



Recherche de facteurs génétiques contrôlant la résistance de lignées de souris consanguines à une infection expérimentale par *Yersinia pestis*, l'agent de la peste.

Lucie Chevallier

► To cite this version:

Lucie Chevallier. Recherche de facteurs génétiques contrôlant la résistance de lignées de souris consanguines à une infection expérimentale par *Yersinia pestis*, l'agent de la peste.. Médecine humaine et pathologie. AgroParisTech, 2012. Français. NNT : 2012AGPT0077 . pastel-00880257

HAL Id: pastel-00880257

<https://pastel.hal.science/pastel-00880257>

Submitted on 5 Nov 2013

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.



Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

**L'Institut des Sciences et Industries
du Vivant et de l'Environnement**

(AgroParisTech)

Spécialité : Génétique

présentée et soutenue publiquement par

Lucie CHEVALLIER

le 5 décembre 2012

Recherche de facteurs génétiques contrôlant la résistance de lignées de souris consanguines à une infection expérimentale par *Yersinia pestis*, l'agent de la peste.

Directeur de thèse : **Dr Xavier MONTAGUTELLI**

Travail réalisé à l'Institut Pasteur, CNRS URA 2578, Unité de Génétique Fonctionnelle de la Souris,
25 rue du Dr Roux, 75724 PARIS Cedex 15

Jury

Mme Diana TRONIK-LE ROUX, Docteur, Laboratoire de Génomique et Hématopoïèse, CEA
M. Claude LIBERT, Professeur, *Mouse Genetics in Inflammation*, Université de Gand
Mme Elisabeth CARNIEL, Chef de laboratoire, Unité des *Yersinia*, Institut Pasteur
Mlle Claire DANE, Docteur, Service de Santé des Armées, DGA
M. Xavier Montagutelli, Chef de laboratoire, Unité de Génétique Fonctionnelle de la Souris
M. Etienne VERRIER, Professeur, UFR Génétique, Elevage et Reproduction, AgroParisTech

Rapporteur
Rapporteur
Examinatrice
Examinatrice
Directeur de thèse
Président du jury

Remerciements

Je remercie tout d'abord la Délégation Générale pour l'Armement qui a financé cette thèse. Ce travail a été réalisé dans l'Unité de Génétique fonctionnelle de la Souris, soutenue financièrement par Merck-Serono et par le projet Initiative d'excellence "Integrative Biology of Emerging Infectious Diseases".

J'adresse tous mes remerciements aux membres du jury qui ont accepté de juger ce travail.

Je remercie le Professeur Etienne Verrier d'avoir accepté de présider ce jury.

Je remercie le Docteur Diana Tronik – Le Roux et le Professeur Claude Libert qui ont pris de leur temps, m'ont prodigué leurs conseils et ont effectué les corrections de mon manuscrit en qualité de rapporteurs.

Je remercie le Docteur Elisabeth Carniel et le Docteur Claire Dane qui ont accepté d'examiner mon travail.

Je remercie le Docteur Xavier Montagutelli, qui a dirigé ma thèse. Merci Xavier de m'avoir proposé ce sujet, de m'avoir encadré pendant ces trois années, de m'avoir tant appris. Je peux l'avouer maintenant, j'ai raté l'examen du CESAM, oui je sais c'est honteux, j'ai eu 9,5/20 et il fallait avoir 10... mais j'ai rien contre les stats c'est promis !

Je tiens également à remercier toute l'équipe de l'unité de Génétique Fonctionnelle de la Souris, dirigée par Jean-Jacques Panthier, pour ces trois années passées au laboratoire. Je remercie particulièrement Charlène Blanchet, Laurent Guillemot, Jean Jaubert, Dominique Simon et Satoko Tokuda pour leur implication dans le projet et leurs discussions scientifiques (et moins scientifiques).

Je remercie également l'unité des *Yersinia*, dirigée par Elisabeth Carniel avec qui nous sommes en étroite collaboration sur ce sujet de recherche. Je remercie tout particulièrement Emilia Pachulec et Christian Demeure pour leurs interactions scientifiques constantes.

Je remercie également l'unité *Infection Genetics* du *Helmholtz Institute* à Braunschweig, pour leur collaboration dans l'expérience de puces à ADN. Je remercie tout particulièrement Claudia Pommerenke qui m'a accueillie très chaleureusement pour analyser les résultats des puces.

Je remercie toute l'équipe de l'animalerie Centrale, en particulier Isabelle Lanctin qui a élevé les souris SEG.

Je remercie également Jean-Louis Guénet pour tous ces conseils avisés et Geneviève Milon qui m'a épaulée au cours de ces trois ans.

Je remercie les membres de mon comité de thèse, Jean-Marc Cavaillon, Sylvie Garcia, Armelle Phalipon et Françoise Ramisse pour leur suivi régulier du projet et leurs conseils.

Je remercie également Hélène Bierne, ma tutrice, qui m'a également suivie pendant ces trois ans. Merci pour tous les conseils et toutes les discussions.

Je remercie également ceux qui m'ont fait aimer la science, Marie Abitbol, Catherine André, Professeur Parodi, Dr Douglas McGregor, Professeur Fraser, un grand merci.

J'en arrive à ma famille, qui n'a rien fait de spécial dans la recherche de facteur génétiques impliqués dans la résistance à la peste, mais que je remercie d'être là.

A Papapierre, pour nos sorties dominicales en scooter à fond dans le parc de Sceaux.

A Mamanfrançoise pour le lapsang-souchang dominical et tes pulls (que je lorgne encore).

A Philippe, pour la physique des particules et le tour au marché dominical d'il y a bien longtemps, mais que je n'ai pas oublié.

A Kaki, pour les rédactions de français et la salade de fruit.

A ma mère, qui me demande depuis 2006 quand est-ce qu'elle est finie cette thèse.

A mon père qui gribouille toujours partout.

A Lili, tu es sûre que tu veux faire une thèse sur la soudure par fusion ?

A Marlou, tu penses que c'est une bonne idée que fufu fasse une thèse sur la soudure par fusion ?

A Gnoc-Gnoc, again.

A mes amis :

Soudeh darling, pour nos let's be realitic eleven o'clock Saturday and Sunday breakfasts. Merci d'être là my personal designer.

Rachel, à nos cafés au lait du matin,

Sophie, en espérant qu'on trouve un moment dans les dix prochaines années pour se prendre un thé !

A Léna, je crois que je prends plein d'ondes négatives depuis qu'il n'y a plus tes cactus dans la pièce d'à côté...

Eloi, oulala gros conflit d'intérêt, je m'arrête.

Ingi, à notre folle aventure d'ambassadeurs ParisTech !

Aurélien, à toutes ces pièces que l'on a jouées ! Contente que tu kiffes grave la recherche, mec (Renan style).

A toute la troupe de théâtre, à tous ces mardis, jeudis, et parfois lundis passés à chanter et jouer.

To the lunch gang, Fernando, Maria², Katie, Katha, Fred, Robin, Marcio, Brian, Jason, Felix, Nat, Mateo, Debo, Ale, Sabah, Toni, Amélita, Emilienne. All of you guys contributed significantly to my happiness in Pasteur (and outside like Istanbul, Berlin, Brastilava... $p < 0.0001$). You made me changed, I think I might be less shy... I love you f***ers. See you for lunch at 12.30.

A mes colocos, Max, Ingrid et Alexis pour ces trois années passées dans notre appart de folie !

A Julie,

And the last ones but not least,

A SEG et B6 et toutes les congéniques (sans oublier les SWISS) dont l'immunité et sûrement d'autres processus ont passionnés les foules pendant trois ans. A NM, OG, OD, ODG, OGG, ODD, GGD, DDG, GGDD, GGGD, DDDG, et même peut-être GGGDDD.

A Minh, à nous, à nos voyages dans la vie, en évitant Vienne, même pour une escale.

Sommaire

Liste des abréviations.....	8
Résumé en français et mots-clés	12
Résumé en anglais et mots-clés	13
Introduction.....	14
ETAT DES CONNAISSANCES	16
Partie 1. Comment un hôte répond-t-il à la rencontre avec un micro-organisme pathogène ?..	17
1. La réponse au niveau de l'organisme.....	18
1.1. La reconnaissance du pathogène	18
1.1.1. « Das ist ja toll ! », les <i>Toll-like</i> récepteurs	18
1.1.2. Les <i>NOD-like</i> Récepteurs, des senseurs cytoplasmiques.....	20
1.1.3. Les senseurs alternatifs.....	21
1.2. La réaction de l'organisme face au pathogène	21
2. La réponse à l'échelle cellulaire.....	22
2.1. Les cellules de l'immunité innée.....	22
2.1.1. Les cellules phagocytaires mononuclées : monocytes et macrophages.....	22
2.1.2. Les neutrophiles.....	23
2.1.3. Les cellules <i>Natural Killer</i> (NK)	24
2.2. Les conséquences, à l'échelle cellulaire, de la reconnaissance du pathogène.....	24
2.2.1. La transduction du signal en aval des TLR	24
2.2.2. Les voies de signalisation des récepteurs cytoplasmiques.....	26
2.2.3. Les médiateurs de l'inflammation induits par l'activation de NF-κB	27
2.2.4. L'activation des <i>Interferon Regulatory Factors</i> (IRFs).....	27
2.2.5. La phagocytose	28
Partie 2. Puissance et limites des études génétiques de la sensibilité aux maladies infectieuses chez l'Homme	30
1. L'approche par gène candidat	31
2. Les analyses criblant l'ensemble du génome	32
2.1. Les analyses de liaison ou <i>Genome Wide Scan</i>	32
2.2. Les études d'association pan-génomiques ou <i>Genome Wide Association Study</i> (GWAS) pour l'étude des maladies communes	33
3. L'étude des variants rares à effet fort.....	35
4. La recherche de variants structuraux	36
Partie 3. La pertinence du modèle souris dans l'étude des maladies infectieuses	38
Partie 4. Les différentes stratégies utilisées pour identifier les gènes impliqués dans la sensibilité aux maladies infectieuses chez la Souris	43
1. La réponse de l'hôte à une infection est un phénomène hautement complexe	43
1.1. Le choix du phénotype étudié pour caractériser un processus infectieux	44
1.1.1. La mort ou la survie.....	44
1.1.2. Le poids	46
1.1.3. La dissémination et la charge microbienne dans les organes cibles	46
1.1.4. Les lésions histologiques	47
1.1.5. La réponse inflammatoire.....	48
2. La mutagenèse à l'ENU : une approche aléatoire mais efficace.....	49
3. Croisements entre deux lignées différant pour leur vulnérabilité	51
4. L'utilisation de populations génétiques de référence.....	53
4.1. Les lignées consanguines	53
4.1.1. Caractéristiques des lignées consanguines	53
4.1.2. La cartographie d'association ou <i>Haplotype Association Mapping</i>	54
4.2. Les lignées recombinantes consanguines, recombinantes congéniques et consomiques	55
4.2.1. Les lignées recombinantes consanguines ou <i>Recombinant Inbred Strains</i> (RI).	55

4.2.2.	Les lignées recombinantes congéniques ou <i>Recombinant Congenic Strains</i> (RC)	57
4.2.3.	Les lignées consomiques ou <i>Chromosome Substitution Strains</i>	58
4.3.	Le <i>Collaborative Cross</i> (CC)	60
5.	Etudier les effets individuels des QTL : les lignées congéniques	61
6.	Comment aller jusqu'au gène ?	67
7.	Bilan : qu'est-ce qui pose encore problème pour aller plus vite ?	70
Partie 5.	L'apport des analyses transcriptomiques à la compréhension des interactions hôte-pathogène chez la Souris	72
1.	Pourquoi étudier le transcriptome ?	72
2.	L'apport des études transcriptomiques <i>in vitro</i>	74
3.	L'apport des études transcriptomiques <i>in vivo</i>	77
4.	Combiner l'analyse transcriptomique et l'approche QTL : un moyen de cibler des gènes candidats ?	80
4.1.	Les différentes stratégies combinant QTL et analyse transcriptomique	81
4.2.	L'approche combinée QTL et analyse de transcriptome est-elle vraiment efficace pour cibler des candidats ?	84
5.	L'apport des analyses transcriptomiques dans la compréhension des interactions entre <i>Yersinia pestis</i> et son hôte	86
Partie 6.	<i>Yersinia pestis</i> , l'agent de la peste	91
1.	La peste	91
1.1.	Le cycle de la peste	92
1.2.	Les différentes formes de la maladie	93
1.3.	L'état actuel des foyers de peste	94
1.4.	Diagnostic	94
1.5.	Traitement	95
1.6.	Vaccination	95
1.7.	Apparition de résistances aux antibiotiques	95
1.8.	Bioterrorisme	96
2.	Les mécanismes de virulence de <i>Yersinia pestis</i>	96
3.	L'apport des modèles animaux dans l'exploration de la pathogénèse des différentes formes de la peste	99
3.1.	Les modèles de la forme bubonique de la peste	100
3.2.	Les modèles de la forme pulmonaire de la peste	101
4.	La recherche de facteurs génétiques déterminant la sensibilité de l'hôte à une infection par <i>Y. pestis</i>	103
Partie 7.	Etat des connaissances sur la résistance à la peste de la lignée SEG/Pas	105
OBJECTIFS DU PROJET		109
Les objectifs de mon travail de thèse		110
MATERIELS ET METHODES		111
1.	Procédures expérimentales utilisées dans l'analyse complémentaire des lignées congéniques	112
2.	Procédures expérimentales utilisées dans l'analyse comparative de transcriptome	112
2.1.	Animaux	112
2.2.	Culture de la bactérie	113
2.3.	Prélèvement de macrophages, culture cellulaire et exposition des macrophages à <i>Y. pestis</i>	113
2.4.	Extraction des ARN	114
2.5.	Expérience de puces à ADN	114
2.6.	Analyse des données de puces à ADN	114
2.7.	RT-PCR quantitative	115
2.7.1.	Synthèse d'ADNc	115
2.7.2.	Pré-amplification des ADNc	116
2.7.3.	PCR quantitative en temps réel	116
2.7.4.	Analyse des résultats	116

RESULTATS117

Introduction.....	118
Partie 1. L'analyse des lignées congéniques permet de disséquer l'architecture de la résistance à la peste chez SEG/Pas	119
1. La résistance à la peste chez SEG/Pas requiert l'action combinée d'au moins quatre facteurs génétiques	119
2. 2. Résultats complémentaires.....	124
2. 2. Résultats complémentaires.....	124
2.1. Les femelles homozygotes pour <i>Yprl1</i> ne survivent pas plus que les femelles B6 ...	124
2.2. Les sous-régions chromosomiques de <i>cg1</i> n'ont pas d'effet sur la survie	124
2.3. Les trois sous-régions de <i>cg1</i> ne présentent pas d'effet, même associées à <i>cg3</i>	125
2.4. Aucun effet de <i>cg1</i> n'est détecté sur un fonds génétique hybride.....	125
2.5. Chez les femelles BSB <i>Yprl1</i> agirait plus tardivement que <i>Yprl2</i> et <i>Yprl3</i> sur le taux de survie	126
Partie 2. Analyse du profil transcriptionnel de macrophages péritonéaux de SEG et de B6 en réponse à <i>Y. pestis</i>	127
1. Le choix du modèle.....	127
1.1. Les études transcriptomiques de l'hôte en réponse à <i>Y. pestis</i> antérieures	127
1.2. Le choix d'un modèle <i>ex vivo</i>	128
1.3. Le choix du type cellulaire	128
2. Description du profil transcriptionnel des macrophages péritonéaux de souris B6 en réponse à une exposition à <i>Y. pestis</i> pendant 3 heures	131
2.1. B6 présente davantage de gènes réprimés qu'activés en réponse à l'exposition à la bactérie	131
2.2. L'analyse fonctionnelle des gènes DE chez B6 en réponse à <i>Y. pestis</i> révèle un impact de la bactérie sur plusieurs voies cellulaires clés	132
2.3. L'analyse des voies de signalisation perturbées chez B6 révèle une faible réponse à <i>Y. pestis</i>	134
3. Description comparative du profil transcriptionnel des macrophages péritonéaux de souris résistantes SEG et de souris sensibles B6 en réponse à une exposition à <i>Y. pestis</i> pendant 3 heures	135
3.1. Les macrophages péritonéaux de SEG réagissent plus fortement que ceux de B6 après trois heures d'exposition à <i>Y. pestis</i>	135
3.2. Les spécificités de lignée à l'homéostasie.....	135
3.3. Description de la fonction des gènes DE chez SEG.....	137
3.4. L'analyse des voies de signalisation perturbées chez SEG révèle des différences majeures dans la mise en place de la défense de l'hôte par rapport à la lignée sensible B6 ..	140
4. Analyse différentielle des gènes exprimés dans les macrophages péritonéaux de SEG et de B6 en réponse à une exposition à <i>Y. pestis</i>	142
5. Les gènes impliqués dans la réponse pro-inflammatoire mise en place chez SEG sont majoritairement des cytokines et des chimiokines	143
6. Etude cinétique de 44 gènes DR chez SEG, B6 et <i>cg2^{B/S}-cg3^{B/S}</i>	144
6.1. Les différences d'expression basale entre les lignées	145
6.2. Comment varie l'expression des 44 gènes chez B6, SEG et <i>cg2^{B/S}-cg3^{B/S}</i> au cours du temps, après exposition à la bactérie ?	146
6.3. Comparaison des trois lignées à chacun des trois temps de la cinétique.....	147
6.4. Comparaison des résultats issus des expériences de puces à ADN et de qRT-PCR	148
DISCUSSION GENERALE	150
Discussion.....	151
1. Le procédé expérimental et l'analyse des données de transcriptome.....	151
1.1. Le choix des macrophages comme modèle d'étude de la réponse infectieuse.....	151
1.2. Quelle peut être la source des discordances observées entre les expériences de puces à ADN et de PCR quantitative ?	152
1.2.1. Les différences de statut immunitaire des souris.....	152
1.2.2. Le polymorphisme des sondes de la puce.....	153

1.3.	Hypothèses quant aux gènes des QTL mis en jeu dans les mécanismes différentiellement activés entre SEG et B6	154
2.	Discussion sur l'architecture génétique de la résistance à la peste chez SEG	160
2.1.	Des lignées congéniques ne suffisent pas toujours pour disséquer des QTL	160
2.1.1.	<i>Ypr11</i> : une vraie fausse association ?	160
2.1.2.	La nécessité d'identifier des sous-phénotypes quantitatifs.....	161
2.1.3.	L'utilisation d'un fonds hybride (F1) pour disséquer <i>Ypr13</i>	163
2.1.4.	Fluctuations du phénotype de résistance des souris SEG	164
2.2.	Plus de gènes que de QTL.....	165
2.3.	Pas d'épistasie ?	166
3.	Conclusion générale.....	166
REFERENCES BIBLIOGRAPHIQUES		169

Liste des abréviations

A

Ac : Anticorps
 ADN : Acide désoxyribonucléique
 ADNc : Acide désoxyribonucléique complémentaire
 Ag : Antigène
 ANOVA : *ANalysis Of VAriance*
 AP1 : *Activator Protein 1*
 Arhgdib : *Rho, GDP dissociation inhibitor (GDI) beta*
 ARN : Acide ribonucléique
 ARNi : ARN interférence
 ASC : *Apoptosis-associated Speck-like protein containing a CARD*
 ATP : Adénosine triphosphate

B

BC : *Backcross*
 BCG : Bacille Calmette-Guérin
 Bcl2 : *B-cell leukaemia/lymphoma 2*
 Bid : *BH3 Interacting Domain death agonist*
 BLAST : *Basic Local Alignment Search Tool*
 BSB : [(C57BL/6J x SEG/Pas)F1 x C57BL/6J]s
 BXD : C57BL/6J x DBA/2

C

C5 : Facteur 5 du complément
 CARD : *Caspase-Activating and Recruitment Domain*
 Carg : *Candida albicans resistance gene*
 Casp : Caspase
 CBP : *CREB-Binding Protein*
 CC : *Collaborative Cross*
 CCDC122 : *Coiled-coil Domain-Containing 122*
 CCL : *Chemokine (C-C motif) Ligand*
 CCR : *C-C chemokine Receptor*
 CD : *Cluster of Differentiation*
 CDC : *Center for Disease Control and prevention*
 Cflar : *CASP8 and FADD-like apoptosis regulator*
 Cg : Congénique
 Char : *Plasmodium chabaudi malaria resistance*
 Chr. : Chromosome
 Chuk : *Conserved helix-loop-helix ubiquitous kinase*
 cIAP : *cellular Inhibitor of Apoptosis*
 Clec : *C-type lectin like*

cM : CentiMorgan

CMH : Complexe Majeur d'Histocompatibilité
 CNV : *Copy Number Variant*
 CO92 : Colorado 92
 CPA : Cellule Présentatrice d'Antigène
 CR3 : *Complement Receptor 3*
 CREB : *Cyclic AMP-Responsive Element-Binding protein*
 CRP : *C-Reactive Protein*
 CSS : *Chromosome Substitution Strain*
 Ct : *Cycle threshold*
 CXCL : *Chemokine (C-X-C motif) Ligand*
 CXCR : *C-X-C chemokine Receptor*
 CYBB : *Cytochrome B-245, Beta polypeptide*

D

DAVID : *Database for Annotation, Visualization and Integrated Discovery*
 Daxx : *Fas death domain-associated protein*
 DE : Différentiellement Exprimé
 DL50 : Dose Létale 50
 DO : *Diversity Outbred*
 DR : Différentiellement Régulé
 Ds : *Double stranded*

E

ENU : Ethyl-Nitro-Urée
 ERK : *Extracellular signal-Regulated Kinase*
 EUCOMM : *European Conditional Mouse Mutagenesis Program*

F

FADD : *Fas-Associated protein with Death Domain*
 FC : *Fold Change*
 FcR : Récepteurs de la partie Fc des immunoglobulines
 FDR : *False-Discovery Rate*
 FHL3 : Lymphohistiocytose hémophagocytaire familiale de type 3

G

G-CSF : *Granulocyte Colony-Stimulating Factor*
 GM-CSF : *Granulocyte-Macrophage Colony Stimulating Factor*
 GO : *Gene Ontology*
 GRO/KC : *Growth-related oncogene/keratinocyte-derived chemokine*
 GWAS : *Genome Wide Association Study*

GWS : *Genome Wide Scan*

H

H : *Heure*
H-2D : *Histocompatibility 2, D*
HAM : *Haplotype Association Mapping*
Hb : *Hémoglobine*
HDAC11 : *Histone déacétylase 11*
HLA : *Human Leucocyte Antigen*
HMGB1 : *High-Mobility Group Box 1*
HPI : *Îlot de Haute Pathogénicité*
HSP : *Heat-Shock Protein*
HZI : *Helmholtz Zentrum für Infektionforschung*

I

IBD : *Identity By Descent*
ICAM-1 : *Intercellular Adhesion Molecule 1*
ID : *Intra-dermique*
iE-DAP : *Acide-gamma-D-glutamyl-meso-diaminopimelic*
IFN : *Interféron (Interféron)*
IFNGR1 : *Récepteur 1 à l'interféron gamma*
Ig : *Immunoglobuline*
Ikbk : *Inhibitor of kappa B kinase*
IKK : *IκB Kinase*
IL : *Interleukin (Interleukine)*
IL-R : *Interleukin receptor*
iNOS ou NOS2 : *Inducible Nitric Oxide Synthetase*
IP : *Intra-péritonéal*
IPA : *Ingenuity Pathways Analysis*
IPAF : *ICE Protease-Activating Factor*
Ipr1 : *Intracellular pathogen resistance 1*
IRAK2 : *Interleukin-1 Receptor-Associated Kinase 2*
IRAK4 : *IL-1 Receptor-Associated Kinase 4*
IRCS : *Interspecific Recombinant Congenic Strains*
IRF : *Interferon Regulatory Factor*
ISG : *Interferon Stimulated Gene*
Itgb2 : *Integrin beta 2*
Ity : *Immunity to Salmonella typhimurium*
IV : *Intra-veineux*
IκBa : *Inhibitor of κ light chain gene enhancer in B cells, alpha*
IκB : *Inhibitor of kappa B*

J

JAK : *Janus kinase*
Jun : *Jun oncogene*

K

Kb : *Kilobase*

KC : *Keratinocyte-derived cytokine*
KEGG : *Kyoto Encyclopedia of Genes and Genomes*
KI : *Knock-in*
Klr : *Killer cell Lectin-like receptor*
KO : *Knock-out*
KOMP : *Knockout Mouse Project*

L

LACC1 : *Laccase (multicopper oxidoreductase) domain-containing 1*
LED : *Lupus érythémateux disséminé*
LRRK2 : *Leucine-Rich Repeat Kinase 2*
Lmr : *Leishmania major response*
LOD : *Likelihood of odds*
LPB : *LPS-Binding Protein*
LPS : *Lipopolysaccharide*
LRR : *Leucine rich repeat*
LTA : *Lipoteichoic acid*
Ltb : *Lymphotoxin B*

M

M. m. : *Mus musculus*
Map3k1 : *Mitogen-activated protein kinase kinase kinase 1*
MAPK : *Mitogen-activated protein kinase*
Mb : *Mégabase*
MCMV : *Mouse cytomegalovirus*
MCP-1 : *Monocyte-Chemoattractant Protein 1*
MDP : *Muramyl dipeptide*
MEF : *Mouse Embryonic Fibroblast*
MGI : *Mouse Genome Informatics*
MIP1α : *Macrophage Inflammatory Protein 1 alpha*
MIT : *Massachusetts Institut of Technology*
Mmp : *Matrix metalloproteinase*
MOI : *Multiplicity of infection*
Mpl : *Montpellier*
MSMD : *Mendelian Susceptibility to Mycobacterial Disease*
Mx1 : *Myxovirus resistance 1*
MyD88 : *Myeloid Differentiation Factor 88*

N

NALP ou NLRP : *Nacht domain-, Leucine-rich repeat-, and PYD-Containing Protein*
Ncf2 : *Neutrophil cytosolic factor 2*
NEMO : *NF-kappa-B essential modulator ou inhibitor of nuclear factor kappa-B kinase subunit gamma (IKK-γ)*
NET : *Neutrophil extracellular traps*
NF-κB : *Nuclear factor -κB*
NK : *Natural killer*
NLR : *NOD-like receptor*

NLRC : *Nacht domain-, Leucine-rich repeat-, and CARD containing domain*
 NO : *Nitric oxide (Oxyde nitrique)*
 NOD : *Nucleotide-binding oligomerization*
 Nramp : *Natural resistance-associated macrophage protein*

O

Oas : *Oligoadenylate synthetase*
 Oas12 : *2'-5' oligoadenylate synthetase like 2*
 OMS : *Organisation Mondiale de la Santé*
 OR : *Odds Ratio*

P

p.i. : *Post-inoculation*
 PA : *Pathway*
 PAMP : *Pathogen-associated molecular pattern*
 Pas : *Pasteur*
 Pb : *Paire de bases*
 PCR : *Polymerase chain reaction*
 Pglyrp : *Peptidoglycan recognition protein*
 pgm : *Pigmentation*
 PGN : *Peptidoglycane*
 Pklr : *Liver and red blood cell-specific pyruvate kinase*
 PN : *Polynucléaire*
 Prl : *Plague resistance locus*
 PRR : *Pattern-recognition receptor*
 Ptg2 : *Prostaglandin-endoperoxide synthase 2*

Q

qPCR : *Quantitative Polymerase Chain Reaction*
 qRT-PCR : *Quantitative Retro Transcription Polymerase Chain Reaction*
 QTG : *Quantitative Trait Gene*
 QTL : *Quantitative Trait Locus*

R

Rag2 : *Recombination activating gene 2*
 RC : *Recombinant congenic strain*
 RI : *Recombinant Inbred Strain*
 RIP1 : *Receptor-interacting protein 1*
 RIPK2 : *Receptor-interacting serine-threonine kinase 2*
 RLR : *Retinoic-acid-inducible gene 1 (RIG)-I-like receptors*
 RNI : *Reactive Nitrogen Intermediates*
 ROI : *Reactive Oxygen Intermediates*
 ROS : *Reactive oxygen species*
 Rpl7 : *Ribosomal protein L7*
 RSV : *Respiratory syncytial virus*

S

Saa : *Sérum amyloïde A*
 SEB : *Spretus Espagne Barcelone*
 SEG : *Spretus Espagne Grenade*
 Serpin : *Serine peptidase inhibitor*
 SFM : *Spretus France Montpellier*
 siRNA : *Short interference RNA*
 Slc11a1 : *Solute carrier family 11 member 1*
 Sle : *Systemic lupus erythmatosus*
 Sles : *Systemic lupus erythmatosus susceptibility*
 SMZ : *Spretus Maroc Azemnour*
 SNP : *Single nucleotide polymorphism*
 Socs : *Suppressor of cytokine signaling*
 SOD : *Superoxyde dismutase*
 SPIA : *Signaling Pathway Impact Analysis*
 Ss : *Single stranded*
 Sst : *Susceptibility to tuberculosis*
 STAT : *Signal transducer and activator of transcription*
 STF : *Spretus Tunisie Fondouk*

T

TAB2 : *TAK1 Binding protein 2*
 TAK1 : *Transforming growth factor- β -Activated Kinase 1*
 TANK : *TRAF family member-associated NF- κ B activator*
 TBK1 : *TANK-binding kinase-1*
 Tbs : *Tuberculosis severity*
 TGF- β : *Transforming growth factor*
 Th1/2 : *Cellules T helper de type 1/2*
 TIR : *Toll/IL-1 Receptor domain*
 TIRAP : *TIR domain containing Adaptor Protein*
 TIRAP/Mal : *TIR domain-containing Adaptor Protein / MyD88-Adaptor Like*
 TLR : *Toll-like receptor*
 Tmevp : *Theiler's murine encephalomyelitis virus persistence*
 TNF : *Tumor necrosis factor*
 Tnfaip : *TNF alpha-induced protein*
 Tnfrsf : *Tumor necrosis factor receptor super Family*
 Tnfsf : *Tumor necrosis factor super family*
 TRADD : *TNK receptor type I-associated death domain protein*
 TRAF : *TNF-receptor associated factor*
 TRAM : *TRIF-related adaptor molecule*
 TRIF : *TIR-domain-containing adapter-inducing interferon- β*
 TRIM : *Tripartite motif*
 Trl : *Tuberculosis resistance locus*
 TRP53 : *Transformation-related protein 53*

TYK : *Tyrosine kinase*

U

Ufc : Unité formant colonies

V

VCAM-1 : *Vascular Cell Adhesion Molecule 1*

VIH : Virus de l'Immunodéficience Humaine

W

WES : *Whole Exome Sequencing*

WT : *Wild Type*

Y

Y. pestis : *Yersinia pestis*

Ymt : *Yersinia Murine Toxin*

Yop : *Yersinia outer protein*

Ysc : *Yop secretion*

Résumé en français et mots-clés

Recherche de facteurs génétiques contrôlant la résistance de lignées de souris consanguines à une infection expérimentale par *Yersinia pestis*, l'agent de la peste.

Yersinia pestis, l'agent de la peste, est une bactérie à Gram-négatif classée comme agent pathogène ré-émergent et potentielle arme de bioterrorisme. De plus, l'apparition d'une souche multi-résistance de cette bactérie souligne la nécessité de mieux comprendre comment cette bactérie hyper-virulente interagit avec son hôte.

Afin d'identifier des facteurs génétiques de vulnérabilité à la peste, notre laboratoire travaille sur la réponse de souris résistantes versus sensibles à *Y. pestis*. Notre stratégie pour identifier les facteurs génétiques impliqués dans la résistance/sensibilité à la peste combine une approche de cartographie de QTL (*Quantitative Trait Locus*) et d'analyse d'expression génique.

Nous avons précédemment décrit la lignée SEG/Pas, issue de *Mus spretus*, comme la première résistante à une souche virulente de *Y. pestis*, alors que la plupart des lignées murines de laboratoire, telle que la lignée C57BL/6J, sont extrêmement sensibles à la bactérie. Des croisements entre SEG/Pas et C57BL/6J nous ont permis d'identifier trois QTL impliqués dans la résistance à *Y. pestis*, localisés sur les chromosomes 3, 4 et 6.

Deux des QTL (situés sur les chromosomes 4 et 6) ont pu être confirmés par l'analyse de lignées congéniques. Plus de 40 % des femelles bi-congéniques hétérozygotes pour ces deux QTL ont survécu à l'infection, alors que tous les témoins C57BL/6J ont succombé.

La dissection de ces deux QTL par l'analyse de lignées sous-congéniques, nous a permis d'affiner l'architecture génétique de la résistance à la peste chez SEG/Pas. Nous avons conclu qu'un minimum de quatre facteurs génétiques, au sein de ces deux QTL, sont nécessaires pour augmenter la résistance à *Y. pestis* chez la Souris. Cependant, la production de plusieurs lignées congéniques portant le QTL situé sur le chromosome 3, dont une lignée triple congénique, ne nous a pas permis de confirmer l'existence de ce QTL.

En parallèle de l'analyse génétique, nous avons déterminé que les macrophages de SEG/Pas et de C57BL/6J présentaient des caractéristiques différentes après exposition à *Y. pestis*. Une analyse différentielle du profil transcriptionnel des macrophages de ces deux lignées a été réalisée à l'aide de puces à ADN. Nos résultats montrent une forte activation de la production cytokinique dans les macrophages de SEG/Pas en réponse à la bactérie, activation qui n'est pas observée dans la lignée C57BL/6J. Ces résultats suggèrent que les souris SEG/Pas sont capables de mettre en place une réponse immune innée plus forte ou peut-être plus précoce que C57BL/6J.

Nous avons ensuite étudié par qRT-PCR l'expression en cinétique de 44 gènes dans des macrophages de SEG/Pas, C57BL/6J et des bi-congéniques portant les QTL sur les chromosomes 4 et 6. Cette étude nous a permis de confirmer que les souris SEG/Pas sont capables de mettre en place une forte réponse inflammatoire lors de l'infection. Cependant, aucune différence significative n'a été observée entre la lignée bicongénique et la lignée parentale C57BL/6J. D'autres expériences seront nécessaires afin de mieux comprendre les mécanismes biologiques impliqués dans la résistance intermédiaire de cette lignée.

La dissection génétique associée à l'analyse de l'expression génique de ces lignées résistante et sensible améliore notre compréhension de la réponse de l'hôte à *Y. pestis*.

Mots -clés : Peste ; Souris ; résistance ; *Yersinia pestis* ; lignée congénique ; analyse de transcriptome.

Résumé en anglais et mots-clés

Abstract and key words

Identification of genetic factors involved in the resistance of inbred strains of mice to an experimental infection with *Yersinia pestis*, the plague agent.

Yersinia pestis, the agent of plague, is a deadly gram-negative bacterium classified as a re-emerging pathogen and class A biological weapon. The appearance of a multi-resistant strain highlights the need to better understand how this pathogen kills its host.

To identify genetic factors of host susceptibility to plague, our laboratory is investigating the response of resistant versus susceptible mice to *Y. pestis*. Our strategy to decipher genetic determinants involved in resistance to plague combines Quantitative Trait Loci (QTL) mapping with gene expression analysis.

We previously described the *Mus spretus*-derived SEG/Pas strain as the first to resist fully virulent *Y. pestis*, while most inbred strains, such as C57BL/6, are highly susceptible. Crosses between these two strains identified three QTLs (located on chromosome 3, 4 and 6) contributing to resistance.

Two of the QTLs (on chromosome 4 and 6) were confirmed through creation of congenic mice. Up to 40% of the congenic mice heterozygous at these two QTLs, on a C57BL/6J background, survived the infection while all C57BL/6J mice died.

Further dissection of these two QTLs, through the use of subcongenic strains, enabled us to refine the genetic architecture of resistance to plague in SEG/Pas mice. We concluded that a minimum of four genetic factors, within these two QTLs, are required to increased resistance to *Y. pestis* in mice. Despite production of numerous congenic strains, including triple congenic mice, we were not able to confirm the existence of the third QTL identified on chromosome 3.

In parallel to genetic studies, we determined that SEG/Pas and C57BL/6J macrophages exhibit distinct characteristics upon *in vitro* exposure to *Y. pestis*. The underlying molecular differences were investigated by using microarrays. Our results show strong activation of cytokines in SEG/Pas macrophages in response to *Y. pestis*, which is not found in C57BL/6J macrophages. These results suggest that SEG/Pas mice are able to better activate innate immune response to *Y. pestis* than C57BL/6J mice.

We further studied the expression of 44 genes in a kinetic study on macrophages *in vitro* of SEG/Pas, C57BL/6J and bicongenic mice (carrying QTLs on chromosome 4 and 6). This study confirmed that SEG/Pas mice are able to build a stronger inflammatory response at early time of infection. Nevertheless no significant differences were observed in the bicongenic strain compared to C57BL/6J. Further studies will be required to understand the mechanisms involved in the intermediate resistance of this strain.

This combination of genetic dissection and gene expression analysis of resistant and susceptible mouse strains will enhance our ability to better understand the host response to plague.

Key words: Plague; mouse; resistance; *Yersinia pestis*; congenic strain; transcriptomic analysis.

Introduction

Lutter contre une maladie infectieuse requiert de connaître l'agent infectieux en cause mais aussi la réponse physio-pathologique de l'hôte confronté au pathogène. Lors d'une épidémie, tous les individus exposés à un même agent infectieux ne développent pas un processus pathologique avec la même gravité. Le développement et le décours d'un processus infectieux dépendent des caractéristiques de l'agent infectieux lui-même, de facteurs environnementaux, et de différences génétiques entre les individus d'une même espèce.

La recherche de facteurs génétiques de l'hôte capables de modifier sa vulnérabilité à un processus infectieux est intéressante à plusieurs titres. Tout d'abord, elle repose sur des différences observées naturellement qui ont une pertinence biologique. Deuxièmement, l'identification de gènes capables de conférer un certain niveau de résistance oriente vers des voies métaboliques, de signalisation ou de régulation qui sont des mécanismes clés pour la compréhension de la physiopathologie de l'infection, et sont autant de cibles thérapeutiques éventuelles. Enfin, les mécanismes identifiés sont souvent partagés entre espèces, ce qui justifie et valide le recours à des modèles animaux soigneusement choisis pour étudier les maladies infectieuses.

Les facteurs génétiques impliqués dans la vulnérabilité aux maladies infectieuses sont difficiles à étudier directement chez l'Homme car l'exploration des patients atteints d'infection est limitée par la difficulté d'évaluer précisément le niveau de vulnérabilité de chaque individu. On ne connaît jamais la dose infectieuse reçue, rarement la date et le site d'infection et de nombreux facteurs tels que l'état physiologique ou la mise en œuvre de traitements à visée thérapeutique, varient entre individus, ce qui perturbe l'analyse du niveau de résistance conféré par le génotype de l'hôte.

L'étude de modèles animaux, dans lesquels tous ces facteurs peuvent être contrôlés et standardisés, se présente alors comme une alternative intéressante, en particulier pour des infections auxquelles l'espèce choisie est naturellement sensible.

L'existence de différents « outils » génétiques chez la Souris, espèce pour laquelle nous disposons d'un grand nombre de lignées consanguines mais aussi de lignées consomiques, recombinantes consanguines ou encore congéniques, permet d'établir des modèles qui présentent des mécanismes physio-pathologiques spécifiques, permettant de décortiquer les régions génomiques identifiées comme participant à la vulnérabilité à une maladie infectieuse. Les lignées congéniques, qui ont fait l'objet principal de mon projet de thèse, rendent possible la répétition des mesures expérimentales contrôlées, augmentant la puissance des études.

Par ailleurs, les techniques d'analyse globale de la structure du génome (comme le génotypage à très haute densité ou le séquençage à haut débit) ou fonctionnelle (comme l'analyse du transcriptome à l'aide de puces à ADN ou par séquençage à haut débit) offrent des possibilités d'exploration inédites qui peuvent contribuer à mieux comprendre la réponse de l'hôte lors d'une infection.

Depuis près de sept ans, notre laboratoire travaille sur l'identification de facteurs génétiques impliqués dans la vulnérabilité à la peste, partant de l'observation originale que la lignée consanguine de souris SEG/Pas est résistante à la peste, à l'inverse des lignées de souris classiques de laboratoires. La souris est un hôte naturel de *Y. pestis*, ce qui en fait un modèle tout à fait pertinent pour étudier les mécanismes de résistance et rechercher d'éventuelles correspondances avec ce qui est observé dans les populations humaines.

Nous menons, conjointement avec le laboratoire des *Yersinia*, l'étude de la lignée résistante SEG/Pas, et d'une lignée sensible, C57BL/6J qui est une lignée de souris de laboratoire de référence. Nous recherchons les facteurs génétiques impliqués dans la résistance des souris SEG/Pas par l'analyse de QTL et l'analyse différentielle du profil transcriptionnel de ces deux lignées. L'Unité des *Yersinia* travaille parallèlement sur les différences physio-pathologiques entre les deux lignées durant l'infection.

Ces deux approches complémentaires ont permis d'identifier plusieurs régions génomiques qui sont impliquées dans la résistance à la peste et différents mécanismes biologiques qui pourraient être à l'origine de la différence de sensibilité observée (Blanchet *et al.*, 2011; Demeure *et al.*, 2012).

Mon travail de thèse a consisté à confirmer les QTL identifiés dans la résistance à *Y. pestis* dans la lignée SEG/Pas sur les bases statistiques de l'analyse d'un backcross entre SEG/Pas (SEG) et C57BL/6J (B6) et à commencer leur dissection. Parallèlement à cette étude génétique, j'ai mené une étude transcriptomique comparative entre ces deux lignées sur une population cellulaire *ex vivo*. Le but de ce projet est d'identifier *in fine* les gènes qui participent directement à la résistance à *Y. pestis*.

Les mécanismes immunitaires innés, qui ont de fortes chances d'être à la base de la résistance observée chez SEG, sont très conservés chez les vertébrés. Il est possible que certains des résultats que nous obtiendrons chez la Souris soient pertinents chez l'humain.

Il est possible que ces gènes soient également impliqués dans la résistance à d'autres infections bactériennes. Nous espérons que cette étude puisse, à long terme, donner des pistes pour concevoir de nouvelles approches thérapeutiques et préventives non seulement contre *Y. pestis* mais aussi contre d'autres bactéries très virulentes.

ETAT DES CONNAISSANCES

« Le matin du 16 avril, le docteur Bernard Rieux sortit de son cabinet et buta sur un rat mort, au milieu du palier. Sur le moment, il écarta la bête sans y prendre garde et descendit l'escalier. Mais, arrivé dans la rue, la pensée lui vint que ce rat n'était pas à sa place et il retourna sur ses pas pour avertir le concierge. »

La peste, Albert Camus

Partie 1. Comment un hôte répond-t-il à la rencontre avec un micro-organisme pathogène ?

Notre organisme mammifère est constamment exposé à des micro-organismes tels que bactéries, virus, champignons ou protozoaires. Les interactions entre ces derniers et l'hôte mammifère présentent une grande variété. Dans certains cas, ces micro-organismes sont indispensables à la physiologie de l'hôte (par exemple pour la digestion chez les herbivores polygastriques). Dans d'autres cas, les interactions entre les deux organismes ressemblent à une "paix armée" dans laquelle le mammifère met en place des mesures (sécrétions, cellules "sentinelles" et effectrices, molécules anti-microbiennes, etc.) qui permettent de contrôler l'invasion éventuelle du micro-organisme (on parle de résistance) ou de tolérer sa présence à long terme, en particulier de limiter les conséquences néfastes de la réponse inflammatoire (on parle de tolérance ou de résilience), (Raberg *et al.*, 2007; Read *et al.*, 2008; Schneider *et al.*, 2008; Warren *et al.*, 2010). Dans d'autres cas enfin, le micro-organisme possède un pouvoir pathogène, c'est-à-dire une capacité intrinsèque à dérégler très profondément l'homéostasie de l'hôte au point de le rendre "malade", voire d'en entraîner la mort. Dans cette dernière situation, on parle de micro-organisme "pathogène".

Le système immunitaire des mammifères est un ensemble fonctionnel très complexe qui associe des tissus, des types cellulaires variés et de nombreuses molécules de signalisation (telles que les cytokines ou les interférons) et/ou effectrices (tels que les anticorps). L'un des rôles du système immunitaire est de lutter contre l'invasion de l'hôte par des micro-organismes pathogènes, pour en préserver l'intégrité et l'homéostasie. La réponse immune déclenchée par la rencontre d'un pathogène comporte deux composantes qui interviennent successivement : la réponse immune innée qui est mise en œuvre immédiatement et de façon non spécifique, et la réponse adaptative qui est spécifiquement dirigée contre le pathogène identifié mais n'intervient généralement qu'environ une semaine après l'infection. C'est cette deuxième réponse qui permet d'acquérir une "mémoire" des pathogènes déjà rencontrés pour réagir plus rapidement et efficacement lors d'une rencontre ultérieure.

Dans les infections bactériennes aiguës telles que la peste, en raison de la multiplication et de la dissémination très rapide du micro-organisme, la réponse innée joue un rôle majeur et souvent critique pour la survie de l'hôte. C'est sur cette dernière que nous allons concentrer la suite de cette partie.

La réponse immune innée fait intervenir un volet cellulaire et un volet humoral. La composante cellulaire comporte des cellules épithéliales (qui empêchent les pathogènes d'atteindre le milieu intérieur), des cellules phagocytaires (neutrophiles, monocytes, macrophages), des cellules *Natural Killer* (NK) et des granulocytes. Ces cellules phagocytent les micro-organismes et produisent des cytokines et chimiokines, qui vont déclencher une réponse inflammatoire, recruter d'autres

cellules de l'immunité sur le site de l'infection et déclencher la mise en place de la réponse immunitaire adaptative. Le volet humoral de l'immunité innée comprend les anticorps naturels, le système du complément et les peptides antimicrobiens (Figure 1), (Dziarski *et al.*, 2010; Kapetanovic *et al.*, 2007).

Cette première partie n'est pas une revue exhaustive des mécanismes qui se mettent en place lors de la rencontre de l'hôte et d'un pathogène, mais a pour but de présenter les acteurs cellulaires et moléculaires pour interpréter les données de transcriptomiques qui seront exposées dans la partie « résultats ».

1. La réponse au niveau de l'organisme

1.1. La reconnaissance du pathogène

Pour identifier les pathogènes, les cellules immunitaires détectent des motifs microbiens, appelés PAMP pour « *Pathogen-Associated Molecular Patterns* », grâce à des récepteurs appelés PRR pour « *Pattern-Recognition Receptors* », tels que les récepteurs membranaires *Toll-like* (TLR) et les récepteurs cytoplasmiques *NOD-like* (NLR). Chaque récepteur est spécifique de certains motifs moléculaires. L'ensemble des PRR permet de reconnaître une grande diversité de molécules provenant de micro-organismes variés.

1.1.1. « Das ist ja toll ! »¹, les *Toll-like* récepteurs

Les TLR sont des senseurs microbiens essentiels dans le déclenchement de la réponse immune innée. Ce sont des glycoprotéines membranaires caractérisées par un domaine extracellulaire comportant des motifs riches en leucine (LRR) et par une région cytoplasmique nommée « domaine TIR » (pour *Toll/IL-1 Receptor domain*) (Figure 2) (Bowie *et al.*, 2000). La partie intracellulaire est conservée et essentielle pour activer différentes voies de signalisation faisant intervenir les protéines adaptatrices MyD88 (*Myeloid Differentiation Factor*), TIRAP/Mal (*TIR domain-containing Adaptor Protein / MyD88-Adaptor Like*) ou TRIF (*TIR domain-containing adaptor inducing IFN β*). La partie extracellulaire est spécialisée dans la reconnaissance du ligand microbien (Takeda *et al.*, 2005).

¹ « C'est super ! » phrase lancée par Christiane Nüsslein-Volhard lors de la découverte en 1985 de l'implication d'un gène dans la polarisation dorso-ventrale lors de l'embryogénèse chez la drosophile, qui fût alors appelé « Toll » Hansson GK, Edfeldt K (2005). Toll to be paid at the gateway to the vessel wall. *Arteriosclerosis, thrombosis, and vascular biology* 25, 1085-1087..

Dix TLR ont été décrits chez l'Homme et 12 chez la Souris (Kawai *et al.*, 2010). Les TLR sont exprimés par de nombreuses cellules immunitaires, mais aussi dans des cellules non immunitaires mais participant à la protection de l'hôte, notamment les fibroblastes ou les cellules épithéliales (Akira *et al.*, 2006).

Les TLR sont localisés soit à la surface des cellules au niveau de la membrane plasmique (TLR 1, 2, 4, 5 et 6) soit dans des membranes d'endosomes (TLR 3, 7, 8 et 9) (Figure 2) (Kawai *et al.*, 2010). Cette disposition permet de détecter les motifs microbiens extracellulaires et les motifs des microbes phagocytés ou provenant de virus dans des cellules infectées (Tableau 1).

1.1.1.1. *TLR4 et la reconnaissance du LPS*

TLR4 permet la détection du lipopolysaccharide (LPS), un composant majeur de la membrane externe des bactéries à gram-négatif. Contrairement à la plupart des TLR, la transduction du signal de TLR4 requiert la présence des co-récepteurs MD-2 et CD14 (Akira *et al.*, 2006; Takeda *et al.*, 2005). D'autres protéines peuvent faciliter la fixation du LPS, comme CD36 ou la LPB (*LPS-Binding Protein*), (Akira *et al.*, 2006; Takeda *et al.*, 2005).

TLR4 est également capable de reconnaître certains motifs viraux ou fongiques, ainsi que des ligands endogènes tels que HMGB1, un médiateur produit par les leucocytes lors d'endotoxémie (Akira *et al.*, 2006; Kawai *et al.*, 2010; Takeda *et al.*, 2005; Yu *et al.*, 2006).

1.1.1.2. *TLR1, 2 et 6*

TLR2 est impliqué dans la reconnaissance de nombreux PAMP parmi lesquels les lipoprotéines bactériennes, comme le peptidoglycane (PGN) et l'acide lipoteichoïque (LTA) des bactéries à gram-positif, le lipoarabinomannane des mycobactéries, le zymosan des champignons, le tGPI-mucine du parasite *Trypanosoma cruzi* ou encore l'hémagglutinine de certains virus. TLR2 peut aussi reconnaître une lipoprotéine synthétique, le Pam3CysSK₄ (Akira *et al.*, 2006; Kawai *et al.*, 2010; Takeda *et al.*, 2005).

TLR2 forme généralement des hétérodimères avec TLR1 ou TLR6. Le complexe TLR2/TLR1 reconnaît les lipopeptides tri-acétylés des bactéries à gram-négatif et des mycoplasmes alors que le complexe TLR2/TLR6 reconnaît des lipopeptides di-acétylés de bactéries à Gram-positif et des mycoplasmes (Tableau 1), (Akira *et al.*, 2006; Kawai *et al.*, 2010; Takeda *et al.*, 2005). Tout comme TLR4, l'activation optimale via TLR2 nécessite l'action de co-récepteurs. Par exemple, CD14 et CD36 sont requis pour que le LTA des bactéries à Gram-positif soit reconnu par TLR2 et induise la production de TNF (Nilsen *et al.*, 2008).

1.1.1.3. Les autres TLR

Le ligand de TLR5 est la flagelline, protéine majeure des flagelles bactériens (Hayashi *et al.*, 2001).

Les TLR 3, 7, 8 et 9 sont exprimés dans des vésicules intracellulaires comme le réticulum endoplasmique, les endosomes, les lysosomes et les endolysosomes. Ils reconnaissent principalement des acides nucléiques microbiens présents dans ces organelles (Figure 2). TLR3 reconnaît l'ARN double brin produit pendant la réplication virale (Alexopoulou *et al.*, 2001). Le TLR7 humain et le TLR8 de souris peuvent détecter l'ARN simple brin retrouvé chez certains virus (Heil *et al.*). TLR9 reconnaît les motifs CpG fréquents chez les bactéries, les parasites et les virus mais plus rares au sein des cellules eucaryotes (Tableau 1) (Bauer *et al.*, 2001). TLR11, identifié chez la Souris, semble jouer un rôle dans la reconnaissance des bactéries uropathogènes, telles qu'*Escherichia Coli* et d'autres motifs microbiens tels que la profiline du parasite *Toxoplasma gondii* (Yarovinsky *et al.*, 2005; Zhang *et al.*, 2004).

1.1.2. Les *NOD-like* Récepteurs, des senseurs cytoplasmiques

Les TLR ne permettent pas la détection des pathogènes dans le cytoplasme des cellules. C'est le rôle des NLR (*NOD-like Receptor*).

La famille des NLR comporte 23 membres chez l'humain et 34 chez la Souris, capables de reconnaître les PAMP, mais aussi des particules non PAMP et des signaux de stress cellulaire. Une cascade de signalisations en aval des NLR induit une réponse pro-inflammatoire par les cellules stimulées (Chen *et al.*, 2009; Kapetanovic *et al.*, 2007). Ces molécules possèdent toutes un domaine LRR qui leur permet de reconnaître les motifs microbiens, un domaine NOD (*Nucleotide-binding oligomerization*) et un domaine de signalisation CARD (*Caspase-activating and recruitment domain*) ou pyrin qui permet la signalisation cellulaire (Figure 2), (Kapetanovic *et al.*, 2007).

Les deux NLR les plus connus sont NOD1 et NOD2. NOD1 reconnaît des motifs de peptidoglycane (PGN) essentiellement présents chez les bactéries à Gram-négatif et l'acide- γ -D-glutamyl-meso-diaminopimelic (Tableau 2), (Chen *et al.*, 2009; Kanneganti *et al.*, 2007). NOD2 détecte le muramyl-dipeptide (MDP), que l'on retrouve chez les bactéries à Gram-négatif et positif (Kanneganti *et al.*, 2007). NOD1 et NOD2 jouent un rôle important dans la reconnaissance des bactéries invasives et permettent de déclencher des mécanismes de défense cellulaire. Il a été montré que NOD1 jouait un rôle important dans la détection de *Shigella flexneri*, *Pseudomonas aeruginosa*, *Helicobacter pylori* ou encore *Listeria monocytogenes* (Girardin *et al.*, 2001; Opitz *et al.*, 2006; Travassos *et al.*, 2005; Viala *et al.*, 2004). Il est à noter qu'*Helicobacter pylori*, bactérie extracellulaire, injecte du PGN dans la cellule cible par l'intermédiaire de son système de sécrétion de

type IV, qui est alors reconnu par NOD1 (Viala *et al.*, 2004). D'autres pathogènes tels que *Salmonella Typhimurium*, *Streptococcus pneumoniae* et *Mycobacterium tuberculosis* sont reconnues par NOD2 (Ferwerda *et al.*, 2005; Hisamatsu *et al.*, 2003; Opitz *et al.*, 2004).

Il existe d'autres récepteurs intra-cytoplasmiques appartenant aux NLR tels que les récepteurs NLRP1 (*Nacht domain-, Leucine-rich repeat-, and PYD-Containing Protein 1*, anciennement NALP1), NLRP3 (anciennement NALP3 ou Cryopyrin) et NLRC4 (*Nacht domain-, Leucine-rich repeat-, and CARD containing domain 4* anciennement IPAF pour *ICE Protease-Activating Factor*), (Figure 2) regroupés au sein de complexes appelés inflammasomes, complexes protéiques impliqués dans l'activation de la caspase 1 qui va cliver la pro-IL-1 β en sa forme extracellulaire active IL-1 β , ainsi que la pro-IL-18 en IL-18 active, (Kanneganti *et al.*, 2007; Mogensen, 2009). En plus de réguler la réponse pro-inflammatoire, la caspase 1 active également des signaux pro-apoptotiques (Fernandes-Alnemri *et al.*, 2007).

1.1.3. Les senseurs alternatifs

La famille des RLR (pour *Retinoic-acid-inducible gene I (RIG)-I-like receptors*) est constituée de 3 membres qui sont connus pour détecter des ARN viraux (Figure 2), (Mogensen, 2009; Yoneyama *et al.*, 2009).

1.2. La réaction de l'organisme face au pathogène

L'entrée du pathogène dans l'organisme déclenche généralement une réaction inflammatoire. Au premier siècle après J. C., Cornelius Celsus caractérisa l'inflammation par quatre phénomènes : rougeur, gonflement, chaleur et douleur. Les facteurs chimiques produits durant l'inflammation (histamine, bradykinine, sérotonine, leucotriènes et prostaglandines) augmentent la sensation de douleur, induisent localement la vasodilatation et la perméabilité vasculaire ainsi que le recrutement de phagocytes, en particulier les neutrophiles. Les neutrophiles produisent des facteurs solubles contribuant à la mobilisation d'autres populations de leucocytes. Le TNF (*Tumor Necrosis Factor*) provoque entre autre l'activation du système de coagulation (Kapetanovic *et al.*, 2007). Cette cytokine induit également l'activation de facteurs de transcription tels que NF-kB et AP-1 qui induisent l'expression de centaines de gènes pro-inflammatoire et pro-apoptotiques. La production de cytokines stimule également les cellules phagocytaires et leur activité microbicide (Chu, 2012).

Ces cytokines participent au recrutement des cellules immunitaires sur le site de l'inflammation. Elles induisent l'augmentation de l'expression de molécules de surface par les cellules endothéliales vasculaires provoquant l'adhérence des cellules de l'immunité. L'expression de

sélectines (L-, P- et E-sélectines) permet aux cellules endothéliales de « capturer » les leucocytes. L'expression d'autres molécules d'adhérence cellulaire telles qu'ICAM-1 (pour *Intracellular Adhesion Molecule 1*) ou VCAM-1 (pour *Vascular Cell Adhesion Molecule 1*) stimule l'infiltration de neutrophiles et de monocytes sur le site de l'inflammation (Weber *et al.*, 2007).

Les cytokines ont également un effet à distance du site inflammatoire. Certaines d'entre elles stimulent la prolifération des cellules hématopoïétiques de la moelle osseuse (comme le *Granulocyte-Macrophage Colony Stimulating Factor* (GM-CSF), qui agit sur la différenciation des cellules hématopoïétiques souches en granulocytes et macrophages) ou la production par le foie de protéines de la phase aigüe (comme la *C-reactive protein*, Sérum amyloïde A ou certains facteurs du complément) (Kapetanovic *et al.*, 2007). Elles agissent aussi sur la partie antérieure de l'hypothalamus, qui régule la température de l'organisme provoquant la fièvre (Klir *et al.*, 1993).

Le système du complément est également activé. La libération d'anaphylatoxines telles que C5a et C3a présente des propriétés chimiotactiques. Les protéines du complément adhèrent aux pathogènes pour induire leur élimination directe ou par opsonisation.

2. La réponse à l'échelle cellulaire

2.1. Les cellules de l'immunité innée

Les cellules épithéliales peuvent détecter une infection et initier une réponse immunitaire. Agissent ensuite les cellules *Natural Killer* (NK), les granulocytes (mastocytes, éosinophiles et basophiles) et les phagocytes (macrophages, neutrophiles et cellules dendritiques). Nous présenterons ici, rapidement, les phagocytes et les cellules NK, dont l'interaction avec *Y. pestis* a été décrite dans la littérature.

2.1.1. Les cellules phagocytaires mononuclées : monocytes et macrophages

Les monocytes sanguins sont les premières cellules à réagir en cas de bactériémie. La migration des monocytes dans les différents tissus induit leur transformation en phagocytes professionnels, comme les macrophages alvéolaires, péritonéaux, les cellules de Kupffer dans le foie, les cellules dendritiques dans le derme, de Langerhans dans l'épiderme, la microglie dans le système nerveux ou encore les ostéoblastes dans les os (Gordon *et al.*, 2005). Ces cellules essentiellement phagocytaires capturent efficacement des particules ou des macromolécules pour les dégrader. Dans certaines conditions liées à l'environnement cellulaire, elles peuvent aussi exercer des fonctions de

présentation d'antigènes. Certaines, comme les cellules dendritiques sont capables d'activer des lymphocytes T naïfs. Elles reconnaissent les antigènes en périphérie (dans le derme par exemple) et les transportent vers les organes lymphoïdes secondaires pour initier la réponse immunitaire adaptative (Rosenberger *et al.*, 2003). Les monocytes qui colonisent le péritoine se différencient en macrophages péritonéaux. Ces macrophages sont particulièrement importants dans un compartiment qui côtoie le tube digestif et sa flore bactérienne. Ils ont un rôle crucial dans l'initiation de l'inflammation et dans l'activation de l'afflux des neutrophiles en cas d'infection de la cavité péritonéale. En raison de leur facilité d'extraction, ils sont aussi très utilisés dans les études *ex vivo*.

Les macrophages résidents sont des sentinelles qui servent de première ligne d'alerte. Lorsqu'ils interagissent avec certains récepteurs des macrophages tels que l'intégrine CD11b/CD18 (CR3, pour *Complement Receptor 3*), les pathogènes sont phagocytés et souvent éliminés. Cependant, certains d'entre eux ont mis au point des stratégies qui leur permettent de moduler l'activité microbicide des macrophages pour y échapper (pour revue voir (Rosenberger *et al.*, 2003)).

Les propriétés fonctionnelles des macrophages varient selon leur distribution tissulaire et leur état d'activation. Les macrophages activés par les cytokines de type 1 et les produits microbiens (voie classique d'activation, ou M1) sont inflammatoires (ils sécrètent des cytokines telles que TNF- α , IL-6, IL-12, IL-1 β) et microbicides en produisant des espèces oxygénés telles que l'oxyde nitrique (NO) (Rosenberger *et al.*, 2003). Les macrophages activés par certains stimuli tels que l'IL-4, l'IL-13, le TGF- β (*Transforming Growth Factor β*) sont appelés macrophages M2. Ils modulent la réponse immune en produisant des cytokines anti-inflammatoires telles que l'IL-10 et sont faiblement microbicides. En réalité, la notion de polarisation M1/M2 des macrophages recouvre un *continuum* d'états fonctionnels. Les macrophages peuvent ainsi être, dans un premier temps, inflammatoires avant de participer à la résolution de l'inflammation, ce qui suggère que les changements fonctionnels des macrophages résultent de modifications de leur microenvironnement (Stout *et al.*, 2005; Stout *et al.*, 2004).

2.1.2. Les neutrophiles

De manière physiologique, les neutrophiles transitent dans la circulation sanguine. En réponse à la détection de produits bactériens ou sous l'influence de molécules attractives (leucotriènes, CXCL8, fragments du complément, etc.), les neutrophiles expriment à leur surface des molécules d'adhésion qui leur permettent de traverser la paroi endothéliale et de migrer vers le site de l'inflammation. L'action anti-infectieuse des neutrophiles repose également sur la phagocytose, la production de formes réactives de l'oxygène (ROS pour *Reactive Oxygen Species*) et la libération de protéines enzymatiques (myéloperoxydase, élastase, lysosyme, gélatinase) et de protéines antimicrobiennes (lactoferrine, défensines) (Dumas *et al.*, 2009). Un autre mécanisme bactéricide mis

en œuvre par les neutrophiles est l'expulsion de filaments appelés NET (*Neutrophil Extracellular Traps*) composés de chromatine et de protéases qui emprisonnent et tuent des microbes hors de la cellule (Brinkmann *et al.*, 2007; Fuchs *et al.*, 2007).

Les neutrophiles sont la source de nombreux médiateurs de l'inflammation. Des cytokines pro-inflammatoires telles qu'IL-1 β et TNF- α amplifient plusieurs fonctions du neutrophile comme l'adhérence aux cellules endothéliales ou la production de ROS par une action autocrine, et permettent également le recrutement et l'activation des macrophages et des cellules dendritiques. Des chimiokines, telles que CXCL8, déclenchent la dégranulation des neutrophiles, ainsi que la production de ROS; d'autres chimiokines telles que CCL3 et 4 (respectivement MIP-1 α et β), CCL5 (RANTES), CCL20 (MIP-3 α) facilitent le recrutement local et l'activation des cellules dendritiques et des lymphocytes T (Chakravarti *et al.*, 2007). Des facteurs de croissance tels que G-CSF (*Granulocyte Colony-Stimulating Factor*), GM-CSF (*Granulocyte/Macrophage Colony-Stimulating Factor*) agissent sur la différenciation des granulocytes et des macrophages à partir de cellules souches hématopoïétiques.

2.1.3. Les cellules *Natural Killer* (NK)

Les cellules NK sont des lymphocytes de l'immunité innée. Elles sont très largement distribuées dans l'organisme, patrouillent notamment en grand nombre dans la circulation sanguine et migrent rapidement vers les zones inflammatoires. Elles exercent deux fonctions effectrices principales : la cytotoxicité (la destruction de cellules cibles par la libération de granules cytotoxiques contenant des protéines telles que la perforine ou les granzymes) et la production de cytokines (Rouzaire *et al.*, 2012). Les cellules NK sont par ailleurs la principale source précoce d'IFN γ capable de potentialiser les effets de l'immunité innée et de l'immunité adaptative (Souza-Fonseca-Guimaraes *et al.*, 2012).

La réponse immunitaire innée fournit ainsi une protection immédiate de l'organisme. Après quelques jours, la réponse adaptative se caractérise par une expansion clonale des lymphocytes B et T possédant des récepteurs spécifiques de l'agent infectieux. Ce type de réponse très spécifique, permet dans certains cas, à l'organisme de garder une mémoire contre le pathogène en question.

2.2. Les conséquences, à l'échelle cellulaire, de la reconnaissance du pathogène

2.2.1. La transduction du signal en aval des TLR

Le mécanisme de transduction du signal par les TLR débute par la dimérisation ou l'oligomérisation des ou du récepteur(s) provoquée par la fixation du ligand. Un nouvel arrangement

des domaines TIR forme alors une structure qui peut fixer les différentes protéines adaptatrices. Quatre protéines adaptatrices interagissent avec les TLR : MyD88 (*Myeloid Differentiation primary response gene 88*), TIRAP (*TIR domain containing Adaptor Protein*), TRIF (*TIR-domain-containing adapter-inducing interferon- β*) et TRAM (*TRIF-Related Adaptor Molecule*), (Figure 3), (Kawai *et al.*, 2007).

Cela initie une cascade de signalisation, aboutissant à l'activation de gènes spécifiques de l'inflammation, tels que des gènes codant des cytokines (telles que TNF, IL6, IL12, etc.), des molécules de co-stimulation (tels que CD80, CD86) ou pour des gènes de la famille des interférons de type I (IFN α et β), (Kawai *et al.*, 2010).

Chaque TLR mène à une réponse biologique spécifique qui peut varier suivant la cellule. Deux voies de signalisation, la voie MyD88-dépendante et la voie TRIF-dépendante, peuvent être différemment activées suivant le TLR stimulé, dépendant des protéines adaptatrices recrutées (Figure 3). TLR4 est le seul TLR à pouvoir activer ces 2 voies de signalisation.

2.2.1.1. La voie MyD88-dépendante

MyD88 est recruté par le TLR grâce à un changement de conformation des domaines TIR du dimère de TLR lors de la fixation du ligand (Figure 3). MyD88 recrute et active une kinase, IRAK4 (*IL-1 Receptor-Associated Kinase 4*), elle-même responsable du recrutement de deux autres kinases : IRAK1 et IRAK2 (Min-Oo *et al.*, 2008; Mogensen, 2009). Lorsque les molécules IRAK sont phosphorylées, elles se détachent de Myd88 et activent TRAF6 (*TNF Receptor-Associated Factor 6*), une E3 ubiquitine ligase. TRAF6 forme alors un complexe avec TAB2 (*TAK1 Binding protein 2*) et TAB3 pour activer TAK1 (*Transforming growth factor- β -Activated Kinase 1*) (Mogensen, 2009). TAK1 peut alors être engagée dans deux voies de signalisation : les voies IKK (IkB Kinase) et MAPK.

TAK1 peut activer le complexe IKK comprenant IKK α , β et γ qui forment un complexe et phosphorylent la protéine I κ B α (*Inhibitor of κ light chain gene enhancer in B cells, alpha*) (Takaesu *et al.*, 2003). Cette phosphorylation permet la translocation de NF- κ B (*Nuclear factor - κ B*) dans le noyau. NF- κ B induit alors l'expression de gènes codant des cytokines pro-inflammatoires et de gènes impliqués dans d'autres fonctions telles que l'inhibition de l'apoptose (Kawai *et al.*, 2010; Mogensen, 2009).

Parallèlement, TAK1 peut aussi activer des *Mitogen-activated Protein kinases* (MAPK), dont Erk1, Erk2, p38 ou JNK. Les facteurs de transcription CREB (*cyclic AMP-responsive element-binding protein*) et AP1 (*Activator protein 1*) sont activés par phosphorylation via la voie des MAP kinases, conduisant entre autres à la synthèse de cytokines pro-inflammatoires (Kawai *et al.*, 2010; Mogensen, 2009).

TRAF6 peut aussi activer le facteur de transcription IRF5 (*Interferon Regulatory Factor 5*) permettant la synthèse de cytokines pro-inflammatoires telles qu'IL-6, IL-12 et TNF α (Takaoka *et al.*, 2005).

2.2.1.2. La voie TRIF-dépendante

Le couple de protéines adaptatrices TRAM/TRIF est responsable de la voie de signalisation MyD88-indépendante (Oshiumi *et al.*, 2003). TRIF possède deux sites d'activation. Cette molécule peut induire la formation d'un complexe avec TRAF3, TBK1 (*TANK-Binding Kinase-1*) et IKK conduisant à la phosphorylation du facteur de transcription IRF3, qui induit l'expression de gènes interféron de type I (IFN α et β) et des gènes dépendants des interférons (Kawai *et al.*, 2007; Taniguchi *et al.*, 2002). TRIF peut aussi, mais dans une moindre mesure, s'associer à TRAF6 et activer TAK1 comme dans la voie MyD88-dépendante, entraînant alors une activation du facteur NF- κ B (Sato *et al.*, 2003).

TRIF recrute aussi RIP1 (*Receptor-Interacting Protein 1*) qui change de conformation et va pouvoir être fixé par TRADD (*TNK Receptor type I-Associated Death Domain protein*), ce qui permet l'activation de TAK1, induisant l'activation du facteur NF- κ B et de la voie des MAP kinases (Figure 3), (Cusson-Hermance *et al.*, 2005).

2.2.2. Les voies de signalisation des récepteurs cytoplasmiques

NOD1 et NOD2 interagissent avec la molécule RIP2, cette association permet d'activer la kinase TAK1, de manière similaire à la voie des TLR, la voie NF- κ B et la production de médiateurs de l'inflammation. TAK1 peut aussi activer MAP kinases Erk, JNK ou p38 (Figure 4), (Kanneganti *et al.*, 2007; Strober *et al.*, 2006).

Les autres membres de la famille NLR, tels que NLRP1, NLRP3 ou NLRC4, appartenant à des complexes protéiques appelés « inflammasome », peuvent activer la caspase-1 *via* la molécule adaptatrice ASC (*Apoptosis-associated Speck-like protein containing a CARD*), (Figure 4). La caspase-1 permet de cliver la pro-IL1 β en IL1 β et stimule sa sécrétion (Kanneganti *et al.*, 2007). Rathinam *et al.* ont montré que les bactéries à Gram-négatif, en activant la voie TRIF, provoquaient l'activation de la caspase 11 *via* la voie de signalisation des IFN de type 1 (figure 4bis). La caspase 11, en synergie avec l'inflammasome incluant NLRP3, régule l'activité de la caspase 1 et conduit également à une forme rapide de mort cellulaire appelée « pyroptose » (Newton *et al.*, 2012; Rathinam *et al.*, 2012). La voie TRIF permet de faire le lien entre les voies des TLR et NLRP3 activées en réponse aux bactéries à Gram-négatif, par l'intermédiaire de la voie IFN (Rathinam *et al.*, 2012).

2.2.3. Les médiateurs de l'inflammation induits par l'activation de NF- κ B

La voie de signalisation NF- κ B joue un rôle important dans les réponses immune et inflammatoire. Les protéines NF- κ B sont des activateurs transcriptionnels présents dans la cellule sous forme de dimères. Ces protéines sont capables d'induire l'expression de plus de 150 gènes, dont des gènes codant des cytokines pro-inflammatoires telles qu'IL-1, IL-6 et TNF- α , ainsi que des chimiokines, dont IL-8 (KC chez la Souris) et RANTES (ou CCL5), (Mogensen, 2009). L'expression de molécules d'adhésion cellulaire, telles que ICAM-1 et E-sélectine est également activée par NF- κ B. De même, la classe des immunorécepteurs, d'importance toute particulière lors de la réponse cellulaire à un pathogène, est activée par NF- κ B. Cette classe comprend les récepteurs aux cytokines et chimiokines, les immunoglobulines, les TLR, les molécules du complexe majeur d'histocompatibilité (CMH) et les molécules co-stimulatrices (Mogensen, 2009). Ensemble, les gènes induits par NF- κ B participent à l'activation et au recrutement des leucocytes sur le site de l'inflammation, à l'activation de la phagocytose, à la lyse médiée par le complément ou les cellules NK (*Natural Killer*) et à la présentation de l'antigène (Mogensen, 2009). De plus, NF- κ B est capable d'induire des gènes anti-apoptotiques (Mogensen, 2009).

2.2.4. L'activation des *Interferon Regulatory Factors* (IRFs)

L'activation du facteur IRF3 par la voie TRIF-dépendante provoque son homo-dimérisation ou sa dimérisation avec IRF7 et va par la suite recruter des co-activateurs transcriptionnels tels que CBP (*CREB-binding protein*) et p300 qui, par leurs fonctions acétyltransférases, vont faciliter sa fixation à l'ADN et induire la transcription de gènes, dont les IFN de type I (Figure 5), (Taniguchi *et al.*, 2002). L'interféron produit va se lier aux deux sous-unités du récepteur IFNAR (IFNAR1 et IFNAR2), menant à l'activation des tyrosines kinases (TYK) : Janus kinase 1 (JAK1) et TYK2, les transducteurs du signal STAT1 (pour *Signal Transducer and Activator of Transcription 1*) et STAT2, qui vont former un hétéro-dimère (Kawai *et al.*, 2007; Taniguchi *et al.*, 2002). Le complexe STAT1/STAT2 s'associe à IRF9 pour former le complexe ISGF3 (*IFN-Stimulated Gene Factor 3*). Ce complexe se fixe sur les promoteurs qui contiennent des éléments de réponse à l'IFN (ou *IFN-Stimulated Response Elements*) ce qui induit alors l'expression des gènes tels que *Isg15*, *Cxcl10*, *Irf7* et *Oas* (Taniguchi *et al.*, 2002). IRF7 peut, seul ou combiné à IRF3, se lier aux promoteurs des gènes interféron de type I, IFN- α et IFN- β (Taniguchi *et al.*, 2002). Ainsi, cette boucle de régulation positive va entraîner la production massive d'IFN- α et IFN- β (Taniguchi *et al.*, 2002). Cette sécrétion abondante d'interféron, de concert avec les autres cytokines et chimiokines produites par les autres voies de signalisation activées en parallèle, déclenche de nombreux effets immunitaires qui vont permettre aux cellules de se défendre contre les pathogènes et recruter les cellules de l'immunité

adaptative, mais qui peuvent également participer à la pathogénèse d'une maladie infectieuse (Patel *et al.*, 2012; Trinchieri, 2010).

2.2.5. La phagocytose

La phagocytose est le mécanisme qui permet la digestion des pathogènes. Le phagocyte peut se lier aux bactéries via les récepteurs du complément, des *scavengers receptors*, les récepteurs de la partie Fc des immunoglobulines (FcR) et des intégrines (Stuart *et al.*, 2005). Des prolongements de membrane, les pseudopodes, s'étendent autour de la particule à phagocyter et l'internalisent dans une vésicule membranaire : le phagolysosome. Le phagolysosome subit divers processus de maturation : acidification, production de composés oxygénés toxiques (comme l'anion superoxyde O_2^- , le peroxyde d'hydrogène H_2O_2 et l'oxyde nitrique NO), production de peptides antimicrobiens (comme les défensines) et d'enzymes (comme le lysozyme ou l'hydrolase), production de protéines compétitrices des facteurs nécessaires à la croissance bactérienne (telle que la lactoferrine qui se lie au fer) (Forman *et al.*, 2002). La phagocytose permet d'induire rapidement l'enzyme iNOS (*inducible Nitric Oxide Synthetase* aussi appelée NOS2) qui va produire du NO. Les espèces réactives à l'oxygène (*Reactive Oxygen Species*, ROS) sont produites par un complexe enzymatique, la NADPH-oxidase, située sur la membrane de la vacuole du phagolysosome (Forman *et al.*, 2002). La superoxyde dismutase (SOD), une enzyme à l'intérieur du phagosome, convertit l'anion superoxyde en peroxyde d'hydrogène (Forman *et al.*, 2002). De plus, des peroxydases peuvent transformer le peroxyde d'hydrogène formé en d'autres radicaux. Cet ensemble de processus est appelé « *oxydative burst* » (Forman *et al.*, 2002).

Vazquez-Torres *et al.* ont montré que les macrophages péritonéaux étaient capables d'éliminer *Salmonella typhimurium* par l'action bactéricide de la NADPH oxydase et de l'enzyme iNOS (Mastroeni *et al.*, 2000; Vazquez-Torres *et al.*, 2000). Gantt *et al.* ont observé que l'anion superoxyde O_2^- et le NO participaient à l'élimination de *Leishmania chagasi* dans les macrophages murins et humains (Gantt *et al.*, 2001). De surcroît, l'acidification du phagosome est un processus essentiel pour stopper la croissance microbienne et permettre la génération des ROS (Lukacs *et al.*, 1991).

