2nd International Electronic Conference on Medicinal Chemistry

1-30 November 2016

chaired by Dr. Jean Jacques Vanden Eynde

sponsored by pharmaceuticals

HIGHLY CONSERVED WNV GENOMIC RNA DOMAINS ARE POTENTIAL TARGETS OF ANTIVIRAL RNA APTAMERS

Cristina Fernández-Molina, Cristina Romero-López* and Alfredo Berzal-Herranz*

[1] Instituto de Parasitología y Biomedicina López-Neyra, IPBLN-CSIC, Av. Conocimiento s/n, PTS Granada, 18016, Armilla, Granada, Spain.

* Corresponding author: aberzalh@ipb.csic.es; cristina_romero@ipb.csic.es

ipbln CSIC

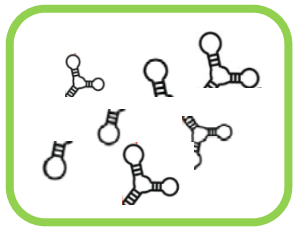
HIGHLY CONSERVED WNV GENOMIC RNA DOMAINS ARE POTENTIAL TARGETS OF ANTIVIRAL RNA APTAMERS

Aptamers

Oligonucleotides obtained by *in vitro* selection that efficiently and specifically bind to a ligand molecule

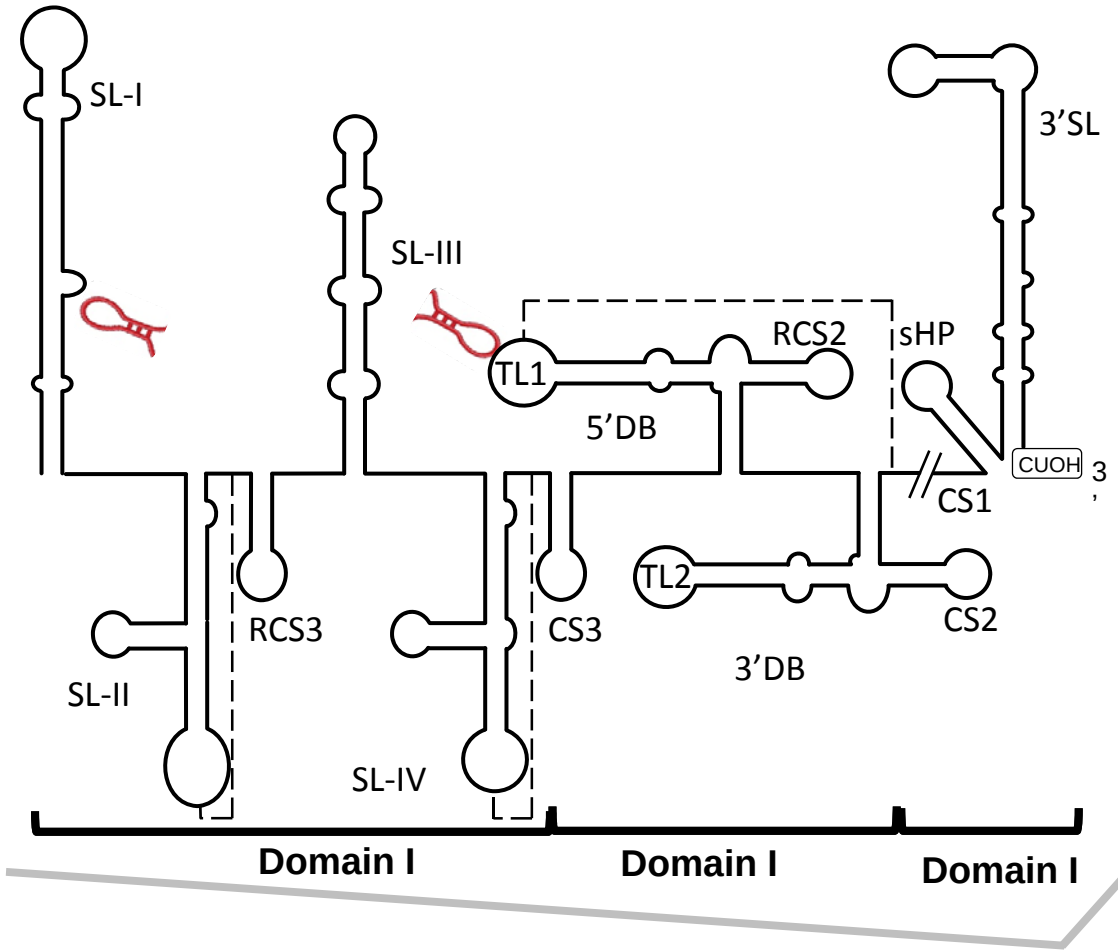
SELEX

Starting population (10^{18} RNA molecules)

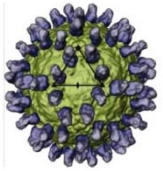


6 SELEX rounds

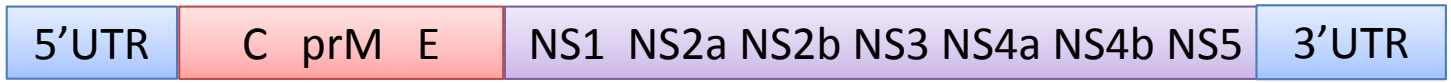
3 families of aptamers



WNV



Single-stranded RNA genome



Abstract:
West Nile virus (WNV) is an enveloped, positive, ssRNA virus belonging to the *Flaviviridae* family. Different WNV strains have been involved in important outbreaks of human and animal diseases, and WNV infection is now considered an emerging world health problem. Like other RNA viruses, WNV has a compact RNA genome that efficiently stores all the information required for the completion of the infectious cycle. The efficiency of this storage system is attributable to supracoding elements, i.e., discrete structural units with essential functions, which overlap and complement the protein coding information. These elements, therefore, offer interesting potential targets for novel therapeutic agents. We have applied a SELEX procedure to isolate RNA aptamers against the essential 3' untranslated region (3'UTR) of the WNV genome. Starting from a theoretical initial population consisting of more than 10^{18} different molecules, we have identified three main groups of aptamers defined by sequence motifs complementary to essential sequence and structural elements of the WNV genome. Current results point out the potential of these essential functional genomic RNA elements to efficiently bind RNA molecules, therefore to be involved in RNA-RNA interactions, offering a potential of being used as targets of antiviral agents based on nucleic acids.
Keywords: RNA aptamer; West Nile Virus; antiviral aptamer; RNA domains

WEST NILE VIRUS (WNV): GENERAL FEATURES (I)

It was isolated in 1937 from the blood of a woman living in the West Nile district of Uganda.

