Journal of Medical Virology 81:511–518 (2009)

Identification of Mutations in a Temperature-Sensitive Mutant (*tsm*5) of Murine Cytomegalovirus Using Complementary Genome Sequencing

Olga Timoshenko,[1] Abdulaziz Al-Ali,[1] Brian A.B. Martin,[2] and Clive Sweet[1]*

[1]*School of Biosciences, University of Birmingham, Edgbaston, Birmingham, United Kingdom*
[2]*Department of Infection and Immunity, Medical School, University of Birmingham, Edgbaston, Birmingham, United Kingdom*

Identification of mutations in mutants derived chemically is a difficult and relatively random process. NimbleGen Comparative Genome Sequencing (CGS) was assessed as an inexpensive, rapid method of identifying mutations in the temperature-sensitive mutant *tsm*5 of the K181 (Birmingham) variant of murine cytomegalovirus (MCMV). This genome resequencing approach requires an established genome sequence as a reference. Comparison of *tsm*5 and the K181 (Birmingham) variant with the published K181 (Perth) MCMV genomic sequence revealed a total of 10 synonymous and 15 non-synonymous SNPs in *tsm*5 and 14 of the latter were confirmed by sequencing. Thus, while CGS cannot be relied upon to identify correctly all mutations it was helpful for identifying a large number of mutations for further investigation that could contribute to the *ts* phenotype of *tsm*5. ***J. Med. Virol. 81:511–518, 2009.*** © 2009 Wiley-Liss, Inc.

KEY WORDS: MCMV; murine; cytomegalovirus; tsm5; CGS

INTRODUCTION

A mutant, *tsm*5, of the murine cytomegalovirus (MCMV) K181 (Birmingham) variant, created by mutagenesis with *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (NTG) and restricted in growth at 40°C, has been described previously [Sammons and Sweet, 1989]. Titers of this mutant were reduced by ~4 $\log_{10}$ at 40°C compared to 33 or 37°C and failed to reach detectable levels in any tissue of 1-week-old mice for up to 21 days following i.p. inoculation with 4×10^3 pfu of virus. It is unlikely that the lack of replication in mice reflected solely the *ts* phenotype of *tsm*5 as it replicated in SCID mice [Morley et al., 2002].

Electron microscopy showed that *tsm*5 had a defect in DNA encapsidation and Southern hybridization indicated that, while DNA synthesis was reduced significantly by ~90% at 40°C, concatameric DNA was cleaved into unit length genomes with close to *wt* efficiency [Sweet et al., 2007]. Sequence analysis of genes involved in DNA replication, processing, and packaging identified three mutations: a C to T mutation leading to a G439R residue change in the large terminase subunit (M56); a C to T mutation (P324S) in the alkaline nuclease (M98); and a C to T mutation (C890Y) in the primase component (M70) of the helicase-primase complex [Sweet et al., 2007]. The significance of these mutations to the *ts* phenotype of *tsm*5 is currently being examined. However, while a chimeric virus containing the central portion of its genome (*M56* to *m138*) from *tsm*5 and the 3′ and 5′ ends from the Smith strain of MCMV was still temperature sensitive it was less than *tsm*5 and now replicated in mice suggesting that other mutations may also be contributing to the *tsm*5 phenotype.

Clearly, these mutations need to be identified. A relatively inexpensive approach, that is a rapid method of screening whole genomes for changes against a reference genome, is the Comparative Genome Sequencing (CGS) service provided by NimbleGen Systems, Inc. (Madison, WI) [Albert et al., 2005]. It requires minimal sample preparation as no libraries need to be generated and DNA is labeled with a standard random prime reaction. The entire genome can be screened and mutations such as single nucleotide polymorphisms

Olga Timoshenko and Abdulaziz Al-Ali contributed equally to this work.

Grant sponsor: Darwin Trust of Edinburgh studentship (to O.T.); Grant sponsor: Saudi Arabia Education Department studentship (to A.A.-A.); Grant sponsor: Rowbotham Bequest University of Birmingham Medical School (to C.S.); Grant sponsor: BBSRC; Grant number: 6/JIF13209.

*Correspondence to: Clive Sweet, School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: c.sweet@bham.ac.uk

Accepted 17 November 2008

DOI 10.1002/jmv.21419

Published online in Wiley InterScience (www.interscience.wiley.com)

© 2009 WILEY-LISS, INC.

(SNPs), amplifications, and deletions can be identified. The procedure uses a two-step approach. Initially, putative mutations are mapped approximately as Regions of Interest (ROIs) when low molecular fragments of the test genome (labeled with 5′-Cy3) and similarly sized fragments of the reference genome (labeled with 5′-Cy5) are hybridized to microarrays comprised of 385,000 29-mer "tiled" oligonucleotides spaced every seven bases on each strand of the reference genome sequence (Fig. 1). Hybridization intensity ratios (reference/test) plotted versus genome position maps possible mutation sites. The second step, referred to as re-sequencing, is when each nucleotide within 48 kb around an ROI is interrogated with eight 29-mer probes, four per genome strand, each probe differing only with respect to the central position (A, C, G, or T) (Fig. 1). When target DNA is hybridized to these probes the perfectly matched probe hybridizes more strongly than the three corresponding mismatched probes for each strand. The differential signal intensity between the perfectly matched probe and the mismatched probes allows the base to be determined precisely.

