Genome Sequence of a Human Tumorigenic Poxvirus: Prediction of Specific Host Response–Evasion Genes

Tatiana G. Senkevich, Joachim J. Bugert, Jerry R. Sisler, Eugene V. Koonin, Gholamreza Darai, Bernard Moss*

Molluscum contagiosum virus (MCV) commonly causes asymptomatic cutaneous neoplasms in children and sexually active adults as well as persistent opportunistic acquired immunodeficiency syndrome (AIDS)–associated disease. Sequencing the 190–kilobase pair genome of MCV has now revealed that the virus potentially encodes 163 proteins, of which 103 have homologs in the smallpox virus. MCV lacks counterparts to 83 genes of the smallpox virus, including those important in suppression of host responses to infection, nucleotide biosynthesis, and cell proliferation. MCV possesses 59 genes that are predicted to encode previously uncharacterized proteins, including major histocompatibility complex class I, chemokine, and glutathione peroxidase homologs, which suggests that there are MCV-specific strategies for coexistence with the human host.

The poxvirus family of large double-stranded DNA viruses is composed of numerous species that infect vertebrates and insects, but only two of which are specific for humans (*1*). Variola virus (VAR), a member of the orthopoxvirus genus, was responsible for smallpox until its recent eradication through immunization with the closely related, and attenuated, vaccinia virus (VV). The other human poxvirus, MCV, has a worldwide distribution and commonly produces 3- to 5-mm papules that may persist in the skin of young children and sexually active adults for months to years before spontaneously regressing (*2*). MCV is also responsible for an extensive and essentially untreatable opportunistic disease that frequently occurs in individuals with AIDS (*2*). MCV infection typically elicits a weak immune response and almost no inflammatory reaction around the hyperplastic, virus-filled epidermal lesions, even in immunocompetent persons. Little is known of the basis for the MCV-host relation, because the virus has not been grown in tissue culture cells or animals and it replicates only inefficiently in human skin grafted to immunodeficient mice (*3*). MCV is the sole member of the molluscipoxvirus genus and is related only distantly to VAR or VV, as revealed by a nearly twofold higher GC content (63% in MCV, 34% in VAR and VV), lack of DNA cross-hybridization or immunological cross-reactivity, and sequence differences in the short segments of MCV DNA that have heretofore been analyzed (*4*).

MCV major subtype 1 DNA (*5*) was sequenced by primer walking with an automatic sequencer (Applied Biosystems, model 373A). Taq polymerase cycle sequencing protocols with inosine triphosphate in place of guanosine triphosphate were used to avoid early termination of sequencing reactions and compressions caused by the high GC content of the MCV DNA (*6*). Each base was determined at least once from each strand, with no discrepancies between the two strands. The absence of frameshifts in the MCV genes that have homologs in orthopoxviruses suggested less than one frameshift error per 100,000 nucleotides. The 190,289–base pair (bp) sequence of MCV DNA, comprising the entire genome with the exception of the covalently closed terminal hairpin loops (*7*, *8*), was deposited in GenBank (accession number U60315). The MCV DNA has a GC content of 63.3%, resulting in a paucity of stop codons and, accordingly, a large number of open reading frames (ORFs). Altogether, the MCV genome contains 588 ORFs longer than 60 codons. With a combination of sequence similarity searches, statistical pre-

