One Year Genome Evolution of Lausannevirus in Allopatric versus Sympatric Conditions

Linda Mueller[1], Claire Bertelli[1,2], Trestan Pillonel[1], Nicolas Salamin[2,3], and Gilbert Greub[1,*]

[1]Center for Research on Intracellular Bacteria (CRIB), Institute of Microbiology, Lausanne University Hospital, and University of Lausanne, Switzerland

[2]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

[3]Department of Ecology and Evolution, Biophore, University of Lausanne, Switzerland

*Corresponding author: E-mail: gilbert.greub@chuv.ch.

Accepted: May 17, 2017

Data deposition: This project has been deposited at European Nucleotide Archive under the accession PRJEB20258 and PRJEB7018.

Abstract

Amoeba-resisting microorganisms raised a great interest during the last decade. Among them, some large DNA viruses present huge genomes up to 2.5 Mb long, exceeding the size of small bacterial genomes. The rate of genome evolution in terms of mutation, deletion, and gene acquisition in these genomes is yet unknown. Given the suspected high plasticity of viral genomes, the microevolution of the 346 kb genome of Lausannevirus, a member of *Megavirales*, was studied. Hence, Lausannevirus was co-cultured within the amoeba *Acanthamoeba castellanii* over one year. Despite a low number of mutations, the virus showed a genome reduction of 3.7% after 12 months. Lausannevirus genome evolution in sympatric conditions was investigated by its co-culture with *Estrella lausannensis*, an obligate intracellular bacterium, in the amoeba *A. castellanii* during one year. Cultures were split every 3 months. Genome sequencing revealed that in these conditions both, Lausannevirus and *E. lausannensis*, show stable genome, presenting no major rearrangement. In fact, after one year they acquired from 2 to 7 and from 4 to 10 mutations per culture for Lausannevirus and *E. lausannensis*, respectively. Interestingly, different mutations in the endonuclease encoding genes of Lausannevirus were observed in different subcultures, highlighting the importance of this gene product in the replication of Lausannevirus. Conversely, mutations in *E. lausannensis* were mainly located in a gene encoding for a phosphoenolpyruvate–protein phosphotransferase (PtsI), implicated in sugar metabolism. Moreover, in our conditions and with our analyses we detected no horizontal gene transfer during one year of co-culture.

Key words: Megavirales, *Marseilleviridae*, Lausannevirus, single nucleotide polymorphisms (SNPs), experimental evolution, *Chlamydiales*.

Introduction

Amoebae are protists resistant to extreme temperature and pH variation. They have a worldwide distribution and are found in both natural and artificial niches, such as soil, rivers, and stagnant water, as well as hospitals and clinics (Greub and Raoult 2004; Thomas et al. 2006, 2008; Boratto et al. 2014). These protozoa are able to phagocyte microorganisms $>0.2\ \mu m$. However, some of these, such as *Chlamydia*-related bacteria, giant viruses (Raoult et al. 2004; Lienard et al. 2011; Thomas et al. 2011) and other amoeba-resisting microorganisms (ARMs), may survive and replicate inside amoebae (Kebbi-Beghdadi and Greub 2014). It is thus possible to find a broad range of microbes in these protists, from human pathogens such as *Legionella* spp., *Mycobacterium avium* or *Pseudomonas aeruginosa* to harmless species such as *Bradyrhizobium japonicum* or *Rhodobacter massiliensis* (Greub and Raoult 2004). Since ARMs may be naturally present together in a single amoeba, as example Lausannevirus and *Legionella* spp. were observed simultaneously in several *Acanthamoeba castellanii* amoebae (Bertelli and Greub 2012), we wondered wether gene exchanges may occur between two different ARMs.

We focused in this study on two amoebal intracellular microorganisms discovered in Lausanne: Lausannevirus and *Estrella lausannensis*. Lausannevirus is a Nucleocytoplasmic Large DNA virus (NCLDV) member of the *Marseilleviridae*

family isolated in 2005 from a water sample of the Seine river (Paris, France) and has a genome of 346 kb (Thomas et al. 2011). *E. lausannensis* (*Chlamydiales*, *Criblamydiaceae*) is an obligate intracellular bacterium (Lienard et al. 2011), which was discovered in a water sample from the Llobregat river (Barcelona, Spain; Corsaro et al. 2009) and has a genome of 2.8 Mb (Bertelli et al. 2015). Both microorganisms were isolated using amoebal co-culture (Jacquier et al. 2013) and showed a lytic cycle between 24 and 48 h in *A. castellanii* (Lienard et al. 2011; Thomas et al. 2011).

The evolution of amoebal intracellular microorganisms has been poorly described until today. Previous investigation showed that ARM's evolution is strongly conditioned by their host that could act as reservoir, vector, replicative niche, and protective armor for the internalized microbes (Winiecka-Krusnell 1999; Greub and Raoult 2004; Kebbi-Beghdadi and Greub 2014; Scheid et al. 2014). In addition, these protozoa might represent an evolutionary crib that allows genetic material to be transferred either between different ARMs or between the ARMs and the amoeba (Greub 2009; Gimenez et al. 2011; Bertelli and Greub 2012; Kebbi-Beghdadi and Greub 2014). For example, it has been suggested that the NCLDV core genes, common to all giant viruses, have been acquired through gene transfer from their eukaryotic hosts (Williams et al. 2011; Forterre et al. 2014). Among intracellular bacteria, members of the order *Chlamydiales* exhibit a very interesting example of adaptation to survive in amoebae that may have selected virulence traits necessary to the survival in these protists and in other professional phagocytes such as macrophages (Kebbi-Beghdadi and Greub 2014). Thus, amoebae are now considered to be a melting pot for gene exchanges between different amoebae-resisting microbes (Moliner et al. 2010).

The single study on giant viruses evolutionary genomics documented the genome reduction from 1.2 Mb to 0.993 Mb of the Mimivirus genome when passaged in culture 150 times (Boyer et al. 2011). All deletions were localized at both terminal ends of the genome and were associated with fiber particles loss and a reduced infectivity. The authors suggested that the core genes of NCLDVs, located in central regions of the genome, are under selective pressures, while the ending regions, in which species specific genes are located, are highly recombinant and more subjected to mutations. However, whether mutations are randomly distributed along the genome of other NCLDVs remains to be determined. In addition, whether the presence of a competitor species could modulate NCLDVs infectivity, contribute to an increase in the virus genome size by gene transfer, or would lead to a balance between gene acquisition and deletion, was not investigated and remains to be clarified.

A milestone in bacteria evolutionary genomics was the report of the genome evolution of *Escherichia coli* over 40,000 generations, when maintained in minimal medium supplemented with limiting glucose during 20 years (Barrick et al. 2009). A constant rate of genomic evolution was observed for 20,000 generations, after which new genotypes presenting an elevated mutation rate appeared. This study is a key example for evolutionary dynamics investigations and illustrated how, even in sympatric and highly controlled stable conditions, important DNA modification occurs, highlighting the complexity of genomic evolution (Bremer 1982; Barrick et al. 2009; Raeside et al. 2014).

Thus, to shed lights into ARMs evolution, we performed a short term (1 year) investigation of the evolution of Lausannevirus, in presence and in absence of *E. lausannensis*. Two independent experiences were performed, in the first one, Lausannevirus was co-cultured over one year in *A. castellanii* amoebae, its natural host. In the second experiment, both ARMs, Lausannevirus and *E. lausannensis*, were co-cultured within *A. castellanii* amoebae, during one year. Full genome sequencing of 3-monthly time-points allowed to determine genetic and functional changes that occurred in the microbial populations, hence providing a one-year snapshot of evolutionary processes driving these microbial genomes in allopatric versus sympatric co-cultures.

Materials and Methods

Amoebal Culture

Acanthamoeba castellanii ATCC 30010, were cultured in peptone yeast extract glucose (PYG) medium, as previously reported (Greub and Raoult 2004; Greub et al. 2004; Jacquier et al. 2013) in 75 cm^2 surface cell culture flasks (Becton Dickinson, Allschwil, Switzerland) at 25 °C.

Bacterial/Viral Cultures and Purification

Lausannevirus was co-cultured with *A. castellanii* ATCC 30010 at 32 °C, in 75 cm^2 surface cell culture flasks (Becton Dickinson), with 30 ml of PYG medium. When complete amoebal lysis was observed, co-cultures were harvested. The same procedure was used for *E. lausannensis* strain CRIB30. Both bacteria and virus were purified with Wizard Genomic DNA purification kit (Promega Corporation, Madison, USA; Thomas et al. 2011).

One Year Amoebal Co-Culture

During one year, representing 144 passages, Lausannevirus and *E. lausannensis* strain CRIB30 were co-cultured with *A. castellanii* ATCC 30010 at 32 °C, in 75 cm^2 surface cell culture flasks. Three times per week infected flask were filtered and fresh confluent amoebal cells were infected (10-fold dilution, as 3 ml/flask). Each three months the culture was split into two subcultures to obtain eight different subcultures after 144 passages (fig. 1). Each month, viruses and bacteria were harvested, as previously described (Thomas et al. 2011). Briefly, after a centrifugation of 15 min at 5,000 × g flasks supernatant was collected and filtered at 5 μm to discard residual amoebal cells. Further, the filtrate was

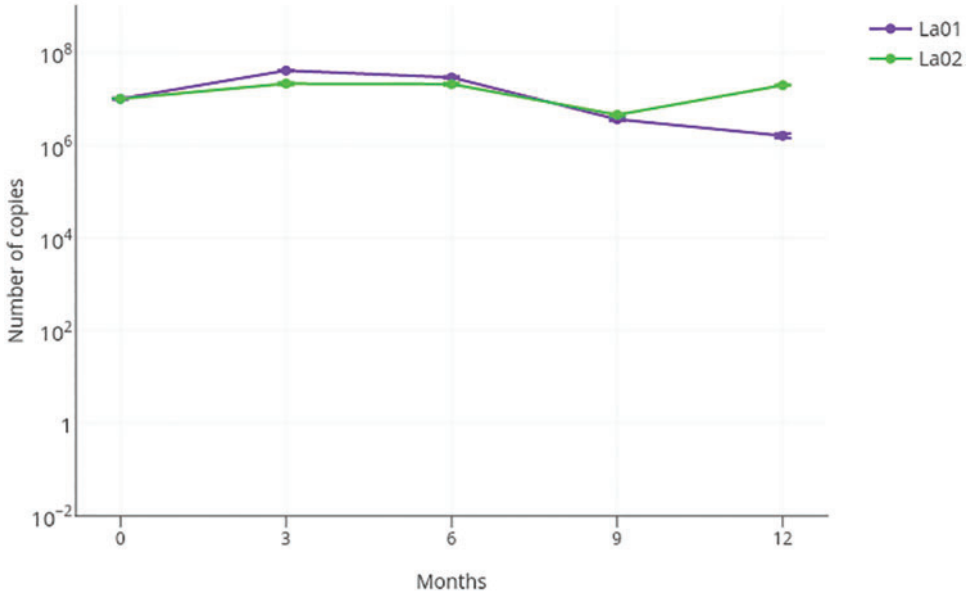

Fig. 1.—Lausannevirus evolution in allopatric conditions: DNA copies number among duplicates. The number of viral DNA copies was assessed by qPCR for both duplicates, La01 and La02. Both subcultures show similar profiles, with a constantly high number of viral DNA copies/l along the year of co-culture.

centrifuged during 1 h at 35,000 × g and the pellet was resuspended in 1 ml of DNA-free water. Samples were stored at −80 °C with 15% glucose.

