

Nachweis und Charakterisierung zirkulärer, einzelsträngiger  
BMMF-Elemente in milchwirtschaftlich genutzten Boviden

von Marie-Thérèse König

Inaugural-Dissertation zur Erlangung der Doktorwürde  
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität  
München

Nachweis und Charakterisierung zirkulärer, einzelsträngiger  
BMMF-Elemente in milchwirtschaftlich genutzten Boviden

von Marie-Thérèse König

aus Leipzig

München 2022



Aus dem Veterinärwissenschaftlichen Department der Tierärztlichen Fakultät  
der Ludwig-Maximilians-Universität München

Lehrstuhl für Hygiene und Technologie der Milch

Arbeit angefertigt unter der Leitung von:

Univ.-Prof. Dr. Dr. h.c. Erwin P. Märtlbauer

Mitbetreuung durch: Priv.-Doz. Dr. Andrea Stockmaier-Didier



Gedruckt mit Genehmigung der Tierärztlichen Fakultät  
der Ludwig-Maximilians-Universität München

Dekan: Univ.-Prof. Dr. Reinhard K. Straubinger, Ph.D.

Berichterstatter: Univ.-Prof. Dr. Dr. h.c. Erwin P. Märtlbauer

Korreferent/en: Univ.-Prof. Dr. Thomas W. Göbel  
Univ.-Prof. Dr. Eckhard Wolf  
Univ.-Prof. Dr. Rolf Mansfeld  
Prof. Dr. Monika Rinder

Tag der Promotion: 30. Juli 2022



*Meiner Familie  
für die liebevolle Unterstützung*

*„Überall geht ein frühes Ahnen dem späteren Wissen voraus.“*

Alexander von Humboldt



## INHALTSVERZEICHNIS

<b>I.</b>	<b>EINLEITUNG .....</b>	<b>1</b>
<b>II.</b>	<b>LITERATURÜBERSICHT .....</b>	<b>3</b>
1.	<b>Einzelstrang-DNA-Viren.....</b>	<b>3</b>
1.1.	Entdeckung und Bedeutung .....	3
1.2.	Der Stamm <i>Cressdnnaviricota</i> .....	3
1.2.1.	Taxonomie .....	3
1.2.2.	Molekulare Charakteristika.....	5
1.2.3.	Die Familie <i>Genomoviridae</i> .....	7
2.	<b>Bovine meat and milk factors .....</b>	<b>9</b>
2.1.	Hintergründe .....	9
2.2.	BMMF – Entdeckung und Gruppierung .....	10
2.3.	BMMF 1 und 2 – Molekulare Charakteristika.....	11
2.4.	Die „BMMF-Hypothese“ .....	12
3.	<b>Detektion zirkulärer Einzelstrang-DNA.....</b>	<b>13</b>
4.	<b>Systematik der <i>Bovidae</i>.....</b>	<b>15</b>
4.1.	Europäisches Hausrind ( <i>Bos primigenius</i> f. <i>taurus</i> ) .....	16
4.2.	Zebu ( <i>Bos primigenius</i> f. <i>indicus</i> ) .....	16
4.3.	Sanga-Rinder.....	17
4.4.	Wasserbüffel ( <i>Bubalus arnee</i> f. <i>bubalis</i> ) .....	18
4.5.	Yak ( <i>Bos mutus</i> f. <i>grunniens</i> ) .....	19
4.6.	Hausziege ( <i>Capra aegagrus</i> f. <i>hircus</i> ) .....	21
4.7.	Hausschaf ( <i>Ovis aries</i> ) .....	22
<b>III.</b>	<b>ZIELSETZUNG DER ARBEIT .....</b>	<b>25</b>
<b>IV.</b>	<b>PUBLIKATIONEN.....</b>	<b>27</b>
1.	<b>Publikation I: Circular Rep-Encoding Single-Stranded DNA Sequences in Milk from Water Buffaloes (<i>Bubalus arnee</i> f. <i>bubalis</i>).....</b>	<b>27</b>
2.	<b>Publikation II: Identification and Characterization of Circular Single-Stranded DNA Genomes in Sheep and Goat Milk .....</b>	<b>47</b>

<b>V.</b>	<b>BISHER UNVERÖFFENTLICHE DATEN – PROBEN</b>	
	<b>WEITERER RINDERSPEZIES .....</b>	<b>67</b>
<b>VI.</b>	<b>DISKUSSION.....</b>	<b>69</b>
1.	<b>BMMF in Probenmaterial unterschiedlicher Spezies .....</b>	<b>69</b>
1.1.	Milchproben von Wasserbüffeln, Schafen und Ziegen.....	69
1.2.	Blutproben weiterer exotischer Rinderspezies.....	74
2.	<b>Kritische Bewertung der eingesetzten Methoden .....</b>	<b>75</b>
2.1.	RCA .....	75
2.2.	PCR <i>versus</i> NGS.....	76
3.	<b>Abschließende Betrachtung und Ausblick.....</b>	<b>77</b>
<b>VII.</b>	<b>ZUSAMMENFASSUNG .....</b>	<b>79</b>
<b>VIII.</b>	<b>SUMMARY .....</b>	<b>81</b>
<b>IX.</b>	<b>LITERATURVERZEICHNIS.....</b>	<b>83</b>
<b>X.</b>	<b>EINSENDER DES PROBENMATERIALS.....</b>	<b>103</b>
1.	<b>Blutproben.....</b>	<b>103</b>
2.	<b>Milchproben .....</b>	<b>103</b>
<b>XI.</b>	<b>DANKSAGUNG.....</b>	<b>105</b>

## ABKÜRZUNGSVERZEICHNIS

BAMI	Bubalus arnee milk isolate
BfR	Bundesinstitut für Risikobewertung
BMMF	Bovine meat and milk factors
Cap	Capsid
CMI	Cow milk isolate
CRESS	Circular Rep-encoding single-stranded
DNA	Deoxyribonucleic acid, Desoxyribonukleinsäure
ds	doppelsträngig
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethyldiamintetraacetat
ELISA	Enzyme-linked Immunosorbent Assay
FAO	Food and Agriculture Organization of the United Nations
FL	Full-length, Volllänge
Fr	Fragments, Fragmente
Gm	Goat milk, Ziegenmilch
HCBI	Healthy cattle blood isolate
ICTV	International Committee on Taxonomy of Viruses
IR	Inverted repeat
MRI	Max Rubner-Institut
MSBI	Multiple sclerosis brain isolate
MSSI	Multiple sclerosis serum isolate
NCBI	National Center for Biotechnology Information
Neu5Gc	N-Glycolylneuraminsäure
NGS	Next-Generation Sequencing
nt	Nukleotide
ORF	Open reading frame, offener Leserahmen
ori	Origin of replication, Replikationsursprung
PAK	Polycyclische aromatische Kohlenwasserstoffe
PCR	Polymerase-Chain-Reaction, Polymerase-Kettenreaktion
RCA	Rolling circle amplification
RCR	Rolling circle replication
Rep	Replikase

## Abkürzungsverzeichnis

---

SF3	Superfamily 3, Superfamilie 3
Sm	Sheep milk, Schafmilch
SNP	Single nucleotide polymorphism
SPHINX	Slow Progressive Hidden INfections of variable (X) latency
ss	single-stranded, einzelsträngig
SsHADV	Sclerotinia sclerotiorum hypovirulence-associated DNA virus
TR	Tandem repeat
TSE	Transmissible spongiforme Enzephalopathie

## I. EINLEITUNG

Lebensmittel tierischer Herkunft wie Kuhmilch und Rindfleisch haben für die menschliche Ernährung eine elementare Bedeutung. Hinweise auf eine potenzielle Gesundheitsgefährdung durch diese Produkte sind daher im Kontext „Public Health“ von fundamentalem Interesse und entsprechende Berichte stoßen regelmäßig auf große mediale Resonanz. Somit ist es nicht überraschend, dass in einer Vielzahl von Studien ein Zusammenhang zwischen dem Milch- und Fleischkonsum und diversen humanen Erkrankungen postuliert wird (Melnik, 2009, Larsson and Orsini, 2013, Sargsyan and Dubasi, 2021).

Im Jahr 2012 veröffentlichte Harald zur Hausen auf Basis globaler epidemiologischer Studien zum Vorkommen von Dickdarmkrebs beim Menschen seine Hypothese über eine mögliche Korrelation zwischen dem Konsum von Produkten europäischer Rinder und der Entstehung von Darm- und Brustkrebs. Diese Theorie basiert vor allem darauf, dass Länder, in denen europäische Rinderrassen keine oder nur eine untergeordnete wirtschaftliche Rolle spielen, für diese Krebsarten besonders niedrige Inzidenzen in der Bevölkerung aufweisen. Folglich müsste es einen spezifisch bovinen, infektiösen Faktor geben, der nur in europäischen Rinderrassen taurinen Ursprungs, nicht aber in Yaks, Zebus oder Wasserbüffeln zu finden sei (zur Hausen, 2012, zur Hausen and de Villiers, 2016). In den darauffolgenden Jahren wurden bei der Untersuchung von Kuhmilch und Kuhseren tatsächlich neuartige DNA-Moleküle entdeckt, die in Aufbau und Struktur bakteriellen Plasmiden und einzelsträngigen/single-stranded (ss), zirkulären DNA-Viren ähneln (Whitley et al., 2014, zur Hausen and de Villiers, 2015). Da diese Elemente keiner der derzeit bekannten Erregerklassen zugeordnet werden konnten, wurden sie ihrer Herkunft nach als „bovine meat and milk factors“ (BMMF) bezeichnet (de Villiers et al., 2019). In weiteren Studien wurden das Vorkommen und die Replikationsfähigkeit von BMMF-DNA in humanen, krankheitsassoziierten Geweben nachgewiesen (Eilebrecht et al., 2018, Bund et al., 2021). Aufgrund dieser und weiterer Daten wird zurzeit eine mögliche Beteiligung von BMMF an der Entstehung verschiedener Karzinome und anderer humaner Erkrankungen diskutiert (de Villiers and zur Hausen, 2021).

Aus bovinen und humanen Proben wurden neben den bisher unbekannten BMMF zusätzlich kleine, zirkuläre Einzelstrang-DNA-Viren isoliert, die der Familie *Genomoviridae* zugeordnet werden konnten (zur Hausen et al., 2017). Diese und andere Familien des kürzlich etablierten Phylums *Cressdnnaviricota* (Krupovic et al., 2020) erfuhren in den vergangenen Jahren durch den Einsatz neuer Techniken starken Zuwachs (Simmonds et al., 2017). Vertreter der Familie *Genomoviridae* sind mit den verschiedensten eukaryotischen Spezies assoziiert (Varsani and Krupovic, 2021). Ob oder welche Bedeutung das Vorkommen dieser DNA-Viren in Rindern für die Tiere selbst und für den Menschen hat, bleibt noch zu klären.

Die in den letzten Jahren aufgestellte Hypothese, BMMF seien spezifische, ausschließlich in taurinen Rindern zu findende infektiöse Faktoren, die an der Entstehung verschiedener humaner Erkrankungen beteiligt sein sollen, lässt aber Fragen offen. Insbesondere ist nach wie vor unklar, ob diese Faktoren tatsächlich nicht auch in anderen lebensmittelliefernden Tieren wie beispielsweise Wasserbüffeln, Yaks und Schafen nachweisbar sind. Die vorliegende Arbeit mit den darin enthaltenen Publikationen widmet sich dieser Fragestellung und liefert somit erstmals einen grundlegenden Beitrag zum Vorkommen von BMMF in weiteren milchwirtschaftlich genutzten Boviden. Die erste Publikation befasst sich dabei mit der Untersuchung von Wasserbüffelmilch, wohingegen die zweite Veröffentlichung die Spezies Schaf und Ziege behandelt. Zusätzlich zu den bereits publizierten Daten wurden Proben weiterer „exotischer“ Rinder (Zebu, Yak, Watussi-Rind) untersucht. Die Ergebnisse dieser Analysen werden in Kapitel V erläutert.

## II. LITERATURÜBERSICHT

### 1. Einzelstrang-DNA-Viren

#### 1.1. Entdeckung und Bedeutung

Schon vor Millionen von Jahren existierten Viren auf der Erde und trugen im Laufe der Zeit bedeutend zur Evolutionsgeschichte aller Lebewesen bei. Ihre Fähigkeit zur Mutation, Transposition und Rekombination machte sie extrem anpassungsfähig und vielgestaltig, sodass Viren in der heutigen Zeit ubiquitär zu finden sind.

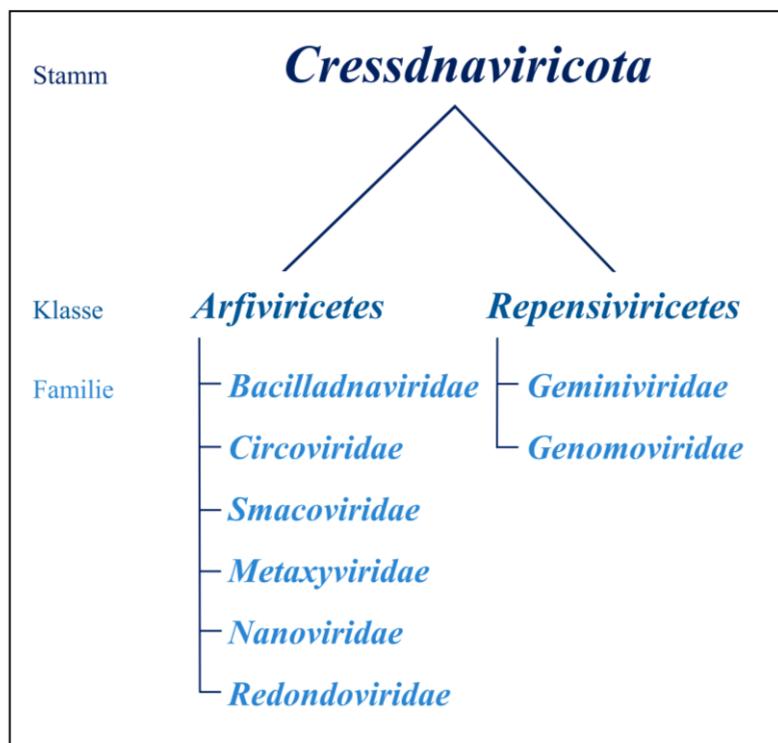
Obwohl wir inzwischen wissen, dass (ss)DNA-Viren einen Großteil der Gesamtheit aller Viren ausmachen und in allen drei Domänen – Archaeen, Bakterien, Eukaryoten – vertreten sind, spielten sie bis in die zweite Hälfte des 20. Jahrhunderts nur eine unwesentliche Rolle in der Wissenschaft (Zhao et al., 2019b). Mit der Entdeckung der pflanzenpathogenen Geminiviren (Goodman, 1977) und den tierpathogenen Circoviren (Tischer et al., 1982) gewannen die zirkulären ssDNA-Viren an Bedeutung. Seit Beginn der 2000er Jahre wurden durch die Anwendung neuer Techniken wie Next-Generation Sequencing sowie durch den Einsatz der Phi29-DNA-Polymerase unzählige, bisher unbekannte zirkuläre DNA-Elemente entdeckt. Die Sequenzen dieser Viren konnten aus einer immensen Fülle verschiedenster Proben wie z. B. Abwasser, Pflanzenmaterial, tierischem Kot, Insekten oder menschlichen Geweben isoliert werden. Innerhalb dieser Gruppe befinden sich sowohl krankheitsassoziierte als auch apathogene Viren (Zhao et al., 2019b).

#### 1.2. Der Stamm *Cressdnnaviricota*

##### 1.2.1. Taxonomie

Neben wenigen Vertretern mit linearem Genom wie z. B. den Parvoviren besitzen die meisten ssDNA-Viren kleine, zirkuläre Genome und infizieren eukaryotische Organismen. Der im Jahr 2019 neu etablierte Stamm „*Cressdnnaviricota*“ umfasst eine Gruppe recht heterogener, zirkulärer ssDNA-Viren, bei denen im Allgemeinen davon ausgegangen wird, dass sie über die sogenannte „rolling circle replication“ (RCR) replizieren. Der Name „*Cressdnnaviricota*“ ist das Portmanteau-Wort für „circular Rep-encoding single-stranded DNA“ (kurz CRESS-DNA) und dem

Suffix *-viricota* für die Stammeshierarchie innerhalb der taxonomischen Stufen und verweist damit auf die gemeinsamen molekularen Merkmale der Mitglieder dieser Gruppe. Nach aktuellen Angaben des International Committee on Taxonomy of Viruses (ICTV) gliedert sich der Stamm derzeit in zwei Klassen (*Arfiviricetes*, *Repensiviricetes*) mit insgesamt acht Familien, bestehend aus knapp 1000 Spezies (Abbildung 1).



**Abbildung 1:** Taxonomie innerhalb des Stammes *Cressdnnaviricota*, bestehend aus zwei Klassen mit insgesamt acht Familien (ICTV).

Drei dieser Familien, die *Nanoviridae*, *Metaxyviridae* und *Geminiviridae* infizieren ausschließlich pflanzliche Wirte und werden durch Vektoren (Insekten) übertragen. Durch das Auslösen verschiedener Krankheitssymptome bei landwirtschaftlich genutzten Pflanzen und den damit verbundenen Ernteverlusten haben sie vielerorts eine große ökonomische Bedeutung (Thomas et al., 2021, Zerbini et al., 2017, Gronenborn et al., 2021). Unter den Vertretern der Familie *Circoviridae* befinden sich bedeutende Tierpathogene wie z. B. die porcinen Circoviren 1 & 2 (PCV1, PCV2) oder das *beak and feather disease virus*, der Erreger der Psittacine Beak and Feather Disease (PBFD) (Rosario et al., 2017). Die Redondo- und Smacoviren sind tier- und humanassoziierte Familienmitglieder, die aus menschlichen bronchoalveolären Lavage-Proben (Redondoviren) oder aus dem Gastrointestinaltrakt

diverser Wirbeltiere (Smacoviren) isoliert werden konnten (Abbas et al., 2021, Varsani and Krupovic, 2021). Die ursprünglichen Wirte der Spezies dieser zwei Familien konnten bislang allerdings nicht eindeutig zugewiesen werden. Neben Eukaryoten werden bei den Smacoviren derzeit auch Archaeen als potenzielle Wirte diskutiert (Díez-Villaseñor and Rodriguez-Valera, 2019). Die Wirte der verbleibenden beiden Familien umfassen bei den wenigen Spezies der Familie *Bacilladnaviridae* verschiedene Algenarten und bei denen der *Genomoviridae* ein deutlich umfangreicheres Spektrum von Pilzen und Pflanzen bis hin zu Säugern und Insekten. Eine detaillierte Beschreibung der Familie *Genomoviridae* folgt in Abschnitt 1.2.3.

Aufgrund der kontinuierlich wachsenden Zahl neuer Viren und der nicht abschließend festgelegten Phylogenie der CRESS-DNA-Viren unterliegt die Taxonomie allerdings einem steten Wandel.

### 1.2.2. Molekulare Charakteristika

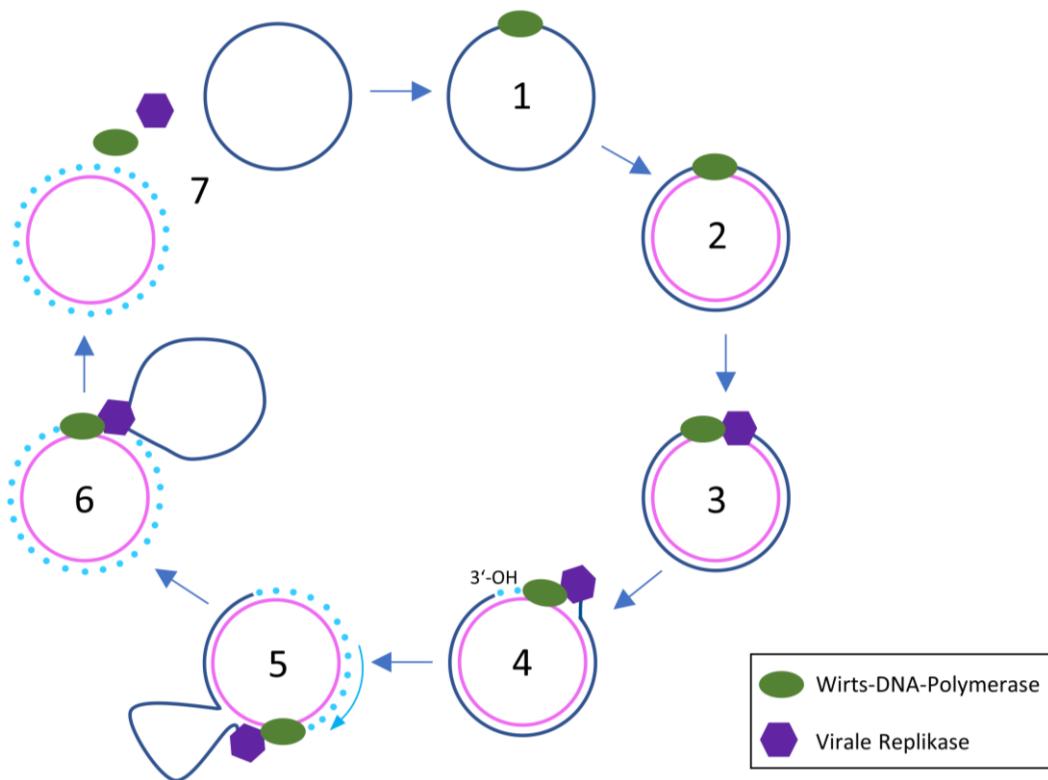
Das Genom der ubiquitären CRESS-DNA-Viren besteht, wie der Name bereits andeutet, aus zirkulärer, einzelsträngiger DNA. Einzelstrang-DNA-Viren zeigen eine den RNA-Viren ähnlich hohe Rekombinationsrate (Duffy et al., 2008, Lefevre et al., 2009), weshalb sie sich sehr schnell weiterentwickeln, ihr Wirtsspektrum ausdehnen und ihre Pathogenität verändern können (Zhou et al., 1997, Padidam et al., 1999). Diese Fähigkeit erklärt die hohe Diversität innerhalb der Familie *Cressnaviricota* und die Schwierigkeit der taxonomischen Klassifizierung.

Dennoch zeigen alle CRESS-DNA-Viren untereinander Ähnlichkeiten in ihrer Genomstruktur, die auf eine gemeinsame Phylogenie hindeuten. Es wird angenommen, dass das virale Genom einheitlich bei allen Vertretern dieser Gruppe durch einen bestimmten Mechanismus, die RCR repliziert wird (Abbildung 2). Dieser Mechanismus, der auch von bakteriellen Plasmiden, Phagen und anderen ssDNA-Viren bekannt ist (del Solar et al., 1998, Stenlund, 2003), wird durch eine Replikase (Rep) initiiert, die auf dem viralen DNA-Einzelstrang kodiert ist. Dieses hoch konservierte Gen dient der Klassifizierung der CRESS-DNA-Viren, da es spezifische Charakteristika aufweist, die Hinweise auf evolutionäre Ereignisse liefern können (Krupovic et al., 2020).

Die Replikase besteht aus zwei funktionellen Domänen, die für ihre enzymatische

Aktivität von entscheidender Bedeutung sind. Die Cofaktoren-abhängige HUH-Endonuklease-Domäne (H = Histidin, U = hydrophober Rest) befindet sich am aminoterminalen Ende des Proteins und ist in der Lage, sogenannte Haarnadelstrukturen (stem-loops) zu erkennen und innerhalb dieser oder in deren Nähe einen Strangbruch zu erzeugen. Während die Replikase dabei eine kovalente Bindung mit dem 5'-Ende des (+) Stranges eingeht, beginnt am freien 3'-OH-Ende die Replikation durch eine DNA-Polymerase des Wirtes. Am Ende eines jeden Replikationszyklus spaltet die HUH-Endonuklease den ursprünglichen (+) Strang vom neu entstandenen Doppelstrang ab und katalysiert abschließend seine Rezyklierung (Chandler et al., 2013) (Abbildung 2). Charakteristisch für die HUH-Endonuklease-Domäne sind die stark konservierten RCR-Motive I, II und III, die sowohl bei der RCR-Initiation als auch der -Termination eine zentrale Rolle spielen. Die zweite funktionelle Domäne, die carboxyterminale „Superfamilie 3 (SF3) Helikase-Domäne“ deutet auf eine nahe Verwandtschaft zu eukaryotischen Viren hin, da dieser Helikase-Typ nicht bei Viren zu finden ist, die Bakterien oder Archaeen infizieren (Krupovic et al., 2020). Sie entwindet die komplexe dreidimensionale Struktur des DNA-Stranges und bereitet sie somit für weitere Schritte wie z. B. die Replikation vor. Die hoch konservierten Motive Walker A, B und C lassen ebenfalls auf eine gemeinsame Abstammung aller CRESS-DNA-Viren schließen (Gorbalenya et al., 1990, Gorbalenya and Koonin, 1993).

Neben der Replikase kodieren alle Viren des Stammes *Cressdnaviricota* auch ein Capsid-Protein (Cap), das der Verpackung des Genoms dient. Die Capsid-kodierenden Gene zeigen im Vergleich zu den *rep*-Genen deutlich geringere Homologien untereinander und sind daher für die Klassifizierung der Isolate weniger geeignet. Allerdings lassen sich anhand der Capside ebenfalls mögliche Rückschlüsse auf die Phylogenie dieser Viren ziehen. So gibt es Hinweise darauf, dass die *cap*-Gene einiger CRESS-DNA-Viren durch RNA-DNA-Rekombination von RNA-Viren erworben wurden, da die Capside teils starke Ähnlichkeiten aufweisen (Krupovic et al., 2009, Diemer and Stedman, 2012).

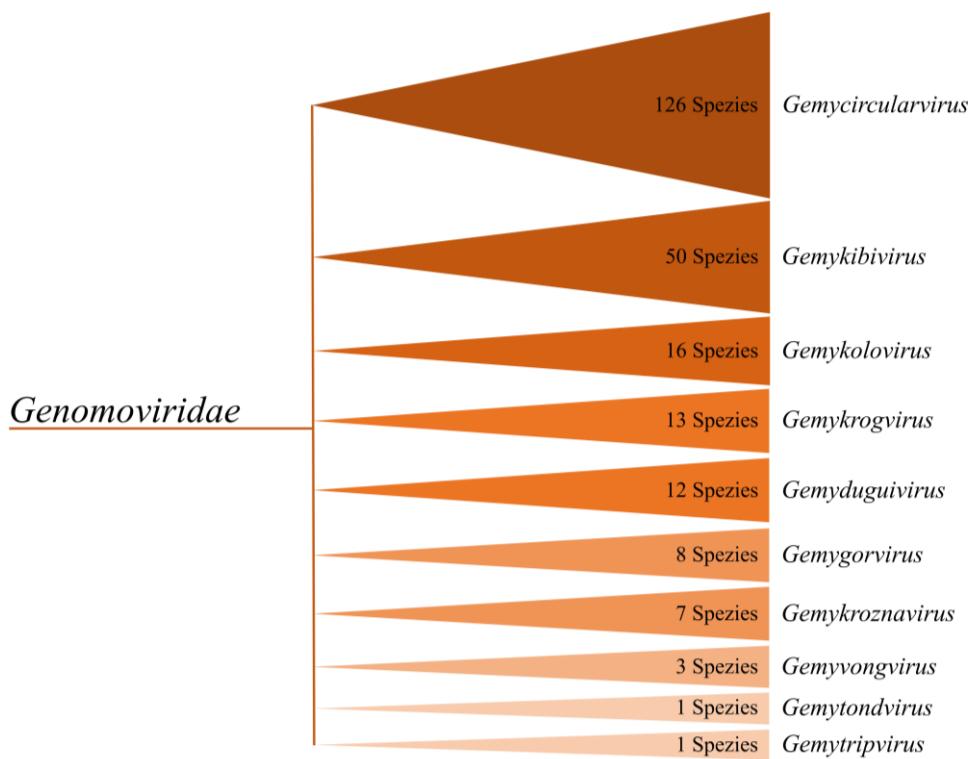


**Abbildung 2:** Schematische Darstellung der rolling circle replication. Nach Eintritt in die Wirtszelle wird die virale Einzelstrang-DNA (dunkelblau) von einer Wirts-DNA-Polymerase gebunden (1) und entlang der Matrize ein komplementärer Strang synthetisiert (rosa), sodass eine doppelsträngige (ds) Replikationsform der DNA entsteht (2). Durch Transkription und Translation entsteht die für die Replikation notwendige virale Replikase, die nun mit der Wirts-DNA-Polymerase einen Enzymkomplex bildet (3). Die Replikase spaltet die dsDNA innerhalb der Haarnadelstruktur, sodass das 3'-OH-Ende des (+) Stranges frei wird. An diesem freien Ende beginnt die Polymerase mit der Synthese eines neuen (+) Stranges (hellblau gepunktete Linie) in 5'-3'-Richtung (4), wobei der ursprüngliche Strang verdrängt wird (5). Am Ende eines jeden Zyklus katalysiert die Replikase die Ligation des ursprünglichen Stranges (6), wodurch dieser von der dsDNA abgespalten wird (7). Neu entstandene virale ssDNA wird entweder in einem Capsid verpackt oder geht in einen weiteren Replikationszyklus ein.

### 1.2.3. Die Familie *Genomoviridae*

Der erste Vertreter der Familie *Genomoviridae* war das infektiöse *Sclerotinia sclerotiorum hypovirulence-associated DNA virus 1* (SsHADV-1), welches aus dem Pilz *Sclerotinia sclerotiorum* isoliert wurde (Yu et al., 2010). Nach Genom- und Strukturanalysen stellte sich heraus, dass dieses Virus zwar einige Ähnlichkeiten zu Mitgliedern der bereits etablierten Familie *Geminiviridae* aufwies, jedoch in seiner Gesamtheit einzigartig war, sodass *Sclerotinia gemycircularvirus 1* als erste Spezies der Gattung *Gemycircularvirus* (Gemini-like myco-infecting circular virus) klassifiziert wurde (Rosario et al., 2012a, Krupovic et al., 2016). Seither

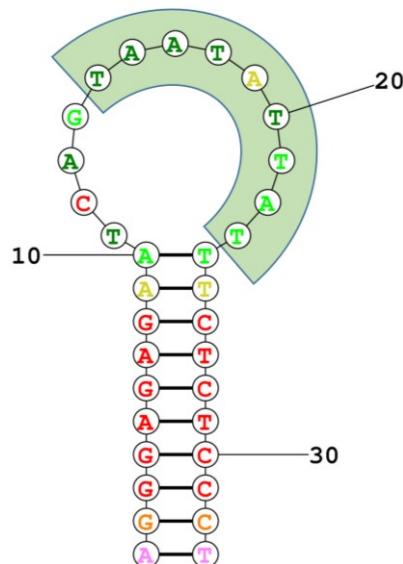
konnten unzählige weitere Genomoviren aus unterschiedlichsten Proben isoliert werden, u. a. aus Abwasser, Flusssedimenten, Pflanzen, Insekten und Säugetieren. Zurzeit besteht die Familie aus 10 Gattungen mit insgesamt 237 Spezies (Abbildung 3), die allesamt kleine, zirkuläre Einzelstrang-Genome einer Größenordnung von ca. 1,8–2,4 Kilobasen besitzen (Varsani and Krupovic, 2021).



**Abbildung 3:** Graphische Darstellung der Familie *Genomoviridae* mit ihren 10 Gattungen nach der derzeit aktuellen ICTV Einteilung (ICTV).

Genomoviren kodieren wie alle CRESS-DNA-Viren mindestens zwei Proteine – ein Rep- und ein Cap-Protein. Typisch für die Mitglieder dieser Familie ist, dass *rep*- und *cap*-Gen in gegenläufiger Richtung auf dem Einzelstrang liegen. In dem Sequenzabschnitt zwischen *rep* und *cap* lässt sich bei allen Mitgliedern die stark konservierte, aber variable „Nonanukleotidstruktur“ finden. Sie wird von einer Palindromsequenz eingeschlossen, die eine Art Stiel oder Stamm (stem) bildet. Das Nonanukleotid formt eine Schleife (loop) auf der Spitze des Stammes. In ihrer Gesamtheit erinnert diese stem-loop-Struktur an eine Haarnadel und stellt ein wichtiges Erkennungsmerkmal für die Replikase dar (Abbildung 4). Dieses Enzym spaltet zwischen Position 7 und 8 der Nonanukleotidstruktur die (ds)DNA (Varsani and Krupovic, 2017).

Während das *cap*-Gen der Genomoviren sehr heterogen ist, weisen ihre Replikasen starke Homologien untereinander auf. Eine wichtige Besonderheit dabei ist, dass viele genomovirale *reps* Introns beinhalten, die dieses Gen durch einen nicht kodierenden Bereich in zwei Teile trennen. Intron-tragende *reps* sind auch bei einigen Vertretern der Familie *Geminiviridae* zu finden (Wright et al., 1997, Varsani et al., 2017), was die nahe Verwandtschaft dieser beiden Familien unterstreicht (Zhao et al., 2019a). Eine weitere interessante Gemeinsamkeit ist ein zusätzliches, stark konserviertes Aminosäuremotiv innerhalb der Endonuklease-Domäne. Dieses „GRS-Motiv“, das sich zwischen den RCR-Motiven II und III befindet, spielt möglicherweise eine Rolle bei der räumlichen Ausrichtung dieser beiden Motive (Nash et al., 2011, Varsani and Krupovic, 2017).



**Abbildung 4:** Beispielhafte Darstellung einer stem-loop-/Haarnadelstruktur der Genomoviren. Die neun Nukleotide umfassende Nonanukleotidstruktur ist grün unterlegt und befindet sich innerhalb des loops. Das Palindrom, bestehend aus jeweils 10 Nukleotiden, befindet sich beiderseits der Nonanukleotidstruktur und bildet dabei den doppelsträngigen Stamm der Haarnadel. (<https://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html>)

## 2. Bovine meat and milk factors

### 2.1. Hintergründe

Seit einigen Jahren wird ein Zusammenhang zwischen dem Langzeitkonsum von rotem Fleisch (u. a. Rind, Schwein, Lamm, Pferd) und verschiedenen humanen Erkrankungen vermutet (Huxley et al., 2009, Richi et al., 2015). Schon in den 1970er Jahren wurde die mutagene und kanzerogene Wirkung einiger

heterozyklischer Amine, die zur Gruppe der polycyclischen aromatischen Kohlenwasserstoffe (PAK) gehören und bei der Zubereitung von Fleisch entstehen, diskutiert und auf mögliche Risiken für die Gesundheit hingewiesen (Sugimura et al., 1977, Nagao et al., 1977). Epidemiologische Studien indizieren eine Korrelation zwischen dem häufigen Verzehr von zubereitetem roten Fleisch und dem Auftreten verschiedener Erkrankungen wie Krebs, Diabetes mellitus Typ 2 oder kardiovaskuläre Erkrankungen (Chan et al., 2011, Pan et al., 2011, Pan et al., 2012, Farvid et al., 2018). Im Jahr 2015 wurde eine weitere Studie veröffentlicht, die erstmalig den Zusammenhang zwischen der Aufnahme von N-Glycolylneuraminsäure (Neu5Gc) und einem erhöhten Krebsrisiko darlegt. Neu5Gc, das vom menschlichen Organismus nicht gebildet werden kann, wird mutmaßlich mit der Nahrung, insbesondere rotem Fleisch, vom Körper aufgenommen und führt im Zielgewebe zu chronisch entzündlichen Veränderungen (Samraj et al., 2015). Daten zu globalen Dickdarmkrebsinzidenzen und dem weltweiten Rindfleischkonsum führten Harald zur Hausen zu der Annahme, dass möglicherweise infektiöse Agenzien in Rindern ursächlich für Dickdarmkrebs sein könnten (zur Hausen, 2012). Seit geraumer Zeit steht neben dem Konsum von rotem Fleisch auch der Milchkonsum in Verdacht, an der Entstehung von neurodegenerativen Erkrankungen wie Multipler Sklerose oder Morbus Parkinson beteiligt zu sein (Winer et al., 2001, Jiang et al., 2014). Laura Manuelidis entdeckte im Jahr 2011 kleine, neuartige, zirkuläre DNA-Moleküle, sogenannte Sphinx-DNA (Slow Progressive Hidden INfections of variable (X) latency) in Geweben von Säugetieren, die an transmissibler spongiformer Enzephalopathie (TSE) erkrankt waren und vermutete deren potenzielle Beteiligung an der TSE-Pathogenese, wodurch sie die „Protein-only-Hypothese“ von Stanley Prusiner (Prusiner, 1982) in Frage stellte. Die als Sphinx 1.76 und Sphinx 2.36 benannten DNA-Moleküle weisen Ähnlichkeiten zu Plasmiden von verschiedenen *Acinetobacter* spp. auf, stellen jedoch eine neue Einheit dar (Manuelidis, 2011). Dies gab Anlass zur Untersuchung humaner und boviner Proben auf bisher unbekannte Pathogene, um die Ätiologie diverser Erkrankungen näher zu beleuchten.

