

RESEARCH ARTICLE

Open Access



Identification of African swine fever virus-like elements in the soft tick genome provides insights into the virus' evolution

Jan H. Forth¹ , Leonie F. Forth¹ , Samantha Lycett², Lesley Bell-Sakyi³ , Günther M. Keil¹ [^], Sandra Blome¹, Sébastien Calvignac-Spencer⁴, Antje Wissgott⁵, Johannes Krause⁵, Dirk Höper¹ , Helge Kampen¹  and Martin Beer^{1*} 

Abstract

Background: African swine fever virus (ASFV) is a most devastating pathogen affecting swine. In 2007, ASFV was introduced into Eastern Europe where it continuously circulates and recently reached Western Europe and Asia, leading to a socio-economic crisis of global proportion. In Africa, where ASFV was first described in 1921, it is transmitted between warhogs and soft ticks of the genus *Ornithodoros* in a so-called sylvatic cycle. However, analyses into this virus' evolution are aggravated by the absence of any closely related viruses. Even ancient endogenous viral elements, viral sequences integrated into a host's genome many thousand years ago that have proven extremely valuable to analyse virus evolution, remain to be identified. Therefore, the evolution of ASFV, the only known DNA virus transmitted by arthropods, remains a mystery.

Results: For the identification of ASFV-like sequences, we sequenced DNA from different recent *Ornithodoros* tick species, e.g. *O. moubata* and *O. porcinus*, *O. moubata* tick cells and also 100-year-old *O. moubata* and *O. porcinus* ticks using high-throughput sequencing. We used BLAST analyses for the identification of ASFV-like sequences and further analysed the data through phylogenetic reconstruction and molecular clock analyses. In addition, we performed tick infection experiments as well as additional small RNA sequencing of *O. moubata* and *O. porcinus* soft ticks.

Conclusion: Here, we show that soft ticks of the *Ornithodoros moubata* group, the natural arthropod vector of ASFV, harbour African swine fever virus-like integrated (ASFLI) elements corresponding to up to 10% (over 20 kb) of the ASFV genome. Through orthologous dating and molecular clock analyses, we provide data suggesting that integration could have occurred over 1.47 million years ago. Furthermore, we provide data showing ASFLI-element specific siRNA and piRNA in ticks and tick cells allowing for speculations on a possible role of ASFLI-elements in RNA interference-based protection against ASFV in ticks. We suggest that these elements, shaped through many years of co-evolution, could be part of an evolutionary virus-vector 'arms race', a finding that has not only high impact on our understanding of the co-evolution of viruses with their hosts but also provides a glimpse into the evolution of ASFV.

Keywords: African swine fever virus, Coevolution, Endogenous viral element, Tick cell line, RNA interference, ASFLI-element

* Correspondence: martin.beer@fli.de

[^]Günther M. Keil is deceased.

¹Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Südufer 10, 17493 Greifswald-Insel Riems, Germany

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The pandemic spread of viral pathogens affecting humankind has been recognised for many years as one of the most dangerous scenarios leading to crises of global proportion. Furthermore, the ‘One Health’ concept of today is arguing that pathogens affecting animals or plants may also have a major impact on the human population [1, 2]. One recent example of such interrelations with an extreme global socio-economic impact is the unprecedented pandemic spread of African swine fever virus (ASFV), one of the most devastating viral diseases of animals [3].

ASFV, a double-stranded DNA (dsDNA) virus, can infect all kinds of suids and leads to a multi-systemic disease, African swine fever (ASF), in non-natural suid hosts with clinical signs of a viral haemorrhagic fever characterised by morbidity and case fatality rates of up to 100% [4, 5]. Despite almost 100 years of intensive research and the occurrence on four continents [6], neither a vaccine nor any treatment is available. Therefore, the introduction of ASFV into the wild boar populations of Eastern Europe in 2007 has led to an unprecedented epidemiological situation that culminated in the introduction of ASFV into Western Europe and Asia [7–9]. Especially in China, the largest producer of pork in the world, extreme socio-economic consequences can be observed, impacting not only Asia, but affecting the global agriculture and food industry and shifting entire markets [3, 10].

Nevertheless, very little is known about ASFV evolution. ASFV appears to be the only member of its genus (*Asfivirus*) and family (*Asfarviridae*) and is the only known DNA arbovirus [4]. First described from, and endemic in, sub-Saharan Africa [11], ASFV is transmitted in a sylvatic cycle between soft ticks of the *Ornithodoros moubata* complex and indigenous wild pigs such as African warthogs (*Phacochoerus africanus*) [12]. Both groups of natural hosts are well adapted to an ASFV infection and, although persistently infected [12], show no obvious pathological signs, which has led to the hypothesis that—similar to other arboviruses—ASFV and its hosts have undergone a long time of co-evolution. Since viruses leave no fossil records, analyses into the evolution of ASFV have been restricted to highly similar genotypes of which only very few whole-genome sequences are available [4, 13]. However, with the recent advances in sequencing technologies, novel ways of analysing virus evolution have been discovered, with one of those being the analysis of endogenous viral elements (EVE)—ancient viral sequences integrated into the host genome [14].

In our study, we identified African swine fever virus-like integrated (ASFLI)-elements in the uncharacterised genome of *O. moubata* complex soft ticks. We show ASFLI-elements in tick cell lines [15], *O. moubata* and *O. porcinus* soft ticks, and in museum-stored

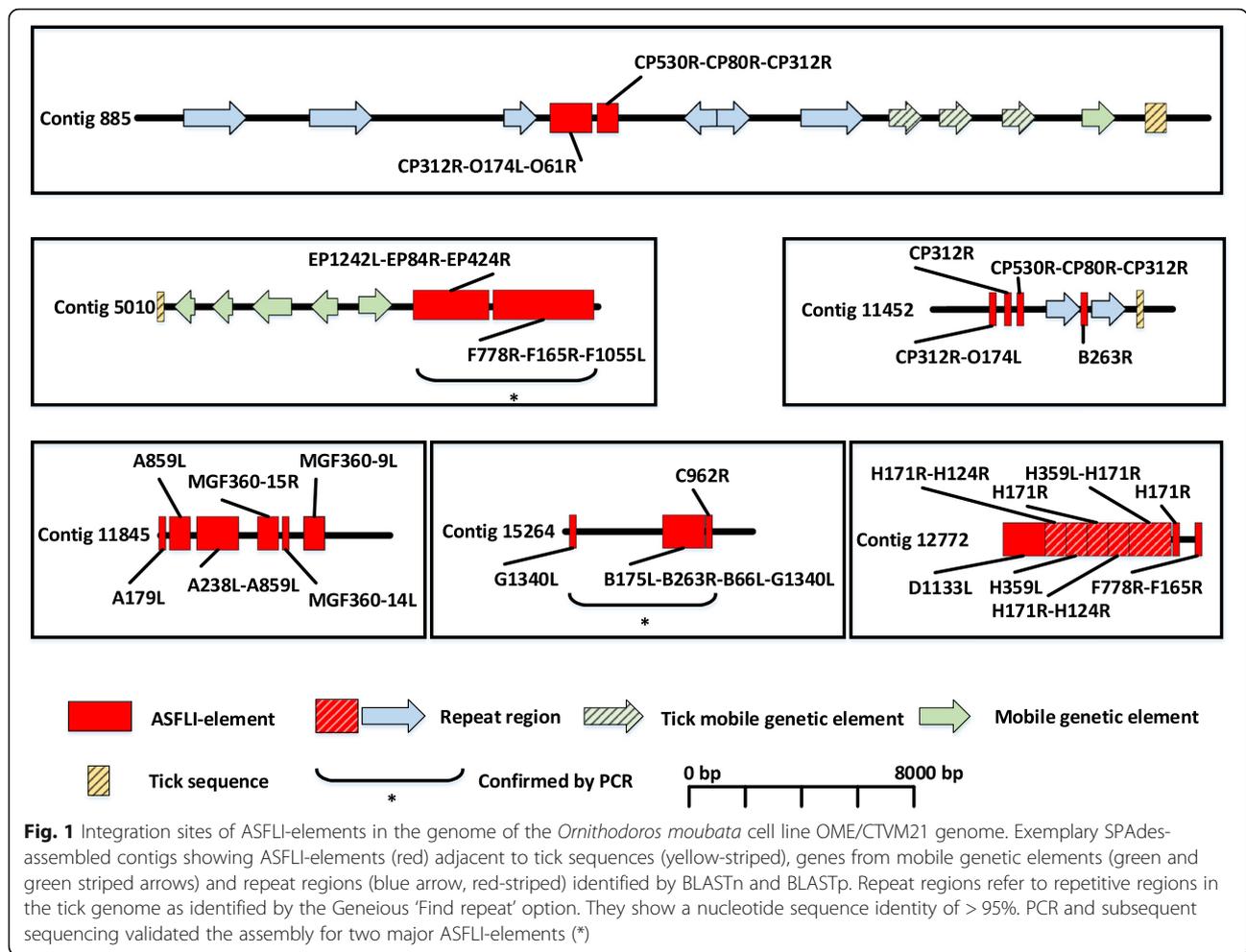
Ornithodoros ticks collected about 100 years ago in the field in East Africa, allowing for phylogenetic reconstruction and orthologous dating. To evaluate whether the ASFLI-elements retained a function, we conducted tick infection experiments, transcriptome analyses and small RNA sequencing and expressed a highly conserved ASFV-like gene from the tick genome for further characterisation.

Results

Evidence of ASFLI-elements in the *O. moubata* tick cell genome

In order to test for the presence of either novel DNA viruses or integrated viral elements, we sequenced DNA from the *O. moubata* cell line OME/CTVM21 [15] and combined reads from different sequencing platforms. In detail, 14.9 million paired-end reads with 300-bp read length (Illumina MiSeq) and 277 million single-end reads with 50-bp read length (Illumina HiSeq) were combined with about 500,000 3rd-generation sequencing single-molecule, ultra-long reads (MinION) (Additional file 1: Table S1).

Using the SPAdes v3.11.1 software [16] for the assembly and BLASTn for the detection of viral sequences, we identified 34 contigs with lengths between 336 and 85,582 bp containing ASFLI-elements (Fig. 1 and Additional files 2, 3 and 4: Tables S2–S4). However, we detected truncated ASFV-like ORFs as well as sequence duplications and reorganisations when comparing ASFLI-elements with the homologous ASFV genes and the ASFV genome organisation (Figs. 1 and 2 and Additional files 3 and 4: Tables S3 and S4). For validation of the assembly, two large ASFLI-elements (6.8 kb and 4.2 kb) were amplified by PCR and subsequently sequenced in a shotgun sequencing approach (Illumina MiSeq) followed by assembly using Newbler 3.0 (Roche) and alignment to the previously assembled SPAdes contigs using MAFFT v7.388 in Geneious (Fig. 1). The BLASTn search as well as the analysis of the >500-bp ORFs by BLASTp resulted in the identification of contigs showing both integrated viral elements and sequences with similarity to the genome of the black-legged tick *Ixodes scapularis*, the only tick species with a fully sequenced and partially annotated genome [17]. Moreover, contigs were identified that cover multiple ORFs with similarity to mobile genetic elements (Fig. 1 and Additional files 3 and 4: Tables S3 and S4). Altogether, we were able to identify ASFLI-elements homologous to 46 ASFV genes, covering 14 genes completely and 32 genes partially, with identities to recent ASFV sequences of up to 86.6% (Fig. 2 and Additional files 3 and 4: Tables S3 and S4).

