scientific reports

OPEN

Identification of candidate genes associated with bacterial and viral infections in wild boars hunted in Tuscany (Italy)

M. C. Fabbri[1]✉, A. Crovetti[1], L. Tinacci[2], F. Bertelloni[2], A. Armani[2], M. Mazzei[2], F. Fratini[2], R. Bozzi[1] & F. Cecchi[2]

Wild boar (*Sus scrofa* L.) is one of the large mammals most spread worldwide, highly adaptable, and its population rapidly increased in many areas in Europe, including Italy, where Tuscany is considered particularly suitable for wild boar. Wild boars are potential hosts for different etiological agents, such as *Brucella* spp., *Leptospira* spp. and Pseudorabies virus and they can contribute to maintain and/or to disseminate some bacterial or viral pathogens to humans and domestic animals, above all-in free-range farms. In order to identify hypothetical genomic regions associated with these infection diseases, 96 samples of wild boars hunted in Tuscany during the 2018–2019 and 2019–2020 hunting seasons were considered. Diagnosis was achieved by serological tests and 42 Pseudorabies, 31 Leptospira and 15 Brucella positive animals were identified. All animals were genotyped with Geneseek Genomic Profiler Porcine HD (70 k) and a genome-wide scan was then performed. Significant markers were highlighted for Pseudorabies (two SNPs), Brucella (seven SNPs), and Leptospira (four SNPs) and they were located within, or nearby, 29 annotated genes on chromosome 6, 9, 12, 13, 14 and 18. Eight genes are implicated in viral (SEC14L1, JMJD6, SRSF2, TMPRSS2, MX1, MX2) or bacterial (COL8A1, SPIRE1) infections, seven genes (MFSD11, METTL23, CTTNBP2, BACE2, IMPA2, MPPE1 and GNAL) are involved in mental disorders and one gene (MGAT5B) is related to the Golgi complex. Results presented here provide interesting starting points for future research, validation studies and fine mapping of candidate genes involved in bacterial and viral infections in wild boar.

The wild boar (*Sus scrofa* L.) is widely distributed throughout Eurasia from Europe to the Far East, including SouthEast Asia, and extending to North Africa[1]; it is considered the second most abundant ungulate species in Europe[2].

In Italy, the wild boar population is widely diffused[3], reaching high-density levels in specific areas[4] and Tuscany is particularly suited to the wild boar. This is evidenced by the high number of animals hunted in this area[3–5].

The high density of wild boar in a particular area is a serious problem for the agricultural economy, causing extensive damage to croplands[6] and may represents a severe hazard for both animals and human health.

In fact, it is known that wild boar can be the host for different etiological agents, thus contributing to maintaining and/or disseminating important zoonotic diseases[7], as well as leptospirosis and brucellosis. Furthermore, other not-zoonotic diseases, such as the Pseudorabies (PrV) or Aujeszky's disease, also spread by wild boar have a large economic impact on the swine industry.

Brucella is a zoonotic Gram-negative bacterium. Among the different species, *B. abortus*, *B. suis*, and rarely, *B. melitensis* can infect swine as well as wild boar (*Sus scrofa*)[8–10]. Tuscany, as many other Italian regions, is free from bovine and ovine brucellosis from several years, thanks to the progress of the eradication plan implemented throughout the country since 1992 and 1994 (D.M. 2/7/92 n. 453; D.M. 27/8/94 n. 651; EFSA 2017). *B. suis* biovars 1 and 3 are rarely reported in Europe, while *B. suis* biovar 2 (bv. 2) is largely diffused in East Europe, and it was recently introduced in Italy where it was isolated from domestic pigs and wild boar[8,10,11]. Wild boar represents one of the main reservoirs of *B. suis* bv. 2, which is responsible for reproductive disorders such as infertility, abortion, stillbirths, decreased litter size, weak piglets, orchitis and epididymitis in males, and focal abscess formation[10].

[1]Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali, Università di Firenze, Firenze, Italy. [2]Dipartimento di Scienze Veterinarie, Università di Pisa, Pisa, Italy. ✉email: mariachiara.fabbri@unifi.it

Infection	Total analysed animals	Positive animals	Prevalence (%)	95% CI
Pseudorabies	96	42	48.9	36.58–50.92
Brucella	96	15	15.6	10.28–20.37
Leptospira	96	31	32.3	25.54–39.05

Table 1. Number of positive animals and prevalence of Pseudorabies, Brucella and Leptospira in the 96 animals collected from Tuscany area.

Recently, *B. suis* bv. 2 has been detected also in cows, in which seroconversion was detected without the presence of clinical signs[12,13]. Human infections by this serovar are rarely reported[11].

Leptospirosis represents a re-emerging worldwide zoonotic disease; it is caused by *Leptospira* spp., a Gram-negative spirochetal bacterium. The genus *Leptospira* is traditionally divided into more than 260 antigenically-different serovars, at present classified as pathogenic, intermediate, and saprophytic, with different levels of pathogenicity for animals and humans[14,15]. Swine are the reservoir host for some serovars (Pomona, Tarassovi, and Bratislava), but it is not excluded that it could be infected by many others serovars[16]. Wild boars, due to their natural behaviour and in relation to the geographical area where they live, are often infected by several serovars such as Icterohaemorrhagiae, Grippotyphosa and Canicola[17–19]. As demonstrated by several studies, Central Italy, and in particular Tuscany, present some environmental and geographic features that promote the *Leptospira* spread. Wild boar and feral pig live in contact with domestic animals and humans, representing one of the most important *Leptospira* reservoir among wildlife. For this reason, wild boar is considered an "indicator" of leptospirosis in those area where many different species are forced to co-exist[20–22].

Pseudorabies or Aujeszky's disease (PrV) is caused by Suid herpesvirus 1 which belongs to the Herpesviridae family, subfamily Alphaherpesvirinae, genus Varicellovirus[23]. This virus, as the other members of the afore mentioned family, is enveloped with a double-stranded DNA genome. Although domestic swine and wild boar represent the natural host of PrV, the virus can also infect numerous wild and domestic mammals including ruminants, carnivores, and rodents[24,25].

Pseudorabies virus circulates in domestic swine and wild boar (*Sus scrofa*) populations in several countries[26]. The PrV wild boar seroprevalence in Europe ranges from 4 to 66%, representing a risk of infection for domestic swine and other susceptible animals[26–29]. Consequently, the role of wild boars in the epidemiology of the Pseudorabies is of primary importance, because it represents a serious threat to the completion of the European Community eradication program.

Further knowledge on alternative methods to control the spread of the disease should therefore be investigated. The use of disease-resistant livestock breeds that reduce infection pressure with a decreased incidence of disease is therefore of considerable importance[30].

Virulence of pathogens and susceptibility of wildlife species to disease are influenced by attributes of the host and the pathogen. Discovering the factors that determine host susceptibility/resistance to disease contributes to prevent diseases in both human and domestic animals and improve "global health". A wide range of genetic variations in disease resistance has been observed in swine regarding different viral, bacterial, and parasitic diseases[30].

Previous studies have reported that genetic diversity of individuals and populations of wildlife, at both neutral loci and functional genes, is related to disease tolerance or resistance. Indeed, high levels of inbreeding[31,32] with a consequent heterozygosity reduction at neutral microsatellite markers can increase susceptibility to disease in wildlife[33–35]. Acevedo-Whitehouse et al.[36] highlighted that a higher genetic heterozygosity was associated with lower probability of infection of bovine tuberculosis in wild boar.

The major histocompatibility complex (MHC), a highly polymorphic family of vertebrate genes involved in initiation and regulation of the immune response, has been the target of considerable investigation of disease resistance and tolerance in wildlife[37,38]. Recent advances in genome sequencing, and particularly in the development of high-density single nucleotide polymorphism (SNPs) arrays, have improved genome-wide screening and, therefore, our ability to detect disease-associated genes[39] and to identify the genetic control of disease resilience[40]. For example, in wild boar, Queirós et al.[39] found candidate genes (i.e. LOC102164072, BDNF/NT-3, NTRK2, CDH8, IGSF21) for host genetic susceptibility to tuberculosis.

