

# Host genetic factors associated with the range limit of a European hantavirus

Moritz Saxenhofer<sup>1,2</sup> | Anton Labutin<sup>1</sup>  | Thomas A. White<sup>1</sup> | Gerald Heckel<sup>1,2</sup> 

<sup>1</sup>Institute of Ecology and Evolution, University of Bern, Bern, Switzerland

<sup>2</sup>Swiss Institute of Bioinformatics, Quartier Sorge – Bâtiment Génopode, Lausanne, Switzerland

## Correspondence

Gerald Heckel, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland  
Email: gerald.heckel@iee.unibe.ch

## Funding information

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A\_149585 and 31003A\_176209

## Abstract

The natural host ranges of many viruses are restricted to very specific taxa. Little is known about the molecular barriers between species that lead to the establishment of this restriction or generally prevent virus emergence in new hosts. Here, we identify genomic polymorphisms in a natural rodent host associated with a strong genetic barrier to the transmission of European Tula orthohantavirus (TULV). We analysed the very abrupt spatial transition between two major phylogenetic clades in TULV across the comparatively much wider natural hybrid zone between evolutionary lineages of their reservoir host, the common vole (*Microtus arvalis*). Genomic scans of 79,225 single nucleotide polymorphisms (SNPs) in 323 TULV-infected host individuals detected 30 SNPs that were consistently associated with the TULV clades CEN.S or EST.S in two replicate sampling transects. Focusing the analysis on 199 voles with evidence of genomic admixture at the individual level (0.1–0.9) supported statistical significance for all 30 loci. Host genomic variation at these SNPs explained up to 37.6% of clade-specific TULV infections. Genes in the vicinity of associated SNPs include *SAHH*, *ITCH* and two members of the *Syngn* gene family, which are involved in functions related to immune response or membrane transport. This study demonstrates the relevance of natural hybrid zones as systems not only for studying processes of evolutionary divergence and speciation, but also for the detection of evolving genetic barriers for specialized parasites.

## KEYWORDS

genome-wide association study, genomic admixture, host-pathogen co-evolution, hybrid zone, *Microtus arvalis*, parasite

## 1 | INTRODUCTION

The evolution of viruses and their respective hosts is usually linked by various dependencies. Patterns of co-evolution can range from highly antagonistic evolutionary arms-races (Daugherty & Malik, 2012) to commensal or even mutualistic relationships (Roossinck,

2011), which in turn impose varying degrees of selective constraints on both the virus and the host (Hall et al., 2011). In the case of infectious diseases, viruses are predominantly under selective pressure to develop and maintain infection pathways to their hosts, as well as coping with their host's immune response (Pybus & Rambaut, 2009). Viruses generally face a tradeoff in fitness between specializing to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

maximize efficiency for a single host (specialists) and infecting multiple host species (generalists; Ciota & Kramer, 2010; Deardorff et al., 2011). In both cases the capability of infecting additional host species is a major opportunity for a virus to increase its infection range (Ciota & Kramer, 2010).

Such opportunities, however, may be limited as virus spillover infections often lead to evolutionary dead ends and transmissions are confined in the new host (Mollentze et al., 2020; Webby et al., 2004). Transmissions to new hosts that allow the virus to proliferate and continue spreading are overall infrequent and mediated by a variety of factors, most importantly the degree of relatedness between the hosts and the frequency of host contacts (Luis et al., 2015; Parrish et al., 2008; Streicker et al., 2010). In recent decades, cross-species transmission, particularly of RNA viruses, has resulted in a large number of disease emergences in humans and livestock, representing a major issue for health systems and economies (Hu et al., 2020; Jones et al., 2008; Lloyd-Smith et al., 2009; Mayer et al., 2017). Understanding the patterns of viral diversity in reservoir species and the identification of factors that affect cross-species virus emergence have become the focus of surveillance and prevention programmes (Lam et al., 2020; Olival et al., 2017).

The constitution of species barriers defining the infection range for a virus has been analysed mostly in the context of zoonotic disease emergence (Longdon et al., 2014; Olival et al., 2017; Plowright et al., 2017; Sironi et al., 2015). Phylogenetic analyses of viruses and their corresponding hosts have suggested that genetic factors play a primary role in determining the infection range of viruses (Faria et al., 2013; Longdon et al., 2014; Streicker et al., 2010). Overcoming a species barrier and establishing an efficient transmission chain in a new host may require multiple adaptive changes in the viral genome, representing a major evolutionary challenge (Kuiken et al., 2006; Parrish et al., 2008; Reperant et al., 2012; Simmonds et al., 2019).

The natural infection ranges of some viruses are restricted to closely related species (Streicker et al., 2010) or even to evolutionary lineages of the same host species (Drewes et al., 2017; Gryseels et al., 2017; Saxenhofer et al., 2019). Experimental studies comparing the infection efficiency of viruses in cells of different host species have allowed the identification of single host genes limiting the species range that certain viruses can infect, such as coronaviruses (van Doremalen et al., 2014), HIV-1 (Stremlau et al., 2004) or influenza A virus (Long et al., 2016). This suggests that relatively few intrinsic differences between otherwise genetically similar hosts can mediate strong barriers to virus emergence. The order of evolutionary events and the relative contribution of adaptive vs. neutral changes is difficult to determine in natural virus–host systems but it is possible that divergence of key genes involved in virus–host interactions could occur early in the process of host speciation (Saxenhofer et al., 2019). Still, it remains generally unclear how a host's genetic factors shape the selective environment of associated viruses and determine their infection range in natural populations.

Here we investigate intrinsic genetic factors delimiting the distribution ranges of Tula orthohantavirus (TULV) among evolutionary lineages of its reservoir host, the common vole (*Microtus arvalis*), using a genome-wide association study (GWAS) approach. In hantaviruses (family Hantaviridae; formerly Bunyaviridae), several immunity-related host genes have been associated with virus replication efficiency and persistence in their rodent reservoirs (see Charbonnel et al., 2014 for review). However, the focus was mostly on hantavirus infections in humans (Liu et al., 2009; Mäkelä et al., 2001; Martínez-Valdebenito et al., 2019; Müller et al., 2019; Mustonen et al., 1996; Wang et al., 2009) with relatively little information available on genes that regulate hantavirus resistance in reservoir species (Guivier, Galan, Salvador, et al., 2010; Rohfritsch et al., 2018). The consequences of candidate host genes for hantaviral infection ranges in nature remain mostly unclear (Dubois et al., 2017; Rohfritsch et al., 2018).

TULV is a single-stranded RNA virus with a three-segmented genome that is horizontally transmitted and causes asymptomatic chronic infections in the rodent reservoir host (Forbes et al., 2018; Vaheri et al., 2013). In Europe, the large-scale geographical distribution of highly diverged phylogenetic clades in TULV is partially associated with the distribution of morphologically cryptic evolutionary lineages in *M. arvalis* (Heckel et al., 2005; Schmidt et al., 2016). A detailed study of an intraspecific hybrid zone in the common vole demonstrated a very tight spatial association between the host lineages Central and Eastern and the virus clades Central South (TULV-CEN.S) and Eastern South (TULV-EST.S; Saxenhofer et al., 2019). A strong evolutionary barrier to effective TULV transmission in the hybrid zone operates at distances that individual voles can travel in very short time (Hahne et al., 2011; Saxenhofer et al., 2019; Schweizer et al., 2007) and where gene flow between host lineages is ongoing (Beysard & Heckel, 2014). Deep genomic divergence (17% sequence difference) and the absence of recombination or reassortment between TULV-CEN.S and TULV-EST.S indicate that these viral clades exceed the stage of speciation of the host lineages (Saxenhofer et al., 2019). The zone of contact between these TULV clades contains no physical barriers, such as rivers or changes in altitude, which might impede host movements or TULV transmission and is thus likely to be driven by intrinsic genetic factors that probably arose in the host lineages before secondary contact after post-glacial recolonization formed the hybrid zone (see Saxenhofer et al., 2019).

In this study, we take advantage of this natural system to interrogate the host genome for genetic polymorphisms associated with the sharp and probably selectively maintained distribution range limit of TULV clades that behave effectively as distinct viral species (see Saxenhofer et al., 2019). We focused our GWAS on the zone of natural admixture where the cosegregation of genetic polymorphisms distinguishing host lineages with the TULV clades is partially broken up by many generations of hybridization after secondary contact (Beysard & Heckel, 2014). Employing replicate sampling transects served to limit the potential of detecting geographically restricted associations and enabled the identification of genomic

polymorphisms and genes in *M. arvalis* that may contribute to confining the range of this European hantavirus.

## 2 | METHODS

### 2.1 | Sample collection

We sampled 547 common voles mostly from the direct contact area between the TULV-CEN.S and the TULV-EST.S clades and combined them for our analyses with TULV-infected samples available from Saxenhofer et al. (2019) (Table 1 and Table S1). Sampling was performed in two replicate transects (Porcelain and Bavaria; ~130 km apart) across the hybrid zone between the Central and Eastern common vole evolutionary lineages (Figure 1). Common voles were trapped using snap traps and stored at  $-20^{\circ}\text{C}$  immediately after collection. Rodent trapping was performed after ethical evaluation and approval by the Bernese cantonal commission on animal experimentation under permits BE-90/10 and BE-33/14. Host DNA was extracted according to a standard phenol-chloroform protocol. DNA concentration was measured for each sample using the Qubit dsDNA BR Assay Kit (Invitrogen) and DNA quality was evaluated on an agarose gel.

### 2.2 | TULV screening and clade assignment

Molecular screening for TULV infection was performed on the 238 adult voles of at least 20 g from 32 sampling sites among the 547 newly sampled individuals in the Porcelain and Bavaria transects. Voles of less than 20 g bodyweight were not assessed for TULV infection as young individuals are typically protected by maternal antibodies (Kallio et al., 2006; Schmidt et al., 2021). TULV infection was detected by amplifying a 540-nucleotide fragment of the nucleocapsid gene on the small TULV genome segment (S-segment) following the RT-PCR (reverse transcriptase PCR) assay described in Essbauer et al. (2006). RNA was extracted from lung tissue, and S-segment fragments of TULV-positive samples were sequenced as described by Schmidt et al. (2016) and Saxenhofer et al. (2019). A phylogenetic analysis was performed with MRBAYES version 3.2.6 (Ronquist et al., 2012) on the CIPRES platform (Miller et al., 2010) to assign TULV sequences to major evolutionary clades using reference sequences from Schmidt et al. (2016) and Saxenhofer et al. (2019) (Table S2). Metropolis-coupled Markov chain Monte Carlo (MCMC) sampling was performed for  $10^7$  generations in four independent

runs comprising four chains. We implemented reversible-jump sampling over the entire general time-reversible substitution model space (Huelsenbeck et al., 2004) and samples were recorded every  $10^3$  generations after discarding a burn-in fraction of 25%.