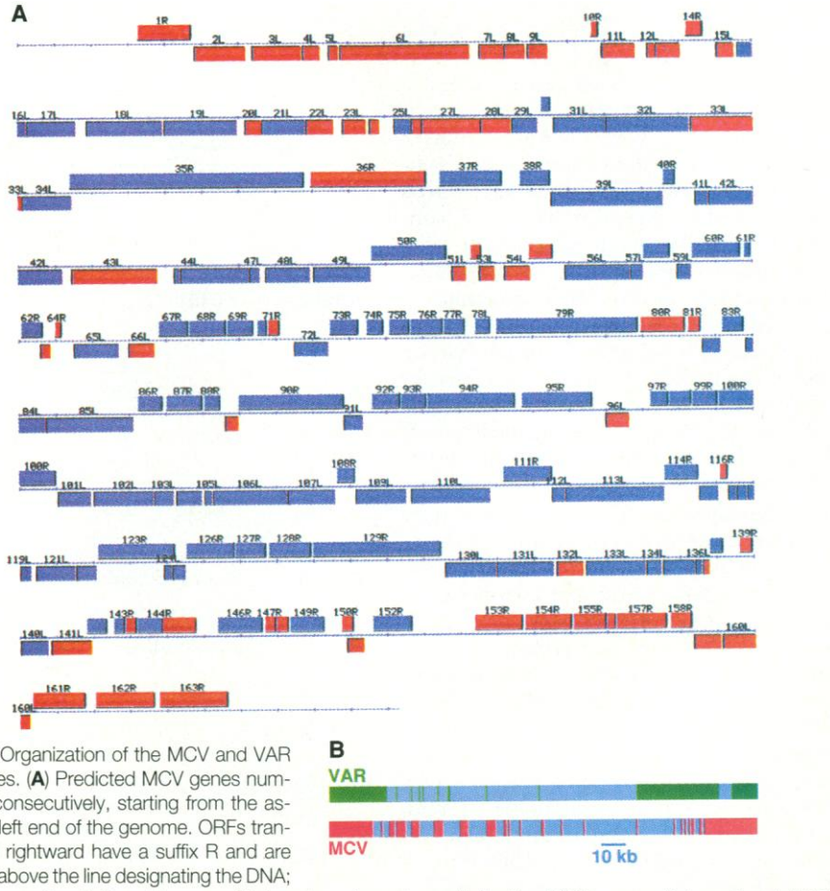

Fig. 1. Organization of the MCV and VAR genomes. (**A**) Predicted MCV genes numbered consecutively, starting from the assigned left end of the genome. ORFs transcribed rightward have a suffix R and are shown above the line designating the DNA; ORFs transcribed leftward have a suffix L and are shown beneath the line. MCV genes with homologs in VAR or VV are shown by blue bars; nonhomologous genes are shown by red bars. The ORFs 123 and 124, and 144 and 145, are completely overlapping. Distances between ticks on the line correspond to 1 kb. (**B**) Comparison of the genome organization of MCV and VAR. Genes common to MCV and VAR are in blue; VAR genes without MCV homologs are in green; MCV genes without VAR homologs are in red.

T. G. Senkevich, J. J. Bugert, J. R. Sisler, B. Moss, Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892–0455, USA.
E. V. Koonin, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20892, USA.
G. Darai, Institut für Medizinische Virologie, Universität Heidelberg, 69120 Heidelberg, Germany.
*To whom correspondence should be addressed.

diction of coding regions, prediction of signal peptides and transmembrane helices in putative proteins, and analysis of putative promoters (9), we identified 163 ORFs (mostly nonoverlapping) as a conservative estimate of the gene complement of MCV (Fig. 1A). The corresponding numbers are 187 and 198 for VAR and VV, respectively (*10*).

Of the 163 predicted MCV proteins, 104 are homologous to proteins of VAR or VV; 55 of these conserved poxvirus genes have a known or predicted function in VV reproduction (*10*). All but one of these functionally characterized proteins either contribute to replication or expression of the viral genome or are structural components of the virion. The sequence identities of the predicted MCV proteins and their orthopoxvirus homologs vary from 15 to 76% (*11*). The genes coding for conserved proteins cluster in the central part of the MCV genome (Fig. 1A). Moreover, the order of the conserved genes is the same in MCV and the orthopoxviruses, even though this region of the MCV genome contains several apparent deletions of poxvirus genes and insertions of other genes (Fig. 1A). The only two rearrangements affect nonessential genes that encode 3β-hydroxysteroid dehydrogenase–steroid D5-D4 isomerase and a large putative membrane protein that is conserved in MCV and VAR but not in VV. All the VV genes known to be essential for virus growth in cell culture, with a single possible exception (*12*), are conserved in MCV, which indicates that the basic mechanisms of genome replication and expression used by MCV are similar to those of orthopoxviruses.