Given the importance of these pathogens in animals and considering the potential risk factor for human diseases, our study is one of the first attempts to identify genomic regions associated with infections from Pseudorabies virus, Brucella spp. and Leptospira spp. in wild boar.

Results

Table 1 reports the prevalence of positive samples over the total of 96 for Pseudorabies, Brucella and Leptospira. Positivity to two etiological agents was detected only in 5/96 samples (5.21%; 95% CI 2.00–8.42%). Eight were the negative samples. To calculate the confidence interval (CI) and to assess the prevalence, a binomial logistic regression was performed.

Genome-wide scan was performed, and 13 significant SNPs were identified, as shown in Table 2.

Two SNPs were significantly associated to Pseudorabies (Fig. 1).

ASGA0084173 is situated on chromosome 12, with *SEC14L1* and *MGAT5B* genes upstream located and *JMJD6*, *MXRA7*, *MFSD11*, *METTL23*, *SRSF2* downstream located; WU_10.2_18_30218795 is on chromosome 18, within *CTTNBP2* gene and upstream to *CFTR* gene.

Infection		Name	Chr	Position	Pvalue	Type of Variant	Within	Upstream	Downstream
Pseudorabies	ASGA0084173	rs81341734	12	4,612,583	4.12E−05	Intergenic variant	–	SEC14L1, MGAT5B	JMJD6, MXRA7, MFSD11, METTL23, SRSF2
	WU_10.2_18_30218795	rs339406967	18	28,399,958	3.04E−05	Intron variant	CTTNBP2	–	CFTR
Brucella	MARC0040908	rs81232015	9	58,693,183	1.55E−05	Intron variant	NTM	–	–
	MARC0029225	rs81224033	9	58,592,360	1.79E−05	Intron variant	NTM	–	–
	ALGA0073505	rs80987534	13	190,711,065	1.86E−05	Intergenic variant	–	–	–
	WU_10.2_13_200912860	rs319979903	13	190,770,452	1.86E−05	Intergenic variant	–	–	–
	ASGA0060211	rs81443008	13	204,883,958	2.44E−05	Intron variant	TMPRSS2	BACE2, FAM3B, MX1, MX2	RIPK4
	MARC0024545	rs80808271	13	159,240,192	4.78E−05	Intron variant	COL8A1	CMSS1, FILIP1L	–
	ALGA0072626	rs80896010	13	159,213,442	4.78E−05	Downstream gene variant	–	CMSS1, FILIP1L	COL8A1
Leptospira	H3GA0053117	rs81338390	6	97,242,874	1.12E−05	Upstream gene variant	–	SPIRE1, PRELID3A, AFG3L2, TUBB6	CIDEA, IMPA2, MPPE1, CHMP1B, GNAL, TRNAG- UCC
	WU 10.2 13 10,381,840	rs324503493	13	9,111,717	2.05E−05	Upstream gene variant	ZNF385D	–	–
	H3GA0042130	rs80982803	14	114,457,304	1.99E−05	Intron variant	NEURL1	INA, PCGF6, TAF5, MIR1307, USMG5, PDCD11, CALHM3	SH3PXD2A
	ASGA0066225	rs80944071	14	114,471,571	1.99E−05	Intron variant	NEURL1	INA, PCGF6, TAF5, MIR1307, USMG5, PDCD11, CALHM3	SH3PXD2A

Table 2. List of significant SNPs identified for each disease. The SNP name, the chromosome and pb positions, p-value, type of variant and the list of the candidate genes in the regions flanking the significant markers (± 250,000 nucleotides) are reported.
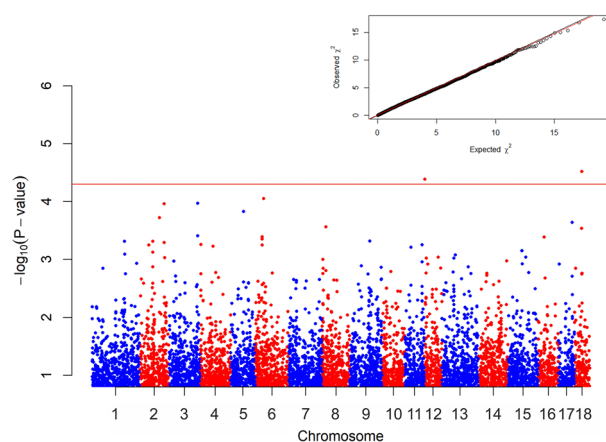
Figure 1. Manhattan plot of the test values obtained, for each marker for Pseudorabies virus. The *horizontal red line* separated the two most significant markers (P < 0.00005).

Several SNPs were identified analysing Brucella infection (Fig. 2), precisely, seven SNPs are found significantly associated with this disease.

Both MARC0040908 and MARC0029225 markers are on chromosome 9 within *NTM gene*. 250 Kbp upstream and 250 Kbp downstream of the aforementioned SNPs no genes are found (Table 2).

The other five significant SNPs are located on chromosome 13: two of them, namely, MARC0024545 and ALGA0072626, are situated in a close region, sharing *CMSS1* and *FILIP1L* genes, which are found upstream to the two SNPs. MARC0024545 is an intron variant because within *COL8A1*, while ALGA0072626 is upstream to the aforementioned gene. ALGA0073505 and WU_10.2_13_200912860 markers are in a genomic region where no characterized genes are present. ASGA0060211 is classified from VeP database as intron variant because it is positioned within *TMPRSS2* gene. *BACE2*, *FAM3B*, *MX1* and *MX2* genes are upstream and *RIPK4* is downstream to the SNPs.

Figure 3 described the four significant associated SNPs for Leptospira.

The H3GA005311 is located on chromosome 6, in a genomic window rich of genes: *SPIRE1*, *PRELID3A*, *AFG3L2* and *TUBB6* are upstream and *CIDEA*, *IMPA2*, *MPPE1*, *CHMP1B*, *GNAL* and *TRNAG-UCC* are
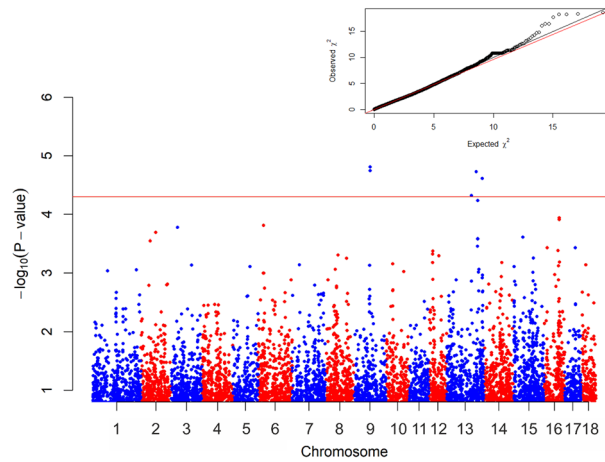
Figure 2. Manhattan plot of the test values obtained, for each marker for Brucella. The *horizontal red line* separated the seven most significant markers (P < 0.00005). ALGA0073505, WU 10.2 13 200912860, MARC0024545 and ALGA0072626 on Chromosome 13 have an extremely closed position, difficult to separate graphically.
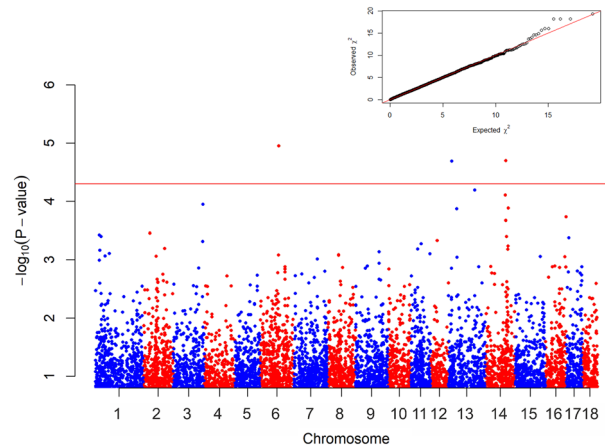

Figure 3. Manhattan plot of the test values obtained for Leptospira. The *horizontal red line* separated the four most significant markers (P < 0.00005). SNP H3GA0042130 and ASGA0066225 on Chromosome 14 have an extremely closed positions, difficult to separate graphically.

downstream of this marker. WU_10.2_13_10381840 has been found within *ZNF385D* gene. H3GA0042130 and ASGA0066225 are close to each other with less than 15 Kbp between the two SNPs. Both are located on *NEURL1* gene, with *INA*, *PCGF6*, *TAF5*, *MIR1307*, *USMG5*, *PDCD11*, *CALHM3* upstream and *SH3PXD2A* downstream genes.

