

A Genome-Wide Comparative Evolutionary Analysis of Herpes Simplex Virus Type 1 and Varicella Zoster Virus

Peter Norberg^{1*}, Shaun Tyler², Alberto Severini², Rich Whitley³, Jan-Åke Liljeqvist¹, Tomas Bergström¹

1 Department of Virology, University of Gothenburg, Gothenburg, Sweden, **2** National Microbiology Laboratory, Public Health Agency of Canada, Canadian Science Centre for Human and Animal Health, Winnipeg, Canada, **3** University of Alabama at Birmingham, Birmingham, Alabama, United States of America

Abstract

Herpes simplex virus type 1 (HSV-1) and varicella zoster virus (VZV) are closely related viruses causing lifelong infections. They are typically associated with mucocutaneous or skin lesions, but may also cause severe neurological or ophthalmic diseases, possibly due to viral- and/or host-genetic factors. Although these viruses are well characterized, genome-wide evolutionary studies have hitherto only been presented for VZV. Here, we present a genome-wide study on HSV-1. We also compared the evolutionary characteristics of HSV-1 with those for VZV. We demonstrate that, in contrast to VZV for which only a few ancient recombination events have been suggested, all HSV-1 genomes contain mosaic patterns of segments with different evolutionary origins. Thus, recombination seems to occur extremely frequent for HSV-1. We conclude by proposing a timescale for HSV-1 evolution, and by discussing putative underlying mechanisms for why these otherwise biologically similar viruses have such striking evolutionary differences.

Citation: Norberg P, Tyler S, Severini A, Whitley R, Liljeqvist J-Å, et al. (2011) A Genome-Wide Comparative Evolutionary Analysis of Herpes Simplex Virus Type 1 and Varicella Zoster Virus. PLoS ONE 6(7): e22527. doi:10.1371/journal.pone.0022527

Editor: Darren P. Martin, Institute of Infectious Disease and Molecular Medicine, South Africa

Received: April 8, 2011; **Accepted:** June 23, 2011; **Published:** July 25, 2011

Copyright: © 2011 Norberg et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Swedish Medical Research Council, the LUA Foundation of Sahlgrenska hospital and the Torsten and Ragnar Söderberg Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: peter.norberg@gu.se

Introduction

Human alpha-herpesviruses comprise three members, herpes simplex virus (HSV) 1 and 2 and varicella zoster virus (VZV). These viruses have a wide host cell range and an efficient and rapid reproductive cell cycle [1]. HSV-1 is typically associated with oral lesions, but primary genital HSV-1 infections have recently become more commonly detected [2]. HSV-2 is the one of the most common sexually transmitted pathogens and is typically associated with genital lesions, although primary oral HSV-2 infections may occur. VZV is the causative agent of chickenpox. After infection, all three viruses have the capacity to establish life-long infections with latency/persistence in sensory neurons, and may induce vesicular lesions during reactivation; HSV-1 and HSV-2 in form of oral or genital ulcers while VZV cause herpes zoster. In addition, these viruses may cause severe complications from the visual and nervous systems. Genetic markers of virulence among circulating strains are largely unknown, but are suggested to be involved in the pathogenesis [3,4,5].

The genomes of HSV and VZV contain a unique long (UL) and a unique short (US) region, each flanked with inverted repeat regions [6,7]. Previous analysis of the three genes US4, US7 and US8, located in the US region and coding for the glycoproteins G (gG), I (gI) and E (gE), respectively, demonstrated a division of HSV-1 strains into three clades, designated as A, B and C [8]. Furthermore, 22% of the strains were classified as recombinants. An RFLP genotyping assay was also developed to classify HSV-1 strains to any of the three clades [9]. This assay has subsequently been utilized in e.g. genotype-phenotype association studies [10], and analysis of geographic stability of

HSV-1 genotypes [11]. Although recombination crossovers have been demonstrated for sub-genomic regions [8,12], the extent of recombination in HSV-1 considering complete genomes has hitherto been unknown.

Studies of whole genome HSV-1 sequences are rare, and only three complete HSV-1 genomes are currently available in GenBank. These are strain 17, which was sequenced over 20 years ago [13,14], and strains F and 129 which were sequenced in 2010 [15]. Similarly, only one HSV-2 strain, HG52 (GenBank accession # NC_001798), is currently available in GenBank. Consequently, no genome-wide phylogenetic analysis based on HSV-1 or HSV-2 has hitherto been presented. In contrast, the complete genome sequence of VZV is well established, and analysis of VZV has demonstrated that globally circulating strains cluster to distinct clades in phylogenetic trees [16,17,18]. Although homologous recombination has contributed to the evolution of VZV, these events were rare and seemed to be relatively old since all members in a specific phylogenetic clade shared the same patterns of recombination crossovers. Therefore, it is possible to genotype a complete VZV genome by analyzing one or a few loci and several such studies have been reported [16,19]. Here, as part of a search for possible genetic traits of neurovirulence in HSV-1, we sequenced nine HSV-1 genomes. In total, ten clinical and two laboratory HSV-1 strains were included in the analysis. We found a striking difference to VZV genomics in that each HSV-1 strain appeared as mosaic recombinants. We propose a time-scale for HSV-1 evolution that demonstrates a slow divergence into different clades starting ~700,000 years before present (BP). The findings allowed a comparison of the genetic evolution of these two related human alpha-herpesviruses.

Results

Genetic diversity

A total of 119,727 nucleotides were analyzed in each HSV-1 genome. A schematic illustration of the genome regions analyzed is given in Fig. 1A. The genetic distance between globally

circulating strains was estimated for HSV-1 and VZV. The genetic distance between the most distantly related HSV-1 strains ranged from approximately 0.4% to 1.7%, and the most distantly related VZV strains ranged from approximately 0.1% to 0.5%. The US region was more variable than the UL region for both viruses.

