Datasets and Figures in 'Construction of three foot-and-mouth disease virus peptide phage display libraries as a tool for the identification of important epitopes' NPB Sekgobela Naledi.

1. Peptide phage library preparation and construction

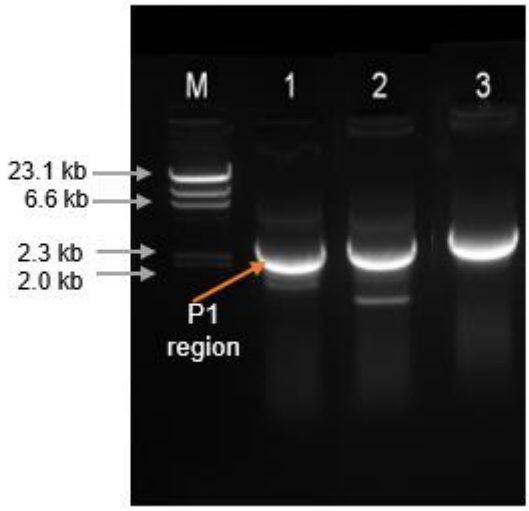


Figure 1.1: A 1% (w/v) agarose gel showing the PCR amplification products of the FMDV SAT1/KNP/196/91, SAT2/KNP/19/89 and SAT3/SAR/1/06 virus strain's, P1 regions that were used in peptide phage display library construction. Lane M in the figure is the Lambda DNA/HindIII marker (Promega) in kilobases (kb). Lanes labelled 1, 2 and 3 represent SAT1, SAT2 and SAT3, respectively. The FMD SAT viruses P1 regions are ~2.2 kilo-bases (kb).

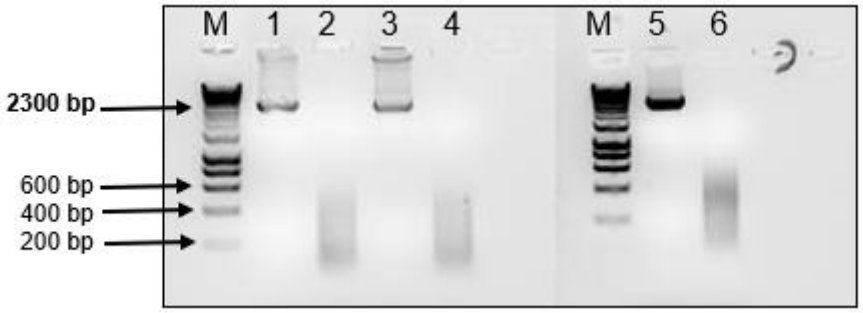


Figure 1.2: A 2.5% (w/v) agarose gel subjected to electrophoresis showing the fragmentation of the FMDV SAT serotypes P1 DNA products. Sheared DNA fragments of 50-200 bp range were observed. Lane M is a 1Kb marker (Bioline). Lanes 1, 3 and 5 represent the unsheared SAT1, SAT2 and SAT3 P1 DNA whilst lanes 2, 4 and 6 represent the sheared SAT1, SAT2 and SAT3 P1 DNA.

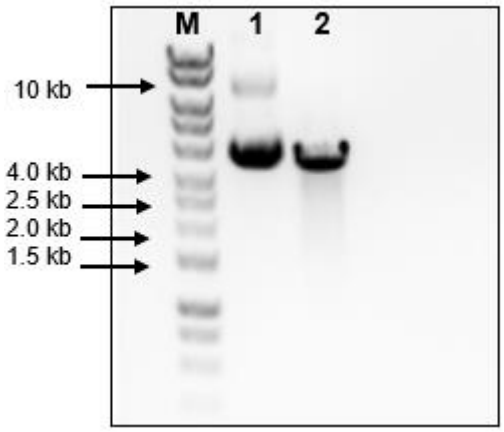


Figure 1.3: A 0.7% agarose gel subjected to electrophoresis showing the vector, pCVEP1585042, restriction enzyme digested with PmeI. Lane M represents the 1 kilobase (kb) marker (Bioline), while lanes 1 and 2 represents the PmeI digested and undigested vector, respectively.

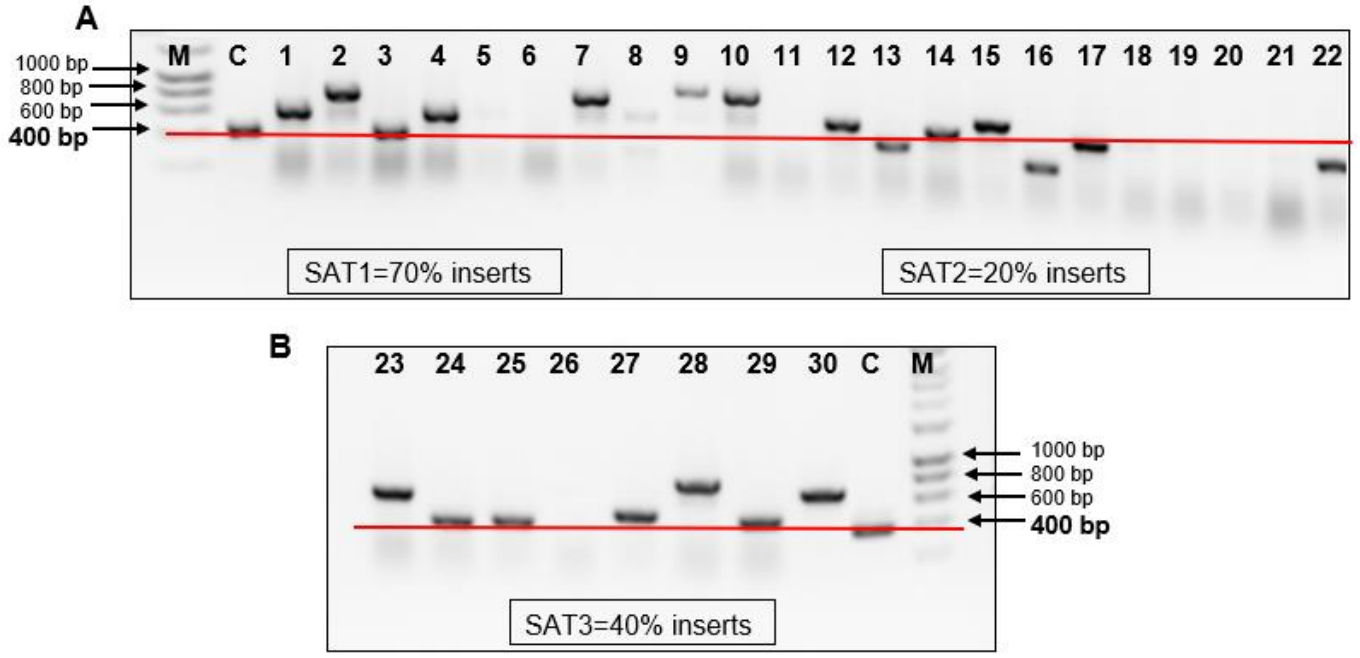


Figure 1.4: A representation of the colony PCR results where the products were separated on a 1% (w/v) agarose gel by electrophoresis (A and B). The insert of interest is present in clones larger than 400 bp in size (red line indicates the 400 bp mark), which included 10 PCR products for each serotype (all results not shown). The control vector is represented by a C (approximately 400 bp), while lane M is the 1kb hyperladder marker (Bioline) in bp. The numbers on the gels represent the lanes, (**A**) lanes 1-10 are SAT1 clones, 11-20 are SAT2 clones; 21-22 are SAT3 clones and (**B**) lanes 23-30 are SAT3 clones.