Approximately 95% of all SNPs are identified by this procedure and false positives occur at <1 in 100,000 bases analyzed (NimbleGen Systems, Inc.). A disadvantage of the CGS approach is that it requires a known genome sequence to facilitate the synthesis of the reference oligonucleotide array. However, this approach became feasible with the recent publication of the K181 (Perth) MCMV genomic sequence [Smith et al., 2008]. This relatively new approach to sequencing has been used for bacterial resequencing [Herring and Palsson, 2007] and to track the evolution of the SARS coronavirus [Wong et al., 2004] but has not been used previously to identify mutations in a mutant virus.

MATERIALS AND METHODS

Viruses

The K181 (Birmingham) variant of MCMV and mutant *tsm*5 has been described previously as has the method of production of primary mouse embryo fibroblasts (MEFs) and virus stocks [Sammons and Sweet, 1989; Sweet et al., 2007]. The K181 (Perth) variant of MCMV was kindly provided by Dr. Alec Redwood, Perth, Australia as a bacterial artificial chromosome (BAC) containing the K181 (Perth) MCMV genome. MCMV BAC-derived virus was produced from this by transfection of NIH 3T3 cells with MCMV BAC DNA using the ExGen 500 in vitro transfection reagent (Fermentas, York, UK) as described by the manufacturer. Individual plaques of recovered virus were passaged in MEFs to eliminate the BAC cassette.

Production of MCMV DNA

Five 162 cm^2 tissue culture flasks were infected with K181 (Perth), K181 (Birmingham), or *tsm*5 at a moi of

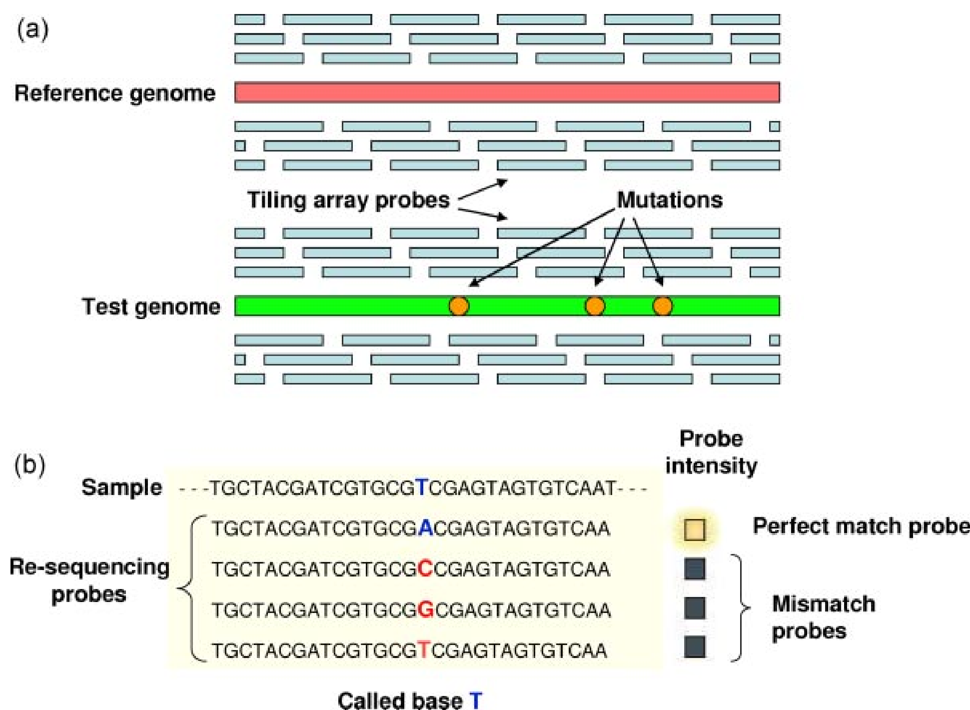

Fig. 1. Diagrammatic representation of the CGS process. **A**: Labeled reference and test genomic DNA is hybridized in parallel to microarrays composed of 29-base-long oligonucleotides tilting array probes tiled with 7-base spacing. **B**: Re-sequencing probes are short oligonucleotides in which every possible variant is represented at the central position of the probe (bold font). Four probes for each strand are used to interrogate each nucleotide position. Hybridization of the sample with perfect match probe will result in the highest intensity (white colored box) [Modified from Albert et al., 2005; Gresham et al., 2008].

0.01 and incubated at 37°C until cpe reached 100% when the supernatant was harvested. Cell debris was pelleted at 850*g* for 10 min at room temperature and the supernatant containing virus filtered through a Whatman filter paper directly into 40 ml Oakridge-style centrifuge tubes. The virus was then centrifuged at 29,000*g* for 30 min at 4°C and the pellet resuspended in 500 μl of DNase I buffer [50 mM Tris–HCL (pH 8.0), 5 mM $MgCl_2$ 0.1 M sodium acetate, and 100 μg bovine serum albumin per ml]. Cellular DNA was digested with 0.2 U of DNase I (Sigma-Alrdich, Dorset, UK) for 1 hr at room temperature and the reaction stopped with 20 μl of 500 mM EDTA (pH 8.0). Virions were treated with 1% sodium dodecyl sulphate and 40 μl of proteinase K (20 mg/ml) for a minimum of 4 hr at 56°C to release viral DNA, which was then purified by phenol chloroform extraction and precipitated in 1 volume of isopropanol [Redwood et al., 2005]. Preparations contained ~10 μg of DNA at a concentration of 250–500 ng/ μl.