Genome Sequencing

To obtain a high quality complete reference sequence of *E. lausannensis*, *E. lausannensis* CRIB-30 DNA was extracted using Gentra Puregene (Qiagen, Hilden, Germany), following gram positive bacteria protocol. To improve bacterial lysis, an extra incubation step with proteinase K during 2 h at 55 °C was performed. Bacterial DNA was eluted in rehydration buffer (1 mM EDTA, 10 mM TRIS–HCl, pH 7.5), quantified with Qubit (LifeTechnologies, Carlsbad, CA) and its quality was assessed using Nanodrop ND-1000 (Witech, Littau, Switzerland). The bacterial DNA was sequenced on PacBio RS (PacBio, USA) using two SMRT cells. The assembly was performed using HGAP2.0 (Chin et al. 2013) leading to a single chromosome sequence as well as a plasmid. The sequence of the complete chromosome has been deposited to the European Nucleotide Archive as an update to WGS sequence set CWGJ01000000 (BioProject Nr: PRJEB7018). The sequence of the plasmid was identical to the previously published sequence (LN867111).

Lausannevirus and *E. lausannensis* libraries were prepared with FC-131-1096 Nextera® XT DNA Sample Preparation Kit (Illumina, Zurich, Switzerland), FC-131-1001 Nextera® XT Index Kit (Illumina) and A63880 Agencourt AMPure XP-PCR Purification 5 ml (Beckman Coulter, Nyon, Switzerland). Samples were sequenced by MiSeq Illumina technology (150 bp paired-end reads), with reagents MS-102-2002 MiSeq Reagent Kit v2.0 (Illumina) and internal control FC-110-3001 PhiX Control V3 Kit (Illumina). To increase the genome coverage up to three independent sequencing were performed for each sample and all fastq files corresponding to a given time point were merged before the analysis (see supplementary table S3, Supplementary Material online).

Paired-end reads were controlled for quality using fastqc version 0.10.1 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/).

Genome Mapping

Each sample sequence was mapped against the reference genome of Lausannevirus (NC_015326.1) and the complete genome of *E. lausannensis*. Mapping was performed using Burrows-Wheeler Aligner (BWA) version 0.7.5a-r405 (Li and Durbin 2009), alignment files were further manipulated with Sequence Alignment/Map (SAM) Tools version 0.1.19-44428 (Li et al. 2009) and quality control was performed with QualiMap v.7.1 (Garcia-Alcalde et al. 2012).

SNP/Indel Calling

SNP and indel calling were performed on each sample using VarScan v2.3.6 (Koboldt et al. 2012) and Freebayes v0.9.9.2 (Garrison and Marth 2012). In both cases, default parameters were used. Variants were retained if detected by at least one of the two programs and further manually curated. All genomes were visualized with Circos (Krzywinski et al. 2009)

Genome Assembly and Lateral Gene Transfer

Reads were assembled with Edena version 130110 (http://www.genomic.ch/edena.php), Velvet version 1.2.09 (Zerbino and Birney 2008) and SPAdes version 2.5.1 (Bankevich et al. 2012), using different k-mer settings. The quality of genome assemblies was assessed with QUAST version 2.2 (Gurevich et al. 2013). The best quality assembly was retained for each sample. Lateral gene transfers were investigated by BLASTN (Altschul et al. 1990) of assembled subcultures genomes against Lausannevirus, *E. lausannensis* and *A. castellanii* strain Neff (GCF_000313135.1) reference genomes. Only sequences blasting at least two reference genomes (cutoff identity > 95%) were retained.

Data Deposition

The annotated Lausannevirus subcultures genomes La01, La02, C12, E12, F12, and H12 have been deposited to the European Nucleotide Archive (Study accession: PRJEB20258).

Mutation Rate

Mutation rate for each subculture were estimated as described in Wielgossa et al. (2012), according to the following equation:

$$\mu = \frac{(n_1 - n_2)}{(T_1 - T_2)}$$

where n_1 and n_2 represents the number of all mutations detected after 1 year (144 passages) of co-culture and the number of total mutations present in the starting culture, respectively. T_1 and T_2 are the ages, estimated in year or in number of passages, of the final and initial subculture.

The probability to observe one mutation in a given gene was estimated with the Poisson distribution.

PCR Analyses

Quantitative real-time PCR were used to quantify viral and bacterial copies present in each samples. The Step One PCR system (Applied Biosystems, Zug, Switzerland) was used. Primers and probe (Eurogentec, Seraing, Belgium) used for *E. lausannensis* detection were EstF (5′-ACACGTGCTACAATGGCCGGT-3′), EstR (5′-CCGGGAACGTATTCACGGCGTT-3′) and EstS (5′-FAM-CAGCCAACCCGTGAGGG-BHQ1-3′) as described previously (Lienard et al. 2011). For the detection of Lausannevirus, primers used were gariF (5′-CAACCCCATAATTCCCTCAATAC-3′′), gariR (5′-CCTGATCCTTATCTTCCCATTG-3′), and gariS (5′-CAACATCAAACACGCCACCCTTCCG-3′) at a concentration of 200 nM for the primers and 100 nM for the probe. The reaction mix contained primers and probe, 10 µl iTaq supermix ROX (Bio-Rad, Reinach, Switzerland), 5 µl of sample DNA and was completed with DNA-free water up to 20 µl. The cycling conditions were 95 °C for 3 min, followed by 45 cycles of 95 °C for 30 s and 60 °C for 1 min. Water was used as negative control and a constructed plasmid was used as positive control. All samples, negative controls and standard curve were analyzed in duplicate and experiments were repeated three times. Data were visualized and represented with iPython (Peréz and Granger 2007)

Functional Domain Analyses and 3D Modeling

Helicases (Accession numbers YP_004346969.1 and YP_004347006.1) functional domains were analyzed using Phyre2 (Kelley and Sternberg 2009). *E. lausannensis* PtsI (ELAC_1197) 3D structure were assessed by Phyre2 based on the available structure model of *E. coli* PtsI (PDB ID: 2HWG) and visualized with Chimera (Pettersen 2004). *E. lausannensis* binding and active sites were based on the *Criblamydia sequanensis* prediction retrieved from Uniprot (UniProt 2015).

PtsI Cloning, Expression, Purification and Enzymatic Activity Analyses

E. lausannensis PtsI sequences were cloned using the Gateway technology (Thermo Fisher Scientific). Open reading frames (ORFs) were amplified by PCR with *att*B primers: ELAC_1197_F_attB1 (5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGAACATGACCAAACTGGCGACG-3′) and ELAC_1197_R_attB2 (5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCCATGCTGGAAATGAAA-3′). The PCR conditions were 98 °C for 1 min, followed by 35 cycles of 98 °C for 10 s, 62 °C for 30 s and 72 °C for 30 s, with a final extension of 72 °C for 10 min. The entry clones were constructed by recombining the purified PCR products with the pDONR201 entry vector following manufacturer's instructions. Entry clones were further transformed in *E. coli* TOP10 competent cells, quantified using Nanodrop ND-1000 (Witech, Littau, Switzerland) and sequenced with Sanger technology, to verify the PtsI cloned sequences (wild type (WT) and variants). Expression clones were constructed by recombining the purified entry clones with the pBAD-DEST49 expression vector, which contain *ara*BAD promoter for protein expression and 6xHis tags for protein detection and purification. Further, expression clones were transformed in TOP10 competent cells. Protein expression was induced O/N at 25 °C with 0.02% L-Arabinose (Sigma–Aldrich, St. Louis, MO, USA). The expressed protein were purified in denaturing condition as follow: Bacterial pellets were resuspended in lysis buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl and 10 mM imidazole, pH 8.0) containing 1 mg/ml lysozyme, 1× Halt Protease Inhibitor (Pierce Biotechnology Inc., Rockford, USA), 5 mg/ml DNAseI and 8 mg/ml RNAseA and lysed by a freeze–defreeze (dry ice and ethanol versus 37 °C bath) step followed by short pulses of sonication on ice. After high speed centrifugation,

bacterial lysate was mixed with Ni-NTA Agarose beads (Invitrogen), loaded on polypropylene columns and let 1 h 30 at 4 °C on a wheel. Columns were washed four times with washing buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl and 20 mM imidazole, pH 8.0) and proteins eluted with elution buffer (50 mM $NaH_2PO_4.H_2O$, 300 mM NaCl and 250 mM imidazole, pH 8.0). Final concentration of purified protein was determined using a Bradford assay (Quick StartTM Bradford Protein Assay, Biorad, Hercules, USA). Purified recombinant proteins were verified by protein size (MW ~65 kDa) through immunodetection using mouse monoclonal anti-His6 antibody (Sigma–Aldrich). Pyruvate kinase activity of purified proteins was measured using Pyruvate Kinase Assay kit (Sigma), following manufacturer's protocol. Fluorescence were measured at $\lambda_{ex} = 540$ and $\lambda_{em} = 580$. PtsI WT and 352605 variant were tested at a concentration of 0.05 nmol. Two sample *t*-test were performed to statistically analyze differences among the different samples. All samples, positive controls and standards were analyzed in duplicate and experiments were repeated three times. Data were visualized and represented with iPython (Peréz and Granger 2007)

Results

Lausannevirus Shows a 13 kb Genome Deletion after One-Year of Co-Culture within *Acanthamoeba castellanii* in Allopatric Conditions

Two culture of Lausannevirus (La01 and La02) were independently propagated during one year (144 passages) within the amoeba *A. castellanii*. The amount of viral particles was estimated every 3 months by real time PCR (qPCR), in order to evaluate the evolution of the population size. Results showed high and stable population size for both replicates along the 12 months (fig. 1)

Both replicate sequenced reads were de novo assembled and mapped against the reference Lausannevirus genome. Data showed a deletion of ~13 kb in both subcultures, from 331,825 bp to 345,347 bp, representing 3.7% of Lausannevirus genome (see supplementary fig. S1, Supplementary Material online). The 12 months' populations lost the thymidylate synthase domain of the dihydrofolate-reductase thymidylate synthase protein (DHFR-TS), a protein involved in folate biosynthesis, and 18 hypothetical proteins located straight after.

