ISSN 0026-8933, Molecular Biology, 2006, Vol. 40, No. 1, pp. 16–19. © Pleiades Publishing, Inc., 2006.
Original Russian Text © I.V. Babkin, S.N. Shchelkunov, 2006, published in Molekulyarnaya Biologiya, 2006, Vol. 40, No. 1, pp. 20–24.

PRIORITY COMMUNICATIONS

UDC 575.86:575.852'113

Time Scale of Poxvirus Evolution

I. V. Babkin and S. N. Shchelkunov

Vector State Research Center of Virology and Biotechnology, Kol'tsovo, Novosibirsk Region, 630559 Russia
e-mail: snshchel@vector.nsc.ru

Received September 1, 2005

Presented for publication by A.A. Mironov

Abstract—Unlike in vertebrates and RNA viruses, the molecular clock has not been estimated so far for DNA viruses. The extended conserved central region (102 kb) of the orthopoxvirus genome and the DNA polymerase gene (3 kb) were analyzed in viruses representing several genera of the family Poxviridae. Analysis was based on the known dating of the variola virus (VARV) transfer from Western Africa to South America and previous data on the phylogenetic relatedness of modern West African and South American isolates of VARV. The mutation accumulation rate was for the first time estimated for these DNA viruses at $(0.9–1.2) \times 10^{-6}$ substitutions per site per year. It was assumed that poxviruses diverged from an ancestor approximately 500,000 years ago to form the recent species and that the ancestor of the genus *Orthopoxvirus* emerged approximately 300,000 years ago and gave origin to the modern species approximately 14,000 years ago.

DOI: 10.1134/S0026893306010031

Key words: DNA viruses, Poxviridae, evolution, molecular clock

INTRODUCTION

The time scale of animal evolution is based on paleontological data. Comparing the nucleotide sequences for genes conserved among different animal taxa, it is possible to estimate the rate of substitutions that have arisen in the given locus during the divergence of the taxa from their ancestor. The mutation rate is assumed to be constant. The resulting estimate is known as the molecular clock [1, 2].

In the case of viruses with single-stranded genomic RNA, the mutation rate is extremely high and the rate of virus evolution can be estimated from the accumulation of changes in the nucleotide sequences of strains isolated within several years [3].

The mutation rate of DNA viruses is far more difficult to estimate. It is hardly possible to determine the molecular clock of DNA viruses by comparing the genome sequence for available isolates of one species; this is because mutations accumulate at a low rate and the isolation dates differ by no more than several decades. There is still no reliable dating of the divergence of virus species or genera from an ancestor of the corresponding family.

Analysis of the evolutionary relationships of DNA viruses belonging to different taxa yields phylogenetic trees lacking a time scale. The phylogenetic trees of some viruses are graphically similar to the trees constructed for particular genes of the host animals [4]. A hypothesis has been advanced on this evidence that viruses emerged during the early evolution of the host organisms and diverged into different taxa in parallel with divergent evolution of the hosts. Proceeding from this hypothesis, three modern Herpesviridae subfamilies of mammalian viruses have been determined to have originated 180–220 million years ago [5]. Yet analysis of a larger set of strains has shown that the phylogenetic relationships among viruses of the family do not always formally coincide with the relationships among their hosts [6].

We consider it incorrect to assume the same molecular clock for viruses and their natural hosts. It is more correct to assume that the ancestor of a particular virus family had initially a broad range of hosts from different taxa and that the progressive specialization of viruses to different hosts took place during their long-term coevolution [7]. The results of such evolutionary specialization are detectable by comparing the nucleotide sequence for viral genes. Yet the molecular clock of DNA viruses is usually impossible to establish.

The variola virus (VARV), which belongs to the genus *Orthopoxvirus* of the family Poxviridae, presents a unique situation. VARV (or its ancestor) caused human epidemics in Asia and Africa from long ago and in Europe from the 6th century. However, VARV was absent in America until the 16th century, when it was brought with slaves from West Africa to South America and caused devastating epidemics among indigenous populations. More recently, variola epidemics with a low mortality have been observed in South America [8].

