



**HAL**  
open science

# Caractérisation génétique et fonctionnelle de la sensibilité et de la résistance innées au virus Zika chez la souris

Marie Bourdon

► **To cite this version:**

Marie Bourdon. Caractérisation génétique et fonctionnelle de la sensibilité et de la résistance innées au virus Zika chez la souris. Sciences agricoles. Université Paris-Saclay, 2023. Français. NNT : 2023UPASB066 . tel-04609286

**HAL Id: tel-04609286**

**<https://theses.hal.science/tel-04609286>**

Submitted on 12 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Genetic and functional characterization of innate susceptibility and resistance to Zika virus in mouse models

*Caractérisation génétique et fonctionnelle de la sensibilité et de la  
résistance innées au virus Zika chez la souris*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 581, agriculture, alimentation, biologie, environnement, santé (ABIES)  
Spécialité de doctorat : Génétique animale  
Graduate School : Biosphera. Référent : AgroParisTech

Thèse préparée dans le laboratoire de **Génétique de la souris (Institut Pasteur)**, sous la  
direction de **Xavier MONTAGUTELLI**, DVM, HDR

Thèse soutenue à Paris, le 30 novembre 2023, par

**Marie BOURDON**

### Composition du jury

Membres du jury avec voix délibérative

<b>Sophie LE PODER</b> Professeure, ENVA	Présidente
<b>Sébastien NISOLE</b> Chargé de recherche (HDR), INSERM (Université de Montpellier)	Rapporteur & Examineur
<b>Binnaz YALCIN</b> Chargée de recherche (HDR), INSERM (Université de Bourgogne)	Rapporteur & Examinatrice
<b>Sabine RIFFAULT</b> Directrice de recherche, INRAE (Université Paris-Saclay)	Examinatrice

**Titre:** Caractérisation génétique et fonctionnelle de la sensibilité et de la résistance innées au virus Zika chez la souris

**Mots clés:** génétique, virus Zika, virologie, réponse immunitaire innée, contrôle génétique de la sensibilité aux infections, réplication virale

**Résumé:** L'infection par le virus Zika (ZIKV) chez l'Homme peut provoquer de multiples conséquences, de l'absence de symptômes cliniques à des complications neurologiques ou des microcéphalies après des infections intra-utérines. En utilisant le Collaborative Cross (CC), une collection de lignées de souris consanguines avec une large diversité génétique, mon laboratoire a précédemment démontré pour la première fois l'importance de la génétique de l'hôte dans la variabilité clinique de la maladie provoquée par ZIKV, et a identifié CC071 comme la lignée la plus sensible, car elle présentait une forte charge virale dans le sang, des symptômes sévères et de la mortalité.

Durant ma thèse, j'ai d'abord étudié l'origine génétique de la sensibilité de la lignée CC071 dans un modèle *in vitro*. Des fibroblastes embryonnaires de souris (MEFs) de CC071 montraient une forte réplication virale associée à une induction des interférons de type I (IFN-I) retardée. Une analyse de quantitative trait locus (QTL) sur des MEFs dérivés d'un backcross entre CC071 et la lignée résistante CC001 a permis l'identification d'une mutation perte de fonction dans le gène *Irf3*, spécifique de la lignée CC071. Des analyses fonctionnelles ont montré que cette mutation est entièrement responsable pour le défaut d'induction des IFN-I et la forte réplication virale dans les MEFs de CC071. Cependant, après une infection *in vivo* par le virus ZIKV, les souris CC071 avaient une plus forte sensibilité que des souris C57BL/6J (B6) *Irf3* KO, démontrant que d'autres gènes de sensibilité sont impliqués dans la pathologie des souris CC071 après une infection *in vivo*.

Afin d'identifier ces gènes, j'ai analysé un backcross entre CC071 et B6 *Irf3* KO, et une F2 entre CC071 et la lignée 129/SvPas (129) *Ifnar1* KO. Leur analyse a permis l'identification de six QTLs

significatifs et vingt-trois QTLs suggestifs, associés avec la charge virale, perte de poids, symptômes cliniques et mortalité.

J'ai aussi étudié les loci qui modifient la sensibilité à l'infection par ZIKV de souris *Ifnar1* KO. Ce KO existe sur deux fonds génétiques, et mon laboratoire a précédemment démontré que les souris de la lignées B6 *Ifnar1* KO sont plus sensibles à l'infection par ZIKV, avec des symptômes sévères et de la mortalité, que les souris 129 *Ifnar1* KO, qui ne montrent que des symptômes légers. Pour identifier les gènes responsables de cette différence, j'ai produit et analysé un backcross entre les lignées B6 *Ifnar1* KO et 129 *Ifnar1* KO. Huit QTLs significatifs et dix-neuf QTLs suggestifs ont été associés avec la charge virale, la mortalité, la perte de poids et les symptômes cliniques des souris backcross.

Pour faciliter l'analyse des quatre croisements étudiés pendant mon projet de thèse, j'ai développé stuart, un package R en accès libre qui permet la curation des données de génotypage génome-entier de croisements à deux générations produits par puce pour une analyse QTL.

Mon travail a permis de mettre en évidence les facteurs génétiques impliqués dans la sensibilité à l'infection par le virus ZIKV chez ces modèles souris. De futures études affineront les intervalles des QTLs pour identifier et prioriser des gènes candidats en utilisant une combinaison d'approches décrites dans cette thèse. L'identification des gènes sous-jacents à ces QTLs apportera de nouvelles connaissances sur les mécanismes de pathogénie de l'infection par ZIKV. Enfin, puisqu'il a été récemment démontré que la lignée CC071 est également sensible à l'infection par d'autres virus, cette lignée sera utile dans le futur pour les études sur les maladies infectieuses.

**Title:** Genetic and functional characterization of innate susceptibility and resistance to Zika virus in mouse models

**Keywords:** genetics, Zika virus, virology, innate immune response, genetic control of susceptibility to infections, viral replication

**Abstract:** Infection of humans by Zika virus (ZIKV) can result in variable outcomes, from the absence of clinical symptoms to severe neurological complication or microcephaly after intra-uterine infection. Using the Collaborative cross, a collection of mouse inbred strains with large genetic diversity, my lab has previously demonstrated for the first time the importance of host genetic factors in this clinical variability, and identified CC071 as the most susceptible strain, with high viral loads in the blood, severe symptoms and mortality.

During my PhD thesis, I have first investigated the genetic origin of CC071's susceptibility using an *in vitro* model. Mouse embryonic fibroblasts (MEFs) from CC071 showed enhanced viral replication associated with delayed induction of type-I interferons (IFN-I). A quantitative trait locus (QTL) mapping analysis on MEFs derived from a backcross between CC071 and the resistant strain CC001 led to the identification of a loss of function mutation in the *Irf3* gene, specific to CC071. Functional analyses demonstrated that this mutation fully explains the defective IFN-I induction and uncontrolled viral replication in CC071 MEFs. However, after *in vivo* ZIKV infection, CC071 mice exhibited higher susceptibility than C57BL/6J (B6) *Irf3* KO mice, demonstrating the involvement of other susceptibility genes in the disease of CC071 mice *in vivo*.

To identify such genes, I analyzed a backcross between CC071 and B6 *Irf3* KO, and an F2 between CC071 and 129/SvPas (129) *Ifnar1*. Their analysis led to the identification of six significant QTLs and twenty-three suggestive QTLs, associated with the

viral load, body weight loss, clinical symptoms, and mortality.

I also investigated the loci which modify the susceptibility to ZIKV infection of *Ifnar1* KO mice. This KO exists on two genetic backgrounds, and my laboratory previously demonstrated that B6 *Ifnar1* KO mice are more susceptible to ZIKV infection, with severe symptoms and mortality, than 129 *Ifnar1* KO mice, which show only mild symptoms. To identify the genes responsible for this difference, I produced and analyzed a backcross between the B6 *Ifnar1* KO and 129 *Ifnar1* KO strains. Eight significant QTLs and nineteen suggestive QTLs were associated with the viral load, mortality, body weight loss and clinical symptoms of backcross mice.

To facilitate genetic analysis of the four crosses analyzed in my PhD project, I have developed stuart, an open-source R package which curates whole-genome genotyping data of second generation crosses produced by arrays for subsequent QTL mapping.

My work highlighted the genetic factors involved in the susceptibility to ZIKV infection in these mouse models. Future work will refine these QTL intervals to identify and prioritize candidate genes using a combination of approaches described in this work. Identifying genes underlying these QTLs will provide new insight into the mechanisms of pathogenesis of ZIKV infection. Lastly, as the CC071 strain was recently reported to be susceptible to the infection by other viruses, this strain will be increasingly useful for the study of infectious diseases.

## Affiliations

**université**  
**PARIS-SACLAY**

**ÉCOLE DOCTORALE**  
Agriculture, alimentation,  
biologie, environnement,  
santé (ABIES)



## Acknowledgements

En premier lieu, je tiens à remercier les membres de mon jury qui ont accepté d'évaluer mon travail. Je remercie mes rapporteurs Sébastien Nisole et Binnaz Yalcin pour avoir pris le temps de relire mon manuscrit. Je remercie Sophie Le Poder et Sabine Riffault pour leur rôle d'examinatrices lors de ma soutenance.

Je remercie l'école doctorale ABIES et en particulier Pierre Larraufie pour son suivi pendant ma thèse et pour son aide dans les démarches administratives.

Je remercie le LabEx IBEID *Integrative Biology of Emerging Infectious Diseases* pour mon financement de thèse.

Je souhaite également remercier les membres de mon comité de suivi individuel de thèse, Antonio Bandeira Ferreira, Nolwenn Jouvenet, Nicole Pavio et Silvia Vidal, pour avoir suivi mes avancées durant ces trois années. Merci pour votre expertise et vos conseils. Nolwenn, merci d'avoir pris le temps en dehors du comité pour répondre à mes questions sur des sujets scientifiques et de poursuite de carrière.

Je souhaite remercier Xavier Montagutelli qui a été un formidable directeur de thèse. Xavier, je me souviens encore de notre rencontre sur Skype pour le stage de M2. Je me souviens avoir eu l'intuition que ce serait génial de travailler avec toi, et je ne me suis pas trompée. Merci pour ton honnêteté et ta confiance en toutes circonstances, et merci de m'avoir toujours poussée à donner mon maximum. Mais aussi, merci pour ta gentillesse et ton sens de l'humour (et pour les tomates du jardin, et les œufs de tes poules). J'espère t'avoir convaincu qu'une boucle n'est pas toujours la meilleure solution dans un script !

Je remercie également tous les tuteurs que j'ai pu avoir pendant mon cursus universitaire : Robin Loesch, Mélanie Migaud, Indra Gupta, Aimee Ryan et Geoffrey Pires. Vous avez tous participé à faire de moi la scientifique que je suis aujourd'hui.

Je remercie tous les membres présents et passés du Laboratoire de Génétique de la Souris. Laurine, tout ça n'aurait pas été possible sans toi. Merci de m'avoir formée à mes débuts, puis d'avoir travaillé avec moi en binôme sur un grand nombre des manips présentées dans ce manuscrit. Caroline, merci pour cette base si solide que tu as construit pendant ta thèse. Tu as suivi le projet de loin pendant quelques temps puis tu es revenue parmi nous, et cela aura été un gros avantage pour moi de t'avoir dans le bureau pendant la rédaction de ce manuscrit. Lucie, merci de m'avoir montré que l'écriture de la thèse était un moment agréable. Merci pour ton soutien et tes conseils. Jean, je te remercie pour tes conseils toujours avisés tout au long de ce projet, et d'avoir pris le temps de prendre des nouvelles de l'écriture de mon manuscrit malgré ton emploi du temps très chargé. Aurélia, Gaetano, vous êtes arrivés à la fin de ce projet mais vous avez tout de suite été mis à contribution ! Merci pour votre aide avec les dernières manips. Je remercie également les anciens membres du labo. Anaïs, merci pour la bonne humeur que tu apportais au laboratoire. Corentin, j'ai adoré t'encadrer pendant ton stage. Merci pour ta participation à ce projet.

Je tiens à remercier tous nos collaborateurs, à Pasteur et ailleurs, qui ont participé de près ou

de loin à ce projet. Merci à Etienne Simon-Loriere, Michel Cohen-Tannoudji, Sandrine Vandormael-Pournin, Matthieu Prot, Maxime Chazal, et Elise Jacquemet pour leurs conseils techniques. Merci à Eliette Bonnefoy pour des discussions enrichissantes qui ont fait avancer ce projet.

Je remercie toute l'équipe de l'animalerie centrale, en particulier Rachid Chennouf, Sylvain Ciret et Mathilde Dubot pour le maintien des souris en animalerie A3, et Tommy Penel pour l'élevage des souris CC.

Je tiens également à remercier mes ami-es qui m'ont soutenue pendant mes études. Merci ma boulette (oui oui, tu as le droit à ce surnom jusque dans les remerciements de ma thèse). L'aventure à Pasteur aura commencé ensemble fin 2019, et malheureusement je ne serai plus là quand tu y reviendras. Merci pour avoir partagé avec moi les galères de manip et les recherches de postdoc. Je ne pars pas si loin l'année prochaine tu sais.

Merci les expeRts de m'avoir accepté dans votre cercle de bioinformaticiens. Ce petit groupe ressemble à quelques-uns de mes scripts : un but précis à l'origine puis on y met tout et n'importe quoi. Vous m'avez soutenue, même quand je ne comprenais pas ce qui était tidy. Je n'aurais pas pu atteindre le niveau que j'ai aujourd'hui sur R (et donc publié un package !) sans vous.

Merci également à Alexa. Je suis si heureuse d'avoir appris à te connaître au Canada. Merci d'avoir appris à être une vraie scientifique à mes côtés et de m'avoir soutenue pendant la thèse, depuis le cours de génétique de la souris !

Enfin, merci aux trois mousquetaires : Philo, Flora et Béa. Philo, merci de m'avoir proposé de m'apporter de la soupe pendant la rédaction de ce manuscrit. Et merci pour Kida, qui m'a apporté beaucoup de soutien émotionnel.

Merci Claire et Nico, pour votre soutien et pour le stéthoscope !

Je remercie également ma petite famille. Maman, Papa, merci d'avoir toujours cru en moi pendant ces 8 années d'étude. Vous m'avez soutenue dans tous mes choix d'orientation. Merci de m'avoir consolée quand je n'ai pas eu la fac que je voulais en première année, et de m'avoir aidé à m'inscrire dans la bonne ensuite. Merci de m'avoir soutenue pendant les révisions des partiels. Merci de m'avoir appelée toutes les semaines malgré le décalage horaire quand j'étais à l'étranger. Merci de m'avoir fait répéter mes oraux (alors que vous ne compreniez pas tout). Vous m'avez poussée à profiter au maximum de cette expérience universitaire qui touche à sa fin.

Merci Mamy d'avoir suivi mon travail de loin, je sais que tu es très fière de moi. J'aurais aimé que Papy soit là pour voir ça tu sais.

Patricia, Jean-Claude, merci pour votre intérêt dans mon sujet, et pour m'avoir permis de prendre des mini vacances au mois d'aout.

Bastien, je ne pouvais finir que par toi. Tu mérites plusieurs paragraphes de mes remerciements, et je pense que tu sais déjà à que point je te suis reconnaissante avant de lire ces lignes. Tu as supporté toutes mes histoires de cellules, de souris, de PCR et d'anticorps. Je suis désolée d'aimer autant te raconter en détail mes journées, mais ça va sûrement continuer. Tu dois au moins avoir un niveau L1 en bio maintenant, on ne va pas s'arrêter là ! Tu auras été mon roc pendant ces trois années. Merci d'avoir calmé mon stress et d'avoir fait passer ma thèse en priorité ces derniers mois. J'ai hâte de commencer cette nouvelle aventure avec toi en fin d'année.

## Contents

List of abbreviations . . . . .	10
List of Figures . . . . .	11
List of Tables . . . . .	12
<b>I Introduction</b>	<b>14</b>
<b>II State of the art</b>	<b>18</b>
II.1 The role of type I interferon response in viral infections . . . . .	19
II.1.1 Sensing of viral infections . . . . .	19
II.1.2 Induction of IFN-I expression . . . . .	19
II.1.3 The antiviral properties of Interferon Stimulated Genes . . . . .	20
II.1.3.1 Viral entry . . . . .	20
II.1.3.2 Nucleoprotein trafficking . . . . .	21
II.1.3.3 Viral protein production . . . . .	22
II.1.3.4 Genome replication and degradation . . . . .	22
II.1.3.5 Viral egress . . . . .	22
II.1.3.6 ISGs not affecting the viral cycle . . . . .	23
II.1.3.7 Regulation of IFN-I expression . . . . .	23
II.1.4 The roles of the genes of the IFN-I induction pathway in viral infections . . . . .	23
II.1.4.1 Reverse genetics studies . . . . .	23
II.1.4.2 The role of mouse and human genes of the IFN-I induction pathway in the susceptibility to viral infections . . . . .	25
II.1.4.3 The role of IRF3 to face viral infection . . . . .	66
II.2 Identification of susceptibility genes in humans . . . . .	83
II.2.1 Human hypothesis-free genetics methods . . . . .	83
II.2.1.1 Sequencing approaches . . . . .	83
II.2.1.2 Genome Wide Association Studies . . . . .	84
II.2.1.3 <i>In vitro</i> human genetic studies . . . . .	84
II.2.2 IFN-I and interferon-independent genes in hypothesis-free methods . . . . .	85
II.3 Mouse genetics of susceptibility to infectious diseases . . . . .	86
II.3.1 Experimental models for studying infectious diseases in mice . . . . .	86
II.3.2 Identifying susceptibility genes in segregating crosses . . . . .	87
II.3.2.1 Production of segregating crosses . . . . .	87
II.3.2.2 QTL mapping in segregating crosses . . . . .	88
II.3.2.3 Identifying candidate susceptibility genes in a QTL . . . . .	89
II.3.2.4 Confirming the causal role of a candidate gene . . . . .	90
II.3.2.5 Advantages of segregating crosses . . . . .	91
II.3.3 Identifying susceptibility genes in recombinant inbred strains . . . . .	92
II.3.3.1 Biparental recombinant inbred strains . . . . .	92
II.3.4 The Collaborative Cross . . . . .	93
II.3.4.1 Genetic analysis in the Collaborative Cross . . . . .	95
II.3.4.2 Infectious diseases studies with the Collaborative Cross . . . . .	96
II.4 Genetic susceptibility to Zika virus . . . . .	99
II.4.1 Zika virus epidemiology . . . . .	99
II.4.1.1 ZIKV emergence and epidemics . . . . .	99
II.4.1.2 ZIKV evolution . . . . .	100
II.4.1.3 ZIKV transmission . . . . .	100



II.4.1.4	ZIKV disease . . . . .	101
II.4.2	Biology of Zika virus . . . . .	101
II.4.2.1	ZIKV tropism and cell cycle . . . . .	101
II.4.2.2	Structure of ZIKV particles . . . . .	102
II.4.3	Host genetic susceptibility to ZIKV infection in human . . . . .	103
II.4.3.1	Sequence analyses . . . . .	103
II.4.3.2	Association analyses . . . . .	103
II.4.3.3	<i>In vitro</i> analyses . . . . .	104
II.4.4	Mouse models of ZIKV infection . . . . .	104
II.4.5	Previous results on ZIKV infection in the Collaborative Cross . . . . .	106
<b>III</b>	<b>Objectives</b>	<b>109</b>
<b>IV</b>	<b>Material and methods</b>	<b>111</b>
<b>V</b>	<b>Results</b>	<b>122</b>
V.1	The delayed induction of IFN-I in CC071 cells is due to a loss-of-function mutation in the <i>Irf3</i> gene . . . . .	123
V.2	Deciphering the <i>in vivo</i> susceptibility to ZIKV of CC071 mice . . . . .	150
V.2.1	Introduction . . . . .	150
V.2.2	Results . . . . .	150
V.2.2.1	Phenotypic analysis of CC071, B6- <i>Irf3</i> , F1 and N2 mice . . . . .	150
V.2.2.2	QTL mapping identified two significant and five suggestive QTLs associated with viral loads of N2- <i>Irf3</i> mice . . . . .	154
V.2.2.3	QTL mapping identified one significant and two suggestive QTLs associated with the body weight loss of N2- <i>Irf3</i> mice . . . . .	161
V.2.2.4	Three suggestive QTLs are associated with the clinical scores of N2- <i>Irf3</i> mice . . . . .	163
V.2.2.5	<i>Irf3</i> does not contribute to the phenotypic variability of the N2- <i>Irf3</i> progeny . . . . .	164
V.2.2.6	Phenotypic analysis of CC071, 129- <i>Irfnar1</i> , and F2 mice . . . . .	164
V.2.2.7	QTL mapping identifies one QTL controlling the plasma viral loads of F2- <i>Irfnar1</i> mice . . . . .	168
V.2.2.8	Two significant QTLs are associated with the survival of F2- <i>Irfnar1</i> mice . . . . .	170
V.2.2.9	QTL mapping identifies seven loci suggestively associated with the body weight loss of F2- <i>Irfnar1</i> mice . . . . .	170
V.2.2.10	Two significant and seven suggestive QTLs control the clinical scores of F2- <i>Irfnar1</i> mice . . . . .	173
V.2.2.11	Could <i>Irf3</i> contribute to the QTLs identified in the F2- <i>Irfnar1</i> cross? . . . . .	173
V.2.3	Discussion . . . . .	176
V.3	Genetic factors modulating ZIKV susceptibility in <i>Irfnar1</i> deficient mice . . . . .	180
V.3.1	Introduction . . . . .	180
V.3.2	Results . . . . .	181
V.3.2.1	Phenotypic analysis of the parental strains and F1s . . . . .	181
V.3.2.2	Phenotypic analysis of the N2 progeny . . . . .	181
V.3.2.3	Three significant QTLs control plasma viral load in N2 mice . . . . .	184
V.3.2.4	QTL mapping identified two loci significantly associated with the survival of N2 mice . . . . .	188
V.3.2.5	One significant and five suggestive QTLs control the body weight loss of N2 mice . . . . .	191
V.3.2.6	Three significant and five suggestive QTLs control the clinical scores of N2 mice . . . . .	191
V.3.2.7	Comparison of the QTLs found in the F2 and N2 crosses . . . . .	194

V.3.3	Discussion . . . . .	195
V.4	Meta-analysis of the three QTL mappings . . . . .	199
V.4.1	Relevance of a meta-analysis . . . . .	199
V.4.2	Opportunities to find common QTLs in the three crosses . . . . .	200
V.4.2.1	Study of common traits . . . . .	200
V.4.2.2	Influence of the experimental model . . . . .	201
V.4.2.3	Investigation of different susceptibility alleles . . . . .	201
V.4.3	Analysis of the colocalization of QTLs between the three crosses . . . . .	204
V.4.4	Refinement of QTLs credible intervals . . . . .	206
V.4.5	Candidate gene analysis . . . . .	206
<b>VI</b>	<b>Discussion</b>	<b>210</b>
VI.1	The contribution of cellular models to mouse genetic studies . . . . .	211
VI.2	Segregating crosses using Collaborative Cross and <i>Ifnar1</i> deficient strains . . . . .	212
VI.3	Perspectives on the identification of susceptibility alleles . . . . .	214
VI.4	CC071: a complex model of viral disease . . . . .	217
VI.5	The Collaborative Cross: a model of the genetic and phenotypic diversity of the human population . . . . .	219
<b>VII</b>	<b>References</b>	<b>221</b>
<b>VIII</b>	<b>Appendix</b>	<b>233</b>
VIII.1	Supplementary Figures . . . . .	234
VIII.2	Supplementary Tables . . . . .	244
VIII.3	Résumé substantiel en français . . . . .	246
VIII.3.1	Introduction . . . . .	246
VIII.3.2	Résultats . . . . .	249
VIII.3.2.1	Le retard d'induction des IFN-I dans les MEFs CC071 est dû à une mutation perte de fonction dans le gène <i>Irf3</i> . . . . .	249
VIII.3.2.2	Décryptage de la sensibilité <i>in vivo</i> des souris CC071 . . . . .	250
VIII.3.2.3	Les facteurs génétiques modifiant la sensibilité des souris <i>Ifnar1</i> KO . . . . .	251
VIII.3.3	Discussion . . . . .	251

## List of abbreviations

<b>129</b> 129S2/SvPas	<b>HIV</b> human immunodeficiency virus	<b>PAMP</b> pathogen associated molecular pattern
<b>129-<i>Ifnar1</i></b> 129S2/SvPas- <i>Ifnar1</i> <sup>-/-</sup>	<b>hpi</b> hour post infection	<b>PFU</b> plaque-forming unit
<b>B10.D1</b> B10.D1-H2q/Sgj	<b>HSV-2</b> herpes simplex virus 2	<b>POWV</b> Powassan virus
<b>B6</b> C57BL/6J	<b>HSV-1</b> herpes simplex virus 1	<b>PRR</b> pathogen recognition receptor
<b>B6-<i>Ifnar1</i></b> C57BL/6J- <i>Ifnar1</i> <sup>-/-</sup>	<b>i.c.</b> intracranial	<b>PVL</b> plasma viral load
<b>B6-<i>Irf3</i></b> C57BL/6J- <i>Irf3</i> <sup>-/-</sup>	<b>i.p.</b> intraperitoneal / intraperitoneally	<b>QTL</b> quantitative trait locus
<b>BMM</b> bone marrow macrophage	<b>IAV</b> influenza A virus	<b>RI</b> recombinant inbred
<b>BVL</b> brain viral load	<b>IFN</b> interferon	<b>RIKA</b> repression of IRF3-mediated NF- $\kappa$ B activity
<b>CC</b> Collaborative Cross	<b>IFN-I</b> type I interferon	<b>RIPA</b> RIG-I-like receptor-induced IRF3 mediated pathway of apoptosis
<b>CC001</b> CC001/Unc	<b>IRF</b> interferon regulatory factor	<b>RLR</b> RIG-I-like receptor
<b>CC071</b> CC071/TauUnc	<b>IRG</b> interferon-repressed gene	<b>RNA-seq</b> RNA sequencing
<b>CTC</b> Complex Trait Consortium	<b>ISG</b> interferon-stimulated gene	<b>RVFV</b> Rift Valley fever virus
<b>CZS</b> congenital Zika syndrome	<b>ISGF3</b> interferon-stimulated gene factor 3	<b>scRNA-seq</b> single-cell RNA sequencing
<b>DENV</b> dengue virus	<b>JEV</b> Japanese encephalitis virus	<b>SGI</b> second generation individual
<b>DO</b> Diversity Outbred	<b>KD</b> knockdown	<b>shRNA</b> small hairpin RNA
<b>dpi</b> day post infection	<b>KO</b> knockout	<b>siRNA</b> small interfering RNA
<b>dsRNA</b> double-stranded RNA	<b>LCMV</b> lymphocytic choriomeningitis virus	<b>SIV</b> simian immunodeficiency virus
<b>EBV</b> Epstein-Barr virus	<b>LOD</b> logarithm of the odds	<b>SNP</b> single nucleotide polymorphism
<b>EMCV</b> encephalomyocarditis virus	<b>Mb</b> megabase	<b>ssRNA</b> single-stranded RNA
<b>ER</b> endoplasmic reticulum	<b>MEF</b> mouse embryonic fibroblast	<b>TBEV</b> tick-borne encephalitis virus
<b>ERV</b> endogenous retrovirus	<b>MGI</b> mouse genome informatics	<b>TLR</b> Toll-like receptor
<b>F1-<i>Ifnar1</i></b> (B6- <i>Ifnar1</i> $\times$ 129- <i>Ifnar1</i> ) F1	<b>MHC</b> major histocompatibility complex	<b>TMEV</b> Theiler's murine encephalomyelitis virus
<b>F2</b> intercross	<b>miRNA</b> micro RNA	<b>tRNA</b> transfer RNA
<b>F2-<i>Ifnar1</i></b> (CC071 $\times$ 129- <i>Ifnar1</i> ) F2	<b>MNV</b> murine norovirus	<b>WES</b> whole exome sequencing
<b>FFU</b> focus-forming unit	<b>MOI</b> multiplicity of infection	<b>WNV</b> West Nile virus
<b>FG15</b> French Guiana 2015	<b>N2</b> backcross	<b>WT</b> wild type
<b>FVB</b> FVB/NJ	<b>N2-<i>Irf3</i></b> (CC071 $\times$ B6- <i>Irf3</i> ) $\times$ CC071 N2	<b>YFV</b> yellow fever virus
<b>GRP</b> genetic reference population	<b>N2-B6.129-<i>Ifnar1</i></b> (B6- <i>Ifnar1</i> $\times$ 129- <i>Ifnar1</i> ) $\times$ B6- <i>Ifnar1</i> N2	<b>ZIKV</b> Zika virus
<b>GWAS</b> genome wide association study	<b>NHEJ</b> non-homologous end joining	<b>Zsl</b> Zika susceptibility locus
<b>HAV</b> hepatitis A virus	<b>NrHV</b> Norway rat hepacivirus	
<b>HBV</b> hepatitis B virus	<b>OAS</b> 2'-5' oligoadenylate	
<b>HCV</b> hepatitis C virus		
<b>HDR</b> homology-directed repair		
<b>hiPSC</b> human induced pluripotent stem cell		

## List of Figures

Figure 1	ISGs targeting different steps of the viral replication cycle. . . . .	21
Figure 2	The CRISPR/Cas9 system for KO and point mutations . . . . .	24
Figure 3	Models to study viral infections in the mouse. . . . .	86
Figure 4	Breeding schemes to produce genetically diverse mouse populations. . . . .	88
Figure 5	Example of QTL mapping results . . . . .	89
Figure 6	Quantitative complementation test method . . . . .	91
Figure 7	Breeding schemes for the production biparental RI strains. . . . .	92
Figure 8	Collaborative Cross strains breeding funnel. . . . .	94
Figure 9	Example of CC strain genome. . . . .	95
Figure 10	CC experiment workflow. . . . .	96
Figure 11	Phylogenic tree of ZIKV strains. . . . .	100
Figure 12	ZIKV replication cycle. . . . .	102
Figure 13	ZIKV structure. . . . .	103
Figure 14	Genetic background impacts clinical symptoms and PVL after ZIKV infection in CC mice. . . . .	106
Figure 15	CC071 MEFs show increased viral replication. . . . .	108
Figure 16	CC071 MEFs show delayed <i>Ifnb1</i> expression . . . . .	108
Figure 17	Objectives of my PhD project. . . . .	110
Figure 18	B6 <i>Irf3</i> KO mice do not develop ZIKV disease. . . . .	151
Figure 19	Phenotypic variation and trait correlation in N2- <i>Irf3</i> mice. . . . .	153
Figure 20	Single QTL mapping for the viral loads in the N2- <i>Irf3</i> progeny. . . . .	155
Figure 21	<i>Zs11</i> and <i>Zs12</i> on chromosome 12 associated with the PVL and BVL in the N2- <i>Irf3</i> progeny. . . . .	156
Figure 22	Chromosome 2 QTLs associated with the PVL in the N2- <i>Irf3</i> progeny. . . . .	157
Figure 23	Chromosome 15 QTL QTLs associated with the PVL in the N2- <i>Irf3</i> progeny. . . . .	158
Figure 24	Additive effects of the three QTLs controlling the PVL at 2 dpi of N2- <i>Irf3</i> mice. . . . .	159
Figure 25	Chromosome 7 QTLs associated with the PVL in the N2- <i>Irf3</i> progeny. . . . .	159
Figure 26	Two-dimensional QTL mapping for the PVL at 6 dpi the N2- <i>Irf3</i> progeny. . . . .	160
Figure 27	Single QTL mapping for the body weight loss in the N2- <i>Irf3</i> progeny showing <i>Zs13</i> and two other suggestive QTLs. . . . .	162
Figure 28	Single QTL mapping for the body weight loss in the N2- <i>Irf3</i> progeny. . . . .	163
Figure 29	Allelic effects for <i>Irf3</i> on body weight loss at 5 dpi in the N2- <i>Irf3</i> progeny. . . . .	164
Figure 30	mAb-treated CC071 and 129- <i>Ifnar1</i> mice have different mortality rate but similar plasma viral loads. . . . .	165
Figure 31	F2- <i>Ifnar1</i> mice show great phenotypic variability. . . . .	167
Figure 32	QTL mapping for the PVL in the F2- <i>Ifnar1</i> progeny. . . . .	169
Figure 33	<i>Zs14</i> on chromosome 2 associated with the PVL in the F2- <i>Ifnar1</i> progeny. . . . .	169
Figure 34	QTL mapping for the susceptibility of the F2- <i>Ifnar1</i> progeny. . . . .	171
Figure 35	<i>Zs15</i> and <i>Zs16</i> on chromosome 7 associated with the survival and day of death of F2- <i>Ifnar1</i> mice. . . . .	171
Figure 36	QTL mapping for the body weight loss in the F2- <i>Ifnar1</i> progeny. . . . .	172
Figure 37	QTL mapping for the clinical scores in the F2- <i>Ifnar1</i> progeny displaying <i>Zs15</i> , <i>Zs16</i> and seven other suggestive QTLs. . . . .	174
Figure 38	Allelic effects for <i>Irf3</i> on the different viral load and clinical manifestations of F2- <i>Ifnar1</i> mice. . . . .	175
Figure 39	Summary of the QTLs found in the two crosses, highlighting potential pleiotropy. . . . .	179
Figure 40	ZIKV disease severity in <i>Ifnar1</i> -deficient mice is controlled by the genetic background. . . . .	182
Figure 41	N2 mice show large phenotypic variability. . . . .	183
Figure 42	QTL mapping for the PVL. . . . .	184
Figure 43	<i>Zs17</i> on chromosome 1 associated with the PVL. . . . .	185
Figure 44	<i>Zs18</i> on chromosome 3 associated with the PVL. . . . .	187
Figure 45	<i>Zs19</i> on chromosome X associated with the PVL. . . . .	188

Figure 46	QTL mapping for the susceptibility of the N2 mice. . . . .	189
Figure 47	<i>Zs/10</i> on chromosome 1 associated with the survival and day of death of N2 mice. . . . .	189
Figure 48	<i>Zs/11</i> on chromosome 4 associated with the survival and day of death of N2 mice. . . . .	190
Figure 49	Additive effects of the two significant QTLs controlling the survival of N2 mice. . . . .	190
Figure 50	QTL mapping for the body weight loss in the N2 progeny displaying <i>Zs/12</i> and five suggestive QTLs. . . . .	192
Figure 51	QTL mapping for the clinical scores in the N2 progeny displaying <i>Zs/11</i> , <i>Zs/13</i> , <i>Zs/14</i> and five suggestive QTLs. . . . .	193
Figure 52	Allelic effects for the QTLs identified in the F2 progeny on the phenotypes of the N2 progeny. . . . .	194
Figure 53	Allelic effects for the QTLs identified in the N2 progeny on the phenotypes of the F2 progeny. . . . .	195
Figure 54	Summary of the QTLs found in the N2 progeny . . . . .	197
Figure 55	Segregating crosses analyzed <i>in vivo</i> in my PhD project. . . . .	199
Figure 56	Summary of the QTLs found in the three <i>in vivo</i> crosses . . . . .	202
Figure 57	Colocalizing QTLs on chromosome 15 . . . . .	204
Figure 58	Colocalizing QTLs controlling the PVL on chromosome 2 . . . . .	205
Figure 59	Expression of <i>Bcl11b</i> and <i>Dlk1</i> in CC071 and B6 MEFs. . . . .	208
Figure 60	Implication of the significant QTLs identified in my PhD on the physiopathology of ZIKV disease. . . . .	213
Figure S1	Effect of sex on the phenotypes of the N2- <i>Irf3</i> progeny. . . . .	234
Figure S2	Marker density maps for the N2- <i>Irf3</i> and F2- <i>Irfnar1</i> progenies . . . . .	234
Figure S3	Allelic effects for <i>Irf3</i> on the different phenotypes tested in the N2- <i>Irf3</i> progeny. . . . .	235
Figure S4	Effect of sex on the phenotypes of the F2 progeny. . . . .	235
Figure S5	Allelic effects for the QTLs controlling the body weights of F2- <i>Irfnar1</i> mice. . . . .	236
Figure S6	Allelic effects for the QTLs controlling the clinical scores of F2- <i>Irfnar1</i> mice. . . . .	237
Figure S7	Allelic effects for <i>Irf3</i> on the different phenotypes tested in the F2- <i>Irfnar1</i> progeny. . . . .	238
Figure S8	Effect of sex on the phenotypes of the N2 progeny. . . . .	239
Figure S9	Marker density map for the B6- <i>Irfnar1</i> × 129- <i>Irfnar1</i> N2 progeny. . . . .	239
Figure S10	Suggestive QTLs associated with the PVL of B6- <i>Irfnar1</i> × 129- <i>Irfnar1</i> N2 mice. . . . .	240
Figure S11	Suggestive QTLs associated with the survival and day of death of B6- <i>Irfnar1</i> × 129- <i>Irfnar1</i> N2 mice. . . . .	241
Figure S12	Allelic effects for the QTLs controlling the body weights of B6- <i>Irfnar1</i> × 129- <i>Irfnar1</i> N2 mice. . . . .	242
Figure S13	Allelic effects for the QTLs controlling the clinical scores of B6- <i>Irfnar1</i> × 129- <i>Irfnar1</i> N2 mice. . . . .	243

## List of Tables

Table 1	Associations between QTLs and phenotypes in the N2- <i>lrf3</i> progeny. . . . .	154
Table 2	Associations between QTLs and phenotypes in the F2- <i>lfnar1</i> progeny. . . . .	168
Table 3	QTLs in the vicinity of <i>lrf3</i> in the F2- <i>lfnar1</i> progeny. . . . .	176
Table 4	Associations between QTLs and phenotypes in the N2 progeny. . . . .	186
Table 5	Effects of the QTLs found in the F2 and N2 crosses . . . . .	196
Table 6	Phenotypes studied in the three crosses . . . . .	200
Table 7	Genotypes found in the three crosses . . . . .	204
Table S1	Genes in <i>Zs1</i> . . . . .	244

# I - Introduction

Communicable diseases are still one of the leading causes of death worldwide, particularly in low-income countries where they account for six of the top ten causes of death, versus only three at the global level (WHO 2020). Viruses are the most diverse group of pathogens. Their genomes can be either DNA or RNA, single- or double-stranded, segmented or not, and enveloped or not. They are responsible for acute (e.g. influenza, enteroviruses) or chronic (e.g. herpesviruses, human immunodeficiency virus (HIV)) infections; for severe (e.g. Ebola, rabies), moderate (e.g. smallpox) or debilitating (e.g. poliovirus) conditions, or induce cancers (e.g. hepatitis B virus (HBV), Epstein-Barr virus (EBV), papillomaviruses). They can be transmitted by various modes: direct contact with an infected person (e.g. herpesviruses), spread of aerosols by sneezing or coughing (e.g. influenza, rhinoviruses), through vehicles such as food and water (e.g. hepatitis A virus (HAV)), vertically from the mother to the fetus (e.g. HIV, herpesviruses) or by vectors such as mosquitoes (e.g. yellow fever virus (YFV), dengue virus (DENV), West Nile virus (WNV)) or ticks (e.g. tick-borne encephalitis virus (TBEV)). Although vaccines are the most effective preventive measures against viral infections, many viruses still lack vaccines such as HIV, hepatitis C virus (HCV), and Chikungunya virus.

Viral genomes are highly mutable, especially those of RNA viruses, whose mutation rates can be a million times higher than those of their vertebrate hosts. Faster genome replication provides a selective advantage to the virus, but faster RNA polymerases make more mistakes and therefore increase mutation rates (Duffy 2018). These mutations allow viruses to expand their host range and to evade post-infection and vaccine-induced immunity, triggering the emergence or re-emergence of viral infection, and leading to epidemics and pandemics. This was recently exemplified with SARS-CoV-2, which most probably originates from a bat coronavirus (Alwine et al. 2023). Successive variants have shown increased host range and transmissivity, but decreased pathogenicity (Bálint et al. 2022; Montagutelli et al. 2021b). Moreover, the vaccine efficacy against COVID-19 depends on the virus variant (Andrews et al. 2022). Other examples of viruses transmitted from animals to humans include HIV, which originates from simian immunodeficiency virus (SIV) (Sharp and Hahn 2011), and influenza A virus (IAV), which transmits from farm animals (poultry and swine) to humans and vice versa (Kessler et al. 2021).

Vector-borne viral diseases are a major concern due to global warming. Rising temperatures are expected to promote the spread of mosquito vectors to new geographic areas and to extend the viral transmission season. Indeed, reported cases of DENV infection have risen from 500,000 in 2000 to more than 5 million in 2019 (WHO 2023). In addition, many of these viruses can infect both humans and various animal species (WNV infects birds, TBEV infects small vertebrates such as rodents), which are reservoirs responsible for spillover events to humans. Climate change will also affect these reservoirs. For example, it is affecting bird diversity, which may favor species of birds with high WNV transmission potential (Rocklöv and Dubrow 2020; Fay et al. 2022).

Zika virus (ZIKV) is a mosquito-borne virus of the genus *flavivirus* which includes DENV, YFV and WNV, among others. ZIKV was first isolated from a rhesus monkey in Uganda in 1947, and the first case of human infection was reported in this country in 1964 (Talero-Gutiérrez et al. 2018). For 50 years, the circulation of the virus was mostly detected by serological surveillance, and only a few cases of ZIKV infection in humans were reported (Musso and Gubler 2016). However, since 2007, several outbreaks have been described, first in the Pacific Islands (Duffy et al. 2009), before a major



epidemic occurred in Brazil in 2015 and eventually spread to 48 countries in the Americas (Ikejezie 2017). While cases of human ZIKV infection have globally declined since 2017, ZIKV transmission continues in endemic regions (Yakob 2022). Moreover, new regions are reporting ZIKV infections: the first cases of autochthonous infection in Europe were reported in 2019 in Southern France (Giron et al. 2019), and two outbreaks occurred in India in 2018 and 2021 (Bardhan et al. 2021).

ZIKV infection is usually asymptomatic or causes mild symptoms such as rash, fever, muscle and joint pain, and headache. In rare cases, ZIKV infection can cause neurological complications such as Guillain-Barré syndrome, which clinical manifestations include bilateral and unilateral facial palsies, limb weakness and paresthesia (i.e. feeling of tingling) (Carod-Artal 2018). During the epidemic in Brazil, ZIKV infection in pregnant women was associated with fetal microcephaly, with more than 8000 cases reported between December 2015 and July 2016 (Carod-Artal 2018). To date, there is no specific treatment or vaccine against ZIKV (WHO 2022).

The variable presentation of ZIKV disease reflects the general case of viral infections which severity often depends on the viral strain, inoculum, and route of infection, as well as host-related factors, including genetics, age, sex, comorbidities, nutrition, stress, immune competence, pre-existing immunity, microbiome, and environmental factors such as hygiene conditions and pollution among others (Zsichla and Müller 2023). Host genetic variants can contribute significantly to the susceptibility to developing severe conditions as shown for example for dengue hemorrhagic fever (Pare et al. 2020) and COVID-19 (Niemi et al. 2022). Variants in immune response genes are often associated with viral disease outcome (Kenney et al. 2017; Manet et al. 2018; Zhang et al. 2020; Bourdon et al. 2020), but genetic studies have also revealed the role of genes involved in other biological processes. Indeed, human genetic studies have found association between viral disease severity and variants in genes involved in protein transport (*SNX8* and *ANKRD27* for measles virus) and post-translational modifications (*PRMT6* for HIV), mitochondrial protein translation (*MRPL10* and *TSMF* for measles virus), chromatin remodeling (*METTL21B* for IAV and measles virus) and lipid metabolism (*SOAT1* for measles virus, *AMACR* for rubella virus, *CPT2* for IAV) (Chen et al. 2005; Le Clerc et al. 2009; Mak et al. 2011; Zhu et al. 2023).

Four studies in humans have identified genetic factors associated with microcephaly in children born to mothers infected with ZIKV during pregnancy, in genes involved in innate immunity (Santos et al. 2019; Aguiar et al. 2020), extracellular matrix organization (Aguiar et al. 2020), mTOR pathway (De O. da Silva et al. 2021) and adenylate cyclase pathway (Rossi et al. 2019), but genetic factors involved in disease severity in infected adults have not been reported. The study of genetic susceptibility to viral infections in humans is hampered by variable infection parameters (viral strain, dose, and mode of infection), and other host factors that impact disease severity such as previous or coinfection, microbiome, and lifestyle (Leist and Baric 2018). Mouse models, on the other hand, allow for experimental infections, in which these factors are controlled. However, mice are not a natural host for ZIKV infection since, unlike humans, they are able to mount a protective type I interferon (IFN-I) response following ZIKV infection. Thus, the study of ZIKV infection in mice requires the use of immunodeficient mice, usually lacking the IFN-I receptor (*Ifnar1* deficient mice), or mice sensitized with a monoclonal antibody blocking this receptor (Sheehan et al. 2006; Lazear et al. 2016).

Previous work from our laboratory has investigated the influence of host genetic diversity on the

susceptibility to ZIKV in mouse models (Manet 2019). Genetically diverse mouse strains were used, either carrying *Ifnar1* deficiency: C57BL/6J-*Ifnar1*<sup>-/-</sup> (B6-*Ifnar1*) and 129S2/SvPas-*Ifnar1*<sup>-/-</sup> (129-*Ifnar1*), or sensitized immunocompetent mice from the Collaborative Cross (CC), a collection of mouse strains with large genetic variability. Differences in susceptibility were observed between these strains: B6-*Ifnar1* mice were more susceptible than 129-*Ifnar1* mice, with severe clinical signs and high mortality rate, despite similar peak viral load in the plasma. In the 35 CC strains tested, CC071/TauUnc (CC071) mice were the most susceptible with severe clinical symptoms, mortality, and high viral load in the plasma, and CC071 fibroblasts displayed high viral replication and delayed IFN-I induction after *in vitro* infection. The aim of my PhD project was to dissect the susceptibility of these strains to identify genes associated with increased severity following ZIKV infection, using genetic and functional method, in *in vitro* and *in vivo* models.

## **II - State of the art**

## II.1 . The role of type I interferon response in viral infections

### II.1.1 . Sensing of viral infections

Innate immunity is the first line of defense against infections, and provides immediate, broad, but non-specific response. It is activated by the recognition of conserved microbial features called pathogen associated molecular patterns (PAMPs). Viral PAMPs, notably DNAs and RNAs, are recognized as non-self as they differ from host molecules. For instance, double-stranded RNA (dsRNA) viral genomes are identified as being of pathogenic origin since mammalian cells do not normally produce long dsRNAs (Takeuchi and Akira 2009). Viral single-stranded RNAs (ssRNAs) are poorly processed and do not carry the same post-transcriptional modifications as cellular mRNAs such as the 5' cap structure (Gebhardt et al. 2017). Viral DNAs are unmethylated which allows to distinguish them from host DNA molecules (Kumagai et al. 2008).

Viral PAMPs are recognized by two families of pathogen recognition receptors (PRRs): RIG-I-like receptors (RLRs) and Toll-like receptors (TLRs). Two out of the three RLRs recognize viral RNA: DDX58 (also known as RIG-I) senses uncapped ssRNA, while IFIH1 (also known as MDA5) senses long dsRNA. The third RLR, DHX58 (also known as LGP2), is a regulator of DDX58 and IFIH1 signaling (Takeuchi and Akira 2009). TLRs recognize viral and bacterial molecules. Among them, three bind viral RNA: TLR3 (dsRNA), TLR7 and TLR8 (ssRNA). TLR9 recognizes unmethylated viral DNAs (Kumagai et al. 2008; Takeuchi and Akira 2009). Once these receptors sense viral molecules, they induce a signaling cascade leading to the expression of interferons (IFNs) (Levy et al. 2011).

### II.1.2 . Induction of IFN-I expression

IFNs were first described in 1957 by Isaacs and Lindenmann as factors capable of interfering with viral replication (Isaacs and Lindenmann 1957). There are 3 families of IFNs, the largest one being type I interferons (IFN-I) with 17 members in humans and 18 in mice (Lazear et al. 2019). IFN-I are highly conserved across animals as they are present in mammals, birds and fish (Zhou et al. 2014; Boudinot et al. 2016). The other types of IFN are type II IFN (IFN $\gamma$ ), and type III IFN (IFN $\lambda$ ). IFN $\gamma$  stimulates adaptive antiviral response by favoring antigen presentation and maturation of dendritic cells and macrophages (Lee and Ashkar 2018). IFN $\lambda$  are more similar to IFN-I as they are induced by the same transcription factors and they induce essentially the same repertoire of antiviral molecules (Donnelly and Kotenko 2010).

The main IFN-I are IFN $\alpha$  (13 subtypes in humans and 14 in mice) and IFN $\beta$ . Genes coding for IFN-I are expressed at low levels in basal condition, but are quickly induced after the recognition of PAMPs by PRRs. Once these receptors bind PAMPs, they recruit adaptor proteins (MAVS for RLRs, MYD88 or TICAM1/2 for TLRs) which activate kinases (IRAK1, TBK1 and IKK $\epsilon$ ). These kinases phosphorylate interferon regulatory factors (IRFs) which dimerize and enter the nucleus to bind the promoters of IFN-I genes and induce their expression (Bourdon et al. 2020). This signaling pathway is more thoroughly detailed in section II.1.4.2.

Once produced, IFN-I are secreted and bind their receptor IFNAR, either on the cell that produced

them or on neighboring cells, thus signaling in an autocrine and paracrine manner (Lazear et al. 2019). IFNAR is composed of two subunits IFNAR1 and IFNAR2, which are associated with the kinases TYK2 and JAK1, respectively. The binding of IFN-I on IFNAR induces the phosphorylation of TYK2 and JAK1, which in turn phosphorylate the transcription factors STAT1 and STAT2 (Zanin et al. 2021). Once activated, STAT1 and STAT2 dimerize and recruit IRF9 to form the interferon-stimulated gene factor 3 (ISGF3) complex. ISGF3 translocates into the nucleus to induce the expression of interferon-stimulated genes (ISGs), which are the actors of this antiviral pathway (Paul et al. 2018).

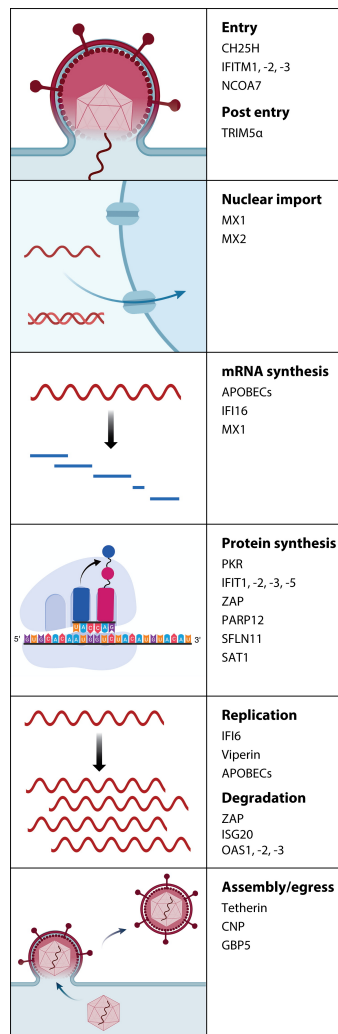
### **II.1.3 . The antiviral properties of Interferon Stimulated Genes**


ISGs are defined as genes upregulated by IFN signaling. ISGs are abundant in mammalian genomes: they are thought to number between 400 and more than 1,000 depending on the species or cell type (Schoggins and Rice 2011; Mostafavi et al. 2016; Shaw et al. 2017; Schoggins 2019). Interestingly, a cross-species RNA sequencing (RNA-seq) study identified 90 genes that were upregulated after IFN treatment in primary fibroblasts of 9 mammalian species, including 62 also present in chicken (Shaw et al. 2017). Moreover, IFN-I signaling also leads to the downregulation of genes called interferon-repressed genes (IRGs). The same study identified 479 IRGs in humans and 157 in mice. These genes are less conserved than ISGs, as none of them were shared by the 10 species studied (Shaw et al. 2017).

ISGs possess various functions to combat viral infection, such as inhibition of viral replication, antigen presentation, cell signaling and apoptosis (Shaw et al. 2017). Most of them are known to inhibit viral replication by targeting different stages of the viral cycle: entry, trafficking, protein production, genome replication, assembly and exit (Schoggins 2019) (Figure 1). Here are several examples of ISGs with different antiviral mechanisms.

#### **II.1.3.1 . Viral entry**

IFITM3 is an endosomal transmembrane protein that inhibits the release of enveloped viruses into the cytoplasm. Enveloped viruses enter host cells by endocytosis and then fuse their envelope with endosomal membranes to release their genome into the cell for subsequent replication and protein production. IFITM3 possesses an amphipathic  $\alpha$ -helix, i.e., with one hydrophilic side and one hydrophobic side, which locates in the cytoplasmic leaflet of the endosomal membrane (Jiménez-Munguía et al. 2022). This helix alters the curvature of the endosomal membrane, thereby inhibiting its fusion with the viral envelope. Since the fusion is blocked, the endosome progresses to the lysosome where the virus is degraded. IFITM3 has been shown to act against several enveloped viruses including IAV, HIV, DENV, ZIKV and SARS-CoV-2 (Das et al. 2021).



 Schoggins JW, 2019.  
*Annu. Rev. Virol.* 6:567–84

**Figure 1:** ISGs targeting different steps of the viral replication cycle. Viral replication steps are depicted on the left boxes. Examples of ISGs targeting each step are indicated on the right (from Schoggins 2019).

### II.1.3.2 . Nucleoprotein trafficking

MXA, encoded by the *Mx1* gene, is the first anti-IAV host restriction factor identified. IAV viral genome replication requires the import of nucleoproteins (i.e., complexes with IAV RNA genome fragments and proteins) into the nucleus. MXA possesses a disordered loop which binds viral nucleoproteins and inhibits their nuclear import, thereby blocking mRNA synthesis and genome replication (Haller and Kochs 2020).

The antiviral properties of MXA were discovered thanks to the use of a rare mouse inbred strain, A2G, which, unlike most common laboratory strain, showed poor viral replication after infection with IAV. Indeed, most laboratory strains possess a defective allele for *Mx1*, making them susceptible to IAV infection, while A2G mice carry a functional *Mx1* allele (Horisberger et al. 1983). Therefore, *Mx1* is a major host genetic factor controlling IAV infection outcome in mice. Polymorphisms in *MX1* in humans have not been associated with IAV susceptibility, but no complete loss-of-function mutations was reported (Ciancanelli et al. 2016; Haller and Kochs 2020). However, a single nucleotide polymorphism (SNP) in *MX1* was associated with symptomatic WNV infection (Bigham et al. 2011).

### **II.1.3.3 . Viral protein production**

SLFN11 is a protein of the Schlafen family with multiple functions related to its three domains: endoribonuclease, protein-interacting, and helicase/ATPase. SLFN11 is known to inhibit cancer cells proliferation by interacting with ribosomal proteins and thereby inhibiting signaling pathways involved in cell growth (Jo and Pommier 2022). SLFN11 was also shown to inhibit HIV protein translation. In infected cells, HIV induces changes in the transfer RNA (tRNA) pool, presumably to cope with the differences in codon usage between host and viral genes. These changes in tRNA composition are inhibited by SLFN11, which binds tRNAs through its endoribonuclease domain. Thus, SLFN11 selectively inhibits viral protein translation (Li et al. 2012). The mouse ortholog of human *SLFN11* is thought to be *Sfn9*, but this remains to be functionally demonstrated (Jo and Pommier 2022).

### **II.1.3.4 . Genome replication and degradation**

2'-5' oligoadenylate (OAS) synthetases are a family of proteins encoded by genes clustered on chromosome 12 in humans, and on chromosome 5 in mice (Mashimo et al. 2008; Choi et al. 2015). OAS proteins are activated by the binding of viral dsRNA, and polymerize ATP into 2'-5' adenosine oligomers. These oligomers can activate the ribonuclease RNase L, capable of degrading viral RNA. In human, the OAS family consists of OAS1, OAS2, OAS3 and 2 OAS-like proteins: OASLa and OASLb, which do not possess oligoadenylate synthetase function (Choi et al. 2015). In mice, the *Oas* cluster contains one *Oas2*, one *Oas3*, and 10 *Oas1* genes designated *Oas1a* to *Oas1j* (Mashimo et al. 2008).

Among the genes encoding OAS proteins, *Oas1b* in mice and *OAS1* in humans are important host genetic factors controlling susceptibility to flaviviruses, in particular WNV. *OAS1* polymorphism was associated with increased risk of encephalitis and paralysis following WNV infection in humans (Bigham et al. 2011).

In mice, WNV infection is lethal in classical laboratory strains, since they carry a mutation in *Oas1b* leading to a premature stop codon and a truncated non-functional protein. On the opposite, mice of wild-derived strains have a functional OAS1B protein and survive WNV infection (Mashimo et al. 2002). Interestingly, a recent study showed that the lack of RNase L in mice carrying a functional *Oas1b* allele did not alter their ability to survive WNV infection. This result indicates that RNase L is dispensable for OAS1B antiviral properties (Madden et al. 2019).

### **II.1.3.5 . Viral egress**

There are less known ISGs targeting the late stages of the viral replication cycle. One example is Tetherin (encoded by the *BST2* gene), a transmembrane protein which prevents enveloped viruses from budding. Tetherin possesses two membrane anchors, one of them attached to the plasma membrane, while the second one can be incorporated in the viral envelope, thereby inhibiting the budding and release of new virions (Sauter 2014). Tetherin was shown to inhibit the budding of multiple enveloped viruses such as HIV (Neil et al. 2008), Lassa virus (Sakuma et al. 2009), Marburg virus (Sakuma et al. 2009), vesicular stomatitis virus (Weidner et al. 2010) and herpes simplex virus 1 (HSV-1) (Blondeau et al. 2013).

### **II.1.3.6 . ISGs not affecting the viral cycle**

Other ISGs do not target the viral replication cycle. For instance, Tumor Necrosis Factor-Related Apoptosis Inducing Ligand (TRAIL) is a protein inducing apoptosis. TRAIL is expressed in immune cells following stimulation by cytokines. Notably, IFN-I stimulate TRAIL expression in monocytes, dendritic cells, T lymphocytes and NK cells. Soluble or membrane TRAIL can bind several receptors (death receptors 4 and 5, and decoy receptors 1 and 2), which activate a signaling cascade leading to the activation of caspases and subsequent cell death (Cummins and Badley 2009). Interestingly, mock-infected cells were shown to be resistant to TRAIL-induced apoptosis. Therefore, TRAIL selectively induces cell-death in infected cells (Sato et al. 2001).

ISG15 is an ubiquitin-like protein which binds substrate proteins in a process called ISGylation. Like ubiquitin, ISG15 is covalently linked to target proteins with a system involving 3 enzymes: the activating enzyme E1, the catalyzing enzyme E2 and the ligase E3 (Zhang and Zhang 2011). It was shown recently that the interferon induced RNF213 protein acts as an ISGylation sensor after being ISGylated by ISG15. Although the fate of ISGylated proteins sensed by RNF213 remains unknown, *RNF213* overexpression was shown to lower viral infection levels (Thery et al. 2021).

### **II.1.3.7 . Regulation of IFN-I expression**

Several genes of the IFN-I induction pathways, such as *DDX58*, *IFIH1* and *IRF7*, are ISGs, and are responsible for an auto-amplification of IFN-I production (Schoggins and Rice 2011).

However, excess in IFN-I signaling is deleterious and can lead to autoinflammatory diseases called interferonopathies (Crow and Stetson 2022). Therefore, inhibitory control of IFN-I expression is critical to maintain immune homeostasis. *Socs1* is an ISG encoding a negative regulator of IFN-I signaling. It binds TYK2, resulting in the internalization of IFNAR1, and, therefore, inhibition of the IFN-I signaling (Piganis et al. 2011).

## **II.1.4 . The roles of the genes of the IFN-I induction pathway in viral infections**

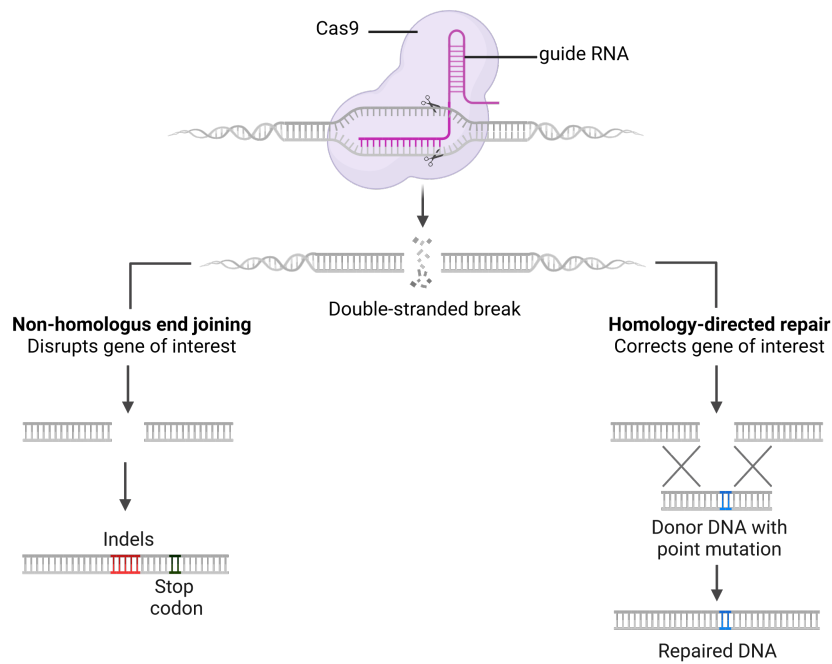
IFN-I are critical actors of the antiviral response through the induction of ISGs (Schoggins 2019). IFN-I production depends on an activation cascade which actors have been studied using several hypothesis-driven approaches, in humans and in mice. In such instances, dysfunction of these genes results in increased susceptibility to viral infection.

### **II.1.4.1 . Reverse genetics studies**

In mice, reverse genetics methods can be used *in vitro* and *in vivo* to study the role of a specific gene. Reverse genetics approaches consist in altering the coding or regulatory sequence of a gene of interest to study the resulting phenotype and gain insight into the function of the gene. Reverse genetics approaches include the induction of gain-of-function mutation (e.g. by insertional transgenesis) or loss-of-function mutation (by deleting part of the coding sequence) also called knockout (KO) mutation (Kherraf et al. 2018).

The first KO methods relied on homologous recombination between the endogenous gene and an homologous DNA sequence injected into mouse embryonic stem cells (Kumar et al. 2009). This technique is tedious and has a low success rate (Kherraf et al. 2018). Generation of KO mice has





**Figure 2:** The CRISPR/Cas9 system for KO and point mutations

After the Cas9 induced a double strand break at the desired site defined by a guide RNA, the non-homologous end joining (NHEJ) mechanism (left) potentially induces indels (red) leading to a frameshift and premature stop codon (black). The homology-directed repair (HDR) mechanism (right) repairs the break using a DNA template with homology arms. Here, the DNA template contained a point mutation (blue) that was inserted in the genome.

become increasingly simple thanks to the CRISPR/Cas9 technique (Hall et al. 2018). This method creates a double strand break at a position of interest by the use of a guide RNA which attracts the Cas9 to the target site. This break can be resolved by a NHEJ event that fixes the break but creates small deletions or insertions. These indels, which are not necessarily multiple of 3 base pairs, can induce a frameshift or a premature stop codon. CRISPR/Cas9 methods have also facilitated the development of models with point mutations, which have been used to study the role of a nucleotide in a gene. An oligonucleotide template with sequence homology with the targeted gene is added to trigger HDR and insert a mutation in the sequence of interest (Figure 2, Inui et al. 2014; Kherraf et al. 2018).

Without modifying the sequence of the genomic DNA, knockdown (KD) methods allow to reduce the amount of mRNA of a specific gene. These methods, which can be used in mouse and human cells, rely on RNA interference with either small interfering RNAs (siRNAs), micro RNA (miRNA) or small hairpin RNAs (shRNAs). These non-coding RNA molecules bind mRNA by their complementary sequence and induce mRNA degradation (Rao et al. 2009; Lam et al. 2015).

Examples of using hypothesis-driven methods in mouse and in human to study the genes of the IFN-I induction pathway are presented in section II.1.4.2.

#### **II.1.4.2 . The role of mouse and human genes of the IFN-I induction pathway in the susceptibility to viral infections**

As an introduction to my experimental work, we published in the journal *Genes and Immunity* a review on the induction pathway of IFN-I and how genes of this pathway contribute to host genetic susceptibility to viral diseases (Bourdon et al. 2020).

This review first details each step of the IFN-I induction from the PAMP recognition by PRRs to the activation of specific transcription factors. Then, it presents the current knowledge on host genetic susceptibilities to viral infections in human and in mice associated with genes of this pathway.

##### **Title**

Host genetic susceptibility to viral infections: the role of type I interferon induction

##### **Abstract**

The innate immune response is the major front line of defense against viral infections. It involves hundreds of genes with antiviral properties which expression is induced by type I interferons (IFNs) and are therefore called interferon stimulated genes (ISGs). Type I IFNs are produced after viral recognition by pathogen recognition receptors, which trigger a cascade of activation events. Human and mouse studies have shown that defective type I IFNs induction may hamper the ability to control viral infections. In humans, moderate to high-effect variants have been identified in individuals with particularly severe complications following viral infection. In mice, functional studies using knock-out alleles have revealed the specific role of most genes of the IFN pathway. Here, we review the role of the molecular partners of the type I IFNs induction pathway and their implication in the control of viral infections and of their complications.



## 10 **Abstract**

11 The innate immune response is the major front line of defense against viral infections. It  
12 involves hundreds of genes with antiviral properties which expression is induced by type I  
13 interferons (IFNs) and are therefore called interferon stimulated genes (ISGs). Type I IFNs  
14 are produced after viral recognition by pathogen recognition receptors which trigger a cascade  
15 of activation events. Human and mouse studies have shown that defective type I IFNs  
16 induction may hamper the ability to control viral infections. In humans, moderate to high-  
17 effect variants have been identified in individuals with particularly severe complications  
18 following viral infection. In mice, functional studies using knock-out alleles have revealed the  
19 specific role of most genes of the IFN pathway. Here, we review the role of the molecular  
20 partners of the type I IFNs induction pathway and their implication in the control of viral  
21 infections and of their complications.

22

## 23 **Introduction**

24 Interferons (IFNs) are cytokines that represent one of the first innate immune barriers  
25 against viruses. They were discovered in 1957 and were named after their capacity to  
26 "interfere" with virus replication. Recognition of non-specific viral molecules such as viral  
27 proteins, DNA and RNA leads to their expression. After recognition of virus components by  
28 pathogen recognition receptors (PRR), an induction cascade leads to the activation of  
29 interferon regulatory factors (IRFs), the transcriptional factors responsible for IFN genes  
30 expression<sup>1</sup>. IFNs are glycoproteins that are secreted into the extracellular medium and act as  
31 autocrine and paracrine factors. The binding to their receptors induces the expression of  
32 interferon stimulated genes (ISGs) with antiviral properties. Non exhaustively, ISGs can  
33 inhibit nuclear import of nucleic acids, synthesis of RNA and proteins, or can enhance virus

34 degradation<sup>1</sup>. Several proteins involved in IFN production and response are inhibited by non-  
35 structural proteins of various viruses which therefore escape host innate defense<sup>2</sup>.

36 IFNs are grouped in three types depending on their sequence, structure and function.  
37 IFN $\gamma$  is the only type II IFN. It is produced by natural killer cells and binds the IFN $\gamma$  receptor  
38 (IFNGR) composed of two subunits (IFNGR1/IFNGR2). This receptor recruits the Janus  
39 kinases 1 (JAK1) and 2 (JAK2), which activate the signal transducer and activator of  
40 transcription 1 (STAT1). STAT1 acts as homodimers<sup>3</sup> and binds gamma-activated sites  
41 present in the target ISGs promoters<sup>1</sup>. Type III IFNs include four IFN lambda numbered  
42 IFN $\lambda$ 1 to IFN $\lambda$ 4. The receptor to type III IFNs is composed of interleukin 28 receptor subunit  
43 alpha (IL-28Ra) and interleukin 10 receptor subunit 2 (IL-10R2). It induces the activation of  
44 the interferon-stimulated gene factor 3 (ISGF3), composed of STAT1, STAT2 and IRF9, that  
45 binds to IFN-stimulated response elements on the promoter of target ISGs<sup>4</sup>. This review  
46 focuses on type I IFNs (IFN-I) which are among the first cytokines produced after viral  
47 infection<sup>3</sup>. IFN-I usually refer to IFN $\alpha$  and IFN $\beta$ , but also include other cell- and species-  
48 specific molecules. All IFN-I signal through the IFN $\alpha$  receptor composed of two subunits  
49 (IFNAR1 and IFNAR2), which recruit JAK1 and non-receptor tyrosine-protein kinase  
50 (TYK2). These kinases activate ISGF3 which binds to IFN-stimulated response elements<sup>3</sup>.

51 Considering the crucial role of IFN-I in host responses to invading viruses, the  
52 inability to induce their expression often leads to severe symptoms. The variable outcome of  
53 viral infections has triggered genetic studies in humans and in mice<sup>5, 6</sup>. Unsurprisingly,  
54 genetic variants or deficiencies in IFN-I induction pathway genes were associated with  
55 susceptibility to diverse viruses. Here we provide a general presentation of the partners of this  
56 pathway, and we review the genetic susceptibilities to viral infections associated with these  
57 genes.

58

## 59 **Molecular mechanisms of type I IFN induction**

### 60 ***Type I interferons***

61 IFN-I is the largest family of IFN proteins. They have a common helical structure  
62 composed of 5  $\alpha$ -helices and are encoded by genes clustered on chromosome 9 in humans and  
63 on chromosome 4 in mice<sup>7</sup>.

64 The two main IFN-I are IFN $\alpha$  and IFN $\beta$ . These proteins are not constitutively  
65 expressed but are up-regulated during viral infection following the activation of the  
66 transcription factors IRF3 and IRF7<sup>8</sup>. Most animal species have multiple IFN $\alpha$  genes, 13  
67 genes with 80% nucleotide identity in human and 14 genes in mice. IFN $\alpha$ s are produced by  
68 plasmacytoid dendritic cells and hematopoietic cells (mostly leucocytes). Each type of IFN $\alpha$   
69 has a different affinity for its receptor, and thus may trigger type-specific responses<sup>9</sup>. IFN $\beta$  is  
70 encoded by a single gene, *IFNB1*, and is also present in most animal species. It is produced by  
71 fibroblasts, dendritic cells and epithelial cells<sup>9</sup>.

72 Other IFNs-I have been described in animal species or in humans. Each of them is  
73 encoded by a single gene<sup>9</sup>. IFN $\epsilon$  is constitutively expressed in the brain, lungs, small intestine  
74 and reproductive tissues. It is regulated by hormones and not during infections. IFN $\kappa$  is  
75 present in a few species including humans and mice. It is constitutively expressed in  
76 keratinocytes and can be up-regulated after exposure to double-stranded RNA (dsRNA).  
77 IFN $\omega$  is present in humans but not in mice. It is expressed mainly in leukocytes. IFN $\zeta$ , also  
78 called limitin, is an IFN-like molecule present only in mice. It is expressed in mature T  
79 lymphocytes, bronchial, epithelial and salivary duct cells. IFN $\tau$  and IFN $\delta$  have been described  
80 but are not expressed in humans or in mice.

81

### 82 ***Virus recognition by PRR***

83 The production of IFN $\beta$  and IFN $\alpha$  is induced by PRRs that recognize molecules  
84 present in pathogens called pathogen-associated molecular patterns. PRRs include Toll-like  
85 receptors (TLRs) and RIG-I-like receptors (RLRs). PRRs recognize components from  
86 bacteria, viruses and fungi and have specific ligands. During viral infections, TLR3  
87 recognizes dsRNA while TLR7 and TLR8 recognize single-stranded RNA (ssRNA) and  
88 TLR9 recognizes DNA molecules. These TLRs are produced in the endoplasmic reticulum  
89 and sense their ligands in endosomes after virus entry into host cells<sup>10</sup>. TLR2 and TLR4 are  
90 present at the cell surface and recognize viral proteins<sup>11, 12</sup>. Replication of viruses with  
91 positive ssRNA genome produces dsRNA which is recognized by TLR3 and RLRs<sup>10</sup>.

92 Three RLRs recognize viral RNA. While DDX58 (also known as RIG-I) senses 5'-  
93 phosphorylated RNA, IFIH1 (MDA5) recognizes long dsRNA. DHX58 (LGP2) facilitates  
94 viral RNA recognition by DDX58 and IFIH1<sup>13</sup> and enhances RLR-dependent IFN induction<sup>14</sup>.

95 Viral recognition by TLRs and RLRs triggers a cascade of molecular activations  
96 which results in the production of IFN-I. This pathway is summarized in Figure 1.

97

### 98 ***TLR pathway***

99 TLR3 recognition of viral RNA induces its own phosphorylation which allows the  
100 recruitment of the adaptor protein Toll-interleukin receptor (TIR) domain-containing adapter  
101 molecule 1 (TICAM1, also called TRIF)<sup>15</sup>. Interaction between TLR3 and TICAM1 is  
102 enabled by the phosphorylation of two TLR3 tyrosine residues<sup>15</sup>. TLR4 also can induce the  
103 expression of IFN-I by recognizing viral proteins present in the extracellular medium and  
104 signalling through the adaptors myeloid differentiation primary response protein (MYD88)  
105 and myelin and lymphocyte protein (MAL). Once activated, TLR4 is endocytosed and  
106 recruits TICAM1 and TIR domain-containing adapter molecule 2 (TICAM2, also called  
107 TRAM) in the endosomes<sup>16</sup>. TICAM1 recruits the TNF receptor associated factor 3

108 (TRAF3)<sup>15</sup>, which then activates the kinases responsible for the activation of the IRFs.  
109 TICAM1 is targeted by the viral 3C protease of hepatitis A virus and coxsackievirus B3  
110 (CVB3) which allows these viruses to escape the host immune response<sup>10</sup>.

111 TLR2 also activates the expression of IFN-I, but the mechanisms are incompletely  
112 understood. Signalling by TLR2 requires MAL, TICAM2 and MYD88 which, once activated,  
113 relocate to the endosomes and induce a signalling cascade resulting in IRF7 activation and  
114 IFN-I expression. Therefore TLR2 and TLR4 likely use similar mechanisms to induce IFN-I  
115 production<sup>12</sup>.

116 TLR7, TLR8 and TLR9 also induce IFN-I expression, but only in plasmacytoid  
117 dendritic cells which are known to produce high levels of IFN after viral infection. These  
118 TLRs use the MYD88 adaptor which, in plasmacytoid dendritic cells, forms a complex with  
119 IRF7. This complex allows the phosphorylation and activation of IRF7 by interleukin 1  
120 receptor associated kinase 1 (IRAK1) and triggers the expression of IFNs<sup>17</sup>.

121

## 122 ***RLR pathway***

123 Viral RNA binding on RLRs DDX58 and IFIH1 induces a conformational change of  
124 these receptors which exposes their caspase activation and recruitment domains (CARD).  
125 These domains interact with the CARD of the mitochondrial antiviral signalling protein  
126 (MAVS, also called IPS-1). Subsequently, DDX58 and IFIH1 promote the formation of  
127 prion-like MAVS aggregates, which induce TRAF3 recruitment<sup>18</sup>. Several proteins of the  
128 RLR pathway are targeted by viruses. Influenza A virus (IAV) NS1 protein and respiratory  
129 syncytial virus NS1 protein bind DDX58 and MAVS, respectively, and block their signalling.  
130 IFIH1 is degraded following poliovirus infection, and encephalomyocarditis virus (EMCV)  
131 3C protease can degrade DDX58<sup>10</sup>.

132



### 133 **Activation of IRFs**

134 TRAF3 recruits two kinases, TANK binding kinase 1 (TBK1) and inhibitor of nuclear  
135 factor kappa-B kinase subunit epsilon (IKKε), to phosphorylate and activate IRF3 and IRF7.  
136 Once phosphorylated, IRF3 and IRF7 form homodimers or heterodimers, translocate to the  
137 nucleus and promote IFN-I transcription<sup>15, 19</sup>. Viral proteins also target these factors. Ebola  
138 virus VP35 protein binds and blocks TBK1 and IKKε. The hepatitis C virus NS3/4A protease  
139 degrades IRF3 while viral homologues of IRFs, such as Kaposi's sarcoma-associated  
140 herpesvirus vIRFs, bind host IRFs and inhibit IFN-I transcription<sup>10</sup>.

141 IFNβ expression is regulated by four positive regulatory domains (PRD). NFκB and  
142 AP1 bind PRDII and PRDIV, respectively, and promote basal expression of *IFNB1*. After  
143 viral infection, IRF3 and IRF7 are activated and bind PRDI and PRDIII to induce *IFNB1*  
144 overexpression<sup>20</sup>. IFNα genes have only PRDI- and PRDIII-like elements and their  
145 expression is therefore controlled exclusively by IRF3 and IRF7. IRF3 has more affinity for  
146 *IFNB1* while IRF7 has more affinity for IFNα genes. IRF3 is constitutively abundant but  
147 inactive while *IRF7* is an ISG present at low levels before infection and up-regulated by IFN-I  
148 signalling. Therefore, in the early phase after infection, IFN-I expression is induced by IRF3,  
149 resulting in predominant IFNβ production. IFNβ signalling induces *IRF7* expression resulting  
150 in IFNα production in a later phase<sup>8</sup>. IRF1 and IRF5 can also induce IFN-I expression,  
151 however both are dispensable and their role remains unclear<sup>8</sup>. Furthermore, TLRs and RLRs  
152 also activate the NFκB pathway after infection through TICAM1, MYD88 and MAVS to  
153 induce the production of inflammatory cytokines<sup>15</sup>.

154

### 155 **Genetic susceptibility to viral infections**

156 Several of the genes described above have been associated with susceptibilities to viral  
157 infections. These studies are summarized in Table 1 and Table 2 for human and mouse genes,

158 respectively. Human studies split into case studies and association studies. Case studies aim to  
159 identify mutations which strongly impact the severity of viral infection but are rare in the  
160 population. Association studies seek common genetic variants generally associated with a  
161 moderate impact. Genome-wide association studies require the analysis of large cohorts  
162 which can rarely be assembled in infectious diseases. However, statistical power is increased  
163 by limiting the variants tested to a reduced set of candidate genes. This approach has led  
164 Zhang *et al.* to identify association between variants at 13 loci governing TLR3- and IRF7-  
165 dependent IFN-I immunity and the severity of COVID-19 by comparing 659 patients with  
166 life-threatening pneumonia and 534 patients with mild or no symptoms<sup>21</sup>. Likewise, Bigham  
167 *et al.* investigated 86 genes regulating immune function and identified association between  
168 three of them and the severity of West Nile virus (WNV) infection<sup>22</sup>.

169 In mice, forward and reverse genetics are used to analyze resistance to viral  
170 infections<sup>23</sup>. Reverse genetics aims at characterizing the function of a given gene by altering  
171 its sequence. Many studies have reported modified susceptibility to viral infections in mice  
172 carrying loss-of-function mutations (gene knock-outs, KO) in IFN-I pathway. Forward  
173 genetics starts with a difference of susceptibility between two strains and aims at identifying  
174 the causal genetic variants. Differences may result from random chemical mutagenesis<sup>24</sup> or  
175 from natural variants between genetically diverse mouse strains such as the Collaborative  
176 Cross<sup>6</sup>. Interestingly, studies performed on the same virus can be compared to assess the  
177 specific or overlapping roles of the genes of the IFN-I cascade in the severity of a viral  
178 infection and in its complications.

179

### 180 ***TLR-TICAM1 pathway***

181 As *TLR3* is the primary TLR involved in IFN-I expression after virus recognition, the  
182 effects of its variants on the susceptibility to viral infections were extensively studied. In

183 humans, association studies and case studies identified *TLR3* variants linked to increased  
184 susceptibility to IAV<sup>25-28</sup>, hepatitis B virus<sup>29</sup>, herpes simplex virus 1 (HSV-1)<sup>30-32</sup>, measles  
185 virus<sup>33</sup> and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)<sup>21</sup>. Two SNPs  
186 associated with susceptibility to IAV are in intronic regions upstream exon 4<sup>25, 26</sup>. Since this  
187 exon contains the signal induction transmembrane protein domain, these SNPs might alter  
188 TLR3 signalling. Other variants are in the luminal leucine-rich repeats of TLR3<sup>21, 27, 28, 30, 32,</sup>  
189 <sup>33</sup>. This region forms a solenoid critical for RNA binding<sup>32</sup> and virus recognition. Lastly,  
190 mutations were identified in the TIR domain of TLR3<sup>21, 30, 31</sup>. In particular, a non-sense  
191 mutation was identified in a case of herpes simplex encephalitis (HSE), a complication of  
192 HSV-1 infection. This mutation removes the TIR domain which is required for the  
193 recruitment of TICAM1 and downstream signalling<sup>30</sup>.

194 In contrast, a common variant in *TLR3* was associated with increased resistance to  
195 human immunodeficiency virus (HIV). The L412F allele, present in approximately 30% of  
196 Europeans and over-represented in a cohort of HIV-exposed seronegative individuals, leads to  
197 reduced viral replication and overexpression of inflammatory cytokines *in vitro*<sup>34</sup>, likely by  
198 increasing TLR3 signalling. However, the same variant showed positive association with  
199 subacute sclerosing panencephalitis, a severe complication of measles virus infection<sup>33</sup>. The  
200 increased inflammatory response due to this mutation may be advantageous in the case of  
201 HIV infection, but deleterious in the case of measles virus infection. Variants in the TLR  
202 adaptor *TICAM1* can also alter susceptibility to viruses. In humans, four mutations were  
203 identified in patients suffering from HSE and three in patients with life-threatening COVID-  
204 19, all of them leading to decreased IFN-I expression<sup>21, 35, 36</sup>.

205 The TLR-TICAM1 pathway has been also extensively studied in mouse viral  
206 infections. Compared to wild-type (WT) mice, *Tlr3*-deficient mice showed a decreased  
207 survival rate with higher viral loads in coxsackievirus B3<sup>37</sup> and EMCV<sup>38</sup> infections and, while

208 they showed an increased serum viral load but unchanged mortality after murine  
209 cytomegalovirus (MCMV) infection<sup>39</sup>. In contrast, *Tlr3*-deficient mice displayed a decreased  
210 mortality following IAV infection<sup>40</sup>. Mice carrying a frameshift-induced deletion in the  
211 *Ticam1* gene showed enhanced susceptibility to MCMV with increased viral load in the  
212 spleen and higher mortality<sup>41</sup>. Another study found that *Ticam1*-deficient mice were more  
213 susceptible to CVB3. Interestingly, these mice presented a decreased IFN-I expression 72  
214 hours post-infection, but an increased expression 7 days after infection<sup>42</sup>, which may result  
215 from an uncontrolled inflammatory response. *Tlr3*-deficient mice also developed cardiac  
216 anomalies, a complication of CVB3 infection, with large myocarditic lesions and increased  
217 heart viral load<sup>37</sup>. Similarly, *Ticam1*-deficient mice presented left ventricular dysfunction and  
218 severe myocardial damage including cardiac fibrosis. These mice also showed increased heart  
219 viral load<sup>42</sup>. The overlapping phenotypes observed in these two studies are consistent with the  
220 direct interactions between *Tlr3* and *Ticam1* in the IFN-I induction cascade.

221 *Tlr3*-deficient mice were also less susceptible to vaccinia virus (VV) infection than  
222 WT mice with higher viral load, while *Ticam1*- and *Tlr4*-deficient mice were more  
223 susceptible<sup>43, 44</sup>. It was hypothesized that abrogating *Tlr3* signaling decreases the  
224 inflammatory response and thus the complications resulting from VV infection. In contrast,  
225 since *Tlr4* signaling activates IRFs and NFκB, *Tlr4* and *Ticam1* KOs block both pathways  
226 and lead to increased susceptibility to VV infection<sup>43, 44</sup>.

227 The outcome of WNV infection in *Tlr3*-deficient mice was investigated in two studies  
228 which used the same mouse strain and two closely related virus strains with contrasted results.  
229 Wang *et al.* reported that *Tlr3*-deficient mice presented a decreased mortality after infection  
230 with WNV isolate 2741, but an increased viral load. Moreover, these mice showed decreased  
231 neuronal inflammation and blood-brain barrier permeability, suggesting that *Tlr3* is involved  
232 in the virus brain entry<sup>45</sup>. Daffis *et al.* who used the WNV strain 3000.0259 reported that

233 *Tlr3*-deficient mice also presented a higher brain viral load but with susceptibility to WNV  
234 infection and mortality than WT mice. Unlike the previous study, blood-brain barrier  
235 permeability and neuroinflammation were not affected, compared with WT mice<sup>46</sup>. These  
236 contrasted results were attributed to the infection route, the viral dose and the cells used to  
237 produce the virus which differed between the two studies<sup>46</sup>. Interestingly, *Tlr3*-deficient mice  
238 produced normal amount of IFN-I in the first study, while they were decreased in the second  
239 study, leading to the hypothesis that IFN-I expression could also have a detrimental effect in  
240 WNV infection<sup>45</sup>. These results illustrate the dual role of *Tlr3* signalling which may lead to  
241 an excessive inflammatory response, while decreased inflammation in *Tlr3* KO mice may  
242 reduce the risk of severe complication.

243

#### 244 ***TLR-MYD88 pathway***

245 In mice, deficiency in *Tlr2*, *Tlr4*, *Tlr7* or *Tlr9* was associated with increased or  
246 decreased susceptibility to viral infections. However, since these receptors signal through the  
247 MYD88 adaptor which also activates the NFκB pathway, their role in the susceptibility to  
248 viruses may not be solely associated with the IFN-I pathway.

249 *Tlr7*-deficient mice were more susceptible to WNV infection<sup>47</sup>. However, they  
250 presented an increased IFN-I expression which could result from the signaling through other  
251 receptors such as *Tlr3* and RLRs. *Tlr9* and *Myd88*-deficient mice were more susceptible to  
252 MCMV with decreased IFN-I production<sup>39</sup>. *Myd88* KO mice had a reduced number of splenic  
253 plasmacytoid dendritic cells which could explain reduced levels of IFN-I. By contrast, *Tlr2*  
254 KO mice were less susceptible to HSV-1 with reduced mortality compared to WT mice. They  
255 also showed decreased NFκB-induced cytokine production which may explain a milder  
256 inflammatory state and the absence of severe complications<sup>48</sup>. IFN-I expression was not  
257 investigated although it could contribute to the pathology.

258 *Myd88* deficiency in mice also resulted in increased susceptibility to chikungunya  
259 virus (CHIKV) and severe acute respiratory syndrome coronavirus (SARS-CoV). Indeed,  
260 *Myd88* KO mice presented higher viral loads following CHIKV infection<sup>49</sup>, and higher  
261 mortality and increased viral load following SARS-CoV infection<sup>50</sup>. Two studies with WNV  
262 led to similar results<sup>47, 51</sup>. Interestingly, in one study, *Myd88* KO mice had higher levels of  
263 IFN-I after infection than WT mice due to an increased expression in bone marrow-derived  
264 macrophages. This might result from high viral replication in these cells and from the  
265 signaling of other pathways, such as RLR- or *Tlr3*-dependent pathways<sup>51</sup>. The role of *Ticam2*  
266 in the susceptibility to SARS-CoV was suspected in an association study using the  
267 Collaborative Cross and was confirmed with a *Ticam2*-deficient strain which showed higher  
268 lung viral loads than WT mice<sup>52</sup>.

269

#### 270 ***RLR pathway***

271 In humans, two variants were identified in *DDX58* in a patient who suffered from  
272 severe IAV infection. The R71H variant is in the CARD protein domain, while the P885  
273 variant is in the regulatory domain involved in viral RNA recognition. These variants lead to  
274 impaired IFN-I expression following IAV infection when expressed in *DDX58* deficient  
275 human embryonic kidney cells 293, but not in the patient's peripheral blood mononuclear  
276 cells where other pathways, such as *TLR7*-dependant signalling might ensure a correct  
277 expression<sup>53</sup>. Variants were identified in *IFIH1* in patients suffering from bronchiolitis  
278 following rhinovirus or respiratory syncytial virus infection and led to decreased expression  
279 of IFN $\beta$ <sup>54</sup>. Moreover, the K365E mutation was identified in a 5-year-old child suffering from  
280 numerous recurrent respiratory virus infections. This mutation prevents *IFIH1* from  
281 interacting with viral RNA, thus inhibiting IFN-I induction<sup>55</sup>. The H843A mutation in *IFIH1*  
282 was also associated with susceptibility to HCV by comparing patients with spontaneously

283 resolved hepatitis or chronic hepatitis<sup>56</sup>. In two studies, *Ddx58*-deficient mice showed similar  
284 mortality after IAV infection compared with WT individuals<sup>57, 58</sup>. Notably, one study showed  
285 that *Ddx58* deficiency also led to defects in adaptive immunity affecting antigen presentation  
286 by dendritic cells and activation of T cell responses<sup>57</sup>. *Ifih1*-deficient mice were more  
287 susceptible to mouse hepatitis virus. Interestingly, they showed decreased expression of IFN-I  
288 but normal induction of ISGs<sup>59</sup>. Moreover, they were found to be more susceptible to human  
289 metapneumovirus<sup>60</sup>, to hepatitis B virus<sup>61</sup>, to murine norovirus 1<sup>62</sup>, and more prone to develop  
290 demyelinating disease following Theiler's murine encephalomyelitis virus infection<sup>63</sup>.

291 In mice, RLRs and MAVS have often been studied together, which has unraveled their  
292 specificity. Using *Ddx58*- and *Ifih1*-deficient mouse embryonic fibroblasts (MEFs) infected  
293 with several viruses, Kato *et al.* found that these two receptors recognize different viruses.  
294 Moreover, they showed that *Ddx58*- and *Ifih1*-deficient mice were more susceptible to  
295 Japanese encephalitis virus than WT mice, and that *Ifih1*- but not *Ddx58*-deficient mice were  
296 more susceptible to EMCV<sup>64</sup>. Susceptibility of *Ifih1*-deficient mice to EMCV infection was  
297 also reported in another study<sup>65</sup>. Furthermore, *Ddx58*-deficient mice showed increased serum  
298 viral load following CHIKV infection, which was not the case for *Ifih1*-deficient mice<sup>49</sup>.  
299 These results show that *Ddx58* and *Ifih1* have complementary roles in the recognition of viral  
300 RNA, consistently with their known differences in molecular pattern recognition.

301 By contrast, both *Ddx58*- and *Ifih1*-deficient mice showed increased susceptibility to  
302 WNV. Double-deficient mice were even more susceptible and invariably died within 8 days  
303 after infection, showing that both receptors are involved in the recognition of WNV. The  
304 phenotype of double-deficient mice was very similar to that of *Mavs*-deficient mice through  
305 which both RLRs signal<sup>66</sup>. Indeed, *Mavs* deficiency resulted in increased susceptibility to  
306 WNV with higher mortality and viral load, and deficient activation of IFN $\beta$ <sup>67</sup>. *Mavs* and *Ifih1*  
307 deficiencies resulted also in increased mortality following CVB3 infection and decreased

308 expression of IFN-I although viral titers were identical to WT mice<sup>68</sup>. *Mavs* deficiency also  
309 resulted in increased susceptibility to EMCV<sup>69</sup>, to vesicular stomatitis virus<sup>69, 70</sup> and to dengue  
310 virus (DENV)<sup>71</sup>, and in increased serum viral load following CHIKV infection<sup>49</sup>.

311 Mice deficient for the auxiliary RLR *Dhx58* gene also showed increased susceptibility  
312 to EMCV<sup>13</sup> and to WNV<sup>14</sup>. WNV-infected, *Dhx58*-deficient mice showed increased mortality  
313 but similar kinetics of IFN $\beta$  production and tissue viral loads compared with WT mice. In the  
314 brain they displayed increased neuronal damage, elevated viral load in a late phase of  
315 infection, low neuroinflammation and decreased recruitment of CD8+ T cells<sup>14</sup>. *Dhx58* is  
316 therefore required for protection against WNV infection. Furthermore, mice overexpressing  
317 *Dhx58* were more resistant to IAV infection<sup>72</sup>. *In vitro*, *Dhx58*-deficient cells exposed to  
318 several RNA viruses produced less IFN-I than WT cells suggesting that *Dhx58* is required for  
319 *Dhx58*- and *Ifih1*-mediated antiviral responses<sup>13</sup>.

320

### 321 ***Activation of IRFs and IFNs***

322 Sequencing of candidate genes in HSE patients identified a heterozygous missense  
323 mutation in *TRAF3* associated with decreased IFN-I expression<sup>73</sup> and two missense mutations  
324 in *TBK1* affecting the kinase domain and thus preventing the phosphorylation of target  
325 proteins<sup>74</sup>. These mutations resulted in reduced IFN-I expression in cells stimulated with  
326 synthetic RNA. Two dominant mutations in *TBK1* were identified in patients with severe  
327 COVID-19 and led to decreased IFN-I expression in HEK293T cells transfected with these  
328 mutant forms of *TBK1*<sup>21</sup>.

329 In mice, no genetic variants in *Traf3* or in *Tbk1* have been associated with altered  
330 susceptibility to viral infections. However, mice deficient for the *Ikk $\epsilon$*  gene (encoding IKK $\epsilon$ ,  
331 a kinase involved in IRFs activation) showed extreme susceptibility to IAV despite normal



332 expression of IFN-I. Mechanistic studies revealed that IKKε indirectly controls the expression  
333 of a subset of ISGs<sup>75</sup>.

334 Mutations in IRFs have been repeatedly associated with susceptibility to viral  
335 infections in humans and in mice. In humans, two missense mutations were found in *IRF3* in  
336 patients suffering from HSE<sup>35, 76</sup>. These mutations are located in the IRF association domain  
337 and might therefore prevent IRF3 dimerization<sup>77</sup>. Two autosomal dominant mutations in *IRF3*  
338 were identified in COVID-19 patients with pneumonia<sup>21</sup> and a non-coding variant was  
339 associated with susceptibility to WNV by comparing asymptomatic and symptomatic infected  
340 individuals<sup>22</sup>. A compound heterozygosity was found in *IRF7* in a patient suffering from life-  
341 threatening infection following IAV infection<sup>78</sup> and seven mutations in *IRF7* were found in  
342 COVID-19 patients leading to decreased IFN-I induction<sup>21</sup>.

343 In mice, *Irf3* and *Irf7* deficiencies have been studied in isolation or in combination.  
344 Both single deficiencies increased mortality following IAV infection and susceptibility was  
345 further enhanced in double deficient mice<sup>79, 80</sup>. Viral load in lungs was not significantly  
346 altered in *Irf7*-deficient mice but was increased in *Irf3*-deficient mice and even more in  
347 double-deficient mice, suggesting that *Irf7* also contributes to controlling viral replication. On  
348 day 2 after infection, IFNα expression was reduced in *Irf3*- but not in *Irf7*-deficient mice  
349 while IFNβ expression was reduced in *Irf7* but not in *Irf3*-deficient mice<sup>79</sup>. This result is  
350 consistent with the distinct affinities of the two IRFs for the IFN-I genes. Mice deficient for  
351 *Irf3* and *Irf7* were also susceptible to WNV with increased mortality rate and viral load,  
352 decreased expression of IFN-I and increased viral load in the brain<sup>81, 82</sup>.

353 However, the consequences of *Irf3* and *Irf7* deficiencies are variable between viruses.  
354 *Irf7*- but not *Irf3*-deficient mice were susceptible to HSV-1 and they were more susceptible to  
355 EMCV than *Irf3*-deficient mice<sup>83</sup>. In the case of CHIKV infection, *Irf3-Irf7* double KO mice  
356 were highly susceptible with increased viremia and mortality, while *Irf3*- and *Irf7*-deficient

357 mice survived and had normal viremia<sup>84</sup>. Following DENV infection, *Irf3-Irf7* double KO  
358 mice and *Irf7*-deficient mice showed increased viral load and decreased IFN-I expression but  
359 survived the infection<sup>85</sup>. In the case of Zika virus infection, *Irf3-Irf5-Irf7* triple KO mice died  
360 with neurological disease signs, while *Irf3*-deficient mice survived<sup>86</sup>. Interestingly, *Irf3-Irf5-*  
361 *Irf7* triple KO mice survived to DENV infection through robust induction of type II IFNs, but  
362 showed increased viremia. This resistance to DENV was abolished when *Irf1* was also  
363 inactivated, which led to the identification of a protective *Irf1*-dependent pathway<sup>87</sup>.

364 Lastly, variants in IFN-I genes themselves were associated with susceptibility to viral  
365 diseases in mice. The role of IFN-I in viral infections has been extensively investigated using  
366 mice deficient for their receptor. *Ifnar1*-deficient mice showed increased susceptibility to a  
367 number of viruses including CHIKV<sup>49</sup>, Zika virus<sup>86</sup>, DENV<sup>71</sup> and Ebola virus<sup>88</sup>. Transient  
368 blockade of IFN $\alpha$  and IFN $\beta$  with monoclonal antibodies resulted in increased mortality after  
369 WNV infection<sup>89</sup>. WNV susceptibility was also studied in *Ifnb1*<sup>-/-</sup> mice and led to similar  
370 results<sup>90</sup>. IFN $\beta$ -deficient mice were also found more susceptible to VV than WT mice<sup>91</sup>.  
371 These two studies led to opposite results regarding IFN $\alpha$  expression. The absence of IFN $\beta$  is  
372 expected to abrogate *Irf7* induction and thus to decrease IFN $\alpha$  expression. This was indeed  
373 observed after VV infection. However, after WNV infection, IFN $\alpha$  was upregulated, which  
374 was hypothesized to result from the high viral load<sup>90</sup>. IFN $\beta$ -deficient mice also showed  
375 increased susceptibility to IAV<sup>92</sup> and CVB3<sup>93</sup>, and increased spleen viral load following  
376 Friend virus infection<sup>94</sup>.

377

## 378 **Discussion**

379 Type I IFNs are critical components of the immediate response against invading  
380 viruses. Indeed, their induction allows the expression of many ISGs which can control viral  
381 infection. The pathway leading to IFN-I production is complex as many genes are involved,

382 and viral proteins target this pathway at multiple levels. Moreover, some of these genes, such  
383 as *IRF7*, are also ISGs, which further complicates the kinetics of IFN-I activation. Despite the  
384 vast number of studies carried out on the induction of IFN-I, not all mechanisms are yet fully  
385 understood.

386 In accordance with the functions of IFN-I, mutations in most genes of the induction  
387 pathway have been associated with increased susceptibility to viral infections in human and  
388 mice. In humans, whole exome or candidate gene sequencing has identified coding and non-  
389 coding variants, primarily in patients with severe forms of infections. It is likely that other  
390 variants are present in the human population but the power to detect them in association  
391 studies depends on their frequency, on their impact on host response to infections and on  
392 cohort size. In mice, most studies have used reverse genetics approaches and have  
393 investigated the consequences of complete loss-of-function mutations in infected mice which,  
394 in most cases, led to higher susceptibility, with mortality and elevated viral load in tissues.  
395 While all mutants reported here were constitutively deficient, tissue-specific conditional  
396 alleles allow investigating the pathway in specific cell lineages. For example, myeloid-  
397 conditional *Tbk1*-deficient mice showed increased survival to IAV infection with reduced  
398 inflammation in the respiratory tract, demonstrating the role of myeloid cells in disease  
399 pathophysiology<sup>95</sup>.

400 Notably, a few genes of the pathways were not tested by reverse genetics. *Tlr2* and  
401 *Tlr4* have been investigated mostly for their role in bacterial infections<sup>11</sup>. For other genes, like  
402 *Tbk1* and *Traf3*, deficiency was only studied *in vitro* on MEFs or macrophages since  
403 homozygous mice die either *in utero* (*Tbk1*) or a few days after birth (*Traf3*)<sup>96</sup>. Interestingly,  
404 Marchlik *et al.* produced a *Tbk1* mutation which resulted in a catalytically inactive protein  
405 and they could obtain homozygous deficient mice with complete ablation of IFN $\beta$   
406 production<sup>97</sup>. This difference in survival of *Tbk1*-deficient mice is likely due to the 129S5

407 genetic background on which this allele was created, compared with the C57BL/6 background  
408 used for most KO alleles. This case highlights the importance of mouse genetic background  
409 when evaluating the phenotype resulting from gene inactivation<sup>98</sup>.

410 The formal description of the IFN-I activation cascade incompletely reflects the  
411 complexity of the mechanisms from viral components recognition to IFN-I-induced effectors.  
412 As exemplified by *Irf3* and *Irf7*, the effect of a host gene variant may be different between  
413 viruses. Moreover, while deficiency of most pathway genes resulted in increased  
414 susceptibility to viral infections due to impaired IFN $\beta$  production, it could also be associated  
415 with reduced susceptibility as in the case of *Tlr3*. This observation underlines the complexity  
416 of immune mechanisms, and the importance of balanced and well-controlled IFN response.  
417 While rapid activation of ISGs is critical to the control of viral replication, excessive or  
418 persistent IFN-I production can be detrimental by triggering inflammatory processes  
419 responsible for tissue damage and organ failure. Notably, dysregulation of immune responses  
420 with delayed expression of IFN-I and robust cytokine response could be at the origin of the  
421 clinical manifestations observed in severe SARS-CoV<sup>99</sup> and SARS-CoV-2 infections<sup>100</sup>.

422 Investigating the role of every gene of the IFN-I induction cascade by gene inactivation has  
423 contributed to dissecting the mechanisms of the pathway. However, a non-functional step may  
424 result from defective interactions between functional but incompatible partner proteins. Such  
425 interactions could occur for example if the two partners were inherited from genetically  
426 distant parents. In mice, investigating strains produced by crosses between founders of  
427 different subspecific origins, like the Collaborative Cross, may identify such situations and  
428 provide new variants for functional analysis<sup>6</sup>. With the growing evidence that microbiota can  
429 also modify the IFN-I response and therefore the susceptibility to infectious diseases<sup>101</sup>, it is  
430 clear that we are still far from understanding the subtle regulations of an essential pathway.

431

432 **Conflict of interest**

433 The authors declare no conflict of interest.

434

435

436 **References**

437

438 1. Fensterl V, Sen GC. Interferons and viral infections. *Biofactors* 2009; **35**(1): 14-20.

439

440 2. Beachboard DC, Horner SM. Innate immune evasion strategies of DNA and RNA  
441 viruses. *Curr Opin Microbiol* 2016; **32**: 113-119.

442

443 3. Lee AJ, Ashkar AA. The Dual Nature of Type I and Type II Interferons. *Front*  
444 *Immunol* 2018; **9**: 2061.

445

446 4. Zhou JH, Wang YN, Chang QY, Ma P, Hu Y, Cao X. Type III Interferons in Viral  
447 Infection and Antiviral Immunity. *Cell Physiol Biochem* 2018; **51**(1): 173-185.

448

449 5. Kenney AD, Dowdle JA, Bozzacco L, McMichael TM, St Gelais C, Panfil AR *et al.*  
450 Human Genetic Determinants of Viral Diseases. *Annu Rev Genet* 2017; **51**: 241-263.

451

452 6. Noll KE, Ferris MT, Heise MT. The Collaborative Cross: A Systems Genetics  
453 Resource for Studying Host-Pathogen Interactions. *Cell Host Microbe* 2019; **25**(4):  
454 484-498.

455

456 7. Hardy MP, Owczarek CM, Jermini LS, Ejdeback M, Hertzog PJ. Characterization of  
457 the type I interferon locus and identification of novel genes. *Genomics* 2004; **84**(2):  
458 331-45.

459

- 460 8. Honda K, Takaoka A, Taniguchi T. Type I interferon [corrected] gene induction by  
461 the interferon regulatory factor family of transcription factors. *Immunity* 2006; **25**(3):  
462 349-60.
- 463
- 464 9. Li SF, Gong MJ, Zhao FR, Shao JJ, Xie YL, Zhang YG *et al.* Type I Interferons:  
465 Distinct Biological Activities and Current Applications for Viral Infection. *Cell*  
466 *Physiol Biochem* 2018; **51**(5): 2377-2396.
- 467
- 468 10. Nan Y, Nan G, Zhang YJ. Interferon induction by RNA viruses and antagonism by  
469 viral pathogens. *Viruses* 2014; **6**(12): 4999-5027.
- 470
- 471 11. Lester SN, Li K. Toll-like receptors in antiviral innate immunity. *J Mol Biol* 2014;  
472 **426**(6): 1246-64.
- 473
- 474 12. Stack J, Doyle SL, Connolly DJ, Reinert LS, O'Keeffe KM, McLoughlin RM *et al.*  
475 TRAM is required for TLR2 endosomal signaling to type I IFN induction. *J Immunol*  
476 2014; **193**(12): 6090-102.
- 477
- 478 13. Satoh T, Kato H, Kumagai Y, Yoneyama M, Sato S, Matsushita K *et al.* LGP2 is a  
479 positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proc Natl Acad*  
480 *Sci U S A* 2010; **107**(4): 1512-7.
- 481
- 482 14. Suthar MS, Ramos HJ, Brassil MM, Netland J, Chappell CP, Blahnik G *et al.* The  
483 RIG-I-like receptor LGP2 controls CD8(+) T cell survival and fitness. *Immunity* 2012;  
484 **37**(2): 235-48.

485

486 15. Kawasaki T, Kawai T. Toll-like receptor signaling pathways. *Front Immunol* 2014; **5**:  
487 461.

488

489 16. Funami K, Matsumoto M, Oshiumi H, Inagaki F, Seya T. Functional interfaces  
490 between TICAM-2/TRAM and TICAM-1/TRIF in TLR4 signaling. *Biochem Soc*  
491 *Trans* 2017; **45**(4): 929-935.

492

493 17. Kawai T, Akira S. Innate immune recognition of viral infection. *Nat Immunol* 2006;  
494 **7**(2): 131-7.

495

496 18. Hou F, Sun L, Zheng H, Skaug B, Jiang QX, Chen ZJ. MAVS forms functional prion-  
497 like aggregates to activate and propagate antiviral innate immune response. *Cell* 2011;  
498 **146**(3): 448-61.

499

500 19. Ning S, Pagano JS, Barber GN. IRF7: activation, regulation, modification and  
501 function. *Genes Immun* 2011; **12**(6): 399-414.

502

503 20. Levy DE, Marie IJ, Durbin JE. Induction and function of type I and III interferon in  
504 response to viral infection. *Curr Opin Virol* 2011; **1**(6): 476-86.

505

506 21. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J *et al.* Inborn errors of  
507 type I IFN immunity in patients with life-threatening COVID-19. *Science* 2020;  
508 **370**(6515).

