Aquavalens Project

"Protecting the health of Europeans by improving methods for the detection of pathogens in drinking water and water used in food preparation."

Grant agreement number: 311846

Deliverable D2.2

Design of variation tolerant primers and probes based on the genomic variation of viruses

Authors: Hongyan Xia, Jonas Blomberg and Jakob Ottoson, Swedish University of Agricultural Sciences

Version 1, 27th January, 2015.

D 2.2. Design of variation tolerant primers and probes based on the genomic variation of viruses.

1. Background

Some of the RNA viruses pose a specific challenge for sequence-based detection, namely a great variation. Many of the water-borne viruses are of this kind. Another problem with detection of water-borne viruses is the high volume of the nucleic acid extracts from the sample, plus the abundance of polymerase inhibitors in them. A procedure which enriches for target nucleic acid is therefore preferred.

1.1. VOCMA-rationale

VOCMA is variation tolerant capture multiplex assay. The VOCMA maximises hybridisation between primers and probe versus target, in the presence of substantial target variation, leading to variation tolerance. VOCMA also provides for target enrichment (Capture) via the avidly binding primer-probes. Due to the low concentration of primer-probes VOCMA can be multiplexed to a greater extent than other PCR systems, because unwanted primer-primer interactions are minimized (table 1).

1.2. VOCMA principle

Variation tolerance is achieved by use of much longer primers and probes than is normally used. This leads to a higher binding strength (high Tm) between target and primers and probe, which allows tolerance for nonmatching nucleotides. A key feature is to include long (>6 nt) perfectly matching stretches, as many as possible. Such stretches can be created by use of a few degenerated nucleotide positions together with a judicious use of universal bases like inosine. Our data show that hybridization then can occur in spite of profound target variation which would normally be considered impossible (Öhrmalm et al. 2012).

A further bioinformatic feature is embodied in the Consort variation analysis program (Blomberg, unpublished). It can decompose target variation into several haplotype sequences, each of which has a smaller variation than the original consensus sequence would have. This is because variants at several sites are often coupled to each other in the form of haplotypes. This allows the number of the desired long uninterrupted perfectly matching stretches to be maximized.

The original VOCMA paper (Öhrmalm et al. 2012) used bead-bound probes. After hybridization of the probe to the amplimers, the results were registered in the Luminex flow meter. A disadvantage of this system is that the tubes have to be opened, creating the possibility of contamination. We therefore wanted to transfer VOCMA to a real-time format in which the readout comes during amplification, without letting he amplimers free after the run. Hence we have adapted the VOCMA system with its advantages to a more sensitive qPCR-platform, however, losing some of the multiplex potential offered by detection in the Luminex.

Feature	Technique	Effect	Outcome
Capture	First specific VT primer-probe coupled to magnetic bead	Removal of inhibitors	less inhibition of PCR
		Enrichment of target	Increases sensitivity and specificity
Variation tolerance	Primer-probes/detection probes designed according NucZip algorithm*	Increased mismatch tolerance	Detection of variable targets e.g RNA virus Detection of whole groups of genes e.g. Antibiotic resistance
	Long target specific part in the 3´end of the primer-probes	Increase the likelihood of long perfectly matching regions (i.e.nucleation sites)	Stronger hybridization Improved variation tolerance
Multiplexity	Usage of both specific and generic primers	Increased specificity	Increased specificity
	Low concentration of target specific primer-probes	Less primer-dimers	High multiplexability, increased specificity and sensitivity
	Few cycles in PCR program at high temp	Limited specific amplification with the long primer-probes	Increased specificity
	High concentration of generic primers	Less primer-dimers	Increased specificity and sensitivity
	High number of cycles in PCR program at low temp	High degree of generic amplification with the short generic primer	Increased specificity and sensitivity
Asymmetric amplification	Unequal conc. of generic primers with biotinylated generic primer in excess	Excess of single stranded biotinylated amplified target	Increased MFI signals
Biotinylation of the second generic primer	The first and second specific VT primer-probes and first generic primer are used in a sequential order to create the ”third copy”		Increased specificity
Single tube amplification	Temperature switch in the PCR-program	Transition from specific to generic primer usage	All primers in same mix Contamination risk minimized
One-step amplification	iScript one-step RT-PCR with both RT Pol and DNA Pol in same reaction	No transfer from Reverse transcription step to the PCR amplification	All enzymes in same mix Contamination risk minimized
Specific detection probe	Probe target region positioned in between primer-probe target regions	The probe introduce a third specificity step No detection of wrongly hybridised and elongated second generic biotinylated primer	Increased specificity Increased specificity