STRING software was used to analyse the possible interactions among proteins encoded by identified genes. Figure 4 showed the results for each infection.

Three of the nine genes (*JMJD6*, *MGAT5B* and *METTL23*, Fig. 4a) for Pseudorabies infection are linked. Brucella phenotype showed the greater interactions among genes: two clusters are identified, the first between *FILIP1L* and *CMSS1*, and the second one among *MX1*, *MX2*, *RIPK4*, *TMPRSS2*, *BACE2* and *FAM3B* genes (Fig. 4b). Although Leptospira was the infection with more genes found in the significant genomic window, few genes are linked to each other (*IMPA2*, *MPPE1*, and *GNAL*—*TAF5* and *USMG5*, Fig. 4c).

Using the PANTHER software, it was possible to summarize the biological processes and molecular functions in which the identified candidate genes are involved (Table 3). For all the infections, the functional genes were enriched in "GO: 0065007, biological regulation", "GO: 0009987, cellular process", "GO: 0051179, localization" and "GO: 0008152, metabolic process". Furthermore, the functional genes of Leptospira were enriched in many other processes including "GO: 0000003, reproduction" and "GO: 0022414, reproductive process". More molecular functions were detected for genes involved in Leptospira infection than for Pseudorabies and Brucella infections.
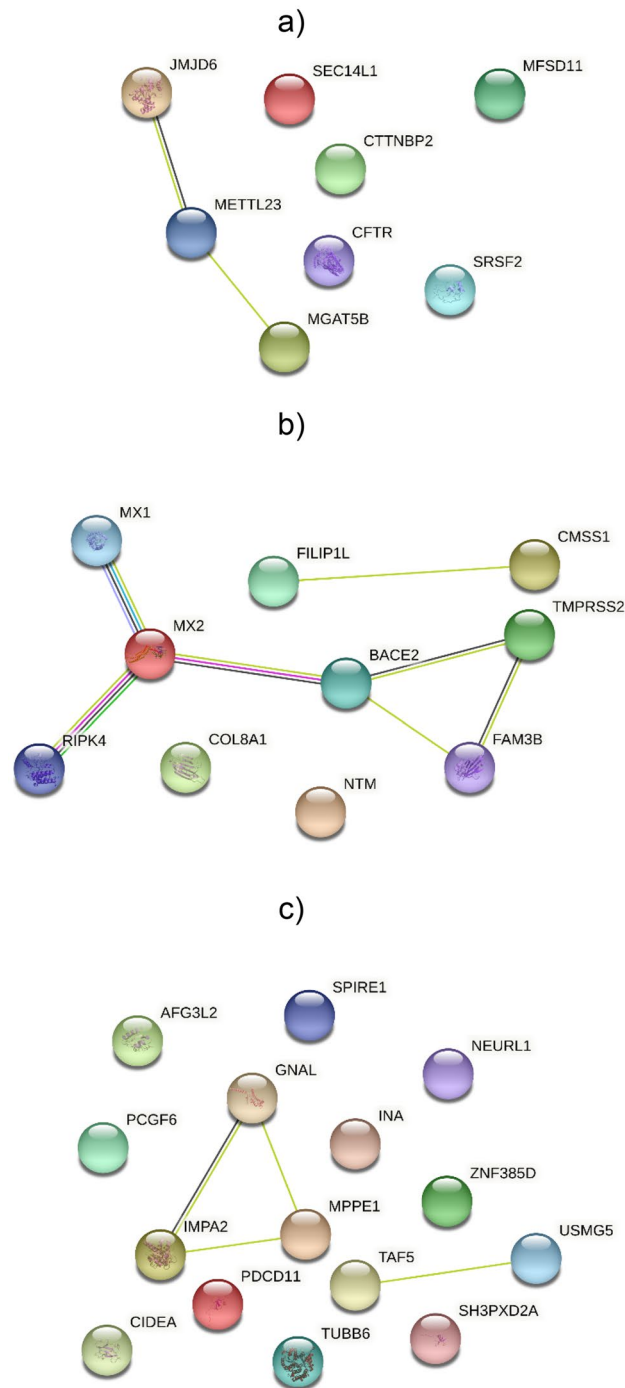
Figure 4. Protein networks of GWAS genes according to STRING database. (**a**) Pseudorabies; (**b**) Brucella; (**c**) Leptospira. Nodes are proteins; lines indicate interactions between proteins with: pink lines for known interactions experimentally determined, sea blue for interactions derived by curated databases. For the predicted interactions, green is for neighbourhood gene, red lines for gene fusions and blue lines for gene co-occurrence. Black lines are for co-expression, light green lines for text mining and light blue lines for protein homology. Protein interactions include direct (physical) and indirect (functional) associations derived from different sources (genomic context, high through-put experiments, conserved co-expression, previous knowledge).

Discussion

The seroprevalence estimate obtained for the three infections were higher than those observed in previous studies. In particular, the presence of PrV (48.9%) in our population is higher than what reported in the same area

Infection	Biological process GO terms	Molecular function GO terms
Pseudorabies	GO: 0065007, biological regulation; GO: 0009987, cellular process; GO: 0051179, localization; GO: 0008152, metabolic process	GO: 000548, binding; GO: 0003824, catalytic activity; GO: 0005215, transporter activity
Brucella	GO: 0065007, biological regulation; GO: 0009987, cellular process; GO: 0051179, localization; GO: 0008152, metabolic process	GO: 000548, binding; GO: 0003824, catalytic activity; GO: 0005215, transporter activity
Leptospira	GO: 0065007, biological regulation; GO: 0009987, cellular process; GO: 0051179, localization; GO: 0008152, metabolic process; GO: 0032502, developmental process; GO: 0051704, multi-organism process; GO: 0032501, multicellular organismal process; GO: 0000003, reproduction; GO: 0022414, reproductive process; GO: 0050896; response to stimulus; GO: 0023052, signalling	GO: 000548, binding; GO: 0003824, catalytic activity; GO: 0005215, transporter activity; GO: 0005198, structural molecule activity; GO: 0098772, molecular function regulator

Table 3. Biological process and molecular function Gene Ontology (GO) terms enrichment analysis results.

by Bertelloni et al.[41] (28.6%) and in previous studies conducted in North-West Italy, indicating the importance of this animal species in the endemicity of Aujeszky's disease[27]. The seroprevalence of PrV in wild boar has been already described in many European countries (France, Germany, Spain, Slovenia, Croatia, and Czech Republic) ranging from 4 to 66%[26,28,42–46].

This investigation, point out also a higher seroprevalence of Brucella than that reported in other studies carried out in past years in Italy (0.00%[47]; 0.53%[41], as well as for Leptospira in Italy (6.0%[47]; 15.3%[48]; 9.5%, 8.8%[41]) and in Europe (3.1%[17]; 10.4%[19]).

The results of the GWAS suggest that some genes might have a role in the resistance/susceptibility to the studied bacterial and viral diseases in wild boar, but no gene was found common to the three infections.

