

Université de Montréal

Identification et regroupement de QTL influençant la pression artérielle en modules épistatiques et analyse de deux gènes candidats chez la souche Dahl Salt-Sensitive

par

Cristina Chauvet

Programme de Biologie Moléculaire

Faculté de Médecine

Thèse présentée à la Faculté de Médecine

en vue de l'obtention du grade de Ph.D.

en Biologie Moléculaire

option générale

Mai, 2015

© Cristina Chauvet, 2015

Résumé

L'hypertension essentielle étant un facteur majeur de morbidité, la compréhension de son l'étiologie est prépondérante. Ainsi, la découverte de nouvelles composantes ou mécanismes de régulation de la PA par l'identification de QTL et l'étude de leurs interactions s'avère une approche prometteuse. L'utilisation de souches congéniques de rats pour l'étude de l'hypertension est une stratégie payante puisqu'elle permet de masquer les effets de l'environnement, tout en gardant le caractère polygénique de la PA.

Longtemps conçu comme un trait issu de l'accumulation des effets minimes des QTL, la PA est régulée par une architecture basée sur l'existence d'interactions épistatiques. L'analyse par paires de QTL individuels a permis d'établir une modularité dans l'organisation des QTL chez le rat Dahl Salt-sensitive en fonction de la présence ou de l'absence d'une interaction épistatique entre eux. Ainsi, deux modules épistatiques ont été établis; EM1 et EM2 où tous les QTL appartenant à EM1 sont épistatiques entre eux et agissent de façon additive avec les membres de EM2. Des hiérarchies dans la régulation peuvent alors être révélées si les QTL d'un même EM ont des effets opposés.

L'identification de la nature moléculaire des candidats *C18QTL4/Hdhd2* et *C18QTL3/Tcof1*, membres du EM1, et de l'interaction épistatique entre ces deux QTL, a permis, en plus, d'élucider une régulation séquentielle au sein du module. Hdhd2 pourrait agir en amont de Tcof1 et réguler ce dernier par une modification post-traductionnelle. Cette interaction est la première évidence expérimentale de la prédiction des relations entre QTL, phénomène établi par leur modularisation.

Le dévoilement du fonctionnement de l'architecture génétique à la base du contrôle de la PA et la découverte des gènes responsables des QTL permettrait d'élargir les cibles thérapeutiques et donc de développer des traitements antihypertenseurs plus efficaces.

Mots-clés : Hypertension, QTL, épistasie, modularisation, génétique épistatique, DSS, Hdhd2, Tcof1, Alpk2, gène candidat, module épistatique, congénique

Abstract

Essential hypertension is a major risk factor for cardiovascular diseases. Understanding the etiology of this pathology is of the utmost importance. Thus, unraveling novel genetic components and mechanisms regulating blood pressure (BP) via QTL identification and QTL-QTL interaction analysis is a promising strategy. Congenic strains establishment is a common and fruitful means for achieving such goal.

A quantitative trait, such as BP, has long been thought to result from the accumulation of infinitesimal effects exerted by multiple QTL. Nevertheless, BP is controlled by an epistasis-based architecture. Pair-wise comparisons of individual QTL based on the existence or lack-of epistatic interaction between them allowed us to establish a modularized organization of BP QTL in the Dahl Salt-sensitive model. Hence, two epistatic modules, namely EM1 and EM2 were constituted. In this fashion, any member of the EM1 is epistatic to all the other members of the same module and is additive to those of EM2. Regulatory hierarchies among BP with paradoxical effects can be revealed within each EM.

The molecular identification of EM1 members, *C18QTL4/Hdhd2* and *C18QTL3/Tcof1*, as well as the revelation of the molecular basis for their epistatic interaction enabled us to suggest a sequential regulation within this EM. *Hdhd2* could act upstream of *Tcof1* and regulate it by post-transcriptional modification. This interaction is the first experimental evidence derived from the predictive model of QTL modularization.

The elucidation of the molecular mechanisms underlying the genetic architecture of BP, as well as the identification of the causal genes for QTL will lay the grounds for expanding therapeutic targets and for developing more efficient antihypertensive treatments.

