View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Virology 438 (2013) 14–19

Contents lists available at SciVerse ScienceDirect

Virology

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

Sequencing approach to analyze the role of quasispecies for classical swine fever

Armin Töpfer [a,b], Dirk Höper [c], Sandra Blome [c], Martin Beer [c], Niko Beerenwinkel [a,b], Nicolas Ruggli [d], Immanuel Leifer [d,*]

[a] Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, CH-4058 Basel, Switzerland
[b] SIB Swiss Institute of Bioinformatics, CH-4058 Basel, Switzerland
[c] Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, Südufer 10, 17493 Greifswald-Insel Riems, Germany
[d] Institute of Virology and Immunoprophylaxis (IVI), Sensemattstrasse 293, CH-3147 Mittelhäusern, Switzerland

ARTICLE INFO

Article history:
Received 8 October 2012
Accepted 28 November 2012
Available online 13 February 2013

Keywords:
Classical swine fever virus
Pestivirus
Virulence
Quasispecies estimation
Haplotype reconstruction
ShoRAH
QuRe
QuasiRecomb

ABSTRACT

Classical swine fever virus (CSFV) is a positive-sense RNA virus with a high degree of genetic variability among isolates. High diversity is also found in virulence, with strains covering the complete spectrum from avirulent to highly virulent. The underlying genetic determinants are far from being understood. Since RNA polymerases of RNA viruses lack any proof-reading activity, different genome variations called haplotypes, occur during replication. A set of haplotypes is referred to as a viral quasispecies. Genetic variability can be a fitness advantage through facilitating of a more effective escape from the host immune response. In order to investigate the correlation of quasispecies composition and virulence *in vivo*, we analyzed next-generation sequencing data of CSFV isolates of varying virulence. Viral samples from pigs infected with the highly virulent isolates "Koslov" and "Brescia" showed higher quasispecies diversity and more nucleotide variability, compared to samples of pigs infected with low and moderately virulent isolates.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Classical swine fever (CSF) is an important contagious disease of pigs causing major economical damage in the pig industries (Edwards, Fukusho et al., 2000;Vandeputte and Chappuis, 1999). The causative agent is the classical swine fever virus (CSFV), a member of the genus *Pestivirus* grouped into the family *Flaviviridae* (Fauquet and Fargette, 2005, King, Lefkowitz et al., 2011). *Pestiviruses* possess a single-stranded positive-sense RNA genome of approximately 12,300 nucleotides, with 5′-terminal and 3′-terminal non-translated regions (5′-NTR and 3′-NTR, respectively) (Meyers and Thiel, 1996). The genome contains one open reading frame encoding a polyprotein that is processed by cellular and viral proteases (Meyers and Thiel, 1996) to the structural (C, E^{rns}, E1, and E2) and the non-structural proteins (N^{pro}, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (Rumenapf, Unger et al., 1993;Tautz, Elbers et al., 1997;Thiel, Stark et al., 1991).

The envelope glycoprotein E2 is highly immunogenic and essential for replication (Van Gennip, Bouma et al., 2002). Moreover, it was shown that it plays a role in viral adsorption to host cells together with other surface proteins, namely E^{rns} and E1 (Hulst and Moormann, 1997;Wang, Nie et al., 2004, Reimann, Depner et al., 2004). The E2 protein shows a high variability among the different CSFV isolates. Despite of cross neutralizing antibodies, an E2 escape from the host immune response seems possible, as it was demonstrated previously by *in vitro* studies (Leifer, Blome et al., 2012). The NS5B protein is the RNA-dependent RNA polymerase (RdRp) and it is more conserved among CSFV isolates (Bjorklund, Lowings et al., 1999). The E2 and NS5B encoding regions were selected for quasispecies analyses to investigate differences between conserved and variable genome regions.

CSF disease can vary from acute hemorrhagic fever to chronic or unapparent infection, which is dependent essentially on the virulence of the viral isolate. Accordingly, CSFV can be classified in highly virulent, moderately to low virulent, and avirulent strains, the latter being mainly vaccine strains (Floegel-Niesmann, Blome et al., 2009;Li, Xu et al., 2006;Mittelholzer, Moser et al., 2000;Tao, Dai et al., 2009). A number of studies have been undertaken to

* Corresponding author. Fax: +41 31 848 9271.
E-mail addresses: Armin.Toepfer@bsse.ethz.ch (A. Töpfer), Dirk.Hoeper@fli.bund.de (D. Höper), Sandra.Blome@fli.bund.de (S. Blome), Martin.Beer@fli.bund.de (M. Beer), Niko.Beerenwinkel@bsse.ethz.ch (N. Beerenwinkel), Nicolas.Ruggli@ivi.admin.ch (N. Ruggli), Immanuelleifer@hotmail.com, Immanuel.Leifer@ivi.admin.ch (I. Leifer).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2012.11.020

investigate genetic markers that may account for the observed differences in CSFV virulence (Risatti, Borca et al., 2005;Risatti, Holinka et al., 2005;Risatti, Holinka et al., 2006;Risatti, Holinka et al., 2007b;Risatti, Holinka et al., 2007a;Sainz, Holinka et al., 2008;Tamura, Sakoda et al., 2012;Van Gennip, Vlot et al., 2004). Despite all these efforts, the viral determinants of virulence are not well understood yet.