2. Material and methods

2.1. Detection on Luminex

Target sequences or whole genomes were collected in BLASTN (NCBI,NIH) were aligned using Consort, visualising the frequency of variation of nucleotides in each base position. Oligo primer-probe design was evaluated *in silico* using Visual OMP to predict hybridisation efficiency. The panel of seven gastrointestinal viruses was built up stepwise in a recursive loop involving re-evaluation and redesign in order to realise an optimal primer-probe design, with as few cross-reactions as possible and minimised primer competition. The synthetic targets used for sensitivity analyses had a natural primer-probe and detection probe complementary sequence derived from GenBank (Tables 2 and 4). Synthetic ssDNA was serially diluted from $0.5 * 10^7$ to $0.5 * 10^0$ targets/µL in DEPC water (Ambion). Patient samples were provided anonymously according to Uppsala University hospital rules and extracted using the EasyMag system (bio-Merieux) according to the manufacturer's instructions.

The primers and probes used for detection using Luminex are included in table 2 and the synthetic targets in table 3. Two µL of synthetic ssDNA target was added to a 23- µL one-step reverse trancriptase PCR (RT-PCR) master mix containing a final concentration of 1 x iScript buffer (BioRad), 0.5 µL iScript reverse transcriptase, 300 nM generic first primer, 500 nM biotinylated generic second primer and 50 nM each of the 14 first and second specific VT primers-probes (table 2) to create the 7-plex VOCMA mixture. Amplification was carried out according to Öhrmalm et al. (2012) with PCR conditions as follow: 50°C 20 min, 95°C 5 min followed by 10 cycles of 95°C for 15 s, ramping at 0,1°C/s from 75°C to 65°C, 65°C for 1 min, followed by 40 cycles of 95°C for 15 s, 52°C 30 s, 60°C 30 s, followed by 60°C for 5 min, 95°C for 1 min and 4°C until the hybridisation step.

Hybridisation: five microliters of biotin-labelled VOCMA amplified target was mixed with hybridisation buffer and probe-coupled xMAP bead. The mixture was heated to 95°C for 2 min followed by hybridisation at 50°C for 30 min with shaking at 600 rpm on a Thermostar microplate incubator (BMG LabTech). After centrifugation, 38 µL buffer was mixed with 2 µL streptavidin-R-phycoerythrin (Qiagen) and incubated for another 15 min at 50°C before analysis on the Luminex 200 flow meter (Luminex corporation). The quantity of the target was measured by median flow intensity (MFI).