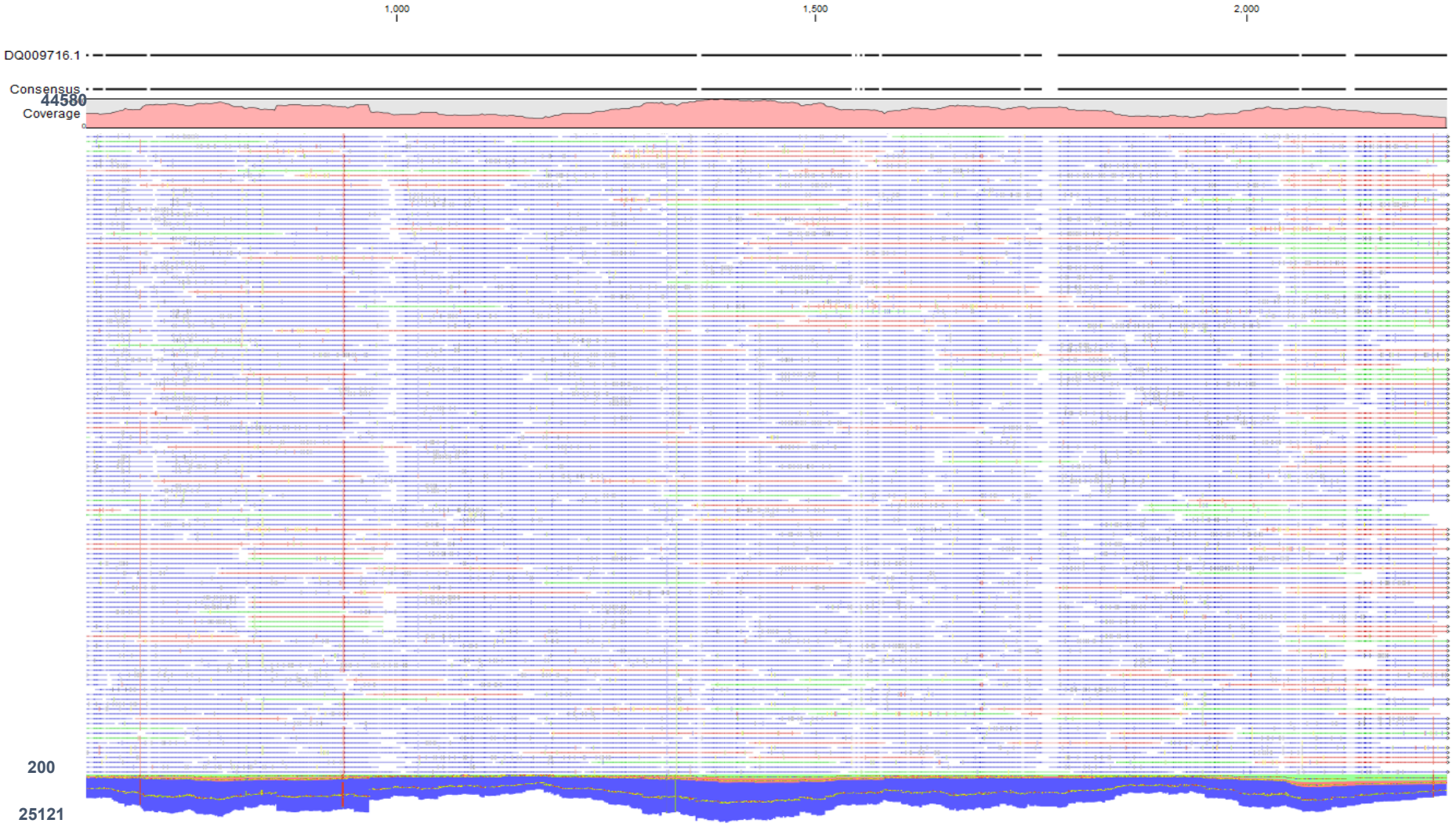
Table 1.1: The three novel FMDV SAT peptide phage library sizes compared to the theoretically required library sizes for a 99.9% probability representation of the genome.

FMDV serotype	Theoretical library sizes required	Actual library sizes (CFU)	*Genome coverage (P1-capsid coding region)
SAT1	5.19×10^2	1.3×10^3	3X
SAT2	1.74×10^3	7.1×10^2	0.41X
SAT3	9.30×10^2	1.1×10^3	1.2X

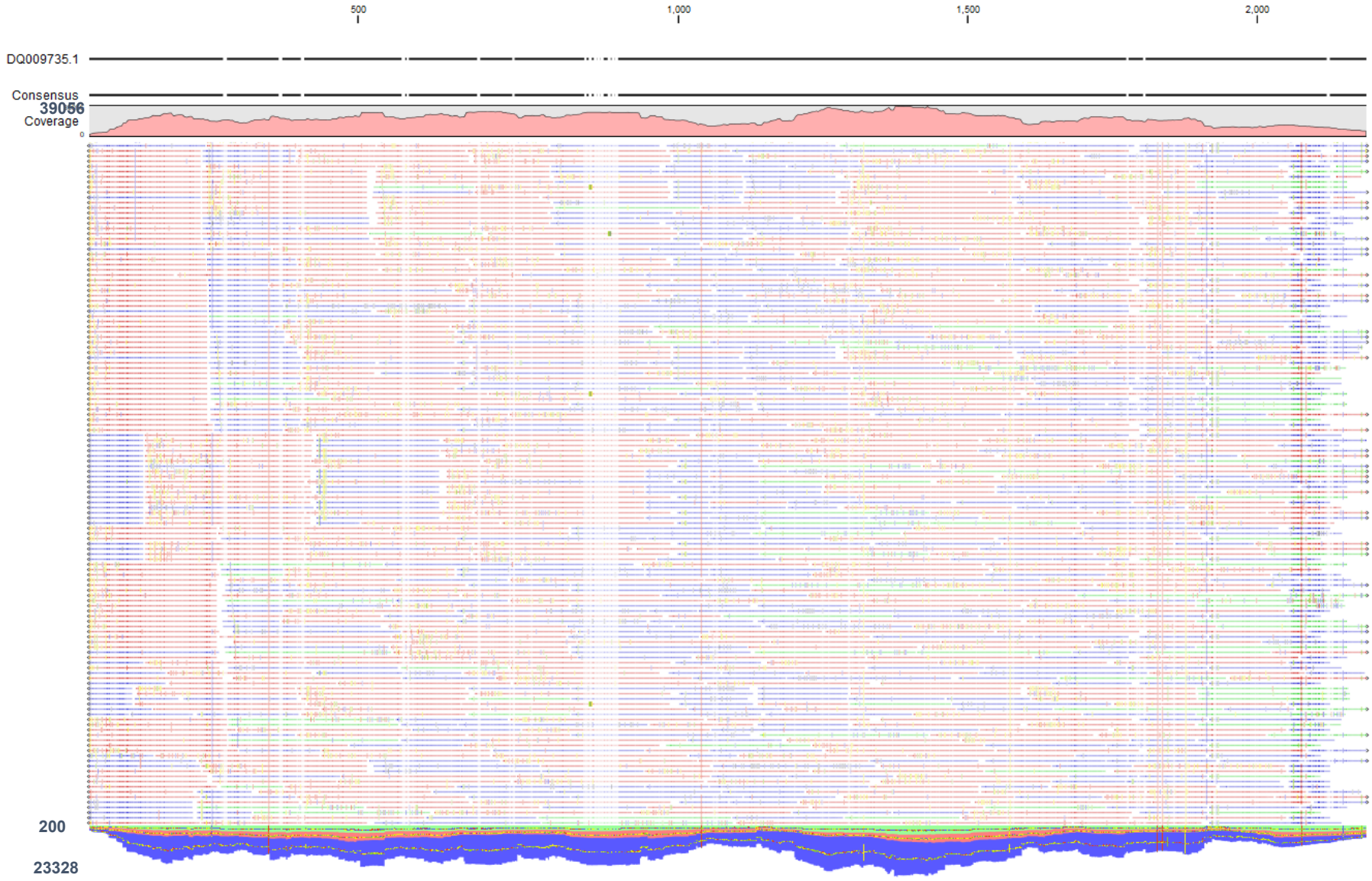
* Genome coverage = Actual library size/ Theoretical library size