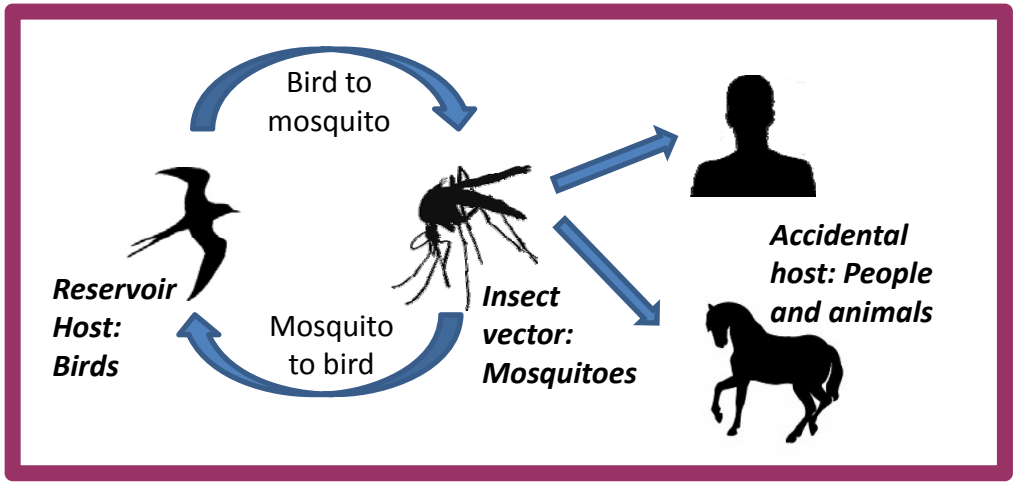
It belongs to the *Flaviviridae* family, within the *Flavivirus* genus.

It is widely distributed in Africa, Europe, Australia, and Asia. Since 1999, it has spread rapidly throughout the western hemisphere.

WEST NILE VIRUS (WNV): GENERAL FEATURES (II)

Though it mainly infects birds, WNV can be transmitted by the bite of infected mosquitoes to humans and horses, in which it causes outbreaks of severe encephalitis and febrile disease. During the 1990s, the virus spread from Asia, Africa and Australia to Europe and America. Since the 1999 outbreak in the US, it has been considered a global health threat, and efficient therapeutic and immunoprophylactic ways to fight WNV infection need to be found.

WNV transmission cycle



Global distribution of WNV

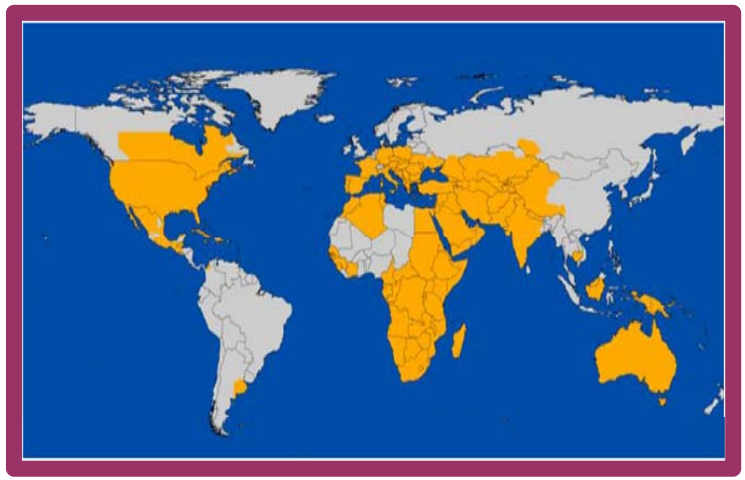
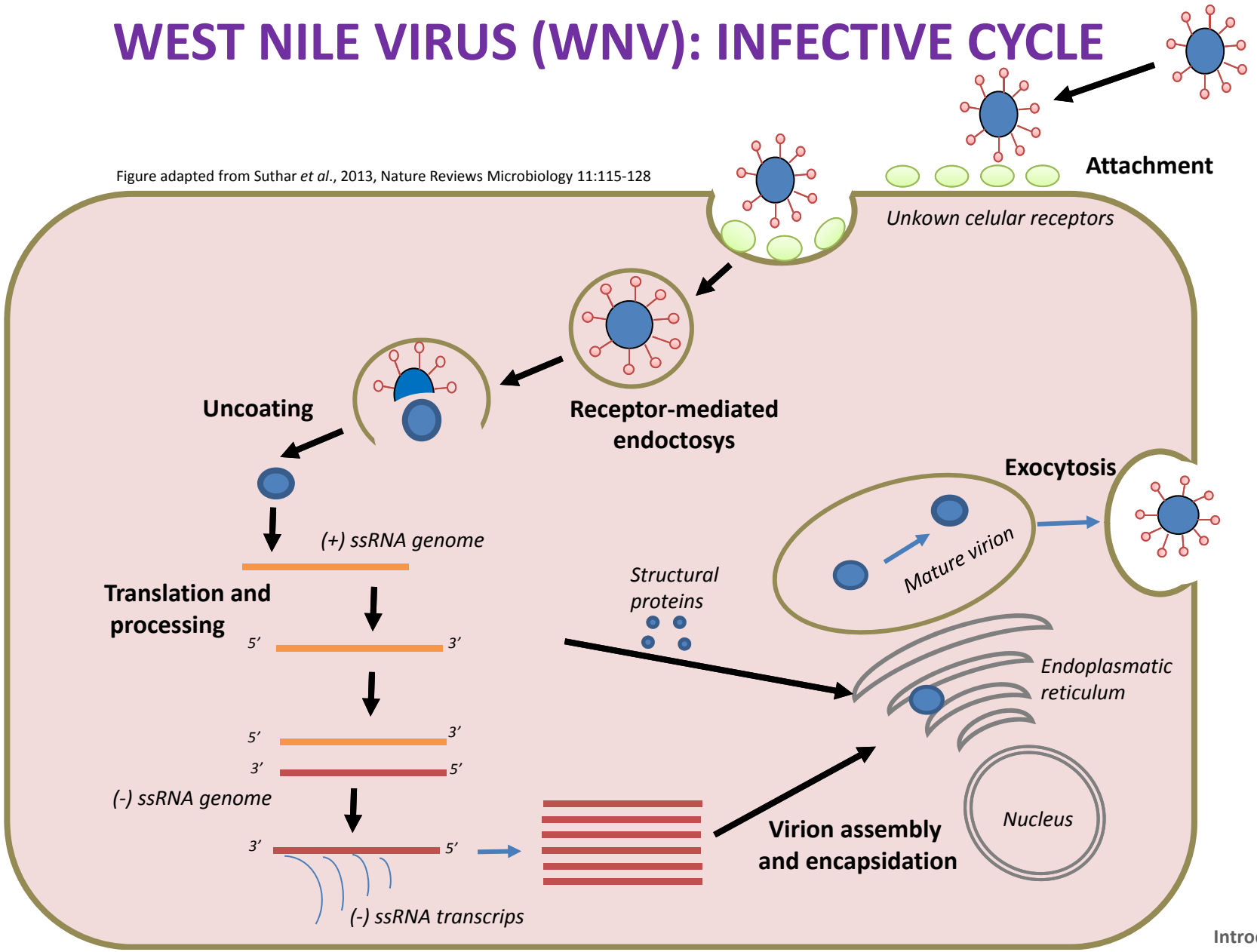


Figure from Reisen, 2013, Viruses, 5(9):2079-2105.