## 2.2. BMMF – Entdeckung und Gruppierung

Die Arbeitsgruppe um Harald zur Hausen entdeckte in Rinderseren und -milch, in menschlichen Blut- und Serumproben sowie in Hirngewebe zirkuläre (ss)DNA-Moleküle einer Größenordnung von ca. 1000-3000 Nukleotiden (nt) (Funk et al.,

2014, Whitley et al., 2014, Gunst et al., 2014, Lamberto et al., 2014). Diese unter dem Oberbegriff „BMMF“ (bovine meat and milk factors) zusammengefassten episomalen Isolate wurden in Abhängigkeit von ihren Sequenzhomologien zu bereits bekannten Organismen in vier Gruppen unterteilt. Die ersten beiden Gruppen umfassen Isolate, die Ähnlichkeiten zu den von Manuelidis entdeckten Sphinx 1.76- und Sphinx 2.36-Elementen aufweisen. Die Gruppe 3 beinhaltet solche, die der Familie *Genomoviridae* angehören. DNA-Moleküle, die Homologien zu Plasmiden von *Psychrobacter* spp. zeigen, werden der Gruppe 4 zugeordnet (de Villiers et al., 2019). Die Nukleotid-Sequenzen der Isolate sind in der Literatur unter verschiedenen Namen zu finden, je nachdem, aus welcher Probe die jeweilige DNA isoliert wurde (Tabelle 1). Unter diesen noch nicht klassifizierten, neuartigen DNA-Elementen besitzen die BMMF-Gruppen 1 und 2 quantitativ die größte Bedeutung.

**Tabelle 1:** Namensgebung der verschiedenen BMMF-Isolate und ihre Abkürzungen.

Cow milk isolate	CMI
Healthy cattle blood isolate	HCBI
Multiple sclerosis brain isolate	MSBI
Multiple sclerosis serum isolate	MSSI

### 2.3. BMMF 1 und 2 – Molekulare Charakteristika

Alle Gruppe 1-Isolate, die sowohl Ähnlichkeiten zu Sphinx 1.76 als auch zu Plasmiden von *Acinetobacter* spp. aufweisen, zeigen eine für diese Gruppe typische Genomstruktur mit charakteristischen Merkmalen. Sie besitzen mindestens einen offenen Leserahmen – englisch: open reading frame (ORF) –, der eine Replikase kodiert. Die weiteren ORFs, die im Rahmen von *in silico*-Analysen vorhergesagt wurden, sind rein hypothetisch. Ihre Expression und eine Funktion konnten bislang nicht gezeigt bzw. bestimmt werden. Ein typisches Strukturmerkmal für alle Vertreter dieser Gruppe ist der sogenannte tandem repeat (TR), der 5'-wärts dem *rep*-Gen vorangestellt ist. Er besteht aus einer Nukleotidabfolge von 22 nt, die sich mindestens dreimal vollständig und ein vieres unvollständiges Mal wiederholt. Der TR ist eine hoch konservierte Region innerhalb des Genoms, die bei den Isolaten nur geringfügig variiert (zur Hausen et al., 2017). Möglicherweise stellt er eine Erkennungsstruktur für die Replikase dar (Chattoraj, 2000). Dem TR wiederum

vorangestellt ist eine palindromische Sequenz von 12 nt Länge, der sogenannte inverted repeat (IR). Auch dieser ist bei den verschiedenen BMMF 1-Isolaten weitestgehend identisch und dient vermutlich der Replikase als Bindungsstelle (zur Hausen et al., 2017).

DNA-Moleküle, die der Gruppe BMMF 2 zugeordnet sind und Ähnlichkeit zu Sphinx 2.36 aufweisen, unterscheiden sich deutlich von denen der Gruppe 1, da sie weder einen TR noch den BMMF 1-typischen IR besitzen. Auch BMMF 2-Elemente kodieren eine potenzielle Replikase, deren Aminosäuresequenzen starke Ähnlichkeiten zu Rep-Proteinen von *Acinetobacter* spp. zeigen (zur Hausen et al., 2017).

#### 2.4. Die „BMMF-Hypothese“

Die unter 2.1. beschriebenen Daten weisen darauf hin, dass die wiederholte Aufnahme von PAK mit dem Verzehr von zubereitetem roten Fleisch ein Risiko in Bezug auf die Entstehung von Kolonkarzinomen darstellt. Allerdings zeigen einige Publikationen auch, dass chemische Kanzerogene nicht nur bei der Zubereitung von rotem Fleisch entstehen, sondern auch beim Braten von Fisch und Huhn, dessen Langzeitkonsum jedoch nicht mit einer erhöhten Dickdarmkrebsinzidenz korreliert (Yano et al., 1988, Kazerouni et al., 2001, Huxley et al., 2009).

Epidemiologische Studien zum weltweiten Vorkommen verschiedener Krebsarten beim Menschen zeigen, dass Regionen wie Indien, Äquatorialafrika und die Mongolei im Vergleich zu europäischen und nordamerikanischen Ländern besonders niedrige Dickdarmkrebsinzidenzen aufweisen (Ferlay et al., 2020, zur Hausen and de Villiers, 2015, zur Hausen and de Villiers, 2016). All diesen Ländern ist gemein, dass dort großteils andere Rinderarten als das europäische Hausrind genutzt werden wie Yak, Zebu, Watussi-Rind oder Wasserbüffel (Grigson, 1991, Leslie and Schaller, 2009, Government of India, 2015, Zhang et al., 2020). Ähnlich wie die weltweiten Inzidenzen für Dickdarmkrebs verhalten sich auch die Brustkrebsinzidenzen, obwohl diese im Allgemeinen etwas höher liegen (Sung et al., 2021). Diese Beobachtungen führten Harald zur Hausen zu der Hypothese, dass die scheinbar ausschließlich in europäischen Hausrindern vorkommenden, infektiösen BMMF eine entscheidende Rolle bei der Entstehung von humanen Krankheitsbildern wie z. B. Brust- und Dickdarmkrebs spielen (zur Hausen and de Villiers, 2015). Insbesondere die Aufnahme von BMMF- und Neu5Gc-haltigen

Nahrungsmitteln im Kindesalter resultiere hierbei in einer lebenslang manifesten Infektion von Darmepithelzellen. Im weiteren Verlauf komme es zur intrazellulären Antigenexpression und zu einer chronischen Inflammation des umliegenden Gewebes (Bund et al., 2021). Die jahrzehntelang persistierende Entzündung und die wiederholte Exposition gegenüber BMMF und anderen Kanzerogenen führe schlussendlich zu Mutationen und damit zur indirekten Karzinogenese (zur Hausen et al., 2019, de Villiers and zur Hausen, 2021).

### 3. Detektion zirkulärer Einzelstrang-DNA

Der Einsatz neuer Techniken vor etwa zwei Jahrzehnten war für die Entdeckung und Erforschung zirkulärer, einzelsträngiger DNA-Elemente ein Meilenstein in der Forschung. Niedrige Abundanzen, die Abhängigkeit von spezifischen Primern und – im Fall von ssDNA-Viren – die mangelnde Kultivierbarkeit der Erreger erschwerten bis dahin die Detektion dieser DNA-Moleküle (Rector et al., 2004). Mit Hilfe der Phi29-DNA-Polymerase und dem Einsatz von Random-Hexamer-Primern gelang mit der sogenannten „rolling circle amplification“ (RCA) die bis zu 10.000-fache Amplifikation von vollständigen, zirkulären Genomen. In der Folge wurden unzählige neue ssDNA-Viren detektiert (Dean et al., 2001, Rosario et al., 2012b). Die Phi29-basierte Prä-Amplifikation der Zielsequenzen ermöglichte die Erhöhung der Sensitivität verschiedener nachgelagerter Detektionsmethoden, die im Folgenden chronologisch aufgeführt werden.

Um geringe Mengen vollkommen unbekannter DNA in Probenmaterial nachzuweisen, nutzte Laura Manuelidis 2011 die RCA als einen zusätzlichen Vorbereitungsschritt der DNA. Das erhaltene Material wurde mit verschiedenen Restriktionsenzymen verdaut, um die entstandenen Phi29-Polymerase-Produkte zu linearisieren. Die auf ein Gel aufgetragenen, verdauten Proben wurden mit unterschiedlichen Kontrollansätzen verglichen und verdächtige Banden aufgereinigt, kloniert und sequenziert. Mit diesem Vorgehen gelang die Detektion der sogenannten Sphinx-DNA aus TSE-Gewebe (Manuelidis, 2011).

Viele Sphinx-ähnliche DNA-Elemente wurden in den darauffolgenden Jahren durch das gleiche methodische Vorgehen aus weiteren Proben isoliert, sodass die Gruppe der BMMF etabliert wurde. Durch Sequenzvergleiche dieser neuartigen BMMF-Genome auf Nukleotidebene konnten stark konservierte, homologe Bereiche identifiziert und spezifische Primer entworfen werden, die die

Volllängenamplifikation der Sequenzen mittels Polymerase-Kettenreaktion (PCR) ermöglichen (de Villiers et al., 2019).

Eine weitere Technik zur Detektion zirkulärer ssDNA stellt die sogenannte „Next-Generation Sequencing-Technologie“ (NGS) dar. Hierbei können simultan Millionen verschiedener DNA-Fragmente sequenziert und anschließend direkt analysiert werden. Mit Hilfe dieser Technik wurde die Nukleinsäureanalytik revolutioniert und die Entdeckung unzähliger neuer Viren in kürzester Zeit ermöglicht. Aufgrund der gigantischen Fülle unterschiedlicher DNA-Sequenzen im Probenmaterial besteht jedoch die Gefahr, dass niedrig abundante DNA bei der Sequenzierung unentdeckt bleibt. Um die Nachweisempfindlichkeit gegenüber dieser DNA zu verbessern, können die Proben zuvor mittels RCA vorbereitet werden (Barzon et al., 2011, Aimone et al., 2022).

In der vorliegenden Arbeit wurde der Ansatz der primären DNA-Vervielfältigung mittels RCA und anschließender Detektion mittels PCR gewählt, da eine NGS-Plattform nicht zur Verfügung stand.

#### 4. Systematik der *Bovidae*

Die äußerst vielfältige Familie der Hornträger (*Bovidae*) gehört zur Ordnung der Paarhufer (*Artiodactyla*) und zur Unterordnung der Wiederkäuer (*Ruminantia*). Eine der beiden Unterfamilien, die *Antilopinae* gliedert sich in neun verschiedene Tribus, wohingegen die Unterfamilie *Bovinae* nur drei Tribus umfasst (Groves and Grubb, 2011). Von den diversen Gattungen sollen nachfolgend nur ausgewählte Spezies der Gattungen *Bos*, *Bubalus*, *Capra* und *Ovis* näher erläutert werden (Abbildung 5), da diese besonders in Regionen mit niedrigen Dickdarmkrebsinzidenzen für die Milch- und/oder Fleischgewinnung gehalten werden und in Bezug auf die „BMMF-Hypothese“ von Interesse sind.

Familie	Unterfamilie	Tribus	Gattung	Spezies
<i>Bovidae</i>	<i>Bovinae</i>	<i>Bovini</i>	<i>Bos</i>	<i>Bos taurus</i> – Hausrind 
				<i>Bos indicus</i> – Zebu 
				<i>Bos grunniens</i> – Hausyak 
				<i>Bos bonasus</i> – Wisent
				<i>Bos gaurus</i> – Gaur
				<i>Bos javanicus</i> – Banteng
				<i>Bison bison</i> – Amerikanischer Bison
		<i>Bubalus</i>	<i>Bubalus bubalis</i>	– Hausbüffel 
			<i>Bubalus depressicornis</i>	– Tiefland-Anoa
			<i>Bubalus quarlesi</i>	– Berg-Anoa
<i>Antilopinae</i>	<i>Caprini</i>	<i>Capra</i>	<i>Capra hircus</i>	– Hausziege 
			<i>Capra ibex</i>	– Alpensteinbock
			<i>Capra falconeri</i>	– Schraubenziege
		<i>Ovis</i>	<i>Ovis aries</i>	– Hausschaf 
			<i>Ovis nivicola</i>	– Schneeschaf
			<i>Ovis canadensis</i>	– Dickhornschaf

**Abbildung 5:** Vereinfachte Darstellung der inneren Systematik der Familie *Bovidae* mit zwei Unterfamilien und ausgewählten Gattungen und Spezies. Diese Abbildung soll zu einem besseren Verständnis der taxonomischen Eingruppierung der in dieser Arbeit thematisierten Spezies beitragen. Sie erhebt keinen Anspruch auf Vollständigkeit, insbesondere in Bezug auf die Auflistung domestizierter Spezies und ihrer Wildformen.

#### 4.1. Europäisches Hausrind (*Bos primigenius f. taurus*)

Das domestizierte europäische Hausrind (*Bos primigenius f. taurus*) stammt vom bereits ausgestorbenen Auerochsen (*Bos primigenius primigenius*) ab. Heute wird angenommen, dass es während des Neolithikums zur unabhängigen Domestikation von zwei verschiedenen Auerochsenpopulationen kam, aus denen sich die heutigen taurinen und zebuinen Rinder entwickelten. Aus Auerochsen im Nahen Osten gingen vor rund 10.500 Jahren die taurinen, oder buckellosen Rinder hervor (Bradley et al., 1996, MacHugh et al., 1997, Helmer et al., 2005). Diese frühen taurinen Rinder breiteten sich Richtung Südwesten und Westen aus und etablierten sich im europäischen Raum. Im Laufe ihrer Domestikation kam es zur Hybridisierung domestizierter Rinder und wilder europäischer Auerochsen, was vermutlich zu einer besseren Anpassung an regionale Umweltbedingungen und zur Diversität der heutigen Rinderrassen beitrug (Schibler et al., 2014, Cubric-Curik et al., 2021). Heute leben taurine Rinder in unzähligen Ländern weltweit und dienen der Erzeugung von Milch, Fleisch und Tierhäuten. Die gezielte Selektion auf ausgewählte Merkmale brachte regional verschiedenste Phänotypen hervor (Hiemstra et al., 2010).

#### 4.2. Zebu (*Bos primigenius f. indicus*)

Namensgebend für die Zebus, oder auch Buckelrinder, ist der auf dem Widerrist befindliche Buckel, der aus dem vergrößerten *Musculus rhomboideus* besteht (Abbildung 6). Zebus entstanden durch Domestikation wilder Auerochsen (*Bos primigenius namadicus*) im Nordwesten des indischen Subkontinents vor ca. 9000 Jahren (Ajmone-Marsan et al., 2010). Entlang der menschlichen Wander- und Handelsrouten breiteten sich die Zebus nach Afrika, China und Südostasien aus. Im 19. und 20. Jh. wurden einige Zeburassen zunächst nach Süd-, später nach Nordamerika eingeführt und dort mit taurinen Rassen gekreuzt (Mukasa-Mugerwa, 1989, Ajmone-Marsan et al., 2010, Felius et al., 2014). Außerhalb Indiens und Pakistans, wo der Großteil der gehaltenen Rinder reinrassige Zebus sind, ist der genetische Zebueintrag in Rinderpopulationen Brasiliens, Chinas, Südostasiens, Zentral- und Ostafrikas besonders hoch (Decker et al., 2014, Manomohan et al., 2021).

In Indien allein werden ca. 37 verschiedene Zeburassen wie etwa Gir, Haryana und Ongole als Arbeitstiere und zur Milchgewinnung gehalten (Government of India, 2020). Zebus und Zebukreuzungen spielen als Fleisch- und Milchlieferanten auch

in vielen süd- und nordamerikanischen Ländern sowie in Teilen Afrikas eine bedeutende Rolle (Cunha et al., 1963, Haile et al., 2011, Menezes et al., 2014).



**Abbildung 6:** Fotografie indischer Buckelrinder/Zebus (*Bos primigenius f. indicus*). Die Fotos der Abbildungen 6-11 können lizenziert von der folgenden Homepage genutzt werden:  
<https://pixabay.com/de/>

#### 4.3. Sanga-Rinder

Die Entstehungsgeschichte der afrikanischen Rinder ist nicht eindeutig geklärt. Zwei mögliche Theorien über ihre Entstehung werden in der Literatur diskutiert. Ursprünglich wurde davon ausgegangen, dass es neben den bereits beschriebenen Domestikationsereignissen im Nahen Osten und auf dem indischen Subkontinent zu einer dritten, unabhängigen Domestikation der Auerochsen in Nordafrika kam (Grigson, 1991). Spätere genetische Analysen wiesen jedoch darauf hin, dass sowohl die Vorfahren afrikanischer als auch europäischer Rinder in der Region des Fruchtbaren Halbmondes anzusiedeln sein könnten (Bradley et al., 1996). Moderne SNP (Single nucleotide polymorphism) -Array-Analysen stützen die Vermutung, dass es nach Einwanderung domestizierter tauriner Rinder nach Afrika zur Hybridisierung mit wilden lokalen Auerochsen kam (Pitt et al., 2019). Um 1500 vor Christus und später mit der arabischen Invasion gelangten Zebus nach Afrika, die ebenfalls mit den taurinen Rassen gekreuzt wurden (Mukasa-Mugerwa, 1989, Loftus et al., 1994). Vom Norden ausgehend verbreiteten sich die Rinder über den Kontinent hinweg und in unterschiedlichen Regionen entwickelten sich verschiedene Sanga-Rassen. Diese findet man heute hauptsächlich in Ost-, Süd-

und Zentralafrika (Mwai et al., 2015). Eine dieser Sanga-Rassen ist das Watussi-Rind (Abbildung 7), das zu den Ankolerindern gehört. Ankolerinder dienen in Afrika der Gewinnung von Milch, Fleisch und Blut (Mbuzza, 1995, Taye et al., 2017).



**Abbildung 7:** Gruppe von Watussi-Rindern (Ankolerinder) mit den typisch langen Hörnern. Quelle: <https://pixabay.com/de/> (Nutzung lizenzfrei)

#### 4.4. Wasserbüffel (*Bubalus arnee f. bubalis*)

Die heute domestizierten Wasserbüffel (*Bubalus arnee f. bubalis*) (Abbildung 8) stammen von der asiatischen Wildform des Wasserbüffels (*Bubalus arnee*) ab und unterscheiden sich somit deutlich von den „eigentlichen Rindern“, die aus Auerochsen hervorgegangen sind. Die Wildform des Wasserbüffels kann in zwei verschiedene Populationen unterteilt werden, die sich geographisch unabhängig voneinander entwickelten (Kumar et al., 2007). Der Flussbüffel, der ursprünglich über den westlichen Teil des indischen Subkontinents verbreitet war, wurde vor ca. 6300 Jahren domestiziert (Colli et al., 2018, Sun et al., 2020). Der Sumpfbüffel wiederum, der in der Grenzregion Chinas und Indochinas vor etwa 3000-7000 Jahren domestiziert wurde, ist heute hauptsächlich in östlicheren Regionen zu finden, die sich von Bangladesch bis zum Yangtse-Tal in China erstrecken (Zhang et al., 2011, Wang et al., 2017). Nach der Domestikation verbreiteten sich die Nachkommen der Sumpfbüffel immer weiter innerhalb Chinas und im südostasiatischen Raum, während die Abkömmlinge der Flussbüffel weit nach

Westen bis nach Südeuropa und Nordafrika vordrangen (Colli et al., 2018). Obwohl der Großteil der Wasserbüffel heute immer noch in Asien lebt, sind verschiedenste Rassen in Ländern aller Kontinente zu finden (Borghese, 2005, Zhang et al., 2020).

Wasserbüffel, deren Zahl in den letzten Jahrzehnten deutlich angestiegen ist, haben weltweit eine große ökonomische Bedeutung und werden als Rohstofflieferanten und Arbeitstiere genutzt. Indien und Pakistan stellen die Top-Produzenten von Wasserbüffelmilch dar und auch in einigen europäischen Ländern werden steigende Tierzahlen und Milchmengen verzeichnet, da sich die aus Büffelmilch gewonnenen Produkte wie Mozzarella immer größerer Beliebtheit erfreuen (Cockrill, 1977, Pasha and Hayat, 2012, Hegde, 2019).



**Abbildung 8:** Fotografie asiatischer Wasserbüffel (*Bubalus arnee f. bubalis*). Sie verdeutlicht das enge Zusammenleben von Tier und Mensch in vielen asiatischen Familien, für die die Tiere häufig das größte Kapital darstellen (Luo et al., 2020). Quelle: <https://pixabay.com/de/> (Nutzung lizenzfrei)

#### 4.5. Yak (*Bos mutus f. grunniens*)

Die genaue Entwicklungsgeschichte des Yaks (Abbildung 9) ist bis heute nicht eindeutig geklärt. Yaks und Rinder trennten sich vermutlich vor ca. 4,9 Millionen Jahren in zwei eigenständige Spezies auf (Qiu et al., 2012). Fossile Funde deuten darauf hin, dass die Vorfahren der wilden Yaks (*Bos mutus*) im späten Tertiär vor etwa 2,5 Millionen Jahren im nordöstlichen Eurasien weit verbreitet waren. Es wird angenommen, dass die Domestikation der wilden Yaks durch nomadisch lebende Bevölkerungsgruppen vor ca. 7300 Jahren auf der Qinghai-Tibet-Hochfläche

stattgefunden hat (Qiu et al., 2015). Vom Hochland Tibets ausgehend verbreiteten sie sich in alle Himmelsrichtungen und besiedelten weitere Regionen Asiens, vornehmlich Höhenlagen zwischen 2000 und 5000 Metern (Wiener et al., 2003). Mitte des 19. Jh. wurden einige Yaks nach Europa und Nordamerika gebracht, um in zoologischen Einrichtungen gehalten oder mit lokalen Rinderrassen gekreuzt zu werden (Wiener et al., 2003).

Die größten Yak-Bestände sind in verschiedenen Provinzen Chinas, in Tibet und der Mongolei zu finden. Yaks werden vor allem als Lastentiere genutzt und sind daher für nomadisch lebende Bevölkerungsgruppen unverzichtbare Nutztiere. Neben Fleisch und Milch werden besonders das Fell und die Häute der Tiere verwertet (Olsen, 1990, Wiener et al., 2003). Um das Yak-Management und die Produktivität der Tiere zu optimieren, werden diese, vor allem in niedrigeren Höhenlagen, schon seit Jahrhunderten mit heimischen taurinen Rinderrassen hybridisiert (Medugorac et al., 2017).



**Abbildung 9:** Fotografie zweier Yaks (*Bos mutus f. grunniens*), die den nomadisch lebenden Bevölkerungsgruppen hauptsächlich als Lastentiere dienen. Quelle: <https://pixabay.com/de/> (Nutzung lizenzfrei)

#### 4.6. Hausziege (*Capra aegagrus f. hircus*)

Die heutigen Hausziegen (Abbildung 10) stammen von wilden Bezoarziegen (*Capra aegagrus*) ab, die während des Neolithikums vor ca. 10.500 Jahren im fruchtbaren Halbmond des Nahen Ostens domestiziert wurden (Zeuner, 1963, Porter, 1996, Naderi et al., 2008). Obwohl dieses Ereignis in derselben Epoche stattfand wie die Domestikation der Rinder, stellt die der Ziegen vermutlich die früheste Form der Viehzucht dar (Zeder, 2008). DNA-Analysen deuten darauf hin, dass es bei den wilden Vorgängern vermutlich mindestens drei genetisch getrennte Ziegenpopulationen gegeben haben muss, die sich im Laufe der Domestikation immer mehr vermischt haben. Schwache phylogeographische Strukturen der heutigen Hausziegen weisen auf einen hohen Genfluss hin und lassen zudem darauf schließen, dass Ziegen seit ihrer Domestikation in hohem Maße interkontinental verbracht wurden und eine wichtige Rolle bei Kolonialisierungen, Migrationen und im Handel gespielt haben müssen (Porter, 1996, Luikart et al., 2001, Zeder et al., 2006).

Besonders in ländlichen und wirtschaftlich weniger entwickelten Regionen sichert die Ziegenhaltung der Bevölkerung den Zugang zu Milch und Fleisch. Weltweit wurden in den letzten Jahrzehnten steigende Tierzahlen verzeichnet, wobei die größten Ziegenpopulationen in Asien und Afrika zu finden sind (Gilbert et al., 2018, Miller and Lu, 2019).



**Abbildung 10:** Eine Ziegenherde (*Capra aegagrus f. hircus*) beim Grasen auf der Weide. Quelle: <https://pixabay.com/de/> (Nutzung lizenzfrei)

#### 4.7. Hausschaf (*Ovis aries*)

Die genaue Abstammung der heutigen Hausschafe (Abbildung 11) ist nicht vollends geklärt. Obwohl angenommen wird, dass wilde asiatische Mufflons die Vorfahren der domestizierten Schafe repräsentieren (Pedrosa et al., 2005), herrscht in der Literatur teils Uneinigkeit bzgl. der Nomenklatur. Daher wird hier der Einfachheit halber für das Hausschaf der wissenschaftliche Name *Ovis aries* verwendet, ohne dabei Bezug auf die Wildform zu nehmen.



**Abbildung 11:** Domestizierte Schafe (*Ovis aries*) der Rassen Coburger Fuchsschaf (links) und Walliser Schwarznasenschaf (rechts). Quelle: <https://pixabay.com/de/> (Nutzung lizenfrei)

Der Ursprung der Gattung *Ovis* ist in Vorderasien anzusiedeln. Untersuchungen mitochondrialer DNA weisen auf mehrere genetisch unterschiedliche Abstammungslinien hin, aus denen unsere heutigen Hausschafe hervorgegangen sind (Hiendleder et al., 2002, Pedrosa et al., 2005, Zeder et al., 2006). Daher wird angenommen, dass es zu verschiedenen Zeitpunkten zu unabhängigen Domestikationsereignissen kam, die vermutlich vor rund 10.500 Jahren ebenfalls im Fruchtbaren Halbmond stattfanden und genau wie die der Ziegen eine der frühesten Formen der Viehhaltung darstellen (Zeder, 2008). Archäologische Funde stützen die Vermutung, dass sich die ersten Domestikationen in den Regionen Mesopotamien, Zentral-Anatolien und dem Zāgros-Gebirge ereigneten. (Luikart et al., 2001, Zeder et al., 2006).

Heute findet man Schafe in unzähligen Ländern aller Kontinente. Sie werden vielseitig genutzt und dienen der Produktion von Milch, Fleisch, Wolle und Tierhäuten. Die weltweite Schafmilchproduktion ist in den vergangenen Jahrzehnten deutlich angestiegen und nimmt weiter an Bedeutung zu (Pulina et al., 2018). Hierbei stellen nach Angaben der Food and Agriculture Organization of the United Nations (FAO) China und die Türkei die größten Produzenten dar. In Ländern des Nahen Ostens und südlich der Sahara gelten Schafe ebenfalls als unverzichtbare Nahrungsquelle und Rohstofflieferanten (Timon and Hanrahan, 1986, Zygogiannis, 2006, Pulina et al., 2018).



### **III. ZIELSETZUNG DER ARBEIT**

Ziel dieser Arbeit war es, verschiedene Tierspezies innerhalb der Familie *Bovidae* auf das Vorkommen bestimmter BMMF-DNA und CRESS-DNA-Viren zu untersuchen. Dabei wurden ausschließlich Spezies inkludiert, die in verschiedenen Regionen weltweit der Fleisch- und Milchproduktion dienen. Da bereits Analysen von Proben europäischer Rinder vorliegen, lag der Schwerpunkt dieser Arbeit auf anderen Spezies als *Bos taurus*. Von besonderer Bedeutung waren dabei diejenigen Tierarten, die in Ländern mit niedrigen Darm- und/oder Brustkrebsinzidenzen genutzt werden. Die Untersuchung von Milch- und Blutproben ermöglichte eine Aussage über die Empfänglichkeit der verschiedenen Arten gegenüber einer Infektion mit den „Erregern“ und gibt Hinweise auf das Vorkommen in deutschen Herden. Da die untersuchten Tierspezies vielerorts als Nahrungsquelle dienen, können die gewonnenen Daten dazu beitragen, die postulierte Korrelation zwischen dem Verzehr dieser Tierprodukte und einem erhöhten Krebsrisiko näher zu beleuchten. Des Weiteren wurden Milchproben aus dem Einzelhandel untersucht, um eine Aussage darüber treffen zu können, inwieweit eine Erhitzung der Milch im Rahmen der molkereitechnischen Bearbeitung Einfluss auf die Detektion kleiner zirkulärer Einzelstrang-DNA hat.

Da besonders die BMMF eine neuartige und nicht näher klassifizierte „Erregergruppe“ darstellen, war ein weiteres Ziel die Isolierung bisher unbekannter BMMF-Isolate und deren detaillierte Genomanalyse, um zu einem besseren Verständnis potenzieller Risiken für den Konsumenten beizutragen. Im Sinne des gesundheitlichen Verbraucherschutzes liefert die vorliegende Arbeit erste Daten, die als Basis für eine Risikobewertung herangezogen werden können. Bereits im Jahr 2019 hatte das Bundesinstitut für Risikobewertung (BfR) in einer Stellungnahme auf die unzureichende Datenlage hingewiesen.



## IV. PUBLIKATIONEN

### 1. Publikation I

#### Titel

Circular Rep-Encoding Single-Stranded DNA Sequences in Milk from Water Buffaloes (*Bubalus arnee f. bubalis*)

#### Autoren

Marie-T. König<sup>1</sup>, Robert Fux<sup>2</sup>, Ellen Link<sup>2</sup>, Gerd Sutter<sup>2</sup>, Erwin Märtlbauer<sup>1</sup>,  
Andrea Didier<sup>1</sup>

<sup>1</sup> Department für Veterinärwissenschaften, Institut für Lebensmittelwissenschaften, Veterinärmedizinische Fakultät, Ludwig-Maximilians-Universität München, Schönleutnerstr. 8, 85764 Oberschleißheim, Deutschland

<sup>2</sup> Department für Veterinärwissenschaften, Lehrstuhl für Virologie, Veterinärmedizinische Fakultät, Ludwig-Maximilians-Universität München, Veterinärstraße 13, 80539 München, Deutschland

#### Journal

Viruses (IF 2021: 5,048)

#### Status

akzeptiert am 03.06.2021

publiziert online am 07.06.2021

Viruses 2021, 13(6), 1088

doi:10.3390/v13061088



Article

## Circular Rep-Encoding Single-Stranded DNA Sequences in Milk from Water Buffaloes (*Bubalus arnee f. bubalis*)

Marie-T. König <sup>1</sup>, Robert Fux <sup>2</sup> , Ellen Link <sup>2</sup>, Gerd Sutter <sup>2</sup> , Erwin Märtblauer <sup>1</sup> and Andrea Didier <sup>1,\*</sup>

<sup>1</sup> Department of Veterinary Sciences, Institute of Food Safety, Faculty of Veterinary Medicine, Ludwig-Maximilians-Universität München, Schönleutnerstraße 8, 85764 Oberschleißheim, Germany; marie-therese.koenig@mh.vetmed.uni-muenchen.de (M.-T.K.); milchhygiene@mh.vetmed.uni-muenchen.de (E.M.)

<sup>2</sup> Department of Veterinary Sciences, Chair of Virology, Faculty of Veterinary Medicine, Ludwig-Maximilians-Universität München, Veterinärstraße 13, 80539 München, Germany; robert.fux@micro.vetmed.uni-muenchen.de (R.F.); ellen.link@micro.vetmed.uni-muenchen.de (E.L.); gerd.sutter@lmu.de (G.S.)

\* Correspondence: a.didier@mh.vetmed.uni-muenchen.de

**Abstract:** Isolation and characterization of circular replicase-encoding single-stranded (ss) DNA from animal, plant and environmental samples are rapidly evolving in virology. We detected 21 circular DNA elements, including one genomoviral sequence, in individual milk samples from domesticated Asian water buffaloes (*Bubalus arnee f. bubalis*). Most of the obtained genomes are related to Sphinx 1.76 and Sphinx 2.36 sequences and share a high degree of similarity to recently published circular DNAs—named BMMF (bovine meat and milk factors)—that have been isolated from commercial milk, as well as from bovine serum. Characteristic features such as *rep* genes, tandem repeats and inverted repeats were detected. These BMMF have recently been found to be present in taurine-type dairy cattle breeds descending from the aurochs (*Bos primigenius*). Importantly, the occurrence of BMMF has been linked to the higher incidence of colorectal and breast cancer in North America and Western Europe compared with Asia. This is the first report of circular ssDNA detected in milk from the domesticated form of the wild Asian water buffalo (*B. arnee*) belonging to the subfamily *Bovinae*. This novelty should be taken into account in view of the above-mentioned cancer hypothesis.

**Keywords:** BMMF; circular ssDNA; colon/breast cancer; replication-associated protein (Rep); rolling circle amplification; rolling circle replication; water buffalo milk



**Citation:** König, M.-T.; Fux, R.; Link, E.; Sutter, G.; Märtblauer, E.; Didier, A. Circular Rep-Encoding Single-Stranded DNA Sequences in Milk from Water Buffaloes (*Bubalus arnee f. bubalis*). *Viruses* **2021**, *13*, 1088. <https://doi.org/10.3390/v13061088>

Academic Editor: Hironori Sato

Received: 3 May 2021

Accepted: 3 June 2021

Published: 7 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

Single-stranded (ss) DNA viruses can be detected throughout all domains of life and are currently classified into thirteen families according to the International Committee on Taxonomy of Viruses (ICTV) [1]. Reports on unclassified virus-like ssDNA have drastically increased in recent years [2,3]. This augmentation is mainly caused by the application of high-throughput metagenomics such as next-generation sequencing (NGS) [4]. Nevertheless, classical approaches comprising rolling circle amplification (RCA) [5] followed by PCR and Sanger sequencing still have their importance for sequence detection and characterization. Recently, these sequences have been assigned to the virus phylum *Cressnaviricota*, where CRESS stands for “circular Rep-encoding single-stranded”, to describe functional features of these isolates, and viricota is the suffix for phylum taxa [6]. Besides their circular ssDNA composition, they depend on a replicase (Rep) for rolling circle replication (RCR) of their genomes. Encoded Reps exhibit a nicking/joining activity at the origin of replication (ori) of the ssDNAs. In addition to this endonuclease domain at the N-terminus, in some cases Reps include a superfamily 3 (SF3) helicase domain at their C-terminus, which helps to unwind the double-stranded DNA replicative intermediate [7,8]. In other cases, helicases encoded by the host are recruited during RCR. Sequence analysis of Reps on the genetic

as well as on the protein level provides information about the evolutionary history and the phylogenetic relatedness of different CRESS isolates [9]. The presence of a replicase is not restricted to CRESS DNA viruses, with sometimes blurred borders between CRESS viruses and plasmid-/phage-derived sequences. A prominent example of this overlap is the co-purification of Sphinx 1.76 and Sphinx 2.36 from transmissible spongiform encephalopathy (TSE) particles [10]. The first isolate shows significant sequence homology to a plasmid isolated from *Acinetobacter baumannii*, while the latter one shows similarities to an *Acinetobacter* sp. phage [11]. The resemblance between these sequences and their common mechanism of replication led to the hypothesis that ssDNA viruses developed from bacterial plasmids via gene transfer with RNA viruses [12,13]. Recently, Sphinx 1.76-related and Sphinx 2.36-related circular DNAs have also been isolated from cattle milk, cattle serum and human brain affected by multiple sclerosis [14–16]. With the ongoing identification of these types of ssDNA from bovine samples, the term BMMF (bovine meat and milk factors) has been introduced, and these factors have been suggested to represent a specific class of infectious agents [17]. According to their sequence homologies to known organisms, the factors are divided into four groups. The first two groups consist of isolates resembling the sequences of Sphinx 1.76 (group 1) or Sphinx 2.36 (group 2). Group 3 contains isolates that have been characterized as novel gemycircularviruses, being one of nine genera within the family *Genomoviridae*. A single isolate with similarity to a *Psychrobacter* species plasmid forms group 4. The terms HCBI (healthy cattle blood isolate) and CMI (cow milk isolate) are utilized to refer to the origin of each isolate [11]. A common attribute on the DNA level of all group 1 isolates is the presence of an iteron-like tandem repeat, a neighbouring inverted repeat, and at least one large open reading frame (ORF)-encoding Rep. Some of the isolates encode a putative second ORF resembling a mobilization element [18]. In contrast to other animal-derived CRESS viruses, these circular DNAs lack the genetic information for a capsid protein.