Labeling of DNA

Genomic DNA was fragmented by sonication and labeled with a random prime reaction. DNA (1 μg) was mixed with 1 O.D. of 5′-Cy5 or 5′-Cy3 labeled random nonamer (TriLink Biotechnologies, San Diego, CA) in 62.5 mM Tris–HCl, 6.25 mM $MgCl_2$, and 0.0875% β mercaptoethanol, denatured at 98°C for 5 min, chilled on ice and incubated with 100 units Klenow fragment (NEB) and dNTP mix (6 mM each in TE buffer) for 2 hr at 37°C. Reactions were terminated with 0.5 M EDTA (pH 8.0), precipitated with isopropanol and resuspended in water.

Mutation Mapping Microarray Design

Mutation mapping microarrays were designed, based on the recently published annotated sequence of the K181 (Perth) MCMV variant [Smith et al., 2008: EMBL accession number am886412], with NimbleGen Systems, Inc. algorithms [Albert et al., 2005] that select a 29-mer oligonucleotide every seven bases on each strand of genome sequence. The 385,000 oligonucleotide array was designed such that the entire genome was covered by the 29 mer oligonucleotides with a 7 bp overlap. All oligonucleotides were synthesized in parallel on an array set using a Digital Light Processor (Texas Instruments, Plano, TX) and photoprotected phosphoramidite chemistry (markless array synthesis) 5,6 (NimbleGen Systems, Inc.) in a random probe layout.

Microarray Hybridization

Labeled genomic DNA was hybridized to arrays in 1× NimbleGen Hybridization Buffer (NimbleGen Systems, Inc.) for 16 hr at 42°C. Genomic DNA from the reference virus (labeled with Cy5) and from each test virus (labeled with Cy3) were co-hybridized on mutation mapping arrays for 16 hr, washed with non-stringent buffer, centrifuged in a NimbleGen custom centrifuge to remove buffer and then stored until scanned.

Analysis of Mapping Array Data and Hybridization of Resequencing Arrays

Microarrays were scanned at 5 μm resolution using a Genepix® 4000b scanner (Axon Instruments, Union City, CA) and pixel intensities were extracted using NimbleScan™ software (NimbleGen Systems, Inc.). Probes that spanned potential mutations were identified using custom software. Probe sequences corresponding to all possible candidate mutation sites were selected for resequencing. The software was designed to be sensitive to the slightest reduction in hybridization of test relative to control so that 48,000 bases were reexamined in each re-sequencing array, utilizing 388,800 oligonucleotides. Over-sampling of potential mutation sites increases the likelihood of identifying mutations that were not reported strongly by the mapping arrays, reducing the overall false negative rate. Eight probes per base position are generated, four per genome strand. These probes contain all possible alleles at a centrally located position of the 29-mer probe. The length, melting temperature, and mismatch position of each probe were optimized. When target DNA is hybridized to these arrays the perfect match probe will hybridize more strongly than the three corresponding mismatch probes for each strand. This differential signal intensity between the perfect probe and mismatch probes allows the base to be determined precisely. These re-sequencing arrays were synthesized, hybridized with labeled genomic DNA from each test line and scanned as above. Sequence base assignments were made using a machine-learning algorithm [Albert et al., 2005].

RESULTS

Mutation Mapping of K181 (Birmingham) MCMV Variant Genome

For mutation mapping a two-step protocol was adopted that first maps the mutations approximately and then determines the mutation sequence precisely as described by Albert et al. [2005]. In the first step, K181 (Perth) DNA labeled with Cy5 (the reference strain) and K181 (Birmingham) DNA labeled with Cy3 (the test strain) was hybridized to the K181 (Perth) microarray. Probes showing differences in signal intensity were flagged as ROIs for further investigation. Two ROIs were identified at positions 10,067 to 10,096 and 40,230 to 40,259 in the K181 (Birmingham) genome. Three SNPs were then identified in the second step using the re-sequencing arrays (Table I).

In the K181 (Birmingham) sequence, two of the putative SNPs were confirmed, the m10 V152G and M32 F393L, but not the m34.2 G80R mutation (Table I). The latter mutation, which was not present in either ROI and was only recognized during resequencing which interrogates 48 kb around an ROI (see Methods

TABLE I. Polymorphisms Identified in K181 (Birmingham) Compared to K181 (Perth) and Mutations Identified in *tsm*5 by CGS