### 2.3 | Host genotyping

Genotyping by Sequencing (GBS; Elshire et al., 2011) was carried out for all TULV-infected common voles in both transects (Figure 1) on the Illumina NextSeq platform at Cornell University. Restriction enzymes *Pst*I and *Msp*I were used to generate the libraries in 96-well plates. Single nucleotide polymorphisms (SNPs) were identified and individuals genotyped simultaneously using the GBS version 2 pipeline (part of the TASSEL 5 software; Glaubitz et al., 2014), using a chromosome-level *Microtus arvalis* genome assembly as the reference sequence (Gouy, A., Wang, X., Neuenschwander, S., Schmid, E., Heckel, G. Excoffier, L.; unpublished data). Default parameters were used for calling genotypes, except that a minimum of five reads were required to identify a unique tag. Only bi-allelic SNPs were retained, and genotypes were only called if individuals had a read depth for that locus of at least five. After SNP calling, individuals with more than 50% of loci with missing data were removed. Then, loci were filtered out if they had a minor allele frequency (MAF) of less than 5%, more than 20% missing data or observed heterozygosity greater than 50%, which may indicate multiple paralogues merged into single loci (White et al., 2013). Three loci with complex indels were also removed. This filtering resulted in a data set of 323 TULV-infected individuals genotyped at 79,225 SNPs, which corresponds to an approximate density of 1 SNP per 50 kb of vole genome.

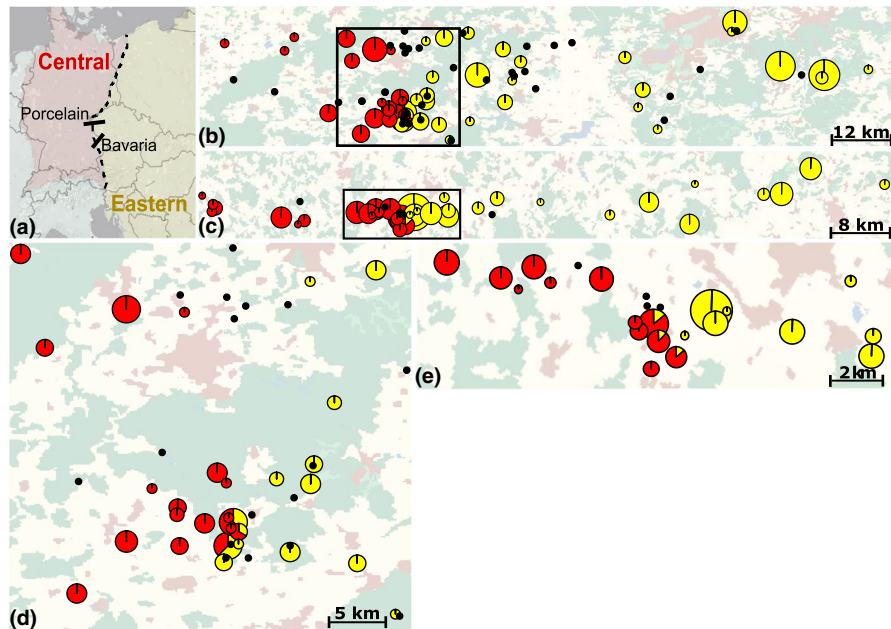
### 2.4 | Host population structure

Previous population structure analyses based on microsatellite markers in the hybrid zone within the Bavaria and Porcelain transects have already shown genetic admixture between the common vole lineages in these regions, supporting  $K = 2$  genetic clusters (Beysard & Heckel, 2014; Saxenhofer et al., 2019). To examine the extent to which these patterns of genetic admixture were also present in the infected individuals of our genomic data set, we analysed the genetic structure in the hybrid zone with the ADMIXTURE software (Alexander et al., 2009; Zhou et al., 2011). The analyses were performed for both transects separately using only TULV-infected individuals. We used initial cross-validation implemented in ADMIXTURE for  $K = 1-10$

	TULV-CEN.S	TULV-EST.S	Uninfected	Total
Porcelain transect	62 (53)	101 (80)	687 (0)	850 (133)
Bavaria transect	98 (89)	123 (101)	479 (0)	700 (190)
Both transects	160 (142)	224 (181)	1166 (0)	1550 (323)

TABLE 1 Number of common vole individuals from two sampling transects across the hybrid zone screened for Tula orthohantavirus

Note: The number of voles infected with the TULV-CEN.S or TULV-EST.S clades is given. The number of individuals included in the GWAS data set is indicated in parentheses.



**FIGURE 1** Sampling locations of common voles (*Microtus arvalis*) and Tula orthohantavirus. (a) Map of Central Europe with the approximate position of the hybrid zone (dashed line) between the Central (red) and Eastern (yellow) evolutionary lineages in *M. arvalis* and the location of the Porcelain and Bavaria sampling transects (full lines). (b,c) Complete view of the Porcelain (b) and Bavaria (c) transects containing the contacts between TULV-CEN.S (red) and TULV-EST.S (yellow) clades. Each circle represents a sampling location with its area proportional to the number of infected individuals. Black dots represent locations where no infected individuals were found. (d, e) Detailed view of the contact zone between TULV-CEN.S and TULV-EST.S for the Porcelain (d) and Bavaria (e) transect. The background in (b)–(e) shows settlements (red), forests (green), water bodies (blue) and cultivated land (yellow)

which supported  $K = 2$  for both transects. We then performed ADMIXTURE bootstrapping with 1000 replicates for each transect to establish genetic cluster membership of each infected individual.

## 2.5 | Genome-wide association analysis

A Genome-wide Efficient Mixed Model Association (GEMMA) analysis was conducted with version 0.98.1 of the software (Zhou & Stephens, 2012). As GEMMA requires a data set without missing data, we used LD-KNNI as implemented in TASSEL 5 (Glaubitz et al., 2014) for the imputation of missing SNPs based on the mean of the 10 closest related neighbours. Neighbours were determined by the 30 physically closest SNPs for each missing site and individual. Imputation was performed independently for the whole data set, as well as for each transect separately, and any SNPs still possessing missing data or a MAF of less than 5% were removed. Correcting for population structure within the data set was initially performed using covariates (principal components 1 and 2) and/or relatedness matrix (calculated via GEMMA version 0.98.1). However, this led to an overcorrection of the GWAS results due to the transition of virus phenotypes (TULV clades) within the host hybrid zone (Saxenhofer et al., 2019). Alternatively, we made use of the replicated nature of our data set to identify genetic polymorphisms in the host that are less dependent on the local population structure in each transect. We ran GEMMA's linear model on default parameters for the whole data set of 323 infected individuals, as well as separately for the Bavaria transect

(190 infected individuals) and Porcelain transect (133 infected individuals). Association strength of individual SNPs was estimated with GEMMA by calculating the Wald test  $p$ -value (Wald, 1945) and corrected for multiple testing with the Bonferroni method (Bonferroni, 1936). Only SNPs which were significantly ( $p < .05$ ) associated with clade-specific TULV infections across all three GWAS (Full data set, Bavaria transect only, Porcelain transect only) were considered for further analyses. To estimate the effect size of loci associated with TULV clades across both transects, a probit model was fitted on the genotype information of individual SNPs using the glm function followed by calculating Nagelkerke's  $R$ -squared (Nagelkerke, 1991) implemented in the FMSB package in R.

## 2.6 | Candidate genes

We analysed a flanking region of 100 kb up- and downstream (based on Brodie, Azaria, & Ofran, 2016; Laurie et al. 2007) of every significant SNP for genes that could serve as candidate genes for restricting the TULV infection range. We identified all genes in these flanking regions based on the presence of start or stop codons within the window around the significant SNP and determined their identity by using BLAST (Altschul et al., 1990) to detect their closest homologue in the well-annotated mouse reference genome. We performed an enrichment analysis of all candidate genes using the PANTHER 16.0 classification system as part of the Gene Ontology (GO) Consortium online resource (Ashburner et al., 2000; Consortium, 2018; Mi et al., 2018).

We used Fisher's exact test to find significantly enriched GO terms for both the biological processes and molecular functions of our candidate genes and corrected for multiple testing with the Bonferroni method (Bonferroni, 1936). For all genes with a GO term relevant for virus infection (i.e., related to immune response or membrane/vesicle transport), we examined further if there was published evidence that either they were involved in virus-related pathways, or a family member of a gene involved these pathways.

### 3 | RESULTS

#### 3.1 | TULV infections of common voles in the sampling area

The screening of 238 new common voles from 32 locations in the zone of admixture identified 78 TULV-infected individuals (33%), adding to the 306 infected individuals detected in the wider area by Saxenhofer et al. (2019) (Table 1 and Table S1; GenBank accession nos.: OK317919–OK317991 and OK356661–OK356665). Phylogenetic inference of TULV sequences from the region together with reference sequences assigned 160 sequences to the TULV-CEN.S and 224 to the TULV-EST.S clade (Figure S1). TULV-CEN.S infections were only detected in the western parts and TULV-EST.S only in the eastern parts of both transects (Figure 1). Common voles collected from the same population were typically infected with virus strains from the same TULV clade except for six locations (three in each transect) where TULV sequences from both clades were detected.