To evaluate Lausannevirus microevolution, SNP and indel calling was performed. All variants supported by at least 10% of the reads were considered. La01 and La02 replicates showed a low number of mutations, eight (two SNPs, six indels) and 10 (two SNPs, eight indels), respectively (table 1). Two mutations in two intergenic regions were observed in both duplicates and all the remaining mutations were non-synonymous. Because of the absence of synonymous mutations, mutation rate was based only on the non-synonymous mutation observed and is of 1.6×10^{-7} mutation per nucleotide site per passage for La01 and 2×10^{-7} mutation per nucleotide site per passage for La02. Interestingly, all mutations observed in La01 replicate were also present in La02, which carry two supplementary indels in two hypothetical protein (table 1). Each replicate exhibited two fixed mutations.

Lausannevirus and *E. lausannensis* Evolution and Population Size Dynamics

We studied the origin, the maintenance and loss of mutations that appeared in mixed populations of Lausannevirus and *E. lausannensis* that were co-cultured for one year. We created independent replicates of the evolutionary process by splitting each mixed populations every 3 months into two different cultures to obtain eight different subcultures after 12 months (fig. 2*A*).

We estimated the evolution of microbial population size over one year by quantifying the amount of viral and bacterial particles present in each culture using real time PCR (qPCR). The Lausannevirus subcultures showed large fluctuations of population size (between 10^2 and 10^6 viral DNA copies/μl; fig. 2*B*), while all subcultures of *E. lausannensis* followed a similar pattern. The bacterial population size was low during the first three months, then increased dramatically at 6 months, and remained high at 9 and 12 months, even though differences up to one log of growth were detected among the subcultures (fig. 2*C*).

In parallel, Lausannevirus was co-cultured in *A. castellanii* with the massive addition of *E. lausannensis* at each passage in order to introduce a strong selection pressure (fig. 2*D*). Here, the viral DNA only was quantified. Population size fluctuated along the year of co-culture and at 12 months 10^6 viral DNA copies/μl were detected.

Striking Genome Stability of Lausannevirus When Co-Cultured One Year with *E. lausannensis*

Sequenced reads from the different subcultures were de novo assembled as well as mapped against the reference Lausannevirus genome. Because coverage issues, successful assemblies were obtained only for subcultures C12, E12, F12, and H12, which showed no genes duplication (see supplementary fig. S2, Supplementary Material online). Additionally, assemblies showed that only subculture E12 present a potential deletion, which was confirmed by mapping. Results showed a mixed subculture, with a large portion (> 98%) of the viral population carrying the same deletion detected in both duplicates of Lausannevirus allopatric cultures.

SNP and indel calling were performed to assess the microevolution of Lausannevirus along passages. Mutation rates, estimated for each subculture based on all synonymous and non-synonymous mutations and indels observed, varied

Table 1

Lausannevirus Mutations After 144 Passages

Position	Locus_Tag	Name	Subcultures[a]								Ref	Var	SYN/NON SYN	Gene Modification
			La01	La02	A12	C12	D12	E12	F12	H12				
119	LAU_0001	Putative helicase	+	+	−	−	−	+	−	−	GAACCAAA	G	NON SYN	Gene reduction
42324	LAU_0038	Putative helicase	+	+	−	−	−	−	−	−	A	AG	NON SYN	Gene reduction
81012	-	Intergenic region	+	+	−	−	−	−	−	−	T	A	-	-
81451	LAU_0099	Hypothetical protein	−	+	−	−	−	−	−	−	CG	C	NON SYN	-
86878	LAU_0110	Conserved putative secreted protein	−	−	+	−	−	−	−	−	A	C	NON SYN	-
94538	LAU_0121	Hypothetical protein	−	−	+	−	−	−	−	−	T	C	NON SYN	-
112448	LAU_0148	Hypothetical protein	−	−	−	−	−	+	−	−	CTGTTCTTCT	C	NON SYN	Gene reduction
117116	LAU_0158	Hypothetical protein	+	+	−	−	−	+	−	−	G	C	NON SYN	-
133481	-	Intergenic region	+	+	−	−	−	+	−	+	T	TC	-	-
1719	LAU_0001	Putative helicase	−	−	−	−	−	+	−	−	G	A	NON SYN	-
41136	LAU_0038	Putative helicase	−	−	−	−	−	-	−	+	T	G	NON SYN	-
193041	LAU_0247	Putative NUDIX hydrolase	−	−	−	−	−	+	−	−	T	C	NON SYN	-
118425	LAU_0160	Putative restriction endonuclease	−	−	−	−	−	+	−	−	T	C	SYN	-
195948	LAU_0251	Putative restriction endonuclease	−	−	−	−	−	+	−	+	AC	A	NON SYN	Gene restoration
195949	LAU_0251	Putative restriction endonuclease	+	+	−	+	+	−	+	−	C	CT	NON SYN	Gene extension
195970	LAU_0252	Putative restriction endonuclease	+	+	−	+	+	−	+	−	CG	C	NON SYN	-
235070	LAU_0299	Putative papain family	+	+	−	−	−	−	−	−	GAAGAGAACCC	G	NON SYN	Gene reduction
235621	LAU_0300	Hypothetical protein	−	+	−	−	−	−	−	−	A	AAAGTCTCCAGGACC	NON SYN	Gene reduction

[a]Subcultures B12 and G12 coverage after sequencing was insufficient to perform SNP/Indel calling.

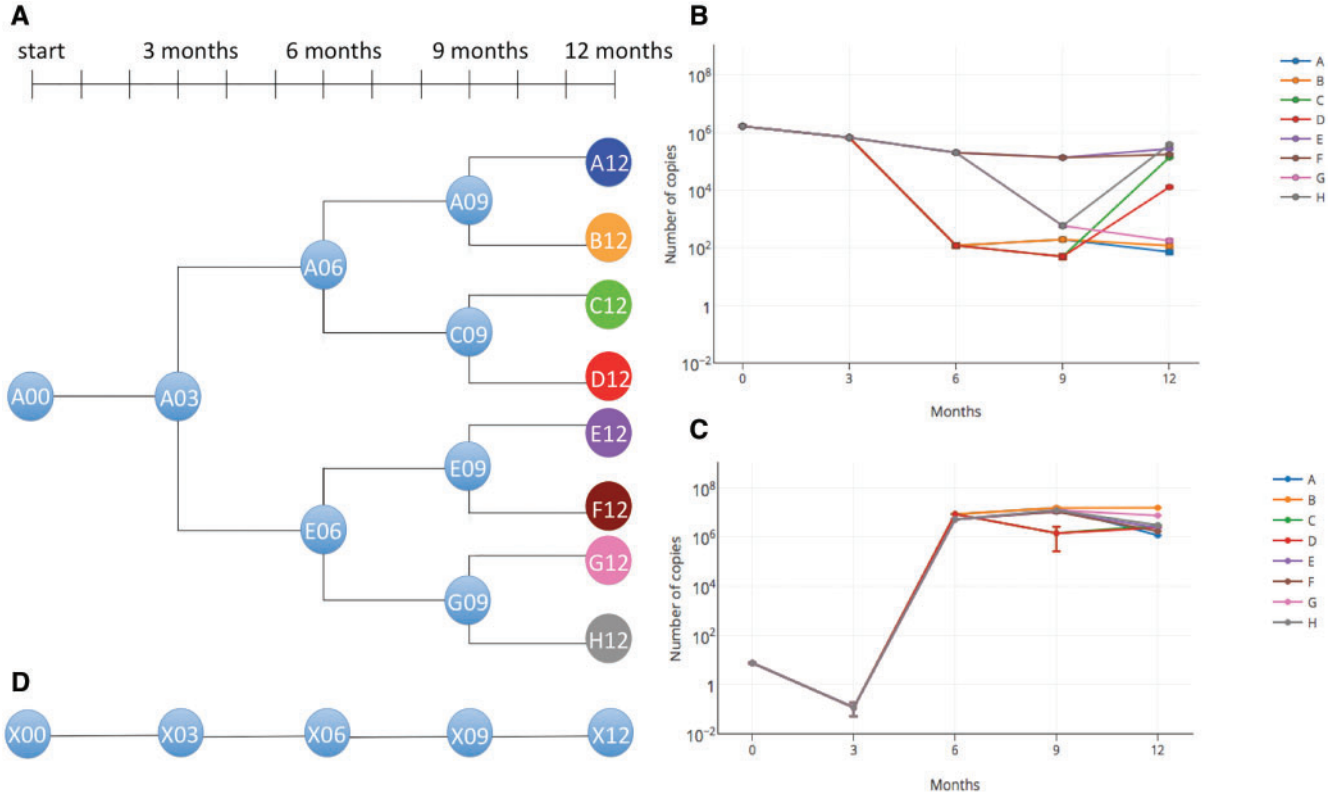

Fig. 2.—Lausannevirus and *E. lausannensis* subcultures during one year of co-culture in *A. castellanii*: schema and DNA copies number among the different subcultures. (*A*) Schema of Lausannevirus and *E. lausannensis* co-culture. Lausannevirus and *E. lausannensis* were maintained in co-culture with *A. castellanii* in 75 cm^2 flasks. Each three months (blue spots) the co-cultures were split into two different subcultures. The virus and the bacteria DNA were extracted and stored. (*B*) Number of Lausannevirus DNA copies in co-cultures. The number of viral DNA copies assessed by qPCR each 3 months before co-culture split is represented here in a log scale for all eight subculture (*A*–*H*) along time. Subcultures present very different number of viral DNA copies/l along the year of co-culture. Two main profile, high and low, are distinct, at 12 months only subcultures C12 show an intermediate profile, subcultures D12, E12, F12, and H12 have high profile while subcultures A12, B12, and G12 show low profiles. (*C*) Number of *E. lausannensis* DNA copies in co-cultures. As for the virus, the number of bacterial DNA copies was assessed by qPCR each 3 months. All the eight subcultures show similar profile along the year of co-culture, with low number of DNA copies/l until 3 months after which all subcultures increase dramatically the number of DNA copies/l and stay high until the end of the year. (*D*) Schema of Lausannevirus co-culture in massive presence of *E. lausannensis*. Lausannevirus was kept in co-culture one year (144 passages) with the addition of *E. Lausannensis* to the co-culture after each passage. Each month the cultures were purified and stored.