Because the RdRp of RNA viruses lack any proof-reading activity, their error rate during genome replication is high. As long as the random mutations are not self-limiting, viruses with genetic differences are maintained in the same host during viral replication. This phenomenon was first described in 1971 and defined as quasispecies (Eigen, 1971). High genome diversity can be a fitness advantage through enabling viral escape from the immune response or from antiviral drug treatment (Thiel, Peters et al., 2002). The influence of viral quasispecies on pathogenesis was reported for several viruses, for instance picornaviruses, foot-and-mouth disease virus, and West Nile fever virus (Jerzak, Bernard et al., 2007;Sanz-Ramos, Diaz-San et al., 2008;Vignuzzi, Stone et al., 2006). This raises the question if the quasispecies composition of CSFV differs between isolates, and if so, whether these differences can be related to fitness or virulence of selected strains.

At present, almost nothing is reported on quasispecies composition of pestiviruses and the possible relevance for virulence. Therefore, the aim of this study was to analyze the quasispecies composition of RNA samples of CSFV isolates differing in virulence in the E2 and NS5B protein encoding genome regions. Sequence data were generated by next-generation sequencing (NGS). Single-site and local diversity estimates can be obtained reliably from NGS data, and it has been shown that local reconstruction can be a good measure of the global diversity, as the majority of the underlying diversity can often be observed in a single window (Zagordi, Däumer et al., 2012). For quasispecies analysis, three different open-source haplotype reconstruction tools, namely ShoRAH (Zagordi, Bhattacharya et al., 2011), QuRe (Prosperi and Salemi, 2012), and QuasiRecomb (Töpfer et al., 2012) were used. Haplotype reconstruction was performed on three different spatial levels: (i) at single sites of the DNA sequence, (ii) in a sliding window of a 300 bp width, called local reconstruction, and (iii) across the full-length of each of the two proteins, referred to as global reconstruction. The diversity of each inferred quasispecies was quantified as the entropy of the respective mutant distribution.

Results and discussion

In this study, we analyzed the intra-strain genomic diversity and estimated the quasispecies composition of the following five CSFV isolates: "Brescia" and "Koslov" (both highly virulent), and "Uelzen", "Paderborn", and "Hennef" (all three moderately virulent). For reliable quasispecies estimation in the E2 and NS5B protein coding regions, including the detection of low-frequency variants, we used a high sequencing coverage of 5333 ± 3244 (mean $\pm$ sd) reads per sequence position for E2 and 5002 ± 2090 for NS5B. Independent of the sample, the coverage varied greatly within the sequenced regions (Fig. 1). During library generation and emulsion PCR, the effectiveness of random primers could have caused these fluctuations, since the primers bind with a different affinity to the corresponding DNA regions. The average read length obtained was 363 bp, except for the isolate "Uelzen" with an average length of 271 bp.

In the E2 and NS5B polyprotein sequence of viable CSFV field isolates, single nucleotide deletions and insertions were never observed. The single-base gaps detected by NGS are likely to reflect sequencing errors. If real, each of these gaps would lead to a frame shift mutation in the polyprotein, which would result in non-functional viral proteins and premature translation stop disabling genome replication due to the *cis* requirement of protein translation for *pestivirus* RNA replication (Frolov, McBride et al., 1998). For CSFV, proteins expressed from alternative reading frames have never been described. Therefore, gaps in the raw sequencing data were not considered for haplotype estimation and quasispecies analysis. Importantly, no insertions or deletions of one or more entire codons have been observed in the data.

As a first level of analysis, Fig. 1 shows, for each isolate, the single-site entropy (*i.e.* diversity) landscape along the E2 and NS5B protein encoding sequences. In general, regions with higher nucleotide variability with clear-cut peaks within E2 and NS5B appear randomly distributed and not restricted to any particular

