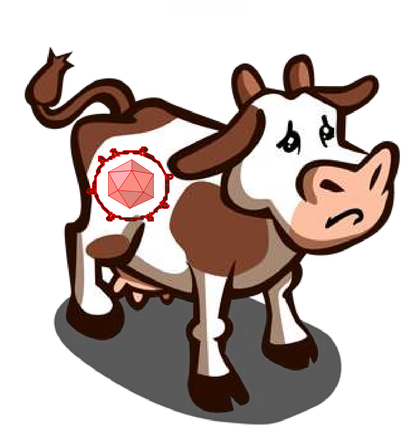
PCR-sequencing of the BVDV genome reflects quasi-species variation defined by next-generation sequencing

George Russell[1], Claudia Bachofen[1,*], Kim Willoughby[1], Ruth Zadoks[2]
[1] Vaccines and [2]Disease Control, Moredun Research Institute, Edinburgh, UK
** Current address: Institute of Virology, Vetsuisse Faculty, University of Zürich, Switzerland*
E-mail: George.Russell@moredun.ac.uk

Introduction

Analysis of RNA virus genomes can be complicated by their error-prone replication, which produces a population of closely related viral sequences in infected individuals. While next-generation sequencing (NGS) allows the quasi-species nature of RNA virus infection to be elucidated, this can be expensive in relation to genome size or require large amounts of material. To study Bovine Viral Diarrhoea Virus (BVDV) sequence variation, we have developed methods to amplify segments of the BVDV genome from very small diagnostic samples. Here we describe a comparison of NGS sequencing with virus-specific RT-PCR and Sanger sequencing, using BVDV from the serum of a persistently-infected animal and the same virus after passage in culture.
NGS analysis of the virus BVDV-MRI103 showed that 276 positions within the assembled BVDV genome sequences were polymorphic in one or both genome assemblies, while the consensus sequence of the cultured virus differed at only 4 positions from the virus in serum.
To develop a generic approach to sequencing BVDV genomes from virus-positive diagnostic samples, primers were designed to amplify 24 PCR products that spanned the entire BVDV genome with a high degree of overlap. While the consensus sequences produced by the two approaches were essentially identical, the chromatogram traces of the PCR-derived sequences also contained evidence of many of the SNP detected by next generation sequencing.

METHODS:



BVDV Persistently Infected animal

Serum sample MRI103

Viral RNA extraction; cDNA synthesis; library preparation; NGS analysis; SNP analysis

(P3) 3 passages in culture, MOI 0.1

(P5) 5 passages in culture, MOI 0.1

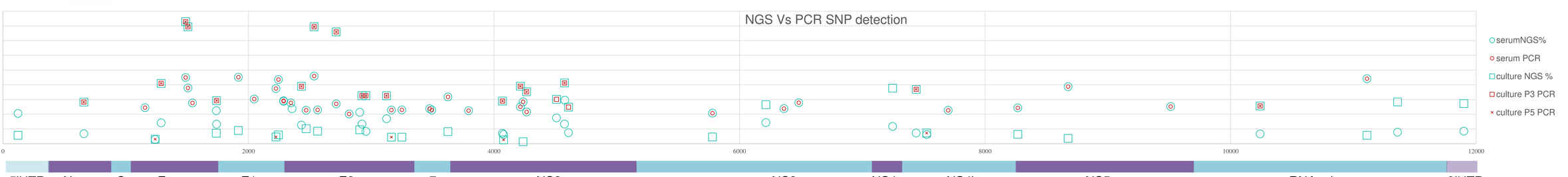
Viral RNA extraction; RT-PCR overlapping fragments; Sanger sequencing; genome assembly

Comparison of genome sequences:

NGS serum	Sanger serum
NGS P3	Sanger P3
	Sanger P5

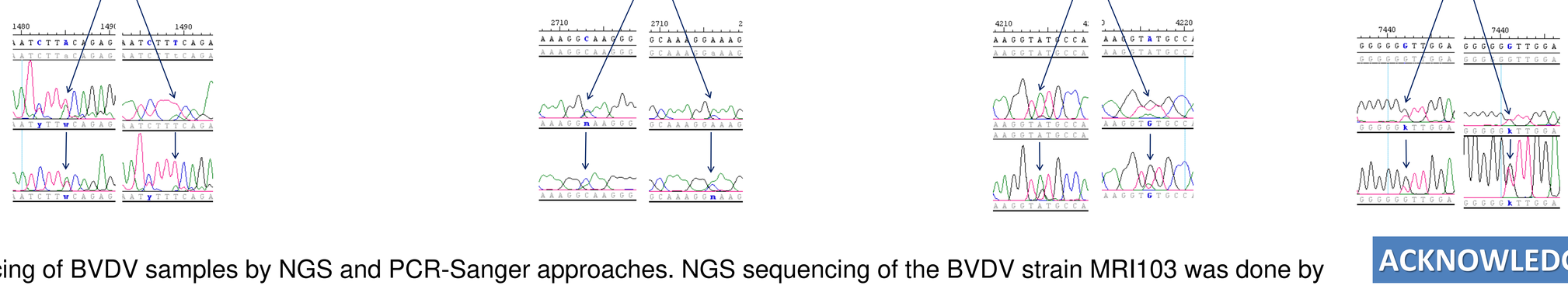
RESULTS:

A



B

Position	123	659	1158	1240	1289	1488	1507	1544	1738	1740	1918	2047	2224	2244	2286	2287	2346	2355	2432	2470	2535	2563	2715	2820	2907	2925	2957	3125	3165	3250	3476	3494	3625	3793	4069	4078	4215	4237	4266	4509	4575	4577	4607	5780	6215	6362	6483	7247	7439	7525	7700	8267	8677	9512	10241	11111	11360	11900
Consensus	G	T	G	C	A	A	G	T	G	A	A	C	G	A	G	A	A	A	G	A	G	T	C	C	A	T	C	G	G	C	G	G	A	G	A	C	G	A	T	G	A	C	A	A	C	G	A	T	G	A	C	A	T	G	A	T	C	A	T
serum NGS variant	A	G	A	T	C	T	A	C	A	G	G	T	A	G	T	C	G	G	A	G	A	C	A	T	G	A	A	A	A	T	T	A	G	G	A	G	A	T	A	T	A	C	A	T	T	A	G	A	T	G	A	C	A	T	G	C	T	G	C
serum %	21%	7%	24%	3%	14%	45%	38%	28%	22%	13%	45%	30%	37%	44%	29%	29%	28%	24%	13%	23%	46%	23%	27%	20%	21%	13%	8%	17%	23%	23%	24%	23%	32%	22%	7%	6%	25%	28%	22%	18%	13%	30%	7%	21%	14%	24%	28%	12%	7%	6%	23%	24%	39%	25%	7%	44%	8%	9%
P3 NGS variant	T	G		T	C	T	A		A	G	G		A	G					A	G	A	C	A			G	A	A	A	A	T			G		A	T	G	T	T							A	T	G		T	G		T	G	T	G	C
P3 %	6%	28%		3%	41%	83%	79%		7%	29%	9%		5%	6%					39%	10%	79%	9%	76%		9%	33%	33%	33%	4%	4%			8%		29%	3%	39%	1%	35%	30%	41%		25%	5%	26%			38%	37%	7%		6%	4%		26%	6%	28%	27%
serum PCR			G/a			T/A	A/G	T/c			A/g	C/t	G/a	A/g	G/t	A/c	A/g			A/g	A/G	T/c	C/a	C/t						G/a	C/t	G/t	G/a	A/g	A/g			A/g	T/c	A/g					A/t		G/a	A/g				C/a	C/t	A/g	T/c		C/t			
P3 PCR		T/g			A/c	T/a	A/g			A/g									A/G		A/g		A/c			T/A	C/a	G/A							G/A		G/A		A/G	T/c	A/g		A/g					G/t						C/t			
P5 PCR		T		C/t	A/c	T/a	A/g			A/t			G/a						A/G		A/G		A/c			A/T	C/A	A/g	G/a						G/a	C/t	G/a		G/A		G/A							G/t	A/g					C/t			



Sequencing of BVDV samples by NGS and PCR-Sanger approaches. NGS sequencing of the BVDV strain MRI103 was done by TGAC by Illumina short-read technology on unamplified cDNA produced from viral RNA from serum or from passage 3 (P3) of culture. PCR-sequencing used RNA purified from the same material (serum and P3) and from P5 cultured virus, amplifying 24 overlapping fragments using primers based on the genomic sequence. (A) Variant sites in the NGS data were scored by read-count as non-consensus base calls represented in >5% of reads, while PCR sequence variants were those with clear trace data for multiple base calls in at least two reads at the identified position. Four examples of trace data representing polymorphic differences between the serum virus and P5 cultured virus are shown.

SUMMARY:

- **Essentially complete BVDV genome sequences were obtained from overlapping PCR products**
- **Sanger sequence data identified sites of quasi-species polymorphism within the BVDV genome**
- **Virus consensus sequence changed even after 3 culture passages**
- **P5 culture fixed some SNPs that were vey low frequency in the serum virus**

ACKNOWLEDGEMENTS:

TGAC The Genome Analysis Centre™
BBSRC Greater Norwich Development Partnership
FNSNF Swiss National Science Foundation
Moredun Research Institute