#### 3.2 | Genetic admixture within the host hybrid zone

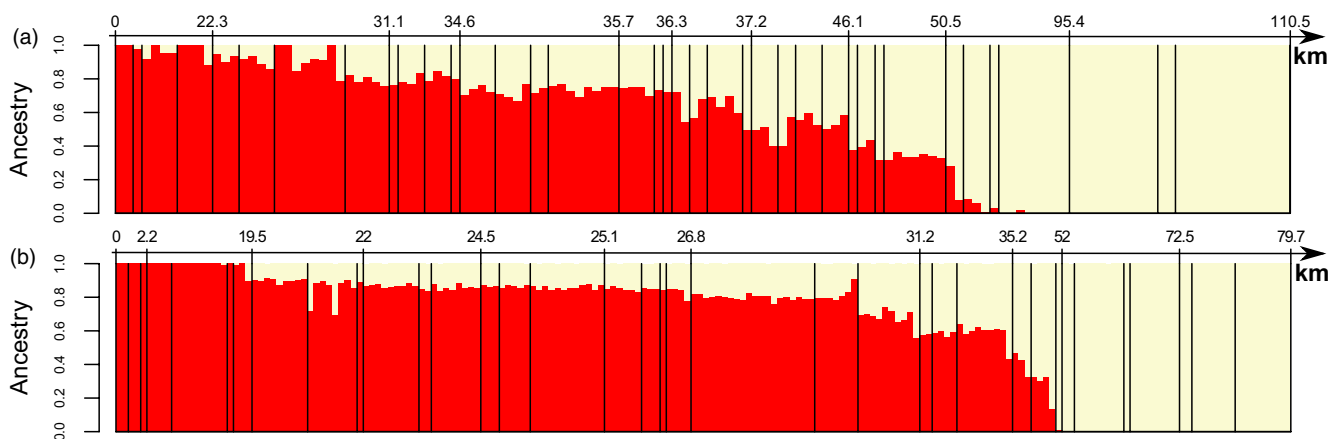
GBS data were successfully obtained for 323 TULV-infected individuals sampled close to the putative zone of admixture in the

transects. This included 190 individuals from the Bavaria transect and 133 individuals from the Porcelain transect (Table 1). The analysis of ancestry based on 79,225 SNPs demonstrated the presence of two major genetic clusters representing the Central and Eastern common vole evolutionary lineages (Beysard & Heckel, 2014; Heckel et al., 2005; Lischer et al., 2013) in both transects, with a clear transition between mostly pure populations at the transect ends (Figure 2). Of 323 infected individuals, 199 showed evidence of admixture with cluster memberships of 0.1–0.9 to either cluster.

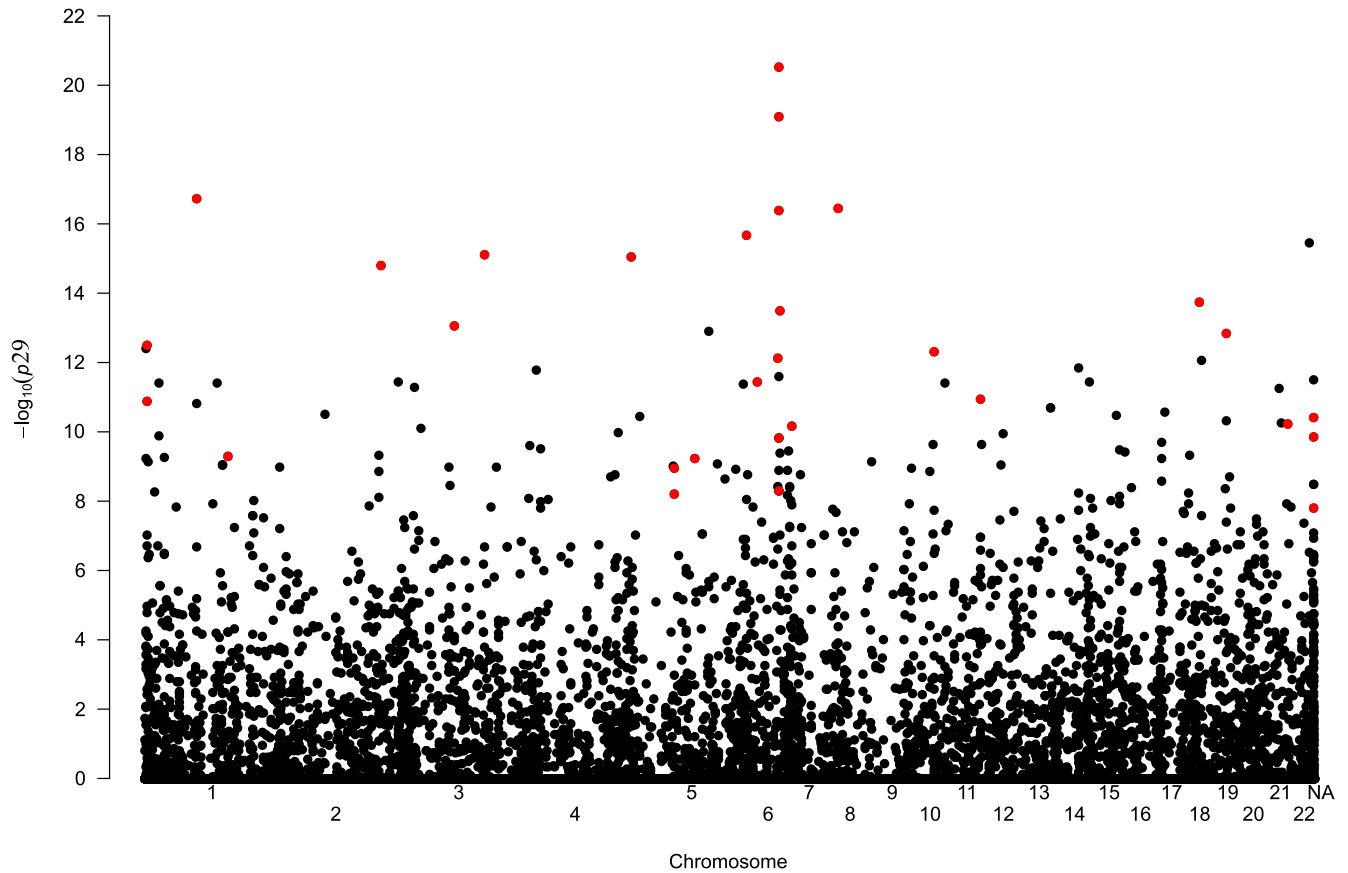
#### 3.3 | Genome-wide association across two replicate transects

Filtering of SNPs with a MAF of less than 5% and imputation of missing data for the GWAS provided a total of 60,471 SNPs for the full data set containing both transects, 57,461 SNPs for the Bavaria transect and 64,490 SNPs for the Porcelain transect. In total, 38,715 SNPs were shared between all three data sets. Among those, we found a total of 32 SNPs that showed a significant association with TULV clades in separate GWAS of the three individual data sets (Figure 3). We observed a particularly strong association on Chromosome 6 across a region of ~1 MB, containing five of the 32 significant SNPs and the two SNPs with the lowest *p*-values.

To assess the potential impact of imputation on the GWAS results, we removed any of the 323 individuals that had any missing data among the 32 significant SNPs, which resulted in 165 remaining individuals. This analysis showed that two of the 32 initially significant SNPs no longer passed the significance threshold for association with the TULV clades. To further assess the possible impact of spatial autocorrelation of vole lineages and virus clades, we ran a separate GWAS including only the 199 admixed individuals with cluster membership between 0.1 and 0.9. All of the 30 significant SNPs were detected again in this conservative analysis and no new SNPs were found to be significantly associated.



**FIGURE 2** Genetic structure of common voles infected with Tula orthohantavirus from the Porcelain (a) and Bavaria (b) transects. Each vertical bar represents the assignment of an individual to the Central (red) and Eastern (yellow) evolutionary lineage determined by the ADMIXTURE software based on 79,225 genetic markers. Black vertical lines separate individuals from different sampling locations. Distances from the western end of the transect are given in kilometres (km)



**FIGURE 3** Results of genome-wide association analyses of 323 common voles infected with Tula orthohantavirus (TULV) plotted for 38,715 SNP markers shared between alternative GWAS data sets. Red dots indicate the 30 SNPs which were significantly associated ( $p < .05$ ) with the infection by a specific TULV clade in each of four complementary data sets (all 323 individuals; Bavaria transect: 190 individuals; Porcelain transect: 133 individuals; no-imputation: 165 individuals; see text for details). The y-axis indicates the  $p$ -value after correction for multiple testing with the Bonferroni method

**TABLE 2** Candidate genes potentially involved in restricting the spatial range of TULV CEN.S and EST.S clades

Ensembl ID	Chromosome	Position	Gene	Distance	$p$ -Value	Function	$R^2$
ENSMUST00000054607	6	58206488	SAHH	85,231	2.88E-21	Methyltransferase regulation	.376
ENSMUST00000029126	6	58585061	ITCH	85,038	7.66E-20	Ubiquitination of target proteins	.371
ENSMUST00000007236	1	87846372	SNG3	0	1.86E-17	Membrane trafficking and transport	.328
ENSMUST00000097373	1	87846372	TSC2	55,654	1.86E-17	Anabolic metabolism regulation	.328
ENSMUST00000007216	3	145514324	AP2M1	29,732	7.56E-16	Clathrin-dependent endocytosis	.27
ENSMUST00000109790	6	56882593	ASXL1	58,637	7.13E-13	DNA and/or histone modification	.277
ENSMUST00000096482	11	43643687	SKP2	39,432	1.11E-11	Microvesicle membrane formation	.265
ENSMUST00000026649	1	4414266	SNG2	63,136	1.25E-11	Membrane trafficking and transport	.241
ENSMUST00000068714	5	66704109	SOS1	20,630	1.05E-09	Ras-guanine exchange factor	.222

Note: Only genes involved in virus-related immune response or membrane/vesicle transport are shown. A complete list of all 105 candidate genes is given in Table S2. Chromosome and position show the location of the SNP in the *Microtus arvalis* reference genome, while distance gives the number of base pairs of the candidate gene and the respective SNP. The  $p$ -value refers to the GWAS of all 323 infected individuals. The  $R^2$  column refers to the phenotypic variance in clade-specific TULV infections explained by the SNP.

