

RNA virus detection and identification using techniques based on DNA hybridization.

Inauguraldissertation

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Ernst-Moritz-Arndt-Universität Greifswald

vorgelegt von

Ariel Viña Rodríguez

geboren am 29 September 1968

in Havanna, Cuba

Greifswald, 2017

Dekan: Prof. Dr. Werner Weitschies

1. Gutachter: Prof. Dr. Winfried Hinrichs

2. Gutachter: Prof. Dr. Reimar Johnen

Tag der Promotion:19. Apr. 2018

**THE MOSQUITO HYPOTHETICALLY CONSIDERED AS AN AGENT
IN THE TRANSMISSION OF YELLOW FEVER POISON,**

By Dr. CHARLES FINLAY,

Member of the Academy of Sciences of Havana, Member of the Clinical Society of
Brussels, etc., etc.

(An extract from the Annals of the Royal Academy of Sciences of Havana.)

Translated by RUDOLPH MATAS, M. D.

Mr. President and Gentlemen—Some years have now elapsed since I had the honor to communicate to you in this Academy

making this announcement here, I desire to justify, to a certain extent, this change in my opinions, by submitting to the appreciative criticism of my distinguished colleagues a new series of experimental studies, which I have undertaken, with the view of ascertaining the mode by which the yellow fever poison may be propagated.

I must state, however, that the subject of this paper has no relation whatsoever with the nature or form in which the morbigenous factor in yellow fever exists; in this regard I will limit my opinions to the following statement: that I admit the existence of a material transportable cause, which may be either an amorphous virus, an animal or vegetable germ, bacterium, etc., etc., but which consists in all cases of a tangible something, which has to be communicated from the sick to the healthy in order that the disease may be propagated. What I propose studying is the medium or agent by which the pathogenic material of yellow fever is carried from the bodies of the infected to be implanted in the bodies of the non infected.

Academy of Sciences of Havana, 1881

*This thesis and my work are dedicated to the Memory of
the noble Cuban Doctor and Scientist Juan Carlos Finlay.*

Contents

List of abbreviations and symbols	v
Summary	vi
Zusammenfassung.....	vii
Aim of the thesis.....	viii
List of publications.....	x
Chapter 1. Abstract	1
1.1. (Re)-emerging diseases	1
1.2. RNA virus detection.....	2
1.2.1. RNA extraction.....	3
1.2.2. Primers and probe design.....	4
1.3. WNV lineage 1 and 2 detection and quantification	6
1.4. Equine encephalitis viruses' detection and quantification.	7
1.5. Ngari virus detection	9
1.6. Detection of HEV in wild German animals and subtyping of HEV genotype 3.....	10
1.7. Flavivirus detection and identification.....	14
1.8. Conclusions	18
1.9. References.....	21
1.10. List of routinely used software.....	23
Chapter 2. Publication I	25
Two New Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction Assays with Unique Target Sites for the Specific and Sensitive Detection of Lineages 1 and 2 <i>West Nile Virus</i> Strains.....	25
Chapter 3. Publication II	33
A Quantitative Real-Time RT-PCR Assay for the Detection of Venezuelan Equine Encephalitis Virus Utilizing a Universal Alphavirus Control RNA.	33
Chapter 4. Publication III	41
Ngari Virus in Goats during Rift Valley Fever Outbreak, Mauritania, 2010.....	41
Chapter 5. Publication IV	45
Hepatitis E Virus Genotype 3 Diversity: Phylogenetic Analysis and Presence of Subtype 3b in Wild Boar in Europe.	45
Chapter 6. Publication V	71
Detection of Hepatitis E Virus in Archived Rabbit Serum Samples, Germany 1989.....	71
Chapter 7. Publication VI	75
A Novel Pan- <i>Flavivirus</i> Detection and Identification Assay, Based on RT-qPCR and Microarray.....	75
Authors contribution to publications	94
Curriculum vitae	97
Acknowledgments	100

List of abbreviations and symbols

AIDS	Acquired Immunodeficiency syndrome
CCHFV	Crimean Congo hemorrhagic fever virus
CG	Complete genome (sequence)
Cq	cycle used for quantification in qPCR
Ct	Throughput cycle, here used as synonym of Cq
DNA	Deoxyribonucleic acid
dsDNA	double strand DNA
EEEV	Eastern equine encephalitis virus
EEV	Equine encephalitis viruses – here: WEEV, EEEV and VEEV
GUI	Graphical user interface
HCV	Hepatitis C virus
HEV	Hepatitis E virus
HIV	Human immunodeficiency virus
JEV	Japanese encephalitis virus
MSA	Multiple sequence alignment
NN	Nearest-neighbor, here used to refer to NN method of calculation of ΔG of DNA hybridization
NS	Nonstructural region – genome region coding for a non-structural protein
PCR	Polymerase chain reaction
qPCR	Quantitative, kinetic or real time PCR
RNA	Ribonucleic acid
RT-PCR	Reverse transcription couplet with PCR
RT-qPCR	Reverse transcription couplet with qPCR
RVFV	Rift Valley fever virus
TBEV	Tick-borne encephalitis virus
UTR	untranslated region
VBA	Visual Basic for applications – a Microsoft programming language embedded in MS Office.
VEEV	Venezuelan equine encephalitis virus
WEEV	Western equine encephalitis virus
WHO	World Health Organization
WNV	West Nile virus
YFV	Yellow fever virus
ΔG	Gibbs free energy of reaction, here of forming a double strand DNA from the individual strands, defined as: $\Delta G = \Delta H - T \Delta S$, is a thermodynamic potential that can be used to calculate the maximum of reversible work that may be performed by a thermodynamic system at a constant temperature and pressure.
ΔH	enthalpy change
ΔS	entropy change
T_m	Melting temperature – at which half of the limiting DNA strand is in the hybridized state.
T_a	Annealing temperature – at which the hybridization experiment take place
SYBRG	SYBR Green I - dye which binds dsDNA, commonly used in implementations of qPCR with no probe.

Summary

Humanity is constantly confronted with the emergence and reemergence of infectious diseases. Many of them produce large or devastating epidemics, like AIDS (HIV) and Ebola. Others have been long neglected, yet pose immediate threats to global public health as evidences the abrupt emergence of Zika virus in South America and its association with microcephaly in babies. The examples illustrate, that many of these diseases are provoked by RNA viruses. One of the first steps in understanding and eliminating those threats is the development of sensitive and rapid diagnostic methods. A general and relatively rapid method is the direct detection and examination of the agent's genome. However, the nature of (re)emerging RNA viruses poses a series of very specific problems for the design of such methods. Therefore, a systematic approach was proposed for the design of DNA-hybridization-base methods to detect and characterize RNA viruses that will have both a high sensitivity and a specificity sufficiently broad to detect, per reaction, down to a single copy of any of the possible variants of the viral genome.

Following this approach a series of assays were designed, developed or adapted and put into use for detection and characterization of important RNA viruses. One of those viruses is West Nile virus (WNV), which after its explosive introduction into USA become the most widespread flavivirus throughout the world and, consequently, many countries began an intensive monitoring. While existing assay detected predominantly the Lineage 1, in Europa Lineage 2 was expected. Two new RT-qPCR for the detection of both lineages were developed, and reportedly used by independent laboratories. Due to more than 50000 associated deaths per year, the Hepatitis E virus also received an increasing attention to elucidate novel routes of transmission. This virus (especially genotype 3) has the zoonotic potential of transmission from pigs and wild boar to humans. RT-qPCR and nested qPCR for detection and characterization of this virus as well as a methodology for subtyping were developed and the first detected case of subtype 3b in a German wild animal was documented. In addition a novel assay for flaviviruses conformed by a RT-qPCR coupled with a low density DNA microarray was developed, which enabled the identification of WNV in mosquitoes from Greece. A RT-qPCR suitable for surveillance and diagnostic of all known variants of Venezuelan equine encephalitis virus was developed too. A causative agent of hemorrhagic infections, the Ngari virus, was detected and characterized in animal samples from Mauritania. These achievements were supported by the development of software applications for selection and visualization of primers and probes from aligned DNA sequences and for modeling of DNA hybridizations using unaligned sequences.

In conclusion a general methodology for rapid development of sensitive diagnostic methods based in DNA-hybridization technics (PCR, sequencing and microarray) was established and successful applications are reported.

Zusammenfassung

Die Menschheit wird ständig durch das Auftreten neuer und neuartiger Infektionskrankheiten wie HIV (AIDS) und Ebolafieber bedroht, die große und verheerende Epidemien auslösen. Andere wurden lange vernachlässigt, doch stellen sie eine unmittelbare Bedrohung für die öffentliche Gesundheit weltweit dar, wie sie durch das Auftreten des Zika-Virus in Südamerika und seine Assoziation von Neugeborenen mit Mikrozephalie belegt wurde. Viele dieser Krankheiten werden durch RNA-Viren verursacht. Daher ist ein erster Schritt zur Beseitigung solcher Bedrohungen die Entwicklung von sensitiven und schnellen diagnostischen Methoden, wobei in der Regel ein Teil des Genoms direkt detektiert wird. Da die Eigenschaften der RNA-Viren für das Design entsprechender Methoden sehr spezifischen Probleme darstellen, wurde ein systematischer Ansatz mit DNA-Hybridisierungs-basierten Methoden gewählt, mit dem hochsensitiv alle Varianten des viralen Genoms detektieren werden können.

Anhand dieses Ansatzes wurden Assays entwickelt oder angepasst und für den Nachweis und die Charakterisierung wichtiger RNA-Viren eingesetzt. Im Rahmen der Arbeit wurde eine RT-qPCR zum Nachweis des West-Nil-Virus (WNV) und seiner Linien 1 und 2 entwickelt. WNV ist weltweit verbreitet und kann beim Menschen zu schweren neurologischen Erkrankungen führen. Daneben wurden RT-qPCR Assays zur Diagnostik von Hepatitis-E-Viren (HEV) und aller Varianten der Venezolanische Pferde-Enzephalomyelitis (VEEV) entwickelt. Das HEV ist, mit weltweit mehr als 50000 Todesfällen im Jahr, der häufigste Erreger akuter Leberentzündungen. Das Virus, insbesondere der Genotyp 3, kommt auch in Deutschland vor, wobei als Hauptwirte Haus- und Wildschweine angesehen werden. Die Subtypisierung der HEV Genotypen 3 wurde aktualisiert und erstmals der Subtyp 3b in einem Wildtier in Europa dokumentiert. VEE ist eine tödlich verlaufende Viruskrankheit der Pferde, die auch auf den Menschen übertragen werden kann. Das Virus wird ebenfalls über Stechmücken übertragen und kommt bislang in Südamerika vor. Zudem konnte mit Hilfe einer RT-qPCR erstmals das Ngarivirus, ein hämorrhagisches Orthobunyavirus, in Ziegen in Mauretanien detektiert und charakterisiert werden. Schließlich wurde ein RT-qPCR Assay entwickelt, der - gekoppelt mit einem DNA-Microarray - Flaviviren spezifisch detektieren und erkennen kann. Mittels diese Assays konnten in Stechmückenproben aus Griechenland WNV der Linie 2 detektiert werden. Diese Ergebnisse wurden durch die Entwicklung von zwei Software-Tools unterstützt: Eines für Visualisierung von DNA-Sequenzalignments zur Selektion von Primern und Sonden und eines für die thermodynamische Modellierung von DNA-Hybridisierungen wobei keine Sequenzalignments notwendig sind.

Zusammenfassend konnte mit Hilfe dieser Methodik eine schnelle Entwicklung von sensitiven diagnostischen Methoden, die auf DNA-Hybridisierungstechniken (PCR, Sequenzierung und Microarray) basieren, etabliert und in einer Reihe von Anwendungen erfolgreich eingesetzt werden.

Aim of the thesis

The unifying goal was to demonstrate the need, convenience and affordability of constructing a systematic approach for the design of DNA-hybridization-based methods of detection and characterization of viruses that are both highly sensitive and sufficiently broad-specific to detect per reaction a minimal number of copies of any of the possible variants of the viral genome. The immediate objectives were to establish diagnostic methods for each of the selected RNA viruses or group of viruses.

An important common constrain was that in spite of the possible high number of known viral genomes sequences (typically with thousands of sequences in data banks) the information may appears incomplete due to biases regarding genomic variants (genotypes, subtypes or even species) and the completeness of most sequences. This concerns in particular the high number of short partial sequences from different, generally non-overlapping regions.

The objectives comprise the following tasks:

- the establishment of a quantitative real-time reverse transcriptase PCR (RT-qPCR) assay for the detection of West Nile virus (WNV) lineages 1 and 2 (Publication 1).
- the establishment of a RT-qPCR for detection of all known Venezuelan equine encephalitis virus (VEEV) variants (Publication 2).
- the establishment of a RT-qPCR for detection and characterization of Hepatitis E virus (HEV) with a revision of the subtype scheme of the genotype 3 (Publication 4).
- the establishment of a RT-qPCR coupled with a low density DNA microarray for detection and identification of flaviviruses (Publication 6).
- the screening of field samples: the detection of Ngari virus in Mauritania (Publication 3), HEV in German wild animals (Publications 4 and 5) and WNV in Greece (Publication 6).
- the development of software tools for: 1) selection and visualization of primers and probes from aligned partial DNA sequences, 2) simple modeling of DNA hybridization using unaligned DNA sequences and 3) automation of RNA extraction.

Notes on the conventions used in structuring and formatting:

This thesis is presented as a collection of published paper (Chapter 2 to Chapter 7) which show all the details of the methods and results, preceded by an introduction (Chapter 1) which expose the achievement of the unifying goal, and provide a broad abstract with some more context added.

Most of the complementary data and software can be acceded on-line through the portal of the author's projects: <http://qpcr4vir.github.io/> but some complementary tables and figures are additionally included here within the corresponding paper.

To avoid duplications of references already used in the included papers they are annotated in the text of the thesis abstract as a direct link in the form: (Chapter X-ref Y), where X is a link to the Chapter including the reference and -ref Y is a link to the reference listed in that Chapter. Figures and tables are also linked this way.

The final version of the publications, or of the author's copy in PDF format were imported into Microsoft Word, and manually edited to eliminate minor formatting differences with the PDF version. Still, some minor differences are present.

The ICTV's [recommendation](#) for how to write a virus name was followed through the abstract (which may introduces some nuances when the name of a virus coincide with the name of the species or genus).

List of publications

Parts of this work are available in the following publications:

1. *Two new real-time quantitative reverse transcription polymerase chain reaction assays with unique target sites for the specific and sensitive detection of lineages 1 and 2 West Nile virus strains.* Eiden M, Vina-Rodriguez A, Hoffmann B, Ziegler U, Groschup MH. J Vet Diagn Invest 2010, 22:748-53. [PubMed](#)
2. *A Quantitative Real-Time RT-PCR Assay for the Detection of Venezuelan equine encephalitis virus Utilizing a Universal Alphavirus Control RNA.* Ariel Vina-Rodriguez, Martin Eiden, Markus Keller, Winfried Hinrichs and Martin H. Groschup. Biomed Res Int. 2016; 2016: 8543204. [PubMed](#).
3. *Ngari virus in goats during Rift Valley fever outbreak, Mauritania, 2010.* Eiden M, Vina-Rodriguez A, El Mamy BO, Isselmou K, Ziegler U, Höper D, Jäckel S, Balkema-Buschmann A, Unger H, Doumbia B, Groschup MH. Emerging infectious diseases. 2014; 20(12):2174-6. [PubMed](#)
4. *Hepatitis E virus genotype 3 diversity: phylogenetic analysis and presence of subtype 3b in wild boar in Europe.* Vina-Rodriguez A, Schlosser J, Becher D, Kaden V, Groschup MH, Eiden M. Viruses 2015 ,7:2704-26. [PubMed](#).
5. *Detection of Hepatitis E Virus in Archived Rabbit Serum Samples.* Eiden M, Vina-Rodriguez A, Schlosser J, Schirrmeier H, Groschup MH. Germany 1989. Food Environ Virol 2016 ,8:105-7. [PubMed](#).
6. *A novel pan-flavivirus detection and identification assay, based on RT-qPCR and microarray.* Vina-Rodriguez A, Sachse K, Ziegler U, Chaintoutis SC, Keller M, Groschup MH and Eiden M. Biomed Res Int. Volume 2017 (2017), Article ID 4248756, 12 pages. [PubMed](#).

Further related publications (not part of this thesis):

7. *Hepatitis E virus in wild rabbits and European brown hares in Germany.* F. Hammerschmidt, K. Schwaiger, L. Dähnert, A. Vina-Rodriguez, D. Höper, M. Gareis, M. H. Groschup, M. Eiden. Zoonoses Public Health. 2017;00:1–11. <https://doi.org/10.1111/zph.12355>
8. *Serologic and Molecular Survey of Hepatitis E Virus in German Deer Populations.* Neumann S, Hackl SS, Piepenschneider M, Vina-Rodriguez A, Dremsek P, Ulrich RG, et al. J Wildl Dis 2016, 52:106-13. PubMed.
9. *Chronically infected wild boar can transmit genotype 3 hepatitis E virus to domestic pigs.* Schlosser J, Vina-Rodriguez A, Fast C, Groschup MH, Eiden M. Vet Microbiol 2015 ,180:15-21. PubMed.
10. *Natural and experimental hepatitis E virus genotype 3-infection in European wild boar is transmissible to domestic pigs.* Schlosser J, Eiden M, Vina-Rodriguez A, Fast C, Dremsek P, Lange E, et al. Vet Res 2014 ,45:121. PubMed
11. *Molecular and serological studies on the Rift Valley fever outbreak in Mauritania in 2010.* Jäckel S, Eiden M, El Mamy BO, Isselmou K, Vina-Rodriguez A, Doumbia B, Groschup MH. Transboundary and emerging diseases. 2013; 60 Suppl 2:31-9. PubMed
12. *Pathogenesis of West Nile virus lineage 1 and 2 in experimentally infected large falcons.* Ziegler U, Angenvoort J, Fischer D, Fast C, Eiden M, Rodriguez AV, Revilla-Fernández S, Nowotny N, de la Fuente JG, Lierz M, Groschup MH. Veterinary microbiology. 2013; 161(3-4):263-73. PubMed
13. *The usefulness of Umelosa hepatitis C virus qualitative kit as supplemental test for confirmation of hepatitis C virus infection.* Gonzalez-Perez I, Gonzalez Gonzalez YJ, Vina-Rodriguez A, Armas CA, Solis RL. Rev Soc Bras Med Trop 2004; 37(1):25-27. PubMed
14. *Validation of a nested PCR assay UMELOSA HCV CUALITATIVO for the detection of Hepatitis C virus.* Gonzalez-Perez I, Gonzalez Gonzalez YJ, Armas CA, Vina-Rodriguez A, Medina CA, Trujillo PN, Perez Guevara MT, Lydia SR. Biologicals 2003; 31(1):55-61. PubMed.
15. *Design of an antisense reverse-transcriptase-polymerase chain reaction primer efficient for all hepatitis C virus genotypes: comparison of its performance vs a commercial primer.* Gonzalez-Perez I, Vina-Rodriguez A, Cayarga AA, Rosa IG, Gonzalez YJ. Anal Biochem 2003; 315(2):281-4. PubMed
16. *The Single Strand Conformational Polymorphism (SSCP) in HCV characterization.* Ariel Viña, Idania González, Odalys García y Juan Morales Grillo. Revista CENIC Ciencias Biológicas, Vol. 31, No. 3, p. 163-7, 2000.

Published viral sequences (GenBank acc. n. CG-complete genome, CS-complete segment):

HEV - [KY436898](#)(CG), [KY436899](#), [KP294371](#)(CG), [KR261083](#), JQ807471-476, JQ807477-524, JQ965993, HE797853-864, KY092451-KY092478;

Ngari virus - [KJ716848](#)(CS), KJ716849(CS), KJ716850(CS);

HCV - [HJ919811](#)-815, [EU257483](#);

RVFV - [JQ974833](#)

Chapter 1. Abstract

1.1. (Re)-emerging diseases

Naturally, infection diseases may cause sporadic appearance of epidemics into unprepared communities. This is usually perceived as emerging or reemerging diseases. Globalizing human activities have the unnatural but logical consequence of sporadic but recurrent dissemination of such diseases. In the past we saw large epidemics moving from Asia to Europe, and others from Europe to America, where they killed large portions of the native population. Recently, a very fast spreading of West Nile virus into USA caused millions of infections in humans and massive dead of wild birds. Not only direct transportation, but also many others factors, including climate change, may boost this undesired effect by allowing new routes of contact of the infection agent, the final host, and possible vectors and reservoirs.

Some of these infections produce large or devastating epidemics, like AIDS (HIV) and Ebola. Others produce intriguing symptoms which trigger world large effort for immediate elucidation, like recently the microcephaly in babies provoked by Zika virus, a largely neglected infection.

As early as 1881 the Cuban Dr. Carlos J Finlay (Chapter 1-ref [1]–[3]) showed the role of mosquitoes as biological vectors for infectious diseases. Through a series of sharp observations and carefully planned experiments he demonstrated that the mosquito *Aedes aegypti* is the agent responsible for the transmission of one of the historically most devastating, and even today, recurrently (re)emerging infectious disease: the Yellow Fever. This infection, together with malaria (also mosquito transmitted) caused more than 20 000 human deaths during the construction of the Canal of Panama - a fact being on the basis of the failure of the French to construct the canal, which was achieved by USA only after eradication of the mosquito following the methods used in a similar and firstly successful campaign organized in Havana under the recommendations of Finlay. Thus, monitoring and controlling those vectors, in particular mosquitos and ticks, is crucial to prevent the spread of many of the most important (re)emerging diseases.

Alongside wild animals, economically important animals are also affected, often leading to the sacrifice of large number of them. Unfortunately some of these agents infect humans too. This zoonotic nature significantly complicates the situation and rises some of the requirements in fighting them in animals to near the same level than in humans (part of the concept of the [One Health Initiative](#)).

Viruses have been proposed as bio weapons (Chapter 1-ref [4][5]), and the possibility of their use by terrorists has added huge urgency to the necessity of been prepared to fight the consequences, firstly by detecting them promptly.

As the examples show many of these diseases are provoked by RNA viruses. One of the first steps in the understanding and elimination of such threats is the development of sensitive and rapid

diagnostic methods. The nature of (re)emerging RNA viruses poses a series of very specific problems for the design of such methods. For example, during the initial diagnostic, and also during the follow-up in patients after treatment it may be required to detect the virus down to the absolute limit of a few copies per sample and also to determine the exact viral (geno)type. In a much broader case we could need to screen large numbers of samples for the presence of some of the variants of the viral genome. In the broadest extreme of requirements, surveillance programs need to detect any viral variant, ideally even unknown ones.

Classical virological methods may not be appropriate in many situations due to insufficient sensitivity, speed or simply because of their absence. PCR based methods, DNA microarray or sequencing among other are often the best choices in such situations. Therefore a systematic approach need to be developed for the design of DNA-hybridization-base methods of detection and characterization of RNA viruses that are both sufficiently broad-specific and high sensitive to detect per reaction a few copies of any of the possible variants of the viral genome.

1.2. RNA virus detection

One characteristic of RNA viruses that affect the diagnostic method is the high variability of the genome. For some viruses the high intra host viral genome variability led to the use of the concept of viral quasispecies. At the level of what is defined as a viral species, the percent of nucleotide substitutions may be in the range of 10-20 % or even more. This variability is by far the major problem during the design of DNA-hybridization mediated diagnostic methods. The problem of choosing a DNA fragment that will hybridize to most (ideally all) variants has no trivial solution.

Tightly linked to variability is the problem of the available amount of sequence information. For many viruses, there are simply too few sequences available. For others, in spite of the existence of large number of sequences, there is a bias in the available amount of sequences for different variants. Typically a vast majority of the sequences could represent a large and important recent outbreak, while others variants are significantly underrepresented. The amount of sequences is not the only limitation, but the quality (of that amount) as well, notably, the length and the genomic region covered by the sequences. This bias may be reinforced by a positive (albeit undesirable) feedback: detection systems designed using majoritarian sequences may detect preferably (or even exclusively) those sequences and contribute to the further absence of sequences of other variants.

The species of most RNA viruses are further subdivided into subgroups called lineages, types, genotypes, subtypes, etc. They may have different pathological, epidemiological, or some other observable biological characteristics, and may ideally serve to break off the original variety into groups for each of which we have to ensure the applicability of the method. They may be also used to set priorities in case of necessity. For example, we may choose to be sure that the more aggressive or lethal

variants are detected, or we may choose to detect just the variants expected to be present in some geographical region or during a concrete outbreak. Nevertheless, we need to keep in mind that those biological classifications may not reflect the technical characteristic that we are trying to use: to group the sequences into groups for which we can guarantee detection with the newly designed method. One special case will be a virus (or some subgroup) with variants distributed in a star-like pattern, with no clear “clades”, but with a “continuous” evolution into many apparently individual variants. Even in those cases we may need to define (and probably some-how name) subgroups that may seem arbitrary for other purposes.

In this thesis are presented the results of the development and partial use of new protocols for the detection and/or identification of WNV and Flavivirus in general, HEV, VEEV and Ngari virus.

1.2.1. RNA extraction

Prior to the application of the DNA-hybridization-based technique, like RT-qPCR, the viral RNA needs to be extracted. We used well established methods and commercially available kits based on columns (RNeasy Mini Kit or QIAamp Viral RNA Mini Kit, QIAGEN, Hilden, Germany) or magnetized particles (NucleoMag® VET kit from MACHEREY-NAGEL, Durel, Germany) to achieve the separation, either automatically, using pipetting robots, or manually.

Especially useful was the combination of magnetized particles with a Freedom EVO universal pipetting robot from TECAN, Mannerdorf, Switzerland. The provided Freedom EVOware software was used to write simple and specific pipetting protocols scripts in a semi visual way. Unfortunately, writing more complex or flexible protocols to accommodate arbitrary number of samples or minor modifications of the protocols was very time consuming and error prone. To overcome these limitations a new software was written in Python, “RobotEvo”, which automatically generates the scripts for the robot. RobotEvo allows a more direct programming of the steps needed in a typical biochemical/biological pipetting protocol like RNA extraction. The implementation adds: a parser and a generator of *instructions* for the script; “modes” of output: as human readable comments, EVOware scripts, etc., in module *EvoMode*; a model of the *robot* to detect possible errors prior to the generation of the script; simple pipetting instructions; a set of *protocol steps*; and, finally, facilities to declare the *reactive* and the *labwares*.

For the protocol for RNA extraction (module *RNAextractionMN_Mag_Vet*) the labwares and reactive are declared first. Immediately an automatically generated check-list is presented in a graphic user interface (*GUI*). After a possible adjustment of the parameters by the operator the program goes through the deduced protocol steps of spreading buffers, mixing, washing, incubating, etc. generating a detailed script to be imported and visualized in the TECAN Freedom EVOware software. The obtained script is very long but structurally very simple and well commented, which facilitates the visual control of each instruction.

1.2.2. Primers and probe design

Although, there are well known software for design of primers and probes or for modeling of DNA hybridizations, we found no simple way to adapt them to the specifics of the RNA virus characteristics influencing the design of primers and probes for PCR, or of probes for a low-density microarray, for virus identification. Even when the goal is to design a high specific (RT-q) PCR the target sequence is always a group of similar but still different sequences, with up to 10 or 20 % of punctual nucleotide differences affecting most positions. Tools that find a set of specifics and compatible primers from only one given sequence are hardly useful. The task becomes even more complicated when trying to design an assay that detect a whole viral group (for example - a given genus), especially if we need to control the detection efficiency for each of the subgroups (for example - species) which form the broader group. The mere possibility of breaking the target group into a classification tree of variable deep is not implemented in most software, and manual workarounds significantly complicate the task.

Further complications are introduced when some incomplete sequences that do not cover the entire target region, are ones of the few representatives of some of the subgroups and therefore can't be ignored. Additionally, to represent all this sequence diversity the number of sequences selected for analysis maybe well in the order of a few hundreds. We did not found any software that simultaneously copes with all these requirements. A solution widely used is to simply print the alignment and manually scan for possible candidates deducing partial consensus. But with the growth of the number of sequences into the hundreds this approach become impracticable and it is always time consuming and error prone. We partially automated this solution using an Excel workbook which is publicly available at [GibHub as VisualOligoDeg](#).

VisualOligoDeg facilitate the visual selection of candidate primer or probes from an existing MSA, by interactively constructing consensus with selected degree of degeneracy, modeling some characteristic (T_m , ΔG , etc.) of the possible hybridization and easily grouping and filtering of sequences. Some implementation details are: its use and installation do not depend on any software other than MS Excel; the truncated and reclassified aligned sequences can be re-exported to a new text/fasta alignment file; the used NN parameters can be modified; part of the workbook cells are regenerated during the import of a fasta file, potentially correcting errors inadvertently introduced by the user; part of the functionality is programmatically expressed in VBA and includes a set of functions to import and export all the code from the workbook, allowing us to track the code with git (a distributed version control system of software source code, <https://git-scm.com/>).

After selection of primers or probe candidates and before or after their validation with other tools, it may be desirable to predict their hybridization characteristics onto a set of additional target sequences originally not present in the MSA used during the design. A question arises: Is it imperative to add the sequence to the MSA prior to test it against a set of probes? This problematic situation also appeared when we tried to predict the result of the hybridization into our low-density microarray of

the amplicon from a viral strain with a known sequence. We decided to use a different model to avoid building or updating the MSA and developed a second software tool: *ThDy_DNAHybrid*.

Most of the algorithms used for sequence alignment, and often reused for the design of primers and probe, originate in the phylogenetic analysis of sequences. There, each position is thought to carry some amount of phylogenetic information, thus, sequences are treated as a text, nucleotide as letters and differences as letter substitutions. But, as primary concept during probe design, we may prefer to refer to the percent of probes or target in hybridized state, which is what determine the efficiency of a PCR or of a microarray detection. The percent of probes hybridized is described with the equilibrium constant K of the reaction, thermodynamically (Chapter 1-ref [6]) related to the reaction $\Delta G = -RT \ln K$, at a given temperature T , where the gas constant $R = 1.9872 \text{ cal/mol K}$. SantaLucia, J. (Chapter 1-ref [6]) provides a “rule of thumb” to illustrate this relation: “every -1.4 kcal/mol in ΔG results in a change in the equilibrium constant by a multiplicative factor of 10”, and “ -4.2 kcal/mol ($= -1.4 \times 3$) equals a K change by 1000”.

Thus, rather than using the phylogenetic or plain text approach we may prefer a chemical or thermodynamic approximation to describe those reactions. Beginning with a rough approximation we may add more and more factors making our predictions more accurate. By using the text approach this incremental approximation is translated into more rules dictated by “experts”, which may be confusing, especially when quantitatively it may be unclear in each situation what rule is predominant. The thermodynamic approach will instead incorporate the new factors into the model itself, into the calculation, potentially making the final interpretation or use of the prediction simpler.

ThDy_DNAHybrid uses the thermodynamic NN method to describe DNA hybridizations. ΔG is calculated through the ΔH and ΔS accumulated by adjacent pair of dinucleotides. Thus, the relative position of the interacting nucleotides at both DNA strand is used, which send us back to an alignment, but now of only two sequences. Instead of trying every possible combination of relative position or annealing, we search for the most stable. This problem was solved by Kadelary (Chapter 1-ref [7]) using the dynamic-programming approach of the Smith–Waterman alignment algorithm modified by changing the alignment objective function with ΔG or T_m based on ΔH and ΔS , introducing the dependencies into pair of adjacent di-nucleotides following the NN model. We have further adapted the algorithm to extract any position at which the stability of the hybridization pass a user defined value: any position in which a probe could produce a measurable signal. The modification can also track all the target sequences that have at least that level of interaction with the given probe.

Additional details of the implementation are: it is written in c++11/14; includes a GUI built with the *Nana C++ GUI library* – a new, simple and modern way to do GUI in c++ for Windows and Linux; the parameters of each runs are saved in a “project” file which can be re-run and/or manually inspected and edited as simple text; a sub-library, *ProgParam*, was developed to managed the parameters and subprograms that form the software and transparently join the core functionality with the user

interface; the sub-library *Unit* make easy the conversion between any physical or chemical units, helping to avoid errors and offering great flexibility for the final user; and results can be saved in a set of text files and presented as interactive tables.

1.3. WNV lineage 1 and 2 detection and quantification

West Nile virus (WNV; family *Flaviviridae*, genus *Flavivirus*) is an arthropod-borne virus grouped in the *Japanese encephalitis virus* (JEV) serocomplex, which includes the *St. Louis encephalitis virus*, JEV and *Murray Valley encephalitis virus* (MVEV), among others (Chapter 2-ref 11). The genome is a linear, single-stranded, plus-sense RNA, which encodes for three structural proteins (C, prM, and E) that form the enveloped virion, and seven nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) proteins (Chapter 2-ref 4).

WNV is transmitted by mosquitoes belonging principally to the *Culex pipiens* complex. A large variety of wild bird species are the natural reservoir. The host range is very broad and encompasses not only humans but also equids, alligators, dogs, sheep, and many other species (Chapter 1-ref [8], Chapter 2-ref 4, 19). Human infections are generally asymptomatic, but near 20% are characterized by flu-like illnesses: headache, high fever, chills, arthralgia, malaise, and retro-orbital pain. Up to 1% of infected human beings develop severe encephalitis, myelitis, and/or meningitis, and of these cases, 1 in 20 dies (Chapter 2-ref 20). There is no human vaccine yet (Chapter 1-ref [8]) and the best method to reduce cases of infections is mosquito control.

WNV has been grouped into at least five lineages. Lineage 1 has a wide distribution and is found in some parts of southern Europe, Africa, the Middle East, Asia, and the Americas. The Kunjin virus, which circulates in Australia, represents the subtype 1b of lineage 1. Lineage 2 strains are found in sub-Saharan Africa and Madagascar where it is endemic, and have also been discovered in Europe: in Hungary (first isolated from a goshawk), Italy (in horses, in the Island Sardinia, 2011 –lineage 1, and lineage 2 in mosquitoes in 2012), Greece, Russia, Rumania and Austria (Chapter 1-ref [8], Chapter 2-ref 2, 3, 5, 23).

The introduction and rapid spread of WNV into USA in 1999 (where it later became endemic and then spread to South America) (Chapter 1-ref [9]) is a dramatic example of the impact of human activity (Chapter 1-ref [10]) in general, and of globalization in particular, on virus distribution. European countries were indeed warned about the possibility of such expansion in their territory too and began to implement large monitoring or surveillance programs. For such countries it was expected than not only the lineage 1 observed in USA may appear, but also the lineage 2. Unfortunately most used RT-qPCR were highly specific for lineage 1, and failed to detect the lineage 2. This situation was reflected in the conclusions of the first international proficiency study on WNV molecular detection published in 2006: “Most participants showed good performance in detecting lineage 1 WNV, the predominant virus

in the Northern Hemisphere. The inability of some laboratories to detect even highly concentrated lineage 2 WNV downgraded the overall outcome”, because “only 27% of participants were able to detect the 6 samples containing $1,8 \times 10^4$ copies/mL” of the lineage 2 strain (Ug37) (Chapter 1-ref [11]).

We established two new RT-qPCR (Chapter 2). Primers and probe of assay 1 target the 5'-UTR, and assay 2 targets the nonstructural region NS2A. This enables an unambiguous and independent WNV diagnosis based on 2 different amplicons. Each assay was designed to guarantee the detection of both lineages. Both assays allow the detection of as few as 2–4 genome copies of WNV strains per reaction (which typically count for 100 to 200 copies/mL of sample, depending on the RNA extraction protocol used). They can be used independently or in combination to improve sensibility and specificity, or to exclude cases of carry-over contamination (the amplicons generated are not mutually amplifiable).

A synthetic RNA corresponding to the 5'-UTR amplicon (Figure 1) was designed containing 6 twist inverted GC base-pair changes at the internal sequence in a way that can be unambiguously recognized by a specially designed probe. This RNA was used as positive control and as external calibrator for quantification.

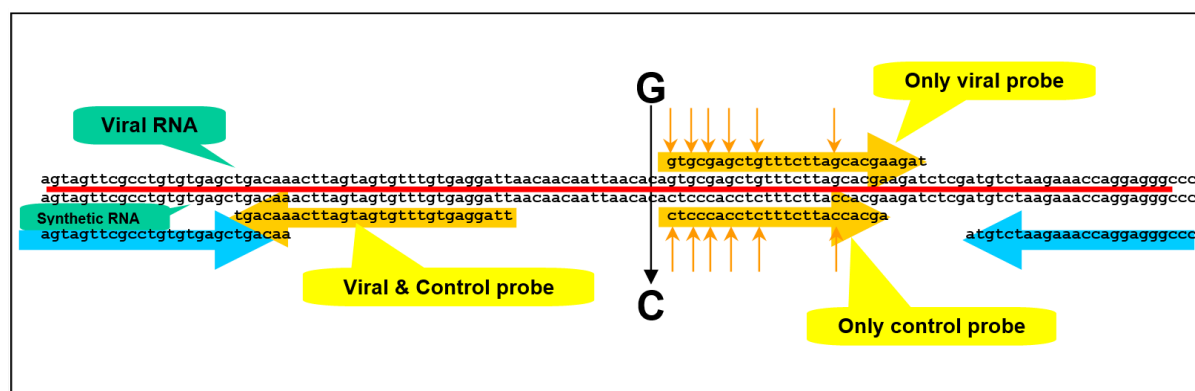


Figure 1. Synthetic WNV RNA control with target regions for both PCR primer, a probe for quantification of the virus and a special region containing 6 twist inverted GC base-pair changes for exclusive detection of this RNA control (but not the virus)

WNV have not been detected in German samples yet, but we were able to detect WNV lineage 2 in mosquito pools from Greece (Chapter 7). The assays have proved to be useful in the practice, and continue to be used not only in our laboratory, but also in many other independent laboratories. The assays have been also further validated by successfully participating in at least four international Ring Test (Chapter 7).

1.4. Equine encephalitis viruses' detection and quantification.

Western equine encephalitis virus (WEEV), Eastern equine encephalitis virus (EEEV), and Venezuelan equine encephalitis virus (VEEV) are arthropod-borne (arbo) viruses of the genus *Alphavirus* of the virus family *Togaviridae*. The single-stranded positive sense RNA genome of VEEVs contains approximately 11,400 bp, and encodes four non-structural proteins (nsP1-4) at the 5'-end and five structural proteins (C, E3, E2, 6K and E1) at the 3'-end (Chapter 1-ref [12]).

These viruses are transmitted by mosquitoes within bird (WEEV, EEEV, and epizootic VEEV (epizootic strains)) or rodent populations (VEEV, enzootic strains). Infections in reservoir hosts do not lead to obvious clinical signs. However, severe diseases can occur when equines and humans are infected with epizootic subtypes. EEEV are classified as **Category B** agent by the Centers for Disease Control and Prevention, Atlanta and are considered potential bioterror weapons (Chapter 1-ref [13]). VEEV can also be transmitted by aerosol exposure (Chapter 3-ref [7][8][9]) and have been weaponized (Chapter 1-ref [4][5]). Intensive research is conducted to obtain a reliable vaccine and an antiviral treatment (Chapter 1-ref [14]).

The three viruses – WEEV, EEEV and VEEV - affect humans, equines and mosquitoes mainly in the Americas. The subtypes (I - VI) of VEEV are also subdivided in varieties or serotypes (e.g. IA/B, IC, IIIA, etc.) (see Table 1. in (Chapter 1-ref [15])). In order to protect the unexposed population in Europe, imported animals are routinely tested for these agents, with the goal of detecting any known variant. While at the moment of the beginning of our work there were sensitive RT-qPCR for the first two (Chapter 3-ref [10]), for VEEV there was a need for a convenient RT-qPCR for the detection of all VEEV variants.

Taking advantage of the few new partial sequences available in public databases (in 2010) we adapted an existing conventional broad range RT-PCR (Chapter 3-ref [14]) to develop an RT-qPCR specific for VEEV (Chapter 3). A total of 33 VEEV sequences were retrieved from the GenBank database. The published broad-range primers, which target the nsP1 region of Alphaviruses, were modified, degenerated bases were inserted (TCCATGCTAATGCYAGAGCGTTTCGCA and TGGCGCACTTCCAATGTCHAGGAT) and a labelled probe (TGATCGARACGGAGGTRGAMCCATCC) was designed to specifically target VEEV sequences (Chapter 3-Table 1). The resulting primer/probe set enabled the application of a quantitative real-time RT-PCR protocol.

A unified synthetic RNA, with the target sequences of the three PCR, was introduced in the assays. A calibrated, external standard curve of this synthetic RNA is used for quantification and as positive RNA control within each of the three assays. The synthetic calibrator comprises a T7 RNA polymerase promoter and the target sequences for the RT-qPCRs of EEEV, WEEV, and VEEV (Chapter 3-Figure 1(a)). The synthetic RNA was obtained by in-vitro transcription. The EEEV and WEEV sequence regions include the targets for the primer and probes adopted unmodified from the literature, but the corresponding probe target sequences were placed on the complementary strand in order to generate a unique (different) amplicon sequence, discriminable from the original virus sequence yet maintaining the same nucleotide composition. In addition, within the VEEV target region the original virus sequence 5'-CTGGCTTCAAAAC-3' was changed to 5'-CTCCGTTCAATAC-3' in order to discriminate unambiguously the synthetic RNA from viral RNA with the special control probe to exclude false positive signals in samples.

Four external standard curves (one for each virus and one for the VEEV-synthetic RNA control probe) were used for quantification. The limit of detection (LOD) in copies/ μ L of RNA corresponded at

least to 70 for VEEV (TC-83 strain), 700 for WEEV (McMillan strain), and 300 for EEEV (New Jersey strain) (Chapter 3-Table 2).

Later the project “Complete genome sequencing of a collection of the Venezuelan equine encephalitis viruses (VEEV)” (Grant ID HHSN272200900007C, National Institutes of Allergy and Infectious Diseases), Accession: PRJNA157941, <https://www.ncbi.nlm.nih.gov/bioproject/157941> yielded 104 new complete genomes. From them, using *ThDy_DNAHybrid* and *VisualOligoDeg*, we extracted and then synthesized 15 sequences, representing the target of our VEEV assay, including all the sequence diversity known by 2014. We were able to experimentally demonstrate that our assay effectively detect all of them (Chapter 3- FIGURE 3).

1.5. Ngari virus detection

Ngari virus is an arthropod-borne virus *recently classified* (Chapter 1-ref [16]) in order *Bunyavirales*, family *Peribunyaviridae*, genus *Orthobunyavirus*, as a member of the species *Bunyamwera orthobunyavirus* (which among others includes the Bunyamwera and Batai viruses).

The genome of members of the *Bunyamwera orthobunyavirus* species comprises 3 segments of mostly negative-sense single stranded RNA: the small (S), medium (M), and large (L) segments, which encode the nucleocapsid (N) protein, the 2 glycoproteins Gn and Gc, and the RNA-dependent RNA-polymerase, respectively. Ngari virus is a reassortant, with segments S and L from Bunyamwera virus and M from Batai virus. Ngari is more virulent than Bunyamwera and Batai and is associated with hemorrhagic fever. It was first isolated from *Aedes simpsoni* mosquitoes in 1979 and from humans in 1993, both in Senegal (Chapter 4-ref 3). During 1997 and 1998, humans were affected with hemorrhagic fever diseases in Kenya and Somalia that were caused by RVFV and by Ngari virus (Chapter 4-ref 2, r4).

Lambert et al. (Chapter 4-ref 6) constructed a series of MSA from the S segment of various viral serogroups and selected promising conserved regions for the design of PCR. The primers were then *selected manually* from those alignments. In essence they created a complex and optimized conventional multiplex-PCR assay with primer pairs for each of the following targets: Family: *Peribunyaviridae*, Genus: *Orthobunyavirus*, species: *California encephalitis/Bwamba* group, *Bunyamwera virus* serogroup, *Wyeomyia* and *Oropouche*; Family: *Nairoviridae*, Genus: *Orthonairovirus - Nairobi group* and Family: *Phenuiviridae*, Genus: *Phlebovirus*.

The RVFV appeared again in Mauritania in 2010 around the Adrar region and we searched for its RNA in serum collected there from various animal species (Chapter 1-ref [17]). RVFV is also a member of the Order *Bunyavirales* but of the Family *Phenuiviridae*, Genus *Phlebovirus*. Since we were committed to find other arboviruses from that Order, we adapted the PCR of Lamberg et al. to be used in separated SYBRG-RT-qPCR protocols. This screening was successful (Chapter 4) and we found goat samples apparently positive to *Bunyamwera virus* with an amplicon's $T_m \approx 78.2^\circ\text{C}$ slightly different from that of our *Batai virus* control (78.0°C). After direct sequencing of the amplicon these samples

appeared to contain *Ngari virus* (Figure 2). Further investigations of the samples, through infection of cell cultures and whole genome sequencing, unambiguously confirmed this result, as shows the phylogenetic tree constructed for the three whole genome segments (Chapter 4-Figure.).

