Propidium monoazide (PMA) and ethidium bromide monoazide (EMA) improve DNA array and high-throughput sequencing of porcine reproductive and respiratory syndrome virus identification.

Christian Bellehumeur,[a,b] Brian Boyle,[c] Steve J. Charette,[b,c,d,e] Josée Harel,[a,b] Yvan L'Homme,[b,f] Luke Masson,[b,g] Carl A. Gagnon[a,b]#*

Groupe de recherche sur les maladies infectieuses du porc (GREMIP)[a] and Swine and Poultry Infectious Diseases Research Center (CRIPA)[b], Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada; Institut de biologie intégrative et des systèmes (IBIS), Université Laval, Québec, QC, Canada[c];

Département de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval, Québec, QC, Canada[d];

Centre de recherche de l'Institut universitaire de cardiologie et de pneumologie de Québec, Québec, QC, Canada[e];

Canadian food inspection agency, Saint-Hyacinthe, QC, Canada[f];

National Research Council Canada, Montréal, QC, Canada[g]

Running Head: Improvement of virus detection from clinical samples

#Address correspondence to carl.a.gagnon@umontreal.ca

*Present mailing address: Faculté de médecine vétérinaire, Université de Montréal, 3200 rue Sicotte, Saint-Hyacinthe, Québec, Canada, J2S 7C6. Phone: 450-773-8521 (8681). Fax: 450-778-8108.

Abstract

Pan-viral DNA array (PVDA) and high-throughput sequencing (HTS) are useful tools to identify novel viruses of emerging diseases. However, both techniques have difficulties to identify viruses in clinical samples because of the host genomic nucleic acid content (hg/cont). Both propidium monoazide (PMA) and ethidium bromide monoazide (EMA) have the capacity to bind free DNA/RNA, but are cell membrane-impermeable. Thus, both are unable to bind protected nucleic acid such as viral genomes within intact virions. However, EMA/PMA modified genetic material cannot be amplified by enzymes. In order to assess the potential of EMA/PMA to lower the presence of amplifiable hg/cont in samples and improve virus detection, serum and lung tissue homogenates were spiked with porcine reproductive and respiratory virus (PRRSV) and were processed with EMA/PMA. In addition, PRRSV RT-qPCR positive clinical samples were also tested. EMA/PMA treatments significantly decreased amplifiable hg/cont and significantly increased the number of PVDA positive probes and their signal intensity compared to untreated spiked lung samples. EMA/PMA treatments also increased the sensitivity of HTS by increasing the number of specific PRRSV reads and the PRRSV percentage of coverage. Interestingly, EMA/PMA treatments significantly increased the sensitivity of PVDA and HTS in two out of three clinical tissue samples. Thus, EMA/PMA treatments offer a new approach to lower the amplifiable hg/cont in clinical samples and increase the success of PVDA and HTS to identify viruses.

Keywords: DNA array; high-throughput sequencing; virus identification; porcine reproductive and respiratory syndrome virus; PRRSV; propidium monoazide; PMA; ethidium bromide monoazide; EMA.

1.0 Introduction

The emergence of new viral diseases represents a constant threat to human and animal health. Fortunately, in the past decade, accesses to novel technologies have improved the detection and identification of unknown viruses in clinical samples. Most of these novel virus identification processes are based on viral genome detection using new technologies like pan-viral DNA microarrays (PVDA) and high-throughput sequencing (HTS).

The first PVDA, which contained 1,600 oligonucleotides probes targeting highly conserved DNA sequences of 140 distinct selected viral genomes, was reported in 2002 [1]. Since then, the PVDA has been further developed and includes, in its latest version, 36,000 oligonucleotides probes targeting approximately 1,500 distinct viral genomes [2]. This technology has been used to rapidly identify viruses involved in human illness, like severe acute respiratory syndrome (SARS) [3], and in animal diseases [2]. Use of this technology is of interest as the results can be generally obtained within a day and does not require other advanced technologies for results interpretation [2]. However, PVDA is dependent on the selected probes it contains and their tolerance to nucleotide mismatch during the DNA hybridization process required for the detection and identification of viruses in a clinical sample [4].

Decreasing costs has made HTS technology more accessible and consequently, its use in identifying novel or unknown viruses affecting humans, animals or plants has increased [5-7]. It has even led to the discovery of unforeseen viruses in clinical samples [8].

Metagenomic sequencing has the potential to determine the entirety of the nucleic acid sequences within a sample, including viral nucleic sequences of interest [9]. The metagenomic DNA sequences obtained with HTS are then compared to a genomic database in order to identify the nucleic acid sequences associated with known viruses [10]. One major benefit to metagenomic sequencing of clinical samples is the potential to detect and assemble the genome of novel viruses [11].

Although both PVDA and HTS have led to the discovery of new viruses in the last years, especially from isolated viruses, both techniques are negatively impacted by the presence of nucleic acid found in clinical samples, mainly host genomic DNA/RNA [10, 12]. The high host to viral DNA ratio in extracted clinical samples greatly decreases PVDA sensitivity since most of the amplified labeled DNA corresponds to host DNA [12]. For sequencing, depending on the method of tissue preparation and viral particle concentration, again the high host to viral DNA ratio decreases the sensitivity of the technology [13]. As more reads must be obtained in order to detect the presence of a virus in a clinical sample, this can increase sequencing costs and lower throughput while creating a potential bioinformatics bottleneck. Thus, in order to improve viral detection in clinical samples with HTS and PVDA, the levels of host genomic DNA must be lowered. This is generally done by treating samples with a combination of ultracentrifugation, filtration and/or nuclease treatment (typically DNase and/or RNase treatment) [9, 14]. As these methods can introduce bias in viral identification (9) the development of alternative methods to lower host genomic material in clinical samples is of interest.

Ethidium bromide monoazide (EMA) and its analog propidium monoazide (PMA), when combined with PCR, allow the quantification of living cells such as bacteria [15-18]. Both are azide-bearing, DNA/RNA-intercalating dyes that only cross damaged lipid membrane barriers. Both dyes can bind and covalently crosslink DNA/RNA when the azide group is converted to a highly reactive nitrene radical upon exposure to bright visible light. Thereafter, they are easily inactivated and the unbound inactivated EMA/PMA remains free in solution. EMA/PMA-generated DNA/RNA cross-linking strongly inhibits reverse-transcription and PCR amplification of the EMA/PMA modified genomes while unmodified genomes from presumptively living bacteria (which possesses intact membranes) can be amplified [17]. Interestingly, EMA/PMA treatments have been used to distinguish infectious from non-infectious viruses such as Hepatitis A virus, coxsakievirus, echovirus, norovirus and poliovirus, suggesting that intact virus particles have the potential to protect their genetic material from EMA/PMA chemicals [19].

In theory, EMA/PMA could be used to prevent host genomic amplification during the PCR steps that are conducted within PVDA and HTS assays, while the viral genome within intact virions are inaccessible to the dyes during treatment before amplification. The main objective of this study was to determine if EMA or PMA treatments can increase the efficacy of PVDA and HTS to detect viruses in clinical samples.

2.0 Materials and methods

2.1. Cell and virus strains

MARC-145 cells were maintained as described previously and were used for virus production [20]. The Porcine reproductive and respiratory syndrome virus (PRRSV) strain used to spike tissue and sera samples was the IAF-Klop reference strain [21]. The PRRSV IAF-Klop strain stock was obtained following three cycles of freeze-thaw of PRRSV MARC-145 infected cells. Afterward, the virus stocks were maintained at -70°C until needed. The infectious dose of the stocks was calculated from MARC-145 infected cells by the Kärber method as described previously [22]. Virus titers were expressed in tissue culture infectious dose 50% per mL ($TCID_{50}$/mL).

2.2 PRRSV spiked tissues and positive clinical samples

Lung and blood samples were collected from negative control and PRRSV experimentally infected piglets. Animals care was done according to the guidelines of the Canadian Council of Animal Care and the protocol approved by the Institutional Animal Care Committee (Protocol 12-Rech-1669). The PRRSV strain involved in this infection was PRRSV FMV12-1425619 (GenBank accession number KJ1888950). Sera of non-infected and infected piglets were collected at different time post-infection (pi) and kept at -70°C until needed. Viral load in samples was determined with a specific PRRSV RT-qPCR assay as previously described [8]. Lung samples were collected at necropsy at 28

days pi and stored at -70°C until needed. Three infected lung samples (PRRSV titers: 1) 4929, 2) 4336 and 3) 9408 $TCID_{50}$/g) were selected. Two sera samples were selected from PRRSV positive clinical sera samples (PRRSV titers of 1) 1059 and 2) 7413 $TCID_{50}$/mL) submitted to the Molecular Diagnostic Laboratory (MDL) of the University of Montreal were selected and stored at -70°C until needed. PRRSV negative swine lung samples or PRRSV negative swine sera samples were spiked with a known quantity of the PRRSV IAF-Klop strain to a final concentration of either 5,000 $TCID_{50}$/mL or 50,000 $TCID_{50}$/mL. Lung tissue samples (spiked samples or PRRSV positive clinical samples; 100 mg of tissue in 1 mL of PBS with glass beads) were homogenized twice for five minutes in a Mini BeadBeater 96 Homogenizer, centrifuged one min at 10 000 rpm in a table top centrifuge and kept at 4°C until used. Serum samples (spiked samples or PRRSV positive clinical samples) were kept at 4°C once thawed.

2.3 Samples processing

2.3.1 Ultracentrifugation

Lung tissue homogenate and serum samples (spiked with PRRSV or clinical samples) were ultracentrifuged for 3h at 25,000 rpm in a Sorvall TH-641 swinging bucket rotor at 4°C through 1mL of a 20% sucrose cushion in TNE buffer (20 mM Tris-HCl (pH 8.0), 150 mM NaCl and 2 mM EDTA). The virus pellets were re-suspended in TNE buffer to the initial sample volume prior to ultracentrifugation. Non-ultracentrifugated sample

aliquots were kept at 4°C for the duration of the ultracentrifugation step. Clinical samples (lung tissue and serum) were assessed by ultracentrifugation only.

2.3.2. Ethidium bromide monoazide and propidium monoazide treatments

EMA and PMA (Biotium, Hayward, CA) were reconstituted according to manufacturer's recommendation. Stock solutions were then diluted in RNase-free water to a working concentration of 2 mM. Both stock and working solutions were kept at -20°C until used. Lung tissue homogenates with and without ultracentrifugation (spiked or clinical samples) and sera samples (spiked or clinical samples) were subsequently treated with EMA (final concentrations of 100 µM), PMA (final concentrations of 100 µM), or with an equivalent volume of water. Treated samples were then incubated in the dark for five minutes at room temperature, five minutes on ice, and then exposed during ten minutes to two 500 watt halogen light sources (at a distance of 20 cm from the light source). Micro-centrifuge tubes were kept on ice during light exposure to avoid excessive heating.