509



- 510 22. Bigam AW, Buckingham KJ, Husain S, Emond MJ, Bofferding KM, Gildersleeve H  
511 *et al.* Host genetic risk factors for West Nile virus infection and disease progression.  
512 *PLoS One* 2011; **6(9)**: e24745.  
513
- 514 23. Beutler B, Eidenschenk C, Crozat K, Imler JL, Takeuchi O, Hoffmann JA *et al.*  
515 Genetic analysis of resistance to viral infection. *Nat Rev Immunol* 2007; **7(10)**: 753-  
516 66.  
517
- 518 24. Crozat K, Georgel P, Rutschmann S, Mann N, Du X, Hoebe K *et al.* Analysis of the  
519 MCMV resistome by ENU mutagenesis. *Mamm Genome* 2006; **17(5)**: 398-406.  
520
- 521 25. Esposito S, Molteni CG, Giliani S, Mazza C, Scala A, Tagliaferri L *et al.* Toll-like  
522 receptor 3 gene polymorphisms and severity of pandemic A/H1N1/2009 influenza in  
523 otherwise healthy children. *Virol J* 2012; **9**: 270.  
524
- 525 26. Lee N, Cao B, Ke C, Lu H, Hu Y, Tam CHT *et al.* IFITM3, TLR3, and CD55 Gene  
526 SNPs and Cumulative Genetic Risks for Severe Outcomes in Chinese Patients With  
527 H7N9/H1N1pdm09 Influenza. *J Infect Dis* 2017; **216(1)**: 97-104.  
528
- 529 27. Hidaka F, Matsuo S, Muta T, Takeshige K, Mizukami T, Nunoi H. A missense  
530 mutation of the Toll-like receptor 3 gene in a patient with influenza-associated  
531 encephalopathy. *Clin Immunol* 2006; **119(2)**: 188-94.  
532

- 533 28. Lim HK, Huang SXL, Chen J, Kerner G, Gilliaux O, Bastard P *et al.* Severe influenza  
534 pneumonitis in children with inherited TLR3 deficiency. *J Exp Med* 2019; **216**(9):  
535 2038-2056.  
536
- 537 29. Al-Qahtani A, Al-Ahdal M, Abdo A, Sanai F, Al-Anazi M, Khalaf N *et al.* Toll-like  
538 receptor 3 polymorphism and its association with hepatitis B virus infection in Saudi  
539 Arabian patients. *J Med Virol* 2012; **84**(9): 1353-9.  
540
- 541 30. Guo Y, Audry M, Ciancanelli M, Alsina L, Azevedo J, Herman M *et al.* Herpes  
542 simplex virus encephalitis in a patient with complete TLR3 deficiency: TLR3 is  
543 otherwise redundant in protective immunity. *J Exp Med* 2011; **208**(10): 2083-98.  
544
- 545 31. Lim HK, Seppanen M, Hautala T, Ciancanelli MJ, Itan Y, Lafaille FG *et al.* TLR3  
546 deficiency in herpes simplex encephalitis: high allelic heterogeneity and recurrence  
547 risk. *Neurology* 2014; **83**(21): 1888-97.  
548
- 549 32. Zhang SY, Jouanguy E, Ugolini S, Smahi A, Elain G, Romero P *et al.* TLR3  
550 deficiency in patients with herpes simplex encephalitis. *Science* 2007; **317**(5844):  
551 1522-7.  
552
- 553 33. Ishizaki Y, Takemoto M, Kira R, Kusuhara K, Torisu H, Sakai Y *et al.* Association of  
554 toll-like receptor 3 gene polymorphism with subacute sclerosing panencephalitis. *J*  
555 *Neurovirol* 2008; **14**(6): 486-91.  
556

- 557 34. Sironi M, Biasin M, Cagliani R, Forni D, De Luca M, Saulle I *et al.* A common  
558 polymorphism in TLR3 confers natural resistance to HIV-1 infection. *J Immunol*  
559 2012; **188**(2): 818-23.  
560
- 561 35. Mork N, Kofod-Olsen E, Sorensen KB, Bach E, Orntoft TF, Ostergaard L *et al.*  
562 Mutations in the TLR3 signaling pathway and beyond in adult patients with herpes  
563 simplex encephalitis. *Genes Immun* 2015; **16**(8): 552-66.  
564
- 565 36. Sancho-Shimizu V, Perez de Diego R, Lorenzo L, Halwani R, Alangari A, Israelsson  
566 E *et al.* Herpes simplex encephalitis in children with autosomal recessive and  
567 dominant TRIF deficiency. *J Clin Invest* 2011; **121**(12): 4889-902.  
568
- 569 37. Negishi H, Osawa T, Ogami K, Ouyang X, Sakaguchi S, Koshiha R *et al.* A critical  
570 link between Toll-like receptor 3 and type II interferon signaling pathways in antiviral  
571 innate immunity. *Proc Natl Acad Sci U S A* 2008; **105**(51): 20446-51.  
572
- 573 38. Hardarson HS, Baker JS, Yang Z, Purevjav E, Huang CH, Alexopoulou L *et al.* Toll-  
574 like receptor 3 is an essential component of the innate stress response in virus-induced  
575 cardiac injury. *Am J Physiol Heart Circ Physiol* 2007; **292**(1): H251-8.  
576
- 577 39. Tabeta K, Georgel P, Janssen E, Du X, Hoebe K, Crozat K *et al.* Toll-like receptors 9  
578 and 3 as essential components of innate immune defense against mouse  
579 cytomegalovirus infection. *Proc Natl Acad Sci U S A* 2004; **101**(10): 3516-21.  
580

- 581 40. Le Goffic R, Balloy V, Lagranderie M, Alexopoulou L, Escriou N, Flavell R *et al.*  
582 Detrimental contribution of the Toll-like receptor (TLR)3 to influenza A virus-induced  
583 acute pneumonia. *PLoS Pathog* 2006; **2(6)**: e53.  
584
- 585 41. Hoebe K, Du X, Georgel P, Janssen E, Tabeta K, Kim SO *et al.* Identification of Lps2  
586 as a key transducer of MyD88-independent TIR signalling. *Nature* 2003; **424(6950)**:  
587 743-8.  
588
- 589 42. Riad A, Westermann D, Zietsch C, Savvatis K, Becher PM, Bereswill S *et al.* TRIF is  
590 a critical survival factor in viral cardiomyopathy. *J Immunol* 2011; **186(4)**: 2561-70.  
591
- 592 43. Hutchens M, Luker KE, Sottile P, Sonstein J, Lukacs NW, Nunez G *et al.* TLR3  
593 increases disease morbidity and mortality from vaccinia infection. *J Immunol* 2008;  
594 **180(1)**: 483-91.  
595
- 596 44. Hutchens MA, Luker KE, Sonstein J, Nunez G, Curtis JL, Luker GD. Protective effect  
597 of Toll-like receptor 4 in pulmonary vaccinia infection. *PLoS Pathog* 2008; **4(9)**:  
598 e1000153.  
599
- 600 45. Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA. Toll-like  
601 receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis.  
602 *Nat Med* 2004; **10(12)**: 1366-73.  
603
- 604 46. Daffis S, Samuel MA, Suthar MS, Gale M, Jr., Diamond MS. Toll-like receptor 3 has  
605 a protective role against West Nile virus infection. *J Virol* 2008; **82(21)**: 10349-58.

606

607 47. Town T, Bai F, Wang T, Kaplan AT, Qian F, Montgomery RR *et al.* Toll-like receptor  
608 7 mitigates lethal West Nile encephalitis via interleukin 23-dependent immune cell  
609 infiltration and homing. *Immunity* 2009; **30**(2): 242-53.

610

611 48. Kurt-Jones EA, Chan M, Zhou S, Wang J, Reed G, Bronson R *et al.* Herpes simplex  
612 virus 1 interaction with Toll-like receptor 2 contributes to lethal encephalitis. *Proc*  
613 *Natl Acad Sci U S A* 2004; **101**(5): 1315-20.

614

615 49. Schilte C, Couderc T, Chretien F, Sourisseau M, Gangneux N, Guivel-Benhassine F *et*  
616 *al.* Type I IFN controls chikungunya virus via its action on nonhematopoietic cells. *J*  
617 *Exp Med* 2010; **207**(2): 429-42.

618

619 50. Sheahan T, Morrison TE, Funkhouser W, Uematsu S, Akira S, Baric RS *et al.* MyD88  
620 is required for protection from lethal infection with a mouse-adapted SARS-CoV.  
621 *PLoS Pathog* 2008; **4**(12): e1000240.

622

623 51. Szretter KJ, Daffis S, Patel J, Suthar MS, Klein RS, Gale M, Jr. *et al.* The innate  
624 immune adaptor molecule MyD88 restricts West Nile virus replication and spread in  
625 neurons of the central nervous system. *J Virol* 2010; **84**(23): 12125-38.

626

627 52. Gralinski LE, Menachery VD, Morgan AP, Totura AL, Beall A, Kocher J *et al.* Allelic  
628 Variation in the Toll-Like Receptor Adaptor Protein Ticam2 Contributes to SARS-  
629 Coronavirus Pathogenesis in Mice. *G3 (Bethesda)* 2017; **7**(6): 1653-1663.

630

- 631 53. Jorgensen SE, Christiansen M, Ryo LB, Gad HH, Gjedsted J, Staeheli P *et al.*  
632 Defective RNA sensing by RIG-I in severe influenza virus infection. *Clin Exp*  
633 *Immunol* 2018; **192**(3): 366-376.  
634
- 635 54. Asgari S, Schlapbach LJ, Anchisi S, Hammer C, Bartha I, Junier T *et al.* Severe viral  
636 respiratory infections in children with IFIH1 loss-of-function mutations. *Proc Natl*  
637 *Acad Sci U S A* 2017; **114**(31): 8342-8347.  
638
- 639 55. Lamborn IT, Jing H, Zhang Y, Drutman SB, Abbott JK, Munir S *et al.* Recurrent  
640 rhinovirus infections in a child with inherited MDA5 deficiency. *J Exp Med* 2017;  
641 **214**(7): 1949-1972.  
642
- 643 56. Hoffmann FS, Schmidt A, Dittmann Chevillotte M, Wisskirchen C, Hellmuth J,  
644 Willms S *et al.* Polymorphisms in melanoma differentiation-associated gene 5 link  
645 protein function to clearance of hepatitis C virus. *Hepatology* 2015; **61**(2): 460-70.  
646
- 647 57. Kandasamy M, Suryawanshi A, Tundup S, Perez JT, Schmolke M, Manicassamy S *et*  
648 *al.* RIG-I Signaling Is Critical for Efficient Polyfunctional T Cell Responses during  
649 Influenza Virus Infection. *PLoS Pathog* 2016; **12**(7): e1005754.  
650
- 651 58. Wu W, Wang X, Zhang W, Tian L, Booth JL, Duggan ES *et al.* RIG-I Signaling via  
652 MAVS Is Dispensable for Survival in Lethal Influenza Infection In Vivo. *Mediators*  
653 *Inflamm* 2018; **2018**: 6808934.  
654

- 655 59. Zalinger ZB, Elliott R, Rose KM, Weiss SR. MDA5 Is Critical to Host Defense during  
656 Infection with Murine Coronavirus. *J Virol* 2015; **89**(24): 12330-40.  
657
- 658 60. Banos-Lara Mdel R, Ghosh A, Guerrero-Plata A. Critical role of MDA5 in the  
659 interferon response induced by human metapneumovirus infection in dendritic cells  
660 and in vivo. *J Virol* 2013; **87**(2): 1242-51.  
661
- 662 61. Lu HL, Liao F. Melanoma differentiation-associated gene 5 senses hepatitis B virus  
663 and activates innate immune signaling to suppress virus replication. *J Immunol* 2013;  
664 **191**(6): 3264-76.  
665
- 666 62. McCartney SA, Thackray LB, Gitlin L, Gilfillan S, Virgin HW, Colonna M. MDA-5  
667 recognition of a murine norovirus. *PLoS Pathog* 2008; **4**(7): e1000108.  
668
- 669 63. Jin YH, Kim SJ, So EY, Meng L, Colonna M, Kim BS. Melanoma differentiation-  
670 associated gene 5 is critical for protection against Theiler's virus-induced  
671 demyelinating disease. *J Virol* 2012; **86**(3): 1531-43.  
672
- 673 64. Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K *et al.* Differential  
674 roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 2006;  
675 **441**(7089): 101-5.  
676
- 677 65. Gitlin L, Barchet W, Gilfillan S, Cella M, Beutler B, Flavell RA *et al.* Essential role of  
678 mda-5 in type I IFN responses to polyriboinosinic:polyribocytidylic acid and

679           encephalomyocarditis picornavirus. *Proc Natl Acad Sci U S A* 2006; **103**(22): 8459-  
680           64.  
681  
682   66.    Errett JS, Suthar MS, McMillan A, Diamond MS, Gale M, Jr. The essential,  
683           nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus  
684           infection. *J Virol* 2013; **87**(21): 11416-25.  
685  
686   67.    Suthar MS, Ma DY, Thomas S, Lund JM, Zhang N, Daffis S *et al.* IPS-1 is essential  
687           for the control of West Nile virus infection and immunity. *PLoS Pathog* 2010; **6**(2):  
688           e1000757.  
689  
690   68.    Wang JP, Cerny A, Asher DR, Kurt-Jones EA, Bronson RT, Finberg RW. MDA5 and  
691           MAVS mediate type I interferon responses to coxsackie B virus. *J Virol* 2010; **84**(1):  
692           254-60.  
693  
694   69.    Kumar H, Kawai T, Kato H, Sato S, Takahashi K, Coban C *et al.* Essential role of  
695           IPS-1 in innate immune responses against RNA viruses. *J Exp Med* 2006; **203**(7):  
696           1795-803.  
697  
698   70.    Sun Q, Sun L, Liu HH, Chen X, Seth RB, Forman J *et al.* The specific and essential  
699           role of MAVS in antiviral innate immune responses. *Immunity* 2006; **24**(5): 633-42.  
700  
701   71.    Perry ST, Prestwood TR, Lada SM, Benedict CA, Shresta S. Cardif-mediated  
702           signaling controls the initial innate response to dengue virus in vivo. *J Virol* 2009;  
703           **83**(16): 8276-81.



704

705 72. Si-Tahar M, Blanc F, Furio L, Choppy D, Balloy V, Lafon M *et al.* Protective role of  
706 LGP2 in influenza virus pathogenesis. *J Infect Dis* 2014; **210**(2): 214-23.

707

708 73. Perez de Diego R, Sancho-Shimizu V, Lorenzo L, Puel A, Plancoulaine S, Picard C *et*  
709 *al.* Human TRAF3 adaptor molecule deficiency leads to impaired Toll-like receptor 3  
710 response and susceptibility to herpes simplex encephalitis. *Immunity* 2010; **33**(3): 400-  
711 11.

712

713 74. Herman M, Ciancanelli M, Ou YH, Lorenzo L, Klaudel-Dreszler M, Pauwels E *et al.*  
714 Heterozygous TBK1 mutations impair TLR3 immunity and underlie herpes simplex  
715 encephalitis of childhood. *J Exp Med* 2012; **209**(9): 1567-82.

716

717 75. Tenoever BR, Ng SL, Chua MA, McWhirter SM, Garcia-Sastre A, Maniatis T.  
718 Multiple functions of the IKK-related kinase IKKepsilon in interferon-mediated  
719 antiviral immunity. *Science* 2007; **315**(5816): 1274-8.

720

721 76. Andersen LL, Mork N, Reinert LS, Kofod-Olsen E, Narita R, Jorgensen SE *et al.*  
722 Functional IRF3 deficiency in a patient with herpes simplex encephalitis. *J Exp Med*  
723 2015; **212**(9): 1371-9.

724

725 77. Lin R, Mamane Y, Hiscott J. Structural and functional analysis of interferon  
726 regulatory factor 3: localization of the transactivation and autoinhibitory domains. *Mol*  
727 *Cell Biol* 1999; **19**(4): 2465-74.

728

- 729 78. Ciancanelli MJ, Huang SX, Luthra P, Garner H, Itan Y, Volpi S *et al.* Infectious  
730 disease. Life-threatening influenza and impaired interferon amplification in human  
731 IRF7 deficiency. *Science* 2015; **348**(6233): 448-53.  
732
- 733 79. Hatesuer B, Hoang HT, Riese P, Trittel S, Gerhauser I, Elbahesh H *et al.* Deletion of  
734 Irf3 and Irf7 Genes in Mice Results in Altered Interferon Pathway Activation and  
735 Granulocyte-Dominated Inflammatory Responses to Influenza A Infection. *J Innate*  
736 *Immun* 2017; **9**(2): 145-161.  
737
- 738 80. Sato M, Suemori H, Hata N, Asagiri M, Ogasawara K, Nakao K *et al.* Distinct and  
739 essential roles of transcription factors IRF-3 and IRF-7 in response to viruses for IFN-  
740 alpha/beta gene induction. *Immunity* 2000; **13**(4): 539-48.  
741
- 742 81. Daffis S, Samuel MA, Keller BC, Gale M, Jr., Diamond MS. Cell-specific IRF-3  
743 responses protect against West Nile virus infection by interferon-dependent and -  
744 independent mechanisms. *PLoS Pathog* 2007; **3**(7): e106.  
745
- 746 82. Daffis S, Samuel MA, Suthar MS, Keller BC, Gale M, Jr., Diamond MS. Interferon  
747 regulatory factor IRF-7 induces the antiviral alpha interferon response and protects  
748 against lethal West Nile virus infection. *J Virol* 2008; **82**(17): 8465-75.  
749
- 750 83. Honda K, Yanai H, Negishi H, Asagiri M, Sato M, Mizutani T *et al.* IRF-7 is the  
751 master regulator of type-I interferon-dependent immune responses. *Nature* 2005;  
752 **434**(7034): 772-7.  
753

- 754 84. Rudd PA, Wilson J, Gardner J, Larcher T, Babarit C, Le TT *et al.* Interferon response  
755 factors 3 and 7 protect against Chikungunya virus hemorrhagic fever and shock. *J*  
756 *Virol* 2012; **86**(18): 9888-98.  
757
- 758 85. Chen HW, King K, Tu J, Sanchez M, Luster AD, Shresta S. The roles of IRF-3 and  
759 IRF-7 in innate antiviral immunity against dengue virus. *J Immunol* 2013; **191**(8):  
760 4194-201.  
761
- 762 86. Lazear HM, Govero J, Smith AM, Platt DJ, Fernandez E, Miner JJ *et al.* A Mouse  
763 Model of Zika Virus Pathogenesis. *Cell Host Microbe* 2016; **19**(5): 720-30.  
764
- 765 87. Carlin AF, Plummer EM, Vizcarra EA, Sheets N, Joo Y, Tang W *et al.* An IRF-3-,  
766 IRF-5-, and IRF-7-Independent Pathway of Dengue Viral Resistance Utilizes IRF-1 to  
767 Stimulate Type I and II Interferon Responses. *Cell Rep* 2017; **21**(6): 1600-1612.  
768
- 769 88. Brannan JM, Froude JW, Prugar LI, Bakken RR, Zak SE, Daye SP *et al.* Interferon  
770 alpha/beta Receptor-Deficient Mice as a Model for Ebola Virus Disease. *J Infect Dis*  
771 2015; **212 Suppl 2**: S282-94.  
772
- 773 89. Sheehan KC, Lazear HM, Diamond MS, Schreiber RD. Selective Blockade of  
774 Interferon-alpha and -beta Reveals Their Non-Redundant Functions in a Mouse Model  
775 of West Nile Virus Infection. *PLoS One* 2015; **10**(5): e0128636.  
776
- 777 90. Lazear HM, Pinto AK, Vogt MR, Gale M, Jr., Diamond MS. Beta interferon controls  
778 West Nile virus infection and pathogenesis in mice. *J Virol* 2011; **85**(14): 7186-94.

779

780 91. Deonarain R, Alcamì A, Alexiou M, Dallman MJ, Gewert DR, Porter AC. Impaired  
781 antiviral response and alpha/beta interferon induction in mice lacking beta interferon.  
782 *J Virol* 2000; **74**(7): 3404-9.

783

784 92. Koerner I, Kochs G, Kalinke U, Weiss S, Staeheli P. Protective role of beta interferon  
785 in host defense against influenza A virus. *J Virol* 2007; **81**(4): 2025-30.

786

787 93. Deonarain R, Cerullo D, Fuse K, Liu PP, Fish EN. Protective role for interferon-beta  
788 in coxsackievirus B3 infection. *Circulation* 2004; **110**(23): 3540-3.

789

790 94. Gerlach N, Schimmer S, Weiss S, Kalinke U, Dittmer U. Effects of type I interferons  
791 on Friend retrovirus infection. *J Virol* 2006; **80**(7): 3438-44.

792

793 95. Hagan RS, Torres-Castillo J, Doerschuk CM. Myeloid TBK1 Signaling Contributes to  
794 the Immune Response to Influenza. *Am J Respir Cell Mol Biol* 2019; **60**(3): 335-345.

795

796 96. Lalani AI, Moore CR, Luo C, Kreider BZ, Liu Y, Morse HC, 3rd *et al.* Myeloid cell  
797 TRAF3 regulates immune responses and inhibits inflammation and tumor  
798 development in mice. *J Immunol* 2015; **194**(1): 334-48.

799

800 97. Marchlik E, Thakker P, Carlson T, Jiang Z, Ryan M, Marusic S *et al.* Mice lacking  
801 Tbk1 activity exhibit immune cell infiltrates in multiple tissues and increased  
802 susceptibility to LPS-induced lethality. *J Leukoc Biol* 2010; **88**(6): 1171-80.

803

- 804 98. Manet C, Simon-Loriere E, Jouvion G, Hardy D, Prot M, Conquet L *et al.* Genetic  
805 Diversity of Collaborative Cross Mice Controls Viral Replication, Clinical Severity,  
806 and Brain Pathology Induced by Zika Virus Infection, Independently of Oas1b. *J Virol*  
807 2020; **94**(3).
- 808
- 809 99. Channappanavar R, Fehr AR, Vijay R, Mack M, Zhao J, Meyerholz DK *et al.*  
810 Dysregulated Type I Interferon and Inflammatory Monocyte-Macrophage Responses  
811 Cause Lethal Pneumonia in SARS-CoV-Infected Mice. *Cell Host Microbe* 2016;  
812 **19**(2): 181-93.
- 813
- 814 100. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Moller R *et al.*  
815 Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*  
816 2020; **181**(5): 1036-1045 e9.
- 817
- 818 101. Steed AL, Christophi GP, Kaiko GE, Sun L, Goodwin VM, Jain U *et al.* The  
819 microbial metabolite desaminotyrosine protects from influenza through type I  
820 interferon. *Science* 2017; **357**(6350): 498-502.

821  
822  
823

824 **Legends**

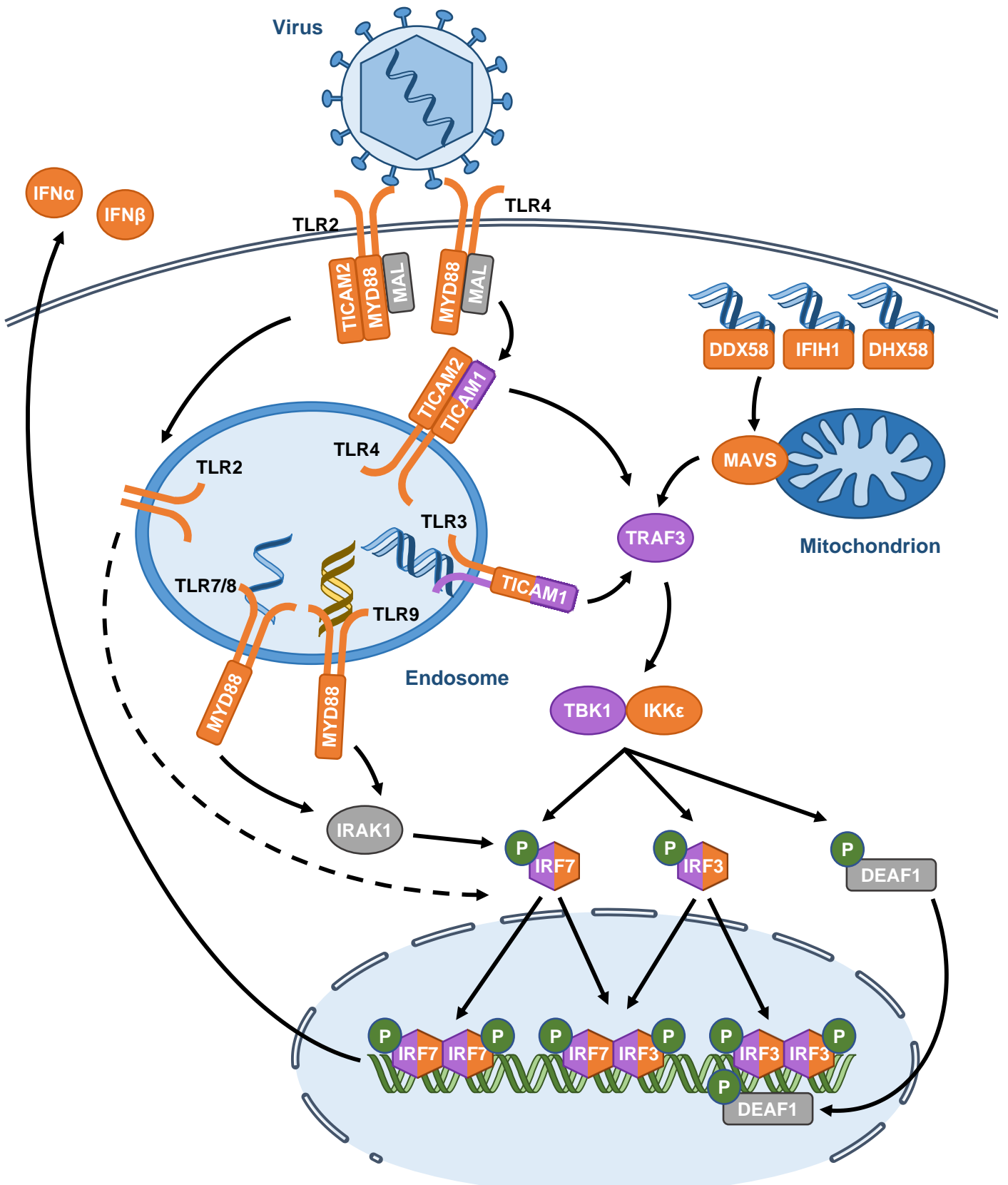
825

826 **Figure 1: Induction of IFN $\alpha$  and IFN $\beta$**

827 Viral molecules (DNA, RNA and proteins) induce the expression of IFN-I after their  
828 recognition by TLRs and RLRs. Signalization leads the activation of kinases, TBK1 and  
829 IKK $\epsilon$  responsible for the activation of the transcription factors IRF3 and IRF7 which induce  
830 the expression of IFN $\alpha$  and IFN $\beta$ . Proteins for which the corresponding gene was associated  
831 with susceptibility to virus infection are indicated in orange for mouse studies and in purple  
832 for human studies. Blue, yellow and green helices depict viral RNA, viral DNA and cellular  
833 DNA molecules, respectively. Gene names are spelled according to the nomenclature rules for  
834 human genes.

835

836



**Table 1: Genes with variants associated with susceptibility to viral infections in humans**

Virus	Type of study	Method	Polymorphism	Impact on gene function	Phenotype	Reference
<b><i>TLR3</i></b>						
HIV	Association study	Genotyping of a candidate gene	L412F [missense]	Increased cytokine production following stimulation of TLR3 and lower viral load	Decreased risk of infection	34
IAV	Association study	Genotyping of candidate genes	rs5743313 [intronic]	ND	Increased risk of pneumonia	25
IAV	Association study	Sequencing of candidate genes	rs5743313 [intronic]	ND	Increased risk of fatal infection	26
IAV	Case study	Sequencing of candidate genes	F303S [missense]	No induction of IFN $\beta$ and no activation of NF $\kappa$ B	Influenza-associated encephalopathy	27
IAV	Case study	WES	P554S [missense], P680L [missense]	Decreased expression of IFN $\beta$ and IFN $\lambda$	Acute respiratory distress syndrome	28
HBV	Association study	Genotyping of a candidate gene	rs1879026 [intronic]	ND	Increased risk of being infected	29
HSV-1	Case study	Sequencing of a candidate gene	P554S [missense], E746X [nonsense], G743D [missense], R811I [missense], L360P [missense]	Decreased induction of cytokines and higher viral replication rate	Herpes simplex encephalitis	30-32
MV	Association study	Genotyping of candidate genes	L412F [missense]	ND	Increased risk of subacute sclerosing panencephalitis	33
SARS-CoV-2	Association study	Sequencing of candidate genes	S339fs [frameshift], P554S [missense], W769X [non sense], M870V [missense]	Decreased expression of IFN $\lambda$	Life-threatening COVID-19	21
<b><i>TICAM1</i></b>						
HSV-1	Case study	WES	A568T, S160F [missense]	Decreased induction of cytokines	Herpes simplex encephalitis	35
HSV-1	Case study	Sequencing of candidate genes	R141X [nonsense], S186L [missense]	Impaired activation of IRF3 and NF $\kappa$ B, decreased induction of cytokines	Herpes simplex encephalitis	36
SARS-CoV-2	Association study	Sequencing of candidate genes	T4I [missense], S60C [missense], Q392K [missense]	Decreased expression of IFN $\beta$	Life-threatening COVID-19	21
<b><i>DDX58</i></b>						
IAV	Case study	WES	R71H + P885S [missense]	Decreased response to ligand and expression of IFN $\beta$	Severe influenza infection	53
<b><i>IFIH1</i></b>						
HCV	Association study	Genotyping of candidate genes	H843A [missense]	Decreased expression of IFN $\beta$ and other cytokines	Chronic hepatitis C	56
HRV	Case study	WES	rs35732034 [intronic]	Lack of exon 14, decreased expression of IFN $\beta$	Bronchiolitis	54
HRV	Case study	WES	E627X [nonsense]	Lack of CTD, decreased expression of IFN $\beta$	Bronchiolitis, pneumonia	54
RSV	Case study	WES	rs35732034 [intronic]	Lack of exon 14, decreased expression of IFN $\beta$	Bronchiolitis, pneumonia	54
RSV	Case study	WES	rs35337543 [intronic]	Lack of exon 8, decreased expression of IFN $\beta$	Bronchiolitis	54
<b><i>TRAF3</i></b>						
HSV-1	Case study	Sequencing of a candidate gene	R118W [missense]	Decreased TRAF3 production	Herpes simplex encephalitis	73
<b><i>TBK1</i></b>						
HSV-1	Case study	Sequencing of a candidate gene	D50A [missense]	Decreased amount of TBK1 mRNA and protein, decreased cytokine production	Herpes simplex encephalitis	74
HSV-1	Case study	Sequencing of a candidate gene	D159A [missense]	No enzyme activity, decreased cytokine production	Herpes simplex encephalitis	74
SARS-CoV-2	Association study	Sequencing of candidate genes	F24S [missense], R308X [nonsense]	Decreased expression of IFN $\beta$	Life-threatening COVID-19	21
<b><i>IRF3</i></b>						
HSV-1	Case study	WES	R285Q [missense]	No phosphorylation and dimerization of IRF3, decreased cytokine production	Herpes simplex encephalitis	35, 76
HSV-1	Case study	WES	A277T [missense]	Decreased cytokine production	Herpes simplex encephalitis	35
SARS-CoV-2	Association study	Sequencing of candidate genes	E49del [deletion], N146K [missense]	Decreased expression of IFN $\beta$	Life-threatening COVID-19	21
WNV	Association study	Genotyping of candidate genes	rs2304207 [intronic]	ND	Increased risk to have a symptomatic infection	22
<b><i>IRF7</i></b>						
IAV	Case study	WES	Q421X [nonsense]	Absence of phosphorylation and nuclear localization in absence of infection, impaired IFN $\alpha$ production (in compound heterozygosity with F410V)	Life-threatening infection	78
IAV	Case study	WES	F410V [missense]	Inability to translocate to the nucleus, impaired IFN $\alpha$ production (in the case of compound heterozygosity with Q421X)	Life-threatening infection	78
SARS-CoV-2	Association study	Sequencing of candidate genes	R7fs [frameshift], F95S [missense], D117N [missense], Q185X [nonsense], P246fs [frameshift], R369Q [missense], M371V [missense]	Decreased expression of IFN $\beta$	Life-threatening COVID-19	21

For exonic variants, the effect is indicated [missense/nonsense/deletion/frameshift]

Abbreviations: HBV: hepatitis B virus; HRV: human rhinovirus; HSV-1: herpes simplex virus type 1; IAV: influenza A virus; MV: measles virus; RSV: respiratory syncytial virus; WNV: West Nile virus, del: deletion, fs: frameshift, X: stop codon, WES: whole exome sequencing, ND: not determined



**Table 2: Genes with variants associated with susceptibility to viral infections in mice**

Virus	Variant*	Mortality**	Viral titer [tissue]**	Type I IFN expression [tissue]**	Reference
<i>Tlr3</i>					
CVB3	KO	+	+ [heart, serum, splenocytes]	= [heart]	37
EMCV	KO	+	+ [heart, liver]	+ [heart]	38
IAV	KO	-	+ [lung]	ND	40
MCMV	KO	=	+ [spleen]	- [serum]	39
VV	KO	-	- [abdomen, lung, chest]	= [lung]	43
WNV	KO	+	+ [brain, spinal cord, spleen]	= [lymph node, serum]	46
WNV	KO	-	+ [blood] - [brain]	- [blood, brain]	45
<i>Ticam1</i>					
CVB3	KO	+	+ [heart]	- 72h + 7 days [heart]	42
MCMV	Point mutation	+	+ [spleen]	- [serum]	41
VV	KO	ND	+ [chest]	ND	44
<i>Tlr2</i>					
HSV-1	KO	-	= [brain]	ND	48
<i>Tlr4</i>					
VV	Point mutation	+	+ [abdomen, chest, head, lung]	= [lung]	44
<i>Tlr7</i>					
WNV	KO	+	+ [blood, brain, spleen]	+ [blood]	47
<i>Tlr9</i>					
MCMV	Point mutation	+	+ [spleen]	- [serum]	39
<i>Myd88</i>					
CHIKV	KO	ND	+ [joint, serum, spleen]	ND	49
MCMV	KO	+	+ [spleen]	- [serum]	39
SARS-CoV	KO	+	+ [lung]	= [lung]	50
WNV	KO	+	+ [blood, brain, spleen]	ND	47
WNV	KO	+	+ [brain, lymph node, spleen]	+ [serum]	51
<i>Ticam2</i>					
SARS-CoV	KO	ND	+ [lung]	ND	52
<i>Ddx58</i>					
IAV	KO	=	+ [lung]	ND	57
CHIKV	KO	ND	+ [serum]	ND	49
JEV	KO	+ §	ND	- [serum] §	64
WNV	KO	+	+ [MEF]	- [MEF]	66
<i>Ifih1</i>					
CVB3	KO	+	= [liver, pancreas, serum]	- [pancreas, serum]	68
EMCV	KO	+ §	+ [heart] §	- [serum] §	64
EMCV	KO	+	ND	- [DC, MP]	65
HBV	KO	ND	+ [liver, serum]	ND	61
hMPV	KO	ND	+ [lung]	- [lung]	60
MHV	KO	+	+ [brain, heart, kidney, lung, spinal cord, spleen]	- [liver]	59
MNV-1	KO	ND	+ [intestine, spleen, lymph node]	- [DC]	62
TMEV	KO	ND	+ [brain, spinal cord]	- [brain, spinal cord]	63
WNV	KO	+	+ [DC, MP]	- [DC, MEF, MP]	66
<i>Dhx58</i>					
EMCV	KO + point mutation	+	+ [heart]	- [serum]	13
IAV	TG	-	= [lung]	- [lung]	72
WNV	KO	+	+ [brain, DC, MP]	- [DC, MP]	14
<i>Mavs</i>					
CHIKV	KO	ND	+ [serum]	ND	49
CVB3	KO	+	= [liver, pancreas, serum]	- [pancreas, serum]	68
DENV	KO	=	+ [bone marrow, lymph node, serum, spleen]	- [bone marrow, lymph node, serum, spleen]	71

EMCV	KO	+	+ [heart] §	- [serum]	69
VSV	KO	+	+ [brain, liver]	ND	69
VSV	KO	+	+ [serum]	= [serum]	70
WNV	KO	+	+ [brain, DC, kidney, MP, serum, spinal cord, spleen]	- [DC, MP]	67
WNV	KO	+	+ [DC]	- [DC]	66
<b><i>Ikbke</i></b>					
IAV	KO	+	+ [lung, MEF]	= [lung, MEF]	75
<b><i>Irf3</i></b>					
IAV	KO	+	+ [lung]	- [lung]	79
WNV	KO	+	+ [brain, kidney, lymph node, serum, spinal cord, spleen]	- [lymph node]	81
<b><i>Irf7</i></b>					
DENV	KO	ND	+ [spleen]	- [serum]	85
IAV	KO	+	ND	- [serum]	80
IAV	KO	+	= [lung]	- [lung]	79
EMCV	KO	+	ND	- [serum]	83
HSV	KO	+	ND	- [serum]	83
WNV	KO	+	+ [brain, kidney, lymph node, serum, spinal cord, spleen]	- [brain, DC, MEF, MP, serum]	83
WNV	KO	+	+ [brain, cortical neurons, DC, kidney, lymph node, MEF, MP, serum, spleen, spinal cord]	- [brain, cortical neurons, DC, MEF, MP]	82
<b><i>Irf3-Irf7</i></b>					
CHIKV	KO	+	+ [blood, brain, liver, muscle, spleen]	- [blood, feet]	84
<b><i>Irf3-Irf5-Irf7</i></b>					
ZIKV	KO	+	ND	ND	86
<b><i>Irf1-Irf3-Irf5-Irf7</i></b>					
DENV	KO	+	+ [MP]	- [MP, serum]	87
<b><i>Iffa</i></b>					
WNV	mAb treated	+	ND	ND	89
<b><i>Iffb1</i></b>					
CVB3	KO	+	+ [liver, spleen]	ND	93
FV	KO	ND	+ [spleen]	= [plasma]	94
IAV	KO	+	+ [lung, MEF]	ND	92
VV	KO	+	+ [lung]	- [MEF]	91
WNV	KO	+	+ [brain, granule cell neurons, kidney, lymph node, myeloid cells, MEF, serum, spinal cord]	+ [serum]	90
WNV	mAb treated	+	ND	ND	89

\* Genetic variant except for 'mAb treated' in which case gene product was transiently inhibited using a monoclonal antibody. The phenotype was compared with wild type mice, except for § (compared with heterozygous mice).

\*\* + : increased; - : decreased; = : unchanged; ND: not determined.

Abbreviations: CVB3: coxsackievirus B3; CHIKV: chikungunya virus; DENV: dengue virus; EMCV: encephalomyocarditis virus; FV: friend virus; hMPV: human metapneumovirus; HSV-1: herpes simplex virus type 1; IAV: influenza A virus; JEV: Japanese encephalitis virus; MCMV: murine cytomegalovirus; MNV-1: murine norovirus 1; SARS-CoV: severe acute respiratory syndrome coronavirus; TMEV: Theiler's murine encephalomyelitis virus; VSV: vesicular stomatitis virus; VV: vaccinia virus; WNV: West Nile virus; KO: knocked-out mice, TG: transgenic mice; MEF: mouse embryonic fibroblasts, DC: dendritic cells, MP: macrophages.

### **II.1.4.3 . The role of IRF3 to face viral infection**

The results of my project led me to study the role of *Irf3* in the susceptibility of a mouse strain to ZIKV both *in vitro* and *in vivo*. After analyzing the literature on this gene, beyond the case of mice and ZIKV, I wrote a review on its role in the response to viral infection. This review is a work in progress, and here I provide its current version.

#### **Title**

IRF3 in viral infections: the interferon response and beyond

#### **Abstract**

IRF3 is the first transcription factor involved in the induction of IFN-I. It is present in the cytoplasm of most cell types in basal conditions, allowing a rapid triggering of the IFN-I pathway after viral infection. IRF3 has often been presented in opposition with IRF7, the other major IFN-I transcription factor, but the role of IRF3 in viral infection is more complex than expected and extends beyond the IFN-I pathway. Here, we review the impact of IRF3 deficiency in infected cells and *in vivo*, in mice and in humans. We discuss the discrepancies between and within studies from the literature, and we highlight how the newly discovered functions of IRF3 may help reconsider its role in viral infection.

# Title: IRF3 in viral infections: the interferon response and beyond

Running title: Multiple roles of IRF3 in viral infections

Marie Bourdon<sup>1</sup>, Caroline Manet<sup>1</sup>, Xavier Montagutelli<sup>#1</sup>

## Affiliation:

<sup>1</sup>Institut Pasteur, Université de Paris, Mouse Genetics Laboratory, F-75015 Paris, France

# Corresponding author: [xavier.montagutelli@pasteur.fr](mailto:xavier.montagutelli@pasteur.fr)

## Abstract

IRF3 is the first transcription factor involved in the induction of type I interferons (IFN-I). It is present in the cytoplasm of most cell types in basal conditions, allowing a rapid triggering of the IFN-I pathway after viral infection. IRF3 has often been presented in opposition with IRF7, the other major IFN-I transcription factor, but the role of IRF3 in viral infection is more complex than expected and extends beyond the IFN-I pathway. Here, we review the impact of IRF3 deficiencies in infected cells and *in vivo*, in mice and in humans. We discuss the discrepancies between and within studies from the literature, and we highlight how the newly discovered functions of IRF3 may help reconsider its role in viral infection.

## Introduction

Induction of type I interferons (IFN-I) by viruses and other pathogens is a critical step in innate immunity and is mediated by the activation of pattern-recognition receptors (PRRs), such as Toll-like receptors, and cytosolic sensors such as RIG-I and MDA5. Recognition of viral molecules by PRRs triggers a signaling pathway leading to the phosphorylation of interferon regulatory factors (IRFs) by kinases IKK $\epsilon$  and TKB1. Activated IRFs form homodimers or heterodimers which translocate to the nucleus and bind regulatory cis elements on the promoters of IFN-I genes [1]. The major IFN-I are IFN $\alpha$  encoded by 13 genes in humans and 14 in mice, and IFN $\beta$  is encoded by the single *Ifnb1* gene in both species [2]. Once secreted, IFN-I bind their receptor which triggers a signaling cascade leading to the expression of interferon stimulated genes (ISGs) with antiviral properties [3].

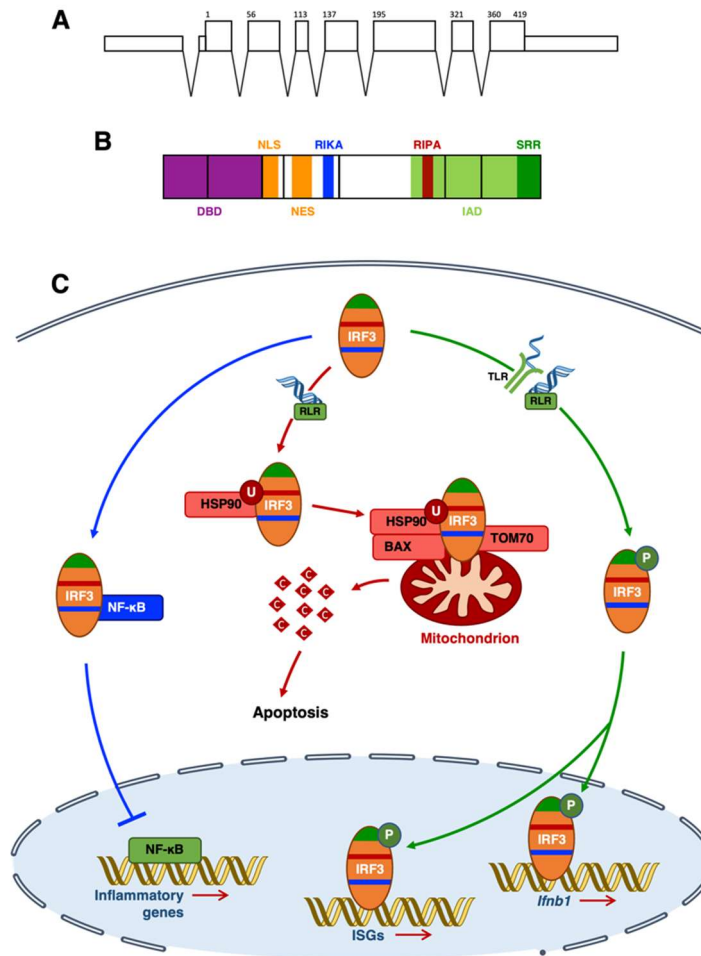
The transcription of type I IFN genes is activated by a family of IRFs containing nine members. IRFs possess diverse functions including the establishment of innate and adaptive immunity and the regulation of growth and differentiation of multiple cell types [4]. Among them, IRF3 and IRF7 are the main transcription activators of IFN-I. IRF3 was identified in 1995 by Au and colleagues, by homology with the other IRFs known at the time, IRF1 and IRF2. IRF3 is expressed in multiple human tissues (spleen, thymus, prostate, ovaries, etc.) and its expression is not modulated by virus infection nor IFN treatment [5].

Since the 2000s, the impact of *Irf3* deficiencies in viral diseases has been extensively investigated in humans and in mice [6–9]. While *in vitro* studies often found a role of *Irf3* in IFN-I levels and control of viral replication [7, 10, 11], *in vivo* studies have led to more contrasting results, as the susceptibility of *Irf3* deficient mice to viral disease depended on the virus and the experimental parameters [7]. Here, we review the multiple functions of IRF3 against viral infections and the consequences of *Irf3* deficiencies in human and mouse, *in vitro* and *in vivo*.

## Structure and functions of IRF3

The most studied function of IRF3 is its role of transcription factor. IRF3 is involved in the expression of IFN-I genes, but IRF3 also directly induces the expression of other genes, notably ISGs, such as *Isg15* [5], *Isg56* [12], *Isg54* [13–15], *Ccl5* (also known as RANTES) [16, 17], and *Cxcl10* (also known as IP-10) [13], by binding to the IFN-stimulated response elements (ISRE) in their promoters. IRF3 also induces the expression of the anti-inflammatory cytokine *Il1rn* [18] and the proapoptotic factors *Tnfsf10* (also known as TRAIL) and *Pmaip1* (also known as NOXA) [15, 19–21].

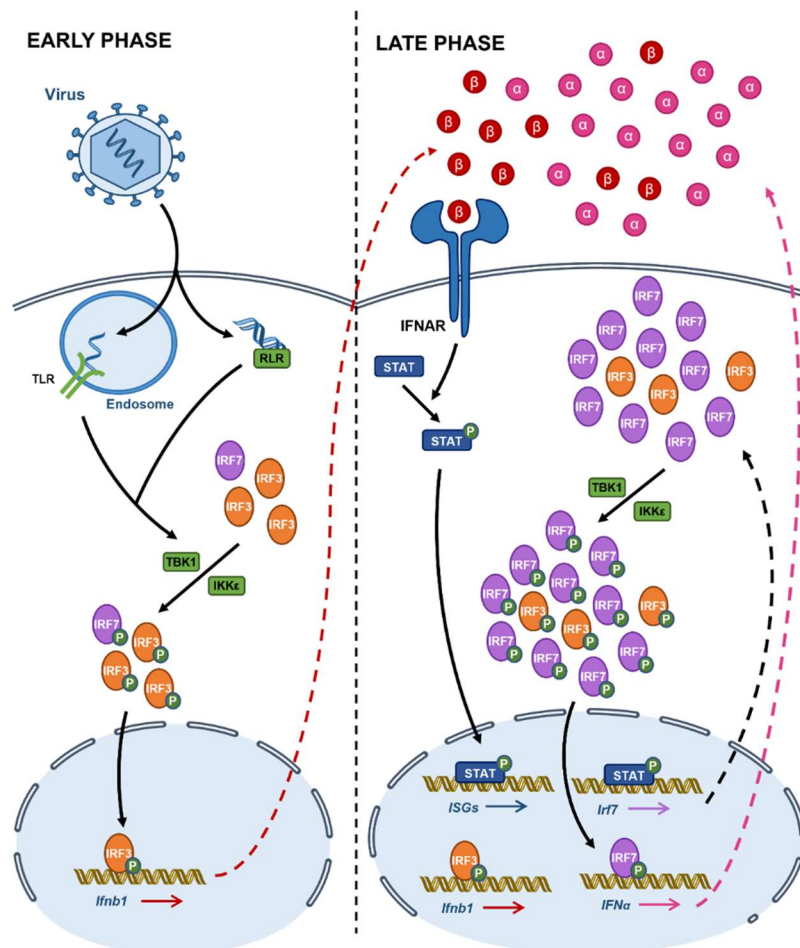
IRF3 possesses five protein domains necessary for its transcriptional functions. It has a serine rich region (SRR), where it is phosphorylated by the kinases IKK $\epsilon$  and TKB1. Once activated, IRF3 forms dimers through its IRF association domain (IAD) in C-terminal. The protein can access and exit the nucleus using its nuclear localization signal (NLS) and nuclear exit signal (NES), respectively. Lastly, IRF3 binds gene promoters with its N-terminal DNA-binding domain (DBD) forming a helix-turn-helix structure (Figure 1A-B) [22, 23].



**Figure 1: Structure and functions of IRF3.** (A) Organization of the mouse *Irf3* gene. Rectangles represent exons. Thin exons represent untranslated exonic regions. Numbers represent the length in amino acids. (B) Protein organization with structural domain. DBD: DNA binding domain. NLS: nuclear localization signal. NES: nuclear exit signal. IAD: IRF association domain. SRR: serin rich region. (C) Representations of the three functions of IRF3: RIKA (left), RIKI (middle), and transcription (right).

IRF3 and IRF7 have complementary functions for IFN-I induction in infected cells. First, the two transcription factors have different affinities for the IFN-I genes. IRF3 binds preferentially to the promoters of *Irf1* and *Irf4* genes, while IRF7 has more affinity for other IFN $\alpha$  genes [24]. IRF3 and IRF7 also induce type III IFNs, a family of four IFN $\lambda$ s in humans and three in mice [25]. As with IFN-I, IRF3 has more affinity for the *Irf1* promoter, while IRF7 binds preferentially to *Irf2* and *Irf3* promoters [26, 27].

Moreover, *Irf3* and *Irf7* have different expression patterns. *Irf3* is constitutively expressed and the IRF3 protein is very stable and present in the cytoplasm in an inactive form in basal conditions. Conversely, *Irf7* is an ISG with a short half-life in most cell types. Besides, the expression of *Irf7* is constitutive in a few immune cell types such as plasmacytoid dendritic cells (pDCs) [28].



**Figure 2: Biphasic induction of IFN-I by IRF3 and IRF7.** During the “early phase”, PRR recognition by PAMPs lead to the activation of IRF3, present in the cytosol in basal conditions, and expression of *Ifnb1*. Once secreted, IFN-I bind the IFNAR receptor which activate the STAT proteins and lead to the expression of ISGs including *Irf7*. In the “late phase”, IRF7 is present in the cytosol. PRRs activation induce the expression of the IFN $\alpha$  genes, which in turn enhance the expression of *Irf7*.

The model of complementary induction of IFN-I by IRF3 and IRF7 was established in 2000 by Sato and colleagues. They produced the first *Irf3* knock-out (KO) mouse model by replacing the region containing the transcription initiation site with a pgk-neo cassette. Thus, the IRF3 protein is absent in *Irf3*<sup>-/-</sup> cells. Using *Irf3* deficient mice, they established the following model. IRF3 allows a rapid induction of IFN $\beta$  after infection, which leads to the induction of *Irf7* expression. In a later phase, IRF7 is active and allows positive feedback between IFN $\alpha$  signaling and *Irf7* expression (Figure 2) [6].

IRF3 also possesses non-transcriptional activities. It has been shown that an apoptotic role of IRF3 depends on RLR signaling after recognition of viral dsRNA but is independent of IRF3’s transcriptional activity. This pathway thus named RLR-induced IRF-3-mediated pathway of apoptosis (RIPA) requires the polyubiquitination of two lysines in the IAD of IRF3 by the LUBAC protein complex, which was previously known to promote ubiquitination and to be recruited by the RLR pathway. Once ubiquitinated, IRF3 interacts with BAX and HSP90, and is recruited to the mitochondria by TOM70, which induces the apoptotic pathway through cytochrome C release (Figure 1C) [29–31].

More recently, a third virus-induced function of IRF3 was described. As the inflammatory response is a significant contributor to viral pathogenesis, and an interaction between IRF3 and the actors of the

proinflammatory NF- $\kappa$ B pathway has been suspected, Popli and colleagues investigated the role of IRF3 in NF- $\kappa$ B-mediated inflammation. They described the repression of IRF3-mediated NF- $\kappa$ B activity (RIKA), a pathway which relies on the interaction between IRF3 and NF- $\kappa$ B and prevents NF- $\kappa$ B-mediated inflammation. Indeed, IRF3 interacts with NF- $\kappa$ B p65 unit in the cytosol, through a domain between the amino acids 210 and 222, distinct from those needed for RIPA or transcriptional activities. This interaction prevents NF- $\kappa$ B translocation to the nucleus to induce the expression of proinflammatory genes (Figure 1C) [32].

These non-transcriptional activities of IRF3 are crucial to fight viral infection. Indeed, mice with functional RIPA and RIKA but no IRF3 transcriptional activity were more susceptible to Sendai Virus (SeV) than *Irf3* WT mice but less than *Irf3* KO mice. Moreover, *Irf3* deficient mice are more susceptible to SeV infection than *Ifnar1* deficient mice, which completely lack IFN-I signaling. These results show that the non-transcriptional activities of IRF3 provide complementary antiviral pathways to the IFN-I induction and are required for a full IRF3-mediated protection against viral infection [30].

## Models for studying the role of IRF3

Multiple models have been used to study the role of IRF3 in mouse and human studies. The *Irf3* KO model described by Sato and colleagues is still mostly used nowadays. Mice of this strain were used to study the role of *Irf3* after *in vivo* infections [6, 7, 10], or primary cells were derived from these mice to study cell-specific effects. MEFs have extensively been used due to their easy production and storage [11, 33–35]. Otherwise, immune cell types have also been used such as bone marrow macrophages (BMMs) and bone marrow dendritic cells (BMDCs) [9, 10, 32]. Some studies have used plasmacytoid dendritic cells [7, 36]. However, as these cells constitutively express *Irf7*, they do not allow to study the transcriptional role of IRF3 in the IFN-I response. Other specific cell types have been used depending on the virus tropism. For example, primary cultured neurons (PCN) have been used to study West Nile virus (WNV) [10] and Western equine encephalitis virus (WEEV) infections [37].

Importantly, the *Irf3* KO strain produced by Sato and colleagues also carries a null mutation in the adjacent gene *Bcl2l12*, an antiapoptotic factor that suppresses DNA damage-induced apoptosis [38]. Recently, the same team developed a Flox *Irf3* mouse strain to study the consequences of an *Irf3* KO without *Bcl2l12* mutation. Two loxP sequences were added to flank exons 2 to 4. *Irf3*-floxed mice were then crossed with CAG-Cre mice to obtain *Irf3*<sup>del/del</sup>-CAG-Cre<sup>+</sup> mice and the CAG-Cre transgene was then removed by crossing with WT mice. *Irf3* deficient and *Irf3* *Bcl2l12* double deficient MEFs showed the same reduced induction of IFN-I after infection with EMCV, HSV-1 and VSV, and single and double deficient mice showed same mortality following EMCV infection [6, 7, 38] showing that the *Bcl2l12* mutation was not responsible for these phenotypes.

Apart from KO models, mouse *Irf3* mutants have been studied. The *Irf3*-S1 mutant has functional RIPA and RIKA but no transcriptional activity, while the *Irf3*-M1 mutant has RIKA activity but no RIPA nor transcriptional activity [30, 32].

In human, known polymorphisms in the *Irf3* gene were tested for the association with viral disease [39–41], or new variants were identified by studying the whole exome sequences of patients with severe illness following viral infection [42, 43], or sequences of candidate genes including *Irf3* [44, 45]. *In vitro*, CRISPR/Cas techniques have been used to ablate the *Irf3* gene [32, 46]. Otherwise,



knockdown (KD) techniques such as siRNA and shRNA have been used to decrease the amount of IRF3 protein in cells [32, 35].

## **IRF3 deficiencies lead to decreased IFN-I induction and increased viral replication *in vitro***

Multiple experimental parameters can have consequences on the readout of viral infection in cellular models, such as the multiplicity of infection (MOI), the cell type used, the time points investigated and of course, the type of virus studied. Despite these differences of models, study of *Irf3* deficiencies in cellular models have often identified a role for IRF3 in the induction of IFN-I and the control of viral replication. The major findings are presented in Table 1.

For instance, Moore and colleagues investigated the consequence of an *Irf3* KO on Theiler's encephalomyelitis virus (TMEV) infection of primary macrophages. They first showed that viral loads in cell lysates were higher in *Irf3* KO macrophages than in WT macrophages, showing that *Irf3* is required to control TMEV genome replication. At 3 hours post infection (hpi), *Irf3* deficient macrophages had lower expression of *Ifnb1* and *Il6* compared with WT macrophages, but higher expression at 24 hpi. The authors hypothesized that the higher replication rate in *Irf3* KO macrophages led to IRF3-independent induction of IFN $\beta$  and IL6, for instance through *Irf1* expression [47]. This increased induction of IL6 could also be explained by the absence of RIKK in *Irf3* deficient cells [32].

Similarly, the role of IRF3 on the control of mouse gammaherpes virus (MGHV) infection was investigated. Infected *Irf3* deficient BMMs showed reduced expression of *Ifnb1* and IFN $\alpha$  genes at 4 hpi and reduced levels of IFN-I in the supernatant at 3 and 4 hpi. Moreover, at 8 hpi, which corresponds to the peak expression of the ISGs *Rsad2* and *Mx1* in WT BMMs, their expression was abolished in *Irf3* deficient macrophages [48].

The role of *Irf3* has also been assessed by using siRNA to reduce the amount of mRNA. This method was used to investigate influenza A virus (IAV) infection in primary human macrophages. *Irf3* KD cells showed lower expression of *Ifnb1* and *Ifnl1* at 3 and 6 hpi. The expression of other cytokines, such as *Cxcl10*, *Tnfa*, *Ccl2*, *Ccl3* and *Ccl5* was also reduced in *Irf3* KD cells. Viral replication was not assessed in this study [49].

Daffis and colleagues have investigated WNV infection in *Irf3* deficient BMMs and PCNs. *Irf3* deficient PCNs had reduced IFN-I gene expression and levels of IFN $\alpha$  and IFN $\beta$  in the supernatants compared with WT cells. *Irf7* induction was also decreased in *Irf3* deficient PCN, which may partially explain the low levels of IFN-I. Viral titers were slightly higher in *Irf3* deficient PCNs than in WT cells. Conversely, in BMMs, while the viral titer was higher in *Irf3* deficient cells than in WT cells, *Irf3* deficient BMMs had earlier and higher level of IFN-I gene expression than WT BMMs. This induction was likely driven by IRF7 as its expression was elevated in *Irf3* deficient BMM. These results demonstrate that IRF3-dependent IFN-I induction *in vitro* may be cell-specific. The time point studied might not have enabled to measure a decrease in IRF3-dependent IFN-I induction. Indeed, IFN-I expression was measured at 24 and 48 hpi while other studies in infected *Irf3* KO BMMs described reduced IFN-I induction at earlier time points, between 1 and 16 hpi [32, 38, 48].

**Table 1: Consequences of *Irf3* deficiencies in cellular models of viral infection**

Virus	Order	Species	Cell	Allele	IFN	Viral load	Other	Reference
AdV	Adenoviridae	Human	A549	<i>Irf3</i> KO <sup>3</sup>	=	=	ND	[46]
MHV	Coronaviridae	Mouse	BMM	<i>Irf3</i> KO <sup>1</sup>	ND	↑	↑ (inf)	[32]
WNV	Flaviviridae	Mouse	MEFs	<i>Irf3</i> KO <sup>1</sup>	NA	↑	NA	[34]
		Mouse	BMM	<i>Irf3</i> KO <sup>1</sup>	↑ (IFN $\alpha$ & $\beta$ )	↑	↑ ( <i>Irf7</i> )	[10]
		Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	NA	↑	NA	[34]
		Mouse	PCN	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\alpha$ & $\beta$ )	↑	↓ ( <i>Irf7</i> )	[10]
		Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	↑	ND	[50]
ZIKV		Human	A549	<i>Irf3</i> KO <sup>3</sup>	= (IFN $\alpha$ & $\beta$ )	=	ND	[46]
HSV	Herpesviridae	Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	ND	ND	[7]
		Mouse	MEF	siRNA	↓ (IFN $\beta$ )	ND	ND	[35]
		Mouse	BMDC	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	↑	ND	[9]
		Mouse	BMDC	<i>Irf3</i> KO <sup>1</sup>	ND	↑	ND	[51]
		Mouse	BMM	<i>Irf3</i> KO <sup>1</sup>	ND	↑	ND	[9]
		Mouse	BMM, BMDC, MEF	<i>Irf3</i> KO <sup>2</sup>	↓ (IFN $\beta$ )	ND	ND	[38]
		Mouse	pDC	<i>Irf3</i> KO <sup>1</sup>	= (IFN $\alpha$ )	ND	ND	[7]
MGHV		Mouse	BMM	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	↑	↓ (ISG)	[48]
IAV	Orthomyxoviridae	Human	A549	<i>Irf3</i> KO <sup>3</sup>	= (IFN $\alpha$ & $\beta$ )	↓	ND	[46]
		Human	PrM	siRNA	↓ (IFN $\beta$ )	ND	↓ (ISG, inf)	[49]
		Mouse	BMM	<i>Irf3</i> KO <sup>1</sup>	ND	↑	↑ (inf)	[32]
NDV		Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\alpha$ & $\beta$ )	ND	ND	[6]
SeV	Paramyxoviridae	Human	HT1080	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	↑	↑ (inf)	[32]
		Mouse	BMM	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	↑	↑ (inf)	[32]
RVFV	Phenuiviridae	Human	A549	<i>Irf3</i> KO <sup>3</sup>	↓ (IFN $\beta$ )	↑	↓ ( <i>IRF7</i> )	[46]
EMCV	Picornaviridae	Mouse	BMM, MEF	<i>Irf3</i> KO <sup>2</sup>	↓ (IFN $\beta$ )	ND	ND	[38]
		Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	ND	ND	[7]
TMEV		Mouse	PM	<i>Irf3</i> KO <sup>1</sup>	↓ (IFN $\beta$ )	↑	↓ ( <i>Irf7</i> ), ↑ (inf)	[47]
Myxoma	Poxviridae	Mouse	pDC	<i>Irf3</i> KO <sup>1</sup>	= (IFN $\alpha$ & $\beta$ )	ND	ND	[36]

		Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	↓ (IFNβ)	ND	ND	[7]
			BMDC,					
VSV	Rhabdoviridae	Mouse	BMM, MEF	<i>Irf3</i> KO <sup>2</sup>	↓ (IFNβ)	ND	ND	[38]
		Mouse	pDC	<i>Irf3</i> KO <sup>1</sup>	= (IFNα)	ND	ND	[7]
CHIKV		Mouse	MEF	<i>Irf3</i> KO <sup>1</sup>	↓ (IFNβ)	ND	ND	[11]
WEEV	Togaviridae	Mouse	PCN	<i>Irf3</i> KO <sup>1</sup>	↓ (IFNβ)	=	↓ (cell viability)	[37]
		Human	BE(2)-C	dn	↓ (IFNβ)	↑	↓ (cell viability)	[37]

A549 : adenocarcinomic human alveolar basal epithelial cell line A549, Adv : adenovirus, BE(2)-C : human neuroblastoma cells, BMDC: bone marrow dendritic cells, BMM: bone marrow macrophages, CHIKV: Chikungunya virus, dn : dominant negative allele, IAV: influenza A virus, inf : inflammation, EMCV: encephalomyocarditis virus, HSV: herpes simplex virus, MHV: murine hepatitis virus, MEF: mouse embryonic fibroblasts, MGHV: mouse gammaherpes virus, ND : not determined, NDV: Newcastle disease virus, PCN: primary cultured neurons, pDC: plasmacytoid dendritic cells, PM : peritoneal macrophages, PrM: primary macrophages, RVFV: Rift Valley fever virus, SeV: Sendai virus, TMEV: Theiler's murine encephalomyelitis virus, VSV: vesicular stomatitis virus, WEEV: Western equine encephalitis virus, WNV: West Nile virus, ZIKV : Zika virus, 1: *Irf3*<sup>tm1Ttg</sup>, 2: *Irf3*<sup>tm2.1Ttg</sup>, 3: CRISPR/Cas9 KO *in vitro*

## ***Irf3* deficiencies *in vivo***

Multiple factors are involved in the susceptibility to viral diseases *in vivo*. As for *in vitro* models, the virus studied and the viral dose can affect the infection outcome. Results may also vary depending on the route of infection. For instance, in mice, routes of experimental infection typically include intraperitoneal, subcutaneous, intravenous, and intranasal routes. Viral replication in specific tissues can be assessed by using more specialized routes such as intracranial or corneal routes. Moreover, infection outcome *in vivo* is more complex than in *in vitro* models as it results from interaction between multiple cell types and tissues. Besides, in human, when the outcome of viral infection is studied in patients, these factors are unknown.

Studies in mice have examined both *in vitro* and *in vivo* phenotypes. For instance, Yanai and colleagues have demonstrated that *Irf3* KO MEFs, BMDCs and BMMs had lower *Ifnb1* expression than WT cells after HSV, VSV, and EMCV infection. Susceptibility of *Irf3* deficient mice to EMCV infection was also assessed *in vivo*. After intravenous infection, *Irf3* deficient mice had a higher mortality rate than WT mice, and lower levels of IFNβ in the serum [38].

Similarly, *Irf3* deficiency in BMMs and BMDC led to higher viral titer in the supernatant than WT BMDC after HSV-1 infection. Moreover, infected *Irf3* deficient BMDC had lower levels of IFNβ in the supernatant than WT BMDC [9]. The authors then studied the survival of *Irf3* deficient mice after corneal and intracranial HSV-1 infection. With the two infection routes, *Irf3* KO mice showed a higher mortality rate than WT mice and higher viral titers in the brain. Moreover, *Irf3* KO mice had increased levels of inflammatory cytokines in the brain after intracranial infection, such as TNFα, IL-10 and CCL5 [52].

In other cases, the susceptibility of *Irf3* deficient cells was not predictive of an increased susceptibility *in vivo*. *Irf3* KO MEFs infected with HSV-1 had a decreased induction of IFN-I, especially IFNβ, compared with WT MEFs. However, *Irf3* KO mice and WT mice survived infection and had the same

level of serum IFN $\alpha$  [7]. Similarly, the expression of *Ifnb1* in *Irf3* KO MEFs infected with Chikungunya virus (CHIKV) was decreased compared with WT, while *Irf3* KO and WT mice had the same viral load in the blood and survived the infection [11].

The infection outcome of *Irf3* deficient mice can differ depending on infection parameters. For instance, Honda and colleagues infected intravenously mice with HSV-1 and observed 100% survival and the same levels of serum IFN $\alpha$  in WT and *Irf3* KO mice [7]. On the other hand, Canivet and colleagues intranasally infected mice with HSV-1 and described a 10% survival rate in *Irf3* deficient individuals while 70% of WT mice survived infection. *Irf3* KO individuals also showed increased body weight loss and viral titers as well as quantity of inflammatory cytokines in the brain compared to WT mice. Levels of IFN $\beta$  in brain homogenates were slightly lower at 3 days post infection (dpi) but higher at 5 dpi in *Irf3* KO mice compared with WT mice. The authors suggested the implication of alternative IFN pathways to explain the increased IFN production at 5 dpi [53]. Corneal and intracranial HSV-1 infection also led to higher mortality rates in *Irf3* deficient mice than in WT mice [9].

Viral dose can also impact the infection outcome. After infection with 120,000 plaque forming unit (pfu) of SeV, 10% of WT mice died, while after infection with 35,000 pfu, none of them did. At both doses, *Irf3* deficient mice showed 100% mortality. With the higher dose, *Irf3* mice had 100-fold more virus titer in the lungs, reduced *Ifnb1* mRNA level and apoptosis compared with WT mice [30, 54].

These findings, along with others [6, 7, 10, 38, 47], describe increased susceptibility in *Irf3* KO mice in the cases where WT mice displayed clinical symptoms and mortality, while studies that concluded in the absence of an effect of IRF3 deficiency on viral infection outcome generally reported the absence of mortality in WT mice [7, 11, 50, 55–58]. As a result, it appears that when the infection parameters lead to some pathology in WT mice, the effect is accentuated in *Irf3* KO mice.

In human, polymorphisms in *IRF3* have been found in patients with severe viral disease. Two heterozygous mutations were found in patients with herpes simplex encephalitis (HSE), c.854G>A (R285Q) and c.829G>A (A277T), and were associated with defective IRF3 phosphorylation and reduced induction of *Ifnb1* and *Cxcl10* compared with controls, respectively [42, 44]. A non-coding *Irf3* variant 925A/G was also associated with HSV induced cirrhosis in Egyptian patients, with the AG genotype being protective [40], while this association was not found in a cohort of Chinese patients [41].

A non-coding variant was found in a patient with severe IAV disease. The patient carrying heterozygous c.1576C>T variant had reduced IRF3 protein in peripheral blood mononuclear cells (PBMCs) and reduced IFN-I and IFN-III expression compared with controls [43]. Similarly, two mutations were found in patients with severe COVID19: p.Glu49del/WT and p.Asn146Lys/WT which led to decreased *Ifnb1* induction [45]. Lastly, in the case of WNV infection, the alternative C allele for the rs2304207 SNP in intron 2 of *Irf3* was overrepresented in symptomatic cases compared with asymptomatic cases [39]. In conclusion, *Irf3* mutations were almost always found in patients with severe disease following infection. Thus, in mice like in human studies, *Irf3* deficiencies are detrimental in the case of severe viral infections.

**Table 2: Consequences of *Irf3* deficiencies *in vivo* in mouse models of viral infection**

<b>Virus</b>	<b>Family</b>	<b>Route</b>	<b>IFN</b>	<b>Inflammation</b>	<b>Viral load</b>	<b>Clinical</b>	<b>Reference</b>
MNV	Calciviridae	O	ND	ND	↑ (liver, LN, ileum, spleen)	ND	[59]
DENV		IV	= (serum)	ND	= (kidney, liver, spleen, serum)	= (S)	[55]
		FP	= (serum)	ND	ND	ND	[60]
WNV	Flaviviridae	FP	↑ (serum), ↓ (LN)	ND	↑ (brain, kidney, LN, serum, spleen, SC)	↓ (S)	[10]
ZIKV		FP	ND	ND	ND	= (BWL, S)	[56]
		IP	ND	ND	= (serum)	= (CS, S)	[50]
		IV	= (serum)	ND	ND	= (S)	[7]
HSV	Herpesviridae	IN	Time dependent (brain)	↑ (brain)	↑ (brain)	↓ (S), ↑ (BWL)	[53]
		C	ND	↑ (brain)	= (brain, eye), ↑ (brain stem)	↓ (S)	[52]
		C	ND	= (serum)	= (brain, cornea, serum, trigeminal ganglia) ↑ (brain stem)	↓ (S), = (CS)	[51]
MGHV		IC	ND	↑ (brain)	↑ (brain)	↓ (S)	[52]
		IN	ND	ND	↑ (lung)	ND	[61]

IAV	Orthomyxoviridae	IN	ND	ND	↑ (lung)	= (S, BWL)	[62]
		IN	↓ (BAL)	ND	↑ (lung)	↓ (S, BWL)	[63]
hMPV	Paramyxoviridae	IN	↓ (lung)	ND	↑ (lung)	ND	[64]
SeV		IN	↓ (lung)	ND	↑ (lung)	↓ (S)	[54]
SeV		IN	ND	↑ (lung)	↑ (lung)	ND	[32]
EMCV	Picornaviridae	IP	↓ (serum)	ND	ND	↓ (S)	[6]
		IP	= (serum)	ND	ND	↓ (S)	[7]
		IV	↓ (serum)	ND	ND	↓ (S)	[38]*
TMEV	Picornaviridae	IC	ND	ND	↑	↓ (S), ↑ (BWL)	[47]
		IN	ND	ND	= (brain)	↑ (CS)	[62]
OROV	Peribunyaviridae	FP	ND	ND	ND	= (BWL, S)	[57]
CHIKV	Togaviridae	FP	↓ (foot), = (blood)	= (blood)	= (serum, foot)	= (S)	[58]
		ID	ND	ND	=	= (S)	[11]
SINV	Togaviridae	IC	↓ (brain)	↑ (brain, spinal cord)	= (brain)	= (S), ↑ (CS, BWL)	[65]

C: corneal, BAL: Bronchoalveolar lavage, BWL: body weight loss, CHIKV: Chikungunya virus, CS: clinical symptoms, DENV dengue virus, EMCV: encephalomyocarditis virus, FP: footpad, hMPV: human metapneumovirus, HSV: herpes simplex virus, IAV: influenza A virus, IC: intracranial, ID: intradermal, IN: intranasal IP: intraperitoneal, IV: intravenous, LN: lymph nodes, MGHV: mouse gammaherpes virus, O: oral, OROV: Oropouche virus, S: survival, SeV: Sendai virus, SINV: Sinbis virus, SP: spinal cord, TMEV: Theiler's murine encephalomyelitis virus, WNV: West Nile virus, ZIKV: Zika virus, All studies used *Irf3<sup>tm1Tg</sup>* mice [6], except \* which used *Irf3<sup>tm2.1Tg</sup>* [38].

## Extending the phenotyping of *Irf3* deficient mice

As *Irf3* KO mice may not show increased mortality rate, viral loads, or decreased IFN-I titers, it may be necessary to analyze other phenotypes to detect the effect of an *Irf3* deficiency on the susceptibility of infected mice. For instance, considering the RIKA function of IRF3, inflammation can be assessed in *Irf3* deficient mice. Several studies have found increased inflammation in *Irf3* KO mice after viral infection (Table 2). For instance, after infection with Sinbis virus (SINV), *Irf3* deficient mice had higher expression of *Ccl2*, *Il1b*, *Tnfa* and *Il10* at 5 dpi in the brain and/or the spinal cord than WT mice. Moreover, the histological inflammation score of *Irf3* KO mice was higher than WT mice [65]. Similarly, *Irf3* deficient mice showed higher levels of IL1 $\alpha$ , IL1 $\beta$ , IL6, IL12p40, IL12p70 and IFN $\gamma$  at 5 dpi after HSV-1 infection than WT mice [53]. Whether these effects are directly due to the absence of RIKA or indirectly caused by the lack of IRF3-dependent IFN-I expression was not established.

The effects of *Irf3* deficiencies could also be observed on more indirect targets of the IFN-I. IRF3 has also been shown to have a role in adaptive immune responses. Indeed, IRF3-mediated IFN-I

signaling has been shown to mediate T cell differentiation in the context of infections [66]. In basal conditions, *Irf3* KO mice have similar lymphocyte population composition to WT mice [6]. Therefore, it may be valuable to investigate lymphocyte differentiation and function in *Irf3* deficient mice following viral infection. For instance, *Irf3* deficient mice show increased infiltration of CD4 T cells and CD19 B cells in the brain after SINV infection [65]. Besides, after TMEV and IAV infection, *Irf3* deficient mice have reduced proportion of CD8 T cells expressing granzyme B, an enzyme involved in the cytotoxicity of CD8 T cells, and impaired development of memory T cell population [62]. Together, these studies demonstrate that the consequences of *Irf3* deficiencies may only be visible by extending the phenotyping of *Irf3* deficient mice beyond the IFN-I response.

## Complementary roles of IRF3 and IRF7

Due to their similar functions, the roles of IRF3 and IRF7 have often been compared [7, 11, 59, 63]. As discussed previously, *Irf7* deficiency were often reported as more severe than *Irf3* deficiency [7, 58, 59, 65]. *Irf3 Irf7* double deficiencies have also been studied and showed that *Irf3* deficiency potentiates the susceptibility observed in *Irf7* single KO mice. For instance, after infection with SINV, *Irf7* deficient mice showed a mortality rate of 100%, while *Irf3* deficient and WT mice all survive. *Irf3 Irf7* double deficient mice have a mortality rate of 100% like *Irf7* KO mice, and die faster [65]. Similarly, after IAV infection, all WT mice survive, while *Irf3* and *Irf7* deficient mice show a mortality rate of 40% and 60%, respectively. *Irf3 Irf7* double deficient mice show a mortality rate of 80% [63].

In other cases, the double deficient model displays a phenotype that was not observed in either single KO strain. For instance, after DENV infection, *Irf3 Irf7* KO mice have increased viral loads in the serum, spleen, liver and kidney, compared with single KO and WT individuals [55]. Other studies have reported that *Irf3 Irf7* double deficient mice succumb to viral infection, while WT, *Irf3* KO and *Irf7* KO mice survive, e.g., after infection with CHIKV or Oropouche virus (OROV) [57, 58].

IRF3 and IRF7 can show complementary roles against viral infection, which leads to different disease signs in *Irf3* deficient mice and in *Irf7* deficient mice. After IAV infection, *Irf3* KO mice had increased viral titer in the lung while *Irf7* KO mice had lung viral titer similar to WT mice. IFN $\alpha$  levels in the bronchoalveolar lavage were more altered in *Irf7* KO mice, which showed abolished induction at 1 and 2 dpi, while *Irf3* KO mice had lower levels than WT mice at 1 dpi and similar levels at 2 dpi. On the opposite, IFN $\beta$  levels were more altered in *Irf3* KO mice which showed lower levels than WT mice at 1 and 2 dpi, while *Irf7* KO mice had reduced levels only at 1 dpi [63]. These results are consistent with the known affinity of the IRFs for the IFN-I gene promoters [24]. Interestingly, IFN-I levels were lower in *Irf7* KO mice than in WT at 1 and 2 dpi, but *Irf3* KO mice had a four-fold increased levels of IFN-I at 2 dpi, suggesting a compensatory mechanism for the reduced levels of IFN-I in *Irf3* deficient mice [63].

*Irf3* and *Irf7* deficiencies may also have different effects depending on the organ. Murine norovirus (MNV) infected *Irf3* deficient mice and *Irf7* deficient mice have increased viral titer in the lung, spleen and lymph node. In the ileum, only *Irf3* deficient mice have an increased viral titer compared to WT mice [59]. Besides, in the case of CHIKV infection, the redundancy between *Irf3* and *Irf7* is age dependent as adult mice lacking either one survives infection while 100% of neonates (9 days old) succumbed. However, in WT mice, the infection outcome also differed between adults and neonates, as adults all survived while about 50% of neonates succumbed [11].

These studies demonstrate *Irf3* and *Irf7* possess both redundant and proper functions to fight viral infection. In some instances, infected *Irf3* deficient mice show disease signs that has not been observed in *Irf7* deficient mice. Besides, *Irf3 Irf7* double deficient mice can show increased susceptibility compared with single deficient mice, or new disease signs that have not been observed in single deficient mice.

## Discussion

IRF3 was extensively studied for its transcriptional activity after its discovery in 1995 [5], notably in the case of viral infections. IRF3 has often been presented in opposition with the other main IFN-I transcription factor, IRF7, but IRF3 possesses other non-transcriptional activities, RIPA and RIKA, extending its range of action against viral infections beyond interferon signaling.

In this review, we have highlighted the duality between the effect of IRF3 deficiency in cellular models and in whole organisms. Several studies reported altered IFN-I signaling and increased viral replication in *Irf3* deficient cellular models, while *Irf3* KO mice showed the same susceptibility to viral infection as WT mice [7, 11, 50]. In other cases, *Irf3* deficiencies were only studied in *in vivo* models, and found no impact of *Irf3* deficiency on the susceptibility to viral diseases, for example to DENV and OROV infections [55, 57]. Thus, it would be worth investigating whether *Irf3* deficient cells infected with these viruses would show a higher susceptibility than WT cells.

The absence of correlation between the results obtained in *in vitro* and *in vivo* studies in mouse models raises the question of the translation of the results obtained *in vitro* in human studies. We have seen several cases of human *Irf3* mutations identified in patients with severe viral disease by sequencing. The *Irf3* loss-of-function is then usually confirmed by functional assay in patients' cells [42–45]. Based on the absence of correlation between *in vitro* phenotype and disease in *Irf3* deficient models in the mouse, it is possible that the altered IFN-I induction in patients' cells may not fully explain their susceptibility to viral infection.

Study of the literature using *Irf3* deficient models emphasized the importance of identifying the proper phenotype to study the role of *Irf3*. First, we have seen that *Irf3* deficiencies lead to enhanced susceptibility, i.e., disease signs and mortality, in KO mice compared with WT mice if the experimental model allowed to observe disease signs in WT mice [7, 30, 53]. Thus, a certain infection route, viral strain or dose may be required to see more severe illness in *Irf3* deficient mice than in WT mice. Otherwise, more subtle phenotyping, such as viral loads in multiple organs, presence of inflammation, and adaptive immune response can be assessed to decipher the consequences of an *Irf3* deficiency [32, 62, 63]. In fact, the discovery of the new functions of IRF3, RIPA and RIKA, demonstrated that this protein is involved in other anti-viral pathways than IFN-I induction. Therefore, future studies may leverage this new knowledge on the diverse functions of IRF3 to give a comprehensive view of its role to face viral infections.



## References

1. **Levy DE, Marie IJ, Durbin JE.** Induction and function of type I and III interferon in response to viral infection. *Curr Opin Virol* 2011;1:476–86.
2. **Li SF, Gong MJ, Zhao FR, Shao JJ, Xie YL, et al.** Type I Interferons: Distinct Biological Activities and Current Applications for Viral Infection. *Cell Physiol Biochem* 2018;51:2377–2396.
3. **Fensterl V, Sen GC.** Interferons and viral infections. *Biofactors* 2009;35:14–20.
4. **Ozato K, Tailor P, Kubota T.** The Interferon Regulatory Factor Family in Host Defense: Mechanism of Action. *Journal of Biological Chemistry* 2007;282:20065–20069.
5. **Au WC, Moore PA, Lowther W, Juang YT, Pitha PM.** Identification of a member of the interferon regulatory factor family that binds to the interferon-stimulated response element and activates expression of interferon-induced genes. *Proc Natl Acad Sci USA* 1995;92:11657–11661.
6. **Sato M, Suemori H, Hata N, Asagiri M, Ogasawara K, et al.** Distinct and essential roles of transcription factors IRF-3 and IRF-7 in response to viruses for IFN- $\alpha$ / $\beta$  gene induction. *Immunity* 2000;13:539–48.
7. **Honda K, Yanai H, Negishi H, Asagiri M, Sato M, et al.** IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* 2005;434:772–7.
8. **Daffis S, Lazear HM, Liu WJ, Audsley M, Engle M, et al.** The Naturally Attenuated Kunjin Strain of West Nile Virus Shows Enhanced Sensitivity to the Host Type I Interferon Response. *J Virol* 2011;85:5664–5668.
9. **Menachery VD, Leib DA.** Control of Herpes Simplex Virus Replication Is Mediated through an Interferon Regulatory Factor 3-Dependent Pathway. *J Virol* 2009;83:12399–12406.
10. **Daffis S, Samuel MA, Keller BC, Gale M Jr, Diamond MS.** Cell-specific IRF-3 responses protect against West Nile virus infection by interferon-dependent and -independent mechanisms. *PLoS Pathog* 2007;3:e106.
11. **Schilte C, Buckwalter MR, Laird ME, Diamond MS, Schwartz O, et al.** Cutting Edge: Independent Roles for IRF-3 and IRF-7 in Hematopoietic and Nonhematopoietic Cells during Host Response to Chikungunya Infection. *J* 2012;188:2967–2971.
12. **Grandvaux N, Servant MJ, tenOever B, Sen GC, Balachandran S, et al.** Transcriptional Profiling of Interferon Regulatory Factor 3 Target Genes: Direct Involvement in the Regulation of Interferon-Stimulated Genes. *J Virol* 2002;76:5532–5539.
13. **Nakaya T, Sato M, Hata N, Asagiri M, Suemori H, et al.** Gene Induction Pathways Mediated by Distinct IRFs during Viral Infection. *Biochemical and Biophysical Research Communications* 2001;283:1150–1156.
14. **Navarro L, Mowen K, Rodems S, Weaver B, Reich N, et al.** Cytomegalovirus Activates Interferon Immediate-Early Response Gene Expression and an Interferon Regulatory Factor 3-Containing Interferon-Stimulated Response Element-Binding Complex. *Mol Cell Biol* 1998;18:3796–3802.
15. **Weaver BK, Ando O, Kumar KP, Reich NC.** Apoptosis is promoted by the dsRNA-activated factor (DRAF1) during viral infection independent of the action of interferon or p53. *FASEB j* 2001;15:501–515.
16. **Génin P, Algarté M, Roof P, Lin R, Hiscott J.** Regulation of RANTES Chemokine Gene Expression Requires Cooperativity Between NF- $\kappa$ B and IFN-Regulatory Factor Transcription Factors. *The Journal of Immunology* 2000;164:5352–5361.
17. **Lin R, Heylbroeck C, Genin P, Pitha PM, Hiscott J.** Essential Role of Interferon Regulatory Factor 3 in Direct Activation of RANTES Chemokine Transcription. *Mol Cell Biol* 1999;19:959–966.
18. **Liu Y, Mo C-F, Luo X-Y, Li H, Guo H-J, et al.** Activation of Toll-Like Receptor 3 Induces Interleukin-1 Receptor Antagonist Expression by Activating the Interferon Regulatory Factor 3. *J Innate Immun* 2020;12:304–320.
19. **Goubau D, Romieu-Mourez R, Solis M, Hernandez E, Mesplède T, et al.** Transcriptional re-programming of primary macrophages reveals distinct apoptotic and anti-tumoral functions of IRF-3 and IRF-7. *Eur J Immunol* 2009;39:527–540.
20. **Kirshner JR, Karpova AY, Kops M, Howley PM.** Identification of TRAIL as an Interferon Regulatory Factor 3 Transcriptional Target. *J Virol* 2005;79:9320–9324.
21. **Peters K, Chattopadhyay S, Sen GC.** IRF-3 Activation by Sendai Virus Infection Is Required for Cellular Apoptosis and Avoidance of Persistence. *J Virol* 2008;82:3500–3508.
22. **Antonczyk A, Krist B, Sajek M, Michalska A, Piaszyk-Borychowska A, et al.** Direct Inhibition of IRF-Dependent Transcriptional Regulatory Mechanisms Associated With Disease. *Front Immunol* 2019;10:1176.
23. **Schwanke H, Stempel M, Brinkmann MM.** Of Keeping and Tipping the Balance: Host Regulation and Viral Modulation of IRF3-Dependent IFN $\beta$ 1 Expression. *Viruses* 2020;12:733.
24. **Marie I.** Differential viral induction of distinct interferon- $\alpha$  genes by positive feedback through interferon regulatory factor-7. *The EMBO Journal* 1998;17:6660–6669.
25. **Lazear HM, Schoggins JW, Diamond MS.** Shared and Distinct Functions of Type I and Type III Interferons. *Immunity* 2019;50:907–923.

26. **Donnelly RP, Kotenko SV.** Interferon-Lambda: A New Addition to an Old Family. *Journal of Interferon & Cytokine Research* 2010;30:555–564.
27. **Zhou JH, Wang YN, Chang QY, Ma P, Hu Y, et al.** Type III Interferons in Viral Infection and Antiviral Immunity. *Cell Physiol Biochem* 2018;51:173–185.
28. **Ning S, Pagano JS, Barber GN.** IRF7: activation, regulation, modification and function. *Genes Immun* 2011;12:399–414.
29. **Chattopadhyay S, Marques JT, Yamashita M, Peters KL, Smith K, et al.** Viral apoptosis is induced by IRF-3-mediated activation of Bax. *EMBO J* 2010;29:1762–1773.
30. **Chattopadhyay S, Kuzmanovic T, Zhang Y, Wetzel JL, Sen GC.** Ubiquitination of the Transcription Factor IRF-3 Activates RIPA, the Apoptotic Pathway that Protects Mice from Viral Pathogenesis. *Immunity* 2016;44:1151–1161.
31. **Wei B, Cui Y, Huang Y, Liu H, Li L, et al.** Tom70 Mediates Sendai Virus-Induced Apoptosis on Mitochondria. *J Virol* 2015;89:3804–3818.
32. **Popli S, Chakravarty S, Fan S, Glanz A, Aras S, et al.** IRF3 inhibits nuclear translocation of NF- $\kappa$ B to prevent viral inflammation. *Proc Natl Acad Sci USA* 2022;119:e2121385119.
33. **Tan YS, Lei YL.** Generation and Culture of Mouse Embryonic Fibroblasts. In: *Mouse Models of Innate Immunity*. New York, NY: Springer New York. pp. 85–91.
34. **Fredericksen BL, Smith M, Katze MG, Shi P-Y, Gale M.** The Host Response to West Nile Virus Infection Limits Viral Spread through the Activation of the Interferon Regulatory Factor 3 Pathway. *J Virol* 2004;78:7737–7747.
35. **Menasria R, Boivin N, Lebel M, Piret J, Gosselin J, et al.** Both TRIF and IPS-1 Adaptor Proteins Contribute to the Cerebral Innate Immune Response against Herpes Simplex Virus 1 Infection. *J Virol* 2013;87:7301–7308.
36. **Dai P, Cao H, Merghoub T, Avogadri F, Wang W, et al.** Myxoma Virus Induces Type I Interferon Production in Murine Plasmacytoid Dendritic Cells via a TLR9/MyD88-, IRF5/IRF7-, and IFNAR-Dependent Pathway. *J Virol* 2011;85:10814–10825.
37. **Peltier DC, Lazear HM, Farmer JR, Diamond MS, Miller DJ.** Neurotropic Arboviruses Induce Interferon Regulatory Factor 3-Mediated Neuronal Responses That Are Cytoprotective, Interferon Independent, and Inhibited by Western Equine Encephalitis Virus Capsid. *J Virol* 2013;87:1821–1833.
38. **Yanai H, Chiba S, Hangai S, Kometani K, Inoue A, et al.** Revisiting the role of IRF3 in inflammation and immunity by conditional and specifically targeted gene ablation in mice. *PNAS* 2018;115:5253–5258.
39. **Bigham AW, Buckingham KJ, Husain S, Emond MJ, Bofferding KM, et al.** Host genetic risk factors for West Nile virus infection and disease progression. *PLoS One* 2011;6:e24745.
40. **Talaat RM, Elsayed SS, Abdel-Hakem NE, El-Shenawy SZ.** Genetic Polymorphism in Toll-Like Receptor 3 and Interferon Regulatory Factor 3 in Hepatitis C Virus-Infected Patients: Correlation with Liver Cirrhosis. *Viral Immunology* 2022;35:609–615.
41. **Yan F, Gao Y-F, Lv F, Zhang T-C, Li X, et al.** No association between IRF3 polymorphism and susceptibility to hepatitis B virus infection in Chinese patients. *World J Gastroenterol* 2012;18:388–392.
42. **Andersen LL, Mork N, Reinert LS, Kofod-Olsen E, Narita R, et al.** Functional IRF3 deficiency in a patient with herpes simplex encephalitis. *J Exp Med* 2015;212:1371–9.
43. **Thomsen MM, Jørgensen SE, Storgaard M, Kristensen LS, Gjedsted J, et al.** Identification of an *IRF3* variant and defective antiviral interferon responses in a patient with severe influenza. *Eur J Immunol* 2019;49:2111–2114.
44. **Mork N, Kofod-Olsen E, Sorensen KB, Bach E, Orntoft TF, et al.** Mutations in the TLR3 signaling pathway and beyond in adult patients with herpes simplex encephalitis. *Genes Immun* 2015;16:552–66.
45. **Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, et al.** Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. Epub ahead of print 24 September 2020. DOI: 10.1126/science.abd4570.
46. **Wüst S, Schad P, Burkart S, Binder M.** Comparative Analysis of Six IRF Family Members in Alveolar Epithelial Cell-Intrinsic Antiviral Responses. *Cells* 2021;10:2600.
47. **Moore TC, Cody L, Kumm PM, Brown DM, Petro TM.** IRF3 helps control acute TMEV infection through IL-6 expression but contributes to acute hippocampus damage following TMEV infection. *Virus Research* 2013;178:226–233.
48. **Wood BM, Mboko WP, Mounce BC, Tarakanova VL.** Mouse gammaherpesvirus-68 infection acts as a rheostat to set the level of type I interferon signaling in primary macrophages. *Virology* 2013;443:123–133.
49. **Hui KPY, Lee SMY, Cheung C, Ng IHY, Poon LLM, et al.** Induction of Proinflammatory Cytokines in Primary Human Macrophages by Influenza A Virus (H5N1) Is Selectively Regulated by IFN Regulatory Factor 3 and p38 MAPK. *The Journal of Immunology* 2009;182:1088–1098.
50. **Bourdon M, Manet C, Conquet L, Ramaugé-Parra C, Kornobis E, et al.** Susceptibility to Zika virus in a Collaborative Cross mouse strain is induced by *Irf3* deficiency in vitro but requires other variants in vivo, not linked to the type I interferon response. Preprint; Microbiology. Epub ahead of print 30 March 2023. DOI: 10.1101/2023.03.27.534491.

51. **Murphy AA, Rosato PC, Parker ZM, Khalenkov A, Leib DA.** Synergistic control of herpes simplex virus pathogenesis by IRF-3, and IRF-7 revealed through non-invasive bioluminescence imaging. *Virology* 2013;444:71–79.
52. **Menachery VD, Pasieka TJ, Leib DA.** Interferon Regulatory Factor 3-Dependent Pathways Are Critical for Control of Herpes Simplex Virus Type 1 Central Nervous System Infection. *J Virol* 2010;84:9685–9694.
53. **Canivet C, Rhéaume C, Lebel M, Piret J, Gosselin J, et al.** Both IRF3 and especially IRF7 play a key role to orchestrate an effective cerebral inflammatory response in a mouse model of herpes simplex virus encephalitis. *J Neurovirol* 2018;24:761–768.
54. **Chattopadhyay S, Fensterl V, Zhang Y, Velepparambil M, Wetzel JL, et al.** Inhibition of Viral Pathogenesis and Promotion of the Septic Shock Response to Bacterial Infection by IRF-3 Are Regulated by the Acetylation and Phosphorylation of Its Coactivators. *mBio* 2013;4:e00636-12.
55. **Chen HW, King K, Tu J, Sanchez M, Luster AD, et al.** The roles of IRF-3 and IRF-7 in innate antiviral immunity against dengue virus. *J Immunol* 2013;191:4194–201.
56. **Lazear HM, Govero J, Smith AM, Platt DJ, Fernandez E, et al.** A Mouse Model of Zika Virus Pathogenesis. *Cell Host Microbe* 2016;19:720–730.
57. **Proenca-Modena JL, Sesti-Costa R, Pinto AK, Richner JM, Lazear HM, et al.** Oropouche Virus Infection and Pathogenesis Are Restricted by MAVS, IRF-3, IRF-7, and Type I Interferon Signaling Pathways in Nonmyeloid Cells. *J Virol* 2015;89:4720–4737.
58. **Rudd PA, Wilson J, Gardner J, Larcher T, Babarit C, et al.** Interferon Response Factors 3 and 7 Protect against Chikungunya Virus Hemorrhagic Fever and Shock. *J Virol* 2012;86:9888–9898.
59. **Thackray LB, Duan E, Lazear HM, Kambal A, Schreiber RD, et al.** Critical Role for Interferon Regulatory Factor 3 (IRF-3) and IRF-7 in Type I Interferon-Mediated Control of Murine Norovirus Replication. *J Virol* 2012;86:13515–13523.
60. **Bourne N, Scholle F, Silva MC, Rossi SL, Dewsbury N, et al.** Early Production of Type I Interferon during West Nile Virus Infection: Role for Lymphoid Tissues in IRF3-Independent Interferon Production. *J Virol* 2007;81:9100–9108.
61. **Mboko WP, Rekow MM, Ledwith MP, Lange PT, Schmitz KE, et al.** Interferon Regulatory Factor 1 and Type I Interferon Cooperate To Control Acute Gammaherpesvirus Infection. *J Virol* 2017;91:e01444-16.
62. **Moore TC, Vogel AJ, Petro TM, Brown DM.** IRF3 deficiency impacts granzyme B expression and maintenance of memory T cell function in response to viral infection. *Microbes and Infection* 2015;17:426–439.
63. **Hatesuer B, Hoang HT, Riese P, Trittel S, Gerhauser I, et al.** Deletion of Irf3 and Irf7 Genes in Mice Results in Altered Interferon Pathway Activation and Granulocyte-Dominated Inflammatory Responses to Influenza A Infection. *J Innate Immun* 2017;9:145–161.
64. **Spann KM, Loh Z, Lynch JP, Ullah A, Zhang V, et al.** IRF-3, IRF-7, and IPS-1 Promote Host Defense against Acute Human Metapneumovirus Infection in Neonatal Mice. *The American Journal of Pathology* 2014;184:1795–1806.
65. **Schultz KLW, Troisi EM, Baxter VK, Glowinski R, Griffin DE.** Interferon regulatory factors 3 and 7 have distinct roles in the pathogenesis of alphavirus encephalomyelitis. *J Gen Virol* 2019;100:46–62.
66. **Ysebrant de Lendonck L, Martinet V, Goriely S.** Interferon regulatory factor 3 in adaptive immune responses. *Cell Mol Life Sci* 2014;71:3873–3883.

## II.2 . Identification of susceptibility genes in humans

I have previously emphasized the importance of the IFN-I induction pathway for fighting viral infections and shown that the ablation of one or several of its genes can adversely affect infection outcome. However, host response to viral infection involves various biological processes beyond innate immunity. Since these biological processes are very diverse, involving genes which function may not be suspected, the identification of these genes relies on hypothesis-free methods.

As a first step, candidate gene approaches can be used to evaluate the impact of natural variants in genes which are known or suspected to contribute to the host response to pathogens. Candidate gene approaches are hypothesis-driven methods since they test the association of one or a few specific variant(s) with a trait, as opposed to genome wide studies. Candidate gene studies compare the frequency of a gene's alleles in patients and controls (susceptible and resistant individuals) to identify alleles overrepresented in affected individuals (Kwon and Goate 2000). These methods provided useful information, as exemplified by variants found in patient with severe COVID-19 (Zhang et al. 2020), and can lead to significant findings with relatively small-sized cohorts.

Searching for gene with unexpected functions require other genome-wide approaches, such as whole exome sequencing (WES) and genome wide association studies (GWAS), which make no assumption on the nature of the genes modulating susceptibility to a viral infection. These hypothesis-free studies provide the opportunity to find unexpected genes, as well as already known genes.

### II.2.1 . Human hypothesis-free genetics methods

#### II.2.1.1 . Sequencing approaches

Disease causing variants showing a Mendelian inheritance are usually rare and include deleterious mutations such as insertions, deletions, nonsense variants or splice variants. They can be detected thanks to sequencing methods, by comparing the variants found in patients with human reference genome sequences from control individuals. One common method in the WES, which focuses only on coding regions and which cost is therefore lower than a whole genome sequencing (Majewski et al. 2011; Rabbani et al. 2014). For instance, WES was performed on patients suffering from recurrent lymphocytic meningitis, a disease characterized by recurrent episodes of fever and meningeal irritation caused by herpes simplex virus 2 (HSV-2). Several rare variants were identified in patients, in genes involved in ubiquitin-proteasome (*RNF126*, *ZNRF2*, *HACE1*, among others) and cell fate (*KMT2D*, *HSPA9*, *ERCC2*, among others) pathways (Hait et al. 2021). However, WES by itself does not establish the causal role of a variant on a disease, which must be demonstrated using functional studies.

### **II.2.1.2 . Genome Wide Association Studies**

GWAS methods assess the association of thousands to millions of genetic variants distributed across the genome with disease or traits. In the case of viral infections, GWAS usually compare patients and uninfected controls. Contrary to sequencing approaches, GWAS investigate the role of common variants with moderate effects to identify overrepresented alleles in affected individuals. These methods are similar to the candidate gene approach detailed previously, but for a large number of SNPs which can be in coding or non-coding sequences (Uffelmann et al. 2021). GWAS is a statistical method which, like WES, does not establish the functional impact of the variants tested. Therefore, functional experiments are required to confirm causality and identify the affected pathways (Uffelmann et al. 2021).

For example, a GWAS of the COVID-19 Human Genetics Initiative identified 23 gene loci associated with critical COVID-19, by comparing patients in intensive care unit with controls. Among them, associated loci were found in the IFN-I pathway (*IFNAR2*) but also genes involved in inflammation (*TYK2* and *CCR2*), mRNA metabolism (*CChCR1*) and protein trafficking (*LZTFL1*) (The GenOMICC Investigators et al. 2021).

### **II.2.1.3 . *In vitro* human genetic studies**

Genetic analyses in humans can also be performed *in vitro*. In CRISPR screens, a pool of cells is transfected with a library of guide RNAs so that each cell receives different combinations of RNAs, leading to the disruption of several genes in presence of the Cas9. In the case of viral infections, cells are infected and cultivated until most of them undergo virus-induced cell-death. Then, remaining cells are sequenced to identify gene KOs enriched in resistant cells (Bock et al. 2022). For instance, a CRISPR screen on DENV-infected Huh7 cells (epithelial-like cells) found that KOs in genes involved in endoplasmic reticulum (ER) translocation, ER-associated protein degradation, and oligosaccharyl transfer were enriched in DENV-resistant cells (Marceau et al. 2016).

Another method to study genetic factors in human *in vitro* is the use of cells from different human donors to perform GWAS on cellular readouts. This method was used to study cellular susceptibility to IAV. EBV-immortalized B cells from donors of European, African, and Asian origin were pooled, infected, and processed by single-cell RNA sequencing (scRNA-seq), which allowed to quantify viral genome copies per cell, and identified variants in the coding regions. A variant in the *ERAP1* gene was associated with viral burden in B cells. This gene encodes an aminopeptidase which trims peptides for antigen presentation in MHC. Functional analyses demonstrated that decrease in *ERAP1* expression led to decrease in the percentage of IAV-positive cells (Schott et al. 2022).

## II.2.2 . IFN-I and interferon-independent genes in hypothesis-free methods

Given the major role of IFN-I in the innate antiviral response, one could expect that GWAS analyzing susceptibility to viral infections would often point at genes of this pathway. However, only 4 of them have been associated with viral diseases in GWAS studies (according to the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>), accessed in June 2023). *TICAM2* was associated with chronic sinus infection (Tian et al. 2017), and *DDX58*, *IKBKE* and *IFNA10* were associated with COVID-19 (Kousathanas et al. 2022; Słomian et al. 2023).

The lack of association with known players of the IFN-I pathway and viral diseases could be explained by the rarity of natural variants with detectable impact in these genes due to natural selection, or by the underexposure of individuals with mutations to the pathogen of interest. SNPs in these genes were associated with other immune traits. For instance, *IRF7* and *TRAF3* polymorphisms were associated with systemic lupus erythematosus (Morris et al. 2016; Wang et al. 2021), while SNPs in *TICAM1* and *IRF3* were associated with vitiligo (Jin et al. 2012; Jin et al. 2016).

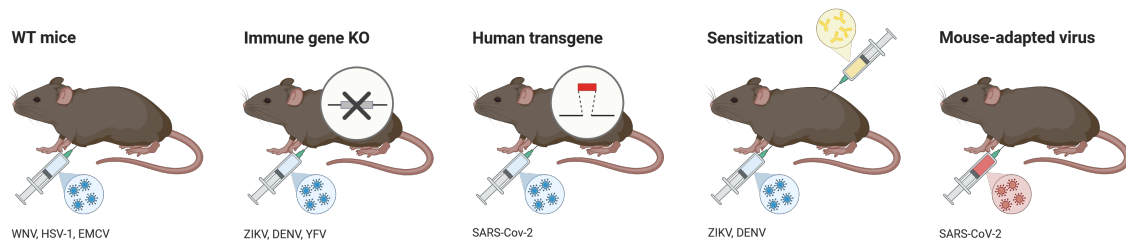
Conversely, as exemplified above, human genetic studies found variants in genes outside the innate immune pathways associated with the susceptibility to viral disease. These genes were involved in other biological processes, such as mRNA metabolism (Hait et al. 2021) protein maturation, trafficking and degradation (Marceau et al. 2016; Hait et al. 2021), and cell fate (Hait et al. 2021). Indeed, viral molecules need to interact with host proteins involved in multiple biological processes, such as transcription, translation, and protein trafficking to replicate their genome and produce new virions (Zhang et al. 2022; Zhou et al. 2023).

In conclusion, while the role of the genes of the IFN-I induction pathway has been clearly demonstrated with reverse genetic approaches, other genes whose role in the host response to pathogens has yet to be determined can be involved in viral disease susceptibility. Hypothesis-free approaches are essential to identify them and to gain more insight into the host genetic control of viral diseases.

## II.3 . Mouse genetics of susceptibility to infectious diseases

### II.3.1 . Experimental models for studying infectious diseases in mice

Mouse models have many advantages in biological studies: they are relatively inexpensive, the housing is easy, including under biosafety confinement, and their short generation time facilitates mouse crosses. Sex, age, and environment can be homogenized within and between experiments, and, in the case of viral infection, viral strain, dose, and infection mode are strictly controlled. Lastly, the existence of inbred strains, for which all individuals within a strain are genetically identical, except for the sex chromosomes, allow performing experiments on identical animals in different laboratories and through time (Sarkar and Heise 2019).



**Figure 3:** Models to study viral infections in the mouse.

Different mouse models exist to study viral infections. WT mice can be used to study viruses that naturally replicate in mice. Otherwise, immunodeficient models, models carrying a human transgene and sensitized mice can be used to study viruses that do not replicate in mice. Alternatively, mouse-adapted viruses can be developed to infect WT mice.

Several types of models are available to study viral infection in the mouse (Figure 3). Standard laboratory mice can be used to study murine and human viruses that are able to replicate in the mouse. For example, wild type (WT) mice can be infected with murine norovirus (MNV) (Thackray et al. 2012), WNV (Daffis et al. 2007), HSV-1 (Canivet et al. 2018) or encephalomyocarditis virus (EMCV) (Yanai et al. 2018).

Genetically modified mouse models may be required to study viruses that do not replicate in the mouse (Sarkar and Heise 2019), depending on the mechanisms of resistance. Some viruses are able to counteract the IFN-I response due to interactions between viral and host proteins. Their interaction may be effective in human cells but not in mouse cells. The use of mouse strains deficient for the *Irfar1* gene can overcome this limitation. *Irfar1* KO mice are used to study flaviviruses such as ZIKV, DENV and YFV (Lazear et al. 2016; Orozco et al. 2012; Marín-Lopez et al. 2019), as well as other viruses such as Chikungunya virus (Schilte et al. 2012). Mice deficient for *Stat1*, which transduces the signal from IFNAR, have also been used to study DENV infection (Krishnakumar et al. 2019). Alternatively, WT mice can be "sensitized" to allow viral replication. In particular, IFNAR1 blocking antibody (Sheehan et al. 2006) can be used in WT mice to block the IFN-I response, allowing ZIKV (Lazear et al. 2016) and DENV replication (Roth et al. 2019; Wilken et al. 2023).

In other cases, the mouse is refractory to viral infection due to the absence of the cellular receptor to the virus (polio, measles) or to reduced affinity (SARS-CoV-2). In such cases, a transgene for the human receptor can confer susceptibility in mice. For example, SARS-CoV-2 spike protein binds the ACE2 receptor to enter host cells. However, the spike protein of SARS-CoV-2 original strain has low affinity for the murine ACE2 receptor, making mice poorly susceptible to SARS-CoV-2 infection. Mice transgenic for human *ACE2* have been used to study SARS-CoV (McCray et al. 2007) and SARS-CoV-2 (Dong et al. 2022). Viral vectors can also be used to induce the transient expression of the receptor. For instance, WT mice sensitized with an adenovirus expressing human ACE2 protein show body weight loss after SARS-CoV-2 infection (Sun et al. 2020).

The last strategy is to adapt the virus to the mouse. Mouse-adapted viruses were produced to increase the affinity of a human virus to mouse homologous receptors or to reduce its capacity to escape the mouse immune system. For instance, SARS-CoV-2 have been adapted to the mouse by genetic engineering (Dinnon et al. 2020), and/or serial passages (Leist et al. 2020; Montagutelli et al. 2021a) where lungs of infected WT mice were collected and used to infect naive mice, for several passages. The resulting mouse adapted viral strains carry specific mutations in the spike protein that allow binding to mouse ACE2 receptor.