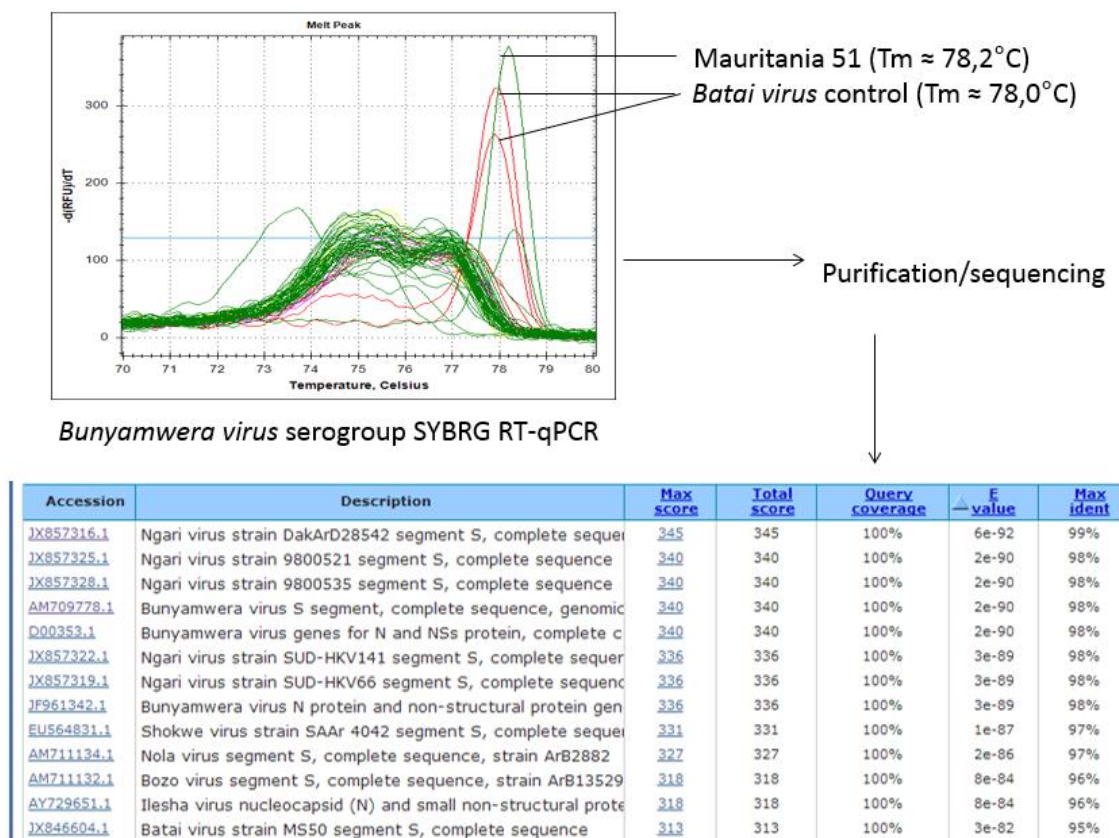


Figure 2. First experimental evidence of the presence of *Ngari virus* in goat samples from the RVFV outbreak in 2010, Mauritania. The melting curve analysis coupled to the RT-qPCR for the *Bunyamwera virus* serogroup adapted from Lambert et al. yielded a signal very similar to *Batai virus*. A simple BLAST-NCBI search of the amplicon sequence pointed to an *Ngari virus*. This was later confirmed by infecting cells cultures and whole genome sequencing.

1.6. Detection of HEV in wild German animals and subtyping of HEV genotype 3.

The Hepatitis E virus is now classified in the family *Hepeviridae*, where four genus are recognized. The genus *Orthohepevirus* contains the species commonly affecting humans. Specifically, the specie *Orthohepavirus A* is proposed to be further divided into genotypes, of which, the four genotypes 1, 2, 3 and 4 are known to infect humans (Chapter 5-ref 68).

The virion is approximately 27–34 nm in diameter and most likely icosahedral. The genome is a positive sense single-stranded RNA of approximately 7.2 kb, which contains a short 5' untranslated region (UTR), a short 3' UTR and three open reading frames (ORF1, ORF2 and ORF3) (Chapter 5-ref 27). The ORF1 encodes non-structural proteins carrying domains with methyl transferase, helicase and replicase activities (Chapter 5-ref 28). The ORF2 codes for the viral capsid protein of about 660 amino acids. The ORF3 is almost completely overlapped by the ORF2 (thus being the more conserved region)

and codes for a small phosphoprotein of about 114 amino acids, which is putatively responsible for the virion egress from infected cells (Chapter 5-ref 29).

Until recently HEV was seen mainly as an endemic causative agent of acute hepatitis in developing countries with poor hygienic conditions, and primarily transmitted via contaminated drinking water. Sporadic cases in industrialized countries were thought to be introduced by travelers from endemic regions. The detection in pigs and some other animals of HEV strains similar to that found in humans triggered investigations about the possible zoonotic and autochthonous origin of infections in humans. Moreover, the consumption of undercooked meat products was found to pose a substantial risk for HEV infection (Chapter 5-ref 4-7). Today it has been established that HEV genotype 3 and 4 have a main reservoir in pigs and wild boars and affects other animals, notably rabbits and deer. In particular, studies revealed, that HEV is ubiquitous in domestic pigs and wild boars throughout Europe (Chapter 5-ref 8). HEV infection in farmed pigs affects up to 80%–100% of the animals worldwide and usually occurs at the age of 2–4 months (Chapter 5-ref 26). While genotype 4 is predominantly present in Asia (Japan, China, India, etc.), genotype 3 appears to be distributed all around the world and also accumulates more diversity. Currently, genotype 3 is recognized as a food-borne zoonosis in developed countries where it usually causes a mild self-limited acute hepatitis (Chapter 1-ref [18]). From 2001 to 2017, 7056 human HEV cases (Chapter 5-ref 38) were reported in Germany (671 in 2014, 1265 in 2015, 1993 in 2016 and 1230 in 2017 until July), which includes a growing number of non-travel associated autochthonous cases.

The role of different HEV-3 genetic variants in the evolution of the disease (Chapter 5-ref 1,7,39,41,42), the possibility of tracking the routes of infection and the influences of human activity (Chapter 5-ref 43-46) are currently under study. To better understand the interplay of these factors, the prevalence of past and current infections has been studied in large collections of samples. A prerequisite for this approach is to optimize the performance of the detection techniques, which is hampered by the high variability of the viral genome, even intra genotypes. We therefore established (Chapter 5) an optimized version of a widely used diagnostic RT-qPCR (Chapter 5-ref 69) that targets the ORF3 of all four genotypes 1-4. With this modified assay the determination of the HEV RNA concentration is carried out using an external standard curve (a synthetic RNA). This calibrator encompasses the 81 bp sequence of the amplicon and includes the T7 promoter sequence at the 5'-end for in vitro transcription.

The high sequence variability potentially makes possible to collect sufficient information to study the routes of transmission. Due to the low viremia levels, high genome variability and low quality of some samples (serum and blood samples are less appropriate than tissue or fecal samples) the technique of choice to obtain most of the sequences have been nested RT-PCR that target relatively short sequences from different genomic regions. Consequently, to compare our viral isolates from German wild animals (Chapter 5, Chapter 6) with other European reports we developed a set of four new nested PCR for genotype 3 (Chapter 5) and applied them to each of our positive samples. This

fragmented nature of the sequence information hinder a strict phylogenetic analysis with direct sequence comparison (as shown by Bayer phylogenetic analysis in Chapter 5-ref 43, 45) and we choose to implement an alternative method. The aim is to obtain the finest possible classification of the sequences into subtypes, which will partially reconstruct the information we cannot obtain due to the impossibility of a direct sequence comparison. We choose the subtyping scheme of Lu *et al.* (Chapter 5-ref 32), which have been commonly used and supported by epidemiological and statistical analysis (Chapter 5-ref 4,43,45,47,48) but which have been recently criticized (Chapter 5-ref 1,30).

Most of the problems with this subtyping scheme are related to the difficulty to find well defined clades due to the apparent continuous or nonhomogeneous pattern of evolution, and also insufficient number of complete genome sequences for most subtypes and to the use of very short genomic regions. There was also an accumulation of small errors and misunderstanding, including confusion with incompatible nomenclatures, ambiguous classification of possible artefactual recombinants and, mainly, insufficient number of widely accepted reference sequences. Additionally the attempt to derive information about subtypes (with 10-18% of nucleotide differences intersubtypes in a genotype) using alignments and phylogenetic analyses optimized for the study of the whole HEV diversity (including all genus and species, with 30-40% of nucleotide differences, and thus eliminating many of the nucleotides positions due to saturation of the phylogenetic signal) could lead to fundamentally erroneous conclusion about the inconsistency of the subtype scheme of a given genotype (Chapter 1-ref [19]). Driven by this situation the subtyping was discouraged at all (Chapter 5-ref 30). Thus, one of our first task was to analyze and actualize the subtype scheme for genotype 3 described by Li and to propose a revised set of reference sequences (Chapter 5).

For primer and probe design, an alignment of 351 HEV sequences was constructed. This alignment was manually curated using both the nucleotide and the deduced amino acid sequences. HEV-1, 2,-3 and -4 genotypes were included (with preference to genotype 3), covering all subtypes, and including 131 complete genomes (CG) (48 of them cited by Lu *et al.* (Chapter 5-ref 32)) and 65 German HEV sequences (Chapter 5-**Supplementary Table S01.**). For genotyping and subtyping, four sets of nested degenerated primers were selected from this alignment, which target different regions of the genome. Previously published primers (Chapter 5-ref 67) were used to amplify a fragment of the RNA-dependent RNA polymerase (RdRp) region. (Chapter 5-**Table 1**). By april-2017 this alignment include 1984 aligned sequences (with the corresponding isolate data), from which 1466 have an assigned subtype. This protocol for subtyping have been used in other five studies on HEV in wild animals in Germany (Chapter 6), (Chapter 1-ref [20]–[23]).

Our general approach for subtyping was to construct a reference tree using CG of genotypes 1-4. This tree defines the genotypes clades, and also shows the four monophyletic groups: “3jab”, “3chi”, “3feg” and “rabbit” that form the genotype 3 (Chapter 5-**Figure 1a**). Adding an unclassified sequence to this tree (or to trees that use partial sequences extracted from the underlying alignment) make possible

a classification onto genotypes (which can be achieved for sequences with lengths down to 200 or even 100 nt (Chapter 1-ref [24])) and a preliminary classification into major clade, group or subtype. By using only sequences of genotype 3 (Chapter 5-**Figure 1b**) or even only of the same clade or group we can obtain a tree with “better resolution” for a possible definitive classification of the sequence (less phylogenetic information is discarded due to saturation of the signal). This approach also permits a better definition of the groups and subtypes, to begin with. To investigate the suitability of smaller partial sequences to be used in subtyping we constructed a large number of trees and checked whether they reproduced the same topology that the corresponding CG tree and whether they offer sufficient (bootstrap) support to each of the interesting clades. We limit the subtyping to those sequence regions that reproduce the corresponding topology of the CG tree, avoiding drawing conclusions from the regions which do not (in particular, do not use it to investigate inconsistencies in the subtype scheme). Trees constructed for sequences more than 1500 nt length retained the same topology and support (with bootstrap values of at least 80% for each of the subtypes with more than one sequence) and are the ones used to define the subtypes and the reference sequences. This has a good agreement with the posteriorly published conclusion of Purdy and Sue of a 1600 nt limit (Chapter 1-ref [24]). They also concluded that “Subtyping may be done, but it requires a careful examination of the region to be used to ensure that it correctly resolves the subtypes”. We found regions which can be used (Chapter 5-**Figure 3, Figure 4** and **Supplementary Figure S03**), and also that the quality degrades with smaller lengths down to regions that provided insufficient information to subtype most sequences, like the ORF3.225nt (Chapter 5-**Supplementary Figure S11.**) or the ORF2.148nt.

The construction of high number of trees was possible by the combination of the software MEGA v6 and a specially developed Excel worksheet [HEVsubtypingMEGAut.xlsm](#). While the alignment [HEV.fas](#) refers the sequences only by the GenBank acc. ID the workbook include information we gained from the NCBI or the original publications about the sequence, isolate or strain, and also the position in the alignment. This permits an easy selection of the sequences (for example only from genotype 3 or major clade I, or from some continent or country, or host) that overlap the selected genomic region and export it to a text “group” file that can be imported in MEGA to quickly select and label the sequences to analyze. All of this is available at <https://github.com/qPCR4vir/HEV> together with detailed instruction of use <https://github.com/qPCR4vir/HEV/wiki> -especially the [Quick Guide of use](#).

While there are still sequences that could not be assigned to any subtype, the proposed subtype scheme provides a framework with which the majority of the sequences can be classified and thus “compared” to other non-overlapping sequences. Subtypes may serve as markers for epidemiological (and transmissions) studies and as preliminary grouping to search for biological differences.

In Mecklenburg-Western Pomerania (Germany) we detected a virus prevalence (HEV RNA) of 3.4% (32 out of 955) and 5.2% (three out of 58) in blood samples of wild boar hunted in 1996/97 and 2005/2006 respectively. In addition, HEV RNA was found in 14 out of 134 wild boar liver samples from the Greifswald region, giving a prevalence rate of about 10.4%. Finally, two wild boar liver samples

(WS03-09 and WS05-09) from individual hunts were also positive. Sequences were recovered from 12 of this samples and were classified (Chapter 5-**Supplementary Table S02**) into genotype 3 in the subtypes: 3a, which has been detected in Germany in Potsdam, Brandenburg and Bavaria (Chapter 5-ref 14,15) and is predominant in USA; 3b, the first isolated from a wild animal outside Japan – from where it probably originated -, and only once detected in Europa (Chapter 5-ref 65); 3jab – with insufficient information for a better subtyping -; and 3i, commonly detected in German wild boars.

In a similar study (Chapter 6) we found HEV RNA in four out of 13 serum samples from wild rabbits collected in the context of a general surveillance program for wildlife pathogens in 1989 in the Greifswald area in Mecklenburg Western Pomerania. From one sample a partial sequence could be recovered that correspond to the HVR and clustered with others rabbit derived HEV sequences. This rabbit sequences form a separate cluster or major clade within the genotype 3. Our sequence was related to two other sequences from France (2007 and 2008) and from China (2008-2011). An additional sequence of this clade was obtained from a human patient in France, which point to a zoonotic potential, which was confirmed by the detection of human derived 3g-like sequence in a German wild rabbit from another study (Chapter 1-ref [23]).

The global distribution of subtypes was tabulated in Chapter 5-**Supplementary Table S01.**, summarized in Chapter 5-**Table 2**, and partially visualized in Chapter 5-**Supplementary Figure S13** which can be interactively viewed at [Tableau-qPCR4vir](#). Subtype 3a appears to be distributed worldwide, but is predominant in USA. 3b is mainly restricted to Japan. Lack of long enough 3chi sequences make it difficult to define its internal structure and we expect that new subtypes will be defined here (3h-like sequences are found in Europe, Latin America and Africa). 3chi is common in Europe, particularly 3c which is predominant in Germany and The Netherlands, but also 3i as it is common in German wild boars. 3feg is also more common in Europe, especially 3f in Spain, France and The Netherlands while 3e in Hungary. In Germany both 3e and 3f are also common. Only one 3f strain was isolated from wild boar (1% from the total of 94 wild boar analyzed), but 204 were isolated from humans and domestic pigs (near 20% of the 998 summing from all subtypes). We found that a reported switch in England and Wales (Chapter 1-ref [25]) from almost only one HEV type in patients in 2005-2007 to another type which becomes majoritarian in 2011-2012 is actually an original subtype 3e being complemented with the introduction of 3c, which was common in neighbor countries.

1.7. Flavivirus detection and identification

Flaviviruses (family: *Flaviviridae*, genus: *Flavivirus*) represent one of the biggest health problems overall. Historically they had caused multiple devastating epidemics each with tens and hundreds of thousands of human deaths. Still today flaviviruses cause serious (re)emerging diseases, e.g. Dengue fever, Japanese encephalitis, Tick borne encephalitis, West Nile fever und Yellow fever among others. Recently, on 1 February 2016, the WHO declared a Public Health Emergency of International Concern

for the spread of Zika virus infections that was presumably provoking cases of microcephaly in babies which caused alarm among pregnant women. That presumption was subsequently confirmed (Chapter 1-ref [26][27]).

For many members of the *Flavivirus* family exist sensitive and specific detection assays, mostly RT-qPCR. In most situations, however there is a need to screen or monitor for several flaviviruses, ideally most of them. But using individual assays in large screening or surveillance programs is very time consuming and both technically and economically inconvenient.

The genomic sequences of the genus are highly heterogeneous, with differences of more than 30% between species. Nevertheless, Pan-*Flavivirus* assays, potentially capable of detecting most of them have been developed. Commonly, these assays target the relatively conserved NS5 which codes for the RNA dependent RNA polymerase. We gathered actualized (2010) sequence information to optimize one of the most used assays (Chapter 7-ref 21) to obtain a more homogeneous sensitivity of detection for most flaviviruses (Chapter 7). Nearly 200 complete *Flavivirus* genome sequences were obtained from GenBank (very similar sequences, with more than 98% identity were excluded). The sequences were aligned automatically with manual correction taking into account the deduced amino acid sequences. Partial sequences were added to reach a total of more than 400 to represent the NS5 gene of all known species, including most subtypes or lineages (a regularly updated version is available at our [Flavivirus GitHub site](#)).

To facilitate visualization and selection of appropriate primers which hybridize equally well to sequences of any group, the fragments within the NS5 (spanning nucleotides 9040 to 9305 of AF196835) of the aligned sequences were imported into *VisualOligoDeg* and manually classified into major groups (MB - mosquito borne, TB - tick borne and Insect-only), virus groups (JEVG, YFVG, TBEVG, etc.), species (WNV, JEV, YFV, TBEV, etc.) and in some cases into lineages (like WNV-1 or WNV-2, etc.) or genotypes. The modified primers PFlav-fAAR (TACAACATGATGGGAAAGAGAGAGAARAA corresponding to nt 9040 to 9068 of AF196835) and PFlav-rKR (GTGTCCCAKCCRGCTGTGTCATC from nt 9305 to 9283 of AF196835) were used in a SYBRG-RT-qPCR protocol with subsequent melting curve analysis and experimentally evaluated for detection of RNA of 26 different flaviviruses, demonstrating for many of them better sensitivity than the original primers. The observed shift in the values for C_q, ranging from of 5 to 13, (Chapter 7, FIGURE 1) for many viruses may be traduced in improvements of more than 10 to 1000 times in the limit of detection.

Detecting and quantifying a flavivirus is a first step, which has to be followed by a test to identify the specie. One universal solution is the sequencing of the amplicon which provides in this case enough information to identify the virus. But this solution has some known technical difficulties and a few limitations. One of them is that without a cloning step it provide only the consensus or majority sequence and possible minor sequences or mixes are not detected.

To overcome these problems and to improve the throughput, microarray based assays have been developed. We continued that path and developed a microarray assay (Chapter 7) based on the commercial, low density, and potentially low cost and high throughput, platform of Alere Technologies GmbH, Jena, Germany (previously Clondiag). The assay is simple to perform in no more than two hours, transferring 1 μ L directly from the PCR reaction into the hybridization buffer provided with the kit. The selection of the probes, in contrary, was challenging. In other applications there is a need to find a set of probes each of which identify one group of viruses (let say, a genotype or specie and do not react to others), selecting a zone with mutations in any other virus. This approach exploits relatively conserved regions with only sufficient variability to detect differences between groups. In our situation, in contrast, due to the extreme high variability of flaviviruses, the major challenge was instead to find a probe or set of probes that just could detect all variants of the given group, e.g. a WNV lineage 2, a Dengue virus 1, JEV, etc. (Chapter 7, **TABLE S01**, with 84 probes with a T_m around 55°C). Following this approach not a unique probe was selected (with *VisualOligoDeg*) to identify each unique virus variant, but a group of probes that collectively identifies a group of viruses. Thus each viral group is characterized by a - “pattern” of probe reactions. Initially we found no software completely appropriated for this situation, and some were slow for the number of test we needed. For the interpretation of the results, we also needed a software to model probe-targets hybridizations, preferably without previous alignment of all the target sequences. The solution was to develop two new software (see *VisualOligoDeg* and *ThDy_DNAHybrid* in section “Primers and probe design”).

To finally identify the virus, new approaches were developed to process the resulting raw data obtained after the initial processing of the image of the experimental hybridization. In the simplest situations a visual inspection of the intensity values using a bar diagram is sufficient to identify the sample (Chapter 7, **FIGURE 3**). While the hybridization patterns are not significantly affected by variations in the quantitative viral load of the sample, changes are, in contrast, readily visible when different strains of the same virus species are examined. Other situations needed more advanced processing. A *new script* was embedded into the Partisan IconoClust® v3.6r0 software (Alere Technologies GmbH) to export the normalized intensity signals in different formats, including a format suitable for import into *Orange* (Chapter 1-ref [28])— software, which allows visual programming and python scripting (<http://orange.biolab.si>).

In *Orange*, the visual program *PanFlavExpStdSampl* (Chapter 7, **FIGURE S01**) was created, which, together with custom modifications of some *parts of Orange* itself, permits an interactive import of the data from experiments used as standards (known samples) for parallel analysis (comparison) with unknown samples for classification. In one approach, an *Orange* “widget” shows a tree (Chapter 7, **FIGURE 4**), in which a cut-off can be interactively selected and groups may be defined to make a report of the proposed classification (this is based on cluster analysis carried out on the distances between the hybridization patterns). This is the preferred method to visualize the identification of viruses in

samples that contain only one single flavivirus. In another approach a heat map-like graphics visualizes calculated distances between intensity patterns (Chapter 7, FIGURE 5), in which the labels and the order of the samples can be interactively selected from a group of pre-options. We have added (directly [modifying the source code of Orange](#)) the option (by mouse clicking the respective cell) of reorganizing the heat map showing the selected sample at the top, followed by the most similar samples. This graphic can also show the samples organized in a tree to reveal clustering and perhaps more importantly, it also permits visualization of mixes of flaviviruses (Chapter 7, FIGURE 5).

The limitation of comparing each new experimental sample only with other known experimental samples could be overcome using virtual hybridizations as standards. *ThDy_DNAHybrid* was used to introduce the sequences of the desired “virtual standard” strains and to calculate the modeled ΔG of hybridization (at 55°C, the temperature used in our experiments) of that sequence with each of the probes present in our microarray. This modeled data is imported into *PanFlavExpVirtStdSampl* a modified *Orange* visual program (Annex, Figure 5) which includes a conversion from ΔG to I (intensity). The last is achieved by the simplified empirical formula:

$$I = I(\Delta G, \Delta G_{sat}, \Delta G_{senc}) = I_{sat} e^{-\frac{\Delta G - \Delta G_{sat}}{\Delta G_{sat} - \Delta G_{senc}} \ln(I_{senc}/I_{sat})}$$

where the values of the saturating intensity and ΔG (I_{sat} and ΔG_{sat}) and of the sensitivity limit (I_{senc} and ΔG_{senc}) were selected empirically. For $\Delta G < \Delta G_{sat}$ I is set to I_{sat} . The final formula used in *VisualOligoDeg* and in *PanFlavExpVirtStdSampl* for the normalized intensity (maximum set to 1) is just a convenient transformation:

$$I_{norm} = I_{norm}(\Delta G, \Delta G_{sat}, \Delta G_{senc}) = \vartheta e^{\Delta G \rho}$$

Where:

$$\rho = \frac{\ln(I_{senc})}{\Delta G_{senc} - \Delta G_{sat}}, \quad \vartheta = e^{-\Delta G_{sat} \rho}$$

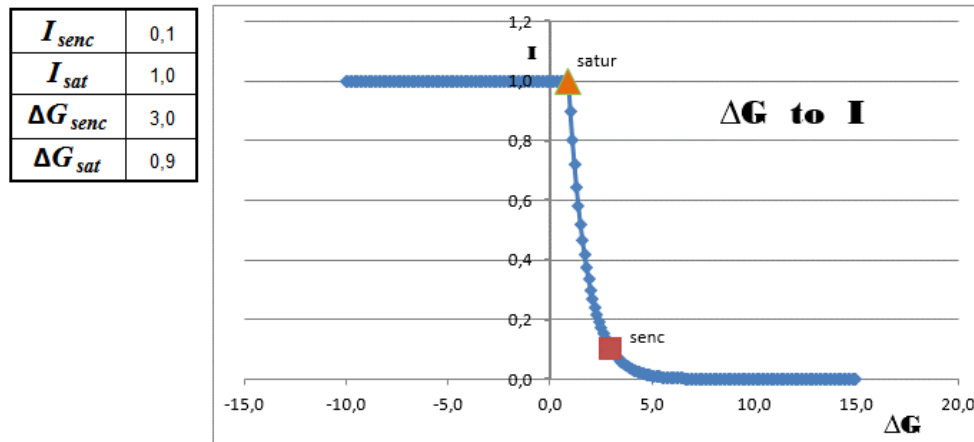


Figure 3. Transformation of the modeled ΔG (for hybridization of the selected target sequence with each probe of the microarray) into a “virtual” microarray by modeling the resulting intensity using an empirical formula with empirically selected values for the parameters. The intensity uses arbitrary units, while ΔG is given in kcal/mol.

The procedure is robust enough to permit significant variations of these parameter and still most of experimental standards cluster together with the corresponding virtual standards.

Flavivirus RT-qPCR screening was conducted (Chapter 7) on 340 mosquito pools collected in Greece in 2012. One pool yielded a strong positive result (Cq – 21.5), and two resulted of medium (Cq – 31.6 and 32.5) and two of weak concentrations (Cq – 37.3 and 38.5). PCR products of 13 mosquito pools (including all positive specimens, as well as six negative) underwent microarray analysis (Chapter 7, TABLE 2), which revealed the presence of WNV lineage 2 sequences, similar to the Austria strain, in five of them. Sequencing of four of these five amplicons revealed full identity to the WNV lineage 2 isolates Hungary/04 (acc. n. DQ116961) and Nea Santa-Greece-2010 (acc. n. HQ537483).

1.8. Conclusions

RNA virus classification and taxonomy directly influences the development of new DNA-hybridization based diagnostic techniques. It affects how the target group is defined and what sequences are selected to be used during the design of the assay. The high sequence variation both intra and inter groups, together with insufficient sequence information and the potentially (only apparently contradicting) high number of those sequences make the task not trivial. We developed two software tools, which conveniently complement other widely used, for design or evaluation of primers and probes. We used then in the development and application of assays for detection and identification of a wide groups of RNA viruses, many of them linked to important (re)emerging animal and human diseases.

Our published assays for detection and quantification of WNV 1 and 2 (Chapter 2) have been successfully used by numerous independent laboratories (Chapter 1-ref [29]–[38]) and contributed to the detection of WNV lineage 2 (Chapter 1-ref [30][37]) in different regions of the world and also to the follow up of vaccine- (Chapter 1-ref [39][40]) and infection experiments (Chapter 1-ref [35], [38], [41], [42]). The assays are also used in WNV monitoring or surveillance programs (Chapter 1-ref [43][44]).

The VEEV posed (Chapter 3) us a particular challenge: we used *VisualOligoDeg* to deduce a set of sequences that comprised all combinations of observed mutations in the target regions of the primers and probe and generated a set of 15 synthetic RNA that includes a total of 10 subtypes from all 6 VEEV types. Experimentally it was demonstrated that the new assay is capable to detect all these subtypes.

The considerable effort of Lambert et al. (Chapter 4-r6) in creating a multiplex PCR assay for the detection of medical important Bunyaviruses (grouped into a classification tree of variable deep) clearly illustrated the necessity and convenience of the tools we created. Interesting, we successfully used (Chapter 4) their Bunyamwera virus primers for the detection of a virus (Ngari) that they could not test. This detection allows us to supports the extension of the range of Ngari virus infection to goats (it had been detected in a human and mosquitoes) (Chapter 4, r8) and demonstrates the occurrence of Ngari virus infection during the 2010 RVFV outbreak in Mauritania. We are aware of only one additional

report of infection, in sheep also in Mauritania in 1988, although no further characterization was conducted (Chapter 4, r9).

We designed (Chapter 5) RT-PCR assays for screening, quantification and genotyping of HEV-3 strains, and detected viral RNA in wild boar samples from Mecklenburg-Western Pomerania, Germany. Twelve strains clustered into subtypes 3a, 3i and, unexpectedly, also 3b, which is a common subtype in Japan, but had never been reported in animals in Europe. The phylogenetic trees based on our partial sequences of ORF1, RdRp, HVR and ORF2 regions reproduced similar topology as obtained from complete genome analysis and were useful for subtyping.

More than 30 different PCR fragments and corresponding genomic regions, without sufficient standardization, have been used for genotyping and subtyping so far, being a source of ambiguous subtyping schemes and inadequate classification. The presented study offers an updated set of reference sequences for the relatively simple and neutral subtype scheme proposed by Lu et al. (Chapter 5-ref 32), which could eliminate most of the existing incongruences and creates the basis for new hypotheses regarding the Hepatitis E epidemiology. In the future, a comprehensive subtyping of all sequenced HEV-3 isolates according to this classification scheme could enable a detailed view of the spread of HEV-3 strains among pigs, wild life and humans, and could allow to determine the consequences of infections with different subtypes on humans and finally help to limit the spread of the disease. Our published findings (Chapter 5) anticipated the publication of the International Committee on the Taxonomy of Viruses, that reconsider their negative opinion about subtyping (Chapter 5-ref 68), recommending now (Chapter 1-ref [45]) a subtype scheme with a set of reference sequences similar to the one we selected. Both sets are not only almost compatible, but also complementary, because while the Committee proposes only one “central” reference for each subtype we aim to identify every sequence which can be used as a reference, thus defining also the current “limits” of each subtype. We also offer the alignments and other tools that facilitate the subtyping.

Our finding in a retrospective study for the first time of HEV in wild rabbit in Germany (Chapter 6) contributed to support the need for of a well- structured wildlife surveillance program in Germany and elsewhere.

There is an urgent and global need for monitoring and surveillance of *Flavivirus*. A sensitive, quick and high throughput assay may add a significant progress in that direction. Our new RT-qPCR-microarray assay have provided a promising starting point (Chapter 7). With an improved set of primers we were able to detect and identify 26 reference strains and to identify *Flavivirus* members in experimental and field samples. An important distinction of this microarray platform from well-known glass-slide arrays used for gene expression studies, is that it is optimized to detect genetic (sequence) variations, rather than the concentration or relative quantity of amplicons. Thus, the present microarray signal intensity values are used solely for identification or classification, while quantification is performed in the preceding RT-qPCR step.

Ours both new software tools were used during the microarray design and interpretation of the results. In particular novel scripts implemented in *Orange* integrate experimental standards and “virtual” ones obtained by modeling DNA hybridizations and permit rapid identification of the virus found in positive PCR samples. Nevertheless the potential use of machine learning algorithms within *Orange* had yet to be explored.

Using the combined assay, five, out of 340 mosquitos pools from Greece (2012) were found to contain WNV lineage 2 similar to the strain previously (2010) found there. This points to the necessity of continue control and monitoring of mosquitos in that country.

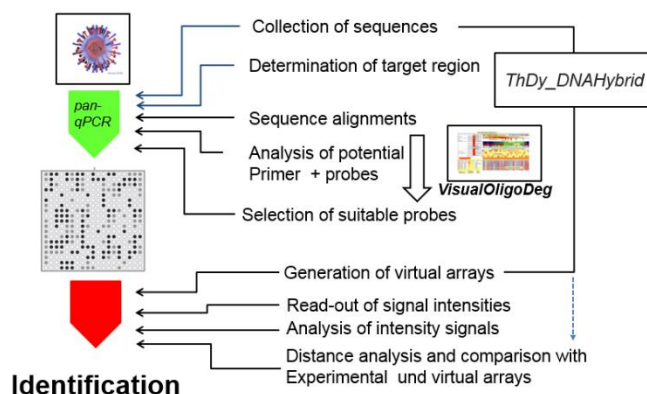


Figure 4. Simplified general schema of the design, use and analyses of results of an assay (similar to Chapter 7) for detection and identification of a given group of RNA viruses and the potential role of the new tools *VisualOligoDeg* and *ThDy_DNAHybrid*. It underlines the importance of the design of primers and probes from a set of highly variable sequences in a complex classification scheme, and also the potential use of modeled hybridizations to cover sequences for which no experimental standard are available.

In summary, we want to stress the importance and complexity of the initial design step of selection of primer and probes candidates that solve the major problem in RNA virus detection: the high sequences variation, both inter and intra target groups. We offer two tools that may help during that selection (Figure 4).

1.9. References

Note: to avoid duplications of references already used in the included papers they are annotated in the text of the thesis abstract as a direct link in the form: (Chapter X-ref Y), where X is a link to the Chapter including the reference and -ref Y is a link to the reference listed in that Chapter.

- [1] C. J. Finlay, "El mosquito hipoteticamente considerado como el agente de transmisión de la fiebre amarilla," *An. la Acad. Ciencias Médicas, Físicas y Nat. la Habana.*, vol. XVIII, pp. 147–69, 1881.
- [2] C. J. Finlay, "The mosquito hypothetically considered as an agent in the transmission of yellow fever poison," *New Orleans Med. Surg. journal.*, vol. 9, no. 8, pp. 601–16, 1882.
- [3] C. J. Finlay, "The Mosquito Hypothetically Considered as an Agent in the Transmission of Yellow Fever Poison," *Yale J. Biol. Med.*, vol. 9, no. 6, pp. 589–604, Jul. 1937.
- [4] R. Roffey, A. Tegnell, and F. Elgh, "Biological warfare in a historical perspective," *Clin. Microbiol. Infect.*, vol. 8, no. 1198–743X (Print), pp. 450–4, 2002.
- [5] E. Croddy, C. Perez-Armendariz, and J. Hart, *Chemical and Biological Warfare*. New York, NY: Springer New York, 2002.
- [6] J. SantaLucia, "Physical Principles and Visual-OMP Software for Optimal PCR Design," in *PCR Primer Design*, vol. 402, A. Yuryev, Ed. Totowa, NJ, 2007, pp. 3–33.
- [7] L. Kaderali, "Primer Design for Multiplexed Genotyping," in *PCR Primer Design*, A. Yuryev, Ed. Totowa, NJ: Humana Press, 2007, pp. 269–85.
- [8] C. Chancey, A. Grinev, E. Volkova, and M. Rios, "The Global Ecology and Epidemiology of West Nile Virus," *Biomed Res. Int.*, vol. 2015, no. Figure 2, pp. 1–20, 2015.
- [9] L. H. Gould and E. Fikrig, "West Nile virus: a growing concern?," *J. Clin. Invest.*, vol. 113, no. 8, pp. 1102–7, Apr. 2004.
- [10] B. L. Johnston and J. M. Conly, "West Nile virus - where did it come from and where might it go?," *Can. J. Infect. Dis.*, vol. 11, no. 4, pp. 175–8, 2000.
- [11] M. Niedrig, S. Linke, H. Zeller, C. Drosten, and B. West, "First International Proficiency Study on West Nile Virus Molecular Detection," *Emerg. Infect. Dis.*, vol. 1854, pp. 1851–4, 2006.
- [12] J. Jose, J. Snyder, and R. Kuhn, "A structural and functional perspective of alphavirus replication and assembly," *Futur. Microbiol.*, vol. 4, p. 30, 2009.
- [13] "CDC. Emergency Preparedness and Response. Bioterrorism Agents/Diseases." [Online]. Available: <https://emergency.cdc.gov/agent/agentlist.asp#v>. [Accessed: 12-May-2017].
- [14] C. Madsen, I. Hooper, L. Lundberg, N. Shafagati, A. Johnson, S. Senina, C. de la Fuente, L. I. Hoover, B. L. Fredricksen, J. Dinman, J. L. Jacobs, and K. Kehn-Hall, "Small molecule inhibitors of Ago2 decrease Venezuelan equine encephalitis virus replication," *Antiviral Res.*, vol. 112, pp. 26–37, 2014.
- [15] S. C. Weaver, M. Anishchenko, R. Bowen, A. C. Brault, J. G. Estrada-Franco, Z. Fernandez, I. Greene, D. Ortiz, S. Paessler, and A. M. Powers, "Genetic determinants of Venezuelan equine encephalitis emergence," in *Emergence and Control of Zoonotic Viral Encephalitides*, no. February 2004, Vienna: Springer Vienna, 2004, pp. 43–64.
- [16] I. Bunyaviridae Study Group, I. Emaravirus Study Group, and I. Emaravirus Study Group, "Create a new order, Bunyavirales, to accommodate nine families (eight new, one renamed) comprising thirteen genera," *ICTV Propos.*, no. 2016.030a-vM, pp. 1–45, 2016.
- [17] M. Eiden, A. Vina-Rodriguez, B. O. El Mamy, K. Isselmou, U. Ziegler, D. Höper, S. Jäckel, A. Balkema-Buschmann, H. Unger, B. Doumbia, and M. H. Groschup, "Ngari Virus in Goats during Rift Valley Fever Outbreak, Mauritania, 2010," *Emerg. Infect. Dis.*, vol. 20, no. 12, pp. 2174–6, 2014.
- [18] R. S. Tedder, S. Ijaz, A. Kitchen, I. Ushiro-Lumb, K. I. Tettmar, P. Hewitt, and N. Andrews, "Hepatitis E risks: pigs or blood-that is the question," *Transfusion*, vol. 57, no. 2, pp. 267–72, 2017.
- [19] E. F. Oliveira-Filho, M. König, and H.-J. Thiel, "Genetic variability of HEV isolates: inconsistencies of current classification," *Vet. Microbiol.*, vol. 165, no. 1–2, pp. 148–54, Jul. 2013.
- [20] S. Neumann, S. S. Hackl, M. Piepenschneider, A. Vina-Rodriguez, P. Dremsek, R. G. Ulrich, M. H. Groschup, and M. Eiden, "Serologic and Molecular Survey of Hepatitis E Virus in German Deer

Populations.,” *J. Wildl. Dis.*, vol. 52, no. 1, pp. 106–13, 2015.

- [21] J. Schlosser, M. Eiden, A. Vina-Rodriguez, C. Fast, P. Dremsek, E. Lange, R. G. R. G. Ulrich, and M. H. M. H. Groschup, “Natural and experimental hepatitis E virus genotype 3 - infection in European wild boar is transmissible to domestic pigs,” *Vet. Res.*, vol. 45, no. 1, pp. 1–13, Jan. 2014.
- [22] J. Schlosser, A. Vina-Rodriguez, C. Fast, M. H. Groschup, and M. Eiden, “Chronically infected wild boar can transmit genotype 3 hepatitis E virus to domestic pigs,” *Vet. Microbiol.*, vol. 180, no. 1–2, pp. 15–21, Oct. 2015.
- [23] F. Hammerschmidt, K. Schwaiger, L. Dähnert, A. Vina-Rodriguez, D. Höper, M. Gareis, M. H. Groschup, and M. Eiden, “Hepatitis E virus in wild rabbits and European brown hares in Germany,” *Zoonoses Public Health*, no. November 2016, pp. 1–11, Mar. 2017.
- [24] M. A. Purdy and A. Sue, “The effect of phylogenetic signal reduction on genotyping of hepatitis E viruses of the species Orthohepevirus A,” *Arch. Virol.*, 2016.
- [25] S. Ijaz, B. Said, E. Boxall, E. Smit, D. Morgan, and R. S. Tedder, “Indigenous hepatitis E in England and Wales from 2003 to 2012: evidence of an emerging novel phylotype of viruses,” *J. Infect. Dis.*, vol. 209, no. 8, pp. 1212–8, Apr. 2014.
- [26] S. A. Rasmussen, D. J. Jamieson, M. A. Honein, and L. R. Petersen, “Zika Virus and Birth Defects — Reviewing the Evidence for Causality,” *N. Engl. J. Med.*, pp. 1–7, 2016.
- [27] M. R. Reynolds, A. M. Jones, E. E. Petersen, E. H. Lee, M. E. Rice, A. Bingham, S. R. Ellington, N. Evert, S. Reagan-Steiner, T. Oduyebo, C. M. Brown, S. Martin, N. Ahmad, J. Bhatnagar, J. Macdonald, C. Gould, A. D. Fine, K. D. Polen, H. Lake-Burger, C. L. Hillard, N. Hall, M. M. Yazdy, K. Slaughter, J. N. Sommer, A. Adamski, M. Raycraft, S. Fleck-Derderian, J. Gupta, K. Newsome, M. Baez-Santiago, S. Slavinski, J. L. White, C. A. Moore, C. K. Shapiro-Mendoza, L. Petersen, C. Boyle, D. J. Jamieson, D. Meaney-Delman, and M. A. Honein, “Vital Signs: Update on Zika Virus—Associated Birth Defects and Evaluation of All U.S. Infants with Congenital Zika Virus Exposure — U.S. Zika Pregnancy Registry, 2016,” *Morb. Mortal. Wkly. Rep.*, vol. 66, no. 13, pp. 366–73, 2017.
- [28] J. Demšar, T. Curk, A. Erjavec, T. Hočevár, M. Milutinovič, M. Možina, M. Polajnar, M. Toplak, A. Starič, M. Stajdohar, L. Umek, L. Zagar, J. Zbontar, M. Zitnik, and B. Zupan, “Orange: Data Mining Toolbox in Python,” *J. Mach. Learn. Res.*, vol. 14, p. 23492353, 2013.
- [29] S. C. Barros, F. Ramos, L. Zé-Zé, M. J. Alves, T. Fagulha, M. Duarte, M. Henriques, T. Luís, and M. Fevereiro, “Simultaneous detection of West Nile and Japanese encephalitis virus RNA by duplex TaqMan RT-PCR,” *J. Virol. Methods*, vol. 193, no. 2, pp. 554–7, 2013.
- [30] G. Savini, G. Puggioni, A. DI Gennaro, G. DI Francesco, a M. Rocchigiani, A. Polci, V. Marini, C. Pinoni, S. Rolesu, G. Marruchella, A. Lorusso, and F. Monaco, “West Nile virus lineage 2 in Sardinian wild birds in 2012: a further threat to public health,” *Epidemiol. Infect.*, vol. 141, no. 11, pp. 2313–6, Nov. 2013.
- [31] M. J. Frost, J. Zhang, J. H. Edmonds, N. A. Prow, X. Gu, R. Davis, C. Hornitzky, K. E. Arzey, D. Finlaison, P. Hick, A. Read, J. Hobson-Peters, F. J. May, S. L. Doggett, J. Haniotis, R. C. Russell, R. A. Hall, A. A. Khromykh, and P. D. Kirkland, “Characterization of virulent West Nile virus Kunjin strain, Australia, 2011,” *Emerg. Infect. Dis.*, vol. 18, no. 5, pp. 792–800, May 2012.
- [32] I. Pem-Novosel, T. Vilibic-Cavlek, I. Gjenero-Margan, N. Pandak, L. Peric, L. Barbic, E. Listes, A. Cvitkovic, V. Stevanovic, and G. Savini, “First Outbreak of West Nile Virus Neuroinvasive Disease in Humans, Croatia, 2012,” *Vector-Borne Zoonotic Dis.*, vol. 14, no. 1, pp. 82–4, 2014.
- [33] J. S. Kumar, D. Saxena, and M. Parida, “Development and comparative evaluation of SYBR Green I-based one-step real-time RT-PCR assay for detection and quantification of West Nile virus in human patients,” *Mol. Cell. Probes*, vol. 28, no. 5–6, pp. 221–7, 2014.
- [34] S. U. Biçeroğlu, E. Karatayli, A. Bayram, A. Turhan, A. Değirmenci, Y. Aydinok, A. M. Bozdayi, and Ş. R. Sertöz, “Investigation of West Nile virus among healthy blood donors in the western part of Turkey,” *Turkish J. Med. Sci.*, vol. 45, no. 1, pp. 84–8, 2015.
- [35] S. Friedrich, T. Schmidt, A. Schierhorn, H. Lilie, G. Szczepankiewicz, S. Bergs, U. G. Liebert, R. P. Golbik, and S.-E. Behrens, “Arginine methylation enhances the RNA chaperone activity of the West Nile virus host factor AUF1 p45,” *RNA*, vol. 22, no. 10, pp. 1574–91, 2016.
- [36] S. Pavone, S. Salamida, I. Pecorelli, E. Rossi, and E. Manuali, “Deadly outbreak of iron storage disease (ISD) in Italian birds of the family Turdidae,” *J. Vet. Med. Sci.*, vol. 76, no. 9, pp. 1209–12, 2014.
- [37] E. Patsoula, A. Vakali, G. Balatsos, D. Pervanidou, S. Beleri, N. Tegos, A. Baka, G. Spanakos, T. Georgakopoulou, P. Tserkezou, W. Van Bortel, H. Zeller, P. Menounos, J. Kremastinou, and C. Hadjichristodoulou, “West Nile Virus Circulation in Mosquitoes in Greece (2010 – 2013),” vol. 2016, 2016.

- [38] M. SPEDICATO, I. CARMINE, A. L. BELLACICCO, G. MARRUCHELLA, V. MARINI, M. PISCIELLA, G. DI FRANCESCO, A. LORUSSO, F. MONACO, and G. SAVINI, “Experimental infection of rock pigeons (*Columba livia*) with three West Nile virus lineage 1 strains isolated in Italy between 2009 and 2012,” *Epidemiol. Infect.*, no. October, pp. 1–11, 2015.
- [39] J. Angenvoort, D. Fischer, C. Fast, U. Ziegler, M. Eiden, J. Garcia de la Fuente, M. Lierz, and M. H. Groschup, “Limited efficacy of West Nile virus vaccines in large falcons (*Falco* spp.),” *Acta Vet. Scand.*, vol. 45, no. 1, p. 41, 2014.
- [40] D. Fischer, J. Angenvoort, U. Ziegler, C. Fast, K. Maier, S. Chabierski, M. Eiden, S. Ulbert, M. H. Groschup, and M. Lierz, “DNA vaccines encoding the envelope protein of West Nile virus lineages 1 or 2 administered intramuscularly, via electroporation and with recombinant virus protein induce partial protection in large falcons (*Falco* spp.),” *Vet. Res.*, vol. 46, no. 1, p. 87, 2015.
- [41] U. Ziegler, J. Angenvoort, D. Fischer, C. Fast, M. Eiden, A. V. Rodriguez, S. Revilla-Fernández, N. Nowotny, J. G. de la Fuente, M. Lierz, and M. H. Groschup, “Pathogenesis of West Nile virus lineage 1 and 2 in experimentally infected large falcons,” *Vet. Microbiol.*, vol. 161, no. 3–4, pp. 263–73, 2013.
- [42] K. Schmidt, M. Keller, B. L. Bader, T. Korytář, S. Finke, U. Ziegler, and M. H. Groschup, “Integrins modulate the infection efficiency of West Nile virus into cells,” *J. Gen. Virol.*, vol. 94, no. Pt 8, pp. 1723–33, Aug. 2013.
- [43] U. Ziegler, D. Seidowski, J. Angenvoort, M. Eiden, K. Müller, N. Nowotny, and M. H. Groschup, “Monitoring of West Nile Virus Infections in Germany,” *Zoonoses Public Health*, vol. 59, no. SUPPL.2, pp. 95–101, 2012.
- [44] U. Ziegler, H. Jöst, K. Müller, D. Fischer, M. Rinder, D. T. Tietze, K.-J. Danner, N. Becker, J. Skuballa, H.-P. Hamann, S. Bosch, C. Fast, M. Eiden, J. Schmidt-Chanasit, and M. H. Groschup, “Epidemic Spread of Usutu Virus in Southwest Germany in 2011 to 2013 and Monitoring of Wild Birds for Usutu and West Nile Viruses,” *Vector-Borne Zoonotic Dis.*, vol. 15, no. 8, pp. 481–8, 2015.
- [45] D. B. Smith, P. Simmonds, J. Izopet, E. F. Oliveira-Filho, R. G. Ulrich, R. Johne, M. Koenig, S. Jameel, T. J. Harrison, X. J. Meng, H. Okamoto, W. H. M. Van Der Poel, and M. A. Purdy, “Proposed reference sequences for hepatitis E Virus subtypes,” *J. Gen. Virol.*, vol. 97, no. 3, pp. 537–42, 2016.