Of the 88 VAR and VV genes that are absent from MCV, 64 are nonessential for replication in cell culture, 1 is conditionally essential (*12*), and 23 are uncharacterized. Most of these VAR and VV genes are located in the terminal regions of the orthopoxvirus genomes (Fig. 1B). One class of nonessential genes that is missing in MCV includes all six orthopoxvirus genes that encode nucleotide biosynthesis enzymes (*13*). Their absence is unusual for large DNA viruses and may in part determine tissue tropism. Another group of orthopoxvirus genes with no homologs in MCV includes ≥20 that encode proteins that modulate the host response to infection (*14*). These proteins include an epidermal growth factor homolog, a complement-binding protein, an interleukin-1 convertase inhibitor, additional protease inhibitors, a double-stranded RNA–binding protein important in inhibition of the interferon-induced double-stranded RNA–dependent protein kinase and ribonuclease, and soluble receptors for tumor necrosis factor, interleukin-1, interferon-α, and interferon-γ. The only orthopoxvirus gene of this group conserved in MCV encodes a 3β-hydroxysteroid dehydrogenase–steroid D5-D4 isomerase homolog.

The orthopoxvirus genes that are missing in MCV are replaced by 59 similarly located but nonhomologous genes (Fig. 1B). Searches of protein and nucleic acid sequence databases (9) showed that some of these 59 genes have cellular homologs with known functions, suggesting that the corresponding viral genes modulate host responses to infection (Table 1) (*15*). The alignment of one MCV protein and its cellular homolog revealed 74% amino acid identity, whereas most such alignments, even though statistically significant, showed 20 to 25% identity (Table 1); the latter relatedness values are similar to those for many orthopoxvirus proteins that contribute to virus-host interactions and whose functions predicted by sequence similarities have been experimentally validated (*14*). Most of the MCV-specific genes are predicted to be expressed early in infection (of the 59 such genes, 45 have early-promoter consensus sequences, 10 have late promoters, and 4 have no recognizable promoter), as are most orthopoxvirus genes that participate in virus-host interactions.

For at least three MCV genes, it is possible to predict a specific function in combating host defenses. An MCV homolog of the major histocompatibility complex (MHC) class I heavy chain, encoded by gene MC080R, although similar to MHC proteins in primary and probably tertiary structure (*16*), lacks most of the conserved amino acids important in peptide binding (Fig. 2A) (*17*) and, therefore, may not present peptides on the infected cell surface. Another unusual feature of the MCV protein is the presence of two predicted

Table 1. MCV genes not present in other poxviruses. Only those genes whose products showed statistically significant sequence similarity to proteins contained in current databases are included. The predicted genes are numbered consecutively from the left end of the genome. L indicates putative genes transcribed leftward and R indicates putative genes transcribed rightward. Database accession numbers are provided for homologs.

MCV		Protein size (amino acids)	Homolog	Homolog size (amino acids)	Identity* (%)
ORF	Genome location (bp)				
MC002L†	4849–6201	451	Human signaling lymphocytic activating molecule (GenBank HSU33017_1)	335	20.4
			MC161R†	478	20.0
			MC162R†	532	22.7
MC003L‡	6401–7735	445	Human carcinoembryonic antigen homolog CGM2 (PIR A55811)	265	16.8
			MC157R‡	452	20.1
MC013L	17,403–18,080	226	Human DnaJ-like protein (SWISS-PROT P31689)	397	25.3
MC026L	30,753–31,001	83	Hypothetical yeast protein with a modified RING finger (PIR S52511)	165	23.6
MC033L	38,362–40,098	579	*Xenopus* class I histocompatibility antigen (PIR S39606)	294	19.1
MC063L	80,621–80,869	83	Murine nonhistone chromatin phosphoprotein HMGI-C (PIR S22597)	108	21.8
MC066L	83,029–83,688	220	Human glutathione peroxidase (SWISS-PROT P07203)	201	74.2
MC080R	96,986–98,170	395	Rat MHC class I protein (PIR S25536)	353	24.5
MC132L	154,658–155,344	229	Human EST§ (GenBank R77178)	109	46.2
MC148R	166,992–167,303	104	Putative chicken CC chemokine (GenBank CHKCYTO_1)	96	25.2
MC163R‖	183,788–185,647	620	Sweet potato superoxide dismutase (SWISS-PROT Q07796)	151	23.2

