Supplemental Information

Nucleotide-Resolution Profiling of RNA Recombination in the Encapsidated Genome of a Eukaryotic RNA Virus by Next-Generation Sequencing

Andrew Routh *[1], Phillip Ordoukhanian[2] and John E. Johnson[1]

1) Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

2) The Next Generation Sequencing Core, The Scripps Research Institute, La Jolla, CA 92037, USA

*corresponding author:

arouth@scripps.edu

Department of Molecular Biology, MB31,

The Scripps Research Institute,

10550 North Torrey Pines Rd,

La Jolla, CA, 92037, USA

Tel: 858-784-8654

Fax: 858-784-8660

Supplemental Figure Legends:

Supplementary Figure 1: Base composition of donor and acceptor sites. The average frequency of each nucleotide as indicated was taken for every junction (limited to 100 events per junction to prevent weighting towards unique and highly represented sequences). **A)** 5' strand junction site in RNA 1 junctions, **B)** 3' strand junction site in RNA 1 junctions, **C)** 5' strand junction site in RNA 2 junctions, and **D)** 3' strand junction site in RNA 2 junctions. The level of expected base composition for each nucleotide is shaded in grey and the position of the junction is marked with a blue line.

Supplementary Figure 2: Read coverage over the FHV genome for FHV-RNAseq and VLP-RNAseq datasets. The number of reads aligning over the wild type FHV genome is indicated for each nucleotide. **A)** FHV-RNAseq to FHV RNA 2. **B)** VLP-RNAseq to FHV RNA 2. **C)** FHV-RNAseq to FHV RNA 1.

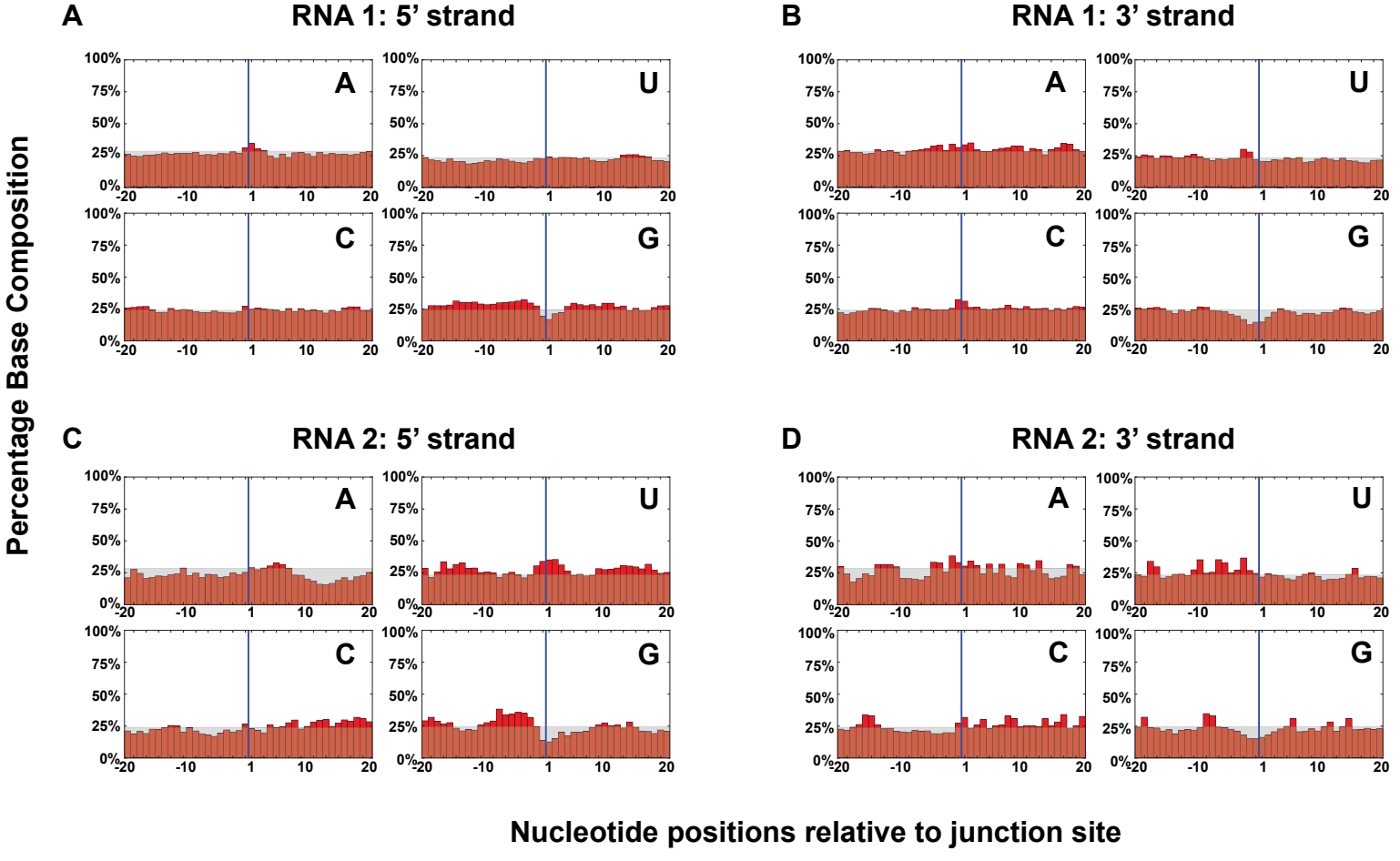
Supplementary Table 1: **Mapping of VLP-RNAseq reads to recombination junctions in genes packaged by VLPs.**