1.10. List of routinely used software

Writing/bibliography: Microsoft Office, [LibreOffice](#), Adobe Acrobat, [Mendeley](#), EndNote, [ResearchGate](#).

Programming:

C++: [Microsoft Visual Studio](#) v 2008-2017, [CLion](#), [CMake](#), [Nana C++ GUI library](#)

Python: [PyCharm](#),

R: [R Studio](#)

VBA: MS Excel

[EvoScript](#): [TECAN](#)

Source control: [SourceTree](#), [GitHub](#), [Bitbucket](#), [git](#), [Git-extensions](#), [Tortoise SVN](#) and [git](#).

Bioinformatics: [Orange](#), [MEGA](#), [Vector NTI](#) (old, free versions), [Geneious](#), [NCBI](#)

Chapter 2. Publication I

Two New Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction Assays with Unique Target Sites for the Specific and Sensitive Detection of Lineages 1 and 2 *West Nile Virus* Strains.

Eiden, Martin, Ariel Vina-Rodriguez, Bernd Hoffmann, Ute Ziegler, and Martin H Groschup.

Journal of Veterinary Diagnostic Investigation : Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc 22 (5): 748–53. 2010.

<http://jvdi.org/cgi/content/full/22/5/748>.

Reproduced with permissions.

Two new real-time quantitative reverse transcription polymerase chain reaction assays with unique target sites for the specific and sensitive detection of lineages 1 and 2 *West Nile virus* strains

Martin Eiden, Ariel Vina-Rodriguez, Bernd Hoffmann, Ute Ziegler, Martin H. Groschup¹

Abstract. Two novel 1-step real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) assays for the simultaneous detection of *West Nile virus* (WNV) lineage 1 and 2 strains were developed. Primers and the probe of assay 1 target the 5'-untranslated region (UTR), whereas the amplicon of assay 2 is located in the nonstructural region NS2A, which enables an unambiguous and independent WNV diagnosis based on 2 different amplicons. Both assays allow the detection of as few as 2–4 genome copies of WNV strains NY99, Uganda B956, Kunjin, and Sarafend (all cultured on Vero cells). A new synthetic RNA mutant of the 5'-UTR amplicon, which contains 6 twist inverted base-pair changes at the probe attachment site, was used as external calibrator control.

Key words: Real-time quantitative reverse transcription polymerase chain reaction; West Nile virus.

West Nile virus (WNV; family *Flaviviridae*, genus *Flavivirus*) was first detected in a woman in the West Nile District of Uganda in 1937.²⁴ *West Nile virus* is an arthropod-borne virus grouped in the *Japanese encephalitis virus* (JEV) serocomplex, which includes the *St. Louis encephalitis virus*, JEV, and *Murray Valley encephalitis virus* (MVEV), among others.¹¹ A large variety of wild bird species are the natural reservoir for WNV¹⁹; however, its host range is very broad and encompasses not only humans but also equids, alligators, dogs, sheep, and many other species.⁴ Human infections are characterized by flu-like illnesses that are associated with headache, high fever, chills, arthralgia, malaise, and retro-orbital pain. Up to 1% of infected human beings develop severe encephalitis, myelitis, and/or meningitis, and of these cases, 1 in 20 dies.²⁰

The introduction of WNV into New York in 1999 and its rapid spread lead to cases in almost all North American states and provinces, in addition to some Central and South American countries.¹⁰ In Europe, WNV was first detected in France¹⁴ and Portugal,⁹ and recent outbreaks have occurred in Romania,²¹ Italy,^{18,22} Hungary,⁸ and Austria.²⁵

West Nile virus consists of a linear, single-stranded, plus-sense RNA, which encodes for 3 structural (C, prM, and E) and 7 nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) proteins.⁴ It has been recently proposed³ that WNV be grouped into 5 lineages. Lineage 1 is found in some parts of southern Europe, Asia, Africa, and North America. The Kunjin virus, which circulates in Australia, represents a subtype of lineage 1.²³ Lineage 2 strains are

found in sub-Saharan Africa and Madagascar⁵ and have also recently been discovered in Hungary and Austria.² Lineage 3 is represented by a virus strain that was isolated from mosquitoes in the Czech Republic, designated the Rabensburg virus¹; lineage 4 was isolated from a tick isolate from the Caucasus.² *West Nile virus* strains from India, which group into a subcluster of lineage 1, are sometimes thought to represent lineage 5.³

The aim of the following study was to develop 2 one-step duplex real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) assays that target different regions of the WNV genome for an unambiguous identification of viral nucleic acid. For this purpose, the WNV strains NY99^a (GenBank accession no. AF196835), Uganda B956^a (GenBank accession no. AY532665), Sarafend^b (GenBank accession no. AY688948), and Kunjin^b (GenBank accession no. D00246) were used, which were grown on Vero (African green monkey kidney epithelial) cells. Virus titers were determined by 10-fold dilution series (in 8 replicates) in 96-well plates (100 μ l/well) on fresh cell monolayers and cytopathologic effects, read after an incubation period of 6–7 days at 37°C. Cells were subsequently fixed and stained with crystal violet. Virus titers (50% tissue culture infective doses [TCID₅₀]) were calculated using the Spearman–Kärber method.¹⁵ Moreover, *Yellow fever virus* (YFV) strain YF-17D,^c JEV,^d Tickborne encephalitis virus (TBEV) strain Langaat,^e and MVEV^b were used in the RT-qPCR specificity studies. Viral RNA was isolated from a cell culture medium using a commercial kit.^e Cell culture supernatant (140 μ l) was added to 560 μ l AVL (lysis) buffer,^e spiked with 5 μ l of an internal control RNA (IC-RNA) containing 2 \times 10⁵ copies/ μ l, and eluted from columns in a final volume of 50 μ l in AVE buffer^e and stored at -70°C until use.

Suitable primers and probes for the RT-qPCR detection of WNV lineage 1 and 2 strains were designed in silico by aligning full-length sequences of 186 flavivirus and 95 WNV isolates (from the National Center for Biotechnology

From the Institute for Novel and Emerging Infectious Diseases (Eiden, Vina-Rodriguez, Ziegler, Groschup) and the Institute for Diagnostic Virology (Hoffmann), Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany.

¹Corresponding Author: MartinH.Groschup, Institute for Novel and Emerging Infectious Diseases at the Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Südufer 10, 17493 Greifswald-Insel Riems, Germany. martin.groschup@fli.bund.de

Table 1. Primers and probes selected for West Nile virus (WNV)–specific real-time quantitative reverse transcription polymerase chain reaction.*

Assay/name	Oligo sequence	Genome position
Assay 1		
INEID f1	AGTAGTTCGCCTGTGTGAGC	(s) 1–20
INEID r1	GCCCTCCTGGTTTCTTAGA	(as) 118–100
INEID probe	FAM-AATCCTCACAACTACTAAGTTGTCA-TAMRA	(as) 40–21
Synthetic control probe	HEX-CTCCACCTCTTTCTTACCACGA-BHQ1	(s) 66–88
Viral control probe	Cy5-GTGCAGCTGTTTCTTAGCACGAAGAT-BHQ1	(s) 66–92
Assay 2		
FLI-WNF5-F	GGGCCTTCTGGTCGTGTTTC	(s) 3558–3576
FLI-WNF6-R	GATCTTGGCYGTCCACCTC†	(as) 3621–3603
FLI-WNF-Probe	FAM-CCACCCAGGAGGTCCTTCGCAA-TAMRA	(s) 3581–3602

* Genome position refers to WNV complete genome NY99 (GenBank accession no. AF196835). (s) = sense orientation; (as) = antisense orientation,

† Y: C/T.

Information database) using Vector NTI Advance primer design software 10.0.[†] The first WNV-specific amplicon site was identified in the highly conserved 5'-untranslated region (UTR) segment (assay 1) and the second in the nonstructural NS2A region (assay 2). The corresponding primers and probes are listed in Table 1. Probes were labeled at the 5' end with the FAM reporter dye and at the 3' end with the quencher dye TAMRA.[§]

The real-time PCR assays were performed with a commercial system^h and kit^e in a total volume of 25 µl. For these assays, 5 µl of RNA, 20.0 pmol of each primer, and 2.5 pmol of each probe were used. An in vitro transcribed green fluorescent protein gene fragment was used as IC-RNA extraction control (2.5 pmol of IC-RNA– specific primers and 1.5 pmol of probe), as described above.¹² Cycling times were as follows: 1 cycle at 50°C for 30 min (reverse transcription), 95°C for 15 min, and 42 cycles at 95°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec. Pure water and a no-template control were used as negative controls in every run. Furthermore, WNV samples were comparatively analyzed using 3 previously published RT-qPCR assays (assays 3–5). Primer and probes of assay 3¹⁶ targeted the WNV genome position 1160–1229, assay 4¹⁷ the genome position 10–153, and assay 5¹³ the genome position 10597–10672.

For the quantification of WNV copy numbers, a synthetic external calibrator was designed (Fig. 1A), which comprised the target sequences of assay 1 and used the same primers and probe. However, the binding site of the second probe (position 66–88) was mutated in this external calibrator by mirror inversions of 6 guanine/cytosine sequences to create a new specific target site, which can be detected only by a synthetic control probe (Fig. 1A; Table 1). An additional viral control probe (position 66–92) detected the corresponding original viral sequence. Using primer INEID f1 and INEID r1 in combination with the synthetic control probe allows the unambiguous detection of the synthetic control RNA, whereas the use of viral control probe together with assay 1 primers (INEID_f1 and INEID_r1) confirms viral RNA. This construct was amplified using vector PCR 2.1[§]; the vector was linearized with *HindIII* and in-vitro transcribed using a

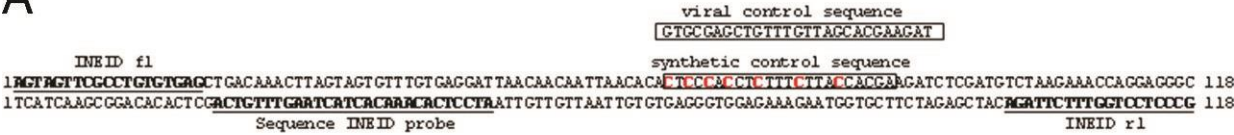
commercial in vitro transcription system.ⁱ The obtained transcripts were purified^e (without carrier RNA), and the amounts of RNA were estimated.^f

In order to determine the minimal copy number, an external calibrator was developed based on the WNV assay 1 target sequence and an authentic target site composition for the probe (Fig. 1A). Serial dilutions of this calibrator yielded copy numbers ranging from 1 to 2.5×10^7 copies/µl and were used to establish a calibration curve depicting mean threshold cycle (Ct) values plotted against the RNA copy numbers (Fig. 1B). The calibrator sequences were amplified in parallel using assay 1 primers and probes as well as assay 1 primers and synthetic control probe. The curve showed a linear progression for the WNV probe assay and a PCR efficiency of 1.0 and displayed (for synthetic control probe– derived assay) a PCR efficiency of 0.97. Both standard curves exhibited a correlation coefficient of .099. Based on this calibration curve, it can be concluded that the 2 new qRT-PCR assays are capable of detecting 2–4 RNA copies of WNV lineage 1 and 2 strains. The analytical sensitivity, as determined by the synthetic calibrator, is based on extractions from pure solutions. No inhibition was observed when this calibrator (100 RNA copies) was extracted from horse plasma (data not shown). In general, the impact of different matrix backgrounds (such as plasma or cell culture medium) was revealed by the internal control (present in all reactions and set to give a Ct value of 25–27).¹²

The 2 novel RT-qPCR assays for WNV were compared to 3 previously published assays (assays 3–5) with regard to their sensitivity and amplification efficiency. To compare the analytical sensitivity of all assays, Ct values were normalized by comparing them with a positive RNA control (WNV strain NY99), which was added to each run. The experimental limit of detection (LOD) was set at the serial dilution corresponding to 3 copies of external calibrator, based on the finding that 3 copies per PCR reaction were detected to 100% (Fig. 1C). This is in accordance to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments Guidelines.⁶

The analytical sensitivity comparison was carried out by determining duplicate Ct values of 10-fold dilutions (10^{-1} – 10^{-7}) of WNV strain NY99-derived and Uganda-derived

A



B

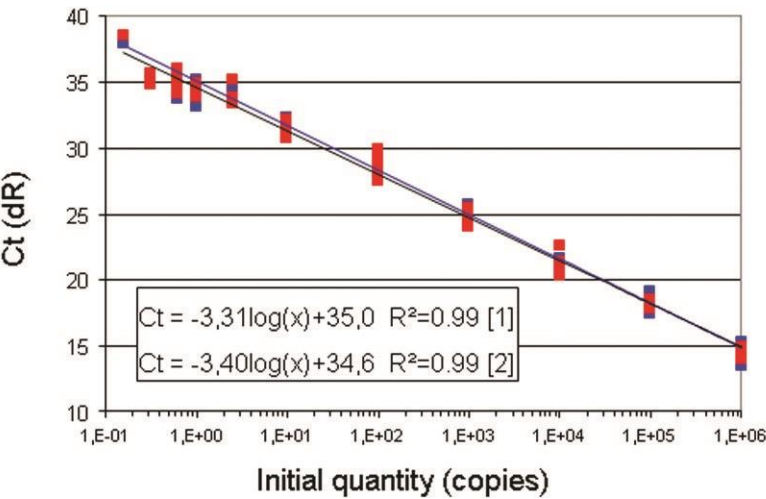


Figure 1. Synthetic calibrator. **A**, composition of the synthetic external calibrator sequence: Cytosine and guanine exchanges of the synthetic calibrator sequence are designated in red. The corresponding viral sequence is shown above. **B**, standard curve of external calibrator real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR). West Nile virus qRT-PCR protocol was carried out using primers and probe of assay 1 (blue line) or synthetic control probe (red line). Ten-fold dilutions of synthetic RNA were subjected to RT-qPCR. Threshold cycle (Ct) values in at least 3 replicates are plotted against synthetic RNA copies on a log scale. The regression function and correlation coefficient (R^2) of RT-qPCR with probe of assay 1 [1] and synthetic control probe [2] are inserted into the plot. **C**, The limit of detection (LOD) was determined by using quantified serial dilutions in at least 3 replicates of external calibrator RNA, which sets the end-point limit of detection for 3 copies per reaction. (copies) = copies per (PCR) reaction.

RNA (Table 2). The LOD for WNV strain NY99 was between 10^{-7} and 10^{-6} dilution, which corresponded to 1.2– 12.2 copies per reaction. All assays detected the 10^{-6} dilution of NY99 RNA with similar analytical sensitivity:

Assay 1 displayed a mean Ct of 34.3 ± 0 , assay 2 a mean Ct of 34.5 ± 0 , assay 4 a mean Ct of 34.5 ± 0.8 , assay 5 a mean Ct of 34.8 ± 0.2 , and assay 3 a mean Ct of 34.9 ± 0.2 . In addition, assays 1, 3, and 5 were able to detect 10^{-7} RNA

Table 2. Comparative real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis of 5 different RT-qPCR primer probes: assay 1 (current study), assay 2 (current study), assay 3,¹⁶ assay 4,¹⁷ and assay 5.¹³

Sample		Threshold cycle										Copies per reaction
Virus	Dilution	Assay 1		Assay 2		Assay 3*		Assay 4		Assay 5		
NY99	10-1	15.9	16.5	17.2	17.2	16.8	16.9	16.2	16.0	16.9	16.6	1,222,811.0
	10-2	19.8	19.9	20.8	20.8	20.5	20.5	19.5	19.1	20.3	20.3	122,281.0
	10-3	23.9	24.1	24.5	24.4	24.3	24.3	23.9	23.7	24.2	24.2	12,228.0
	10-4	28.0	28.0	28.0	28.0	28.0	28.0	28.0	28.0	27.9	28.1	1,223.0
	10-5	31.5	31.7	31.4	31.6	31.6	31.6	30.6	30.5	31.5	31.5	122.0
	10-6	34.3	34.3	34.5	34.5	34.8	35.0	35.1	33.9	34.7	35.0	12.2
	10-7	36.3	35.9	>42	>42	40.0	38.7	No Ct†	No Ct	37.6	38.5	1.2
Uganda	10-1	18.0	18.3	19.2	19.1			19.0	19.0	21.5	21.6	668,139.0
	10-2	22.0	22.0	22.8	22.7			22.4	22.5	25.0	25.1	66,814.0
	10-3	25.1	25.1	26.2	26.2			26.0	25.9	28.4	28.7	6,681.0
	10-4	28.6	28.9	29.7	29.6			28.6	28.7	32.1	32.4	668.0
	10-5	31.6	32.2	32.5	32.7			31.8	31.8	35.2	35.9	66.8
	10-6	34.9	35.0	37.7	36.9			No Ct	39.1	40.3	37.1	6.7
	10-7	36.7	36.1	No Ct	.42			>42	41.1	No Ct	No Ct	0.7

* Assay 3 is only designed for the detection of lineage 1 strains.

† Ct = threshold cycle.

Table 3. Sensitivity of the real-time quantitative reverse transcription polymerase chain reaction assays for *West Nile virus* (WNV) strains NY99, Uganda 956, Kunjin, and Sarafend.*

WNV strain	Dilution	Virus titer (TCID ₅₀ /ml)	Assay 1		Assay 2		Copies per reaction
NY99 10E8.5	10E-1	10E7.5	16.7	17.0	17.4	15.5	1,074,525.0
	10E-2	10E6.6	20.5	21.2	20.8	20.5	107,453.0
	10E-3	10E5.5	24.1	24.6	24.6	24.6	10,745.0
	10E-4	10E4.5	27.0	27.9	27.8	28.0	1,075.0
	10E-5	10E3.5	30.0	31.8	30.5	32.2	107.0
	10E-6	10E2.5	33.5	34.1	33.8	33.5	11.0
	10E-7	10E1.5	37.7	35.8	39.9	37.0	1.1
	10E-8	10E0.5	No Ct	>42	No Ct	No Ct	0.1
	10E-9	10E-0.5	>42	40.5	>42	>42	0.0
Uganda 10E5.0	10E-1	10E4.0	18.5	18.5	19.7	20.0	298,427.0
	10E-2	10E3.0	22.3	22.7	23.9	23.8	29,843.0
	10E-3	10E2.0	26.4	26.2	27.9	28.0	2,984.0
	10E-4	10E1.0	29.5	30.21	30.9	30.91	298.0
	10E-5	10E0.0	32.65	33.06	33.4	33.62	30.0
	10E-6	10E-1.0	36.21	35.1	37.2	37.7	3.0
	10E-7	10E-2.0	No Ct	No Ct	>42	41.9	0.3
	10E-8	10E-3.0	No Ct	No Ct	No Ct	No Ct	0.03
	10E-9	10E-4.0	>42	No Ct	41.6	No Ct	0.0
Kunjin 10E6.9	10E-1	10E5.9	17.8	18.1	17.3	17.8	462,655.0
	10E-2	10E4.9	22.0	21.2	21.1	20.7	46,265.0
	10E-3	10E3.9	25.7	25.7	24.7	23.9	4,627.0
	10E-4	10E2.9	29.4	29.7	28.6	28.3	463.0
	10E-5	10E1.9	37.3	35.5	33.0	33.6	46.0
	10E-6	10E0.9	37.2	36.0	35.8	36.6	4.6
	10E-7	10E-0.1	39.2	39.0	40.4	39.1	0.46
	10E-8	10E-1.1	40.9	38.7	40.4	>42	0.05
	10E-9	10E-2.1	40.1	>42	No Ct	No Ct	0.0
Sarafend 10E8.8	10E-1	10E7.8	19.5	19.9	18.1	18.1	320,931.0
	10E-2	10E6.8	23.0	22.0	21.4	21.6	32,093.0
	10E-3	10E5.8	26.5	26.3	25.3	25.4	3,209.0
	10E-4	10E4.8	28.1	28.4	27.9	27.8	321.0
	10E-5	10E3.8	33.7	33.6	32.3	32.6	32.0
	10E-6	10E2.8	35.1	34.4	34.6	34.2	3.2
	10E-7	10E1.8	36.7	36.1	36.2	35.9	0.3
	10E-8	10E0.8	>42	No Ct	42.0	>42	0.0
	10E-9	10E-0.2	>42	No Ct	No Ct	No Ct	0.0

* TCID₅₀/ml = 50% tissue culture infective doses per milliliter; Ct = threshold cycle.

dilution (mean Ct of 36.1 ± 0.3 for assay 1, a mean Ct of 38.1 ± 0.6 for assay 5, and a mean Ct of 39.3 ± 1.0 for assay 3). The LOD for WNV strain Uganda was between 10^{-7} and 10^{-6} RNA dilution, which corresponded to 0.7–6.7 copies per reaction. The 10^{-6} Uganda RNA dilution was detected by assays 1 and 2 and harbored mean Ct values of 35.0 ± 0.1 and 37.3 ± 0.6 , respectively. Assay 5 displayed a mean Ct of 38.6 ± 2.3 , while assay 4 exhibited only a single Ct of 39.1. In addition, assay 1 detected 10^{-7} Uganda RNA dilution, with a mean Ct of 36.4 ± 0.4 . Assay 3 is designed only for WNV lineage 1 strains. The results demonstrate that the new assays 1 and 2 are suitable for sensitive detection of both WNV lineages.

In addition, both of the new RT-qPCR assays were evaluated by testing the four WNV strains for which the infectivity titers in the tissue culture supernatants were determined: WNV strain NY99 contained $10^{8.5}$ TCID₅₀/ml, Uganda B956 contained $10^{5.0}$ TCID₅₀/ml, Kunjin

contained $10^{6.9}$ TCID₅₀/ml, and Sarafend contained $10^{8.8}$ TCID₅₀/ml. The analytical sensitivity of each assay was evaluated by comparing duplicate Ct values of RNA extracts of 10-fold supernatant dilutions (Table 3). Both assays detected WNV strains of lineage 1 (NY99, Kunjin) and lineage 2 (Uganda B956, Sarafend) with an extremely high and comparable sensitivity. The LOD for WNV strain NY99 was in the range of $10^{1.5}$ – $10^{2.5}$ TCID₅₀/ml, which corresponded to 1.1–11 copies per reaction. Assay 1 exhibited a similar sensitivity compared to assay 2, yielding a mean Ct of 36.75 ± 1.34 versus 38.45 ± 2.05 for a solution containing $10^{1.5}$ TCID₅₀/ml. The LOD for WNV strain Uganda was at a virus titer of 10^{-1} TCID₅₀/ml, which corresponded to 3 copies per reaction. Again, both assays 1 displayed a comparable sensitivity (mean Ct of 35.66 ± 0.78 vs. 37.45 ± 0.35). For WNV strains Kunjin and Sarafend, the LODs were at a titer of $10^{0.9}$ TCID₅₀/ml (4.6 copies per reaction) and $10^{2.8}$ TCID₅₀/ml (3.2 copies per reaction). Even for

Table 4. Specificity of real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) assays for *Murray Valley encephalitis virus* (MVEV), *Japanese encephalitis virus* (JEV), *Yellow fever virus* (YFV), and *Tick-borne encephalitis virus* (TBEV).*

Species	Threshold cycle		
	pan-Flavi (SYBR Green)	Assay 1	Assay 2
WNV (NY99)	22	23	21.7
WNV (Kunjin)	22.7	27.8	24.9
WNV (Sarafend)	23.3	25.8	23.3
WNV (Uganda)	24.2	25.2	26.9
MVEV	23.6	33.8	No Ct
JEV	23.9	No Ct	28.4
YFV	23.9	No Ct	No Ct
TBEV	24.3	No Ct	No Ct

* WNV 5 West Nile virus; Ct = threshold cycle. *Flavivirus* species-derived RNA is calibrated with a SYBR Green real-time RT-PCR based on pan-Flavi primer set.

these strains, assays 1 and 2 were compatible in terms of sensitivity for the detection of WNV strains Sarafend (mean Ct of 34.74 ± 0.52 vs. 34.40 ± 0.28) and Kunjin (mean Ct of 36.63 ± 0.86 vs. 36.2 ± 0.56).

The 2 new RT-qPCR assays were eventually tested with other species of genus *Flavivirus*, such as YFV, JEV, TBEV, and MVEV, in order to verify the specificity of the detection. For the adjustment of comparable RNA amounts for the different species, a SYBR Green real-time RT-PCR based on the pan-Flavi primer set (as previously published⁷) was used, which targeted the conserved NS5 region of this genus. As shown in Table 4, assay 1 detected MVEV (Ct = 33.8), and assay 2 detected JEV (Ct = 28.4). All other analyzed species were not detected. Assay 2 may therefore be used under certain conditions for additional detection of JEV. Thus, the combined use of assays 1 and 2 helps one to avoid false-positive results for MVEV or JEV and enables the specific detection of WNV.

In summary, 2 new RT-qPCR assays were established for the detection of WNV genome, with excellent analytical sensitivity and good specificity. The high sensitivity of both assays for WNV lineage 2 allows the efficient detection of these emerging infections in Europe (i.e., Hungary since 2004⁸ and Austria since 2008 [Department for Environment, Food Rural Affairs: 2008, West Nile virus: Austria. Reference: VITT 1200/WNV-Austria. Available at <http://www.defra.gov.uk/foodfarm/farmanimal/diseases/monitoring/documents/wnv-austria.pdf>. Accessed on April 15, 2010]). The 2 assays can be used in parallel or sequentially to mutually reconfirm the obtained results, since the 2 primer sets anneal at different regions of the WNV genome (highly conserved 5'UTR segment [assay 1] and NS2A region [assay 2]). Moreover, a new kind of external calibrator was designed, which relies solely on mirror-inversion mutations that can be discerned from the original viral sequences by a calibrator-specific RNA probe.

Acknowledgements. This work was funded by the Federal Ministry of Education and Research in the "Research on Zoonotic Infectious Diseases" program.

Sources and manufacturers

- Kindly provided by M. Niedrig, Robert-Koch-Institute, Berlin, Germany.
- Kindly provided by A. Muellbacher, John Curtin School of Medical Research, Canberra, Australia.
- Health Protection Agency, Salisbury, United Kingdom.
- Kindly provided by F. T. Hufert and M. Weidmann, Institute for Virology, Göttingen, Germany.
- QIAamp®, QuantiTect®, RNeasy® MinElute®, Qiagen GmbH, Hilden, Germany.
- Quant-IT™, Invitrogen Corp., Carlsbad, CA.
- Eurofins MWG Operon, Ebersberg, Germany.
- Mx3000P® QPCR system, Stratagene Inc., La Jolla, CA.
- Riboprobe®, Promega Corp., Madison, WI.

References

- Bakonyi T, Hubálek Z, Rudolf I, Nowotny N: 2005, Novel flavivirus or new lineage of West Nile virus, central Europe. *Emerg Infect Dis* 11:225–231.
- Bakonyi T, Ivanics E, Erdélyi K, et al.: 2006, Lineage 1 and 2 strains of encephalitic West Nile virus, central Europe. *Emerg Infect Dis* 12:618–623.
- Bondre VP, Jadhav RS, Mishra AC, et al.: 2007, West Nile virus isolates from India: evidence for a distinct genetic lineage. *J Gen Virol* 88:875–884.
- Braut AC: 2009, Changing patterns of West Nile virus transmission: altered vector competence and host susceptibility. *Vet Res* 40:43.
- Burt FJ, Grobbelaar AA, Leman PA, et al.: 2002, Phylogenetic relationships of southern African West Nile virus isolates. *Emerg Infect Dis* 8:820–826.
- Bustin SA, Benes V, Garson JA, et al.: 2009, The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622.
- Chao DY, Davis BS, Chang GJ: 2007, Development of multiplex real-time reverse transcriptase PCR assays for detecting eight medically important flaviviruses in mosquitoes. *J Clin Microbiol* 45:584–589.
- Erdélyi K, Ursu K, Ferenczi E, et al.: 2007, Clinical and pathologic features of lineage 2 West Nile virus infections in birds of prey in Hungary. *Vector Borne Zoonotic Dis* 7: 181–188.
- Esteves A, Almeida AP, Galão RP, et al.: 2005, West Nile virus in Southern Portugal, 2004. *Vector Borne Zoonotic Dis* 5:410–413.
- Glaser A: 2004, West Nile virus and North America: an unfolding story. *Rev Sci Tech* 23:557–568.
- Heinz FX, Collett MS, Purcell RH, et al.: 1999, Family: Flaviviridae. In: *Virus taxonomy: classification and nomenclature of viruses*. Seventh report of the International Committee on Taxonomy of Viruses, ed. Van Regenmortel MHV, Fauquet CM, Bishop DHL, et al., pp. 859–878. Academic Press, San Diego, CA.
- Hoffmann B, Depner K, Schirrmeier H, Beer M: 2006, A universal heterologous control system for duplex real-time RT-PCR assays used in a detection system for pestiviruses. *J Virol Methods* 136:200–209.

13. Jiménez-Clavero MA, Agüero M, Rojo G, Gómez-Tejedor C: 2006, A new fluorogenic real-time RT-PCR assay for detection of lineage 1 and lineage 2 West Nile viruses. *J Vet Diagn Invest* 18:459–462.
14. Jourdain E, Schuffenecker I, Korimbocus J, et al.: 2007, West Nile virus in wild resident birds, Southern France, 2004. *Vector Borne Zoonotic Dis* 7:448–452.
15. Kärber G: 1931, Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche [Contribution to collective treatment of pharmacological serial tests]. *Arch Exp Pathol Pharmacol* 162:480–483. In German.
16. Lanciotti RS, Kerst AJ, Nasci RS, et al.: 2000, Rapid detection of West Nile virus from human clinical specimens, field-collected mosquitoes, and avian samples by a TaqMan reverse transcriptase–PCR assay. *J Clin Microbiol* 38:4066–4071.
17. Linke S, Ellerbrok H, Niedrig M, et al.: 2007, Detection of West Nile virus lineages 1 and 2 by real-time PCR. *J Virol Methods* 146:355–358.
18. Macini P, Squintani G, Finarelli AC, et al.: 2008, Detection of West Nile virus infection in horses, Italy, September 2008. *Euro Surveill* 13(39):18990.
19. Marra PP, Griffing SM, McLean RG: 2003, West Nile virus and wildlife health. *Emerg Infect Dis* 9:898–899.
20. Petersen LR, Roehrig JT: 2001, West Nile virus: a reemerging global pathogen. *Emerg Infect Dis* 7:611–614.
21. Popovici F, Sarbu A, Nicolae O, et al.: 2008, West Nile fever in a patient in Romania, August 2008: case report. *Euro Surveill* 13(39):18989.
22. Rossini G, Cavrini F, Pierro A, et al.: 2008, First human case of West Nile virus neuro invasive infection in Italy, September 2008—case report. *Euro Surveill* 13(41):19002.
23. Scherret JH, Mackenzie JS, Hall RA, et al.: 2002, Phylogeny and molecular epidemiology of West Nile and Kunjin viruses. *Curr Top Microbiol Immunol* 267:373–390.
24. Smithburn KC, Hughes TP, Burke AW, Paul JH: 1940, Aneurotropic virus isolated from the blood of a native of Uganda. *Am J Trop Med Hyg* 20:471–492.
25. Weissenböck H, Hubálek Z, Bakonyi T, Nowotny N: 2010, Zoonotic mosquito-borne flaviviruses: worldwide presence of agents with proven pathogenicity and potential candidates of future emerging diseases. *Vet Microbiol* 140:271–280.

Chapter 3. Publication II

A Quantitative Real-Time RT-PCR Assay for the Detection of Venezuelan Equine Encephalitis Virus Utilizing a Universal Alphavirus Control RNA.

Ariel Vina-Rodriguez¹, Martin Eiden¹, Markus Keller, Winfried Hinrichs², and Martin H. Groschup¹.

¹ *Institute for Novel and Emerging Infectious Diseases, Friedrich-Loeffler-Institut, Greifswald, Insel Riems, Germany*

² *Department of Molecular Structural Biology, Institute for Biochemistry, University of Greifswald, Greifswald, Germany*

BioMed Research International

Volume 2016: 1–7, Article ID 8543204, 7 pages

<http://dx.doi.org/10.1155/2016/8543204>

Research Article

A Quantitative Real-Time RT-PCR Assay for the Detection of *Venezuelan equine encephalitis virus* Utilizing a Universal Alphavirus Control RNA

Ariel Vina-Rodriguez,¹ Martin Eiden,¹ Markus Keller,¹
Winfried Hinrichs,² and Martin H. Groschup¹

¹ Institute for Novel and Emerging Infectious Diseases, Friedrich-Loeffler-Institut, Greifswald, Insel Riems, Germany

² Department of Molecular Structural Biology, Institute for Biochemistry, University of Greifswald, Greifswald, Germany

Correspondence should be addressed to Martin H. Groschup; martin.groschup@fli.de

Received 4 August 2016; Revised 27 September 2016; Accepted 25 October 2016

Academic Editor: Charles Spencer

Copyright © 2016 Ariel Vina-Rodriguez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Venezuelan equine encephalitis virus (VEEV) is an *Alphavirus* from the family *Togaviridae* that causes epizootic outbreaks in equids and humans in Central and South America. So far, most studies use conventional reverse transcriptase PCR assays for the detection of the different VEEV subtypes. Here we describe the development of a TaqMan quantitative real-time reverse transcriptase PCR assay for the specific detection and quantitation of all VEEV subtypes which uses in parallel a universal equine encephalitis virus control RNA carrying target sequences of the three equine encephalitis viruses. The control RNA was used to generate standard curves for the calculation of copy numbers of viral genome of *Eastern equine encephalitis virus* (EEEV), *Western equine encephalitis virus* (WEEV), and VEEV. The new assay provides a reliable high-throughput method for the detection and quantitation of VEEV RNA in clinical and field samples and allows a rapid differentiation from potentially cocirculating EEEV and WEEV strains. The capability to detect all known VEEV variants was experimentally demonstrated and makes this assay suitable especially for the surveillance of VEEV.

1. Introduction

Western equine encephalitis virus (WEEV), *Eastern equine encephalitis virus* (EEEV), and *Venezuelan equine encephalitis virus* (VEEV) are arthropod-borne (arbo) viruses of the genus *Alphavirus* of the virus family *Togaviridae*. To date these viruses are restricted to the Americas but due to worldwide travelling and trade they might be introduced also to other parts of the world in the future. All three equine encephalitis viruses are classified as Category B agent by the Centers for Disease Control and Prevention, Atlanta (<https://emergency.cdc.gov/agent/agentlist-category.asp>). They are transmitted by sanguivorous mosquitoes within bird (WEEV, EEEV, and epizootic VEEV (epizootic strains)) or rodent populations (VEEV, enzootic strains), respectively. Infections in reservoir hosts do not lead to obvious clinical signs. However, severe diseases can occur when equines and humans are

infected with epizootic subtypes by biting mosquitoes. In the last decade 662 equine cases with 302 fatalities of VEE were reported to the OIE and 75,000–100,000 human cases with more than 300 fatalities were counted in the most recent outbreak in Venezuela and Colombia in 1995 [1]. For some epizootic VEEV strains a productive replication followed by successful intra- and interspecies transmission cycles was observed in horses and humans [2–6].

In the time period 2007 to 2012 a total of 1926 EEEV and 3 WEEV associated equine cases were diagnosed in the United States (<http://www.oie.int/>). In the time period 1964 to 2009 the CDC registered 639 human WEEV and 260 EEEV cases with only few fatalities. It was shown recently in an experimental animal model that all three viruses are transmissible by aerosols [7–9].

In general, PCR-diagnostic for emerging viruses address different questions: for the monitoring during an epidemic

TABLE 1: Primers and probes selected for equine encephalitis virus-specific quantitative reverse transcription polymerase chain reaction.

Target	Primer or probe	Sequence (5'→ 3')	Genome position	Reference
<i>Eastern equine encephalomyelitis virus</i> (EEEV)	EEE9391	ACACCGCACCCCTGATTTTACA	9391–9411 (s)	[10]
	EEE9459c	CTTCCAAGTGACCTGGTCGTC	9459–9439 (as)	
	EEE.9414probe	FAM-TGCACCCGGACCATCCGACCT-TAMRA	9414–9434 (s)	
<i>Western equine encephalomyelitis virus</i> (WEEV)	WEE10,248	CTGAAAGTCGGCCTGCGTAT	10,248–10,267 (s)	[10]
	WEE 10,314c	CGCCATTGACGAACGTATCC	10,314–10,295 (as)	
	WEE 10,271probe	FAM-ATACGGCAATACCACCGCGCACC-TAMRA	10,271–10,293 (s)	
<i>Venezuelan equine encephalomyelitis virus</i> (VEEV) and synthetic calibrator	AlphaVIR966F	TCCATGCTAATGCGYAGAGCGTTTTTCGCA	151–178 (s)	Modified [14]
	AlphaVIR966R	TGGCGCACTTCCAATGTCHAGGAT	248–225 (as)	
	INEID-VEEV probe	FAM-TGATCGARACGGAGGTRGAMCCATCC-TAMRA	193–218 (s)	This study
	VEEV-Coprobe	VIC-CTCCGTTCAATAC-MGB-NFQ*	180–192 (s)	This study

The synthetic calibrator RNA is specifically detected by the VEEV-Coprobe in combination with the AlphaVIR966F and AlphaVIR966R primers. Y, H, R, and M are designed for degenerative bases, where Y = C/T, H = A/C/T, R = A/G, and M = A/C. Modifications compared to the original sequence as well as novel sequences were indicated in italic font. *MGB: minor groove binder; NFQ: Nonfluorescent quencher.

or for the confirmation of infections from a known source, a PCR specific for some virus variant may be sufficient. In contrast to prevent the introduction of the virus into a corresponding region or country, it is necessary the use of an assay with an experimentally proven capability to detect every known virus variant. According to this several conventional RT-PCRs were developed to qualitatively detect VEEV, EEEV, and WEEV genome sequences and real-time reverse transcriptase PCRs (RT-qPCR) for EEEV and WEEV were published [10–13]. However, no RT-qPCR assay was available at the date of our study for the specific detection of different VEEV subtypes. A VEEV diagnosis is presently confirmed mostly by conventional RT-PCR using broad-range primer pairs covering the whole genus *Alphavirus* followed by subsequent amplicon sequencing [14]. Recent publications experimentally demonstrated RT-qPCR assays for detection of the VEEV vaccine strain TC-83 but without proven experimental demonstration of the assay's sensitivity and efficiency regarding other VEEV subtypes [15, 16]. In this study we are introducing a general purpose, rapid, one-step quantitative RT-qPCR assay for the sensitive and specific detection of all VEEV subtypes in combination with an internal calibrator construct which in turn can be used in the quantification of the three equine encephalitis viruses.

2. Materials and Methods

2.1. Primer Design. Multiple sequence alignments of VEEV sequences were performed using Vector NTI Advanced v.10 (Invitrogen, Carlsbad, CA, USA) and MEGA Software [17] to reveal primers, as well as a probe. For this purpose, a total of 33 VEEV sequences were retrieved from the GenBank database. Published broad-range primers, which target the nsP1 region of

Alphaviruses and previously used within a conventional RT-PCR protocol [14], were modified by the insertion of a degenerated base in each of the forward and the reverse primer and complemented with a FAM- (6-carboxyfluorescein-) labelled probe, which specifically targets VEEV sequences (Table 1) and enables the application of a quantitative real-time RT-PCR protocol.

2.2. Quantitative Real-Time RT-PCR (RT-qPCR). RT-qPCR was carried out by using a commercial kit (QuantiTect RTPCR kit, Qiagen, Germany). After the reverse transcription (50°C for 30 minutes) the DNA was denatured (95°C for 15 min). Amplification cycles included denaturation (95°C for 15 sec), annealing (55°C for 30 sec), and elongation (72°C for 30 sec) steps. Ct values were determined by the CFX96 software (Bio-Rad, USA).

2.3. Synthetic Calibrator. To determine the copy number of viral genomes a synthetic calibrator was developed, which comprises a T7 RNA polymerase promoter and the target sequences for the RT-qPCRs of EEEV, WEEV, and VEEV (Figure 1(a)) cloned into the pCR2.1 vector (Eurofins MWG Operon, Germany). The EEEV and WEEV sequences include targets for primer and probes adopted unmodified from the literature [10] (Table 1), but the corresponding probe target sequences were placed on the complementary strand in order to generate a unique (different) amplicon sequence, discriminable from the original virus sequence yet maintaining the same nucleotide composition. In addition, within the VEEV target region the original virus sequence 5'-CTGGCTTCAAAAC-3' was changed to 5'-CTCCGTTCAATAC-3' in order to discriminate unambiguously the synthetic RNA from viral RNA and to exclude false positive signals in samples potentially contaminated with synthetic RNA. This specific synthetic RNA sequence section can be detected only by

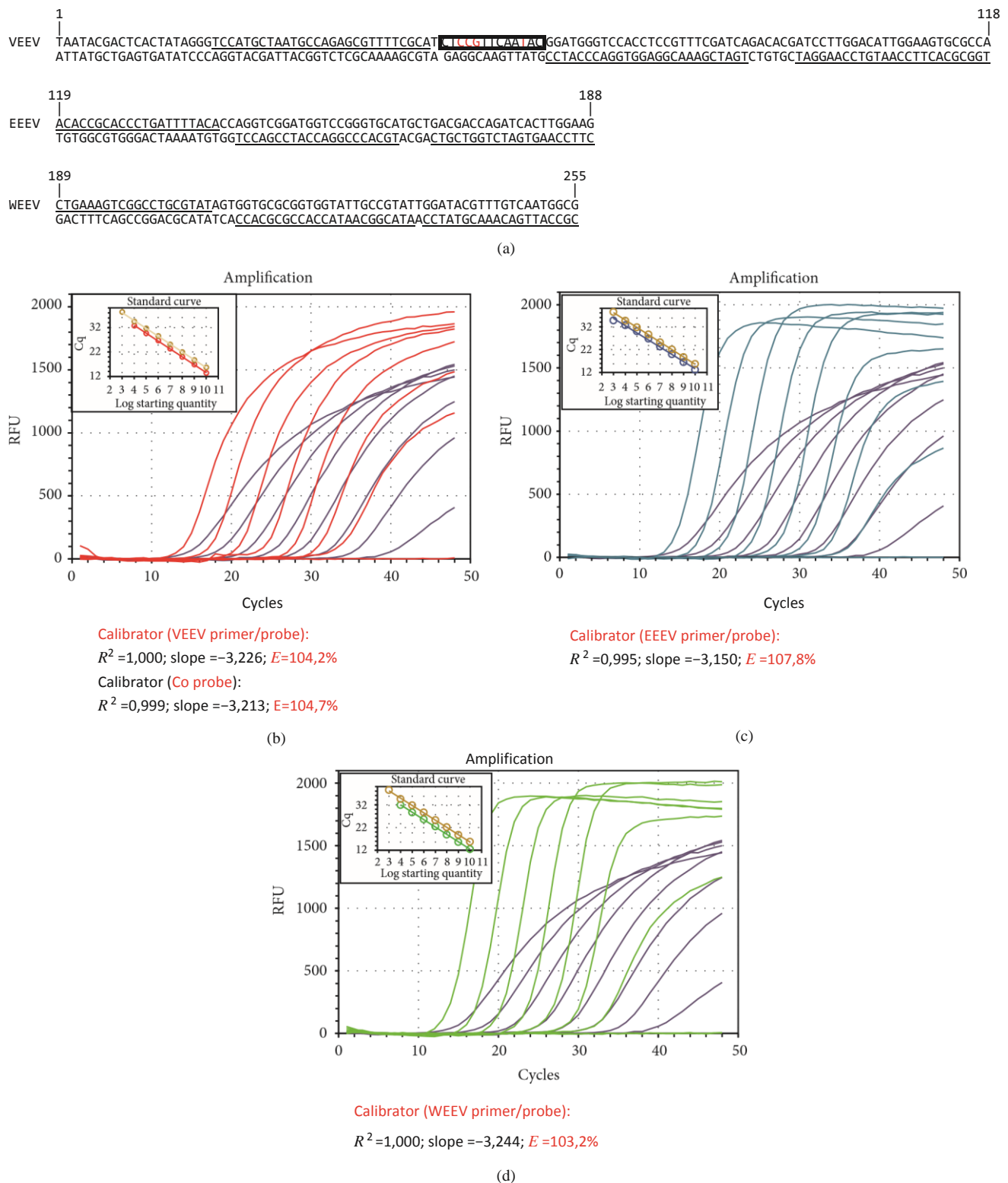


Figure 1: The nucleotide sequence of the synthetic construct used for calibration of the EEV-specific RT-qPCRs (a). The target sequences (underlined) for the specific RT-qPCRs are cloned into the vector pCR2.1-TOPO. Within the VEEV target region a modified sequence (framed) allows the differentiation of the synthetic calibrator from viral sequences. Nucleotide exchanges are indicated in red. Amplification curves of the RT-qPCRs specific for VEEV (b), EEEV (c), and WEEV (d) using the synthetic calibrator template. Amplification curve of the synthetic calibrator targeted with the control probe is indicated in olive, respectively. Standard curves (see enclosed boxed figures) were obtained by Ct values plotted against the log of the starting quantity. Calculated correlation coefficients (R^2), slopes, and amplification efficiencies (E) are depicted below the corresponding figures.