WEST NILE VIRUS (WNV): INFECTIVE CYCLE

Figure adapted from Suthar *et al.*, 2013, Nature Reviews Microbiology 11:115-128



WEST NILE VIRUS (WNV): STRUCTURAL FEATURES

The virion is a spherical particle of~ 50 nm diameter.

It consists of a host-derived lipid bilayer membrane surrounding a nucleocapsid core that contains single-stranded positive-sense RNA genome of approximately 11 000 nucleotides.

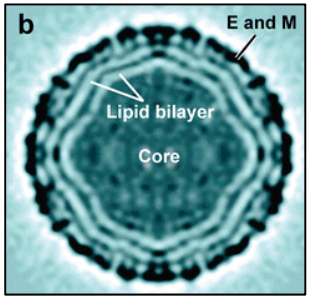
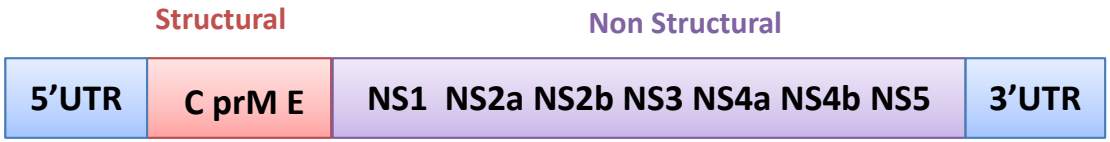
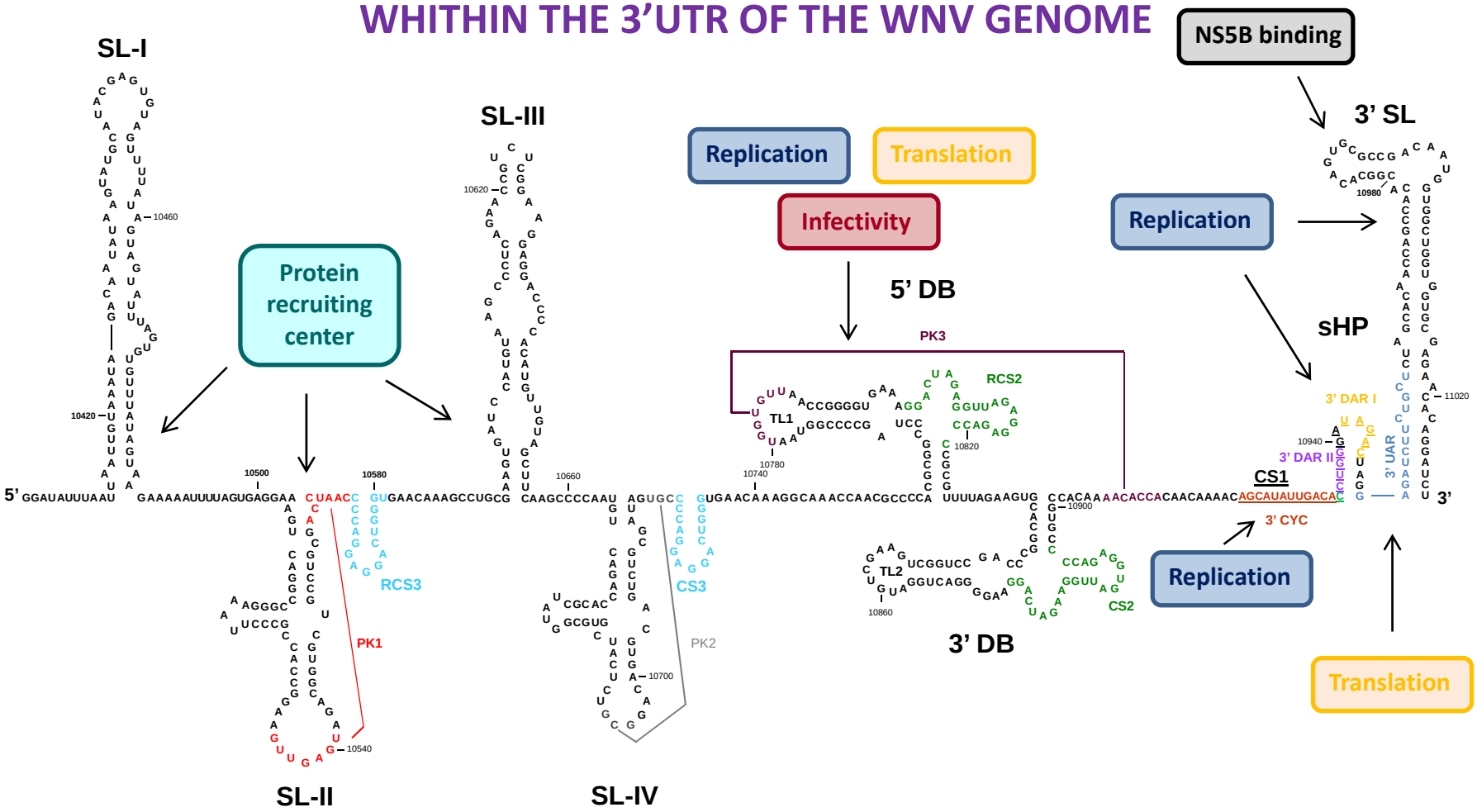


Figure reproducted from Mukhopadhyay *et al*., 2003, *Science,* 302 (5643):248

The WNV genome encodes a single ORF flanked by untranslated regions (UTRs). These regions are defined by discrete, functionally active structural elements that play important roles in the viral cycle. Such functional RNA elements appear as highly conserved, complex folding regions across flaviviruses, despite the lack of extensive sequence complementarity