The genomic window 250 Kbp downstream and upstream to the significant SNPs contained a variable number of genes, ranged from 9 (Pseudorabies) to 20 (Leptospira), as reported in Table 2. Interesting is that several SNPs are close to each other, suggesting being in linkage disequilibrium and probably fixed. Two pairs of SNPs were found very close in Brucella infection, ALGA0073505—WU_10.2_13_200912860 (~ 182 Kbp of distance) and MARC0024545—ALGA0072626 (~ 26 Kbp of distance). The closest SNPs are H3GA0042130 and ASGA0066225, identified in Leptospira analysis, distant less than 15 Kbp.

Seven genes associated with Pseudorabies and involved in multiple biological processes are located on chromosome 12. For instance, the *MGAT5B* gene (alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyl-transferase B) is related to the Golgi apparatus operating at the intersection of the secretory, lysosomal, and endocytic pathways. Several studies[49,50] indicated that viral PrV envelopment and tegumentation occurs at late Golgi or post-Golgi compartments, suggesting that this gene may have a role in PrV virulence and dissemination. As reported in Fig. 4a, this gene is linked to other two genes of interest, the *JMJD6* (JmjC domain-containing protein) and the *METTL23* (methyltransferase like 23)**.** The *JMJD6* gene has many functions, ranging from a cell surface receptor for recognition of apoptotic cells to a nuclear factor responsible for lysine hydroxylation and arginine demethylation[51]. Recent reports indicating a multifactorial role in foot-and-mouth disease virus (FMDV) infection[52], in tumorigenesis and virological interactions[53]. *METTL23* belongs to a family of methyltransferase like proteins (*METTL*) that transfer methyl group to various substrates and it is involved in human intellectual disability[54]. The same functionality has been attributed to the *MFSD11* (major facilitator superfamily domain containing 11) gene[55].

On the same chromosome *SEC14L1* gene was interestingly detected (SEC14 like lipid binding, also called *PRELID4A*) and it interacts with RIG-I, a cytosolic pattern recognition receptor, which has found to be required for the activation of anti-PrV activity[56].

MXRA7 (matrix remodelling associated 7) is highly expressed in murine and human ocular tissues and might play a role in pathological processes or diseases involving injury, neovascularization and wound healing[57]. Another study[58] reports evidence for a protective role in the mouse psoriatic epidermis. Moreover, *MXRA7* gene might be involved in bone marrow mesenchymal stem cells (BMSCs) functions[59].

SRSF2 (serine and arginine rich splicing factor 2; also known as SFRS2) gene plays vital roles in a number of biological and pathological processes and it is associated in humans with the progression of a variety of diseases, including viral infection and tumorigenesis[60].

On chromosome 18, *CFTR* (Cystic fibrosis transmembrane conductance regulator) and *CTTNBP2* (cortactin binding protein 2) genes were detected. It is interesting the role of CTTNBP2 in cattle and humans. In Brown Swiss cattle breed this gene is involved in a recessive neurological disease[61], i.e. the Bovine Progressive Degenerative Myeloencephalopathy (Weaver Syndrome). It is considered as a good candidate gene in humans for a role in the pathogenesis of mental retardation[62] such as autism-like behaviours[63].

Genes linked to nervous system are probably expected to be found, because neurological signs predominate with Pseudorabies disease progression, e.g. ataxia, circling, paresis and paralysis[64].

Several genes associated with host resistance or susceptibility to Brucella spp. have been identified in cattle[65], buffalo[66], goats[67,68], humans and pigs[69], from those different type of Collagenases were found. In this research the Collagen type VIII Alpha 1 chain (*COL8A1*) is identified. *COL28A1* has been previously associated with antibody response in feral swine (Sus scrofa) infected with Brucella suis[69], suggesting that *COL8A1* might play an analogue role in wild board.

Genes implicated in viral infections are *MX1* (Interferon-induced GTP-binding protein Mx1) and *MX2* (Interferon-induced GTP-binding protein Mx2)[70]. Mx proteins are interferon (IFN)-induced dynamin-like GTPases that are present in all vertebrates and are known to inhibit the multiplication of several viruses[71], including vesicular stomatitis virus (VSV)[72], influenza A virus (FLUAV)[73] and classical swine fever virus (CSFV)[74–76].

Furthermore, recent studies have shown that MX1 inhibits the replication of foot-and-mouth disease virus (FMDV) (as *JMJD6* found associated with Pseudorabies), and bovine viral diarrhoea virus (BVDV)[77]. Finally, porcine *MX2* was also found to have the antiviral activity against Porcine reproductive and respiratory syndrome virus (PRRSV)[78].

As STRING software highlighted, the Mx genes are flanked by *BACE2*, *TMPRSS2*, *FAM3B* and *RIPKA* genes. The *TMPRSS2* (transmembrane serine protease 2) gene encodes a serine protease that can process the influenza A virus hemagglutinin into its fusion-competent state in human airway epithelial cells[79] and it is mainly involved in SARS-CoV and SARS-CoV-2 infections[80].

The second cluster that STRING identified for Brucella, was formed by *CMSS1* and *FILIP1L* genes, but no studies describing these two genes were found.

NTM (Neurotrimin) gene encodes a member of the IgLON (*LAMP*, *OBCAM*, *NTM*) family of immuno-globulin (Ig) domain-containing glycosylphosphatidylinositol (*GPI*)-anchored cell adhesion molecules. A study performed in humans suggested that *NTM* gene is associated with the level of the intelligence quotient (IQ) and genome wide association studies identified an association between *NTM* variation and cognitive function performances in humans[81,82].

Among the genes associated with Leptospira disease here highlighted, several of them are previously found to be associated with human neurological disorders: *IMPA2*, *GNAL*, *MPPE1* and *AFG3L2*. *IMPA2* (CIDE-N domain-containing protein) gene has been associated with bipolar disorder, schizophrenia[83] and febrile seizure[73]. *GNAL* (G protein subunit alpha L) gene has been linked to bipolar disorder and schizophrenia[84]. *MPPE1* (Metal-lophosphoesterase 1) clustered with *IMPA2* and *GNAL* genes through STRING software (Fig. 4c). Dysregulation of protein phosphorylation and subsequent abnormal cellular signalling has been postulated to be involved in neuropsychiatric disorders, thus making *MPPE1* a plausible biological candidate gene for bipolar disorder (BPD)[85]. *AFG3L2* (AFG3 Like Matrix AAA Peptidase Subunit 2) gene is a candidate gene for hereditary spastic paraplegias or neurodegenerative disorders (https://www.genecards.org/).

These results are interesting because human patients affected from Leptospira reported neurological mani-festations in only 10–15% of cases[86]. The associations here identified suggest that nervous system could be more involved in wild boar Leptospira infections than in humans.

Another interesting gene is *SPIRE1* (KIND domain-containing protein) which belongs to the SPIRE family that emerged as a class of host cell factors that may affect the invasion process. Interestingly, *SPIRE1* has been implicated in the infection of *Salmonella typhimurium*[87].

Other two genes of interest are the *PRELID3A* (PRELI/MSF1 domain-containing protein, also called *SLMO1*) and *TUBB6* (Tubulin Beta 6 Class V). In particular, the PRELI-like family proteins acted as lipid transporters and play an important role of embryonic and development lymphocyte differentiation. The PRELI-like family proteins have been proposed to involve many cellular functions including apoptosis, cellular lipid metabolism and cellular signalling and were correlated with several types of diseases and malignant tumours[88]. *TUBB6* gene was found associated with muscle differentiation and regeneration[89].

In mouse, *CIDEA* (Cell Death Inducing DFFA Like Effector A) gene regulates thermogenesis, lipolysis, and conservation of energy and it is considered to be a proapoptotic factor[90]. The *SH3PXD2A* gene was studied in mouse and humans defining it as a potential risk gene for orofacial clefting, indeed, Cejudo-Martin et al.[91] argued that disruption of the mouse *SH3PXD2A* gene was associated with complete cleft of the secondary palate in 50–90% of mutant mice.