2.4. Total nucleic acid extraction

Following treatments, total DNA and RNA were extracted using a phenol-chloroform-isoamyl alcohol. Briefly, 200 µL of a phenol solution at pH 7.6-7.8 (UltraPure™ buffer-saturated phenol; Invitrogen, Burlington, ON), 200 µL of molecular grade chloroform (Fisher scientific, Ottawa, ON) and 20 µL of isoamyl alcohol (Fisher scientific) were added to each sample. Samples were then homogenized and centrifuged in a table-top

microcentrifuge for 1 min at 13,000 rpm. The supernatant was kept and assessed for a second phenol-chloroform step. Finally, the supernatant was treated twice with 200 µL of chloroform and total DNA and RNA precipitated by adding 500 µL of ethanol and 20 µL of sodium acetate (3M, pH 5.2) and incubation at -70°C overnight followed by centrifugation for 30 min at 4°C in a table-top microcentrifuge at 13,000 rpm. The pellet representing total nucleic acid was resuspended in 50 µL of RNase-free water and stored in a freezer at -70°C until used.

2.5. PRRSV and host genome quantification by qPCR and RT-qPCR

PRRSV and swine β-Actin (representing swine host genomic DNA) were quantified in extracted DNA/RNA by RT-qPCR and qPCR, respectively. Equal sample volumes were used for each test to ensure comparable results. PRRSV was quantified in DNA/RNA extracted from tested samples using the commercial EZ-PRRSV™ MPX 4.0 Real Time RT-PCR kit (Tetracore, Rockville, Maryland, USA), following the manufacturer's recommendations. The β-actin quantification was done by qPCR using the SsoFast™ EvaGreen® Supermix kit (Bio-rad, Hercules, CA, USA) in order to evaluate host genome in spiked and clinical samples following each treatment. Samples were diluted in RNase free water (1:16) prior to β-actin qPCR tests. The PCR amplification program for β-actin quantification consisted of an enzyme activation step of 3 min at 98°C followed by 40 cycles of a denaturing step (2 s at 98°C) and an annealing/extension step (5 s at 58°C) using the following primers: forward primer (5'- ATCTTCATGAGGTAGTCGGTCAGG - 3') and reverse primer (5'- ACCACTGGCATTGTCATGGACTCT -3'). Both primers

were selected to achieve amplification efficiency between 90 and 110% (data not shown) and were designed from the NCBI GenBank mRNA sequences using web-based software primerquest from Integrated DNA technologies. All amplification steps were done on a Bio-Rad CFX 96 apparatus with results expressed as Ct values.

2.6. DNA/RNA samples amplification

Following treatment and extraction, total nucleic acid samples used for all experiments (i.e. both array detection and PMA/EMA treatment) were amplified using a modified random PCR protocol [1, 2]. Random-amplified samples for PVDA testing were spiked with 225 pg of purified pUC19 plasmid DNA and used both as a positive control and localization marker on array slides.

2.7. DNA array

2.7.1 DNA array development

Probe sequences targeting PRRSV American strains were selected from the PRRSV probes used on the ViroChip developed by Wang *et al.* (2002)[1] and were deduced from sequences alignment of full and partial PRRSV genomic sequences gathered from the National Center for Biotechnology Information (NCBI) GenBank database using Geneious pro software, version 5.6.6 (Biomatters, Auckland, New Zealand [http://www.geneious.com/]). The PRRSV homology of candidate probes was verified

with BLASTN. Each selected oligonucleotide probe (70-mers) was unique and was targeting specific PRRSV conserved regions. Reverse and forward sequences of 17 conserved regions were selected for the PRRSV genotype 2 strains. Two probes were also selected to target a specific region of pUC19 plasmid DNA as an array positioning control and one probe was selected as a negative hybridization control. A total of 37 probes were selected and were synthesized by Eurofins MWG Operon (Huntsville, AL, USA). These probes are reported in the Gene Expression Omnibus (GEO) NCBI database (accession number GSE62910). DNA array spotting was done at the National Research Council Canada, as previously described [23].

2.7.2. DNA array hybridization and analysis

After RT and PCR random-amplification steps, PCR products were incubated with aminoallyl (aa)-dUTP (Invitrogen, Burlington, ON, Canada) in the presence of Klentaq (Clontech) as previously described [1, 2]. The generated aa-DNA was purified with the QIAquick PCR purification kit (QIAGEN, Toronto, ON, Canada), re-suspended in 30 µL of RNase-free water and supplemented with 3 µL of 1M sodium bicarbonate. Thereafter, the aa-DNA was incubated for 1h in the presence of 1:10 DMSO-reconstituted Cy3 Mono-Reactive Dye (GE Healthcare Life Sciences, Pittsburgh, PA, USA). Labeled DNA was then purified using a QIAquick PCR purification kit and assessed for quality using a NanoDrop 1000 spectrophotometer (Fisher Scientific, Toronto, ON, Canada).

Microarray slides were pre-hybridized at 50°C for 1h with 85.5 µL of DIG easy Hyb buffer (Hoffmann-La Roche Limited, Mississauga, ON, Canada) supplemented with 4.5 µL of 10% (w/v) bovine serum albumin (Invitrogen, Burlington, ON, Canada) under 22mm x 60mm x 0.25mm Grace Bio-Labs Hybrislip™ coverslips (Sigma Aldrich, Oakville, ON, Canada). Subsequently, cover slips were removed by dipping the glass slides into 0.1x SSC (15mM NaCl, 1.5mM sodium citrate) and the slides dried by a quick centrifugation. Total Cy3-labeled DNA (typically 3 µg) was dried in a Speedvac (Fisher scientific) then suspended in 7 µL of DIG Easy Hyb buffer. Afterwards, the DNA was denatured for five min in a boiling water bath followed by five min of incubation on ice. Samples were hybridized overnight in a water bath at 50°C under 22 mm x 22 mm Grace Bio-Labs Hybrislip™ (Sigma Aldrich). Finally, coverslips were removed in 0.1x SSC, 0.1% (V/V) sodium dodecyl sulfate (SDS) and the slides washed three times in 0.1x SSC, 0.1% (v/v) SDS and once in 0.1x SSC for five minutes per wash. Hybridized arrays were imaged using a fluorescence scanner (ScanArray; Perkin Elmer, Mississauga, ON, Canada) and ScanArray software version 1.1. Fluorescent spot intensities were scanned at a laser fluorescent intensity of 80 to 100 and quantified using ScanArray software version 1.1. DNA array fluorescent intensity results were analysed with Microsoft Excel™. The intensity of each spotted probe was compared to the average intensity of the two negative control spots. For a probe to be considered positive, the average of signal-to-noise fluorescence ratios of their duplicate spots had to be ≥ 2.0.

2.8. High-throughput sequencing

271 The random amplified samples were end-repaired and A-tailed using KAPA High
272 Throughput Library Preparation Kit with SPRI solution and Standard PCR Library
273 Amplification/Illumina series (Kapa Biosystems, Wilmington, MA, USA). Illumina
274 TruSeq HT dual indexed adapters (Illumina, SanDiego, CA, USA) were ligated to the
275 amplified samples and the libraries were amplified with the KAPA kit. After the final
276 cleanup, the quality of the libraries were assessed on High Sensitivity DNA Chips
277 (Agilent, Santa Clara, CA, USA) using a 2100 BioAnalyzer (Agilent). Equal amounts of
278 each library were pooled and sequenced on an Illumina MiSeq (2 x 300 paired-end reads,
279 dual-indexed) at the Plateforme d'Analyses Génomiques de l'institut de Biologie
280 Intégrative et des Systèmes de l'Université Laval (Quebec, QC, Canada). Raw sequencing
281 reads were trimmed for the random amplification primers and mapped to the *Sus scrofa*
282 genome v10.2 and the PRRSV IAF-Klop viral genome sequence using the gsMapper
283 application of Newbler v2.9.
284
285 *2.9. Statistical analysis*
286
287 A parametric one-way ANOVA model, followed by Tukey's Multiple Comparison tests
288 (GraphPad PRISM Version 5.03 software) was used to determine if a statistically
289 significant difference exists between the quantification of targeted genes (β-actin or
290 PRRSV) for each treated and untreated samples, as evaluated by qPCR and RT-qPCR. A
291 non-parametric one-way ANOVA model, followed by Dunn's Multiple Comparison Test
292 (GraphPad PRISM software), was used to determine if a statistically significant
293 difference exists between the mean relative fluorescent intensity of each PRRSV

294 detection probe on the DNA array for each treated and untreated samples. The ratio of
295 positive PRRSV reads compared to the total amount of reads following each treatment
296 was analysed with a mixed linear model, with trial number as a random factor and
297 treatment as a fixed factor, followed by Tukey's Multiple Comparison tests (SAS version
298 9.3 software, Cary, NC, USA). HTS results were also evaluated individually using chi-
299 square tests in order to determine if the odds to obtained positive PRRSV reads were
300 different following each selected treatments, compared to the total amount of reads not
301 related to PRRSV (GraphPad PRISM software). Finally, a Spearman's non-parametric
302 correlation was also used to evaluate the relation between the presence of DNA
303 contaminant reads and positive PRRSV reads as well as the ratio of PRRSV reads and the
304 PRRSV percent of coverage in HTS results (SAS version 9.3 software). Differences were
305 considered statistically significant with a $P<0.05$, with the exception of chi-square tests
306 results where only $P<0.003$ ($P<0.05/15$) were considered significant.

3.0. Results

3.1. EMA or PMA treatment effects on host-genomic DNA and PRRSV detection in spiked samples

In order to select an effective concentration for EMA and PMA treatment, a preliminary experiment was done with PRRSV spiked lung samples (5,000 $TCID_{50}$/mL; data not shown) and a final EMA and PMA concentration of 100 µM was selected for the realisation of subsequent experiments. Using this concentration, the effect of EMA or PMA treatments on virus genome's presence and host genomic DNA was further evaluated in tissue samples spiked with known quantities of PRRSV (5,000 $TCID_{50}$/mL or 50,000 $TCID_{50}$/mL) but also in serum samples spiked with the same PRRSV quantities. Results with lung homogenates spiked with both concentration of PRRSV indicate that EMA and PMA treatments significantly lowered PRRSV detection compared to non-treated samples (Fig. 1A; $P<0.001$ and $P<0.05$ respectively). However, PRRSV detection was more negatively affected following EMA treatment (Fig. 1A; $P<0.01$), indicating that PMA treatments have less negative impact on RT-qPCR PRRSV detection as observed in the preliminary experiment. Ultracentrifugation had a significant positive impact on PRRSV detection in lung tissue homogenates spiked with either concentration of PRRSV following EMA treatment (Fig. 1A; $P<0.001$). In lung tissue homogenates spiked only with the highest concentration of PRRSV, ultracentrifugation had also increased significantly PRRSV detection following no treatment and following PMA treatment (Fig. 1A; $P<0.01$ and $P<0.05$ respectively). In spiked serum samples,

only EMA treatment significantly decreased PRRSV detection (Fig. 1A; $P<0.05$). Ultracentrifugation significantly improved PRRSV detection in non-treated spiked sera samples (Fig. 1A; $P<0.001$) and significantly decreased PRRSV RT-qPCR detection in EMA treated spiked sera samples (Fig. 1A; $P<0.001$).