From a disease-related point of view, BMMF and Sphinx-like sequences have entered the spotlight due to a potential connection with colorectal and breast cancer as well as neurodegenerative illnesses such as multiple sclerosis [19]. BMMF DNA and Rep proteins have been found in tissue of colon cancer patients and are proposed to cause chronic inflammation [20], thus linking the cancer hypothesis to the consumption of milk and meat from taurine cattle. These are descendants from the European aurochs and are potential carriers of BMMF [21]. In North America and Western Europe, cattle breeds are largely taurine in origin (*Bos taurus taurus*). In contrast, the Asian cattle population largely consists of cattle breeds derived from the indicine lineage (*Bos taurus indicus*) [22,23], yaks (*Bos mutus*) [24] and water buffaloes (*Bubalus arnee*) [25], which are not supposed to be carriers of BMMF [26]. Especially on the Indian subcontinent and in Mongolia, the above-mentioned cancer incidence is lower than in Western countries [21,27]. However, studies on samples of “Asian-type” cattle investigating the presence of circular replication competent ssDNA sequences are missing in the literature. For this reason, we initiated an in-depth analysis of individual milk samples obtained from Asian water buffaloes (*B. arnee f. bubalis*).

## 2. Methods

### 2.1. Sample Collection

In total, 30 individual milk samples (50 mL) were collected from two different buffalo flocks in Germany, one kept in Lower Saxony (A) and one in Baden-Württemberg (B). A total of 23 samples were obtained from herd A and 7 from herd B. Immediately after collection, the milk was cooled to 4 °C and cold-shipped to the Institute of Food Safety within 24 h where it was immediately processed. Samples from flock A were numbered 8 through 30 and from flock B 42 through 48. In the case of obtaining full-length circular DNA sequences, sample numbers were completed with the term BAMI, which refers to “*Bubalus Arnee Milk Isolate*” and a number indicating the sequence length.

## 2.2. DNA Extraction and Rolling Circle Amplification (RCA) of Circular ssDNA

The QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) was applied to isolate total DNA from 200  $\mu$ L of whole milk according to the manufacturer's instructions. Approx. 40 ng of purified DNA was subjected to RCA using the TempliPhi Amplification Kit (Cytiva, Marlborough, MA, USA) with random primers according to the instruction manual of the supplier. Isothermal amplification was carried out at 30  $^{\circ}$ C for 18 h.

## 2.3. PCR Screening

RCA products were initially screened for the presence of *rep*-sequences with a set of 10 primer pairs designed via the *rep* genes of Sphinx 1.76 (GenBank Acc. No. HQ444404.1), Sphinx 2.36 (GenBank Acc. No. HQ444405.1) and related sequences obtained by Blast search. A similar approach was chosen with primers for *Genomoviridae*. The total volume of all PCR reactions was 50  $\mu$ L applied on a TProfessional Gradient 96 Thermocycler (Biometra, Jena, Germany). ThermoPrime Plus Polymerase (ThermoScientific, Waltham, MA, USA) was used at 1.25 U/reaction. The dNTPs (ThermoScientific, Waltham, MA, USA) were added to a final concentration of 200  $\mu$ M each. Detailed information on the primer sequences as well as the individual amplification programmes is summarized in Tables S1 and S2. After agarose-gel electrophoresis, bands were purified with the HighYield PCR Purification/Gel Extraction Kit (SLG, Südlaborbedarf, Gauting, Germany). Sequencing was performed by MWG Eurofins (Ebersberg, Germany). The sequences obtained from the PCR screening were aligned, and primer pairs matching homologous regions of the alignments were designed for inverse PCR in order to obtain full-length sequences of the circular DNA elements (for detailed information see Table S3). Additionally, abutting primers for amplification of BMMF2 sequences published by de Villiers et al. [18] and for BMMF1 sequences published by Whitley et al. [16] were applied.

## 2.4. Cloning of Inverse PCR Products

Amplification products obtained from inverse PCR were cloned into pCR2.1 TopoTA vector (Invitrogen, Carlsbad, CA, USA). Briefly, the gel-purified PCR product was mixed with the vector at a molar ratio of approx. 3:1 in a total volume of 6  $\mu$ L and allowed to rest for 30 min at room temperature for ligation. Afterwards, chemically competent *Escherichia coli* DH5 alpha were transformed with the ligation reaction for 30 min on ice. After a heat shock at 42  $^{\circ}$ C for 1 min, 1 mL of LB medium was added to the reaction and cells were allowed to recover for 1 h at 37  $^{\circ}$ C with gentle agitation. Next, *E. coli* were centrifuged at 6100  $\times$  g for 5 min. The pellet was suspended in 200  $\mu$ L LB medium and plated on LB agar with ampicillin (100  $\mu$ g/mL). After overnight incubation in LB medium with ampicillin, five clones from each plate were subjected to MiniPrep Plasmid isolation via the GenElute Plasmid Miniprep Kit (Sigma Aldrich, Taufkirchen, Germany). A control digest of 500 ng plasmid DNA with EcoRI (New England BioLabs, Ipswich, MA, USA) was performed, and insert-bearing clones were sequenced via the M13for and M13rev primer binding sites present in the vector.

## 2.5. Data Analysis

Sequences obtained from MWG Eurofins were subjected to BlastSearch against the nucleotide database at NCBI [28]. For multiple sequence alignments and cross-comparison, the DNASTAR Software (DNASTAR Inc., Madison, WI, USA, Version 15.0.0) was used [29]. Phylogenetic trees were calculated in MEGA software (version 5.03) by using the maximum likelihood method, and bootstrap values were computed with 500 replicates [30]. Furthermore, all sequences were checked at Emboss explorer for tandem repeats [31] and inverted repeats [32] upstream the *rep* gene sequence. Maximum size of tandem repeats was restricted to 30 nucleotides (nt), and the minimum size of inverted repeats was set to 4 nt. In silico detection of potential open reading frames was performed with the ORFfinder software at NCBI [33]. ATG or TTG as an alternative start codon was allowed. The minimum ORF size was restricted to 75 amino acids (aa). For subsequent comparative and

functional analysis, only ORFs  $\geq 95$  aa were included. All full-length genomes and their annotations detected in this study were deposited in the GenBank database (Accession numbers MW828657–MW828677).

### 3. Results

#### 3.1. PCR Screening and Detection of Circular Single-Stranded DNA Elements

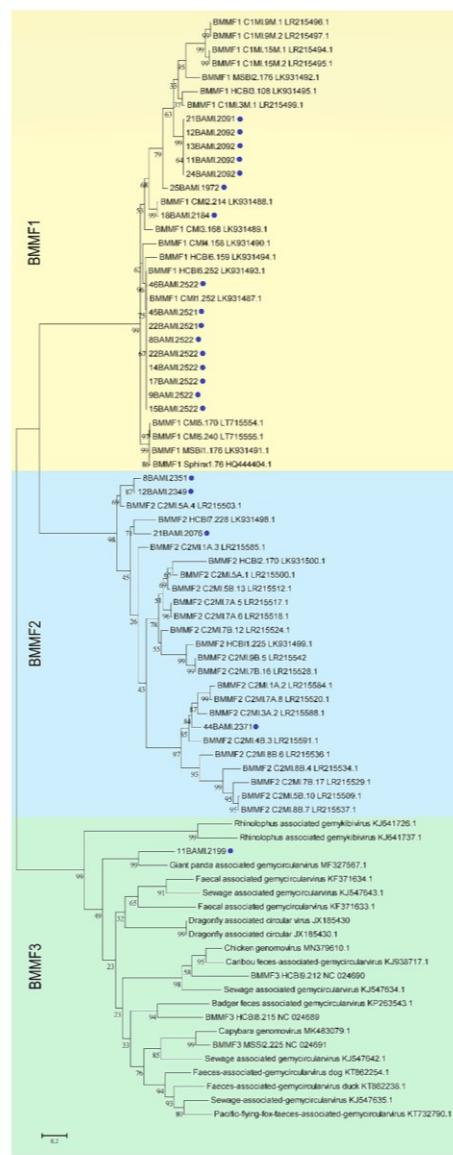
By using RCA followed by several screening PCRs, 83% (19/23) of the individual milk samples from flock A and 100% (7/7) from flock B reacted positive in at least one PCR applied for BMMF and CRESS DNA detection. Inverse PCRs employing novel primers from this study resulted in retrieval of four circular DNA elements, all from flock A (8BAMI.2351, 11BAMI.2199, 12BAMI.2349, 21BAMI.2076). Application of abutting primers published by Whitley et al. [16] and de Villiers et al. [18] yielded 17 additional circular DNAs from samples of both flocks. Interestingly, five of the positive milk samples (8BAMI, 11BAMI, 12BAMI, 21BAMI, 22BAMI) contained two different circular DNA elements. Altogether, 21 unique complete circular ssDNA genomes were detected in the present study, one of them later assigned as gemycircularvirus.

#### 3.2. Characterization of the Sequences on the DNA Level

The overall sequence length of the circular DNA elements ranged from 1972 to 2522 nt. After subjecting all DNAs to BlastSearch, the best two matches were summarized, as shown in Table S4. Nine of the 21 detected circular DNAs are nearly identical to the BMMF1 isolates HCBI6.252 and CMI1.252, showing a query coverage of 100% and more than 99% identity. Seven sequences are similar to C1MI GenBank entries with a query coverage of 71–98% and 77–97% identity, respectively. Four of the DNA elements revealed similarities to C2MI isolates, belonging to the BMMF2 group. One sequence from the present study was assigned to a gemycircularvirus detected in a giant panda with 92.94% identity in BlastSearch but 48% query coverage only and therefore belonged to the BMMF group 3. None of the herein detected isolates belongs to BMMF group 4.

Homologies among sequences from the present study were calculated by additional multiple sequence alignment. Sequences 8BAMI.2351, 12BAMI.2349, 21BAMI.2076 and 44BAMI.2371 are the most remote ssDNA elements showing a maximum of 41% identity to the other sequences, but 59–99% identity among each other. Some of the other circular DNAs (8BAMI.2522, 9BAMI.2522, 15BAMI.2522, 17BAMI.2522, 22BAMI.2521, 22BAMI.2522, 45BAMI.2521, 46BAMI.2522) share  $\geq 99\%$  identity despite originating from different flocks. These isolates differ in 6 to 14 nt positions only. A similar observation was made for isolates 11BAMI.2092, 12BAMI.2092, 13BAMI.2092, 21BAMI.2091 and 24BAMI.2092, which possess  $\geq 99\%$  identity as well.

Next, a phylogenetic distance tree was generated including Sphinx 1.76, Sphinx 2.36, representatives from the BMMF1 and 2 groups and from the family *Genomoviridae* as references (Figure 1). This approach allowed the affiliation of the novel sequences to the BMMF1 and 2 groups. Within the BMMF1 group, the isolates 11BAMI.2092, 12BAMI.2092, 13BAMI.2092, 21BAMI.2091 and 24BAMI.2092 form a cluster distinct from any of the sequences published by the group of zur Hausen so far. The Gemycircularvirus-like element detected in sample 11 is the only representative of the BMMF3 group according to the nomenclature of zur Hausen et al. [11].



**Figure 1.** Results of the molecular phylogenetic analysis including novel full-length nucleotide sequences and representatives of the BMMF groups 1 to 3. The latter also includes members of the family *Genomoviridae*. The evolutionary history was inferred by using the maximum likelihood method and a JTT matrix-based model [34]. The bootstrap consensus tree was calculated from 500 replicates. Bootstrap values at the branch points indicate the percentage of replicate trees in which respective isolates clustered [35]. Branch support values lower than 70% were not included. Blue dots mark all sequences found in this study. The scale bar represents the number of substitutions per site. All taxa are indicated by name followed by the corresponding GenBank accession number.

Further analysis of the 21 circular sequences for typical replication-relevant ssDNA elements included: (i) search for tandem repeats (TR) [31], (ii) search for inverted repeats (IR)/stem loops as a potential origin of replication [32], (iii) search for replicase genes and (iv) search for additional putative ORFs [33].

Except for the circular DNA sequences 8BAMI.2351, 12BAMI.2349, 21BAMI.2076 and 44.BAMI.2371, all isolates exhibited typical tandem repeats that can also be found in related BMMF [11]. The DNA elements lacking a tandem repeat belonged either to the BMMF2 or BMMF3 group. Detailed information on the tandem repeat features is illustrated in Figure 2.



**Figure 2.** Clustal alignment of the three different TR motifs identified. Each of the upper two boxes comprises isolates with identical TR sequences, except for one nucleotide substitution in 14BAMI.2522. The lower box shows one further TR represented by a single isolate only. Vertical black lines indicate the end of a single repeat. The TR size is 22 nt, and the copy number ranges from 3.5 to 3.9.

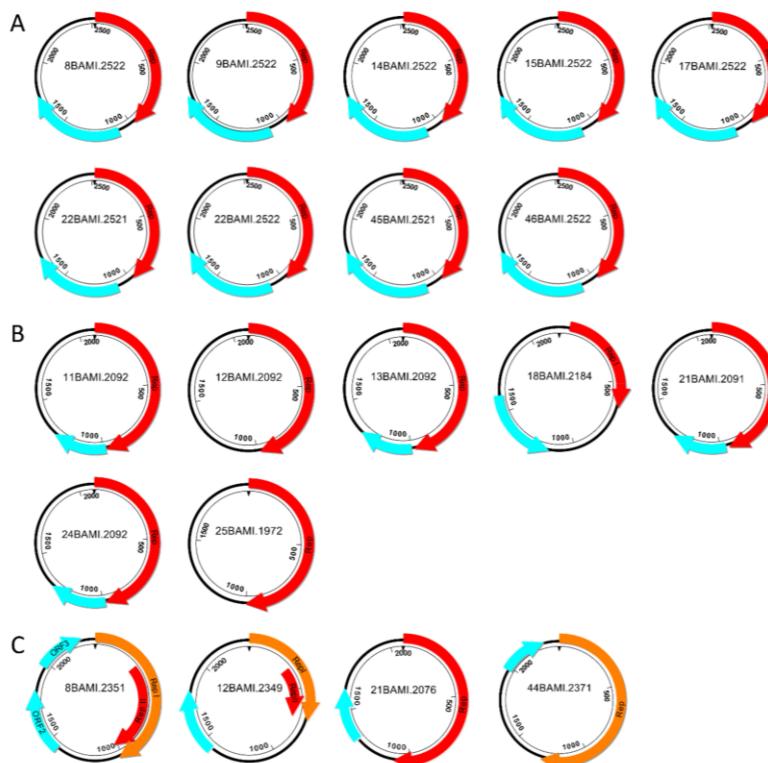
Furthermore, the circular DNAs were analysed for the presence of inverted repeats in proximity to the tandem repeat region, thus potentially able to form a double strand and to function as a replicase docking site. All BMMF1 sequences with TR motifs contain the same palindromic sequence TAAATGCTTTA (the first and the last 4 nt build the inverted repeat region separated by a 4 nt gap), which is located 49 to 56 nt upstream of the TR region (Table 1) and has also been described by zur Hausen et al. [11].

**Table 1.** Inverted repeat motif as a potential origin of replication. Sequences were detected by EMBOSS inverted repeat finder and additionally compared with the sequences described by zur Hausen et al. [11].

Isolate	Size	Inverted Repeat	nt between IR and TR	nt between IR and rep
8BAMI.2522	12		50	193
9BAMI.2522	12		50	193
11BAMI.2092	12		56	160
12BAMI.2092	12		56	160
13BAMI.2092	12		56	160
14BAMI.2522	12		50	193
15BAMI.2522	12		50	193
17BAMI.2522	12		50	193
18BAMI.2184	12	TAAATGCTTTA	49	133
21BAMI.2091	12		56	192
22BAMI.2522	12		50	193
22BAMI.2521	12		50	193
24BAMI.2092	12		56	160
25BAMI.1972	12		56	160
45BAMI.2521	12		50	193
46BAMI.2522	12		50	193

### 3.3. In Silico Characterization of BMMF-like DNA Elements on the Protein Level

In addition to the analyses on the nucleotide level, we performed an in silico characterization in order to detect open reading frames (ORFs) [33] and the deduced proteins including a subsequent BlastSearch. Basic features of all circular DNA elements detected in this study are depicted in Figure 3. The only gemycircularvirus found herein is described in Section 3.4.

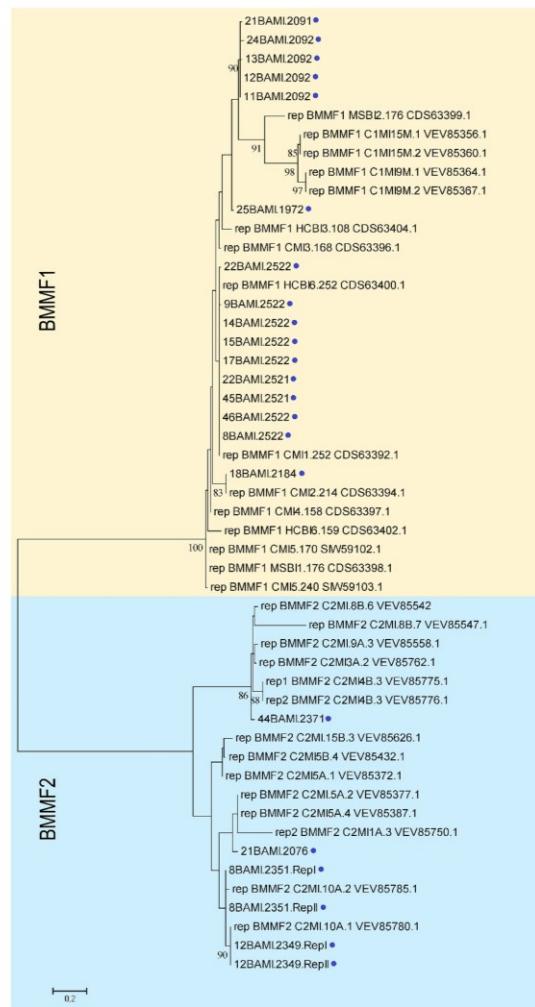


**Figure 3.** Schematic illustration of BMMF-like DNA elements recovered from buffalo milk. Red arrows indicate *rep* genes starting with ATG; orange arrows show *rep* genes starting with TTG. The light blue arrows represent further potential ORFs. According to the results of the in silico analysis, circular DNA elements were assigned to three groups (A–C). Group (A) contains nine related isolates, which could be assigned to the GenBank entries CMI1.252 and HCBI6.252; group (B) comprises seven elements, which exhibit the greatest homologies to other BMMF1 isolates; group (C) includes four circular DNAs belonging to the BMMF2 group.

Each circular DNA element possesses an ORF coding for Rep. Depending on the settings in the ORFfinder tool (“ATG only” versus “ATG and alternative initiation codons”), results slightly differ. Therefore, two isolates with overlapping or nested *rep*-ORFs were found, starting with either ATG or TTG. In this case, respective protein sequences were named as follows: 8BAMI.2351.RepI and 8BAMI.2351.RepII. Results of BlastSearch analysis are included in Table S5.

The majority (13) of the 22 putative Reps found in this study showed the best Blast hits for BMMF-related replicases. The remaining nine Reps scored closer to *Acinetobacter*-related replicases or those from even more distant organisms. Remarkably, the *rep* translation

products of 12BAMI.2349 and 18BAMI.2184 are shorter (up to 150 aa) than their best matches after BlastSearch (Table S5). Phylogenetic analysis of the Rep proteins revealed that they belong to the established BMMF groups (Figure 4). Again, the isolates 11BAMI.2092, 12BAMI.2092, 13BAMI.2092, 21BAMI.2091 and 24BAMI.2092 build their own cluster.

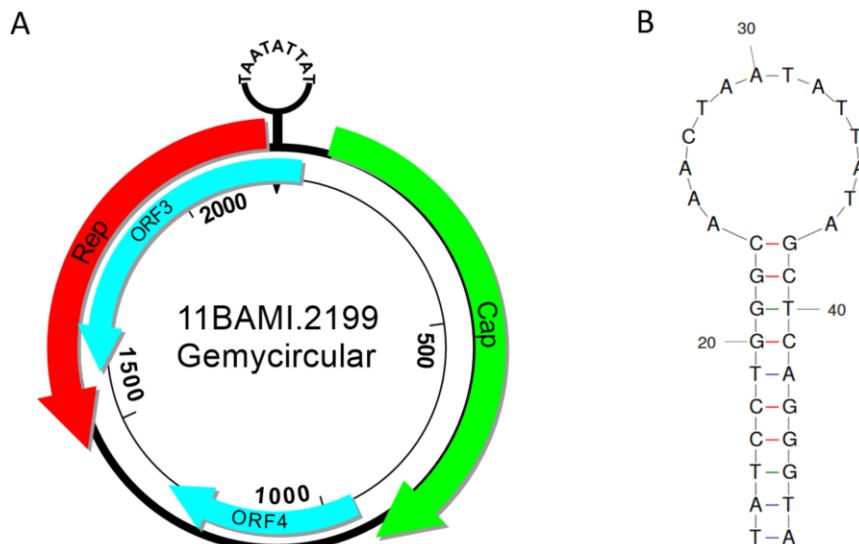


**Figure 4.** Maximum likelihood phylogenetic tree of Rep amino acid sequences and representative Reps from BMMF1 and 2 isolates calculated as described in Figure 1 [36]. The bootstrap consensus tree is based on 500 replicates. Bootstrap values at the branch points indicate the percentage of replicate trees in which respective isolates clustered [35]. Branch support values lower than 70% were not included. Blue dots mark all sequences found in this study. The scale bar represents the number of substitutions per site. All taxa are indicated by name followed by the corresponding GenBank accession number.

All DNA elements were analysed for the potential occurrence of further proteins. Although the software tool detected up to 22 putative ORFs per DNA sequence, the length of the potential proteins was restricted to a minimum of 95 aa to be included in the Results section of the manuscript. For additional putative proteins, the same approach as for the Reps was carried out. Thirteen putative proteins found in the novel sequences were identical or nearly identical to uncharacterized proteins from BMMF1 or BMMF2 isolates. Detailed results of the BlastSearch are summarized in Table S6. 12BAMI.2092 and 25BAMI.1972 do not code for an additional, hypothetical protein. A phylogenetic tree based on the putative protein sequences is given in Figure S1.

#### 3.4. In Silico Characterization of a Novel Gemycircularvirus on the Protein Level

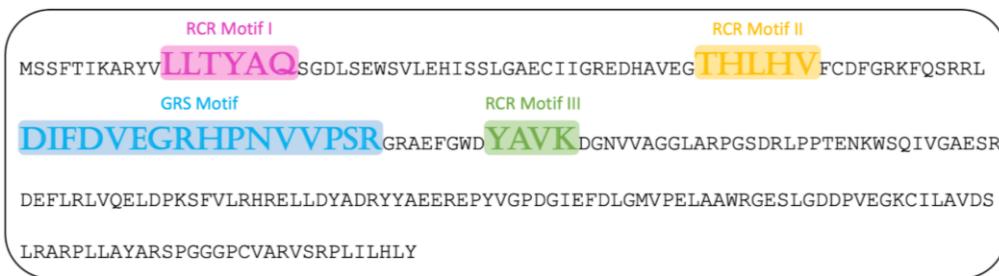
The only CRESS DNA virus isolated herein was recovered from a buffalo milk sample also containing one more circular DNA element belonging to the BMMF1 group. After in silico analyses, this new CRESS DNA virus was named 11BAMI.2199 (GenBank accession number MW828677). Typical characteristics of a gemycircularvirus could be identified. Besides a stem-loop structure at the origin of replication, it includes a *capsid (cap)* gene on the sense strand and a *rep* gene on the antisense strand. The putative stem-loop structure in the intergenic region between *rep* and *cap* features a nonanucleotide motif (TAATATTAT) that is present at the loop-tip and therefore a further characteristic of *Genomoviridae* [37–39] (Figure 5). Additionally, the isolate comprises two putative proteins, one in sense, the other one in antisense direction.



**Figure 5.** (A) Schematic map of the gemycircularvirus 11BAMI.2199. The in silico generated genes encoding Rep and Cap are shown as red and green arrows, respectively. Further putative ORFs are shown as light blue arrows. The nonanucleotide motif (TAATATTAT) within the potential stem-loop structure at position 1 is depicted at the top. (B) Enlarged view of the stem-loop structure between *rep* and *cap* as a potential origin of replication as calculated by the ProbKnot server at the University of Rochester. The nucleic acid type was set to DNA (<http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/ProbKnot/ProbKnot.html>; accessed on 17 March 2021).

BlastSearches of the proteins encoded by this virus revealed 60–93% identity to giant panda-associated and plant-associated proteins, inter alia (GenBank entries: QCX29436.1, ASH99166.1, ASH99165.1).

The typical replication mechanism for all CRESS DNA viruses is the rolling circle replication initiated by the Rep. 11BAMI.2199 encodes a Rep that contains well-conserved RCR motifs (Figure 6). Motif I is supposed to have a sequence recognition function [38]. Motif II is characteristic of HUH endonucleases [40]. Motif III is possibly involved in dsDNA nicking [38,41]. The GRS motif has initially been described in geminiviruses by Nash et al. [42].



**Figure 6.** Conserved motifs in the Rep of the *Gemycircularvirus* found in this study in accordance with data by Varsani et al. [38].

#### 4. Discussion

BMMF initially described by zur Hausen et al. [11,17] are suspected to correlate with colon and breast cancer incidences as well as neurodegenerative diseases in the “Western” world. BMMF DNA and Rep proteins have been detected in colorectal cancer peritumour and tumour tissues of human patients by immunohistochemistry and Western blot [20]. Thereby and by the means of transfection assays, transcriptional activity of BMMF can be assumed [43]. Lower cancer incidences in countries traditionally breeding cattle evolved from other bovine lineages such as water buffaloes, yaks and zebras suggest that circular ssDNA elements in milk from taurine cattle might be one of several food-borne risk factors that contribute to the aforementioned diseases.

The present study was initiated to screen milk samples from non-taurine cattle for the presence of BMMF and CRESS viruses. Although screening PCR results suggested that 86% (26/30) of the water buffalo milk samples were positive, circular full-length sequences could only be recovered from 53% (16/30). This discrepancy might be caused by the formation of linear by-products, which occur to some extent during RCA. Thus, a PCR screening with primers, amplifying solely subgenomic fragments, might not be a suitable marker for the potential presence of full-length genomes. Of the 21 circular ssDNA elements detected and characterized herein, 20 belong to the BMMF groups 1 and 2. Inverse PCR additionally revealed one sequence belonging to the genus *Gemycircularvirus* (BMMF3). The further in silico analysis of genomes described herein revealed that two sequences (12BAMI.2349 and 18BAMI.2184) contained Reps substantially reduced in length (Table S5). Their replication capacity might thus be questionable due to a potentially dysfunctional Rep. On the other hand, five of the 21 positive samples harbour two ssDNA elements. Therefore, one must also consider that a functional Rep from one element might be able to complement a dysfunctional Rep from a second sequence, thus enabling RCR of both. However, to answer this question, the functional activity and replication behaviour of these DNA elements will have to be tested in transfection assays, as described by Eilebrecht et al. [43].

From a nutritional point of view, water buffalo milk is of great importance in some Asian countries. Although the multipurpose domestic water buffalo is widely distributed all over the world [44], the largest number of animals kept for dairy production is recorded

in India, contributing to approx. 68% of the world's buffalo milk production [45]. Buffalo milk production in India exceeded 86 million tonnes in 2017–2018; at 49%, this accounts for the largest part of total Indian milk production, compared with cow's milk at only 47%. The per capita availability of milk adds up to 337 g per day, and the domestic demand for milk and milk products is steadily increasing [46]. From the centre of domestication in the Indian subcontinent, domestic water buffaloes spread westward to Egypt, the Balkans and Italy around the 7th century [25]. Nowadays, significant flocks in Europe are mainly kept in Italy, Romania and Turkey, but they represent only a minor proportion of total livestock [47,48].

Apart from one study in Switzerland that found one *gemycircularvirus* in the plasma of water buffaloes [49], reports on "Asian-type" cattle harbouring CRESS viruses or even BMMF are scarce. Either the cattle under study descend from the taurine lineage [50,51] or authors do not provide substantial information on the breeds [52–54]. None of the previous studies address the occurrence of BMMF in "Asian-type" cattle such as water buffaloes. Hence, this is the first profound survey investigating the appearance of circular ssDNA, particularly that of BMMF.

Although data from the present study underpin that BMMF are not restricted to taurine cattle, one has to keep in mind that water buffaloes from "Western" flocks might have acquired these DNA elements due to their coexistence with taurine cattle by yet unknown routes of transmission. As water buffaloes in Germany typically have access to pasture, the transmission from neighbouring, taurine cattle via vectors such as black flies, for example, cannot be excluded. To clarify this issue, future studies on samples from water buffaloes domiciled in Asia are needed.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/v13061088/s1>, Figure S1: Maximum likelihood phylogenetic tree of novel hypothetical protein sequences and uncharacterized BMMF1 and 2 proteins; Table S1: Information on the primer sets applied in the PCR screening for the amplification of Sphinx-like sequences. Table S2: Information on the primer sets applied in the PCR screening for the amplification of *Genomoviridae* family members. Table S3: Information on the primer sets applied in the inverse PCR for amplification of full-length DNA. Table S4: Summary of the BLASTn results for the individual isolates. Table S5: Summary of the BLASTp results of the in silico generated Reps. Table S6: Summary of the BLASTp results of the in silico generated additional proteins.

**Author Contributions:** Conceptualization—E.M. and G.S.; Methodology—E.L. and R.F.; Investigation and Visualization—M.-T.K.; Validation and Writing—A.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision by the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme (funding reference: 2819110217).

**Institutional Review Board Statement:** Ethical review and approval was waived for this study because milk was collected during the daily milking routine on the farm.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

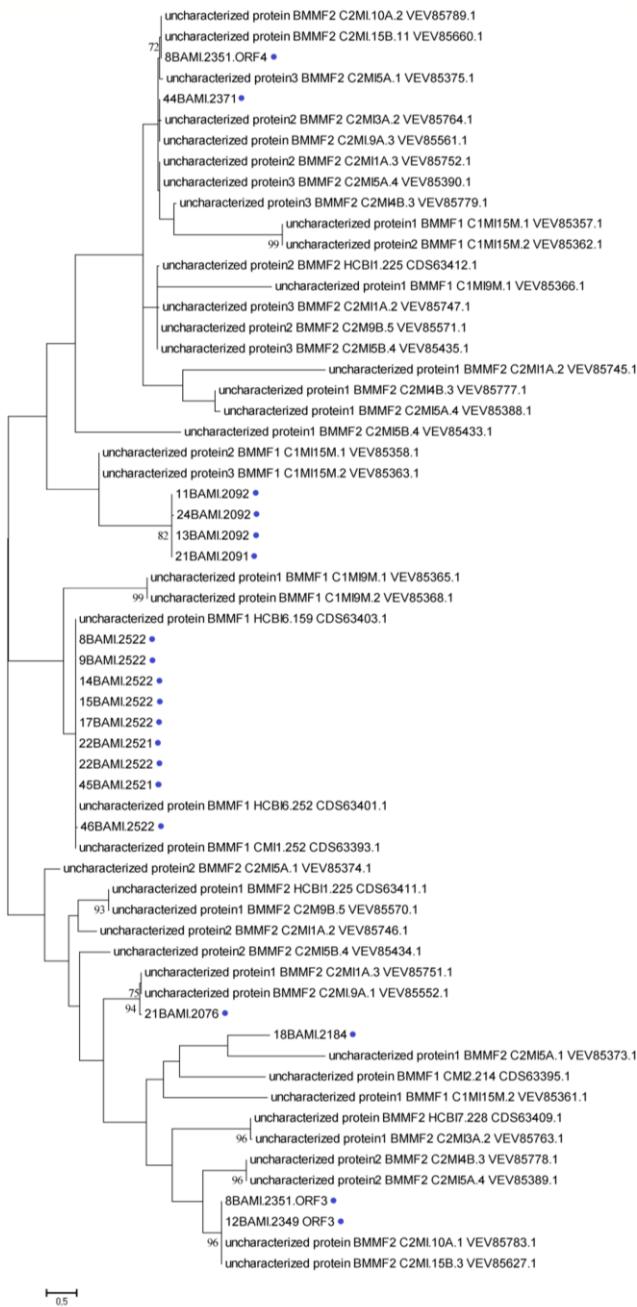
## References

1. International Committee on Taxonomy of Viruses (ICTV). Available online: <https://talk.ictvonline.org> (accessed on 28 October 2020).
2. Rosario, K.; Duffy, S.; Breitbart, M. A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. *Arch. Virol.* **2012**, *157*, 1851–1871. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Tisza, M.J.; Pastrana, D.V.; Welch, N.L.; Stewart, B.; Peretti, A.; Starrett, G.J.; Pang, Y.S.; Krishnamurthy, S.R.; Pesavento, P.A.; McDermott, D.H.; et al. Discovery of several thousand highly diverse circular DNA viruses. *Elife* **2020**, *9*. [\[CrossRef\]](#)
4. Dayaram, A.; Galatowitsch, M.L.; Argüello-Astorga, G.R.; Van Bysterveldt, K.; Kraberger, S.; Stainton, D.; Harding, J.S.; Roumagnac, P.; Martin, D.P.; Lefevre, P.; et al. Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. *Infect. Genet. Evol.* **2016**, *39*, 304–316. [\[CrossRef\]](#) [\[PubMed\]](#)

