

RESEARCH ARTICLE

Open Access



A novel locus on mouse chromosome 7 that influences survival after infection with tick-borne encephalitis virus

Martin Palus^{1,2†} , Yahya Sohrabi^{3†}, Karl W. Broman⁴, Hynek Strnad⁵, Matyáš Šíma³, Daniel Růžek^{1,2}, Valeriya Volkova³, Martina Slapničková³, Jarmila Vojtíšková³, Lucie Mrázková^{3,6}, Jiří Salát² and Marie Lipoldová^{3,6*}

Abstract

Background: Tick-borne encephalitis (TBE) is the main tick-borne viral infection in Eurasia. Its manifestations range from inapparent infections and fevers with complete recovery to debilitating or fatal encephalitis. The basis of this heterogeneity is largely unknown, but part of this variation is likely due to host genetic. We have previously found that BALB/c mice exhibit intermediate susceptibility to the infection of TBE virus (TBEV), STS mice are highly resistant, whereas the recombinant congenic strain CcS-11, carrying 12.5% of the STS genome on the background of the BALB/c genome is even more susceptible than BALB/c. Importantly, mouse orthologs of human TBE controlling genes *Oas1b*, *Cd209*, *Tlr3*, *Ccr5*, *Irf1* and *Irf3*, are in CcS-11 localized on segments derived from the strain BALB/c, so they are identical in BALB/c and CcS-11. As they cannot be responsible for the phenotypic difference of the two strains, we searched for the responsible STS-derived gene-locus. Of course the STS-derived genes in CcS-11 may operate through regulating or epigenetically modifying these non-polymorphic genes of BALB/c origin.

Methods: To determine the location of the STS genes responsible for susceptibility of CcS-11, we analyzed survival of TBEV-infected F₂ hybrids between BALB/c and CcS-11. CcS-11 carries STS-derived segments on eight chromosomes. These were genotyped in the F₂ hybrid mice and their linkage with survival was tested by binary trait interval mapping. We have sequenced genomes of BALB/c and STS using next generation sequencing and performed bioinformatics analysis of the chromosomal segment exhibiting linkage with TBEV survival.

Results: Linkage analysis revealed a novel suggestive survival-controlling locus on chromosome 7 linked to marker D7Nds5 (44.2 Mb). Analysis of this locus for polymorphisms between BALB/c and STS that change RNA stability and genes' functions led to detection of 9 potential candidate genes: *Cd33*, *Klk1b22*, *Siglece*, *Klk1b16*, *Fut2*, *Grwd1*, *Abcc6*, *Otog*, and *Mkrn3*. One of them, *Cd33*, carried a nonsense mutation in the STS strain.

Conclusions: The robust genetic system of recombinant congenic strains of mice enabled detection of a novel suggestive locus on chromosome 7. This locus contains 9 candidate genes, which will be focus of future studies not only in mice but also in humans.

Keywords: Tick-borne encephalitis virus (TBEV), Mouse model, Survival, Susceptibility locus, Chromosome 7, Candidate gene

*Correspondence: lipoldova@img.cas.cz

†Martin Palus and Yahya Sohrabi contributed equally to this work

³ Department of Molecular and Cellular Immunology, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Vídeňská 1083, 14220 Prague, Czech Republic

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



Background

Tick-borne encephalitis (TBE) is the main tick-borne viral infection in Eurasia. It is prevalent across the entire continent from Japan to France [1]. The disease is caused by tick-borne encephalitis virus (TBEV), a flavivirus of the family *Flaviviridae*, which besides TBEV includes West Nile virus (WNV), dengue virus (DENV), Zika virus (ZIKV), yellow fever virus (YFV), Japanese encephalitis virus (JEV), and several other viruses causing extensive morbidity and mortality in humans. Ticks act as both the vector and reservoir for TBEV. The main hosts are small rodents, with humans being accidental hosts. In Europe and Russia between 5000 and 13,000 clinical cases of TBE are reported annually, with a large annual fluctuation [2]. The highest incidence of TBE is reported in western Siberia, in the Czech Republic, Estonia, Slovenia and Lithuania, but the prevalence of the disease is believed to be higher than actually reported [1, 2]. TBEV may produce a variety of clinical symptoms, from an asymptomatic disease to a fever and acute or chronic progressive encephalitis. The outcome of infection depends on the strain of virus [1], as well as on the genotype [3], sex and age of the host [4], and on the environmental and social factors [1]. Environmental and social factors influence also risk of infection.

Genetic influence on susceptibility to TBEV-induced disease has been analyzed by two main strategies: a hypothesis-independent phenotype-driven approach and a hypothesis-driven approach. Application of a genome-wide search (hypothesis-independent approach) in mouse led to identification of the gene *Oas1* (2'-5'-oligoadenylate synthetase gene) [5, 6]. A stop codon in exon 4 of the gene *Oas1b* (a natural knockout) present in majority of mouse laboratory strains causes production of protein lacking 30% of the C terminal sequence [5]. This part of molecule seems to be critical for tetramerization required for OAS1B activity leading to degradation of viral RNA. Thus, this mutation makes majority of mouse laboratory strains susceptible to flaviviruses [6, 7]. Human ortholog to this gene (*OAS1*) also modifies susceptibility to other flaviviruses (WNV) [8, 9], whereas *OAS2* and *OAS3* localized in the same cluster on chromosome 12q24.2 influence response to TBEV [3]. The polymorphic sites associated in *OAS2* and *OAS3* with susceptibility to TBEV did not resulted in amino acid changes, thus mechanisms of susceptibility control is not known [3]. The hypothesis-driven approach has focused on genes that encode molecules indicated to be involved in antiviral response by mechanistic studies [9]. These candidate genes studies revealed that polymorphisms in *CD209/DC-SIGN* [10], *CCR5* [11, 12], *TLR3* [12, 13], *IL10* [14] and *IFNL3/IL28B* [14] influence susceptibility to TBEV in humans.

Our previous study has shown that both after subcutaneous and intracerebral inoculation of European prototypic TBEV, BALB/c mice exhibited intermediate susceptibility to the infection, STS mice were highly resistant, whereas the strain CcS-11, which carries 12.5% of the STS genome on the background of the genome of the strain BALB/c [15], is even more susceptible than its two parents—BALB/c and STS [16]. Importantly, mouse orthologs of human TBEV controlling genes: *Oas1b*, *Cd209*, *Tlr3*, *Ccr5*, *Il10* and *Ifnl3* are in CcS-11 localized on segments derived from the strain BALB/c (Fig. 1), so they are identical in both BALB/c and CcS-11 and hence cannot be responsible for the phenotypic difference of the two strains. Therefore, the difference must be due to a presently unknown locus, which could be detected by a linkage study of a cross between BALB/c and CcS-11. Thus, we have generated a F₂ intercross between BALB/c and CcS-11 and performed a linkage and bioinformatics analysis. These studies revealed a novel suggestive locus on mouse chromosome 7 containing 9 potential candidate genes.