Both EMA and PMA treatments, combined or not with ultracentrifugation, were equally efficient in lowering β-actin amplification in PRRSV spiked lung tissue homogenates (Fig. 1B; $P<0.001$). Ultracentrifugation of non-treated spiked lung tissue samples also slightly lowered β-actin amount (Fig. 1B; $P<0.05$). In spiked sera samples, β-actin was already at the limit of detection in untreated samples, which indicates that hg/cont is much lower in sera compared to lung tissue homogenates (Fig. 1B).

3.2. DNA array sensitivity with EMA or PMA treated PRRSV spiked tissues.

After nucleic acid extraction, random amplification and DNA labeling were done on each spiked samples and the fluorescence of each PRRSV probe was measured following the hybridization of labeled samples on DNA arrays and compared to the negative probe fluorescence intensity (see accession number GSE62910 in the GEO NCBI database for raw data). Surprisingly, probes signals intensity and positivity were varying between experiments. For lung samples spiked with PRRSV (5,000 $TCID_{50}$/mL), a low number of slightly positive probes (probes relative signal intensity < 5) were detected in untreated samples (Fig. 2A). Interestingly, multiple probes with relative high fluorescence

intensity (>5) were detected following PMA treatment when combined with ultracentrifugation and were associated with a significantly higher number of positive probes compared to untreated samples with ultracentrifugation and the PMA treated samples without ultracentrifugation (Fig. 2A; $P<0.001$). In contrast, the fluorescence intensity of detected positive probes of EMA-treated lung homogenates spiked with PRRSV (5 000 $TCID_{50}$/mL) were significantly lower than all other experimental groups, indicating a lower chance to detect positive probes following treatment with EMA ($P<0.05$; Fig. 2A).

In lung samples spiked with a higher concentration of PRRSV (50,000 $TCID_{50}$/mL), a small number of low intensity positive probes (fluorescence relative signal <5) or no positive probes (fluorescence relative signal <1) were detected in untreated samples (Fig. 2B). A significant increase in the number of high intensity positive probes (fluorescence relative signal >5) were found in samples treated with PMA, with or without ultracentrifugation, when compared to untreated samples without ultracentrifugation (Fig. 2B; $P<0.001$). Surprisingly, only a few positive probes were found in samples treated with ultracentrifugation and EMA, similar to untreated samples. However, the number and intensity of probes in samples treated with EMA without ultracentrifugation were significantly higher compared to untreated samples without ultracentrifugation ($P<0.001$).

In spiked sera samples, it was interesting to observe that all experimental groups had large numbers of probes with high intensity fluorescence signal against PRRSV,

including untreated samples (Fig. 2C), indicating overall that all treatments (ultracentrifugation versus PMA/EMA) did not improve DNA array sensitivity when used with sera samples. Interestingly, PMA treatment compared to untreated sera reduced the PVDA sensitivity even if a high number of high intensity positive probes were observed in PMA treated sera (Fig. 2C; $P<0.05$).

3.3. HTS efficiency with EMA or PMA treated PRRSV spiked tissues.

DNA sequences obtained from sequencing experiments were compared to the swine mitochondrial (GenBank accession numbers NC_000845) and chromosomal genomic DNA sequences (GenBank accession numbers NC_010443-NC_010462) as well as the full genetic sequence of the PRRSV strain used in this study. All HTS reads associated with PRRSV sequence were also considered to evaluate the virus coverage obtained following each treatment combination.

In two out of three experiments, PMA and EMA treated PRRSV spiked lung tissues (5,000 $TCID_{50}$/mL) had a significant increase in the number of PRRSV reads when compared to untreated samples, as revealed by the chi-square analysis (Fig. 3A; $P<0.001$). A strong correlation was found between PRRSV percent of coverage and the higher number of reads following treatment ($r=0.81$, $P<0.001$), indicating that the increase in PRRSV coverage is related to the augmentation in PRRSV reads. However, when all three HTS experiments were combined together in the statistical analyses, no significant differences were obtained between the amount of PRRSV specific reads

399 against the total amount of reads (Fig. 3A; *P*=0.4358) or the PRRSV percentage of
400 coverage (Fig. 3B; *P*=0.5585). Noteworthy, the experiment showing no improvement in
401 the amount of PRRSV reads was associated with a higher level of hg/cont as revealed in
402 Fig. 3C. In fact, a strong negative correlation was found between the ratio of PRRSV
403 reads (number of PRRSV reads / total number of reads) and the ratio of host genomic
404 reads (number of genomic reads / total number of reads) (r=-0.66, *P*<0.005) in these three
405 experiments. The reason why this experiment indicates a low effect of treatments on the
406 percentage of host genomic DNA and possibly its transcripts is currently unknown but
407 might be caused by a higher rate of host genome released during the tissue preparation.
408
409 In tissue homogenates spiked with higher amounts of PRRSV (50,000 $TCID_{50}$/mL),
410 although not statistically significant, statistical analyses revealed a tendency for an
411 increase of the ratio of PRRSV reads when the two HTS experiments were taken into
412 account (Fig. 3D; *P*=0.0772). Moreover, a significant variation was found in PRRSV
413 coverage in those samples (Fig. 3E; *P*=0.0103) and a significant increase in PRRSV
414 coverage in PMA treated samples without ultracentrifugation was observed (Fig. 3F;
415 *P*=0.0082). Ultracentrifugation of samples containing the higher amount of PRRSV did
416 not improve either the PRRSV number of reads or its coverage (Fig. 3D and E).
417
418 In PRRSV spiked sera, there was an important increase in the ratio of PRRSV reads and
419 coverage in untreated samples compared to results obtained with lung homogenates (Fig.
420 3). This was in accordance with the lower amount of host genomic DNA detected in sera
421 of untreated samples (Fig. 1B, 3iC and 5F). According to the chi-square analysis, a slight

increase in PRRSV ratio of reads was observed following PMA and EMA treatments. However, ultracentrifugation lowered the ratio of PRRSV reads when combined with EMA treatment (Fig. 3G) and was associated with a lower PRRSV coverage (Fig. 3H).

3.4. High-throughput sequencing efficiency with EMA or PMA treated PRRSV positive clinical samples.

In order to confirm the effectiveness of EMA and PMA treatments to increase the sensitivity of PVDA and HTS to detect viruses, both techniques were evaluated with different PRRSV positive clinical samples gathered from PRRSV experimentally infected piglets and clinical samples submitted to the MDL. All clinical samples were ultracentrifuged and treated with PMA or untreated. PMA was selected because overall, our previous findings indicated that PMA treatment was more efficient than EMA with regards to PRRSV RT-qPCR, PVDA and HTS detection.

PRRSV and the amplifiable host genomic DNA copy numbers were significantly reduced in PMA treated lung tissues (Fig. 4A and Fig. 4B; $P<0.001$). The important PRRSV decrease in PMA treated samples was probably caused by the presence of damaged virions and/or non-encapsidated viral genomes within clinical samples. PRRSV was undetectable in all untreated samples by DNA array (Fig. 4C). Interestingly, several PRRSV probes were found positive for lung #1 in PMA treated samples (Fig. 4C; $P<0.001$). However, PRRSV was not detected by DNA array in lung #3 and few PRRSV probes were found positive with lung #2 following PMA treatment (Fig. 4C). For HTS

results, when the three clinical cases are taken into account together, results revealed that PMA tends to increase the number of reads obtained from lung tissues compared to untreated samples (Fig. 4D; P=0.0529). Nonetheless, PMA treatment significantly improved the percentage of PRRSV HTS reads in two out of three clinical cases tested (Fig 4D, P<0.001). Although these experiments showed an important increase in PRRSV genomic coverage (3 to 6-fold), no overall significant differences were detected (Fig. 4E; P=0.1117). No increase in PRRSV coverage was detected for lung #1 (Fig. 4E). Although lung #3 showed no increase in the percentage of PRRSV reads, an important increase (6-fold) of PRRSV coverage was detected. However, a higher host genomic DNA content is observed in this sample as suggested by the lower Ct values detected in treated and untreated sample (Fig. 4B) and the higher genomic percentage of reads (Fig. 4F), compared to lungs #1 and #2.

PMA treatment of clinical sera had no significant effect on amplifiable PRRSV (Fig. 5A) or host genomeic DNA levels (Fig. 5B, Ct values >40). In agreement with spiked sera samples, PRRSV was strongly detected with all experimental conditions tested from both clinical sera samples by PVDA and HTS (Fig. 5C, D and E). However, the number of PRRSV positive probes was significantly lower in PMA treated samples for both sera samples (Fig. 5C; P<0.01). HTS results revealed that the PMA treatment also lowers the ratio of PRRSV reads in one out of two cases (Fig. 5D; P<0.001). PRRSV coverage was high in all cases (over 60%), except for serum 2 treated with PMA where the PRRSV percent of coverage was slightly higher than 40%.

4.0 Discussion and conclusion

Although generally preferable, direct virus isolation and identification from clinical samples are not always possible and can sometimes lead to mis-identification of the etiological agent (29). Consequently direct identification within clinical samples, using newly available genomic technologies, is desirable. During the past decade, PVDA and HTS technologies have led to the discovery of new viruses [24-31]. However, the sensitivity of these technologies can suffer from excessive amounts of contaminating host DNA and RNA. Any methodology to decrease the masking effect that these contaminants have on viral detection is important.

A new approach to lower the host genomic DNA within clinical samples is presented in this report. Our results support an increase in sensitivity of PVDA and HTS in lung tissue samples treated with PMA and, to a lesser extent, with EMA. This treatment can be done as a standalone treatment or in combination with other treatments like ultracentrifugation. Surprisingly, adding an ultracentrifugation step before the EMA or PMA treatment sometimes lowers the sensitivity of both PVDA and HTS. This is probably caused by a physical degradation of viral particles during ultracentrifugation, making these particles more sensitive to EMA or PMA treatment.