Keywords : Hypertension, QTL, epistasis, modularization, epistatic genetics, DSS, *Hdhd2*, *Tcof1*, *Alpk2*, epistatic module, congenics

Table des matières

Résumé.....	i
Abstract.....	ii
Table des matières.....	iii
Liste des tableaux.....	xi
Liste des figures	xii
Liste des sigles et abréviations.....	xiii
Remerciements.....	xix
Introduction.....	
1. La pression sanguine.....	2
1.1 Définition et mesures	2
1.2 Régulation de la pression artérielle.....	3
1.2.1 Régulation neuronale	3
1.2.1.1 Voies afférentes	3
1.2.1.2 Voies efférentes	4
1.2.1.2.1 Le SNPS.....	4
1.2.1.2.2 Le SNS.....	4
1.2.2 Régulation hormonale.....	5
1.2.2.1 Système rénine-angiotensine	5
1.2.2.1.1 Système rénine-angiotensine-aldostérone.....	5
1.2.2.1.2 Autres composantes du RAS	7
1.2.2.2 Autres voies de régulation endocrine.....	8
1.2.2.2.1 L'endothéline	8
1.2.2.2.2 Le monoxyde d'azote.....	9
1.2.2.2.3 L'adrénomedulline.....	9
1.2.2.2.4 Les kinines	10
1.2.2.2.5 Les prostaglandines.....	10

1.2.2.2.6 Les hormones natriurétiques	10
1.2.3 Le rôle des ions dans l'homéostasie de la pression artérielle	11
1.2.4 Caractère quantitatif et polygénique de la pression artérielle.....	12
2. L'hypertension	13
2.1 Définition et concepts épidémiologiques.....	13
2.2 La génétique de l'hypertension.....	14
2.2.1 L'hypertension secondaire.....	14
2.2.1.1 Formes Mendéliennes d'hypertension	15
2.2.1.1.1 L'aldostéronisme rémissible par les glucocorticoïdes.....	15
2.2.1.1.2 Le syndrome d'excès apparent en minéralocorticoïdes	16
2.2.1.1.3 Le syndrome d'hyperkaliémie et d'hypertension de Gordon	16
2.2.1.1.4 Le syndrome de Liddle	17
2.2.2 L'hypertension essentielle	17
2.2.2.1 Les facteurs génétiques.....	18
2.2.2.2 Les facteurs environnementaux	18
2.2.2.3 Les interactions gènes-environnement.....	18
3. Les modèles animaux.....	21
3.1 Utilisation de modèles murins pour l'étude de l'hypertension	21
3.2 Le modèle Dahl <i>Salt-Sensitive</i>	22
3.2.1 Pression artérielle.....	22
3.2.2 Fonction rénale.....	22
3.2.3 Balance des systèmes vasoconstricteur et vasodilatateur	23
4. Études génétiques et identification de QTL.....	24
4.1 Principes d'identification de QTL influençant la PA:	24
4.1.1 Définition de QTL.....	24
4.1.2 Approche gène candidat.....	24
4.1.3 Outils génétiques.....	25

4.1.3.1 Marqueurs génétiques	25
4.1.3.2 Analyses génétiques.....	26
4.1.3.2.1 Analyses de liaison génétique.....	26
4.1.3.2.2 Genome-wide association studies (GWAS).....	27
4.1.4 Utilisation du rat dans la découverte de QTL	28
4.1.4.1 Souches congéniques de rat	29
4.1.4.2 Sous souches congéniques	29
4.1.4.3. Doubles congéniques	31
4.1.5 Identification du gène et de la variante de séquence à la base du QTL	32
5. Interactions des QTL.....	33
5.1 Interactions alléliques et effets du fond génétique.....	33
5.2 Interactions géniques	34
5.2.1 Interactions additives	34
5.2.2 Interactions épistatiques.....	34
6. Régions QTL chez le rat chez.....	36
6.1 Identification de régions QTL par analyses de liaison génétique chez le rat Dahl Salt-Sensitive.....	36
6.2 Construction de congéniques pour cibler les régions QTL.....	37
6.3 QTL sur le chromosome 18	39
6.3.1 <i>ADRB2</i> et hypertension.....	39
6.3.2 <i>NEDD4L</i> et hypertension	40
7. Hypothèse et objectifs.....	41
Méthodes.....	43
8. Cartographie de QTL et mesures de pression artérielle.....	44