### 3.4 | Immune-system related candidate genes

The examination of a 100-kb up- and downstream flanking region for each of the 30 SNPs significantly associated with TULV-CEN.S or TULV-EST.S infection revealed a total of 105 candidate genes with

homologues in the *Mus musculus* genome (Table S3). We detected no significant enrichment of GO-terms among these genes (all  $p > .3$ ). Among our candidate genes we identified nine with published evidence of involvement in viral response or infection pathways or that are a family member of such a gene (Table 2). The flanking regions of

the three SNPs with the lowest  $p$ -values in all GWAS contained four genes with known relevance in viral infection pathways: *S*-adenosyl homocysteinase (*SAHH*, SNP  $p = 2.88 \times 10^{-21}$ ), E3 ubiquitin-protein ligase Itchy (*ITCH*, SNP  $p = 7.66 \times 10^{-20}$ ) and the two genes Synaptogyrin-3 and Tuberous Sclerosis Complex 2 near the third locus (*Syngn03/TSC2*, SNP  $p = 1.8 \times 10^{-17}$ ; Table 2). Both *SAHH* and *ITCH* are located within the highly associated region on chromosome 6 between the two top scoring SNPs that are ~150 kb distant from each other. *Syngn03* corresponds to a region on chromosome 1, with the associated SNP located in the 4<sup>th</sup> exon of *Syngn03*. *TSC2* corresponds to a region of 33 kb that is 56 kb downstream of the very same SNP. For readability and because *Syngn03* is closer to the associated SNP, we refer to the *Syngn03/TSC2* locus in the following simply as the *Syngn03* locus. Fitting a probit regression model demonstrated an effect size of 37.6% for the *SAHH* locus, 37.1% for the *ITCH* locus and 32.8% for the *Syngn03* locus, which represents the proportion of clade-specific TULV infections explained by the associated host genotype (Table 2).

### 3.5 | Clade-specific TULV infections and associated host genotypes

Individual-level and spatial patterns of the SNPs at the candidate loci emphasize a strong association with infections with TULV-CEN.S or TULV-EST.S (Figure 4). For both *SAHH* and *ITCH*, most individuals homozygous for the alleles predominant in the Central or Eastern lineage were infected by the respective virus clade, while heterozygous hosts were infected by the two TULV clades at similar frequencies (Figure 4a,b). In contrast, the homozygous Central genotype at *Syngn03* was more strongly associated with restriction of TULV-EST.S infection than the Eastern genotype for TULV-CEN.S (Figure 4c). In the spatial context of the hybrid zone, the allele frequencies at the candidate loci showed a clear transition along both transects (Figure 4d-i). The spatial distribution of homozygous genotypes for the Central and Eastern *SAHH*- or *ITCH*-associated SNPs was highly consistent with the geographical range of TULV-CEN.S and TULV-EST.S (Figure 4d-g), thus potentially contributing to the limitations of the nonadapted virus clade. For *Syngn03*, the distribution of the homozygous Central genotype was highly consistent with the range of the TULV-CEN.S clade. TULV-EST.S was almost absent from hosts with homozygous Central lineage *Syngn03* genotypes, while the Eastern lineage allele extended far into the range of TULV-CEN.S (Figure 4h, i).

## 4 | DISCUSSION

Our analyses demonstrate very clear spatial patterns of local genomic admixture between common vole lineages and corroborate the extremely abrupt transition between parapatric TULV clades within the hybrid zone. Relatively few SNPs in the rodent host were consistently associated with infection by specific TULV clades across

replicate sampling transects, explaining up to 37.6% of the variance. Even at the very fine geographical scale of our analyses, the effects of genetic polymorphisms appear to limit the range of the two TULV clades, providing strong evidence for the evolution of an infection barrier within a host species.

The geographical distribution of pathogens within a host species is impacted by many factors, including landscape features (Guivier et al., 2011) or polymorphisms in resistance against pathogen infections (Alves et al., 2019; Magwire et al., 2012; Rohfritsch et al., 2018; White et al., 2011). Spatially narrow contacts between TULV-CEN.S and TULV-EST.S clades (Figure 1c,d) and an absence of potential dispersal barriers, such as rivers or forests (see also Saxenhofer et al., 2019), render a major effect of landscape connectivity (Gryseels et al., 2017), climatic differences (Gloria-Soria et al., 2017) or other extrinsic factors in the prevention of virus transmission unlikely.

The infection range limits of the TULV clades are probably caused by differences in TULV fitness in the two host lineages because hantavirus infection is largely asymptomatic in common voles and other reservoir hosts (Forbes et al., 2018). The detection of two individuals infected by both TULV clades (Hiltbrunner & Heckel, 2020; Saxenhofer et al., 2019) rules out a full resistance phenotype. Instead, we suggest that host gene polymorphisms that arose in allopatric glacial refugia of the evolutionary lineages of *Microtus arvalis* may act as a partial transmission barrier to the individual TULV clades (see Saxenhofer et al., 2019). Maintenance of a local transmission chain is essential for virus persistence in a host population (Biek & Real, 2010; Forbes et al., 2018; Kuiken et al., 2006). Less efficient virus transmission among individuals of the opposite host lineage (e.g., due to longer transmission intervals) may form part of the barrier for differentially adapted TULV clades (Saxenhofer et al., 2019). The very narrow contact area of TULV-CEN.S and TULV-EST.S clades (Figure 1) implies virus transmission primarily among adjacent vole populations (see also Saxenhofer et al., 2017; Schmidt-Chanasit et al., 2010; Weber de Melo et al., 2015). Direct competition for susceptible hosts among viruses from different clades could only occur in mixed populations or when viruses would be transferred across the clade contact by a dispersing vole host. TULV prevalence in local populations varies between 0% and 45% (Maas et al., 2017; Schmidt et al., 2016, 2021; Schmidt-Chanasit et al., 2010), which would potentially provide enough noninfected local hosts for transmission across the clade contact. However, the rarity of host populations with both TULV clades and double-infected individuals (see Saxenhofer et al., 2019) suggests only limited potential for direct within-host competition between clades to contribute to establishing the range limits.

The GWAS and the spatial context show a strong association between particular SNPs in the host and clade-specific TULV infections. In principle, a narrow and consistent contact zone of the TULV clades in both transects suggests the involvement of only a single gene of major effect or a small number of tightly linked genes in restricting the infection range of the TULV clades. The transmission barrier might be mediated by a matching-allele type of virus-host interaction (Dybdahl et al., 2014) where changes in allele frequencies

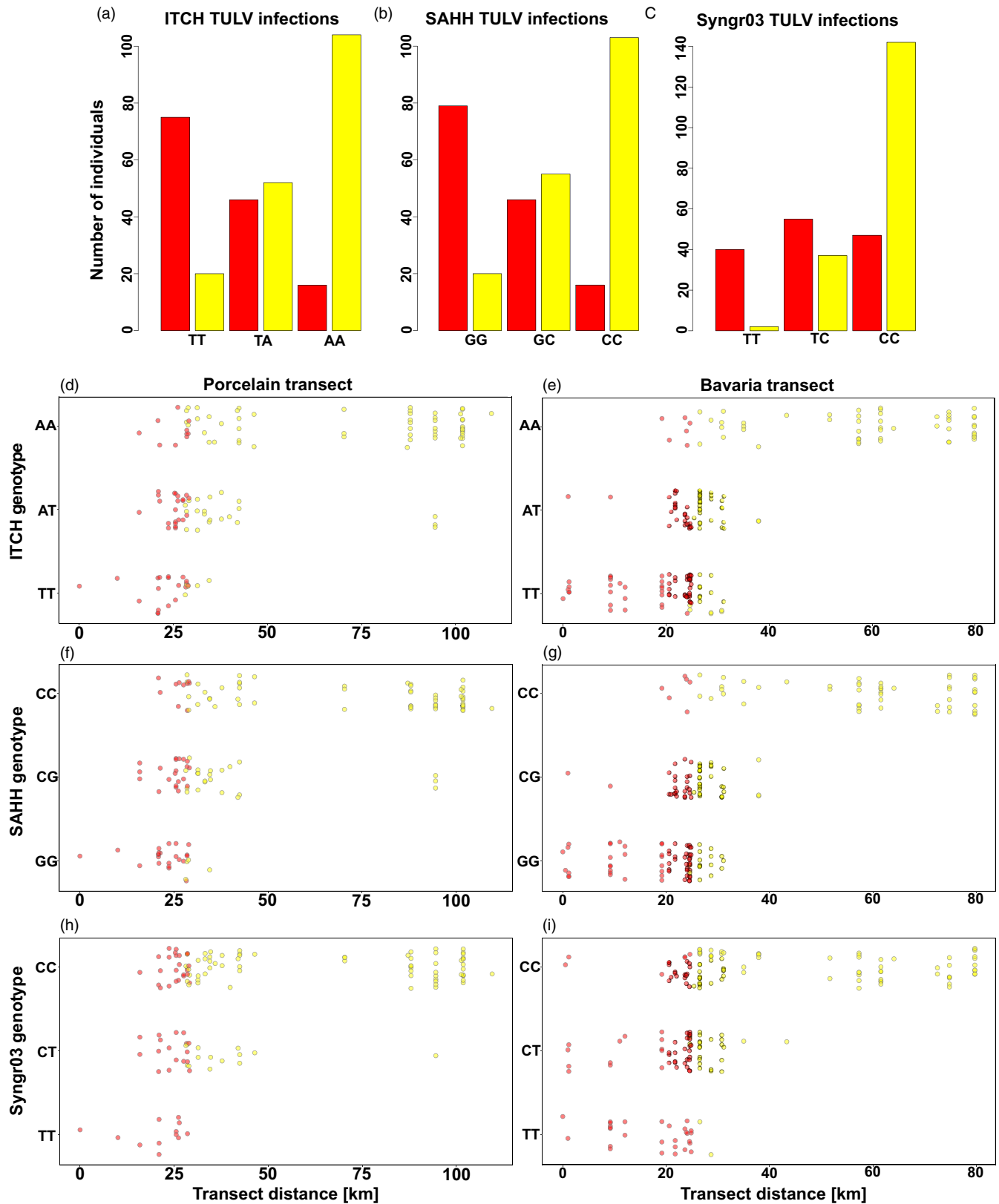


FIGURE 4 Genotypes at SNPs associated with candidate loci *SAHH* (a), *ITCH* (b) and *Syng03* (c) and clade-specific TULV infections. The number of individuals infected with TULV-CEN.S (red) and TULV-EST.S (yellow) is given for homozygous Central lineage genotypes, heterozygotes and homozygous Eastern lineage genotypes. (d–i) Spatial distributions of genotypes along the Porcelain (d, f, h) and Bavaria (e, g, i) transects. Vertical jitter was added for each genotype class for better visibility of individual data points



of these host genes along the transects represent a major barrier for transmission of differentially adapted TULV clades (Saxenhofer et al., 2019). Single host factors that restrict the infection range and impede virus emergence have been described so far predominantly between relatively distantly related host species with highly divergent genetic and ecological backgrounds (Long et al., 2016; Stremlau et al., 2004; van Doremalen et al., 2014). In hantaviruses, studies on the interactions with the host immune system have revealed several host factors involved in virus infection, persistence and replication (Charbonnel et al., 2014; Easterbrook et al., 2007; Guivier, Galan, Male, et al., 2010; Martínez-Valdebenito et al., 2019; Müller et al., 2019). However, infection barriers for hantaviruses have been characterized mostly in the context of human infections (Liu et al., 2009; Mäkelä et al., 2001; Martínez-Valdebenito et al., 2019; Mustonen et al., 1996; Wang et al., 2009). The rodent host genes identified in this study may present new candidates involved in restricting viral infection ranges and in limiting host shifts of hantaviruses between closely related species.