Table 2: Sensitivity of RT-qPCR assays for VEEV, WEEV, and EEEV strains and determination of copy number.

RNA	Dilution	Ct	Copies/ μ L
VEEV	10^{-2}	27,94	70200
	10^{-3}	30,69	8840
	10^{-4}	33,75	894
	10^{-5}	35,51	240
	10^{-6}	37,24	66
	10^{-7}	no Ct	0
WEEV	10^{-2}	21,76	2880000
	10^{-3}	25,13	216000
	10^{-4}	28,64	13780
	10^{-5}	31,1	2060
	10^{-6}	32,38	736
	10^{-7}	no Ct	0
EEEV	10^{-3}	29,67	12760
	10^{-4}	32,55	1212
	10^{-5}	34,22	316
	10^{-6}	no Ct	0
	10^{-7}	no Ct	0

a control probe (Table 1, VEEV-Coprobe). The plasmid was linearized with *Xba*I and subsequently transcribed into RNA and the DNA degraded using the Riboprobe® Combination System, T3/T7 RNA Polymerase (Promega Corporation's, Madison, WI, USA), and the QIAamp Viral RNA Mini Kit (Qiagen) was used for RNA isolation (without carrier RNA). The RNA concentration was estimated with the QuantIt™ RNA Assay Kit, Broad-Range (Invitrogen). The copy number of the synthetic RNA was calculated from the RNA concentration and the molecular mass of the RNA transcript.

3. Results and Discussion

In the here presented study, we are introducing a rapid, sensitive, and reliable one-step quantitative RT-PCR assay for VEEV as well as a synthetic RNA construct which can be used as calibrator for the quantification of alpha virus associated equine encephalitis viruses. PCR was carried out on serial dilutions of the synthetic RNA in a one-step RT-qPCR and Ct values were eventually plotted proportionally to the logarithm of the input copy numbers to produce standard quantitation curves. Negative controls were included in each run. The synthetic RNA was concurrently amplified using primer and probes for VEEV, WEEV, and EEEV, respectively (Figures 1(b)–1(d)). In addition the synthetic calibrator was amplified with primer for VEEV and detected by the VEEV-Coprobe. Both assays run independently in a single-plex format. All four standard curves exhibit a correlation coefficient >99% and an amplification efficiency of about 103–107% over a linear range of 10^2 to 10^{10} copies. Based on the corresponding standard curves the sensitivity and viral load for different EEEV strains in concurrent runs could be determined: the

limit of detection (LOD) corresponded to 66,2 copies per μ L (Ct = 37,24) for VEEV (TC-83 strain), at 736 copies (Ct = 32,38) for WEEV (McMillan strain), and at 316 copies per μ L (Ct = 34,21) for EEEV (New Jersey strain) (Table 2). Since standard curves form part of every run, the copy number from each analysed sample can be determined. Therefore samples with copy numbers above 1 copy per μ L (the theoretical detection limit) are considered to be positive. VEEV specific primer/probe combination did not detect other equine encephalomyelitis viruses (EEEV, WEEV) (Figure 2) or closely related species (Chikungunya virus, Sindbis virus, and Ross River virus; data not shown).

To further assess the performance of the VEEV specific RT-qPCR we used 15 synthetic RNA constructs (sVEEV) encompassing the target region and representing 10 different VEEV subtypes (Figure 3(a)). This included all combinations of observed mutations in the primers and probe target regions. sVEEV were designed as oligonucleotides with a 5' T7 RNA polymerase promoter sequence and were transcribed *in vitro* as aforementioned. All VEEV subtypes were successfully detected by the novel RT-qPCR assay with a suitable sensitivity and high performance as demonstrated by linear standard curves over 5 logs (Figure 3(b)). R^2 values and slope indicate good precision and high efficiency (Table 3). To evaluate the effect of nucleotide changes to the PCR amplification efficiency we applied the relative threshold cycle (RTC) method, which refers to mean Ct-differences (mean Δ Ct) of the corresponding sVEEV template compared to the unmodified template [18]. The data indicate that most nucleotide changes exhibited only small or moderate reduction of the RTC efficiency. Only sVEEV-14, representing one variant of

TABLE 3: Relative threshold cycle (RTC) amplification efficiencies of synthetic VEEV (sVEEV) RNA constructs.

Template	sVEEV -1	sVEEV -2	sVEEV -3	sVEEV -4	sVEEV -5	sVEEV -6	sVEEV -7	sVEEV -8	sVEEV -9	sVEEV -10	sVEEV -11	sVEEV -12	sVEEV -13	sVEEV -14	sVEEV -15
Dilution	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct	Ct
10 ⁻⁴	17,3	20,6	19,4	19,6	17,4	21,1	18,6	17,8	18,3	18,5	20,9	19,6	20,2	24,6	21,5
10 ⁻⁵	21,4	24,4	23,4	23,7	21,3	25,0	22,3	21,3	21,7	22,2	24,5	23,1	24,1	28,2	25,0
10 ⁻⁶	25,2	28,3	27,1	27,3	25,2	28,7	26,5	25,3	25,4	25,8	28,0	27,1	27,8	31,7	28,6
10 ⁻⁷	28,9	31,8	30,8	31,2	29,0	32,4	30,2	28,9	28,7	29,3	31,6	30,9	31,9	35,5	32,4
10 ⁻⁸	31,8	34,6	34,2	34,7	32,7	36,1	33,7	32,2	32,2	32,8	35,3	34,9	35,6	39,1	36,1
Slope R ²	-2,99 0,9969	-2,91 0,9959	-3,05 0,9992	-3,1 0,9995	-3,16 0,9998	-3,08 0,9998	-3,16 0,9992	-2,98 0,9993	-2,85 0,9999	-2,92 0,9999	-2,95 1	-3,17 0,9997	-3,18 0,9999	-2,99 0,9999	-3,00 0,9998
Mean ΔCt	0,0	-3,0	-2,1	-2,4	-0,2	-3,7	-1,4	-0,2	-0,3	-0,8	-3,2	-2,2	-3,0	-6,9	-3,8
Mean RTC	1,0	0,12	0,24	0,19	0,87	0,08	0,39	0,89	0,80	0,57	0,11	0,22	0,12	0,01	0,07

ΔCt is calculated as mean difference of corresponding Ct values compared to unmodified reference template sVEEV-1 across all template dilutions. RTC is calculated according $RTC = 2^{\Delta Ct}$.

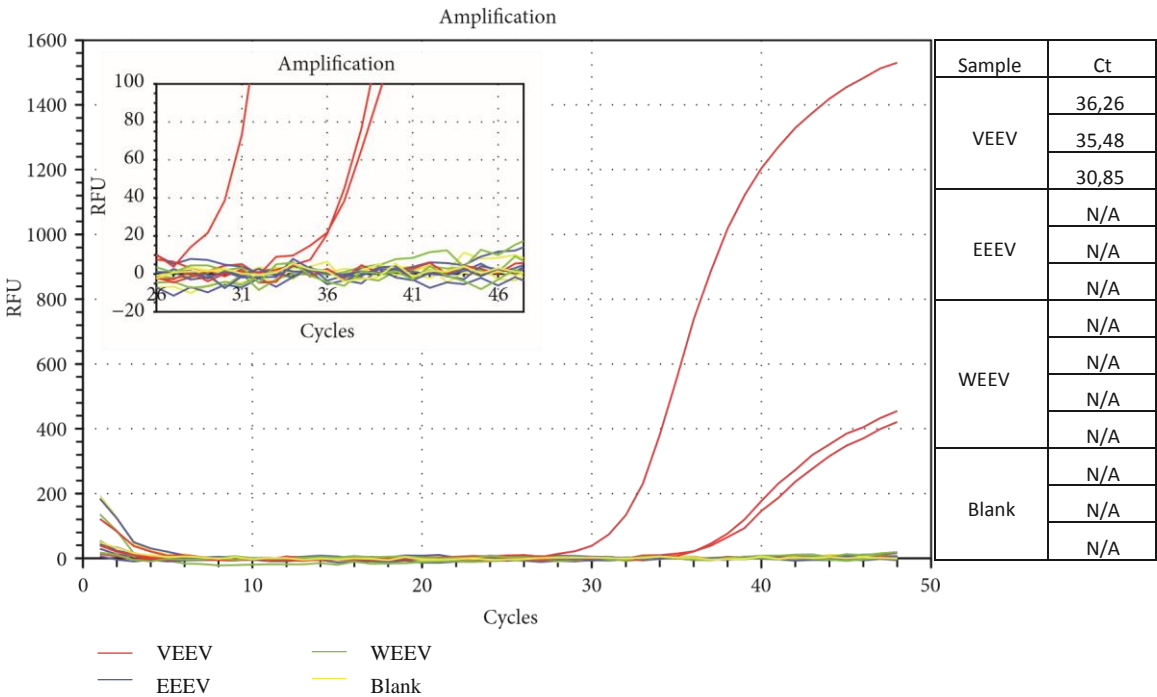


FIGURE 2: Specificity of the VEEV specific primer-probe combination. Specific amplification of VEEV derived RNA (red) by RT-qPCR. No amplification of EEEV (blue) and WEEV (green) derived RNA. Insert shows the boxed region at higher magnification.

subtype VI, showed a stronger decline in RTC efficiency which is probably caused by 13 nucleotide exchanges compared to the reference template. In summary this assay can be used whenever a sensitive and high-throughput detection or quantification of VEEV RNA is needed, for example, for confirmation of virus presence in patients, during infection experiments or large screening of field probes. But it is particularly useful when a proven application for the detection of all known VEEV variants is required, for example, to prevent

the introduction of any virus variant into a so far virus free region or country.

4. Conclusions

We report here the first experimental evidence of a quantitative real-time RT-PCR assay for the sensitive and specific detection of all known VEEV subtypes or sequence variants. The synthetic calibrator RNA allows the determination of

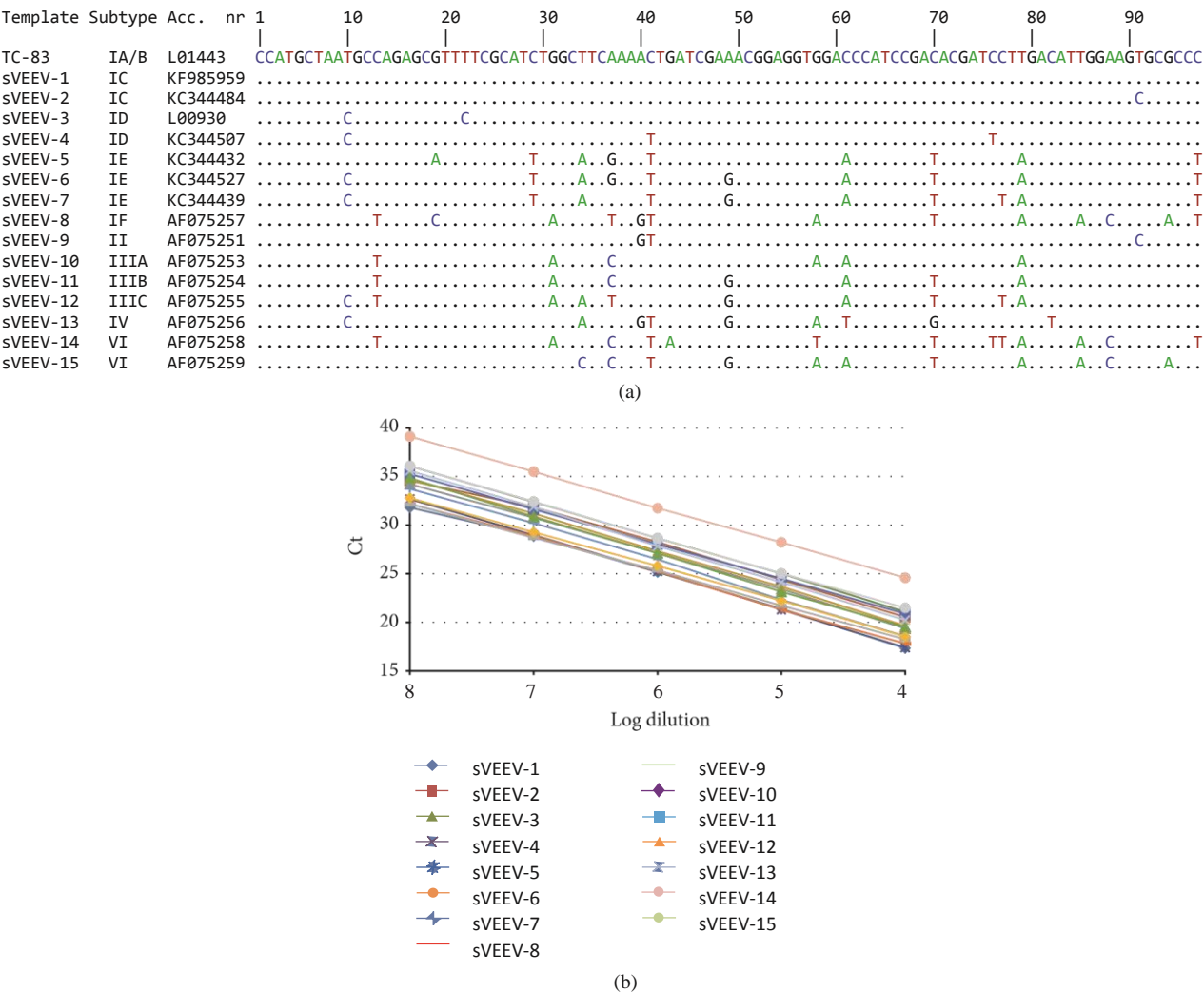


FIGURE 3: Comparison of the consensus sequences of different VEEV subtypes. (a) Sequences of synthetic RNA constructs (sVEEV) encompass the target region of the VEEV specific RT-qPCR. Nucleotides with mismatch to the reference sequence are indicated. (b) Standard curves of serially diluted sVEEV were obtained by Ct values plotted against the log of diluted template.

viral genome equivalents of VEEV as well as WEEV and EEEV by a one-step RT-qPCR reaction.

Competing Interests

The authors declare that they have no competing interests.

References

[1] S. Paessler and M. Pfeffer, "Togaviruses causing encephalitis," in *Encyclopedia of Virology*, B. W. J. Mahy and M. H. V. Van Regenmortel, Eds., pp. 76–82, Elsevier, Oxford, UK, 3rd edition, 2008.

[2] T. E. Walton and M. A. Grayson, "Venezuelan equine encephalitis," in *The Arboviruses: Epidemiology and Ecology*, T. P. Monath, Ed., vol. 4, pp. 203–231, CRC Press, Boca Raton, Fla, USA, 1988.

[3] S. C. Weaver and A. D. T. Barrett, "Transmission cycles, host range, evolution and emergence of arboviral disease," *Nature Reviews Microbiology*, vol. 2, no. 10, pp. 789–801, 2004.

[4] S. C. Weaver, M. Anishchenko, R. Bowen et al., "Genetic determinants of Venezuelan equine encephalitis emergence," in *Emergence and Control of Zoonotic Viral Encephalitides*, C. H. Calisher and D. E. Griffin, Eds., vol. 18 of *Archives of Virology, Supplementa*, pp. 43–64, Springer, Berlin, Germany, 2004.

[5] O. M. Suarez and G. H. Bergold, "Investigations of an outbreak of Venezuelan equine encephalitis in towns of eastern Venezuela," *American Journal of Tropical Medicine and Hygiene*, vol. 17, no. 6, pp. 875–880, 1968.

[6] R. E. Kissling, "The arthropod-borne viruses of man and other animals," *Annual Review of Microbiology*, vol. 14, pp. 261–282, 1960.

[7] D. S. Reed, M. G. Lackemeyer, N. L. Garza et al., "Severe encephalitis in cynomolgus macaques exposed to aerosolized eastern equine encephalitis virus," *Journal of Infectious Diseases*, vol. 196, no. 3, pp. 441–450, 2007.

[8] D. S. Reed, T. Larsen, L. J. Sullivan et al., "Aerosol exposure to western equine encephalitis virus causes fever and encephalitis in cynomolgus macaques," *Journal of Infectious Diseases*, vol. 192, no. 7, pp. 1173–1182, 2005.

[9] D. S. Reed, C. M. Lind, L. J. Sullivan, W. D. Pratt, and M. D. Parker, "Aerosol infection of cynomolgus macaques with enzootic strains of Venezuelan equine encephalitis viruses," *Journal of Infectious Diseases*, vol. 189, no. 6, pp. 1013–1017, 2004.

- [10] A. J. Lambert, D. A. Martin, and R. S. Lanciotti, "Detection of North American eastern and western equine encephalitis viruses by nucleic acid amplification assays," *Journal of Clinical Microbiology*, vol. 41, no. 1, pp. 379–385, 2003.
- [11] B. Linssen, R. M. Kinney, P. Aguilar et al., "Development of reverse transcription-PCR assays specific for detection of equine encephalitis viruses," *Journal of Clinical Microbiology*, vol. 38, no. 4, pp. 1527–1535, 2000.
- [12] D.J.Johnson,E.N.Ostlund,andB.J.Schmitt,"Nested multiplex RT-PCR for detection and differentiation of West Nile virus and eastern equine encephalomyelitis virus in brain tissues," *Journal of Veterinary Diagnostic Investigation*, vol. 15, no. 5, pp. 488–493, 2003.
- [13] K. Grywna, B. Kupfer, M. Panning et al., "Detection of all species of the genus *Alphavirus* by reverse transcription-PCR with diagnostic sensitivity," *Journal of Clinical Microbiology*, vol. 48, no. 9, pp. 3386–3387, 2010.
- [14] M. W. Eshoo, C. A. Whitehouse, S. T. Zoll et al., "Direct broadrange detection of alphaviruses in mosquito extracts," *Virology*, vol. 368, no. 2, pp. 286–295, 2007.
- [15] S. Qian, B. He, Z. Tu, H. Guo, and C. Tu, "Establishment of a one-step real-time RT-PCR method for the detection of Venezuelan equine encephalitis virus," *Bing Du Xue Bao*, vol. 31, no. 2, pp. 107–113, 2015 (Chinese).
- [16] J. G. Julander, R. Skirpstunas, V. Siddharthan et al., "C3H/HeN mouse model for the evaluation of antiviral agents for the treatment of Venezuelan equine encephalitis virus infection," *Antiviral Research*, vol. 78, no. 3, pp. 230–241, 2008.
- [17] K. Tamura, J. Dudley, M. Nei, and S. Kumar, "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0," *Molecular Biology and Evolution*, vol. 24, no. 8, pp. 1596– 1599, 2007.
- [18] J. A. Sikorsky, D. A. Primerano, T. W. Fenger, and J. Denvir, "Effect of DNA damage on PCR amplification efficiency with the relative threshold cycle method," *Biochemical and Biophysical Research Communications*, vol. 323, no. 3, pp. 823–830, 2004.

Chapter 4. Publication III

Ngari Virus in Goats during Rift Valley Fever Outbreak, Mauritania, 2010

Martin Eiden , **Ariel Vina-Rodriguez**, Bezeid O. El Mamy, Katia Isselmou, Ute Ziegler, Dirk Höper, Susanne Jäckel, Anne Balkema-Buschmann, Hermann Unger, Baba Doumbia, Martin H. Groschup

Emerging Infectious Diseases 20 (12): 2174–76.

doi:10.3201/eid2012.140787.

Copyright: *Emerging Infectious Diseases* is published by the Centers for Disease Control and Prevention, a U.S. Government agency. Therefore, materials published in *Emerging Infectious Diseases*, including text, figures, tables, and photographs are in the public domain and can be reprinted or used without permission with [proper citation](#).

Ngari Virus in Goats during Rift Valley Fever Outbreak, Mauritania, 2010

To the Editor: Ngari virus (NRIV) is a single-stranded RNA virus belonging to the family *Bunyaviridae*, genus *Orthobunyavirus*. The genome comprises 3 segments, the small (S), medium (M), and large (L) segments, which encode the nucleocapsid (N) protein, the 2 glycoproteins Gn and Gc, and the RNA-dependent RNA-polymerase, respectively. Sequence analysis showed that NRIV is a reassortant between Bunyamwera virus (BUNV) and Batai virus (BATV), both from the genus *Orthobunyavirus*. S and L segments derived from BUNV, and the M segment derived from BATV (1,2). NRIV is more virulent than BUNV and BATV and is associated with hemorrhagic fever. NRIV was first isolated from *Aedes simpsoni* mosquitoes in 1979 and from humans in 1993, both in Senegal (3). During 1997 and 1998, humans were affected with hemorrhagic fever diseases in Kenya and Somalia that were caused by Rift Valley fever virus (RVFV) and by NRIV (2,4).

In 2010, during an ongoing RVFV outbreak in Mauritania, we collected

163 serum samples (62 from camels, 8 from cattle, and 93 from small ruminants) (5). RVFV RNA was isolated from serum samples as described previously (5). Further molecular testing of the samples was conducted by a SYBRGreen-based real-time reverse transcription PCR (RT-PCR) adapted from a conventional RT-PCR and based on generic primers (bun_group_forw 5'-CTGCTAACAC-CAGCAGTACTTTTGAC-3' and bun_group_rev 5'-TGGAGGGTAAG-ACCATCGTCAGGAAC-3') that target a 250-nt sequence of the S segment of Bunyamwera serogroup members (6). Real-time RT-PCR was performed in a CFX 96 real-time PCR system (Bio-Rad, Hercules, CA, USA) by using 5 µL RNA with a QuantiTect SYBR Green RT-PCR Kit (QIAGEN, Hilden Germany) in a final volume of 25 µL. Cycling conditions included RT at 50°C for 30 min and 95°C for 15 min, followed by amplification with 44 cycles of 95°C for 15 s, 55°C for 25 s, 72°C for 30 s, and 77°C for 5 s. A melting curve analysis was then performed starting with 95°C for 60 s, and a temperature gradient was conducted from 68°C to 94°C in increments of 0.2°C.

Of the 163 serum samples tested, 2 samples from goats resulted in a positive signal with cycle thresholds of 23 (sample 51) and 28 (sample 65), respectively. Both samples showed similar melting peaks at ≈78.2°C and shared the identical partial nucleotide sequence of the S segment. The sequence belongs to the Bunyamwera serogroup, but the short partial sequence was not sufficient for accurate virus determination and identification. For this reason, both serum samples were used to inoculate cell monolayers of Vero E6 cells that were assayed for virus replication. Only sample 51 displayed a cytopathic effect after 72 h and was further analyzed. We isolated the viral RNA from cell culture with TRIzol reagent

(Invitrogen, Carlsbad, CA, USA) and used it to prepare a sequencing library according to a recently published protocol (7) but using Illumina adaptors (Illumina, San Diego, CA, USA). We sequenced the resulting library using the Illumina MiSeq instrument with v2 chemistry.

We recovered full-length genome sequences of the S, M, and L segments of the virus and deposited them in GenBank (accession nos. KJ716848–716850). Phylogenetic analysis of complete genome sequences indicated that the virus belongs to the Ngari virus group and showed high homology to previous NRIV isolates in all 3 segments (Figure). As for all previous NRIV strains, the new isolate was highly similar to BUNV regarding the S and the L segment (Figure, panels A, C); the M segment was highly similar to BATV (Figure, panel B).

This evidence supports the extension of the range of NRIV infection to goats (complete sequences already had been derived from a human and from mosquitoes [8]) and demonstrates the occurrence of NRIV during the 2010 RVFV outbreak in Mauritania. We are aware of only 1 additional report of NRIV-infected sheep (in 1988), also in Mauritania, although no further characterization or isolation has been conducted (9). Both NRIV-positive samples were negative for RVFV RNA but positive for RVFV-specific IgG. In addition, sample 51 contained IgM against RVFV (5), indicating possible coinfection of RVFV and NRIV. Because both ELISAs rely on detection of antibodies against RVFV N protein, which is highly divergent to the deduced NRIV N sequence, cross-reactivity is highly unlikely but needs to be substantiated. Both samples originated from the Adrar region, which was the center of an unusual RVFV outbreak in Mauritania in 2010 (10).

The possible clinical importance to livestock and the circulation of NRIV among mosquitoes, livestock,

and humans needs to be clarified. No further information about clinical signs of sampled animals or reports of human NRIV cases is available. Because infection with both RVFV and NRIV induces hemorrhagic fe-

ver, affected humans also should be tested for NRIV infection. Further development of specific molecular and serologic diagnostic tools for NRIV should be pursued to obtain more information about NRIV distribution in

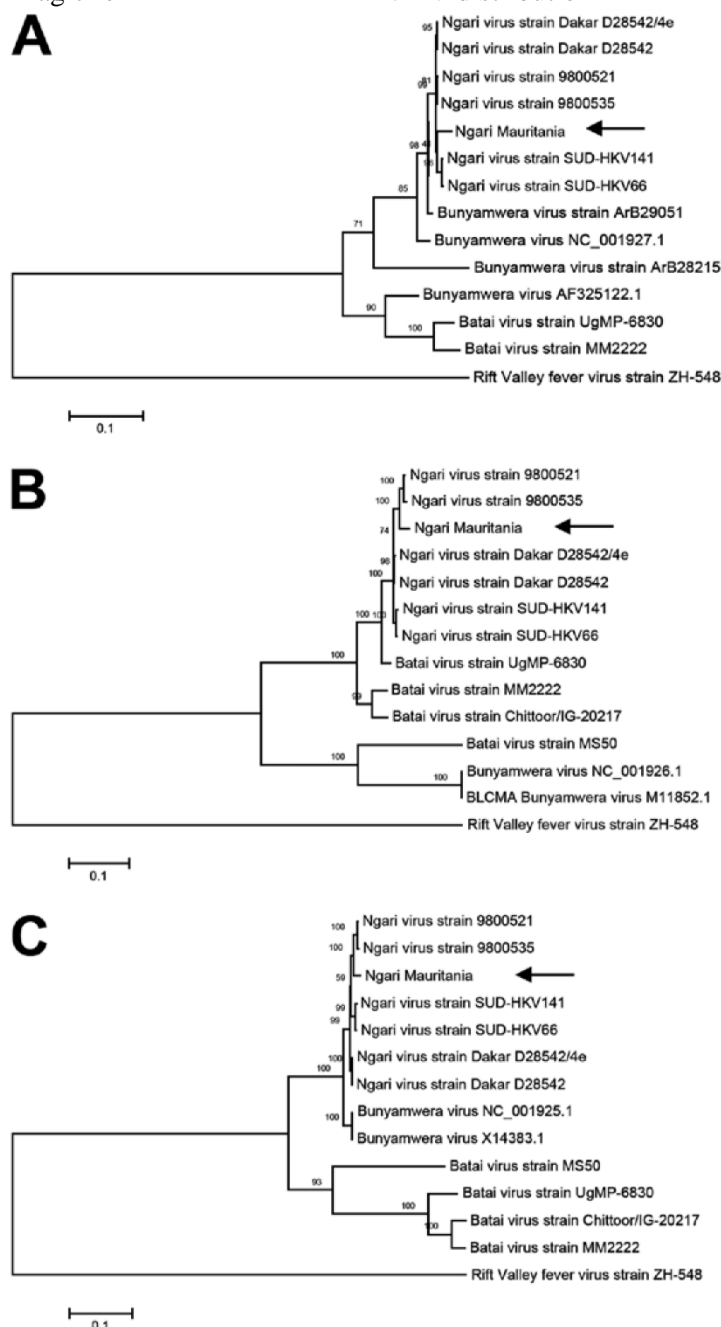


Figure. Phylogenetic tree of Ngari virus–derived A) small (975 bp), B) medium (4,507 bp), and C) large (6,887) segment sequences of Bunyamwera and Batai viruses compared with isolate obtained from a goat in Mauritania in 2010 (arrows). The tree was constructed on the basis of the nucleotide sequences of the 3 complete segments by using the neighborjoining method (1,000 bootstrap replications). The tree was rooted to the sequence of Rift Valley fever virus strain ZH-548. Scale bars indicate substitutions per nucleotide position.

humans and livestock in Mauritania and other African countries.

**Martin Eiden,
Ariel Vina-Rodriguez,
Bezeid O. El Mamy,
Katia Isselmou, Ute Ziegler,
Dirk Höper, Susanne Jäckel,
Anne Balkema-Buschmann,
Hermann Unger, Baba Doumbia,
and Martin H. Groschup**

Author affiliations: Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany (M. Eiden, A. Vina-Rodriguez, U. Ziegler, D. Höper, S. Jäckel, A. Balkema-Buschmann, M.H. Groschup); Centre National de l'Élevage et de Recherches Vétérinaires, Nouakchott, Mauritania (B.O. El Mamy, K. Isselmou, B. Doumbia); Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, Vienna, Austria (H. Unger); and Ministère du Développement Rural, Nouakchott (B. Doumbia)

DOI: <http://dx.doi.org/10.3201/eid2012.140787>

References

- Gerrard SR, Li L, Barrett AD, Nichol ST. Ngari virus is a Bunyamwera virus reassortant that can be associated with large outbreaks of hemorrhagic fever in Africa. *J Virol*. 2004;78:8922–6. <http://dx.doi.org/10.1128/JVI.78.16.8922-8926.2004>
- Briese T, Bird B, Kapoor V, Nichol ST, Lipkin WI. Batai and Ngari viruses: M segment reassortment and association with severe febrile disease outbreaks in East Africa. *J Virol*. 2006;80:5627–30. <http://dx.doi.org/10.1128/JVI.02448-05>
- Zeller HG, Diallo M, Angel G, TraoréLamizana M, Thonnon J, Digoutte JP, et al. Ngari virus (Bunyaviridae: Bunyavirus). First isolation from humans in Senegal, new mosquito vectors, its epidemiology [in French]. *Bull Soc Pathol Exot*. 1996;89:12–6.
- Bowen MD, Trappier SG, Sanchez AJ, Meyer RF, Goldsmith CS, Zaki SR, et al. RVF Task Force. A reassortant bunyavirus isolated from acute hemorrhagic fever cases in Kenya and Somalia. *Virology*. 2001;291:185–90. <http://dx.doi.org/10.1006/viro.2001.1201>
- Jäckel S, Eiden M, El Mamy BO, Isselmou K, Vina-Rodriguez A, Doumbia B, et al. Molecular and serological studies on the Rift Valley fever outbreak in Mauritania in 2010. *Transbound Emerg Dis*. 2013;60(Suppl 2):31–9. <http://dx.doi.org/10.1111/tbed.12142>
- Lambert AJ, Lanciotti RS. Consensus amplification and novel multiplex sequencing method for S segment species identification of 47 viruses of the *Orthobunyavirus*, *Phlebovirus*, and *Nairovirus* genera of the family *Bunyaviridae*. *J Clin Microbiol*. 2009;47:2398–404. <http://dx.doi.org/10.1128/JCM.00182-09>
- Steglich C, Grund C, Ramp K, Breithaupt A, Höper D, Keil G, et al. Chimeric Newcastle disease virus protects chickens against avian influenza in the presence of maternally derived NDV immunity. *PLoS ONE*. 2013;8:e72530. <http://dx.doi.org/10.1371/journal.pone.0072530>
- Groseth A, Weisend C, Ebihara H. Complete genome sequencing of mosquito and human isolates of Ngari virus. *J Virol*. 2012;86:13846–7. <http://dx.doi.org/10.1128/JVI.02644-12>
- Gonzalez JP. Activités du laboratoire d'écologie virale. Rapp sur la fonct tech de l'Institut Pasteur de Dakar. 1988;89:100–18.
- El Mamy AB, Baba MO, Barry Y, Isselmou K, Dia ML, El Kory MO, et al. Unexpected Rift Valley fever outbreak, northern Mauritania. *Emerg Infect Dis*. 2011;17:1894–6. <http://dx.doi.org/10.3201/eid1710.110397>

Address for correspondence: Martin H. Groschup, Friedrich-Loeffler-Institut, Südufer 10, 17493 Greifswald-Insel Riems, Germany; email: martin.groschup@fli.bund.de

Chapter 5. Publication IV

Hepatitis E Virus Genotype 3 Diversity: Phylogenetic Analysis and Presence of Subtype 3b in Wild Boar in Europe.

Ariel Vina-Rodriguez, Josephine Schlosser, Dietmar Becher, Volker Kaden, Martin H Groschup, and Martin Eiden.

Viruses 7 (5). Multidisciplinary Digital Publishing Institute: 2704–26. 2015.

doi:10.3390/v7052704.

Article

Hepatitis E Virus Genotype 3 Diversity: Phylogenetic Analysis and Presence of Subtype 3b in Wild Boar in Europe

Ariel Vina-Rodriguez ¹, Josephine Schlosser ¹, Dietmar Becher ², Volker Kaden ³,
Martin H. Groschup ¹ and Martin Eiden ^{1,*}

¹ Institute for Novel and Emerging Infectious Diseases, Friedrich-Loeffler-Institut, Südufer 10, 17493 Greifswald-Insel Riems, Germany; E-Mail: Ariel.Vina.Rodriguez@fli.bund.de (A.V.-R.); josephine.schlosser@fli.bund.de (J.S.); martin.groschup@fli.bund.de (M.H.G.)

² Micromun GmbH, Greifswald, Walther-Rathenau-Straße 49A, 17489 Greifswald, Germany; E-Mail: becher@micromun.de

³ Institute of Infectology, Friedrich-Loeffler-Institut, Südufer 10, 17493 Greifswald-Insel Riems, Germany; E-Mail: kaden@gmx.de

* Author to whom correspondence should be addressed; E-Mail: martin.eiden@fli.bund.de; Tel.: +49-38351-7-1182; Fax: +49-38351-7-1188.

Academic Editor: Andrew Mehle

Received: 14 April 2015 / Accepted: 20 May 2015 / Published: 22 May 2015

Abstract: An increasing number of indigenous cases of hepatitis E caused by genotype 3 viruses (HEV-3) have been diagnosed all around the world, particularly in industrialized countries. Hepatitis E is a zoonotic disease and accumulating evidence indicates that domestic pigs and wild boars are the main reservoirs of HEV-3. A detailed analysis of HEV-3 subtypes could help to determine the interplay of human activity, the role of animals as reservoirs and cross species transmission. Although complete genome sequences are most appropriate for HEV subtype determination, in most cases only partial genomic sequences are available. We therefore carried out a subtype classification analysis, which uses regions from all three open reading frames of the genome. Using this approach, more than 1000 published HEV-3 isolates were subtyped. Newly recovered HEV partial sequences from hunted German wild boars were also included in this study. These sequences were assigned to genotype 3 and clustered within subtype 3a, 3i and, unexpectedly, one of them within the subtype 3b, a first non-human report of this subtype in Europe.

Keywords: hepatitis E virus; HEV; genotype; subtype; phylogenetic analysis

1. Introduction

The Hepatitis E virus (HEV) is a causative agent of acute hepatitis in developing countries in Asia, Africa and Latin America where it is transmitted primarily via contaminated drinking water. Sporadic cases of HEV are reported in developed countries, partially imported by travelers from endemic areas, but there are also an increasing number of reports of autochthonous HEV infections. The transmission route of most of the autochthonous infections in industrialized countries still remains unclear. Reports of transfusion and transplant related infections exist [1–3], but accumulating evidence suggests that hepatitis E is a zoonotic disease with domestic pigs and wild boars being the main reservoirs. Moreover, the consumption of undercooked meat products poses a risk for HEV infection [4–7].

Studies on HEV RNA detection in animals revealed, that HEV is ubiquitous in domestic pigs and wild boars throughout Europe [8]. This includes the United Kingdom [9], France [10], Germany [11–15], Hungary [16], Italy [17,18], The Netherlands [19,20], Belgium [21], Spain [22], Slovenia [23], Czech Republic [24] and Sweden [25]. HEV infection in farmed pigs affects up to 80%–100% of the animals worldwide and usually occurs at the age of 2–4 months [26].

The virion is approximately 27–34 nm in diameter and most likely icosahedral. HEV has a positive sense single-stranded RNA genome of approximately 7.2 kb, which contains a short 5' untranslated region (UTR), a short 3' UTR and three open reading frames (ORF1, ORF2 and ORF3) [27]. The ORF1 encodes for viral non-structural proteins carrying domains with methyl transferase, helicase and replicase activities [28]. The ORF2 codes for the viral capsid protein of about 660 amino acids. The ORF3 is almost completely overlapped by the ORF2 and codes for a small phosphoprotein of about 114 amino acids, which is putatively responsible for the virion egress from infected cells [29].

A new proposed consensus for the HEV classification [30] divides the *Hepeviridae* family in two genera: *Orthohepevirus* and *Piscihepevirus*. The latter includes only isolates from cutthroat trout so far. The genus *Orthohepevirus* is further subdivided into four species: *Orthohepevirus A* with isolates from human, pig, wild boar, deer, mongoose, rabbit and camel, and *Orthohepevirus B*, C and D with avian and other mammal isolates. *Orthohepevirus A* is subsequently divided into at least six genotypes (HEV-1, HEV-2, etc.). HEV-1 and 2 include exclusively human HEV strains, whereas HEV-3 and 4 can also infect other animal species, particularly domestic pigs and wild boar. The separation in anthrotopropic (HEV-1 and -2) and enzoonotic (HEV-3 and -4) forms may have occurred more than 500 year ago [31].

HEV-1 is found in Asia and Africa [32], whereas HEV-2 was first isolated in Mexico [33] and later in Africa [34]. HEV-4 includes strains from sporadic human HEV cases in Asia [35]. HEV-3 was isolated initially from human cases in the USA [36] and has been detected in all continents including Europe [37].

Since 2001, 2703 human HEV cases [38] have been reported in Germany, which include an increasing number of non-travel associated autochthonous cases. The origin of infection remains unclear for most of the autochthonous cases, however, often the suspected HEV sources are domestic pigs and wild boars [39]. Similar increases are also reported in other European countries [40].

The role of different HEV-3 genetic variants in the evolution of the disease [1,7,39,41,42], the possibility of tracking the routes of infection and the influences of human activity on it [43–46] are currently under study. The direct comparison of isolates is still hampered by the limited number of complete genome (CG) sequences available. Due to this limitation, the subtyping scheme proposed by Lu *et al.* [32] has been commonly used and have been supported by epidemiological and/or statistical analysis [4,43,45,47,48] but questions have arisen, partially due to the lack of commonly accepted reference sequences for some subtypes [1]. In response to this, and to the increasing number of partial sequences, the subtyping of genotype 3 strains was actualized in order to provide an update of the subtyping scheme of HEV-3 and of the set of reference sequences. We carried out the classification of newly recovered HEV isolates from German wild boar and detected HEV subtype 3b strains for the first time in animals in Europe and, possibly, the first from a wild animal outside Japan.

2. Materials and Methods

2.1. Samples and RNA Extraction

Blood samples were collected from wild boar hunted in Mecklenburg-Western Pomerania during the seasons 1996/1997 (955 samples) and 2005/2006 (58 samples). Liver samples were collected in 2009/2010 from 134 animals hunted in the region of Greifswald and from another five from Western Pomerania. All samples were stored at –80 °C prior to their use. RNA was extracted with the RNeasy Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. A synthetic RNA (IC) was used as internal extraction control [49].

2.2. Primers and Probe Design

For primer and probe design, an alignment of 351 HEV sequences was constructed using the Vector NTI Advanced v.10 (Invitrogen, Carlsbad, CA, USA), BioEdit v.7.0.5.3 [50] and MEGA v6 [51] software. This alignment was manually curated using both the nucleotide and the deduced amino acid sequences. Very similar sequences were not included (more than 99% identity). HEV-1, -2, -3 and -4 genotypes were included (with preference to genotype 3), covering all subtypes, and including 131 CG (48 of them cited by Lu *et al.* [32]) and 65 German HEV sequences (the accession numbers are included in Supplementary Table S01). For genotyping and subtyping, four sets of nested degenerated primers were selected from this alignment, which target different regions of the genome. Previously published primers [14] were used to amplify an RNA-dependent RNA polymerase (RdRp) region. A novel diagnostic quantitative real-time RT-PCR assay (qRT-PCR) that targets ORF3 was also designed, which we already used in a recently published work [52]. Primer and probes used are listed in Table 1 (nucleotide positions refer to FJ705359, strain wbGER27, a German wild boar isolate [14]) and were included in the GenBank sequences entries.

2.3. PCR

The diagnostic/screening RT-qPCR was performed using the QuantiTec Probe RT-PCR kit (QIAGEN) in 25 µL reaction volume. In all reactions, the final concentration of each primer was 0.8 µM, and of the probe 0.1 µM if present. A volume of 5 µL of the RNA eluate was added. The reverse transcription (RT) was carried out at 50 °C for 30 min, followed by denaturation/activation at 95 °C for 15 min. DNA was amplified immediately with 45 cycles at 95 °C (10 s), 55 °C (25 s) and 72 °C (25 s).

Table 1. Primer and probes used in this study. All nucleotide positions refer to FJ705359 (strain wbGER27, a German wild boar isolate). Abbreviations: open reading frame (ORF), hyper variable region (HVR), RNA dependent RNA polymerase (RdRp).

Region Name and Internal Length (nt)	Primer Name	Position	Sequence	Step	Product Length (bp)
ORF1 -5' This study	HEV.ORF1_F1	33–58	CCCAYCAGTTYATWAAGGCTCCTGGC	RT-PCR	493
	HEV.ORF1_R1	497–525	TGCARDGARTANARRGCNAYNCCNGTCTC		
	HEV.ORF1_F2	98–126	AAYTCYGCCYTGGCGAATGCTGTGGTGGT	nested PCR	302
	HEV.ORF1_R2	377–399	CCVCGRGTNGGRGCRGWRTACCA		
HVR (for genotype 3) This study	HEV.HVR_F1	2069–2091	TTYTCYCCTGGGCAYMTYTGGGA	RT-PCR	401
	HEV.HVR_R1	2441–2469	TTAACCARCCARTCACARTCYGAYTCAAA		
	HEV.HVR_F2a	2135–2157	ACYTGGTCHACATCTGGYTTYTC	nested PCR	263 or 293
	HEV.HVR_F2b	2165–2184	TTYTCCCCYCCTGAGGCGGC		
RdRp 280 (4312–4591) (Johne <i>et al.</i> 2010)	HEV-cs	4181–4203	TCGCGCATCACMTTYTTCCARAA	RT-PCR	469
	HEV-cas	4628–4650	GCCATGTTCCAGACDGTCTTCCA		
	HEV-csn	4287–4311	GTGCTCTGTTTGCCCNCTGGTTYMG	nested PCR	330
	HEV-casn	4592–4617	CCAGGCTCACCRGARTGYTTCTTCCA		
ORF3 225 (5205–5429) This study	HEV.ORF3_F1	5126–5145	MGGKTRGAATGAATAACATG	RT-PCR	326 or 362
	HEV.ORF3_R2	5430–5451	GGCGCTGGGAYTGGTCRCGCCA		
	HEV.ORF3_R1	5467–5487	CAGYTGGGGYAGRTCGACGRC		
	HEV.ORF3_F2	5182–5204	GGGCTGTTCTGTTKYTGTYTCYTC	nested PCR	219 or 269
	HEV.ORF3_R2	5430–5451	GGCGCTGGGAYTGGTCRCGCCA		
	HEV.ORF3_R2a	5382–5401	CGAGGGCGAGCTCCAGCCCC		
modified Diagnostic-qPCR This study and (Schlosser <i>et al.</i> 2014)	HEV.Fa	5278–5294	GTGCCGCGCGGTGGTTTC	RT-qPCR	81
	HEV.Fb	5278–5296	GTGCCGCGCGGTGGTTTCTG		
	HEV.R	5340–5359	GCGAAGGGGTTGGTTGGATG		
	HEV.P	5300–5320	FAM-TGACMGGGTTGATTCTCAGCC-BHQ1		
ORF2 187 (6277–6488) This study	HEV.ORF2_F1	6205–6223	CDGCNACYCGBTTYATGAA	RT-PCR	393
	HEV.ORF2_R1a	6573–6598	GTKAGRGARAGCCAWAGYACATCATT		
	HEV.ORF2_R1b	6573–6598	GTRAGNGADAGCCACARRACATCATT		
	HEV.ORF2_F2	6276–6301	GCBYTHACNYTRTTYAACTTGCTGA	nested PCR	241
	HEV.ORF2_R2	6489–6517	TGYTCRTGYTGRTRTCTARTCYTG DAT		

The determination of the HEV RNA concentration was carried out using a standard curve according to a synthetic external calibrator. This calibrator encompassed the 81 bp sequence of the diagnostic qRT-PCR amplicon and included the T7 promoter sequence at the 5'-end for *in vitro* transcription. The RNA synthesis and DNA degradation were carried out by the Riboprobe® Combination System— T3/T7 RNA Polymerase (Promega Corporation's, Madison, WI, USA); and the QIAamp Viral RNA Mini Kit (QIAGEN) kit was used for RNA isolation (without carrier RNA). The RNA concentration was

estimated with the Quant-It™ RNA Assay Kit, Broad Range (Invitrogen) and confirmed by endpoint dilution PCR.

For genotyping, the initial RT-qPCR was performed with the QuantiTect SYBR Green RT-PCR kit (QIAGEN) in 25 µL reaction volume using 5 µL of the sample RNA. The thermal profile applied was: 30 min at 50 °C for RT, 15 min 95 °C denaturation/activation followed by 45 cycles of 95 °C for 10 s, 55 °C for 25 s, 72 °C for 25 s and 80 °C for 5 s (with fluorescence reading). A final dissociation curve generation step was also included. Two microliters of the resulting solution was added to 23 µL of the Maxima™ SYBR Green/ROX qPCR Master Mix kit (Fermentas, Canada) containing the primers and the PCR was carried out under similar conditions (45 cycles of: 95 °C for 10 s, 55 °C for 25 s, 72 °C for 25 s and 80 °C for 5 s with a final dissociation curve generation step).

2.4. Sequencing, Phylogenetic Analysis and Classification

RT-PCR or nested-PCR products were directly sequenced with the corresponding forward and reverse PCR primers using the BigDye Terminator v1.1 Cycle Sequencing Kit on the DNA sequencer “3130 Genetic Analyzer” (Applied Biosystems, Waltham, MA, USA).