between 2.40×10^{-7} and 1.44×10^{-6} mutation per nucleotide site per month, corresponding to 2×10^{-8} and 1.20×10^{-7} mutation per nucleotide site per passage. After one year of co-culture, a total of 12 different positions harbored genetic changes among the different subcultures (table 1), including seven SNPs, two insertions and three deletions. A single mutation was localized in an intergenic region in two cultures, E12 and H12 (see supplementary table S1, Supplementary Material online). Among the 11 other unique mutations, 10 were non-synonymous whereas 1 was synonymous and were distributed in 10 protein-encoding sequences. The non-synonymous mutations arose in the sequences coding for three hypothetical proteins, two putative restriction endonucleases, two putative helicases, a conserved putative secreted protein and a putative NUDIX hydrolase, while the synonymous mutation occurred in a third putative restriction endonuclease (table 1).

In the parallel experiment, during which Lausannevirus was co-cultured in *A. castellanii* in massive presence of *E. lausannensis* (fig. 2*D*), only two indels (one insertion and one deletion) and no SNPs were detected after one year. However, these two indels were located in one of the three putative restriction endonuclease encoding genes detected in the previous experiment (position 195948 and 195949; fig. 3) and carry the same formerly found indels. Nevertheless, these two mutations were already present in the starting culture A00 at a low level, with a variant allele reads percentage of 15.97% and 7.54%, respectively (table 2). For both experiences, no Lausannevirus mutations reached fixation within the year of culture.

When looking at mutation distributions along the genome, we observed that most mutations occurred in gene present in all *Marseilleviridae*, but not conserved among all giant viruses (fig. 4)

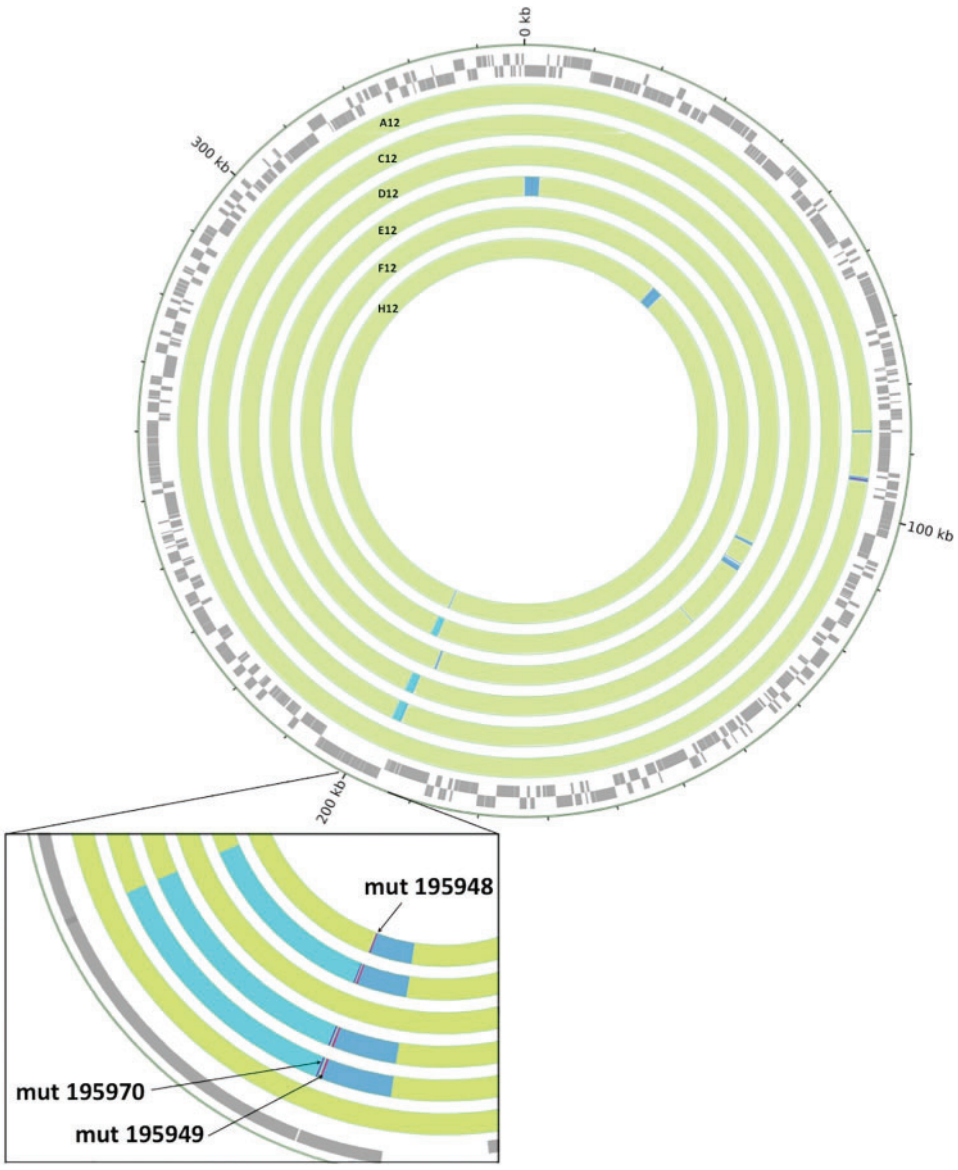

Fig. 3.—Circular representation of Lausannevirus mutations. The external circle represents Lausannevirus reference genome. Green circles from the outside to the inside: culture A12, culture C12, culture D12, culture E12, culture F12, culture H12. In blue and sky blue are genes in which a mutation (red) was detected. Zoom on the 195–200 kb region: mutations on two putative restriction endonucleases: mutations in positions 195949 (blue gene) and 195970 (sky blue gene) in subcultures C12, D12, and F12, mutation in position 195948 in subculture H12.

Lateral gene transfer was investigated by BLASTN of assembled subculture genomes against Lausannevirus and *E. lausannensis* reference genomes. No regions blasting both references with a high percentage of identity (cutoff $>$ 95%) were observed, suggesting that no gene transfer between Lausannevirus and *E. lausannensis* occurred during the year of co-culture. Moreover, lateral gene transfer between Lausannevirus and its host, *A. castellanii*, was investigated using the same strategy and the same cutoff. Data showed that no gene transfer occurred among Lausannevirus and *A. castellanii* during the 144 passages.

Putative Restriction Endonucleases and Helicases of Lausannevirus as Hot Spots for Mutations

Restriction endonucleases, also called restriction enzymes, are involved in the cleavage of sugar–phosphate bonds of DNA. With the exception of subculture A, all Lausannevirus sympatric subcultures showed at least one mutation located in a putative restriction endonuclease (table 2). A total of four mutations appeared during the sympatric year-long evolution experiment in three out of the seven putative restriction endonucleases present in the Lausannevirus genome (table 1). One mutation is synonymous (position 118425), while the

Table 2

Mutations in Putative Restriction Endonuclease Among Lausannevirus Cultures: Number of Variant Mapping Reads/Depth

LAU_0251 195949	Months of co-culture					subculture
	0	3	6	9	12	
C->CT	185/2454	1/544	-	-	1/27	A
					-	B
				-	98/248	C
					11/32	D
			-/31	-/16	2/131	E
					421/961	F
				-	-	G
					31/717	H
LAU_0252 195970	**Months of co-culture**					**subculture**
	0	3	6	9	12	
CG->C	178/1833	-/540	-	-	-/17	A
					-	B
				-	104/156	C
					15/25	D
			-/31	-/14	2/123	E
					427/544	F
				-	-	G
					26/658	H
LAU_0251 195948	**Months of co-culture**					**subculture**
	0	3	6	9	12	
AC -> A	364/2279	508/540	+	+	1/27	A
					+	B
				+	3/150	C
					-/21	D
			29/31	16/16	120/126	E
					2/538	F
				+	+	G
					430/689	H

PCR +/-	0% mapping reads	1-9.9%	10-24.9%	25-49.9%	>50%

+/- When the mapping coverage was insufficient, presence (+)/absence (-) of the mutation were verified by PCR

remaining three are non-synonymous and detected in two distinct proteins (two mutations in LAU_0251 and one in LAU_0252; table 1). At the end of the experiment, however, four subcultures (B, E, G, and H) carry only one of the two mutations detected in LAU_0251, the one in position 195948. In contrast, the other mutation of the same protein, in position 195949, was always detected together with the mutation in LAU_0252 in three subcultures (C, D, and F; fig. 3 and table 2). Both putative restriction endonucleases, LAU_0251 and LAU_0252, have the putative restriction endonuclease ISTM_159 of Insectomime virus as their closest homologue with, respectively, 95% and 94% amino acid identity. This result indicates that Lausannevirus probably carried a split gene (fig. 5). Lausannevirus deletion in position 195948, observed in the four subcultures mentioned above (table 2), caused a frameshift that partially restored the ancestral gene, while the three subcultures carrying the two linked indels (the insertion in position 195949 and the

