Mapping and Sequencing of a Significant Quantitative Trait Locus Affecting Resistance to Koi Herpesvirus in Common Carp

G3: Genes|Genomes|Genetics Early Online, published on August 27, 2018 as doi:10.1534/g3.118.200593

Mapping and sequencing of a significant quantitative trait locus affecting resistance to Koi herpesvirus in common carp

Christos Palaiokostas[1,2], Diego Robledo[1], Tomas Vesely[4], , Martin Prchal[3], Dagmar Pokorova[4], Veronika Piackova[3], Lubomir Pojezdal[4], Martin Kocour[3], Ross D. Houston[1*]

[1] The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland, United Kingdom

[2] Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7090, 750 07 Uppsala, Sweden

[3] University of South Bohemia in Ceske Budejovice, Faculty of Fisheries and Protection of Waters, South Bohemian Research Centre of Aquaculture and Biodiversity of Hydrocenoses, Zatisi 728/II, 389 25 Vodnany, Czech Republic

[4] Veterinary Research Institute, Hudcova 70, Brno 62100, Czech Republic

Abstract

Cyprinids are the most highly produced group of fishes globally, with common carp being one of the most valuable species of the group. Koi herpesvirus (KHV) infections can result in high levels of mortality, causing major economic losses, and is listed as a notifiable disease by the World Organisation for Animal Health. Selective breeding for host resistance has the potential to reduce morbidity and losses due to KHV. Therefore, improving knowledge about host resistance and methods of incorporating genomic data into breeding for resistance may contribute to a decrease in economic losses in carp farming. In the current study, a population of 1,425 carp juveniles, originating from a factorial cross between 40 sires and 20 dams was challenged with KHV. Mortalities and survivors were recorded and sampled for genotyping by sequencing using Restriction Site-Associated DNA sequencing (RADseq). Genome-wide association analyses were performed to investigate the genetic architecture of resistance to KHV. A genome-wide significant QTL affecting resistance to KHV was identified on linkage group 44, explaining approximately 7 % of the additive genetic variance. Pooled whole genome resequencing of a subset of resistant (n = 60) and susceptible animals (n = 60) was performed to characterize QTL regions, including identification of putative candidate genes and functional

annotation of associated polymorphisms. The TRIM25 gene was identified as a promising positional and functional candidate within the QTL region of LG 44, and a putative premature stop mutation in this gene was discovered.

Introduction

Common carp (*Cyprinus carpio* and *Cyprinus rubrofuscus*), is one of the most highly produced aquaculture fish species globally (FAO, 2015), being farmed in a wide variety of environments and production systems (Balon 1995). However, in common with many aquaculture species, only a minority of farmed carp are derived from family-based selective breeding programs (Vandeputte 2003; Janssen *et al.* 2017). The potential for selective breeding to enhance production in carp is highlighted by several studies, but much of the production of commercial stock is still generated via intraspecific crossbreeding (Kocour *et al.* 2005, 2007; Vandeputte *et al.* 2008; Nielsen *et al.* 2010; Prchal *et al.* 2018).

Koi herpesvirus (KHV), also known as Cyprinid herpesvirus-3 (CyHV-3), is one of the main threats to carp production. The first major outbreaks were recorded in 1998 (Hedrick *et al.* 2000), and subsequent outbreaks in many carp producing countries were reported worldwide (Haenen *et al.* 2004). The seriousness of the KHV threat is highlighted by its listing as a notifiable disease by the European Union (Taylor *et al.* 2010) and the World Organization for Animal Health (OIE 2018). Selective breeding is a valuable tool for contributing to sustainable food production through the prevention and management of infectious outbreaks in a wide range of species (Bishop and Woolliams 2014). This may be particularly true in aquaculture species, due to moderate to high heritabilities of disease resistance documented in numerous cases (Ødegård *et al.* 2011; Houston 2017), and successful examples of disease control using

marker-assisted breeding, e.g. the case of the IPN virus in Atlantic salmon; *Salmo salar* (Houston *et al.* 2008; Moen *et al.* 2009).

Several studies have investigated the genetic basis of KHV resistance in carp (utilizing data and samples collected from disease challenge trials), showing encouraging results with large variation in survival both between families (Dixon *et al.* 2009; Tadmor-Levi et al. 2017) and between strains (Shapira *et al.* 2005; Piačková *et al.* 2013). Results from candidate gene association studies have suggested a possible role for polymorphism in MHC loci (Rakus *et al.* 2009) and Interleukin-11 (Kongchum *et al.* 2011) in host resistance to KHV. Taken together, these studies indicate that selective breeding has the potential to increase resistance to KHV, with potential downstream benefits for the carp aquaculture industry and fish welfare. However, to date, genome-wide polymorphisms have not been applied to investigate the genetic architecture of resistance to KHV.

Restriction-site associated DNA sequencing (RADseq) (Baird *et al.* 2008) and similar genotyping by sequencing techniques have been widely applied to generated genome-wide SNP markers due to their cost-efficiency in a wide range of aquaculture species (Robledo *et al.* 2017), including common carp ((Palaiokostas *et al.* 2018a). Various genome wide association studies (GWAS) using this technique have been published in aquaculture species (e.g. Campbell *et al.* 2014; Palti *et al.* 2015). GWAS have been used to study disease resistance in various aquaculture species including salmonids (Correa *et al.* 2015, 2017; Vallejo *et al.* 2017; Barría *et al.* 2018; Robledo *et al.* 2018), catfish (Zhou *et al.* 2017), European sea bass (Palaiokostas *et al.* 2018b) and Pacific oyster (Gutierrez *et al.* 2018) amongst others. With the notable exception of the aforementioned case of IPN resistance in salmon, the GWAS results have pointed to a polygenic or oligogenic architecture for disease resistance in aquaculture

species. The main aim of this study was to investigate genetic resistance to KHV in common carp using a RADseq approach. Classical genome wide association study (GWAS) and weighted genomic best linear unbiased predictor (WGBLUP) approaches were taken to examine the genetic architecture of resistance. Finally, pooled whole genome sequencing (PWGS) was performed in a subset of samples with divergent resistance and susceptibility to characterize and annotate QTL regions, and to identify potential gene candidates and polymorphisms involved in KHV resistance.