The newly generated sequences were manually inserted in the multi-alignment previously used for primer design. This multi-alignment was updated with new HEV sequences (NCBI, 2014-12-15), up to more than 1400 sequences (Supplementary Table S01), mainly genotype 3 (near 1300 sequences). This included all HEV-3 sequences longer 1000 nt, all HEV NCBI nucleotide entries with the keyword “Germany”, the 135 sequences cited by Lu *et al.* [32] and other sequences from around the world, but particularly from other European countries. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model [53]. The trees with the highest log are shown. The percentage of trees (bootstrap values for 500 replicates for the first CG tree—and 100 replicates for all others trees) in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining (NJ) method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories). The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. Codon positions included were 1st + 2nd + 3rd. Evolutionary analyses were conducted in MEGA6 [51]. The alignment in FASTA format and the auxiliary worksheet for classification, selection and automatic labeling of sequences in MEGA are provided in Supplementary Files S01, S02 (updated versions are planned to be available from the authors). The sheet enables a quick selection of all the sequences spanning a given genomic region, which can be set in alignment coordinates (nt position) or by referencing to the sequences M73218 (Burma) or FJ705359 (wbGER27). Starting with the CG, and followed by the longest sequences, we built phylogenetic trees, and labeled each sequence as a “reference” to be used in subsequence classifications only if the tree reproduced the same topology as the tree for CG and if the clade was supported by bootstrap values of more than 70%. That is: HEV-3 sequences were labeled as “reference” only if the subtype was unambiguously determined. We assumed that all sequences from one strain represent the same genome sequence, and if one of them was labeled “reference”, all the others were also labeled.

3. Results

3.1. HEV RNA Detection

HEV RNA was detected in 32 out of 955 blood samples from 1996/97 and three out of 58 blood samples from 2005/2006, which suggests a virus prevalence of about 3.4% and 5.2%, respectively. In addition, HEV RNA was found in 14 out of 134 wild boar derived liver samples from the Greifswald region, giving a prevalence rate of about 10.4%. Finally, two wild boar liver samples (WS03-09 and WS05-09) from individual hunts were also positive. All HEV RNA positive samples were re-tested with the PCR for genotyping, and partial sequences from 12 animals could be recovered and subjected to phylogenetic analysis.

3.2. Phylogenetic Analyses

A reference phylogenetic tree was constructed based on 166 CG sequences, including eight German HEV isolates and 98 strains of genotype 3 (Figure 1a, which corresponds to Figure 4 in [32]). The hypervariable region (HVR) (2146–2358 nt) was excluded from this analysis. All nucleotide positions refer to sequence M73218. This tree confirmed a good separation of the HEV-3 from all other HEV genotypes. The sequences clustered into four monophyletic groups: “3jab”, “3chi”, “3feg” and “rabbit”. A detailed overview of the HEV-3 clade is shown in Figure 1b.

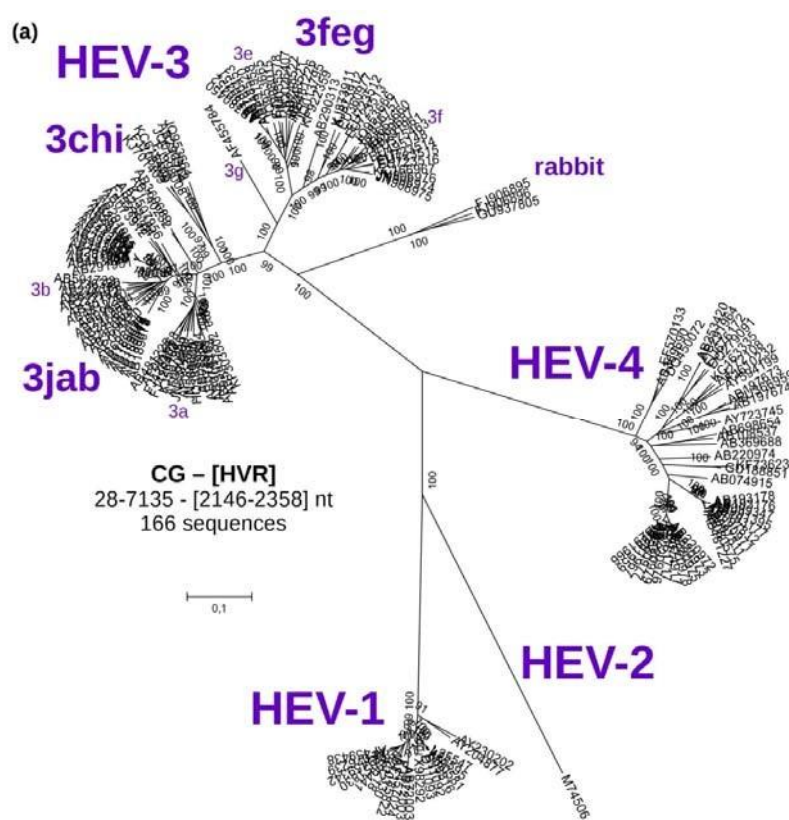


Figure 1. Cont.

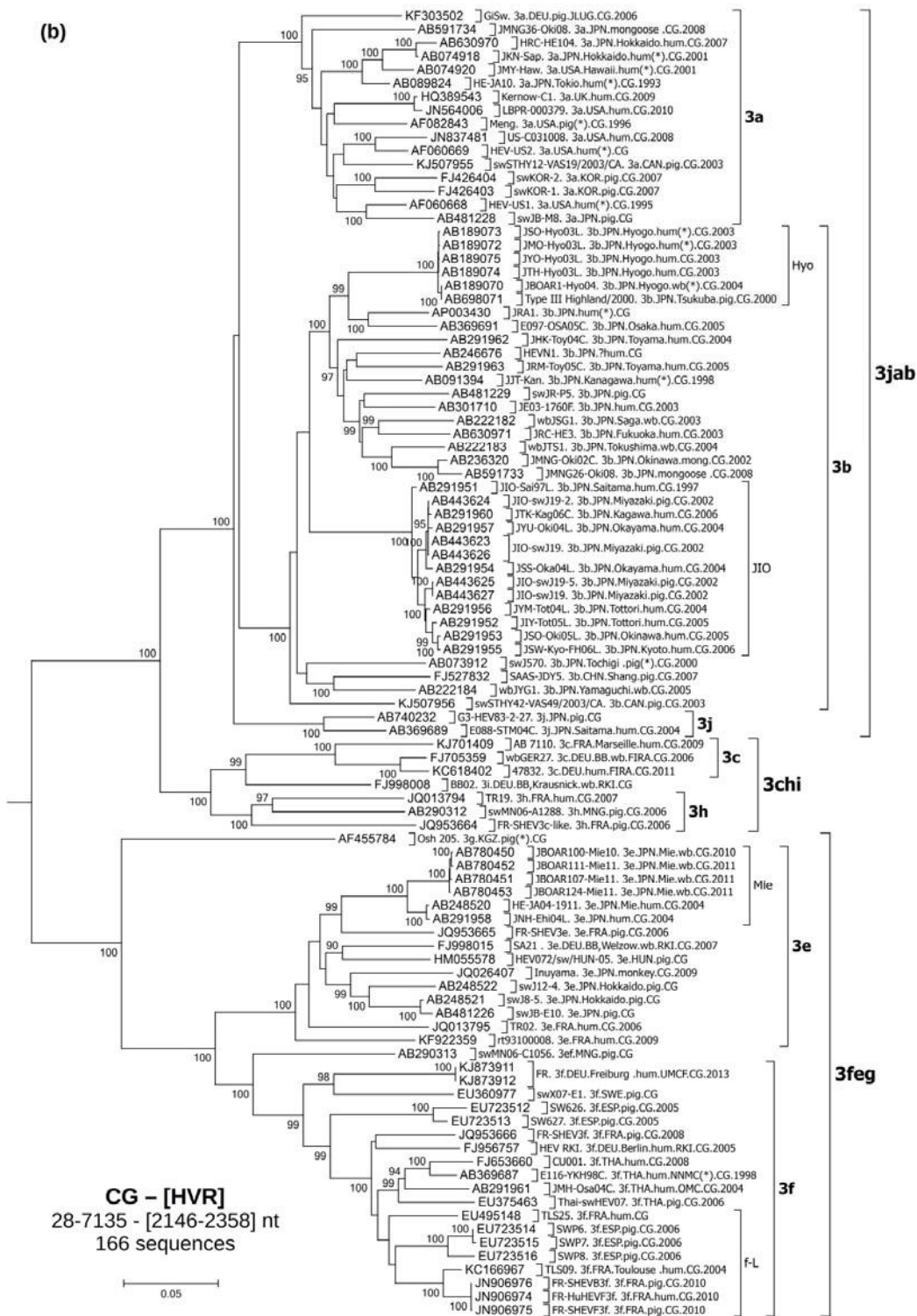


Figure 1. Molecular Phylogenetic analysis of 166 complete HEV genomes by Maximum Likelihood method based on the Kimura 2-parameter model. The percentage of trees (from 500 bootstrap replicates) in which the associated taxa clustered together is shown next to the branches when over 70%. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. A discrete Gamma distribution was used to model evolutionary rate differences among sites (five categories (+G, parameter = 0.5255)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated. There were a total of 6868 positions in the final dataset. Evolutionary analyses were conducted in MEGA6. (a) Global view of the unrooted tree, which corresponds to Figure 4 in Lu *et al.*

[32]; (b) Detailed view of the HEV-3 clade. (*)—references sequences cited by Lu; HRC-HE104—strain used in the HEV RNA WHO standard; wb—wild boar. See Supplementary Table S01 for more information.

The analysis was continued only with HEV-3, excluding the rabbit sequences, which form a well separated clade.

Comparisons amongst complete HEV-3 genomes display different levels of diversity within this genotype (Figure 2). The analysis of the frequency of corrected distances between sequences shows a possible separation around 0.14 substitutions per site. Sequences with lower differences belong to the same subtype. The graphic of the frequencies depicts two additional intermediary peaks, which are the basis for the definition of groups (3jab, 3chi and 3feg) and of major clades 3-I (3jab and 3chi) and 3-II (3feg) [54] (Figure 2).

The newly obtained tree (Figure 3a) for complete HEV-3 genomes further segregated the sequences in subtypes clades 3j, 3a, 3b, 3c, 3h, 3e and 3f. This classification was supported by bootstrap values of 99%–100%. Subtypes 3i and 3g were represented by individual isolates. Ninety-six of these sequences were marked as “reference”. Two strains were marked 3ef. German strains were grouped in the 3a, 3c, 3i, 3e and 3f subtypes. This Figure represents our best approximation to the true topology and displays the reference structure for all subsequent trees.

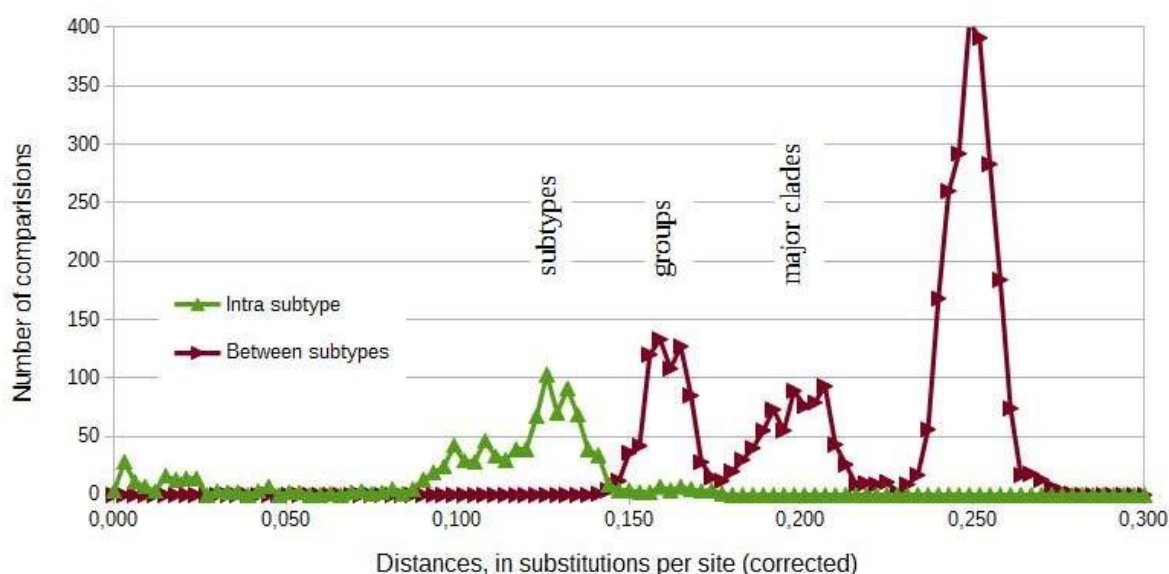
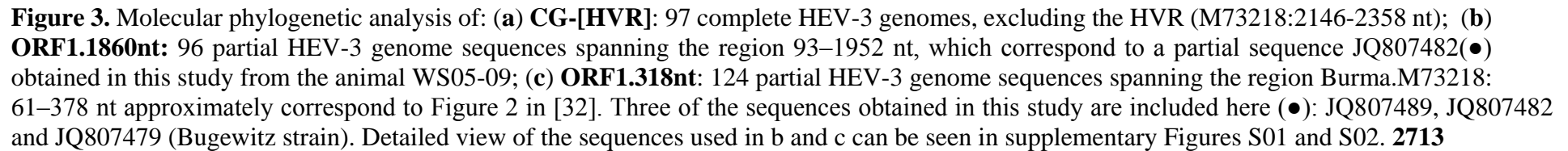


Figure 2. Frequency diagram of corrected distances between sequences from Figure 3a. It shows a possible separation of subtypes and suggests a possible definition of groups (3chi, 3jab and 3feg) and of major clades 3-I and 3-II (Norder, 2009). Pairwise distances were estimated using the Maximum Composite Likelihood (MCL) approach and grouped in intervals of 0.003. A discrete Gamma distribution was used to model evolutionary rate differences among sites (five categories). All positions with less than 95% site coverage were eliminated.



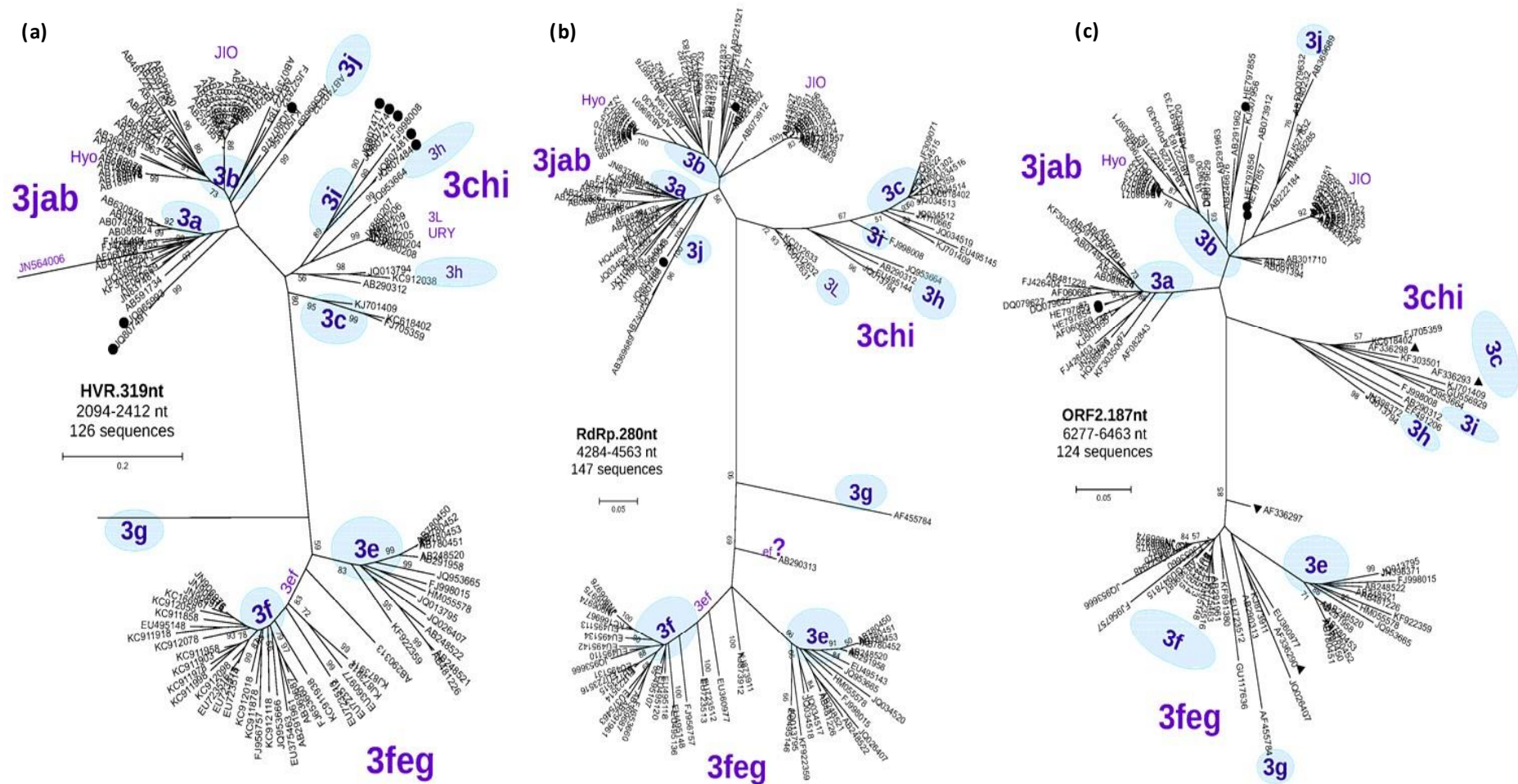


Figure 4. Molecular phylogenetic analysis of: **(a) HVR.319nt:** 126 partial HEV-3 genome sequences spanning the region 2094–2412 nt. The branches for sequences AF455784-3g and JN564006-3a are very large and have been truncated, including eight of the sequences obtained in this study (●); **(b) RdRp.280nt:** 147 partial HEV-3 genome sequences spanning the region 4284–4563 nt, including three of the sequences obtained in this study (●); **(c) ORF2.187nt:** 124 partial HEV-3 genome sequences spanning the region 6277–6463 nt, including five of the sequences obtained in this study (●), and stool pools from The Netherlands: 3c (▲) NLSW36 and NLSW105 and 3c+3f: (▼)NLSW20 and NLSW99 (see discussion). Detailed view of the sequences can be seen in supplementary Figures S07–S09.

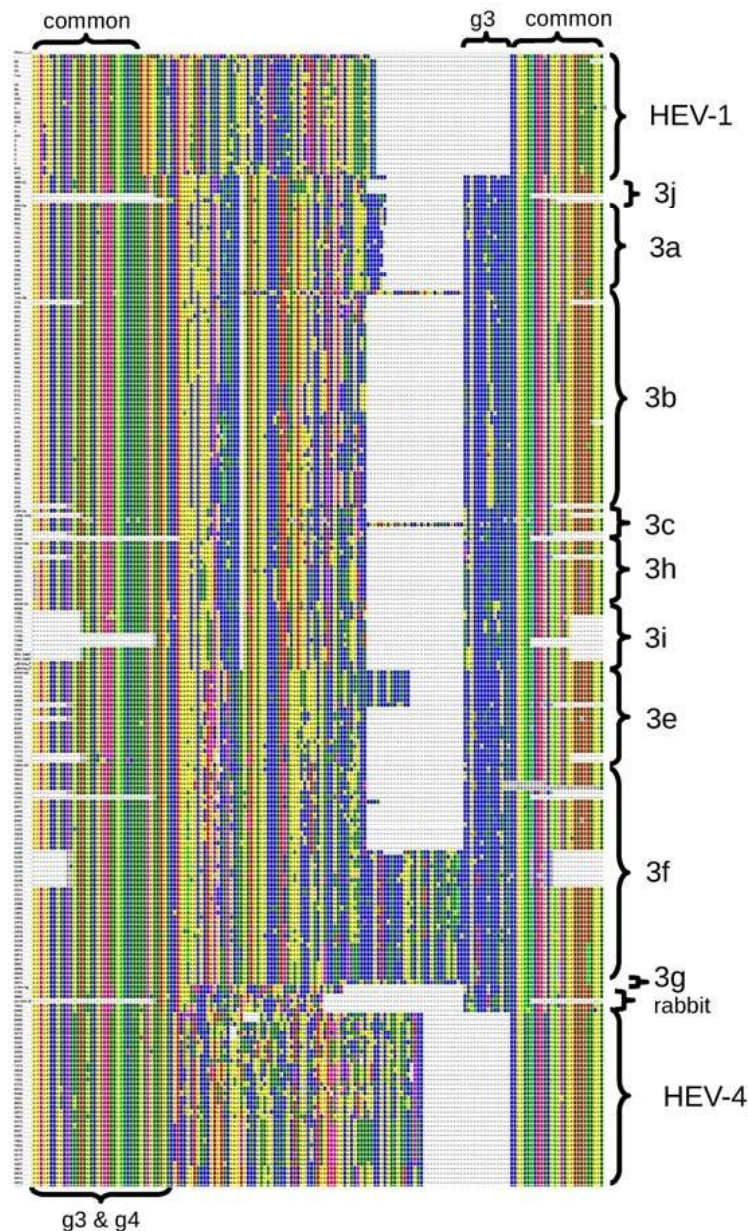


Figure 5. Scheme of the alignment of the deduced amino acid sequences including the HVR, which is flanked by conserved regions (some sequences only partially span this region and gaps in the conserved flanks are not part of the alignment). Obtained from the Alignment Explorer of MEGA6 (screenshot).

Each sequence in the alignment was analyzed by constructing a tree based on its exact (or near equal) length and including all the other sequences that span the region. More than 50 trees were analyzed, and as a result 68 more sequences were selected as “reference”, 48 of them longer than 1000 nt. Trees constructed for sequences more than 1500 nt length retained the same topology and support, with bootstrap values of at least 80% and each of the subtypes with more than one sequence. A representative tree based on 96 partial sequences spanning a region of 1860 nucleotide from position 93 to 1925 in the ORF1 (ORF1.1860nt) is depicted in Figure 3b.

Reduced sequence length maintains the same basic tree topology, but leads to a reduced or no bootstrap value support. This goes along with an increasing number of available sequences, better reflecting the HEV-3 diversity (Supplementary Figure S03). The tree based on a 318-nucleotide region from the 5' section of the genome (position 61–378, ORF1.318nt), which includes the alignment of 124 partial sequences, is shown in Figure 3b. Newly recovered sequences from German wild boar isolates

could be classified as subtype 3b (isolate WS03-09) and as subtype 3i (isolate WS05-09). Isolate WS34-10 was tentatively assigned to subtype 3b.

An overview of all the newly classified isolates from Mecklenburg-Western Pomerania and the corresponding accession numbers are depicted in Supplementary Table S02. A phylogenetic tree based on a 242-nucleotide region (positions 125–366 nt, ORF2.242nt) including 294 isolates is deposited at Supplementary Figure S03a.

The HVR (nucleotide position 2146–2358) was found to be particularly variable and was therefore manually aligned considering the deduced amino acid sequences. A scheme of a protein alignment of this region is shown in Figure 5. The longest stable HVR was found in strains of the subtype 3f, which were isolated from humans in France and Spanish pigs. In comparison, all other isolates contained shorter sequences without changing the coding frame. A tree based on a 319-nucleotide region (position 2094–2412) with 126 partial HEV-3 nucleotide sequences is shown in Figure 4a. This grouping assigned the strain WS03/09 to subtype 3b, the two isolates WS34-10 and WS35-10 were associated only weakly with subtype 3a and five isolates (WS05-09, WS21-10, 5160, 5304, 4322) clustered within subtype 3i.

A tree for 147 partial sequences within the RdRp region (280 nucleotide length, position 4284–4563, RdRp.280nt, Figure 4b, which approximately corresponds to Figure 6 presented by Lu *et al.* [32]) reproduced a similar topology compared to the CG. This tree confirmed the assignment of the strain WS03/09 to subtype 3b and a weak association of two isolates WS34-10 and WS28-10 to subtype 3a.

Finally, we carried out a phylogenetic analysis with sequences of the ORF3 (position 5180–5404, ORF3.225nt, Supplementary Figure S10/S11) and ORF2 (position 6277–6463, ORF2.187nt in Figure 4c, which partially correspond to Figure 5 in [32]) including 114 and 124 sequences, respectively. In both cases, monophyletic groups were not well supported leading to bootstrap values below 70% and in the case of the ORF3 without clear separation into subtypes. Thus, the subtypes were proposed (if possible) assuming the subtype of the nearest sequences. Using ORF3 derived sequences, we assigned three isolates to subtype 3a (isolate 8603, isolate 4701 and 4973), isolate WS03-09 to 3b, three isolates (WS28-10, WS34-10, WS35-10) weakly to 3b and two isolates to 3i (WS05-09, WS21-10). In the case of partial ORF2 derived sequences, two isolates segregated to subtype 3a (8603, 4701) and three isolates into subtype 3b (WS34-10, WS35-10, WS03-09).

In summary, twelve new isolates from Germany were analyzed using partial sequences from different regions of the genome (Supplementary Table S02). More than 1200 publicly available sequences, representing more than 1100 isolates, were subtyped accordingly and are listed in Supplementary Table S01. The classified strains are summarized by geographic region, subtype and host in Table 2.

Table 2. Subtype distribution of the 1109 isolates of genotype 3, represented by the 1283 sequences analyzed. Summarized by continent, country, and source—shown in order: human‘pig‘wild animal‘other. For example 1’2’3’4 means: 1 human isolate; 2 isolates from pig; 3 isolates from wild animals; and 4 from others sources (environment samples, water, *etc.*). As some isolates have undetermined subtypes and are not shown, not all rows or columns sum the real total. Due to various types of biases in the selection of the biological samples by the original authors and in our selection of sequences, the only approximate analysis possible was the comparison of the proportion of subtypes by host in different countries or regions. An alternative, graphical view is available at Supplementary Figure S13, which can also be explored interactively.

Origen	3	3jab				3chi				3feg					Total
		3a	3b	3d	3j	3h	3L	3i	3c	3e	3ef	3f	3g	3k	
Africa						7									7
CMR						0’4									0’4
MDG						0’3									0’3
America		27	2		2		15								47
ARG					1		3’1								5’1
URY							10								10
BRA							0’1								0’1
CUB		7’6													7’6
CAN		2’1	1’1												3’2
MEX					0’1										0’1
USA		6’4’0’1													6’4’0’1
Asia		18	67	3	3	3				13	1	5	1		114
CHN			1’1	0’3											1’4
JPN		10’3’2	30’12’23		1’2					4’3’4					46’20’29
KOR		0’2													0’2
KGZ													0’1		0’1
MNG					0’1						0’1				0’2
NZL					1’1					1					2’1
THA												4’1			4’1
Europe	8	59	2			10	25	277		326	2	208	9		940
ESP		0’0’0’1				0’0’0’1				0’2		7’27’0’13			7’29’0’15
FRA			1			3’1			3’1	4’1’0’1		24’9			35’12’0’1
GRC										1		1			2’3
ITA						1			0’1			0’2			1’3
AUT		0’4					1’1								1’5
NLD		1’10						16’29’1		2’2’0’1		5’33’0’1			24’74’1’2
DEU	4	9’2’11	0’0’1			0’0’4	0’0’21	40’6’14		11’8’6	2	15’1’1	1		82’17’61
HUN		0’8’3								3’12’2		2			5’20’5
CZE										2		1’4	1’4’2		4’8’2
SRB		0’4											0’1’2		0’5
SVN														0’3	0’3
GBR	3	6						165		268		57			507
SWE	1						2	0’0’1				4’1			7’0’1’1
Total	8	104	71	3	5	20	15	25	277	339	3	213	10	3	1109

4. Discussion

The high sequence variability of HEV genomes is the central problem affecting the screening and diagnostic methods for the detection and quantification of the viral RNA. Therefore, the conserved ORF3 region offers a promising target for PCR assays (as reported by Jothikumar *et al.* [55]). Within this region, considering potential secondary structures, the high cg content, and trying to avoid the frequently observed non-specific signals with high Ct-values, we developed a new real-time RT-PCR assay (Table 1).

The HEV RNA detection rate of 3.4% in nearly 1000 sera collected from swine in 1996/97 and of 5.2% collected in 2005/2006 in the here presented study was similar to a previously reported 5.3% prevalence rate found in 189 samples collected in 1995/96 from the same region [12]. This indicates a constant circulation of HEV in this region. In addition, the prevalence of HEV in liver samples was 10.4%, which corresponds to the 14.9% [14] and 18.1% [13] found in wild boar derived liver samples from other regions in Germany.

The use of different and short genome regions for genotyping can lead to incongruences and provides insufficient evidence for establishing or refuting phylogenetic hypotheses [56]. Considering the given restriction, Lu *et al.* [32] proposed a comprehensive subtype scheme for the phylogenetic analysis of Hepatitis E virus, which has been commonly used. Nevertheless, Lu pointed to some incongruence in this scheme due to the use of different regions and to the small number of sequences within some subtypes that were available.

A major source of inconsistency during subtyping is the combined use of short sequences and the pooling of samples with a subsequent *in silico* concatenation of sequences. For example, Lu pointed out that the Arkell strain isolated from a pool of pig feces in Canada [57] is probably an artificial mixture of sequences, which could explain the inconsistent classification of this strain using different regions. Similarly, ORF1.242nt and ORF2.301nt sequences derived from pools of 20 to 60 pig faces [19] were used by Lu to define the 3c subtype (in the major clade I), but if only the ORF.148nt is analyzed, two of them distantly cluster 3f in the major clade II. Both 3c and 3f subtypes are common in Dutch pigs. Most of these inconsistencies can be avoided using the original set (or a new consensus set) of reference sequences. Nevertheless, in general, this effect could appear without pooling of samples, due to co-infections and to true recombination between distant strains that are presumably rare events.

Another source of inconsistency is the lack, or insufficient number, of strains in some subtypes. The introduction of a new subtype based only on one single or a few short sequences can be error prone, due to laboratory artifacts, insufficient phylogenetic information, recombination, *etc.* For example, a new Hepatitis C virus (HCV) subtype (among other requisites) is created only when one complete genome (CG) and two other sufficiently informative sequences are available [58]. Basically, the CG will serve as a reference along the whole genome and the other two will determine the cluster, or prove the existence of a relatively recent common ancestor. This is an obvious problem within the HEV group 3chi: only 12 sequences longer than 1500 nt are available, from which only seven comprise CG (Supplementary Figure S12). Within the group 3chi, the best-described subtype appears to be subtype 3c (in the set of sequences we had already analyzed), with three CG and a large number of partial sequences. However, the 3i subtype is represented by only one CG, thus making it difficult to compare the sequences from different genomic regions that could be assigned to this subtype. Three

CG were assigned to the subtype 3h, but they are highly divergent. It is important to note that the current poor structure of the group 3chi is not due to a rare detection of 3chi sequences, but rather to a relatively limited effort to obtain CG or nearly complete sequences. Most of the long HEV-3 sequences have been obtained in Japan (54 CG of 97 and 58 of 117 sequences longer 1500 nt), where 3chi apparently does not circulate.

Based on Lu's classification scheme, we generated an updated phylogenetic tree with all newly available CG of genotype 3 and used the corresponding structure (Figure 3a) as template for subsequent classification of other strains based on partial sequences only. Our experimentally recovered partial sequences from different genomic regions, covering the 5' ORF1, HVR, ORF3 and ORF2 (target regions selected in this study) and the RdRp regions [59] were originally selected to match that of the majority of the European sequences. The results show that partial sequences from our ORF1, HVR, RdRp and ORF2 regions generated trees with similar structures compared to the reference tree and can be used to subtype most sequences. In contrast, sequences from ORF3.225 are only partially suitable for classification up to the subtype level. In this context, the tree topology of Figures 3b and 5 correspond to previously published trees (Figures 2 and 6, Lu *et al.* [32]).

The HVR is not a typical hypervariable region, but rather a genotypically diverse sequence [31]. The variability of this region has two components: (1) a higher mutation rate, and (2) insertions and/or deletions of one or two triplets or of much longer sequences (but maintaining the same reading frame) [60,61]. Taking this into account, it is almost trivial to find the right alignment manually and to decide whether the region should be included or not in the phylogenetic analyses. This alignment alone (Figure 5) allows an approximate reconstruction of the evolutionary history of the HEV genotypes and subtypes.

Based on the analysis of the 1652 nt part of the ORF2 not overlapping the ORF3 region, Purdy *et al.* [31] (including near 55 HEV-3 sequences) calculated the Time of the Most Recent Common Ancestor (TMRCA) of the four genotypes to be around the year 1475, and for HEV-3 and -4 around 1595. The TMRCA of major clades 3-I and 3-II (HEV-3 excluding the rabbit sequences) was determined near the year 1790, and for the clade 3-I (corresponding to 3chi and 3jab together) in 1865. Each group (3jab, 3chi and 3feg) has evolved roughly from 1900. While subtypes 3a and 3e have a TMRCA in 1945, the 3b and 3f+3ef are approximately 15 year older (without the 3ef, it is reduced for 3f to around 1960). Interestingly, the whole HEV-1 is only 100 years old. In a more recent analysis [43] using only the ORF2.301nt (including 208 HEV-3 sequences) the TMRCA for both major clades I and II together was dated to 1810 and for the clade I alone (3chi + 3jab groups) to 1895. Each group was correspondingly estimated: 3jab—1920; 3chi—1919; and 3feg—1889. For subtypes: 3a—1959; 3b—1944; 3f—1935; and 3e—1917 (include sequences not included by Purdy). The subtype 3d was the last separation, in 2002. Another study [45] shows compatible results and possible sources of minor discordance are discussed [43].

Not all possible methods of evolutionary tree reconstruction were thoroughly evaluated, but we noted that modeling evolutionary rate differences among sites have a major impact on the consistency of the results and that the tree generated with a ML method have longer internal branches and shorter terminal branches than with the NJ method, which is considered a good characteristic [62].

Partial sequences from twelve field isolates could be recovered and they all clustered within genotype 3. Three strains (4701, 4973, 8603) from the retrospective samples segregated to subtype 3a, which has been already detected in German autochthonous human infections [63], in wild boars around the city of Potsdam, in Brandenburg [14], as well as in human and pig samples from Bavaria

[15]. Subtype 3a appears to be worldwide represented in samples from humans, pigs and wild animals (especially boar). In American samples, 3a could be the predominant and potential indigenous subtype (Supplementary Table S02), and import of USA pigs has been pointed to as a source of infection in South Korea and Japan [45].

Unexpectedly, the strain WS03-09 collected from an animal hunted in Western Pomerania clustered within subtype 3b for the four regions analyzed. In Europe, subtype 3b has not been detected in wild boars or domestic pig populations so far. This subtype probably originated from Japan [45] and has been mainly identified in humans, wild boars, domestic pigs and deer from that country. It has also been isolated from one Canadian pig and (reportedly) from humans and swine in Brazil [64]. In Europe, only one human isolate from France was grouped into subtype 3b [65]. We report here the first non-Japanese 3b isolate obtained from a wild animal.

WS 34/10, WS 35/10, and WS 28/10 could be classified 3jab, but could not be unambiguously subtyped and further investigations using longer sequences are needed to define if they cluster into existing subtypes (no -a or -b) or whether they define a new subtype within the 3jab group.

No sequences within subtype 3c were discovered in this study, although 3c appears to be specific for Central Europe and is the major subtype in Germany, Netherlands and recently United Kingdom, detected in humans, pigs and wild boars.

Sequences from animals WS 05/09, WS 21/10, 5160, 5304 and 4322 clustered within subtype 3i, which is closely related to the 3c, and could have similar distribution. Curiously, none of the 101 analyzed strains from The Netherlands was classified 3i, but 21 (all from wild board) out of 162 Germans strains were classified 3i. Until now, this subtype has been detected in Germany in only wild boars, but in Austria and Argentina has been also detected in humans.

Other subtypes were not detected, although especially subtype 3e and 3f are widely distributed in Europe (Table 2). The subtype 3e appears to be more widely distributed than 3c, including clusters of sequences from Japan and West Europe but it is more represented in Central Europe. In contrast, 3f sequences are more frequently found in Spain and France, and also found in other European countries. 3f has been also detected outside Europe in Thailand (two Japan patients were infected with this subtype after a trip to this country). Interesting, only one 3f strain was isolated from wild boar (from 94 total wild boar analyzed), but 204 were isolated from humans and domestic pigs (out of 998 from all subtypes). Finally, we recommend the use of partial sequences only when the obtained tree reproduces the same structure compared to the CG tree. Ideally, sequences with more than 1000 nt should be used for classification. In contrast, sequences below 200 nt should be avoided for subtyping. In particular, the commonly used ORF2.148nt, and the OFR2.171nt generate poorly structured trees. ORF3 sequences are sufficient for genotype, but not for subtype determination. HVR sequences should only be used for intra genotype comparisons, and alignments have to be checked manually, especially in the case of sequences with long insertions, which are impossible to be compared with the reference sequences. Do not define or modify subtypes based only on a single CG or only on short sequences (less than 1500 nt).

5. Conclusions

We designed RT-PCR assays for screening, quantification and genotyping of HEV-3 strains, and detected viral RNA in wild boar samples from Mecklenburg-Western Pomerania, Germany. Twelve strains clustered into subtypes 3a, 3i and, unexpectedly, also 3b, which is a common subtype in Japan, but has not been reported in animals in Europe. The phylogenetic trees based on our partial sequences of ORF1, RdRp, HVR and ORF2 regions reproduced similar topology as obtained from complete genome analysis and were useful for subtyping.

More than 30 different PCR fragments and the corresponding genomic regions have been used for genotyping and subtyping so far, which is a source of ambiguous subtyping schemes and inadequate classification. The presented study offers an updated set of reference sequences for the relatively simple and neutral subtype scheme proposed by Lu *et al.* [32], which could eliminate most of the existing incongruences and creates the basis for new hypotheses regarding the Hepatitis E epidemiology. A comprehensive subtyping of HEV-3 according to this classification scheme could enable a detailed view of the spread of HEV-3 strains among pigs, wild life and humans, and could allow determining the consequences of infections with different subtypes on humans and finally help limit the potential spread of the disease.

Acknowledgments

We thank Birke Boettcher and Gina Lucht for excellent technical assistance.

Author Contributions

Conceived and designed the experiments: AVR, MHG, ME. VK and DB performed sample collection and characterization. JS was involved in RNA isolation and sequencing. AVR, JS, ME performed data analysis. AVR and ME wrote the paper. AVR, JS, MHG, ME edited the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Supplementary Information: (zip file: <http://www.mdpi.com/1999-4915/7/5/2704/s1>)

Tables

Supplementary Table S01. Global classification of Hepatitis E virus isolates using complete genomes (CG) or partial sequences from open reading frame 1 (ORF1), hyper variable region (HVR), RNA dependent RNA polymerase (RdRp), open reading frame 3 (ORF3) and open reading frame 2 (ORF2). All nucleotide positions refer to FJ705359 (wbGer27). Abbreviations: (*) reference sequences as cited by Lu *et al.* [32]; (**) strain used in the HEV RNA WHO standard (WHO, 2013); 3b.JIO—virulent strain from Japan; the sequences selected as “reference” are marked with an +; wb—wild boar; 3-letter country code after ISO 3166-1 alpha-3. When the information was available, this country code reflects the country from where the infection was imported. German isolates derived from Federal Institute for Risk Assessment (FIRA), Robert-Koch-Institut (RKI), Friedrich-LoefflerInstitut (FLI), Germany, University of Regensburg, Institute of Medical Microbiology and Hygiene (UReg), Bernhard Nocht Institute for Tropical Medicine (BNI) and University of Veterinary Medicine Hannover (TiHo). Burma

(Myanmar, MMR) was conserved due to the widespread use in HEV related literature. mac-hum— isolated from macaques original inoculated with human feces, or from both. (Supplem.Table S01.xlsx)

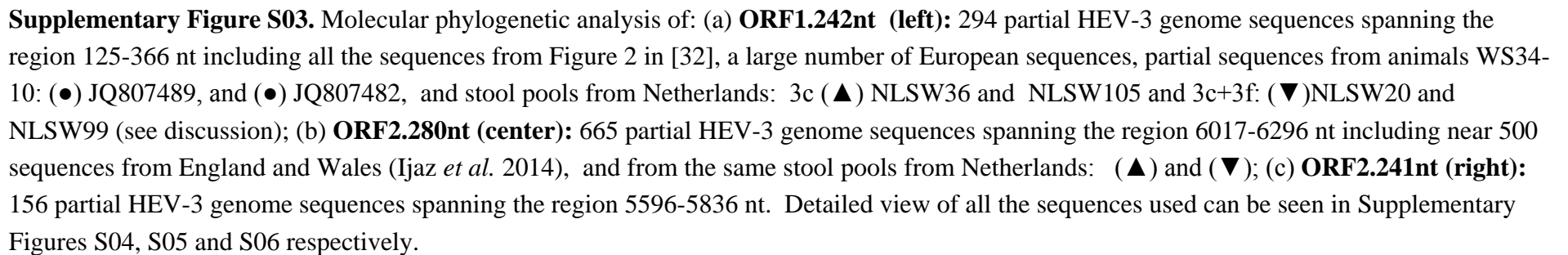
Region / year	Name	ORF1	HVR	RdRp	ORF3	ORF2	copies / reaction
Mecklenburg-Western Pomerania 2005	8603	n.a.	n.a.	n.a.	3a	3a	11.000
	accession nr.				JQ807476	HE797854	
Mecklenburg-Western Pomerania 1997	4701	n.a.	n.a.	n.a.	3a	3a	2
	accession nr.				JQ807472	HE797853	
	4973	n.a.	n.a.	n.a.	3a	n.a.	1
	accession nr.				JQ807473		
Mecklenburg-Western Pomerania, Greifswald, Diedrichshagen 2010	WS 34-10	3b (3jab)*	3a (3jab)*	3a (3jab)*	3b ?	3b ?	1.750
	accession nr.	JQ807489	JQ965993	JQ807488	JQ807490	HE797856	
	WS 35-10	n.a.	3a (3jab)*	n.a.	3b ?	3b ?	16.300
	accession nr.		JQ807491		JQ807492	HE797857	
	WS 28-10	n.a.	n.a.	3a (3jab)*	3b ?	n.a.	640
	accession nr.			JQ807486	JQ807487		
Mecklenburg-Western Pomerania, Bugewitz 2009	WS 03-09	3b	3b	3b	3b	3b	63.000
	accession nr.	JQ807479	JQ807478	JQ807477	JQ807480	HE797855	
	WS 05-09	3i	3i	n.a.	3i	n.a.	1.100
	accession nr.	JQ807482	JQ807481		JQ807483		
Mecklenburg-Western Pomerania, Greifswald, Steffenshagen 2010	WS 21-10	n.a.	3i	n.a.	3i	n.a.	1.240
	accession nr.		JQ807484		JQ807485		
Mecklenburg-Western Pomerania 1997	5160	n.a.	3i	n.a.	n.a.	n.a.	560
	accession nr.		JQ807474				
	5304	n.a.	3i	n.a.	n.a.	n.a.	400
	accession nr.		JQ807475				
	4322	n.a.	3i	n.a.	n.a.	n.a.	99
	accession nr.		JQ807471				

Supplementary Table S02. Subtype designation and accession numbers of sequences from isolates recovered in this study. *tentatively assignments.

Figures

Supplementary Figure S01. Detailed view of the sequences used in Figure 3b.

Supplementary Figure S02. Detailed view of the sequences used in Figure 3c.





Supplementary Figure S13. Geographical distribution of subtypes by host (human, domestic pig, or wild animals). This information can be also interactively explored at: https://public.tableau.com/views/HEV-g3/Story3?:embed=y&:showTabs=y&:display_count=yes.

Supplementary Figure S04. Detailed view of the sequences used in Supplementary Figure S03a.

Supplementary Figure S05. Detailed view of the sequences used in Supplementary Figure S03b.

Supplementary Figure S06. Detailed view of the sequences used in Supplementary Figure S03c.

Supplementary Figure S07. Detailed view of the sequences used in Figure 4a.

Supplementary Figure S08. Detailed view of the sequences used in Figure 4b.

Supplementary Figure S09. Detailed view of the sequences used in Figure 4c.

Supplementary Figure S11. Molecular phylogenetic analysis of: (a) **ORF3.225nt**: 114 partial HEV-3 genome sequences spanning the region 5180–5404 nt, including sequences obtained in this study (●).

Supplementary Figure S10. Detailed view of the sequences used in in Supplementary Figure S11.

Supplementary Figure S12. Detailed view of the 3chi group from three partial genomic regions.

Supplementary Files

Supplementary File S01: HEV-A.fasta.zip. Multialignment in FASTA format, compressed.

Supplementary File S02: HEVsubtypingMEGAut.xlsx. Auxiliary worksheet for classification, selection and automatic labeling of sequences in MEGA.