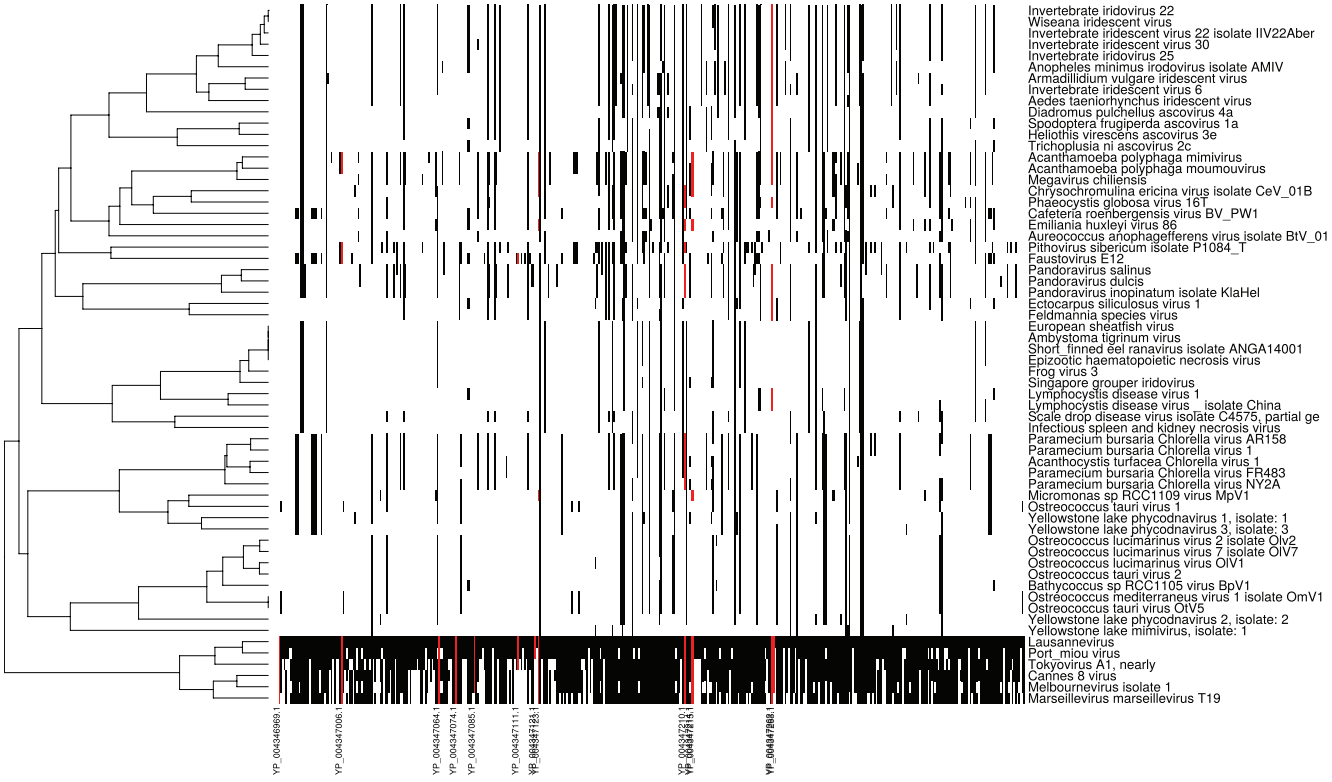

Fig. 4.—Mutations distribution along giant viruses. Representation of orthologue groups among giant viruses. Red colors are the protein in which at least one mutation was observed in Lausannevirus allopatric or sympatric co-cultures. Most mutations occurred in regions relatively conserved among *Marseilleviridae*, but not conserved among all giant viruses.

deletion in position 195970) elongated the LAU_0251 protein (fig. 5). The three subcultures carrying the elongated endonuclease (C, D, and F) showed an intermediate viral population size, while half of the cultures carrying the restored endonuclease (E12 and H12) exhibit a large population size and the other half have a small population size (B12 and G12; fig. 2*B* and table 2). Therefore, no link between these mutations and the population size of the virus could be observed, preventing any conclusion on the effect of these mutations on the activity of Lausannevirus putative restriction endonucleases. Curiously, the linked indels causing an elongation of the LAU_0251 protein, were observed also in the two replicates of Lausannevirus when co-cultured one year in allopatric conditions, suggesting that this mutations might not randomly appear. In addition to this mutation, we also observed four and five gene length reduction in allopatric and sympatric co-cultures, respectively (table 1).

In Lausannevirus sympatric subcultures, two SNPs (position 1719 and 41136) leading to non-synonymous amino acid replacements were detected in two different putative helicases (table 3). After 144 passages, mutation in position 41136 is present only in the subculture H (table 3), which is the only subculture that undergoes a dramatic growth between 9 and 12 months of culture (fig. 2*B*). Mutation in position 1719 is detected in three subcultures (B, E, and G), presenting high and low population sizes. Thus, a link between the detected SNPs and their effect on protein activity remained unclear. The same putative helicases showed a deletion (position 119) and an insertion (position 42324) also in the allopatric replicates, supporting the observation that these proteins are hot spots for mutations.

Estrella lausannensis Genome Microevolution

As performed for Lausannevirus, sequenced reads were mapped against *E. lausannensis* genome to enable SNP and indel calling. Total synonymous and non-synonymous mutation rates ranged between 8.66×10^{-8} and 2.89×10^{-7} mutation per nucleotide site per month, corresponding to 7.21×10^{-9} and 2.4×10^{-8} mutation per nucleotide site per passage, among the eight subcultures. After 12 months of co-culture a total of 21 different mutated positions were detected among the eight subcultures (table 4). The 16 SNPs and five deletions were distributed over two intergenic regions and 14 proteins, out of which 11 are hypothetical proteins. The others are a phosphoenolpyruvate–protein phosphotransferase (PtsI), a serine/threonine phosphatase and a ribosomal protein L11 methyltransferase (table 4).

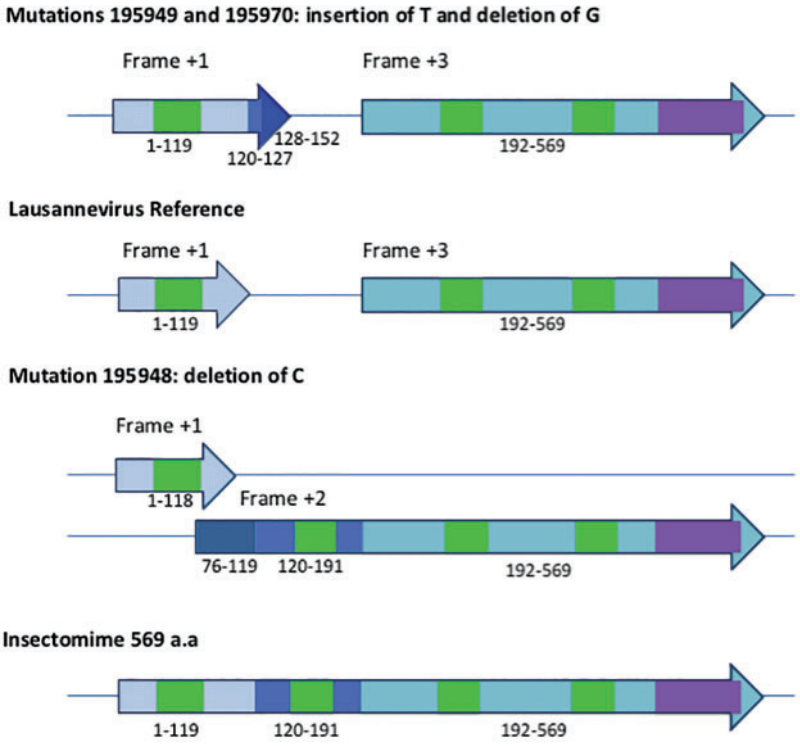

Fig. 5.—Translation of Lausannevirus variants putative restriction endonucleases and conserved domains detection. Homologous parts are highlighted by the same dotted blue part. The top translated variant, with mutations 195949/195970 shows no frameshift and carry the same conserved domain as the Lausannevirus reference genome. The deletion in position 195948 shows a correction of frameshift and encode for a longer protein compared with the Lausannevirus reference, with a larger region of homology with Insectomime putative restriction endonuclease (ISTM_159). In this variant, as in Insectomime, 4 DUF4379 domains and a restriction endonuclease domain were detected.

Out of the 19 mutations found in coding regions, 17 were non-synonymous, and only two were synonymous. While mutations were distributed along the whole genome, all subcultures at 12 months carried at least one mutation in the PtsI protein and in at least one of the different hypothetical proteins (table 4 and fig. 6). However, BLAST analyses against the non-redundant database showed that all the hypothetical proteins in which mutations were found have no homologs in the non-redundant database and no known domains. Additionally, the translated protein carrying variants showed no frameshift, suggesting that they are specific to *E. lausannensis* and may have essential role in its biology. Contrary to the viral genome where no mutations reached fixation, at least one mutation was fixed (cutoff $>=99\%$) in 6 out of 8 subculture of *E. lausannensis* (see supplementary table S2, Supplementary Material online). The fixed mutations (positions 1756009, 2149513, and 351939) were located in two different hypothetical protein-encoding genes and in the *ptsI* gene, respectively (table 4).

Lateral gene transfer among *E. lausannensis* and *A. castellanii* was investigated by BLASTN as described previously. No regions blasting both microorganisms references genome with a high percentage of identity (cutoff $>95\%$) were observed, suggesting that no gene transfer between *E. lausannensis* and its host occurred during the year of co-culture.