Methods

Mice

417 female F₂ offspring of an intercross between strains CcS-11 and BALB/c (mean and median age 9.5 and 9 weeks, respectively, at the time of infection) were produced at the Institute of Molecular Genetics AS CR. Mice were tested in three successive experimental groups at the Institute of Parasitology, AS CR. When used for these experiments, strain CcS-11 had undergone more than 90 generations of inbreeding. Experiments Nr. 1, 2, and 3 comprised 120, 121 and 176 F₂ mice, respectively. Sterilized pellet diet and water were supplied ad libitum. The mice were housed in plastic cages with wood-chip bedding, situated in a specific pathogen-free room with a constant temperature of 22 °C and a relative humidity of 65%.

Virus infection and disease phenotype

Experiments were performed with European prototypic TBEV strain Neudoerfl (a generous gift from Professor F. X. Heinz, Medical University of Vienna). This strain was passaged five times in brains of suckling mice before the use in this study [16]. Mice were infected subcutaneously with 10⁴ pfu of the virus.

Mice were scored for mortality for a period of 35 days post-infection (p.i.) with TBEV, as well as presence of ruffled fur and paresis in three independent successive experiments at the Institute of Parasitology AS CR.

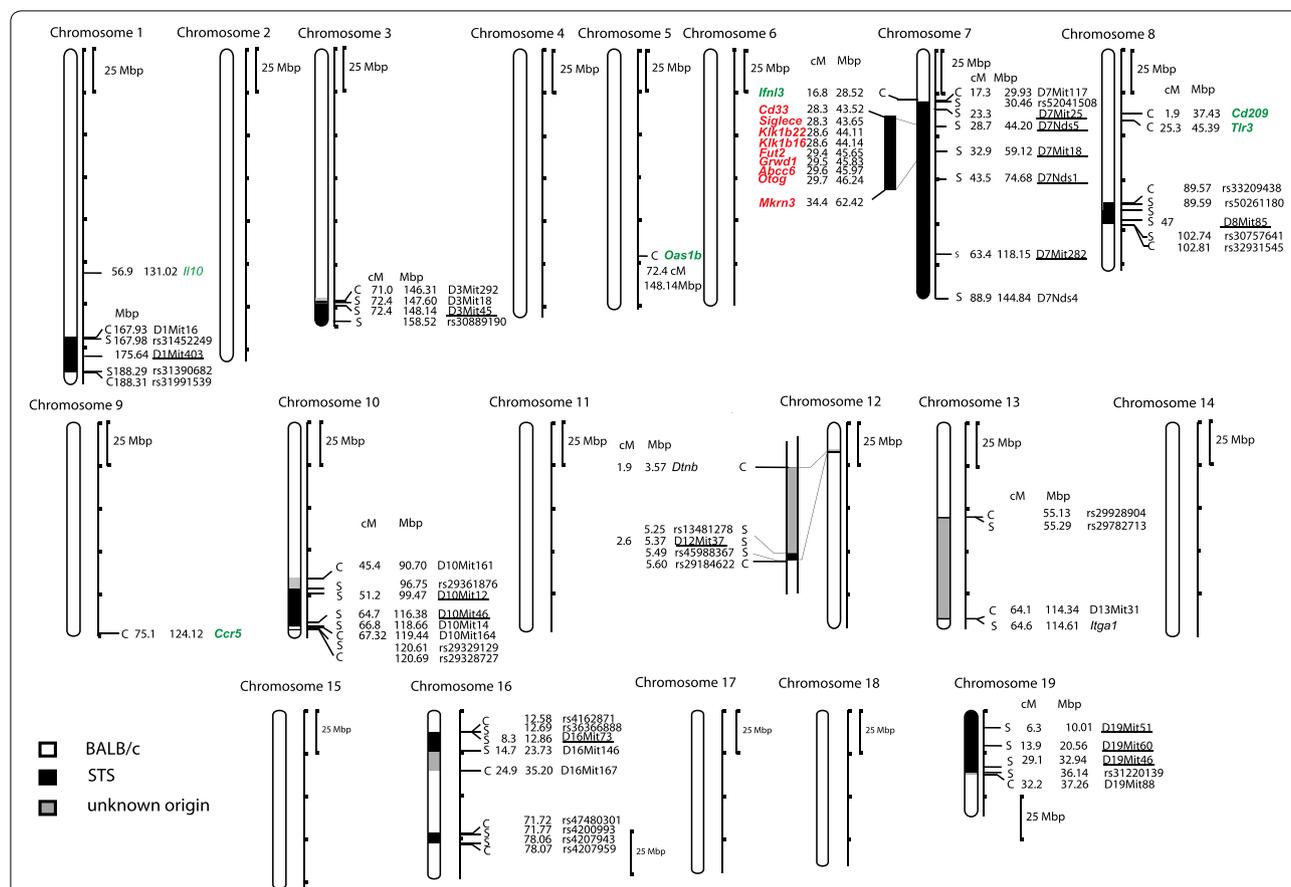


Fig. 1 Genetic composition of the strain CcS-11. The regions of STS and BALB/c origin are represented as dark and white, respectively, the boundary regions of undetermined origin are shaded. Only the markers or SNPs defining the boundaries of STS-derived segment and markers that were tested for linkage (underlined) are shown. Genes *Oas1b*, *Cd209*, *Tlr3*, *Ccr5*, *Ifnl3* and *Il10*, known to control susceptibility to TBEV are shown in green, potential candidate genes *Cd33*, *Klk1b22*, *Siglece*, *Klk1b16*, *Fut2*, *Grwd1*, *Abcc6*, *Otag*, and *Mkrn3* detected in current study are shown in red

Genotyping of F₂ mice

DNA was isolated from tails using a standard proteinase procedure. The strain CcS-11 differs from BALB/c at STS-derived regions on eight chromosomes [17]. These differential regions were genotyped in the F₂ hybrid mice between CcS-11 and BALB/c using 16 microsatellite markers (Generi Biotech, Hradec Králové, Czech Republic): D1Mit403, D3Mit45, D7Mit25, D7Nds5, D7Mit18, D7Nds1, D7Mit282, D7Mit259, D8Mit85, D10 Mit12, D10Mit46, D12Mit37, D16Mit73, D19Mit51, D19Mit60, D19Mit46 (Fig. 1) as described in [17].

Statistical analysis

Survival, ruffled fur and paresis were treated as binary phenotypes (death/survival; presence/absence of symptom), and binary trait interval mapping was performed [18, 19]. A permutation test [20] was used to assess significance. This takes account of the limited genetic difference between the strains BALB/c and CcS-11. On the basis of 10,000 permutation replicates, the 5%

significance LOD threshold was 2.56; the 10% threshold was 2.23. The Pearson correlation coefficient between presence of death and paresis was computed by the program Statistica for Windows 12.0 (StatSoft, Inc., Tulsa, OK).