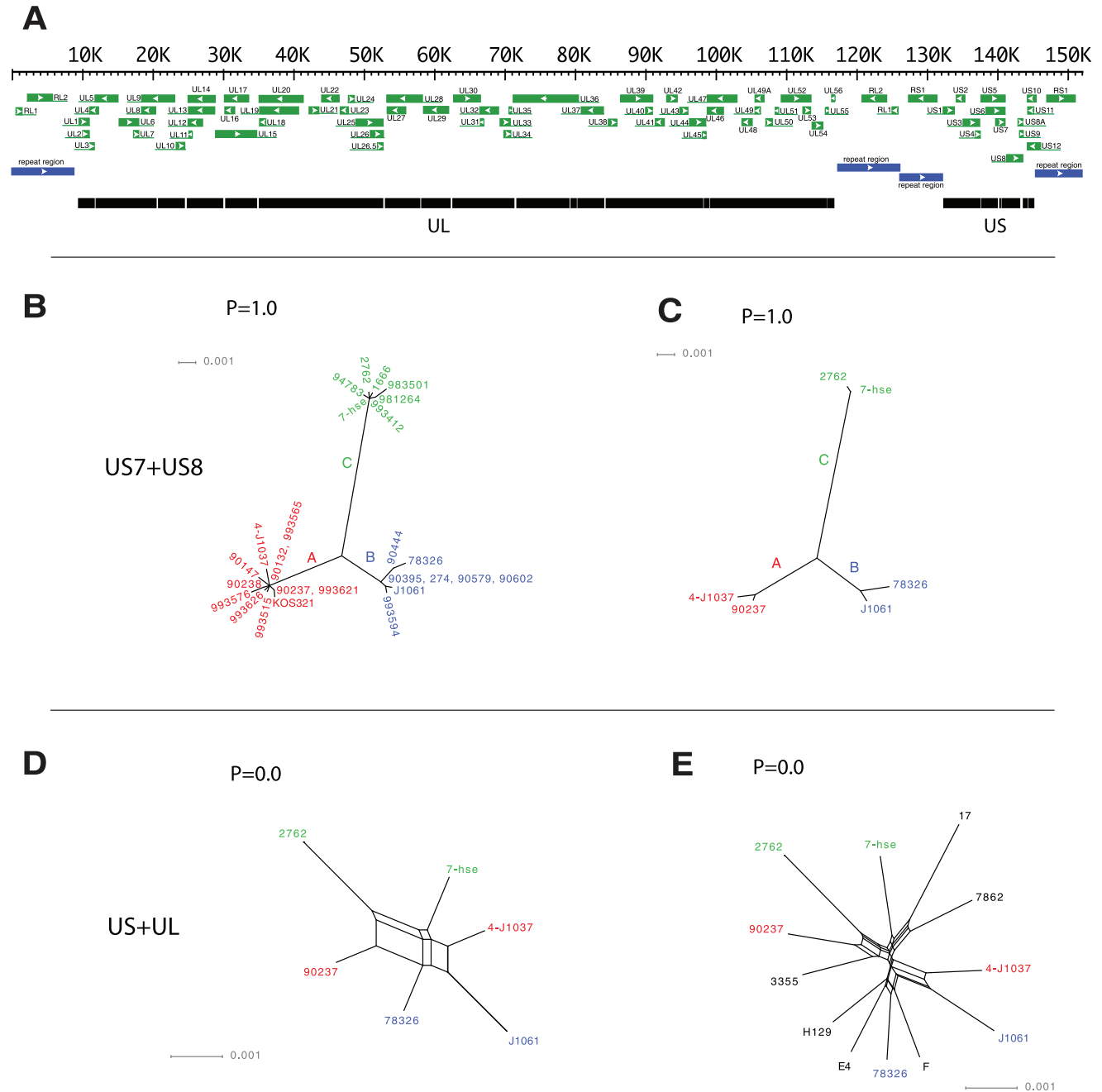


Figure 1. HSV-1 phylogeny. (A) Schematic illustration of the HSV-1 genome. The thick black bar in the bottom depicts the regions that were included in the analysis after all repeat regions and gaps were removed. Nucleotide positions refer to the laboratory strain 17. (B, C, D and E) Phylogenetic networks (splits networks). Networks B and C are based only on US7 and US8, which were the genes analyzed previously by Norberg et al [8]. All strains with recombination crossovers in this region were excluded in order to illustrate the division of the non-recombinant strains into the three distinct evolutionary clades A, B and C. Network B also includes the strains used by Norberg et al, and network C includes only the strains sequenced for this study. Network D is based on the complete genome, including only the strains used in network C (i.e. without recombination crossovers in US7 or US8). Here, recombination crossovers in other parts of the genome prohibit distinct classification of the strains into three distinct clades. Finally, network E is based on the complete genome, including all strains. The network reveals massive recombination and it is irrelevant to attempt to classify a complete genome into any of the three phylogenetic clades A, B or C. The statistical significance for recombination is shown for each dataset. doi:10.1371/journal.pone.0022527.g001

Phylogenetic and recombination analysis

Two splits networks were initially constructed based on the US7 to US8 region. The first network (Fig. 1B) includes all strains without recombination crossovers in the US7 to US8 region previously analyzed by Norberg et al [8]. The second network (Fig. 1C) includes only strains which were sequenced in this study, and where no recombination crossovers were detected in the US7–US8 region. These two networks clearly divide the strains into the three previously described clades A, B and C with no statistical significance for recombination ($p=1.0$). Two genome-wide networks were subsequently constructed. The first (Fig. 1D) only includes strains without recombination crossovers in the US7–US8 region. There was, however, a high statistical significance for recombination in the dataset ($p=0.0$), and the recombination crossovers prohibited a distinct classification of the strains into clades A, B and C. The genome-wide network based on all strains (Fig. 1E) presented frequent recombination crossovers in the genomes that entirely erased the classification into distinct clades, leaving a star-shaped phylogeny and a high statistical significance for recombination ($p=0.0$).