Among the candidate genes, *SAHH*, *ITCH* and *Syngn03* as well as its close neighbour *TSC2* stand out as particularly promising candidates (Table 2). *SAHH* enables methylation of a variety of both DNA and RNA motifs, which is crucial for the replication of several virus species such as vaccinia virus, yellow fever and vesicular stomatitis (Borchardt et al., 1984; Tseng et al., 1989). *ITCH* is an E3 ubiquitin ligase and a common target for viral hijacking to recruit host proteins necessary for virus development and budding in multiple viruses including Ebola (Han et al., 2016), Herpes (Koshizuka et al., 2018) and Influenza A. *Syngn03* is a member of the Synaptogyrin gene family, which is involved in vesicular transport and endo- and exocytosis crucial for virus replication (Sessions et al., 2009; Sun et al., 2016; Walker et al., 2018). Among our candidate genes is also *Syngn02* with a spatial allele distribution resembling that of *Syngn03* (Table 2 and Figure S2). Finally, *TSC2* is involved in anabolic metabolism in cells and a crucial target of the Human Cytomegalovirus UL38 protein, facilitating efficient viral replication (Bai et al., 2015; Moorman et al., 2008). None of these genes has been implicated in hantavirus–host interactions so far, but comparable genomic analyses at similar scale are lacking for other natural systems except for an investigation into the genetics of bank vole tolerance to Puumala hantavirus (PUUV; Rohfritsch et al., 2018).

Further characterization of our candidate genes in the hantavirus context is necessary to identify potential interactions with the investigated TULV clades. Signatures of positive selection on a protein binding signal peptide region of the TULV M-Segment (Saxenhofer et al., 2019) indicate adaptation towards specific host genes. Many years of research on human-pathogenic hantaviruses and others have shown the difficulty in identifying key genes or proteins conferring resistance or differences in susceptibility to infection in dead-end hosts such as humans but also in reservoir hosts (Charbonnel et al., 2014; Martínez-Valdebenito et al., 2019; Müller et al., 2019; Rohfritsch et al., 2018). The well-defined ecological, evolutionary and spatial context of this natural system analysed here holds the potential for using even more refined genomic approaches

for complementing and prioritizing the list of candidate host factors (see Atkinson et al., 2021; Kwok et al., 2021), which may then feed back into research on pathogenic systems.

Cross-species transmission of viruses is related to ecological and evolutionary diversity in many mammalian and avian host taxa (Allen et al., 2017; Mollentze & Streicker, 2020), and intrinsic species barriers for viruses may arise mostly as a by-product of host diversification (Cuypers et al., 2020; de Bellocq et al., 2015; Gryseels et al., 2017; Martinů et al., 2020). In the TULV system, Central and Eastern are the evolutionarily closest lineages in the common vole (Beysard & Heckel, 2014; Lischer et al., 2013; Sutter et al., 2013) and the restrictions in the range of TULV-CEN.S and TULV-EST.S have probably arisen only after the host lineages have diverged in allopatry and established secondary contact in the hybrid zone (Saxenhofer et al., 2019). It would thus be interesting to examine the factors and dynamics of the range limits between (hanta-)virus clades within single host species or lineages further. These could represent transient situations related to range expansion processes of viruses within hosts or be maintained in place by selection related to host polymorphisms (see also Martinů et al., 2020; Theodosopoulos et al., 2019). The TULV system potentially allows a direct test of these alternatives in the future with multiple additional TULV clade contacts in Central Europe, both within and between evolutionary lineages of its rodent host (see Saxenhofer et al., 2019; Schmidt et al., 2016; Schlegel et al., 2012). Studying the genomic barriers to virus transmission across TULV contact zones of different evolutionary divergence analogous to their hosts (Beysard & Heckel, 2014; Beysard et al., 2015) may offer insights into the evolutionary mechanics that drive the divergence of hantaviruses and potentially even the generation of new virus species. For closely related and human-pathogenic PUUV, many virus clades have been described across Europe and partially associated with particular evolutionary lineages in the bank vole host and the regional absence of the disease in humans (Drewes et al., 2017), but the direct link to genomic polymorphisms in the rodent host has not been established.

## 5 | CONCLUSIONS

Extant pathogen populations in animal reservoirs are the most common source of outbreaks of infectious diseases in humans and livestock (Jones et al., 2008; Mollentze & Streicker, 2020). Understanding the factors that affect cross-species virus emergence is the focus of research and prevention programmes but the combat against outbreaks is often impeded by very limited knowledge about the reservoir hosts (Groseth et al., 2007; Shi et al., 2018). As RNA virus evolution is mostly driven by patterns of co-evolution and co-divergence with host taxa (Lin et al., 2012; Mélade et al., 2016; Switzer et al., 2005), combining evolutionary analyses of host and virus divergence directly may allow insights that are difficult to obtain in the nonequilibrium situation of disease outbreaks (Cuypers et al., 2020; Schneider et al., 2021). The explicit consideration of the spatial context of the association of

many hantaviruses with their host taxa may be particularly beneficial for clarifying the relationships and succession of events in evolutionary adaptation or host-species switches (see also de Bellocq et al., 2015, 2018; Bennett et al., 2014; Cuyppers et al., 2020; Gryseels et al., 2017; Guo et al., 2013; Martinů et al., 2020; Saxenhofer et al., 2017, 2019; Worobey et al., 2010). Our analyses demonstrate that detailed examination of natural hybrid zones between host taxa—or more generally admixture between hosts—has the potential to aid in identifying not only genetic polymorphisms relevant for developing and maintaining species barriers among the hosts but also those loci that contribute to these processes in tightly associated parasites.

## ACKNOWLEDGEMENTS

We thank Susanne Tellenbach for laboratory support, and Melanie Hiltbrunner, Susanne Hein, Till Meyer zu Westram, Marco Felber, Johan Kantsjö, Daniel Abbühl and Remco Folkertsma for assistance with sample collection. G.H. would like to thank Rainer G. Ulrich for long-term collaboration on hantaviruses that provided the foundation for developing the present work. Constructive comments of the reviewers helped to improve the manuscript. This study was supported by grants 31003A\_149585 and 176209 from the Swiss National Science Foundation to G.H.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## AUTHOR CONTRIBUTIONS

G.H. conceptualized the study. M.S., G.H. and T.A.W. performed fieldwork and generated the data for this study. M.S., A.L. and T.A.W. analysed the data. A.L., M.S. and G.H. wrote the original draft. All authors contributed to the reviewing and editing of the final manuscript.

## DATA AVAILABILITY STATEMENT

The Tula virus S-segment sequence data are openly available in GenBank under accession nos. OK317919–OK317991 and OK356661–OK356665. Raw GBS sequence data are available at the NCBI Sequence Read Archive (SRA) at <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA767008>. The keyfile for matching GBS reads to individual samples together with the GWAS results are available at the Dryad repository under <https://doi.org/10.5061/dryad.5dv41ns6p>.