Poxviruses whose nucleotide sequences were analyzed in this work

Genus	Species	Strain	Abbreviation	GenBank accession no.
Orthopoxvirus	Variola virus	India-1967	VARV-IND	X69198
		Bangladesh-1975	VARV-BSH	L22579
		Garcia-1966	VARV-GAR	Y16780
	Camelpox virus	CMS	CMLV-CMS	AY009089
		M-96	CMLV-M96	AF438165
	Cowpox virus	GRI-90	CPXV-GRI	X94355
		Brighton Red	CPXV-BRI	AF482758
	Ectromelia virus	Naval	ECTV-NAV	*
		Moscow	ECTV-MOS	AF012825
	Monkeypox virus	Zaire-96-I-16	MPXV-ZRE	AF380138
Yatapoxvirus	Yaba monkey tumor virus	Roswell Park-Yohn	YMTV-RPY	NC_005179
		YLD	YMTV-YLD	AJ293568
Capripoxvirus	Lumpy skin disease virus	Neethling 2490	LSDV-2490	AF325528
		Neethling vaccine LW 1959	LSDV-1959	AF409138
	Sheeppox virus	A	SPPV-A	AY077833
		TU-V02127	SPPV-TU	NC_004002
	Goatpox virus	G20-LKV	GTPV-G20	AY077836
		Pellor	GTPV-Pellor	AY077835
Suipoxvirus	Swinepox virus	17077-99	SWPV-99	AF410153
Leporipoxvirus	Myxoma virus	Lausanne	MYXV-LAU	AF170726
	Shope fibroma virus	Kasza	SFV-KAS	AF170722
Avipoxvirus	Fowlpox virus	FCV	FWPV-FCV	AF198100
	Canarypox virus	Wheatley C93	CNPV-WC93	NC_005309
Molluscipoxvirus	Molluscum contagiosum virus	Subtype 1	MOCV-SB1	MCU60315
Parapoxvirus	Orf virus	OV-IA82	ORFV-IA82	AY386263
	Bovine papular stomatitis virus	BV-AR02	BPSV-BV	NC_005337

* The sequence is available at www.sanger.ac.uk/Projects/Ectromelia_virus.

The mortality during variola epidemics varied broadly with time and geographic region, from 0.2 to 40%. Two subtypes of VARV are distinguished accordingly: VARV major, causing epidemics with a high (from 5 to 30–40%) mortality, and VARV minor, causing a lower mortality. South American VARV is known as VARV minor alastrim. Laboratory tests have shown that viruses of this subtype differ in several properties from African isolates of VARV minor [8]. Moreover, VARV minor alastrim clearly differs from other VARV strains from various regions of Asia and Africa on evidence of phylogenetic analysis of individual genes [9] and extended genome regions [10, 11]. Yet the results obtained so far do not allow a timing of evolutionary changes in the VARV genome.

We were the first to introduce the time factor in phylogenetic analysis of 63 VARV genomes, which were tested for restriction fragment length polymorphism [12]. The South American strains proved to be closely related to the West African strains. Assuming that the divergent evolution of these geographically isolated variants of VARV started in the 16th century (approximately 400 years ago), we deduced that the West African VARV diverged from its ancestor approximately 1100–1300 years ago.

In this work, we applied the resulting time scale to the estimation of the poxvirus molecular clock by analyzing the conserved central region (approximately 102 kb) of the orthopoxvirus genome and the DNA polymerase gene (2967–3039 bp) in different genera of the family Poxviridae.

EXPERIMENTAL

The **Poxvirus nucleotide sequences** used in this work were extracted from GenBank (table).

Phylogenetic analysis was performed and the mutation accumulation rate was estimated using the programs ClustalX v. 1.81 [13], BioEdit v. 7.0.0 [14],

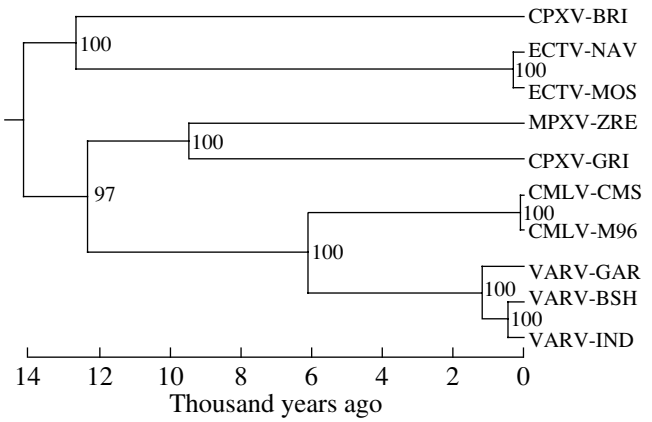

Fig. 1. Phylogenetic tree constructed by analyzing the genome region highly conserved in the genus *Orthopoxvirus* and bounded by *C8L* and *A24R* according to the nomenclature of VARV strain India-1967 [8]. Here and in Fig. 2: Analysis was performed by the maximum likelihood method with a molecular clock [15]. The significance of knots is indicated. Strains are designated as in the table.

SEQBOOT, CONSENSE, and DnaMLK with an optional equalizing of the probabilities of sequence input orders from the PHYLIP package v. 3.6 [15].