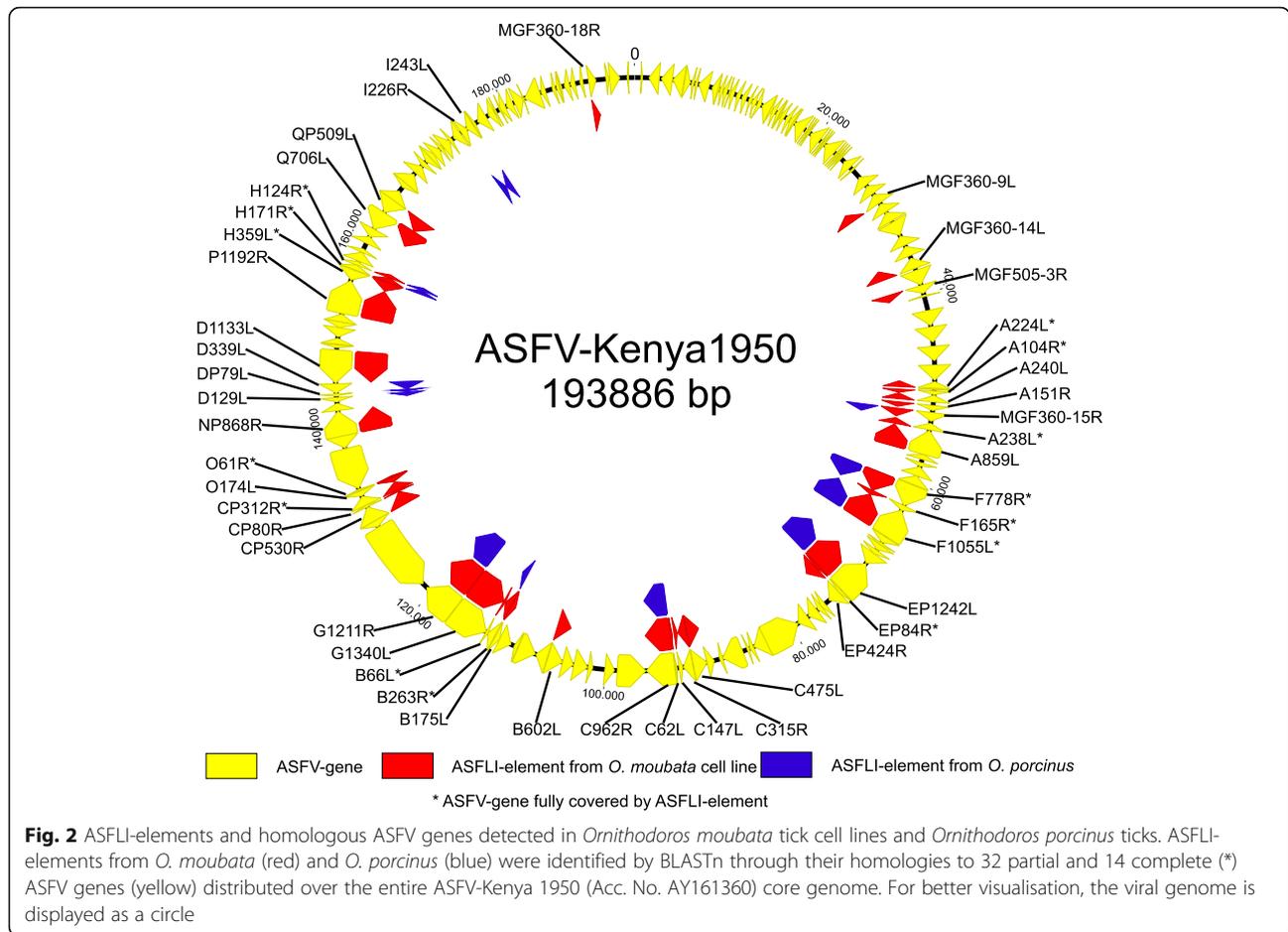


Phylogenetic analysis shows ASFLI-elements are close relatives of ASFV sequences

In order to place the ASFLI-elements in the context of other nucleocytoplasmic large DNA viruses (NCLDV), the ASFLI-elements corresponding to ASFV proteins pF1055L (Helicase) and pEP1242L (RNA polymerase subunit 2) were chosen because (a) they were present in the samples and quite complete, (b) are quite long and (c) these genes are present in the Megavirales background. Both were aligned with homologous NCLDV amino acid sequences including sequences of ASFV (Additional file 5: Supplementary Appendix). The amino acid sequence trees of the NCLDV are diverse, but both ASFLI-elements, pF1055L (Fig. 3) and pEP1242L (Additional file 5: Supplementary Appendix), group with the ASFV amino acid sequences, thus forming well-supported monophyletic clades (bootstrap value = 1). The ASFLI-elements are, however, placed as a close outgroup to the other ASFV sequences (Fig. 3 and Additional file 5: Supplementary Appendix).

ASFLI-elements are present in recently sampled *O. moubata*, *O. porcinus* and approx. 100-year-old *O. moubata* and *O. porcinus* field-collected ticks from Africa

To evaluate if the integration of the ASFLI-elements is specific for the *O. moubata* cell line OME/CTVM21, we tested all six available *Ornithodoros* cell lines as well as laboratory-reared *O. moubata* ticks from various origins using NGS and quantitative PCR (qPCR) for ASFLI-elements. Furthermore, we analysed *O. porcinus* ticks collected from the field in Kenya in 2017 and from a French laboratory-reared colony, *Ornithodoros erraticus* field ticks from Portugal and *Ornithodoros savignyi* from Nigeria. While all tick cell lines and the *O. moubata* ticks tested positive for ASFLI-elements, *O. porcinus*, *O. savignyi* and *O. erraticus* (Additional file 6: Table S5) tested negative. However, by sequencing DNA from the *O. porcinus* ticks (Additional file 1: Table S1) and assembly, we were able to identify 18 contigs containing ASFLI-elements with lengths over 150 bp which were, due to the differences in the nucleotide sequence



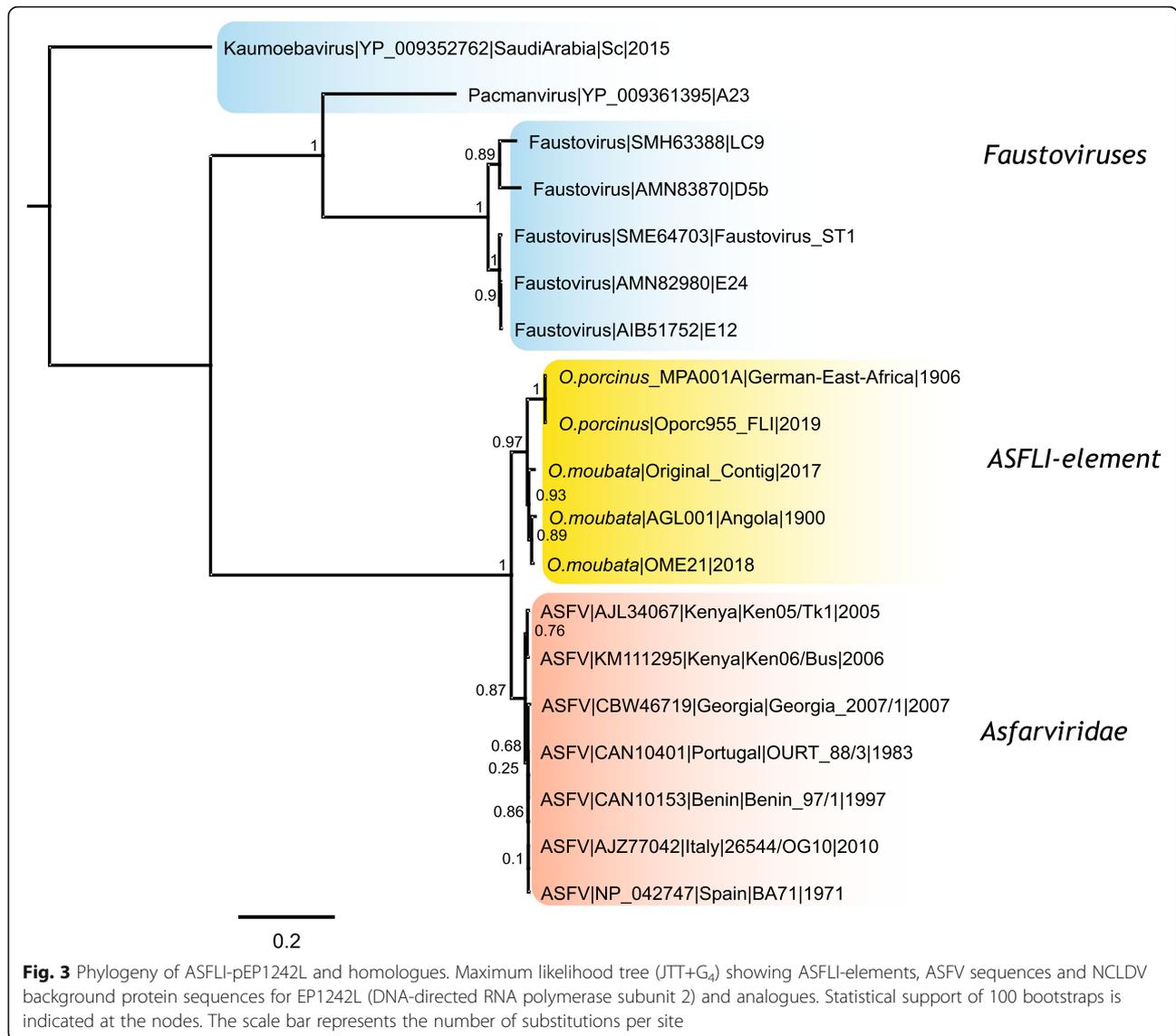
identity to the *O. moubata* ASFLI-elements, not detected by qPCR. In total, these *O. porcinus* ASFLI-elements relate to 14 ASFV genes (Fig. 2 and Additional files 1 and 7: Tables S1 and S6) and for five, corresponding ASFLI-elements were only detected in *O. porcinus* (Fig. 2 and Additional files 1 and 7: Tables S1 and S6). One possible orthologous element, including the ASFV-like EP1242L gene, was discovered in *O. moubata* and *O. porcinus* (98.5% nucleotide sequence identity flanking a piggyBac transposable element with 90.3% nucleotide sequence identity) (Additional file 8: Figure S1).

Through the analysis of published transcriptome data from *Ornithodoros* ticks, we identified ASFLI-element-specific RNA in *O. moubata* ticks from laboratory colonies from Spain [18] and Japan [19] (Fig. 4 and Additional files 1 and 2: Tables S1 and S2), while in pertinent data from *O. erraticus* [21] and *Ornithodoros turicata* (PRJNA447876), no ASFLI-specific RNA was detected (Fig. 4). Furthermore, we examined nine *Ornithodoros* specimens collected between 1906 and 1913 in the former German colonies of East and Southern Africa that had been stored in ethanol at the Berlin Museum for Natural History (MfN) by sequencing (Fig. 4 and Additional files 1

and 9: Tables S1 and S7). In all libraries relating to *O. moubata* and *O. porcinus* ticks, we detected ASFLI-element-specific reads or even contigs while in one library relating to *Ornithodoros pavementosus*, no ASFLI-element-specific DNA could be detected (Fig. 4, Additional file 9: Table S7).

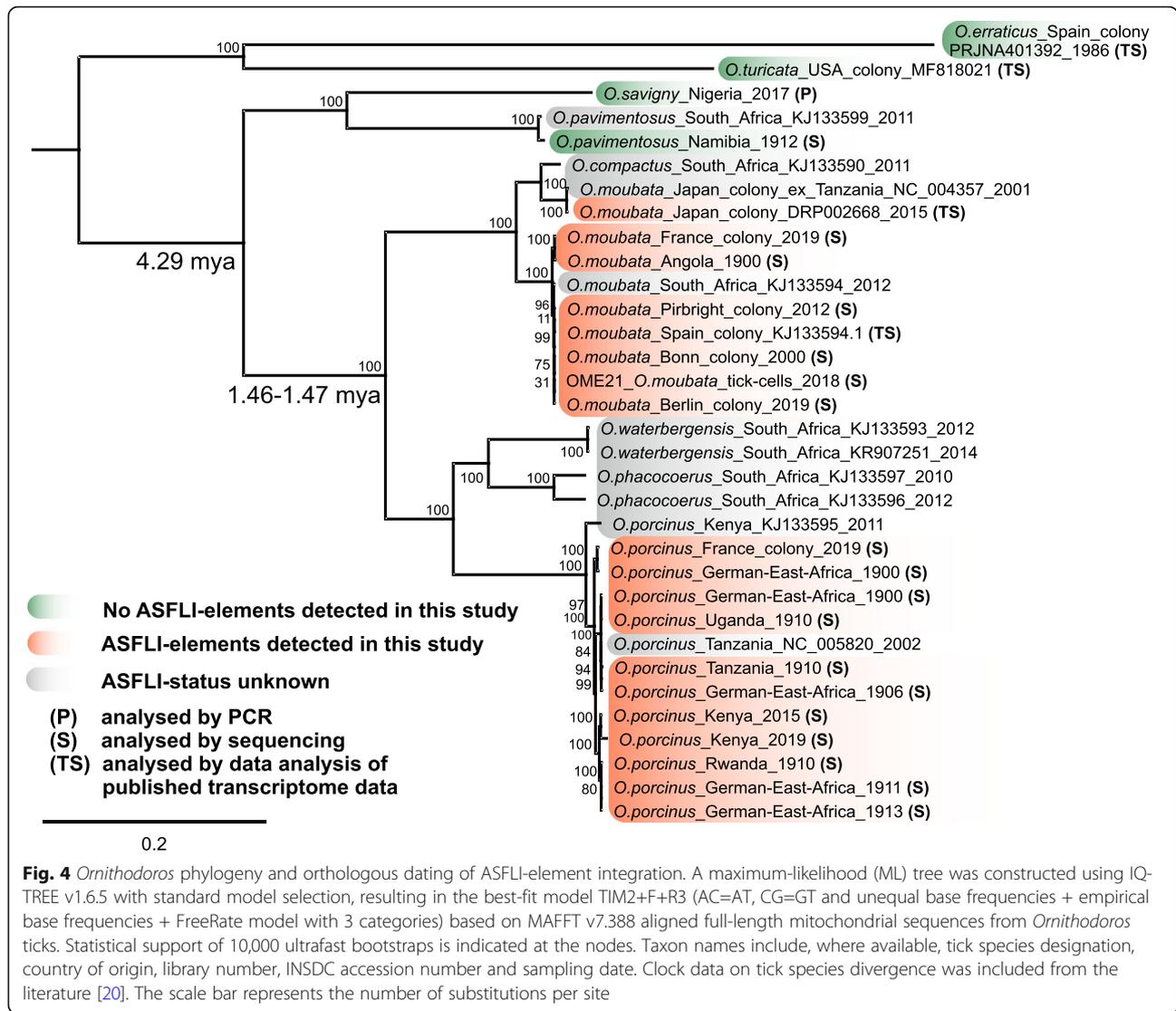
Phylogenetic reconstruction using full-length mitochondrial genomes of soft ticks reveals a possible integration of an ASFLI-element might have occurred over 1.46–1.47 million years ago (mya)