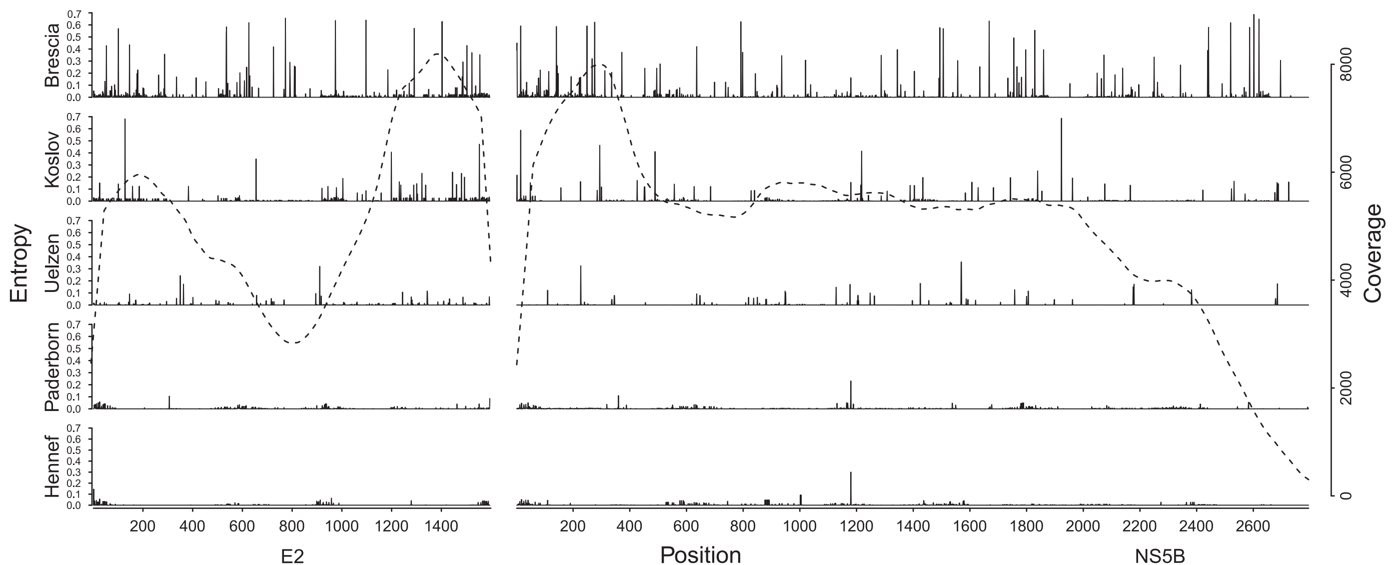

Fig. 1. Position-wise entropy of nucleotide distributions in the E2 and NS5B protein encoding regions. For each strain (Brescia, Koslov, Uelzen, Paderborn, Hennef), the entropy (left axis) computed from error-corrected reads is shown at each single genomic site of the E2 (left) and NS5B (right) encoding sequences. The mean coverage (right axis) of the next-generation sequencing of the five isolates data is plotted as a dotted lined.

regions. For the moderately virulent isolates ‘‘Hennef’’ and ‘‘Paderborn’’, a few regions with slightly elevated nucleotide variability were found (Fig. 1). We found higher mean nucleotide diversity in the highly virulent isolates compared to the moderately virulent isolates [Table 1, $p=0.027$ for both proteins, with the Mann–Kendall trend test (Mann, 1945;Kendall, 1975)].

Next, using different computational tools, we reconstructed for each isolate and each protein the viral quasispecies to analyze its diversity on a local and global scale (Fig. 2). For both proteins, the average diversity increases with virulence on both the local and global scale ($p=0.027$, Mann–Kendall). In general, estimated diversities of NS5B are slightly, but not significantly, higher than those of E2. This is surprising, since conventional Sanger sequencing showed higher conservation of the NS5B protein compared with the E2 protein (Leifer et al., 2010). Furthermore, it is assumed that the RdRp is more conserved compared to the E2 envelope glycoprotein, which is highly immunogenic and responsible for effective host antibody response. The difference in diversity may be due to length differences, since NS5B is 75% longer than E2 and therefore contains more variable positions.

This study focuses primarily on the inter-strain quasispecies diversity with respect to virulence, which implies assessment of the intra-strain mutant distribution for each individual isolate. Analysis of the raw data from NGS with different computational haplotype reconstruction tools can lead to different results. This is shown for instance with strain ‘‘Koslov’’ in Fig. 3. The estimated haplotype distributions display for the two genes and the three reconstruction methods the typical quasispecies distribution, in which a few haplotypes are dominant and many others are of very low frequency. The reconstruction tool ShoRAH estimates quasispecies which are approximately 20 times larger in the number of different mutants than predictions of the more conservative tool QuRe. Quasispecies predictions of QuasiRecomb are 50 times larger than those of ShoRAH. Thus, the differences in size of the estimated haplotype distributions are due to different methodologies (see Methods). Even though the quasispecies estimates differ, they are consistent regarding the inter-strain trend of increasing diversity with increasing virulence observed across all three levels of the analysis.

Table 1
Single-site means of the entropy landscapes.

	Brescia	Koslov	Uelzen	Paderborn	Hennef
E2	0.0120	0.0054	0.0021	0.0014	0.0010
NS5B	0.0109	0.0041	0.0032	0.0024	0.0012

The means of the position-wise entropy landscapes are shown for each strain (column-wise) and for both proteins (E2, first row; NS5B, second row).

Based on the finding that viruses rescued from the CSFV cDNA clones of the highly virulent isolates ‘‘Eystrup’’ and ‘‘Brescia’’ triggered severe clinical picture in animal experiments (Mayer, Thayer et al., 2003;Risatti, Borca et al., 2005), the quasispecies diversity is probably not *per se* responsible for the manifestation and severity of clinical signs. This is supported by the finding that low-fitness viruses, i.e. viruses with little quasispecies diversity, also can have a highly virulent phenotype (Ojosnegros, Beerenwinkel et al., 2010). Contrary, single clones can rapidly generate high diversity after few replication cycles (Ojosnegros, Beerenwinkel et al., 2010). As in this latter study, virulence is often defined as the capability to kill cells *in vitro* (Herrera, Garcia-Arriaza et al., 2007). In the present study however, virulence is defined as the ability to cause clinical signs after virus infections *in vivo*. With CSFV, acquisition of cytopathogenicity *in vitro* was associated with loss of virulence (Gallei, Blome et al., 2008).