Materials and Methods

Sample collection and disease challenge

A population of Amur Mirror Carp was created at the University of South Bohemia in České Budějovice, Czech Republic in May 2014 using artificial insemination (Vandeputte *et al.* 2004) involving four factorial crosses of five dams x ten sires (20 dams and 40 sires in total). Incubation of eggs was performed in 9 L Zugar jars at 20°C. At the first swimming stage, randomly sampled progeny from each mating (of approximately equal total volume) were pooled and stocked into several nursery earthen ponds at stocking density of 150,000 larvae / ha and reared under semi-intensive pond conditions throughout the growing season (from May to September). Before the challenge test a random sample of 1,500 fish described above were tagged and fin clipped for DNA extraction. These fish were the same as those described in Palaiokostas et al. (2018a). These animals were acclimatized for five days at water temperature of 22 °C and bathed in FMC solution (formalin, malachite green, methylene blue using a dose of 2 mL per 100 L of water) to eliminate ectoparasites. Subsequently, the fish were transferred to Veterinary Research Institute (VRI) in Brno (Czech Republic) to perform the KHV disease challenge test. A small (n = 215) sample of koi carp were challenged alongside the Amur mirror carp as a positive control, since Koi carp are highly susceptible to KHV.

A cohabitation challenge was performed in a 1,400 L tank equipped with recirculation and biological filtration. Koi carp received an intraperitoneal injection with 0.2 mL culture medium containing 10^4 TCID 50 / mL KHV at day 0 and were added into the tank with challenged fish. Mortality of individual fish was recorded twice a day for a period of 35 days post infection (dpi). Presence of KHV on a sample of dead fish (n = 100) was confirmed by PCR according to guidelines by the Centre for Environment, Fisheries & Aquaculture Science, UK (Cefas) (Pokorova et al. 2010). The experiment was run until mortalities were negligible, implying that survivors were resistant. The entire experiment was conducted in accordance with the law on the protection of animals against cruelty (Act no. 246/1992 Coll. of the Czech Republic) upon its approval by Institutional Animal Care and Use Committee (IACUC) of the VRI and appropriate state authority. All people conducting the experiment hold a certificate about qualification to conduct experiments on the live animals, and the VRI is accredited for the culture of experimental animals according to the aforementioned law.

Library preparation and sequencing

The RAD library preparation protocol followed the methodology originally described in Baird *et al*. (2008) and presently in detail in Palaiokostas et al (2018a). Briefly, template DNA was digested using the *Sbf*I (recognizing the CCTGCA|GG motif) high fidelity restriction enzyme (New England Biolabs; NEB). DNA shearing was conducted with a Pico bioruptor (Diagenode). Following a final gel elution step into 20 µL EB buffer (MinElute Gel Purification Kit, Qiagen), 66 libraries (24 animals each) were sent to BMR Genomics (Italy), for quality control and high-throughput sequencing. RAD libraries were run in fourteen lanes of an Illumina NextSeq 500, using 75 base paired-end reads (v2 chemistry).

Whole genome sequencing libraries (n = 4) from pooled DNA samples (30 animals each library) of susceptible and resistant animals were constructed using the Illumina TruSeq DNA PCR free kit (350bp insert). Sequencing was performed in Edinburgh Genomics facilities using two lanes of Illumina HiSeq 4000.

SNP discovery and genotyping

The process of obtaining the SNP genotype data from the RADSeq reads was described in detail in Palaikostas *et al* (2018a). Briefly, sequenced reads were aligned to the common carp reference genome assembly version *GCA_000951615.2* (Xu *et al.* 2014) using bowtie2 (Langmead and Salzberg 2012). The aligned reads were sorted into RAD loci and SNPs were identified using the Stacks software 1.4 (Catchen *et al.* 2011). The SNPs were detected using a minimum stack depth of at least ten or five for the parental and offspring samples respectively. SNPs with minor allele frequency (MAF) below 0.01, greater than 20 % missing data, and deviating from expected Hardy-Weinberg equilibrium in the parental samples ($P < 1e-06$) were discarded. R/hsphase (Ferdosi *et al.* 2014) software was used for parentage assignment allowing for a maximum genotyping error of 4 %. The pedigree obtained was further validated for possible erroneous assignments using FImpute (Sargolzaei *et al.* 2014). In total, 1,214 offspring were uniquely assigned, forming 195 full-sib families (40 sires, 20 dams). Since the carp reference genome assembly is currently very fragemented, a medium density linkage map of 12,311 SNPs grouped in 50 linkage groups was created (Palaiokostas *et al.* 2018b), and used to orientate the results from the GWAS.

Heritability estimation

The probit link function was used to connect the observed binary phenotype (0 = dead, 1 = alive) with the underlying liability scale. Variance components were estimated using the

R/BGLR (Pérez and de Los Campos 2014) software with the following animal model:

$$\mathbf{l} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}, \quad (1)$$

where **l** is the vector of latent variables, **b** is the vector of the fixed effects (cross, standard length), **X** is the incidence matrix relating phenotypes with the fixed effects, **Z** is the incidence matrix relating phenotypes with the random animal effects, **u** is the vector of random animal effects ~ $N(0, \mathbf{A}\sigma_g^2)$ [where **A** corresponds to the pedigree-based relationship matrix and is replaced by G for analyses using the genomic relationship matrix (VanRaden 2008) and σ_g^2 is the additive genetic variance], **e** the vector of residuals ~$N(0, \mathbf{I}\sigma_e^2)$ where σ_e^2 is the residual variance.

The parameters of this model were estimated through Markov chain Monte Carlo (MCMC) using Gibbs sampling (11 M iterations; burn-in: 1 M; thin: 1,000). Convergence of the resulting posterior distributions was assessed both visually (inspecting the resulting MCMC plots) and analytically using R/coda v0.19-1 (Plummer *et al.* 2006). Heritability for the trait of survival during the KHV challenge (on the underlying liability scale) was estimated using the following formula:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2}$$

where σ_g^2 is the previous estimated additive genetic variance and σ_e^2 the residual variance. Residual variance on the underlying scale is not identifiable in threshold models (Goldstein *et al.* 2002; Nakagawa and Schielzeth 2013) and was therefore fixed to 1.

Genome wide association analysis (GWAS)

To test the association between individual SNPs and resistance to KHV, a classical genome wide association study (CGWAS) was performed using R/gaston (Perdry and Dandine-Roulland 2016). The mixed model applied for overall survival had the same format as in

(1) with the addition of including each SNP as a fixed effect. The variance components were estimated using the penalized quasi-likelihood approach (Chen *et al.* 2016). The genome-wide significance threshold was calculated using a Bonferroni correction (0.05 / N), where N represents the number of tested SNPs.