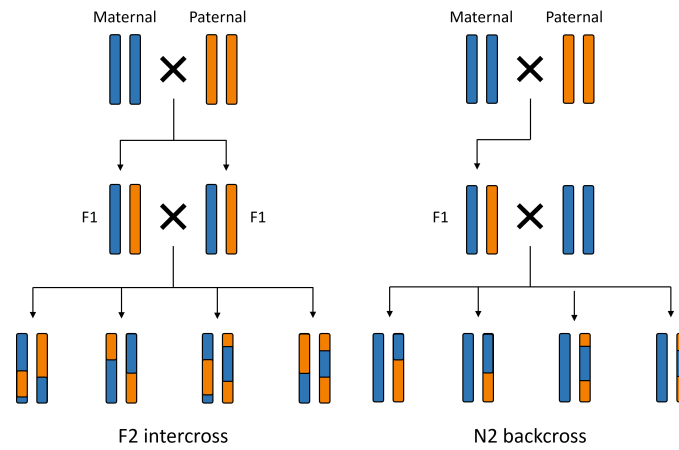
### **II.3.2 . Identifying susceptibility genes in segregating crosses**

#### **II.3.2.1 . Production of segregating crosses**

Mouse genetic studies require a population of individuals with genetic variations. A classical approach starts with identifying a pair of inbred strains showing differences in their susceptibility to the viral infection of interest. To identify host genes responsible for these differences, intercrosses (F2s) and backcrosses (N2s) are produced (Figure 4). F2 is preferentially used when the susceptibility of F1 mice is intermediate between the two parental strains, while N2 is generally used when F1 mice have a susceptibility close to that of one of the parental strains, in order to maximize the amount of phenotypic diversity in second generation individual (SGI) (Kane and Golovkina 2019). The SGIs are then phenotyped for their susceptibility to the infection and genotyped for markers distributed across the genome to perform association analyses (Kane and Golovkina 2019).

SGIs can be used to study Mendelian traits, like in the case of WNV infection where some mice could be unequivocally classified as susceptible or resistant (see section II.1.3.4, Mashimo et al. 2002). They can also be used to study quantitative traits (Mott and Flint 2013), which are generally controlled by several genes and non-genetic factors (Panthier and Montagutelli 2012; Noll et al. 2019). Indeed, the phenotypes studied after viral infection in mice are usually quantitative: number of viral genome copies, histological features, body weight loss, day of death, etc. (Gralinski et al. 2015; Manet 2019; Cartwright et al. 2022; Jasperse et al. 2023).





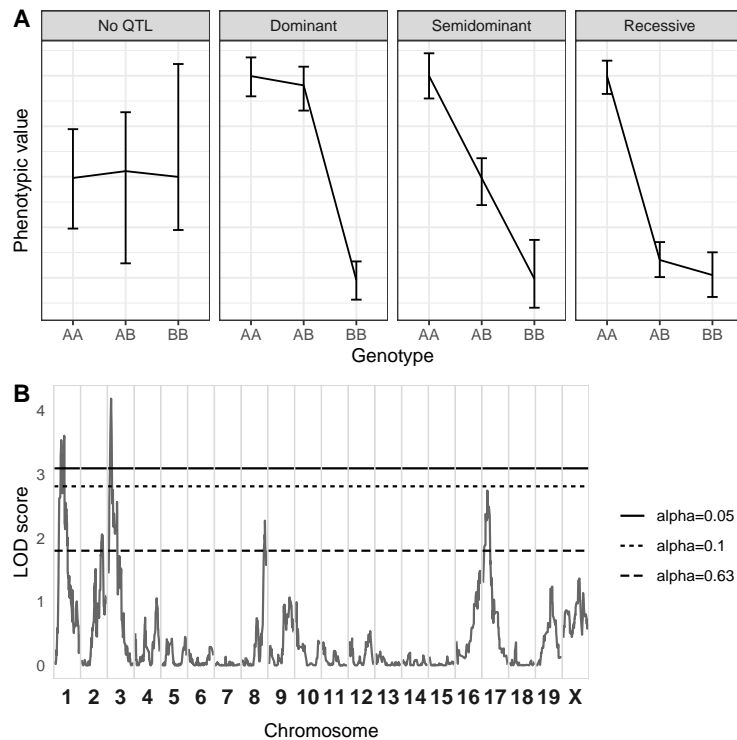
**Figure 4:** Breeding schemes to produce genetically diverse mouse populations. F2 individuals are produced by crossing F1s together (left panel), while N2 individuals are produced by crossing F1s with one of the parental strains (right panel).

### II.3.2.2 . QTL mapping in segregating crosses

Quantitative trait locus (QTL) mapping studies investigate genomic regions contributing to the variability of the trait of interest. The principle of QTL mapping is the following: the SGI population is split into 2 (for N2s) or 3 (for F2s) groups by their genotype at a given marker. If the mean phenotype differs between groups, the locus examined is associated with the phenotype (Figure 5A). This analysis is performed for markers spread across the whole genome to find genomic regions associated with the trait of interest. In addition, interval mapping allows to perform association analysis for regions of the genome with low genotyping coverage. This method calculates the probability for a mouse to have a certain genotype at a position, depending on the genotypes at the nearest flanking markers. Therefore, QTL mapping can be performed for positions where the genotypes of the SGI are lacking (Broman 2001). The statistical value usually used to represent the results is the logarithm of the odds (LOD), which represents the strength of evidence of the presence of a QTL at a given position (Figure 5B, Broman 2001).

A peak of the LOD score curve is considered a QTL when it reaches the significance thresholds established by permutation tests performed on the experimental data. In each permutation, the phenotypic values of the SGIs are randomly reallocated, and LOD scores are calculated to measure signals of spurious association. Using a large number of permutations ( $n > 1,000$ ), a significance threshold of  $p = 0.05$  is given as a LOD score observed in only 5% of them. After finding a significant QTL, it is possible to calculate the percentage of the phenotypic variability that is explained using the following formula:  $1 - 10^{-2LOD/n}$ , with  $LOD$  the LOD score at the peak position of the QTL and  $n$  the number of phenotyped individual (Broman and Sen 2009).

Once a QTL is identified, it is possible to investigate its inheritance pattern to determine which allele is dominant, or whether the two alleles are semidominant (Figure 5A). This analysis can give new insights on the variants associated with the susceptibility. Indeed, loss-of-function variants are usually recessive, while gain-of-function variants are usually dominant (Backwell and Marsh 2022).



**Figure 5:** Example of QTL mapping results

(A) Allelic effects for a locus with no association ("No QTL"), for an associated locus for which the A allele is dominant, for an associated locus for which the two alleles are semidominant, and for an associated locus for which the A allele is recessive. (B) Example of genome scan. The y-axis represents the LOD score. The x-axis represents genomic positions. Horizontal lines represent the significance thresholds calculated by 1000 permutations.

### II.3.2.3 . Identifying candidate susceptibility genes in a QTL

A QTL is not a single position, but a genomic region with a confidence interval (or credible interval if calculated with Bayesian method) which may contain a large number of genes. Thus, follow-up analyses and experiments are required to identify the gene(s) underlying the QTL. The first approach is to reduce the size of the confidence interval.

If two or more SGI studies found a QTL for the same phenotype at the same genomic position using different parental strains, one could make the assumption that these QTLs are controlled by the same gene. Combine cross analysis is a statistical method allowing to reduce the confidence interval size of a QTL shared by several studies. This method requires to reencode the genotypes of the SGIs to a phenotype-specific code, such as "S" for susceptible, "R" for resistant, and "H" for heterozygous, depending on the phenotype of the parental strains. Then, QTL mapping analysis is performed on the combined cross data. For example, DiPetrillo and colleagues have reduced the interval of a QTL controlling high-density lipoprotein concentration from 30 cM to 10 cM by combining data from 4 crosses (DiPetrillo et al. 2005).

Another method to reduce a QTL interval is the development of congenic strains. A congenic mouse strain is a strain carrying a genomic fragment from a "donor" strain and the rest of the genome from a "recipient" strain. This genomic fragment is transferred into the recipient strain by repeated backcrosses to the recipient strain (Montagutelli and Abitbol 2004). To reduce the size of a QTL

identified in a cross between a susceptible and a resistant strains, it is possible to produce congenic strains carrying different fragment of the QTL interval from the susceptible strain. Therefore, the gene responsible for the QTL is in the fragment received by susceptible congenic strains, but absent in resistant congenic strains (Montagutelli and Abitbol 2004).

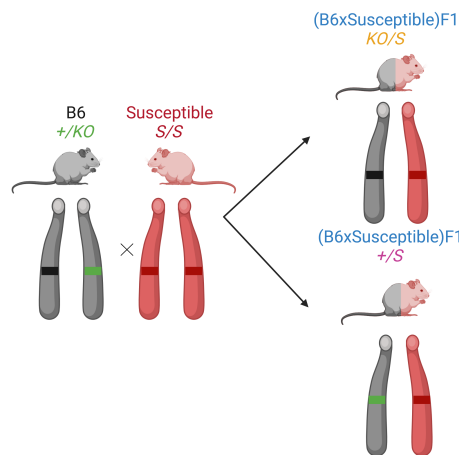
Once the QTL interval is reduced, it is possible to filter the remaining genes with several criteria. Known function of the genes can be used, for example, genes with immune functions can be favored in the search for genes controlling susceptibility to viral infections (Manet 2019). Moreover, DNA sequences of the two parental stains can be compared to identify genes with variants which may modify gene expression or protein function (DiPetrillo et al. 2005). In addition, gene expression data can be used to exclude genes that are not expressed in a tissue of interest, for example, genes only expressed in the liver may not control susceptibility to respiratory infection.

For instance, Bieber and colleagues studied the ability of mice to repair Theiler's murine encephalomyelitis virus (TMEV)-induced demyelination in the central nervous system. They identified a susceptible strain, B10.D1-H2q/SgJ (B10.D1), and a resistant strain, FVB/NJ (FVB). B10.D1 mice showed only minimal myelination repair at 300 day post infection (dpi), while FVB mice had extensive repair. F1 mice had the same resistant phenotype as FVB mice, therefore, the authors produced (B10.D1 × FVB) × B10.D1 N2 mice. The remyelination percentage in N2 mice spread continuously between the values of the two parental strains. QTL mapping was performed to identify loci controlling the remyelination percentage and two significant peaks were identified on chromosomes 3 and 9, explaining 18 and 30% of the variance of this phenotype, respectively. In the chromosome 3 QTL, they selected *Egf* as a candidate gene among the 112 genes in the confidence interval, as it is expressed in the central nervous system. The authors found four polymorphic variants between the two parental strains which altered EGF protein maturation (Bieber et al. 2010).

#### **II.3.2.4 . Confirming the causal role of a candidate gene**

Once a few candidate genes are selected, their role in the susceptibility to viral infection can be tested. Phenotyping KO strains can give insight into the potential role of a gene to face viral infection, however, this method does not confirm that the susceptibility of the SGI parental strain is due to this gene. Congenic strain containing only one gene from the susceptible strain can confirm its causal role in the susceptible strain, but congenic strains for small regions require a large number of generations (Rapp and Joe 2012).

The other method that can confirm the role of a gene in the susceptibility of a specific strain in the quantitative complementation test. This test consists in crossing the susceptible strain with a mouse heterozygous for a KO for the gene on interest (Figure 6). This KO can be found on different genetic background, notably C57BL/6J (B6) mouse. B6 WT mice must have a resistant phenotype, while B6 KO mice must have a susceptible phenotype. As a KO is usually recessive, heterozygous mice have a resistant phenotype. The cross of B6 heterozygous mice with mice of the susceptible strain produces individuals with a (B6 × Susceptible) F1 genetic background. 50% of them received the allele from the susceptible strain (named "S") and the KO allele (*KO/S*), while the other 50% received the allele from the susceptible strain and the WT allele (*+/S*). If the susceptible strain is susceptible due to the gene studied in the test, the S allele will not "complement" the effect of the KO, and *KO/S* individuals



**Figure 6:** Quantitative complementation test method

The susceptible strain (red) is crossed with a mouse heterozygous for a KO of the candidate gene, usually on a B6 background. The black bar represents the KO allele and the green bar represents the WT allele. The mice obtained all have the same genetic background: B6 × Susceptible F1, but have different allele combinations for the gene of interest. 50% received the allele from the susceptible strain ("S", in red) and the KO allele, and the other 50% received the allele from the susceptible strain and the WT allele.

will have a susceptible phenotype. On the opposite, +/S individuals will show a resistant phenotype. Conversely, if the susceptible strain is susceptible due to another gene, the S allele will "complement" the effect of the KO, and both KO/S and +/S individuals will have a resistant phenotype (Turner 2014).

### II.3.2.5 . Advantages of segregating crosses

Two-generation segregating crosses have many advantages to identify susceptibility alleles. They allow to study any mouse strains with contrasting phenotypes. The choice of the parental strains can be made to take into account known genetic factors which effects could hide the role of other susceptibility alleles. For instance, the *H-2D* locus of the major histocompatibility complex (MHC) was shown to be involved in the genetic control of the susceptibility to TMEV infection in mice, along with other loci. Therefore, Bureau and colleagues have used parental strains with contrasting phenotypes sharing the same *H-2D* haplotype. They used SJL/J, which was susceptible to persistent infection, and B10.S, which was resistant. They produced a backcross by mating F1 mice with B10.S mice. In this progeny, two loci on chromosomes 10 and 18 were associated with viral persistence (Bureau et al. 1993).

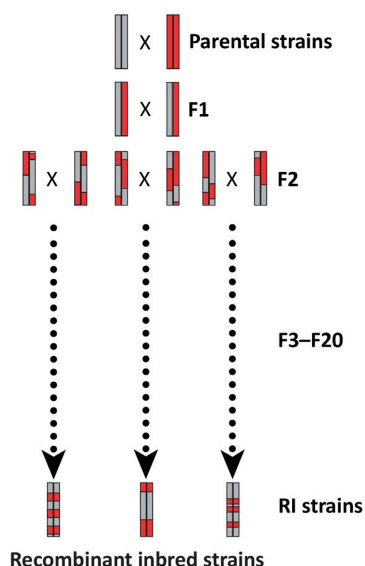
The size and type of cross can be chosen by the experimenter. The choice of F2 or N2 depending on the phenotype of the F1s is made in order to maximize the phenotypic variability in the SGI progeny, which is necessary to identify QTLs. Furthermore, the size of the cross can be increased until significant associations are found.

Any phenotype of interest in the parental strains can be studied in SGIs, and several phenotypes can be measured on the same individual to investigate trait correlations. For example, viral loads can be correlated with body weights post infection. Kinetic studies that require the euthanasia of the mice at different time points cannot be investigated in SGIs. For instance, in the study of Bieber *et al.* described previously, the kinetics of remyelination could not be studied in the N2 as they require the sampling of the brain (Bieber et al. 2010).

### II.3.3 . Identifying susceptibility genes in recombinant inbred strains

#### II.3.3.1 . Biparental recombinant inbred strains

Mouse genetic reference populations (GRPs) are collections of mouse recombinant inbred (RI) strains generated by crossing two or more parental strains and then inbreeding by sibling matings (Figure 7). Several GRPs are biparental RI strains derived from classical laboratory strains: CXB (BALB/cj × B6), AXB/BXA (A/J × B6, B6 × A/J), BXH (B6 × C3H/HeJ) and BXD (B6 × DBA/2J), the latter being the largest biparental GRP (Peirce et al. 2004; Noll et al. 2019).



**Figure 7:** Breeding schemes for the production biparental RI strains. Biparental RI strains are generated by mating intercross animals for at least 20 generations to obtain homozygous individuals (from Leist and Baric 2018)

Like SGIs, biparental RI strains allow to investigate the difference in susceptibility of the two parental strains. For instance, DBA/2J were shown to be highly susceptible to IAV, while B6 were resistant. Indeed, DBA/2J mice rapidly lost weight and died within 7 dpi while B6 mice lost weight between 6 and 8 dpi but regained weight by 14 dpi (Srivastava et al. 2009). Thus, Nedelko and colleagues infected mice of 53 BXD strains with IAV. BXD showed large phenotypic variation. Based on the observed phenotypes, the authors identified 3 groups of BXD strains. In the first group, all the infected mice of each strain survived the infection. In the second group, a majority of mice died but some survived, and in the third group, all mice died (Nedelko et al. 2012).

QTL mapping can be performed on biparental RI strains. For instance, Nedelko and colleagues identified a QTL on chromosome 5 controlling the body weight loss of the BXD strains. Interestingly, this QTL was "transgressive", as the allele inherited from DBA/2J was associated with lower body weight loss, while DBA/2J was the susceptible parental strain (Nedelko et al. 2012).

Biparental RI strains share the advantages of inbred strains. Performing the analyses on several individuals that are genetically identical allow investigating the interindividual variability of the studied traits. Moreover, contrary to SGIs, RI strains allow performing kinetic studies on phenotypes that require to euthanize the individuals and correlating traits measured in different individuals.

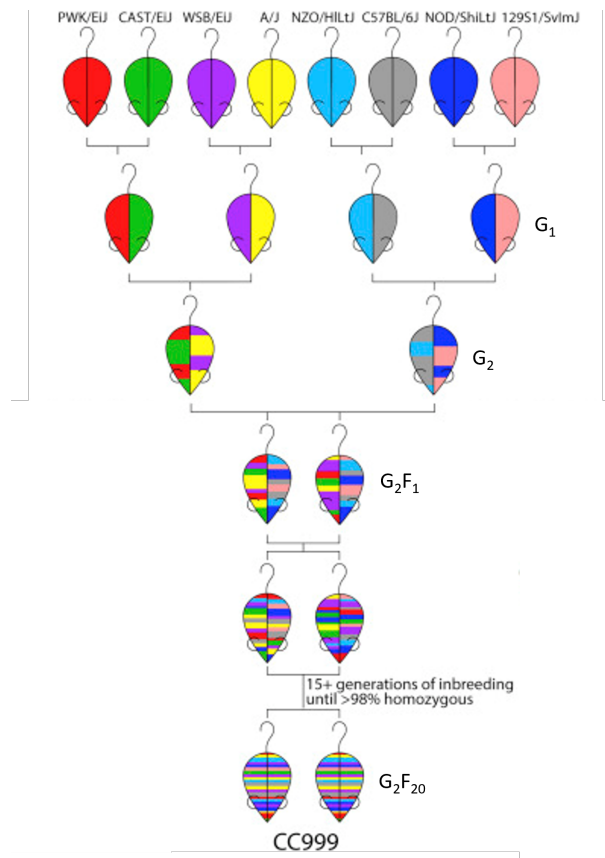
Biparental RI strains were produced from classical laboratory strains, which do not represent a vast amount of genetic diversity. This can make it more difficult to find new susceptibility genes, however, it can also be beneficial. For instance, as discussed previously, most inbred strains possess a deletion in the *Mx1* gene and are susceptible to IAV infection (Staeheli et al. 1988), and a nonfunctional allele for *Oas1b* conferring them WNV susceptibility (Mashimo et al. 2008). Indeed, B6 and DBA/2J mice carry nonfunctional alleles for these two genes, and so do all the BXD strains (Moritoh et al. 2009). This allows studying other susceptibility genes than *Mx1* and *Oas1b* in BXD strains (Nedelko et al. 2012).

### II.3.4 . The Collaborative Cross

To maximize the concept of RI strains by enhancing the genetic diversity segregating in the panel, the Complex Trait Consortium (CTC), a community of mouse geneticists, introduced the idea of creating a multiparental GRP in the early 2000s (Threadgill et al. 2002). The concept of CC was presented in 2004 (The Complex Trait Consortium 2004).

The CC is a GRP derived from eight founder strains: five classical laboratory strains including three extensively studied strains, A/J, B6 and 129S1/SvImJ and two disease-model strains, NOD/ShiLtj which is a model for type 1 diabetes and NZO/HiLtj which is used to study obesity. The three other strains are wild-derived strains from three different *Mus musculus* subspecies, CAST/Eij (*Mus musculus castaneus*), PWK/Phj (*Mus musculus musculus*) and WSB/Eij (*Mus musculus domesticus*), which were selected to maximize genetic diversity (Collaborative Cross Consortium 2012; Panthier and Montagutelli 2012; Noll et al. 2019).

The eight founder strains were crossed according to a funnel breeding scheme leading to third-generation individuals with genetic contribution from the eight parental strains. The positions of the CC founders in the funnel schemes were rotated between each CC to allow balanced contribution from the mitochondrial and Y chromosome genomes. Individuals were then inbred by sibling mating (Figure 8). CC strains were produced in three institutions: Oak Ridge National Laboratory in Tennessee (USA), then relocated to University of North Carolina ("Unc" strains), International Livestock Research Institute in Kenya, then relocated to Tel Aviv University ("Tau" strains) and Western Australia by Geniad Ltd. ("Geni" strains) (UNC 2023). While the goal was to produce over 1000 CC strains, the extinction rate was extremely high (Shorter et al. 2017) and there are nowadays less than 80 available CC strains, all bred at the University of North Carolina (UNC 2023), and available from the Jackson Laboratory (Bar Harbor, Maine, USA). About 40 of them were imported and are bred at the Institut Pasteur.



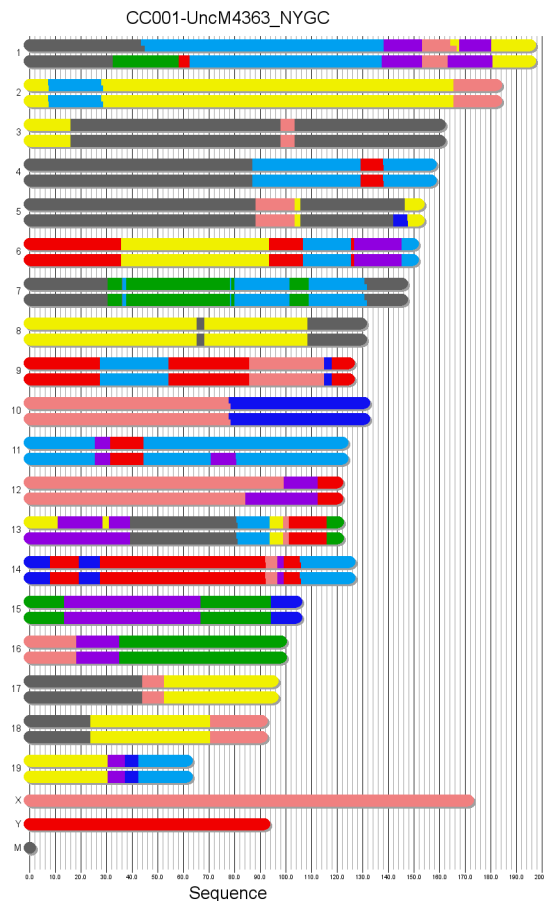
**Figure 8:** Collaborative Cross strains breeding funnel.

The eight founder strains were crossed in funnel breeding scheme to obtain a progeny with genetic contribution from each one ( $G_2F_1$ ), then individuals were inbred to reach homozygosity ( $G_2F_{20}$ ) (modified from Noll et al. 2019).

The breeding scheme resulted in the genomes of the CC strains being a mosaic of haplotypes of the 8 founder strains (Figure 9, Srivastava et al. 2017). Haplotypes and genotypes are encoded with specific colors and letters corresponding to each founder strain: A/J, yellow (A); B6, gray (B); 129S1/SvImJ, pink (C); NOD/ShiLtJ, dark blue (D); NZO/HILtJ, light blue (E); CAST/Eij, green (F); PWK/PhJ, red (G); and WSB/Eij, purple (H). Apart from genetic polymorphisms present in the founder strains, CC strains may also carry spontaneous mutations which got fixed during the inbreeding process, named "private variants" (Srivastava et al. 2017).

The CC combines the advantages of inbred strains with a broad genetic diversity, similar to that observed in the human population (Roberts et al. 2007), which results in very large phenotypic diversity. The CC has been used to study many complex traits such as Alzheimer's disease (Uyar et al. 2021), diabetes (Ghnam et al. 2023), cancers (Nellis et al. 2021; Yang et al. 2023a), addiction (Schoenrock et al. 2022) and infection diseases (see below), among others.

During the development of the CC, a collection of partly inbred individuals (pre-CC) were used to derive another diverse mouse population: the Diversity Outbred (DO). Contrary to the CC, the DO is an outbred stock maintained by randomized breeding strategies allowing each DO individual to be genetically unique, heterozygous, with new recombination breakpoints accumulating at each generation. This population provides high-resolution mapping, which is only limited by the number of mice analyzed. While the genomes of CC mice are fixed and known, each DO individual is unique



**Figure 9:** Example of CC strain genome.

Schematic representation of CC001 genome, from <https://csbio.unc.edu/CCstatus/CCGenomes/>. Colors represent the CC ancestral haplotype: A/J, yellow; B6, gray; 129S1/SvImJ, pink; NOD/ShiLtJ, dark blue; NZO/HILtJ, light blue; CAST/Eij, green; PWK/PhJ, red; and WSB/Eij, purple.

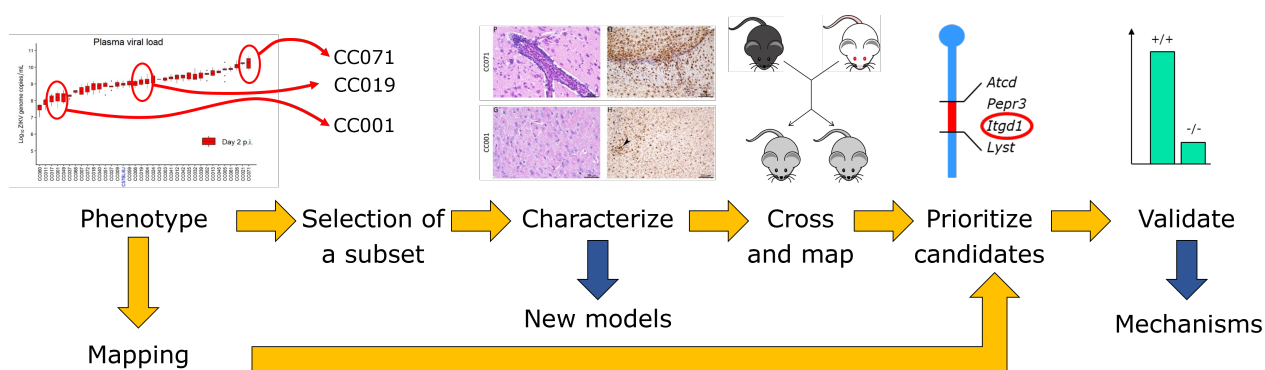
and genotyping is required to perform genetic mapping (Svenson et al. 2012; Keele 2023).

#### II.3.4.1 . Genetic analysis in the Collaborative Cross

In the Mouse Genetics Laboratory of the Institut Pasteur, we use the following workflow when studying CC strains (Figure 10). First, we screen the CC strains available on site for a trait of interest, in this case viral infection. We investigate many relevant parameters on infected mice, such as viral load in blood and/or tissues, body weight loss, clinical symptoms, mortality, histological severity features, etc. For more in-depth studies, we restrict analyses to a subset of strains sampling the range of phenotypes and/or presenting outstanding features. Deep phenotyping of these strains allows the characterization of new models to study viral infection (Manet 2019).

One of our main goals is to identify genes responsible for differences in the resistance or the susceptibility to viral diseases. To do so, QTL mapping can be performed directly on CC strains by looking for association between post-infection phenotypes and their known genotypes (Gralinski et al. 2017; Martin et al. 2020). The power to detect significant associations may vary depending on the complexity of the genetic control of each trait (number of genes, strength of each gene's effect). It is limited by the number of strains available, and their complex genetic architecture. In particular, the number of alleles segregating at every locus. Moreover, the mapping resolution is limited by





**Figure 10:** CC experiment workflow.

After screening of multiple CC strains, strains with diverse phenotypes are selected and characterized. Pairs of strains are crossed to perform genetic mapping. Mapping can also be done directly on CC screening results. After mapping, candidate genes are validated with functional analyses.

the number of CC strains studied and the recombination breakpoints they carry (Manet et al. 2020). To reduce this complexity, F2s and N2s can be produced by crossing two CC strains with extreme phenotypes, or a CC strain and a classical laboratory mouse strain (Gralinski et al. 2017; Schäfer et al. 2022). The progeny resulting from these two-generation crosses are analyzed as described above.

#### II.3.4.2 . Infectious diseases studies with the Collaborative Cross

The CC have several advantages for studying viral infection. Like RI strains, they allow assessing phenotypes of genetically identical individuals, and replication experiments over time and by different research teams. Additionally, their genetic diversity results in broader ranges of phenotypes than classical inbred strains. For example, values of immune cell composition in the blood of 47 CC strains ranged beyond phenotypic values of the classical laboratory strains tested, B6 and BALB/c mice (Martin et al. 2020). Allelic combinations in the CC strains can also result in novel disease types. For instance, Norway rat hepatitis virus (NrHV) causes chronic infection in rats, but common laboratory mice clear the infection in 14 to 21 days. In contrast, 4 out of the 10 CC strains tested showed persistent viremia 4 weeks post infection (Brown et al. 2023).

Among experimental models of viral infection in mice presented in Figure 3, CC strains can be used to study viruses for which mice are naturally susceptible, mouse-adapted viruses, and viruses that require sensitization. For example, CC strains have been used to study mouse-adapted SARS-CoV and SARS-CoV-2 (Schäfer et al. 2022), Rift Valley fever virus (RVFV), which is lethal in WT B6 mice (Cartwright et al. 2022), and ZIKV which requires prior sensitization with IFNAR1 blocking antibody (Manet et al. 2020). CC strains can also be used to study viruses requiring a transgene by crossing each CC with the transgenic strain. The resulting F1 mice have one fixed chromosome from the transgenic strain, and one variable chromosome from the CC strains. This method was used to study SARS-CoV-2 infection in F1s between CC founders and mice transgenic for human ACE2 receptor (Robertson et al. 2023).

Viral diseases were investigated when the CC was still under development, first on founder strains, then on incipient strains (pre-CC) and finally on established CC strains starting in the 2010s.

Gralinski and colleagues first tested mice of the 8 founder strains and 147 pre-CC mice infected with mouse adapted SARS-CoV. They measured body weight loss, lung viral titer, inflammation and

lung pathology at 4 dpi. They observed wide variability in these phenotypes with body weight loss ranging from 0 to 25% and lung viral titer from  $1e3$  to  $1e7$  plaque-forming unit (PFU)/g, but they showed that the two traits did not correlate across the pre-CC mice. PWK/PhJ mice had the lowest lung viral titer of the 8 founders, but were among those that lost the most body weight after infection. Genetic mapping on the CC strains identified 4 QTLs. Of these, one QTL on chromosome 3 controlling vascular cuffing (i.e., immune cell infiltration) in the lung contained only one coding gene, *Trim55*, after filtering for genes with known polymorphisms between founder strains. This gene had only been investigated in the case of muscle development and cardiac function prior to this study. Upon infection, *Trim55* KO mice had similar body weight loss to WT mice but reduced vascular cuffing in the lungs (Gralinski et al. 2015). Therefore, *Trim55* loss-of-function impacted the severity of SARS-CoV infection, but this experiment does not prove that it is responsible for the QTL detected in the CC study.

Following this study, they selected two CC strains with extreme phenotypes: CC003/Unc, which showed mild body weight loss but high viral titer after infection; and CC053/Unc which showed increased body weight loss and mortality but low lung viral titer. They produced and infected 264 F2s between these two strains. F2s showed high phenotypic diversity beyond the range observed in the parental strains and a merely positive correlation between lung viral titer and body weight loss. Genetic analysis revealed 5 QTLs including one on chromosome 18 associated with multiple phenotypes: body weight loss, viral titer and pulmonary hemorrhage. *Ticam2* was selected as a candidate gene for further analysis. *Ticam2*-deficient mice lost more body weight than WT mice and higher viral titer at 2 but not at 4 dpi (Gralinski et al. 2017).

Similarly, Schäfer and colleagues have studied SARS-CoV and SARS-CoV-2 infection in CC strains. Among the 5 strains infected with mouse-adapted SARS-CoV MA15, CC011/Unc was resistant (100% survival rate, less than 5% body weight loss at 4 dpi) and CC074/Unc was susceptible (80% mortality and more than 15% body weight loss), but the two strains had similar lung viral loads. After infection with mouse adapted SARS-CoV-2, CC074/Unc mice showed increased titer in the lung at 2 dpi compared with CC011/Unc, and all died while CC011/Unc mice survived (Schäfer et al. 2022).

More than 400 F2 individuals were then produced and infected with MA15, and showed high phenotypic variability in survival, lung viral titer and lung pathology. Genetic mapping in the F2 population revealed 5 significant QTLs including one on chromosome 9 associated with mortality, lung congestion, altered pulmonary function, and peripheral lymphocyte and neutrophil levels. Among the candidate genes in this region, *Ccr9* and *Cxcr6* were investigated. *Ccr9*-deficient mice showed significantly higher body weight loss and increased lung pathology compared with WT mice, while *Cxcr6*-deficient mice showed increased mortality and lung viral titer after infection with either mouse-adapted SARS-CoV or SARS-CoV2 infections (Schäfer et al. 2022).

Other studies have examined immunologic phenotypes. Martin and colleagues studied immune cell compositions in the blood and found that the percentage of immune cell type (B cells, CD4 and CD8 T cells, NK cells, granulocytes and monocytes) is highly variable among CC strains, exceeding values observed in classical laboratory strains B6 and BALB/c. After lymphocytic choriomeningitis virus (LCMV) infection, CC strains showed variable body weight loss, circulating cytokine levels (IFN $\alpha$  and IFN $\gamma$ ) and percentages of CD4 and CD8 T cells, and subtypes of CD8 T cells, demonstrating that genetic background influences the composition of immune cell types before and after infection. QTL mapping on subclasses of CD8 T cells after infection identified two loci on chromosomes 18 and 19 associated with the amount of central memory T cells, which are protective against chronic infection. Nine candidate genes were identified in these two regions based on known polymorphisms in founder strains with potential effects on protein function or gene regulation (Martin et al. 2020).

In conclusion, the CC has proven valuable for studying of the variability of response to viral infection and for identifying genes responsible for the differences in susceptibility to viral disease.

## II.4 . Genetic susceptibility to Zika virus

Since 2016, the Mouse Genetics Laboratory has been investigating host genetic control of the susceptibility to ZIKV using CC strains. This was the topic of my PhD project.

### II.4.1 . Zika virus epidemiology

#### II.4.1.1 . ZIKV emergence and epidemics

ZIKV is a flavivirus, belonging to the *Flaviviridae* family, like DENV, WNV, YFV and Japanese encephalitis virus (JEV). ZIKV was first isolated in 1947 from a rhesus monkey in Zika Forest, Uganda. Then, the virus was isolated from *Aedes Africanus* mosquitoes in the same forest in 1948. The first case of human infection with ZIKV was reported in 1964 in Uganda (Talero-Gutiérrez et al. 2018).

Only a few cases of ZIKV infection were reported in humans until the 2000s (Musso and Gubler 2016). In 2007, the first significant outbreak of ZIKV occurred in Yap Islands where a survey of 173 households identified 414 out of 557 persons with antibodies against ZIKV, including 37.7% of symptomatic infections. In 3 years, 72.6% of the population was estimated to be infected (Musso and Gubler 2016).

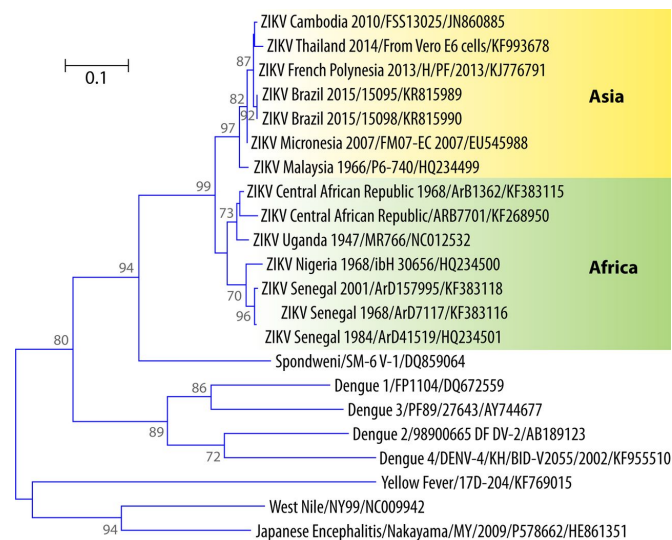
In 2013, a ZIKV outbreak occurred in French Polynesia and spread in 2014 to other Pacific islands, New Caledonia, Cook Islands and Easter Islands (Weaver et al. 2016). The estimated proportion of ZIKV-infected people at the end of the outbreak was 11.5% in French Polynesia, and 0.8% in New Caledonia (Musso and Gubler 2016).

The epidemic in Brazil started in late 2014 and the number of cases increased during the first months of 2015. The Brazilian Ministry of Health estimates that between 497,593 and 1,482,701 cases of ZIKV infection occurred between the beginning of the epidemic and 2016 (Weaver 2017). Different hypotheses exist as to the introduction of ZIKV. It was suggested that ZIKV was introduced during the 2014 World Cup soccer competition or the 2014 World Spring Canoe championship (Musso and Gubler 2016). However, phylogenetic studies suggest an introduction in 2013 (Weaver 2017). The wide distribution of *Aedes aegypti* and *Aedes albopictus* in Brazil probably facilitated the swift spread of ZIKV. From Brazil, ZIKV eventually disseminated to 48 countries in the Americas (Ikejezie 2017).

The first ZIKV outbreak in India occurred in 2018 with 159 cases tested positive including 64 pregnant women (Agarwal and Chaurasia 2021), but no increase of microcephaly cases or neurological complication were reported (Biswas et al. 2020). A second outbreak occurred in the country in 2021 during the COVID-19 epidemic (Bardhan et al. 2021). As of December 2021, a total of 89 countries have reported cases of ZIKV infection (WHO 2022).

### II.4.1.2 . ZIKV evolution

Phylogenetic studies have identified two lineages of ZIKV: the African lineage and the Asian lineage (Figure 11). The African lineage contains the original Uganda strain, and strains from other African countries where the virus is endemic (i.e., Senegal, Nigeria; Musso and Gubler 2016). The Asian lineage contains the viruses responsible for the epidemics in the Pacific islands and the Americas (Hung and Huang 2021). Viruses from this lineage have been extensively studied as they account for approximately 97% of the published ZIKV sequences (Beaver et al. 2018). It has been shown that recent African ZIKV strains were more transmissible by mosquitoes than Asian strains, and induced increased plasma viremia and body weight loss and faster development of disease symptoms in mice (Aubry et al. 2021).



**Figure 11:** Phylogenetic tree of ZIKV strains.

Viral strains are grouped in two lineages: the African lineage and the Asian lineage (from Musso and Gubler 2016).

### II.4.1.3 . ZIKV transmission

ZIKV is primarily transmitted through *Aedes* mosquito bite. Several species of mosquitoes can transmit ZIKV: *Aedes aegypti*, *Aedes albopictus*, *Aedes hensilli* and *Aedes polynesiensis* (Bhardwaj et al. 2021). Mosquitoes get infected by biting an infected individual, then the virus replicates into the mosquito and can be transmitted to another individual during a following blood meal (Bhat et al. 2023). ZIKV can adopt a sylvatic transmission cycle, where the virus is transmitted between mosquitoes and non-human primates, or an urban transmission cycle, where the transmission occurs between mosquitoes and humans (Vasilakis et al. 2011; Rather et al. 2017).

ZIKV can also be transmitted sexually as it can persist in men's semen for more than 400 days, and both men-to-woman and men-to-men transmissions have been described (Sookaromdee and Wiwanitkit 2018). Cases of woman-to-man transmissions are less frequent (Bhat et al. 2023). Vertical transmission can occur from infected mother to fetus during pregnancy as the virus can cross the blood-placental-barrier. Mothers can transmit ZIKV to their children through breastfeeding, but newborns are usually asymptomatic with this mode of transmission (Bhardwaj et al. 2021).

#### **II.4.1.4 . ZIKV disease**

ZIKV infection in humans can lead to various outcomes. Approximately 80% of infected individuals are asymptomatic (Javed et al. 2018). Otherwise, infected individuals experience a mild flu-like illness with fever, muscle pain, headache, fatigue, rash and/or conjunctivitis. Less frequently, ZIKV can cause neurological complications in adults, including encephalitis, meningitis and Guillain-Barré syndrome. Guillain-Barré syndrome is an immune-mediated neuropathy characterized by palsies, ataxia (i.e., defective muscle coordination) and areflexia (i.e., absence of muscle response to stimuli). In Columbia between 2015 and 2017, the frequency of Guillain-Barré cases was 41.9 per 10,000 cases of ZIKV disease (Carod-Artal 2018). In rare instances, ZIKV infection has led to mortality in children with sickle cell disease and adults with cancer (Pierson and Diamond 2018; Bhardwaj et al. 2021).

Infection of fetuses through ZIKV vertical transmission can induce congenital Zika syndrome (CZS), the most severe clinical manifestations of which include fetal loss and microcephaly. Indeed, a 20-fold increase in the number of newborns with reduced head circumference and 8201 cases of microcephaly were recorded in Brazil between December 2015 and July 2016. The link between ZIKV infection of pregnant mothers and microcephaly in newborns was suspected after ZIKV genome was detected in brain samples from deceased neonates with microcephaly, consistent with the increase of incidence of microcephaly observed during the ZIKV epidemic. The incidence of microcephaly was between 6% in and 11% ZIKV-infected pregnant women in the USA (Carod-Artal 2018).

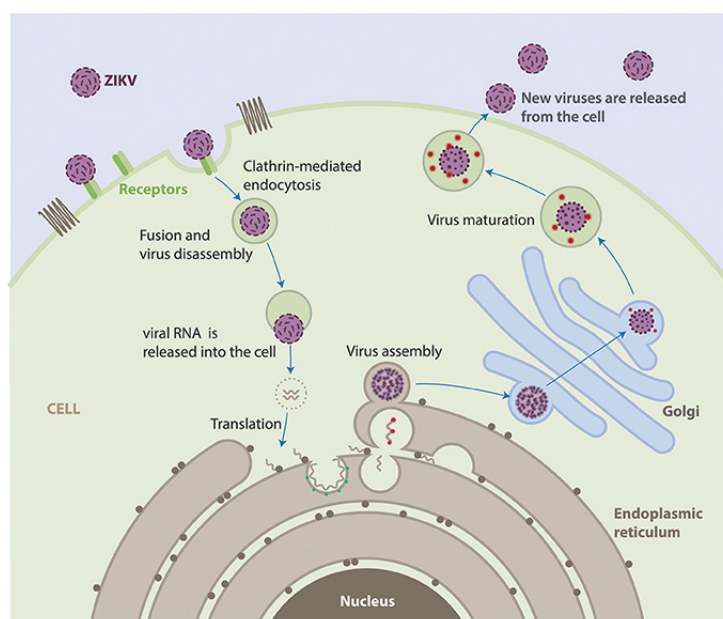
Infants with CZS may show other symptoms for months after birth, such as irritability, altered movements, seizures, sleeping disorders, hearing and vision alteration, etc. (Pierson and Diamond 2018; Bhardwaj et al. 2021). Long term consequences of ZIKV infection during fetal development are becoming evident as children become older.

#### **II.4.2 . Biology of Zika virus**

##### **II.4.2.1 . ZIKV tropism and cell cycle**

ZIKV tropism includes several cell types. ZIKV infects skin cells such as fibroblasts, keratinocytes, and skin dendritic cells, cells from the optic nerve, cornea, iris and retinal cells, neural precursor cells, and fetal endothelial cells (Laureti et al. 2018). Indeed, several receptors have been shown to mediate ZIKV entry into host cells: the immune receptor DC-SIGN, the transmembrane proteins TIM-1 and the TAM receptors TYRO3, AXL and MER (Laureti et al. 2018). However, whether any of these receptors is necessary for viral entry remains unclear. For example, ZIKV-infected *AXL* KD human cells produced reduced ZIKV titer. Likewise, ZIKV gene expression in human *AXL* KO HeLa cells was reduced, but restored by the expression of murine *Axl*. On the other hand, *Axl* KO mice had similar viral RNA levels in the brain and in the spleen as WT mice. Thus, further studied will elucidate the role of AXL and other receptors in ZIKV entry into host cells.

After binding its receptors, ZIKV enters host cells by endocytosis. Once internalized, the viral and endosomal membranes fuse to release the viral genome into the cytoplasm. The viral genome is replicated in the cytoplasm and proteins are translated in the ER where immature virions are formed. Then, they mature during the transit to the Golgi network from which they bud as vesicles. Mature virions are released by exocytosis (Figure 12, Acosta-Ampudia et al. 2018; Hasan et al. 2018).



**Figure 12:** ZIKV replication cycle.

ZIKV enters into host cells through endocytosis. Inside the cell, viral RNA is released, replicated and translated. New virions are assembled in the ER, transported through the Golgi network and released by exocytosis (from Acosta-Ampudia et al. 2018).

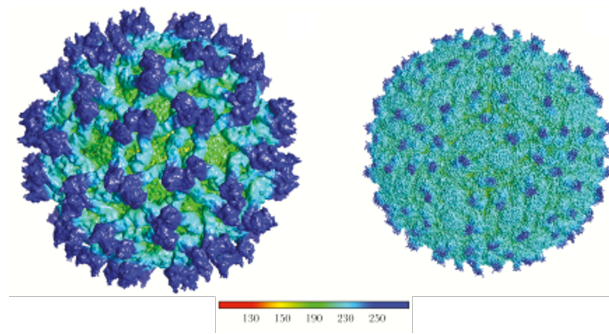
#### II.4.2.2 . Structure of ZIKV particles

ZIKV is a small enveloped virus with an 11 kilobase single positive strand RNA genome. The genome is translated into a single polypeptide which is cleaved by proteases to yield three structural proteins: capsid (C), envelope (E) and precursor membrane (prM) proteins, and seven non-structural proteins: NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 (Bhardwaj et al. 2021).

The precursor membrane protein and the envelope protein are embedded in the viral membrane. The E protein is involved in receptor binding and fusion (Javed et al. 2018). prM and E proteins are assembled in the ER membrane as groups of 3 prM-E heterodimers, with the pr side of the prM proteins on top, forming a spike and preventing fusion by E protein. In the Golgi network, the pH allows a conformational change of the E protein and cleavage of the pr-M junction, leading to a smooth virion (Figure 13, Sirohi and Kuhn 2017).

The capsid proteins bind ZIKV genome to form the nucleocapsid (Bhardwaj et al. 2021). Capsid proteins possess one side with basic residues that bind viral RNA, thus positioning inside the nucleocapsid, and a hydrophobic side which positions towards the viral envelope (Javed et al. 2018). The capsid protects viral RNAs from the innate immune response by hiding them from host RNA sensors and RNases (Javed et al. 2018).

The nonstructural proteins of ZIKV are involved in different steps of the viral cycle. NS1 facilitates ZIKV replication by forming a replication compartment for ZIKV in the ER (Ci et al. 2019), NS2B is involved in viral assembly, NS3/NS2B are involved in ZIKV polypeptide cleavage (Javed et al. 2018; Guo et al. 2021), and NS5 in viral RNA synthesis. ZIKV nonstructural proteins also inhibit the host immune response. NS1, NS2A, NS2B and NS4B block TBK1 phosphorylation, NS4A inhibits IRF3 phosphorylation and associates with MAVS to block its interactions with RLRs. NS5 mediates STAT2 degradation through the proteasome, thus blocking signal transduction from IFNAR and inhibiting



**Figure 13:** ZIKV structure.

Left image shows ZIKV immature virion (9.1 Å resolution) and right image shows ZIKV mature virion after conformational change of the E protein (3.8 Å resolution). Images are surface-shaded colored views of virions (modified from Sirohi and Kuhn 2017).

ISGs production. This mechanism is active in human but not in mice, which explains why this species is refractory to ZIKV infection (Cumberworth et al. 2017; Guo et al. 2021).

### II.4.3 . Host genetic susceptibility to ZIKV infection in human

A few studies have investigated host genetic factors controlling susceptibility to ZIKV infection in human, mostly focusing on CZS.

#### II.4.3.1 . Sequence analyses

De O. da Silva and colleagues performed a WES study on discordant dizygotic twins. As they previously demonstrated that the expression of genes of the mTOR pathway are differently expressed between affected and unaffected twins (Caires-Júnior et al. 2018), they searched for variants in genes of this pathway. Several homozygous variants in the *MTOR* gene were identified in the affected twin, while the unaffected twin was heterozygous. However, the study did not confirm the causal role of these variants (De O. da Silva et al. 2021).

Similarly, WES on postmortem brains of 5 CZS neonates identified 23 variants in genes involved in central nervous system development, immune system (including *TLR2* and *IRF3*), and extracellular matrix organization, notably collagen-encoding genes. Since only CZS patients were sequenced, these variants were not genetically associated with the disease. However, collagen fibers were less abundant in the brains CZS patients compared with ZIKV-negative controls, suggesting a role of collagen in the CZS pathogenesis (Aguar et al. 2020).

#### II.4.3.2 . Association analyses

Five studies used association analyses to identify variants controlling CZS susceptibility. Three of them did not identify variants significantly associated with CZS in affected children (Caires-Júnior et al. 2018; Borda et al. 2021; Gomes et al. 2023).

Conversely, two other studies found significant associations. Rossi and colleagues performed association analyses on 28 CZS babies and 24 control babies from mothers infected with ZIKV during pregnancy. They focused on genes of the adenylate cyclase pathway which were previously shown to modulate HBV and HIV entry (Zhang et al. 2016; Moreno-Fernandez et al. 2011). 22 variants in 17



genes in mothers were associated with CZS in their children, and the most highly associated SNPs were in the *ADCY3* and *ADCY7* genes (Rossi et al. 2019).

In another study, SNPs in 8 genes of the innate immunity and inflammation pathways were investigated. These SNPs were genotyped in 68 CZS children and their mothers, as well as in healthy 43 children and mothers. Alleles of *TLR3* variants were differently distributed between mothers of CZS patients and healthy control mothers; and a variant in *TNF* was associated with CZS in children (Santos et al. 2019).

No genome wide analysis has identified alleles associated with susceptibility to ZIKV disease in humans.

#### **II.4.3.3 . *In vitro* analyses**

ZIKV host susceptibility factors have also been studied *in vitro*. CRISPR/Cas9 screen was performed in HeLa cells infected with ZIKV. After 8 dpi, most cells were dead but a small population of mutant cells survived. Analysis of the mutations in the surviving cells showed that silencing of *AXL* and genes coding proteins of the ER complex was associated with increased cell survival (Savidis et al. 2016). The same screening strategy was then performed in neural progenitors derived from human induced pluripotent stem cell (hiPSC). Silencing of genes associated with ER processing (*EMC1*, *EMC2*, *EMC6*, etc.), endosome-lysosome acidification (*ATP6V1C1*, *ATP6V1F*), heparan sulfation (*TM9SF2*, *EXTL2*, *EXTL3*, etc.) and negative regulators of the IFN-I response (*SOCS3*, *STAT3*, etc.) were protective against infection-induced cell mortality. The identification of host restriction factors in the endosome-lysosome pathway was consistent with ZIKV viral entry by endocytosis (Li et al. 2019).

Another study used undifferentiated hiPSC from 77 individuals of different ethnicities infected with ZIKV. Depending on the level of viral replication *in vitro*, hiPSC were classified in "permissive" and "low permissive". GWAS was performed and identified association of SNPs in the regulatory regions of the *NDUFA4*. This gene encodes a mitochondrial protein involved in electron transport chain. Moreover, *NDUFA4* deficient cells showed lower percentage of virus-positive cells after infection with DENV or SARS-CoV-2 (Han et al. 2022). Then, the authors differentiated hiPSC into trophectoderm cells, as they represent a major route for mother-to-fetus ZIKV transmission. In this cellular model also, reduction of *NDUFA4* expression led to reduced ZIKV titers (Yang et al. 2023b).

#### **II.4.4 . Mouse models of ZIKV infection**

As discussed in section II.3.1, *Ifnar1* KO mice are the most widely used model of ZIKV infection. *Ifnar1* KO mice have been used to study ZIKV vertical and sexual transmission (Winkler et al. 2017), vaccine and treatment development (Gambino et al. 2021; Matz et al. 2021), immune response (Huang et al. 2017), and variants between viral strains (Carbaugh et al. 2020), for example.

The first *Ifnar1* deficient mouse model was produced in 1994 to describe the role of IFN-I in the antiviral response. Homologous recombination using a vector targeting the exons 3 and 4 of the gene was used in 129S2/SvPas (129) ES cells, to obtain 129-*Ifnar1* mice, also called A129 (Müller et al. 1994). *Ifnar1* KO mice were then backcrossed to B6 in order to obtain a congenic strain carrying the *Ifnar1* KO on a B6 genetic background (<https://www.jax.org/strain/010830>).

These two mouse strains were described as models of ZIKV pathogenesis in 2016. Lazear

and colleagues subcutaneously infected B6-*Ifnar1* mice with  $10^2$  focus-forming unit (FFU) of ZIKV H/PF/2013 (Asian lineage) or MR 766 (African lineage). B6-*Ifnar1* showed 100% mortality with the Asian strain and 80% with the African strain, with a body weight loss ranging between 15 and 25% (Lazear et al. 2016). Rossi and colleagues infected intraperitoneally (i.p.) 129-*Ifnar1* mice with the Asian ZIKV strain FSS13025. 3-week-old mice showed a 100% mortality rate, while only 50% of 5-week-old mice died and all 11-week-old mice survived (Rossi et al. 2016).

Previous work from our group have investigated the impact of the genetic background of *Ifnar1* deficient mice on ZIKV pathogenesis. Age-matched B6-*Ifnar1* and 129-*Ifnar1* were infected i.p. with  $10^7$  FFUs of French Guiana 2015 (FG15) Asian ZIKV strain and followed for 14 days. B6-*Ifnar1* started to develop symptoms, including body weight loss, ruffled fur, ataxia, and hind limb paralysis, from 4 dpi, and were all moribund or dead by 7 dpi. On the opposite, 129-*Ifnar1* developed only mild symptoms (ruffled fur and hunched back) from 6 dpi and only one out of 7 mice died at 9 dpi. B6-*Ifnar1* and 129-*Ifnar1* mice had similar plasma viral load (PVL) at 2 dpi but B6-*Ifnar1* mice had increased PVL at 6 dpi (Manet et al. 2020).

Other immunodeficient models have been used and showed clinical symptoms after infection. *Irf3 Irf5 Irf7* triple KO B6 mice lack the three main IFN-I transcription factors and are therefore unable to mount a IFN-I response. They showed similar mortality and body weight loss as B6-*Ifnar1* mice, while *Irf3* simple KO mice did not develop illness (Lazear et al. 2016). Mice deficient for both IFN-I and type II IFN receptors on the 129 genetic background showed the same disease severity as 129-*Ifnar1* mice (Rossi et al. 2016).

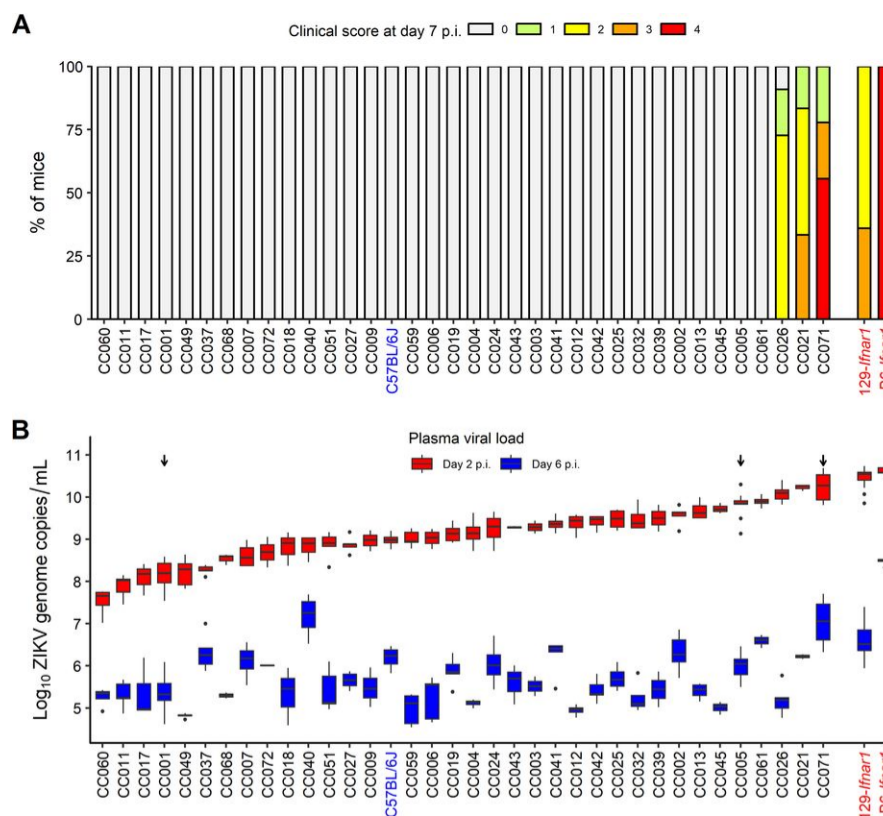
ZIKV infection has been also studied in WT in mice with the use of MAR1-5A3, a monoclonal antibody which blocks the IFNAR1 receptor to inhibit the binding of IFN-I and subsequent induction of ISGs. Compared to B6-*Ifnar1* deficient mice, WT B6 mice treated with MAR1-5A3 antibody did not succumb or lose weight after infection, but developed higher viral load compared to WT mice untreated or treated with control isotype antibody (Lazear et al. 2016; Manet et al. 2020).

A mouse adapted strain of ZIKV was developed in 2018 by Gorman and colleagues. This strain was obtained by serial passages of the African Dakar strain of ZIKV on *Rag1* deficient mice. These mice lack mature B and T cells but have an intact IFN-I induction and were therefore chosen to select viral mutants escaping the IFN-I response. As the virus is able to invade the brain, brain homogenates of infected *Rag1* deficient mice were used to infect naive *Rag1* deficient mice by subcutaneous injection for 4 passages to obtain the mouse adapted ZIKV-Dak-MA strain. This strain carries a mutation in NS4B which leads to increased viral replication and reduced IFN-I production in WT B6 mice after subcutaneous infection. However, this strain still requires that mice are pretreated with MAR1-5A3 antibody to efficiently replicate (Gorman et al. 2018).

#### II.4.5 . Previous results on ZIKV infection in the Collaborative Cross

In their 2019 preprint, Mattocks and colleagues have reported the study of 21 CC strains with 3 ZIKV strains including the mouse adapted Dakar strain, in the search for a strain that could be a model for ZIKV pathogenesis without the need to block the IFN-I response. However, none of the strains tested showed body weight loss following infection with either virus and ZIKV was rapidly cleared in the blood (Mattocks et al. 2019).

C. Manet, a former PhD student from my laboratory, has studied CC mice infected with ZIKV. CC mice were infected i.p. with  $10^7$  FFUs of FG15 Asian strain preceded by MAR1-5A3 antibody treatment. The screening of 35 CC strains showed that only mice from 3 strains developed symptoms after infection. While CC021/Unc and CC026/GeniUnc mice recovered, 78% of CC071 mice were moribund or dead at 7 dpi (Figure 14A). PVL peak, which occurred at 2 dpi, showed a 3-log variation among the CC strains tested and did not correlate with the PVL at 6 dpi. Again, the CC strain with the highest PVL at 2 dpi was CC071 (Figure 14B, Manet et al. 2020). Of note, CC071 mice were not tested in Mattocks and colleagues' study (Mattocks et al. 2019), although they showed low viral titer and no symptoms in the absence of MAR1-5A3 (Manet et al. 2020).

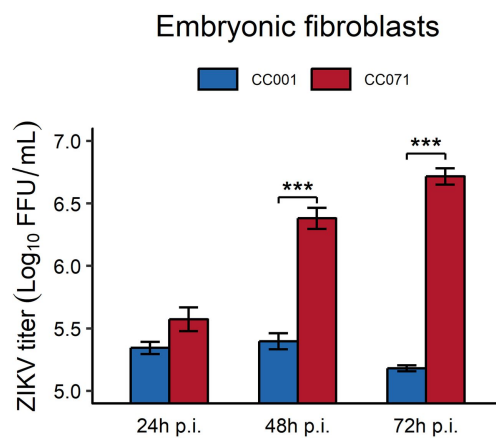


**Figure 14:** Genetic background impacts clinical symptoms and PVL after ZIKV infection in CC mice. 35 CC strains (between 2 and 9 mice per strain) and *Ifnar1* KO were infected i.p. with  $10^7$  FFUs of ZIKV FG15. A: Clinical scores at 7 dpi as follows: 0, no symptoms; 1, ruffled fur; 2, emaciation, hunched posture, and/or hypoactivity; 3, hind limb weakness, prostration, and/or closed eyes; and 4, moribund or dead. B: Plasma viral load at 2 and 6 dpi quantified by RT-qPCR (from Manet et al. 2020).

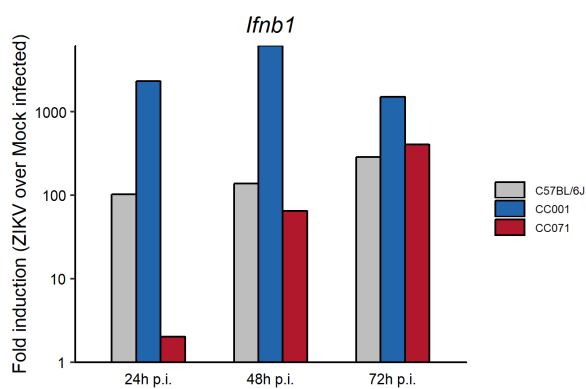
Several strains with divergent phenotypes were selected for deeper phenotyping, in particular CC071 as the most susceptible strain, and CC001/Unc (CC001) as a resistant strain with low peak plasma viral load and no illness after infection (Manet et al. 2020). Brain pathology following ZIKV infection was investigated. CC071 showed higher brain viral load after i.p. and intracranial (i.c.) infection, with severe inflammatory lesions, immune cell infiltration and microglial reactivity (Manet et al. 2020). The difference of susceptibility between CC001 and CC071 mice was correlated with other flaviviruses. CC071 mice were highly susceptible to the HD78788 strain of the African lineage of ZIKV with 100% mortality, whereas all CC001 mice survived. After infection with WNV, all CC001 and CC071 mice succumbed but CC071 mice died more rapidly. The plasma viral load of CC071 mice was significantly higher than that of CC001 mice after infection with DENV.

*In vitro* infection in mouse embryonic fibroblasts (MEFs) was investigated to assess whether the differences in plasma viral load correlate with differences in levels of cellular viral replication. CC MEFs were infected at a multiplicity of infection (MOI) of 5 without MAR1-5A3 treatment, and viral titers in the supernatants were measured at 24, 48 and 72 hours post infection (hpi). While in CC001 MEFs, viral titer remained low and decreased between 48 and 72 hpi, CC071 MEFs showed increasing titer over time (Figure 15). Thus, CC071 MEFs were unable to control viral replication (Manet 2019; Manet et al. 2020).

Since IFN-I represent a first line of defense against viral infection, *Ifnb1* expression was measured in CC071, CC001 and B6 MEFs. While in CC001 and B6 MEFs, *Ifnb1* was highly expressed from 24 hpi, its induction was delayed in CC071 MEFs with a very low expression at 24 hpi (Figure 16). This delayed IFN-I induction is likely to explain the increased viral replication in CC071 cells. RNA sequencing analysis of infected CC071, CC001 and B6 MEFs showed that genes of the immune response were not upregulated in CC071, consistent with the delayed expression of *Ifnb1* (Manet 2019). Together, these results suggest a contribution of the delayed induction of IFN-I and the increased viral replication in infected cells to the susceptibility of the CC071 strain.



**Figure 15:** CC071 MEFs show increased viral replication. 3 biological replicates of CC001 and CC071 MEFs were infected with ZIKV at a MOI of 5. Viral titers in supernatants were quantified by focus-forming assay (from Manet et al. 2020).



**Figure 16:** CC071 MEFs show delayed *Ifnb1* expression. 3 biological replicates of CC001 and CC071 MEFs were infected with ZIKV at a MOI of 5. *Ifnb1* expression was measured with RT-qPCR (from Manet 2019).

## **III - Objectives**

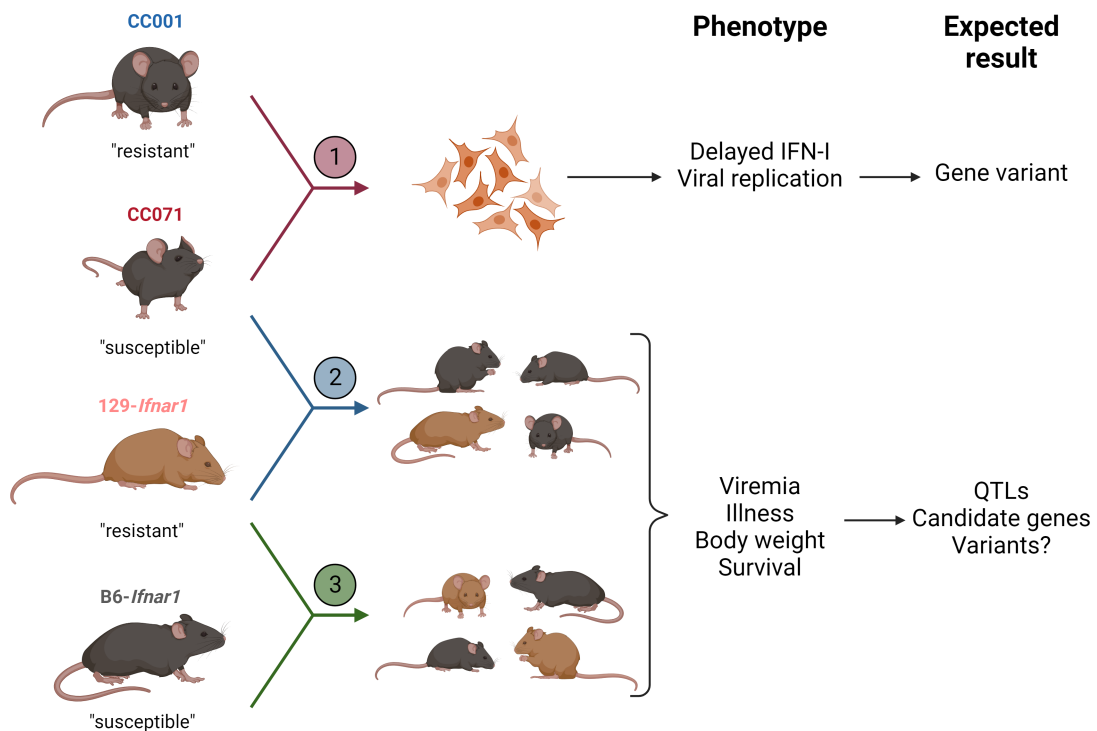
The objectives of my PhD project focused on elucidating the genetic control and the mechanisms supporting the differences in susceptibility to ZIKV in the mouse models studied in my laboratory.

The first objective was to identify the genetic factor(s) that control the delayed induction of IFN-I and uncontrolled viral replication in CC071 MEFs, using a combination of genetic and functional approaches.

The second objective was to assess whether the variant(s) identified in an *in vitro* model could explain the *in vivo* susceptibility of CC071 mice and, if not, to search for additional genes using genetic analysis of mouse crosses.

The third objective was to uncover modifiers of the *Ifnar1* KO that explain the different infection outcomes observed in B6-*Ifnar1*, which develop severe symptoms and show a high mortality rate post infection, and 129-*Ifnar1* mice, which develop only mild symptoms.

To meet these objectives, three crosses were produced and analyzed as illustrated in Figure 17. The experimental and analytical details, as well as the results from these experiments are presented in the following sections.



**Figure 17:** Objectives of my PhD project. Arrows represent the 2-generation crosses made between inbred strains. The numbers (1, 2 and 3) represent the three objectives as detailed above. Cells represent the objective studied *in vitro*, while mice represent objectives studied *in vivo*.

## **IV -Material and methods**



Material and methods of part V.1 of the results are detailed in the article "Susceptibility to Zika virus in a Collaborative Cross mouse strain is induced by *Irf3* deficiency *in vitro* but requires other variants *in vivo*". Material and methods of parts V.2, V.3, and V.4 complementary to those presented in part V.1 are detailed here.

## Mouse infection

### CC071 x (CC071 x B6-*Irf3*) N2

CC071 and B6 *Irf3* KO mice were bred at the Institut Pasteur. N2 mice were bred in all possible directions. Mice were phenotyped in 8 experimental batches. 7 to 9 week-old, male or female mice received an IP injection of 2 mg of MAR1-5A3 anti-IFNAR1 antibody (Euromedex, Cat#BX-BE0241) and were injected i.p. with  $10^7$  FFU of ZIKV FG15 the following day. Mice were followed for 7 dpi. Blood samples were collected at 2 and 6 dpi and brains were collected at 7 dpi to measure plasma and brain viral load, respectively.

### CC071 x 129-*Ifnar1* F2

CC071 and 129-*Ifnar1* mice were bred at the Institut Pasteur. Experiments were performed on ((129-*Ifnar1* × CC071) × (129-*Ifnar1* × CC071)) F2 mice (crosses are described as female × male). Mice were phenotyped in 11 experimental batches. 6 to 9 week-old, male or female mice were injected i.p. with  $10^7$  FFU of ZIKV FG15 and followed for 14 dpi. Blood samples were collected at 2 and 6 dpi. Only *Ifnar1*<sup>-/-</sup> (CC071 × 129-*Ifnar1*) F2 mice were used for this cross. Genotyping procedure is detailed below.

### (129-*Ifnar1* x B6-*Ifnar1*) x B6-*Ifnar1* N2

B6-*Ifnar1* and 129-*Ifnar1* mice were bred at the Institut Pasteur. Experiments were performed on (129-*Ifnar1* × B6-*Ifnar1*) × B6-*Ifnar1* backcross mice (crosses are described as female × male). Mice were phenotyped in 7 experimental batches. 6 to 8 week-old, male or female mice were injected i.p. with  $10^7$  FFU of ZIKV FG15 and followed for 14 dpi. Blood samples were collected at 2 and 6 dpi.

## Brain viral load quantification

Brain tissue samples were homogenized in 1mL of TRIzol reagent (Invitrogen Cat#15596026), using ceramic beads and an automated homogenizer (Qiagen Tissue Lyser II). RNAs were extracted with chloroform according to standard protocols. ZIKV RNAs were quantified by RT-qPCR using Luna Universal One-Step Kit (New England Biolabs Cat#E3005L) and the following primers: forward, 5'-CCG CTG CCC AAC ACA AG- 3', reverse, 5'-CCA CTA ACG TTC TTT TGC AGA CAT-3' and the probe 5'-6FAM-AGC CTA CCT TGA CAA GCA ATC AGA CAC TCA A-MGBEQ-3' (Eurofins).

### ***Ifnar1* genotyping**

(CC071 x 129-*Ifnar1*) F2 mice were genotyped for the *Ifnar1* KO by PCR. Tails were lysed in a buffer containing 1M of Tris, 1.5mM of NaCl and 10% of Tween20. PCR was performed on lysates using the following primers: WT allele forward, 5'-AAG ATG TGC TGT TCC CTT CCT CTG CTC TGA-3', WT allele reverse, 5'-ATT ATT AAA AGA AAA GAC GAG GCG AAG TGG-3', KO allele forward 5'-TCA GCG CAG GGG CGC CCG GTT CTT T-3', KO allele reverse 5'-ATC GAC AAG ACC GGC TTC CAT CCG A-3'.

### **Whole genome genotyping**

Genomic DNA was prepared from mouse tails using proteinase K digestion, phenol-chloroform extraction and ethanol precipitation according to standard protocols. Whole-genome genotyping was performed on 109 (CC071 x (CC071 x B6-*Irf3*)) N2s, 164 (CC071 x 129-*Ifnar1*) F2s and 186 ((129-*Ifnar1* x B6-*Ifnar1*) x B6-*Ifnar1*) N2s at Neogen (Neogen/Geneseek, Inc, Lincoln, NE, USA) using the MiniMUGA array containing 11,125 SNP markers. Raw genotypes were curated using the stuart package (Bourdon and Montagutelli 2022, see below).

### **QTL mapping**

QTL mapping was performed using R/qtl (Broman et al. 2003) using the EM algorithm in the `scanone` function. Survival and day of death were analyzed together with the two-part model. Plasma viral load, body weight loss and clinical scores were analyzed as normal quantitative traits. Credible intervals were calculated using the Bayesian method. The percentage of variance explained (PVE) was calculated with the following formula:  $1 - 10^{-2LOD/n}$ , with *LOD* the LOD score at the peak position of the QTL and *n* the number of phenotyped individual. QTLs were considered significant if the p-value was below 0.05, and suggestive if the p-value was below 0.63.

### **Candidate gene analysis**

Allelic data in the regions of interest were recovered from GenomeMUSter (Ball et al. 2023). The effects of the variants were then assessed using the Ensembl Variant Effect Predictor (VEP) tool (McLaren et al. 2016). Genes were filtered using the Phenotypes/Diseases criterion "abnormal immune system physiology" MP:0001790.

## Statistical analyses

Statistical analyses were performed using R version 4.1.0 (R Core Team 2021). Viremia were analyzed with mixed two-way ANOVA using the `anova_test` function of the `rstatix` package (Kassambara 2023). Strains were compared using the `between` argument and times post infection were compared using the `within` argument. Body weight and clinical score curves were analyzed with linear mixed models using the `lme4` (Bates et al. 2015) and `emmeans` (Lenth 2023) packages. Correlation matrixes were created using the `ggcorrplot` package.

## Development of the stuart R package

To facilitate the analysis of F2 and N2 crosses, I have developed `stuart`, an open-source R package. `stuart` curates whole-genome genotyping data from second-generation individuals produced by SNP arrays, based on informativeness, Mendelian inheritance and consistency with parental genotypes. `stuart` then formats the data for subsequent QTL mapping with widely used R package `R/qtl`. `stuart` is available at <https://gitlab.pasteur.fr/mouselab/stuart/>. We have published a paper describing its functionalities in the journal *G3* (Bourdon and Montagutelli 2022).

### Title

`stuart`: an R package for the curation of SNP genotypes from experimental crosses

### Abstract

Genetic mapping in 2-generation crosses requires genotyping, usually performed with single nucleotide polymorphism markers arrays which provide high-density genetic information. However, genetic analysis on raw genotypes can lead to spurious or unreliable results due to defective single nucleotide polymorphism assays or wrong genotype interpretation. Here, we introduce `stuart`, an open-source R package, which analyzes raw genotyping data to filter single nucleotide polymorphism markers based on informativeness, Mendelian inheritance pattern, and consistency with parental genotypes. The functions of this package provide a curation pipeline and formatting adequate for genetic analysis with the `R/qtl` package. `stuart` is available with detailed documentation from <https://gitlab.pasteur.fr/mouselab/stuart/>.

# stuart: an R package for the curation of SNP genotypes from experimental crosses

Marie Bourdon , Xavier Montagutelli \*

Mouse Genetics Laboratory, Institut Pasteur, Université Paris Cité, F-75015 Paris, France

\*Corresponding author: Mouse Genetics Laboratory, Institut Pasteur, Université Paris Cité, F-75015 Paris, France. Email: [xavier.montagutelli@pasteur.fr](mailto:xavier.montagutelli@pasteur.fr)

## Abstract

Genetic mapping in 2-generation crosses requires genotyping, usually performed with single nucleotide polymorphism markers arrays which provide high-density genetic information. However, genetic analysis on raw genotypes can lead to spurious or unreliable results due to defective single nucleotide polymorphism assays or wrong genotype interpretation. Here, we introduce stuart, an open-source R package, which analyzes raw genotyping data to filter single nucleotide polymorphism markers based on informativeness, Mendelian inheritance pattern, and consistency with parental genotypes. The functions of this package provide a curation pipeline and formatting adequate for genetic analysis with the R/qtl package. stuart is available with detailed documentation from <https://gitlab.pasteur.fr/mouselab/stuart/>.

**Keywords:** R package; genetic analysis; SNP genotypes

## Introduction

Genetic mapping of Mendelian or quantitative traits in inbred strains is classically achieved in 2-generation crosses such as intercrosses (F<sub>2</sub>) and backcrosses (N<sub>2</sub>), in which the inheritance of the trait is compared with the genotypes at multiple genetic markers encompassing the genome map. Variations of a quantitative trait are controlled by one or more quantitative trait loci (QTLs). A QTL is defined as a marker at which individuals carrying different genotypes show different average trait values. QTL mapping searches for QTLs by testing the association between trait values and genotypes at markers spanning the genome map. The statistical significance of the association is expressed as logarithm of the odds (LOD) score which is calculated for each genotyped marker and, at intermediate positions, for pseudo-markers created by interval mapping, generating an LOD score curve (Broman 2001). The curve peaks at regions potentially associated with the trait. These peaks are called QTLs if they reach predefined statistical thresholds established either from general statistical models (Lander and Kruglyak 1995) or by permutation tests performed on the cross data. For each permutation, phenotypes are shuffled between individuals to break real associations, and LOD scores are calculated to identify peaks, which are all false positives. The distribution of the peak LOD scores over a large number (>1,000) of permutations provides statistical thresholds: if a LOD score of 3.8 or higher is observed in 5% of the permutations, this value will be taken as the  $P = 0.05$  threshold (Doerge and Churchill 1996). QTL mapping on F<sub>2</sub>s and N<sub>2</sub>s can be conducted with R packages such as R/qtl (Broman et al. 2003) and R/qtl2 (Broman et al. 2019).

With genome sequencing, single nucleotide polymorphisms (SNPs) have become the standard across species for their very

high frequency, low cost, and high-throughput analysis using various genotyping platforms. In mice, several generations of Mouse Universal Genotyping Arrays (MUGA) have been developed, the most recent being GigaMUGA (143k SNPs; Morgan et al. 2015) and MiniMUGA (10.8k SNPs; Sigmon et al. 2020). GigaMUGA provides high-density coverage for the fine characterization of inbred strains or outbred populations such as the Diversity Outbred (Svenson et al. 2012), while the modest number of SNPs in MiniMUGA is largely sufficient to genotype intercrossed or backcrossed individuals. However, SNP reliability is affected by the performance of genotyping platforms and polymorphism between and within inbred strains. Spurious or unreliable mapping outputs can result from defective SNP assays or wrong genotype interpretations. Therefore, raw data obtained from genotyping services must be curated before performing genetic analyses.

Several tools exist for quality control of SNP genotyping arrays, including Illumina's GenomeStudio. R packages such as argyle (Morgan 2015) analyze hybridization intensity signals from MUGA arrays. The simple genetic structure of 2-generation crosses provides specific and efficient means for identifying spurious genotyping data, such as consistency with parental genotypes and expected Mendelian proportions. The R/qtl package includes functions to build genetic maps and check for genotype consistency (<https://rqt.org/tutorials/geneticmaps.pdf>). However, this control is performed once genotypes have been imported and involves multiple steps of manual curation. To provide a more automated process of data curation before genetic analysis, we have developed stuart, an R package that implements a pipeline for automatic filtering and curation of SNP genotyping data from 2-generation crosses based on simple rules. This package formats raw SNP allele calls from Illumina files into genotypes ready for importation in R/qtl. Using 3 intercross datasets, we illustrate the consequences of inconsistent

Received: July 5, 2022. Accepted: August 19, 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of Genetics Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

genotypes on the estimated marker map and QTL mapping, and how the curation achieved by each function in *stuart* leads to trustworthy results.

## Materials and methods

*stuart* is a tidyverse (Wickham et al. 2019) based R package requiring R version 3.5.0 or later. Its open source is available on Institut Pasteur's GitLab: <https://gitlab.pasteur.fr/mouselab/stuart/> and can be installed with devtools (Wickham et al. 2021). *stuart*'s vignette provides detailed descriptions of data import and of each function. *stuart* imports SNP allele calls from MUGA Illumina platform or other sources using the same file format. The central object of *stuart* is the marker table which summarizes for each marker, the alleles found in the population, the number of individuals of each genotype and the exclusion status resulting from the curation steps. *stuart* exports curated data to an R/qtl compatible format. The SNP annotation file used was downloaded from [https://raw.githubusercontent.com/kbroman/MUGAarrays/master/UWisc/mini\\_uwisc\\_v2.csv](https://raw.githubusercontent.com/kbroman/MUGAarrays/master/UWisc/mini_uwisc_v2.csv) (last accessed August 29, 2022).

Three datasets were used to test the package. This article presents the results from 176 (CC001/Unc X C57BL/6J-*Ifnar1* KO) F2 mice (dataset 1). The analysis of 2 other data sets, 94 (C57BL/6J-*Ifnar1* KO X 129S2/SvPas-*Ifnar1* KO) F2 mice (dataset 2) and 89 (C57BL/6NCRl X CC021/Unc) F2 mice (dataset 3) is presented as [Supplementary data](#). Quantitative traits were studied in the 3 F2s. Phenotype distributions are presented in [Supplementary Fig. 1](#). Genotyping was performed by Neogen (Auchincruive, Scotland) with MiniMUGA on DNA prepared from tail biopsies using standard phenol-chloroform extraction. Genotype call rate was 0.927, 0.931, and 0.948 for dataset 1, dataset 2, and dataset 3, respectively. QTL mapping was performed using R/qtl. Statistical significance of phenotype-genotype association was computed by data permutation (Doerge and Churchill 1996), which provides genome-wide thresholds accounting for multiple testing. The following thresholds were used, as commonly accepted (Members of the Complex Trait Consortium 2003):  $P = 0.05$  for significant association,  $P = 0.1$  and  $P = 0.63$  for the suggestive association. All figures were designed with ggplot2 (Wickham 2016) or R/qtl.

## Results and discussion

### Consequences of inconsistent genotypes

SNP data delivered by the Illumina platform are base alleles that need to be translated into genotypes for genetic analysis. From our experience on multiple 2-generation crosses, we identified several types of genotype inconsistencies that were responsible for distorted marker maps and spurious QTL mapping results. Recombination fraction (RF), which measures the genetic distance between 2 markers, is estimated in a cross by analyzing the proportion of recombinants between adjacent markers in all individuals. The map of markers calculated from the cross data should be consistent with their known positions. The R/qtl `est-map()` and `plotMap()` functions produce a graphical comparison of the 2 maps (Fig. 1a and [Supplementary Fig. 2, a and b](#)). For each chromosome, the known position of each marker provided in the annotation file (left) is connected with the estimated position (right) based on observed RF. With minimally curated genotypes (exclusion of nonpolymorphic markers and markers with over 50% missing genotypes), large RF was found in many instances between closely linked markers, resulting in fan-like patterns. To further describe these distortions, we computed the

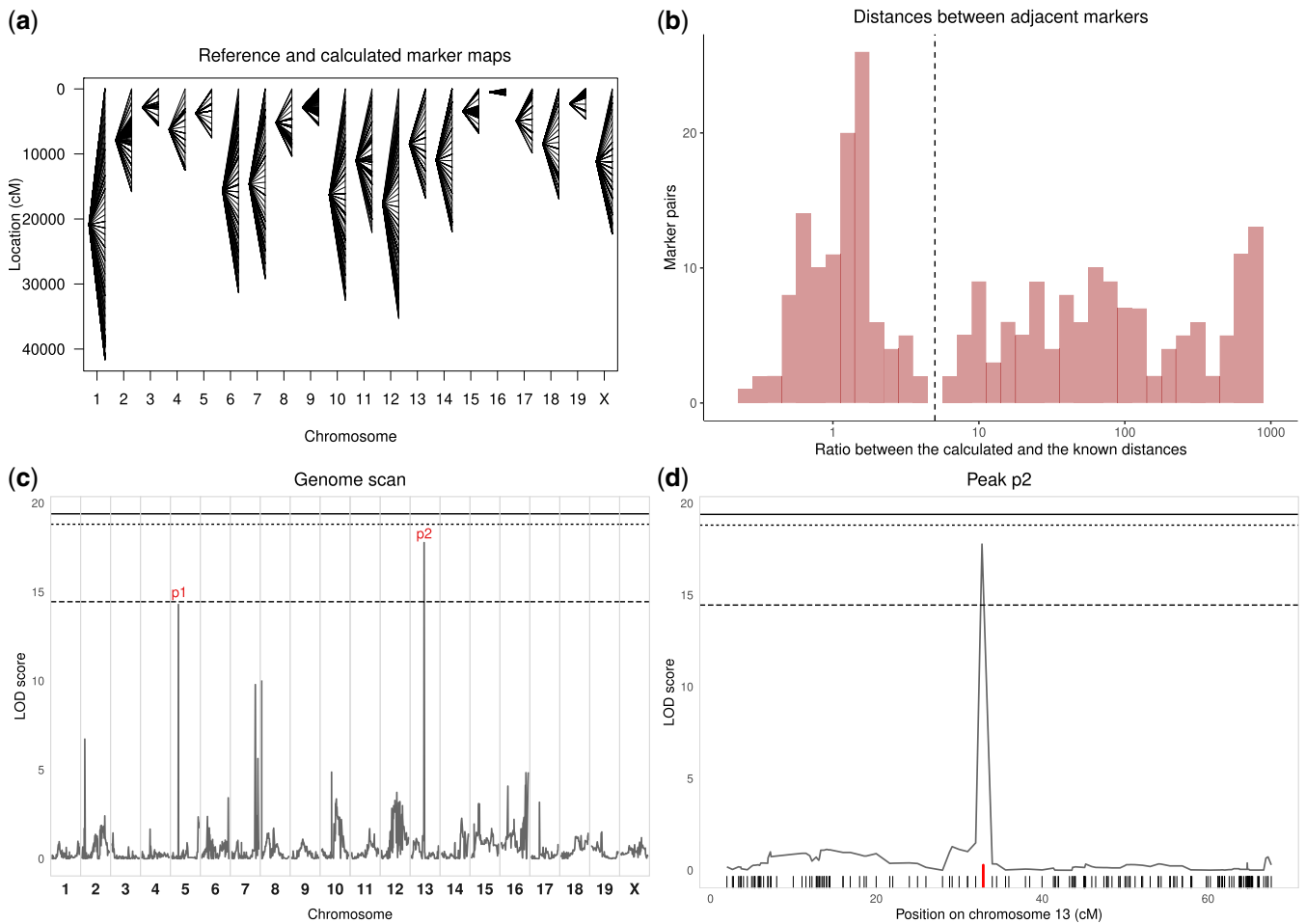
distribution of the ratio between the calculated and the known genetic distances between adjacent markers (Fig. 1b and [Supplementary Fig. 2, c and d](#)); to avoid exaggerated ratios, we considered only markers with a known distance of 1 cM or more). This analysis revealed 2 groups of markers. In dataset 1, for 43% of them, the ratio was below 5 and followed a Gaussian distribution with mean = 1.31 and SD = 0.77. The other markers (57%) showed a ratio between 5 and 981.87 (Fig. 1b) which necessarily results from incorrect genotypes, as only a few individuals should show recombination between adjacent markers. On chromosome 1, while the known marker positions spanned ~100 cM, the cumulated genetic distance estimated from observed RF was ~40,000 cM. As QTL mapping relies on coherent genotypes at a series of markers encompassing a genetic interval, problematic genotypes at a given marker will perturb the analysis and, in some cases, may result in peaks of the LOD score curve in the absence of true association (Cheung et al. 2014). Such false positives increase significance thresholds calculated by data permutation.

These 2 consequences of genotyping inconsistencies are illustrated in Fig. 1c and [Supplementary Fig. 3, a and b](#) which were obtained using the R/qtl `scanone()` function on a quantitative trait from the uncured F2 datasets. For dataset 1, the  $P = 0.05$  significance threshold was estimated at 19.4 (Fig. 1c), while it usually ranges between 3.3 and 4.3 depending on the inheritance model for crosses of this type and size (Lander and Kruglyak 1995). Several peaks were detected although none reached  $P = 0.05$  significance. Moreover, their narrow profile was highly unexpected in F2 crosses. Indeed, these peaks involved only 1–3 markers, and the LOD score curve felt abruptly between these and adjacent markers on both sides (Fig. 1d), while genetic linkage between closely linked markers should result in progressive decrease of the LOD score curve on both sides of a peak (Guénet et al. 2015). Among the 3 datasets, we identified 4 narrow peaks reaching suggestive significance level ( $P < 0.63$ ): 2 were located at a marker with non-Mendelian allelic proportions and 2 were located at 1–3 pseudomarkers adjacent to a marker with non-Mendelian proportions ([Supplementary Fig. 3, c, d and e, f](#), respectively). We identified 5 other narrow peaks (LOD score between 6.72 and 10.03) out of which 4 resulted from the same situations as above and one was located on a pseudomarker and a marker with non-Mendelian proportions.

Inconsistent marker maps may also originate from the wrong assignment of markers to their chromosome and position provided to the mapping program. Indeed, R/qtl developer K. Broman identified errors in MUGA arrays annotation files affecting marker positions, probe sequences mapping to several locations, and unmappable markers. We recommend using K. Broman's corrected annotation files available on GitHub. The conversion of SNP alleles (A, C, T, G) observed in second-generation individuals (SGIs) to genotypes encoded according to the parental alleles may also create genotype errors. Reference SNP alleles established for many mouse strains may be used to infer the SGI genotypes. However, we recommend genotyping individuals of the parental strains used in the cross since they could differ from the reference panel. In our example dataset, the 2 parental strains used in the cross showed allelic differences with their reference panel counterpart at 200 markers.

### Data control and curation performed in *stuart*

Although each of *stuart*'s functions can be called independently, we present a logical analysis workflow appropriate for 2-generation crosses. [Table 1](#) summarizes the data curation and filtering



**Fig. 1.** Analysis of the dataset 1 illustrating the consequences of genotyping errors and inconsistencies on QTL mapping. Nonpolymorphic markers and markers with more than 50% missing genotypes were excluded to avoid excessive calculation time. a) Comparison of the known marker map (left) and the genetic map estimated from observed RF (right), as calculated by `est.map()` and represented by `plotMap()` functions of R/qtl. Lines connect the positions of each marker in the 2 maps. The estimated map is considerably expanded because of multiple genotype inconsistencies. b) Distribution of the ratio between estimated and known distances between adjacent markers. Markers with known and calculated distances below 1 cM were removed as they may lead to extremely small or large ratios. The expansion of the estimated map leads to a distribution tail of high ratios. The y-axis is in logarithmic scale. Fifty-seven percent of markers have a ratio above 5 (dashed line). c) Output of the scanone function of R/qtl showing the identification of narrow LOD score peaks. Genome-wide significance thresholds computed by data permutation are shown as plain ( $P = 0.05$ ), dotted ( $P = 0.1$ ), and dashed ( $P = 0.63$ ) lines. d) Magnification of the scanone plot restricted to chromosome 13 (peak p2). The LOD score peak is located on one marker (red tick) distant by 1.728 and 1.24 cM from the proximal and distal markers, respectively, on the known marker map, but by 1,001.582 and 1,001.506 cM based on calculated RF.

performed by each function, and the number of markers of dataset 1 retained after each step.

### Data importation

Genetic mapping requires both genotype and phenotype data. Required formats and instructions are detailed in the vignette (see example of phenotype data in [Supplementary Table 1](#)). Parental strains' genotyping data can be loaded from the same genotyping results as the SGI, from a previous genotyping file or from a reference file. Annotation data from K. Broman can be imported directly from GitHub. The `geno_strains()` function formats parental genotypes from a 2-allele encoding in Illumina format into a single letter encoding, and merges these data with the annotation table into a table with parental allele and marker positions.

### Consistency between parents and SGI alleles and genotypes

Several generations of MUGA arrays have been developed (Mega, Giga, Mini), each with successive versions differing by multiple

SNP markers. If parental and SGI data were produced on different versions, the marker lists must be compared to retain only common SNPs. This is achieved by the `mark_match()` function.

Converting alleles into genotypes requires that SGI segregate for the 2 parental alleles, and that each allele is found only in one parent. The aim of the `mark_allele()` function is to control consistency of allele's origin at multiple levels.

First, this function excludes markers with missing data in both parents. If allele data are available for only one parent and this allele is also found in SGI, the other allele present in SGI will be assigned to the parent with missing allele. However, this imputation is not error-free since we have observed, in rare occasions, markers which alleles were identical in the parental strains but were polymorphic in the SGI (Table 2 for such SNPs in dataset 1). This situation may occur when the parental strains used in the cross have diverged from those of the reference panel, or if one parent is heterozygous. Such markers will be excluded by the `mark_allele()` function but they could escape detection if allele information was missing in one parent.

**Table 1.** stuart analysis pipeline and application to dataset 1.

Steps	Function	Excluded markers	Number of markers retained
1. Import SGI alleles from MUGA arrays	read.table()/read_tsv()	–	11,125
2. Add data from parental strain	geno_strains()	–	–
Genotyped with SGI: make consensus	read.table()/read_tsv(), geno_strains	–	–
Imported from another dataset: import and make consensus	read.table()/other readr function depending on the format	–	–
3. Filter on allele consistency between parents and SGI	mark_match()	Not present in both parents and SGI	11,125
Same set of markers between parents and SGI	mark_allele	Missing alleles in both parents	10,375
Alleles consistent between parents and SGI		Not polymorphic in parents but polymorphic in SGI Different alleles in parents and SGI In backcrosses: homozygotes for the wrong allele Optional: one parent missing or heterozygous	
4. Exclude markers with high proportion of missing genotypes	mark_na()	>50% of missing genotypes by default	9,918
5. Exclude nonpolymorphic markers in SGI	mark_poly()	Nonpolymorphic in SGI	2,738
6. Verify Mendelian proportions	mark_prop()	Departure from expected Mendelian segregation (proportion of each class or statistical threshold)	2,254
7. Verify RF between markers	est.map() followed by mark_estmap()	High RFs with adjacent markers	2,251

**Table 2.** Markers of dataset 1 non polymorphic between parental strains but polymorphic in SGI.

Marker	Allele parent 1	Allele parent 2	Allele SGI 1	Allele SGI 2
S6J017555686	C	C	T	C
S6J113080150	G	G	A	G
gAX00038569	C	C	T	C
mUNC21540855	C	C	A	C
gUNC21555204	T	T	T	C
gUNC21596600	A	A	A	G

Adding the `parNH = FALSE` argument to the `mark_allele()` function will exclude markers missing one parental allele or for which one parent is heterozygous. However, while preventing rare errors, this option will also exclude a number of truly informative markers.

The `mark_allele()` function also discards markers at which parents and SGI carry different alleles, and, for backcrosses, markers for which some SGI are homozygous for the wrong allele.

### Nonpolymorphic markers

Genetic analysis requires polymorphic markers, i.e. for which parents carry different alleles which segregate in the SGI. The `mark_poly()` function excludes markers for which all genotyped SGI carry the same allele, which saves computation time.

### Missing genotypes

Reliable QTL mapping results depend on markers with medium to high rate of successful genotyping. Figure 2a shows markers distribution based on the proportion of missing genotypes. For over 95% of markers genotyping rate was above 50%. Genotyping

failures may result from poor-quality genotyping assay. The `mark_na()` function excludes such poorly genotyped markers.

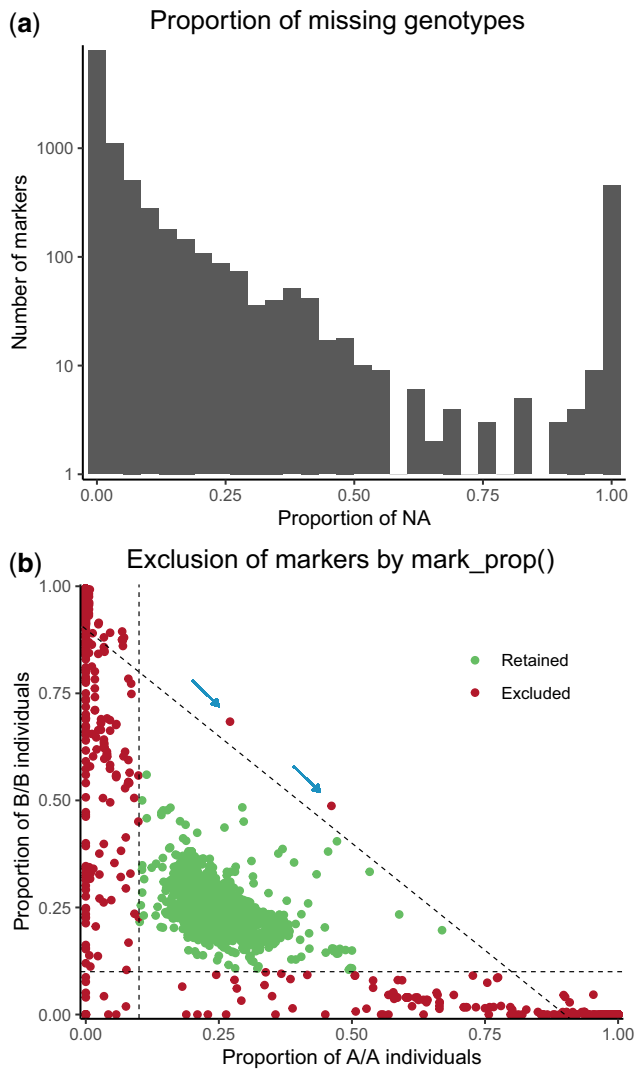
### Mendelian proportions

In 2-generation crosses between inbred strains, the proportions of the 2 or 3 classes of genotypes are predictable, i.e. for autosomes, 25% of each type of homozygotes and 50% of heterozygotes in an intercross, and 50% of homozygotes and 50% of heterozygotes in a backcross. Comparing the observed proportions with these expectations provides another criterion of filtering.

The `mark_prop()` function filters markers based either on a minimum proportion of each genotype or on the statistically significant departure from the expected proportions (Chi<sup>2</sup> test, with a *P*-value threshold). Figure 2b shows the exclusions of the autosomal markers depending on the proportion of each genotype. X chromosome genotypic proportions differ from autosomes, therefore, different arguments of `mark_prop()` function are used to filter X-linked markers for more precise curation.

### Filtering report and impact on QTL mapping results

At every step, the markers filtered out are annotated in a marker table which can be exported for further inspection. The last column of Table 1 shows the number of markers retained after each step in the example dataset 1. Most of the starting markers (7,180/11,125 = 65%) which were eventually removed by stuart's functions were removed by `mark_poly()` as nonpolymorphic, a ratio expected for crosses between 2 standard mouse inbred strains (Frazer et al. 2007). `mark_allele()` rejected 750 markers, `mark_na()` 457 and `mark_prop()` 484. Across the 3 datasets, we found 1,546 markers with either non-Mendelian proportions or allele inconsistencies between parental strains and SGIs. Overall, 619 of them were retained by stuart's filtering in at least one of the



**Fig. 2.** a) Distribution of the markers by their proportion of missing genotype (NA) in dataset 1. The y-axis is in logarithmic scale. 4.63% of markers have >50% missing genotypes. b) Exclusion of markers depending on genotypic proportions in dataset 1. Markers on X and Y chromosomes and mitochondrial DNA are not represented. The 2 axes represent the proportions of the 2 types of homozygous individuals in the intercross: AA and BB. Each dot represents a marker. Markers were excluded if the proportion of at least one of the 3 genotypes (AA, AB, and BB) was less than 10%, i.e. outside the triangle defined by the 3 dashed lines ( $AA = 0.1$ ,  $BB = 0.1$ , and  $AA + BB = 0.9$ ). Arrows point at 2 markers excluded due to a proportion of heterozygotes <10%.

other crosses, ruling out their misassignment to the genetic map. Out of the residual markers, 85 were removed from all datasets for another criterion than absence of polymorphism and were therefore considered as unreliable.

At this step, the dataset may still contain markers showing high RFs with adjacent markers either for a reason not tested by the current version of *stuart* or due to the parameters used in `mark_na()` and `mark_prop()` functions. These markers can be identified by calculating the estimated map using `R/qtl est.map()` and using *stuart*'s `mark_estmap()` function which excludes markers presenting high RFs with adjacent markers. Over the 3 datasets, 9 markers were removed by `mark_estmap()`. Five of them were retained in at least one other dataset, indicating the problem was dataset specific. Finally, for dataset 1, 2,251 markers passed all steps resulting in an average genetic interval between

adjacent markers lower than 2 cM, which is largely sufficient to perform QTL mapping (Darvasi et al. 1993). After curation, phenotype and genotype data are combined and exported in the R/qtl format using the `write_rqtl()` function. The `qtl2convert` package (Broman 2021) converts this output into the adequate format required by the more recent R/qtl2 package.

Figure 3a and Supplementary Fig. 4, a and b show the marker maps calculated after data curation with *stuart*. The known marker map and the estimated genetic map are consistent, with minimal expansions or contractions. Large ratios between the calculated and the known genetic distances between adjacent markers have been eliminated (Fig. 3b, Supplementary Fig. 4, c and d). QTL mapping analysis on curated dataset 1 is shown on Fig. 3c (to be compared with Fig. 1c; see Supplementary Fig. 5 for datasets 2 and 3). LOD thresholds are in the expected range for an F2, and the LOD score curve reveals broader peaks than in Fig. 1b, with progressive LOD score decrease on both sides of the peak marker. One significant and 3 suggestive QTLs were identified on chromosomes 12 ( $P$ -value=0.037, Fig. 3d), 5 ( $P$ -value=0.460), 10 ( $P$ -value=0.157), and 15 ( $P$ -value=0.244) which were not visible using noncurated data due to very high LOD score thresholds.

Being very simple to use and efficient at curating genotyping errors, *stuart* will facilitate the use of genotyping arrays for genetic mapping purposes in 2-generation crosses, bridging the gap between raw allele data produced by SNP platforms and genetic analysis software. Moreover, its functions can be used independently to analyze inbred strains genotypes. For example, `geno_strain()` creates a genotype consensus between 2 or more individuals of the same strain suitable for further inspection, which can be useful when genotyping or re-genotyping a strain of interest. Comparing genotyping results of an inbred strain after several generations of breeding with `mark_allele()` will readily identify variants that have emerged or been selected over time. Likewise, this function will help identifying genetic variants between substrains.

## Web resources

The source code of the *stuart* package and the code used for the figures of this article are publicly available from <https://gitlab.pasteur.fr/mouselab/stuart/>.

## Data availability

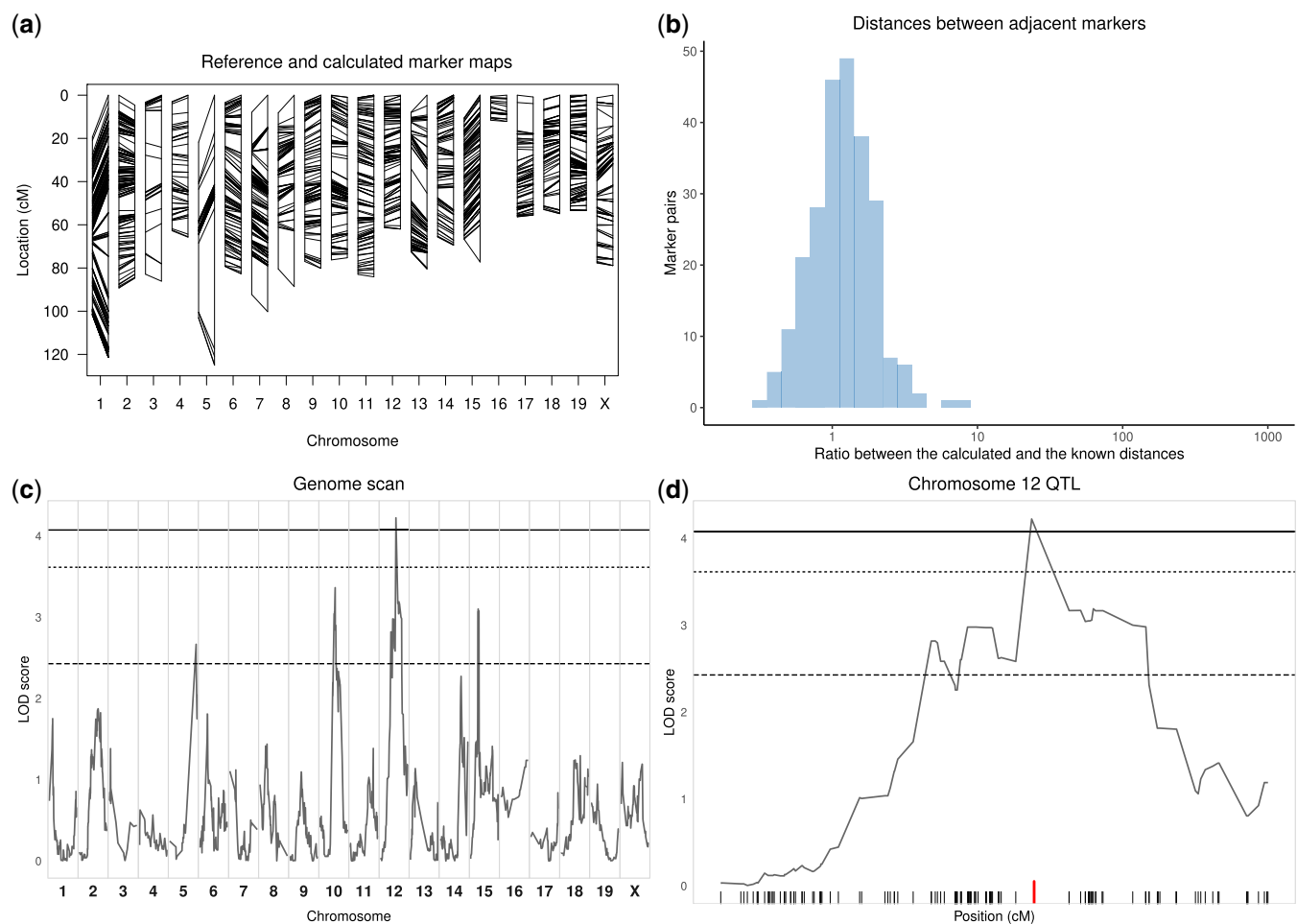
All datasets used as examples in this article are available from <https://gitlab.pasteur.fr/mouselab/stuart/>. Dataset 1 is included in the package and can be loaded once the package is loaded (see the vignette for details). The 2 other datasets are available from GitLab in the "article" directory in separate folders (i.e. "data2" and "data3"). Each folder contains the genotypes of the SGIs in file "geno\_dataX.csv," the phenotypes of the SGIs in file "pheno\_dataX.csv," the parental strains' genotypes in file "parents\_dataX.csv" and the reference genotypes for the parental strains in file "ref\_gen\_dataX.csv." Analysis of each cross is in each folder in an R markdown file ("dataX.Rmd").

Supplemental material is available at G3 online.

## Acknowledgments

We thank Elise Jacquemet of the Pasteur Institute Bioinformatics and Biostatistics HUB for helping with the use of GitLab.





**Fig. 3.** Analysis of dataset 1 after curation of genotyping data by Stuart using the `mark_match()`, `mark_allele()`, `mark_na()`, `mark_poly()`, `mark_prop()`, and `mark_estmap()` functions. Refer to Fig. 1 for comparison with original data. a) The estimated marker map is now consistent with the known marker map. Despite some contraction or expansion of specific intervals, the genome length of the observed marker map for each chromosome is consistent with the known map (ratio between the calculated and the known length of the genome: 1.12). b) The distribution of the ratio between estimated and known distance between adjacent markers. Markers with known and calculated distances below 1 cM were removed as they may lead to extremely small or large ratios. Ratios are normally distributed with mean = 1.33 and SD = 0.81 showing consistency between the known and estimated maps. c) The LOD score curve shows several peaks, one of which is significant at  $P < 0.05$  (plain line, genome-wide significance computed by data permutation). Note that the significance thresholds are much lower than in Fig. 1c. None of the peaks shown in Fig. 1c were confirmed after data curation. Conversely, none of the peaks above  $P = 0.63$  (dashed line) found after data curation had been detected in Fig. 1c. d) The magnification of the QTL peak identified on chromosome 12, showing progressive decrease of the LOD score curve over a large genetic interval. The marker with the highest LOD score is identified with a thick (red) tick.

## Funding

This project was funded by the French Government's Investissement d'Avenir programme, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant n° ANR-10-LABX-62-IBEID).