2. MiSEQ of unpanned libraries

A



B


500
1,000
1,500
2,000
DQ009735.1
Consensus
39056
Coverage
0
200
23328

C

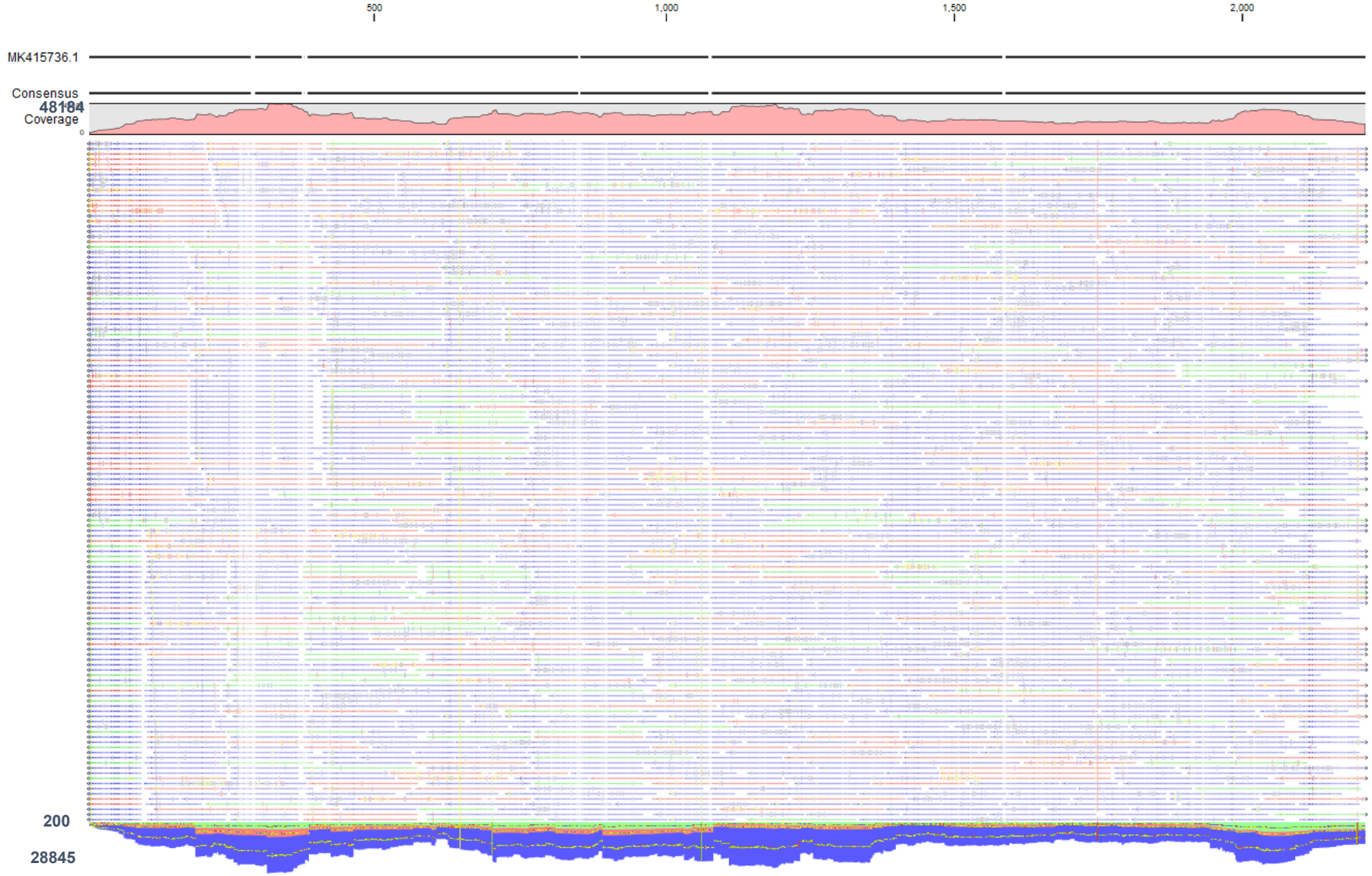

500
1,000
1,500
2,000
MK415736.1
Consensus
48184
Coverage
0
200
28845

Figure 2.1: A CLC genomic workbench read mapping representation of the unpanned FMDV constructed peptide phage libraries aligned against their respective FMDV SAT P1 reference sequence, (**A**) FMDV SAT1, (**B**) FMDV SAT2 and (**C**) FMDV SAT3. The P1 reference sequence is indicated as a bold line at the top. The blue colour indicates a paired-end read, while the green and red colours are the forward and reverse read match, respectively. The sequences from all the libraries revealed library coverage ranging from 39056-48184 and sequence depth of 23328-28845 across the entire SAT P1 regions. Numbers in the top left indicates library coverage, while those in the bottom left represents sequence depth.

Table 2.1: Unpanned FMDV SAT libraries mapping summary report.

FMDV SAT libraries mapping summary	SAT1	SAT2	SAT3
Percentage of mapped reads (%)	58	43	51
Percentage of reads in pairs (%)	53	35	43
Average mapped reads length (bp)	204	217	215

3. SPCE and NSP ELISA

Table 3.1: SPCE and NSP ELISA results of the respective FMDV SAT bovine sera samples. The percentage inhibition, (PI) >60 was considered positive.

