JVI Accepts, published online ahead of print on 25 September 2013
J. Virol. doi:10.1128/JVI.02060-13
Copyright © 2013, American Society for Microbiology. All Rights Reserved.

JVI Accepts published online ahead of print

1 # Genome scale evolution of myxoma virus (MYXV) reveals
2 host-pathogen adaptation and rapid geographic spread

3
4
5 Peter J. Kerr[1], Matthew B. Rogers[2,3], Adam Fitch[2], Jay V. DePasse[2#], Isabella M. Cattadori[4],
6 Alan C. Twaddle[2], Peter J. Hudson[4], David C. Tscharke[5], Andrew F. Read[4], Edward C.
7 Holmes[6,7], Elodie Ghedin[2,3*]
8
9
10 [1] CSIRO Ecosystem Sciences, Canberra, Australian Capital Territory, Australia.
11 [2] Center for Vaccine Research, University of Pittsburgh School of Medicine, Pittsburgh,
12 Pennsylvania, USA.
13 [3] Department of Computational and Systems Biology, University of Pittsburgh School of
14 Medicine, Pittsburgh, Pennsylvania, USA.
15 [4] Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State
16 University, University Park, Pennsylvania, USA.
17 [5] Research School of Biology, The Australian National University, Canberra, Australian
18 Capital Territory, Australia.
19 [6] Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Biological
20 Sciences and Sydney Medical School, The University of Sydney, Sydney, Australia.
21 [7] Fogarty International Center, National Institutes of Health, Bethesda, Maryland, USA.
22
23 *Corresponding author: Dr Elodie Ghedin, Department of Computational and Systems
24 Biology and Center for Vaccine Research, BST3, room 9048, 3501 Fifth Avenue, University
25 of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, 15261. USA.
26
27 # Present address: Jay V. DePasse, depasse@psc.edu, Pittsburgh Supercomputing Center,
28 Pittsburgh, PA
29
30 **Running Title:** Genome scale evolution of myxoma virus
31
32 **Key words:** Myxoma virus, poxvirus, genome, host-pathogen coevolution, rabbit
33

JVI Accepts published online ahead of print

34 **ABSTRACT**

35 The evolutionary interplay between myxoma virus (MYXV) and the European rabbit
36 (*Oryctolagus cuniculus*) following release of the virus in Australia in 1950 as a biological
37 control is a classic example of host-pathogen coevolution. We present a detailed genomic
38 and phylogeographic analysis of 30 strains of MYXV, including the Australian progenitor
39 strain SLS, 24 Australian viruses isolated from 1951 to 1999, and three isolates from the
40 early radiation in Britain from 1954 and 1955. We show that in Australia MYXV has spread
41 rapidly on a spatial scale, with multiple lineages co-circulating within individual localities, and
42 that both highly virulent and attenuated viruses were still present in the field through the
43 1990s. In addition, the detection of closely related virus lineages at sites 1000 km apart
44 suggests that MYXV moves freely in geographic space with mosquitoes, fleas, and rabbit
45 migration all providing means of transport. Strikingly, despite multiple introductions, all
46 modern viruses appear to be ultimately derived from the original introductions of SLS. The
47 rapidity of MYXV evolution was also apparent at the genomic scale, with gene duplications
48 documented in a number of viruses. Duplication of potential virulence genes may be
49 important in increasing the expression of virulence proteins, and provides the basis for the
50 evolution of novel functions. Mutations leading to loss of open reading frames were
51 surprisingly frequent and in some cases may explain attenuation, but no common mutations
52 that correlated with virulence or attenuation were identified.

53

54

JVI Accepts published online ahead of print

55 **INTRODUCTION**

56 The experimental introduction of myxoma virus (MYXV) into the European rabbit 57 (*Oryctolagus cuniculus*) population of Australia and its unprecedented and unanticipated 58 spread initiated one of the great natural experiments in evolution (15). The subsequent 59 emergence of slightly attenuated viruses that were more efficiently transmitted, and the 60 natural selection of rabbits with genetic resistance to MYXV, were carefully documented in 61 real time (17). Sixty years later these studies continue to inform theory and practice in host-62 parasite co-evolution and particularly the complex relationship between virulence and 63 transmissibility.

64 MYXV is a poxvirus and the type species of the *Leporipoxvirus* genus. MYXV is 65 native to South America where its natural host is the tapeti ('forest rabbit'; *Sylvilagus* 66 *brasiliensis*) in which the virus causes a largely innocuous, localized, cutaneous fibroma. 67 MYXV is transmitted by mosquitoes or other biting arthropods probing through the fibroma 68 and picking up virus on their mouthparts. Transmission is passive as MYXV does not 69 replicate in the vector. In European rabbits, which are not native to the Americas, MYXV 70 causes the generalized lethal disease myxomatosis. As such, this represents a classic 71 example of a pathogen that is highly virulent in a new host species with no evolutionary 72 history of adaptation to that pathogen. Viruses closely related to MYXV are found in *S.* 73 *bachmani* (brush rabbit) on the west coast of the USA and the Baja peninsula of Mexico 74 (Californian myxoma viruses), and in *S. floridanus* (eastern cottontail) in eastern and central 75 parts of North America (rabbit fibroma virus; RFV) (17).

76 European rabbits were introduced into Australia with European settlement in 1788, 77 but the continent-wide spread of rabbits was initiated in 1859 by the introduction of 18-24 78 wild rabbits for hunting. Within 50 years these rabbits had spread over most of Australia with 79 the exception of the wet tropics and the far north (51). The European rabbit became 80 Australia's worst vertebrate pest responsible for enormous ecological destruction and 81 agricultural losses. Field trials in 1950 to assess MYXV as a biological control resulted in the 82 mosquito-driven epizootic spread of the virus throughout much of south-eastern Australia in 83 the summer of 1950/51, and it re-emerged the following spring (50). Assisted by large-scale 84 inoculation campaigns, MYXV spread and established over the rabbit-infested areas of 85 Australia during the next five years (17).

86 The MYXV introduced into Australia, termed the Standard Laboratory Strain (SLS), 87 was derived from an isolate made in Brazil, probably in 1910 (17, 40) and subsequently 88 maintained by rabbit passage. Importantly, the original virus used to initiate the epizootic 89 was available to serve as a reference for subsequent field isolates. SLS had a case-fatality 90 rate estimated at 99.8% in infected wild rabbits and similar lethality in laboratory rabbits, 91 which are domestic breeds of *Oryctolagus cuniculus*.

JVI Accepts published online ahead of print

92 It quickly became apparent that viruses with slightly lower case fatality rates were
93 emerging in the field and outcompeting ongoing releases of the virulent SLS (36, 37, 42).
94 Fenner and Marshall (16) classified the virulence of MYXV into 5 grades based on average
95 survival times, case fatality rates and symptomatology of groups of 4-6 laboratory rabbits
96 infected with very low doses of virus. The predominant viruses in the field were of grade 3
97 virulence (case fatality rates of 70-95%), and with average survival times that were
98 prolonged compared to SLS (17-28 days versus <13 days). Mosquito transmission is a
99 function of the titres of virus in the skin lesions induced by the virus and the time for which
100 the rabbit survives. By allowing the infected rabbit to survive for longer with high titres of
101 virus, the moderately attenuated viruses had a selection advantage over more virulent
102 strains. Highly attenuated grade 5 viruses (<50% case fatality rates) tended to be poorly
103 transmitted because the infected rabbits controlled virus replication, in turn reducing
104 transmissibility (14). Importantly, the emergence of more attenuated virus strains may have
105 facilitated the rapid selection of rabbits with genetic resistance to MYXV (17).

106 A separate strain of MYXV was released in France in 1952; the virus was obtained
107 from the Laboratory of Bacteriology in Lausanne, Switzerland and has hence been termed
108 the Lausanne strain (Lu), although like SLS it was originally isolated in Brazil (in Campinas in
109 1949). Unlike SLS, Lu had undergone relatively few rabbit passages. Lu and SLS have
110 indistinguishable virulence in laboratory rabbits; however, Lu is of considerably higher
111 virulence than SLS when tested in genetically resistant rabbits. Despite the differences in
112 starting virus, environmental conditions and insect vectors, the outcome of MYXV-rabbit co-
113 evolution in Europe was remarkably similar to Australia with the emergence of attenuated
114 viruses and the selection of rabbits with genetic resistance (25).

115 The Lu strain of MYXV is considered the reference genome. It has a double-stranded
116 (ds) DNA genome of 161,777 bp with inverted terminal repeats (TIR) 11,577 bp in length. It
117 encodes 158 unique open reading frames, 12 of which are duplicated in the TIRs. Genes
118 located toward the centre of the genome tend to be conserved between poxviruses and are
119 essential for replication and structure, whereas those toward the termini tend to be involved
120 in subversion of the host-immune response or have host-range functions and are less
121 conserved across poxviruses (6).

122 We have recently outlined the evolutionary patterns and dynamics of the Australian
123 progenitor SLS virus and 19 Australian isolates sampled between 1951 to 1999, as well as
124 two isolates of grade 1 and grade 5 virulence from the early radiation of MYXV in the UK
125 following the introduction of MYXV there in 1953 (26). To reveal the genetic basis to the
126 phenotypic differences between these viruses, and particularly their profound differences in
127 virulence, we report here the detailed genome sequences of these viruses plus an additional
128 five Australian viruses. In addition, we sequenced and analysed a second strain of KM13

JVI Accepts published online ahead of print

129 (KM13 2A) and the Lu virus strain produced by the Commonwealth Serum Laboratories for
130 release in Australia, as well as a grade 3 virus isolated in the UK in 1954. Such a rich
131 genomic data set enabled us to obtain a more detailed picture of the evolution and
132 geographic spread of this virus through Australia and particularly the broad range of genes
133 involved in this evolutionary process.
134
135 **METHODS**

136 *Virus isolates*
137 The isolates of MYXV used in this study are described in **Table 1**.
138
139 *Preparation of DNA*
140 Viruses were passaged twice in RK13 cells to prepare working stocks; viral DNA was
141 prepared from infected RK13 cells as previously described (26).
142
143 *Sequencing, assembly and comparative analyses*
144 The seven virus samples newly reported here were sequenced on the Illumina HiSeq 2000
145 platform. Demultiplexed and trimmed sequence reads were assembled with the Velvet de-
146 novo assembler (58) using a range of k-mer values from 59 to 77, and an expected
147 coverage of 600x. Contigs containing MYXV genomic DNA were identified by BlastX
148 searches and were ordered into a single scaffold against the Lu genome (accession #
149 AF170726) using the Abacas.pl script (3). The quality of each scaffold was verified by
150 remapping the untrimmed reads to the assembly using Smalt
151 (www.sanger.ac.uk/resources/software/smalt/); the resulting BAM files were converted to
152 pileup format to verify the read coverage at each site. Read coverage line plots for scaffolds
153 at each k-mer value were generated in R and examined by eye. In general, we found that
154 scaffolds generated at high k-mers (greater than 65) resulted in single contig assemblies of
155 the MYXV genomes, but inspection of coverage plots revealed many low coverage regions.
156 Further examination of these low coverage areas revealed that these were large insertions
157 unique to the strain in question when compared to the 23 previously sequenced strains of
158 MYXV (26). Assemblies at lower k-mer values (51-65) were often fragmented into multiple
159 contigs, but showed even read coverage across contigs corresponding to MYXV segments.
160 Further, these were of expected lengths relative to the 23 previously sequenced strains (26).
161 Gaps, SNPs and indels of interest were closed by Sanger sequencing of PCR products. In
162 every case, only one complete, or near complete, copy of the terminal inverted repeat (TIR)
163 was assembled at either the 5' or the 3' end, though up to a full read length of the
164 complementary TIR was observed at the opposite end, allowing easy identification of the TIR

JVI Accepts published online ahead of print

165 junction. To further verify the position of the TIR junction, we duplicated the complete TIR,
166 generated a reverse-complement of the sequence that was added on the opposite end, and
167 re-mapped the sequence reads to that assembled portion of the genome.
168 Genome annotation was transferred from the Lu strain to the newly sequenced
169 MYXV genomes using the Rapid Annotation Transfer Tool (45). EMBL flatfiles of transferred
170 gene models were then inspected and compared to Lu using the Artemis Comparison Tool
171 (8); incorrect models were corrected, and new gene models added where transfer had not
172 occurred. Genes are numbered based on their location in the MYXV genome, with the
173 direction of transcription indicated by L or R (e.g. *M010L*). Genes in the TIR are identified by
174 L/R (e.g. *M007L/R*). Proteins are identified by the same number as the gene with the
175 transcription direction omitted i.e. M010.
176 To generate the heatmaps for the comparative analyses of each gene to the SLS or
177 Lu strains, we used a custom perl script to produce multi-fasta files containing all taxa in
178 which this gene was present. Sequence alignments were generated using ClustalW (32),
179 and PAUP* 4.0b10 (54) was used to remove ambiguous and gapped sites from the
180 alignments and generate the number of SNP mutations in each gene. Columns from the
181 distance matrix comparing viral taxa to SLS were parsed, and two subsequent matrices were
182 generated, one for European strains compared to Lu, and one for Australian strains
183 compared to SLS.
184
185 *Evolutionary analysis*
186 A total of 30 genome sequences of MYXV were subjected to phylogenetic analysis, with a
187 total alignment length of 163,555 nt. Sequences were aligned by MAFFT (23), then
188 inspected by eye. Phylogenetic analysis employed the maximum likelihood (ML) method
189 available in PhyML 3.0 (22). Because of the very low numbers of substitutions separating
190 these sequences we employed the HKY85 model of nucleotide substitution (46) with subtree
191 pruning and regrafting (i.e. SPR) branch-swapping. To assess the robustness of each node
192 on the tree, a bootstrap resampling analysis was undertaken (1000 replicates) employing the
193 parameters described above. To determine whether these 30 MYXV genomes contain any
194 recombinant regions we utilized the RDP, GENECOV and BOOTSCAN methods available
195 within the RDP4 package (38) and assuming the default parameters. As with our previous
196 study (26) no recombination was observed.
197 To estimate the rates of evolutionary change and times to common ancestry in these
198 data (including those of two key nodes shown on **Fig. 1**) we employed the Bayesian Markov
199 chain Monte Carlo (MCMC) method available in the BEAST package (10). This analysis
200 utilized both strict and relaxed (uncorrelated lognormal) molecular clocks, a Bayesian skyline
201 coalescent prior, and the HKY85 nucleotide substitution mode. The MCMC was run for 100