*Percentage identity in the aligned region includes the entire length of the shorter protein. †MC002L, MC161R, and MC162R constitute a family of putative membrane receptors. ‡MC003L and MC157R constitute a second family of putative membrane receptors. §EST, expressed sequence tag; this clone encodes a partial sequence of the respective hypothetical protein. ‖Predicted membrane protein with a superoxide dismutase–related domain.

signal peptides that precede the one that aligns with the MHC signal peptide. This viral protein may compete with the assembly, transport, or function of host MHC proteins, thus precluding the presentation of MCV-specific peptides. Alternatively, a viral MHC homolog expressed on the cell surface may deceive natural killer cells (*18*). An MHC protein homolog is also encoded by human cytomegalovirus (*19*); in contrast to the MCV protein, the cytomegalovirus protein contains the conserved amino acid residues required for peptide binding (*20*).

The protein encoded by MCV gene MC148R belongs to the CC family of chemokines and is shown aligned with human macrophage inflammatory protein 1β (MIP-1β) in Fig. 2B. Although the MCV protein is predicted to retain the disulfide bonding pattern and the general structure of chemokines, the NH_2-terminal region implicated in monocyte activation (*21*) is deleted. Truncated analogs of CC chemokines have been shown to bind their receptors but have no activity (*21*). Therefore, the MCV protein is predicted to be a chemokine antagonist and may thereby inhibit inflammation.

MCV gene MC066L encodes a protein that is 74% identical to human glutathione peroxidase and contains the TGA opal codon that probably specifies the active site selenocysteine (Fig. 2C). This protein may protect the virus and components of infected cells from oxidative damage by peroxides, whose formation may be stimulated by infection (*22*). Virus-encoded enzymes of this class have not previously been described.

It was proposed that the MCV-induced proliferation of keratinocytes is mediated by a homolog of vaccinia virus growth factor, a member of the epidermal growth factor family (*8*, *23*). Our analysis of the MCV genome did not reveal a protein related to any known growth factor. However, one or more of the predicted secreted proteins encoded by MCV may represent previously uncharacterized growth factors.

Among the putative MCV proteins whose roles in virus-host interaction are not readily predictable, three are of particular interest (Table 1). A DnaJ chaperone homolog (MC013L) is the second virus-encoded member of this ubiquitous protein family, after the DnaJ-related domain of polyomavirus T antigen (*24*). A role of the MCV-encoded chaperone in a late event, such as virion morphogenesis, is suggested by the late-promoter consensus sequence of the gene. Another MCV gene (MC026L) encodes a small protein that belongs to a previously undescribed family with a modified RING finger domain, a domain that is conserved throughout eukaryotes (Fig. 2D). This MCV protein does not show any specific similarity to the proteins with classic RING fingers encoded by orthopoxviruses and leporipoxviruses (*10*, *25*), with the exception of the motif proper. RING finger proteins are encoded by most of the large DNA viruses and, in some instances, appear to be important in transcriptional regulation or replication (*26*). MCV also contains a gene (MC063L) that encodes a protein related to the nonhistone chromatin protein HMGI-C. A role for the virus protein in modulating host cell transcription seems possible, although sequence differences in the DNA binding motif suggest a different DNA binding specificity than that of HMGI-C (*27*).