CALHM3 (Calcium Homeostasis Modulator 3) gene together with *CALHM1*, form a complex to mediate rapid taste neurotransmission; indeed, genetic deletion of *CALHM3* abolished sweet, bitter, and human taste perception[92]. Also, *NEURL1* (neuralized E3 ubiquitin protein ligase 1) gene has many functions, in particularly it was related to the cellular process involved in reproduction in multicellular organism[93]. Moreover, *NEURL1* has been associated with fat content in Nordic cattle breeds[94], while in humans this gene was related to survival in Oesophageal adenocarcinoma (EAC) patients[70].

On the same chromosome *PCGF6* (polycomb group ring finger 6) was found, which plays an essential role in embryonic development of mice and in mouse fertility[95].

In summary, significant SNPs were detected to be associated with viral and bacterial disease. Furthermore, among the 29 genes highlighted, 18 genes could be considered candidate genes for genetic resistance or sus-ceptibility to diseases. Indeed, identified genes are implicated in viral (*SEC14L1*, *JMJD6*, *SRSF2*, *TMPRSS2*, *MX1*, *MX2*) bacterial (*COL8A1*, *SPIRE1*), and neurological disorders (*MFSD11*, *METTL23*, *CTTNBP2*, *BACE2*, *IMPA2*, *MPPE1* and *GNAL*), or in the functions of the Golgi complex (*MGAT5B*), organelle where viral envelope is occurred. No candidate genes related to reproduction system were identified for Leptospirosis and Brucel-losis, but it could be hypothesized that wild boar responses are slightly different from those reported on reared pigs and only with a greater sample size it would be possible to individuate the association. In addition, the interesting findings of genes not directly related to infection symptoms, are intriguing, suggesting that further studies are needed to better clarify the pathways of these diseases. Results presented here represent interesting areas for future research, validation studies and fine mapping of candidate genes involved in bacterial and viral infections in wild boar.

Methods

Statement of animal rights.
The wild boars were not hunted for the purpose of this study and none of the authors were involved with the hunting. Animals were hunted following regional hunting laws (Regolamento di attuazione della legge regionale 12 gennaio 1994 n 3 DPGR 48/R/2017). Thus, in accordance to the 2010/63/EU guide and the adoption of the Law D.L. 04/03/2014, n.26 by the Italian Government, an ethical approval was not required for this study.

Sample collection. 96 wild boars (54 females and 42 males) hunted in Tuscany during the 2018–2019 and 2019–2020 hunting seasons (from November to January) were sampled. The study was part of the project PRA_2018_56 financed by University of Pisa and entitled "Evaluation of hygienic-sanitary and qualitative parameters of wild boars hunted in Tuscany and Liguria"[14,29,41], which had the purpose to investigate the role of the wild boar in the epidemiology of some infectious diseases for livestock and humans.

Animals were hunted in different areas in the provinces of Pisa (34, from 5 different localities), Siena (20, from 2 localities), Grosseto (35, from 11 localities), and Livorno (7, from 5 localities), characterised by the abundant presence of wild boars and other wild ungulates[96]. At postmortem examination, samples did not present relevant lesions related to infectious disease. During necroscopy, the kidneys were collected, and serum was extracted from the infraorbital cavities[97].

Serological analysis for Pseudorabies, Brucella and Leptospira infections. Serum samples were analysed by ID Screen Aujeszky gB competitive kit detecting anti-gB PrV antibodies (ID.vet, Grabels, France). Test procedures and interpretation of results were performed according to the manufacturer's instructions, adopting the short serum incubation protocol. The optical density was measured by a plate reader (Multiscan FC; Thermo Scientific, Waltham, MA, USA) at 450 nm wave-length.

Rose Bengal Test (RBT) and complement fixation test (CFT) were employed to detect anti-brucella antibodies. RBT and CFT were performed as described by World Organization for Animal Health (OIE) (OIE 2016); antigens used in both tests were obtained from the "Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise G. Caporale, Teramo".

Leptospira antibodies were detected by microscopic agglutination test (MAT) as previously described[11], titers of 1:100 were considered positive. The serovars were employed as live antigens in MAT (Supplementary Information). Anti-serum has been used as positive control for each investigated serogroup provided by "Istituto Zooprofilattico Sperimentale della Lombardia ed Emilia Romagna, Brescia", while sterilized saline water was used as a negative control.

DNA extraction and SNPs quality control (QC). Ninety-six wild boars were genotyped using the "Geneseek Genomic Profiler Porcine HD (70 k)", containing 62,330 SNPs. SNP genotyping was outsourced at the Science and Technology Park of Sardinia (Porto Conte Ricerche; https://www.portocontericerche.it/it).

Total DNA extraction for each sample of kidney was performed, starting from about 100 μl of homogenised tissue, according to the salting out procedure proposed by Armani et al.[98], further modified and applied as Salting out reference protocol by Armani et al.[99]. Final DNA concentration and purity were assessed with Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, US) by two subsequent measurements of the absorbance value at 260 nm and calculation of A260/A280 and of A260/230 ratios. 260/280 and 260/230 values ≥ 2 were considered indicative of nucleic acid purity according to the manufacturer's indications (https://tools.thermofisher.com/content/sfs/brochures/TN52646-E-0215M-NucleicAcid.pdf). DNA samples were prepared in a volume of 30 ul per concentration of 50 ng/ul in 96 plates and sent to the laboratory for genotyping.

The SNPs quality control (QC) has performed with PLINK v.1.07 (http://zzz.bwh.harvard.edu/plink/) and only autosomal SNPs with a call rate higher than 95%, a minor allele frequency (MAF) > 1% and with no extreme deviation from Hardy–Weinberg equilibrium (P value > 0.00001) were included in the analysis. Animals with more than 5% of missing genotypes were discarded. After QC, 42,431 SNPs mapped on the 18 porcine autosomes and 93 individuals were retained. The number of SNPs per chromosome is reported in Table 4.

Genome wide association study and gene set enrichment analysis. The analysis was carried out for each infection separately, evaluating healthy vs infected animals. The association analysis was carried out with GenABEL[100], which performs a simple linear regression marker-phenotype analysis. Firstly, the genomic relationship matrix was calculated with the function *ibs* (https://rdrr.io/cran/GenABEL/man/ibs.html), where for a given pair of individuals i and j, the identical by state coefficients ($f_{i,j}$) is calculated as follows:

$$f_{i,j} = \frac{1}{N} \sum_{k} \frac{(x_{i,k} - p_k)(x_{i,k} - p_k)}{p_k(1 - p_k)},$$

where N is the number of markers used, $x_{i,k}$ is the genotype of the ith individual at the kth SNP (coded as 0, ½ and 1), p_k is the frequency of the reference allele and k = 1,…, N.

Then, the additive polygenic model described below was applied:

$$Y = X\beta + a + e.$$

Each phenotype has been analysed separately (affected: Pseudorabies 42; Leptospira 31; Brucella 15); β was a vector with the fixed sex effect and **X** was the incidence matrix that associated each observation to levels of factor in β. The random effects in the model were the animal and the residual, which were assumed normally distributed as $\alpha \sim N\left(0, G\sigma_g^2\right)$ and $e \sim N\left(0, I\sigma_e^2\right)$, where **G** was the genomic relationship matrix, **I** is an identity matrix, and σ_g^2 and σ_e^2 are the additive genomic and residual variances, respectively. Regression was performed using the GenABEL function *mmscore* and the associations between marker and phenotype with a *P* value ≤ 5 × 10⁻⁵ were considered significant[101]. The aforementioned threshold was used considering that the study was carried out on wild boar, which has not a species-specific SNP chip. Moreover, this level of significance for association is called "suggestive", and it was introduced by Lander e Kruglyak[101] and it is widely used in GWAS[102–105]. For each trait, a Manhattan plot and a quantile–quantile (Q–Q) plot were produced using the R

Chromosome	Total markers pre-QC	Total markers post-QC	%Markers post-QC
1	5926	4054	9.55
2	4141	2546	6.00
3	3575	2600	6.13
4	3807	2626	6.19
5	3009	2103	4.95
6	4360	2870	6.76
7	3943	2823	6.66
8	3533	2376	5.60
9	3856	2730	6.43
10	2873	1967	4.63
11	2344	1623	3.82
12	2513	1526	3.59
13	4435	3088	7.28
14	4087	2556	6.02
15	3607	2534	5.97
16	2386	1748	4.12
17	2178	1350	3.18
18	1757	1311	3.09

Table 4. Total number of SNPs before quality control (pre-QC), post quality control (post-QC) and the percentage of SNPs retained for each autosomal chromosome.

software[103] uploading chromosome position from *Sus Scrofa 10.2 assembly* (https://may2017.archive.ensembl.org/Sus_scrofa/Info/Index) to the recent *Sus* Scrofa 11.1 (https://www.ensembl.org/Sus_scrofa/Info/Annotation).