Le réseau d'actine a une importance toute particulière dans la phagocytose. L'actine est fortement polymérisée sous le point d'adhérence et permet au macrophage d'entourer la particule à phagocyter (Champion *et al.*, 2006; Stuart *et al.*, 2005). La fixation du micro-organisme à phagocyter couplée à des événements de signalisation, induit un réarrangement des filaments d'actine par l'activation de Rho GTPases, dont Rho, Rac et Cdc42 (Rosenberger *et al.*, 2003; Stuart *et al.*, 2005). Ces protéines sont des cibles idéales pour les facteurs de virulence bactériens car l'inactivation de ces voies de signalisation bloque la phagocytose et permet la survie de la bactérie dans la cellule (Rosenberger *et al.*, 2003).

Certains micro-organismes ont évolué et élaboré des stratégies afin d'échapper aux conséquences de la phagocytose. C'est le cas de *Yersinia pestis* qui, à 37°C, synthétise un polymère de protéines F1, lui permettant de s'encapsuler et de résister à la phagocytose (Du *et al.*, 2002). De plus, cette bactérie empêche son internalisation en désorganisant le réseau d'actine qui permet l'expansion de pseudopodes (Cornelis, 2002 A). D'autres micro-organismes tels que *Listeria monocytogenes* ou *Shigella flexneri* sont capables de détruire la membrane du phagosome (Goetz *et al.*, 2001). Il semblerait que *Mycobacterium tuberculosis* et *Toxoplasma gondii*, par exemple, neutralisent l'acidification du phagosome essentielle à sa maturation et à la fusion avec le lysosome (Gordon *et al.*, 1980).

Les mécanismes de défense majeurs mis en place à l'échelle cellulaire suite à la reconnaissance d'un pathogène sont représentés dans un macrophage, figure 6.

En conclusion, la mise en place et le déroulement de l'immunité font intervenir de nombreux types cellulaires et acteurs moléculaires qui interagissent de manière complexe. D'un individu à l'autre, la réponse à un pathogène peut être très différente, sous l'influence de facteurs environnementaux, physiologiques mais aussi génétiques. Des variations dans la structure de certaines molécules de signalisation ou effectrices, mais aussi dans le niveau d'expression et les mécanismes de régulations peuvent être à l'origine de grandes différences de vulnérabilité. La recherche des facteurs génétiques expliquant des différences de vulnérabilité entre individus peut donc permettre de mieux comprendre les mécanismes essentiels dans la réponse d'un hôte à un agent infectieux. C'est ce que nous verrons dans la partie suivante.

Partie 2. Puissance et limites des études génétiques de la sensibilité aux maladies infectieuses chez l'Homme

Une caractéristique fréquente des maladies infectieuses, que ce soit chez l'Homme ou chez l'animal, est qu'une partie seulement des individus exposés à un pathogène développe des signes cliniques de la maladie (Casanova *et al.*, 2004; Hill, 2012). La présence d'un micro-organisme est nécessaire mais pas nécessairement suffisante au développement d'une maladie infectieuse chez un individu. Les caractéristiques génétiques de l'hôte et son état physiologique jouent un rôle important dans le développement d'une infection. Louis Pasteur avait souligné l'importance des facteurs non-microbiens, dont des facteurs héréditaires, dans la sensibilité aux maladies infectieuses (Casanova *et al.*, 2004). Chez l'humain, l'influence de facteurs génétiques de l'hôte sur sa sensibilité à des maladies infectieuses a été démontrée, dès les années 1950, par des études sur des jumeaux² (par exemple pour la tuberculose, la lèpre et l'hépatite B (Hill, 1996, 2001)) ainsi que sur des enfants adoptés³. Une étude a ainsi montré que des enfants adoptés avaient 5,8 fois plus de chances de mourir d'une maladie infectieuse si leurs parents biologiques étaient décédés avant l'âge de 50 ans de cette cause. En revanche, la mort d'un parent adoptant des suites d'une maladie infectieuse n'avait pas d'effet significatif sur les risques de mortalité de l'enfant adopté (Sorensen *et al.*, 1988).

L'exemple le plus connu de facteur génétique humain associé à la résistance à une maladie infectieuse est celui des hémoglobinopathies et de la malaria. L'anémie falciforme est causée par un variant du gène bêta-globine (HbS) qui induit une polymérisation de la globine et une anomalie des globules rouges. Les individus homozygotes (HbS/S) souffrent d'une grave anémie mais les hétérozygotes présentent un risque 10 fois moindre que les individus HbA/A de contracter le paludisme (Hill, 2001; Hill *et al.*, 1991; Weatherall *et al.*, 1988). Cet avantage sélectif des hétérozygotes explique la persistance de l'allèle délétère HbS dans les populations exposées à *Plasmodium falciparum*.

Cet exemple de gène ayant un effet fort est toutefois exceptionnel, car bien souvent, les facteurs génétiques identifiés ne contrôlent qu'une fraction de la sensibilité. Dès lors, le contrôle génétique d'une maladie infectieuse revêt un caractère polygénique. La recherche de ces facteurs génétiques chez l'Homme repose principalement sur des approches dites de « gène candidat » et des

² Les études sur les jumeaux sont basées sur le fait que si une maladie a une composante génétique, des jumeaux monozygotes, qui partagent la même séquence d'ADN, présenteront un taux de concordance supérieur à une maladie que des jumeaux dizygotes. Le taux de concordance est le rapport entre le nombre de paires de jumeaux où les deux jumeaux sont atteints et le nombre total de paires. Plus la différence entre les taux de concordance des jumeaux monozygotes et dizygotes est élevée, plus la part génétique dans le déterminisme la maladie est importante.

³ Les études portant sur des enfants adoptés, c'est-à-dire séparés de leurs parents biologiques et intégrés dans un environnement différent de celui de leurs parents biologiques, permettent de dissocier composante génétique et composante environnementale familiale. On compare la fréquence de la maladie chez les enfants adoptés dont les parents biologiques ou les parents adoptants présentent la maladie.

approches qui criblent l'ensemble du génome sans à priori : les études de liaison (*Genome Wide Scan*) et d'associations pangénomiques (*Genome Wide Association Studies* ou GWAS), (Figure 7). Les études portant sur les variants rares à effet fort et les variants structuraux, qui sont aujourd'hui en expansion, seront aussi mentionnées (Chapman *et al.*, 2012; Hill, 2006).

1. L'approche par gène candidat

L'approche « gène candidat » repose sur l'hypothèse qu'un gène est impliqué dans un phénotype et vise à confirmer cette hypothèse en recherchant des variants de ce gène associés à des différences de phénotype. Elle est fondée sur le génotypage de polymorphismes dans des gènes dont la fonction biologique est plausiblement impliquée dans le phénotype étudié. Une fois le gène choisi et les polymorphismes identifiés, deux approches sont possibles. Une étude d'association cas-témoins consiste à comparer la fréquence de chaque variant entre un groupe de patients et un groupe témoin. Un test de déséquilibre de transmission (*Transmission Disequilibrium Test*) étudie la fréquence de transmission des variants dans des familles où ségrège la maladie. Si il y a une différence de fréquence significative ou que le variant est significativement plus transmis au sein d'une famille, on peut alors conclure que le variant participe probablement à la maladie ou bien que le variant est en déséquilibre de liaison avec le facteur génétique en cause.

Un exemple de cette approche « gène candidat » est l'étude du gène TLR1 dans la sensibilité à la lèpre. Le dimère des protéines TLR1/TLR2 participe à la reconnaissance de *Mycobacterium leprae*, l'agent de la lèpre (Krutzik *et al.*, 2003). Les patients qui portent l'allèle G du SNP T1805G ont une protéine TLR1 non fonctionnelle (Hawn *et al.*, 2007). Hawn *et al.* ont montré un défaut d'activation de la voie NF- κ B et de la production d'IL6 dans un modèle cellulaire transfecté avec la séquence de TLR1 issue de patients homozygotes 1805G, alors que la séquence de TLR1 issue de patients 1805T présentait une activation normale (Hawn *et al.*, 2007). Le polymorphisme T1805G ayant une fréquence allélique très différente entre les différentes régions du globe, Hawn *et al.*, ont émis l'hypothèse que ce variant pourrait être à l'origine d'une variation de sensibilité aux maladies infectieuses dont la lèpre (Revue dans (Misch *et al.*, 2010)). Barreiro *et al.*, ont montré que le SNP 1805G est soumis à une pression de sélection positive, suggérant que cet allèle, qui confère une faible réponse à *M. leprae* et peut-être à d'autres pathogènes, pourrait offrir un avantage sélectif dans certaines populations (Barreiro *et al.*, 2009). A partir de ces observations, trois études ont recherché une association entre les allèles T et G du SNP 1805 de *TLR1* et la sensibilité à la lèpre. Johnson *et al.*, ont mis en évidence, sur un échantillon de 57 patients et 90 contrôles en Turquie, une association entre 1805G et une protection accrue contre la lèpre (OR = 0,48; $P = 0,004$), (Johnson *et al.*, 2007). L'étude

de Misch *et al.*, portant sur 933 cas de lèpre au Népal, n'a pas confirmé cette association mais a néanmoins montré que l'apparition d'une forte réponse Th1 lors de la maladie était associée à l'allèle 1805G (OR = 0.51; P = 0.01) (Misch *et al.*, 2008). Une autre étude menée sur 842 patients et 543 contrôles au Bangladesh, n'a pas mis en évidence d'association entre 1805G et une sensibilité à la lèpre (Schuring *et al.*, 2009). Néanmoins, cette étude a montré qu'un autre SNP du gène TLR1 (A743G) est associé, chez les patients homozygotes 743G à une sensibilité accrue à la maladie (OR = 1.34; P = 0.02) (Schuring *et al.*, 2009). Ces études illustrent que le déterminisme génétique de la vulnérabilité à la lèpre, comme certainement à beaucoup d'autres maladies infectieuses, présente une hétérogénéité entre les populations (Chapman *et al.*, 2012; Misch *et al.*, 2010; Stein, 2011).

Les études basées sur cette approche de gène candidat ont permis d'établir un lien de causalité entre différents gènes et une variation de sensibilité à des pathogènes. Des gènes candidats peuvent également être suggérés par les résultats de travaux sur des modèles animaux comme nous le verrons ultérieurement.

2. Les analyses criblant l'ensemble du génome

2.1. Les analyses de liaison ou Genome Wide Scan

Les analyses de liaison chez l'Homme sont basées sur la co-ségrégation, au cours des générations, des allèles de marqueurs polymorphes avec le locus responsable de la maladie, dans une famille ou une collection de familles (Feingold, 1998).

Un exemple d'étude de liaison est le travail de Marquet *et al.* sur la vulnérabilité à une infection par *Schistosoma mansoni*. Cette étude menée sur 142 sujets brésiliens appartenant à 11 familles, a permis d'identifier une seule région génomique (5q31-q33) présentant un locus fort contrôlant environ 66% de la variance observée. L'étude de cette région a mené les auteurs à retenir 8 candidats basés sur leur fonction biologique dans la réponse inflammatoire : *IL3*, *IL4*, *IL5*, *IL9*, *IL13*, *Irf1*, *Csf2* et *Csf1r* (Marquet *et al.*, 1996).

Un autre exemple est l'identification du locus 10p13 dans le contrôle de la sensibilité à la lèpre, dans une étude portant sur 245 frères et sœurs du Sud de l'Inde dont au moins deux étaient atteints par famille (Siddiqui *et al.*, 2001). Une autre étude sur la lèpre menée dans le sud du Viêt-Nam, sur 205 frères et sœurs, a identifié une liaison entre la région chromosomique 6q25 et une sensibilité accrue à la maladie (Mira *et al.*, 2003). Une liaison avec deux marqueurs a ensuite été confirmée par une étude d'association sur des familles indépendantes de la même région du Viêt-Nam (Mira *et al.*, 2003).

Ces exemples illustrent la diversité des régions chromosomiques mises en évidence en fonction des populations étudiées. De plus, les variations des caractéristiques phénotypiques des

patients inclus dans ces différentes études (forme de la maladie, âge d'apparition, sévérité des symptômes, par exemple) influent sur la puissance de détection des régions en cause et ne permettent pas toujours de comparer les études et leurs résultats.

Les études de *Genome Wide Scan* participent à la compréhension de l'architecture génétique des caractères complexes chez l'humain : certains sont gouvernés par un ou plusieurs variants à effet fort alors que d'autres (ce sont de loin les plus fréquents) sont gouvernés par une multitude de variants génétiques ayant tous un effet mineur. Toutefois, l'analyse de liaison a l'inconvénient de manquer de puissance pour détecter des gènes ayant un effet mineur. En effet, elle ne permet généralement pas de localiser un gène avec précision. Les régions identifiées couvrent souvent plusieurs dizaines de mégabases en raison de déséquilibres de liaison. L'identification de locus de vulnérabilité est en général suivie d'une cartographie fine à l'aide des marqueurs plus rapprochés dans les locus identifiés. Une fois la taille du locus réduite, les gènes de l'intervalle sont étudiés par une approche de candidat positionnel. La confirmation par une analyse d'association sur un grand nombre d'individus est ensuite nécessaire pour valider l'implication du locus dans la maladie.

2.2. Les études d'association pan-génomiques ou *Genome Wide Association Study* (GWAS) pour l'étude des maladies communes

Avec le génotypage à haut débit sont apparues, dans les années 2000, les études d'associations pangénomiques qui sont un complément aux analyses de liaison pour identifier les gènes de prédisposition aux maladies complexes chez l'Homme. La réalisation d'une GWAS nécessite de constituer un groupe de personnes affectées par la maladie étudiée (cas) et un groupe témoin indemne (témoins) aussi proche que possible du groupe atteint tant au niveau de l'ascendance que du mode de vie, afin de réduire l'hétérogénéité génétique et environnementale au maximum. Chaque individu est génotypé pour un grand nombre de SNP, en général entre 500 000 et un million. Une analyse des allèles significativement sur- ou sous-représentés dans la population atteinte par rapport aux sujets témoins permet de cibler des régions du génome dans lesquelles se trouvent des gènes (ou des séquences régulatrices) qui influencent la vulnérabilité des sujets à la maladie étudiée. L'hypothèse « *common disease, common variant* », sur laquelle reposent l'approche GWAS, suppose que les maladies multifactorielles fréquentes reposeraient sur des variants génétiques fréquents (c'est-à-dire, dont la fréquence de l'allèle minoritaire est supérieure à 5%), (Chapman *et al.*, 2012; McCarthy *et al.*, 2008; WTCCC, 2007). Une étude de liaison est nécessaire pour valider l'existence d'un caractère héréditaire, puis une étude approfondie de la région chromosomique autour des allèles identifiés ainsi que des tests fonctionnels sont requis pour déterminer les facteurs génétiques en cause et leurs effets biologiques (Jordan, 2009).

Un exemple de GWAS est l'étude menée par Zhang *et al.*, en 2009 en Chine qui a permis de découvrir des gènes impliqués dans la sensibilité à la lèpre et de mieux comprendre les mécanismes de réponse de l'hôte face à cette maladie (Zhang *et al.*, 2009). Cette étude, menée sur 706 cas et 1225 témoins génotypés pour environ 500 000 SNP, rapporte une association significative ($P < 10^{-10}$) avec des SNP dans la région *HLA-DR-DQ* et dans les gènes *RIPK2* (*Receptor-interacting serine/threonine-Protein Kinase 2*), *TNFSF15* (*Tumor Necrosis Factor SuperFamily member 15*), *LACC1* anciennement *C13orf31*, (*LACCase (multicopper oxidoreductase) domain-containing 1*), *CCDC122* (*Coiled-Coil Domain-Containing 122*) et *NOD2* (*Nucleotide-binding Oligomerization Domain-containing 2*) (Zhang *et al.*, 2009). Une association plus faible a également été détectée avec le gène *LRRK2* (*Leucine-Rich Repeat Kinase2*).

Des associations entre la région *HLA-DR-DQ*, les gènes *LACC1* et *CCDC122* ont ensuite été retrouvées dans une autre étude menée sur des patients indiens (Wong *et al.*, 2010). L'association trouvée avec *NOD2* n'a pas été retrouvée dans la cohorte indienne, alors que ce gène avait été présenté par Zhang *et al.* comme un senseur bactérien primordial pour la reconnaissance du pathogène (Wong *et al.*, 2010; Zhang *et al.*, 2009). Une autre étude sur la lèpre menée au Viêt-Nam avait identifié des associations dans les gènes *PARK2-PARK* et *LTA*, associations qui n'ont pas été répliquées dans d'autres populations (Mira *et al.*, 2004; Mira *et al.*, 2003).

Ces études d'association portant sur la lèpre montrent la difficulté de répliquer des résultats. Selon Wong *et al.* cette difficulté est due à l'hétérogénéité génétique entre les populations et ces auteurs recommandent que les études fonctionnelles soient réalisées sur les mêmes populations que celles où l'association génétique a été caractérisée (Wong *et al.*, 2010). En effet, les études d'association se heurtent à un problème de variabilité entre les individus tant au niveau génétique qu'environnemental, qui crée un bruit de fond, diminuant la puissance de détection. Le seul recours pour gagner en puissance de détection semble être actuellement d'augmenter la taille des cohortes analysées pour inclure, non plus des milliers, mais plusieurs dizaines de milliers d'individus (Jordan, 2009). Chapman et Hill soulignent également, dans leur revue de 2012, que des tailles d'échantillons différentes, l'utilisation de différentes puces de génotypage ainsi que des différences en termes de fréquence allélique et de structure haplotypique des populations étudiées sont autant de raisons qui font varier la puissance de détection des facteurs de sensibilité entre les études et contribuent à affaiblir le taux de réplification des études d'association chez l'Homme (Chapman *et al.*, 2012).

Les GWAS ont permis de se forger une idée de l'architecture génétique des caractères complexes de l'Homme. Le risque de développer une maladie commune est presque toujours contrôlé par un large panel de gènes aux effets très faibles, parfois par quelques gènes ayant des effets moyens, et rarement par des gènes à effets forts (Manolio *et al.*, 2009; Panthier *et al.*, 2012; WTCCC, 2007). La méthode GWAS ne permet pas d'identifier des variants génétiques rares. Cependant, certains variants

généétiques rares jouent un rôle non négligeable dans la sensibilité aux maladies infectieuses (Casanova *et al.*, 2007; Chapman *et al.*, 2012).

3. L'étude des variants rares à effet fort

Les GWAS ont permis de découvrir de nombreux polymorphismes impliqués dans les maladies infectieuses, mais ces variants n'expliquent généralement qu'une faible part de l'héritabilité observée (Chapman *et al.*, 2012; Manolio *et al.*, 2009; McCarthy *et al.*, 2008; WTCCC, 2007). La question est alors de savoir comment et où trouver la part non expliquée d'héritabilité des caractères complexes (Manolio *et al.*, 2009). Grâce à l'avènement du séquençage à haut débit, la recherche de variants rares (variants individuels ou familiaux dont l'effet est fort, s'apparentant parfois à des caractères mendéliens) est aujourd'hui une des approches dominantes pour détecter les variants génomiques à l'origine de la différence de sensibilité aux maladies (Bamshad *et al.*, 2011; Chapman *et al.*, 2012; Cirulli *et al.*, 2010).

La méthode de capture et séquençage de l'ensemble des exons (*Whole Exome Sequencing* ou WES) est adaptée à l'identification de ces variants rares (dont la fréquence allélique minoritaire est inférieure à 1%) que ne pouvaient pas identifier les GWAS (Bamshad *et al.*, 2011; Manolio *et al.*, 2009). Cette méthode permet d'avoir accès à l'ensemble de la séquence codante d'un individu. Les gènes codants des protéines constituent environ 1% du génome humain, mais contiennent 85% des mutations à effet fort reliées à des maladies (Choi *et al.*, 2009). Ainsi le séquençage de l'ensemble des exons a le potentiel de contribuer à la compréhension de l'effet des variants communs et rares dans les maladies humaines (Choi *et al.*, 2009). Cette méthode doit permettre d'identifier tous les variants entre des sujets atteints et sains afin de les comparer.

L'intérêt croissant pour le rôle des mutations rares à effet fort est très pertinent pour l'étude des maladies infectieuses. L'identification de mutations privées (propres à quelques individus) dans des gènes impliqués dans le système immunitaire associées à une sensibilité accrue aux infections a amené à définir le concept d'« immunodéficience à déterminisme génétique », transmise sur un mode mendélien (Casanova *et al.*, 2007). Un exemple est la sensibilité extrême aux mycobactéries, dont le déterminisme est clairement mendélien (appelée *Mendelian Susceptibility to Mycobacterial Disease* ou MSMD) (Casanova *et al.*, 2007; Chapman *et al.*, 2012; Fortin *et al.*, 2007). Différentes mutations ont été identifiées chez des enfants atteints d'infections sévères dues à des mycobactéries. Des mutations détectées dans les gènes *IFNGR1*, *IFNGR2* (récepteurs 1 et 2 à l'interféron gamma) et *STAT1* provoquent une réponse altérée à l'interféron gamma (IFN γ) et des mutations dans les gènes

IL12B, *IL12RB1* et *NEMO* provoquent une production altérée d'IFN γ . Des défauts de réponse des macrophages et de la différenciation des cellules dendritiques ont également été reliés à des mutations dans les gènes *CYBB* et *IRF8*, respectivement chez des patients atteints de MSMD (Bustamante *et al.*, 2011; Hambleton *et al.*, 2011; van de Vosse *et al.*, 2009).

Une des difficultés inhérentes à la méthode WES pour l'identification d'allèles rares est de pouvoir distinguer les allèles morbides parmi l'ensemble des polymorphismes et parmi les erreurs de séquençage (Bamshad *et al.*, 2011; Cooper *et al.*, 2011). Le séquençage du génome d'individus d'une même famille dans laquelle des individus sont atteints permet de comparer la transmission des allèles et celle de la maladie. En prenant des individus apparentés les plus distants possibles, qui présentent la même maladie, le nombre de variants communs (et donc potentiellement causaux) est diminué. Cette stratégie, avec celle des « phénotypes extrêmes » (qui consiste à séquencer le génome des individus présentant les phénotypes les plus extrêmes d'une maladie) et la recherche de variants *de novo* dans des trios parents-enfant, sont utilisées dans la recherche de gènes candidats (Chapman *et al.*, 2012; Hill, 2001). Une part de l'hérédité manquante qui n'a pas été détectée par les GWAS pourrait résider dans ces allèles rares à effet fort (Bamshad *et al.*, 2011).

4. La recherche de variants structuraux

Les variants de grandes taille, comme les CNV (*Copy Number Variants*) qui sont des segments stables et transmis d'ADN d'une taille supérieure à 1kb et qui présentent un nombre de copies variables dans une population, pourraient aussi participer au déterminisme génétique des maladies complexes humaines (Almasy, 2012; Eichler *et al.*, 2010; Manolio *et al.*, 2009). La détection des CNV est possible en utilisant des logiciels adaptés à partir de données de génotypage à haut débit ou de séquençage. Il a d'ailleurs été montré que les duplications segmentaires du génome humain présentent une sur-représentation de gènes impliqués dans des fonctions immunitaires (Gonzalez *et al.*, 2005). Les conséquences physiologiques de cet enrichissement ne sont pour l'instant pas connues mais il pourrait jouer un rôle dans la sensibilité aux maladies infectieuses (Gonzalez *et al.*, 2005).

Par exemple, le nombre de copies du gène *CCL3L1* est impliqué dans la sensibilité au VIH (Gonzalez *et al.*, 2005). Le gène *CCL3L1*, qui code la chimiokine MIP-1 α P, est le co-récepteur majeur de CCR5 (ligand du VIH-1). Une augmentation du nombre de copies de ce gène (par référence au nombre moyen dans la population européenne) diminue la sensibilité à une infection par le VIH-1 (Gonzalez *et al.*, 2005). Une variation de dose des gènes impliqués dans l'immunité pourrait participer aux bases génétiques de la réponse de l'hôte aux maladies infectieuses.

Les études de liaison, les GWAS ou les études de mutations rares ont permis d'identifier quelques éléments de l'architecture génétique de la sensibilité aux maladies infectieuses chez l'Homme. Ces études montrent que, pour un même phénotype, les variants en cause peuvent différer entre les populations. Des résultats discordants sont toujours délicats à interpréter et l'absence de réplication incite souvent à douter d'un résultat. En revanche, l'hétérogénéité génétique des populations humaines peut donner accès à une vision globale de l'ensemble des régions du génome qui participent à la vulnérabilité à un pathogène. Lorsque les gènes sont connus, cette diversité peut aider à avoir une idée des différents mécanismes cellulaires avec lesquels un pathogène interfère. Toutefois, les techniques actuelles n'ont permis de déceler qu'une petite partie des facteurs héréditaires et peu de gènes ont été mis en cause formellement dans la sensibilité aux maladies infectieuses.

Par ailleurs, les études chez l'Homme présentent des limites importantes liées à la nature des maladies infectieuses. En effet, de nombreux paramètres peuvent varier d'un individu à l'autre et entre populations : souche microbienne, dose infectante, date de l'infection, état physiologique de l'individu, composantes de l'environnement. Tous ces paramètres ont une grande influence sur le déroulement du processus infectieux. De plus, les patients reçoivent généralement un traitement (lorsqu'il existe) qui ne permet pas d'évaluer précisément leur capacité innée à contrôler le processus infectieux. Enfin, il faut noter que la plus grande partie des études génétiques ont porté sur des maladies chroniques (mycobactéries, parasites, etc.) et rarement sur des maladies aiguës.

On le voit, s'il peut être préférable de rechercher des facteurs génétiques de vulnérabilité directement chez l'Homme, l'étude des maladies infectieuses dans des modèles animaux permet de contrôler strictement un grand nombre de paramètres et d'augmenter la puissance des études. Même si ces modèles ne reflètent pas exactement ce qu'il se passe dans les populations humaines, différents exemples démontrent qu'ils sont un outil essentiel pour mener des expériences afin de caractériser les interactions qui se produisent entre un hôte et un pathogène, expériences qui ne sont pas réalisables chez l'Homme. Cette expérimentation, réglementée et soumise à l'appréciation de comités d'éthique, est cruciale pour identifier les facteurs génétiques qui participent à la réponse d'un hôte à un pathogène.

Partie 3. La pertinence du modèle souris dans l'étude des maladies infectieuses

La compréhension des mécanismes biologiques nécessite d'avoir recours à des modèles expérimentaux *in vitro* et *in vivo* reproductibles. L'unique étude des populations humaines ne suffit pas à elle seule pour analyser et décrire la complexité des mécanismes biologiques sous-jacents aux maladies humaines. Les raisons majeures sont l'hétérogénéité génétique présente dans et entre les populations humaines, la variation des facteurs environnementaux et des contraintes éthiques qui rendent impossibles certaines investigations (Aitman *et al.*, 2011; Casanova *et al.*, 2004; Guenet, 2011). Les modèles animaux (spontanés ou induits) de maladies humaines fournissent des approches complémentaires indispensables. Ces modèles permettent de standardiser le fonds génétique des animaux, les facteurs de l'environnement et, dans le cas des maladies infectieuses, les modalités d'infection. Le contrôle de ces paramètres apporte un surcroît de puissance considérable par rapport aux études sur l'Homme. Bien sûr, chaque espèce est différente et un modèle animal ne représentera jamais exactement le déroulement de l'infection chez l'humain (Casanova *et al.*, 2004; Fortin *et al.*, 2007). Mais l'étude des modèles animaux permet de simplifier la complexité observée dans des populations naturelles humaines quand aucun paramètre n'est fixé (Guenet, 2011). Certains objecteront qu'aucun modèle animal ne saurait refléter la réalité complexe des populations humaines. Or un modèle a pour objectif de réduire la complexité afin d'augmenter la probabilité de pouvoir conclure, en particulier sur la valeur causale d'un paramètre (Guenet, 2011). Il est vrai que souris et humains ont divergé il y a plus de 70 millions d'années et ont été soumis depuis à une pression de sélection microbienne complètement différente (Tuite *et al.*, 2006). Des critiques ont été formulées quant à la pertinence et l'utilité du modèle souris suite à l'observation de mutations de gènes impliqués dans la réponse immunitaire qui affectaient la sensibilité à un groupe de micro-organismes beaucoup plus large chez la Souris que chez l'Homme (Casanova *et al.*, 2007). Par exemple, les mutations dans les voies IL12/IFN γ /STAT1, TIR/IRAK4/NF-KB et TLR3/UNC93B/IFN α affectent la sensibilité à un grand nombre de pathogènes chez la Souris, mais ont un impact réduit chez l'Homme (Fortin *et al.*, 2007). La voie IL12/IL23-INF α est essentielle seulement dans la réponse contre les infections mycobactériennes et par *Salmonella* chez l'Homme, mais n'a pas été montrée comme cruciale pour d'autres pathogènes comme c'est le cas chez la Souris (Casanova *et al.*, 2004; Fieschi *et al.*, 2003; MacLennan *et al.*, 2004). Ceci peut-être dû à une spécialisation accrue des systèmes de reconnaissance des pathogènes chez l'Homme et au développement de mécanismes immuns compensatoires résultant des mélanges de populations humaines au cours de l'évolution (Casanova *et al.*, 2004; Fortin *et al.*, 2007).

L'utilisation d'un modèle a ses limites, ses conditions d'utilisation. Les modèles animaux n'ont pas besoin d'être complètement fidèles pour être utiles, comme le souligne Jean-Louis Guenet dans une récente revue (Guenet, 2011). Il souligne également que les modèles animaux doivent être

considérés comme des outils capables de répondre à des questions précises plutôt que d'être des simples répliques, à plus petite échelle, d'une maladie humaine (Guenet, 2011).

Les modèles animaux, lorsqu'ils sont utilisés dans un contexte biologique pertinent, sont un complément, parfois un préambule, aux études génétiques chez l'Homme. Ils permettent de répondre à des questions précises, donnent des pistes d'études, proposent des gènes candidats à valider chez l'humain. Ils apportent une connaissance générale des voies de signalisation et des mécanismes biologiques qu'il n'est pas toujours possible d'acquérir directement dans les populations humaines où l'hétérogénéité génétique et environnementale créent un « bruit de fond » important (Hunter *et al.*, 2008; Nadeau, 2005).

L'étude des modèles animaux et plus particulièrement de la Souris a permis une avancée majeure dans la connaissance du fonctionnement du système immunitaire. Un exemple historique d'une grande découverte effectuée sur la Souris est le travail de George Snell qui a isolé les déterminants génétiques de la compatibilité tissulaire (aujourd'hui appelé Complexe Majeur d'Histocompatibilité, CMH), molécules qui ont un rôle clé dans la reconnaissance des antigènes et le rejet de greffe. Les études chez la Souris ont aussi fourni des gènes candidats pour les études d'association chez l'Homme et ont permis une avancée significative dans la connaissance des mécanismes de pathogénie et de la réponse de l'hôte dans différents types d'infection (Vidal *et al.*, 2008). La pertinence du modèle Souris pour la découverte de gènes impliqués dans la réponse immunitaire sera illustrée ici par trois exemples : la découverte de *Nramp1* (maintenant appelé *Scl11a1*), *Tlr4* et *Oas1b*.

Le clonage et la caractérisation fonctionnelle de *Nramp1* illustrent une des premières démonstrations de la puissance des études génétiques portant sur la sensibilité aux maladies infectieuses menées sur la Souris. Il a été montré que la sensibilité à différentes mycobactéries, dont *Mycobacterium bovis*, était modulée par le locus *Bcg* localisé sur le chromosome 1. Des locus de sensibilité à *Salmonella typhimurium*, responsable de la fièvre typhoïde, et *Leishmania donovani*, l'agent de la leishmaniose, respectivement dénommés *Ity* et *Lsh* ont aussi été cartographiés sur la même région du chromosome 1 que *Bcg*. Cette observation suggérait que ces trois locus étaient liés et correspondaient au même gène. Par ailleurs, la physiopathologie des trois infections suggérait un rôle clé des macrophages. Une carte physique de la région proximale du chromosome 1 où se situait le locus *Bcg* a été construite, révélant six gènes candidats. Les données d'expression des six gènes ont permis de sélectionner un gène dont l'expression était très forte dans les macrophages. Ce gène a été dénommé *Nramp1* pour *Natural resistance associated macrophage protein 1* (aujourd'hui renommé *Scl11a1*) (Blackwell *et al.*, 1994 B; Malo *et al.*, 1994; Skamene, 1994). Des analyses de séquence révélèrent ensuite la présence d'une substitution d'acides aminés (Gly169Asp) dans la protéine dans les lignées de souris sensibles. La preuve de l'identité entre le gène *Nramp1* et les locus *Ity*, *Bcg* et *Lsh* fut obtenue par la création d'un allèle perte de fonction du gène *Nramp1* qui abrogea la résistance des

lignées de souris aux trois pathogènes *M. bovis*, *S. typhimurium* et *L. donovani*. De plus, une complémentation fonctionnelle par le transfert d'un allèle sauvage *Nramp1*^{Gly169} chez des souris sensibles portant l'allèle muté *Nramp1*^{Asp169} conférait une résistance aux souris sensibles (revu dans (Fortin *et al.*, 2007; Lam-Yuk-Tseung *et al.*, 2003; Tuite *et al.*, 2006; Vidal *et al.*, 2008)).

L'implication de *Nramp1* dans la sensibilité à différents pathogènes a depuis été démontrée dans d'autres espèces animales comme le poulet ou le cheval (Beaumont *et al.*, 2003; Hu *et al.*, 1997). Chez l'Homme, des études ont montré que des polymorphismes du gène *SLC11A1* sont associés ou liés à une sensibilité accrue à la tuberculose dans des populations africaines, asiatiques, autochtones canadiennes et chez des enfants dans le sud de l'Afrique (Greenwood *et al.*, 2000; Li *et al.*, 2006; Malik *et al.*, 2005). Le gène *SLC11A1* est aussi impliqué dans la sensibilité à deux autres espèces de mycobactéries, *Mycobacterium leprae*, l'agent de la lèpre, et *Mycobacterium ulcerans* responsable de l'ulcère de Buruli (Alcais *et al.*, 2000; Stienstra *et al.*, 2001).

Un autre exemple est la découverte de *Tlr4* par le groupe de Bruce Beutler, qui a été faite sur des souris C3H/HeJ résistantes au LPS⁴ mais très sensibles aux bactéries telles que *S. typhimurium* ou *Escherichia coli* (revu dans (Hoebe *et al.*, 2006)). Des recherches sur les causes génétiques de la résistance au LPS des souris C3H/HeJ mirent en évidence un locus unique alors dénommé *Lps*, qui fut cartographié en 1978 par J. Watson *et al.* sur le chromosome 4 de la souris (Watson *et al.*, 1978 A; Watson *et al.*, 1978 B). La cartographie fine de la région a révélé *Tlr4* comme un bon candidat qui a alors été séquencé dans différentes lignées de souris. La lignée C3/HeJ présentait une mutation dans le domaine cytoplasmique très conservé de la protéine TLR4 abolissant l'induction du signal produit par la détection du LPS (Poltorak *et al.*, 1998).

Chez l'Homme, deux polymorphismes de TLR4, Asp299Gly et Thr399Ile, ont été associés à une réponse altérée au LPS inhalé chez des volontaires sains. Ces polymorphismes ont aussi été associés à un risque accru de contracter des infections à bactéries à Gram-négatif, une septicémie, un syndrome de réponse inflammatoire systémique, un paludisme sévère, la brucellose ou encore des infections respiratoires dues au virus respiratoire syncytial (revu dans (Miller *et al.*, 2005)).

Un autre exemple de la contribution des modèles murins dans l'identification des gènes impliqués dans la réponse aux pathogènes est la découverte du rôle du gène *Oas1b* dans la sensibilité au virus *West Nile*⁵. Les souris de laboratoire sont très sensibles à l'infection par ce virus et meurent en moins de 15 jours, contrairement aux lignées dérivées de souris sauvages qui sont complètement résistantes (Mashimo *et al.*, 2002). Des analyses génétiques ont mis en évidence un locus *Wnv*, sur le chromosome 5 (Mashimo *et al.*, 2002). La région chromosomique identifiée a été finement

⁴ Le LPS (lipopolysaccharide) est un composant essentiel de la membrane externe des bactéries à Gram-négatif.

⁵ Le virus West Nile est un Flavivirus (de la même famille que le virus de la fièvre jaune) transmis par un moustique et capable de provoquer une encéphalite parfois mortelle chez l'homme et certains mammifères (comme les rongeurs et les chevaux).

caractérisée. Mashimo *et al.*, ont mis en évidence un cluster de gènes codant des enzymes 2'-5'-oligoadenylate synthetases (2'-5'-OAS) (Mashimo *et al.*, 2002). Ces gènes, qui sont induits par l'interféron et activés par des molécules d'ARN double brin, ont un effet antiviral (Castelli *et al.*, 1998). L'analyse de la séquence génomique des isoformes des gènes 2'-5'-OAS dans plusieurs lignées de souris de laboratoire et de souris sauvages a révélé une corrélation parfaite entre la sensibilité ou la résistance des souris et la présence d'un codon STOP dans l'exon 4 du gène *Oas1b* (isoforme 1B de la 2'-5'-oligoadénylate synthétase) (Mashimo *et al.*, 2002). Ce codon STOP est responsable d'une protéine OAS1B tronquée et inactive qui contribue à la sensibilité des souris de laboratoire au virus *West Nile* (Mashimo *et al.*, 2002). La même année, Perelygin *et al.*, ont également identifié *Oas1b* comme responsable de la variation de sensibilité de diverses lignées de souris aux *flaviviridae* (dont fait partie le virus *West Nile*), (Perelygin *et al.*, 2002). Le rôle de *Oas1b* a ensuite été confirmé chez des souris de laboratoire par l'introduction d'un allèle résistant dans des souris sensibles par des techniques de « knock-in » et de transgénèse (Scherbik *et al.*, 2007; Simon-Chazottes *et al.*, 2011). Ces résultats obtenus chez la souris ont incité des généticiens humains à étudier le rôle éventuel du gène homologue dans les infections humaines. Deux groupes ont identifié des variants du gène OAS1 associés soit à une plus grande réceptivité à l'infection (Lim *et al.*, 2009), soit à un risque accru de développer une forme grave (Yakub *et al.*, 2005).

Si les gènes impliqués dans des maladies humaines et animales similaires ne sont pas toujours les mêmes, les mécanismes moléculaires et cellulaires ont des chances d'être conservés entre les mammifères (Guenet, 2011; Hoebe *et al.*, 2006; Richer *et al.*, 2008). Les études sur les animaux, à défaut d'identifier les gènes responsables des pathologies humaines, donnent des informations cruciales quant aux voies biochimiques et physiologiques en jeu dans la pathogénie d'une maladie (Mohan *et al.*, 1998). Ainsi, l'étude chez la Souris, par exemple, va faciliter l'identification de gènes candidats chez l'Homme, qui pourront être testés sur des cohortes bien plus petites que les milliers d'individus nécessaires lors d'une étude d'association pangénomique.

Les modèles animaux permettent de diminuer la part de la variance d'origine génétique et environnementale dans des modèles d'infection par rapport aux populations humaines. Des données phénotypiques fiables et reproductibles sont obtenues sur des lignées consanguines de souris (génétiquement homogènes) entretenues dans des conditions environnementales standardisées (Hunter *et al.*, 2008; Nadeau, 2005). De plus, les approches de génétique directe (*forward genetics*) et inverse (*reverse genetics*) chez la Souris, alliées à la connaissance de la séquence de 17 lignées et à l'accumulation massive de données de phénotypage, de génotypage et d'expression produites chez la Souris faciliteront probablement la dissection des interactions entre les gènes, et entre les gènes et les facteurs environnementaux, indispensable à la compréhension des maladies et caractères complexes humains.

Les stratégies efficaces pour mettre en évidence l'architecture du déterminisme génétique, et les gènes eux-mêmes, de la sensibilité aux maladies infectieuses chez la Souris sont détaillées ci-après.

Partie 4. Les différentes stratégies utilisées pour identifier les gènes impliqués dans la sensibilité aux maladies infectieuses chez la Souris

1. La réponse de l'hôte à une infection est un phénomène hautement complexe

La réponse de l'hôte à un agent pathogène met en œuvre un ensemble hautement complexe de phénomènes qui intègrent de nombreux mécanismes physiologiques et immunitaires. L'infection d'un animal fait appel à des réseaux de différents types cellulaires dans, et entre lesquels, des molécules interagissent en cascade pour activer des voies de signalisation. Il n'est donc pas étonnant que le déterminisme génétique de la réponse d'un système aussi complexe et aussi redondant que le système immunitaire implique un nombre important de gènes (Zinkernagel, 2007 communication personnelle; (Rose *et al.*, 2012)). L'étude de Johnson *et al.* sur la sensibilité à *Staphylococcus aureus* chez la Souris est un exemple, parmi de nombreux autres, de cette complexité. L'infection par *S. aureus* provoque chez l'humain une large gamme de manifestations cliniques allant d'une colonisation asymptomatique à une endocardite mortelle. S'il est clair que des facteurs bactériens et environnementaux influencent la sévérité de l'infection par *S. aureus*, des études de cas familiaux ont également démontré l'influence des facteurs génétiques chez l'Homme (Hill, 2001). Chez la souris, on observe une sensibilité variable entre différentes lignées consanguines, entretenues dans les mêmes conditions environnementales, et ayant reçu le même inoculum de la même souche de *S. aureus*, ce qui suggère fortement que des facteurs génétiques de l'hôte influence cette sensibilité. Johnson *et al.* ont mesuré quatre paramètres quantitatifs sur 13 lignées de souris consanguines : la médiane de survie, la charge bactérienne dans le rein, la charge bactérienne dans la cavité péritonéale et le taux d'IL6 dans le sérum. Sur la base de la médiane de survie, ils ont défini trois catégories d'animaux : les résistants (qui ne meurent pas, c'est le cas de 3 lignées), les hyper-sensibles (qui meurent tous en moins de 24h : 4 lignées) et les intermédiaires (6 lignées), (Johnson *et al.*, 2012). Ils ont mesuré les valeurs des trois autres phénotypes pour chacune des 13 lignées et ont remarqué que les valeurs phénotypiques étaient variables entre les 3 lignées résistantes, et l'étaient également entre les 4 lignées hyper-sensibles. Les auteurs ont par exemple montré qu'il n'y avait pas de corrélation entre la charge bactérienne dans le péritoine et la résistance, une lignée résistante pouvant avoir la même charge qu'une lignée sensible. La même observation a été effectuée chez les lignées sensibles (Johnson *et al.*, 2012). Cet exemple illustre que deux lignées qui survivent à un pathogène ne survivent pas forcément pour les mêmes raisons, de même deux lignées sensibles qui meurent, ne meurent pas forcément pour les mêmes raisons. Deux phénotypes similaires peuvent être dus à des facteurs génétiques sous-jacents différents et plusieurs

facteurs génétiques peuvent contribuer au même phénotype (illustré par l'étude de (Lipoldova *et al.*, 2000)).

1.1. Le choix du phénotype étudié pour caractériser un processus infectieux

Le choix du phénotype à étudier pour suivre le déroulement et la gravité d'un processus infectieux dépend évidemment de l'agent pathogène et de la maladie qu'il induit. A chaque interaction entre un hôte et un pathogène particulier correspond une pathogénèse différente, un déroulement spécifique du processus infectieux. Le choix du phénotype à étudier pour rechercher les déterminants génétiques de la sensibilité de l'hôte nécessite de connaître avec le plus de détail possible le déroulement du processus infectieux. Pour faire un choix pertinent, il est nécessaire de connaître les symptômes, les différentes phases de la maladie, la dissémination du pathogène, les organes atteints, et même si possible l'évolution des populations cellulaires immunitaires et des molécules qui entrent en jeu dans la réponse de l'hôte. Les gènes n'agissent que très indirectement sur la capacité d'un individu à survivre à une infection. C'est pourquoi, connaître en détail les perturbations des fonctions biologiques de l'hôte est indispensable si l'on veut identifier des "sous-phénotypes" (mesurés au niveau cellulaire, tissulaire ou fonctionnel) qui sont sous un contrôle génétique plus direct.

Ainsi, dans le cadre de l'étude des maladies infectieuses, différents phénotypes peuvent être étudiés suivant le type de pathogène et le déroulement du processus infectieux. De plus, en fonction des connaissances de la physiopathologie de l'infection, des phénotypes plus ou moins globaux peuvent être mesurés (Figure 8).

Le choix du phénotype dépend aussi de la nécessité de garder l'animal vivant pour une utilisation ultérieure (étude d'expression ou reproduction). Des phénotypes tels que la survie ou la mesure de la charge de pathogène dans les organes cibles requièrent le sacrifice de l'animal. Des phénotypes tels que la variation du poids corporel, de la température ou de marqueurs biochimiques sanguins permettent de garder l'animal en vie.

1.1.1. La mort ou la survie

Le phénotype de mort ou de survie des animaux infectés est largement utilisé dans les études portant sur le déterminisme génétique des maladies infectieuses. C'est le phénotype le plus global qui soit en terme de réponse de l'hôte à un pathogène. Ce phénotype reflète l'issue globale de la maladie puisqu'il inclut l'évolution de l'infection *in vivo* dans l'organisme entier et la mise en place d'une réponse immunitaire complexe dont l'issue est soit la défaillance des organes vitaux conduisant à la mort, soit la survie de l'animal.

Le phénotype mort/survie se décompose en deux paramètres : le taux de survie des animaux et la durée de survie (ou latence) des animaux qui sont morts de l'infection. Selon l'agent pathogène, ces deux paramètres présentent une variabilité plus ou moins grande à l'intérieur d'un groupe d'animaux génétiquement homogènes analysés dans des conditions parfaitement standardisées. Cette variabilité s'explique par le fait que la mort d'un individu résulte d'interactions complexes entre le micro-organisme et l'hôte, qui sont chacune sujettes à des fluctuations stochastiques. Une forte variabilité phénotypique diminue la puissance de détection des locus recherchés. Il faut noter aussi que, si l'on s'en tient à la mort ou la survie de l'animal, chaque individu fournit une information binaire et qu'il faut des effectifs très importants pour mettre en évidence une association entre le génotype à un locus et une variation de survie. A titre d'exemple, il faut environ 300 souris *backcross* pour mettre en évidence un locus associé à une variation de 20% de la survie. Dans ces conditions, on ne peut détecter que des locus à effet fort. C'est par exemple le cas dans l'étude de Diez *et al.* qui ont identifié un QTL associé à la survie des souris C57BL/6J dans un *backcross* (C3HxC57BL/6J)xC3H suite à l'infection par *Citrobacter rodentium*. Ce QTL a ensuite été confirmé à l'aide de lignées congéniques (Diez *et al.*, 2011).

Le phénotype de mort ou survie a été utilisé dans de nombreux travaux sur la sensibilité à *Mycobacterium tuberculosis* réalisés à partir de différentes lignées consanguines. Mitsos *et al.* ont mis en évidence trois locus de résistance, *Trl-1*, *Trl-2* et *Trl-3*, en analysant la survie de 95 souris F2 entre les lignées C57BL/6J et DBA/2J (Mitsos *et al.*, 2000). Les locus *Tbs1* et *Tbs2* (Sanchez *et al.*, 2003) et *Sst1* (Kramnik *et al.*, 2000) ont également été identifiés en utilisant ce phénotype. Le locus *Sst1* s'est avéré influencer également la survie lors d'une infection par *Listeria monocytogenes* (Pan *et al.*, 2005). Un gène candidat a été mis en évidence pour ce locus, *Ipr1* (*Intracellular pathogen resistance 1*), surexprimé uniquement dans les macrophages résistants. Il limite la multiplication de *M. tuberculosis* et *L. monocytogenes* (Pan *et al.*, 2005).

La survie est aussi le phénotype qui a été utilisé dans notre étude. La survie des souris SEG/Pas à une infection par *Yersinia pestis* a été associée à trois QTL, *Ypr11* à *Ypr13* (Blanchet *et al.*, 2011).

Comme le soulignent Roy *et al.* dans une de leurs études sur les facteurs de sensibilité à *Salmonella typhimurium*, la variance du phénotype de survie est grande et implique des facteurs additionnels autres que génétiques. Il est possible, pour les lignées qui présentent une résistance intermédiaire, qu'une variation (même faible) de certains facteurs tels que la dose de pathogène injectée ou l'état du système immunitaire puissent exercer une influence plus forte sur le phénotype que chez les lignées très résistantes ou très sensibles moins sujettes aux fluctuations (Roy *et al.*, 2006).

1.1.2. Le poids

L'évolution du poids au cours de l'infection est un caractère quantitatif qui résulte de l'interaction de nombreux paramètres. Il reflète l'impact du processus infectieux sur l'état général, et est donc également un phénotype très global. Les animaux les plus sensibles sont généralement ceux qui perdent le plus de poids, alors que les résistants peuvent perdre du poids dans la phase symptomatique de l'infection, mais le reprennent en guérissant. Des observations préliminaires sont nécessaires pour connaître la fenêtre temporelle durant laquelle la différence de poids est significative entre individus sensibles et résistants. Le suivi du poids sur plusieurs jours permet d'étudier les variations de façon cumulée et peut aboutir à l'identification de différents locus en fonction de la phase de l'infection.

La mesure de la perte de poids est un phénotype très utilisé dans les études sur le virus grippal (*Influenza*) – voir <http://www.ncbi.nlm.nih.gov/pubmed/19293935>.

La mesure du poids a également été utilisée par Miyairi *et al.* comme caractère quantitatif pour mettre en évidence des locus de sensibilité à *Chlamydia psittaci* dans une collection de 40 lignées recombinantes BXD (Miyairi *et al.*, 2012). Un locus sur le chromosome 11 a été identifié grâce à ce phénotype. L'implication de ce locus a été confirmée dans la même étude par la mesure d'autres phénotypes comme l'activation des macrophages et des neutrophiles de la cavité péritonéale ou le dosage de cytokines (Miyairi *et al.*, 2012). Les autres phénotypes mesurés ont permis d'identifier un autre locus sur le chromosome 1, illustrant l'importance de mesurer plusieurs phénotypes pour décrire les facteurs génétiques intervenant dans un modèle aussi complexe que l'interaction d'un hôte et d'un pathogène.

1.1.3. La dissémination et la charge microbienne dans les organes cibles

La charge bactérienne, virale ou parasitaire des organes de l'animal infecté est un caractère quantitatif fréquemment utilisé. En effet, un agent infectieux prolifère généralement davantage dans un organisme sensible que chez un individu résistant. La mesure de ce caractère nécessite que les organes cibles soient connus et que la fenêtre temporelle où l'on peut caractériser une différence significative entre les lignées étudiées soit bien définie. Ce caractère permet d'évaluer la multiplication ainsi que la dissémination du pathogène dans l'organisme au cours de l'infection. C'est un bon phénotype pour caractériser la capacité de l'hôte à contrôler la prolifération du pathogène ou à l'éliminer. Ce phénotype peut être complété par une étude histologique des lésions de l'organe.

La collecte des organes dans le but de mesurer la charge de pathogène nécessite d'euthanasier l'animal. En revanche, l'étude de la bactériémie, ne nécessitant qu'une prise de sang, ne présente pas cet inconvénient. Les prélèvements de sang ont de surcroît l'avantage de pouvoir être répétés pour des

études cinétiques et permettent de détecter des locus associés aux différentes phases de l'infection. Néanmoins la prise de sang, acte stressant et affaiblissant pour l'animal, peut interférer avec le déroulement de l'infection.

Il faut noter que les progrès considérables de l'imagerie *in vivo* associée à l'utilisation de micro-organismes rendus luminescents par modification génétique ouvre des horizons très prometteurs dans l'étude des processus infectieux. Elle permet de suivre longitudinalement un même animal au long du développement du processus et d'évaluer de façon semi-quantitative la dissémination de l'agent pathogène et la charge de certains organes. Si elle n'a pas encore été utilisée pour rechercher des facteurs génétiques de l'hôte, nul doute qu'elle le sera prochainement.

Dans une étude sur les facteurs de sensibilité à *Candida albicans*, Radovanovic *et al.* ont mesuré la charge du pathogène dans le rein de 23 lignées consanguines de souris afin de réaliser une analyse haplotypique pour trouver les locus contrôlant la capacité des animaux à contenir la multiplication de *C. albicans* (Radovanovic *et al.*, 2011). Ils ont montré qu'un allèle du gène *C5* (facteur 5 du complément, impliqué dans le complexe d'attaque membranaire qui participe à la lyse des cellules infectées) était associé à une moindre charge bactérienne dans le rein (Radovanovic *et al.*, 2011). Un croisement en deux générations de type *intercross* entre une lignée très sensible A/J et une autre résistante AKR/J a permis de compléter cette analyse génétique en identifiant deux locus supplémentaires : *Carg3* (*Candida albicans* resistance gene 3) et *Carg4*, indépendants de l'allèle de résistance au gène *C5*. Ces deux locus influencent aussi la charge de *C. albicans* dans le rein (Radovanovic *et al.*, 2011).

1.1.4. Les lésions histologiques

L'analyse histologique permet de quantifier les lésions dans les organes cibles.

Les mécanismes en jeu dans les interactions hôte-agent pathogène conduisent fréquemment à des modifications de l'architecture tissulaire ou de la morphologie cellulaire dans certains organes (inflammation, prolifération, nécrose, apoptose, etc.). L'étude de ces lésions permet d'enregistrer des paramètres aussi bien qualitatifs (type de lésions) que semi-quantitatifs (grade histologique).

Les lésions cutanées ont été utilisées comme phénotype pour identifier les déterminants génétiques contrôlant la sensibilité à *Leishmania major* (Lipoldova *et al.*, 2000; Roberts *et al.*, 1997). Deux régions chromosomiques, *Lmr4* et *Lmr5* (pour *Leishmania major* resistance locus), ont ainsi pu être liées à ce caractère. Le foie et la rate étant également des organes cibles lors de leishmaniose, la splénomégalie et l'hépatomégalie ont aussi été mesurées, permettant d'identifier un autre locus, *Lmr3*, et de constater que *Lmr5* influence aussi la splénomégalie (Lipoldova *et al.*, 2000).

L'étude de lésions de démyélinisation de la moelle épinière chez des souris infectées par le virus de Theiler a permis de mettre en évidence le rôle du locus *H-2D* (Lipton *et al.*, 1995). L'allèle déficient du gène *H-2D* est associé à une élimination réduite et plus lente du virus, ce qui augmente durée de la réponse inflammatoire entraînant des lésions de démyélinisation et une paralysie postérieure (Brahic *et al.*, 1998).

1.1.5. La réponse inflammatoire

Le suivi de l'évolution des populations cellulaires de l'immunité ou le dosage de biomarqueurs de l'inflammation (interleukines, cytokines, etc.) permettent d'évaluer la réponse inflammatoire de l'hôte. Le niveau d'expression de ces biomarqueurs peut être suivi en concentration de protéines ou en ARNm. Dans ce cas, les gènes identifiés s'apparentent à des eQTL (*expression-QTLs*) qui peuvent être situés à proximité immédiate de la séquence codante, par exemple dans le promoteur (*cis-eQTL*) ou à distance (*trans-eQTL*).

L'étude de Lipoldova *et al.* sur la sensibilité à une infection par *Leishmania major* a combiné différents phénotypes, certains relativement globaux comme la splénomégalie et l'hépatomégalie ou encore la taille des lésions cutanées mais aussi des sous-phénotypes très précis comme le taux d'IgE, d'IL4, d'IL12 et d'IFN γ dans le sérum et l'expansion des populations lymphocytaires des souris infectées (Lipoldova *et al.*, 2000). Cinq locus, *Lmr3* à *Lmr7*, ont été détectés grâce à la mesure de ces phénotypes. *Lmr3* contrôle la splénomégalie, l'hépatomégalie et les taux sériques d'IgE et d'IFN γ . *Lmr4* influence la taille des lésions cutanées, et, en interaction avec *Lmr5*, le taux sérique d'IFN γ . *Lmr5* contrôle la taille des lésions, la splénomégalie et les taux sériques d'IL12, d'IFN γ et d'IgE. *Lmr6* contrôle le taux sérique d'IL4 et enfin *Lmr7* agit sur la prolifération de lymphocytes dans le sang des souris infectées (Lipoldova *et al.*, 2000). La majorité des paramètres mesurés dans cette étude sont sous le contrôle de plusieurs locus (caractères multigéniques) et la majorité des locus contrôlent plusieurs de ces paramètres (locus pléiotropes). Cette étude donne une vue assez globale de l'architecture génétique de la réponse à *Leishmania major* grâce à la mesure de nombreux paramètres précis pour lesquels les lignées parentales diffèrent.

Les phénotypes aussi précis que le dosage de cytokines sont généralement plus robustes et reproductibles que des phénotypes globaux, à condition de bien maîtriser les paramètres d'environnement (en particulier l'environnement microbiologique). En revanche, ils ne reflètent qu'une partie de la réponse de l'hôte à un pathogène. La mesure de plusieurs biomarqueurs est nécessaire pour avoir une vue plus globale de la réponse à l'infection chez l'hôte étudié.

La mesure de différents phénotypes multiplie les « angles de vue » sur l'interaction entre l'hôte et le pathogène, comme illustré dans l'étude de Lipoldova *et al.* (Lipoldova *et al.*, 2000). Les

phénotypes tels que la survie ou la charge de l'agent pathogène dans un organe donnent une vision globale de cette interaction mais sont plus éloignés de l'action directe des gènes. En mesurant des phénotypes plus discrets on se rapproche de l'action des gènes mais on perd la vision globale des interactions uniques qui caractérisent la réponse d'un hôte particulier à un pathogène particulier (Figure 8). Il s'agit donc d'approches complémentaires.

Mesurer un phénotype global présente l'avantage de ne faire aucun à priori sur les mécanismes en jeu lors de l'infection. Cela se révèle utile quand la pathogénie de la maladie est mal connue. La diversité des phénotypes mesurables permet de caractériser différentes étapes du processus inflammatoire et immunitaire et de mettre en évidence à la fois les locus communs, qui contrôlent plusieurs paramètres à une période donnée de l'infection, mais aussi les locus qui agissent sur des phases spécifiques du processus infectieux (Lipoldova *et al.*, 2000).

Le déterminisme génétique des variables quantitatives est étudié par une approche QTL, qui lie deux types d'informations : les données phénotypiques et des données génotypiques (Falconer *et al.*, 1996; Kearsey, 1998; Lynch *et al.*, 1998). L'analyse QTL part d'une différence de sensibilité existante entre des lignées pour trouver les gènes qui la contrôlent (génétique directe). Elle permet d'identifier le nombre de régions chromosomiques en jeu, d'étudier l'effet de chaque région sur le phénotype étudié et les interactions entre ces régions. L'approche QTL a largement été utilisée dans l'étude des maladies infectieuses et c'est sur cette approche que se focalise très largement cette introduction. Néanmoins, les approches de génétique inverse, comme la mutagenèse à l'ENU et l'inactivation de gènes (*Knock-out* ou KO), ont permis l'étude de mutations qui altèrent un phénotype de résistance d'une lignée de souris à un agent infectieux. Je ne présenterai ici que la mutagenèse à l'ENU en raison de sa contribution importante à l'étude des maladies infectieuses.

2. La mutagenèse à l'ENU : une approche aléatoire mais efficace

L'ENU (N-ethyl-N-nitrosourea) est un mutagène puissant aussi bien chez les plantes, les invertébrés (tels que la drosophile) ou les mammifères (comme la Souris) (Russell *et al.*, 1979). L'ENU provoque un fort taux de mutations ponctuelles au hasard dans le génome. Lorsqu'elles surviennent dans des spermatogonies, le mâle peut les transmettre à sa descendance, dans laquelle on peut étudier leurs effets (Justice *et al.*, 1999). Ces descendants sont soumis à des tests phénotypiques (criblage) pour identifier les individus "phénotévants" éventuellement porteurs d'une mutation qui est ensuite caractérisée sur le plan moléculaire. En créant de nombreuses mutations aléatoires, on peut espérer muter peu à peu tous les gènes impliqués dans un phénotype, comme la résistance à une

maladie infectieuse. Cet ensemble de gènes impliqués dans la résistance à un agent infectieux, a été défini comme le « résistome » (Crozat *et al.*, 2007; Marathe *et al.*, 2004).

Crozat *et al.* ont identifié des gènes impliqués dans le résistome au cytomégalo virus murin (*Mouse Cytomegalovirus* ou MCMV) (Crozat *et al.*, 2006). Crozat *et al.* définissent le résistome au MCMV comme étant le nombre de gènes non redondants qui permettent à une souris C57BL/6 de résister à l'infection par le MCMV. Ces auteurs ont criblé les descendants de 3500 mâles mutagénisés et recherché des individus présentant des signes de maladie ou mourant à la suite d'une infection par le MCMV. Huit mutations transmissibles ont été trouvées comme induisant une sensibilité accrue au MCMV (Crozat *et al.*, 2006). Une des mutations, *Domino*, est une mutation ponctuelle qui altère le site de liaison de STAT1 à l'ADN, ce qui conduit à un défaut de transcription des gènes ciblés par STAT1 (Crozat *et al.*). Une autre mutation, *Jinx*, dans le gène *Unc13d*, abolit la cytotoxicité des cellules *natural killer* (NK) (Crozat *et al.*, 2007). Les cellules NK et les lymphocytes T cytotoxiques des souris portant la mutation sont toujours capables de produire des cytokines mais présentent un défaut de dégranulation, ce qui conduit à une multiplication virale non maîtrisée et à la mort de l'hôte (Crozat *et al.*, 2007). La mutation *Jinx* représente un modèle pour la maladie humaine appelée lymphohistiocytose hémophagocytaire familiale de type 3 (FHL3). En effet, le gène *Unc13d* est l'homologue du gène humain *MUNC13-4* pour lequel des mutations sont responsables de la FHL3 (Crozat *et al.*, 2007). Une autre mutation, *Joker*, également identifiée dans ce criblage, modifie le site accepteur d'épissage entre les exons 5 et 6 du gène *Itgb2*, conduisant à une délétion totale de l'exon 6 dans l'ARNm issu de la transcription de ce gène. Cette mutation pourrait être à l'origine d'une absence d'expression de molécules adhésion à la surface des cellules NK qui affecte leur développement (Thèse de Karine Crozat, 2007). Trois autres mutations identifiées, *Goodnight*, *Solitaire* et *Slumber*, conduisent à la mort rapide des animaux infectés par le MCMV. *Slumber* est co-localisé avec une autre mutation, *Mayday*, qui est un allèle nul du gène *Kcnj8*. Ce gène code la protéine Kir6.1 qui se combine avec SUR2 pour former un canal potassium ATP dépendant (K_{ATP}) exprimé dans les muscles lisses de l'artère coronaire et dans les cellules endothéliales (Croker *et al.*, 2007). L'étude de cette mutation montre que K_{ATP} possède une fonction cruciale de maintien de l'homéostasie pendant l'infection et prévient la vasoconstriction de l'artère coronaire induite par les cytokines produites durant l'infection (Croker *et al.*, 2007). Cette mutation qui provoque une hypersensibilité à l'infection par le MCMV, a permis de mettre en évidence l'existence d'une nouvelle catégorie de gènes qui n'agissent pas directement sur le processus infectieux mais sur le rétablissement des perturbations de l'homéostasie provoquées par l'infection (Croker *et al.*, 2007).

Cette étude de mutagenèse à l'ENU a permis d'estimer qu'environ 480 gènes sont probablement impliqués dans le résistome au MCMV chez la Souris (Crozat *et al.*, 2006). Cet exemple montre qu'en produisant un grand nombre de mutations aléatoires, on peut espérer inventorier tous les gènes impliqués dans la résistance à un pathogène et ouvrir la voie à la recherche de nouvelles voies de signalisation (Crozat *et al.*, 2006; Hoebe *et al.*, 2006).

Une limitation de la mutagenèse à l'ENU est qu'un gène nécessaire à la réponse de l'hôte lors d'une infection peut aussi posséder une fonction vitale durant le développement et ne pas être identifié à cause d'une mortalité embryonnaire (Richer *et al.*, 2008).

Malgré cela, les groupes de Bruce Beutler et de Sylvie Vidal ont prouvé l'efficacité de cette technique pour disséquer l'architecture génétique des caractères complexes (Croker *et al.*, 2007; Crozat *et al.*, 2006; Crozat *et al.*, 2007; Vidal *et al.*, 2008).

A la différence des approches de génétique inverse où l'effet d'une mutation sur un phénotype est analysé, l'approche QTL consiste à disséquer les facteurs génétiques responsables d'une différence de phénotype observé entre des lignées. Différentes populations génétiques de référence et différents types de croisements au sein de ces populations sont couramment utilisées dans l'approche QTL.

3. Croisements entre deux lignées différant pour leur vulnérabilité

Après avoir caractérisé une différence de sensibilité entre deux lignées consanguines, l'une des approches les plus utilisées pour détecter des QTL, c'est-à-dire des régions génomiques qui participent au contrôle d'un caractère complexe, est de produire une population en ségrégation entre ces deux lignées à l'aide de croisements en deux générations. Deux schémas de croisement sont utilisés : le « backcross » (croisement en retour) qui donne une population nommée N2 ou l'« intercross » qui donne une population F2 (Figure 9), (Darvasi, 1998).

Les individus de la population N2 ou F2 sont phénotypés et génotypés avec des marqueurs répartis sur l'ensemble du génome. On recherche ensuite une association significative entre le génotype à une région génomique et une différence de phénotype. Les croisements en deux générations entre deux lignées consanguines de souris de laboratoire ont été largement employés comme stratégie expérimentale dans la cartographie des QTL pour un grand nombre de phénotypes.

Sebastiani *et al.* ont étudié la durée de survie de différentes lignées de souris après infection par *Salmonella typhimurium*. Les souris de la lignée MOLF/Ei présentaient un taux de mortalité de 100%, comme les souris C57BL/6J, malgré la présence chez MOLF/Ei d'un allèle fonctionnel au gène *Nramp1* (maintenant appelé *Slc11a1*) alors que C57BL/6J porte un allèle déficient pour ce gène (Vidal *et al.*, 1995). Le fait que la lignée MOLF/Ei était toutefois très sensible à *S. typhimurium* suggérait que d'autres facteurs génétiques que *Nramp1* étaient capables d'induire une sensibilité à ce pathogène (Sebastiani *et al.*, 1998). L'analyse de la survie de 191 animaux issus d'un croisement de type F2 entre les lignées C57BL/6J et MOLF/Ei a permis d'identifier trois locus. Le premier sur le chromosome 1 était co-localisé avec *Nramp1*, et l'allèle protecteur MOLF/Ei présentait un mode de transmission co-

dominant additif à ce locus. Le second (*Ity2*), localisé sur le chromosome 11, avait également un effet protecteur additif et le troisième (*Ity3*), localisé sur le chromosome 1 en position distale par rapport à *Nramp1*, conférait une sensibilité accrue sur un mode récessif. A eux trois, ces QTL expliquaient environ 55% de la variance phénotypique observée entre les lignées parentales (Sebastiani *et al.*, 1998).

Le choix entre N2 et F2 est ouvert car on ne connaît pas, à priori, le mode de transmission (récessif, co-dominant, dominant, additif, épistatique) des allèles aux différents QTL qui ségrégent dans le croisement. Pour un même nombre de descendants, le nombre d'animaux par classe de génotype est plus petit dans la F2, donc la puissance d'identification des QTL est plus faible. Néanmoins, la présence de trois classes génotypiques augmente la probabilité de mettre en évidence des interactions épistatiques entre les QTL. Produire et analyser une F2 présente l'avantage d'avoir une vision plus globale de l'architecture génétique du caractère étudié (nombre de QTL, effets des QTL). L'analyse d'une population N2 est préférable si l'objectif de l'étude est de mettre en évidence un ou des QTL à effet majeur, car ce type de croisement est plus puissant (Darvasi, 1998; Panthier *et al.*, 2003).

Les deux avantages majeurs de ce type de croisements sont la liberté de choisir les lignées parentales et l'obtention rapide d'animaux N2 ou F2 à partir des lignées choisies. Quel que soit le phénotype observé, plusieurs croisements en deux générations peuvent être effectués à partir de différentes lignées présentant les mêmes caractéristiques de départ dans le but d'identifier des régions chromosomiques communes et de diminuer les intervalles de confiance (et donc le nombre de gènes présents) ou d'identifier de nouveaux locus.

Chaque individu analysé représente un réassortiment unique et non fixé des deux génomes parentaux. Tous les individus qui présentent le même génotype à un marqueur donné portent des génotypes différents aux autres marqueurs du génome. Ainsi lorsque l'on compare les valeurs phénotypiques moyennes des homozygotes et des hétérozygotes à ce marqueur, l'hétérogénéité génétique des individus à l'intérieur des deux groupes de génotypes est responsable d'une variabilité phénotypique intra-groupe qui affaiblit la puissance de la comparaison (Panthier *et al.*, 2003). De plus, chaque animal N2 ou F2 est unique, donc le phénotype associé à un génotype particulier ne peut être répliqué. Il est donc très important que tous les individus soient phénotypés avec le plus grand soin et dans les mêmes conditions expérimentales pour réduire au maximum l'ampleur des variations liées aux conditions environnementales. De plus, un autre inconvénient des croisements en deux générations est la taille de l'intervalle de confiance du QTL identifié. Il n'est pas rare que celui-ci couvre plusieurs dizaines de mégabases et contienne des centaines de gènes. En effet, les individus analysés résultent d'une ou deux méioses informatives qui n'ont pas pu fragmenter finement le génome (Panthier *et al.*, 2003).

Une alternative aux croisements en deux générations pour détecter des QTL est l'utilisation de populations génétiques de référence.

4. L'utilisation de populations génétiques de référence

Il existe de nombreuses populations de référence qui permettent d'identifier des QTL. Les lignées consanguines, les lignées recombinantes consanguines, les lignées recombinantes congéniques ou encore les lignées consomiques, sont autant de populations stables génétiquement et polymorphes qui permettent de détecter des QTL de sensibilité aux maladies infectieuses (Figure 10).

4.1. Les lignées consanguines

4.1.1. Caractéristiques des lignées consanguines

Les lignées consanguines sont des populations d'individus génétiquement identiques et homozygotes à tous les locus de leur génome. Chaque lignée représente une population de référence génétiquement stable dans le temps. La composition génétique étant figée, les caractéristiques phénotypiques obtenues sur une lignée sont également stables dans le temps, à condition que les facteurs environnementaux et les caractéristiques du pathogène restent les mêmes (Panthier *et al.*, 2003).

L'étude de l'origine des souris de laboratoire a montré que leur génome est une mosaïque des sous-espèces *M. m. domesticus*, *M. m. musculus* et pour quelques pourcents seulement *M. m. castaneus* (Yang *et al.*, 2011 B). L'appareillement des lignées de laboratoire, qui résulte d'un nombre restreint de géniteurs qui ont contribué à leur développement, est reflété par un fort taux de conservation entre les lignées pour environ deux tiers des régions génomiques (Wade *et al.*, 2002). Cette observation suggère que les lignées de laboratoire couvrent une diversité génétique réduite en regard de celle qui existe entre les espèces du genre *Mus*, auquel appartiennent les souris. Néanmoins, du polymorphisme peut être apporté grâce à l'existence de lignées dérivées de souris sauvages capturées dans diverses régions du globe (Bonhomme *et al.*, 1987). Les lignées de souris sauvages présentent un haut degré de polymorphisme génétique entre elles et avec les lignées de laboratoire. Ce polymorphisme se traduit par des différences phénotypiques bien plus grandes que les différences observées au sein des souris de laboratoire. Si l'utilisation de ces lignées présente un certain nombre de difficultés, elles présentent un atout majeur dans la recherche de QTL (Dejager *et al.*, 2009; Guenet *et al.*, 2003). Les caractéristiques des souris sauvages seront détaillées dans la partie 7 de cette introduction.

Les études de Sebastiani *et al.* et Turcotte *et al.* ont utilisé des croisements interspécifiques entre des lignées de laboratoire et MOLF/Ei et SPRET/EiJ, dérivées respectivement de la sous-espèce *Mus musculus molossinus* et de l'espèce *Mus spretus*, pour mettre en évidence de nouveaux locus de sensibilité à *S. typhimurium* et *M. bovis*, respectivement (Sebastiani *et al.*, 1998; Turcotte *et al.*, 2006).

4.1.2. La cartographie d'association ou *Haplotype Association Mapping*

L'*Haplotype Association Mapping* (HAM) chez la souris utilise les données de génotypage pour les millions de marqueurs qui couvrent l'ensemble du génome des différentes lignées consanguines, afin de rechercher des associations entre un caractère et un génotype en considérant les différentes lignées consanguines comme les individus d'une population (Pletcher *et al.*, 2004; Tsaih *et al.*, 2009). La base de données du *Mouse Phenome Project* (www.jax.org/phenome) met à disposition des milliers de données phénotypiques, de génotypage et d'expression entre différentes lignées consanguines. De plus 17 lignées ont été entièrement séquencées et sont disponibles à ce jour (données du Jackson lab : <http://research.jax.org>). La base de données du *Mouse Genome Informatics* (MGI) met également à disposition des données de génotypage de SNP très denses. La collecte de ces données a permis une avancée rapide des études de cartographie d'association, aussi connue sous le nom d'*in silico* mapping (Grupe *et al.*, 2001). Depuis plusieurs années, l'HAM est utilisé pour identifier de nouveaux QTL, identifier quelles lignées consanguines sont les plus informatives pour initier de nouvelles analyses QTL ou encore pour affiner la localisation de QTL trouvés dans des études antérieures (Liu *et al.*, 2007).

L'étude de Johnson *et al.* sur les facteurs de sensibilité à une infection par *Staphylococcus aureus* a utilisé une approche d'HAM sur 13 lignées consanguines dont le génotype a été obtenu sur la base de données de *Perlegen SNP* (disponible sur le site web du *Jackson Laboratory*). Les 13 lignées ont été infectées par *S. aureus* et quatre phénotypes ont été mesurés : la survie, la charge bactérienne dans le rein et dans le liquide péritonéal 24h post-infection et le taux sérique d'IL6 (Johnson *et al.*, 2012). L'HAM a mis en évidence un bloc haplotypique de 1,9 kb sur le chromosome 7, significativement associé à la charge bactérienne dans le rein à 24h. Ils montraient que 9 des 13 lignées qui ont une charge bactérienne faible (<100 ufc/g) présentaient trois haplotypes différents (GTC, GCC et AAT) pour ce bloc. Parmi les 4 lignées restantes, trois partageaient l'haplotype GTT et avaient une charge bactérienne plus élevée (autour de 250 ufc/g). La dernière lignée portait l'haplotype ACT et présentait une charge bactérienne bien supérieure aux autres (>1000 ufc/g). Le bloc haplotypique détecté ne contenait pas de gène mais était situé dans un cluster de 26 protéines kallikréines dont le rôle dans l'immunité innée a été démontré antérieurement (Sotiropoulou *et al.*, 2010). Pour compléter l'analyse d'haplotype, une étude d'expression pangénomique dans les lignées présentant les phénotypes les plus extrêmes en terme de survie et de charge bactérienne dans le rein, a

mis en évidence que les protéines kallikréines étaient différenciellement régulées entre les lignées sensibles et résistantes en réponse à l'infection (Johnson *et al.*, 2012).

La méthode d'HAM présente différents avantages par rapport à l'approche traditionnelle de détection des QTL : un gain de temps, des coûts réduits (le génotype – voire la séquence – des lignées est déjà connu) et la possibilité de détecter des régions génomiques plus petites.

Néanmoins l'étude de Johnson *et al.* met aussi en évidence deux limitations de la méthode d'HAM pour la détection des régions génomiques associées à un caractère complexe. La première est que l'HAM présume que des similarités phénotypiques entre des lignées consanguines reposent sur des variants génétiques communs. Ceci est particulièrement inexact lorsque des lignées très divergentes sont étudiées, ce qui nécessite de prendre en compte la structure de la population. Un phénotype peut résulter de l'interaction entre de nombreux facteurs génétiques, chacun d'eux étant partagé par une partie des lignées qui présentent le même phénotype. La puissance de l'analyse d'haplotype est dans ce cas réduite et peut conduire à la détection de fausses associations. De plus, dans cette étude, le nombre relativement faible de lignées utilisées peut aussi participer à un manque de puissance nécessaire pour identifier des QTL (Johnson *et al.*, 2012).

Malgré ces limitations, l'analyse d'haplotype a permis d'identifier des régions fines qui ont ensuite été confirmées par d'autres méthodes.

4.2. *Les lignées recombinantes consanguines, recombinantes congéniques et consomiques*

Ces lignées issues de différents croisements entre deux ou plusieurs lignées consanguines ont le point commun d'avoir une composition génétique fixée et sont donc elles-mêmes des lignées consanguines. Elles représentent des outils d'une grande valeur pour la détection des QTL.

4.2.1. Les lignées recombinantes consanguines ou *Recombinant Inbred Strains* (RI)

Les lignées recombinantes consanguines sont produites à partir de deux lignées consanguines de laboratoire. Elles sont développées sous forme de jeu de lignées qui présentent un réassortiment unique des allèles provenant des deux lignées parentales, fixés à l'état homozygote. Chaque lignée RI a hérité en moyenne de la moitié des allèles de chacune des lignées parentales. L'exemple le plus abouti de lignées RI est la population BXD constituée de plus d'une centaine de lignées qui ont reçu chacune en moyenne 50% de la lignée C57BL/6J (B) et 50% de la lignée DBA/2 (D).