RESULTS AND DISCUSSION

The family Poxviridae includes subfamilies of vertebrate (Chordopoxvirinae) and insect (Entomopoxvirinae) viruses. The Chordopoxvirinae are divided into eight genera: one (*Avipoxvirus*) containing avian viruses and seven containing mammalian viruses. Of the latter, viruses of the genus *Orthopoxvirus* have been studied most extensively, because this genus includes the human pathogens VARV, monkeypox virus, and cowpox virus, along with the vaccinia virus, which serves as a live vaccine against variola and other orthopoxvirus infections [8].

Phylogenetic analysis of data on restriction fragment length polymorphism in a large set of VARV strains has revealed a close relationship between West African and South American isolates [11, 12]. As VARV was brought from West Africa to South America in the 16th century, we deduced that the West African variant diverged about 1100–1300 years ago from a common ancestor of modern VARV [12].

With the resulting time scale of VARV evolution, we phylogenetically analyzed the extended central conserved (genus-specific) region in the sequenced genomes of several species of the genus *Orthopoxvirus*. Using the ClustalX [13] and BioEdit [14] programs, a 102,374-bp sequence (102 genes) was aligned for ten strains of different orthopoxvirus genes. The alignment was analyzed by the maximum likelihood method with a molecular clock, using the PHYLIP package [15]. The significance of the resulting tree was tested by bootstrap analysis. The results showed with a high significance that the modern orthopoxvirus species separated from their ancestor approximately 14,000 years ago (Fig. 1).

The orthopoxvirus strains grouped by species on the phylogenetic tree (Fig. 1). The only exception was the cowpox virus: one strain clustered together with the monkeypox virus and the other, with the ectromelia virus. These findings and our previous phylogenetic analysis of the chemokine-binding protein in 85 orthopoxvirus strains [9] make it possible to assume that the cowpox virus species consists of at least two subspecies.

The camelpox virus is the closest relative of VARV. The two species, each strongly host-specific, diverged from a common ancestor (which was probably zoonotic) approximately 6000 years ago. The divergence coincided with the appearance of large human settlements, the domestication of ungulates, and their accumulation in large herds, which would have expedited the spread of infection among humans and animals.

It is unfeasible to compare extended genome regions for poxviruses of different genera because of considerable intergeneric differences in their organization. To date the evolutionary divergence of vertebrate poxvirus genera, phylogenetic analysis was performed with the nucleotide sequence (2967–3039 bp) of the conserved DNA polymerase gene, using the same method as with the genome region conserved among orthopoxviruses (Fig. 2). We analyzed the nucleotide sequences of 26 strains of different vertebrate orthopoxviruses representing eight genera (table). The strains proved to cluster by genera. Most branches of the resulting tree were highly significant, with the exception of viruses of the genus *Yatapoxvirus*. The separation of species of the genus *Orthopoxvirus* was less reliable with the single DNA polymerase gene (Fig. 2) than with the genome region which included 102 genes (Fig. 1).

The results shown in Fig. 2 demonstrate that two evolutionary branches diverged from an ancestral virus approximately 500,000 years ago. One of these branches is now represented by two genera—*Parapoxvirus* and *Molluscipoxvirus*—whose DNA is characterized by a high (64.0–64.5%) GC content. The second branch combines the other genera and is characterized by a low (25.0–43.6%) GC content in viral DNA. Avian viruses of the genus *Avipoxvirus* diverged from mammalian viruses with a low GC content approximately 420,000 years ago. Although the separation of the modern orthopoxvirus species from a common ancestor was relatively recent (Fig. 1), the genus *Orthopoxvirus* diverged from the other poxivurses approximately 300,000 years ago (Fig. 2).

Estimation of the molecular clock showed that mutations accumulate in poxviruses at a rate of $(0.9–1.2) \times 10^{-6}$ substitutions per site (nucleotide) per year. This estimate is three orders of magnitude lower than the mutation rate of viruses having single-stranded genomic RNA [3] and three to four orders of magnitude higher than the