VeP database (https://www.ensembl.org/info/docs/tools/vep/index.html) was used to investigate the type of significant SNPs. A genomic window of 250 Kbp upstream and downstream from the significant SNP for each trait was investigated using the R package biomaRt[106,107], which accesses the data available in Ensembl database (https://www.ensembl.org). The genes identification was based on Scrofa 11.1 assembly (https://www.ensembl.org/Sus_scrofa/Info/Annotation). For the Gene set enrichment analysis, the lists of protein coding genes were uploaded to STRING 11.5[108] and PANTHER v.16.0[109].

Data availability

The datasets analyzed during the current study are available online on the following link: https://osf.io/bz89n/.

References

1. Choi, S. K. *et al.* Asia-wide phylogeography of wild boar (*Sus scrofa*) based on mitochondrial DNA and Y-chromosome: Revising the migration routes of wild boar in Asia. *PLoS ONE* **15**, e0238049 (2020).
2. Putman, R., Apollonio, M. & Andersen, R. *Ungulate Management in Europe: Problems and Practices* (Cambridge University Press, 2011).
3. Massei, G. *et al.* Wild boar populations up, numbers of hunters down? A review of trends and implications for Europe. *Pest Manage. Sci.* **71**, 492–500 (2015).
4. Pittiglio, C., Khomenko, S. & Beltran-Alcrudo, D. Wild boar mapping using population-density statistics: From polygons to high resolution raster maps. *PLoS ONE* **13**, e0193295 (2018).
5. Santilli, F. & Varuzza, P. Factors affecting wild boar (*Sus scrofa*) abundance in southern Tuscany. *Hystrix Ital. J. Mammal.* **24**, 169–173 (2013).
6. Lombardini, M., Meriggi, A. & Fozzi, A. Factors influencing wild boar damage to agricultural crops in Sardinia (Italy). *Curr. Zool.* **63**, 507–514 (2017).
7. Blanchong, J. A., Robinson, S. J., Samuel, M. D. & Foster, J. T. Application of genetics and genomics to wildlife epidemiology. *J. Wildl. Manage.* **80**, 593–608 (2016).
8. Olsen, S. C. & Tatum, F. M. Swine brucellosis: Current perspectives. *Vet. Med. Auckl. N. Z.* **8**, 1–12 (2017).
9. Díaz Aparicio, E. Epidemiology of brucellosis in domestic animals caused by *Brucella melitensis*, *Brucella suis* and *Brucella abortus*. *Rev. Sci. Tech. Off. Int. Épizooties* **32**, 53 (2013).
10. Cilia, G. *et al.* Genital *Brucella suis* biovar 2 infection of wild boar (*Sus scrofa*) hunted in Tuscany (Italy). *Microorganisms* **9**, 582 (2021).
11. Bertelloni, F. *et al.* Serological survey on Leptospira infection in slaughtered swine in North-Central Italy. *Epidemiol. Infect.* **146**, 1275–1280 (2018).
12. Fretin, D. *et al.* Unexpected *Brucella suis* biovar 2 infection in a dairy cow, Belgium. *Emerg. Infect. Dis.* **19**, 2053–2054 (2013).
13. Szulowski, K., Iwaniak, W., Weiner, M. & Złotnicka, J. Brucella suis biovar 2 isolations from cattle in Poland. *Ann. Agric. Environ. Med.* **20**, 672–675 (2013).
14. Cilia, G., Bertelloni, F., Angelini, M., Cerri, D. & Fratini, F. Leptospira survey in wild boar (*Sus scrofa*) hunted in Tuscany, Central Italy. *Pathogens* **9**, 377 (2020).
15. Adler, B. & de la Peña Moctezuma, A. Leptospira and leptospirosis. *Vet. Microbiol.* **140**, 287–296 (2010).
16. Ellis, W. A. Animal leptospirosis. In *Leptospira and Leptospirosis* (ed. Adler, B.) 99–137 (Springer, 2015).