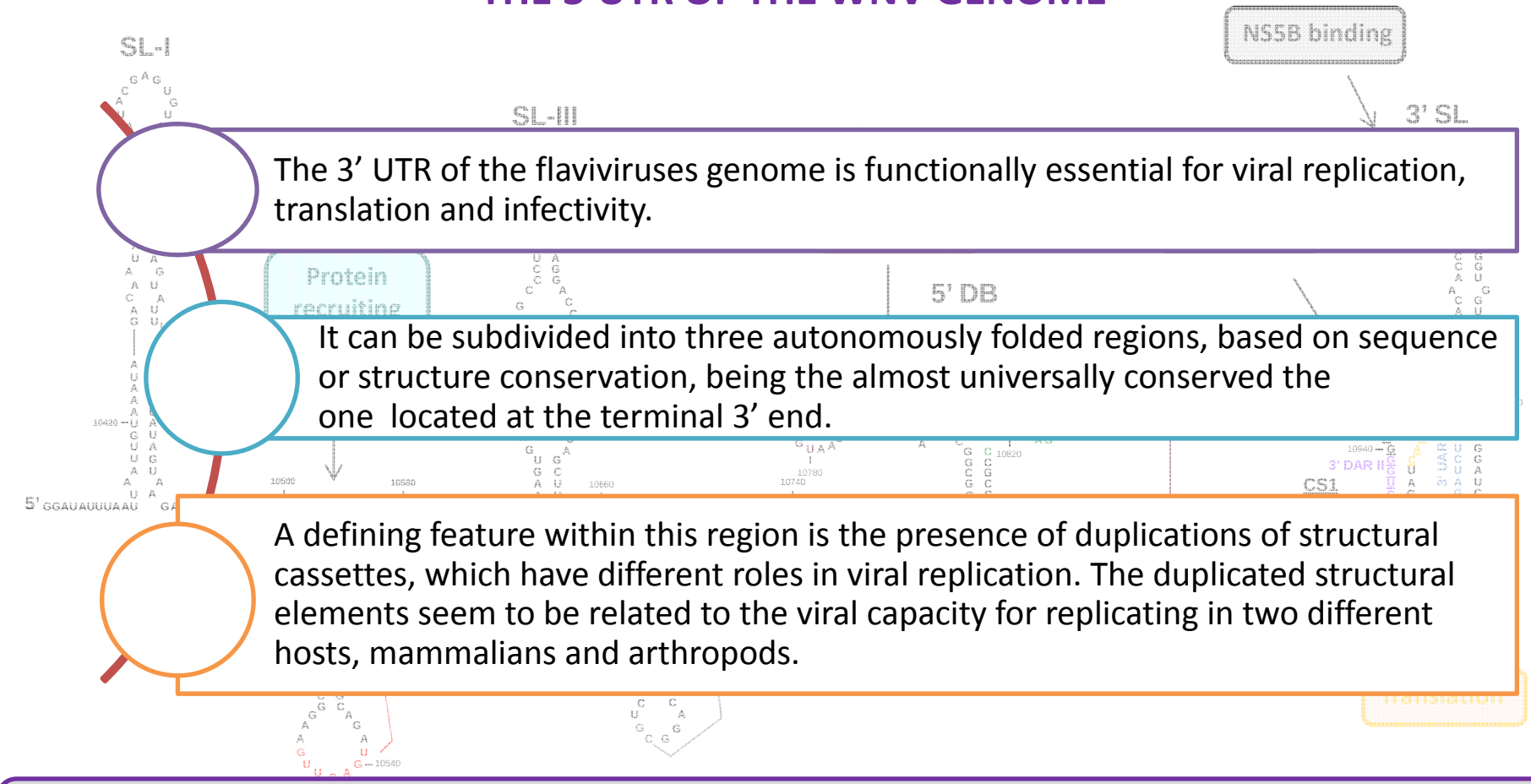


FUNCTIONAL RNA DOMAINS AND LONG-RANGE RNA-RNA INTERACTIONS WHITHIN THE 3'UTR OF THE WNV GENOME



The functional importance of the highly conserved structural genomic RNA domains in different RNA viruses renders them therapeutic targets for new antiviral drugs

FUNCTIONAL RNA DOMAINS AND LONG-RANGE RNA-RNA INTERACTIONS WHITHIN THE 3'UTR OF THE WNV GENOME



The 3' UTR of the flaviviruses genome is functionally essential for viral replication, translation and infectivity.

It can be subdivided into three autonomously folded regions, based on sequence or structure conservation, being the almost universally conserved the one located at the terminal 3' end.

A defining feature within this region is the presence of duplications of structural cassettes, which have different roles in viral replication. The duplicated structural elements seem to be related to the viral capacity for replicating in two different hosts, mammalians and arthropods.

Therefore, the 3'UTR of the WNV genome is an excellent candidate target for the development of novel therapeutic and biotechnological tools aimed to study and fight the viral infection