Weighted genomic best linear unbiased predictor (WGBLUP) was performed (Wang *et al.* 2012) using direct genomic values (DGV) (Lourenco *et al.* 2015; Zhang *et al.* 2016). The weighted genomic relationship matrix was initially created following VanRaden (2008) as:

$$\mathbf{G^* = ZDZ'q}$$

where **Z** *is* the design matrix relating genotypes of each locus, **D** is a weight matrix for all SNPs, and **q** is a weighting vector derived from observed SNP frequencies. SNP weights were calculated using the nonlinearA method (VanRaden 2008). Briefly the steps for performing WGBLUP were as follows (Wang *et al.* 2012):

a) Initialize **D** = **I** and t=1, where **I** the identity matrix and t is the iteration number.
b) Calculate $\mathbf{G}^*$.
c) Estimate DGVs.
d) Estimate SNP effects from GEBVs : $\hat{\boldsymbol{\alpha}} = \mathbf{qDZ'}\boldsymbol{G}^*\hat{\mathbf{u}}$, where $\hat{\boldsymbol{\alpha}}$ the vector of SNP effects and $\hat{\mathbf{u}}$ the vector of DGV
e) Calculate the weight for each SNP: $d_{ii}^{(t+1)} = 1.125^{\frac{|\hat{a}_i|}{sd(\hat{a})}-2}$, where $\hat{a}_i$ the estimated SNP effect (VanRaden 2008).
f) Normalize SNP weights so the total genetic variance remains constant.
g) Loop to step d) until convergence (10^{-14}).

Convergence of SNP weights was tested using the convergence criterion BLUPF90 uses for variance components estimation

$$C = \frac{\sum_i (\theta_i - \xi_i)^2}{\sum_i \theta_i^2}$$

Percentage of additive genetic variance was estimated by non-overlapping windows of 10 adjacent SNPs as follows:

$$\frac{\mathrm{Var}(\alpha_i)}{\sigma_g^2} \times 100\% = \frac{\mathrm{Var}(\sum_i^{i=10} z_i \widehat{\alpha_i})}{\sigma_g^2} \times 100\%$$

where var(a_i) the additive genetic variance of the tested window of adjacent SNPs and σ_g^2 the total additive genetic variance. The weighted GBLUP analyses were performed using THRGIBBSF90 for estimating DGVs (Misztal *et al.* 2002) combined with iterations of PreGSF90 and PostGSF90 (Aguilar *et al.* 2011) until convergence (10^{-14}).

Pooled whole genome sequencing analysis

Pools of genomic DNA (25 ng / ul) from 60 survivors and 60 mortalities from the disease challenge experiment were prepared. These animals originated from 20 full-sib families, and the family structure was balanced between the resistant and susceptible pools. Libraries were prepared using the TruSeq DNA PCR free kit (350 bp insert size) and sequenced in two lanes of an Illumina HiSeq 4000 using paired-end sequencing by Edinburgh Genomics.

Reads were QC-filtered (phred score above 30) and trimmed to 140 bp long using Trimmomatic v0.36 (Bolger *et al.* 2014). Reads were aligned to the carp reference genome *GCA_000951615.2* (Xu *et al.* 2014) using bowtie2 (Langmead and Salzberg 2012). SNP identification was performed using Burrows-Wheeler Aligner v0.7.8 (BWA-mem, Li 2013). Pileup files describing the base-pair information at each genomic position were generated from the alignment files using the mpileup function of Samtools v1.6 (Li et al. 2009) requiring minimum mapping and base quality of 20. A Cochran-Mantel-Haenszel test was performed to test the significance of the allele frequency differences using Popoolation 2 v1.201 (Kofler et al. 2011). Only those genomic positions with at least 6 reads of the alternative allele across all pools and a maximum coverage of 50 reads and a

minimum of 8 in all pools were considered SNPs. All QC-filtered SNPs were annotated using SNPeff (Cingolani *et al.* 2012).

Data availability

Raw reads were deposited in the National Centre for Biotechnology Information (NCBI) repository under project ID PRJNA414021. Table S1 contains the phenotypic data. Table S2 contains the pedigree. Table S3 contains the genotypic data.

Results

Disease challenge

Mortalities began at 12 dpi reaching a maximum between 21 and 24 dpi (98 – 130 mortalities per day) decreasing thereafter with no mortalities observed after 35 dpi (Figure 1). The overall mortality in the KHV challenge experiment for the Amur Mirror Carp was 66 %. All observed mortalities displayed typical KHV symptoms (e.g. weakness, lethargy, loss of equilibrium, erratic swimming, sunken eyes, excessive mucous production, increased respiratory rate, discoloration, and hemorrhagic lesions on the skin and gills). The presence of KHV was confirmed in all tested samples (n = 100).

Heritability estimation

There was marked between-family variation in survival rate for both sires (6 – 83 %) and dams (0 – 52 %), suggesting the existence of considerable genetic variation for host resistance. Heritability estimates of overall survival for the pedigree and genomic relationship matrix on the underlying scale were 0.61 (HPD interval 95%: 0.42 – 0.80) and 0.50 (HPD interval 95%: 0.38 – 0.63) respectively.

Genome wide association approaches - SNP annotation in QTL region

The three SNPs with the highest association according to classical GWAS were located on linkage group 44 (chromosome 33; P < 1e-05; denoted by stars in Fig 2). This QTL was also identified using the WGBLUP approach (Figure 3) suggesting it accounted for approximately 7 % (convergence obtained after 5 iterations) of the additive genetic variance on the underlying scale. In addition the WGBLUP identified QTLs explaining more than 1% of the additive genetic variance in linkage groups 34 (~ 2.5%) and 42 (~ 1.1%). Whole genome sequencing data from the pools of resistance and susceptible animals was used to discover and annotate additional SNPs in the QTL region (Figure 4), and potential candidate genes were identified. Further, SNPs with significant allele frequency differences (P-value < 0.05) between the two groups were identified. A SNP coding for a putative premature stop codon was identified in gene TRIM25 (Glu258*), an E3 ubiquitin ligase with a major role in initiation of intracellular antiviral response to herpesviruses (Gupta *et al.* 2018).

Discussion

In the current study, high throughput sequencing was applied to study genetic resistance of common carp to KHV. While genomic data in the form of genetic markers can be a valuable addition to selective breeding for disease resistance, how to apply this data depends on the underlying genetic architecture. In the case of traits controlled by a major QTL, it may be most effective to use marker-assisted selection, while in the case of polygenic traits genomic selection is likely to be preferable. Modern genomic tools also facilitate high resolution study of the genomic regions underpinning genetic resistance, facilitating identification and annotation of promising functional candidate genes which may play a direct role in differential host response to infection.