## Conflicts of interest

None declared.

## Literature cited

Broman KW. Review of statistical methods for QTL mapping in experimental crosses. *Lab Anim (NY)*. 2001;30(7):44–52.  
 Broman KW. `qtl2convert`: Convert Data among QTL Mapping Packages. 2021. [accessed 2022 May 15]. <https://CRAN.R-project.org/package=qtl2convert>.

Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen S, Yandell BS, Churchill GA. `R/qtl2`: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics*. 2019;211(2):495–502. doi:10.1534/genetics.118.301595.  
 Broman KW, Wu H, Sen S, Churchill GA. `R/qtl`: QTL mapping in experimental crosses. *Bioinformatics*. 2003;19(7):889–890. doi:10.1093/bioinformatics/btg112.  
 Cheung CYK, Thompson EA, Wijsman EM. Detection of Mendelian consistent genotyping errors in pedigrees: detection of genotyping errors. *Genet Epidemiol*. 2014;38(4):291–299. doi:10.1002/gepi.21806.  
 Darvasi A, Weinreb A, Minke V, Weller JI, Soller M. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics*. 1993;134(3):943–951. doi:10.1093/genetics/134.3.943.  
 Doerge RW, Churchill GA. Permutation tests for multiple loci affecting a quantitative character. *Genetics*. 1996;142(1):285–294. doi:10.1093/genetics/142.1.285.  
 Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, Gupta RV, Montgomery J, Morenzoni MM, Nilsen GB, et al. A sequence-

- based variation map of 8.27 million SNPs in inbred mouse strains. *Nature*. 2007;448(7157):1050–1053. doi:[10.1038/nature06067](https://doi.org/10.1038/nature06067).
- Guénet JL, Benavides F, Panthier J-J, Montagutelli X. 2015. Genetics of the Mouse. [accessed 2022 Jun 24]. <https://link.springer.com/book/10.1007/978-3-662-44287-6>.
- Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 1995; 11(3):241–247. doi:[10.1038/ng1195-241](https://doi.org/10.1038/ng1195-241).
- Members of the Complex Trait Consortium. The nature and identification of quantitative trait loci: a community's view. *Nat Rev Genet*. 2003;4(11):911–916. doi:[10.1038/nrg1206](https://doi.org/10.1038/nrg1206).
- Morgan AP. argyle: an R package for analysis of illumina genotyping arrays. *G3 (Bethesda)*. 2015;6(2):281–286. doi:[10.1534/g3.115.023739](https://doi.org/10.1534/g3.115.023739).
- Morgan AP, Fu C-P, Kao C-Y, Welsh CE, Didion JP, Yadgary L, Hyacinth L, Ferris MT, Bell TA, Miller DR, et al. The mouse universal genotyping array: from substrains to subspecies. *G3 (Bethesda)*. 2015;6(2):263–279. doi:[10.1534/g3.115.022087](https://doi.org/10.1534/g3.115.022087).
- Sigmon JS, Blanchard MW, Baric RS, Bell TA, Brennan J, Brockmann GA, Burks AW, Calabrese JM, Caron KM, Cheney RE, et al. Content and performance of the MiniMUGA genotyping array: a new tool to improve rigor and reproducibility in mouse research. *Genetics*. 2020;216(4):905–930. doi:[10.1534/genetics.120.303596](https://doi.org/10.1534/genetics.120.303596).
- Svenson KL, Gatti DM, Valdar W, Welsh CE, Cheng R, Chesler EJ, Palmer AA, McMillan L, Churchill GA. High-resolution genetic mapping using the mouse diversity outbred population. *Genetics*. 2012;190(2):437–447. doi:[10.1534/genetics.111.132597](https://doi.org/10.1534/genetics.111.132597).
- Wickham H. ggplot2: elegant Graphics for Data Analysis. 2016. [accessed 2022 May 15]. <https://ggplot2.tidyverse.org>.
- Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. Welcome to the tidyverse. *J Open Source Softw*. 2019;4(43):1686. doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wickham H, Hester J, Chang W, Bryan J. devtools: Tools to Make Developing R Packages Easier. 2021. [accessed 2022 May 15]. <https://devtools.r-lib.org>.

Communicating editor: F. P.-M. de Villena

## V - Results

## V.1 . The delayed induction of IFN-I in CC071 cells is due to a loss-of-function mutation in the *Irf3* gene

The results of the first objective of my thesis were the subject of a paper accepted for publication in *PLOS Pathogens* (Bourdon et al. 2023).

### Title

Susceptibility to Zika virus in a Collaborative Cross mouse strain is induced by *Irf3* deficiency *in vitro* but requires other variants *in vivo*.

### Abstract

Zika virus (ZIKV) is a Flavivirus responsible for recent epidemics in Pacific Islands and in the Americas. In humans, the consequences of ZIKV infection range from asymptomatic infection to severe neurological disease such as Guillain-Barré syndrome or fetal neurodevelopmental defects, suggesting, among other factors, the influence of host genetic variants. We previously reported similar diverse outcomes of ZIKV infection in mice of the Collaborative Cross (CC), a collection of inbred strains with large genetic diversity. CC071/TauUnc (CC071) was the most susceptible CC strain with severe symptoms and lethality. Notably, CC071 has been recently reported to be also susceptible to other flaviviruses including dengue virus, Powassan virus, West Nile virus, and to Rift Valley fever virus. To identify the genetic origin of this broad susceptibility, we investigated ZIKV replication in mouse embryonic fibroblasts (MEFs) from CC071 and two resistant strains. CC071 showed uncontrolled ZIKV replication associated with delayed induction of type-I interferons (IFN-I). Genetic analysis identified a mutation in the *Irf3* gene specific to the CC071 strain which prevents the protein phosphorylation required to activate interferon beta transcription. We demonstrated that this mutation induces the same defective IFN-I response and uncontrolled viral replication in MEFs as an *Irf3* knock-out allele. By contrast, we also showed that *Irf3* deficiency did not induce the high plasma viral load and clinical severity observed in CC071 mice and that susceptibility alleles at other genes, not associated with the IFN-I response, are required. Our results provide new insight into the *in vitro* and *in vivo* roles of *Irf3*, and into the genetic complexity of host responses to flaviviruses.

## RESEARCH ARTICLE

# Susceptibility to Zika virus in a Collaborative Cross mouse strain is induced by *Irf3* deficiency *in vitro* but requires other variants *in vivo*

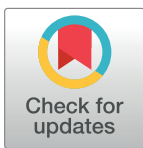
Marie Bourdon<sup>1</sup>✉‡, Caroline Manet<sup>1</sup>✉‡, Laurine Conquet<sup>1</sup>, Corentin Ramaugé Parra<sup>1</sup>, Etienne Kornobis<sup>2</sup>, Eliette Bonnefoy<sup>3</sup>, Xavier Montagutelli<sup>1</sup>\* 

**1** Institut Pasteur, Université Paris Cité, Mouse Genetics Laboratory, Paris, France, **2** Institut Pasteur, Université Paris Cité, Biomics, Paris, France, **3** Université Paris Cité, Institut Cochin, Inserm, CNRS, Paris, France

✉ These authors contributed equally to this work.

‡ These authors share first authorship on this work.

\* [xavier.montagutelli@pasteur.fr](mailto:xavier.montagutelli@pasteur.fr)



## OPEN ACCESS

**Citation:** Bourdon M, Manet C, Conquet L, Ramaugé Parra C, Kornobis E, Bonnefoy E, et al. (2023) Susceptibility to Zika virus in a Collaborative Cross mouse strain is induced by *Irf3* deficiency *in vitro* but requires other variants *in vivo*. PLoS Pathog 19(9): e1011446. <https://doi.org/10.1371/journal.ppat.1011446>

**Editor:** Ashley L. St. John, Duke-National University of Singapore, SINGAPORE

**Received:** May 30, 2023

**Accepted:** September 5, 2023

**Published:** September 21, 2023

**Copyright:** © 2023 Bourdon et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** RNAseq primary data are available from the European Nucleotide Archive (<https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-12765>). All other experimental data that support the findings of this study are included as [supplemental information](#).

**Funding:** This work was supported by grants from the Agence Nationale de la Recherche (ANR) NeuroZika (ANR-20-CE16-0017 to XM and EB) and from the French Government's Investissement

## Abstract

Zika virus (ZIKV) is a Flavivirus responsible for recent epidemics in Pacific Islands and in the Americas. In humans, the consequences of ZIKV infection range from asymptomatic infection to severe neurological disease such as Guillain-Barré syndrome or fetal neurodevelopmental defects, suggesting, among other factors, the influence of host genetic variants. We previously reported similar diverse outcomes of ZIKV infection in mice of the Collaborative Cross (CC), a collection of inbred strains with large genetic diversity. CC071/TauUnc (CC071) was the most susceptible CC strain with severe symptoms and lethality. Notably, CC071 has been recently reported to be also susceptible to other flaviviruses including dengue virus, Powassan virus, West Nile virus, and to Rift Valley fever virus. To identify the genetic origin of this broad susceptibility, we investigated ZIKV replication in mouse embryonic fibroblasts (MEFs) from CC071 and two resistant strains. CC071 showed uncontrolled ZIKV replication associated with delayed induction of type-I interferons (IFN-I). Genetic analysis identified a mutation in the *Irf3* gene specific to the CC071 strain which prevents the protein phosphorylation required to activate interferon beta transcription. We demonstrated that this mutation induces the same defective IFN-I response and uncontrolled viral replication in MEFs as an *Irf3* knock-out allele. By contrast, we also showed that *Irf3* deficiency did not induce the high plasma viral load and clinical severity observed in CC071 mice and that susceptibility alleles at other genes, not associated with the IFN-I response, are required. Our results provide new insight into the *in vitro* and *in vivo* roles of *Irf3*, and into the genetic complexity of host responses to flaviviruses.

d'Avenir program, Laboratoire d'Excellence: IBEID (Integrative Biology of Emerging Infectious Diseases, ANR-10-LABX-62-IBEID to XM, CM and MB). CM and MB were supported by doctoral fellowships from grant ANR-10-LABX-62-IBEID. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Author summary

Recent ZIKV outbreaks led to millions of infected people, with rare but severe complications such as Guillain-Barré syndrome and encephalitis in adults suggesting that host genes influence the susceptibility to severe forms of infection. We previously reported the importance of host genes in ZIKV pathogenesis using a panel of genetically diverse mouse strains and identified CC071 as the most susceptible strain. Importantly, this mouse strain has been shown by others to be also susceptible to several other RNA viruses. Through a combination of functional and genetic approaches in a cellular model, we identified a mutation in the *Irf3* gene which plays a key role in activating the expression of interferon beta to induce the type I interferon response, the first line of host defense against the virus. This mutation fully explains the high viral replication observed in CC071 cells. However, it was not able to induce the elevated viremia and the disease signs displayed by CC071 ZIKV-infected mice, unraveling the implication of other host genes which are not associated with the type I interferon response. Because of the broad susceptibility of CC071 to multiple viruses, our results have implications beyond ZIKV infection and contribute to shedding light on the plurality of host mechanisms fighting infectious diseases.

## Introduction

Zika virus (ZIKV) is a mosquito-borne virus of the *Flaviviridae* family identified in 1947 in Uganda. The first noticeable human outbreaks occurred in Micronesia in 2007 and in French Polynesia and New Caledonia in 2013–2014. In 2015–2016, ZIKV caused an epidemic in Brazil which rapidly spread across the Americas and the Caribbean. To date, 89 countries have reported evidence of mosquito-transmitted Zika virus infection ([https://www.who.int/health-topics/zika-virus-disease#tab=tab\\_1](https://www.who.int/health-topics/zika-virus-disease#tab=tab_1)).

While most people infected with ZIKV remain asymptomatic, some develop non-specific symptoms including rash, fever, conjunctivitis, muscle and joint pain, malaise and headache. Neurological complications have been described in adults such as Guillain-Barré syndrome [1] and encephalitis [2]. Infection of pregnant women was associated with congenital Zika syndrome in the fetus, which can lead to neurodevelopmental deficiencies, brain malformation [3] or in some cases to fetal loss [4].

Many factors may contribute to this variable severity, including the viral strain, the infection route and dose, and the host genetic background [5,6]. Indeed, mouse and human studies have shown that host genes influence flaviviral infections' outcomes [7]. While human genetic studies are hampered by the variability of these multiple factors, they can be controlled in mouse models which have proven very valuable to identify susceptibility variants [8,9]. Relevant ZIKV infection models have been developed in mice either using *Ifnar1* knock-out (KO) mice in which the IFNAR receptor to IFN-I has been inactivated [10,11], or by blocking this receptor using a monoclonal antibody targeting the IFNAR1 receptor subunit (MAR1-5A3 [12]).

We have previously explored the role of mouse natural genetic variants on ZIKV susceptibility in the Collaborative Cross (CC), a panel of recombinant inbred mice encompassing a genetic diversity similar to that of the human population and capturing approximately 90% of the mouse natural genetic variants [13,14]. We reported that the CC genetic diversity enabled large variations in the clinical severity of ZIKV disease, plasma viral load and intensity of brain pathology, comparable to those observed in human cases [15]. We specifically identified CC071/TauUnc (CC071) mice as very susceptible, with high mortality and high peak plasma

viral load. Notably, the CC071 strain has been recently reported as also susceptible to other flaviviruses including dengue virus [15], Powassan virus [16], West Nile virus (WNV) [15], and to Rift Valley fever virus [17], emphasizing its value to decipher genetic factors controlling host responses to RNA viruses.

We previously demonstrated that genetic background influenced ZIKV replication in CC mouse embryonic fibroblasts (MEFs). Here, we investigated the mechanisms driving high viral replication in CC071 MEFs. We found that, compared with the more resistant C57BL/6J (B6) and CC001/Unc (CC001) strains, ZIKV-infected CC071 MEFs displayed a delayed expression of the IFN-I response genes. Genetic and functional analyses identified a strain-specific variant in the *Irf3* gene, the first transcription factor involved in interferon (IFN)  $\beta$  expression, which mimics the effects of a null allele *in vitro* and fully explains the delayed IFN-I expression and uncontrolled viral replication. By contrast, we showed, from the *in vivo* comparison of CC071, *Irf3*-deficient and backcross mice, that the CC071 *Irf3* mutation is not sufficient to explain the high susceptibility of CC071 mice to ZIKV infection and that other genes, not associated with the IFN-I response, are involved. These findings provide new insights into the roles of *Irf3* in viral diseases and exemplify how the study of CC strains allows deciphering the role of host genes in viral pathogenesis.

## Results

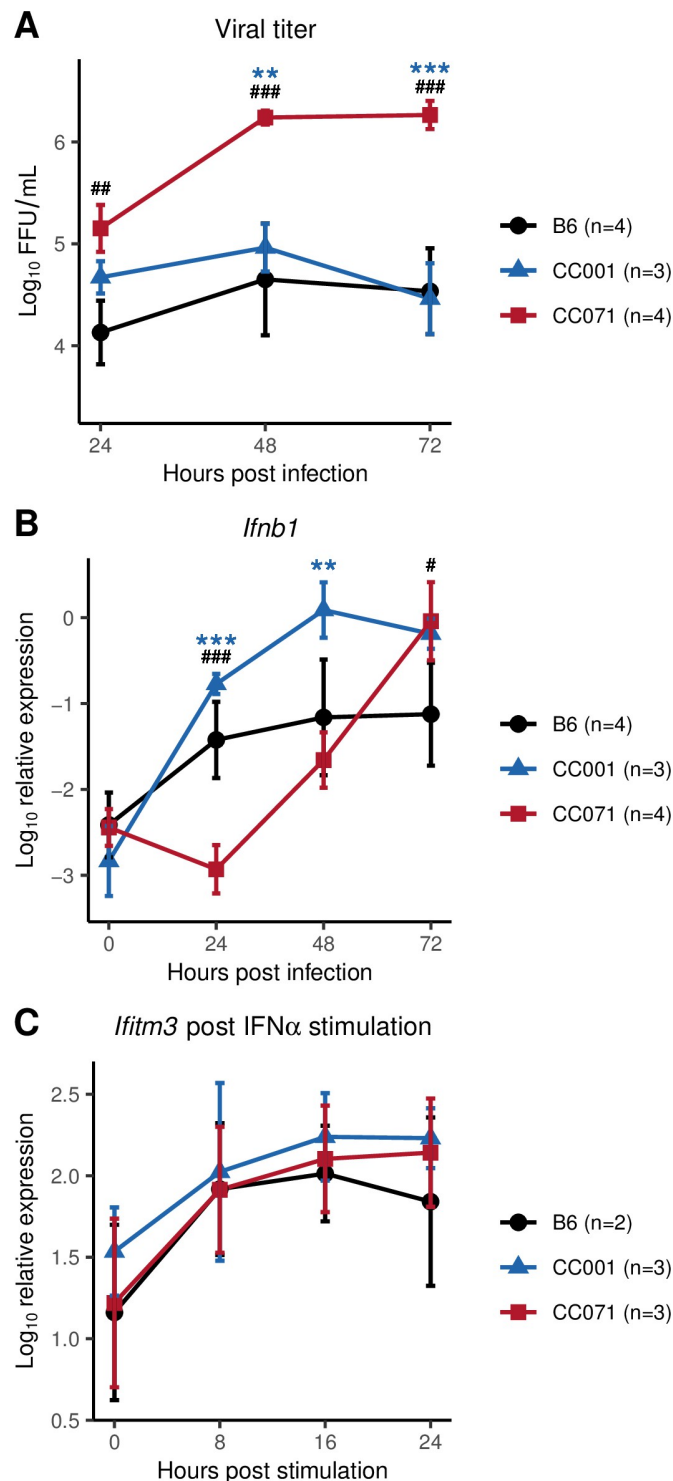
### CC071 MEFs show defective control of viral replication and delayed IFN-I expression, but normal response to IFN-I stimulation

We previously reported that, unlike CC001 MEFs, CC071 MEFs produced increasing quantities of viral particles during the first 72 hours post-infection (hpi) [15]. Here, we confirmed and expanded this observation by infecting B6, CC001 and CC071 MEFs and by quantifying viral particles by FFA. After ZIKV infection, CC071 MEFs displayed high and increasing viral titers between 24 and 72 hpi, while CC001 and B6 MEFs showed stable and lower titers (Fig 1A).

To investigate the origin of the defective control of viral replication in CC071 MEFs, we measured the expression level of the *Ifnb1* gene coding for IFN $\beta$  in ZIKV-infected CC071, CC001 and B6 MEFs. *Ifnb1* expression is induced very rapidly after virus detection by sensors and triggers the innate antiviral response which is essential for limiting viral replication. In CC001 and B6 MEFs, *Ifnb1* expression was significantly induced at 24 hpi and remained stable and high until at least 72 hpi (Fig 1B). In contrast, its expression in CC071 MEFs was low at 24 hpi and reached the level of CC001 only at 72 hpi. Similar results were obtained for *Ifna4* which encodes one of the IFN $\alpha$  proteins (S1 Fig). Notably, *Ifnb1* expression in CC071 MEFs at 72 hpi was significantly higher than in B6 MEFs, showing that CC071 MEFs were delayed but not intrinsically hampered in their ability to induce strong IFN-I expression.

To test whether this defective induction of *Ifnb1* expression was specific to ZIKV infection, MEFs were then transfected with the influenza A virus-derived 3-phosphate-hairpin-RNA (3p-hpRNA), an agonist of the RIG-I/MDA5-MAVS pathway, or treated with polyinosine-polycytidylic acid (poly (I:C)), that activates both Toll-like receptor (TLR) 3 (TLR3) and the RIG-I/MDA5-MAVS pathway [18]. Here again, CC071 MEFs showed a delayed expression of *Ifnb1* after both stimulations by comparison with B6 and CC001 MEFs (S2 Fig), indicating that the defect in IFN-I genes expression in CC071 MEFs was not specific to ZIKV infection. This result suggested a defect in the molecular cascade between cellular sensors of pathogen-associated molecular patterns (PAMP) and the *Ifnb1* gene transcription machinery.

To evaluate the capacity of CC071 MEFs to respond to IFN-I, they were treated with recombinant IFN $\alpha$ . The expression of IFN-stimulated genes (ISGs) such as *Ifitm3* was induced with



**Fig 1. CC071 MEFs fail to control viral replication, with delayed *Ifnb1* expression but normal response to IFN-I.** (A, B) MEFs derived from B6 (black circles), CC001 (blue triangles) and CC071 (red squares) were infected with ZIKV at a MOI of 5 and analyzed 24, 48 and 72 hpi. (A) Viral titer in supernatants was quantified by FFA. (B) *Ifnb1* expression was normalized to *Tbp* reference gene. Data are mean  $\pm$  sem from 3 to 4 biological replicates per strain (MEFs derived from individual embryos). (C) MEFs were stimulated with recombinant IFN $\alpha$ . *Ifitm3* relative expression normalized to *Tbp* reference gene is shown as an example of ISG. Data are mean  $\pm$  sem from 2 to 3



biological replicates. Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by Tukey HSD, \*/#  $p < 0.05$ , \*\*/##  $p < 0.01$ , \*\*\*/###  $p < 0.001$ ).

<https://doi.org/10.1371/journal.ppat.1011446.g001>

the same kinetics and level as in B6 and CC001 MEFs (Fig 1C), showing that CC071 MEFs are able to respond normally to IFN-I stimulation and that their defect is limited to the induction of *Ifnb1* gene expression.

### CC071's delayed *Ifnb1* expression is strain-specific

To gain insight into the mechanisms responsible for defective *Ifnb1* induction, we investigated the expression levels of genes involved in *Ifnb1* expression on CC071, B6 and CC001 MEFs at 16, 24 and 32 hpi. Mock-infected MEFs were analyzed at 24 hours as controls. Expression levels were measured by RNA sequencing (RNAseq) which provided a comprehensive analysis of transcriptomic changes. In CC001 MEFs, the expression of many genes rapidly increased after infection (160 at 16h, 821 at 24h and 971 at 32h; log<sub>2</sub> fold-change > 1, FDR = 0.05), reflecting a robust innate antiviral response (S3A Fig). A similar pattern was observed in B6 MEFs. By contrast, the expression of only 38 genes was increased in CC071 MEFs at 32hpi (34 of which were also activated in CC001), consistent with the delayed induction of *Ifnb1* expression. Among the genes that are involved in the pathway between PAMP sensors and *Ifnb1* transcription, ISGs such as *Tlr3*, *Ddx58* (coding for RIG-I sensor) or *Irf7* were not induced upon infection in CC071, while constitutively expressed genes such as *Mavs*, *Ticam1* (coding for the TRIF adaptor), *Traf3* or *Irf3*, showed comparable levels of expression in the three strains (S3B Fig). Therefore, this analysis did not provide new clues for identifying the gene responsible for the defect observed in CC071.

We then leveraged the genetic architecture of the CC which genomes are patchworks of haplotypes inherited from the eight founder strains [19]. Although CC071 was the only strain with severe ZIKV disease, we hypothesized that, if the delayed activation of *Ifnb1* resulting in uncontrolled viral replication observed in CC071 MEFs was due to an allele at one of the genes involved in the *Ifnb1* induction pathway inherited from a parental strain, ZIKV-infected MEFs of CC strains carrying the same allele would present similarly high viral titers. We therefore derived MEFs from each CC strain available to us carrying the same ancestral haplotype as CC071 at one of the 13 genes of the pathway (Fig 2A). Upon ZIKV infection, none of these CC MEFs showed viral titer kinetics resembling that observed in CC071 MEFs (Fig 2B). These results suggested two alternative hypotheses. Either the delayed *Ifnb1* activation involved two members of the pathway with a CC071-specific allelic combination leading to a non-functional interaction, or CC071 was carrying a strain-specific allele at one of these genes, resulting from a CC071-specific mutation that probably arose on an ancestral haplotype during the CC071 inbreeding. However, the sequencing of one male of each CC strain reported in 2017 [20] did not identify such "private" variants with high predicted impact in CC071 for any of these genes. Whatever the molecular mechanism, our results indicated that it was specific to CC071.

### Genetic analysis identifies *Irf3* as a candidate gene in a haplotype shared between CC071 and CC001

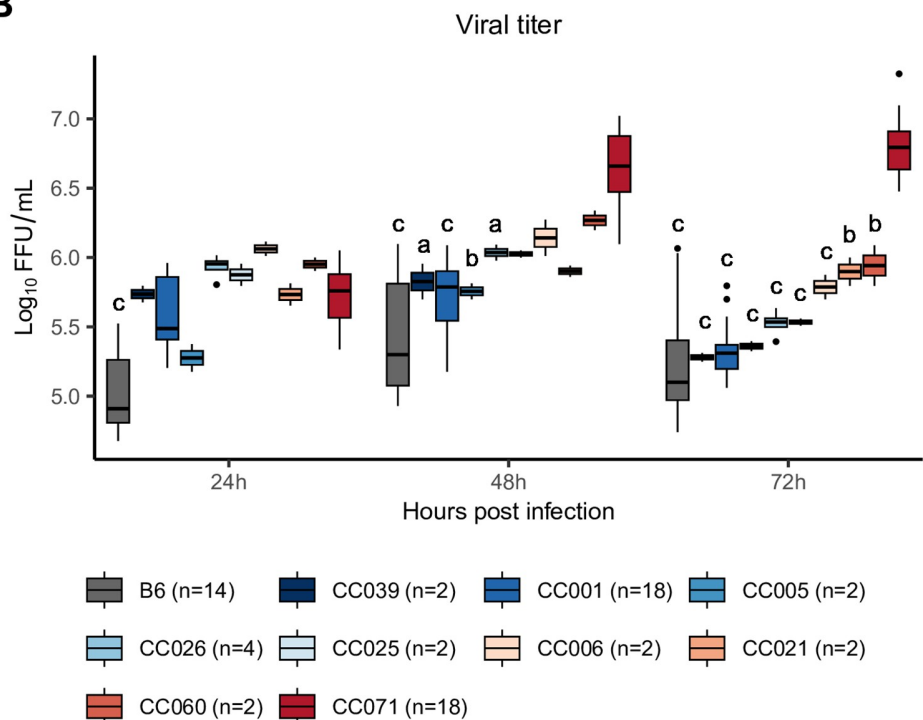
We then turned to a genetic mapping approach. We first established that (CC001xCC071)F1 MEFs responded to infection with as rapid induction of *Ifnb1* expression as CC001 MEFs (S4 Fig), suggesting that this CC071 trait was recessively inherited. F1 mice were therefore backcrossed with CC071. MEFs were produced from each of 51 backcross (BC) embryos, infected with ZIKV, and analyzed for viral titer and *Ifnb1* expression as above. One CC001 and one

**A**

CC strains haplotypes

Gene	Protein	CC071	CC001	CC005	CC026	CC021	CC006	CC025	CC039	CC060
<i>Ddx58</i>	RIG-I	HH	BB	EE	CC	HH	AA	HH	AD	EE
<i>Irf1</i>	LGP2	GG	AA	DD	HH	GG	BB	AA	BB	CC
<i>Dhx58</i>	MDA-5	FF	EE	DD	HH	FF	DD	FF	FF	EE
<i>Tlr3</i>	TLR3	AA	AA	BB	CC	EE	AA	BB	FF	CC
<i>Mavs</i>	MAVS	CC	AA	DD	HH	GG	CC	DD	HH	DD
<i>Ticam1</i>	TRIF	CC	AA	EE	BB	CC	DD	HH	AA	EE
<i>Azi2</i>	NAP1	EE	DD	FF	BB	FF	EE	GG	CC	EE
<i>Traf3</i>	TRAF3	HH	HH	GG	CC	FF	CC	CC	AA	BB
<i>Tbk1</i>	TBK1	HH	DD	GG	HH	AA	FF	AA	GG	BB
<i>Ikbke</i>	IKK $\epsilon$	CC	EE	CC	AA	EE	CC	CC	GH	AA
<i>Irf3</i>	IRF3	FF	FF	HH	BB	CC	DD	EE	HH	BB
<i>Irf7</i>	IRF7	HH	BB	DD	DD	HH	EE	DD	HH	DD
<i>Irfb1</i>	IFN $\beta$	HH	EE	DD	CC	HH	BB	HH	CC	DE

**B**

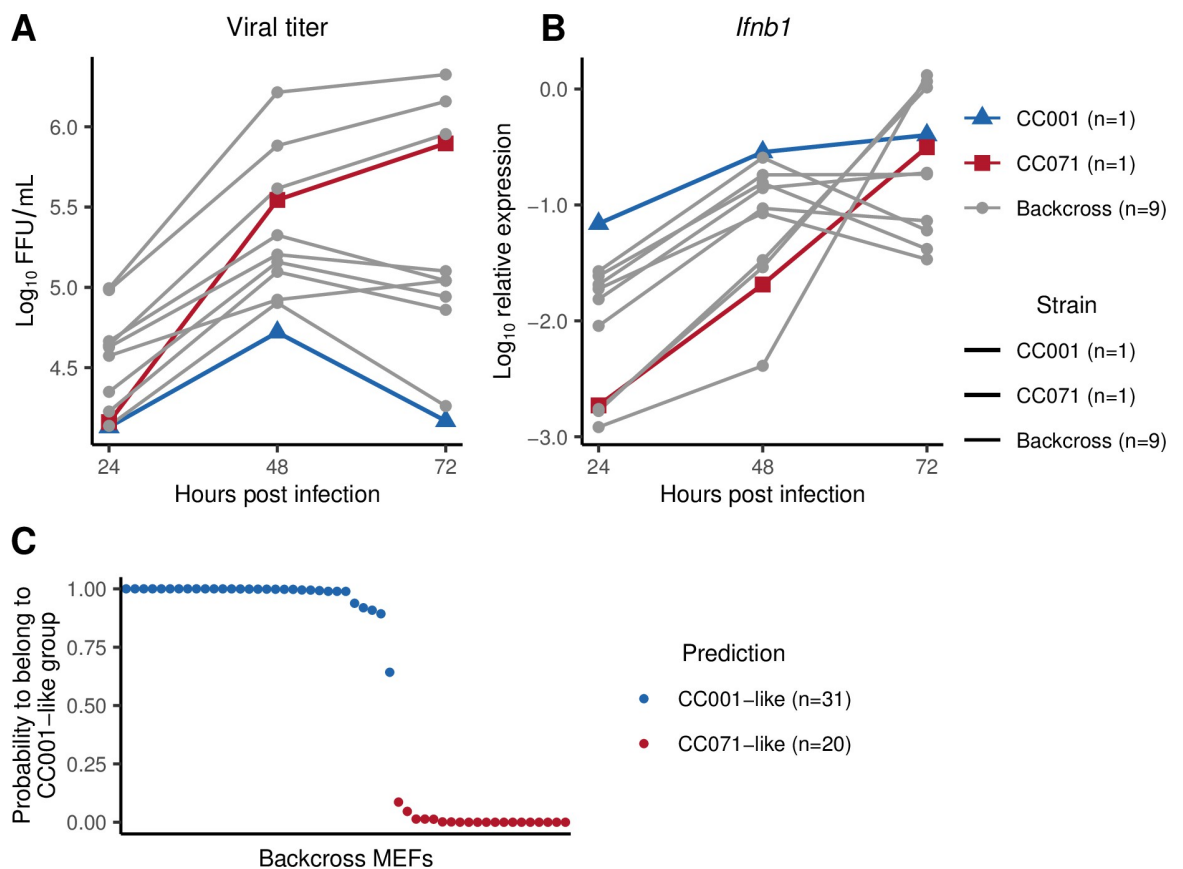


**Fig 2. Haplotype analysis fails to identify a gene from the *Irfb1* induction pathway associated with uncontrolled viral replication.** (A) Identification of CC strains which carry the same ancestral haplotype as CC071 at the genes involved in the pathway leading to *Irfb1* expression. Colored boxes indicate matched haplotypes between CC071 and other CC strains. Letters and colors designate the eight CC founder strains. A: A/J (yellow); B: C57BL/6J; C: 129S1/SvImJ (pink); D: NOD/ShiLtJ; E: NZO/HILtJ (light blue); F: CAST/Eij (green); G: PWK/PhJ (red); H: WSB/Eij (purple). Doubled letters (eg AA) indicate homozygous genotypes. Heterozygous genotypes are indicated by the two corresponding letters (eg AD). (B) Kinetics of viral titer in MEFs from CC071 and the 8 CC strains shown in (A). Experimental conditions were as in Fig 1A n: number of technical replicates for each strain. Data are mean +/- sem from the technical replicates. Letters show statistical significance between CC071 and other strains. (ANOVA followed by Tukey HSD, a: p < 0.05, b: p < 0.01, c: p < 0.001).

<https://doi.org/10.1371/journal.ppat.1011446.g002>

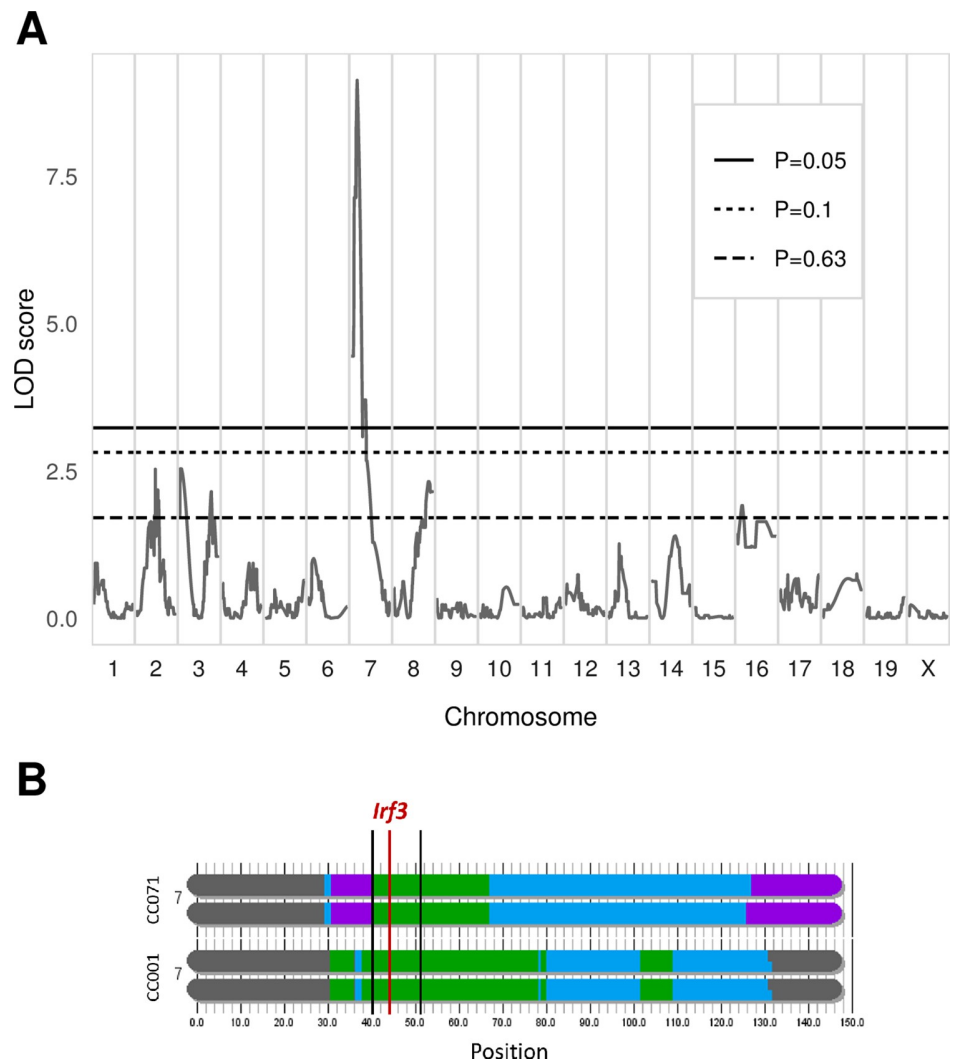
CC071 MEF lines were included in each infection experiment as controls. BC MEFs displayed either rapid and high *Ifnb1* expression with low viral titer (like CC001), or high viral titer and delayed *Ifnb1* expression (like CC071, Fig 3A and 3B), showing that these two traits correlated across the BC diverse genetic backgrounds.

To confirm this apparently binary distribution, we conducted linear discriminant analysis (LDA) on CC001 and CC071 MEFs using *Ifnb1* expression at the three time points as variables. Applying the LDA coefficients to backcross MEFs data classified individuals either in a CC001-like group (n = 31; 61%) or in a CC071-like group (n = 20; 39%), with a mean probability of prediction of 0.975 and 0.991, respectively (Fig 3C). Quantitative trait locus (QTL) mapping was performed using LDA classification as a binary trait. Genome scan identified a peak on chromosome 7 with a LOD (logarithm of the odd) score of 9.138 ( $p < 0.001$ , Fig 4A) located in a region centered on the *Irf3* gene which, given its main role in the regulation of *Ifnb1* expression, appeared as an obvious candidate. However, both CC001 and CC071 inherited the CAST/EiJ haplotype in this region (Fig 4B), strongly suggesting that CC071's susceptibility was caused by a variant proper to this strain.



**Fig 3. Backcross MEFs display either a CC001-like or a CC071-like phenotype.** MEFs derived from CC001 (blue triangles), CC071 (red squares) and backcross (gray circles) embryos were infected with ZIKV at a MOI of 5. (A) Viral titer and (B) *Ifnb1* expression in ZIKV-infected MEFs from 9 backcross embryos. Experimental conditions were as in Fig 1A and 1B. Red and blue curves show the results for CC071 and CC001 MEFs, respectively, from the same infection experiment. (C) Results of LDA on the backcross MEFs. LDA coefficients were calculated from *Ifnb1* expression data in CC001 and CC071 infected MEFs, and applied to backcross MEFs. The graph shows the probability of each BC MEF to belong to the "CC001-like" group, resulting in two distinct populations shown in blue ("CC001-like") and in red ("CC071-like"). n: number of BC MEFs assigned to each group.

<https://doi.org/10.1371/journal.ppat.1011446.g003>

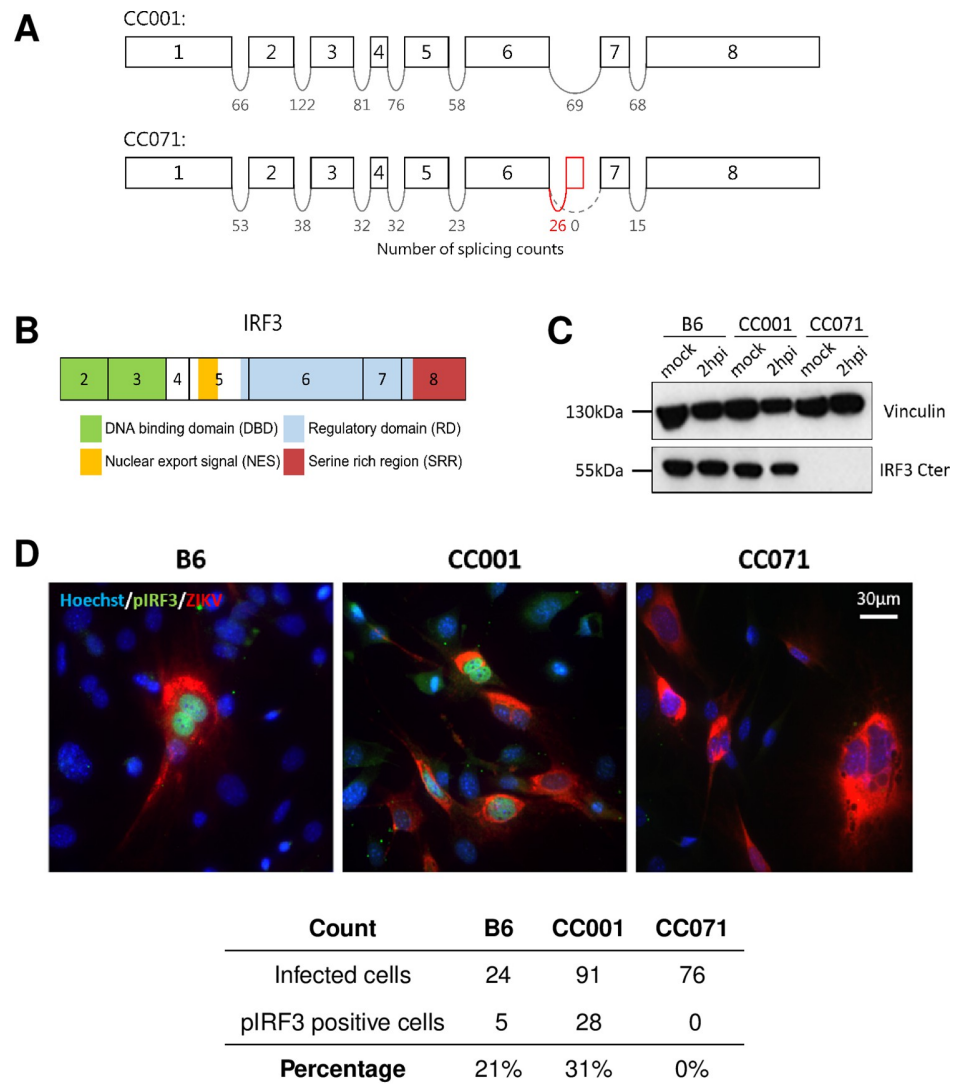


**Fig 4. Genetic analysis of *Ifnb1* expression in backcross MEFs identifies a major determinant mapping to the *Irf3* locus.** (A) Genome-wide linkage analysis of the LDA classification of backcross MEFs performed with R/qtl. X-axis: genomic location. Y-axis: LOD score of the phenotype-genotype association. Genome-wide significance thresholds ( $P = 0.05$ ,  $P = 0.1$  and  $P = 0.63$ ) were computed from 1000 data permutation. The chromosome 7 peak has a LOD score of 9.138. (B) Schematic representation of CC001 and CC071 chromosome 7 haplotypes, from <https://csbio.unc.edu/CCstatus/CCGenomes/>. Colors represent the CC ancestral haplotypes (same colors as in Fig 2A). Thick vertical black lines show the peak's 95% Bayesian confidence interval (25.9–31.3 cM, corresponding to 40.1–50.6 Mb). The red line shows the position of *Irf3* gene.

<https://doi.org/10.1371/journal.ppat.1011446.g004>

### Abnormal splicing of *Irf3* mRNA in CC071 leads to a loss of IRF3 transcriptional function

To identify the CC071-specific mutation, we re-analyzed the RNAseq data and investigated the splicing events between *Irf3* exons. As shown in Fig 5A, no splicing was observed between exons 6 and 7 in CC071 MEFs, while a short cryptic exon was added to exon 6 (red box). This aberrant splicing resulted in an mRNA lacking the last two exons. Notably, exon 8 encodes the serine-rich region of the protein with the phosphorylation sites necessary for IRF3 activation and nuclear translocation leading to *Ifnb1* transcription (Fig 5B). Neither long-range PCRs nor sequencing could identify the exact nature of *Irf3* genetic alteration in CC071 but



**Fig 5. CC071 *Irf3* mRNAs show abnormal splicing, resulting in defective IRF3 protein.** (A) Schematic representation of the exons of the *Irf3* gene with the number of reads spanning successive exons in the CC001 and CC071 RNAseq data (one sample of each strain). The red box between CC071's exons 6 and 7 depicts a novel exon resulting from abnormal splicing. (B) Schematic representation of the IRF3 protein structural domains (exon 1 is untranslated). Exon 8 encodes the serine rich region containing the phosphorylation sites for IRF3 activation. (C) Western blot using an anti-C-terminal IRF3 antibody from mock-infected and ZIKV-infected B6, CC001 and CC071 MEFs at 2 hpi, showing the absence of full-length IRF3 in CC071 MEFs. Vinculin was used as a loading control. (D) Immunofluorescence using an anti-phosphorylated IRF3 (pIRF3, green) in ZIKV-infected B6, CC001 and CC071 MEFs at 24 hpi, showing the absence of pIRF3 in the nucleus of CC071 MEFs upon infection. Red-labeled 4G2 antibody labels ZIKV-infected cells. Cell nuclear DNA labeled by Hoechst (blue). Quantification of the number of infected and pIRF3 positive cells is presented in the table. Proportions were established on 420, 428 and 551 cells for CC001, CC071 and B6, respectively.

<https://doi.org/10.1371/journal.ppat.1011446.g005>

suggested the insertion of a repeated sequence between exons 6 and 7. Nevertheless, the functional consequence of this mutation was confirmed by Western blot using a specific C-terminal IRF3 antibody which showed that full-length IRF3 protein was absent in CC071 MEFs (Fig 5C). Moreover, immunofluorescence using an antibody directed against phosphorylated IRF3 detected a positive signal in the nucleus of many ZIKV-infected CC001 and B6 MEFs, but not

in CC071 MEFs (Fig 5D). Altogether, these results show that CC071 carries a mutation in *Irf3* that prevents IRF3 phosphorylation which is required to induce *Ifnb1* expression. Whether the altered mRNA sequence prevents the production of the protein or alters its activity, this mutation results in the loss of IRF3 transcriptional function.

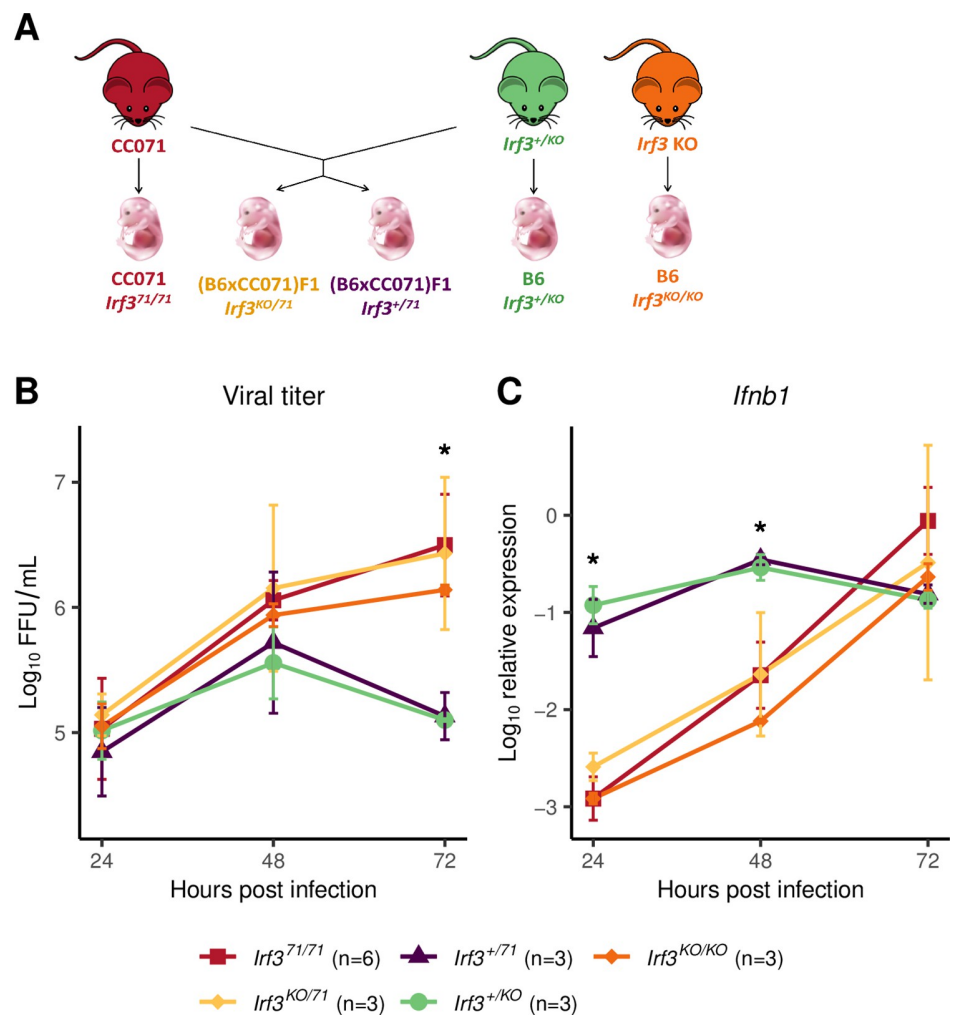
### CC071 *Irf3* mutation is responsible for uncontrolled viral replication in MEFs

To test if the delayed *Ifnb1* expression resulting in uncontrolled viral replication in CC071 MEFs was caused exclusively by the *Irf3* mutation, we performed a quantitative complementation test by producing compound heterozygous MEFs carrying a knockout *Irf3* allele (*Irf3*<sup>KO</sup>) and a CC071 allele (*Irf3*<sup>71</sup>). These *Irf3*<sup>KO/71</sup> MEFs were compared with CC071 and B6-*Irf3*<sup>KO/KO</sup> MEFs, and with heterozygous MEFs carrying a B6 wildtype allele and either an *Irf3*<sup>KO</sup> or an *Irf3*<sup>71</sup> allele (*Irf3*<sup>+/KO</sup> or *Irf3*<sup>+/71</sup>, respectively, Fig 6A). While *Irf3*<sup>+/KO</sup> and *Irf3*<sup>+/71</sup> ZIKV-infected MEFs showed the same pattern as CC001 or B6 MEFs (rapid induction of *Ifnb1* expression and controlled viral replication, see Fig 1A for comparison), CC071 (*Irf3*<sup>71/71</sup>), B6-*Irf3*<sup>KO/KO</sup> and *Irf3*<sup>KO/71</sup> MEFs carrying two defective alleles at *Irf3* showed similar results (Fig 6B and 6C). These data demonstrate that, since the *Irf3*<sup>KO</sup> did not complement the *Irf3*<sup>71</sup> allele, the *Irf3* mutation in CC071 contributes to the defects observed in ZIKV-infected MEFs. Moreover, since the data obtained on CC071 and on *Irf3*<sup>KO/71</sup> MEFs were identical, we conclude that the CC071 *Irf3* mutation is sufficient to induce the defects observed in CC071 MEFs.

### CC071 *Irf3* mutation is not sufficient to explain susceptibility *in vivo*

We investigated whether this *Irf3* mutation was also responsible for the high susceptibility to ZIKV of CC071 mice. We first evaluated its effects in a context allowing IFN-I response. To this aim, we compared B6, B6-*Irf3* KO and CC071 mice infected without prior MAR1-5A3 treatment, which results in moderate and short-lasting plasma viral loads without clinical signs, as we previously reported [15]. Similar viral loads were observed in B6 and B6-*Irf3* KO mice at all days p.i. (dpi), while CC071 mice showed significantly higher viral loads at days 1 and 2 p.i. (Fig 7A). This result indicates that, unlike in MEFs, the *Irf3* null mutation does not induce elevated viral replication *in vivo* (as measured by plasma viral load) and suggests that CC071 mice carry susceptibility alleles at other loci.

We then blocked IFN-I response by pre-treating mice with MAR1-5A3 and compared peak plasma viral load and clinical signs in ZIKV-infected B6, B6-*Irf3* KO and CC071 mice. While no differences were observed between B6 and B6-*Irf3* KO mice (moderate plasma viral loads without clinical signs), as expected if the IFN-I response is neutralized, CC071 mice developed significantly higher viral loads at day 2 p.i. and signs of disease (ruffled fur, hunched posture and body weight loss) around 7 days p.i. (Fig 7B). This contrast could result either from functional differences between the *Irf3*<sup>KO</sup> and *Irf3*<sup>71</sup> alleles or from CC071 alleles at other genes not linked with the IFN-I response. F1 mice between CC071 and B6-*Irf3* KO mice gave similar results as B6-*Irf3* KO mice, suggesting that CC071 susceptibility alleles were recessive. Therefore, we generated a small cohort of (CC071 x B6-*Irf3* KO) x CC071 backcross mice and genotyped them for the *Irf3* gene by PCR using primers that amplify intron 6 of the *Irf3*<sup>KO</sup> allele, but not of the *Irf3*<sup>71</sup> allele. (Figs 5A and S5). We found no differences in plasma viral load and clinical scores between *Irf3*<sup>KO/71</sup> and *Irf3*<sup>71/71</sup> backcross mice, indicating that the two alleles are functionally equivalent *in vivo*. Moreover, *Irf3*<sup>KO/71</sup> backcross mice were more severely affected than *Irf3*<sup>KO/71</sup> F1 mice, indicating the strong effects of recessive alleles from CC071 which likely explain the variability of plasma viral load and clinical signs among backcross mice. Altogether, our results demonstrate that the CC071 mutation is a loss-of-function (LOF) variant

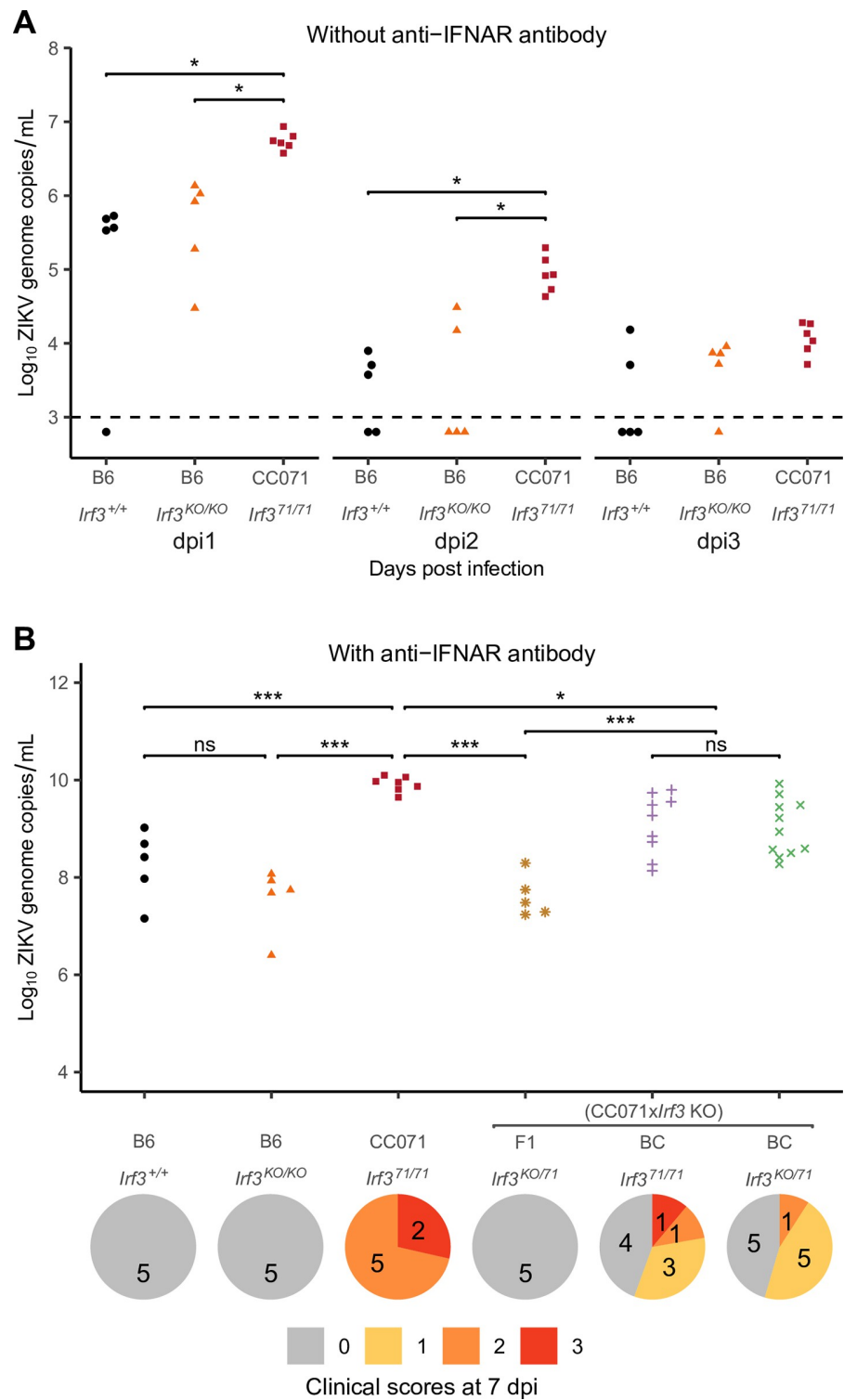


**Fig 6. Quantitative complementation test confirms that *Irf3* LOF in CC071 is responsible for uncontrolled viral replication and delayed *Ifnb1* expression.** (A) Mice heterozygous for an inactivated *Irf3* allele (B6- $Irf3^{+/KO}$ ) were mated with CC071 mice to produce four types of embryos from which MEFs were derived. The genetic background and *Irf3* genotype is shown below each type of embryos. (B) Viral titer and (C) *Ifnb1* expression in ZIKV-infected MEFs. Experimental conditions were as in Fig 1A and 1B. Data are mean  $\pm$  sem from 6 biological replicates for CC071 and 3 for the other groups. The asterisks represent the results of ANOVA test between all groups (\*  $p < 0.05$ ). Results of the Tukey HSD post-hoc are as follows. Viral titer at 72h: CC071 vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ :  $p < 0.001$ .  $Irf3^{KO/71}$  vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ :  $p < 0.01$ .  $Irf3^{KO/KO}$  vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ :  $p < 0.05$ . *Ifnb1* expression at 24hpi: CC071 vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ ,  $Irf3^{KO/71}$  vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ ,  $Irf3^{KO/KO}$  vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ :  $p < 0.001$ . *Ifnb1* expression at 48hpi: CC071 vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ ,  $Irf3^{KO/71}$  vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ ,  $Irf3^{KO/KO}$  vs  $Irf3^{+/71}$  and vs  $Irf3^{+/KO}$ :  $p < 0.01$ . Fig 6 includes a drawing of a mouse which was taken from <https://publicdomainvectors.org/fr/gratuitement-des-vecteurs/Dessin-de-souris-de-dessin-anim%C3%A9-avec-longue-moustache-vectoriel/22268.html> and a picture of a mouse embryo taken from [https://commons.wikimedia.org/wiki/File:201309\\_mouse\\_embryo.png](https://commons.wikimedia.org/wiki/File:201309_mouse_embryo.png). <https://doi.org/10.1371/journal.ppat.1011446.g006>

which does not confer susceptibility to ZIKV disease in the absence of other recessive alleles carried by CC071 at genes not associated with the IFN-I response.

## Discussion

An increasing number of publications have reported how the genetic diversity provided by the CC translates into an extended range of resistant and susceptible phenotypes in various infectious models. Since our first description of the high susceptibility of CC071 mice to ZIKV



**Fig 7. *Irf3* LOF is not sufficient to explain CC071 susceptibility to ZIKV infection *in vivo*.** (A) B6, B6-*Irf3* KO and CC071 mice were infected IP with 10<sup>7</sup> FFUs of ZIKV and monitored for 7 days, without prior IP injection of MAR1-5A3 IFNAR-blocking monoclonal antibody. Plasma viral load was quantified at days 1 to 3 p.i. by RT-qPCR. The two lines of the group legend (X-axis) indicate the genetic background of the mice and their genotype at the *Irf3* locus, respectively. Each dot represents one mouse. Groups were compared by Kruskal-Wallis followed by Wilcoxon test with



Holm correction for multiple testing (\*  $p < 0.05$ ). (B) B6, B6-*Irf3* KO, CC071, (CC071 x *Irf3* KO) F1 and (CC071 x *Irf3* KO) x CC071 BC mice were infected IP with  $10^7$  FFUs of ZIKV after IP injection of 2 mg of MAR1-5A3 IFNAR-blocking monoclonal antibody 24 h before infection, and monitored for 7 days. The graph shows plasma viral loads quantified at day 2 p.i. by RT-qPCR. The two lines of the group legend (X-axis) indicate the genetic background of the mice and their genotype at the *Irf3* locus, respectively. BC mice are separated into two groups depending on the genotype at the *Irf3* gene: homozygous for the CC071 mutant allele (*Irf3*<sup>71/71</sup>), or heterozygous for the CC071 and the KO alleles (*Irf3*<sup>KO/71</sup>). *Irf3* genotyping results are presented in S5 Fig. Each dot represents one mouse. Groups were compared by ANOVA followed by Tukey HSD (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ). Below the graph is shown the distribution of clinical scores at day 7 p.i. in the same groups of mice as above (0, no symptoms; 1, slight hunched posture; 2, ruffled fur, hunched posture and/or mild ataxia; 3: prostration, ataxia, partial paralysis).

<https://doi.org/10.1371/journal.ppat.1011446.g007>

infection, with disease severity and peak plasma viral load almost as high as those of B6-*Ifnar1* KO mice, and higher clinical signs and mortality than 129-*Ifnar1* KO mice [15], several groups have reported similar observations for other viruses. High viral load was found in CC071 mice after dengue infection [15]. Lethality after infection with West Nile virus [15] or Powassan virus [16] were also described, although there is likely contribution from the defective *Oas1b* allele that CC071 has inherited from 129S1/SvImJ. CC071 was also one of the most susceptible CC strains to Rift Valley Fever virus infection [17] and to hepatitis C virus with long-term viral persistence [21]. Because of its unique genetic background, this strain will be increasingly useful for infectious diseases studies, which underlines the importance of deciphering the mechanisms of its susceptibility to each pathogen.

In this study, we leveraged from our previous observation that CC071 MEFs failed to control ZIKV replication by comparison with resistant strains. Although they cannot recapitulate the complex interactions between multiple pathways, cell types and tissues of a whole organism, MEFs are a convenient cellular model which can be easily derived from any mouse genetic background. We identified a delayed activation of the *Ifnb1* gene in CC071 MEFs, resulting in delayed stimulation of ISGs. We then used a combination of genetic approaches to find the causative gene defect. The observation that MEFs were normally responsive to IFN-I stimulation was consistent with the higher susceptibility of mice treated by the MAR1-5A3 antibody compared with untreated mice which did not develop symptoms and showed much lower plasma viral load [15]. These results pointed at the pathway from PAMP sensors to *Ifnb1* transcription factors. Transcriptional analysis did not identify reduced expression of non-ISGs of this pathway. CC strains' genomic structure allows searching for haplotypes inherited from the same founder in CC strains showing similar phenotypes. However, we did not identify such haplotypes for genes of the *Ifnb1* induction pathway, suggesting that the defects observed in CC071 MEFs were strain specific. It is finally the analysis of MEFs derived from backcross embryos that established a monogenic inheritance, and genetic linkage unambiguously pointed at the causative *Irf3* gene. The MEF experimental model was particularly appropriate since we could derive cell lines from every backcross embryo. Our RNAseq data showed an RNA splicing defect in the *Irf3* gene in CC071, which functional consequences could be validated *in vitro*. The formal proof that the CC071 *Irf3* mutation was necessary and sufficient to cause *Ifnb1* delayed activation and uncontrolled viral replication came from a quantitative complementation test in which the *Irf3*<sup>KO</sup> allele was combined either with a CC071 or a B6 (functional) allele. Altogether, these results show that the *Irf3* defect in CC071 abrogates its transcriptional activity. Unfortunately, none of the antibodies recognizing the N-terminal protein domain that we tested worked in our hands in western blot. Therefore, we could not establish whether the *Irf3* mRNA detected by RNAseq is translated into an abnormal protein or not. However, the *in vivo* observation that *Irf3*<sup>KO/71</sup> and *Irf3*<sup>71/71</sup> backcross mice displayed the same phenotypes upon ZIKV infection shows that the two alleles are functionally equivalent in this context.

*Irf3* maps to a chromosomal region that CC071 inherited from CAST/EiJ, like the resistant CC001 strain (and CC046/Unc, not available to us). The CC071-specific mutation may have arisen during the inbreeding generations leading to this strain. Alternatively, the mutation might have been segregating in the CAST/EiJ founders and transmitted by chance only to the CC071 breeding funnel. Since all CAST/EiJ founders used in CC crosses originated from The Jackson Laboratory colony, genotyping past and present breeders of this colony could not only unravel the origin of the mutation but also allow eliminating this mutation to avoid unwanted effects.

Notably, this is not the first example of a CC-strain-specific mutation. We [22] and others [23] previously reported the extreme susceptibility of CC042/GeniUnc (CC042) to *Salmonella* Typhimurium and to *Mycobacterium tuberculosis*, respectively, as a consequence of a *de novo* 15-nucleotide deletion in the *Irgal1* gene. In the case of *Salmonella* Typhimurium, CC042 was standing out, with bacterial loads up to 1000 times higher than other CC and the susceptible B6 strains. In our previous study on ZIKV, CC071 was the most susceptible strain, but its peak viral load was just the highest in a continuous distribution of values. Additionally, this is not the first example of a *Irf3* spontaneous variant identified in mice modulating susceptibility to bacterial [24] or viral infections [25].

*Irf3* is an important transcription factor involved in the innate immune response. It is constitutively expressed and, at rest, inactive IRF3 is present in the cytoplasm. Upon viral entry (or other stimuli that activate TLRs such as TLR3 and 4 or RIG-I-like receptors), signal transduction leads to the phosphorylation of IRF3, resulting in its dimerization and translocation to the nucleus where it binds to the *Ifnb1* promoter [26]. This mechanism leads to the very fast production of IFN $\beta$  which is secreted by the cell and triggers the immediate response to viral infection through the activation of ISGs with diverse antiviral functions [27]. In addition to its transcriptional activity, IRF3 has been shown to counter viral infection by two additional mechanisms. IRF3 induces apoptosis through the RLR-induced IRF3-mediated pathway of apoptosis (RIPA) in infected cells [28] and inhibits NF- $\kappa$ B-mediated inflammation through the repression of IRF3-mediated NF- $\kappa$ B activity (RIKA) [29]. These newly described roles of IRF3 depend on other protein domains than those involved in its transcriptional activity. Whether these activities are also abrogated by the CC071 *Irf3* mutation remains to be established.

In mice, studies that used *Irf3*-deficient cellular models have consistently reported decreased IFN-I production and/or increased viral replication. For example, after infection with WNV, viral replication was increased in B6-*Irf3* KO bone marrow macrophages (BMMs) and moderately in primary neurons [30]. Higher viral replication was observed in herpes simplex virus 1 (HSV-1)-infected B6-*Irf3* KO bone marrow-derived dendritic cells (BMDCs) and BMMs, and IFN $\beta$  was reduced in BMDCs supernatants [31]. Similarly, *Ifnb1* expression was reduced in HSV-1-infected [32] and CHIKV-infected [33] B6-*Irf3* KO MEFs. In line with these findings, our study provides, to our knowledge, the first evidence for a role of *Irf3* in the infection of murine cells by ZIKV. Moreover, we recently reported that primary cultured neurons (PCNs) show a delayed activation of the *Ifnb1* expression upon ZIKV exposure compared with MEFs, and that this delay is even longer in PCNs derived from CC071 compared with CC001 [34]. The identification of the *Irf3* mutation in CC071 provides an explanation for this observation, although it would require confirmation in B6-*Irf3* KO PCNs.

In contrast, *in vivo* studies have reported inconsistent consequences of *Irf3* deficiency between viral infections. Here, we did not observe any differences in clinical signs nor in plasma viral loads between B6 and B6-*Irf3* KO mice (pre-treated or not with the MAR1-5A3 antibody), consistently with a previous study that reported neither mortality, nor body weight loss after ZIKV infection [10]. Likewise, B6-*Irf3* KO mice have been reported to show no

mortality and low virus load in the circulation following dengue virus [35] and CHIKV infection [33]. Contrastingly, WNV infection was lethal in all B6-*Irf3* KO mice with increased viral burdens in peripheral and central nervous system tissues, while 65% of infected WT mice survived [30]. Differences were also associated with the route of infection. For example, all WT and B6-*Irf3* KO mice survived after intravenous inoculation with HSV-1 [32], while intranasal inoculation led to 90% mortality in B6-*Irf3* KO but only 30% in WT mice [36]. In humans, LOF mutations in *Irf3* have been associated with increased susceptibility to diseases caused by WNV [37,38], HSV-1 [39,40] and more recently by SARS-CoV-2 [41]. Discrepancies between the consequences of *Irf3* deficiency in *in vitro* and *in vivo* experiments likely reflect differences between a single cell type model with an intrinsic capacity to mount a more or less IRF3-dependent IFN-I response, and a multicellular organism in which the paracrine effect of IFN-I produced by the most reactive cells induces protection of other, less reactive, cell types. In fact, we have recently reported such cooperation between neurons and microglia cells [34].

Our *in vivo* data, showing high plasma viral loads and clinical signs in all CC071 and half of backcross mice after MAR1-5A3 treatment, demonstrated that the high susceptibility of CC071 mice requires the contribution of other recessive alleles. Since their effect was observed in mice which IFNAR receptor had been blocked, we conclude that their mode of action is not dependent on an intact IFN-I response. Moreover, the comparison between the results of the complementation test on MEFs, with CC071 and *Irf3*<sup>KO/71</sup> yielding very similar data, and the *in vivo* observation that CC071 showed higher viremia than *Irf3*<sup>KO/71</sup> F1 mice, suggests that these other susceptibility alleles may not act at the cell level, at least during the first 72 hpi, but rather at a more integrated or systemic level. Whether or not these alleles require *Irf3* deficiency to induce the susceptibility observed in the CC071 mice has yet to be determined. While our limited understanding of the severe disease developing in CC071 ZIKV-infected mice does not point at potential mechanisms, identifying these alleles will allow disentangling the genetic factors controlling ZIKV and other viruses' pathogenesis.

## Material and methods

### Ethics statement

All mouse experiments were approved by the Institut Pasteur Ethics Committee (CETEA, project number dap190107) and authorized by the French Ministry of Research (project #19469), in compliance with French and European regulations.

### Mice and crosses

C57BL/6J (B6) mice were purchased from Charles River Laboratories France. Collaborative Cross strains (CC001/Unc, CC071/TauUnc, CC005/TauUnc, CC011/Unc, CC026/GeniUnc, CC061/GeniUnc, CC021/Unc, CC006/TauUnc, CC025/GeniUnc, CC039/Unc, CC060/Unc) were purchased from the Systems Genetics Core Facility, University of North Carolina and bred at the Institut Pasteur. *Irf3 Irf7* double KO mice (C57BL/6J-*Bcl2l12/Irf3*<sup>tm1Ttg</sup> *Irf7*<sup>tm1Ttg</sup>, [32,42]) were bred at the Institut Pasteur and backcrossed to B6 mice to generate *Irf3* single KO mice (B6-*Irf3* KO). Genetic mapping was performed on MEFs derived from (CC001 x CC071) x CC071 and CC071 x (CC001 x CC071) backcross embryos. For the quantitative complementation test, MEFs were derived from CC071, B6-*Irf3* KO and (B6-*Irf3*<sup>+ / KO</sup> x CC071) embryos. *In vivo* experiments were performed on B6, B6-*Irf3* KO, CC071, (CC071 x *Irf3* KO) F1 and CC071 x (CC071 x *Irf3* KO) backcross mice. All crosses are described as female x male. All mice were maintained as described previously [15].

## ZIKA virus

The FG15 Asian Zika virus (ZIKV) strain, isolated from a patient during ZIKV outbreak in French Guiana in December 2015, was obtained from the Virology Laboratory of the Institut Pasteur of French Guiana. Viral stock (passage 5) was prepared from supernatant of infected C6/36 cells, clarified by centrifugation at 800g and titrated on Vero cells by focus-forming assay.

## Mouse infection

All infection experiments were performed in a biosafety level 3 animal facility and mice were kept in isolators. Six- to 10-week-old, male or female mice were injected intraperitoneally with  $10^7$  PFU of ZIKV FG15. In some *in vivo* experiments, mice received an IP injection of 2 mg of MAR1-5A3 anti-IFNAR antibody (Euromedex, Cat#BX-BE0241) one day prior infection. Mouse numbers are indicated in figure legends. Both males and females were used since no differences between sexes were detected in our previous and present experiments. Clinical signs and body weight loss were recorded for up to seven days post infection. Blood samples were collected on EDTA from the retromandibular vein for plasma viral load assessment. Quantification of ZIKV viral copies by RT-qPCR was previously described [15].

## MEFs isolation

Pregnant females were euthanized at day 13.5–15.5 of gestation. For B6, CC001, CC071 and crosses used for genetic mapping and the complementation test, MEFs were isolated from individual fetuses to obtain biological replicates. For other CC strains, MEFs were derived from individual or pooled fetuses. Fetus bodies were chopped and digested with trypsin (Gibco Cat#25300054), then cultured at 37°C and 5% CO<sub>2</sub> in complete medium (DMEM Gibco Cat# 31966047, 10% fetal bovine serum PAA Laboratories Cat#A15-101, 1% penicillin/streptomycin Sigma Cat#P4333). MEFs were used until passage 2. For the backcross experiment, MEF lines were isolated from 51 backcross fetuses. Heads were used to prepare DNA for whole-genome genotyping.

## MEFs infection

MEFs were seeded at  $5 \times 10^4$  cells per well in 24-well plates the day before infection. They were exposed to ZIKV FG15 strain at a MOI of 5 for 2 hours after which the inoculum was replaced with fresh complete medium and MEFs were incubated for up to 72 hours. For kinetics studies, different wells were used for each time point. Backcross MEFs were infected in 6 infection experiments, each of which included one CC001 and one CC071 MEF lines.

## MEFs IFN $\alpha$ stimulation

MEFs were seeded at  $5 \times 10^4$  cells per well in 24-well plates one day before stimulation and treated with 300 IU/mL IFN $\alpha$  (Miltenyi Biotec Cat#130-093-131) and incubated for up to 24 hours.

## MEFs stimulation with poly(I:C) and 3p-hpRNA

For intracellular stimulation with Poly(I:C) or with 3p-hpRNA, MEFs were seeded at  $1.1 \times 10^5$  cells per well in 12-well plates the day before stimulation, transfected with 1  $\mu$ g/mL Poly(I:C) (InvivoGen Cat#vac-pic) or 0.5  $\mu$ g/mL 3p-hpRNA (InvivoGen Cat#tlrl-hprna) using 5  $\mu$ L Lipofectamine LTX and 1  $\mu$ L for poly(I:C) stimulation or 0.5  $\mu$ L for 3p-hpRNA stimulation of

Plus Reagent (ThermoFischer Scientific Cat#15338100), according to the manufacturer's instructions. After stimulation, MEFs were incubated for 8 to 24 hours.

### Focus forming assay

Quantification of ZIKV particles was performed by focus forming assay on Vero cells (ATCC CRL-1586) as previously described [15].

### RNA extraction from cells

MEFs were lysed in 350 $\mu$ L of RLT buffer (Qiagen) with 1%  $\beta$ -mercaptoethanol. RNA was extracted using RNeasy Mini Kit (Qiagen Cat#74104) according to the manufacturer's instructions, with addition of DNase I (Qiagen Cat# 79254) to prevent genomic DNA contamination.

### Reverse transcription and qPCR

Reverse-transcription was performed on 200ng of RNA using Superscript II polymerase (Invitrogen Cat#18064022) and RNaseOUT ribonuclease inhibitor (Invitrogen Cat#10777019). qPCR was performed on 20ng of cDNAs using Power SYBR Green PCR Master Mix (Applied Biosystems Cat#4367659) and 6pmol of each primer, on a QuantStudio 12K Flex or a ViiA 7 (ThermoFisher Scientific). Primers (Eurofins) sequences are provided in [S1 Table](#). Gene expression was expressed on a Log10 scale of relative expression to the reference Tbp gene.

### Genotyping

Genomic DNA was prepared from backcross fetuses' heads by proteinase K digestion, phenol-chloroform extraction and ethanol precipitation according to standard protocols. Whole-genome genotyping was performed at Neogen (Neogen/Geneseek, Inc, Lincoln, NE, USA) using the MiniMUGA array containing 11,125 SNP markers. For the quantitative complementation test, *Irf3* and *Irf7* genotyping was performed by Transnetyx (Cordova, TN) by real-time PCR on fetuses' heads. (CC071 x *Irf3* KO) x CC071 backcross individuals were genotyped by PCR using primers amplifying the intron 6 of *Irf3*<sup>+</sup> and *Irf3*<sup>KO</sup> but not *Irf3*<sup>71</sup> alleles. Primers (Eurofins) sequences are provided in [S1 Table](#).

### Immunofluorescence

MEFs were plated on glass coverslips before infection, fixed with 4% paraformaldehyde for 20 min and permeabilized with 100% methanol for 10 min at -20°C. Cells were incubated with blocking buffer (5% FBS 0.3% triton in PBS) for 1 hour, with primary antibodies diluted in antibody incubation buffer (AIB: 1% BSA 0.3% triton in PBS) overnight at 4°C and with secondary antibodies and Hoechst (dilution 1:1000) diluted in AIB for 1 hour. Coverslips were mounted on slides and imaged with a widefield microscope (Zeiss Axio Observer.Z1 with a Plan-Apochromat 20x/0.8 M27 objective and a Hamamatsu sCMOS ORCA-Flash 4.0 v3 camera). ZEN blue 2012 software (ZEISS) imaging software was used for image capture and Image J software (National Institutes of Health) to adjust brightness and contrast. Primary and secondary antibodies are indicated in [S2 Table](#).

### Western blot

MEFs were trypsinized for 5 min, washed and lysed in a protein extraction buffer (10mM TrisHCl pH7.5, 5mM EDTA, 150mM NaCl, 30mM Na<sub>2</sub>HPO<sub>4</sub>, 50mM NaF, 10% glycerol, 1% NP40, 1X cOmplete (Roche #11873580001), 1X PhosSTOP (Roche #4906845001), 1/1000 benzozonaz (Sigma Cat#E1014)) for 30 min at 4°C. Proteins diluted in Laemmli were resolved on

4–12% Bis-Tris gels (Invitrogen Cat#NP0323BOX) in MOPS buffer (Invitrogen Cat#NP0001) and transferred to nitrocellulose membranes (Bio-Rad Cat#1620112) in a 25mM Tris 200mM glycine 20% ethanol buffer. Blots were blocked in 5% milk in TBS-T (0.1% Tween20 in Tris Base Sodium), incubated with primary antibodies diluted in 3% milk in TBST overnight at 4°C, and incubated with secondary antibodies diluted in 3% milk in TBST for 90 min. Blots were revealed with ECL substrate (Thermo Scientific Cat#32132) and imaged with X-ray films. Primary and secondary antibodies are indicated in [S2 Table](#).

### Statistical analysis

Statistical analyses were performed using R version 4.1.0 [43]. Viral titers, gene expression and genome copies were log-transformed for graphs and statistical tests. One way ANOVA followed by Tukey HSD were used for testing multiple comparisons. For *in vivo* studies without MAR1-5A3 treatment, non-parametric Kruskal-Wallis followed by Wilcoxon tests with Holm correction for multiple testing were used to handle values below the limit of detection. Linear discriminant analysis (LDA) was conducted using the MASS package [44]. The LDA was trained on the phenotypes of the two parental CC001 and CC071 strains from each infection batch. LDA coefficients were applied to backcross mice for assignment to "CC071-like" or "CC001-like" groups.

### Genetic analysis

Raw genotypes were curated using the stuart package [45]. QTL mapping was performed using R/qtl [46]. LDA prediction was used as a binary trait. Statistical significance thresholds were computed by data permutation ( $n = 1000$ ). 95% confidence interval was estimated using the Bayesian method.

### RNA sequencing

MEF RNA was prepared as described in the main text. RNA integrity and quantification were assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). One microgram of high-quality RNA samples ( $RIN > 9.2$ ) representing biological triplicates were submitted to Novogene for RNA-sequencing (Novogene Beijing, China). Poly-A selected RNA was used for paired-end library preparation and transcriptome sequencing. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's instructions. The library preparations were sequenced on an Illumina platform and paired-end reads were generated. The RNA-seq analysis was performed with Sequana 0.9.8 [47]. In particular, we used the RNA-seq pipeline [https://github.com/sequana/sequana\\_rnaseq](https://github.com/sequana/sequana_rnaseq) built on top of Snakemake 6.1.1 [48]. Briefly, reads were trimmed from adapters using Cutadapt 2.7 then mapped to the Mus musculus genome assembly GCA\_000001635.8 from NCBI using STAR 2.7.3a [49]. FeatureCounts 1.6.4 [50] was used to produce the count matrix, assigning reads to features using corresponding annotation v92 from NCBI with strand-specificity information. Quality control statistics were summarized using MultiQC 1.6 [51]. Clustering of transcriptomic profiles were controlled using a Principal Component Analysis (PCA). Differential expression testing was conducted using DESeq2 library 1.24.0 [52] scripts indicating the significance (Benjamini-Hochberg adjusted p-values, false discovery rate  $FDR < 0.05$ ) and the effect size (fold-change) for each comparison. Splicing analysis of the *Irf3* gene was performed using Majiq 2.4 [53] with default parameters to investigate alternative transcripts between genotypes.

## Supporting information

**S1 Fig. *Ifna4* expression is delayed in ZIKV-infected CC071 MEFs.** MEFs derived from B6 (gray circles), CC001 (blue triangles) and CC071 (red squares) were infected with ZIKV at a MOI of 5. *Ifna4* expression was determined by RT-qPCR on MEFs total RNA by normalizing to *Tbp* housekeeping gene. Data are mean +/- sem from 3 biological replicates. For one CC001, one B6 and one CC071 replicates at 0 hpi and one CC071 replicate at 24 hpi, gene expression was below the limit of detection. Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by post-hoc Tukey HSD, \*/# p < 0.05, \*\*\* p < 0.001).  
(TIF)

**S2 Fig. Delayed IFN-I expression in CC071 MEFs results from a constitutive defect in the transcription activation cascade.** MEFs were transfected with either (A) poly(I:C), which activates both TLR and RIG-I pathways, or (B) 3p-hpRNA, a RIG-I agonist. *Ifnb1* expression was determined as in Fig 1B. Data are mean +/- sem from 3 biological replicates for CC001 and CC071 (2 at 4 hours) or 2 biological replicates for B6 (1 at 4 hours). Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by post-hoc Tukey HSD, \*/# p < 0.05).  
(TIF)

**S3 Fig. Expression of genes involved in the *Ifnb1* induction pathway in ZIKV-infected MEFs of B6, CC001 and CC071 strains.** MEFs derived from B6 (black circles), CC001 (blue triangles) and CC071 (red squares) were infected with ZIKV at a MOI of 5. mRNA expression levels were measured by RNAseq in non-infected (NI) and ZIKV-infected MEFs. (A) Number of upregulated genes per strain at 16, 24 and 32 hpi (log2 fold-change > 1, FDR = 0.05). (B-G) Genes constitutively expressed. (H-N) Genes which expression is induced by the IFN-I response (ISGs). Expression levels are shown on a logarithmic scale. For *Ifnb1* expression, null counts were transformed to 1. Data are mean +/- sem from 3 biological replicates. Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by post-hoc Tukey HSD, \*/# p < 0.05, \*\*/## p < 0.01, \*\*\*/### p < 0.001).  
(TIF)

**S4 Fig. (CC001xCC071)F1 MEFs show normal induction of *Ifnb1*.** *Ifnb1* expression upon ZIKV infection in CC001, CC071 and (CC001xCC071)F1 MEFs determined as in Fig 1B. Data are mean +/- sem from 3 biological replicates. Blue asterisks and orange hashes show statistical significance of CC071 compared to CC001 and to F1, respectively (ANOVA followed by post-hoc Tukey HSD, \*\*/## p < 0.01).  
(TIF)

**S5 Fig. Genotyping of (CC071 x *Irf3* KO) x CC071 backcross individuals.** Backcross individuals were genotyped by PCR with primers amplifying the intro 6 of *Irf3* (sequences in S1 Table). The *Irf3*<sup>KO</sup> allele results in a band at 877pb while the *Irf3*<sup>71</sup> allele results in no band. Individuals 1, 5, 11, 14, 15, 16, 17, 19, 21, 22 and 23 are *Irf3*<sup>KO/71</sup> and individuals 2, 6, 7, 8, 9, 10, 12, 13 and 24 are *Irf3*<sup>71/71</sup>.  
(TIF)

**S1 Table. Sequence of primers used for genotyping or qPCR.**  
(PDF)

**S2 Table. Antibodies used for immunofluorescence or western blot assays.**

(PDF)

**S1 Data. Excel spreadsheet containing, in separate sheets, the underlying numerical data for Figs 1A, 1B, 1C, 2B, 3A, 3B, 3C, 4A, 5D, 6A, 6B, 7A, 7B, S1, S2A, S2B, S3A, S3B–S3N and S4.**

(XLSX)

**Acknowledgments**

We thank Matthieu Prot, Maxime Chazal and Sandrine Vandormael-Pournin for technical advice and for providing reagents. We are grateful to Tommy Penel and Rachid Chennouf of the Institut Pasteur Central Animal Facility of the C2RA (Center for Animal Resources and Research) for the careful breeding of CC strains and for the maintenance of mice in the BSL-3 animal facility, respectively. We thank Etienne Simon-Loriere and Nolwenn Jouvenet for scientific advice and continuous support along this project.

**Author Contributions**

**Conceptualization:** Marie Bourdon, Caroline Manet, Eliette Bonnefoy, Xavier Montagutelli.

**Data curation:** Marie Bourdon, Caroline Manet, Etienne Kornobis, Xavier Montagutelli.

**Formal analysis:** Marie Bourdon, Caroline Manet, Etienne Kornobis, Xavier Montagutelli.

**Funding acquisition:** Eliette Bonnefoy, Xavier Montagutelli.

**Investigation:** Marie Bourdon, Caroline Manet, Laurine Conquet, Corentin Ramaugé Parra.

**Methodology:** Marie Bourdon, Caroline Manet, Etienne Kornobis, Xavier Montagutelli.

**Project administration:** Eliette Bonnefoy, Xavier Montagutelli.

**Resources:** Marie Bourdon, Caroline Manet, Laurine Conquet.

**Software:** Marie Bourdon, Etienne Kornobis.

**Supervision:** Xavier Montagutelli.

**Validation:** Xavier Montagutelli.

**Visualization:** Marie Bourdon, Caroline Manet, Etienne Kornobis.

**Writing – original draft:** Marie Bourdon, Caroline Manet, Xavier Montagutelli.

**Writing – review & editing:** Marie Bourdon, Caroline Manet, Laurine Conquet, Corentin Ramaugé Parra, Etienne Kornobis, Eliette Bonnefoy, Xavier Montagutelli.

**References**

1. Cao-Lormeau VM, Blake A, Mons S, Lastère S, Roche C, Vanhomwegen J, et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet*. 9 avr 2016; 387(10027):1531–9.
2. Carod-Artal FJ. Neurological complications of Zika virus infection. *Expert Review of Anti-infective Therapy*. 4 mai 2018; 16(5):399–410. <https://doi.org/10.1080/14787210.2018.1466702> PMID: 29668332
3. Oliveira Melo AS, Malinger G, Ximenes R, Szejnfeld PO, Alves Sampaio S, Bispo de Filippis AM. Zika virus intrauterine infection causes fetal brain abnormality and microcephaly: tip of the iceberg?: *Physician Alert. Ultrasound Obstet Gynecol*. janv 2016; 47(1):6–7.

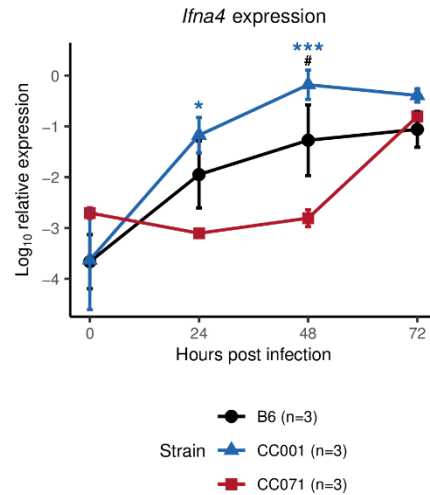


4. Paixao ES, Cardim LL, Costa MCN, Brickley EB, de Carvalho-Sauer RCO, Carmo EH, et al. Mortality from Congenital Zika Syndrome—Nationwide Cohort Study in Brazil. *N Engl J Med*. 24 févr 2022; 386(8):757–67. <https://doi.org/10.1056/NEJMoa2101195> PMID: 35196428
5. Rossi SL, Ebel GD, Shan C, Shi PY, Vasilakis N. Did Zika Virus Mutate to Cause Severe Outbreaks? *Trends in Microbiology*. 1 oct 2018; 26(10):877–85. <https://doi.org/10.1016/j.tim.2018.05.007> PMID: 29903417
6. Liu ZY, Shi WF, Qin CF. The evolution of Zika virus from Asia to the Americas. *Nat Rev Microbiol*. mars 2019; 17(3):131–9.
7. Manet C, Roth C, Tawfik A, Cantaert T, Sakuntabhai A, Montagutelli X. Host genetic control of mosquito-borne Flavivirus infections. *Mamm Genome*. 2018; 29(7–8):384–407. <https://doi.org/10.1007/s00335-018-9775-2> PMID: 30167843
8. Rasmussen AL, Okumura A, Ferris MT, Green R, Feldmann F, Kelly SM, et al. Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. *Science*. 21 nov 2014; 346(6212):987–91. <https://doi.org/10.1126/science.1259595> PMID: 25359852
9. Gralinski LE, Menachery VD, Morgan AP, Totura AL, Beall A, Kocher J, et al. Allelic Variation in the Toll-Like Receptor Adaptor Protein Ticam2 Contributes to SARS-Coronavirus Pathogenesis in Mice. *G3 (Bethesda)*. 7 juin 2017; 7(6):1653–63. <https://doi.org/10.1534/g3.117.041434> PMID: 28592648
10. Lazear HM, Govero J, Smith AM, Platt DJ, Fernandez E, Miner JJ, et al. A Mouse Model of Zika Virus Pathogenesis. *Cell Host Microbe*. 11 mai 2016; 19(5):720–30. <https://doi.org/10.1016/j.chom.2016.03.010> PMID: 27066744
11. Grant A, Ponia SS, Tripathi S, Balasubramaniam V, Miorin L, Sourisseau M, et al. Zika Virus Targets Human STAT2 to Inhibit Type I Interferon Signaling. *Cell Host & Microbe*. juin 2016; 19(6):882–90.
12. Sheehan KCF, Lai KS, Dunn GP, Bruce AT, Diamond MS, Heutel JD, et al. Blocking Monoclonal Antibodies Specific for Mouse IFN- $\alpha$  /  $\beta$  Receptor Subunit 1 (IFNAR-1) from Mice Immunized by In Vivo Hydrodynamic Transfection. *Journal of Interferon & Cytokine Research*. nov 2006; 26(11):804–19.
13. Roberts A, Pardo-Manuel de Villena F, Wang W, McMillan L, Threadgill DW. The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics. *Mamm Genome*. juill 2007; 18(6–7):473–81. <https://doi.org/10.1007/s00335-007-9045-1> PMID: 17674098
14. Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. sept 2011; 477(7364):289–94. <https://doi.org/10.1038/nature10413> PMID: 21921910
15. Manet C, Simon-Lorière E, Jouvion G, Hardy D, Prot M, Conquet L, et al. Genetic Diversity of Collaborative Cross Mice Controls Viral Replication, Clinical Severity, and Brain Pathology Induced by Zika Virus Infection, Independently of Oas1b. *J Virol [Internet]*. 17 janv 2020; 94(3). Disponible sur: <https://journals.asm.org/doi/10.1128/JVI.01034-19> PMID: 31694939
16. Jasperse BA, Mattocks MD, Noll KM, Ferris MT, Heise MT, Lazear HM. Neuroinvasive flavivirus pathogenesis is restricted by host genetic factors in Collaborative Cross mice, independently of Oas1b [Internet]. *bioRxiv*; 2022 [cité 26 oct 2022]. Disponible sur: <https://www.biorxiv.org/content/10.1101/2022.10.24.513634v1>
17. Cartwright HN, Barbeau DJ, Doyle JD, Klein E, Heise MT, Ferris MT, et al. Genetic diversity of collaborative cross mice enables identification of novel rift valley fever virus encephalitis model. *PLOS Pathogens*. 14 juill 2022; 18(7):e1010649. <https://doi.org/10.1371/journal.ppat.1010649> PMID: 35834486
18. McCartney S, Vermi W, Gilfillan S, Cella M, Murphy TL, Schreiber RD, et al. Distinct and complementary functions of MDA5 and TLR3 in poly(I:C)-mediated activation of mouse NK cells. *Journal of Experimental Medicine*. 21 déc 2009; 206(13):2967–76. <https://doi.org/10.1084/jem.20091181> PMID: 19995959
19. Collaborative Cross Consortium. The Genome Architecture of the Collaborative Cross Mouse Genetic Reference Population. *Genetics*. 1 févr 2012; 190(2):389–401. <https://doi.org/10.1534/genetics.111.132639> PMID: 22345608
20. Srivastava A, Morgan AP, Najarian ML, Sarsani VK, Sigmon JS, Shorter JR, et al. Genomes of the Mouse Collaborative Cross. *Genetics*. 1 juin 2017; 206(2):537–56. <https://doi.org/10.1534/genetics.116.198838> PMID: 28592495
21. Brown AJ, Won JJ, Wolfisberg R, Fahnøe U, Catanzaro N, West A, et al. Host genetic variation guides hepatitis C virus clearance, chronicity, and liver fibrosis in mice [Internet]. *Microbiology*; 2023 mars [cité 24 mars 2023]. Disponible sur: <http://biorxiv.org/lookup/doi/10.1101/2023.03.18.533278>
22. Zhang J, Teh M, Kim J, Eva MM, Cayrol R, Meade R, et al. A Loss-of-Function Mutation in the Integrin Alpha L (Itgal) Gene Contributes to Susceptibility to Salmonella enterica Serovar Typhimurium Infection in Collaborative Cross Strain CC042. *Infect Immun*. 17 2019; 88(1). <https://doi.org/10.1128/IAI.00656-19> PMID: 31636138

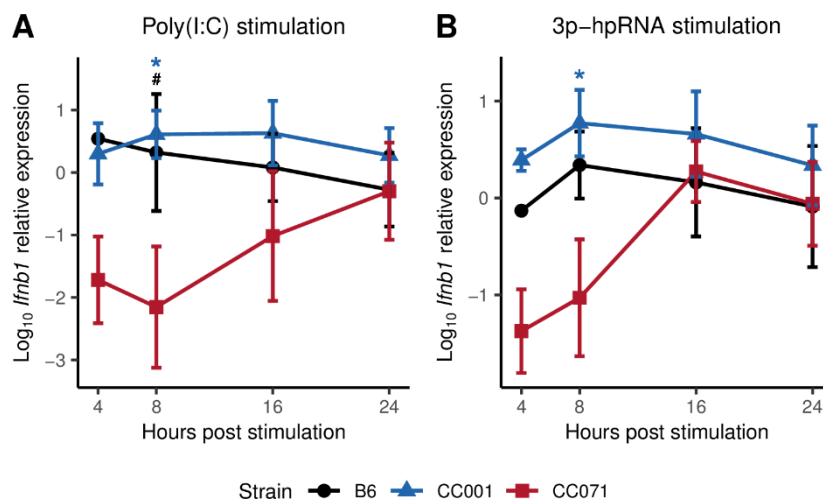
23. Smith CM, Proulx MK, Lai R, Kiritsy MC, Bell TA, Hock P, et al. Functionally Overlapping Variants Control Tuberculosis Susceptibility in Collaborative Cross Mice. Vance R, éditeur. *mBio*. 24 déc 2019; 10(6):e02791–19.
24. Garifulin O, Qi Z, Shen H, Patnala S, Green MR, Boyartchuk V. Irf3 Polymorphism Alters Induction of Interferon Beta in Response to *Listeria monocytogenes* Infection. *PLoS Genetics*. 7 sept 2007; 3(9): e152. <https://doi.org/10.1371/journal.pgen.0030152> PMID: 17845078
25. Moore TC, Al-Salleh FM, Brown DM, Petro TM. IRF3 polymorphisms induce different innate anti-Theiler's virus immune responses in RAW264.7 macrophages. *Virology*. sept 2011; 418(1):40–8. <https://doi.org/10.1016/j.virol.2011.06.028> PMID: 21810534
26. Bourdon M, Manet C, Montagutelli X. Host genetic susceptibility to viral infections: the role of type I interferon induction. *Genes Immun*. déc 2020; 21(6–8):365–79. <https://doi.org/10.1038/s41435-020-00116-2> PMID: 33219336
27. Schoggins JW. Interferon-Stimulated Genes: What Do They All Do? *Annu Rev Virol*. 29 sept 2019; 6(1):567–84.
28. Chattopadhyay S, Kuzmanovic T, Zhang Y, Wetzel JL, Sen GC. Ubiquitination of the Transcription Factor IRF-3 Activates RIPA, the Apoptotic Pathway that Protects Mice from Viral Pathogenesis. *Immunity*. 17 mai 2016; 44(5):1151–61. <https://doi.org/10.1016/j.immuni.2016.04.009> PMID: 27178468
29. Popli S, Chakravarty S, Fan S, Glanz A, Aras S, Nagy LE, et al. IRF3 inhibits nuclear translocation of NF- $\kappa$ B to prevent viral inflammation. *Proc Natl Acad Sci USA*. 13 sept 2022; 119(37):e2121385119.
30. Daffis S, Samuel MA, Keller BC, Gale M Jr, Diamond MS. Cell-specific IRF-3 responses protect against West Nile virus infection by interferon-dependent and -independent mechanisms. *PLoS Pathog*. 27 juill 2007; 3(7):e106. <https://doi.org/10.1371/journal.ppat.0030106> PMID: 17676997
31. Menachery VD, Leib DA. Control of Herpes Simplex Virus Replication Is Mediated through an Interferon Regulatory Factor 3-Dependent Pathway. *J Virol*. déc 2009; 83(23):12399–406. <https://doi.org/10.1128/JVI.00888-09> PMID: 19759149
32. Honda K, Yanai H, Negishi H, Asagiri M, Sato M, Mizutani T, et al. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature*. 7 avr 2005; 434(7034):772–7. <https://doi.org/10.1038/nature03464> PMID: 15800576
33. Schilte C, Buckwalter MR, Laird ME, Diamond MS, Schwartz O, Albert ML. Cutting Edge: Independent Roles for IRF-3 and IRF-7 in Hematopoietic and Nonhematopoietic Cells during Host Response to Chikungunya Infection. *Jl*. 1 avr 2012; 188(7):2967–71. <https://doi.org/10.4049/jimmunol.1103185> PMID: 22371392
34. Manet C, Mansuroglu Z, Conquet L, Bortolin V, Comptdaer T, Segrt H, et al. Zika virus infection of mature neurons from immunocompetent mice generates a disease-associated microglia and a tauopathy-like phenotype in link with a delayed interferon beta response. *J Neuroinflammation*. 20 déc 2022; 19(1):307. <https://doi.org/10.1186/s12974-022-02668-8> PMID: 36539803
35. Chen HW, King K, Tu J, Sanchez M, Luster AD, Shresta S. The roles of IRF-3 and IRF-7 in innate antiviral immunity against dengue virus. *J Immunol*. 15 oct 2013; 191(8):4194–201. <https://doi.org/10.4049/jimmunol.1300799> PMID: 24043884
36. Canivet C, Rhéaume C, Lebel M, Piret J, Gosselin J, Boivin G. Both IRF3 and especially IRF7 play a key role to orchestrate an effective cerebral inflammatory response in a mouse model of herpes simplex virus encephalitis. *J Neurovirol*. déc 2018; 24(6):761–8.
37. Bigam AW, Buckingham KJ, Husain S, Emond MJ, Bofferding KM, Gildersleeve H, et al. Host genetic risk factors for West Nile virus infection and disease progression. *PLoS One*. 2011; 6(9):e24745. <https://doi.org/10.1371/journal.pone.0024745> PMID: 21935451
38. Cahill ME, Conley S, DeWan AT, Montgomery RR. Identification of genetic variants associated with dengue or West Nile virus disease: a systematic review and meta-analysis. *BMC Infect Dis*. déc 2018; 18(1):282. <https://doi.org/10.1186/s12879-018-3186-6> PMID: 29929468
39. Mork N, Kofod-Olsen E, Sorensen KB, Bach E, Ormtoft TF, Ostergaard L, et al. Mutations in the TLR3 signaling pathway and beyond in adult patients with herpes simplex encephalitis. *Genes Immun*. déc 2015; 16(8):552–66. <https://doi.org/10.1038/gene.2015.46> PMID: 26513235
40. Andersen LL, Mork N, Reinert LS, Kofod-Olsen E, Narita R, Jorgensen SE, et al. Functional IRF3 deficiency in a patient with herpes simplex encephalitis. *J Exp Med*. 24 août 2015; 212(9):1371–9. <https://doi.org/10.1084/jem.20142274> PMID: 26216125
41. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* [Internet]. 24 sept 2020; Disponible sur: <https://www.ncbi.nlm.nih.gov/pubmed/32972995> <https://science.sciencemag.org/content/sci/370/6515/eabd4570.full.pdf>

42. Sato M, Suemori H, Hata N, Asagiri M, Ogasawara K, Nakao K, et al. Distinct and essential roles of transcription factors IRF-3 and IRF-7 in response to viruses for IFN-alpha/beta gene induction. *Immunity*. oct 2000; 13(4):539–48. [https://doi.org/10.1016/s1074-7613\(00\)00053-4](https://doi.org/10.1016/s1074-7613(00)00053-4) PMID: 11070172
43. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. 2021. Disponible sur: <https://www.R-project.org/>
44. Venables WN, Ripley BD. Modern Applied Statistics with S. Fourth Edition [Internet]. Fourth. Springer; 2002. Disponible sur: <https://www.stats.ox.ac.uk/pub/MASS4/>
45. Bourdon M, Montagutelli X. stuart: an R package for the curation of SNP genotypes from experimental crosses. *G3 (Bethesda)*. 24 août 2022;jkac219.
46. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. *Bioinformatics*. 1 mai 2003; 19(7):889–90. <https://doi.org/10.1093/bioinformatics/btg112> PMID: 12724300
47. Cokelaer T, Desvillechabrol D, Legendre R, Cardon M. « Sequana »: a Set of Snakemake NGS pipelines. *JOSS*. 30 août 2017; 2(16):352.
48. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*. 1 oct 2012; 28(19):2520–2. <https://doi.org/10.1093/bioinformatics/bts480> PMID: 22908215
49. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 1 janv 2013; 29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635> PMID: 23104886
50. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 1 avr 2014; 30(7):923–30. <https://doi.org/10.1093/bioinformatics/btt656> PMID: 24227677
51. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 1 oct 2016; 32(19):3047–8. <https://doi.org/10.1093/bioinformatics/btw354> PMID: 27312411
52. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. déc 2014; 15(12):550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
53. Vaquero-Garcia J, Barrera A, Gazzara MR, González-Vallinas J, Lahens NF, Hogenesch JB, et al. A new view of transcriptome complexity and regulation through the lens of local splicing variations. *eLife*. 1 févr 2016; 5:e11752. <https://doi.org/10.7554/eLife.11752> PMID: 26829591

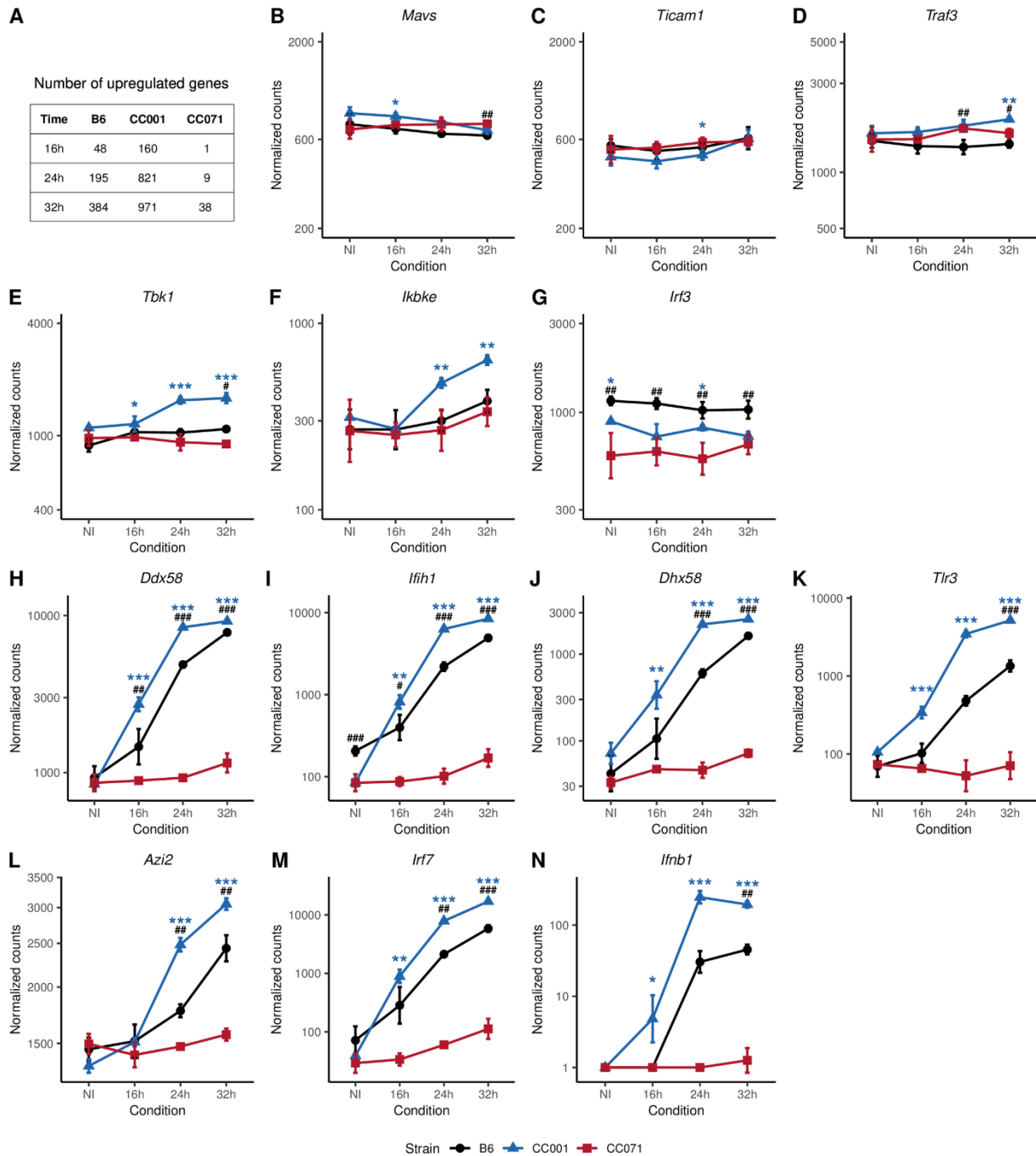
## Supplementary figures



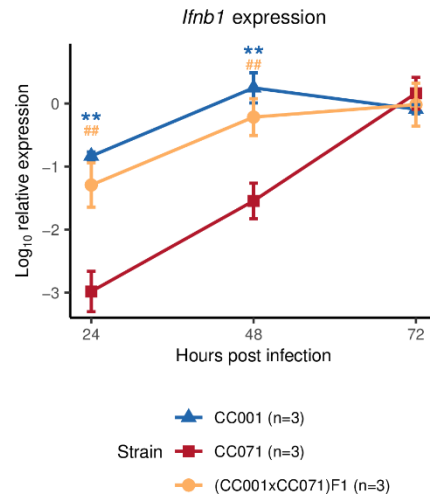
**S1 Fig. *Ifna4* expression is delayed in ZIKV-infected CC071 MEFs.** MEFs derived from B6 (gray circles), CC001 (blue triangles) and CC071 (red squares) were infected with ZIKV at a MOI of 5. *Ifna4* expression was determined by RT-qPCR on MEFs total RNA by normalizing to *Tbp* housekeeping gene. Data are mean +/- sem from 3 biological replicates. For one CC001, one B6 and one CC071 replicates at 0 hpi and one CC071 replicate at 24 hpi, gene expression was below the limit of detection. Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by post-hoc Tukey HSD, \*/#  $p < 0.05$ , \*\*\*  $p < 0.001$ ).