8.1 Marqueurs génétiques	44
8.2 Séquençage de génomes complets de DSS et Lewis	44
8.3 Mesures de pression artérielle.....	45
Publications scientifiques.....	47
9. Modularization and epistatic hierarchy determine homeostatic actions of multiple blood pressure quantitative trait loci	48
9.1 Abstract	49
9.2 Introduction.....	50
9.3 Materials and Methods.....	51
9.3.1 Animals	51
9.3.2 Construction of new congenic strains or substrains.....	51
9.3.3 Systematic assessment of QTL-QTL interactions by congenic combinations	52
9.3.4 Animal protocols and BP measurements	52
9.3.5 Statistical analysis.....	52
9.4 Results.....	53
9.4.1 Study design in grouping QTLs.....	53
9.4.2 Modularization of BP QTLs	54
9.4.3 Biological significance of QTL modularization	55
9.4.4 Implications of modularization for human QTLs	56
9.5 Discussion.....	57
9.5.1 Epistatic modularization, not quantitative accumulation, is the <i>modus operandi</i> in biological BP determination by QTLs.	57
9.5.2 Pathway of epistatic modules.....	59
9.6 Conclusion	60
9.7 References.....	61
9.8 Tables, figures and legends.....	65
9.10 Supplementary data.....	75
9.11 Addendum.....	96

10. Alpha kinase 2 is a novel candidate gene for inherited hypertension in Dahl rats	97
10.1 Abstract	98
10.2 Introduction.....	99
10.3 Methods.....	99
10.3.1 Animals	99
10.3.2 Construction of new congenic sub-strains	100
10.3.3 Animal protocols, BP phenotyping and statistical analyses	100
10.3.4 Mutation screening in <i>Adrb2</i> and <i>Nedd4l</i> genes	101
10.3.5 Analysis of gene expressions of <i>Adrb2</i> and <i>Nedd4l</i>	101
10.3.6 Sequencing of additional candidate genes	102
10.4. Results.....	102
10.4.1 QTL placements and assessment of QTL interactions	102
10.4.2 Evaluation of candidate genes <i>Adrb2</i> and <i>Nedd4l</i>	103
10.4.2.1 <i>Adrb2</i>	103
10.4.2.2 <i>Nedd4l</i> coding regions	103
10.4.2.3 <i>Nedd4l</i> exon-intron junctions.....	103
10.4.2.4 <i>Nedd4l</i> 5' upstream domain.....	104
10.4.2.5 Analysis of <i>Adrb2</i> and <i>Nedd4l</i> gene expressions.....	104
10.4.3 Alpha kinase 2 (<i>Alpk2</i>), not gastrin-releasing peptide (<i>Grp</i>) gene, as a leading candidate for C18QTL3	104
10.5 Discussion.....	105
10.5.1 Fine congenic resolution of BP QTLs	105
10.5.2 Alpha kinase 2 (<i>Alpk2</i>) gene is a standout candidate to be <i>C18QTL3</i>	106
10.6 Conclusion	108
10.7 Reference List	108
10.8 Tables, figures and legends.....	111
10.9 Supplementary data.....	116
10.10 Addendum.....	166