According to the data presented here, highly virulent CSFV field samples showed a more peaked single-site entropy landscape, meaning higher diversity in the nucleotide sequences and therefore in the estimated quasispecies in the E2 and NS5B encoding regions. The higher replication rates of the highly virulent isolates can be one reason for the higher genome variability and quasispecies diversity. This diversity was slightly more prominent in the NS5B mRNA. It has to be mentioned that samples of the highly virulent isolates were taken 5 to 7 days post infection from blood, when animals were euthanized for animal welfare reasons at that time point. All other samples were taken between 10 to 16 days post infection. If the time of sampling influences the quasispecies composition, we would suspect increase in the quasispecies variability during the course

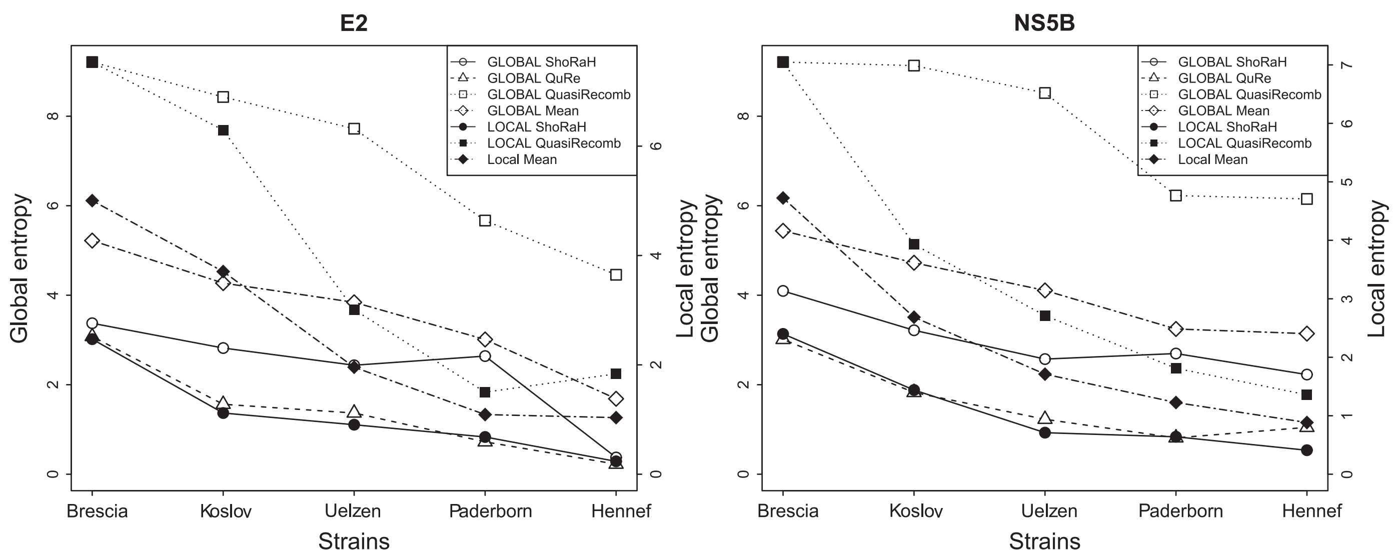

Fig. 2. Local and global diversity of E2 and NS5B haplotype distributions using different reconstruction methods. For each strain (Brescia, Koslov, Uelzen, Paderborn, Hennef) and each protein encoding sequence (E2, left; NS5B, right), the entropy of the predicted haplotype distribution is shown for the local (filled symbols, right axes) and global (unfilled symbols, left axis) analysis. Global haplotype reconstructions were performed using the computational tools ShoRAH, QuasiRecomb, and QuRe; local haplotypes were reconstructed with ShoRAH and QuasiRecomb. In addition, for each strain, protein encoding sequence, and scale, the mean entropy across all methods is shown.