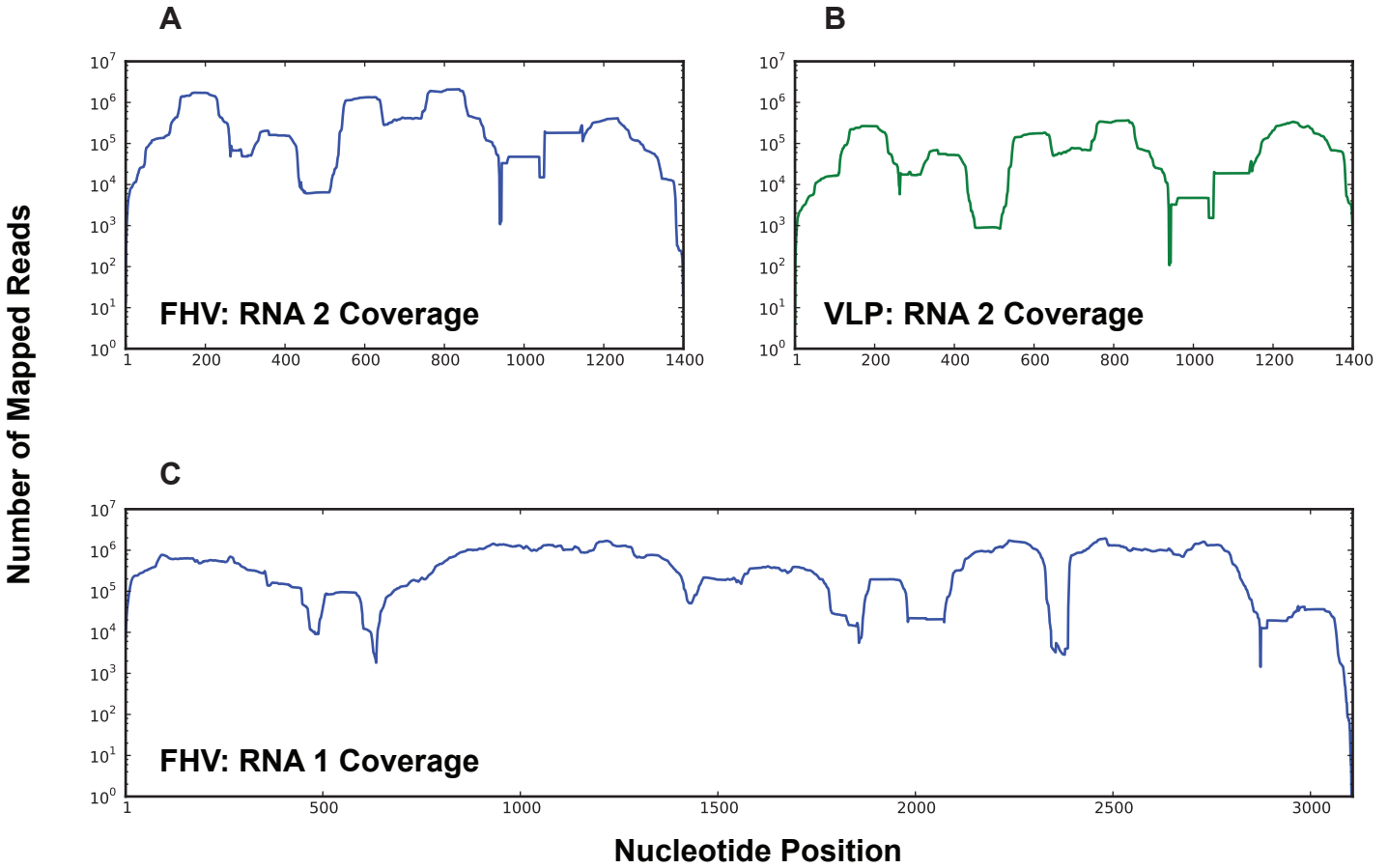
	R2-R2	**Ar-Ar**	**Ac-Ac**	**R2-Ac**	**R2-Ar**	**Ar-Ac**	**Ac-Ar**	**Ar-R2**	**Ac-R2**
Total	2,168	17,462	801	124	82	12	85	51	427
MicroIndels	*1,812*	*17,241*	*583*	-	-	-	-	-	-
Recombination	*356*	*221*	*218*	*124*	*82*	*12*	*85*	*51*	*427*

Ac = ORF 1629 from AcMNPV;

Ar = 18S rRNA from *Attacus Ricini* rRNA;

R2 = FHV RNA 2


A
RNA 1: 5' strand
B
RNA 1: 3' strand
C
RNA 2: 5' strand
D
RNA 2: 3' strand
A
U
C
G
100%
75%
50%
25%
0%
-20
-10
1
10
20
Percentage Base Composition
Nucleotide positions relative to junction site


A
B
C
FHV: RNA 2 Coverage
VLP: RNA 2 Coverage
FHV: RNA 1 Coverage
Number of Mapped Reads
Nucleotide Position