Virology 512 (2017) 124–131

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Shared ancestry of herpes simplex virus 1 strain Patton with recent clinical isolates from Asia and with strain KOS63

Aldo Pourchet[a,1], Richard Copin[b,1], Matthew C. Mulvey[c], Bo Shopsin[a,b], Ian Mohr[a], Angus C. Wilson[a,*]

[a] Department of Microbiology, New York University School of Medicine, New York, NY, USA
[b] Department of Medicine, New York University School of Medicine, New York, NY, USA
[c] BeneVir Biopharm, Inc., Gaithersburg, MD, USA

ARTICLE INFO

Keywords:
Herpes simplex virus
HSV-1
Viral genomic sequences
Latency
Illumina sequencing
Strain origins
Phylogenetics
Reactivation
Neurovirulence
Viral pathogenesis

ABSTRACT

Herpes simplex virus 1 (HSV-1) is a widespread pathogen that persists for life, replicating in surface tissues and establishing latency in peripheral ganglia. Increasingly, molecular studies of latency use cultured neuron models developed using recombinant viruses such as HSV-1 GFP-US11, a derivative of strain Patton expressing green fluorescent protein (GFP) fused to the viral US11 protein. Visible fluorescence follows viral DNA replication, providing a real time indicator of productive infection and reactivation. Patton was isolated in Houston, Texas, prior to 1973, and distributed to many laboratories. Although used extensively, the genomic structure and phylogenetic relationship to other strains is poorly known. We report that wild type Patton and the GFP-US11 recombinant contain the full complement of HSV-1 genes and differ within the unique regions at only eight nucleotides, changing only two amino acids. Although isolated in North America, Patton is most closely related to Asian viruses, including KOS63.

1. Introduction

An important and widespread human pathogen, herpes simplex virus 1 (HSV-1 or alternatively, Human alphaherpesvirus 1) is considered the archetypal member of the neurotrophic *Alphaherpesvirinae* subfamily (Roizman et al., 2013). More than 70% of the human population become seropositive for HSV-1 during their lifetime, with most establishing a lifelong latent infection punctuated by reactivation events that produce new infectious viruses (Nahmias et al., 1990). Recurring reactivation is responsible for many of the clinical manifestations associated with HSV-1, including painful oral lesions (herpetic lesions), blindness, lasting nerve pain, cognitive impairment and life-threatening viral encephalitis. Despite decades of study there are currently no licensed vaccines against HSV-1 or antiviral drugs that can deplete the latent reservoir (Coen and Whitley, 2011). This deficit provides a strong impetus to better understand the mechanisms of pathogenesis and the control of latency and reactivation (Bloom, 2016; Wilson and Mohr, 2012).

Live-animal infection models have been invaluable in identifying viral determinants of latency and pathogenesis and in exposing the key role of host immunity (Wagner and Bloom, 1997). However, as studies of virus-host interactions become more detailed there is increased need for robust *in vitro* models of latency that simplify the complex biology of natural infections and at the same time allow for more nuanced experimental manipulation of host factors (Bloom, 2016; Wilson and Mohr, 2012). The recent development of several *in vitro* models has benefitted from recombinant viruses that express markers that can be visualized in live cells (Camarena et al., 2010; Cliffe et al., 2015; Kuhn et al., 2012; Pourchet et al., 2017; Thellman et al., 2017). One example is HSV-1 GFP-US11 (strain Patton), a recombinant virus that expresses enhanced green fluorescent protein (EGFP) as an in-frame fusion to the viral US11 protein (Benboudjema et al., 2003). The GFP-US11 fusion protein is expressed with true-late (γ2) kinetics and accumulates at levels sufficient for detection by light microscopy after the onset of viral DNA replication. This provides an easily scored indicator of acute infection and reactivation (Camarena et al., 2010). In all other measures, HSV-1 GFP-US11 replicates with the same efficiency and kinetics as the parental wild type Patton virus (Benboudjema et al., 2003). Despite extensive usage in studies of productive replication, latency, and pathogenesis (Bonneau and Jennings, 1989; Karp et al., 1997; Robey et al., 1976), the complete sequence and gene composition of HSV-1 strain Patton has not been reported.

* Corresponding author.
E-mail address: angus.wilson@med.nyu.edu (A.C. Wilson).
[1] A.P. and R.C. contributed equally to this work.

http://dx.doi.org/10.1016/j.virol.2017.09.016
Received 27 June 2017; Received in revised form 16 September 2017; Accepted 18 September 2017