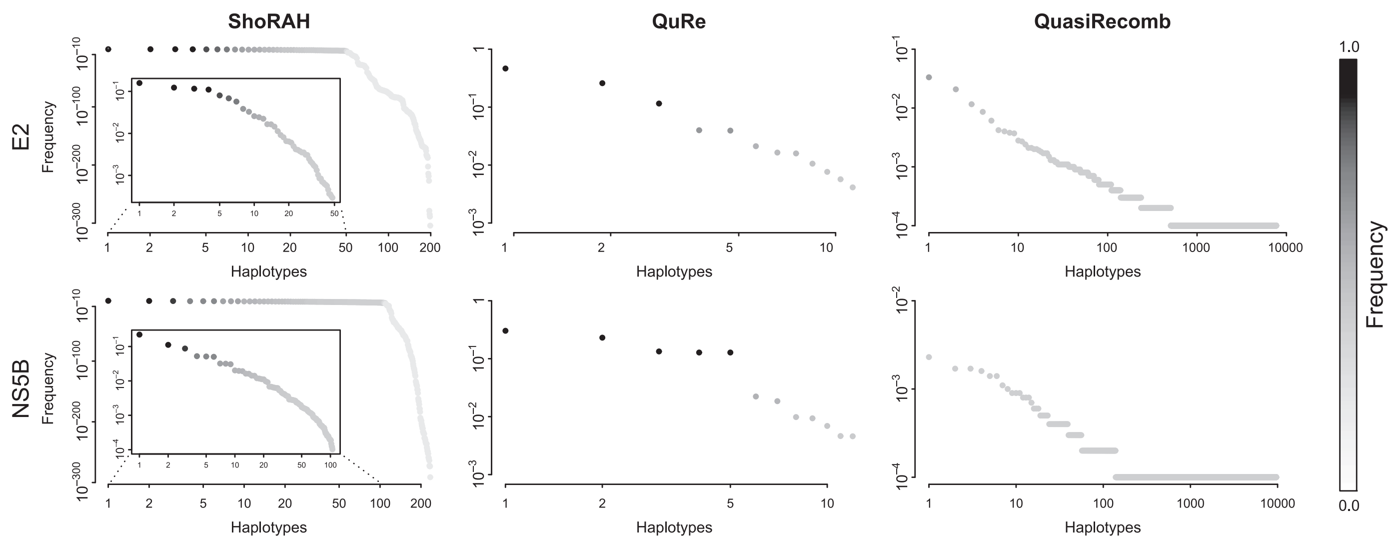

Fig. 3. Global haplotype distribution of the highly virulent strain Koslov. Displayed are the haplotype distributions on a log-log scale, *i.e.*, the estimated relative frequencies of all reconstructed haplotypes, for each gene (E2, first row; NS5B, second row) and for three different computational haplotype prediction methods (ShoRAH, first column; QuRe, second column; and QuasiRecomb, third column). For the first column, a zoom-in is included to show the distribution for haplotypes with frequency greater than 10^{-4}. Even though the number of called haplotypes differs by orders of magnitude, the shape of the distributions is similar. See material and methods for further explanations of the methods and their conceptual differences.

of disease. As little is known about quasispecies variation over time in pestiviruses, it is not known if time of sampling would influence the quasispecies distribution.

The quasispecies diversity for the highly virulent strain "Koslov" differs only slightly from the low and moderately virulent isolates "Hennef" and "Uelzen". This fits with the error catastrophe theory, stating that the variability of the viral quasispecies is limited (Eigen, 1971). According to this theory, high polymerase replication error rates are self-limiting due to high mutation rates leading to abortive virus infections. Interestingly however, the highly virulent isolate "Brescia" showed markedly higher quasispecies diversity than "Koslov", which altogether suggests that CSFV does apparently tolerate a large range of quasispecies diversity. The high genome variability of the isolate "Brescia" shows indeed that the higher fitness advantage of this viral strain does not imply loss of virulence.

Conclusions

In summary, this is the first analysis of CSFV genome quasispecies composition and diversity using *ex-vivo* samples of acute CSF. *In vivo*, highly virulent CSFV isolates showed higher genetic variability and quasispecies diversity in the E2 and NS5B encoding regions compared with low and moderately virulent CSFV isolates, indicating an association between virulence and quasispecies diversity in CSFV. This association needs to be validated or falsified in future studies analyzing additional isolates under controlled conditions. Higher quasispecies diversity and increased virulence may have common unknown causes and represent a fitness advantage for CSFV isolates.

In future studies, a larger number of different CSFV isolates and more replicates per isolate need to be analyzed to gain more statistical power to draw conclusions. In order to ensure that the quasispecies is not biased by different setups, animal infections and sample extractions should be synchronized. The number of passages (animal experiments, field, or cell culture) the virus isolates were passed through prior to quasispecies analysis will influence the results as well. The quasispecies distribution may vary depending on the time after infection and the site of replication, which should also be considered. To find out more about the quasispecies differences in the field, samples from different outbreaks should sequenced and analyzed directly.

It is also not known how fast and to what degree of diversity quasispecies can emerge from a homogenous virus preparation, and how this would relate to virulence. This would imply systematic analysis *in vivo* using cDNA-derived CSFV of different virulence.

Material and methods

Preparation of viral cDNA libraries

Viral RNA was extracted from whole blood or organ material of field samples, or from animal trials, according to a protocol published previously (Rasmussen, Reimann et al., 2009). Briefly, RNA was extracted using a combination of Trizol®LS (Life Technologies), chloroform, and RNeasy column (Quiagen, Hilden, Germany). RNA of the following samples were extracted for sequencing: CSFV isolate "Koslov" (highly virulent, genotype 1.1, first occurrence in Czechoslovakia, year unknown) from blood collected 6 days post infection (dpi) from infected piglets of animal trial 17/2005 at the Friedrich Loeffler Institute (FLI); CSFV "Brescia" (highly virulent, genotype 1.2, first occurrence in Italy, year unknown) from blood collected 6 dpi from animal trial at IVI 1/1998; CSFV "Hennef" (moderately virulent, genotype 2.3, German field sample K47/09-2, isolated from outbreak in 2009 from spleen suspension prepared from very small piglet 10–12 dpi, next related isolate Euskirchen [2006]); CSFV 0277 "Paderborn" (moderately virulent, genotype 2.1, first occurrence in Germany in 1997) from spleen collected from a pig 16 dpi of animal trial at FLI; CSFV 0634 "Uelzen" (moderately virulent, genotpye 2.3, first occurrence in Germany 1999, next related isolate Uelzen 504, isolated in Germany in 1992) from lymph nodes collected from a pig 14 dpi of animal trial at FLI.