Following pedigree reconstruction using the RAD SNP data, the heritability of resistance as measured by survival on the underlying scale was estimated to be 0.61 (pedigree) and 0.50 (genomic). This is an unusually high heritability estimate, but is comparable to the estimated of 0.79 that was previously documented for this trait (Ødegård *et al.* 2010). These independent high estimates of heritability of resistance to KHV highlight that selective breeding has major potential for producing carp with increased resistance. Additionally, in a recent study of introgression of KHV resistance from a wild carp strain to a farmed carp strain, significant additive genetic variation in resistance was detected (Tadmor-Levi *et al.* 2017). Furthermore, the authors showed that resistant carp do become infected, implying that resistance is due to an effective host response to infection (Tadmor-Levi *et al.* 2017). Early stage host response to KHV infection is likely to have a major interferon pathway component, with Interferon αβ, and interleukin 12 suggested to play a major role in Koi and Red common carp (Hwang *et al.* 2017).

The CGWAS resulted in the identification of genome-wide significant QTL on linkage group 44. While this test is the most commonly used association analysis, it fails to utilize all available information since it does not consider linkage disequilibrium between adjacent SNPs, resulting in reduced statistical power as opposed to methods where all SNPs are used simultaneously (Wang *et al.* 2012). The WGBLUP approach incorporates multiple SNPs and combines the computational efficiency of GBLUP with an increased statistical power for QTL detection (Zhang *et al.* 2016). However, WGBLUP has limitations as well like the heuristic influence regarding optimal number of iterations and the difficulty to determine appropriate significance levels for the identified QTL (Wang *et al.* 2012; Lourenco *et al.* 2015; Zhang *et al.* 2016). The recent implementation of nonlinearA (VanRaden 2008) in PostGSF90 (Misztal *et al. 2018*) may help circumvent the issue of optimal number of iterations due to its better convergence properties. NonlinearA benefits particularly in situations where a non normal prior distribution

more accurately describes the trait under study (VanRaden 2008). In the current study, both CGWAS and WGBLUP provided significant evidence for the existence of a QTL associated with resistance to KHV on linkage group 44, explaining approximately 7 % of the genetic variation in a highly heritable trait.

The SNP with highest association in the CGWAS was located ~6.5 Kb upstream of TRIM25, an E3 ubiquitin ligase with a major role in initiation of intracellular antiviral response to herpesviruses. Auto-ubiquitinisation of TRIM25 is a viral strategy for functional inactivation of the pattern recognition protein RIG1, and subsequent cellular interferon response (Gack *et al.* 2008). In the PWGS, the majority of the SNPs with significant allele frequency differences between the resistant and susceptible pools were annotated as 'intergenic'. However, interestingly, a putative premature stop mutation in position 258 of the carp TRIM25 protein was identified. TRIM25 has 649 - 682 amino acids (isoform dependant), and therefore this stop mutation is highly likely to result in loss of function. The premature stop causing allele is rare in the population, but reads of this allele were more common in the susceptible (n = 11) than the resistance (n = 3) pools, albeit the Cochran-Mantel-Haenszel test p-value for this SNP was only nominaly significant (0.049). This may fit with a loss of function of TRIM25 in susceptible fish, being unable to trigger an appropriate antiviral response.

It will be interesting to study whether this single genome-wide significant QTL for resistance to KHV has an effect in other carp populations and strains. Follow up functional studies of candidate genes in the QTL region, including assessment of gene expression response to infection and the differential response between alternate QTL types, may be a fruitful avenue to shortlist functional candidate genes. Currently, TRIM25 and its premature stop mutation seem to be the most promising candidates, and additional genotyping of this SNP alongside

directed functional studies may help to test if it may be causative for the QTL. While the QTL identified in the current study was highly significant, the proportion of genetic variation explained was relatively moderate, implying multifactorial causal mechanisms underlying host resistance. Nonetheless, it is plausible that genetic markers within the QTL region may have value for marker-assisted selection, either directly or via a genomic prediction strategy with increased weighting on QTL-region SNPs.

Conclusions

In conclusion, the results from the current study demonstrate that SNP markers generated via RADseq are effective at studying the genetic variation in resistance to KHV in a common carp breeding population. The RAD-derived SNPs facilitated the identification of a genome-wide significant QTL on LG 44 affecting resistance to KHV. The sequencing and annotation of the QTL regions provided candidate functional genes and polymorphisms for future study to understand the mechanisms underlying the QTL. This QTL may have value for selective breeding via incorporation into marker-assisted or genomic selection, albeit genetic resistance to KHV in common carp appears to be multifactorial in nature.

Competing interests

The authors declare that they have no competing interests

Author contributions

TV, MK, MP, VP, RH conceived the study, contributed to designing the experimental structure. MK, MP, VP share on establishing and on-growing the experimental stock, P.I.T. tagging and fin clipping the fish. TV, DP, LP carried out the own KHV challenge experiment. CP carried out DNA extractions, RAD library preparation and sequence data processing. CP. and RH carried out parentage assignment and the quantitative genetic analyses. All authors contributed to drafting the manuscript.

ACKNOWLEDGEMENTS

The authors are supported by funding from the European Union's Seventh Framework Programme (FP7 2007-2013) under grant agreement no. 613611 (FISHBOOST). CP and RH gratefully acknowledge Institute Strategic Funding Grants to The Roslin Institute (Grant Nos. BBS/E/D/20002172, BBS/E/D/30002275, and BBS/E/D/10002070). MK, MP, VP were also supported by the Ministry of Education, Youth, and Sports of the Czech Republic (projects CENAKVA – CZ.1.05/2.1.00/01.0024 and CENAKVA II – LO1205 under the NPU I program) and the Grant Agency of the University of South Bohemia (project no. 125/2016/Z).

References

Aguilar, I., I. Misztal, A. Legarra, and S. Tsuruta, 2011 Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J. Anim. Breed. Genet. 128: 422–428.

Baird, N. a, P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver *et al.*, 2008 Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3: e3376.

Balon, E. K., 1995 Origin and domestication of the wild carp, *Cyprinus carpio*: from Roman gourmets to the swimming flowers. Aquaculture 129: 3–48.

Barría, A., K. A. Christensen, G. M. Yoshida, K. Correa, A. Jedlicki *et al.*, 2018 Genomic Predictions and Genome-Wide Association Study of Resistance Against Piscirickettsia salmonis in Coho Salmon (*Oncorhynchus kisutch*) Using ddRAD Sequencing. G3 (Bethesda) 8(4): 1183 - 1194.