*LABORATORY ANIMAL IDENTIFICATION NUMBER	FMD SPCE DIAGNOSTIC RESULT			FMDV NSP DIAGNOSTIC RESULT
	SEROTYPE	PI	RESULT	RESULT
15-0840	SAT1	65	**POS**	**POS**
15-0847	SAT1	66	**POS**	**POS**
15-0856	SAT1	65	**POS**	**POS**
15-0804	SAT1	62	**POS**	**POS**
15-0897	SAT1	68	**POS**	**POS**
12-0852	SAT2	81	**POS**	**POS**
12-0891	SAT2	78	**POS**	**POS**
12-0833	SAT2	79	**POS**	**POS**
1070	SAT3	83	**POS**	**POS**
1067	SAT3	77	**POS**	**POS**
1060	SAT3	78	**POS**	**POS**
1051	SAT3	78	**POS**	**NEG**

***** Animals were experimentally infected (intradermolingually, previous research study) with the following FMDV strain:
SAT1: SAT1/KNP196/91; SAT2: SAT2/KNP/19/89 and SAT3: SAT3/BOT/6/98. Biobank sera stocks of SAT3/SAR/1/06 were not available at the time of biopanning, thus the available SAT3/BOT/6/98 sera stocks were used instead for biopanning of the SAT3 library.

4. Bovine sera and IgG purification

Table 4.1: Total IgG concentrations of bovine sera after purification. The sera were obtained from animals experimentally infected with the FMDV SAT serotypes and collected at 28 days post infection (dpi) for SAT1 and 21 dpi for SAT2 and SAT3.

*Virus strains	FMDV SAT bovine sera lab animal identification number	Total IgG concentration (mg/ml)	FMDV infected (intradermolingually)
SAT1/KNP/ 196/91	15-0840	2	✓
	15-0847	1.17	
	15-0856	1.21	
	15-0804	1.14	
	15-0897	1.08	
SAT2/KNP/ 19/89	12-0852	2.28	✓
	12-0891	1.67	
	12-0833	3.09	
SAT3/BOT/ 6/98	1070	1.86	✓
	1067	2.85	
	1060	2.44	
	1051	1.84	

*The FMD SAT virus strains used to infect the animals (previous research study), where the subsequent sera used to obtain IgG for the biopanning process.

5. Peptide phage titres

Table 5.1: FMDV SAT peptide phage titres calculated in phages/ml.

FMDV SAT peptide phage library	Phage titre (phage/ml)
SAT1	3.5×10^{12}
SAT2	4.2×10^{12}
SAT3	4.0×10^{12}

6. Biopannings of FMDV SATs peptide phage libraries

Table 6.1: Output phage titres (phage/ml) from biopanning rounds of FMDV SATs peptide phage libraries against the respective purified IgGs. Three to five biopanning selection rounds were performed with each FMDV SAT library, against each IgG.

IgGs Biopanned	Output Phage Titres (Phage/ml) at each biopanning round				
	Round 1	Round 2	Round 3	Round 4	Round 5
SAT1 IgG 15-0840	2.00×10^{7}	1.50×10^{8}	1.61×10^{8}	**	**
SAT1 IgG 15-0847	1.50×10^{4}	1.04×10^{8}	4.34×10^{7}	**	**
SAT1 IgG 15-0856	1.52×10^{8}	1.46×10^{8}	2.56×10^{8}	**	**
SAT1 IgG 15-0804	5.23×10^{7}	6.60×10^{7}	6.42×10^{7}	**	**

SAT1 IgG 15-0897	6.44×10^8	6.90×10^8	2.14×10^8	**	**
SAT2 IgG 12-0852	1.30×10^2	1.24×10^2	1.33×10^2	**	**
SAT2 IgG 12-0891	2.54×10^2	3.12×10^2	3.61×10^2	**	**
SAT2 IgG 12-0833	1.56×10^2	1.50×10^2	1.62×10^2	**	**
SAT3 IgG 1070	1.68×10^6	1.20×10^6	3.50×10^6	3.72×10^6	4.23×10^6
SAT3 IgG 1067	2.43×10^3	1.95×10^2	1.87×10^3	**	**
SAT3 IgG 1060	3.91×10^2	2.58×10^2	3.10×10^2	**	**
SAT3 IgG 1051	1.30×10^2	1.77×10^2	1.93×10^2	**	**

**Biopanning rounds not performed.

7. Polyclonal phage ELISA

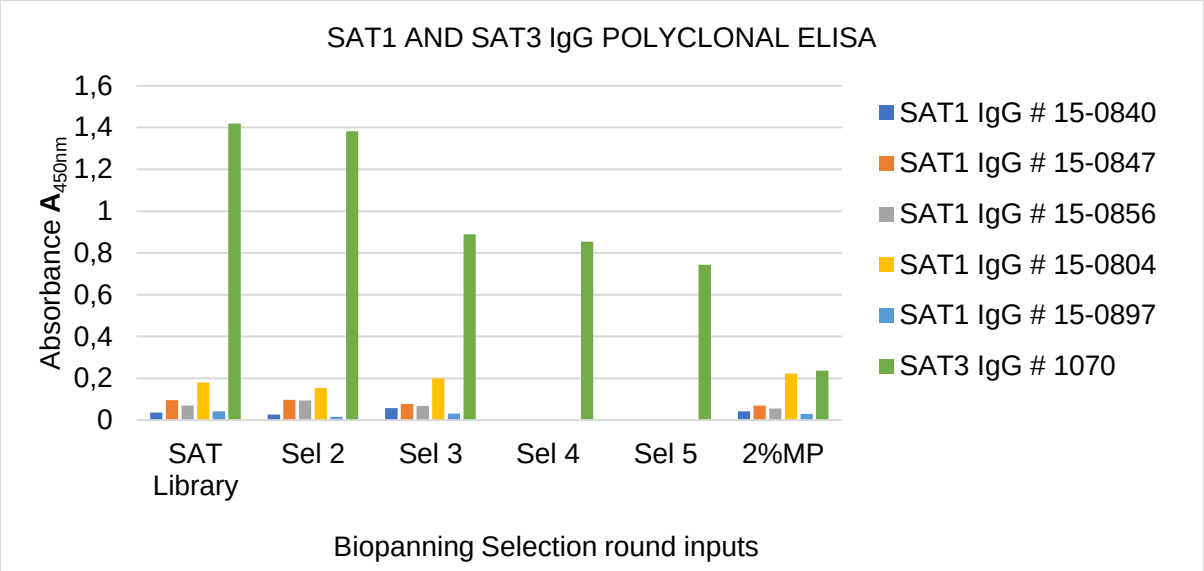


Figure 7.1: Enrichment of phage-displayed FMDV peptides using a polyclonal ELISA. Phage displayed FMDV peptides from the input phages that bound to corresponding IgGs (SAT library biopanning round) and was amplified in subsequent affinity selection (labelled sel) rounds are shown (colour key indicated). Five SAT1 IgGs screened for three affinity selection rounds is shown, as well as one SAT3 IgG screened for five affinity selection rounds. The unpanned aliquot of either the SAT1 or SAT3 peptide phage library (labelled SAT library) was a non-enriched control. The negative control used was 2% milk powder (labelled 2% MP). The ELISA signals were measured at absorbance A_{450nm}.