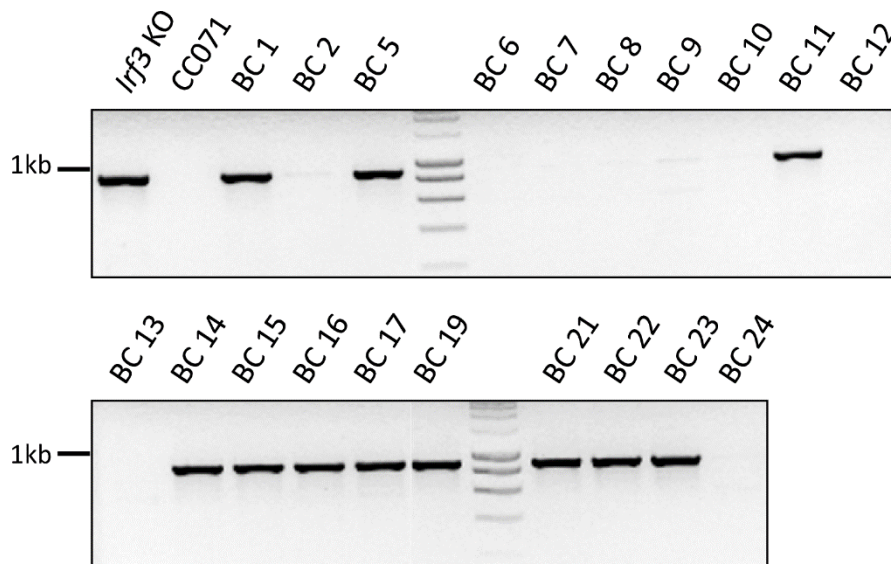
**S2 Fig. Delayed IFN-I expression in CC071 MEFs results from a constitutive defect in the transcription activation cascade.** MEFs were transfected with either (A) poly(I:C), which activates both TLR and RIG-I pathways, or (B) 3p-hpRNA, a RIG-I agonist. *Ifnb1* expression was determined as in Figure 1B. Data are mean +/- sem from 3 biological replicates for CC001 and CC071 (2 at 4 hours) or 2 biological replicates for B6 (1 at 4 hours). Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by post-hoc Tukey HSD, \*/#  $p < 0.05$ ).



**S3 Fig. Expression of genes involved in the *Ifnb1* induction pathway in ZIKV-infected MEFs of B6, CC001 and CC071 strains.** MEFs derived from B6 (black circles), CC001 (blue triangles) and CC071 (red squares) were infected with ZIKV at a MOI of 5. mRNA expression levels were measured by RNAseq in non-infected (NI) and ZIKV-infected MEFs. (A) Number of upregulated genes per strain at 16, 24 and 32 hpi (log<sub>2</sub> fold-change > 1, FDR = 0.05). (B-G) Genes constitutively expressed. (H-N) Genes which expression is induced by the IFN-I response (ISGs). Expression levels are shown on a logarithmic scale. For *Ifnb1* expression, null counts were transformed to 1. Data are mean +/- sem from 3 biological replicates. Blue asterisks and black hashes show statistical significance of CC071 compared to CC001 and to B6, respectively (ANOVA followed by post-hoc Tukey HSD, \*/# p < 0.05, \*\*/### p < 0.01, \*\*\*/#### p < 0.001).



**S4 Fig. (CC001xCC071)F1 MEFs show normal induction of *Ifnb1*.** *Ifnb1* expression upon ZIKV infection in CC001, CC071 and (CC001xCC071)F1 MEFs determined as in Figure 1B. Data are mean +/- sem from 3 biological replicates. Blue asterisks and orange hashes show statistical significance of CC071 compared to CC001 and to F1, respectively (ANOVA followed by post-hoc Tukey HSD, \*\*/## p < 0.01).



**S5 Fig. Genotyping of (CC071 x *Irf3* KO) x CC071 backcross individuals.** Backcross individuals were genotyped by PCR with primers amplifying the intro 6 of *Irf3* (sequences in S1 Table). The *Irf3*<sup>KO</sup> allele results in a band at 877pb while the *Irf3*<sup>71</sup> allele results in no band. Individuals 1, 5, 11, 14, 15, 16, 17, 19, 21, 22 and 23 are *Irf3*<sup>KO/71</sup> and individuals 2, 6, 7, 8, 9, 10, 12, 13 and 24 are *Irf3*<sup>71/71</sup>.

## V.2 . Deciphering the *in vivo* susceptibility to ZIKV of CC071 mice

### V.2.1 . Introduction

In the PLOS Pathogens article, we demonstrated that CC071 mice carry a loss-of-function mutation in *Irf3* responsible for delayed IFN-I induction and high viral replication in MEFs, but that other alleles are required to explain the susceptibility of CC071 mice to ZIKV infection *in vivo* (Bourdon et al. 2023). To identify these alleles, here, we performed genetic mapping analyses on two crosses.

First, as *Irf3* was responsible for the high viral replication in MEFs, but not for the high PVL in CC071 mice *in vivo*, I analyzed more mice from our (CC071 × C57BL/6J-*Irf3*<sup>-/-</sup> (B6-*Irf3*)) × CC071 N2 cohort (hereinafter designated as N2-*Irf3*) to identify QTLs associated with the PVL. Since all N2-*Irf3* mice carry two non-functional *Irf3* alleles, this cross maximizes the power to detect other susceptibility alleles.

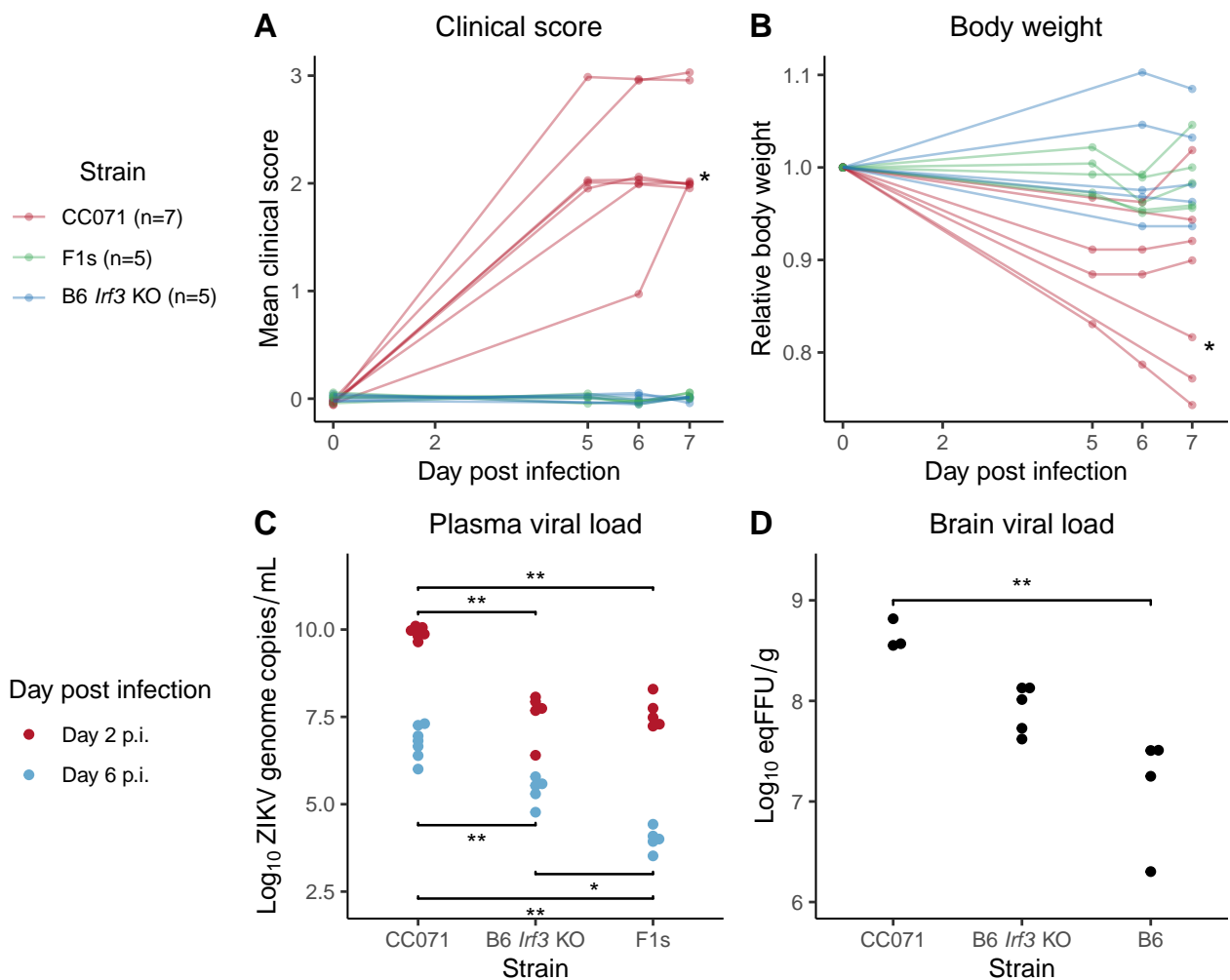
Separately, I searched for QTLs explaining the severe illness observed in CC071 mice after ZIKV infection. Thus, to favor the apparition of symptoms in infected mice, I have used a more severe model by crossing CC071 mice with *Ifnar1* deficient mice. An F2 population was produced between CC071 and 129-*Ifnar1* and only *Ifnar1*<sup>-/-</sup> F2s mice were analyzed. We chose the 129-*Ifnar1* strain as it was more resistant to ZIKV infection (Manet et al. 2020). Besides, since all mice had an abolished IFN-I response, it facilitated the identification of susceptibility genes outside this pathway.

### V.2.2 . Results

#### V.2.2.1 . Phenotypic analysis of CC071, B6-*Irf3*, F1 and N2 mice

To get a first insight into the inheritance of susceptibility alleles, CC071, B6-*Irf3* and CC071 × B6-*Irf3* F1 males and females were infected with prior MAR1-5A3 treatment. Mice were followed for 7 days. As described previously, B6-*Irf3* and F1 mice did not develop signs of illness and did not lose weight, while CC071 mice showed ruffled fur and hunched posture, with a clinical score of 2 or 3 (out of 5) at 7 dpi (Figure 18A). Most CC071 mice lost between 5 and 25% of their original body weight by 7 dpi, although one CC071 mouse lost only 4% of its initial body weight between 5 and 6 dpi and returned to its original weight by 7 dpi (Figure 18B). This difference was not explained by the sex of mice as only 1 female CC071 was tested. CC071 mice were between 6 and 9 week-old, and older mice tend to lose less weight than younger mice within this age range. Thus, the interindividual variability of body weight loss in CC071 may be explained by the age of mice. Clinical scores and body weight loss of CC071 were not correlated, as the body weight loss at 7 dpi of mice with a clinical score of 2 ranged between 0 and 30%. It may be due to the infection dose which may not be high enough to induce competent susceptibility in CC071 mice. This phenotypic diversity within the CC071 strain may hamper further QTL mapping analyses in the N2-*Irf3* progeny by adding noise, thus reducing the power to identify host factors controlling body weight loss.

As reported above, B6-*Irf3* and F1 mice had similar lower PVL than CC071 at 2 dpi. At 6 dpi, B6-*Irf3* mice had also lower PVL than CC071, and the PVL of F1 mice was even lower (Figure 18C). As a result, the decrease between the PVLs at 2 and 6 dpi was similar in F1 and CC071 mice. The absence of



**Figure 18:** B6 *lrf3* KO mice do not develop ZIKV disease.

(A-C) Six to 9-week-old CC071 (n=7), B6-*lrf3* (n=5) and F1 (n=5) mice were infected i.p. with  $10^7$  FFUs of FG15 one day after treatment with 2 mg of MAR1-5A3 monoclonal antibody and monitored for 7 days. (A) Curves of individual clinical scores expressed with numerical values given as follows: 1, slight hunched posture; 2, ruffled fur, hunched posture and/or mild ataxia; 3, prostration, ataxia, partial lib paralysis (endpoint for euthanasia); 4, ataxia, total limb paralysis or moribund; 5, dead. (B) Curves of individual body weight loss, expressed as the proportion of the starting weight at the day of infection. Mice were weighed at day 5, 6, and 7 post infection. (C) Plasma viral loads, quantified at days 2 and 6 post infection by RT-qPCR. (D) 5 B6, 3 CC071 (presented in A - C), and 5 B6-*lrf3* mice (presented in A - C) were euthanized at 7 dpi and brain was collected. BVL were quantified by RT-qPCR. (A-B) Groups were compared by linear mixed model analysis. The \* represents the significant differences between CC071 and B6-*lrf3*, and between CC071 and F1. (C) Groups were compared by mixed two-way ANOVA testing for the effect of the time post infection and the strain. Then, groups were compared by pairwise Wilcoxon test with Bonferroni correction. (D) Groups were compared by Kruskal-Wallis test followed by Dunn test (\*  $p < 0.05$ , \*\*  $p < 0.001$ ).



correlation between peak PVL and PVL decrease in the three groups, which was previously observed in CC strains (Manet et al. 2020), suggests distinct genetic control of these two traits in B6-*Irf3* and CC071 strains

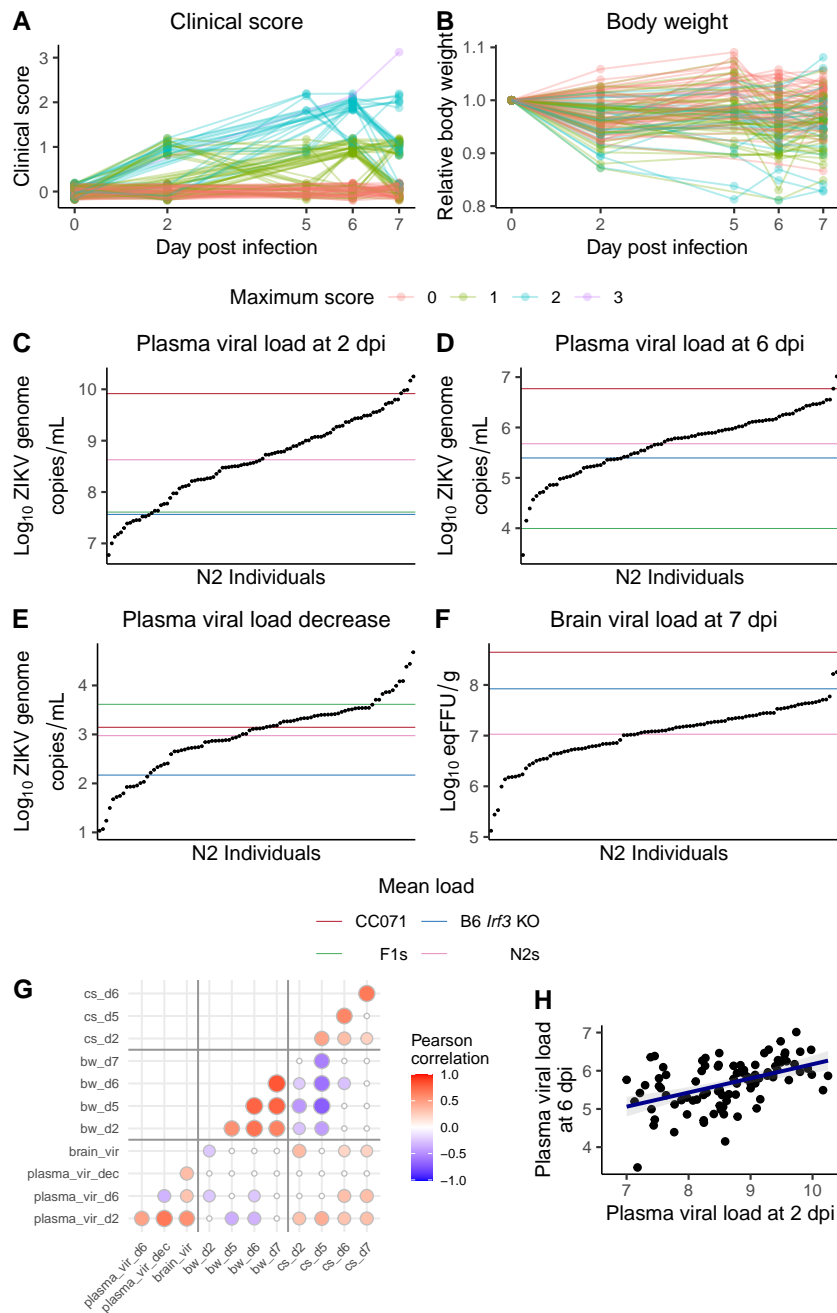
CC071 mice were previously reported with elevated brain viral loads (BVLs) after ZIKV infection (Manet et al. 2020). To assess whether the *Irf3* mutation could contribute to this trait, we measured BVLs at 7 dpi in B6-*Irf3*, B6 WT and CC071 mice. B6-*Irf3* mice had higher loads compared with B6 WT mice, but lower than CC071 mice, although the results of the post-hoc tests after comparisons by ANOVA were not significant (Figure 18D). This result suggests that the *Irf3* mutation in CC071 mice may contribute to their high brain viral loads, along with other susceptibility alleles.

One hundred and four N2-*Irf3* mice were produced and infected under the same experimental conditions. More mice will be infected and added to the analysis to increase the sample size and the QTL detection power. N2-*Irf3* mice had variable clinical score kinetics (Figure 19A). 41.2% of them did not show any sign of illness from 0 to 7 dpi. 36.8% had a maximum clinical score of 1, with different profiles. While some reached their maximum score at 2 dpi and then recovered, others peaked at later time points, between 5 and 7 dpi. Three N2 mice had a clinical score of 1 from 2 to 7 dpi. 21% of N2-*Irf3* mice had a maximum clinical score of 2. This score was reached between 5 and 6 dpi, and most mice then showed reduced symptoms at 7 dpi. However, 6 of them remained at a score of 2 at 7 dpi. Only one N2-*Irf3* mouse had a maximum score of 3, reached at 7 dpi. The mice that had a clinical score of 0 for 7 dpi lost less weight than the other mice (linear mixed model,  $p = 0.005$  for the comparisons with mice having a maximum score of 1 and  $p = 0.007$  with mice having a maximum score of 2), but the same body weight kinetics were obtained for the mice with a maximum score of 1 or 2 ( $p = 0.97$ ). The mouse that reached a score of 3 showed only a weight loss of 0.93 (Figure 19B). Sex impacted the clinical score kinetics, as males had significantly higher scores than females, but not body weight loss (Figures S1C-D).

PVLs at 2 dpi of most N2-*Irf3* mice ranged between the average values observed in the parental strains, with only a few mice exceeding these values (Figure 19C). PVLs at 6 dpi were comprised between the values observed in CC071 and F1 mice (Figure 19D). Notably, the PVL decrease (defined as the difference between the viral loads at 2 and 6 dpi) in N2-*Irf3* mice extended beyond that of the parental strains (Figure 19E). Most N2-*Irf3* mice had values of BVL lower than the average values of CC071 or B6-*Irf3*. No F1 mice have been phenotyped for this trait yet, but I expect that as for the PVL at 6 dpi, F1 mice will have lower BVL than the parental strains and that the values of BVL in N2-*Irf3* mice range between the values observed in CC071 and F1 mice.

Correlations between the different phenotypes measured in the N2-*Irf3* progeny were tested. Groups of similar traits (i.e., viral loads, body weights, clinical scores, separated by horizontal and vertical lines on Figure 19G), showed significant intra-group correlation. PVLs at 2 and 6 were moderately but significantly correlated (Pearson correlation coefficient,  $R = 0.48$ ,  $p = 1.1e-6$ , Figure 19H). However, mice with similar PVLs at 2 dpi, but differing at 6 dpi by two logs were observed, suggesting different genetic control of the early and late viral loads (Figure 19G). As expected from its calculation, the PVL decrease was correlated to both the PVLs at 2 dpi and at 6 dpi.

Body weights between 2 and 7 dpi were strongly correlated, and significantly correlated with clinical scores, mostly at 5 dpi at which this parameter often peaked. Notably, clinical scores at all



**Figure 19:** Phenotypic variation and trait correlation in N2-*Irf3* mice.

Seven to 9-week-old N2 mice (n=104) were infected i.p. with  $10^7$  FFUs of FG15 one day after treatment with 2 mg of MAR1-5A3 monoclonal antibody and monitored for 7 days. (A) Curves of clinical scores, expressed with numerical values as in Figure 18A. (B) Body weight loss curves, expressed as the percent of the starting weight (day of infection). Mice were weighed at day 2, 5, 6, and 7 post infection. Curves were colored by the maximum clinical score of each mouse. (C - E) PVLs quantified at days 2 (C) and 6 (D) post infection by RT-qPCR, and PVL decrease calculated as the difference of log values between the load at 2 dpi and the load at 6 dpi (E). (F) BVLs quantified at 7 dpi by RT-qPCR. (C - F) Horizontal lines show the average PVL of CC071 (red), B6-*Irf3* (blue), F1s (green), and N2s (pink). (G) Correlation matrix between the phenotypes of the N2 mice. Horizontal and vertical lines separate categories of phenotypes. The strength of correlation is represented by both circle sizes and colors. Small white circles show nonsignificant correlations. (H) Correlation between viral loads at 2 and 6 dpi (Pearson correlation coefficient:  $R = 0.48$ ,  $p = 1.1 \times 10^{-6}$ ).

time points also showed some correlation with PVLs, mostly at 2 dpi. The clinical score at 5 dpi was not correlated with the score at 7 dpi, but this is likely due to the small number of mice (n = 39) which were scored at 5 dpi due to variations in the follow-up schedule.

### V.2.2.2 . QTL mapping identified two significant and five suggestive QTLs associated with viral loads of N2-*Irf3* mice

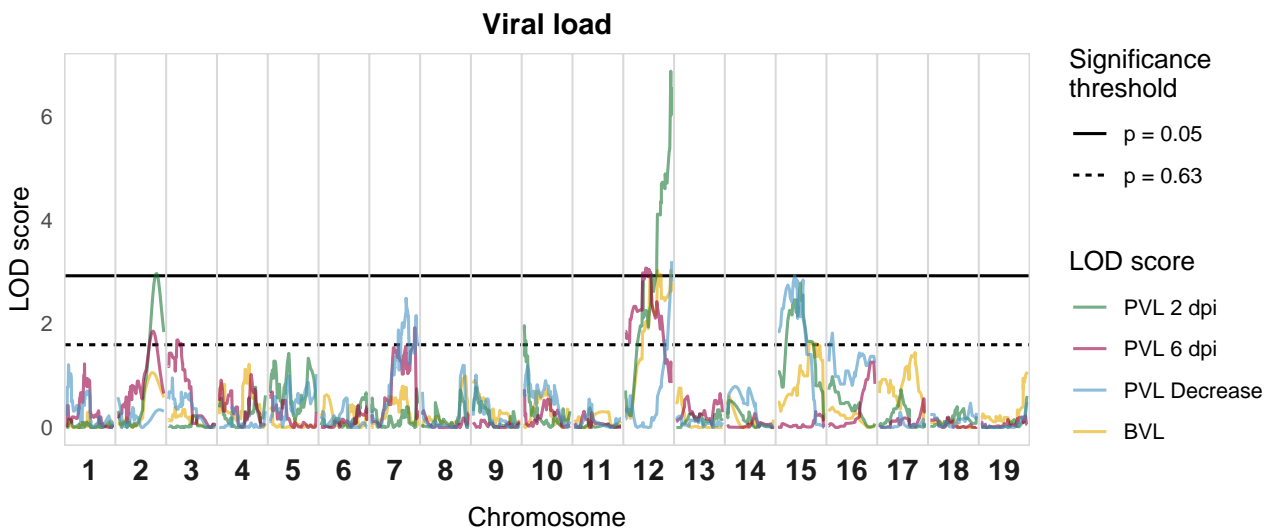
In order to find susceptibility alleles involved in the illness observed in CC071 mice after ZIKV infection, N2-*Irf3* mice were genotyped with the MiniMUGA array containing 11,125 markers spread across the genome. The stuart package (Bourdon and Montagutelli 2022) was used to curate the raw genotypes and retained 1,873 markers for QTL mapping (Figure S2A). Multiple QTLs were found in this cross and are presented in Table 1. For each QTL is indicated the LOD score, the p-value, the position with the Bayesian credible interval, and the direction of the allelic effect (which allele leads to a higher phenotype value).

**Table 1:** Associations between QTLs and phenotypes in the N2-*Irf3* progeny.

QTL	Phenotype	Chr	Position	LOD	p-value	CI (Mb)	PVE (%)	Effect
<i>Zs1</i>	PVL (2 dpi)	12	114.26	6.610	<0.001	105.46 - 119.37	28.04	71 > B6
<i>Zs1</i>	PVL (decrease)	12	119.37	3.205	0.025	100.9 - 119.4	15.13	71 > B6
<i>Zs1</i>	BVL	12	114.26	3.029	0.039	73.65 - 119.5	13.5	71 > B6
<i>Zs2</i>	PVL (6 dpi)	12	75.51	3.071	0.048	31.8 - 100.9	14.39	71 > B6
<i>Zs3</i>	BW (5 dpi)	7	29.65	3.558	0.005	24.0 - 44.2	28.92	71 > B6
-	PVL (2 dpi)	2	≈ 165	2.996	0.067	132.7 - 180.1	13.26	71 > B6
-	PVL (6 dpi)	2	≈ 160	1.857	0.503	50.1 - 180.1	8.97	71 > B6
-	PVL (6 dpi, scantwo)	3 & 12	68.72/75.51	7.7	NA	NA	NA	71 > B6
-	PVL (6 dpi)	7	134.39	1.920	0.458	39.2 - 137.5	9.23	71 > B6
-	PVL (decrease)	7	119.3	2.486	0.153	66.7 - 139.8	11.94	71 < B6
-	PVL (2 dpi)	15	64.76	2.783	0.089	21.3 - 90.3	12.50	71 > B6
-	PVL (decrease)	15	55.52	2.928	0.060	10.3 - 74.2	13.91	71 > B6
-	BW (7 dpi)	2	114.95	1.869	0.482	38.5 - 101.6	7.87	71 > B6
-	BW (5 dpi)	12	66.23	1.736	0.516	25.2 - 116.1	15.34	71 < B6
-	BW (6 dpi)	12	95.71	1.690	0.601	53.2 - 119.3	7.08	71 < B6
-	BW (7 dpi)	12	66.23	2.207	0.265	34.1 - 119.4	9.23	71 < B6
-	CS (7 dpi)	4	117.63	1.848	0.472	55.5 - 156.1	7.71	71 > B6
-	CS (6 dpi)	10	7.83	2.240	0.255	5.3 - 118.3	9.27	71 > B6
-	CS (5 dpi)	16	97.30	2.185	0.347	66.6 - 97.3	9.27	71 < B6

CI: credible interval, PVE: percentage of variance explained. BW: body weight. CS: clinical score. scantwo: two-dimensional genome scan. The p-value, CI, and PVE were not assessed for the interaction between the chromosome 3 and 12 loci associated with PVL at 6 dpi.

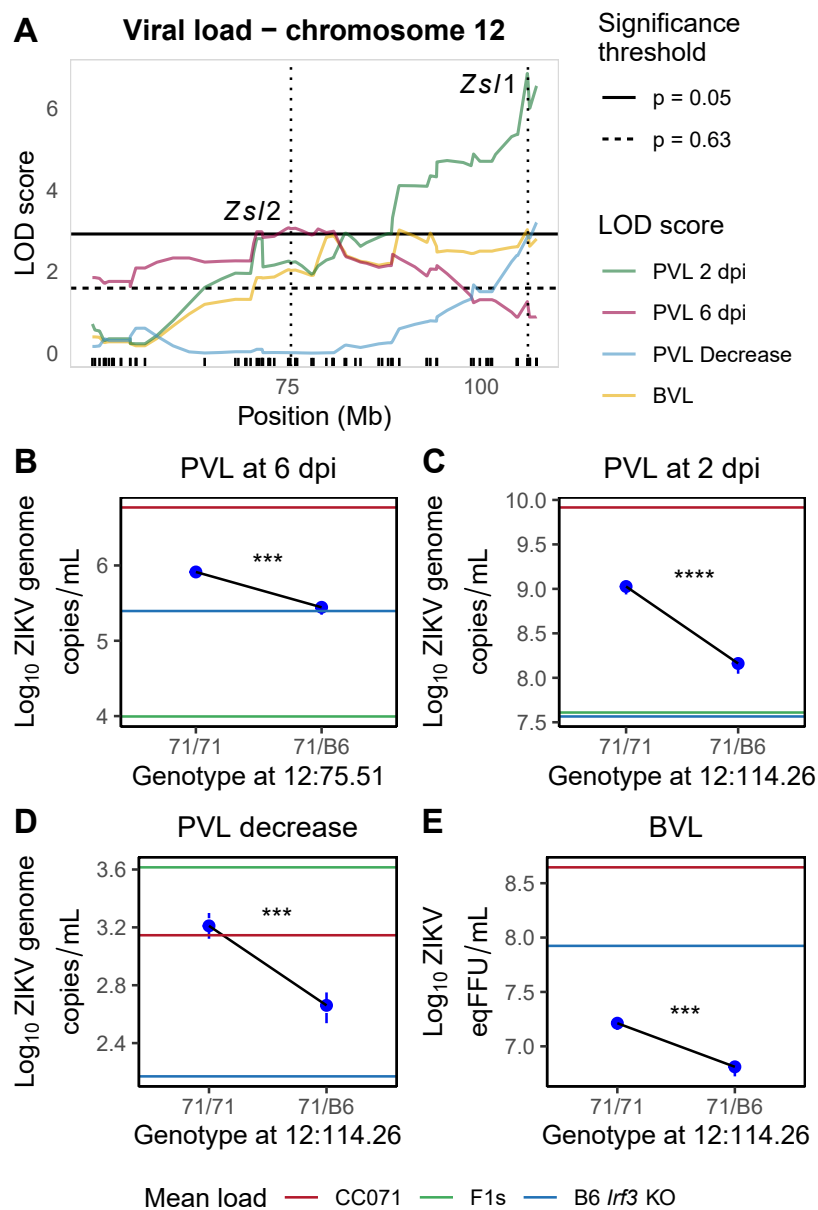
Using a single QTL model, I found one significant QTL and two suggestive QTLs controlling the PVL at 2 dpi. The significant QTL, named *Zika susceptibility locus 1 (Zs1)*, was positioned on distal chromosome 12 and peaked close to the end of the chromosome (Figures 20 and 21A). This QTL was highly significant, with a LOD score of 6.61 (Table 1). This was reflected in the allelic effects for the marker at the peak of the QTL. The difference in PVL at 2 dpi between the two genotype groups was about one log, as CC071/CC071 mice had a PVL of 9.02 while CC071/B6 mice had an average PVL of 8.16. This is more than a third of the difference between the PVL of B6-*Irf3* and CC071 mice (Figure 21B).



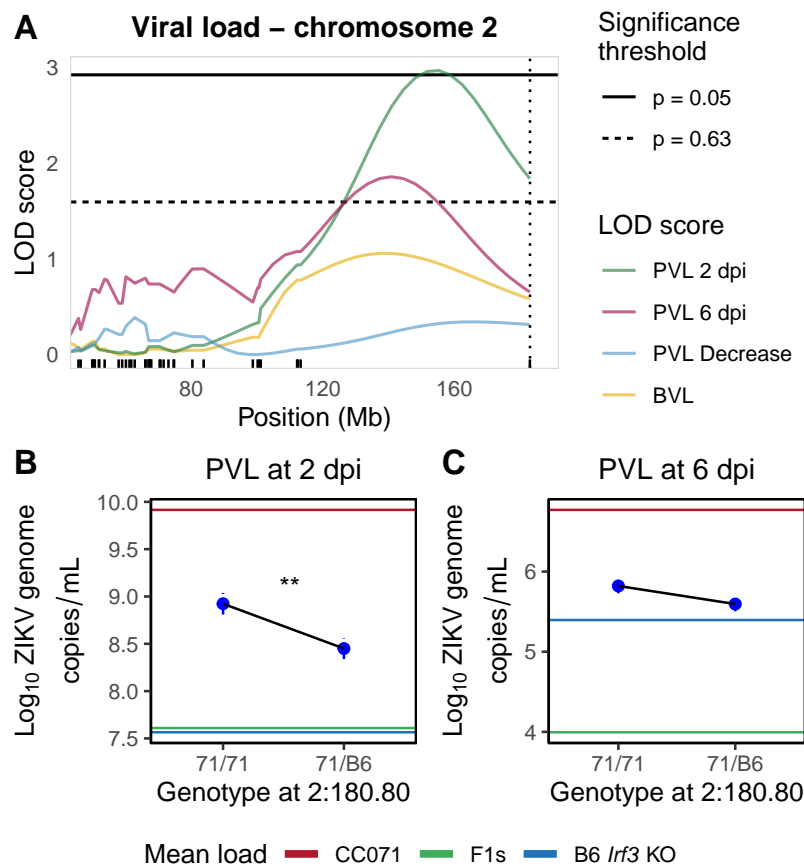
**Figure 20:** Single QTL mapping for the viral loads in the N2-*Irf3* progeny. QTL mapping was performed on PVLs at 2 and 6 dpi, on PVL decrease, and on BVL. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing (n=1000). For better visibility, and as the significant thresholds were very close for the three traits ( $p = 0.05$ , 3.11 for the PVL at 2 dpi, 3.06 for 6 dpi, 3.01 for the decrease, and 2.92 for the BVL;  $p = 0.63$ , 1.69 for the PVL at dpi, 1.70 for 6 dpi, 1.69 for the decrease, and 1.59 for the BVL), only the lower value was displayed.

The two suggestive but close to significance QTLs for the PVL at 2 dpi were detected on chromosomes 2 and 15, with a LOD score of 2.996 and 2.783, respectively (Figures 22 and 23). For these two QTLs, mice with the CC071/CC071 genotype had higher PVLs than CC071/B6 mice. Notably, the QTL on chromosome 2 was located in a region where no marker was retained for the QTL mapping analysis (Figure S2A), as CC071 inherited the B6 haplotype between 134 and 180 megabase (Mb) (Figure S2C). Therefore, since this region is not expected to contain any polymorphism between B6-*Irf3* and CC071, this QTL could be an artefact of interval mapping due to the lack of genotyped marker over a large interval. Alternatively, it could be a real QTL due to a private variant in CC071 or a small region of non-B6 origin missed during the CC haplotype reconstruction based on genotyping. These QTLs, as for the other suggestive QTLs detailed below, were not named and will be given a *Zs/1* number if the analysis of a larger number of mice confirms significant associations.

As each of the 3 QTLs detected for the PVL at 2 dpi explained a high proportion of the variance of this trait (more than 10% each, and 28.04% for *Zs/1*, see Table 1), we investigated the PVL of N2-*Irf3* for each of the eight possible genotype combination at the three QTL peak positions (Figure 24). Mice with the CC071/CC071 genotype for the three QTLs had similar PVL as CC071 mice, while mice with the CC071/B6 genotype for the three QTLs had similar PVL as B6-*Irf3* mice. Otherwise, on average, mice with increasing number of CC071/CC071 genotypes had increasing PVL levels. The major effect of *Zs/1* was visible, for instance, mice with the CC071/CC071 genotype for *Zs/1* but heterozygous for the two other QTLs had slightly higher PVLs on average than mice heterozygous for *Zs/1* but carrying the CC071/CC071 genotype for the two other QTLs (4<sup>th</sup> and 5<sup>th</sup> groups on Figure 24). This analysis allows to conclude that the three QTLs have additive effects on the PVL at 2 dpi of N2-*Irf3* mice. Moreover, as the whole range of PVL between the values of the parental strains is covered, we may have identified the major QTLs contributing to the difference of PVL between CC071 and B6-*Irf3* mice.



**Figure 21:** *Zs/1* and *Zs/2* on chromosome 12 associated with the PVL and BVL in the N2-*Irf3* progeny. (A) Zoom on the chromosome 12 QTLs, *Zs/1* and *Zs/2*, associated with the PVLs and BVL. Horizontal dashed lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 20. The vertical dotted lines represent the positions of the markers used to display allelic effects in B, C, D, and E. (B - E) Allelic effects assessed at the marker with the highest LOD score for the PVL at 6 dpi: SFT123018710, at 75.51 Mb on the PVL at 6 dpi (B), and at the marker with the highest LOD score for the PVL at 2 dpi on the PVL at 2 dpi (C), PVL decrease (D) and BVL (E). Results are represented as mean  $\pm$  sem. Groups were compared by t test (\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ ). Horizontal lines show the average PVL of CC071 (red), B6-*Irf3* (blue), and F1s (green).

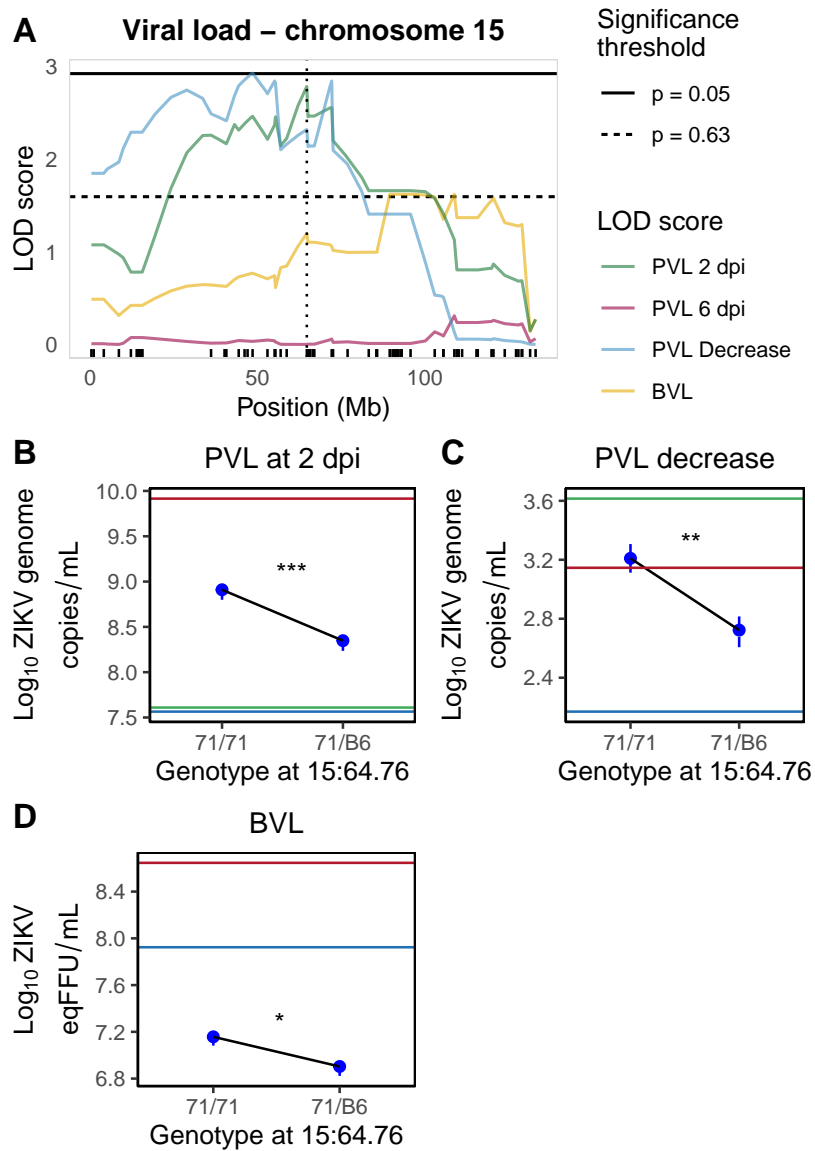


**Figure 22:** Chromosome 2 QTLs associated with the PVL in the N2-*Irf3* progeny.

(A) Zoom on the chromosome 2 QTLs associated with the viral loads at 2 and 6 dpi and the viral load decrease. Horizontal dashed lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 20. The vertical dotted line represent the position of the marker used to display allelic effects in B and C. (B - C) Allelic effects assessed at the marker with the highest LOD score for the PVL at 2 and 6 dpi: S2T027246659, at 180.80 Mb. Results are represented as mean  $\pm$  sem. Groups were compared by t test (\*\*  $p < 0.01$ ).

One significant QTL was associated with the PVL at 6 dpi, located on chromosome 12 (Figure 21A). The peak LOD score of this QTL named *Zs/2* was 3.071 and was reached at 75.51 Mb. This QTL explained 14.39% of the variance of the PVL at 6 dpi. Notably, the CC071/CC071 genotype was associated with viral loads higher by 0.47 log, which is about 1 sixth of the viral load difference between CC071 and F1 mice (Figure 21C). The credible interval of this QTL spanned between 31.81 Mb and 100.90 Mb, which did not overlap with *Zs/1*, controlling the PVL at 2 dpi (Table 1), thus suggesting that different genes on chromosome 12 control early and late PVL.

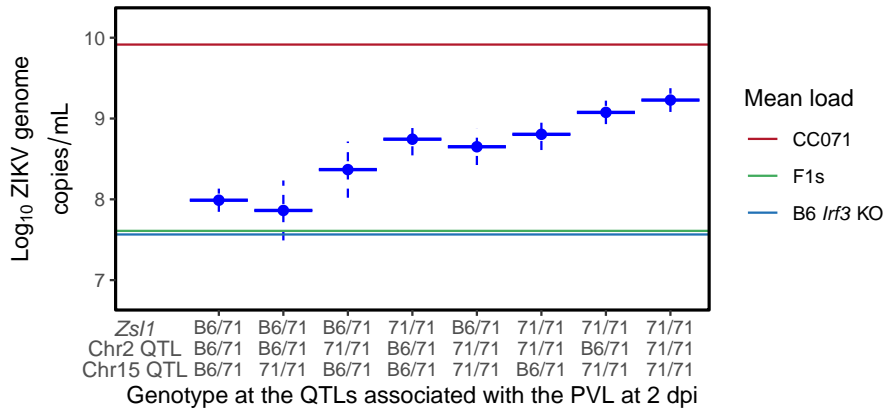
Two suggestive QTLs were associated with the PVL at 6 dpi. First, a QTL on chromosome 2 reached a LOD score of 1.857 at approximately 160 Mb (Figure 22A). This QTL colocalized with the chromosome 2 QTL found for the PVL at 2 dpi, and thus, was also located in the supposedly non-polymorphic region. The allelic effects for the closest marker showed that CC071/CC071 mice had higher PVL (Figure 22C). This effect was not significant, probably as this marker is too distant from the peak LOD score.



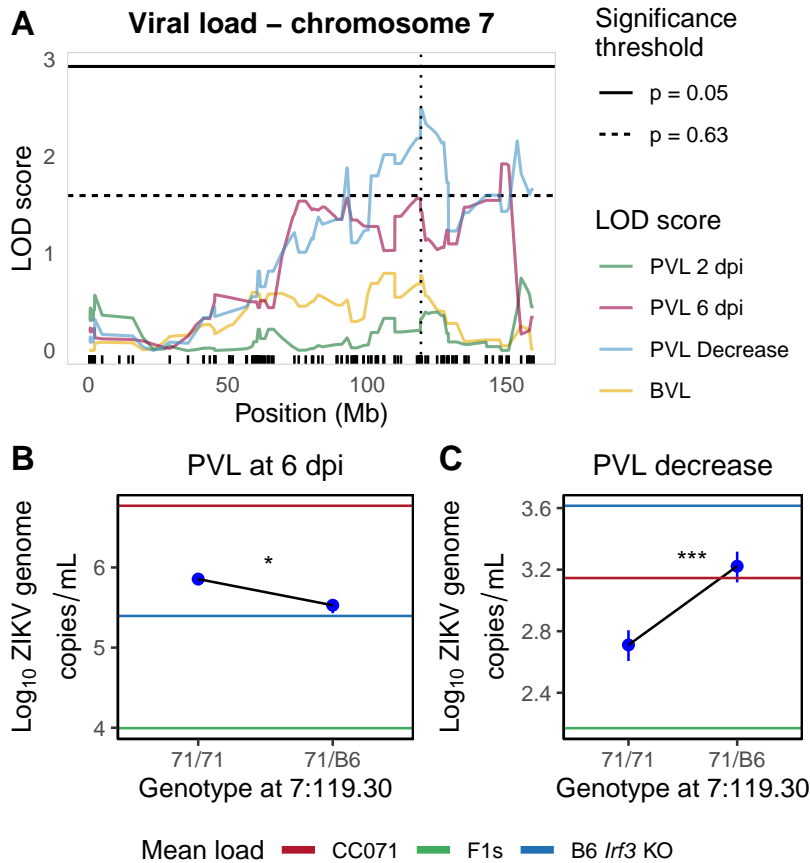
**Figure 23:** Chromosome 15 QTL QTLs associated with the PVL in the N2-*Irf3* progeny. Zoom on the chromosome 15 QTLs associated with the PVL at 2, PVL decrease and BVL. Horizontal dashed lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 20. The vertical dotted line represents the position of the marker used to display allelic effects in B and C. (B - C) Allelic effects assessed at the marker with the highest LOD score for the PVL at 2 dpi: gUNC25757731, at 64.76 Mb. Results are represented as mean  $\pm$  sem. Groups were compared by t test (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).

The second suggestive QTL was located on chromosome 7, where the LOD score reached 1.920 at 134.39 Mb (Figures 20 and 25A). The large credible interval of this QTL suggests, if this QTL is confirmed, that several genes in this region are involved in the PVL of N2-*Irf3* mice. Further analysis of a higher number of mice might confirm this QTL and increase resolution.

The single QTL analysis identified loci which effect can be detected independently. However, genetically controlled variation may result from a combination of alleles at two loci with non-additive interactions, which may be missed if each locus is analyzed in isolation. To identify such cases, two-dimensional genome scan was performed. This analysis explores a pairwise combination of marker and computes the additive and interactive effects. For the PVL at 2 dpi, additive associations were

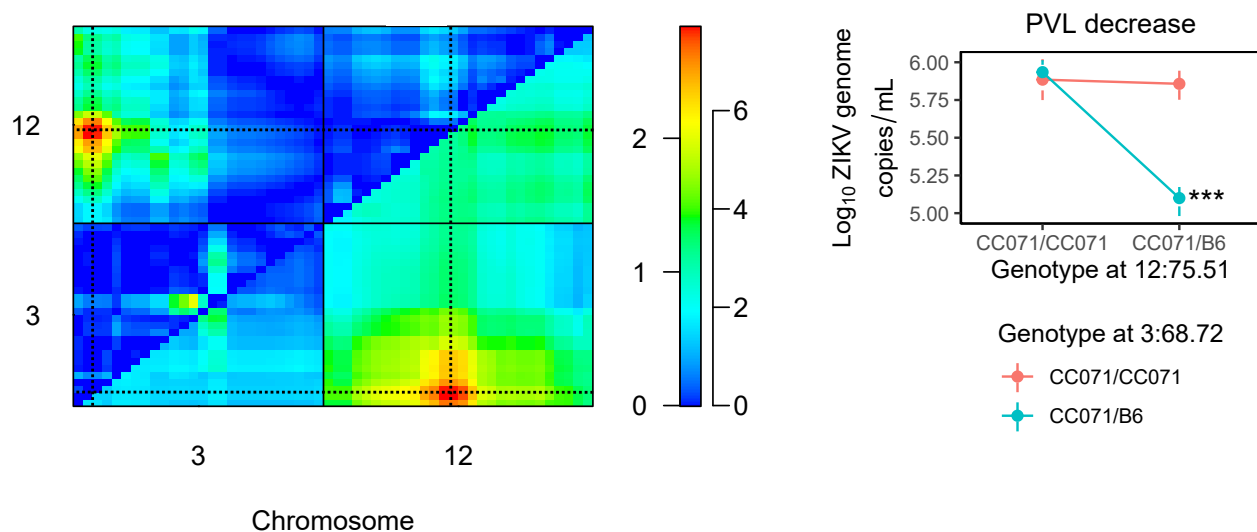


**Figure 24:** Additive effects of the three QTLs controlling the PVL at 2 dpi of N2-*Irf3* mice. Allelic effects assessed at gUNC21939504, S2T027246659, and gUNC25757731 (see Figures 21, 22, and 23). Groups were sorted by increasing proportion of CC071 alleles.



**Figure 25:** Chromosome 7 QTLs associated with the PVL in the N2-*Irf3* progeny. Zoom on the chromosome 7 QTLs associated with the viral loads at 2 and 6 dpi and the viral load decrease. Horizontal dashed lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 20. The vertical dotted line represents the position of the marker used to display allelic effects in B and C. (B - C) Allelic effects assessed at the marker with the highest LOD score for the PVL decrease: gJAX00156302, at 119.30 Mb, on the PVL at 6 dpi (B) and the PVL decrease (C). Results are represented as mean  $\pm$  sem. Groups were compared by t test (\*  $p < 0.05$ , \*\*\*  $p < 0.001$ ).





**Figure 26:** Two-dimensional QTL mapping for the PVL at 6 dpi the N2-*Irf3* progeny.

(A) Two-dimensional QTL mapping was performed on the PVL at 6 dpi. The lower right triangle represents results of the full QTL model (considering additive and interactive effects) and the upper left triangle represent results of the interactive QTL model. Each square represents the positions of two markers. The color of each square represents the LOD score. The vertical dotted lines represent the positions of the markers used to display allelic effects in B. (B) Allelic effects were assessed at the marker previously identified on chromosome 12: SFT123018710, at 75.51 Mb on chromosome 12, and the loci on chromosome 3 with the highest LOD score in the two-dimensional QTL mapping analysis: gUNCHS008973, at 68.72 Mb on chromosome 3, for PVL at 6 dpi. Results are represented as mean  $\pm$  sem. The effects of the genotypes at the two markers were assessed by two-way ANOVA. The effects of each marker and the interactive effect of the two markers were significant. Groups were compared in pairs by t test with Bonferroni correction for multiple testing. Double heterozygous significantly differed from the other groups (\*\*\*)  $p < 0.001$ ).

observed for *Zs1* and the suggestive QTL on chromosome 2, but no new association was detected.

For the PVL at 6 dpi, when looking at the full model (i.e., additive and interactive effects), I found a high LOD score of 7.70 for markers on chromosomes 3 and 12 (Figure 26A lower-right triangle). This position on chromosome 12 corresponds to *Zs2*. On chromosome 3, the peak LOD score was found at 68.72 Mb. No LOD score peak was observed in this region with a single QTL model (Figure 20). Notably, this position is the first polymorphic marker on chromosome 3 at 68.72 Mb, since CC071 inherited the B6 haplotype for the proximal region of the chromosome 3 (Figures S2A and C).

The interactive model showed that there is an epistatic interaction between these two markers (Figure 26A upper-left triangle). The effect plot revealed that mice having the CC071/CC071 genotype for at least one of the two loci had higher PVL than double heterozygous mice (Figure 26B). This observation is consistent with the results obtained on F1 mice, which are heterozygous for all the genome and showed lower viral loads than CC071 and B6-*Irf3* mice (Figure 18).

Single QTL mapping for the PVL decrease identified one significant QTL at the end of chromosome 12, colocalizing with *Zs1*, which suggests that these two traits are controlled by the same gene, which is expected from the calculation of the PVL decrease (Figure 21). A suggestive QTL was found on chromosome 15, colocalized with the suggestive QTL associated with the PVL at 2 dpi (Figure 23). For these two QTLs, the CC071/CC071 genotype was associated with higher PVL decrease. A second

suggestive QTL was found on chromosome 7 and colocalized with the suggestive QTL controlling the PVL at 6 dpi (Figure 25). Unlike the previous QTLs, the CC071/CC071 genotype was associated with lower PVL decrease.

Therefore, all the QTLs obtained for the PVLs were associated with two of the three PVL traits (2 dpi, 6 dpi and decrease). When the QTL was associated with the PVLs at 2 and 6 dpi, the CC071/CC071 genotype was associated with higher PVL at both time points (Figures 22). When the QTL was associated with the PVL at 2 dpi and the decrease, the CC071/CC071 genotype was associated with higher PVL at 2 dpi and with higher decrease (Figures 21 and 23). Lastly, when the QTL was associated with the PVL at 6 dpi and the decrease, the CC071/CC071 genotype was associated with higher PVL at 6 dpi and lower decrease (Figures 25). This suggests that different mechanisms of control of the viral load levels are involved for each of the three situations.

Single QTL mapping for the BVL identified one significant and one suggestive QTLs. The significant QTL was located on chromosome 12, with a maximum LOD score of 3.029, colocalizing with *Zs/1* (Figure 21). The Bayesian credible interval for this QTL ranged between 73.65 and 119.37 Mb, thus, it also colocalized with *Zs/2*. This result suggests that the two loci controlling the PVL at 2 and 6 dpi also control the BVL. Mice with the CC071/CC071 genotype had higher BVL than mice with the CC071/B6 genotype (Figure 21E).

The suggestive QTL was located on chromosome 15 (Figures 20 and 23A) with a LOD score of 1.621. The CC071/CC071 genotype was associated with higher BVL (Figure 23E). The peak LOD score did not colocalize with the peak of the QTL identified with the PVL at 2 dpi and the PVL decrease, suggesting that a locus is specifically controlling the BVL.

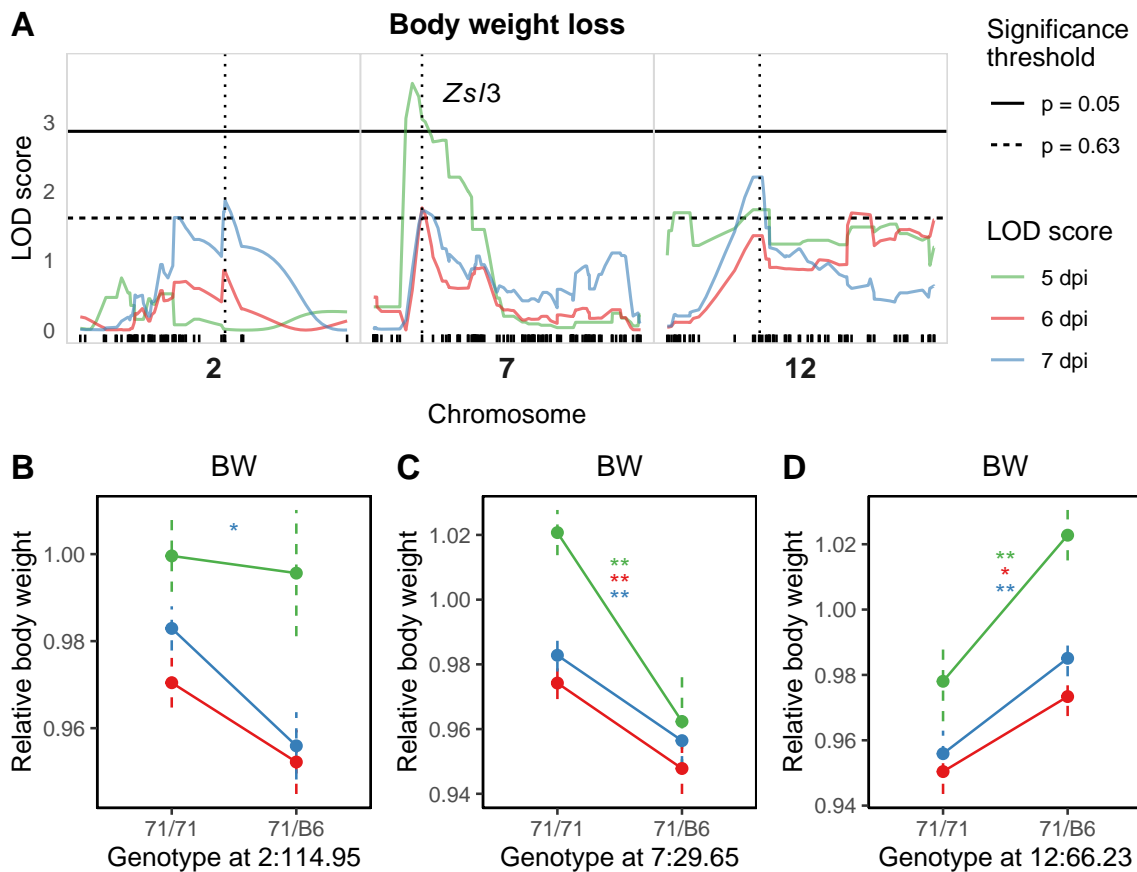
### **V.2.2.3 . QTL mapping identified one significant and two suggestive QTLs associated with the body weight loss of N2-*Irf3* mice**

The experimental models used in this cross did not allow to observe severe disease signs in infected mice. QTL mapping was still performed on the body weights and clinical scores of N2-*Irf3* mice to identify loci that may be involved in mild or moderate signs of illness.

I performed QTL mapping for the weights at 5, 6 and 7 dpi. *Zs/3* was found on chromosome 7 and was significantly associated with the body weight at 5 dpi, with a LOD score of 3.558 (Figure 27A). Interestingly, the CC071/CC071 genotype was associated with higher relative body weights at 5 dpi (Figure 27C). Suggestive associations were also found for the body weights at 6 and 7 dpi (Figure 27A).

The body weights also showed two suggestive associations on chromosomes 2 and 12 (Figure 27A). On chromosome 2, suggestive association was observed for the body weight only at 7 dpi and the CC071/CC071 genotype was associated with higher relative body weight at 7 dpi. At 6 dpi, the difference in body weight between these two groups of mice was visible but not significant, while the effect was not visible at 5 dpi (Figure 27B). Therefore, this QTL controlled only body weights at later time points.

Lastly, the body weights at 5, 6 and 7 dpi showed suggestive associations with loci on chromosome 12 (Figure 27A). The LOD scores were elevated across the chromosome and peaked at 66.23 Mb for the weights at 5 and 7 dpi, reaching scores of 1.736 and 2.207, respectively. For the body weight loss at 6 dpi, the maximum LOD score was 1.690 and was reached at 95.71 Mb. The shape of these



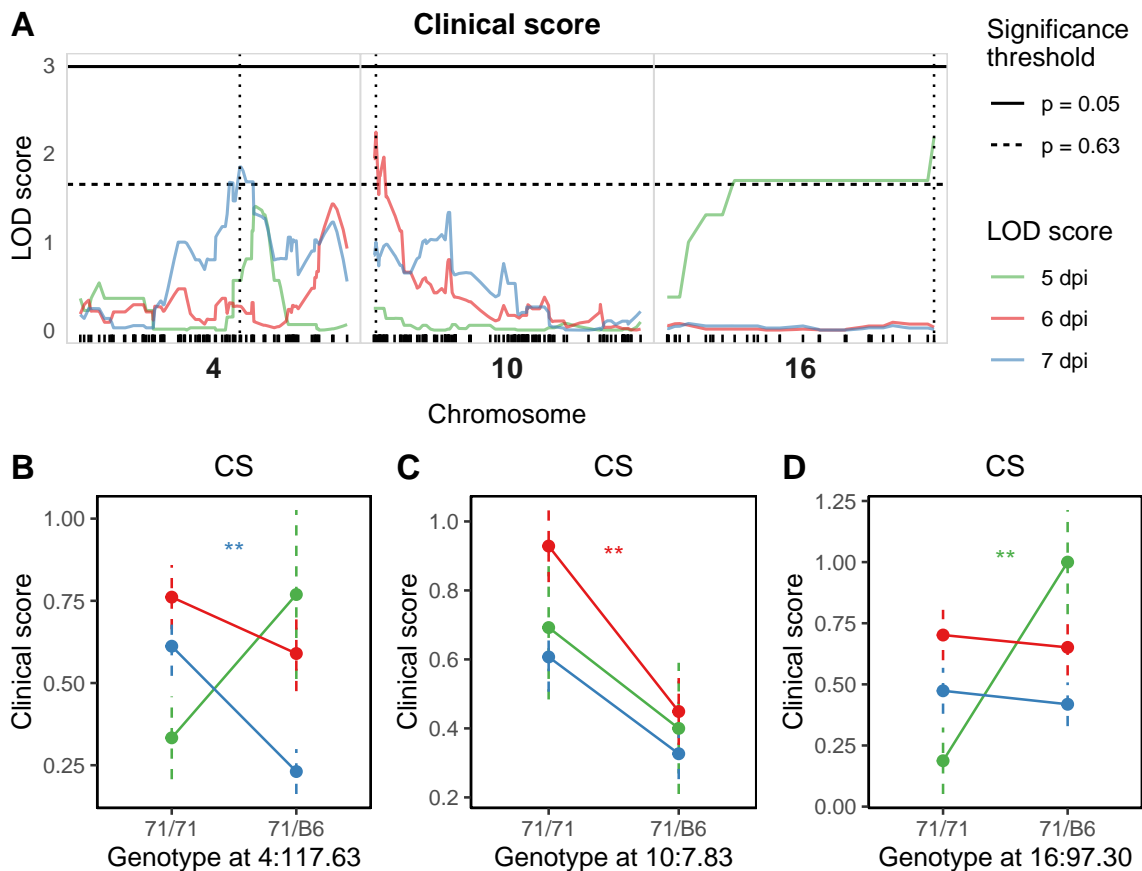
**Figure 27:** Single QTL mapping for the body weight loss in the *N2-lrf3* progeny showing *Zs/3* and two other suggestive QTLs.

(A) QTL mapping was performed on relative body weight at 5, 6 and 7 dpi. Only the chromosomes with suggestive or significant associations were represented. *Zs/3* on chromosome 7 was associated with the body weight at 5 dpi. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ). For better visibility, and as the significant thresholds are highly similar for the body weight loss at different time points ( $p = 0.05$ , 2.86 for 5 dpi, 3.00 for 6 dpi, 2.93 for 7 dpi,  $p = 0.63$ , 1.62 for 5 dpi, 1.64 for 6 dpi, 1.66 for 7 dpi), only the minimum threshold was displayed. The vertical dotted lines represent the positions of the markers used to display allelic effects in B, C and D. (B) Allelic effects assessed at the marker with the highest LOD score on chromosome 2 for the body weight at 7 dpi: S3C122647423, at 66.23 Mb. (C) Allelic effects assessed at the marker with the highest LOD score on chromosome 7 for the body weight at 5 dpi: gUNC12563949, at 29.65 Mb. Allelic effects assessed at the marker with the highest LOD score on chromosome 12 for the body weight at 7 dpi: S3C122647423, at 66.23 Mb. Results are represented as mean  $\pm$  sem. Groups were compared by t test (\*  $p < 0.05$ , \*\*  $p < 0.01$ ). BW = body weight.

QTLs, spreading on a long genomic region, suggests that several loci on chromosome 12 control the body weights of *N2-lrf3*. Contrary to the two other body weight QTLs, the CC071/CC071 genotype was associated with lower body weights than the CC071/B6 genotype (Figure 27D).

### V.2.2.4 . Three suggestive QTLs are associated with the clinical scores of N2-*Irf3* mice

Lastly, I analyzed the clinical scores of N2-*Irf3* mice. QTL mapping was performed on the clinical scores at 5, 6 and 7 dpi, and three suggestive QTLs were obtained (Figure 28A).



**Figure 28:** Single QTL mapping for the body weight loss in the N2-*Irf3* progeny.

(A) QTL mapping was performed on the clinical scores at 5, 6 and 7 dpi. Only the chromosomes with suggestive or significant associations were represented. Horizontal dashed lines indicate genome-wide significance thresholds ( $p = 0.05$  and  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ). For better visibility, only the minimum threshold was displayed ( $p = 0.05$ , 3.32 for 5 dpi, 2.98 for 6 dpi, 3.07 for 7 dpi,  $p = 0.63$ , 1.70 for 5 dpi, 1.71 for 6 dpi, 1.65 for 7 dpi). The vertical dotted lines represent the positions of the markers used to display allelic effects in B, C and D. (B) Allelic effects assessed at the marker with the highest LOD score on chromosome 4 for the clinical score at 7 dpi: B10040117767, at 117.63 Mb. (C) Allelic effects assessed at the marker with the highest LOD score on chromosome 10 for the clinical score at 6 dpi: mUNC17374593, at 78.30 Mb. (D) Allelic effects assessed at the marker with the highest LOD score on chromosome 16 for the clinical score at 5 dpi: gUNCHS043484, at 97.30 Mb. Groups were compared by Kruskal Wallis test. Results are represented as mean  $\pm$  sem. CS: clinical score.

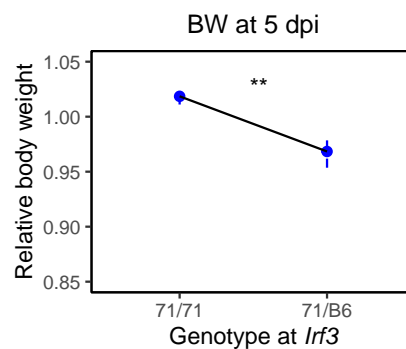
A suggestive association was observed for the clinical score at 7 dpi and loci on chromosome 4 (Figure 28A). The peak LOD score was 1.848 and was reached at 117.63 Mb. CC071/CC071 mice had higher clinical scores at 7 dpi than CC071/B6 mice (Figure 28B). Notably, at 5 dpi CC071/CC071 mice seemed to have lower clinical scores than CC071/B6 mice (Figures 28B), suggesting different genetic control of the clinical score at 5 and 7 dpi.

A suggestive QTL was obtained on chromosome 10 and was associated with the clinical score at 6 dpi (LOD = 2.24, Figure 28A). CC071/CC071 mice had significantly higher scores than CC071/B6 mice, and this effect was visible but not significant at 5 and 6 dpi (Figure 28C). Lastly, a suggestive QTL was found on chromosome 16 for the clinical score at 6 dpi (Figure 28A). However, the shape of this QTL

curve was unusual, being flat over most of the chromosome (Figure 28A), thus, this QTL requires further investigation.

#### V.2.2.5 . *Irf3* does not contribute to the phenotypic variability of the N2-*Irf3* progeny

In the N2-*Irf3* progeny, we did not expect an effect of *Irf3* on the phenotypes of the N2s as they all carry two non-functional *Irf3* alleles. Indeed, the genotype at the closest marker to *Irf3*, SAH071783312 at 44.23 Mb on chromosome 7 (*Irf3* is positioned at 44.65 Mb), had no significant effect on the PVLs, BVL, body weights at 6 and 7 dpi and clinical scores at 5, 6 and 7 dpi (Figure S3). By contrast, the body weight at 5 dpi was influenced by the genotype at this marker and CC071/CC071 mice showed an increased relative body weight compared with CC071/B6 mice ( $p < 0.01$ , Figure 29).



**Figure 29:** Allelic effects for *Irf3* on body weight loss at 5 dpi in the N2-*Irf3* progeny. Allelic effects were assessed at the closest marker to *Irf3*, SAH071783312, at 44.23 Mb on chromosome 7, for relative body weight loss at 5 dpi. Groups were compared by t test (\*\*  $p < 0.01$ ).

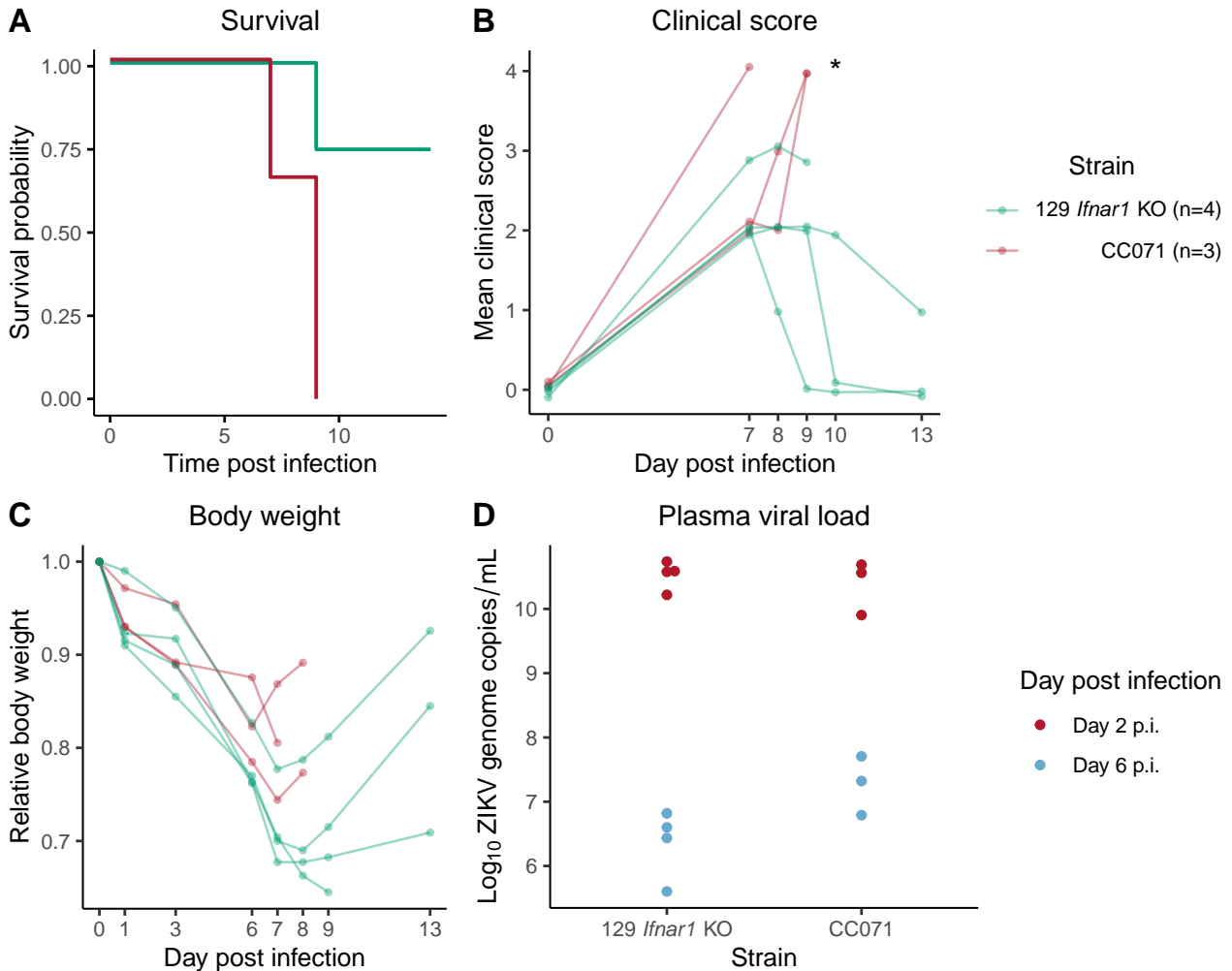
However, as discussed previously, we found a significant QTL, *Zs13*, close to *Irf3* for the body weight at 5 dpi (Figures 27A and C). The peak LOD score for this QTL was 3.558 and was observed for a marker located at 29.65 Mb, which is 15 Mb from *Irf3*. The LOD score at the marker closest to *Irf3* was only 2.20. Moreover, *Irf3* was not contained in the credible interval of this QTL (CI = 23.99 - 44.23 Mb, *Irf3* is located at 44.65 Mb), thereby *Irf3* is unlikely to influence the body weight loss at 5 dpi.

#### V.2.2.6 . Phenotypic analysis of CC071, 129-*Ifnar1*, and F2 mice

To identify QTLs involved in the severe disease signs of CC071 mice, we have used a more severe model by analyzing a cross between CC071 and *Ifnar1* deficient mice. Indeed, *Ifnar1* KO mice were previously shown to be more susceptible than antibody-treated WT mice (Lazear et al. 2016; Manet et al. 2020).

Two *Ifnar1* strain exist and could be crossed with CC071. C. Manet previously showed that 129-*Ifnar1* mice develop much less severe symptoms than B6-*Ifnar1* mice despite similar viral loads at 2 dpi. Besides, it was previously demonstrated that antibody treated WT B6 mice are less susceptible to ZIKV infection than antibody treated CC071 mice (Manet et al. 2020). Thus, an F2 cross was produced between CC071 and 129-*Ifnar1* mice.

Pretreated CC071 and untreated 129-*Ifnar1* mice were infected and followed for 13 days. All CC071 mice died in 9 dpi, while 129-*Ifnar1* mice showed a mortality rate of 25% (Figure 30A). CC071 and 129-*Ifnar1* mice had similar average clinical scores until 8 dpi, then all CC071 mice were dead by 9 dpi, while the scores of most 129-*Ifnar1* mice decreased between 8 and 13 dpi (Figure 30B). In this



**Figure 30:** mAb-treated CC071 and 129-*Irfar1* mice have different mortality rate but similar plasma viral loads. Six to 8-week-old CC071 (n=3) and 129-*Irfar1* (n=4) mice were infected i.p. with  $10^7$  FFUs of FG15 and monitored for 14 days. CC071 mice were treated with 2 mg of IMAR1-5A3 monoclonal antibody one day before the infection. (A) Kaplan-Meier survival curves showing 100% lethality in CC071 and 75% survival in 129-*Irfar1* mice. (B) Curves of individual clinical scores expressed with numerical values given as follows: 0, no symptoms; 1, ruffled fur; 2, emaciation, hunched posture, and/or hypoactivity; 3, hind limb weakness, prostration, and/or closed eyes; and 4, moribund or dead. (C) Curves of individual body weight loss, expressed as the percent of the starting weight (day of infection). Mice were weighed at day 1, 3, 6, 7, 8, 9, and 13 post infection. (D) PVL, quantified at days 2 and 6 post infection by RT-qPCR. (B, C) Groups were compared by linear mixed model analysis. (\*  $p < 0.05$ ) (C) Groups were compared by mixed two-way ANOVA testing the effect of the time post infection and the strain. There were no significant differences between strains.

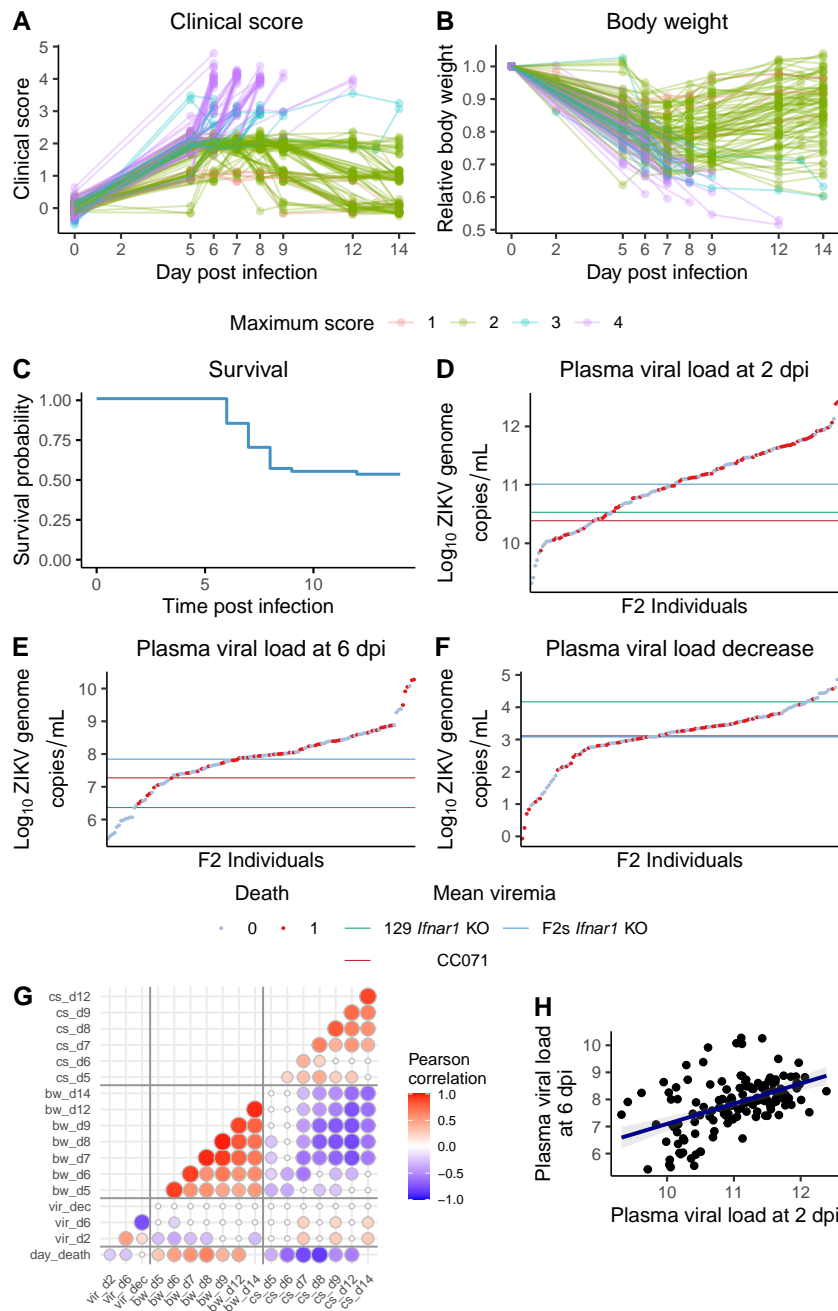
experiment, CC071 mice showed less interindividual variability of signs of illness than in the previous experiment (Figure 18), but only three mice were phenotyped. 129-*Ifnar1* mice lost more weight than CC071 mice between 6 and 8 dpi. However, 3 out of 4 regained weight between 9 and 13 dpi and survived. At 2 dpi, PVLs were similar between CC071 and 129-*Ifnar1* mice (Figure 30D) At 6 dpi, they were slightly higher in CC071 (7.27 log<sub>10</sub> genome copies per mL versus 6.63 in 129-*Ifnar1*), although not significantly due to the small sample size. In conclusion, even if the MAR1-5A3 treatment results in milder severe infection, antibody-treated CC071 mice were more severely affected than 129-*Ifnar1* mice, justifying the relevance of the cross between these two strains.

F1s were intercrossed to generate F2 mice, hereinafter designated as F2-*Ifnar1*, which were genotyped for the *Ifnar1* KO allele, and 166 *Ifnar1*<sup>-/-</sup> mice were infected and monitored for 14 days. Seventy-seven out of the 166 mice (46.4%) of the F2-*Ifnar1* progeny were euthanized as moribund or died between 6 and 14 dpi (Figure 31C). These most susceptible mice showed severe symptoms including early body weight loss, hypoactivity and limb paralysis (Figures 31A-B). Surviving mice displayed more variable signs of illness, from light symptoms (ruffled fur and hunched posture) to more severe symptoms such as tremors and partial paralysis. Of note, the clinical score scale differed from that of the N2-*Irf3* mice. Indeed, as N2-*Irf3* mice had less severe clinical manifestations than *Ifnar1* deficient mice, we extended the clinical score scale for N2-*Irf3* mice to better distinguish mild from moderate symptoms. Surviving mice lost weight between the day of infection and 7 dpi, and then regained weight between 7 and 14 dpi. However, some mice still showed low body weight at 14 dpi, with stable clinical scores between 9 and 12 dpi (Figures 31B-C). Sex impacted the clinical scores of mice, as males had on average higher clinical scores from 8 dpi, but not their body weights, similarly to what was observed in N2-*Irf3* mice (Figure S4). Males and females showed not significantly different mortality rates (50.6% in males, 41.6% in females), but females died earlier than males (Figure S4).

This variability of disease signs in a genetically variable population is an occasion to refine the humane endpoints used for *Ifnar1* deficient mice infected with ZIKV. The body weight of infected mice did not seem to be an appropriate limit as we observed large variability in the body weight curves of F2-*Ifnar1* mice, especially in mice that survived. Indeed, 54 F2-*Ifnar1* mice lost more than 20% of their original body weight in the course of the experiment but did not succumb, including 7 which reached back a weight superior to 90% of their original weight by 14 dpi. We found that the disease signs of the mice were better indicators of the probability of survival of these mice. The motility was a good predictor of their susceptibility to the infection. For some of these mice, this lack of activity was explained by severe paralysis of the limbs, which we considered as a sign of significant susceptibility. Thus, mice that showed hypoactivity were euthanized.

The PVL of F2-*Ifnar1* mice ranged over 3 logs at 2 dpi and over 5 logs at 6 dpi, and the PVL was slightly higher in mice that succumbed (t test, p = 0.0046 at 2 dpi, p = 0.0022 at 6 dpi). The average PVL in the F2-*Ifnar1* mice was higher than those of the parental strains, especially at 6 dpi, and the PVL decrease was lower (Figures 31D-F). High PVL in the F2-*Ifnar1* probably resulted from the synergy between the susceptibility alleles from CC071 and the *Ifnar1* deficiency, which leads to higher loads than after antibody treatment.

Correlations between the different phenotypes measured in F2-*Ifnar1* mice were calculated (Figures 31G). As observed with the N2-*Irf3* progeny, groups of similar traits (i.e., PVL, clinical scores



**Figure 31:** F2-*Ifnar1* mice show great phenotypic variability.

Six to 9-week-old F2 mice ( $n=166$ ) were infected i.p. with  $10^7$  FFUs of FG15 and monitored for 14 days. (A) Curves of clinical scores, expressed with numerical values as in Figure 30. (B) Body weight loss curves, expressed as the percent of the starting weight (day of infection). Mice were weighed at day 2, 5, 6, 7, 8, 9, 12 and 14 post infection. Curves were colored by the maximum clinical score of each mouse. (C) Survival curve of F2 mice. (D - F) Plasma viral loads, quantified at days 2 (D) and 6 (E) post infection by RT-qPCR. (F) PVL decrease was calculated as the difference between the load at 2 dpi and the load at 6 dpi. Horizontal lines show the average PVL of CC071 (red), 129-*Ifnar1* (green), and F2s (blue). Mice that survived the infection ("0") are represented in blue and mice that died due to the infection ("1") are represented in red. (G) Correlation matrix between of the phenotypes of the N2 mice. The strength of correlation is represented by both circle sizes and colors. Small white circles show nonsignificant correlations (H) Correlation between viral loads at 2 and 6 dpi (Pearson correlation coefficient:  $R = 0.5$ ,  $p = 4.1 \times 10^{-10}$ ).



and body weights) showed intra-group correlations. The values of PVL at 2 and 6 dpi were more significantly correlated in the F2-*Ifnar1* progeny (Pearson correlation coefficient,  $R = 0.5$ ,  $p = 4.1 \times 10^{-10}$ , Figure 31H), but mice with low PVL at 2 dpi showed variable loads at 6 dpi ranging over 3 logs. The PVL decrease was more correlated with the PVL at 6 dpi than with the PVL at 2 dpi, which is consistent with the higher variability of this trait at 6 dpi.

Early clinical score values were not significantly correlated with late values of clinical scores and body weights. This could be explained by the fact that mice which had high scores at early time points died before the end of the experiment. As expected, the day of death was inversely correlated with the clinical scores, and mice that succumbed lose more weight (Figure 31G).

### V.2.2.7 . QTL mapping identifies one QTL controlling the plasma viral loads of F2-*Ifnar1* mice

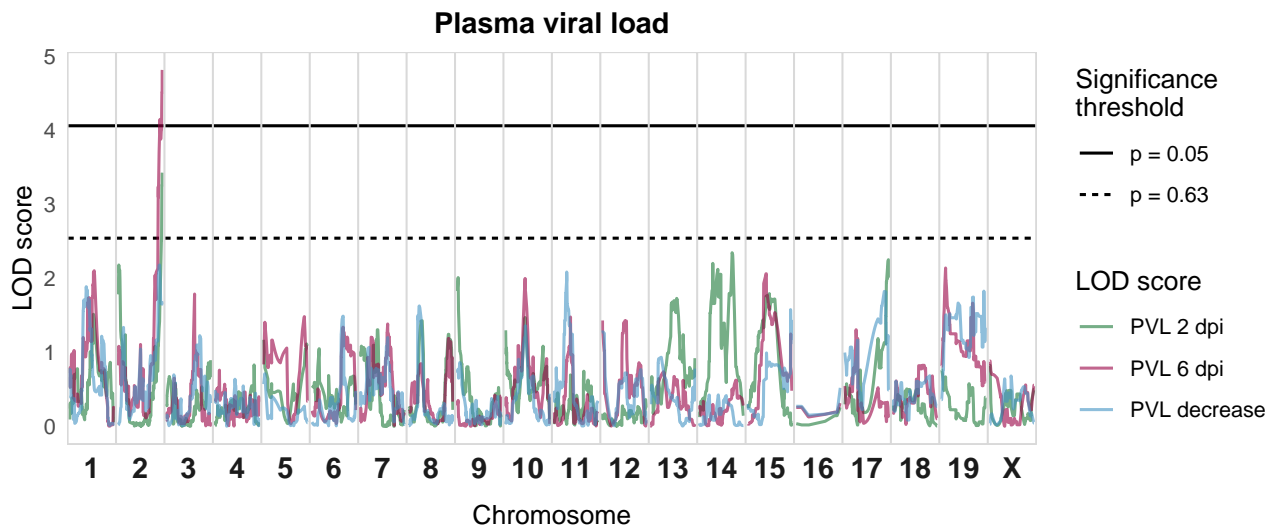
The F2-*Ifnar1* progeny was genotyped with the MiniMUGA array containing 11,125 markers spread across the genome. After curation with the stuart package (Bourdon and Montagutelli 2022), 2692 markers were retained for QTL mapping (Figure S2B). The QTLs detected in this cross are presented in Table 2.

**Table 2:** Associations between QTLs and phenotypes in the F2-*Ifnar1* progeny.

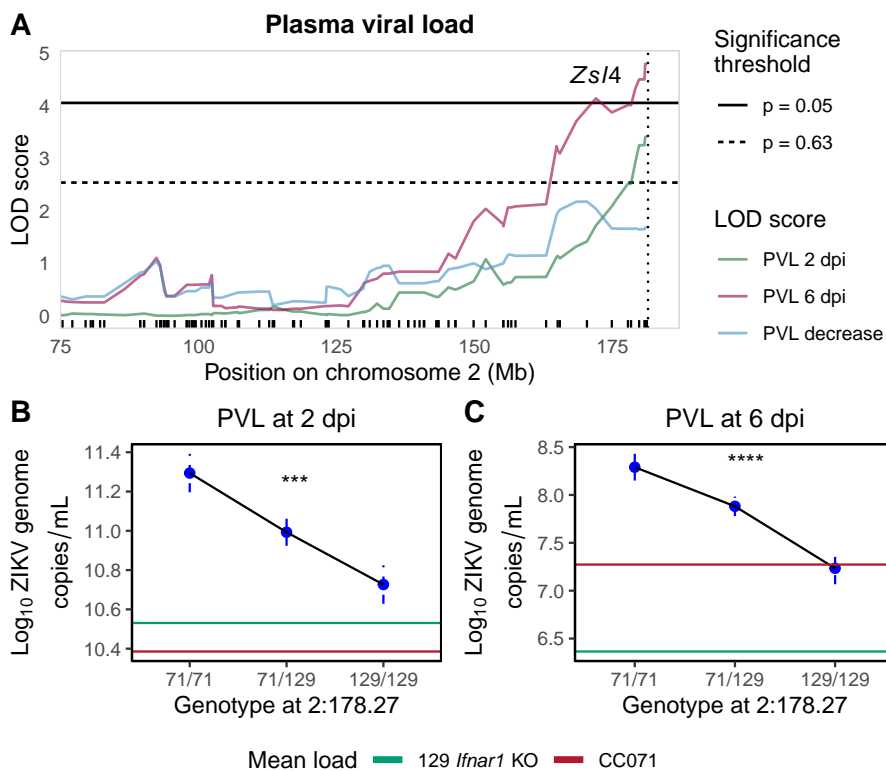
QTL	Phenotype	Chr	Position	LOD	p-value	CI (Mb)	PVE (%)	Effect
<i>Zsl4</i>	PVL (6 dpi)	2	178.27	4.237	0.033	165.6 - 181.4	13.10	71 > 129
<i>Zsl4</i>	PVL (2 dpi)	2	181.43	3.424	0.188	4.0 - 181.5	9.3	71 > 129
<i>Zsl5</i>	Mortality	7	69.76	4.325	0.035	29.08 - 115.1	11.4	71 > 129
<i>Zsl5</i>	CS (7 dpi)	7	69.76	5.256	0.007	49.1 - 83.1	19.44	71 > 129
<i>Zsl5</i>	CS (12 dpi)	7	69.76	4.217	0.041	43.5 - 71.4	22.04	71 > 129
<i>Zsl6</i>	Day of death	7	116.64	4.440	0.033	79.4 - 125.7	11.72	71 > 129
<i>Zsl6</i>	CS (6 dpi)	7	114.17	4.989	0.012	33.0 - 122.3	16.09	71 > 129
-	BW (5 dpi)	2	150.74	2.814	0.396	3.9 - 181.4	9.1	71 < 129
-	BW (5 dpi)	3	108.80	2.631	0.538	52.8 - 112.2	8.5	71 < 129
-	BW (6 dpi)	3	108.80	3.809	0.081	68.2 - 126.6	13.7	71 < 129
-	BW (7 dpi)	3	108.80	2.672	0.547	52.9 - 122.2	12.0	71 < 129
-	BW (7 dpi)	7	64.76	2.959	0.348	49.1 - 144.6	13.2	71 < 129
-	BW (5 dpi)	7	128.13	2.846	0.379	117.2 - 139.4	9.2	71/129 > 71/71 = 129/129
-	BW (14 dpi)	11	17.13	2.892	0.379	6.3 - 121.6	14.2	71/129 > 71/71 = 129/129
-	BW (12 dpi)	14	124.84	2.850	0.402	3.3 - 124.9	15.9	71/129 < 71/71 = 129/129
-	BW (7 dpi)	18	76.33	2.591	0.605	67.3 - 89.6	11.7	71/129 < 71/71 = 129/129
-	CS (7 dpi)	1	143.44	2.946	0.380	88.6 - 189.7	11.4	71 > 129
-	CS (7 dpi)	2	170.01	2.726	0.503	4.0 - 180.1	10.6	71/129 > 71/71 = 129/129
-	CS (6 dpi)	10	34.97	3.497	0.161	12.8 - 46.8	11.6	71 > 129
-	CS (8 dpi)	10	117.24	2.904	0.359	5.3 - 128.0	9.7	71 > 129
-	CS (7 dpi)	12	30.36	3.672	0.090	12.1 - 35.4	14.0	71 > 129
-	CS (12 dpi)	17	93.47	3.712	0.094	53.2 - 93.5	20.3	71 < 129
-	CS (14 dpi)	17	93.47	3.192	0.243	52.5 - 93.5	16.6	71 < 129
-	CS (12 dpi)	18	80.47	3.685	0.099	75.6 - 89.6	20.2	71/129 > 71/71 = 129/129
-	CS (14 dpi)	18	80.47	3.568	0.150	76.3 - 84.7	18.4	71/129 > 71/71 = 129/129

Survival and day of death were studied together as a two-part model (here referred as Susceptibility phenotype). CI: credible interval, PVE: percentage of variance explained. BW: body weight. CS: clinical score.

This cross was design to identify QTLs associated with the disease signs of infected mice, though, as PVL was measured on F2-*Ifnar1* mice, QTL mapping was also performed on this trait to investigate QTL that may control PVL using an *Ifnar1* deficient model.



**Figure 32:** QTL mapping for the PVL in the F2-*Ifnar1* progeny. QTL mapping was performed on the plasma viral load at 2 and 6 dpi and on the PVL decrease. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing ( $n=1000$ ). For better visibility, and as the significant thresholds were very close for the three traits ( $p = 0.05$ , 4.01 for the PVL at 2 dpi, 4.14 for 6 dpi, and 4.08 for the decrease;  $p = 0.63$ , 2.58 for the PVL at 2 dpi, 2.54 for 6 dpi, and 2.54 for the decrease), only the lower value was displayed.



**Figure 33:** *Zs14* on chromosome 2 associated with the PVL in the F2-*Ifnar1* progeny. (A) Zoom on the chromosome 2 QTL associated with the PVL at 6 dpi. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$  and  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 32. The vertical dotted line represents the positions of the marker used to display allelic effects in B and C. (B - D) Allelic effects assessed at the marker with the highest LOD score for the PVL at 6 dpi: mUNC4608754, at 178.27 Mb, on the PVL at 2 dpi (B) and 6 dpi. Groups were compared by ANOVA (\*\*\*)  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ). Horizontal lines show the average PVL of CC071 (red) and 129-*Ifnar1* (green).

QTL mapping analysis of the PVL at 2 dpi did not identify a significant QTL (Figure 32). However, a suggestive QTL was found on the distal region of chromosome 2. Interestingly, the PVL at 6 dpi was associated with a significant QTL in the same region, which was named *Zs/4* (Figure 33). The CC071/CC071 genotype was associated with higher PVL at 6 dpi and the CC071 allele was semidominant as CC071/129 had intermediate PVL (Figure 33C). It is possible that the suggestive QTL found in the same region of chromosome 2 for the PVL at 2 dpi is just the consequence of the correlation between the PVLs at 2 and 6 dpi (Figure 31H).

#### **V.2.2.8 . Two significant QTLs are associated with the survival of F2-*Ifnar1* mice**

QTL mapping was performed on the survival and day of death using a two-part model (Broman 2003, Figure 34). This model evaluates separately and in combination the probability of survival and the average day of death for mice which succumbed, and gives 3 LOD score values for each position tested. These results allow the identification of loci that control either or both phenotypes. The combined trait is referred to as "susceptibility" hereinafter.

These traits showed significant associations on chromosome 7 (Figure 35A). Analysis of the three LOD score curves revealed two QTLs. First, the association with the survival of F2-*Ifnar1* mice peaked at 69.76 Mb. This QTL, named *Zs/5*, reached a LOD score of 5.06. Mice with CC071/CC071, CC071/129 and 129/129 genotype at this locus had a mortality rate of 76.7%, 35.4% and 33.3%, respectively. Thus, the CC071 allele for this QTL was recessive (Figure 35B). The maximum LOD score for the day of death was reached more distally in chromosome 7, at 117.21 Mb. This QTL, named *Zs/6*, had a LOD score of 4.584. CC071/CC071 and CC071/129 mice had an earlier time of death than 129/129 mice, thus, the CC071 allele was dominant (Figure 35C).

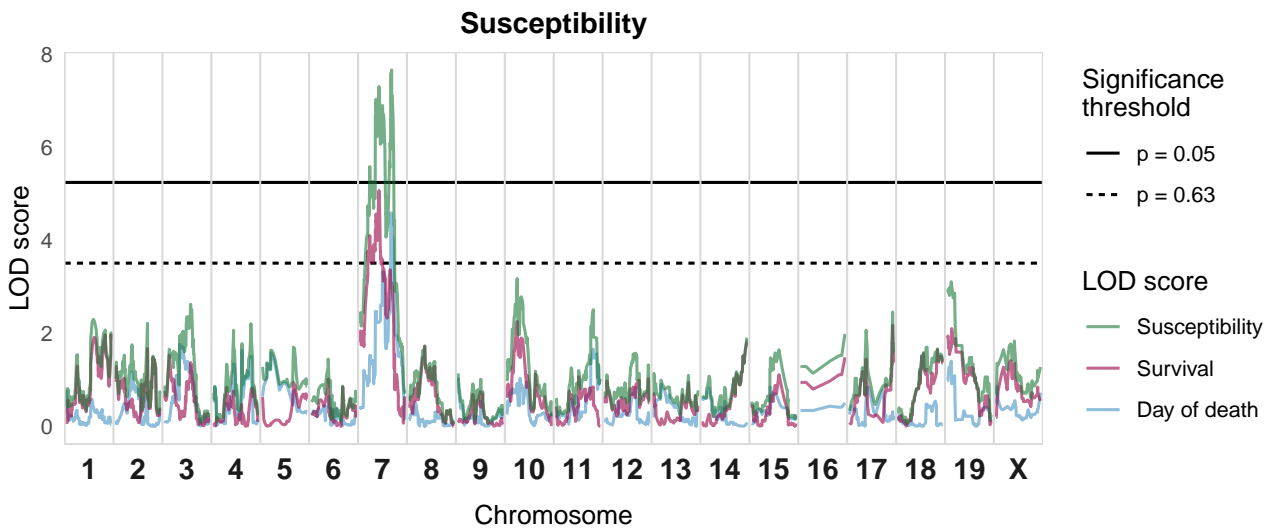
As a consequence, the "susceptibility" trait, which combines the survival and the day of death, was associated with a long interval spanning most of chromosome 7 with a complex curve reflecting the two separate QTLs (Figure 35A).

#### **V.2.2.9 . QTL mapping identifies seven loci suggestively associated with the body weight loss of F2-*Ifnar1* mice**

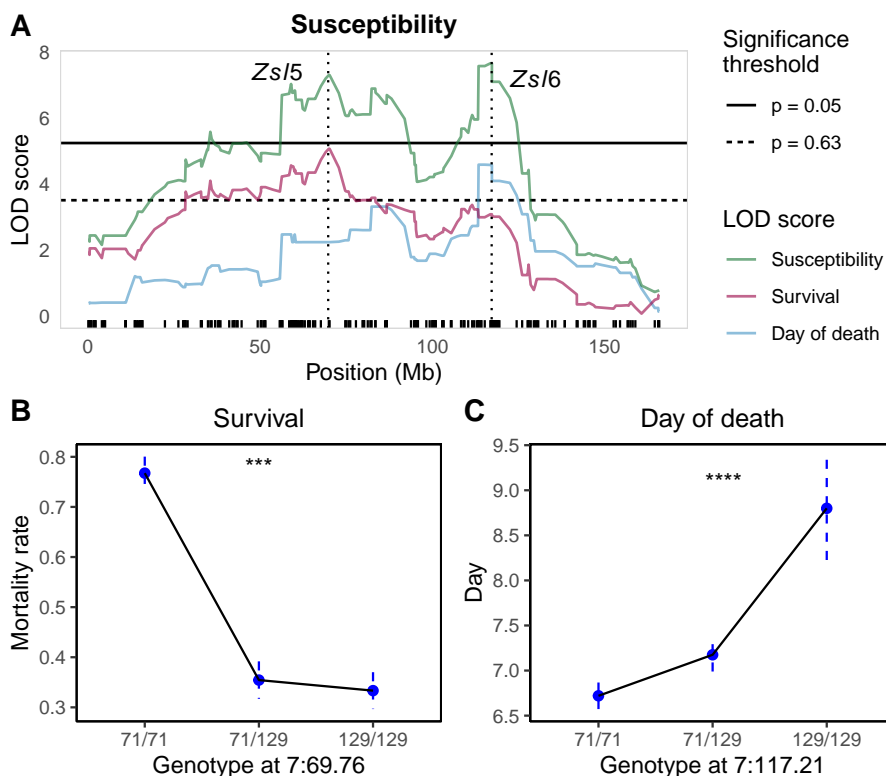
QTL mapping was performed on the relative body weights of F2-*Ifnar1* mice at 5, 6, 7, 8, 9, 12 and 14 dpi and identified seven suggestive QTLs on 6 chromosomes (Figure 36A). As these QTLs did not reach the significance threshold, they were not given a *Zs/* number. Some of these QTLs were associated with body weights at early time points, between 5 and 8 dpi, such as the QTLs on chromosomes 2, 3, 7, and 18, while others were associated with the body weights at later time points, between 12 and 14 dpi (Figure 36B), suggesting that different loci are involved in the control of the body weight during early and late infection phase.

Some of these QTLs showed suggestive associations with the weights at several successive days. For instance, the QTL on chromosome 3 had high LOD scores for the body weights at 5, 6 and 7 dpi (in red in Figure 36B). For this QTL, the CC071 allele was associated with lower body weight (Figure S5B).

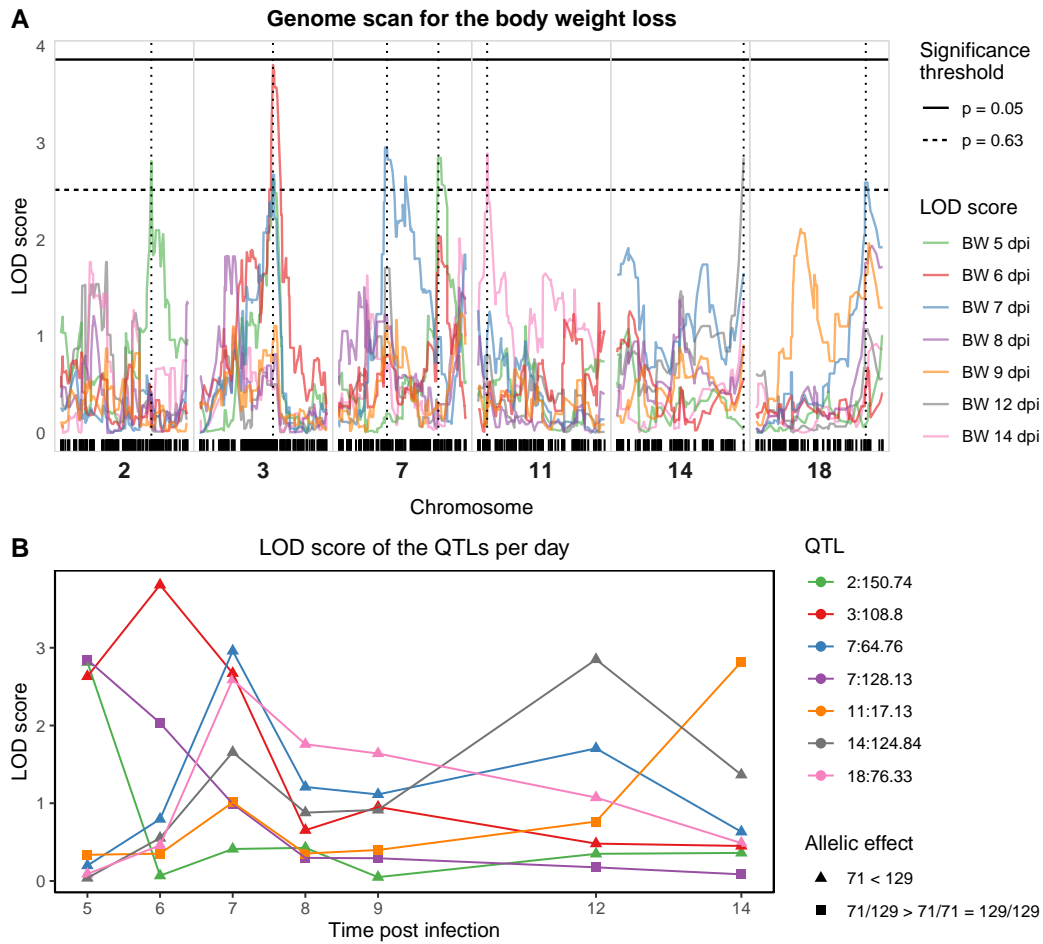
By contrast, the QTL on chromosome 2 was only associated with the body weights at 5 dpi, and the allelic effects revealed phenotypic differences between mice carrying different genotypes at this locus only at this time point (Figures 36 and S5A). It is unlikely that a locus impacts the body weight of



**Figure 34:** QTL mapping for the susceptibility of the F2-*Ifnar1* progeny. QTL mapping was performed on survival load and day of death studied with a two-part model. The combined trait for the analysis of the combination of the survival and the day of death is referred as "susceptibility". The horizontal lines indicate genome-wide significance thresholds determined by permutation testing (n=1000). For better visibility, only the thresholds for the susceptibility are displayed (p = 0.05, 5.24 for the susceptibility, 3.92 for the survival, and 3.99 for the day of death; p = 0.63, 3.51 for the susceptibility, 2.52 for the survival, and 2.48 for the day of death).



**Figure 35:** *Zs/5* and *Zs/6* on chromosome 7 associated with the survival and day of death of F2-*Ifnar1* mice. (A) Zoom on the chromosome 7 QTLs associated with the survival and day of death. Horizontal lines indicate genome-wide significance thresholds (p = 0.05 and p = 0.63) determined by permutation testing (n=1000), as in Figure 34. The vertical dotted line represents the positions of the markers used to display allelic effects in B and C. (B) Allelic effects assessed at the marker with the highest LOD score for the survival: gUNC13076624, at 69.76 Mb. Results are represented as mean  $\pm$  sd. Groups were compared by  $\chi^2$ . (C) Allelic effects assessed at the marker with the highest LOD score for the survival: S3J074704299, at 117.21 Mb. Results are represented as mean  $\pm$  sem. Groups were compared by ANOVA (\*\*\* p < 0.001, \*\*\*\* p < 0.0001).



**Figure 36:** QTL mapping for the body weight loss in the F2-*Ifnar1* progeny.

(A) QTL mapping was performed on relative body weight at 5, 6, 7, 8, 9, 12 and 14 dpi. Only the chromosomes with suggestive or significant associations were represented. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing ( $n=1000$ ). For better visibility, and as the significant thresholds are highly similar for the body weight loss at different time points ( $p = 0.05$ , 3.86 for 5 dpi, 4.03 for 6 dpi, 3.99 for 7 dpi, 3.99 at 8 dpi, 4.07 at 9 dpi, 4.12 at 12 dpi, 4.07 at 14 dpi,  $p = 0.63$ , 2.52 for 5 dpi, 2.60 for 6 dpi, 2.55 for 7 dpi, 2.56 at 8 dpi, 2.58 at 9 dpi, 2.53 at 12 dpi, 2.52 at 14 dpi), only the minimum threshold was displayed. Vertical dotted lines represent the positions of the markers with the highest LOD scores used to display the LOD score per day in (B) and the allelic effects in Figure S5. (B) LOD score for each QTL identified at each time post infection tested. BW = relative body weight.

F2 mice only at one precise day post infection, thus, this QTL will not be favored for future candidate gene search.

The suggestive QTL on chromosome 11 was only associated with the body weight at 14 dpi. The effect plot for this QTL revealed unusual allelic effects. Indeed, CC071/129 mice showed a higher body weight than CC071/CC071 and 129/129 mice (Figure S5E). Similar effects were obtained for the QTL on chromosome 14 (Figure S5F). These allelic effects reveal a more complex mechanism than a simple loss-of-function or gain-of-function variant which necessitates to be further investigated.

#### **V.2.2.10 . Two significant and seven suggestive QTLs control the clinical scores of F2-*Ifnar1* mice**

QTL mapping was performed on the clinical scores of F2-*Ifnar1* mice at 5, 6, 7, 8, 9, 12 and 14 dpi (Figure 37). Two significant QTLs were detected on chromosome 7 and colocalized with *Zs15* and *Zs16*, that were associated with the survival and day of death, respectively (Figure 35). As the clinical scores at 6 and 7 dpi are highly correlated with the day of death and the survival of mice, I hypothesized that *Zs15* and *Zs16* are also associated with the clinical scores of F2-*Ifnar1* mice.

*Zs15* was associated with the clinical scores at 7 and 12, with a LOD score of 5.256 and 4.217, respectively, while *Zs16* was associated with the clinical score at 6 dpi. For the two QTLs, the CC071 allele was associated with higher scores, consistent with the allelic effects observed for the survival and day of death (Figure S6). Moreover, the CC071 allele for *Zs15* was recessive for the clinical score, as for the survival (Figure 35). For *Zs16*, the CC071 allele was semidominant, while it seemed rather dominant for the day of death (Figure 35).

In addition to these two QTLs, seven suggestive QTLs were obtained for the clinical scores. Most of them were associated with the clinical scores at early time points, between 6 and 8 dpi, but *Zs15* and the QTLs on chromosome 17 and 18 were associated with the scores at 12 and 14 dpi, suggesting that they control the scores at late time points post infection (Figure 37B).