5. Johnne, R.; Müller, H.; Rector, A.; van Ranst, M.; Stevens, H. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends Microbiol.* **2009**, *17*, 205–211. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Krupovic, M.; Varsani, A.; Kazlauskas, D.; Breitbart, M.; Delwart, E.; Rosario, K.; Yutin, N.; Wolf, Y.I.; Harrach, B.; Zerbini, F.M.; et al. Cressdnaviricota: A Virus Phylum Unifying Seven Families of Rep-Encoding Viruses with Single-Stranded, Circular DNA Genomes. *J. Virol.* **2020**, *94*, 12. [\[CrossRef\]](#)
7. Kazlauskas, D.; Varsani, A.; Koonin, E.V.; Krupovic, M. Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nat. Commun.* **2019**, *10*, 3425. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Zhao, L.; Rosario, K.; Breitbart, M.; Duffy, S. Eukaryotic Circular Rep-Encoding Single-Stranded DNA (CRESS DNA) Viruses: Ubiquitous Viruses With Small Genomes and a Diverse Host Range. *Adv. Virus Res.* **2019**, *103*, 71–133. [\[CrossRef\]](#)
9. Rosario, K.; Mettel, K.A.; Benner, B.E.; Johnson, R.; Scott, C.; Yusseff-Vanegas, S.Z.; Baker, C.C.M.; Cassill, D.L.; Storer, C.; Varsani, A.; et al. Virus discovery in all three major lineages of terrestrial arthropods highlights the diversity of single-stranded DNA viruses associated with invertebrates. *PeerJ* **2018**, *6*, e5761. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Manuelidis, L. Nuclease resistant circular DNAs copurify with infectivity in scrapie and CJD. *J. Neurovirol.* **2011**, *17*, 131–145. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Hausen, H.Z.; Bund, T.; de Villiers, E.-M. Infectious Agents in Bovine Red Meat and Milk and Their Potential Role in Cancer and Other Chronic Diseases. In *Viruses, Genes, and Cancer*; Hunter, E., Bister, K., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 407, pp. 83–116. ISBN 978-3-319-61803-6.
12. Krupovic, M. Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. *Curr. Opin. Virol.* **2013**, *3*, 578–586. [\[CrossRef\]](#)
13. Koonin, E.V.; Dolja, V.V.; Krupovic, M. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* **2015**, *479*, 2–25. [\[CrossRef\]](#)
14. Funk, M.; Gunst, K.; Lucansky, V.; Muller, H.; Hausen, H.Z.; De Villiers, E.-M. Isolation of Protein-Associated Circular DNA from Healthy Cattle Serum. *Genome Announc.* **2014**, *2*, e00846-14. [\[CrossRef\]](#)
15. Lamberto, I.; Gunst, K.; Müller, H.; Hausen, H.Z.; De Villiers, E.-M. Mycovirus-Like DNA Virus Sequences from Cattle Serum and Human Brain and Serum Samples from Multiple Sclerosis Patients. *Genome Announc.* **2014**, *2*, e00848-14. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Whitley, C.; Gunst, K.; Müller, H.; Funk, M.; Hausen, H.Z.; De Villiers, E.-M. Novel Replication-Competent Circular DNA Molecules from Healthy Cattle Serum and Milk and Multiple Sclerosis-Affected Human Brain Tissue. *Genome Announc.* **2014**, *2*, e00849-14. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Hausen, H.Z.; Bund, T.; De Villiers, E. Specific nutritional infections early in life as risk factors for human colon and breast cancers several decades later. *Int. J. Cancer* **2018**, *144*, 1574–1583. [\[CrossRef\]](#) [\[PubMed\]](#)
18. De Villiers, E.-M.; Gunst, K.; Chakraborty, D.; Ernst, C.; Bund, T.; Hausen, H.Z. A specific class of infectious agents isolated from bovine serum and dairy products and peritumoral colon cancer tissue. *Emerg. Microbes Infect.* **2019**, *8*, 1205–1218. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Hausen, H.Z. What do breast and CRC cancers and MS have in common? *Nat. Rev. Clin. Oncol.* **2015**, *12*, 569–570. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Bund, T.; Nikitina, E.; Chakraborty, D.; Ernst, C.; Gunst, K.; Boneva, B.; Tessmer, C.; Volk, N.; Brobeil, A.; Weber, A.; et al. Analysis of chronic inflammatory lesions of the colon for BMMF Rep antigen expression and CD68 macrophage interactions. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2025830118. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Hausen, H.Z.; De Villiers, E.-M. Dairy cattle serum and milk factors contributing to the risk of colon and breast cancers. *Int. J. Cancer* **2015**, *137*, 959–967. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Decker, J.E.; McKay, S.; Rolf, M.; Kim, J.; Alcalá, A.M.; Sonstegard, T.S.; Hanotte, O.; Gotherstrom, A.; Seabury, C.M.; Praharani, L.; et al. Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle. *PLoS Genet.* **2014**, *10*, e1004254. [\[CrossRef\]](#)
23. Amer, P.R.; Kemp, R.A.; Smith, C. Genetic differences among the predominant beef cattle breeds in Canada: An analysis of published results. *Can. J. Anim. Sci.* **1992**, *72*, 759–771. [\[CrossRef\]](#)
24. Qi, X.B.; Jianlin, H.; Wang, G.; Rege, J.E.O.; Hanotte, O. Assessment of cattle genetic introgression into domestic yak populations using mitochondrial and microsatellite DNA markers. *Anim. Genet.* **2010**, *41*, 242–252. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Zhang, Y.; Colli, L.; Barker, J.S.F. Asian water buffalo: Domestication, history and genetics. *Anim. Genet.* **2020**, *51*, 177–191. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Hausen, H.Z.; De Villiers, E.-M. Risikofaktoren für Krebs und MS in Kuhmilch und Rindfleisch? *Biol. Unserer Zeit* **2016**, *46*, 26–31. [\[CrossRef\]](#)
27. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **2015**, *136*, E359–E386. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [\[CrossRef\]](#)
29. Burland, T.G. DNASTAR's Lasergene sequence analysis software. *Methods Mol. Biol.* **2000**, *132*, 71–91. [\[CrossRef\]](#)

30. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [\[CrossRef\]](#) [\[PubMed\]](#)
31. The European Molecular Biology Open Software Suite (EMBOSS) A. Available online: <http://emboss.bioinformatics.nl/cgi-bin/emboss/equicktandem> (accessed on 26 October 2020).
32. The European Molecular Biology Open Software Suite (EMBOSS) B. Available online: <https://www.bioinformatics.nl/cgi-bin/emboss/palindrome> (accessed on 26 October 2020).
33. National Center for Biotechnology Information (NCBI). ORFFinder. Available online: <https://www.ncbi.nlm.nih.gov/orffinder/> (accessed on 26 October 2020).
34. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [\[CrossRef\]](#)
36. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **1992**, *8*, 275–282. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Krupovic, M.; Ghabrial, S.A.; Jiang, D.; Varsani, A. Genomoviridae: A new family of widespread single-stranded DNA viruses. *Arch. Virol.* **2016**, *161*, 2633–2643. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Varsani, A.; Krupovic, M. Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family Genomoviridae. *Virus Evol.* **2017**, *3*, vew037. [\[CrossRef\]](#)
39. Heyraud-Nitschke, F.; Schumacher, S.; Laufs, J.; Schaefer, S.; Schell, J.; Gronenborn, B. Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic Acids Res.* **1995**, *23*, 910–916. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Chandler, M.; de la Cruz, F.; Dyda, F.; Hickman, A.B.; Moncalian, G.; Ton-Hoang, B. Breaking and joining single-stranded DNA: The HUH endonuclease superfamily. *Nat. Rev. Microbiol.* **2013**, *11*, 525–538. [\[CrossRef\]](#)
41. Ilyina, T.V.; Koonin, E.V. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeabacteria. *Nucleic Acids Res.* **1992**, *20*, 3279–3285. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Nash, T.E.; Dallas, M.B.; Reyes, M.I.; Buhrman, G.K.; Ascencio-Ibañez, J.T.; Hanley-Bowdoin, L. Functional Analysis of a Novel Motif Conserved across Geminivirus Rep Proteins. *J. Virol.* **2011**, *85*, 1182–1192. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Eilebrecht, S.; Hotz-Wagenblatt, A.; Sarachaga, V.; Burk, A.; Falida, K.; Chakraborty, D.; Nikitina, E.; Tessmer, C.; Whitley, C.; Sauerland, C.; et al. Expression and replication of virus-like circular DNA in human cells. *Sci. Rep.* **2018**, *8*, 2851. [\[CrossRef\]](#)
44. Borghese, A. Situation and perspectives of buffalo in the world, Europe and Macedonia. *Maced. J. Anim. Sci.* **2011**, *1*, 281–296.
45. Habib, M.R.; Haque, M.N.; Rahman, A.; Aftabuzzaman, M.; Ali, M.M.; Shahjahan, M. Dairy buffalo production scenario in Bangladesh: A review. *Asian J. Med. Biol. Res.* **2017**, *3*, 305–316. [\[CrossRef\]](#)
46. Government of India. Department of Animal Husbandry and Dairying. Annual Report (2018–2019). Available online: <https://dahd.nic.in/documents/reports> (accessed on 13 November 2020).
47. Borghese, A.; Moioli, B. Buffalo: Mediterranean Region. In *Reference Module in Food Science*; Elsevier: Amsterdam, The Netherlands, 2016. [\[CrossRef\]](#)
48. Borghese, A.; Mazzi, M.; Moioli, B.; Barile, V.L.; Terzano, G.M.; Terramoccia, S.; Bartocci, S.; Puppo, S.; Tripaldi, C.; Allegrini, S.; et al. *Buffalo Production and Research*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 2005; ISSN 1020-3737.
49. Lechmann, J.V. A Pilot Study Contributing to the Virome of Water Buffaloes in Switzerland. Ph.D. Thesis, Universität Zürich, Zürich, Switzerland, 2017.
50. Guo, Z.; He, Q.; Tang, C.; Zhang, B.; Yue, H. Identification and genomic characterization of a novel CRESS DNA virus from a calf with severe hemorrhagic enteritis in China. *Virus Res.* **2018**, *255*, 141–146. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Ling, Y.; Zhang, X.; Qi, G.; Yang, S.; Jingjiao, L.; Shen, Q.; Wang, X.; Cui, L.; Hua, X.; Deng, X.; et al. Viral metagenomics reveals significant viruses in the genital tract of apparently healthy dairy cows. *Arch. Virol.* **2019**, *164*, 1059–1067. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Wang, H.; Li, S.; Mahmood, A.; Yang, S.; Wang, X.; Shen, Q.; Shan, T.; Deng, X.; Li, J.; Hua, X.; et al. Plasma virome of cattle from forest region revealed diverse small circular ssDNA viral genomes. *Virol. J.* **2018**, *15*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Wen, L.; Mao, A.; Fan, Z.; Li, W.; Xiao, Q.; Liu, Q.; Xie, J.; He, K. Porcine circovirus-like virus P1 in cattle, goats and rabbits in China. *Transbound. Emerg. Dis.* **2018**, *65*, e217–e218. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Kim, H.K.; Park, S.J.; Nguyen, V.G.; Song, D.S.; Moon, H.J.; Kang, B.K.; Park, B.K. Identification of a novel single-stranded, circular DNA virus from bovine stool. *J. Gen. Virol.* **2012**, *93*, 635–639. [\[CrossRef\]](#)

## Supplementary Material



**Figure S1.** Maximum Likelihood phylogenetic tree of novel hypothetical protein sequences and uncharacterized BMMF1 and 2 proteins calculated as described in Fig. 1. The bootstrap consensus tree is based on 500 replicates. Bootstrap values at the branch points indicate the percentage of replicate trees in which respective isolates clustered. Branch support values lower than 70 % were not included. Blue dots mark all sequences found in this study. The scale bar represents the number of substitutions per site. All taxa are indicated by name followed by the corresponding GenBank accession number.

**Table S1.** Information on the primer-sets applied in the PCR-screening for the amplification of Sphinx-like sequences. Wobble bases are indicated according the universal code: R = A/G, Y = C/T, M = A/C, K = G/T, S = C/G, W = A/T, B = C/G/T, D = A/G/T, H = A/C/T, V = A/C/G, and N = A/C/G/T. The annealing temperature was 55 °C and the PCR comprised 35 cycles. bp = base pairs. \* = this study

Amplon name	Primer-set*	Amplon length in bp	For (F) Rev (R)	Sequence 5'-3'	Sphinx-type
S1A	S1F + S1.1R + S1.2R	233	S1F	AGATAACGCCCTAATGAACGCT	Sphinx 1.76
			S1.1.R	TCGTCCACGTTCTCGCTTT	
			S1.2.R	TCGTCCACGCTCACGTTTT	
S1B	S1F + S1.3R + S1.4R	314	S1F	AGATAACGCCCTAATGAACGCT	Sphinx 1.76
			S1.3R	AACGGTTGCTGTATCGTCCA	
			S1.4R	AACTGTTGCCGTTTCATCAA	
S1C	S1F + S1.5R + S1.6R	339	S1F	AGATAACGCCCTAATGAACGCT	Sphinx 1.76
			S1.5R	CAACCGCAGGGGCAAAATAA	
			S1.6R	CAACCGCAGGGGCAAATATTA	
S2A	S2F + S2.1R + S2.2R	346	S2F	AACGCCCTAATGAACGCTAGTT	Sphinx 2.36
			S2.1R	AATCAGAGGAACAACCGCAGG	
			S2.2R	AATCARWGGAAACAACCGCAGG	
S3A	S3F + S1.3R + S1.4R	308	S3F	CGCCCTAATGAACGCTAGTT	Sphinx 1.76
			S1.3R	AACGGTTGCTGTATCGTCCA	
			S1.4R	AACTGTTGCCGTTTCATCAA	
S3B	S3F + S3.1R + S3.2R	454	S3F	CGCCCTAATGAACGCTAGTT	Sphinx 1.76
			S3.1R	CCTGTGCTACGCCAACAAAT	
			S3.2R	CCTGTTGTTCTCCAGCAGAT	
S4A	S4F + S4.1R + S4.2R	351	S4F	GAAGCGAGAGAAAACAGGCAAAG	Sphinx 2.36
			S4.1R	CGAACAGCATAAGCACTCGATA	
			S4.2R	CGRACAGCATADGCACTCGATA	
S5A	S5.1F + S5.2F + S5.1R + S5.2R	334	S5.1F	GGSAAGGGATTAATGCMIAATGATC	Sphinx 2.36
			S5.2F	GGSAAGGAATTAAATGCAAATGAC	
			S5.1R	TAAGCGAATKGCTAACGCRC	
			S5.2R	TAGACGGACGGCATAGGCAC	
S6A	S6F + S6R	310	S6F	TTGGCGWASTACAGGYAAA	Sphinx 2.36
			S6R	TWCGGAATAATGMCGTTGAG	
S7A	S7F + S7.1R + S7.2R + S7.3R	152	S7F	ARTGGCATCCGCATATYCAT	
			S7.1R	CAGCTTGCTATAGCCGTGT	
			S7.2R	CAGCCTTGCTATACCGTGT	
			S7.3R	CCGCTTGCTATAACCGTAA	

**Table S2.** Information on the primer-sets applied in the PCR-screening for the amplification of *Genomoviridae* family members. Primers highlighted in green are the only ones resulting in amplicons during screening PCR and subsequently in the generation of one full length sequence. Wobble bases are indicated according the universal code: R = A/G, Y = C/T, M = A/C, K = G/T, S = C/G, W = A/T, B = C/G/T, D = A/G/T, H = A/C/T, V = A/C/G, and N = A/C/G/T. The annealing temperature was 55 °C and the PCR comprised 35 cycles. bp = base pairs. \* = this study

Amplicon name	Primer-set*	Amplicon length in bp	For (F) Rev (R)	Sequence 5'-3'	Genus
G1A	G1F + G1R	414	G1F	TGGAGGTACTCAYCTCCAYGYTTT	<i>Gemycircularvirus</i>
			G1R	CCARCTSAGGTACCATCCM	
G2A	G2F + G2R	692	G2F	GAGCTSGAAWKGTCACCTGA	<i>Gemycircularvirus</i>
			G2R	CTACKCTRTATTGGCACGAAARAT	
G5A	G5F + G5R	643	G5F	GGGATTGGTCCATGAACGT	<i>Gemycircularvirus</i>
			G5R	GSGGAGTTWACCTCACRAAAATACAAT	
G6A	G6F + G5R	643	G6F	GGGATYGGTRCATGASTTG	<i>Gemygorvirus</i>
			G5R	GSGGAGTTWACCTCACRAAAATACAAT	
G8A	G8F + G8R	359	G8F	TCCATGGAGAGGCTACGACTAT	<i>Gemykibivirus</i>
			G8R	CTGACGGAGTACCTATGAGTG	
G12A	G12F + G11.2R	737	G12F	ACCCGAACATTRCRCCATGC	<i>Gemygorvirus</i>
			G11.2R	CACCCYTTGTAvgAAGGRAA	
G13A	G13F + G13R	541	G13F	GATGGYGGAACATCTRCACG	<i>Gemykibivirus</i>
			G13R	TAGGGTTAGGGTTAGGGTTAGG	
G14A	G14F + G13R	541	G14F	GATGGTGGGACGCATCTACATG	<i>Gemykibivirus</i>
			G13R	TAGGGTTAGGGTTAGGGTTAGG	
G17A	G17F + G17R	969	G17F	CGGRTTGGTCGAGAGCTACA	<i>Gemykolovirus</i>
			G17R	CTTGGATCCYTGTGCAKRT	
G18A	G18F + G18R	352	G18F	ACTGCGTTGACTACTTGGAG	<i>Gemykrogvirus</i>
			G18R	CACTCGGAGATCTCAGGATAGC	
G19A	G19F + G19R	428	G19F	TATTCAAGGCTCTTCGAGT	<i>Gemykronavirus</i>
			G19R	GCCCGACAGTAAGAGACGAG	
G20A	G20F + G20R	347	G20F	CTTGGTTCAGAGAGCGGGTCC	<i>Gemytondavirus</i>
			G20R	GGGCCGTATTCTGAGACGTA	
G21A	G21F + G21R	794	G21F	AGTCTCGCGAACCTCGTTA	<i>Gemyvongvirus</i>
			G21R	AACACCTCCATCCACTCTG	

**Table S3.** Information on the primer-sets applied in the inverse PCR for amplification of full-length DNA. \* = this study

Amplicon name	Primer-set*	For (F) Rev (R)	Sequence 5'-3'	Annealing temperature	Target-DNA
S7A_RSL	S7A_RSLfor + S7A_RSLrev	S7A_RSLfor	CCAAAGAACACGGCTATAG	48 °C	<i>Sphinx 2.36</i>
		S7A_RSLrev	AAAATATGAATATGCGGATGC	48 °C	
G_RSL1	G_RSLfor1 + G_RSLrev1	G_RSLfor1	AACATGGAGATGAGTACC	49°C	<i>Gemycircularvirus</i>
		G_RSLrev1	AGTTTCTCGACTGTAC	49°C	
G_RSL_Rest	G_RSL_Rest_F + G_RSL_Rest_R	G_RSL_Rest_F	GATCATCTCCAAGAGACTC	53°C	<i>Gemycircularvirus</i>
		G_RSL_Rest_R	ATCAAGTGAGGAGGATACTG	53°C	

**Table S4.** Summary of the BLASTn results for the individual isolates. Only the two best matching results according to the query coverage and the percentage of identity are listed. The last column indicates the number of missing (-) or additional (+) nucleotides in the novel sequences compared to the GeneBank hits. Background colours highlight the affiliation to the BMMF groups 1 to 3. bp = base pairs; nt = nucleotides.

Isolate	Accession Number	Length in bp	Best Hit(s) in BlastSearch	Gene Bank Acc. No. of Hit	Length in bp	% Query Coverage	% ID	Δ nt to blast Hit
8BAMI.2522	MW828657	2522	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,6	0
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,2	-1
9BAMI.2522	MW828659	2522	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,7	0
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,3	-1
11BAMI.2092	MW828660	2092	BMMF1 DNA sequence, isolate C1MI.15M.2	LR215495.1	2041	72	77,2	51
			BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499.1	1765	71	85,2	327
12BAMI.2092	MW828661	2092	BMMF1 DNA sequence, isolate C1MI.15M.2	LR215495.1	2041	72	77,2	51
			BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499.1	1765	71	85,2	327
13BAMI.2092	MW828663	2092	BMMF1 DNA sequence, isolate C1MI.15M.2	LR215495.1	2041	72	77,2	51
			BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499.1	1765	71	85,2	327
14BAMI.2522	MW828664	2522	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,4	0
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,1	-1
15BAMI.2522	MW828665	2522	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,7	0
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,4	-1
17BAMI.2522	MW828666	2522	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,6	0
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,3	-1
18BAMI.2184	MW828667	2184	<i>Acinetobacter</i> sp. WCHAc010034 plasmid p10_010034, complete sequence	CP032267.1	3470	100	98,5	-1286
			Sphinx1.76-related DNA, replication competent episomal DNA CMI2.214	LK931488.1	2148	75	97,2	36
21BAMI.2091	MW828668	2091	BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499.1	1765	71	85,4	326
22BAMI.2522	MW828670	2522	Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,8	-1
			Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,3	0
22BAMI.2521	MW828671	2521	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493.1	2522	100	99,6	-1
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487.1	2523	100	99,3	-2
24BAMI.2092	MW828672	2092	BMMF1 DNA sequence, isolate C1MI.15M.2	LR215495.1	2041	72	77,1	51
			BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499.1	1765	71	85,3	327
25BAMI.1972	MW828673	1972	Sphinx1.76-related DNA, replication competent episomal DNA CMI2.214	LK931488.1	2148	84	82,5	-176
			<i>Acinetobacter</i> sp. WCHAc010034 plasmid p10_010034, complete sequence	CP032267.1	3470	84	83,2	-1498
8BAMI.2351	MW828658	2351	BMMF2 DNA sequence, isolate C2MI.15B.12	LR215564.1	2362	100	95,4	-11
			BMMF2 DNA sequence, isolate C2MI.15B.3	LR215555.1	2362	100	95,4	-11
12BAMI.2349	MW828662	2349	BMMF2 DNA sequence, isolate C2MI.15B.3	LR215555.1	2362	100	92,6	-13
			BMMF2 DNA sequence, isolate C2MI.15B.4	LR215556.1	2362	100	92,5	-13
21BAMI.2076	MW828669	2076	BMMF2 DNA sequence, isolate C2MI.9A.1	LR215538.1	2405	92	93,1	-327
			BMMF2 DNA sequence, isolate C2MI.1A.1	LR215583.1	2293	91	93,9	-215
44BAMI.2371	MW828674	2371	BMMF2 DNA sequence, isolate C2MI.4B.3	LR215591.1	2661	78	84,3	-290
			BMMF2 DNA sequence, isolate C2MI.4A.2	LR215590.1	2661	78	84,1	-290
11BAMI.2199	MW828677	2199	Giant panda associated gemycircularvirus strain ggpe010, complete genome	MF327567.1	2216	52	92,8	-17
			Blackfly gemnovirus 4 isolate SF02_836, complete genome	MK433236.1	2181	55	80,2	18

**Table S5.** Summary of the BLASTp results of the *in silico* generated Reps. Only the two best matching results according to the query coverage and the percentage of identity are listed. The last column indicates the number of missing (-) or additional (+) amino acids in the novel proteins compared to the GeneBank hits. Background colours highlight the affiliation to the BMMF groups 1 to 3. Isolates with overlapping rep-ORFs substantially shorter than all best Blast Hits and thus potentially incapable of replication are in red. aa = amino acids.

Isolate	Start codon	Length in aa	Best Hit(s) in BlastSearch	Gene Bank Acc. No. of Hit	Length in aa	% Query Coverage	% ID	Δ aa to blast Hit
8BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,69 98,36	0 10
9BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,08 98,03	0 10
11BAMI.2092	ATG	327	unnamed protein product [Brugia timori] MULTISPECIES: replication initiation protein [Acinetobacter]	VDO15958.1 WP_087739066.1	318 318	96 96	85,67 85,67	9 9
12BAMI.2092	ATG	327	unnamed protein product [Brugia timori] MULTISPECIES: replication initiation protein [Acinetobacter]	VDO15958.1 WP_087739066.1	318 318	96 96	85,13 85,13	9 9
13BAMI.2092	ATG	327	unnamed protein product [Brugia timori] MULTISPECIES: replication initiation protein [Acinetobacter]	VDO15958.1 WP_087739066.1	318 318	96 96	85,13 85,13	9 9
14BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,69 98,36	0 10
15BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	100 98,68	0 10
17BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	100 98,68	0 10
18BAMI.2184	ATG	197	MULTISPECIES: replication initiation protein [Acinetobacter] replication initiation protein [Acinetobacter ursingii]	WP_120385393.1 WP_151825680.1	324 324	100 100	94,42 94,42	-127 -127
21BAMI.2091	ATG	316	MULTISPECIES: replication initiation protein [Acinetobacter] unnamed protein product [Brugia timori]	WP_087739066.1 VDO15958.1	318 318	100 100	85,13 85,13	-2 -2
22BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,39 98,03	0 10
22BAMI.2521	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,39 98,03	0 10
24BAMI.2092	ATG	327	unnamed protein product [Brugia timori] MULTISPECIES: replication initiation protein [Acinetobacter]	VDO15958.1 WP_087739066.1	318 318	96 96	84,81 84,81	9 9
25BAMI.1972	ATG	334	replication initiation protein [Acinetobacter junii] RepB family plasmid replication initiator protein [Acinetobacter sp. YH16058]	WP_129264380.1 WP_179991797.1	324 324	96 96	87,04 87,35	10 10
45BAMI.2521	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,69 98,36	0 10
46BAMI.2522	ATG	326	replication protein [Sphinx1.76-related DNA] replication protein [Sphinx1.76-related DNA]	CDS63400.1 CDS63392.1	326 316	100 93	99,69 98,68	0 10
8BAMI.2351.Repl	TTG	341	replication protein [BMMF2 DNA sequence] replication protein [BMMF2 DNA sequence]	VEV85780.1 VEV85785.1	278 341	63 92	99,08 98,09	63 0
8BAMI.2351.ReplI	ATG	231	replication protein [BMMF2 DNA sequence] replication protein [BMMF2 DNA sequence]	VEV85786.1 VEV85785.1	307 341	100 100	98,7 98,7	-76 -110
12BAMI.2349.Repl	TTG	238	protein rep [Acinetobacter cumulans] protein rep [Acinetobacter sp. YH01009]	WP_121595097.1 WP_180190657.1	341 344	97 97	90,09 90,09	-103 -106
12BAMI.2349.ReplI	ATG	128	replication protein [BMMF2 DNA sequence] replication protein [BMMF2 DNA sequence]	VEV85781.1 VEV85780.1	244 278	95 95	100 100	-116 -150
21BAMI.2076	ATG	365	protein rep [Acinetobacter sp. YH01006] protein rep [Acinetobacter sp. YH01024]	WP_180191664.1 WP_180174131.1	346 349	92 92	96,45 85,25	19 16
44BAMI.2371	TTG	434	protein rep [Acinetobacter baumannii] replication protein [BMMF2 DNA sequence]	MBE2397913.1 VEV85558.1	293 434	67 93	97,27 95,06	141 0
11BAMI.2199	ATG	234	putative ORF3 protein [Giant panda associated gemycircularvirus] replication-associated protein [Giant panda associated gemycircularvirus]	ASH99165.1 ASH99166.1	201 277	40 83	93,68 91,79	33 -43

**Table S6.** Summary of the BLASTp results of the *in silico* generated additional proteins. At most the two best matching results according to the query coverage and the percentage of identity are listed. The last column indicates the number of missing (-) or additional (+) amino acids in the novel proteins compared to the GeneBank hits. Background colours highlight the affiliation to the BMMF groups 1 to 3. aa = amino acids.

Isolate	Start codon	Length in aa	Best Hit(s) in BlastSearch	Gene Bank Acc. No. of Hit	Length in aa	% Query Coverage	% ID	Δ aa to blast Hit
8BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1	219	100	100	0
9BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1	219	100	100	0
11BAMI.2092	TTG	101	hypothetical protein [Acinetobacter sp. ANC 3813] hypothetical protein [Acinetobacter guerrei]	WP_086212311.1 WP_151979748.1	102 98	94 94	69,47 64,21	-1 3
12BAMI.2092			n. d.					
13BAMI.2092	TTG	106	hypothetical protein [Acinetobacter sp. ANC 3813] hypothetical protein [Acinetobacter guerrei]	WP_086212311.1 WP_151979748.1	102 98	95 94	68,32 64	4 8
14BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA] MULTISPECIES: hypothetical protein [unclassified Psychrobacter]	CDS63393.1 WP_025643765.1	219 221	100	99,54 69,41	0 -2
15BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA] uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1 CDS63403.1	219 115	100 52	99,54 99,13	0 104
17BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA] uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1 CDS63403.1	219 115	100 52	99,54 99,13	0 104
18BAMI.2184	ATG	149	MULTISPECIES: PH domain-containing protein [Acinetobacter] PH domain-containing protein [Acinetobacter sp. WCHAc010034]	WP_086231555.1 WP_068913323.1	149 135	100 90	100 99,26	0 14
21BAMI.2091	TTG	106	hypothetical protein [Acinetobacter sp. ANC 3813] hypothetical protein [Acinetobacter guerrei]	WP_086212311.1 WP_151979748.1	102 98	95 93	68,32 63,64	4 8
22BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA] uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1 CDS63403.1	219 115	100 52	100 100	0 104
22BAMI.2521	ATG	206	uncharacterized protein [Sphinx1.76-related DNA] uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1 CDS63403.1	219 115	100 49	100 100	-13 91
24BAMI.2091	TTG	106	hypothetical protein [Acinetobacter sp. ANC 3813] hypothetical protein [Acinetobacter guerrei]	WP_086212311.1 WP_151979748.1	102 98	95 94	68,32 63	4 8
25BAMI.1972			n. d.					
45BAMI.2521	ATG	219	uncharacterized protein [Sphinx1.76-related DNA] uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1 CDS63403.1	219 115	100 52	100 100	0 104
46BAMI.2522	ATG	219	uncharacterized protein [Sphinx1.76-related DNA] uncharacterized protein [Sphinx1.76-related DNA]	CDS63393.1 CDS63403.1	219 115	100 52	99,54 99,13	0 104
8BAMI.2351.ORF2	ATG	131	uncharacterized protein [BMMF2 DNA sequence] uncharacterized protein [BMMF2 DNA sequence]	VEV85783.1 VEV85627.1	138 138	100 100	100 100	-7 -7
8BAMI.2351.ORF3	TTG	97	uncharacterized protein [BMMF2 DNA sequence] uncharacterized protein [BMMF2 DNA sequence]	VEV85789.1 VEV85660.1	97 97	100 98	100 100	0 0
12BAMI.2349	ATG	131	uncharacterized protein [BMMF2 DNA sequence] uncharacterized protein [BMMF2 DNA sequence]	VEV85783.1 VEV85627.1	138 138	100 100	100 100	-7 -7
21BAMI.2076	ATG	98	hypothetical protein [Acinetobacter schindleri] hypothetical protein [Acinetobacter baumannii]	WP_108619725.1 WP_198064358.1	98 98	100 100	100 98,98	0 0
44BAMI.2371	ATG	97	uncharacterized protein [BMMF2 DNA sequence] uncharacterized protein [BMMF2 DNA sequence]	VEV85561.1 VEV85557.1	99 99	100 100	98,97 98,97	-2 -2
11BAMI.2199.ORF3	TTG	214	putative ORF3 protein [Giant panda associated gemycircularvirus] replication-associated protein [Giant panda associated gemycircularvirus]	ASH99165.1 ASH99166.1	201 277	40 83	93,68 91,79	13 -63
11BAMI.2199.ORF4	TTG	131	PspC domain-containing protein [Streptomyces albidochromogenes] PspC domain-containing protein [Streptomyces chrysaeus]	WP_138901707.1 WP_138906643.1	449 449	15 15	60 60	-318 -318

## 2. **Publikation II**

### **Titel**

Identification and Characterization of Circular Single-Stranded DNA Genomes in  
Sheep and Goat Milk

### **Autoren**

Marie-T. König<sup>1</sup>, Robert Fux<sup>2</sup>, Ellen Link<sup>2</sup>, Gerd Sutter<sup>2</sup>, Erwin Märtlbauer<sup>1</sup>,  
Andrea Didier<sup>1</sup>

<sup>1</sup> Department für Veterinärwissenschaften, Institut für Lebensmittelwissenschaften, Veterinärmedizinische  
Fakultät, Ludwig-Maximilians-Universität München, Schönleutnerstr. 8, 85764 Oberschleißheim,  
Deutschland

<sup>2</sup> Department für Veterinärwissenschaften, Lehrstuhl für Virologie, Veterinärmedizinische Fakultät, Ludwig-  
Maximilians-Universität München, Veterinärstraße 13, 80539 München, Deutschland

### **Journal**

Viruses (IF 2021: 5,048)

### **Status**

akzeptiert am 27.10.2021

publiziert online am 28.10.2021

Viruses 2021, 13(11), 2176

doi:10.3390/v13112176



Article

## Identification and Characterization of Circular Single-Stranded DNA Genomes in Sheep and Goat Milk

Marie-Thérèse König <sup>1</sup>, Robert Fux <sup>2</sup>, Ellen Link <sup>2</sup>, Gerd Sutter <sup>2</sup>, Erwin Märthbauer <sup>1</sup> and Andrea Didier <sup>1,\*</sup>

<sup>1</sup> Department of Veterinary Sciences, Faculty of Veterinary Medicine, Institute of Food Safety, Ludwig-Maximilians-Universität München, 85764 Oberschleißheim, Germany; marie-therese.koenig@mh.vetmed.uni-muenchen.de (M.-T.K.); e.maerthbauer@mh.vetmed.uni-muenchen.de (E.M.)

<sup>2</sup> Chair of Virology, Department of Veterinary Sciences, Faculty of Veterinary Medicine, Ludwig-Maximilians-Universität München, 80539 München, Germany; Robert.Fux@viro.vetmed.uni-muenchen.de (R.F.); Ellen.Link@viro.vetmed.uni-muenchen.de (E.L.); sutter@viro.vetmed.uni-muenchen.de (G.S.)

\* Correspondence: A.Didier@mh.vetmed.uni-muenchen.de; Tel.: +49-89-2180-78571; Fax: +49-89-2180-78602

**Abstract:** In recent years, a variety of circular replicase-encoding single-stranded (CRESS) DNA viruses and unclassified virus-like DNA elements have been discovered in a broad range of animal species and environmental samples. Key questions to be answered concern their presence in the human diet and their potential impact on disease emergence. Especially DNA elements termed bovine meat and milk factors (BMMF) are suspected to act as co-factors in the development of colon and breast cancer. To expand our knowledge on the occurrence of these potential pathogens in human nutrition, a total of 73 sheep and 40 goat milk samples were assayed by combining rolling circle amplification (RCA), PCR and Sanger sequencing. The present study further includes retail milk from the aforementioned species. We recovered 15 single stranded (ss) circular genomes. Of those, nine belong to the family *Genomoviridae* and six are members of the unclassified group of BMMF. Thus, dairy sheep and goats add to dispersal of CRESS viruses and circular ssDNA elements, which enter the food chain via milk. The presence of these entities is therefore more widespread in *Bovidae* than initially assumed and seems to be part of the common human nutrition.

**Keywords:** BMMF; circular ssDNA; colon/breast cancer; *Cressnaviricota*; *Genomoviridae*; milk; small ruminants



**Citation:** König, M.-T.; Fux, R.; Link, E.; Sutter, G.; Märthbauer, E.; Didier, A. Identification and Characterization of Circular Single-Stranded DNA Genomes in Sheep and Goat Milk. *Viruses* **2021**, *13*, 2176. <https://doi.org/10.3390/v13112176>

Academic Editor:  
Thomas Stamminger

Received: 7 October 2021

Accepted: 27 October 2021

Published: 28 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

In recent years, a considerable number of small circular replicase-encoding single stranded DNA (CRESS) viruses [1–3] and unclassified virus-like DNA molecules [4–6] have been recovered from numerous animal species as well as from environmental samples. The *Cressnaviricota* encompassing eight families represent a novel, rapidly growing phylum of these DNA viruses [7]. One of them—the *Genomoviridae*—comprises 10 genera isolated from all domains of life as well as from environmental samples [8]. The number of species within this family steadily increases and there are numerous uncultivated isolates in the databases, which are still unclassified. Their genomes are approximately 2–2.4 kb in length and encode at least two proteins, which are: (i) a rolling circle replication initiation protein (Rep) and (ii) a capsid protein (CP). Pairwise comparison of CP sequences revealed a higher degree of divergence compared to Rep. Assignment to a genus is therefore mainly based on the amino acid (aa) sequence of Rep [9]. These Rep proteins initiate rolling circle replication (RCR) of the viral genomes [10]. In addition to studies focussing on the molecular mechanism of replication, some studies describe the occurrence of circular single stranded (ss) DNAs in livestock and foods of animal origin. Investigations from China reported on the characterization of CRESS viral genomes from the genital tract and blood of cattle [11,12]. Viral genomes and CRESS DNAs have been found in various meat samples in Porto Alegre,

Brazil and in San Francisco, USA. However, in these reports the detection was not linked to human diseases [13,14]. In contrast, multiple unclassified circular viral-like DNA elements, isolated from raw milk and cow milk at retail in Germany are suspected to contribute to the emergence of colon and breast cancer and neurodegenerative diseases [5,15–17]. These DNAs have been termed “BMMF” (bovine meat and milk factors) and are currently assigned to BMMF groups 1–4 according to their molecular characteristics [15]. Often a certain strain is synonymously described by abbreviations like HCBI (healthy cattle blood isolate) and CMI (cow milk isolate) thus pointing to the matrix from which it has been originally isolated. The majority of isolates is assorted to BMMF group 1 and 2, showing similarities to Sphinx 1.76 and Sphinx 2.36 DNA initially described by Manuelidis [4]. A few recovered isolates have been identified as genomoviruses and were assigned to BMMF group 3. Group 4 embodies a single isolate exhibiting similarities to a *Psychrobacter* sp. plasmid. BMMF group 1 and 2 members lack a CP and therefore differ significantly from group 3. All BMMF contain *rep*-genes and apparently replicate via RCR. In addition to a potential hairpin structure as a putative origin of replication, all BMMF1 genomes feature tandem repeats (TR) in proximity to *rep* [17]. These attributes underpin the theory on the phylogeny of circular ssDNAs, which are supposed to have evolved from bacterial plasmids via RNA virus gene transfer [18,19]. Therefore, borders between CRESS viruses, plasmids and phages sometimes blur.

From a disease related point of view, latest studies demonstrated the presence of BMMF encoded antigens in colorectal peritumor and tumor tissue and therefore support the idea of their involvement in colorectal cancerogenesis [20]. Thus far, the isolation of *Cressdnnaviricota* and BMMF has been reported from meat, milk and dairy products of “taurine” cattle descending from the European aurochs. Despite the potential correlation between these DNAs and human cancer [21–23], data from other *Bovidae* are still missing. Beside “taurine” cattle, “zebuine” cattle, water buffaloes and small ruminants contribute to the worldwide milk supply [24]. In a recently published study, we showed the occurrence of circular ssDNA and a gemycircularvirus in water buffalo milk [25]. Data from studies focussing on domestic livestock are scarce and display a still narrow basis for a risk assessment concerning potential adverse effects on human health.