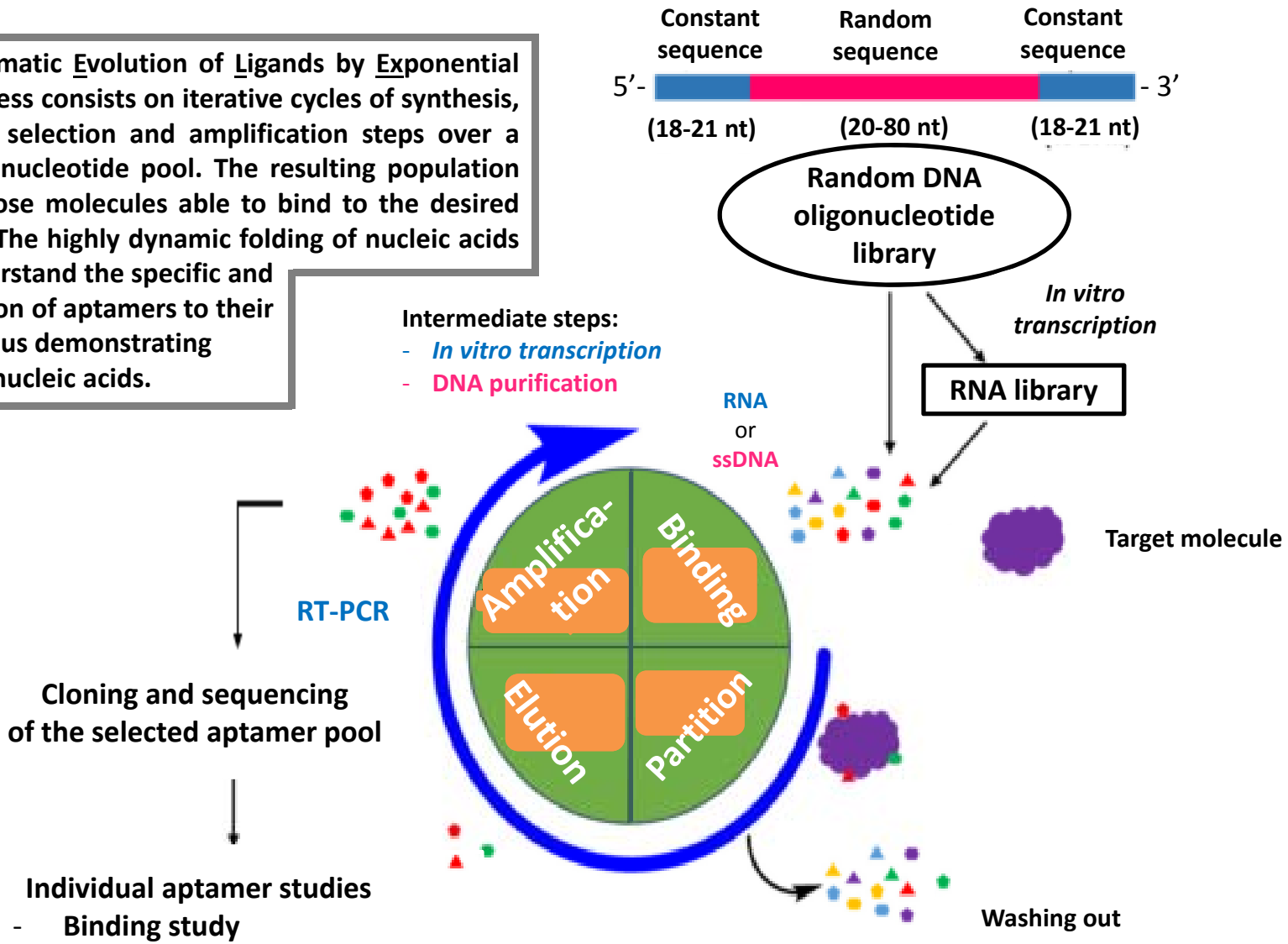
APTAMERS

- Single-stranded DNA or RNA molecules.
- Recognize their target with great affinity and specificity due to their three-dimensional folding.
- Potential use in both research and therapy.

Isolated from oligonucleotide libraries by the so-called **SELEX** (Systematic Evolution of Ligands by Exponential Enrichment) *in vitro* process

SELEX

The SELEX (Systematic Evolution of Ligands by Exponential enrichment) process consists on iterative cycles of synthesis, binding, positive selection and amplification steps over a randomized oligonucleotide pool. The resulting population is enriched in those molecules able to bind to the desired target molecule. The highly dynamic folding of nucleic acids is the key to understand the specific and efficient interaction of aptamers to their cognate target, thus demonstrating the versatility of nucleic acids.



The main goal of this work was to identify RNA aptamers that efficiently and specifically bind to the 3'UTR of the WNV genome

RNA LIBRARY DESIGN

POPULATION

Random region refers to the 25 nts-long randomized region. PBS, constant sequence used as primer binding site. T7P1-EcoRI and P3-KpnI indicates the hybridization sites of the oligonucleotides used for the RT-PCR step during the SELEX process.

NAME	OLIGONUCLEOTIDE SEQUENCE
25N (2)	5'- **TAA TAC GAC TCA CTA TA**G GGA *GAA TTC* TCG AAG CTA GCA TNN NNN NNN NNN NNN NNN NNN NNN NNA TCG ATG ACA GT*G GTA CC*A AC -3'
P3-KpnI	5'- GTT *GGT ACC* ACT GTC ATC GAT -3'
T7P1- EcoRI	5'- **TAA TAC GAC TCA CTA TA**G GGA *GAA TTC* TCG AAG CTA GCA T -3'

Oligonucleotides used in this study. The promoter sequence for the RNA polymerase derived from the T7 bacteriophage is shown in bold. Restriction sites are underlined and highlited in italics.

ANALYSIS OF THE BINDING ABILITY

During the course of the selection procedure, the stringency was increased for the fourth generation by raising the temperature of the association step (to 37ºC), diminishing the ratio target:aptamer pool and reducing the incubation time. After six rounds of selection, the efficiency of the SELEX method was monitored by gel mobility shift assays. It is noteworthy that the population derived from the sixth round of selection shows a significant enrichment in aptamers able to efficiently bind to the 3'UTR of the WNV genome.

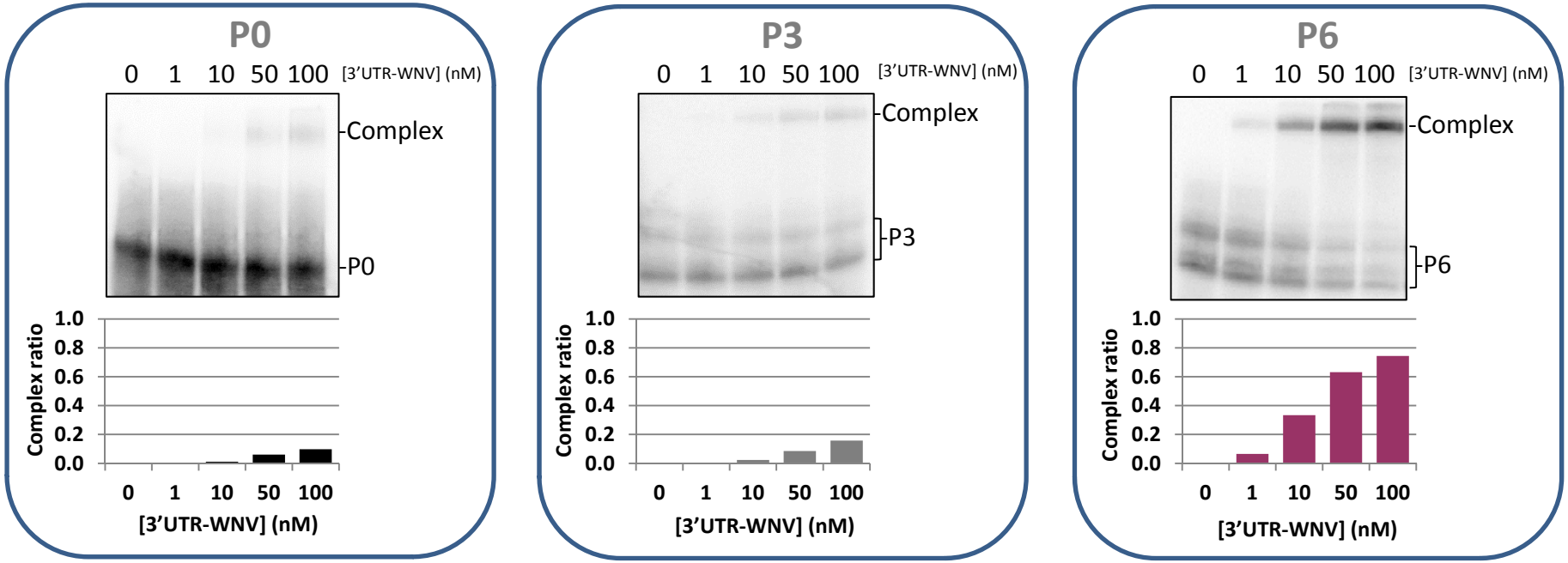


Figure shows representative autoradiograms of the *in vitro* binding assays. The internally ^{32}P-labeled RNA pools P0, P3 and P6 (initial population, RNA molecules derived from the third round and RNA molecules derived from the sixth round, respectively) were incubated with increasing amounts of the target 3'UTR-WNV RNA construct. The pool of RNA variants corresponding to the selection cycle is shown. Complexes showed a reduced electrophoretic mobility.