Interestingly, our results indicate a higher sensitivity of PVDA and HTS with sera samples, compared to tissue homogenates. This may be explained by two phenomena. Firstly, there is a lower concentration of hg/cont in sera samples compared to tissue

490 homogenates, as demonstrated by β-actin DNA measurements for both untreated sera
491 samples and tissue homogenates. Secondly, the random amplification method used to
492 increase the amount of DNA for both HTS and PVDA techniques amplify all nucleic
493 DNA and RNA sequences found in a sample, including hg/cont. Thus, the sensitivity of
494 both PVDA and HTS will be highly affected by the ratio of the viral genetic material of
495 interest over the total amount of hg/cont found in samples. This indicates that the initial
496 amount of hg/cont in clinical samples has a deep impact on PVDA and HTS sensitivity
497 for the identification of viruses. The variation in hg/cont content within nucleic acid
498 extraction from clinical samples may explain the high standard deviation obtained within
499 the results of PVDA and HTS with lung tissue homogenates spiked with the lowest
500 PRRSV concentration. This variation can be explained by multiple factors like the
501 extraction method, the tissue quality and the homogenization process. Also, the virus
502 integrity in clinical samples will have to be taken into account to avoid virus particle
503 degradation during the tissue manipulation, storage and homogenization, which can
504 sensitize viral particles to EMA or PMA treatment. This is especially true for clinical
505 samples where the virus type, the integrity of tissue (dead animal, sample conservation at
506 room temperature, etc.), the amount of hg/cont or the presence of PCR inhibitors will
507 have an critical impact on PVDA and HTS sensitivity. In addition, the efficiency of the
508 random amplification process could be affected by the viral genome itself (its sequence
509 and its secondary structures), making some viral genomes less compatible with the use of
510 random amplification prior to HTS and PVDA [32]. It is important to note that our work
511 utilized a unique virus, and does not take into account the efficiency of EMA and PMA
512 treatments on other types of viruses. However, the outer structure of the virions of several

513 viruses can protect the viral genetic material from PMA treatment [19], suggesting that
514 this treatment should reduce the influence of hg/cont when used for clinical samples with
515 other virus types.
516
517 Multiple reports have revealed differential PVDA and HTS sensitivities for viral
518 detection and identification in clinical samples following the use of different treatment
519 combinations, including ultracentrifugation and nuclease treatment [24, 33]. The lowest
520 viral load in spiked serum samples was shown by Nicholson and collaborators (2011)
521 where they determined the limit of PVDA detection for PRRSV in spiked serum samples,
522 not subjected to nuclease treatment, to be 10,000 $TCID_{50}$/mL [34]. Similar results were
523 obtained in the present study but with lower virus concentration (5 000 $TCID_{50}$/mL).
524
525 Previous studies have reported the ratio of viral sequencing reads to total number of reads
526 ranged between 0.00019% and 2.8% from non-nuclease treated serum and
527 nasopharyngeal aspirates [25, 26]. However, Mishra and collaborators (2014) showed a
528 ratio (viral reads/ total number of reads) of 0.00012% by HTS from muscle tissue
529 samples [24]. In this study, rRNA was depleted and a DNase treatment was done
530 following the RNA extraction, in order to lower hg/cont in that clinical sample [24].
531 These results are in accordance with our work since the percentage of PRRSV specific
532 reads ranged between 0% and 0.00059% in spiked untreated samples and between 0%
533 and 0.19785% in EMA and PMA-treated spiked samples. In contrast, Djikeng and
534 collaborators (2008) have reported much higher ratios of host genomic content to isolated
535 virus ranging between 3 and 40% [33] while Nakamura and collaborators (2009) have

reported higher genomic content ranging between 90.0% and 94.6% [26]. The variation of the host genomic DNA content between those studies is probably explained by the difference in purification and treatment methods used for each virus, like the use of gradient density centrifugation to obtained highly purified viruses with lower hg/cont [33]. The ratio of host genomic DNA content reported in our study varied between 67.50% and 77.40% in untreated samples and between 40.91% and 72.76% in EMA and PMA-treated samples. This represents an improvement over the results reported by Nakamura and collaborators (2009) where no purification methods were used, and less than the ratio of hg/cont reported by Djikeng and collaborators (2008). However, in the last case, the viruses were isolated from cell culture and the virions were subsequently purified.

Thus, our results indicates that EMA and PMA treatments improve the sensitivity of PVDA and HTS to detect viruses in clinical samples contaminated with hg/cont. Furthermore, EMA/PMA treatments are faster and easier to perform than a nuclease treatment. Firstly, they require shorter incubation times compared to nuclease treatment which would result in faster processing of clinical samples in a diagnostic laboratory, and should increase the robustness of the method by exerting less stress on temperature sensitive viral particles. Secondly, EMA and PMA are easy to inactivate through light exposure, subsequently leaving EMA/PMA molecules unable to destroy newly exposed viral genomic material following nucleic acid extraction, unlike the case when residual nucleases may still be present after nucleic acid purification steps. This is especially true if RNase or DNase treatment is being used in the presence of RNA or DNA viruses,

559 respectively. PMA/EMA treatments were more effective in tissue samples compared to
560 sera samples which was presumably due to lower hg/cont in the latter. PMA was more
561 effective than EMA in improving HTS and PVDA sensitivity for viral detection in
562 clinical samples. This would be explained by the fact that EMA can leak through
563 phospholipid bilayer membrane [35]. Moreover, ultracentrifugation appears to lower the
564 sensitivity of HTS and PVDA following some treatments, possibly because of direct
565 physical degradation of the viral particle. In conclusion, pre-treatment of clinical samples
566 with EMA, and especially PMA, represents an interesting novel approach that improves
567 PVDA and HTS sensitivity for the identification of viruses from clinical samples.
568

569 **5.0 Acknowledgements**

570 This manuscript was financially supported by the Canadian Swine Health Board (CSHB).
571 C.A. Gagnon was financially supported by the Natural Sciences and Engineering
572 Research Council of Canada (NSERC). The authors wish to thank Guy Beauchamp and
573 Philippe Garneau for scientific support, and Miria Elias for technical support. The authors
574 have no conflict of interest.
575

576 **6.0. References**

577

578 [1] Wang, D., Coscoy, L., Zylberberg, M., Avila, P. C., *et al.*, Microarray-based detection
579 and genotyping of viral pathogens. *Proceedings of the National Academy of Sciences of*
580 *the United States of America* 2002, *99*, 15687-15692.

[2] Chen, E. C., Miller, S. A., DeRisi, J. L., Chiu, C. Y., Using a pan-viral microarray assay (Virochip) to screen clinical samples for viral pathogens. *Journal of visualized experiments : JoVE* 2011.

[3] Wang, D., Urisman, A., Liu, Y. T., Springer, M., *et al.*, Viral discovery and sequence recovery using DNA microarrays. *PLoS biology* 2003, *1*, E2.

[4] Delwart, E., Animal virus discovery: improving animal health, understanding zoonoses, and opportunities for vaccine development. *Current opinion in virology* 2012, *2*, 344-352.

[5] Massart, S., Olmos, A., Jijakli, H., Candresse, T., Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus research* 2014, *188C*, 90-96.

[6] Lipkin, W. I., Firth, C., Viral surveillance and discovery. *Current opinion in virology* 2013, *3*, 199-204.

[7] Quinones-Mateu, M. E., Avila, S., Reyes-Teran, G., Martinez, M. A., Deep sequencing: Becoming a critical tool in clinical virology. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2014, *61*, 9-19.

[8] Bellehumeur, C., Boyle, B., Mandeville, I., Gagnon, C. A., High-throughput sequencing revealed the presence of an unforeseen parvovirus species in Canadian swine: the porcine partetravirus. *The Canadian veterinary journal. La revue veterinaire canadienne* 2013, *54*, 787-789.

[9] Hall, R. J., Wang, J., Todd, A. K., Bissielo, A. B., *et al.*, Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *Journal of virological methods* 2014, *195*, 194-204.

604 [10] Firth, C., Lipkin, W. I., The genomics of emerging pathogens. *Annual review of*
605 *genomics and human genetics* 2013, *14*, 281-300.
606 [11] Barzon, L., Lavezzo, E., Costanzi, G., Franchin, E., *et al.*, Next-generation
607 sequencing technologies in diagnostic virology. *Journal of clinical virology : the official*
608 *publication of the Pan American Society for Clinical Virology* 2013, *58*, 346-350.
609 [12] Kang, X., Qin, C., Li, Y., Liu, H., *et al.*, Improvement of the specificity of a pan-
610 viral microarray by using genus-specific oligonucleotides and reduction of interference
611 by host genomes. *Journal of medical virology* 2011, *83*, 1624-1630.
612 [13] Yang, J., Yang, F., Ren, L., Xiong, Z., *et al.*, Unbiased parallel detection of viral
613 pathogens in clinical samples by use of a metagenomic approach. *Journal of clinical*
614 *microbiology* 2011, *49*, 3463-3469.
615 [14] Clem, A. L., Sims, J., Telang, S., Eaton, J. W., Chesney, J., Virus detection and
616 identification using random multiplex (RT)-PCR with 3'-locked random primers.
617 *Virology journal* 2007, *4*, 65.
618 [15] Nocker, A., Sossa-Fernandez, P., Burr, M. D., Camper, A. K., Use of propidium
619 monoazide for live/dead distinction in microbial ecology. *Applied and environmental*
620 *microbiology* 2007, *73*, 5111-5117.
621 [16] Gatteschi, D., Fittipaldi, M., Sangregorio, C., Sorace, L., Exploring the no-man's
622 land between molecular nanomagnets and magnetic nanoparticles. *Angewandte Chemie*
623 2012, *51*, 4792-4800.
624 [17] Codony, F., Perez, L. M., Adrados, B., Agusti, G., *et al.*, Amoeba-related health risk
625 in drinking water systems: could monitoring of amoebae be a complementary approach to
626 current quality control strategies? *Future microbiology* 2012, *7*, 25-31.

[18] Nocker, A., Mazza, A., Masson, L., Camper, A. K., Brousseau, R., Selective detection of live bacteria combining propidium monoazide sample treatment with microarray technology. *Journal of microbiological methods* 2009, *76*, 253-261.

[19] Elizaquivel, P., Aznar, R., Sanchez, G., Recent developments in the use of viability dyes and quantitative PCR in the food microbiology field. *Journal of applied microbiology* 2014, *116*, 1-13.