17. Boqvist, S., Bergström, K. & Magnusson, U. Prevalence of antibody to six leptospira servovars in swedish wild boars. *J. Wildl. Dis.* **48**, 492–496 (2012).
18. Vale-Gonçalves, H. M. *et al.* Prevalence of Leptospira antibodies in wild boars (*Sus scrofa*) from Northern Portugal: Risk factor analysis. *Epidemiol. Infect.* **143**, 2126–2130 (2015).
19. Żmudzki, J. *et al.* First overall report of Leptospira infections in wild boars in Poland. *Acta Vet. Scand.* **58**, 3 (2016).
20. Ebani, V., Bertelloni, F., Pinzauti, P. & Cerri, D. Seroprevalence of *Leptospira* spp. and *Borrelia burgdorferi* sensu Lato in Italian horses. *Ann. Agric. Environ. Med.* **19**, 237–240 (2012).
21. Fratini, F. *et al.* The presence of leptospira in coypus (*Myocastor coypus*) and rats (*Rattus norvegicus*) living in a protected wetland in Tuscany (Italy). *Vet. Arh.* **85**, 407–414 (2015).
22. Fratini, F. *Leptospira Infection in Wild Animals*. https://onesearch.unipi.it (Accessed November 2021).
23. Mettenleiter, T. C. Aujeszky's disease (pseudorabies) virus: The virus and molecular pathogenesis–state of the art, June 1999. *Vet. Res.* **31**, 99–115 (2000).
24. Boadella, M., Gortázar, C., Vicente, J. & Ruiz-Fons, F. Wild boar: An increasing concern for Aujeszky's disease control in pigs? *BMC Vet. Res.* **8**, 7 (2012).
25. Ruiz-Fons, F., Segalés, J. & Gortázar, C. A review of viral diseases of the European wild boar: Effects of population dynamics and reservoir rôle. *Vet. J. Lond. Engl.* **1997**(176), 158–169 (2008).
26. Müller, T. *et al.* Characterization of pseudorabies virus of wild boar origin from Europe. *Epidemiol. Infect.* **138**, 1590–1600 (2010).
27. Verin, R., Varuzza, P., Mazzei, M. & Poli, A. Serologic, molecular, and pathologic survey of pseudorabies virus infection in hunted wild boars (*Sus scrofa*) in Italy. *J. Wildl. Dis.* **50**, 559–565 (2014).
28. Moreno, A. *et al.* Detection and molecular analysis of Pseudorabies virus strains isolated from dogs and a wild boar in Italy. *Vet. Microbiol.* **177**, 359–365 (2015).
29. Pacini, M. I. *et al.* Detection of Pseudorabies virus in wild boar foetus. *Anim. Open Access J. MDPI* **10**, 366 (2020).
30. Reiner, G. Investigations on genetic disease resistance in swine—A contribution to the reduction of pain, suffering and damage in farm animals. *Appl. Anim. Behav. Sci.* **118**, 217 (2009).
31. Spielman, D., Brook, B. W., Briscoe, D. A. & Frankham, R. Does inbreeding and loss of genetic diversity decrease disease resistance? *Conserv. Genet.* **5**, 439–448 (2004).
32. Whiteman, N. K., Matson, K. D., Bollmer, J. L. & Parker, P. G. Disease ecology in the Galápagos Hawk (*Buteo galapagoensis*): Host genetic diversity, parasite load and natural antibodies. *Proc. R. Soc. B Biol. Sci.* **273**, 797–804 (2006).
33. Coltman, D. W., Pilkington, J. G., Smith, J. A. & Pemberton, J. M. Parasite-mediated selection against inbred soay sheep in a free-living Island populaton. *Evolution* **53**, 1259–1267 (1999).
34. Cassinello, J., Gomendio, M. & Roldan, E. R. S. Relationship between coefficient of inbreeding and parasite burden in endangered Gazelles. *Conserv. Biol.* **15**, 1171–1174 (2001).
35. Acevedo-Whitehouse, K., Gulland, F., Greig, D. & Amos, W. Disease susceptibility in California sea lions. *Nature* **422**, 35–35 (2003).
36. Acevedo-Whitehouse, K. *et al.* Genetic resistance to bovine tuberculosis in the Iberian wild boar. *Mol. Ecol.* **14**, 3209–3217 (2005).
37. Bernatchez, L. & Landry, C. MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *J. Evol. Biol.* **16**, 363–377 (2003).
38. Acevedo-Whitehouse, K. & Cunningham, A. A. Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol. Evol.* **21**, 433–438 (2006).
39. Queirós, J., Alves, P. C., Vicente, J., Gortázar, C. & de la Fuente, J. Genome-wide associations identify novel candidate loci associated with genetic susceptibility to tuberculosis in wild boar. *Sci. Rep.* **8**, 1980 (2018).
40. Bai, X. *et al.* Investigating the genetic architecture of disease resilience in pigs by genome-wide association studies of complete blood count traits collected from a natural disease challenge model. *BMC Genomics* **22**, 535 (2021).
41. Bertelloni, F. *et al.* Serological survey on bacterial and viral pathogens in wild boars hunted in Tuscany. *EcoHealth* **17**, 85–93 (2020).
42. Vicente, J. *et al.* Antibodies to selected viral and bacterial pathogens in European wild boars from southcentral Spain. *J. Wildl. Dis.* **38**, 649–652 (2002).
43. Sedlak, K., Bartova, E. & Machova, J. Antibodies to selected viral disease agents in wild boars from the Czech Republic. *J. Wildl. Dis.* **44**, 777–780 (2008).
44. Vengust, G., Lindtner-Knific, R., Zele, D. & Bidovec, A. Leptospira antibodies in wild boars (*Sus scrofa*) in Slovenia. *Eur. J. Wildl. Res.* **54**, 749–752 (2008).
45. Cano-Manuel, F. J. *et al.* Long-term monitoring of 10 selected pathogens in wild boar (*Sus scrofa*) in Sierra Nevada National Park, southern Spain. *Vet. Microbiol.* **174**, 148–154 (2014).
46. Caruso, C. *et al.* Serological and virological survey of hepatitis E virus in wild boar populations in northwestern Italy: Detection of HEV subtypes 3e and 3f. *Arch. Virol.* **160**, 153–160 (2015).
47. Ebani, V. V., Cerri, D., Poli, A. & Andreani, E. Prevalence of leptospira and brucella antibodies in wild boars (*Sus scrofa*) in Tuscany, Italy. *J. Wildl. Dis.* **39**, 718–722 (2003).
48. Chiari, M. *et al.* Seroprevalence and risk factors of leptospirosis in wild boars (*Sus scrofa*) in northern Italy. *Hystrix Ital. J. Mammal.* **27**, 145–149 (2016).
49. Omar, O. S., Simmons, A. J., Andre, N. M., Wilson, D. W. & Gross, S. T. Pseudorabies virus and herpes simplex virus type 1 utilize different tegument-glycoprotein interactions to mediate the process of envelopment. *Intervirology* **56**, 50–54 (2013).
50. Nanbo, A., Noda, T. & Ohba, Y. Epstein-Barr virus acquires its final envelope on intracellular compartments with golgi markers. *Front. Microbiol.* **9**, 454 (2018).
51. Chang, B., Chen, Y., Zhao, Y. & Bruick, R. K. JMJD6 is a histone arginine demethylase. *Science* **318**, 444–447 (2007).
52. Lawrence, P. & Rieder, E. Insights into Jumonji C-domain containing protein 6 (JMJD6): A multifactorial role in foot-and-mouth disease virus replication in cells. *Virus Genes* **53**, 340–351 (2017).
53. Vangimalla, S. S., Ganesan, M., Kharbanda, K. K. & Osna, N. A. Bifunctional enzyme JMJD6 contributes to multiple disease pathogenesis: New twist on the old story. *Biomolecules* **7**, 41 (2017).
54. Almannai, M. *et al.* Further delineation of METTL23-associated intellectual disability. *Am. J. Med. Genet. A* **182**, 785–791 (2020).
55. Anazi, S. *et al.* Expanding the genetic heterogeneity of intellectual disability. *Hum. Genet.* **136**, 1419–1429 (2017).
56. Chen, X. *et al.* Pseudorabies virus UL24 antagonizes OASL-mediated antiviral effect. *Virus Res.* **295**, 198276 (2021).
57. Jia, C., Zhang, F., Zhu, Y., Qi, X. & Wang, Y. Public data mining plus domestic experimental study defined involvement of the old-yet-uncharacterized gene matrix-remodeling associated 7 (MXRA7) in physiopathology of the eye. *Gene* **632**, 43–49 (2017).
58. Ning, J. *et al.* Altered expression of matrix remodelling associated 7 (MXRA7) in psoriatic epidermis: Evidence for a protective role in the psoriasis imiquimod mouse model. *Exp. Dermatol.* **27**, 1038–1042 (2018).
59. Zhou, Z. *et al.* Matrix remodeling associated 7 promotes differentiation of bone marrow mesenchymal stem cells toward osteoblasts. *J. Cell. Physiol.* **234**, 18053–18064 (2019).
60. Li, K. & Wang, Z. Splicing factor SRSF2-centric gene regulation. *Int. J. Biol. Sci.* **17**, 1708–1715 (2021).
61. McClure, M. *et al.* Fine mapping for weaver syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. *PLoS ONE* **8**, e59251 (2013).

62. Patel, C. *et al.* Translocation breakpoint at 7q31 associated with tics: Further evidence for IMMP2L as a candidate gene for Tourette syndrome. *Eur. J. Hum. Genet.* **19**, 634–639 (2011).

63. Shih, P.-Y. *et al.* CTTNBP2 controls synaptic expression of zinc-related autism-associated proteins and regulates synapse formation and autism-like behaviors. *Cell Rep.* **31**, 107700 (2020).

64. Sehl, J. & Teifke, J. P. Comparative pathology of pseudorabies in different naturally and experimentally infected species—A review. *Pathogens* **9**, 633 (2020).

65. Yapan, S., Liping, S., Aizhen, G. & Liguo, Y. Effects of toll-like receptor 2 gene mutation on resistance to bovine brucellosis. *Livest. Sci.* **170**, 30–34 (2014).

66. Borriello, G. *et al.* Genetic resistance to *Brucella abortus* in the water buffalo (*Bubalus bubalis*). *Infect. Immun.* **74**, 2115–2120 (2006).

67. Rossi, U. A. *et al.* A haplotype at intron 8 of PTPRT gene is associated with resistance to Brucella infection in Argentinian creole goats. *Vet. Microbiol.* **207**, 133–137 (2017).

68. Rossi, U. A. *et al.* Association of an IRF3 putative functional uORF variant with resistance to Brucella infection: A candidate gene based analysis of InDel polymorphisms in goats. *Cytokine* **115**, 109–115 (2019).

69. Pierce, C. F. *et al.* Loci associated with antibody response in feral swine (*Sus scrofa*) infected with *Brucella suis*. *Front. Vet. Sci.* **7**, 957 (2020).