SELECTED SEQUENCES

After six rounds of selection, three groups of variants were defined by bearing common sequence motifs. Some of them were classifiable in several groups since they had more than one consensus domain. Sequence of the 25-nucleotide-long RNA motif (aptamer) is shown. Representative variants for each group were chosen for further analyses, indicated with an asterisk. In addition, molecule P6-36 was identified to carry a unique sequence motif targeting the highly conserved apical loop within the 3'SL element.

Group I = 5' UAACAC 3'

*>P6-2	CACUAACACCCCACUGCCGAUGAGU
>P6-22	UUAACACUAACACCACACGCAUUGGG
>P6-23	UAACACCCACGAUUACGGUUGCCC
*>P6-37	UAACACCAUAGUCCUGUUCUUGGGU
>P6-39	UAACACCAUCUAAGCUGCUGUCAGG
>P6-53	GACACUAACACCUUCCUGGACGGCA
>P6-57	UUACACUAACACCUUGUCGUACCGC
>P6-62	UAUUUGCAUUUAACACACUAACCAU
>P6-63	GACAUAAUAUAACACUAGUUGCGCCA
>P6-70	ACACUAACACCGGUACACGCAAGC
>P6-77	UCACUAACACCUGCACUCUAUAGCG

Group II = 5' YACAC 3'

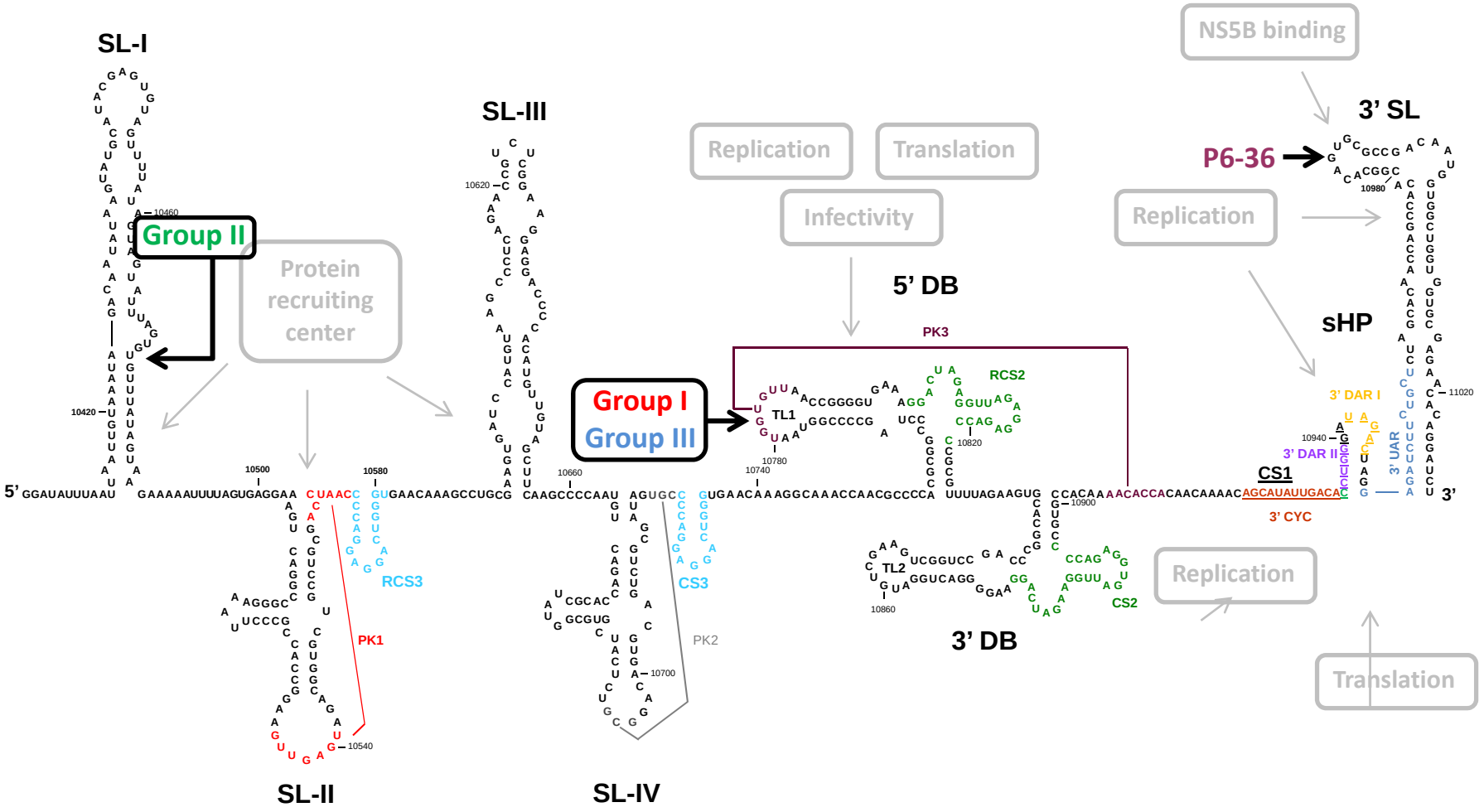
*>P6-3	CACUACACUACACUCGACACACGGC
>P6-6	UACACUAUACACAGCGUAUUCCACC
>P6-7	CACUCAUUACACUAUGGCAAACCACA
>P6-17	CACACUACAACACUACACUCCAGGU
>P6-22	UUAACACUAACACCACACGCAUUGGG
*>P6-32	UAACUAUUUACACAUUUCGAGCCUC
*>P6-44	ACUAUAAGACUGCACACUGCACCAU
>P6-48	UUACACUAACCCAUCCGCAGCUCCA
>P6-50	UUUACACUAUCUACUACAGACCACU
*>P6-51	GACUUAACUAUACACUACGCUCCAC
>P6-57	UUACACUAACACCUUGUCGUACCGC
>P6-58	UUUACACUAAACCAUACCACGGGCCA
>P6-59	CAUGAACUACAAUCUACACACACCAGU
>P6-62	UAUUUGCAUUUAACACACUAACCAU
>P6-70	ACACUAACACCCGGUACACGCAAGC
>P6-75	AUUUACACAACACCUUGCCUCUGUC
>P6-76	GACUAUUAUUACACUACACCACGCG
>P6-78	CAUACACAUUACACCUCGUACCGUC
>P6-79	UUACACUAACCCACUUAAUCUUCCU