[20] Beura, L. K., Sarkar, S. N., Kwon, B., Subramaniam, S., *et al.*, Porcine reproductive and respiratory syndrome virus nonstructural protein 1beta modulates host innate immune response by antagonizing IRF3 activation. *Journal of virology* 2010, *84*, 1574-1584.

[21] Gagnon, C. A., Lachapelle, G., Langelier, Y., Massie, B., Dea, S., Adenoviral-expressed GP5 of porcine respiratory and reproductive syndrome virus differs in its cellular maturation from the authentic viral protein but maintains known biological functions. *Archives of virology* 2003, *148*, 951-972.

[22] Gagnon, C. A., del Castillo, J. R., Music, N., Fontaine, G., *et al.*, Development and use of a multiplex real-time quantitative polymerase chain reaction assay for detection and differentiation of Porcine circovirus-2 genotypes 2a and 2b in an epidemiological survey. *Journal of veterinary diagnostic investigation : official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc* 2008, *20*, 545-558.

[23] Vuong, N. M., Villemur, R., Payment, P., Brousseau, R., *et al.*, Fecal source tracking in water using a mitochondrial DNA microarray. *Water research* 2013, *47*, 16-30.

[24] Mishra, N., Pereira, M., Rhodes, R. H., An, P., *et al.*, Identification of a Novel Polyomavirus in a Pancreatic Transplant Recipient With Retinal Blindness and Vasculitic Myopathy. *The Journal of infectious diseases* 2014.

650 [25] Yozwiak, N. L., Skewes-Cox, P., Stenglein, M. D., Balmaseda, A., *et al.*, Virus
651 identification in unknown tropical febrile illness cases using deep sequencing. *PLoS*
652 *neglected tropical diseases* 2012, *6*, e1485.
653 [26] Nakamura, S., Yang, C. S., Sakon, N., Ueda, M., *et al.*, Direct metagenomic
654 detection of viral pathogens in nasal and fecal specimens using an unbiased high-
655 throughput sequencing approach. *PloS one* 2009, *4*, e4219.
656 [27] Blomstrom, A. L., Belak, S., Fossum, C., McKillen, J., *et al.*, Detection of a novel
657 porcine boca-like virus in the background of porcine circovirus type 2 induced
658 postweaning multisystemic wasting syndrome. *Virus research* 2009, *146*, 125-129.
659 [28] Mihindukulasuriya, K. A., Wu, G., St Leger, J., Nordhausen, R. W., Wang, D.,
660 Identification of a novel coronavirus from a beluga whale by using a panviral microarray.
661 *Journal of virology* 2008, *82*, 5084-5088.
662 [29] Victoria, J. G., Kapoor, A., Dupuis, K., Schnurr, D. P., Delwart, E. L., Rapid
663 identification of known and new RNA viruses from animal tissues. *PLoS pathogens*
664 2008, *4*, e1000163.
665 [30] Hang, J., Forshey, B. M., Kochel, T. J., Li, T., *et al.*, Random amplification and
666 pyrosequencing for identification of novel viral genome sequences. *Journal of*
667 *biomolecular techniques : JBT* 2012, *23*, 4-10.
668 [31] Rosseel, T., Lambrecht, B., Vandenbussche, F., van den Berg, T., Van Borm, S.,
669 Identification and complete genome sequencing of paramyxoviruses in mallard ducks
670 (Anas platyrhynchos) using random access amplification and next generation sequencing
671 technologies. *Virology journal* 2011, *8*, 463.

672 [32] Codony, F., Fittipaldi, M., Lopez, E., Morato, J., Agusti, G., Well water as a possible
673 source of Waddlia chondrophila infections. *Microbes and environments / JSME* 2012, *27*,
674 529-532.
675 [33] Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., *et al.*, Viral genome sequencing
676 by random priming methods. *BMC genomics* 2008, *9*, 5.
677 [34] Nicholson, T. L., Kukielka, D., Vincent, A. L., Brockmeier, S. L., *et al.*, Utility of a
678 panviral microarray for detection of swine respiratory viruses in clinical samples. *Journal*
679 *of clinical microbiology* 2011, *49*, 1542-1548.
680 [35] Nocker, A., Cheung, C. Y., Camper, A. K., Comparison of propidium monoazide
681 with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal
682 of DNA from dead cells. *Journal of microbiological methods* 2006, *67*, 310-320.
683
684
685

7.0. Figures legends

Figure 1

PRRSV and host genomic detection in spiked tissues following EMA or PMA treatment using qPCR. Effect of EMA and PMA treatments, with or without ultracentrifugation, on A) PRRSV quantification and B) host genomic DNA (β-Actin) in lung tissue homogenates spiked with PRRSV (5,000 $TCID_{50}$/mL or 50,000 $TCID_{50}$/mL) or in serum spiked with PRRSV (5,000 $TCID_{50}$/mL). Results are expressed as Ct and were obtained from two to seven independent experiments. The results of each independent experiment (trial) are illustrated in Supplemental Figure 1. Sample ($TCID_{50}$/mL) represents the type of tissue spiked with PRRSV. Numbers in brackets represent the PRRSV concentration for each spiked sample expressed in $TCID_{50}$/mL. Open bars represent results obtained from samples processed without an ultracentrifugation step while filled bars represent results obtained from samples treated with an ultracentrifugation step. A Ct value of 37 (dashed line) represents the limit of detection of each qPCR test. Labeling of two sets of data with different letters indicates that these two sets of data are statistically different ($P<0.05$). Sets of data using letters with the same superscript number must be compared only together.

Figure 2

PRRSV detection in spiked samples by DNA array following EMA or PMA treatments. DNA array probes relative intensity from A) lung tissue homogenates spiked with PRRSV (5,000 $TCID_{50}$/mL), B) lung tissue homogenates spiked with PRRSV

709 (50,000 $TCID_{50}$/mL) and C) serum samples spiked with PRRSV (5,000 $TCID_{50}$/mL).
710 Dots are relative fluorescence intensity mean values of two identical probes gathered
711 from two to seven independent experiments (each experiment consisting of a duplicate of
712 34 PRRSV specific probes) and was calculated as followed: [(PFL-BFL)/BFL] where
713 PFL represents a PRRSV probe fluorescence intensity and BFL represents the basal
714 fluorescence level (negative control probe fluorescence). The results of each independent
715 experiment (trial) are illustrated in Supplemental Figure 2. The line represents the
716 fluorescence mean value of all probes. Open dot circles represent results obtained from
717 samples processed without an ultracentrifugation step while filled dot circles represent
718 results obtained from samples treated with an ultracentrifugation step. Labeling of two
719 sets of data with different letters indicates that these two sets of data are statistically
720 different (P<0.05). A probe relative intensity of 1 (dashed line) represents the lowest
721 limit of DNA array positive results.
722
723 ***Figure 3***
724 **PRRSV and host genomic detection efficiency in spiked tissue samples following**
725 **EMA or PMA treatment by high-throughput sequencing**. HTS results gathered from
726 A), B) and C) lung tissue homogenates spiked with PRRSV (5,000 $TCID_{50}$/mL); from D),
727 E) and F) lung tissue homogenates spiked with PRRSV (50,000 $TCID_{50}$/mL); and from
728 G), H) and I) serum samples spiked with PRRSV (5,000 $TCID_{50}$/mL). The amounts of
729 PRRSV specific reads compared to the total number of reads gathered from each HTS
730 run (expressed as %) are reported in panels A), D) and G) while the percentage coverage
731 of PRRSV recovered from the total number of PRRSV specific reads are reported in

panel B), E) and H. The host genomic specific reads compared to the total number of reads gathered from each HTS run (expressed as %) are reported in panels C), F) and I). Open bars represent results obtained from samples processed without an ultracentrifugation step while filled bars represent results obtained from samples treated with an ultracentrifugation step. The results from each experiment are expressed separately in each graphic. Labeling of two sets of data with different letters indicates that these two sets of data are statistically different ($P<0.05$). Sets of data using letters with the same superscript number must be compared only together. The overall *P*-values shown in the boxes represent the statistical analysis of treatments effects taking into account all the experimental groups.

Figure 4

Detection of PRRSV and of host genomic DNA in clinical lung samples by RT-qPCR, DNA array and high-throughput sequencing following PMA treatment. Results obtained from three clinical lung samples by A) PRRSV RT-qPCR; B) swine host genomic quantification (β-Actin) qPCR; C) DNA array; D) HTS PRRSV specific reads compared to the total amount of reads (expressed as %); E) HTS PRRSV percentage coverage recovered from the total number of PRRSV specific reads; and F) host genomic specific reads compared to the total amount of reads (expressed as %). Open bars or open circles represent results obtained from untreated samples while filled bars or filled circles represent results obtained from samples treated with PMA. A Ct value of 37 (dashed line) represents the limit of detection of each qPCR test. Dots are relative fluorescence intensity mean values of two identical probes gathered from three

independent experiments (each experiment consisting of a duplicate of 34 PRRSV specific probes) and was calculated as followed: [(PFL-BFL)/BFL] where PFL represents a PRRSV probe fluorescence intensity and BFL represents the basal fluorescence level (negative control probe fluorescence). The line represents the fluorescence mean value of all probes. A probe relative intensity of 1 (dashed line) represents the lowest limit of DNA array positive results. Results obtained from each clinical case are expressed separately in each panel. The overall *P* values shown in boxes represents the statistical analysis of treatments effects taking into account all the experimental groups. When two sets of data or group of data are labeled with an asterisk, it indicates that these two sets of data and group are statistically different (*** $P<0.001$). Labeling of two sets of data with different letters indicates that these two sets of data are statistically different ($P<0.05$). Only sets of data using letters with the same superscript number should be compared together.

Figure 5

Detection of PRRSV and of host genomic DNA in clinical sera samples by RT-qPCR, DNA array and high-throughput sequencing following PMA treatment. Results obtained from two clinical serum samples by A) PRRSV RT-qPCR; B) swine host genomic quantification (β-Actin) qPCR; C) DNA array; D) HTS PRRSV specific reads compared to the total amount of reads (expressed as %); E) HTS PRRSV percentage coverage recovered from the total number of PRRSV specific reads; and F) host genomic specific reads compared to the total amount of reads (expressed as %). Open bars or open circles represent results obtained from untreated samples while filled

bars or filled circles represent results obtained from samples treated with PMA. A Ct value of 37 (dashed line) represents the limit of detection of each qPCR test. Dots are relative fluorescence intensity mean values of two identical probes gathered from two independent experiments (each experiment consisting of a duplicate of 34 PRRSV specific probes) and was calculated as followed: [(PFL-BFL)/BFL] where PFL represents a PRRSV probe fluorescence intensity and BFL represents the basal fluorescence level (negative control probe fluorescence). The line represents the fluorescence mean value of all probes. A probe relative intensity of 1 (dashed line) represents the lowest limit of DNA array positive results. The results from each clinical case are expressed separately in each panel. The overall *P* values shown in boxes represent the statistical analysis of treatments effects taking into account all the experimental groups. When two sets of data or group of data are labeled with an asterisk, it indicates that these two sets of data and group are statistically different (*** $P<0.001$; ** $P<0.01$). Labeling of two sets of data with different letters indicates that these two sets of data are statistically different ($P<0.05$). Only sets of data using letters with the same superscript number should be compared together.