PtsI Showed Lower Activity in the Variant 352605 When Compared with the WT

Mutations in the sequence coding for the PtsI, a protein member of the phosphotransferase system (PTS) responsible for the transport of sugars across the cell membrane and their phosphorylation, were observed in all the eight subcultures at 12 months despite being absent at time 0 (table 4 and fig. 6). At 12 months, four variants were observed in this protein (table 5), which contains three phosphoenolpyruvate (PEP)-dependent domains: the PEP-utilizers superfamily N-terminal domain, the PEP-utilizers superfamily mobile element domain and a pyruvate kinase superfamily domain. All mutations are non-synonymous and localized in the pyruvate kinase domain. PtsI protein is present in all member of the *Chlamydiales* order with amino-acids sequence identities to *E. lausannensis* PtsI ranging from 31% (*Chlamydia trachomatis*) to 66% (*C. sequanensis*).

None of the known active or binding sites exhibit mutations; however all SNPs are closely located to them. Protein structure predictions of the PtsI WT sequence and of the four

Table 3
Putative Helicases Mutations Distribution Among Lausannevirus Cultures: Number of Variant Mapping Reads/Depth

LAU_0038 41136	Months of co-culture					subculture
	0	3	6	9	12	
T -> G	-/2729	14/1950	-	-	-/26	A
					-	B
				-	2/45	C
					-/12	D
			-/13	-/18	-/76	E
					-/179	F
				-	-	G
					517/1659	H
LAU_0001 1719	**Months of co-culture**					**subculture**
	0	3	6	9	12	
G -> A	3/3118	207/610	+	+	-/28	A
					+	B
				+	1/398	C
					-/56	D
			24/28	8/21	119/206	E
					1/2377	F
				+	+	G
					-/916	H

PCR +/-	0% mapping reads	1-9.9%	10-24.9%	25-49.9%	>50%

variants sequences detected in the eight subcultures were performed. Analyses revealed conformational changes in the 3D structure of the variant proteins that could further lead to changes in the protein activity.

In the prediction, 96% of the sequences were modeled with 100% confidence while the 4% remain poorly resolved. To identify which part of the 3D structure was modeled with low confidence, protein structure prediction of the same WT sequence were performed five times and the resulting structure were superposed. The four variants of PtsI protein carrying SNPs were modeled and the 3D structures were superposed to the five WT structures previously predicted, for a total of nine structures superposed (see supplementary fig. S3, Supplementary Material online). All binding sites had a conserved 3D structure in the nine aligned sequences, suggesting that a SNP effect on their conformation is unlikely. On the contrary, both active sites showed conformational differences in two independent predictions of the WT sequence. Thus, a mutation effect on the conformational changes of these sites could not be accurately estimated (see supplementary fig. S3, Supplementary Material online). Hence, cloning, expression and purification of the original PtsI and the four variants were performed. The variant 352605 and the WT PtsI were successfully expressed and purified, while in the same conditions the remaining three variants did not expressed the protein. Enzymatic test allowed the quantification (nmol) of pyruvate generated by the PtsI WT, the PtsI variant 352605, and a positive control along 60 min (fig. 7). Pyruvate kinase activity was evaluated during the linear rate (0–20 min) and showed similar results for the WT PtsI (0.22 mU, as nmol/min/ml) and the positive control (0.27 mU), whereas the PtsI variant 352605 exhibited a reduced activity (0.11 mU). Altogether these results suggest that the non-synonymous nucleotide substitution in position 352605 of *E. lausannesis* genome negatively affect the PtsI activity. At 12 months, only the variant 351939, became fixed in the subculture G (table 5). The other three variants are present in at least one subculture with a frequency higher than 50%. Altogether, these observations suggest that PtsI protein is under a positive selection pressure, since (1) all mutations are non-synonymous, (2) one mutation reached fixation, and (3) mutations occurred in the same gene in different subcultures.

Discussion and Perspectives

We investigated the evolution of Lausannevirus in presence and absence of a competitor. To the best of our knowledge, this is the first time that the evolution of a virus and an

Table 4

E. lausannensis Mutations After 144 Passages

Subculture	Position	Locus_Tag	Name	Ref	Var	SYN/NON-SYN
A12						
	352327	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	A	NON-SYN
	352605	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	G	NON-SYN
	1313553	ELAC_1521	Hypothetical protein	A	T	NON-SYN
[a]	1756009	ELAC_1407	Hypothetical protein	C	G	NON-SYN
[a]	2149513	ELAC_0517	Hypothetical protein	A	T	NON-SYN
	2751923	ELAC_0799	Hypothetical protein	A	G	NON-SYN
B12						
	352327	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	A	
	352605	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	G	NON-SYN
	1313553	ELAC_1521	Hypothetical protein	A	T	NON-SYN
[a]	1756009	ELAC_1407	Hypothetical protein	C	G	NON-SYN
[a]	2149513	ELAC_0517	Hypothetical protein	A	T	NON-SYN
C12						
	352327	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	A	NON-SYN
	352605	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	G	NON-SYN
	608804	ELAC_1987	Hypothetical protein	A	T	NON-SYN
	1313553	ELAC_1521	Hypothetical protein	A	T	NON-SYN
	1466692	ELAC_0123	Hypothetical protein	C	T	NON-SYN
[a]	1756009	ELAC_1407	Hypothetical protein	C	G	NON-SYN
	1784527	ELAC_1381	Hypothetical protein	A	T	NON-SYN
[a]	2149513	ELAC_0517	Hypothetical protein	A	T	NON-SYN
D12						
	352327	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	A	NON-SYN
	1313553	ELAC_1521	Hypothetical protein	A	T	NON-SYN
[a]	1756009	ELAC_1407	Hypothetical protein	C	G	NON-SYN
[a]	2149513	ELAC_0517	Hypothetical protein	A	T	NON-SYN
	2561900	ELAC_0665	Hypothetical protein	G	A	NON-SYN
E12						
	352411	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	G	A	NON-SYN
	807003	ELAC_1842	Ribosomal protein L11 methyltransferase	AG	A	NON-SYN
	1033975	ELAC_1731	Hypothetical protein	CA	C	NON-SYN
	2149627	ELAC_0517	Hypothetical protein	GC	G	NON-SYN
F12						
	352327	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	A	NON-SYN
	352605	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	C	G	NON-SYN
	1313553	ELAC_1521	Hypothetical protein	A	T	NON-SYN
	1466692	ELAC_0123	Hypothetical protein	C	T	NON-SYN
[a]	1756009	ELAC_1407	Hypothetical protein	C	G	NON-SYN
	1784527	ELAC_1381	Hypothetical protein	A	T	NON-SYN
[a]	2149513	ELAC_0517	Hypothetical protein	A	T	NON-SYN
	2751923	ELAC_0799	Hypothetical protein	A	G	NON-SYN
	608653	ELAC_1987	Hypothetical protein	CA	C	NON-SYN
	917764		Intergenic	TA	T	–
G12						
[a]	351939	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	A	G	NON-SYN
[a]	2149513	ELAC_0517	Hypothetical protein	A	T	NON-SYN
	2217706	ELAC_2079	Serine/threonine phospathase stp	G	T	NON-SYN
	917764		Intergenic	TA	T	–
H12						
	352411	ELAC_1197	Phosphoenolpyruvate–protein phosphotransferase	G	A	NON-SYN
	2507202	ELAC_0616	Hypothetical protein	C	T	SYN
	2624167	ELAC_0705	Hypothetical protein	C	T	SYN

(continued)

Table 4 Continued

Subculture	Position	Locus_Tag	Name	Ref	Var	SYN/NON-SYN
	2674969		Intergenic	C	T	–
	807003	ELAC_1842	Ribosomal protein L11 methyltransferase	AG	A	NON-SYN
	1033975	ELAC_1731	Hypothetical protein	CA	C	NON-SYN
	2149627	ELAC_0517	Hypothetical protein	GC	G	NON-SYN

[a]Fixed mutations.