In conclusion, MCV lacks nearly all of the genes of VAR that encode proteins that mediate virus-host interactions, but has its own set of genes that likely play such a role. The different gene repertoires almost certainly contribute to the vastly different infection strategies of MCV and VAR. VAR causes a fulminant, systemic disease that is either quickly controlled by

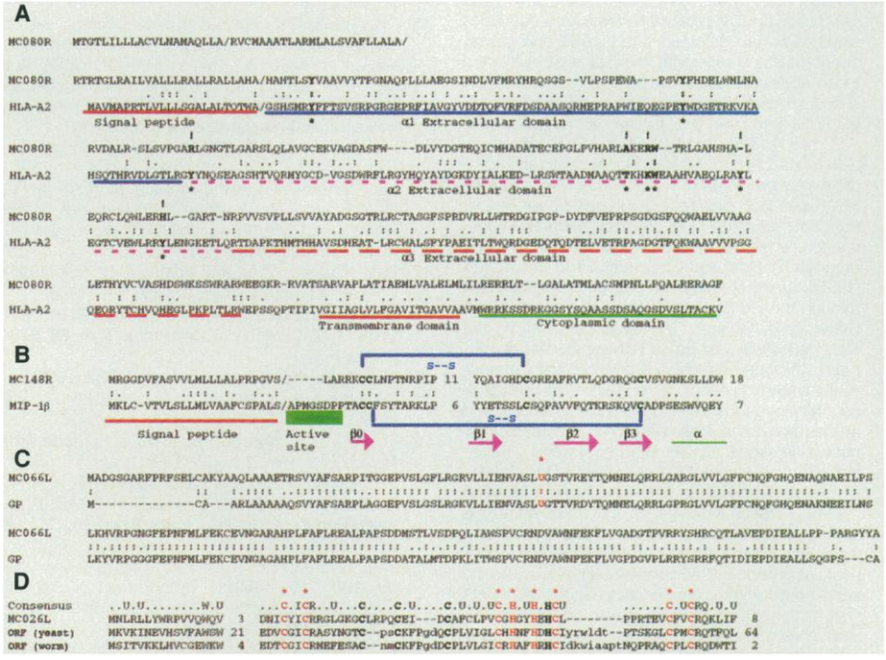

Fig. 2. Alignments of four MCV-specific proteins with their nonviral homologs. Alignments were constructed with the ALIGN program (*11*). Identities are designated by colons, and conservative replacements by dots; dashes represent gaps introduced to optimize alignment. (**A**) The MHC class I protein homolog (MC080R) aligned with human lymphocyte antigen (HLA) A2 (1A02_HUMAN, SWISS-PROT). The domains are based on the x-ray structure of the A2 antigen (*28*). The amino acids of HLA-A2 that participate directly in peptide binding are indicated by asterisks, with those that are replaced in the viral protein indicated by exclamation points. In the MCV protein, two additional predicted signal peptides precede the conserved signal peptide designated by a bar; the respective putative cleavage sites are shown by slashes. Even though the putative transmembrane domain in the MCV protein is less hydrophobic than its counterpart in the A2 antigen, it was confidently predicted by the PHDhtm program (*9*) as a transmembrane helix. (**B**) The CC chemokine homolog (MC148R) aligned with human MIP-1β (MI1B_HUMAN, SWISS-PROT). The disulfide bonding pattern, the signal peptide, and the region immediately downstream of the signal peptide implicated in chemokine activity (*21*) are indicated. The secondary structure elements (β0 to β3, α) are shown based on the x-ray structure of MIP-1β (*29*). Distances between the aligned segments and those from the protein COOH-termini are indicated by numbers. (**C**) The putative MCV-encoded glutathione peroxidase (MC066L) aligned with its human homolog (GP, GSHC_HUMAN, SWISS-PROT). The active site selenocysteine encoded by an opal codon (TGA) is designated **U** (asterisk). (**D**) MCV-encoded member (MC026L) of the previously undescribed family of small proteins with a modified RING finger. ORF (yeast) and ORF (worm) indicate uncharacterized open reading frames SCCHRIV42_20 and CEF35G12_8 (both from GenBank), respectively. The modified motif consists of a $C_3H_2C_3$ pattern as compared to the C_3HC_4 pattern in the classical RING finger motif (*26*). Regions of uncertain alignment are shown in lowercase. Distances between the aligned segments and those from the protein COOH-termini are indicated by numbers. Conserved cysteines and histidines are in bold, and those that constitute the motif are shown in red and are indicated by asterisks. The consensus line shows conserved residues, with U indicating a bulky hydrophobic residue. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y. Tyr.