Supplemental Figure 1

Colour coded individual trial results of PRRSV and host genome detection in spiked tissues following EMA or PMA treatment using qPCR. Effect of EMA and PMA treatments, with or without ultracentrifugation (UC), on PRRSV quantification in lung tissue homogenates spiked with A) PRRSV (5,000 $TCID_{50}$/mL) or with C) PRRSV (50,000 $TCID_{50}$/mL) and E) serum spiked with PRRSV (5,000 $TCID_{50}$/mL); and on host

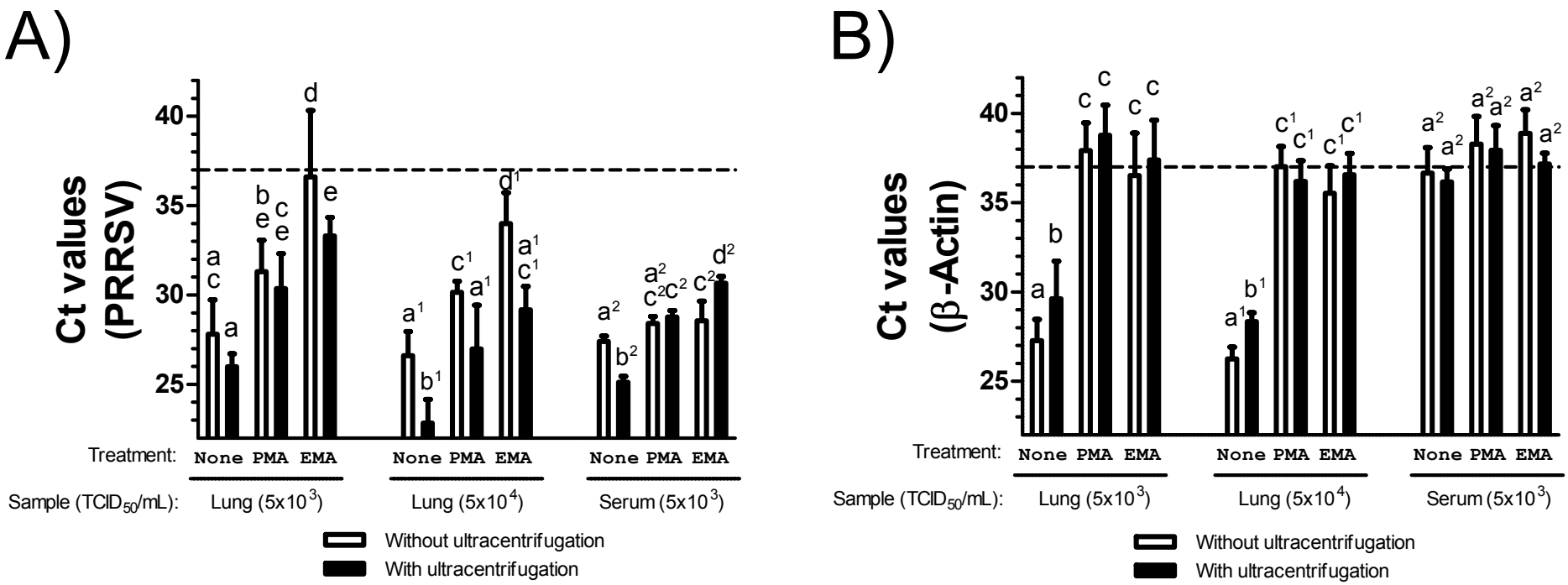
genomic DNA quantification (β-Actin) in lung tissue homogenates spiked with B) PRRSV (5,000 $TCID_{50}$/mL) or with D) PRRSV (50,000 $TCID_{50}$/mL) and F) serum spiked with PRRSV (5,000 $TCID_{50}$/mL). Each trial identification number represents a unique experiment. All trials are the same than those illustrated in Supplemental Figure 2. Results are expressed as Ct. A Ct value of 37 (dashed line) represents the limit of detection of each the PRRSV qPCR assay. See Figure 1 for the combined statistical analyses of all trials. N.D.: not determined.

Supplemental Figure 2

Colour coded individual trial results of PRRSV detection in spiked samples by DNA array following EMA or PMA treatments. DNA array probes relative intensity from A) lung tissue homogenates spiked with PRRSV (5,000 $TCID_{50}$/mL), B) lung tissue homogenates spiked with PRRSV (50,000 $TCID_{50}$/mL) and C) serum samples spiked with PRRSV (5,000 $TCID_{50}$/mL). Dots are relative fluorescence intensity mean values of two identical probes from each trials (each experiment consisting of a duplicate of 34 PRRSV specific probes) and was calculated as followed: [(PFL-BFL)/BFL] where PFL represents a PRRSV probe fluorescence intensity and BFL represents the basal fluorescence level (negative control probe fluorescence). Each trial identification number represents a unique experiment. All trials are the same than those illustrated in Supplemental Figure 1. A probe relative intensity of 1 (dashed line) represents the lowest limit of DNA array positive results. See Figure 2 for combined statistical analyses of all trials.

Figure(s)

Figure 1



Figure(s)

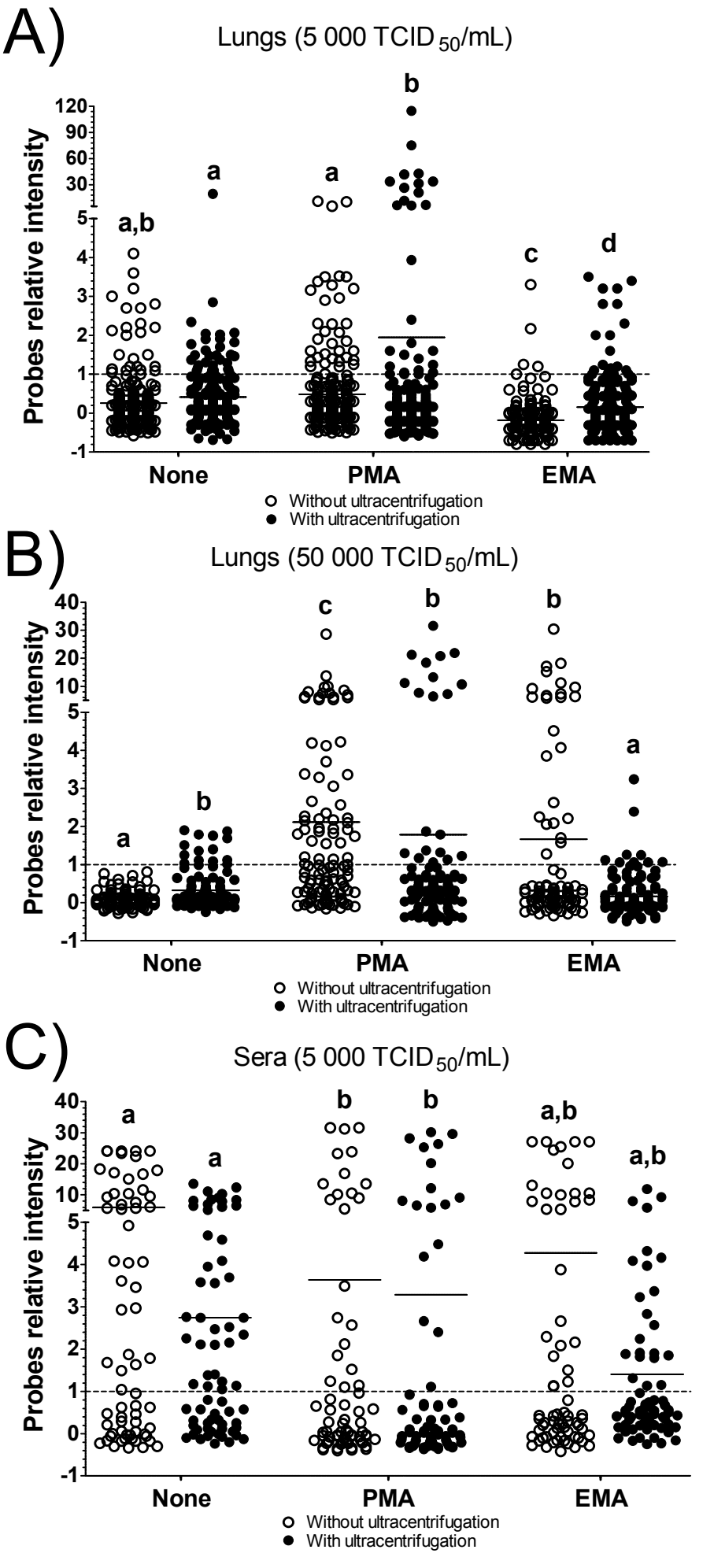


Figure 2

Figure(s)