JVI Accepts published online ahead of print

202 million generations and convergence was observed in all parameters. Statistical uncertainly
203 is presented as values for the 95% highest probability density (HPD).
204
205 **RESULTS**
206
207 *Evolution and phylogeography of MYXV*
208 Our phylogenetic analysis of 30 complete MYXV genomes, including five new Australian
209 isolates sampled during 1993-1999 and an early attenuated isolate from the UK sampled in
210 1954, depicted the major division between the Australian and European epidemics as
211 observed previously (**Fig. 1**; (26)) and with no evidence of recombination. In addition, that all
212 the recently sampled Australian viruses (1991-1999) are clearly distinct from both SLS and
213 Lu indicates that these two viruses made no significant contribution to the later evolution of
214 MYXV in Australia even though they were introduced multiple times over many years.
215 Hence, these data suggest that all (sampled) Australian MYXV strains have their ancestry in
216 the initial introduction of SLS in 1950, although the close phylogenetic relationship among
217 the sequences means that we cannot determine whether the Glenfield (Gv) strain, which
218 was also widely released in NSW and Victoria, made any contribution to the spread of
219 MYXV. Our estimates of rates of nucleotide substitution—at 0.8-1.1 x 10^{-5} nucleotide
220 substitutions per site, per year (95% HPD values)—and times to common ancestry were also
221 essentially identical to those observed previously (26). Hence, these data again indicate that
222 the evolution of MYXV is both relatively rapid (for a dsDNA virus) and remarkably clock-like.
223 A visual overview of genome-scale genetic variation, manifest as the genetic
224 distance of each gene from the progenitor strain—SLS for the Australian isolates and Lu for
225 the European isolates—is represented by heatmaps (**Fig. 2A** and **2B**, respectively). These
226 maps reveal that the majority of genes remain highly conserved, with a few genes exhibiting
227 more diversity. An example of the latter is *M017L*. Although the function of this gene is
228 unknown, it has acquired mutations in the majority of the Australian strains compared to SLS
229 (**Fig. 2A**, **Table 2**). Multiple genes (*M003.1L/R*, *M103L*, *M105L* and *M132L*) have acquired
230 mutations in OB3/1120/1996 and WS6/1071/1995, which are linked to the other MYXV
231 strains by a relatively long branch (**Fig. 1**). However, of these, only *M103L* encodes a protein
232 with a predicted function (structural membrane protein), while the majority of mutations
233 involved are either commonplace and/or synonymous ones exhibiting no clear association
234 with changing virulence. Similarly, with the exception of the attenuated Spanish isolate 6918,
235 which appears as genetically distant based on this and the phylogenetic analysis, the
236 European isolates have very few mutations as compared to Lu (**Fig. 2B**), reflecting their
237 sampling early in the epidemic.

JVI Accepts published online ahead of print

238 To reveal aspects of the phylogeography of MYXV we coded the Australian isolates
239 by their State of origin (**Fig. 1**), in which 'CD' delineates viruses that were sampled in close
240 proximity to each other (within 10-15 km) in the 'Canberra District' that straddles the
241 NSW/ACT border in South-eastern Australia (see below). Strikingly, BD23 and BD44
242 sampled from hot, dry rangelands at Bulloo Downs in South West Queensland in 1999, are
243 very closely related to viruses (OB2/W60/1995 and SWH/1209/1996) sampled 3-4 years
244 earlier from the cool climate, higher rainfall Canberra district approximately 1000 km away.
245 Also of interest is the Meby strain sampled from Tasmania, and which is separated from
246 mainland Australia by the Bass Strait, which is up to 240 km wide. Although SLS was
247 released in Tasmania in the early 1950's following its spread on the mainland, Meby is
248 clearly descended from a mainland virus that diverged in the late 1960s and has then
249 remained isolated since this time (**Fig. 1**). It is therefore possible that the virus reached
250 Tasmania from the mainland on a mosquito inadvertently transported by ship or plane. The
251 majority (15) of the sequenced viruses were isolated between 1993 and 1996 from a set of
252 seven closely situated study sites (WS1, WS6, OB1, OB2, OB3, SWH, BRK) in the Canberra
253 district (27) (57). From the phylogenetic analysis (**Fig. 1**) it is obvious that viral lineages have
254 co-circulated at a single locality during a specific time period. In general, these results
255 highlight the relative rapidity of MYXV movement, likely aided by mosquito transmission,
256 including a dispersal of over 1000km during 1950.
257
258 *Comparison of the SLS and Lu sequences*
259 SLS was the original virus released in Australia in 1950. We compared the complete
260 genome sequence of SLS to the Lu strain. These two progenitor strains have differences in
261 symptomatology, virulence and passage history prior to release. Overall, there are 80
262 nucleotide differences (0.05% difference), including indels, between SLS and Lu (72 if TIRs
263 are only counted in one copy) (**Table 3**). However, frame-shifts in *M005L/R, M083L* and
264 *M152R* due to indels also produce multiple amino acid changes in SLS compared to the Lu
265 sequence: *M005L/R*, which codes for an E3 Ub ligase/apoptosis regulator, is disrupted by a
266 C insert at nucleotide 34. It is likely that translation is occurring from an alternative ATG from
267 nucleotide 17 that does not change the ANK repeats and the C-terminal F-Box domain of the
268 M005 protein. *M083L* is disrupted by a C deletion in a homopolymer tract at nucleotide 513.
269 M083 is homologous to rabbit carbonic anhydrase (6) and is probably a virion structural
270 protein. Finally, there is a T deletion in a homopolymer towards the 3' end of the *M152R*
271 (Serp 3) gene at nucleotide 782; read-through of the Lu stop codon leads to a predicted
272 protein of 273 amino acids in SLS rather than 266 amino acids in Lu. These indels are also
273 present in Australian isolates of MYXV sampled between 1951-1953, confirming that the
274 mutations were present in the progenitor virus. It is likely that one of these frame-shift

JVI Accepts published online ahead of print

275 mutations explains the reduced virulence of SLS compared to Lu.

276

277 *Comparisons of SLS with subsequent Australian isolates*

278 We sequenced three isolates of MYXV sampled within the first three years of the initial 279 epizootic of myxomatosis in Australia and which had been previously characterized in terms 280 of virulence (16): the Gv strain (Dubbo /Feb 1951; grade 1 virulence), KM13 (Corowa /Dec 281 1952; the prototype grade 3 virus) and Uriarra (Ur) (Uriarra/Feb 1953; grade 5 virulence 282 (28)). Amino acid sequence changes and gene disruptions between SLS and these three 283 viruses are summarized in **Table 4**.

284 Three of the nonsynonymous mutations in Gv, which is more virulent than SLS, are 285 in enzymes involved in viral transcription and replication—M044 (RNA helicase: R606H; the 286 R is conserved in chordopoxvirus sequences); M108 (DNA helicase: F18I; only MYXV and 287 RFV have F at this position, other chordopoxviruses have I, M or L); and M114 (RNA 288 polymerase: A686V; the A is completely conserved at this position in chordopoxvirus 289 sequences)—each of which could affect replication efficiency. *M014L*, *M130R* and *M153R* 290 all have single nucleotide indels that disrupt the reading frame. The single nucleotide indel in 291 *M014L* causes premature termination at residue 477 rather than the 517 amino acids of the 292 SLS protein. This indel is also present in Ur and KM13 indicating that the mutation arose 293 early on. M014 has an N-terminal BTB motif and C-terminal kelch motifs and is predicted to 294 form an E3 ubiquitin ligase complex that targets cellular protein(s) to the proteosome for 295 destruction (59). This truncation would delete the final kelch domain potentially altering 296 target protein recognition. However, the role of M014 in virulence is unknown as is that of 297 *M130R*. The large number of gene disruptions in Gv suggests that this virus may be a 298 variant selected during previous plaque purification (52) from which this virus was obtained. 299 The likely explanation for the attenuation of KM13 is the disruption to *M014L*. Although this 300 mutation is also present in the virulent Gv there is no other obvious mutation that might lead 301 to the attenuation of KM13. To further assist in the documentation of virulence determinants 302 we also sequenced a laboratory variant of KM13 reported to have a lower case fatality rate 303 (KM13 2A) (17, 35). The only difference between KM13 and KM13 2A appears to be an 304 extra A in non-coding sequence of KM13.

305 Also of note is that Ur has an extra C inserted after nucleotide 30 in *M005L/R*. This 306 means that the alternative ATG, which we predict to be used by SLS and all other Australian 307 viruses sequenced here, does not create a sense open reading frame (ORF) in Ur. The only 308 downstream ATG that is compatible with an ORF is at nt 308; translation from this ATG 309 would produce a 382 residue protein with 5 ANK repeats and the C-terminal F-box 310 compared to 7 ANK repeats in the 478 residue SLS protein. However, there is no convincing 311 promoter sequence upstream of this ATG. This is likely the main attenuating mutation in Ur.

JVI Accepts published online ahead of print

312 Ur also has an A insert in a homopolymer tract towards the 3' end of *M134R* at nucleotide
313 5911, the same location as the 3A insert in Gv and KM13 (**Table 4**). This leads to a
314 predicted truncated protein of 1973 amino acid residues rather than the 2000 residues in
315 SLS and Lu which retains the predicted C-terminal transmembrane domain that is conserved
316 across the chordopoxviridae.
317
318 *Recent Australian isolates*
319 In total, we determined sequences for 21 Australian viruses isolated between 1991 and
320 1999, six of which had been characterized by virulence assays: Bendigo, Wellington, BRK
321 (grade 1), SWH, Gung (grade 4) and Meby (grade 5) (53). All of these viruses have the C
322 insertion at 35 in *M005L/R* and the T deletion in *M152R* seen in SLS. However, the indel in
323 *M083L* present in SLS, Ur, Gv and KM13 has reverted in every Australian isolate sequenced
324 from the 1990s. Similarly, the indel disrupting the *M014L* gene found in Gv, Ur and KM13 is
325 not present in any of the more recent isolates. All the isolates have the 3A indel in *M134R*
326 seen in Gv and KM13. However, the underlying sequence reads that map to that genomic
327 region indicate that there is a subpopulation of viruses in OB3/1120 that have a 2A insertion,
328 rather than a 3A insertion; this 2A indel would lead to disruption of the *M134R* ORF. A
329 similar subpopulation with 2A insert was seen in Ur, which has a majority population with a
330 single A insert. Homopolymer sequences such as in *M134R* are common in MYXV, and poly
331 A or poly T tracts are common at the 3' end of genes and in intergenic sequence, where they
332 are frequently part of promoter structures for the downstream gene or the T_5NT early
333 transcription termination signal. In the Australian isolates, 13 of 16 single base indels that
334 occur in coding sequences (**Table 5**) occur in homopolymer tracts of 4 or more bases and
335 there are 17 positions with single base indels in intergenic homopolymers involving one or
336 more viruses (positions in TIRs have only been counted once). Polymerase slippage leading
337 to read-throughs or premature termination (e.g. SLS *M152R*) may facilitate evolutionary
338 plasticity, allowing slight changes in protein sequences. Indels either in homopolymers or
339 repeat sequence can also lead to gene disruption, in turn affecting virulence (1) (12), and
340 also function to repair ORFs, as in the case of *M083L*.
341 All but one of the recent isolates have a frameshift mutation due to a single
342 nucleotide insertion in a homopolymer tract in *M009L*, a member of a three-gene family
343 (*M006L/R*, *M008L/R*, *M009L*) (6) that are predicted to encode E3 ubiquitin ligases with N-
344 terminal BTB domains followed by kelch motifs (59). The insertion at nucleotide 420
345 produces a truncated protein of 146 rather than 509 residues. In addition, four viruses have
346 further mutations that disrupt the reading frame, and *M009L* is also disrupted in viruses that
347 have gene duplications from the right hand (RH) end of the genome (see below), implying
348 that this gene is non-essential. BRK has a 92 bp deletion in the *M036L* gene, which leads to

JVI Accepts published online ahead of print

349 a truncated protein of only 212 residues rather than 680 in SLS. The function of this gene
350 (an orthologue of VACV *O1L*) is unknown in MYXV, but in VACV the O1 protein enhances
351 signalling via Erk1/2 by the viral epidermal growth factor homologue (VGF) and increases
352 virulence (Scheneker et al 2012). As BRK is of grade 1 virulence (53), *M036L* is unlikely to
353 be crucial for virulence in this virus. In this respect, the attenuated UK isolates Sussex (1954;
354 grade 3) and Nottingham (1955; grade 5) also have a common indel that disrupts the *M036L*
355 ORF and the attenuated Spanish isolate 6918 has an independent disruption to *M036L* (39).
356 Both Nottingham and 6918 possess other mutations that explain attenuation. However, the
357 disruption to *M036L* is the only one in Sussex, suggesting that it may play some role in
358 virulence.