For library generation, cDNA was prepared from viral RNA by reverse transcription using the Super Script III First Strand synthesis kit (Life Technologies). The number of CSFV RNA genomes were calculated by qRT-PCR (Hoffmann, Beer et al., 2005). For each sample, 10^5 RNA templates were used for cDNA generation. The cDNA was amplified by PCR using primers

CSFV2003F (5′-ATG GGC GGC CAC CTA TCA GAA TT-3′) and CSFV3613R (5′-GTG TGG GTR ATT ARG TTC CCT AT-3′) for the E2 gene, and primers CSFV9448F 5′-GGG GTC TGA TAA RGC CAC CTC-3′) and CSFV12297Ra (5′-GGG CCG TTA GGA AAT TAC CTT AGT-3′) for NS5B. For DNA amplification, a mastercycler gradient (Eppendorf) with the following cycler program was used: initial step of 94 °C for 15 s and 25 cycles of 94 °C for 15 s (denaturation), 65 °C for 30 s (annealing), and 68 °C for 3 min (elongation). For DNA amplification, the AccuPrime Taq DNA Polymerase High Fidelity kit (Life Technologies) was used according to the manufacturer's recommendations.

Next-generation sequencing

Nucleotide sequence data of the CSFV E2 and NS5B protein encoding regions were generated by pyrosequencing with a 454 GS FLX Genome Sequencer (Roche Diagnostics, Mannheim, Germany) as described previously (Leifer et al., 2010). Briefly, CSFV DNA fragments (obtained by RT-PCR) were separated by agarose gel electrophoresis and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Orange, *CA*, USA) prior to library generation. To this end, DNA libraries were prepared according to published protocols (Wiley, Macmil et al., 2009) followed by binding to library capture beads and recovery of the single-stranded template DNA library. These libraries were used for sequencing in 1/16 of a picotiter plate according to the manufacturer's instructions.

Sequence data analysis

The InDelFixer (2012) software (http://www.cbg.ethz.ch/software/InDelFixer/) was used to extract reads from the standard flowgram format (sff) files, to perform quality clipping with respect to the annotations in the sff files, to create a multiple alignment of the reads by pairwise alignment with a reference genome, and to correct frame shift-causing insertions and deletions in the reads. For quasispecies reconstruction, three computational tools were used. ShoRAH (Zagordi, Geyrhofer et al., 2010) corrects sequencing errors by locally clustering reads in small windows of the alignment and reconstructs a minimal set of global haplotypes that explain the reads. QuRe (Prosperi and Salemi, 2012) uses a Poisson error correction and a read graph approach to find the most probable set of overlapping variants as an estimate of the quasispecies. Both methods reconstruct global haplotypes as paths in a graph whose vertices correspond to the observed reads. QuasiRecomb (Töpfer, Zagordi et al., 2012) implements a probabilistic model, which assumes that the underlying diversity is generated from a small set of sequences by mutation and recombination. It estimates the position-wise sequencing error rate and predicts the distribution of haplotypes. This model-based parametric distribution is conceptually different from the output of ShoRAH and QuRe, which aim at enumerating all haplotypes that can be explained by the observed reads.

Quasispecies diversity was quantified as the Shannon entropy (Cover and Thomas, 2006) of the mutant distribution at three different spatial scales, namely single-site, local, and global. For the single-site analysis, the position-wise entropy of the nucleotide distributions was computed from reads error-corrected by ShoRAH. For the local analysis, QuasiRecomb and ShoRAH were used to reconstruct local haplotypes in sliding windows of 300 bp size from all reads covering the respective window. According to previous studies, the window with the highest entropy correlates well with the overall entropy (Zagordi, Däumer et al., 2012), and was therefore selected for quasispecies estimation. In the global analysis, ShoRAH, QuasiRecomb, and QuRe were used to reconstruct full-length E2 and NS5B haplotypes from overlapping reads. For testing the association of diversity with virulence, entropy estimates were averaged across computational reconstruction tools, for each isolate, protein coding sequence, and spatial scale.

Author contributions

AT performed sequence analysis and haplotype reconstruction, and wrote the manuscript. IL planned and supervised the project, prepared the samples, and was involved in drafting the manuscript and data analysis. NR was involved in supervision of the project and correcting the manuscript. DH performed the sequencing and was involved in preliminary data analysis. NB was involved in drafting and revising the manuscript, planning sequencing data analyses, and discussions of results. MB and SB carefully revised the manuscript and were involved in discussions of the data and performing animal experiments for viral sample collection.

Funding

This study was supported by the Swiss Federal Veterinary Office (BVET) (Grant no. 1.10.13) and by the Swiss National Science Foundation [Grant no. CR32I2_127017].

Acknowledgments

We thank Moctezuma Reimann for excellent technical assistance, Thomas Tschannen-Müller for help with computer and software problems, and the CSFV community reference laboratory (CRL Hannover, Germany) for providing CSFV samples.

References

Bjorklund, H., Lowings, P., Stadejek, T., Vilcek, S., Greiser-Wilke, I., Paton, D., Belak, S., 1999. Phylogenetic comparison and molecular epidemiology of classical swine fever virus. Virus Genes. 19, 189–195.

Cover, T.M., Thomas, J.A., Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing). Publisher: Wiley-Interscience | ISBN: 0471241954 | edition 2006.

Edwards, S., Fukusho, A., Lefevre, P., Lipowski, A., Pejsak, Z., Roehe, P., Westergaard, J., 2000. Classical swine fever: the global situation. Vet. Microbiol. 73, 103–119.