ORF	ORF position (nt)[a]		Position of nt mutation	Nucleotide change	Amino acid change
	From	To			
K181 polymorphisms					
m10	9,624	10,499	10,078	T to G	V152G
M32	39,261	41,426	40,250	A to G	F393L
M34	43,073	45,646	45,607	C to G	A846A[c]
m34.2	45,512	45,844	45,607	C to G	G80R[c]
*tsm*5 mutations[b]					
m20	20,575	23,028	20,949	C to T	A694T
			22,698	C to T	E111K
m21	22,628	23,317	22,698	C to T	S24F
M25	25,995	287,739	27,523	G to A	R510K
			28,697	G to A	Q920Q[d]
M34	43,073	45,646	44,687	C to T	P539S
			44,715	T to G	V548G[c]
m34.2	45,512	45,844	45,740	G to A	P35P[d]
M36 exon 1	49,031	49,267	49,108	C to T	V54I
M45	59,515	63,039	61,705	G to A	I445I[d]
M47	63,924	67,046	65,612	G to A	Q563Q[d]
			66,658	C to T	A912V
M53	78,456	79,454	79,272	G to A	V273M
M54	79,694	82,987	82,379	G to A	S203S[d]
M69	96,243	98,771	98,442	C to T	R110R[d]
M70	99,060	101,954	99,286	C to T	C890Y
M82	115,757	117,556	115,946	C to T	G537G[d]
M87	127,433	130,213	127,756	C to T	V108V[d]
M97	140,185	142,116	141,258	C to T	G358G[d]
M98	142,145	143,830	143,114	C to T	P324S
m127	185,267	185,668	185,594	C to T	R25R[d]
m132.1	188,672	189,499	188,683	C to A	G273C
m139	194,162	196,096	194,402	G to T	Y565X
m141	197,787	199,313	198,731	C to T	V195M
m143	201,008	202,639	201,944	C to T	M232I
Intergenic			142,121	C to T	Promoter of M98?

[a]ORF positions are based on the K181 (Perth) sequence (EMBL Acc. No. AM886412).
[b]The K181 (Birmingham) SNPs were also present in *tsm*5.
[c]These mutations were not confirmed by sequencing.
[d]Synonymous mutations were not sequenced.

Section), would also have created a synonymous A846A mutation in ORF M34 (Table I). Of more concern than this false positive were the false negatives. Comparing previous sequencing of a number of genes in K181 (Birmingham) [Sweet et al., 2007] with the newly annotated K181 (Perth) genomic sequence a number of synonymous and non-synonymous polymorphisms were identified, all of which were not evident in the CGS analysis: an A (K181 Perth) to C (K181 Birmingham) polymorphism at nucleotide 41,958 in the M33 ORF leading to a change to leucine in K181 (Birmingham) from a methionine in K181 (Perth) in residue 107; a T to C at nucleotide 82,062 (D309G) in M54; a G to A at nucleotide 89,124 (A919V) in M57; a C to A at 202,998 (Q314H), a C to T at 203,055 (R295R), an A to G at 203,070 (G290G), a TT to CC at 203,182 and 203,183 (K253G), a G to A at 203,238 (Y234Y) and a C to G at position 203,515 (S142T), all in m144.

Mutation Mapping of Mutant *tsm*5

To identify mutations in *tsm*5, DNA prepared from K181 (Birmingham) labeled with Cy5 (as reference strain) and *tsm*5 DNA labeled with Cy3 (as test strain) was hybridized to the K181 (Perth) microarray. Fifteen ROIs were identified in *tsm*5, which revealed a total of 10 synonymous and 15 non-synonymous SNPs and 1 mutation characterized as intergenic (Table I). Identified non-synonymous changes were then examined by conventional ABI3700 capillary sequencing.

Fourteen of the 15 non-synonymous SNPs identified in *tsm*5 were confirmed by sequencing, the false positive being the T to G change (V548G) in M34 (Table I). Importantly, mutations observed previously from sequencing were confirmed by CGS, namely the C890Y mutation in M70, the P324S mutation in M98 and the V54I mutation in M36 [Sweet et al., 2007]. Again, two false negative mutations were observed, the M56 G439R and M27 A658S mutations shown in Figure 2. It is possible that these were missed because it is now evident that *tsm*5 is polymorphic in at least five loci and therefore contains a mixture of related viruses (Fig. 2). However, the relative proportions of each mutant within the stock vary with the stock preparation. All except two of the non-synonymous SNPs identified by CGS were single peaks on sequencing,