70. Wang, Y., Liang, N., Xue, Z. & Xue, X. Identifying an eight-gene signature to optimize overall survival prediction of esophageal adenocarcinoma using bioinformatics analysis of ceRNA network. *OncoTargets Ther.* **13**, 13041–13054 (2020).

71. Sasaki, K. *et al.* A single nucleotide polymorphism of porcine MX2 gene provides antiviral activity against vesicular stomatitis virus. *Immunogenetics* **66**, 25–32 (2014).

72. Zhang, X. *et al.* In vitro inhibition of vesicular stomatitis virus replication by purified porcine Mx1 protein fused to HIV-1 Tat protein transduction domain (PTD). *Antiviral Res.* **99**, 149–157 (2013).

73. Nakajima, E. *et al.* A naturally occurring variant of porcine Mx1 associated with increased susceptibility to influenza virus in vitro. *Biochem. Genet.* **45**, 11–24 (2007).

74. Zhao, Y. *et al.* Human MxA protein inhibits the replication of classical swine fever virus. *Virus Res.* **156**, 151–155 (2011).

75. He, D. *et al.* In vitro inhibition of the replication of classical swine fever virus by porcine Mx1 protein. *Antiviral Res.* **104**, 128–135 (2014).

76. Zhou, J. *et al.* Porcine Mx1 protein inhibits classical swine fever virus replication by targeting nonstructural protein NS5B. *J. Virol.* **92**, e02147-e2217 (2018).

77. Yuan, B., Fang, H., Shen, C. & Zheng, C. Expression of porcine Mx1 with FMDV IRES enhances the antiviral activity against foot-and-mouth disease virus in PK-15 cells. *Arch. Virol.* **160**, 1989–1999 (2015).

78. Wang, H. *et al.* The interferon-induced Mx2 inhibits porcine reproductive and respiratory syndrome virus replication. *J. Interferon Cytokine Res.* **36**, 129–139 (2016).

79. Böttcher, E. *et al.* Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *J. Virol.* **80**, 9896–9898 (2006).

80. Bovo, S. *et al.* Describing variability in pig genes involved in coronavirus infections for a One Health perspective in conservation of animal genetic resources. *Sci. Rep.* **11**, 3359 (2021).

81. Pan, Y., Wang, K.-S. & Aragam, N. NTM and NR3C2 polymorphisms influencing intelligence: Family-based association studies. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **35**, 154–160 (2011).

82. Liu, F. *et al.* A genomewide screen for late-onset Alzheimer disease in a genetically isolated Dutch population. *Am. J. Hum. Genet.* **81**, 17–31 (2007).

83. Yoshikawa, T. *et al.* Evidence for association of the myo-inositol monophosphatase 2 (IMPA2) gene with schizophrenia in Japanese samples. *Mol. Psychiatry* **6**, 202–210 (2001).

84. Vuoristo, J. T. *et al.* Sequence and genomic organization of the human G-protein Golfα gene (GNAL) on chromosome 18p11, a susceptibility region for bipolar disorder and schizophrenia. *Mol. Psychiatry* **5**, 495–501 (2000).

85. Lohoff, F. W., Ferraro, T. N., Brodkin, E. S., Weller, A. E. & Bloch, P. J. Association between polymorphisms in the metallophosphoesterase (MPPE1) gene and bipolar disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153**, 830–836 (2010).

86. Mathew, T. *et al.* Neuroleptospirosis—Revisited: Experience from a tertiary care neurological centre from south India. *Indian J. Med. Res.* **124**, 155–162 (2006).

87. Andritschke, D. *et al.* A genome-wide siRNA screen implicates spire1/2 in SipA-driven *Salmonella typhimurium* host cell invasion. *PLoS ONE* **11**, e0161965 (2016).

88. Zhu, Y. *et al.* Lipid metabolism-related proteins of relevant evolutionary and lymphoid interest (PRELI) domain containing family proteins in cancer. *Am. J. Transl. Res.* **12**, 6015–6026 (2020).

89. Randazzo, D. *et al.* Persistent upregulation of the β-tubulin tubb6, linked to muscle regeneration, is a source of microtubule disorganization in dystrophic muscle. *Hum. Mol. Genet.* **28**, 1117–1135 (2019).

90. Viswakarma, N. *et al.* Transcriptional regulation of Cidea, mitochondrial cell death-inducing DNA fragmentation factor alpha-like effector A, in mouse liver by peroxisome proliferator-activated receptor alpha and gamma. *J. Biol. Chem.* **282**, 18613–18624 (2007).

91. Cejudo-Martin, P. *et al.* Genetic disruption of the Sh3pxd2a gene reveals an essential role in mouse development and the existence of a novel isoform of Tks5. *PLoS ONE* **9**, e107674 (2014).

92. Ma, Z. *et al.* calhm3 is essential for rapid ion channel-mediated purinergic neurotransmission of GPCR-mediated tastes. *Neuron* **98**, 547-561.e10 (2018).

93. Sadowska, A. *et al.* Transcriptional profiling of porcine granulosa cells exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Chemosphere* **178**, 368–377 (2017).

94. Iso-Touru, T., Sahana, G., Guldbrandtsen, B., Lund, M. S. & Vilkki, J. Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants. *BMC Genet.* **17**, 55 (2016).

95. Liu, M. *et al.* The polycomb group protein PCGF6 mediates germline gene silencing by recruiting histone-modifying proteins to target gene promoters. *J. Biol. Chem.* **295**, 9712–9724 (2020).

96. Ungulati, B. D. Status, distribuzione, consistenza, gestione e prelievo venatorio delle popolazioni di Ungulati in Italia. *Istituto Superiore per la Protezione e la Ricerca Ambientale*. https://www.isprambiente.gov.it/it/pubblicazioni/documenti-tecnici/banca-dati-ungulati-status-distribuzione (Accessed November 2021).

97. Arenas-Montes, A. *et al.* Blood sampling by puncture in the cavernous sinus from hunted wild boar. *Eur. J. Wildl. Res.* **59**, 299–303 (2013).

98. Armani, A., Castigliego, L., Tinacci, L., Gianfaldoni, D. & Guidi, A. Molecular characterization of icefish, (Salangidae family), using direct sequencing of mitochondrial cytochrome b gene. *Food Control* **22**, 888–895 (2011).

99. Andrea, A. *et al.* Development of a simple and cost-effective bead-milling method for DNA extraction from fish muscles. *Food Anal. Methods* **7**, 946–955 (2014).

100. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: An R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–1296 (2007).

101. Lander, E. & Kruglyak, L. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**, 241–247 (1995).

102. Burton, P. R. *et al.* Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
103. Mousa, M. *et al.* Genome-wide association study of hospitalized COVID-19 patients in the United Arab Emirates. *EBioMedicine* **74**, 103695 (2021).
104. Schut, C. H. *et al.* Identification of single-nucleotide variants associated with susceptibility to Salmonella in pigs using a genome-wide association approach. *BMC Vet. Res.* **16**, 138 (2020).
105. Hammond, R. K. *et al.* Biological constraints on GWAS SNPs at suggestive significance thresholds reveal additional BMI loci. *Elife* **10**, e62206 (2021).
106. Durinck, S. *et al.* BioMart and bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinform. Oxf. Engl.* **21**, 3439–3440 (2005).
107. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
108. Szklarczyk, D. *et al.* The STRING database in 2017: Quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
109. Mi, H. *et al.* PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **49**, D394–D403 (2021).

Acknowledgements

Author contributions

M.C.F. performed the formal analysis and wrote the original draft. A.C. reviewed the manuscript and supervised. F.F. reviewed the manuscript and funding acquisition; L.T., F.B., A.A., F.F., M.M. found resources and reviewed the manuscript. R.B. conceptualized and reviewed the manuscript. F.C. elaborated the methodology and wrote the original draft.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-12353-8.

Correspondence and requests for materials should be addressed to M.C.F.

Reprints and permissions information is available at www.nature.com/reprints.

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