To assess the phylogenetic relationship between the different tick species carrying ASFLI-elements and perform orthologous dating, we generated full-length mitochondrial genomes from the tick sequence data generated in this study (Additional file 1: Table S1). These genomes were aligned together with additional tick mitochondrial genome sequences from the literature [20] using MAFF T v7.388 in Geneious, and a phylogenetic tree (Fig. 4) was constructed using IQ-TREE v1.6.5. Molecular clock data from the literature on tick species divergence, which was calibrated using fossil data [20], suggests that the *O. moubata* group including *O. moubata*, *O.*



porcinus, *O. compactus*, *O. waterbergensis* and *O. phaco-coeris* originated 4.29 mya and that *O. moubata* and *O. porcinus* probably diverged 1.46–1.47 mya [20]. We detected a major ASFLI-element flanking a mobile genetic element showing the same orientation and high nucleotide sequence identities in both *O. moubata* and *O. porcinus* (Additional file 8: Figure S1). Since no more comparable ASFLI-element containing loci could be defined, we analysed if multiple copies of this mobile genetic element could exist per genome, e.g. if this combination of ASFLI-element and mobile genetic element might have happened independently by chance. Therefore, we mapped DNA sequencing data from the libraries lib02151-2 and lib02339 against Contig 22316_OME and lib03101-2 against Contig 955_OPF using Bowtie2 with default parameters in the ‘very sensitive’

mode. In the case that multiple copies of the transposable element exist, a strong increase in coverage would be expected when compared to the coverage of the ASFLI-element. As displayed (Additional file 5: Figure S5), no significant difference could be detected in *O. moubata* and in *O. porcinus*. However, a BLASTn search of the transposable element sequence against the SPAdes assembled contigs of *O. moubata* and *O. porcinus* using only hits with an *e* value lower than 0.0001 revealed no additional contigs including this transposable elements nucleotide sequence. Therefore, we suggest that this ASFLI-element might have been integrated in a shared integration event that could have occurred before both species diverged 1.46–1.47 mya (Fig. 4). Because we did not discover any ASFLI-element-specific DNA in *O. pavementosus* (Fig. 4), which diverged from *O. moubata*



and *O. porcinus* around 4.29 mya, it could further be hypothesised that the maximum time to the integration is 4.29 mya.

Molecular clock analyses using ASFLI-elements from different *Ornithodoros* species provide an estimate for a time to the most recent common ancestor consistent with orthologous dating

In addition to the mitochondrial genomes, we analysed the orthologous ASFLI-element containing ASFV-EP1242L from five soft tick specimens (three *O. moubata* and two *O. porcinus*). These sequences contained deletions and one insertion in comparison with the homologous ASFV-EP1242L. Maximum likelihood phylogenies from the corresponding amino acid sequences revealed that all samples clustered as an outgroup close to the *Asfarviridae* (Additional file 5: Supplementary

Appendix). Due to missing data points over time that would be useful for calibration, inferring time-scaled phylogenies on the EP1242L containing ASFLI-element sequences from the soft tick samples was more challenging than for the mitochondrial genomes. However, by using a strict clock with molecular clock rate priors of 5e-7 to 1e-8 substitutions per site per year (commensurate with the expected substitution rate for host DNA), a time to most recent common ancestor (TMRC A) of 0.08 to 4.7 million years was inferred (Additional file 5: Supplementary Appendix). Although the confidence intervals on these TMRCAs are very wide, the TMRCAs of the EP1242L ASFLI-element is consistent with the overall integration estimate of 1.46–4.29 million years obtained from orthologous dating using the host mitochondrial genomes (Additional file 5: Supplementary Appendix).

Ornithodoros tick species and tick cell lines show differences in the infectability with various ASFV genotype isolates

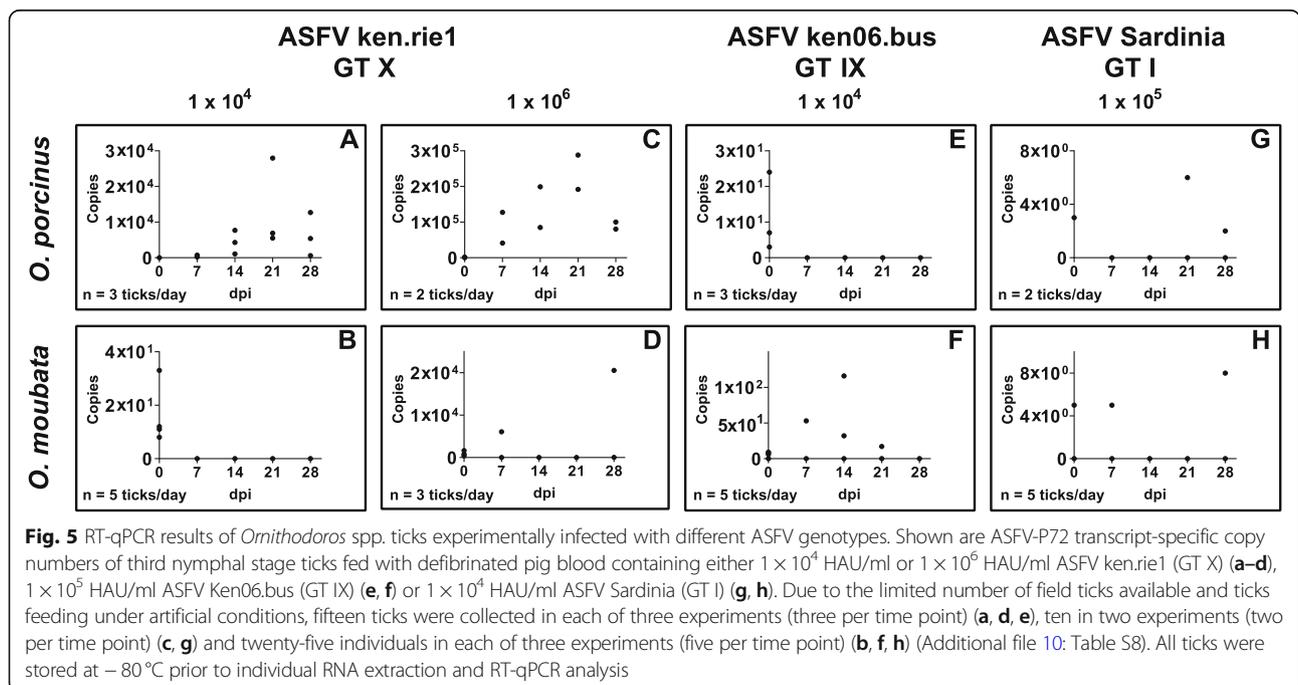
Since we detected different repertoires of ASFLI-elements in *O. porcinus* and *O. moubata* (Fig. 2), we investigated the differences in the infectability of both tick species with different ASFV strains. Therefore, we infected *O. porcinus* ticks from Kenya and *O. moubata* ticks from the Berlin colony by membrane-feeding with 1×10^4 and 1×10^6 haemadsorbing units (HAU)/ml of the ASFV ken.rie1 P72 genotype X strain that was originally isolated from a Kenyan *O. porcinus* tick. To analyse the influence of different ASFV genotypes on the infection rate of the ticks, *O. porcinus* and *O. moubata* ticks were furthermore infected with 1×10^5 HAU/ml of an ASFV-Ken06.bus genotype IX strain or with 1×10^4 HAU/ml of an ASFV-Sardinia genotype I isolate. As a marker for a successful early phase and transition into the late phase of infection, all tick specimens were tested for the accumulation of the late-expressed *p72*-specific viral transcripts by (RT-qPCR), and the virus titre of the blood-virus solution used for oral infection was determined by titration on porcine macrophages.

In the initial experiments (Fig. 5a, c), all *O. porcinus* ticks showed a clear accumulation of viral RNA while from the *O. moubata* ticks, only two accumulated viral transcripts after infection with the higher virus dose (Fig. 5b, d and Additional file 10: Table S8). While among the ticks infected with ASFV Ken06.bus, a minimal *p72* transcript accumulation was detected only in four *O. moubata* ticks (Fig. 5e, f), two *O. moubata* and

one *O. porcinus* tick infected with ASFV Sardinia showed a minimal increase of *p72* transcripts (Fig. 5g, h). Furthermore, infection of the tick cell lines OME/CTVM21, OME/CTVM22, OME/CTVM24 and OME/CTVM27 with ASFV ken.rie1 did not result in any detectable viral transcription (data not shown).

RNA sequencing demonstrates ASFLI-element-specific mRNA—small-interfering and piwi-interacting RNAs in tick cells

Interference with small RNAs (RNAi), e.g. small-interfering RNA (siRNA) or piwi-interacting RNA (piRNA), could be responsible for the observed differences in ASFV infection rates of the *Ornithodoros* ticks (as shown in insects [22]); we therefore investigated hypothetical ASFLI-element-specific small RNAs in the ticks and performed small RNA sequencing (Additional files 1, 11, 12, 13, 14 and 15: Figure S2-S4 and Tables S1 and S9). The distribution of the read lengths showed a distinct bimodal distribution in all libraries with one peak at 22 nt, representing siRNAs, and one at 28–29 nt, representing piRNAs (as shown by the typical U-bias at the 5'-end) (Additional files 11 and 15: Figure S2 and Table S9) [23]. This is in accordance with small RNA sequencing data from hard ticks [24, 25]. By mapping the resulting data (siRNA and piRNA fractions) against the ASFLI-element-containing contigs, we identified siRNAs and piRNAs homologous to different ASFLI-elements (Additional files 12 and 13: Figure S3 and S4). Sequence analysis of these piRNAs (mapped to ASFLI-elements from both strands) also revealed a clear