To analyze and visualize recombination crossovers and the degree of fragmentation and shifting phylogenetic relationships between genomes, we used the bootscan and SimPlot methods. Results from these analyses demonstrated highly fragmented genomes. For example, the two strains 2762 and 4-J1037, which were classified as non-recombinants when analyzing only US7 and US8, presented highly shifting phylogenetic relationships to the other strains in the genome-wide analysis (Fig. 2A and 2B). The SimPlot analysis of the same strains demonstrated a high (near 100%) sequence similarity to at least four and two putative parental strains, respectively. In addition, the bootscan analysis of the laboratory strains F and 17 revealed highly shifting phylogenetic relationships indicating massive recombination (Fig. 2C and 2D). Similar results were obtained by the bootscan analysis of the other strains (data not shown).

Timescale of HSV-1 evolution

Due to the high frequency of recombination crossovers in the genomes, the timescale analysis of HSV-1 was restricted to the US7 to US8 genomic region, for which the three clades A, B and C are well characterized. This approach also allowed us to include additional strains that were previously presented [8]. Only strains with no recombination crossovers in this region were included in the analysis. The mean substitution rate, averaged over the whole tree in this region, was estimated to 1.82×10^{-8} nucleotide substitution per site per year. The estimated times since the most recent common ancestors were as follows: clades A, B and C: $\sim 710,000$ years; clades A and B: $\sim 410,000$ years; clade B: $\sim 137,000$ years; clade A: $\sim 113,000$ years; and clade C: $\sim 72,000$ years (Fig. 3). Most of the branch points within each clade occurred between 70,000 and 10,000 years BP (shaded in green in Fig. 3). It should be noted that the standard deviation was rather high (purple bars in Fig. 3), indicating that the true times since divergences may differ slightly from these estimations.

Additional phylogenetic analysis of VZV

To extend previous analysis of VZV, we constructed a phylogenetic network based on all available complete VZV genome sequences (Fig. 4). In addition to a deep lineage reticulate pattern demonstrated previously [16,17,20,21], the network constructed here, and previously by Schmidt-Chanasit and Sauerbrei [21], also demonstrates a reticulate pattern within clade 1, indicating that strain SVETA is involved in an intraclade recombination event. To analyze this further, we also applied the

phi-statistics on strain SVETA and its putative parental strains Kel, Dumas, SD and MSP (the sequences of the passaged strain 32 were excluded from this analysis to minimize influences of mutations introduced during cell-passages). The result demonstrates a high statistical significance for recombination ($p=0.038$) within clade 1.