Abdeltawab *et al.*, dans une étude portant sur les facteurs de sensibilité à la souche 5448 d'un streptocoque de groupe A (GAS), ont phénotypé plusieurs lignées consanguines de souris et ont observé que la lignée DBA/2 était très sensible, alors que la lignée C57BL/6J était résistante. Ils ont alors tiré parti du jeu de lignées BXD existant et ont phénotypé entre 5 et 26 souris de 30 lignées recombinantes BXD pour trois phénotypes : la survie, la bactériémie et la charge bactérienne dans la rate 24h après inoculation de 10^7 bactéries en IV (Abdeltawab *et al.*, 2008). Ils ont identifié deux QTL, un commun aux trois phénotypes mesurés localisé sur le chromosome 2 et couvrant 12 Mb et un autre plus distal sur le même chromosome, commun à la bactériémie et à la charge bactérienne dans la rate, couvrant 25 Mb (Abdeltawab *et al.*, 2008). Le premier QTL a été confirmé dans plusieurs lignées qui n'avaient pas été testées au départ. Les lignées qui portaient le QTL sous forme allélique B étaient résistantes alors que celles qui portaient le QTL sous forme allélique D étaient sensibles. La taille de ce QTL de 12 Mb a ensuite été réduite à moins de 10 Mb grâce à l'utilisation de 6 lignées RI qui présentaient une recombinaison dans cet intervalle. La présence d'une lignée restée hétérozygote à ce locus a montré que l'allèle D conférait une sensibilité accrue sur un mode dominant (Abdeltawab *et al.*, 2008). Cette étude illustre l'intérêt des lignées RI dans la détection des QTL et leur cartographie fine. Un des atouts non négligeable de ces jeux de lignées RI est que le génotype précis de chaque lignée est déjà connu, ce qui évite l'étape de génotypage que nécessitent les croisements en deux générations. La puissance de détection et la précision de localisation des QTL sont améliorées en utilisant des lignées RI par rapport aux croisements en deux générations, car la variabilité phénotypique intra-groupe est réduite, puisque chaque lignée est une lignée consanguine.

De plus, Abdeltawab *et al.* soulignaient que la variation des phénotypes observée dans un jeu de lignées recombinantes était largement supérieure à celle observée entre les deux lignées parentales. Certaines lignées présentaient une sensibilité significativement supérieure à DBA/2 alors que d'autres étaient plus résistantes que C57BL/6J, ce qui illustre que la création de combinaisons originales de gènes polymorphes permet de détecter de nouveaux gènes impliqués dans la sensibilité à un pathogène (Abdeltawab *et al.*, 2008). Abdeltawab *et al.* ont associé cette recherche de QTL à une étude d'expression sur les gènes présents dans cet intervalle pour identifier des gènes candidats et des voies de signalisation potentiellement impliquées dans le phénotype (Abdeltawab *et al.*, 2008).

L'utilisation des lignées RI est néanmoins limitée par le nombre de lignées existantes, les jeux comportent en général 15 à 30 lignées tout au plus, à l'exception de la population BXD qui en comporte une centaine. La précision de détection d'un QTL dépendant du nombre de lignées analysées : si celui-ci est trop faible la puissance de détection sera médiocre.

Lorsqu'un caractère est contrôlé par plusieurs QTL, chaque lignée RI porte les allèles d'une des deux lignées parentales en moyenne pour la moitié de ces QTL, ce qui rend difficile d'établir les effets individuels de chaque QTL, surtout lorsque certains sont liés par des interactions épistatiques.

Pour résoudre cette difficulté, Demant *et al.* ont développé des lignées recombinantes congéniques (Demant *et al.*, 1986).

4.2.2. Les lignées recombinantes congéniques ou *Recombinant Congenic Strains* (RC)

Les lignées recombinantes congéniques diffèrent des lignées RI par la proportion de chaque lignée parentale (Demant *et al.*, 1986). Les RC sont établies à partir d'un croisement entre deux lignées consanguines, suivi de deux *backcross* successifs de la lignée donneuse avec la lignée receveuse, une succession de croisements frère-sœur permet ensuite de fixer les allèles de la lignée donneuse à l'état homozygote afin d'obtenir une lignée stable génétiquement. Chaque lignée RC diffère de la lignée receveuse pour $1/8^{\text{ème}}$ (12,5%) de son génome (Panthier *et al.*, 2003). La probabilité est grande qu'une lignée RC ait hérité d'un seul QTL de la lignée donneuse. Il est ainsi possible d'isoler les différents QTL dans des lignées RC distinctes. La puissance d'analyse est augmentée dans le cas des caractères contrôlés par de nombreux QTL.

Roy *et al.* ont phénotypé un jeu complet de lignées recombinantes congéniques AcB et BcA, issues des lignées parentales A/J et C57BL/6J, pour leur sensibilité à une infection par *S. typhimurium* (Roy *et al.*, 2006). Des études précédentes avaient mis en évidence le rôle majeur du gène *Nramp1* dans la résistance à une infection par *S. typhimurium* (Vidal *et al.*, 1995). Néanmoins, *Nramp1* ne contrôle qu'une partie de la différence de phénotype et d'autres gènes, à effet plus faible, participent au contrôle d'une infection par *S. typhimurium* (Roy *et al.*, 2006; Sebastiani *et al.*, 1998). Les souris C57BL/6J portent un allèle non fonctionnel pour le gène *Nramp1* et sont très sensibles à l'infection contrairement à la lignée A/J qui meurt plus tard (Roy *et al.*, 2006). Trente-six lignées AcB/BcA ont été phénotypées pour différents caractères : la mortalité, la charge bactérienne dans le foie et la rate 4 jours après l'injection IV de 10^3 ufc de *S. typhimurium* (Roy *et al.*, 2006). L'analyse du comportement du jeu de lignées RC a confirmé que le seul génotype au locus de *Nramp1* ne permettait pas de prédire l'issue de l'infection chez toutes les lignées. Par exemple, la lignée recombinante congénique AcB61 est extrêmement sensible à l'infection en dépit d'un allèle fonctionnel au locus *Nramp1*. En revanche, la lignée AcB64 présente une survie supérieure à la lignée parentale résistante A/J.

Ces deux lignées ont été utilisées, indépendamment l'une de l'autre, pour produire des individus F2 afin d'identifier de nouveaux locus de résistance à *S. typhimurium*. La lignée AcB61 a été croisée avec des individus résistants 129S6/SvEvTac pour produire 247 individus F2 pour lesquels le jour de mort a été relevé. L'analyse de cette population a permis de mettre en évidence deux QTL : *Ity4* et *Ity5* qui viennent compléter les trois premiers locus identifiés antérieurement dans un croisement *backcross* entre C57BL/6J et MOLF/Ei (Sebastiani *et al.*, 1998). *Ity4* a ensuite été disséqué finement et le gène *Pklr* (*Liver and red blood cell-specific pyruvate kinase*) a été mis en évidence

comme sous-tendant ce QTL (Roy *et al.*, 2007). On remarquera que ce ne sont pas directement les lignées recombinantes congéniques qui ont été utilisées pour détecter les QTL, mais un croisement F2 de celles-ci avec une lignée différente des parentales. En effet, l'introduction d'une nouvelle lignée consanguine augmente le polymorphisme contenu dans le croisement et maximise les chances d'identifier des QTL.

Trois autres locus *Ity6*, *Ity7* et *Ity8* ont été identifiés dans un second croisement F2 entre la lignée AcB64, qui présente une forte résistance, et la lignée DBA/2J (sensible).

Ces travaux sur des lignées recombinantes congéniques ont permis d'identifier de nouvelles régions chromosomiques impliquées dans la résistance à une infection par *S. typhimurium* en s'affranchissant de l'effet de *Nramp1* (Roy *et al.*, 2006).

L'étude de Lipoldova *et al.* illustre également la puissance des lignées RC pour disséquer le déterminisme génétique de la sensibilité *Leishmania major*. Cette étude utilise un jeu de lignées RC entre la lignée BALB/c sensible à *L. major* et la lignée STS/A plus résistante. Parmi ce jeu de lignée, la lignée CcS-5, qui a reçu en moyenne 12,5% des gènes de STS/A, présente les mêmes caractéristiques que la lignée STS/A pour de nombreux paramètres (tels que les lésions cutanées, le dosage d'IgE ou de cytokines) (Lipoldova *et al.*, 2000). La dissection génétique de la sensibilité à *L. major* a été grandement facilitée par l'apport de la lignée CcS-5 qui ne comporte que 12,5% du génome de la lignée STS/A. L'étude de cette lignée a permis d'identifier 5 locus impliqués dans différents phénotypes caractéristiques de la maladie (Lipoldova *et al.*, 2000).

Les RC partagent de nombreux avantages avec les RI, en particulier un génotype homozygote fixé et déterminé une fois pour toutes, des populations permanentes, une bonne puissance de détection et de résolution des QTLs. Comme pour les lignées RI, le nombre de RC détermine la puissance de détection du QTL et conditionne aussi le pourcentage du génome couvert par le jeu. Ainsi 20 lignées de 12,5% permettent de couvrir environ 93% du génome de la lignée donneuse. Si une des lignées montre une différence phénotypique intéressante, d'autres *backcross* sont effectués dans le but de réduire la part du génome issue de la lignée donneuse afin d'identifier les régions chromosomiques responsables. Une fois identifiées, les régions d'intérêt peuvent être isolées dans des lignées congéniques en réalisant au maximum 7 *backcross* successifs.

4.2.3. Les lignées consomiques ou *Chromosome Substitution Strains*

Des jeux de lignées consomiques ont aussi été établis pour accélérer la détection et la caractérisation moléculaire des QTL (Gregorova *et al.*, 2008; Nadeau *et al.*, 2000; Shao *et al.*, 2008;

Takada *et al.*, 2008; Xiao *et al.*, 2010). Ces lignées sont établies en transférant un chromosome d'une lignée donneuse dans un fonds génétique receveur par des *backcross* répétés associés à une sélection par marqueurs des individus intéressants, puis ces lignées sont conservées à l'état homozygote par des croisements frère-sœur. Elles ne diffèrent donc que pour un chromosome de la lignée receveuse.

Un exemple d'utilisation des lignées consomiques est l'étude de Ahn *et al.* sur la sensibilité à l'infection par *Staphylococcus aureus* (Ahn *et al.*, 2010). Parmi les lignées consanguines, la lignée A/J est très sensible à une infection par *S. aureus* alors que la lignée C57BL/6J est résistante. Ahn *et al.* ont tiré parti d'un jeu de lignées consomiques existantes portant chacun un chromosome de la lignée A/J dans un fond C57BL/6J (Nadeau *et al.*, 2000). L'analyse du jeu complet de lignées a permis de mettre en évidence que les chromosomes 8, 11 et 18 portaient des gènes qui rendaient les souris significativement plus sensibles à *S. aureus* que la lignée receveuse C57BL/6J. Le chromosome 18 de la lignée CSS18 a ensuite été fragmenté par un *backcross*, permettant l'identification de deux QTL sur ce chromosome (Ahn *et al.*, 2010).

L'utilisation des lignées consomiques présente un avantage en termes de temps par rapport à l'approche qui consiste à détecter des QTL dans un croisement en deux générations puis à créer des lignées congéniques pour confirmer et disséquer les QTL. Le phénotypage de l'ensemble du jeu de lignées permet directement de savoir sur quel(s) chromosome(s) sont localisés des gènes de la lignée donneuse qui influent sur le phénotype. Une fois que la ou les lignées qui présente(nt) un phénotype déviant de la lignée receveuse sont identifiées, un croisement en deux générations est suffisant pour cartographier finement le QTL. L'obtention de lignées congéniques à partir de lignées consomiques est aussi facilitée. Utiliser ce type de lignées réduit le niveau de complexité due à la co-ségrégation de plusieurs QTL (Ahn *et al.*, 2010; Nadeau *et al.*, 2000; Xiao *et al.*, 2010). Cependant c'est aussi une limitation majeure de ce type de lignée. Un phénotype dû à des interactions épistatiques de deux gènes ou plus localisés sur des chromosomes différents de la lignée donneuse ne sera pas retrouvé en utilisant des lignées consomiques.

Les deux principales limites des populations génétiques de référence sont une densité de recombinaison assez faible et surtout le fait que la plupart ont été dérivées de souris de laboratoire qui présentent un taux de polymorphisme peu élevé. Pour surmonter ces deux défauts, des généticiens de la souris réunis lors de la conférence de l'*International Mammalian Genome Society* en 2001, ont imaginé un nouveau système expérimental idéal dont les principales propriétés sont une très grande diversité génétique pour augmenter les chances de découvrir des phénotypes contrastés, une excellente résolution de localisation pour faciliter l'identification des gènes responsables des QTL et une parfaite connaissance du génome des individus de la population pour assurer une puissance maximale aux études d'association (Panthier *et al.*, 2012). De ces discussions est né le *Collaborative Cross*.

4.3. Le Collaborative Cross (CC)

Le Collaborative Cross constitue une super-extension du concept de lignées recombinantes consanguines. L'idée du CC était de créer un panel de 1000 lignées consanguines recombinantes en croisant huit lignées consanguines. Cinq sont des lignées de laboratoire (A/J, C57BL/6J, 129S1/SvImJ, NOD/ShiLtJ, NZO/H1LtJ) et les trois autres sont dérivées de sous-espèces de *Mus m. musculus* (PWK/PhJ), *Mus m. domesticus* (WSB/EiJ) et *Mus m. castaneus* (CAST/EiJ).

Un schéma strict de croisement a été mis en place pour que les allèles des lignées fondatrices soient représentés de façon équilibrée dans la population finale (Figure 11 et 12). Celles-ci renferment le maximum de la diversité génétique du croisement avec du matériel génétique de chacune des huit lignées parentales à part égale ($1/8 = 12,5\%$) à l'état homozygote (Panthier *et al.*, 2012). L'origine génétique variée des lignées entraîne un fort polymorphisme, qui se répercute en une variabilité phénotypique élevée entre les différentes lignées, une aubaine pour la recherche de QTL.

Un exemple d'utilisation des lignées du CC pour l'étude des maladies infectieuses est une étude sur la sensibilité à *Aspergillus fumigatus* (Durrant *et al.*, 2011). 371 souris issues de 66 lignées ont été testées pour leur sensibilité au pathogène *A. fumigatus*, champignon responsable de l'aspergillose, une maladie nosocomiale qui touche surtout les individus immunodéprimés. Le temps de survie après infection a été associé à sept QTL, localisés sur les chromosomes 2, 3, 8, 10 (deux QTL), 15, and 18 (Durrant *et al.*, 2011). Les QTL couvrent entre 6 et 17Mb et contiennent entre 39 et 139 gènes. Ces premiers résultats confirment la très bonne résolution de détection des QTL permise par le CC, si on tient compte du nombre restreint de souris testées dans cette étude ($n = 371$) et le choix de mesurer un phénotype très global, donc peu robuste, comme le temps de survie.

L'analyse de différents caractères phénotypiques a apporté la preuve de l'intérêt des trois lignées d'origine sauvage, source d'une grande variabilité. L'excellente résolution de localisation des QTL apportée par le CC vient du fait qu'il se produit en moyenne sept fois plus de recombinaisons entre deux marqueurs physiquement proches lors de l'établissement des lignées que dans une population N2 ou F2 (où il ne se produit qu'une seule méiose informative) (Panthier *et al.*, 2012).

L'architecture génétique des lignées du CC a été présentée lors du *Complex Trait Consortium 2012*. Un site web permet de disposer de toutes les informations génétiques connues sur les lignées : <http://csbio.unc.edu/CCstatus/index.py> (Welsh *et al.*, 2012).

5. Etudier les effets individuels des QTL : les lignées congéniques

La caractérisation d'un QTL identifié, soit à partir de l'analyse d'un croisement F2 ou *backcross*, soit à partir de l'analyse d'un jeu de lignées RI, RC ou consomiques, passe généralement par le développement de lignées congéniques (Panthier *et al.*, 2003). L'utilisation des lignées congéniques permet d'extraire une région chromosomique de son environnement génomique pour le placer dans un contexte génomique différent, afin de confirmer l'existence d'un QTL mais aussi d'affiner la localisation du QTL et d'étudier les interactions entre les différents QTL en croisant les lignées congéniques (Rogner *et al.*, 2003).

L'introduction d'un allèle ou d'un segment chromosomique provenant d'une lignée « donneuse » (A) dans une lignée « receveuse » (B) est réalisée par des croisements en retour répétés. La lignée A est croisée avec la lignée B pour obtenir une F1 (première génération ou N1). Un individu F1 est croisé en retour (*backcross*) avec un individu de la lignée receveuse B. Les individus de la deuxième génération (N2) sont génotypés, et ceux qui possèdent la région chromosomique d'intérêt issue de la lignée donneuse A sont accouplés avec des individus de la lignée A, et ainsi de suite jusqu'à N10. Dix générations de *backcross* sont nécessaires pour obtenir une lignée congénique ne possédant plus qu'une infime fraction du génome de la lignée donneuse A en dehors de la région chromosomique sélectionnée. En effet, la proportion du génome de la lignée A est divisée par deux à chaque génération. Elle est ainsi de 50 % chez les individus N1, 25 % chez les N2, 12,5 % chez les N3, etc. Au bout de 10 générations de *backcross*, cette proportion théorique est de $(\frac{1}{2})^{10}$, soit moins de 0,1% du génome (Tableau 3). En réalité, il est pratiquement certain que la lignée congénique ne porte plus aucun allèle de la lignée donneuse en dehors du locus transféré, par contre, du fait de la liaison génétique, ces individus hétérozygotes pour le locus transféré sont également hétérozygotes aux locus situés à proximité du locus transféré (Montagutelli *et al.*, 2004). Après dix générations de croisements ces locus adjacents transférés couvrent en moyenne 40 Mb autour du QTL, et contiennent en moyenne 400 gènes qui peuvent parfois influencer le phénotype étudié (Figure 13) (Montagutelli *et al.*, 2004).

Pour l'étude des QTL, le segment transféré dans une lignée congénique correspond au moins à l'intervalle de confiance du QTL déterminé à partir des données du croisement initial. Cet intervalle couvre plusieurs dizaines de mégabases. On confirme la présence d'un QTL dans ce segment en mettant en évidence une différence phénotypique entre la lignée congénique et la lignée receveuse. Ensuite, la localisation précise du QTL est affinée en fractionnant ce segment, par des *backcross* supplémentaires, en une série d'intervalles plus petits, partiellement chevauchants et recouvrant l'ensemble de l'intervalle initial. Ces sous-segments sont fixés indépendamment dans des lignées dites sous-congéniques (Figure 14). Lors de l'analyse des lignées sous-congéniques, il est fréquent

d'observer que l'effet attribué à un QTL est dû non pas à un gène, mais à plusieurs gènes physiquement liés (situés à proximité sur le même chromosome) (Bihl *et al.*, 1999; Grattan *et al.*, 2002; Montagutelli *et al.*, 2004; Rogner *et al.*, 2003). Lorsque le QTL est constitué de nombreux gènes à effet individuel faible, il est fréquent que le phénotype s'estompe à mesure que la région est disséquée en plus petits segments (Legare *et al.*). Il est alors important d'avoir identifié des sous-phénotypes qui caractérisent plus précisément le processus infectieux et qui pourront être mesurés si le phénotype global n'est pas retrouvé chez les congéniques ou sous-congéniques (Sissons *et al.*, 2009).

Les lignées congéniques constituent un outil de choix pour étudier les effets individuels des QTL (Panthier *et al.*, 2003). Elles permettent d'étudier l'effet de chaque QTL en comparant les caractéristiques de la lignée congénique et celles de la lignée receveuse. Le fait que le phénotype de la lignée congénique diffère significativement de celui de la lignée receveuse permet de confirmer la présence du QTL, en gardant à l'esprit qu'un QTL n'explique qu'une partie de la différence phénotypique observée entre les lignées parentales. Il arrive ainsi fréquemment que la présence du QTL ne puisse pas être confirmée en mesurant le phénotype qui avait permis d'identifier le QTL, mais seulement en analysant une composante du phénotype global, appelée sous-phénotype.

Les travaux de E. Wakeland et L. Morel sur une maladie auto-immune, le lupus érythémateux disséminé (LED), illustrent de façon remarquable la dissection d'un caractère complexe multigénique à l'aide de lignées congéniques et sous-congéniques. Le LED est caractérisé par la production d'immunoglobuline G spécifiques de constituants nucléaires, qui entraîne de nombreux dommages tissulaires (lésions inflammatoires cutanées, articulaires et rénales). Cette maladie présente une très grande variabilité phénotypique au sein des individus atteints en termes de manifestations cliniques et de niveau d'auto-anticorps produits (Vyse *et al.*, 1996). Plus de 12 locus (non-associés au complexe majeur d'histocompatibilité, CMH) ont été associés à l'étiologie du LED chez la Souris. Le locus *Sle1* (pour *Systemic lupus erythematosus*) situé sur le chromosome 1, a été identifié comme fortement lié à la maladie et constitue le locus majeur de sensibilité dans la lignée recombinante NZM2410, dans laquelle l'incidence de LED atteint 80% chez les mâles à un an (Morel *et al.*, 1994). Les souris de la lignée congénique C57BL/6.*Sle1* portant *Sle1* sous forme allélique NZM à l'état homozygote sur fond C57BL/6 (lignée qui ne développe pas de LED), présentent un fort taux d'IgG spécifiques de la chromatine et une activation élevée des cellules B et T. *Sle1* peut à lui seul provoquer la perte de la tolérance envers la chromatine (Sobel *et al.*, 2002). La lignée congénique C57BL/6.*Sle1* a permis de confirmer l'existence du QTL *Sle1* et d'étudier son effet dans un fonds génétique différent du fonds NZM.

Une analyse génétique détaillée du locus *Sle1* à l'aide de sous-congéniques a montré que *Sle1* est en fait composé de trois locus dénommés *Sle1a*, *Sle1b* et *Sle1c* présents dans le même intervalle congénique. Chaque lignée sous-congénique portant indépendamment *Sle1a* à *Sle1c* présente un profil sérologique différent et des caractéristiques cellulaires particulières (Morel *et al.*, 2001). Les lignées sous-congéniques ont ainsi permis de disséquer l'architecture de *Sle1*.

Un deuxième locus, *Sle2*, diminue le seuil d'activation des cellules B et un troisième locus, *Sle3*, entraîne la dérégulation des cellules T CD4⁺ (Morel *et al.*, 2000a; Morel *et al.*, 1997).

La force des lignées congéniques réside dans la possibilité de produire et d'analyser un grand nombre d'animaux ayant strictement le même génotype dans des conditions identiques pour ce qui concerne les facteurs environnementaux contrôlés, ce qui accroît considérablement la puissance de détection de différences phénotypiques et la possibilité de démontrer le rôle causal de chaque région chromosomique étudiée. Toutefois, les caractères qui résultent d'interactions additives ou épistatiques entre plusieurs QTL ne sont pas forcément reproductibles dans des lignées congéniques où chaque QTL est isolé. Croiser entre elles différentes lignées congéniques pour créer des lignées multi-congéniques permet d'étudier les interactions additives ou épistatiques entre ces QTL (Panthier *et al.*, 2003). Les croisements entre lignées congéniques permettent d'identifier les combinaisons de QTL nécessaires et suffisantes pour reproduire le phénotype global étudié.

Les lignées congéniques ont permis de décortiquer un caractère complexe multigénique, le LED, en une série de caractères monogéniques. Morel *et al.* ont ensuite croisé ces lignées entre elles afin d'obtenir des lignées bi- et des tri-congéniques, qui leur ont permis de reconstruire la complexité du caractère premièrement étudié. Les trois lignées congéniques portant les locus *Sle1*, *Sle2* et *Sle3* ont été croisées selon différentes combinaisons montrant que *Sle1* est indispensable pour que se développe une glomérulonéphrite lupique, avec un effet aggravant additif des autres QTL (Morel *et al.*, 2000b). L'observation de lésions modestes dans les lignées congéniques par rapport au phénotype fort trouvé dans la lignée bicongénique portant *Sle1* et *Sle3* suggère un rôle critique des interactions épistatiques dans la pathogénèse du LED (Wakui *et al.*, 2005).

Si les souris bicongéniques pour *Sle1* et *Sle3* présentent un phénotype aggravé par rapport aux congéniques simples (splénomégalie, expansion des populations de lymphocytes B et T CD4⁺, taux élevé d'auto-anticorps IgG spécifiques de la chromatine et glomérulonéphrite à forte pénétrance), la lignée tri-congénique portant les trois locus dans un fonds C57BL/6 développe une auto-immunité sévère à pénétrance complète, ainsi qu'une glomérulonéphrite létale (Mohan *et al.*, 1998; Mohan *et al.*, 1999; Morel *et al.*, 2000a). Ces résultats montrent très élégamment que les QTL, isolés et caractérisés

fonctionnellement dans différentes lignées congéniques, entraînent finalement, une fois combinés, l'expression complète de la maladie.

Un autre exemple de la caractérisation, à l'aide de lignées congéniques, de facteurs génétiques impliqués dans la sensibilité à une maladie infectieuse chez la Souris est illustré par le travail du groupe de Danielle Malo sur *Salmonella typhimurium*. Des approches QTL ont permis d'identifier 8 locus qui affectent la réponse de l'hôte dans un modèle de fièvre typhoïde (*Ity2* à *Ity8* et *Ity13*) (Khan *et al.*, 2012; Roy *et al.*, 2006; Sancho-Shimizu *et al.*, 2007) et 10 locus corrélés à la persistance de la bactérie (*Ses1* et *Ses1.1* puis *Ses3* à *Ses10*) (Caron *et al.*, 2002; Caron *et al.*, 2005).

L'étude d'une population F2 résultant du croisement des lignées C57BL/6J (sensible) et MOLF/Ei (résistante) a permis d'identifier trois locus impliqués dans la sensibilité à une infection par *S. typhimurium* (*Ity*, *Ity2* et *Ity3*) (Sebastiani *et al.*, 1998). Le premier locus *Ity* dont la base moléculaire est le gène *Slc11a1*, anciennement *Nramp1*, avait déjà été identifié dans des études antérieures (Vidal *et al.*, 1995). Des lignées congéniques portant les trois locus indépendamment, sur un fonds génétique C57BL/6J, ont permis de confirmer l'existence de ces locus et leur mode d'action : l'allèle MOLF/Ei conférait un effet protecteur aux locus *Ity* et *Ity2* sur un mode dominant alors qu'il augmentait la sensibilité de manière récessive au locus *Ity3* (Sebastiani *et al.*, 1998).

L'analyse de deux lignées sous-congéniques, l'une portant une partie proximale et l'autre une partie distale d'*Ity2* (en présence de l'allèle MOLF/Ei au locus *Nramp1* correspondant à *Ity*) ont permis de montrer que ce QTL était en réalité composé de deux locus distincts sur le chromosome 11 (Sancho-Shimizu *et al.*, 2007). Une analyse transcriptomique pan-génomique de la lignée bi-congénique portant *Ity* et *Ity2* (résistante) comparée à la lignée portant *Ity* et la partie distale d'*Ity2* (sensible) ont permis de mettre en évidence le rôle de la voie de signalisation de l'interféron de type I et de la voie TRP53 comme impliquées dans la pathogénie de l'infection par *S. typhimurium* (Khan *et al.*, 2012).

La création de lignées congéniques qui portent l'allèle MOLF/Ei aux locus *Ity* et *Ity3*, a également confirmé la sensibilité conférée par cet allèle au locus *Ity3*. L'analyse de quatre lignées sous-congéniques a permis d'affiner l'intervalle d'*Ity3*, d'une taille supérieure à 100 Mb à un intervalle de 34 Mb contenant environ 322 gènes (Sancho-Shimizu *et al.*, 2007). Une approche gène candidat a montré que *Ncf2* (*Neutrophil cytosol factor 2*), qui participe à la réponse immune innée, est un candidat majeur d'*Ity3* (Sancho-Shimizu *et al.*, 2006). Le séquençage de ce gène a révélé la modification d'un codon provoquant un changement d'acide aminé dans un domaine fonctionnel très conservé de la protéine. De plus l'expression de *Ncf2* est réduite chez les bi-congéniques sensibles portant *Ity* et *Ity3* par rapport aux congéniques portant seulement *Ity*. Des analyses fonctionnelles ont montrées que les macrophages péritonéaux de ces souris produisaient des quantités significativement

plus faibles de superoxyde (qui possède une activité anti-microbienne) en réponse à l'infection *in vitro* par *S. typhimurium* (Sancho-Shimizu *et al.*, 2006).

La dissection fine des locus *Ity2* et *Ity3* qui a été réalisée illustre la puissance de l'utilisation des lignées congéniques pour mettre en évidence l'architecture génétique d'un caractère étudié et d'affiner la localisation des QTL.

Un autre exemple de dissection d'un caractère complexe à l'aide de lignées congéniques est le travail du groupe de J.F. Bureau qui porte sur la sensibilité à la maladie neurodégénérative induite par le virus de Theiler, qui est considérée comme un modèle pour l'étude de la physiopathologie de la sclérose en plaques. Les différences de sensibilité sont principalement liées au contrôle de la charge virale par le système immunitaire au cours de l'infection persistante. Un premier locus incluant la région *H2-D* du CMH a été impliqué dans la variation de sensibilité de différentes lignées de souris au virus de Theiler (Brahic *et al.*, 1998). Néanmoins, les lignées SJL/J et B10.S portent le même allèle de résistance intermédiaire au locus *H2-D*, mais présentent respectivement une forte sensibilité et une sensibilité intermédiaire, révélant que d'autres locus sont impliqués dans la résistance à ce virus (Bihl *et al.*, 1999). Un *backcross* entre ces lignées a mis en évidence deux QTL sur les chromosomes 10 et 18 (Bihl *et al.*, 1999). Quatre lignées congéniques portant séparément différentes parties du QTL localisé sur le chromosome 18 sous forme B10.S ont été créées mais n'ont pas montré de différence significative par rapport à la lignée receveuse SJL/J en ce qui concerne la quantité de virus dans la moelle épinière (Bihl *et al.*, 1999). Une lignée congénique réciproque portant le même QTL, dans sa totalité, mais sous forme SJL/J dans un fond B10.S n'a pas non plus montré de différence par rapport à la lignée receveuse. Les auteurs ont conclu à la détection d'une fausse association pour ce QTL (Bihl *et al.*, 1999).

En revanche, une lignée congénique portant l'ensemble du QTL identifié sur le chromosome 10 sous forme B10.S présentait une différence de sensibilité significative par rapport à la lignée receveuse SJL/J. L'étude de deux lignées sous-congéniques qui ne se chevauchaient pas a montré que le phénotype de résistance était conservé chez ces deux lignées, indiquant que le QTL identifié dans le *backcross* était en réalité constitué de deux locus distincts, alors dénommés *Tmevp2* et *Tmevp3*. La dissection du locus *Tmevp3* à l'aide de 6 lignées sous-congéniques portant chacune une fraction de ce locus a mené à la réduction de la région (Bihl *et al.*, 1999). La cartographie fine de *Tmevp3* a permis de réduire l'intervalle à 4,1 Mb, ne contenant plus que trois gènes candidats : *Il22*, *Ifng* et *Tmevp3* (Levillayer *et al.*, 2007). Aucun polymorphisme fonctionnel n'a été trouvé dans ces gènes, ce qui suggère que *Tmevp3* en régulerait l'expression (Levillayer *et al.*, 2007). *Tmevp3*, un gène codant un ARN non-codant exprimé dans les lymphocytes T faisait figure de bon candidat pour le locus *Tmevp3* (Vigneau *et al.*, 2003). Cette étude a mené à considérer le rôle du locus humain homologue de *Tmevp3* comme candidat pour le déterminisme génétique de la sclérose en plaques (Vigneau *et al.*, 2003).

Ce travail illustre, là encore, l'utilité des lignées congéniques pour disséquer un QTL constitué de plusieurs locus. Il montre également que certains QTL détectés dans un croisement en deux générations ne sont pas confirmés par les lignées congéniques. Cet exemple sera discuté à la lumière de nos propres résultats. En effet, les auteurs ont créé des lignées congéniques réciproques et n'ont pas détecté d'effet (Bihl *et al.*, 1999). Ils en concluent que le QTL détecté dans le *backcross* est une fausse association. Cette étude suffit-elle à évincer ce QTL ? L'existence d'interactions additives ou épistatiques qui pourraient être nécessaires à l'expression de ce QTL (uniquement testé de manière isolé) a-t-elle pu mener à des conclusions prématurées ? Nous discuterons également de l'intérêt de croiser les lignées congéniques avec la lignée donneuse afin d'augmenter les chances de voir l'effet d'un QTL.

Les lignées congéniques permettent également d'identifier des QTL à effet faible. Lorsqu'un QTL à effet majeur, qui explique une forte proportion de la variance phénotypique, ségrège dans un croisement, il est souvent difficile d'identifier d'autres QTL dont les effets sont plus faibles. Si deux lignées consanguines A et B diffèrent pour un QTL à effet majeur, et que l'on souhaite rechercher d'autres QTL à effet plus faible différenciant ces deux lignées il suffit d'établir une lignée congénique (le QTL sous la forme allélique parentale A est transféré dans le fonds génétique B, par exemple) et de croiser ensuite cette lignée congénique avec la lignée parentale A (Figure 15). On obtient ainsi des F1, hybrides A/B sur l'ensemble du génome excepté au locus du QTL pour lequel les individus sont B/B. Les individus F1 sont croisés entre eux pour établir une F2 (ou une N2) qui sera alors analysée avec une probabilité supérieure de détecter des QTL à effet faible, l'effet du QTL majeur étant neutralisé puisque tous les individus sont identiques à ce locus (Panthier *et al.*, 2003).

Un exemple d'utilisation des lignées congéniques pour détecter des QTL à effet faible sont les études de Yan *et al.* et de Sissons *et al.* sur les facteurs de sensibilité à *Mycobacterium tuberculosis* (Sissons *et al.*, 2009; Yan *et al.*, 2006). Un locus sur le chromosome 1, *Sst1*, avait été identifié antérieurement comme impliqué dans l'hyper-sensibilité de la lignée C3HeB/FeJ à *M. tuberculosis* (Kramnik *et al.*, 2000). Des lignées congéniques réciproques portant l'allèle de sensibilité *Sst1*^{C3H} sur fonds C57BL/6 ou l'allèle de résistance *Sst1*^{B6} sur fonds C3H (B6.C3H-*Sst1* et C3H.B6-*Sst1*) présentaient un phénotype intermédiaire entre les deux lignées parentales, démontrant que *Sst1* n'était pas le seul locus contrôlant le phénotype. La production d'une population F2 entre la lignée congénique C3H.B6-*Sst1* et la lignée parentale C57BL/6J, dont tous les individus étaient homozygotes pour l'allèle B6 au locus *Sst1*, a permis d'identifier quatre autres QTL sur les chromosomes 7, 12, 15 et 17 (Yan *et al.*, 2006). Des croisements multiples ont permis de confirmer et d'affiner le QTL présent sur le chromosome 7 (Sissons *et al.*, 2009). La création d'une lignée consanguine portant le chromosome 7 de C57BL/6J dans un fonds sensible C3HeB/FeJ a montré un effet individuel assez

faible de ce QTL. Le croisement de congéniques C3H.B6-*SstI* avec cette lignée consomique, pour créer des « bi-congéniques » a permis de mettre en évidence des interactions additives de ces deux locus, qui une fois combinés confèrent une très forte protection. *SstI* agit à un stade précoce en empêchant la nécrose des granulomes pulmonaires, alors que le QTL localisé sur le chromosome 7 contrôle la multiplication intracellulaire de *M. tuberculosis* dans les lésions inflammatoires à un stade tardif de l'infection (Sissons *et al.*, 2009). Des lignées sous-congéniques portant *SstI*, sous sa forme résistante, et différents fragments chevauchants du chromosome 7 ont permis de déterminer un intervalle de 8,7 cM sur le chromosome 7 qui porte les facteurs génétiques impliqués dans la sensibilité de la lignée C3HeB/FeJ (Sissons *et al.*, 2009).

Cette étude montre que ces deux locus, identifiés tous deux dans la population F2 initiale comme influençant la durée de survie, agissent à des étapes différentes de l'infection et n'ont pas d'effet sur la durée de survie une fois isolés dans des lignées congéniques. Cette étude illustre l'importance d'identifier des sous phénotypes lors de l'étude du processus infectieux et les limites d'un phénotype aussi global que la survie. Il illustre aussi l'intérêt de combiner différentes lignées congéniques pour décortiquer le rôle de chaque locus dans le déroulement du processus infectieux.

Les lignées congéniques sont un outil puissant pour confirmer l'existence des QTL, étudier leurs effets dans un fonds génétique donné, et comprendre l'architecture génétique d'un caractère complexe et les interactions qui existent entre les locus responsables. Néanmoins, elles couvrent en générale des régions chromosomiques étendues. Des approches complémentaires sont nécessaires pour réduire le nombre de gènes candidats encore en jeu et faire des tests fonctionnels (Figure 16). L'utilisation des données de séquençage du génome complet des lignées parentales permet d'identifier, entre autres, les polymorphismes présents dans la région congénique. Les gènes dont la séquence codante ou les promoteurs sont mutés entre les deux lignées parentales sont ciblés comme des candidats prioritaires. Par exemple, des analyses bioinformatiques permettent d'évaluer l'impact théorique des mutations sur la fonction biologique d'un gène et permettent de hiérarchiser les gènes candidats au sein de l'intervalle congénique. Lorsque les lignées parentales sont très divergentes, il peut être difficile, dans la multitude des polymorphismes identifiés, de détecter les mutations ayant des chances d'être causales. Toutefois, des tests fonctionnels restent nécessaires pour valider l'effet des polymorphismes identifiés.

6. Comment aller jusqu'au gène ?

La découverte des facteurs génétiques sous-jacents à un QTL repose généralement sur une combinaison d'approches. L'analyse de l'expression des gènes au niveau transcriptionnel et/ou protéique, la comparaison de séquence, l'analyse combinée de plusieurs croisements, l'analyse

d'haplotypes, la génomique comparative permettent soit de réduire drastiquement la taille des régions chromosomique identifiées soit de sélectionner des gènes candidats à l'intérieur de la région. L'analyse des candidats repose ensuite sur des tests fonctionnels *in vitro* et sur la caractérisation fonctionnelle de systèmes génétiques qui permettent d'inactiver (*Knock-out*), de réduire l'expression (*Knock-down*, ARN interférence), de remplacer des allèles (*Knock-in*) ou de sur-exprimer (transgéniques) différents gènes (Clee *et al.*, 2006; DiPetrillo *et al.*, 2005; Rogner *et al.*, 2003; Sullivan *et al.*, 1997).

L'étude d'Ahn *et al.* précédemment citée comme exemple d'utilisation des lignées consomiques est aussi un exemple d'étude intégrant plusieurs outils à la disposition des généticiens pour disséquer des QTL jusqu'à mettre en évidence les gènes responsables (Figure 17). Dans cette étude Ahn *et al.* ont établi, en analysant des lignées consomiques, que les chromosomes 8, 11 et 18 portaient des gènes de sensibilité à *S. aureus* (Ahn *et al.*, 2010). Ils ont ensuite comparé le profil transcriptionnel des souris A/J et C57BL/6J infectées et non-infectées, en analysant l'ARNm extrait du sang. Ils ont mis en évidence que 191 gènes étaient différentiellement exprimés suite à l'infection dans la lignée A/J, dont respectivement 56, 100 et 35 étaient localisés sur les chromosomes 8, 11 et 18 (Ahn *et al.*, 2010). Ils ont ensuite disséqué le QTL du chromosome 18. Un *backcross* de la lignée CSS18 (lignée consomique portant le chromosome 18 de la lignée sensible A/J sur un fonds C57BL/6J résistant) par la lignée C57BL/6J a mis en évidence deux QTL sur ce chromosome associés au temps de survie. Ces deux locus de 13 et 10 Mb, respectivement, contenaient à eux deux environ 317 gènes. Dix gènes candidats différentiellement régulés entre les lignées A/J et C57BL/6J en réponse à l'infection, et localisés dans les deux locus, ont été retenus. Parmi ces 10 candidats, 6 ont été confirmés comme significativement différentiellement régulés par RT-PCR quantitative. Pour les 10 candidats, une expérience de « knock-down » par *siRNA* (*short interference RNA*) dans un système cellulaire *in vitro* a pu évaluer l'effet d'une sous-expression de ces gènes sur la production de 32 cytokines impliquées dans la réponse pro-inflammatoire. La sous-expression de seulement deux gènes, *Tnfaip8* et *Seh1l*, avait un impact sur la production cytokinique (Ahn *et al.*, 2010). La sous-expression de *Tnfaip8* réduisait la production d'IL1 β et la sous-expression de *Seh1l* augmentait la production de GM-CSF. Cette observation *in vitro* a été confirmée *in vivo*, les souris A/J présentant une expression diminuée des gènes *Tnfaip8* et *Seh1l* dans le sang. De plus, les macrophages péritonéaux de la lignée CSS18 produisaient moins d'IL1 β et plus de GM-CSF suite à l'exposition *in vitro* à *S. aureus* que les macrophages de C57BL/6J, confirmant les observations de « knock-down » par *siRNA* (Ahn *et al.*, 2010). Si un test fonctionnel *in vivo* (par transgénèse ou *knock-in*) manque à cette étude pour confirmer formellement l'implication de ces deux gènes dans le phénotype de sensibilité de la lignée A/J à *S. aureus*. Cette approche combinée utilisant plusieurs outils de génétique et de biologie moléculaire est un exemple intéressant de la dissection fine d'un caractère hautement complexe. Ainsi, à partir d'environ 4200 gènes présents sur les trois chromosomes de départ, Ahn *et al.* ont identifié

deux gènes qui ont de fortes chances d'être associés à la sensibilité à une infection par *S. aureus* (Ahn *et al.*, 2010).

Un autre exemple d'approche globale est celui de la découverte des gènes *Ncf2* et *Pklr* dans la sensibilité à une infection par *Salmonella typhimurium*.

Différents croisements ont permis de mettre en évidence plus d'une dizaine de locus dans la sensibilité à une infection par *S. typhimurium*. Parmi ces locus, *Ity3* a pu être restreint, grâce à l'utilisation de lignées congéniques, à un intervalle de 84 Mb où le gène *Ncf2* faisait figure d'excellent candidat. Une approche combinant une comparaison de la séquence de *Ncf2* entre plusieurs lignées consanguines résistantes et sensibles à *S. typhimurium*, une analyse de l'expression de *Ncf2* dans les organes cibles et des tests de complémentation fonctionnelle ont permis de valider *Ncf2* comme gène causal au locus *Ity3* (Sancho-Shimizu *et al.*, 2006).

Le locus *Ity4* a également été impliqué dans la différence de sensibilité à *S. typhimurium*. Le pic de LOD score associé à ce QTL était localisé au niveau du gène *Pklr*. La lignée congénique recombinante AcB61 qui avait été utilisée dans le croisement pour détecter *Ity4* est connue pour porter une mutation dans le gène *Pklr* qui a un effet protecteur contre *Plasmodium chabaudi* mais une entraîne une sensibilité accrue à une infection par *S. typhimurium* (Roy *et al.*, 2007). En utilisant une approche de cartographie fine combinée à des tests de complémentation, des études comparatives de séquence et une analyse fonctionnelle, Roy *et al.* ont démontré que *Pklr* était bien le gène causal de *Ity4* (Roy *et al.*, 2007).

Un autre exemple d'approche combinant différentes techniques pour réduire une région génomique de plusieurs dizaines de mégabases à quelques gènes et l'étude de Radovanovic *et al.* sur *Candida albicans* précédemment citée (Radovanovic *et al.*, 2011). Cette étude a identifié plusieurs locus à l'aide d'une analyse d'association haplotypique sur 23 lignées consanguines. Les auteurs ont ensuite sélectionné deux lignées qui présentaient des phénotypes très différents, AKR/J résistante et A/J sensible, pour rechercher des QTL. Ces deux lignées présentaient un allèle déficient pour le gène *C5* (antérieurement mis en évidence pour influencer la sensibilité à une infection par *C. albicans*). A l'aide d'un croisement de type F2 entre ces lignées, ils ont mis en évidence deux locus *Carg3* et *Carg4* indépendants du gène *C5*, d'une taille respective de 40 et 30 Mb (Radovanovic *et al.*, 2011). Afin d'affiner ces régions, ils ont combiné les données d'analyse de liaison (QTL) avec les résultats d'HAM. Ils ont ciblé pour chaque locus un intervalle de 10 Mb autour du SNP qui présentaient la valeur d'association la plus forte dans l'analyse d'haplotype. Cette analyse leur a permis de réduire le nombre de gènes à 59 et 147 pour les deux locus. Une caractéristique de la lignée résistante étudiée

dans ce croisement, AKR/J, est une production importante d'IFN γ qui pourrait mener à une activation accrue de la voie Jak-STAT. Radovanovic *et al.* ont sélectionné dans la liste de gènes candidats des deux QTL ceux dont l'expression était modulée par l'IFN γ (Seifert *et al.*, 2010) et ceux susceptibles de se lier à STAT-1 (Robertson *et al.*, 2007). La liste de gènes candidats dans *Carg3* et *Carg4* a ainsi été réduite à 4 et 11 gènes respectivement (Radovanovic *et al.*, 2011). Cette sélection a permis d'obtenir une liste raisonnable de candidats qui pourront être testés fonctionnellement.

7. Bilan : qu'est-ce qui pose encore problème pour aller plus vite ?

Le nombre de gènes qui ont été identifiés à la suite de la découverte d'un QTL reste restreint (Roff, 2007). En 2005, parmi les 2000 QTL qui avaient été découverts chez les rongeurs et chez l'humain, les gènes en cause étaient connus pour moins d'un pourcent seulement (DiPetrillo *et al.*, 2005; Hunter *et al.*, 2008). L'obstacle majeur de l'approche QTL reste la réduction du QTL à quelques candidats qui pourront être testés fonctionnellement (DiPetrillo *et al.*, 2005). Une approche QTL efficace nécessite une haute résolution de cartographie, ce qui fait encore fréquemment défaut. De nombreux QTL détectés par des croisements en deux générations couvrent de grandes régions génomiques (souvent supérieures à 20 cM, soit environ 40 Mb) qui contiennent de nombreux locus (Roff, 2007). Les autres facteurs qui ralentissent la découverte des facteurs génétiques sous-jacents aux QTL sont (1) la faiblesse des effets des locus individuels sur le phénotype observé, (2) l'absence de sous-phénotypes connus pour caractériser l'effet d'un locus, (3) l'existence d'interactions épistatiques entre les locus et (4) la complexité de chaque QTL dont l'effet peut être dû à plusieurs gènes liés (Fortin *et al.*, 2007; Vidal *et al.*, 2008). La réduction de ces régions chromosomiques est longue et coûteuse et est fréquemment un frein à l'identification de gènes (Price, 2006). Aujourd'hui, la disponibilité des lignées RI, RC et de plusieurs centaines de lignées du *Collaborative Cross* ajoutée à la disponibilité de séquences génomiques complètes et de données de génotypage très denses sont autant d'outils qui permettent de réduire l'intervalle des QTL (Hunter *et al.*, 2008). L'utilisation de populations de référence comme les RI et RC ont l'avantage de pouvoir tester plusieurs animaux de chaque lignée pour mesurer un phénotype. Ces réplicats biologiques qui ne peuvent pas être effectués dans des populations issues de croisements en deux générations sont une source importante de puissance statistique pour l'approche QTL (Hunter *et al.*, 2008). La résolution de cartographie peut aussi être augmentée par l'utilisation du *Diversity Outbred* (DO) *stock*. Cette population a été développée à partir du croisement randomisé de 160 lignées du *CC* et est maintenue non-consanguine par des croisements entre individus non apparentés (Mott *et al.*, 2002). Cette population présente une grande diversité génétique qui permet de se rapprocher des populations humaines et d'obtenir une résolution de cartographie des QTL bien supérieure aux autres croisements en raison de l'accumulation de nombreux événements de recombinaison au fil des générations (Hunter *et al.*, 2008).

Augmenter le ratio du nombre de gènes détectés par rapport au nombre de QTL identifiés nécessite aussi de recouper les informations et les études (Hubner *et al.*, 2005). Par exemple, le croisement d'une analyse QTL avec des études d'expression génique permet de réduire la liste de gènes candidats rendant possibles les études fonctionnelles. L'utilisation de bases de données publiques peut aussi s'avérer utile pour trouver les gènes communs qui ressortent en recoupant les études portant sur le même sujet mais sur des populations génétiques de référence différentes ou provenant d'un environnement différent (Wayne *et al.*, 2002). Les données de QTL obtenues *in vivo* sur des lignées ou *in vitro* sur un type cellulaire dans un environnement contrôlé, peuvent être utilisées pour orienter d'autres types d'études telles que celles menées sur des populations naturelles (Stinchcombe *et al.*, 2008).

Seule, la stratégie consistant à identifier des QTL dans des croisements, à les confirmer, à en comprendre l'architecture et les interactions à l'aide de lignées congéniques, ne permet pas de réduire suffisamment le nombre de gènes candidats pour pouvoir effectuer des tests fonctionnels sur chacun des gènes potentiellement en cause. Des observations sur la localisation et l'effet des SNP dans les QTL, suggèrent que la majorité des QTL ont plus de chances d'être dus à des variations subtiles du niveau d'expression des gènes qu'à des mutations non-synonymes (ou *nonsense*) (Hunter *et al.*, 2008). Cette hypothèse met en exergue l'importance des études d'expression pangénomiques dans l'élucidation des mécanismes moléculaires contrôlant les caractères complexes.

Partie 5. L'apport des analyses transcriptomiques à la compréhension des interactions hôte-pathogène chez la Souris

La rencontre entre un agent infectieux et son hôte déclenche de part et d'autre une multitude d'événements moléculaires et physiologiques complexes. L'hôte détecte la présence du pathogène et mobilise des mécanismes de défense non-spécifiques puis spécifiques. Le pathogène à son tour peut modifier sa propre physiologie pour s'adapter au microenvironnement de l'hôte, et perturber les mécanismes de défense de l'hôte afin de pouvoir survivre et se multiplier. L'étude du statut transcriptionnel de l'hôte et/ou du pathogène permet d'identifier les mécanismes moléculaires mis en place lors d'une infection. Les analyses transcriptomiques permettent d'avoir accès simultanément au niveau d'expression de milliers de gènes dans un tissu ou dans une population cellulaire, *in vitro*, *ex vivo* ou *in vivo* (Waddell *et al.*, 2007).

Nous nous pencherons tout d'abord sur l'apport des analyses transcriptomiques dans des modèles de maladies infectieuses avant de s'intéresser à l'approche combinée de cartographie des QTL et analyse du transcriptome. Enfin nous étudierons plus spécifiquement les études transcriptionnelles portant sur l'interaction entre *Yersinia pestis* et son hôte.

1. Pourquoi étudier le transcriptome ?

Si la séquence primaire d'ADN est fixe dans l'ensemble des cellules d'un organisme⁶, son expression, par contre, varie au cours du temps, suivant le type cellulaire et en fonction du contexte physiologique. Le transcriptome représente l'ensemble des molécules d'acide ribonucléique (ARN) incluant les ARN messagers, ribosomaux, de transfert et autres ARN non-codants produits par une cellule ou une population de cellules. Par extension, on appellera « transcriptome » les ARN extraits à partir d'un type cellulaire ou d'un tissu.

Plusieurs techniques sont utilisées pour mesurer simultanément le niveau d'expression de l'ensemble des ARN produits par une cellule. Les techniques seront seulement citées ici mais pas détaillées.

Les études pangénomiques font principalement appel à deux techniques. Les puces à ADN (*DNA microarrays*), utilisées dans notre travail, sont des supports solides sur lesquelles sont fixées des

⁶ A quelques exceptions près, notamment les cellules immunitaires qui subissent des réarrangements profonds des gènes d'immunoglobulines ou des TCR.

sondes complémentaires auxquelles viennent s'hybrider les ADN complémentaires (ADNc) synthétisés à partir des ARN étudiés par rétro-transcription. Dans la technique de *RNA-seq*, les ADNc sont séquencés à très haut débit et les séquences obtenues sont analysées pour identifier les gènes auxquels elles correspondent et pour mesurer l'abondance de chaque transcrit.

La RT-PCR quantitative permet également de quantifier l'expression des gènes, mais pour un nombre limité de gènes (quelques centaines au maximum). Elle n'est donc pas une méthode d'analyse pangénomique mais elle est souvent utilisée pour valider les données issues des études pangénomiques. Nous nous intéresserons ici à l'étude du transcriptome à l'aide de la technologie par puces à ADN.

Ce sont les variations de l'état de transcription du génome que cherche à mettre en évidence l'analyse transcriptomique. Elle consiste généralement à comparer deux états du transcriptome pour en identifier les changements. La variation de l'expression des gènes est mesurée par rapport à une condition définie ou par rapport à des gènes définis qui dont le niveau d'expression est invariant. Le choix des gènes témoins et de la condition de référence, ou témoin, est une étape cruciale dans l'analyse des variations du transcriptome (Sturdevant *et al.*, 2010).

L'ensemble des analyses statistiques qui permettent de générer les résultats d'expression à partir des données brutes ne seront pas détaillés ici, tout d'abord parce qu'elles sont très variables d'une étude à l'autre, ensuite parce qu'une méthode d'analyse, celle que nous avons utilisée, sera détaillée dans la partie « Matériels et méthodes ».

Nous nous focaliserons ici sur l'étude de la réponse de l'hôte durant la progression de l'infection, mais le transcriptome du pathogène peut aussi être quantifié au cours de l'infection par les techniques de puce à ADN ou *RNA seq*. Connaître la réponse transcriptionnelle d'un pathogène offre une vision du statut « physiologique » du pathogène et des mécanismes que celui-ci met en place pour survivre aux défenses de l'hôte (Waddell *et al.*, 2007). En 2010, très peu de publications avaient analysé simultanément les transcriptomes de l'hôte et du pathogène au cours d'une infection. De telles études ont l'avantage de suivre la co-évolution de l'hôte et du pathogène et de pouvoir déceler les interactions spécifiques qui s'établissent lors d'une infection. Ces études de co-expression permettent d'obtenir une compréhension globale du système d'interaction hôte-pathogène (Liu *et al.*, 2009; Lovegrove *et al.*, 2006; Motley *et al.*, 2004).

2. L'apport des études transcriptomiques *in vitro*

La plupart des études s'intéressent aux changements d'expression des gènes qui se produisent dans un type cellulaire après contact avec un pathogène spécifique *in vitro*. Ces études suivent différents protocoles, qui donnent chacun un aspect spécifique des interactions hôte-pathogène. Certaines études utilisent une souche sauvage du pathogène et quantifient les variations du transcriptome à différents temps post-infection (Rodriguez *et al.*, 2004; Shapira *et al.*, 2009). D'autres comparent l'effet de différents pathogènes sur le programme transcriptionnel de cellules identiques, permettant de distinguer la réponse commune de l'hôte à un pathogène, de la réponse spécifique à chaque pathogène ou type de pathogène (Boldrick *et al.*, 2002; Chaussabel *et al.*, 2003; Huang *et al.*, 2001; Nau *et al.*, 2002). L'exposition de cellules à des pathogènes mutés, en particulier déficients pour un facteur de virulence ou un composant spécifique, permet de mettre en évidence comment chaque composant du pathogène agit sur la réponse de l'hôte (Detweiler *et al.*, 2001; Grassl *et al.*, 2003; Hoffmann *et al.*, 2004; Sauvonnnet *et al.*, 2002). L'exposition à des pathogènes tués, inactivés ou avirulents, comparés à une souche virulente vivante, permet de distinguer l'effet de l'activité du pathogène sur l'expression génique de l'hôte et la réponse de l'hôte provoquée par le seul contact avec le pathogène (Boldrick *et al.*, 2002; Huang *et al.*, 2001; Wang *et al.*, 2011a). Une alternative est l'exposition à des protéines ou des composants isolés du pathogène (Huang *et al.*, 2001; Nau *et al.*, 2002; Shapira *et al.*, 2009). La mesure de l'activité transcriptionnelle suite à un traitement à l'aide de petites molécules (IFN γ par exemple) permet d'identifier les voies cellulaires spécifiques de la réponse immune de l'hôte (Nau *et al.*, 2003; Scherbik *et al.*, 2007; Shapira *et al.*, 2009).

L'utilisation de cellules mutées ou d'un modèle animal déficient pour une protéine ou pour un type cellulaire aide à identifier les mécanismes moléculaires impliqués dans la réponse de l'hôte au pathogène (Probst *et al.*, 2012). De même, la comparaison du transcriptome de lignées de souris présentant une sensibilité différente à un pathogène est utilisé pour mettre en évidence les mécanismes différentiellement activés ou réprimés entre des lignées sensibles ou résistantes (do Valle *et al.*, 2010; Probst *et al.*, 2012). Des études combinant différentes approches telles que protéomique, transcriptomique, ou d'inactivation par ARN interférence permettent d'avoir une approche globale, à différents niveaux, de l'interaction entre un hôte et un pathogène (Goldmann *et al.*, 2007; Shapira *et al.*, 2009).

L'étude comparée du transcriptome issu d'une population cellulaire *in vitro* provenant de lignées de souris présentant une variation de sensibilité à un pathogène est la méthode d'étude que nous avons choisie. C'est aussi celle qui avait été réalisée dans l'étude de do Valle *et al.* qui comparait le transcriptome de MEF (*Mouse Embryonic Fibroblasts*) de souris BALB/cByJ modérément sensibles

à une infection par le virus de la fièvre de la Vallée du Rift et de souris MBT/Pas extrêmement sensibles à ce virus. A la suite de leur infection, les MEF issus de MBT/Pas produisent des titres viraux supérieurs aux MEF de BALB/cByJ. Une analyse transcriptomique a montré que les MEF de BALB/cByJ présentaient une activation de la voie interféron de type I, alors que les MEF de MBT/Pas montraient une réponse interféron incomplète et retardée (do Valle *et al.*, 2010). Des gènes tels que *Irf7*, *Isg15* ou *Oasl2*, caractéristiques de la réponse antivirale étaient induits rapidement dans les MEF de BALB/cByJ mais très tardivement chez MBT/Pas (do Valle *et al.*, 2010). Pour valider l'hypothèse que la différence de titre viral observé entre les deux lignées de souris était corrélée à cette déficience de la voie interféron chez MBT/Pas, ces gènes ont été inactivés par ARN interférence dans des MEF issus de BALB/cByJ. L'inhibition de ces gènes entraînait effectivement une augmentation de la production virale dans les cellules (do Valle *et al.*, 2010). Les auteurs concluaient que l'incapacité des souris MBT/Pas à produire une réponse interféron complète et précoce contribuait à la sensibilité extrême de cette lignée à une infection par le virus de la fièvre de la Vallée du Rift (do Valle *et al.*, 2010).

Cette étude illustre l'utilité de l'analyse du transcriptome *in vitro* pour comprendre les mécanismes moléculaires qui sont responsables des différents types de réponse de l'hôte.

La comparaison du transcriptome de l'hôte exposé à un pathogène virulent (noté WT pour *wild type*) ou à des mutants atténués de ce pathogène permet de comprendre le rôle et les mécanismes d'action des facteurs de virulence. Par exemple, *Yersinia enterocolitica* possède un plasmide, nommé pYV, qui code le *Yop virulon*, lequel comprend le système de sécrétion de type III et les effecteurs de la bactérie, les Yop (pour *Yersinia outer protein*, ces facteurs de virulence seront détaillés dans la partie 6). Le rôle du *Yop virulon* a été étudié par l'analyse transcriptomique de macrophages murins exposés à *Y. enterocolitica* WT, au mutant pYV⁻ délété pour le plasmide pYV, au mutant YopP⁻ délété pour l'effecteur YopP et au mutant YopM⁻ délété pour l'effecteur YopM (Sauvonnet *et al.*, 2002). L'altération du transcriptome a été mesurée sur 6657 gènes. En comparant le profil d'expression des macrophages infectés par la bactérie WT ou avirulente pYV⁻, Sauvonnet *et al.* ont mis en évidence que l'expression de 857 gènes étaient affectés par l'infection et 339 spécifiquement par le *Yop virulon*. L'analyse de ces 339 gènes, essentiellement réprimés, a montré qu'il s'agissait principalement de gènes impliqués dans la réponse immune. Ils en concluent qu'un des effets du *Yop virulon* est de contrecarrer la réponse inflammatoire mise en place par la cellule. Le transcriptome des macrophages infectés par les mutants YopP⁻ et YopM⁻ a permis de définir les cibles spécifiques des deux effecteurs YopP et YopM parmi les 339 gènes réprimés par le plasmide pYV (Sauvonnet *et al.*, 2002). Cinquante-neuf gènes étaient réprimés par l'effecteur YopP dans les macrophages dont *Batf*, *Junb*, *Nfkb1a*, *Tnf*, *Tnfaip2*, *Tnfaip3* et *Trafi1* qui sont des gènes communément activés par l'hôte lors d'une infection (Jenner *et al.*, 2005; Sauvonnet *et al.*, 2002). L'analyse des gènes ciblés par YopM a permis

de montrer que cet effecteur participe à la régulation du cycle cellulaire et de la croissance cellulaire (Sauvonnet *et al.*, 2002).

La répression de la réponse inflammatoire induite par l'effecteur YopP de *Y. enterocolitica* a aussi été mise en évidence par Bohn *et al.* Ils ont montré qu'un mutant déficient pour YopP induit l'activation des gènes impliqués dans la réponse pro-inflammatoire, tels que *Atf3*, *Ccl20*, *Cxcl1*, *Cxcl2*, *Cxcl3*, *Icam1*, *Ier3*, *Il1a*, *Il8*, *Nfkb1a* et *Ptgs2*, dans des cellules épithéliales intestinales (Bohn *et al.*, 2004). Comme Sauvonnet *et al.*, Bohn *et al.* concluent que *Y. enterocolitica* est capable de réprimer l'expression de nombreux gènes afin de déjouer la réponse immune innée mise en place dans les cellules épithéliales (Bohn *et al.*, 2004).

Les résultats de ces deux études ont été corroborés par ceux de Hoffmann *et al.*, qui ont réalisé une autre étude transcriptomique sur la réponse de macrophages à différentes souches de *Y. enterocolitica* WT et mutantes (Hoffmann *et al.*, 2004). Cette étude leur a permis d'identifier deux mécanismes indépendants mis en place par la bactérie pour détourner de la réponse immune. Le premier est la suppression de l'expression des gènes impliqués dans la réponse inflammatoire. Le deuxième est l'induction de gènes ayant une fonction de « *silencing* ». Ils ont confirmé que YopP était le facteur principal qui provoque la suppression de la réponse inflammatoire en réprimant les voies NF- κ B et MAP kinase (Hoffmann *et al.*, 2004).

La comparaison du profil transcriptionnel d'un hôte soumis à un pathogène WT ou à une souche atténuée ou déficiente pour un facteur de virulence fournit des informations clés sur les interactions hôte-pathogène et permet de faire le lien avec la pathogénèse de l'infection. Un profil transcriptionnel inhibé en réponse à une souche virulente par rapport à une souche atténuée suggère un mode de pathogénicité basé sur l'inhibition de la réponse immune (Hossain *et al.*, 2006; Jenner *et al.*, 2005). A l'inverse, une réponse transcriptionnelle augmentée suggère que la pathogénicité réside dans la capacité du pathogène à stimuler une réponse immunitaire excessive ou inappropriée (Jenner *et al.*, 2005).

De plus, des études transcriptomiques *in vitro* ont montré que les micro-ARN, petits ARN de 21 à 25 nucléotides, jouaient un rôle majeur dans les processus infectieux. Plusieurs études ont montré des variations du niveau d'expression de micro-ARN suite à une infection. C'est le cas de l'article de Schulte *et al.*, qui étudie l'expression, par *RNA-seq*, des micro-ARN suite à une infection par *Salmonella enterica* dans des macrophages murins et humains. Cette étude révèle que *S. enterica* induit, en quelques heures, la production de plusieurs micro-ARN tels que miR-155, miR146a et miR-21, dans des macrophages murins infectés (Schulte *et al.*, 2012). Ces micro-ARN ont déjà été décrits comme étant sur-exprimés par l'activation des TLR (O'Neill *et al.*, 2011) et sont connus pour être transcrits en réponse à une infection par différents pathogènes tels qu'*Helicobacter pylori* (Tang *et al.*, 2010), le virus d'Epstein Barr (Cameron *et al.*, 2008; Motsch *et al.*, 2007) ou encore le virus de la stomatite vésiculaire (Hou *et al.*, 2009). L'étude de Schulte *et al.*, montre également que plusieurs

membres de la famille de micro-ARN let-7 étaient rapidement inhibés par *S. enterica*. A l'aide d'expériences complémentaires, les auteurs ont montré que l'inhibition de let-7 était accompagnée d'une diminution de synthèse d'IL-6 et d'IL-10, et à l'inverse, le traitement des cellules avec un inhibiteur de let-7 avant infection était corrélé avec des niveaux plus forts d'IL-6 et d'IL-10 (Schulte *et al.*, 2012). Les auteurs concluent que la répression de let-7 permet aux cellules de produire une réponse inflammatoire équilibrée en réponse au pathogène (Schulte *et al.*, 2012). Le nombre d'études croissant portant sur l'expression et la fonction des micro-ARN lors d'infection montre que ces molécules sont une strate de régulation supplémentaire de la réponse immune et ouvre de nouvelles cibles thérapeutiques (O'Connell *et al.*, 2010; O'Neill *et al.*, 2011).

L'interaction *in vitro* entre une population cellulaire et un pathogène est un modèle simplifié de la complexité des interactions qui se produisent lors d'une infection et ne reflète ni le micro-environnement moléculaire qui existe *in vivo*, ni l'hétérogénéité des types cellulaires présents dans un organe, ni la complexité des interactions au sein d'un organe. Les études *in vitro* ou *ex vivo* ne peuvent donc pas prétendre refléter la complexité des interactions multifactorielles qui ont lieu *in vivo* (Jenner *et al.*, 2005). Néanmoins, ces modèles sont essentiels pour distinguer des interactions spécifiques ou communes entre un pathogène et un type cellulaire particulier. L'étude isolée de différentes populations cellulaires permet de disséquer des interactions complexes qui existent au sein d'un organe, en une série d'interactions moins complexes pour comprendre l'implication de chaque type cellulaire dans la réponse à un pathogène (Waddell *et al.*, 2007). En effet, comprendre la complexité du réseau d'interactions qui se produit lors d'une infection peut poser problème quand il s'agit d'interpréter des données d'expression issues d'un mélange de cellules qui répondent de manière différente à un pathogène.

Les données obtenues *in vivo* fournissent une perspective plus globale du changement d'expression génique au cours de l'infection dans un tissu ou organe (Waddell *et al.*, 2007).

3. L'apport des études transcriptomiques *in vivo*

Une maladie infectieuse résulte de l'interaction d'un pathogène avec différentes populations cellulaires, différents tissus et différents organes. L'étude de modèles *in vivo* permet d'étudier l'interaction entre un pathogène et un environnement complexe constitué de différentes populations cellulaires créant un micro-environnement moléculaire et cellulaire particulier. Le nombre élevé de variables dans un système *in vivo* peut compliquer l'identification de la source de variation des transcrits et diminuer la puissance de détection des gènes affectés par l'infection. Par ailleurs, les réponses transcriptionnelles peuvent être différentes d'un type cellulaire à l'autre dans un même tissu, la réponse spécifique d'un type cellulaire peu représenté risquant alors de passer inaperçue. Cependant,

ce type d'étude donne une vision intégrée de tous les processus qui se déroulent lors d'une infection au sein d'un organe ou d'un tissu.

Les études cinétiques explorent la nature dynamique de la réponse immunitaire. Un exemple d'étude transcriptionnelle cinétique *in vivo* est le travail du groupe de Klaus Schughart, qui a quantifié les variations d'expression des gènes dans les poumons de souris C57BL/6J pendant 60 jours durant une infection non-létale par le virus de l'Influenza A (H1N1) (Pommerenke *et al.*, 2012). Une analyse de *cluster* a été réalisée afin de déterminer des groupes de gènes qui variaient de manière similaire au cours de l'infection. Les auteurs ont mis en évidence un *cluster* de 445 gènes dont l'activation maximale était située entre les 2^{ème} et 5^{ème} jours post-infection, les 50 gènes les plus exprimés de ce cluster étant des gènes codant des cytokines et chimiokines tels que *Cxcl10*, *Ccl2*, *Ccl7*, et des gènes activés par l'interféron tels que *Ifi205* et *Irf7* (Pommerenke *et al.*, 2012). Un autre *cluster* comprenant 998 gènes présentait une activation maximale au 8^{ème} jour de l'infection. Une analyse d'enrichissement fonctionnel a mis en évidence une surreprésentation de gènes impliqués dans l'activation des cellules T, des cellules NK, dans l'induction de l'apoptose et dans l'induction d'une réponse interféron de type II (IFN γ). Deux autres clusters étaient activés plus tardivement dans l'infection, leur activation débutant entre le 3^{ème} et 5^{ème} jour pour atteindre un maximum au 14^{ème} jour avant de décroître plus ou moins lentement suivant le cluster. Les gènes de ces deux clusters étaient majoritairement impliqués dans l'activation des lymphocytes B (Pommerenke *et al.*, 2012). Peu de changements ont été observés entre le 18^{ème} et le 60^{ème} jour post-infection. Cependant, deux mois après l'infection, plus de 500 gènes étaient encore différenciellement exprimés par rapport à des souris non-infectées gardées 60 jours dans les mêmes conditions. Une analyse d'enrichissement fonctionnel de ces 500 gènes a mis en évidence une surreprésentation de gènes impliqués dans la réparation tissulaire (Pommerenke *et al.*, 2012).

De plus, Pommerenke *et al.* ont évalué la capacité de l'analyse transcriptionnelle à détecter la déficience d'une population cellulaire chez un animal. Pour cela, ils ont comparé le transcriptome de souris « knock-out » pour le gène *Rag2*, souris qui ne produisent pas de lymphocyte T et B, avec des souris WT au cours de l'infection par l'Influenza A. Le transcriptome de ces souris était similaire à celui des souris WT dans la phase précoce de l'infection. Cependant, au 8^{ème} jour, les souris *Rag2*^{-/-} ne présentaient pas de signature transcriptionnelle de l'activation des cellules T comme les WT, même si elles présentaient une activation similaire des cellules NK. De même, la déficience en cellules B était détectable au niveau transcriptionnel dans la phase plus tardive de l'infection (Pommerenke *et al.*, 2012).

Afin d'affiner leurs analyses et d'avoir une vision plus précise du profil d'activation des populations cellulaires qui constituent ou infiltrent les poumons, Pommerenke *et al.* ont réalisé une étude transcriptomique sur des populations isolées de cellules NK, et de cellules T et B issues de poumons infectés. La comparaison des gènes exprimés dans les populations de cellules isolées et dans

le tissu pulmonaire total a permis de mettre en évidence une signature transcriptionnelle spécifique de chacune de ces populations cellulaires (Pommerenke *et al.*, 2012).

Cette étude, basée uniquement sur l'étude du transcriptome du poumon, donne une vision globale du processus infectieux (activation, infiltration et déclin des populations de cellules immunitaires impliquées) engendré par le virus de l'Influenza A. Elle illustre comment l'étude *in vivo* du transcriptome, combinée à une dissection *in vitro*, permet de comprendre les mécanismes de défense de l'hôte au cours d'une infection.

Une limite des études *in vivo* est toutefois une faible puissance de détection des variations d'expression en raison d'un fort bruit de fond. L'hétérogénéité de la progression de l'infection entre les animaux nécessite d'analyser des groupes d'animaux de taille suffisante. L'utilisation d'animaux consanguins infectés dans des conditions très standardisées avec le même inoculum réduit cette variabilité, mais une hétérogénéité apparaît malgré tout (Sturdevant *et al.*, 2010). Un exemple de cette perte de puissance de détection des variations transcriptionnelles mesurées *in vivo* est l'étude de Rogers *et al.* portant sur un modèle d'infection par *Yersinia pestis* (qui sera détaillée ci-après) (Rogers *et al.*, 2007). Leur étude transcriptionnelle a porté sur la rate, prélevée deux jours post-infection, sur des souris BALB/c infectées par voie intra-péritonéale avec la souche CO92 de *Y. pestis*, à faible dose (8 ufc) ou à forte dose (257 ufc). La variabilité intra-groupe des résultats chez les souris infectées à faible dose était tellement forte qu'aucun gène n'a pu être considéré comme significativement affecté par l'infection (Rogers *et al.*, 2007). Cette étude illustre que le déroulement d'un processus infectieux, même dans des conditions très standardisées (environnementales, génétiques et d'infection) varie d'un individu à l'autre. C'est une limite des études *in vivo*, qui peut être minimisée si le processus infectieux est caractérisé chez l'ensemble des animaux par une évaluation de paramètres tel qu'un grade histologique des lésions ou le dosage de cytokines, qui permet de créer des groupes homogènes (Comer *et al.*, 2010). L'équipe d'Elisabeth Carniel a montré que lors d'une infection par *Y. pestis*, le temps que met la bactérie à migrer du site d'infection sous-cutané au nœud lymphatique drainant était la phase qui présentait la plus grande variabilité inter-individuelle. En revanche, une fois que la bactérie avait atteint le nœud lymphatique, le processus d'infection était très homogène (Nham *et al.*, 2012).

La prise en compte de cette variabilité au cours du processus infectieux est essentielle pour réduire la dispersion intra-groupe dans les études *in vivo*.

4. Combiner l'analyse transcriptomique et l'approche QTL : un moyen de cibler des gènes candidats ?

Comme nous l'avons vu précédemment, l'approche QTL qui est utilisée pour identifier des régions génomiques associées à un caractère, détecte généralement des régions de plusieurs dizaines de mégabases qui contiennent des centaines de gènes parmi lesquels il est difficile d'identifier des candidats. Seul un petit nombre de gènes pourra être validé fonctionnellement, en raison du coût et du temps que prennent de telles études, ce qui nécessite d'identifier au préalable un petit nombre de candidats robustes. Les analyses d'expression génique peuvent alors être d'un grand secours. Tout d'abord parce que la variation de régulation des gènes est un mécanisme important dans la détermination des caractères complexes, dans différentes espèces (Hunter *et al.*, 2008; Mackay, 2001 A; Wayne *et al.*, 2002). Mais aussi parce que les études transcriptomiques permettent de mettre en évidence des voies de signalisation ou voies métaboliques (*pathways*) qui diffèrent entre deux lignées, ce qui permet ensuite d'identifier des gènes candidats localisés dans le QTL qui ont une fonction en rapport avec les voies de signalisation identifiées.

Le raisonnement sous-tendant cette approche combinée est que le déterminisme génétique d'un phénotype peut correspondre à des polymorphismes dans les régions de régulation en *cis* ou en *trans* de gènes dont les variations du niveau d'expression influencent le phénotype étudié. Il y a plusieurs manières par lesquelles un polymorphisme peut provoquer une variation du niveau d'expression d'un gène. Une mutation dans le site de fixation d'un facteur de transcription peut, par exemple, diminuer l'affinité du facteur pour son site (GuhaThakurta, 2006 A; GuhaThakurta *et al.*, 2006 B). Une mutation dans un facteur de transcription lui-même peut aussi affecter la reconnaissance et la fixation sur ses cibles. Une mutation non-sens dans la région codante d'un gène peut également diminuer le niveau de transcription par la formation d'ARNm instables rapidement dégradés (Chang *et al.*, 2007).

L'approche combinant la localisation génétique des QTL et une analyse transcriptomique ne doit pas être confondue avec la recherche de QTL d'expression, dans laquelle le niveau d'expression d'un gène est considéré comme un caractère quantitatif et soumis à une recherche de QTL. Cette approche permet d'identifier des QTL d'expression ou e-QTL (Alberts *et al.*, 2011; Bottomly *et al.*, 2012; Farrall, 2004; Pomp *et al.*, 2004).

4.1. Les différentes stratégies combinant QTL et analyse transcriptomique

La méthodologie de l'approche combinée QTL et transcriptomique varie d'une étude à l'autre. La première étape est l'identification d'un QTL et le recensement des gènes contenus dans l'intervalle de confiance. La deuxième étape est la comparaison du transcriptome en réponse à l'infection entre les lignées parentales ou entre lignées parentales et lignées congéniques. C'est la troisième étape, qui consiste à combiner ces deux approches pour cibler des candidats, qui varie suivant les études.

On distinguera ici trois stratégies principales, qui sont (1) l'étude des lignées parentales, (2) l'étude des lignées congéniques et (3) la recherche de gènes situés dans un QTL et impliqués dans les voies de signalisation perturbées.

(1) Certaines études croisent la liste des gènes présents dans le QTL et celle des gènes différentiellement exprimés entre les lignées parentales afin de distinguer les gènes présents dans le QTL dont l'expression varie entre les lignées étudiées. Cette stratégie est basée sur l'hypothèse que les gènes à l'origine du QTL sont différentiellement régulés entre les lignées parentales en réponse à l'infection et que cette variation est régulée en *cis*. Si cette approche a permis la détection de candidats (Figure 17 pour rappel) (Ahn *et al.*, 2010), elle est néanmoins réductrice, car elle ne tient pas compte des effets en *trans* des QTL (un QTL localisé sur un chromosome peut modifier l'expression de gènes situés sur d'autres chromosomes). De plus, les lignées parentales présentent fréquemment de nombreux polymorphismes sans relation avec le phénotype étudié. Cette approche fournit donc des pistes à explorer pour détecter des effets en *cis*. En revanche, l'étude des mécanismes biologiques différentiellement activés pendant l'infection entre les lignées parentales peut procurer des pistes quant aux gènes du QTL qui ont un effet en *cis* ou en *trans* et permettre de réduire le nombre de gènes candidats dans le QTL (Verdugo *et al.*, 2010).