Table 2 Oligonucleotides used in the Gastro VOCMA

Target	Primer/Probe	Modification	Sequence 5´to 3´
Generic first			TTGGATAAGTGGGATA
Generic second		Biotin	AAGATATCGTAAGGAT
Norovirus GG1	Second		AAGATATCGTAAGGATAAGACAGGGGGICTTGAAATGTATGTICCAGGITGGCAGGCCATGTTCCGITGGATGCG
	First	AminoC12	TTGGATAAGTGGGATACCGGCGCCAITIGCGCCATCIRIGITTGGIGTIGCGTCCTTAGACGCCAT
	Detection	AminoC12	CATCATCATTTACIAATTCGGGCAGIAGATTGCGATCTCCTGTCCACA
Norovirus GG2	Second		AAGATATCGTAAGGATAGATACAAGTACCACTATGATGCIGAITACTCTCGGTGGGA
	First	AminoC12	TTGGATAAGTGGGATATCAGAGAGCGCACAGAGAGTGAGAAGCCAGTGGGCGATGGAGTTCCATTGGGAGGTGCAGGGC ACCCCAGA
	Detection	AminoC12	CCCTCGTTGATTGATATTGTGAAGTCACCCACATCCACCACGCTAGGAGAAAGAAGGTCTTCTGC
Adenovirus	Second		AAGATATCGTAAGGATTTTGCCATGGAIATIAATITGGCGGCIA+AII+TIT+GGIGIAGCTTCTTITAYTCMAATGTAGC
	First	AminoC12	TTGGATAAGTGGGATAATGTTTACGTAGGTATCIAGGGCGCTIGGIACIGCIACCCGACCGTTCATGTAGGCATA
	Detection	AminoC12	TTGTTKTCGGGIAIIGTAATGTTITCIGGIGTIATITTITAIGAGTCAGGCAA
Rotavirus A	Second		AAGATATCGTAAGGATTATTCITTGTCAAAAACTCTTAAAGATGCTAGGG+AIA+ARATTGTYG+AAGG
	First	AminoC12	TTGGATAAGTGGGATACCAAAATCAAAAGTCCAA+TTTCTAATIGGTAGAT+TGCCAA+TTCCTCCAGT
	Detection	AminoC12	CCIATICCTCCAGTYTGRAAHTCATTYCCATTCATRGTAAYTATCATTTGRTTAAAITGITGAATIA
Sapovirus	Second		AAGATATCGTAAGGATGGTACAGTACITGACCIAGIGGGTCICACTGAAGGIACCCICAAAWTAGTGTTTGAGATGGA
	First	AminoC12	TTGGATAAGTGGGATAGCAAAGCAGTTGCGTATTGCITCAGGGACATTKGATTGGATTGCACCIGTIGCAACAGCCA
	Detection	AminoC12	TCCAGGCGYTGIGCGGSCCCAITGGGTTGITCIGGATTAGCAACAACAACGTGGGATGTGGTCGGICC
Enterovirus	Second		AAGATATCGTAAGGATGGTGGGAAGAGGCTATTGAGCTACITGIGAGTCCTCCGGCCCCTGAATGCGGCTAATCC
	First	AminoC12	TTGGATAAGTGGGATACACCGGATGGCCAATCCAATIGCTITATGGTAACAATCTITGIATTGTCACCATAAGCAGCCA

	Detection	AminoC12	ACGGACACCCAAAGTAGTCGGTTCCGCYGCAGAITTGCICGTTACGACAIGCIACICACTGGTTTGTG
Astrovirus	Second		AAGATATCGTAAGGATGAATCICTCCATGGGIAICTCCTITGTTATCAGTTGC+TYK+CTG+CITTTATGG
	First	AminoC12	TTGGATAAGTGGGATACTAGCCATTGCITTTTTTTGGTCCTCCCCTCCAGATGGRATGGAGTTGCTCTTC
	Detection	AminoC12	CCAGAITCACGAAGCTGCTTIGCAGTCCKIGAIAGGCAGTGITCIAIGTA
	Seq_Fw_GG2		TAGATACAAGTACCACTATGATGC
	Seq_Rev_GG2		CAGAGAGTGAGAAGCCAGTG