359 ORF-disrupting mutations were also common in *M153R* which encodes a protein with
360 an N-terminal RING-CH domain that is predicted to form an E3 Ub ligase complex and
361 downregulates MHC-1, CD4, ALCAM/CD166 and Fas/CD95 on the membrane of infected
362 cells potentially inhibiting $CD8^+$ T lymphocyte recognition and death signalling. Deletion of
363 this gene in the T1 Lu-derived strain reduced the case fatality rate from 100% to 30% (9, 21,
364 34). Meby, a grade 5 virus has a 73 bp deletion between repeat sequence blocks (AATACG)
365 in the C-terminal conserved region (CR) of *M153R* (9, 53) which leads to read-through of the
366 normal stop signal and a completely changed C-terminal protein sequence after residue 168
367 of the 206 amino acid protein. A single nt deletion at nt 469 in WS6/1071 and OB3/1120
368 leads to a stop after residue 161 and removes the CR region but retains the N-terminal
369 RING-CH domain and two putative transmembrane domains at 95-115 and 135-153 (9). Gv
370 and BD44 both have a truncated protein of 118 and 124 residues due to an independent
371 indel at nt 329. Other indels causing significant disruption to ORFs were only found in single
372 viruses: *M000.5L/R* (BD44; unknown function), *M008.1L/R* (BD44; secreted serine
373 proteinase inhibitor; virulence function), *M005L/R* (WS6/346; host-range; virulence function),
374 *M012L* (OB3/Y317; dUTPase) and *M147R* (BD23; Ser/Thr protein kinase) (**Table 5**). The
375 disruptions to *M008.1L/R* in BD44 and *M005L/R* in WS6/346 might be expected to attenuate
376 these viruses (33, 41).
377
378 *Gene duplications*
379 The inverted terminal repeat regions of poxviruses contain non-coding terminal regions
380 essential for replication but also encode different numbers of genes, depending on the
381 location of the TIR boundary. These genes typically have virulence or host-range functions
382 but the number of genes can vary greatly. This region also appears to be a potential
383 recombination hotspot, as shown by recombination and deletions including those in
384 malignant rabbit virus—a recombinant between MYXV and RFV (55)—or the MYXV SG33
385 vaccine strain (7).

JVI Accepts published online ahead of print

386 Two genes, *M154L* and *M156R,* which are normally found as single copies outside
387 the RH TIR with *M156R* overlapping the TIR boundary, have been duplicated at the left hand
388 (LH) TIR in the common ancestor of SWH/1209, OB2/W60, BD44 and BD23. In addition, the
389 *M153R* gene has been partially duplicated (**Fig. 3**). This duplication is essentially an
390 expansion of the TIR by 1635 bp, 36 nt downstream of the *M153R* ATG start codon,
391 meaning that *M153R* now overlaps the TIR junction (27). At the LH end of the genome this
392 duplicated sequence has replaced 923 bp of the *M009L* gene, leaving only the 5' 608 bp;
393 however, the *M009L* ORF is disrupted after amino acid 146 due to a T insert at nt 420. M156
394 is an orthologue of the VACV K3 protein and is predicted to inhibit the action of type 1 IFN
395 (49). M154 is an orthologue of VACV M2 and so may inhibit NFκB (6, 20). At the LH end,
396 *M153R* is lacking the 5' 36 nucleotides and upstream promoter, and the insert is not in frame
397 with the *M009L* sequence. Interestingly, this duplication was observed in the Canberra
398 region in viruses isolated in 1995 and 1996 and is also present in viruses isolated from SW
399 Queensland in 1999; that viruses with this deletion occupy such a wide geographic area
400 means that it is unlikely to have an adverse effect on fitness. Whether this duplication of two
401 potential virulence genes increases virulence or compensates for other mutations by
402 increasing expression of these proteins is not known.

403 Finally, two intergenic repeat sequence regions have been defined as being variable
404 in Australian field isolates (27, 53); one between *M017L* and *M018L*, with 2 to 8 extra copies
405 of a GTATGTAG repeat compared to SLS and 1 to 2 extra copies of an AGTTTAGT repeat
406 (**Fig. 4A**), and the other immediately upstream of the M002 gene in the TIR with 27 or 39
407 nucleotides deleted in 10 recent Australian isolates (**Fig. 4B**). That this latter duplication
408 occurs on multiple branches of the phylogenetic tree indicates that it has been gained or lost
409 in different viral lineages.

410

411 *Promoter sequences*

412 Alterations in gene expression and potentially virulence could occur due to changes in
413 promoters. The poxvirus early (E), intermediate (I) and late (L) promoter sequences are
414 conserved in the leporipoxviruses (56). Six viruses have mutations with a potential impact on
415 putative promoters. WS6/346 has a T deletion in the upstream T tract of the *M008.1L/R* L
416 promoter (**Fig. 5A**); OB1/406 has an extra T in the upstream T tract of the L promoter for
417 *M057L* (**Fig. 5B**). G91 has a mutation in a putative weak L promoter for *M00.5L/R*, but
418 whether this ORF is expressed has not been determined (**Fig. 5C**). SWH/8-2-93 has an
419 extra A inserted in the potential E promoter for *M138L* (**Fig. 5D**) which might be predicted to
420 enhance the promoter structure based on consensus early promoter sequences (56).
421 WS6/1071 and OB3/1120 have an A deleted in the 3' end of the potential E promoter of

JVI Accepts published online ahead of print

422 *M153R* (**Fig. 5E**) which could have an impact on the promoter activity. However, both
423 viruses also have a deletion at nt 321 in *M153R*, which disrupts the ORF.
424
425 *Pathways to attenuation and virulence in Australian isolates*
426 Overall, nine viruses derived from SLS and sequenced here have a previously defined
427 virulence phenotype. The coding changes from SLS in these viruses are summarized in
428 **Table 6**. Three attenuated viruses were sequenced from the 1990s, of which only the Grade
429 5 Meby has a probable explanation for its attenuated phenotype. Strikingly, the grade 4
430 SWH/9/1992 is closely related to the grade 1 BRK. Excluding the disruption of the *M036L*
431 ORF in BRK, only three coding differences exist to explain the attenuated phenotype of
432 SWH: P227S mutation at the C terminus of M004; a P33L mutation in M087 (mRNA capping
433 enzyme; P at this position in conserved in most poxviruses) and a VP duplication in the
434 M093 viral core protein. BRK has one unique mutation outside M036, A47V in M112 a
435 Holliday junction resolvase. The A in this position is not conserved outside the
436 leporipoxviruses. A similar analysis with the three grade 1 viruses sequenced revealed only
437 two shared mutations for two viruses: Bendigo and Wellington both share I481V in M032 and
438 Y302H in M099 the major core protein precursor. This suggests that attenuation and
439 virulence mutations may be subtle or involve multiple epistatic effects.
440 An example of the complexity of possible virulence determinants involves two genes
441 that encode proteins that are functionally conserved in poxviruses and inhibit type 1
442 interferon responses; *M029L*, an orthologue of VACV *E3L*, and *M156R*, an orthologue of
443 VACV *K3L* (**Fig. 1**). Eleven of the recent Australian isolates have a A17V mutation in *M029L*,
444 the only one in this gene; based on the molecular clock dating analysis this mutation was
445 fixed between 1969 and 1975 (**Fig. 1**), and coincides with the introduction and spread of the
446 European rabbit flea, which altered the epizootology of myxomatosis in temperate Australia.
447 Interestingly, this mutation has reverted to the original sequence twice on independent
448 branches of the tree, and in viruses isolated from widely separated geographic regions.
449 M029 has been shown to function similarly to VACV E3, binding double strand RNA,
450 inhibiting protein kinase R (PKR) activation and inhibiting IFN-β, TNF-α and IL-6 expression
451 (44); it is a critical virulence factor in rabbit infections but also has a second function binding
452 RNA helicase A (RHA/DHX9), which promotes virus replication in some cell lines and so has
453 been described as a dual function virulence and host-range factor (48).The amino acid
454 sequence from the related leporipoxvirus RFV is conserved from amino acid 1 to 8 and is
455 identical to MYXV from amino acid 20-55 but is poorly conserved between residues 9-19;
456 this region is also divergent in the Californian MSW strain of MYXV (31). If we consider the
457 mutation and reversion in MYXV and the divergence in 3 leporipoxviruses with different

JVI Accepts published online ahead of print

458 natural hosts, then it is possible that this region is involved in species-specificity and host
459 adaptation.
460 M156 is a homologue of the alpha subunit of eukaryotic translation initiation factor 2
461 (eIF2α) and competes with eIF2α for phosphorylation by PKR (49). The L98P mutation in
462 M156 is present in 13 of the modern isolates, and most likely has appeared twice on
463 independent branches of the tree (**Fig. 1**) (although a lack of bootstrap support at key nodes
464 means that we cannot formally exclude a single origin of this mutation). Two of the viruses
465 encoding this mutation, WS6/346 and OB3/Y317, which are phylogenetically distinct,
466 possess an additional T deletion in a homopolymer at the 3' end of M156R that allows read
467 through of the stop codon and the addition of EG at the C-terminus of the protein. Based on
468 NMR structure, the 102 amino acid M156 protein is predicted to be a 5-stranded antiparallel
469 β barrel (49). The L98P mutation occurs in the β5 sheet with the L side-chain predicted to
470 form part of the interior of the barrel, while residues D97 and R99 have been proposed to be
471 involved in binding PKR (49). Interestingly, *M156R* has been duplicated in the common
472 ancestor of four of the viruses sequenced here, all of which have the L98P mutation. In
473 VACV undergoing artificial selection, *K3L* has been shown to expand and reduce in copy
474 numbers while acquiring adaptive mutations (11).
475
476 *The virulent LuCSL virus released in Australia during 1970s to 1990s*
477 The Lu virus was widely released in Australia from the 1970s to the 1990s. However, our
478 phylogenetic analysis, coupled with previous studies (4, 27, 53), demonstrates that
479 Australian field isolates are derived from SLS and that if Lu has left descendent viruses in
480 Australia they have not been sampled. The Lu virus sequenced here is from a vial supplied
481 for release and has only a single difference from the originally published Lu sequence (6,
482 39)—a C insert at nucleotide 142 in a homopolymer tract in *M127L* causing a frame-shift
483 mutation. The mutation may have been present in a single plaque or pock used to produce
484 the original seed virus for release (47).
485
486 **DISCUSSION**
487 MYXV evolution is characterized by a relatively rapid rate of nucleotide substitution, frequent
488 changes of virulence, and a rapid spread in geographic space. This was exemplified by the
489 initial mosquito-borne epizootic in 1950/51, during which SLS spread across an area
490 approximately 1600 km south to north and 1800 km east to west in 3 months (50). Indeed,
491 our phylogeographic analysis clearly shows that viruses from geographically disjunct regions
492 of Australia can still be remarkably closely related, indicative of frequent viral traffic. The
493 success of MYXV and, subsequently, rabbit haemorrhagic disease virus (RHDV) as
494 biological controls, combined with changes in land management, means that modern rabbit

JVI Accepts published online ahead of print

495 populations are likely to be less connected than in 1950 (18, 19). The key vectors for viral
496 transmission are the mosquito, which is predominantly a spring to autumn vector and
497 requires water for breeding, and the rabbit fleas, *Spilopsyllus cuniculi* (in temperate
498 Australia) and *Xenopsyllus cunicularis* (in arid Australia), which were introduced into
499 Australia in 1970 and 1994, respectively. Fleas provide the potential for local transmission
500 year round, whereas mosquitoes are seasonal but have the potential for longer distance
501 spread. Virus may also be spread by dispersing migrating rabbits—predominantly juvenile
502 males—that are either incubating the disease or immune and carrying fleas with the virus. In
503 addition, large scale rabbit migrations out of dry country during droughts may bring high
504 numbers of susceptible animals into contact with virus providing opportunities for spread.
505 Accidental or deliberate translocation of infected rabbits could possibly also occur. Work in
506 the Canberra district also suggests that viral spread is rapid as shown by the multiple viral
507 lineages that can co-circulate within a single community, with no apparent dominance of one
508 lineage over any other. Such lineage co-circulation also tentatively suggests that these
509 viruses do not differ greatly in long-term fitness despite their possible differences in
510 virulence, although this will need to be confirmed with additional data. Indeed, in our analysis
511 as a whole, there was no obvious signal for major fitness differences across multiple
512 genotypes within a small geographic range.

513 The outcome of infection with MYXV depends on the interaction of multiple viral
514 immune evasion and immunosuppression proteins, and proteins and cells of the host innate
515 and adaptive immune system, together with the proteins required for virus replication,
516 assembly and infection. The emergence of slightly attenuated viruses during the early
517 radiation in Australia and Europe means that mutations were selected that enhanced
518 transmission because the infected rabbit survived longer than rabbits infected with grade 1
519 strains.