Group III = 5' RCACUAA 3'

>P6-22	UUAACACUAACACCACACGCAUUGGG
>P6-28	GCACUAACAACACUAUUUAGCCACC
>P6-48	UUACACUAACCCAUCCGCAGCUCCA
*>P6-51	GACUUAACUAUACACUACGCUCCAC
>P6-52	UAACUAUUUGCACUAACGACGUUGG
>P6-53	GACACUAACACCUUCCUGGACGGCA
>P6-57	UUACACUAACACCUUGUCGUACCGC
>P6-58	UUUACACUAAACCAUACCACGGGCCA
>P6-70	ACACUAACACCCGGUACACGCAAGC
>P6-79	UUACACUAACCCACUUAAUCUUCCU

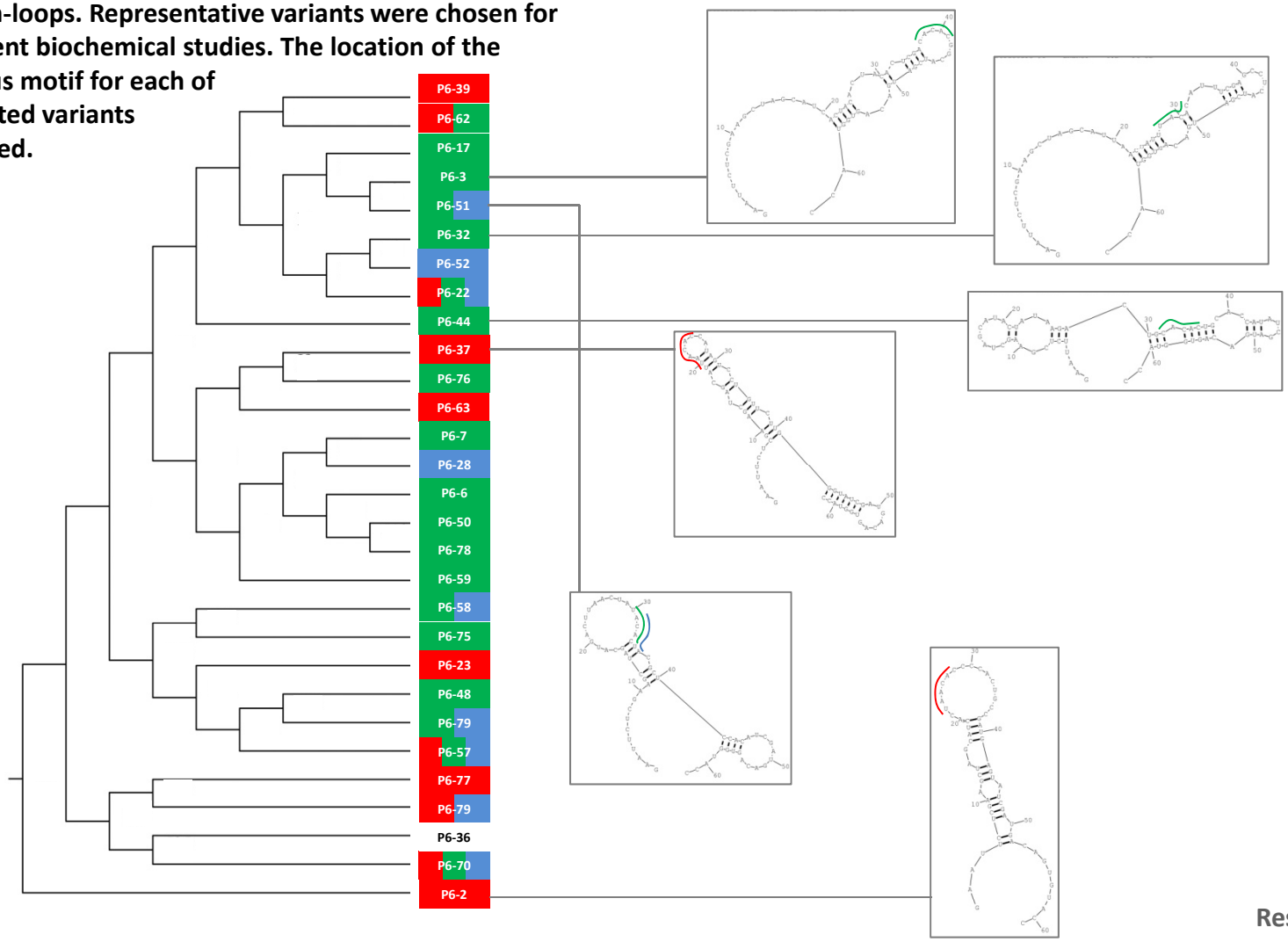
>P6-36	ACAACACGACAGUGGCGCACAAAAGG

TARGET IDENTIFICATION



In silico STRUCTURAL ANALYSIS

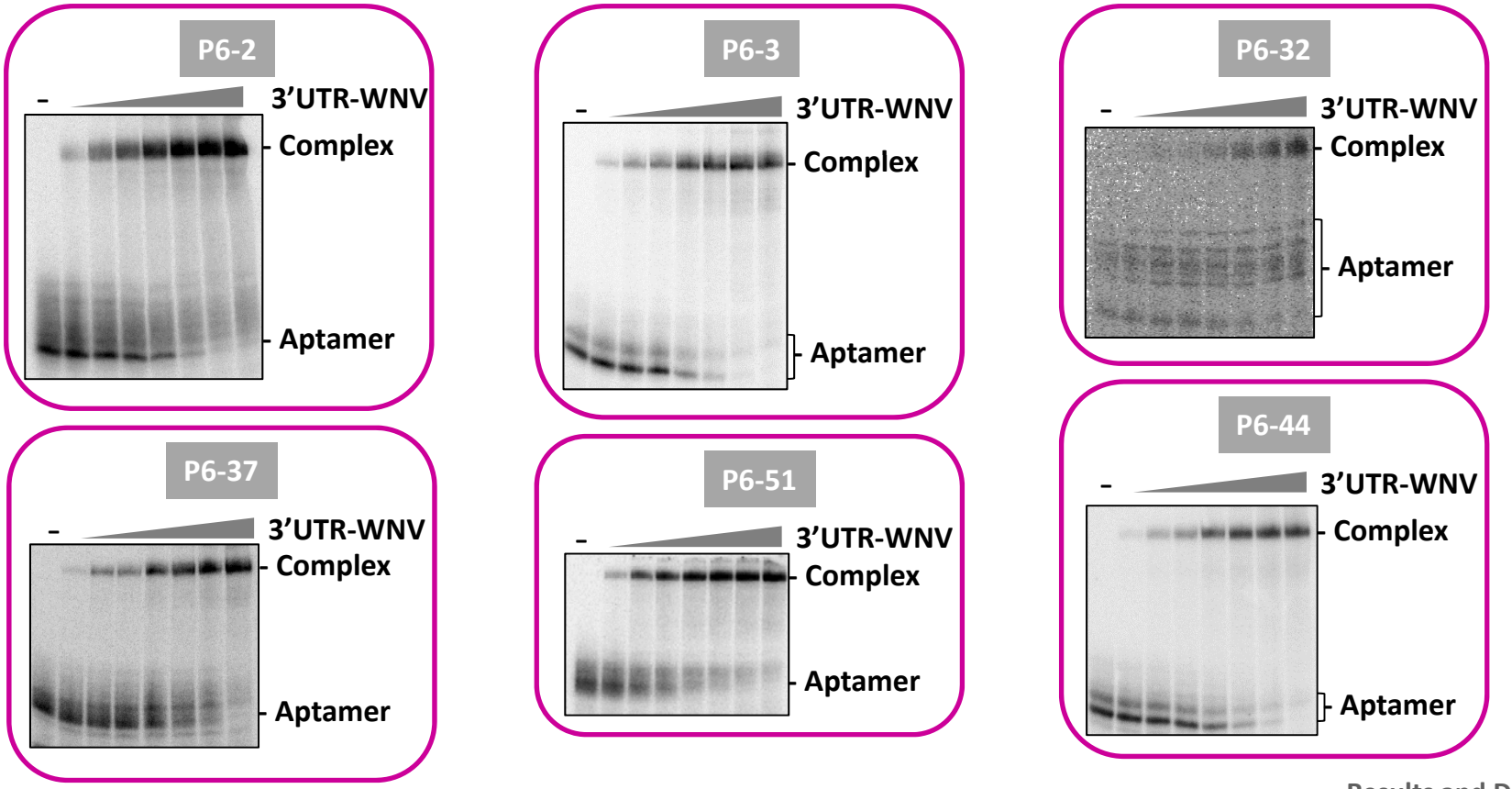
The pairwise LocARNA algorithm was used for aligning the RNA variants derived from the sixth round of selection. LocARNA performs simultaneous alignment of both sequence and folding. Two main structural clusters were identified, defined by the acquisition of one or two stem-loops. Representative variants were chosen for subsequent biochemical studies. The location of the consensus motif for each of the selected variants is indicated.



Results and Discussion

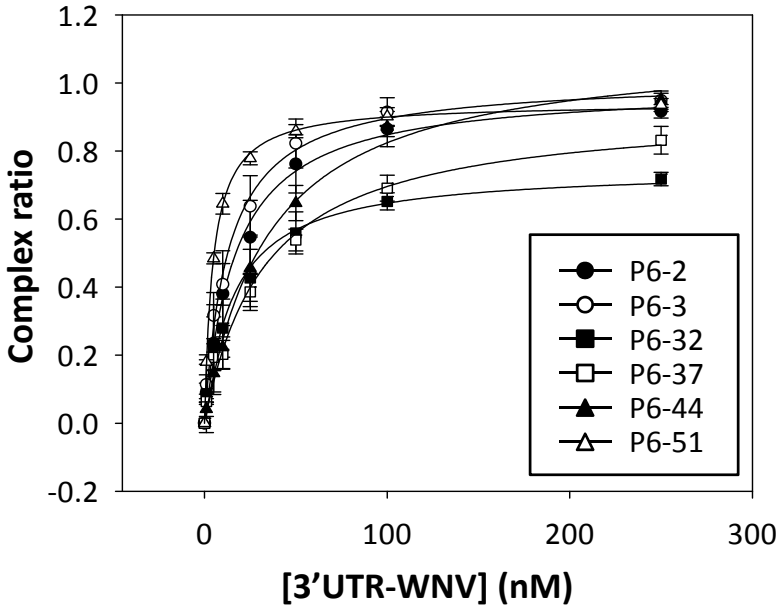
THE SELECTED MOLECULES EFFICIENTLY BIND TO THE 3'UTR OF THE WNV GENOME

Binding assays were performed by incubating trace amounts of the ^{32}P 5'-end-labelled aptamer, P6-2, P6-3, P6-32, P6-37, P6-44 or P6-51, with increasing concentrations of the RNA substrate 3'UTR-WNV in the presence of binding buffer. The resulting complexes were resolved by native polyacrylamide gel electrophoresis.



THE SELECTED MOLECULES EFFICIENTLY BIND TO THE 3'UTR OF THE WNV GENOME

Complex formation was quantified using the Image Quant 5.2© software. K_d values were calculated using the Sigma Plot 8.02© tool according to the equation $y = (B_{max} \cdot x)/(K_d + x)$, where y is the percentage of complexed inhibitory RNA, B_{max} is the amplitude of the reaction, x is the concentration of the substrate RNA and K_d is the dissociation constant.