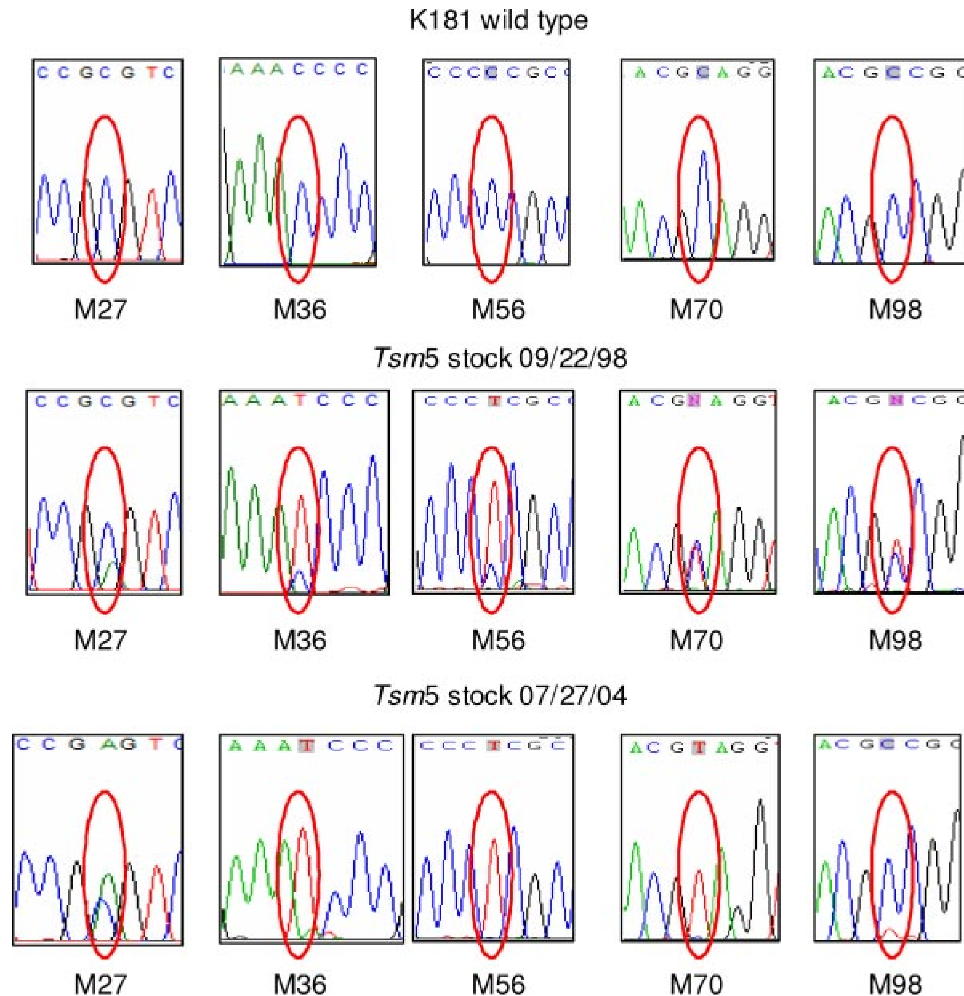

Fig. 2. Sequences of five ORFs in two *tsm*5 virus stocks produced 09/22/98 and 07/27/04 and in the wt K181 (Birmingham) variant. The presence of the *wt* nucleotide (C) for ORFs M27, M36, M56, M70, and M98 or mutant nucleotides A for ORF M27 or T for ORFs M36, M56, M70, and M98 are indicated in the ovals.

the exceptions being the G to A mutation in ORF M53 and the G to T mutation in m139, but in both the much more dominant peak was the mutant nucleotide (data not shown). This confirms that *tsm*5 is a mixture of mutants and not contaminated with *wt* virus. Similarly, passage of *tsm*5 eleven times in tissue culture did not yield *wt* virus but a mixture of mutants, the relative contribution of the *wt* and mutant nucleotide varying with passage as with virus stocks (data not shown).

DISCUSSION

The CGS method appears to be a useful, cost effective approach to identify quickly mutations in a relatively large genome like a herpes virus where it would be costly to identify mutations by conventional sequencing and where SNPs may go unrecognized. False positive SNPs are reported to occur at a frequency of 1 in 100,000–1,000,000 nucleotides sequenced (NimbleGen Systems, Inc.). In the present study false positive non-synonymous SNPs occurred at the rate of 4 per 100,000 nucleotides (i.e., 1 per 23,928 bp sequenced) in both K181 (Birmingham) and *tsm*5. In the context of this study these false positive SNPs were easily recognized by sequencing. Of more concern is the lack of sensitivity, that is, the large number (9 in 23,928 bp sequenced) of known SNPs that were not recognized in the K181 (Birmingham) variant. The reasons for this are unclear as a re-examination of the arrays showed no evidence of these SNPs. A similar error frequency may be observed by 454 sequencing [Moore et al., 2006; Roche, personal communication]; however, this may be improved by increasing the depth of coverage although this would increase costs.

Interestingly, the non-synonymous T to C (D309G) SNP identified in M54 ORF of K181 (Birmingham) was identified in a previous sequence of the M54 ORF of the K181 (Perth) variant [Scott et al., 2005] but not in any natural isolate (Fig. 3). Thus, it appears that some polymorphic alleles may vary with virus stock. This was evident also in the m144 SNPs where the synonymous SNPs are conserved in several MCMV isolates whilst the

```
K181P'th   AFPAAENVGDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
K181P'thR  AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
K181B'ham  AFPAAENVGDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
tsm5       AFPAAENVGDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
Smith      AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
WP15B      AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
K17A       AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
N1         AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
C4A        AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
G4         AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
W2         AFPAAENVDDIIIQISCVCFGVGEMVSHAYDVHADLSTPAVPENHLFTIGPCAPIPDTKI 360
W5         AFPAAENVDDIIIQISCVCFGVGEMVSHAYDVHADLSTPAVPENHLFTIGPCAPIPDTKI 360
W9         AFPAAENVDDIIIQISCVCFGVGEMVSHAYDVHADLSTPAVPENHLFTIGPCAPIPDTKI 360
G3A        AFPAAENVDDIIIQISCVCFGVGEMVSHAYDVHADLSTPAVPENHLFTIGPCAPIPDTKI 360
G2         AFPAAENVDDIIIQISCVCFGVGEMVSHAYDVHADLSTPAVPENHLFTIGPCAPIPDTKI 360
W8         AFPAAENVDDIIIQISCVCFGVGEMVHHAYDVHADLSTPAVPENHLFTIGPCAPIPDVKI 360
K17G       AFPAAENVDDIIIQISCVCFGVGEMVSHAYDVHADLSTPAVPENHLFTIGPCAPIPDTKI 360
           ********.***************** ******************************.**
```