(2) Une approche intéressante pour distinguer les effets en *cis* et en *trans* des QTL est la comparaison du transcriptome entre une lignée congénique et la lignée receveuse. Les variations d'expression des gènes localisés dans la région du QTL sont dues à des facteurs génétiques situés dans les limites de l'intervalle transféré (Verdugo *et al.*, 2010). En revanche, les différences d'expression de gènes situés en dehors de la région transférée sont dues à l'effet direct ou indirect, en *trans*, de gènes contenus dans l'intervalle congénique.

Le travail de Marquis *et al.*, sur les facteurs génétiques de sensibilité à *Mycobacterium tuberculosis*, est un exemple d'étude employant une stratégie d'approche traditionnelle des QTL combinée à une analyse transcriptomique comparative entre des lignées parentales et congéniques (Marquis *et al.*, 2009). Un croisement F2 a montré que le contrôle de la sensibilité à *M. tuberculosis* chez DBA/2J (lignée sensible) et C57BL/6J (lignée résistante) était sous le contrôle d'au moins quatre

locus (dénommé *Trl1* à *Trl4* pour *Tuberculosis resistance locus*) dont *Trl3* a été plus amplement étudié (Mitsos *et al.*, 2000; Mitsos *et al.*, 2003). *Trl3* participe au contrôle de la survie et de la quantité de bactéries dans les poumons aux 30^{ème} et 70^{ème} jours post-infection (Mitsos *et al.*, 2000; Mitsos *et al.*, 2003). Une lignée congénique D2.B6-*Trl3* homozygote C57BL/6J pour *Trl3* dans le fonds sensible DBA/2J a été créée (Marquis *et al.*, 2009). Cette lignée présentait un taux de survie et une charge bactérienne intermédiaires entre les lignées parentales, confirmant l'effet de *Trl3*. De plus, une étude des lésions histologiques et du niveau de réplication de *M. tuberculosis* dans les macrophages pulmonaires a montré que *Trl3* permettait de limiter la réplication bactérienne. Afin de mieux comprendre les mécanismes cellulaires et moléculaires qui se jouaient dans cette différence de sensibilité entre DBA/2J, C57BL/6J et la lignée congénique, Marquis *et al.* ont réalisé une étude transcriptomique sur les poumons de ces trois lignées aux 30^{ème} et 70^{ème} jour post-infection.

Ils ont tout d'abord établi la liste des gènes dont le profil transcriptionnel était identique aux 30^{ème} et au 70^{ème} jour post-infection entre les trois lignées. Cette liste, qui comprend des gènes impliqués dans la réponse inflammatoire tels que *Ifn γ* , *Stat1*, *Irg1*, *Iigp1*, et des ligands et récepteurs de chemokine tels que *Cxcl9*, *Cxcl10*, *Ccl8* et *Cxcr6*, permet de définir un profil transcriptionnel mis en place en réponse à *M. tuberculosis* dans le poumon quelque soit la sensibilité de l'hôte (Marquis *et al.*, 2009). Dans un deuxième temps, les auteurs ont cherché à identifier un groupe fonctionnel de gènes dont la variation d'expression était similaire entre C57BL/6J et la congénique, mais différente de la lignée sensible DBA/2J. Ils ont trouvé 565 gènes différentiellement régulés au 30^{ème} jour de l'infection entre C57BL/6J et DBA/2J, et 69 gènes entre DBA/2J et D2.B6-*Trl3* (Marquis *et al.*, 2009). Un fort enrichissement de transcrits régulés en *cis* était observé dans la lignée congénique D2.B6-*Trl3* (34,8% des gènes différentiellement régulés étaient des gènes de l'intervalle congénique). La majorité des gènes de cette liste étaient impliqués dans la réponse immunitaire. Sept de ces gènes étaient localisés dans l'intervalle congénique transféré dans la lignée D2.B6-*Trl3*. Une revue de la littérature sur la fonction de ces gènes a permis à Marquis *et al.* de cibler deux candidats. *Cd22*, localisé au niveau du pic du QTL, est un inhibiteur de la voie de signalisation induite par le BCR (immunoglobuline impliquée dans la reconnaissance de l'antigène à la surface des lymphocytes B), et *Coro1a* qui code une protéine de la membrane du phagosome dans lequel se trouvent les mycobactéries (Marquis *et al.*, 2009). L'analyse transcriptomique constitue dans cet exemple une première étape de criblage pour identifier des gènes candidats dans le QTL, qui feront ensuite l'objet d'expérimentations plus ciblées.

Cependant, cette étude s'est concentrée sur les effets en *cis* des QTL. Mais les modifications du niveau d'expression des gènes situés dans l'intervalle congénique ne sont pas forcément causales. De plus, plusieurs études menées dans différentes espèces ont montré que les effets en *cis* sont plus forts et plus nombreux, mais ont moins de chance d'être causaux que les effets en *trans* (Brem *et al.*, 2002; Dixon *et al.*, 2007; Hubner *et al.*, 2005; Morley *et al.*, 2004; Schadt *et al.*, 2003; Verdugo *et al.*, 2010).

Une limite de l'étude transcriptomique des lignées congéniques est la présence potentielle de fragments résiduels de la lignée donneuse en dehors de l'intervalle transféré qui peut amener à la détection de faux effets en *trans* (Verdugo *et al.*, 2010). Pour éviter ce problème, il faut utiliser des lignées congéniques résultant d'un minimum de dix *backcross*. Il est maintenant peu coûteux de les caractériser pour plusieurs milliers de marqueurs pour s'assurer de leur pureté génétique. Une autre solution consiste à randomiser l'effet d'éventuelles régions contaminantes en analysant des animaux issus d'un croisement F2 entre la lignée congénique et la lignée receveuse (Verdugo *et al.*, 2010).

(3) Une autre stratégie pour cibler des candidats consiste à identifier les gènes communs aux QTL et aux voies de signalisation sur-représentées dans l'analyse transcriptomique. Les données de transcriptome permettent d'identifier des mécanismes biologiques qui diffèrent entre les lignées étudiées. Les gènes présents dans les QTL qui pourraient être à l'origine des différences biologiques observées sont ciblés en priorité.

Cette approche a été utilisée par Fisher *et al.* dans une étude sur la sensibilité à *Trypanosoma congolense* dans la lignée sensible A/J et la lignée résistante C57BL/6 (Figure 18), (Fisher *et al.*, 2007). Cette étude souligne l'importance de connaître les mécanismes biologiques qui sous-tendent un caractère, mais aussi l'avantage d'avoir à disposition différents logiciels et bases de données qui permettent d'évaluer la régulation des différentes voies de signalisation à partir des données de transcriptome. L'analyse des *pathways* activés ou réprimés donne des pistes pour sélectionner des candidats à tester fonctionnellement parmi l'ensemble des gènes candidats positionnels. L'étude de Fisher *et al.* menée avec le logiciel *KEGG Pathways* leur a permis de mettre en évidence quatre *pathways* différenciellement activés dans l'analyse transcriptomique entre les deux lignées de souris étudiées : « *complement and coagulation cascade* », « *arachidonic acid metabolism* », « *Toll-like receptor signalling pathway* » et « *MAPK signalling pathway* » (Fisher *et al.*, 2007). Ils avaient antérieurement mis en évidence que la survie à *T. congolense* était sous contrôle de plusieurs locus dont un à effet fort nommé *Tir1* (pour *Trypanosoma infection response 1*), (Kemp *et al.*, 1997). L'étude de la fonction des gènes présents dans *Tir1* leur a permis de mettre en évidence que plusieurs gènes étaient impliqués dans ces quatre *pathways*. Une analyse de la littérature les a menés à s'intéresser particulièrement à quatre gènes du QTL impliqués dans la voie des MAP kinases : *Daxx*, *TNF*, *Mapk13* et *Map14* (Fisher *et al.*, 2007). Le gène *Daxx*, activateur de l'apoptose, présentait une forte différence d'expression en réponse à l'infection entre la lignée résistante et la lignée sensible (Fisher *et al.*, 2007). Le séquençage de ce gène dans la lignée A/J a mis en évidence une délétion de trois nucléotides dans l'exon 5 du gène, conduisant à la suppression d'un acide aspartique dans le domaine de fixation DAXX à p53 (protéine majeure de contrôle de l'apoptose cellulaire). Cette délétion a été retrouvée chez une autre lignée sensible (Fisher *et al.*, 2007).

Ce travail illustre l'intérêt des études de transcriptome combinées avec l'approche QTL pour donner des pistes de gènes candidats. Les gènes présents dans un QTL qui appartiennent à plusieurs voies de signalisation surreprésentées sont des candidats à tester en priorité.

De plus, Fisher *et al.* soulignent aussi l'intérêt de comparer les résultats des analyses de *pathways* et d'enrichissement fonctionnel des gènes à l'aide de plusieurs logiciels ou bases de données (par exemple *KEGG Pathway*, *Ingenuity Pathway Analysis*, *GO term enrichment analysis*, ou encore *DAVID*).

4.2. *L'approche combinée QTL et analyse de transcriptome est-elle vraiment efficace pour cibler des candidats ?*

Pour évaluer la performance de cette approche, Verdugo *et al.* ont réalisé une méta-analyse de 37 études combinant approche QTL et étude de transcriptomique pour différents caractères complexes chez la souris et le rat. Ils ont complété cette méta-analyse par une étude originale sur des lignées congéniques qui leur a permis de tirer des conclusions sur les avantages et les limites de cette approche combinée (Verdugo *et al.*, 2010). Ils ont montré dans leur étude de trois lignées congéniques comparées à la lignée receveuse, que sélectionner les candidats en ne gardant que les gènes différentiellement régulés et présents dans le QTL permettait de réduire la liste des candidats positionnels de plus de 92% (Verdugo *et al.*, 2010). Néanmoins, ils soulignent également que ce filtre crée de nombreux faux négatifs, car l'ensemble des gènes de l'intervalle ne sont pas couverts par la puce et les gènes qui présentent du polymorphisme dans la région d'hybridation des sondes entre les lignées étudiées doivent être rajoutés à la liste des candidats. Une fois ces paramètres pris en compte, la réduction effective de la liste de gènes n'est plus, en moyenne, que de 64% (Verdugo *et al.*, 2010). En effet, une limite bien connue de l'approche combinée QTL et étude transcriptomique est le polymorphisme qui peut exister entre la séquence génomique des lignées étudiées et la séquence de la lignée pour laquelle les sondes des puces à ADN ont été conçues. Ce polymorphisme peut affecter significativement l'hybridation des ADNc sur les sondes et être responsable de faux positifs parmi les gènes différentiellement exprimés entre deux lignées. Ce problème a un impact d'autant plus fort que les lignées comparées sont divergentes (Alberts *et al.*, 2005; Alberts *et al.*, 2007). Une recommandation de Verdugo *et al.* pour l'éviter est d'utiliser des puces à haute densité dont les sondes ciblent différents domaines des molécules d'ADNc, et d'éliminer de l'analyse les sondes qui présentent un polymorphisme entre les lignées étudiées (Verdugo *et al.*, 2010). La technique de *RNA-seq* permet de dépasser cette limite technique des puces à ADN, puisque cette technique est directement basée sur le séquençage des ADNc (Wang *et al.*, 2009). De plus, la technique de *RNA-seq* permet de dépasser d'autres limites que présentent les méthodes basées sur l'hybridation. Par exemple, elle permet d'identifier les différentes isoformes issues d'un ARNm suite à un épissage alternatif. Ces

isoformes ne sont détectées par les puces que lorsque des sondes ont été designées pour, alors que toutes les isoformes d'un même ARNm sont détectées par *RNA-seq*, qui permet même d'identifier de nouveaux exons (Wang *et al.*, 2009). Cet avantage de la technique de *RNA-seq* est particulièrement intéressant lorsque deux lignées divergentes sont étudiées, car il est probable qu'une fraction non négligeable du polymorphisme touche des sites d'épissage. L'expression de différentes isoformes d'un ARNm entre deux lignées peut être à l'origine d'une variation phénotypique (Wang *et al.*, 2008).

Verdugo *et al.* mettent également en garde contre l'utilisation du critère d'IBD « *identity by descent* » pour sélectionner les gènes candidats. Ce critère de sélection repose sur l'hypothèse que deux régions provenant d'un ancêtre commun aux deux lignées étudiées ont moins de chance de porter des gènes causaux responsables de la différence phénotypique que des régions ayant une origine ancestrale distincte (Doss *et al.*, 2005). Les auteurs soulignent que des mutations ont pu apparaître depuis la divergence des lignées étudiées. De plus, des gènes modificateurs localisés hors des régions d'IBD peuvent réguler l'expression de gènes localisés dans la région d'IBD (Verdugo *et al.*, 2010).

L'étude de Verdugo *et al.* démontre que la stratégie de sélection des gènes candidats sur le seul critère d'une expression différentielle dans les lignées parentales, ou entre une lignée congénique et la lignée receveuse, présente de sérieuses limitations. En effet, la conclusion de leur méta-analyse est que moins de 2% des gènes candidats d'un QTL sont différentiellement exprimés (Verdugo *et al.*, 2010). Les études transcriptomiques révèlent néanmoins des informations précieuses sur les processus biologiques mis en place pendant une infection entre différentes lignées, comme il est montré dans l'étude de Fisher *et al.* (Fisher *et al.*, 2007). Les gènes des QTL impliqués dans ces processus biologiques sont des candidats robustes.

Enfin, il est toujours utile de confirmer que la différence de niveau d'expression en ARNm se traduit par un niveau différent d'expression protéique pour prendre en compte les différences de régulation post-transcriptionnelles et post-traductionnelles.

Cette approche combinée donne des pistes de gènes candidats, mais ne permet cependant pas d'établir un lien de causalité entre des gènes différentiellement exprimés et un phénotype. Des tests complémentaires, tels que l'inactivation de gènes par ARN interférence sont nécessaires pour passer d'un QTL à un QTG (*Quantitative Trait Gene*).

5. L'apport des analyses transcriptomiques dans la compréhension des interactions entre *Yersinia pestis* et son hôte

Une autre manière de hiérarchiser les gènes candidats d'un QTL est de réaliser une méta-analyse des gènes qui ressortent comme différentiellement exprimés entre différentes études portant sur le même caractère.

Même si de nombreux facteurs varient entre les études (type de puce utilisée, conditions environnementales, génétique de l'hôte, types cellulaires ou tissus, modes d'analyses, critères statistiques appliqués, etc.), les gènes qui ressortent de plusieurs analyses de transcriptome ont des chances d'être de robustes biomarqueurs de l'interaction hôte-pathogène étudiée.

Jenner et Young, en 2005, ont réalisé une méta-analyse des données de 32 études de transcriptome, qui impliquait 77 interactions hôte-pathogène différentes. De cette méta-analyse, ils ont mis en évidence 511 gènes dont l'expression était commune à différents types cellulaires en réponse à un large panel de plus de 60 bactéries, virus, levures et parasites (Jenner *et al.*, 2005). Ce groupe de gènes contenait aussi des gènes préférentiellement induits dans un type cellulaire ou par un pathogène spécifique. Jenner et Young suggèrent que ce programme transcriptionnel commun pourrait constituer un « signal d'alarme » non spécifique des différents types cellulaires qui maximiserait les chances que l'infection soit détectée (Jenner *et al.*, 2005). Ils ont montré que ces 511 gènes étaient majoritairement associés à la réponse immunitaire et particulièrement dans la réponse pro-inflammatoire (cytokines et chimiokines). Certains facteurs de transcription étaient aussi sur-représentés, ainsi que des gènes impliqués dans la transduction du signal et des facteurs pro- et anti-apoptotiques (Tableau 4) (Jenner *et al.*, 2005).

Six études transcriptomiques portant sur *Y. pestis* ont été publiées à ce jour (Tableau 5) (Comer *et al.*, 2010; Das *et al.*, 2007; Galindo *et al.*, 2009; Liu *et al.*, 2009; Ng *et al.*, 2003; Rogers *et al.*, 2007). J'ai comparé les gènes dont l'expression est affectée par *Y. pestis* dans ces études, afin de faire ressortir ceux qui sont retrouvés de façon récurrente (Tableau disponible en version informatique seulement).

Sur les six études, deux présentent les données de transcriptome produites sur un modèle de macrophages cultivés *in vitro* et exposés à *Y. pestis* cultivée à 26°C. Cultivée à cette température, qui est celle de la puce, *Y. pestis* est intracellulaire (contenue dans les phagolysosomes des macrophages) et n'exprime pas le *Yop virulon*, qui code entre autre les effecteurs Yop. Les études de Ng *et al.* et de Das *et al.* ont utilisé ce modèle pour mimer la phase très précoce de l'infection qui précède l'expression par *Y. pestis* de ces facteurs de virulence induits par le passage dans un hôte mammifère à

37°C (Das *et al.*, 2007; Ng *et al.*, 2003). L'étude de Ng *et al.* ne permettant pas d'avoir accès à la méthode utilisée, ni à une liste assez exhaustive de gènes (seuls 22 gènes surexprimés sont décrits dans une très courte publication), seule l'étude de Das *et al.*, apparaît dans la comparaison que j'ai réalisée.

Les quatre autres études décrites ont utilisé un modèle *in vivo*. Deux études présentent le niveau d'expression des gènes dans un modèle de peste bubonique et/ou septicémique. Comer *et al.* ont inoculé la bactérie à des rats par voie intra-dermique et étudié le transcriptome dans le nœud lymphatique drainant à 36 et 60 h (Comer *et al.*, 2010). Les auteurs ne trouvent aucun gène différentiellement exprimé par rapport aux témoins non-infectés à 36 h (phase bubonique de l'infection, mise en évidence par une analyse histologique concomitante du nœud lymphatique). Par contre, une forte réponse inflammatoire était décelée au niveau transcriptomique à 60 h, phase de l'infection où la bactérie était passée dans la circulation générale, créant une septicémie (Comer *et al.*, 2010).

Rogers *et al.* ont inoculé la bactérie à des souris par voie intra-péritonéale et ont étudié la réponse transcriptionnelle dans la rate après 48h (Rogers *et al.*, 2007). Ils ont comparé le transcriptome de souris infectées avec une faible dose (8 ufc) et une forte dose (257 ufc) de *Y. pestis*. Les auteurs ont considéré les 76% de survie des souris ayant reçu une faible dose constituait la signature d'une réponse de l'hôte efficace contre la bactérie, alors que les 100% de mortalité des souris infectées à forte dose témoignaient d'une réponse inefficace contre *Y. pestis*. Ils n'ont mis en évidence aucun gène différentiellement exprimé en réponse à l'infection dans le groupe infecté par une faible dose. Par contre 534 gènes sont apparus comme différentiellement exprimés après une infection à forte dose. La majorité de ces gènes était réprimée (384 sous-exprimés contre 150 sur-exprimés). Les gènes sur-exprimés étaient impliqués dans des processus immunitaires et d'apoptose. Les gènes sous-exprimés étaient impliqués dans les processus de « réponse immunitaire », « cytosquelette », « cycle cellulaire », « dégradation protéique » et « apoptose » (Rogers *et al.*, 2007).

Les deux autres études ont étudié un modèle d'inoculation intra-nasale afin de déclencher une forme pulmonaire de la maladie. L'étude de Liu *et al.* a utilisé une bactérie avirulente chez l'humain, mais fortement virulente chez la Souris. Cette étude décrit les variations transcriptomiques dans les poumons, le foie et la rate 12 et 48 h après une infection par aérosol (Liu *et al.*, 2009). Les auteurs montrent qu'à 12 h, seul le foie produit une forte réponse inflammatoire, alors qu'une réponse n'est détectée dans les poumons et la rate qu'après 48 h (Liu *et al.*, 2009).

L'étude de Galindo *et al.* décrit également les variations transcriptomiques dans les poumons, le foie et la rate 12 et 48 h après une infection par aérosol. Les auteurs montrent que peu de gènes (quelques dizaines tout au plus) sont affectés 12 h après infection, alors qu'une forte réponse est présente à 48 h (plusieurs centaines de gènes affectés). Ils montrent que les processus activés lors de l'infection sont très similaires dans ces trois organes, où ils détectent une modification des processus inflammatoires, apoptotiques, de prolifération et de migration cellulaires (Galindo *et al.*, 2009).

Ces quatre études de transcriptome *in vivo* montrent clairement que *Y. pestis* arrive à inhiber la réponse de l'hôte afin de survivre et se multiplier dans les premiers temps de l'infection avant de provoquer une réponse inflammatoire forte lors de la phase septicémique aboutissant à un choc septique et à la mort de l'animal.

Les données d'expression de ces différentes études n'étant pas déposées dans des bases de données publiques, la comparaison de gènes communs que j'ai effectuée ne porte que sur les résultats présentés dans les publications et les informations complémentaires (724 entrées) (Comer *et al.*, 2010; Das *et al.*, 2007; Galindo *et al.*, 2009; Liu *et al.*, 2009; Rogers *et al.*, 2007). Il est donc clair qu'une telle comparaison est à considérer avec précaution et que d'autres gènes communs à la réponse de l'hôte à *Y. pestis* ne sont pas identifiés dans cette analyse.

Cent vingt-huit gènes sont activés ou réprimés de façon récurrente en réponse à l'infection à *Y. pestis* dans différents organes. Huit classes fonctionnelles très vastes sont représentées par les 128 gènes identifiés. Le groupe fonctionnel majoritaire est « réponse immunitaire », dont font partie deux sous-catégories majoritaires impliquées dans la réponse inflammatoire « cytokines et récepteurs » et « chimiokines et récepteurs ». Les groupes fonctionnels « apoptose », « cycle cellulaire », « complément », « dégradation protéique » et « réponse au stress » sont également représentés.

Les familles majoritaires de gènes communs aux cinq études sont :

- les membres de la famille de Bcl2, impliqués dans l'apoptose;
- différents « *insulin-like growth factors* » qui ont également un rôle dans l'apoptose;
- des cyclines impliquées dans le cycle cellulaire;
- des « *CD antigens* » impliqués dans diverses fonctions telles que l'activation des lymphocytes T, l'adhésion cellulaire, la transduction du signal. Parmi cette famille, le gène *Cd14*, co-récepteur qui participe à la reconnaissance du LPS avec TLR4 et MD2, est sur-exprimé dans tous les organes testés des quatre études menées *in vivo*;
- des chimiokines et des récepteurs de chimiokines, dont *Ccl20*, *Cxcl2* et *Cxcl6* sont sur-exprimés dans trois études;
- des composants du complément, impliqués dans la réponse inflammatoire et le recrutement de cellules de l'immunité;
- des facteurs de croissance impliqués dans la prolifération et la différenciation des cellules souches hématopoïétiques en macrophages et granulocytes (*Granulocyte and/or Macrophage Colony stimulating factor*), dont *Csf3* qui est sur-exprimé dans deux études et son récepteur dans une autre;

- l'interféron gamma (*Ifn γ*), impliqué dans la réponse inflammatoire, sur-exprimé dans deux études *in vivo* et sous-exprimé dans une autre;

- des interleukines, majoritairement impliquées dans la réponse inflammatoire, dont *Il1 β* sur-exprimée dans tous les organes des trois études *in vivo* et de celle *in vitro*. Néanmoins, la surexpression de ce gène au niveau transcriptionnel n'est pas très informative, car la protéine pro-IL1 β synthétisée à partir du transcrit doit ensuite être activée au sein de l'inflammasome pour présenter un effet biologique ;

- les lectines de type C (Clec), impliquées dans divers processus, tels que l'inflammation, l'adhésion cellulaire, l'endocytose ou encore la fixation du complément;

- *Icam1* et *Icam2*, codant des protéines d'adhésion cellulaire et impliqués dans différents processus dont l'inflammation, l'adhésion et la migration cellulaire;

- *Ifitm1*, 2, 3 et 6 qui codent des protéines induites par l'interféron et impliquées dans la réponse inflammatoire. Ces quatre gènes sont sur-exprimés dans tous les organes examinés par les 4 études *in vivo*;

- des « *matrix metalloproteinases* » (MMP), gènes codant des enzymes impliquées dans de nombreux processus biologiques tels que la dégradation et le remodelage de la matrice extracellulaire, la prolifération cellulaire, la différenciation cellulaire, l'adhésion et la migration cellulaire, l'angiogénèse, l'apoptose, ainsi que dans plusieurs phases de l'inflammation, telles que le recrutement de cellules pro-inflammatoires, l'activation et l'inactivation des cytokines, mais aussi la résolution de l'infection et la réparation tissulaire (Van Lint *et al.*, 2007; Vanlaere *et al.*, 2009). Néanmoins, la sur-expression des gènes codant les MMP est difficile à interpréter, car même si ces protéines sont exprimées, plusieurs d'entre elles doivent être clivées pour acquérir leur potentiel enzymatique (Van Lint *et al.*, 2007). Des « *Timp* », inhibiteurs des « *matrix metalloproteinases* », sont également sur-exprimés ;

- *Mt1* et *Mt2*, « *metallothionin* » 1 et 2, sur-exprimées dans les 4 études *in vivo*; elles sont impliquées dans la fixation des métaux lourds;

- des Rho GTPases, impliquées dans des processus cellulaires variés tels que la dynamique des filaments d'actine, la prolifération cellulaire ou encore l'apoptose;

- des « *Serpin* » pour « *Serine peptidase inhibitor* » dont le rôle n'est pas encore très bien défini;

- la famille du *Tgf β* , qui participe à un grand nombre de fonctions cellulaires telles que le contrôle de la prolifération cellulaire et l'apoptose, ou encore la réponse immunitaire;

- la famille du *Tnf*, dont les membres sont impliqués dans une grande variété de processus, tels que l'inflammation et son contrôle, l'apoptose, l'activation des lymphocytes. Les gènes *Tnfaip2*, *3b*, *6* et *8* impliqués dans l'apoptose et l'inhibition de l'inflammation sont sur-exprimés dans trois études *in vivo*;

- les gènes *Gadd45a*, *b* et *g*, senseurs du stress, impliqués dans différents processus tels que la réparation de l'ADN, le contrôle du cycle cellulaire ou encore l'apoptose via l'activation des voies JNK, p38 et MAP kinases; ces trois gènes sont sur-exprimés dans trois des études *in vivo*.

J'ai comparé cette « réponse commune » de l'hôte à *Y. pestis* à celle établie par Jenner et Young dans leur méta-analyse qui impliquait 77 interactions hôte-pathogène différentes (Tableau 4). Tous les gènes, ou familles de gènes, présentés ici sont également rapportés par Jenner et Young dans leur étude. La réponse à *Y. pestis* semble être typique d'une réponse classique à un agent infectieux, si ce n'est que cette liste commune de gènes est réalisée à un stade avancé de l'infection, car dans la phase précoce seule une très faible réponse est détectable.

Cette méta-analyse n'est pas exhaustive à cause de l'accès à un nombre limité de gènes pour chaque étude. Des groupes fonctionnels de gènes qui ne sont pas identifiés comme significativement enrichis dans les études prises isolément (et ne sont donc pas cités dans les données disponibles) pourraient en réalité être affectés par *Y. pestis* si toutes les données brutes étaient analysées simultanément. De plus, j'ai conscience que le petit nombre d'études disponibles ne permet de faire que des hypothèses sur les perturbations transcriptionnelles générales induites par *Y. pestis*.

Quatre études ont étudié les interactions entre le protéome de *Y. pestis* et les protéines humaines *in vitro* (Chromy *et al.*, 2004; Dyer *et al.*, 2010; Yang *et al.*, 2011 A; Zhang *et al.*, 2005a). Les données qui découlent de ces analyses montrent que la bactérie cible en priorité des voies cellulaires centrales dans la mise en place de la réponse immunitaire, telles que la voie des TLR, la voie des MAP kinases, ou encore la voie d'activation des leucocytes. Les protéines clés dans ces voies sont ciblées par la bactérie afin de réprimer la mise en place de la réponse inflammatoire (Dyer *et al.*, 2010; Zhou *et al.*, 2012).

Pour conclure, les études transcriptomiques, confirmées par les études protéomiques, montrent que *Yersinia pestis* est capable de déjouer la réponse de l'hôte pour survivre et se multiplier durant la phase précoce de l'infection. Par la suite, une vigoureuse réponse pro-inflammatoire apparaît lorsque la bactérie dissémine et atteint l'ensemble des organes.

Partie 6. *Yersinia pestis*, l'agent de la peste

« Les bubons contiennent en abondance et à l'état de pureté, un bacille très petit, court, à bouts arrondis, ne se teignant pas par la méthode de Gram, mais se colorant par le violet de Gentiane. Chez huit malades, j'ai retrouvé le bacille dans les bubons. A l'autopsie de deux pestiférés, j'ai rencontré le même microbe. [...] Des souris inoculées avec une trace de la pulpe d'un bubon meurent en vingt-quatre heures avec des bacilles dans les ganglions, dans les organes et dans le sang où ils sont plus longs et grêles; elles succombent à une véritable septicémie. » A. Yersin dans une lettre à E. Duclaux, 1894, lettre communiquée à l'Académie des Sciences le 30 juillet 1894.

La peste est une maladie bactérienne causée par le bacille Gram-négatif *Yersinia pestis*. Le nom de *Yersinia* lui a été donné en hommage à Alexandre Yersin, médecin d'origine suisse, élève de Louis Pasteur, qui isola, en 1894 à Hong-Kong, le bacille de la peste sur des cadavres humains et de rongeurs.

1. La peste

Depuis le début de l'ère chrétienne, *Y. pestis* a été la cause d'au moins trois grandes pandémies. La première rapportée dans l'histoire, dite peste de Justinien, date du milieu du VI^{ème} siècle. Cette pandémie a sévi de 541 à 767 dans tout le bassin méditerranéen. Elle aurait tué plusieurs dizaines de millions de personnes durant ces deux siècles (Demeure *et al.*, 2009). La Peste Noire, ou peste médiévale, est la seconde grande pandémie connue. Venue d'Asie Centrale, elle toucha l'Europe au XIV^{ème} siècle et provoqua la mort d'un quart de la population de l'Europe Occidentale (soit 25 millions d'êtres humains) entre 1348 et 1352 (Demeure *et al.*, 2009). Des villes comme Paris, Venise ou Hambourg ont perdu jusqu'à la moitié de leur population. La peste noire eut des conséquences durables sur la civilisation européenne, d'autant que la maladie refit régulièrement son apparition en Europe pendant plus de trois siècles (Stenseth *et al.*, 2008). La troisième pandémie prit naissance à Hong-Kong en 1894. C'est par la voie maritime, notamment grâce à la navigation à vapeur, que des continents qui avaient été épargnés jusqu'alors furent atteints (Amérique du Nord et du Sud, Madagascar, Afrique du Sud, Australie). Cette troisième pandémie toucha principalement l'Inde où elle fit 12 millions de morts entre 1898 et 1928 (Demeure *et al.*, 2009).

1.1. Le cycle de la peste

Y. pestis est habituellement transmise de rongeur à rongeur par des puces c'est avant tout une maladie des rongeurs. Néanmoins, des puces de rongeurs peuvent, dans certaines circonstances, contaminer l'homme par leur piqûre. On distingue schématiquement trois cycles majeurs (Figure 19):

(1) La peste sylvatique qui sévit chez les rongeurs sauvages. Sa transmission s'effectue par les puces qui passent d'un rongeur à l'autre ; la contamination de l'homme est exceptionnelle.

(2) La peste rurale résulte de la contamination d'une population d'un village à partir des rongeurs vivant à proximité. C'est le cycle de transmission qui est le plus souvent à l'origine d'une épidémie humaine. L'épizootie décime la population de rongeurs et les puces, affamées, cherchent alors à se nourrir sur d'autres mammifères, dont l'Homme (Perry *et al.*, 1997).

(3) La peste urbaine, importée en ville à partir d'un foyer rural ou sauvage, se propage rapidement chez les rongeurs urbains (rats).

Alors que de nombreuses espèces de rongeurs sont très sensibles au bacille de la peste, d'autres sont relativement résistantes et constituent un réservoir naturel de la bactérie. A. Yersin avait déjà décrit le rôle essentiel des rats dans la propagation de la peste; il écrivait, dans une lettre à A. Calmette datée du 15 septembre 1894, « Les rats sont certainement les grands propagateurs de l'épidémie; ces animaux habitent les égouts où ils entretiennent la maladie à l'état sporadique; impossible de les désinfecter ni de les chasser de là. ».

Trois ans plus tard, en 1897, le pasteurien Paul-Louis Simond démontra que le bacille pesteux se transmettait par l'intermédiaire des puces (Simond, 1898). Les puces présentent une spécificité d'hôte, même si celle-ci n'est pas absolue. *Xenopsylla cheopis* est la puce du rat, mais elle peut accidentellement piquer l'homme. La puce s'infecte en prélevant du sang sur un animal en phase septicémique. La bactérie se multiplie dans le proventricule de la puce et finit par en obstruer la lumière. Lorsque la puce change d'hôte, le sang ingéré dans le proventricule est régurgité après avoir été en contact avec le bouchon bactérien, ce qui infecte l'animal piqué. Le bacille circule ainsi en permanence dans les populations de rongeurs (Perry *et al.*, 1997).

Les rongeurs peuvent aussi contracter la peste en ingérant le cadavre de leurs congénères infectés par *Y. pestis*. Yersin l'avait également observé : « En faisant ingérer aux animaux soit des cultures soit des fragments de rate ou de foie d'animaux morts de la peste, on tue souvent les souris, presque toujours les rats. » (Yersin, 1894). Des contaminations humaines par ingestion de viande de chameau infecté ont également été rapportées en Jordanie en 1997 (Arbaji *et al.*, 2005).

Les rongeurs peuvent aussi se contaminer en inhalant des bactéries, en particulier lors du fouissage de la terre. En effet, le bacille de la peste peut persister pendant plusieurs semaines dans le

sol (Eisen *et al.*, 2008), voire plusieurs années dans les terriers de rongeurs (Mollaret *et al.*, 1963). Des études expérimentales ont également montré que le bacille pouvait être transmis par un sol contaminé depuis plusieurs semaines (Boegler *et al.*, 2012). La contamination des rongeurs sauvages fouisseurs peut ainsi créer des foyers de résurgence (Perry *et al.*, 1997).

1.2. Les différentes formes de la maladie

La peste humaine résulte le plus souvent de la piqûre accidentelle de l'homme par une puce ayant été au préalable contaminée par un rongeur infecté. La symptomatologie chez l'homme, peut revêtir deux formes majeures : bubonique et pulmonaire.

La forme bubonique est la plus fréquente. Une fois le bacille entré dans l'organisme par voie intradermique, il se dissémine vers le nœud lymphatique drainant. Sa multiplication intense et la réaction inflammatoire qu'il produit aboutissent à la formation d'un bubon, qui est un gonflement très douloureux du nœud lymphatique pouvant aller jusqu'à son ulcération ou sa rupture. L'apparition d'un ou plusieurs bubons est un élément de diagnostic de la forme bubonique de la peste. Une très forte fièvre accompagnée d'une grande fatigue, des céphalées, des troubles digestifs (vomissements et parfois diarrhées) voire des troubles du comportement (agitation, délire ou prostration), accompagnent généralement l'apparition de bubons. Cette forme bubonique est mortelle dans 40 à 70% des cas en moins d'une semaine en l'absence de traitement antibiotique précoce (Demeure *et al.*, 2009; Perry *et al.*, 1997).

Lors d'une peste bubonique, il arrive que la bactérie reste contenue au niveau du bubon et ne se dissémine pas dans l'organisme. La maladie évolue alors vers la fistulisation du bubon et à la guérison spontanée du patient. Mais dans la plupart des cas, la bactérie se dissémine par voie lymphatique et sanguine, créant une forme septicémique de la maladie. Différents organes comme la rate, le foie et éventuellement les poumons présentent de fortes charges bactériennes. La dissémination sanguine est associée à un état de choc ainsi qu'à une coagulation intra-vasculaire disséminée. Cette extension de l'infection à l'ensemble de l'organisme est presque toujours mortelle en l'absence de traitement (Demeure *et al.*, 2009; Perry *et al.*, 1997; Pollitzer, 1960).

Lors de la phase septicémique, la dissémination de la bactérie peut atteindre les poumons et provoquer une peste pulmonaire. Le patient atteint est alors extrêmement contagieux en raison de la production d'aérosols chargés de bactéries. L'inhalation de ces aérosols est responsable d'une infection primaire du poumon et déclenche une forme pulmonaire de la peste qui est la plus virulente. La phase initiale de l'infection (< 24 h) est marquée par une dégradation rapide et générale de l'état du patient, accompagnée d'une forte fièvre, d'une cyanose des extrémités, d'une tachycardie, d'une prostration et de céphalées souvent intenses. Les premiers symptômes pulmonaires apparaissent après 48 h et consistent en une toux intense et douloureuse, une détresse respiratoire et parfois de signes neurologiques. En l'absence d'une antibiothérapie précoce, la mort du patient survient dans 100 % des

cas entre 2 et 4 jours après la contamination. Les aérosols infectés produits durant cette phase sont responsables d'une contamination interhumaine extrêmement efficace (Demeure *et al.*, 2009; Perry *et al.*, 1997).

D'autres formes cliniques, plus rares, peuvent être observées. Une peste septicémique foudroyante est observée lorsque l'infection s'effectue au-niveau d'une blessure cutanée ou de la conjonctive. Le bacille atteint directement la circulation sanguine conduisant à la mort du patient en quelques heures à quelques jours après l'apparition des premiers signes cliniques (Demeure *et al.*, 2009).

1.3. *L'état actuel des foyers de peste*

La peste est souvent considérée comme une maladie du passé mais elle est toujours présente dans certaines régions du monde. Plus de 50 000 cas humains ont été déclarés par l'Organisation Mondiale de la Santé depuis 1990, répartis dans 26 pays (WHO, 2009). Les deux tiers des cas ont été recensés sur le continent africain avec en particulier des foyers endémiques à Madagascar, en République Démocratique du Congo et en Tanzanie (Bertherat *et al.*, 2005; Demeure *et al.*, 2009). Le dernier tiers des cas a été recensé en Asie du Sud-Est (Vietnam, Chine), en Asie Centrale (Mongolie, Kazakhstan, Inde) et sur le continent américain (ouest des Etats-Unis), (Demeure *et al.*, 2009; Ramalingaswami, 1995).

Le risque d'infection de l'Homme dépend à la fois de la prévalence de l'infection parmi les rongeurs, de la promiscuité de ceux-ci avec la population humaine, de l'état des populations de vecteurs et du climat (Figure 20), (Samia *et al.*, 2011).

1.4. *Diagnostic*

En raison de la rapidité d'évolution de la maladie, un diagnostic rapide est l'élément clé pour réduire la mortalité et la morbidité de la peste. Le diagnostic de la maladie chez l'Homme ne peut être affirmé que par l'isolement de *Y. pestis* à partir de ponction ganglionnaire (en cas de peste bubonique), d'expectorations (en cas de peste pulmonaire) ou de prélèvement sanguin (lors de la phase septicémique).

Le diagnostic peut être effectué soit par des techniques bactériologiques classiques (qui nécessitent toutefois au moins 48 h de culture), soit par des techniques rapides comme l'immunochromatographie en bandelettes (qui révèle si le patient est infecté en 15 minutes, très utile sur le terrain car il ne nécessite aucun matériel particulier), (Chanteau *et al.*, 2003 A; Chanteau *et al.*, 2003 B) ou la PCR basée sur l'amplification spécifique du génome de *Y. pestis*, qui permet un diagnostic même si les bactéries prélevées sont mortes (Higgins *et al.*, 1998; Hinnebusch *et al.*, 1993; Loiez *et al.*, 2003; Rahalison *et al.*, 2000; Tomaso *et al.*, 2003).

1.5. *Traitement*

Le traitement de la peste est basé sur l'administration d'antibiotiques. L'OMS recommande d'utiliser de la streptomycine pendant les premiers jours de la maladie, relayée par des sulfamides, des tétracyclines ou du chloramphénicol en monothérapie. Plus récemment, des alternatives ont été proposées comme l'utilisation des fluoroquinolones et de la gentamycine (site CDC <http://www.bt.cdc.gov>). Une antibiothérapie prophylactique est administrée en cas de contact avec un malade pesteux.

1.6. *Vaccination*

L'histoire de la vaccination contre la peste a commencé en 1895, quand Yersin, Calmette et Borrel démontrèrent que le lapin pouvait être vacciné par des injections répétées de microbes tués par la chaleur (Annales de l'Institut Pasteur), (Pollitzer, 1960). En 1934, Girard et Robic ont préparé, à Madagascar, un vaccin à partir d'une souche vivante de virulence atténuée. C'est le vaccin qui a été, et qui est encore, le plus utilisé. Il a permis de diminuer significativement l'incidence de la peste à Madagascar (Girard, 1955). Cependant, la courte durée de protection, un manque d'efficacité sur la forme pulmonaire de la maladie et des effets secondaires parfois sévères ne permettent pas d'utiliser ce vaccin à grande échelle (Demeure *et al.*, 2009; Girard, 1955; Meyer, 1970). Il n'y a pas à l'heure actuelle de vaccin contre la peste disponible pour le public. L'antibiothérapie ayant une très bonne efficacité, elle est devenue le principal mode de traitement à la fois curatif et préventif de la peste (Demeure *et al.*, 2009).

Toutefois, des vaccins acellulaires constitués de protéines recombinantes purifiées sont en cours de développement, les antigènes les plus ciblés étant l'antigène de capsule F1 et la protéine LcrV (Demeure *et al.*, 2009). Une autre stratégie comme l'administration par voie orale d'une souche de *Y. pseudotuberculosis* de virulence atténuée est également à l'étude (Blisnick *et al.*, 2008).

1.7. *Apparition de résistances aux antibiotiques*

Deux souches de *Y. pestis* résistantes à des antibiotiques ont été isolées sur des patients atteints de peste bubonique en 1995 à Madagascar (Guiyoule *et al.*, 2001). La première souche possédait une forte résistance à la streptomycine et la seconde résistait à sept antibiotiques dont quatre recommandés par l'OMS pour les traitements curatifs et prophylactiques (Galimand *et al.*, 2006; Galimand *et al.*, 1997). L'émergence de nouvelles souches multi-résistantes aux antibiotiques pourrait constituer une menace pour la santé publique.

1.8. Bioterrorisme

Y. pestis fait partie des six « armes biologiques » de classe A recensées par le CDC (<http://www.bt.cdc.gov>, le CDC a défini trois classes, la classe A étant réservée aux agents microbiens les plus dangereux). En effet la dispersion d'aérosols de *Y. pestis* pourrait être responsable d'une grande épidémie de forme pulmonaire primaire dans une population urbaine. La bactérie pourrait également être utilisée pour infecter des rongeurs, responsable secondairement d'une dissémination dans la population humaine (Inglesby *et al.*, 2001; Levison, 2000 B). Le risque de bioterrorisme est d'autant plus important et critique que des souches multi-résistantes aux antibiotiques pourraient être utilisées (Hinnebusch *et al.*, 2002 A).

2. Les mécanismes de virulence de *Yersinia pestis*

La première ligne de défense de l'organisme, chez les mammifères, est la réponse immune innée, composée de cellules qui phagocytent les pathogènes et produisent des molécules qui activent la réponse inflammatoire et recrutent les cellules de l'immunité adaptative. Pour survivre et se disséminer, *Y. pestis* utilise toute une gamme de mécanismes et d'effecteurs pour échapper à la réponse immunitaire de son hôte et la déjouer.

Nous décrivons ici les facteurs de virulence de *Y. pestis* et leur mode d'action, car ces facteurs interagissent avec des voies de signalisation de l'hôte (Figures 21 et 22). La perturbation des voies de signalisation cellulaire par *Y. pestis* a été analysée au niveau transcriptionnel dans notre modèle et sera présentée dans la partie « résultats ».

La virulence de *Y. pestis* est avant tout liée à sa capacité à se multiplier rapidement au sein des tissus de l'hôte infecté. Tous les facteurs de virulence de *Y. pestis* concourent à favoriser la réplication extracellulaire de ce pathogène, en inhibant tout particulièrement la réponse immunitaire innée de l'hôte. Le génome de *Y. pestis* est constitué d'un chromosome et de 3 plasmides nommés pPla (ou pPCP1), pFra (ou pMT1) et pYV (ou pCD1), (Perry *et al.*, 1997). Les gènes participant à la virulence de *Y. pestis* peuvent être localisés sur le chromosome bactérien ou être plasmidiques.

Le plasmide pPla (9,5 kb) code une protéase Pla qui possède une activité coagulase à 28°C et fibrinolytique (activateur du plasminogène) à 37°C. Ce plasmide joue principalement un rôle dans la dissémination de *Y. pestis* après son inoculation (Hinnebusch *et al.*, 2002 B). Les souches de *Y. pestis* exemptes de pPla sont incapables de coloniser les nœuds lymphatiques (Sebbane *et al.*, 2006) et pPla est nécessaire au développement de la peste pulmonaire (Lathem *et al.*, 2007).

Le plasmide pFra porte notamment le matériel génétique impliqué dans la production de la pseudo-capsule peptidique constituée de polymères de l'antigène F1. La capsule est synthétisée *in vivo* à 37°C et permet à la bactérie d'acquérir une résistance contre la phagocytose par les cellules de l'hôte

(Du *et al.*, 2002) et contre certains agents antibactériens de l'hôte (Galvan *et al.*, 2008). Néanmoins, des souches dépourvues de capsule sont aussi virulentes que les souches encapsulées (Davis *et al.*, 1996; Drozdov *et al.*, 1995). Ce plasmide porte également le matériel génétique impliqué dans la synthèse de la phospholipase D, anciennement dénommée toxine Ymt (pour *Yersinia Murine Toxin*) qui joue un rôle dans le blocage du provenricule de la puce (Hinnebusch *et al.*, 2002 B).

Le plasmide de virulence pYV confère aux *Yersinia* qui le possèdent une cytotoxicité essentielle à l'expression de leur virulence. Ce plasmide code le système de sécrétion de type III (TTSS), composé de l'appareil de sécrétion Ysc (pour *Yop secretion*) et des toxines Yop (pour *Yersinia Outer Proteins*) transloquées par Ysc dans le cytoplasme des cellules cibles (Cornelis, 2002; Ramamurthi *et al.*, 2002). L'expression des Yop est médiée par le passage de la bactérie à 37-38°C (température de l'hôte mammifère).

L'injection des Yop inhibe la phagocytose et l'induction de la réponse inflammatoire à de multiples niveaux (Ramamurthi *et al.*, 2002). Cette paralysie de la réponse immunitaire innée inhibe également la mise en place d'une réponse adaptative, favorisant la prolifération de la bactérie à l'intérieur de l'hôte (Marketon *et al.*, 2005).

La machinerie de sécrétion, nommée injectisome, est composée d'une vingtaine de protéines Ysc qui permettent la formation d'une aiguille ancrée dans la membrane bactérienne. Les Yop transitent à travers cette aiguille avec l'aide de protéines chaperons. La translocation des Yop dans le cytoplasme de la cellule cible nécessite la présence de YopB, YopD et LcrV qui forment un canal dans la membrane cytoplasmique eucaryote à la suite d'un signal déclenché par YopN, lors du contact avec la cellule eucaryote (Cornelis, 2002 A).

Les protéines effectrices Yop, transloquées dans le cytoplasme des cellules eucaryotes cibles, permettent à *Y. pestis* d'échapper à la phagocytose et donc de rester extracellulaire et permettent aussi de réprimer la réponse inflammatoire.

Au moins quatre Yop sont impliquées dans l'inhibition de la phagocytose, soit en régulant l'activité des Rho GTPases qui contrôlent la polymérisation de l'actine dans ces cellules (YopE, YopT et YopO), soit en inhibant les protéines de l'hôte qui transmettent le signal du récepteur membranaire aux molécules d'actine (YopH), (Cornelis, 2002 B; Li *et al.*, 2008).

YopM est une protéine acide qui forme un complexe avec les sérine/thréonine kinases PRK2 et RSK1 (McDonald *et al.*, 2003) et semble interférer dans la réponse immunitaire en déplaçant les cellules NK (Kerschen *et al.*, 2004a). La déplétion des cellules NK induit une diminution de la synthèse d'interféron gamma (IFN- γ) dont résulte une production réduite de ROS (*Reactive Oxygen Species*), (Kerschen *et al.*, 2004b; Li *et al.*, 2008). YopM agit également au niveau nucléaire en ciblant les gènes du cycle et de la prolifération cellulaire (Li *et al.*, 2008).

YopH est une tyrosine phosphatase qui agit en déphosphorylant des protéines impliquées dans la transmission du signal des cellules eucaryotes. Par ailleurs, la production de cytokines inflammatoires (notamment le TNF- α) et de ROS par l'hôte infecté est en partie inhibée par YopH (Li *et al.*, 2008). De plus, YopH paralyse l'activation des cellules T en induisant un programme d'apoptose. La mort cellulaire a deux effets : les cellules T ne peuvent ni détruire la bactérie ni mettre en place une réponse cellulaire adaptative protectrice (Bruckner *et al.*, 2005).

YopJ inhibe la mise en place de la réponse inflammatoire en bloquant les voies NF- κ B et MAP kinases (Cornelis, 2002 B). Elle induit également l'apoptose des phagocytes via l'activation de caspases et l'inhibition de la voie NF- κ B (Li *et al.*, 2008). De plus, YopJ réduit l'expression de molécules d'adhésion telles que ICAM-1 et E-selectin, dans les cellules endothéliales et de l'épithélium bronchial, inhibant le recrutement des polymorphonucléaires au niveau du site d'infection (Li *et al.*, 2008). De récentes études ont montré que YopJ avait une fonction de sérine/thréonine acétyl-transférase (Paquette *et al.*, 2012). Elle acétyle les sérines et thréonines de diverses MAP2 kinases (*Mitogen-activated protein kinase kinase* ou MAP2K). L'étude de Paquette *et al.* a démontré chez la drosophile que YopJ ciblait entre autres la protéine TAK1 (*TGF β -activated kinase*) aussi appelée MAP3K7, kinase de la voie des MAP kinases qui intervient dans l'activation de NF- κ B (Paquette *et al.*, 2012). Cette inhibition est suffisante pour bloquer la signalisation en aval de TAK1, empêchant l'activation des facteurs de transcription p38, JNK et NF- κ B. L'activité générale d'inhibition de la réponse immunitaire de YopJ résulterait de sa capacité à inactiver TAK1 ainsi que d'autres MAP2 kinases (telles que MKK4, 6 et 7, IKK α et IKK β), (Paquette *et al.*, 2012).

La protéine LcrV (anciennement appelé antigène V) également codée par le plasmide pYV, intervient dans la translocation des Yop à travers la membrane plasmique. De plus, LcrV détourne la voie de signalisation de TLR2/TLR6/CD14 pour déclencher la synthèse de la cytokine anti-inflammatoire IL-10 dans les cellules dendritiques et cellules T, ce qui a pour conséquence d'abolir la réponse pro-inflammatoire et de rendre ces cellules « tolérantes » vis-à-vis du pathogène (DePaolo *et al.*, 2008; Pouillot *et al.*, 2005). DePaolo *et al.* ont montré qu'une souche de *Y. pestis* déficiente pour le facteur LcrV induit une réponse inflammatoire en 24 h après infection sous-cutanée, alors que cette réponse inflammatoire est bien plus tardive avec la souche sauvage (DePaolo *et al.*, 2008).

En plus de ces facteurs de virulence plasmidiques, *Y. pestis* possède d'autres mécanismes pour échapper aux défenses de l'hôte, tels que la capture du fer, la modification structurale du LPS ou la synthèse de la protéine Ail.

Y. pestis partage avec les souches les plus pathogènes de *Y. pseudotuberculosis* et *Y. enterocolitica*, un facteur de virulence essentiel appelé « îlot de haute pathogénicité » (HPI). Cet îlot porte un ensemble de gènes impliqués dans la capture du fer par l'intermédiaire d'un sidérophore spécifique, la yersiniabactine (Perry *et al.*, 1990). La bactérie peut ainsi trouver le fer dont elle a besoin pour sa croissance, malgré le très faible taux de fer libre dans les tissus de l'hôte. L'HPI est situé

dans un locus appelé *pgm* (pour *pigmentation*) car il contient également des gènes conférant à la bactérie le caractère de pigmentation par capture d'hémine. La perte du locus *pgm* provoque une forte atténuation de virulence. Une souche KIM *pgm*⁻ injectée par voie sous-cutanée ou intra-péritonéale est avirulente dans les lignées de souris de laboratoires (DL50>10⁷ ufc) et rapidement éliminée de l'organisme contrairement à la souche sauvage (Une *et al.*, 1984).

Le LPS, composant majeur de la membrane externe des bactéries à Gram-négatif, est reconnu par certains types cellulaires de l'hôte par le récepteur TLR4. Cette interaction permet d'initier la réponse immune. L'hexa-acétylation du lipide A du LPS est nécessaire à une stimulation maximale de la réponse inflammatoire chez les mammifères. Une modification de la structure du LPS diminue la stimulation de TLR4 et donc la réponse cellulaire qui en découle. *Y. pestis* produit un LPS hexa-acétylé (typique) lors de son passage dans la puce (ou lorsque la bactérie est cultivée entre 21 et 26°C). La transition de température à 37°C lors de son passage dans l'hôte mammifère provoque la synthèse d'un LPS tétra-acétylé qui ne stimule pas TLR4 et qui de plus agit comme un antagoniste du LPS hexa-acétylé (Dziarski, 2006; Kawahara *et al.*, 2002; Matsuura *et al.*, 2010; Montminy *et al.*, 2006). Ce changement dans la structure du LPS réduit l'activation des macrophages, la sécrétion de cytokines pro-inflammatoires (telles que TNF, IL-1, IL-6 ou encore IL-8), et l'activation et la maturation des cellules dendritiques nécessaires à l'induction d'une réponse immune adaptative (Dziarski, 2006; Li *et al.*, 2008; Montminy *et al.*, 2006).

Y. pestis présente une résistance à la lyse médiée par les facteurs du complément, ce qui lui permet de survivre et de se multiplier dans le sang, que ce soit dans la puce ou lors de la phase septicémique chez l'hôte mammifère. Cette résistance serait due au moins en partie à la protéine Ail exprimée dans la membrane extracellulaire de la bactérie (Bartra *et al.*, 2008). Cette protéine serait également impliquée dans l'adhésion aux cellules cibles, à la prévention du recrutement des cellules immunitaires dans les nœuds lymphatiques et dans la répression de la réponse inflammatoire (Kolodziejek *et al.*, 2012).

3. L'apport des modèles animaux dans l'exploration de la pathogénèse des différentes formes de la peste

Les rongeurs étant les hôtes naturels de *Y. pestis*, il est pertinent d'étudier le déroulement du processus infectieux dans ces espèces. De plus, la symptomatologie de la peste chez les rats et les souris infectés par des puces est similaire à ce qui est observé dans la forme bubonique chez l'humain. A. Yersin soulignait déjà en 1894 que les rats « présent[ai]ent de véritables bubons. » (Yersin, 1894) comme ce qui est observé chez l'Homme. Des modèles murin de peste pulmonaire primaire ont également montré que la broncho-pneumonie exsudative et purulente observée chez les animaux

infectés était très proche de ce qui est observé chez l'Homme (Lathem *et al.*, 2005; Sebbane *et al.*, 2005). L'étude des modèles rat et souris est à l'origine d'une part importante des connaissances acquises sur les interactions entre *Y. pestis* et son hôte, tant d'un point de vue du déroulement du processus infectieux que d'un point de vue moléculaire. L'intérêt de ces modèles animaux est entre autre de pouvoir mesurer les variations des acteurs de la réponse immune, comme les cytokines, et de les corréler à des lésions histo-pathologiques dans le but de mieux comprendre la hiérarchisation des processus physiologiques mis en place en réponse à l'infection par *Y. pestis*.

3.1. Les modèles de la forme bubonique de la peste

Chez le rat et la souris, la forme bubonique de la peste a été induite par inoculation de la bactérie par voie intradermique dans l'oreille (Sebbane *et al.*, 2005) ou par injection sous-cutanée (Demeure *et al.*, 2012; Gonzalez *et al.*, 2012; Lukaszewski *et al.*, 2005; Nham *et al.*, 2012).

Sebbane *et al.* ont été les premiers à décrire un modèle d'infection intradermique chez le rat. Ils ont montré qu'après injection de 600 ufc de la souche virulente 195/P de *Y. pestis*, les ganglions lymphatiques drainants proximaux étaient rapidement colonisés (6 h) (Sebbane *et al.*, 2005). La bactérie n'est retrouvée dans les ganglions distaux, le sang et la rate qu'au bout de 36 h, puis la charge bactérienne augmente de façon exponentielle dans l'ensemble des organes colonisés jusqu'à 72 h. Les taux de TNF- α et d'IL-10 mesurés dans le sérum entre 48 et 72 h après inoculation sont très faibles alors que le taux d'IFN- γ est élevé et est corrélé avec la charge bactérienne sérique. Ce modèle animal reproduit bien les symptômes observés chez les patients atteints de peste bubonique et peut donc être utilisé pour comprendre la réponse mise en place par l'hôte, notamment durant la phase précoce de colonisation des ganglions lymphatiques. Une dissémination similaire à ce qui est décrit chez le rat a également été observée chez des souris infectées par voie intradermique avec la souche 195/P de *Y. pestis* (Bosio *et al.*, 2012). De plus, Bosio *et al.* ont montré que le recrutement des neutrophiles et des macrophages est inhibé au site d'injection ainsi que dans le nœud lymphatique drainant chez des souris infectées par une souche virulente de *Y. pestis* par comparaison avec une injection de LPS (Bosio *et al.*, 2012). Ces résultats concordent avec ceux du groupe d'Elisabeth Carniel dans un modèle similaire, qui montrent que *Y. pestis* est capable d'infiltrer massivement le nœud lymphatique drainant sans provoquer de réaction organisée des leucocytes polynucléaires (Guinet *et al.*, 2008). Ce n'est qu'après dissémination de la bactérie dans l'organisme entier qu'était observée un recrutement massif de ces cellules, toutefois incapables de contrôler la prolifération bactérienne massive (Guinet *et al.*, 2008).

Les études utilisant l'infection sous-cutanée présentent des résultats de dissémination bactérienne similaire. Deux études utilisant une bactérie bioluminescente ont corroboré ces études (Gonzalez *et al.*, 2012; Nham *et al.*, 2012). Ces études ont l'avantage non négligeable de pouvoir

suivre tout le déroulement de l'infection sur un même animal. L'étude de Nham *et al.* a montré que 74% des souris infectées présentent une dissémination comparable de la bactérie après injection en sous-cutané. La colonisation du nœud lymphatique drainant s'effectuait en moyenne au 3^{ème} jour post-infection. La phase de migration du point d'infection au nœud lymphatique drainant était la plus variable dans la durée. Cependant, Nham *et al.* ont montré qu'une fois le nœud lymphatique drainant (nœud lymphatique inguinal dans cette étude) atteint et colonisé, le déroulement de la dissémination bactérienne était très homogène entre les animaux. La bactérie migrait ensuite jusqu'au nœud lymphatique axillaire en quelques heures avant d'envahir le foie et la rate un jour plus tard (4^{ème} jour). L'ensemble de l'organisme était colonisé au 5^{ème} jour aboutissant à la mort de l'animal au 6^{ème} jour (Nham *et al.*, 2012). Des modes de dissémination atypiques ont été observés (26 % des cas). Pour certains animaux, la bactérie semblait rester localisée au point d'injection jusqu'à la mort de l'animal ou se disséminait directement du point d'injection dans l'ensemble de l'organisme, ce dernier mode de dissémination pouvant être dû à un passage direct de la bactérie dans le sang (Nham *et al.*, 2012).

3.2. Les modèles de la forme pulmonaire de la peste

Lathem *et al.* ont modélisé une infection pulmonaire primaire dans la lignée C56BL/6 à l'aide d'une instillation intra-nasale de la souche virulente CO92 de *Y. pestis* (Lathem *et al.*, 2005). Ils ont observé un syndrome bi-phasique dans lequel l'infection débute par un état anti-inflammatoire dans les premières 24-36 h post-infection, et progresse ensuite rapidement vers un état hyper-inflammatoire (48 h), s'achevant par la mort de l'animal le troisième jour (Lathem *et al.*, 2005). Ces résultats ont été corroborés par l'étude de Bubeck *et al.* qui ont démontré que le recrutement des neutrophiles dans les poumons de souris C57BL/6J et CD1 (non-consanguines) n'apparaissait qu'à partir de 48 h post-infection, accompagné par une production tardive de cytokines (IL-6, TNF α , IFN γ , IL-12p70, MCP-1) et de chimiokines (KC, MIP-2). En comparaison, ces deux types de souris produisaient une forte réponse inflammatoire précoce (< 12 h) à *Klebsiella pneumoniae*, démontrant encore une fois que *Y. pestis* est capable de retarder la réponse inflammatoire (Bubeck *et al.*, 2007). Price *et al.* soulignent également la capacité de *Y. pestis* à établir rapidement un environnement anti-inflammatoire qui est également permissif pour la prolifération d'autres micro-organismes (même habituellement non pathogènes) (Price *et al.*, 2012). Lorsque la bactérie est finalement reconnue par le système immunitaire, la charge bactérienne est telle qu'elle provoque une très forte réponse inflammatoire. Cette forte réponse pourrait même participer au déclin de la fonction pulmonaire (Price *et al.*, 2012).

Agar *et al.* ont également décrit un modèle d'infection pulmonaire, mais chez le rat. Leur étude, recoupant histologie, mesure de la charge bactérienne et mesure d'une vingtaine de cytokines pro- et anti-inflammatoires dans les poumons, le sang, le foie et la rate, montre également une latence avant la mise en place de la réponse pro-inflammatoire. En revanche, la réponse pro-inflammatoire augmente exponentiellement à partir de 48 h. Les cytokines les plus exprimées au niveau protéique étaient G-CSF (*Granulocyte Colony Stimulating Factor*) et KC (*Keratinocyte-derived cytokine*), dans

tous les organes étudiés (Agar *et al.*, 2009). Ce résultat vient corroborer les études transcriptionnelles qui montraient que G-CSF, codé par le gène *Csf3*, était fréquemment surexprimé dans les organes en réponse à une infection par *Y. pestis* (Agar *et al.*, 2009; Comer *et al.*, 2010; Galindo *et al.*, 2009; Liu *et al.*, 2009; Rogers *et al.*, 2007). D'autres cytokines comme MIP1 α (*Macrophage Inflammatory Protein 1 alpha*), MCP-1 (*Monocyte-Chemoattractant Protein-1*), IL-1 α et l'IL-6 étaient significativement élevés 48h après inoculation et augmentaient jusqu'à 72 h. Ces chimiokines et interleukines ont pour rôle de recruter les cellules inflammatoires au niveau des tissus atteints et d'induire la sécrétion d'autres cytokines pro-inflammatoires en réponse à l'infection. Cependant, cette forte réponse inflammatoire ne permettait pas de contrôler la prolifération de la bactérie chez le rat qui succombait entre 72 et 96 h après inoculation.

Plusieurs études ont montré que les neutrophiles jouaient un rôle important dans la l'élimination des bactéries lors d'une infection pulmonaire primaire par *Y. pestis*. Des souris déplétées en neutrophiles circulants présentaient une charge bactérienne significativement plus élevée après infection (Eisele *et al.*, 2011; Laws *et al.*, 2010). De plus, Laws *et al.* ont montré que la co-culture de *Y. pestis* avec des neutrophiles provoquait une diminution de 5 fois du nombre de bactéries, alors qu'il n'était pas significativement modifié lors d'une co-culture de *Y. pestis* avec un mélange de macrophages, monocytes et lymphocytes (Laws *et al.*, 2010).

La dissémination de la bactérie depuis les poumons, site d'infection primaire, jusqu'à l'ensemble des organes a été suivie à l'aide d'une bactérie bioluminescente (Gonzalez *et al.*, 2012). L'étude de Gonzalez *et al.* a montré que, suite à une instillation nasale, *Y. pestis* est d'abord contenue dans les poumons pendant les premières 48 h avant de se disséminer dans tout l'organisme entre 48 et 72 h post-infection, conduisant à la mort rapide de l'animal (Gonzalez *et al.*, 2012).

L'ensemble de ces études, qu'elles soient *in vivo* ou *in vitro*, dans un modèle pulmonaire ou bubonique de la peste confirment qu'une caractéristique majeure de l'infection par *Y. pestis* est l'inhibition précoce des défenses de l'hôte (phagocytose, production de cytokines, de ROS, recrutement des cellules de l'immunité innée puis adaptative, etc.) jusqu'à une phase tardive de l'infection où la bactérie est en nombre massif et provoque alors une réponse inflammatoire forte (Bubeck *et al.*, 2007; Demeure *et al.*, 2012; Guinet *et al.*, 2008; Lukaszewski *et al.*, 2005; Price *et al.*, 2012). Ces observations sont en accord avec les études transcriptomiques qui ont été effectuées *in vivo* (Comer *et al.*, 2010; Galindo *et al.*, 2009; Liu *et al.*, 2009; Rogers *et al.*, 2007).

4. La recherche de facteurs génétiques déterminant la sensibilité de l'hôte à une infection par *Y. pestis*

Le modèle murin a aussi été utilisé pour identifier des facteurs génétiques de l'hôte contrôlant la vulnérabilité face à l'infection. Turner *et al.* ont montré que les souris BALB/cJ étaient résistantes à l'injection par voie intraveineuse d'une souche de *Y. pestis* KIM5 *pgm*⁻ (souche de virulence très atténuée) contrairement aux souris C57BL/6J (Turner *et al.*, 2009). Les souris BALB/cJ présentaient une charge bactérienne dans la rate et le foie significativement inférieure aux souris C57BL/6J ainsi qu'une absence de thrombose vasculaire fibrineuse et un taux sanguin plus faible d'IL-6 (Turner *et al.*, 2008). Un QTL sur le chromosome 17, dénommé *Prll* (pour *Plague locus resistance 1*), a été identifié dans une population F2 issue d'un croisement C57BL/6J × BALB/c et confirmé par un *backcross* (Turner *et al.*, 2008). Ce QTL a été identifié comme contrôlant la charge bactérienne dans la rate 48 h après inoculation. Les auteurs ont par la suite produit des souris congéniques pour ce locus sous forme allélique BALB/cJ dans un fonds sensible C57BL/6J. La lignée congénique présentait une charge bactérienne splénique 48 h post-inoculation identique à la lignée parentale BALB/cJ, confirmant l'existence de *Prll*. Ce locus comprend entre autres les gènes codant le Complexe Majeur d'Histocompatibilité (CMH), et de nombreux autres gènes impliqués dans la réponse immunitaire, comme des gènes codant des facteurs du complément ou des membres de la famille du TNF (Turner *et al.*, 2008). A cause de son rôle central dans la réponse immune, cette région a été souvent associée à des différences de sensibilité à divers pathogènes tels que *Plasmodium berghei*, *Chlamydia pneumoniae* ou encore *Streptococcus pyogenes* (Goldmann *et al.*, 2005; Goncalves *et al.*, 2008; Min-Oo *et al.*, 2008).

Plusieurs études ont montré que certaines sous-lignées de souris 129 sont résistantes à une injection intraveineuse d'une souche de *Y. pestis* *Pgm*⁻ (KIM5) mais sont sensibles à une infection par une souche virulente CO99-3015 (Congleton *et al.*, 2006). Philipovski *et al.* avaient montré en 2005 que des souris *knock-out* pour l'*IL-10* (*IL-10*^{-/-}) étaient résistantes à une infection par la souche KIM5 de *Y. pestis* (Philipovski *et al.*, 2005). Néanmoins, Turner *et al.* ont mis en évidence quelques années plus tard, que ce phénotype de résistance était en partie conféré par un allèle dominant d'origine 129/P2J provenant d'une région de 30 cM autour du gène *IL-10* hérité du fonds génétique des cellules embryonnaires souches utilisées pour la délétion de ce gène (Turner *et al.*, 2009).

Par ailleurs, une étude d'association portant sur le gène candidat *Ccr5* (*Chemokine receptor 5*) qui a été menée sur les populations de rats sauvages à Madagascar suggère qu'un polymorphisme dans une région conservée de ce gène pourrait être lié à une différence de sensibilité à *Y. pestis* entre une population de rats sauvages relativement résistants à l'infection et une population de rats sensibles (Tollenaere *et al.*, 2008).

A l'heure actuelle, la seule lignée de souris décrite comme résistante à une souche virulente de *Y. pestis* est la lignée SEG/Pas (Blanchet *et al.*, 2011). Les travaux réalisés par notre laboratoire portant sur les facteurs génétiques de résistance à *Y. pestis* dans la lignée SEG/Pas, ainsi que les travaux détaillant la réponse physiopathologique de cette lignée comparée à la lignée sensible C57BL/6J, sont détaillés ci-après.

Partie 7. Etat des connaissances sur la résistance à la peste de la lignée SEG/Pas

La recherche de facteurs génétiques de résistance à la peste chez la Souris a débuté dans notre laboratoire avec le travail initié par Marek Szatanik en 2006. Il a tiré parti de la collection de lignées consanguines de l'Animalerie Centrale de l'Institut Pasteur provenant de différentes espèces et sous-espèces de *Mus* dont *M. m. musculus*, *M. m. domesticus* et *Mus spretus* pour maximiser la probabilité d'identifier des différences de vulnérabilité à la souche virulente CO92 de *Y. pestis*. Il a montré que les souris de la lignée SEG/Pas (dérivées de *Mus spretus*) résistaient à 90% à une infection par environ 100 ufc de bactéries injectées par voie sous-cutanée alors que toutes les lignées de laboratoire testées mouraient, dont la lignée C57BL/6J. SEG/Pas était la première lignée de souris décrite comme résistante à l'inoculation d'une souche virulente de *Y. pestis* par voie sous-cutanée (Blanchet *et al.*, 2011). D'autres études avaient rapporté des différences de vulnérabilité entre lignées de souris, mais elles avaient utilisé des souches bactériennes atténuées (Turner *et al.*, 2008; Turner *et al.*, 2009).

La lignée SEG/Pas a été constituée en réalisant des croisements consanguins à partir d'individus de l'espèce *Mus spretus* capturés en Espagne (Spretus Espagne Grenade). Cette espèce a divergé de l'espèce *Mus musculus* depuis environ 1,5 M d'années (Guenet *et al.*, 2003), accumulant un polymorphisme génétique très élevé. Des analyses de séquence et de SNP indiquent qu'il y aurait un variant toutes les 50 à 130 paires de base (Mahler *et al.*, 2008; Zhang *et al.*, 2005b), contre un toutes les 250 à 300 paires de bases entre les lignées de laboratoire (Frazer *et al.*, 2007). Ce polymorphisme génétique se traduit par de nombreuses différences phénotypiques par rapport aux lignées de laboratoire (Dejager *et al.*, 2009; Mott *et al.*, 2008) et a permis de disséquer l'architecture génétique de la sensibilité à différentes maladies infectieuses ou au LPS (Dejager *et al.*, 2010a; Dejager *et al.*, 2010b; Turcotte *et al.*, 2006). Plusieurs lignées consanguines ont été produites à partir d'individus capturés en différents points de la région de distribution géographique de l'espèce : SPRET/Ei, SEG/Pas, SEB/Pas (Espagne), STF/Pas (Tunisie), SMZ/Mpl (Maroc), SFM/Mpl (France) (Dejager *et al.*, 2009). Plusieurs de ces lignées ont été testées au laboratoire et ont montré des différences de vulnérabilité à *Y. pestis* (J. Jaubert, résultats non publiés).

Néanmoins, l'utilisation des souris *Mus spretus* présente des inconvénients. Ces lignées se reproduisent difficilement et donnent de petites portées, certaines sont peu dociles voire agressives, les mâles F1 obtenus par croisement avec des lignées de laboratoire sont stériles, interdisant la production de F2, et des incompatibilités génomiques peuvent empêcher l'obtention de certains génotypes (Blanchet *et al.*, 2011; Burgio *et al.*, 2007; Chevallier *et al.*, 2012; Dejager *et al.*, 2009; Guenet *et al.*, 2003; L'Hote *et al.*, 2010). En outre, la grande densité de polymorphismes entre les lignées constituées à partir de l'espèce *Mus spretus* et les lignées de laboratoire peut rendre difficile l'identification de différences de séquence causales lors de la recherche de gène fonctionnel (DiPetrillo *et al.*, 2005).

Les résultats préliminaires obtenus par Marek Szatanik ont servi à initier un programme collaboratif entre plusieurs unités de l'Institut Pasteur (*Yersinia*, Cytokines & Inflammation, Immunophysiologie & Parasitisme Intracellulaire et Histo-technologie & Pathologie) afin d'explorer le contrôle génétique et les caractéristiques physiopathologiques de la différence de réponse à *Y. pestis* des lignées C57BL/6J (B6) et SEG/Pas (SEG).

Au cours de sa thèse, Charlène Blanchet a entrepris la recherche des facteurs génétiques impliqués dans la résistance à la peste chez SEG. Ses résultats ont donné lieu à une publication dans *Genes and Immunity* (Blanchet *et al.*, 2011). Elle a confirmé que les souris femelles SEG survivaient à hauteur de 90% à une infection par 100 bactéries de la souche virulente CO92 de *Y. pestis*, inoculées par voie sous-cutanée. Les femelles des lignées FVB/N, BALB/cByJ, SLJ/J et CAST/Ei ne résistaient qu'entre 0 et 23% et les femelles B6 résistaient à 6,3%. La DL50 calculée était de 3,2 ufc chez B6 alors qu'elle était supérieure à 10^7 chez SEG. Les femelles SEG résistaient également à l'infection par une autre souche virulente (6/69) de *Y. pestis*, suggérant que le phénotype de résistance n'était pas spécifique de la souche bactérienne CO92. Les femelles SEG étant significativement plus résistantes que les mâles à l'infection, l'analyse génétique a été restreinte aux femelles.

Un *backcross* a été réalisé entre SEG et B6, produisant une population BSB (pour (C57BL/6 × SEG)F1 × C57BL/6), afin de rechercher des locus de résistance. L'analyse de 322 femelles BSB a mis en évidence l'existence de 3 QTL localisés sur les chromosomes 3, 4 et 6 contrôlant 70% de la différence de taux de survie entre les lignées parentales. Ces QTL ont été appelés respectivement *Ypr11*, *Ypr12* et *Ypr13* pour *Yersinia pestis* *resistance* *locus*. Dans les trois cas, l'allèle SEG conférait une résistance accrue sur un mode dominant.

L'analyse des souris BSB portant des combinaisons de deux ou trois des locus *Ypr1* a montré que les trois QTL agissaient de manière additive sur la survie des souris. Chaque QTL augmentait la survie d'environ 20% et les femelles BSB hétérozygotes pour les trois locus présentaient un taux de survie identique à celui de la lignée parentale SEG. Les femelles BSB ne portant aucun des trois QTL survivaient significativement plus que les B6 (27 % versus 6,3 %), laissant supposer que d'autres facteurs génétiques, hors des régions identifiées, participaient aussi au contrôle du taux de survie.

Charlène Blanchet a également analysé 217 mâles BSB et détecté un unique QTL co-localisé avec *Ypr13*. Comme chez les femelles, ce QTL contrôle uniquement la survie. Les mâles hétérozygotes B/S au locus *Ypr13* présentent un taux de survie comparable à celui des mâles SEG. Les individus BSB homozygotes B/B pour *Ypr13* ont un taux de survie significativement supérieur à celui des mâles B6 (48,5 % versus 6%), ce qui indique l'existence d'autres gènes impliqués dans la résistance à *Y. pestis*.

Aucun autre QTL n'a été mis en évidence chez les mâles, que ce soit pour le taux de survie ou pour la latence. En particulier, les locus *Yprl1* et *Yprl2* n'ont pas été retrouvés et semblaient donc spécifiques des femelles dans l'analyse de ce *backcross*. Ces données ont été reprises dans l'article présenté dans la partie « résultats ».

En parallèle de l'analyse génétique, un projet mené par Christian Demeure de l'unité des *Yersinia* a comparé les mécanismes physiopathologiques mis en place pendant l'infection chez les souris SEG et B6 (Demeure *et al.*, 2012). Les résultats ont montré que les deux lignées de souris étaient efficacement colonisées par *Y. pestis*. Toutefois, chez les souris SEG, la bactériémie augmentait rapidement pour atteindre un pic au 2^{ème} jour, puis diminuait jusqu'à s'annuler au 4^{ème} jour. Chez les souris B6, elle augmentait progressivement pour atteindre, au 4^{ème} jour, le même niveau que SEG au 2^{ème} jour, mais elle continuait d'augmenter et menait à la mort des souris dans les jours suivants. La charge bactérienne dans le foie, la rate et les ganglions lymphatiques était comparable chez B6 et SEG durant les trois premiers jours, puis régressait significativement le 4^{ème} jour chez SEG, alors qu'elle continuait à augmenter chez B6.

Une augmentation significative du nombre de leucocytes sanguins a été mise en évidence chez les deux lignées en réponse à l'infection. Cependant au 4^{ème} jour, le nombre de leucocytes augmentait significativement chez SEG par rapport à B6, chez qui la population leucocytaire restait stable. Le même type d'évolution a été observé pour les plaquettes sanguines. Une étude plus précise de la population leucocytaire montra que les lymphocytes, les monocytes et les polynucléaires (PN) augmentaient dans les deux lignées de souris durant les deux premiers jours de l'infection. Néanmoins, une augmentation significative des PN, monocytes et lymphocytes était notée chez SEG au 4^{ème} jour qui n'avait pas lieu chez B6. Afin de savoir si cette différence était due à une hématopoïèse plus forte chez SEG, les auteurs ont comparé les populations cellulaires de la moelle osseuse de SEG et de B6, avant et au cours de l'infection. Les populations cellulaires hématopoïétiques étaient similaires avant infection et jusqu'au 4^{ème} jour post-infection, à partir duquel une augmentation significative des populations de PN et de monocytes était observée chez SEG, suggérant une hématopoïèse plus efficace chez SEG à un stade avancé de l'infection.