The present study highlights the recovery and characterization of circular DNA genomes in sheep and goat milk, thereby demonstrating a considerably more frequent occurrence of genomoviruses and BMMF in milk of dairy *Bovidae* than assumed so far.

## 2. Methods

### 2.1. Sample Collection

A total of 113 samples of sheep and goat milk were collected from eight different flocks in Germany. Five different sheep farmers contributed altogether 73 samples (flock A:  $n = 20$ , flock B:  $n = 15$ , flock C:  $n = 15$ , flock D:  $n = 3$ , flock E:  $n = 20$ ) and three goat farmers added 40 further milk samples to the study (flock F:  $n = 4$ , flock G:  $n = 20$ , flock H:  $n = 16$ ). Animals were kept in Hesse, Lower Saxony, North Rhine-Westphalia, Baden-Württemberg and Bavaria. After collection, that was carried out during the daily milking routine, the milk was cooled to 4 °C and cold-shipped to the Institute of Food Safety within 48 h. Additionally, commercially available sheep milk ( $n = 6$ ) and goat milk ( $n = 6$ ) produced in Germany, Austria and Spain was bought in German supermarkets.

### 2.2. DNA Extraction and RCA

DNA was extracted from 200 µL milk using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. DNA concentration and purity (i.e. OD260/OD280 ratio) were determined on a DeNovix DS-11 FX (DeNovix, Wilmington, DE, USA) spectrophotometer. Rolling circle amplification (RCA) of approx. 40 ng DNA was performed by the TempliPhi Amplification Kit (Cytiva, Marlborough, MA, USA) with random primers according to the manufacturer’s instructions. Isothermal amplification time at 30 °C was set to 18 h.

### 2.3. Recovery of Viral-like DNA Elements

Initial screening of RCA products for the presence of Sphinx-like DNA included three primer pairs designed via the *rep*-gene of Sphinx 1.76 (GenBank Acc. No. HQ444404.1; accessed in September 2021) and one designed via the *rep*-gene of Sphinx 2.36 (GenBank Acc. No. HQ444405.1; accessed in September 2021). A similar approach included 13 primer pairs for detection of genomoviruses. All PCRs were run in a volume of 50  $\mu$ L on a TProfessional Gradient 96 Thermocycler (Biometra, Jena, Germany). ThermoPrime Plus Polymerase (ThermoScientific, Waltham, MA, USA) at 3 U/reaction was applied. The final concentration of dNTPs (ThermoScientific, Waltham, MA, USA) of 200  $\mu$ M each. Primer sequences as well as the individual amplification programs are summarized in Supplementary Table S1 (sheets 1 + 2). PCR products were purified from agarose gels with HighYield PCR Purification/Gel Extraction Kit (SLG, Südlaborbedarf, Gauting, Germany) and subjected to Sanger sequencing on both strands. MWG Eurofins (Ebersberg, Germany) operated all sequencing reactions performed in the present study. PCR screening sequencing results were used for inverted primer design to obtain full-length sequences of the circular DNA elements (Supplementary Table S1, sheet 3). Additional PCRs including abutting primers for amplification of BMMF2 sequences published by de Villiers et al. [15] and for BMMF1 sequences published by Whitley et al. [16] were carried out.

### 2.4. Cloning and Full-Length Sequencing

Inverse, gel-purified PCR products were ligated with pCR2.1-TopoTA vector (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Afterwards, chemically competent *Escherichia coli* DH5 alpha were transformed and the resulting pellet was plated on LB-agar with ampicillin (100  $\mu$ g/mL). Five clones from each plate were subjected to MiniPrep Plasmid isolation via GenElute Plasmid Miniprep Kit (Sigma Aldrich, Taufkirchen, Germany). After control digest of 500 ng plasmid DNA with EcoRI (New England BioLabs, Ipswich, MA, USA) insert-bearing clones were sequenced via the M13rev-29 and M13uni-21 primer binding sites present in the vector.

### 2.5. Data Analysis

All full-length sequences were subjected to BlastSearch against the nucleotide database at NCBI [26]. Multiple sequence alignments and phylogenetic trees were computed in MEGA software (v. 10.2.6) by the means of the Maximum likelihood method. Bootstrap values were computed with 500 replicates [27]. SDT v1.2 was utilised for displaying pairwise genome identity scores of novel genomoviruses calculated from pairwise alignments generated by MUSCLE [28]. Tandem repeats [29] and inverted repeats [30] upstream the *rep*-gene sequences were checked at Emboss explorer. Maximum size of tandem repeats was restricted to 30 nucleotides (nt) and the minimum size of inverted repeats was set to 4 nt. ORFfinder software (version 1.3.0) at NCBI [31] served for in silico detection of potential open reading frames. Illustrations were plotted with SeqBuilder Pro of the DNASTAR software (Lasergene Inc., v. 17.1. DNASTAR. Madison, WI, USA) [32]. ATG or TTG as alternative start codons were allowed. The minimum ORF size was limited to 75 amino acids (aa). Only ORFs  $\geq$  95 aa were included in further comparative and functional analysis. Putative intron acceptor-/donor sites and conserved amino acid motifs were identified manually. All annotated genomes from this study were deposited in the GenBank database (Accession numbers OK148616-OK148630).

## 3. Results and Discussion

### 3.1. General Data Analysis and Interpretation

The present study aimed to detect and characterize circular single stranded DNAs (CRESS viruses and BMMF) in sheep and goat milk, thus expanding the current knowledge on the occurrence of these factors in animal-based foods. To cover this topic, individual milk samples from flocks kept in different German provinces as well as milk available at retail were assayed. Grocery milk packages originate from Germany, Austria and Spain.

Table 1 summarizes the overall information on the origin and basic features of circular DNA elements detected.

**Table 1.** Overview of all circular genomes detected in this study with accession numbers, sequence length, DNA-types, species and origin of isolation. Interestingly, two genomoviral sequences were isolated from sheep No. 115.

Isolate	Accession Number	Length in Nucleotides (nt)	DNA-Type	Species	Origin
SmGV1	OK148616	2191	Genomovirus	Sheep No. 109	Flock B
SmGV2	OK148617	2192	Genomovirus	Sheep No. 110	Flock B
SmGV3	OK148618	2190	Genomovirus	Sheep No. 111	Flock B
SmGV4	OK148619	2190	Genomovirus	Sheep No. 113	Flock B
SmGV5	OK148620	2191	Genomovirus	Sheep No. 115	Flock B
SmGV6	OK148621	2124	Genomovirus	Sheep No. 115	Flock B
SmGV7	OK148622	2190	Genomovirus	Sheep No. 117	Flock B
SmGV8	OK148623	2125	Genomovirus	Sheep No. 151	Flock C
GmGV9	OK148624	2122	Genomovirus	Goat	Retail
SmI1	OK148625	2036	Sphinx 1.76-like (BMMF1)	Sheep No. 153	Flock C
GmI2	OK148626	2092	Sphinx 1.76-like (BMMF1)	Goat	Retail
SmI3	OK148627	2522	Sphinx 1.76-like (BMMF1)	Sheep	Retail
SmI4	OK148628	2093	Sphinx 1.76-like (BMMF1)	Sheep	Retail
GmI5	OK148629	1765	Sphinx 1.76-like (BMMF1)	Goat	Retail
GmI6	OK148630	2522	Sphinx 1.76-like (BMMF1)	Goat	Retail

Abbreviations: SmGV—sheep milk genomovirus; GmGV—goat milk genomovirus; SmI—sheep milk isolate; GmI—goat milk isolate.

Circular ssDNA elements that were detected in individual milk samples originated from two out of five sheep flocks. In flock B, 40% (6/15) of the sheep milk samples contained circular full-length sequences with one animal bearing two different sequences. In contrast, flock C comprised 13% (2/15) positive individual samples only. Results obtained from sheep and goat milk on the flock level thus markedly differ from the recently published data on water buffalo milk: (i) both water buffalo herds under study were affected and (ii) 56% of the animals from herd 1 and 42% from herd 2 tested positive for full-length sequences [25]. Although three out of 40 individual goat milk samples yielded amplicons of the expected size and sequence with at least one primer pair during PCR screening, we failed to retrieve full-length circular genomes from any of them. This discrepancy has already been observed when water buffalo milk was assayed [25] and again underlines the importance not to rely on the presence of sub-genomic amplicons to categorize an individual as positive. In contrast to the individual samples from the goat farms, four full-length genomes were recovered from goat milk at retail (GmGV9; GmI2, GmI5 and GmI6).

Overall, 50% of sheep and goat milk at retail contained either full-length CRESS-DNA or BMMF-like DNA. Interestingly, goat milk at retail tested positive although none of the individual samples under study contained circular ssDNA elements. Unfortunately, milk at the grocery cannot be traced back to the delivering farms especially those brands produced in Austria and Spain. To interpret this outcome, one has to keep in mind that due to blending at the dairy plant a milk package available at retail contains bulk milk from numerous individuals kept on different farms. Although the number of herds and animals assayed herein is far too low for a comprehensive prevalence estimation, a cautious interpretation of results could read as follows: There are positive and negative herds of sheep and goats kept in Germany. The percentage of affected animals in a positive herd is low to medium. Processing at the dairy plant leads to a medium percentage of packages positive for ssDNA elements at retail.

Furthermore, one has to consider that individual milk samples from the farms have been assayed as raw milk, whereas milk at retail has been pasteurized. As both sample types enabled the recovery of full-length ssDNA heat treatment at the dairy plant does not seem to negatively affect the isolation and detection procedure. Plenty other studies demonstrated the occurrence of CRESS-DNA in various samples like e.g., blood, feces, insects and plant material. Thus, a low-level contamination of milk with those entities cannot be fully excluded. Furthermore, one has to keep in mind, that all full-length genomes were recovered after RCA of input DNA. Thus, we assume that the “viral load” is rather low. A classical qPCR approach to determine viral copy number is questionable, because qPCR is based on the amplification of sub-genomic amplicons. In the present study, we occasionally faced diverging results after PCR screening and full-length amplification due to unknown reasons. Thus, a quantification of the sub-genomic fragments might be misleading. Future studies would benefit from a coherent definition of “BMMF/genomovirus positive” individuals and samples.

With regard to the cancer hypothesis one has to keep in mind that sheep and goats have a considerable impact on human nutrition worldwide [33,34]. Remarkably, small ruminants play a significant role in some Asian and African countries [35], like for example India, West and Central Africa, where we observe low colorectal cancer incidences [36].

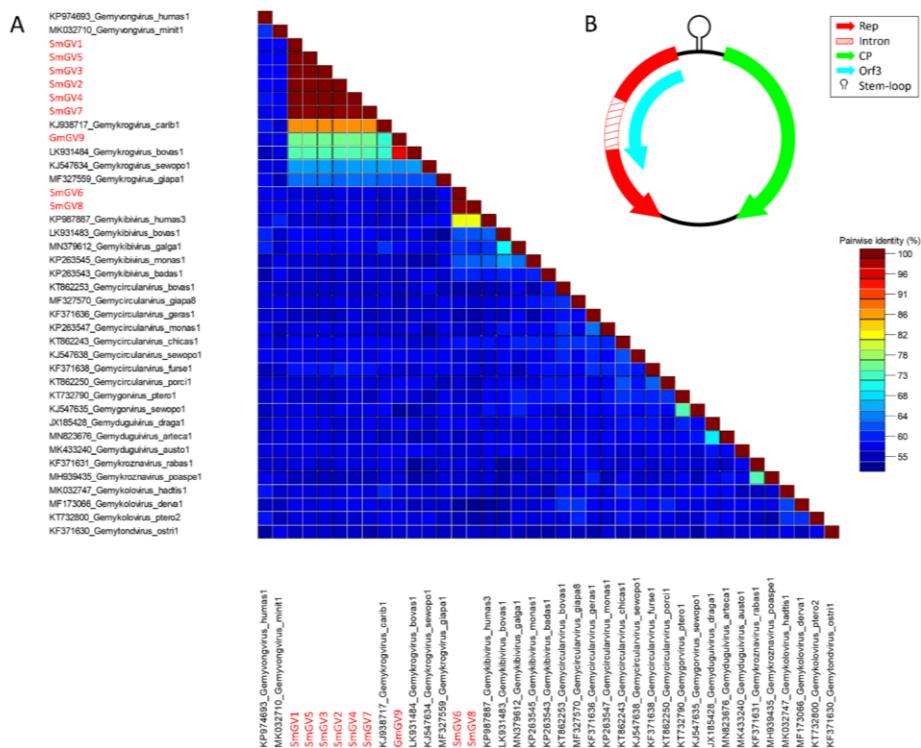
### 3.2. Characterization of Genomoviral Sequences

In the present study, nine genomoviral sequences, eight from individual sheep milk samples and one from goat milk at retail, were detected. The isolates were termed sheep milk genomovirus 1–8 (SmGV1–8) and goat milk genomovirus 9 (GmGV9). Their length ranged from 2122 to 2192 nt. BlastSearch hits assigned the sequences to the family *Genomoviridae* (Supplementary Table S2, sheet 1). A genome-wide pairwise similarity comparison of the isolates from this study with representatives from nine genomoviral genera listed by the International Committee for Taxonomy of Viruses (ICTV) was performed to classify SmGV1–GmGV9 (Figure 1A). By setting the species demarcation threshold to 78% according to the suggestions given by Varsani and Krupovic in 2017 [9] the isolates SmGV1–5 and SmGV7 represent variants of the ICTV assigned species “*Gemykrogvirus carib1*” (KJ938717) isolated from caribou feces with 88% genome pairwise identity. These six isolates show >99% similarity among each other thus representing one strain. Isolate GmGV9 also fits the genus *Gemykrogvirus*. The highest sequence similarity (95%) is given to “*Gemykrogvirus bovas1*” (LK931484) which has been detected in taurine cattle serum. With 82% pairwise similarity, isolates SmGV6 and SmGV8 are variants of the species “*Gemykibivirus humas3*” (KP987887) and therefore belong to the genus *Gemykibivirus*. In order to obtain a depiction of results better comparable to the later presented BMMF isolates, a phylogenetic tree was calculated based on representative sequences from nine genomoviral genera and genomoviral sequences from this study (Figure 2).

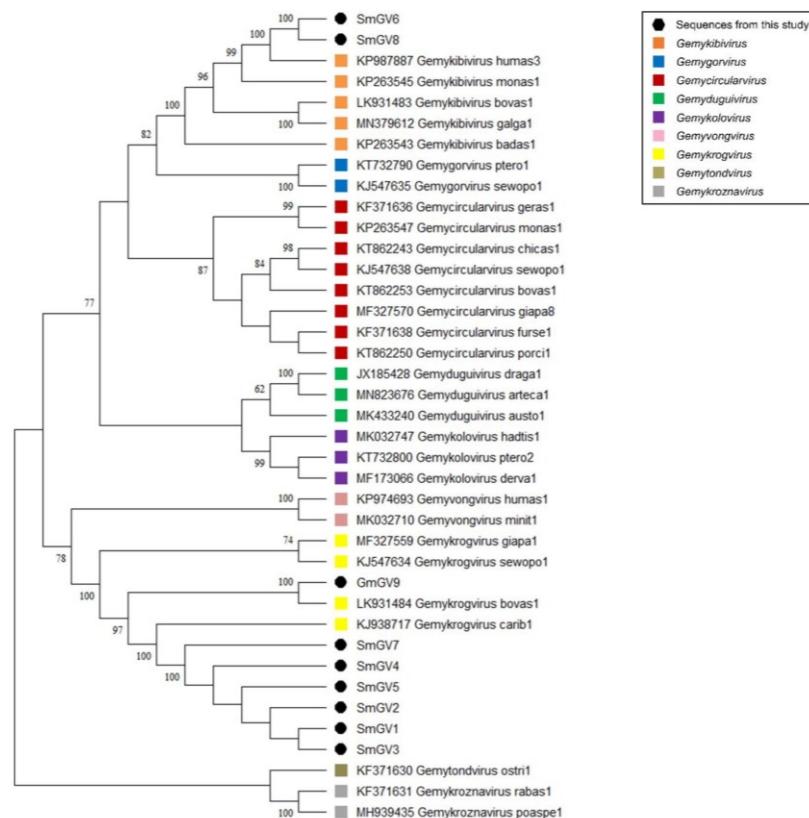
In summary, seven out of nine genomoviral isolates under study cluster within species of the genus *Gemykrogvirus*, while the remaining two genomes belong to a species of genus *Gemykibivirus*. According to the similarity calculations and the prerequisites given by Varsani and Krupovic in 2017 [9], none of our isolates represents a novel species.

All isolates identified in this study exhibit typical features of *Genomoviridae* family members upon in silico analysis (Figure 1B). Beside a potential stem-loop structure at the origin of replication with a conserved nonanucleotide motif (TAATATTAT) at the loop-tip [39–42] genomes encode a Rep protein in complementary sense direction separated by a putative intron. These intron sequences have firstly been described in plant-infecting geminiviruses [43]. Like diverse ssDNA viruses, bacterial plasmids and phages, CRESS DNA viruses replicate through rolling circle replication (RCR) [44,45]. This mechanism is initiated by the Rep protein containing a distinct RCR endonuclease domain and a superfamily 3 (SF3) helicase domain [46]. Genomoviral Reps contain characteristic motifs associated with the endonuclease and helicase function. These motifs slightly differ between genera [9]. All conserved motifs identified in the new isolates are summarized in Table 2 and the variation

in the amino acid motifs of all gemykibivirus- and gemykrogvirus-isolates including the ICTV-listed ones is illustrated in Figure 3. In addition to the genome-wide sequence comparison, amino acid motifs confirm the assignment of SmGV1–5, SmGV7 and SmGV9 to the genus *Gemykrogvirus* and SmGV6 and 8 to the genus *Gemykibivirus*. RCR motif I is thought to represent an essential part for recognition of the origin of replication and is located at the N-terminus of the Rep. RCR Motif II comprises two histidines ‘u-His-Tyr-His-u’ potentially coordinating metal ions that function as cofactors for endonuclease activity (u denotes hydrophobic residues). The third RCA motif seems to play an important role in dsDNA cleavage, covalent attachment of Rep and positioning during catalysis. GRS motifs separating motifs II and III have been identified for the first time in members of the family *Geminiviridae* by Nash et al. [47]. Their presence in isolates from this study indicates an involvement in the replication mechanism of genomoviruses. Furthermore all nine isolates exhibit typical SF3 helicase motifs (Walker A, B and C), which are necessary for dsDNA intermediate unwinding [48,49]. Interestingly, isolate SmGV2 contains a modified Walker C motif where a neutral amino acid (N) is replaced by an acidic one (D). Whether this substitution has a functional influence on the helicase activity has to be assayed. However, according to Varsani and Krupovic, 2017 [9] an acidic amino acid at the end of Walker C does not occur in any other genomoviral sequence.



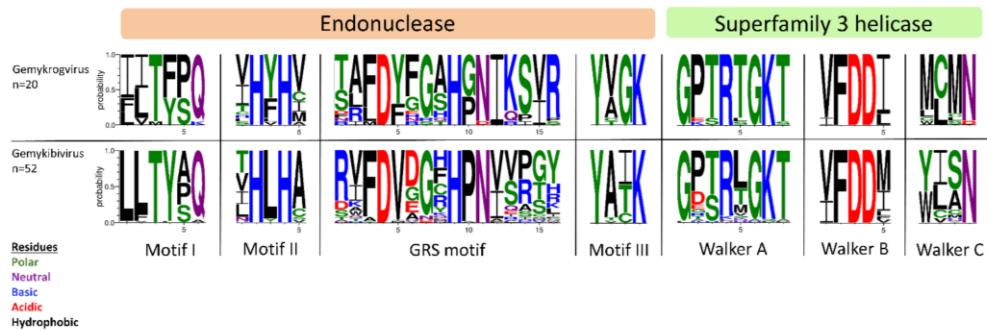
**Figure 1.** (A) Genome-wide pairwise nucleotide similarity score matrix including full-length genomoviruses from this study and representative members from nine genera of the family *Genomoviridae*. Sequences from this study are highlighted in red. (B) Schematic genome organization of the herein identified genomoviral sequences shows at least two major open reading frames and a stem-loop structure in the intergenic region.



**Figure 2.** Maximum Likelihood phylogenetic tree of representative sequences from nine genomoviral genera and genomoviral sequences from this study. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [37]. The bootstrapping consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed [38]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [38]. Branch support values lower than 60% were not included. This analysis involved 39 nucleotide sequences. There were a total of 3641 positions in the final dataset. Evolutionary analyses were conducted in MEGA software. All taxa are indicated by name and the corresponding GenBank accession number.

**Table 2.** Summary of conserved motifs in genomoviral sequences after intron clearing and in silico translation.

Isolate	Nonanucleotide	Motif I	Motif II	GRS Motif	Motif III	Walker A	Walker B	Walker C
SmGV1	TAATATTAT	IITFPQ	VHYHV	TAFDYFGAHGNIKSVR	YVGK	GPTRTGKT	VFDDI	MCMN
SmGV2	TAATATTAT	IITFPQ	VHYHV	TAFDYFGAHGNIKSVR	YVGK	GPTRTGKT	VFDDI	MCMN
SmGV3	TAATATTAT	IITFPQ	VHYHV	TAFDYFGAHGNIKSVR	YVGK	GPTRTGKT	VFDDI	MCMN
SmGV4	TAATATTAT	IITFPQ	VHYHV	TAFDYFGAHGNIKSVR	YVGK	GPTRTGKT	VFDDI	MCMN
SmGV5	TAATATTAT	IITFPQ	VHYHV	TAFDYFGAHGNIKSVR	YVGK	GPTRTGKT	VFDDI	MCMN
SmGV6	TAATATTAC	LFTYSQ	THLHA	RKFDFVGFHPNIISTI	YATK	GPSRTGKT	VFDDI	WLSN
SmGV7	TAATATTAT	IITFPQ	VHYHV	TAFDYFGAHGNIKSVR	YVGK	GPTRTGKT	VFDDI	MCMN
SmGV8	TAATATTAC	LFTYSQ	THLHA	RKFDFVGFHPNIISTI	YATK	GPSRTGKT	VFDDI	WLSN
GmGV9	TAATATAT	IIMFPQ	IHYHI	TAFDYFGAHGNIKSIR	YVGK	GPTRTGKT	VFDDI	MCMN



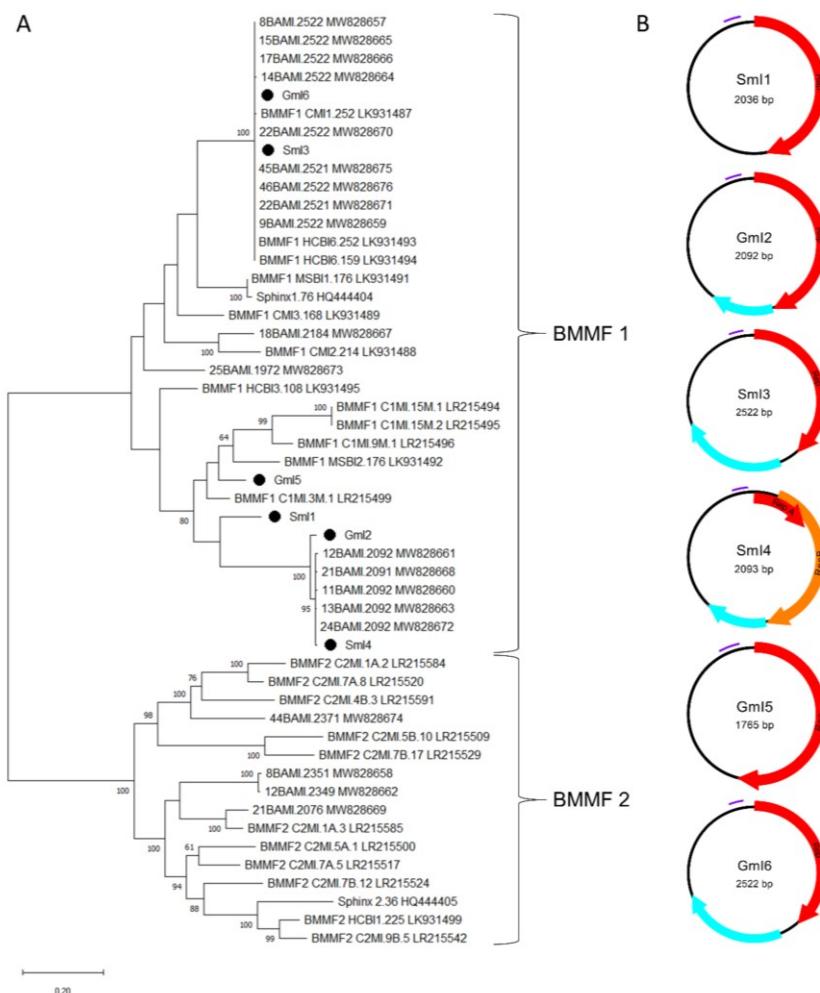
**Figure 3.** Illustration of amino acid variations in conserved Rep motifs using WebLogo3 [50]. Reps of ICTV-listed gemykrog- and gemykibiviruses as well as from the nine genomoviral isolates identified herein have been included in the alignment.

In addition to the species assignment based on a genome-wide comparison, a further approach is based on the analysis of the Rep amino acid sequences. Results might slightly differ from whole genome analysis because of intra-familiar recombination [51]. In terms of the sequences retrieved herein, Rep comparison did not differ from the genome-wide assessment and thus did not hint to recombination events (Supplementary Figure S1A). Genomoviruses further encode CPs, which in general exhibit higher sequence divergence than Reps. This divergence can be seen in a similarity score matrix based on amino acid sequences of CPs (Supplementary Figure S1B). While the percentage identities of CPs from SmGV1–5, SmGV7 and SmGV9 to those from other gemykrogviruses are still higher than 83%, CPs from SmGV6 and SmGV8 share 72% similarity to their closest relative “Gemykibivirus humas3” only. BlastSearch hits for Rep, CP and putative ORF3 translation products are summarized in Supplementary Table S2 (sheets 2–4).

### 3.3. Characterization of BMMF-Related Sequences

In addition to nine genomoviral sequences, we retrieved another six circular ssDNA elements. Of those, three were recovered from sheep milk (two from milk at retail, one from herd C) and three from goat milk at retail. According to the BlastSearch hits, all reveal similarities to already published BMMF elements (Supplementary Table S2, sheet 1). A phylogenetic tree based on present sequences, previously published sequences found in water buffalo milk, BMMF from database entries, Sphinx 1.76 and Sphinx 2.36 is depicted in Figure 4A.

Sml3 and Gml6 belong to a cluster containing water buffalo milk isolates and other well characterized BMMF, such as CMI1.252 and HCBI6.252 (Figure 4A). Gml2 and Sml4 built a cluster with five sequences recently detected in water buffalo milk. Sml1 is only distantly related to the latter cluster. Gml5 is the most remote of our sequences showing relatedness only to BMMF1 isolates MSBI2.176, C1MI.9M.1 and C1MI.15M.1. Results depicted in the phylogenetic tree are consistent with the BlastSearch hits summarized in Supplementary Table S2 (sheet 1). Open reading frame prediction revealed the common presence of a *rep*-gene (Figure 4B). Sml4 bears a first short and a second larger *rep* that starts on ‘ATG’ or ‘TTG’ respectively. BlastSearch of in silico translated Reps showed highest similarities to BMMF1 or *Acinetobacter* sp. replication proteins (Supplementary Table S2, sheet 3). Sml1 and Gml5 are the only sequences lacking an additional, putative ORF.



**Figure 4.** (A) Maximum Likelihood phylogenetic tree of novel full-length Sphinx-like DNA genomes, Sphinx 1.76, Sphinx 2.36, selected members of BMMF group 1 and 2 and recently published Sphinx-like DNA genomes isolated from water buffalo milk. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [37]. The tree with the highest log likelihood (“-50216.13”) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches [38]. Branch support values lower than 60% were not included. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 51 nucleotide sequences. There were a total of 4500 positions in the final dataset. Black dots mark sequences found in this study. All taxa are indicated by name followed by the corresponding GenBank accession number. (B) Genomic organization of the six new BMMF-like genomes from this study. Red and orange arrows indicate *rep*-genes, blue arrows show further potential ORFs and purple bars represent tandem repeat regions.

All sequences from the present study could be assigned to the BMMF1 group according to their features. This group is mainly distinguishable from BMMF2 by the presence of a tandem repeat region that is present in the sheep and goat milk isolates as shown in

**Table 3.** Tandem repeats that are typically located upstream the *rep*-gene possibly function as replicase recognition sites. The RCR mechanism of these DNA elements is further supported by inverted repeats, which are located 49–56 nt upstream the tandem repeat region. All BMMF1-related isolates exhibit such a palindromic sequence similar to a motif in BMMF1 elements published by zur Hausen et al. [17].

**Table 3.** BMMF-like DNA isolates and inherent tandem repeat regions.

Isolate	Period Size	Copy Number	Sequence	Nt between TR and RepA	Nt between TR and RepB
SmI1	22	3.5	CCTACGTTACCCATCAATACC	60	-
GmI2	22	3.5	ACACCGTTACCCATCAATATG	59	-
SmI3	22	3.8	ATACCCCTACGTTACCGATCA	60	-
SmI4	22	3.5	CACCGTTACCCATCAATATGA	27	151
GmI5	22	3.8	ATACTCCTAGGTTACCTACCA	59	-
GmI6	22	3.8	ATACCCCTACGTTACCGATCA	60	-

#### 4. Conclusions

Results recorded in the present study substantiate our hypothesis that the occurrence of Cressnaviricota and BMMF-related DNA elements in milk is not as narrowly restricted to “taurine” cattle as previously assumed. Sheep and goats kept for dairy purposes also shed those entities with milk. Thus, the present study contributes to the growing body of evidence on ssDNA elements frequently detectable in food of animal origin. From an epidemiological point of view, latest data published by Lechmann and colleagues in 2021 support the idea of interindividual and even interspecies transmission for gemykgroviruses [52]. Due to the marked discrepancies between the occurrence of viral DNA elements on herd level and in grocery packages an in depth prevalence assessment is needed. These future investigations should also include putative environmental CRESS sources like forage or water supply that might cause an affection of dairy ruminants and thus milk. Knowledge on affection levels of flocks and individuals could provide the first step for implementation of potential countermeasures. With regard to the circular ssDNA linked cancer hypothesis, further research efforts on underlying biological mechanisms and risk assessment are urgently needed.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/v13112176/s1>, Figure S1: SDT-matrices of Reps and CPs, Table S1: Primer-sets, Table S2: BlastSearch hits.

**Author Contributions:** Conceptualization—E.M. and G.S.; Methodology—E.L. and R.F.; Investigation and Visualization—M.-T.K.; Validation and Writing—M.-T.K. and A.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision by the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program (funding reference: 2819110217).

**Institutional Review Board Statement:** Ethical review and approval was waived for this study because milk was collected during the daily milking routine on the farm.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All annotated genomes from this study were deposited in the GenBank database (Accession numbers OK148616-OK148630).

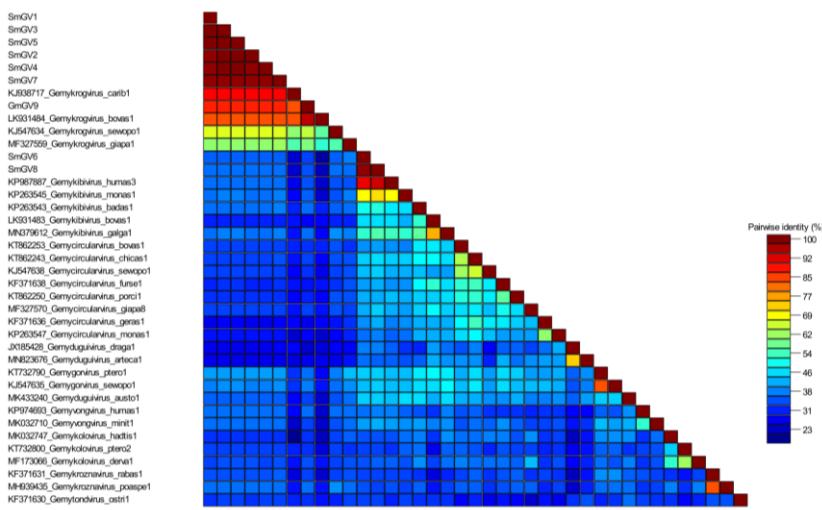
**Conflicts of Interest:** The authors declare no conflict of interest.