the immune system or results in death. In contrast, MCV produces slow-growing epidermal neoplasms that persist for long periods of time with little immune response. The availability of the MCV genome sequence may allow studies of previously unidentified host defense mechanisms, provide new diagnostic reagents to determine the true incidence of this worldwide infection, and lead to the development of therapeutics to treat the increasingly frequent and persistent MCV disease associated with AIDS and other immunodeficiency conditions.

REFERENCES AND NOTES

1. B. Moss, in *Fields Virology,* B. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, 1966), pp. 2637–2672; F. Fenner, *ibid.*, pp. 2703–2737; _____, D. A. Henderson, I. Arita, J. Jezek, I. D. Ladnyi, *Smallpox and Its Eradication* (World Health Organization, Geneva, 1988).
2. C. D. Porter *et al.*, *Mol. Cell. Biol. Hum. Dis. Series* **1**, 233 (1992); S. L. Gottlieb and P. L. Myskowski, *Int. J. Dermatol.* **33**, 453 (1994).
3. R. M. L. Buller, J. Burnett, W. Chen, J. Kreider, *Virology* **213**, 655 (1995).
4. N. W. Blake, C. D. Porter, L. C. Archard, *J. Virol.* **65**, 3583 (1991); J. J. Bugert *et al.*, *Virology* **192**, 391 (1993); R. P. Hadasch *et al.*, *Intervirology* **36**, 32 (1993); K. C. Sonntag *et al.*, *Virology* **210**, 471 (1995).
5. Previously cloned MCV DNA fragments have been described [G. Darai *et al.*, *J. Med. Virol.* **18**, 29 (1986); J. J. Bugert, A. Rosen-Wolff, G. Darai, *Virus Genes* **3**, 159 (1989); J. J. Bugert and G. Darai, *J. Med. Virol.* **33**, 211 (1991)]. Digestion of MCV DNA with mung bean nuclease, polymerase chain reaction (PCR) amplification, and cloning of near-terminal fragments will be described elsewhere. (T. G. Senkevich *et al.*, in preparation).
6. MCV DNA cloned in pAT153 or pACYC184 was extracted with the use of Promega Miniprep or Qiagen Midiprep kits. Oligonucleotide primers of 16 to 22 nucleotides were designed manually and tested for dimer formation and nonspecific hybridization with the AMPLIFY program (B. Engels, University of Wisconsin). Sequencing reactions were performed with the Prism Ready Reaction Dyedeoxy Terminator kit (Applied Biosystems) or an improved version containing AmpliTaq FS. Annealing and extension of primers were performed at 60°C. Approximately 5% of the MCV DNA with the highest local GC content was sequenced in the presence of 5% dimethylsulfoxide and, in several instances, also a fourfold increased concentration of the AmpliTaq polymerase and a twofold increased concentration of the deoxyribonucleoside triphosphate substrates.
7. R. P. Parr, J. W. Burnett, C. F. Garon, *Virology* **81**, 247 (1977).
8. C. D. Porter and L. C. Archard, *J. Gen. Virol.* **68**, 673 (1987).
9. Initial screening of the nonredundant protein sequence database (National Center for Biotechnology Information, NIH) was performed with the BLASTX program, which translates the test nucleotide sequence in six reading frames; such analysis not only identifies protein-coding genes but also allows the detection of frameshift errors [W. Gish and D. J. States, *Nature Genet.* **3**, 266 (1993)]. Subsequently, all ORFs containing >60 codons were translated and compared with the protein database with the BLASTP program [S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990)] and with the nucleotide database translated in six frames and including the current database of expressed sequence tags (dbEST) with the TBLASTN program [S. F. Altschul, M. S. Boguski, W. Gish, J. C. Wootton, *Nature Genet.