Fig. 3. Clustal W alignment of M54 amino acid sequences from several isolates of murine cytomegalovirus. The boxed amino acid indicates the non-synonymous D309G SNP identified in K181 (Birmingham) and K181 (Perth) as indicated in EMBL Acc. No. (AAS47480) [Scott et al., 2005] compared to K181 (PerthR) (EMBL Accession Number AM886412) [Smith et al., 2008], tsm5 [Sweet et al., 2007], MCMV Smith strain (Smith) (Acc. No. U68299) and various Australian isolates WP15B (Acc. No. ACE95404), K17A (Acc. No. AAS47478), N1 (Acc. No. AAS47481), C4A (Acc. No. ACE95568), G4 (Acc. No. ACE95421), W2 (Acc. No. AAS47482), W5 (Acc. No. AAS47483), W9 (Acc. No. AAS47485), G3A (Acc. No. AAS47477), G2 (Acc. No. AAS47476), W8 (Acc. No. AAS47484), and K17G (Acc. No. AAS47479).

non-synonymous SNPs show the same amino acid variations in different isolates suggesting they could be hot spots for mutation (Fig. 4). Thus, a possible explanation for the relatively poor specificity might be nucleotide variation in hot spots for change between different stocks of the virus.

Nevertheless, identification of additional candidate mutations in *tsm*5 is a considerable advance although the significance of these newly identified mutations (Table I) remains to be determined. Little information is available on ORFs *m20*, *M25*, or *M47* other than that they are tegument proteins found in virions [Katterhorn et al., 2004]. Interruption of *M25* by transposon mutagenesis demonstrated its non-essential function both in vitro and in vivo [Zhan et al., 2000] while gene knockout of *m143* demonstrated that it was an essential gene [Menard et al., 2003]. Disruption of m141 and m139 severely limits virus replication in certain macrophage cell lines but not in MEFs or endothelial cells [Menard et al., 2003]. M53 is an essential protein that binds to M50 to form the nuclear egress complex, which forms at the inner nuclear membrane and mediates primary envelopment and nuclear egress of nucleocapsids [Lotzerich et al., 2006]. There are several important conserved regions in the C-terminal two thirds of the protein of herpesvirus M53 homologues but V273 is not conserved. As *tsm*5 lacks ability to replicate to detectable levels in immunocompetent mice, candidates to examine initially include the C to T mutation at residue 22,698 which leads to a non-synonymous mutation in two ORFs, m20 and m21; and the mutation in m139 which produces a stop codon resulting in a carboxy terminal truncated protein of 564 amino acids instead of the *wt* 644 amino acid protein. MCMV specific genes are believed to have important roles in vivo.

Interestingly, all 27 identified SNPs were point mutations with single base substitutions and no deletions, frame shifts or insertions. Four were transversion mutations while the remaining twenty-three were transition mutations, 16 of which were G:C to A:T mutations. Nitrosoguanidine is an alkylating agent and exerts its mutagenic effects by methylation of DNA. While spontaneous mutations are biased towards transitions recent characterization of NTG mutations in a *Corynebacterium glutamicum* strain also showed that 47 of 50 mutations were G:C to A:T transitions, the remaining three being A:T to G:C transitions with no transversions [Ohnishi et al., 2008]. This suggests that most, if not all, of the mutations identified in *tsm*5 resulted from the mutagenic activity of NTG.

That *tsm*5 is a mixture of mutants was initially surprising as originally the K181 (Birmingham) variant was cloned twice at limiting dilution prior to mutagenesis with NTG and mutagenized virus was ultrasonicated and filtered before examining for *ts* virus. Furthermore, mutants were then twice plaque picked under agar following temperature shift from 33°C to 39°C before stocks were produced, the temperature shift being used to enhance the likelihood of *ts* virus being selected [Sammons and Sweet, 1989]. This might suggest that reversion to the *wt* nucleotide has occurred subsequently in the production of virus stocks and that there is significant evolutionary pressure to change. However, this cannot be proven as the oldest virus stock still available produced in 1990 already was pleomorphic at these alleles (data not shown). It is more likely that the original mutagenized *tsm*5 virus stock was a mixture of mutants. Several isolates of human cytomegalovirus have proved to be phenotypically and genotypically heterogeneous and difficult to clone even after several passages [Dolan et al., 2004; Sinzger et al., 2008]. Furthermore, MCMV is well known to produce multicapsid virions [Hudson et al., 1976; Weiland et al., 1986]. BAC cloning may be a suitable method of