References

1. Moal, V.; Gerolami, R.; Ferretti, A.; Purgus, R.; Devichi, P.; Burtey, S.; Colson, P. Hepatitis e virus of subtype 3i in chronically infected kidney transplant recipients in southeastern france. *J. Clin. Microbiol.* **2014**, *52*, 3967–3972.
2. Moal, V.; Ferretti, A.; Devichi, P.; Colson, P. Genome sequence of a hepatitis e virus of genotype 3e from a chronically infected kidney transplant recipient. *Genome Announc.* **2014**, *2*, e01156–e0115613.
3. Marano, G.; Vaglio, S.; Pupella, S.; Facco, G.; Bianchi, M.; Calizzani, G.; Candura, F.; Catalano, L.; Farina, B.; Lanzoni, M.; *et al.* Hepatitis E: An old infection with new implications. *Blood Transfus. = Trasfus. Sangue* **2015**, *13*, 6–20.
4. Colson, P.; Borentain, P.; Queyriaux, B.; Kaba, M.; Moal, V.; Gallian, P.; Heyries, L.; Raoult, D.; Gerolami, R. Pig liver sausage as a source of hepatitis e virus transmission to humans. *J. Infect. Dis.* **2010**, *202*, 825–834.
5. Li, T.C.; Chijiwa, K.; Sera, N.; Ishibashi, T.; Etoh, Y.; Shinohara, Y.; Kurata, Y.; Ishida, M.; Sakamoto, S.; Takeda, N.; *et al.* Hepatitis e virus transmission from wild boar meat. *Emerg. Infectious Dis.* **2005**, *11*, 11–13.
6. Tei, S.; Kitajima, N.; Takahashi, K.; Mishihiro, S. Zoonotic transmission of hepatitis e virus from deer to human beings. *Lancet* **2003**, *362*, 371–373.
7. Renou, C.; Pariente, A.; Cadranet, J.F.; Nicand, E.; Pavio, N. Clinically silent forms may partly explain the rarity of autochthonous genotype 3c hepatitis e infection in france. *J. Clin. Virol.* **2011**, *51*, 139–141.
8. Lewis, H.C.; Wichmann, O.; Duizer, E. Transmission routes and risk factors for autochthonous hepatitis e virus infection in europe: A systematic review. *Epidemiol. Infect.* **2010**, *138*, 145–166.
9. Meader, E.; Thomas, D.; Salmon, R.; Sillis, M. Seroprevalence of hepatitis e virus in the UK farming population. *Zoonoses Public Health* **2010**, *57*, 504–509.
10. Kaba, M.; Davoust, B.; Marie, J.L.; Colson, P. Detection of hepatitis e virus in wild boar (*sus scrofa*) livers. *Vet. J.* **2010**, *186*, 259–261.
11. Preiss, J.C.; Plentz, A.; Engelmann, E.; Schneider, T.; Jilg, W.; Zeitz, M.; Duchmann, R. Autochthonous hepatitis e virus infection in germany with sequence similarities to other european isolates. *Infection* **2006**, *34*, 173–175.
12. Kaci, S.; Nöckler, K.; Johne, R. Detection of hepatitis e virus in archived german wild boar serum samples. *Vet. Microbiol.* **2008**, *128*, 380–385.
13. Adlhoch, C.; Wolf, A.; Meisel, H.; Kaiser, M.; Ellerbrok, H.; Pauli, G. High hev presence in four different wild boar populations in east and west germany. *Vet. Microbiol.* **2009**, *139*, 270–278.
14. Schielke, A.; Sachs, K.; Lierz, M.; Appel, B.; Jansen, A.; Johne, R. Detection of hepatitis e virus in wild boars of rural and urban regions in germany and whole genome characterization of an endemic strain. *Virol. J.* **2009**, *6*, e58.
15. Wenzel, J.J.; Preiß, J.; Schemmerer, M.; Huber, B.; Plentz, A.; Jilg, W. Detection of hepatitis e virus (hev) from porcine livers in southeastern germany and high sequence homology to human hev isolates. *J. Clin. Virol.* **2011**, *52*, 50–54.
16. Reuter, G.; Fodor, D.; Forgách, P.; Kátai, A.; Szucs, G. Characterization and zoonotic potential of endemic hepatitis e virus (hev) strains in humans and animals in hungary. *J. Clin. Virol.* **2009**, *44*, 277–281.
17. Ponterio, E.; di Bartolo, I.; Orru, G.; Liciardi, M.; Ostanello, F.; Ruggeri, F.M. Detection of serum antibodies to hepatitis e virus in domestic pigs in italy using a recombinant swine hev capsid protein. *BMC Vet. Res.* **2014**, *10*, e133.
18. Caruso, C.; Modesto, P.; Bertolini, S.; Peletto, S.; Acutis, P.L.; Dondo, A.; Robetto, S.; Mignone, W.; Orusa, R.; Ru, G.; *et al.* Serological and virological survey of hepatitis e virus in wild boar populations in northwestern italy: Detection of hev subtypes 3e and 3f. *Arch. Virol.* **2015**, *160*, 153–160.

19. Van der Poel, W.H.; Verschoor, F.; van der Heide, R.; Herrera, M.I.; Vivo, A.; Kooreman, M.; de Roda Husman, A.M. Hepatitis e virus sequences in swine related to sequences in humans, the netherlands. *Emerg. Infect. Dis.* **2001**, *7*, 970–976.
20. Rutjes, S.A.; Lodder, W.J.; Lodder-Verschoor, F.; van den Berg, H.H.; Vennema, H.; Duizer, E.; Koopmans, M.; de Roda Husman, A.M. Sources of hepatitis e virus genotype 3 in the netherlands. *Emerg. Infect. Dis.* **2009**, *15*, 381–387.
21. Hakze-van der Honing, R.W.; van Coillie, E.; Antonis, A.F.G.; van der Poel, W.H.M. First isolation of hepatitis e virus genotype 4 in europe through swine surveillance in the netherlands and belgium. *PLoS ONE* **2011**, *6*, e22673.
22. De Deus, N.; Seminati, C.; Pina, S.; Mateu, E.; Martín, M.; Segalés, J. Detection of hepatitis e virus in liver, mesenteric lymph node, serum, bile and faeces of naturally infected pigs affected by different pathological conditions. *Vet. Microbiol.* **2007**, *119*, 105–114.
23. Steyer, A.; Naglič, T.; Močilnik, T.; Poljšak-Prijatelj, M.; Poljak, M. Hepatitis e virus in domestic pigs and surface waters in slovenia: Prevalence and molecular characterization of a novel genotype 3 lineage. *Infect. Genet. Evol.* **2011**, *11*, 1732–1737.
24. Vasickova, P.; Psikal, I.; Widen, F.; Smitalova, R.; Bendova, J.; Pavlik, I.; Kralik, P. Detection and genetic characterisation of hepatitis e virus in czech pig production herds. *Res. Vet. Sci.* **2009**, *87*, 143–148.
25. Widén, F.; Sundqvist, L.; Matyi-Toth, A.; Metreveli, G.; Belák, S.; Hallgren, G.; Norder, H. Molecular epidemiology of hepatitis e virus in humans, pigs and wild boars in sweden. *Epidemiol. Infect.* **2011**, *139*, 361–371.
26. Meng, X.J. Recent advances in hepatitis e virus. *J. Viral Hepat.* **2010**, *17*, 153–161.
27. Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.-C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.R. Hepatitis e virus (heV): Molecular cloning and sequencing of the full-length viral genome. *Virology* **1991**, *185*, 120–131.
28. Koonin, E.V.; Gorbalenya, A.E.; Purdy, M.A.; Rozanov, M.N.; Reyes, G.R.; Bradley, D.W. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis e virus: Delineation of an additional group of positive-strand rna plant and animal viruses. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8259–8263.
29. Okamoto, H. Hepatitis e virus cell culture models. *Virus Res.* **2011**, *161*, 65–77.
30. Smith, D.B.; Simmonds, P.; members of the International Committee on the Taxonomy of Viruses *Hepeviridae* Study Group; Jameel, S.; Emerson, S.U.; Harrison, T.J.; Meng, X.J.; Okamoto, H.; van der Poel, W.H.; Purdy, M.A. Consensus proposals for classification of the family hepeviridae. *J. Gen. Virol.* **2014**, *95*, 2223–2232.
31. Purdy, M.A.; Khudyakov, Y.E. Evolutionary history and population dynamics of hepatitis e virus. *PLoS ONE* **2010**, *5*, e14376.
32. Lu, L.; Li, C.; Hagedorn, C.H. Phylogenetic analysis of global hepatitis e virus sequences: Genetic diversity, subtypes and zoonosis. *Rev. Med. Virol.* **2006**, *16*, 5–36.
33. Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.; Reyes, G.R. Molecular cloning and sequencing of the mexico isolate of hepatitis e virus (heV). *Virology* **1992**, *191*, 550–558.
34. Van Cuyck-Gandré, H.; Zhang, H.Y.; Tsarev, S.A.; Clements, N.J.; Cohen, S.J.; Caudill, J.D.; Buisson, Y.; Coursaget, P.; Warren, R.L.; Longer, C.F. Characterization of hepatitis e virus (heV) from algeria and chad by partial genome sequence. *J. Med. Virol.* **1997**, *53*, 340–347.
35. Wang, Y.; Ling, R.; Erker, J.C.; Zhang, H.; Li, H.; Desai, S.; Mushahwar, I.K.; Harrison, T.J. A divergent genotype of hepatitis e virus in chinese patients with acute hepatitis. *J. Gen. Virol.* **1999**, *80*, 169–177.
36. Kwo, P.Y.; Schlauder, G.G.; Carpenter, H.A.; Murphy, P.J.; Rosenblatt, J.E.; Dawson, G.J.; Mast, E.E.; Krawczynski, K.; Balan, V. Acute hepatitis e by a new isolate acquired in the united states. *Mayo Clin. Proc. Mayo Clin.* **1997**, *72*, 1133–1136.
37. Aggarwal, R.; Naik, S. Epidemiology of hepatitis e: Current status. *J. Gastroenterol. Hepatol.* **2009**, *24*, 1484–1493.
38. Robert Koch-Institut. Available online: <http://www.rki.de> (accessed on 7 April 2015)

39. Scharn, N.; Ganzenmueller, T.; Wenzel, J.J.; Dengler, R.; Heim, A.; Wegner, F. Guillain-barre syndrome associated with autochthonous infection by hepatitis e virus subgenotype 3c. *Infection* **2014**, *42*, 171–173.
40. Manka, P.; Bechmann, L.P.; Coombes, J.D.; Thodou, V.; Schlattjan, M.; Kahraman, A.; Syn, W.K.; Saner, F.; Gerken, G.; Baba, H.; *et al.* Hepatitis e virus infection as a possible cause of acute liver failure in europe. *Clin. Gastroenterol. Hepatol.* **2015**, doi: 10.1016/j.cgh.2015.04.014.
41. Takahashi, K.; Okamoto, H.; Abe, N.; Kawakami, M.; Matsuda, H.; Mochida, S.; Sakugawa, H.; Sugino, Y.; Watanabe, S.; Yamamoto, K.; *et al.* Virulent strain of hepatitis e virus genotype 3, japan. *Emerg. Infect. Dis.* **2009**, *15*, 704–709.
42. Pfefferle, S.; Frickmann, H.; Gabriel, M.; Schmitz, N.; Gunther, S.; Schmidt-Chanasit, J. Fatal course of an autochthonous hepatitis e virus infection in a patient with leukemia in germany. *Infection* **2012**, *40*, 451–454.
43. Zehender, G.; Ebranati, E.; Lai, A.; Luzzago, C.; Paladini, S.; Tagliacarne, C.; Galli, C.; Galli, M.; Ciccozzi, M.; Zanetti, A.R.; *et al.* Phylogeography and phylodynamics of european genotype 3 hepatitis e virus. *Infect. Genet. Evol.* **2014**, *25*, 138–143.
44. Nakano, T.; Okano, H.; Kobayashi, M.; Ito, K.; Ohmori, S.; Nomura, T.; Kato, H.; Ayada, M.; Nakano, Y.; Akachi, S.; *et al.* Molecular epidemiology and genetic history of european-type genotype 3 hepatitis e virus indigenized in the central region of japan. *Infect. Genet. Evol.* **2012**, *12*, 1524–1534.
45. Nakano, T.; Takahashi, K.; Pybus, O.G.; Hashimoto, N.; Kato, H.; Okano, H.; Kobayashi, M.; Fujita, N.; Shiraki, K.; Takei, Y.; *et al.* New findings regarding the epidemic history and population dynamics of japan-indigenous genotype 3 hepatitis e virus inferred by molecular evolution. *Liver Int.* **2012**, *32*, 675–688.
46. Nakano, T.; Takahashi, K.; Arai, M.; Okano, H.; Kato, H.; Ayada, M.; Okamoto, H.; Mishi, S. Identification of european-type hepatitis e virus subtype 3e isolates in Japanese wild boars: Molecular tracing of hev from swine to wild boars. *Infect. Genet. Evol.* **2013**, *18*, 287–298.
47. Colson, P.; Borentain, P.; Motte, A.; Lagrange, X.; Kaba, M.; Henry, M.; Tamalet, C.; Gerolami, R. First human cases of hepatitis e infection with genotype 3c strains. *J. Clin. Virol.* **2007**, *40*, 318–320.
48. Kamar, N.; Dalton, H.R.; Abravanel, F.; Izopet, J. Hepatitis e virus infection. *Clin. Microbiol. Rev.* **2014**, *27*, 116–138.
49. Hoffmann, B.; Depner, K.; Schirrmeier, H.; Beer, M. A universal heterologous internal control system for duplex real-time rt-pcr assays used in a detection system for pestiviruses. *J. Virol. Methods* **2006**, *136*, 200–209.
50. Hall, T.A. Bioedit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt. *Nucl. Acids Symp. Ser.* **1999**, *41*, 95–98.
51. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. Mega6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729.
52. Schlosser, J.; Eiden, M.; Vina-Rodriguez, A.; Fast, C.; Dremsek, P.; Lange, E.; Ulrich, R.G.; Groschup, M.H. Natural and experimental hepatitis e virus genotype 3-infection in european wild boar is transmissible to domestic pigs. *Vet. Res.* **2014**, *45*, e121.
53. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **1980**, *16*, 111–120.
54. Norder, H.; Sundqvist, L.; Magnusson, L.; Ostergaard Breum, S.; Löfdahl, M.; Larsen, L.E.; Hjulsgaard, C.K.; Magnius, L.; Böttiger, B.E.; Widen, F. Endemic hepatitis e in two nordic countries. *Eurosurveillance* **2009**, *14*, 1–9.
55. Jothikumar, N.; Cromeans, T.L.; Robertson, B.H.; Meng, X.J.; Hill, V.R. A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis e virus. *J. Virol. Methods* **2006**, *131*, 65–71.
56. Rokas, A.; Williams, B.L.; King, N.; Carroll, S.B. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* **2003**, *425*, 798–804.

57. Pei, Y.; Yoo, D. Genetic characterization and sequence heterogeneity of a canadian isolate of swine hepatitis e virus. *J. Clin. Microbiol.* **2002**, *40*, 4021–4029.
58. Simmonds, P.; Bukh, J.; Combet, C.; Deleage, G.; Enomoto, N.; Feinstone, S.; Halfon, P.; Inchauspe, G.; Kuiken, C.; Maertens, G.; *et al.* Consensus proposals for a unified system of nomenclature of hepatitis c virus genotypes. *Hepatology* **2005**, *42*, 962–973.
59. Johne, R.; Plenge-Bönig, A.; Hess, M.; Ulrich, R.G.; Reetz, J.; Schielke, A. Detection of a novel hepatitis e-like virus in faeces of wild rats using a nested broad-spectrum RT-PCR. *J. Gen. Virol.* **2010**, *91*, 750–758.
60. Purdy, M.A. Evolution of the hepatitis e virus polyproline region: Order from disorder. *J. Virol.* **2012**, *86*, 10186–10193.
61. Lhomme, S.; Abravanel, F.; Dubois, M.; Sandres-Saune, K.; Mansuy, J.M.; Rostaing, L.; Kamar, N.; Izopet, J. Characterization of the polyproline region of the hepatitis e virus in immunocompromised patients. *J. Virol.* **2014**, *88*, 12017–12025.
62. Yang, Z. *Computational Molecular Evolution*; Oxford University Press: Oxford, UK, 2006; p. xvi, p. 357.
63. Wichmann, O.; Schimanski, S.; Koch, J.; Kohler, M.; Rothe, C.; Plentz, A.; Jilg, W.; Stark, K. Phylogenetic and case-control study on hepatitis e virus infection in germany. *J. Infect. Dis.* **2008**, *198*, 1732–1741.
64. Lopes Dos Santos, D.R.; Lewis-Ximenez, L.L.; da Silva, M.F.; de Sousa, P.S.; Gaspar, A.M.; Pinto, M.A. First report of a human autochthonous hepatitis e virus infection in brazil. *J. Clin. Virol.* **2010**, *47*, 276–279.
65. Legrand-Abravanel, F.; Mansuy, J.M.; Dubois, M.; Kamar, N.; Peron, J.M.; Rostaing, L.; Izopet, J. Hepatitis e virus genotype 3 diversity, france. *Emerg. Infect. Dis.* **2009**, *15*, 110–114.

© 2015 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 license (<http://creativecommons.org/licenses/by/4.0/>).

Chapter 6. Publication V

Detection of Hepatitis E Virus in Archived Rabbit Serum Samples, Germany 1989.

Eiden, M., Ariel Vina-Rodriguez, J. Schlosser, H. Schirrmeier, and M.H. Groschup.
Food and Environmental Virology. 2015.
doi:10.1007/s12560-015-9222-4

Reproduced with permission to be included in the thesis for University use.

BRIEF COMMUNICATION

Detection of Hepatitis E Virus in Archived Rabbit Serum Samples, Germany 1989

Martin Eiden¹ • Ariel Vina-Rodriguez¹ • Josephine Schlosser¹ • Horst Schirrmeier² • Martin H. Groschup¹

Received: 11 August 2015/Accepted: 26 October 2015
©Springer Science+Business Media New York 2015

Abstract: We detected Hepatitis E virus in serum samples of wild rabbits that were hunted in 1989 around the city of Greifswald, Germany. The recovery of one partial sequence and subsequent phylogenetic analysis indicates a close relationship to rabbit HEV sequences from France and suggests a long-established circulation of rabbit HEV in Europe.

Keywords: Hepatitis E virus • Rabbit • Germany • Phylogenetic analysis

Hepatitis E virus is a positive-sense single-stranded nonenveloped RNA virus which has been classified in the Hepeviridae family, genus Hepevirus. The viral genome contains three open reading frames (ORFs), whereas ORF1 codes for several non-structural proteins, ORF2 for the capsid protein, and ORF3 for a small phosphoprotein. Hepatitis E virus is a zoonotic pathogen, transmitted primarily by the fecal–oral route causing large numbers of human cases throughout Asia, the Middle East, and Africa. In Europe and Northern America, an increasing number of autochthonous cases of hepatitis E are reported. According to a new classification scheme, these strains belong to genus

Orthohepevirus, genotype HEV-3, which infects humans, pigs, and other mammals (Smith et al. 2014). Rabbit-derived isolates belong to HEV-3 but form a separated clade in this genotype (Vina-Rodriguez et al. 2015).

After the first identification of a novel rabbit-derived HEV strain in farmed rex rabbits (Zhao et al. 2009), subsequent findings from farmed rabbits were reported in other regions of China (Geng et al. 2011; Wang et al. 2013; Geng et al. 2013), Mongolia (Jirintai et al. 2012), and the USA (Cossaboom et al. 2011, Cossaboom et al. 2012a, b). Recently, laboratory rabbits from US vendors were tested positive for antibodies against HEV (Birke et al. 2014). Current findings in Europe are only reported from France including farmed and wild rabbits (Izopet et al. 2012) and Italy (Caruso et al. 2015) so far. A recent study showed that rabbit HEV belongs to the same serotype as human HEV (Wang et al. 2013). Moreover, under experimental conditions, rabbit HEV has been shown to infect non-human primates, which developed typical hepatitis (Liu et al. 2013). In a French human strain, a unique 93-nucleotide insertion 3' adjacent to the hyper variable region (HVR) within the X domain was identified which was also found in French rabbit HEV isolates (Izopet et al. 2012). A similar 90-nucleotide insertion observed in a rabbit isolate from the United States (Cossaboom et al. 2012). The influence of this insertion on host range and pathogenesis remains to be elucidated. Thus, recovered rabbit sequences demonstrated a close relationship to (human) HEV-3 strains indicating zoonotic infections in humans.

Therefore, additional information is needed to assess the distribution of rabbit-derived HEV and the importance of this virus reservoir for human health.

In our study, 13 serum samples from wild rabbits were analyzed which were collected in the context of a general surveillance for wildlife pathogens in 1989 in the larger

Electronic supplementary material The online version of this article (doi:10.1007/s12560-015-9222-4) contains supplementary material, which is available to authorized users.

✉ Martin Eiden
martin.eiden@fli.bund.de

¹ Institute of Novel and Emerging Infectious Diseases at the Friedrich-Loeffler-Institut, Südufer 10, Greifswald-Insel-Riems, Germany

² Institute for Virus Diagnostic, at the Friedrich-LoefflerInstitut, Greifswald-Insel-Riems, Germany

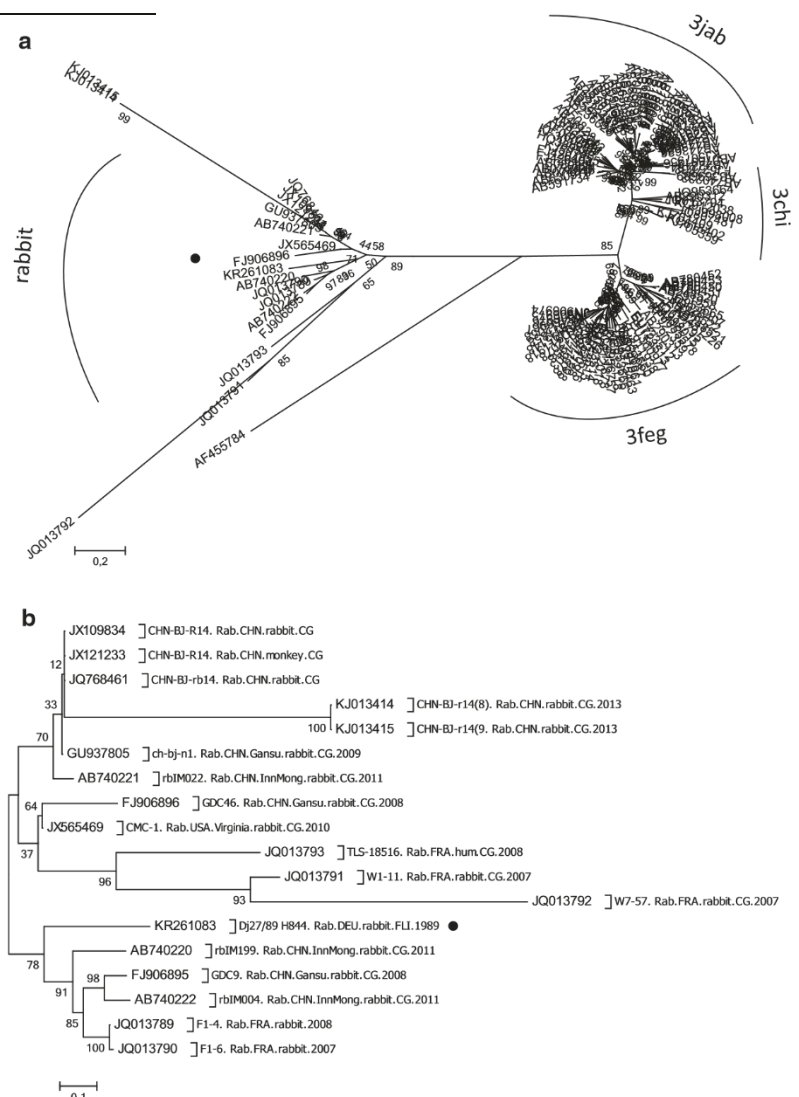
Published online: 30 October 2015

Table 1 Results of testing for HEV RNA by quantitative real-time RT-PCR and HEV antibodies by AXIOM ELISA

Laboratory ID	RT-qPCR [Ct]	Axiom ELISA	Greifswald area in Mecklenburg Western Pomerania. The presence of antibodies was analyzed with the HEV-Ab ELISA, which detects total antibodies in sera irrespectively of the species (Axiom solutions, Germany). All 13 sera tested were diluted 1:1 in specimen diluent according to manufacturer's instructions. RNA from serum was isolated via QIAamp viral kit (Qiagen, Germany) and tested by quantitative real-time RT-PCR (RT-qPCR), which targets a sequence within ORF2 (Vina-Rodriguez et al. 2015). In total, 4 out of 13 sera samples were positive in the ELISA and one sample borderline. Interestingly, four samples were PCR positive, indicating a recent infection (Table 1). One sample (Laboratory ID: Dj27/89 H844) was seropositive as well as HEV RNA positive. From this sample, a partial sequence could be recovered by RT-PCR assay which targets the hypervariable region (HVR): RT-PCR step was performed using primers HEV.HVR_F1
Dj27/89 H815	N/A	neg	
Dj27/89 H816	N/A	neg	
Dj27/89 H817	N/A	neg	
Dj27/89 H819	N/A	pos	
Dj27/89 H820	N/A	doubtful	
Dj27/89 H836	N/A	neg	
Dj27/89 H839	35, 82	neg	
Dj27/89 H841	N/A	pos	
Dj27/89 H843	N/A	pos	
Dj27/89 H844	29, 22	pos	
Dj27/89 H846	36, 01	neg	
Dj27/89 H847	36, 6	Neg	
Dj27/89 H848	N/A	neg	

Positive results are depicted in bold

Fig. 1 Phylogenetic analysis of rabbit HEV isolates based on partial HVR sequences (300 nt). The phylogenetic tree was generated by the maximumlikelihood method based on the Kimura 2-parameter model by using MEGA6 software. All positions with less than 80 % site coverage were eliminated. GenBank accession numbers are shown for each HEV strain used in the phylogenetic analysis. Scale bar indicates nucleotide substitutions per site. a The phylogenetic tree consists of 127 nucleotide sequences and representative strains of HEV-3 genotype including major groups 3jab, 3chi, 3feg, and the rabbit clade. Isolates are characterized by accession number, isolate designation, classification (subtype), country (3-letter code), host, and sampling year. b The phylogenetic tree of the rabbit clade with 17 rabbit-derived sequences and one human sequence (JQ013793). The German isolate KR261083 is indicated by a closed circle



(5'-TTYTCYCCTGGGCAYMTYTGGGA-3') and HEV.HVR_R1 (5'-TTAACCARCCARTCACARTCYG-AYTCAAA-3'), referring to nucleotide position 2069–2469 of wild boar isolate FJ705359 (Vina-Rodriguez et al. 2015). Phylogenetic analysis of the 300-bp sequence (accession no. KR261083) demonstrated the clustering of the German isolate within the rabbit clade of the HEV-3 genotype (Fig. 1a, b). A detailed overview of Fig. 1a including used accession numbers is shown in supplemental figure S1. Nucleotide sequence alignment assigned the smallest differences to French rabbit HEV isolates JQ013789 and JQ013790 collected in 2007/2008 (Izopet et al. 2012) displaying differences of about 30 % compared to the German isolate. Due to the short partial sequence, an extended comprehensive molecular analysis of the whole genome including the potential 93-nucleotide insertion was not possible. The general high percentage of substitutions per site is a prerequisite of the high variability within this region.

The seroprevalences in farmed rabbits differed between 0 and 55 % in distinct Chinese Provinces (Geng et al. 2011) and 30–52 % in US (Cossaboom et al. 2011). The corresponding RNA prevalence ranged from 0 to 11.6 % and 3.3–48 %, respectively. The striking differences can be attributed to a just started or ongoing fecal oral transmission of the virus under cage management conditions and are also related to the mostly unknown age of the animals. Correspondingly, virus prevalence rates for French wild rabbits, established in warrens, ranged from 0 to 100 % (Izopet et al. 2012). Free-living rabbits from our study harbored a seroprevalence of about 31 % (4 out of 13) and a RNA prevalence of 7.7 % (1 out of 13) which is in line to the previous findings; however, due to the low sample number, only limited conclusions can be made. Therefore, the prevalence of rabbit HEV in wild rabbit populations and the impact for wild life need to be examined more closely.

Since the positive samples date back to the year 1989, this detection of HEV in wild rabbits in Germany is a first hint for the presence of HEV in the German wild rabbit population and most probably also in wider Europe. To determine its current importance in the wild rabbit population, a well-structured wildlife surveillance program would be desirable in Germany and elsewhere.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

- Birke, L., Cormier, S. A., You, D., Stout, R. W., Clement, C., Johnson, M., & Thompson, H. (2014). Hepatitis E antibodies in laboratory rabbits from 2 US vendors. *Emerging Infectious Diseases*, 20, 693–696.
- Cossaboom, C. M., Cordoba, L., Cao, D., Ni, Y. Y., & Meng, X. J. (2012a). Complete genome sequence of hepatitis E virus from rabbits in the United States. *Journal of Virology*, 86, 13124–13125.
- Cossaboom, C. M., Cordoba, L., Dryman, B. A., & Meng, X. J. (2011). Hepatitis E virus in rabbits, virginia, USA. *Emerging Infectious Diseases*, 17, 2047–2049.
- Cossaboom, C. M., Cordoba, L., Sanford, B. J., Pineyro, P., Kenney, S. P., Dryman, B. A., et al. (2012b). Cross-species infection of pigs with a novel rabbit, but not rat, strain of hepatitis E virus isolated in the united states. *Journal of General Virology*, 93, 1687–1695.
- Geng, Y., Zhang, H., Li, J., Huang, W., Harrison, T. J., Zhao, C., et al. (2013). Comparison of hepatitis E virus genotypes from rabbits and pigs in the same geographic area: No evidence of natural cross-species transmission between the two animals. *Infection Genetics and Evolution*, 13, 304–309.
- Geng, Y., Zhao, C., Song, A., Wang, J., Zhang, X., Harrison, T. J., et al. (2011). The serological prevalence and genetic diversity of hepatitis E virus in farmed rabbits in China. *Infection Genetics and Evolution*, 11, 476–482.
- Izopet, J., Dubois, M., Bertagnoli, S., Lhomme, S., Marchandeau, S., Boucher, S., et al. (2012). Hepatitis E virus strains in rabbits and evidence of a closely related strain in humans, France. *Emerging Infectious Diseases*, 18, 1274–1281.
- Jirintai, S., Jinshan, Tanggis, Manglai, D., Mulyanto, Takahashi, M., et al. (2012). Molecular analysis of hepatitis E virus from farm rabbits in inner Mongolia, China and its successful propagation in A549 and PLC/PRF/5 cells. *Virus Research*, 170, 126–137.
- Liu, P., Bu, Q. N., Wang, L., Han, J., Du, R. J., Lei, Y. X., et al. (2013). Transmission of hepatitis E virus from rabbits to cynomolgus macaques. *Emerging Infectious Diseases*, 19, 559–565.
- Smith, D. B., Simmonds, P., Members of the International Committee on the Taxonomy of Viruses Hepeviridae Study Group, Jameel, S., Emerson, S. U., Harrison, T. J., et al. (2014). Consensus proposals for classification of the family hepeviridae. *Journal of General Virology*, 95, 2223–2232.
- Vina-Rodriguez, A., Schlosser, J., Becher, D., Kaden, V., Groschup, M. H., & Eiden, M. (2015). Hepatitis E virus genotype 3 diversity: A subtyping update and first detection of Genotype 3b in animals in Europe. *Viruses*, 7, 2704–2726.
- Wang, S., Dong, C., Dai, X., Cheng, X., Liang, J., Dong, M., et al. (2013). Hepatitis E virus isolated from rabbits is genetically heterogeneous but with very similar antigenicity to human HEV. *Journal of Medical Virology*, 85, 627–635.
- Zhao, C., Ma, Z., Harrison, T. J., Feng, R., Zhang, C., Qiao, Z., et al. (2009). A novel genotype of hepatitis E virus prevalent among farmed rabbits in China. *Journal of Medical Virology*, 81, 1371–1379.

Chapter 7. Publication VI

A Novel Pan-*Flavivirus* Detection and Identification Assay, Based on RT-qPCR and Microarray

Ariel Vina-Rodriguez¹, Konrad Sachse^{2#}, Ute Ziegler¹, Serafeim C. Chaintoutis^{1,3}, Markus Keller¹, Martin H. Groschup¹ and Martin Eiden¹

¹ *Institute for Novel and Emerging Infectious Diseases, Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany.*

² *Institute of Molecular Pathogenesis, Friedrich-Loeffler-Institut, Jena, Germany.*

³ *Diagnostic Laboratory, Department of Clinical Sciences, School of Veterinary Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece.*

[#] *Present address: Dept. Bioinformatics for High-Throughput Analysis, Faculty of Mathematics and Computer Science, Friedrich-Schiller-Universität Jena, Germany*

BioMed Research International

Volume 2017 (2017), Article ID 4248756

<https://doi.org/10.1155/2017/4248756>

<https://www.hindawi.com/journals/bmri/2017/4248756/>

Note: the author's manuscript was used because at the moment of preparing this thesis the final version at BioMed Research International was not yet ready. The content is the same and some effort was made to moderate formatting differences. Part of the supporting information available only *on-line* is included here.

Research Article

A Novel Pan-*Flavivirus* Detection and Identification Assay, Based on RT-qPCR and Microarray

Ariel Vina-Rodriguez¹, Konrad Sachse^{2#}, Ute Ziegler¹, Serafeim C. Chaintoutis^{1,3}, Markus Keller¹, Martin H. Groschup¹ and Martin Eiden¹

¹ Institute for Novel and Emerging Infectious Diseases, Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany.

² Institute of Molecular Pathogenesis, Friedrich-Loeffler-Institut, Jena, Germany.

³ Diagnostic Laboratory, Department of Clinical Sciences, School of Veterinary Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece.

[#] Present address: Dept. Bioinformatics for High-Throughput Analysis, Faculty of Mathematics and Computer Science, Friedrich-Schiller-Universität Jena, Germany

Correspondence should be addressed to Martin Eiden: martin.eiden@fli.bund.de

Received 20 December 2016; Accepted 7 February 2017

Academic Editor: Satish B. Nimse

Copyright © 2016 Ariel Vina-Rodriguez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The genus *Flavivirus* includes arthropod-borne viruses responsible for a large number of infections in humans and economically important animals. While RT-PCR protocols for specific detection of most *Flavivirus* species are available, there has been also a demand for a broad-range *Flavivirus* assay covering all members of the genus. It is particularly challenging to balance specificity at genus level with equal sensitivity towards each target species.

In the present study, a novel assay combining a SYBR Green-based RT-qPCR with a low-density DNA microarray has been developed. Validation experiments confirmed that the RT-qPCR exhibited roughly equal sensitivity of detection and quantification for all flaviviruses tested. These PCR products are subjected to hybridization on a microarray carrying 84 different oligonucleotide probes that represent all known *Flavivirus* species. This assay has been used as a screening and confirmation tool for *Flavivirus* presence in laboratory and field samples, and it performed successfully in international External Quality Assessment of NAT studies. Twenty-six *Flavivirus* strains were tested with the assay, showing equivalent or superior characteristics compared with the original- or even with species-specific RT-PCRs. As an example, test results on West Nile virus detection in a panel of 340 mosquito pool samples from Greece are presented.

1. Introduction

The genus *Flavivirus* contains nearly 70 recognized viruses, many of which infect humans and economically important animals [1]. Flaviviruses, such as Dengue virus (DENV) [2] and Yellow fever virus (YFV) [3], have been a common cause of devastating diseases in tropical and less developed countries, but in recent years the emergence of flaviviral zoonoses was observed worldwide. Examples include the occurrence of West Nile virus (WNV) in the United States [4], Japanese encephalitis virus (JEV) in Australia [5], as well as Usutu virus (USUV) [6], WNV [7] and DENV [2] in Europe. Recently, Zika virus (ZIKV) also expanded into Southern America, with reports of detection in Europe and USA [8].

Large surveillance and early warning systems commonly applied in European countries and around the world could benefit from a more sensitive and broader range screening method. Both mosquito pools and (sentinel) birds are common targets of massive screening for arbovirus, particularly for flaviviruses like WNV or USUV [6, 9, 10]. Rapid virus identification and quantification are crucial for accurate diagnosis of ongoing infections, treatment selection and follow-up, as well as for selection and timely introduction of control measures in outbreaks scenarios. In this context, highly parallel detection technologies, such as DNA microarrays, are gaining importance [11-20].

Like RNA viruses in general, flaviviruses are distinguished by extensive genetic heterogeneity, which implies classification into subunits, e.g. genotypes, lineages, etc., each with distinct epidemiological or clinical significance. This heterogeneity represents a major challenge in primer and probe design for PCR and DNA microarray assay development.

In the present study, we have optimized a pan-*Flavivirus*-specific SYBR Green-based real-time reverse-transcription polymerase chain reaction (RT-qPCR) [21]. The amplification is subsequently complemented with hybridization on a low-density DNA microarray, which exploits the genetic heterogeneity contained in the internal segment of the PCR amplicon, thus allowing rapid identification of flaviviruses from clinical or field samples.

2. Materials and Methods

2.1. Viral samples and RNA extraction.

The 26 *Flavivirus* strains to be used as reference were collected and stored at -80 °C until processed as following:

African green monkey kidney (Vero) cells (Collection of Cell Lines in Veterinary Medicine, Friedrich-Loeffler-Institut, Greifswald–Insel Riems, Germany) were infected with viruses in biosafety L3** laboratory facilities, and cell culture supernatants were collected and inactivated in Buffer AVL (QIAGEN, Hilden, Germany) as previously described [22, 23]. The WNV strains lineage 1a - NY99 (ac. AF196835) and Dakar, lineage 1b - Kunjin (ac. D00246) and lineage 2 - Uganda 1937 (ac. M12294), B956 (ac. AY532665), Sarafend (ac. AY688948) and goshawk Austria 361/10 (2009) (ac. HM015884) were reused from our previous works [22-24]. The strains USUV Germany 2011 (BH/65) (ac. HE599647), YFV 17D, JEV, TBEV Langat, Malaysia, 1956 (ac. AF253419) and Murray Valley encephalitis virus (MVEV) were reused from another work [25]. USUV Austria 2001 (939/01) and Germany 2011 (939/01) were taken from [26] and JEV Nakayama (ac. EF571853) and TBEV Neudoerfl (ac. U27495) from [27].

Additionally, the following strains were obtained from the Health Protection Agency, Salisbury, United Kingdom: WNV lineage 2, MB 1957; SLEV; YFV “French Neurotrop” and Lolli (#780). JEV Nakamura was kindly provided by A. Müllbacher (John Curtin School of Medical Research, Canberra, Australia). RNA from ZIKV and DENV 1, 2, 3 and 4 viruses was provided by P. Despres (Flavivirus Unit, Institute Pasteur, Paris, France), while F. Hufert and M. Weidmann (Institute for Virology, Göttingen, Germany) kindly provided the viruses TBEV Absettarov, 1951 (ac. AF091005), TBEV Aina, 1963 (ac. AF091006) and TBEV Hypr, 1953 (ac. U39292).

RNA was extracted with the RNeasy Mini Kit, or the QIAamp Viral RNA Mini Kit (QIAGEN, Hilden, Germany), according to the manufacturer’s protocol. A synthetic RNA was used as internal extraction control (IC) [28]. The NY99 RNA was used to create an external calibrator curve, which in turn was calibrated using the 5’UTR WNV-specific RT-qPCR [22].

2.2. Selection of primers and probes.

Nearly 200 complete *Flavivirus* genome sequences were obtained from the NCBI Nucleotide database by the end of 2010. Very similar sequences (more than 98% identity) were excluded. The sequences were aligned using CLUSTAL X [29] (accessed from BioEdit software v.7.0.5.3) [30] and VectorNTI Advanced v.10 (Invitrogen, Carlsbad, CA, USA). This alignment was manually curated using both the nucleotide and the deduced amino acid sequences, and extended by adding partial sequences to reach a total of more than 400, so as to represent the NS5 gene of all known species, including most subtypes or lineages (a regularly updated version is available at [Flavivirus GitHub site](https://github.com/qPCR4vir/VisualOliDeg)).

VisualOligoDeg (<https://github.com/qPCR4vir/VisualOliDeg>) was used to facilitate visualization and selection of appropriate oligonucleotides hybridizing to sequences of a given. The NS5 regions (spanning nucleotides 9040 to 9305 of AF196835) of the aligned sequences were imported into this workbook and were manually classified into major groups (MB - mosquito borne, TB - tick borne and Insect-only), virus groups (JEVG, YFVG, TBEVG, etc.), species (WNV, JEV, YFV, TBEV, etc.) and in some cases into lineages (like WNV-1 or WNV-2, etc.) or genotypes.

Alongside species-specific RT-PCRs [31], broader, genus-specific RT-PCR protocols have been also reported [19, 21, 32, 33]. A genus-specific Taq-Man RT-PCR for *Flavivirus* detection was modified [21], with primer sequences optimized using *VisualOligoDeg*. It targets conserved flanking regions with an internal region with sufficient variability to enable virus identification by sequencing the amplicon. The original sequences were adapted to consistently amplify most of the *Flavivirus* members resulting in degenerate primers PFlav-fAAR (TACAACATGATGGGAAAGAGA-GAGAA~~R~~AA from 9040 to 9068 of AF196835) and PFlav-rKR (GTGTCCCA~~K~~CC~~R~~GC~~T~~TGTGTCATC from position 9305 to 9283 of AF196835).

A total of 50 probes with a T_m around 55 °C were selected. A second set of sequences was also selected as a replacement, in case of failure of first-set sequences. The candidate sequences were submitted to the manufacturer for a

final *in silico* evaluation of properties (homogenous hybridization, discriminatory potential, etc.). As a result, all 84 sequences were found suitable for inclusion in the production of the microarray (TABLE S01).

2.3 Detection and quantification using RT-qPCR.

A one-step SYBRGreen-based RT-qPCR with melting curve analysis was developed. The QuantiTect SYBR® Green RT-PCR Kit (QIAGEN, Hilden, Germany) was used following the manufacturer's instructions. Briefly, each primer (PFlav-fAAR and PFlav-rKR) was 5'-biotinylated during the initial synthesis (Eurofins Genomics, Ebersberg, Germany), and used at a final concentration of 0.8 μ M, in a final reaction volume of 25 μ L, including 5 μ L of RNA sample solution. Real-time RT-qPCR was carried out on a CFX96 real-time PCR detection system (Bio-Rad Laboratories, Hercules, USA). The thermal cycling profiles are presented in TABLE 1. Species-specific RT-qPCR were also used to determine the sensitivity of the *Flavivirus* RT-qPCR. Reference RNA samples included WNV [22], USUV [26, 34] and TBEV [35]. In order to have comparable Cq values across experiments, when possible, we set a separate fluorescence cut-off value for each target (PCR primer mix) in each 96-well plate run, such that Cq \approx 28 for the control RNA WNV NY99 diluted 10^{-4} . This control was previously calibrated using a synthetic RNA [22]. In order to perform *Flavivirus* quantification, an external standard curve consisting of four dilutions of RNA WNV NY99: 10^{-2} , 10^{-3} , 10^{-4} and 10^{-5} was included in each run. These were equivalent to 4×10^4 , 4×10^3 , 4×10^2 and 40 genome copies/ μ L of RNA solution or to 1.4×10^7 , 1.4×10^6 , 1.4×10^5 , and 1.4×10^4 copies per mL of homogenized sample, respectively. To analyze the *Flavivirus* SYBR Green RT-PCR, the fluorescence was measured at step 6 of the standard protocol (TABLE 1). Similarly, for species-specific RT-qPCR, fluorescence measurement was conducted at step 4.

TABLE 1: Thermal cycling profiles used in the *Flavivirus* RT-qPCR.

Standard (4.5 h, 25 μ L)	Fast (2.5 h, 25 μ L or 10 μ L)
1. 50 °C for 30 min	1. 50 °C for 30 min
2. 95 °C for 15 min	2. 95 °C for 15 min
3. 95 °C for 15 s	3. 95 °C for 15 s
4. 55 °C for 25 s + Plate Read	4. 55 °C for 20 s
5. 72 °C for 25 s + Plate Read	5. 72 °C for 20 s + Plate Read
6. 80 °C for 1 s + Plate Read	6. GOTO 3; 44 more times
7. GOTO 3; 44 more times	7. 95 °C for 1 min
8. 95 °C for 1 min	8. Melting Curve 68 to 86 °C, increment 0.2 °C, 1 s + Plate Read
9. Melting Curve 68 to 88 °C, increment 0.1 °C, 1 s + Plate Read	END
10. 4 °C Forever (optional)	
END	

2.4. Sequence analysis.

Twenty μ L of non-purified PCR product were sent to Eurofins Genomics (Ebersberg, Germany) for direct DNA sequencing by the Sanger dideoxy method, using the amplification primers. A BLAST (NCBI) search of the obtained sequences was usually sufficient to identify the virus strain. The (internal) amplicon of approximately 240 nt contains sufficient phylogenetic information to re-construct phylogenetic trees, and allows classification of more distantly related and unknown strains and even new species.

2.5 Microarray analysis.

Flavivirus-specific oligonucleotide probes selected using the *VisualOligoDeg* were spotted onto a low-density microarray of the commercially available ArrayStrip™ (AS) platform (Alere Technologies GmbH, Jena, Germany). The 84 probes (spots 1-84 in S01 File) were spotted either in triplicate (Chip Wildtech Virology-Mycob 01, from 2011-01-13, assay ID-10610) [36] or in quintuplicate (Chip Wildtech Virology 02, from 2012-09-10, assay ID-16050). The Alere Hybridization Kit was used following previously published instructions [37, 38]. Briefly, only positive RT-PCR reactions were routinely analyzed, from which 1 μ L was directly denatured in 100 μ L of the hybridization buffer at 95 °C for 5 min and then placed for cooling at 4 °C for 5 min in a thermocycler (Bio-Rad Laboratories). This solution was transferred to the AS vessel (previously conditioned with water and hybridization buffer) and incubated at 55 °C for 1 h upon shaking at 550 rpm on a BioShake iQ heatable shaker (Quantifoil Instruments, Jena, Germany). The AS vessel was subsequently washed twice at 50 °C for 10 min, incubated with 100 μ L of a peroxidase conjugate solution at 30 °C for 10 min, washed

and incubated at room temperature with 100 μ L of the substrate solution (Seramun Grün; Seramun Diagnostica GmbH, Heidesee, Germany) for 5 min. Images of processed microarrays were saved in bitmap (.bmp) format, using the ArrayMate transmission Reader (Alere Technologies GmbH).

2.6. Microarray data processing and Flavivirus identification.

The web-based database *Pionir – The ExperimentNavigator of PARTISAN ArrayLIMS* (Alere Technologies GmbH) was used for visualization and analysis of images and complete experiments, as well as for additional backup. Alongside, the Partisan IconoClust® v3.6r0 software (Alere Technologies GmbH) was used for local analysis of microarray images, generating for each spot of each picture the background-corrected signal intensities $NI = 1 - M/BG$, with NI being normalized intensity, M average (mean) spot intensity, and BG local background intensity. Spot intensities are measured as light transmission, with M values ranging from 1 for complete transmission (background, weak spots) to 0 for complete absorption (dark spots). Thus, normalized signal intensities range between 0 and 1. Previous evaluation of the signal to background ratio (SBR) in assays using the same technology have set the cut-off value for positive signals to 0.1 [37]. A custom python script was embedded in this software to enable visualization of individual hybridizations or batch analysis of series of experiments. This script exports experimental data in different formats, including a format suitable for import into *Orange* – software, which allows visual programming and python scripting for data mining and visualization (v2.7.6.dev – installed 2014-06-06 from <http://orange.biolab.si>).