Bishop, S. C., and J. A. Woolliams, 2014 Genomics and disease resistance studies in livestock. Livest. Sci. 166: 190–198.

Bolger, A. M., M. Lohse, and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–2120.

Campbell, N. R., S. E. LaPatra, K. Overturf, R. Towner, and S. R. Narum, 2014 Association mapping of disease resistance traits in rainbow trout using restriction site associated DNA sequencing. G3 (Bethesda) 4: 2473–81.

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait, 2011 Stacks: building and genotyping Loci de novo from short-read sequences. G3 (Bethesda) 1: 171–82.

Chen, H., C. Wang, M. P. Conomos, A. M. Stilp, Z. Li *et al.*, 2016 Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. Am. J. Hum. Genet. 98: 653–666.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin). 6: 80–92.

Correa, K., J. P. Lhorente, L. Bassini, M. E. López, A. Di Genova *et al.*, 2017 Genome wide association study for resistance to Caligus rogercresseyi in Atlantic salmon (*Salmo salar* L.) using a 50K SNP genotyping array. Aquaculture 472: 61–65.

Correa, K., J. P. Lhorente, M. E. López, L. Bassini, S. Naswa *et al.*, 2015 Genome-wide association analysis reveals loci associated with resistance against Piscirickettsia salmonis in two Atlantic salmon (*Salmo salar* L.) chromosomes. BMC Genomics 16: 854.

Dixon, P. F., C. L. Joiner, K. Way, R. A. Reese, G. Jeney *et al.*, 2009 Comparison of the resistance of selected families of common carp, *Cyprinus carpio* L., to koi herpesvirus: preliminary study. J. Fish Dis. 32: 1035–1039.

Ferdosi, M. H., B. P. Kinghorn, J. H. J. van der Werf, S. H. Lee, and C. Gondro, 2014 hsphase: An R package for pedigree reconstruction, detection of recombination events, phasing and imputation of half-sib family groups. BMC Bioinformatics 15:172.

Gack, M. U., A. Kirchhofer, Y. C. Shin, K.-S. Inn, C. Liang *et al.*, 2008 Roles of RIG-I N-terminal tandem CARD and splice variant in TRIM25-mediated antiviral signal transduction. Proc. Natl. Acad. Sci. U. S. A. 105: 16743–8.

Goldstein, H., W. Browne, and J. Rasbash, 2002 Partitioning Variation in Multilevel Models. Underst. Stat. 1: 223–231.

Gupta, S., P. Ylä-Anttila, S. Callegari, M. H. Tsai, H. J. Delecluse *et al.*, 2018 Herpesvirus deconjugases inhibit the IFN response by promoting TRIM25 autoubiquitination and functional inactivation of the RIG-I signalosome. PLoS Pathog. 14(1): e1006852.

Gutierrez, A. P., T. P. Bean, C. Hooper, C. A. Stenton, M. B. Sanders *et al.*, 2018 A Genome-

Wide Association Study for Host Resistance to Ostreid Herpesvirus in Pacific Oysters (*Crassostrea gigas*). G3 (Bethesda). 8: 1273–1280.
Haenen, O. L. M., K. Way, S. M. Bergmann, and E. Ariel, 2004 The emergence of koi herpesvirus and its significance to European aquaculture. Bull. Eur. Ass. Fish Pathol 24(6).
Hedrick, R. P., O. Gilad, S. Yun, J. V. Spangenberg, G. D. Marty *et al.*, 2000 A herpesvirus associated with mass mortality of juvenile and adult koi, a strain of common carp. J. Aquat. Anim. Health 12: 44–57.
Hervé Perdry and Claire Dandine-Roulland gaston: Genetic Data Handling (QC, GRM, LD, PCA); Linear Mixed Models version 1.5 from CRAN. Available from: https://rdrr.io/cran/gaston/
Houston, R. D., and R. Brasileira De Zootecnia, 2017 Invited Review Future directions in breeding for disease resistance in aquaculture species. Bras. Zootec 46: 545–551.
Houston, R. D., C. S. Haley, A. Hamilton, D. R. Guy, A. E. Tinch *et al.*, 2008 Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). Genetics 178: 1109–1115.
Hwang, J.-A., J. E. Kim, H.-S. Kim, and J.-H. Lee, 2017 Immune Response to Koi Herpesvirus (KHV) of Koi and Koi × Red Common Carp (*Cyprinus carpio*). Dev. Reprod. 21: 361–370.
Janssen, K., H. Chavanne, P. Berentsen, and H. Komen, 2017 Impact of selective breeding on European aquaculture. Aquaculture 472: 8–16.
Kocour, M., D. Gela, M. Rodina, and O. Linhart, 2005 Testing of performance in common carp Cyprinus carpio L. under pond husbandry conditions I: top-crossing with Northern mirror carp. Aquac. Res. 36: 1207–1215.
Kocour, M., S. Mauger, M. Rodina, D. Gela, O. Linhart *et al.*, 2007 Heritability estimates for processing and quality traits in common carp (*Cyprinus carpio* L.) using a molecular pedigree. Aquaculture 270: 43–50.
Kongchum, P., E. Sandel, S. Lutzky, E. M. Hallerman, G. Hulata *et al.*, 2011 Association between IL-10a single nucleotide polymorphisms and resistance to cyprinid herpesvirus-3 infection in common carp (*Cyprinus carpio*). Aquaculture 315: 417–421.
Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357–359.
Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, Y. Masuda, I. Aguilar *et al.*, 2015 Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. J. Anim. Sci. 93: 2653–2662.
Misztal, I., Shogo Tsuruta, Daniela Lourenco, I. Aguilar, A. Legarra *et al.* Manual for BLUPF90 family of programs. Univ. Georg. Athens, USA.
Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet *et al.*, 2002 BLUPF90 and related programs (BGF90), pp. 21–22 in *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*,.
Moen, T., M. Baranski, A. K. Sonesson, and S. Kjøglum, 2009 Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. BMC Genomics 10: 368.
Nakagawa, S., and H. Schielzeth, 2013 A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4: 133–42.
Nielsen, H. M., J. Ødegård, I. Olesen, B. Gjerde, L. Ardo *et al.*, 2010 Genetic analysis of common carp (*Cyprinus carpio*) strains. I: Genetic parameters and heterosis for growth traits and survival. Aquaculture 304: 14–21.
Ødegård, J., M. Baranski, B. Gjerde, and T. Gjedrem, 2011 Methodology for genetic