8. Potential FMDV binders

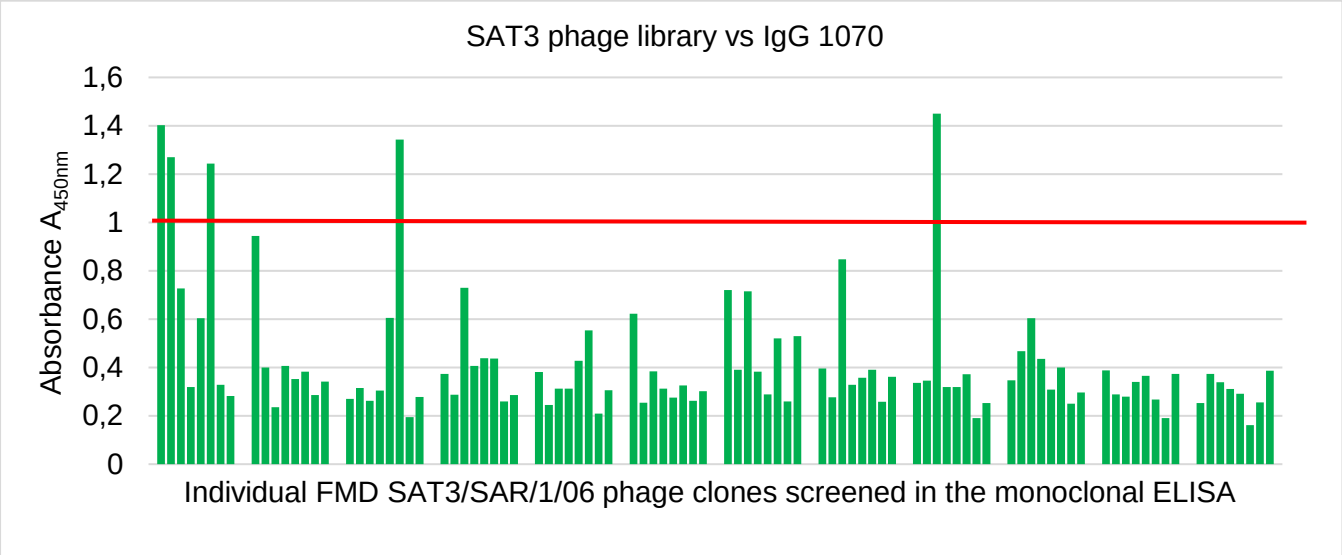


Figure 8.1: Monoclonal ELISA of single phage clones randomly selected following biopanning round five and four titration plates (only one ELISA plate result shown) to identify monoclonal SAT3/SAR/1/06-specific binders. Clones with ELISA signals of ≥1 were chosen as potential positive clones. ELISA signals were measured at absorbance A_{450nm}. These clones were named according to the respective position (well) on the 96-well ELISA plate.

9. Confirmation of positive binders

Figure 9.1: Alignment of FMDV SAT3 binders *i.e.,* G2, C8 and B1 sequences to the pVCEPI585042 phagemid vector (BioEdit Alignment Sequence Editor). All three SAT3 binder sequences had the same sequence. The red box shows the sequence of the SAT3 FMDV peptide inserted into the pVCEPI585042 vector during cloning. The sequences flanking the red box for each binder indicates the vector sequence where the restriction enzyme (*PmeI*) cleavage site is highlighted in yellow.

Figure 9.2: A graphic overview of the NCBI database of FMDV SAT3 sequences matching and aligning to the FMDV SAT3φ1 sequence when conducting a blast search (NCBI). The horizontal-coloured coded bars represent sequence alignment by % scores and shows the extent match of the alignment to the SAT3φ1 sequence. Alignment scores are colour coded and are shown above the sequence distribution overview.

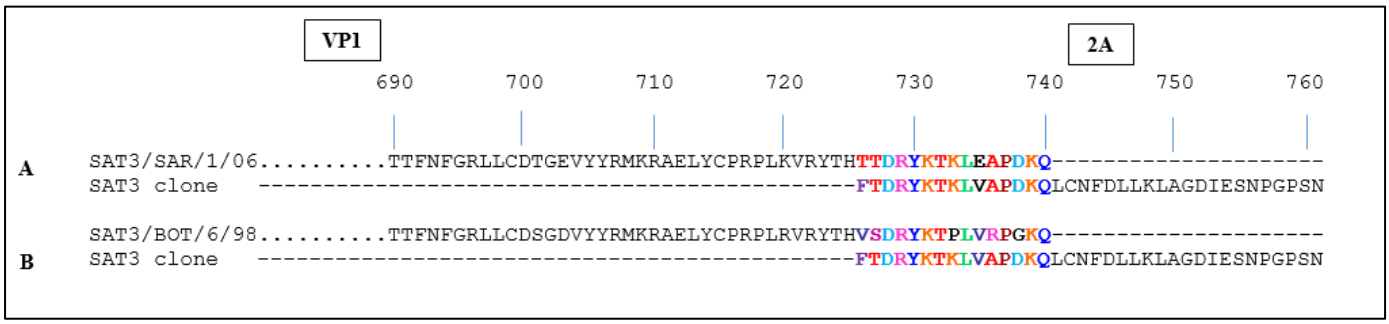


Figure 9.3: Partial FMDV genome P1/2A amino acid sequence alignment of the FMDV SAT3φ1 to published FMD SAT3 virus strains. The SAT3φ1alignment to **(A)** SAT3/SAR/1/06 and **(B)** SAT3/BOT/6/98 virus strains are shown (partial C-terminus P1 sequences, N-terminus 2A sequence not observed). The dots (.) represents a continuous sequence, while the dash (-) represents absent amino acids.