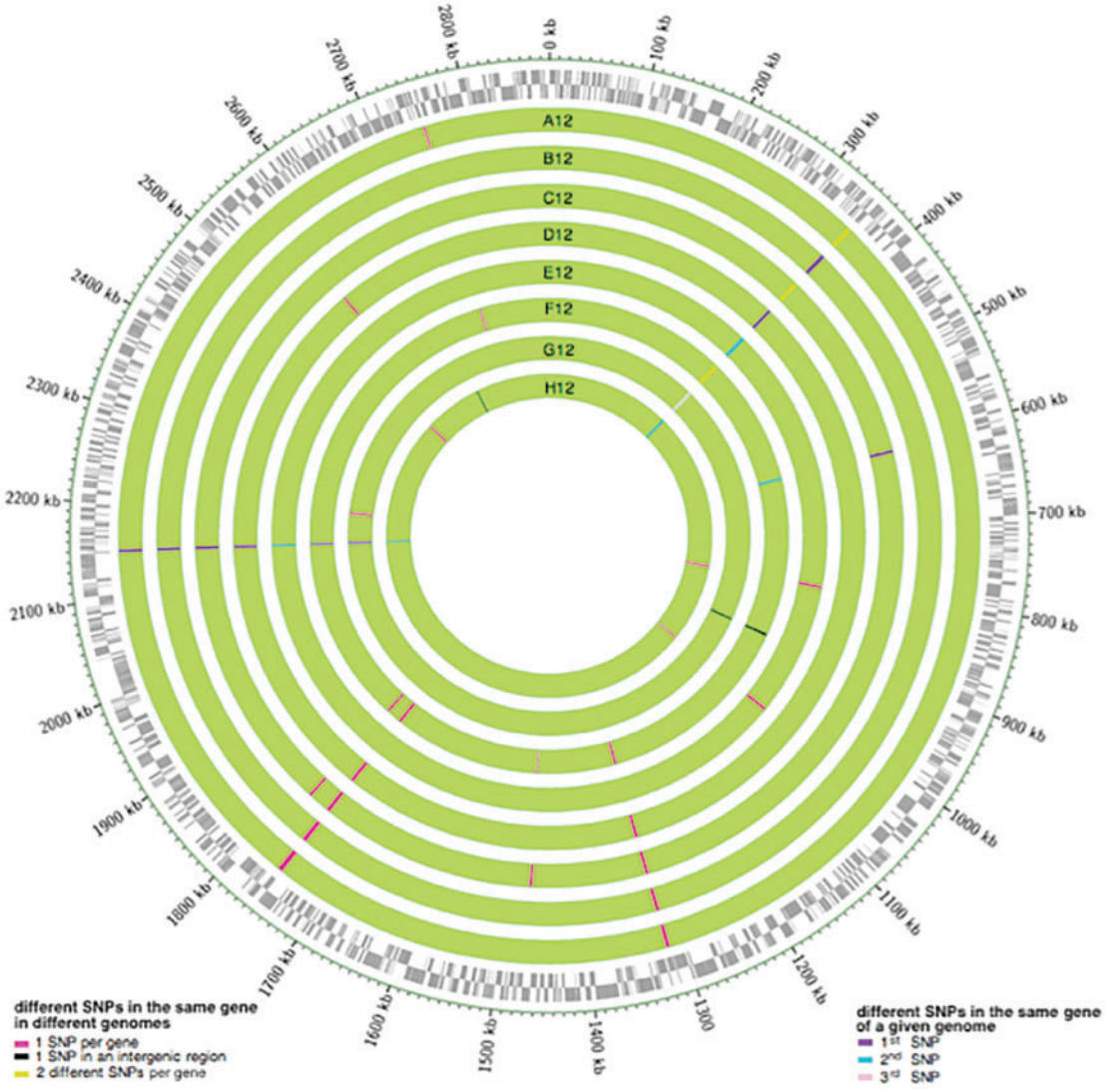

Fig. 6.—Circular representation of *E. lausannensis* mutations. Mutations among the eight *E. lausannensis* subcultures are represented. The external circle represents *E. lausannensis* reference genome. Green circles from the outside to the inside: culture A12, culture B12, culture C12, culture D12, culture E12, culture F12, culture G12, and culture H12. Mutations are distributed along the whole genome and several subcultures present mutations in the same gene.

intracellular bacterium co-cultured in the same host was reported. No horizontal gene transfer between Lausannevirus and *E. lausannensis* or between these two ARMs and their host *A. castellanii* were identified during one year of co-culture. Instead, a low number of mutations were observed in all, allopatric and sympatric subcultures. Lausannevirus exhibited a range from 1.6×10^{-7} to 2×10^{-7} and from 2×10^{-8} to 1.20×10^{-7} mutation per nucleotide site per passage in allopatric and sympatric co-cultures, respectively. A range from 7.21×10^{-9} to 2.4×10^{-8} mutation per

Table 5

Phosphoenolpyruvate–protein Phosphotransferase Mutations in *E. lausannensis*: Number of Variant Mapping Reads/Depth

ELAC_1197 352327	Months of co-culture 0	3	6	9	12	subculture
C->A	reference	-	8/557	20/294	295/492	A
					264/326	B
				164/226	160/253	C
					236/273	D
			-/20	-/305	-/240	E
					20/49	F
				1/442	-/283	G
					-/159	H
ELAC_1197 352411	**Months of co-culture 0**	**3**	**6**	**9**	**12**	**subculture**
G->A	reference	-	1/497	1/293	2/404	A
					-/314	B
				-/199	-/230	C
					1/244	D
			-/7	76/303	206/234	E
					-/48	F
				121/434	1/255	G
					137/152	H
ELAC_1197 352605	**Months of co-culture 0**	**3**	**6**	**9**	**12**	**subculture**
C->G	reference	-	45/415	110/243	150/374	A
					51/276	B
				46/216	70/228	C
					23/237	D
			-/27	1/297	-/196	E
					24/47	F
				-/380	-/228	G
					-/146	H
ELAC_1197 351939	**Months of co-culture 0**	**3**	**6**	**9**	**12**	**subculture**
A->G	reference	-	1/599	-/297	1/554	A
					-/337	B
				-/219	-/266	C
					-/319	D
			-/24	2/336	-/258	E
					-/42	F
				3/500	231/233	G
					9/174	H
ELAC_1197 352312	**Months of co-culture 0**	**3**	**6**	**9**	**12**	**subculture**
G->A	reference	+	118/512	90/264	9/463	A
					1/304	B
				6/213	-/238	C
					-/257	D
			-/18	1/280	-/226	E
					-/48	F
				-/417	-/271	G
					-/146	H

PCR +/-	0% mapping reads	1-9.9%	10-24.9%	25-49.9%	>50%

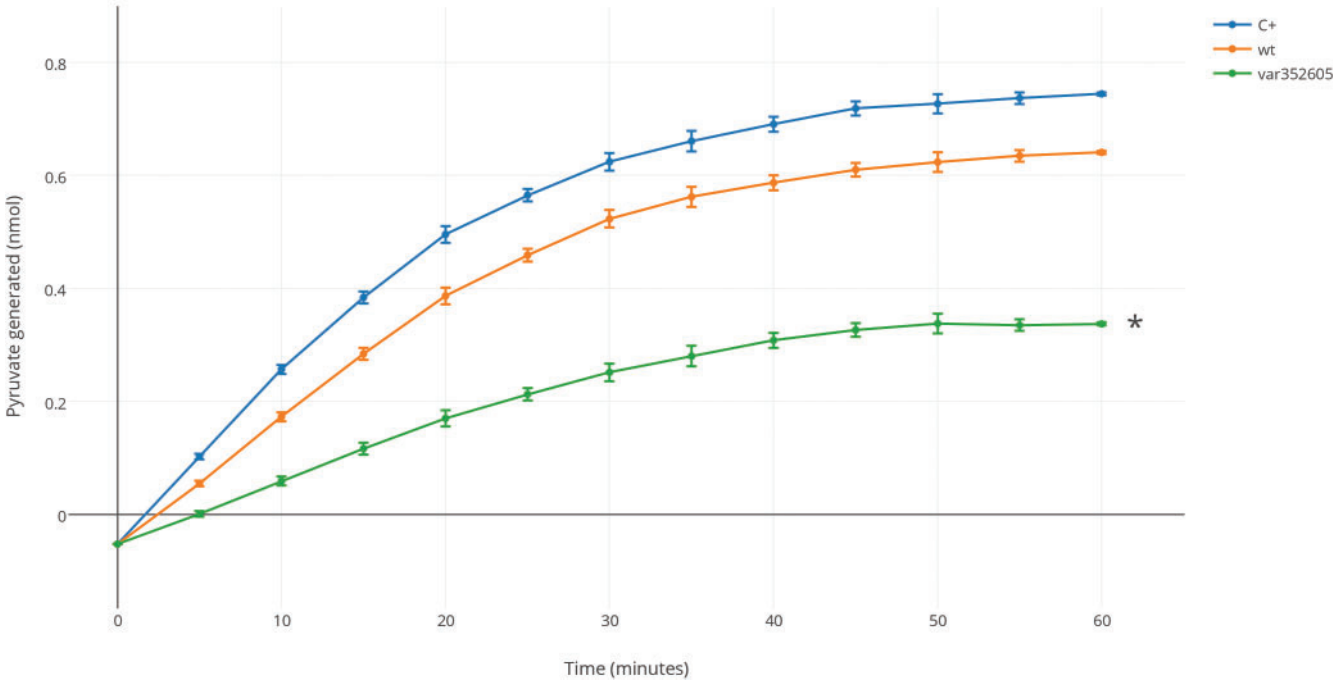

Fig. 7.—Pyruvate generated between 0 and 60 min by the PtsI WT and the PtsI variant 352605. Amount (nmol) of pyruvate generated between 0 and 60 min. Among the WT (orange) and the positive control (blue) no significant difference were detected (*P*-value = 0.37). PtsI variant 352605 (green) showed significant (*) difference in pyruvate generation when compared with the WT (*P*-value = 0.01).

nucleotide site per passage were observed in *E. lausannensis*. In Lausannevirus sympatric subcultures no mutations became fixed, whereas among the 49 observed mutations in *E. lausannensis*, 12 reached the fixation.

E. lausannensis genome showed a range of 4–10 mutations per subculture per year, similar to the six mutations observed in *E. coli* after approximately one year of subculture (2,000 generations; Barrick et al. 2009). Mutations in the PtsI gene (ELAC_1197), encoding the enzyme I (EI) appeared at different positions in *E. lausannensis* along the passages. This enzyme is part of the PTS system, a key player of sugar metabolism, and initiates the phosphorylation chain of sugars during their uptake across the cell membrane. Enzymatic activity tests revealed a 50% reduced activity of the PtsI variant 352605 compared with the PtsI WT, demonstrating that this mutation in the pyruvate kinase domain of the PtsI protein may impact the aerobic glycolysis and the replicative fitness of *E. lausannensis*. The presence of several different nonsynonymous mutations in the same PtsI gene in all eight different subcultures at 12 months reinforce the likely importance of this enzyme in defining the fitness of *E. lausannensis*. Furthermore, only one mutation (position 351939) was fixed, and a large majority of nonsynonymous over synonymous mutations was observed along the genome, suggesting that *E. lausannensis* genome evolution might be subjected to positive selection.

Theoretically, each culture passage—performed three times per week—should create a bottleneck effect. The small population remaining after each passage would be subjected to genetic drift leading to the possible loss of less conserved alleles and the observation of more fixed mutations. However, this hypothesis is in opposition with the little number of substitutions observed.

Interestingly, allopatric replicates showed no synonymous mutations, a single synonymous mutation was observed in Lausannevirus sympatric subcultures (table 1) and two synonymous mutations in *E. lausannensis* (table 4), suggesting that (1) neutral mutations are rapidly eliminated or (2) that both genomes might be under positive selection. Nevertheless, the variant detection cutoff was set to 10% of the mapping reads and, therefore, in a population of 10^6 DNA copies/μl a mutant has to be present in at least 10^5 DNA copies/μl to be detected. Because the sensitivity varies with the sequencing depth at each time point, mutants present in low frequency would escape detection, and some mutations may be missed. Altogether these observations might suggest that the large populations observed are little sensitive to genetic drift and might evolve under positive selection, as previously observed for some RNA viruses, like foot and mouth disease virus (FMDV; Haydon et al. 2001), H3N2 human influenza A virus (Suzuki 2006) and Human Immunodeficiency Virus (HIV) type 1 (Chen et al. 2004). These observations are in contrast with a previous study conducted on Melbournevirus, another member of the *Marseilleviridae* family, for which purifying selection were reported for most of its genes (Doutre et al. 2014).