517 evaluation of disease resistance in aquaculture species: Challenges and future prospects.
518 Aquac. Res. 42: 103–114.
519 Ødegård, J., I. Olesen, P. Dixon, Z. Jeney, H. M. Nielsen *et al.*, 2010 Genetic analysis of
520 common carp (*Cyprinus carpio*) strains. II: Resistance to koi herpesvirus and
521 Aeromonas hydrophila and their relationship with pond survival. Aquaculture 304: 7–
522 13.
523 OIE, 2018 OIE-Listed diseases, infections and infestations in force in 2018. World Organ.
524 Anim. Heal.
525 Palaiokostas, C., M. Kocour, M. Prchal, and R. D. Houston, 2018a Accuracy of genomic
526 evaluations of juvenile growth rate in common carp (*Cyprinus carpio*) using genotyping
527 by sequencing. Front. Genet. 9: 82.
528 Palaiokostas, C., S. Cariou, A. Bestin, J.-S. Bruant, P. Haffray *et al.*, 2018b Genome-wide
529 association and genomic prediction of resistance to viral nervous necrosis in European
530 sea bass (*Dicentrarchus labrax*) using RAD sequencing. Genet. Sel. Evol. 50: 30.
531 Palti, Y., R. L. Vallejo, G. Gao, S. Liu, A. G. Hernandez *et al.*, 2015 Detection and
532 Validation of QTL Affecting Bacterial Cold Water Disease Resistance in Rainbow Trout
533 Using Restriction-Site Associated DNA Sequencing (H. Wang, Ed.). PLoS One 10:
534 e0138435.
535 Pérez, P., and G. de Los Campos, 2014 Genome-Wide Regression and Prediction with the
536 BGLR Statistical Package. Genetics 198(2):483-495.
537 Piačková, V., M. Flajšhans, D. Pokorová, S. Reschová, D. Gela *et al.*, 2013 Sensitivity of
538 common carp, *Cyprinus carpio* L., strains and crossbreeds reared in the Czech Republic
539 to infection by cyprinid herpesvirus 3 (CyHV-3; KHV). J. Fish Dis. 36: 75–80.
540 Plummer, M., N. Best, K. Cowles, and K. Vines, 2006 CODA: convergence diagnosis and
541 output analysis for MCMC. R News 6: 7–11.
542 Prchal, M., A. Kause, M. Vandeputte, D. Gela, J. M. Allamellou *et al.*, 2018 The genetics of
543 overwintering performance in two-year old common carp and its relation to performance
544 until market size. PLoS One 13(1):e0191624.
545 Rakus, K. Ł., G. F. Wiegertjes, M. Adamek, A. K. Siwicki, A. Lepa *et al.*, 2009 Resistance of
546 common carp (*Cyprinus carpio* L.) to Cyprinid herpesvirus-3 is influenced by major
547 histocompatibility (MH) class II B gene polymorphism. Fish Shellfish Immunol. 26:
548 737–43.
549 Robledo, D., O. Matika, A. Hamilton, and R. D. Houston, 2018 Genome-Wide Association
550 and Genomic Selection for Resistance to Amoebic Gill Disease in Atlantic Salmon. G3
551 (Bethesda). 8: 1195–1203.
552 Robledo, D., C. Palaiokostas, L. Bargelloni, P. Martínez, and R. Houston, 2017 Applications
553 of genotyping by sequencing in aquaculture breeding and genetics. Rev. Aquac 0:1-13.
554 Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel, 2014 A new approach for efficient
555 genotype imputation using information from relatives. BMC Genomics 15: 478.
556 Shapira, Yechiam, Magen Yigal, Zak Tatyana, Kotler Mosche, Hulata Gideon, L.-S. B., 2005
557 Differential resistance to koi herpes virus (KHV)/carp interstitial nephritis and gill
558 necrosis virus (CNGV) among common carp (*Cyprinus carpio* L.) strains and
559 crossbreds. Aquaculture 245: 1–11.
560 Tadmor-Levi, R., E. Asoulin, G. Hulata, and L. David, 2017 Studying the Genetics of
561 Resistance to CyHV-3 Disease Using Introgression from Feral to Cultured Common
562 Carp Strains. Front. Genet. 8: 24.
563 Taylor, N. G. H., P. F. Dixon, K. R. Jeffery, E. J. Peeler, K. L. Denham *et al.*, 2010 Koi
564 herpesvirus: distribution and prospects for control in England and Wales. J. Fish Dis. 33:
565 221–230.
566 Vallejo, R. L., S. Liu, G. Gao, B. O. Fragomeni, A. G. Hernandez *et al.*, 2017 Similar

Genetic Architecture with Shared and Unique Quantitative Trait Loci for Bacterial Cold Water Disease Resistance in Two Rainbow Trout Breeding Populations. Front. Genet. 8: 156.

Vandeputte, M., 2003 Selective breeding of quantitative traits in the common carp (*Cyprinus carpio*): A review. Aquatic Living Resources 16(5):399-407.

Vandeputte, M., M. Kocour, S. Mauger, M. Dupont-Nivet, D. De Guerry *et al.*, 2004 Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). Aquaculture 235: 223–236.

Vandeputte, M., M. Kocour, S. Mauger, M. Rodina, A. Launay *et al.*, 2008 Genetic variation for growth at one and two summers of age in the common carp (*Cyprinus carpio* L.): Heritability estimates and response to selection. Aquaculture 277: 7–13.

VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci. 91: 4414–23.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir, 2012 Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. (Camb). 94: 73–83.

Xu, P., X. Zhang, X. Wang, J. Li, G. Liu *et al.*, 2014 Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. Nat Genet 46(11):1212-9.

Zhang, X., D. Lourenco, I. Aguilar, A. Legarra, and I. Misztal, 2016 Weighting strategies for single-step genomic BLUP: An iterative approach for accurate calculation of GEBV and GWAS. Front. Genet. 7:151.

Zhou, T., S. Liu, X. Geng, Y. Jin, C. Jiang *et al.*, 2017 GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance. Mol. Genet. Genomics 292: 231–242.