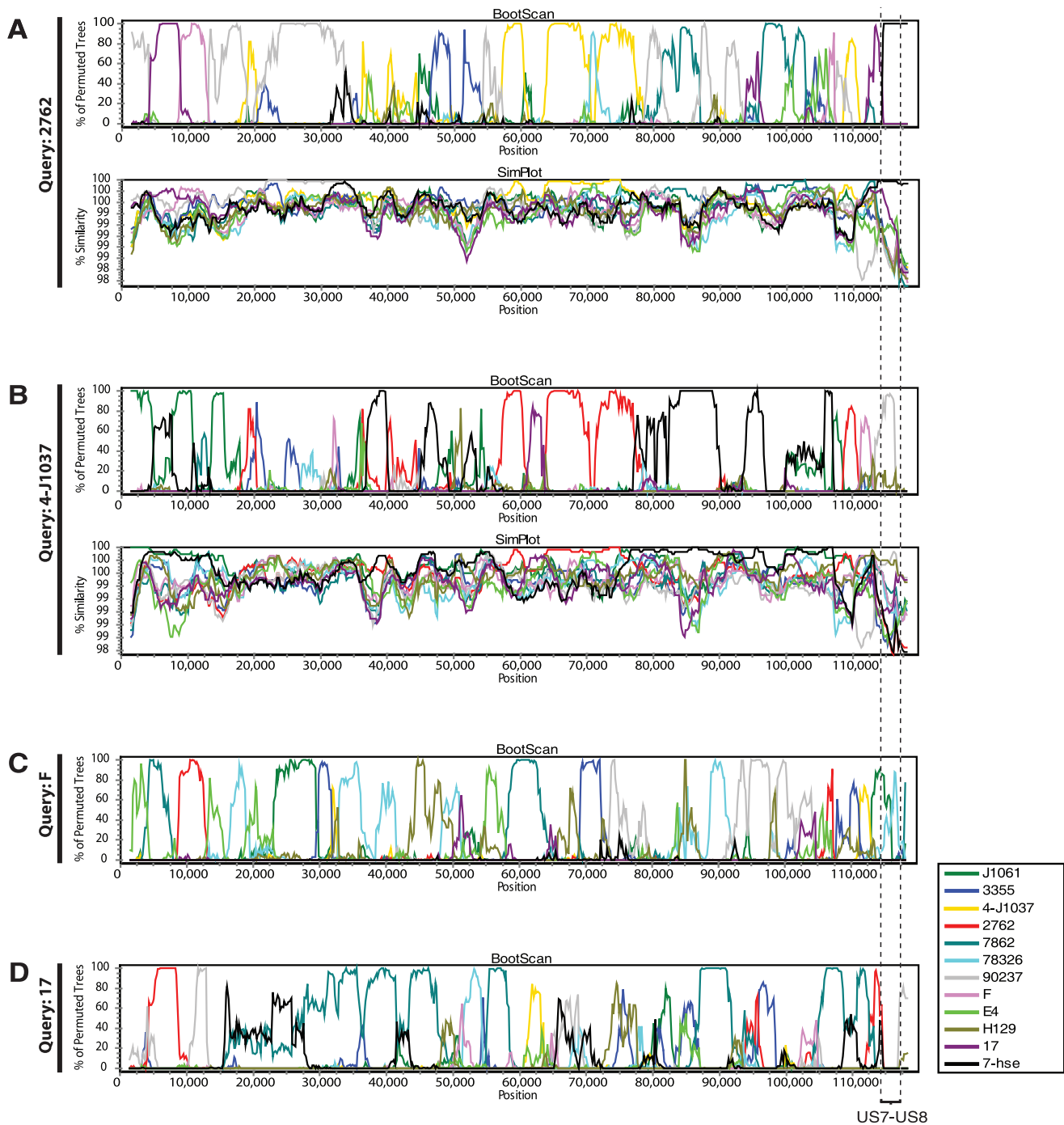
Discussion

The genome-wide analysis performed here reveals an evolutionary history of HSV-1 including an extremely high frequency of recombination events, and that all HSV-1 genomes consist of mosaic patterns of genetic segments of different evolutionary origins. The frequency of recombinants and recombination crossovers in each HSV-1 genome is hence considerably higher than previously demonstrated by the analysis of the US4 to US8 regions, supporting previous suggestions that most, or all, of the globally circulating HSV-1 strains are recombinants [8]. However, it is worth noting that the bootscan method applied here was designed to demonstrate recombination crossovers in a recombinant strain as compared to its non-recombinant parental strains. If, however, the putative parental strains are also recombinants, some of the recombination crossovers demonstrated in the query strain might instead be crossovers in any of the other strains. The frequency of recombination crossovers in each specific strain might therefore be overestimated if all strains are recombinants, which may be the case here with HSV-1. Nevertheless, the method provides an informative way to illustrate changes in phylogenetic relationships between the query strain and the other strains as a result of recombination.

In contrast to VZV, most of the detected recombination crossovers in HSV-1 are unique for each strain. It is thus likely that these recombination events are relatively recent events, except for at least one ancient recombination crossover previously demonstrated in US4 [8]. The majority of the branch points of the analyzed strains was estimated to have occurred 70,000 to 10,000 years BP (highlighted in green in Fig. 3) implying that the vast majority of the detected recombination events in HSV-1 occurred 0 to 70,000 years ago. This feature, in combination with the fact that three distinct phylogenetic clades can still be detected in short genomic stretches, suggest that strains from the three clades have been geographically separated for significant time. This hypothesis fits well with the suggested time since divergence of the different HSV-1 clades of $\sim 700,000$ to 70,000 years BP, and the putative geographic isolation of *Homo sapiens* and other *Homo*-species during that time. Therefore, most of the HSV-1 recombination events are probably results of more recent mixtures of viral populations due to human migrations and explorer expeditions.

The high frequency of HSV-1 recombinants may appear somehow surprising since only a few ancient recombination events have been demonstrated for VZV. Complete VZV genomes can be classified into distinct phylogenetic clades, and five clades (and two additional clades that needs to be confirmed) have hitherto been demonstrated [16,17,19,22]. In contrast, the extremely high frequency of HSV-1 recombinants demonstrated here emphasizes the irrelevance of assigning a complete HSV-1 genome to a specific phylogenetic clade as can be done with VZV. This is an important feature that needs to be considered in future genotyping strategies and viral genotype/phenotype association studies.

Another difference between the two viruses is that HSV-1 is at least four times more diverged than VZV. Such differences in diversity for similar viruses are commonly explained by different times since divergence and thus in this case, that HSV-1 has had significantly longer time to evolve into different clades than VZV.



Bootscan parameters: Window: 3000 bp, Step: 200 bp, GapStrip: On, Reps: 100, Kimura (2-parameter), T/t: 2.0, Neighbor-Joining
 SimPlot parameters: Window : 3000 bp, Step: 200 bp, GapStrip: On, Kimura (2-parameter), T/t: 2.0

Figure 2. HSV-1 recombination analysis. Recombination analysis using BootsCan and SimPlot. A and B illustrate the bootsCan and SimPlot analysis of query sequences 2762 and 4-J1037, respectively, which were classified as non-recombinants based on the US7-US8 regions. BootsCan plots demonstrate highly fragmented genomes as a result of recombination. Similarity plots demonstrate the sequence similarity between the query sequence and the other sequences. C and D depicts the bootsCan analysis of the reference strains F and 17. Nucleotide positions refer to the sequence alignment excluding gaps and repeat regions.
 doi:10.1371/journal.pone.0022527.g002

This assumption is based on another assumption, namely that these viruses evolve with a similar molecular clock. Since HSV-1 and VZV both are human alpha-herpesviruses with similar

biology including persistency in sensory neurons, such a difference in time since divergence is paradoxical. The hypothetical shorter time scale of VZV divergence might suggest either an isolation of