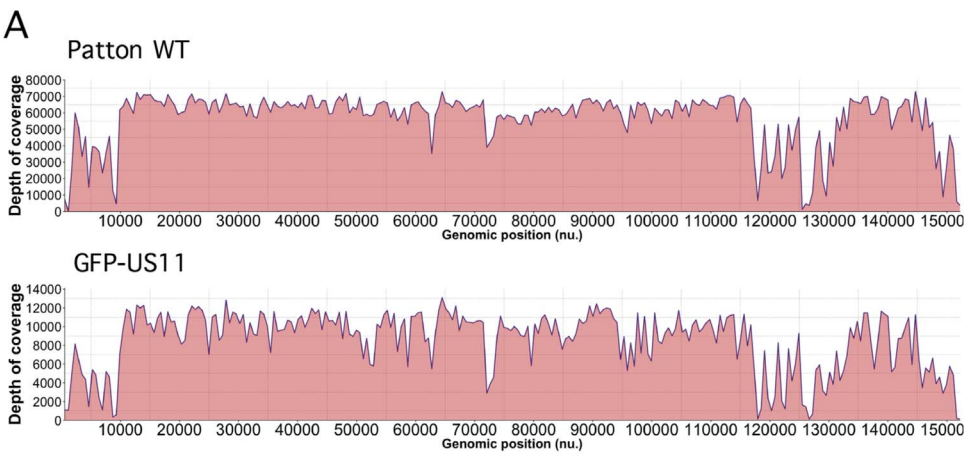

Fig. 1. Sequencing of parental HSV-1 strain Patton and the GPF-US11 recombinant. (A) Summary of sequencing coverage for the wild type and recombinant viruses. Pre-processed sequence reads were aligned to the new draft genomes to assess the coverage depth of each assembly. This is plotted (y axis) against the length of the HSV-1 genome (x axis). (B) List of high confidence single nucleotide polymorphisms (SNPs) between parental Patton virus and its derivative GFP-US11. Additional polymorphisms in regions of low sequence complexity have been excluded. Two SNPs are predicted to change the amino acid sequence of the UL17 capsid protein and UL36 large tegument protein.

B

Coding sequence single nucleotide polymorphisms (SNPs): Patton WT versus GFP-US11

Genomic location	Position of SNP in ORF	ORF name	ORF length	Position of amino acid change	Protein, function
30,956	C341T	UL17	703	Ala-114-Val	capsid protein, DNA packaging
37,053	G3477A	UL18	318	synonymous	VP23, capsid closure
54,254	G1542A	UL28	785	synonymous	terminase subunit, DNA packaging
66,776	C604T	UL31	306	synonymous	nuclear protein, unknown function
71,796	G8673A	UL36	3,164	synonymous	tegument protein, deubiquitinase
79,837	C632T	UL36	---	Ala-211-Val	---
104,925	A156G	UL48	490	synonymous	VP16, tegument/transactivator
137,175	C429A	US4	238	synonymous	envelope glycoprotein G

The importance of HSV-1 strain differences is well recognized, especially in animal pathogenesis models, and may reflect differences in their capacity to cause disease in humans (Dix et al., 1983; Sedarati and Stevens, 1987). For a number of years, the only complete HSV-1 genome sequence was for strain 17 (Brown et al., 1973; McGeoch et al., 1988), but with recent advances in high-throughput sequencing technology, the complete or near-complete sequences of many laboratory strains and clinical isolates have been determined (Bowden et al., 2006; Colgrove et al., 2015; Macdonald et al., 2012a, 2012b; Pfaff et al., 2016; Szpara et al., 2014; Watson et al., 2012). This very substantial increase in sequence information has revealed extensive genotypic variation between isolates, mostly involving single nucleotide polymorphisms (SNPs) along with small insertions and deletions. Some viruses exhibit larger deletions and frame shifts that may be deleterious for growth *in vivo*, however, for the most part individual strains appear relatively stable even after repeated passage in culture (Colgrove et al., 2015). Using sequence variation, individual HSV-1 strains can be grouped into three major clades or phylogroups of African, Asian and European/North American origin (Bowden et al., 2006; Parsons et al., 2015; Szpara et al., 2014, 2010). This diversity is thought to reflect the extended co-evolution of HSV-1 with humans and for the most part the distribution of genotypes mirrors the migration of our species across the globe (Bowden et al., 2006; Sakaoka et al., 1994).

Strain Patton was isolated from a recurrent oral lesion and distributed to other researchers by William E. Rawls M.D., a member of the faculty in the Department of Virology and Epidemiology at Baylor College of Medicine in Houston, Texas (Duff and Rapp, 1973). Unfortunately, the ethnicity, travel history and health status of the donor was not published. Over successive years the genome of the virus has been characterized by restriction endonuclease mapping (Denniston et al., 1981; Enquist et al., 1979; Graham et al., 1978) and small portions have been sequenced (Umene et al., 1984; Watson et al., 1982). However, the exact composition and arrangement of viral genes, especially those involved in important traits such as neurovirulence, remain unknown. Likewise, there is no information on the phylogenetic relationship of Patton to other HSV-1 strains.

The approximately 152-kb double-stranded DNA genome of sequenced HSV-1 strains share a common structure and relatively high (68%) G/C content (McGeoch et al., 1988; Roizman et al., 2013). There are 80 or so single copy open reading frames (ORFs) and non-coding RNA genes distributed across the unique long (U_L, 106.5 kb) and unique short (U_S, 13.5 kb) regions, which are flanked by terminal and internal long inverted (TR_L/IR_L, 9.2 kb) and short inverted (TR_S/IR_S, 6.6 kb) repeats. There are also small microsatellite repeats (less than 100 bp each) and short reiterated sequences (less than 50 bp each) that together account for much of the sequence variation between HSV-1 strains.

Here we have used Illumina sequencing of viral DNA to derive complete genome sequences for both the parental HSV-1 strain Patton and recombinant derivative GFP-US11, the goals being to determine the gene repertoire and sequence of each virus, and to establish the evolutionary relationship to other laboratory strains. We identified the full complement of HSV-1 open reading frames (ORFs) and non-coding RNA genes in each viral genome. Using strain 17 as reference, no large deletions or rearrangements in gene order were identified. Despite extensive engineering and passage within the laboratory, only eight single nucleotide polymorphisms (SNPs) were found in non-repetitive regions of the recombinant virus compared to the parental sequence and of these, only two SNPs are predicted to change the corresponding protein sequence. Unexpectedly, lineage analysis showed that strain Patton is related to strain KOS (referred to here as KOS63, (Bowen et al., 2016)) another widely used laboratory strain first isolated at Baylor College of Medicine, Texas in the early 1960s (Smith, 1964). Numerous differences in the genomic sequences as well as previously reported differences in biological properties such as neurovirulence confirm that KOS63 and Patton are distinct viruses. We conclude that wild type