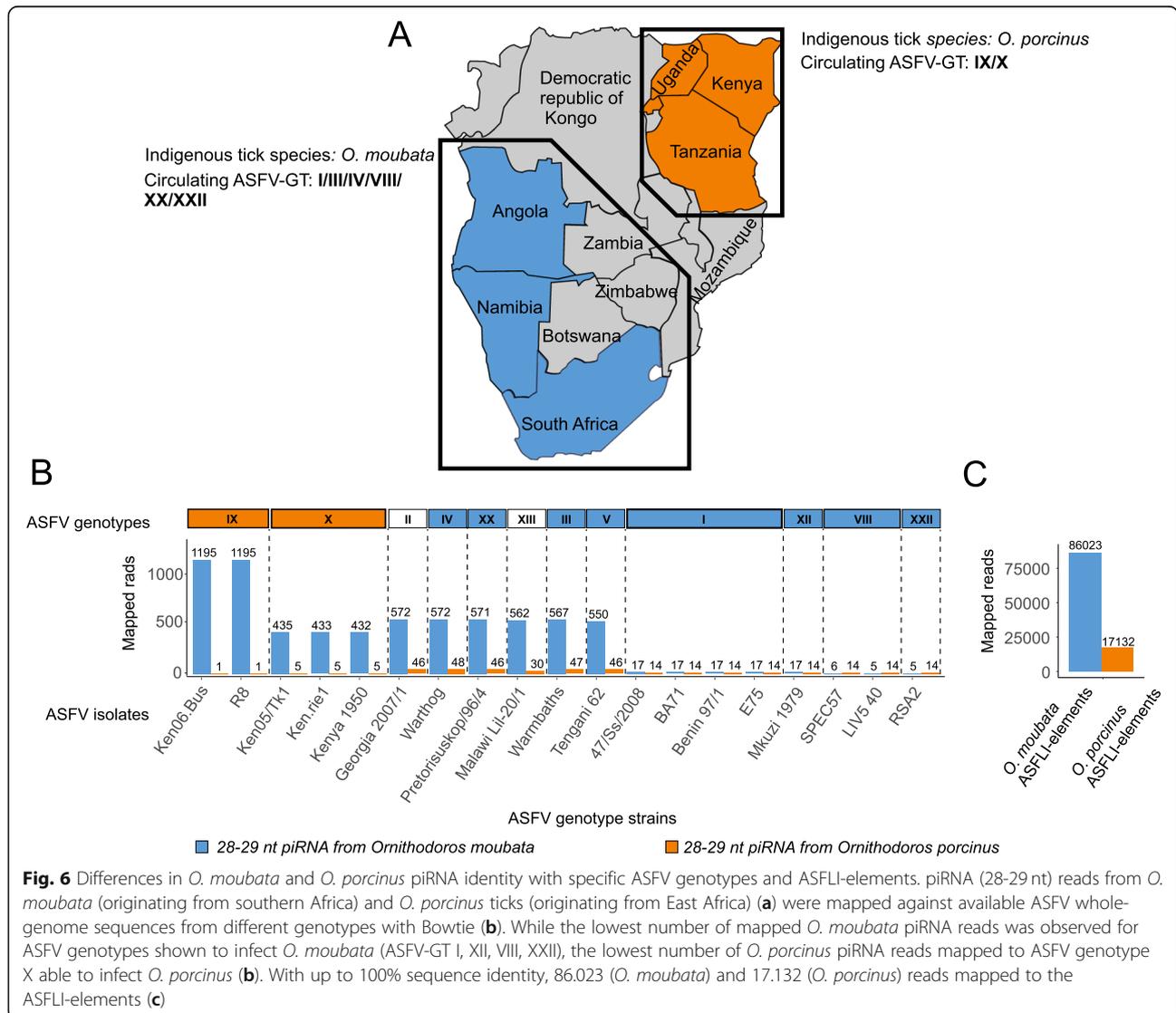


5'-U bias (Additional file 11: Figure S2C). While for some elements, none or only single homologous siRNA/piRNA molecules were identified, elements related to the ASFV genes EP1242L and F778R showed multiple corresponding siRNAs/piRNAs. Furthermore, sequence analysis of piRNA mapped to ASFLI-elements from *O. moubata* and *O. porcinus* revealed piRNAs with ping-pong signatures indicative of ping-pong cycle amplification (Additional files 16 and 17: Table S10 and S11) [26]. Following the mapping of all siRNAs and piRNAs against available ASFV whole-genome sequences from different viral genotypes, we observed matching RNAs but identified differences in the number of mapped reads from the different tick species. While the highest number of siRNA and piRNA from *O. moubata* mapped to ASFV genotypes IX and X, piRNA from *O. porcinus* mapped best to ASFV genotypes II, III, IV and V (Fig. 6

and Additional files 12, 13, 14 and 15: Figure S3-S5 and Table S9), the siRNA fraction of *O. porcinus*, however, mapped best to genotype IX and X isolates. The mapped piRNA reads also showed a clear 5'-U-bias (Additional file 11: Figure S2C). In addition to the small RNA, we detected ASFLI-element-specific mRNA in the OME/CTVM21 cells as well as in *O. moubata* transcriptome data from the literature (Additional files 1, 2: Tables S1-S2).

The reconstructed ASFLI-A104R protein is highly similar to its ASFV homologue but is not expressed in tick cell lines

Among the complete ASFLI-element genes discovered by NGS, ASFLI-A104R is one of the most conserved with nucleotide sequence identities of 83.0–85.0% and amino acid sequence identities of 94.0–95.0%. However,



the insertion of an 'A' nucleotide in a homopolymer region of eight 'A' nucleotides, starting at position 9 of the gene, leads to a frameshift and stop codon integration. Since no small RNA molecules matching ASFLI-A104R were identified, we investigated whether the ASFLI-A104R protein is expressed in the cultured tick cells. We repaired the *alA104R*-ORF in silico by removing the additional 'A', expressed the synthetic gene as both a native protein and a flag-tagged variant in *Escherichia coli* and raised rabbit sera against the purified *alA104R* proteins. Using Western blot analysis, we observed specific signals for the native protein as well as for the flag-tagged *alA104R* protein, and also for the ASFV-Kenya 1033 protein. However, no ASFLI-A104R-specific signals were obtained with protein extracts from the different tick cell lines (Additional file 18: Figure S6).

Discussion

We here report the existence of endogenous ASFLI-elements in the genome of *O. moubata* complex soft ticks. The detection of these ancient integrated sequences will allow us, for the first time, to gain deeper insights into the evolution of ASFV, one of the most dangerous animal pathogens known to date, and into the role of endogenous viral elements in the co-evolution of ticks and viruses.

Generally, very little is known about tick genome organisation or how and how often viral elements become integrated into a tick genome [27] and for what reasons. We identified ORFs related to mobile genetic elements as well as repeat regions and possible target site duplications formed by the integration of mobile genetic elements adjacent to the ASFLI-elements. Therefore, it might be hypothesised that the integration of viral elements into large hyper-variable regions of the tick genome [17, 28–30] is assisted by the interaction with mobile genetic elements as discussed for endogenous Borna-like N (EBLN) elements [31] and various arthropod viruses [28, 30, 32, 33].

Further analyses showed that *O. moubata* and *O. porcinus* ticks, in addition to the *O. moubata* tick cell lines, harbour ASFLI-elements. This observation suggests the fixation of these elements in the *O. moubata* and *O. porcinus* genomes. While it remains unclear how the fixation might have occurred, the initial event leading to germline integration might be explained by the fact that ASFV infects the tick's reproductive tissues in order to become transovarially and venereally transmitted. Furthermore, although ASFV replication takes place in the cytoplasm, viral mRNAs were demonstrated to enter the nucleus of infected Vero cells and primary porcine macrophages [34, 35], where they could have been integrated. However, more individual *O. moubata* and *O. porcinus* ticks from different geographical locations need

to be analysed to proof the fixation of the ASFLI-elements in the *O. moubata* and *O. porcinus* genomes.

Although we detected matching ASFLI-elements in different *Ornithodoros* tick species, the occurrence of multiple integration events over time cannot be ruled out. However, the identification of a highly identical ASFLI-element including the ASFV-like EP1242L gene flanking a highly identical piggyback transposable element sequence in the same orientation in both *O. moubata* and *O. porcinus* might suggest that one integration event might have taken place over 1.46–1.47 million years ago. Since we were not able to detect ASFLI-elements in *Ornithodoros* species other than *O. moubata* and *O. porcinus*, we can further estimate a maximum time to the integration of 4.29 million years ago, when these species diverged. However, a more detailed analysis of *Ornithodoros* ticks in and outside of the *O. moubata* group including DNA sequencing are needed to further substantiate this assumption. Although the molecular clock analysis using the orthologous ASFLI-elements resulted in a TMRCA consistent in range with the orthologous dating, missing data through time hindered calibration, and thus, model selection was difficult and the confidence intervals are very wide. Therefore, and since the pattern and rate of substitutions of the ASFLI-elements could have been furthermore influenced by selection, the results should be interpreted with caution and more data on ASFLI-elements in different *Ornithodoros* tick species is needed to define the possible time point of integration more precisely. Whether the integration of non-retroviral sequences into a host genome is a directed process and how often it leads to improved fitness is controversially discussed [30, 36, 37]. However, protection against homologous viruses was demonstrated after integration of retroviral elements in mice and sheep [37] and is also discussed for EBLN elements [28, 31]. Furthermore, the transcriptional activity of EVEs and protection against viral infection was demonstrated to occur in crustaceans [30, 32, 33] and mosquitoes [38, 39].

In total, we identified ASFLI-elements in the genomes of *O. moubata* and *O. porcinus* ticks homologous to 46 ASFV genes; most of them represent genes with properties linked to viral DNA replication or RNA transcription. These observations may be explained by the fact that these genes in particular are highly conserved between viral isolates (90–100%); the genes coding for structural proteins are more variable; therefore, we cannot exclude that the latter, if also integrated, are simply not detectable by identity-based search methods like BLAST. However, the integration of major viral structure protein genes might have been harmful for the tick cell, so that they were just not maintained in the genome. Furthermore, it might be possible that that the

integration of the conserved viral core genes, essential for the early phases of viral infection, is in particular linked to a beneficial function, such as protection by RNAi.

During infection experiments using *O. moubata*, *O. porcinus* and different ASFV genotypes, we observed differences in the ability of ASFV to infect the various *Ornithodoros* species. While a genotype X isolate from Kenya was able to replicate in *O. porcinus* ticks from Kenya, no infection was observed in *O. moubata* (phylogenetically closest to southern African ticks). However, when using a higher viral titre, successful infection could be observed in a small number of *O. moubata* ticks indicating that, in principle, infection is possible and that low-level infection might be blocked by a tick defence mechanism, such as RNAi.

Through the identification of ASFLI-element-specific siRNA and piRNA in the ticks and tick cells, we provide data suggesting that the ASFLI-elements could be involved in RNAi-based defence against ASFV infection in *Ornithodoros* ticks. Both siRNA and piRNA are known as major regulation factors of transcription, transposon suppression and defence against viral pathogens [40–43]. Although for ticks the known arsenal of antiviral mechanisms includes protection by RNAi [44], the role of EVEs in RNAi and protection against viral pathogens has never been reported.

The results of our infection experiments with different ASFV genotypes using the *O. porcinus* and *O. moubata* ticks are in agreement with observations from other studies in which clear differences were shown in tick infection rates using different ASFV genotype isolates [45–49]. Together with our data on ASFLI-element-specific siRNA and piRNA in these ticks, it could be hypothesised that the ASFLI-elements might play a role in determining which ASFV can infect which tick species. Therefore, it could be speculated that during millions of years of co-evolution, where ASFV hosts were regionally isolated by geographical factors (e.g. mountains, rivers), ASFV and its soft tick hosts have adapted to co-exist as displayed by the regional distribution of ASFV genotypes and the lack of any pathological findings in infected soft ticks, despite generalised virus infection and high viral titres. However, additional immunological or genetic factors determining which ASFV strain can infect a specific *Ornithodoros* tick species need to be considered, and more infection experiments using different *Ornithodoros* tick species and ASFV genotypes and data on small RNA are needed to substantiate this hypothesis.

The possible role of EVEs in the antiviral response of ticks might also explain why some EVEs have been protected from accumulating mutations. However, it remains unclear why no siRNA or piRNA could be detected for the ASFLI-A104R protein, which is one of

the most abundant ASFV proteins in infected wild boar lung cells and infected Vero cells [50], highly conserved between ASFV isolates [51] and whose viral counterpart is crucial for virus replication [51]. While our data might lead to the conclusion that the ASFLI-A104R protein is not translated, the amount of protein might just have been too low for detection or cellular machinery might degrade the protein. Due to the high identity to modern ASFV-A104R sequences, a more recent integration event has also to be considered.

However, further analyses including mass spectrometry analyses, the cloning of the ASFLI-A104R element into a recent ASFV isolate and the biochemical and functional characterisation of the protein should be performed to elucidate a possible function and the nature of ASFLI-A104R.

Conclusion

In conclusion, we present evidence for endogenous viral elements of the only known DNA arbovirus, ASFV, which might have been integrated into soft tick genomes over 1.46–1.47 million years ago. They serve as templates for siRNA and piRNA thereby possibly protecting the tick against viral infection. We believe that this discovery is an exceptional starting point to study the evolution of ASFV, one of the most dangerous threats to sustainable pig production in Africa, Europe and Asia, as well as the mechanisms by which organisms adapt to the ancient struggle with viral pathogens.

Methods

Virus strains

All work with active ASFV was carried out in the FLIs high-security facilities under BSL3+ conditions.