520 Most of these early (slightly) attenuated viruses still had case fatality rates of 90-99%,
521 but with prolonged survival times compared to SLS and Lu (43); (16, 36). Experimentally,
522 grade 4 viruses (case fatality rate of 50-70%; average survival times 29-50 days) had the
523 highest rates of mosquito transmission (14), but in field surveys from 1951-1981, these
524 viruses were always less prevalent than grade 3 viruses (13). The rapid selection of rabbits
525 with resistance to myxomatosis—which appears to operate through an enhanced innate
526 immune response rather than resistance to infection (5, 24)—is likely to have driven virus
527 evolution towards increased virulence, and hence to maintain transmissibility and
528 competitiveness, and this may explain the preponderance of grade 3 viruses since all these
529 virulence measurements were done in laboratory rabbits with no resistance. Viruses with a
530 grade 1 phenotype in laboratory rabbits appear as grade 4 or 5 in wild rabbits with genetic

JVI Accepts published online ahead of print

531 resistance, while some such as BRK are more virulent than the progenitor SLS when tested
532 in wild rabbits (29, 30).
533 The pathway to virulence reversion and enhancement could involve reversal of
534 attenuating mutations. For example, reversal of the indel in *M083L* has occurred in the
535 common ancestor to all the modern isolates we sequenced, while that in *M014L* was
536 common to all three early viruses sequenced, although whether this is a reversal is not clear.
537 Similarly, mutations could compensate for attenuating mutations, such as the disruption of
538 M036L, which appears to be attenuating in Sussex but not in BRK, or mutations could
539 increase virulence by new pathways. The duplication of virulence genes and the
540 fragmentation of some reading frames also provides the raw material for further evolution of
541 new functions as has occurred; for example, in cowpox virus where a gene fragment has
542 evolved a new function in immunosuppression (2). While this might suggest that field
543 isolates should now be of higher virulence for laboratory rabbits, the reality appears more
544 nuanced with grade 4 and 5 viruses present in our samples, indicating that many factors at
545 the local level influence the effective virulence and successful transmission. In addition, the
546 widespread establishment of the European rabbit flea in Australia, which was credited with
547 enhancing the impact of myxomatosis by providing a year-round vector and increasing
548 transmission, may have altered the selection pressures on both virus and rabbit.
549 The large and complex genome of MYXV has provided the plasticity for multiple
550 routes to attenuation and multiple and complex routes back to virulence. The accumulation
551 of mutations in more recent virus isolates makes it difficult to identify single mutations that
552 are critical for phenotype, whether virulent or attenuated. In particular, we have shown here
553 that it is difficult to define possible roles for single amino acid changes or potentially even
554 synonymous changes in this evolutionary process. Indeed, there has been remarkably little
555 characterization of Australian field viruses in rabbits or even in cell culture since the 1980s.
556 Importantly, characterization of the sequenced viruses in rabbits will provide opportunities for
557 matching virulent and attenuated viruses that are phylogenetically closely related and using
558 reverse genetics to define these pathways.
559
560
561 **DATA ACCESS**
562 The seven new Myxoma genome assemblies have been deposited on Genbank under
563 accession numbers KC660079-KC660085.
564
565 **ACKNOWLEDGEMENTS**

566 This work was funded in part by grant R01 AI093804 from the National Institute of Allergy
567 and Infectious Diseases, National Institutes of Health. ECH is funded by an NHMRC
568 Australia Fellowship. DCT is funded by an ARC Future Fellowship.

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

569 **REFERENCES**

570 1. **Alcami, A., and G. L. Smith.** 1996. A mechanism of the inhibition of fever by a virus.
571 Proc, Natl. Acad. Sci. USA **93:**11029-11034.
572 2. **Alzhanova, D., D. M. Edwards, E. Hammarlund, I. G. Scholz, D. Horst, M. J.**
573 **Wagner, C. Upton, E. J. Wiertz, M. K. Slifka, and K. Fruh.** 2009. Cowpox virus
574 inhibits the transporter associated with antigen processing to evade T cell
575 recognition. Cell host & microbe **6:**433-445.
576 3. **Assefa, S., T. M. Keane, T. D. Otto, C. Newbold, and M. Berriman.** 2009.
577 ABACAS: algorithm-based automatic contiguation of assembled sequences.
578 Bioinformatics **25:**1968-1969.
579 4. **Berman, D., P. J. Kerr, R. Stagg, B. H. van Leeuwen, and T. Gonzalez.** 2006.
580 Should the 40-year-old practice of releasing virulent myxoma virus to control rabbits
581 (*Oryctolagus cuniculus*) be continued? Wildl. Res. **33:**549-556.
582 5. **Best, S. M., and P. J. Kerr.** 2000. Coevolution of host and virus: the pathogenesis of
583 virulent and attenuated strains of myxoma virus in resistant and susceptible
584 European rabbits. Virology **267:**36-48.
585 6. **Cameron, C., S. Hota-Mitchell, L. Chen, J. Barrett, J. X. Cao, C. Macaulay, D.**
586 **Willer, D. Evans, and G. McFadden.** 1999. The complete DNA sequence of
587 myxoma virus. Virology **264:**298-318.
588 7. **Camus-Bouclainville, C., M. Gretillat, R. Py, J. Gelfi, J. L. Guerin, and S.**
589 **Bertagnoli.** 2011. Genome sequence of SG33 strain and recombination between
590 wild-type and vaccine myxoma viruses. Emerging infectious diseases **17:**633-638.
591 8. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell,**
592 **and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics **21:**3422-
593 3423.
594 9. **Collin, N., J. L. Guerin, I. Drexler, S. Blanie, J. Gelfi, S. Boullier, G. Foucras, G.**
595 **Sutter, and F. Messud-Petit.** 2005. The poxviral scrapin MV-LAP requires a
596 myxoma viral infection context to efficiently downregulate MHC-I molecules. Virology
597 **343:**171-178.
598 10. **Drummond, A. J., and A. Rambaut.** 2007. BEAST: Bayesian evolutionary analysis
599 by sampling trees. BMC evolutionary biology **7:**214.
600 11. **Elde, N. C., S. J. Child, M. T. Eickbush, J. O. Kitzman, K. S. Rogers, J.**
601 **Shendure, A. P. Geballe, and H. S. Malik.** 2012. Poxviruses deploy genomic
602 accordions to adapt rapidly against host antiviral defenses. Cell **150:**831-841.
603 12. **Esposito, J. J., S. A. Sammons, A. M. Frace, J. D. Osborne, M. Olsen-**
604 **Rasmussen, M. Zhang, D. Govil, I. K. Damon, R. Kline, M. Laker, Y. Li, G. L.**
605 **Smith, H. Meyer, J. W. LeDuc, and R. M. Wohlhueter.** 2006. Genome sequence
606 diversity and clues to the evolution of Variola (smallpox) virus. Science **313:**807-812.
607 13. **Fenner, F.** 1983. The Florey lecture, 1983. Biological control, as exemplified by
608 smallpox eradication and myxomatosis. Proc R Soc Lond B Biol Sci **218:**259-285.
609 14. **Fenner, F., M. F. Day, and G. M. Woodroofe.** 1956. Epidemiological consequences
610 of the mechanical transmission of myxomatosis by mosquitoes. J. Hyg. (Camb.)
611 **54:**284-303.
612 15. **Fenner, F., and B. Fantini.** 1999. Biological Control of vertebrate Pests: The History
613 of Myxomatosis, An Experiment in Evolution. Stylus Pub Llc.
614 16. **Fenner, F., and I. D. Marshall.** 1957. A comparison of the virulence for European
615 rabbits (Oryctolagus cuniculus) of strains of myxoma virus recovered in the field in
616 Australia, Europe and America. The Journal of hygiene **55:**149-191.
617 17. **Fenner, F., and F. N. Ratcliffe.** 1965. Myxomatosis. Cambridge University Press,
618 Cambridge.
619 18. **Fuller, S. J., P. B. Mather, and J. C. Wilson.** 1996. Limited genetic differentiation
620 among wild Oryctolagus cuniculus L. (rabbit) populations in arid eastern Australia.
621 Heredity **77 (Pt 2):**138-145.

JVI Accepts published online ahead of print

622 19. **Fuller, S. J., J. C. Wilson, and P. B. Mather.** 1997. Patterns of differentiation among
623 wild rabbit populations Oryctolagus cuniculus L. in arid and semiarid ecosystems of
624 north-eastern Australia. Molecular ecology **6:**145-153.
625 20. **Gedey, R., X.-L. Jin, O. Hinthong, and J. L. Shisler.** 2006. Poxviral regulation of
626 the host-NF-kB response: the vaccinia virus M2L protein inhibits induction of NF-kB
627 activation via an ERK2 pathway in virus-infected human embryonic kidney cells. J.
628 Virol. **80:**8676-8685.
629 21. **Guerin, J. L., J. Gelfi, S. Boullier, M. Delverdier, F. A. Bellanger, S. Bertagnoli, I.**
630 **Drexler, G. Sutter, and F. Messud-Petit.** 2002. Myxoma virus leukemia-associated
631 protein is responsible for major histocompatibility complex class I and Fas-CD95
632 down-regulation and defines scrapins, a new group of surface cellular receptor
633 abductor proteins. Journal of virology **76:**2912-2923.
634 22. **Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O.**
635 **Gascuel.** 2010. New algorithms and methods to estimate maximum-likelihood
636 phylogenies: assessing the performance of PhyML 3.0. Systematic biology **59:**307-
637 321.
638 23. **Katoh, K., G. Asimenos, and H. Toh.** 2009. Multiple alignment of DNA sequences
639 with MAFFT. Methods Mol Biol **537:**39-64.
640 24. **Kerr, P., and G. McFadden.** 2002. Immune responses to myxoma virus. Viral
641 immunology **15:**229-246.
642 25. **Kerr, P. J.** 2012. Myxomatosis in Australia and Europe: a model for emerging
643 infectious diseases. Antiviral research **93:**387-415.
644 26. **Kerr, P. J., E. Ghedin, J. V. Depasse, A. Fitch, I. M. Cattadori, P. J. Hudson, D. C.**
645 **Tscharke, A. F. Read, and E. C. Holmes.** 2012. Evolutionary history and
646 attenuation of myxoma virus on two continents. PLoS pathogens **8:**e1002950.
647 27. **Kerr, P. J., J. Hone, L. Perrin, N. French, and C. K. Williams.** 2010. Molecular and
648 serological analysis of the epidemiology of myxoma virus in rabbits. Veterinary
649 microbiology **143:**167-178.
650 28. **Kerr, P. J., R. J. Jackson, A. J. Robinson, J. Swan, L. Silvers, N. French, H.**
651 **Clarke, D. F. Hall, and M. K. Holland.** 1999. Infertility in female rabbits (Oryctolagus
652 cuniculus) alloimmunized with the rabbit zona pellucida protein ZPB either as a
653 purified recombinant protein or expressed by recombinant myxoma virus. Biology of
654 reproduction **61:**606-613.
655 29. **Kerr, P. J., J. C. Merchant, L. Silvers, G. M. Hood, and A. J. Robinson.** 2003.
656 Monitoring the spread of myxoma virus in rabbit Oryctolagus cuniculus populations
657 on the southern tablelands of New South Wales, Australia. II. Selection of a strain of
658 virus for release. Epidemiology and infection **130:**123-133.
659 30. **Kerr, P. J., H. D. Perkins, B. Inglis, R. Stagg, E. McLaughlin, S. V. Collins, and B.**
660 **H. Van Leeuwen.** 2004. Expression of rabbit IL-4 by recombinant myxoma viruses
661 enhances virulence and overcomes genetic resistance to myxomatosis. Virology
662 **324:**117-128.
663 31. **Kerr, P. J., M. B. Rogers, A. Fitch, J. V. Depasse, I. M. Cattadori, P. J. Hudson,**
664 **D. C. Tscharke, E. C. Holmes, and E. Ghedin.** 2013. Comparative analysis of the
665 complete genome sequence of the Californian MSW strain of myxoma virus reveals
666 potential host adaptations. Journal of virology.
667 32. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H.**
668 **McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J.**
669 **Gibson, and D. G. Higgins.** 2007. Clustal W and Clustal X version 2.0.
670 Bioinformatics **23:**2947-2948.
671 33. **Macen, J. L., C. Upton, N. Nation, and G. McFadden.** 1993. SERP1, a serine
672 proteinase inhibitor encoded by myxoma virus, is a secreted glycoprotein that
673 interferes with inflammation. Virology **195:**348-363.
674 34. **Mansouri, M., E. Bartee, K. Gouveia, B. T. Hovey Nerenberg, J. Barrett, L.**
675 **Thomas, G. Thomas, G. McFadden, and K. Früh.** 2003. The PHD/LAP-domain