10. Illumina sequencing

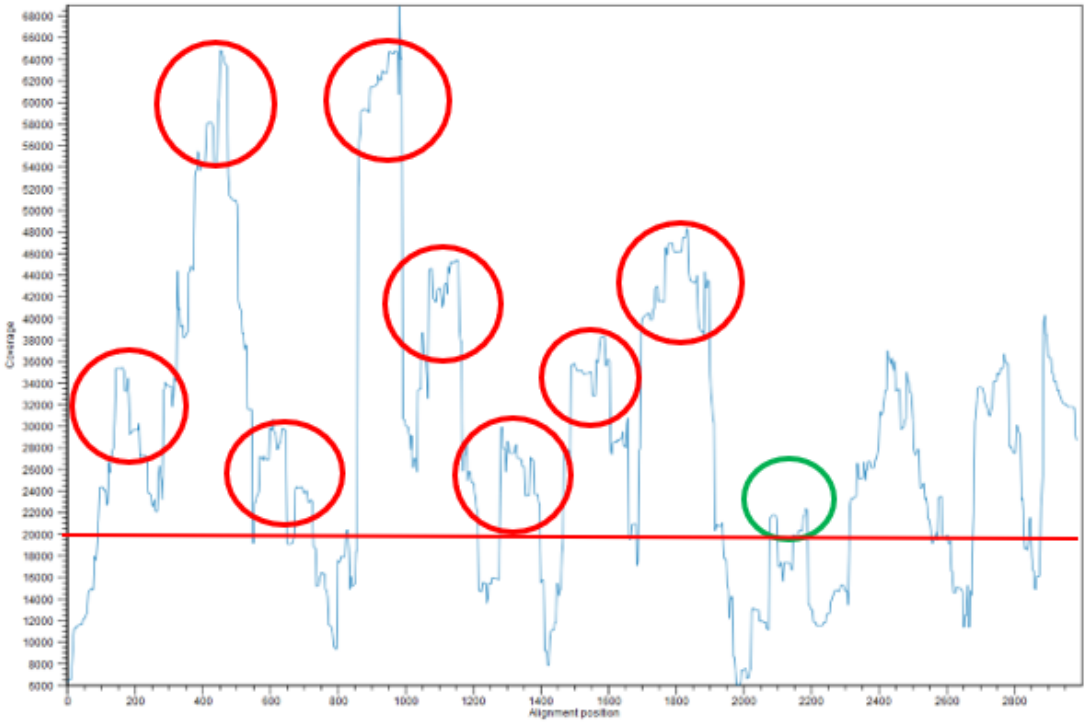


Figure 10.1: The CLC analysis of the Illumina data, showing the coverage across the FMDV SAT3 P1/2A region from the output selection round. Regions with the highest coverage (≥20000) indicate potential epitopic regions. This includes the sequence for SAT3φ1, nucleotide positions at 2178 to 2220 (green circle). Other potential epitopic regions (red circle) include nucleotide position regions (bp) 105 to 200, 401 to 450, 550 to 701, 900 to 1001, 1050 to 1101, 1301 to 1401, 1500 to 1600 and 1701 to 1900 of the FMDV SAT3 P1 region.

Table 10.1: Potential antigenic sites located on the FMDV SAT3 P1 region.

Genome position	Nucleotide/alignment positions (bp)	Amino acid positions	Number of amino acid
*VP1	1701-1900	567-633	66
VP1	2178-2220	726-740	14
VP2	401-450	134-150	16
VP2	550-701	183-234	51
VP2 to VP3	900-1001	300-334	34
VP3	1050-1101	350-367	17
VP3	1301-1401	434-467	33
VP3 to VP1	1500-1600	500-533	33
VP4	105-200	35-67	32