Une étude histologique a montré que les souris SEG présentaient moins de lésions de la rate que B6, et qu'elles étaient capables de circonscrire les foyers bactériens. Alors que des lésions nécrotiques s'étendaient chez B6, des micro-abcès infiltrés de PN séquestraient des bactéries chez SEG. Afin de mieux comprendre la différence histologique observée dans la rate, les populations cellulaires spléniques ont été analysées avant infection et deux jours après infection. Les souris SEG présentaient un plus grand nombre de macrophages résidant dans la pulpe rouge de la rate avant infection et cette population se développait considérablement durant l'infection. Ces résultats

suggéraient qu'une différence physiologique constitutive entre SEG et B6 pouvait être à l'origine de la réaction cellulaire précoce et massive observée chez les souris SEG.

Le dosage de dix cytokines impliquées dans la réponse immune a mis en évidence que G-CSF (*Granulocyte colony-stimulating factor*), KC (*Keratinocyte-derived chemokine*), MCP-1 (*Macrophage cationic peptide-1*), IL-1 α et IL-6 présentaient des différences significatives de production entre les deux lignées dans le plasma, la rate et les ganglions lymphatiques inguinaux. L'évolution de la production de ces cytokines chez SEG était caractérisée par un pic au 2^{ème} jour dans le plasma et les ganglions lymphatiques drainant le site d'injection et par un pic au 3^{ème} jour dans la rate, puis la production déclinait à partir du 4^{ème} jour. Ces cytokines étaient significativement moins produites chez B6 aux 2^{ème} et 3^{ème} jours, mais augmentaient progressivement, pour atteindre au 4^{ème} jour un niveau similaire voire plus élevé par rapport au pic observé au 2^{ème} jour chez SEG. L'infection semblait caractérisée par un pic cytokinique précoce et transitoire chez SEG alors qu'il était plus tardif mais soutenu chez B6. Une corrélation était observée entre le niveau de ces cytokines et la charge bactérienne dans les trois organes étudiés aussi bien chez SEG que chez B6 (schéma récapitulatif figure 23).

Ces résultats indiquent que SEG et B6 ont la capacité de développer la même réponse cytokinique lorsqu'elles sont exposées à une charge bactérienne similaire. La différence majeure entre SEG et B6 ne réside pas dans la qualité de la réponse cytokinique mais dans la chronologie de cette réponse. De manière inattendue, la dissémination précoce de *Y. pestis* chez SEG pourrait être critique pour le déclenchement de la réponse immunitaire précoce et efficace qui est observée dans cette lignée.

En conclusion, la recherche simultanée de facteurs génétiques et d'éléments physiopathologiques associés à la différence de résistance à la peste entre les lignées SEG et B6 a fourni des informations complémentaires. Les régions génomiques identifiées permettront, à condition d'être confirmées et réduites, d'établir une liste de gènes candidats par leur position. Les différences physiopathologiques permettent d'établir un second crible. En se basant sur des critères fonctionnels, ces différences permettent de sélectionner, parmi des listes de plusieurs centaines de gènes, ceux dont le rôle dans des voies de signalisation ou des processus physiologiques les désigne comme des candidats de choix.

L'objectif de mon travail de thèse était d'une part de confirmer l'existence des trois QTL identifiés dans le croisement BSB et de mieux en définir la structure, d'autre part d'identifier des différences de niveau d'expression associées aux différences de résistance entre les deux lignées pour établir un troisième crible de sélection des gènes candidats.

OBJECTIFS DU PROJET

« Peu de jours après la mort des rats
Les hommes tombèrent aussi
Comme des murs qui s'écroulent. »
Shih Toanen, 1792

Les objectifs de mon travail de thèse

Malgré le fléau qu'a représenté *Y. pestis* au cours de l'histoire humaine et la menace qu'elle constitue toujours pour la santé publique dans certains pays, les mécanismes de défense de l'hôte ciblés par cette bactérie sont loin d'être tous élucidés. Nous ne savons pas exactement comment *Y. pestis* réussit à déjouer le système immunitaire inné et adaptatif de l'hôte, ni quels sont les gènes qui participent aux variations de vulnérabilité que l'on observe que ce soit dans une population humaine ou entre les différentes lignées de souris.

Les travaux de notre laboratoire ont pour but d'identifier ces facteurs génétiques. C'est pourquoi, dans la continuité du travail réalisé par Charlene Blanchet, j'ai tout d'abord cherché à confirmer, à l'aide de lignées congéniques, l'existence des trois QTL qu'elle avait mis en évidence sur la base d'analyses statistiques. Puis j'ai tenté de réduire ces régions génomiques encore très larges et de les disséquer afin de comprendre leur architecture génétique, à l'aide de lignées congéniques et sous-congéniques.

Par ailleurs, dans l'objectif de déceler des mécanismes biologiques impliqués dans la différence de résistance entre les deux lignées que nous étudions, j'ai entrepris l'analyse du profil transcriptomique de macrophages exposés *ex vivo* à la bactérie.

L'objectif de cette démarche associant des analyses génétique et transcriptionnelle était de cibler des gènes candidats, sur la base de leur localisation génomique, et de leur appartenance aux voies de signalisation cellulaires différenciellement activées entre la lignée sensible et la résistante.

MATERIELS ET METHODES

« Si tu vois un rat tomber
d'un toit et errer comme un
homme ivre, sauve-toi car
la peste est à portée de ta
main. »

Baghavata Purana, texte
sacré indien

Matériels et méthodes

1. Procédures expérimentales utilisées dans l'analyse complémentaire des lignées congéniques

Seuls sont présentés ici les matériels et méthodes qui ne figurent pas dans l'article (voir section « résultats »).

Analyses statistiques

Le logiciel J/QTL version 1.3.1 a été utilisé pour la recherche de QTL chaque jour après infection (3 à 14). Le taux de survie a été analysé comme un caractère quantitatif. Les seuils de significativité à 5% et 10%, calculés par 1000 permutations des données, étaient de 2,77 et 2,43, respectivement.

Les proportions des souris survivantes ou mortes à chaque jour post-infection pour chaque QTL ont été comparées par un test de Chi2 (seuil de significativité à 5%), à l'aide du logiciel Prism 5.0 (GraphPad, La Jolla, CA).

2. Procédures expérimentales utilisées dans l'analyse comparative de transcriptome

2.1. Animaux

Les souris SEG/Pas, bi-congéniques $cg2^{B/S}$ - $cg3^{B/S}$ ainsi que le lot de souris femelles C57BL/6J utilisées dans l'expérience de qRT-PCR ont été élevées dans l'animalerie centrale de l'Institut Pasteur. Le lot de souris C57BL/6J utilisées dans l'expérience de *micro-array* ont été achetées à l'âge de 6 semaines chez Charles River (L'Arbresle, France) et ont été acclimatées pendant 3 à 4 semaines dans l'animalerie centrale de l'Institut Pasteur.

Toutes les souris utilisées pour l'extraction des macrophages péritonéaux pour les expériences de puces à ADN et de qRT-PCR étaient des femelles âgées de 8 à 12 semaines.

2.2. Culture de la bactérie

La souche virulente CO92 de *Y. pestis* a été cultivée à 28°C sur des boîtes gélosées Luria-Bertani supplémentées avec 0,002% d'hémine (LBH) pendant 48 heures, suivi d'une culture sur la nuit dans un milieu liquide LB à 37°C sous agitation. La culture de *Y. pestis* a été réalisée dans un laboratoire P3.

2.3. Prélèvement de macrophages, culture cellulaire et exposition des macrophages à *Y. pestis*

Les souris ont été euthanasiées par dislocation cervicale. Les cellules péritonéales ont été collectées en injectant 5 mL de RPMI par voie intrapéritonéale et par ré-aspiration du liquide. Les cellules ont été comptées et ajustées à la concentration de 1×10^6 cellules/mL (1 mL/puits). Après 1h d'incubation à 37°C + 5% de CO₂, les cellules non adhérentes ont été éliminées par lavage avec 2 ml de RPMI. Les cellules adhérentes consistaient en une population cellulaire de 100% de cellules F4/80+, exprimant CD11b à des intensités variables (Pachulec *et al.*, en préparation). Ces cellules adhérentes ont été cultivées en milieu RPMI contenant 5% de sérum de veau fœtal (SVF).

Ces cellules ont soit été exposées avec *Yersinia pestis* à 10 MOI (*Multiplicity Of Infection*, ratio du nombre de bactéries par rapport au nombre de cellules cibles, on compte ici 10 bactéries pour un macrophage) soit laissées non-exposées. Le temps d'exposition était de 3 heures dans l'expérience de puces à ADN et de 1h30, 3h ou 4h30 dans l'expérience de qRT-PCR.

Deux expériences indépendantes ont été réalisées pour l'expérience de puces à ADN et de RT-PCR quantitative.

Dans l'expérience de puces à ADN, six femelles de chaque lignée (C57BL/6J et SEG/Pas) ont été utilisées. Le liquide de lavage péritonéal a été poolé pour les six animaux d'une même lignée. Trois puits ont été attribués par condition (exposé 3 heures / non-exposé à *Y. pestis*) et par lignée, soit $3 \times 2 \times 2 = 12$ échantillons.

Dans l'expérience de qRT-PCR, quatre femelles de chaque lignée (C57BL/6J, SEG/Pas et $cg2^{B/S}$ - $cg3^{B/S}$) ont été utilisées. Le liquide de lavage péritonéal a été poolé pour les quatre animaux d'une même lignée. Trois puits ont été attribués par condition (non-exposé, exposé 1h30, 3h et 4h30) et par lignée, soit $3 \times 4 \times 3 = 36$ échantillons.

2.4. *Extraction des ARN*

Après exposition, les cellules adhérentes ont été récoltées par trypsinisation et l'ARN a été isolé à l'aide du kit RNeasy Plus Mini Kit (Qiagen) selon les instructions du fournisseur, incluant un traitement à la DNase afin d'éliminer d'éventuelles traces d'ADN génomique. Une étape supplémentaire d'homogénéisation a été réalisée à l'aide du kit Qias shredder (Qiagen). La concentration des ARN a été mesurée à l'aide du Nanodrop 1000 (Thermo Scientific). Leur qualité a été vérifiée en faisant migrer 2 μ L de solution sur un gel d'agarose ultra-pur à 1%.

2.5. *Expérience de puces à ADN*

500 ng d'ARN totaux ont été utilisés pour la réaction de marquage par le fluorochrome Cy-3 en utilisant le protocole « one color Quick Amp Labeling » (Agilent Technologies, Waldbronn, Germany). L'ARNc marqué a été hybridé à des puces de type Agilent 4x44k mouse_GeneChips (Agilent) pendant 16 heures à 68°C. Les 12 échantillons ont été hybridés sur 3 puces. Les puces ont été scannées par le scanner « Agilent DNA Microarray Scanner ». Les valeurs d'expression ont été calculées par le *package* Feature Extraction 10.5.1.1 (Agilent Technologies; Waldbronn, Germany). L'expérience de puces à ADN a été réalisée par la plateforme de transcriptomique du Helmholtz Institute, Braunschweig, Allemagne.

2.6. *Analyse des données de puces à ADN*

Les données ont été analysées avec le logiciel R, plusieurs *packages* de *Bioconductor* (Gentleman *et al.*, 2004) ont été utilisés incluant les *packages* Agi4x44PreProcess, affycoretools, annotate, RankProd, GOstats, SPIA, KEGGSOAP, Cairo, psych, gplots, RColorBrewer, mgug4122a.db et gtools. Le prétraitement des données a inclus des étapes de correction du bruit de fond (qui estime la portion du signal due au bruit de fond et le soustrait en conséquence), (Ritchie *et al.*, 2007), de normalisation par les quantiles (normalisation intra- et inter-puces, permet la comparaison entre des valeurs d'expression provenant de différentes puces), de récapitulation des

différentes sondes ciblant un gène à la médiane de l'ensemble (*probe summarization*) et une transformation log2. La valeur médiane des trois réplicats a été attribuée à chaque gène. Sur un total de 44 000 sondes, 34 196 ont passé les critères de qualité, correspondant à 20 837 gènes. Une analyse en composantes principales sur l'ensemble des 20 837 gènes a montré que les réplicats étaient très homogènes et les groupes d'échantillons se distinguaient clairement entre les lignées et le statut (exposé/non-exposé).

Les gènes différentiellement exprimés (DE) ont été définis comme des gènes qui présentaient une variation du niveau d'expression minimale d'un facteur 2 et des changements significatifs par rapport au témoin non-infecté dépassant le seuil de FDR (*false-discovery rate*) de 10% (Breitling *et al.*, 2004; van Iterson *et al.*, 2009).

Une analyse de variance (ANOVA) à deux facteurs croisés (facteur « lignée » et facteur « exposition à la bactérie ») a été appliquée aux 20 837 gènes. Il permet de calculer les *P*-value associées à l'effet propre de la lignée, à l'effet propre du statut (exposé ou non à la bactérie) et une troisième valeur, nommée *P*(interaction) qui reflète l'importance de l'interaction (non additive) de ces deux facteurs. Les gènes différentiellement régulés (DR) ont été sélectionnés sur les critères suivants : *P*(interaction) < 0,01 et facteur de variation du niveau d'expression entre les conditions exposées et non-exposées d'un minimum de 2 (*fold change* ou FC).

Afin de caractériser fonctionnellement les gènes DE et DR, GOstats a été appliqué pour déterminer la sur-représentation des *Gene Ontology terms* (Ashburner *et al.*, 2000; Falcon *et al.*, 2007), ainsi que l'algorithme SPIA (*Signaling Pathway Impact Analysis*) (Tarca *et al.*, 2009) qui a permis d'examiner les voies de signalisation de KEGG pathways (*Kyoto Encyclopedia of Genes and Genomes*) (Kanehisa *et al.*, 2000; Kanehisa *et al.*, 2012). Le logiciel IPA (*Ingenuity Pathways Analysis*, IngenuitySystems) a également été utilisé.

L'analyse de la sur-représentation des classes fonctionnelles dans un groupe de gènes a été effectuée à l'aide de DAVID (*Database for Annotation, Visualization and Integrated Discovery*, version 6.7) (Huang da *et al.*, 2009 A; Huang da *et al.*, 2009 B) et AmiGo *Term enrichment* (version 2012) (Boyle *et al.*, 2004).

2.7. RT-PCR quantitative

2.7.1. Synthèse d'ADNc

Les réactions de RT-PCR ont été réalisées sur 0,6 µg d'ARN extraits dans les conditions citées ci-dessus. L'étape de RT-PCR a été réalisée à l'aide du kit RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) suivant les instructions du fournisseur.

2.7.2. Pré-amplification des ADNc

Nous avons réalisé une étape de pré-amplification des ADNc à l'aide du Taqman PreAmp Master Mix (Invitrogen) et des *primers* spécifiques des 47 gènes sélectionnés (44 gènes cibles et 3 contrôles, Tableau 1). Les *primers* ont été conçus sur Primer-BLAST (Primer-BLAST <http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Le tableau 1 présente la liste des couples de *primers* utilisés. L'efficacité de PCR de l'ensemble des 47 couples de primers utilisés était de 100% (+/-5%).

Les ADNc ont été pré-amplifiés sur 14 cycles de PCR, suivi d'un traitement à l'exonucléase I (Fermentas).

2.7.3. PCR quantitative en temps réel

La réaction de PCR quantitative en temps réel a été réalisée à l'aide d'une puce nanofluidique (48.48 *dynamic array*, Fluidigm) permettant l'analyse simultanée de 48 échantillons pour 48 gènes sur un thermocycleur BioMark™ HD Real-time PCR System (Fluidigm, San Francisco, Calif). Nous avons utilisé un protocole utilisant l'EVA Green Master Mix (Fisher Molecular Biology) et des *primers* spécifiques pour chaque gène (Tableau 1).

Le prétraitement des données et l'analyse de la qualité de réplicats ont été évalués à l'aide du logiciel *Fluidigm Real-Time PCR Analysis* (Fluidigm). La spécificité d'amplification a été confirmée par l'évaluation des courbes de fusion à l'aide du logiciel *Fluidigm Melting Curve Analysis* (Fluidigm).

2.7.4. Analyse des résultats

L'expression de 44 gènes cibles et 3 contrôles (*Gapdh*, *Rpl7* et *Ald*), dont l'expression ne variait ni entre les lignées ni en fonction du statut dans l'expérience de puces à ADN, a été mesurée (Tableau 1). Le meilleur gène de référence a été choisi, sur des critères d'homogénéité d'expression dans les différentes conditions testées, en employant le logiciel Genorm (Vandesompele *et al.*, 2002). Les valeurs de Ct ont été converties en expression normalisée relative (par rapport à la médiane de l'expression du gène *Rpl7* dans la condition de référence) par la méthode des $2^{-\Delta\Delta CT}$ (Livak *et al.*, 2001). La médiane des valeurs d'expression a été retenue pour chaque gène.

RESULTATS

« Me voilà de nouveau à Saigon mais pas longtemps car dans douze jours, je dois partir pour le Tonkin afin d'aller de là étudier la peste au Yunnan ! Tu vois que je ne perds pas mon temps. J'ai abouti à Tourane le 10 mai et de là je suis rentré à Saigon par le paquebot. Pendant mon absence, des ordres étaient venus de Paris afin de m'envoyer au Yunnan pour étudier la peste. Cette terrible maladie existe à l'état sporadique au Yunnan, avec une violente épidémie chaque année. Le Tonkin est menacé, c'est pour cela qu'on m'envoie là-bas. L'épidémie s'est étendue cette année en Chine. A Hong Kong, il y a quarante décès par jour. J'ai donc l'idée d'aller d'abord là-bas pour commencer mon travail.[...] La première chose que j'aurai à faire dans l'étude de la peste sera de rechercher son microbe. »

Alexandre Yersin, 1894, dans une lettre à sa mère.

Résultats

Introduction

Afin de confirmer l'implication des trois QTL identifiés dans le croisement BSB, Charlène Blanchet avait débuté le développement de lignées congéniques portant chacune l'un des QTL identifiés sous forme SEG/Pas, dans un fonds génétique sensible C57BL/6J. Pour ce faire, elle a croisé en retour la lignée SEG/Pas avec la lignée C57BL/6J et a sélectionné par génotypage, à chaque génération, les souris hétérozygotes pour une large région génomique contenant l'intervalle de confiance du QTL. Lorsque j'ai pris la suite de son travail, les lignées congéniques pour *Yprl1* et *Yprl3* avaient atteint la 6^{ème} génération et *Yprl2* la 5^{ème} génération.

Le premier volet de mon travail a consisté à continuer ces croisements jusqu'à la dixième génération, à fixer les lignées congéniques à l'état homozygote quand cela était possible, et à produire des lignées bi-congéniques, tri-congéniques et sous-congéniques. Ces lignées ont ensuite été caractérisées pour leur résistance à une infection par *Y. pestis*. Les résultats de ces expériences ont fait l'objet d'une publication, sous presse, dans la revue *Genes and Immunity*.

Le deuxième volet de mon projet de thèse portait sur la comparaison du transcriptome des macrophages péritonéaux de SEG et de B6 en réponse à une exposition *in vitro* à *Y. pestis*. Les résultats de ce travail sont présentés dans une deuxième partie et feront l'objet d'une publication en cours d'écriture.

Partie 1. L'analyse des lignées congéniques permet de disséquer l'architecture de la résistance à la peste chez SEG/Pas

1. La résistance à la peste chez SEG/Pas requiert l'action combinée d'au moins quatre facteurs génétiques

L'article présenté ci-après décrit l'architecture génétique de la résistance à une infection par *Y. pestis* chez SEG.

Nous montrons que, pris indépendamment, aucun des trois QTL identifiés ne permet d'augmenter le niveau de survie de souris sensibles. En revanche, nous montrons qu'une fois combinés *Yprl2* et *Yprl3* augmentent significativement le niveau de survie des souris, aussi bien chez les femelles que chez les mâles. En revanche, nous n'avons pas détecté d'effet de *Yprl1*, seul ou associé à *Yprl2* ou *Yprl3*. L'association détectée dans le *backcross* pourrait être une fausse association, ou d'autres facteurs génétiques présents chez les souris, BSB mais perdus chez les congéniques, pourraient être nécessaires à l'observation de son effet.

La production de lignées sous-congéniques pour les locus *Yprl2* et *Yprl3* m'a permis de mettre en évidence qu'au moins quatre facteurs génétiques présents dans ces régions étaient requis pour augmenter la survie des souris. Enfin, le croisement des lignées congéniques avec la lignée parentale SEG m'a permis de déterminer le mode de transmission des allèles de résistance présents chez SEG et de proposer une méthode permettant une dissection plus fine de ces locus.

Par ce travail, j'ai pu confirmer l'existence de deux régions génomiques complexes qui agissent chez les femelles et les mâles pour contribuer à une résistance accrue à la peste.

ORIGINAL ARTICLE

Resistance to plague of *Mus spretus* SEG/Pas mice requires the combined action of at least four genetic factors.

L Chevallier^{1,2}, C Blanchet^{1,2}, J Jaubert^{1,2}, E Pachulec³, C Demeure³, E Carniel³, J-J Panthier^{1,2} and X Montagutelli^{1,2}

We have previously described SEG/Pas as the first mouse inbred strain able to survive subcutaneous injection of virulent *Yersinia pestis*, the agent of plague, and we identified *Yprl1*, *Yprl2* and *Yprl3* as three quantitative trait loci (QTLs) controlling this exceptional phenotype in females from a backcross between SEG/Pas and C57BL/6 strains. We have now developed congenic strains to further characterize the extent and effect of these genomic regions. In this study, we confirm the importance of two of these regions, both in males and females, while the third one may well be a spurious association. We show that no genomic region alone is able to increase the survival of C57BL/6 mice, but that C57BL/6 mice carrying both *Yprl2* and *Yprl3* exhibit intermediate resistance. Each of these two QTLs contains at least two subregions, which are required to increase survival. Finally, through the analysis of congenic strains in an F1 background, we establish the mode of inheritance of the SEG-derived resistance alleles. Altogether, this study has clarified and enhanced our understanding of the genetic architecture of resistance to plague in SEG/Pas mice.

Genes and Immunity advance online publication, 15 November 2012; doi:10.1038/gene.2012.50

Keywords: plague; mouse; resistance; *Yersinia pestis*; congenic strain

INTRODUCTION

Plague is an acute bacterial infection caused by the Gram-negative bacterium *Yersinia pestis*, which can develop in several forms, such as bubonic, septicemic or pulmonary. The rapid multiplication of the bacteria from the entry point and the extremely efficient mechanisms by which *Y. pestis* can escape host innate immune responses are responsible for very high mortality rate in the absence of effective antibiotherapy.^{1–3}

However, the existence of host genetic factors controlling variable levels of susceptibility is strongly suggested by at least three lines of evidence. First, during past plague pandemics, a fraction of the individuals who developed bubonic plague were able to survive. Second, some mammalian species, such as rodents, cats, camels or monkeys, are susceptible, while others, such as dogs, cows and horses, are naturally resistant.⁴ The third line of evidence was provided by our recent report describing for the first time a mouse inbred strain demonstrating a high level of resistance to virulent strains of *Y. pestis*, while all mouse strains tested so far are susceptible.⁵

Taking advantage of a collection of mouse inbred strains sampling the genetic diversity found in the *Mus* genus, we evaluated their susceptibility to subcutaneous inoculation of 100 colony-forming units (CFU) of the virulent CO92 *Y. pestis* strain. Although all classical laboratory strains, including C57BL/6 (B6), as well as strains derived from *M. musculus musculus*, *M. m. domesticus* or *M. m. castaneus* progenitors, showed 0 to 23% survival, 75% of the male mice and 90% of the female mice of the *Mus spretus* derived SEG/Pas (SEG) strain were able to survive *Y. pestis* infection.⁵

Starting from this observation, we developed a forward genetics approach⁶ to identify the genetic factors underlying the resistant phenotype. Quantitative trait locus (QTL) analysis of a large (B6 × SEG)F1 × B6 (BSB) backcross cohort of 322 females detected three genomic regions, on chromosomes 3, 4 and 6, significantly

associated with differential survival.⁵ In the mixed B6-SEG background of the backcross population, individuals heterozygous (B6/SEG) for either of the three QTLs showed ~20% increase in survival compared with individuals homozygous for the B6 allele at the same QTL. Interestingly, we demonstrated that the three QTLs, named *Yprl1* to *Yprl3*, acted in an additive manner, and that mice heterozygous at the three QTLs showed a survival rate similar to that of SEG mice. Altogether, we concluded that the resistance to *Y. pestis* of SEG female mice was under the control of at least three dominant, SEG-derived, alleles accounting for most of the difference in susceptibility between the two parental strains.⁵

In this report, we used congenic strains for each of the three *Yprl* regions to confirm the existence of these QTL regions, to better describe the genetic architecture of resistance to *Y. pestis* in the mouse and to refine the confidence intervals of underlying genes. We show that none of the three QTLs alone is able to increase survival in the susceptible B6 background, but that 25% of the males and 42% of the females carrying both *Yprl2* and *Yprl3* survived an inoculation of *Y. pestis*. We further demonstrate that each QTL region contains two distinct subregions, and that increased resistance is observed only in mice carrying all four subregions. Finally, we show that the (B6 × SEG)F1 background may provide the proper genomic environment to reveal the effect on survival of individual QTLs, as exemplified with *Yprl3*. Altogether, our data provide strong evidence that resistance to plague in SEG mice is conferred by the combined action of at least four genes, with different modes of inheritance.

RESULTS

The resistance locus *Yprl3* is detected in BSB males, but neither *Yprl1* nor *Yprl2*

Males ($n = 217$) from the BSB backcross described previously were inoculated with 100 CFU of *Y. pestis* CO92 and genotyped

¹Mouse Functional Genetics Unit, Institut Pasteur, Paris, France; ²CNRS URA2578, Paris, France and ³*Yersinia* Unit, Institut Pasteur, Paris, France. Correspondence: Dr X Montagutelli, Mouse Functional Genetics Unit and CNRS URA2578, Institut Pasteur, 25 rue du Docteur Roux, Paris 75724, France.

E-mail: xavier.montagutelli@pasteur.fr

Received 9 August 2012; revised 19 September 2012; accepted 24 September 2012

for the same 721 single-nucleotide polymorphism markers. In all, 140 mice survived the infection, giving a survival rate of 64%. QTL analysis of survival as a binary trait revealed only one significant peak above the 0.05 threshold ($P = 1.9 \times 10^{-5}$; logarithm of odds (LOD) = 4.136; Figure 1). This peak is located at 93 Mb on chromosome 6, at a position very close to that of *Yprl3* (95 Mb), with a 95% confidence interval spanning from 73 to 108 Mb. At this position, survival was 48.5% in *B/B* homozygous males, while it rose to 76.7% in *B/S* heterozygotes, a level similar to that of SEG males (Supplementary Figure 1). This 28.2% difference in the survival of BSB males is greater than the 20.4% difference observed in females,⁵ suggesting a stronger effect of *Yprl3* QTL in males.

No other peak reached the 0.1 significance level. In particular, no association was found in BSB males between survival and markers on chromosomes 3 and 4, where resistance loci *Yprl1* and *Yprl2* have been mapped in females. The search for interacting loci using the 'scantwo' feature of the J/qtl software (<http://churchill.jax.org/software/jqtl.shtml>) in BSB males detected no additional loci significantly associated with survival. These results suggest that two of the three *Yprl* QTLs could be female-specific.

Homozygous *Yprl3* congenic male mice showed delayed mortality. To confirm the existence and to characterize the effect of QTLs in a pure inbred background, we developed a series of B6.SEG-*Yprl1*, B6.SEG-*Yprl2* and B6.SEG-*Yprl3* congenic strains by introgressing large regions of chromosomes 3, 4 and 6, encompassing the confidence intervals of *Yprl1* to 3, from SEG in the susceptible B6 background. Figure 2 shows, for each QTL, the genomic segment transferred in the congenic strains with reference to the QTL peak and confidence interval. The congenic strains for the QTL regions of chromosome 3, 4 and 6 (*Yprl1*, 2 and 3, respectively) are thereafter designated as cg1, cg2 and cg3, respectively, with the

haplotype in superscript. For example, mice heterozygous for the *Yprl1* congenic segment are designated as $cg1^{B/S}$, while homozygotes are named $cg1^{S/S}$.

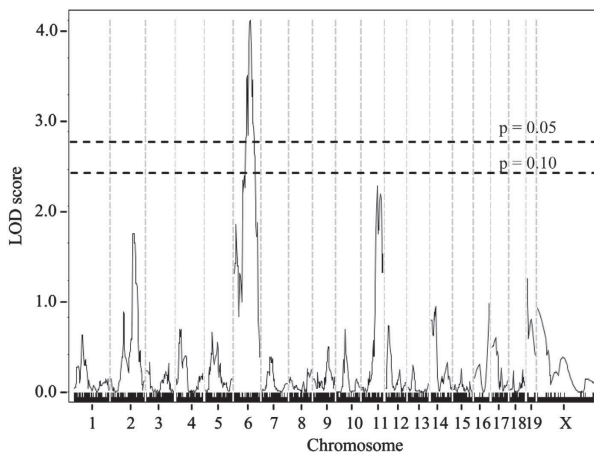
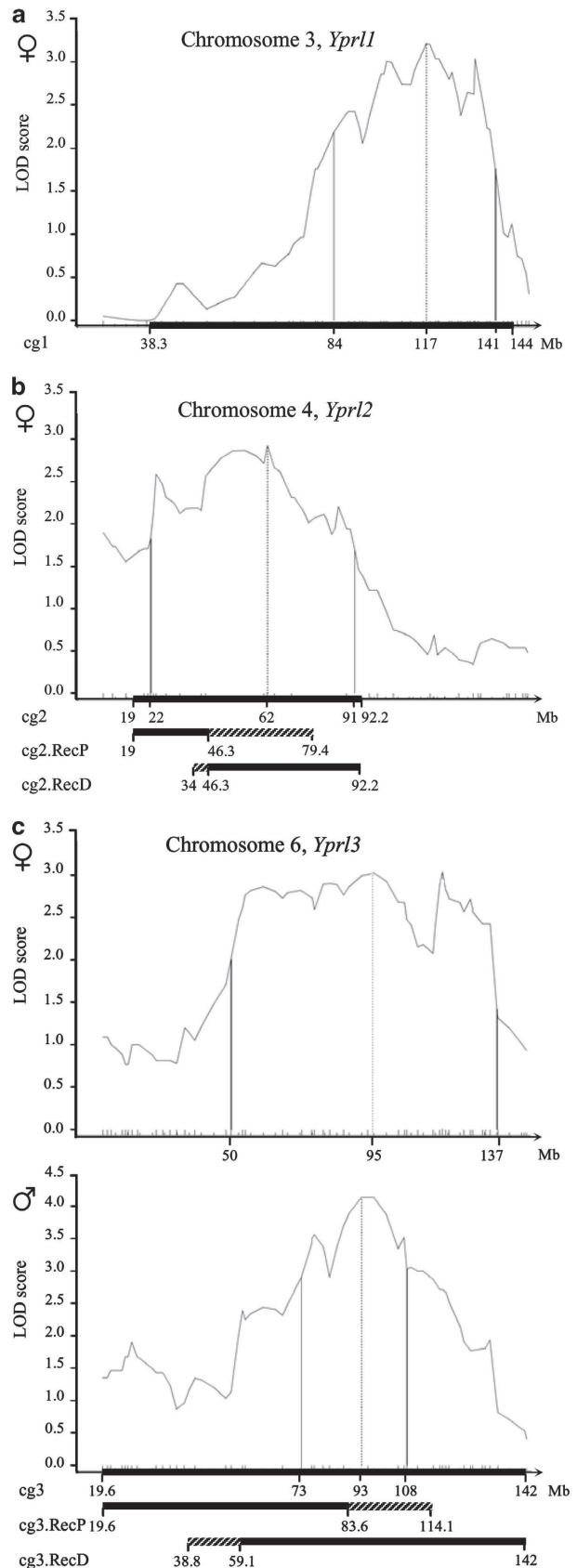


Figure 1. QTL mapping of survival to a *Y. pestis* challenge in 217 BSB males, showing a single significant QTL on chromosome 6 (*Yprl3*). Dashed lines denote thresholds of significance ($P = 0.1$ and 0.05, permutation test).

Figure 2. LOD score curves illustrating the position of each QTL peak (vertical dashed line), the 95% confidence interval (vertical solid lines) and the genomic segment transferred in congenic and subcongenic strains (black boxes) on (a) chromosome 3 (females), (b) chromosome 4 (females) and (c) chromosome 6 (upper graph: females; lower graph: males). Hatched boxes depict the undefined boundaries of the congenic regions inherited from SEG. Positions are indicated in Mb.

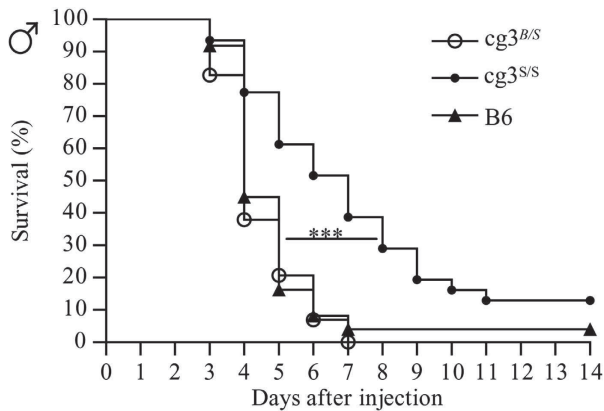


Figure 3. Survival over a 14-day period of B6 ($n=49$), heterozygous $cg3^{B/S}$ ($n=29$) and homozygous $cg3^{S/S}$ ($n=31$) male mice following subcutaneous inoculation of 10^2 CFU of strain CO92 *Y. pestis*. Mortality of $cg3^{S/S}$ males was delayed compared with B6 and $cg3^{B/S}$ males (log rank, $P<0.0001$).

We produced $cg3^{B/S}$ heterozygous and $cg3^{S/S}$ homozygous male mice and assessed their resistance to plague by comparison with B6 (Figure 3). The survival curves were analyzed by log-rank tests, to take into account both the final survival rate and the mortality time. Neither heterozygous $cg3^{B/S}$ nor homozygous $cg3^{S/S}$ males survived at a higher rate than B6, which indicates that the intrinsic effect of this genomic region is too weak to promote survival in a fully susceptible background. However, $cg3^{S/S}$ males died significantly later than both B6 and $cg3^{B/S}$ males (log rank, $P<0.0001$). This result suggests that the *Yprl3*^{SEG} resistance allele acts recessively on time to death in the congenic $cg3$ strain, while it was detected as a dominant survival-associated allele in the BSB population.

None of the *Yprl* loci are self-sufficient to increase resistance in congenic females

Females of the three congenic strains were inoculated with *Y. pestis*. As for males, both $cg3^{B/S}$ and $cg3^{S/S}$ mice were tested. However, we failed to breed $cg1$ and $cg2$ congenic mice to homozygosity because of sterility of homozygotes in at least one sex. This is most likely due to genomic incompatibility reasons that are frequent in $B6 \times SEG$ crosses.⁷ Thus, only heterozygous $cg1^{B/S}$ and $cg2^{B/S}$ females could be tested.

The survival rate and mortality time of *Y. pestis*-infected $cg1^{B/S}$ and $cg2^{B/S}$ females did not differ from that of B6 females (Figure 4). Moreover, unlike $cg3^{S/S}$ males, $cg3^{S/S}$ females did not show any difference with either B6 or $cg3^{B/S}$ females. These results could alternatively suggest that (1) one or more of these resistance loci was spuriously detected in the QTL analysis of the BSB cross; (2) each QTL alone did not have a strong enough effect to increase resistance of susceptible B6 mice; (3) the detection of QTLs' effects would require a genetic background exhibiting a resistance intermediate between that of the B6 and SEG parental strains. Further crosses were performed to discriminate these hypotheses.

Yprl2 and *Yprl3* together confer increased survival in bicongenic females

To address the first two hypotheses, we crossed the three congenic strains to produce $cg1^{B/S}$ - $cg2^{B/S}$, $cg1^{B/S}$ - $cg3^{B/S}$ and $cg2^{B/S}$ - $cg3^{B/S}$ bicongenic females. We also produced $cg1^{B/S}$ - $cg2^{B/S}$ - $cg3^{B/S}$ tricongenic females. Groups of mice carrying these compound genotypes were assessed for their resistance to *Y. pestis* (Figure 5a).

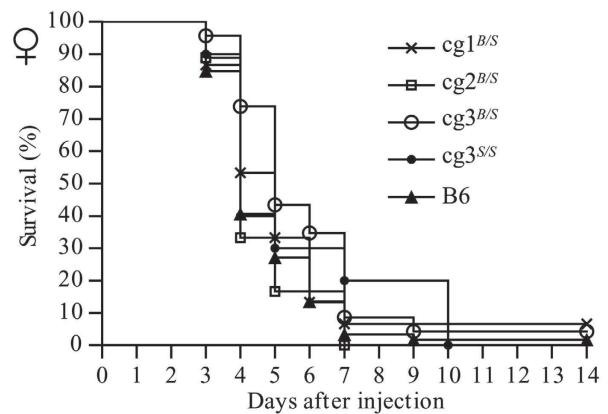


Figure 4. Survival curves of B6 ($n=59$), heterozygous $cg1^{B/S}$ ($n=15$), $cg2^{B/S}$ ($n=18$) and $cg3^{B/S}$ ($n=23$) females and homozygous $cg3^{S/S}$ females ($n=10$) to subcutaneous inoculation of 10^2 CFU. None of the congenic strains was found to be more resistant than B6.

First, we observed that 41.6% (10/24) of $cg2^{B/S}$ - $cg3^{B/S}$ bicongenic females survived the infection (versus 1/60 in B6, $P=6.5 \times 10^{-6}$). This survival rate is much higher than that found in $cg2^{B/S}$ and $cg3^{B/S}$ congenic mice, which demonstrates that both *Yprl2* and *Yprl3* genomic regions contain at least one resistance locus each, and that these two QTL must be present simultaneously to overcome the susceptible B6 background.

On the other hand, neither $cg1^{B/S}$ - $cg2^{B/S}$ nor $cg1^{B/S}$ - $cg3^{B/S}$ bicongenic mice were able to resist significantly better than B6 or congenic strains ($cg1^{B/S}$ - $cg2^{B/S}$: 2/12 = 16.7%; $cg1^{B/S}$ - $cg3^{B/S}$: 0/15 = 0%), which suggests that either the *Yprl1* region does not contain any resistance gene or *Yprl1* has a weaker effect compared with *Yprl2* and *Yprl3*.

Despite the difficulties of producing mice heterozygous for three large genomic regions,^{7,8} we produced a small group of tricongenic mice ($n=9$). These mice showed a resistance level lower but not significantly different from that of $cg2^{B/S}$ - $cg3^{B/S}$ mice (2/9 = 22.2%; $P=0.43$). Taken together with the $cg1^{B/S}$ - $cg2^{B/S}$ and $cg1^{B/S}$ - $cg3^{B/S}$ data, these results led us to consider that *Yprl1* could be a spurious QTL and to discard it from further analysis.

Yprl2 participates in resistance to *Y. pestis* also in males

We further tested whether *Yprl2* was indeed not involved in male resistance, or whether it had been missed in the analysis of BSB males. To this aim, we infected $cg2^{B/S}$, $cg3^{B/S}$ (same as above, Figure 3) and $cg2^{B/S}$ - $cg3^{B/S}$ males. Neither $cg3^{B/S}$ (as reported previously) nor $cg2^{B/S}$ males exhibited increased resistance compared with B6 males (1/18 = 5.5% and 2/51 = 3.9%, respectively; $P=1$). By contrast, bicongenic $cg2^{B/S}$ - $cg3^{B/S}$ males survived to a higher level than B6 males (11/44 = 25% and 2/51 = 3.9%, respectively; $P=0.0053$; Figure 5b), indicating that although no association was found between survival and chromosome 4 markers in BSB males, *Yprl2* was able to increase the effect of *Yprl3*. We concluded that *Yprl2* participates in resistance to *Y. pestis* in both sexes.

Yprl2 and *Yprl3* dissection identifies at least four loci required for resistance

As both $cg2^{B/S}$ and $cg3^{B/S}$ congenic strains carry large genomic regions (approximately 73 and 123 Mb, respectively), we selected recombinants and bred them to produce subcongenic strains carrying either the proximal or the distal part of *Yprl2* and *Yprl3* (see Figure 2 for detailed composition). The resulting subcongenics were intercrossed to produce combinations of the four subregions. All mice were heterozygous at the subregions.

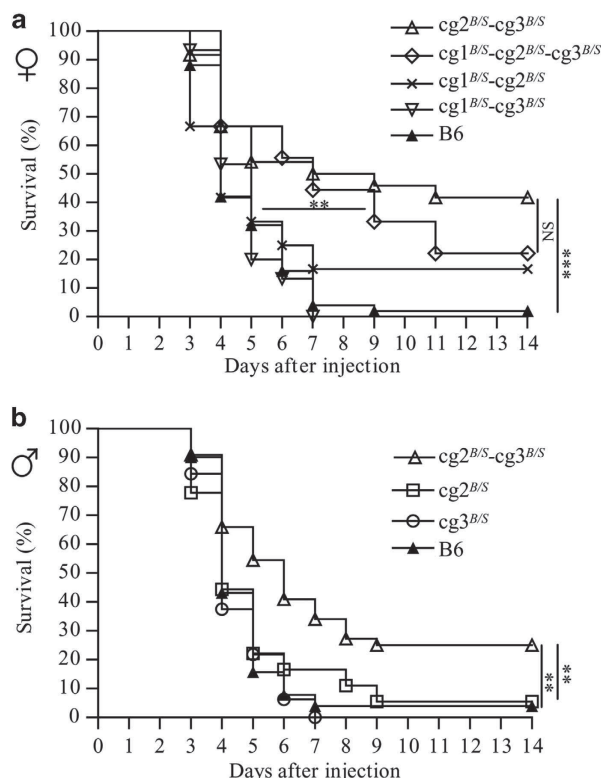


Figure 5. Survival curves of bi- and tricongenic females and congenic and bicongenic males. **(a)** Females: B6 ($n=50$), heterozygous bicongenic $cg1^{B/S}-cg2^{B/S}$ ($n=12$), $cg1^{B/S}-cg3^{B/S}$ ($n=15$) and $cg2^{B/S}-cg3^{B/S}$ ($n=24$) and heterozygous tricongenic females $cg1^{B/S}-cg2^{B/S}-cg3^{B/S}$ ($n=9$). $Cg1^{B/S}-cg2^{B/S}$ and $cg1^{B/S}-cg3^{B/S}$ were not different from B6, while the survival rate of $cg2^{B/S}-cg3^{B/S}$ (10/24 = 41.6%) was significantly higher than that of B6 (1/60 = 1.6%, $P = 6.5 \times 10^{-6}$). $Cg1^{B/S}-cg2^{B/S}-cg3^{B/S}$ mice did not survive in higher percentage than B6 (2/9 = 22.2% and 1/60 = 1.6%, respectively; $P=0.059$), but died significantly later (log rank, $P=0.0057$). $Cg2^{B/S}-cg3^{B/S}$ and $cg1^{B/S}-cg2^{B/S}-cg3^{B/S}$ did not significantly differ in their survival rate. **(b)** Males: B6 ($n=51$), heterozygous $cg2^{B/S}$ ($n=18$), $cg3^{B/S}$ ($n=29$, same as on Figure 3) and heterozygous bicongenic $cg2^{B/S}-cg3^{B/S}$ ($n=44$). Survival rate of $cg2^{B/S}-cg3^{B/S}$ (11/44 = 25.0%) was higher than that of B6 and $cg2^{B/S}$ (2/51 = 3.9%, $P = 0.0053$ and 1/18 = 5.5%, $P = 0.0019$, respectively).

Data from males and females were pooled as both sexes gave similar results.

As anticipated from the result of $\text{cg2}^{B/S}$ and $\text{cg3}^{B/S}$ congenic strains, mice carrying only the proximal or distal region of either QTL did not show any difference from B6 in their susceptibility to *Y. pestis* (data not shown). In both males and females, mice carrying the full region at one QTL and either the proximal or the distal subregion at the other QTL did not show increased survival (Figure 6). However, $\text{cg2.RecP}^{B/S}$ - $\text{cg3}^{B/S}$ males and females died later than B6 ($P < 0.0001$; Figure 7). We concluded that each of the four subregions contains a resistance locus necessary to confer the survival rate observed in $\text{cg2}^{B/S}$ - $\text{cg3}^{B/S}$ bicongenic mice.

Yprl3 increases survival of (B6 × SEG)F1 mice

Our results from congenic and bicongenic strains suggested that resistance is dependent on the combined effect of multiple SEG-derived alleles at the *Yprl2* and *Yprl3* loci. Moreover, in the BSB backcross population, we observed that mice homozygous for the B6 allele at the three QTLs still resist at a higher level than B6 (27%, $P = 0.0017$),⁵ suggesting that even a low proportion (25%) of SEG alleles dispersed throughout the genome is able to confer minimal

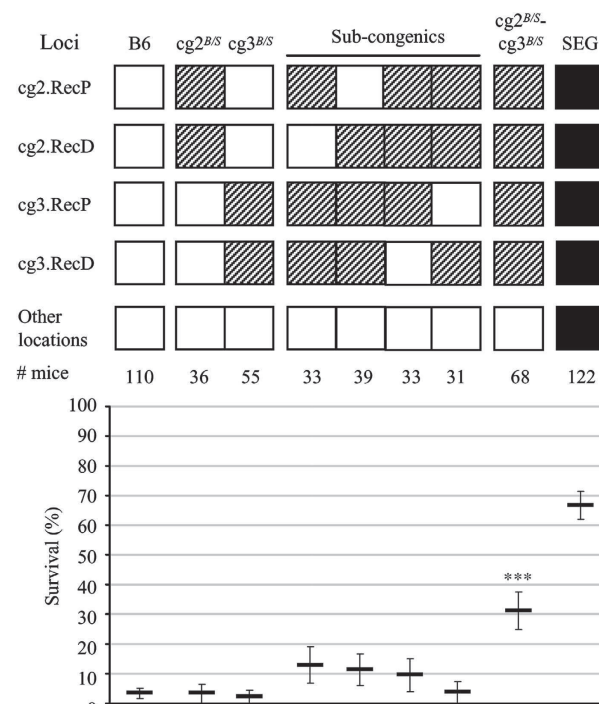


Figure 6. Survival rate of male and female B6 ($n=110$), $cg2^{B/S}$ ($n=36$), $cg3^{B/S}$ ($n=55$), subcongenics of $cg2$ and $cg3$: $cg2.RecP^{B/S}$ - $cg3^{B/S}$ ($n=33$), $cg2.RecD^{B/S}$ - $cg3^{B/S}$ ($n=39$), $cg2^{B/S}$ - $cg3.RecP^{B/S}$ ($n=33$) and $cg2^{B/S}$ - $cg3.RecD^{B/S}$ ($n=31$), $cg2^{B/S}$ - $cg3^{B/S}$ ($n=68$) and SEG ($n=122$). The top table gives the genotype of each group for the proximal and distal interval of *Yprl2* and *Yprl3*. White box: B/B ; hatched box: B/S ; and black box: S/S . The graph shows the survival rate of each group (with error bars indicating 1 s.d.). Only $cg2^{B/S}$ - $cg3^{B/S}$ mice survived at a significantly higher rate than B6 ($21/68=30.8\%$ and $3/110=2.7\%$, respectively; $P=1.32 \times 10^{-7}$).

level of resistance to *Y. pestis*. We hypothesized that studying single QTLs on a mixed B6-SEG background could increase the power to detect small effects.

We have previously reported that (B6 \times SEG)F1 mice show intermediate resistance between B6 and SEG.⁵ This hybrid background is easy to obtain and is very reproducible. Therefore, we set up crosses between heterozygous congenic females and SEG males to obtain both *Ypr*^{B/S} heterozygotes and *Ypr*^{S/S} homozygotes on the (B6 \times SEG)F1 background.

In a preliminary experiment, we infected SEG and F1 females. Although F1 mice resisted at a lower level than SEG mice, the difference was not statistically significant, which was not a favorable condition to detect an increase in survival possibly conferred by a resistance QTL. The difference was larger in males and we therefore investigated the effect of *Yprl2* and *Yprl3* in the F1 background. F1-cg3 (*Yprl3*^{S/S}) males resisted at a much higher level than F1 (*Yprl3*^{B/S}) males (11/21 = 52.4% and 6/28 = 21.4%, respectively; *P* = 0.035; Figure 8), which clearly demonstrates that an effect of the *Yprl3* locus on survival could be detected, provided an appropriate genetic background is used. It implies that the *Yprl3*^{SEG} allele is inherited codominantly since *Yprl3*^{S/S} and *Yprl3*^{B/S} showed different levels of resistance on the same genetic background. We further tested F1-cg3.RecP^{S/S} mice, which are homozygous for the SEG haplotype in the proximal part of the *Yprl3* QTL, and found that they were much more resistant than F1 males (20/32 = 62.5% and 6/28 = 21.4%, respectively; *P* = 0.0018; Figure 8) and did not differ in survival from F1-cg3^{S/S} mice. This result indicates that this proximal region is sufficient to recapitulate the increased resistance observed in F1-cg3^{S/S} mice. We also tested F1-cg3.RecD^{S/S}, which are homozygous for the

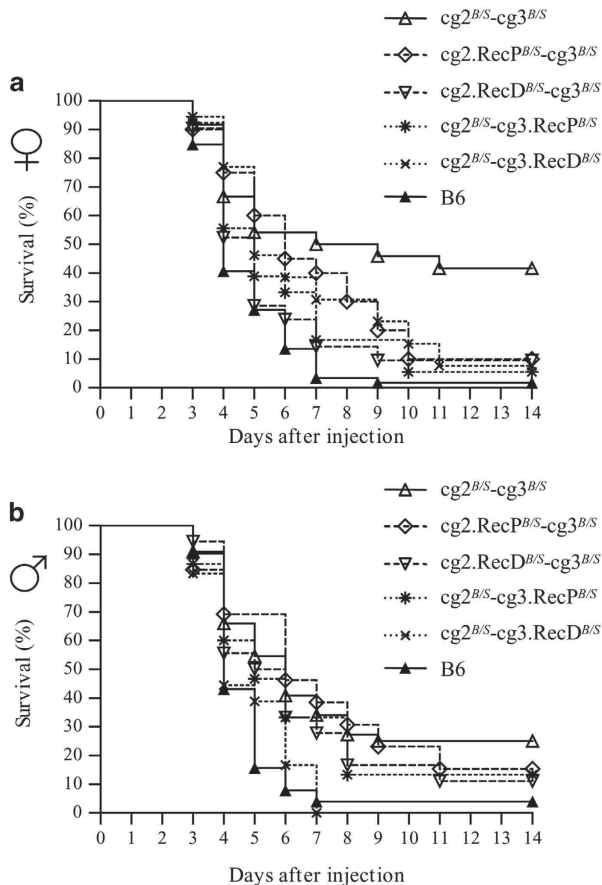


Figure 7. Survival curves of males and females subcongenic for *Yprl2* and *Yprl3*. **(a)** Survival curves of females: B6 ($n=59$), $cg2.RecP^{B/S}-cg3^{B/S}$ ($n=20$), $cg2.RecD^{B/S}-cg3^{B/S}$ ($n=21$), $cg2^{B/S}-cg3.RecP^{B/S}$ ($n=18$), $cg2^{B/S}-cg3.RecD^{B/S}$ ($n=13$) and $cg2^{B/S}-cg3^{B/S}$ ($n=24$). Only $cg2^{B/S}-cg3^{B/S}$ survived to a significantly higher rate than B6 ($10/24=41.6\%$ and $1/60=1.6\%$, respectively; $P=6.5 \times 10^{-6}$). Females: $cg2.RecP^{B/S}-cg3^{B/S}$ died significantly later than B6 ($P<0.0001$). **(b)** Survival curves of males: B6 ($n=51$), $cg2.RecP^{B/S}-cg3^{B/S}$ ($n=13$), $cg2.RecD^{B/S}-cg3^{B/S}$ ($n=18$), $cg2^{B/S}-cg3.RecP^{B/S}$ ($n=15$), $cg2^{B/S}-cg3.RecD^{B/S}$ ($n=18$) and $cg2^{B/S}-cg3^{B/S}$ ($n=44$). Only $cg2^{B/S}-cg3^{B/S}$ survived to a significantly higher rate than B6 ($11/44=25.0\%$ and $2/51=3.9\%$, respectively; $P=0.0053$). Males: $cg2.RecP^{B/S}-cg3^{B/S}$ died significantly later than B6 (log rank, $P<0.0001$).

distal part of the *Yprl3* region. Although only a small number of mice could be produced because of low fertility, F1- $cg3.RecD^{S/S}$ survived to the same level as F1 mice ($2/8=25\%$ and $6/28=21.4\%$, respectively). These results suggest that the SEG allele at the proximal locus of the *Yprl3* interval is codominant (as F1- $cg3.RecP^{S/S}$ survived better than F1), while the SEG allele at the distal locus is fully dominant (as F1- $cg3.RecP^{S/S}$ and F1- $cg3^{S/S}$ survive similarly and F1- $cg3.RecD^{S/S}$ and F1 also survived similarly).

Conversely, we detected no difference in survival between F1- $cg2^{S/S}$ (homozygous for *Yprl2*) and F1 males ($6/16=37.5\%$ and $6/29=20.6\%$, respectively; $P=0.29$; Figure 8), which may result from *Yprl2* SEG allele(s) acting in a fully dominant manner.

DISCUSSION

Y. pestis is one of the most virulent bacteria for humans and cause for extremely acute infections. Many rodents, including mouse laboratory strains, are highly susceptible to *Y. pestis*.^{5,9,10} We previously reported that almost all inbred strain tested succumbed in less than a week following experimental subcutaneous

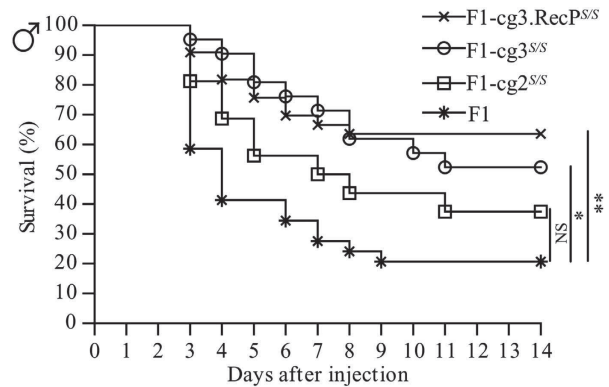


Figure 8. Survival curves of male F1 ($n=28$), F1- $cg2^{S/S}$ ($n=16$), F1- $cg3^{S/S}$ ($n=21$) and F1- $cg3.RecP^{S/S}$ ($n=32$) mice. F1- $cg3^{S/S}$ and F1- $cg3.RecP^{S/S}$ mice survived at a higher rate than F1 ($11/21=52.4\%$, $20/32=62.5\%$ and $6/28=21.4\%$, respectively; $P=0.035$ and 0.0018 , respectively). The difference between F1- $cg2^{S/S}$ and F1 was not significant.

inoculation of 100 CFU of a virulent strain such as CO92. The only outstanding exception was the *Mus spretus*-derived SEG/Pas strain, which showed high survival rate under these conditions. Through a forward genetics approach, we identified three candidate genomic regions controlling survival in BSB females. In each region, SEG alleles act dominantly to increase survival by 20%, and BSB mice heterozygous for the three regions were found to resist to the same level as SEG mice.⁵ In parallel, we explored a number of clinical and immunological parameters in search for pathophysiological mechanisms associated with resistance to plague.¹¹

This study aimed at confirming the existence of these QTLs and at refining our understanding of the genetic architecture of resistance to *Y. pestis* in mice. First, we analyzed males from the same BSB backcross used in our initial study. We found a major QTL at a position very close to that of *Yprl3*, suggesting that *Yprl3* is efficient in both sexes. Notably, its effect on survival was stronger in males than in females (28.2% and 20.4%, respectively). No other suggestive or significant QTL was found in males. In particular, no association was found at the *Yprl1* and *Yprl2* genomic locations. This could have at least three reasons. First, the number of males analyzed in the backcross population was one-third smaller than the number of females (217 and 322, respectively), resulting in reduced power for QTL detection. However, if *Yprl1* and *Yprl2* were efficient in males, they would likely show at least weak association with survival, which was not observed. Second, we reported that the difference in survival rate between SEG and B6 was greater in females than in males, which led us to analyze only females in the first study. *Yprl1* and *Yprl2* genes could indeed have no effect in males, resulting in the reduced survival observed in SEG males compared with females. The third hypothesis is that one or both QTLs were spuriously detected in females and not replicated in males. This is plausible as their LOD score was just above the significance threshold.

Congenic strains are used in mouse complex traits to confirm and further characterize QTLs identified by association in segregating crosses.¹² We therefore developed congenic strains by introgressing putative SEG-derived resistance alleles in the susceptible B6 background. As the confidence intervals observed in the BSB cross were quite large, suggesting the presence of several linked genes, we transferred genomic regions encompassing 73 to 123 Mb. We could obtain homozygous congenic mice only for *Yprl3*, probably because of interspecific allelic incompatibilities that are common in B6 \times SEG crosses.^{7,8}

The phenotyping of congenic and bicongenic strains led to important and unexpected conclusions. First of all, none of the

QTLs identified in females was able to increase the very low level of resistance of B6 mice. Although this was surprising considering the 20% increase in survival that each of them was able to induce in BSB mice, this is not an uncommon observation when a late and general phenotype such as death is considered. For example, in their dissection of the genetic control of systemic lupus erythematosus, Morel and co-workers¹³ identified four genomic regions associated with the development of glomerulonephritis, a late complication of autoimmunity.¹³ When they developed congenic strains on the resistant background, they observed the development of mild kidney lesions in only one of them, while the three other strains remained unaffected.¹⁴

Our second observation was that $cg2^{B/S}$ - $cg3^{B/S}$ bicongenic mice showed a survival rate approximately half that of SEG mice. Unexpectedly, the same observation was made in males and females, although the *Yprl2* locus had not been identified in BSB males. Congenic strains provide stronger experimental evidence than association studies in segregating populations since genetic composition is controlled and homogeneous between groups and within groups. Therefore, we can conclude that the previously described *Yprl2* and *Yprl3* are both true QTLs controlling resistance to *Y. pestis*.

On the other hand, congenic strains have failed to provide confirmation of the existence of *Yprl1* identified in BSB females. This locus was associated with the same level of resistance as the other two, with an LOD score just above the genome-wide threshold for significance. It is therefore possible that it was a spurious QTL. In fact, literature reports many examples of QTL described in a segregating population and not replicated in congenic strains.¹⁵ Another hypothesis could be that this locus acts epistatically with another locus. However, we did not detect any such interaction in the analysis of BSB females. Therefore, we decided not to pursue the analysis of *Yprl1*.

Congenic strains are also invaluable in refining the location of QTLs, through the fragmentation of the congenic interval.^{16–20} Because the *Yprl2* and *Yprl3* peaks were quite broad, $cg2$ and $cg3$ encompass 50 to 80% of the entire chromosome length. As an initial step towards QTL fine mapping, we split each congenic strain in two subregions. However, removing any of the four subregions abolished the survival observed with $cg2^{B/S}$ - $cg3^{B/S}$ bicongenic mice. This indicates that each of these four regions contains at least one gene critical to induce a detectable level of resistance in the B6 background. Moreover, SEG alleles at these four regions act under a dominant (or semidominant) mode of inheritance.

The comparison of the BSB cross with congenic strains led us to postulate that the genetic background as a whole was crucial in revealing the effect of individual QTLs. Each QTL likely provides a small increase in the capacity to overcome bacterial multiplication and dissemination. QTL effects on survival could be detected in the BSB cross since mice lacking the QTLs (but yet having inherited 25% of their genome from SEG) survived at an intermediate rate of 30%.⁵ Under these conditions, even mild QTL effects on resistance mechanisms could induce a detectable variation of survival rate. Conversely, on a pure, highly susceptible B6 background, the small increase in resistance controlled by QTLs would have no detectable impact on the ability to survive. This led us to compare mice differing for the QTLs, in an F1 background. This experiment was set up by crossing congenic females with SEG males. We could not use this approach in females because of the high resistance level of F1s. It was successful in males with *Yprl3*. Our results suggest that such crosses between a congenic strain and its donor strain could be very helpful in unraveling the effect of QTL regions, which are missed in the recipient background.

Assessing the mode of inheritance of QTLs is important as an indication of their mode of action. According to the BSB data (in which mice are either *B/B* or *B/S*), *Yprl2* and *Yprl3* must be

inherited dominantly or semidominantly. Owing to the difficulties inherent to interspecific crosses, $cg3$ is the only congenic strain that could be bred to homozygosity. In males, we showed that, although they did not survive better than B6 and $cg3^{B/S}$ mice, $cg3^{S/S}$ mice died later, which suggests a semidominant mode of inheritance. Moreover, in F1 crosses, we found that F1- $cg3$.RecP^{S/S} mice, which are homozygous for the SEG haplotype in the proximal part of the *Yprl3* region, did not differ in survival from F1- $cg3^{S/S}$ mice and survived better than F1 mice. Altogether, these results suggest that SEG alleles are codominant at the proximal locus of *Yprl3* and dominant at the distal locus. Conversely the absence of difference between F1- $cg2^{S/S}$ and F1 males suggests that SEG alleles at both regions of *Yprl2* are fully dominant.

In conclusion, this study has clarified and enhanced our understanding of the genetic architecture of resistance to plague in SEG/Pas mice. We have definitely confirmed the existence of two complex genomic regions, acting in both males and females. Although neither of them is able by itself to promote survival, mice carrying both regions resist at an intermediate level. These two QTLs encompass at least four genes. For three of them, the dominant inheritance of SEG alleles suggests that they are gain-of-function variants. The last one is semidominantly inherited and could be a quantitative variant. There must be other resistance genes as *Yprl2* and 3 together fail to provide the same resistance level as SEG. Finally, the use of a hybrid background exhibiting an intermediate level of resistance was capable of revealing *Yprl3* effect on survival and will be instrumental in the fine mapping of the $cg3$.RecP region.

The identification of the genes underlying the effects of the four chromosomal regions will require a combination of approaches, including the fine genetic mapping of each region, the identification of more subtle and earlier phenotypes than death to compare the progression of the disease in congenic and subcongenic strains and whole-genome expression studies to detect pathways that are differentially regulated in response to *Y. pestis* exposure.

MATERIALS AND METHODS

Mice and crosses

The SEG/Pas strain was developed as previously described²¹ and is maintained at the Institut Pasteur. C57BL/6 (B6) mice were purchased at 6 weeks of age from Charles River Laboratory (L'Arbresle, France). They were maintained for at least 3 weeks in the Institut Pasteur animal facilities before inoculation. Male BSB backcross mice were part of the same cross described in our previous study.⁵

Congenic strains were developed by backcrossing 10 times SEG onto B6. Specifically, SEG males were mated with B6 females. F1 females were then backcrossed to B6 males. To overcome hybrid sterility, which affect 7/8 backcross males,⁷ the resulting BC1 (or N2) females that had retained either of the QTL regions (see genotyping below) were mated with B6 males. This cross was repeated once more. For further generation, males heterozygous for the genomic regions of interest were crossed with B6 females. After 10 generations of backcrossing (N10), attempts were made to establish homozygous congenic strains. However, most congenic strains had to be maintained by continuously backcrossing heterozygous carriers. To facilitate reading, a simplified nomenclature is used throughout the article. B6.SEG-*Yprl1* to B6.SEG-*Yprl3* are referred to as $cg1$ to $cg3$. The haplotype is added in superscript (for example, $cg1^{B/S}$ describes a mouse heterozygous for the congenic segment).

Subcongenic strains were produced by selecting recombinants from the progeny of N10 (or more) congenic mice. RecP and RecD are used to designate the subcongenic strains carrying the proximal and the distal region of the original congenic region, respectively (for example, $cg2$.RecD).

Bicongenic individuals were obtained by intercrossing heterozygous mice of congenic strains (for example, $cg1^{B/S}$ and $cg2^{B/S}$). Tricongenic individuals were obtained by intercrossing $cg1^{B/S}$ males with $cg2^{B/S}$ - $cg3^{B/S}$ bicongenic females. Mice homozygous for *Yprl2* (or *Yprl3*) on a (B6 × SEG)F1 background were produced by crossing SEG males with heterozygous $cg2^{B/S}$ (or $cg3^{B/S}$) females and were named F1- $cg2^{S/S}$ (or F1- $cg3^{S/S}$).

Notably, the continuous backcrossing of congenic strains to B6 (beyond 10 generations), and the genotypic selection of bi- and tricongenic mice in

the segregating progeny of crosses between congenic strains precluded the potential effect of residual SEG alleles.

All animal experiments were approved and conducted in accordance with the Institut Pasteur Biosafety Committee, and in compliance with French and European regulation on protection of animals used for scientific purposes. Mice were kept in isolators after inoculation. They were given food and water *ad libitum* and were kept at $22 \pm 2^\circ\text{C}$ with alternating 12 h periods of light and dark.

Culture of *Y. pestis* strains and experimental infections of mice

Experimental infections were performed using the virulent *Y. pestis* strain CO92.²² Cultures were carried out at 28°C for 48 h on LB agar medium supplemented with 0.002% (WV^{-1}) hemin (LBH). Bacteria were collected and suspended in saline. Bacterial concentration was estimated by spectrometry at 600 nm ($1\text{ OD} \sim 10^9\text{ CFU ml}^{-1}$) and confirmed by CFU counts on LBH agar plates. Mice were injected subcutaneously between 8 and 12 weeks of age, in the ventral region, with a $100\text{ }\mu\text{l}$ inoculum, and were then monitored daily for 14 days. For practical reasons, congenic strains were tested over multiple infection experiments. To control for consistency and allow merging data, B6 and SEG control mice ($n \geq 5$) were included in every experiment.

Genotyping

For congenic strains, mice were genotyped using DNA prepared from tail biopsies collected at weaning age. Microsatellite markers were genotyped according to standard PCR protocols, and using 4% agarose gels. DNA preparation and genotyping of BSB mice was described previously.⁵ The markers used for genotyping congenic mice are listed in Supplementary Table 1. The interval between neighbor markers varied between 3 and 40 Mb ($\sim 2\text{--}20\text{ cM}$), which precluded missing double crossing-overs, as interference is known to encompass 30–40 cM in the mouse.²³

Statistical and QTL analysis

Survival rates were compared by Fisher's exact test and survival curves were compared by a log-rank (Mantel–Cox) test using StatView 5.0 (SAS Institute Inc., Cary, NC, USA) or Prism 4.0 statistical software (GraphPad, La Jolla, CA, USA). QTL analysis was performed by using the J/qtl software version 1.3.1. The survival rate was analyzed as a binary trait. Significance thresholds of LOD scores were estimated by 1000 permutations of experimental data, and were 2.77 and 2.43 for 5% and 10% thresholds, respectively. In all, 95% confidence intervals were estimated in J/qtl using the Bayesian Credible Interval function.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Simon Foote, Beverley Paigen and Ken Paigen for fruitful discussions. We are grateful to Sofia Filali and Nicolas Cabanel for technical expertise in culture of *Yersinia pestis* CO92, and to Laurent Guillemot for technical assistance in experimental infections of mice. We thank Isabelle Lanctin for careful breeding of the SEG/Pas colony. LC is a recipient of a DGA (*Délégation Générale pour l'Armement*) fellowship. The Mouse functional Genetics Unit is supported by Merck Serono and by the 'Initiative d'excellence' Project Integrative Biology of Emerging Infectious Diseases.

Supplementary Information accompanies the paper on Genes and Immunity website (<http://www.nature.com/gene>)

REFERENCES

- 1 Lawrenz MB, Lenz JD, Miller VL. A novel autotransporter adhesin is required for efficient colonization during bubonic plague. *Infect Immun* 2009; **77**: 317–326.
- 2 Nakajima R, Motin VL, Brubaker RR. Suppression of cytokines in mice by protein A–V antigen fusion peptide and restoration of synthesis by active immunization. *Infect Immun* 1995; **63**: 3021–3029.
- 3 Titball RW, Williamson ED. Vaccination against bubonic and pneumonic plague. *Vaccine* 2001; **19**: 4175–4184.
- 4 Perry RD, Fetherston JD. *Yersinia pestis*—etiologic agent of plague. *Clin Microbiol Rev* 1997; **10**: 35–66.
- 5 Blanchet C, Jaubert J, Carniel E, Fayolle C, Milon G, Szatanik M *et al*. *Mus spretus* SEG/Pas mice resist virulent *Yersinia pestis*, under multigenic control. *Genes Immun* 2011; **12**: 23–30.
- 6 Vidal SM, Malo D, Marquis JF, Gros P. Forward genetic dissection of immunity to infection in the mouse. *Annu Rev Immunol* 2008; **26**: 81–132.
- 7 Burgio G, Szatanik M, Guenet JL, Arnau MR, Panthier JJ, Montagutelli X. Inter-specific recombinant congenic strains between C57BL/6 and mice of the *Mus spretus* species: a powerful tool to dissect genetic control of complex traits. *Genetics* 2007; **177**: 2321–2333.
- 8 Dejager L, Libert C, Montagutelli X. Thirty years of *Mus spretus*: a promising future. *Trends Genet* 2009; **25**: 234–241.
- 9 Pollitzer R. A review of recent literature on plague. *Bull World Health Organ* 1960; **23**: 313–400.
- 10 Russell P, Eley SM, Hibbs SE, Manchee RJ, Stagg AJ, Titball RW. A comparison of plague vaccine, USP and EV76 vaccine induced protection against *Yersinia pestis* in a murine model. *Vaccine* 1995; **13**: 1551–1556.
- 11 Demeure CE, Blanchet C, Fitting C, Fayolle C, Khun H, Szatanik M *et al*. Early systemic bacterial dissemination and a rapid innate immune response characterize genetic resistance to plague of SEG mice. *J Infect Dis* 2012; **205**: 134–143.
- 12 Rogner UC, Avner P. Congenic mice: cutting tools for complex immune disorders. *Nature reviews. Immunology* 2003; **3**: 243–252.
- 13 Morel L, Rudofsky UH, Longmate JA, Schiffenbauer J, Wakeland EK. Polygenic control of susceptibility to murine systemic lupus erythematosus. *Immunity* 1994; **1**: 219–229.
- 14 Morel L, Yu Y, Blenman KR, Caldwell RA, Wakeland EK. Production of congenic mouse strains carrying genomic intervals containing SLE-susceptibility genes derived from the SLE-prone NZM2410 strain. *Mamm Genome* 1996; **7**: 335–339.
- 15 Marquis JF, Lacourse R, Ryan L, North RJ, Gros P. Genetic and functional characterization of the mouse Trl3 locus in defense against tuberculosis. *J Immunol* 2009; **182**: 3757–3767.
- 16 Mott R. Finding the molecular basis of complex genetic variation in humans and mice. *Philos Trans R Soc Lond Ser* 2006; **361**: 393–401.
- 17 Cuda CM, Zeumer L, Sobel ES, Croker BP, Morel L. Murine lupus susceptibility locus Sle1a requires the expression of two sub-loci to induce inflammatory T cells. *Genes Immun* 2010; **11**: 542–553.
- 18 Lindvall T, Karlsson J, Holmdahl R, Andersson A. Dissection of a locus on mouse chromosome 5 reveals arthritis promoting and inhibitory genes. *Arthritis Res Ther* 2009; **11**: R10.
- 19 Sancho-Shimizu V, Khan R, Mostow S, Larivière L, Wilkinson R, Riendeau N *et al*. Molecular genetic analysis of two loci (Ity2 and Ity3) involved in the host response to infection with *Salmonella typhimurium* using congenic mice and expression profiling. *Genetics* 2007; **177**: 1125–1139.
- 20 Zeumer L, Sang A, Niu H, Morel L. Murine lupus susceptibility locus Sle2 activates DNA-reactive B cells through two sub-loci with distinct phenotypes. *Genes Immun* 2011; **12**: 199–207.
- 21 Guenet JL, Bonhomme F. Wild mice: an ever-increasing contribution to a popular mammalian model. *Trends Genet* 2003; **19**: 24–31.
- 22 Doll JM, Zeitz PS, Ettestad P, Bucholtz AL, Davis T, Gage K. Cat-transmitted fatal pneumonic plague in a person who traveled from Colorado to Arizona. *Am J Trop Med Hyg* 1994; **51**: 109–114.
- 23 Davies PO, Melanitou E, Asano M, Avner PR, Montagutelli X. An anchored molecular map of mouse chromosome 6 with an analysis of interference. *Mamm Genome* 1995; **6**: 738–740.

2. Résultats complémentaires

Les résultats présentés dans l'article montrent que les femelles congéniques B6.SEG-*Ypr11*^{B/S} (notées *cg1*^{B/S}) ne survivent pas significativement plus que les souris B6, contrairement à ce que laissaient supposer les résultats du croisement BSB. *Ypr11* ne permet pas non plus d'augmenter la survie des souris lorsqu'il est associé à *Ypr12* ou *Ypr13*. De plus, les femelles tri-congéniques *cg1*^{B/S}-*cg2*^{B/S}-*cg3*^{B/S} présentent un taux de survie moindre (mais néanmoins pas différent statistiquement) par rapport aux souris bi-congéniques *cg2*^{B/S}-*cg3*^{B/S}.

J'ai réalisé d'autres expériences afin de tester, de différentes manières, l'effet éventuel de *Ypr11* sur la survie des souris.

2.1. Les femelles homozygotes pour *Ypr11* ne survivent pas plus que les femelles B6

L'effet du QTL *Ypr11* n'a été testé qu'à l'état hétérozygote. Comme il est mentionné dans l'article, les femelles *cg1*^{S/S} sont complètement infertiles, ce qui m'a empêchée de fixer la lignée congénique *cg1* à l'état homozygote. Il m'a donc été impossible de produire en grand nombre des femelles *cg1*^{S/S} pour analyser leur survie. Néanmoins, dans la multitude des croisements effectués, j'ai obtenu un petit nombre de femelles possédant la région chromosomique *cg1* à l'état homozygote (*cg1*^{S/S}) et des segments variables de *cg3* à l'état hétérozygote. Ni le taux de survie, ni la durée de survie de ces femelles *cg1*^{S/S} ne différaient de celles des femelles témoins B6 (1/13 = 7,7 % et 1/38 = 2,6% de survie, respectivement; Figure 1). Donc, même à l'état homozygote, *Ypr11* ne semble pas capable d'augmenter la survie des souris sensibles.

2.2. Les sous-régions chromosomiques de *cg1* n'ont pas d'effet sur la survie

La région chromosomique transférée dans la lignée congénique *cg1* est beaucoup plus grande (106 Mb) que l'intervalle de confiance calculé pour ce QTL (environ 50 Mb). Il serait possible que le segment chromosomique transféré dans la lignée congénique *cg1* contienne plusieurs facteurs génétiques, situés de part et d'autre du QTL, possédant des effets contradictoires sur la résistance des souris. Dans ce cas, l'effet du QTL détecté dans le *backcross* pourrait être supprimé par un allèle de sensibilité lié venant de SEG et co-transféré dans la lignée congénique *cg1*. La création de lignées sous-congéniques portant différents intervalles du QTL peut permettre de distinguer ces facteurs. J'ai ainsi produit et analysé trois lignées sous-congéniques portant soit une région proximale (*cg1*.RecP), soit une région distale (*cg1*.RecD), soit la région couvrant seulement l'intervalle de confiance de *Ypr11*

(cg1.RecM) (Figure 2). Aucune de ces sous-lignées ne présentait de différence significative de taux ou de durée de survie par rapport aux femelles B6 (0/7 = 0% de survie pour les cg1.RecP^{B/S}, 0/7 = 0% pour les cg1.RecD^{B/S}, et 0/9 = 0% pour les cg1.RecM^{B/S} versus 1/59 = 1.7 % pour les femelles B6; Figure 3). Ce résultat m'a permis d'écarter l'hypothèse de plusieurs gènes liés à effets contradictoires.

2.3. Les trois sous-régions de cg1 ne présentent pas d'effet, même associées à cg3

Comme des interactions avec les autres QTL pourraient être nécessaires pour observer un effet de ces trois sous-régions de cg1 (comme pour cg2), j'ai produit des souris hétérozygotes à la fois pour l'une des trois sous-régions de cg1 et pour cg3. J'ai ainsi analysé la survie de souris femelles cg1.RecP^{B/S}-cg3^{B/S}, cg1.RecD^{B/S}-cg3^{B/S} et cg1.RecM^{B/S}-cg3^{B/S} par rapport aux femelles B6 et aux femelles cg3^{B/S} (Figure 4). Aucune de ces lignées ne présentait de différence significative de taux ou de durée de survie par rapport aux deux lignées témoins.

Le nombre trop restreint de souris obtenues dans certains groupes associant les sous-régions de cg1 avec cg2 ne nous a pas permis de conclure quant à d'éventuelles interactions entre *Ypr11* et *Ypr12* (ces données ne seront pas présentées ici).

L'absence d'effet du locus *Ypr11* dans la lignée cg1 n'est donc pas liée à la nécessité d'une potentialisation par la région *Ypr13*.