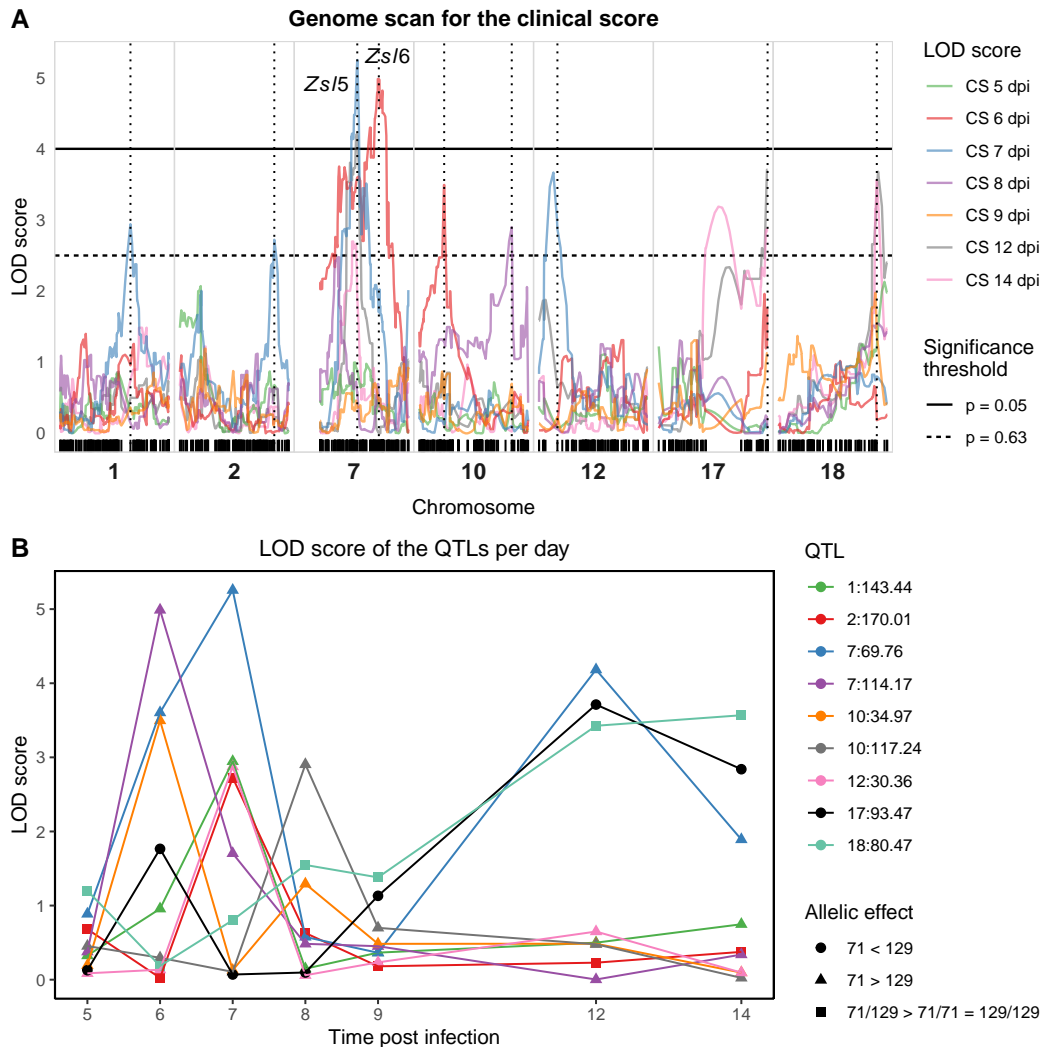
For instance, a QTL on chromosome 17 was suggestively associated with the clinical scores at 12 and 14 dpi. The effect plot for this locus revealed that the CC071 allele was associated with lower clinical scores at 9, 12 and 14 dpi. Therefore, this locus may control the clinical scores in the late phase of infection (Figure S6H).

The QTLs on chromosome 1, 2, 10 and 12 were only associated with the clinical score at one time point post infection (Figures 37A and S6). As for the body weight, it is unlikely that a locus is involved in the clinical score of F2 mice only at a specific time points.

#### **V.2.2.11 . Could *Irf3* contribute to the QTLs identified in the F2-*Ifnar1* cross?**

As all the F2-*Ifnar1* mice are homozygous for the *Ifnar1* KO, this cross only allows to identify a role of *Irf3* independent of IFN-I production (i.e., direct transcription of ISGs, transcription of IFN $\lambda$ , RIPA and RIKA). Therefore, we investigated the influence of the genotype of F2-*Ifnar1* mice at the closest marker to the *Irf3* locus, SX1071791093 (at 44.42 Mb on chromosome 7, *Irf3* is located at 44.65 Mb) on the viral loads and illness of F2-*Ifnar1* mice.

Genotype at the SX1071791093 marker did not influence the PVLs at 2 and 6 dpi, PVL decrease, day of death, body weights at 5, 6, 8, 9, 12 and 14 dpi, and clinical scores at 5, 8, 9 and 14 dpi (Figure S7). By contrast, the survival, body weight at 7 dpi, and clinical scores at 6, 7, and 12 dpi were



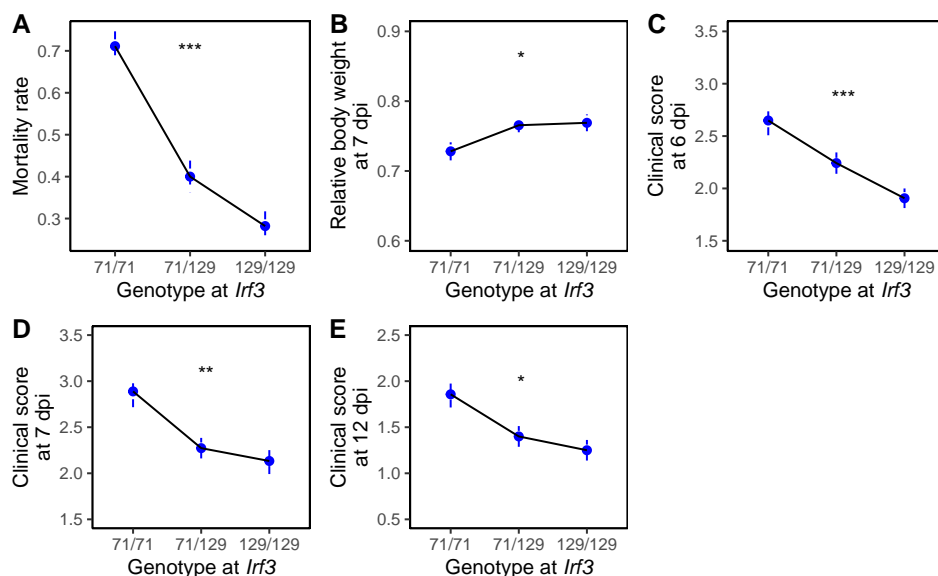
**Figure 37:** QTL mapping for the clinical scores in the F2-*Ifnar1* progeny displaying *Zs15*, *Zs16* and seven other suggestive QTLs. (A) QTL mapping was performed on the clinical scores at 5, 6, 7, 8, 9, 12 and 14 dpi. Only the chromosomes with suggestive or significant associations were represented. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing (n=1000). For better visibility, and as the significant thresholds are highly similar for the clinical scores at different time points (p = 0.05, 4.36 for 5 dpi, 4.23 for 6 dpi, 3.96 for 7 dpi, 3.91 at 8 dpi, 3.89 at 9 dpi, 4.10 at 12 dpi, 4.24 at 14 dpi, p = 0.63, 2.56 for 5 dpi, 2.54 for 6 dpi, 2.53 for 7 dpi, 2.51 at 8 dpi, 2.51 at 9 dpi, 2.55 at 12 dpi, 2.53 at 14 dpi), only the minimum threshold was displayed. Vertical dotted lines represent the positions of the markers with the highest LOD scores used to display the LOD score per day in (B) and the allelic effects in Figure S6. (B) LOD score for each QTL identified at each time post infection tested. CS = clinical score.

influenced by the genotype at the SX1071791093 marker, with CC071/CC071 mice showing more susceptible phenotypes than the other genotypes. These effects can be explained by *Irf3* or another gene genetically linked to *Irf3*. Therefore, we searched for LOD score peaks in the vicinity of *Irf3* for these traits in the F2-*Ifnar1* progeny.

As presented in Figure 36, the body weight at 7 dpi showed a suggestive QTL on chromosome 7 with a maximum LOD score of 2.96 at 64.76 Mb. The LOD score at the marker closest to *Irf3* was only 1.23 (Table 3), and *Irf3* was not contained in the Bayesian credible interval, which spanned between 49.1 and 86.8 Mb, while *Irf3* is located at 44.65 Mb.

Besides, *Zsl6* was found on chromosome 7 and was significantly associated with the survival and the clinical scores at 7 and 12 dpi (Figures 35 and 37). The LOD score curves profiles were investigated to determine whether these traits are controlled by *Irf3*. The LOD score highly decreased between the peak and the *Irf3* locus for the clinical scores at 7 dpi (5.256 at the peak versus 2.88 at the *Irf3* locus), and decreased by 1 or more for the other phenotypes (Table 3).

In conclusion, several QTLs were obtained in the vicinity of *Irf3* in the F2-*Ifnar1* cross. The allelic effect could be consistent with an effect of the *Irf3* mutation, as the CC071/CC071 genotype was associated with increased susceptibility. By contrast, the QTL peaks were localized far from the *Irf3* locus, and the LOD score at the *Irf3* locus was decreased compared to the LOD scores at the peaks of the QTLs. Thus, *Irf3* was not located in the credible intervals of most of these QTLs. Moreover, the *Ifnar1* deficiency likely masks a mild effect of the *Irf3* loss-of-function. Therefore, we cannot conclude as for the influence of *Irf3* on the disease signs of F2-*Ifnar1* mice, but the results suggest that other genes are better contributors of their severity.



**Figure 38:** Allelic effects for *Irf3* on the different viral load and clinical manifestations of F2-*Ifnar1* mice. Allelic effects were assessed at the closest marker to *Irf3*: SX1071791093, at 44.42 Mb on chromosome 7. (A) Groups were compared by  $\chi^2$ . (B) Groups were compared by ANOVA. (C - E) Groups were compared by Kruskal-Wallis test. (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).



**Table 3:** QTLs in the vicinity of *Irf3* in the F2-*Ifnar1* progeny.

Phenotype	Peak (Mb)	LOD score	Distance from <i>Irf3</i> (Mb)	$\Delta$ LOD score
Survival	69.76	4.325	25.11	1.170
Clinical score (6 dpi)	114.17	4.440	69.52	0.999
Clinical score (7 dpi)	69.76	5.256	25.11	2.372
Clinical score (12 dpi)	69.76	4.217	25.11	0.682

Mb: megabases.  $\Delta$  LOD score: difference of LOD score between the QTL peak and the *Irf3* locus

### V.2.3 . Discussion

The CC071 strain is becoming increasingly interesting to model viral disease in mice with its recently reported enhanced susceptibility to several flaviviruses, including ZIKV, DENV, WNV, Powassan virus (POWV), but also viruses from other families such as NrHV and RVFV (Manet et al. 2020; Cartwright et al. 2022; Brown et al. 2023; Jasperse et al. 2023). We previously demonstrated that this strain carries a loss-of-function mutation in *Irf3*, but that this mutation is not sufficient to induce elevated viral load and disease signs in ZIKV-infected CC071 mice (Bourdon et al. 2023).

This observation led us to address two questions. First, as the *Irf3* deficiency explains a high viral replication in CC071 MEFs, but not the elevated PVL in CC071 mice after *in vivo* infection, the other genes involved in this trait were investigated. For this purpose, a cohort of (CC071  $\times$  B6-*Irf3*)  $\times$  CC071 mice was infected after treatment with MAR1-5A3 and analyzed.

Second, the genetic determinants of the illness and mortality of CC071 mice after ZIKV infection were assessed. Thus, a more severe model of ZIKV infection was used as we investigated *Ifnar1* deficient mice. CC071 mice were crossed with 129-*Ifnar1* mice to produce an F2 cohort, and only *Ifnar1*<sup>-/-</sup> mice were infected.

This difference in the severity of ZIKV infection of antibody treated and *Ifnar1* KO mice was previously reported by us and others. Indeed, Lazear and colleagues reported that antibody-treated B6 mice survive and do not lose weight after ZIKV infection, while *Ifnar1* deficient B6 mice had a 100% mortality rate (Lazear et al. 2016). Furthermore, we previously showed that B6-*Ifnar1* mice have higher PVLs at 2 dpi than antibody treated B6 mice by more than one log (Manet et al. 2020).

The difference of outcome of these two models could be explained by the antibody kinetics, as only one injection of MAR1-5A3 is done one day prior to the infection. However, we previously showed that additional antibody injections post infection did not modify the susceptibility of infected mice (Manet et al. 2020).

As planned, the N2-*Irf3* cohort allowed to identify QTLs associated with the PVL. Two significant QTL were found on chromosome 12, *Zs1* and *Zs2*, associated with the PVL at 2 and 6 dpi, respectively, and two suggestive QTLs on chromosomes 2 and 17 were close to significance for the PVL at 2 dpi (Figure 20). The CC071 allele was associated with high PVL for these four QTLs (Figures 21, 22, and 23), thus, the loci underlying these QTLs may be involved in the high PVL observed in CC071 mice. As the CC071 allele was recessive, the variants explaining these QTLs are likely loss-of-function variants from CC071. As presented in Figure 24, the 3 QTLs associated with the PVL at 2 dpi have additive effects and together account for most of the difference of PVL observed between CC071 and B6-*Irf3* mice.

QTLs controlling the illness of N2-*Irf3* mice were also investigated. As expected by the mild severity of the model used for this cross, only a small number of QTLs were obtained for the body weight loss and the clinical scores, and only one, *Zsl3*, was significant (Figures 27 and 28).

By contrast, the F2-*Ifnar1* cohort was produced to investigate the genetic control of the disease signs observed in infected CC071 mice. Two significant QTLs, *Zsl5* and *Zsl6* were obtained for the susceptibility and clinical scores of F2-*Ifnar1* mice, and the CC071 allele was associated with higher mortality rate, earlier time to death, and higher clinical score (Figures 35 and 37). Notably, the mortality rate of mice carrying the CC071/CC071 genotype for *Zsl5* was more than twice that of mice carrying the 129/129 genotype, showing that this QTL is a major contributor to the survival of F2-*Ifnar1* mice (Figure 35). For most of the QTLs associated with the illness of F2-*Ifnar1* mice, the susceptibility allele was inherited from CC071 (Table 2), thus, many loci involved in the susceptibility to ZIKV of CC071 mice may have been identified in this cross.

We did not expect to identify a QTL controlled by *Irf3* in the N2-*Irf3* progeny, as all mice carried two non-functional alleles for this gene. In the F2-*Ifnar1* progeny, since only *Ifnar1*<sup>-/-</sup> mice were infected, the absence of IFN-I gene expression induced by IRF3 in *Irf3*<sup>71/71</sup> mice should not impact their susceptibility to ZIKV infection. By contrast, the IFN-I-independent functions of IRF3 could contribute to the susceptibility of F2-*Ifnar1* mice. Indeed, IRF3 can directly induce the expression of ISGs, including *Isg15*, *Ccl5* and *Isg54*, without the requirement for IFN-I signaling (Au et al. 1995; Génin et al. 2000; Navarro et al. 1998). IRF3 also induces the expression of IFN $\lambda$  (Donnelly and Kotenko 2010), and possesses the nontranscriptional RIPA and RIKA activities (Chattopadhyay et al. 2016; Popli et al. 2022).

One QTL on chromosome 7 was identified in the N2-*Irf3* progeny, but *Irf3* was not contained in its Bayesian credible interval, and the shape of the LOD score curve suggested that this QTL was controlled by another gene (Figure 27). Two QTLs were also obtained on chromosome 7 in the F2-*Ifnar1* progeny, but the peaks were distant from *Irf3*, thus, these QTLs are likely driven by other genes. To confirm that *Irf3* was not involved in the control of the susceptibility of N2-*Irf3* and F2-*Ifnar1* mice, congenic strains could be produced using CC071 as a donor strain. In conclusion, the susceptibility of N2-*Irf3* and F2-*Ifnar1* mice was primarily driven by other loci than *Irf3*.

An open question is whether the *Irf3* loss-of-function is necessary for the other susceptibility alleles of CC071 to induce high PVL and severe disease signs. In order to address this question, genetic engineering of the CC071 strain could be performed to obtain a CC071 strain with a functional *Irf3* locus, but this method requires to know the nature of the *Irf3* mutation in CC071.

Pardo Manuel de Villena and colleagues have investigated endogenous retrovirus (ERV) insertions in CC strains and identified about a hundred of fixed ERVs. In particular, they found an ERV inserted in *Irf3* in CC071, resulting in a chimeric RNA containing the beginning of the *Irf3* gene and the ERV (results presented at the 36<sup>th</sup> International Mammalian Genome Conference, Pardo Manuel De Villena et al. 2023). This insertion is about 200 base pairs long. Thus, a CC071 strain with a functional *Irf3* could be produced by CRISPR/Cas9 by excising the ERV insertion. Indeed, this method has been used to excise exonic regions of about 500 base pairs (Loesch et al. 2022), and CRISPR/Cas9 has already been successfully used on a CC strain (Jasperse et al. 2023).

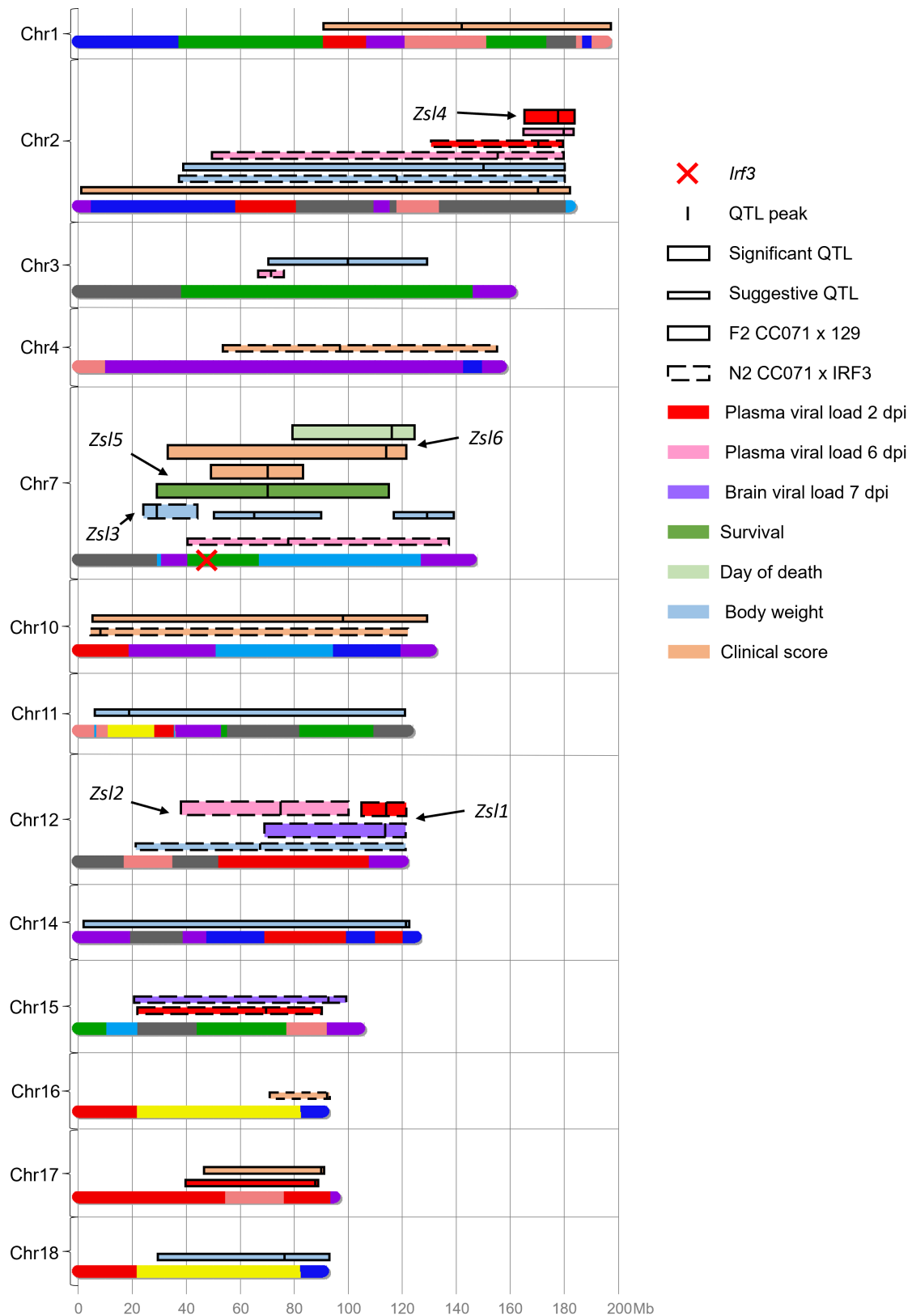
The analysis of two different segregating crosses involving CC071 as a parental strain allows

to investigate whether common QTLs were found. Two different situations could be obtained, first, colocalizing QTLs associated with the same trait in the two crosses, which we can define as "homologous" QTLs, or colocalizing QTL associated with different traits in the two crosses, that we may refer as "heterologous" QTLs. Indeed, as different experimental models were used for each cross, a single loci could affect the susceptibility of N2-*Irf3* and F2-*Ifnar1* mice in different ways.

The position of the QTLs identified in the two crosses is presented in Figure 39, which shows cases of colocalizing QTLs between the two crosses. For instance, I found a QTL on chromosome 2 controlling the PVL at 2 dpi of N2-*Irf3* mice in a region where CC071 inherited the B6 haplotype, thus where there is supposedly no polymorphism between B6-*Irf3* and CC071 (Figure 39). This QTL was obtained thanks to interval mapping in this region, but no genotyped marker showed a significant association with the PVL (Figure 22). Therefore, this QTL will require to be validated with further analysis. Besides, I also found a QTL on chromosome 2, *Zs14*, controlling the PVL in the F2-*Ifnar1* progeny (Figure 39). These two QTLs were located in the distal region of the chromosome, and the CC071 allele was associated with high viral loads, suggesting that the same homologous QTL was found in the two crosses, and reinforcing the hypothesis that the QTL found in the N2-*Irf3* progeny is a true association.

On chromosome 7, QTLs were associated with the survival, clinical scores and body weights of F2-*Ifnar1* mice, and the CC071 allele was associated with increased susceptibility (Figure 39 and Table 2). Besides, one QTL was suggestively associated with the PVL at 6 dpi of N2-*Irf3* mice, and the CC071 allele was associated with higher PVL. This may be a case of heterologous QTL, controlled by a variant inducing higher PVL in N2-*Irf3* mice and more severe disease signs in F2-*Ifnar1* mice.

In conclusion, analysis of the two crosses gave complementary results and both allowed the identification of QTLs involved in the susceptibility of SGIs. The use of different experimental models allowed to investigate different aspects of the pathology of ZIKV-infected CC071 mice, the PVL in antibody treated N2-*Irf3* mice, and the disease signs in F2-*Ifnar1* mice. Multiple QTLs were obtained, and most of them have large Bayesian credible intervals (Figure 39). Increasing the size of the N2-*Irf3* cohort may help reduce the intervals for some of the QTLs identified in this cross, but we may focus on the most significant and narrowest QTLs to identify candidate susceptibility genes, such as *Zs11* (Figure 21).



**Figure 39:** Summary of the QTLs found in the two crosses, highlighting potential pleiotropy. Chromosomes colors represent the CC founder haplotypes inherited by CC071 mice (A/J: yellow, B6: grey, 129S1/SvImj: pink, NOD/ShiLtj: dark blue, NZO/HILtj: light blue, CAST/Eij: green, PWK/Phj: red, WSB/Eij: purple). The colored bars represent the Bayesian credible intervals of the QTLs. The QTL interval color represents the phenotype. The outline represents the cross in which the QTL was identified. Thick bars represent significant QTLs while thin bars represent suggestive QTLs. Vertical bars represent the position of the peak LOD score of each QTL. Arrows indicate the names of significant QTLs.

## V.3 . Genetic factors modulating ZIKV susceptibility in *Ifnar1* deficient mice

### V.3.1 . Introduction

Mice are not natural hosts of ZIKV infection. Therefore, the study of ZIKV infection in mouse models requires the use of immunodeficient or sensitized mice (Lazear et al. 2016; Rossi et al. 2016). *Ifnar1* deficient mice, which lack one subunit of the IFN-I receptor, have extensively been used as a model of ZIKV pathogenesis (Lazear et al. 2016; Rossi et al. 2016; Winkler et al. 2017; Huang et al. 2017; Carbaugh et al. 2020; Gambino et al. 2021; Matz et al. 2021). Previous results from our laboratory showed that while B6-*Ifnar1* were highly susceptible to ZIKV infection with disease and mortality, 129-*Ifnar1* showed mild clinical manifestations, although the two strains had similar peak PVL (Manet et al. 2020).

In order to identify genes which modulate the severity of ZIKV disease in *Ifnar1* deficient mice (so-called modifier genes), C. Manet produced and analyzed an F2 cross between B6-*Ifnar1* and 129-*Ifnar1* mice. One hundred ninety-two F2 mice were infected i.p. with  $10^7$  FFUs of ZIKV FG15 and followed for 14 days. The majority of mice displayed mild, short-lasting symptoms (moderate body weight loss, ruffled fur and diminished activity), and 18.2% died in 14 days. The F2 progeny displayed a 2-log variation in PVL at 2 dpi, exceeding the mean values of the parental strains (Manet 2019).

Ninety-four F2 mice were genotyped using the miniMUGA array to perform QTL mapping. Two significant QTLs were identified. First, a QTL on chromosome 5 was associated with the PVL at 2 dpi, with a LOD score of 4.33. This QTL peaked at 84.66 Mb and its Bayesian credible interval spanned between 79.12 and 132.32 Mb. The B6 allele was recessive and the B6/B6 genotype was associated with lower viral PVL. The second QTL was associated with the day of death, with a LOD score of 3.69. This QTL was positioned on chromosome 12 at 12.0 Mb (credible interval: 2.79-61.58 Mb). The B6 allele was recessive and the B6/B6 genotype was associated with earlier death. Due to their large credible intervals, these two QTLs contained too many genes (546 and 230, respectively) to select candidates.

The recessivity of the B6 allele for these two QTLs could explain why the proportion of highly susceptible mice was low in the F2 progeny, and led us to produce a (B6-*Ifnar1* × 129-*Ifnar1*) × B6-*Ifnar1* N2, to increase the proportion of B6/B6 susceptible mice and facilitate the identification of loci associated with significant phenotypic differences between B6/B6 and B6/129 mice.

### V.3.2 . Results

#### V.3.2.1 . Phenotypic analysis of the parental strains and F1s

B6-*Ifnar1*, 129-*Ifnar1* and (B6-*Ifnar1* × 129-*Ifnar1*) F1 (F1-*Ifnar1*) mice were infected with ZIKV and followed for 14 dpi. B6-*Ifnar1* mice showed increasingly severe symptoms, including body weight loss, hunched posture, ruffled fur, ataxia and partial paralysis, with 60% mortality. By contrast, 129-*Ifnar1* and F1-*Ifnar1* mice showed only mild symptoms (hunched posture and ruffled fur) and quickly recovered (Figure 40A-B). The mortality rate of B6-*Ifnar1* mice in this study was lower than previously reported by our group (Log Rank test,  $p = 0.03$ , Manet et al. 2020). This can be explained by the smaller number of mice studied here (5 here versus 10 in the last study).

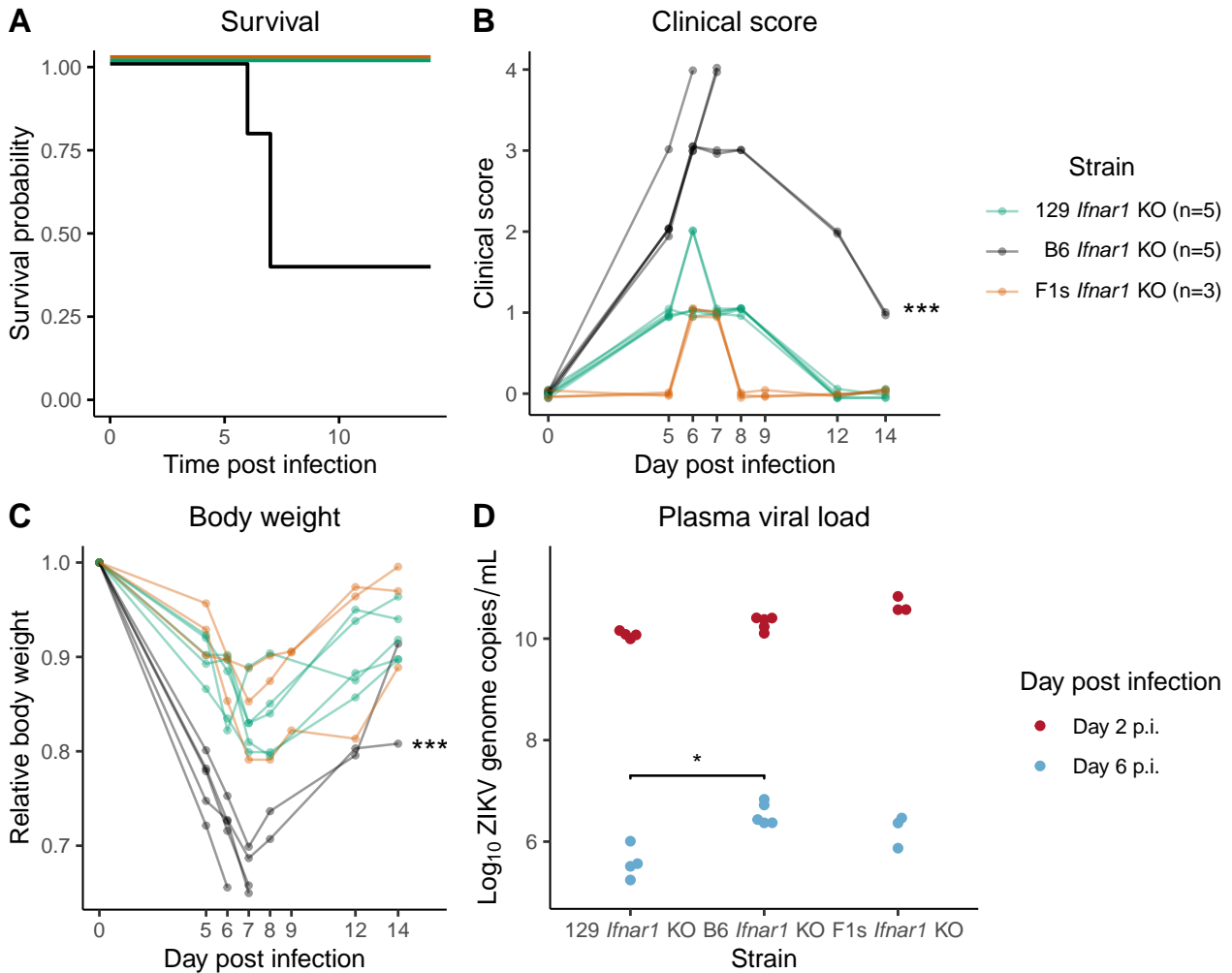
B6-*Ifnar1* mice uniformly lost body weight between the infection day and 7 dpi, and the surviving mice regained weight between 7 and 14 dpi, reaching 80% of their original weight (Figure 40C). B6-*Ifnar1* and 129-*Ifnar1* mice had similar PVL at 2 dpi, and F1-*Ifnar1* had slightly higher PVL, although this difference was not significant. At 6 dpi, B6-*Ifnar1* mice had significantly increased PVL compared with 129-*Ifnar1* mice, and F1-*Ifnar1* mice had intermediate PVL (Figure 40D). Overall, the phenotype of F1-*Ifnar1* mice was similar to that of 129-*Ifnar1* mice, consistent with the hypothesis that loss-of-function variants explain at least partially the susceptibility of B6-*Ifnar1* mice.

#### V.3.2.2 . Phenotypic analysis of the N2 progeny

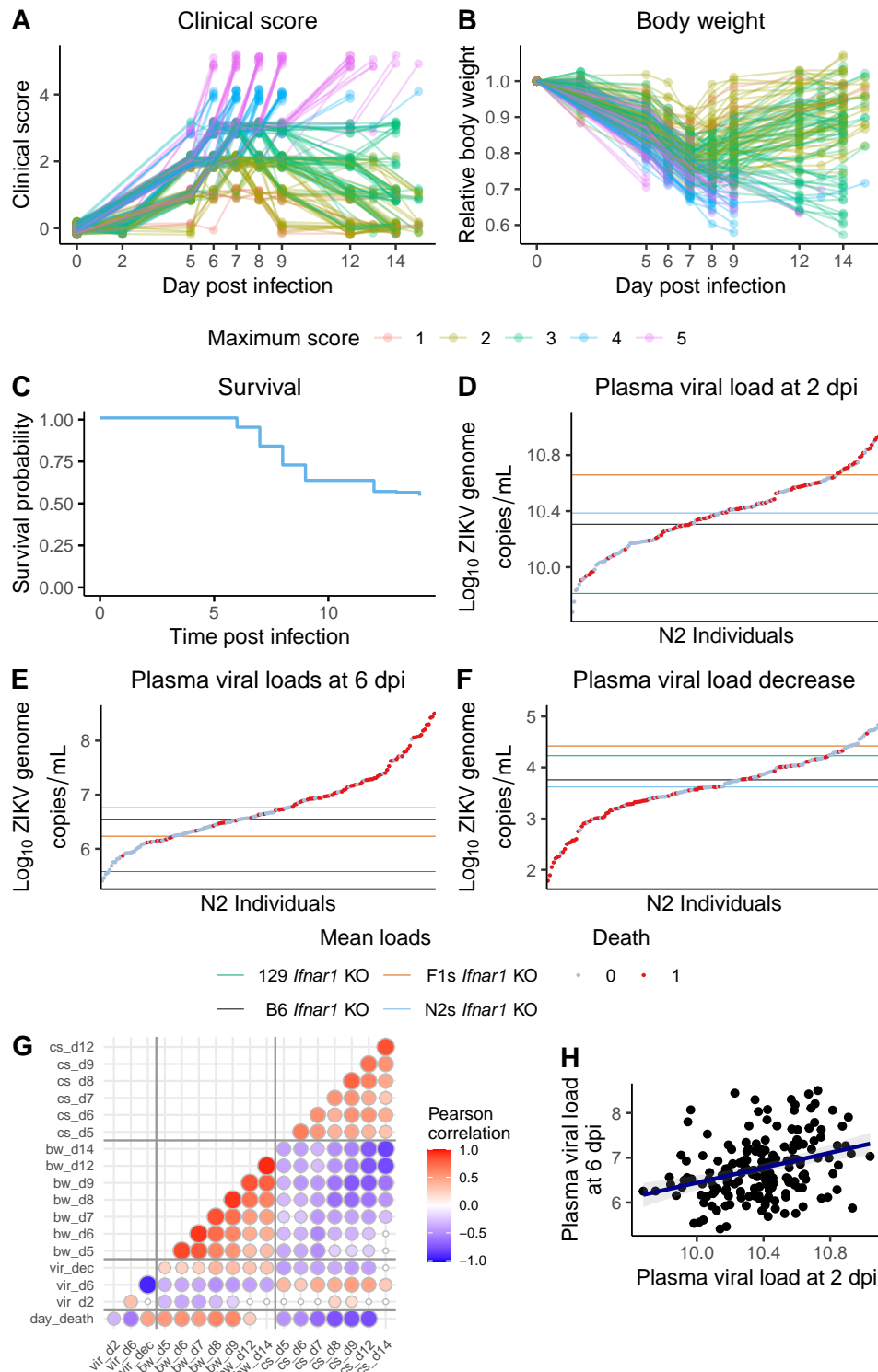
F1s were crossed with B6-*Ifnar1* mice to generate a cohort of N2 mice. 197 6 to 9-week-old (B6-*Ifnar1* × 129-*Ifnar1*) × B6-*Ifnar1* N2 mice were infected i.p. with  $10^7$  FFUs of FG15 and monitored for 14 - 15 days. Eighty-nine out of the 197 mice (45.2%) of the N2 progeny mice did not survive the infection and died between 6 and 14 dpi (Figure 41C). These mice showed severe symptoms including body weight loss, ataxia, total limb paralysis and hypoactivity (Figures 41A and B). Mice that survived the infection showed a variability of clinical manifestations, from slightly hunched posture to hypoactivity, with maximum clinical scores ranging between 1 and 3, and lost weight between 5 and 7 dpi (Figures 41A and B). Most of the surviving mice regained weight between 9 and 14 dpi, but 21 of them (19.4% of the surviving mice) continued to lose weight and kept elevated clinical scores during that period (Figure 41A and B). Therefore, a combination of B6-*Ifnar1* and 129-*Ifnar1* alleles may lead to long ZIKV disease.

At 2 dpi, the PVL showed moderate variation and ranged over 1.3 log, consistent with the reduced variability between B6-*Ifnar1* and 129-*Ifnar1* mice (Figure 41D), and was higher superior in mice that succumbed (t test,  $p = 6.1e-05$ ). At 6 dpi the PVL ranged over 3 logs, surprisingly exceeding the values observed in the parental strains (Figure 41E). The PVL at 6 dpi was significantly higher in mice that succumbed (t test,  $p = 3.22e-14$ ).

We then tested the effect of sex on these phenotypes. The survival rate was similar between males and females, but females died slightly faster than males, although this was not significant (Figure S8A-B). Both sexes had the same PVL at 2 dpi, but males showed higher PVL at 6 dpi (Figure S8C). Males and females had similar body weight curves (Figure S8D) but males had higher clinical scores from 8 dpi (Figure S8E). These results suggest that female N2 mice died faster than males, but female mice that survived recovered faster than males.



**Figure 40:** ZIKV disease severity in *Ifnar1*-deficient mice is controlled by the genetic background. Six to 8-week-old B6-*Ifnar1* (n=5), 129-*Ifnar1* (n=5) and F1-*Ifnar1* (n=3) mice were infected i.p. with  $10^7$  FFUs of FG15 and monitored for 14 days. (A) Kaplan-Meier survival curves showing 60% lethality in B6-*Ifnar1* and 100% survival in 129-*Ifnar1* and F1-*Ifnar1*. (B) Individual curves of clinical score, with numerical values given as follows: 1, slight hunched posture; 2, ruffled fur, hunched posture and/or mild ataxia; 3, prostration, ataxia, partial limb paralysis; 4, ataxia, total limb paralysis or moribund; 5, dead. (C) Individual curves of body weight loss, expressed as the percent of the starting weight (day of infection). Mice were weighed at day 5, 6, 7, 8, 9, 12 and 14 post infection. (D) PVLs, quantified at days 2 and 6 post infection by RT-qPCR. Groups were compared by ANOVA followed by Tukey HSD (\*  $p < 0.05$ , \*\*\*  $p < 0.001$ ).



**Figure 41:** N2 mice show large phenotypic variability.

Six to 9-week-old N2 mice ( $n=197$ ) were infected i.p. with  $10^7$  FFUs of FG15 and monitored for 14 days. (A) Curves of clinical scores, expressed with numerical values as in Figure 40. (B) Body weight loss curves, expressed as the percent of the starting weight (day of infection). Mice were weighed at day 2, 5, 6, 7, 8, 9, 12 and 14 post infection. Curves were colored by the maximum clinical score of each mouse. (C) Survival curve of F2 mice. (D - F) PVLs, quantified at days 2 (D) and 6 (E) post infection by RT-qPCR. (F) PVL decrease was calculated as the difference between the load at 2 dpi and the load at 6 dpi. Horizontal lines show the average PVL of 129-*Ifnar1* (brown), B6-*Ifnar1* (black), F1-*Ifnar1* (green). Mice that survived the infection ("0") are represented in blue and mice that died due to the infection ("1") are represented in red. (G) Correlation matrix between of the phenotypes of the N2 mice. The strength of correlation is represented by both circle sizes and colors. Small white circles show nonsignificant correlations (H) Correlation between PVLs at 2 and 6 dpi (Pearson correlation coefficient:  $R = 0.33$ ,  $p = 6.2e-6$ ).



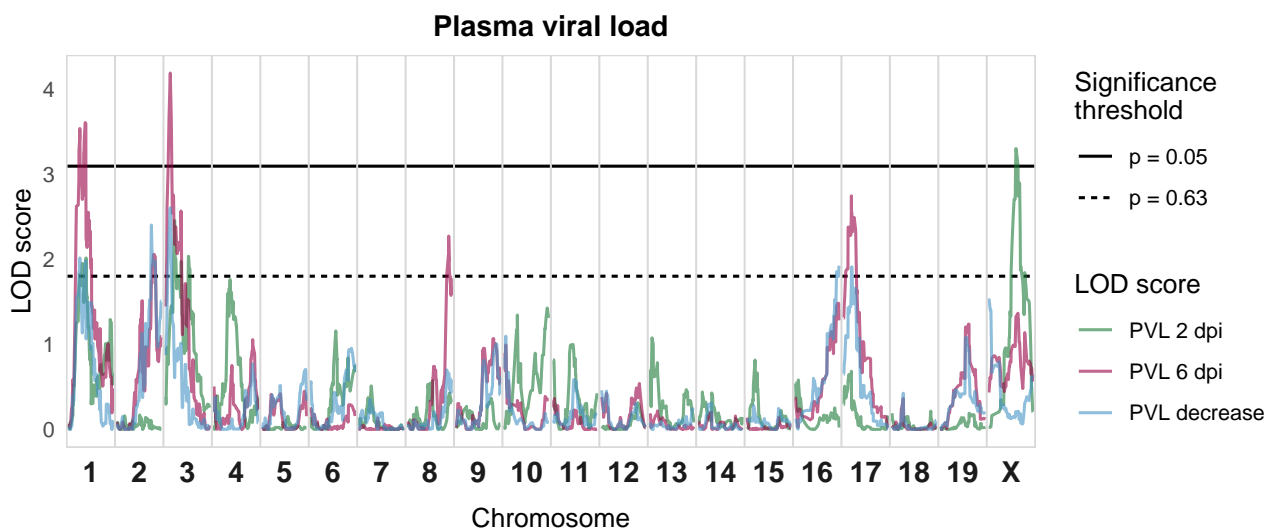
The correlations between the different traits measured in the N2 progeny were assessed (Figure 41G). Groups of similar traits (i.e., PVL, body weights, clinical scores), showed significant intra-group correlation. PVLs at 2 and 6 dpi were slightly correlated (Pearson correlation coefficient, 0.23,  $p = 0.0021$ , Figure 41H). This correlation was significant due to the number of mice analyzed, but mice with similar PVL at 2 dpi showed a high variability of PVL at 6 dpi, ranging over more than 2 logs. PVLs at 2 and especially at 6 dpi were globally correlated with the day of death, body weights and clinical scores, suggesting that the illness of N2 mice is influenced by the levels of circulating viruses.

### V.3.2.3 . Three significant QTLs control plasma viral load in N2 mice

The N2 progeny was genotyped with the MiniMUGA array containing 11,125 markers spread across the genome. The stuart R package (Bourdon and Montagutelli 2022) was used to clean the raw genotype data and 2,882 markers were retained for genetic analyses (Figure S9). Multiple QTLs were found in this cross and are presented in Table 4.

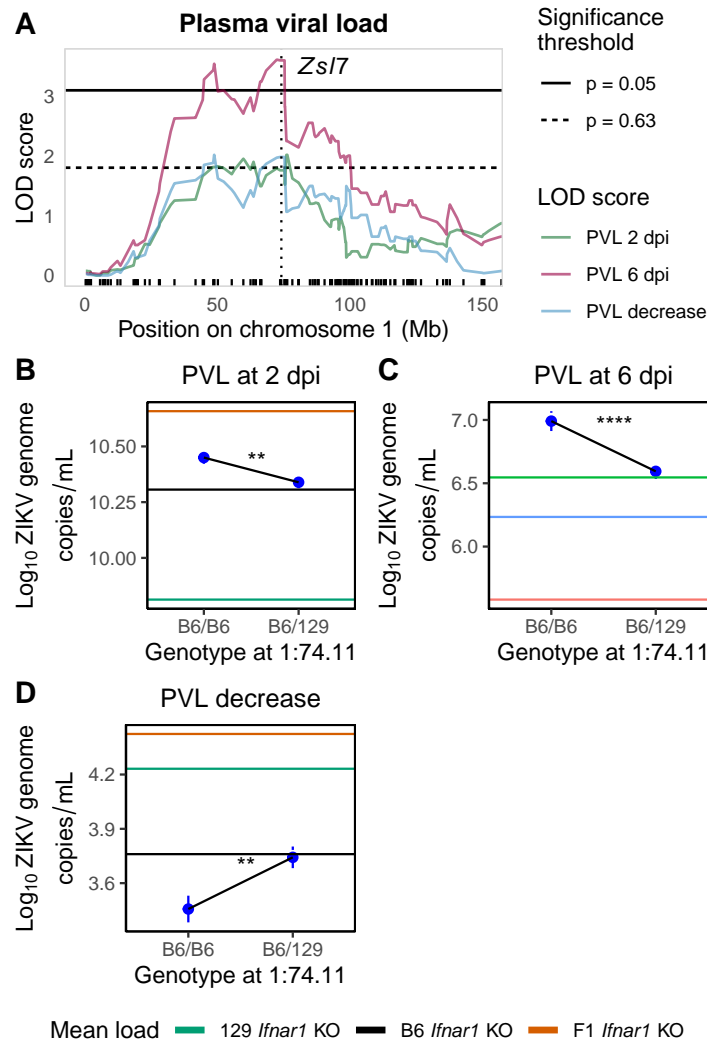
Three significant and 4 suggestive QTLs were associated with PVLs (Figure 42). Out of them, six were associated with the PVL at 6 dpi, consistent with the larger variability of this trait compared with the PVL at 2 dpi in N2 mice (Figure 41D).

Zs/7 on chromosome 1 was significantly associated with the PVL at 6 dpi (Figure 43A). Its peak LOD score was 3.607 at 74.11 Mb. This QTL explained 9% of the variance of this trait (Table 4). This QTL also showed suggestive associations with the PVL at 2 dpi and the PVL decrease, and the B6/B6 genotype was associated with higher PVLs at 2 and 6 dpi, and lower PVL decrease (Figure 43B - D).



**Figure 42:** QTL mapping for the PVL.

QTL mapping was performed on the PVL at 2 and 6 dpi and on the PVL decrease. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing ( $n=1000$ ). For better visibility, and as the significant thresholds were very close for the three traits ( $p = 0.05$ , 3.24 for the PVL at 2 dpi, 3.09 for 6 dpi, and 3.13 for the decrease;  $p = 0.63$ , 1.81 for the PVL at 2 dpi, 1.80 for 6 dpi, and 1.80 for the decrease), only the lower value was displayed.



**Figure 43:** *Zs17* on chromosome 1 associated with the PVL.

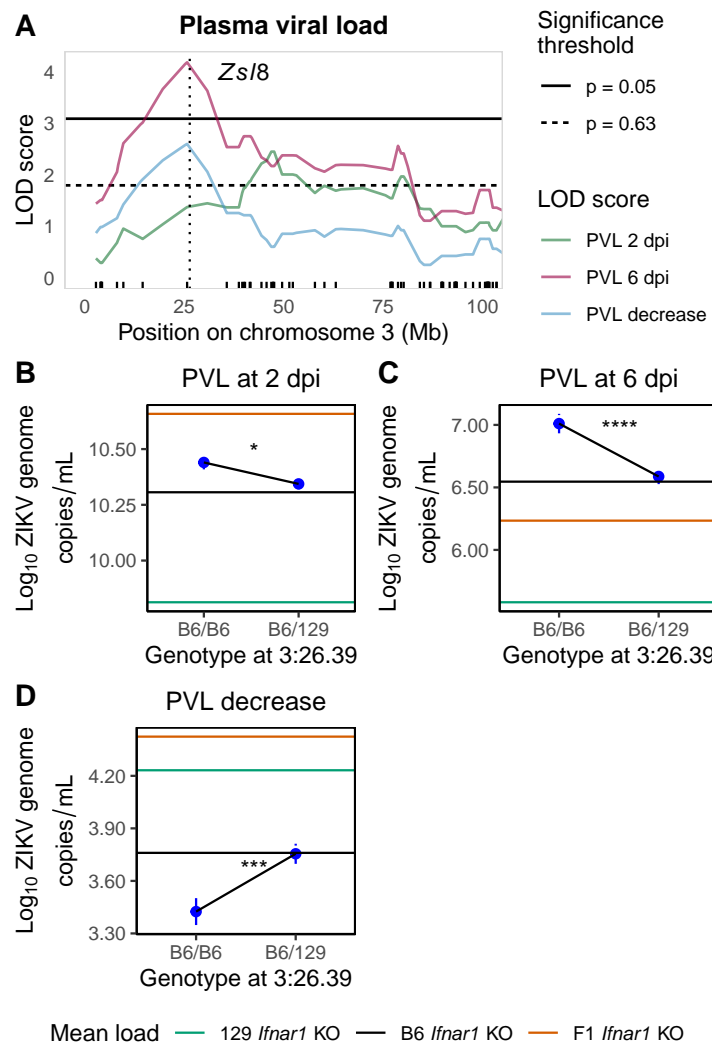
(A) Zoom on the chromosome 1 QTL associated with the PVL at 6 dpi. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 42. The vertical dotted line represents the positions of the marker used to display allelic effects in B, C, and D. (B - D) Allelic effects assessed at the marker with the highest LOD score for the PVL at 6 dpi: mUNC010142203, at 74.11 Mb. Results are represented as mean  $\pm$  sem. Groups were compared by t-test (\*\*  $p < 0.01$ , \*\*\*\*  $p < 0.0001$ ). Horizontal lines show the average PVL of B6-*Ifnar1* (black), 129-*Ifnar1* (green), and F1-*Ifnar1* (brown).

**Table 4:** Associations between QTLs and phenotypes in the N2 progeny.

QTL	Phenotype	Chr	Position	LOD score	p-value	CI (Mb)	PVE (%)	Effect
<i>Zsl7</i>	PVL (6 dpi)	1	74.10	3.607	0.018	37.68 - 99.42	9.00	B6 > 129
<i>Zsl7</i>	PVL (2 dpi)	1	77.58	2.015	0.449	37.5 - 194.6	5.0	B6 > 129
<i>Zsl7</i>	PVL (decrease)	1	50.50	2.015	0.471	33.8 - 151.0	5.1	B6 < 129
<i>Zsl8</i>	PVL (6 dpi)	3	26.38	4.189	0.007	17.06 - 65.42	10.38	B6 > 129
<i>Zsl8</i>	PVL (2 dpi)	3	38.20	2.454	0.208	17.9 - 121.6	6.1	B6 > 129
<i>Zsl8</i>	PVL (decrease)	3	26.39	2.604	0.168	10.2 - 68.2	6.6	B6 < 129
<i>Zsl9</i>	PVL (2 dpi)	X	116.01	3.301	0.043	92.43 - 151.83	8.10	B6 < 129
<i>Zsl10</i>	Mortality	1	39.67	3.193	0.048	37.49 - 76.12	7.64	B6 > 129
<i>Zsl11</i>	Mortality	4	153.04	3.832	0.012	137.10 - 154.96	9.10	B6 > 129
<i>Zsl11</i>	Day of death	4	148.24	2.905	0.088	137.10 - 156.00	6.98	B6 < 129
<i>Zsl11</i>	CS (7 dpi)	4	150.54	3.231	0.036	133.74 - 154.96	8.11	B6 > 129
<i>Zsl12</i>	BW (5 dpi)	1	174.82	5.130	0.001	169.90 - 187.12	11.99	B6 < 129
<i>Zsl12</i>	BW (6 dpi)	1	177.71	7.358	<0.001	170.27 - 187.12	17.16	B6 < 129
<i>Zsl12</i>	BW (7 dpi)	1	178.10	6.249	<0.001	168.90 - 189.46	15.83	B6 < 129
<i>Zsl12</i>	BW (8 dpi)	1	170.27	4.382	0.007	39.67 - 189.46	12.99	B6 < 129
<i>Zsl13</i>	CS (6 dpi)	9	65.74	3.145	0.048	34.98 - 106.08	7.49	B6 > 129
<i>Zsl14</i>	CS (8 dpi)	X	144.50	3.650	0.019	139.43 - 162.83	10.28	B6 < 129
-	PVL (6 dpi)	2	165.63	2.057	0.447	77.2 - 181.4	5.2	B6 > 129
-	PVL (decrease)	2	160.56	2.402	0.255	102.5 - 181.4	6.1	B6 < 129
-	PVL (6 dpi)	8	122.77	2.272	0.286	81.7 - 129.3	5.8	B6 > 129
-	PVL (decrease)	16	80.72	1.922	0.529	140.7 - 80.3	4.9	B6 > 129
-	PVL (6 dpi)	17	24.5	2.745	0.118	3.3 - 47.6	6.9	B6 > 129
-	PVL (decrease)	17	26.3	1.910	0.540	3.3 - 67.0	4.9	B6 < 129
-	Mortality	3	40.17	2.301	0.308	17.1 - 104.3	5.6	B6 > 129
-	Mortality	8	129.33	2.403	0.250	95.3 - 129.4	5.8	B6 > 129
-	Mortality	9	117.2	1.897	0.541	51.8 - 123.9	4.6	B6 > 129
-	Day of death	14	101.5	2.257	0.306	32.4 - 119.8	5.5	B6 > 129
-	BW (8 dpi)	1	39.67	2.760	-	-	8.4	B6 < 129
-	BW (9 dpi)	1	39.67	2.095	-	-	7.4	B6 < 129
-	BW (8 dpi)	3	26.38	2.117	0.433	10.2 - 156.1	6.5	B6 < 129
-	BW (9 dpi)	3	26.33	2.079	0.406	10.1 - 146.6	7.3	B6 < 129
-	BW (8 dpi)	4	55.45	1.970	0.520	10.6 - 91.44	6.1	B6 > 129
-	BW (9 dpi)	4	57.24	2.491	0.209	35.8 - 91.4	8.7	B6 > 129
-	BW (12 dpi)	4	55.45	2.187	0.335	13.4 - 106.2	11.0	B6 > 129
-	BW (12 dpi)	13	39.17	2.137	0.374	3.6 - 110.8	10.8	B6 > 129
-	BW (8 dpi)	X	153.54	3.045	0.072	143.7 - 164.0	9.2	B6 > 129
-	BW (9 dpi)	X	153.54	2.759	0.117	140.5 - 167.4	9.6	B6 > 129
-	BW (14 dpi)	X	162.83	2.741	0.119	5.5 - 168.2	13.7	B6 > 129
-	CS (9 dpi)	1	45.63	2.341	0.260	25.7 - 136.4	7.7	B6 > 129
-	CS (12 dpi)	1	81.30	2.277	0.308	37.5 - 172.7	10.3	B6 > 129
-	CS (14 dpi)	1	162.9	1.913	0.524	43.2 - 190.3	9.8	B6 > 129
-	CS (12 dpi)	2	181.44	2.595	0.167	136.6 - 181.5	11.7	B6 > 129
-	CS (12 dpi)	5	150.33	1.871	0.562	3.2 - 150.4	8.6	B6 > 129
-	CS (6 dpi)	8	31.62	2.226	0.320	13.2 - 118.2	5.4	B6 > 129
-	CS (7 dpi)	8	23.83	2.159	0.375	3.5 - 129.3	5.4	B6 > 129
-	CS (12 dpi)	15	80.76	1.837	0.586	27.8 - 102.3	8.4	B6 < 129

Survival and day of death were studied together as a two-part model (here referred as "susceptibility" phenotype). CI: credible interval, PVE: percentage of variance explained.

A second significant QTL for the PVL at 6 dpi was found on chromosome 3 (Figure 44A). This QTL, named *Zsl8*, had a peak LOD score of 4.189 at 26.39 Mb, with a large credible interval (between 17.86 and 65.42 Mb), which could suggest a second, distal, QTL. Suggestive associations were also observed for the PVL at 2 dpi and the PVL decrease. For the decrease, the maximum LOD score was at the peak of *Zsl8*, while the LOD score for the PVL at 2 dpi peaked more distally on the chromosome (Figure 44A), supporting the hypothesis of two distinct QTLs on this chromosome. The B6/B6 genotype for *Zsl8* was associated with higher PVL at 2 and 6 dpi, and lower PVL decrease (Figure 44B - C).

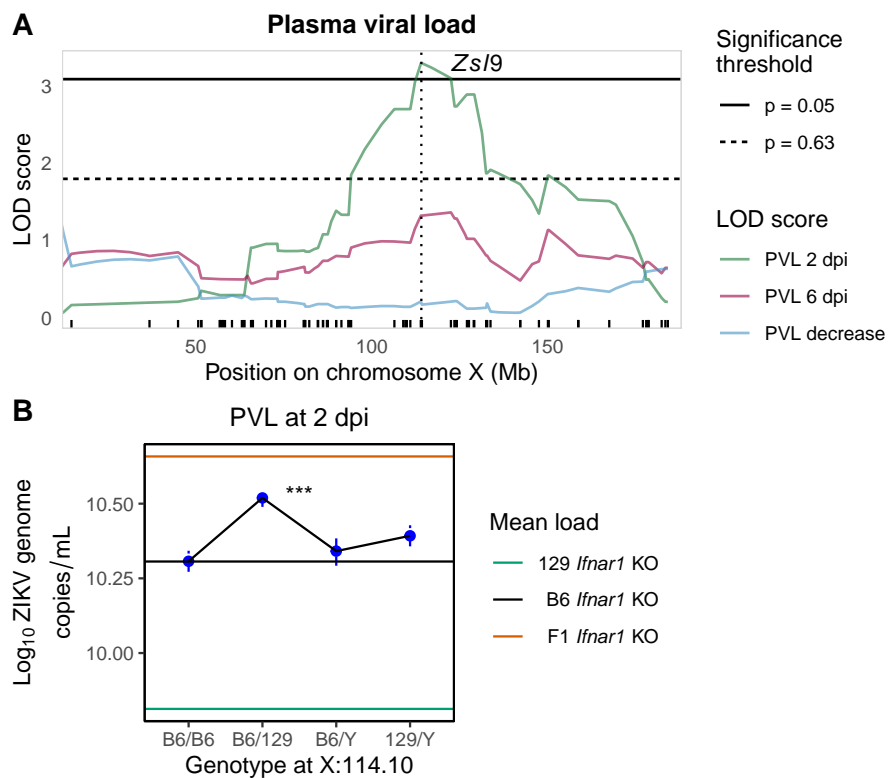


**Figure 44:** *Zs/8* on chromosome 3 associated with the PVL.

(A) Zoom on the chromosome 3 QTL associated with the PVL at 6 dpi. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 42. The vertical dotted line represents the positions of the marker used to display allelic effects in B and C. (B - D) Allelic effects assessed at the marker with the highest LOD score for the PVL at 6 dpi: gUNC4904534, at 23.39 Mb, on the PVL at 2 dpi (B) and 6 dpi (C). Results are represented as mean  $\pm$  sem. Groups were compared by t-test (\*  $p < 0.05$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ). Horizontal lines show the average PVL of B6-*Ifnar1* (black), 129-*Ifnar1* (green), and F1-*Ifnar1* (brown).

One QTL was significantly associated with the PVL at 2 dpi on chromosome X. This QTL was named *Zs/9* and its peak was at 114.10 Mb with a LOD score of 3.30 (Figure 45A) and explained 8.1% of the variance (Table 4). The effect of this QTL was transgressive (B6/B6 mice showing lower PVL than B6/129 mice) in female N2 mice (Figure 45B).

Four other suggestive QTLs were obtained for the PVLs of N2 mice. These QTLs were not given a *Zs/* number as they did not reach the significance threshold. Two suggestive QTLs were suggestively associated with the PVL at 6 dpi and with the PVL decrease on chromosomes 2 and 17. A suggestive QTL was suggestively associated with the PVL at 6 dpi on chromosome 8, but was not associated with the PVL at 2 dpi or the decrease. Lastly, a suggestive QTL on chromosome 16 was suggestively associated with the PVL decrease. For these suggestive QTLs, the B6/B6 genotype was associated with higher PVL and/or lower decrease (Figure S10 and Table 4).



**Figure 45:** *Zs/9* on chromosome X associated with the PVL.

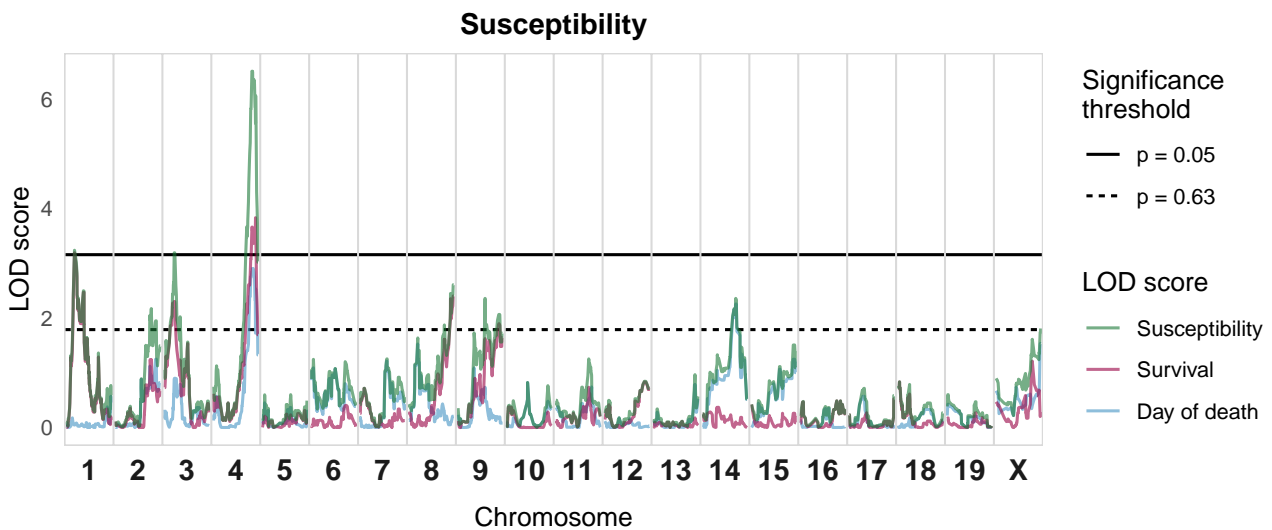
(A) Zoom on the chromosome X QTL associated with the PVL at 2 dpi. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 42. The vertical dotted line represents the positions of the marker used to display allelic effects in B. (B) Allelic effects assessed at the marker with the highest LOD score for the PVL at 2 dpi: S3R204607613, at 114.10 Mb. Results are represented as mean  $\pm$  sem. Groups were compared by ANOVA (\*\*\*)  $p < 0.001$ ). Horizontal lines show the average PVL of B6-*Ifnar1* (black), 129-*Ifnar1* (green), and F1-*Ifnar1* (brown).

### V.3.2.4 . QTL mapping identified two loci significantly associated with the survival of N2 mice

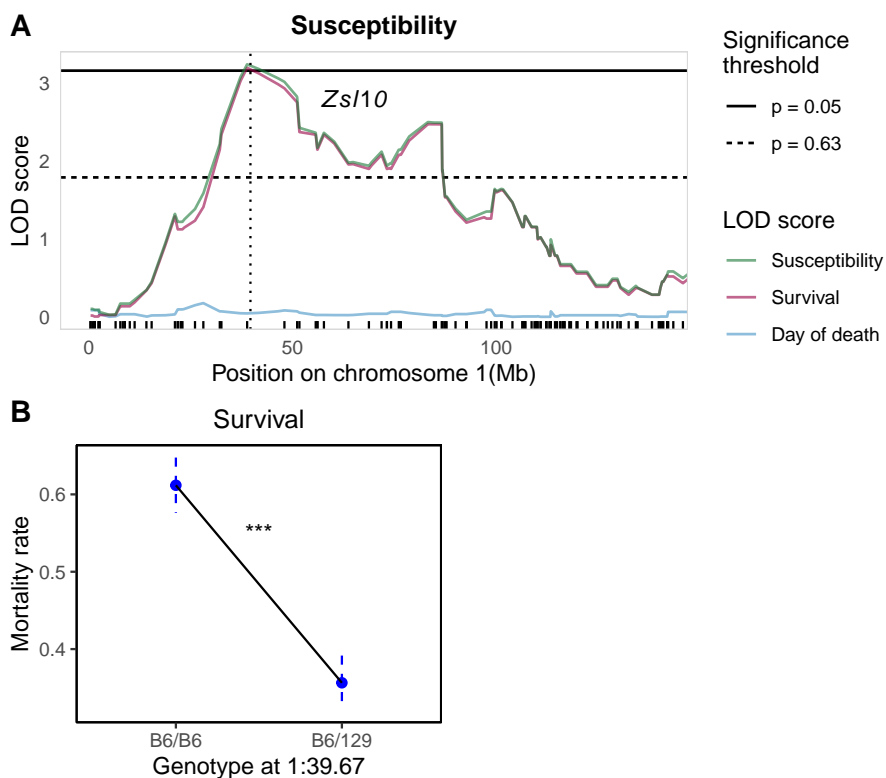
QTL mapping was performed on the survival and day of death using a two-part model, as described previously in section V.2.2.8 of the Results (Broman 2003). Two significant QTLs were obtained (Figure 46). A QTL named *Zs/10* on chromosome 1 was associated with the survival, with a LOD score of 3.19 at 39.67 Mb. Mice carrying the B6/B6 and the B6/129 genotypes at this locus a mortality rate of 61.2% and 35.6%, respectively (Figure 47B).

*Zs/11*, located distally on chromosome 4 was associated with both the survival and the day of death (Figure 48A). The resulting LOD score for the "susceptibility" trait was 6.508 at 143.02 Mb. *Zs/11* explained 9.1% of the variance of the survival rate and 7% of the variance of the day of death (Table 4). The B6 allele was associated with higher mortality rate and earlier death (Figure 48B - C). Therefore, for both QTLs, the effects were consistent with the parental phenotypes.

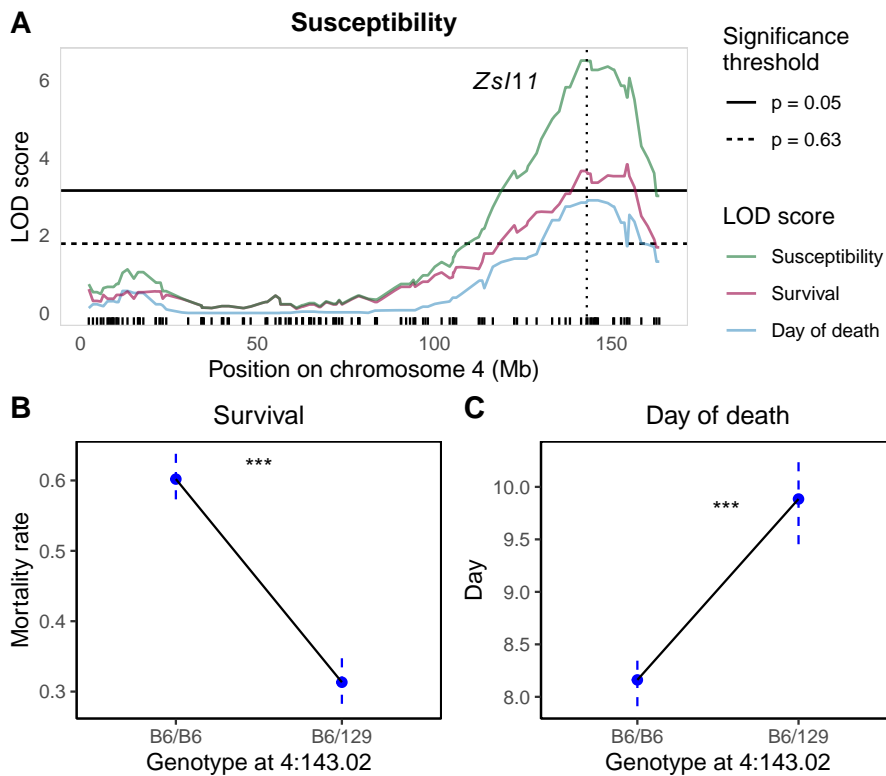
We investigated the mortality rate of N2 mice grouped by genotypes at *Zs/10* and *Zs/11* loci. Double B6/B6 homozygous mice had the highest mortality rate (72.7%), while double B6/129 heterozygous mice had the lowest rate (14.3%). Mice homozygous for either *Zs/10* or *Zs/11* had intermediate mortality rate, around 50% (Figure 49). These results that *Zs/10* and *Zs/11* have additive effects on the survival of N2 mice and together majorly contribute to this trait.



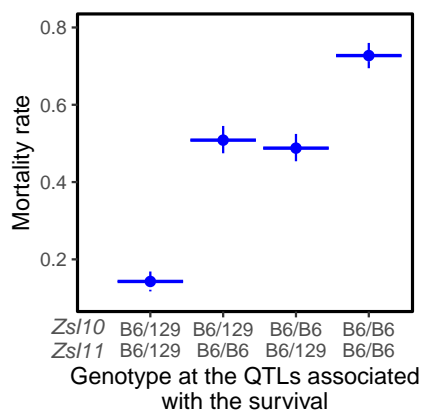
**Figure 46:** QTL mapping for the susceptibility of the N2 mice. QTL mapping was performed on survival load and day of death studied with a two-part model. The combined trait for the analysis of the combination of the survival and the day of death is referred as "susceptibility". The horizontal lines indicate genome-wide significance thresholds determined by permutation testing (n=1000). For better visibility, only the thresholds for the susceptibility are displayed (p = 0.05, 4.0 for the susceptibility, 3.18 for the survival, and 3.16 for the day of death; p = 0.63, 2.5 for the susceptibility, 1.79 for the survival, and 1.82 for the day of death).



**Figure 47:** *Zs10* on chromosome 1 associated with the survival and day of death of N2 mice. (A) Zoom on the chromosome 1 QTLs associated with the survival and day of death. Horizontal lines indicate genome-wide significance thresholds (p = 0.05, p = 0.63) determined by permutation testing (n=1000), as in Figure 46. The vertical dotted line represents the positions of the marker used to display allelic effects in B. (B) Allelic effects assessed at the marker with the highest LOD score for the survival: gUNC496797, at 39.67 Mb. Results are represented as mean  $\pm$  sd. Groups were compared by  $\chi^2$  (\*\*\*) p < 0.001).



**Figure 48:** *Zs11* on chromosome 4 associated with the survival and day of death of N2 mice. (A) Zoom on the chromosome 3 QTLs associated with the survival and day of death. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 46. The vertical dotted lines represent the positions of the marker used to display allelic effects in B. (B - C) Allelic effects assessed at the marker with the highest LOD score for the susceptibility trait: mUNC8397172, at 143.02 Mb. Results are represented as mean  $\pm$  sd (B) or mean  $\pm$  sem (C). Groups were compared by  $\chi^2$  (B) or t-test (C) (\*\*\*)  $p < 0.001$ .



**Figure 49:** Additive effects of the two significant QTLs controlling the survival of N2 mice. Allelic effects assessed at gUNC496797 and mUNC8397172 (see Figures 47, and 48). Groups were sorted by increasing proportion of B6 alleles.

Three suggestive QTLs were also detected for the survival (Figure S11A), on chromosomes 3 (LOD = 2.301), 8 (LOD = 2.403) and 9 (LOD = 1.897). For these three suggestive QTLs, the B6/B6 genotype was associated with higher mortality rate (Figure S11B - D). A fourth suggestive QTL with a LOD score of 2.257 was observed for the day of death on chromosome 14. This QTL was transgressive since the B6/B6 genotype was associated with later time to death than the B6/129 genotype (Figure S11E).

#### **V.3.2.5 . One significant and five suggestive QTLs control the body weight loss of N2 mice**

QTL mapping was performed on body weights at 5, 6, 7, 8, 9, 12, and 14 dpi (Figure 50). One significant QTL was detected, *Zs/12*, and was significantly associated with the weights at 5, 6, 7, and 8 dpi, and was suggestively associated with the weights at 9 and 12 dpi, probably due to the correlation between body weights at early and late time points (Figure 41G). The maximum LOD score was obtained for the body weight at 6 dpi (7.358, at 177.72 Mb). Mice with the B6/B6 genotype at *Zs/12* had significantly lower relative body weights at all time points post infection (Figure S12B).

Five suggestive QTLs were also obtained (Figure 50A and Table 4). For four of them (located on chromosomes 1, 3, 4, and X), the peak of LOD score reached its maximum between 8 and 9 dpi, but one suggestive QTL on chromosome 13 was associated with the weight at 12 dpi (Figure 50B), suggesting that the latter QTL is associated with the weight in the late infection phase. The QTLs on chromosomes 4, 13 and X showed transgressive alleles, as B6/B6 mice had higher body weights than B6/129 mice (Figure S12D - F).

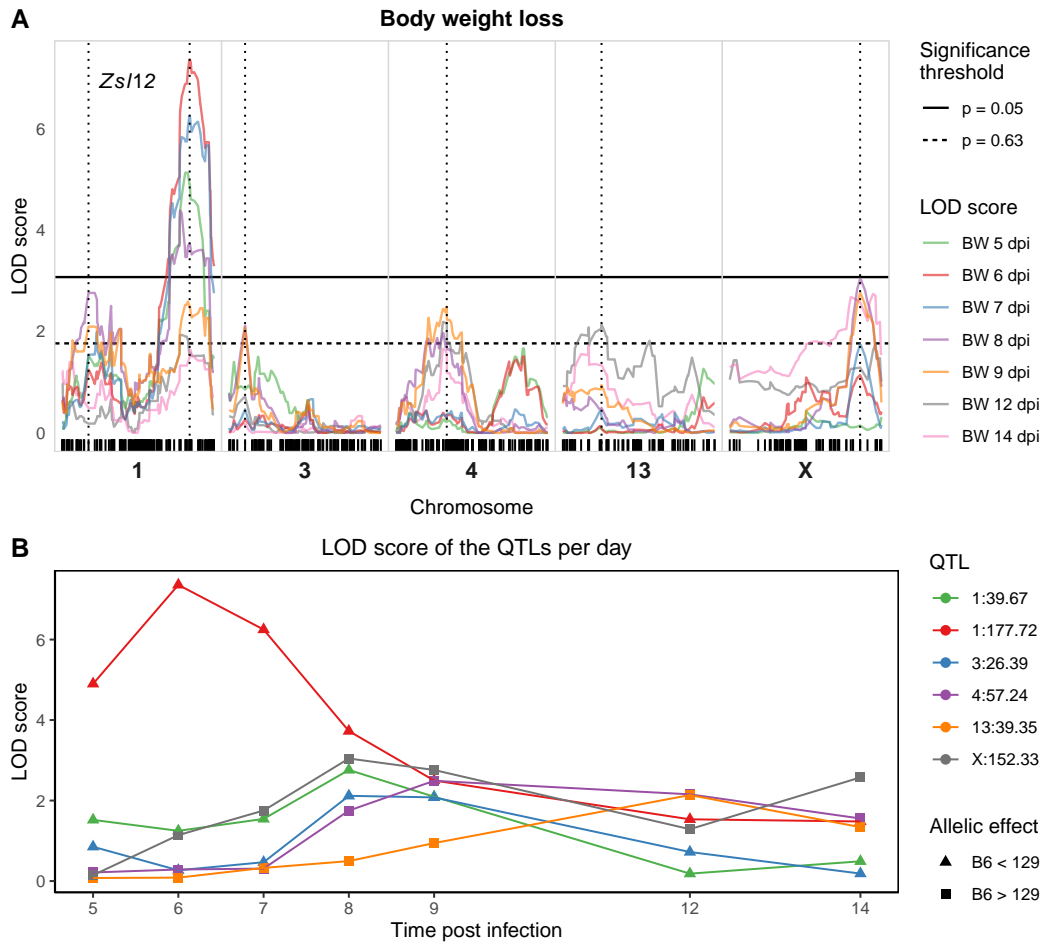
#### **V.3.2.6 . Three significant and five suggestive QTLs control the clinical scores of N2 mice**

QTL mapping was performed on the clinical scores of N2 mice at 5, 6, 7, 8, 9, 12 and 14 dpi (Figure 51). Three significant QTLs were obtained. First, a QTL on chromosome 4 was significantly associated with the clinical score at 7 dpi (Figure 51A). The peak LOD score was 3.231 ( $p = 0.036$ ). This QTL explained 8.1% of the variance of this trait (Table 4). Mice with the B6/B6 genotype had significantly higher clinical scores than mice with the B6/129 genotype, but only at 7 dpi (Figure 51D). This QTL colocalized with *Zs/11*, a QTL controlling the survival and day of mice of N2 mice identified previously in section V.3.2.4, and similar allelic effects were obtained for these two QTLs (B6/B6 mice had higher clinical scores and mortality rate, and earlier time to death, Figures 48 and S13C). These results suggest that a single genetic variant controls both the clinical score at 7 dpi and the survival. The two traits could also be controlled by closely linked but distinct genes.

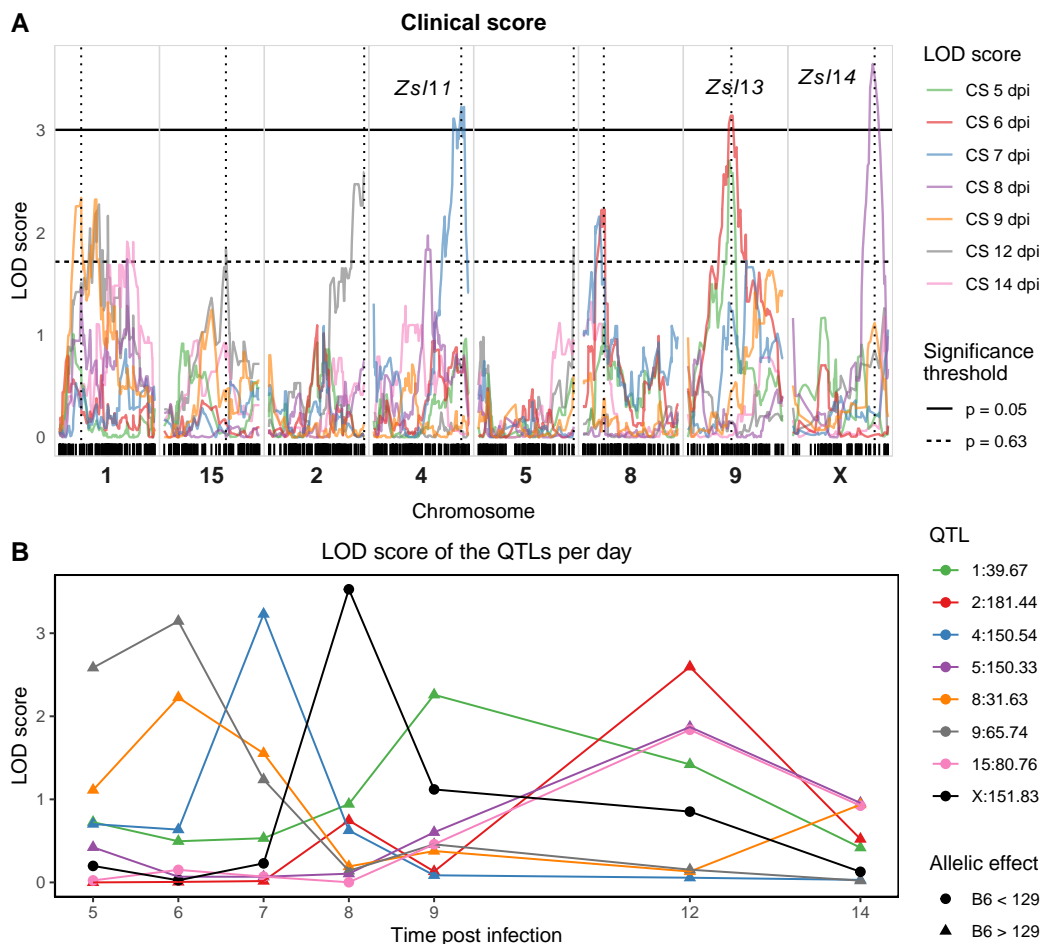
The second significant QTL was found on chromosome 9. This QTL, named *Zs/13*, was significantly associated with the clinical score at 6 dpi, and suggestively associated with the score at 5 dpi (Figure 51A). The LOD score at the QTL peak was 3.145 ( $p = 0.048$ ), and this QTL explained 7.5% of the variance of the clinical score at 6 dpi (Table 4). Mice with the B6/B6 genotype for *Zs/13* had significantly higher scores at 5, 6, and 7 dpi (Figure S13F).

The third significant QTL, named *Zs/14*, was located on chromosome X. This QTL was significantly associated with the clinical score at 8 dpi, and was not associated with the scores at the other time points post infection (Figure 51A). B6/B6 females had lower clinical scores at B6/129 females (Tukey USD,  $p = 4e-4$ ), while B6/Y and 129/Y males had on average the same score (Tukey HSD,  $p = 0.86$ , Figure S13H).





**Figure 50:** QTL mapping for the body weight loss in the N2 progeny displaying *Zs/12* and five suggestive QTLs. (A) QTL mapping was performed on relative body weight at 5, 6, 7, 8, 9, 12 and 14 dpi. Only the chromosomes with suggestive or significant associations were represented. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing ( $n=1000$ ). For better visibility, and as the significant thresholds are highly similar for the body weight loss at different time points ( $p = 0.05$ , 3.08 for 5 dpi, 3.21 for 6 dpi, 3.19 for 7 dpi, 3.19 at 8 dpi, 3.13 at 9 dpi, 3.07 at 12 dpi, 3.13 at 14 dpi,  $p = 0.63$ , 1.80 for 5 dpi, 1.80 for 6 dpi, 1.79 for 7 dpi, 1.83 at 8 dpi, 1.77 at 9 dpi, 1.80 at 12 dpi, 1.79 at 14 dpi), only the minimum threshold was displayed. Vertical dotted lines represent the positions of the markers with the highest LOD scores used to display the LOD score per day in (B) and the allelic effects in Figure S12. (B) LOD score for each QTL identified at each time post infection tested. BW = relative body weight.



**Figure 51:** QTL mapping for the clinical scores in the N2 progeny displaying *Zsl11*, *Zsl13*, *Zsl14* and five suggestive QTLs.

(A) QTL mapping was performed on the clinical scores at 5, 6, 7, 8, 9, 12 and 14 dpi. Only the chromosomes with suggestive or significant associations were represented. Horizontal lines indicate genome-wide significance thresholds determined by permutation testing ( $n=1000$ ). For better visibility, and as the significant thresholds are highly similar for the clinical scores at different time points ( $p = 0.05$ , 3.32 for 5 dpi, 3.12 for 6 dpi, 3.09 for 7 dpi, 3.25 at 8 dpi, 3.09 at 9 dpi, 3.17 at 12 dpi, 3.01 at 14 dpi,  $p = 0.63$ , 1.81 for 5 dpi, 1.79 for 6 dpi, 1.81 for 7 dpi, 1.82 at 8 dpi, 1.75 at 9 dpi, 1.79 at 12 dpi, 1.72 at 14 dpi), only the minimum threshold was displayed. Vertical dotted lines represent the positions of the markers with the highest LOD scores used to display the LOD score per day in (B) and the allelic effects in Figure S13. (B) LOD score for each QTL identified at each time post infection tested. CS = clinical score.

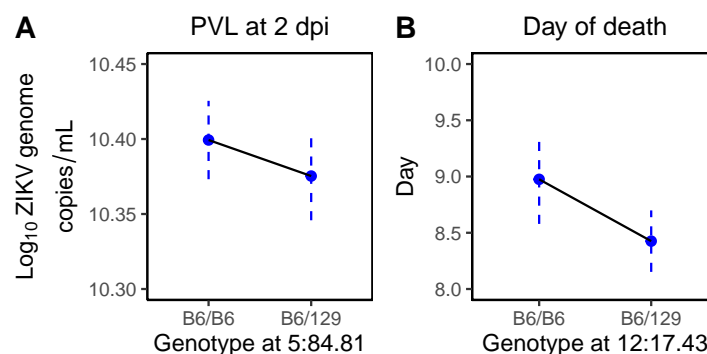
In addition to these three significant QTLs, five suggestive QTLs were obtained for the clinical scores on chromosome 1, 2, 5, 8, and 15 (Figure 51A and Table 4).

The Figure 51B shows that some QTLs are involved in the clinical scores at early time points, between 5 and 9 dpi (*Zsl11*, *Zsl13*, *Zsl14* and the QTLs on chromosomes 2 and 8), while others are associated with the scores later after infection (QTLs on chromosomes 1, 5 and 15). For *Zsl11* and the QTLs on chromosomes 2 and X, the LOD score was high only for one specific dpi, while it is not likely that a locus impacts the scores of N2 mice only at a one precise day post infection, as mentioned in the section V.2.2.10.

### V.3.2.7 . Comparison of the QTLs found in the F2 and N2 crosses

In the analysis of B6-*Ifnar1* × 129-*Ifnar1* F2 mice, two significant QTLs were identified. The first QTL on chromosome 5 was associated with the PVL at 2 dpi. Mice with the B6/B6 genotype had lower PVL than mice with the B6/129 and 129/129 genotypes. The second QTL was located on chromosome 12 and was associated with the day of death. The B6/B6 genotype was associated with earlier time to death compared with the B6/129 and 129/129 genotypes.

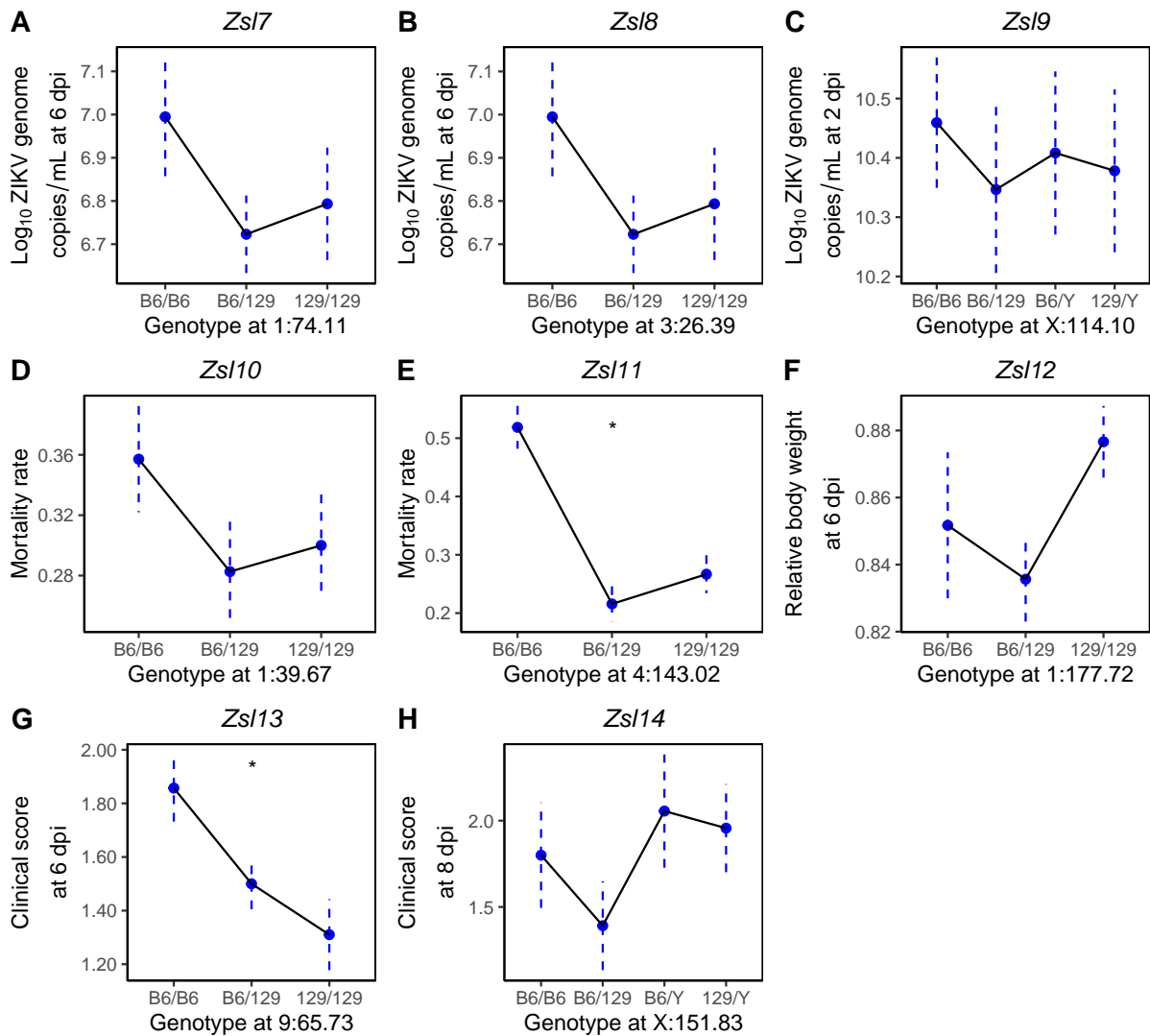
In the N2 progeny studied here, no LOD score peak were found at these positions on chromosomes 5 and 12 for PVL at 2 dpi, day of death, or any other phenotype studied (Figures 42, 46, 50, and 51). Indeed, B6/B6 and B6/129 mice for the peak position of the chromosome 5 QTL had the same average PVL at 2 dpi (Figure 52A). Similarly, B6/B6 and B6/129 mice for the peak position of the chromosome 12 QTL had the same average day of death (Figure 52B). Thus, the QTLs found in the B6-*Ifnar1* × 129-*Ifnar1* F2 progeny were not observed in the N2 progeny.



**Figure 52:** Allelic effects for the QTLs identified in the F2 progeny on the phenotypes of the N2 progeny. Allelic effects were assessed at the marker with the highest LOD score in the analysis of the F2 progeny: mUNC050135252 (A) and gUNCHS032900 (B). X-axis indicates the genotype of the SNP marker. Results are represented as mean ± sem. Groups were compared by t test and results were not significant.

We then looked at the allelic effects for the significant QTLs found in this study on the phenotypes observed in the F2 progeny (Figure 53). For instance, *Zs17* was associated with the PVL at 6 dpi of N2 mice, and B6/B6 mice had higher PVL than B6/129 mice (Figure 43). In F2s, this allelic effect was also observed, but the result of the statistical test was not significant (Figure 53A). Similar effects were obtained for *Zs18* and *Zs10* (Figures 53B and D). For *Zs11* and *Zs13*, the comparison between B6/B6 and B6/129 mice was also significant in F2s (Figures 53E and G). Thus, *Zs17*, *Zs18*, *Zs10*, *Zs11* and *Zs13* show consistent allelic effects between the F2 and N2 progenies (Table 5).

Conversely, for *Zs12*, the B6/B6 genotype was associated with lower body weight in N2s (Figure S12B), while in F2s, no significant differences were obtained between B6/B6 and B6/129 mice, and B6/B6 tended to have higher weights (Figure 53F). Similarly, the two chromosome X QTLs, *Zs19* and *Zs14* did not show consistent allelic effects between the N2 and the F2 progeny. For *Zs19*, B6/B6 N2 females had lower PVL at 2 dpi than B6/129 N2 females (Figure 45B). In F2s, no significant differences were obtained between the genotype groups (Figure 53C). Similarly, for *Zs14*, B6/B6 N2 females showed lower clinical scores than B6/129 N2 females (Figure S13H). This effect was not found in F2 mice, as B6/B6 females showed slightly higher values of clinical scores, although this difference was not significant (Figure 53H).



**Figure 53:** Allelic effects for the QTLs identified in the N2 progeny on the phenotypes of the F2 progeny. Allelic effects were assessed for the significant QTLs found in the N2 progeny at the marker with the highest LOD score (see Figures 42 to 51). X-axis indicates the genotype of the SNP marker. Y-axis displays the phenotypic quantification as mean  $\pm$  sem. Groups were compared by ANOVA (A - C, F - H). In D and E, significances were assessed for each sex/age group and the same results were obtained.

### V.3.3. Discussion

*Ifnar1* deficient mice are a common model of ZIKV infection and have been used for instance to study of the virus pathogenesis or the development of treatments (Nazneen et al. 2023; Zhu et al. 2023; Pattnaik et al. 2023; Mancini et al. 2023). While this KO exists on two different genetic backgrounds, the impact of this factor on the study of ZIKV in mouse models is rarely discussed (Bradley and Nagamine 2017; Alves dos Santos and Fink 2018; Li et al. 2023). Yet, genetic background is a major contributor of the infection outcome in mice (Manet et al. 2020).

Here, we investigated the genetic control of the susceptibility of *Ifnar1* deficient mice to ZIKV infection by producing and analyzing a (B6-*Ifnar1*  $\times$  129-*Ifnar1*)  $\times$  129-*Ifnar1* N2. Eight significant and nineteen suggestive QTLs were obtained (Figure 54). This large number of QTLs obtained demonstrate that analyzing a backcross allowed to increase the power to detect loci involved in the susceptibility of B6-*Ifnar1* mice.

**Table 5:** Effects of the QTLs found in the F2 and N2 crosses

QTL	Phenotype	Chr	Position	Effect in F2s	Effect in N2s
-	PVL (2 dpi)	5	84.81	B6/B6 < B6/129 = 129/129	B6/B6 = B6/129
-	Day of death	12	17.43	B6/B6 < B6/129 = 129/129	B6/B6 = 129/129
Zs/7	PVL (6 dpi)	1	74.11	B6/B6 > B6/129 = 129/129	B6/B6 > B6/129
Zs/8	PVL (6 dpi)	3	26.39	B6/B6 > B6/129 = 129/129	B6/B6 > B6/129
Zs/9	PVL (2 dpi)	X	114.10	B6/B6 = B6/129 & B6/Y = 129/Y	B6/B6 < B6/129 & B6/Y = 129/Y
Zs/10	Mortality	1	39.67	B6/B6 > B6/129 = 129/129	B6/B6 > B6/129
Zs/11	Mortality	4	143.02	B6/B6 > B6/129 = 129/129	B6/B6 > B6/129
Zs/12	BW (6 dpi)	1	177.72	B6/B6 = B6/129 < 129/129	B6/B6 < B6/129
Zs/13	CS (6 dpi)	9	65.74	B6/B6 > B6/129 > 129/129	B6/B6 > B6/129
Zs/14	CS (8 dpi)	X	151.83	B6/B6 > B6/129 & B6/Y = 129/Y	B6/B6 < B6/129 & B6/Y = 129/Y

The first trait for which B6-*Ifnar1* and 129-*Ifnar1* mice differ is the PVL at 6 dpi. Thus, QTLs associated with the PVL of N2 mice were investigated. As expected, most QTLs for the PVL were obtained at 6 dpi. For all these QTLs, the B6 allele was associated with higher PVL. Thus, we have identified multiple QTLs that may be involved in the high PVL observed in B6-*Ifnar1* mice at 6 dpi.

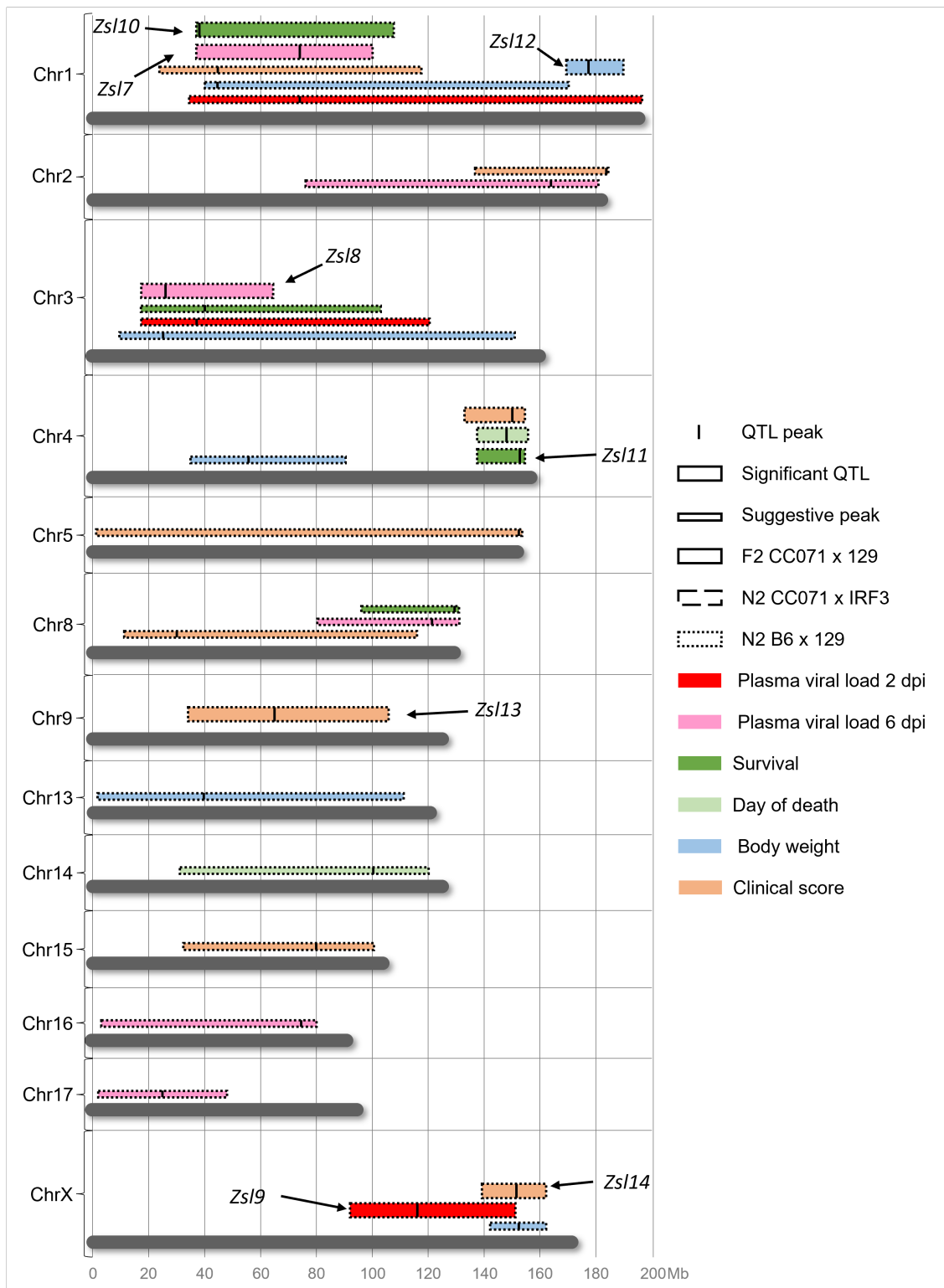
The QTLs associated with the PVL at 6 dpi also showed suggestive associations with the PVL at 2 dpi. This result suggests that the mechanism responsible for the high PVL of B6-*Ifnar1* at 6 dpi also influences the PVL at 2 dpi, although the suggestive associations for the PVL at 2 dpi could only be a consequence of the small correlation between the PVLs at 2 and 6 dpi.

The second aspect of ZIKV-disease which differs between B6-*Ifnar1* and 129-*Ifnar1* mice is the severity of the illness observed in infected mice, as measured by the body weight loss, clinical scores and mortality. A large variability of these symptoms were obtained in the N2 progeny, which allowed to map QTLs controlling these traits.

For the survival and day of death, two significant QTLs were obtained, for which the B6 allele was the susceptibility allele, and for which the difference of mortality rate between B6/129 and B6/B6 could be twofold. Thus, this analysis allowed to find QTLs which greatly contribute to the probability of survival in the N2 progeny. Here also, as the B6 allele was recessive, the variants underlying these QTLs are likely to be loss-of-function.

For the illness of N2 mice, transgressive QTLs, i.e., for which the illness of B6/B6 mice was less severe than that of B6/129 mice, were also obtained. While these loci may not contribute to the severity of the symptoms of B6-*Ifnar1* mice after infection with ZIKV, they may contribute to the differences of infection outcome between B6-*Ifnar1* and 129-*Ifnar1* mice.

Analysis of the QTLs identified in this cross led to colocalizing QTLs associated with multiple traits, that we can name "heterologous" QTLs (Figure 54). For instance, two significant QTLs were obtained on proximal chromosome 1, *Zs/7* controlling the PVL at 6 dpi, and *Zs/10*, controlling the survival of N2 mice (Figure 54). These two traits were correlated in the N2 progeny (Figure 41E and G), but this correlation does not demonstrate that the two traits are controlled by the same genes. Thus, these two QTLs were given different *Zs/* numbers until it is confirmed that they are in fact the same QTL.



**Figure 54:** Summary of the QTLs found in the N2 progeny

The colored bars represent the Bayesian credible intervals of the QTLs. The QTL interval color represent the phenotype associated with the QTL. Thick bars represent significant QTLs while thin bars represent suggestive QTLs. Vertical bars represent the position of the peak LOD score of each QTL. Arrows indicate the names of significant QTLs.

*Zs17* and *Zs110* also colocalized with suggestive QTLs on chromosome 1, associated with the PVL at 2 dpi, the body weights and the clinical scores of N2 mice (Figure 54). The allelic effects of all these QTLs were concordant, as B6/B6 mice had higher PVL, mortality rate, body weight loss and clinical scores as B6/129 mice (Figure 43B - C, 47B, 50 and 51). Therefore, it is possible that the genes responsible for *Zs17* and *Zs110* are involved in all the phenotypes studied in the N2 progeny (i.e., the PVLs, survival, body weights and clinical scores). If it is confirmed that these QTLs are actually one pleiotropic QTL, it would suggest a single mechanism is responsible for a high PVL and severe illness following infection.

As genetic analysis was performed in both an F2 and an N2 between B6-*Ifnar1* and 129-*Ifnar1* mice, results of the two crosses were compared. The identification of a QTL in the two crosses would strengthen the hypothesis of its role in the difference of susceptibility of B6-*Ifnar1* and 129-*Ifnar1* mice. In the F2 study, 192 mice were infected but only 92 were genotyped for QTL mapping. Thus, this cross had less power to detect QTLs. The remaining F2 mice will be genotyped and added to the QTL mapping analysis.

The two QTLs detected in the F2 cross and three of the QTLs identified here were not retrieved in the N2 or in the F2, respectively (Table 5). Conversely, five QTLs, *Zs17*, *Zs18*, *Zs110*, *Zs111* and *Zs113*, identified in the N2 progeny showed consistent allelic effects in the F2s. Thus, investigating the candidate genes underlying these five QTLs may identify modifier genes of the *Ifnar1* KO.

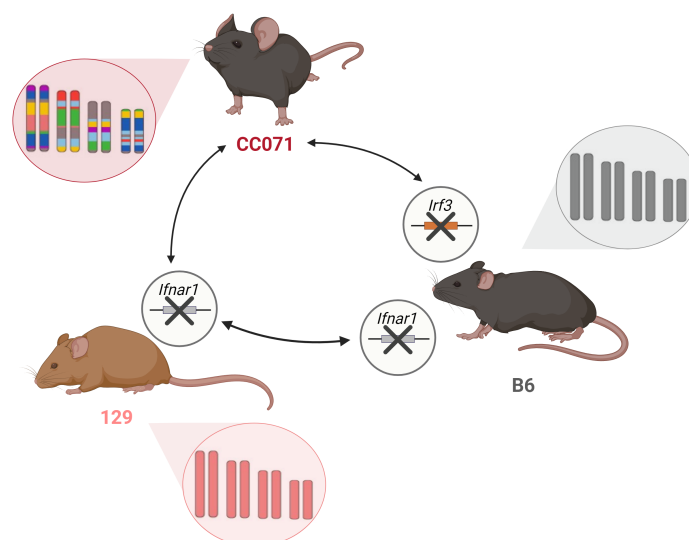
In conclusion, my results showed that several loci are involved in the differences of susceptibility of B6-*Ifnar1* and 129-*Ifnar1* mice. QTL mapping in the N2 progeny allowed to address both the differences of PVL at 6 dpi and the differences of clinical manifestations between these two mouse strains. As the credible intervals of these QTLs remain large, we will focus on the QTLs with the highest LOD scores and the narrowest intervals, and which showed consistent allelic effects between the F2s and the N2s, such as *Zs111* on chromosome 4 (Figure 54). Future validation studies will shed light on the mechanisms responsible for the increased susceptibility of B6-*Ifnar1* mice, providing new clues on how to use *Ifnar1* deficient strains to better model ZIKV disease in animal research.

## V.4 - Meta-analysis of the three QTL mappings

### V.4.1 . Relevance of a meta-analysis

In my objectives, I presented a figure with the three segregating crosses that I planned to analyze during my PhD, one cross which was analyzed *in vitro*, between CC001 and CC071, and two crosses for which SGI mice were produced and analyzed *in vivo* (Figure 17). These two crosses share one parental strain, 129-*Ifnar1*, which, on one hand, was crossed with CC071 to obtain a CC071  $\times$  129-*Ifnar1* F2 progeny, and on the other hand, was crossed with B6-*Ifnar1* to obtain a (B6-*Ifnar1*  $\times$  129-*Ifnar1*)  $\times$  B6-*Ifnar1* N2 progeny.

The identification of the *Irf3* loss-of-function mutation in CC071 led us to produce and analyze a third progeny of (CC071  $\times$  B6-*Irf3*)  $\times$  CC071 N2. Therefore, we crossed mice of three different genetic backgrounds in pairs, as represented in Figure 55, setting aside the different KOs that these mice carry. This circular breeding scheme is favorable for discovering common QTLs between these crosses.



**Figure 55:** Segregating crosses analyzed *in vivo* in my PhD project.

Arrows represent the 2-generation crosses made between inbred strains. Bubbles contain the chromosomes of each strain with colors representing the CC ancestral haplotypes: (A/J): yellow, B6: grey, 129S1/SvImJ: pink, NOD/ShiLtJ: dark blue, NZO/HiLtJ: light blue, CAST/EiJ: green, PWK/PhJ: red, WSB/EiJ: purple). The 129-*Ifnar1* strain has a 129S2/SvPas genetic background which is different from 129S1/SvImJ, but was represented with the same pink color.

The possibility to identify common QTLs is influenced by several criteria. First, QTL mapping in a segregating cross between two inbred strains only allows to identify QTLs if different alleles for the underlying gene are found between the parental strains. For instance, the strain A must have a different allele than the strain B for a susceptibility gene to be able to see the effect of this gene in a  $A \times B$  F2. When studying several crosses, it is possible that this condition is satisfied in one cross, but not in the other. If strain A has the same allele as a strain C, the QTL would be obtained in the  $A \times B$  F2, but not in the  $A \times C$  N2.



Then, common QTLs are more likely if they are associated with the same trait. Therefore, identification of such homologous QTLs requires that the same phenotypes are measured in the crosses studied, for instance the survival in the A × B F2 and in the A × C N2. Another, less likely, situation, is the identification of multitrait QTLs in the same cross, or in two crosses. For instance, a QTL associated with the body weights of A × B F2 mice could be associated with the clinical score of the A × C N2 mice.

The experimental model also influences the genetic control of the susceptibility of mice to the infection. In the case of ZIKV infection, two modalities of inhibition of the IFN-I pathway can be used. If A × B F2 mice carry an *Ifnar1* deficiency, while A × C N2 were treated with the IFNAR1-blocking antibody, different genetic factors could be involved in the susceptibility of mice of each cross.

Lastly, the ability to detect common QTLs depends on the genotypes present in each cross, and on the allelic effects of the QTLs. If we consider a A × B F2, mice can have the AA, AB and BB genotypes at each locus in the genome. (A × C) × C N2 mice can only carry the AC and CC genotypes. If a QTL is obtained in the A × B F2, for which the A allele is the susceptibility allele and is recessive, we have AA < AB = BB. In the (A × C) × C N2, the AA genotype is absent, while AC and CC will show the same phenotypes as the A allele is recessive. Thus, this QTL will be obtained in the A × B F2 but not in the (A × C) × C N2.

#### V.4.2 . Opportunities to find common QTLs in the three crosses

##### V.4.2.1 . Study of common traits

The exact same phenotypes were studied in F2-*Ifnar1* and (B6-*Ifnar1* × 129-*Ifnar1*) × B6-*Ifnar1* N2s (hereinafter designated as N2-B6.129-*Ifnar1*), which were the survival, PVLs at 2 and 6 dpi, body weights and clinical scores from 5 to 14 dpi. Some of these phenotypes were also studied in N2-*Irf3*, namely the PVLs, body weights and clinical scores until 7 dpi (Table 6). By contrast, the body weights and clinical scores after 7 dpi were not investigated in N2-*Irf3* mice, and the BVL was only investigated in N2-*Irf3* mice.

