*, 2001. **34**(4 Pt 1): p. 809-16.

86. Burbelo, P.D., et al., *Serology-enabled discovery of genetically diverse hepaciviruses in a new host*. J Virol, 2012. **86**(11): p. 6171-8.
87. Corman, V.M., et al., *Highly divergent hepaciviruses from African cattle*. J Virol, 2015. **89**(11): p. 5876-82.
88. Drexler, J.F., et al., *Evidence for novel hepaciviruses in rodents*. PLoS Pathog, 2013. **9**(6): p. e1003438.
89. Quan, P.L., et al., *Bats are a major natural reservoir for hepaciviruses and pegiviruses*. Proc Natl Acad Sci U S A, 2013. **110**(20): p. 8194-9.
90. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets*. Mol Biol Evol, 2016. **33**(7): p. 1870-4.
91. Suchard, M.A., et al., *Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10*. Virus Evolution, 2018. **4**(1).
92. Martin, D.P., et al., *RDP4: Detection and analysis of recombination patterns in virus genomes*. Virus Evol, 2015. **1**(1): p. vev003.
93. Bielejec, F., et al., *SpreaD3: Interactive Visualization of Spatiotemporal History and Trait Evolutionary Processes*. Mol Biol Evol, 2016. **33**(8): p. 2167-9.
94. Rambaut, A., et al., *Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen)*. Virus Evol, 2016. **2**(1): p. vew007.
95. Durricha, D., et al., *jModelTest 2: more models, new heuristics and parallel computing*. Nat Methods, 2012. **9**(8): p. 772.
96. Condit, R., et al., *The status of the Panama Canal watershed and its biodiversity at the beginning of the 21st century*. Bioscience, 2001. **51**(5): p. 389-398.
97. Rompré, G., W. Robinson, and A. Desrochers, *Causes of habitat loss in a Neotropical landscape: The Panama Canal corridor*. Landscape and Urban Planning, 2008. **87**: p. 129-139.
98. Adler, G.H., *Tropical Forest Fragmentation and Isolation Promote Asynchrony among Populations of a Frugivorous Rodent*. Journal of Animal Ecology, 1994. **63**(4): p. 903-911.
99. Schmid, J., et al., *Ecological drivers of Hepacivirus infection in a neotropical rodent inhabiting landscapes with various degrees of human environmental change*. Oecologia, 2018.
100. Paraskevopoulou, S., et al., *Mammalian deltavirus without hepadnavirus coinfection in the neotropical rodent Proechimys semispinosus*.

- Proceedings of the National Academy of Sciences of the United States of America, 2020. **117**(30): p. 17977-17983.
101. Goodacre, N., et al., *A Reference Viral Database (RVDB) To Enhance Bioinformatics Analysis of High-Throughput Sequencing for Novel Virus Detection*. mSphere, 2018. **3**(2).
  102. Buchfink, B., C. Xie, and D.H. Huson, *Fast and sensitive protein alignment using DIAMOND*. Nat Methods, 2015. **12**(1): p. 59-60.
  103. Cohen, J., *Statistical power analysis for the behavioral sciences*. 2013: Academic press.
  104. Gray, R.R., et al., *The mode and tempo of hepatitis C virus evolution within and among hosts*. BMC Evolutionary Biology, 2011. **11**(1): p. 131.
  105. Pollock, H.S., et al., *Long-term monitoring reveals widespread and severe declines of understory birds in a protected Neotropical forest*. Proc Natl Acad Sci U S A, 2022. **119**(16): p. e2108731119.
  106. Ibanez, R., et al., *An ecosystem report on the Panama Canal: monitoring the status of the forest communities and the watershed*. Environ Monit Assess, 2002. **80**(1): p. 65-95.

## 9. Erfolgte Publikationen

Paraskevopoulou S, **Pirzer F**, Goldmann N, Schmid J, Corman VM, Gottula LT, Schroeder S, Rasche A, Muth D, Drexler JF, Heni AC, Eibner GJ, Page RA, Jones TC, Müller MA, Sommer S, Glebe D, Drosten C.

Mammalian deltavirus without hepadnavirus coinfection in the neotropical rodent *Proechimys semispinosus*.

Proc Natl Acad Sci U S A. 2020 Jul 28;117(30):17977-17983. doi:

10.1073/pnas.2006750117. Epub 2020 Jul 10. PMID: 32651267; PMCID:

PMC7395443.