## ORCID

Anton Labutin  <https://orcid.org/0000-0002-5706-2822>

Gerald Heckel  <https://orcid.org/0000-0002-0162-323X>

## REFERENCES

- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Allen, T., Murray, K. A., Zambrana-Torrel, C., Morse, S. S., Rondinini, C., Di Marco, M., Breit, N., Olival, K. J., & Daszak, P. (2017). Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications*, 8(1), 1–10. <https://doi.org/10.1038/s41467-017-00923-8>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Alves, J. M., Carneiro, M., Cheng, J. Y., de Matos, A. L., Rahman, M. M., Loog, L., & Jiggins, F. M. (2019). Parallel adaptation of rabbit populations to myxoma virus. *Science*, 363(6433), 1319–1326
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25. <https://doi.org/10.1038/75556>
- Atkinson, E. G., Maihofer, A. X., Kanai, M., Martin, A. R., Karczewski, K. J., Santoro, M. L., Ulirsch, J. C., Kamatani, Y., Okada, Y., Finucane, H. K., Koenen, K. C., Nievergelt, C. M., Daly, M. J., & Neale, B. M. (2021). Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power. *Nature Genetics*, 53(2), 195–204. <https://doi.org/10.1038/s41588-020-00766-y>
- Bai, Y., Xuan, B., Liu, H., Zhong, J., Yu, D., & Qian, Z. (2015). Tuberosclerosis complex protein 2-independent activation of mTORC1 by human cytomegalovirus pUL38. *Journal of Virology*, 89(15), 7625–7635. <https://doi.org/10.1128/JVI.01027-15>
- Bennett, S. N., Gu, S. H., Kang, H. J., Arai, S., & Yanagihara, R. (2014). Reconstructing the evolutionary origins and phylogeography of hantaviruses. *Trends in Microbiology*, 22(8), 473–482. <https://doi.org/10.1016/j.tim.2014.04.008>
- Beysard, M., & Heckel, G. (2014). Structure and dynamics of hybrid zones at different stages of speciation in the common vole (*Microtus arvalis*). *Molecular Ecology*, 23(3), 673–687.
- Beysard, M., Krebs-Wheaton, R., & Heckel, G. (2015). Tracing reinforcement through asymmetrical partner preference in the European common vole *Microtus arvalis*. *BMC Evolutionary Biology*, 15, 170.
- Biek, R., & Real, L. A. (2010). The landscape genetics of infectious disease emergence and spread. *Molecular Ecology*, 19(17), 3515–3531. <https://doi.org/10.1111/j.1365-294X.2010.04679.x>
- Brodie, A., Azaria, J. R., & Ofran, Y. (2016). How far from the SNP may the causative genes be? *Nucleic Acids Research*, 44, 6046–6054. <https://doi.org/10.1093/nar/gkw500>
- Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni Del R Istituto Superiore Di Scienze Economiche E Commerciali Di Firenze*, 8, 3–62.
- Borchardt, R. T., Keller, B. T., & Patel-Thombre, U. (1984). Neplanocin A. A potent inhibitor of S-adenosylhomocysteine hydrolase and of vaccinia virus multiplication in mouse L929 cells. *Journal of Biological Chemistry*, 259(7), 4353–4358.
- Charbonnel, N., Pagès, M., Sironen, T., Henttonen, H., Vapalahti, O., Mustonen, J., & Vaheri, A. (2014). Immunogenetic factors affecting susceptibility of humans and rodents to hantaviruses and the clinical course of hantaviral disease in humans. *Viruses*, 6(5), 2214–2241. <https://doi.org/10.3390/v6052214>
- Ciota, A. T., & Kramer, L. D. (2010). Insights into arbovirus evolution and adaptation from experimental studies. *Viruses*, 2(12), 2594–2617. <https://doi.org/10.3390/v2122594>
- Consortium, G. O. (2018). The Gene Ontology resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1), D330–D338.
- Cuyppers, L. N., Baird, S. J. E., Hánová, A., Locus, T., Katakweba, A. S., Gryseels, S., Bryja, J., Leirs, H., & Goüy de Bellocq, J. (2020). Three arenaviruses in three subspecific natal multimammate mouse taxa in Tanzania: same host specificity, but different spatial genetic structure? *Virus Evolution*, 6(2), veaa039. <https://doi.org/10.1093/ve/veaa039>

- Daugherty, M. D., & Malik, H. S. (2012). Rules of engagement: Molecular insights from host-virus arms races. *Annual Review of Genetics*, *46*, 677–700. <https://doi.org/10.1146/annurev-genet-110711-155522>
- de Bellocq, J. G., Baird, S. J. E., Albrechtová, J., Sobeková, K., & Piálek, J. (2015). Murine cytomegalovirus is not restricted to the house mouse *Mus musculus domesticus*: prevalence and genetic diversity in the European house mouse hybrid zone. *Journal of Virology*, *89*(1), 406–414. <https://doi.org/10.1128/JVI.02466-14>
- de Bellocq, J. G., Wasimuddin, Ribas, A., Bryja, J., Piálek, J., & Baird, S. J. E. (2018). Holobiont suture zones: Parasite evidence across the European house mouse hybrid zone. *Molecular Ecology*, *27*, 5214–5227.
- Deardorff, E. R., Fitzpatrick, K. A., Jerzak, G. V. S., Shi, P.-Y., Kramer, L. D., & Ebel, G. D. (2011). West Nile virus experimental evolution in vivo and the trade-off hypothesis. *PLoS Pathogens*, *7*(11), e1002335. <https://doi.org/10.1371/journal.ppat.1002335>
- Drewes, S., Ali, H. S., Saxenhofer, M., Rosenfeld, U. M., Binder, F., Cuyppers, F., Schlegel, M., Röhrs, S., Heckel, G., & Ulrich, R. G. (2017). Host-associated absence of human Puumala virus infections in Northern and Eastern Germany. *Emerging Infectious Diseases*, *23*(1), 83. <https://doi.org/10.3201/eid2301.160224>
- Dubois, A., Galan, M., Cosson, J.-F., Gauffre, B., Henttonen, H., Niemimaa, J., Razzauti, M., Voutilainen, L., Vitalis, R., Guivier, E., & Charbonnel, N. (2017). Microevolution of bank voles (*Myodes glareolus*) at neutral and immune-related genes during multiannual dynamic cycles: Consequences for Puumala hantavirus epidemiology. *Infection, Genetics and Evolution*, *49*, 318–329. <https://doi.org/10.1016/j.meegid.2016.12.007>
- Dybdahl, M. F., Jenkins, C. E., & Nuismer, S. L. (2014). Identifying the molecular basis of host-parasite coevolution: Merging models and mechanisms. *The American Naturalist*, *184*(1), 1–13. <https://doi.org/10.1086/676591>
- Easterbrook, J. D., Zink, M. C., & Klein, S. L. (2007). Regulatory T cells enhance persistence of the zoonotic pathogen Seoul virus in its reservoir host. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(39), 15502–15507. <https://doi.org/10.1073/pnas.0707453104>
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, *6*(5), e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Essbauer, S., Schmidt, J., Conraths, F. J., Friedrich, R., Koch, J., Hautmann, W., Pfeffer, M., Wölfel, R., Finke, J., Dobler, G., & Ulrich, R. (2006). A new Puumala hantavirus subtype in rodents associated with an outbreak of *Nephropathia epidemica* in South-East Germany in 2004. *Epidemiology & Infection*, *134*(6), 1333–1344. <https://doi.org/10.1017/S0950268806006170>
- Faria, N. R., Suchard, M. A., Rambaut, A., Streicker, D. G., & Lemey, P. (2013). Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1614), 20120196. <https://doi.org/10.1098/rstb.2012.0196>
- Forbes, K. M., Sironen, T., & Plyusnin, A. (2018). Hantavirus maintenance and transmission in reservoir host populations. *Current Opinion in Virology*, *28*, 1–6. <https://doi.org/10.1016/j.coviro.2017.09.003>
- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One*, *9*(2), e90346. <https://doi.org/10.1371/journal.pone.0090346>
- Gloria-Soria, A., Armstrong, P. M., Powell, J. R., & Turner, P. E. (2017). Infection rate of *Aedes aegypti* mosquitoes with dengue virus depends on the interaction between temperature and mosquito genotype. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1864), 20171506.
- Groseth, A., Feldmann, H., & Strong, J. E. (2007). The ecology of Ebola virus. *Trends in Microbiology*, *15*(9), 408–416. <https://doi.org/10.1016/j.tim.2007.08.001>
- Gryseels, S., Baird, S. J. E., Borremans, B., Makundi, R., Leirs, H., & de Bellocq, J. G. (2017). When viruses don't go viral: The importance of host phylogeographic structure in the spatial spread of arenaviruses. *PLoS Pathogens*, *13*(1), e1006073. <https://doi.org/10.1371/journal.ppat.1006073>
- Guivier, E., Galan, M., Chaval, Y., Xuéreb, A., Ribas salvador, A., Poulle, M.-L., Voutilainen, L., Henttonen, H., Charbonnel, N., & Cosson, J. F. (2011). Landscape genetics highlights the role of bank vole metapopulation dynamics in the epidemiology of Puumala hantavirus. *Molecular Ecology*, *20*(17), 3569–3583. <https://doi.org/10.1111/j.1365-294X.2011.05199.x>
- Guivier, E., Galan, M., Male, P.-J.G., Kallio, E. R., Voutilainen, L., Henttonen, H., Olsson, G. E., Lundkvist, A., Tersago, K., Augot, D., Cosson, J.-F., & Charbonnel, N. (2010). Associations between MHC genes and Puumala virus infection in *Myodes glareolus* are detected in wild populations, but not from experimental infection data. *Journal of General Virology*, *91*(10), 2507–2512. <https://doi.org/10.1099/vir.0.021600-0>
- Guivier, E., Galan, M., Salvador, A. R., Xuéreb, A., Chaval, Y., Olsson, G. E., Essbauer, S., Henttonen, H., Voutilainen, L., Cosson, J.-F., & Charbonnel, N. (2010). Tnf- $\alpha$  expression and promoter sequences reflect the balance of tolerance/resistance to Puumala hantavirus infection in European bank vole populations. *Infection, Genetics and Evolution*, *10*(8), 1208–1217. <https://doi.org/10.1016/j.meegid.2010.07.022>
- Guo, W.-P., Lin, X.-D., Wang, W., Tian, J.-H., Cong, M.-L., Zhang, H.-L., Wang, M.-R., Zhou, R.-H., Wang, J.-B., Li, M.-H., Xu, J., Holmes, E. C., & Zhang, Y.-Z. (2013). Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. *PLoS Pathogens*, *9*(2), e1003159. <https://doi.org/10.1371/journal.ppat.1003159>
- Hahne, J., Jenkins, T., Halle, S., & Heckel, G. (2011). Establishment success and resulting fitness consequences for vole dispersers. *Oikos*, *120*(1), 95–105. <https://doi.org/10.1111/j.1600-0706.2010.18589.x>
- Hall, A. R., Scanlan, P. D., Morgan, A. D., & Buckling, A. (2011). Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecology Letters*, *14*(7), 635–642. <https://doi.org/10.1111/j.1461-0248.2011.01624.x>
- Han, Z., Sagum, C. A., Bedford, M. T., Sidhu, S. S., Sudol, M., & Harty, R. N. (2016). ITCH E3 ubiquitin ligase interacts with Ebola virus VP40 to regulate budding. *Journal of Virology*, *90*(20), 9163–9171. <https://doi.org/10.1128/JVI.01078-16>
- Heckel, G., Burri, R., Fink, S., Desmet, J.-F., & Excoffier, L. (2005). Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution*, *59*(10), 2231–2242. <https://doi.org/10.1111/j.0014-3820.2005.tb00931.x>
- Hiltbrunner, M., & Heckel, G. (2020). Assessing genome-wide diversity in European hantaviruses through sequence capture from natural host samples. *Viruses*, *12*(7), 749. <https://doi.org/10.3390/v12070749>
- Hu, B., Guo, H., Zhou, P., & Shi, Z.-L. (2020). Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, *19*(3), 141–154. <https://doi.org/10.1038/s41579-020-00459-7>
- Huelsenbeck, J. P., Larget, B., & Alfaro, M. E. (2004). Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution*, *21*(6), 1123–1133. <https://doi.org/10.1093/molbev/msh123>
- Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990. <https://doi.org/10.1038/nature06536>
- Kallio, E. R., Poikonen, A., Vaheri, A., Vapalahti, O., Henttonen, H., Koskela, E., & Mappes, T. (2006). Maternal antibodies postpone hantavirus infection and enhance individual breeding success. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1602), 2771–2776.
- Koshizuka, T., Kobayashi, T., Ishioka, K., & Suzutani, T. (2018). Herpesviruses possess conserved proteins for interaction with