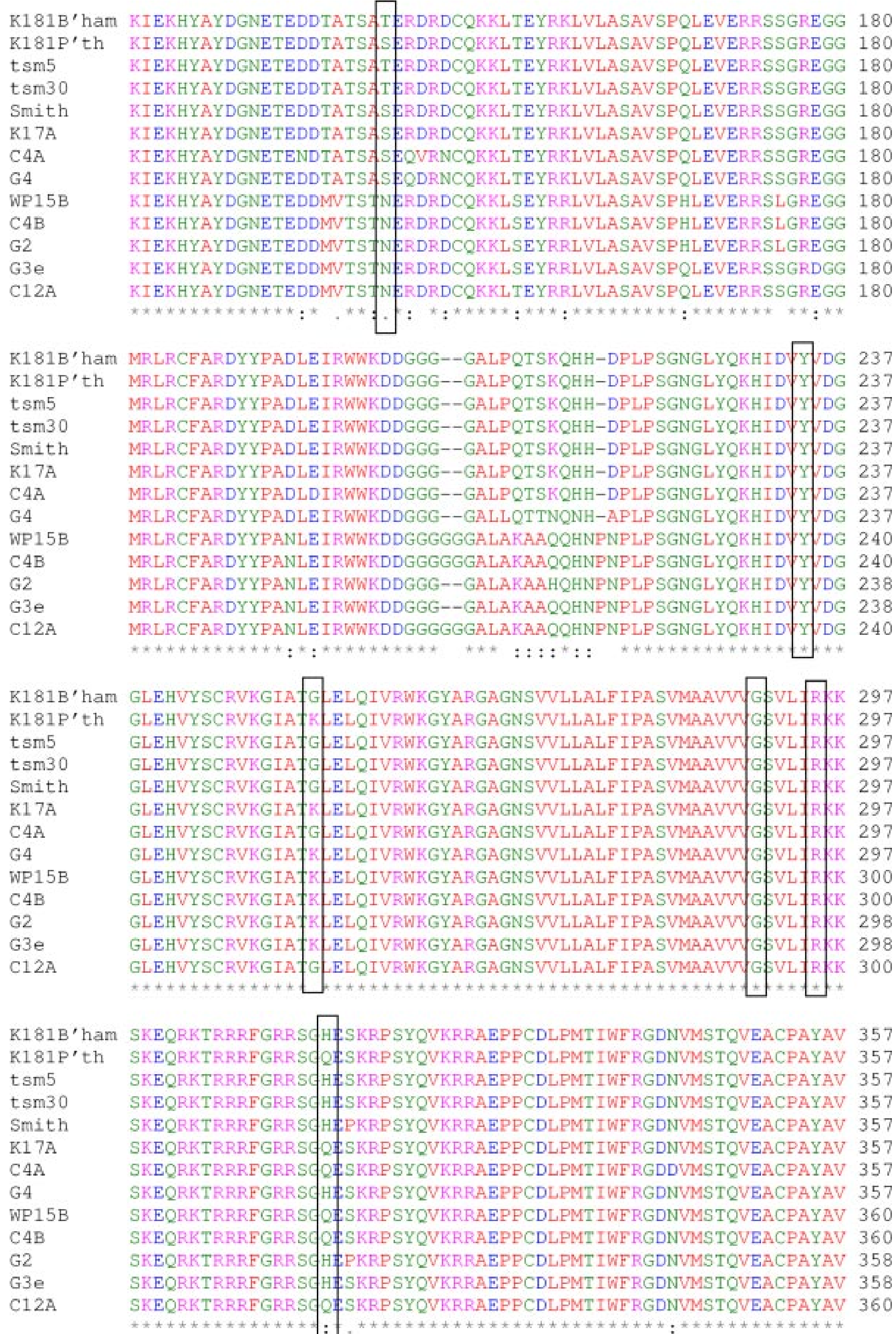

Fig. 4. Clustal W alignment of m144 amino acid sequences from several isolates of murine cytomegalovirus. Amino acids in boxes indicate non-synonymous 142ST or N, K253G and Q314H and synonymous Y234Y, G290G, and R295R SNPs identified in K181 (Birmingham) (K181B'ham) compared to K181 (Perth) (EMBL Accession Number AM886412), tsm5 [Sweet et al., 2007], tsm30 [Sweet et al., 2007], MCMV Smith strain (Smith) (Acc. No. U68299), and various Australian isolates K17A (Acc. No. AM236120), C4A (Acc. No. AM236115), G4 (Acc. No. AM236118), WP15B (Acc. No. AM236116), C4B (Acc. No. AM236121), G2 (Acc. No. AM236114), G3e (Acc. No. AM236119), and C12A (Acc. No. AM236117).

generating genetically defined virus clones as used for human cytomegalovirus isolate TB40/E [Sinzger et al., 2008]. CGS would then be a useful method for rapidly identifying the differences between the clones. Nevertheless, interaction between mutations appears to maintain mutant populations as in none of these stocks produced over 18 years has *wt* virus emerged, perhaps because there are too many mutations. This allelic

polymorphism is a feature of the current varicella-zoster vaccine [Quinlivan et al., 2006, 2007] and could make an interesting vaccine approach for cytomegalovirus.

Thus, in conclusion CGS has proved useful for identifying single-point mutations in a temperature-sensitive mutant but it also identified false positive SNPs and, more importantly, it failed to recognize some SNPs and thus cannot be relied upon to identify all possible mutations present.

ACKNOWLEDGMENTS

We thank Jigar Patel at Roche NimbleGen for providing details of the CGS method.