676 protein M153R of myxoma virus is a ubiquitin ligase that induces the rapid
677 internalization and lysosomal destruction of CD4. J. Virol. **77:**1427-1440.
678 35. **Marshall, I. D.** 1959. The influence of ambient temperature on the course of
679 myxomatosis in rabbits. J. Hyg. (Camb.) **57:**484-497.
680 36. **Marshall, I. D., A. L. Dyce, W. E. Poole, and F. Fenner.** 1955. Studies in the
681 epidemiogy of infectious myxomatosis of rabbits. IV. Observations of disease
682 behaviour in two localities near the northern limit of rabbit infestation in Australia,
683 May 1952 to April 1953. J. Hyg. (*Camb.*) **53:**12-25.
684 37. **Marshall, I. D., and F. Fenner.** 1960. Studies in the epidemiology of infectious
685 myxomatosis of rabbits. VII. The virulence of strains of myxoma virus recovered from
686 Australian wild rabbits between 1951 and 1959. The Journal of hygiene **58:**485-488.
687 38. **Martin, D. P., P. Lemey, M. Lott, V. Moulton, D. Posada, and P. Lefeuvre.** 2010.
688 RDP3: a flexible and fast computer program for analyzing recombination.
689 Bioinformatics **26:**2462-2463.
690 39. **Morales, M., M. A. Ramirez, M. J. Cano, M. Parraga, J. Castilla, L. I. Perez-**
691 **Ordoyo, J. M. Torres, and J. Barcena.** 2009. Genome comparison of a
692 nonpathogenic myxoma virus field strain with its ancestor, the virulent Lausanne
693 strain. Journal of virology **83:**2397-2403.
694 40. **Moses, A.** 1911. O virus do myxoma dos coelhos. Mem Inst Osw Cruz **3:**46-53.
695 41. **Mossman, K., S. F. Lee, M. Barry, L. Boshkov, and G. McFadden.** 1996.
696 Disruption of M-T5, a novel myxoma virus gene member of poxvirus host range
697 superfamily, results in dramatic attenuation of myxomatosis in infected European
698 rabbits. Journal of virology **70:**4394-4410.
699 42. **Myers, K.** 1954. Studies in the epidemiology of infectious myxomatosis of rabbits. II.
700 Field experiments, August-November 1950, and the first epizootic of myxomatosis in
701 the Riverine Plain of south-eastern Australia. The Journal of hygiene **52:**47-59.
702 43. **Myers, K., I. D. Marshall, and F. Fenner.** 1954. Studies in the epidemiology of
703 infectious myxomatosis of rabbits. III. Observations on two succeeding epizootics in
704 Australian wild rabbits on the Riverine plain of south-eastern Australia. J. Hyg.
705 (*Camb.*) **52:**337-360.
706 44. **Myskiw, C., J. Arsenio, C. Hammett, R. van Bruggen, Y. Deschambault, N.**
707 **Beausoleil, S. Babiuk, and J. Cao.** 2011. Comparative analysis of poxvirus
708 orthologues of the vaccinia virus E3 protein: modulation of protein kinase R activity,
709 cytokine responses, and virus pathogenicity. Journal of virology **85:**12280-12291.
710 45. **Otto, T. D., G. P. Dillon, W. S. Degrave, and M. Berriman.** 2011. RATT: Rapid
711 Annotation Transfer Tool. Nucleic acids research **39:**e57.
712 46. **Page, R. D. M., and E. C. Holmes.** 1998. Molecular Evolution: A Phylogenetic
713 Approach. Blackwell Science Ltd, Oxford.
714 47. **Parer, I., W. R. Sobey, D. Conolly, and R. Morton.** 1994. Virulence of strains of
715 myxoma virus and the resistance of wild rabbits (*Oryctolagus cuniculus* L.), from
716 different locations in Australasia. Australian Journal of Zoology **42:**347-362.
717 48. **Rahman, M. M., J. Liu, W. M. Chan, S. Rothenburg, and G. McFadden.** 2013.
718 Myxoma Virus Protein M029 Is a Dual Function Immunomodulator that Inhibits PKR
719 and Also Conscripts RHA/DHX9 to Promote Expanded Host Tropism and Viral
720 Replication. PLoS pathogens **9:**e1003465.
721 49. **Ramelot, T. A., J. R. Cort, A. A. Yee, F. Liu, M. B. Goshe, A. M. Edwards, R. D.**
722 **Smith, C. H. Arrowsmith, T. E. Dever, and M. A. Kennedy.** 2002. Myxoma virus
723 immunomodulatory protein M156R is a structural mimic of eukaryotic translation
724 initiation factor eIF2alpha. Journal of molecular biology **322:**943-954.
725 50. **Ratcliffe, F. N., K. Myers, B. V. Fennessy, and J. H. Calaby.** 1952. Myxomatosis in
726 Australia. A step towards the biological control of the rabbit. Nature **170:**7-19.
727 51. **Rolls, E. C.** 1969. They All Ran Wild. Angus and Robertson, Melbourne.
728 52. **Russell, R. J., and S. J. Robbins.** 1989. Cloning and molecular characterization of
729 the myxoma virus genome. Virology **170:**147-159.

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

730 53. **Saint, K. M., N. French, and P. Kerr.** 2001. Genetic variation in Australian isolates
731 of myxoma virus: an evolutionary and epidemiological study. Archives of virology
732 **146:**1105-1123.
733 54. **Swofford, D. L.** 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and other
734 methods), 4 ed. Sinauer Associates, Sunderland, Mass.
735 55. **Upton, C., J. L. Macen, R. A. Maranchuk, A. M. DeLange, and G. McFadden.**
736 1988. Tumorigenic poxviruses: fine analysis of the recombination junctions in
737 malignant rabbit fibroma virus, a recombinant between Shope fibroma virus and
738 myxoma virus. Virology **166:**229-239.
739 56. **Willer, D. O., G. McFadden, and D. H. Evans.** 1999. The complete genome
740 sequence of shope (rabbit) fibroma virus. Virology **264:**319-343.
741 57. **Williams, C. K., C. C. Davey, R. J. Moore, L. A. Hinds, L. E. Silvers, P. J. Kerr, N.**
742 **French, G. M. Hood, R. P. Pech, and C. J. Krebbs.** 2007. Populations responses to
743 sterility imposed on female European rabbits. J. Appl. Ecol. **44:**291-301.
744 58. **Zerbino, D. R., G. K. McEwen, E. H. Margulies, and E. Birney.** 2009. Pebble and
745 rock band: heuristic resolution of repeats and scaffolding in the velvet short-read de
746 novo assembler. PloS one **4:**e8407.
747 59. **Zhang, L., N. Y. Villa, and G. McFadden.** 2009. Interplay between poxviruses and
748 the cellular ubiquitin/ubiquitin-like pathways. FEBS Lett **583:**607-614.
749
750
751

JVI Accepts published online ahead of print

752 **TABLE LEGENDS**

753

754 **Table 1.** Origin of strains of MYXV sequenced here and in Kerr et al. (2012). The 7 new
755 sequences are shaded grey.

756

757 **Table 2.** Mutations from SLS conserved in Australian MYXV isolates.

758

759 **Table 3.** Genome changes in SLS compared to Lu.

760

761 **Table 4.** Coding changes and indels in viruses from 1951-53 compared to SLS.

762

763 **Table 5.** Insertions and deletions in coding regions of Australian MYXV isolates.

764

765 **Table 6.** Coding changes from SLS in viruses of defined virulence.

766

JVI Accepts published online ahead of print

767 **FIGURE LEGENDS**

768

769 **Figure 1. Phylogeny and phylogeography of MYXV isolates.** Samples are color-coded
770 according to place of sampling (BRK [Brooklands] = green, OB = pink, SWH [Southwell Hill]
771 = red, WS [Woodstock] = Blue), while the state/region of sampling is noted in parenthesis
772 (ACT = Australian Capital Territory, CD = Canberra District, NSW = New South Wales, QLD
773 = Queensland, TA = Tasmania, VIC = Victoria). Viruses newly sequenced here are marked
774 with * symbol. The phylogenetic distribution of mutation and reversion in the *M029* gene, and
775 of mutation in the *M156* gene, is also shown. Bootstrap values are shown for key nodes and
776 all horizontal branches are drawn according to the number of nucleotide substitutions per
777 year. Divergence times (95% HPD values) for two key nodes in the Australian part of the
778 phylogeny were inferred from the BEAST analysis (see Methods).

779

780 **Figure 2. Heatmaps showing the number of SNP mutations in each gene from the (A)**
781 **SLS or (B) Lu strains.** Genes are organized in rows according to their order in the
782 reference genome, and taxa are ordered along columns by their branching order in the
783 MYXV phylogeny.

784

785 **Figure 3. Expansion of the TIR boundaries in SWH 1209. OB2 W60, BD23 and BD24.**
786 Duplication of the 1635 nucleotide region outside the right hand (RH) TIR encoding *M153R*-
787 *M156R* and inversion and insertion of this sequence at the left hand (LH) TIR. This replaces
788 the 3' 923 nucleotides of *M009L* at the LH end; there are now complete copies of *M156* and
789 *M154* at each end of the genome but *M153* has lost the 5' 36 nucleotides and promoter at
790 the LH end of the genome. *M156R* originally spanned the TIR boundary at the RH end of the
791 genome. The sequence in the TIR was present at the LH end as non-coding DNA. Note that
792 the Lu genome annotation does not use the M155 gene number.

793

794 **Figure 4. Insertions and deletions in non-coding repeat sequence regions in MYXV**
795 **isolates.**

796

797 **Figure 5. Sequence mutations in 5 potential promoter regions**. **(A)** <u>*M008.1L/R* late</u>
798 <u>promoter</u>: There is a TAAAT late promoter motif italicized incorporating the ATG start codon
799 (in red). This is preceded by a 6 nt spacer and then a run of 8 T (underlined) which is typical
800 of strong late promoters (spacer of 4-10 nts and then a T rich tract of 5-15 nts). A number of
801 isolates have a mutation at the underlined upstream T (T>C) but this seems unlikely to affect
802 the promoter structure. WS6 346 and OBY317 have a T deletion in the 8 T tract. At the LH
803 end of the genome this promoter is in a non-coding sequence, but at the RH end it is within

804 the 3' end of M156R. **(B)** *M057L* late promoter: TAAAT motif incorporating the ATG, a 4
805 nucleotide spacer and then 8 T. OB1/406 has an extra T in the T tract making 9 T. **(C)**
806 *M000.5L/R* possible late promoter: Putative promoter structure is italicized–the 2 C residues
807 are almost invariably A in late promoters. There is no upstream T rich domain for 100 nts
808 upstream. G-91 has a C>T mutation in the putative promoter structure → TTCATG. There is
809 no sign of an upstream A rich region that could act as an early promoter. **(D)** *M138L* early
810 promoter: SWH 8/2/93 has an extra A where the dash is in the poly A tract of the italicized
811 potential promoter. **(E)** M153R early promoter: Potential early promoter sequence is
812 italicized. All Australian isolates lack the upstream T underlined, this seems unlikely to have
813 any impact on the promoter. WS6 1071 and OB31120 lack an A in the homopolymer tract –
814 highlighted in bold. A possible alternative promoter is underlined but seems too close to the
815 ATG.
816

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

Table 1. Origin of strains of MYXV sequenced here and in Kerr et al. (2012)

Virus	Formal name	Geographic origin	Source	Reference	Virulence grade	Region sequenced[4]	Accession #
SLS (Moses strain/strain B)	None given	Brazil	Rabbit tissue stock [1](Fenner)	Fenner and Marshall 1957	1	1-161777 (161763)	JX565574
Glenfield (Gv)	Aust/Dubbo/2-51/1	Central NSW	[2]CV-1 cell stock	Russell and Robbins 1989	1	15-161763 (161742)	JX565567
KM13	Aust/Corowa/12-52/2	Southern NSW	rabbit tissue stock (Fenner)	Fenner and Marshall 1957	3	1-161777 (161771)	JX565569
KM13_2A	Aust/Corowa/12-52/2A	Southern NSW	rabbit tissue stock (Fenner)	Marshall 1959	3	1-161777 (161769)	KC660080
Uriarra (Ur)	Aust/Uriarra/2-53/1	Canberra district	CV-1 cell stock	Russell and Robbins 1989	5	1-161777 (161768)	JX565577
SWH	Aust/Southwell Hill/9-92/1	Canberra district	Wild rabbit	Saint et al 2001	4	1-161777 (161797)	JX565576
BRK	Aust/Brooklands/4-93	Canberra district	Wild rabbit	Saint et al 2001	1	1-161777 (161701)	JX565562
Bendigo	Aust/Bendigo/7-92	Central Victoria	Wild rabbit	Saint et al 2001	1	1-161777 (161738)	JX565565
Meby	Aust/Meby/8-91	Tasmania	Wild rabbit	Saint et al 2001	5	87-161691 (161542)	JX565571
Lu CSL	Brazil/Campinas/1949/1	Brazil	[3]Commonwealth Serum Laboratories 1973		1	1-161777 (161778)	JX565570
Cornwall	England/Cornwall/4-54/1	Cornwall UK	Rabbit tissue stock (Fenner)	Fenner and Marshall 1957	1	1-161777 (161775)	JX565566
Sussex	England/Sussex/9-54/1	Sussex UK	Rabbit tissue stock (Fenner)	Fenner and Marshall 1957	3	1-161777 (161778)	KC660084
Nottingham attenuated	England/Nottingham/4-55/1	Nottingham UK	Rabbit tissue stock (Fenner)	Fenner and Marshall 1957	5	1-161777 (161777)	JX565572
Gung/91 (G-91)	Aust/Gungahlin/1-91	Canberra district	Wild rabbit	Saint et al 2001	4	151-161627 (161443)	JX565568
Wellington	Aust/Wellington/1-91	Central NSW	Wild rabbit	Saint et al 2001	1	29-161749 (161688)	JX565582
BRK/12-2-93	Aust/Brooklands/2-93	Canberra district	Wild rabbit	Kerr et al 2010	ND	140-161638 (161496)	JX565563
BD23	Aust/Bulloo Downs/11-99	SW Queensland	Wild rabbit	Berman et al 2006	ND	285-161555 (161971)	JX565584
BD44	Aust/Bulloo Downs/12-99	SW Queensland	Wild rabbit	Berman et al 2006	ND	1-161777 (162847)	KC660079
BRK/897	Aust/Brooklands/1-95	Canberra district	Wild rabbit	Kerr et al 2010	ND	103-161675 (161545)	JX565564
OB1/406	Aust/OB1/Hall/3-94	Canberra district	Wild rabbit	Kerr et al 2010	ND	87-161691 (161612)	JX565573
OB2/W60	Aust/OB2/Hall/11-95	Canberra district	Wild rabbit	Kerr et al 2010	ND	1-161777 (162483)	KC660081
OB3/Y317	Aust/OB3/Hall/2-94	Canberra district	Wild rabbit	Kerr et al 2010	ND	1-161777 (161748)	KC660083
OB3/1120	Aust/OB3/Hall/2-96	Canberra district	Wild rabbit	Kerr et al 2010	ND	1-161777 (161722)	KC660082
WS1/234	Australia/Woodstock 1 /3-94	Canberra district	Wild rabbit	Kerr et al 2010	ND	1-161777 (161754)	JX565578
WS6/1071	Aust/Woodstock 6 /11-95	Canberra district	Wild rabbit	Kerr et al 2010	ND	41-161737 (161752)	JX565580
WS1/328	Aust/Woodstock 1 /3-94	Canberra district	Wild rabbit	Kerr et al 2010	ND	156-161622 (161483)	JX565579
WS6/346	Aust/Woodstock 6 /3-95	Canberra district	Wild rabbit	Kerr et al 2010	ND	140-161638 (161430)	JX565581
SWH/8-2-93	Aust/Southwell Hill/2-93	Canberra district	Wild rabbit	Kerr et al 2010	ND	1-161777 (161740)	JX565575
SWH/805	Aust/Southwell Hill/11-93	Canberra district	Wild rabbit	Kerr et al 2010	ND	1-161777 (161780)	KC660085
SWH/1209	Aust/Southwell Hill/2-96	Canberra district	Wild rabbit	Kerr et al 2010	ND	33-161745 (162413)	JX565583