Detection of polymorphisms that change RNA stability and genes' functions

We have sequenced the genomes of strains BALB/c and STS using next generation sequencing (NGS) system HiSeq 2500 (Illumina) (12× coverage). NGS data was preprocessed using software Trimmomatic [21] and overlapped paired reads were joined by software Flash [22]. Alignment—reference mouse sequence mm10 (build GRCm38)—was performed using BWA (Burrows-Wheeler Aligner) [23] program. Mapped reads were sorted and indexed, duplicated reads were marked. Segment covering peak of linkage on chromosome 7 from 36.2 to 74.5 Mb was inspected for polymorphisms between BALB/c and STS that change RNA stability and

genes' functions. Local realignment around indels, base recalibration and variants filtration were performed using software GATK (The Genome Analysis Toolkit) [24]. Variant annotation and effect prediction was performed by software SnpEff [25]. IGV (Integrated Genome Viewer) was used for visualization of results [26].

Results

Binary trait linkage analysis revealed a suggestive locus on chromosome 7 near D7Nds5 affecting the binary trait (death/survival) (LOD=2.15), with a corresponding genome-scan-adjusted *P* value=0.12 (Fig. 2a). The 1-LOD support interval spans from D7Mit25 to D7Nds1. The STS allele both in homozygotes and heterozygotes was associated with a higher death rate in each of the three separate experimental groups (Fig. 2b), and in the pooled data (Fig. 3a), so its presence in CcS-11 enhances even more the overall susceptibility determined by the BALB/c background. Ruffled fur was observed in only 8%

of mice, so it was not suitable for statistical analysis. Paresis was less frequent than mortality (n=60 vs. 102) and not all paretic mice died, but the two phenotypes were positively correlated (Pearson correlation 0.53). Moreover, frequency of paresis in the three D7Nds5 genotypes (Fig. 3b), although not significantly different, was biologically compatible with the survival data, because D7Nds5 CC homozygotes had the highest survival rate, and the lowest percentage of paresis.

We have sequenced genomes of BALB/c and STS and analyzed the segment covering peak of linkage on chromosome 7 from 36.2 to 74.5 Mb for polymorphisms between BALB/c and STS that change RNA stability and genes' functions. This revealed 9 potential candidate genes: *Cd33* (CD33 antigen), *Klk1b22* (kallikrein 1-related peptidase b22), *Siglece* (sialic acid binding Ig-like lectin E), *Klk1b16* (kallikrein 1-related peptidase b16), *Fut2* (fucosyltransferase 2), *Grwd1* (glutamate-rich WD repeat containing 1), *Abcc6* (ATP-binding cassette,

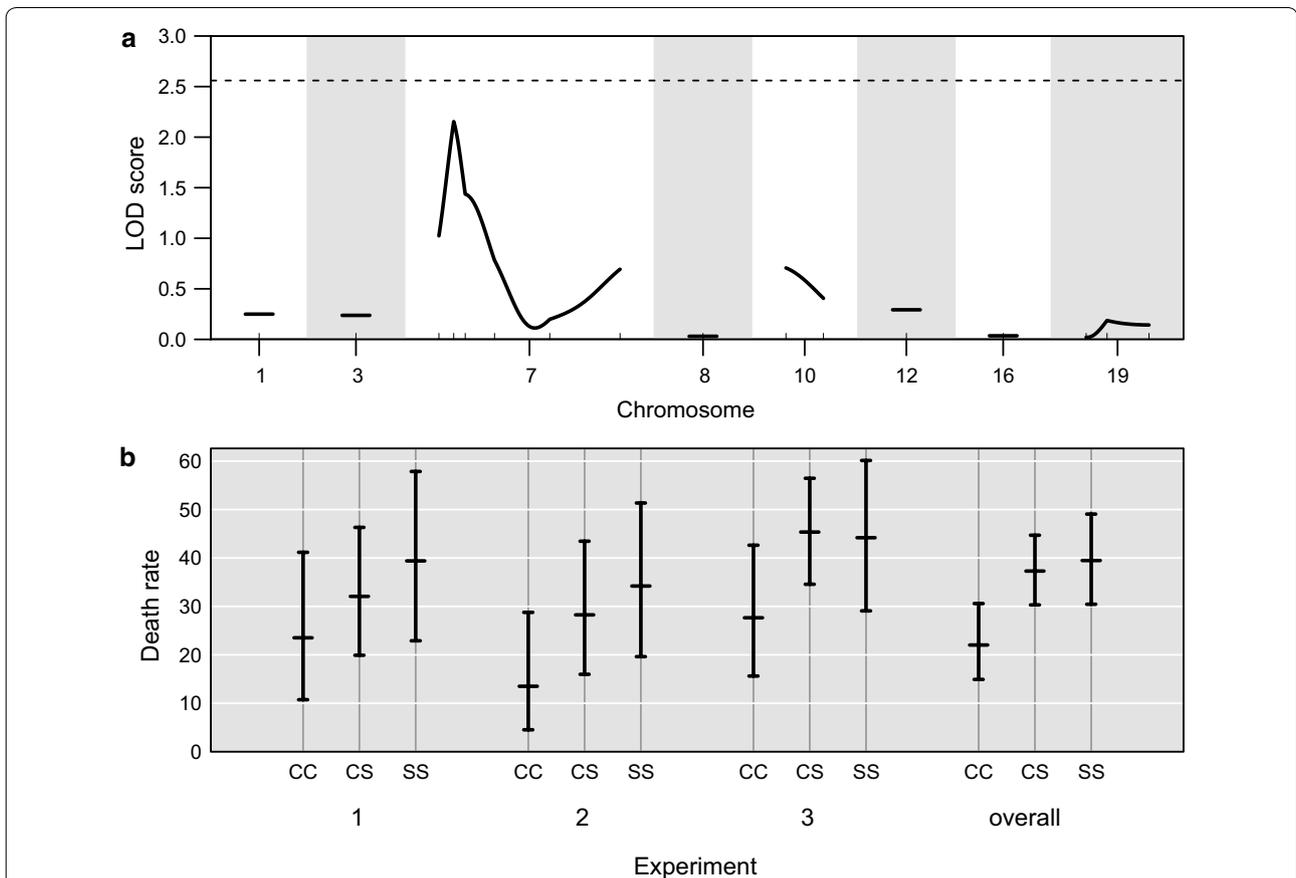
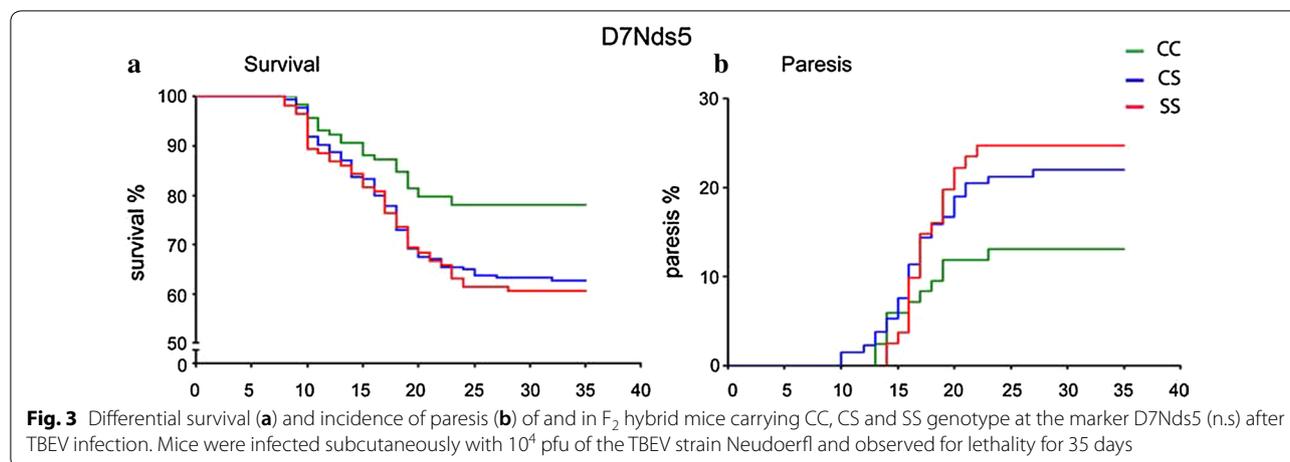