* **6**, 119 (1994)]. Alignments with relatively low significance (probability of occurring by chance of $>10^{-3}$) were additionally analyzed by motif search with the CAP and MoST programs [R. L. Tatusov, S. F. Altschul, E. V. Koonin, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12091 (1994)] and by construction of multiple alignments with the MACAW program [G. D. Schuler, S. F. Altschul, D. J. Lipman, *Prot. Struct. Funct. Genet.* **9**, 180 (1991)]. The putative MCV proteins for which no sequence similarities were detected with BLAST and subsequent motif and alignment analyses were subjected to an additional database screening with a highly sensitive version (ktuple = 1) of the FASTA program [W. R. Pearson and D. J. Lipman, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444 (1988); W. R. Pearson, *Genomics* **11**, 635 (1991)]. Signal peptides in proteins were predicted with the Signalp V1.0 program (H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, *Prot. Struct. Funct. Genet.* **24**, 165 (1996), and transmembrane helices were predicted with the PHDhtm program [B. Rost and C. Sander, *Prot. Sci.* **4**, 521 (1995)]; both of these programs were accessed through the ExPasy World Wide Web server at the University of Geneva. The statistical prediction of MCV genes was performed with the GeneMark program, which derives nonhomogeneous Markov models for a learning set of coding sequences and ordinary Markov models for noncoding sequences, and applies them to gene identification in uncharacterized nucleotide sequences [M. Borodovsky and J. McIninch, *Comput. Chem.* **17**, 123 (1993);M. Borodovsky, K. E. Rudd, E. V. Koonin, *Nucleic Acids Res.* **22**, 4756 (1994)]. Putative MCV genes that have homologs among orthopoxvirus genes were used as the learning set of coding regions, and MCV DNA sequences containing no long ORFs were used as the learning set of noncoding regions. Additionally, the MCV DNA was screened with models derived for GC-rich human coding sequences (M. Borodovsky, personal communication).
10. S. J. Goebel *et al.*, *Virology* **179**, 247 (1990); *ibid.*, p. 517 (appendix); R. F. Massung *et al.*, *Nature* **366**, 748 (1993); R. F. Massung *et al.*, *Virology* **201**, 215 (1994); S. N. Shchelkunov, V. M. Blinov, L. S. Sandakhchiev, *FEBS Lett.* **319**, 80 (1993); S. N. Shchelkunov, R.F. Massung, J. J. Esposito, *Virus Res.* **36**, 107 (1995).
11. Pairwise alignments were constructed with the ALIGN program (W. R. Pearson, University of Virginia) based on a modified algorithm of E. Myers and W. Miller [*Comput. Appl. Biosci.* **4**, 11 (1988)]. In most instances, MCV proteins and their homologs from orthopoxviruses aligned along their entire lengths, but several MCV proteins had long extensions at either the NH_2- or COOH-terminus. The sequences of these extensions showed a biased amino acid composition and contained short repeats, which is typical for coding sequences of high GC content. To rule out the possibility that these extensions resulted from cloning artifacts, we verified the sequences for two of them (MC034R and MC123R; homologs of VV genes E4L and A18R, respectively) by sequencing the appropriate regions of the MCV genomic DNA by PCR. Both sequences were identical to those determined from the plasmid clones, thus confirming the authenticity of the extensions.