REFERENCES

Albert TJ, Dailidiene D, Dailide G, Norton J, Kalia A, Richmond TA, Molla M, Singh J, Green RD, Berg DE. 2005. Mutation discovery in bacterial genomes: Metronidazole resistance in Helicobacter pylori. Nature Meth 2:951–953.

Dolan A, Cunningham C, Hector RD, Hassan-Walker AF, Lee L, Addison C, Dargan DJ, McGeoch DJ, Gatherer D, Emery VC, Griffiths PD, Sinzger C, McSharry BP, Wilkinson GWG, Davison AJ. 2004. Genetic content of wild-type human cytomegalovirus. J Gen Virol 85:1301–1312.

Gresham D, Dunham MJ, Botstein D. 2008. Comparing whole genomes using DNA microarrays. Nat Rev Genet 9:291–302.

Herring CD, Palsson BØ. 2007. An evaluation of comparative genome sequencing (CGS) by comparing two previously-sequenced bacterial genomes. BMC Genomics 8:274.

Hudson JB, Misra V, Mosmann TR. 1976. Properties of the multicapsid virions of murine cytomegalovirus. Virology 72:224–234.

Katterhorn LM, Mills R, Wagner M, Lomsadze A, Makeev V, Boradovsky M, Ploegh HL, Kessler BM. 2004. Identification of proteins associated with murine cytomegalovirus virions. J Virol 78:11187–11197.

Lotzerich M, Ruszics Z, Koszinowski UH. 2006. Functional domains of murine cytomegalovirus nuclear egress protein M53/p38. J Virol 80:73–84.

Menard C, Wagner M, Ruzsics Z, Holak K, Brune W, Campbell AE, Koszinowski UH. 2003. Role of murine cytomegalovirus US22 gene family members in replication in macrophages. J Virol 77:5557–5570.

Moore MJ, Dhingra A, Soltis PS, Shaw R, Farmerie WG. 2006. Rapid and accurate pyrosequencing of angiosperm plastid genomes. BMC Plant Biol 6:17. 10.1186/1471–2229–6–17.

Morley PJ, Ertl P, Sweet C. 2002. Immunisation of Balb/c mice with severely attenuated murine cytomegalovirus (MCMV) mutants induces protective cellular and humoral immunity. J Med Virol 67: 187–199.

Ohnishi J, Mizoguchi H, Takeno S, Ikeda M. 2008. Characterization of mutations induced by N-methyl-N'-nitro-N-nitrosoguanidine in an industrial *Corynebacterium glutamicum* strain. Mutat Res 649: 239–244.

Quinlivan MA, Gershon AA, Nichols RA, LaRussa P, Steinberg SP, Breuer J. 2006. Vaccine Oka Varicella-Zoster genotypes are monomorphic in single vesicles and polymorphic in respiratory tract secretions. J Infect Dis 193:927–930.

Quinlivan MA, Gershon AA, Al Bassam MM, Steinberg SP, LaRussa P, Nichols RA, Breuer J. 2007. Natural selection for rash-forming genotypes of the varicella-zoster vaccine virus detected within immunized human hosts. Proc Nat Acad Sci USA 104:208–212.

Redwood AJ, Messerle M, Harvey NL, Hardy CM, Koszinowski UH, Lawson MA, Shellam GR. 2005. Use of a murine cytomegalovirus K181-derived bacterial artificial chromosome as a vaccine vector for immunocontraception. J Virol 79:2998–3008.

Sammons CC, Sweet C. 1989. Isolation and preliminary characterization of temperature-sensitive mutants of mouse cytomegalovirus of differing virulence for 1-week-old mice. J Gen Virol 70:2373–2381.

Scott GM, Ng HL, Morton CJ, Parker MW, Rawlinson WD. 2005. Murine cytomegalovirus resistant to antivirals has genetic correlates with human cytomegalovirus. J Gen Virol 86:2141–2151.

Sinzger C, Hahn G, Digel M, Katona R, Sampaio KL, Messerle M, Hengel H, Koszinowski UH, Brune W, Adler B. 2008. Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E. J Gen Virol 89:359–368.

Smith LM, McWhorter AC, Masters LL, Shellam GR, Redwood AJ. 2008. Laboratory strains of murine cytomegalovirus are genetically similar but phenotypically distinct from wild type strains. J Virol 82:6689–6696.

Sweet C, Ball K, Morley PJ, Guilfoyle K, Kirby M. 2007. Mutations in the temperature-sensitive murine cytomegalovirus (MCMV) mutants tsm5 and tsm30: A study of genes involved in immune evasion, DNA packaging and processing, and DNA replication. J Med Virol 79:285–299.

Weiland F, Keil GM, Reddehase MJ, Koszinowski UH. 1986. Studies on the morphogenesis of murine cytomegalovirus. Intervirology 26:192–201.

Wong CW, Albert TJ, Vega VB, Norton JE, Cutlet DJ, Richmond TA, Stanton LW, Liu ET, Miller LD. 2004. Tracking the evolution of the SARS coronavirus using high-throughput, high density resequencing arrays. Genome Res 14:398–405.

Zhan X, Lee M, Abenes G, Van Reis I, Kittinunvorakoon C, Ross-Macdonald P, Snyder M, Liu F. 2000. Mutagenesis of murine cytomegalovirus using a Tn3-based transposon. Virol 266:264–274.