Fig. 2 Genetic influence on susceptibility to TBEV in an F₂ intercross between BALB/c and CcS-11. **a** LOD curves from binary trait interval mapping for death/survival. A dashed horizontal line is plotted at the 5% significance threshold, adjusting for the genome scan. **b** A plot of the death rate as a function of genotype at marker D7Nds5 and experiment, with 95% confidence intervals. C and S indicate the presence of BALB/c and STS allele, respectively. The S allele is associated with a higher death rate. The numbers of mice in the experiments 1, 2, and 3 were 120, 121 and 176, respectively



sub-family C (CFTR/MRP), member 6), *Otog* (otogelin), and *Mkrn3* (makorin, ring finger protein, 3) (Table 1, Fig. 1).

One of these genes, *Cd33*, carried in the strain STS a nonsense mutation. Other changes in the strain STS in comparison with BALB/c (and the reference strain C57BL/6) represented single amino acids change in Siglec E, KLK1B22, KLK1B16, FUT2 and OTOG. The BALB/c strain had in comparison with STS (and the reference strain C57BL/6) deletion of two amino acids in GRWD1, insertion of four amino acids in MKRN3 and single amino acid change in and ABCC6 (Table 1).

Discussion

CD33 and Siglec E belong to family of CD33-related sialic-acid-binding immunoglobulin-like lectins (CD33r-Siglecs). They are ITIM-containing inhibitory receptors, which are involved in regulation of inflammatory and immune responses [27]. Gene *Cd33* carried in the strain STS a nonsense mutation (Table 1). Product of this gene is in mouse expressed on myeloid precursors and cells of myeloid origin [28] and on microglial cells [29]. It can inhibit response to amyloid plaques and its deletion leads to protection in the mouse model of Alzheimer disease (AD) [29] and in humans some *CD33* genetic variants are associated with late-onset AD [30]; its potential role in pathology of TBEV might be associated with its regulatory role in inflammatory responses. Gene *Siglece* carried in the strain STS a single amino acid change. Siglec E is expressed on microglia and inhibits neurotoxicity triggered by neural debris [31], which might have protective role against damage induced by flaviviruses.

A single amino acid change was present in KLK1B22 and KLK1B16. Kallikreins are serine proteases that might both help to fight infection by activating complement system [32], as well as aggravate disease symptoms by

releasing bradykinin, which causes alterations in vascular permeability [33]. Their role in defense against flaviviruses has not been described. Kallikrein-bradykinin system have been described to contribute to protection against *Leishmania* [34] and *Trypanosoma cruzi* [35] parasites in mice. Interestingly, on the mouse chromosome 7 were in the strain CcS-11 mapped loci *Lmr21* and *Tbbr3* that control susceptibility to *L. major* [36] and *T. b. brucei* [17], respectively. However, both loci are mapped on a long chromosomal segment, thus other gene(s) might be responsible for their effect.

FUT2 have been described to influence control of a wide range of pathogens such as noroviruses [37], rotaviruses [38], HIV [39], and *Escherichia coli* [40] in humans, and to *Helicobacter pylori* in mouse [41], but its role in resistance to flaviviruses is not known.

Makorin 1 induces degradation of WNV capsid which might protect host cells [42]. The E3 ligase domain responsible for MKRN1 effect is present also in MKRN3 [43]. Thus, gene *Mkrn3* might have relationship to defense against flaviviruses. Similarly, possible role of *Otog*, *Grwd1* and *Abcc6* in resistance to TBEV remains to be elucidated.

Public database BioGPS shows that all the candidate genes are in uninfected mice expressed in tissues such as brain, spleen and liver (Table 2). Brain is the main target for the virus; however, during the extraneural phase of the infection, several tissues and organs in the body are infected, including spleen and liver [44]. Highest expression in these tissues exhibits *Cd33* and *Siglece* with expression in microglia ten times higher than median value ($>10M$), *Cd33* and *Klk1b22* are highly expressed in spleen ($>3M$), $>10M$ expression of these two genes is also observed in bone marrow; *Siglece* is also highly expressed in bone ($>3M$) and bone marrow macrophages ($>3M$), whereas *Cd33* is highly expressed

Table 1 List of candidate genes in TBEV susceptibility locus

Position Bp	Reference genotype C57BL/6	Genotype BALB/c	Genotype STS	Protein position of amino acid	Reference amino acid	Alteration	Type of change	Gene symbol	Transcription status	Gene name	Gene ID: MGI	Gene ID: NCBI
43,528,893	C/C	C/C	T/T	353	G	K	Single AA Change	<i>Cd33</i>	KNOWN	CD33 antigen	99,440	12,489
43,532,167	G/G	G/G	A/A	190	R	*	Nonsense Mutation	<i>Cd33</i>	KNOWN	CD33 antigen	99,440	12,489
43,659,827	G/G	G/G	T/T	102	D	E	Single AA Change	<i>Siglece</i>	KNOWN	Sialic acid binding Ig-like lectin E	1,932,475	83,382
44,115,970	A/A	A/A	C/A	115	L	Y	Single AA Change	<i>Klkb22</i>	KNOWN	Kallikrein 1-related peptidase b22	95,291	13,646
44,140,534	G/G	G/G	C/C	76	G	A	Single AA Change	<i>Klkb16</i>	KNOWN	Kallikrein 1-related peptidase b16	891,982	16,615
45,650,779	G/G	G/G	A/A	190	R	W	Single AA Change	<i>Fut2</i>	KNOWN	Fucosyltransferase 2	109,374	14,344
45,830,054	CTCTTC A/CTC TTCA	C/C	CTCTTC/CTCTCA	129	ED	.	Deletion	<i>Grwd1</i>	KNOWN	Glutamate-rich WD repeat containing 1	2,141,989	101,612
45,977,290	C/C	A/A	C/C	1448	V	L	Single AA Change	<i>Abcc6</i>	KNOWN	ATP-binding cassette, sub-family C (CFTR/MRP), member 6	1,351,634	27,421
46,262,804	C/C	C/C	T/T	748	R	W	Single AA Change	<i>Otog</i>	KNOWN	Otogelin	1,202,064	18,419
62,419,214	C/C	CGGCATTGG CACT/CGG CATTGG CACT	C/C	275	P	PVPMP	Insertion	<i>Mkrm3</i>	KNOWN	Makorin, ring finger protein, 3	2,181,178	22,652