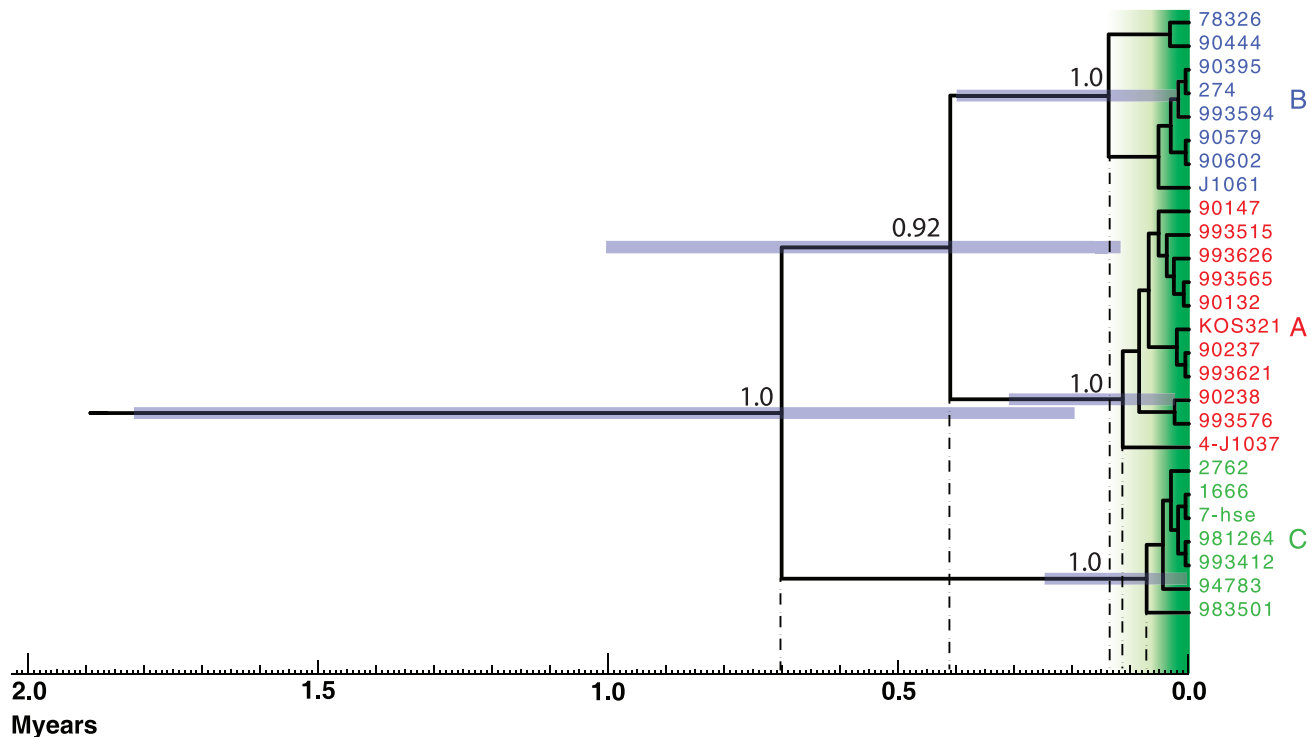


Figure 3. HSV-1 timescale estimation based on the genes US7 and US8. Estimated timescale for the HSV-1 evolution. Each major branch point in the phylogenetic tree is marked with a dotted line to the scale-bar below. The 95% confidence intervals for each predicted time since divergence are denoted with purple bars. The increased number of branch points in each clade is highlighted with green background. Posterior probabilities for the major clades are shown.
doi:10.1371/journal.pone.0022527.g003

the virus in a subpopulation of human hosts, or a zoonotic transfer of VZV from another species. Since VZV shows a high global seroprevalence and a low mutation rate, neither of these explanations seems plausible. There might however, be another more likely explanation for the differences in divergence of HSV-1 and VZV. We suggest that these viruses may instead have started their divergence into different clades at approximately the same time-point. The formation of the different clades hence probably reflects similar co-evolution with the human host. This explanation does not necessarily exclude similar mutation rates in terms of mutations per replication cycle, but instead assumes different numbers of replication cycles (and hence also mutations) per year.

After primary infection, both viruses can reactivate from a state of latency or persistency in sensory ganglia. In a comparative study on viral presence in trigeminal ganglia from cadavers, VZV was detected in 97% and HSV-1 in 65% of the subjects [23]. While HSV-1 may reactivate frequently (up to monthly in some cases) to cause oral herpes, the reactivation of VZV in form of herpes zoster seldom occurs more than once in a lifetime. Furthermore, while HSV-1 is shed asymptotically in saliva several times per week in almost all studied subjects, such shedding of VZV is rarely detected [24,25]. Taken together, it is probable that HSV-1 undergoes a higher number of replication cycles per year, as compared with VZV. In addition, an early (mean 13 months) acquisition of HSV-1 as suggested by seroconversion data [26] is compatible with a spread from parents while outbreaks of chickenpox may originate from herpes zoster in the grandparental generation [27], further slowing down the evolution of VZV in comparison to HSV-1.

Despite the lower frequency of recombinants reported for VZV than for HSV-1, it is likely that the inherent ability to recombine is

similar for those. As an example, during simultaneous cell culture infection with two distinct strains of another varicellovirus, the bovine herpesvirus 1, 40% of the progeny were found to be recombinants already after one round of replication [28]. These bovine herpesvirus 1 recombinants presented a highly fragmented pattern similar to the fragmentation presented here for HSV-1. Likewise, experimental infection with several VZV strains resulted in recombinants of this virus [29]. Furthermore, in the present study we demonstrate that intra-clade recombination occurs also for VZV, exemplified by strain SVETA. Thus, there may be an alternative explanation for the (in comparison to HSV-1) lower frequency of recombinant VZV strains. As suggested above, a putative lower frequency of reactivations and replication cycles per year for VZV most likely diminish the possibilities for recombination. That is, even if two VZV strains infect the same host, the probability for simultaneous replication in the same cell decreases if the number of replications and reactivations is low. Another possibility is that a primary infection of several VZV strains simultaneously is relatively uncommon as compared to HSV-1, due to the low degree of asymptomatic shedding of the former virus.

Furthermore, while all three HSV-1 clades were found to co-exist in the same geographic locations [8,11], the VZV clades have been demonstrated to follow a specific geographic distribution [30,31,32,33]. Such geographic separation of strains from different clades would be an efficient barrier to prevent intra-clade recombination.

Interestingly, recent studies indicate that this geographic separation of VZV clades is also about to fade out. An investigation of VZV in Germany demonstrates a higher genetic diversity between clinical isolates from patients with varicella than