Eigen, M., 1971. Molecular self-organization and the early stages of evolution. Q. Rev. Biophys. 4, 149–212.

Fauquet, C.M., Fargette, D., 2005. International committee on taxonomy of viruses and the 3142 unassigned species. J.Virol 2, 64.

Floegel-Niesmann, G., Blome, S., Gerss-Dulmer, H., Bunzenthal, C., Moennig, V., 2009. Virulence of Classical Swine Fever virus isolates from Europe and other areas during 1996 until 2007. Vet. Microbiol. 139, 165–169.

Frolov, I., McBride, M.S., Rice, C.M., 1998. cis-acting RNA elements required for replication of bovine viral diarrhea virus-hepatitis C virus 5′ nontranslated region chimeras. RNA 4, 1418–1435.

Gallei, A., Blome, S., Gilgenbach, S., Tautz, N., Moennig, V., Becher, P., 2008. Cytopathogenicity of Classical Swine Fever Virus Correlates with Attenuation in the Natural Host. J. Virol. 82, 9717–9729.

Herrera, M., Garcia-Arriaza, J., Pariente, N., Escarmis, C., Domingo, E., 2007. Molecular basis for a lack of correlation between viral fitness and cell killing capacity. PLoS. Pathogens 3, e53.

Hoffmann, B., Beer, M., Schelp, C., Schirrmeier, H., Depner, K., 2005. Validation of a real-time RT-PCR assay for sensitive and specific detection of classical swine fever. J. Virological Methods 130, 36–44.

Hulst, M.M., Moormann, R.J., 1997. Inhibition of pestivirus infection in cell culture by envelope proteins E(rns) and E2 of classical swine fever virus: E(rns) and E2 interact with different receptors. J. Gen. Virol. 78, 2779–2787.

InDelFixer, 〈http://www.cbg.ethz.ch/software/InDelFixer/〉, September 2012.

Jerzak, G.V., Bernard, K., Kramer, L.D., Shi, P.Y., Ebel, G.D., 2007. The West Nile virus mutant spectrum is host-dependant and a determinant of mortality in mice. Virology 360, 469–476.

Kendall, M.G., 1975. Rank Correlation Methods. Griffin, London, UK.

King, A.M.Q., Lefkowitz, E., Adams, M.J., Carstens, E.B., 2011. Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses, Oxford.