Table shows differences between BALB/c and STS in DNA and protein sequences in potential candidate genes. Table shows also sequences of the reference mouse strain C57BL/6

Table 2 Expression of potential candidate genes in organs and cells of uninfected mice

Gene symbol	Cd33	Siglece	Klk1b22	Klk1b16	Fut2	Grwd1	Abcc6	Otog	Mkrn3
Gene ID: MGI	99,440	1,932,475	95,291	891,982	109,374	2,141,989	1,351,634	1,202,064	2,181,178
Gene ID: NCBI	12,489	83,382	13,646	16,615	14,344	101,612	27,421	18,419	22,652
Median	5.2	4.7	4.6	4.6	4.9	163	4.6	4.95	4.8
Organs									
Bone	77.05	> 10M	20.82	> 3M	4.64	~ M	4.64	~ M	~ M
Bone marrow	114.05	> 10M	51.63	> 10M	4.64	~ M	4.64	~ M	~ M
Brain									
Amygdala	4.97	~ M	4.67	~ M	4.64	~ M	4.64	~ M	~ M
Cerebellum	4.91	~ M	4.76	~ M	5.24	~ M	4.64	~ M	~ M
Cerebral cortex	4.91	~ M	4.67	~ M	4.64	~ M	4.64	~ M	~ M
Hippocampus	4.91	~ M	4.67	~ M	4.64	~ M	4.64	~ M	~ M
Olfactory bulb	4.89	~ M	4.67	~ M	4.64	~ M	4.64	~ M	> 3M
Eye									
Eye cup	13.71	~ 3M	4.67	~ M	5.80	~ M	4.64	~ M	> M
Lens	9.26	> M	6.18	~ M	5.80	~ M	7.54	> M	> M
Retina	5.72	~ M	4.67	~ M	4.86	~ M	4.64	~ M	> 10M
Intestine large	4.86	< M	4.67	~ M	91.36	> 3M	6.13	> M	~ M
Intestine small	4.91	~ M	4.68	~ M	6.45	~ M	4.66	~ M	~ M
Kidney	4.87	~ M	4.67	~ M	10.44	> M	15.57	~ 3M	~ M
Lacrimal gland	4.92	~ M	4.67	~ M	15.28	> M	187.21	> 30M	~ M
Liver	4.89	~ M	5.09	~ M	4.86	~ M	6.14	> M	~ M
Lymph nodes	85.8	> 3M	4.68	~ M	4.99	~ M	4.64	~ M	~ M
Pancreas	7.40	~ M	5.15	~ M	10.44	> M	22.47	> 3M	~ M
Pituitary	4.89	~ M	4.67	~ M	7.07	> M	114.09	~ 30M	~ M
Prostate	5.25	~ M	4.75	~ M	42.15	> 3M	4.64	~ M	~ M
Salivary gland	5.62	~ M	4.67	~ M	5.95	~ M	36,542.06	> 1000M	~ M
Spleen	15.39	~ 3M	9.02	> M	4.91	~ M	13.11	~ 3M	~ M
Stomach	4.89	~ M	4.65	~ M	117.61	> 10M	4.64	~ M	~ M
Testis	5.40	~ M	4.67	~ M	16.28	~ 3M	30.89	> 3M	~ M
Uterus	4.91	~ M	4.67	~ M	222.61	> 30M	4.64	~ M	~ M
Cells									
B cells_marginal_zone	25.26	> 3M	4.67	~ M	4.86	~ M	4.64	~ M	~ M
Common myeloid progenitor	11.3	> M	6.18	~ M	5.20	~ M	5.05	~ M	~ M
Dendritic lymphoid cells	9.83	> M	4.67	~ M	4.86	~ M	4.64	~ M	~ M

Table 2 (continued)

Gene symbol	Cd33	Siglece	Klk1b22	Klk1b16	Fut2	Grwd1	Abcc6	Otog	Mkrn3
Dendritic cells myeloid CD8a-	77.51 > 10M	4.96 ~ M	4.64 ~ M	4.64 ~ M	4.86 ~ M	299.19 > M	4.64 ~ M	4.95 ~ M	4.64 ~ M
Dendritic plasmacytoid B220+	140.59 ~ 30M	4.74 ~ M	88.52 > 10M	4.64 ~ M	4.86 ~ M	238.95 > M	4.64 ~ M	5.14 ~ M	4.64 ~ M
Granulocytes mac1 + gr1+	651.27 > 30M	6.87 > M	4.64 ~ M	4.64 ~ M	5.04 ~ M	53.79 < M	4.64 ~ M	4.95 ~ M	4.64 ~ M
Hematopoietic stem cells	42.14 > 3M	4.8 ~ M	4.96 ~ M	4.64 ~ M	6.41 > M	1445.80 ~ 10M	4.64 ~ M	4.95 ~ M	5.19 ~ M
Macrophage_bone_marrow	11.64 > M	14.68 > 3M	4.64 ~ M	4.64 ~ M	4.86 ~ M	308.83 > M	4.64 ~ M	4.95 ~ M	4.64 ~ M
Mast cells	214.54 > M	4.67 ~ M	4.64 ~ M	4.64 ~ M	4.83 ~ M	55.07 < M	4.64 ~ M	4.95 ~ M	4.64 ~ M
Mega_erythrocyte progenitor	5.31 ~ M	4.67 ~ M	7.57 > M	4.64 ~ M	6.47 > M	2842.61 > 10M	4.64 ~ M	8.07 > M	4.64 ~ M
Microglia	99.81 > 10M	79.64 > 10M	4.64 ~ M	4.64 ~ M	4.86 ~ M	269.08 > M	4.64 ~ M	4.95 ~ M	4.64 ~ M
Osteoclasts	363.46 > 30M	4.77 ~ M	7.67 > M	4.64 ~ M	4.99 ~ M	327.07 > M	4.64 ~ M	5.31 ~ M	4.64 ~ M
T-cellsCD4+	4.91 ~ M	5.00 ~ M	7.57 > M	4.82 ~ M	4.86 ~ M	621.74 > 3M	4.64 ~ M	5.06 ~ M	4.64 ~ M
T-cellsCD8+	4.91 ~ M	4.67 ~ M	13.55 ~ 3M	4.64 ~ M	5.29 ~ M	715.19 > 3M	4.64 ~ M	6.13 > M	4.64 ~ M
T-cells FoxP3+	25.67 > 3M	4.67 ~ M	4.64 ~ M	4.96 ~ M	4.93 ~ M	584.63 ~ 3M	4.64 ~ M	4.95 ~ M	4.64 ~ M