2.4. Aucun effet de cg1 n'est détecté sur un fonds génétique hybride

Comme il est décrit dans l'article, nous avons croisé les trois lignées congéniques établies pour *Ypr11*, 2 et 3 avec la lignée parentale SEG afin d'augmenter les chances de détecter l'effet d'un seul QTL. Les courbes de survie des souris femelles F1 portant *Ypr11* (F1-cg1^{S/S}), *Ypr12* (F1-cg2^{S/S}) et *Ypr13* (F1-cg3^{S/S}) ne diffèrent pas significativement de celle des femelles hybrides F1. Toutefois, comme le montre la figure 5, les souris F1-cg1^{S/S} ont un taux de survie plus bas que les F1 (1/9 = 11,1% versus 44,4%, respectivement) alors que les femelles F1-cg2^{S/S} et F1-cg3^{S/S} présentent un taux de survie légèrement plus élevé que les F1 (12/20 = 60% et 18/29 = 62%, respectivement). De plus, le taux de survie de ces deux derniers groupes est significativement plus élevé ($p < 0,02$) que celui des femelles F1-cg1^{S/S}. Ces résultats montrent que, même dans un fonds hybride qui se rapproche du fonds BSB dans lequel il a été détecté, *Ypr11* n'est pas capable d'accroître la résistance des souris sensibles.

2.5. Chez les femelles BSB *Yprl1* agirait plus tardivement que *Yprl2* et *Yprl3* sur le taux de survie

Les trois QTL rapportés dans l'article de Blanchet *et al.* ont été détectés en utilisant comme phénotype la survie à 14 jours. Ce phénotype ne prend pas en compte le moment auquel meurent les souris. Afin de rechercher s'il existe des différences de contrôle génétique entre les animaux qui meurent précocement ou plus tardivement, j'ai réalisé une analyse QTL à chaque jour de l'infection. Un test de permutation sur les données des 322 femelles *backcross* a été effectué à partir du 3^{ème} jour (jour à partir duquel les souris commencent à mourir) jusqu'au 14^{ème} jour afin de suivre l'apparition des QTL et leur évolution au cours de l'infection (Figure 6).

Cette analyse montre que les QTL *Yprl2* et *Yprl3*, localisés respectivement sur les chromosomes 4 et 6, atteignent le seuil de significativité ($P = 0.05$) au 6^{ème} jour. En revanche, la signification de *Yprl1* localisé sur le chromosome 3 croît plus lentement au cours de l'infection et atteint ce seuil au 10^{ème} jour (Figure 6).

De manière complémentaire j'ai réalisé une analyse de Chi2 en comparant la proportion de femelles survivantes dans les souris B/B et B/S à chaque QTL. Le taux de survie chez les femelles hétérozygotes B/S pour *Yprl3* devient statistiquement supérieur à celui des femelles B/B dès le 4^{ème} jour et dépasse nettement le seuil de significativité au 6^{ème} jour. Une différence significative est observée au 6^{ème} jour pour *Yprl2* et seulement au 9^{ème} jour pour *Yprl1*. La différence observée au 14^{ème} jour de l'infection est la même pour les trois QTL, mais cette différence est apparue plus tard pour *Yprl1* que pour les deux autres QTL. *Yprl1* semble avoir un effet plus tardif sur la survie.

Les différents croisements que j'ai réalisés ne m'ont pas permis de détecter un effet de *Yprl1* dans les lignées congéniques et sous-congéniques, alors que les résultats du *backcross* montraient que ce QTL expliquait 20% de la variance comme *Yprl2* et *Yprl3*. Devant l'ensemble de ces résultats négatifs, nous avons décidé de ne pas poursuivre l'étude du QTL *Yprl1*.

DISCUSSION GENERALE

« 1348. Les hommes et les femmes tombaient comme des mouches. La grande peste avait gagné l'ensemble du sud de la France. Le pape, alors, vivait en Avignon. Au cours de l'été 1348, soixante mille avignonnais allaient mourir du même mal. Plus de la moitié de la population de la ville. A l'abri des hauts murs de son palais, le pape Clément VI priait, lui aussi, pour le salut de son âme. »

Le Porteur d'histoire, Alexis Michalik

2. Discussion sur l'architecture génétique de la résistance à la peste chez SEG

2.1. Des lignées congéniques ne suffisent pas toujours pour disséquer des QTL

La démarche classique d'analyse des QTL contrôlant des caractères complexes consiste à développer des lignées congéniques dans lesquelles l'intervalle de confiance de chaque QTL identifié dans un croisement (F2 ou N2) est introduit dans l'un des deux fonds génétiques pour analyser son effet individuel.

Les résultats de mon travail montrent qu'un seul QTL est insuffisant pour induire, même partiellement, un phénotype de résistance, et qu'il est nécessaire d'associer deux QTL de résistance dans la même lignée pour induire une augmentation de survie détectable. Selon le mode d'action additif que suggèrent les résultats du *backcross*, on peut faire l'hypothèse que chaque QTL augmente modérément les capacités de survie des souris, mais insuffisamment pour vaincre le fort niveau de vulnérabilité de B6. Il a été nécessaire de produire et d'analyser des lignées bi-congéniques pour pouvoir conclure à l'existence réelle de ces QTL. Une telle situation n'est certainement pas exceptionnelle et il est probable qu'elle se soit produite dans une fraction non négligeable des travaux publiés qui ont conclu, par l'analyse de lignées mono-congéniques uniquement, que certains QTL sont de fausses associations (Bihl *et al.*, 1999).

2.1.1. *Ypr11* : une vraie fausse association ?

L'analyse des femelles issues du *backcross* de SEG par B6 avait mis en évidence trois QTL qui semblaient agir de manière additive. La création de lignées congéniques nous a permis d'en confirmer deux et de montrer qu'au moins deux gènes sont nécessaires dans chacun de ces QTL. Par contre, l'effet du QTL *Ypr11* n'a pas pu être confirmé malgré la production de lignées congéniques portant ce QTL seul à l'état hétérozygote ou homozygote, ou en association avec *Ypr13*, ni même dans un fonds hybride (F1). Compte tenu des LOD-scores obtenus dans l'analyse du *backcross* (juste au seuil de significativité), la détection d'une fausse association est tout à fait possible.

Cependant, en menant une analyse de QTL sur la survie à chaque jour de l'infection, nous avons remarqué que *Ypr11* (qui présente un LOD score similaire aux deux autres QTL au 14^{ème} jour) atteint le seuil de significativité 4 jours après *Ypr12* et *Ypr13*. Cet effet retardé a-t-il une raison biologique ou statistique ? *Ypr11* pourrait avoir une implication fonctionnelle dans une phase plus tardive de l'infection (comme la mise en place de la réponse adaptative ou la réparation des lésions),

alors que les deux autres QTL pourraient intervenir dans des processus plus précoces (tels que la reconnaissance du pathogène, la mise en place de l'inflammation, le recrutement des cellules de l'immunité innée, etc.).

Une implication biologique tardive pourrait être la raison pour laquelle on ne retrouve pas l'effet de *Yprl1* sur la survie chez les congéniques qui le portent de manière isolée. En effet, ces souris mourraient avant d'avoir pu bénéficier des avantages conférés par *Yprl1*. Cependant, si *Yprl1* avait un effet protecteur dans une phase tardive de l'infection, nous aurions dû voir son effet dans la lignée tri-congénique $cg1^{B/S}-cg2^{B/S}-cg3^{B/S}$, qui présente un bon potentiel de résistance conféré par l'action combinée de *Yprl2* et *Yprl3*. Or les souris tri-congéniques ne survivaient pas plus que les souris bi-congéniques $cg2^{B/S}-cg3^{B/S}$, ce qui suggère que *Yprl1* pourrait être simplement une fausse association détectée dans la population *backcross*.

2.1.2. La nécessité d'identifier des sous-phénotypes quantitatifs

La survie est un phénotype global et fruste, pour lequel chaque souris infectée apporte une information binaire, ce qui requiert de tester un grand nombre d'animaux pour avoir des chances de voir un effet. Par exemple, il faut environ une centaine d'animaux par groupe (sujets et témoins) pour détecter une différence significative de la survie de 15 %. Il est possible que nous n'ayons pas pu mettre en évidence des effets faibles chez les congéniques car nous n'avons pas testé un assez grand nombre d'animaux. Malheureusement ces animaux n'étaient pas toujours disponibles. En effet, je me suis heurtée, dans la production des souris congéniques, à une faible fertilité des reproducteurs (voire une infertilité pour certaines lignées), un inconvénient qui est bien connu dans les croisements interspécifiques (Burgio *et al.*, 2007; Dejager *et al.*, 2009; Guenet *et al.*, 2003). L'obtention des lignées congéniques a souvent été compliquée (petites portées, stérilité dans un des deux sexes, problèmes de développement ou encore de dystocie). Ces difficultés d'élevage expliquent les petits nombres de souris analysées pour certaines lignées et l'absence d'animaux homozygotes pour les QTL *Yprl1* et *Yprl2*.

L'analyse de paramètres quantitatifs augmente l'informativité de chaque animal qui est caractérisé par une ou plusieurs variables continues. De tels paramètres permettent de définir plus finement le phénotype d'un individu que l'issue (mort ou survie) de l'infection. Par ailleurs, cette issue finale ne reflète pas la complexité des mécanismes mis en jeu en réponse à l'infection par la bactérie. Ainsi, deux souris peuvent succomber le même jour alors que l'une des deux a développé une réponse inflammatoire, certes inefficace, mais supérieure à l'autre souris. Cette différence peut témoigner d'un certain niveau de résistance qui passe inaperçu si l'on ne fait qu'enregistrer la mort des animaux. Ces paramètres discrets qui participent au phénotype global de résistance (ou vulnérabilité) sont appelés sous-phénotypes.

Mon analyse des lignées congéniques et sous-congéniques a porté uniquement sur la survie des animaux, mise à part l'étude de l'apoptose et de la production des cytokines par les macrophages des lignées bi-congénique $cg2^{B/S}-cg3^{B/S}$ et tri-congénique $cg1^{B/S}-cg2^{B/S}-cg3^{B/S}$. L'analyse de sous-phénotypes quantitatifs permettrait peut-être de mettre en évidence un effet individuel et spécifique des QTL *Yrpl2* et *Yprl3*, et peut-être un effet pour *Yprl1*. En effet, dans l'état actuel de nos données, chacune des 4 régions dont nous avons démontré l'effet sur la survie apparaît indistinctement associée au même phénotype, alors qu'il est très vraisemblable que chacune agit sur une composante spécifique de la réponse de l'hôte. C'est ce qu'il sera nécessaire d'étudier dans l'avenir.

Le travail qu'a réalisé le groupe d'E. Wakeland et L. Morel dans la dissection du lupus érythémateux disséminé illustre parfaitement l'utilisation de sous-phénotypes pour associer une fonction à chaque QTL. Partant d'un phénotype global (le développement de la glomérulonéphrite qui est une complication habituelle du lupus), ils ont identifié quatre locus (Morel *et al.*, 1994). Ils ont créé des lignées congéniques pour ces locus et les ont été analysées pour le développement d'une glomérulonéphrite lupique mais également pour différents paramètres immunologiques reflétant le niveau d'activation non spécifique de la réponse immunitaire et la perte de tolérance aux antigènes du soi (taux d'IgM et d'IgG, IgG dirigées contre des antigènes nucléaires, par exemple). Seule une lignée congénique montrait une augmentation de la fréquence de développement d'une glomérulonéphrite lupique par rapport à la lignée receveuse. En revanche, toutes les lignées présentaient des anomalies immunologiques variées qui ont permis d'attribuer un rôle propre à chaque région (Morel *et al.*, 1997). En réunissant les trois régions dans une seule lignée tri-congénique, les auteurs sont parvenus à reconstruire le phénotype de glomérulonéphrite lupique avec une fréquence identique à celle de la lignée parentale (Morel *et al.*, 2000a).

Dans notre étude, des paramètres tels que la charge bactérienne dans différents organes, les taux de cytokines, les niveaux d'activation de différentes populations cellulaires (sous-phénotypes qui présentent des différences dans les lignées parentales (Demeure *et al.*, 2012)) pourront être analysés pour tenter caractériser l'effet biologique précis de chaque QTL.

De plus, Toan Nham, dans l'unité des *Yersinia*, a créé une bactérie luminescente qui permet de suivre la dissémination bactérienne par imagerie *in vivo* (Nham *et al.*, 2012). Demeure *et al.* ont montré que l'apparition d'un pic précoce de bactériémie chez SEG pourrait être impliquée dans la survie de ces souris (Demeure *et al.*, 2012). Ainsi, caractériser la dissémination bactérienne chez SEG, B6 et nos congéniques avec cette bactérie luminescente pourrait permettre d'identifier des différences de cinétique.

L'unité des *Yersinia* a montré que les macrophages péritonéaux de SEG et B6 exposés *ex vivo* à *Y. pestis* présentaient des caractéristiques distinctes d'apoptose et de production des cytokines (Pachulec *et al.*, en préparation). Toutefois, à ce jour aucune différence n'a été observée pour ces

caractères entre les lignées bi-congéniques $cg2^{B/S}-cg3^{B/S}$ et tri-congéniques $cg1^{B/S}-cg2^{B/S}-cg3^{B/S}$ et la lignée B6 (données non présentées).

Il pourrait également être intéressant de tester le phénotype de survie des souris congéniques avec des doses plus faibles de bactéries (ou avec une souche de virulence atténuée). Pris individuellement l'effet de chaque QTL est trop faible pour induire une augmentation de la survie dans les conditions expérimentales que nous avons utilisées depuis le début du projet, choisies pour montrer une différence nette entre les lignées parentales B6 et SEG. Il n'est pas exclu que ces QTL soient capable d'induire une augmentation détectable de la survie en réponse à une dose infectieuse plus faible. Il est possible que certaines combinaisons de sous-congéniques, telles que la lignée $cg2.RecP^{B/S}-cg3^{B/S}$, qui présentent une latence significativement augmentée par rapport à la lignée parentale B6, présentent une survie accrue dans des conditions d'infection plus modérées.

2.1.3. L'utilisation d'un fonds hybride (F1) pour disséquer *Yprl3*

L'observation des données du *backcross* suggéraient que des facteurs génétiques en dehors des QTL identifiés intervenaient dans le phénotype de survie des souris SEG puisque les souris homozygotes B/B aux trois QTL détectés survivaient à un taux supérieur à celui des B6. Les congéniques ont confirmé cette hypothèse puisque les souris bi-congéniques $cg2^{B/S}-cg3^{B/S}$ présentaient un taux de survie inférieur à celui des SEG. L'absence d'effet sur la survie de ces deux QTL pris isolément dans un fonds B6 suggère que ce fonds porte trop d'allèles de vulnérabilité pour qu'un seul gène de résistance apporté par SEG soit capable d'induire une augmentation de survie détectable. A l'inverse, l'effet individuel des QTL a pu être détecté dans le *backcross* car ces souris portent un fonds génétique contenant en moyenne 25% d'allèles SEG. Pour essayer de détecter un effet des QTL isolés dans les lignées congéniques, nous avons cherché à travailler sur un fonds comportant "une certaine proportion" d'allèles SEG, apportant un niveau de survie basal supérieur à celui de B6.

C'est pourquoi j'ai croisé les lignées congéniques avec des souris SEG pour étudier l'effet des QTL dans un fonds hybride ($B6 \times SEG$)F1 qui présente une résistance intermédiaire. Cependant, les femelles F1- $cg2^{S/S}$ et F1- $cg3^{S/S}$ présentaient un niveau de survie trop proche de celui des femelles F1 pour pouvoir distinguer une augmentation significative de la survie. Il en était de même pour *Yprl2* chez les mâles. En revanche, le niveau de survie des mâles portant *Yprl3* à l'état homozygote était significativement supérieur à celui des mâles F1, montrant que ce locus confère un effet détectable à lui seul à condition d'être dans un contexte génomique favorable. J'ai ainsi pu montrer que ce QTL agit de façon semi-dominante puisque les trois génotypes conduisent à des phénotypes différents. Cette stratégie d'étude des QTL à l'aide d'un fonds génétique hybride entre les lignées parentales sensible et résistante s'est donc révélée utile et pourrait être appliquée à d'autres études de QTL.

Nous avons débuté la dissection du QTL *Yprl3* sur un fonds F1 en sélectionnant par génotypage les mâles F1 recombinants pour l'intervalle congénique cg3. Nous avons ainsi pu obtenir une dizaine de groupes d'animaux (entre 5 et 15 animaux par génotype) présentant des recombinaisons différentes entre les mêmes marqueurs dans l'intervalle cg3. L'analyse de ces souris ne nous a toutefois pas permis de réduire l'intervalle de confiance de *Yprl3* chez les mâles à cause d'une variabilité de survie des souris entre les dernières expériences d'infection qui a mené à l'obtention de résultats discordants.

2.1.4. Fluctuations du phénotype de résistance des souris SEG

Le degré de résistance des différentes lignées congéniques étudiées a été mesuré au cours de plusieurs expériences qui ont été effectuées à mesure que de nouveaux animaux étaient disponibles, tout au long de mon travail de thèse. Dans chaque expérience d'infection, j'ai inclus un groupe de souris B6 et de souris SEG pour valider l'efficacité de l'infection.

Au cours des trois années, alors que les souris B6 se sont montrées uniformément sensibles à l'infection, j'ai observé des fluctuations parfois importantes du taux de survie des souris SEG. Ce taux a parfois chuté jusqu'à 30% (alors qu'il se situe entre 80 et 95% chez les femelles) pour des raisons inexpliquées. Dans ce cas, je n'ai pas incorporé les données des congéniques provenant de ces expériences. Ce problème a rendu difficile l'interprétation de certaines données concernant les congéniques sur fonds F1.

L'évolution globale montre une diminution du niveau de résistance des souris SEG au fil du temps. Plusieurs hypothèses peuvent être avancées pour tenter d'expliquer cette sensibilité accrue et fluctuante des souris SEG. La première est une contamination génétique éventuelle de la lignée SEG par une lignée sensible. Les modalités pratiques de son élevage, et la stérilité résultant des croisements entre cette lignée et les lignées de laboratoire (par exemple) rendent cette hypothèse très improbable. La seconde concerne la préparation de la bactérie utilisée pour les infections. La numération bactérienne varie inévitablement entre les expériences successives mais la mortalité de souris SEG n'était pas corrélée à ces variations (données non présentées). La troisième hypothèse concerne des variations du statut microbiologique ou immunitaire des souris SEG. Toutefois, ces souris ont toujours été élevées dans la même animalerie, par la même technicienne, et les analyses microbiologiques de souris sentinelles de l'animalerie n'a pas révélé d'infection. Ces variations de survie n'ont donc pas été élucidées pour le moment, mais présentent malheureusement un impact important sur nos études.

2.2. Plus de gènes que de QTL

En fragmentant les grands intervalles portés par les lignées *cg2* et *cg3*, il est apparu que chacun contenait au moins deux gènes indispensables au phénotype de survie partielle observé dans les souris bi-congéniques (respectivement dans les sous-régions *Yprl2*.RecP et *Yprl2*.RecD, et *Yprl3*.RecP et *Yprl3*.RecD). Cette observation s'ajoute aux études, déjà nombreuses, qui indiquent clairement qu'un QTL identifié dans un croisement en deux générations à l'intérieur d'un large intervalle s'avère fréquemment, lors de l'analyse de lignées congéniques, résulter de l'action combinée de plusieurs gènes liés (Ahn *et al.*, 2010; Bihl *et al.*, 1999; Morel *et al.*, 2001; Roy *et al.*, 2002; Sancho-Shimizu *et al.*, 2007). Deux hypothèses peuvent être avancées pour expliquer ces observations. La première est que notre connaissance actuelle des QTL est biaisée par notre capacité à les détecter. Lorsque deux gènes ou davantage, liés génétiquement, agissent de façon concertée pour contrôler un caractère, leurs effets combinés se traduisent, dans une étude d'association, par un LOD score généralement plus élevé que dans le cas d'un gène isolé. Ces QTL sont donc plus facilement et fréquemment détectés. Dans ce cas, le pic de LOD score du QTL ne se trouve généralement pas à la position de l'un des gènes contribuant au QTL mais à une position intermédiaire (Figure 7) (Thèse de Marie Abitbol, 2004). La forme en dôme plat des courbes de LOD score de *Yprl2* et *Yprl3* dans le *backcross* est cohérente avec l'observation, grâce aux lignées sous-congéniques, que chaque QTL correspond à l'action d'au moins deux gènes.

La deuxième hypothèse vient de l'observation que les gènes qui participent à une même grande fonction (cytokines, canaux ioniques, récepteurs, etc.) sont fréquemment regroupés dans le génome en clusters et combinent leurs effets. Sur le plan évolutif, cette organisation favorise la transmission de combinaisons alléliques (des haplotypes) associées à des phénotypes avantageux. L'observation qu'un QTL recouvre fréquemment plusieurs gènes causaux ne relèverait plus d'un biais de détection des QTL mais reflèterait une organisation particulière des gènes contrôlant certaines grandes fonctions physiologiques.

La figure 4, présente la position des principaux gènes candidats dans *Yprl2* et *Yprl3*. Elle montre que les quatre intervalles contiennent des gènes capables de coopérer pour apporter un phénotype de résistance : *Tlr4* et *Map3k7* pour *Yprl2*, l'un étant dans l'intervalle *Yprl2*.RecP et l'autre dans *Yprl2*.RecD ; *Irf5* et *Nod1* dans *Yprl3*.RecP ; *Hdac11* et *Irak2* dans *Yprl3*.RecD.

La résistance à la peste chez SEG est un caractère complexe où plusieurs gènes à effet modéré ou faible agissent ensemble pour augmenter le niveau de résistance. Les effets individuels ne sont pas détectables en utilisant un phénotype aussi global et grossier que la survie. C'est pourquoi, la suite de cette dissection génétique passera par l'identification de sous-phénotypes quantitatifs, qui permettront, en utilisant un nombre limité de souris, de caractériser l'action biologique des quatre locus identifiés. Aller jusqu'aux gènes sous-jacents aux QTL *Yprl2* et *Yprl3* nécessitera certainement de combiner différentes approches comme celles qui ont été décrites dans l'introduction.

2.3. Pas d'épistasie ?

Il peut paraître surprenant d'arriver pratiquement au terme de ce document et d'évoquer pour la première fois l'épistasie dont on ne cesse de mettre en évidence l'importance majeure dans le contrôle génétique des caractères complexes.

Pourtant, nos résultats peuvent être expliqués sans l'intervention d'interactions épistatiques entre gènes non liés. Plus encore, nous avons recherché en vain une éventuelle épistasie dans le *backcross* et n'avons trouvé que des interactions additives (Blanchet *et al.*, 2010). Les données des lignées congéniques simples ou multiples, peuvent également s'expliquer dans un modèle purement additif, même si la survie requiert l'intervention d'un nombre minimal de quatre facteurs génétiques, l'absence d'un seul d'entre eux faisant disparaître complètement le potentiel de résistance. Je fais l'hypothèse que la survie des seules double congéniques cg2-cg3 est due à un seuil de résistance qu'il est nécessaire de franchir pour acquérir une certaine capacité à résister à l'infection, seuil qui est atteint par l'apport cumulé des deux QTL, mais pas par l'action d'un seul. En revanche, j'ai pu montrer que *Yprl3* était capable d'apporter, en l'absence de *Yprl2*, un faible niveau de résistance, que ce soit par un retard à la mort chez les mâles congéniques homozygotes ou par une survie accrue chez les mâles congénique sur fonds F1.

3. Conclusion générale

L'étude croisée des données de transcriptome et des résultats obtenus sur les lignées congéniques m'a conduit à proposer plusieurs gènes qui pourraient être responsables d'une partie de la différence de phénotype observée entre SEG et B6. Le rôle éventuel de ces gènes devrait dans un premier temps être testé en comparant les lignées parentales, puis les lignées congéniques. Compte tenu de nos observations, il est indispensable de veiller à ce que les souris comparées soient dans un état microbiologique et immunologique aussi homogène que possible.

Il ne faudra pas pour autant négliger une approche purement génétique qui peut permettre de réduire les intervalles génomiques dans lesquels se trouvent les gènes sous-jacents aux QTLs. La dissection génétique de *Yprl3* dans un fonds génétique F1 semble réalisable, à condition toutefois de parvenir à mieux contrôler la reproductibilité des infections et à identifier si possible un paramètre quantitatif pour tirer le meilleur parti de chaque souris disponible.

Ensuite, il sera nécessaire de valider fonctionnellement les candidats les plus convaincants. Une approche possible serait de réaliser un test de complémentation quantitative (Su *et al.*, 2009) en utilisant des souris KO pour le gène à tester. Les grands consortium Européen (EUCOMM) et Nord-Américain (KOMP) produisent des allèles nuls pour tous les gènes du génome de la souris. Ces allèles sont créés sur un fonds B6. En croisant des souris hétérozygotes (+/-) avec des souris SEG, on pourra produire des souris portant soit l'allèle B6, soit allèle inactivé en face de l'allèle SEG. En comparant le phénotype des deux types de souris en réponse à l'infection par *Y. pestis*, on pourra démontrer l'implication de ce gène. Là encore, il sera précieux de pouvoir analyser des sous-phénotypes (y compris transcriptionnels) pour augmenter la puissance de détection d'effets faibles.

Des tests de complémentation fonctionnelle d'autre nature pourraient aussi être envisagés, par exemple la transfection de macrophages de souris B6 avec un gène candidat sous forme allélique SEG, suivi de la mesure de différentes cytokines pro-inflammatoires.

Une autre approche serait d'inhiber *in vitro* ou *ex vivo* l'expression des gènes candidats identifiés dans des macrophages de SEG et B6 par ARN interférence. Il serait par exemple intéressant d'inhiber l'action de l'histone déacétylase 11 dans des macrophages de SEG et de B6 et de comparer les niveaux d'IL-10 (et de cytokines pro-inflammatoires synthétisées) par rapport à des macrophages non inhibés.

Deux autres projets en cours au laboratoire pourraient nous aider à discerner la valeur des candidats que nous avons identifiés.

Le premier est l'analyse des données de séquençage de SEG qui nous permettra très prochainement d'avoir un inventaire complet des variations de séquence entre SEG et B6. Cependant, le taux élevé de polymorphisme nucléotidique entre ces lignées risque fort de rendre difficile l'identification des polymorphismes causaux. C'est en effet une autre limite de l'utilisation de lignées hautement divergentes, chez lesquelles distinguer une mutation causale parmi la multitude de variants n'est pas une tâche aisée.

Le deuxième projet, prévu pour l'année à venir, est de réaliser l'étude simultanée des transcriptomes de l'hôte et de *Y. pestis*, par séquençage de l'ensemble des ARN, dans la rate de souris SEG et B6 après infection. La comparaison *in vivo* du profil transcriptionnel de la rate révélera peut-être des gènes communs avec nos candidats, ou des mécanismes biologiques communs.

Mon travail illustre la difficulté d'identifier les facteurs génétiques à l'origine des QTL. C'est l'étape la plus compliquée de l'analyse QTL. Seul 1% des QTG (*Quantitative Trait Gene*) à l'origine des QTL ont été caractérisés (que ce soit chez l'animal de laboratoire, chez l'Homme, ou encore chez les plantes), (DiPetrillo *et al.*, 2005; Hunter *et al.*, 2008; Korstanje *et al.*, 2002; Roff, 2007).

L'identification des gènes à effet mineur est une étape difficile, car il est très fréquent de ne pas retrouver les effets phénotypiques lorsque le gène causal est isolé de son contexte génétique naturel, soit parce que son effet propre est trop faible, soit en raison d'interactions épistatiques avec d'autres gènes dont il a été séparé. Pour les mêmes raisons, l'utilisation de souris KO pour tester un gène candidat peut également se révéler décevante dans les modèles hautement polygéniques.

L'identification de gènes à effet très fort, tel qu'*Oas1b* qui contrôle de manière pratiquement binaire la survie au virus *West Nile* et aux autres *flaviviridae* (Mashimo *et al.*, 2002; Pereygin *et al.*, 2002; Scherbik *et al.*, 2007; Simon-Chazottes *et al.*, 2011) ou *Mx1*, qui contrôle la survie à une infection par des *orthomyxoviridae* (Vanlaere *et al.*, 2008) est une situation exceptionnelle dans le contrôle génétique de la vulnérabilité aux maladies infectieuses.

En conclusion, les résultats de notre travail illustrent la contribution que l'étude d'un modèle murin peut apporter à la compréhension et à l'identification de gènes de vulnérabilité à une maladie infectieuse telle que la peste. Ces travaux permettront peut-être à terme la caractérisation de facteurs génétiques capables de conférer une résistance exceptionnelle à *Y. pestis* ou à d'autres maladies bactériennes aiguës. Même si les variants alléliques qui seront identifiés chez la souris ne seront pas nécessairement retrouvés chez l'Homme, ils fourniront des indices précieux sur les voies de signalisation et les mécanismes physiopathologiques qui, eux, présentent un degré de conservation élevé et pourront servir de pistes pour des thérapeutiques complémentaires et adjuvantes aux traitements antibiotiques de la peste.

REFERENCES BIBLIOGRAPHIQUES

Références

- Abdeltawab NF, Aziz RK, Kansal R, Rowe SL, Su Y *et al.* (2008). An unbiased systems genetics approach to mapping genetic loci modulating susceptibility to severe streptococcal sepsis. *PLoS pathogens* 4, e1000042.
- Agar SL, Sha J, Foltz SM, Erova TE, Walberg KG *et al.* (2009). Characterization of the rat pneumonic plague model: infection kinetics following aerosolization of *Yersinia pestis* CO92. *Microbes and infection / Institut Pasteur* 11, 205-214.
- Ahn SH, Deshmukh H, Johnson N, Cowell LG, Rude TH *et al.* (2010). Two genes on A/J chromosome 18 are associated with susceptibility to *Staphylococcus aureus* infection by combined microarray and QTL analyses. *PLoS pathogens* 6, e1001088.
- Aitman TJ, Boone C, Churchill GA, Hengartner MO, Mackay TF *et al.* (2011). The future of model organisms in human disease research. *Nature reviews. Genetics* 12, 575-582.
- Akira S, Uematsu S, Takeuchi O (2006). Pathogen recognition and innate immunity. *Cell* 124, 783-801.
- Alberts R, Lu L, Williams RW, Schughart K (2011). Genome-wide analysis of the mouse lung transcriptome reveals novel molecular gene interaction networks and cell-specific expression signatures. *Respiratory research* 12, 61.
- Alberts R, Terpstra P, Bystrykh LV, de Haan G, Jansen RC (2005). A statistical multiprobe model for analyzing cis and trans genes in genetical genomics experiments with short-oligonucleotide arrays. *Genetics* 171, 1437-1439.
- Alberts R, Terpstra P, Li Y, Breitling R, Nap JP *et al.* (2007). Sequence polymorphisms cause many false cis eQTLs. *PloS one* 2, e622.
- Alcais A, Sanchez FO, Thuc NV, Lap VD, Oberti J *et al.* (2000). Granulomatous reaction to intradermal injection of lepromin (Mitsuda reaction) is linked to the human NRAMP1 gene in Vietnamese leprosy sibships. *The Journal of infectious diseases* 181, 302-308.
- Alexopoulou L, Holt AC, Medzhitov R, Flavell RA (2001). Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* 413, 732-738.
- Almasy L (2012). The role of phenotype in gene discovery in the whole genome sequencing era. *Human genetics*
- Arbaji A, Kharabsheh S, Al-Azab S, Al-Kayed M, Amr ZS *et al.* (2005). A 12-case outbreak of pharyngeal plague following the consumption of camel meat, in north-eastern Jordan. *Annals of tropical medicine and parasitology* 99, 789-793.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H *et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25, 25-29.
- Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ *et al.* (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nature reviews. Genetics* 12, 745-755.
- Barreiro LB, Ben-Ali M, Quach H, Laval G, Patin E *et al.* (2009). Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS genetics* 5, e1000562.
- Bartra SS, Styer KL, O'Bryant DM, Nilles ML, Hinnebusch BJ *et al.* (2008). Resistance of *Yersinia pestis* to complement-dependent killing is mediated by the Ail outer membrane protein. *Infection and immunity* 76, 612-622.
- Bauer S, Kirschning CJ, Hacker H, Redecke V, Hausmann S *et al.* (2001). Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. *Proceedings of the National Academy of Sciences of the United States of America* 98, 9237-9242.

- Beaumont C, Protais J, Pitel F, Leveque G, Malo D *et al.* (2003). Effect of two candidate genes on the Salmonella carrier state in fowl. *Poultry science* 82, 721-726.
- Bertherat E, Lamine KM, Formenty P, Thuier P, Mondonge V *et al.* (2005). [Major pulmonary plague outbreak in a mining camp in the Democratic Republic of Congo: brutal awakening of an old scourge]. *Medecine tropicale : revue du Corps de sante colonial* 65, 511-514.
- Bihl F, Brahic M, Bureau JF (1999). Two loci, Tmevp2 and Tmevp3, located on the telomeric region of chromosome 10, control the persistence of Theiler's virus in the central nervous system of mice. *Genetics* 152, 385-392.
- Blackwell JM, Roberts CW, Roach TI, Alexander J (1994 B). Influence of macrophage resistance gene Lsh/Ity/Bcg (candidate Nramp) on Toxoplasma gondii infection in mice. *Clinical and experimental immunology* 97, 107-112.
- Blanchet C, Jaubert J, Carniel E, Fayolle C, Milon G *et al.* (2011). Mus spretus SEG/Pas mice resist virulent Yersinia pestis, under multigenic control. *Genes and immunity* 12, 23-30.
- Bliska JB (2006). Yersinia inhibits host signaling by acetylating MAPK kinases. *ACS chemical biology* 1, 349-351.
- Blisnick T, Ave P, Huerre M, Carniel E, Demeure CE (2008). Oral vaccination against bubonic plague using a live avirulent Yersinia pseudotuberculosis strain. *Infection and immunity* 76, 3808-3816.
- Boegler KA, Graham CB, Montenieri JA, Macmillan K, Holmes JL *et al.* (2012). Evaluation of the Infectiousness to Mice of Soil Contaminated with Yersinia pestis-Infected Blood. *Vector borne and zoonotic diseases (Larchmont, N.Y.)*
- Bohn E, Muller S, Lauber J, Geffers R, Speer N *et al.* (2004). Gene expression patterns of epithelial cells modulated by pathogenicity factors of Yersinia enterocolitica. *Cellular microbiology* 6, 129-141.
- Boldrick JC, Alizadeh AA, Diehn M, Dudoit S, Liu CL *et al.* (2002). Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 99, 972-977.
- Bonhomme F, Guénet JL (1987). The polyphyletic origin of laboratory inbred mice and their rate of evolution. *Biol J Lin Soc* 30, 51-58.
- Bosio CF, Jarrett CO, Gardner D, Hinnebusch BJ (2012). Kinetics of the Innate Immune Response to Yersinia pestis after Intradermal Infection in a Mouse Model. *Infection and immunity*
- Bottomly D, Ferris MT, Aicher LD, Rosenzweig E, Whitmore A *et al.* (2012). Expression quantitative trait Loci for extreme host response to influenza a in pre-collaborative cross mice. *G3 (Bethesda, Md.)* 2, 213-221.
- Bowie A, O'Neill LA (2000). The interleukin-1 receptor/Toll-like receptor superfamily: signal generators for pro-inflammatory interleukins and microbial products. *Journal of leukocyte biology* 67, 508-514.
- Boyle EI, Weng S, Gollub J, Jin H, Botstein D *et al.* (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics (Oxford, England)* 20, 3710-3715.
- Brahic M, Bureau JF (1998). Genetics of susceptibility to Theiler's virus infection. *BioEssays : news and reviews in molecular, cellular and developmental biology* 20, 627-633.
- Breitling R, Armengaud P, Amtmann A, Herzyk P (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS letters* 573, 83-92.
- Brem RB, Yvert G, Clinton R, Kruglyak L (2002). Genetic dissection of transcriptional regulation in budding yeast. *Science (New York, N.Y.)* 296, 752-755.
- Brinkmann V, Zychlinsky A (2007). Beneficial suicide: why neutrophils die to make NETs. *Nature reviews. Microbiology* 5, 577-582.

Bruckner S, Rhamouni S, Tautz L, Denault JB, Alonso A *et al.* (2005). Yersinia phosphatase induces mitochondrially dependent apoptosis of T cells. *The Journal of biological chemistry* 280, 10388-10394.

Bubeck SS, Cantwell AM, Dube PH (2007). Delayed inflammatory response to primary pneumonic plague occurs in both outbred and inbred mice. *Infection and immunity* 75, 697-705.

Burgio G, Szatanik M, Guenet JL, Arnau MR, Panthier JJ *et al.* (2007). Interspecific recombinant congenic strains between C57BL/6 and mice of the *Mus spretus* species: a powerful tool to dissect genetic control of complex traits. *Genetics* 177, 2321-2333.

Bustamante J, Arias AA, Vogt G, Picard C, Galicia LB *et al.* (2011). Germline CYBB mutations that selectively affect macrophages in kindreds with X-linked predisposition to tuberculous mycobacterial disease. *Nature immunology* 12, 213-221.

Cameron JE, Yin Q, Fewell C, Lacey M, McBride J *et al.* (2008). Epstein-Barr virus latent membrane protein 1 induces cellular MicroRNA miR-146a, a modulator of lymphocyte signaling pathways. *Journal of virology* 82, 1946-1958.

Caron J, Loredó-Osti JC, Laroche L, Skamene E, Morgan K *et al.* (2002). Identification of genetic loci controlling bacterial clearance in experimental *Salmonella enteritidis* infection: an unexpected role of Nramp1 (*Slc11a1*) in the persistence of infection in mice. *Genes and immunity* 3, 196-204.

Caron J, Loredó-Osti JC, Morgan K, Malo D (2005). Mapping of interactions and mouse congenic strains identified novel epistatic QTLs controlling the persistence of *Salmonella Enteritidis* in mice. *Genes and immunity* 6, 500-508.

Casanova JL, Abel L (2004). The human model: a genetic dissection of immunity to infection in natural conditions. *Nature reviews. Immunology* 4, 55-66.

Casanova JL, Abel L (2007). Human genetics of infectious diseases: a unified theory. *The EMBO journal* 26, 915-922.

Castelli J, Wood KA, Youle RJ (1998). The 2-5A system in viral infection and apoptosis. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* 52, 386-390.

Chakravarti A, Allaey I, Poubelle PE (2007). [Neutrophils and immunity: is it innate or acquired?]. *Medecine sciences : M/S* 23, 862-867.

Champion JA, Mitragotri S (2006). Role of target geometry in phagocytosis. *Proceedings of the National Academy of Sciences of the United States of America* 103, 4930-4934.

Chang YF, Imam JS, Wilkinson MF (2007). The nonsense-mediated decay RNA surveillance pathway. *Annual review of biochemistry* 76, 51-74.

Chanteau S, Nato F, Migliani R (2003 A). [Interest in rapid immunochromatography tests for surveillance of characteristic diseases epidemic in developing countries: the example of plague in Madagascar]. *Medecine tropicale : revue du Corps de sante colonial* 63, 574-576.

Chanteau S, Rahalison L, Ralafiarisoa L, Foulon J, Ratsitorahina M *et al.* (2003 B). Development and testing of a rapid diagnostic test for bubonic and pneumonic plague. *Lancet* 361, 211-216.

Chapman SJ, Hill AV (2012). Human genetic susceptibility to infectious disease. *Nature reviews. Genetics* 13, 175-188.

Chaussabel D, Semnani RT, McDowell MA, Sacks D, Sher A *et al.* (2003). Unique gene expression profiles of human macrophages and dendritic cells to phylogenetically distinct parasites. *Blood* 102, 672-681.

Chen G, Shaw MH, Kim YG, Nunez G (2009). NOD-like receptors: role in innate immunity and inflammatory disease. *Annual review of pathology* 4, 365-398.

Chevallier L, Blanchet C, Jaubert J, Pachulec E, Demeure C *et al.* (2012). Resistance to plague of *Mus spretus* SEG/Pas mice requires the combined action of at least four genetic factors. *Genes and immunity* in press

- Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR *et al.* (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 106, 19096-19101.
- Chromy BA, Perkins J, Heidbrink JL, Gonzales AD, Murphy GA *et al.* (2004). Proteomic characterization of host response to *Yersinia pestis* and near neighbors. *Biochemical and biophysical research communications* 320, 474-479.
- Chu WM (2012). Tumor necrosis factor. *Cancer letters*
- Cirulli ET, Singh A, Shianna KV, Ge D, Smith JP *et al.* (2010). Screening the human exome: a comparison of whole genome and whole transcriptome sequencing. *Genome biology* 11, R57.
- Clee SM, Yandell BS, Schueler KM, Rabaglia ME, Richards OC *et al.* (2006). Positional cloning of *Sorcs1*, a type 2 diabetes quantitative trait locus. *Nature genetics* 38, 688-693.
- Comer JE, Sturdevant DE, Carmody AB, Virtaneva K, Gardner D *et al.* (2010). Transcriptomic and innate immune responses to *Yersinia pestis* in the lymph node during bubonic plague. *Infection and immunity* 78, 5086-5098.
- Congleton YH, Wulff CR, Kerschen EJ, Straley SC (2006). Mice naturally resistant to *Yersinia pestis* Delta pgm strains commonly used in pathogenicity studies. *Infection and immunity* 74, 6501-6504.
- Conner JR, Smirnova, II, Poltorak A (2009). A mutation in *Irak2c* identifies IRAK-2 as a central component of the TLR regulatory network of wild-derived mice. *The Journal of experimental medicine* 206, 1615-1631.
- Cooper GM, Shendure J (2011). Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nature reviews. Genetics* 12, 628-640.
- Cornelis GR (2002). The *Yersinia* YSC-YOP 'type III' weaponry. *Nature Reviews Molecular Cell Biology* 3, 742-752.
- Cornelis GR (2002 A). The *Yersinia* Ysc-Yop 'type III' weaponry. *Nature reviews. Molecular cell biology* 3, 742-752.
- Cornelis GR (2002 B). *Yersinia* type III secretion: send in the effectors. *The Journal of cell biology* 158, 401-408.
- Cornelis GR (2002 C). The *Yersinia* Ysc-Yop virulence apparatus. *International journal of medical microbiology : IJMM* 291, 455-462.
- Croker B, Crozat K, Berger M, Xia Y, Sovath S *et al.* (2007). ATP-sensitive potassium channels mediate survival during infection in mammals and insects. *Nature genetics* 39, 1453-1460.
- Crozat K, Georgel P, Rutschmann S, Mann N, Du X *et al.* (2006). Analysis of the MCMV resistome by ENU mutagenesis. *Mammalian genome : official journal of the International Mammalian Genome Society* 17, 398-406.
- Crozat K, Hoebe K, Ugolini S, Hong NA, Janssen E *et al.* (2007). Jinx, an MCMV susceptibility phenotype caused by disruption of *Unc13d*: a mouse model of type 3 familial hemophagocytic lymphohistiocytosis. *The Journal of experimental medicine* 204, 853-863.
- Cusson-Hermance N, Khurana S, Lee TH, Fitzgerald KA, Kelliher MA (2005). *Rip1* mediates the Trif-dependent toll-like receptor 3- and 4-induced NF- κ B activation but does not contribute to interferon regulatory factor 3 activation. *The Journal of biological chemistry* 280, 36560-36566.
- Darvasi A (1998). Experimental strategies for the genetic dissection of complex traits in animal models. *Nature genetics* 18, 19-24.
- Das R, Dhokalia A, Huang XZ, Hammamieh R, Chakraborty N *et al.* (2007). Study of proinflammatory responses induced by *Yersinia pestis* in human monocytes using cDNA arrays. *Genes and immunity* 8, 308-319.

- Davis KJ, Fritz DL, Pitt ML, Welkos SL, Worsham PL *et al.* (1996). Pathology of experimental pneumonic plague produced by fraction 1-positive and fraction 1-negative *Yersinia pestis* in African green monkeys (*Cercopithecus aethiops*). *Arch. Pathol. Lab. Med.* 120, 156-163.
- Dejager L, Libert C, Montagutelli X (2009). Thirty years of *Mus spretus*: a promising future. *Trends in genetics : TIG* 25, 234-241.
- Dejager L, Pinheiro I, Bogaert P, Huys L, Libert C (2010a). Role for neutrophils in host immune responses and genetic factors that modulate resistance to *Salmonella enterica* serovar typhimurium in the inbred mouse strain SPRET/Ei. *Infection and immunity* 78, 3848-3860.
- Dejager L, Pinheiro I, Puimege L, Fan YD, Gremeaux L *et al.* (2010b). Increased glucocorticoid receptor expression and activity mediate the LPS resistance of SPRET/EI mice. *The Journal of biological chemistry* 285, 31073-31086.
- Demant P, Hart AA (1986). Recombinant congenic strains--a new tool for analyzing genetic traits determined by more than one gene. *Immunogenetics* 24, 416-422.
- Demeure CE, Blanchet C, Fitting C, Fayolle C, Khun H *et al.* (2012). Early systemic bacterial dissemination and a rapid innate immune response characterize genetic resistance to plague of SEG mice. *The Journal of infectious diseases* 205, 134-143.
- Demeure CE, Carniel E (2009). *Yersinia pestis*. *Revue Francophone des laboratoires* 415, 33-39.
- Depaolo RW, Tang F, Kim I, Han M, Levin N *et al.* (2008). Toll-like receptor 6 drives differentiation of tolerogenic dendritic cells and contributes to LcrV-mediated plague pathogenesis. *Cell host & microbe* 4, 350-361.
- Detweiler CS, Cunanan DB, Falkow S (2001). Host microarray analysis reveals a role for the *Salmonella* response regulator phoP in human macrophage cell death. *Proceedings of the National Academy of Sciences of the United States of America* 98, 5850-5855.
- Diez E, Zhu L, Teatero SA, Paquet M, Roy MF *et al.* (2011). Identification and characterization of Cri1, a locus controlling mortality during *Citrobacter rodentium* infection in mice. *Genes and immunity* 12, 280-290.
- DiPetrillo K, Wang X, Stylianou IM, Paigen B (2005). Bioinformatics toolbox for narrowing rodent quantitative trait loci. *Trends in genetics : TIG* 21, 683-692.
- Dixon AL, Liang L, Moffatt MF, Chen W, Heath S *et al.* (2007). A genome-wide association study of global gene expression. *Nature genetics* 39, 1202-1207.
- do Valle TZ, Billecocq A, Guillemot L, Alberts R, Gomet C *et al.* (2010). A new mouse model reveals a critical role for host innate immunity in resistance to Rift Valley fever. *Journal of immunology (Baltimore, Md. : 1950)* 185, 6146-6156.
- Doss S, Schadt EE, Drake TA, Lusis AJ (2005). Cis-acting expression quantitative trait loci in mice. *Genome research* 15, 681-691.
- Drozdzov IG, Anisimov AP, Samoilova SV, Yezhov IN, Yereimin SA *et al.* (1995). Virulent non-capsulate *Yersinia pestis* variants constructed by insertion mutagenesis. *J. Med. Microbiol.* 42, 264-268.
- Du YD, Rosqvist R, Forsberg A (2002). Role of fraction 1 antigen of *Yersinia pestis* in inhibition of phagocytosis. *Infect. Immun.* 70, 1453-1460.
- Dumas A, Pouliot M (2009). [Neutrophil: foe or friend?]. *Medecine sciences : M/S* 25, 699-704.
- Durrant C, Tayem H, Yalcin B, Cleak J, Goodstadt L *et al.* (2011). Collaborative Cross mice and their power to map host susceptibility to *Aspergillus fumigatus* infection. *Genome research* 21, 1239-1248.
- Dyer MD, Neff C, Dufford M, Rivera CG, Shattuck D *et al.* (2010). The human-bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. *PloS one* 5, e12089.

- Dziarski R (2006). Deadly plague versus mild-mannered TLR4. *Nature immunology* 7, 1017-1019.
- Dziarski R, Gupta D (2010). Review: Mammalian peptidoglycan recognition proteins (PGRPs) in innate immunity. *Innate immunity* 16, 168-174.
- Eichler EE, Flint J, Gibson G, Kong A, Leal SM *et al.* (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nature reviews. Genetics* 11, 446-450.
- Eisele NA, Lee-Lewis H, Besch-Williford C, Brown CR, Anderson DM (2011). Chemokine receptor CXCR2 mediates bacterial clearance rather than neutrophil recruitment in a murine model of pneumonic plague. *The American journal of pathology* 178, 1190-1200.
- Eisen RJ, Petersen JM, Higgins CL, Wong D, Levy CE *et al.* (2008). Persistence of *Yersinia pestis* in soil under natural conditions. *Emerging infectious diseases* 14, 941-943.
- Falcon S, Gentleman R (2007). Using GOstats to test gene lists for GO term association. *Bioinformatics (Oxford, England)* 23, 257-258.
- Falconer D, Mackay T (1996). *Introduction to Quantitative Genetics*, 4th ed. London
- Farrall M (2004). Quantitative genetic variation: a post-modern view. *Human molecular genetics* 13 Spec No 1, R1-7.
- Feingold J (1998). [Determining the genetic component of common diseases]. *Pathologie-biologie* 46, 701-704.
- Fernandes-Alnemri T, Wu J, Yu JW, Datta P, Miller B *et al.* (2007). The pyroptosome: a supramolecular assembly of ASC dimers mediating inflammatory cell death via caspase-1 activation. *Cell death and differentiation* 14, 1590-1604.
- Ferwerda G, Girardin SE, Kullberg BJ, Le Bourhis L, de Jong DJ *et al.* (2005). NOD2 and toll-like receptors are nonredundant recognition systems of *Mycobacterium tuberculosis*. *PLoS pathogens* 1, 279-285.
- Fieschi C, Casanova JL (2003). The role of interleukin-12 in human infectious diseases: only a faint signature. *European journal of immunology* 33, 1461-1464.
- Fisher P, Hedeler C, Wolstencroft K, Hulme H, Noyes H *et al.* (2007). A systematic strategy for large-scale analysis of genotype phenotype correlations: identification of candidate genes involved in African trypanosomiasis. *Nucleic acids research* 35, 5625-5633.
- Forman HJ, Torres M (2002). Reactive oxygen species and cell signaling: respiratory burst in macrophage signaling. *American journal of respiratory and critical care medicine* 166, S4-8.
- Fortin A, Abel L, Casanova JL, Gros P (2007). Host genetics of mycobacterial diseases in mice and men: forward genetic studies of BCG-osis and tuberculosis. *Annual review of genomics and human genetics* 8, 163-192.
- Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA *et al.* (2007). A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* 448, 1050-1053.
- Fritz JH, Girardin SE, Fitting C, Werts C, Mengin-Lecreulx D *et al.* (2005). Synergistic stimulation of human monocytes and dendritic cells by Toll-like receptor 4 and NOD1- and NOD2-activating agonists. *European journal of immunology* 35, 2459-2470.
- Fuchs TA, Abed U, Goosmann C, Hurwitz R, Schulze I *et al.* (2007). Novel cell death program leads to neutrophil extracellular traps. *The Journal of cell biology* 176, 231-241.
- Galimand M, Carniel E, Courvalin P (2006). Resistance of *Yersinia pestis* to antimicrobial agents. *Antimicrobial agents and chemotherapy* 50, 3233-3236.
- Galimand M, Guiyoule A, Gerbaud G, Rasoamanana B, Chanteau S *et al.* (1997). Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid. *The New England journal of medicine* 337, 677-680.
- Galindo CL, Moen ST, Kozlova EV, Sha J, Garner HR *et al.* (2009). Comparative Analyses of Transcriptional Profiles in Mouse Organs Using a Pneumonic Plague Model after Infection with Wild-

Type *Yersinia pestis* CO92 and Its Braun Lipoprotein Mutant. Comparative and functional genomics 2009, 914762.

Galvan EM, Lasaro MAS, Schifferli DM (2008). Capsular antigen fraction 1 and pla modulate the susceptibility of *Yersinia pestis* to pulmonary antimicrobial peptides such as cathelicidin. Infection and Immunity 76, 1456-1464.

Gantt KR, Goldman TL, McCormick ML, Miller MA, Jeronimo SM *et al.* (2001). Oxidative responses of human and murine macrophages during phagocytosis of *Leishmania chagasi*. Journal of immunology (Baltimore, Md. : 1950) 167, 893-901.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome biology 5, R80.

Girard G (1955). Plague. Annual review of microbiology 9, 253-276.

Girardin SE, Tournebize R, Mavris M, Page AL, Li X *et al.* (2001). CARD4/Nod1 mediates NF-kappaB and JNK activation by invasive *Shigella flexneri*. EMBO reports 2, 736-742.

Goetz M, Bubert A, Wang G, Chico-Calero I, Vazquez-Boland JA *et al.* (2001). Microinjection and growth of bacteria in the cytosol of mammalian host cells. Proceedings of the National Academy of Sciences of the United States of America 98, 12221-12226.

Goldmann O, Lengeling A, Bose J, Bloecker H, Geffers R *et al.* (2005). The role of the MHC on resistance to group A streptococci in mice. Journal of immunology (Baltimore, Md. : 1950) 175, 3862-3872.

Goldmann O, von Kockritz-Blickwede M, Holtje C, Chhatwal GS, Geffers R *et al.* (2007). Transcriptome analysis of murine macrophages in response to infection with *Streptococcus pyogenes* reveals an unusual activation program. Infection and immunity 75, 4148-4157.

Goncalves LA, Almeida P, Mota MM, Penha-Goncalves C (2008). Malaria liver stage susceptibility locus identified on mouse chromosome 17 by congenic mapping. PloS one 3, e1874.

Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R *et al.* (2005). The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science (New York, N.Y.) 307, 1434-1440.

Gonzalez RJ, Weening EH, Frothingham R, Sempowski GD, Miller VL (2012). Bioluminescence imaging to track bacterial dissemination of *Yersinia pestis* using different routes of infection in mice. BMC microbiology 12, 147.

Gordon AH, Hart PD, Young MR (1980). Ammonia inhibits phagosome-lysosome fusion in macrophages. Nature 286, 79-80.

Gordon S, Taylor PR (2005). Monocyte and macrophage heterogeneity. Nature reviews. Immunology 5, 953-964.

Grassl GA, Kracht M, Wiedemann A, Hoffmann E, Aepfelbacher M *et al.* (2003). Activation of NF-kappaB and IL-8 by *Yersinia enterocolitica* invasin protein is conferred by engagement of Rac1 and MAP kinase cascades. Cellular microbiology 5, 957-971.

Grattan M, Mi QS, Meagher C, Delovitch TL (2002). Congenic mapping of the diabetogenic locus Idd4 to a 5.2-cM region of chromosome 11 in NOD mice: identification of two potential candidate subloci. Diabetes 51, 215-223.

Greenwood CM, Fujiwara TM, Boothroyd LJ, Miller MA, Frappier D *et al.* (2000). Linkage of tuberculosis to chromosome 2q35 loci, including NRAMP1, in a large aboriginal Canadian family. American journal of human genetics 67, 405-416.

Gregorova S, Divina P, Storchova R, Trachtulec Z, Fotopulosova V *et al.* (2008). Mouse consomic strains: exploiting genetic divergence between *Mus m. musculus* and *Mus m. domesticus* subspecies. Genome research 18, 509-515.

Grupe A, Germer S, Usuka J, Aud D, Belknap JK *et al.* (2001). In silico mapping of complex disease-related traits in mice. Science (New York, N.Y.) 292, 1915-1918.

Guenet JL (2011). Animal models of human genetic diseases: do they need to be faithful to be useful? *Molecular genetics and genomics* : MGG 286, 1-20.

Guenet JL, Bonhomme F (2003). Wild mice: an ever-increasing contribution to a popular mammalian model. *Trends in genetics* : TIG 19, 24-31.

GuhaThakurta D (2006 A). Computational identification of transcriptional regulatory elements in DNA sequence. *Nucleic acids research* 34, 3585-3598.

GuhaThakurta D, Xie T, Anand M, Edwards SW, Li G *et al.* (2006 B). Cis-regulatory variations: a study of SNPs around genes showing cis-linkage in segregating mouse populations. *BMC genomics* 7, 235.

Guinet F, Ave P, Jones L, Huerre M, Carniel E (2008). Defective innate cell response and lymph node infiltration specify *Yersinia pestis* infection. *PloS one* 3, e1688.

Guiyoule A, Gerbaud G, Buchrieser C, Galimand M, Rahalison L *et al.* (2001). Transferable plasmid-mediated resistance to streptomycin in a clinical isolate of *Yersinia pestis*. *Emerging infectious diseases* 7, 43-48.

Hambleton S, Salem S, Bustamante J, Bigley V, Boisson-Dupuis S *et al.* (2011). IRF8 mutations and human dendritic-cell immunodeficiency. *The New England journal of medicine* 365, 127-138.

Hansson GK, Edfeldt K (2005). Toll to be paid at the gateway to the vessel wall. *Arteriosclerosis, thrombosis, and vascular biology* 25, 1085-1087.

Hardy MP, O'Neill LA (2004). The murine IRAK2 gene encodes four alternatively spliced isoforms, two of which are inhibitory. *The Journal of biological chemistry* 279, 27699-27708.

Hawn TR, Misch EA, Dunstan SJ, Thwaites GE, Lan NT *et al.* (2007). A common human TLR1 polymorphism regulates the innate immune response to lipopeptides. *European journal of immunology* 37, 2280-2289.

Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC *et al.* (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 410, 1099-1103.

Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C *et al.* (2004). Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science (New York, N.Y.)* 303, 1526-1529.

Higgins JA, Ezzell J, Hinnebusch BJ, Shipley M, Henchal EA *et al.* (1998). 5' nuclease PCR assay to detect *Yersinia pestis*. *Journal of clinical microbiology* 36, 2284-2288.

Hill AV (1996). Genetics of infectious disease resistance. *Current opinion in genetics & development* 6, 348-353.

Hill AV (2001). The genomics and genetics of human infectious disease susceptibility. *Annual review of genomics and human genetics* 2, 373-400.

Hill AV (2006). Aspects of genetic susceptibility to human infectious diseases. *Annual review of genetics* 40, 469-486.

Hill AV (2012). Evolution, revolution and heresy in the genetics of infectious disease susceptibility. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 367, 840-849.

Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, Twumasi P *et al.* (1991). Common west African HLA antigens are associated with protection from severe malaria. *Nature* 352, 595-600.

Hinnebusch BJ, Rosso ML, Schwan TG, Carniel E (2002 A). High-frequency conjugative transfer of antibiotic resistance genes to *Yersinia pestis* in the flea midgut. *Molecular microbiology* 46, 349-354.

Hinnebusch BJ, Rudolph AE, Cherepanov P, Dixon JE, Schwan TG *et al.* (2002 B). Role of *Yersinia* murine toxin in survival of *Yersinia pestis* in the midgut of the flea vector. *Science (New York, N.Y.)* 296, 733-735.

- Hinnebusch J, Schwan TG (1993). New method for plague surveillance using polymerase chain reaction to detect *Yersinia pestis* in fleas. *Journal of clinical microbiology* 31, 1511-1514.
- Hisamatsu T, Suzuki M, Reinecker HC, Nadeau WJ, McCormick BA *et al.* (2003). CARD15/NOD2 functions as an antibacterial factor in human intestinal epithelial cells. *Gastroenterology* 124, 993-1000.
- Hoebe K, Jiang Z, Tabeta K, Du X, Georgel P *et al.* (2006). Genetic analysis of innate immunity. *Advances in immunology* 91, 175-226.
- Hoffmann R, van Erp K, Trulzsch K, Heesemann J (2004). Transcriptional responses of murine macrophages to infection with *Yersinia enterocolitica*. *Cellular microbiology* 6, 377-390.
- Hossain H, Tchatalbachev S, Chakraborty T (2006). Host gene expression profiling in pathogen-host interactions. *Current opinion in immunology* 18, 422-429.
- Hou J, Wang P, Lin L, Liu X, Ma F *et al.* (2009). MicroRNA-146a feedback inhibits RIG-I-dependent Type I IFN production in macrophages by targeting TRAF6, IRAK1, and IRAK2. *Journal of immunology* (Baltimore, Md. : 1950) 183, 2150-2158.
- Hu J, Bumstead N, Barrow P, Sebastiani G, Olien L *et al.* (1997). Resistance to salmonellosis in the chicken is linked to NRAMP1 and TNC. *Genome research* 7, 693-704.
- Huang da W, Sherman BT, Lempicki RA (2009 A). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 4, 44-57.
- Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T *et al.* (2009 B). Extracting biological meaning from large gene lists with DAVID. *Current protocols in bioinformatics / editorial board, Andreas D. Baxeavanis ... [et al.] Chapter 13, Unit 13 11.*
- Huang Q, Liu D, Majewski P, Schulte LC, Korn JM *et al.* (2001). The plasticity of dendritic cell responses to pathogens and their components. *Science (New York, N.Y.)* 294, 870-875.
- Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H *et al.* (2005). Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nature genetics* 37, 243-253.
- Hunter KW, Crawford NP (2008). The future of mouse QTL mapping to diagnose disease in mice in the age of whole-genome association studies. *Annual review of genetics* 42, 131-141.
- Inglesby TV, Grossman R, O'Toole T (2001). A plague on your city: observations from TOPOFF. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 32, 436-445.
- Jenner RG, Young RA (2005). Insights into host responses against pathogens from transcriptional profiling. *Nature reviews. Microbiology* 3, 281-294.
- Jo EK (2008). Mycobacterial interaction with innate receptors: TLRs, C-type lectins, and NLRs. *Current opinion in infectious diseases* 21, 279-286.
- Johnson CM, Lyle EA, Omuetti KO, Stepensky VA, Yegin O *et al.* (2007). Cutting edge: A common polymorphism impairs cell surface trafficking and functional responses of TLR1 but protects against leprosy. *Journal of immunology* (Baltimore, Md. : 1950) 178, 7520-7524.
- Johnson NV, Ahn SH, Deshmukh H, Levin MK, Nelson CL *et al.* (2012). Haplotype Association Mapping Identifies a Candidate Gene Region in Mice Infected With *Staphylococcus aureus*. *G3 (Bethesda, Md.)* 2, 693-700.
- Jordan B (2009). [The decline of genome-wide association studies]. *Medecine sciences : M/S* 25, 537-539.
- Justice MJ, Noveroske JK, Weber JS, Zheng B, Bradley A (1999). Mouse ENU mutagenesis. *Human molecular genetics* 8, 1955-1963.
- Kanehisa M, Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28, 27-30.

- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research* 40, D109-114.
- Kanneganti TD, Lamkanfi M, Nunez G (2007). Intracellular NOD-like receptors in host defense and disease. *Immunity* 27, 549-559.
- Kapetanovic R, Cavaillon JM (2007). Early events in innate immunity in the recognition of microbial pathogens. *Expert opinion on biological therapy* 7, 907-918.
- Kawahara K, Tsukano H, Watanabe H, Lindner B, Matsuura M (2002). Modification of the structure and activity of lipid A in *Yersinia pestis* lipopolysaccharide by growth temperature. *Infection and immunity* 70, 4092-4098.
- Kawai T, Akira S (2007). TLR signaling. *Seminars in immunology* 19, 24-32.
- Kawai T, Akira S (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature immunology* 11, 373-384.
- Kearsey M (1998). The principles of QTL analysis (a minimal mathematics approach). *Journal of Experimental Botany* 49, 1619-1623.
- Kemp SJ, Iraqi F, Darvasi A, Soller M, Teale AJ (1997). Localization of genes controlling resistance to trypanosomiasis in mice. *Nature genetics* 16, 194-196.
- Kerschen EJ, Cohen DA, Kaplan AM, Straley SC (2004a). The plague virulence protein YopM targets the innate immune response by causing a global depletion of NK cells. *Infection and Immunity* 72, 4589-4602.
- Kerschen EJ, Cohen DA, Kaplan AM, Straley SC (2004b). The plague virulence protein YopM targets the innate immune response by causing a global depletion of NK cells. *Infection and immunity* 72, 4589-4602.
- Khan R, Sancho-Shimizu V, Prendergast C, Roy MF, Loredó-Osti JC *et al.* (2012). Refinement of the genetics of the host response to *Salmonella* infection in MOLF/Ei: regulation of type 1 IFN and TRP3 pathways by Ity2. *Genes and immunity* 13, 175-183.
- Klir JJ, Roth J, Szelenyi Z, McClellan JL, Kluger MJ (1993). Role of hypothalamic interleukin-6 and tumor necrosis factor- α in LPS fever in rat. *The American journal of physiology* 265, R512-517.
- Kolodziejek AM, Hovde CJ, Minnich SA (2012). *Yersinia pestis* Ail: multiple roles of a single protein. *Frontiers in cellular and infection microbiology* 2, 103.
- Korstanje R, Paigen B (2002). From QTL to gene: the harvest begins. *Nature genetics* 31, 235-236.
- Kramnik I, Dietrich WF, Demant P, Bloom BR (2000). Genetic control of resistance to experimental infection with virulent *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* 97, 8560-8565.
- Krausgruber T, Blazek K, Smallie T, Alzabin S, Lockstone H *et al.* (2011). IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses. *Nature immunology* 12, 231-238.
- Krutzik SR, Ochoa MT, Sieling PA, Uematsu S, Ng YW *et al.* (2003). Activation and regulation of Toll-like receptors 2 and 1 in human leprosy. *Nature medicine* 9, 525-532.
- Kurt-Jones EA, Popova L, Kwinn L, Haynes LM, Jones LP *et al.* (2000). Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nature immunology* 1, 398-401.
- L'Hôte D, Laissue P, Serres C, Montagutelli X, Veitia RA *et al.* (2010). Interspecific resources: a major tool for quantitative trait locus cloning and speciation research. *BioEssays : news and reviews in molecular, cellular and developmental biology* 32, 132-142.
- Lam-Yuk-Tseung S, Gros P (2003). Genetic control of susceptibility to bacterial infections in mouse models. *Cellular microbiology* 5, 299-313.

- Lathem WW, Crosby SD, Miller VL, Goldman WE (2005). Progression of primary pneumonic plague: a mouse model of infection, pathology, and bacterial transcriptional activity. *Proceedings of the National Academy of Sciences of the United States of America* 102, 17786-17791.
- Lathem WW, Price PA, Miller VL, Goldman WE (2007). A plasminogen-activating protease specifically controls the development of primary pneumonic plague. *Science* 315, 509-513.
- Laws TR, Davey MS, Titball RW, Lukaszewski R (2010). Neutrophils are important in early control of lung infection by *Yersinia pestis*. *Microbes and infection / Institut Pasteur* 12, 331-335.
- Legare ME, Bartlett FS, 2nd, Frankel WN (2000). A major effect QTL determined by multiple genes in epileptic EL mice. *Genome research* 10, 42-48.
- Leonard TA, Hurley JH (2011). Regulation of protein kinases by lipids. *Current opinion in structural biology* 21, 785-791.
- Levillayer F, Mas M, Levi-Acobas F, Brahic M, Bureau JF (2007). Interleukin 22 is a candidate gene for Tmevp3, a locus controlling Theiler's virus-induced neurological diseases. *Genetics* 176, 1835-1844.
- Levison ME (2000 B). Lessons Learned from History on Mode of Transmission for Control of Pneumonic Plague. *Current infectious disease reports* 2, 269-271.
- Li B, Yang R (2008). Interaction between *Yersinia pestis* and the host immune system. *Infection and immunity* 76, 1804-1811.
- Li HT, Zhang TT, Zhou YQ, Huang QH, Huang J (2006). SLC11A1 (formerly NRAMP1) gene polymorphisms and tuberculosis susceptibility: a meta-analysis. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 10, 3-12.
- Lim JK, Lisco A, McDermott DH, Huynh L, Ward JM *et al.* (2009). Genetic variation in OAS1 is a risk factor for initial infection with West Nile virus in man. *PLoS pathogens* 5, e1000321.
- Lipoldova M, Svobodova M, Krulova M, Havelkova H, Badalova J *et al.* (2000). Susceptibility to *Leishmania major* infection in mice: multiple loci and heterogeneity of immunopathological phenotypes. *Genes and immunity* 1, 200-206.
- Lipton HL, Melvold R, Miller SD, Dal Canto MC (1995). Mutation of a major histocompatibility class I locus, H-2D, leads to an increased virus burden and disease susceptibility in Theiler's virus-induced demyelinating disease. *Journal of neurovirology* 1, 138-144.
- Liu H, Wang H, Qiu J, Wang X, Guo Z *et al.* (2009). Transcriptional profiling of a mice plague model: insights into interaction between *Yersinia pestis* and its host. *Journal of basic microbiology* 49, 92-99.
- Liu P, Vikis H, Lu Y, Wang D, You M (2007). Large-scale in silico mapping of complex quantitative traits in inbred mice. *PloS one* 2, e651.
- Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)* 25, 402-408.
- Loiez C, Herwegh S, Wallet F, Armand S, Guinet F *et al.* (2003). Detection of *Yersinia pestis* in sputum by real-time PCR. *Journal of clinical microbiology* 41, 4873-4875.
- Lovegrove FE, Pena-Castillo L, Mohammad N, Liles WC, Hughes TR *et al.* (2006). Simultaneous host and parasite expression profiling identifies tissue-specific transcriptional programs associated with susceptibility or resistance to experimental cerebral malaria. *BMC genomics* 7, 295.
- Lukacs GL, Rotstein OD, Grinstein S (1991). Determinants of the phagosomal pH in macrophages. In situ assessment of vacuolar H(+)-ATPase activity, counterion conductance, and H+ "leak". *The Journal of biological chemistry* 266, 24540-24548.
- Lukaszewski RA, Kenny DJ, Taylor R, Rees DG, Hartley MG *et al.* (2005). Pathogenesis of *Yersinia pestis* infection in BALB/c mice: effects on host macrophages and neutrophils. *Infection and immunity* 73, 7142-7150.

- Lynch M, Walsh B (1998). *Genetics and Analysis of Quantitative Traits*. MA: Sinauer.
- Mackay TF (2001 A). The genetic architecture of quantitative traits. *Annual review of genetics* 35, 303-339.
- MacLennan C, Fieschi C, Lammas DA, Picard C, Dorman SE *et al.* (2004). Interleukin (IL)-12 and IL-23 are key cytokines for immunity against *Salmonella* in humans. *The Journal of infectious diseases* 190, 1755-1757.
- Mahler KL, Fleming JL, Dworkin AM, Gladman N, Cho HY *et al.* (2008). Sequence divergence of *Mus spretus* and *Mus musculus* across a skin cancer susceptibility locus. *BMC genomics* 9, 626.
- Malik S, Abel L, Tooker H, Poon A, Simkin L *et al.* (2005). Alleles of the NRAMP1 gene are risk factors for pediatric tuberculosis disease. *Proceedings of the National Academy of Sciences of the United States of America* 102, 12183-12188.
- Malissen B, Ewbank JJ (2005). 'TaiLoRing' the response of dendritic cells to pathogens. *Nature immunology* 6, 749-750.
- Malo D, Vogan K, Vidal S, Hu J, Cellier M *et al.* (1994). Haplotype mapping and sequence analysis of the mouse Nramp gene predict susceptibility to infection with intracellular parasites. *Genomics* 23, 51-61.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA *et al.* (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747-753.
- Marathe R, Guan Z, Anandalakshmi R, Zhao H, Dinesh-Kumar SP (2004). Study of *Arabidopsis thaliana* resistome in response to cucumber mosaic virus infection using whole genome microarray. *Plant molecular biology* 55, 501-520.
- Marketon MM, DePaolo RW, DeBord KL, Jabri B, Schneewind O (2005). Plague bacteria target immune cells during infection. *Science (New York, N.Y.)* 309, 1739-1741.
- Marquet S, Abel L, Hillaire D, Dessein H, Kalil J *et al.* (1996). Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Nature genetics* 14, 181-184.
- Marquis JF, Lacourse R, Ryan L, North RJ, Gros P (2009). Genetic and functional characterization of the mouse Trl3 locus in defense against tuberculosis. *Journal of immunology (Baltimore, Md. : 1950)* 182, 3757-3767.
- Mashimo T, Lucas M, Simon-Chazottes D, Frenkiel MP, Montagutelli X *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. *Proceedings of the National Academy of Sciences of the United States of America* 99, 11311-11316.
- Mastroeni P, Vazquez-Torres A, Fang FC, Xu Y, Khan S *et al.* (2000). Antimicrobial actions of the NADPH phagocyte oxidase and inducible nitric oxide synthase in experimental salmonellosis. II. Effects on microbial proliferation and host survival in vivo. *The Journal of experimental medicine* 192, 237-248.
- Matsuura M, Takahashi H, Watanabe H, Saito S, Kawahara K (2010). Immunomodulatory effects of *Yersinia pestis* lipopolysaccharides on human macrophages. *Clinical and vaccine immunology : CVI* 17, 49-55.
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J *et al.* (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature reviews. Genetics* 9, 356-369.
- McDonald C, Vacratsis PO, Bliska JB, Dixon JE (2003). The *Yersinia* virulence factor YopM forms a novel protein complex with two cellular kinases. *Journal of Biological Chemistry* 278, 18514-18523.
- Meyer KF (1970). Effectiveness of live or killed plague vaccines in man. *Bulletin of the World Health Organization* 42, 653-666.

- Miller SI, Ernst RK, Bader MW (2005). LPS, TLR4 and infectious disease diversity. *Nature reviews. Microbiology* 3, 36-46.
- Min-Oo G, Lindqvist L, Vaglenov A, Wang C, Fortin P *et al.* (2008). Genetic control of susceptibility to pulmonary infection with *Chlamydia pneumoniae* in the mouse. *Genes and immunity* 9, 383-388.
- Mira MT, Alcais A, Nguyen VT, Moraes MO, Di Flumeri C *et al.* (2004). Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature* 427, 636-640.
- Mira MT, Alcais A, Van Thuc N, Thai VH, Huong NT *et al.* (2003). Chromosome 6q25 is linked to susceptibility to leprosy in a Vietnamese population. *Nature genetics* 33, 412-415.
- Misch EA, Berrington WR, Vary JC, Jr., Hawn TR (2010). Leprosy and the human genome. *Microbiology and molecular biology reviews* : MMBR 74, 589-620.
- Misch EA, Macdonald M, Ranjit C, Sapkota BR, Wells RD *et al.* (2008). Human TLR1 deficiency is associated with impaired mycobacterial signaling and protection from leprosy reversal reaction. *PLoS neglected tropical diseases* 2, e231.
- Mitsos LM, Cardon LR, Fortin A, Ryan L, LaCourse R *et al.* (2000). Genetic control of susceptibility to infection with *Mycobacterium tuberculosis* in mice. *Genes and immunity* 1, 467-477.
- Mitsos LM, Cardon LR, Ryan L, LaCourse R, North RJ *et al.* (2003). Susceptibility to tuberculosis: a locus on mouse chromosome 19 (Tr1-4) regulates *Mycobacterium tuberculosis* replication in the lungs. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6610-6615.
- Miyairi I, Ziebarth J, Laxton JD, Wang X, van Rooijen N *et al.* (2012). Host genetics and *Chlamydia* disease: prediction and validation of disease severity mechanisms. *PloS one* 7, e33781.
- Mogensen TH (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical microbiology reviews* 22, 240-273, Table of Contents.
- Mohan C, Alas E, Morel L, Yang P, Wakeland EK (1998). Genetic dissection of SLE pathogenesis. Sle1 on murine chromosome 1 leads to a selective loss of tolerance to H2A/H2B/DNA subnucleosomes. *The Journal of clinical investigation* 101, 1362-1372.
- Mohan C, Yu Y, Morel L, Yang P, Wakeland EK (1999). Genetic dissection of Sle pathogenesis: Sle3 on murine chromosome 7 impacts T cell activation, differentiation, and cell death. *Journal of immunology* (Baltimore, Md. : 1950) 162, 6492-6502.
- Mollaret HH, Karimi Y, Eftekhari M, Baltazard M (1963). [BURROWING PLAGUE]. *Bulletin de la Societe de pathologie exotique et de ses filiales* 56, 1186-1193.
- Montagutelli X, Abitbol M (2004). [Applications of congenic strains in the mouse]. *Medecine sciences : M/S* 20, 887-893.
- Montminy SW, Khan N, McGrath S, Walkowicz MJ, Sharp F *et al.* (2006). Virulence factors of *Yersinia pestis* are overcome by a strong lipopolysaccharide response. *Nature immunology* 7, 1066-1073.
- Morel L, Blenman KR, Croker BP, Wakeland EK (2001). The major murine systemic lupus erythematosus susceptibility locus, Sle1, is a cluster of functionally related genes. *Proceedings of the National Academy of Sciences of the United States of America* 98, 1787-1792.
- Morel L, Croker BP, Blenman KR, Mohan C, Huang G *et al.* (2000a). Genetic reconstitution of systemic lupus erythematosus immunopathology with polycongenic murine strains. *Proc Natl Acad Sci U S A* 97, 6670-6675.
- Morel L, Mohan C, Yu Y, Croker BP, Tian N *et al.* (1997). Functional dissection of systemic lupus erythematosus using congenic mouse strains. *Journal of immunology* (Baltimore, Md. : 1950) 158, 6019-6028.
- Morel L, Rudofsky UH, Longmate JA, Schiffenbauer J, Wakeland EK (1994). Polygenic control of susceptibility to murine systemic lupus erythematosus. *Immunity* 1, 219-229.