Using *Orange*, the visual program *PanFlavExpStdSampl* (FIGURE S01) was created, which, together with custom modifications of some parts of *Orange* itself, permits an interactive import of the experiments used as standards (known samples) for parallel analysis with unknown samples for classification. This program calculates the distances between pairs of signal intensity patterns (two microarray experiments), which can serve as a measure of similarity between samples. Several distance formulas are available interactively from *Orange* and their meaning is explained elsewhere [39]. These distances are visualized in one of the widgets as a heat map-like graphics (FIGURE 5 in Results), in which the labels and the order of the samples can be interactively selected from a group of pre-options. We have added (directly modifying the source code of *Orange*) the option to reorganize the heat map showing the selected sample at the top followed by the most similar samples by mouse clicking the respective cell. This graphic can also show the samples organized in a tree to reveal clustering. Another *Orange* widget uses the distances to construct a tree (FIGURE 4 in Results), in which a cut-off can be interactively selected and/or groups defined to make a report of the proposed classification.

2.7. Field samples.

After development, initial evaluation and optimization, the new assay was further evaluated by testing RNA extracts from mosquito pools from Greece. Mosquito trapping was performed within the framework of a surveillance program, which was implemented in the region of Central Macedonia during the 2012 arbovirus transmission period. The aims of this program were to molecularly characterize WNV and assess population dynamics of the major arbovirus vector species, to timely notify public health authorities on increased risk [40]. Female adult mosquitoes were collected at dry ice-baited CDC mosquito traps, which were set in areas with previous indications of high arbovirus prevalence [40].

Identified mosquitoes (50 individuals of the same genus) were placed in 2 mL microcentrifuge tubes with two 4 mm sterile glass balls and 1 mL of phosphate-buffered saline (PBS) containing 1% of heat-inactivated fetal bovine serum (FBS; Sigma-Aldrich, Steinheim, Germany). Disruption was performed for 30 s at speed 4.0 using a RiboLyser homogenizer (Hybaid, Ltd., Teddington, UK). The homogenates were centrifuged ($16,000 \times g$, 5 min at 4 °C), and 150 μ L of supernatant from each mosquito pool underwent RNA isolation, based on the previously described RNA extraction-method.

In total, 340 mosquito pools were tested with the *Flavivirus* RT-qPCR protocol. Out of them, 180 pools were represented by *Culex* mosquitoes, 75 were comprised by *Aedes* and 85 by *Anopheles* mosquitoes.

3. Results

Following preliminary observations that the sensitivity and efficiency of amplification with the primers from reference [21] were not homogeneous for different flaviviruses, we decided to design a modified set of degenerate primers. Using *VisualOligoDeg* we selected primers for a modified RT-qPCR for *Flavivirus* detection and quantification. In the initial

experiments, we compared the new procedure with published RT-qPCRs using reference viruses. All 26 *Flavivirus* reference strains described in 2.1 were tested (FIGURE 1).

The optimized *Flavivirus* RT-qPCR was also compared in terms of detection limit with species-specific RT-qPCRs for WNV, USUV and TBEV by testing fresh RNA solutions (from supernatants of infected cells) in a series of end-point dilution experiments (FIGURE S02, S03).

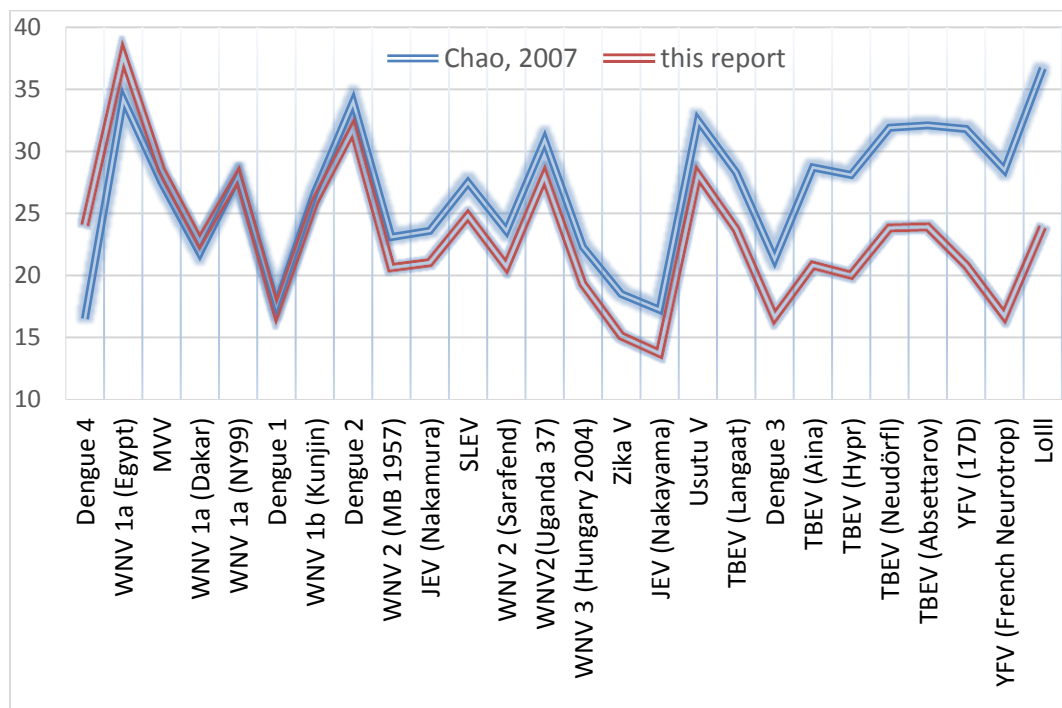


FIGURE 1: Cq values for 26 viruses. Comparison of Cq values obtained for different viruses using the protocol of Chao et al.[21] (blue line) with those of our own procedure (red line). The same virus RNA preparation was tested in the same PCR plate with both RT-qPCR protocols. The order of viruses from left to right is according to increasing Cq difference.

Subsequently, we conducted melting curve analysis, which can provide useful information to determine positivity and sequence differences (as shown in FIGURE 2), but which has limited value for identification. A well-defined peak between 79 and 84 °C is a strong indication of positivity, and differences in T_m indicate sequence differences, possibly representing different species or lineages. The results from end-point dilution experiments and the calculations using the externally calibrated standard curve suggest that, for most members of *Flavivirus* genus, this RT-qPCR assay is highly sensitive, capable of detecting a few viral RNA copies per reaction. However, this also means that cross-sample contamination and DNA carryover are a major concern. Therefore, it is crucial to organize the laboratory work accordingly and include sufficient controls to validate the results of each experiment.

To evaluate the accuracy and performance of the developed RT-qPCR assay, we successfully participated in five international ring trials for quality assessment of nucleic acid amplification tests, i.e. ANSES 2013 (melting curves of the samples shown in FIGURE 2), Quality Control for Molecular Diagnostics or QCMD-2010 QCMD-2011 [41], QCMD-2012, and QCMD-2013 (results summarized in TABLE S02) (<http://www.qcmd.org/>). These results complemented the information obtained from WNV-specific RT-qPCRs. Both WNV and non-WNV *Flavivirus* strains were quantified using the *Flavivirus* RT-qPCR with the WNV calibration curve, and were subsequently identified by microarray analysis.

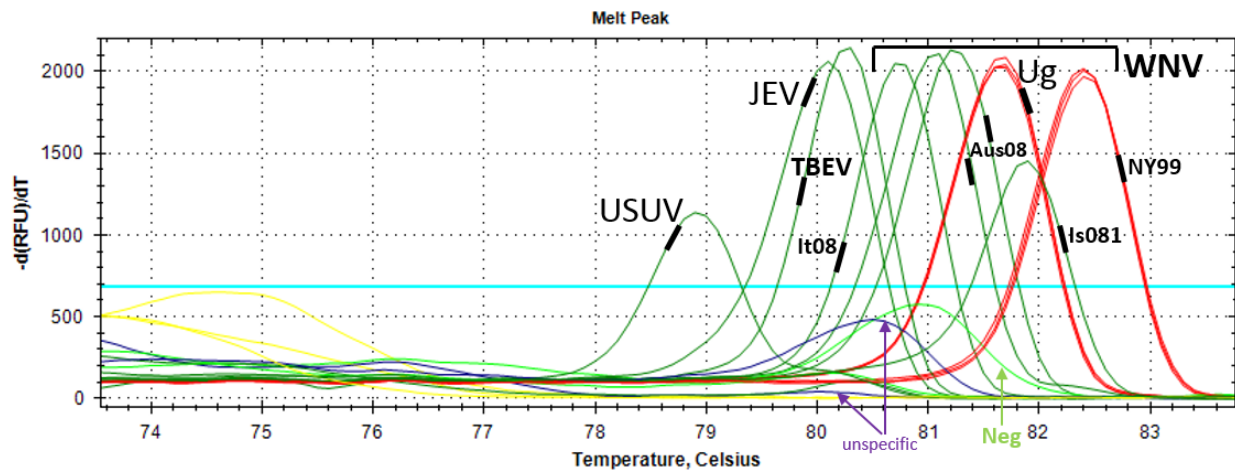


FIGURE 2: Melting curves from different flaviviruses. Melting curve analysis of the amplicons from different flaviviruses (samples from the ANSES WNV Proficiency Test 2013). NY99, It08 and Is985 are WNV lineage 1a strains, while Aus08 is a lineage 2 strain. The viruses were identified using the microarray, with confirmation by DNA sequencing.

The combined RT-qPCR/microarray procedure was applied on the 26 reference virus strains and RT-qPCR positive field samples of diverse origin, in more than 300 hybridization experiments. Each analyzed reference virus produced a specific hybridization pattern that allowed discrimination. Most flaviviruses could be identified at the level of species, genotype or even strain, following comparison of their hybridization patterns with those of reference samples. A compilation of four *Flavivirus* isolates each examined at two different RNA dilutions is shown in FIGURE 3. This figure reveals an important distinction of this microarray platform from well-known glass-slide arrays used for gene expression studies, meaning that the developed microarray is optimized to detect genetic (sequence) variations, rather than the concentration or relative quantity of amplicons. Thus, the present microarray signal intensity values are used solely for identification or classification, while quantification is performed in the preceding RT-qPCR step. As shown in FIGURE 3, the hybridization patterns are not significantly affected by quantitative variations in the viral load of the sample, but qualitative changes are readily visible when different strains of the same virus species are examined.

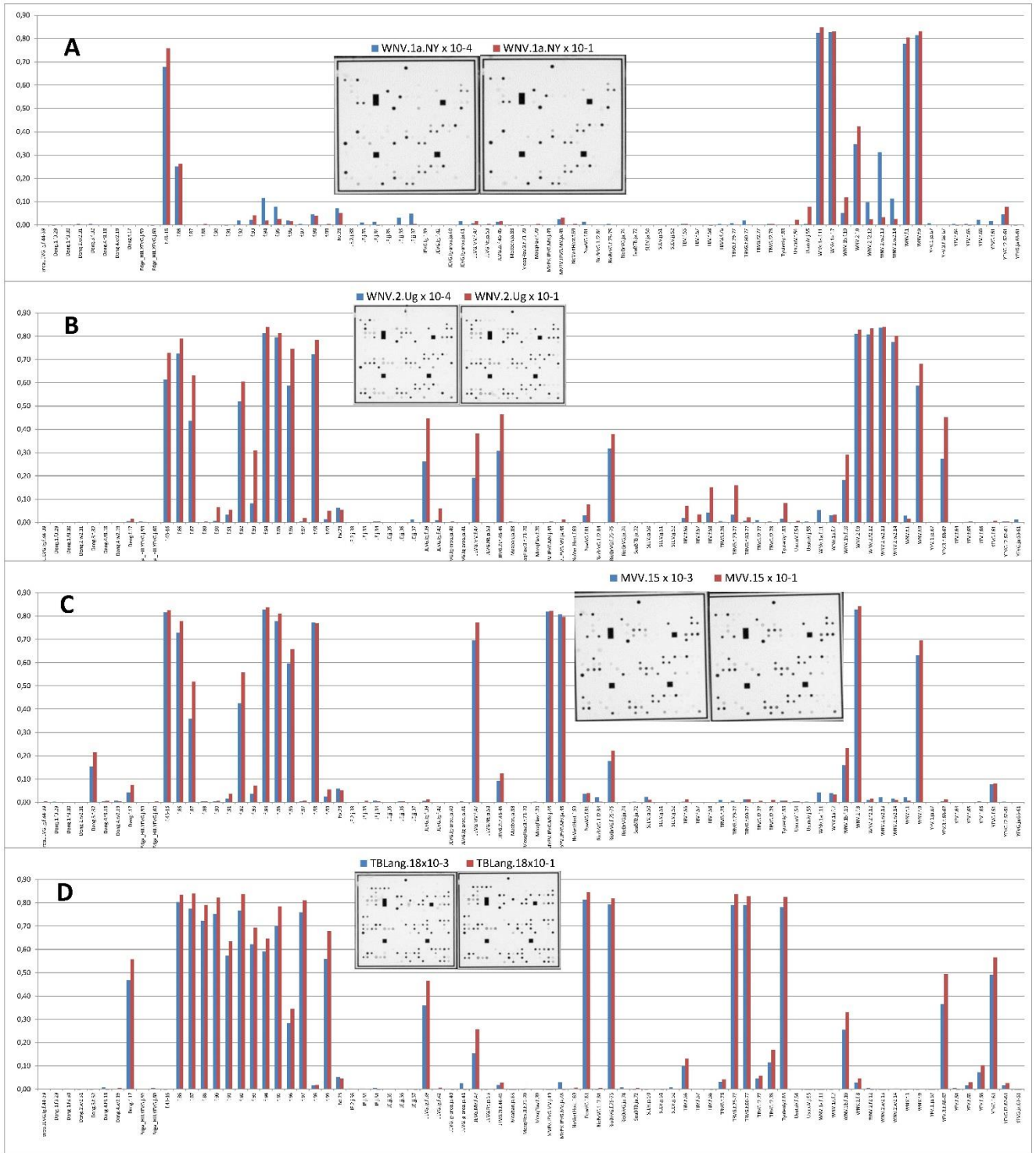


FIGURE 3: Hybridization patterns for different dilutions of viral samples. (cell culture supernatants) (A) West Nile virus lineage 1a, strain NY99 (WNV.1a.NY) - red bars: sample diluted 10⁻¹ and blue bars: diluted 10⁻⁴; (B) West Nile virus lineage 2, strain Uganda 1937 (WNV.2.Ug) - red bars: 10⁻¹ and blue bars: 10⁻⁴; (C) Murray Valley encephalitis virus (MVEV) - red bars: 10⁻¹ and blue bars: 10⁻³; (D) Tick-borne encephalitis virus, strain Langaat (TBLang) - red bars: 10⁻¹ and blue bars: 10⁻³. These data show that hybridization patterns are very sensitive to variation in the viral sequence, while remaining stable in a broad concentration range of viral RNA.

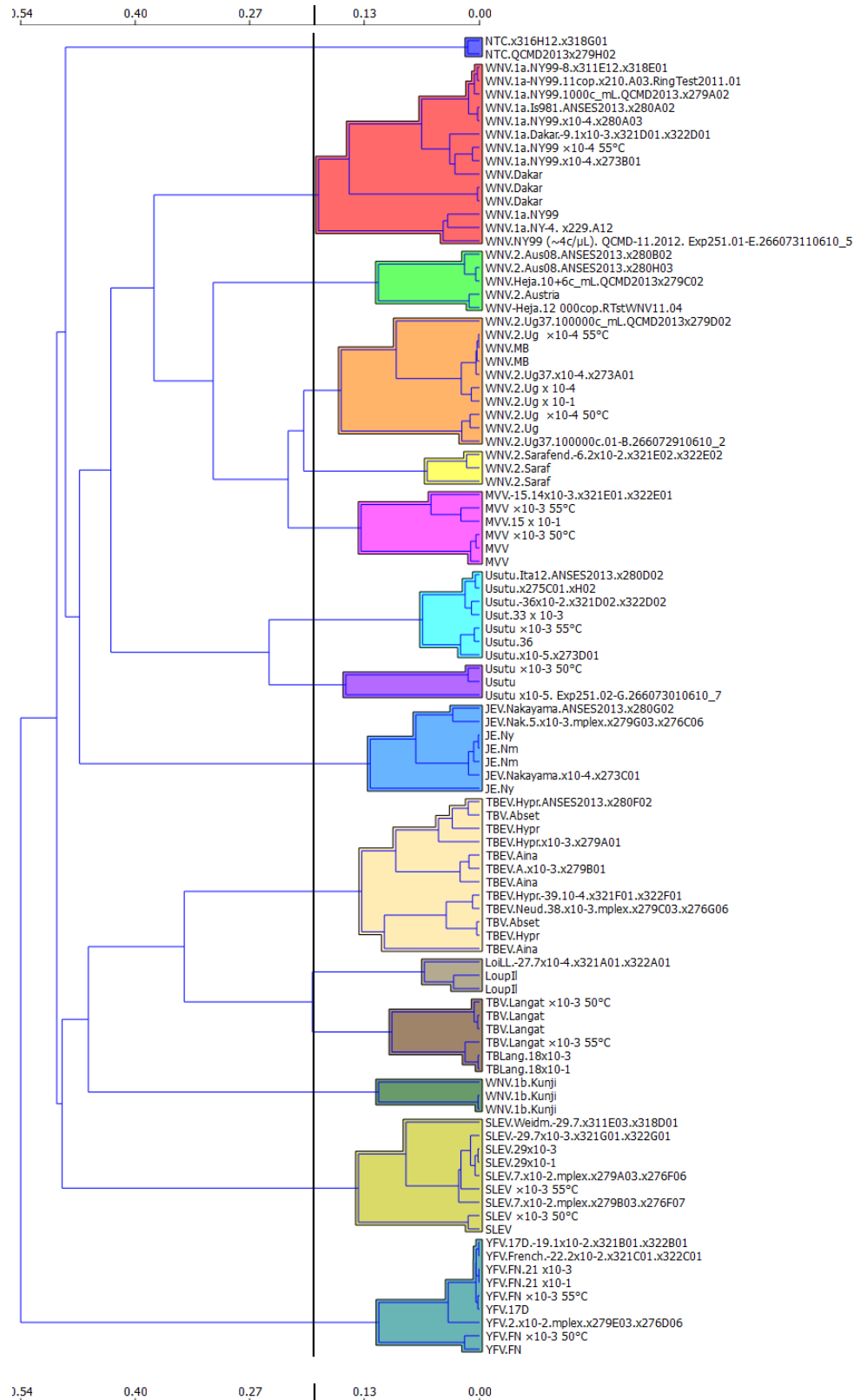


FIGURE 4: Cluster analysis of tested samples. Output of the tree widget in *Orange* software. A cluster analysis was conducted using the distances between the hybridization patterns of experimentally tested samples. This is the preferred method to visualize the identification of viruses in samples that contain only one *Flavivirus*.

In some cases, virus identification is already possible by visually comparing the signal patterns of the bar diagrams (e.g. in FIGURE 3). However, given the complexity of the signals, it was necessary to include a computer-based solution for data processing. *Orange* was used to create a visual program (FIGURE S01) to import the raw data from the *Icono Clust* software, to define a set of experimental standards and identify viral samples by clustering (FIGURE 4). The procedure also permits visualization of mixed of *Flavivirus* infections in heat-map-like graphics (FIGURE 5). This capability represents another major advantage of the microarray compared to direct sequencing. Mixes of RNA from

closely related *Flavivirus* strains were tested to explore the possibility of detection of co-infections in a given sample, or detection of the presence of more than one virus in tested pools. We have selected USUV and WNV-1a and WNV-2 (NY99 and Ug37) viruses to show that it is possible to unambiguously detect each component in a mix even of related viruses, such as lineages 1 and 2 of the WNV (FIGURE 5 b and c).

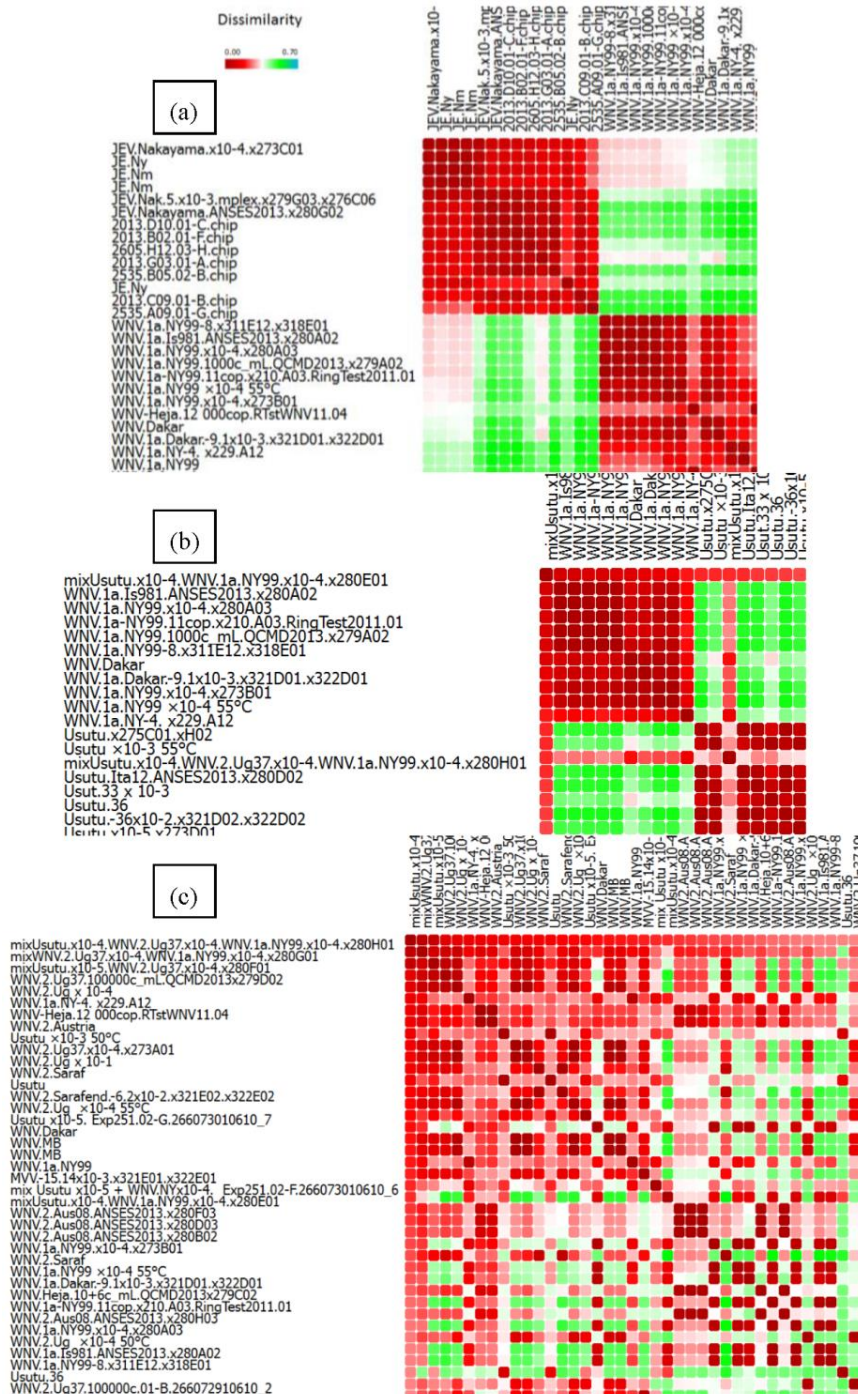


FIGURE 5: Presentation of experimental results processed by the *Orange's* heat-map-like widget showing an all vs. all sample comparison. Calculated distances between intensity patterns of each pair of microarray results obtained from hybridization of sample amplicons are shown, with each cell representing the comparison of two microarray experiments. The dissimilarity color scale is shown at the top. Columns and rows are organized in the same order, making the diagonal an "identity" sample comparison. When one sample is selected by mouse-clicking, the map is immediately reorganized to show this sample in the upper row and left column followed by the most similar samples. (a) The selection of the "JEV.Nakayama.x10-4.x273C01" sample reveals a good separation of JEV samples from all other viral species. (b) The same heap-map after selecting a sample containing a mix of diluted USUV and WNV1a RNA, which is easily recognized by a **framed** mosaic-like map. (c) A more complex mix, including USUV, WNV1 and WNV-2 RNA, was selected showing a more complex pattern that still permits the identification of the components (which can be additionally confirmed by the successive selection of each of the three standard WNV-1, WNV-2 and USUV to check that this mix clusters together with each of them).

The *Flavivirus* RT-qPCR screening was conducted on the 340 mosquito pools from Greece. PCR products of 13 mosquito pools (including all positive specimens, as well as six RT-qPCR-negative pools) underwent microarray analysis (TABLE 2), which revealed the presence of WNV lineage 2 sequences similar to the Austria strain in five of them

(MPGr.01-MPGr.05, for ~1.5% of the total number of pools tested). *Culex* mosquitoes comprised all five WNV-positive pools. One pool yielded a strong positive result (Cq = 21.5), two were of mediums (Cq = 31.6 and 32.5) and two of weak concentrations (Cq = 37.3 and 38.5). Sequencing of these five amplicons revealed identity to WNV isolates goshawk-Hungary/04 DQ116961, Nea Santa-Greece-2010 HQ537483[42] and Italy 2013 KF647248 in four cases, while the fifth (the one with the highest Cq value) was not sequenceable.

TABLE 2: Screening of mosquito pools collected in Greece. Review of the screening of the 340 mosquito pools (50 mosquitoes each) collected in Greece. All RT-qPCR positive and 6 RT-qPCR negatives pools were tested further by microarray and direct sequencing. All other pools tested showed a Cq > 38 and Tm < 78 °C.

Mosquito pool	Cq	Tm (°C)	Microarray	DNA sequencing result
MPGr.01	37.3	81.4	WNV.2 Austria	Nea Santa-Greece-2010 HQ537483
MPGr.02	31.6	81.6	WNV.2 Austria	Nea Santa-Greece-2010 HQ537483
MPGr.03	21.5	81.7	WNV.2 Austria	Nea Santa-Greece-2010 HQ537483
MPGr.04	32.5	81.9	WNV.2 Austria	Nea Santa-Greece-2010 (with 1 mutation)
MPGr.05	38.5	81.6	WNV.2 Austria	Negative
MPGr.06	35.6	80.4	?	Marisma mosquito virus (93% identity)
MPGr.07	39.6	81.0	Culex Flav	New Mosq Culex Flavivirus (78% identity)
MPGr.08	39.8	79.4	Negative	Negative
MPGr.09	41.4	83.0	Negative	Negative
MPGr.10	38.6	81.2	Negative	Negative (residual sequence: <i>Salmonella</i> sp.)
MPGr.11	40.1	82.4	Negative	Negative
MPGr.12	40.6	79.2	Negative	Negative
MPGr.13	39.6	80.7	Negative	Negative (residual sequence: <i>Pseudomonas</i> sp.)

The microarray also detected weak signals of insect-specific flaviviruses (ISFVs) in two of the pools tested. DNA sequencing revealed the presence of sequences with similarities to the isolate HU4528/07 of Marisma mosquito virus (JN603190, 93% nucleotide identity) in one *Aedes* sp. pool (MPGr.06). Additionally, a presumably new mosquito *Flavivirus* (with only 78% identity with GQ165809, the Nakiwogo virus strain from Uganda) was also detected in an *Anopheles* sp. pool (MPGr.07). *Flavivirus* RT-qPCR/microarray screening of the present panel was completed in less than three working days.

Additionally, other specimens were analyzed, including organs (such as brain, heart, liver, lung, kidney, spleen) and blood from falcons [24], pheasants, blackbird, great grey owl, common kingfisher, and nearly 70 other avian species [6], as well as from mice, camels, horses, donkeys, cattle, humans, mosquitoes, and ticks. Only horse samples have caused artifacts in the RT-qPCR, showing low non-specific signals (Cq > 33), thus effectively lowering the detection limit of RT-qPCR alone in horse samples to over 50 copies per reaction (data not shown).

4. Discussion

By designing the present combined RT-qPCR/microarray assay for detection, quantification and identification of flaviviruses, a number of methodological problems have been solved. Using *VisualOligoDeg*, we selected primers and probes for the newly developed assay. The ability of a relatively accurate quantification during the RT-qPCR phase of the assay is one of its major advantages. Technically, the combined RT-qPCR/microarray assay is easy to handle, as only standard experience with real-time PCR and ELISA-like tests is required. We regularly achieved complete testing of samples in one working day, from RNA extraction to final visualization of the tree and interactive heat-map-like graphics.

The present assay permits classification and/or identification up to the (sub-) lineage level, avoiding in most cases the need for sequencing. It has been shown to be as sensitive as species-specific RT-qPCRs, and suitable for broad-range *Flavivirus* screening, as well as a confirmatory assay in both laboratory and field samples. The present study has also demonstrated that the assay can be efficiently used in arbovirus surveillance programs, for rapid screening and discrimination of flaviviruses e.g. in mosquito or animal specimens. In areas where numerous arbovirus strains of different virulence co-circulate, such as in Greece and other European countries, molecular identification of the circulating viruses is a necessity [10]. Especially for flaviviral zoonoses, phylogeography and identification of virulent strains are of utmost importance. The assay is also capable of detecting insect-specific flaviviruses. A report on the presence of ISFVs in *Culex* mosquitoes of Central Macedonia-Greece already exists [43]. Application of RT-qPCR/microarray testing revealed the presence of a virus strain with sequence similarities to Marisma mosquito virus, as well as a presumably novel mosquito

Flavivirus sequence, in the same area of Greece. The detection of these ISFVs via the combined RT-qPCR/microarray protocol extends our knowledge on the presence of mosquito flaviviruses in Greece. The broad range of flaviviruses that are being tested simultaneously in this assay, in combination with its convenience and the minimum time required for obtaining the results, make it a useful tool that potentially can be applied widely for surveillance and epidemiological surveys.

While our assay has been proven to be suitable for diagnostic and research laboratories, it represents an open system that can be further improved. In case of newly emerging pathogenic and genetically distinct strains, primer sequences could be further optimized using newly available genome sequences and/or improving probes of those species that were not in the focus of the present study (possibly DENV 4). The number of probes on the microarray is currently at 84, but could be increased to 500, using the present technology. Since we rely on the specifications of the commercial hybridization kit we do not performed extensive optimization of the incubation times and temperature, and will evaluate the possibility of using future versions of that kit [44]. A limitation of the present assay is the requirement of experimental hybridizations of known samples, which does not allow direct classification of unknowns. Nevertheless, the methodology is robust enough to allow the use of theoretical hybridization patterns, or those obtained in other laboratories, which could be a solution when a particular virus reference is not available. It should be emphasized that the tree constructed by the *Orange* tree widget using the distances between hybridization patterns is not a phylogenetic tree. The graphics just help to group samples according to sequence-based relatedness, and thereby facilitates their identification. The *Orange* software also provides possibilities for classification using machine-learning methods, which have yet to be explored, but they have the potential to significantly improve the accuracy of the final report.

Compared to the diagnostic assays for *Flavivirus* detection published so far, our procedure is distinguished by its high degree of parallelity in detection of a wide range of virus species, strains and their variations, which cannot be achieved through “one-dimensional” RT-PCR assays [21, 32, 45]. Previously published microarray or diagnostic chip approaches [20] lack the ease of operation of the ArrayStrip platform used here and cover only part of the range of viruses that we can identify [12, 14, 19, 46].

5. Conclusion

We have developed a combined RT-qPCR-microarray assay for high-throughput screening and identification of flaviviruses, including mixed infections of different species or strains. Our experience in analyzing field samples (from ticks and mosquito vectors, and from human and animal samples of different sources), shows that the assay allows rapid and highly sensitive screening and identification of *Flavivirus* strains within one day. The assay has helped to overcome limitations in virological diagnosis due to lack of specificity or sensitivity in conventional and real-time RT-PCR protocols. Even if direct sequencing is used as genotyping tool, the developed microarray can be used as a rapid complementary test to detect mixtures of different *Flavivirus* strains. The good performance of the assay was also confirmed, by correctly quantifying and identifying members of the *Flavivirus* genus in samples from international ring trials for quality assessment of nucleic acid amplification tests.

Acknowledgement

Alexandra Chaskopoulou (European Biological Control Laboratory, USDA-ARS, Thessaloniki, Greece) is acknowledged for providing identified mosquitoes.

References

1. E. A. Gould and T. Solomon, "Pathogenic flaviviruses," *Lancet*, vol. 371, no. 9611, pp. 500-509, 2008.
2. M. G. Guzman and E. Harris, "Dengue," *Lancet*, vol. 385, no. 9966, pp. 453-465, 2015.
3. T. P. Monath and P. F. Vasconcelos, "Yellow fever," *J Clin Virol*, vol. 64, pp. 160-173, 2015.
4. R. S. Lanciotti, J. T. Roehrig, V. Deubel, J. Smith, M. Parker, K. Steele, B. Crise, K. E. Volpe, M. B. Crabtree, J. H. Scherret, R. A. Hall, J. S. MacKenzie, C. B. Cropp, B. Panigrahy, E. Ostlund, B. Schmitt, M. Malkinson, C. Banet, J. Weissman, N. Komar, H. M. Savage, W. Stone, T. McNamara and D. J. Gubler, "Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States," *Science*, vol. 286, no. 5448, pp. 2333-2337, 1999.
5. J. N. Hanna, S. A. Ritchie, D. A. Phillips, J. M. Lee, S. L. Hills, A. F. van den Hurk, A. T. Pyke, C. A. Johansen and J. S. Mackenzie, "Japanese encephalitis in north Queensland, Australia, 1998," *Med J Aust*, vol. 170, no. 11, pp. 533-536, 1999.
6. U. Ziegler, H. Jost, K. Muller, D. Fischer, M. Rinder, D. T. Tietze, K. J. Danner, N. Becker, J. Skuballa, H. P. Hamann, S. Bosch, C. Fast, M. Eiden, J. Schmidt-Chanasit and M. H. Groschup, "Epidemic Spread of Usutu Virus in Southwest Germany in 2011 to 2013 and Monitoring of Wild Birds for Usutu and West Nile Viruses," *Vector Borne Zoonotic Dis*, vol. 15, no. 8, pp. 481-488, 2015.

7. M. Marcantonio, A. Rizzoli, M. Metz, R. Rosa, G. Marini, E. Chadwick and M. Neteler, "Identifying the environmental conditions favouring West Nile Virus outbreaks in Europe," *PLoS One*, vol. 10, no. 3, pp. e0121158, 2015.
8. A. S. Fauci and D. M. Morens, "Zika Virus in the Americas--Yet Another Arbovirus Threat," *N Engl J Med*, vol. 374, no. 7, pp. 601-604, 2016.
9. S. C. Chaintoutis, C. I. Dovas, M. Papanastassopoulou, S. Gewehr, K. Danis, C. Beck, S. Lecollinet, V. Antalis, S. Kalaitzopoulou, T. Panagiotopoulos, S. Mourelatos, S. Zientara and O. Papadopoulos, "Evaluation of a West Nile virus surveillance and early warning system in Greece, based on domestic pigeons," *Comp Immunol Microbiol Infect Dis*, vol. 37, no. 2, pp. 131-141, 2014.
10. A. Chaskopoulou, C. I. Dovas, S. C. Chaintoutis, J. Kashefi, P. Koehler and M. Papanastassopoulou, "Detection and early warning of West Nile virus circulation in Central Macedonia, Greece, using sentinel chickens and mosquitoes," *Vector-Borne and Zoonotic Diseases*, vol. 13, no. 10, pp. 723-732, 2013.
11. L. G. Glushakova, A. Bradley, K. M. Bradley, B. W. Alto, S. Hoshika, D. Hutter, N. Sharma, Z. Yang, M. J. Kim and S. A. Benner, "High-throughput multiplexed xMAP Luminex array panel for detection of twenty two medically important mosquito-borne arboviruses based on innovations in synthetic biology," *J Virol Methods*, vol. 214, pp. 60-74, 2015.
12. J. J. Tan, M. Capozzoli, M. Sato, W. Watthanaworawit, C. L. Ling, M. Mauduit, B. Malleret, A. C. Gruner, R. Tan, F. H. Nosten, G. Snounou, L. Renia and L. F. Ng, "An integrated lab-on-chip for rapid identification and simultaneous differentiation of tropical pathogens," *PLoS Negl Trop Dis*, vol. 8, no. 7, pp. e3043, 2014.
13. N. D. Grubaugh, L. N. Petz, V. R. Melanson, S. S. McMenamy, M. J. Turell, L. S. Long, S. E. Pisarcik, A. Kengluocha, B. Jaichapor, M. L. O'Guinn and J. S. Lee, "Evaluation of a field-portable DNA microarray platform and nucleic acid amplification strategies for the detection of arboviruses, arthropods, and bloodmeals," *Am J Trop Med Hyg*, vol. 88, no. 2, pp. 245-253, 2013.
14. N. D. Grubaugh, S. S. McMenamy, M. J. Turell and J. S. Lee, "Multi-gene detection and identification of mosquito-borne RNA viruses using an oligonucleotide microarray," *PLoS Negl Trop Dis*, vol. 7, no. 8, pp. e2349, 2013.
15. D. D. Singh and A. Jain, "Multipurpose instantaneous microarray detection of acute encephalitis causing viruses and their expression profiles," *Curr Microbiol*, vol. 65, no. 3, pp. 290-303, 2012.
16. J. Teo, P. Di Pietro, F. San Biagio, M. Capozzoli, Y. M. Deng, I. Barr, N. Caldwell, K. L. Ong, M. Sato, R. Tan and R. Lin, "VereFlu: an integrated multiplex RT-PCR and microarray assay for rapid detection and identification of human influenza A and B viruses using lab-on-chip technology," *Arch Virol*, vol. 156, no. 8, pp. 1371-1378, 2011.
17. X. Kang, C. Qin, Y. Li, H. Liu, F. Lin, Y. Li, J. Li, Q. Zhu and Y. Yang, "Improvement of the specificity of a pan-viral microarray by using genus-specific oligonucleotides and reduction of interference by host genomes," *J Med Virol*, vol. 83, no. 9, pp. 1624-1630, 2011.
18. K. Xiao-Ping, L. Yong-Qiang, S. Qing-Ge, L. Hong, Z. Qing-Yu and Y. Yin-Hui, "Development of a consensus microarray method for identification of some highly pathogenic viruses," *J Med Virol*, vol. 81, no. 11, pp. 1945-1950, 2009.
19. H. Nordstrom, K. I. Falk, G. Lindegren, M. Mouzavi-Jazi, A. Walden, F. Elgh, P. Nilsson and A. Lundkvist, "DNA microarray technique for detection and identification of seven flaviviruses pathogenic for man," *J Med Virol*, vol. 77, no. 4, pp. 528-540, 2005.
20. J. Korimbocus, N. Scaramozzino, B. Lacroix, J. M. Crance, D. Garin and G. Vernet, "DNA probe array for the simultaneous identification of herpesviruses, enteroviruses, and flaviviruses," *J Clin Microbiol*, vol. 43, no. 8, pp. 3779-3787, 2005.
21. D. Y. Chao, B. S. Davis and G. J. Chang, "Development of multiplex real-time reverse transcriptase PCR assays for detecting eight medically important flaviviruses in mosquitoes," *J Clin Microbiol*, vol. 45, no. 2, pp. 584-589, 2007.
22. M. Eiden, A. Vina-Rodriguez, B. Hoffmann, U. Ziegler and M. H. Groschup, "Two new real-time quantitative reverse transcription polymerase chain reaction assays with unique target sites for the specific and sensitive detection of lineages 1 and 2 West Nile virus strains," *J Vet Diagn Invest*, vol. 22, no. 5, pp. 748-753, 2010.
23. K. Schmidt, M. Keller, B. L. Bader, T. Korytar, S. Finke, U. Ziegler and M. H. Groschup, "Integrins modulate the infection efficiency of West Nile virus into cells," *J Gen Virol*, vol. 94, no. Pt 8, pp. 1723-1733, 2013.
24. U. Ziegler, J. Angenvoort, D. Fischer, C. Fast, M. Eiden, A. V. Rodriguez, S. Revilla-Fernandez, N. Nowotny, J. G. de la Fuente, M. Lierz and M. H. Groschup, "Pathogenesis of West Nile virus lineage 1 and 2 in experimentally infected large falcons," *Vet Microbiol*, vol. 161, no. 3-4, pp. 263-273, 2013.
25. N. Becker, H. Jost, U. Ziegler, M. Eiden, D. Hoper, P. Emmerich, E. Fichet-Calvet, D. U. Ehichioya, C. Czajka, M. Gabriel, B. Hoffmann, M. Beer, K. Tenner-Racz, P. Racz, S. Gunther, M. Wink, S. Bosch, A. Konrad, M. Pfeffer, M. H. Groschup and J. Schmidt-Chanasit, "Epizootic emergence of Usutu virus in wild and captive birds in Germany," *PLoS One*, vol. 7, no. 2, pp. e32604, 2012.
26. H. Jost, A. Bialonski, D. Maus, V. Sambri, M. Eiden, M. H. Groschup, S. Gunther, N. Becker and J. Schmidt-Chanasit, "Isolation of usutu virus in Germany," *Am J Trop Med Hyg*, vol. 85, no. 3, pp. 551-553, 2011.
27. U. Ziegler, A. Skrypnyk, M. Keller, C. Staubach, M. Bezymenyyi, A. M. Damiani, N. Osterrieder and M. H. Groschup, "West Nile virus antibody prevalence in horses of Ukraine," *Viruses*, vol. 5, no. 10, pp. 2469-2482, 2013.
28. B. Hoffmann, K. Depner, H. Schirrmeier and M. Beer, "A universal heterologous internal control system for duplex real-time RT-PCR assays used in a detection system for pestiviruses," *J Virol Methods*, vol. 136, no. 1-2, pp. 200-209, 2006.
29. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins, "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools," *Nucleic Acids Res*, vol. 25, no. 24, pp. 4876-4882, 1997.
30. T. A. Hall, "BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.," *Nucleic Acids Symp Ser.*, vol. 41, pp. 95-98, 1999.
31. C. Domingo, P. Patel, S. Linke, K. Achazi and M. Niedrig, "Molecular diagnosis of flaviviruses," *Future Virology*, vol. 6, no. 9, pp. 1059-1074, 2011.
32. P. Patel, O. Landt, M. Kaiser, O. Faye, T. Koppe, U. Lass, A. A. Sall and M. Niedrig, "Development of one-step quantitative reverse transcription PCR for the rapid detection of flaviviruses," *Virol J*, vol. 10, pp. 58, 2013.

33. M. P. Sanchez-Seco, D. Rosario, C. Domingo, L. Hernandez, K. Valdes, M. G. Guzman and A. Tenorio, "Generic RT-nested-PCR for detection of flaviviruses using degenerated primers and internal control followed by sequencing for specific identification," *Journal of Virological Methods*, vol. 126, no. 1-2, pp. 101-109, 2005.
34. F. Cavrini, M. E. Della Pepa, P. Gaibani, A. M. Pierro, G. Rossini, M. P. Landini and V. Sambri, "A rapid and specific real-time RT-PCR assay to identify Usutu virus in human plasma, serum, and cerebrospinal fluid," *J Clin Virol*, vol. 50, no. 3, pp. 221-223, 2011.
35. M. Schwaiger and P. Cassinotti, "Development of a quantitative real-time RT-PCR assay with internal control for the laboratory detection of tick borne encephalitis virus (TBEV) RNA," *J Clin Virol*, vol. 27, no. 2, pp. 136-145, 2003.
36. A. Ruettinger, J. Nieter, A. Skrypnik, I. Engelmann, A. Ziegler, I. Moser, S. Monecke, R. Ehrlich and K. Sachse, "Rapid spoligotyping of Mycobacterium tuberculosis complex bacteria by use of a microarray system with automatic data processing and assignment," *J Clin Microbiol*, vol. 50, no. 7, pp. 2492-2495, 2012.
37. C. Schnee, S. Schulsse, H. Hotzel, R. D. Ayling, R. A. Nicholas, E. Schubert, M. Heller, R. Ehrlich and K. Sachse, "A novel rapid DNA microarray assay enables identification of 37 Mycoplasma species and highlights multiple Mycoplasma infections," *PLoS One*, vol. 7, no. 3, pp. e33237, 2012.
38. C. Schnee and K. Sachse, "DNA microarray-based detection of multiple pathogens: Mycoplasma spp. and Chlamydia spp," *Methods Mol Biol*, vol. 1247, pp. 193-208, 2015.
39. A. Urisman, K. F. Fischer, C. Y. Chiu, A. L. Kistler, S. Beck, D. Wang and J. L. DeRisi, "E-Predict: a computational strategy for species identification based on observed DNA microarray hybridization patterns," *Genome Biol*, vol. 6, no. 9, pp. R78, 2005.
40. S. C. Chaintoutis, A. Chaskopoulou, T. Chassalevris, P. G. Koehler, M. Papanastassopoulou and C. I. Dovas, "West Nile virus lineage 2 strain in Greece, 2012," *Emerg Infect Dis*, vol. 19, no. 5, pp. 827-829, 2013.
41. S. Linke, W. G. Mackay, C. Scott, P. Wallace and M. Niedrig, "Second external quality assessment of the molecular diagnostic of West Nile virus: are there improvements towards the detection of WNV?," *J Clin Virol*, vol. 52, no. 3, pp. 257-260, 2011.
42. A. Papa, T. Bakonyi, K. Xanthopoulou, A. Vazquez, A. Tenorio and N. Nowotny, "Genetic characterization of West Nile virus lineage 2, Greece, 2010," *Emerg Infect Dis*, vol. 17, no. 5, pp. 920-922, 2011.
43. A. Papa, E. Papadopoulou, S. Kalaitzopoulou, K. Tsioka and S. Mourelatos, "Detection of West Nile virus and insect-specific flavivirus RNA in Culex mosquitoes, central Macedonia, Greece," *Trans R Soc Trop Med Hyg*, vol. 108, no. 9, pp. 555-559, 2014.
44. I. V. Jani, B. Meggi, A. Vubil, N. E. Siteo, N. Bhatt, O. Tobaiwa, J. I. Quevedo, O. Loquiha, J. D. Lehe, L. Vojnov and T. F. Peter, "Evaluation of the Whole-Blood Alere Q NAT Point-of-Care RNA Assay for HIV-1 Viral Load Monitoring in a Primary Health Care Setting in Mozambique," *J Clin Microbiol*, vol. 54, no. 8, pp. 2104-2108, 2016.
45. M. W. Gaunt and E. A. Gould, "Rapid subgroup identification of the flaviviruses using degenerate primer E-gene RT-PCR and site specific restriction enzyme analysis," *J Virol Methods*, vol. 128, no. 1-2, pp. 113-127, 2005.
46. A. J. Foord, V. Boyd, J. R. White, D. T. Williams, A. Colling and H. G. Heine, "Flavivirus detection and differentiation by a microsphere array assay," *J Virol Methods*, vol. 203, pp. 65-72, 2014.