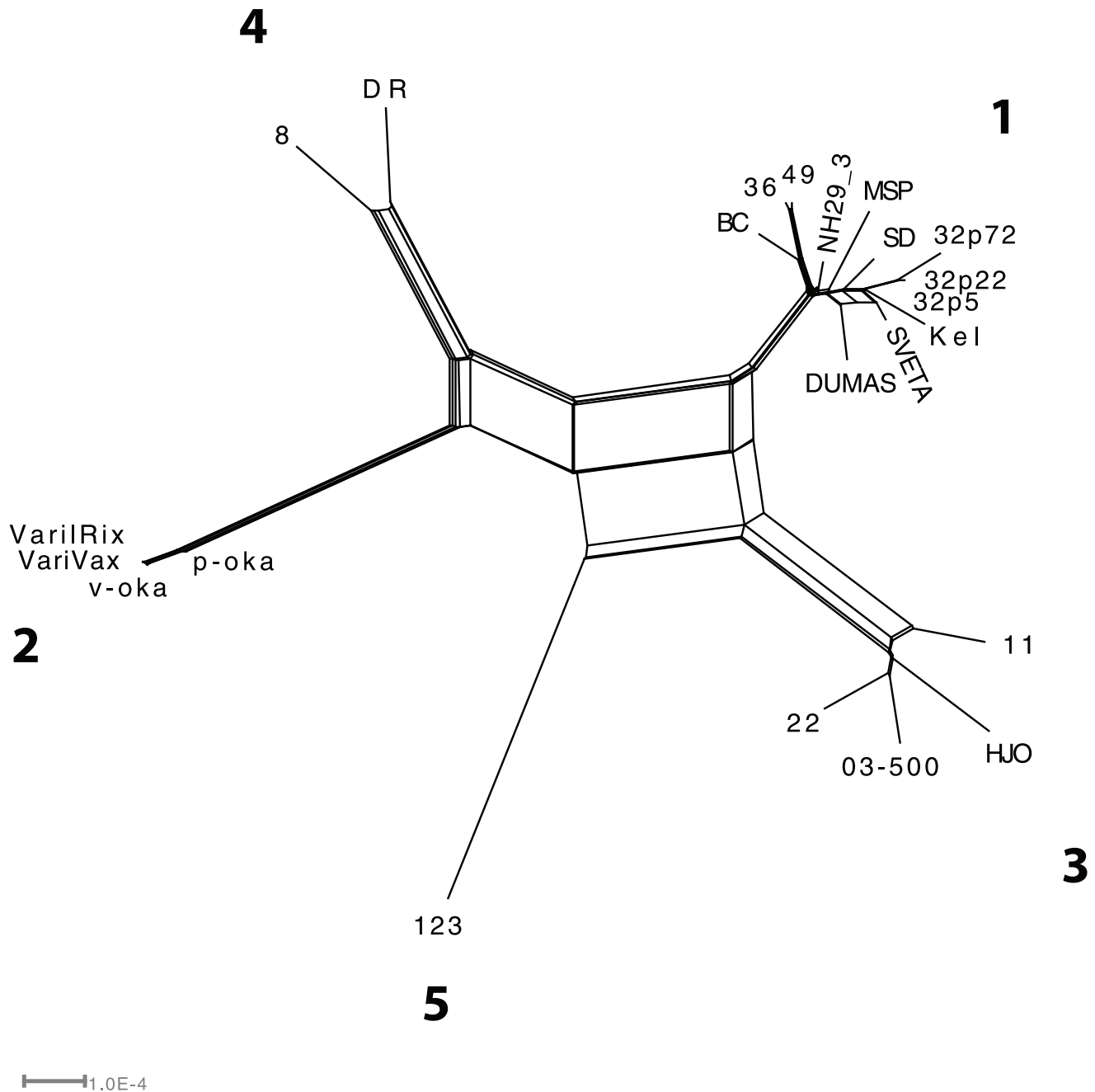


Figure 4. VZV phylogeny. Phylogenetic network (splits network) of available complete genome VZV sequences. The strains are divided into clades 1 to 5, and the deep lineage reticulate topology is caused by a few ancient recombination events described previously. The network also presents a reticulate pattern within clade 1 caused by a putative recent recombination event involving strain SVETA. doi:10.1371/journal.pone.0022527.g004

from those with zoster, an observation which the authors suggest is due to recent import of genotypes from other countries [34,35]. Furthermore, coinfection of two VZV clades, both circulating in the same area, was recently demonstrated in a child with varicella [36]. Another evidence of geographic integration of VZV clades is that in immigrant patients with herpes zoster in East London, 30% of the strains belonged to European clades despite a history of varicella in the countries of origin of these individuals (14). This finding clearly indicates that reinfection of VZV might be common, which in turn may allow for recombination events.

Another novel possibility of VZV recombination events is the world-wide distribution of an attenuated vaccine strain [37]. This live VZV vaccine, which clusters to clade 2, is currently routinely used in for example Japan, Korea, the USA, Canada, Australia, Germany, Costa Rica, Uruguay, and Qatar, where strains from other VZV clades are predominant. This increased geographic mixing of VZV strains from different clades will undoubtedly weaken the geographic barrier for intra-clade recombination of VZV.

In conclusion we demonstrate that the evolution of HSV-1 and VZV, on a genome-wide scale, show two major differences. First,

the genetic diversity of globally circulating strains is significantly higher for HSV-1 than for VZV and second; the frequency of recombinant strains is much higher for HSV-1 than for VZV. We propose that these two differences are related to biological properties regarding replication and spread of the respective infection rather than to distinct molecular clocks and different time-scales of evolution in humans. Owing to the increased mixing of previously geographically separated VZV strains, the frequency of inter-clade VZV recombinants may drastically increase in the future. Thus, the mosaic genetics of HSV-1 can be a plausible indicator on future evolutionary characteristics of VZV, provided that migration and/or vaccination alter previous geographic restriction of the spread of this virus.

Materials and Methods

HSV-1 strains

Clinical isolates of HSV-1 ($n = 9$) from USA (4 strains of which 2 were isolated from oral lesions and 2 from brains of patients with encephalitis) and from Sweden (5 strains of which one was isolated from lip and four from brains) were used to infect Vero cells at a multiplicity of infection of approximately 0.01. Cells were harvested when showing 100% cytopathic effect and lysed in 1% SDS, Tris-EDTA buffer at pH 8.0. Total DNA was purified by Proteinase K digestion and phenol/chloroform extraction and HSV-1 DNA was separated from the host DNA by NaI gradient centrifugation, as previously described [38]. Gradient-purified HSV-1 DNA was subjected to sequencing, performed via pyrosequencing on a Roche GS FLX instrument using the Titanium sequencing kits. All strains received a minimum depth of coverage of 40 \times . The data was assessed using the GS Assembler software as a de novo assembly as well as by mapping to the reference sequence of HSV-1 strain 17 (NC_001806) using the GS Mapper software with similar results. Each strain produced a number of high quality contigs (Q40 Plus Bases greater than 99.6%) even though the coverage greatly exceeded the 25 \times coverage necessary for high quality de novo assemblies. In general it was observed that the unique long and unique short regions produced single contigs whereas the repeat regions were highly fragmented and represented by numerous, relatively short contigs with multiple gaps, some spanning more than 1 kb. This did not appear to be an assembly issue per se as the mapping results produced similar results. It would appear that the sequencing libraries were underrepresented with reads in this region which may be due to a PCR bias as a result of the high G+C content of the repeat regions. The sequencing data for this study has been deposited to the short read archive in Genbank under accession number SRA035240.1.