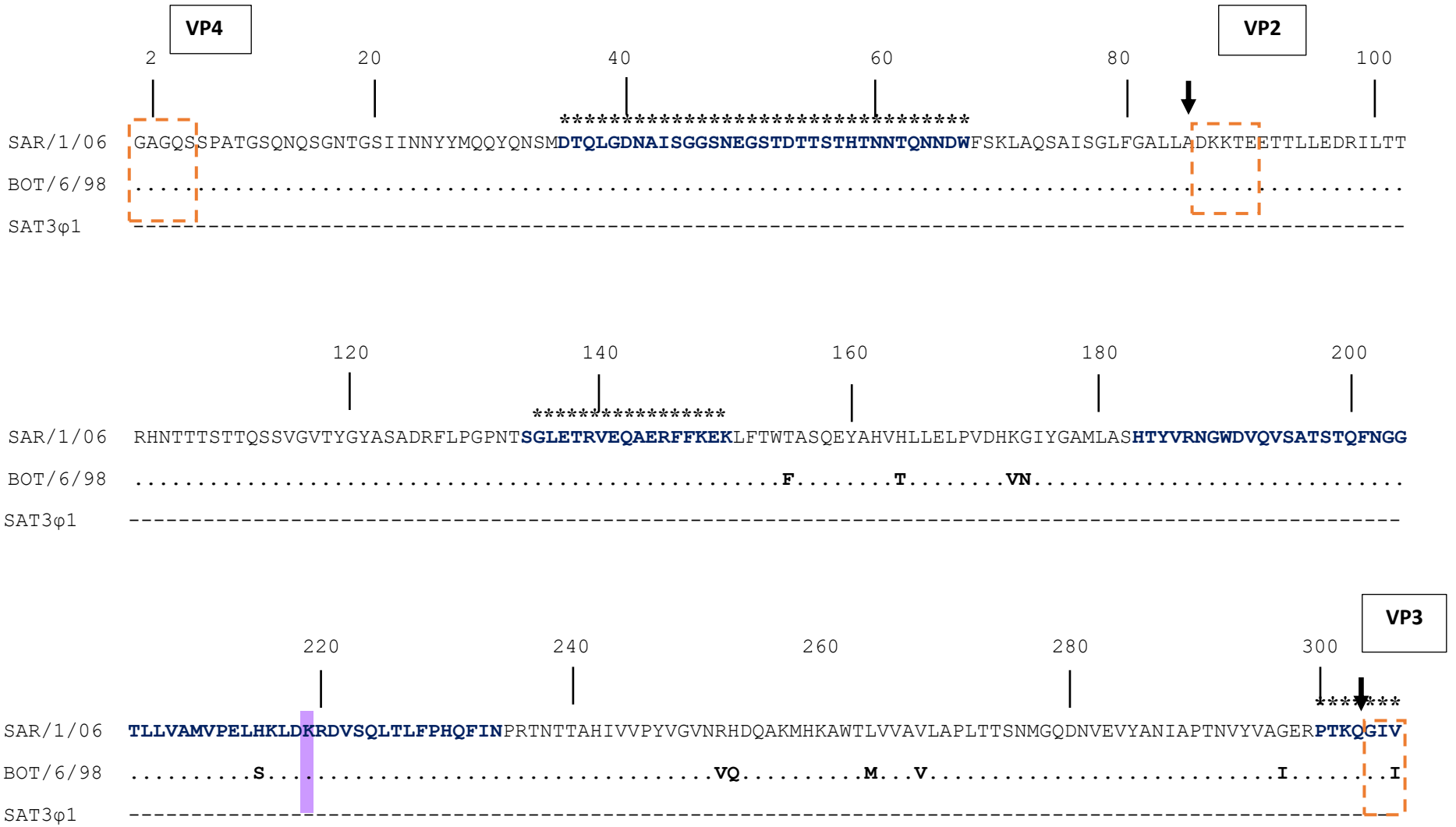
*VP=Viral protein

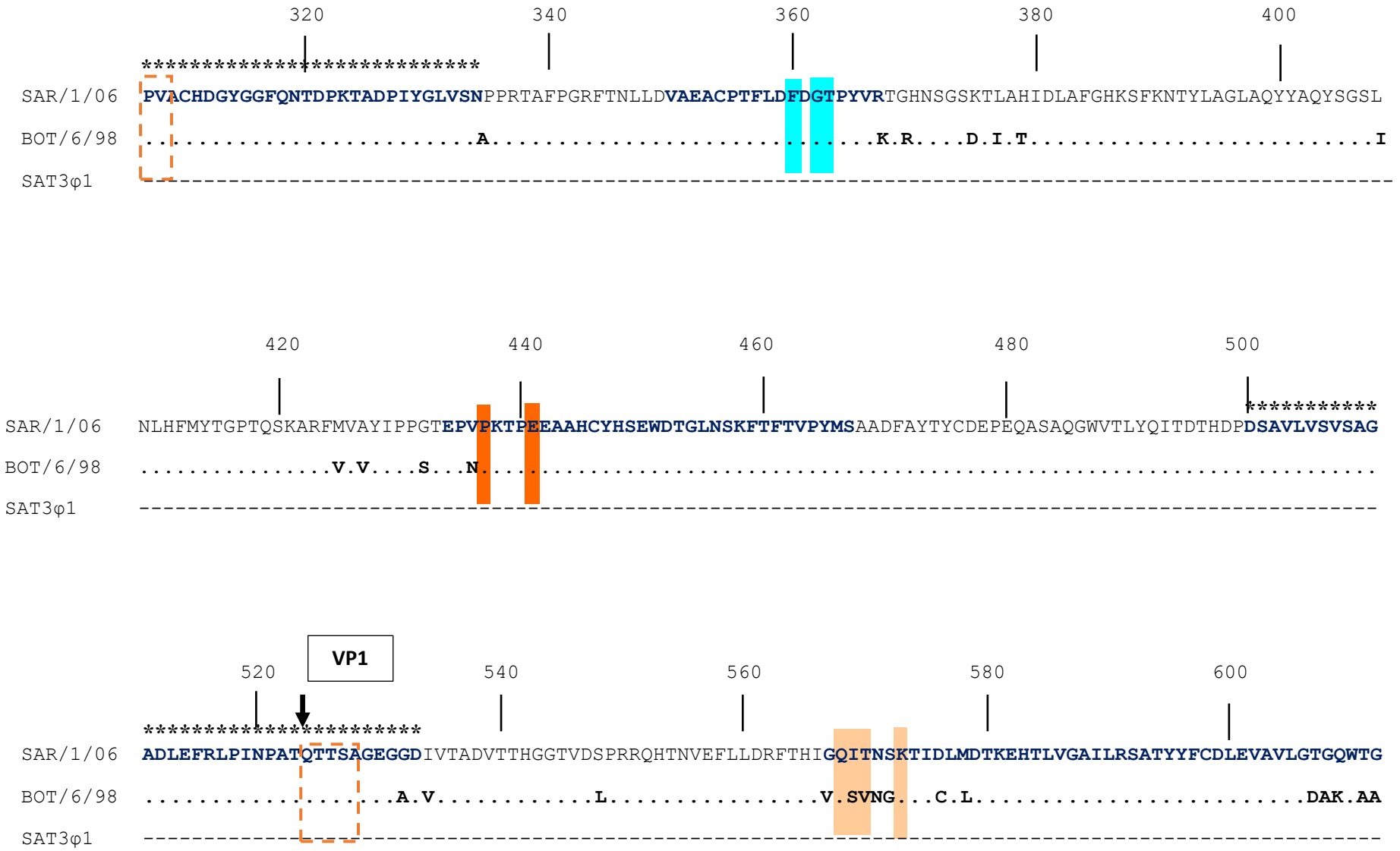
Table 10.2: Comparison of identified potential FMDV SAT3 antigenic sites to known published antigenic sites for FMDV.

*Potential antigenic sites and amino acid positions	Known antigenic sites	FMDV serotypes
Amino acids 567-633 of P1 (42-108 of VP1)	Site 3 amino acid residues 43-45 and 48 of VP1 (Barnett *et al.,* 1998; Kitson *et al.,* 1990; Grazioli *et al.,* 2013).	O
Amino acids 726-740 of P1 (201-215 of VP1)	Residues 200-213 of the C-terminus of VP1 for serotype O (Xie *et al.,* 1987; Parry *et al.,* 1989; Grazoili *et al.,* 2013) and serotype C (Mateu *et al.,* 1990; Grazoili *et al.,* 2013). Amino acid 208 of VP1 (Mahapatra *et al.,* 2012; Opperman, 2013).	O, C
Amino acids 183-234 of P1 (98-149 of VP2)	VP2 residue 134 (SAT3) (Maake *et al.,* 2020; Mukonyora, 2015). SAT2 and serotype O (Crowther *et al.,* 1993; Grazioli *et al.,* 2013; Maake *et al.,* 2020).	O, SAT2 and SAT3
Amino acids 350-367 of P1 (46-63 of VP3)	Amino acid residues 56 and 58-59, located at the β-B “knob” of VP3 (McCullough *et al.,* 1987; Barnett *et al.,* 1998) of serotype O and serotype A, Asia1 and site D3 of serotype C (Thomas *et al.,* 1988; Kitson *et al.,* 1990; Lea *et al.,* 1994; Grazioli *et al.,* 2013).	O, A, Asia1 and C
Amino acids 434-467 of P1 (132-165 of VP3)	Amino acid residue 139, was for Serotype A, (Thomas *et al.,* 1988) and amino acid residue at 135 of VP3 for SAT1 viruses (Grazioli *et al.,* 2013; Maake *et al.,* 2020).	A and SAT1

* Detailed P1/2A amino acid alignment positions are shown in Appendix C.

APPENDIX C: SEQUENCE ALIGNMENT




320
340
360
380
400
SAR/1/06 PVACHDGYGGFQNTDPKTADPIYGLVSNPPRTAFPGRFTNLLDVAEACPTFLDFDGTPYVRTGHNSGSKTLAHIDLAFGHKSFKNTYLAGLAQYYAQYSGSL
BOT/6/98 .....A.....K.R....D.I.T.....I
SAT3φ1 ----------
420
440
460
480
500
SAR/1/06 NLHFMYTGPTQSKARFMVAYIPPGTEPVPKTPEEAAHCYHSEWDTGLNSKFTFTVPYMSAADFAYTYCDEPEQASAQGWVTLYQITDTHDPDSAVLVSVSAG
BOT/6/98 .....V.V....S...N.....
SAT3φ1 ----------
VP1
520
540
560
580
600
SAR/1/06 ADLEFRLPINPATQTTSAGEGGDIVTADVTTHGGTVDSPRRQHTNVEFLLDRFTHIGQITNSKTIDLMDTKEHTLVGAILRSATYYFCDLEVAVLGTGQWTG
BOT/6/98 .....A.V.....L.....V.SVNG...C.L.....DAK.AA
SAT3φ1 ----------