- Nedd4 family ubiquitin E3 ligases. *Scientific Reports*, 8(1), 4447. <https://doi.org/10.1038/s41598-018-22682-2>
- Kuiken, T., Holmes, E. C., McCauley, J., Rimmelzwaan, G. F., Williams, C. S., & Grenfell, B. T. (2006). Host species barriers to influenza virus infections. *Science*, 312(5772), 394–397.
- Kwok, A. J., Mentzer, A., & Knight, J. C. (2021). Host genetics and infectious disease: new tools, insights and translational opportunities. *Nature Reviews Genetics*, 22(3), 1–17. <https://doi.org/10.1038/s41576-020-00297-6>
- Lam, T.-T.-Y., Jia, N., Zhang, Y.-W., Shum, M.-H.-H., Jiang, J.-F., Zhu, H.-C., & Cao, W.-C. (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, 583(7815), 282–285.
- Laurie, C. C., Nickerson, D. A., Anderson, A. D., Weir, B. S., Livingston, R. J., Dean, M. D., Smith, K. L., Schadt, E. E., & Nachman, M. W. (2007). Linkage disequilibrium in wild mice. *PLoS Genetics*, 3, e144. <https://doi.org/10.1371/journal.pgen.0030144>
- Lin, X.-D., Wang, W., Guo, W.-P., Zhang, X.-H., Xing, J.-G., Chen, S.-Z., Li, M.-H., Chen, Y., Xu, J., Plyusnin, A., & Zhang, Y.-Z. (2012). Cross-species transmission in the speciation of the currently known murinae-associated hantaviruses. *Journal of Virology*, 86(20), 11171–11182. <https://doi.org/10.1128/JVI.00021-12>
- Lischer, H. E. L., Excoffier, L., & Heckel, G. (2013). Ignoring heterozygous sites biases phylogenomic estimates of divergence times: Implications for the evolutionary history of *Microtus voles*. *Molecular Biology and Evolution*, 31(4), 817–831. <https://doi.org/10.1093/molbev/mst271>
- Liu, Z., Gao, M., Han, Q., Lou, S., & Fang, J. (2009). Platelet glycoprotein IIb/IIIa (HPA-1 and HPA-3) polymorphisms in patients with hemorrhagic fever with renal syndrome. *Human Immunology*, 70(6), 452–456. <https://doi.org/10.1016/j.humimm.2009.03.009>
- Lloyd-Smith, J. O., George, D., Pepin, K. M., Pitzer, V. E., Pulliam, J. R. C., Dobson, A. P., & Grenfell, B. T. (2009). Epidemic dynamics at the human-animal interface. *Science*, 326(5958), 1362–1367.
- Long, J. S., Giotis, E. S., Moncorgé, O., Frise, R., Mistry, B., James, J., Morisson, M., Iqbal, M., Vignal, A., Skinner, M. A., & Barclay, W. S. (2016). Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature*, 529(7584), 101. <https://doi.org/10.1038/nature16474>
- Longdon, B., Brockhurst, M. A., Russell, C. A., Welch, J. J., & Jiggins, F. M. (2014). The evolution and genetics of virus host shifts. *PLoS Pathogens*, 10(11), e1004395. <https://doi.org/10.1371/journal.ppat.1004395>
- Luis, A. D., O'Shea, T. J., Hayman, D. T. S., Wood, J. L. N., Cunningham, A. A., Gilbert, A. T., Mills, J. N., & Webb, C. T. (2015). Network analysis of host-virus communities in bats and rodents reveals determinants of cross-species transmission. *Ecology Letters*, 18(11), 1153–1162. <https://doi.org/10.1111/ele.12491>
- Maas, M., De Vries, A., Van Roon, A., Takumi, K., van der Giessen, J., & Rockx, B. (2017). High prevalence of Tula hantavirus in common voles in the Netherlands. *Vector-Borne and Zoonotic Diseases*, 17(3), 200–205. <https://doi.org/10.1089/vbz.2016.1995>
- Magwire, M. M., Fabian, D. K., Schweyen, H., Cao, C., Longdon, B., Bayer, F., & Jiggins, F. M. (2012). Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. *PLoS Genetics*, 8(11), e1003057. <https://doi.org/10.1371/journal.pgen.1003057>
- Mäkelä, S., Hurme, M., Ala-Houhala, I., Mustonen, J., Koivisto, A.-M., Partanen, J., Vapalahti, O., Vaheri, A., & Pasternack, A. (2001). Polymorphism of the cytokine genes in hospitalized patients with Puumala hantavirus infection. *Nephrology Dialysis Transplantation*, 16(7), 1368–1373. <https://doi.org/10.1093/ndt/16.7.1368>
- Martínez-Valdebenito, C., Angulo, J., Le Corre, N., Marco, C., Vial, C., Miquel, J., Cerda, J., Mertz, G., Vial, P., Lopez-Lastra, M., & Ferrés, M. (2019). A single-nucleotide polymorphism of  $\alpha\beta 3$  integrin is associated with the Andes virus infection susceptibility. *Viruses*, 11(2), <https://doi.org/10.3390/v11020169>
- Martinů, J., Štefka, J., Poosakkannu, A., & Hypša, V. (2020). "Parasite turnover zone" at secondary contact: A new pattern in host-parasite population genetics. *Molecular Ecology*, 29(23), 4653–4664. <https://doi.org/10.1111/mec.15653>
- Mayer, S. V., Tesh, R. B., & Vasilakis, N. (2017). The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers. *Acta Tropica*, 166, 155–163. <https://doi.org/10.1016/j.actatropica.2016.11.020>
- Mélade, J., Wieseke, N., Ramasindrazana, B., Flores, O., Lagadec, E., Gomard, Y., Goodman, S. M., Dellagi, K., & Pascalis, H. (2016). An eco-epidemiological study of Morbilli-related paramyxovirus infection in Madagascar bats reveals host-switching as the dominant macro-evolutionary mechanism. *Scientific Reports*, 6(1), 1–12. <https://doi.org/10.1038/srep23752>
- Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2018). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, 47(D1), D419–D426.
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE), 1–8.
- Mollentze, N., & Streicker, D. G. (2020). Viral zoonotic risk is homogeneous among taxonomic orders of mammalian and avian reservoir hosts. *Proceedings of the National Academy of Sciences of the United States of America*, 117(17), 9423–9430. <https://doi.org/10.1073/pnas.1919176117>
- Mollentze, N., Streicker, D. G., Murcia, P. R., Hampson, K., & Biek, R. (2020). Virulence mismatches in index hosts shape the outcomes of cross-species transmission. *Proceedings of the National Academy of Sciences of the United States of America*, 117(46), 28859–28866. <https://doi.org/10.1073/pnas.2006778117>
- Moorman, N. J., Cristea, I. M., Terhune, S. S., Rout, M. P., Chait, B. T., & Shenk, T. (2008). Rapid proteomic profiling reveals that human cytomegalovirus UL38 protein antagonizes the tuberous sclerosis protein complex. *Cell Host & Microbe*, 3(4), 253.
- Müller, A., Baumann, A., Essbauer, S., Radosa, L., Krüger, D. H., Witkowski, P. T., Zeier, M., & Krautkrämer, E. (2019). Analysis of the integrin  $\beta 3$  receptor for pathogenic orthohantaviruses in rodent host species. *Virus Research*, 267, 36–40. <https://doi.org/10.1016/j.virusres.2019.04.009>
- Mustonen, J., Partanen, J., Kanerva, M., Pietilä, K., Vapalahti, O., Pasternack, A., & Vaheri, A. (1996). Genetic susceptibility to severe course of nephropathia epidemica caused by Puumala hantavirus. *Kidney International*, 49(1), 217–221. <https://doi.org/10.1038/ki.1996.29>
- Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), 691–692. <https://doi.org/10.1093/biomet/78.3.691>
- Olival, K. J., Hosseini, P. R., Zambrana-Torrel, C., Ross, N., Bogich, T. L., & Daszak, P. (2017). Host and viral traits predict zoonotic spillover from mammals. *Nature*, 546(7660), 646. <https://doi.org/10.1038/nature22975>
- Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E.-C., Burke, D. S., Calisher, C. H., Laughlin, C. A., Saif, L. J., & Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews*, 72(3), 457–470. <https://doi.org/10.1128/MMBR.00004-08>
- Plowright, R. K., Parrish, C. R., McCallum, H., Hudson, P. J., Ko, A. I., Graham, A. L., & Lloyd-Smith, J. O. (2017). Pathways to zoonotic spillover. *Nature Reviews Microbiology*, 15(8), 502. <https://doi.org/10.1038/nrmicro.2017.45>
- Pybus, O. G., & Rambaut, A. (2009). Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, 10(8), 540. <https://doi.org/10.1038/nrg2583>
- Reperant, L. A., Kuiken, T., & Osterhaus, A. D. M. E. (2012). Adaptive pathways of zoonotic influenza viruses: From exposure to