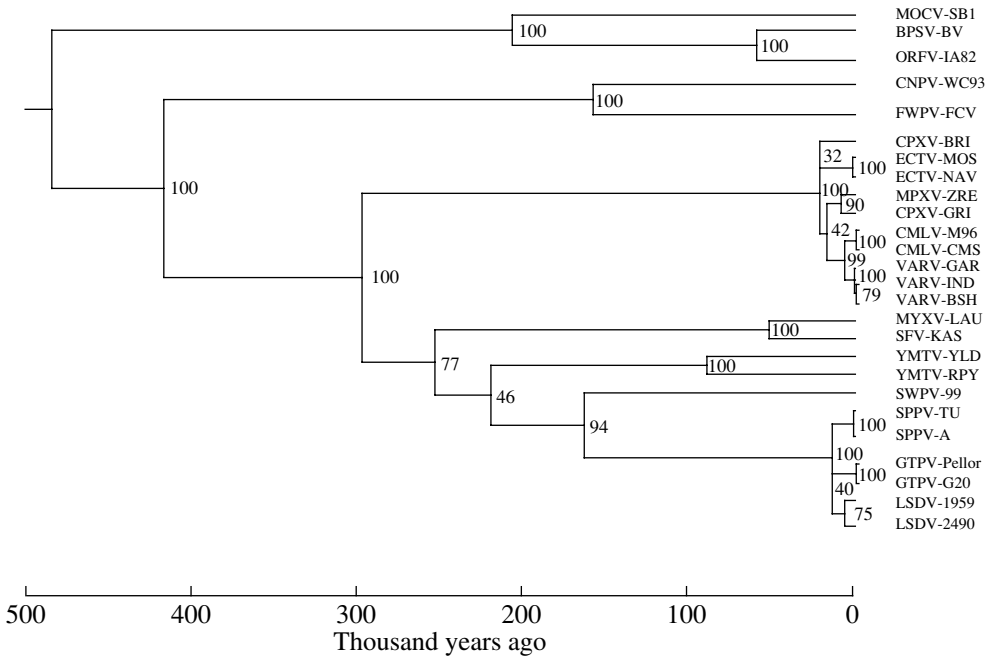

Fig. 2. Phylogenetic tree constructed by analyzing the nucleotide sequence of the DNA polymerase gene in vertebrate poxviruses.

evolutionary rate of animal chromosomal genes [2]. The poxvirus molecular clock objectively reflects the fact that the alternation of generations proceeds at a far greater rate in viruses than in their hosts.

Thus, we were the first to estimate the molecular clock for DNA viruses exemplified by the family Poxviridae. Our data call into question the hypothesis that modern viruses originate from ancestors that emerged and started divergent evolution hundreds or tens of million years ago, simultaneously with the development of new animal taxa [4, 5].

REFERENCES

1. Kumar S., Hedges S.B. 1998. A molecular timescale for vertebrate evolution. *Nature*. **392**, 917–920.
2. Russell P. 1998. *Genetics*. 5th ed. Menlo Park, CA: Addison Wesley Longman Inc.
3. Jenkins G.M., Rambaut A., Pybus O.G., Holmes E.C. 2002. Rates of molecular evolution in RNA viruses: A quantitative phylogenetic analysis. *J. Mol. Evol*. **54**, 152–161.
4. McGeoch D.J., Cook S. 1994. Molecular phylogeny of the alphaherpesvirinae subfamily and a proposed evolutionary timescale. *J. Mol. Biol*. **238**, 9–22.
5. McGeoch D.J., Cook S., Dolan A., Jamieson F.E., Telford E.A.R. 1995. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J. Mol. Biol*. **247**, 443–458.
6. McGeoch D.J., Gatherer D. 2005. Integrating reptilian herpesviruses into the family Herpesviridae. *J. Virol*. **79**, 725–731.
7. Herniou E.A., Olszewski J.A., O'Reilly D.R., Cory J.S. 2004. Ancient coevolution of baculoviruses and their insect hosts. *J. Virol*. **78**, 3244–3251.
8. Shchelkunov S.N., Marennikova S.S., Moyer R.W. 2005. *Orthopoxviruses Pathogenic for Humans*. Berlin: Springer.
9. Mikheev M.V., Feshchenko M.V., Shchelkunov S.N. 2004. Phylogenetic analysis of the gene for the orthopoxvirus chemokin-binding protein. *Mol. Genet. Mikrobiol. Virusol*. **1**, 29–36.
10. Shchelkunov S.N., Totmenin A.V., Babkin I.V., Safronov P.F., Ryazankina O.I., Petrov N.A., Gutorov V.V., Uvarova E.A., Mikheev M.V., Sisler J.R., Esposito J.J., Jahrling P.B., Moss B., Sandakhchiev L.S. 2001. Human monkeypox and smallpox viruses: Genomic comparison. *FEBS Lett*. **509**, 66–70.
11. Babkina I.N., Babkin I.V., Marennikova S.S., Sandakhchiev L.S., Shchelkunov S.N. 2004. Comparative restriction enzyme analysis of the genome in variola virus strains from the Russian collection. *Mol. Biol*. **38**, 429–436.
12. Babkina I.N., Babkin I.V., Li Yu., Ropp S., Klein R., Damon I., Esposito J., Sandakhchiev L.S., Shchelkunov S.N. 2004. Phylogenetic comparison of the genome in different variola virus strains. *Dokl. Akad. Nauk*. **398**, 818–822.
13. Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. **25**, 4876–4882.
14. Hall T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser*. **41**, 95–98.
15. Felsenstein J. 1989. PHYLIP: Phylogeny Inference Package (Version 3.2). *Cladistics*. **5**, 164–166.