Data were compiled from public database BioGps (<http://biogps.org>) May 6, 2018. First column: relative units; second column: relationship to median (M); M = median value across all samples for a single probe set

in granulocytes (>30M), plasmacytoid dendritic cells (>30M), osteoclasts (>30M), myeloid dendritic cells (>10M), in spleen (>3M), lymph nodes (>3M), eyecup (>3M), B cells in marginal zone (>3M) and in FoxP⁺ T cells (>3M). For both *Klk1b22* and *Klk1b16* is characteristic very high expression in salivary gland (>1000M) and high expression in lacrimal gland (>30M). *Klk1b22* is also highly expressed in large intestine (>30M), kidney (>30M), pancreas (>30M), testis (>30M), stomach (>10M), plasmacytoid dendritic cells (>10M), small intestine (>3M), spleen (>3M) and CD8⁺ T cells (>3M), whereas *Klk1b16* is also highly expressed in pituitary (>30M), kidney (>3M), pancreas (>3M) and testis (>30M). Highest expression of *Fut2* was observed in uterus (>30M), and in stomach (>10M), it was also highly expressed in large intestine (>3M), prostate (>3M) and in testis (>3M). GRWD1 was described to play a role in ribosome biogenesis and during myeloid differentiation [45]. High expression level in hematopoietic stem cells (>10M), mega erythrocyte progenitors (>10M), granulocytes (>10M), common myeloid progenitor (>3M) supports this finding, but it is also expressed in several T cell subpopulations (>3M), B cells in marginal zone (>3M), as well as in lacrimal gland. *Abcc6* is highly expressed in liver (≫30M) and in lens (>10M) and *Mrkn3* is highly expressed in retina (10M) and in olfactory bulb (>3M). The expression data further support a potential role of detected candidate genes in defense against TBEV, but they must be in the future complemented with data describing gene expression after TBEV infection.

We have found a susceptibility allele of a locus on chromosome 7 in the resistant strain STS. This apparent paradox is likely caused by the fact that most inbred mouse strains were produced without an intentional selective breeding for a specific quantitative phenotype (like susceptibility to specific infections). Therefore they inherited randomly from their non-inbred ancestors susceptible alleles at some loci and resistant alleles at others, so that their overall susceptibility phenotype depends on the relative number of both types of alleles. Such finding is not unique, as susceptibility alleles originating from resistant strains were found in susceptibility studies of other infectious diseases [17, 46, 47] and colon cancer [48]. Similarly, in different in vitro tests of immune responses a low-responder allele was identified in a high responding strain [49] or vice versa [50]. Another explanation might be presence of BALB/c allele interacting with STS allele on chromosome 7. Demonstration of such interaction would require further experiments. We have already observed interaction of STS and BALB/c alleles leading to extreme phenotypes in susceptibility to *L. major* [51] and *L. tropica* [47].

Conclusion

Mapping of TBEV controlling genes in mice is not easy due to presence of a strong TBEV controlling gene *Oas1b*, which is identical both in BALB/c and CcS-11, as well as in majority of laboratory mouse strains [6, 7] and masks effects of other controlling genes. Therefore using a powerful genetic system—recombinant congenic strains, we succeeded in mapping novel TBEV susceptibility locus on chromosome 7 and identified 9 potential candidate genes. Products of some of these genes have been described to participate in defense against flaviviruses, the role of the others is unknown. The genes detected here will be focus of future studies that will include characterization of candidate gene(s) products in BALB/c and CcS-11, introducing modification to candidate genes and study their influence on disease outcome in mouse, and study influence of polymorphisms in human orthologs of candidate genes on susceptibility to TBEV in humans.

List of abbreviations

Abcc6 (ATP-binding cassette, sub-family C (CFTR/MRP), member 6), mouse gene; *ABCC6*-human gene; *ABCC6*-protein (gene product); *Ccr5*: chemokine (C–C motif) receptor 5; *Cd209*: CD209 antigen; *Cd33* (CD33 antigen); *Fut2* (fucosyltransferase 2); *Grwd1* (glutamate-rich WD repeat containing 1); *Ifnl3*: interferon lambda 3 (synonym *Il28b*); *Il10*: interleukin 10; *Klk1b16* (kallikrein 1-related peptidase b16); *Klk1b22* (kallikrein 1-related peptidase b22); *Mkrn3* (makorin, ring finger protein, 3); *Oas1b*: 2′-5′-oligoadenylate synthetase gene; *Otog* (otogelin); pfu: Plaque-forming unit; RC: Recombinant congenic; s.c: Subcutaneous; *Siglece* (sialic acid binding Ig-like lectin E); TBE: Tick-borne encephalitis; TBEV: Tick-borne encephalitis virus; *Tlr3*: toll-like receptor 3.

According to current gene and protein nomenclature, mouse gene symbols are italicized, with only the first letter in upper-case (e.g. *Cd33*). Protein symbols are not italicized, and all letters are in upper-case (e.g. CD33). Human gene symbols are in upper-case and are italicized (e.g. *CD33*). Protein symbols are identical to their corresponding gene symbols except that they are not italicized (e.g. CD33).

Authors' contributions

ML conceived and designed the research, analyzed the experiments, coordinated collaborations; and wrote the manuscript. MP, YS, MŠ, DR, VV, JV, LM, JS and MS performed the experiments and participated in the design of experiments and analysis of data. HS performed sequence analysis. KWB performed statistical analysis. HS and JV analyzed bioinformatics data. All authors read and approved the final manuscript.

Author details

¹ Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, Branišovská 31, 37005 České Budějovice, Czech Republic. ² Department of Virology, Veterinary Research Institute, Hudcova 70, 62100 Brno, Czech Republic. ³ Department of Molecular and Cellular Immunology, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Vídeňská

1083, 14220 Prague, Czech Republic. ⁴ Department of Biostatistics and Medical Informatics, 6770 Medical Sciences Center, 1300 University Avenue, Madison, WI 53706-1532, USA. ⁵ Department of Genomics and Bioinformatics, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Videňská 1083, 14220 Prague, Czech Republic. ⁶ Department of Natural Sciences, Faculty of Biomedical Engineering, Czech Technical University in Prague, Sítňá 3105, 272 01 Kladno, Czech Republic.