**Table 6:** Phenotypes studied in the three crosses

Phenotypes	Survival	PVL	BVL	BW		CS	
				≤ 7 dpi	> 7 dpi	≤ 7 dpi	> 7 dpi
CC071 × B6- <i>Irf3</i> N2s		×	×	×		×	
CC071 × 129- <i>Ifnar1</i> F2	×	×		×	×	×	×
B6- <i>Ifnar1</i> × 129- <i>Ifnar1</i> N2	×	×		×	×	×	×

BVL: brain viral load, BW: body weight, CS: clinical score, PVL: plasma viral load.

#### **V.4.2.2 . Influence of the experimental model**

Two different experimental models of ZIKV infection were used in these three crosses. N2-B6.129-*Ifnar1* and F2-*Ifnar1* mice were all homozygous for the *Ifnar1* KO. Conversely, N2-*Irf3* do not carry this deficiency, and were treated with MAR1-5A3 antibody to block the IFN-I response and allow for efficient viral replication. This model was less severe than the *Ifnar1* deficiency, as N2-*Irf3* mice had lower clinical scores and body weight loss than F2-*Ifnar1* and did not succumb.

The susceptibility of antibody treated and *Ifnar1* mice could result from different mechanisms. The IFNAR1 blockade is not as complete in antibody-treated mice due to the presence of sites such as the brain to which the antibody cannot easily diffuse (Iwasaki 2017). Thus, variants in gene of the IFN-I pathway could impact the immune response in these sites and modify the susceptibility of antibody treated mice, but not of *Ifnar1* deficient mice.

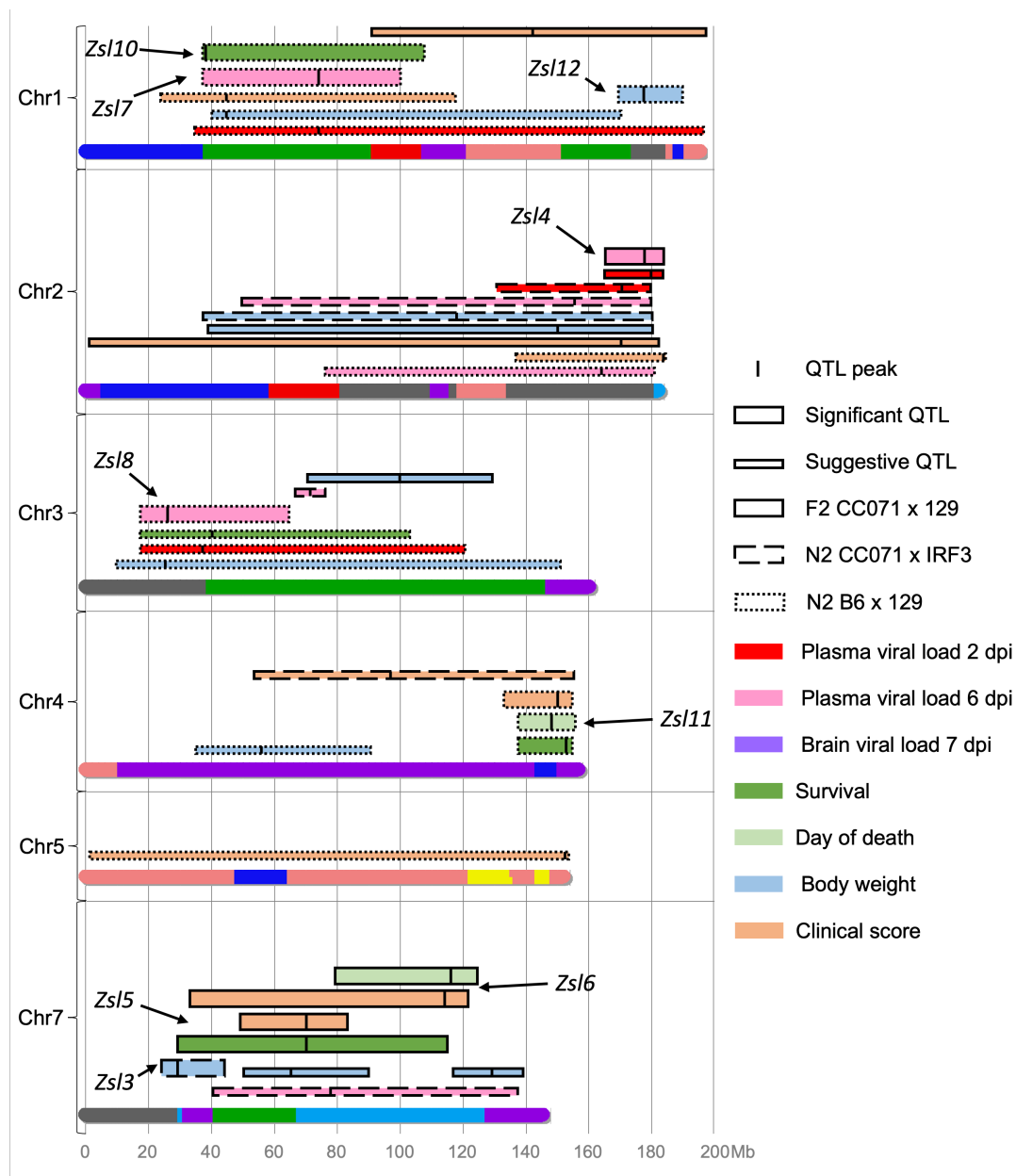
Moreover, it is possible that the association with susceptibility alleles with small effects could only be observed in a severe model, such as in *Ifnar1* deficient mice but not in antibody treated mice. Therefore, different QTLs may be identified in N2-B6.129-*Ifnar1* and F2-*Ifnar1* on one side, and N2-*Irf3* on the other side.

#### **V.4.2.3 . Investigation of different susceptibility alleles**

The three crosses were not used to meet a single objective. The crosses involving CC071, e.g., N2-*Irf3* and F2-*Ifnar1*, were used to identify CC071 susceptibility alleles. By contrast, the goal of the N2-B6.129-*Ifnar1* was to identify alleles explaining the difference of severity of ZIKV disease between the two B6-*Ifnar1* and 129-*Ifnar1*. Common QTLs in the two crosses with CC071 as a parental strain are likely as they involve the same susceptibility alleles. Otherwise, as the susceptibility of CC071 mice and B6-*Ifnar1* may not be controlled by the same genes, it was less likely to obtain the same QTLs between the N2-B6.129-*Ifnar1* and either the N2-*Irf3* or the F2-*Ifnar1*.

The exceptions to this line of reasoning are the genomic regions for which CC071 inherited the B6 haplotype. Indeed, due to the design used to create the CC strains, about one eighth of the genome of CC071 originates from B6 (as represented in Figure S2C). It is possible that part of the susceptibility of CC071 mice is explained by alleles in these regions, which are in fact B6 alleles. In this case, common QTLs would be obtained between the F2-*Ifnar1* and the N2-B6.129-*Ifnar1* crosses.

B6 was a parental strain of the N2-B6.129-*Ifnar1* and the N2-*Irf3* crosses (setting aside the different KOs in the two crosses). To produce the N2-B6.129-*Ifnar1* progeny, B6-*Ifnar1* × 129-*Ifnar1* F1 mice were crossed with B6-*Ifnar1*, thus, mice of this cross carry either the B6/B6 or the B6/129 genotype. To produce the N2-*Irf3* progeny, CC071 × B6-*Irf3* F1 mice were crossed with CC071, thus, in this progeny, we find the CC071/CC071 and CC071/B6 genotypes (Table 7). Thus, the N2-B6.129-*Ifnar1* cross only allows to identify QTLs for which the B6 allele is recessive or semidominant, i.e., for which a phenotypic difference between B6/B6 and B6/129 mice. If the B6 allele is recessive, the effect of the QTL could not be detected in the N2-*Irf3*, as there would be no phenotypic difference between mice carrying the CC071/CC071 genotype and mice carrying the CC071/B6 genotype.



**Figure 56:** Summary of the QTLs found in the three *in vivo* crosses

As the bottom of each panel is represented the chromosome of the CC071 strain with colors corresponding to haplotypes inherited by CC071 mice (A/J: yellow, B6: grey, 129S1/Svlmj: pink, NOD/ShiLtj: dark blue, NZO/HILtj: light blue, CAST/Eij: green, PWK/Phj: red, WSB/Eij: purple). The colored bars represent the Bayesian credible intervals of the QTLs. The QTL interval color represent the phenotype associated with the QTL. The outline represents the cross in which the QTL was identified. Thick bars represent significant QTLs while thin bars represent suggestive QTLs. Vertical bars represent the position of the peak LOD score of each QTL. Arrows indicate the names of significant QTLs.

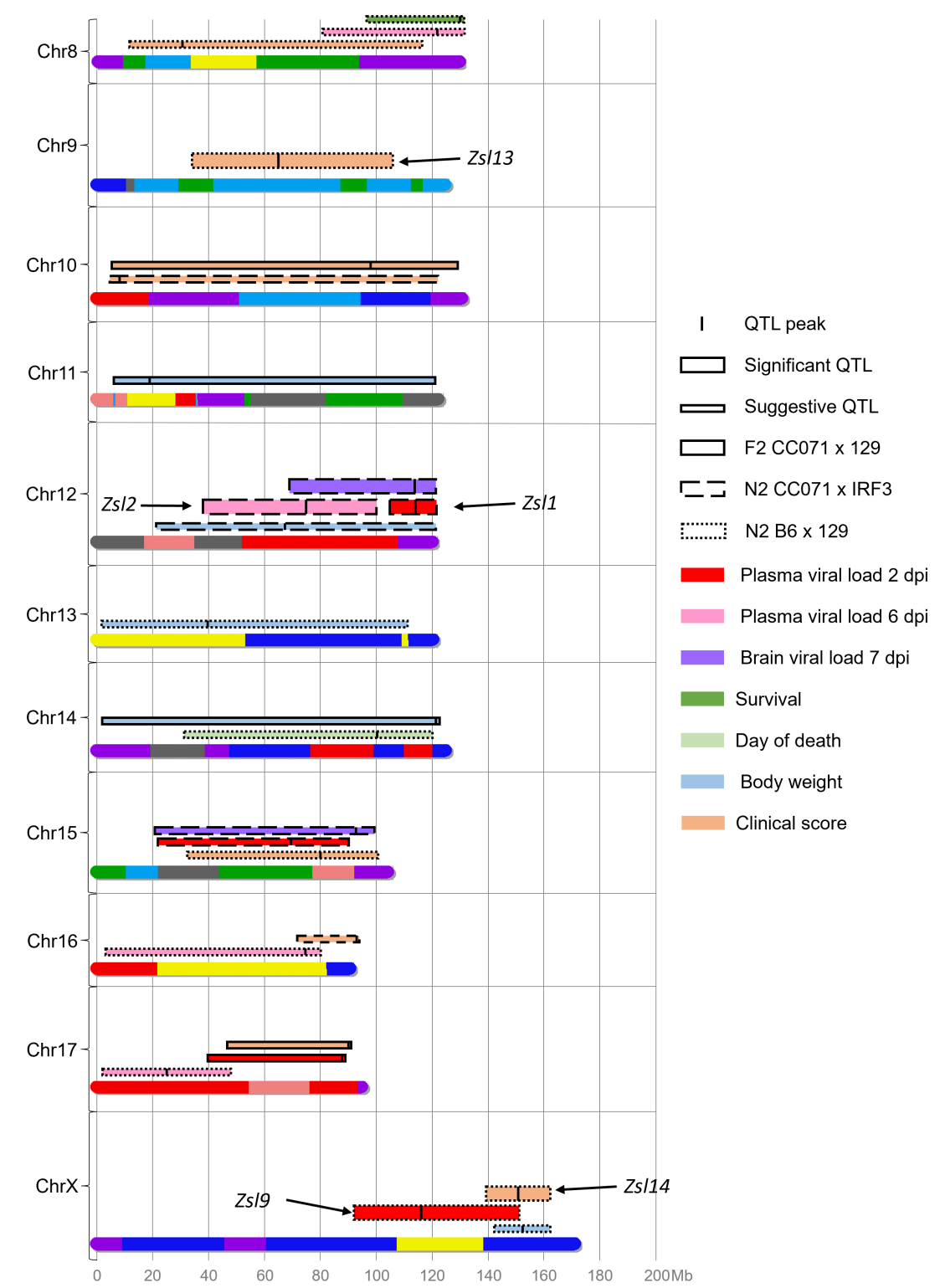


Figure 56: (cont.)

**Table 7:** Genotypes found in the three crosses

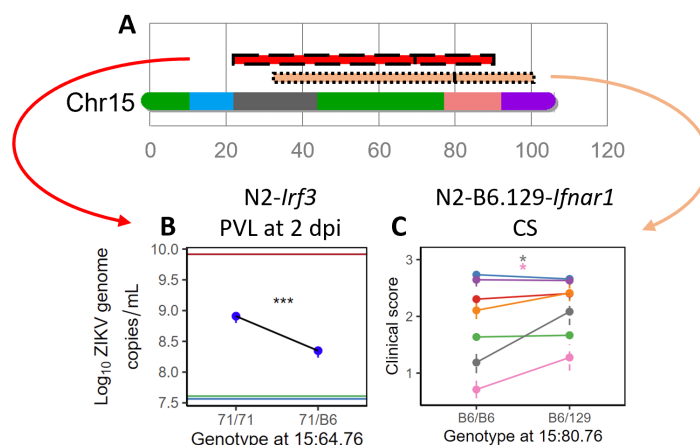
Genotypes	71/71	71/B6	71/129	129/129	B6/B6	B6/129
CC071 × B6- <i>Irf3</i> N2	×	×				
CC071 × 129- <i>Ifnar1</i> F2	×		×	×		
B6- <i>Ifnar1</i> × 129- <i>Ifnar1</i> N2					×	×

### V.4.3 . Analysis of the colocalization of QTLs between the three crosses

Over the three crosses, I identified 14 significant QTLs and more than 40 suggestive QTLs on 17 chromosomes, for all the phenotypes studies, e.g., PVL at 2 and 6 dpi, BVL, survival, day of death, body weight, and clinical score (Figure 56). No significant QTLs was detected for the body weights and the clinical scores at late time points (between 9 and 14 dpi). This could have been expected as most susceptible mice died before 9 dpi, thus, the number of mice with values of body weight and clinical score after 9 dpi was reduced and perhaps not sufficient to obtain significant associations.

Colocalizing QTLs were obtained between the N2-*Irf3* and F2-*Ifnar1* progenies (Figure 56). Notably, on chromosome 2, a significant QTL was associated with the PVL at 2 dpi of F2-*Ifnar1* mice and colocalized with a suggestive QTL associated with the PVL of N2-*Irf3* mice. More detailed analysis of the colocalization of the QTLs found in these two crosses was previously reported in the section V.2.3.

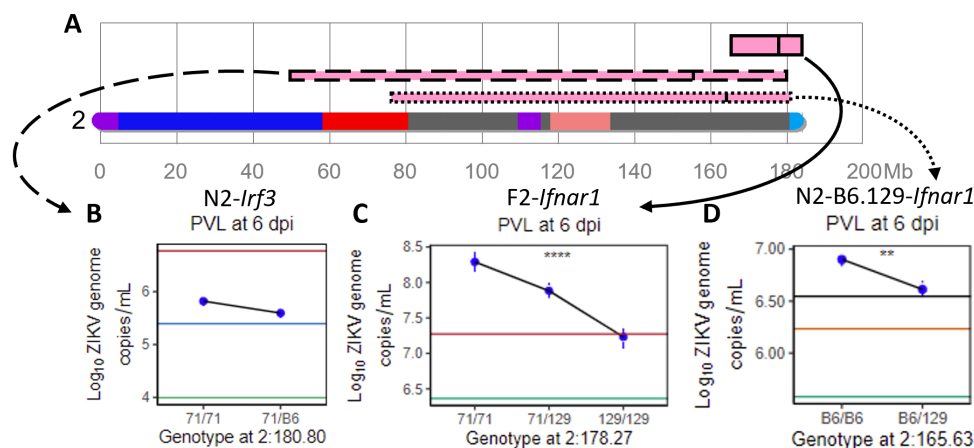
Here, I will focus on common QTLs between the N2-B6.129-*Ifnar1* cross and either the N2-*Irf3* or the F2-*Ifnar1* cross. For instance, on chromosome 15, one suggestive QTL was associated with the PVL at 2 dpi of N2-*Irf3* mice, and a suggestive QTL was associated with the clinical scores of N2-B6.129-*Ifnar1* mice (Figure 57A). It is unlikely that a single pleiotropic QTL is responsible for these two associations. First, the B6 allele was dominant for the QTL found in the N2-*Irf3* progeny, but recessive for the QTL found in the N2-B6.129-*Ifnar1* progeny (Figure 57B - C). Moreover, the PVL and the body weight are likely to be controlled by different genes. Thus, it is unlikely that these two QTLs are controlled by the same gene.

**Figure 57:** Colocalizing QTLs on chromosome 15

(A) Example of QTLs found on chromosome 15 on the N2-*Irf3* and N2-B6.129-*Ifnar1* progenies, as in Figure 56. (B) Allelic effect for the QTL associated with the PVL in the N2-*Irf3* progeny, from Figure 23. (C) Allelic effect for the QTL associated with the clinical score in the N2-B6.129-*Ifnar1* progeny, from Figure S13.

I also found colocalized QTLs between the N2-B6.129-*Ifnar1* and the F2-*Ifnar1* crosses on chromosomes 1, 3, 14 and 17, but the maximum LOD score was not localized in the same chromosomal region in the two crosses (Figure 56). Furthermore, it was more likely for the same allele to be responsible for a QTL in these two crosses if they were located in regions where CC071 inherited the B6 haplotype, which was not the case for these QTLs. Therefore, these associations are unlikely to be driven by the same genes in the two crosses.

On the distal region of chromosome 2, I found QTLs controlling the PVL in the two crosses involving CC071, N2-*Irf3* and F2-*Ifnar1* (Figure 58). I previously hypothesized that the same susceptibility allele from CC071 was responsible for these QTLs (see section V.2.3). As this QTL was found in a cross between CC071 and B6-*Irf3*, it was more likely that this susceptibility allele was located in the region where CC071 inherited the NZO/HILtj haplotype and not in the region where the strain inherited the B6 haplotype. I also identified a suggestive QTL on the distal region of chromosome 2 controlling the PVL of N2-B6.129-*Ifnar1* mice (Figure 58). Therefore, it is also possible that the QTL observed in F2-*Ifnar1* mice results from a susceptibility allele in the region where CC071 inherited the B6 haplotype, and that the same allele is responsible for the QTL found in N2-B6.129-*Ifnar1*. Indeed, the allelic effects are consistent between the two crosses as the CC071/CC071 and the B6/B6 genotypes are associated with higher PVL in the F2-*Ifnar1* cross and N2-B6.129-*Ifnar1* cross, respectively (Figure 58).



**Figure 58:** Colocalizing QTLs controlling the PVL on chromosome 2

(A) Example of QTLs found on chromosome 2 on the N2-*Irf3*, F2-*Ifnar1* and N2-B6.129-*Ifnar1* progenies, as in Figure 56. (B) Allelic effect for the QTL associated N2-*Irf3* progeny, from Figure 22. (C) Allelic effect for the QTL associated F2 progeny, from Figure 33. (D) Allelic effect for the QTL in the N2-B6.129-*Ifnar1* progeny, from Figure S10.

Two hypotheses can be established to explain this situation. First, there are two different susceptibility alleles in this region involved in the high PVL of CC071 mice. The first one is located in the NZO/HILtj haplotype, and leads to a QTL in the N2-*Irf3* and in the F2-*Ifnar1* progenies. The second allele is inherited from B6 and is located in the B6 haplotype, leads to a QTL in the F2-*Ifnar1* progeny, but could not lead to a QTL in the N2-*Irf3* progeny as there would not be polymorphism between B6-*Irf3* and CC071. Besides, as this susceptibility allele is a B6 allele, it would also lead to a QTL in the N2-B6.129-*Ifnar1* progeny.

As a result, it is likely that the QTLs on chromosome 2 are common to F2-*Ifnar1* and N2-B6.129-*Ifnar1* crosses. Thus, combine cross analyses could be used to increase the power of the QTL mapping and reduce the size of the QTL interval, which would facilitate the search for candidate genes (see section II.3.2.3).

The second hypothesis is based on the observation that the QTL identified on chromosome 2 in the CC071 × B6-*Irf3* N2 results from interval mapping, as no polymorphic marker between CC071 and B6-*Irf3* was found in the region (Figure 22). Thus, this QTL must be confirmed by future investigation. If this QTL is a false positive association, there would be only one susceptibility allele in this region, responsible for the QTLs identified in the F2-*Ifnar1* and in the N2-B6.129-*Ifnar1* progenies.

#### V.4.4 . Refinement of QTLs credible intervals

The analysis of the QTL map presented in Figure 56 can help refine the credible intervals of the different QTLs identified in the three crosses. First, in the N2-*Irf3* cross, it is unlikely to identify QTLs in regions where CC071 inherited the B6 haplotype, as these regions are supposedly identical between CC071 and B6-*Irf3*. The exception is the possible but rare occurrence of private variants carried by CC071 in these regions. Therefore, the search for candidate genes focuses on regions where CC071 carries a non-B6 haplotype. For instance, *Zsl3* was associated with the body weight of N2-*Irf3* mice and was located on chromosome 7 in a region where CC071 inherited haplotypes from B6, NZO/HILtJ, WSB/Eij and CAST/Eij (Figure 56). The same strategy applies for the two QTLs on chromosome 12 and the QTL on chromosome 15 controlling the PVL at 6 dpi, the body weight and the PVL at 2 dpi, respectively (Figure 56).

The QTLs identified in the N2-B6.129-*Ifnar1* progeny would be also found in the F2-*Ifnar1* cross if they resulted from susceptibility alleles located in regions where CC071 inherited the B6 haplotype. For instance, the credible intervals of the QTLs found on chromosomes 3 and 15 in the N2-B6.129-*Ifnar1* cross contain such regions and do not colocalize with QTLs identified in the F2-*Ifnar1* cross, therefore the susceptibility alleles are likely located in the non-B6 regions. The reciprocal reasoning can be applied to the QTLs identified in the F2-*Ifnar1* cross. For instance, *Zsl5* and *Zsl6* on chromosome 7 control the survival and the day of death of F2-*Ifnar1* mice, respectively. As no QTL was found in the N2-B6.129-*Ifnar1* cross on this chromosome, the susceptibility alleles of CC071 is likely outside the B6 region.

#### V.4.5 . Candidate gene analysis

Each of the QTLs identified in this study must be investigated further to identify candidate genes which role in the susceptibility to ZIKV disease can be functionally validated. I will detail the analysis of one of the QTLs identified in the N2-*Irf3* cross.

*Zsl1* was found on chromosome 12 and was associated with the PVL at 2 dpi of N2-*Irf3* mice. The peak of this QTL was located distally on the chromosome, and its credible interval spanned between 105 Mb and the end of the chromosome. This interval contains 108 protein coding genes according to the mouse genome informatics (MGI) database (Table S1).

To filter this list of candidate genes, I have used three criteria. First, I investigated sequence differences between B6 and CC071 in the QTL interval using GenomeMUSter (Ball et al. 2023). As a high number of polymorphisms were obtained, I focused on those affecting the coding sequence and predicted as "moderate" or "deleterious" by the Ensembl Variant Effect Predictor (<https://www.ensembl.org/info/docs/tools/vep/>). This filter is not infallible as polymorphisms in the regulatory sequences of a gene can impact its expression.

Second, expression levels of genes in the QTL interval were investigated. I took advantage of the RNA-seq data from B6 and CC071 MEFs. This data allowed to investigate the expression of genes in B6 and in CC071 cells in the same experimental conditions, in non-infected and ZIKV infected cells. The limitation of this analysis is that the MEF model does not represent the complexity of the response to the infection at the organism level, and that some genes are not expressed in MEFs. Other resources can be used, for instance the transcriptomes of heart, liver, and kidney tissues of CC strains at the basal state were investigated by Zhang and colleagues (Zhang et al. 2023). Expression levels of CC071 mice could be compared with expression data of B6 mice from other studies (Song et al. 2012), but this would require more advanced bioinformatic analyses. Besides, genes could have similar expression levels between B6 and CC071 and still be involved in the difference of susceptibility of these two strains. For instance, B6 and CC071 MEFs had similar levels of expression of *Irf3* (Figure S3G in section V.1), while this gene was responsible for the delayed IFN-I and high viral replication in CC071 MEFs.

Lastly, known functions of the genes located in the QTL interval were assessed. As we studied susceptibility to a viral infection, it is likely that the gene responsible for the association of the PVL with *Zs1* possesses immune functions. Therefore, the list of candidates was filtered using the "abnormal immune system physiology" (Mouse Phenotype MP:0001790) criterion on the MGI database. This criterion allowed to filter genes for which a mutant mouse showed an altered immune phenotype. It is also possible that a gene involved in differences of PVL between B6-*Irf3* and CC071 does not have a known function in immunity or susceptibility to infections.

Each of the three criteria is not adamant, as they all have some limitations discussed above. They were used to facilitate the search for candidate genes for *Zs1*, and the genes matching the highest number of criteria will be investigated in priority.

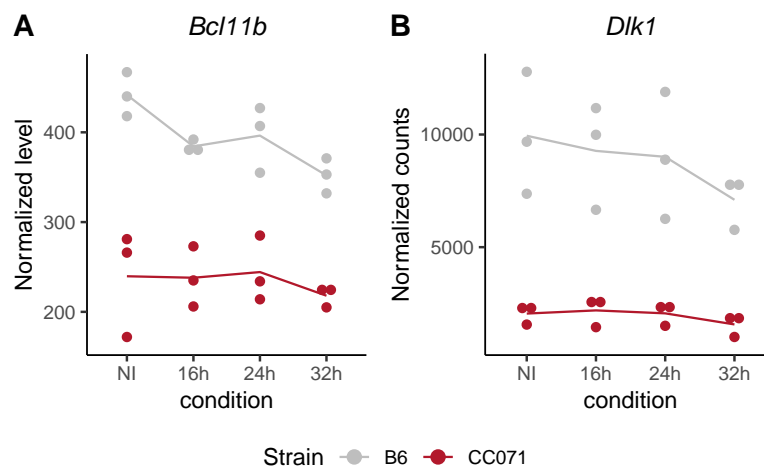
Allelic data of B6 and CC071 in this region showed that among the 108 genes, 58 contained polymorphisms between the two strains locating in the coding sequence of genes, including 23 with variants in CC071 predicted as "deleterious" (Table S1). Second, the levels of expression of the 108 genes in MEFs showed that among the 81 genes expressed in MEFs, 5 were differentially expressed between B6 and CC071. CC071 MEFs showed lower expression levels of *Ahnak2*, *Bcl11b* and *Dlk1*, and higher expression levels of *Atg2b* and *Hhip1* (Table S1). Lastly, 13 genes were associated with an "abnormal immune system physiology" (Table S1).

None of the 108 genes showed simultaneously a deleterious variant in CC071, a different expression level in CC071 MEFs compared with B6 MEFs and has a known immune function. Thus, genes with two of these three criteria were prioritized, and four such genes were identified. First, CC071 possesses a deleterious variant in *Ak7*, which was associated with an immune function in MGI. *Ak7* deficient mice show signs of primary ciliary dyskinesia, a respiratory disease characterized by widening and obstruction of the bronchi due to a lack of ciliary motility (Fernandez-Gonzalez et al.



2009). *Ak7* deficient mice were also analyzed by the International Mouse Phenotyping Consortium and showed an abnormal thymus morphology, but it was not associated with more precise immune defects. Therefore, I did not retain *Ak7* as a priority candidate gene.

Similarly, CC071 has a deleterious variant in *Brf1*, which was associated with an immune function in MGI. *Brf1* loss-of-function was associated with disruption of homeostasis in the intestine, liver, and pancreas. This gene is associated with immune function in the MGI database as immune cell infiltration was observed in mice with *Brf1* deficient hepatocytes (Liko et al. 2019). As this gene was not shown to be involved in other immune functions, I did not retain this gene as a priority candidate for the present study.



**Figure 59:** Expression of *Bcl11b* and *Dlk1* in CC071 and B6 MEFs.

Data from Bourdon et al. 2023. MEFs derived from B6 (grey) and CC071 (red) were infected with ZIKV at a MOI of 5. *Bcl11b* (A) and *Dlk1* (B) mRNA expression levels were measured by RNA-seq in non-infected (NI) and ZIKV-infected MEFs. Data from 3 biological replicates.

The expression of *Bcl11b* was lower in CC071 MEFs than in B6 MEFs at basal condition and after ZIKV infection (Figure 59A). This gene was also associated with immune functions in mice. Indeed, *Bcl11b* deficient thymocytes show differentiation defects and are blocked at the CD4- CD8- double-negative stage due to impaired V(D)J rearrangement of the T cell receptor (Wakabayashi et al. 2003). Low *Bcl11b* in CC071 thymocytes could alter T cell differentiation in CC071 mice.

Lastly, *Dlk1* expression in CC071 MEFs was lower than in B6 MEFs (Figure 59B). This gene was shown to be involved in B cell development. Indeed, *Dlk1* deficiency resulted in reduced number of mature B cells in mice (Raghunandan et al. 2008). Thus, low expression of *Dlk1* in CC071 cells could result in altered B cell development.

All in all, this analysis of candidate genes in *Zs1* did not lead to an obvious candidate gene. Two genes which met the highest number of criteria were identified, *Bcl11b* and *Dlk1*. There is no known polymorphism in the coding sequence of these two genes between B6 and CC071, but polymorphisms in regulatory regions may impact the expression levels of these genes in CC071.

Additional criteria could be used to identify additional candidate genes which may be responsible for the association of the PVL with *Zs1*. As discussed above, open-source data of CC071 and B6 transcriptomes could be compared to identify differentially expressed genes. Protein levels could also be assessed to identify genes which are not differentially expressed between B6 and CC071 but for which the protein levels differ. Another approach would be to search for colocalizing QTLs associated with similar traits in other studies. For instance, a QTL on chromosome 12 was associated with lung viral titers after SARS-CoV infection (Gralinski et al. 2017). Combine crosses analyses could be performed using data from our segregating crosses and data from this study to refine the QTL interval. Besides, consequences of mutations in human or other species could be investigated to identify those leading to immune defects.

## **VI - Discussion**

My PhD project has investigated the genetic control of the susceptibility to ZIKV in mouse models, using genetic and functional tools, *in vitro* and *in vivo*. Here, I discuss the benefits of using this collection of various tools and the relevance of the use of genetically diverse strains for the study of viral diseases in mouse models.

## VI.1 . The contribution of cellular models to mouse genetic studies

Cellular models have extensively been used in virology, in humans (Najafi Fard et al. 2021; Kann et al. 1997; Reeves et al. 2005), and in mice (Yoneyama et al. 2004; Kawai et al. 2005). In particular, KO cell lines have been used to study the role of a specific gene, as exemplified for *Irf3* in section II.1.4.3.

Here, we have used cellular models for a forward genetic approach. We have used MEFs, a cellular model which is often used in mouse studies to investigate the innate antiviral response. Indeed, these cells are able to mount an efficient innate immune response as they express several PRRs which induce antiviral activities such as the IFN-I and NF- $\kappa$ B pathways (Tan and Lei 2019; O'Dea et al. 2007).

We produced a collection of (CC071  $\times$  CC001)  $\times$  CC071 backcross MEF lines and perform genetic mapping on *in vitro* phenotypes measured in 51 cell lines. This method has proven very efficient, as it resulted in a very significant LOD score, demonstrating that only a small number of individuals are required to map susceptibility alleles in this *in vitro* model. This was explained by the low complexity of the *in vitro* trait studied, which resulted from the effects of a small number of genes.

Examples of segregating cross analyzed in cellular models in the literature include a (C57BL/6ByJ  $\times$  BALB/cByJ)  $\times$  C57BL/6ByJ N2 used to investigate response to *Listeria monocytogenes* infection in bone marrow macrophages (BMMs) (Garifulin et al. 2007). While BMMs, as immune cells, may be more relevant to study the response to infection, they cannot be established as cell lines, and experiments cannot be repeated on the same individuals.

By contrast, the production of MEFs lines was separated from their phenotyping. MEFs can be frozen and cultured for a few passages. Thus, we have established a collection of 90 MEF lines from the N2 cross. Thus, like a GRP, this collection allows reproducibility, and several types of experiments can be performed on a cell line originating from a single embryo.

For instance, I would use this collection of N2 MEFs to investigate the response of CC001 MEFs to ZIKV infection. Our *in vitro* RNA-seq experiment using CC001, CC071 and B6 MEFs showed that while the expression of *Tbk1* and *Ikbke*, two members of the IFN-I induction cascade (see section II.1.4.2), does not vary after infection, their expression is upregulated in infected CC001 MEFs. The genetic control of the upregulation of these genes could be investigated in our collection of N2 MEFs, taking advantage of their already known genotypes.

In addition, C. Manet previously demonstrated that the percentage of ZIKV positive cells at early time points post infection (i.e., 16 and 24 hpi) was higher in CC001 MEFs than in CC071 and B6 MEFs (Manet 2019). The genetic control of these traits could also be investigated in N2 MEFs.

The number of experiments that can be performed on N2 MEFs remains limited by the number of samples produced per embryo. As illustrated in Figure 2B in section V.1, MEFs can also be derived from CC strains, which allows an infinite number of experiments on genetically identical cells. In this case, other cell types could also be considered. As review by Swanzey and colleagues, *in vitro*

investigation of genetically diverse mouse populations would allow a systems genetics approach by studying multiple cellular and molecular traits, while limiting animal use and economic cost (Swanzy et al. 2021). This strategy has been little exploited for now, but the development of next-generation sequencing methods has facilitated the investigation of molecular traits, such as RNA expression, protein levels, DNA methylation and chromatin accessibility, which will probably favor the investigation and genetic mapping of these molecular traits in mouse *in vitro* GRPs (Swanzy et al. 2021).

While *in vitro* mouse GRPs derived from CC strains have not yet been used to investigate the genetic control of the susceptibility to infections, results in *ex vivo* models has proven efficient. As presented in section II.3.4.2, Martin and colleagues investigated the composition of peripheral blood leucocytes of 47 CC strains, before and after *ex vivo* infection with LCMV. QTL mapping on immune cell counts revealed association with genetic factors involved in the development of subpopulations of T cells (Martin et al. 2020). This example suggests that similar approaches on *in vitro* cell lines derived from CC strains could be used to investigate the genetic control of cellular and molecular traits in response to infection.

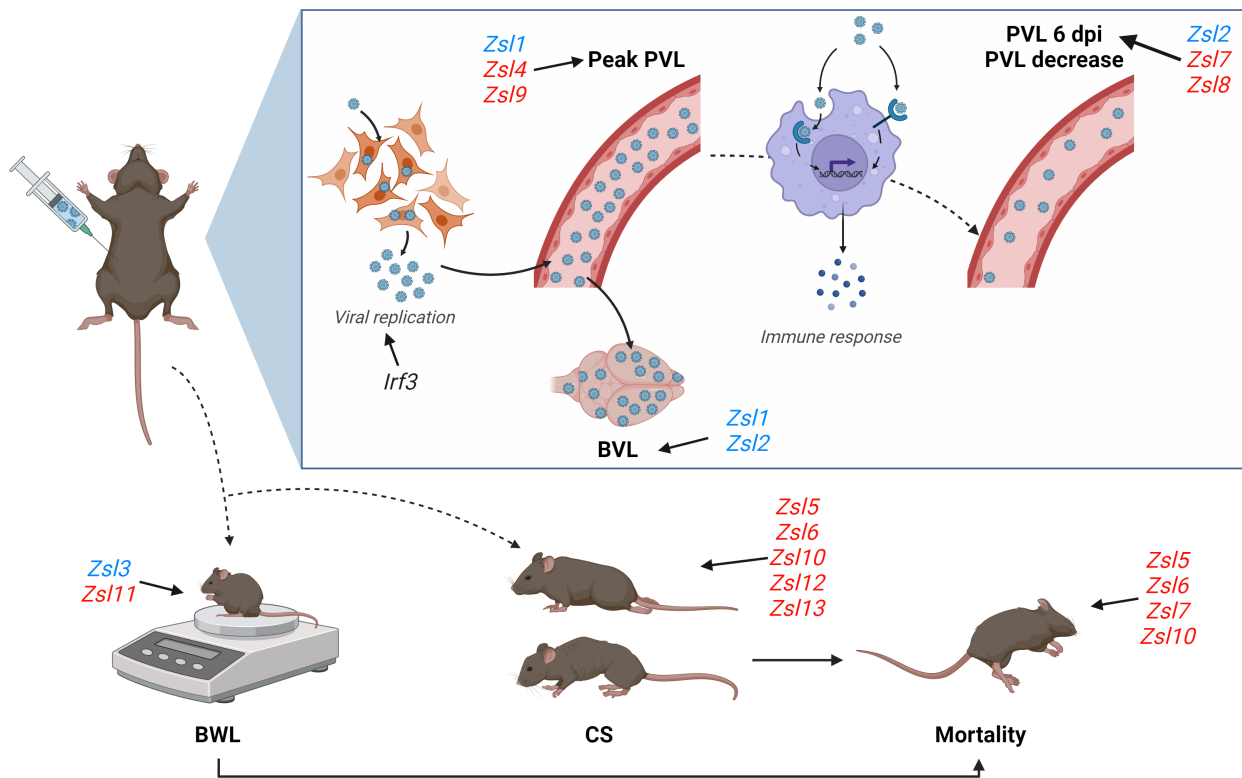
These approaches would be similar to the studies in human genetics which investigate response to microbial challenges in cells from healthy donors (Piasecka et al. 2018). Development of collections of cells have allowed investigating immune responses in varied genetic backgrounds and their genetic control, using a combination of omics methods (Aquino et al. 2023). As this field is expanding, new methods of molecular QTL mapping are developed, which could be applied to mouse genetic studies *in vitro*.

## **VI.2 . Segregating crosses using Collaborative Cross and *Ifnar1* deficient strains**

My PhD project was primarily based on the analysis of segregating crosses. Four crosses led to the identification of one causal variant and 14 QTLs (Figure 60). This method was used with varied experimental models. The first cross presented in my thesis was analyzed *in vitro*, as discussed in section VI.1, and two models of ZIKV infection in mice were used in the crosses analyzed *in vivo*, i.e., *Ifnar1* deficiency and the use of MAR1-5A3 antibody.

The model of treatment with MAR1-5A3 mainly identified QTLs associated with the PVLs, as antibody treated mice only showed light or mild symptoms, but a high variability of PVL (in blue in Figure 60). By contrast, the *Ifnar1* model led to more severe illness and allowed the identification of QTLs associated with disease signs (in red in Figure 60). Therefore, the use of these two complementary models allowed to investigate different aspects of the susceptibility to ZIKV in mice.

These examples demonstrate the benefit of multiplying the phenotypes studied in a segregating cross. As our goal was to identify the different mechanisms of susceptibility to ZIKV infection, we investigated various aspects of the disease by measuring multiple readouts in infected mice, which were the viral loads in the plasma and in the brain, body weights, clinical symptoms, and mortality (Figure 60). Analysis of the genetic control of each of these traits will allow to better characterize the mechanisms of susceptibility, which will provide new information on ZIKV biology in the infected host, and on the host response to viral infection. This analysis will be complete once candidate genes



**Figure 60:** Implication of the significant QTLs identified in my PhD on the physiopathology of ZIKV disease. QTLs identified in mice treated with MAR1-5A3 are displayed in blue, while QTLs identified in *Irfnar1* deficient mice are displayed in red.

are found and validated, but analysis of the QTLs found already gives us insight of the mechanisms into the physiopathology of ZIKV disease. For instance, the colocalization of QTLs identified in the N2-B6.129-*Irfnar1* controlling the PVL at 6 dpi and the survival suggested a common mechanism involved in the control of these two traits. Further studies will determine whether the PVL of infected *Irfnar1* deficient mice is predictive of their survival.

Moreover, we could obtain multiple QTLs because we have used different parental strains for these crosses. The susceptibility of CC071 mice was assessed by crossing them with three other strains, i.e., CC001, B6-*Irf3* and 129-*Irfnar1*. In each of these three crosses, the two parental strains carried chromosomal regions identical by descent. Each CC strain has one eighth of its genome identical by descent to each CC founder strain, and each other CC strain, due to the breeding scheme used to create the CC (Figure 8). This was particularly visible in the marker maps for the N2-*Irf3* and F2-*Irfnar1* progenies presented in Figures S2A and B, showing that some chromosomal regions did not contain any marker usable for the QTL mapping analysis, such as the proximal region of chromosome 3 in the N2-*Irf3* progeny. This may lead to an incapacity to detect QTLs in these regions, but this effect could be avoided by crossing the CC strain of interest with a non-CC and non-CC founder strain. For instance, this was proposed by Alugupalli and colleagues, after identifying among 9 CC strains the CC003/Unc and CC053/Unc as permissive to *Salmonella Typhi* infection. To identify new susceptibility genes, the authors suggested to cross one of these strains with BALB/cj mice, known as resistant, to produce an F2 progeny (Alugupalli et al. 2023).

This is not the only example from the literature where authors screened a small number of mice and plan to produce segregating crosses to identify susceptibility genes. The 6 studies screening CC strains for their susceptibility to infectious diseases published in 2023 have used 8, 9, 10, 11, 17 and 23 strains, and only the latter performed QTL mapping (Lone et al. 2023b; Alugupalli et al. 2023; Brown et al. 2023; Lone et al. 2023a; Jasperse et al. 2023; Karmakar et al. 2023). As for Alugupalli and colleagues, Jasperse and colleagues did not report QTL mapping results obtained on CC strains. By contrast, they identified strains with extreme phenotypes, in particular CC045/GeniUnc, the only resistant strain to POWV infection, and CC071, which was one of the most susceptible strains, and plan to intercross these two strains to perform QTL mapping on the F2s (Jasperse et al. 2023).

The switch from the strategy of screening of a large number of CC strains (Ferris et al. 2013; Gralinski et al. 2015; Nashef et al. 2018; Zhang et al. 2018) to a smaller screen followed by the analysis of segregating crosses may have become more used as it was demonstrated to efficiently identify susceptibility genes. Over the 9 studies which has validated the role of a gene in the susceptibility to infectious diseases (Ferris et al. 2013; Gralinski et al. 2015; Gralinski et al. 2017; Green et al. 2017; Smith et al. 2019; Zhang et al. 2019; Lorè et al. 2020; Schäfer et al. 2022; Bourdon et al. 2023), five have used segregating crosses. Therefore, this strategy can be an efficient method to more rapidly detect QTLs and validate the role of underlying susceptibility genes.

### VI.3 . Perspectives on the identification of susceptibility alleles

As presented in Figure 60, I have identified during my PhD project fourteen significant QTLs involved in the susceptibility of CC071 and B6-*Ifnar1* mice. Each QTL may include one or several gene(s). Therefore, the next step of this project will be the identification of candidate genes, and the validation of their role in the susceptibility of these two mouse strains.

Several methods can be used for selecting candidate genes, some of which were presented in section V.4.5. The methods to favor depend on the specificities of each QTL, such as the phenotype associated, the size of its credible interval, and the allelic effects. For instance, a QTL with a small interval could more easily be dissected by looking at variants between parental strains than a QTL with a large interval, which should be first reduced using congenic strains. Generally speaking, it is the combination of these multiple approaches that will identify the best candidate genes.

Genetic approaches can be employed. Genomic sequences of laboratory mouse strains including CC strains are available (Srivastava et al. 2017) and new bioinformatic tools allow to easily search for polymorphisms (Ball et al. 2023). The quantity of variants in a QTL interval can be large. For instance, 817,152 sequence variants were found in *Zs1* credible interval before filtering for those with moderate or deleterious predicted effect, which is why this method is more appropriate for QTLs with small intervals.

Another approach is to search for colocalizing QTLs associated with similar traits found in other studies. As discussed in section V.4.5, we could search for colocalizing QTLs in mouse studies, and colocalizing QTLs can also be found in other species. As homologous segments, e.g., with the same linear organization of genes, exist between species, QTLs found in other species, such as in human or in rat, could be used to refine a QTL interval, assuming that the QTLs are controlled by orthologous

genes (DiPetrillo et al. 2005).

One unequivocal method to decrease the number of candidate genes is to reduce the QTL interval by using congenic strains. Congenic strains allow investigating a QTL independently of the genetic background of the susceptible strain in which it was detected, and, as a result, identify the precise phenotypes in which it is involved (Morel et al. 1997).

Congenic strains also allow to investigate the interactive effects of multiple QTLs. If several QTLs are involved in the phenotype of a susceptible strain, the use of polycongenic strains can determine which loci are necessary to induce the complete susceptibility phenotype. For instance, Morel and colleagues identified three loci involved in systemic lupus erythematosus in the NZM2410 mouse strain. Then, they analyzed a polycongenic strain carrying the three locus on a B6 background. This triple congenic strain fully recapitulated the phenotype observed in NZM2410 mice, demonstrating that the three loci were sufficient to induce lupus. The authors elegantly compare this approach with Koch's postulate. Indeed, the three loci were "isolated" from susceptible NZM2410 strain, and together induce lupus in a resistant B6 genetic background (Morel et al. 2000). This approach could be used to investigate the additive role of the QTLs associated with the high PVL of CC071 mice (Figure 24).

Lastly, the use of congenic strains allow fine-mapping of the genetic location of a QTL. In contrast with the Bayesian credible interval which is a statistical approach, the interval of a QTL transferred in a congenic strain is a definite interval. With this method, the size of the QTL interval can be reduced by producing subcongenic strains carrying different fragments of the QTL interval. Subcongenic strains may also identify a polygenic control of a QTL. For instance, Morel and colleagues have used subcongenic strains to dissect one of the lupus susceptibility loci mentioned above. They found that three different loci within the interval were independently involved in the lupus phenotype. Analysis of each of the three subcongenic strains identified the specific aspect of lupus pathogenesis in which each locus is involved (Morel et al. 2001).

Therefore, congenic strains allow to precisely dissect the genetic control underlying QTL and investigate the interactions between severity loci. This method is particularly relevant to study QTLs with large credible intervals, because it would either reduce the QTL interval or identify a polygenic control of susceptibility.

Concurrently with genetic methods, functional approaches can be used to filter the list of candidate genes. These approaches will refine the susceptible phenotype of CC071 and B6-*Irfnar1* mice, and identify mechanisms that may be altered in these strains and which may explain their high PVL and severe illness after ZIKV infection.

*Irf3* was a good example of the interest of identifying phenotypes closer to gene effects, unlike body weight loss or death. Starting from the *in vivo* susceptibility of CC071 mice, we reduced the complexity by phenotyping MEFs and eventually identified a delayed IFN-I induction. This molecular trait was more easily linked to a small number of genes. Therefore, when we identified a QTL on chromosome 7, *Irf3* was immediately identified as a very credible candidate (see section V.1).

One aspect of the physiopathology that I would further investigate is the identification of the tissues responsible for the high PVL observed in CC071 and B6-*Irfnar1*. Viral loads in various organs could be assessed to detect the tissues responsible for the elevated viral loads in the plasma. The



tissue tropism of ZIKV in B6-*Ifnar1* mice have already been investigated by Lazear and colleagues and revealed that multiple organs are infected in this model, such as the spleen, liver, kidney and brain (Lazear et al. 2016). The same approach could be used on the CC071 strain and compared with resistant strains, such as 129-*Ifnar1* or CC001, or with other strains which showed a high PVL, such as the CC005/TauUnc strain, to identify mechanisms that specifically explain the high viral loads observed in CC071 mice. Viral dissemination dynamics could be measured using a bioluminescent reporter virus (Wang et al. 2020).

I would also investigate histological profiles in these tissues, such as the presence of inflammatory lesions. For instance, brain pathology has already been examined in ZIKV-infected CC071 mice and revealed inflammatory lesions and activation of microglial cells (Manet et al. 2020). Similar approaches could be used for other tissues of infected mice. Thus, if a tissue with particularly high viral replication or inflammatory phenotype is identified, we could study *ex vivo* or *in vitro* cells from this tissue to decipher the underlying mechanisms, similar to our study in MEF.

Immune phenotyping could also be performed on CC071 and B6-*Ifnar1* mice to investigate any anomalies, at the basal state or after infection. Cytokines and chemokines could be assessed longitudinally in the serum of mice after ZIKV infection using multiplex analysis, and blood cell counts could be measured. This strategy of broad phenotyping of a CC strain have already been used. After identifying the CC057/Unc strain as a "divergent" CC strain after RVFV infection, as it died later than other strains, CC057/Unc and B6 mice were compared for viral loads in and multiple organs, tissue pathology, blood cell counts, cytokine and chemokine levels. The authors have identified specific signatures in CC057 mice, such as anemia, high levels of inflammatory cytokines, and high neutrophil and lymphocyte counts, which could be used as biomarkers for RVFV disease (Cartwright et al. 2022).

Immune phenotyping of the CC071 and B6-*Ifnar1* strains could be assessed through a collaboration with mouse clinics specialized in the phenotyping of mouse inbred strains. For instance, my laboratory is collaborating with the Institut Clinique de la Souris in Strasbourg. We are also considering a collaboration with the Center for Immunophenomics of Marseille, which performs high-throughput immunity profiling of mice in basal condition and post infection. These analyses of immune traits in CC071 will help determine whether *Bcl11b* and *Dlk1*, identified in section V.4.5 as candidate genes for *Zs1*, might be involved in the susceptibility of CC071 mice. As these genes are involved in immune cell development and differentiation, analysis of immune cell composition and activation in CC071 might detect anomalies which would be consistent with potential mutations in one of these genes

To conclude, the identification of candidate genes for a QTL requires that multiple approaches are used in combination. Each tool provides specific insight on the genes that may be involved in the susceptibility, and the intersection between these different hints will determine the best candidates. As each approach for this analysis requires different expertise which are brought by scientists in different fields, the search for candidate genes is an occasion to build collaborative projects with research teams outside the field of mouse genetics.

#### VI.4 . CC071: a complex model of viral disease

The results presented in this thesis have provided new information on the susceptibility of the CC071 strain to the infection by ZIKV virus. This strain carries a loss-of-function mutation in *Irf3*, and segregating crosses have identified multiple QTLs involved in its susceptibility to *in vivo* infections. Therefore, the unique genetic background of CC071, which involves several factors involved in its susceptibility to viral disease, makes it a relevant model for studying ZIKV infection in mice.

ZIKV infection in CC071 mice was studied after treatment with MAR1-5A3. This method is one of the two most used models for ZIKV infection in mice, the second one being the use of *Ifnar1* deficiency. These two models have proven effective for the study of ZIKV in mice, as demonstrated by many studies in the literature (Lazear et al. 2016; Matz et al. 2021; Oh et al. 2021; Li et al. 2023), and also by my results of segregating crosses using these two models, described in the section VI.2.

These models have been criticized as they do not closely reflect the pathogenesis in human hosts (Gorman et al. 2018). While the use of immunodeficient or sensitized models should not be excluded for the study of ZIKV in mice, the use of models in which the IFN-I response is not inhibited could be used to better reflect infection in humans (Lazear et al. 2016; Morrison and Diamond 2017). Inhibiting the IFN-I response may hamper the study of certain aspects of ZIKV disease. For instance, vaccine testing in *Ifnar1* deficient mice could be obstructed by the lack of IFN-I signaling, which is involved in B and T cell responses (Lazear et al. 2016). Similarly, the antibody response of MAR1-5A3 treated mice differs from untreated mice as they produce antibodies that recognize different epitopes of ZIKV proteins (Lee et al. 2020). Thus, for these specific aspects of ZIKV studies in mice, a model without a fully abolished IFN-I response would be valuable.

Besides, our results have shown that, compared with resistant strains, CC071 mice have higher PVL, even without MAR1-5A3 treatment (Manet et al. 2020; Bourdon et al. 2023), suggesting that ZIKV is able to replicate at low levels in CC071, without the necessity to block the IFN-I response. Thus, a more acute model of ZIKV infection might result in permissive infection in CC071 mice, which would represent a new lead for a model of ZIKV infection in mice.

For instance, a mouse-adapted strain of ZIKV was developed, but still requires the use of MAR1-5A3 treatment in WT mice to observe clinical symptoms (Gorman et al. 2018). Considering that ZIKV is able to replicate at low levels in CC071, infection of CC071 mice with the mouse-adapted strain of ZIKV could be permissive and result in an elevated viral load, or even in clinical symptoms. The mouse-adapted ZIKV strain was tested on twenty-one CC strains, for the purpose of finding an immunocompetent model of ZIKV infection in mice, but CC071 was not tested. Moreover, this study used a subcutaneous infection, while we use an i.p. route (Mattocks et al. 2019). Thus, i.p. infection of CC071 mice could be tested to potentially identify a model of ZIKV infection which does not require MAR1-5A3 treatment.

Besides, compared with natural infection, the i.p. route bypasses a series of local events which precede systemic dissemination of the virus after transmission by mosquito bite, while the subcutaneous route is more representative of natural transmission by mosquitoes. Different

infection routes may involve different susceptibility mechanisms. Indeed, while CC001 and CC042/GeniUnc have similar PVL after subcutaneous ZIKV infection (Mattocks et al. 2019), the PVL of CC001 mice is about ten times higher than that of CC042/GeniUnc mice after i.p. infection (Manet et al. 2020), suggesting that the susceptibility to ZIKV in mice may depend on the infection route. Therefore, it would be relevant to infect CC071 mice via a subcutaneous route to know whether the high susceptibility of this strain is specific to the i.p. route.

The use of CC071 mice, which susceptibility involves multiple genetic factors, as a model for ZIKV infection would better reflect the susceptibility to viral infections in humans, which relies on multiple genetic (and non-genetic) factors. Besides, the *Irf3* deficiency in CC071, altering IFN-I induction, could mimic ZIKV infection in human, as the virus is able to inhibit but not fully abolish the IFN-I response (Hu et al. 2023). This model would be beneficial to study aspects of ZIKV infection and pathogenesis requiring that the IFN-I is not fully inhibited, such as the humoral response.

CC071 could also be a model for other viral infections in mice. Indeed, my laboratory demonstrated that the CC071 strain also showed increased susceptibility to other flaviviruses, DENV and WNV. By contrast, compared with CC001, CC005/TauUnc and BALB/c mice, CC071 did not show significantly higher mortality rate (Manet et al. 2020). However, a more recent study screening 20 CC strains with Rift Valley fever virus infection showed that the CC071 was one of the most susceptible CC strain and that all CC071 mice succumbed by 3 dpi (Cartwright et al. 2022). More recently, CC071 mice were shown to be susceptible to POWV and NrHV (Jasperse et al. 2023; Brown et al. 2023). Thus, the CC071 strain could also be used as a model for the infection with these viruses.

Because CC071 is a complex genetic model, its multiple susceptibility alleles may contribute differently to infection with each virus. For instance, while the *Irf3* mutation may not be involved in its susceptibility to *in vivo* infection with ZIKV, it may be sufficient to explain its susceptibility to other viruses. *Irf3* KO mice are more susceptible to WNV than WT mice (Daffis et al. 2007), thus, the *Irf3* loss-of-function of CC071 mice may be involved in their susceptibility to WNV, although the defective *Oas1b* allele of B6-*Irf3* and CC071 mice likely contributes to their susceptibility. Besides, *Irf3* KO mice are not susceptible to DENV infection (Chen et al. 2013), thus, other susceptibility genes than *Irf3* are likely involved in the susceptibility of CC071 mice to DENV. Whether *Irf3* or the other CC071 susceptibility alleles explain the susceptibility of this strain to POWV and NrHV remains to be understood.

To conclude, the CC071 strain carries multiple susceptibility alleles involved in its susceptibility to ZIKV, and which may be involved in its susceptibility to other viruses. Further investigations using this strain could shed light on the mechanisms of susceptibility that are virus-specific or that are shared by multiple viral families.

## VI.5 . The Collaborative Cross: a model of the genetic and phenotypic diversity of the human population

The human population presents a large level of genetic variation, which has not always been reflected in biological studies (Oh et al. 2015; Bentley et al. 2017; Hindorff et al. 2018; Passmore et al. 2022; Smith et al. 2022). Lack of genetic diversity in research may lead to detrimental consequences on the treatment of minorities. For instance, the African American population in the United States is particularly at risk of HCV infection. While they represent 13% of the population, they account for 23% of the cases of hepatitis C. However, this population is underrepresented in clinical trials for HCV treatments (Wilder et al. 2016). Similarly, while Black and Hispanic populations are more at risk to develop Alzheimer's disease, they only accounted for 20% over about 1,800 participants for the clinical trial approving the latest drug, lecanemab, in July 2023 (Reardon 2023; Dyck et al. 2023). Yet, the efficacy of treatments can vary across ethnic groups. For instance, 67.0% of Puerto Rican asthmatic children do not respond to albuterol, a drug commonly used to treat asthma, compared with 42.1% for Mexican children (Naqvi et al. 2007). These are examples of applied research, but genetic diversity should be considered at all levels of biological research, starting with mouse models.

Mice have been used as a model organism of human phenotypes and disease for over 120 years, thanks to the development of inbred strains (Phifer-Rixey and Nachman 2015), but have often been criticized for the difficulties to translate results obtained in mice to human (Perlman 2016). Many factors have been highlighted to explain these difficulties, such as the differences of metabolic rate, diet, microbiome, and pathogen exposure (Perlman 2016). New approaches have tried to tackle these issues, such as the development of humanized mice, which are grafted with human cells or tissues (Fujiwara 2018), or human fecal microbiota transplant (Burz et al. 2021). The other aspect that may contribute to this difficulty to translate the results obtained in mice to humans is the use of a single genetic background to model humans, ignoring the vast genetic diversity of the human population. I believe that using genetically diverse mouse population such as the CC can resolve this issue.

The level of genetic diversity of the CC was maximized by using founder strains of three *Mus musculus* species, resulting in the capture of about 90% of the genetic diversity of laboratory mice (Roberts et al. 2007), allowing to map genes associated with complex traits (The Complex Trait Consortium 2004). This high level of genetic diversity of the CC resulted in large phenotypic variations across this population. Indeed, all the studies which have used the CC cited in this thesis have succeeded in identifying phenotypic differences across the strains tested. This phenotypic diversity is the raw material for genetic studies, which makes the CC a powerful tool for mouse geneticists. Furthermore, the phenotypic diversity observed in CC strains can be utilized more generally in studies using mouse models.

The CC allows taking into account the effect of the genetic diversity on drugs efficacy from the stage of testing on animal models (Ryan et al. 2021). Generalization of this method will greatly improve the capacity to select treatments which are effective on a variety of genetic backgrounds for clinical trials in humans. Adding to this a better representation of the genetic variability of the human

population in clinical trials, cases of treatments that are poorly efficient in certain ethnic group would be decreased.

Besides, the high level of genetic diversity in the CC allowed identifying strains which develop illness that can be used as models of human diseases. For instance, CC011/Unc mice spontaneously develops colitis. This strain can be used as a model of inflammatory bowel disease to replace standard models which require induction of colitis by chemical treatment, bacterial infection and/or genetic engineering (Rogala et al. 2014). Thus, the CC011/Unc can be used to decipher the mechanisms leading to inflammatory bowel disease in humans and test treatments. Similarly, this thesis has demonstrated that the CC071 strain is getting increasingly useful to study viral infections.

Discovery of strains with particular phenotypes such as the CC071 and the CC011/Unc are only possible if the CC strains are extensively phenotyped. For the study of viral diseases and more broadly of immunity, high throughput immune phenotyping of CC strains may identify strains with particular phenotypes. These strains could be investigated by mouse geneticists, to decipher the genetic control of these phenotypes, and would also be relevant for functional studies by virologists and immunologists.

To conclude, the phenotypic diversity observed in CC strains has the potential to reflect that of the human population, which may facilitate the translation between mouse and human studies. Besides, the best possible use of the CC for the study of immune traits including viral diseases is through a collaborative approach between mouse geneticists, immunologist and virologists, which will allow to better decipher the mechanisms of susceptibility in mouse models.

## VII - References

- Acosta-Ampudia, Yeny et al. (2018). "Autoimmune Neurological Conditions Associated With Zika Virus Infection". In: *Frontiers in Molecular Neuroscience* 11, p. 116. ISSN: 1662-5099. DOI: [10.3389/fnmol.2018.00116](https://doi.org/10.3389/fnmol.2018.00116).
- Agarwal, Ankita and Deepti Chaurasia (2021). "The expanding arms of Zika virus: An updated review with recent Indian outbreaks". In: *Reviews in Medical Virology* 31.1, pp. 1–9. ISSN: 1099-1654. DOI: [10.1002/rmv.2145](https://doi.org/10.1002/rmv.2145).
- Aguiar, Renato S. et al. (2020). "Molecular alterations in the extracellular matrix in the brains of newborns with congenital Zika syndrome". In: *Science Signaling* 13.635, eaay6736. ISSN: 1937-9145. DOI: [10.1126/scisignal.aay6736](https://doi.org/10.1126/scisignal.aay6736).
- Alugupalli, Kishore R. et al. (2023). "Identification of collaborative cross mouse strains permissive to *Salmonella enterica* serovar Typhi infection". In: *Scientific Reports* 13.1, p. 393. ISSN: 2045-2322. DOI: [10.1038/s41598-023-27400-1](https://doi.org/10.1038/s41598-023-27400-1).
- Alves dos Santos, Eduardo and Katja Fink (2018). "Animal Models for Dengue and Zika Vaccine Development". In: *Dengue and Zika: Control and Antiviral Treatment Strategies*. Ed. by Rolf Hilgenfeld and Subhash G. Vasudevan. Advances in Experimental Medicine and Biology. Singapore: Springer, pp. 215–239. ISBN: 978-981-10-8727-1. DOI: [10.1007/978-981-10-8727-1\\_16](https://doi.org/10.1007/978-981-10-8727-1_16).
- Alwine, James C. et al. (2023). "A Critical Analysis of the Evidence for the SARS-CoV-2 Origin Hypotheses". In: *mBio* 14.2. Publisher: American Society for Microbiology, e00583–23. DOI: [10.1128/mbio.00583-23](https://doi.org/10.1128/mbio.00583-23).
- Andrews, Nick et al. (2022). "Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant". In: *New England Journal of Medicine* 386.16, pp. 1532–1546. ISSN: 0028-4793. DOI: [10.1056/NEJMoa2119451](https://doi.org/10.1056/NEJMoa2119451).
- Aquino, Yann et al. (2023). "Dissecting human population variation in single-cell responses to SARS-CoV-2". In: *Nature*. Publisher: Nature Publishing Group, pp. 1–9. ISSN: 1476-4687. DOI: [10.1038/s41586-023-06422-9](https://doi.org/10.1038/s41586-023-06422-9).
- Au, W C et al. (1995). "Identification of a member of the interferon regulatory factor family that binds to the interferon-stimulated response element and activates expression of interferon-induced genes." In: *Proceedings of the National Academy of Sciences* 92.25, pp. 11657–11661. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.92.25.11657](https://doi.org/10.1073/pnas.92.25.11657).
- Aubry, Fabien et al. (2021). "Recent African strains of Zika virus display higher transmissibility and fetal pathogenicity than Asian strains". In: *Nature Communications* 12.1, p. 916. ISSN: 2041-1723. DOI: [10.1038/s41467-021-21199-z](https://doi.org/10.1038/s41467-021-21199-z).
- Backwell, Lisa and Joseph A. Marsh (2022). "Diverse Molecular Mechanisms Underlying Pathogenic Protein Mutations: Beyond the Loss-of-Function Paradigm". In: *Annual Review of Genomics and Human Genetics* 23.1, pp. 475–498. ISSN: 1527-8204, 1545-293X. DOI: [10.1146/annurev-genom-111221-103208](https://doi.org/10.1146/annurev-genom-111221-103208).
- Bálint, Gábor, Barbara Vörös-Horváth, and Aleksandar Széchenyi (2022). "Omicron: increased transmissibility and decreased pathogenicity". In: *Signal Transduction and Targeted Therapy* 7.1. Number: 1 Publisher: Nature Publishing Group, pp. 1–3. ISSN: 2059-3635. DOI: [10.1038/s41392-022-01009-8](https://doi.org/10.1038/s41392-022-01009-8).
- Ball, Robyn L. et al. (2023). *GenomeMUSter mouse genetic variation service enables multi-trait, multi-population data integration and analyses*. Pages: 2023.08.08.552506 Section: New Results. DOI: [10.1101/2023.08.08.552506](https://doi.org/10.1101/2023.08.08.552506).
- Bardhan, Mainak et al. (2021). "Dual burden of Zika and COVID-19 in India: challenges, opportunities and recommendations". In: *Tropical Medicine and Health* 49.1, p. 83. ISSN: 1349-4147. DOI: [10.1186/s41182-021-00378-0](https://doi.org/10.1186/s41182-021-00378-0).
- Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67, pp. 1–48. ISSN: 1548-7660. DOI: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Beaver, Jacob T. et al. (2018). "Evolution of Two Major Zika Virus Lineages: Implications for Pathology, Immune Response, and Vaccine Development". In: *Frontiers in Immunology* 9, p. 1640. ISSN: 1664-3224. DOI: [10.3389/fimmu.2018.01640](https://doi.org/10.3389/fimmu.2018.01640).
- Bentley, Amy R., Shawneequa Callier, and Charles N. Rotimi (2017). "Diversity and inclusion in genomic research: why the uneven progress?" In: *Journal of Community Genetics* 8.4, pp. 255–266. ISSN: 1868-310X. DOI: [10.1007/s12687-017-0316-6](https://doi.org/10.1007/s12687-017-0316-6).
- Bhardwaj, Utkarsh et al. (2021). "Gist of Zika Virus pathogenesis". In: *Virology* 560, pp. 86–95. ISSN: 00426822. DOI: [10.1016/j.virol.2021.04.008](https://doi.org/10.1016/j.virol.2021.04.008).
- Bhat, Eijaz Ahmed et al. (2023). "Insights into the structure, functional perspective, and pathogenesis of ZIKV: an updated review". In: *Biomedicine & Pharmacotherapy* 165, p. 115175. ISSN: 07533322. DOI: [10.1016/j.biopha.2023.115175](https://doi.org/10.1016/j.biopha.2023.115175).
- Bieber, Allan J. et al. (2010). "Allelic variation in the Tyk2 and EGF genes as potential genetic determinants of CNS repair". In: *Proceedings of the National Academy of Sciences* 107.2. Publisher: Proceedings of the National Academy of Sciences, pp. 792–797. DOI: [10.1073/pnas.0906589107](https://doi.org/10.1073/pnas.0906589107).
- Bigham, A. W. et al. (2011). "Host genetic risk factors for West Nile virus infection and disease progression". In: *PLoS One* 6.9, e24745. ISSN: 1932-6203 (Electronic) 1932-6203 (Linking). DOI: [10.1371/journal.pone.0024745](https://doi.org/10.1371/journal.pone.0024745).
- Biswas, Ashutosh et al. (2020). "Zika outbreak in India in 2018". In: *Journal of Travel Medicine* 27.4, taaa001. ISSN: 1195-1982, 1708-8305. DOI: [10.1093/jtm/taaa001](https://doi.org/10.1093/jtm/taaa001).
- Blondeau, Caroline et al. (2013). "Tetherin Restricts Herpes Simplex Virus 1 and Is Antagonized by Glycoprotein M". In: *Journal of Virology* 87.24. Publisher: American Society for Microbiology, pp. 13124–13133. DOI: [10.1128/jvi.02250-13](https://doi.org/10.1128/jvi.02250-13).
- Bock, Christoph et al. (2022). "High-content CRISPR screening". In: *Nature Reviews Methods Primers* 2.1. Number: 1 Publisher: Nature Publishing Group, pp. 1–23. ISSN: 2662-8449. DOI: [10.1038/s43586-021-00093-4](https://doi.org/10.1038/s43586-021-00093-4).
- Borda, Victor et al. (2021). "Whole-exome sequencing reveals insights into genetic susceptibility to Congenital Zika Syndrome". In: *PLOS Neglected Tropical Diseases* 15.6. Ed. by Helene Dutartre, e0009507. ISSN: 1935-2735. DOI: [10.1371/journal.pntd.0009507](https://doi.org/10.1371/journal.pntd.0009507).
- Boudinot, Pierre et al. (2016). "The Peculiar Characteristics of Fish Type I Interferons". In: *Viruses* 8.11. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute, p. 298. ISSN: 1999-4915. DOI: [10.3390/v8110298](https://doi.org/10.3390/v8110298).
- Bourdon, Marie, Caroline Manet, and Xavier Montagutelli (2020). "Host genetic susceptibility to viral infections: the role of type I interferon induction". In: *Genes & Immunity* 21.6, pp. 365–379. ISSN: 1466-4879, 1476-5470. DOI: [10.1038/s41435-020-00116-2](https://doi.org/10.1038/s41435-020-00116-2).
- Bourdon, Marie and Xavier Montagutelli (2022). "stuart: an R package for the curation of SNP genotypes from experimental crosses". In: *G3 (Bethesda, Md.)*, jkac219. ISSN: 2160-1836. DOI: [10.1093/g3journal/jkac219](https://doi.org/10.1093/g3journal/jkac219).

- Bourdon, Marie et al. (2023). "Susceptibility to Zika virus in a Collaborative Cross mouse strain is induced by Irf3 deficiency in vitro but requires other variants in vivo". In: *PLoS Pathogens* 19.9. Publisher: Public Library of Science, e1011446. ISSN: 1553-7374. DOI: [10.1371/journal.ppat.1011446](https://doi.org/10.1371/journal.ppat.1011446).
- Bradley, Michael P and Claude M Nagamine (2017). "Animal Models of Zika Virus". In: *Comparative Medicine* 67.3, pp. 242–252.
- Broman, K. W. (2001). "Review of statistical methods for QTL mapping in experimental crosses". In: *Lab Animal* 30.7, pp. 44–52. ISSN: 0093-7355.
- Broman, K. W. et al. (2003). "R/qrtl: QTL mapping in experimental crosses". In: *Bioinformatics* 19.7, pp. 889–890. ISSN: 1367-4803, 1460-2059. DOI: [10.1093/bioinformatics/btg112](https://doi.org/10.1093/bioinformatics/btg112).
- Broman, Karl and Saunak Sen (2009). *A Guide to QTL Mapping with R/qrtl: Online complements*.
- Broman, Karl W (2003). "Mapping Quantitative Trait Loci in the Case of a Spike in the Phenotype Distribution". In: *Genetics* 163.3, pp. 1169–1175. ISSN: 1943-2631. DOI: [10.1093/genetics/163.3.1169](https://doi.org/10.1093/genetics/163.3.1169).
- Brown, Ariane J. et al. (2023). "Host genetic variation guides hepatitis C virus clearance, chronicity, and liver fibrosis in mice". In: *Hepatology*. ISSN: 0270-9139. DOI: [10.1097/HEP.0000000000000547](https://doi.org/10.1097/HEP.0000000000000547).
- Bureau, Jean-François et al. (1993). "Mapping loci influencing the persistence of Theiler's virus in the murine central nervous system". In: *Nature Genetics* 5.1. Number: 1 Publisher: Nature Publishing Group, pp. 87–91. ISSN: 1546-1718. DOI: [10.1038/ng0993-87](https://doi.org/10.1038/ng0993-87).
- Burz, Sebastian D. et al. (2021). "Fecal Microbiota Transplant from Human to Mice Gives Insights into the Role of the Gut Microbiota in Non-Alcoholic Fatty Liver Disease (NAFLD)". In: *Microorganisms* 9.1, p. 199. ISSN: 2076-2607. DOI: [10.3390/microorganisms9010199](https://doi.org/10.3390/microorganisms9010199).
- Caires-Júnior, Luiz Carlos et al. (2018). "Discordant congenital Zika syndrome twins show differential in vitro viral susceptibility of neural progenitor cells". In: *Nature Communications* 9.1, p. 475. ISSN: 2041-1723. DOI: [10.1038/s41467-017-02790-9](https://doi.org/10.1038/s41467-017-02790-9).
- Canivet, Coraline et al. (2018). "Both IRF3 and especially IRF7 play a key role to orchestrate an effective cerebral inflammatory response in a mouse model of herpes simplex virus encephalitis". In: *Journal of Neurovirology* 24.6, pp. 761–768. ISSN: 1538-2443. DOI: [10.1007/s13365-018-0666-9](https://doi.org/10.1007/s13365-018-0666-9).
- Carbaugh, Derek L. et al. (2020). "Two Genetic Differences between Closely Related Zika Virus Strains Determine Pathogenic Outcome in Mice". In: *Journal of Virology* 94.20. Ed. by Julie K. Pfeiffer, e00618–20. ISSN: 0022-538X, 1098-5514. DOI: [10.1128/JVI.00618-20](https://doi.org/10.1128/JVI.00618-20).
- Carod-Artal, Francisco Javier (2018). "Neurological complications of Zika virus infection". In: *Expert Review of Anti-infective Therapy* 16.5, pp. 399–410. ISSN: 1478-7210, 1744-8336. DOI: [10.1080/14787210.2018.1466702](https://doi.org/10.1080/14787210.2018.1466702).
- Cartwright, Haley N. et al. (2022). "Genetic diversity of collaborative cross mice enables identification of novel rift valley fever virus encephalitis model". In: *PLoS Pathogens* 18.7, e1010649. ISSN: 1553-7374. DOI: [10.1371/journal.ppat.1010649](https://doi.org/10.1371/journal.ppat.1010649).
- Chattopadhyay, Saurabh et al. (2016). "Ubiquitination of the Transcription Factor IRF-3 Activates RIP1, the Apoptotic Pathway that Protects Mice from Viral Pathogenesis". In: *Immunity* 44.5, pp. 1151–1161. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2016.04.009](https://doi.org/10.1016/j.immuni.2016.04.009).
- Chen, H. W. et al. (2013). "The roles of IRF-3 and IRF-7 in innate antiviral immunity against dengue virus". In: *J Immunol* 191.8, pp. 4194–201. ISSN: 1550-6606 (Electronic) 0022-1767 (Linking). DOI: [10.4049/jimmunol.1300799](https://doi.org/10.4049/jimmunol.1300799).
- Chen, Y. et al. (2005). "Thermolabile phenotype of carnitine palmitoyltransferase II variations as a predisposing factor for influenza-associated encephalopathy". In: *FEBS Letters* 579.10, pp. 2040–2044. ISSN: 1873-3468. DOI: [10.1016/j.febslet.2005.02.050](https://doi.org/10.1016/j.febslet.2005.02.050).
- Choi, Un Yung et al. (2015). "Oligoadenylate synthase-like (OASL) proteins: dual functions and associations with diseases". In: *Experimental & Molecular Medicine* 47.3. Number: 3 Publisher: Nature Publishing Group, e144–e144. ISSN: 2092-6413. DOI: [10.1038/emm.2014.110](https://doi.org/10.1038/emm.2014.110).
- Ci, Yali et al. (2019). "Zika NS1-induced ER remodeling is essential for viral replication". In: *Journal of Cell Biology* 219.2, e201903062. ISSN: 0021-9525. DOI: [10.1083/jcb.201903062](https://doi.org/10.1083/jcb.201903062).
- Ciancanelli, Michael J et al. (2016). "Host genetics of severe influenza: from mouse Mx1 to human IRF7". In: *Current Opinion in Immunology*. Innate immunity 38, pp. 109–120. ISSN: 0952-7915. DOI: [10.1016/j.coi.2015.12.002](https://doi.org/10.1016/j.coi.2015.12.002).
- Collaborative Cross Consortium (2012). "The Genome Architecture of the Collaborative Cross Mouse Genetic Reference Population". In: *Genetics* 190.2, pp. 389–401. ISSN: 1943-2631. DOI: [10.1534/genetics.111.132639](https://doi.org/10.1534/genetics.111.132639).
- Crow, Yanick J. and Daniel B. Stetson (2022). "The type I interferonopathies: 10 years on". In: *Nature Reviews Immunology* 22.8. Number: 8 Publisher: Nature Publishing Group, pp. 471–483. ISSN: 1474-1741. DOI: [10.1038/s41577-021-00633-9](https://doi.org/10.1038/s41577-021-00633-9).
- Cumberworth, Stephanie L. et al. (2017). "Inhibition of type I interferon induction and signalling by mosquito-borne flaviviruses". In: *Cellular Microbiology* 19.5, e12737. ISSN: 1462-5814. DOI: [10.1111/cmi.12737](https://doi.org/10.1111/cmi.12737).
- Cummins, Nathan and Andrew Badley (2009). "The TRAIL to Viral Pathogenesis: The Good, the Bad and the Ugly". In: *Current molecular medicine* 9.4, pp. 495–505. ISSN: 1566-5240.
- Daffis, S. et al. (2007). "Cell-specific IRF-3 responses protect against West Nile virus infection by interferon-dependent and -independent mechanisms". In: *PLoS Pathog* 3.7, e106. ISSN: 1553-7374 (Electronic) 1553-7366 (Linking). DOI: [10.1371/journal.ppat.0030106](https://doi.org/10.1371/journal.ppat.0030106).
- Das, Tandrita et al. (2021). "S-palmitoylation and sterol interactions mediate antiviral specificity of IFITM isoforms". In: *Research Square*, rs.3.rs-1179000. DOI: [10.21203/rs.3.rs-1179000/v1](https://doi.org/10.21203/rs.3.rs-1179000/v1).
- De O. da Silva, Luciana Reboredo et al. (2021). "Zika Virus Congenital Syndrome and MTOR gene variants: insights from a family of dizygotic twins". In: *Heliyon* 7.4, e06878. ISSN: 24058440. DOI: [10.1016/j.heliyon.2021.e06878](https://doi.org/10.1016/j.heliyon.2021.e06878).
- Dinnon, Kenneth H. et al. (2020). "A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures". In: *Nature* 586.7830. Number: 7830 Publisher: Nature Publishing Group, pp. 560–566. ISSN: 1476-4687. DOI: [10.1038/s41586-020-2708-8](https://doi.org/10.1038/s41586-020-2708-8).



- DiPetrillo, Keith et al. (2005). "Bioinformatics toolbox for narrowing rodent quantitative trait loci". In: *Trends in Genetics* 21.12, pp. 683–692. ISSN: 01689525. DOI: [10.1016/j.tig.2005.09.008](https://doi.org/10.1016/j.tig.2005.09.008).
- Dong, Wenjuan et al. (2022). "The K18-Human ACE2 Transgenic Mouse Model Recapitulates Non-severe and Severe COVID-19 in Response to an Infectious Dose of the SARS-CoV-2 Virus". In: *Journal of Virology* 96.1. Ed. by Tom Gallagher, e00964–21. ISSN: 0022-538X, 1098-5514. DOI: [10.1128/JVI.00964-21](https://doi.org/10.1128/JVI.00964-21).
- Donnelly, Raymond P. and Sergei V. Kotenko (2010). "Interferon-Lambda: A New Addition to an Old Family". In: *Journal of Interferon & Cytokine Research* 30.8, pp. 555–564. ISSN: 1079-9907, 1557-7465. DOI: [10.1089/jir.2010.0078](https://doi.org/10.1089/jir.2010.0078).
- Duffy, Mark R. et al. (2009). "Zika Virus Outbreak on Yap Island, Federated States of Micronesia". In: *New England Journal of Medicine* 360.24, pp. 2536–2543. ISSN: 0028-4793, 1533-4406. DOI: [10.1056/NEJMoa0805715](https://doi.org/10.1056/NEJMoa0805715).
- Duffy, Siobain (2018). "Why are RNA virus mutation rates so damn high?" In: *PLOS Biology* 16.8. Publisher: Public Library of Science, e3000003. ISSN: 1545-7885. DOI: [10.1371/journal.pbio.3000003](https://doi.org/10.1371/journal.pbio.3000003).
- Dyck, Christopher H. van et al. (2023). "Lecanemab in Early Alzheimer's Disease". In: *New England Journal of Medicine* 388.1. Publisher: Massachusetts Medical Society \_eprint: <https://doi.org/10.1056/NEJMoa2212948>, pp. 9–21. ISSN: 0028-4793. DOI: [10.1056/NEJMoa2212948](https://doi.org/10.1056/NEJMoa2212948).
- Fay, Rachel L., Alexander C. Keyel, and Alexander T. Ciota (2022). "Chapter Three - West Nile virus and climate change". In: *Advances in Virus Research*. Ed. by Marilyn J. Roossinck. Vol. 114. Viruses and Climate Change. Academic Press, pp. 147–193. DOI: [10.1016/bs.aivir.2022.08.002](https://doi.org/10.1016/bs.aivir.2022.08.002).
- Fernandez-Gonzalez, Angeles et al. (2009). "Mutation of Murine Adenylate Kinase 7 Underlies a Primary Ciliary Dyskinesia Phenotype". In: *American Journal of Respiratory Cell and Molecular Biology* 40.3, pp. 305–313. ISSN: 1044-1549, 1535-4989. DOI: [10.1165/rcmb.2008-0102OC](https://doi.org/10.1165/rcmb.2008-0102OC).
- Ferris, Martin T. et al. (2013). "Modeling host genetic regulation of influenza pathogenesis in the collaborative cross". In: *PLoS pathogens* 9.2, e1003196. ISSN: 1553-7374. DOI: [10.1371/journal.ppat.1003196](https://doi.org/10.1371/journal.ppat.1003196).
- Fujiwara, Shigeyoshi (2018). "Humanized mice: A brief overview on their diverse applications in biomedical research". In: *Journal of Cellular Physiology* 233.4, pp. 2889–2901. ISSN: 1097-4652. DOI: [10.1002/jcp.26022](https://doi.org/10.1002/jcp.26022).
- Gambino, Frank et al. (2021). "A vaccine inducing solely cytotoxic T lymphocytes fully prevents Zika virus infection and fetal damage". In: *Cell Reports* 35.6, p. 109107. ISSN: 22111247. DOI: [10.1016/j.celrep.2021.109107](https://doi.org/10.1016/j.celrep.2021.109107).
- Garifulin, Oleg et al. (2007). "Irf3 Polymorphism Alters Induction of Interferon Beta in Response to *Listeria monocytogenes* Infection". In: *PLOS Genetics* 3.9, e152. ISSN: 1553-7404. DOI: [10.1371/journal.pgen.0030152](https://doi.org/10.1371/journal.pgen.0030152).
- Gebhardt, Anna, Beatrice T. Laudenbach, and Andreas Pichlmair (2017). "Discrimination of Self and Non-Self Ribonucleic Acids". In: *Journal of Interferon & Cytokine Research* 37.5. Publisher: Mary Ann Liebert, Inc., publishers, pp. 184–197. ISSN: 1079-9907. DOI: [10.1089/jir.2016.0092](https://doi.org/10.1089/jir.2016.0092).
- Génin, Pierre et al. (2000). "Regulation of RANTES Chemokine Gene Expression Requires Cooperativity Between NF- $\kappa$ B and IFN-Regulatory Factor Transcription Factors". In: *The Journal of Immunology* 164.10, pp. 5352–5361. ISSN: 0022-1767, 1550-6606. DOI: [10.4049/jimmunol.164.10.5352](https://doi.org/10.4049/jimmunol.164.10.5352).
- Ghnaim, Aya et al. (2023). "Unraveling the Host Genetic Background Effect on Internal Organ Weight Influenced by Obesity and Diabetes Using Collaborative Cross Mice". In: *International Journal of Molecular Sciences* 24.9, p. 8201. ISSN: 1422-0067. DOI: [10.3390/ijms24098201](https://doi.org/10.3390/ijms24098201).
- Giron, Sandra et al. (2019). "Vector-borne transmission of Zika virus in Europe, southern France, August 2019". In: *Eurosurveillance* 24.45. Publisher: European Centre for Disease Prevention and Control, p. 1900655. ISSN: 1560-7917. DOI: [10.2807/1560-7917.ES.2019.24.45.1900655](https://doi.org/10.2807/1560-7917.ES.2019.24.45.1900655).
- Gomes, Julia A. et al. (2023). "Investigation of the impact of *AXL*, *TLR3*, and *STAT2* in congenital Zika syndrome through genetic polymorphisms and protein–protein interaction network analyses". In: *Birth Defects Research*, bdr2.2232. ISSN: 2472-1727, 2472-1727. DOI: [10.1002/bdr2.2232](https://doi.org/10.1002/bdr2.2232).
- Gorman, Matthew J. et al. (2018). "An Immunocompetent Mouse Model of Zika Virus Infection". In: *Cell Host & Microbe* 23.5, 672–685.e6. ISSN: 19313128. DOI: [10.1016/j.chom.2018.04.003](https://doi.org/10.1016/j.chom.2018.04.003).
- Gralinski, L. E. et al. (2017). "Allelic Variation in the Toll-Like Receptor Adaptor Protein Ticam2 Contributes to SARS-Coronavirus Pathogenesis in Mice". In: *G3 (Bethesda)* 7.6, pp. 1653–1663. ISSN: 2160-1836 (Electronic) 2160-1836 (Linking). DOI: [10.1534/g3.117.041434](https://doi.org/10.1534/g3.117.041434).
- Gralinski, Lisa E. et al. (2015). "Genome Wide Identification of SARS-CoV Susceptibility Loci Using the Collaborative Cross". In: *PLoS genetics* 11.10, e1005504. ISSN: 1553-7404. DOI: [10.1371/journal.pgen.1005504](https://doi.org/10.1371/journal.pgen.1005504).
- Green, Richard et al. (2017). "Oas1b-dependent Immune Transcriptional Profiles of West Nile Virus Infection in the Collaborative Cross". In: *G3 (Bethesda, Md.)* 7.6, pp. 1665–1682. ISSN: 2160-1836. DOI: [10.1534/g3.117.041624](https://doi.org/10.1534/g3.117.041624).
- Guo, Moujian et al. (2021). "ZIKV viral proteins and their roles in virus-host interactions". In: *Science China Life Sciences* 64.5, pp. 709–719. ISSN: 1674-7305, 1869-1889. DOI: [10.1007/s11427-020-1818-4](https://doi.org/10.1007/s11427-020-1818-4).
- Hait, Alon Schneider et al. (2021). "Whole-Exome Sequencing of Patients With Recurrent HSV-2 Lymphocytic Mollaret Meningitis". In: *The Journal of Infectious Diseases* 223.10, pp. 1776–1786. ISSN: 0022-1899, 1537-6613. DOI: [10.1093/infdis/jiaa589](https://doi.org/10.1093/infdis/jiaa589).
- Hall, Bradford et al. (2018). "Genome Editing in Mice Using CRISPR/Cas9 Technology". In: *Current protocols in cell biology* 81.1, e57. ISSN: 1934-2500. DOI: [10.1002/cpcb.57](https://doi.org/10.1002/cpcb.57).
- Haller, Otto and Georg Kochs (2020). "Mx genes: host determinants controlling influenza virus infection and trans-species transmission". In: *Human Genetics* 139.6, pp. 695–705. ISSN: 0340-6717, 1432-1203. DOI: [10.1007/s00439-019-02092-8](https://doi.org/10.1007/s00439-019-02092-8).
- Han, Yuling et al. (2022). "A human iPSC-array-based GWAS identifies a virus susceptibility locus in the NDUFA4 gene and functional variants". In: *Cell Stem Cell* 29.10, 1475–1490.e6. ISSN: 19345909. DOI: [10.1016/j.stem.2022.09.008](https://doi.org/10.1016/j.stem.2022.09.008).
- Hasan, S. Saif et al. (2018). "Structural biology of Zika virus and other flaviviruses". In: *Nature Structural & Molecular Biology* 25.1, pp. 13–20. ISSN: 1545-9985. DOI: [10.1038/s41594-017-0010-8](https://doi.org/10.1038/s41594-017-0010-8).

- Hindorff, Lucia A. et al. (2018). "Prioritizing diversity in human genomics research". In: *Nature Reviews. Genetics* 19.3, pp. 175–185. ISSN: 1471-0064. DOI: [10.1038/nrg.2017.89](https://doi.org/10.1038/nrg.2017.89).
- Horisberger, M A, P Staeheli, and O Haller (1983). "Interferon induces a unique protein in mouse cells bearing a gene for resistance to influenza virus." In: *Proceedings of the National Academy of Sciences* 80.7. Publisher: Proceedings of the National Academy of Sciences, pp. 1910–1914. DOI: [10.1073/pnas.80.7.1910](https://doi.org/10.1073/pnas.80.7.1910).
- Hu, Huan, Yaxiu Feng, and Ming-Liang He (2023). "Targeting Type I Interferon Induction and Signaling: How Zika Virus Escapes from Host Innate Immunity". In: *International Journal of Biological Sciences* 19.10, pp. 3015–3028. ISSN: 1449-2288. DOI: [10.7150/ijbs.83056](https://doi.org/10.7150/ijbs.83056).
- Huang, Huarong et al. (2017). "CD8<sup>+</sup> T Cell Immune Response in Immunocompetent Mice during Zika Virus Infection". In: *Journal of Virology* 91.22. Ed. by Michael S. Diamond, e00900–17. ISSN: 0022-538X, 1098-5514. DOI: [10.1128/JVI.00900-17](https://doi.org/10.1128/JVI.00900-17).
- Hung, Su-Jhen and Sheng-Wen Huang (2021). "Contributions of Genetic Evolution to Zika Virus Emergence". In: *Frontiers in Microbiology* 12, p. 655065. ISSN: 1664-302X. DOI: [10.3389/fmicb.2021.655065](https://doi.org/10.3389/fmicb.2021.655065).
- Ikejezie, Juniorcaius (2017). "Zika Virus Transmission — Region of the Americas, May 15, 2015–December 15, 2016". In: *MMWR. Morbidity and Mortality Weekly Report* 66. ISSN: 0149-2195/1545-861X. DOI: [10.15585/mmwr.mm6612a4](https://doi.org/10.15585/mmwr.mm6612a4).
- Inui, Masafumi et al. (2014). "Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system". In: *Scientific Reports* 4.1. Number: 1 Publisher: Nature Publishing Group, p. 5396. ISSN: 2045-2322. DOI: [10.1038/srep05396](https://doi.org/10.1038/srep05396).
- Isaacs, A. and J. Lindenmann (1957). "Virus interference. I. The interferon". In: *Proceedings of the Royal Society of London. Series B, Biological Sciences* 147.927, pp. 258–267. ISSN: 0950-1193. DOI: [10.1098/rspb.1957.0048](https://doi.org/10.1098/rspb.1957.0048).
- Iwasaki, Akiko (2017). "Immune Regulation of Antibody Access to Neuronal Tissues". In: *Trends in Molecular Medicine* 23.3, pp. 227–245. ISSN: 1471-4914. DOI: [10.1016/j.molmed.2017.01.004](https://doi.org/10.1016/j.molmed.2017.01.004).
- Jasperse, Brittany A. et al. (2023). "Neuroinvasive Flavivirus Pathogenesis Is Restricted by Host Genetic Factors in Collaborative Cross Mice, Independently of Oas1b". In: *Journal of Virology*, e0071523. ISSN: 1098-5514. DOI: [10.1128/jvi.00715-23](https://doi.org/10.1128/jvi.00715-23).
- Javed, Farakh et al. (2018). "Zika virus: what we need to know?" In: *Journal of Basic Microbiology* 58.1, pp. 3–16. ISSN: 1521-4028. DOI: [10.1002/jobm.201700398](https://doi.org/10.1002/jobm.201700398).
- Jiménez-Munguía, I. et al. (2022). "Interferon-induced transmembrane protein 3 (IFITM3) and its antiviral activity". In: *Current Opinion in Structural Biology* 77, p. 102467. ISSN: 0959440X. DOI: [10.1016/j.sbi.2022.102467](https://doi.org/10.1016/j.sbi.2022.102467).
- Jin, Ying et al. (2012). "Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo". In: *Nature Genetics* 44.6, pp. 676–680. ISSN: 1061-4036, 1546-1718. DOI: [10.1038/ng.2272](https://doi.org/10.1038/ng.2272).
- Jin, Ying et al. (2016). "Genome-wide association studies of autoimmune vitiligo identify 23 new risk loci and highlight key pathways and regulatory variants". In: *Nature Genetics* 48.11, pp. 1418–1424. ISSN: 1061-4036, 1546-1718. DOI: [10.1038/ng.3680](https://doi.org/10.1038/ng.3680).
- Jo, Ukhyun and Yves Pommier (2022). "Structural, molecular, and functional insights into Schlafen proteins". In: *Experimental & Molecular Medicine* 54.6. Number: 6 Publisher: Nature Publishing Group, pp. 730–738. ISSN: 2092-6413. DOI: [10.1038/s12276-022-00794-0](https://doi.org/10.1038/s12276-022-00794-0).
- Kane, Melissa and Tatyana V. Golovkina (2019). "Mapping Viral Susceptibility Loci in Mice". In: *Annual Review of Virology* 6.1, pp. 525–546. ISSN: 2327-056X, 2327-0578. DOI: [10.1146/annurev-virology-092818-015544](https://doi.org/10.1146/annurev-virology-092818-015544).
- Kann, M, A Bischof, and W H Gerlich (1997). "In vitro model for the nuclear transport of the hepadnavirus genome". In: *Journal of Virology* 71.2. Publisher: American Society for Microbiology, pp. 1310–1316. DOI: [10.1128/jvi.71.2.1310-1316.1997](https://doi.org/10.1128/jvi.71.2.1310-1316.1997).
- Karmakar, Moumita et al. (2023). "Baseline Gait and Motor Function Predict Long-Term Severity of Neurological Outcomes of Viral Infection". In: *International Journal of Molecular Sciences* 24.3, p. 2843. ISSN: 1422-0067. DOI: [10.3390/ijms24032843](https://doi.org/10.3390/ijms24032843).
- Kassambara, Alboukadel (2023). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. R package version 0.7.2.
- Kawai, Taro et al. (2005). "IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction". In: *Nature Immunology* 6.10. Number: 10 Publisher: Nature Publishing Group, pp. 981–988. ISSN: 1529-2916. DOI: [10.1038/ni1243](https://doi.org/10.1038/ni1243).
- Keele, Gregory R (2023). "Which mouse multiparental population is right for your study? The Collaborative Cross inbred strains, their F1 hybrids, or the Diversity Outbred population". In: *G3 Genes | Genomes | Genetics* 13.4, jkad027. ISSN: 2160-1836. DOI: [10.1093/g3journal/jkad027](https://doi.org/10.1093/g3journal/jkad027).
- Kenney, Adam D. et al. (2017). "Human Genetic Determinants of Viral Diseases". In: *Annual Review of Genetics* 51, pp. 241–263. ISSN: 1545-2948. DOI: [10.1146/annurev-genet-120116-023425](https://doi.org/10.1146/annurev-genet-120116-023425).
- Kessler, Susanne et al. (2021). "Influenza A Viruses and Zoonotic Events—Are We Creating Our Own Reservoirs?" In: *Viruses* 13.11, p. 2250. ISSN: 1999-4915. DOI: [10.3390/v13112250](https://doi.org/10.3390/v13112250).
- Kherraf, Zine-Eddine et al. (2018). "Creation of knock out and knock in mice by CRISPR/Cas9 to validate candidate genes for human male infertility, interest, difficulties and feasibility". In: *Molecular and Cellular Endocrinology* 468, pp. 70–80. ISSN: 03037207. DOI: [10.1016/j.mce.2018.03.002](https://doi.org/10.1016/j.mce.2018.03.002).
- Kousathanas, Athanasios et al. (2022). "Whole-genome sequencing reveals host factors underlying critical COVID-19". In: *Nature* 607.7917, pp. 97–103. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/s41586-022-04576-6](https://doi.org/10.1038/s41586-022-04576-6).
- Krishnakumar, Vinodhini et al. (2019). "Recent Updates on Mouse Models for Human Immunodeficiency, Influenza, and Dengue Viral Infections". In: *Viruses* 11.3, p. 252. ISSN: 1999-4915. DOI: [10.3390/v11030252](https://doi.org/10.3390/v11030252).
- Kumagai, Yutaro, Osamu Takeuchi, and Shizuo Akira (2008). "TLR9 as a key receptor for the recognition of DNA". In: *Advanced Drug Delivery Reviews*. Toll-like Receptor and Pattern Sensing for Evoking Immune Response 60.7, pp. 795–804. ISSN: 0169-409X. DOI: [10.1016/j.addr.2007.12.004](https://doi.org/10.1016/j.addr.2007.12.004).
- Kumar, T. Rajendra et al. (2009). "Transgenic Mouse Technology: Principles and Methods". In: *Molecular Endocrinology: Methods and Protocols*. Ed. by Ok-Kyong Park-Sarge and Thomas E. Curry. Methods in Molecular Biology. Totowa, NJ: Humana Press, pp. 335–362. ISBN: 978-1-60327-378-7. DOI: [10.1007/978-1-60327-378-7\\_22](https://doi.org/10.1007/978-1-60327-378-7_22).

- Kwon, Jennifer M. and Alison M. Goate (2000). "The Candidate Gene Approach". In: *Alcohol Research & Health* 24.3, pp. 164–168. ISSN: 1535-7414.
- Lam, Jenny K W et al. (2015). "siRNA Versus miRNA as Therapeutics for Gene Silencing". In: *Molecular Therapy - Nucleic Acids* 4, e252. ISSN: 21622531. DOI: [10.1038/mtna.2015.23](https://doi.org/10.1038/mtna.2015.23).
- Lauretì, Mathilde et al. (2018). "Flavivirus Receptors: Diversity, Identity, and Cell Entry". In: *Frontiers in Immunology* 9, p. 2180. ISSN: 1664-3224. DOI: [10.3389/fimmu.2018.02180](https://doi.org/10.3389/fimmu.2018.02180).
- Lazear, Helen M., John W. Schoggins, and Michael S. Diamond (2019). "Shared and Distinct Functions of Type I and Type III Interferons". In: *Immunity* 50.4, pp. 907–923. ISSN: 1074-7613. DOI: [10.1016/j.immuni.2019.03.025](https://doi.org/10.1016/j.immuni.2019.03.025).
- Lazear, Helen M. et al. (2016). "A Mouse Model of Zika Virus Pathogenesis". In: *Cell Host & Microbe* 19.5, pp. 720–730. ISSN: 1934-6069. DOI: [10.1016/j.chom.2016.03.010](https://doi.org/10.1016/j.chom.2016.03.010).
- Le Clerc, Sigrid et al. (2009). "Genomewide Association Study of a Rapid Progression Cohort Identifies New Susceptibility Alleles for AIDS (ANRS Genomewide Association Study 03)". In: *The Journal of Infectious Diseases* 200.8, pp. 1194–1201. ISSN: 0022-1899. DOI: [10.1086/605892](https://doi.org/10.1086/605892).
- Lee, A. J. and A. A. Ashkar (2018). "The Dual Nature of Type I and Type II Interferons". In: *Front Immunol* 9, p. 2061. ISSN: 1664-3224 (Electronic) 1664-3224 (Linking). DOI: [10.3389/fimmu.2018.02061](https://doi.org/10.3389/fimmu.2018.02061).
- Lee, Cheryl Yi-Pin et al. (2020). "Type I interferon shapes the quantity and quality of the anti-Zika virus antibody response". In: *Clinical & Translational Immunology* 9.4, e1126. ISSN: 2050-0068. DOI: [10.1002/cti2.1126](https://doi.org/10.1002/cti2.1126).
- Leist, Sarah R. and Ralph S. Baric (2018). "Giving the Genes a Shuffle: Using Natural Variation to Understand Host Genetic Contributions to Viral Infections". In: *Trends in Genetics* 34.10, pp. 777–789. ISSN: 01689525. DOI: [10.1016/j.tig.2018.07.005](https://doi.org/10.1016/j.tig.2018.07.005).
- Leist, Sarah R. et al. (2020). "A Mouse-Adapted SARS-CoV-2 Induces Acute Lung Injury and Mortality in Standard Laboratory Mice". In: *Cell* 183.4, 1070–1085.e12. ISSN: 0092-8674. DOI: [10.1016/j.cell.2020.09.050](https://doi.org/10.1016/j.cell.2020.09.050).
- Lenth, Russell V (2023). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.8.7.
- Levy, D. E., I. J. Marie, and J. E. Durbin (2011). "Induction and function of type I and III interferon in response to viral infection". In: *Curr Opin Virol* 1.6, pp. 476–86. ISSN: 1879-6265 (Electronic) 1879-6257 (Linking). DOI: [10.1016/j.coviro.2011.11.001](https://doi.org/10.1016/j.coviro.2011.11.001).
- Li, Manqing et al. (2012). "Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11". In: *Nature* 491.7422, pp. 125–128. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/nature11433](https://doi.org/10.1038/nature11433).
- Li, Qin Hui, Kenneth Kim, and Suján Shresta (2023). "Mouse models of Zika virus transplacental transmission". In: *Antiviral Research*. Special Issue in Honor of Dr. Mike Bray on his retirement as the Editor-in-Chief of Antiviral Research 210, p. 105500. ISSN: 0166-3542. DOI: [10.1016/j.antiviral.2022.105500](https://doi.org/10.1016/j.antiviral.2022.105500).
- Li, Yun et al. (2019). "Genome-wide CRISPR screen for Zika virus resistance in human neural cells". In: *Proceedings of the National Academy of Sciences* 116.19, pp. 9527–9532. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1900867116](https://doi.org/10.1073/pnas.1900867116).
- Liko, Dritan et al. (2019). "Brl1 loss and not overexpression disrupts tissues homeostasis in the intestine, liver and pancreas". In: *Cell Death & Differentiation* 26.12, pp. 2535–2550. ISSN: 1350-9047, 1476-5403. DOI: [10.1038/s41418-019-0316-7](https://doi.org/10.1038/s41418-019-0316-7).
- Loesch, Robin et al. (2022). "Deleting the  $\beta$ -catenin degradation domain in mouse hepatocytes drives hepatocellular carcinoma or hepatoblastoma-like tumor growth". In: *Journal of Hepatology* 77.2, pp. 424–435. ISSN: 0168-8278. DOI: [10.1016/j.jhep.2022.02.023](https://doi.org/10.1016/j.jhep.2022.02.023).
- Lone, Iqbal M. et al. (2023a). "High-fat diet and oral infection induced type 2 diabetes and obesity development under different genetic backgrounds". In: *Animal Models and Experimental Medicine* 6.2, pp. 131–145. ISSN: 2576-2095, 2576-2095. DOI: [10.1002/ame2.12311](https://doi.org/10.1002/ame2.12311).
- Lone, Iqbal M. et al. (2023b). "Intestinal cancer development in response to oral infection with high-fat diet-induced Type 2 diabetes (T2D) in collaborative cross mice under different host genetic background effects". In: *Mammalian Genome*. ISSN: 0938-8990, 1432-1777. DOI: [10.1007/s00335-023-09979-y](https://doi.org/10.1007/s00335-023-09979-y).
- Lorè, Nicola Ivan et al. (2020). "Collaborative Cross Mice Yield Genetic Modifiers for Pseudomonas aeruginosa Infection in Human Lung Disease". In: *mBio* 11.2. Ed. by Joanna B. Goldberg, e00097–20. ISSN: 2161-2129, 2150-7511. DOI: [10.1128/mBio.00097-20](https://doi.org/10.1128/mBio.00097-20).
- Madden, J. C., Dan Cui, and M. A. Brinton (2019). "RNase L Antiviral Activity Is Not a Critical Component of the Oas1b-Mediated Flavivirus Resistance Phenotype". In: *Journal of Virology* 93.22. Publisher: American Society for Microbiology, 10.1128/jvi.00946–19. DOI: [10.1128/jvi.00946-19](https://doi.org/10.1128/jvi.00946-19).
- Majewski, Jacek et al. (2011). "What can exome sequencing do for you?" In: *Journal of Medical Genetics* 48.9, pp. 580–589. ISSN: 0022-2593, 1468-6244. DOI: [10.1136/jmedgenet-2011-100223](https://doi.org/10.1136/jmedgenet-2011-100223).
- Mak, Chloe Miu et al. (2011). "Fatal viral infection-associated encephalopathy in two Chinese boys: a genetically determined risk factor of thermolabile carnitine palmitoyltransferase II variants". In: *Journal of Human Genetics* 56.8. Number: 8 Publisher: Nature Publishing Group, pp. 617–621. ISSN: 1435-232X. DOI: [10.1038/jhg.2011.63](https://doi.org/10.1038/jhg.2011.63).
- Mancini, Maria Vittoria et al. (2023). "Evaluation of an Engineered Zika Virus-Like Particle Vaccine Candidate in a Mosquito-Mouse Transmission Model". In: *mSphere* 8.2. Publisher: American Society for Microbiology, e00564–22. DOI: [10.1128/msphere.00564-22](https://doi.org/10.1128/msphere.00564-22).
- Manet, Caroline (2019). "Genetic control of susceptibility to Zika virus in the mouse using strains of the Collaborative Cross". These de doctorat. Paris, Institut agronomique, vétérinaire et forestier de France.
- Manet, Caroline et al. (2018). "Host genetic control of mosquito-borne Flavivirus infections". In: *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 29.7, pp. 384–407. ISSN: 1432-1777. DOI: [10.1007/s00335-018-9775-2](https://doi.org/10.1007/s00335-018-9775-2).
- Manet, Caroline et al. (2020). "Genetic Diversity of Collaborative Cross Mice Controls Viral Replication, Clinical Severity, and Brain Pathology Induced by Zika Virus Infection, Independently of Oas1b". In: *Journal of Virology* 94.3. ISSN: 1098-5514. DOI: [10.1128/JVI.01034-19](https://doi.org/10.1128/JVI.01034-19).