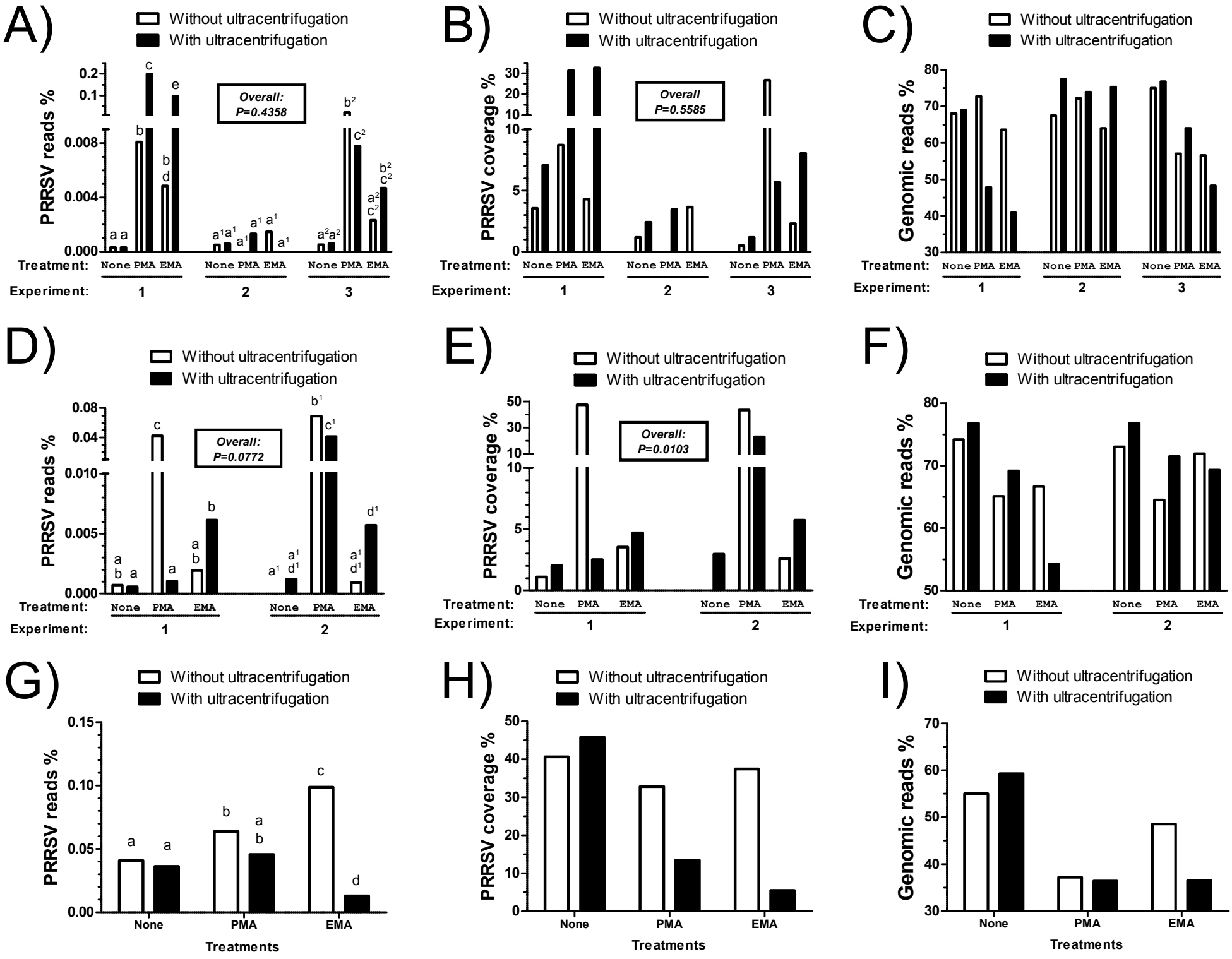
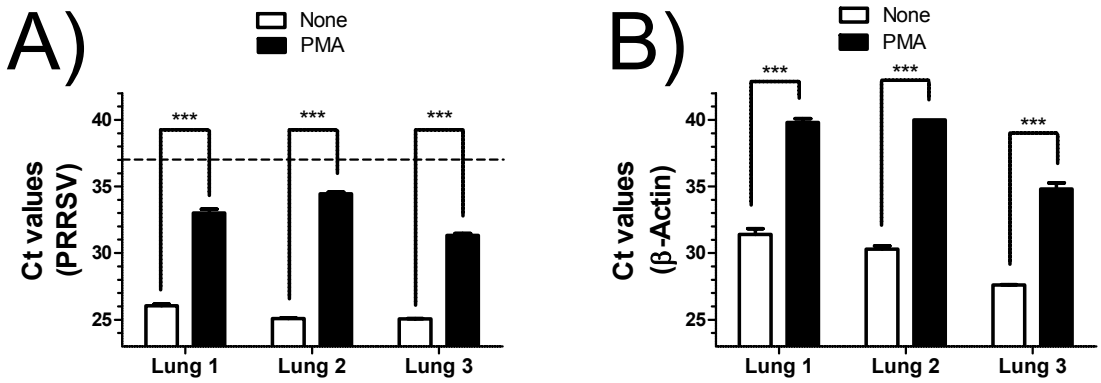
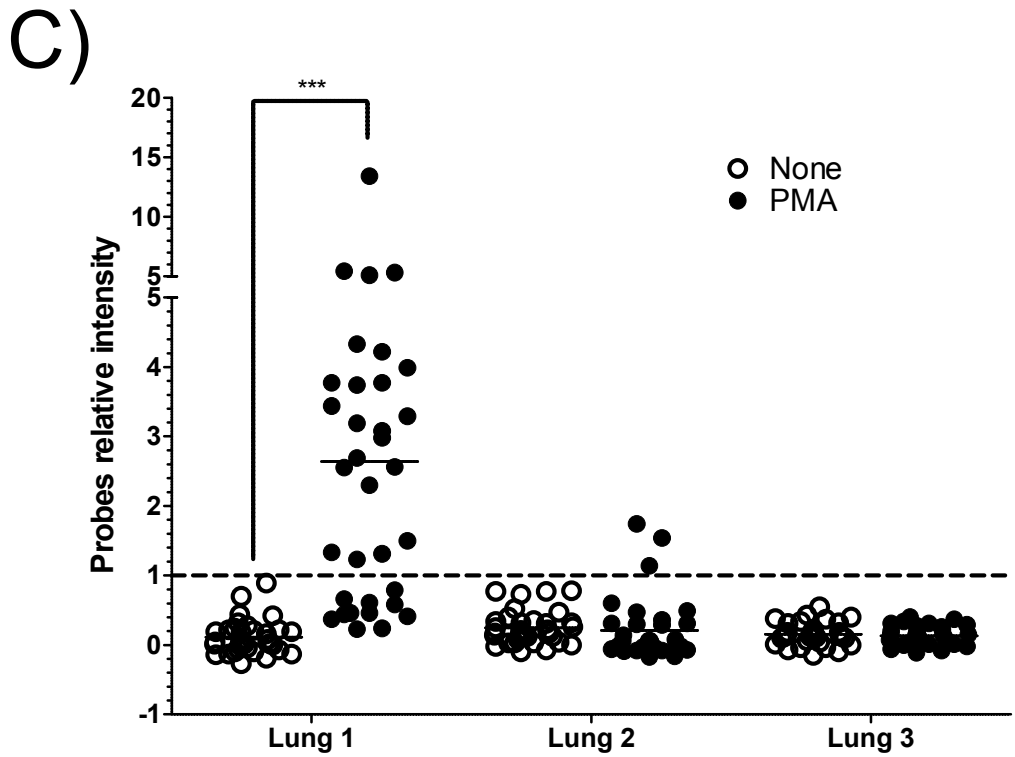


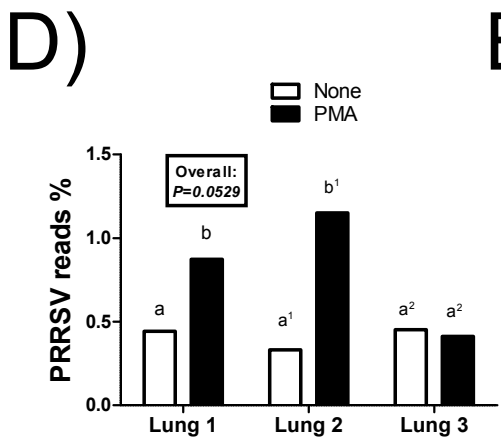
Figure 3

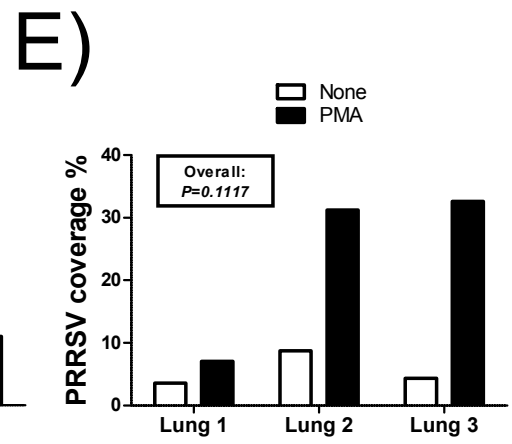
Figure(s)