[1] Virus stocks were originally obtained as freeze dried rabbit tissue from Professor Frank Fenner, John Curtin School of Medical Research, Australian National University, Canberra, ACT.
[2] Virus stocks were from viruses plaque purified as described in Russell and Robbins 1989.
[3] Virus was from an ampoule of freeze dried rabbit tissue powder prepared by the Commonwealth Serum Laboratories for rabbit control.
[4] Based on the Lu sequence from Cameron et al 1999 1-161777 as corrected by Morales et al 2009; the actual sequence length is shown in brackets.
Gray rows indicate isolates sequenced for this paper

Table 2. Mutations from SLS conserved in Australian MYXV isolates[1]

Position SLS	Gene	Function	Mutation	No. of viruses
1968	*M002L/R*	TNF binding /apoptosis inhibition	A226V	21 (all recent)
2576	*M002L/R*	TNF binding /apoptosis inhibition	synonymous GCG>GCA	24 (all Australian)
3168	*M003.1L/R*	PRR signal inhibition?	A37V	21 (all recent)
5082	*M005L/R*	Apoptosis inhibition /Ub ligase	R434W	21 (all recent)
5756	*M005L/R*	Apoptosis inhibition /Ub ligase	S209Y	21 (all recent)
11484	Intergenic	-	A>G	13
12348	*M009L*	Putative Ub ligase	A261V	[2]16
12715	*M009L*	Putative Ub ligase	Frameshift T insert 420	20 (including [3]Meby)
16042	*M014L*	Putative Ub ligase	V175I	21 (all recent)
16201	*M014L*	Putative Ub ligase	G122W	21 (all recent)
16478	*M014L*	Putative Ub ligase	Synonymous GTC>GTT	18
16615	Intergenic	-	A deletion	13 (includes Ur)
16923	*M015L*	Ribonucleotide reductase	Synonymous GAA>GAG	21 (all recent)
17332	*M015L*	Ribonucleotide reductase	V85A	21 (all recent)
17877	*M017L*	?	E71K	21 (all recent)
18236	intergenic	-	multiple GTAGGTAG insert	21
18250	Intergenic	-	multiple AGTTTAGT insert	17
18277	Intergenic	-	T>C	21 (all recent)
21578	*M021L*	EV maturation; vacF12L orthologue	D315N	14
23608	*M022L*	EEV protein	Synonymous GTC>GTT	21 (all recent)
24933	*M025L*	VACV F16 orthologue	M11I	20
28185	*M028L*	EEV formation	S244L	18
36832	*M034L*	DNA polymerase	Synonymous TTC>TTT	21 (all recent)
38437	*M036L*	VACV O1 orthologue	C270Y	22(all recent & KM13)
38987	*M036L*	VACV O1 orthologue	Synonymous CTG>TTG	21 (all recent)
41406	Intergenic	-	T insert	20
47167	*M045L*	Virion morphogenesis	D263N	21 (all recent)
50515	*M049R*	VACV G5 orthologue	Synonymous CTG>TTG	21 (all recent)
52256	*M052L*	Core structural protein	S29N	21 (all recent)
55900	*M057L*	Core protein	L90V	21 (all recent)
57398	*M060R*	Virion protein	Synonymous GCG>GCA	21 (all recent)
58821	*M062R*	Host range	K142T	20
59512	*M063R*	Host range	S195C	21 (all recent)
59585	Intergenic	-	T deletion	20
60567	*M065R*	Poly A pol subunit	T98M	21 (all recent)
64305	*M068R*	RNA pol subunit	Synonymous ACG>ACA	22
70115	*M072L*	RNA pol assoc transcription factor	I150T	22
82120	*M083L*	CA homologue/virion protein	G insert	21 (all recent)
85496	*M086L*	DNA helicase	Synonymous CGA>CGG	21 (all recent)
99168	*M099L*	Core protein precursor	Synonymous GCG>GCA	20
112683	*M114R*	RNA pol subunit	P1147H	12
115902	*M121R*	CLECT EEV protein	S21F	20
119939	*M127L*	photolyase	Syn 1107 AGC>AGT	15
123409	*M132L*	?	C133Y	18
125935	*M134R*	Membrane protein	S84P	21 (all recent)
128748	*M134R*	Membrane protein	Synonymous GCG>GCA	21 (all recent)
131595	*M134R*	Membrane protein	AAA insert (K)	22
133151	*M137R*	VACV A51 orthologue	Synonymous GGC>GGT	20
135593	*M140R*	Putative Ub ligase	P76H	24
142764	*M148R*	Putative Ub ligase	L383F	21 (all recent)
145699	*M150R*	Putative Ub ligase	P173S	21 (all recent)
147192	*M151R*	SERPIN (Serp 2)	R173G	21 (all recent)
148711	*M153R*	Ub ligase MHC-1 downregulation	synonymous TGT>TGC	17
149127	*M153R*	Ub ligase MHC-1 downregulation	L204S	[4]19 (24) (all Australian)
149717	*M154R*	VACVM2 orthologue / NF-κB inhibition?	Y53C	19
149836	*M154L*	VACVM2 orthologue / NF-κB inhibition?	Syn 39 GTC>GTT	13
150280	*M156R*	e-IF2α homologue (IFN resistance)	L98P	13

[1] Mutations shared by 12 or more viruses; [2] SWH/1209, OB3/1120, BD44 and BD23 have deleted this region of *M009L* . [3] Meby was isolated in Tasmania not on mainland Australia. All other Australian isolates were from the mainland. There are 21 isolates from 1991-1999; [4] All Australian isolates sequenced have the nucleotide mutation at SLS 149127 but Gv, WS6/1071, BD44 and Meby have frameshift mutations that alter the reading frame of M153 and Wellington has a 9 nucleotide deletion which shifts the amino acid mutation to L201S. Mutations duplicated in TIR are indicated by L/R in the gene name and are only shown at the LH end of the genome.

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

Table 3. Genome changes in SLS compared to Lu*

Lu position	Lu sequence	SLS sequence	SLS position	Gene	mutation	Protein function
22	A	-	21	non coding		
621	A	C	621	non coding		
2577	C	T	2577	M002L	syn GCG>GCA (A) nt 69	TNF binding/apoptosis regulator
2794	G	-	2793	intergenic	bw M003.1 and M002	
6092	C	T	6091	M005L	D98N GAC>AAC nt 297	E3 Ub ligase/apoptosis regulator
6349	-	G	6349	M005L	frameshift	E3 Ub ligase/apoptosis regulator
6351	A	G	6351	M005L	syn CCT>CCC (P) nt 34	E3 Ub ligase/apoptosis regulator
9370	G	A	9370	M008L	Syn GAC>GAT (D) nt 1005	Putative E3 Ub ligase
13169	G	T	13169	intergenic	bw M009L and M10L	
18237-18244	GGTATGTA	-	18235	intergenic	bw17L and 18L tandem repeat	
20976	C	T	20967	M021L	Syn CAG>CAA (Q) nt 1554	EV maturation
22645	C	T	22636	M022L	syn ACG>ACA (T) nt 1032	EV protein
25039	G	A	25030	M026R	A9T GCA>ACA nt 25	DNA binding phosphoprotein
30013	T	C	30004	M030L	T10A ACA>GCA nt 28	RNA pol subunit
36188	T	C	36179	M034L	Y227C TAT>TGT nt 680	DNA pol
38319	G	A	38310	M036L	syn TTC>TTT (F) nt 936	VACV O1L orthologue
40605-40607	TTC	-	40595	M040L	E 258 deleted	DNA binding phosphoprotein
48097	G	A	48085	M046L	syn CCC>CCT (P) nt 201	membrane protein
48780	A	G	48768	M047R	T164A ACA>GCA nt 490	late gene expression regulator
53704	T	C	53692	M054R	syn CGT>CGC (R) nt 519	Membrane fusion complex
54952	A	G	54940	intergenic	immediately 5' to M056R	
55308	C	T	55296	M057L	S291N AGT>AAT nt 872	core protein
56431	C	T	56419	M058R	Syn ATC>ATT(I) nt 228	core protein
57922	A	T	57910	M061R	K41N AAA>AAT (K) nt 123	Thymidine kinase
60376	A	G	60364	M065R	Syn ACA>ACG (T) nt 90	Poly A pol regulatory subunit
62205	.	T	62194	intergenic	immediately 5' to M068R	
67449	C	T	67438	M071L	syn TCG>TCA (S) nt 735	membrane protein
74445	G	A	74434	M076R	syn TCG>TCA (S) nt 1740	mRNA capping enzyme
80489	A	G	80478	M081R	Q371R CAG>CGG nt 1112	Early transcription factor subunit
82131	G	-	82119	M083L	C del at 513 >reading frame shift	Carbonic anhydrase homology/membrane protein
82179	C	T	82167	M083L	syn CTG>CTA (L) nt 462	Carbonic anhydrase homology/membrane protein
83974	A	G	83962	M085R	syn GTA>GTG (V) nt 669	VACV D10R orthologue
87056	A	C	87044	M088L	syn TCT>TCG (S) nt 1494	virion protein
90140	T	C	90128	M092L	syn GCA>GCG (A) nt 1790	core protein
118290	T	C	118278	M124R	syn GGT>GGC (G) nt774	unknown
123774	T	C	123762	M132L	syn GTA>GTG (V) nt 45	unknown
128789	C	T	128777	M134R	S1031L TCG>TTG nt 3092	membrane protein
129085	G	T	129073	M134R	A1130S GCT>TCT nt 3388	membrane protein
130326	C	T	130314	M134R	syn GGC>GGT (G) nt 4629	membrane protein
131079	C	T	131067	M134R	syn GAC>GAT (D) nt 5382	membrane protein
131133	T	C	131121	M134R	syn GCG>GCC (A) nt 5436	membrane protein
131176	C	T	131164	M134R	syn GTG>TTG (L) nt 5479	membrane protein
131187	C	A	131175	M134R	syn ACC>ACA (T) nt 5490	membrane protein
131230	A	G	131218	M134R	T1845A ACG>GCG nt 5533	membrane protein
131238	G	C	131226	M134R	E1847D GAG>GAC nt 5541	membrane protein
131259	T	C	131247	M134R	syn GAT>GAC (D) nt 5562	membrane protein
131316	A	G	131304	M134R	syn GCA>GCG (A) nt 5619	membrane protein
131328	G	A	131316	M134R	syn CCC>CCA (P) nt 5631	membrane protein
131377	A	G	131365	M134R	T1894A ACA>GCA nt 5680	membrane protein
131424	C	T	131412	M134R	syn GAC>GAT (D) nt 5727	membrane protein
131487	G	A	131475	M134R	M1930I ATG>ATA nt 5790	membrane protein
131550	G	A	131538	M134R	syn GGG>GGA (G) nt 5853	membrane protein
132122	G	A	132110	M135R	Syn GCG>GCA (A) nt 420	immune modulation/virulence
133197	G	A	133185	M137R	D96N GAC>AAC nt 286	VACV A51
133552	A	G	133540	M137R	D214G GAC>GGC nt 641	VACV A51
134435	C	T	134423	M138L	D106N GAC>AAC nt 316	α-2,3 sialyltransferase
141046	T	C	141034	M147R	S115P TCG>CCG nt 3343	S/T protein kinase
147887	G	A	147875	M152R	A66T GCA>ACA nt 196	Serp 3
148316	A	G	148304	M152R	T209A ACA>GCA nt 625	Serp 3
148375	A	G	148363	M152R	syn GCG>GCA (A) nt 684	Serp 3
148472	T	-	148459	M152R	nt 782	Serp 3
149140	C	T	149127	M153	S204L TCA>TTA nt 611	E3 Ub ligase MHC 1 downregulation
149864	C	T	149851	M154L	syn GTG>GTA (A) nt 24	NFκB regulation