Table 3. Sequences for synthetic targets for VOCMA on Luminex platform

Target	Origin	Sequence 5´to 3´ of synthetic target
Norovirus GG1	M87661	AAGACTGGTGGATTGGAAATGTATGTCCCAGGATGGCAGGCCATGTTCCGCTGGATGCGCTTCCATGACCTCGGATTGTGGACAGGAGATCGCGATCTTCTGCCCGAATTCGTA AATGATGATGGCGTCTAAGGACGCTACATCAAGCGTGGATGGCGCTAGTGGCGCTGG
Norovirus GG2	Consensus	AGATACAAGTACCACTATGATGCGGAGTACTCUCGGTGGGAGCAGAAGACCTTCTTUCTCCTAGCGTGGTGGATGTGGGTGACTUCACAATATCAATCAACGAGGGTCTGGGG TGCCCTGCACCTCCCAAUGGAACTCCATCGCCCACTGGCTTCTCACTCTCTGTGCGCTCTCTGA
Adenovirus	Ad 41	TTTGCCATGGAAATCAATTTGGCGGCCAATCTCTGGCGCAGCTTCTTATACTCCAATGTAGCTTTGTACTTGCCTGACTCATACAAGATTACGCCAGACAACATTACACTGCCCGA AAACAAGAACACCTATGCCTACATGAACGGTCGGGTGGCGGTTCCTAGCGCCCTCGATACCTACGTAAACAT
Rotavirus A	K02086	TACTCACTGTCAAAAACTCTTAAAGATGCTAGGGACAAAATTGTTGAAGGTACATTATATTCTAATGTTAGCGATCTTATTCAGCAATTCAATCAAATGATAGTAACTATGAATGG AAATGATTTTCAGACTGGAGGAATTGGTAATTTACCTGTTAGAAATTGGACTTTCGATTTTGG
Sapovirus	NC_006269	GGTACAGTACCTGACCCAGTGGGTCACACTGAAGGAACCCACAAAATAGTGTTTGAGATGGAGAGGTCCGACCACATCCCACGTTGTTGTTGCTAATCCGGAGCAACCCAATGG GGCCGCACAGCGCCTGGATTGGCTGTTGCCACTGGTGCAATCCAATCCAATGTCCCTGAGGCAATACGCAACTGCTTTGC
Enterovirus	Poliovirus NC_002058	GGTGTGAAGAGCCTATTGAGCTACATAAGAATCCTCCGGCCCCTGAATGCGGCTAATCCCACAAACCAGTGATTGGCCTGTCGTAACGCGCAAGTCCGTGGCGGAACCGACTAC TTTGGGTGTCCGTTGGCTGCTTATGGTGACAATCACAGATTGTTATCATAAAGCGAATTGGATTGGCCATCCGGTG
Astrovirus	FJ755402	GAATCACTCCATGGGGAGCTCCTATGCTATCAGTTGCTTGCTGCGTTCATGGCAGAAGATCACCCTTTTAAGGTGTATGTAGAACACTGCCTATCACGGACTGCGAAGCAGCTTC GTGACTCTGGCCTTCCGGCCAGACTCACAGAAGAGCAACTCCATCGCATTTGGAGGGGAGGACCAAAGAAGTGTGATGGCTAG

2.2. Detection using real-time PCR

The design of oligo primers –probes used in the qPCR was performed as described above and sequences shown in Table 4. The synthetic target (table 4) was used for optimization of the real-PCR condition and tested for sensitivity of the assays. A 10-fold serial dilution of 10^7 to 10^0 copies of the synthetic target was used in the sensitivity test. A total of seventy-seven clinical samples from routine diagnostic group at academic hospital in Uppsala were used in this study. Among those patient samples, forty-six samples were norovirus genotype II, eleven samples were genotype I and twenty samples were genotype I and II negative (determined by routine diagnostic group). The samples were sequenced if they did not agree with the result from routine diagnostic lab.

The real-time PCR was performed on a CFX 96 touch real-time PCR detection system (Bio-Rad) using Path-ID™ Multiplex One-Step RT-PCR Kit. The concentrations of the norovirus genotype II specific primer, generic primers and probes were 40 nM, 300 nM and 100 nM respectively. Real-time PCR condition was: 45°C 10 min, 95°C 10 min followed by 25 cycles of 95°C 15 sec, 61°C 45 sec and 40 cycles of 95°C 15 sec, 52°C 45 sec.

Table 4. Primers, Probes and synthetic target for norovirus genotype II for detection on a real time PCR platform

Primers/Probes	Sequences (5’ to 3’)
Generic first	TTGGATAAGTGGGATA
Generic second	AAGATATCGTAAGGAT
NVFw	AAGATATCGTAAGGATAATYAGYAARYTRGTYATTGCAGAGYTDAARGARGGTGGYATGGATTTTTAC
NVRev	TTGGATAAGTGGGATAAYTCGACGCCATCTTCATTCACAAARCTGGGAGCCAGATTGCGATCGCCCTC
NVprobeTint	FAM/CCACGTGCTCAGRTCYGAGAA[T(BHQ-1)]CTCATCCAYCTRAACATYGGCTCTTGYCTGGGCAC/spacer C3
NVprobe1	FAM/CCACGTGCTCAGRTCYGAGA/iBHQ1/ATCTCATCCAYCTRAACATYGGCTCTTGYCTGGGCAC/3Spacer C3
NVprobtry	FAM/CCACGTGCT/zen/CAGRTCYGAGAATCTCATCC/zen/AYCTRAACATYGGCTCTTGYCTGGGCAC/3lowa black RQ-Sp
NVprobtryA	FAM/CCACGTGCT/zen/CAGRTCYGAGAATCTCATCCAYCTRAACATYGGCTCTTGYCTGGGCAC/3lowa black RQ-Sp
Synthetic target(4929-5102)	AATTAGCAAGCTAGTCATTGCAGAACTGAAGGAAGGTGGCATGGATTTTTACGTGCCCAGACAAGAGCCAATGTTCAGATGGATGAGATTCTCAGATCT GAGCACGTGGGAGGGCGATCGCAATCTGGCTCCCAGCTTTGTGAATGAAGATGGCGTCGAAT