For infection studies, we used an ASFV isolate designated ‘ASFV ken.rie1’. The virus was isolated from *O. porcinus* ticks from Kenya on swine peripheral blood mononuclear cells (PBMCs), passaged twice and titrated on PBMCs as described elsewhere [52]. By sequencing the *p72*, *B602L* and *p54* genes, we compared ‘ASFV ken.rie1’ to known ASFV isolates from GenBank and grouped it into *p72* genotype X, which is known to circulate in East Africa between ticks and indigenous pigs [53]. ‘ASFV-Sardinia’ and ‘ASFV ken06.bus’ (originally received from the European reference laboratory for ASFV (CISA-INIA)) were stored at the German National Reference Laboratory for ASFV as reference material. Both had initially been passaged and titrated on PBMCs.

Tick rearing, tick infection and tick cell cultures

O. moubata ticks, sourced from laboratory colonies maintained at HU-Berlin, Germany, and CIRAD-France as well as *O. porcinus* ticks collected from the field in Kenya and *O. erraticus* ticks collected from the field in

Portugal were held at 25 °C and 85% relative humidity in the dark and fed approximately every 2 months on defibrinated pig blood without any additives (heated to 37 °C) through a parafilm membrane without any stimulants using a polystyrene feeding chamber mounted on a heating plate as described elsewhere [54]. Laboratory-reared *O. moubata* from The Spanish National Research Council (CSIC) Salamanca in Spain and The Pirbright Institute in the UK were received and stored in 70% ethanol. DNA from Nigerian *O. savignyi* was provided by the University of East London. Tick infection was carried out using third nymphal stage ticks from an ASFV-negative laboratory-reared colony (*O. moubata*) or field ticks (*O. porcinus*). To ensure the absence of ASFV in the field ticks, adult females were fed as described above and all produced eggs were separated. After the death of the adult females, they were tested for ASFV P72-specific DNA by qPCR as described below, and in case of negative results, all offspring were designated ASFV-free, maintained on ASFV free defibrinated pig blood (as described above) and used for infection experiments. For infection, the same feeding system than for colony maintenance [54] was used with ASFV-spiked blood. Ticks of the same species were fed together until repletion and spontaneous detachment. The ticks were manually collected using forceps and held as described above for 28 days. Random ticks were collected on days 0, 7, 14, 21 and 28 post-infection (dpi) and stored at -80 °C prior to RNA extraction and RT-qPCR analysis as described below. The *O. moubata* cell lines OME/CTVM21, OME/CTVM22, OME/CTVM24, OME/CTVM25, OME/CTVM26 and OME/CTVM27 were maintained in sealed flat-sided culture tubes (Nunc, Thermo Scientific, Waltham, MA) at 28 °C in either L15, L15/H-Lac or L15/MEM medium as described elsewhere [15]. OME/CTVM25 and OME/CTVM26 grow extremely slowly and were therefore only maintained briefly for NGS and qPCR analysis. OME/CTVM21, OME/CTVM22, OME/CTVM24 and OME/CTVM27 were infected with ASFV by adding 0.22×10^5 HAU directly to 2.2-ml cell culture without washing. On days 1, 7, 14, 21 and 28 post-infection, 125 µl cell culture was removed and added to 125 µl PBS and 750 µl TRIzol LS (Thermo Scientific) for RNA extraction. An equal volume of fresh culture medium was added to the culture tubes. All samples were stored at 4 °C until RNA extraction. The medium was changed weekly as described elsewhere [15].

Nucleic acid extraction

Ticks were homogenised in a 2-ml reaction tube with two 5-mm steel beads and 500 µl PBS using a Tissue Lyser II (Qiagen, Hilden, Germany) for 3 min at 30 Hz. The homogenate was centrifuged at 10,000×g for 5 min, and DNA was extracted from the supernatant using the

High Pure Template Prep Kit (Roche) according to the manufacturer's instructions. Tick cell line DNA was extracted using the same kit according to the manufacturer's instructions.

High-molecular weight DNA was isolated from tick cells using a salting-out protocol [55], slightly modified by adding RNase A (10 mg/ml) for incubation at 37 °C for 60 min in the first step and proteinase K (10 mg/ml) for incubation at 56 °C for another 60 min in the second step.

For RNA extraction, ticks were homogenised in a 2-ml reaction tube with two 5-mm steel beads and 1 ml TRIzol Reagent (Thermo Scientific) using a Tissue Lyser II (Qiagen) for 3 min at 30 Hz. For efficient cell lysis, the homogenate was incubated for 10 min at room temperature. After addition of 200 µl trichloromethane (Sigma Aldrich, St. Louis, MO), samples were thoroughly mixed, incubated for 10 min at room temperature and centrifuged at 10,000×g for 10 min at 4 °C. The clear aqueous phase was removed, mixed with 600 µl of 100% ethanol and loaded onto an RNeasy-spin column (Qiagen) for RNA extraction and additional on-column DNA digestion following the manufacturer's instructions for total RNA extraction. For small RNA extraction, single N3-nymphs were homogenised in 1 ml TRIzol reagent as described above. Small RNA was extracted from the resulting clear aqueous phase using the miRNeasy Mini kit (Qiagen) in combination with the RNeasy MinElute® Cleanup Kit (Qiagen) according to the manufacturer's protocol.

RNA extraction from tick cell lines was performed using TRIzol LS and RNeasy-spin columns (Qiagen) according to the manufacturer's protocol. Prior to qRT-PCR analysis, all RNA samples were treated with TurboDNase (Life Technologies) for removal of residual DNA.

DNA extraction from museum-stored ticks was done in a DNA cleanroom under ancient-DNA (aDNA) conditions as described elsewhere [56, 57].

Oligonucleotide design

All primers and probes were designed using the Genious (v.10.0.9) software suite and are listed in Additional file 19: Table S12 with additional primer sequences obtained from the literature.

PCR

For PCR reactions, Phusion Green Hot Start II High Fidelity PCR Master Mix (Thermo Scientific) was used according to the manufacturer's instructions on a C1000 Thermo Cycler (Bio-Rad, Hercules, CA). PCR products were visualised, depending on their length, on 1% or 1.5% agarose gels and visualised by ethidium bromide staining under UV light.

qPCR

All qPCR reactions were done using the QuantiTect Multiplex PCR NoROX Kit (Qiagen) in a 12.5- μ l reaction including 2.5 μ l extracted DNA, on a C1000 Thermo Cycler (Bio-Rad) in combination with a CFX96™ Real-Time PCR Detection System (Bio-Rad) according to the manufacturer's instructions. Primers and probes were premixed (10 pmol primer/ μ l + 1.25 pmol probe/ μ l), and 1 μ l of the mixture was used per reaction. For ASFV DNA amplification, an OIE-listed qPCR assay was used targeting the viral B646L gene showing a detection limit of 5–10 copies as described elsewhere [58].

Control qPCR systems targeted *O. moubata* mitochondrial (16S rDNA) modified from [59], *O. moubata* chromosomal (enolase) and *O. erraticus* chromosomal (subolesin) housekeeping genes (designed in this study).

RT-qPCR

All RT-qPCR reactions were performed using the QuantiTect Probe RT-PCR Kit (Qiagen) on a C1000 Thermo Cycler (Bio-Rad) in combination with a CFX96™ Real-Time PCR Detection System (Bio-Rad) according to the manufacturer's instructions. RNA concentration was measured using the NanoDrop™ 2000 (Thermo Fisher Scientific), and 100 ng RNA in 4 μ l (25 ng/ μ l) was used in a 12.5- μ l reaction according to the manufacturer's protocol. For P72 transcript quantification, the OIE listed qPCR assay [58] was used with the following cycling conditions: 50 °C for 30 min, 95 °C for 15 min, 45 cycles of 95 °C for 1 min, 60 °C for 30 s and 72 °C for 30 s. For all samples tested by RT-qPCR, controls without reverse transcriptase were included to demonstrate the absence of DNA. For absolute quantification, a standard curve was generated, using samples of known copy number produced from cloned gene fragments, and measured by the Quant-iT dsDNA Assay Kit HS (Thermo Fisher Scientific).

Sanger sequencing

For Sanger sequencing, PCR products were extracted from the agarose gel with the QIAQuick Gel Extraction Kit (Qiagen). The sequencing PCR was done using the BigDye Terminator Kit V1.1 (Thermo Scientific). After purification with NucleoSEQ spin columns (Macherey Nagel, Düren, Germany), samples were sequenced on a Genetic Analyzer 3031XL (Life Technologies, Carlsbad, CA), followed by chromatogram editing with the Geneious (v.10.0.9) software.

Next-generation sequencing

For IonTorrent PGM sequencing, the libraries lib01543-lib01546, lib01610 and lib01611 were prepared from the six *O. moubata* cell lines as detailed in Additional file 1: Table S1 and sequenced as described elsewhere [60].

The libraries from museum-stored ticks (lib03076-3086) were prepared under aDNA conditions in a DNA clean-room and sequenced as described elsewhere [56, 57].

The libraries lib02151 and lib02152 were prepared from OME/CTVM21 cells for sequencing on an Illumina MiSeq platform in 300-bp paired-end mode as previously described [61] (Additional file 1: Table S1). The library lib02339 was prepared from OME/CTVM21 cells for sequencing on an Illumina HiSeq platform in 50-bp paired-end mode (Additional file 1: Table S1). Extracted DNA was fragmented to a size of 150 bp (M220 Focused-ultrasonicator™, Covaris, Woburn, MA), and libraries were prepared using NEXTflex™ Dual-Index adapters (Bioo Scientific, Austin, TX) in an automated fashion (SPRIworks Fragment Library System II, Beckman Coulter, Brea, CA). Size exclusion was performed manually aiming at a fragment length peak of 350 bp (Agencourt® AMPure® XP Beads, Beckman Coulter). The library was then amplified (AccuPrime™ Taq DNA Polymerase, high fidelity, Invitrogen, Carlsbad, CA), purified (Agencourt® AMPure® XP Beads) and quality-checked (High Sensitivity DNA Kit, Agilent Technologies). The final library was quantified (KAPA Library Quantification Kit, Roche) and sequenced (Illumina) according to the manufacturer's instructions.

For MinION sequencing, high-molecular weight tick cell DNA was fragmented using G-Tubes (Covaris), followed by library preparation (libMinION_1) using the 1D² Sequencing Kit (R9.5) (Oxford Nanopore, Oxford, UK) according to the manufacturer's protocol. Sequencing was performed on the MinION MIN-101B system (Oxford Nanopore).

For small RNA sequencing, library preparation from enriched small RNA samples was performed using the Ion Total RNA-Seq Kit v2 (Thermo Scientific). The libraries lib02965 (mRNA), lib02966, lib02967, lib03460 and lib03506 (small RNA) were prepared for, and sequenced on, an IonTorrent S5 (Thermo Fisher Scientific) as described elsewhere [62].

Amplicon sequencing for assembly validation

For the validation of the two ASFLI element containing contigs (4.3 and 6.8 kb), PCR reactions were performed as described above. PCR products were extracted from the agarose gel with the QIAQuick Gel Extraction Kit (Qiagen) and sequenced (lib02035) on an Illumina MiSeq platform in 300-bp paired-end mode as previously described [61]. The data was analysed by assembly using Newbler 3.0 and alignment with the originally assembled sequences using MAFFT v7.388 in Geneious.

Data analysis

The metagenomic classification software pipeline RIEMS [63] was used on lib01543-01546 and lib01610-01611

(Additional file 1: Table S1). For further data analysis and proof of integration of ASFV-like sequences into the tick cell genome, overlapping read pairs in the MiSeq data sets (lib02151 and lib02152) were merged to single reads using FLASH v1.2.11 [64]. Subsequently, the merged reads were assembled together with data from lib02339 and libMinION_1 using SPAdes v3.11.1 [16] in the mode of read error correction prior to assembly. In total, 66,745 contigs were assembled with standard assembly parameters and automatically determined *K*-mer length of 21, 33, 55, 77, 99 and 127. The resulting contigs were blasted (BLASTn, NCBI, v2.6.0+) against a customised database comprising all sequences with the NCBI taxonomy ID 10497 (African swine fever virus) (as of 16 January 2018). Hits were filtered using a cut-off *e* value of 1×10^{-4} and a minimum alignment length of 150 bp, resulting in 34 contigs. These were then blasted against the complete NCBI database (the non-redundant nucleotide collection) to reliably identify and annotate ASFV-like sequences and areas of the host genome using default parameters. BLASTp search of > 500 bp ORFs was performed against the 'non-redundant protein sequences' database using default parameters (Additional file 20: Figure S7).

ASFLI-element sequences from mRNA sequencing data and additional *O. moubata* transcriptome data [18] downloaded from GenBank were identified by mapping using the Newbler software 3.0 (Roche) against a database containing all known ASFV sequences and previously obtained ASFLI element sequences.