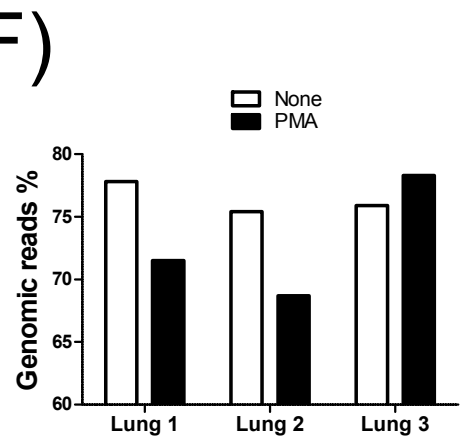
Figure 4





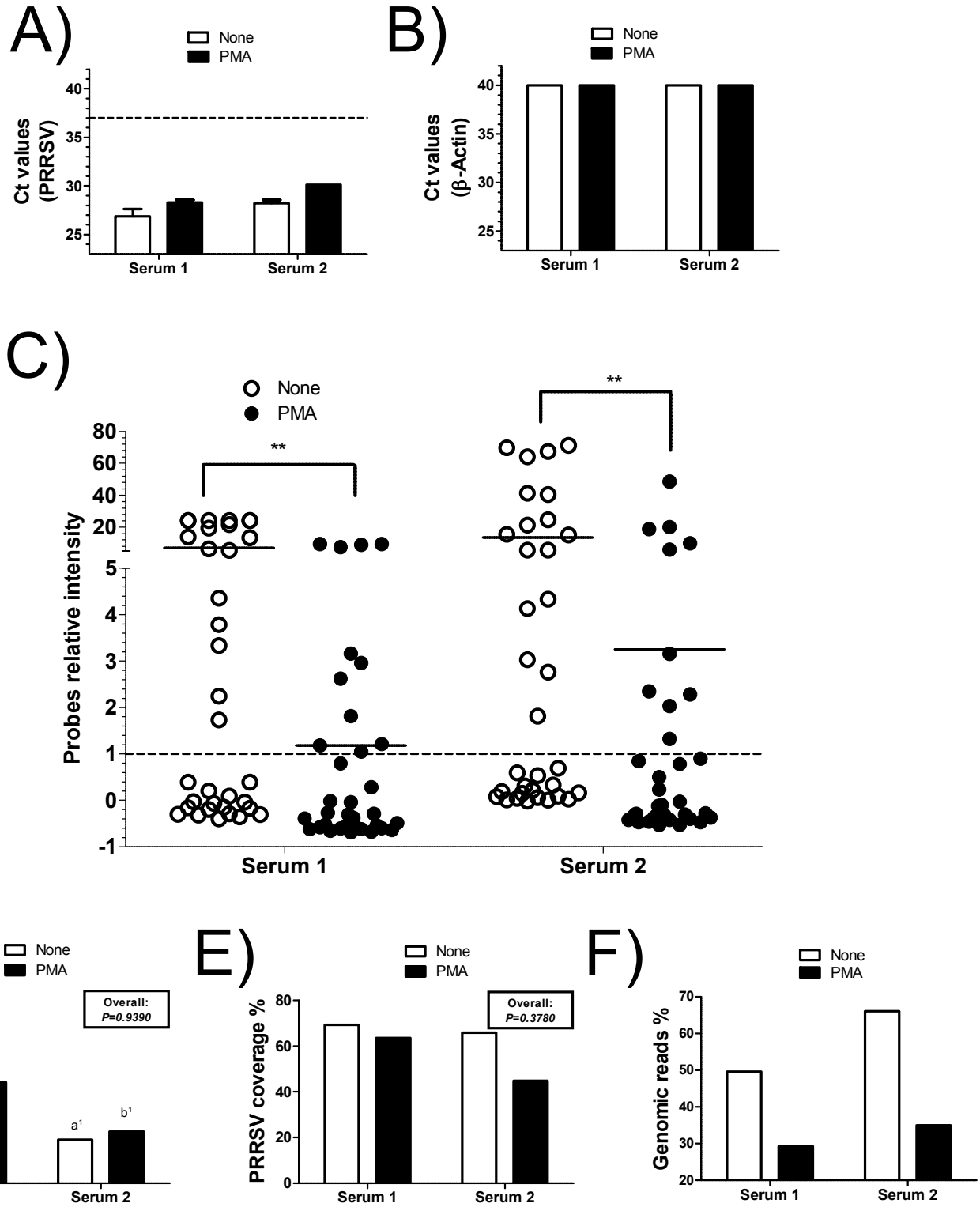






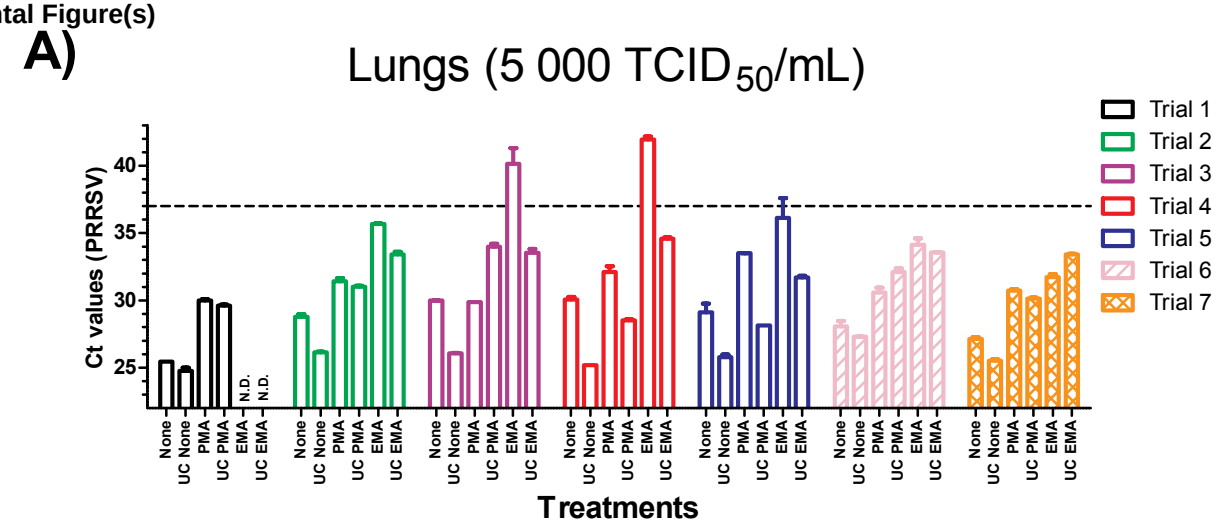
Figure(s)

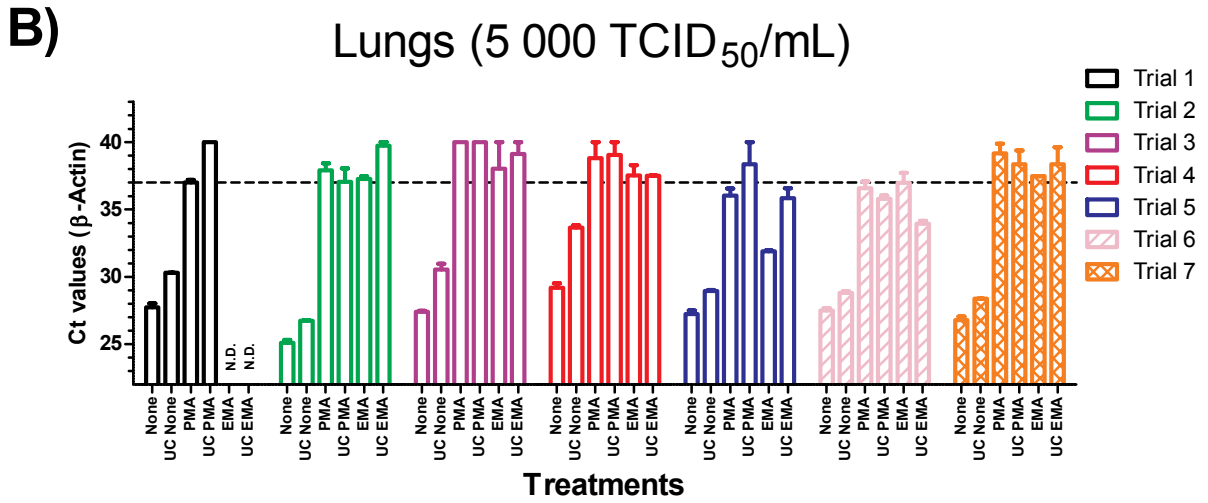
Figure 5

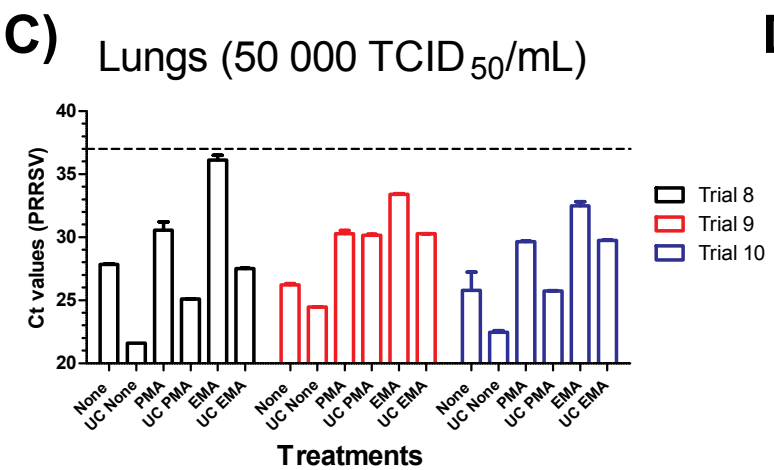


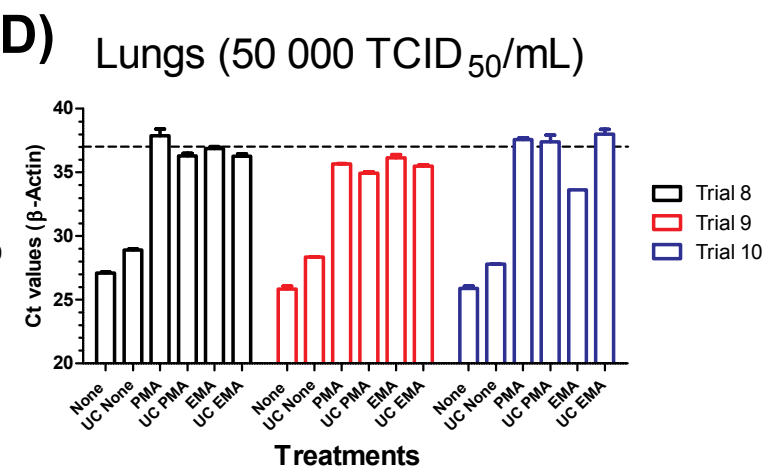
Supplemental Figure(s)

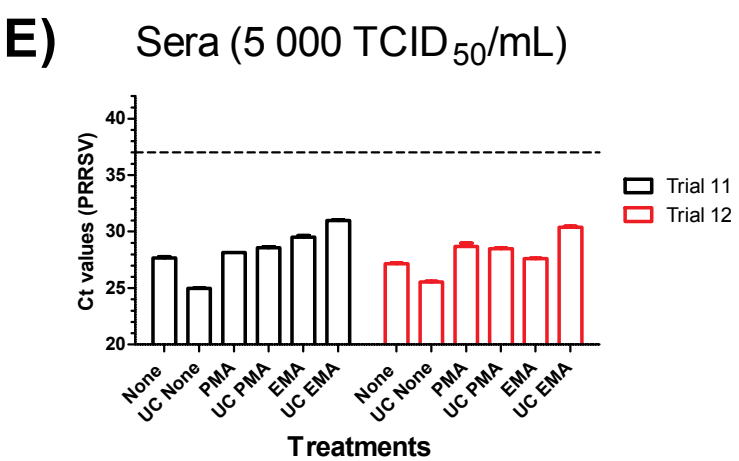
Supplemental Figure 1

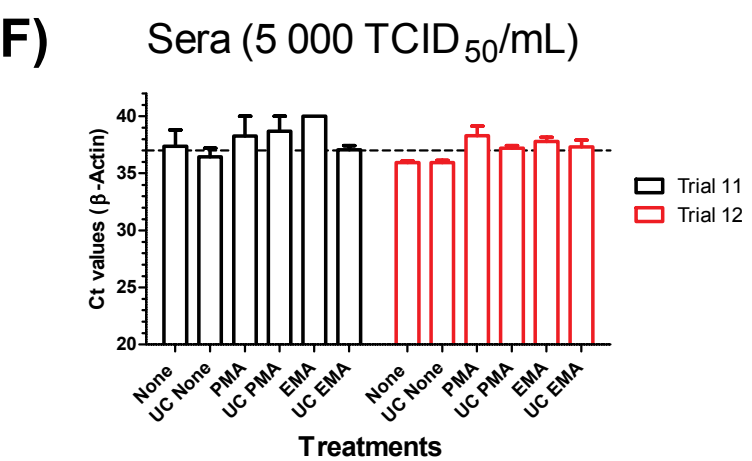












Supplemental Figure(s)

Supplemental Figure 2