Supporting Information:

TABLE S01: *Flavivirus*-specific oligonucleotide probes, including spot number, position, probe name and probe sequence.

# save NORM results of experiment(s); use only VALID spots (FLAGS=0) # array: Chip_Wildtech-Virology_Mycob-01_130111 (id=216) substance_class: DNA oligonucleotide (id=140) for probe job: all probe jobs # exported by ArielVina at 24-Jun-14					
spot_id	spot_x_pos	spot_y_pos	substance_id	substance_name	sequence
1	0	0	7740	JE.j.34	TGAGCCGAGAGAATTTCAGGAGGTGG
2	.05	0	7741	SLEV.ja.50	GGAGAATTTGGGAAGGCTAAAGGAAGCA
3	.1	0	7742	WNV.2.w.2.13	GACCACTGGCTGGGTAGGAAGAACT
4	.15	0	7743	YFV.1.ja.67	TCAGAGTTTGGGAAAGCAAAGGGAAGC
5	.2	0	7744	JE.2.j.38	TCAGGAGGTGGGGTGAAGGTTCA
6	.25	0	7745	MVEV.JEVG.MV.ja.48	GGAGAATTTGGAAAGGCCAAGGGAAG
7	.3	0	7746	WNV.1a.f.11	TGGTTCATGTGGCTCGGAGCTCG
8	.35	0	7747	TBV.f.58	ACATGTGGCTGGGGAGCCGCTT
9	.4	0	7725	Deng.4.f3.18	CCAAGGGAAGCCGAGCAATCTGGTA
10	.45	0	7726	JEV.G.MV.f.47	GGAAGCAGAGCCATCTGGTACATGTG
11	.5	0	7727	WNV.f.1	GCCATTTGGTTCATGTGGCTCGG
12	.55	0	7728	RioBrVG.0.ja.85	TGCCAAGGGTAGCCGTACCATATGGTA
13	.6	0	7729	UsutuV.j.55	CTTCTGAATGAGGACCATTGGTTAGGAAG
14	.65	0	7730	TBVG.f.76	AGCCGTGCCATCTGGTACATGTGGTT
15	.7	0	7731	f.87	GCCATCTGGTACATGTGGCTGGG
16	.75	0	7732	Deng.1.f3.29	GCAAAAGGAAGTCGTGCAATATGGTACATG
17	.8	0	7733	Deng.3.f.32	ATATGGTACATGTGGTTGGGAGCCAGG
18	.85	0	7734	JEV.G.Nt.ja.53	GAATTTGGAAAAGCAAAGGAAGCAGGC
19	.9	0	7735	PowVG.f.81	AGCAGGGCCATCTGGTACATGTGG

20	.95	0	7736	f.97	GGGCCATCTGGTACATGTGGCTTGG
21	1	0	7737	f.96	GGGCCATCTGGTTCATGTGGCTTGG
22	0	.05	7738	f.95	GGGCCATCTGGTTCATGTGGCTGGG
23	.05	.05	7739	f.94	GAGCCATCTGGTTCATGTGGCTGGG
24	.1	.05	7710	Deng.2.w2.31	GGAGTGAAGGAGAAGGGCTGC
25	.15	.05	7711	JEVG.lg.f.39	ATCTGGTACATGTGGCTGGGGGC
26	.2	.05	7712	WNV.1b.f.10	GCAGGGCCATCTGGTTTATGTGGCT
27	.25	.05	7713	YFV.3.f.68-67	ATCTGGTACATGTGGCTGGGGGC
28	.3	.05	7714	TBV.f.57	ATGTGGCTGGGGAGCCGCTT
29	.35	.05	7715	YFVG.ja.63-61	GAAACCATCTGAATTTGGGAAAGCCAAAGG
30	.4	.05	7716	WNV.f.9	GCAGAGCCATTTGGTTCATGTGGCT
31	.45	.05	7717	f.98	GAGCCATCTGGTTCATGTGGCTTGG
32	.5	.05	7718	f.92	GGGCCATCTGGTACATGTGGCTGGG
33	.55	.05	7719	YFVG.f2.62-61	TTTCTGGAGTTTGAGGCACTGGG
34	.6	.05	7720	MosqFlav.3.f.71-70	TCCAGAACGATATGGTATATGTGGCTAGG
35	.65	.05	7721	Tyulenyi.f.83	CCATTTGGTACATGTGGCTGGGCAG
36	.7	.05	7722	Meaban.ja.83	GGAGAGTTTGGCAAGGCAAAAGGAAG
37	.75	.05	7723	Edge_Hill.YFVG.j.60	GTTTGGGAGAGAGAACTCATTGGCTGG
38	.8	.05	7724	NoVertHost.f.69	GAACTATCTGGTACATGTGGCTCGGAAG
39	.85	.05	7692	TBVG.f2.77	CTGGAATTTGAGGCACTGGGTTTCCTC
40	.9	.05	7693	WNV.2.f2.12	TTCTGGAGTTTGAAGCTCTCGGATT
41	.95	.05	7694	YFVG.f.61	GCCATCTGGTACATGTGGCTAGGAGC
42	1	.05	7695	TBV.f.56	TGGTACATGTGGCTGGGGAGCCG
43	0	.1	7696	WNV.1a.f.7	GAGCCATTTGGTTCATGTGGCTCGG
44	.05	.1	7697	RioBrVG.f.75-75	GGACCATCTGGTACATGTGGCTGGG
45	.1	.1	7698	TBVG.f.79-77	AAGGGCAGCAGGGCCATCTGG
46	.15	.1	7699	JEVG.Zi.f.46-45	GCCATCTGGTTCATGTGGTTAGGGGC
47	.2	.1	7700	SLEV.jj.52	GGATACATCCTGCAAGAAATCTCCCAAAT
48	.25	.1	7701	JEVG.lg-aroa.ja.41	GGAGAGTTTGGAAAAGCGAAAGGAAGCC
49	.3	.1	7702	aroa.JEVG.lg.f.44-39	ATATGGTACATGTGGCTAGGCGCAC
50	.35	.1	7703	JEVG.lg.f.42	TGGTACATGTGGCTGGGGGC
51	.4	.1	7704	f.86	GAGCCATCTGGTACATGTGGCTGGG
52	.45	.1	7705	RioBrVG.1.f2.84	TAGAATTTGAAAGCTTTGGCTTTCTGAA
53	.5	.1	7706	TBVG.f.80-77	CAGGGCCATCTGGTACATGTGGCT
54	.55	.1	7707	TBVG.f2.78	GAATTTGAGGCACTGGGTTTCCTCAATGA
55	.6	.1	7708	JEVG.Zi.f.45	GAGCCATCTGGTACATGTGGCTGGG
56	.65	.1	7709	NoVertHost.0.f.73	GGACCATCTGGTACATGTGGCTCGG
57	.7	.1	7671	Deng.4.w2.19	TCATGGAGTGGAGTGAAGGGGAAG
58	.75	.1	7672	WNV.2.w2.14	GGTAGGAAGAACTCAGGAGGAGGAG
59	.8	.1	7673	UsutuV.f.54	GCCATATGGTTCATGTGGCTAGGCG
60	.85	.1	7674	Deng.f.17	GCCATATGGTACATGTGGCTGGGAG
61	.9	.1	7675	SeaBTB.ja.72	GGAGAGTTTGGCAAGGCCAAAGGAAG
62	.95	.1	7676	JE.j.33	TGGCTGAGCCGAGAGAATTCAGGAG
63	1	.1	7677	Deng.1.f3.30	AAAGGAAGTCGTGCAATATGGTACATGTGG
64	0	.15	7678	MVEV.JEVG.MV.j.49	ATTGGATGAGTAGAGAGAATTCAGGAGGAG
65	.05	.15	7679	JE.jj.36	ATAGCAGGAAAGCAAGGAGGGAAAATGTA
66	.1	.15	7680	SLEV.jj.51	AAACTCGGATACATCCTGCAAGAAATCTC
67	.15	.15	7681	MosqFlav.f.70	TCGAGAACTATCTGGTACATGTGGTT
68	.2	.15	7682	JEVG.lg-aroa.ja.40	GGAGAATTTGGAAAAGCGAAAGGAAGCC
69	.25	.15	7683	JE.jj.37	TCCGTGACATAGCAGGAAAGCAAGGAG
70	.3	.15	7684	JE.jj.35	GGATACACCCTCCGTGACATAGCAG
71	.35	.15	7685	YFV.f.66	CGTGCCATATGGTACATGTGGCTGG
72	.4	.15	7686	Edge_Hill.YFVG.j.59	GACCATTGGTTTGGGAGAGAGAACTCATTG
73	.45	.15	7687	YFV.f.65	CGTGCCATATGGTATATGTGGCTGGG
74	.5	.15	7688	RioBrVG.ja.74	GAGAAAAGAAACCAGGTGAGTTTGGAGT
75	.55	.15	7689	YFV.f.64	GGAAGCCGTGCCATATGGTATATGTGG
76	.6	.15	7690	WNV.2.f.8	GCAGAGCCATCTGGTTCATGTGGCT
77	.65	.15	7691	f.88	AGCAGGGCCATATGGTACATGTGG
78	.7	.15	7650	f.43-16	GGAAGCAGAGCCATATGGTTCATGTGG
79	.75	.15	7651	fw.23	TACAACATGATGGGAAAGAGAGAGAAGAA

80	.8	.15	7652	f.43-16-2	GGAAGCAGAGCCATATGGTTCATGTGG
81	.85	.15	7653	f.93	GAGCCATTTGGTACATGTGGCTGGG
82	.9	.15	7654	f.99	GAGCCATCTGGTACATGTGGCTTGG
83	.95	.15	7655	f.90	GGCCATATGGTACATGTGGCTGGG
84	1	.15	7656	f.91	GAGCCATATGGTACATGTGGCTGGG

TABLE S02: Quantification of flaviviruses in samples from the QCMD-2013 ring trial using the new Flavivirus RT-qPCR and a calibrated WNV NY99 RNA standard curve.

Cop/mL. provided by the organizers		Cq, 2013-04-29		Copies/mL		Tm
WNV-NY99	12.000.000 Cop/mL	21,05	21,34	5.100.000 ± 610.000 Cop/mL		81,9 °C
WNV-NY99	1.200.000 Cop/mL	25,05	25,17	490.000 ± 25.000 Cop/mL		81,9 °C
WNV-NY99	120.000 Cop/mL	29,12	29,03	47.000 ± 1.800 Cop/mL		81,8 °C
WNV-NY99	120.000 Cop/mL	29,99	29,75	29.000 ± 3.000 Cop/mL		81,8 °C
WNV-NY99	12.000 Cop/mL	33,51	33,46	3.400 ± 71 Cop/mL		81,9 °C
WNV-NY99	1.200 Cop/mL	37,15	37,61	340 ± 65 Cop/mL		81,8 °C
WNV-Heja	7.300.000 Cop/mL	21,04	20,32	7.000.000 ± 2.100.000 Cop/mL		80,9 °C
WNV-Heja	730.000 Cop/mL	24,22	24,27	830.000 ± 18.000 Cop/mL		80,9 °C
WNV-Ug37	110.000 Cop/mL	21,84	21,33	4.100.000 ± 860.000 Cop/mL		81,2 °C
noWNV Flav-each	1.000.000 Cop/mL	23,55	23,54	1.300.000 ± 3.300 Cop/mL		80,2 °C
noWNV Flav-each	1.000.000 Cop/mL	23,61	23,54	1.200.000 ± 36.000 Cop/mL		80,1 °C
Neg-VTM	0 Cop/mL	N/A	N/A			77,1 °C
Our WNV RNA standard curve	14.000.000 Cop/mL	19,33				82,4 °C
	1.400.000 Cop/mL	23,11				82,3 °C
	140.000 Cop/mL	28,02				82,1 °C
	14.000 Cop/mL	30,59				82,0 °C
	0 Cop/mL	N/A	N/A			72,7 °C

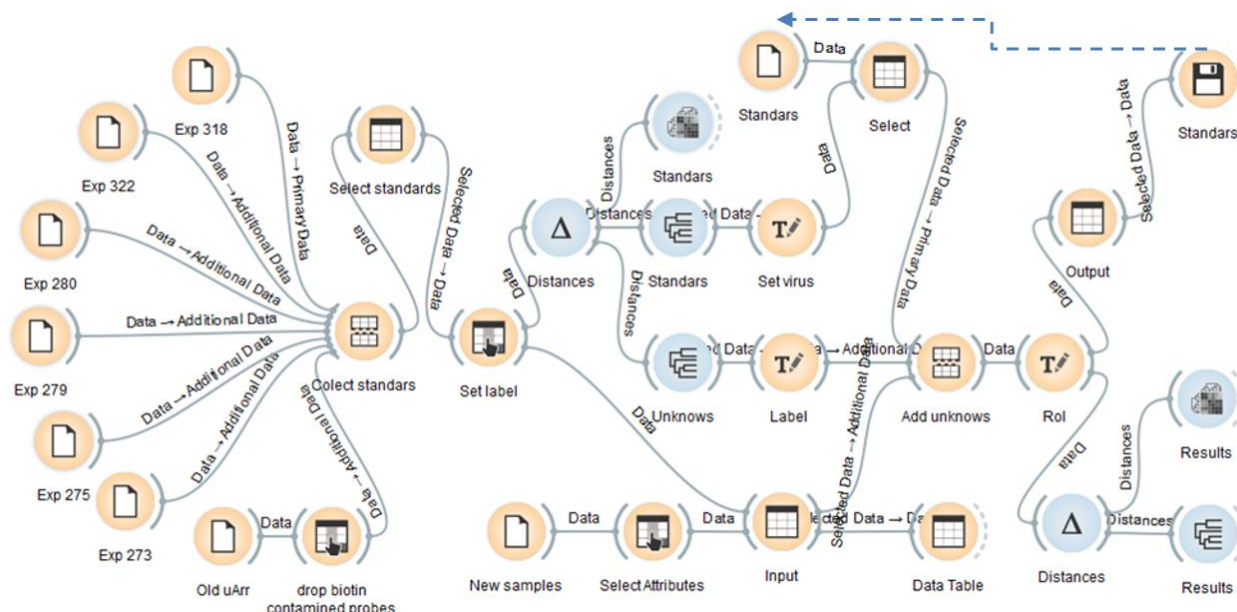


FIGURE S01: The visual program *PanFlavExpStdSampl*. Schematic workflow - the visual program *PanFlavExpStdSampl* for the Orange software package. The scheme permits an interactive import of the experiments used as standards (known samples), which are being subsequently used to identify new or unknown samples.

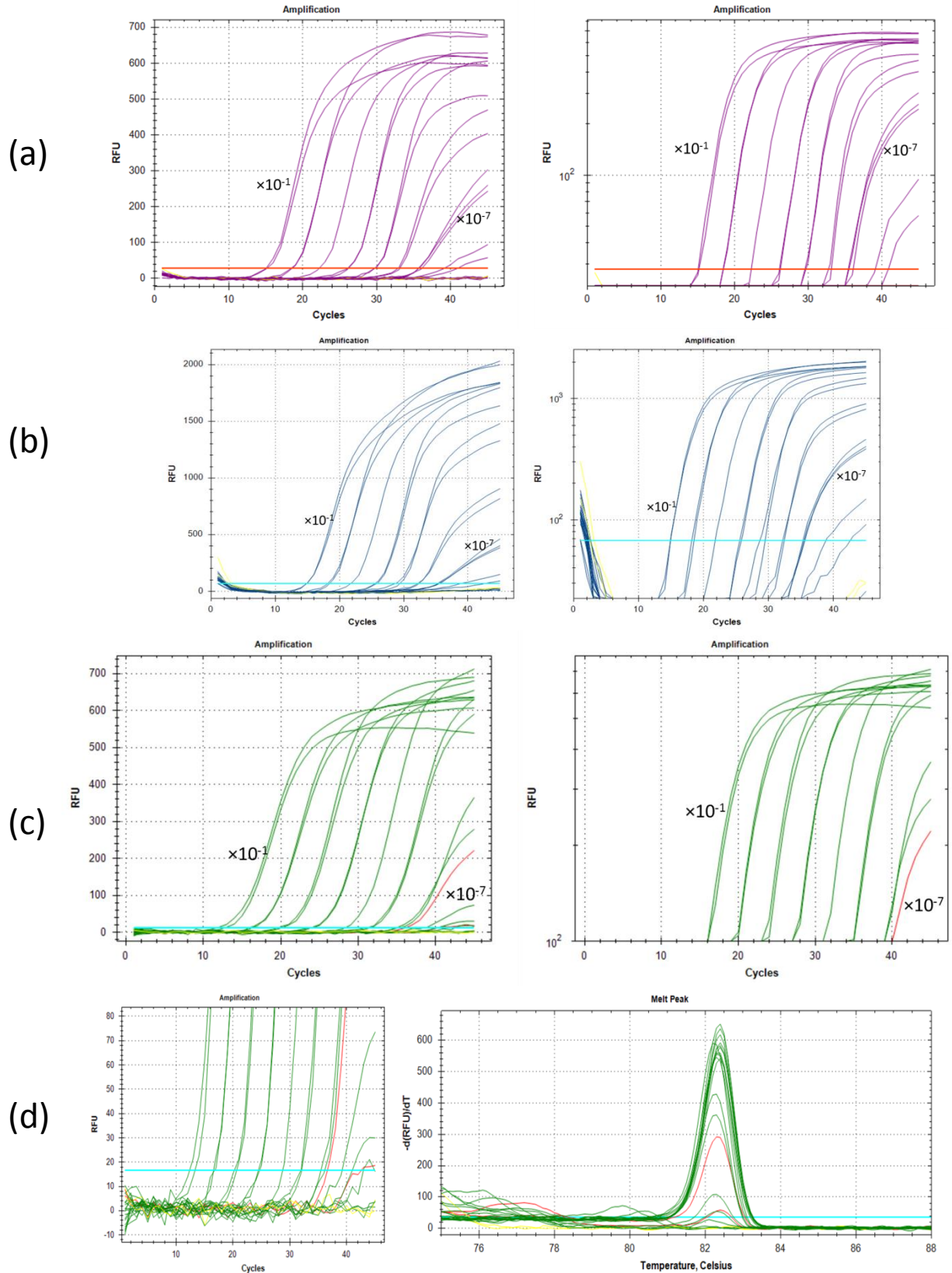


FIGURE S02: WNV titration. (a) WNV titration with specific RT-qPCR (Eiden, 2010). WNV.NY99 RNA new titration curve (right – logarithm scale). WNV $\times 10^{-7}$ is correctly three time positive. WNV $\times 10^{-8}$ is once positive and once extremely low positive, and once negative. WNV $\times 10^{-9}$ and WNV $\times 10^{-10}$ are negative. (b) WNV titration with specific RT-qPCR (Eiden, 2010). WNV.NY99 RNA new titration curve (right – logarithm scale). WNV $\times 10^{-7}$ is correctly three time positive. WNV $\times 10^{-8}$ is once positive and once extremely low positive, and once negative. WNV $\times 10^{-9}$ and WNV $\times 10^{-10}$ are negative. (c) WNV titration with Flavivirus RT-qPCR. WNV.NY99 RNA new titration curve (right – logarithm scale). WNV $\times 10^{-7}$ is correctly three time positive. WNV $\times 10^{-8}$ is three-time low positive. WNV $\times 10^{-9}$ is once extremely low positive, and once negative. WNV $\times 10^{-10}$ are negative. See the corresponding melting curves in panel d. (d) WNV titration with Flavivirus RT-qPCR (melting curves). WNV.NY99 RNA new titration curve (right – logarithm scale). WNV $\times 10^{-7}$ is correctly three time positive. WNV $\times 10^{-8}$ is three-time low positive. WNV $\times 10^{-9}$ is once extremely low positive, and once negative. WNV $\times 10^{-10}$ are negative.

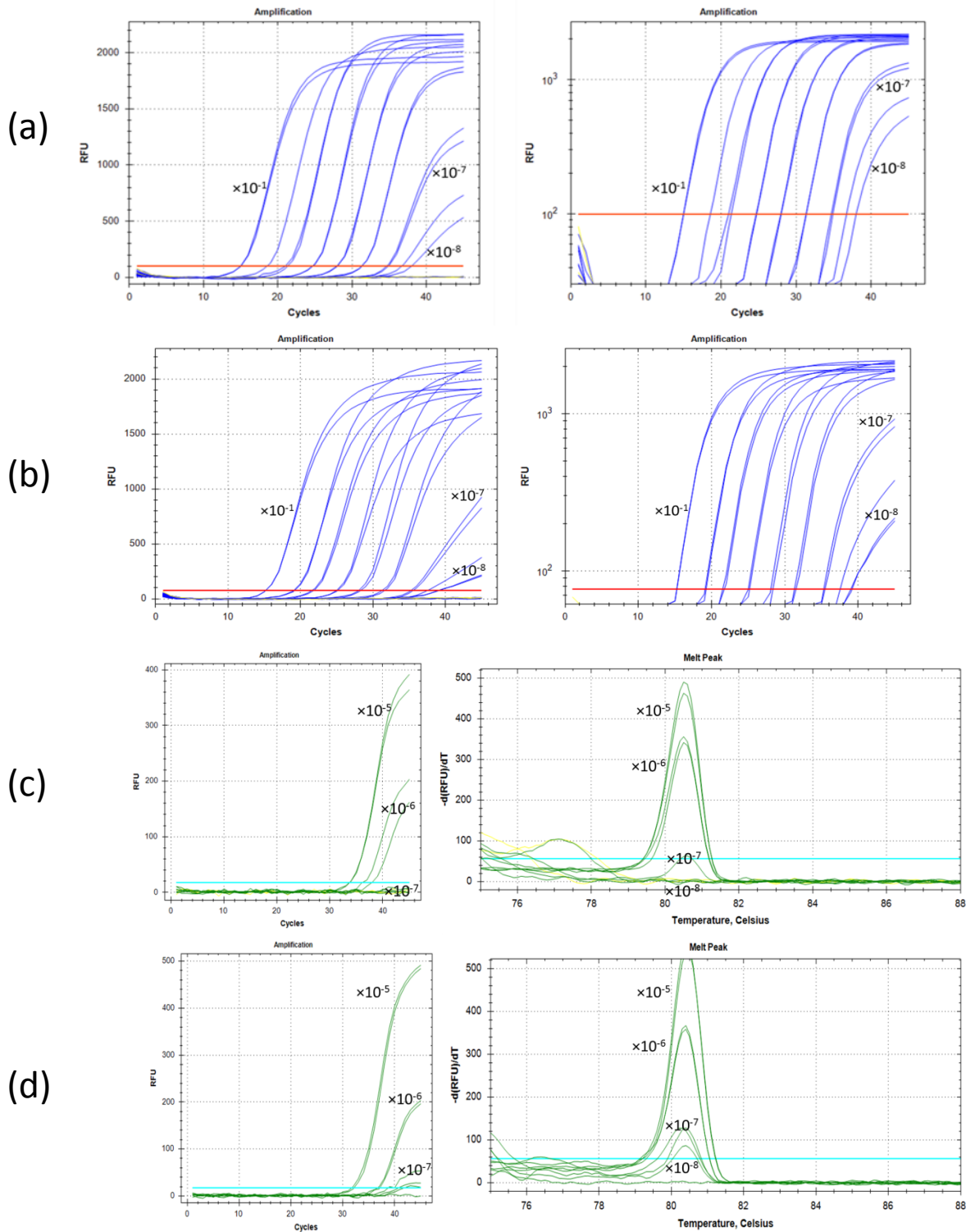


FIGURE S03: USUV titration. (a) USUV titration with specific RT-qPCR (Cavrini, 2011). USUV RNA new titration curve (right – logarithm scale). USUV $\times 10^{-7}$ is correctly twice positive. WNV $\times 10^{-8}$ is once low positive and once extremely low positive. WNV $\times 10^{-9}$ and WNV $\times 10^{-10}$ are negative. (b) USUV titration with specific RT-qPCR (Jost, 2011). USUV RNA new titration curve (right – logarithm scale). USUV $\times 10^{-7}$ is twice positive. USUV $\times 10^{-8}$ is once low positive and once extremely low positive. WNV $\times 10^{-9}$ is once extremely low positive and once negative. USUV $\times 10^{-10}$ are negative. (c) USUV titration with Flavivirus RT-qPCR. USUV RNA new titration curve (right – logarithm scale). USUV $\times 10^{-6}$ is correctly twice positive. USUV $\times 10^{-7}$ is once extremely low positive and once negative. USUV $\times 10^{-8}$ are negative. It need more than 30 copies/reaction or 5 copies/ μ L to be detected. Both USUV specific PCR used were possibly even more sensitive (Figure S3 (a) and (b)), with Ct of 28 for the 10^{-5} USUV RNA dilution, 32 for the 10^{-6} , 35 for the 10^{-7} and 38-40 for the 10^{-8} dilution. This is more efficient, with Ct up to 8 cycles smaller than with the new Flavivirus PCR, which showed in this experiments Ct of 33 for the 10^{-5} USUV RNA dilution and 36-37 for the 10^{-6} dilution. We observed using the smelt curve (Figure S3 (c)) that the sensitivity (last detected dilution) was at best 10 times smaller (with a low signal for the 10^{-7} dilution), to allow maximal detection of an estimated 30 copies/reaction. (d) SLEV titration with Flavivirus RT-PCR. SLEV RNA new titration curve (right – logarithm scale). SLEV $\times 10^{-6}$ is correctly twice positive. SLEV $\times 10^{-7}$ and once SLEV $\times 10^{-8}$ are extremely low positive. For TBEV.A and TBEV.H viruses (control RNA) the specific PCR got Ct of 28 and 26 while the Flavivirus PCR got 25.4 and 23.7 respectively.

Annex

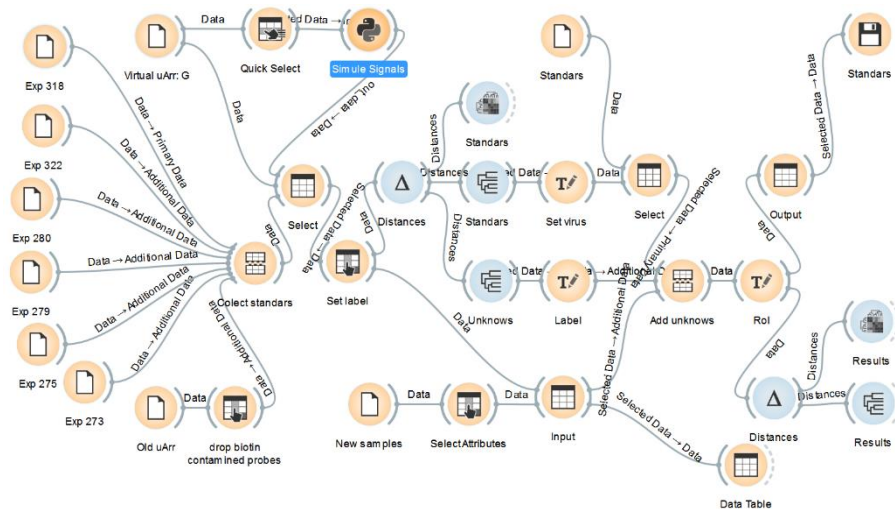


Figure 5. *PanFlavExpVirtStdSampl* is a modification of the *Orange* visual program *PanFlavExpStdSampl* which permit the import of modeled G from *DNAHydrid* virtual hybridizations and convert it into compatible modeled intensity signals. These virtual hybridization are then mixed with experimental and used as standard to classify unknown samples.

Authors contribution to publications

Publication I (Chapter 2): **Two New Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction Assays with Unique Target Sites for the Specific and Sensitive Detection of Lineages 1 and 2 *West Nile Virus* Strains.** Eiden, M.: envisioned the paper; supervised experiments and directly contributed to them; analyzed the results; wrote the manuscript. Vina-Rodriguez, A.: designed and experimentally established assay 1, including the synthetic RNA control; extracted RNA from samples; prepared the standard curves; experimentally compared the assays; contributed to the writing of results, tables and figure. Hoffmann, B.: designed and experimentally established assay 2. Ziegler, Z.: prepared and provided the virus samples. Groschup, MH.: supervised and coordinated the whole study, corrected the manuscript. All authors discussed the results and approved the manuscript.

Publication II (Chapter 3): **A Quantitative Real-Time RT-PCR Assay for the Detection of Venezuelan Equine Encephalitis Virus Utilizing a Universal Alphavirus Control RNA.** Vina-Rodriguez, A.: designed and analyzed MSA, probes, primers and the set of synthetic RNA for experimental demonstration of detection of all known variants; contributed to the writing of the manuscript. Eiden, M.: designed the synthetic calibrator; supervised the experiments; analyzed the results; took part in the writing of the manuscript. Keller, M.: envisioned the paper; carried out the initial experiments; wrote the first version of the manuscript. Hinrichs, W.: helped in the analysis of the results; corrected the manuscript. Groschup, MH.: supervised and coordinated the study, corrected the manuscript. All authors discussed the results and approved the manuscript.

Publication III (Chapter 4): **Ngari Virus in Goats during Rift Valley Fever Outbreak, Mauritania, 2010.** Eiden, M.: envisioned the paper; supervised experiments and contributed to them; analyzed the results; wrote the manuscript. Vina-Rodriguez, A.: experimentally established the PCR procedure and participated on the screening; sequenced the amplicons and first spotted the Ngari isolate. Jäckel, S.: processed the samples. El Mamy, BO., Isselmou, K., Unger, H., Doumbia, B.: provided the field samples. Ziegler, U.: carried out virus replication. Balkema-Buschmann, A.: analyzed and interpreted the results. Höper, D.: carried out full genome sequencing; provided NGS data. Groschup, MH.: supervised and coordinated the study, corrected the manuscript. All authors discussed the results and approved the manuscript.

Publication IV (Chapter 5): **Hepatitis E Virus Genotype 3 Diversity: Phylogenetic Analysis and Presence of Subtype 3b in Wild Boar in Europe.** Vina-Rodriguez, A.: Conceived and designed all experiments including phylogenetic analysis and subtyping; designed MSA, primers, probes, protocols and synthetic and viral controls; extracted RNA from samples; carried out screening and genotyping; updated the subtyping scheme; wrote the manuscript. Schlosser, J.: collected

and processed samples, participated in screening and sequencing, corrected the manuscript. Becher, D. and Kaden V.: collected and processed samples. Groschup, MH.: supervised and coordinated the study, corrected the manuscript. Eiden, M.: supervised all experiments and directly contributed to them; analyzed the results; collaborated in the writing of the manuscript. All authors discussed the results and approved the manuscript.

Publication V (Chapter 6): **Detection of Hepatitis E Virus in Archived Rabbit Serum Samples, Germany 1989.** Eiden, M.: supervised all experiments and directly contributed to them; analyzed the results; wrote the manuscript. Vina-Rodriguez, A.: participated in the screening; processed the sequences, and performed the phylogenetic analysis. Schlosser, J.: collected and processed samples, participated in sequencing, corrected the manuscript. Schirrmeier, H.: provided the samples. Groschup, MH.: supervised and coordinated the study, corrected the manuscript. All authors discussed the results and approved the manuscript.

Publication VI (Chapter 7): **A Novel Pan-Flavivirus Detection and Identification Assay, Based on RT-qPCR and Microarray.** Vina-Rodriguez, A.: designed MSA, primers, probes and protocols for PCR and microarray; designed and performed the PCR, microarray and sequencing experiments; designed and coded the solutions for processing and interpretation of microarray results; wrote the manuscript. Sachse, K.: envisioned the paper; participated in the original design of the microarray; coordinated the project; participated in the analysis of the results and the writing of the manuscript. Ziegler, Z. and Keller, M.: infected cell cultures and provided samples; participated in RNA extraction. Chaintoutis, SC.: provided mosquito samples; participated in experiments; analyzed results; collaborated in the writing of the manuscript. Groschup, MH.: supervised and coordinated the study, corrected the manuscript. Eiden, M.: supervised all experiments and directly contributed to them; analyzed the results; edited and participated in the writing of the manuscript. All authors discussed the results and approved the manuscript.

In the previous list the contributions of the author of this thesis (Vina-Rodriguez, A.) are better detailed than the contributions of the rest of the authors. Additionally, he designed and coded the new software tools, including the visual programs for Orange, *RobotEvo*, *VisualOligoDeg* and *ThDy_DNAHybrid* and maintained the source code repositories and the corresponding documentation.

In agreement:

Ariel Vina-Rodriguez

Professor, Dr. rer. nat. Winfried Hinrichs

Eigenständigkeitserklärung

(Diese Erklärung ist zu unterschreiben und in die Dissertationen einzubinden.)

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Ernst-Moritz-Arndt-Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Unterschrift des Promovenden

Curriculum vitae

Name: Ariel Viña Rodríguez.
(ORCID.org/0000-0003-4632-6287)

WWW: ResearchGate, GitHub-qPCR4vir, Google Scholar

Citizenship: Cuban

Date of birth: September 29, 1968; **Sex:** Male

Marital status: Married, has children



Professional career:

2017 Writing papers and the PhD thesis.

2008-2016 **Friedrich-Loeffler-Institut**, Institute of Novel and Emerging Infectious Diseases (FLI-INNT, Greifswald-Riems, Germany). RNA-virus detection.

1993-2005 **Center for Genetic Engineering and Biotechnology** (CIGB, Havana, Cuba). HCV vaccine Department. "Added Researcher" (Investigador Agregado).

1998-2005 Collaboration in the foundation of the Molecular Biology Laboratory of the Immuno Assay Center (CIE, Havana, Cuba), creating a diagnostic kit for detection of HCV viral load.

1996-1997 Molecular laboratory of Liver Unit of the Hospital Clinic i Provincial de Barcelona: AECl-Spanish Agency for International Cooperation- project: immune response against HCV.

1993-2005 Annual Awards of the Cuban Academy of Science: 1993- "Characterization of the infection by the Hepatitis C Virus in Cuba". 2000- "Definition of a vaccine candidate against the HCV as result of the pre-clinical studies". 2001- "Obtaining and assembling of particles similar to virus of the HCV capsid and their applications". 2002- "Qualitative UMELOSA®HCV. An RT-PCR-in plate hybridization kit for RNA-HCV detection". 2005- "Expression and processing of hepatitis C virus structural proteins in Pichia pastoris yeast."

2005 General Organizer and webmaster of the I International Conference "Immunotherapy in Viral Hepatitis B and C", Varadero 2005.

Education:

2007-2008 German as foreign language at the Volkshochschule-Greifswald with an B1 level.

1990-1993 Diploma/Master of Science thesis: "The Fractals properties of genetics texts". Laboratory of Computer Analysis of Biopolymers. Institute "Engelgart" of Molecular Biology of the Russian Academy of Sciences.

1987-1993 **Moscow State University "Lomonosov" (Russia), Faculty of Physics, Chair of Biophysics.**

1983-1986 **National Special High School of Exact Sciences, Havana, Cuba. (IPECE – "Humboldt 7").** Biology and Physics specialties. Firsts places in National Competitions of Physics and computing/programming.

Other publications:

17. *Toxicology and biodistribution study of CIGB-230, a DNA vaccine against hepatitis C virus.* Bacardí D, Amador-Cañizares Y, Cosme K, Urquiza D, Suárez J, Marante J, Viña A, Vázquez A, Concepción J, Pupo M, Aldana L, Soria Y, Romero J, Madrigal R, Martínez L, Hernández L, González I, Dueñas-Carrera S. Human & experimental toxicology. 2009; 28(8):479-91. [PubMed](#)
18. *New alternatives for the development of vaccine preparations: Contributions to the knowledge on the interaction of recombinant protein viral antigens with nucleic acids.* Santiago Dueñas-Carrera, Liz Alvarez-Lajonchere, Alexis Musacchio, Nelson Acosta-Rivero, Viviana Falcón, Gillian Martínez, Yalena Amador-Cañizares, Ivis Guerra, Julio C Alvarez-Obregón, Angel Pérez, Marbelis Linares, Miladys Limonta, Odalis Ruiz, Dania Bacardí, [Ariel Vina-Rodriguez](#), Juan Morales-Grillo, Dinorah Torres, Gabriel Márquez, Jeny Marante, Maria C de la Rosa, Maribel Vega, Julio C Aguilar, Yordanka Soria, Dagmara Pichardo, Eduardo Martínez, Verena Muzio, Mariela Vázquez, Boris Acevedo Castro Gerardo Guillen, Cosme Karelía, Marisel Quintana, Pedro Antonio Lopez-Saura, Luis Herrera. Biotecnología Aplicada 24(3-4):311-4 · July 2007
19. *HCV core protein localizes in the nuclei of nonparenchymal liver cells from chronically HCV-infected patients.* Falcon V, Acosta-Rivero N, Shibayama M, Chinea G, Gavilondo JV, de la Rosa MC, Menendez I, Gra B, Dueñas-Carrera S, [Vina A](#), Garcia W, Gonzalez-Bravo M, Luna-Munoz J, Miranda-Sanchez M, Morales-Grillo J, Kouri J, Tsutsumi V. Biochem Biophys Res Commun. 2005;329(4):1320-8. [PubMed](#)
20. *Ultrastructural evidences of HCV infection in hepatocytes of chronically HCV-infected patients.* Falcon V, Acosta-Rivero N, Chinea G, Gavilondo J, de la Rosa MC, Menendez I, Duenas-Carrera S, [Vina A](#), Garcia W, Gra B, Noa M, Reytor E, Barcelo MT, Alvarez F, Morales-Grillo J. Biochem Biophys Res Commun 2003; 305(4):1085-90. [PubMed](#)
21. *Expression and processing of hepatitis C virus structural proteins in Pichia pastoris yeast.* Martinez-Donato G, Acosta-Rivero N, Morales-Grillo J, Musacchio A, [Vina A](#), Alvarez C, Figueroa N, Guerra I, Garcia J, Varas L, Muzio V, Duenas-Carrera S. Biochem Biophys Res Commun 2006; 342(2):625-31. [PubMed](#)
22. *Antigenicity and immunogenicity of the hepatitis C virus envelope E2 protein.* Gillian Martínez-Donato, Santiago Dueñas-Carrera, Liz Alvarez-Lajonchere, Juan Morales, Nelson Acosta-Rivero, Eduardo Martínez, Ariel Viña, Ivis Guerra, Angel Pérez, Alexis Musacchio, José García, Osvaldo Reyes, Hilda E Garay, Luis J González, Julio C Alvarez, Yordanka Soria. Biotecnología Aplicada 23(1):60-36 · January 2006.
23. *Hepatitis C virus (HCV) core protein enhances the immunogenicity of a co-delivered DNA vaccine encoding HCV structural antigens in mice.* Alvarez-Lajonchere L, Gonzalez M, Alvarez-Obregon JC, Guerra I, [Vina A](#), Acosta-Rivero N, Musacchio A, Morales J, Duenas-Carrera S. Biotechnol Appl Biochem 2006; 44(Pt 1):9-17. [PubMed](#)
24. *Immunization with a DNA vaccine encoding the hepatitis-C-virus structural antigens elicits a specific immune response against the capsid and envelope proteins in rabbits and Macaca irus (crab-eating macaque monkeys).* Duenas-Carrera S, [Vina A](#), Martinez R, Alvarez-Lajonchere L, Alvarez-Obregon JC, Marante J, Perez A, Mosqueda O, Martinez G, Morales J. Biotechnol Appl Biochem 2004; 39(Pt 2):249-55. [PubMed](#)
25. *Desarrollo de un sistema de diagnóstico molecular para la detección cualitativa del ARN del virus de la Hepatitis C.* Yaimé Josefina González González, Idania González Pérez, [Ariel Viña Rodríguez](#), Anny Armas Cayarga, Iria García De La Rosa, Rosa Lydia Solís Rodríguez. Biotecnología Aplicada 20(2):122-5 · January 2003
26. *Enhancement of the immune response generated against hepatitis C virus envelope proteins after DNA vaccination with polyprotein-encoding plasmids.* Duenas-Carrera S, Alvarez-Lajonchere L, Cesar Alvarez-Obregon J, Perez A, Acosta-Rivero N, Vazquez DM, Martinez G, [Vina A](#), Pichardo D, Morales J. Biotechnol Appl Biochem 2002; 35(Pt 3):205-12. [PubMed](#).
27. *Additives and Protein-DNA Combinations Modulate the Humoral Immune Response Elicited by a Hepatitis C Virus Core-encoding Plasmid in Mice.* Alvarez-Lajonchere L, Dueñas-Carrera S, [Viña A](#), Ramos T, Pichardo D, Morales J. Memorias do Instituto Oswaldo Cruz. 2002;97(1) :95-9. [PubMed](#)
28. *Definition of a vaccinal candidate against the hepatitis C virus from pre-clinic studies results.* Dueñas-Carrera S, Morales J, Alvarez-Lajonchere L, Alvarez JC, Lorenzo LJ, Acosta-Rivero N,

- Martinez G, Vina A, Guerra I, Pichardo D, Herrera A, Martinez, R, Vaquez DM, Silva R, Cosme K. *Biotechnología Aplicada* 2001; 8: 99-100. PDF, Bioline-Full
29. *Immunological evaluation of Escherichia coli-derived hepatitis C virus second envelope protein (E2) variants*. Duenas-Carrera S, Vina A, Garay HE, Reyes O, Alvarez-Lajonchere L, Guerra I, Gonzalez LJ, Morales J. *J Pept Res* 2001; 58(3):221-8. PubMed.
 30. *Characterization of the HCV core virus-like particles produced in the methylotrophic yeast Pichia pastoris*. Acosta-Rivero N, Aguilar JC, Musacchio A, Falcon V, Vina A, de la Rosa MC, Morales J. *Biochem Biophys Res Commun* 2001; 287(1):122-5. PubMed.
 31. *Humoral Immune Response against a Hepatitis C Virus Envelope E2 Variant Expressed in Escherichia coli*". Gillian Martínez, Ariel Viña, Madeline Borges, Eduardo Martínez y Juan Morales Grillo. *Biotechnología Aplicada* 2000; 17:231-4. PDF, Bioline-Full,
 32. *Cloning and Purification of the Hydrophilic Fragment of Hepatitis C Virus E2 Protein Fused to Choline-binding Domain of the Major Autolysin of Streptococcus pneumoniae: Evaluation of the Humoral Immune Response in Rabbits*, Gillian Martínez, Ariel Viña, Jose Luis García, Juan Morales Grillo *Biotechnología Aplicada*, 2000; 17:85-8. PDF , Bioline-Full
 33. *Fractal properties of DNA sequences.(*)* S. V. Korolov, A. R. Viña (*), V. G. Tumanian, N. G. Esipova. pg 221-8. *Fractals in the Nature and Applied Sciences (A-41)*. M.M. Novak (Editor). 1994 IFIP.

Patents:

1. *Vaccine formulation enhanced by the combination of a DNA and an antigen*. 2001. Dueñas-Carrera S, Morales J, Alvarez-Lajonchere L, Musacchio A, Pajón R, Viña A, Alvarez-Obregón JC, Acosta- Rivero N, Martínez G. Cuban application: 2001-0171. International application: July, 12, 2002. PCT/CU02/00005. EP1417973
2. *Sequences derived from the genome of the hepatitis C virus, and use thereof*. 1998. Morales J, Viña A, Garcia C, Acosta-Rivero N, Dueñas-Carrera, S, Garcia O.,Guerra I. WO 98/25960. PCT/CU97/00007
3. Certificado de Autor de Invención. *Highly sensible and specific primers and probes from the 5'NCR of the HCV*. Idania Gonzalez Perez, Ariel Vina Rodriguez, Yaime Josefina Gonzalez Gonzalez, Anny Armas Cayarga, Iria Garcia de la Rosa, Yenitse Perea Hernandez. CU23530(A1) 2010-06-17

Acknowledgments

First at all, I would like to thank Professors Martin H. Groshup and Thomas C. Mettenleiter for giving me the opportunity to work at the Friedrich-Loeffler-Institut in the exciting times of the establishment of the new laboratory buildings, which coincided with the expansion of the investigation subject of the Institute of Novel and Emerging Infectious Diseases (INNT).

I deeply appreciate the confidence in me and the support of Professor Groshup and of our Laboratory Head, and my personal mentor, Dr. Martin Eiden. They let me participate in important projects while at the same time they gave me full freedom to materialize my own ideas. I specially appreciate the kindness and patience with which Eiden discussed with me all the details of my work.

For years I enjoined a very friendly atmosphere from all the member of our laboratory and Institute, which made it a pleasure to work with them. I was also fortunate to experience the enthusiasm and knowledge in the field of identification of pathogens by microarray of Professor Dr. Konrad Sachse who made a fundamental contribution to that important part of this work.

I am very grateful to Professor and Chair of Biochemistry at the University of Greifswald Dr. Winfried Hinrichs, which agreed to tutor my promotion and continued to do so even after he was retired.

Finally, I must say that there are no words to express my acknowledgment to my wife, to my parents and to my children, who patiently supported my endless hours of work for many years.