* mutations in the TIR are only shown at the left hand TIR

Table 4. Coding changes and indels in viruses from 1951-53 compared to SLS

Gene	Protein function (No. of aa)	Virus		
		Glenfield	KM13	Ur
		1951 Grade 1	1952 Grade 3	1953 Grade 5
M005L/R	Host range /E3 Ub ligase (478)			ORF disrupted by C nt ins at 30
M014L	E3 ubiquitin ligase (517)	ORF disrupted by C nt ins at 1405	ORF disrupted by C nt ins at 1405	ORF disrupted by C nt ins at 1405
M036L	VACV O1 orthologue (680)		C270Y*	F293L
M044R	RNA helicase (678)	R606H		
M071L	Virion protein (324)			E172K
M072L	RNA pol assoc transcription factor (796)		I150T*	
M108R	DNA helicase (478)	F18I		
M114R	RNA pol subunit (1155)	A686V		
M130R	unknown (122)	ORF disrupted by G ins nt 30 f		
M134R	surface glycoprotein (2000)	AAA nt insert* K	AAA insert* K	A nt ins>premature stop at 1973
M137R	Orthology to VACV A51 (310)			A308T
M140R	E3 Ub ligase (553)	P76H*	P76H*	P76H*
M141R	OX-2 homologue (218)	S45 insert		
M153R	RING CH E3 Ub ligase (206)	ORF disrupted by G nt del at 329	L240S*	L240S*

* Present in all modern isolates from Australia

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

Table 5. Insertion and deletions in coding regions of Australian MYXV isolates

[1]SLS Position	Gene	Function	Mutation and context	Effect	Virus (virulence grade if known)
408	*M000.5L/R*	undetermined	G del	Frameshift from aa 58 and readthrough stop codon	BD44
5533	*M005L/R*	apoptosis inhibition/host range	C insert (homopol)	Stop after aa 317	WS6/346
6352	*M005L/R*	apoptosis inhibition/host range	C insert (homopol)	ORF disruption	Ur (5)
10663	*M008.1L/R*	Serpin	CC insert (homopol)	ORF disruption	BD44
11626	*M009L*	Putative Ub ligase	A del (homopol)	ORF disruption	SWH/8-2-93
12170	*M009L*	Putative Ub ligase	TA insert (TA rpt)	ORF disruption	BRK (1)
12715	*M009L*	Putative Ub ligase	A insert (homopol)	ORF disruption	All recent Australian except Bendigo (1)
12809	*M009L*	Putative Ub ligase	A deletion	ORF disruption	WS6/1071; OB3/1120
14397	*M012L*	dUTP pyrophosphatase	13 nt deletion	ORF disruption	OB3/Y317
15164	*M014L*	Putative Ub ligase	G insert (homopol)	ORF disruption	Gv (1); KM13 (3); Ur (5)
18324	*M018L*	cytoplasmic protein; VACF8L orthologue	TT insert	Frameshift from aa 60; readthrough adds 20 aa	OB3/Y317
22511	*M021L*	EV maturation; VACF12L orthologue	9 nt insert (duplication)	Duplicates LLG aa 4-6	OB3/Y317
38589-38680	*M036L*	VACV O1 orthologue	92 base deletion	ORF disruption	BRK (1)
40596	*M040L*	DNA binding phosphoprotein	TCT duplication	E258 inserted	Bendigo (1)
58322	*M061R*	Thymidine kinase	T insert (homopol)	Readthrough adds LKY to C terminus	WS1/234
59143	*M063R*	Host range	ACC duplication	H72 duplicated	Gung/91 (4)
59554-59568	*M063R*	Host range	15 base deletion of repeat sequence	Deletes TEEEE from a repeat at the C terminus	WS6/346
60122-60124	*M064R*	Host range?	AGA deleted (tandem repeat)	E deleted	OB1/406; BRK/12-2-93
60122-60124	*M064R*	Host range?	AGA inserted (tandem repeat)	E 168 inserted	Well (1); WS1/234; BRK/897; SWH/805; OB3/Y317
82120	*M083L*	CA homologue/structural	G insert (homopol)	Corrects G deletion in SLS	All recent Australian
92163	*M093L*	Core protein	GGAAAC duplication	VP duplication	SWH (4)
113080	*M115L*	Fusion protein	CTTCGT del	66D 67E deleted	Gung/91 (4)
122397	*M130R*	?	G insert (homopol)	ORF disrupted	Gv (1)
131595	*M134R*	Transmembrane protein	AAA insert (homopol)	K insert	All Australian except Ur (A) and SLS; OB3/1120 has 2A subpopulation
131595	*M134R*	Transmembrane protein	A insert (homopol)	Truncates ORF – early stop	Ur (5)
137195	*M141R*	OX-2 homologue; downregulation of macrophage activation	AGT insert (tandem repeats)	S insert in repeat sequence	Gv (1)
141092	*M147R*	S/T protein kinase	GT deletion (tandem repeat)	Repeat sequence disrupts ORF	BD23
148485	*M152R*	Serp 3	A del (homopol)	Premature stop 271 (SLS 273)	WS6/1071; OB3/1120
148845	*M153R*	Ub ligase/MHC-1 downregulation	G del (homopol)	ORF disruption early stop after aa 118	GV (1)
148845	*M153R*	Ub ligase/MHC-1 downregulation	G insert (homopol)	ORF disruption early stop after aa 124	BD44
148985	*M153R*	Ub ligase/MHC-1 downregulation	T del	ORF disruption stop after aa 161	WS6/1071; OB3/1120
149018	*M153R*	Ub ligase/MHC-1 downregulation	73 bp deletion between repeats	Sequence read through replaces CR domain of M153	Meby (5)
149062-149070	*M153R*	Ub ligase/MHC-1 downregulation	9 bp deletion in duplicated sequence	VEE repeat deleted from CR domain	Well (1)
150294	*M156R*	eIF2α homologue; IFN resistance	T del (homopol)	Read through stop – extra EG at C terminus	WS6/346; OB3/Y317

[1] Sequence positions are for the Australian progenitor SLS; mutations in genes within the TIRs are only shown for the left hand TIR

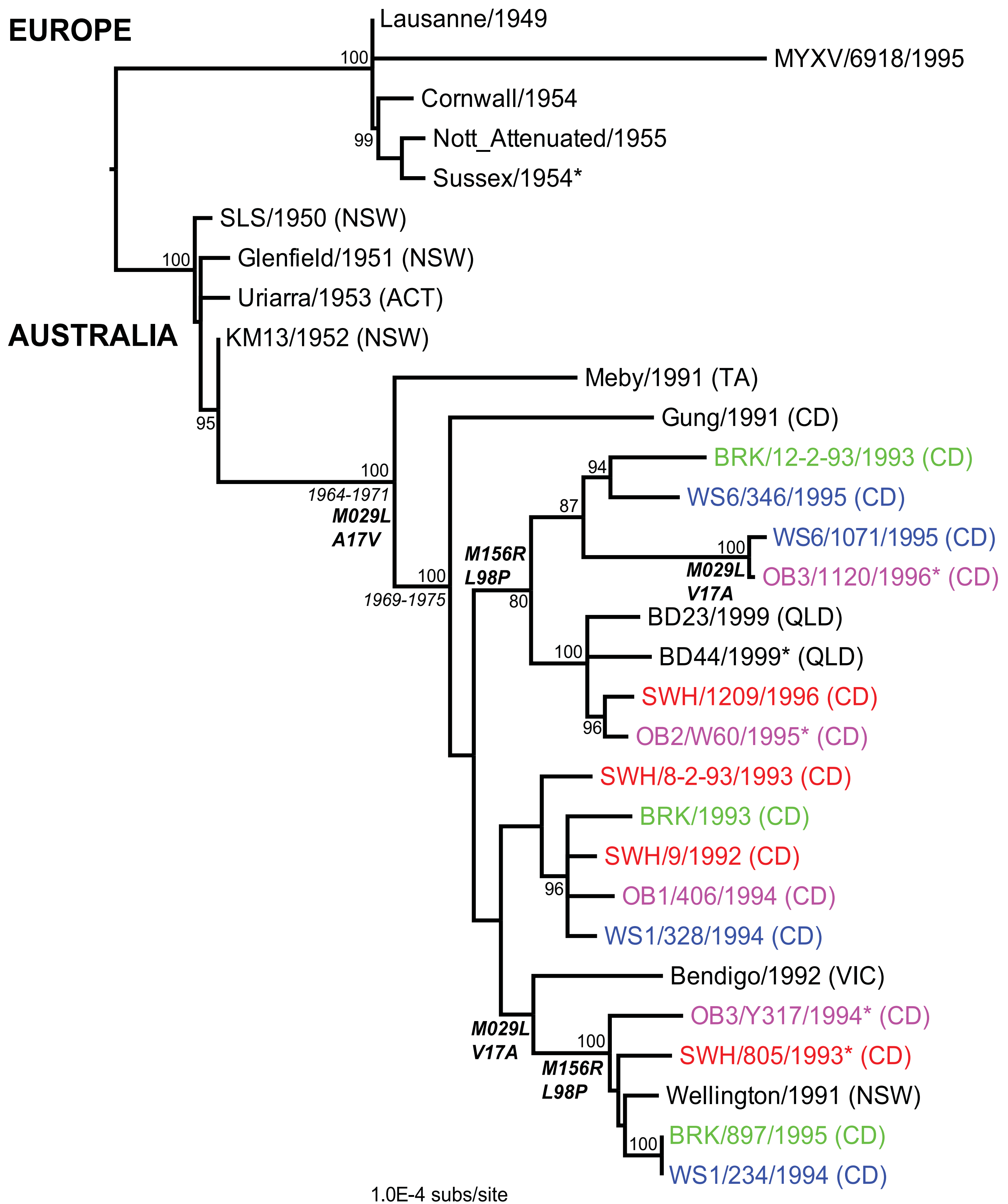
JVI Accepts published online ahead of print