- Marceau, Caleb D. et al. (2016). "Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens". In: *Nature* 535.7610. Number: 7610 Publisher: Nature Publishing Group, pp. 159–163. ISSN: 1476-4687. DOI: [10.1038/nature18631](https://doi.org/10.1038/nature18631).
- Marín-Lopez, Alejandro et al. (2019). "Modeling Arboviral Infection in Mice Lacking the Interferon Alpha/Beta Receptor". In: *Viruses* 11.1, p. 35. ISSN: 1999-4915. DOI: [10.3390/v11010035](https://doi.org/10.3390/v11010035).
- Martin, Matthew D. et al. (2020). "Diverse CD8 T Cell Responses to Viral Infection Revealed by the Collaborative Cross". In: *Cell Reports* 31.2, p. 107508. ISSN: 22111247. DOI: [10.1016/j.celrep.2020.03.072](https://doi.org/10.1016/j.celrep.2020.03.072).
- Mashimo, T., D. Simon-Chazottes, and J.-L. Guénet (2008). "Innate resistance to flavivirus infections and the functions of 2'-5' oligoadenylate synthetases". In: *Current Topics in Microbiology and Immunology* 321, pp. 85–100. ISSN: 0070-217X. DOI: [10.1007/978-3-540-75203-5\\_4](https://doi.org/10.1007/978-3-540-75203-5_4).
- Mashimo, Tomoji et al. (2002). "A nonsense mutation in the gene encoding 2'-5'-oligoadenylate synthetase/L1 isoform is associated with West Nile virus susceptibility in laboratory mice". In: *Proceedings of the National Academy of Sciences* 99.17. Publisher: Proceedings of the National Academy of Sciences, pp. 11311–11316. DOI: [10.1073/pnas.172195399](https://doi.org/10.1073/pnas.172195399).
- Mattocks, Melissa D. et al. (2019). *Zika virus infection in Collaborative Cross mice*. preprint. Microbiology. DOI: [10.1101/695510](https://doi.org/10.1101/695510).
- Matz, Keesha et al. (2021). "Favipiravir (T-705) Protects IFNAR-/- Mice against Lethal Zika Virus Infection in a Sex-Dependent Manner". In: *Microorganisms* 9.6, p. 1178. ISSN: 2076-2607. DOI: [10.3390/microorganisms9061178](https://doi.org/10.3390/microorganisms9061178).
- McCray, Paul B. et al. (2007). "Lethal Infection of K18-hACE2 Mice Infected with Severe Acute Respiratory Syndrome Coronavirus". In: *Journal of Virology* 81.2. Publisher: American Society for Microbiology, pp. 813–821. DOI: [10.1128/jvi.02012-06](https://doi.org/10.1128/jvi.02012-06).
- McLaren, William et al. (2016). "The Ensembl Variant Effect Predictor". In: *Genome Biology* 17.1, p. 122. ISSN: 1474-760X. DOI: [10.1186/s13059-016-0974-4](https://doi.org/10.1186/s13059-016-0974-4).
- Montagutelli, Xavier and Marie Abitbol (2004). "Utilisation des lignées congéniques chez la souris". In: *médecine/sciences* 20.10. Number: 10 Publisher: EDK, pp. 887–893. ISSN: 0767-0974, 1958-5381. DOI: [10.1051/medsci/20042010887](https://doi.org/10.1051/medsci/20042010887).
- Montagutelli, Xavier et al. (2021a). *A mouse-adapted SARS-CoV-2 strain replicating in standard laboratory mice*. preprint. Microbiology. DOI: [10.1101/2021.07.10.451880](https://doi.org/10.1101/2021.07.10.451880).
- Montagutelli, Xavier et al. (2021b). *The B1.351 and P.1 variants extend SARS-CoV-2 host range to mice*. Pages: 2021.03.18.436013 Section: New Results. DOI: [10.1101/2021.03.18.436013](https://doi.org/10.1101/2021.03.18.436013).
- Morel, L et al. (1997). "Functional dissection of systemic lupus erythematosus using congenic mouse strains." In: *The Journal of Immunology* 158.12, pp. 6019–6028. ISSN: 0022-1767, 1550-6606. DOI: [10.4049/jimmunol.158.12.6019](https://doi.org/10.4049/jimmunol.158.12.6019).
- Morel, Laurence et al. (2000). "Genetic reconstitution of systemic lupus erythematosus immunopathology with polycongenic murine strains". In: *Proceedings of the National Academy of Sciences* 97.12, pp. 6670–6675. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.97.12.6670](https://doi.org/10.1073/pnas.97.12.6670).
- Morel, Laurence et al. (2001). "The major murine systemic lupus erythematosus susceptibility locus, *Sle1*, is a cluster of functionally related genes". In: *Proceedings of the National Academy of Sciences* 98.4, pp. 1787–1792. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.98.4.1787](https://doi.org/10.1073/pnas.98.4.1787).
- Moreno-Fernandez, Maria E. et al. (2011). "Regulatory T cells control HIV replication in activated T cells through a cAMP-dependent mechanism". In: *Blood* 117.20, pp. 5372–5380. ISSN: 0006-4971. DOI: [10.1182/blood-2010-12-323162](https://doi.org/10.1182/blood-2010-12-323162).
- Moritoh, Kanako et al. (2009). "Generation of congenic mouse strains by introducing the virus-resistant genes, *Mx1* and *Oas1b*, of feral mouse-derived inbred strain MSM/Ms into the common strain C57BL/6J". In: *The Japanese Journal of Veterinary Research* 57.2, pp. 89–99. ISSN: 0047-1917.
- Morris, David L et al. (2016). "Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus". In: *Nature Genetics* 48.8, pp. 940–946. ISSN: 1061-4036, 1546-1718. DOI: [10.1038/ng.3603](https://doi.org/10.1038/ng.3603).
- Morrison, Thomas E. and Michael S. Diamond (2017). "Animal Models of Zika Virus Infection, Pathogenesis, and Immunity". In: *Journal of Virology* 91.8. Ed. by Ted C. Pierson, e00009–17. ISSN: 0022-538X, 1098-5514. DOI: [10.1128/JVI.00009-17](https://doi.org/10.1128/JVI.00009-17).
- Mostafavi, Sara et al. (2016). "Parsing the Interferon Transcriptional Network and Its Disease Associations". In: *Cell* 164.3, pp. 564–578. ISSN: 0092-8674. DOI: [10.1016/j.cell.2015.12.032](https://doi.org/10.1016/j.cell.2015.12.032).
- Mott, Richard and Jonathan Flint (2013). "Dissecting Quantitative Traits in Mice". In: *Annual Review of Genomics and Human Genetics* 14.1, pp. 421–439. ISSN: 1527-8204, 1545-293X. DOI: [10.1146/annurev-genom-091212-153419](https://doi.org/10.1146/annurev-genom-091212-153419).
- Müller, Ulrike et al. (1994). "Functional Role of Type I and Type II Interferons in Antiviral Defense". In: *Science* 264.5167, pp. 1918–1921. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.8009221](https://doi.org/10.1126/science.8009221).
- Musso, Didier and Duane J. Gubler (2016). "Zika Virus". In: *Clinical Microbiology Reviews* 29.3. Publisher: American Society for Microbiology, pp. 487–524. DOI: [10.1128/cmr.00072-15](https://doi.org/10.1128/cmr.00072-15).
- Najafi Fard, Saeid et al. (2021). "In Vitro Models for Studying Entry, Tissue Tropism, and Therapeutic Approaches of Highly Pathogenic Coronaviruses". In: *BioMed Research International* 2021. Publisher: Hindawi, e8856018. ISSN: 2314-6133. DOI: [10.1155/2021/8856018](https://doi.org/10.1155/2021/8856018).
- Naqvi, Mariam et al. (2007). "Ethnic-specific differences in bronchodilator responsiveness among African Americans, Puerto Ricans, and Mexicans with asthma". In: *The Journal of Asthma: Official Journal of the Association for the Care of Asthma* 44.8, pp. 639–648. ISSN: 1532-4303. DOI: [10.1080/02770900701554441](https://doi.org/10.1080/02770900701554441).
- Nashef, A. et al. (2018). "Integration of Murine and Human Studies for Mapping Periodontitis Susceptibility". In: *Journal of Dental Research* 97.5, pp. 537–546. ISSN: 0022-0345, 1544-0591. DOI: [10.1177/0022034517744189](https://doi.org/10.1177/0022034517744189).
- Navarro, Lorena et al. (1998). "Cytomegalovirus Activates Interferon Immediate-Early Response Gene Expression and an Interferon Regulatory Factor 3-Containing Interferon-Stimulated Response Element-Binding Complex". In: *Molecular and Cellular Biology* 18.7, pp. 3796–3802. ISSN: 0270-7306, 1098-5549. DOI: [10.1128/MCB.18.7.3796](https://doi.org/10.1128/MCB.18.7.3796).

- Nazneen, Farzana et al. (2023). "An effective live-attenuated Zika vaccine candidate with a modified 5' untranslated region". In: *npj Vaccines* 8.1. Number: 1 Publisher: Nature Publishing Group, pp. 1–12. ISSN: 2059-0105. DOI: [10.1038/s41541-023-00650-w](https://doi.org/10.1038/s41541-023-00650-w).
- Nedelko, Tatiana et al. (2012). "Distinct gene loci control the host response to influenza H1N1 virus infection in a time-dependent manner". In: *BMC Genomics* 13.1, p. 411. ISSN: 1471-2164. DOI: [10.1186/1471-2164-13-411](https://doi.org/10.1186/1471-2164-13-411).
- Neil, Stuart J. D., Trinity Zang, and Paul D. Bieniasz (2008). "Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu". In: *Nature* 451.7177. Number: 7177 Publisher: Nature Publishing Group, pp. 425–430. ISSN: 1476-4687. DOI: [10.1038/nature06553](https://doi.org/10.1038/nature06553).
- Nellis, Mary et al. (2021). "Lung metabolome of 1,3-butadiene exposed Collaborative Cross mice reflects metabolic phenotype of human lung cancer". In: *Toxicology* 463, p. 152987. ISSN: 0300483X. DOI: [10.1016/j.tox.2021.152987](https://doi.org/10.1016/j.tox.2021.152987).
- Niemi, Mari E. K., Mark J. Daly, and Andrea Ganna (2022). "The human genetic epidemiology of COVID-19". In: *Nature Reviews Genetics* 23.9, pp. 533–546. ISSN: 1471-0056, 1471-0064. DOI: [10.1038/s41576-022-00478-5](https://doi.org/10.1038/s41576-022-00478-5).
- Noll, Kelsey E., Martin T. Ferris, and Mark T. Heise (2019). "The Collaborative Cross: A Systems Genetics Resource for Studying Host-Pathogen Interactions". In: *Cell Host & Microbe* 25.4, pp. 484–498. ISSN: 1934-6069. DOI: [10.1016/j.chom.2019.03.009](https://doi.org/10.1016/j.chom.2019.03.009).
- O'Dea, Ellen L et al. (2007). "A homeostatic model of I $\kappa$ B metabolism to control constitutive NF- $\kappa$ B activity". In: *Molecular Systems Biology* 3, p. 111. ISSN: 1744-4292. DOI: [10.1038/msb4100148](https://doi.org/10.1038/msb4100148).
- Oh, Hyun-Seok et al. (2021). "A purified inactivated vaccine derived from Vero cell-adapted zika virus elicits protection in mice". In: *Virology* 560, pp. 124–130. ISSN: 00426822. DOI: [10.1016/j.virol.2021.05.003](https://doi.org/10.1016/j.virol.2021.05.003).
- Oh, Sam S. et al. (2015). "Diversity in Clinical and Biomedical Research: A Promise Yet to Be Fulfilled". In: *PLOS Medicine* 12.12. Publisher: Public Library of Science, e1001918. ISSN: 1549-1676. DOI: [10.1371/journal.pmed.1001918](https://doi.org/10.1371/journal.pmed.1001918).
- Orozco, Susana et al. (2012). "Characterization of a model of lethal dengue virus 2 infection in C57BL/6 mice deficient in the alpha/beta interferon receptor". In: *Journal of General Virology* 93.10, pp. 2152–2157. ISSN: 0022-1317, 1465-2099. DOI: [10.1099/vir.0.045088-0](https://doi.org/10.1099/vir.0.045088-0).
- Panthier, Jean-Jacques and Xavier Montagutelli (2012). "Le Collaborative Cross: Un outil révolutionnaire à l'assaut des caractères complexes". In: *médecine/sciences* 28.1, pp. 103–108. ISSN: 0767-0974, 1958-5381. DOI: [10.1051/medsci/2012281024](https://doi.org/10.1051/medsci/2012281024).
- Pardo Manuel De Villena, Fernando et al. (2023). "Fixation of de novo ERV insertions lead to emergent strain-dependent phenotypes in the CC population". Poster. International Mammalian Genome Conference 2023. Tsukuba.
- Pare, Guillaume et al. (2020). "Genetic risk for dengue hemorrhagic fever and dengue fever in multiple ancestries". In: *EBioMedicine* 51, p. 102584. ISSN: 23523964. DOI: [10.1016/j.ebiom.2019.11.045](https://doi.org/10.1016/j.ebiom.2019.11.045).
- Passmore, Susan Racine et al. (2022). "There's not much we can do..." researcher-level barriers to the inclusion of underrepresented participants in translational research". In: *Journal of Clinical and Translational Science* 6.1, e4. ISSN: 2059-8661. DOI: [10.1017/cts.2021.876](https://doi.org/10.1017/cts.2021.876).
- Pattnaik, Aryamav et al. (2023). "A Ferritin Nanoparticle-Based Zika Virus Vaccine Candidate Induces Robust Humoral and Cellular Immune Responses and Protects Mice from Lethal Virus Challenge". In: *Vaccines* 11.4. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, p. 821. ISSN: 2076-393X. DOI: [10.3390/vaccines11040821](https://doi.org/10.3390/vaccines11040821).
- Paul, Alvin, Thean Hock Tang, and Siew Kit Ng (2018). "Interferon Regulatory Factor 9 Structure and Regulation". In: *Frontiers in Immunology* 9. ISSN: 1664-3224.
- Peirce, Jeremy L et al. (2004). "A new set of BXD recombinant inbred lines from advanced intercross populations in mice". In: *BMC Genetics* 5.1, p. 7. ISSN: 14712156. DOI: [10.1186/1471-2156-5-7](https://doi.org/10.1186/1471-2156-5-7).
- Perlman, Robert L. (2016). "Mouse models of human disease". In: *Evolution, Medicine, and Public Health* 2016.1, pp. 170–176. ISSN: 2050-6201. DOI: [10.1093/emph/eow014](https://doi.org/10.1093/emph/eow014).
- Phifer-Rixey, Megan and Michael W Nachman (2015). "Insights into mammalian biology from the wild house mouse *Mus musculus*". In: *eLife* 4. Publisher: eLife Sciences Publications, Ltd, e05959. ISSN: 2050-084X. DOI: [10.7554/eLife.05959](https://doi.org/10.7554/eLife.05959).
- Piasecka, Barbara et al. (2018). "Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges". In: *Proceedings of the National Academy of Sciences* 115.3. Publisher: Proceedings of the National Academy of Sciences, E488–E497. DOI: [10.1073/pnas.1714765115](https://doi.org/10.1073/pnas.1714765115).
- Pierson, Theodore C. and Michael S. Diamond (2018). "The emergence of Zika virus and its new clinical syndromes". In: *Nature* 560.7720, pp. 573–581. ISSN: 1476-4687. DOI: [10.1038/s41586-018-0446-y](https://doi.org/10.1038/s41586-018-0446-y).
- Piganis, Rebecca A. R. et al. (2011). "Suppressor of Cytokine Signaling (SOCS) 1 Inhibits Type I Interferon (IFN) Signaling via the Interferon  $\alpha$  Receptor (IFNAR1)-associated Tyrosine Kinase Tyk2". In: *Journal of Biological Chemistry* 286.39, pp. 33811–33818. ISSN: 0021-9258. DOI: [10.1074/jbc.M111.270207](https://doi.org/10.1074/jbc.M111.270207).
- Popli, Sonam et al. (2022). "IRF3 inhibits nuclear translocation of NF- $\kappa$ B to prevent viral inflammation". In: *Proceedings of the National Academy of Sciences* 119.37, e2121385119. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.2121385119](https://doi.org/10.1073/pnas.2121385119).
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*.
- Rabbani, Bahareh, Mustafa Tekin, and Nejat Mahdih (2014). "The promise of whole-exome sequencing in medical genetics". In: *Journal of Human Genetics* 59.1, pp. 5–15. ISSN: 1434-5161, 1435-232X. DOI: [10.1038/jhg.2013.114](https://doi.org/10.1038/jhg.2013.114).
- Raghunandan, Ramadevi et al. (2008). "Dlk1 Influences Differentiation and Function of  $\beta$  Lymphocytes". In: *Stem Cells and Development* 17.3, pp. 495–508. ISSN: 1547-3287, 1557-8534. DOI: [10.1089/scd.2007.0102](https://doi.org/10.1089/scd.2007.0102).
- Rao, Donald D. et al. (2009). "siRNA vs. shRNA: Similarities and differences". In: *Advanced Drug Delivery Reviews. Towards Therapeutic Application of RNA-Mediated Gene Regulation* 61.9, pp. 746–759. ISSN: 0169-409X. DOI: [10.1016/j.addr.2009.04.004](https://doi.org/10.1016/j.addr.2009.04.004).
- Rapp, John P. and Bina Joe (2012). "Use of contiguous congenic strains in analyzing compound QTLs". In: *Physiological Genomics* 44.2, pp. 117–120. ISSN: 1094-8341. DOI: [10.1152/physiolgenomics.00136.2011](https://doi.org/10.1152/physiolgenomics.00136.2011).

- Rather, Irfan A. et al. (2017). "Zika Virus: An Emerging Worldwide Threat". In: *Frontiers in Microbiology* 8, p. 1417. ISSN: 1664-302X. DOI: [10.3389/fmicb.2017.01417](https://doi.org/10.3389/fmicb.2017.01417).
- Reardon, Sara (2023). "Alzheimer's drug trials plagued by lack of racial diversity". In: *Nature* 620.7973. Bandiera\_abtest: a Cg\_type: News Number: 7973 Publisher: Nature Publishing Group Subject\_term: Medical research, Alzheimer's disease, pp. 256–257. DOI: [10.1038/d41586-023-02464-1](https://doi.org/10.1038/d41586-023-02464-1).
- Reeves, M. B. et al. (2005). "An in vitro model for the regulation of human cytomegalovirus latency and reactivation in dendritic cells by chromatin remodelling". In: *Journal of General Virology* 86.11. Publisher: Microbiology Society, pp. 2949–2954. ISSN: 1465-2099. DOI: [10.1099/vir.0.81161-0](https://doi.org/10.1099/vir.0.81161-0).
- Roberts, Adam et al. (2007). "The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics". In: *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 18.6, pp. 473–481. ISSN: 0938-8990. DOI: [10.1007/s00335-007-9045-1](https://doi.org/10.1007/s00335-007-9045-1).
- Robertson, Shelly J. et al. (2023). "Genetically diverse mouse models of SARS-CoV-2 infection reproduce clinical variation in type I interferon and cytokine responses in COVID-19". In: *Nature Communications* 14.1, p. 4481. ISSN: 2041-1723. DOI: [10.1038/s41467-023-40076-5](https://doi.org/10.1038/s41467-023-40076-5).
- Rocklöv, Joacim and Robert Dubrow (2020). "Climate change: an enduring challenge for vector-borne disease prevention and control". In: *Nature Immunology* 21.5. Number: 5 Publisher: Nature Publishing Group, pp. 479–483. ISSN: 1529-2916. DOI: [10.1038/s41590-020-0648-y](https://doi.org/10.1038/s41590-020-0648-y).
- Rogala, Allison R. et al. (2014). "The Collaborative Cross as a Resource for Modeling Human Disease: CC011/Unc, a New Mouse Model for Spontaneous Colitis". In: *Mammalian Genome* 25.3, pp. 95–108. ISSN: 1432-1777. DOI: [10.1007/s00335-013-9499-2](https://doi.org/10.1007/s00335-013-9499-2).
- Rossi, Á D. et al. (2019). "Variations in maternal adenylate cyclase genes are associated with congenital Zika syndrome in a cohort from Northeast, Brazil". In: *Journal of Internal Medicine* 285.2, pp. 215–222. ISSN: 1365-2796. DOI: [10.1111/joim.12829](https://doi.org/10.1111/joim.12829).
- Rossi, Shannan L. et al. (2016). "Characterization of a Novel Murine Model to Study Zika Virus". In: *The American Journal of Tropical Medicine and Hygiene* 94.6, pp. 1362–1369. ISSN: 0002-9637, 1476-1645. DOI: [10.4269/ajtmh.16-0111](https://doi.org/10.4269/ajtmh.16-0111).
- Roth, Claude et al. (2019). "A Modified mRNA Vaccine Targeting Immunodominant NS Epitopes Protects Against Dengue Virus Infection in HLA Class I Transgenic Mice". In: *Frontiers in Immunology* 10, p. 1424. ISSN: 1664-3224. DOI: [10.3389/fimmu.2019.01424](https://doi.org/10.3389/fimmu.2019.01424).
- Ryan, Nathan M. et al. (2021). "Onchocerca volvulus bivalent subunit vaccine induces protective immunity in genetically diverse collaborative cross recombinant inbred intercross mice". In: *npj Vaccines* 6.1, p. 17. ISSN: 2059-0105. DOI: [10.1038/s41541-020-00276-2](https://doi.org/10.1038/s41541-020-00276-2).
- Sakuma, Toshie et al. (2009). "Inhibition of Lassa and Marburg Virus Production by Tetherin". In: *Journal of Virology* 83.5. Publisher: American Society for Microbiology, pp. 2382–2385. DOI: [10.1128/jvi.01607-08](https://doi.org/10.1128/jvi.01607-08).
- Santos, Camilla N O et al. (2019). "Association Between Zika Virus Microcephaly in Newborns With the rs3775291 Variant in Toll-Like Receptor 3 and rs1799964 Variant at Tumor Necrosis Factor- $\alpha$  Gene". In: *The Journal of Infectious Diseases* 220.11, pp. 1797–1801. ISSN: 0022-1899, 1537-6613. DOI: [10.1093/infdis/jiz392](https://doi.org/10.1093/infdis/jiz392).
- Sarkar, Sanjay and Mark T. Heise (2019). "Mouse Models as Resources for Studying Infectious Diseases". In: *Clinical Therapeutics* 41.10, pp. 1912–1922. ISSN: 0149-2918. DOI: [10.1016/j.clinthera.2019.08.010](https://doi.org/10.1016/j.clinthera.2019.08.010).
- Sato, Kojiro et al. (2001). "Antiviral response by natural killer cells through TRAIL gene induction by IFN- $\alpha/\beta$ ". In: *European Journal of Immunology* 31.11, pp. 3138–3146. ISSN: 1521-4141. DOI: [10.1002/1521-4141\(200111\)31:11<3138::AID-IMMU3138>3.0.CO;2-B](https://doi.org/10.1002/1521-4141(200111)31:11<3138::AID-IMMU3138>3.0.CO;2-B).
- Sauter, Daniel (2014). "Counteraction of the multifunctional restriction factor tetherin". In: *Frontiers in Microbiology* 5. ISSN: 1664-302X. DOI: [10.3389/fmicb.2014.00163](https://doi.org/10.3389/fmicb.2014.00163).
- Savidis, George et al. (2016). "Identification of Zika Virus and Dengue Virus Dependency Factors using Functional Genomics". In: *Cell Reports* 16.1, pp. 232–246. ISSN: 2211-1247. DOI: [10.1016/j.celrep.2016.06.028](https://doi.org/10.1016/j.celrep.2016.06.028).
- Schäfer, Alexandra et al. (2022). "A Multitrait Locus Regulates Sarbecovirus Pathogenesis". In: *mBio* 13.4. Ed. by John A. Lednický, e01454–22. ISSN: 2150-7511. DOI: [10.1128/mbio.01454-22](https://doi.org/10.1128/mbio.01454-22).
- Schilte, Clémentine et al. (2012). "Cutting Edge: Independent Roles for IRF-3 and IRF-7 in Hematopoietic and Nonhematopoietic Cells during Host Response to Chikungunya Infection". In: *The Journal of Immunology* 188.7, pp. 2967–2971. ISSN: 0022-1767, 1550-6606. DOI: [10.4049/jimmunol.1103185](https://doi.org/10.4049/jimmunol.1103185).
- Schoenrock, Sarah A. et al. (2022). "The collaborative cross strains and their founders vary widely in cocaine-induced behavioral sensitization". In: *Frontiers in Behavioral Neuroscience* 16, p. 886524. ISSN: 1662-5153. DOI: [10.3389/fnbeh.2022.886524](https://doi.org/10.3389/fnbeh.2022.886524).
- Schoggins, J. W. (2019). "Interferon-Stimulated Genes: What Do They All Do?" In: *Annu Rev Virol* 6.1, pp. 567–584. ISSN: 2327-0578 (Electronic) 2327-056X (Linking). DOI: [10.1146/annurev-virology-092818-015756](https://doi.org/10.1146/annurev-virology-092818-015756).
- Schoggins, John. W. and C. M. Rice (2011). "Interferon-stimulated genes and their antiviral effector functions". In: *Current Opinion in Virology* 1.6, pp. 519–525. ISSN: 18796257. DOI: [10.1016/j.coviro.2011.10.008](https://doi.org/10.1016/j.coviro.2011.10.008).
- Schott, Benjamin H. et al. (2022). "Single-cell genome-wide association reveals that a nonsynonymous variant in ERAP1 confers increased susceptibility to influenza virus". In: *Cell Genomics* 2.11, p. 100207. ISSN: 2666-979X. DOI: [10.1016/j.xgen.2022.100207](https://doi.org/10.1016/j.xgen.2022.100207).
- Sharp, Paul M. and Beatrice H. Hahn (2011). "Origins of HIV and the AIDS Pandemic". In: *Cold Spring Harbor Perspectives in Medicine*: 1.1, a006841. ISSN: 2157-1422. DOI: [10.1101/cshperspect.a006841](https://doi.org/10.1101/cshperspect.a006841).
- Shaw, Andrew E. et al. (2017). "Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses". In: *PLOS Biology* 15.12. Publisher: Public Library of Science, e2004086. ISSN: 1545-7885. DOI: [10.1371/journal.pbio.2004086](https://doi.org/10.1371/journal.pbio.2004086).

- Sheehan, Kathleen C.F. et al. (2006). "Blocking Monoclonal Antibodies Specific for Mouse IFN-  $\alpha$  /  $\beta$  Receptor Subunit 1 (IFNAR-1) from Mice Immunized by *In Vivo* Hydrodynamic Transfection". In: *Journal of Interferon & Cytokine Research* 26.11, pp. 804–819. ISSN: 1079-9907, 1557-7465. DOI: [10.1089/jir.2006.26.804](https://doi.org/10.1089/jir.2006.26.804).
- Shorter, John R et al. (2017). "Male Infertility Is Responsible for Nearly Half of the Extinction Observed in the Mouse Collaborative Cross". In: *Genetics* 206.2, pp. 557–572. ISSN: 1943-2631. DOI: [10.1534/genetics.116.199596](https://doi.org/10.1534/genetics.116.199596).
- Sirohi, Devika and Richard J Kuhn (2017). "Zika Virus Structure, Maturation, and Receptors". In: *The Journal of Infectious Diseases* 216 (suppl\_10), S935–S944. ISSN: 0022-1899, 1537-6613. DOI: [10.1093/infdis/jix515](https://doi.org/10.1093/infdis/jix515).
- Słomian, Dawid et al. (2023). "Better safe than sorry—Whole-genome sequencing indicates that missense variants are significant in susceptibility to COVID-19". In: *PLOS ONE* 18.1. Ed. by Chika Kingsley Onwuamah, e0279356. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0279356](https://doi.org/10.1371/journal.pone.0279356).
- Smith, Clare M. et al. (2019). "Functionally Overlapping Variants Control Tuberculosis Susceptibility in Collaborative Cross Mice". In: *mBio* 10.6. Ed. by Russell Vance, e02791–19. ISSN: 2161-2129, 2150-7511. DOI: [10.1128/mBio.02791-19](https://doi.org/10.1128/mBio.02791-19).
- Smith, E.M. et al. (2022). "Reviewing fair subject selection considerations for the unique case of post sequelae COVID-19 translational studies". In: *Journal of Clinical and Translational Science* 6.1, e91. ISSN: 2059-8661. DOI: [10.1017/cts.2022.425](https://doi.org/10.1017/cts.2022.425).
- Song, Hong Ki et al. (2012). "Deep RNA sequencing reveals novel cardiac transcriptomic signatures for physiological and pathological hypertrophy". In: *PloS One* 7.4, e35552. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0035552](https://doi.org/10.1371/journal.pone.0035552).
- Sookaromdee, Pathom and Viroj Siwanitkit (2018). "Homosexual and Zika virus infection: Is it a new start of new problem for gay?" In: *Indian Journal of Sexually Transmitted Diseases and AIDS* 39.1. Company: Medknow Publications and Media Pvt. Ltd. Distributor: Medknow Publications and Media Pvt. Ltd. Institution: Medknow Publications and Media Pvt. Ltd. Label: Medknow Publications and Media Pvt. Ltd. Publisher: Medknow Publications, p. 61. ISSN: 2589-0557. DOI: [10.4103/ijstd.ijstd\\_65\\_16](https://doi.org/10.4103/ijstd.ijstd_65_16).
- Srivastava, Anuj et al. (2017). "Genomes of the Mouse Collaborative Cross". In: *Genetics* 206.2, pp. 537–556. ISSN: 1943-2631. DOI: [10.1534/genetics.116.198838](https://doi.org/10.1534/genetics.116.198838).
- Srivastava, Barkha et al. (2009). "Host Genetic Background Strongly Influences the Response to Influenza A Virus Infections". In: *PLOS ONE* 4.3. Publisher: Public Library of Science, e4857. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0004857](https://doi.org/10.1371/journal.pone.0004857).
- Staheli, Peter et al. (1988). "Influenza Virus-Susceptible Mice Carry *Mx* Genes with a Large Deletion or a Nonsense Mutation". In: *Molecular and Cellular Biology* 8.10, pp. 4518–4523. ISSN: 1098-5549. DOI: [10.1128/mcb.8.10.4518-4523.1988](https://doi.org/10.1128/mcb.8.10.4518-4523.1988).
- Sun, Jing et al. (2020). "Generation of a Broadly Useful Model for COVID-19 Pathogenesis, Vaccination, and Treatment". In: *Cell* 182.3, 734–743.e5. ISSN: 00928674. DOI: [10.1016/j.cell.2020.06.010](https://doi.org/10.1016/j.cell.2020.06.010).
- Svenson, Karen L et al. (2012). "High-Resolution Genetic Mapping Using the Mouse Diversity Outbred Population". In: *Genetics* 190.2, pp. 437–447. ISSN: 1943-2631. DOI: [10.1534/genetics.111.132597](https://doi.org/10.1534/genetics.111.132597).
- Swaney, Emily, Callan O'Connor, and Laura G. Reinholdt (2021). "Mouse Genetic Reference Populations: Cellular Platforms for Integrative Systems Genetics". In: *Trends in Genetics* 37.3. Publisher: Elsevier, pp. 251–265. ISSN: 0168-9525. DOI: [10.1016/j.tig.2020.09.007](https://doi.org/10.1016/j.tig.2020.09.007).
- Takeuchi, Osamu and Shizuo Akira (2009). "Innate Immunity to Virus Infection". In: *Immunological reviews* 227.1, pp. 75–86. ISSN: 0105-2896. DOI: [10.1111/j.1600-065X.2008.00737.x](https://doi.org/10.1111/j.1600-065X.2008.00737.x).
- Talero-Gutiérrez, C. et al. (2018). "Zika virus epidemiology: from Uganda to world pandemic, an update". In: *Epidemiology and Infection* 146.6, pp. 673–679. ISSN: 1469-4409. DOI: [10.1017/S0950268818000419](https://doi.org/10.1017/S0950268818000419).
- Tan, Yee Sun and Yu L. Lei (2019). "Generation and Culture of Mouse Embryonic Fibroblasts". In: *Mouse Models of Innate Immunity*. Vol. 1960. Series Title: Methods in Molecular Biology. New York, NY: Springer New York, pp. 85–91. ISBN: 978-1-4939-9166-2 978-1-4939-9167-9. DOI: [10.1007/978-1-4939-9167-9\\_7](https://doi.org/10.1007/978-1-4939-9167-9_7).
- Thackray, Larissa B. et al. (2012). "Critical Role for Interferon Regulatory Factor 3 (IRF-3) and IRF-7 in Type I Interferon-Mediated Control of Murine Norovirus Replication". In: *Journal of Virology* 86.24, pp. 13515–13523. ISSN: 0022-538X, 1098-5514. DOI: [10.1128/JVI.01824-12](https://doi.org/10.1128/JVI.01824-12).
- The Complex Trait Consortium (2004). "The Collaborative Cross, a community resource for the genetic analysis of complex traits". In: *Nature Genetics* 36.11, pp. 1133–1137. ISSN: 1061-4036, 1546-1718. DOI: [10.1038/ng1104-1133](https://doi.org/10.1038/ng1104-1133).
- The GenOMICC Investigators et al. (2021). "Genetic mechanisms of critical illness in COVID-19". In: *Nature* 591.7848, pp. 92–98. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/s41586-020-03065-y](https://doi.org/10.1038/s41586-020-03065-y).
- Thery, Fabien et al. (2021). "Ring finger protein 213 assembles into a sensor for ISGylated proteins with antimicrobial activity". In: *Nature Communications* 12.1. Number: 1 Publisher: Nature Publishing Group, p. 5772. ISSN: 2041-1723. DOI: [10.1038/s41467-021-26061-w](https://doi.org/10.1038/s41467-021-26061-w).
- Threadgill, David W., Kent W. Hunter, and Robert W. Williams (2002). "Genetic dissection of complex and quantitative traits: from fantasy to reality via a community effort". In: *Mammalian Genome* 13.4, pp. 175–178. ISSN: 0938-8990, 1432-1777. DOI: [10.1007/s00335-001-4001-y](https://doi.org/10.1007/s00335-001-4001-y).
- Tian, Chao et al. (2017). "Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections". In: *Nature Communications* 8.1, p. 599. ISSN: 2041-1723. DOI: [10.1038/s41467-017-00257-5](https://doi.org/10.1038/s41467-017-00257-5).
- Turner, Thomas L. (2014). "Fine-mapping natural alleles: quantitative complementation to the rescue". In: *Molecular ecology* 23.10, pp. 2377–2382. ISSN: 0962-1083. DOI: [10.1111/mec.12719](https://doi.org/10.1111/mec.12719).
- Uffelmann, Emil et al. (2021). "Genome-wide association studies". In: *Nature Reviews Methods Primers* 1.1, p. 59. ISSN: 2662-8449. DOI: [10.1038/s43586-021-00056-9](https://doi.org/10.1038/s43586-021-00056-9).
- UNC, Systems Genetics Core Facility (2023). *CompGen Tool Suite*. URL: <https://csbio.unc.edu/CCstatus/index.py> (visited on 07/26/2023).
- Uyar, Asli et al. (2021). "Incorporating genetic diversity to improve alignment of mouse models to human Alzheimer's disease". In: *Alzheimer's & Dementia* 17 (S3). ISSN: 1552-5260, 1552-5279. DOI: [10.1002/alz.056498](https://doi.org/10.1002/alz.056498).

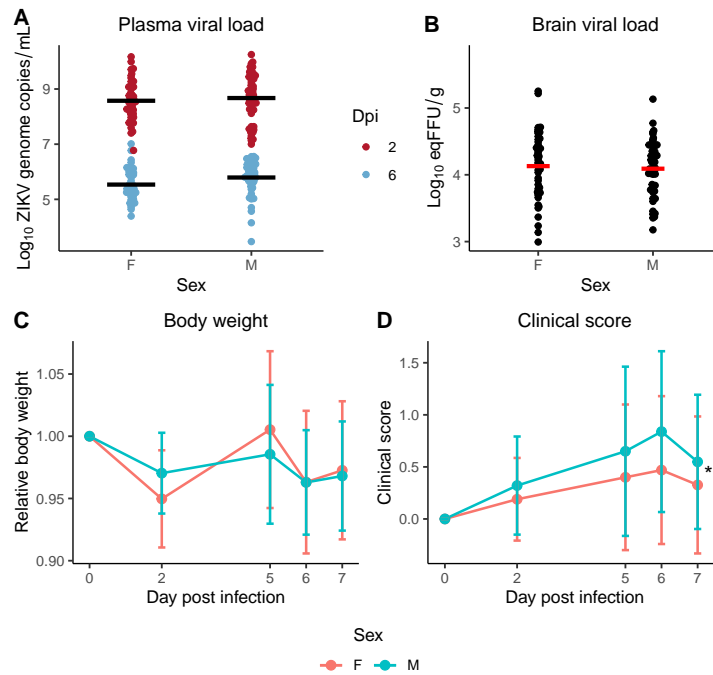
- Vasilakis, Nikos et al. (2011). "Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health". In: *Nature Reviews Microbiology* 9.7. Number: 7 Publisher: Nature Publishing Group, pp. 532–541. ISSN: 1740-1534. DOI: [10.1038/nrmicro2595](https://doi.org/10.1038/nrmicro2595).
- Wakabayashi, Yuichi et al. (2003). "Bcl11b is required for differentiation and survival of  $\alpha\beta$  T lymphocytes". In: *Nature Immunology* 4.6, pp. 533–539. ISSN: 1529-2908, 1529-2916. DOI: [10.1038/ni927](https://doi.org/10.1038/ni927).
- Wang, Ting et al. (2020). "In vivo imaging of Zika virus reveals dynamics of viral invasion in immune-sheltered tissues and vertical propagation during pregnancy". In: *Theranostics* 10.14, pp. 6430–6447. ISSN: 1838-7640. DOI: [10.7150/thno.43177](https://doi.org/10.7150/thno.43177).
- Wang, Yong-Fei et al. (2021). "Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups". In: *Nature Communications* 12.1, p. 772. ISSN: 2041-1723. DOI: [10.1038/s41467-021-21049-y](https://doi.org/10.1038/s41467-021-21049-y).
- Weaver, Scott C. (2017). "Emergence of Epidemic Zika Virus Transmission and Congenital Zika Syndrome: Are Recently Evolved Traits to Blame?" In: *mBio* 8.1, e02063–16. ISSN: 2161-2129, 2150-7511. DOI: [10.1128/mBio.02063-16](https://doi.org/10.1128/mBio.02063-16).
- Weaver, Scott C. et al. (2016). "Zika virus: History, emergence, biology, and prospects for control". In: *Antiviral Research* 130, pp. 69–80. ISSN: 01663542. DOI: [10.1016/j.antiviral.2016.03.010](https://doi.org/10.1016/j.antiviral.2016.03.010).
- Weidner, Jessica M. et al. (2010). "Interferon-Induced Cell Membrane Proteins, IFITM3 and Tetherin, Inhibit Vesicular Stomatitis Virus Infection via Distinct Mechanisms". In: *Journal of Virology* 84.24. Publisher: American Society for Microbiology, pp. 12646–12657. DOI: [10.1128/jvi.01328-10](https://doi.org/10.1128/jvi.01328-10).
- WHO (2020). *The top 10 causes of death*. URL: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (visited on 07/12/2023).
- (2022). *Zika virus*. URL: <https://www.who.int/news-room/fact-sheets/detail/zika-virus> (visited on 07/13/2023).
- (2023). *Dengue and severe dengue*. URL: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue> (visited on 07/12/2023).
- Wilder, Julius et al. (2016). "A Systematic Review of Race and Ethnicity in Hepatitis C Clinical Trial Enrollment". In: *Journal of the National Medical Association* 108.1, pp. 24–29. ISSN: 0027-9684. DOI: [10.1016/j.jnma.2015.12.004](https://doi.org/10.1016/j.jnma.2015.12.004).
- Wilken, Lucas et al. (2023). "Transient Blockade of Type I Interferon Signalling Promotes Replication of Dengue Virus Strain D2Y98P in Adult Wild-Type Mice". In: *Viruses* 15.4, p. 814. ISSN: 1999-4915. DOI: [10.3390/v15040814](https://doi.org/10.3390/v15040814).
- Winkler, Clayton W. et al. (2017). "Sexual and Vertical Transmission of Zika Virus in anti-interferon receptor-treated Rag1-deficient mice". In: *Scientific Reports* 7.1, p. 7176. ISSN: 2045-2322. DOI: [10.1038/s41598-017-07099-7](https://doi.org/10.1038/s41598-017-07099-7).
- Yakob, Laith (2022). "Zika Virus after the Public Health Emergency of International Concern Period, Brazil - Volume 28, Number 4—April 2022 - Emerging Infectious Diseases journal - CDC". In: DOI: [10.3201/eid2804.211949](https://doi.org/10.3201/eid2804.211949).
- Yanai, Hideyuki et al. (2018). "Revisiting the role of IRF3 in inflammation and immunity by conditional and specifically targeted gene ablation in mice". In: *Proceedings of the National Academy of Sciences* 115.20, pp. 5253–5258. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1803936115](https://doi.org/10.1073/pnas.1803936115).
- Yang, Hui et al. (2023a). "Thirdhand tobacco smoke exposure increases the genetic background-dependent risk of pancreatic tumor development in Collaborative Cross mice". In: *Environment International* 174, p. 107876. ISSN: 01604120. DOI: [10.1016/j.envint.2023.107876](https://doi.org/10.1016/j.envint.2023.107876).
- Yang, Liuliu et al. (2023b). "Isogenic human trophectoderm cells demonstrate the role of NDUFA4 and associated variants in ZIKV infection". In: *iScience* 26.7, p. 107001. ISSN: 25890042. DOI: [10.1016/j.isci.2023.107001](https://doi.org/10.1016/j.isci.2023.107001).
- Yoneyama, Mitsutoshi et al. (2004). "The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses". In: *Nature Immunology* 5.7. Number: 7 Publisher: Nature Publishing Group, pp. 730–737. ISSN: 1529-2916. DOI: [10.1038/ni1087](https://doi.org/10.1038/ni1087).
- Zanin, Natacha et al. (2021). "Interferon Receptor Trafficking and Signaling: Journey to the Cross Roads". In: *Frontiers in Immunology* 11. ISSN: 1664-3224.
- Zhang, Dongxian and Dong-Er Zhang (2011). "Interferon-Stimulated Gene 15 and the Protein ISGylation System". In: *Journal of Interferon & Cytokine Research* 31.1. Publisher: Mary Ann Liebert, Inc., publishers, pp. 119–130. ISSN: 1079-9907. DOI: [10.1089/jir.2010.0110](https://doi.org/10.1089/jir.2010.0110).
- Zhang, Jing et al. (2019). "A Loss-of-Function Mutation in the Integrin Alpha L (Itgal) Gene Contributes to Susceptibility to Salmonella enterica Serovar Typhimurium Infection in Collaborative Cross Strain CC042". In: *Infection and Immunity* 88.1. ISSN: 1098-5522. DOI: [10.1128/IAI.00656-19](https://doi.org/10.1128/IAI.00656-19).
- Zhang, Liang et al. (2018). "Association of CD1 and Fc $\gamma$ R gene polymorphisms with Guillain-Barré syndrome susceptibility: a meta-analysis". In: *Neurological Sciences* 39.12, pp. 2141–2149. ISSN: 1590-3478. DOI: [10.1007/s10072-018-3563-3](https://doi.org/10.1007/s10072-018-3563-3).
- Zhang, Q. et al. (2020). "Inborn errors of type I IFN immunity in patients with life-threatening COVID-19". In: *Science*. ISSN: 1095-9203 (Electronic) 0036-8075 (Linking). DOI: [10.1126/science.abd4570](https://doi.org/10.1126/science.abd4570).
- Zhang, Shaojun et al. (2022). "Comparison of viral RNA–host protein interactomes across pathogenic RNA viruses informs rapid antiviral drug discovery for SARS-CoV-2". In: *Cell Research* 32.1, pp. 9–23. ISSN: 1001-0602, 1748-7838. DOI: [10.1038/s41422-021-00581-y](https://doi.org/10.1038/s41422-021-00581-y).
- Zhang, Tian et al. (2023). "Multi-omics analysis identifies drivers of protein phosphorylation". In: *Genome Biology* 24.1, p. 52. ISSN: 1474-760X. DOI: [10.1186/s13059-023-02892-2](https://doi.org/10.1186/s13059-023-02892-2).
- Zhang, Xiaoyong et al. (2016). "Epigenetically regulated miR-449a enhances hepatitis B virus replication by targeting cAMP-responsive element binding protein 5 and modulating hepatocytes phenotype". In: *Scientific Reports* 6.1. Number: 1 Publisher: Nature Publishing Group, p. 25389. ISSN: 2045-2322. DOI: [10.1038/srep25389](https://doi.org/10.1038/srep25389).
- Zhou, Hao et al. (2014). "Interferons and Their Receptors in Birds: A Comparison of Gene Structure, Phylogenetic Analysis, and Cross Modulation". In: *International Journal of Molecular Sciences* 15.11. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute, pp. 21045–21068. ISSN: 1422-0067. DOI: [10.3390/ijms151121045](https://doi.org/10.3390/ijms151121045).



- Zhou, Yadi et al. (2023). "A comprehensive SARS-CoV-2–human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets". In: *Nature Biotechnology* 41.1, pp. 128–139. ISSN: 1087-0156, 1546-1696. DOI: [10.1038/s41587-022-01474-0](https://doi.org/10.1038/s41587-022-01474-0).
- Zhu, Xiaobo et al. (2023). "Using multi-tissue transcriptome-wide association study to identify candidate susceptibility genes for respiratory infectious diseases". In: *Frontiers in Genetics* 14, p. 1164274. ISSN: 1664-8021. DOI: [10.3389/fgene.2023.1164274](https://doi.org/10.3389/fgene.2023.1164274).
- Zsichla, Levente and Viktor Müller (2023). "Risk Factors of Severe COVID-19: A Review of Host, Viral and Environmental Factors". In: *Viruses* 15.1, p. 175. ISSN: 1999-4915. DOI: [10.3390/v15010175](https://doi.org/10.3390/v15010175).

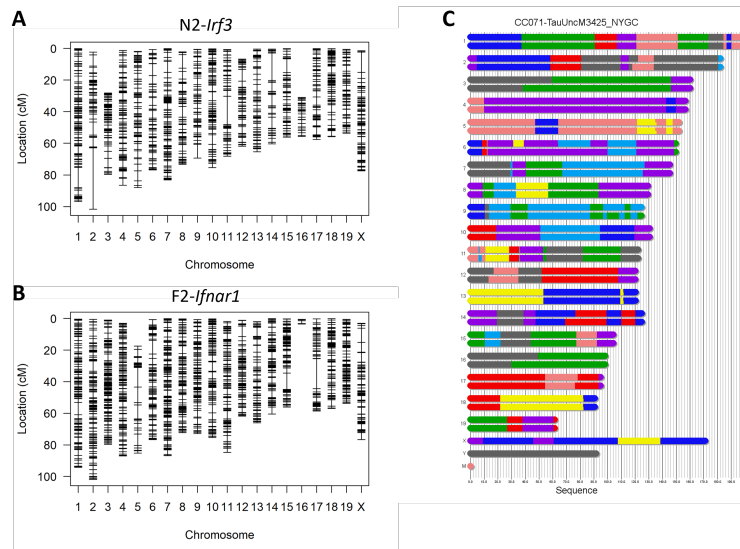
## VIII - Appendix

## VIII.1 . Supplementary Figures



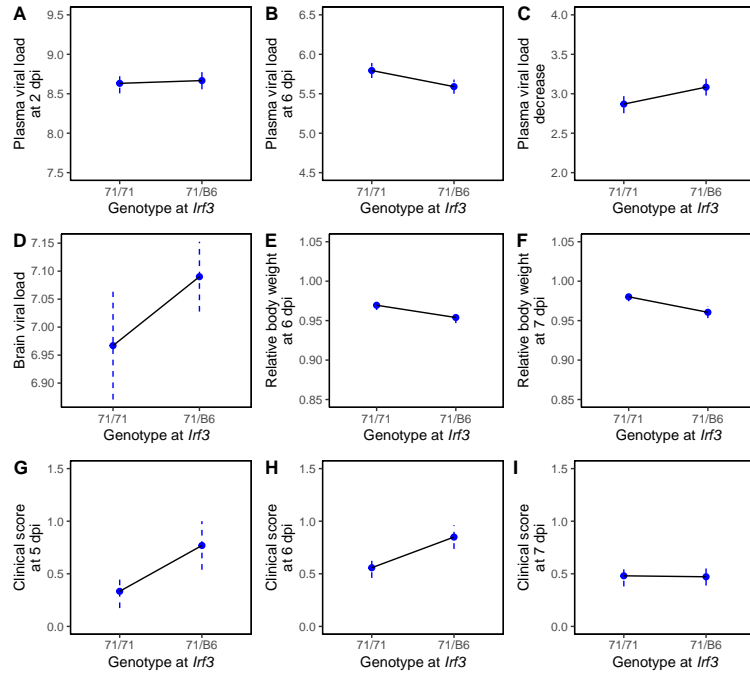
**Figure S1:** Effect of sex on the phenotypes of the N2-*Irf3* progeny.

Effect of sex on plasma viral load (A), brain viral load (B), body weight loss (C) and clinical score (D). Phenotypic means are depicted as black bars in C and as red bars in B. Differences between males and females were assessed with mixed two-way ANOVA (A), t test (B), or linear mixed models (C - D).

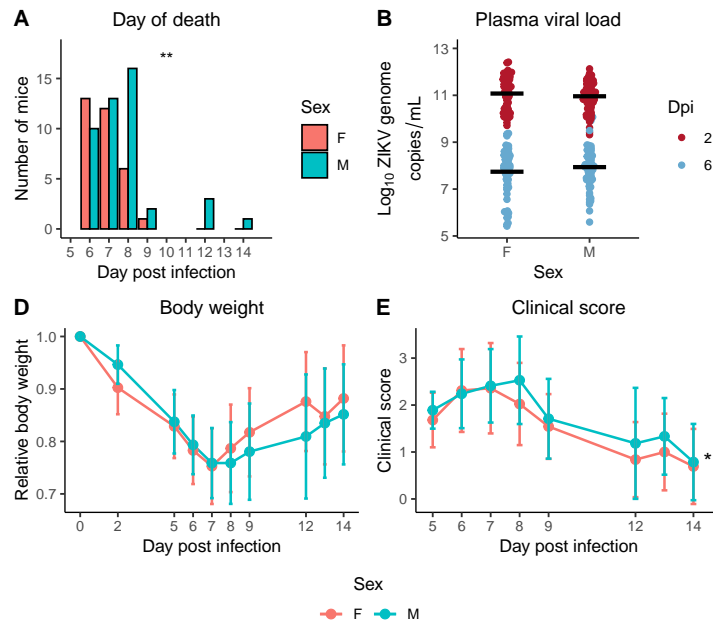


**Figure S2:** Marker density maps for the N2-*Irf3* and F2-*Ifnar1* progenies

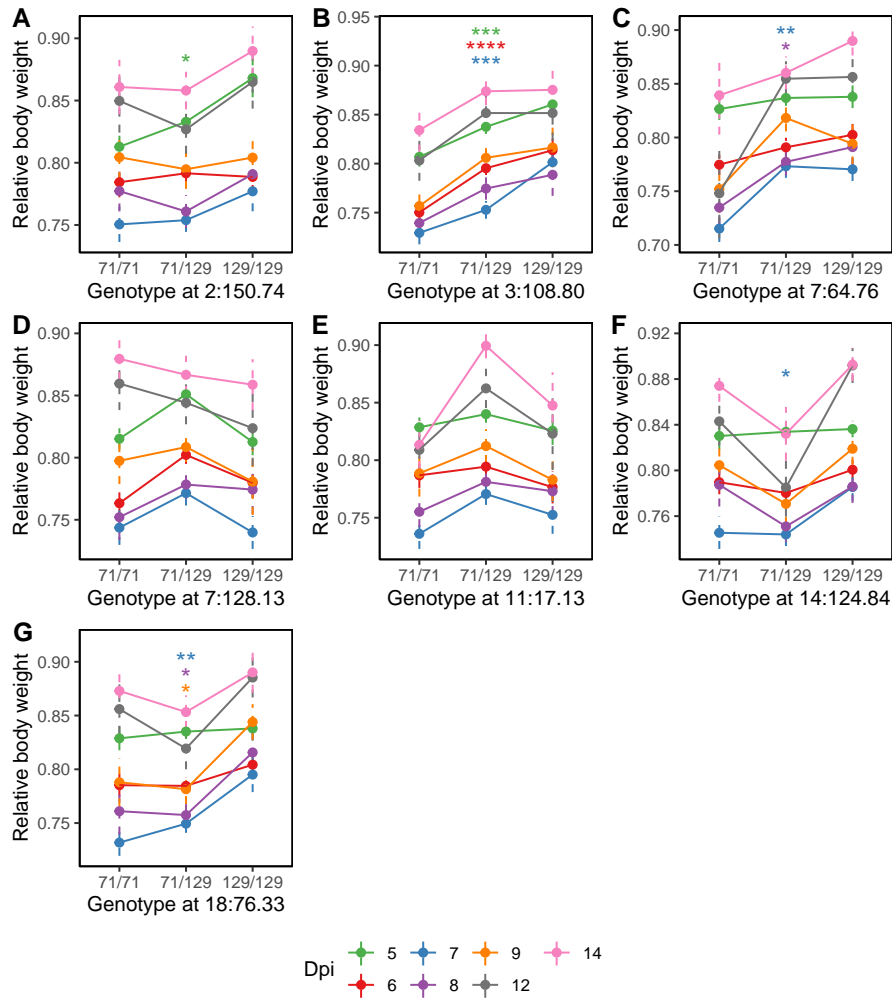
Genetic location along the chromosomes (in cM) of informative markers used for QTL mapping in N2-*Irf3* (A) and F2-*Ifnar1* (B) mice. (C) Schematic representation of CC071 haplotypes, from <https://csbio.unc.edu/CCstatus/CCGenomes/>. Colors represent the CC ancestral haplotypes (A/J: yellow, B6: grey, 129S1/SvImj: pink, NOD/ShiLtj: dark blue, NZO/HILtj: light blue, CAST/Eij: green, PWK/Phj: red, WSB/Eij: purple).



**Figure S3:** Allelic effects for *Irf3* on the different phenotypes tested in the N2-*Irf3* progeny. Allelic effects were assessed at the closest marker to *Irf3*, SAH071783312, at 44.23 Mb, (A - D) Groups were compared by t test and the results were not significant. (E - I) Groups were compared by Wilcoxon test and the results were not significant.

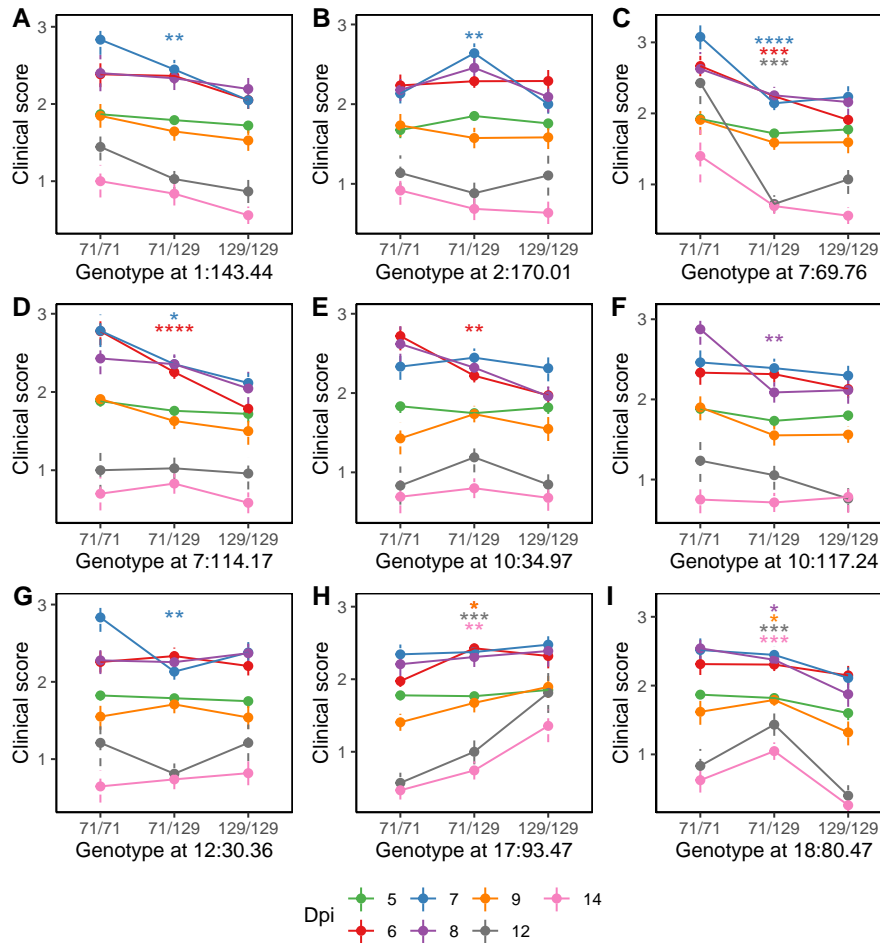


**Figure S4:** Effect of sex on the phenotypes of the F2 progeny. Effect of sex on day of death (A), survival (B), PVL (C), body weight loss (D) and clinical score (E). Phenotypic means are depicted as black bars in C and as red bars in D and E. Differences between males and females were assessed with t test (A), mixed two-way ANOVA (B), or linear mixed models (C - D).



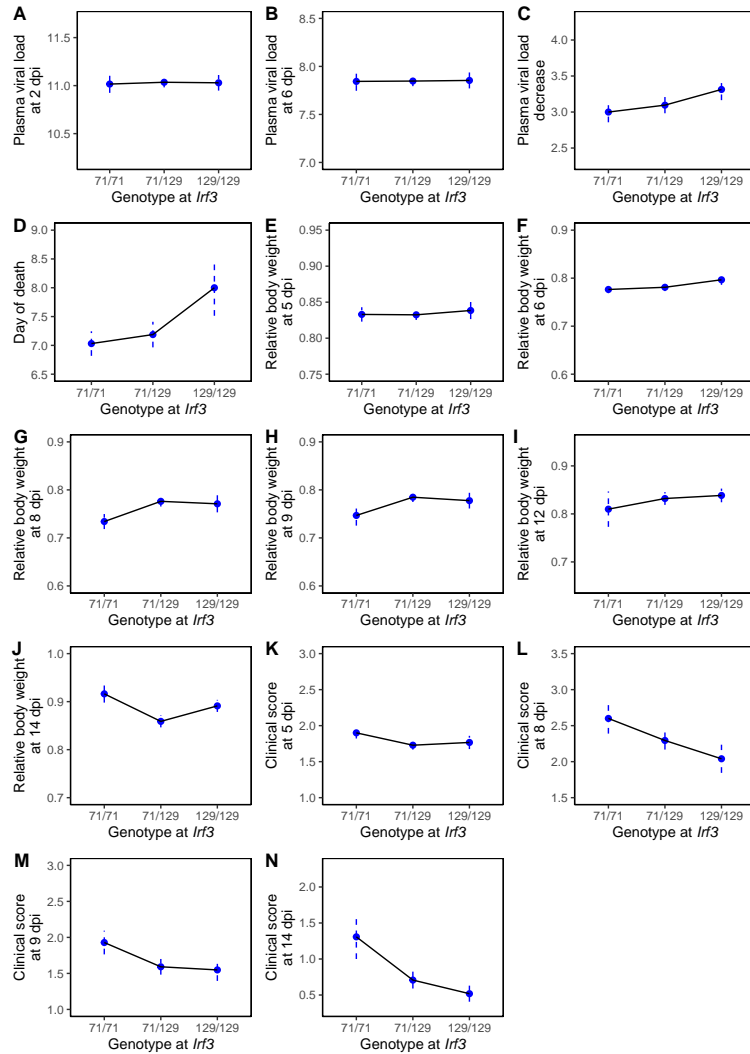
**Figure S5:** Allelic effects for the QTLs controlling the body weights of F2-*Ifnar1* mice.

Allelic effects assessed at the marker with the highest LOD scores: gJAX00101614, at 150.74 Mb on chromosome 2 (A), gUNC5933912, at 108.8 Mb on chromosome 3 (B), gUNC13007707, at 64.76 Mb on chromosome 7 (C), UNC13823755, at 128.13 Mb on chromosome 7 (D), SBR110687040, at 17.13 Mb on chromosome 11 (E), mUNC24942178, at 124.874 Mb on chromosome 14 (F), and gbackupUNC180311503, at 76.33 Mb on chromosome 18. Results are represented as mean  $\pm$  sem. Groups were compared by ANOVA (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).

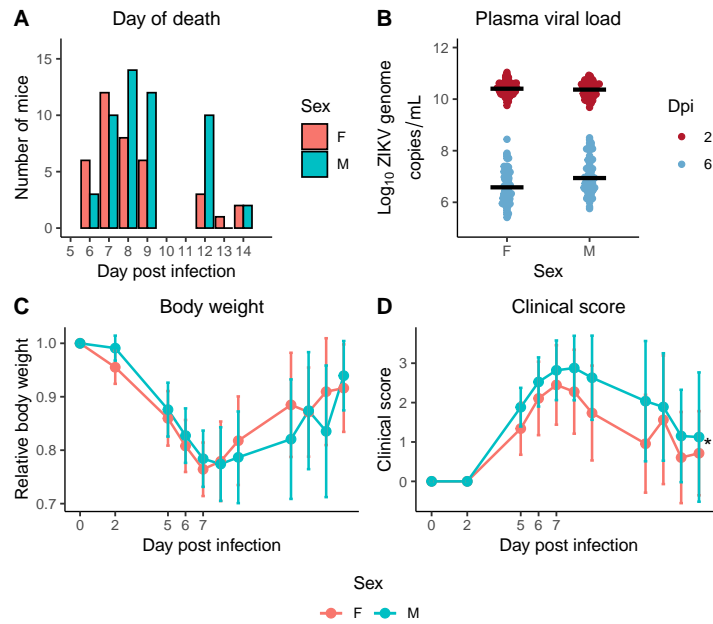


**Figure S6:** Allelic effects for the QTLs controlling the clinical scores of F2-*Ifnar1* mice.

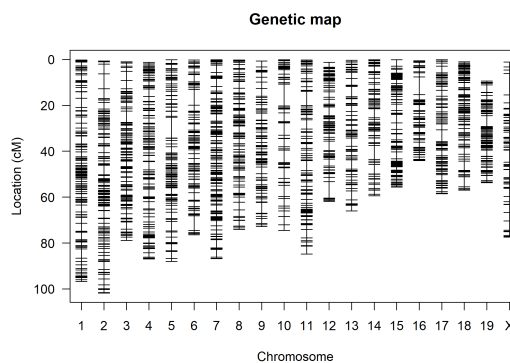
Allelic effects assessed at the marker with the highest LOD scores: DSB015742401, at 143.44 Mb on chromosome 1 (A), gUNC4486158, at 170.01 Mb on chromosome 2 (B), gUNC13076624, at 69.76 Mb on chromosome 7 (C), gUNC13659520, at 114.17 Mb on chromosome 7 (D), mUNC17738774, at 34.97 Mb on chromosome 10 (E), S2T104696211, at 117.24 Mb on chromosome 10 (F), SS1121212211, at 30.36 Mb on chromosome 12 (G), UNC28626374, at 93.47 Mb on chromosome 17 (H), and gUNC29691584, at 80.41 Mb on chromosome 18 (I). Results are represented as mean  $\pm$  sem. Groups were compared by Kruskal Wallis (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).



**Figure S7:** Allelic effects for *Irf3* on the different phenotypes tested in the F2-*Ifnar1* progeny. Allelic effects were assessed at the closest marker to *Irf3*: SX1071791093, at 44.42 Mb on chromosome 7. (A - J) Groups were compared by ANOVA and the results were not significant. (K - N) Groups were compared by Kruskal Wallis and the results were not significant.

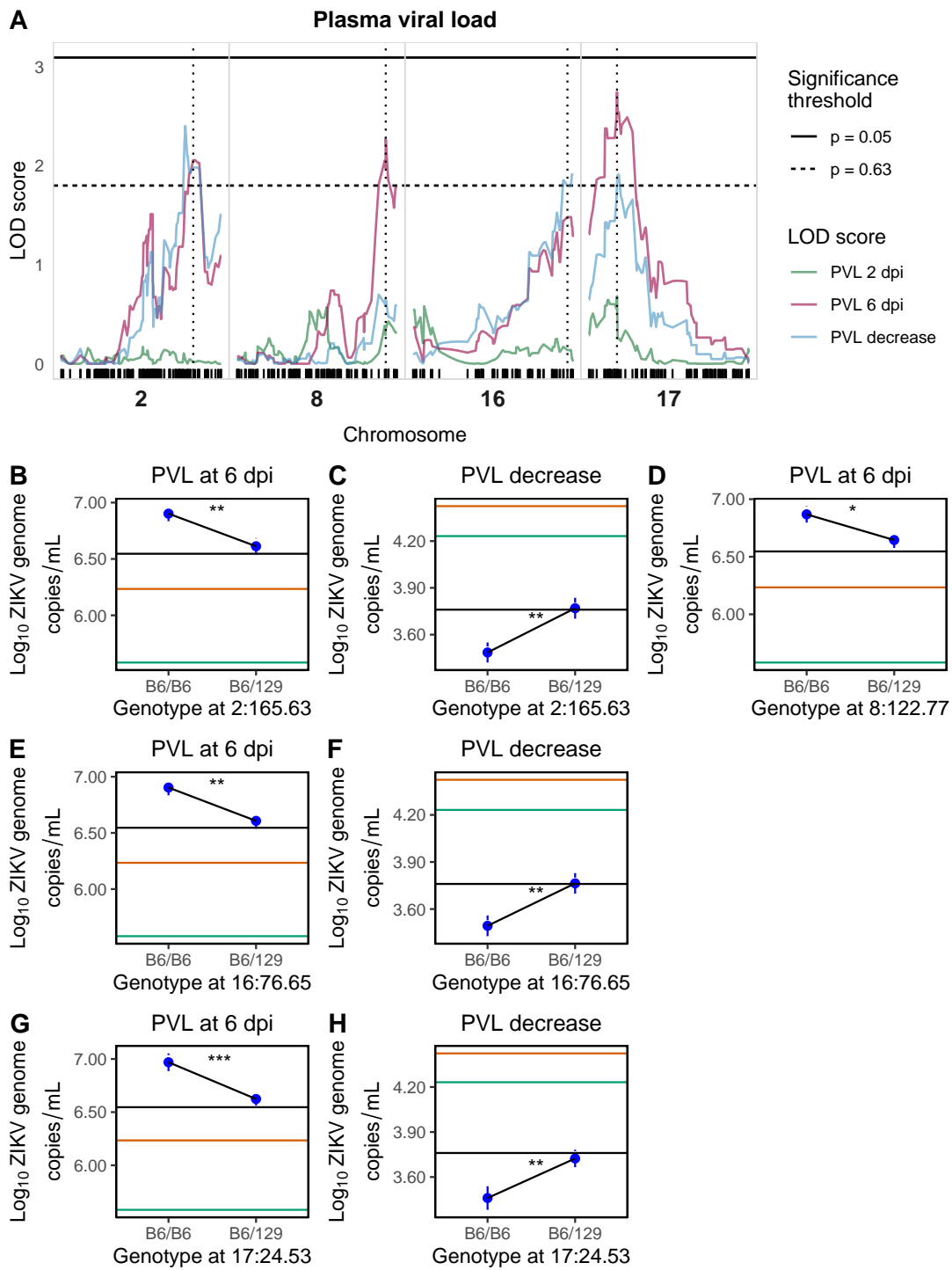


**Figure S8:** Effect of sex on the phenotypes of the N2 progeny. Effect of sex on day of death (A), survival (B), plasma viral load (C), body weight loss (D) and clinical score (E). Phenotypic means are depicted as black bars in C and as red bars in D and E. Differences between males and females were assessed with t test (A), mixed two-way ANOVA (B), or linear mixed models (C - D).



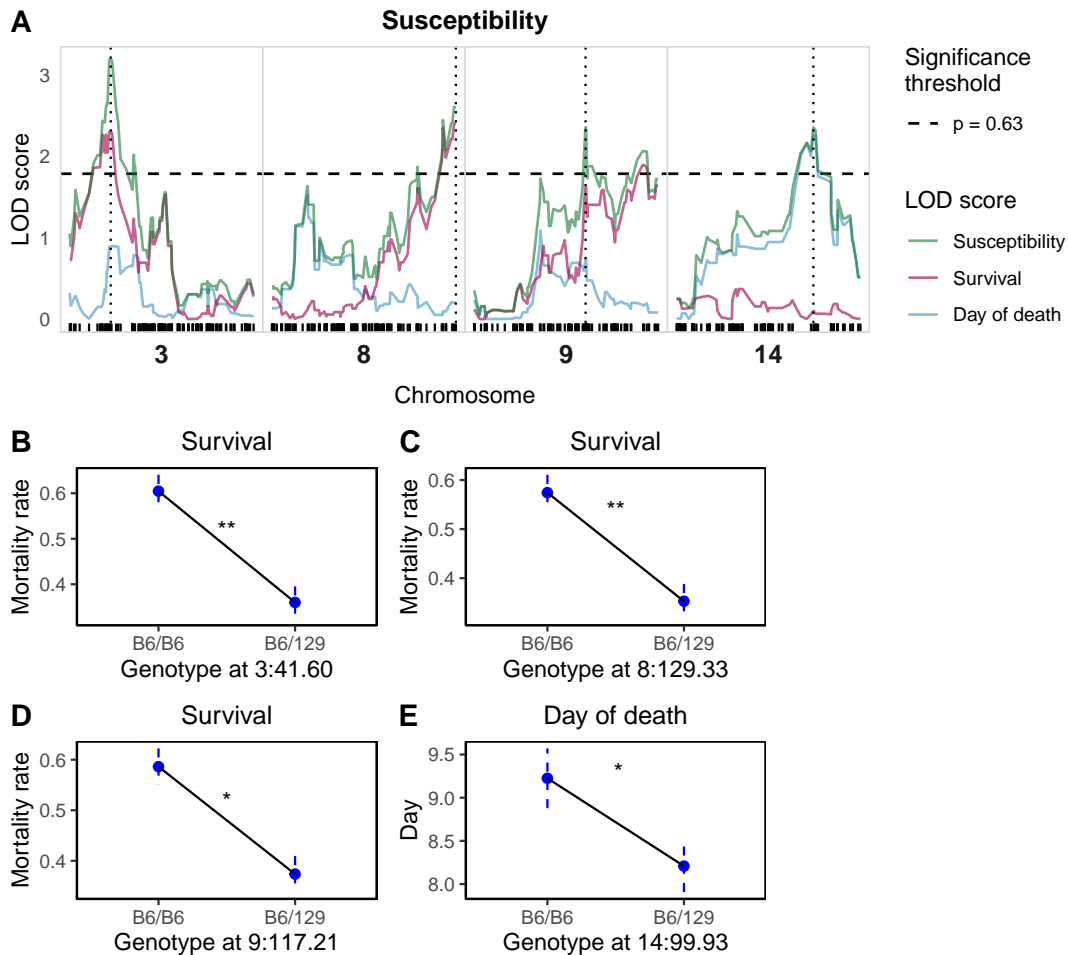
**Figure S9:** Marker density map for the B6-*Irfar1* × 129-*Irfar1* N2 progeny. Genetic location along the chromosomes (in cM) of informative markers used for QTL mapping.



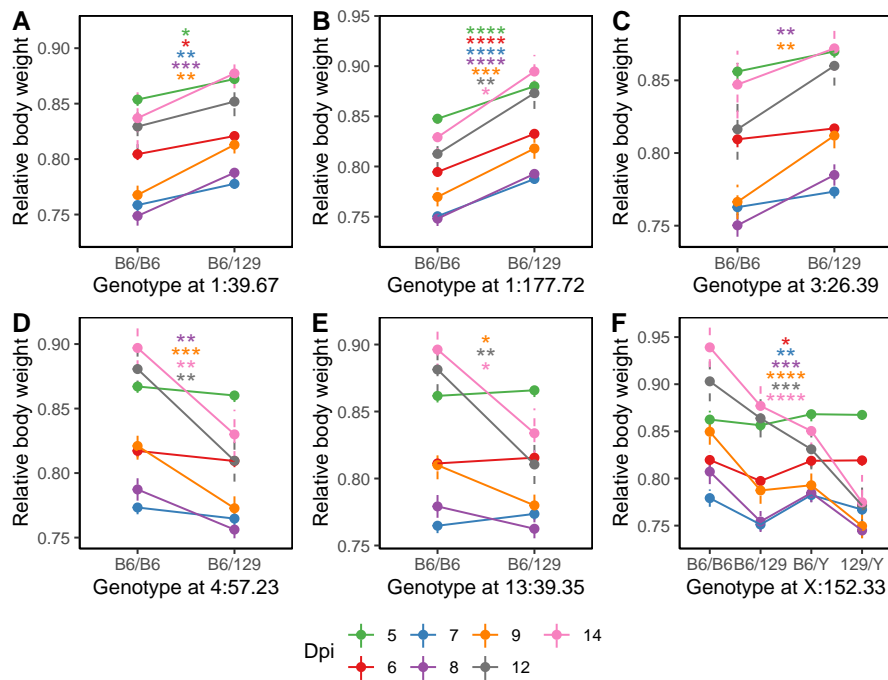


**Figure S10:** Suggestive QTLs associated with the PVL of B6-*Ifnar1* × 129-*Ifnar1* N2 mice.

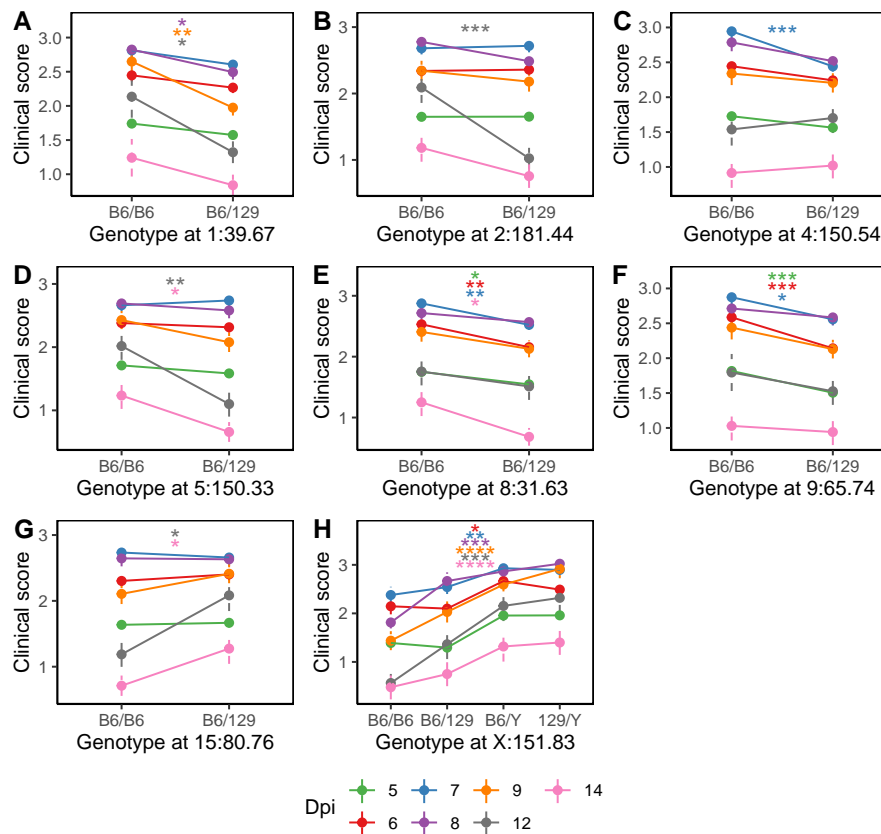
(A) Zoom on the chromosomes 2, 8, 16 and 17 QTL associated with the PVL. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.05$ ,  $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 42. The vertical dotted lines represent the positions of the markers used to display allelic effects in B - H. Allelic effects assessed at the marker with the highest LOD scores: gUNC4422942 at 165.63 Mb on chromosome 2 (B - C), gUNC27161197, at 76.65 Mb on chromosome 16 (D - E), gUNC27686840, at 24.53 Mb on chromosome 17 (F - G). Results are represented as mean  $\pm$  sem. Groups were compared by t test (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ). Horizontal lines show the average PVL of B6-*Ifnar1* (black), 129-*Ifnar1* (green), and F1-*Ifnar1* (brown).



**Figure S11:** Suggestive QTLs associated with the survival and day of death of B6-*Ifnar1* × 129-*Ifnar1* N2 mice. (A) Zoom on the chromosomes 3, 8, 9 and 14 QTL associated with the survival and the day of death. Horizontal lines indicate genome-wide significance thresholds ( $p = 0.63$ ) determined by permutation testing ( $n=1000$ ), as in Figure 46. The vertical dotted lines represent the positions of the markers used to display allelic effects in B - E. Allelic effects assessed at the marker with the highest LOD scores: gUNCHS008529 at 41.60 Mb on chromosome 3 (B), UNC15748820, at 129.33 Mb on chromosome 8 (C), gUNCHS026514, at 85.61 Mb on chromosome 9 (D), gUNC24598868, at 99.93 Mb on chromosome 14 (E). Results are represented as mean  $\pm$  sd (B - D) or mean  $\pm$  sem (E). Groups were compared by  $\chi^2$  (B) or t-test (C) (\*  $p < 0.05$ , \*\*  $p < 0.01$ ).



**Figure S12:** Allelic effects for the QTLs controlling the body weights of B6-*Irfnar1* × 129-*Irfnar1* N2 mice. Allelic effects assessed at the marker with the highest LOD scores: gUNC496797, at 39.67 Mb on chromosome 1 (B), gUNCHS003416, at 177.72 Mb on chromosome 1 (C), gUNC4904534, at 26.38 Mb on chromosome 3 (D), gJAX00551855, at 57.23 Mb on chromosome 4 (E), gUNC22457961, at 39.35 Mb on chromosome 13 (F) and SXX206141922, at 152.33 Mb on chromosome X (G). Results are represented as mean ± sem. Groups were compared by ANOVA (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).



**Figure S13:** Allelic effects for the QTLs controlling the clinical scores of B6-*Ifnar1* × 129-*Ifnar1* N2 mice. Allelic effects assessed at the marker with the highest LOD scores: gUNC496797, at 39.67 Mb on chromosome 1 (B), mUNC4608754, at 181.44 Mb on chromosome 2 (C), D6H046018203, at 150.54 Mb on chromosome 4 (D), gUNC10448854, at 150.33 Mb on chromosome 5 (E), UNCHS022552, at 31.63 Mb on chromosome 8 (F), mUNC16527703, at 65.74 Mb on chromosome 10 (G), gUNC25967395, at 80.76 Mb on chromosome 15 (H), and gUNC31470943, at 151.93 Mb on chromosome X (I). Results are represented as mean ± sem. Groups were compared by Wilcoxon test (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).

## VIII.2 . Supplementary Tables

**Table S1:** Genes in *Zs1*

Symbol	Position (Mb)	Variant	Expression	Immune function
<i>1700001K19Rik</i>	110.63			
<i>4930595D18Rik</i>	111.13			
<i>A530016L24Rik</i>	112.46	Deleterious		
<i>Abcb5</i>	118.83	Deleterious		
<i>Adam6a</i>	113.51			
<i>Adam6b</i>	113.45			
<i>Adss1</i>	112.59		nd	
<i>Ahnak2</i>	112.74		↓	
<i>Ak7</i>	105.67	Deleterious		Yes
<i>Akt1</i>	112.62			
<i>Amn</i>	111.24	Tolerated		
<i>Ankrd9</i>	110.94	Tolerated		
<i>Aspg</i>	112.07	Deleterious		
<i>Atg2b</i>	105.58	Tolerated	↑	
<i>Atp5mj</i>	111.93		nd	
<i>Bag5</i>	111.68	Tolerated		
<i>Bcl11</i>	107.88		↓	Yes
<i>Bdkrb1</i>	105.57	Tolerated		Yes
<i>Bdkrb2</i>	105.53			
<i>Begain</i>	109	Tolerated		
<i>Brf1</i>	112.92	Deleterious		Yes
<i>Btbd6</i>	112.94			
<i>Ccdc85c</i>	108.17	Tolerated		
<i>Ccnk</i>	108.15			
<i>Cdc42bpb</i>	111.26	Tolerated		
<i>Cdca4</i>	112.78	Tolerated		
<i>Cdca7l</i>	117.81	Tolerated		
<i>Cep170b</i>	112.69	Deleterious		
<i>Cinp</i>	110.84	Deleterious		
<i>Ckb</i>	111.64			
<i>Clba1</i>	112.77	Deleterious	nd	
<i>Coa8</i>	111.68	Deleterious	nd	
<i>Crip1</i>	113.12			
<i>Crip2</i>	113.1			
<i>Cyp46a1</i>	108.3	Tolerated		
<i>Degs2</i>	108.65			
<i>Dio3</i>	110.25			
<i>DIK1</i>	109.42		↓	Yes
<i>Dnah11</i>	117.84	Deleterious		
<i>Dync1h1</i>	110.57	Tolerated		
<i>Dync2i1</i>	116.17	Deleterious	nd	
<i>Eif5</i>	111.5	Tolerated		
<i>Eml1</i>	108.34			
<i>Esyt2</i>	116.24	Tolerated		
<i>Evl</i>	108.52			Yes
<i>Exoc3l4</i>	111.38			
<i>Gm10427</i>	108.21			
<i>Gm11027</i>	116.24		nd	
<i>Gm266</i>	111.45	Tolerated		
<i>Gm30599</i>	113.11		nd	
<i>Gm34220</i>	108.76		nd	
<i>Gm46320</i>	106.6		nd	
<i>Gm46382</i>	112.68		nd	
<i>Gm52009</i>	112.79		nd	
<i>Gm53767</i>	108.8		nd	
<i>Gpr132</i>	112.81			Yes
<i>Gskip</i>	105.65			
<i>Hhip1l</i>	108.27	Tolerated	↑	
<i>Hsp90aa1</i>	110.66			Yes
<i>Inf2</i>	112.56	Deleterious		
<i>Itgb8</i>	119.12	Tolerated		
<i>Jag2</i>	112.87	Tolerated		
<i>Kif26a</i>	112.11	Deleterious		
<i>Klc1</i>	111.73	Tolerated		

Symbol	Position (Mb)	Variant	Expression	Immune function
<i>Lbhd2</i>	111.37		nd	
<i>Macc1</i>	119.35	Deleterious		
<i>Mark3</i>	111.54			
<i>Mok</i>	110.77	Deleterious		
<i>Mta1</i>	113.06			
<i>Ncapg2</i>	116.37			
<i>Nudt14</i>	112.9			
<i>Pacs2</i>	112.98	Deleterious		
<i>Papola</i>	105.75	Tolerated		
<i>Pld4</i>	112.73	Tolerated		Yes
<i>Ppp1r13b</i>	111.79	Tolerated		Yes
<i>Ppp2r5c</i>	110.41	Deleterious		
<i>Ptprn2</i>	116.45	Tolerated		
<i>Rapgef5</i>	117.48	Tolerated		
<i>Rcor1</i>	111.01			
<i>Rd3l</i>	111.95			
<i>Rtl1</i>	109.56			
<i>Setd3</i>	108.07			Yes
<i>Siva1</i>	112.61	Tolerated		
<i>Slc25a29</i>	108.79	Tolerated		
<i>Slc25a47</i>	108.82	Tolerated		
<i>Sp4</i>	118.2	Deleterious		
<i>Sp8</i>	118.81			
<i>Tdrd9</i>	111.94	Deleterious		
<i>Tecpr2</i>	110.86	Tolerated		
<i>Tedc1</i>	113.12		nd	
<i>Tex22</i>	113.04			
<i>Tmem121</i>	113.15			
<i>Tmem179</i>	112.47	Tolerated		
<i>Tmem196</i>	119.91			
<i>Tnfaip2</i>	111.41	Deleterious		
<i>Traf3</i>	111.13			Yes
<i>Trmt61a</i>	111.64	Tolerated		
<i>Vipr2</i>	116.04	Tolerated		Yes
<i>Vrk1</i>	105.98	Tolerated		
<i>Wars1</i>	108.83		nd	
<i>Wdr20</i>	110.7	Deleterious		
<i>Wdr25</i>	108.86	Tolerated		
<i>Xrcc3</i>	111.77	Deleterious		
<i>Yy1</i>	108.76			
<i>Zbtb42</i>	112.65	Tolerated		
<i>Zfp386</i>	116.01	Tolerated		
<i>Zfp839</i>	110.82	Deleterious		
<i>Zfyve21</i>	111.78			

The list of genes was obtained from MGI website (<https://www.informatics.jax.org/>). Polymorphisms were obtained from GenomeMUSter ([muster.jax.org/](https://muster.jax.org/)) and their effects from the Ensembl Variant Effect Predictor (<https://www.ensembl.org/info/docs/tools/vep/>). The Expression column indicates whether a gene was differentially expressed between CC071 and B6 in our MEF RNAseq data (↑: higher expression in CC071, ↓: lower expression in CC071, nd: non-determined, the gene was not expressed in MEFs). The immune function was assessed in MGI using the Phenotypes/Diseases criterion: MP:0001790. Green cells indicate genes with deleterious mutations, genes which are differentially expressed between CC071 and B6, and genes with known immune functions.

## VIII.3 . Résumé substantiel en français

### VIII.3.1 . Introduction

Les maladies transmissibles sont un facteur majeur de mort dans le monde. Elles représentent 6 des 10 causes principales de décès dans les pays à faible revenu, contre 3 à l'échelle mondiale. Parmi elles, les maladies virales sont provoquées par un groupe très divers de pathogènes. En effet, les virus peuvent provoquer des maladies aiguës ou chroniques, des conditions sévères ou modérées, et peuvent conduire à l'apparition de cancers. Au cours de l'évolution, des mécanismes sont apparus chez les organismes hôtes pour contrer les infections virales, et ont été conservés dans de nombreuses espèces.

Parmi ces mécanismes, l'immunité innée est la première ligne de défense contre les infections virales. Celle-ci est activée lorsque de molécules antivirales, appelées Pathogen Associated Molecular Patterns (PAMPs), sont reconnus par des récepteurs cellulaires appelés Pathogen Recognition Receptors (PRRs). La reconnaissance des PAMPs par les PRRs active des cascades de signalisation qui conduiront à la production de cytokines antivirales, dont les interférons (IFNs).

Les IFNs sont des cytokines connues pour « interférer » avec l'infection virale. En effet, la fixation des IFNs sur leurs récepteurs conduit à l'expression de gènes aux fonctions antivirales appelés Interferon Stimulated Genes (ISGs). Il existe trois familles d'interférons : type I, II et III. Chaque type d'IFNs a son propre récepteur et conduit à la production d'un ensemble spécifique d'ISGs. Mon travail s'est focalisé sur les IFNs de type I (IFN-I). Les IFN-I majeurs sont les IFN $\alpha$  (13 chez l'homme et 14 chez la souris) et l'IFN $\beta$ .

La production des IFN-I est induite à la suite de la reconnaissance des PAMPs viraux par les PRRs. Il existe deux types de PRRs pouvant reconnaître des molécules virales. Les Toll-Like Receptors (TLRs) qui peuvent reconnaître des protéines virales (TLR2, TLR4) ou des acides nucléiques viraux (TLR3, TLR7, TLR8, TLR9). Les RIG-I-Like Receptors (RLRs) reconnaissent des ARN viraux. Une fois la fixation de molécules virales sur les PRRs, ceux-ci activent des protéines adaptatrices qui activent la protéine TRAF3. TRAF3 active ensuite les kinases TBK1 et IKK $\epsilon$ . Ces kinases phosphorylent les facteurs de transcription responsables de l'expression des gènes des IFN-I appelés Interferon Regulatory Factors (IRFs). Les IRFs majeurs pour l'induction des IFN-I sont IRF3 et IRF7, qui, une fois phosphorylés, forment des homodimères ou des hétérodimères, transloquent dans le noyau et activent la transcription des gènes des IFN-I.

Il a été montré que des variants dans les gènes de la cascade d'induction des IFN-I sont associés à des sensibilités accrues aux maladies virales, grâce à des études chez l'homme et chez la souris. Deux types d'approches peuvent être utilisées pour démontrer qu'un gène est associé à un trait : l'approche gène candidat et l'approche génome entier. L'approche gène candidat démarre par l'identification d'un gène qui, étant donné sa fonction, peut être un bon candidat pour le phénotype d'intérêt. Ce gène est alors génotypé dans une population humaine, ou un knock-out (KO) peut être utilisé dans une étude chez la souris. L'association entre le génotype et le phénotype est alors testée. Dans l'approche génome entier, il n'y a pas d'hypothèse au préalable sur les gènes qui pourraient

être associés au phénotype d'intérêt. L'étape de génotypage se fait donc à l'échelle génomique grâce à des marqueurs répartis sur l'ensemble du génome. Des analyses d'associations génome entier (GWAS) sont alors réalisées et permettent d'identifier plusieurs associations. Les gènes de la cascade d'induction des IFN-I ont été associés à des sensibilités aux maladies virales principalement grâce à des approches gène candidat. En conséquence, afin d'identifier de nouveaux gènes de sensibilité, il est nécessaire d'utiliser des approches génome entier, également appelées approches de génétique directe.

Une première méthode de génétique directe utilisant des lignées consanguines de souris est le croisement à deux générations. Cette méthode démarre par l'identification de deux lignées de souris présentant des sensibilités différentes à une maladie virale. Ces deux lignées sont alors croisées pour donner des individus F1. Si la sensibilité des individus F1 est intermédiaire entre la sensibilité des lignées parentales, les individus F1 sont croisés pour donner une population d'individus intercross (F2). Autrement, si la sensibilité des individus F1 est similaire à celle d'une des deux lignées parentales, les F1 sont croisés avec l'autre lignée parentale pour donner une population d'individus backcross (N2). Dans les deux cas, les individus de seconde génération sont d'une part phénotypés pour leur sensibilité à la maladie virale, et de l'autre génotypés pour des marqueurs répartis sur l'ensemble du génome. Des analyses d'associations génotype/phénotype, également appelées analyses de Quantitative Trait Locus (QTL) mapping sont alors réalisées.

Pour chaque marqueur génotypé, le phénotype (quantité de virus dans un tissu, perte de poids post infection, jour de mort, etc.) est comparé pour chaque groupe de génotype. Si la valeur de phénotype varie en fonction du génotype à un marqueur, alors il y a un QTL à ce marqueur. Cette analyse est faite pour tous les marqueurs génotypés, pour lesquels le Logarithm of the Odds (LOD) score est calculé. Plus ce score est élevé, plus la probabilité de présence d'un QTL à cette position est forte. Cela permet d'obtenir des pics de LOD score dans le génome. Si ces pics dépassent les seuils de significativité ( $p=0,05$ ) ou suggestivité ( $p=0,63$ ), un QTL est présent à la position correspondante.

Les populations d'individus de seconde génération, F2 ou N2, sont chacun génétiquement unique, ce qui ne permet pas de répétabilité des expériences. A l'inverse, les populations génétiques de référence de souris sont des collections de lignées consanguines de souris qui ont donc l'avantage de pouvoir répéter les expériences sur des individus identiques. Une population génétique de référence que nous utilisons dans notre laboratoire est le Collaborative Cross (CC).

Le CC est une population de référence de souris dérivée de huit lignées fondatrices : cinq lignées de laboratoire dont trois lignées utilisées très largement utilisées en recherche, A/J, C57BL/6J (B6), 129S1/SvImJ, and deux lignées modèles de maladie, NOD/ShiLtj qui est un modèle de diabète de type 1 et NZO/HILtj qui est utilisées pour étudier l'obésité. Les trois autres lignées appartiennent aux trois différentes sous-espèces de *Mus musculus* : CAST/Eij (*Mus musculus castaneus*), PWK/Phj (*Mus musculus musculus*) et WSB/Eij (*Mus musculus domesticus*), et ont été sélectionnées pour maximiser la diversité génétique dans les lignées fondatrices. Les huit lignées fondatrices ont été croisées selon un schéma de croisement en entonnoir puis les individus ont été croisés par des croisements frère-sœur afin d'obtenir des lignées consanguines. Les lignées CC ont alors des génomes en mosaïque des génomes des huit lignées fondatrices. Les génotypes et haplotypes sont encodées par des lettres et couleurs qui correspondent à chaque lignée fondatrice : A/J, jaune (A) ; B6, gris (B) ; 129S1/SvImJ,



rose (C) ; NOD/ShiLtj, bleu foncé (D) ; NZO/HlLtj, bleu clair (E) ; CAST/Eij, vert (F) ; PWK/Phj, rouge (G) ; WSB/Eij, violet (H).

Mon laboratoire a utilisé les lignées du CC pour étudier la sensibilité au virus Zika (ZIKV). ZIKV appartient au genre *Flavivirus*, comme le virus de la dengue ou le virus de la fièvre jaune. ZIKV est transmis par les moustiques du genre *Aedes*, en particulier *Aedes aegypti* et *Aedes albopictus*. Chez l'humain adulte, l'infection par ZIKV est le plus souvent asymptomatique, et peut mener dans environ 20% des cas à des symptômes grippaux. Dans de plus rares cas, des complications neurologiques peuvent survenir comme le syndrome de Guillain-Barré ou des encéphalites. Si l'infection se produit pendant la grossesse, le virus peut être transmis de façon verticale de la mère au fœtus, ce qui peut provoquer des fausses couches ou des syndromes congénitaux. Ces syndromes peuvent être symptomatiques à la naissance et provoquer par exemple des microcéphalies. Autrement, les enfants peuvent naître asymptomatiques et développer des symptômes en grandissant comme des déficits neurodéveloppementaux.

Les souris ne sont pas naturellement sensibles à l'infection par ZIKV. En effet, chez l'homme, l'induction des ISGs est inhibée par une protéine virale, NS5, qui cible le facteur de transcription des ISGs, STAT2, et provoque sa dégradation. Chez la souris, NS5 n'est pas capable de cibler la protéine STAT2 murine, donc l'induction de ISGs est efficace. Pour étudier ZIKV chez la souris, il est possible d'utiliser un modèle de souris *Ifnar1* KO. Ces souris ne possèdent pas le récepteur aux IFN-I ce qui abolit la production des ISGs. Il est également possible d'utiliser un modèle de blocage pharmacologique grâce à un anticorps (MAR1-5A3) dirigé contre le récepteur IFNAR. Cet anticorps empêche la fixation des IFN-I sur leur récepteur, ce qui inhibe la production des ISGs. Ce second modèle est moins sévère que le modèle *Ifnar1* KO.

Le modèle *Ifnar1* KO existe sur deux fonds génétiques : B6 et 129S2/SvPas (129). Mon laboratoire a précédemment démontré l'influence du fond génétique sur la sensibilité à ZIKV en utilisant ces deux lignées. En effet, les souris B6 *Ifnar1* KO développaient des symptômes sévères, dont de la perte de poids, une fourrure ébouriffée, de l'ataxie et une paralysie des membres postérieurs, et étaient toutes moribondes ou mortes à 7 jours post-infection (dpi). A l'inverse, les souris 129 *Ifnar1* KO ne développaient que des symptômes modérés et seulement une souris sur 7 était morte à 9 dpi. Les souris B6 *Ifnar1* KO et 129 *Ifnar1* KO avaient la même charge virale plasmatique (PVL) à 2 dpi mais les souris B6 *Ifnar1* KO avaient une PVL plus élevée à 6 dpi.

La sensibilité des lignées CC à ZIKV a été étudiée en utilisant l'anticorps MAR1-5A3. Les souris ont été traitées par l'anticorps la veille de l'infection puis ont été suivies pendant 14 jours. A 2 et 6 jours post-infection, du sang a été prélevé pour mesurer la PVL. La plupart des lignées du CC, ainsi que la lignée B6, n'ont montré aucun symptôme après l'infection. Trois lignées du CC étaient symptomatiques, dont la lignée CC071 qui était la plus sensible et montrait de la mortalité post-infection. La PVL à 2 dpi avait une variabilité de 3 logs dans les lignées du CC, et n'était pas corrélée à la PVL à 6 jours post-infection. Deux lignées avec des phénotypes extrêmes ont été sélectionnées pour poursuivre les études, la lignée CC001 comme lignée « résistante », avec une faible charge virale plasmatique et pas de symptôme ; et la lignée CC071 comme lignée « sensible », avec une forte charge virale plasmatique et des symptômes sévères.

Ces deux lignées ont été étudiées *in vitro* en utilisant un modèle de fibroblastes embryonnaires

de souris (MEFs). Les MEFs de lignées résistantes, CC001 et B6 avaient un titre viral faible et constant post-infection, alors que les MEFs CC071 montraient un titre viral élevé et croissant entre 24 et 72 heures post-infection (hpi). En outre, l'expression des IFN-I était forte dès 24 hpi dans les MEFs CC001 et B6, alors que leur expression était retardée dans les MEFs CC071, ce qui peut expliquer le fort titre viral dans ces cellules.

L'objectif général de ma thèse était alors d'élucider le contrôle génétique de la sensibilité à ZIKV. Mon premier objectif était d'identifier le(s) facteur(s) génétique(s) qui contrôlent le retard d'induction des IFN-I dans les MEFs CC071, en utilisant une combinaison de méthodes génétiques et fonctionnelles. Mon second objectif était de déterminer si les variants identifiés *in vitro* pouvaient expliquer la sensibilité *in vivo* des souris CC071, et, si non, d'identifier les facteurs génétiques responsables grâce à des analyses génétiques de croisements. Enfin, mon troisième objectif était d'identifier des gènes modificateurs du KO *Ifnar1* qui expliquent les différences de sensibilité des souris B6 et 129 *Ifnar1* KO.

### VIII.3.2 . Résultats

#### VIII.3.2.1 . Le retard d'induction des IFN-I dans les MEFs CC071 est dû à une mutation perte de fonction dans le gène *Irf3*

Afin d'identifier les facteurs génétiques contrôlant le retard d'induction des IFN-I dans les MEFs CC071, j'ai utilisé une approche génétique de croisement à deux générations entre la lignée résistante CC001 et la lignée sensible CC071. Les MEFs (CC001 x CC071) F1 montraient une expression des IFN-I dès 24 hpi comme les MEFs CC001, donc les F1 ont été backcrossés avec CC071 pour produire des embryons N2. Chaque embryon backcross a été utilisé pour produire une lignée unique des MEFs. Les MEFs N2 montraient soit une induction rapide des IFN-I avec un faible titre viral ou un fort titre viral et une expression retardée des IFN-I. Une analyse de QTL mapping a révélé un QTL sur le chromosome 7, dans une région centrée sur le gène *Irf3*. Par ailleurs, CC001 et CC071 ont hérité de l'haplotype CAST/Eij dans cette région, suggérant la présence d'un variant privé dans le génome de CC071.

Des analyses de données transcriptomiques sur les MEFs CC001 et CC071 ont révélé un épissage anormal des exons du gène *Irf3* dans les cellules CC071. En effet, l'exon 6 du gène n'est pas épissé avec l'exon 7 mais avec une courte région intronique, ce qui mène à la production d'un ARNm qui ne contient pas les deux derniers exons du gène (exons 7 et 8). En effet, un Western blot utilisant un anticorps spécifique de la région C-terminal de la protéine IRF3 a montré que la protéine IRF3 entière est absente des cellules CC071. Cependant, la potentielle présence d'une forme tronquée de la protéine reste à être démontrée. La protéine IRF3 étant phosphorylée dans sa région C-terminale, cette forme phosphorylée de la protéine était donc absente des cellules CC071.

Le rôle de cette mutation dans le gène *Irf3* dans le retard d'induction des IFN-I et le fort titre viral des MEFs CC071 a été démontré par un test de complémentation fonctionnel. Cependant, le phénotypage de souris B6 *Irf3* KO a montré qu'une déficience en *Irf3* n'est pas suffisante pour induire une sensibilité à l'infection par ZIKV *in vivo*, suggérant la présence d'autres facteurs de sensibilité dans la lignée CC071, expliquant la forte PVL et les symptômes cliniques observés chez ces souris.

### VIII.3.2.2 . Décryptage de la sensibilité *in vivo* des souris CC071

Afin d'identifier les allèles de sensibilité de la lignée CC071 expliquant sa sensibilité *in vivo* à l'infection ZIKV, j'ai réalisé des analyses génétiques sur deux croisements. Premièrement, j'ai analysé une population de souris backcross (CC071 x B6 *Irf3* KO) x CC071 (désigné ci-dessous comme N2-*Irf3*). Vu que toutes les souris N2-*Irf3* possèdent deux allèles *Irf3* non fonctionnels, ce croisement maximise la puissance de détecter d'autres allèles de sensibilité. Ce croisement avait pour but principal d'identifier des QTLs associés à la PVL.

Séparément, j'ai recherché des QTLs expliquant les symptômes sévères observés dans les souris CC071 après infection. Pour cela, j'ai utilisé un modèle plus sévère en croisant CC071 avec 129 *Ifnar1* KO pour produire une population de souris F2 (désignée ci-dessous comme F2-*Ifnar1*). Nous avons choisi la lignée 129 *Ifnar1* KO car elle est moins sensible à l'infection que la lignée B6 *Ifnar1* KO. De plus, comme toutes les souris F2 avaient une réponse IFN-I abolie, cela favorisait l'identification de gènes de sensibilité en dehors de cette voie.

Les souris N2-*Irf3* montraient une faible diversité de signes cliniques et perte de poids, comme attendu. Inversement, elles montraient une forte diversité dans les PVL à 2 et 6 dpi qui s'étendaient sur 3 logs. L'analyse de QTL mapping a permis l'identification de 3 QTLs significatifs.

*Zika susceptibility locus 1 (Zs1)* était positionné sur le chromosome 12 et était associé à la PVL à 2 dpi avec un LOD score de 6,610 et une p-value inférieure à 0,001. Deux QTLs suggestifs, sur les chromosomes 2 et 15 étaient également associés à la PVL à 2 dpi, et les 3 QTLs montraient des effets additifs. *Zs2* était positionné sur le chromosome 12 et était associé à la PVL à 6 dpi. Deux QTLs suggestifs sur les chromosomes 2 et 7 ont été également associés à la PVL à 6 dpi. *Zs3* sur le chromosome 7 était associé à la perte poids à 5 dpi. Pour le score clinique, aucun QTL significatif a été identifié.

Les souris F2-*Ifnar1* quant à elles ont montré une forte diversité de signes cliniques post infection, en termes de scores cliniques et perte de poids, avec un taux de mortalité de 46,4%. Contrairement aux lignées parentales de ce croisement qui ne montraient pas de différence significative de PVL, les souris F2-*Ifnar1* montraient une variabilité de 2 logs à 2 dpi et de 4 logs à 6 dpi. L'analyse de QTL mapping a permis l'identification de 3 QTLs significatifs.

*Zs4* était positionné sur le chromosome 2 et était associé à la PVL à 6 dpi. De façon intéressante, un QTL suggestif sur le chromosome 2 avait également été identifié dans les N2-*Irf3* dans cette région du génome. Il est donc possible que le même QTL ait été identifié dans les deux croisements. *Zs5* et *Zs6* étaient localisés sur le chromosome 7 et associés aux phénotypes de mortalité des souris. En effet, *Zs5* était associé au phénotype binaire de survie, alors que *Zs6* était associé au jour de mort des souris ayant succombé à l'infection. L'analyse de QTL mapping pour les scores cliniques a permis l'identification de 9 QTLs dont 2 significatifs, *Zs5* et *Zs6*. Le suivi des scores cliniques ayant été réalisé pendant 14 jours, la cinétique des QTLs a pu être analysés. Alors que la plupart des QTLs étaient associés au score clinique en phase précoce après infection, certains étaient associés au score clinique en phase tardive. Le même résultat a pu être obtenu pour l'analyse de la perte de poids.

### VIII.3.2.3 . Les facteurs génétiques modifiant la sensibilité des souris *Ifnar1* KO

Pour identifier les facteurs génétiques modifiant la sensibilité des souris *Ifnar1* KO, une F2 ente B6 *Ifnar1* KO et 129 *Ifnar1* KO a précédemment été réalisée dans mon laboratoire. Le taux de mortalité dans la population était faible, probablement car les allèles de sensibilité de B6 *Ifnar1* KO sont récessifs. Deux QTLs, sur les chromosomes 5 et 12, associés respectivement à la PVL à 2 dpi et le jour de mort, ont été identifiés.

Afin d'augmenter le taux de mortalité dans la population de seconde génération, j'ai produit et analysé une population de souris (B6 *Ifnar1* KO x 129 *Ifnar1* KO) x B6 *Ifnar1* KO. Les souris N2 ont montré une grande diversité de signes cliniques et de perte de poids, avec un taux de mortalité de 45,2%. Les souris N2 montraient une faible variabilité de PVL à 2 dpi, mais des PVL s'étendant sur environ 3 logs à 6 dpi, de façon cohérente avec les valeurs de PVL des lignées parentales. L'analyse de QTL mapping a permis d'identifier 8 QTLs significatifs.

*Zs/7* (sur le chromosome 1) et *Zs/8* (sur le chromosome 3) étaient associés avec la PVL à 6 dpi, et *Zs/9* (sur le chromosome X) était associé avec la PVL à 2 dpi. *Zs/10* était associé avec la survie des souris et colocalisait avec *Zs/7* sur le chromosome 1, ce qui pourrait suggérer que le même gène contrôle la PVL à 6 dpi et la mortalité des souris *Ifnar1* KO. *Zs/11* était localisé sur le chromosome 4 et était associé à la survie, au jour de mort et au score clinique des souris. De plus, *Zs/10* et *Zs/11* avaient des effets additifs sur la survie des souris. *Zs/12* était associé au poids de souris entre 5 et 8 dpi. Enfin, *Zs/13* et *Zs/14* étaient associés aux scores cliniques des souris.

Étonnamment, l'effet des deux QTLs identifiés dans les F2s n'a pas été retrouvé dans la population de N2s. Inversement, parmi les 8 QTLs identifiés dans les N2s, 5 montraient des effets alléliques cohérents dans les F2s, les plaçant comme prioritaires pour de futures analyses.

### VIII.3.3 . Discussion

Mon projet a premièrement démontré l'intérêt des modèles cellulaires dans les études de génétique chez la souris grâce au backcross entre CC001 et CC071. En effet, 90 lignées de MEFs N2 avaient été produites, donc 51 ont été utilisées pour cette étude. Ces lignées pourront être réutilisées pour de futures expériences, mais en nombre limité car pour chaque embryon N2, génétiquement unique, un nombre fini de tubes de MEFs a été produit. Cette stratégie pourrait alors être étendue en utilisant une population génétique de référence, en produisant des MEFs de CC, ce qui permettrait alors un nombre infini d'expérience sur des MEFs génétiquement identiques. De plus, l'étape de génotypage ne serait pas nécessaire.

J'ai ensuite analysé des croisements de lignées de souris *in vivo*, en me focalisant sur deux types de phénotypes : la charge virale et les symptômes cliniques. J'ai pour cela utilisé deux modèles expérimentaux d'infection par ZIKV. Premièrement, j'ai analysé une population de souris N2 traitées par l'anticorps bloquant le récepteur IFNAR, puis j'ai analysé deux populations de souris *Ifnar1* KO. Ces deux modèles expérimentaux ont permis l'identification de QTLs contrôlant différents aspects de la maladie Zika. En effet, le modèle de traitement par anticorps a particulièrement été utile pour identifier des QTLs associés à la charge virale, alors que le modèle *Ifnar1* KO a permis d'identifier des QTLs associés aux symptômes cliniques.

L'étape suivant l'identification de QTLs est l'identification de gène(s) candidat(s) dans l'intervalle du QTL afin de valider leur rôle dans la sensibilité à la maladie virale. Pour ce faire, plusieurs critères peuvent être utilisés afin de filtrer les gènes présents dans l'intervalle. Par exemple, j'ai trié les gènes dans l'intervalle de *Zs11* en fonction de la présence de variants délétères dans une des deux lignées parentales du croisement, une différence d'expression du gène entre les lignées parentales, et une fonction immunitaire connue du gène. Aucun des 108 gènes de l'intervalle ne remplissait ces trois critères. Cela s'explique par les biais de chacun de ces filtres. En effet, le QTL peut être provoqué par un variant non délétère ou non codant, peut ne pas provoquer de différence d'expression du gène, et le gène peut ne pas avoir de fonction immunitaire connue. Cela montre qu'il n'existe pas une seule méthode infaillible pour identifier le bon gène candidat, mais qu'il faut multiplier les méthodes afin de trouver et de tester les meilleurs candidats.

Pour finir, mon projet a montré l'intérêt du CC pour étudier les maladies virales. Tester une collection de lignées permet d'identifier de nouveaux modèles de sensibilité, comme la lignée CC071. Cette lignée est sensible à ZIKV mais également à d'autres virus du genre des Flavivirus (virus de la dengue, virus Powassan, virus du Nil occidental) mais également des virus d'autres familles comme le virus de la Vallée du Rift. Cette lignée est un modèle complexe de sensibilité aux maladies virales, comme démontré par le grand nombre de QTLs identifiés dans ce projet. Il est alors possible que chaque facteur génétique contribue de façon différente à la sensibilité des souris CC071 à chaque virus. De façon plus générale, utiliser les lignées du CC permet de mieux représenter la diversité génétique humaine. En effet, l'utilisation du modèle murin est parfois critiquée à causes de difficulté à transposer les résultats obtenus chez la souris à l'homme. Un des facteurs participant à cette difficulté est l'utilisation d'un seul fond génétique pour représenter la vaste diversité humaine. Cela peut mener à des observations anecdotiques spécifiques d'un fond génétique. Utiliser les lignées du CC permet d'éviter ce problème en étudiant non pas un phénotype mais une étendue de phénotype. En conclusion, le CC est un modèle de la diversité génétique de la population humaine qui facilite la transposition des études souris à l'homme.