References

- Roizman B, Sears AE (1996) Herpes simplex viruses and their replication. In: Fields BN, Knipe DM, Howley PM, et al. (1996) Virology. Philadelphia: Lippincott-Raven, pp 2231–2295.
- Lowhagen GB, Tunback P, Bergstrom T (2002) Proportion of herpes simplex virus (HSV) type 1 and type 2 among genital and extragenital HSV isolates. *Acta Derm Venereol* 82: 118–120.
- Bergstrom T, Alestig K, Svennerholm B, Horal P, Skoldenberg B, et al. (1990) Neuroirulence of herpes simplex virus types 1 and 2 isolates in diseases of the central nervous system. *Eur J Clin Microbiol Infect Dis* 9: 751–757.
- Dix RD, McKendall RR, Baringer JR (1983) Comparative neuroirulence of herpes simplex virus type 1 strains after peripheral or intracerebral inoculation of BALB/c mice. *Infect Immun* 40: 103–112.
- Whitley R, Lakeman AD, Nahmias A, Roizman B (1982) Dna restriction-enzyme analysis of herpes simplex virus isolates obtained from patients with encephalitis. *N Engl J Med* 307: 1060–1062.
- Sheldrick P, Berthelot N (1975) Inverted repetitions in the chromosome of herpes simplex virus. *Cold Spring Harb Symp Quant Biol* 39 Pt 2: 667–678.
- Wadsworth S, Jacob RJ, Roizman B (1975) Anatomy of herpes simplex virus DNA. II. Size, composition, and arrangement of inverted terminal repetitions. *J Virol* 15: 1487–1497.
- Norberg P, Bergstrom T, Rekabdar E, Lindh M, Liljeqvist JA (2004) Phylogenetic analysis of clinical herpes simplex virus type 1 isolates identified three genetic groups and recombinant viruses. *J Virol* 78: 10755–10764.
- Norberg P, Bergstrom T, Liljeqvist JA (2006) Genotyping of clinical herpes simplex virus type 1 isolates by use of restriction enzymes. *J Clin Microbiol* 44: 4511–4514.
- Duan R, van Dun JM, Remeijer L, Siemerink M, Mulder PG, et al. (2008) Prevalence of herpes simplex virus type 1 glycoprotein G (gG) and gI genotypes in patients with herpetic keratitis. *Br J Ophthalmol* 92: 1195–1200.
- Schmidt-Chanasit J, Bialonski A, Heinemann P, Ulrich RG, Gunther S, et al. (2009) A 10-year molecular survey of herpes simplex virus type 1 in Germany demonstrates a stable and high prevalence of genotypes A and B. *J Clin Virol* 44: 235–237.
- Bowden R, Sakaoka H, Donnelly P, Ward R (2004) High recombination rate in herpes simplex virus type 1 natural populations suggests significant co-infection. *Infect Genet Evol* 4: 115–123.
- McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC, et al. (1988) The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J Gen Virol* 69 (Pt 7): 1531–1574.
- McGeoch DJ, Dolan A, Donald S, Brauer DH (1986) Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. *Nucleic Acids Res* 14: 1727–1745.

DNA sequence analysis

The viral DNA sequences were aligned by using ClustalW included in eBioTools and BioX 1.5, using default settings, followed by manual inspection of the complete alignment. All gaps and tandem repeat regions were excluded prior to the analysis. Genetic distances were calculated using SimPlot as a mean value in a 3000 bp window, which were slid through the genomes with a step size of 200 bp. To calculate a measure of statistical significance for recombination we used the phi-test, which has been proven to give reliable results for conserved DNA sequences [39]. The splits networks were created by using SplitsTree 4.10 [40], using the uncorrected P characters transformation (other models of characters transformation yielded only negligible differences in results). The bootscan analysis and similarity plots were performed using the SimPlot program [41] with a window size of 3000 nt and a 200 nt step size.

Calculating time scale for HSV-1

To estimate a substitution rate for the US7 to US8 genomic region in HSV-1, and dates for the divergences into the three clades, we used the BEAST software [42]. ModelTest [43] was used to estimate the appropriate evolutionary model (Substitution model: GTR; Site heterogeneity model: Gamma; Number of gamma parameters: 4). The clock model was set to relaxed clock with uncorrected log-normal so as to account for lineage specific heterogeneity. To account for the exponential population growth, the tree prior was set to ‘Coalescent: exponential growth’. HSV-2 was included as outgroup, and the mean time since divergence of HSV-1 and HSV-2 was set to 8.45 Myears with a standard deviation of 25,000 years, which gives a central 95% range of about 4.5 to 4.4 Myears since divergence, which corresponds well to earlier suggestions of HSV-1/HSV-2 divergence [44,45]. The MCMC was run with a chain length of 10,000,000 with sampling of parameters every 200 step. The final tree was constructed as a maximum clade credibility tree with 10% burning (the first 10% of the trees were removed prior to the final tree was built), with no posterior probability limit (all nodes annotated), and the node heights were calculated as mean heights.

Author Contributions

Conceived and designed the experiments: PN JL TB. Performed the experiments: PN ST AS RW. Analyzed the data: PN. Contributed reagents/materials/analysis tools: ST AS RW TB. Wrote the paper: PN JL TB.