DNA sequence data from lib03101 and lib03102 (*O. porcinus* ticks) were assembled using SPAdes v3.11.1 [16] in the mode of 'read error correction prior to assembly' with standard assembly parameters and automatically determined *K*-mer length of 21, 33, 55, 77, 99 and 127. The resulting contigs were blasted against a customised database comprising all sequences with the NCBI taxonomy ID 10497 (African swine fever virus) (as of 10 May 2019). Obtained hits were filtered as described above (Additional file 20: Figure S7).

DNA sequence data from the museum ticks (lib03076-03086 and lib03376-03377) were mapped against a database comprising all available ASFLI-element sequences either from the *O. moubata* cell lines or from *O. porcinus* ticks using Bowtie (v.1.1.2) with default parameters. Mapped reads were assembled using SPAdes v3.11.1. Small RNA data were used with and without deduplication using dedupe from the BBTtools package [65]. Reads were mapped against ASFV genomes or ASFLI-element-containing databases using Bowtie (1.1.2) in Geneious with 22-nt seed length for siRNA and 28-nt seed length for piRNA and one allowed mismatch. The piRNA 'ping-pong signature' was analysed using PingPongPro v1.0 [66] in default mode with the -b option (creates

additional browser track files, which are suitable for display in common genome browsers), and results were analysed in Geneious.

For the generation of full-length tick mitochondrial genomes, sequence data was mapped against a database containing all available *Ornithodoros* full-length mitochondrial sequences (as of 20 June 2019) using Newbler 3.0 and Bowtie 2 (2.3.0), and mapped reads were assembled using Newbler 3.0 and SPAdes v.3.11.1 with default parameters and automatically chosen *K*-mer length of 21, 33 and 55.

Phylogenetic analysis

For the identification of homologous genes from other viral genera, BLASTp (protein-protein BLAST) was used with amino acid sequences from the translated ASFV Ken06/Bus F1055L (helicase) and EP1242L (RNA polymerase subunit 2) genes using the non-redundant protein sequences database. All viral sequences identified by BLASTp, covering at least 60% of the query, were used for further analysis. This yielded sequences from the NCLDV, including Faustoviruses, Pacmanvirus, Kaumoebavirus and Marseillevirus as well as African swine fever virus. These sequences together with the corresponding translation of the ASFLI-elements were aligned using MUSCLE in MEGA [67, 68]. Nucleotide sequences of the EP1242L gene and equivalents from ASFV, Faustovirus, Pacmanvirus and Kaumoebavirus were also downloaded from GenBank. These nucleotide sequences together with the ASFLI-element were again aligned using MUSCLE in MEGA. A maximum likelihood tree was calculated from an alignment of NCLDV, ASFV and ASFV-like sequences in MEGA. The Jones-Taylor-Thornton (JTT) model allowing gamma-distributed rates between sites (four categories), considering sites with > 95% data, and 100 bootstraps was used [69]. More details about the methods used for the phylogenetic analyses are presented in Additional file 5: Supplementary Appendix.

Tick mitochondrial genomes were aligned using MAFFT v7.388 in Geneious. The phylogenetic tree was constructed using IQ-TREE v1.6.5 with standard model selection, resulting in the best-fit model TIM2+F+R3 (AC=AT, CG=GT and unequal base frequencies + empirical base frequencies + FreeRate model with 3 categories).

Clock rate estimates and Bayesian time-scaled trees

To obtain an initial clock rate estimate, the ASFV nucleotide sequences were analysed using BEAST 1.8.4 [70, 71]. Several different clock models were employed, and results were evaluated with the marginal likelihood estimation using path sampling and the stepping-stone sampling feature within BEAST [72]. The clock models

were: 'strict', 'uncorrelated relaxed log-normal', 'uncorrelated relaxed gamma', 'uncorrelated relaxed exponential', 'random local clock' and 'fixed local clock' [73–75]. Apart from the clock models, the substitution model TN93 with gamma-distributed site-to-site rate variation over four categories and constant population size tree prior model (effective population size) were used. Each BEAST run used a Markov chain Monte Carlo (MCMC) chain with a length of 10,000,000 steps, sampling every 1000 steps and discarding the first 10% as burn-in.

All estimations of the time to the most recent common ancestor (TMRCA) are given in millions of years ago (mya). More details about the methods used for the molecular clock analyses are presented in Additional file 5: Supplementary Appendix.

Protein expression and purification in *E. coli* and rabbit immunisation

The reconstructed ASFLI-A104R gene was custom-synthesised and inserted into vector pMA-T (Genieart, Invitrogen). It was then amplified by PCR, and the products were inserted into vector pMAL-p2X (New England Biolabs, Ipswich, MA). After transformation of *E. coli* (TB1) with the plasmid and induction with isopropyl- β -D-thiogalactopyranosid (IPTG), the maltose-binding protein (MBP) fusion protein was purified by affinity solid-phase extraction on an amylose resin as suggested by the manufacturer (New England Biolabs). Rabbits were immunised by intramuscular injections of 400 μ g purified MBP–A104R fusion protein every 4 weeks over a period of 4 months. Blood was taken monthly and stored at 4 °C overnight after collection, followed by centrifugation for 10 min at 453 \times g at room temperature. The resulting sera were stored at – 20 °C.

Transfection

Approximately 10⁶ HEK 293T (ATCC® Number: CRL-11268™) cells were cultivated for 24 h in 6-well plates. Using 5 μ g of polyethylenimine (PEI) [76], the cells were transfected with 2.5 μ g plasmid pCAGGS-Zecke-A104R in which transgene expression is directed by a hybrid HCMV major immediate-early enhancer/chicken-actin promoter [77]. The plasmid/PEI mix was added to the cell culture after pre-incubation at room temperature for 20 min and centrifuged at 600 \times g for 60 min at 26 °C to enhance transfection efficacy. After 5 h of incubation at 37 °C, the medium was replaced by fresh medium, and after 48 h of further incubation, cells were harvested and lysed in 500 μ l sample buffer for SDS-PAGE.

SDS-PAGE and immunoblotting

SDS-PAGE was performed with a Mini-Protean Tetra System (Bio-Rad) using hand-cast SDS gels with 15% polyacrylamide according to Laemmli [78]. After

separation, proteins were blotted onto nitrocellulose membranes (Whatman, Maidstone, UK) using a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad) under conditions suggested by the manufacturer. Membranes were blocked by incubation in PBS supplemented with 6% milk powder (Sigma Aldrich) and 10% horse serum (Sigma Aldrich) overnight at 4 °C. After consecutive washes with 0.3% and 0.1% Tween20 in PBS, the blot was incubated with a 1:2500 dilution of the rabbit serum (collected following the third immunisation), washed again with PBS-Tween20 as described above, and finally incubated with a 1:20,000 dilution of a peroxidase-conjugated secondary goat anti-rabbit antibody (Dianova, Hamburg, Germany). Signals were visualised using the ECL kit (Clarity Western ECL, Bio-Rad, Hercules, CA) as suggested by the manufacturer.

Statistical analysis

To calculate the number of ticks needed for the infection experiment, we conducted sample size calculation using R-studio (<https://www.rstudio.com>) as described elsewhere [79] with an assumed prevalence of 0.9 in one population and 0.3 in the second population. In order to obtain a power of 80% ($\beta = 0.2$) and a confidence level of 99% ($\alpha = 0.01$), the required sample size was calculated to be ten animals per group.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12915-020-00865-6>.

Additional file 1: Table S1. Overview of NGS experiments conducted in this study and data obtained from the literature. *Ornithodoros moubata* cell line names have been abbreviated by removal of "/CTVM" and *Ornithodoros porcinus* and *Ornithodoros moubata* have been abbreviated for convenience.

Additional file 2: Table S2. ASFV-like transcripts detected in data from six *Ornithodoros moubata* cell lines and *O. moubata* ticks by mRNA-sequencing and data analysis from published sequences ASFV-like sequences from tick cell libraries lib01543-46 und lib01610-11 (total RNA) were identified using RIEMS. The resulting reads were blasted against the NCBI BLASTn database for confirmation and annotation. *Ornithodoros moubata* transcriptome data from a previous study⁴² was downloaded from GenBank, and ASFV-like sequences were identified by mapping the data against a tailored database containing all known ASFV sequences and all previously obtained ASFV-like sequences. The resulting reads and contigs were annotated after a BLASTn search against the entire BLASTn database. Sequencing data from the mRNA enriched OME21 library lib02965 was mapped against a tailored database containing all known ASFV sequences and all previously obtained ASFV-like sequences, and resulting contigs were blasted against the entire BLASTn database. Per sequence, only the hit with the lowest e-value is shown.

Additional file 3: Table S3. BLASTn results of SPAdes-assembled contigs containing ASFLI- and other endogenous viral elements detected in the OME/CTVM21 genome. Shown are hits with the lowest e-value for viral, tick and mobile genetic element-related genes, as obtained by Blastn against the entire NCBI database.

Additional file 4: Table S4. BLASTp results of all ORFs >500 bp on SPAdes-assembled contigs containing ASFLI- and other endogenous viral elements detected in the OME/CTVM21 genome. Shown are hits with the

lowest e-value for viral-, tick and mobile genetic element-related proteins, as obtained by Blastp of ORFs >500 bp against the entire NCBI database.

Additional file 5: Supplementary Appendix. Phylogenetic and molecular clock analyses of ASFLI-elements from different tick genomes and ASFV using different clock-rates and substitution models implemented in BEAST. **Supplementary Appendix. Figure A1-A5, Tables A1-A7.** FigA1- Phylogenetic tree of ASFV and ASFLI-elements. FigA2 – Phylogenetic tree of NCLDV including ASFV and ASFLI-elements. FigA3 – Time-scaled tree for partial EP1242L sequences from tick samples and reference ASFV (non-integrated). FigA4 - Time-scaled tree for partial EP1242L sequences from tick samples only (5e-7 substitutions per site per year). FigA5 - Time-scaled tree for partial EP1242L sequences from tick samples (1e-8 substitutions per site and year). Table A1 - Summary of EP1242L fragments derived from ticks and the tick cell line OME/CTVM21 (OME21) used in the analysis. Table A2 - Indels in the tick sample sequences relative to the start of EP1242L in ASFV|KM111295|Kenya|Ken06/Busj2006. Table A3 - Estimated root height and overall mean clock rate for strict clocks with fixed priors. Table A4 - Estimated root height and overall (averaged over all branches) mean clock rate for relaxed clocks with fixed priors of the rate of the relaxed clock (some variation). Table A5 - Estimated root height and overall (averaged over all branches) mean clock rate for relaxed clocks with normal or log-normal priors on the rate of the relaxed clock (most variation). Table A6 - Estimated root height and overall (averaged over all branches) mean clock rate for strict clocks with log-normal priors on the rate of the strict clock for the five ASFV-like sequences from tick samples only. Table A7 - Marginal likelihood estimation using Path Sampling and Stepping Stone Sampling, showing that the log-normal clock rate prior with mean = 5e-7 is best, but not significantly better than the other clock rates.

Additional file 6: Table S5. Results of tick and tick cell line screening for ASFLI-elements by qPCR. DNA from single ticks and every tick cell line ($n=1$) was extracted and analysed by qPCR for six ASFV-like genes as described in the 'Material and methods' section. All samples were tested with a qPCR control targeting a tick housekeeping gene to demonstrate successful DNA extraction and presence of tick DNA. Absence of ASFV in samples producing false positive results was proven by an OIE-listed qPCR.

Additional file 7: Table S6. BLASTn results of SPAdes-assembled contigs from *Ornithodoros porcinus* France and Kenya17 containing ASFLI-elements. Shown are hits with the lowest e-value for ASFV-genes as obtained by Blastn against the entire NCBI database.

Additional file 8: Figure S1. Similarity between ASFLI-element contigs identified in *Ornithodoros moubata* cells and *Ornithodoros porcinus* ticks. Contigs were assembled using SPAdes 3.13, on sequencing data obtained from *O. porcinus* ticks (OPF) and and OME/CTVM21 cells (OME) using default parameters. After contig identification by BLASTn and BLASTp search against the entire NCBI database and annotation, contigs were aligned using MAFFT v7.388 in Geneious. To analyse the possible existence of multiple copies of the piggyback transposable element, coverage was calculated in Geneious from mapping of libs 02151-2 and lib02339 against Contig 22316_OME and lib03101-2 against Contig 955_OPF using Bowtie2 with default parameters in the "very sensitive" mode.