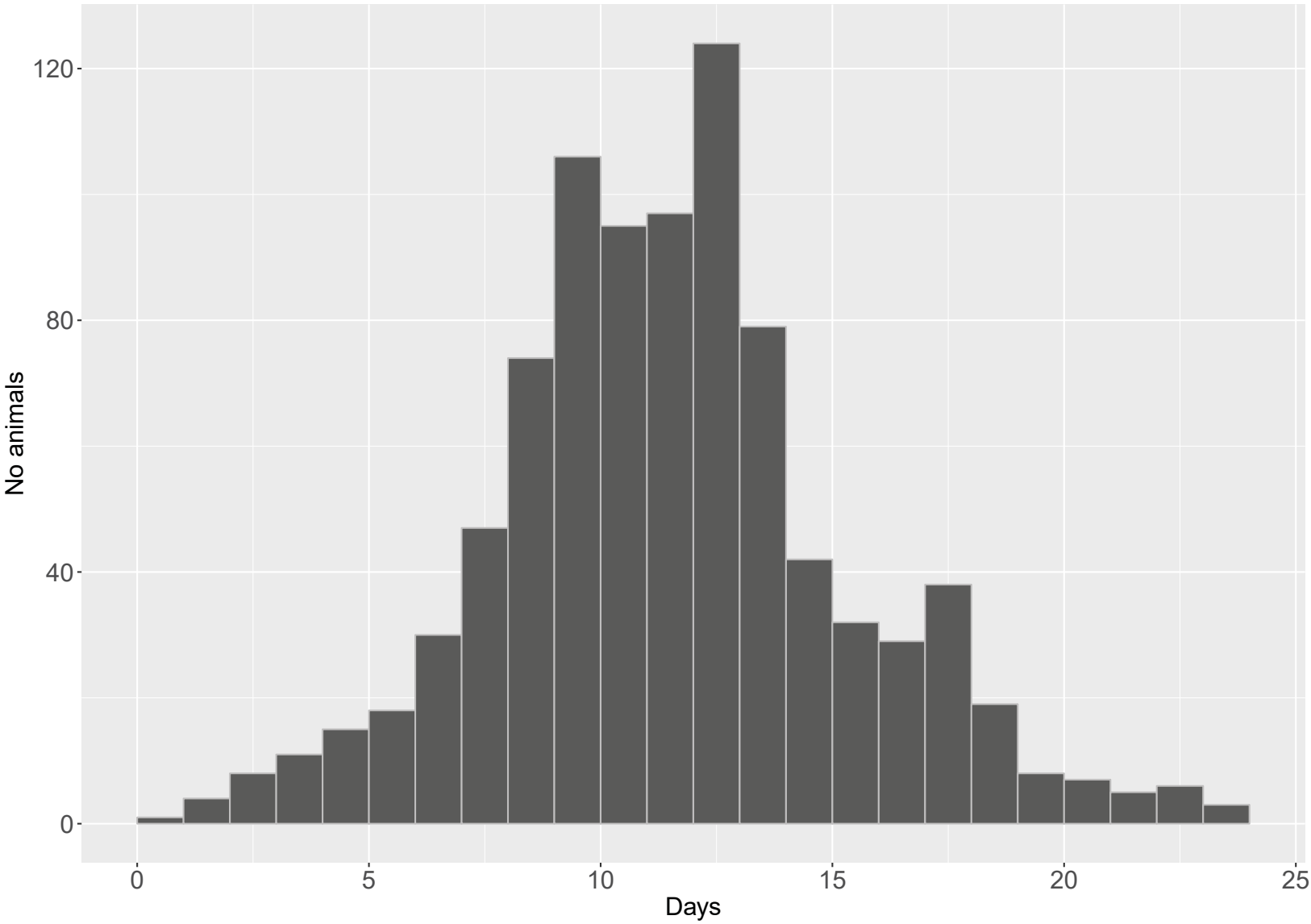
Figures legends

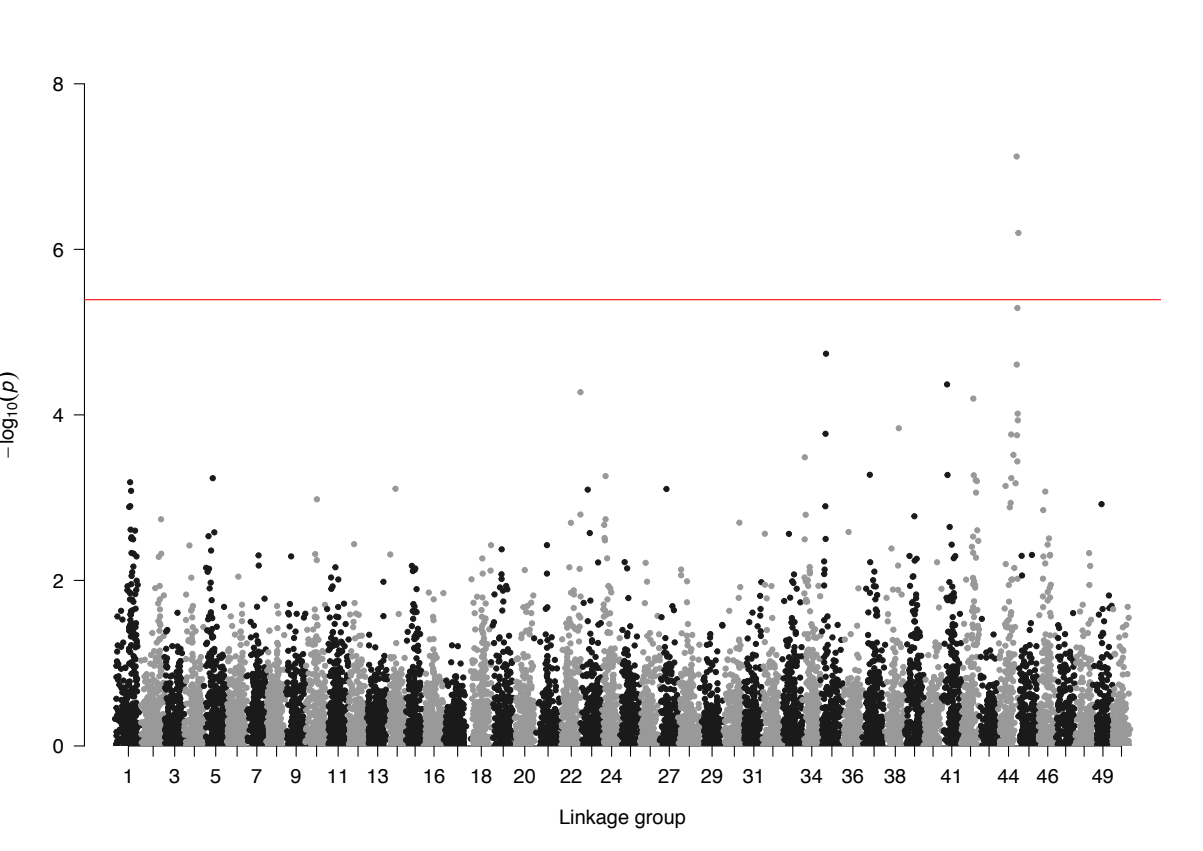
Figure 1. Daily mortality levels of fish during the KHV challenge experiment.