## References

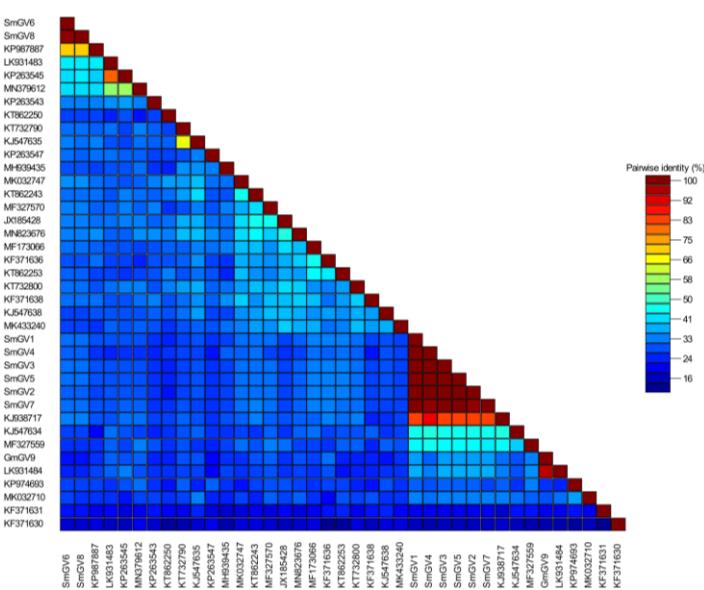
1. Shulman, L.; Davidson, I. Viruses with Circular Single-Stranded DNA Genomes Are Everywhere! *Annu. Rev. Virol.* **2017**, *4*, 159–180. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Tisza, M.J.; Pastrana, D.V.; Welch, N.L.; Stewart, B.; Peretti, A.; Starrett, G.J.; Pang, Y.-Y.S.; Krishnamurthy, S.R.; Pesavento, P.A.; McDermott, D.H.; et al. Discovery of several thousand highly diverse circular DNA viruses. *eLife* **2020**, *9*, 9. [\[CrossRef\]](#)
3. Dayaram, A.; Galatowitsch, M.L.; Argüello-Astorga, G.R.; van Bysterveldt, K.; Kraberger, S.; Stainton, D.; Harding, J.S.; Roumagnac, P.; Martin, D.P.; Lefevre, P.; et al. Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. *Infect. Genet. Evol.* **2016**, *39*, 304–316. [\[CrossRef\]](#)
4. Manuelidis, L. Nuclease resistant circular DNAs copurify with infectivity in scrapie and CJD. *J. NeuroVirology* **2010**, *17*, 131–145. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Funk, M.; Gunst, K.; Lucansky, V.; Müller, H.; Hausen, H.Z.; de Villiers, E.-M. Isolation of Protein-Associated Circular DNA from Healthy Cattle Serum. *Genome Announc.* **2014**, *2*, e00846-14. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Gunst, K.; Hausen, H.Z.; de Villiers, E.-M. Isolation of Bacterial Plasmid-Related Replication-Associated Circular DNA from a Serum Sample of a Multiple Sclerosis Patient. *Genome Announc.* **2014**, *2*, e00847-14. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Krupovic, M.; Varsani, A.; Kazlauskas, D.; Breitbart, M.; Delwart, E.; Rosario, K.; Yutin, N.; Wolf, Y.I.; Harrach, B.; Zerbini, F.M.; et al. Cressdnaviricota: A Virus Phylum Unifying Seven Families of Rep-Encoding Viruses with Single-Stranded, Circular DNA Genomes. *J. Virol.* **2020**, *94*. [\[CrossRef\]](#)
8. Zhao, L.; Rosario, K.; Breitbart, M.; Duffy, S. Eukaryotic Circular Rep-Encoding Single-Stranded DNA (CRESS DNA) Viruses: Ubiquitous Viruses With Small Genomes and a Diverse Host Range. *Adv. Clin. Chem.* **2019**, *103*, 71–133. [\[CrossRef\]](#)
9. Varsani, A.; Krupovic, M. Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family Genomoviridae. *Virus Evol.* **2017**, *3*, vew037. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Rosario, K.; Duffy, S.; Breitbart, M. A field guide to eukaryotic circular single-stranded DNA viruses: Insights gained from metagenomics. *Arch. Virol.* **2012**, *157*, 1851–1871. [\[CrossRef\]](#)
11. Ling, Y.; Zhang, X.; Qi, G.; Yang, S.; Jingjiao, L.; Shen, Q.; Wang, X.; Cui, L.; Hua, X.; Deng, X.; et al. Viral metagenomics reveals significant viruses in the genital tract of apparently healthy dairy cows. *Arch. Virol.* **2019**, *164*, 1059–1067. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Wang, H.; Li, S.; Mahmood, A.; Yang, S.; Wang, X.; Shen, Q.; Shan, T.; Deng, X.; Li, J.; Hua, X.; et al. Plasma virome of cattle from forest region revealed diverse small circular ssDNA viral genomes. *Virol. J.* **2018**, *15*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Zhang, W.; Li, L.; Deng, X.; Kapusinszky, B.; Delwart, E. What is for dinner? Viral metagenomics of US store bought beef, pork, and chicken. *Virol.* **2014**, *468–470*, 303–310. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Cibulski, S.; de Lima, D.A.; dos Santos, H.F.; Teixeira, T.F.; Tochetto, C.; Mayer, F.Q.; Roehe, P.M. A plate of viruses: Viral metagenomics of supermarket chicken, pork and beef from Brazil. *Virol.* **2021**, *552*, 1–9. [\[CrossRef\]](#) [\[PubMed\]](#)
15. De Villiers, E.-M.; Gunst, K.; Chakraborty, D.; Ernst, C.; Bund, T.; Hausen, H.Z. A specific class of infectious agents isolated from bovine serum and dairy products and peritumoral colon cancer tissue. *Emerg. Microbes Infect.* **2019**, *8*, 1205–1218. [\[CrossRef\]](#)
16. Whitley, C.; Gunst, K.; Müller, H.; Funk, M.; Hausen, H.Z.; de Villiers, E.-M. Novel Replication-Competent Circular DNA Molecules from Healthy Cattle Serum and Milk and Multiple Sclerosis-Affected Human Brain Tissue. *Genome Announc.* **2014**, *2*, 2. [\[CrossRef\]](#)
17. Falida, K.; Eilebrecht, S.; Gunst, K.; Hausen, H.Z.; de Villiers, E.-M. Isolation of Two Virus-Like Circular DNAs from Commercially Available Milk Samples. *Genome Announc.* **2017**, *5*, e00266-17. [\[CrossRef\]](#)
18. Krupovic, M. Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. *Curr. Opin. Virol.* **2013**, *3*, 578–586. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Koonin, E.V.; Dolja, V.V.; Krupovic, M. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virol.* **2015**, *479–480*, 2–25. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Bund, T.; Nikitina, E.; Chakraborty, D.; Ernst, C.; Gunst, K.; Boneva, B.; Tessmer, C.; Volk, N.; Brobeil, A.; Weber, A.; et al. Analysis of chronic inflammatory lesions of the colon for BMMF Rep antigen expression and CD68 macrophage interactions. *Proc. Natl. Acad. Sci. USA* **2021**, *118*. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Hausen, H.Z.; Bund, T.; De Villiers, E. Specific nutritional infections early in life as risk factors for human colon and breast cancers several decades later. *Int. J. Cancer* **2018**, *144*, 1574–1583. [\[CrossRef\]](#)
22. Manuelidis, L. Prokaryotic SPHINX 1.8 REP protein is tissue-specific and expressed in human germline cells. *J. Cell. Biochem.* **2018**, *120*, 6198–6208. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Lamberto, I.; Gunst, K.; Müller, H.; Hausen, H.Z.; de Villiers, E.-M. Mycovirus-Like DNA Virus Sequences from Cattle Serum and Human Brain and Serum Samples from Multiple Sclerosis Patients. *Genome Announc.* **2014**, *2*, 2. [\[CrossRef\]](#)
24. Pulina, G.; Milán, M.J.; Lavin, M.P.; Theodoridis, A.; Morin, E.; Capote, J.; Thomas, D.; Francesconi, A.; Caja, G. Invited review: Current production trends, farm structures, and economics of the dairy sheep and goat sectors. *J. Dairy Sci.* **2018**, *101*, 6715–6729. [\[CrossRef\]](#) [\[PubMed\]](#)
25. König, M.-T.; Fux, R.; Link, E.; Sutter, G.; Märtlbauer, E.; Didier, A. Circular Rep-Encoding Single-Stranded DNA Sequences in Milk from Water Buffaloes (*Bubalus arnee f. bubalis*). *Viruses* **2021**, *13*, 1088. [\[CrossRef\]](#)
26. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [\[CrossRef\]](#) [\[PubMed\]](#)

27. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. *PLoS ONE* **2014**, *9*, e108277. [\[CrossRef\]](#)
29. The European Molecular Biology Open Software Suite (EMBOSS) A. Available online: <http://emboss.bioinformatics.nl/cgi-bin/emboss/equicktandem> (accessed on 20 September 2021).
30. The European Molecular Biology Open Software Suite (EMBOSS) B. Available online: <https://www.bioinformatics.nl/cgi-bin/emboss/palindrome> (accessed on 22 September 2021).
31. National Center for Biotechnology Information (NCBI). ORFfinder. Available online: <https://www.ncbi.nlm.nih.gov/orffinder/> (accessed on 16 September 2021).
32. Burland, T.G. DNASTAR's Lasergene Sequence Analysis Software. *Bioinform. Methods Protoc.* **2000**, *132*, 71–91. [\[CrossRef\]](#)
33. Zygogiannis, D. Sheep production in the world and in Greece. *Small Rumin. Res.* **2006**, *62*, 143–147. [\[CrossRef\]](#)
34. Skapetas, B.; Bampidis, V. Goat production in the World: Present situation and trends. *Livest Res Rural Dev* **2016**, *28*, 200.
35. Food and Agriculture Organization of the United Nations: Gateway to dairy production and products. Available online: <http://www.fao.org/dairy-production-products/production/dairy-animals/small-ruminants/en/> (accessed on 20 September 2021).
36. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J. Clin.* **2021**, *71*, 209–249. [\[CrossRef\]](#)
37. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526. [\[CrossRef\]](#)
38. Felsenstein, J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution* **1985**, *39*, 783–791. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Heyraud-Nitschke, F.; Schumacher, S.; Laufs, J.; Schaefer, S.; Schell, J.; Gronenborn, B. Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic Acids Res.* **1995**, *23*, 910–916. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Rosario, K.; Dayaram, A.; Marinov, M.; Ware, J.; Kraberger, S.; Stainton, D.; Breitbart, M.; Varsani, A. Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J. Gen. Virol.* **2012**, *93*, 2668–2681. [\[CrossRef\]](#)
41. Laufs, J.; Traut, W.; Heyraud, F.; Matzeit, V.; Rogers, S.G.; Schell, J.; Gronenborn, B. In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3879–3883. [\[CrossRef\]](#)
42. Timchenko, T.; de Kouchkovsky, F.; Katul, L.; David, C.; Vetten, H.J.; Gronenborn, B. A Single Rep Protein Initiates Replication of Multiple Genome Components of Faba Bean Necrotic Yellows Virus, a Single-Stranded DNA Virus of Plants. *J. Virol.* **1999**, *73*, 10173–10182. [\[CrossRef\]](#)
43. Gutierrez, C. Geminivirus DNA replication. *Cell. Mol. Life Sci.* **1999**, *56*, 313–329. [\[CrossRef\]](#)
44. Ilyina, T.V.; Koonin, E.V. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Res.* **1992**, *20*, 3279–3285. [\[CrossRef\]](#)
45. Gilbert, W.; Dressler, D. DNA Replication: The Rolling Circle Model. *Cold Spring Harb. Symp. Quant. Biol.* **1968**, *33*, 473–484. [\[CrossRef\]](#)
46. Gorbatenko, A.; Koonin, E.V.; Wolf, Y.I. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett.* **1990**, *262*, 145–148. [\[CrossRef\]](#)
47. Nash, T.E.; Dallas, M.B.; Reyes, M.I.; Buhrman, G.K.; Ascencio-Ibáñez, J.T.; Hanley-Bowdoin, L. Functional Analysis of a Novel Motif Conserved across Geminivirus Rep Proteins. *J. Virol.* **2011**, *85*, 1182–1192. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Choudhury, N.R.; Malik, P.S.; Singh, D.K.; Islam, M.N.; Kaliappan, K.; Mukherjee, S.K. The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. *Nucleic Acids Res.* **2006**, *34*, 6362–6377. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Clerot, D.; Bernardi, F. DNA Helicase Activity Is Associated with the Replication Initiator Protein Rep of Tomato Yellow Leaf Curl Geminivirus. *J. Virol.* **2006**, *80*, 11322–11330. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Crooks, G.E.; Hon, G.; Chandonia, J.-M.; Brenner, S.E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188–1190. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Kraberger, S.; Argüello-Astorga, G.R.; Greenfield, L.G.; Galilee, C.; Law, D.; Martin, D.P.; Varsani, A. Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. *Infect. Genet. Evol.* **2015**, *31*, 73–86. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Lechmann, J.; Ackermann, M.; Kaiser, V.; Bachofen, C. Viral infections shared between water buffaloes and small ruminants in Switzerland. *J. Veter. Diagn. Investig.* **2021**, *33*, 894–905. [\[CrossRef\]](#) [\[PubMed\]](#)

Supplementary Material



A



B

**Supplementary Figure S1.** Genome-wide pairwise amino acid similarity score matrices including Reps (A) and CPs (B) of all genomoviruses from this study and representative members from nine genera of the family *Genomoviridae* using SDT v1.2.

**Table S1.** Sheet 1: Information on the primer-sets applied in the PCR-screening that resulted in the amplification of Sphinx-like sequences. Wobble bases are indicated according the universal code: R = A/G, Y = C/T, M = A/C, K = G/T, S = C/G, W = A/T, B = C/G/T, D = A/G/T, H = A/C/T, V = A/C/G, and N = A/C/G/T. The annealing temperature was 55 °C and the PCR comprised 35 cycles. bp = base pairs.

Amplicon name	Forward (F), Reverse (R)	Sequence 5'-3'	Amplicon length in bp	Sphinx-type	
S1C	S1F	AGATAACGCCCTAATGAACGCT	233	Sphinx 1.76	
	S1.5R	CAACCGCAGGGGCAAAATAA			
	S1.6R	CAACCGCAGGGGCAAATATTA			
S5A	S5.1F	GGSAAAGGGATTAATGCMAATGATC	334		
	S5.2F	GGSAAGGAATTAATGCAAATGAC			
	S5.1R	TAAGCGAATKGCTAAGCRC			
	S5.2R	TAGACGGACGGCATAGGCAC			
S6A	S6F	TTGGCGWASTACAGGYAAA	310	Sphinx 2.36	
	S6R	TWGCGAATAATGMCGTTGAG			
S7A	S7F	ARTGGCATCCGCATATYCAT	152		
	S7.1R	CAGCTTGCTATAGCCGTGT			
	S7.2R	CAGCCTTGCTATCCGTGT			
	S7.3R	CCGCTTGCTATAACCGTAA			

Sheet 2: Information on the primer-sets applied in the PCR-screening, that resulted in the amplification of genomoviral amplicons. Wobble bases are indicated according the universal code: R = A/G, Y = C/T, M = A/C, K = G/T, S = C/G, W = A/T, B = C/G/T, D = A/G/T, H = A/C/T, V = A/C/G, and N = A/C/G/T. The annealing temperature was 55 °C and the PCR comprised 35 cycles. bp = base pairs.

Amplicon name	Forward (F), Reverse (R)	Sequence 5'-3'	Amplicon length in bp	Genus
G13A	G13F	GATGGYGGAACTCATCTRCACG	541	Gemykibivirus
	G13R	TAGGGTTAGGGTTAGGGTTAGG		
G14A	G14F	GATGGTGGGACGCATCTACATG	541	
	G13R	TAGGGTTAGGGTTAGGGTTAGG		
G18A	G18F	ACTGCGTTCGACTACTTGGAG	352	Gemykrogvirus
	G18R	CACTCGGAGATCTCAGGATAGC		

Sheet 3: Primers for the amplification of full length genomes. Primers from this study are not abutting and were designed via the sequences gained from PCR screening. The annealing temperature was 55°C for primers from this study, 49,5°C for primers from Whitley *et al.* and the PCR comprised 40 cycles. \* = this study. <sup>1</sup> = Whitley, C. *et al.* 2014.

Isolate	Accession Number	Primer Forward*	Primer Reverse*
SmGV1	OK148616	GCTCCAAGAGATTGGTTATCC	CATAATCGAAGACCTTGTGTG
SmGV2	OK148617	GCTCCAAGAGATTGGTTATCC	GATTCTTAGTGTGCGAGCTGATG
SmGV3	OK148618	CTGTCAGCGGTACGTGAGC	GTCAAGTCGGATGTCTCCATC
SmGV4	OK148619	GCTCCAAGAGATTGGTTATCC	GATTCTTAGTGTGCGAGCTGATG
SmGV5	OK148620	CTGTCAGCGGTACGTGAGC	GTCAAGTCGGATGTCTCCATC
SmGV6	OK148621	CCAATCTGTTGATCTTACAC	GACAATAAGGTGATCGATCCACTC
SmGV7	OK148622	GCTCCAAGAGATTGGTTATCC	GATTCTTAGTGTGCGAGCTGATG
SmGV8	OK148623	CCAATCTGTTGATCTTACAC	GACAATAAGGTGATCGATCCACTC
GmGV9	OK148624	GCTCCAAGAGATTGGTTATCC	GATTCTTAGTGTGCGAGCTGATG
Sml1	OK148625	GAATACCTGGTTGTCACCG	CGTGTCTAGGTCTTGTGACG
Gml2	OK148626	CAAGTCGATGGGTATCACAAATTG	GATATGCCGTTGTTCTGTGACG
Sml4	OK148628	CAAGTCGATGGGTATCACAAATTG	GATATGCCGTTGTTCTGTGACG
Gml5	OK148629	CGGTTATCGAGTCCTATGC	GATAGCCAAGAGAATTAACGCTG

Isolate	Accession Number	Primer Forward <sup>1</sup>	Primer Reverse <sup>1</sup>
Sml3	OK148627	GAGGACGAATTAATATTACAAGTC	GTTCTCGCTTTCTTGGTAA
Gml6	OK148630	GAGGACGAATTAATATTACAAGTC	GTTCTCGCTTTCTTGGTAA

**Table S2.** Sheet 1: Summary of the best two BLASTn results for each isolate. The last column indicates the number of missing (-) or additional (+) nucleotides in the novel sequences compared to the GenBank hits. Green background colour highlights the affiliation to the family *Genomoviridae*, the yellow background indicates the affiliation to BMMF. nt = nucleotides.

Isolate	Accession Number	Length in nt	Hit in BlastSearch	GenBank Accession No	Length in nt	% Query Coverage	% ID	Δ nt to blast Hit
SmGV1	OK148616	2191	Gila monster-associated gemycircularvirus isolate gila_283, complete genome	MH378453	2191	99	97,6	0
			Genomoviridae sp. isolate ctbd374, complete genome	MK032706	2150	92	96,0	41
SmGV2	OK148617	2192	Gila monster-associated gemycircularvirus isolate gila_283, complete genome	MH378453	2191	99	97,4	1
			Genomoviridae sp. isolate ctbd374, complete genome	MK032706	2150	92	96,0	42
SmGV3	OK148618	2190	Gila monster-associated gemycircularvirus isolate gila_283, complete genome	MH378453	2191	99	97,4	-1
			Genomoviridae sp. isolate ctbd374, complete genome	MK032706	2150	92	96,02	40
SmGV4	OK148619	2190	Gila monster-associated gemycircularvirus isolate gila_283, complete genome	MH378453	2191	99	97,1	-1
			Genomoviridae sp. isolate ctbd374, complete genome	MK032706	2150	92	95,9	40
SmGV5	OK148620	2191	Gila monster-associated gemycircularvirus isolate gila_283, complete genome	MH378453	2191	99	97,3	0
			Genomoviridae sp. isolate ctbd374, complete genome	MK032706	2150	92	95,9	41
SmGV6	OK148621	2124	Mongoose feces-associated gemycircularvirus c strain 541c, complete genome	KP263546	2123	99	99,2	1
			Chicken genomovirus mg2_77, complete genome	MN379602	2127	99	92,6	-3
SmGV7	OK148622	2190	Gila monster-associated gemycircularvirus isolate gila_283, complete genome	MH378453	2191	99	96,9	-1
			Genomoviridae sp. isolate ctbd374, complete genome	MK032706	2150	92	95,8	40
SmGV8	OK148623	2125	Mongoose feces-associated gemycircularvirus c strain 541c, complete genome	KP263546	2123	99	99,3	2
			Chicken genomovirus mg2_77, complete genome	MN379602	2127	99	92,7	-2
GmGV9	OK148624	2122	HCBI9.212 virus complete sequence	LK931484	2121	100	94,4	1
			Murine feces-associated gemycircularvirus 2 isolate MFAGCV2/NYC/Manhattan/poolF3, partial genome	MF416388	2418	89	79,0	-296
SmI1	OK148625	2036	BMMF1 DNA sequence, isolate C1MI.15M.1	LR215494	2041	99	84,1	-5
			BMMF1 DNA sequence, isolate C1MI.15M.2	LR215495	2040	99	84,0	-4
GmI2	OK148626	2092	BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499	1765	71	88,8	327
			Sphinx1.76-related DNA, replication competent episomal DNA MSB1.176	LK931491	1766	55	83,4	326
SmI3	OK148627	2522	Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493	2522	100	99,6	0
			Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487	2523	100	99,4	-1
SmI4	OK148628	2093	BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499	1765	71	88,4	328
			Uncultured prokaryote from Rat gut metagenome metamobilome, plasmid pGRRH0296	LN852968	2203	46	83,4	-110
GmI5	OK148629	1765	BMMF1 DNA sequence, isolate C1MI.3M.1	LR215499	1765	100	91,1	0
			Mediopippa subpectinata	OC882703	1897	99	86,7	-132
GmI6	OK148630	2522	Sphinx1.76-related DNA, replication competent episomal DNA CMI1.252	LK931487	2523	100	99,5	-1
			Sphinx1.76-related DNA, replication competent episomal DNA HCBI6.252	LK931493	2522	100	99,8	0

Sheet 2: Summary of the best two BLASTp results for CPs of each genomoviral isolate. The last column indicates the number of missing (-) or additional (+) amino acids in the novel sequences compared to the GenBank hits. aa = amino acids.

Isolate	Orientation	Start codon	Length in aa	Best Hit(s) in BlastSearch	GenBank Accession No of Hit	Length in aa	% Query Coverage	% ID	Δ aa to blast Hit
SmGV1	sense	ATG	302	capsid protein [Genomoviridae sp.]	QCW23629	302	100	98,7	0
				capsid protein [Gila monster-associated gemycircularvirus]	AWX63613	302	100	98,3	0
SmGV2	sense	ATG	302	capsid protein [Genomoviridae sp.]	QCW23629	302	100	98,3	0
				capsid protein [Gila monster-associated gemycircularvirus]	AWX63613	302	100	98,0	0
SmGV3	sense	ATG	302	capsid protein [Genomoviridae sp.]	QCW23629	302	100	98,7	0
				capsid protein [Gila monster-associated gemycircularvirus]	AWX63613	302	100	98,3	0
SmGV4	sense	ATG	257	capsid protein [Genomoviridae sp.]	QCW23629	302	100	98,8	-45
				capsid protein [Gila monster-associated gemycircularvirus]	AWX63613	302	100	98,4	-45
SmGV5	sense	ATG	302	capsid protein [Genomoviridae sp.]	QCW23629	302	100	98,7	0
				capsid protein [Gila monster-associated gemycircularvirus]	AWX63613	302	100	98,3	0
SmGV6	sense	ATG	305	capsid [Mongoose feces-associated gemycircularvirus c]	YP_009130627	305	100	99,3	0
				capsid protein [Chicken genomovirus mg2_77]	QIR82228	307	100	85,3	-2
SmGV7	sense	ATG	302	capsid protein [Genomoviridae sp.]	QCW23629	302	100	98,0	0
				capsid protein [Gila monster-associated gemycircularvirus]	AWX63613	302	100	97,7	0
SmGV8	sense	ATG	305	capsid [Mongoose feces-associated gemycircularvirus c]	YP_009130627	305	100	99,3	0
				capsid protein [Chicken genomovirus mg2_77]	QIR82228	307	100	85,3	-2
GmGV9	sense	ATG	294	capsid protein [HCBI9.212 virus]	YP_0090518331	294	84	93,1	0
				Cap [Murine feces-associated gemycircularvirus 2]	AWB14659	295	84	90,3	-1

Sheet 3: Summary of the best two BLASTp results for Reps of each isolate. Green background colour highlights the affiliation to the family *Genomoviridae*, the yellow background indicates the affiliation to BMMF. The last column indicates the number of missing (-) or additional (+) amino acids in the novel sequences compared to the GenBank hits. aa = amino acids.

Isolate	Orientation	Start codon	Length in aa	Best Hit(s) in BlastSearch	GenBank Accession No of Hit	Length in aa	% Query Coverage	% ID	Δ aa to blast Hit
SmGV1	antisense	ATG	324	replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	100	95,7	-5
				replication-associated protein [Genomoviridae sp.]	QCW23628	329	100	93,3	-5
SmGV2	antisense	ATG	324	replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	100	95,7	-329
				replication-associated protein [Genomoviridae sp.]	QCW23628	329	100	93,3	-5
SmGV3	antisense	ATG	324	replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	100	96,1	-329
				replication-associated protein [Genomoviridae sp.]	QCW23628	329	100	93,6	-5
SmGV4	antisense	ATG	324	replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	100	95,7	0
				replication-associated protein [Genomoviridae sp.]	QCW23628	329	100	93,3	-5
SmGV5	antisense	ATG	324	replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	100	96,1	-5
				replication-associated protein [Genomoviridae sp.]	QCW23628	329	100	93,6	-5
SmGV6	antisense	ATG	318	replication associated protein [Chicken genomovirus mg2_77]	QIR82229.1	326	100	96,2	-8
				rep [Mongoose feces-associated gemycircularvirus c]	YP_009130628.1	306	100	80,1	12
SmGV7	antisense	ATG	321	replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	100	95,7	-8
				replication-associated protein [Genomoviridae sp.]	QCW23628	329	100	93,3	-8
SmGV8	antisense	ATG	300	replication associated protein [Chicken genomovirus mg2_77]	QIR82229.1	326	100	91,5	-26
				rep [Mongoose feces-associated gemycircularvirus c]	YP_009130628.1	306	100	82,3	-6
GmGV9	antisense	ATG	337	replication-associated protein [Genomoviridae sp.]	QCW23628	329	96	87,0	8
				replication-associated protein [Gila monster-associated gemycircularvirus]	AWX63615	329	96	86,7	8
Sml1	sense	ATG	319	replication protein [BMMF1 DNA sequence]	VEV85356	327	100	85,3	-8
				hypothetical protein [uncultured prokaryote]	CRY95586	323	100	85,3	-4
Gml2	sense	ATG	316	replication initiation protein RepM [Acinetobacter sp. NIPH 284]	WP_005155574	324	72	92,1	-8
				replication initiation protein RepM [Acinetobacter sp. YH12231]	WP_180074503	324	69	90,4	-8
Sml3	sense	ATG	326	replication protein [BMMF1 DNA sequence]	QWQ66138	326	100	99,4	0
				replication protein [BMMF1 DNA sequence]	QWQ66157	326	100	99,1	0
Sml4 Rep A	sense	ATG	115	RepB family plasmid replication initiator protein [Acinetobacter junii]	WP_151834693	136	68	96,2	-21
				unnamed protein product [Brugia timori]	VDO49810	148	78	95,6	-33
Sml4 Rep B	sense	TTG	286	MULTISPECIES: replication initiation protein RepM [Acinetobacter]	WP_180172971	320	81	71,2	-34
				replication initiation protein RepM [Acinetobacter sp. YH12058]	WP_180080578	312	81	71,2	-26
Gml5	sense	ATG	318	MULTISPECIES: replication initiation protein RepM [Acinetobacter]	WP_121774444	318	100	93,1	0
				MULTISPECIES: replication initiation protein RepM [Acinetobacter]	WP_087739066	318	100	92,5	0
Gml6	sense	ATG	326	replication protein [Sphinx1.76-related DNA]	CDS63400	326	100	100,0	0
				replication protein [Sphinx1.76-related DNA]	CDS63392	316	93	98,7	10

Sheet 4: Summary of the best two BLASTp results for further hypothetical ORFs. Green background colour highlights the affiliation to the family *Genomoviridae*, the yellow background indicates the affiliation to BMMF. The last column indicates the number of missing (-) or additional (+) amino acids in the novel sequences compared to the GenBank hits. aa = amino acids.

Isolate	Orientation	Start codon	Length in aa	Best Hit(s) in BlastSearch	GenBank Accession No of Hit	Length in aa	% Query Coverage	% ID	Δ aa to blast Hit
SmGV1	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	88	96,8	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	89	57,7	-146
SmGV2	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	88	96,8	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	89	57,7	-146
SmGV3	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	88	96,2	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	89	57,7	-146
SmGV4	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	88	96,2	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	89	57,1	-146
SmGV5	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	88	96,8	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	89	57,7	-146
SmGV7	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	88	96,2	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	89	57,1	-146
GmGV9	antisense	ATG	211	hypothetical protein [bubaline-associated gemykgrovirus]	QOQ72540	230	94	80,4	-19
				Rep [Murine feces-associated gemycircularvirus 2]	AWB14658	357	98	52,1	-146
Gml2	sense	TTG	106	hypothetical protein [Acinetobacter sp. ANC 3813]	WP_086212311	102	95	68,3	4
				hypothetical protein [Acinetobacter guerae]	WP_151979748	98	94	63,0	8
Sml3	sense	ATG	219	hypothetical protein [Acinetobacter johnsonii]	WP_213073224	219	100	98,6	0
				hypothetical protein [BMMF1 DNA sequence]	QWQ66141	219	100	98,2	0
Sml4	sense	TTG	106	hypothetical protein [Acinetobacter sp. ANC 3813]	WP_086212311	102	95	69,3	4
				hypothetical protein [Acinetobacter guerae]	WP_151979748	98	94	63,0	8
Gml6	sense	ATG	219	uncharacterized protein [Sphinx1.76-related DNA]	CDS63393	219	100	99,1	0
				MULTISPECIES: hypothetical protein [unclassified Psychrobacter]	WP_025643765	221	100	69,0	-2



## V. BISHER UNVERÖFFENTLICHE DATEN – PROBEN WEITERER RINDERSPEZIES

In den beiden vorgestellten Publikationen wurde das Vorkommen von BMMF in Wasserbüffel-, Schaf- und Ziegenmilch untersucht. Doch noch weitere Spezies spielen in Ländern mit niedrigen Krebsinzidenzen bei der Versorgung der Bevölkerung mit Milch und Fleisch eine wichtige Rolle. Um die Frage zu klären, ob die im westeuropäischen Raum als exotisch geltenden Rinderspezies nicht tauriner Herkunft ebenfalls potenzielle BMMF-Träger darstellen könnten, wurden insgesamt 60 Proben von Zebus (n=23), Yaks (n=5), Wasserbüffeln (n=4) und Watussi-Rindern (n=28) aus neun deutschen Zoos und Tierparks auf BMMF-DNA untersucht. Da diese Tiere nicht gemolken werden, musste auf anderes Material zurückgegriffen werden. In Anlehnung an die bereits im Literaturteil erwähnten Arbeiten von Harald zur Hausen und Kollegen wurden Blut- und Serumproben verwendet. Frisches EDTA-Blut wurde im Rahmen der jährlichen Routine-Blutentnahmen für die BHV1-Diagnostik entnommen. Bei den Seren handelte es sich um tiefgekühlte Rückstellproben, die zum Teil bis zu 10 Jahre alt waren. Die neun Zoos bzw. Tierparks wurden anonymisiert mit den Buchstaben A-I benannt und die einzelnen Proben mit einer individuellen Kennzeichnung versehen (Tabelle 2).

Die 60 Proben von verschiedenen Zoos und Tierparks wurden gekühlt (EDTA-Blut n=24) oder gefroren (Serum n=36) innerhalb von 24 Stunden zum Lehrstuhl für Hygiene und Technologie der Milch geschickt. Gekühltes EDTA-Blut wurde in vier verschiedene Fraktionen (Vollblut, Leukozyten, Blutplasma, Blutzellen) separiert. Die DNA-Extraktion und anschließende RCA der Proben erfolgte nach demselben methodischen Vorgehen, wie zuvor in den Publikationen beschrieben. Alle RCA-Produkte wurden mittels PCR auf das Vorkommen von BMMF-Elementen untersucht. Im Falle positiver PCR-Reaktionen, die Banden auf der zu erwartenden Höhe lieferten, wurden diese aus dem Gel ausgeschnitten, die DNA aufgereinigt und nach erfolgreicher Klonierung sequenziert (methodische Details und Informationen zu den Sequenzanalysen siehe Publikationen I und II „Methods“).

Bei der Untersuchung der Blut- und Serumproben von Zebus, Watussi-Rindern, Wasserbüffeln und Yaks konnten 3 der insgesamt 60 Proben in einem ersten PCR-Screening positiv getestet werden. Der Einsatz eines Primer-Paars für den

Nachweis von BMMF 1-DNA resultierte bei zwei Yak-EDTA-Blutproben vom Standort H (75.Y, 76.Y) in der Amplifikation von BMMF-ähnlichen Fragmenten. Durch den Einsatz von Ringschluss-Primern gelang allerdings nur bei einem der beiden Tiere der Nachweis eines vollständigen zirkulären Genoms. Zusätzlich wurde aus einer einzelnen Wasserbüffel-EDTA-Blutprobe (122.WB) aus der Herde G eine vollständige zirkuläre Sequenz isoliert. Somit konnten insgesamt in 2 von 60 Einzeltierproben „exotischer“ Rinderspezies zirkuläre Gesamtgenome detektiert werden, die bei näherer Charakterisierung starke Ähnlichkeiten zu bekannten BMMF-Isolaten aufweisen.

**Tabelle 2:** Auflistung der neun Standorte und des jeweils eingesendeten Probenmaterials.

Standort	Spezies	Anzahl	Probenmaterial	Positive Tiere	Proben-Kennung positive Tiere
A	Watussi	2	Serum	0	–
B	Yak	1	EDTA-Blut	0	–
	Zebu	2	EDTA-Blut	0	–
C	Watussi	2	EDTA-Blut	0	–
D	Watussi	11	Serum	0	–
E	Zebu	2	EDTA-Blut	0	–
F	Zebu	13	Serum	0	–
	Watussi	9	Serum	0	–
G	Yak	1	EDTA-Blut	0	–
	Wasserbüffel	4	EDTA-Blut	1	122.WB
H	Yak	3	2x EDTA-Blut, 1x Serum	2	75.Y, 76.Y
I	Watussi	4	EDTA-Blut	0	–
	Zebu	6	EDTA-Blut	0	–
Gesamt		60		3	

## VI. DISKUSSION

### 1. BMMF in Probenmaterial unterschiedlicher Spezies

Den Ausgangspunkt für die vorliegende Dissertationsschrift bildet die von Harald zur Hausen formulierte Hypothese, dass die verhältnismäßig hohen Dickdarm- und Brustkrebsinzidenzen in westlichen Industrienationen im Zusammenhang mit der Aufnahme von Milch und Fleisch von Rindern tauriner Abstammung stehen (zur Hausen and de Villiers, 2015). Diese Hypothese basiert vor allem darauf, dass in Ländern, die im Gegensatz zur westlichen Hemisphäre vorwiegend „exotische“ Rinderarten wie Yak, Zebu oder Wasserbüffel nutzen, die Inzidenzen für die zuvor genannten Krebsarten auffallend niedrig sind. Aus Proben taurinen Ursprungs wiederum konnten die sogenannten BMMF isoliert werden, die seither in Verdacht stehen, an einer indirekten Karzinogenese und weiteren humanen Erkrankungen beteiligt zu sein (de Villiers et al., 2019, de Villiers and zur Hausen, 2021).

Die von zur Hausen und Kollegen aufgestellte Hypothese ist sowohl spannend als auch gewagt. Einerseits gilt es zu bedenken, dass Gesundheitsdaten in den Entwicklungs- und Schwellenländern in niedrigerem Umfang erhoben werden, als dies in den westlichen Industrienationen der Fall ist, sodass ein Vergleich der beiden Datensätze nur bedingt möglich ist. Andererseits wurden mit der Isolierung der BMMF tatsächlich neuartige Agenzien identifiziert, die möglicherweise nicht nur mit einer indirekten Karzinogenese, sondern auch mit weiteren häufig vorkommenden Zivilisationskrankheiten der westlichen Welt in Zusammenhang stehen könnten. Grundsätzlich fehlten allerdings bisher aussagekräftige Vergleichsdaten zum BMMF-Vorkommen bei „exotischen“ Rinderarten, die ein wesentlicher Baustein für die Untermauerung der oben genannten Hypothese sind.

In der vorliegenden Arbeit wurden erstmals grundlegende Datensätze zum Vorkommen von BMMF in „exotischen“ Rindern und weiteren lebensmittel-liefernden Tieren wie Schafe und Ziegen generiert.

#### 1.1. Milchproben von Wasserbüffeln, Schafen und Ziegen

Bei der Untersuchung von Einzelgemelksproben von in Deutschland lebenden Wasserbüffeln konnten zirkuläre DNA-Moleküle isoliert werden, die den bereits beschriebenen BMMF-Gruppen von Harald zur Hausen zugeordnet werden

können. In 16 von insgesamt 30 untersuchten Milchproben wurden 20 BMMF 1- als auch BMMF 2-ähnliche Isolate und zusätzlich ein Gemycircularvirus nachgewiesen. Durch *in silico*-Analysen konnten bei den Isolaten ein Replikase-kodierender ORF und zum Teil weitere hoch konservierte Strukturen identifiziert werden, die charakteristisch für BMMF 1-Moleküle sind (s. Publikation I). Die beprobten Tiere waren klinisch unauffällig und stammten aus zwei verschiedenen Herden mit Weidezugang an unterschiedlichen Standorten. Beide Herden umfassten nach dem PCR-Screening sowohl BMMF-positive als auch -negative Individuen.

Die Produktion von Wasserbüffelmilch spielt in vielen asiatischen Ländern neben der von Kuhmilch eine bedeutende Rolle. In Indien und Pakistan – Länder, in denen vergleichsweise niedrige Inzidenzen von Brust- und Darmkrebs feststellbar sind – werden die größten Mengen Wasserbüffelmilch produziert und konsumiert (Ferlay et al., 2020, FAO). Da alle in Europa gehaltenen Wasserbüffel von asiatischen Wasserbüffeln abstammen, die auf dem indischen Subkontinent domestiziert und gezüchtet wurden (Zhang et al., 2020), lässt sich mit Hilfe der hier erhobenen Daten zeigen, dass die Spezies *Bubalus arnee* f. *bubalis* als Träger von BMMF fungiert und diese somit potenziell auf den Menschen übertragen könnte. Obwohl die Untersuchung von nur zwei Herden nicht ausreichend ist, um eine umfassende Prävalenzsituation abzubilden, konnte zweifelsfrei gezeigt werden, dass im Gegensatz zu der von zur Hause aufgestellten Theorie BMMF auch in der Milch von Wasserbüffeln in einem nicht unerheblichen Umfang vorkommen.