15. Szpara M, Parsons L, Enquist LW (2010) Sequence variability in clinical and lab isolates of Herpes Simplex Virus 1 reveals new mutations. *J Virol*.
16. Norberg P, Liljeqvist JA, Bergstrom T, Sammons S, Schmid DS, et al. (2006) Complete-genome phylogenetic approach to varicella-zoster virus evolution: genetic divergence and evidence for recombination. *J Virol* 80: 9569–9576.
17. Peters GA, Tyler SD, Grose C, Severini A, Gray MJ, et al. (2006) A full-genome phylogenetic analysis of varicella-zoster virus reveals a novel origin of replication-based genotyping scheme and evidence of recombination between major circulating clades. *J Virol* 80: 9850–9860.
18. McGeoch DJ (2009) Lineages of varicella-zoster virus. *J Gen Virol* 90: 963–969.
19. Loparev VN, Rubtcova EN, Bostik V, Govil D, Birch CJ, et al. (2007) Identification of five major and two minor genotypes of varicella-zoster virus strains: a practical two-amplicon approach used to genotype clinical isolates in Australia and New Zealand. *J Virol* 81: 12758–12765.
20. Norberg P (2010) Divergence and genotyping of human alpha-herpesviruses: an overview. *Infect Genet Evol* 10: 14–25.
21. Schmidt-Chanasit J, Sauerbrei A (2011) Evolution and world-wide distribution of varicella-zoster virus clades. *Infect Genet Evol* 11: 1–10.
22. Breuer J, Grose C, Norberg P, Tipples G, Schmid DS (2010) A proposal for a common nomenclature for viral clades that form the species varicella-zoster virus: summary of VZV Nomenclature Meeting 2008, Barts and the London School of Medicine and Dentistry, 24–25 July 2008. *J Gen Virol* 91: 821–828.
23. Inoue H, Motani-Saitoh H, Sakurada K, Ikegaya H, Yajima D, et al. (2010) Detection of varicella-zoster virus DNA in 414 human trigeminal ganglia from cadavers by the polymerase chain reaction: a comparison of the detection rate of varicella-zoster virus and herpes simplex virus type 1. *J Med Virol* 82: 345–349.
24. Kaufman HE, Azcuy AM, Varnell ED, Sloop GD, Thompson HW, et al. (2005) HSV-1 DNA in tears and saliva of normal adults. *Invest Ophthalmol Vis Sci* 46: 241–247.
25. Wang CC, Yepes LC, Danaher RJ, Berger JR, Mootoor Y, et al. (2010) Low prevalence of varicella zoster virus and herpes simplex virus type 2 in saliva from human immunodeficiency virus-infected persons in the era of highly active antiretroviral therapy. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 109: 232–237.
26. Tunback P, Bergstrom T, Andersson AS, Nordin P, Krantz I, et al. (2003) Prevalence of herpes simplex virus antibodies in childhood and adolescence: a cross-sectional study. *Scand J Infect Dis* 35: 498–502.
27. Gershon AA, Gershon MD, Breuer J, Levin MJ, Oaklander AL, et al. (2010) Advances in the understanding of the pathogenesis and epidemiology of herpes zoster. *J Clin Virol* 48 Suppl 1: S2–7.
28. Muylkens B, Farnir F, Meurens F, Schynts F, Vanderplasschen A, et al. (2009) Coinfection with two closely related alphaherpesviruses results in a highly diversified recombination mosaic displaying negative genetic interference. *J Virol* 83: 3127–3137.
29. Dohner DE, Adams SG, Gelb LD (1988) Recombination in tissue culture between varicella-zoster virus strains. *J Med Virol* 24: 329–341.
30. LaRussa PS, Gershon AA (2001) Biologic and geographic differences between vaccine and clinical varicella-zoster virus isolates. *Arch Virol Suppl*. pp 41–48.
31. Loparev VN, Gonzalez A, Deleon-Carnes M, Tipples G, Fickenscher H, et al. (2004) Global identification of three major genotypes of varicella-zoster virus: longitudinal clustering and strategies for genotyping. *J Virol* 78: 8349–8358.
32. Muir WB, Nichols R, Breuer J (2002) Phylogenetic analysis of varicella-zoster virus: evidence of intercontinental spread of genotypes and recombination. *J Virol* 76: 1971–1979.
33. Quinlivan M, Hawrami K, Barrett-Muir W, Aaby P, Arvin A, et al. (2002) The molecular epidemiology of varicella-zoster virus: evidence for geographic segregation. *J Infect Dis* 186: 888–894.
34. Sauerbrei A, Wutzler P (2007) Different genotype pattern of varicella-zoster virus obtained from patients with varicella and zoster in Germany. *J Med Virol* 79: 1025–1031.
35. Sauerbrei A, Zell R, Philipps A, Wutzler P (2008) Genotypes of varicella-zoster virus wild-type strains in Germany. *J Med Virol* 80: 1123–1130.
36. Quinlivan M, Sengupta N, Breuer J (2009) A case of varicella caused by co-infection with two different genotypes of varicella-zoster virus. *J Clin Virol* 44: 66–69.
37. Takahashi M, Otsuka T, Okuno Y, Asano Y, Yazaki T (1974) Live vaccine used to prevent the spread of varicella in children in hospital. *Lancet* 2: 1288–1290.
38. Walboomers JM, Schegget JT (1976) A new method for the isolation of herpes simplex virus type 2 DNA. *Virology* 74: 256–258.
39. Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172: 2665–2681.
40. Huson DH (1998) SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* 14: 68–73.
41. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, et al. (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 73: 152–160.
42. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.
43. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
44. McGeoch DJ, Cook S (1994) Molecular phylogeny of the alphaherpesvirinae subfamily and a proposed evolutionary timescale. *J Mol Biol* 238: 9–22.
45. McGeoch DJ, Cook S, Dolan A, Jamieson FE, Telford EA (1995) Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J Mol Biol* 247: 443–458.