Table . Coding changes from SLS in viruses of defined virulence

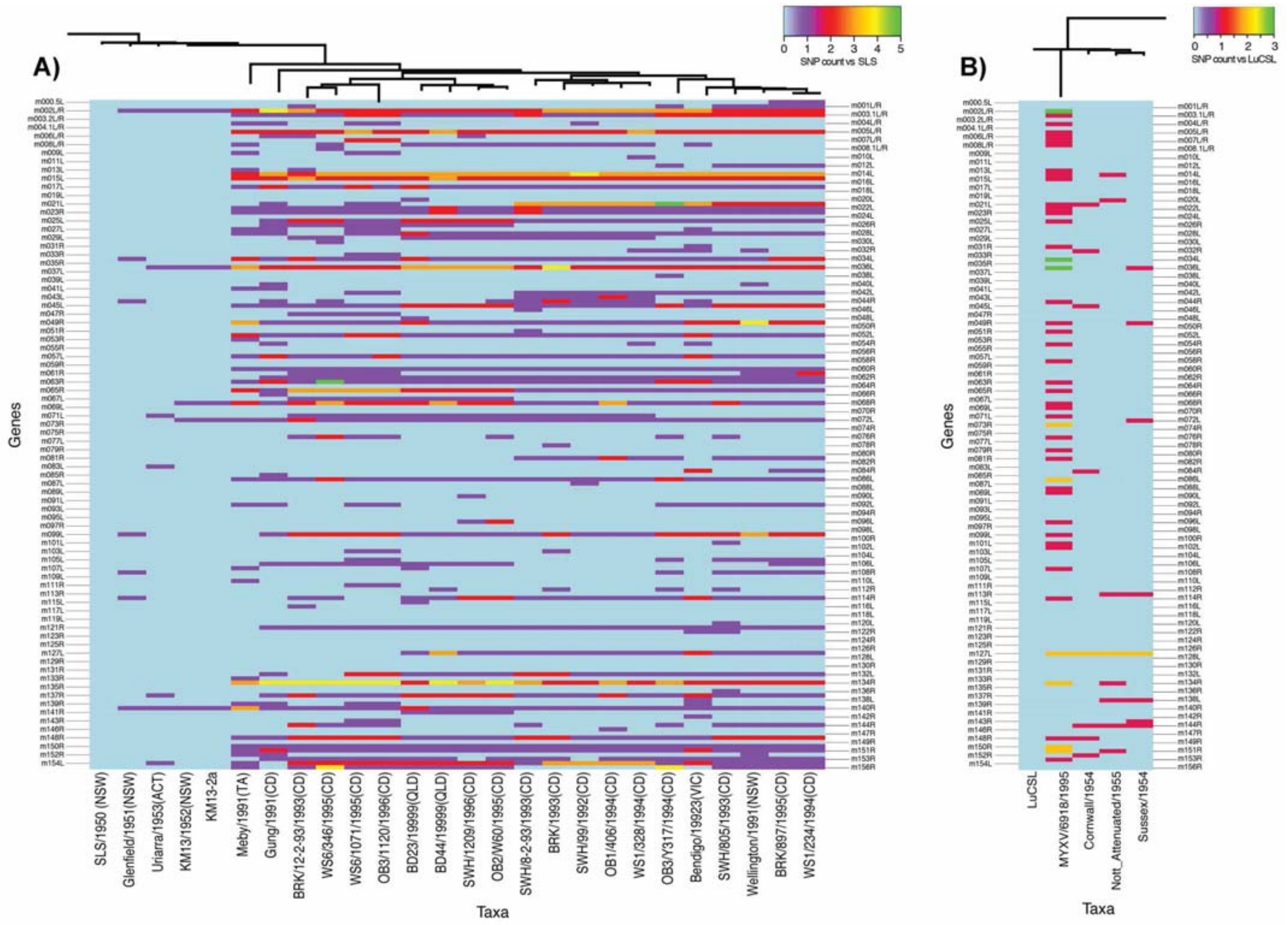
		Virus isolate/year of isolation (virulence rade)								
Protein	**Function**	**Glenfield/1951 (1)**	**KM1 /1952 ( )**	**Uriarra/195 (5)**	**Gun /1991 (4)**	**Wellin ton/1991 (1)**	**Meby/1991 (5)**	**Bendi o/1992 (1)**	**SWH/9/1992 (4)**	**BRK/199 (1)**
M001	Chemokine binding					S213N				
M002	TNF inhibition/anti-apoptosis				A226V T188A Q117R	A226V	A226V	A226V Q117R	A226V	A226V
M003.1	PRR signal inhibition?				A37V	A37V	A37V	A37V L76V	A37V	A37V
M004	Anti apoptosis						N138K		P227S	
M005	Anti-apoptosis			G insert disrupts reading frame	R434W S209Y	R434W S209Y	R434W S209Y	R434W S209Y	R434W S209Y	R434W S209Y
M006	E3 Ub ligase?				D171G					
M009	E3 Ub ligase?				A261V T insert at 420	A261V T insert at 420	P343L T insert at 420; TGG>TGA codon 348 A140T	A261V	A261V T insert at 420	A261V T insert at 420 TA insert nt 968
M012	dUTP nucleotidohydrolase					S129N				
M014	E3 Ub ligase?	C insert at 1406	C insert at 1406	C insert at 1406	V175I G122W	V175I G122W	V175I G122W	V175I G122W	V175I G122W	V175I G122W
M015	Ribonucleotide reductase small subunit				V85A	V85A	V85A	V85A	V85A	V85A
M017	?				E71K	E71K	E71K	E71K	E71K	E71K
M021	VACV F12/EEV maturation				D315N	R328C D315N		R328C D315N	R328C D315N	R328C D315N
M025	VACV F16 ?				M11I	M11I		M11I	M11I	M11I
M027	Poly A pol catalytic subunit				D143N			D143N		
M028	VACV E2 EEV formation				S244L	S244L	R33C	S244L	S244L	S244L
M029	IFN resistance PKR inhibitor								A17V	A17V
M032	Virion protein?					I481V		I481V		
M034	DNA pol						H222Y			
M036	VACV O1L/leu zipper motif		C270Y	F293L	C270Y	C270Y	P278S C270Y	C270Y	C270Y	92 nts deleted [1]C270Y E125K
M040	DNA binding phosphoprotein/virion							E 258 ins		
M041	VACV I5 structural				V49A					
M043	VACV I7 core/cyst proteinase?								A299T	A299T
M044	RNA helicase	R606H			S361L			S361L	V350A	V350A
M045	Core enzyme-morphogenesis				D263N	D263N	D263N	D263N	D263N	D263N
M049	Core protein					E21K	T314A			
M052	Fusion complex				S29N	S29N	S29N	D290G S29N	S29N	S29N
M057	Core protein				L90V	L90V	L90V	L90V	L90V	L90V
M062	Host range				K142T	K142T		K142T	K142T	K142T
M063	Host range				S195C H72 insert	S195C	S195C	S195C	S195C	S195C
M064	Host range					E163 insert				
M065	Poly A pol regulatory subunit				T98M	T98M	T98M	T98M	T98M	T98M
M068	RNA pol subunit						D1012G			
M071	VACV H3L membrane protein			E172K					V113I	V113I
M072	RNA pol assoc transcription factor		I150T		I150T	I150T	I150T	I150T	I150T	I150T
M081	VETF-1					R234C				
M083	Carbonic anhydrase homolog/structural protein				C nt insert 513 restores ORF	C nt insert 513 restores ORF	C nt insert 513 restores ORF	C nt insert 513 restores ORF	C nt insert 513 restores ORF	C nt insert 513 restores ORF
M084	VACV D9							S109L		
M087	mRNA capping enzyme/VITF								P33L	
M092	Major core protein						M119I			
M093	Core protein								V92P93 insert	
M099	Major core protein precursor					Y302H		Y302H		
M106	Fusion complex				M136I					
M107	VACV A17 membrane protein						A172T			
M108	DNA helicase	F18I				P38S				
M112	Holiday junction resolvase									A47V
M114	RNA pol subunit	A686V						P1147H	P1147H	P1147H
M115	IMV surface protein				66D67E deleted					
M121	EEV gp CLECT family; NK evasion?				S21F	S21F		S21F	S21F	S21F
M127	DNA photolyase							A343T		
M130	?	G nt 30 inserted ORF disrupted								
M132	?				C133Y	C133Y		C133Y	C133Y	C133Y
M134	membrane? EEV? VARV B22R	K1970 insert	K1970 insert	A nt inserted at 5911 premature stop	S84P K1970 insert R600K	S84P K1970 insert	S84P K1970 insert E1763A	S84P K1970 insert	S84P K1970 insert	S84P K1970 insert
M137	VACV A51			A308T				A201T		
M138	Sialyltransferase							M25L		
M139	TLR signal inhibition?						R210S			
M140	E3 Ub ligase?	P76H	P76H	P76H	P76H	P76H	P76H; T371A	P76H	P76H	P76H
M141	OX-2 homologue	S45 insert								
M142	Ser/Thr protein kinase							Y183C		
M144	VACV B5R orthologue					D106G				
M148	E3 Ub ligase?				L383F	R41Q L383F	L383F	L383F	L383F	L383F
M150	NFκ B inhibition				P173S	P173S	P173S	P173S	P173S	P173S
M151	SERP 2				P140S R173G	R173G	R173G	P140S R173G	R173G	R173G
M152	SERP 3				V106A	R181H				
M153	MHC downregulation	G nt 329 deleted ORF disrupted	L204S	L204S	R40L L204S	V182E183E184 deleted L204S	nt 509-573 deleted	L204S	L204S	L204S
M154	NFκB inhibition?					Y53C		Y53C	Y53C	Y53C
M156	IFN resistance					L98P				

[1] BRK has the nucleotide mutation at this position but the earlier 92 bp deletion in M036L means that the reading frame is disrupted here.

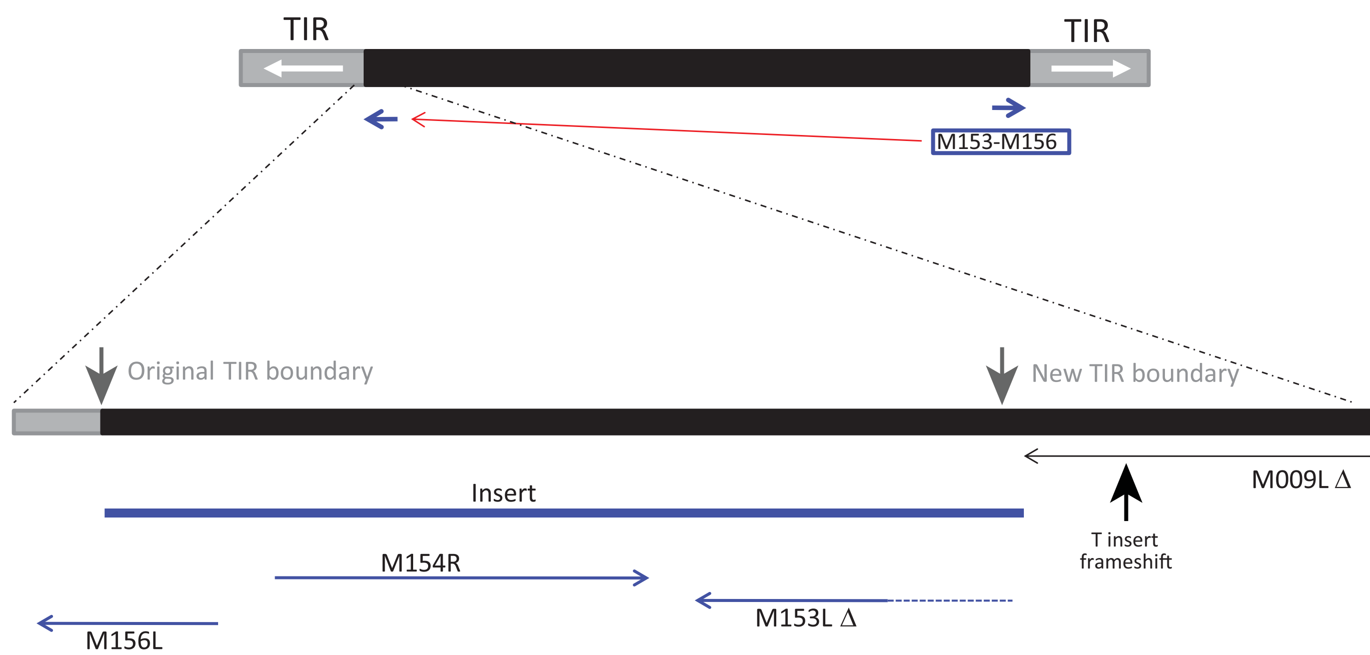
JVI Accepts published online ahead of print



JVI Accepts published online ahead of print



JVI Accepts published online ahead of print



A) Intergenic repeat sequence between M017L and M018L

```
Lu  (18237)   GTATGTAGGTATGTAG--------------------------------------------------------CTTAGTAGTTTAGT----------------GTAA (18263)
SLS (18236)   GTATGTAG----------------------------------------------------------------CTTAGTAGTTTAGT----------------GTAA (18254)
GV            GTATGTAG----------------------------------------------------------------CTTAGTAGTTTAGT----------------GTAA
KM13          GTATGTAG----------------------------------------------------------------CTTAGTAGTTTAGT----------------GTAA
Ur            GTATGTAG----------------------------------------------------------------CTTAGTAGTTTAGT----------------GTAA
BRK           GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
BEND          GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
WELL          GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGTAGTTTAGTGTAA
GUNG          GTATGTAGGTATGTAGGTATGTAGGTATGTAG----------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
SWH           GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
Meby          GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGT----------------GTAA
BRK/12-2-93   GTATGTAGGTATGTAG--------------------------------------------------------CTTAGTAGTTTAGT----------------GTAA
BRK/897       GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGTAGTTTAGTGTAA
OB1/406       GTATGTAGGTATGTAG--------------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
SWH/1209      GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
BD23          GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
SWH/8-2-93    GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
WS1/234       GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGTAGTTTAGTGTAA
WS1/328       GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
WS6/1071      GTATGTAGGTATGTAGGTATGTAGGTATGTAGGTATGTAGGTATGTAGGTATGTAGGTATGTAGGTATGTAGCTTAGTAGTTTAGT----------------GTAA
WS6 346       GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGTAGTTTAGTGTAA
BD44          GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
OB2/W60       GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGT--------GTAA
OB3/Y317      GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGTAGTTTAGTGTAA
OB3/1120      GTATGTAGGTATGTAGGTATGTAGGTATGTAGGTATGTAG--------------------------------CTTAGTAGTTTAGT----------------GTAA
SWH/805       GTATGTAGGTATGTAGGTATGTAG------------------------------------------------CTTAGTAGTTTAGTAGTTTAGTAGTTTAGTGTAA
```

B) Intergenic deletions between M002L/R and M003L/R

```
Lu  (2727)    CCGAATTAGAGAGGAGCCAATTAGGAGCCGAATTAGGAGCCGAATTAGAGAGGAGCCGAATTAGGAGCCGAATTAGAGAGGAGCCA (2822)
SLS (2736)    CCGAATTAGAGAGGAGCCAATTAGGAGCCGAATTAGGAGCCGAATTAGAGAGGAGCC-AATTAGGAGCCGAATTAGAGAGGAGCCA (2820)
BEND          CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
WELL          CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
GUNG          CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
BRK/807       CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
SWH/8-2-93    CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
WS1/234       CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
SWH/805       CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
OB3/Y317      CCGAATTAGAGAGGAGCCAATTAGGAGCCG----------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
WS6/346       CCGAATTAGAGAGGAGCC----------------------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
OB3/1120      CCGAATTAGAGAGGAGCC----------------------------------------AATTAGGAGCCGAATTAGAGAGGAGCCA
```

JVI Accepts published online ahead of print

JVI Accepts published online ahead of print

A) *M008.1L/R* late promoter

5′ TGCGACGTTTTTTTTGAGGGT*TAA****ATG***A 3′

B) *M057L* late promoter

5′ TTTTTTTTGTGA*TAA****ATG*** 3′

C) *M000.5L/R* possible late promoter

5′ ATTCTACGCGGACC*TCC****ATG***G 3′

D) *M138L* early promoter

5′ GTAGAC*TAAAACAC-AAAAAAAA*TCTTGCTTCTGCGAT**ATG** 3′

E) *M153R* early promoter

5′ CTTTTTGTTTATGGGG*AAACTCTAAAAAAAA****A***TTGTCAATTAAAGTAAATAGGTTGTGTAAAC**ATG** 3′