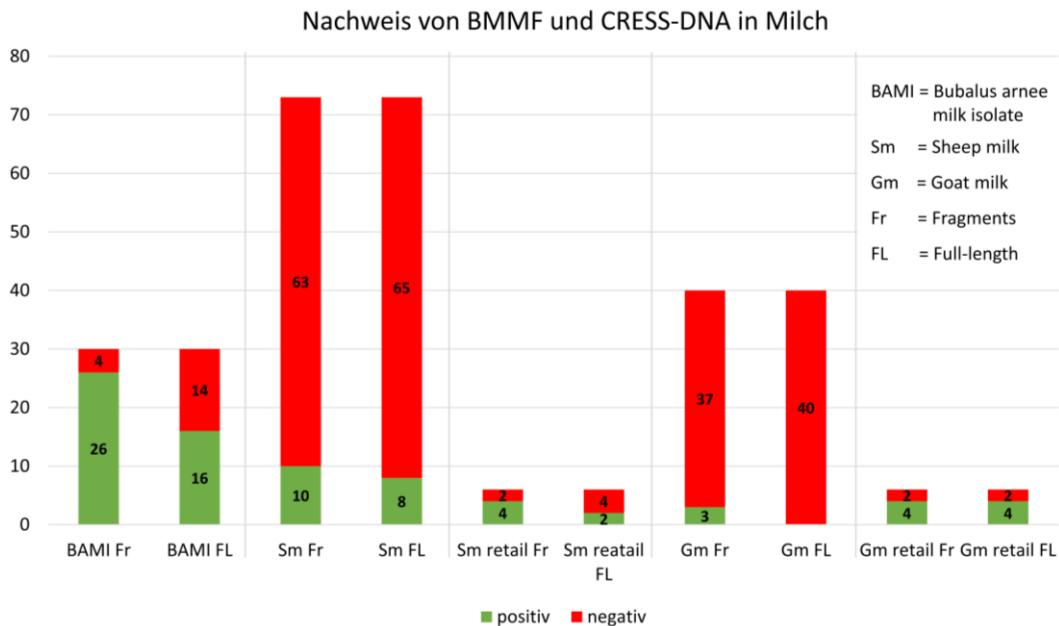
Einschränkend muss allerdings angemerkt werden, dass für diese Arbeit keine Proben von Wasserbüffeln aus den ursprünglichen Verbreitungs- oder Zuchtgebieten wie Indien und Pakistan zur Verfügung standen. Daher bleibt die Frage offen, ob die episomalen BMMF-Elemente über die ursprünglichen Importtiere nach Deutschland gebracht oder ob sie am Standort Deutschland erworben wurden. Bisher gibt es keine ausreichenden Informationen darüber, wo BMMF ursprünglich zu finden und welche Formen der Übertragung möglich sind. Die Übertragungswege und welche Spezies als ursprüngliche Wirte angesehen werden könnten, sind auch bei einigen CRESS-DNA-Viren bis *dato* nicht sicher geklärt. Der kürzliche Fund eines neuen Gemykrogvirus, sowohl bei Wasserbüffeln als auch bei Schafen eines landwirtschaftlichen Betriebes in der Schweiz, lässt vermuten, dass diese Erreger zwischen verschiedenen Spezies übertragen werden

können (Lechmann et al., 2021). Interessanterweise waren einige BMMF-Isolate aus den zwei hier untersuchten Wasserbüffelherden nahezu identisch, obwohl die Tiere mehrere hundert Kilometer getrennt voneinander gehalten wurden. Diese Genome zeigten zudem große Homologien (nahezu 100 %) zu einigen aus Kuhmilch isolierten BMMF-Molekülen. Eine mögliche Erklärung dafür wäre eine BMMF-Übertragung von infizierten europäischen Rindern anderer Höfe auf Wasserbüffel durch Vektoren. Der Eintrag über solche Vektoren wie blutsaugende Insekten oder Zecken, die für die Übertragung vieler Infektionserreger verantwortlich und nachweislich Träger von diversen CRESS-DNA-Viren sind (Rosario et al., 2012a, Waits et al., 2018, Krabberger et al., 2019), scheint durchaus möglich, wurde bisher jedoch noch nicht untersucht. Da fast alle in Deutschland gehaltenen Wasserbüffel von wenigen nach Europa importierten Tieren abstammen, wäre eine vertikale Übertragung der BMMF-Elemente von den Elterntieren auf deren Nachkommen ebenfalls denkbar. Weitere Studien zum Vorkommen von BMMF bei Rindern unter Einbeziehung der Verwandtschaftsbeziehungen und der Genetik wären daher wünschenswert.

Als weitere, in der Hypothese von zur Hausen und Kollegen nicht berücksichtigte Spezies, wurden im Rahmen dieser Arbeit zusätzlich Schafe und Ziegen untersucht. Diese werden weltweit in bedeutender Zahl gehalten und für die Milch- und Fleischgewinnung genutzt (Dubeuf et al., 2004, Zygogiannis, 2006), wobei Pakistan zu den Top 10 Produzenten von Ziegenmilch zählt (FAO). Bei der Untersuchung von Schafmilch konnten in zwei von fünf Herden Genomovirus-positive Tiere detektiert werden. Das Screening zeigte deutlich, dass nie alle Schafe einer Herde positiv waren, sondern nur einzelne Individuen. Der prozentuale Anteil positiver Tiere war bei den Schafen jedoch erheblich geringer (11 % aller untersuchten Tiere) als bei den Wasserbüffeln (53 % aller untersuchten Tiere). Aus der Ziegenmilch gelang bei keiner der drei untersuchten Herden ein Nachweis vollständiger zirkulärer Genome.

Neben den individuellen Einzelgemelken wurde Schaf- und Ziegenmilch aus dem Einzelhandel getestet, wobei hier in 50 % aller untersuchten Proben entweder BMMF- oder CRESS-DNA nachgewiesen wurde. Im Unterschied zu der Analyse auf Einzeltierebene wurden bei der Ziegenmilch aus dem Supermarkt interessanterweise aus 4 von 6 Proben zirkuläre BMMF-Genome isoliert (Abbildung 12). Ähnliches konnte bei der Beprobung eines Milchschafbetriebes festgestellt werden:

20 untersuchte Tiere dieses Betriebes wurden negativ getestet. Im Gegensatz dazu befanden sich in einer aus dem Einzelhandel erworbenen Sammelmilch desselben Betriebes vollständige zirkuläre BMMF-Elemente. Diese, sich grundsätzlich widersprechenden Befunde, werfen natürlich einige Fragen auf. Prinzipiell wäre es möglich, dass BMMF mit der Milch nicht kontinuierlich, sondern intermittierend ausgeschieden werden. Positive, aber intermittierend ausscheidende Tiere, die nur zu einem einzigen Zeitpunkt beprobt wurden, könnten somit unentdeckt bleiben. Folglich besteht die Möglichkeit, dass die 20 Schafe zu einem „falschen“ Zeitpunkt getestet wurden, an dem keine BMMF-Ausscheidung nachweisbar war. Interessant wären in diesem Zusammenhang weitere Studien, in denen die Tiere zu unterschiedlichen Laktationszeitpunkten mehrfach beprobt werden. So stünden aussagekräftige Daten zur BMMF-Infektion und -Ausscheidung zur Verfügung.



**Abbildung 12:** Vergleichende Darstellung der Diskrepanz zwischen dem PCR-Screening zum Nachweis von DNA-Fragmenten (Fr) und dem finalen Nachweis von vollständigen zirkulären Genomen (FL). Auf der X-Achse sind die verschiedenen Spezies aufgetragen, die Y-Achse zeigt die Anzahl der untersuchten Tiere bzw. der getesteten Milchsorten aus dem Einzelhandel. In Rot und Grün sind jeweils die negativ und positiv getesteten Proben dargestellt.

Da in der vorliegenden Arbeit allerdings nur ein Teil der gesamten Schafherde beprobt wurde, könnten ebenso die nicht untersuchten Schafe den Eintrag in die Milch verursacht haben. Es wäre auch denkbar, dass eine Herde wenige „Superausscheider“ beherbergt, die große Mengen BMMF-DNA sezernieren und trotz des Verdünnungseffektes der Herdensammelmilch zu einer nachweislichen

Kontamination der Tankmilch führen. Eine mögliche „Verunreinigung“ der Sammelmilch mit BMMF zu einem späteren Zeitpunkt (z. B. innerhalb der Molkerei) wäre ebenso eine plausible Erklärung für das oben beschriebene Szenario.

Obwohl die Isolierung von vollständigen BMMF-Genomen aus Schaf- und Ziegenmilch in deutlich geringerem Ausmaß gelang als aus Wasserbüffelmilch, konnte trotzdem erstmals gezeigt werden, dass BMMF auch in diesen Tierarten verbreitet zu sein scheinen. Die Empfänglichkeit von Schafen und Ziegen für BMMF verdeutlicht, dass die DNA-Elemente scheinbar in sehr viel mehr Spezies vorkommen als bislang angenommen. Allerdings standen auch in dieser Untersuchungsreihe keine Proben von Tieren aus Ländern mit niedrigen Brust- oder Darmkrebsinzidenzen zur Verfügung und es bleibt daher unklar, ob sich das BMMF-Vorkommen innerhalb der Tierpopulation in den Herkunftsländern ähnlich verhält wie in deutschen Herden.

Der Nachweis von BMMF aus technologisch verarbeiteten Milcherzeugnissen, den bereits zur Hausen und Kollegen für Kuhmilch erbracht haben (de Villiers et al., 2019), konnte mit den hier durchgeführten Untersuchungen der Schaf- und Ziegenmilch aus dem Einzelhandel bestätigt werden. Jedoch gibt es bisher keine Informationen darüber, ob der Grad der Wärmebehandlung oder eine Veränderung im Fettgehalt eine Auswirkung auf den BMMF-Nachweis hat. Daher wurden verschiedene Milchtypen wie ultrahocherhitzte Milch (UHT), Extended-Shelf-Life-Milch (ESL) und klassisch kurzzeiterhitzte Milch der Spezies Rind (unveröffentlichte Daten), Schaf und Ziege aus dem Einzelhandel untersucht. Dabei schien weder die Art des Herstellungsprozesses noch der Fettgehalt einen Einfluss auf die Detektierbarkeit von BMMF zu haben, da bei allen Herstellungsvarianten vollständige zirkuläre BMMF-Genome nachgewiesen werden konnten. Auch andere Publikationen deuten darauf hin, dass weitere Maßnahmen wie Trocknung, Erhitzung oder Fermentation keinen Einfluss auf die Detektion von BMMF zu haben scheinen (de Villiers et al., 2019, Pohl et al., 2022).

Der in dieser Arbeit erbrachte Nachweis von BMMF in allen drei boviden Spezies, die in Deutschland für die Milchgewinnung von Bedeutung sind, zeigt eindeutig, dass zumindest in Mitteleuropa diese Faktoren nicht wie ursprünglich propagiert nur bei taurinen Rindern vorkommen.

## 1.2. Blutproben weiterer exotischer Rinderspezies

Zur Abrundung der Untersuchungsreihe wurden zudem Blut- und Serumproben von in deutschen Zoos und Tierparks lebenden Zebus, Watussi-Rindern, Yaks und Wasserbüffeln untersucht (s. Kapitel V). Bei der Analyse dieser Proben gelang die Isolierung vollständiger, zirkulärer BMMF-ähnlicher Isolate aus einem Yak und einem Wasserbüffel. Bis auf einen weiteren Yak konnten bei den restlichen Tieren – insbesondere den Zebus und Watussi-Rindern – weder Full-length-Genome noch BMMF-DNA-Fragmente nachgewiesen werden. Aufgrund der geringen Anzahl untersuchter Herden sind valide Rückschlüsse auf die tatsächliche BMMF-Verbreitung allerdings nur bedingt möglich.

Es besteht die Möglichkeit, dass die untersuchten Zebu- und Watussi-Herden nicht infiziert waren und daher eine Detektion von BMMF oder CRESS-DNA unmöglich war. Die Ergebnisse könnten andererseits auch darauf hindeuten, dass Zebus und Watussi-Rinder *per se* tatsächlich nicht empfänglich für eine Infektion mit BMMF oder CRESS-DNA-Viren sind, da sie sich genetisch von den empfänglichen Rinderspezies unterscheiden. Diese Tatsache würde demnach die „BMMF-Hypothese“ von zur Hausen stützen. Wie in VI. 2. noch detaillierter diskutiert, könnten aber auch methodische Gründe ursächlich sein. So reduziert der Einsatz spezifischer Primer die Bandbreite detektierbarer DNA-Elemente und könnte daher zu falsch negativen Ergebnissen geführt haben. Schlussendlich ist nicht auszuschließen, dass der Zeitpunkt der Blutentnahme in einer Latenz-ähnlichen Periode der BMMF-Infektion stattgefunden haben könnte und somit ein „Erreger“-Nachweis nur schwer möglich gewesen wäre.

Grundsätzlich bleibt aber festzuhalten, dass sowohl im Tierartvergleich und innerhalb der Herden als auch bei der Untersuchung der beiden Probenmatrices Blut und Milch ein starker Unterschied in der Häufigkeit des BMMF-Nachweises zu beobachten ist. Es stellt sich daher die Frage, ob es zwischen den Spezies oder sogar den Individuen möglicherweise genetische Unterschiede gibt, die eine Aufnahme oder Weitergabe der BMMF begünstigen oder verhindern. Dies könnte in Studien, die das Genom positiver und negativer Tiere miteinander vergleichen, abgeklärt werden, um so eine plausible Erklärung dafür zu finden, weshalb nicht alle Tiere der untersuchten Milchvieh-Herden infiziert waren oder warum der BMMF-Nachweis aus Zebus und Watussi-Rindern bisher gar nicht gelang. Einige Studien belegen, dass die Empfänglichkeit für Vektor-übertragene Infektionen innerhalb

verschiedener bovider Spezies und Rassen sehr stark variiert und genetisch bedingt ist (Murray et al., 1984, Kemp et al., 1996). Eine potenzielle genetische BMMF-Resistenz könnte somit auch in Bezug auf die bereits diskutierte Übertragung durch Vektoren von Bedeutung sein.

## 2. Kritische Bewertung der eingesetzten Methoden

### 2.1. RCA

Die RCA bietet bei der Suche nach kleinen zirkulären ssDNA-Molekülen große Vorteile und erleichtert deren Detektion auch bei niedriger Abundanz. Durch den Einsatz dieser Technik ist die Entdeckung einer Vielzahl bisher unbekannter Genome in beeindruckender Geschwindigkeit vorangeschritten (Dean et al., 2001, Rosario et al., 2012b). Neben diesen Errungenschaften stellt die RCA Wissenschaftler jedoch auch vor neue Herausforderungen und Fragen (Johne et al., 2009).

Aufgrund der vermuteten geringen Abundanz der Ziel-DNA wurden im Rahmen dieser Studie alle DNA-Proben vor dem PCR-Screening der RCA unterzogen, um gezielt zirkuläre DNA zu vervielfältigen. Sowohl beim Screening der Wasserbüffelproben als auch bei dem der Schafe und Ziegen fiel auf, dass der Nachweis vollständiger, zirkulärer Genome bei deutlich weniger Tieren gelang als die Detektion kurzer BMMF- oder CRESS-DNA-ähnlicher Fragmente (Abbildung 12). Diese Ergebnisse deuten somit darauf hin, dass es bei der RCA nicht nur zur Amplifikation zirkulärer DNA kommt, sondern zusätzlich kurze subgenomische Nebenprodukte entstehen können. Bei der Detektion dieser Nebenprodukte in der anschließenden Screening-PCR könnte fälschlicherweise der Eindruck erweckt werden, es handle sich um positive Proben mit den gesuchten vollständigen, zirkulären DNA-Elementen. Unterzieht man die bei dem Screening gefundenen Fragmente einer Datenbanksuche (NCBI), so werden immer wieder hoch homologe Treffer für verschiedene *Acinetobacter baumannii* und andere bakterielle Plasmide angezeigt. Diese Sequenzhomologien wurden bereits beschrieben (zur Hausen et al., 2017). Allerdings sind viele der genannten Plasmide deutlich länger als die gesuchten BMMF-Elemente, sodass eine eindeutige Zuordnung auf Basis kurzer Fragmente aus der Screening-PCR nicht sicher möglich ist. Demnach ist es von entscheidender Bedeutung, wann ein Tier oder Lebensmittel als „positiv“ kategorisiert wird. Reicht der Nachweis kurzer genomicscher Fragmente von wenigen 100 Basenpaaren aus oder muss ein vollständiges zirkuläres Genom

isoliert werden? Aufgrund der obigen Überlegungen zur Ähnlichkeit der BMMF mit bakteriellen Plasmiden erscheint es geboten, ein Tier erst dann als BMMF-positiv zu kategorisieren, wenn ein vollständiges zirkuläres Genom isoliert werden konnte.

Eine weitere Herausforderung bei der Anwendung der RCA stellt die Reproduzierbarkeit der Ergebnisse dar. Nach wiederholter RCA einiger Milchproben kam es bei den Untersuchungen teilweise zu unterschiedlichen Ergebnissen. War beispielsweise nach erstmaliger RCA der Nachweis eines bestimmten vollständigen BMMF-Genoms aus einer Milchprobe möglich, so kam es in einigen wenigen Fällen vor, dass nach Wiederholung der RCA derselben Probe bei gleichem Vorgehen dieser Nachweis nicht mehr möglich war. Dieses Phänomen könnte darin begründet liegen, dass die zirkulären ssDNA-Elemente durch das Auftauen und Wiedereinfrieren der Probe instabil werden. Käme es durch die Lagerung und Handhabung der Proben zu Strangbrüchen innerhalb der gesuchten DNA-Genome, könnte dies ebenfalls eine mögliche Begründung für den Nachweis von DNA-Fragmenten in der PCR sein, ohne dass ein vollständiger Ringschluss generiert werden konnte. Zur Abklärung einer möglichen Instabilität der zirkulären DNA könnten multimodale chromatographische Verfahren dienen, bei denen potenziell entstandene DNA-Fragmente ihrer Länge nach aufgetrennt werden. Da die Ziel-DNA im Probenmaterial allerdings in einem Hintergrund aus genomischer DNA vorliegt, könnten diese Verfahren nur mit Hilfe zirkulärer Kontroll-DNA vergleichbarer Größe durchgeführt werden (Matos et al., 2014).

## 2.2. PCR *versus* NGS

In der vorliegenden Arbeit erfolgte der DNA-Nachweis ausschließlich mittels PCR. Aufgrund der Abhängigkeit von Primern, die auf bereits publizierten Genom-Sequenzen basieren, können so nur Genome detektiert werden, die mit den bekannten Isolaten Ähnlichkeiten aufweisen. Der Einsatz von Primern, die sogenannte Wobble-Basen enthalten, kann diesen Effekt abmildern, jedoch nicht vollständig ausschließen. Um ein potenziell breiteres BMMF- und CRESS-DNA-Spektrum zu detektieren, wäre eine Analyse der Proben mittels NGS von Vorteil. NGS-basierte Methoden werden bereits in vielen Studien erfolgreich eingesetzt, um CRESS-DNA-Viren in verschiedensten Matrices wie beispielsweise tierischem Serum, Kot und Abwasser nachzuweisen (Krabberger et al., 2015, Steel et al., 2016, Tochetto et al., 2020). Die Proben können entweder direkt in der

Sequenzierreaktion eingesetzt oder durch einen Amplifikationsschritt (z. B. RCA) vorbereitet werden. Grundsätzlich bietet die Anwendung der NGS-Technologie die Möglichkeit, Genome unabhängig von bekannten Sequenzen zu erfassen und zu assemblieren. Die Geräteausstattung ist jedoch sehr kostspielig und spezialisierten Laboren mit hochqualifizierten und erfahrenen Mitarbeitern vorbehalten. Ein Nachteil dieser Technologie ist darin zu sehen, dass Sequenzen, die im Probenmaterial mit niedriger Abundanz vorkommen, unzureichend oder nicht vollumfänglich erfasst werden und somit bei der Assemblierung von BMMF- und CRESS-DNA-Gesamtgenomen verloren gehen (Barzon et al., 2011, Cheval et al., 2011, Datta et al., 2015).

### **3. Abschließende Betrachtung und Ausblick**

Spätestens seit der Isolierung und Beschreibung der Sphinx 1.76- und Sphinx 2.36-DNA aus prionenhaltigem Material (Manuelidis, 2011) wird zum einen die „Protein-only-Hypothese“ von Stanley Prusiner hinterfragt (Prusiner, 1982), zum anderen wurde eine potenzielle neue Erregergruppe beschrieben, zu der die BMMF-Sequenzen deutliche Ähnlichkeiten aufweisen.

Bei der Fülle der verschiedenen BMMF-Isolate und CRESS-DNA-Viren, die in den letzten Jahren in diversen Probenmaterialien nachgewiesen werden konnte, scheint die ursprüngliche Hypothese einer obligaten Pathogenität nicht mehr haltbar. Neben den hier vorgestellten Untersuchungsergebnissen, die sich mit dem Nachweis von BMMF und CRESS-DNA befassen, zeigt auch eine weitere aktuelle Studie, dass BMMF scheinbar deutlich weiter verbreitet sind als ursprünglich angenommen. Bei der Untersuchung von diversem tierischen und pflanzlichen Material konnten BMMF-ähnliche Fragmente nachgewiesen werden, die große Homologien zu Molekülen aufweisen, die bereits aus Kuh- und Wasserbüffelmilch isoliert wurden (Pohl et al., 2022). Somit scheint die Aufnahme von BMMF durch unterschiedlichste Nahrungsmittel möglich. Es ist allerdings nicht auszuschließen, dass bestimmte BMMF-Varianten, die eventuell in einigen Tierspezies oder Regionen gehäuft vorkommen, ein Risiko für die menschliche Gesundheit darstellen. Hinweise auf eine mögliche Pathogenität der BMMF wurden erst kürzlich bei immunhistochemischen Untersuchungen von humanem Kolongewebe gesunder und an Dickdarmkrebs erkrankter Probanden erhalten. Hierbei wurde mittels monoklonaler Antikörper das BMMF 1-Rep-Antigen in deutlich größerem

Ausmaß in der *Lamina propria* von erkrankten, als von gesunden Personen detektiert. Zudem konnten massive Entzündungsreaktionen in den betroffenen Darmarealen nachgewiesen werden (Bund et al., 2021).

Im Rahmen dieser Arbeit konnte die ursprüngliche Hypothese zum ausschließlichen Vorkommen von BMMF in taurinen Rindern zweifelsfrei widerlegt werden. Die hier erhobenen und vorgestellten Daten belegen, dass neben taurinen Rindern auch Wasserbüffel, Schafe, Ziegen und Yaks für eine Infektion mit BMMF empfänglich sind. Im Hinblick auf den hohen methodischen Aufwand zum Nachweis vollständiger zirkulärer DNA-Sequenzen wäre es mittelfristig ein wünschenswertes Ziel, alternative Detektionssysteme zu etablieren, die methodisch einfacher und schneller durchzuführen sind als DNA-basierte Verfahren. Eine mögliche Strategie wäre ein antikörperbasierter Nachweis BMMF-assozierter Replikasen im Western-Blot oder ELISA. Diese weniger aufwendigen Techniken ermöglichen die Untersuchung großer Probenzahlen und liefern gleichzeitig Informationen über die Replikation der BMMF in den betroffenen Tieren. Voraussetzung für den Rep-Nachweis auf Proteinebene ist allerdings die Generierung Rep-spezifischer Antikörper. Ein weiterer potenzieller Ansatz wäre die Etablierung einer Serodiagnostik basierend auf der Prämisse, dass die BMMF-assoziierte Replikase möglicherweise eine Antikörperantwort im betroffenen Tier auslöst. Voraussetzung für dieses Vorgehen wäre die rekombinante Expression der BMMF-Replikase. Derartige Methoden würden es zudem erlauben, ein umfassenderes Verständnis für die Biologie dieser neuartigen DNA-Elemente zu gewinnen. Auf die Notwendigkeit weiterer Forschungsarbeiten zu diesem Thema hatte auch bereits das BfR in einer Stellungnahme von 2019 hingewiesen (BfR and MRI, 2019). Erste Schritte zur Etablierung der obengenannten Strategien wurden bereits in die Wege geleitet.

## VII. ZUSAMMENFASSUNG

In den vergangenen zehn Jahren gelang dank neuer Techniken die Entdeckung einer Vielzahl neuer, potenziell pathogener CRESS-DNA-Viren und anderer bisher unbekannter ssDNA-Elemente. Im Hinblick auf die Lebensmittelsicherheit erlangten vor allem neuartige zirkuläre ssDNA-Elemente, sogenannte bovine meat and milk factors (BMMF), die aus Kuhmilch, -seren und Kuhmilchprodukten isoliert werden konnten, besondere Aufmerksamkeit. Diese vermeintlich nur in europäischen Rindern taurinen Ursprungs vorkommenden Moleküle gerieten auf Basis globaler epidemiologischer Studien in Verdacht, indirekt an der Entstehung humaner Erkrankungen wie Kolon- und Mammakarzinomen und Multipler Sklerose beteiligt zu sein.

Um Klarheit über das Vorkommen dieser ssDNA-Elemente in anderen Spezies als *Bos taurus* zu erlangen, die weltweit ebenfalls der Gewinnung von Lebensmitteln dienen, wurden in der vorliegenden Arbeit individuelle Milchproben der Spezies Wasserbüffel (*Bubalus arnee f. bubalis*), Schaf (*Ovis aries*) und Ziege (*Capra aegagrus f. hircus*) und zusätzlich Schaf- und Ziegenmilch aus dem Einzelhandel untersucht. Bei der Untersuchung der Milch mittels PCR-basierter Techniken und anschließender Sequenzierung gelang erstmalig die Isolierung verschiedener BMMF-artiger DNA-Elemente aus allen drei Spezies. Insgesamt konnten aus 53 % der Wasserbüffelmilchproben vollständige zirkuläre Genome isoliert werden, wohingegen dies nur aus 11 % aller individuellen Schafmilchproben gelang. Während die Einzelgemelksproben der Ziegen negativ getestet wurden, konnten aus der Hälfte aller aus dem Einzelhandel stammenden Schaf- und Ziegenmilchproben BMMF-ähnliche Elemente isoliert werden. Des Weiteren standen Blutproben von Yaks, Zebus und Watussi-Rindern aus deutschen Zoos und Tierparks zur Verfügung. Im Blut dieser Spezies konnte lediglich bei einem Yak zirkuläre ssDNA nachgewiesen werden.

Im Rahmen dieser Studie konnte somit gezeigt werden, dass BMMF in mehr boviden Spezies vorkommen, als ursprünglich angenommen. Ob diese DNA-Elemente und CRESS-DNA-Viren Auswirkungen auf die Gesundheit von Tieren oder Menschen haben können, muss allerdings noch weiter erforscht werden.



## VIII. SUMMARY

### **Detection and characterization of circular single-stranded BMMF elements in dairy bovids**

Over the past decade, new techniques have enabled the discovery of a variety of novel, potentially pathogenic CRESS DNA viruses and other previously unknown ssDNA elements. In terms of food safety, novel circular ssDNA elements, so-called bovine meat and milk factors (BMMF), which could be isolated from cow milk, sera and cow milk products, received special attention. On the basis of global epidemiological studies, these molecules, which are supposed to be exclusively present in European cattle of taurine origin, are suspected of being indirectly involved in the development of human diseases such as colon and breast carcinomas and multiple sclerosis.

To gain clarity on the occurrence of these ssDNA elements in species other than *Bos taurus*, which are also kept for food production worldwide, individual milk samples from the species water buffalo (*Bubalus arnee f. bubalis*), sheep (*Ovis aries*) and goat (*Capra aegagrus f. hircus*) and additionally sheep and goat milk at retail were investigated in this study. By PCR-based analyses of the milk and subsequent sequencing, the isolation of different BMMF-like DNA elements was achieved for the first time from all three species. In total, complete circular genomes could be isolated from 53 % of the water buffalo milk samples, whereas this was only possible from 11 % of all individual sheep milk samples. While the individual goat milk samples tested negative, BMMF-like DNA elements could be isolated from half of all sheep and goat milk samples at retail. Furthermore, blood samples from yaks, zebus and watussi cattle from German zoos were available. In the blood of these species, circular ssDNA was detected in one yak only.

Overall, this study showed that BMMF are present in more bovid species than originally assumed. Whether these DNA-elements and CRESS-DNA-viruses can impact animal or human health still needs further research.



## IX. LITERATURVERZEICHNIS

ABBAS, A., TAYLOR, L. J., COLLMAN, R. G., BUSHMAN, F. D. & ICTV REPORT CONSORTIUM 2021. ICTV Virus Taxonomy Profile: Redondoviridae. *Journal of general virology*, 102, jgv001526.

AIMONE, C. D., HOYER, J. S., DYE, A. E., DEPPONG, D. O., DUFFY, S., CARBONE, I. & HANLEY-BOWDOIN, L. 2022. An experimental strategy for preparing circular ssDNA virus genomes for next-generation sequencing. *Journal of Virological Methods*, 300, 114405.

AJMONE-MARSAN, P., GARCIA, J. F. & LENSTRA, J. A. 2010. On the origin of cattle: How aurochs became cattle and colonized the world. *Evolutionary Anthropology: Issues, News, and Reviews*, 19, 148-157.

BARZON, L., LAVEZZO, E., MILITELLO, V., TOPPO, S. & PALÙ, G. 2011. Applications of Next-Generation Sequencing Technologies to Diagnostic Virology. *International Journal of Molecular Sciences*, 12, 7861-7884.

BORGHESE, A. 2005. Buffalo production and research. In: BORGHESE, A. (ed.) *REU Technical Series 67*. Rome, Italy: FAO Regional Office for Europe.

BRADLEY, D. G., MACHUGH, D. E., CUNNINGHAM, P. & LOFTUS, R. T. 1996. Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci U S A*, 93, 5131-5.

BUND, T., NIKITINA, E., CHAKRABORTY, D., ERNST, C., GUNST, K., BONEVA, B., TESSMER, C., VOLK, N., BROBEIL, A., WEBER, A., HEIKENWALDER, M., ZUR HAUSEN, H. & DE VILLIERS, E. M. 2021. Analysis of chronic inflammatory lesions of the colon for BMMF Rep antigen expression and CD68 macrophage interactions. *Proc Natl Acad Sci U S A*, 118.

BUNDESINSTITUT FÜR RISIKOBEWERTUNG (BFR) & MAX RUBNER-INSTITUT (MRI). 2019. *Neuartige Erreger in Rind und Kuhmilchprodukten: Weitere Forschung notwendig* [Online]. Available: <https://www.bfr.bund.de/cm/343/neuartige-erreger-in-rind-und-kuhmilchprodukten-weitere-forschung-notwendig.pdf> [Accessed 28 January 2022].

CHAN, D. S., LAU, R., AUNE, D., VIEIRA, R., GREENWOOD, D. C., KAMPMAN, E. & NORAT, T. 2011. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS One*, 6, e20456.

CHANDLER, M., DE LA CRUZ, F., DYDA, F., HICKMAN, A. B., MONCALIAN, G. & TON-HOANG, B. 2013. Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat Rev Microbiol*, 11, 525-38.

CHATTORAJ, D. K. 2000. Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol*, 37, 467-76.

CHEVAL, J., SAUVAGE, V., FRANGEUL, L., DACHEUX, L., GUIGON, G., DUMEY, N., PARIENTE, K., ROUSSEAU, C., DORANGE, F. & BERTHET, N. 2011. Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. *Journal of clinical microbiology*, 49, 3268-3275.

COCKRILL, W. R. 1977. The Water Buffalo: Domestic Animal of the Future. *The Bovine Practitioner*, 1977, 92-98.

COLLI, L., MILANESI, M., VAJANA, E., IAMARTINO, D., BOMBA, L., PUGLISI, F., DEL CORVO, M., NICOLAZZI, E. L., AHMED, S. S. E., HERRERA, J. R. V., CRUZ, L., ZHANG, S., LIANG, A., HUA, G., YANG, L., HAO, X., ZUO, F., LAI, S. J., WANG, S., LIU, R., GONG, Y., MOKHBER, M., MAO, Y., GUAN, F., VLAIC, A., VLAIC, B., RAMUNNO, L., COSENZA, G., AHMAD, A., SOYSAL, I., ÜNAL, E., KETUDAT-CAIRNS, M., GARCIA, J. F., UTSUNOMIYA, Y. T., BARUSELLI, P. S., AMARAL, M. E. J., PARNPAI, R., DRUMMOND, M. G., GALBUSERA, P., BURTON, J., HOAL, E., YUSNIZAR, Y., SUMANTRI, C., MOIOLI, B., VALENTINI, A., STELLA, A., WILLIAMS, J. L. & AJMONE-MARSAN, P. 2018. New Insights on Water Buffalo Genomic Diversity and Post-Domestication Migration Routes From Medium Density SNP Chip Data. *Front Genet*, 9, 53.

CUBRIC-CURIK, V., NOVOSEL, D., BRAJKOVIC, V., ROTA STABELLI, O., KREBS, S., SÖLKNER, J., ŠALAMON, D., RISTOV, S., BERGER, B., TRIVIZAKI, S., BIZELIS, I., FERENČAKOVIĆ, M., ROTHAMMER, S., KUNZ, E., SIMČIĆ, M., DOVČ, P., BUNEVSKI, G., BYTYQI, H., MARKOVIĆ, B., BRKA, M., KUME, K., STOJANOVIĆ, S., NIKOLOV, V., ZINOVIEVA, N., SCHÖNHERZ, A. A., GULDBRANDTSEN, B., ČAČIĆ, M., RADOVIĆ, S., MIRACLE, P., VERNESI, C., CURIK, I. & MEDUGORAC, I. 2021. Large-scale mitogenome sequencing reveals consecutive expansions of domestic taurine cattle and supports sporadic aurochs introgression. *Evolutionary Applications*.

CUNHA, T. J., KOGER, M. & WARNICK, A. G. 1963. *Crossbreeding Beef Cattle*, Gainesville, University of Florida Press.

DATTA, S., BUDHAULIYA, R., DAS, B., CHATTERJEE, S., VANLALHMUAKA & VEER, V. 2015. Next-generation sequencing in clinical virology: Discovery of new viruses. *World journal of virology*, 4, 265-276.

DE VILLIERS, E. M., GUNST, K., CHAKRABORTY, D., ERNST, C., BUND, T. & ZUR HAUSEN, H. 2019. A specific class of infectious agents isolated from bovine serum and dairy products and peritumoral colon cancer tissue. *Emerg Microbes Infect*, 8, 1205-1218.

DE VILLIERS, E. M. & ZUR HAUSEN, H. 2021. Bovine Meat and Milk Factors (BMMFs): Their Proposed Role in Common Human Cancers and Type 2 Diabetes Mellitus. *Cancers (Basel)*, 13.

DEAN, F. B., NELSON, J. R., GIESLER, T. L. & LASKEN, R. S. 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res*, 11, 1095-9.

DECKER, J. E., MCKAY, S. D., ROLF, M. M., KIM, J., MOLINA ALCALA, A., SONSTEGARD, T. S., HANOTTE, O., GOTHERSTROM, A., SEABURY, C. M., PRAHARANI, L., BABAR, M. E., CORREIA DE ALMEIDA REGITANO, L., YILDIZ, M. A., HEATON, M. P., LIU, W. S., LEI, C. Z., REECY, J. M., SAIF-UR-REHMAN, M., SCHNABEL, R. D. & TAYLOR, J. F. 2014. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet*, 10, e1004254.

DEL SOLAR, G., GIRALDO, R., RUIZ-ECHEVARRÍA, M. J., ESPINOSA, M. & DÍAZ-OREJAS, R. 1998. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev*, 62, 434-64.

DIEMER, G. S. & STEDMAN, K. M. 2012. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biology Direct*, 7, 13.

DÍEZ-VILLASEÑOR, C. & RODRIGUEZ-VALERA, F. 2019. CRISPR analysis suggests that small circular single-stranded DNA smacoviruses infect Archaea instead of humans. *Nature Communications*, 10, 294.

DUBEUF, J. P., MORAND-FEHR, P. & RUBINO, R. 2004. Situation, changes and future of goat industry around the world. *Small Ruminant Research*, 51, 165-173.

DUFFY, S., SHACKELTON, L. A. & HOLMES, E. C. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*, 9, 267-276.

EILEBRECHT, S., HOTZ-WAGENBLATT, A., SARACHAGA, V., BURK, A., FALIDA, K., CHAKRABORTY, D., NIKITINA, E., TESSMER, C., WHITLEY, C., SAUERLAND, C., GUNST, K., GREWE, I. & BUND, T. 2018. Expression and replication of virus-like circular DNA in human cells. *Sci Rep*, 8, 2851.

FARVID, M. S., STERN, M. C., NORAT, T., SASAZUKI, S., VINEIS, P., WEIJENBERG, M. P., WOLK, A., WU, K., STEWART, B. W. & CHO, E. 2018. Consumption of red and processed meat and breast cancer incidence: A systematic review and meta-analysis of prospective studies. *Int J Cancer*, 143, 2787-2799.

FELIUS, M., BEERLING, M.-L., BUCHANAN, D. S., THEUNISSEN, B., KOOLMEES, P. A. & LENSTRA, J. A. 2014. On the History of Cattle Genetic Resources. *Diversity*, 6, 705-750.