As observed for Modified vaccinia Ankara (MVA; Mayr et al. 1975, 1978) and Mimivirus (Boyer et al. 2011), Lausannevirus also exhibit a genome reduction event in allopatric condition. A deletion being observed in both replicates with a coverage of zero in that region, as compared with a mean coverage of 9,426 for La01 and 5,758 for La02 (see supplementary fig. S1, Supplementary Material online). However, in sympatric conditions, the same deletion was observed in a single subculture (out of 8) (see supplementary fig. S2, Supplementary Material online), with ~1.2% of viral particles still exhibiting the full length genome. This might suggest that the presence of *E. lausannensis* in the same co-culture might slow down Lausannevirus genetic loss. Interestingly, MVA was obtained after 530 passages of Vaccinia virus strain Ankara (CVA WT) in chicken embryo fibroblast (CEFs; Mayr et al. 1975). MVA undergoes several mutations affecting the host interactive proteins (Antoine et al. 1998) and showed six major deletions as well as severe host range restriction (Meyer et al. 1991). Nevertheless, the very narrow host range of Lausannevirus was already known before our experiences (Thomas et al. 2011). The loss of the TS domain of Lausannevirus DHFR-TS protein might suggest that this region is not essential for a successful replication inside *A. castellanii*.

The current experiments showed that mutations were distributed all along Lausannevirus genome and not clustered at genome ends. Additionally, they were observed in regions relatively conserved among *Marseilleviridae* but not among core genes shared by most giant viruses (fig. 4). However, mutations occurred in specific gene functions. 50% of mutations were detected in putative restriction endonucleases, 16.6% occurred in hypothetical proteins, 11.1% in putative helicases, 11.1% in intergenic regions, 5.6% in putative NUDIX hydrolase, and 5.6% in a conserved putative secreted protein.

Despite 11.1% of mutations were observed in putative helicases, none of them were localized in the core D6/11-like helicase, being well conserved and thus less subjected to mutations.

Furthermore, among the seven putative restriction endonucleases present in Lausannevirus, three presented mutations after 12 months of sympatric co-culture. Those mutations were observed in all the subcultures after 144 passages, except for subculture A. Altogether, these observations suggest a non-random appearance of mutations in Lausannevirus putative restriction endonuclease-encoding genes. Interestingly, in Chloroviruses restriction endonucleases may play a major role in host DNA degradation during the early stages of infection (Agarkova et al. 2006) highlighting the relevant role of such enzymes in viral replication.

In sympatric subcultures, two opposite types of viral populations were observed: high (10^6 viral copies/μl) or low (10^2 viral copies/μl) population size. High viral load leads to a rapid burst in replication with, as a consequence, a rapid host cell lysis, while low viral load allows the host to survive longer, resulting in slower viral replication. Lausannevirus frequent change in population size among the different subcultures suggests that this virus may possesses the two different replication strategies. The mutation in a putative Nudix hydrolase could be a factor explaining the high viral load observed in subculture E12. Indeed, the Nudix hydrolase of African Swine fever virus has been shown to have a mRNA decapping activity (Parrish 2009), therefore enabling the virus to control viral/host mRNA ratio and to bias this ratio to produce large amount of viral mRNA. However, we could not evidence the presence of a single genetic change associated with low and high viral loads in all subcultures. Thus, we hypothesize that the choice of replication strategy may be most likely tightly regulated by environmental factors such as for example nutrients availability, or under the model of bistable genetic switches. Because viral populations rather than single clones were passed in culture, an alternative hypothesis forecast the presence of two or more variants among the same subcultures, able to inhibit or enhance reciprocally the replication of the other population type, resulting in population size variation. Further studies are needed to test these hypotheses.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

The computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311). Finally, we thank Sébastien Aeby for his precious technical help. This project was mainly supported by a interdisciplinary grant from the University of Lausanne to G. Greub and N. Salamin.

Literature Cited

Agarkova IV, Dunigan DD, Van Etten JL. 2006. Virion-associated restriction endonucleases of chloroviruses. J Virol. 80:8114–8123.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 3:403–410.

Antoine G, Scheiflinger F, Dorner F, Falkner FG. 1998. The complete genomic sequence of the modified Vaccinia Ankara strain: comparison with other orthopoxviruses. Virology 244:365–396.

Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19:455–477.

Barrick JE, et al. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature 461:1243–1247.

Bertelli C, et al. 2015. Sequencing and characterizing the genome of *Estrella lausannensis* as an undergraduate project: training students and biological insights. Front Microbiol. 6:101.

Bertelli C, Greub G. 2012. Lateral gene exchanges shape the genomes of amoeba-resisting microorganisms. Front Cell Infect Microbiol. 2:110.
Boratto PV, et al. 2014. Amoebas as mimivirus bunkers: increased resistance to UV light, heat and chemical biocides when viruses are carried by amoeba hosts. Arch Virol. 159:1039–1043.
Boyer MA, et al. 2011. Mimivirus shows dramatic genome reduction after intraamoebal culture. Proc Natl Acad Sci U S A. 108:10296–10301.
Bremer H. 1982. Variation of generation times in Escherichia co6 populations: its cause and implications. J Gener Microbiol. 128:2865–2876.
Chen L, Perlina A, Lee CJ. 2004. Positive selection detection in 40,000 human immunodeficiency virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations in HIV protease and reverse transcriptase. J Virol. 78:3722–3732.
Chin CS, et al. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 10:563–569.
Corsaro D, et al. 2009. Novel Chlamydiales strains isolated from a water treatment plant. Environ Microbiol. 11:188–200.
Doutre G, Philippe N, Abergel C, Claverie JM. 2014. Genome analysis of the first *Marseilleviridae* representative from Australia indicates that most of its genes contribute to virus fitness. J Virol. 88:14340–14349.
Forterre P, Krupovic M, Prangishvili D. 2014. Cellular domains and viral lineages. Trends Microbiol. 22:554–558.
Garcia-Alcalde F, et al. 2012. Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 28:2678–2679.
Gimenez GB, et al. 2011. Insight into cross-talk between intra-amoebal pathogens. BMC Genomics. 12:542.
Greub G. 2009. Parachlamydia acanthamoebae, an emerging agent of pneumonia. Clin Microbiol Infect. 15:18–28.
Greub G, Raoult D. 2004. Microorganisms resistant to free-living amoebae. Clin Microbiol Rev. 17:413–433.
Greub G, La Scola B, Raoult D. 2004. Amoebae-resisting bacteria isolated from human nasal swabs by amoebal coculture. Emerg Infect Dis. 8:470–477.
Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.
Haydon DT, Bastos AD, Knowles NJ, Samuel AR. 2001. Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. Genetics 157:7–15.
Jacquier N, Aeby S, Lienard J, Greub G. 2013. Discovery of new intracellular pathogens by amoebal coculture and amoebal enrichment approaches. J Vis Exp 80:51055.
Kebbi-Beghdadi C, Greub G. 2014. Importance of amoebae as a tool to isolate amoeba-resisting microorganisms and for their ecology and evolution: the Chlamydia paradigm. Environ Microbiol Rep. 4:309–324.
Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc. 4:363–371.
Koboldt DC, et al. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 22:568–576.
Krzywinski M, et al. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.
Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.
Lienard J, Croxatto A, Prod'hom G, Greub G. 2011. *Estrella lausannensis*, a new star in the Chlamydiales order. Microbes Infect. 13:1232–1241.
Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. https://arxiv.org/abs/1207.3907v2.
Mayr A, Hochstein-Mintzel V, Stickl H. 1975. Abstammung, eigenschaften und verwendung des attenuierten vaccinia-stammes MVA. Infection 3:6–14.
Mayr A, Stickl H, Müller HK, Danner K, Singer H. 1978. Der pockenimpfstamm MVA: Marker, genetische struktur, erfahrungen mit der parenteralen schutzimpfung und verhalten im abwehrgesch- wächten organismus. Zbl Bakt Hyg. 167:375–390.
Meyer H, Sutter G, Mayr A. 1991. Mapping of deletions in the genome of the highly attenuated vaccinia virus MVA and their influence on virulence. J Gener Virol. 72:1031–1038.
Moliner C, Fournier P-E, Raoult D. 2010. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. FEMS Microbiol Rev. 34:281–294.
Parrish S, Hurchalla M, Liu SW, Moss B. 2009. The African Swine Fever Virus g5R Protein Possesses mRNA Decapping Activity. Virology. 1:177–182.
Peréz F, Granger BE. 2007. IPython: a system for interactive scientific computing. Computing in Science & Engineering. 9:21–29.
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem. 13:1605–1612.
Raeside C, et al. 2014. Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. MBio 5:e01377–e01314.
Raoult DA, et al. 2004. The 1.2-megabase genome sequence of Mimivirus. Science 306:1344–1350.
Scheid P, Balczun C, Schaub GA. 2014. Some secrets are revealed: parasitic keratitis amoebae as vectors of the scarcely described pandoraviruses to humans. Parasitol Res. 113:3759–3764.
Suzuki Y. 2006. Natural selection on the influenza virus genome. Mol Biol Evol. 23:1902–1911.
Thomas V, Herrera-Rimann K, Blanc DS, Greub G. 2006. Biodiversity of amoebae and amoeba-resisting bacteria in a hospital water network. Appl Environ Microbiol. 72:2428–2438.
Thomas V, Loret JF, Jousset M, Greub G. 2008. Biodiversity of amoebae and amoebae-resisting bacteria in a drinking water treatment plant. Environ Microbiol. 10:2728–2745.
Thomas V, et al. 2011. Lausannevirus, a giant amoebal virus encoding histone doublets. Environ Microbiol. 13:1454–1466.
UniProt C. 2015. UniProt: a hub for protein information. Nucleic Acids Res. 43:D204–D212.
Wielgossa SB, et al. 2012. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. Proc Natl Acad Sci U S A. 110:222–227.
Williams TA, Embley TM, Heinz E. 2011. Informational gene phylogenies do not support a fourth domain of life for nucleocytoplasmic large DNA viruses. PLoS ONE. 6:e21080.
Winiecka-Krusnell J, Linder E. 1999. Free-living amoebae protecting legionella in water: tip of an iceberg? Scand J Infect Dis. 31:383–385.
Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18:821–829.

Associate editor: Chantal Abergel