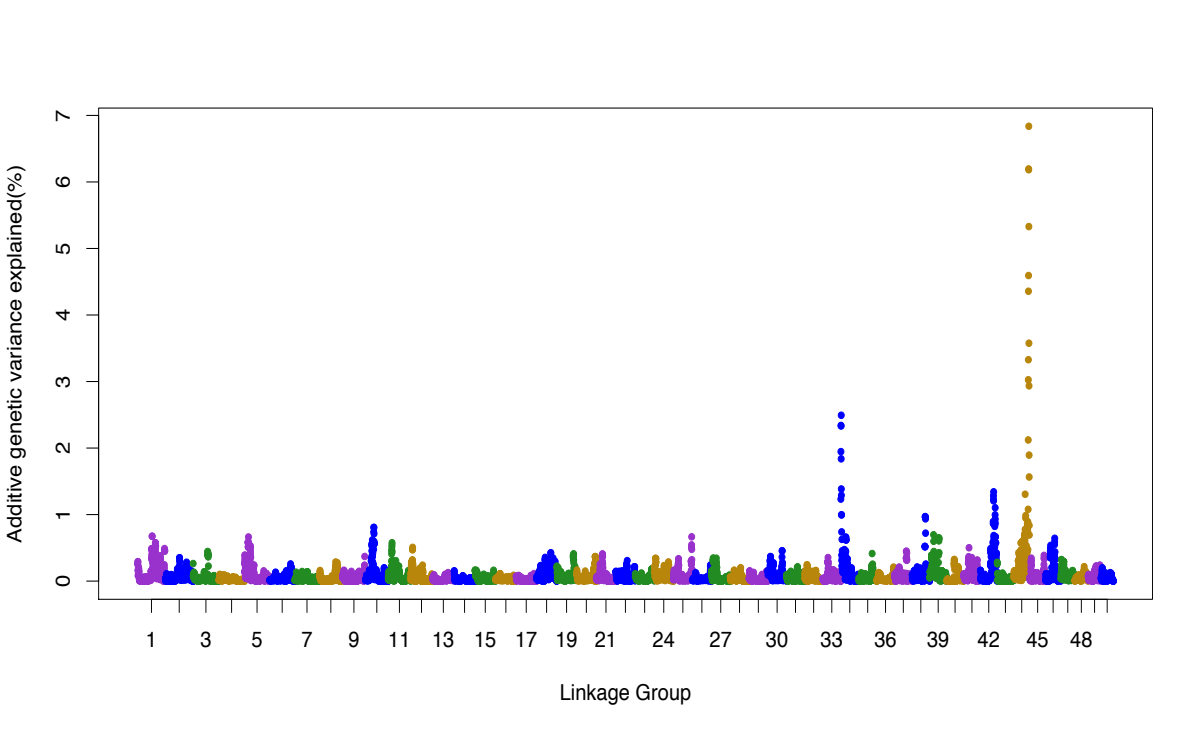
Figure 2. Classical Genome wide association plot for overall survival during the KHV challenge.

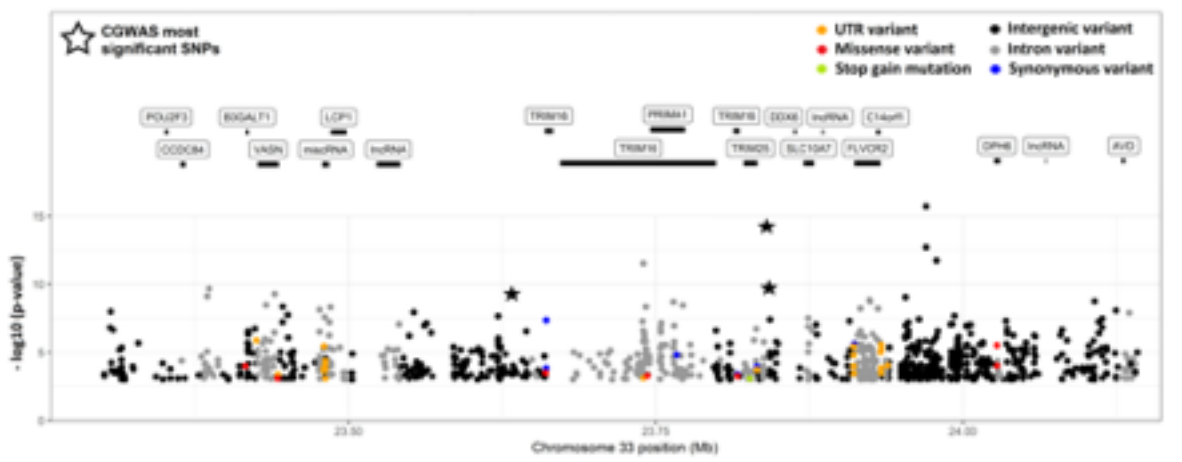
Figure 3. WGBLUP for resistance to KHV. The additive genetic variance explained was calculated using windows of 10 adjacent SNPs.

Figure 4. Annotation of the QTL region on LG 44 including identification of putative genes in the region, functional annotation of SNPs.


120
80
40
0
No animals
0
5
10
15
20
25
Days


8
6
4
2
0
$-\log_{10}(p)$
1
3
5
7
9
11
13
16
18
20
22
24
27
29
31
34
36
38
41
44
46
49
Linkage group


Additive genetic variance explained(%)
0
1
2
3
4
5
6
7
1
3
5
7
9
11
13
15
17
19
21
24
27
30
33
36
39
42
45
48
Linkage Group


CGWAS most significant SNPs
UTR variant
Missense variant
Stop gain mutation
Intergenic variant
Intron variant
Synonymous variant
POU2F3
B3GALT1
LCP1
CCDC84
VASN
miscRNA
lncRNA
TRIM16
PRIM1
TRIM16
TRIM16
DDX6
lncRNA
C14orf1
TRIM25
SLC10A7
FLVCR2
OPHS
lncRNA
AVD
- log10 (p-value)
23.50
23.75
24.00
Chromosome 33 position (Mb)