Additional file 9: Table S7. SPAdes-assembled contigs from deep sequencing data from libraries AGL001 and MPA001, as generated from museum-stored ticks. Data from libraries from museum-stored tick were mapped against ASFLI-element-containing databases of *Ornithodoros moubata* (AGL001) and *Ornithodoros porcinus* (MPA001) using Bowtie2 (v.2.3.4.3) with default parameters. Subsequently, mapped reads were assembled using SPAdes and aligned to the corresponding ASFLI- element-contigs using MAFFT (v. 7.388) in Geneious.

Additional file 10: Table S8. qRT-PCR results of *Ornithodoros* ticks experimentally infected with different ASFV-genotype isolate. Shown are ASFV-P72 transcript-specific Cq-values of third nymphal stage ticks fed with defibrinated pig blood, containing either 1×10^4 HAU/ml or 1×10^6 HAU/ml ASFV-ken.rie1 (GT X) (A-D), 1×10^5 HAU/ml ASFV-Ken06.bus (GT IX) (E-F) or 1×10^4 HAU/ml ASFV-Sardinia (GT I) (G-H). Due to the limited number of field ticks available and feeding under artificial conditions,

fifteen *Ornithodoros porcinus* ticks were collected in each of three experiments and ten in two experiments while for the laboratory-reared *Ornithodoros moubata*, twenty-five individuals were collected in each of three experiments and fifteen in one experiment. All samples were stored at -80°C until RNA-extraction and ASFV transcript-specific qRT-PCR analysis as described in the 'Material and methods' section.

Additional file 11: Figure S2. Read length distribution and piRNA nucleotide bias of small RNA libraries. The length distribution is shown from small RNA libraries before and after removal of duplicate reads. Blue colored bars indicate siRNA (22nt) and piRNA (28-29nt) fractions (A). Reads of the piRNA fraction (28-29nt) were extracted from raw and deduplicated libraries, trimmed at the 3' end to 28 nt for analysis and nucleotide frequencies for every position were calculated using R-studio (<https://www.rstudio.com>) for the entire dataset (B) and for non-duplicated reads mapped to all ASFLI-elements and all available ASFV whole-genome sequences using Bowtie in Geneious with 28 nt seed length and one allowed mismatch (C). piRNA from *O. porcinus* did only map to the ASFV forward strand.

Additional file 12: Figure S3. siRNA/piRNA mapping against ASFV-R8 (GT IX) and two ASFLI-elements. siRNA (22 nt) and piRNA (29 nt) fractions from *Ornithodoros moubata* were individually mapped against available ASFV whole-genome sequences (the one showing the most mapped reads, ASFV-R8 (GT IX) is shown) and two exemplary *O. moubata* ASFLI-element-containing contigs (NODE5010 and NODE28076) using Bowtie (1.1.2) in Geneious. Marked (arrows) are the most abundant small RNA molecules related to the ASFV genes and ASFLI-elements.

Additional file 13: Figure S4. siRNA/piRNA mapping against ASFV-Kenya1950 (GT X), Warthog (GT IV) and two ASFLI-elements. siRNA (22 nt) and piRNA (28-29 nt) fractions from *Ornithodoros porcinus* were individually mapped against available ASFV whole-genome sequences (the ones showing the most mapped reads, ASFV-Kenya1950 (GT X), ASFV Warthog (GT IV) are shown) and two exemplary *O. porcinus* ASFLI-element containing contigs (NODE955 and NODE1089) using Bowtie 1 (1.1.2) in Geneious. Marked (arrows) are the most abundant small RNA molecules relating to the ASFV genes and ASFLI-elements.

Additional file 14: Figure S5. siRNA/piRNA mapping against ASFV and ASFLI-element database. siRNA (22 nt) and piRNA (28-29 nt) fractions from *Ornithodoros porcinus* and *Ornithodoros moubata*, before and after deduplication, were individually mapped against ASFV whole-genome sequence and the *O. moubata* or *O. porcinus* ASFLI-element containing datasets using Bowtie 2 (2.3.0) in Geneious.

Additional file 15: Table S9. Results of small RNA sequencing and mapping against ASFV and ASFLI-elements. Small RNA was sequenced from *Ornithodoros porcinus* and *Ornithodoros moubata* nymphal stage ticks. After deduplication using BBMap, 22 nt siRNA and 28-29 nt piRNA fractions were extracted and mapped against ASFV and ASFLI-elements using Bowtie.

Additional file 16: Table S10. *O. moubata* piRNA ping-pong signature analysed by PingPongPro v1.0 with default parameters and -b option (creates additional browser track files, which are suitable for display in common genome browsers). The closer the Score-value (1 minus the FDR-value: estimated fraction of signatures that have the same combination of properties, but that are not true ping-pong signatures) which is calculated from adenine bias at pos. 10 of the piRNA, stack height - e.g. the number of reads that make up the stack and independence of local coverage is to 1, the more likely is a true ping-pong signature.

Additional file 17: Table S11. *O. porcinus* piRNA ping-pong signature analysed by PingPongPro v1.0 with default parameters and -b option (creates additional browser track files, which are suitable for display in common genome browsers). The closer the Score-value (1 minus the FDR-value: estimated fraction of signatures that have the same combination of properties, but that are not true ping-pong signatures) which is calculated from adenine bias at pos. 10 of the piRNA, stack height - e.g. the number of reads that make up the stack and independence of local coverage is to 1, the more likely is a true ping-pong signature.

Additional file 18: Figure S6. The reconstructed ASFV-like A104R protein is highly similar to its ASFV homologue (A) A rabbit antiserum raised against the reconstructed A104 gene recognised a flag-tagged and an

untagged version of A104R protein (lanes A104-Flag and A104, respectively), but showed no specific reaction with extracts of tick cell lines OME/CTVM21, OME/CTVM22, OME/CTVM24, and OME/CTVM27. In extracts of WSL-HP cells infected with ASFV Kenya 1033, the serum reacted with a single band of 12 kDa which is similar to the calculated molecular weight of ASFV A104R (11.6 kDa). **(B)** The Coomassie stained gel confirms equal loading with proteins.

Additional file 19: Table S12. Primer and probe sequences used in this study.

Additional file 20: Figure S7. BLAST-analysis for identification of ASFV-like element containing contigs and annotation. 66,745 SPAdes assembled contigs were blasted (BLASTn, NCBI, v2.6.0+) against a customised database comprising all sequences with the NCBI taxonomy ID 10497 (African swine fever virus) (as of 16 January 2018). Hits were filtered using a cut off e-value of 1×10^{-4} and a minimum alignment length of 150 bp, resulting in 34 contigs. These were then blasted against the complete NCBI database (The non-redundant nucleotide collection) to reliably identify and annotate ASFV-like sequences and areas of the host genome using default parameters. BLASTp search of >500 bp ORFs was performed against the "Non-redundant protein sequences" database using default parameters.

Acknowledgements

This work is dedicated to our friend and colleague Dr. Günther Keil who died on February 18, 2020.

The authors thank Carola Sauter-Louis (Friedrich-Loeffler-Institut, Germany) for the help with the statistical analysis; Florian Pfaff (Friedrich-Loeffler-Institut, Germany) for the help with RNA sequencing; the Tick Cell Biobank for the tick cell line provision; Ariane Dux (Robert-Koch-Institut, Germany) for the help with the museum-stored field ticks; Guido Brandt and Cäcilia Freund (MPI-Jena, Germany) for their help with HiSeq sequencing; Richard Bishop and Naftaly Githaka (ILRI-Kenya), Richard Lucius (Humboldt University Berlin-Germany), Fernando Boinas (University of Lisbon, Portugal), Sally Cutler (University of East London, UK), Cornelia Silaghi (Friedrich-Loeffler-Institut, Germany), Simon Carpenter (The Pirbright Institute, UK, under BBSRC project code BBS/E/1/00007039), Ricardo Pérez Sánchez (CSIC, Spain), Jason Dunlop (Berlin Museum for Natural History, Germany) and Laurence Vial (CIRAD, France) for supplying tick specimens; and furthermore, Juliane Horenk, Kathrin Giesow and Patrick Zitzow for the technical assistance and René Klein for the tick rearing.

Authors' contributions

JF, LF, HK, SB, DH and MB designed the experiments and did the comparative data analyses. JF conducted the qPCR, tick phylogeny and infection experiments and coordinated the study. JF, LF, SCS, AW, JK and DH conducted the NGS experiments and analysed the data. GMK did the cloning and expression of ASFV-like A104R. SL and JF did the phylogeny and SL the BEAST analysis. LBS generated the tick cell lines and provided nucleic acid samples. JF, LF, SB and MB drafted the manuscript, and LBS critically revised it. All authors commented on and approved the final manuscript.

Funding

JF was funded intramurally through the FLI 'ASF-Research Network' project. SL was funded by a BBSRC Roslin Institute Strategic Research programme BBS/E/D/20002173 and University of Edinburgh Chancellor's Fellowship. LBS was funded by UK BBSRC grant numbers BBS/E/1/00001741 and BB/P024270/1. AW and JK were funded by the Max Planck Society. This work was in part financed by COMPARE under the European Union's Horizon 2020 programme, grant no. 643476, and by the German Federal Ministry of Education and Research-funded project DetektiVir, grant no. 13N13783. Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The datasets supporting the conclusions of this article are available from the International Nucleotide Sequence Database Collaboration (INSDC) under study accession number PRJEB26739, <https://www.ebi.ac.uk/ena/data/search?query=PRJEB26739> [80].

Ethics approval and consent to participate

All parts of the study involving animals were carried out following all applicable German animal welfare regulations and institutional guidelines. Immunisation of rabbits was reported to the competent authority under file accession LALLF-Nr. 7221.3-2.5-04/10.

Consent for publication

Not applicable

Competing interests

The authors declare no competing financial interest.

Author details

¹Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Südufer 10, 17493 Greifswald-Insel Riems, Germany. ²The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, Scotland, UK. ³Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, 146 Brownlow Hill, Liverpool L3 5RF, UK. ⁴Robert Koch-Institut Berlin, Seestraße 10, 13353 Berlin, Germany. ⁵Max Planck Institute for the Science of Human History, Kahlaische Str. 10, 07745 Jena, Germany.

Received: 17 August 2020 Accepted: 4 September 2020

Published online: 08 October 2020

References

- Destoumieux-Garzon D, Mavingui P, Boetsch G, Boissier J, Darriet F, Duboz P, et al. The One Health concept: 10 years old and a long road ahead. *Front Vet Sci*. 2018;5:14.
- Kahn LH. Perspective: the one-health way. *Nature*. 2017;543(7647):S47.
- Mallapaty S. Spread of deadly pig virus in China hastens vaccine research. *Nature*. 2019;569(7754):13–4.
- Alonso C, Borca M, Dixon L, Revilla Y, Rodriguez F, Escribano JM, et al. ICTV virus taxonomy profile: Asfarviridae. *J Gen Virol*. 2018;99:613–4.
- Dixon LK, Chapman DA, Netherton CL, Upton C. African swine fever virus replication and genomics. *Virus Res*. 2013;173(1):3–14.
- Sanchez-Vizcaino JM, Mur L, Martinez-Lopez B. African swine fever: an epidemiological update. *Transbound Emerg Dis*. 2012;59(Suppl 1):27–35.
- Beltran-Alcrudo D, Lubroth J, Depner K, De La Rocque S. African swine fever in the Caucasus. *FAO Empres Watch*. 2008:1–8.
- Zhou X, Li N, Luo Y, Liu Y, Miao F, Chen T, et al. Emergence of African swine fever in China, 2018. *Transbound Emerg Dis*. 2018;65(6):1482–4.
- Linden A, Licoppe A, Volpe R, Paternostre J, Lesenfans C, Cassart D, et al. Summer 2018: African swine fever virus hits North-western Europe. *Transbound Emerg Dis*. 2018;66(1):54–5.
- Bao J, Wang Q, Lin P, Liu C, Li L, Wu X, et al. Genome comparison of African swine fever virus China/2018/AnhuiXCGQ strain and related European p72 genotype II strains. *Transbound Emerg Dis*. 2019;66(3):1167–76.
- Montgomery RE, On A. Form of swine fever occurring in British East Africa (Kenya Colony). *J Comp Pathol*. 1921;34:159–91.
- Jori F, Bastos AD. Role of wild suids in the epidemiology of African swine fever. *EcoHealth*. 2009;6(2):296–310.
- Alkhamis MA, Gallardo C, Jurado C, Soler A, Arias M, Sanchez-Vizcaino JM. Phylodynamics and evolutionary epidemiology of African swine fever p72-CVR genes in Eurasia and Africa. *PLoS One*. 2018;13(2):e0192565.
- Katzourakis A, Gifford RJ. Endogenous viral elements in animal genomes. *PLoS Genet*. 2010;6(11):e1001191.
- Bell-Sakyi L, Ruzek D, Gould EA. Cell lines from the soft tick *Ornithodoros moubata*. *Exp Appl Acarol*. 2009;49(3):209–19.
- Antipov D, Korobeynikov A, McLean JS, Pevzner PA. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. *Bioinformatics*. 2016;32(7):1009–15.
- Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nat Commun*. 2016;7:1–13.
- Oleaga A, Obolo-Mvoulouga P, Manzano-Roman R, Perez-Sanchez R. Functional annotation and analysis of the *Ornithodoros moubata* midgut genes differentially expressed after blood feeding. *Ticks Tick Borne Dis*. 2017;8(5):693–708.
- Ogihara MH, Hikiba J, Suzuki Y, Taylor D, Kataoka H. Ovarian ecdysteroidogenesis in both immature and mature stages of an Acari, *Ornithodoros moubata*. *PLoS One*. 2015;10(4):e0124953.