3. Results

3.1. Detection on Luminex

The sensitivity and specificity of the assay was determined using synthetic target. The VOCMA assay on Luminex format showed high specificity with no cross-hybridization (figure 1).

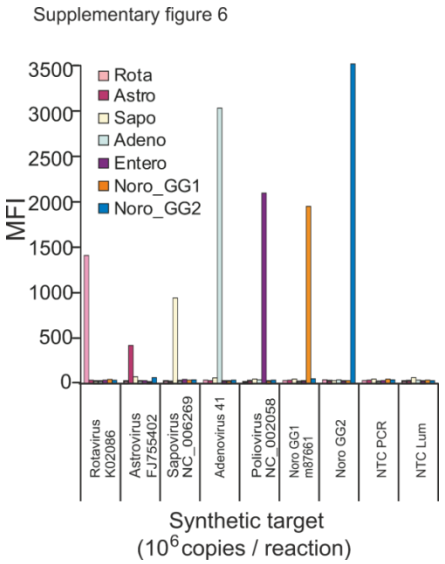

Figure 1. A 7-plex VOCMA detecting different synthetic ssDNA targets (10^6 copies / reaction). The synthetic targets (Supplementary Table 3) have different degrees of mismatch towards the primer-probes and the detecting probes. The panel of seven detection probes is indicated by the different colors in the figure. The X-axis displays the different targets, with seven differently colored bars for the indicated detection probes. The Y-axis displays the Median Fluorescent Intensity, MFI. Signals are specific, with a very low background.

The analytic sensitivity of the assay for rotavirus and adenovirus is 1-10 copies/reaction, 100 copies/reaction for astrovirus and poliovirus, 100-1000 copies/reaction for norovirus genogroup II, and 10000 copies/reaction for norovirus genogroup I and sapovirus (Figure 2).

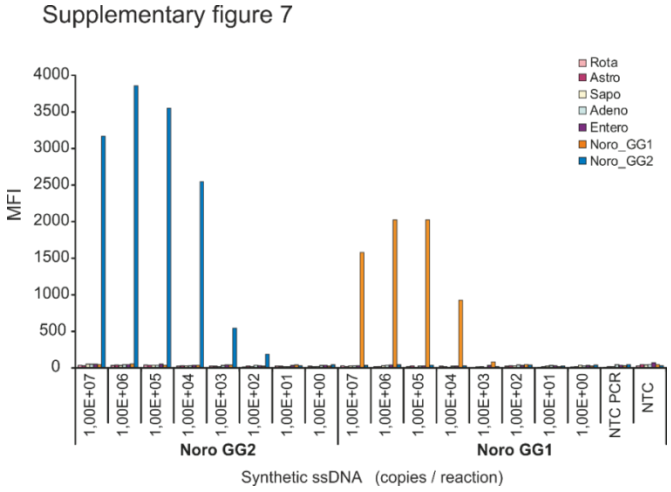

Figure 2. A 7-plex VOCMA detecting synthetic ssDNA targets of norovirus genotype 1 (orange bars) and genotype 2 (dark blue bars) in titers ranging from 10^7 to 10^0 copies per sample. The panel of the seven detection probes is indicated by the different colors in the figure. The X-axis displays the different targets, with seven differently colored bars for the indicated detection probes. The Y-axis displays the Median fluorescent intensity, MFI.

3.2. Detection with qPCR

The sensitivity of the assays using differently labelled probes was tested using synthetic target according to table 4. One to 10 copies/µL could be detected with the VOCMA system on the real time PCR system (Table 5).