- Morel L, Wakeland EK (2000b). Lessons from the NZM2410 model and related strains. *International reviews of immunology* 19, 423-446.
- Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG *et al.* (2004). Genetic analysis of genome-wide variation in human gene expression. *Nature* 430, 743-747.
- Mosser DM, Edwards JP (2008). Exploring the full spectrum of macrophage activation. *Nature reviews. Immunology* 8, 958-969.
- Motley ST, Morrow BJ, Liu X, Dodge IL, Vitiello A *et al.* (2004). Simultaneous analysis of host and pathogen interactions during an in vivo infection reveals local induction of host acute phase response proteins, a novel bacterial stress response, and evidence of a host-imposed metal ion limited environment. *Cellular microbiology* 6, 849-865.
- Motsch N, Pfuhl T, Mrazek J, Barth S, Grasser FA (2007). Epstein-Barr virus-encoded latent membrane protein 1 (LMP1) induces the expression of the cellular microRNA miR-146a. *RNA biology* 4, 131-137.
- Mott R, Flint J (2002). Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks. *Genetics* 160, 1609-1618.
- Mott R, Flint J (2008). Prospects for complex trait analysis in the mouse. *Mammalian genome : official journal of the International Mammalian Genome Society* 19, 306-308.
- Mukherjee S, Keitany G, Li Y, Wang Y, Ball HL *et al.* (2006). Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science (New York, N.Y.)* 312, 1211-1214.
- Nadeau JH (2005). Listening to genetic background noise. *The New England journal of medicine* 352, 1598-1599.
- Nadeau JH, Singer JB, Matin A, Lander ES (2000). Analysing complex genetic traits with chromosome substitution strains. *Nature genetics* 24, 221-225.
- Nau GJ, Richmond JF, Schlesinger A, Jennings EG, Lander ES *et al.* (2002). Human macrophage activation programs induced by bacterial pathogens. *Proceedings of the National Academy of Sciences of the United States of America* 99, 1503-1508.
- Nau GJ, Schlesinger A, Richmond JF, Young RA (2003). Cumulative Toll-like receptor activation in human macrophages treated with whole bacteria. *Journal of immunology (Baltimore, Md. : 1950)* 170, 5203-5209.
- Newton K, Dixit VM (2012). Signaling in innate immunity and inflammation. *Cold Spring Harbor perspectives in biology* 4
- Ng LC, Forslund O, Koh S, Kuoppa K, Sjostedt A (2003). The response of murine macrophages to infection with Yersinia pestis as revealed by DNA microarray analysis. *Advances in experimental medicine and biology* 529, 155-160.
- Nham T, Filali S, Danne C, Derbise A, Carniel E (2012). Imaging of bubonic plague dynamics by in vivo tracking of bioluminescent Yersinia pestis. *PloS one* 7, e34714.
- Nilsen NJ, Deininger S, Nonstad U, Skjeldal F, Husebye H *et al.* (2008). Cellular trafficking of lipoteichoic acid and Toll-like receptor 2 in relation to signaling: role of CD14 and CD36. *Journal of leukocyte biology* 84, 280-291.
- O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D (2010). Physiological and pathological roles for microRNAs in the immune system. *Nature reviews. Immunology* 10, 111-122.
- O'Neill LA, Sheedy FJ, McCoy CE (2011). MicroRNAs: the fine-tuners of Toll-like receptor signalling. *Nature reviews. Immunology* 11, 163-175.
- Opitz B, Puschel A, Beermann W, Hocke AC, Forster S *et al.* (2006). Listeria monocytogenes activated p38 MAPK and induced IL-8 secretion in a nucleotide-binding oligomerization domain 1-dependent manner in endothelial cells. *Journal of immunology (Baltimore, Md. : 1950)* 176, 484-490.

- Opitz B, Puschel A, Schmeck B, Hocke AC, Rosseau S *et al.* (2004). Nucleotide-binding oligomerization domain proteins are innate immune receptors for internalized *Streptococcus pneumoniae*. *The Journal of biological chemistry* 279, 36426-36432.
- Oshiumi H, Sasai M, Shida K, Fujita T, Matsumoto M *et al.* (2003). TIR-containing adapter molecule (TICAM)-2, a bridging adapter recruiting to toll-like receptor 4 TICAM-1 that induces interferon-beta. *The Journal of biological chemistry* 278, 49751-49762.
- Pan H, Yan BS, Rojas M, Shebzukhov YV, Zhou H *et al.* (2005). Ipr1 gene mediates innate immunity to tuberculosis. *Nature* 434, 767-772.
- Panthier JJ, Montagutelli X (2012). [The Collaborative Cross, a groundbreaking tool to tackle complex traits]. *Medecine sciences : M/S* 28, 103-108.
- Panthier JJ, Montagutelli X, Guénet JL (2003). *Génétique de la souris. Les organismes modèles*, Belin Biologie ed
- Paquette N, Conlon J, Sweet C, Rus F, Wilson L *et al.* (2012). Serine/threonine acetylation of TGFbeta-activated kinase (TAK1) by *Yersinia pestis* YopJ inhibits innate immune signaling. *Proceedings of the National Academy of Sciences of the United States of America* 109, 12710-12715.
- Patel AA, Lee-Lewis H, Hughes-Hanks J, Lewis CA, Anderson DM (2012). Opposing Roles for Interferon Regulatory Factor-3 (IRF-3) and Type I Interferon Signaling during Plague. *PLoS pathogens* 8, e1002817.
- Perelygin AA, Scherbik SV, Zhulin IB, Stockman BM, Li Y *et al.* (2002). Positional cloning of the murine flavivirus resistance gene. *Proceedings of the National Academy of Sciences of the United States of America* 99, 9322-9327.
- Perry RD, Fetherston JD (1997). *Yersinia pestis*--etiologic agent of plague. *Clinical microbiology reviews* 10, 35-66.
- Perry RD, Pendrak ML, Schuetze P (1990). Identification and Cloning of a Hemin Storage Locus Involved in the Pigmentation Phenotype of *Yersinia pestis*. *Journal of Bacteriology* 172, 5929-5937.
- Philipovskiy AV, Cowan C, Wulff-Strobel CR, Burnett SH, Kerschen EJ *et al.* (2005). Antibody against V antigen prevents Yop-dependent growth of *Yersinia pestis*. *Infection and immunity* 73, 1532-1542.
- Pletcher MT, McClurg P, Batalov S, Su AI, Barnes SW *et al.* (2004). Use of a dense single nucleotide polymorphism map for in silico mapping in the mouse. *PLoS biology* 2, e393.
- Pollitzer R (1960). A review of recent literature on plague. *Bulletin of the World Health Organization* 23, 313-400.
- Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C *et al.* (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science (New York, N.Y.)* 282, 2085-2088.
- Pommerenke C, Wilk E, Srivastava B, Schulze A, Novoselova N *et al.* (2012). Global transcriptome analysis in influenza-infected mouse lungs reveals the kinetics of innate and adaptive host immune responses. *PloS one* 7, e41169.
- Pomp D, Allan MF, Wesolowski SR (2004). Quantitative genomics: exploring the genetic architecture of complex trait predisposition. *Journal of animal science* 82 E-Suppl, E300-312.
- Ponnusamy D, Clinkenbeard KD (2012). *Yersinia pestis* Intracellular Parasitism of Macrophages from Hosts Exhibiting High and Low Severity of Plague. *PloS one* 7, e42211.
- Pouillot F, Derbise A, Kukkonen M, Foulon J, Korhonen TK *et al.* (2005). Evaluation of O-antigen inactivation on Pla activity and virulence of *Yersinia pseudotuberculosis* harbouring the pPla plasmid. *Microbiology (Reading, England)* 151, 3759-3768.
- Price AH (2006). Believe it or not, QTLs are accurate! *Trends in plant science* 11, 213-216.

Price PA, Jin J, Goldman WE (2012). Pulmonary infection by *Yersinia pestis* rapidly establishes a permissive environment for microbial proliferation. *Proceedings of the National Academy of Sciences of the United States of America* 109, 3083-3088.

Probst CM, Silva RA, Menezes JP, Almeida TF, Gomes IN *et al.* (2012). A comparison of two distinct murine macrophage gene expression profiles in response to *Leishmania amazonensis* infection. *BMC microbiology* 12, 22.

Raberg L, Sim D, Read AF (2007). Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science (New York, N.Y.)* 318, 812-814.

Radovanovic I, Mullick A, Gros P (2011). Genetic control of susceptibility to infection with *Candida albicans* in mice. *PloS one* 6, e18957.

Rahalison L, Vololonirina E, Ratsitorahina M, Chanteau S (2000). Diagnosis of bubonic plague by PCR in Madagascar under field conditions. *Journal of clinical microbiology* 38, 260-263.

Ramalingaswami V (1995). Plague in India. *Nature medicine* 1, 1237-1239.

Ramamurthi KS, Schneewind O (2002). Type iii protein secretion in *Yersinia* species. *Annual review of cell and developmental biology* 18, 107-133.

Rathinam VA, Vanaja SK, Waggoner L, Sokolovska A, Becker C *et al.* (2012). TRIF licenses caspase-11-dependent NLRP3 inflammasome activation by gram-negative bacteria. *Cell* 150, 606-619.

Read AF, Graham AL, Raberg L (2008). Animal defenses against infectious agents: is damage control more important than pathogen control. *PLoS biology* 6, e4.

Richer E, Qureshi ST, Vidal SM, Malo D (2008). Chemical mutagenesis: a new strategy against the global threat of infectious diseases. *Mammalian genome : official journal of the International Mammalian Genome Society* 19, 309-317.

Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D *et al.* (2007). A comparison of background correction methods for two-colour microarrays. *Bioinformatics (Oxford, England)* 23, 2700-2707.

Roberts LJ, Baldwin TM, Curtis JM, Handman E, Foote SJ (1997). Resistance to *Leishmania major* is linked to the H2 region on chromosome 17 and to chromosome 9. *The Journal of experimental medicine* 185, 1705-1710.

Robertson SA, Care AS, Skinner RJ (2007). Interleukin 10 regulates inflammatory cytokine synthesis to protect against lipopolysaccharide-induced abortion and fetal growth restriction in mice. *Biology of reproduction* 76, 738-748.

Rodriguez NE, Chang HK, Wilson ME (2004). Novel program of macrophage gene expression induced by phagocytosis of *Leishmania chagasi*. *Infection and immunity* 72, 2111-2122.

Roff DA (2007). A centennial celebration for quantitative genetics. *Evolution; international journal of organic evolution* 61, 1017-1032.

Rogers JV, Choi YW, Giannunzio LF, Sabourin PJ, Bornman DM *et al.* (2007). Transcriptional responses in spleens from mice exposed to *Yersinia pestis* CO92. *Microbial pathogenesis* 43, 67-77.

Rogner UC, Avner P (2003). Congenic mice: cutting tools for complex immune disorders. *Nature reviews. Immunology* 3, 243-252.

Rose AM, Bell LC (2012). Epistasis and immunity: the role of genetic interactions in autoimmune diseases. *Immunology*

Rosenberger CM, Finlay BB (2003). Phagocyte sabotage: disruption of macrophage signalling by bacterial pathogens. *Nature reviews. Molecular cell biology* 4, 385-396.

Rouzaire P, Mayol K, Viel S, Bienvenu J, Walzer T (2012). [Homeostasis of natural killer cells]. *Medecine sciences : M/S* 28, 403-408.

Roy MF, Malo D (2002). Genetic regulation of host responses to *Salmonella* infection in mice. *Genes and immunity* 3, 381-393.

- Roy MF, Riendeau N, Bedard C, Helie P, Min-Oo G *et al.* (2007). Pyruvate kinase deficiency confers susceptibility to *Salmonella typhimurium* infection in mice. *The Journal of experimental medicine* 204, 2949-2961.
- Roy MF, Riendeau N, Loredó-Osti JC, Malo D (2006). Complexity in the host response to *Salmonella Typhimurium* infection in AcB and BcA recombinant congenic strains. *Genes and immunity* 7, 655-666.
- Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC *et al.* (1979). Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* 76, 5818-5819.
- Samia NI, Kausrud KL, Heesterbeek H, Ageyev V, Begon M *et al.* (2011). Dynamics of the plague-wildlife-human system in Central Asia are controlled by two epidemiological thresholds. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14527-14532.
- Sanchez F, Radaeva TV, Nikonenko BV, Persson AS, Sengul S *et al.* (2003). Multigenic control of disease severity after virulent *Mycobacterium tuberculosis* infection in mice. *Infection and immunity* 71, 126-131.
- Sancho-Shimizu V, Khan R, Mostowy S, Lariviere L, Wilkinson R *et al.* (2007). Molecular genetic analysis of two loci (*Ity2* and *Ity3*) involved in the host response to infection with *Salmonella typhimurium* using congenic mice and expression profiling. *Genetics* 177, 1125-1139.
- Sancho-Shimizu V, Malo D (2006). Sequencing, expression, and functional analyses support the candidacy of *Ncf2* in susceptibility to *Salmonella typhimurium* infection in wild-derived mice. *Journal of immunology (Baltimore, Md. : 1950)* 176, 6954-6961.
- Sato S, Sugiyama M, Yamamoto M, Watanabe Y, Kawai T *et al.* (2003). Toll/IL-1 receptor domain-containing adaptor inducing IFN-beta (TRIF) associates with TNF receptor-associated factor 6 and TANK-binding kinase 1, and activates two distinct transcription factors, NF-kappa B and IFN-regulatory factor-3, in the Toll-like receptor signaling. *Journal of immunology (Baltimore, Md. : 1950)* 171, 4304-4310.
- Sauvonnet N, Pradet-Balade B, Garcia-Sanz JA, Cornelis GR (2002). Regulation of mRNA expression in macrophages after *Yersinia enterocolitica* infection. Role of different Yop effectors. *The Journal of biological chemistry* 277, 25133-25142.
- Schadt EE, Monks SA, Drake TA, Lusk AJ, Che N *et al.* (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297-302.
- Scherbik SV, Kluetzman K, Perelygin AA, Brinton MA (2007). Knock-in of the *Oas1b(r)* allele into a flavivirus-induced disease susceptible mouse generates the resistant phenotype. *Virology* 368, 232-237.
- Schneider DS, Ayres JS (2008). Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. *Nature reviews. Immunology* 8, 889-895.
- Schulte LN, Westermann AJ, Vogel J (2012). Differential activation and functional specialization of miR-146 and miR-155 in innate immune sensing. *Nucleic acids research*
- Schuring RP, Hamann L, Faber WR, Pahan D, Richardus JH *et al.* (2009). Polymorphism N248S in the human Toll-like receptor 1 gene is related to leprosy and leprosy reactions. *The Journal of infectious diseases* 199, 1816-1819.
- Sebastiani G, Olien L, Gauthier S, Skamene E, Morgan K *et al.* (1998). Mapping of genetic modulators of natural resistance to infection with *Salmonella typhimurium* in wild-derived mice. *Genomics* 47, 180-186.
- Sebbane F, Gardner D, Long D, Gowen BB, Hinnebusch BJ (2005). Kinetics of disease progression and host response in a rat model of bubonic plague. *The American journal of pathology* 166, 1427-1439.
- Sebbane F, Lemaitre N, Sturdevant DE, Rebeil R, Virtaneva K *et al.* (2006). Adaptive response of *Yersinia pestis* to extracellular effectors of innate immunity during bubonic plague. *Proceedings of the National Academy of Sciences of the United States of America* 103, 11766-11771.

- Seifert U, Bialy LP, Ebstein F, Bech-Otschir D, Voigt A *et al.* (2010). Immunoproteasomes preserve protein homeostasis upon interferon-induced oxidative stress. *Cell* 142, 613-624.
- Shao H, Burrage LC, Sinasac DS, Hill AE, Ernest SR *et al.* (2008). Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. *Proceedings of the National Academy of Sciences of the United States of America* 105, 19910-19914.
- Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM *et al.* (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* 139, 1255-1267.
- Shoham S, Huang C, Chen JM, Golenbock DT, Levitz SM (2001). Toll-like receptor 4 mediates intracellular signaling without TNF- α release in response to *Cryptococcus neoformans* polysaccharide capsule. *Journal of immunology* (Baltimore, Md. : 1950) 166, 4620-4626.
- Siddiqui MR, Meisner S, Tosh K, Balakrishnan K, Ghei S *et al.* (2001). A major susceptibility locus for leprosy in India maps to chromosome 10p13. *Nature genetics* 27, 439-441.
- Simon-Chazottes D, Frenkiel MP, Montagutelli X, Guenet JL, Despres P *et al.* (2011). Transgenic expression of full-length 2',5'-oligoadenylate synthetase 1b confers to BALB/c mice resistance against West Nile virus-induced encephalitis. *Virology* 417, 147-153.
- Simond P (1898). La propagation de la peste. *Annales de l'Institut Pasteur* 12, 625-687.
- Sissons J, Yan BS, Pichugin AV, Kirby A, Daly MJ *et al.* (2009). Multigenic control of tuberculosis resistance: analysis of a QTL on mouse chromosome 7 and its synergism with *sst1*. *Genes and immunity* 10, 37-46.
- Skamene E (1994). The Bcg gene story. *Immunobiology* 191, 451-460.
- Sobel ES, Satoh M, Chen Y, Wakeland EK, Morel L (2002). The major murine systemic lupus erythematosus susceptibility locus Sle1 results in abnormal functions of both B and T cells. *Journal of immunology* (Baltimore, Md. : 1950) 169, 2694-2700.
- Sorensen TI, Nielsen GG, Andersen PK, Teasdale TW (1988). Genetic and environmental influences on premature death in adult adoptees. *The New England journal of medicine* 318, 727-732.
- Sotiropoulou G, Pampalakis G (2010). Kallikrein-related peptidases: bridges between immune functions and extracellular matrix degradation. *Biological chemistry* 391, 321-331.
- Souza-Fonseca-Guimaraes F, Adib-Conquy M, Cavaillon JM (2012). Natural killer (NK) cells in antibacterial innate immunity: angels or devils? *Molecular medicine* (Cambridge, Mass.) 18, 270-285.
- Stein RA (2011). Super-spreaders in infectious diseases. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 15, e510-513.
- Stenseth NC, Atshabar BB, Begon M, Belmain SR, Bertherat E *et al.* (2008). Plague: past, present, and future. *PLoS medicine* 5, e3.
- Stienstra Y, van der Graaf WT, te Meerman GJ, The TH, de Leij LF *et al.* (2001). Susceptibility to development of *Mycobacterium ulcerans* disease: review of possible risk factors. *Tropical medicine & international health : TM & IH* 6, 554-562.
- Stinchcombe JR, Hoekstra HE (2008). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158-170.
- Stout RD, Jiang C, Matta B, Tietzel I, Watkins SK *et al.* (2005). Macrophages sequentially change their functional phenotype in response to changes in microenvironmental influences. *Journal of immunology* (Baltimore, Md. : 1950) 175, 342-349.
- Stout RD, Suttles J (2004). Functional plasticity of macrophages: reversible adaptation to changing microenvironments. *Journal of leukocyte biology* 76, 509-513.
- Straley SC, Harmon PA (1984). *Yersinia pestis* grows within phagolysosomes in mouse peritoneal macrophages. *Infection and immunity* 45, 655-659.

- Strober W, Murray PJ, Kitani A, Watanabe T (2006). Signalling pathways and molecular interactions of NOD1 and NOD2. *Nature reviews. Immunology* 6, 9-20.
- Stuart LM, Ezekowitz RA (2005). Phagocytosis: elegant complexity. *Immunity* 22, 539-550.
- Sturdevant DE, Virtaneva K, Martens C, Bozinov D, Ogundare O *et al.* (2010). Host-microbe interaction systems biology: lifecycle transcriptomics and comparative genomics. *Future microbiology* 5, 205-219.
- Su Z, Ishimori N, Chen Y, Leiter EH, Churchill GA *et al.* (2009). Four additional mouse crosses improve the lipid QTL landscape and identify Lipg as a QTL gene. *Journal of lipid research* 50, 2083-2094.
- Sullivan PM, Mezdour H, Aratani Y, Knouff C, Najib J *et al.* (1997). Targeted replacement of the mouse apolipoprotein E gene with the common human APOE3 allele enhances diet-induced hypercholesterolemia and atherosclerosis. *The Journal of biological chemistry* 272, 17972-17980.
- Takada T, Mita A, Maeno A, Sakai T, Shitara H *et al.* (2008). Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits. *Genome research* 18, 500-508.
- Takaesu G, Surabhi RM, Park KJ, Ninomiya-Tsuji J, Matsumoto K *et al.* (2003). TAK1 is critical for IkappaB kinase-mediated activation of the NF-kappaB pathway. *Journal of molecular biology* 326, 105-115.
- Takaoka A, Yanai H, Kondo S, Duncan G, Negishi H *et al.* (2005). Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* 434, 243-249.
- Takeda K, Akira S (2005). Toll-like receptors in innate immunity. *International immunology* 17, 1-14.
- Tang B, Xiao B, Liu Z, Li N, Zhu ED *et al.* (2010). Identification of MyD88 as a novel target of miR-155, involved in negative regulation of Helicobacter pylori-induced inflammation. *FEBS letters* 584, 1481-1486.
- Taniguchi T, Takaoka A (2002). The interferon-alpha/beta system in antiviral responses: a multimodal machinery of gene regulation by the IRF family of transcription factors. *Current opinion in immunology* 14, 111-116.
- Tarca AL, Draghici S, Khatri P, Hassan SS, Mittal P *et al.* (2009). A novel signaling pathway impact analysis. *Bioinformatics (Oxford, England)* 25, 75-82.
- Tollenaere C, Rahalison L, Ranjalaly M, Rahelinirina S, Duplantier JM *et al.* (2008). CCR5 polymorphism and plague resistance in natural populations of the black rat in Madagascar. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 8, 891-897.
- Tomaso H, Reisinger EC, Al Dahouk S, Frangoulidis D, Rakin A *et al.* (2003). Rapid detection of Yersinia pestis with multiplex real-time PCR assays using fluorescent hybridisation probes. *FEMS immunology and medical microbiology* 38, 117-126.
- Travassos LH, Carneiro LA, Girardin SE, Boneca IG, Lemos R *et al.* (2005). Nod1 participates in the innate immune response to Pseudomonas aeruginosa. *The Journal of biological chemistry* 280, 36714-36718.
- Trinchieri G (2010). Type I interferon: friend or foe? *The Journal of experimental medicine* 207, 2053-2063.
- Tsaih SW, Korstanje R (2009). Haplotype association mapping in mice. *Methods in molecular biology (Clifton, N.J.)* 573, 213-222.
- Tuite A, Gros P (2006). The impact of genomics on the analysis of host resistance to infectious disease. *Microbes and infection / Institut Pasteur* 8, 1647-1653.
- Turcotte K, Loredi-Osti JC, Fortin P, Schurr E, Morgan K *et al.* (2006). Complex genetic control of susceptibility to Mycobacterium bovis (Bacille Calmette-Guerin) infection in wild-derived Mus spretus mice. *Genes and immunity* 7, 684-687.

- Turner JK, McAllister MM, Xu JL, Tapping RI (2008). The resistance of BALB/cJ mice to *Yersinia pestis* maps to the major histocompatibility complex of chromosome 17. *Infection and immunity* 76, 4092-4099.
- Turner JK, Xu JL, Tapping RI (2009). Substrains of 129 mice are resistant to *Yersinia pestis* KIM5: implications for interleukin-10-deficient mice. *Infection and immunity* 77, 367-373.
- Une T, Brubaker RR (1984). In Vivo Comparison of Avirulent Vwa⁻ and Pgm⁻ or Pst^f Phenotypes of *Yersiniae*. *Infection and Immunity* 43, 895-900.
- van de Vosse E, van Dissel JT, Ottenhoff TH (2009). Genetic deficiencies of innate immune signalling in human infectious disease. *The Lancet infectious diseases* 9, 688-698.
- van Iterson M, t Hoen PA, Pedotti P, Hooiveld GJ, den Dunnen JT *et al.* (2009). Relative power and sample size analysis on gene expression profiling data. *BMC genomics* 10, 439.
- Van Lint P, Libert C (2007). Chemokine and cytokine processing by matrix metalloproteinases and its effect on leukocyte migration and inflammation. *Journal of leukocyte biology* 82, 1375-1381.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N *et al.* (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* 3, RESEARCH0034.
- Vanlaere I, Libert C (2009). Matrix metalloproteinases as drug targets in infections caused by gram-negative bacteria and in septic shock. *Clinical microbiology reviews* 22, 224-239, Table of Contents.
- Vanlaere I, Vanderrijst A, Guenet JL, De Filette M, Libert C (2008). Mx1 causes resistance against influenza A viruses in the *Mus spretus*-derived inbred mouse strain SPRET/Ei. *Cytokine* 42, 62-70.
- Vazquez-Torres A, Jones-Carson J, Mastroeni P, Ischiropoulos H, Fang FC (2000). Antimicrobial actions of the NADPH phagocyte oxidase and inducible nitric oxide synthase in experimental salmonellosis. I. Effects on microbial killing by activated peritoneal macrophages in vitro. *The Journal of experimental medicine* 192, 227-236.
- Verdugo RA, Farber CR, Warden CH, Medrano JF (2010). Serious limitations of the QTL/microarray approach for QTL gene discovery. *BMC biology* 8, 96.
- Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE *et al.* (2004). Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori* cag pathogenicity island. *Nature immunology* 5, 1166-1174.
- Vidal S, Tremblay ML, Govoni G, Gauthier S, Sebastiani G *et al.* (1995). The Ity/Lsh/Bcg locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the Nramp1 gene. *The Journal of experimental medicine* 182, 655-666.
- Vidal SM, Malo D, Marquis JF, Gros P (2008). Forward genetic dissection of immunity to infection in the mouse. *Annual review of immunology* 26, 81-132.
- Vigneau S, Rohrlisch PS, Brahic M, Bureau JF (2003). Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *Journal of virology* 77, 5632-5638.
- Villagra A, Cheng F, Wang HW, Suarez I, Glozak M *et al.* (2009). The histone deacetylase HDAC11 regulates the expression of interleukin 10 and immune tolerance. *Nature immunology* 10, 92-100.
- Vyse TJ, Morel L, Tanner FJ, Wakeland EK, Kotzin BL (1996). Backcross analysis of genes linked to autoantibody production in New Zealand White mice. *Journal of immunology (Baltimore, Md. : 1950)* 157, 2719-2727.
- Waddell SJ, Butcher PD, Stoker NG (2007). RNA profiling in host-pathogen interactions. *Current opinion in microbiology* 10, 297-302.
- Wade CM, Kulbokas EJ, 3rd, Kirby AW, Zody MC, Mullikin JC *et al.* (2002). The mosaic structure of variation in the laboratory mouse genome. *Nature* 420, 574-578.

- Wakui M, Morel L, Butfiloski EJ, Kim C, Sobel ES (2005). Genetic dissection of systemic lupus erythematosus pathogenesis: partial functional complementation between Sle1 and Sle3/5 demonstrates requirement for intracellular coexpression for full phenotypic expression of lupus. *Journal of immunology* (Baltimore, Md. : 1950) 175, 1337-1345.
- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L *et al.* (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470-476.
- Wang F, Hu S, Liu W, Qiao Z, Gao Y *et al.* (2011a). Deep-sequencing analysis of the mouse transcriptome response to infection with *Brucella melitensis* strains of differing virulence. *PloS one* 6, e28485.
- Wang H, Cheng F, Woan K, Sahakian E, Merino O *et al.* (2011b). Histone deacetylase inhibitor LAQ824 augments inflammatory responses in macrophages through transcriptional regulation of IL-10. *Journal of immunology* (Baltimore, Md. : 1950) 186, 3986-3996.
- Wang Z, Gerstein M, Snyder M (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* 10, 57-63.
- Warren HS, Fitting C, Hoff E, Adib-Conquy M, Beasley-Topliffe L *et al.* (2010). Resilience to bacterial infection: difference between species could be due to proteins in serum. *The Journal of infectious diseases* 201, 223-232.
- Watson J, Kelly K, Largen M, Taylor BA (1978 A). The genetic mapping of a defective LPS response gene in C3H/HeJ mice. *Journal of immunology* (Baltimore, Md. : 1950) 120, 422-424.
- Watson J, Largen M, McAdam KP (1978 B). Genetic control of endotoxic responses in mice. *The Journal of experimental medicine* 147, 39-49.
- Wayne ML, McIntyre LM (2002). Combining mapping and arraying: An approach to candidate gene identification. *Proceedings of the National Academy of Sciences of the United States of America* 99, 14903-14906.
- Weatherall DJ, Bell JI, Clegg JB, Flint J, Higgs DR *et al.* (1988). Genetic factors as determinants of infectious disease transmission in human communities. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 321, 327-348.
- Weber C, Fraemohs L, Dejana E (2007). The role of junctional adhesion molecules in vascular inflammation. *Nature reviews. Immunology* 7, 467-477.
- Welsh CE, Miller DR, Manly KF, Wang J, McMillan L *et al.* (2012). Status and access to the Collaborative Cross population. *Mammalian genome : official journal of the International Mammalian Genome Society* 23, 706-712.
- WHO (2009). Human plague: review of regional morbidity and mortality, 2004-2009. *Releve epidemiologique hebdomadaire / Section d'hygiene du Secretariat de la Societe des Nations = Weekly epidemiological record / Health Section of the Secretariat of the League of Nations* 85, 40-45.
- Wong SH, Hill AV, Vannberg FO (2010). Genomewide association study of leprosy. *The New England journal of medicine* 362, 1446-1447; author reply 1447-1448.
- WTCCC (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661-678.
- Xiao J, Liang Y, Li K, Zhou Y, Cai W *et al.* (2010). A novel strategy for genetic dissection of complex traits: the population of specific chromosome substitution strains from laboratory and wild mice. *Mammalian genome : official journal of the International Mammalian Genome Society* 21, 370-376.
- Yakub I, Lillibridge KM, Moran A, Gonzalez OY, Belmont J *et al.* (2005). Single nucleotide polymorphisms in genes for 2'-5'-oligoadenylate synthetase and RNase L in patients hospitalized with West Nile virus infection. *The Journal of infectious diseases* 192, 1741-1748.
- Yan BS, Kirby A, Shebzukhov YV, Daly MJ, Kramnik I (2006). Genetic architecture of tuberculosis resistance in a mouse model of infection. *Genes and immunity* 7, 201-210.

Yang H, Ke Y, Wang J, Tan Y, Myeni SK *et al.* (2011 A). Insight into bacterial virulence mechanisms against host immune response via the *Yersinia pestis*-human protein-protein interaction network. *Infection and immunity* 79, 4413-4424.

Yang H, Wang JR, Didion JP, Buus RJ, Bell TA *et al.* (2011 B). Subspecific origin and haplotype diversity in the laboratory mouse. *Nature genetics* 43, 648-655.

Yarovinsky F, Zhang D, Andersen JF, Bannenberg GL, Serhan CN *et al.* (2005). TLR11 activation of dendritic cells by a protozoan profilin-like protein. *Science (New York, N.Y.)* 308, 1626-1629.

Yersin A (1894). La peste bubonique à Hong Kong. *Annales de l'Institut Pasteur* VIII, p. 662-667.

Yoneyama M, Fujita T (2009). RNA recognition and signal transduction by RIG-I-like receptors. *Immunological reviews* 227, 54-65.

Yu M, Wang H, Ding A, Golenbock DT, Latz E *et al.* (2006). HMGB1 signals through toll-like receptor (TLR) 4 and TLR2. *Shock (Augusta, Ga.)* 26, 174-179.

Zhang CG, Chromy BA, McCutchen-Maloney SL (2005a). Host-pathogen interactions: a proteomic view. *Expert review of proteomics* 2, 187-202.

Zhang D, Zhang G, Hayden MS, Greenblatt MB, Bussey C *et al.* (2004). A toll-like receptor that prevents infection by uropathogenic bacteria. *Science (New York, N.Y.)* 303, 1522-1526.

Zhang FR, Huang W, Chen SM, Sun LD, Liu H *et al.* (2009). Genomewide association study of leprosy. *The New England journal of medicine* 361, 2609-2618.

Zhang J, Wheeler DA, Yakub I, Wei S, Sood R *et al.* (2005b). SNPdetector: a software tool for sensitive and accurate SNP detection. *PLoS computational biology* 1, e53.

Zhou L, Ying W, Han Y, Chen M, Yan Y *et al.* (2012). A proteome reference map and virulence factors analysis of *Yersinia pestis* 91001. *Journal of proteomics* 75, 894-907.

Figures « Etats des connaissances »

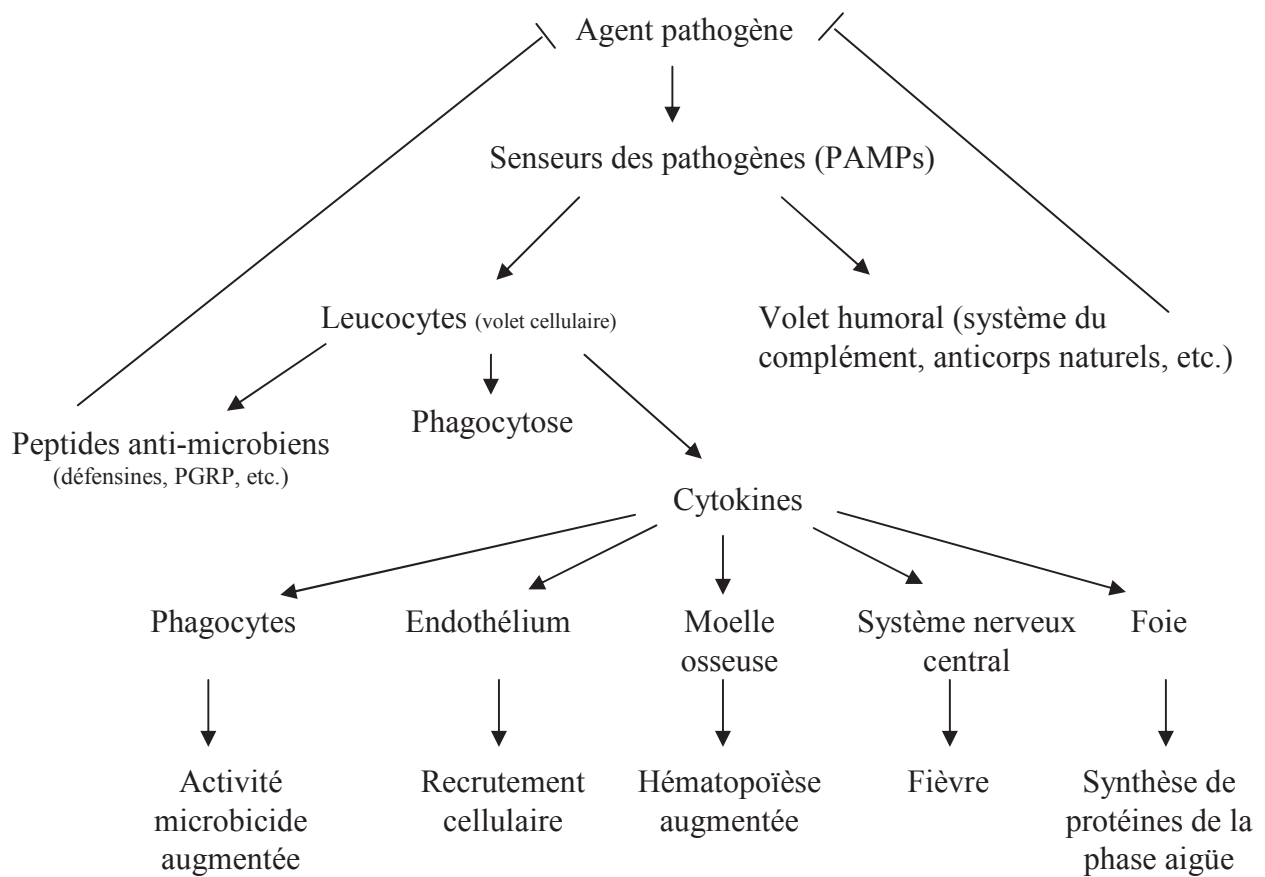


Figure 1. **Les différents acteurs de la réponse immunitaire innée.** Lors de l'infection, les agents microbiens sont reconnus par l'organisme. La réponse immunitaire innée se met en place impliquant un volet cellulaire et un volet humoral. Les leucocytes vont produire des peptides anti-microbiens, activer la phagocytose et sécréter des cytokines (pro- et anti-inflammatoires) qui vont agir sur l'ensemble de l'organisme. PAMPs : *Pathogen-Associated Molecular Patterns* ; PGRPs : *Peptidoglycane Recognition Proteins*, (D'après Kapetanovic et Cavaillon, 2007).

TLR / complexe	Localisation cellulaire	Ligands
TLR1-2	Surface cellulaire	lipopeptides tri-acétylés, Pam3CSK4
TLR2	Surface cellulaire	PSA
TLR2-6	Surface cellulaire	Peptidoglycane, phospholipomannane, hemagglutinines, porines
TLR3	Endosome	ARN double brin, Poly(I:C)
TLR4	Surface cellulaire	LPS, mannane, glycolipides, HSP60, HSP70, fibrinogène, HMGB1, nickel
TLR4-6	Surface cellulaire	OxLDL, fibrilles amyloid-b
TLR5	Surface cellulaire	Flagelline
TLR7	Endosome	ARN simple brin
TLR8	Endosome	ARN simple brin
TLR9	Endosome	ADN, CpG-A, CpG-B, CpG-C
TLR10 (h)	Surface cellulaire	ND
TLR11 (s)	Endosome	Profiline
TLR12 (s)	Endosome	ND
TLR13 (s)	Endosome	ND

Tableau 1. **Les *Toll-like* récepteurs et leurs ligands.** PSA = Polysaccharide A, LPS = lipopolysaccharide, HSP (60 or 70) = *heat shock protein* (60 or 70), HMGB1 = *high-mobility group box 1*, OxLDL = *oxidized low density lipoprotein*, ND = non déterminé, (h) = humain uniquement, (s) = souris uniquement (Akira *et al.*, 2006).

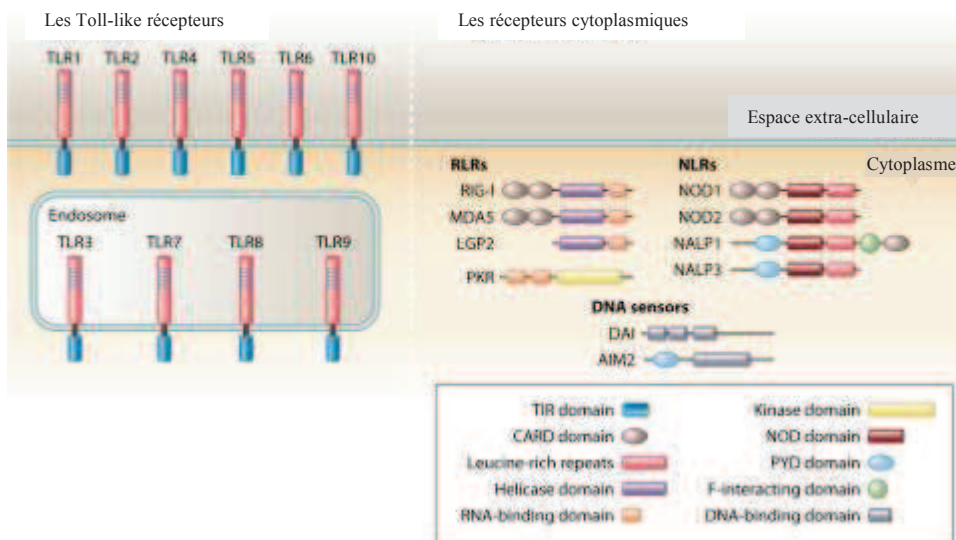


Figure 2. **Représentation structurale simplifiée des senseurs microbiens cellulaires principaux.** Les TLR sont transmembranaires, localisés sur la membrane cellulaire ou dans les endosomes. Parmi les senseurs intracytoplasmiques, les RLR se lient à l'ARN. DAI et AIM2 sont des senseurs de l'ADN. Les NLR reconnaissent différents motifs bactériens et viraux (voir Tableau 2), (d'après Mogensen, 2009).

Récepteur	Agoniste	Bactérie	Voie de signalisation activée
NOD1	GM-tripeptide	<i>Helicobacter pylori</i>	NF-κB
	meso-lanthionine, meso-DAP	<i>Shigella flexneri</i>	MAPK
	γ-d-Glu-DAP(iEDAP)	<i>Listeria monocytogenes</i> <i>Campylobacter jejuni</i>	
	d-lactyl-l-Ala-γ-Glu-meso-DAP-Gly (FK156)	Enteropathogenic <i>Escherichia coli</i>	
	heptanoly-l-γ-Glu-meso-DAP-Ala (FK565)	<i>Chlamydia pneumoniae</i> <i>Pseudomonas aeruginosa</i> <i>Bacillus</i> spp.	
		<i>Streptococcus pneumoniae</i>	NF-κB
NOD2	Muramyl dipeptide	<i>Listeria monocytogenes</i>	MAPK
	MurNAc-l-Ala-g-d-Glu-l-Lys (M-TRILys)	<i>Mycobacterium tuberculosis</i> <i>Salmonella</i> Typhimurium <i>Staphylococcus aureus</i> <i>Shigella flexneri</i>	

Tableau 2. Ligands, bactéries reconnues et voies de signalisation cellulaires activées par NOD1 et NOD2, (adapté de Chen *et al.*, 2009)

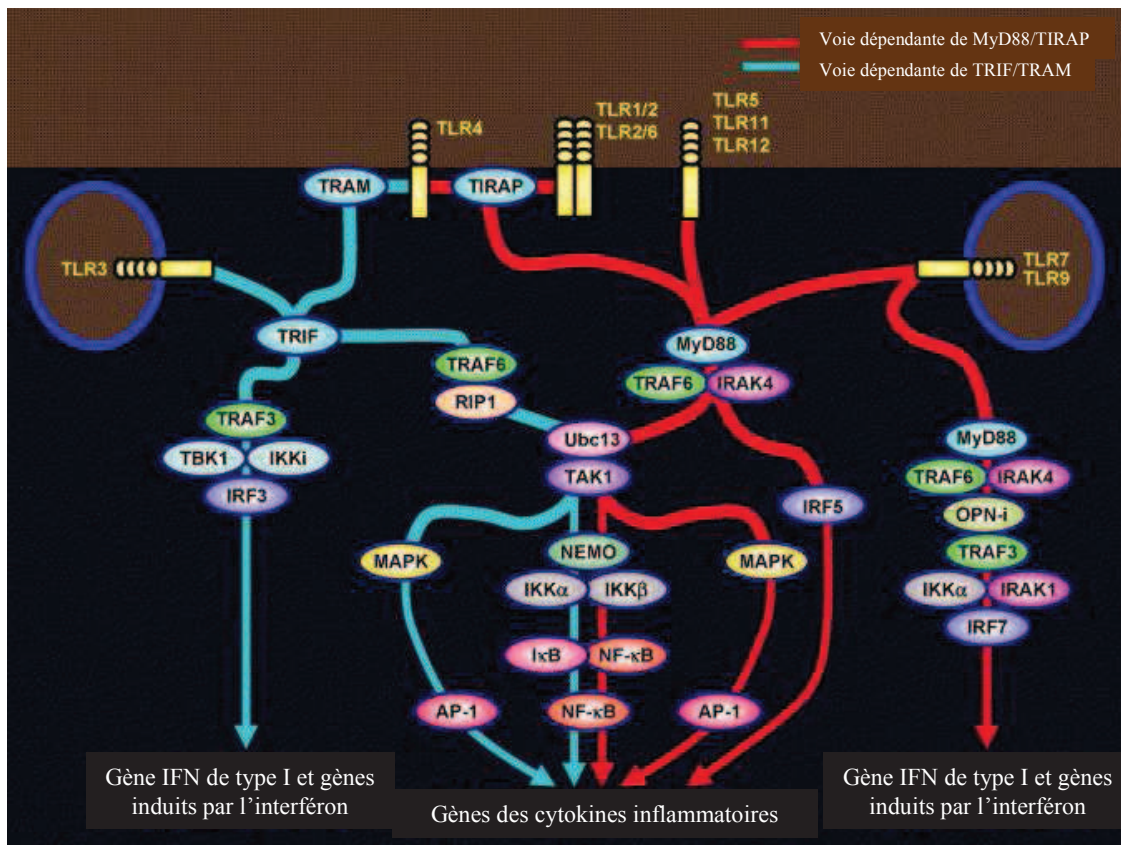


Figure 3. **Cascade de signalisation des TLR.** Cette cascade peut être divisée en deux grande voies. Voies de signalisation MyD88/TIRAP-dépendant (en rouge) et TRIF/TRAM-dépendant (en bleu). La majorité des TLR impliquent la voie MyD88; TLR3 et TLR4 peuvent activer la voie MyD88-indépendante via TRIF. Après stimulation, tous les TLR, hormis, TLR3, recrutent les protéines MyD88, IRAK et TRAF6. TAK1 active le complexe IKK afin de phosphoryler la sous-unité IκB. NF-κB est alors libéré dans le cytoplasme puis transloqué dans le noyau pour activer les gènes responsables de la production de cytokines inflammatoires (Kawai et Akira, 2007).

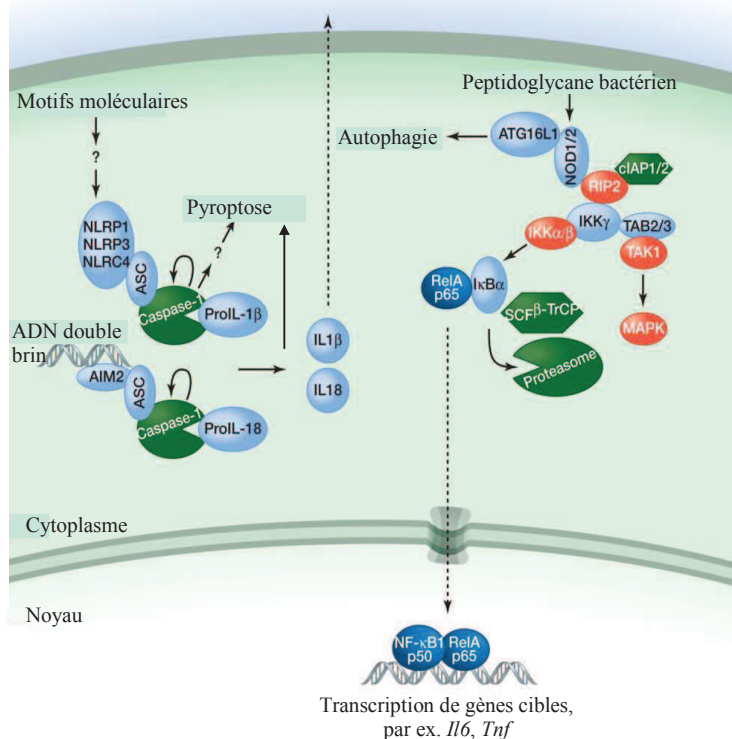


Figure 4. Voie de signalisation des NLR. NLRP1, NLRP3 et NLRP4 répondent à différents motifs moléculaires (d'agents pathogènes ou de lésion cellulaire) en engageant la protéine adaptatrice ASC qui permet l'activation de la caspase 1. AIM2 provoque la même réponse suite à la reconnaissance d'ADN double brin cytoplasmique. L'activation de la caspase 1 au sein des inflammasomes permet le clivage des pro-IL1β et pro-IL18 et la sécrétion de leurs formes biologiquement actives. L'activation de la caspase 1 provoque également une forme de mort cellulaire rapide, appelée « pyroptose ». L'activation de la caspase 1 peut aussi être effectuée par la caspase 11, qui agit alors en synergie avec l'inflammasome (voir figure 5, Rathinam *et al.*, 2012). NOD1 et NOD2 reconnaissent différents composants des peptidoglycane bactériens et stimulent soit le processus d'autophagie soit la transcription de gènes via l'activation de NF-κB et de la voie des MAP kinases. Cette dernière nécessite l'interaction de NOD1 ou NOD2 avec la kinase RIP2. L'« ubiquitination » de RIP2 par cIAP1/2 permet de recruter TAB2/3 et IKKγ ayant pour conséquence d'activer TAK1 et IKK (d'après Newton et Dixit, 2012).

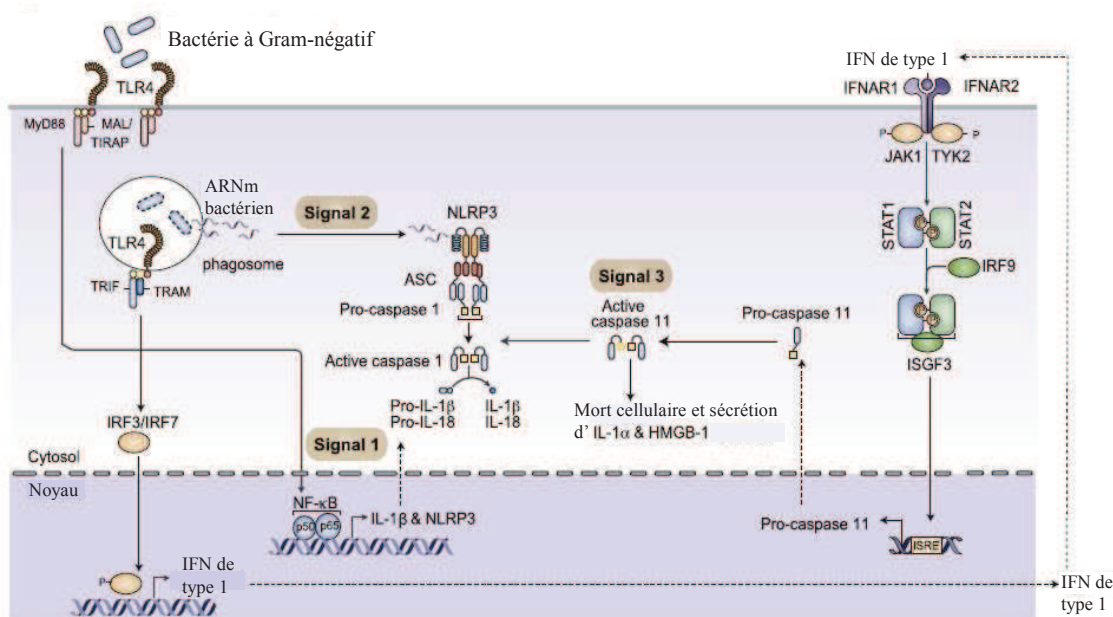


Figure 4bis. Model intégratif de l'activation caspase 11-dépendante de la caspase 1, par les voies TLR4-TRIF-IFN et NLRP3 suite à la reconnaissance d'une bactérie à Gram-négatif. En activant la voie Myd88-dépendante en aval de TLR4, les bactéries à Gram-négatif induisent la transcription des gènes *Il1β*, *Il18* et *Nlrp3* (signal 1). La dégradation des bactéries phagocytées provoque le relargage d'ARNm bactérien dans le cytosol, ce qui induit l'assemblage de l'inflammasome incluant NLRP3 (signal 2). De manière simultanée, l'activation de la voie TRIF en aval de TLR4 provoque la transcription du gène codant pour la pro-caspase 11 par l'intermédiaire des IFN de type 1 (signal 3). La caspase 11, en synergie avec l'inflammasome incluant NLRP3, active la caspase 1, qui clive les pro-IL1β et pro-IL18 en leur forme mature IL-1β et IL-18, médiateurs de l'inflammation (d'après Rathinam *et al.*, 2012).

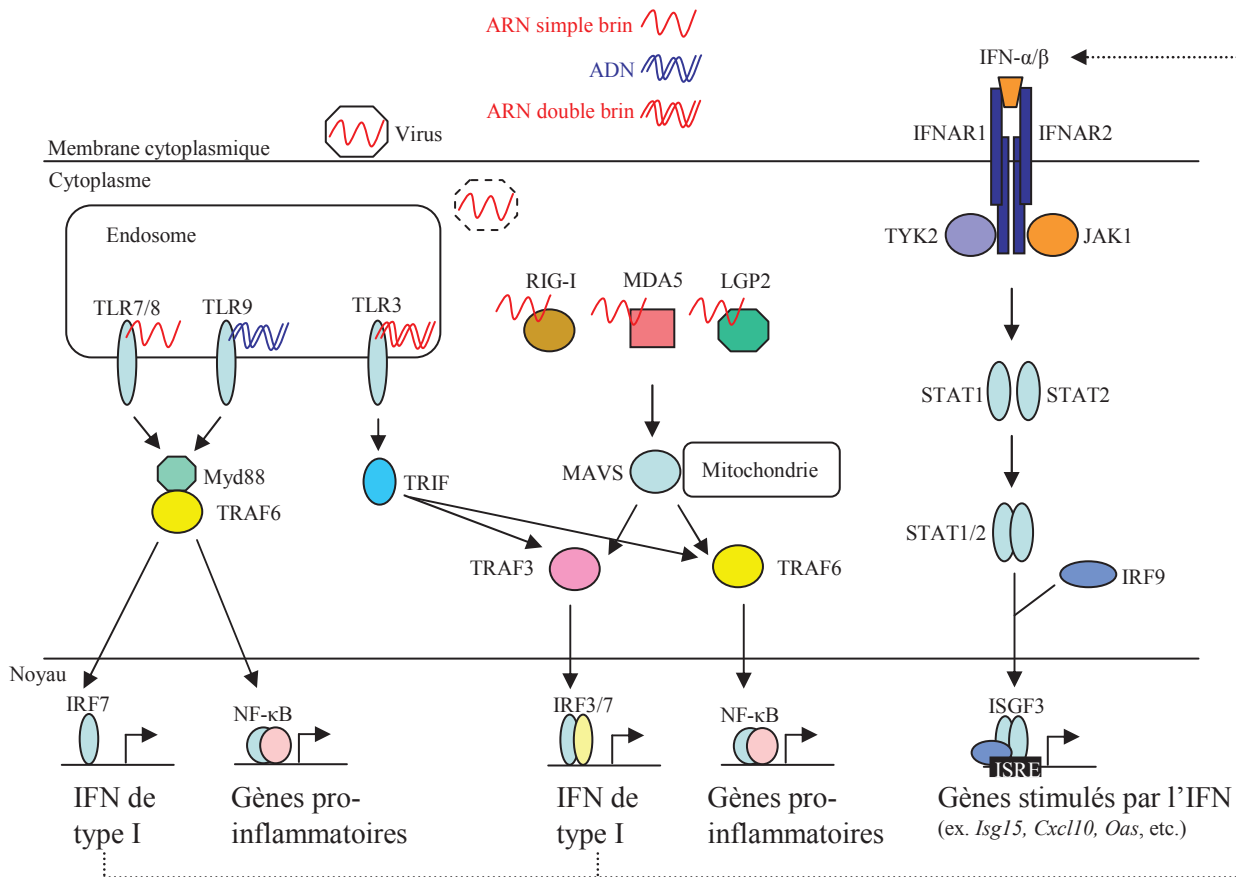


Figure 5. **Schéma des principales voies d'induction de l'interféron par les senseurs antiviraux cellulaires et transduction du signal par les interférons de type I.** Lors d'une infection virale, TLR3 reconnaît l'ARN double brin, TLR7/8 l'ARN simple brin et TLR9 certaines formes d'ADN. L'activation de la voie Myd88-dépendante active les facteurs de transcription IRF7 et NF-κB, menant à la synthèse d'interférons de type I et de gènes pro-inflammatoires. La transduction du signal *via* TRIF, active TRAF3 ou TRAF6, induisant l'expression des gènes interféron de type I (*via* IRF3 et IRF7) et de gènes pro-inflammatoires (*via* NF-κB). Les hélicases cytoplasmiques telles RIG-I, LGP2 ou MDA5 reconnaissent les ARN simple brin. La transduction du signal *via* l'adaptateur mitochondrial MAVS active également le facteur de transcription IRF3, qui peut se dimériser avec IRF7, et NF-κB. En retour, les IFN de type I (α et β) synthétisés se fixent sur leur récepteurs qui activent les kinases JAK1 et TYK2, conduisant à la formation du complexe de transcription ISG3, qui active à son tour la synthèse des gènes induits par l'IFN. ISRE, *Interferon-sensitive response element*; MAVS, *Mitochondrial Antiviral Signaling protein*; MAD5, *Melanoma differentiation-associated gene 5*; LGP2, *Laboratory of Genetics and Physiology 2* (d'après Kawai et Akira 2008 et Takoaka *et al.*, 2006).

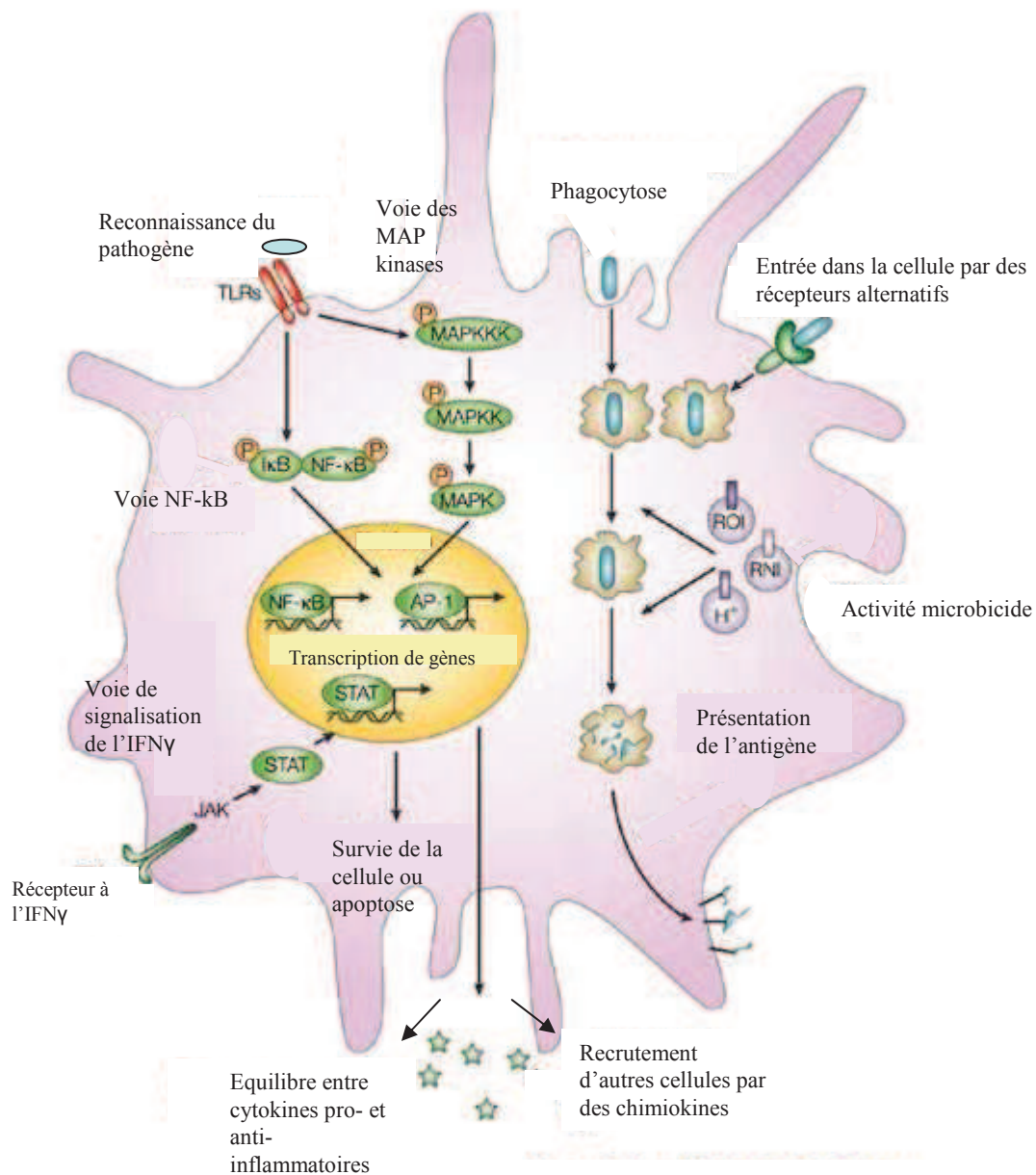


Figure 6. Schématisation des grandes voies cellulaires et des mécanismes mis en place suite à la reconnaissance d'un pathogène (ici une bactérie à Gram-négatif) dans un macrophage. AP-1, *activator protein-1*; IFN, interféron; IκB, *inhibitor of NF-κB*; JAK, Janus kinase; MAPK, *mitogen-activated protein kinase*; MAPKK, MAPK kinase; MAPKKK, MAPKK kinase; NF-κB, nuclear factor-kappaB; STAT, *signal transducer and activator of transcription*; TLR, *Toll-like receptor*; ROI, *Reactive Oxygen Intermediates*; RNI, *Reactive Nitrogen Intermediates* (d'après Rosenberger *et al.*, 2003).

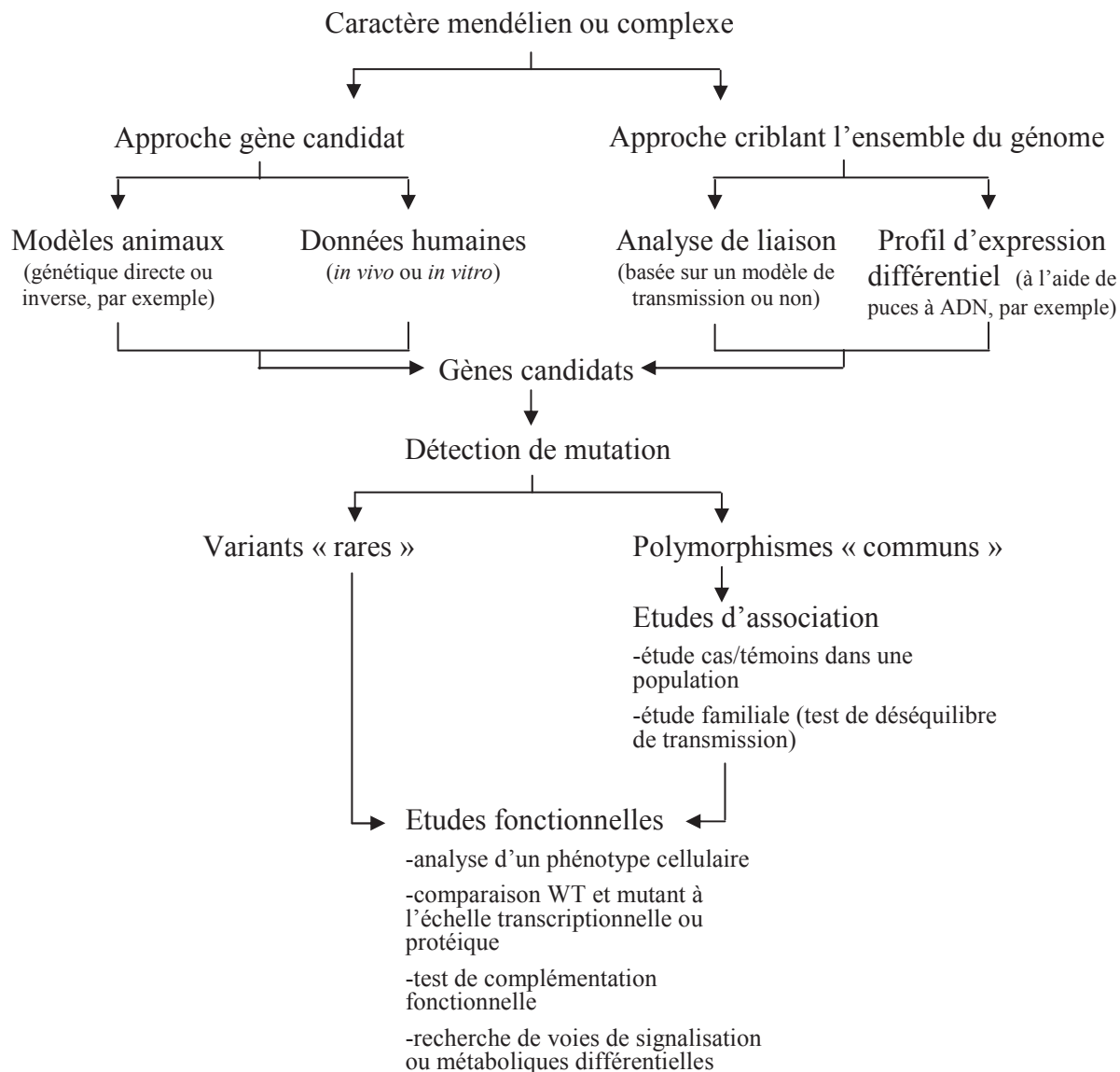


Figure 7. **Méthodes d'étude des maladies infectieuses chez l'Homme.** La recherche de gènes humains qui contribuent à la vulnérabilité aux maladies infectieuses peut être réalisée sans à priori par une approche pan-génomique ou être guidée par une hypothèse. Dans ce dernier cas, le choix des gènes candidats est basé sur la connaissance des modèles animaux, des données obtenues *in vitro*, des connaissances existantes sur la maladie en question ou une maladie proche. Des variants potentiellement causaux sont recherchés dans le gène candidat choisi. Les mutations rares ont généralement un impact fonctionnel important et pour lesquelles la corrélation entre génotype et phénotype peut-être établie facilement. En revanche, pour les polymorphismes communs à effet faible, la validation fonctionnelle des polymorphismes identifiés est plus difficile, l'effet d'une mutation isolée ayant des chances d'être beaucoup plus subtile. L'approche pangénomique (études de liaison et d'association) peut-être appliquée à des caractères complexes ou mendéliens pour définir de nouveaux candidats. Des analyses complémentaires (transcriptomique, protéomique, séquençage) peuvent être utilisées pour réduire la région chromosomique mise en évidence par les études pangénomiques. Des études fonctionnelles sont ensuite nécessaires pour mettre en évidence le lien causal entre la mutation identifiée et le phénotype observé (D'après Casanova et Abel, 2004).

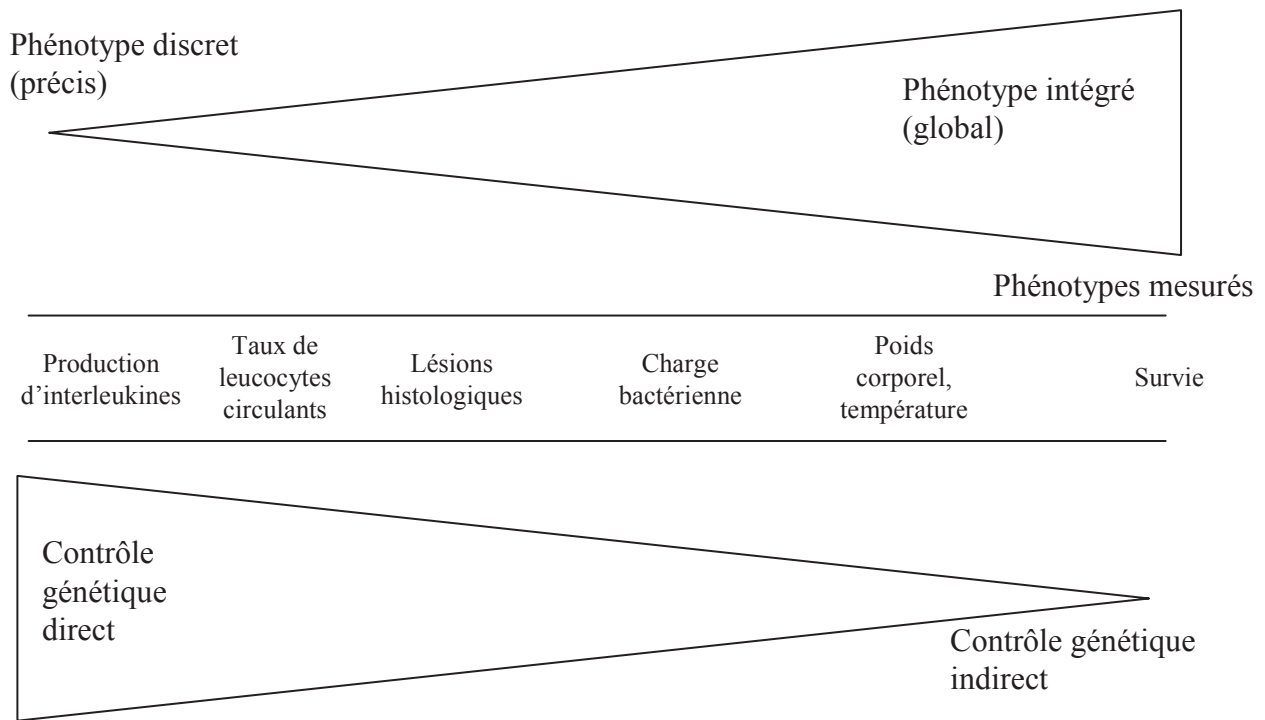
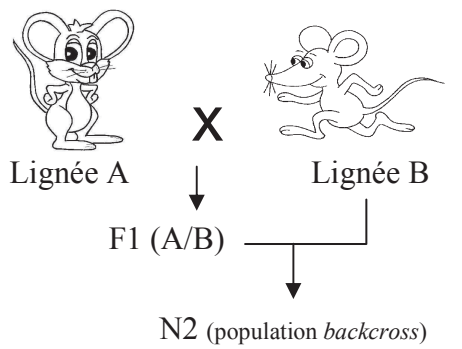


Figure 8. **Précision et contrôle génétique de différents phénotypes mesurés.** Les phénotypes précis tels que le dosage de cytokines dans un organe, sont « près des gènes » (contrôle génétique direct) alors que les phénotypes globaux comme la survie sont « loin des gènes » (contrôle génétique indirect).

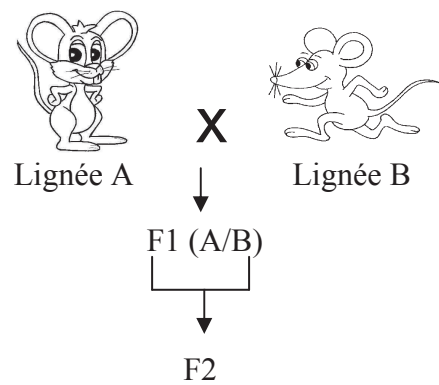
(A)



Deux génotypes possibles:

A/B et B/B

(B)



Trois génotypes possibles:

A/A, A/B et B/B

Figure 9. Croisement en deux générations de type N2 (*backcross*) (A) et F2 (*intercross*) (B). On trouve, dans une population F2, trois génotypes différents à chaque locus (A/A, A/B, B/B pour des lignées parentales A et B), alors que dans une population N2 issue d'un *backcross*, seulement deux catégories génotypiques sont présentes (A/B et B/B, si on a croisé en retour la lignée A avec la lignée B).

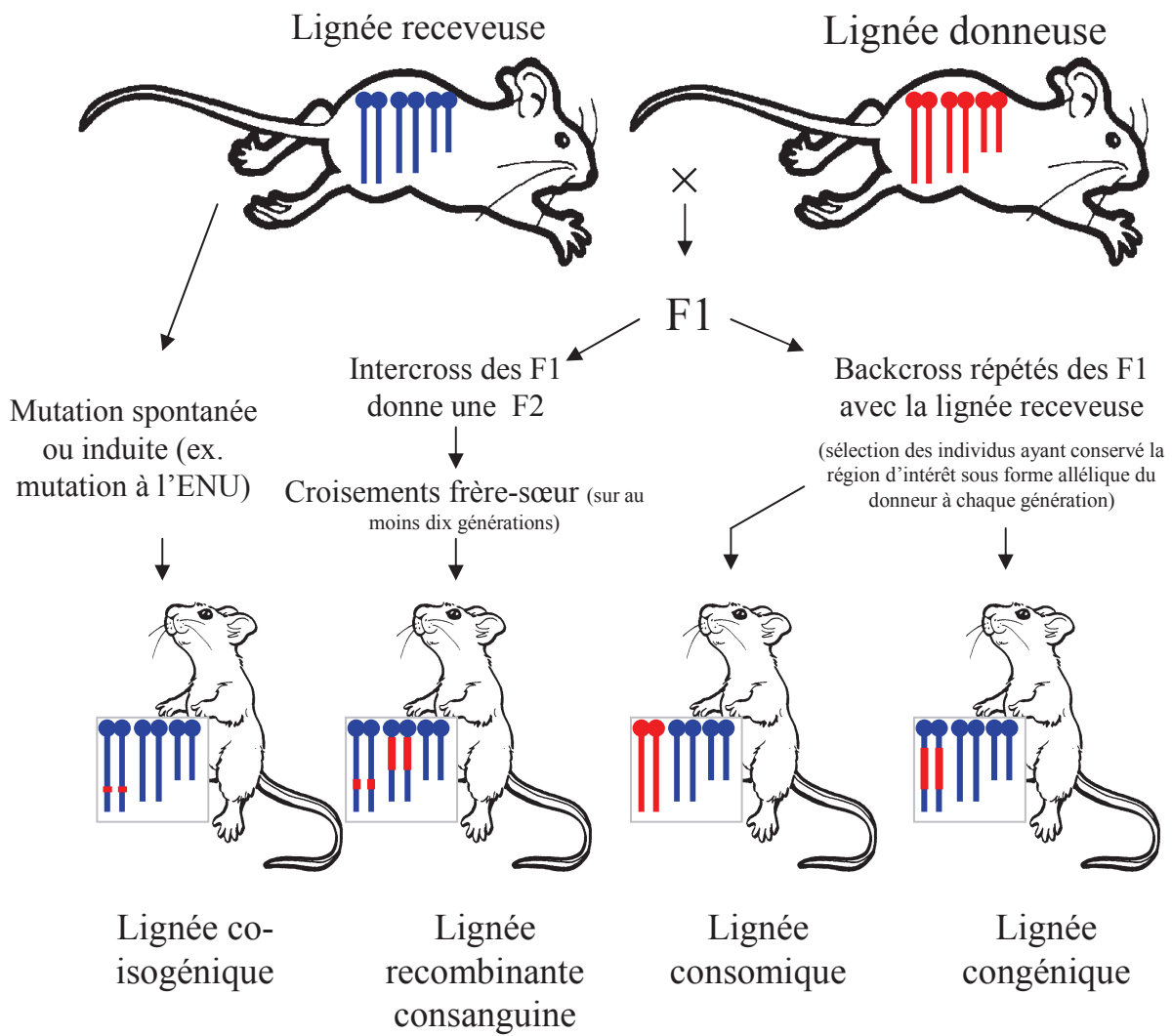


Figure 10. **Schéma d'obtention de différentes populations de référence de souris.** La production de lignées co-isogéniques, recombinantes consanguines, consomiques et congéniques sont représentées ici. D'après Rogner et Avner, *Nature Reviews Immunology*, 2003.

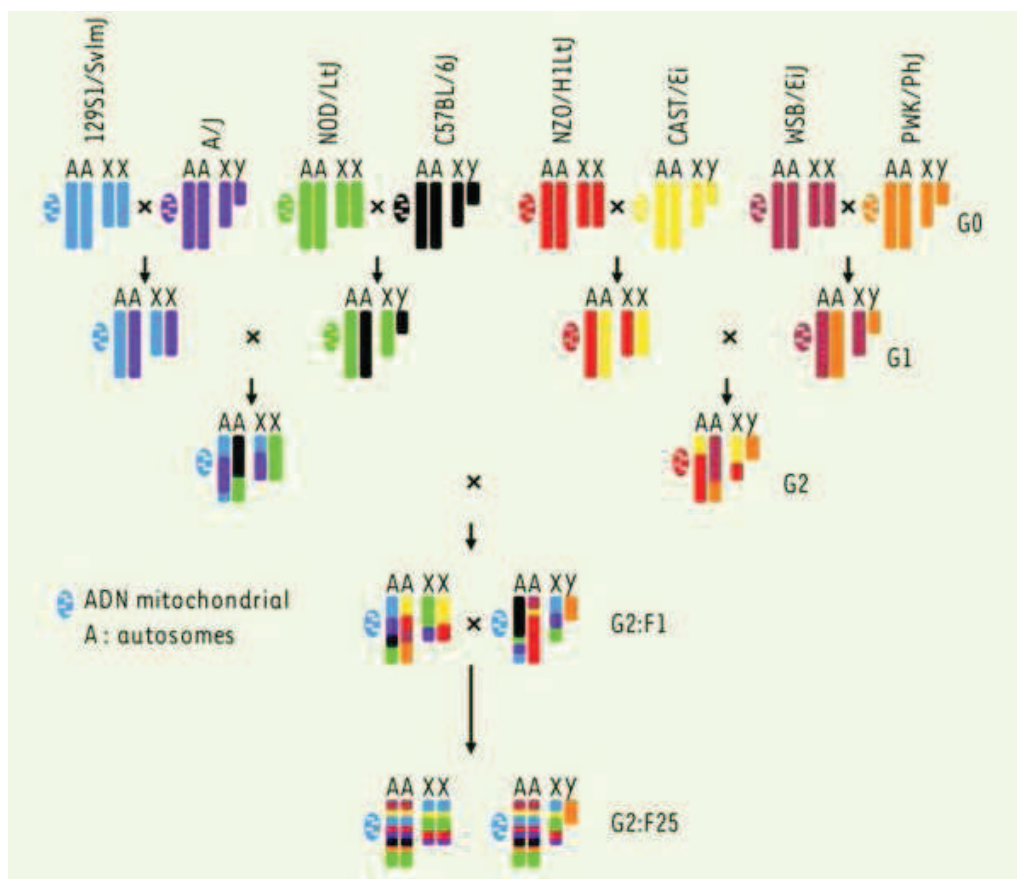


Figure 11. **Schéma de croisement ayant conduit aux lignées consanguines du CC.** Les croisements réalisés en première génération déterminent l'origine parentale des chromosomes X et Y et de l'ADN mitochondrial dans la lignée produite (d'après Chesler *et al.*, 2008 ; figure issue de Panthier et Montagutelli, 2012).