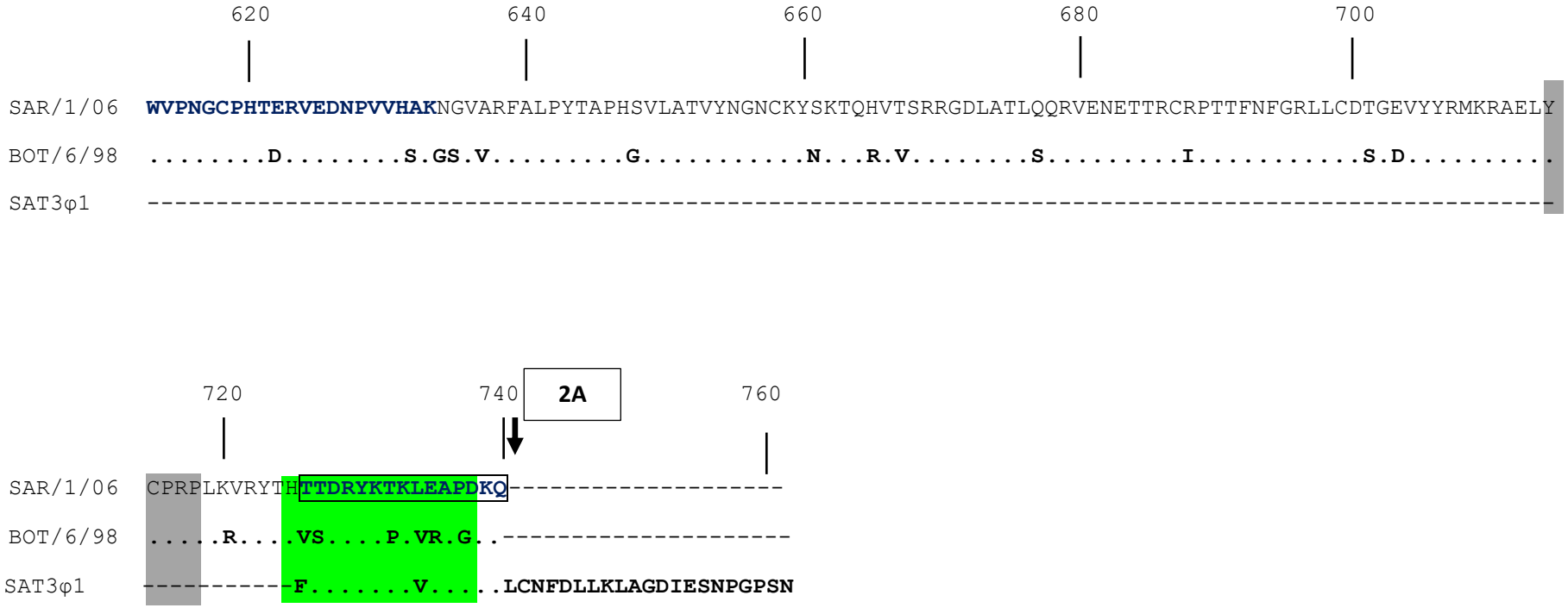

620
640
660
680
700
SAR/1/06 WVPNGCPHTERVEDNPVVHAKNGVARFALPYTAPHSVLATVYNGNCKYSKTQHVTSRRGDLATLQQRVENETTRCRPTTFNFGRLLCDTGEVYYRMKRAELY
BOT/6/98 ........D.........S.GS.V.........G............N...R.V.........S.........I............S.D.........
SAT3φ1 ----------
720
740
2A
760
SAR/1/06 CPRPLKVRYTHTTDRYKTKLEAPDKQ----------
BOT/6/98 .....R....VS....P.VR.G..----------
SAT3φ1 ----------F.......V.....LCNFDLLKLAGDIESNPGPSN

Figure A1: Amino acid alignment of the P1 region of two FMDV SAT3 virus strains and the FMDV SAT3 clone (SAT3φ1), including the 2A junction overlap of SAT3φ1. Identified potential FMDV SAT3 antigenic sites are indicated in **bold blue**. Known FMDV antigenic sites that correlates to potential antigenic sites from this study are shown (colour-coded). The dot (**.**) represents similarity of amino acids, whilst a dash (**-**) represents the absence of amino acids. The ⬚ represents the amino acid residues at the N-terminus of the FMD virus structural proteins of P1 *i.e.,* VP4, VP2, VP3 and VP1 respectively.

<u>The identified potential antigenic sites with known FMDV antigenic sites (Refer to alignment sequence):</u>

Potential antigenic site [Amino acids 350-367 of P1 (46-63 of VP3)]

Amino acid residues 56 and 58-59, located at the β-B "knob" of VP3 constitutes as site 4 (McCullough *et al.,* 1987; Barnett *et al.,* 1998) of serotype O was identified as an antigenic site. This antigenic site corresponds to site 4 of serotype A, Asia1 and site D3 of serotype C (Thomas *et al.,* 1988; Kitson *et al.,* 1990; Lea *et al.,* 1994; Grazioli *et al.,* 2013).

Potential antigenic site [Amino acids 434-467 of P1 (132-165 of VP3)]

A distant position amino acid residue 139, was also an antigenic site identified for Serotype A, forming part of site 4 (Thomas *et al.,* 1988).
Using MAb resistant (MAR) mutants, amino acid residue at 135 of VP3 was identified to be of importance for the antigenicity of SAT1 viruses (Grazioli *et al.,* 2013; Maake *et al.,* 2020).

Potential antigenic site [Amino acids 183-234 of P1 (98-149 of VP2)]

VP2 residue 134 has been identified as an antigenic site associated with SAT3 (KNP/1/03 and ZIM/5/91) (Maake *et al.,* 2020). The same site was identified by Mukonyora (2015) using in silico predictions programmes. This site was previously identified as an antigenic site for FMDV SAT2 and serotype O (Crowther *et al.,* 1993; Grazioli *et al.,* 2013; Maake *et al.,* 2020).

Potential antigenic site [Amino acids 567-633 of P1 (42-108 of VP1)]

Serotype O site 3 was mapped at amino acid residues 43-45 and 48 of VP1 (Barnett *et al.,* 1989; Kitson *et al.,* 1990; Grazioli *et al.,* 2013). Based on their location (H-I and B-C loop respectively), these sites were suggested to be related (Grazioli *et al.,* 2013).

Potential antigenic site [Amino acids 726-740 of P1 (201-215 of VP1)]

A conformationally dependent and trypsin-sensitive site (site 1 of Serotype O), which is located on the βG-βH loop of VP1 also involves amino acid residue 208 (Mahapatra *et al.,* 2012; Opperman, 2013). Residues 200-213 of the C-terminus of VP1 has been reported to form part of the discontinuous site, with site 1 of serotype O (Xie *et al.,* 1987; Parry *et al.,* 1989; Grazoili *et al.,* 2013). The same site was identified in serotype C (site C) (Mateu *et al.,* 1990; Grazoili *et al.,* 2013).

The conserved motif (Amino acid 185-190 of VP1 C-terminus)

An important motif for correct structure maintenance of picornavirus capsid precursors (P1-2A), before processing and subsequent capsid assembly. This site may represent interacts with cellular chaperones (Kristensen and Belsham, 2019).

*****No known published research data for amino acid positions **35-67**, **134-150**, **300-334** and **500-533** of the FMDV SAT3 P1 region could be correlated to FMDV antigenic sites. Thus, this study proposes these amino acids as novel potential antigenic sites for FMDV SAT3.