Acknowledgements

We thank Dr. Peter Demant from the Department of Molecular and Cellular Biology of the Roswell Park Cancer Institute in Buffalo, New York, USA for useful discussion. We appreciate help of Mrs. Helena Havelková with preparation of figures.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The research had complied with all relevant European Union guidelines for work with animals and was in accordance with the Czech national law and guidelines on the use of experimental animals and protection of animals against cruelty (the Animal Welfare Act Number 246/1992 Coll.). The protocol was approved by the Committee on the Ethics of Animal Experiments of the Institute of Parasitology and of the Departmental Expert Committee for the Approval of Projects of Experiments on Animals of the Academy of Sciences of the Czech Republic (Permit Number: 165/2010).

Funding

This work was supported by the by the Czech Science Foundation [Grants 16-20054S, 14-30186S, 16-22346S], the Academy of Sciences of the Czech Republic [RVO 68378050], the Ministry of Education, Youth, and Sports of the Czech Republic, under the NPU I program [Grant LO1218], by Project "FIT" [Pharmacology, Immunotherapy, nanoToxicology], which was funded by the European Regional Development Fund, the Czech Technical University [SGS15/173/OHK4/2T/17] and by the United States National Institutes of Health [R01 GM074244, to KWB]. Y. Sohrabi and M. Šíma are PhD students of Faculty of Science, Charles University, Prague, Czech Republic. M. Slapničková is a PhD student of Faculty of Science, The University of South Bohemia, České Budějovice, Czech Republic, Lucie Mrázková is a PhD student of the Department of Natural Sciences, Faculty of Biomedical Engineering, Czech Technical University in Prague.

Publisher's Note

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

Received: 12 January 2018 Accepted: 20 June 2018

Published online: 06 July 2018

References

- Dobler G, Gniel D, Petermann R, Pfeffer M. Epidemiology and distribution of tick-borne encephalitis. *Wien Med Wochenschr*. 2012;162:230–8.
- Süss J. Tick-borne encephalitis 2010: epidemiology, risk areas, and virus strains in Europe and Asia—an overview. *Ticks Tick Borne Dis*. 2011;2:2–15.
- Barkhash AV, Perelygin AA, Babenko VN, Myasnikova NG, Pilipenko PI, Romaschenko AG, Voevoda MI, Brinton MA. Variability in the 2'-5'-oligoadenylate synthetase gene cluster is associated with human predisposition to tick-borne encephalitis virus-induced disease. *J Infect Dis*. 2010;202:1813–8.
- Rushton JO, Lecollinet S, Hubálek Z, Svobodová P, Lussy H, Nowotny N. Tick-borne encephalitis virus in horses, Austria. *Emerg Infect Dis*. 2011;19:635–7.
- Perelygin AA, Scherbik SV, Zhulin IB, Stockman BM, Li Y, Brinton MA. Positional cloning of the murine flavivirus resistance gene. *Proc Natl Acad Sci USA*. 2002;99(14):9322–7.
- Mashimo T, Lucas M, Simon-Chazottes D, Frenkiel MP, Montagutelli X, Ceccaldi PE, Deubel V, Guenet JL, Despres P. A nonsense mutation in the gene encoding 2'-5'-oligoadenylate synthetase/L1 isoform is associated with West Nile virus susceptibility in laboratory mice. *Proc Natl Acad Sci USA*. 2002;99:11311–6.
- Darnell MB, Koprowski H, Lagerspetz K. Genetically determined resistance to infection with group B arboviruses. I. Distribution of the resistance gene among various mouse populations and characteristics of gene expression in vivo. *J Infect Dis*. 1974;129(3):240–7.
- Lim JK, Lisco A, McDermott DH, Huynh L, Ward JM, Johnson B, Johnson H, Pape J, Foster GA, Krysztof D, Follmann D, Stramer SL, Margolis LB, Murphy PM. Genetic variation in *OAS1* is a risk factor for initial infection with West Nile virus in man. *PLoS Pathog*. 2009;5(2):e1000321.
- Turtle L, Griffiths MJ, Solomon T. Encephalitis caused by flaviviruses. *QJM*. 2012;105:219–23.
- Barkhash AV, Perelygin AA, Babenko VN, Brinton MA, Voevoda MI. Single nucleotide polymorphism in the promoter region of the *CD209* gene is associated with human predisposition to severe forms of tick-borne encephalitis. *Antivir Res*. 2012;93:64–8.
- Kindberg E, Mickiene A, Ax C, Akerlind B, Vene S, Lindquist L, Lundkvist A, Svensson L. A deletion in the chemokine receptor 5 (*CCR5*) gene is associated with tickborne encephalitis. *J Infect Dis*. 2008;197:266–9.
- Barkhash AV, Voevoda MI, Romaschenko AG. Association of single nucleotide polymorphism rs3775291 in the coding region of the *TLR3* gene with predisposition to tick-borne encephalitis in a Russian population. *Antivir Res*. 2013;99(2):136–8.
- Mickiene A, Pakalnienė J, Nordgren J, Carlsson B, Hagbom M, Svensson L, Lindquist L. Polymorphisms in chemokine receptor 5 and toll-like receptor 3 genes are risk factors for clinical tick-borne encephalitis in the Lithuanian population. *PLoS ONE*. 2014;9(9):e106798.
- Barkhash AV, Babenko VN, Voevoda MI, Romaschenko AG. Association of *IL28B* and *IL10* gene polymorphism with predisposition to tick-borne encephalitis in a Russian population. *Ticks Tick Borne Dis*. 2016;7(5):808–12.
- Démant P, Hart AA. Recombinant congenic strains—a new tool for analyzing genetic traits determined by more than one gene. *Immunogenetics*. 1986;24(6):416–22.
- Palus M, Vojtíšková J, Salát J, Kopecký J, Grubhoffer L, Lipoldová M, Demant P, Růžek D. Mice with different susceptibility to tick-borne encephalitis virus infection show selective neutralizing antibody response and inflammatory reaction in the central nervous system. *J Neuroinflammation*. 2013;10:77.
- Šíma M, Havelková H, Quan L, Svobodová M, Jarošíková T, Vojtíšková J, Stassen AP, Demant P, Lipoldová M. Genetic control of resistance to *Trypanosoma brucei brucei* infection in mice. *PLoS Negl Trop Dis*. 2011;5:e1173.
- Xu S, Atchley WR. Mapping quantitative trait loci for complex binary diseases using line crosses. *Genetics*. 1996;143:1417–24.
- Broman KW. Mapping quantitative trait loci in the case of a spike in the phenotype distribution. *Genetics*. 2003;163(3):1169–75.
- Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics*. 1994;138:963–71.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
- Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*. 2011;27(21):2957–63.
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589–95.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303.
- Cingolani P, Platts A, Le Wang L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80–92.

26. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24–6.
27. Macauley MS, Crocker PR, Paulson JC. Siglec-mediated regulation of immune cell function in disease. *Nat Rev Immunol*. 2014;14(10):653–66.
28. Brinkman-Van der Linden EC, Angata T, Reynolds SA, Powell LD, Hedrick SM, Varki A. CD33/Siglec-3 binding specificity, expression pattern, and consequences of gene deletion in mice. *Mol Cell Biol*. 2003;23(12):4199–206.
29. Gričič A, Serrano-Pozo A, Parrado AR, Lesinski AN, Asselin CN, Mullin K, Hooli B, Choi SH, Hyman BT, Tanzi RE. Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. *Neuron*. 2013;78(4):631–43.
30. Jiang T, Yu JT, Hu N, Tan MS, Zhu XC, Tan L. CD33 in Alzheimer's disease. *Mol Neurobiol*. 2014;49(1):529–35.
31. Claude J, Linnartz-Gerlach B, Kudin AP, Kunz WS, Neumann H. Microglial CD33-related Siglec-E inhibits neurotoxicity by preventing the phagocytosis-associated oxidative burst. *J Neurosci*. 2013;33(46):18270–6.
32. Kant Upadhyay R. Biomarkers in Japanese encephalitis: a review. *Biomed Res Int*. 2013;2013:591290.
33. Rust NM, Papa MP, Scovino AM, da Silva MM, Calzavara-Silva CE, Marques ET Jr, Peçanha LM, Scharfstein J, Arruda LB. Bradykinin enhances Sindbis virus infection in human brain microvascular endothelial cells. *Virology*. 2012;422(1):81–91.
34. Nico D, Feijó DF, Maran N, Morrot A, Scharfstein J, Palatnik M, Palatnik-de-Sousa CB. Resistance to visceral leishmaniasis is severely compromised in mice deficient of bradykinin B2-receptors. *Parasit Vectors*. 2012;5:261.
35. Monteiro AC, Schmitz V, Morrot A, de Arruda LB, Nagajyothi F, Granato A, Pedreiro JB, Müller-Esterl W, Tanowitz HB, Scharfstein J. Bradykinin B2 Receptors of dendritic cells, acting as sensors of kinins proteolytically released by *Trypanosoma cruzi*, are critical for the development of protective type-1 responses. *PLoS Pathog*. 2007;3(11):e185.
36. Kurey I, Kobets T, Havelková H, Slapnicková M, Quan L, Trtková K, Grekov I, Svobodová M, Stassen AP, Hutson A, Demant P, Lipoldová M. Distinct genetic control of parasite elimination, dissemination, and disease after *Leishmania major* infection. *Immunogenetics*. 2009;61(9):619–33.
37. Lindesmith L, Moe C, Marionneau S, Ruvoen N, Jiang X, Lindblad L, Stewart P, LePendu J, Baric R. Human susceptibility and resistance to Norwalk virus infection. *Nat Med*. 2003;9(5):548–53.
38. Payne DC, Currier RL, Staat MA, Sahni LC, Selvarangan R, Halasa NB, Englund JA, Weinberg GA, Boom JA, Szilagyi PG, Klein EJ, Chappell J, Harrison CJ, Davidson BS, Mijatovic-Rustempasic S, Moffatt MD, McNeal M, Wikswo M, Bowen MD, Morrow AL, Parashar UD. Epidemiologic association between FUT2 secretor status and severe rotavirus gastroenteritis in children in the United States. *JAMA Pediatr*. 2015;169(11):1040–5.
39. Kindberg E, Hejdeman B, Bratt G, Wahren B, Lindblom B, Hinkula J, Svensson L. A nonsense mutation (428G → A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection. *Aids*. 2006;20(5):685–9.
40. Mottram L, Wiklund G, Larson G, Qadri F, Svennerholm AM. FUT2 non-secretor status is associated with altered susceptibility to symptomatic enterotoxigenic *Escherichia coli* infection in Bangladeshis. *Sci Rep*. 2017;7(1):10649.
41. Magalhães A, Rossez Y, Robbe-Masselot C, Maes E, Gomes J, Shevtsova A, Bugaytsova J, Borén T, Reis CA. Muc5ac gastric mucin glycosylation is shaped by FUT2 activity and functionally impacts *Helicobacter pylori* binding. *Sci Rep*. 2016;6:25575.
42. Ko A, Lee EW, Yeh JY, Yang MR, Oh W, Moon JS, Song J. MKRN1 induces degradation of West Nile virus capsid protein by functioning as an E3 ligase. *J Virol*. 2010;84(1):426–36.
43. <http://www.informatics.jax.org/genes.shtml>. Accessed 5 Sept 2017.
44. Hayasaka D, Aoki K, Morita K. Development of simple and rapid assay to detect viral RNA of tick-borne encephalitis virus by reverse transcription-loop-mediated isothermal amplification. *Virus J*. 2013;10:68.
45. Gratenstein K, Heggstad AD, Fortun J, Notterpek L, Pestov DG, Fletcher BS. The WD-repeat protein GRWD1: potential roles in myeloid differentiation and ribosome biogenesis. *Genomics*. 2005;85(6):762–73.
46. Vladimirov V, Badalová J, Svobodová M, Havelková H, Hart AA, Blazková H, Demant P, Lipoldová M. Different genetic control of cutaneous and visceral disease after *Leishmania major* infection in mice. *Infect Immun*. 2003;71:2041–6.
47. Sohrabi Y, Havelková H, Kobets T, Šíma M, Volkova V, Grekov I, Jarošíková T, Kurey I, Vojtíšková J, Svobodová M, Demant P, Lipoldová M. Mapping the genes for susceptibility and response to *Leishmania tropica* in mouse. *PLoS Negl Trop Dis*. 2013;7(7):e2282.
48. van Wezel T, Stassen AP, Moen CJ, Hart AA, van der Valk MA, Demant P. Gene interaction and single gene effects in colon tumour susceptibility in mice. *Nat Genet*. 1996;14:468–70.
49. Lipoldová M, Havelková H, Badalova J, Vojtíšková J, Quan L, Krulova M, Sohrabi Y, Stassen AP, Demant P. Loci controlling lymphocyte production of interferon γ after alloantigen stimulation in vitro and their co-localization with genes controlling lymphocyte infiltration of tumors and tumor susceptibility. *Cancer Immunol Immunother*. 2010;59:203–13.
50. Kosařová M, Havelková H, Krulová M, Demant P, Lipoldová M. The production of two Th2 cytokines, interleukin-4 and interleukin-10, is controlled independently by locus *Cypr1* and by loci *Cypr2* and *Cypr3*, respectively. *Immunogenetics*. 1999;49:134–41.
51. Havelková H, Badalová J, Svobodová M, Vojtíšková J, Kurey I, Vladimirov V, Demant P, Lipoldová M. Genetics of susceptibility to leishmaniasis in mice: four novel loci and functional heterogeneity of gene effects. *Genes Immun*. 2006;7(3):220–33.