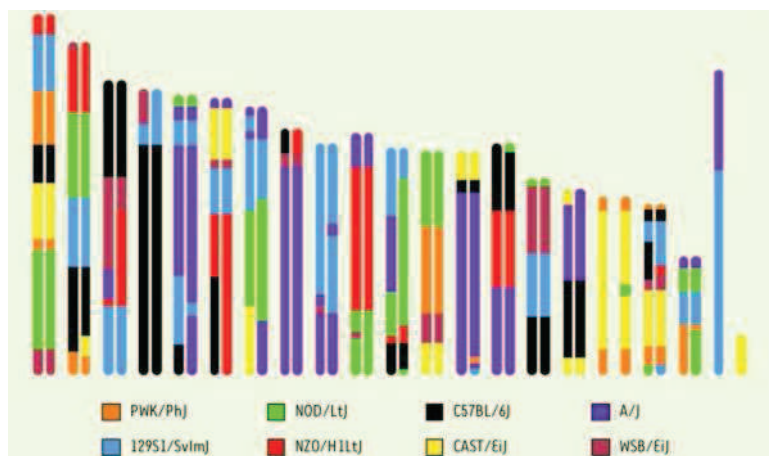


Figure 12. **Exemple de la composition génétique d'une lignée pré-CC.** Elle comprend en moyenne 143 segments chromosomiques provenant chacun d'une des huit lignées parentales. On remarque que certains segments ne sont pas encore fixés à l'état homozygote. On remarque également dans cette lignée, que la composition génétique du chromosome X provient seulement de deux lignées parentales. En théorie, entre une et cinq lignées participent à la composition génétique du chromosome X (d'après Aylor *et al.*, 2011 ; figure issue de Panthier et Montagutelli, 2012).

Génération	% moyen segments hét.	% génome donneur	% génome receveur
F1	100	50	50
N2	50	25	75
N3	25	12,5	87,5
N4	12,5	6,25	93,7
N5	6,3	3,125	96,7
N6	3,1	1,56	98,4
N7	1,6	0,78	99,2
N8	0,8	0,39	99,6
N9	0,4	0,2	99,8
N10	0,2	0,1	99,9

Tableau 3. **Proportions théoriques de provenance des génomes donneurs et receveurs dans les lignées congéniques au cours des générations de *backcross*.** Hét. : hétérozygote (d'après Montagutelli et Abitbol, 2004).

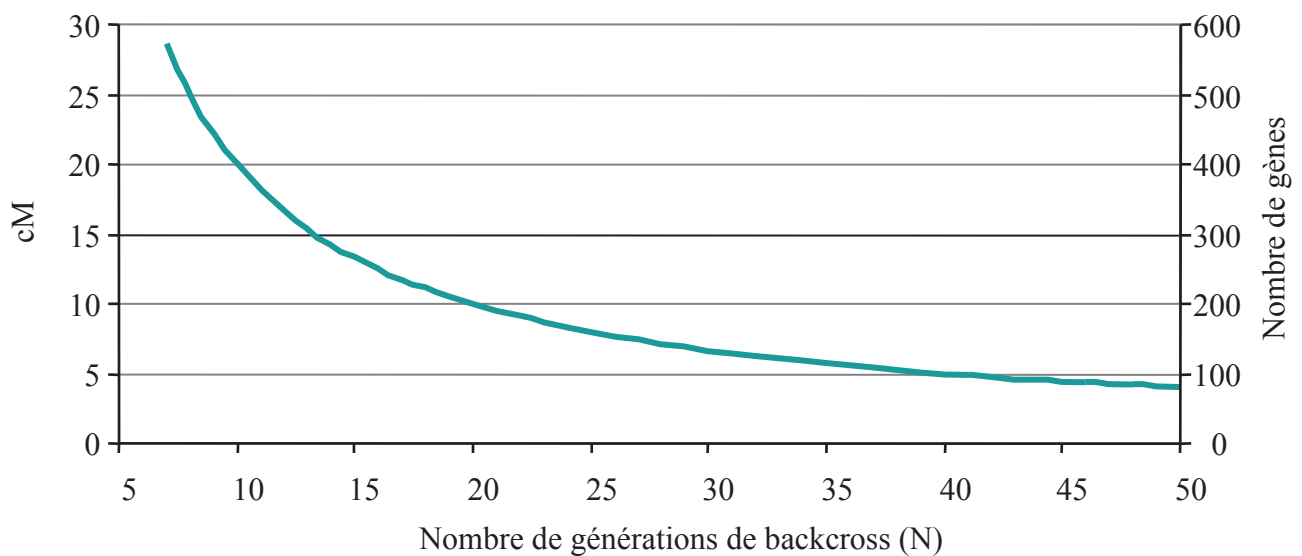


Figure 13. **Longueur du segment génomique transféré (en cM) de part et d'autre de l'intervalle congénique au cours des générations de *backcross*.** Le nombre de gènes moyen dans ce segment génomique est donné à droite (d'après Montagutelli et Abitbol, 2004).

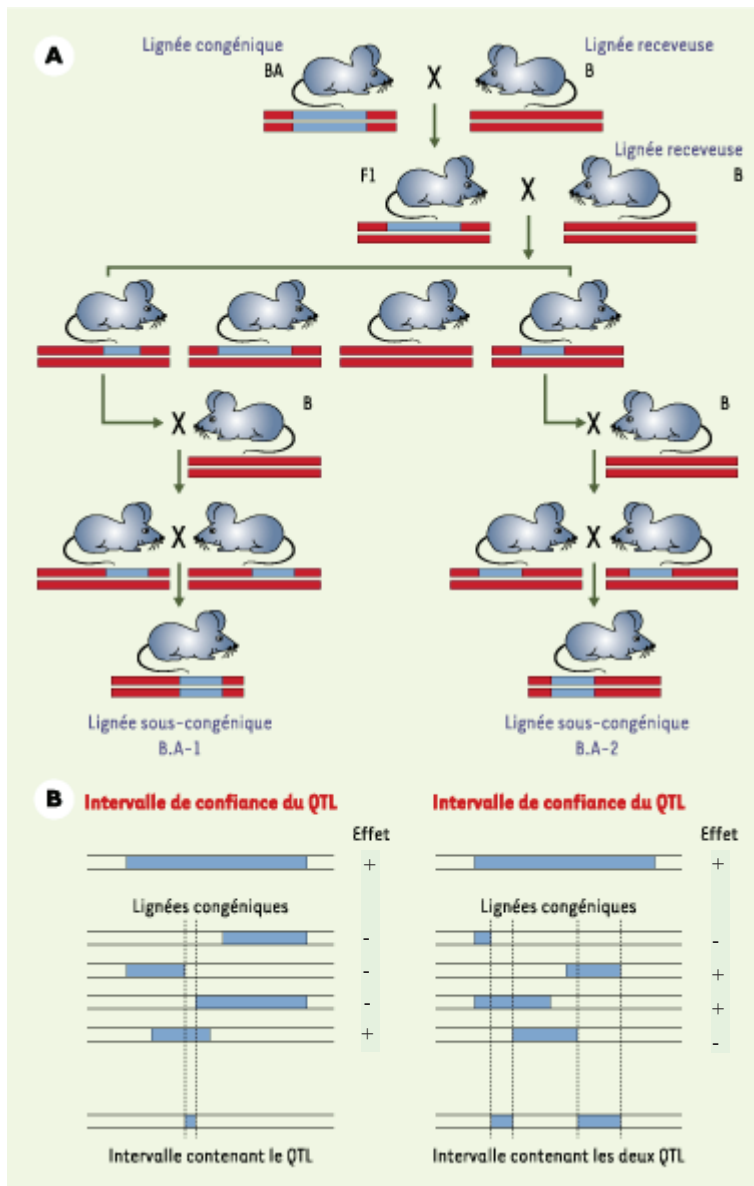
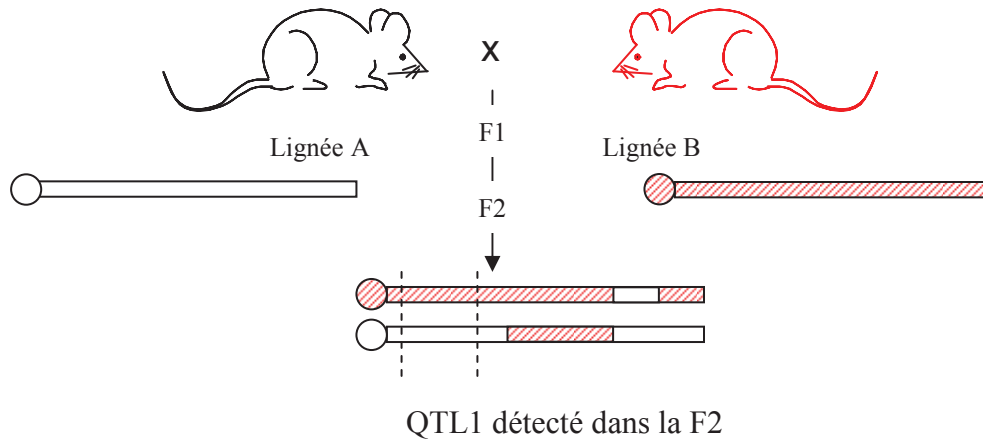
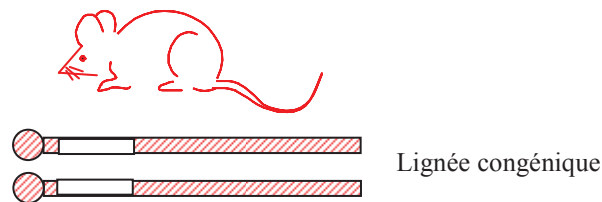


Figure 14. **Utilisation de lignées sous-congéniques pour affiner la localisation d'un QTL (quantitative trait locus).** **A.** À partir d'une lignée congénique porteuse d'un grand segment chromosomique, on dérive plusieurs lignées sous-congéniques portant des segments partiellement chevauchants : la lignée congénique (BA) est croisée deux fois avec la lignée receveuse (B) pour produire une population de quelques dizaines d'individus qui sont génotypés pour des marqueurs couvrant le segment initial. Les événements de recombinaison chez le parent F1 ont produit des chromosomes contenant un segment congénique plus court. Les individus qui les portent sont accouplés avec la lignée receveuse et leurs descendants hétérozygotes pour le même segment sont croisés entre eux pour obtenir une lignée sous-congénique homozygote. **B.** L'étude de lignées sous-congéniques permet de confirmer les résultats d'une analyse QTL effectuée sur un croisement, en comparant la composition génétique de chaque lignée avec ses propriétés pour le caractère étudié (« effet »). À gauche : les résultats permettent de confirmer l'existence d'un QTL et de réduire l'intervalle dans lequel il doit être recherché. À droite : l'analyse des lignées montre que l'intervalle défini initialement contient en réalité deux QTL liés, combinant leurs effets (aucun sous-intervalle ne permettant à lui seul d'expliquer l'ensemble des résultats, l'effet global est dû à l'effet de l'un et/ou de l'autre QTL), qui ont pu être isolés dans deux lignées sous-congéniques (Issu de Montagutelli et Abitbol, 2004).

(A) Etape 1. *Identification d'un QTL*



(B) Etape 2. *Création d'une lignée congénique pour le QTL sous forme allélique A sur fonds B*



(C) Etape 3. *Détection de QTL à effets faibles*

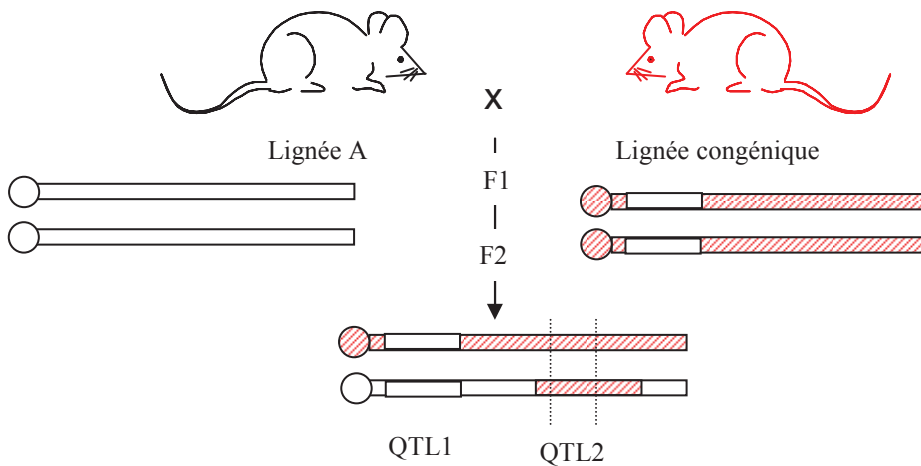


Figure 15. **Utilisation d'une lignée congénique pour identifier des QTL à effets faibles.** Un premier QTL à effet fort a été identifié dans la F2 entre les lignées A et B (A). Pour identifier des QTL à effets faibles on crée une lignée congénique portant l'allèle A du QTL sur fonds B (B) puis on effectue une F2 entre cette lignée congénique et la lignée A (C), (d'après Panthier *et al.*, 2003).

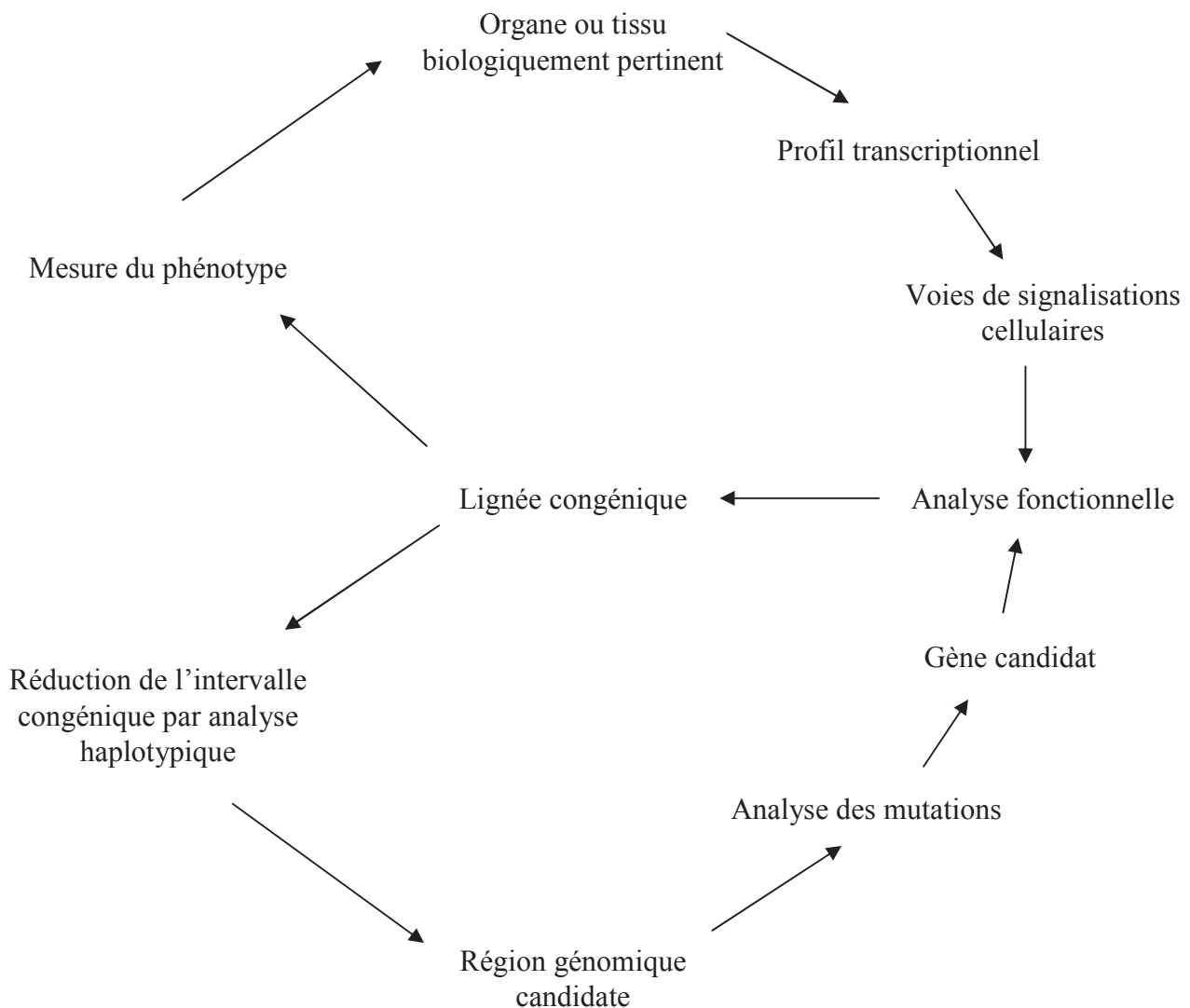


Figure 16. **Rôle pivot occupé par les lignées congéniques dans l'analyse des caractères complexes.** Un phénotypage exhaustif permet de déterminer les organes ou tissus qu'il est pertinent d'étudier et qui pourront être utilisés pour réaliser une étude transcriptionnelle. La région génomique candidate définie par l'analyse QTL peut être réduite par l'analyse des haplotypes portés par les lignées sous-congéniques. L'analyse des différences de séquence entre les lignées parentales dans le segment congénique peut servir à filtrer des gènes candidats. De même, l'analyse des gènes différenciellement exprimés entre la lignée congénique et la lignée receveuse, ainsi que l'analyse des voies de signalisation différenciellement activées, va permettre de réaliser un crible supplémentaire pour filtrer les gènes candidats. Les gènes identifiés seront sujet à des analyses fonctionnelles utilisant les congéniques (d'après Rogner et Avner, 2003).

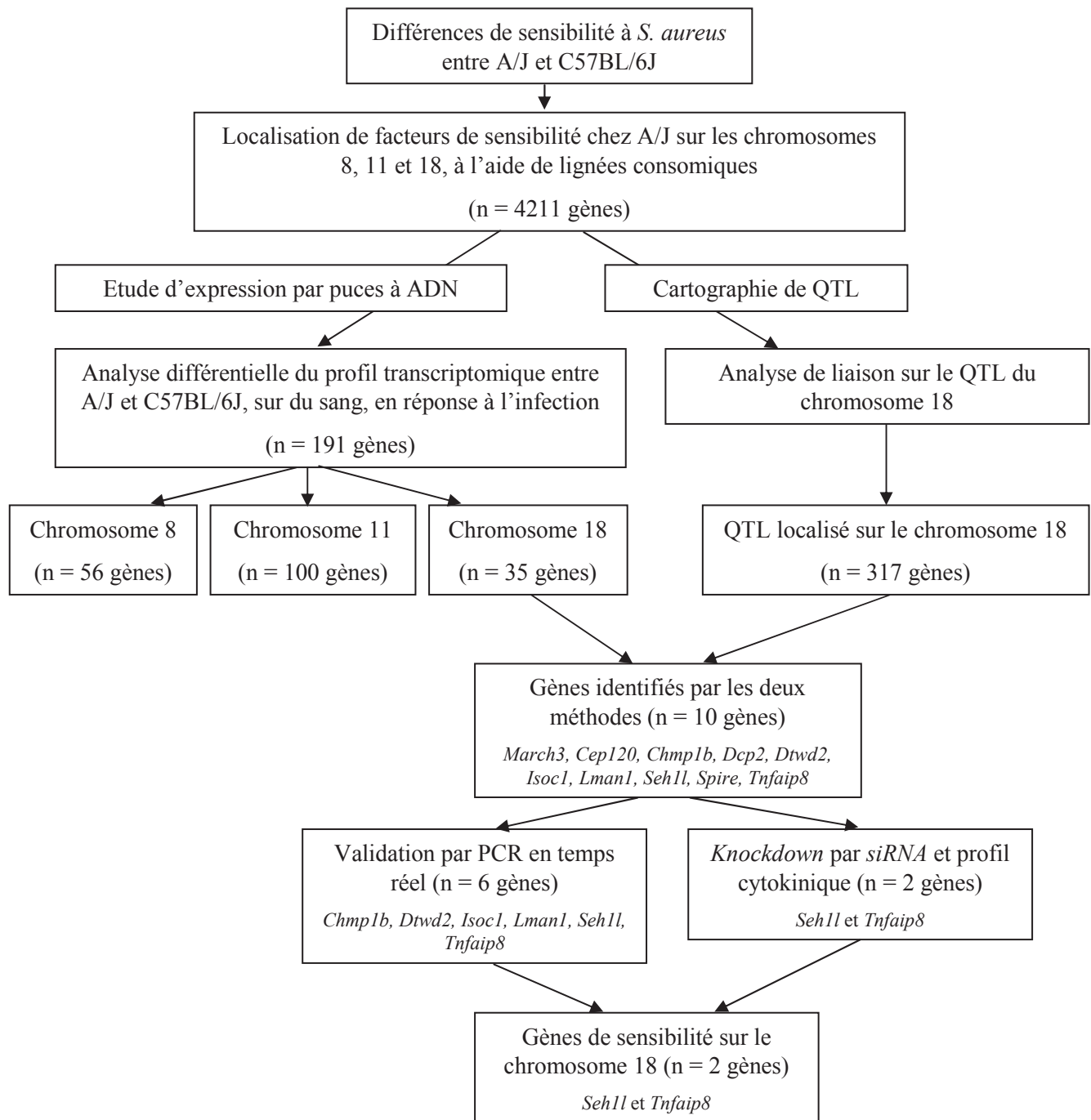


Figure 17. Stratégie globale utilisée dans l'étude de Ahn *et al.*, afin d'identifier des gènes associées à la sensibilité des souris A/J à une infection par *S. aureus* (d'après Ahn *et al.*, 2010).

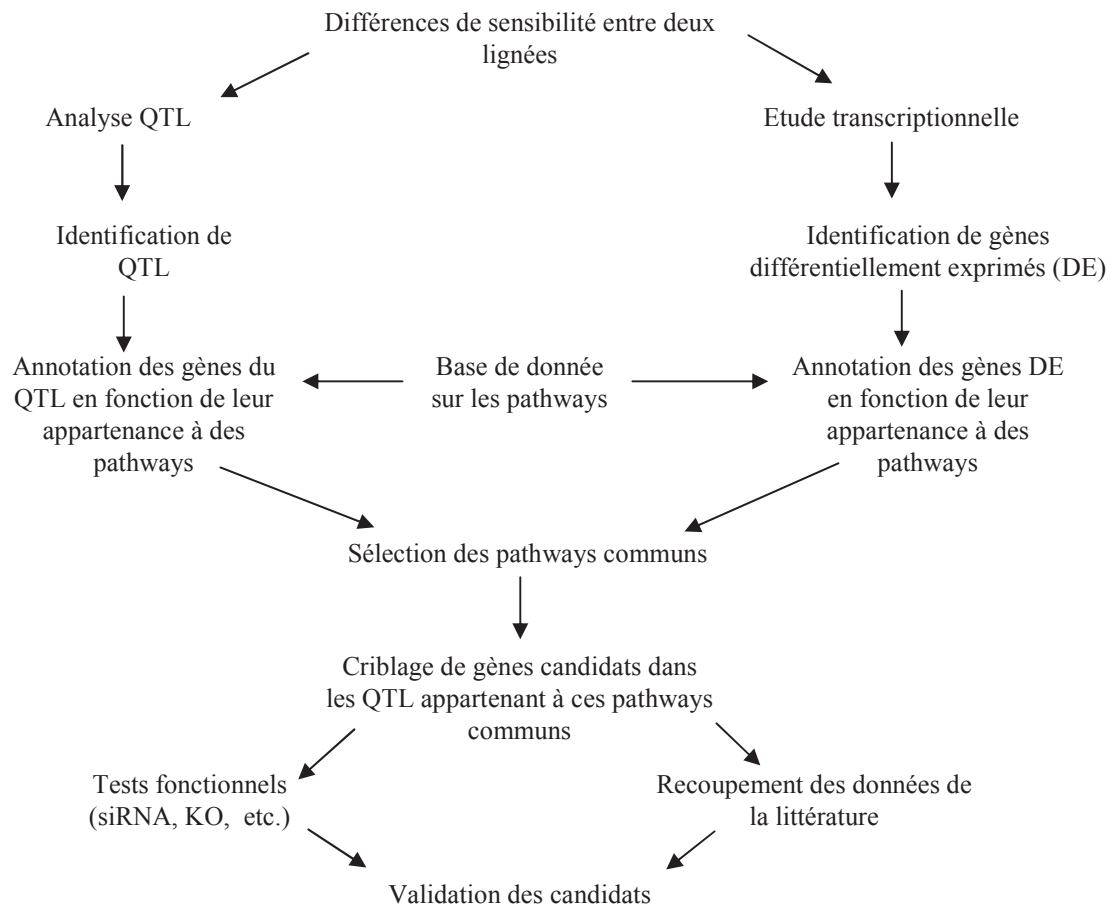


Figure 18. Stratégie de double criblage des gènes candidats pour la vulnérabilité à une maladie infectieuse, basée sur l'identification des gènes localisés dans les QTL et appartenant aux pathways différentiellement activés mis en évidence dans l'analyse transcriptionnelle entre deux lignées (d'après Fisher *et al.*, 2007).

Catégories fonctionnelles	Exemples de gènes
Gènes pro-inflammatoires	
Médiateurs de l'inflammation	<i>TNF, IL1b, IL6, IL8, CSF3 (GCSF)</i>
Chemokines	<i>CCL3 (MIP1a), CCL4 (MIP1b), CCL20 (MIP3a/LARC), CXCL1 (GRO1), CXCL2 (MIP2a/GRO2), CXCL3 (MIP2b/GRO3)</i>
Production de prostaglandines	<i>PTGS2 (cyclooxygenase-2)</i>
Gènes stimulés par l'interféron	
Gènes induits par l'IFN	<i>PRKR (protein kinase R), PML (promyelocytic leukaemia protein), OAS (2',5'-oligoadenylate synthetase)1, OAS2 et OASL, MX1 et MX2, GIP2 (ISG15), IFI16, ISG20 (HEM45), IFIM1</i>
Chemokines induites par l'IFN	<i>CCL8 (MCP2), CXCL9 (MIG), CXCL10 (IP10), CXCL11 (I-TAC)</i>
Gènes codant pour des "metallothionein"	<i>MITE, MT1G, MT1H, MT2A</i>
Gènes activateurs de la réponse immune	
Facteurs de transcription	<p>membres de la familles NF-κB et leur co-activateur <i>BCL3</i></p> <p>les composants de la protéine AP1 (activator protein-1) : <i>JUN, JUNB</i> et <i>FOSL2</i></p> <p>médiateurs des effets de l'interféron : <i>IRF1, IRF4, IRF7, ISGF3G, STAT1, STAT4, STAT5A</i> et leur co-activateur <i>NMI</i></p> <p><i>cAMP-response-element binders</i> : <i>ATF3, ATF4, BATF</i> et <i>CREM</i></p> <p><i>BCL6, CEBPG, EGR4, HIF1, NFATC1, SMAD7, XBP1</i></p>
Adatateurs cytoplasmiques dans la transduction du signal	<i>MYD88, TRADD, TRAF1, TRAF6</i>
Gènes modulateurs/qui limitent la réponse immune	
Régulation négative de la voie NF-κB	<i>NFKBIA, NFKBIE, TANK, TNFAIP3</i> (ou A20) et <i>TNIP1</i>
Kinases	<i>MAP2K3, MAP3K4</i> et <i>MAP3K8</i>
Phosphatases	<i>DUSP1, DUSP2, DUSP4, DUSP5, DUSP6</i> et <i>DUSP8</i>
Gènes impliqués dans l'apoptose	
Gènes pro-apoptotiques	<i>BBC3, CASP1, CASP3, CASP4, CASP5, CASP7, TRAIL</i> et <i>TRADD</i>
Facteurs anti-apoptotiques	<i>BIRC2, BIRC3, CFLAR</i> (ou FLIP) et la famille <i>BCL2</i> dont <i>BCL2A1</i> et <i>MCL1</i>
Autres gènes de la réponse commune de l'hôte	
Activation des lymphocytes	<i>CD80 (B7-1), CD83, CD86 (B7-2), TNFRSF5 (CD40)</i>
Présentation de l'antigène	<i>HLA-E, HLA-F, HLA-G, PSMA4, PSMB8, PSMB9, PSMB10, PSME2, TAP1</i> et <i>TAP2</i>
Adhésion cellulaire	<i>ADRM1, CD6, CD38, CD44, CD53, ICAM1, ITGA5, ITGAX, ITGB8, LGALS9, LGALS3BP</i>
Invasion tissulaire, processus inflammatoires (par ex. activation et inactivation des cytokines)	<i>MMP1, MMP7, MMP10, MMP12, MMP14, MMP19</i>

Tableau 4. Exemples des gènes appartenant à la réponse commune de l'hôte mise en évidence dans la méta-analyse de Jenner et Young (d'après Jenner et Young, 2005).

Etude	Das <i>et al.</i> , 2007	Rogers <i>et al.</i> , 2007	Comer <i>et al.</i> , 2010	Galindo <i>et al.</i> , 2009	Liu <i>et al.</i> , 2009
Modèle	In vitro, 26°C	Bubonique/ septicémique	Bubonique/ septicémique	Pulmonaire	Pulmonaire
Espèce	Humain	Souris	Rat	Souris	Souris
Donneur/lignée	Donneurs sains	BALB/c	Brown Norway	Swiss-Webster	BALB/c
Organe/cellules	Monocytes	Rate	Nœud lymphatique	Poumon, foie et rate	Poumon, foie et rate
Souche de <i>Y. p.</i>	KIM5	CO92	195/P	CO92	201
Inoculation	-	Intrapéritonéale	Intradermale	Intranasale	Intranasale
Dose	10MOI	257 ufc	10 ³ ufc	5 LD50	10 ⁶ ufc
Témoin	Non-infecté	Non-infecté	Non-infecté	Non-infecté	Non-infecté
Temps (h)	1 à 8	48	60	12 et 48	12 et 48
# de gènes testés	1185	34.000	28.000	34.000	13.000
Type de données	Filtrées	Filtrées	Filtrées	Filtrées	Filtrées
Type de puce à ADN utilisée	Clontech human	Affymetrix GeneChip Mouse Genome 430 2.0	Rat 230 2.0 Affymetrix GeneChip	Affymetrix GeneChip Mouse Genome 430 2.0	Mouse PancChip 6.25

Tableau 5. Tableau récapitulatif des caractéristiques des cinq études de transcriptome réalisées sur l'hôte en réponse à *Y. pestis*. # de gènes testés: nombre de gènes testés.

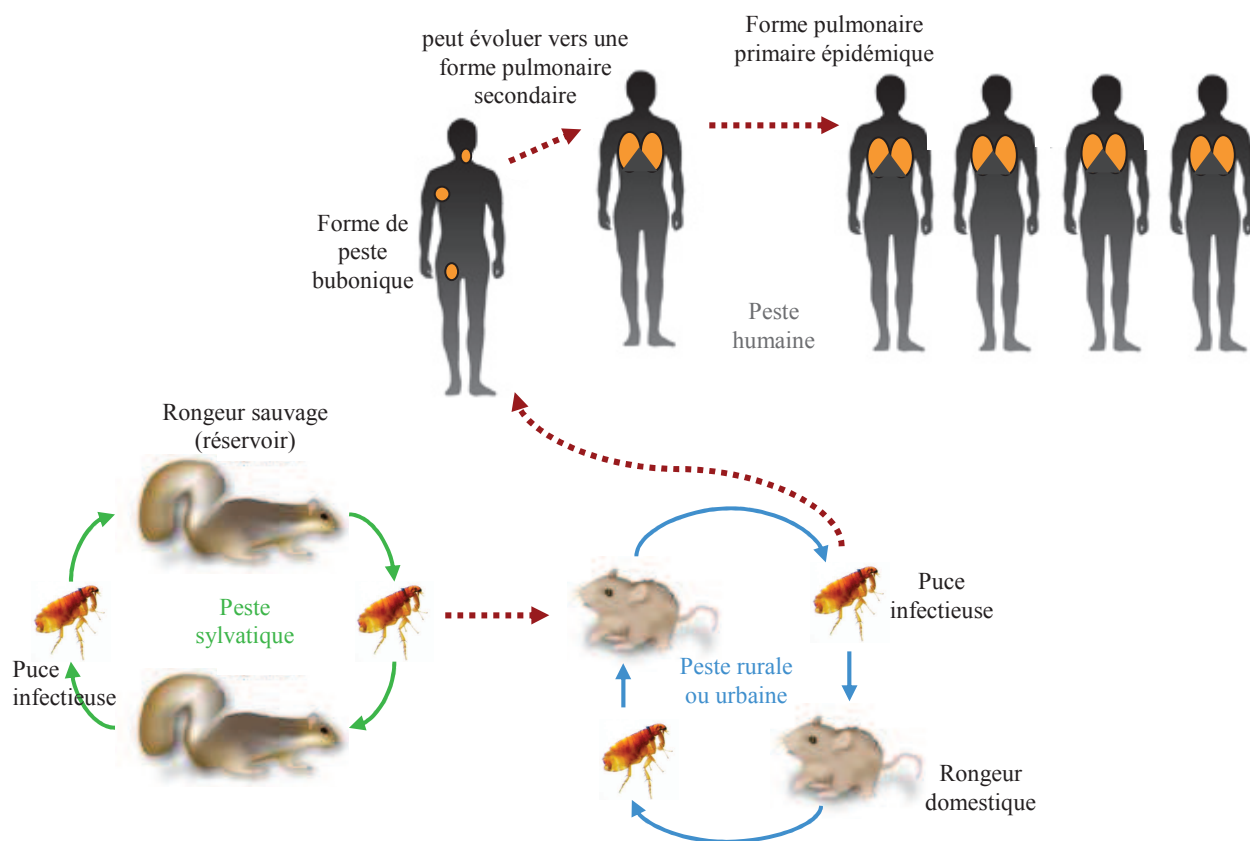


Figure 19. **Cycles et voies de transmission de la peste.** Les flèches pleines sont les voies usuelles de transmission, les flèches en pointillés sont les voies occasionnelles de transmission. D'après le site de l'OMS (<http://www.who.int/fr/>).

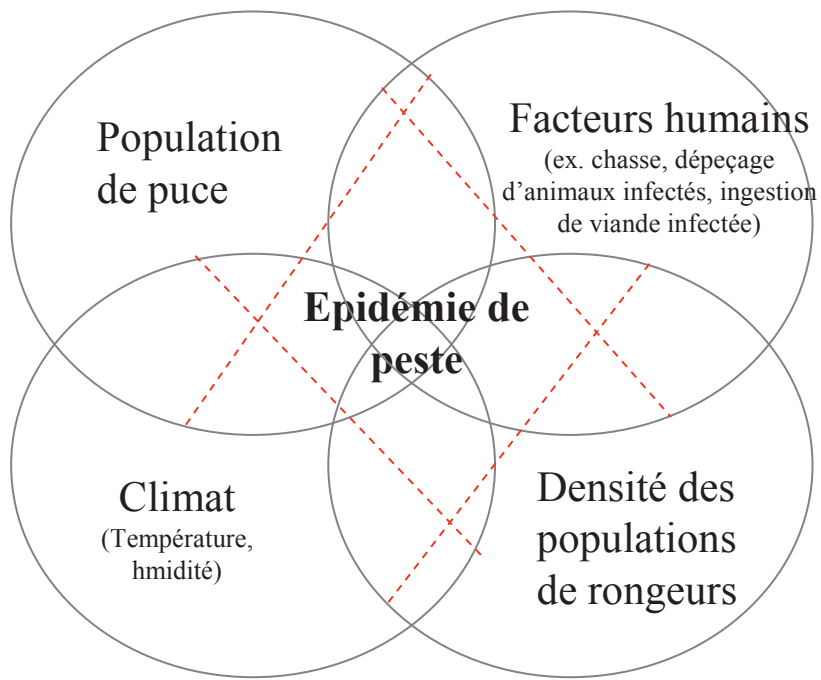


Figure 20. Schématisation de l'action des facteurs environnementaux, de l'état des populations de vecteurs, de la densité de rongeurs et des facteurs humains prédisposant à une épidémie de peste. La ligne rouge pointillée représente le seuil au-delà duquel la probabilité de voir une épidémie apparaître est augmentée (d'après l'étude de Samia *et al.*, 2011).

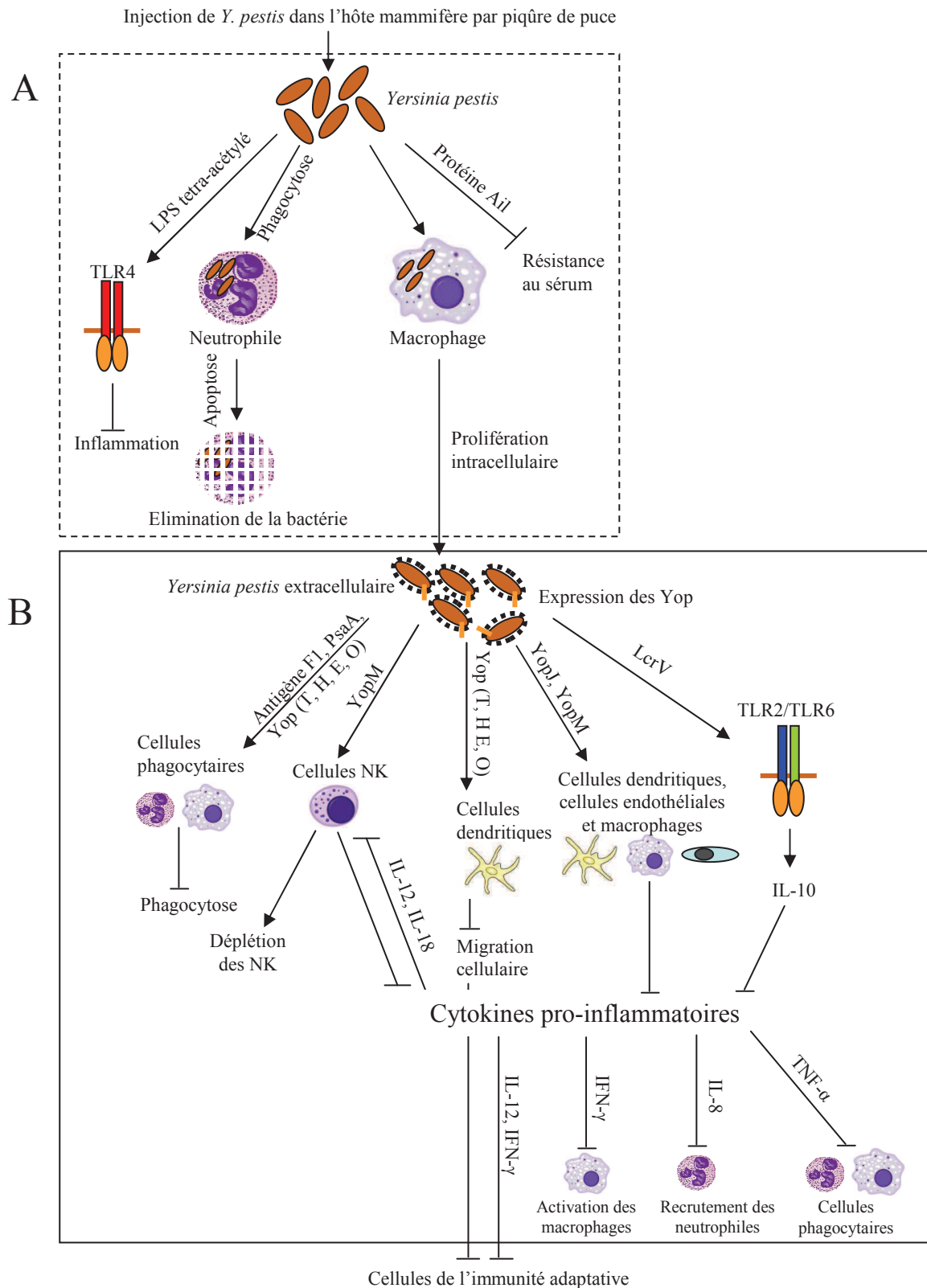


Figure 21. **Mécanismes de défense de *Y. pestis* contre la réponse immune inné de l'hôte.** (A) Mécanismes de défense dans le stade précoce de l'infection. Le passage de la bactérie à 37°C dans l'hôte mammifère provoque l'expression chez *Y. pestis* d'un LPS tétra-acétylé qui ne stimule pas TLR4 et confère à la bactérie une résistance à la lyse médiée par le complément. La bactérie internalisée dans les macrophages prolifère et exprime différents facteurs de virulence qui lui permettent de déjouer la réponse immune de l'hôte. (B) Mécanismes de défense après le relargage des bactéries hors des macrophages. *Y. pestis* exprime des Yop qui lui permette de résister à la phagocytose et d'inhiber la production de cytokines pro-inflammatoires, ayant pour conséquence d'atténuer la mise en place d'une réponse immune adaptative (d'après Li et Yang, 2008).

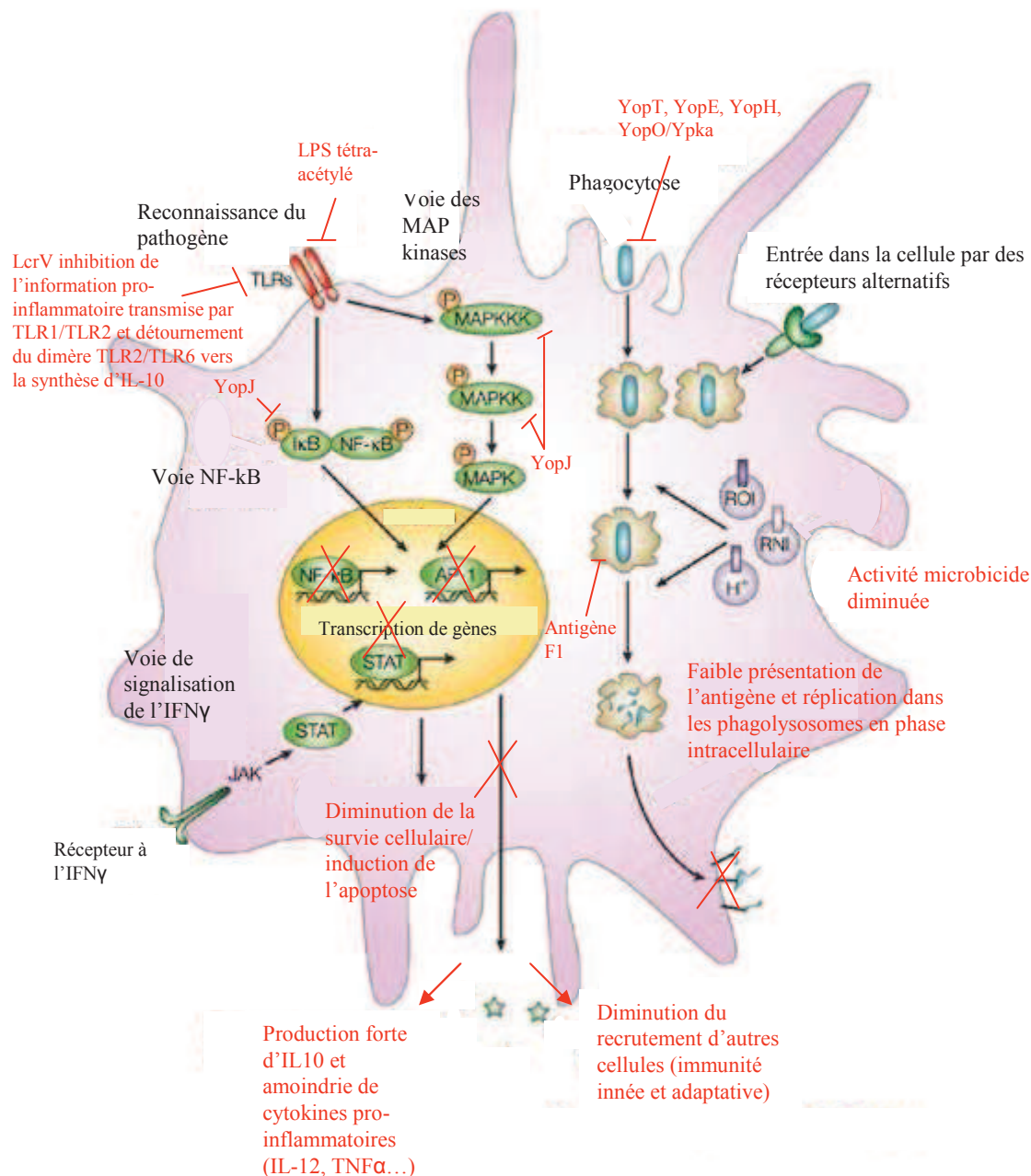


Figure 22. **Mécanismes de détournement de la réponse immunitaire par *Y. pestis* dans un macrophage.** Inhibition de l'internalisation par phagocytose (phase extracellulaire) par YopT, H, E et O; et inhibition de la dégradation (phase intracellulaire) par la production de l'antigène de capsule F1. Inhibition des voies MAPK et NF-κB à plusieurs niveaux par YopJ, résultant en une synthèse amoindrie des gènes pro-inflammatoires et anti-apoptotiques. Induction par LcrV de la synthèse d'IL-10, engageant les macrophages ciblés vers un profil d'activation anti-inflammatoire (M2). Inhibition de l'activation de TLR4 par une tétra-acétylation du LPS. AP-1, *activator protein-1*; IFN, interféron; IκB, *inhibitor of NF-κB*; JAK, Janus kinase; MAPK, *mitogen-activated protein kinase*; MAPKK, MAPK kinase; MAPKKK, MAPKK kinase; NF-κB, nuclear factor-kappaB; STAT, *signal transducer and activator of transcription*; TLR, *Toll-like receptor*; ROI, *Reactive Oxygen Intermediates*; RNI, *Reactive Nitrogen Intermediates* (schéma de base issu de Rosenberg *et al.*, 2003).

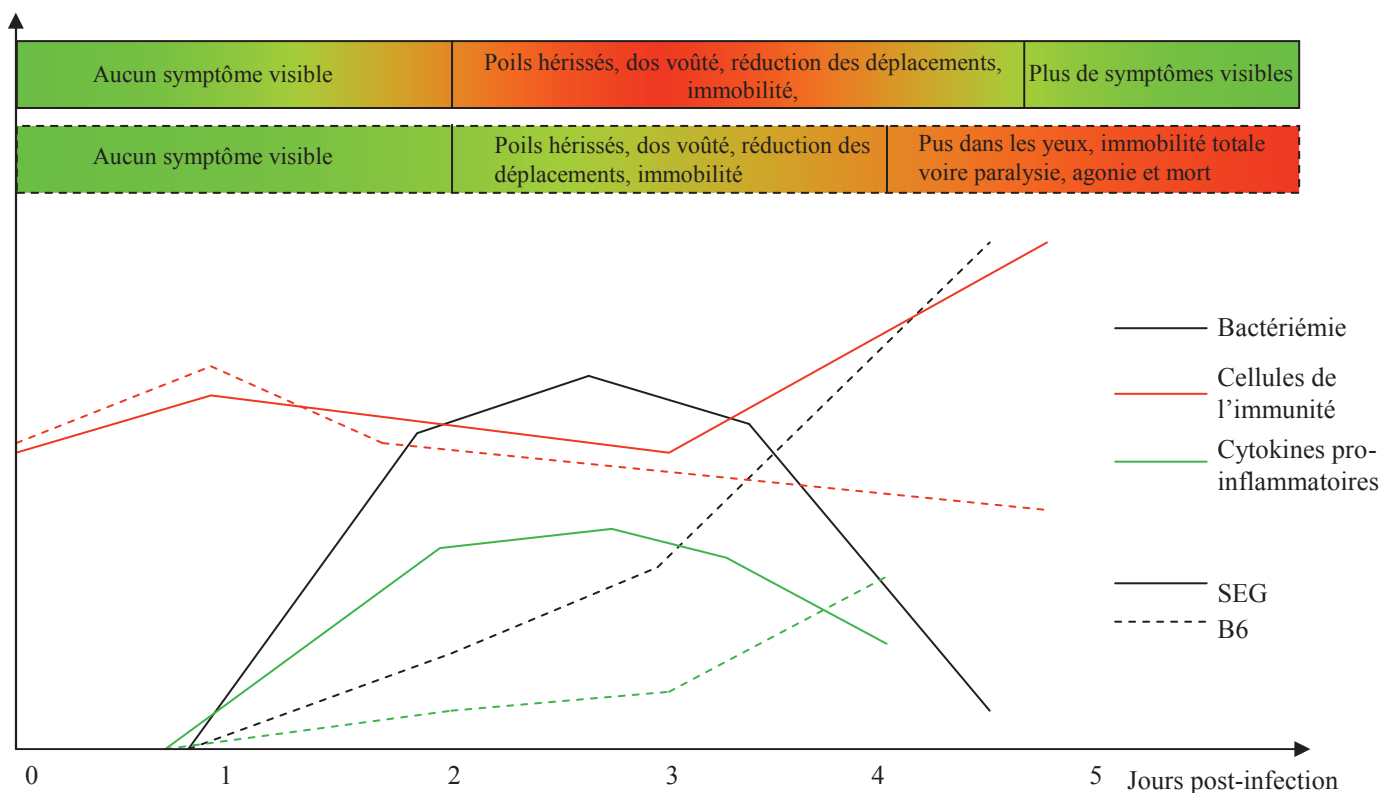


Figure 23. Résumé schématique des différences physio-pathologiques mises en évidence par Demeure *et al.* entre SEG et B6 au cours d'une infection expérimentale par 10^2 bactéries de la souche CO92 de *Y. pestis*. Les symptômes observés lors de l'infection sont présentés dans la partie supérieure du graphe. Des courbes schématiques de la bactériémie, de la quantification des populations cellulaires immunes et des cytokines pro-inflammatoires différenciellement exprimées sont présentées pour SEG (ligne pleine) et B6 (ligne pointillée), (d'après Demeure *et al.*, 2011).

Figures « Matériels et Méthodes »

Figures « Résultats »

Partie 1

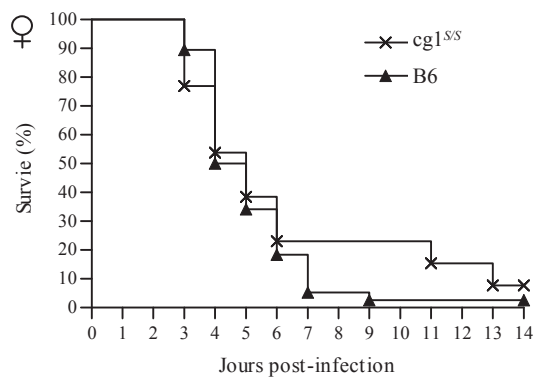


Figure 1. Courbe de survie sur 14 jours des femelles *cg1^{S/S}* (n=13) et B6 (n=38) après injection de 10² ufc de la souche CO92 de *Y. pestis*. Les souris *cg1^{S/S}* ne diffèrent pas des témoins B6 ni pour le taux de survie, ni pour la durée de survie.

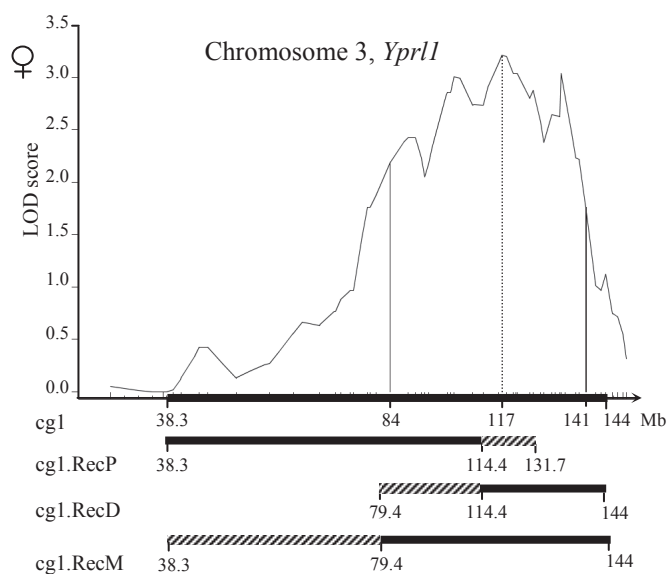


Figure 2. Courbe de LOD score illustrant la position du pic du QTL (ligne verticale en pointillés fins), l'intervalle de confiance du QTL à 95% (lignes verticales pleines), le segment génomique transféré dans la lignée congénique *cg1* et les lignées sous-congéniques *cg1.RecP*, *cg1.RecD* et *cg1.RecM* (lignes épaisses horizontales noires). La lignée *cg1.RecP* couvre la partie proximale de l'intervalle congénique transféré, *cg1.RecD* couvre la partie distale et *cg1.RecM* couvre l'intervalle de confiance à 95% du QTL. Les lignes horizontales hachurées représentent les limites non-définies des segments transférés, dont le point de recombinaison est situé entre les deux positions génomiques indiquées. Les positions génomiques sont indiquées en Mb.

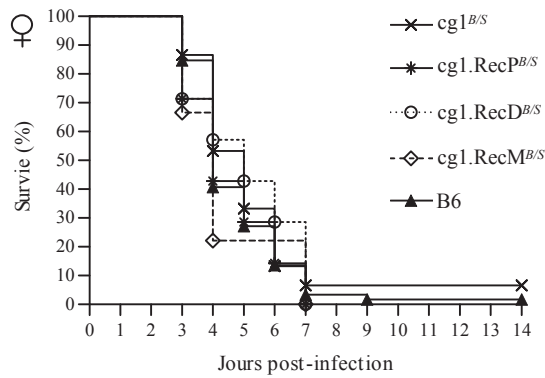


Figure 3. Courbe de survie des femelles $cg1^{B/S}$ ($n=15$), $cg1.RecP^{B/S}$ ($n=7$), $cg1.RecD^{B/S}$ ($n=7$), $cg1.RecM^{B/S}$ ($n=9$) et B6 ($n=58$) après inoculation de 10^2 ufc de *Y. pestis*. Aucune des lignées n'était différente de B6 ou de la lignée congénique $cg1^{B/S}$, que ce soit pour le taux de survie ou la durée de survie.

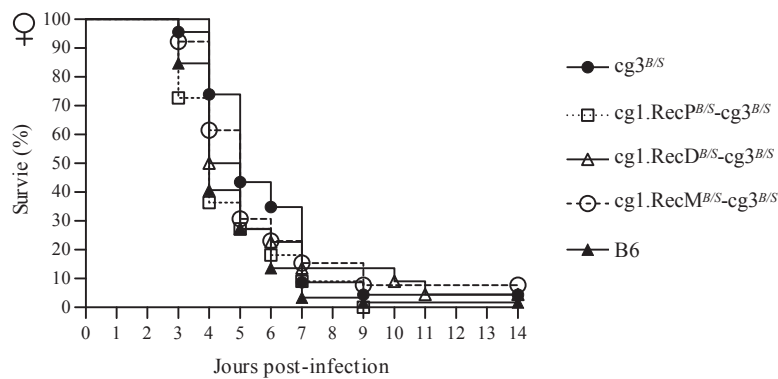


Figure 4. Courbe de survie des femelles $cg3^{B/S}$ ($n=23$), $cg1.RecP^{B/S}-cg3^{B/S}$ ($n=11$), $cg1.RecD^{B/S}-cg3^{B/S}$ ($n=22$), $cg1.RecM^{B/S}-cg3^{B/S}$ ($n=13$) et B6 ($n=59$). Aucune des lignées n'était différente de B6 que ce soit pour le taux de survie ou la durée de survie. Les trois lignées portant un segment sous-congénique de $cg1$ et le segment $cg3$ ne différaient pas non plus de la lignée $cg3$.

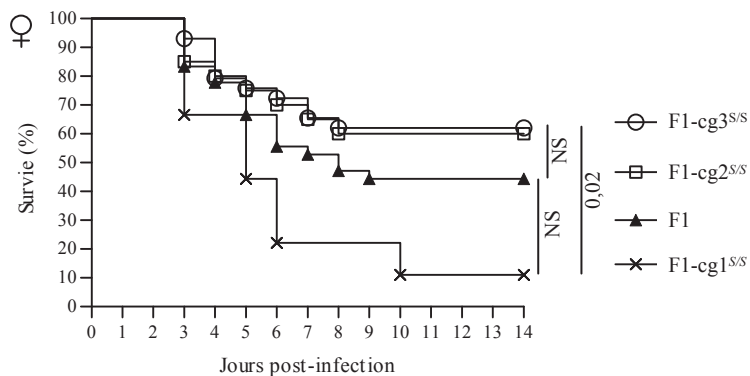
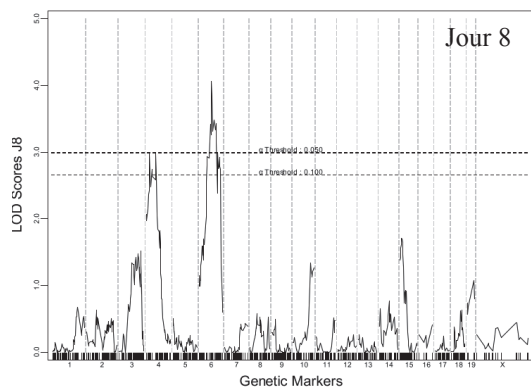
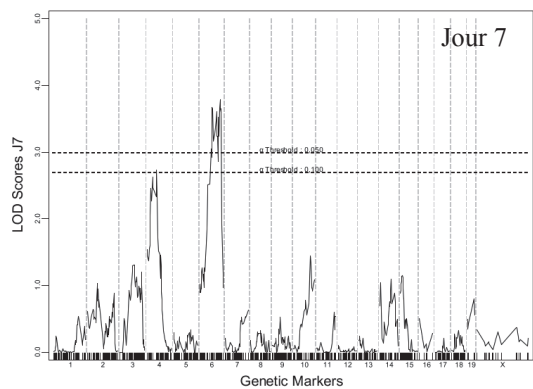
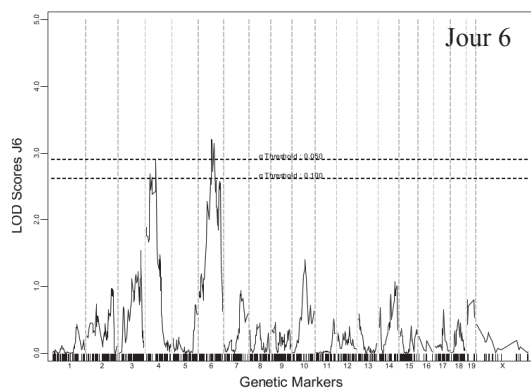
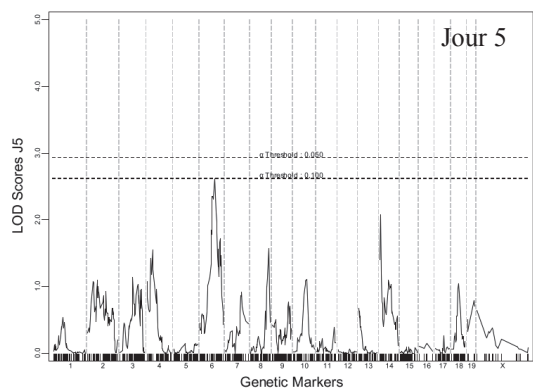
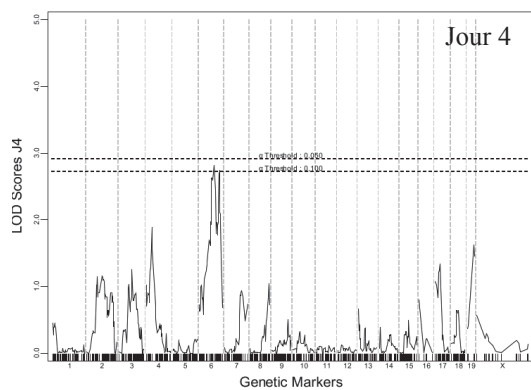
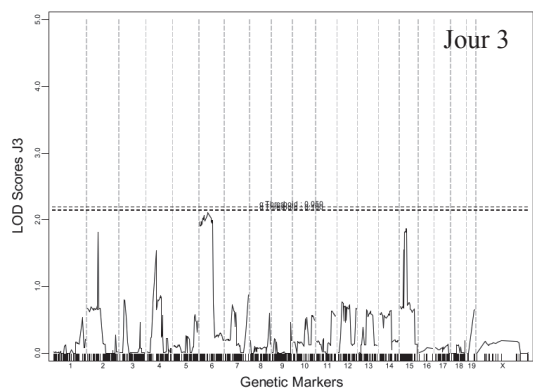


Figure 5. Courbes de survie des femelles F1- $cg1^{S/S}$ ($n=9$), F1- $cg2^{S/S}$ ($n=20$), F1- $cg3^{S/S}$ ($n=29$) et F1 ($n=36$). Le taux de survie des femelles F1- $cg1^{S/S}$ n'était pas significativement différent de celui des femelles F1 ($P=0,12$), mais significativement plus faible que celui des lignées F1- $cg2^{S/S}$ et F1- $cg3^{S/S}$ ($P<0,02$).



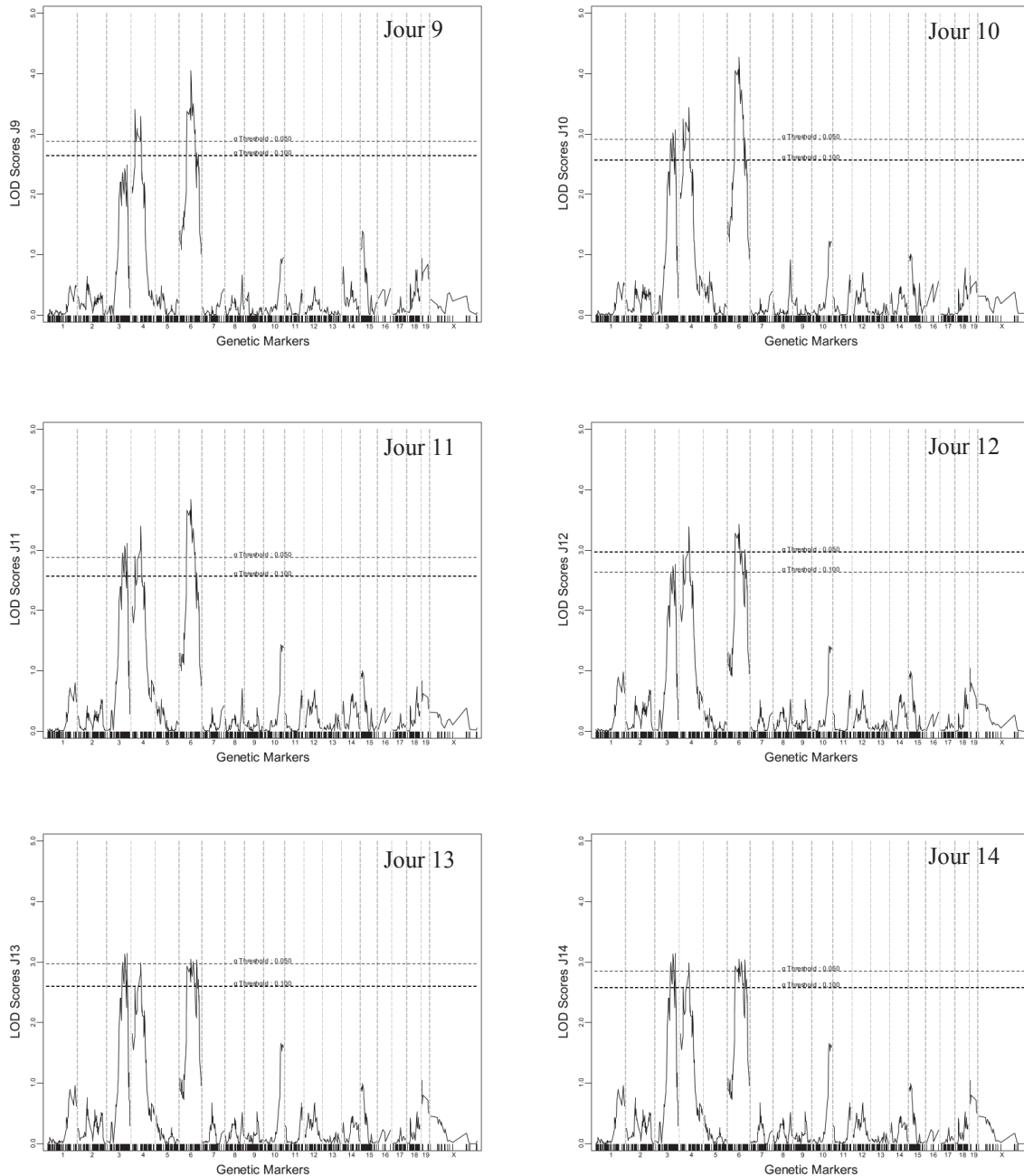


Figure 6. Recherche de QTL associés à la survie sur les 322 femelles BSB, du 3^{ème} au 14^{ème} jour post-infection. Les QTL localisés sur les chromosomes 4 (*Ypr12*) et 6 (*Ypr13*) atteignent le seuil de significativité $P = 0.05$ au 6^{ème} jour. Le QTL sur le chromosome 3 (*Ypr11*) atteint le seuil de significativité $P = 0.05$ au 10^{ème} jour post-infection. Aucun autre QTL n'a atteint le seuil de significativité dans cette analyse. Les lignes horizontales pointillées représentent les seuils de significativité pour $P = 0,1$ et $P = 0,05$.

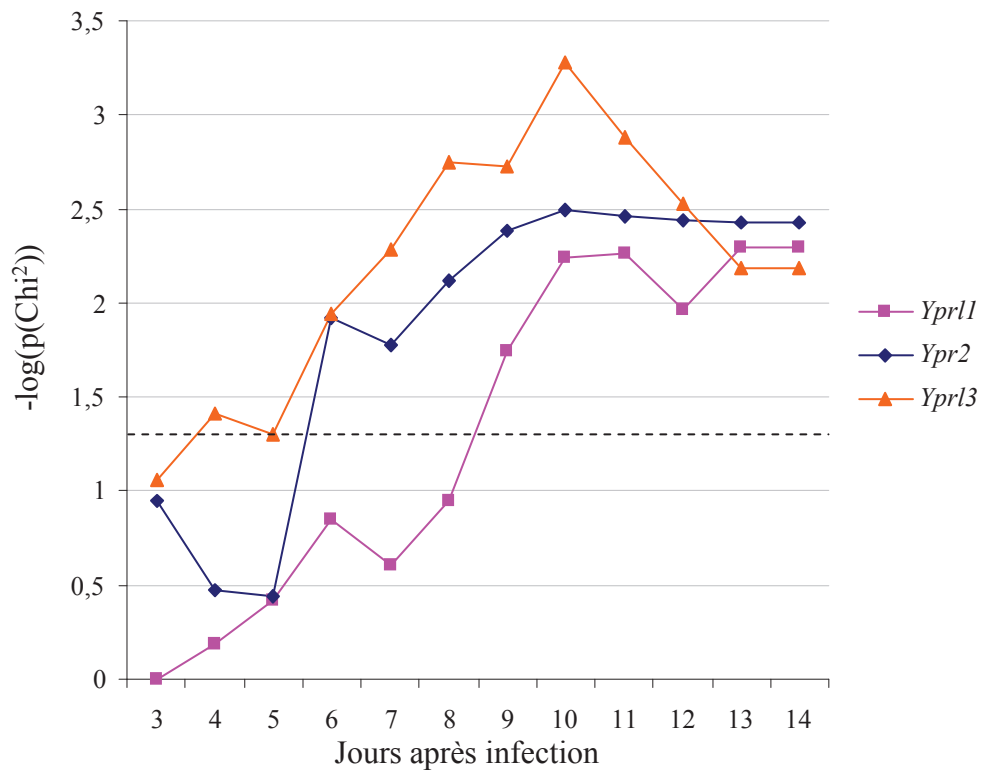


Figure 7. Significativité de chacun des trois QTLs chez les 322 femelles BSB, en fonction du jour d'analyse. Le test réalisé (Chi2) compare la proportion de survivantes dans les souris B/S et B/B au QTL. L'axe des ordonnées indique la valeur de $-\log(p(\text{Chi}^2))$. La ligne horizontale pointillée représente le seuil de significativité à 5%. Le QTL *Yprl3* franchit ce seuil au 6^{ème} jour de même que *Yprl2*, alors que *Yprl1* ne devient significatif qu'au 9^{ème} jour.

Figures « Résultats »

Partie 2

Figures « Discussion »

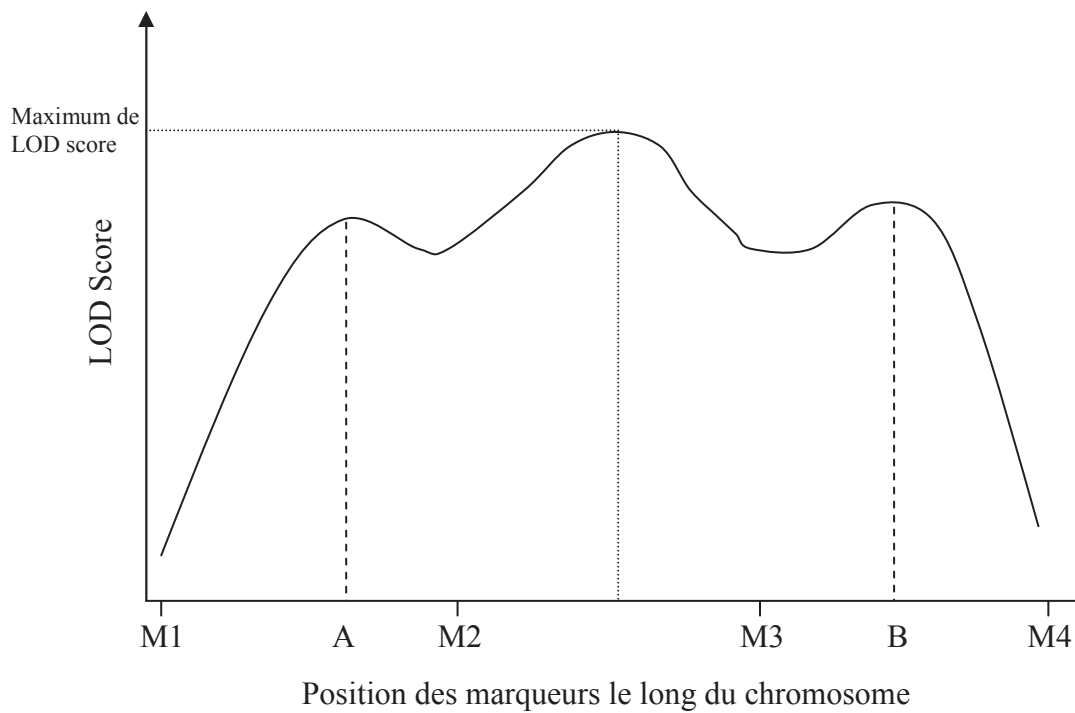


Figure 7. Détection d'un seul « dôme » sur la courbe de LOD score résultant de deux QTL proches. Lorsque deux QTL A et B sont localisés dans deux intervalles disjoints M1-M2 et M3-M4. Ils additionnent leurs effets de sorte que le LOD score maximum se situe à une position intermédiaire. On peut alors ne détecter qu'un seul QTL dans l'intervalle M2-M3, alors qu'il y en a en réalité deux sous-jacents (d'après la thèse de Marie Abitbol).

Recherche de facteurs génétiques contrôlant la résistance de lignées de souris consanguines à une infection expérimentale par *Yersinia pestis*, l'agent de la peste.

Yersinia pestis, l'agent de la peste, est une bactérie à Gram-négatif classée comme agent pathogène ré-émergent et potentielle arme de bioterrorisme. De plus, l'apparition d'une souche multi-résistance de cette bactérie souligne la nécessité de mieux comprendre comment cette bactérie hyper-virulente interagit avec son hôte.

Afin d'identifier des facteurs génétiques de vulnérabilité à la peste, notre laboratoire travaille sur la réponse de souris résistantes versus sensibles à *Y. pestis*. Notre stratégie pour identifier les facteurs génétiques impliqués dans la résistance/sensibilité à la peste combine une approche de cartographie de QTL (*Quantitative Trait Locus*) et d'analyse d'expression génique.

Nous avons précédemment décrit la lignée SEG/Pas, issue de *Mus spretus*, comme la première résistante à une souche virulente de *Y. pestis*, alors que la plupart des lignées murines de laboratoire, telle que la lignée C57BL/6J, sont extrêmement sensibles à la bactérie. Des croisements entre SEG/Pas et C57BL/6J nous ont permis d'identifier trois QTL impliqués dans la résistance à *Y. pestis*, localisés sur les chromosomes 3, 4 et 6.

Deux des QTL (situés sur les chromosomes 4 et 6) ont pu être confirmés par l'analyse de lignées congéniques. Plus de 40 % des femelles bi-congéniques hétérozygotes pour ces deux QTL ont survécu à l'infection, alors que tous les témoins C57BL/6J ont succombé.

La dissection de ces deux QTL par l'analyse de lignées sous-congéniques, nous a permis d'affiner l'architecture génétique de la résistance à la peste chez SEG/Pas. Nous avons conclu qu'un minimum de quatre facteurs génétiques, au sein de ces deux QTL, sont nécessaires pour augmenter la résistance à *Y. pestis* chez la Souris. Cependant, la production de plusieurs lignées congéniques portant le QTL situé sur le chromosome 3, dont une lignée triple congénique, ne nous a pas permis de confirmer l'existence de ce QTL.

En parallèle de l'analyse génétique, nous avons déterminé que les macrophages de SEG/Pas et de C57BL/6J présentaient des caractéristiques différentes après exposition à *Y. pestis*. Une analyse différentielle du profil transcriptionnel des macrophages de ces deux lignées a été réalisée à l'aide de puces à ADN. Nos résultats montrent une forte activation de la production cytokinique dans les macrophages de SEG/Pas en réponse à la bactérie, activation qui n'est pas observée dans la lignée C57BL/6J. Ces résultats suggèrent que les souris SEG/Pas sont capables de mettre en place une réponse immune innée plus forte ou peut-être plus précoce que C57BL/6J.

Nous avons ensuite étudié par qRT-PCR l'expression en cinétique de 44 gènes dans des macrophages de SEG/Pas, C57BL/6J et des bi-congéniques portant les QTL sur les chromosomes 4 et 6. Cette étude nous a permis de confirmer que les souris SEG/Pas sont capables de mettre en place une forte réponse inflammatoire lors de l'infection. Cependant, aucune différence significative n'a été observée entre la lignée bicongénique et la lignée parentale C57BL/6J. D'autres expériences seront nécessaires afin de mieux comprendre les mécanismes biologiques impliqués dans la résistance intermédiaire de cette lignée.

La dissection génétique associée à l'analyse de l'expression génique de ces lignées résistante et sensible améliore notre compréhension de la réponse de l'hôte à *Y. pestis*.

Mots-clés : Peste ; Souris ; résistance ; *Yersinia pestis* ; lignée congénique ; analyse de transcriptome.

Identification of genetic factors involved in the resistance of inbred strains of mice to an experimental infection with *Yersinia pestis*, the plague agent.

Yersinia pestis, the agent of plague, is a deadly gram-negative bacterium classified as a re-emerging pathogen and class A biological weapon. The appearance of a multi-resistant strain highlights the need to better understand how this pathogen kills its host.

To identify genetic factors of host susceptibility to plague, our laboratory is investigating the response of resistant versus susceptible mice to *Y. pestis*. Our strategy to decipher genetic determinants involved in resistance to plague combines Quantitative Trait Loci (QTL) mapping with gene expression analysis.

We previously described the *Mus spretus*-derived SEG/Pas strain as the first to resist fully virulent *Y. pestis*, while most inbred strains, such as C57BL/6, are highly susceptible. Crosses between these two strains identified three QTLs (located on chromosome 3, 4 and 6) contributing to resistance.

Two of the QTLs (on chromosome 4 and 6) were confirmed through creation of congenic mice. Up to 40% of the congenic mice heterozygous at these two QTLs, on a C57BL/6J background, survived the infection while all C57BL/6J mice died.

Further dissection of these two QTLs, through the use of subcongenic strains, enabled us to refine the genetic architecture of resistance to plague in SEG/Pas mice. We concluded that a minimum of four genetic factors, within these two QTLs, are required to increased resistance to *Y. pestis* in mice. Despite production of numerous congenic strains, including triple congenic mice, we were not able to confirm the existence of the third QTL identified on chromosome 3.

In parallel to genetic studies, we determined that SEG/Pas and C57BL/6J macrophages exhibit distinct characteristics upon *in vitro* exposure to *Y. pestis*. The underlying molecular differences were investigated by using microarrays. Our results show strong activation of cytokines in SEG/Pas macrophages in response to *Y. pestis*, which is not found in C57BL/6J macrophages. These results suggest that SEG/Pas mice are able to better activate innate immune response to *Y. pestis* than C57BL/6J mice.

We further studied the expression of 44 genes in a kinetic study on macrophages *in vitro* of SEG/Pas, C57BL/6J and bicongenic mice (carrying QTLs on chromosome 4 and 6). This study confirmed that SEG/Pas mice are able to build a stronger inflammatory response at early time of infection. Nevertheless no significant differences were observed in the bicongenic strain compared to C57BL/6J. Further studies will be required to understand the mechanisms involved in the intermediate resistance of this strain.

This combination of genetic dissection and gene expression analysis of resistant and susceptible mouse strains will enhance our ability to better understand the host response to plague.

Key words: Plague; mouse; resistance; *Yersinia pestis*; congenic strain; transcriptomic analysis.