FERLAY, J., ERVIK, M., LAM, F., COLOMBET, M., MERY, L., PIÑEROS, M., ZNAOR, A., SOERJOMATARAM, I. & BRAY, F. 2020. *Global Cancer Observatory: Cancer Today* [Online]. Lyon, France: International Agency for Research on Cancer. Available: <https://gco.iarc.fr/today/fact-sheets-cancers> [Accessed 10 January 2022].

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). *Food and agriculture data* [Online]. Available: <https://www.fao.org/faostat/en/> [Accessed 10 January 2022].

FUNK, M., GUNST, K., LUCANSKY, V., MULLER, H., ZUR HAUSEN, H. & DE VILLIERS, E. M. 2014. Isolation of protein-associated circular DNA from healthy cattle serum. *Genome Announc*, 2.

GILBERT, M., NICOLAS, G., CINARDI, G., VAN BOECKEL, T. P., VANWAMBEKE, S. O., WINT, G. R. W. & ROBINSON, T. P. 2018. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data*, 5, 180227.

GOODMAN, R. M. 1977. Infectious DNA from a whitefly-transmitted virus of *Phaseolus vulgaris*. *Nature*, 266, 54-55.

GORBALENYA, A. E. & KOONIN, E. V. 1993. Helicases: amino acid sequence comparisons and structure-function relationships. *Current Opinion in Structural Biology*, 3, 419-429.

GORBALENYA, A. E., KOONIN, E. V. & WOLF, Y. I. 1990. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett*, 262, 145-8.

GOVERNMENT OF INDIA – DEPARTMENT OF ANIMAL HUSBANDRY AND DAIRYING. 2020. *Basic Animal Husbandry Statistics-2020* [Online]. Krishi Bhawan, New Delhi. Available: <https://dahd.nic.in/sites/default/filess/Book%20BAHS%202020-Final.pdf> [Accessed 09 February 2022].

GOVERNMENT OF INDIA – DEPARTMENT OF ANIMAL HUSBANDRY, DAIRYING AND FISHERIES. 2015. *Estimated Livestock Population Breed Wise - Based on Breed Survey 2013* [Online]. Krishi Bhawan, New Delhi. Available: [https://dahd.nic.in/sites/default/filess/Breeding%20Survey%20Book\\_0.pdf](https://dahd.nic.in/sites/default/filess/Breeding%20Survey%20Book_0.pdf) [Accessed 09 February 2022].

GRIGSON, C. 1991. An African origin for African cattle? — some archaeological evidence. *African Archaeological Review*, 9, 119-144.

GRONENBORN, B., RANDLES, J. W., VETTEN, H. J. & THOMAS, J. E. 2021. Create one new family (Metaxyviridae) with one new genus (Cofodevirus) and one species (Coconut foliar decay virus) moved from the family Nanoviridae (Mulgavirales). *International Committee for Taxonomy of Viruses proposal (Taxoprop)*, number 2020.022P.

GROVES, C. & GRUBB, P. 2011. *Ungulate Taxonomy*, Baltimore, Johns Hopkins University Press.

GUNST, K., ZUR HAUSEN, H. & DE VILLIERS, E. M. 2014. Isolation of bacterial plasmid-related replication-associated circular DNA from a serum sample of a multiple sclerosis patient. *Genome Announc*, 2.

HAILE, A., WORKNEH, A., TEGEGNE, A., KEBEDE, N. & DESSIE, T. 2011. Breeding Strategy to Improve Ethiopian Boran Cattle for Meat and Milk Production. *Working Paper 26*. Nairobi, Kenya: ILRI: International Livestock Research Institute.

HEGDE, N. G. 2019. Buffalo Husbandry for Sustainable Development of Small Farmers in India and other Developing Countries. *Asian Journal of Research in Animal and Veterinary Sciences*, 1-20.

HELMER, D., GOURICHON, L., MONCHOT, H., PETERS, J. & SAÑA SEGUI, M. 2005. Identifying early domestic cattle from Pre-Pottery Neolithic sites on the Middle Euphrates using sexual dimorphism. *The first steps of animal domestication: new archaeozoological approaches*. London (GB).

HIEMSTRA, S. J., DE HAAS, Y., MÄKI-TANILA, A. & GANDINI, G. 2010. *Local cattle breeds in Europe: development of policies and strategies for self-sustaining breeds*, Netherlands, Wageningen Academic Publishers.

HIENDLEDER, S., KAUPE, B., WASSMUTH, R. & JANKE, A. 2002. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc Biol Sci*, 269, 893-904.

HUXLEY, R. R., ANSARY-MOGHADDAM, A., CLIFTON, P., CZERNICHOW, S., PARR, C. L. & WOODWARD, M. 2009. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int J Cancer*, 125, 171-80.

INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES (ICTV). *Taxonomy* [Online]. Available: <https://talk.ictvonline.org/taxonomy/> [Accessed 10 January 2022].

JIANG, W., JU, C., JIANG, H. & ZHANG, D. 2014. Dairy foods intake and risk of Parkinson's disease: a dose-response meta-analysis of prospective cohort studies. *Eur J Epidemiol*, 29, 613-9.

JOHNE, R., MÜLLER, H., RECTOR, A., VAN RANST, M. & STEVENS, H. 2009. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends Microbiol*, 17, 205-11.

KAZEROUNI, N., SINHA, R., HSU, C. H., GREENBERG, A. & ROTHMAN, N. 2001. Analysis of 200 food items for benzo[a]pyrene and estimation of its intake in an epidemiologic study. *Food Chem Toxicol*, 39, 423-36.

KEMP, S. J., DARVASI, A., SOLLER, M. & TEALE, A. J. 1996. Genetic control of resistance to trypanosomiasis. *Veterinary Immunology and Immunopathology*, 54, 239-243.

KRABERGER, S., ARGÜELLO-ASTORGA, G. R., GREENFIELD, L. G., GALILEE, C., LAW, D., MARTIN, D. P. & VARSANI, A. 2015. Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. *Infect Genet Evol*, 31, 73-86.

KRABERGER, S., SCHMIDLIN, K., FONTENELE, R. S., WALTERS, M. & VARSANI, A. 2019. Unravelling the Single-Stranded DNA Virome of the New Zealand Blackfly. *Viruses*, 11, 532.

KRUPOVIC, M., GHABRIAL, S. A., JIANG, D. & VARSANI, A. 2016. Genomoviridae: a new family of widespread single-stranded DNA viruses. *Archives of Virology*, 161, 2633-2643.

KRUPOVIC, M., RAVANTTI, J. J. & BAMFORD, D. H. 2009. Geminiviruses: a tale of a plasmid becoming a virus. *BMC Evolutionary Biology*, 9, 112.

KRUPOVIC, M., VARSANI, A., KAZLAUSKAS, D., BREITBART, M., DELWART, E., ROSARIO, K., YUTIN, N., WOLF, Y. I., HARRACH, B., ZERBINI, F. M., DOLJA, V. V., KUHN, J. H. & KOONIN, E. V. 2020. Cressdnnaviricota: a Virus Phylum Unifying Seven Families of Rep-Encoding Viruses with Single-Stranded, Circular DNA Genomes. *J Virol*, 94.

KUMAR, S., NAGARAJAN, M., SANDHU, J. S., KUMAR, N., BEHL, V. & NISHANTH, G. 2007. Mitochondrial DNA analyses of Indian water buffalo support a distinct genetic origin of river and swamp buffalo. *Anim Genet*, 38, 227-32.

LAMBERTO, I., GUNST, K., MULLER, H., ZUR HAUSEN, H. & DE VILLIERS, E. M. 2014. Mycovirus-like DNA virus sequences from cattle serum and human brain and serum samples from multiple sclerosis patients. *Genome Announc*, 2.

LARSSON, S. C. & ORSINI, N. 2013. Red Meat and Processed Meat Consumption and All-Cause Mortality: A Meta-Analysis. *American Journal of Epidemiology*, 179, 282-289.

LECHMANN, J., ACKERMANN, M., KAISER, V. & BACHOFEN, C. 2021. Viral infections shared between water buffaloes and small ruminants in Switzerland. *Journal of Veterinary Diagnostic Investigation*, 33, 894-905.

LEFEUVRE, P., LETT, J. M., VARSANI, A. & MARTIN, D. P. 2009. Widely conserved recombination patterns among single-stranded DNA viruses. *J Virol*, 83, 2697-707.

LESLIE, D. M., JR. & SCHALLER, G. B. 2009. Bos grunniens and Bos mutus (Artiodactyla: Bovidae). *Mammalian Species*, 1-17.

LOFTUS, R. T., MACHUGH, D. E., BRADLEY, D. G., SHARP, P. M. & CUNNINGHAM, P. 1994. Evidence for two independent domestications of cattle. *Proc Natl Acad Sci U S A*, 91, 2757-61.

LUIKART, G., GIELLY, L., EXCOFFIER, L., VIGNE, J. D., BOUVET, J. & TABERLET, P. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci U S A*, 98, 5927-32.

LUO, X., ZHOU, Y., ZHANG, B., ZHANG, Y., WANG, X., FENG, T., LI, Z., CUI, K., WANG, Z., LUO, C., LI, H., DENG, Y., LU, F., HAN, J., MIAO, Y., MAO, H., YI, X., AI, C., WU, S., LI, A., WU, Z., ZHUO, Z., DA GIANG, D., MITRA, B., VAHIDI, M. F., MANSOOR, S., AL-BAYATTI, S. A., SARI, E. M., GORKHALI, N. A., PRASTOWO, S., SHAFIQUE, L., YE, G., QIAN, Q., CHEN, B., SHI, D., RUAN, J. & LIU, Q. 2020. Understanding divergent domestication traits from the whole-genome sequencing of swamp- and river-buffalo populations. *National Science Review*, 7, 686-701.

MACHUGH, D. E., SHRIVER, M. D., LOFTUS, R. T., CUNNINGHAM, P. & BRADLEY, D. G. 1997. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (Bos taurus and Bos indicus). *Genetics*, 146, 1071-86.

MANOMOHAN, V., SARAVANAN, R., PICHLER, R., MURALI, N., SIVAKUMAR, K., SUDHAKAR, K., NACHIAPPAN, R. K. & PERIASAMY, K. 2021. Legacy of draught cattle breeds of South India: Insights into population structure, genetic admixture and maternal origin. *PLoS One*, 16, e0246497-e0246497.

MANUELIDIS, L. 2011. Nuclease resistant circular DNAs copurify with infectivity in scrapie and CJD. *J Neurovirol*, 17, 131-45.

MATOS, T., QUEIROZ, J. A. & BÜLOW, L. 2014. Plasmid DNA purification using a multimodal chromatography resin. *Journal of Molecular Recognition*, 27, 184-189.

MBUZA, F. M. B. 1995. The indigenous domestic animal genetic resources of Uganda. *Animal Genetic Resources Information*, 15, 23-42.

MEDUGORAC, I., GRAF, A., GROHS, C., ROTHAMMER, S., ZAGDSUREN, Y., GLADYR, E., ZINOVIEVA, N., BARBIERI, J., SEICHTER, D., RUSS, I., EGGEN, A., HELLENTHAL, G., BREM, G., BLUM, H., KREBS, S. & CAPITAN, A. 2017. Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. *Nature Genetics*, 49, 470-475.

MELNIK, B. C. 2009. Milk – The promoter of chronic Western diseases. *Medical Hypotheses*, 72, 631-639.

MENEZES, G. C. C., VALADARES FILHO, S. C., RUAS, J. R. M., DETMANN, E., MENEZES, A. C., ZANETT, D., MARIZ, L. D. S., RENNÓ, L. N. & SILVA JUNIOR, J. M. 2014. Meat Production in a Feedlot System of Zebu—Holstein Steers and Heifers with Dairy Genetics: Productive and Biological Analyses. *The Scientific World Journal*, 2014, 371968.

MILLER, B. A. & LU, C. D. 2019. Current status of global dairy goat production: an overview. *Asian-Australasian Journal of Animal Sciences*, 32.

MUKASA-MUGERWA, E. 1989. *A Review of Reproductive Performance of Female Bos Indicus (Zebu) Cattle*, Addis Ababa, Ethiopia, International Livestock Centre for Africa.

MURRAY, M., TRAIL, J. C. M., DAVIS, C. E. & BLACK, S. J. 1984. Genetic Resistance to African Trypanosomiasis. *The Journal of Infectious Diseases*, 149, 311-319.

MWAI, O., HANOTTE, O., KWON, Y. J. & CHO, S. 2015. African Indigenous Cattle: Unique Genetic Resources in a Rapidly Changing World. *Asian-Australas J Anim Sci*, 28, 911-21.

NADERI, S., REZAEI, H.-R., POMPANON, F., BLUM, M. G. B., NEGRINI, R., NAGHASH, H.-R., BALKIZ, Ö., MASHKOUR, M., GAGGIOTTI, O. E., AJMONE-MARSAN, P., KENCE, A., VIGNE, J.-D. & TABERLET, P. 2008. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences*, 105, 17659-17664.

NAGAO, M., HONDA, M., SEINO, Y., YAHAGI, T. & SUGIMURA, T. 1977. Mutagenicities of smoke condensates and the charred surface of fish and meat. *Cancer Letters*, 2, 221-226.

NASH, T. E., DALLAS, M. B., REYES, M. I., BUHRMAN, G. K., ASCENCIO-IBAÑEZ, J. T. & HANLEY-BOWDOIN, L. 2011. Functional Analysis of a Novel Motif Conserved across Geminivirus Rep Proteins. *Journal of Virology*, 85, 1182-1192.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION (NCBI). Bethesda (MD): National Library of Medicine (US). Available: <https://www.ncbi.nlm.nih.gov/> [Accessed 08 February 2022].

OLSEN, S. J. 1990. Fossil Ancestry of the Yak, Its Cultural Significance and Domestication in Tibet. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 142, 73-100.

PADIDAM, M., SAWYER, S. & FAUQUET, C. M. 1999. Possible Emergence of New Geminiviruses by Frequent Recombination. *Virology*, 265, 218-225.

PAN, A., SUN, Q., BERNSTEIN, A. M., SCHULZE, M. B., MANSON, J. E., STAMPFER, M. J., WILLETT, W. C. & HU, F. B. 2012. Red meat consumption and mortality: results from 2 prospective cohort studies. *Arch Intern Med*, 172, 555-63.

PAN, A., SUN, Q., BERNSTEIN, A. M., SCHULZE, M. B., MANSON, J. E., WILLETT, W. C. & HU, F. B. 2011. Red meat consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. *Am J Clin Nutr*, 94, 1088-96.

PASHA, T. & HAYAT, Z. 2012. Present situation and future perspective of buffalo production in Asia. *Journal of Animal and Plant Sciences*, 22, 250-256.

PEDROSA, S., UZUN, M., ARRANZ, J. J., GUTIÉRREZ-GIL, B., SAN PRIMITIVO, F. & BAYÓN, Y. 2005. Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. *Proc Biol Sci*, 272, 2211-7.

PITT, D., SEVANE, N., NICOLAZZI, E. L., MACHUGH, D. E., PARK, S. D. E., COLLI, L., MARTINEZ, R., BRUFORD, M. W. & OROZCO-TERWENGEL, P. 2019. Domestication of cattle: Two or three events? *Evolutionary Applications*, 12, 123-136.

POHL, S., HABERMANN, D., LINK, E. K., FUX, R., BOLDT, C. L., FRANZ, C. M. A. P., HÖLZEL, C. & KLEMPT, M. 2022. Detection of DNA sequences attributed to bovine meat and milk factors (BMMF/SPHINX) in food-related samples. *Food Control*, 135, 108779.

PORTER, V. 1996. *Goats of the world*, Ipswich, UK, Farming Press.

PRUSINER, S. B. 1982. Novel Proteinaceous Infectious Particles Cause Scrapie. *Science*, 216, 136-144.

PULINA, G., MILÁN, M. J., LAVÍN, M. P., THEODORIDIS, A., MORIN, E., CAPOTE, J., THOMAS, D. L., FRANCESCONI, A. H. D. & CAJA, G. 2018. Invited review: Current production trends, farm structures, and economics of the dairy sheep and goat sectors. *J Dairy Sci*, 101, 6715-6729.

QIU, Q., WANG, L., WANG, K., YANG, Y., MA, T., WANG, Z., ZHANG, X., NI, Z., HOU, F., LONG, R., ABBOTT, R., LENSTRA, J. & LIU, J. 2015. Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nature Communications*, 6, 10283.

QIU, Q., ZHANG, G., MA, T., QIAN, W., WANG, J., YE, Z., CAO, C., HU, Q., KIM, J., LARKIN, D. M., AUVEL, L., CAPITANU, B., MA, J., LEWIN, H. A., QIAN, X., LANG, Y., ZHOU, R., WANG, L., WANG, K., XIA, J., LIAO, S., PAN, S., LU, X., HOU, H., WANG, Y., ZANG, X., YIN, Y., MA, H., ZHANG, J., WANG, Z., ZHANG, Y., ZHANG, D., YONEZAWA, T., HASEGAWA, M., ZHONG, Y., LIU, W., ZHANG, Y., HUANG, Z., ZHANG, S., LONG, R., YANG, H., WANG, J., LENSTRA, J. A., COOPER, D. N., WU, Y., WANG, J., SHI, P., WANG, J. & LIU, J. 2012. The yak genome and adaptation to life at high altitude. *Nature Genetics*, 44, 946-949.

RECTOR, A., TACHEZY, R. & RANST, M. V. 2004. A Sequence-Independent Strategy for Detection and Cloning of Circular DNA Virus Genomes by Using Multiply Primed Rolling-Circle Amplification. *Journal of Virology*, 78, 4993-4998.

RICHI, E. B., BAUMER, B., CONRAD, B., DARIOLI, R., SCHMID, A. & KELLER, U. 2015. Health Risks Associated with Meat Consumption: A Review of Epidemiological Studies. *Int J Vitam Nutr Res*, 85, 70-8.

ROSARIO, K., BREITBART, M., HARRACH, B., SEGALÉS, J., DELWART, E., BIAGINI, P. & VARSANI, A. 2017. Revisiting the taxonomy of the family Circoviridae: establishment of the genus *Cyclovirus* and removal of the genus *Gyrovirus*. *Arch Virol*, 162, 1447-1463.

ROSARIO, K., DAYARAM, A., MARINOV, M., WARE, J., KRABERGER, S., STAINTON, D., BREITBART, M. & VARSANI, A. 2012a. Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J Gen Virol*, 93, 2668-2681.

ROSARIO, K., DUFFY, S. & BREITBART, M. 2012b. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch Virol*, 157, 1851-71.

SAMRAJ, A. N., PEARCE, O. M., LÄUBLI, H., CRITTENDEN, A. N., BERGFELD, A. K., BANDA, K., GREGG, C. J., BINGMAN, A. E., SECREST, P., DIAZ, S. L., VARKI, N. M. & VARKI, A. 2015. A red meat-derived glycan promotes inflammation and cancer progression. *Proc Natl Acad Sci U S A*, 112, 542-7.

SARGSYAN, A. & DUBASI, H. B. 2021. Milk Consumption and Prostate Cancer: A Systematic Review. *World J Mens Health*, 39, 419-428.

SCHIBLER, J., ELSNER, J. & SCHLUMBAUM, A. 2014. Incorporation of aurochs into a cattle herd in Neolithic Europe: single event or breeding? *Scientific Reports*, 4, 5798.

SIMMONDS, P., ADAMS, M. J., BENKÓ, M., BREITBART, M., BRISTER, J. R., CARSTENS, E. B., DAVISON, A. J., DELWART, E., GORBALENYA, A. E., HARRACH, B., HULL, R., KING, A. M., KOONIN, E. V., KRUPOVIC, M., KUHN, J. H., LEFKOWITZ, E. J., NIBERT, M. L., ORTON, R., ROOSSINCK, M. J., SABANADZOVIC, S., SULLIVAN, M. B., SUTTLE, C. A., TESH, R. B., VANDER VLUGT, R. A., VARSANI, A. & ZERBINI, F. M. 2017. Consensus statement: Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol*, 15, 161-168.

STEEL, O., KRABERGER, S., SIKORSKI, A., YOUNG, L. M., CATCHPOLE, R. J., STEVENS, A. J., LADLEY, J. J., CORAY, D. S., STAINTON, D., DAYARAM, A., JULIAN, L., VAN BYSTERVELDT, K. & VARSANI, A. 2016. Circular replication-associated protein encoding DNA viruses identified in the faecal matter of various animals in New Zealand. *Infect Genet Evol*, 43, 151-64.

STENLUND, A. 2003. Initiation of DNA replication: lessons from viral initiator proteins. *Nat Rev Mol Cell Biol*, 4, 777-85.

SUGIMURA, T., NAGAO, M., KAWACHI, T., HONDA, M., YAHAGI, T., SEINO, Y., SATO, S., MATSUKURA, N., MATSUSHIMA, T., SHIRAI, A., SAWAMURA, M. & MATSUMOTO, H. 1977. Mutagen-carcinogens in food, with special reference to highly mutagenic pyrolytic products in broiled foods. In: HIATT, H. H., WATSON, J. D. & WINSTEN, J. A. (eds.) *Origins of Human Cancer. Book C*. New York: Cold Spring Harbor Laboratory.

SUN, T., SHEN, J., ACHILLI, A., CHEN, N., CHEN, Q., DANG, R., ZHENG, Z., ZHANG, H., ZHANG, X., WANG, S., ZHANG, T., LU, H., MA, Y., JIA, Y., CAPODIFERRO, M. R., HUANG, Y., LAN, X., CHEN, H., JIANG, Y. & LEI, C. 2020. Genomic analyses reveal distinct genetic architectures and selective pressures in buffaloes. *GigaScience*, 9.

SUNG, H., FERLAY, J., SIEGEL, R. L., LAVERSANNE, M., SOERJOMATARAM, I., JEMAL, A. & BRAY, F. 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*, 71, 209-249.

TAYE, M., KIM, J., YOON, S. H., LEE, W., HANOTTE, O., DESSIE, T., KEMP, S., MWAI, O. A., CAETANO-ANOLLES, K. & CHO, S. 2017. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. *BMC genetics*, 18, 1-14.

THOMAS, J. E., GRONENBORN, B., HARDING, R. M., MANDAL, B., GRIGORAS, I., RANDLES, J. W., SANO, Y., TIMCHENKO, T., VETTEN, H. J., YEH, H. H., ZIEBELL, H. & ICTV REPORT, C. 2021. ICTV Virus Taxonomy Profile: Nanoviridae. *J Gen Virol*, 102.

TIMON, V. M. & HANRAHAN, J. P. 1986. FAO Animal production and health paper 58. *Small ruminant production in the developing countries*. Food and Agriculture Organization of the United Nations, Rome.

TISCHER, I., GELDERBLOM, H., VETTERMANN, W. & KOCH, M. A. 1982. A very small porcine virus with circular single-stranded DNA. *Nature*, 295, 64-66.

TOCHETTO, C., MUTERLE VARELA, A. P., ALVES DE LIMA, D., LOIKO, M. R., MENGUE SCHEFFER, C., PINTO PAIM, W., CERVA, C., SCHMIDT, C., CIBULSKI, S. P., CANO ORTIZ, L., CALLEGARI JACQUES, S. M., FRANCO, A. C., QUOOS MAYER, F. & ROEHE, P. M. 2020. Viral DNA genomes in sera of farrowing sows with or without stillbirths. *PLoS One*, 15, e0230714.

VARSANI, A. & KRUPOVIC, M. 2017. Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family Genomoviridae. *Virus Evolution*, 3.

VARSANI, A. & KRUPOVIC, M. 2021. Family Genomoviridae: 2021 taxonomy update. *Archives of virology*, 166, 2911-2926.

VARSANI, A., ROUMAGNAC, P., FUCHS, M., NAVAS-CASTILLO, J., MORIONES, E., IDRIS, A., BRIDDON, R. W., RIVERA-BUSTAMANTE, R., MURILO ZERBINI, F. & MARTIN, D. P. 2017. Capulavirus and Grablovirus: two new genera in the family Geminiviridae. *Archives of Virology*, 162, 1819-1831.

WAITS, K., EDWARDS, M. J., COBB, I. N., FONTENELE, R. S. & VARSANI, A. 2018. Identification of an anellovirus and genomoviruses in ixodid ticks. *Virus Genes*, 54, 155-159.

WANG, S., CHEN, N., CAPODIFERRO, M. R., ZHANG, T., LANCIONI, H., ZHANG, H., MIAO, Y., CHANTHAKHOUN, V., WANAPAT, M., YINDEE, M., ZHANG, Y., LU, H., CAPORALI, L., DANG, R., HUANG, Y., LAN, X., PLATH, M., CHEN, H., LENSTRA, J. A., ACHILLI, A. & LEI, C. 2017. Whole Mitogenomes Reveal the History of Swamp Buffalo: Initially Shaped by Glacial Periods and Eventually Modelled by Domestication. *Sci Rep*, 7, 4708.

WHITLEY, C., GUNST, K., MULLER, H., FUNK, M., ZUR HAUSEN, H. & DE VILLIERS, E. M. 2014. Novel replication-competent circular DNA molecules from healthy cattle serum and milk and multiple sclerosis-affected human brain tissue. *Genome Announc*, 2, e00849-14.

WIENER, G., JIANLIN, H. & LONG, R. 2003. The Yak. 2 ed. Bangkok, Thailand: FAO Regional Office for Asia and the Pacific.

WINER, S., ASTSATUROV, I., CHEUNG, R. K., SCHRADE, K., GUNARATNAM, L., WOOD, D. D., MOSCARELLO, M. A., O'CONNOR, P., MCKERLIE, C., BECKER, D. J. & DOSCH, H. M. 2001. T cells of multiple sclerosis patients target a common environmental peptide that causes encephalitis in mice. *J Immunol*, 166, 4751-6.

WRIGHT, E. A., HECKEL, T., GROENENDIJK, J., DAVIES, J. W. & BOULTON, M. I. 1997. Splicing features in maize streak virus virion- and complementary-sense gene expression. *Plant J*, 12, 1285-97.

YANO, M., WAKABAYASHI, K., TAHIRA, T., ARAKAWA, N., NAGAO, M. & SUGIMURA, T. 1988. Presence of nitrosable mutagen precursors in cooked meat and fish. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 202, 119-123.

YU, X., LI, B., FU, Y., JIANG, D., GHABRIAL, S. A., LI, G., PENG, Y., XIE, J., CHENG, J., HUANG, J. & YI, X. 2010. A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc Natl Acad Sci U S A*, 107, 8387-92.

ZEDER, M. A. 2008. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proceedings of the National Academy of Sciences*, 105, 11597-11604.

ZEDER, M. A., EMSHWILLER, E., SMITH, B. D. & BRADLEY, D. G. 2006. Documenting domestication: the intersection of genetics and archaeology. *Trends in Genetics*, 22, 139-155.

- ZERBINI, F. M., BRIDDON, R. W., IDRIS, A., MARTIN, D. P., MORIONES, E., NAVAS-CASTILLO, J., RIVERA-BUSTAMANTE, R., ROUMAGNAC, P., VARSANI, A. & ICTV REPORT, C. 2017. ICTV Virus Taxonomy Profile: Geminiviridae. *J Gen Virol*, 98, 131-133.
- ZEUNER, F. E. 1963. *A history of domesticated animals*, London, Hutchinson.
- ZHANG, Y., COLLI, L. & BARKER, J. S. F. 2020. Asian water buffalo: domestication, history and genetics. *Anim Genet*, 51, 177-191.
- ZHANG, Y., VANKAN, D., ZHANG, Y. & BARKER, J. S. 2011. Genetic differentiation of water buffalo (*Bubalus bubalis*) populations in China, Nepal and south-east Asia: inferences on the region of domestication of the swamp buffalo. *Anim Genet*, 42, 366-77.
- ZHAO, L., LAVINGTON, E. & DUFFY, S. 2019a. A comprehensive genealogy of the replication associated protein of CRESS DNA viruses reveals a single origin of intron-containing Rep. *bioRxiv*, 687855.
- ZHAO, L., ROSARIO, K., BREITBART, M. & DUFFY, S. 2019b. Eukaryotic Circular Rep-Encoding Single-Stranded DNA (CRESS DNA) Viruses: Ubiquitous Viruses With Small Genomes and a Diverse Host Range. *Adv Virus Res*, 103, 71-133.
- ZHOU, X., LIU, Y., CALVERT, L., MUÑOZ, C., OTIM-NAPE, G. W., ROBINSON, D. J. & HARRISON, B. D. 1997. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology*, 78, 2101-2111.
- ZUR HAUSEN, H. 2012. Red meat consumption and cancer: reasons to suspect involvement of bovine infectious factors in colorectal cancer. *Int J Cancer*, 130, 2475-83.

ZUR HAUSEN, H., BUND, T. & DE VILLIERS, E. M. 2017. Infectious Agents in Bovine Red Meat and Milk and Their Potential Role in Cancer and Other Chronic Diseases. In: HUNTER, E., BISTER, K. & COMPANS, R. W. (eds.) *Viruses, Genes, and Cancer*. 2017/03/30 ed. Switzerland: Springer.

ZUR HAUSEN, H., BUND, T. & DE VILLIERS, E. M. 2019. Specific nutritional infections early in life as risk factors for human colon and breast cancers several decades later. *Int J Cancer*, 144, 1574-1583.

ZUR HAUSEN, H. & DE VILLIERS, E. M. 2015. Dairy cattle serum and milk factors contributing to the risk of colon and breast cancers. *Int J Cancer*, 137, 959-67.

ZUR HAUSEN, H. & DE VILLIERS, E. M. 2016. Risikofaktoren für Krebs und MS in Kuhmilch und Rindfleisch? *Biologie in unserer Zeit*, 46, 26-31.

ZYGOYIANNIS, D. 2006. Sheep production in the world and in Greece. *Small Ruminant Research*, 62, 143-147.



## **X. EINSENDER DES PROBENMATERIALS**

### **1. Blutproben**

Alle in dieser Doktorarbeit untersuchten Blutproben von Zebus, Watussi-Rindern, Wasserbüffeln und Yaks stammen von Tieren aus deutschen Zoos und Tierparks. Für das Entnehmen, Aliquotieren, Beschriften und Versenden der Proben geht mein herzlicher Dank an die/den:

- **Arche Warder** und Frau Dr. Anabell Jandowsky
- **Tierpark Berlin**, Herrn Dr. Andreas Pauly und Herrn Dr. Marco Roller
- **Tierpark Gera** und Herrn Steffen Horn
- **Tierpark Hagenbeck** und Herrn Dr. Michael Flügger
- **Wilhelma** in Stuttgart und Herrn Dr. Tobias Knauf-Witzens
- **Zoo Dresden** und Herrn Dr. Dimitri Widmer
- **ZOOM Erlebniswelt** in Gelsenkirchen und Frau Dr. Pia Krawinkel
- **Zoologischen Stattgarten Karlsruhe** und Frau Julia Heckmann
- **Zoo Stralsund** und Herrn Dr. Christoph Langner

Ich möchte mich bei allen beteiligten Einrichtungen und Personen ganz herzlich bedanken. Ohne ihre Unterstützung wäre meine Arbeit in dieser Form nicht realisierbar gewesen.

### **2. Milchproben**

Alle untersuchten Milchproben wurden von deutschen Wasserbüffel-, Schaf- und Ziegenhaltern bei der regelmäßigen Melkroutine gewonnen. Bei allen Einsendern bedanke ich mich von ganzem Herzen für das Zusenden der Proben, da diese Arbeit ohne ihre Beteiligung und Mühe nicht in diesem Umfang zustande gekommen wäre

.

## **XI. DANKSAGUNG**

Zu Beginn möchte ich mich ganz herzlich bei Prof. Dr. Dr. h.c. Erwin Märtlbauer dafür bedanken, dass er mir am Lehrstuhl die Möglichkeit gegeben hat, an diesem brandaktuellen und höchst spannenden Thema zu forschen, meine eigenen Ideen und Vorstellungen einbringen und umsetzen zu dürfen. Vielen Dank für die Unterstützung meiner Arbeit und meines weiteren beruflichen Werdeganges.

Ein ganz besonderer Dank gilt meiner Betreuerin Dr. Andrea Stockmaier-Didier. Dein fundiertes und breitgefächertes Wissen zusammen mit Deiner hilfsbereiten und sympathischen Art haben mir geholfen, stets das Beste aus mir und meinem Thema herauszuholen. Für jede Frage und jedes Problem hattest Du zwei offene Ohren, hast Dir immer Zeit für meine Belange genommen und warst für jede meiner Ideen offen. Ich hätte mir keine bessere Betreuung wünschen können.

Außerdem möchte ich mich bei Dr. Richard Dietrich bedanken. Ohne Deine Fachkenntnisse und wegweisenden Ratschläge hätte meine Arbeit nicht diese Qualität erlangt. Danke für Deine Anregungen und konstruktive Kritik. Danke auch an Dr. Nadja Jeßberger und Dr. Kristina Schauer für die gute Zusammenarbeit und für alle Hilfestellungen im Labor.

Ganz herzlich bedanken möchte ich mich auch bei Prof. Dr. Dr. h.c. Gerd Sutter und Dr. Robert Fux vom Lehrstuhl für Virologie. Die ersten Wochen meiner Promotionszeit unter Ihrer Aufsicht waren außerordentlich lehrreich und der Beginn einer erfolgversprechenden und zukunftsweisenden Forschungsarbeit in Kooperation beider Institute. Ein großes Dankeschön geht dabei auch an Dr. Ellen Link für die ausgezeichnete Einarbeitung im Labor. Bedanken möchte ich mich außerdem bei Dr. Michael Lehmann ohne dessen ansteckende Begeisterung für die Forschung und die stets wohlgesonnenen Ratschläge ich diesen Weg vermutlich nie eingeschlagen hätte.

Des Weiteren möchte ich mich nochmals herzlich bei allen in Kapitel „X. Einsender des Probenmaterials“ aufgelisteten Zoos, Tierparks und Tierärzten bedanken, die meine Doktorarbeit mit der Bereitstellung von Probenmaterial verschiedener Rinderspezies unterstützt haben. Zusätzlich geht ein großes Dankeschön an alle Personen, die mir in mühevoller Arbeit Milchproben ihrer Wasserbüffel, Schafe und Ziegen zugesandt und somit meine Untersuchungen ermöglicht haben.

Natürlich möchte ich mich außerdem bei allen Mitarbeiterinnen des Lehrstuhls für Milchhygiene bedanken. Durch das fortwährend harmonische Arbeitsklima und die freundliche gegenseitige Unterstützung habe ich meinen Arbeitsplatz schnell lieben gelernt. Ein großer Dank geht an Julia, die mich so tatkräftig im Labor unterstützt und dabei hilft, die Menge der vielen Proben zu bewältigen. Bei Bruni möchte ich mich besonders dafür bedanken, dass sie mich in neue Techniken eingearbeitet hat und mir stets mit Rat und Tat zur Seite steht. Vielen Dank auch an Maria, für die liebevollen und anerkennenden Worte, die mir immer wieder Selbstvertrauen und Zuversicht gegeben haben.

Meinen Doktoranden-Kollegen Valerie, Katalin, Clara, Klement und Jonas, die ich im Institut kennengelernt habe, möchte ich ebenfalls danken. Durch die Gründung des „Milchkunde-Kegelclubs“ – auch wenn er leider nur einen sehr kurzen Aktionszeitraum hatte –, fühlte ich mich sofort herzlichst aufgenommen. Valerie, Dir möchte ich ganz besonders danken, für Deine herzenswarme Freundschaft, die mir in jeder Mittagspause – ob beim Spaziergang ums Schloss oder bei einem Kaffee im Büro – immer wieder Kraft und Motivation gespendet hat.

Mein größter Dank geht an meine Familie. Nur durch Eure jahrelange Unterstützung, die liebevollen, aufmunternden Worte und Euren Rückhalt war es mir möglich, diesen Weg zu gehen. Euer Vertrauen in mich hat mir dabei geholfen, meinen eigenen Weg zu finden und diesen mit Mut und voller Lebensfreude zu gehen. Vielen Dank!

Mein lieber Quirin, zu guter Letzt möchte ich Dir meinen tiefen Dank aussprechen. Du hast mich all die Jahre auf diesem Weg begleitet, warst mein Fels in der Brandung und mein Ruhepol. In allem, was ich tat, hast Du mich stets unterstützt und mir so die Möglichkeit gegeben, mich selbst zu verwirklichen. Danke!