- establishment in humans. *Vaccine*, 30(30), 4419–4434. <https://doi.org/10.1016/j.vaccine.2012.04.049>
- Rohfritsch, A., Galan, M., Gautier, M., Gharbi, K., Olsson, G., Gschloessl, B., Zeimes, C., VanWambeke, S., Vitalis, R., & Charbonnel, N. (2018). Preliminary insights into the genetics of bank vole tolerance to Puumala hantavirus in Sweden. *Ecology and Evolution*, 8(22), 11273–11292. <https://doi.org/10.1002/ece3.4603>
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), 539–542.
- Roossinck, M. J. (2011). The good viruses: Viral mutualistic symbioses. *Nature Reviews Microbiology*, 9(2), 99. <https://doi.org/10.1038/nrmicro2491>
- Saxenhofer, M., de Melo, V., Ulrich, R. G., & Heckel, G. (2017). Revised time scales of RNA virus evolution based on spatial information. *Proceedings of the Royal Society B: Biological Sciences*, 284(1860), 20170857. <https://doi.org/10.1098/rspb.2017.0857>
- [Dataset] Saxenhofer, M., Labutin, A., White, T. A., & Heckel, G. (2021). Host genetic factors associated with the range limit of a European hantavirus. *Dryad digital repository* <https://doi.org/10.5061/dryad.5dv41ns6p>
- Saxenhofer, M., Schmidt, S., Ulrich, R. G., & Heckel, G. (2019). Secondary contact between diverged host lineages entails ecological speciation in a European hantavirus. *PLoS Biology*, 17(2), e3000142. <https://doi.org/10.1371/journal.pbio.3000142>
- Schlegel, M., Kindler, E., Essbauer, S. S., Wolf, R., Thiel, J., Groschup, M. H., Heckel, G., Oehme, R. M., & Ulrich, R. G. (2012). Tula virus infections in the Eurasian water vole in Central Europe. *Vector-Borne and Zoonotic Diseases*, 12(6), 503–513. <https://doi.org/10.1089/vbz.2011.0784>
- Schmidt, S., Reil, D., Jeske, K., Drewes, S., Rosenfeld, U. M., Fischer, S., Spierling, N. G., Labutin, A., Heckel, G., Jacob, J., Ulrich, R. G., & Imholt, C. (2021). Spatial and temporal dynamics and molecular evolution of Tula orthohantavirus in German vole populations. *Viruses*, 13, 1132. <https://doi.org/10.3390/v13061132>
- Schmidt, S., Saxenhofer, M., Drewes, S., Schlegel, M., Wanka, K. M., Frank, R., Klimpel, S., von Blanckenhagen, F., Maaz, D., Herden, C., Freise, J., Wolf, R., Stubbe, M., Borkenhagen, P., Ansorge, H., Eccard, J. A., Lang, J., Jourdain, E., Jacob, J., ... Ulrich, R. G. (2016). High genetic structuring of Tula hantavirus. *Archives of Virology*, 161(5), 1135–1149. <https://doi.org/10.1007/s00705-016-2762-6>
- Schmidt-Chanasit, J., Essbauer, S., Petraityte, R., Yoshimatsu, K., Tackmann, K., Conraths, F. J., Sasnauskas, K., Arikawa, J., Thomas, A., Pfeiffer, M., Scharninghausen, J. J., Splettstoesser, W., Wenk, M., Heckel, G., & Ulrich, R. G. (2010). Extensive host sharing of central European Tula virus. *Journal of Virology*, 84(1), 459–474. <https://doi.org/10.1128/JVI.01226-09>
- Schneider, J., Hoffmann, B., Fevola, C., Schmidt, M. L., Imholt, C., Fischer, S., & Ulrich, R. G. (2021). Geographical distribution and genetic diversity of bank vole hepaciviruses in Europe. *Viruses*, 13, 1258. <https://doi.org/10.3390/v13071258>
- Schweizer, M., Excoffier, L., & Heckel, G. (2007). Fine-scale genetic structure and dispersal in the common vole (*Microtus arvalis*). *Molecular Ecology*, 16(12), 2463–2473. <https://doi.org/10.1111/j.1365-294X.2007.03284.x>
- Sessions, O. M., Barrows, N. J., Souza-Neto, J. A., Robinson, T. J., Hershey, C. L., Rodgers, M. A., Ramirez, J. L., Dimopoulos, G., Yang, P. L., Pearson, J. L., & Garcia-Blanco, M. A. (2009). Discovery of insect and human dengue virus host factors. *Nature*, 458(7241), 1047. <https://doi.org/10.1038/nature07967>
- Shi, M., Lin, X.-D., Chen, X., Tian, J.-H., Chen, L.-J., Li, K., & Zhang, Y.-Z. (2018). The evolutionary history of vertebrate RNA viruses. *Nature*, 556(7700), 197–202.
- Simmonds, P., Aiewsakun, P., & Katzourakis, A. (2019). Prisoners of war – host adaptation and its constraints on virus evolution. *Nature Reviews Microbiology*, 17(5), 321–328. <https://doi.org/10.1038/s41579-018-0120-2>
- Sironi, M., Cagliani, R., Forni, D., & Clerici, M. (2015). Evolutionary insights into host-pathogen interactions from mammalian sequence data. *Nature Reviews Genetics*, 16(4), 224. <https://doi.org/10.1038/nrg3905>
- Streicker, D. G., Turmelle, A. S., Vonhof, M. J., Kuzmin, I. V., McCracken, G. F., & Rupprecht, C. E. (2010). Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science*, 329(5992), 676–679.
- Strelau, M., Owens, C. M., Perron, M. J., Kiessling, M., Autissier, P., & Sodroski, J. (2004). The cytoplasmic body component TRIM5 $\alpha$  restricts HIV-1 infection in Old World monkeys. *Nature*, 427(6977), 848–853. <https://doi.org/10.1038/nature02343>
- Sun, Q., Qi, X., Zhang, Y., Wu, X., Liang, M., Li, C., Li, D., Cardona, C. J., & Xing, Z. (2016). Synaptogyrin-2 promotes replication of a novel tick-borne bunyavirus through interacting with viral nonstructural protein NSs. *Journal of Biological Chemistry*, 291(31), 16138–16149. <https://doi.org/10.1074/jbc.M116.715599>
- Sutter, A., Beysard, M., & Heckel, G. (2013). Sex-specific clines support incipient speciation in a common European mammal. *Heredity*, 110, 398–404. <https://doi.org/10.1038/hdy.2012.124>
- Switzer, W. M., Salemi, M., Shanmugam, V., Gao, F., Cong, M., Kuiken, C., & Heneine, W. (2005). Ancient co-speciation of simian foamy viruses and primates. *Nature*, 434(7031), 376–380.
- Theodosopoulos, A. N., Hund, A. K., & Taylor, S. A. (2019). Parasites and host species barriers in animal hybrid zones. *Trends in Ecology & Evolution*, 34(1), 19–30. <https://doi.org/10.1016/j.tree.2018.09.011>
- Tseng, C. K. H., Marquez, V. E., Fuller, R. W., Goldstein, B. M., Haines, D. R., McPherson, H., Parsons, J. L., Shannon, W. M., & Arnett, G. (1989). Synthesis of 3-deazaneplanocin A, a powerful inhibitor of S-adenosylhomocysteine hydrolase with potent and selective in vitro and in vivo antiviral activities. *Journal of Medicinal Chemistry*, 32(7), 1442–1446. <https://doi.org/10.1021/jm00127a007>
- Vaheri, A., Strandin, T., Hepojoki, J., Sironen, T., Henttonen, H., Mäkelä, S., & Mustonen, J. (2013). Uncovering the mysteries of hantavirus infections. *Nature Reviews Microbiology*, 11(8), 539. <https://doi.org/10.1038/nrmicro3066>
- van Doremalen, N., Miazgowicz, K. L., Milne-Price, S., Bushmaker, T., Robertson, S., Scott, D., Kinne, J., McLellan, J. S., Zhu, J., & Munster, V. J. (2014). Host species restriction of Middle East respiratory syndrome coronavirus through its receptor, dipeptidyl peptidase 4. *Journal of Virology*, 88(16), 9220–9232. <https://doi.org/10.1128/JVI.00676-14>
- Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2), 117–186. <https://doi.org/10.1214/aoms/1177731118>
- Walker, L. R., Engle, T. B., Vu, H., Tosky, E. R., Nonneman, D. J., Smith, T. P. L., Borza, T., Burkey, T. E., Plastow, G. S., Kachman, S. D., & Ciobanu, D. C. (2018). Synaptogyrin-2 influences replication of Porcine circovirus 2. *PLoS Genetics*, 14(10), e1007750. <https://doi.org/10.1371/journal.pgen.1007750>
- Wang, M. L., Lai, J. H., Zhu, Y., Zhang, H. B., Li, C., Wang, J. P., Li, Y. M., Yang, A. G., & Jin, B. Q. (2009). Genetic susceptibility to haemorrhagic fever with renal syndrome caused by Hantaan virus in Chinese Han population. *International Journal of Immunogenetics*, 36(4), 227–229. <https://doi.org/10.1111/j.1744-313X.2009.00848.x>
- Webby, R., Hoffmann, E., & Webster, R. (2004). Molecular constraints to interspecies transmission of viral pathogens. *Nature Medicine*, 10(12s), S77. <https://doi.org/10.1038/nm1151>
- Weber de Melo, V., Sheikh Ali, H., Freise, J., Kuhnert, D., Essbauer, S. S., Mertens, M., & Heckel, G. (2015). Spatiotemporal dynamics of Puumala hantavirus associated with its rodent host, *Myodes glareolus*. *Evolutionary Applications*, 8, 545–559. <https://doi.org/10.1111/eva.12263>

- White, B. J., Lawniczak, M. K. N., Cheng, C., Coulibaly, M. B., Wilson, M. D., Sagnon, N., & Besansky, N. J. (2011). Adaptive divergence between incipient species of *Anopheles gambiae* increases resistance to Plasmodium. *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), 244–249. <https://doi.org/10.1073/pnas.1013648108>
- White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: The invasive bank vole (*Myodes glareolus*) in Ireland. *Molecular Ecology*, 22(11), 2971–2985.
- Worobey, M., Telfer, P., Souquière, S., Hunter, M., Coleman, C. A., Metzger, M. J., Reed, P., Makuwa, M., Hearn, G., Honarvar, S., Roques, P., Apetrei, C., Kazanji, M., & Marx, P. A. (2010). Island biogeography reveals the deep history of SIV. *Science*, 329(5998), 1487. <https://doi.org/10.1126/science.1193550>
- Zhou, H., Alexander, D., & Lange, K. (2011). A quasi-Newton acceleration for high-dimensional optimization algorithms. *Statistics and Computing*, 21(2), 261–273. <https://doi.org/10.1007/s11222-009-9166-3>
- Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44(7), 821–824. <https://doi.org/10.1038/ng.2310>

#### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Saxenhofer, M., Labutin, A., White, T. A., & Heckel, G. (2022). Host genetic factors associated with the range limit of a European hantavirus. *Molecular Ecology*, 31, 252–265. <https://doi.org/10.1111/mec.16211>