Table 5. Analytical sensitivity for detection of norovirus genogroup II synthetic target using different probes. Results from seven replicates at low copy numbers each

Probe	Analytic sensitivity (copies/µL)
NVprobetry	1-10
NVprobetryA	1-10
NOprobe1	10-100
NVprobetin	1-10

The specificity of the assays was tested using clinical samples (table 6). All the eleven norovirus genogroup 1 samples determined by the routine diagnostic lab were genotype II negative. It demonstrated that VOCMA has a high specificity. Among the forty-six norovirus genogroup II, we detected forty-four samples as norovirus genogroup II positive by all four probes. Ct values were very low (from 2 – 15 typically) and in some cases there was even amplification in the first portion of the single-tube nested PCR of VOCMA (data not shown).

In two of the negative samples the PCR product could be readily seen on a gel. However, direct sequencing of PCR product showed several peaks at the same position, which may indicate that a population of norovirus genotype II existed in those two samples. Subcloning of the PCR product to vector and sequencing of the individual colons will be performed to clarify the sequence (sequence data has not arrived yet).

For the twenty genogroup I and II negative samples, our assays detected one sample as genogroup II positive by probetry and probetryA. The remaining samples were all detected as genogroup II negative by all probes.

Table 6. Real-time PCR result from clinical samples

Samples	Norovirus II	Norovirus I	Norovirus II and I negative
NVprobeTint (Number of Norovirus II positive samples/Number of tested samples)	44/46	0/11	0/13
NVprobe1	44/46	0/11	0/13
NVprobtry	44/46	0/11	0/13
NVprobtryA	44/46	0/11	0/13

4. Discussion

As explained above the design of VOCMA primer-probes maximises hybridisation between primers and probe versus target, in the presence of substantial target variation, leading to variation tolerance. This fact also enables this PCR system to address more variable regions of the genomes and, hence, target genes involved in pathogenicity with minimised risk of false negative results. Further, due to the low concentration of primer-probes VOCMA can be multiplexed to a greater extent than other PCR systems, because unwanted primer interactions are minimized. Another advantage, not discussed above, is that the products generated are typically longer than for many other PCR systems. The products may therefore have enough information potentially of use for typing during outbreak situations without having to reanalyse your sample using another primer pair.

The original VOCMA was designed for detection using Luminex. This is advantageous because it provides a highly multiplex potential. One problem with this system, however, is the potential for contamination to occur when the PCR tubes were opened before the hybridization step. Due to the use of generic primers in the final amplification, such a contamination could ruin all VOCMA components which used the same generic primers. Despite this threat, it has not happen during two years of developmental work. The sensitivity could also be questioned, even though for some targets down to ten copies could be detected, for others almost a thousand copies in the original sample are needed. Because of these reasons, VOCMA has been developed to a real time PCR platform within Aquavalens.

VOCMA in the real time-PCR format proved to be highly sensitive, and still multiplexable. The low concentration of the primary (virus specific) primers minimizes primer-primer interactions, a major cause of primer-dimers which diminishes sensitivity. According to our results, VOCMA in the real time format is also quantitative. Another novelty is that the VOCMA primers now can be biotinylated. This was not possible before, because biotin was used later in the hybridisation procedure. With biotinylated primers, the capture process can be made more efficient, letting biotinylated primers catch the target nucleic acids, then capturing them on streptavidin-coated magnetic beads. Capture efficiencies of >50% are feasible as an alternative secondary concentration procedure that also is efficient in removing inhibiting substances that do not bind as efficiently to magnetic beads as the do to *e.g.* silica.

Based on the results from the experiments reported herein we expect to be able to use the previous primers (table 2) for the VOCMA real time PCR and with only minor modifications of the probes (Xia et al. 2015). Therefore we will be able to offer multiplex panels for the detection of enteric and hepatitis viruses respectively within a couple of months to be ready for the validation taking place within work package 9 of Aquavalens (www.aquavalens.org).

5. References

Öhrmalm C, Eriksson R, Jobs M, Simonsson M, Stromme M, Bondesson K, Herrman B, Melhus A, Blomberg J (2012). Variation-tolerant capture and multiplex detection of nucleic acids: application to detection of microbes. *J Clin Microbiol* 50: 3208-3215.

Xia H, Öhrmalm C, Ottoson J, Blomberg J (2015). Internal quenchers allow long variation tolerant TaqMan probes for improved detection of molecular targets in problematic matrixes (manuscript in preparation).