Leifer, I., Blome, S., Blohm, U., Konig, P., Kuster, H., Lange, B., Beer, M., 2012. Characterization of C-strain "Riems" TAV-epitope escape variants obtained through selective antibody pressure in cell culture. Vet. Res. 43, 33.
Leifer, I., Hoffmann, B., Hoper, D., Bruun, R.T., Blome, S., Strebelow, G., Horeth-Bontgen, D., Staubach, C., Beer, M., 2010. Molecular epidemiology of current classical swine fever virus isolates of wild boar in Germany. J. Gen. Virol. 91, 2687–2697.
Li, X., Xu, Z., He, Y., Yao, Q., Zhang, K., Jin, M., Chen, H., Qian, P., 2006. Genome comparison of a novel classical swine fever virus isolated in China in 2004 with other CSFV strains. Virus Genes 33, 133–142.
Mann, H.B., 1945. Nonparametric tests against trend. Econometrica 13, 245–259.
Mayer, D., Thayer, T.M., Hofmann, M.A., Tratschin, J.D., 2003. Establishment and characterisation of two cDNA-derived strains of classical swine fever virus, one highly virulent and one avirulent. Virus Res. 98, 105–116.
Meyers, G., Thiel, H.J., 1996. Molecular characterization of pestiviruses. Adv. Virus Res. 47, 53–118.
Mittelholzer, C., Moser, C., Tratschin, J., Hofmann, M.A., 2000. Analysis of classical swine fever virus replication kinetics allows differentiation of highly virulent from avirulent strains. Vet. Microbiol. 74, 293–308.
Ojosnegros, S., Beerenwinkel, N., Antal, T., Nowak, M.A., Escarmis, C., Domingo, E., 2010. Competition-colonization dynamics in an RNA virus. Proc. Natl. Acad. Sci. USA 107, 2108–2112.
Prosperi, M.C., Salemi, M., 2012. QuRe: software for viral quasispecies reconstruction from next-generation sequencing data. Bioinformatics 28, 132–133.
Rasmussen, T.B., Reimann, I., Uttenthal, A., Leifer, I., Depner, K., Schirrmeier, H., Beer, M., 2009. Generation of recombinant pestiviruses using a full-genome amplification strategy. Vet. Microbiol. 142, 13–17.
Reimann, I., Depner, K., Trapp, S., Beer, M., 2004. An avirulent chimeric Pestivirus with altered cell tropism protects pigs against lethal infection with classical swine fever virus. Virology 322, 143–157.
Risatti, G.R., Borca, M.V., Kutish, G.F., Lu, Z., Holinka, L.G., French, R.A., Tulman, E.R., Rock, D.L., 2005. The E2 glycoprotein of classical swine fever virus is a virulence determinant in swine. J. Virol. 79, 3787–3796.
Risatti, G.R., Holinka, L.G., Carrillo, C., Kutish, G.F., Lu, Z., Tulman, E.R., Sainz, I.F., Borca, M.V., 2006. Identification of a novel virulence determinant within the E2 structural glycoprotein of classical swine fever virus. Virology 355, 94–101.
Risatti, G.R., Holinka, L.G., Fernandez, S., Carrillo, I., Kutish, C., Lu, G.F., Zhu, Z., Rock, J., Borca, M. V., D.L., 2007a. Mutations in the carboxyl terminal region of E2 glycoprotein of classical swine fever virus are responsible for viral attenuation in swine. Virology 364, 371–382.
Risatti, G.R., Holinka, L.G., Fernandez, S., Carrillo, I., Lu, C., Borca, M. V., Z., 2007b. N-linked glycosylation status of classical swine fever virus strain Brescia E2 glycoprotein influences virulence in swine. J. Virol. 81, 924–933.
Risatti, G.R., Holinka, L.G., Lu, Z., Kutish, G.F., Tulman, E.R., French, R.A., Sur, J.H., Rock, D.L., Borca, M.V., 2005. Mutation of E1 glycoprotein of classical swine fever virus affects viral virulence in swine. Virology 343, 116–127.
Rumenapf, T., Unger, G., Strauss, J.H., Thiel, H.J., 1993. Processing of the envelope glycoproteins of pestiviruses. J. Virol. 67, 3288–3294.
Sainz, I.F., Holinka, L.G., Lu, Z., Risatti, G.R., Borca, M.V., 2008. Removal of a N-linked glycosylation site of classical swine fever virus strain Brescia Erns glycoprotein affects virulence in swine. Virology 370, 122–129.
Sanz-Ramos, M., Diaz-San, S.F., Escarmis, C., Domingo, E., Sevilla, N., 2008. Hidden virulence determinants in a viral quasispecies *in vivo*. J. Virol. 82, 10465–10476.
Tamura, T., Sakoda, Y., Yoshino, F., Nomura, T., Yamamoto, N., Sato, Y., Okamatsu, M., Ruggli, N., Kida, H., 2012. Selection of classical swine fever virus with enhanced pathogenicity reveals synergistic virulence determinants in E2 and NS4B. J. Virol. 86, 8602–8613.
Tao, P., Dai, L., Luo, M., Tang, F., Tien, P., Pan, Z., 2009. Analysis of synonymous codon usage in classical swine fever virus. Virus Genes 38, 104–112.
Tautz, N., Elbers, K., Stoll, D., Meyers, G., Thiel, H.J., 1997. Serine protease of pestiviruses: determination of cleavage sites. J. Virol. 71, 5415–5422.
Thiel, H.J., Stark, R., Weiland, E., Rumenapf, T., Meyers, G., 1991. Hog cholera virus: molecular composition of virions from a pestivirus. J. Virol. 65, 4705–4712.
Thiel, J., Peters, T., Mas, M.A., Rosler, B., Peter, H.H., Weiner, S.M., 2002. Kinetics of hepatitis C (HCV) viraemia and quasispecies during treatment of HCV associated cryoglobulinaemia with pulse cyclophosphamide. Ann. Rheum. Dis. 61, 838–841.
Töpfer A., Zagordi O., Prabhakaran S., Roth V., Halperin E., Beerenwinkel N., 2012. Probabilistic inference of viral quasispecies subject to recombination. J. Comput. Biol. 20(2), http://dx.doi.org/10.1089/cmb.2012.0232.
Van Gennip, H.G., Bouma, A., van Rijn, P.A., Widjojoatmodjo, M.N., Moormann, R.J., 2002. Experimental non-transmissible marker vaccines for classical swine fever (CSF) by trans-complementation of E(rns) or E2 of CSFV. Vaccine 20, 1544–1556.
Van Gennip, H.G., Vlot, A.C., Hulst, M.M., de Smit, A.J., Moormann, R.J., 2004. Determinants of virulence of classical swine fever virus strain Brescia. J. Virol. 78, 8812–8823.
Vandeputte, J., Chappuis, G., 1999. Classical swine fever: the European experience and a guide for infected areas. Rev. Sci. Tech. 18, 638–647.
Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R., 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439, 344–348.
Wang, Z., Nie, Y., Wang, P., Ding, M., Deng, H., 2004. Characterization of classical swine fever virus entry by using pseudotyped viruses: E1 and E2 are sufficient to mediate viral entry. Virology 330, 332–341.
Wiley, G., Macmil, S., Qu, C., Wang, P., Xing, Y., White, D., Li, J., White, J.D., Domingo, A., Roe, B.A., 2009. Methods for generating shotgun and mixed shotgun/paired-end libraries for the 454 DNA sequencer. Curr. Protocols Hum. Genet., Chapter 18, Unit18.
Zagordi, O., Bhattacharya, A., Eriksson, N., Beerenwinkel, N., 2011. ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. BMC. Bioinformatics. 12, 119.
Zagordi,O., Däumer, M., Beisel, C., Beerenwinkel, N., Read length versus depth of coverage for viral quasispecies reconstruction. PLoS One. 7(10). :e47046. doi: 10.1371/journal.pone.0047046. Epub 2012 Oct 3.
Zagordi, O., Geyrhofer, L., Roth, V., Beerenwinkel, N., 2010. Deep sequencing of a genetically heterogeneous sample: local haplotype reconstruction and read error correction. J. Comput. Biol. 17, 417–428.