20. Mans BJ, Featherston J, Kvas M, Pillay KA, de Klerk DG, Pienaar R, et al. Argasid and ixodid systematics: implications for soft tick evolution and systematics, with a new argasid species list. *Ticks Tick Borne Dis.* 2019;10(1):219–40.
21. Oleaga A, Obolo-Mvoulouga P, Manzano-Roman R, Perez-Sanchez R. *De novo* assembly and analysis of midgut transcriptome of the argasid tick *Ornithodoros erraticus* and identification of genes differentially expressed after blood feeding. *Ticks Tick Borne Dis.* 2018;9(6):1537–54.
22. Gammon DB, Mello CC. RNA interference-mediated antiviral defense in insects. *Curr Opin Insect Sci.* 2015;8:111–20.
23. Stein CB, Genzor P, Mitra S, Elchert AR, Ipsaro JJ, Benner L, et al. Decoding the 5' nucleotide bias of PIWI-interacting RNAs. *Nat Commun.* 2019;10(1):828.
24. Schnettler E, Tykalova H, Watson M, Sharma M, Sterken MG, Obbard DJ, et al. Induction and suppression of tick cell antiviral RNAi responses by tick-borne flaviviruses. *Nucleic Acids Res.* 2014;42(14):9436–46.
25. Zhou J, Zhou Y, Cao J, Zhang H, Yu Y. Distinctive microRNA profiles in the salivary glands of *Haemaphysalis longicornis* related to tick blood-feeding. *Exp Appl Acarol.* 2013;59(3):339–49.
26. Czech B, Hannon GG. One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. *Trends Biochem Sci.* 2016;41(4):324–37.
27. Bell-Sakyi L, Attoui H. Virus discovery using tick cell lines. *Evol Bioinformatics Online.* 2016;12(Suppl 2):31–4.
28. Belyi VA, Levine AJ, Skalka AM. Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes. *PLoS Pathog.* 2010;6(7):e1001030.
29. Whitfield ZJ, Dolan PT, Kunitomi M, Tassetto M, Seetin MG, Oh S, et al. The diversity, structure, and function of heritable adaptive immunity sequences in the *Aedes aegypti* genome. *CurrB.* 2017;27(22):3511–9 e7.
30. Flegel TW. Hypothesis for heritable, anti-viral immunity in crustaceans and insects. *Biol Direct.* 2009;4:32.
31. Horie M, Kobayashi Y, Suzuki Y, Tomonaga K. Comprehensive analysis of endogenous bornavirus-like elements in eukaryote genomes. *Philos Trans R Soc B.* 2013;368(1626):1–11.
32. Utari HB, Soowannayan C, Flegel TW, Whityachumnarnkul B, Kruatrachue M. Variable RNA expression from recently acquired, endogenous viral elements (EVE) of white spot syndrome virus (WSSV) in shrimp. *Dev Comp Immunol.* 2017;76:370–9.
33. Flegel T. Update on viral accommodation, a model for host-viral interaction in shrimp and other arthropods. *Dev Comp Immunol.* 2007;31(3):217–31.
34. Tabares E, Sanchez BC. Synthesis of DNA in cells infected with African swine fever virus. *Arch Virol.* 1979;61(1–2):49–59.
35. Simoes M, Martins C, Ferreira F. Early intranuclear replication of African swine fever virus genome modifies the landscape of the host cell nucleus. *Virus Res.* 2015;210:1–7.
36. Frank JA, Feschotte C. Co-option of endogenous viral sequences for host cell function. *Curr Opin Virol.* 2017;25:81–9.
37. Aswad A, Katzourakis A. Paleovirology and virally derived immunity. *Trends Ecol Evol.* 2012;27(11):627–36.
38. Suzuki Y, Frangeul L, Dickson LB, Blanc H, Verdier Y, Vinh J, et al. Uncovering the repertoire of endogenous flaviviral elements in *Aedes* mosquito genomes. *J Virol.* 2017;91(15):e00571–17.
39. Palatini U, Miesen P, Carballar-Lejarazu R, Ometto L, Rizzo E, Tu Z, et al. Comparative genomics shows that viral integrations are abundant and express piRNAs in the arboviral vectors *Aedes aegypti* and *Aedes albopictus*. *BMC Genomics.* 2017;18(1):512.
40. Iwasaki YW, Siomi MC, Siomi H. PIWI-interacting RNA: Its biogenesis and functions. *Annu Rev Biochem.* 2015;84:405–33.
41. Barrero RA, Keeble-Gagnere G, Zhang B, Moolhuijzen P, Ikeo K, Tateno Y, et al. Evolutionary conserved microRNAs are ubiquitously expressed compared to tick-specific miRNAs in the cattle tick *Rhipicephalus (Boophilus) microplus*. *BMC Genomics.* 2011;12:328.
42. Hackenberg M, Langenberger D, Schwarz A, Erhart J, Kotsyfakis M. *In silico* target network analysis of de novo-discovered, tick saliva-specific microRNAs reveals important combinatorial effects in their interference with vertebrate host physiology. *Rna.* 2017;23(8):1259–69.
43. Dietrich I, Shi X, McFarlane M, Watson M, Blomstrom AL, Skelton JK, et al. The antiviral RNAi response in vector and non-vector cells against Orthobunyaviruses. *PLoS Negl Trop Dis.* 2017;11(1):e0005272.
44. Kurscheid S, Lew-Tabor AE, Rodriguez Valle M, Bruyeres AG, Doogan VJ, Munderloh UG, et al. Evidence of a tick RNAi pathway by comparative genomics and reverse genetics screen of targets with known loss-of-function phenotypes in *Drosophila*. *BMC Mol Biol.* 2009;10:1–21.
45. de Carvalho Ferreira HC, Tudela Zuquete S, Wijnveld M, Weesendorp E, Jongejan F, Stegeman A, et al. No evidence of African swine fever virus replication in hard ticks. *Ticks Tick Borne Dis.* 2014;5(5):582–9.
46. Kleinboeker SBS, G.A.; Burrage,T.G.; and Sur J.-H. African swine fever virus replication in the midgut epithelium is required for infection of *Ornithodoros* Ticks. *J Virol.* 1999;73(10):8587–98.
47. Pereira de Oliveira R, Hutet E, Paboef F, Duhayon M, Boinas F, Perez de Leon A, et al. Comparative vector competence of the Afrotropical soft tick *Ornithodoros moubata* and Palearctic species, *O. erraticus* and *O. verrucosus*, for African swine fever virus strains circulating in Eurasia. *PLoS one.* 2019; 14(11):e0225657.
48. Rennie L, Wilkinson PJ, Mellor PS. Effects of infection of the tick *Ornithodoros moubata* with African swine fever virus. *Med Vet Entomol.* 2000;14:355–60.
49. Rennie L, Wilkinson PJ, Mellor PS. Transovarial transmission of African swine fever virus in the argasid tick *Ornithodoros moubata*. *Med Vet Entomol.* 2001;15:140–6.
50. Keßler C, Forth JH, Keil GM, Mettenleiter TC, Blome S, Karger A. The intracellular proteome of African swine fever virus. *Sci Rep.* 2018;8(1):14714.
51. Frouco G, Freitas FB, Coelho J, Leitao A, Martins C, Ferreira F. DNA-binding properties of African swine fever virus pA104R, a histone-like protein involved in viral replication and transcription. *J Virol.* 2017;91(12):e02498–16.
52. Carrascosa AL, Bustos MJ, de Leon P. Methods for growing and titrating African swine fever virus: field and laboratory samples. *Curr Protoc Cell Biol.* 2011;53:1–25.
53. Onzere CK, Bastos AD, Okoth EA, Lichoti JK, Bochere EN, Owido MG, et al. Multi-locus sequence typing of African swine fever viruses from endemic regions of Kenya and eastern Uganda (2011–2013) reveals rapid B602L central variable region evolution. *Virus Genes.* 2018;54(1):111–23.
54. Kröber T, Guerin PM. In vitro feeding assays for hard ticks. *Trends Parasitol.* 2007;23(9):445–9.
55. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16(3):1215.
56. Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A.* 2013;110(39):15758–63.
57. Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philos T R Soc B-Biol Sci.* 2014;370(1660):20130624.
58. Tignon M, Gallardo C, Iscaro C, Hutet E, Van der Stede Y, Kolbasov D, et al. Development and inter-laboratory validation study of an improved new real-time PCR assay with internal control for detection and laboratory diagnosis of African swine fever virus. *J Virol Methods.* 2011;178(1–2):161–70.
59. Bastos AD, Arnot LF, Jacquier MD, Maree S. A host species-informative internal control for molecular assessment of African swine fever virus infection rates in the African sylvatic cycle *Ornithodoros* vector. *Med Vet Entomol.* 2009;23(4):399–409.
60. Pfaff F, Schlottau K, Scholes S, Courtenay A, Hoffmann B, Höper D, et al. A novel astrovirus associated with encephalitis and ganglionitis in domestic sheep. *Transbound Emerg Dis.* 2017;64(3):677–82.
61. Hanke D, Pohlmann A, Sauter-Louis C, Höper D, Stadler J, Ritzmann M, et al. Porcine epidemic diarrhea in Europe: in-detail analyses of disease dynamics and molecular epidemiology. *Viruses.* 2017;9(7):1–13.
62. Wylezich C, Papa A, Beer M, Höper D. A versatile sample processing workflow for metagenomic pathogen detection. *Sci Rep.* 2018;8(1):13108.
63. Scheuch M, Höper D, Beer M. RIEMs: a software pipeline for sensitive and comprehensive taxonomic classification of reads from metagenomics datasets. *BMC Bioinformatics.* 2015;16:69.
64. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics.* 2011;27(21):2957–63.
65. Bushnell B, Rood J, Singer E. BBMerge - accurate paired shotgun read merging via overlap. *PLoS One.* 2017;12(10):e0185056.
66. Uhrig S, Klein H. PingPongPro: a tool for the detection of piRNA-mediated transposon-silencing in small RNA-Seq data. *Bioinformatics.* 2019;35(2):335–6.
67. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7.
68. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 2013;30(12):2725–9.
69. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci.* 1992;8(3):275–82.

70. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012;29(8):1969–73.
71. Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst Biol.* 2012;61(1):170–3.
72. Baele G, Li WL, Drummond AJ, Suchard MA, Lemey P. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Mol Biol Evol.* 2013;30(2):239–43.
73. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 2006;4(5):699–710.
74. Drummond AJ, Suchard MA. Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 2010;8:1–12.
75. Yoder AD, Yang ZH. Estimation of primate speciation dates using local molecular clocks. *Mol Biol Evol.* 2000;17(7):1081–90.
76. Bauer A, Nolden T, Nemitz S, Perlon E, Finke S. A dynein light chain 1 binding motif in rabies virus polymerase L protein plays a role in microtubule reorganization and viral primary transcription. *J Virol.* 2015; 89(18):9591–600.
77. Keil GM, Pollin R, Müller C, Giesow K, Schirmer H. BacMam platform for vaccine antigen delivery. *Methods Mol Biol.* 2016;1349:105–19.
78. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970;227(5259):680–5.
79. Chow S, Shao J, Wang H. Sample size calculations in clinical research (Chapman & Hall/CRC Biostatistics Series), 978–0824709709, Shein-Chung Chow, CRC Press; 2 Rev Exp edition.
80. Forth JH, Forth LF, Lycett S, Bell-Sakyi L, Keil GM, Blome S et al. Identification of African swine fever virus-like elements in soft tick genome provides insights into the virus' evolution. Supplementary Datasets. 2020. INSDC ENA accession: PRJEB26739. <https://www.ebi.ac.uk/ena/data/search?query=PRJEB26739>. Accessed 15 Sept 2020.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