Variant	K_d ± sd (nM)	Amplitude (nM)
P6-2	16.79 ± 1.61	0.99 ± 0.02
P6-3	12.99 ± 1.36	1.01 ± 0.03
P6-32	15.93 ± 2.14	0.75 ± 0.03
P6-37	33.03 ± 5.09	0.92 ± 0.05
P6-44	35.31 ± 3.26	1.11 ± 0.03
P6-51	4.66 ± 0.16	0.94 ± 0.01

All the tested variants showed efficient binding ability to their target RNA 3'UTR-WNV, with K_d values in the range of low-medium nanomolar and binding yield close to 100%

CONCLUSIONS

- An *in vitro* selection method has been performed for the isolation of RNA aptamers targeting the highly conserved 3'UTR region of the WNV genome.
- After six rounds of selection, 29 different variants have been isolated for their ability to bind the 3'UTR-WNV RNA construct.
- All the selected variants can be classified into three different groups defined by consensus sequence motifs.
- The identified consensus motifs present a complementary sequence to conserved regions in the target 3'UTR-WNV.
- Groups I and III share the same target region located in the highly conserved domain 5'DB, which is essential for viral translation, replication and infectivity; while group II maps in the SL-I element, involved in the recruitment of host and viral protein factors.
- The isolated sequences can be also clustered by virtue of their conformation in two groups, depending on they fold into one or two stem-loops.
- A number of representative variants for each group or cluster have shown to efficiently bind to the 3'UTR-WNV RNA, thus confirming the proper functioning of the selection procedure.

ACKNOWLEDGEMENTS

Vicente Augustin-Vacas
Beatriz Berzal-Herranz

Dr. Juan Carlos Sáiz

FUNDING

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
INIA
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria

JUNTA DE ANDALUCIA