12. The VV gene B1R, for which there is no counterpart in MCV, encodes a protein kinase; mutations in this gene result in a temperature-sensitive DNA^- phenotype, the severity of which is host cell–dependent [R. E. Rempel and P. Traktman, *J. Virol.* **66**, 4413 (1992)].
13. These enzymes are thymidine kinase, thymidylate kinase, guanylate kinase, deoxyuridine triphosphatase, and two ribonucleotide reductase subunits. The guanylate kinase gene in VV and VAR contains a frameshift mutation [G. L. Smith, Y. S. Chan, S. T. Howard, *J. Gen. Virol.* **72**, 1349 (1991); B. Aguado, I. P. Selmes, G. L. Smith, *ibid.*, **73**, 2887 (1992)].
14. A. Alcami and G. L. Smith, *Immunol. Today* **16**, 474 (1995).
15. MCV encodes several putative membrane or secreted proteins (MC004L, MC009L, MC011L, MC024L, MC054L, MC089L, MC116R, MC156R, and MC158R) that do not show statistically significant similarity to any proteins in databases but that may be important in host-response modulation.
16. Homology modeling of the three-dimensional structure of the MHC class I homolog encoded by MCV was performed with the ProMod program [M. C. Peitsch and C. V. Jongeneel, *Int. Immunol.* **5**, 233 (1993)] and the SWISS-MODEL server [M. C. Peitsch, *Biotechnology* **13**, 658 (1995)].
17. D. R. Madden, *Annu. Rev. Immunol.* **13**, 587 (1995).
18. D. H. Raulet and W. Held, *Cell* **82**, 697 (1995).
19. S. Beck and B. G. Barrell, *Nature* **331**, 269 (1988); H. Browne, M. Churcher, T. Minson, *J. Virol.* **66**, 6784 (1992).
20. M. L. Fahnestock *et al.*, *Immunity* **3**, 583 (1995).
21. I. Clark-Lewis *et al.*, *J. Leukocyte Biol.* **57**, 703 (1995).
22. C. Michiels, M. Raes, O. Toussaint, J. Remacle, *Free Radical Biol. Med.* **17**, 235 (1994).
23. J. P. Brown *et al.*, *Nature* **313**, 491 (1985).
24. S. Zhang *et al.*, *EMBO J.* **11**, 3787 (1992); W. L. Kelley and S. J. Landry, *Trends Biochem. Sci.* **19**, 277 (1994).
25. T. G. Senkevich, E. V. Koonin, R. M. L. Buller, *Virology* **198**, 118 (1994); C. Upton *et al.*, *J. Virol.* **68**, 4186 (1994); T. G. Senkevich, E. J. Wolffe, R. M. L. Buller, *ibid.* **69**, 4103 (1995).
26. P. S. Freemont, *Ann. N.Y. Acad. Sci.* **684**, 174 (1993).
27. D. Thanos and T. Maniatis, *Cell* **83**, 1091 (1995); J. V. Falvo, D. Thanos, T. Maniatis, *ibid.*, p. 1101.
28. M. A. Saper, P. J. Bjorkman, D. C. Wiley, *J. Mol. Biol.* **219**, 277 (1991).
29. P. J. Lodi *et al.*, *Science* **263**, 1762 (1994).

30. We thank M. Borodovsky and W. S. Hayes for performing the analysis of the MCV DNA sequence with the GeneMark program; R. L. Tatusov for writing several programs that facilitated sequence analysis; R. M. L. Buller for MCV DNA and helpful discussions; M. Turner for assistance in obtaining MCV samples; S. H. Bryant, D. Landsman, M. Merchlinsky, and A. Grunhaus for helpful discussions; and J. R. Bennink and J. W. Yewdell for critical reading of the manuscript. Supported in part by a fellowship from the Alexander von Humboldt Stiftung to J.J.B. and a grant from the Deutsche Forschungsgemeinschaft (DA-142/10-1) to G.D.

1 April 1996; accepted 2 May 1996