11. Two candidate genes for 2 quantitative trait loci epistatically attenuate hypertension in a novel pathway	167
11.1 Abstract	168
11.2 Introduction	169
11.3 Methods	170
11.3.1 Hypothesis and Study design	170
11.3.2 Animals	170
11.3.2 Whole genome sequencing of DSS and Lewis rats	170
11.3.3 In vitro function analysis of 2 isoforms for Haloacid dehalogenase-like hydrolase domain containing 2 (Hdhd2)	171
11.4 Results	172
11.4.2 Restricting chromosome intervals harboring BP QTLs	172
11.4.3 <i>Treacher Collins-Franceschetti syndrome 1</i> gene (<i>Tcof1</i>) is a unique function candidate for <i>C18QTL3</i>	172
11.4.4 <i>Hdhd2</i> is a foremost candidate for <i>C18QTL4</i>	174
11.4.5 Function predictions of missense mutations	174
11.4.6 Function analysis of <i>Hdhd2</i>	174
11.4.7 The Val113Leu substitution alters the phosphatase function of Hdhd2	175
11.4.8 Hdhd2 dephosphorylates Tcof1	175
11.5 Discussion	176
11.5.1 Elucidating the regulatory relationship in controlling BP between <i>Hdhd2</i> and <i>Tcof1</i> by synergizing epistasis and molecular mechanisms	176
11.5.2 <i>Hdhd2/C18QTL4</i> as a post-translational regulator of <i>Tcof1/C18QTL3</i>	177
11.5.3 <i>Tcof1/C18QTL3</i> may be a downstream target of <i>Hdhd2/C10QTL4</i>	178
11.5.4 Mechanistic connection to blood pressure regulation	179
11.5.5 The BP QTLs are not the prominent genes physiologically known to be involved in BP regulation	180
11.5.6 Implications in human essential hypertension	180
11.6 Conclusion	181
11.7 References	182

11.8 Tables, figures and legends.....	186
11.9 Supplementary data.....	197
11.10 Addendum.....	263
11.10.1 Résultats complémentaires.....	265
Discussion.....	267
12. Nouvelles découvertes dans l'architecture génétique de la pression artérielle chez le rat DSS.....	268
12.1 Évidence de l'existence d'interactions épistatiques des QTL.....	268
12.2 Classification des QTL par approche de génétique épistatique.....	270
12.2.1 Modularisation des QTL.....	270
12.2.2 Établissement d'une hiérarchie dans l'organisation des QTL.....	271
12.3 Signification biologique de la modularisation des QTL.....	274
13. Identification de QTL au niveau du chromosome 18 du rat.....	275
13.1 Analyse des régions C18QTL3 et C18QTL4.....	275
13.1.1 Cartographie fine des C18QTL.....	275
13.1.2 Identification de gènes candidats.....	276
13.1.2.1 QTL influençant la PA et gènes liés à sa régulation.....	278
13.2 Modularisation de <i>C18QTL3</i> et <i>C18QTL4</i> et étude de leur interaction.....	279
14. Implications et applications chez l'humain.....	282
14.1 Effets de l'hétérogénéité génétique sur l'identification de QTL.....	282
14.2 Relation entre les QTL influençant la PA chez l'humain et les QTL de PA chez DSS.....	283
14.3 Pertinence de la modularisation des QTL.....	284
Conclusion.....	285

Annexes.....	i
1. Analyse de l'activité d'une phosphatase et identification du substrat	i
Bibliographie.....	i

Liste des tableaux

Tableau I. Liste des régions QTL influençant la PA identifiés par analyse de liaison génétique dans des générations F2 et leurs effets.....	36
---	----

Liste des figures

Figure 1. Schéma simplifié du système rénine-angiotensine (RAS).	8
Figure 2. Schéma représentant la construction d'une lignée congénique	30
Figure 3. Carte de substitution	31
Figure A1. Résultats complémentaires pour l'activité enzymatique de Hdhd2.....	265
Figure A2. Détection par immunobuvardage de type Western de la déphosphorylation de Tcof1 par les isoformes de Hdhd2 et contrôles.	266
Figure 4. Schéma de la détermination hiérarchique de deux QTL par génétique épistatique.	272
Figure 5. Schéma de la modularisation des QTL.....	273
Figure 6. Analyse de l'interaction épistatique entre C18QTL3 et C18QTL4.....	280
Figure 7. Schéma récapitulatif de modularisation des QTL et des hiérarchies existantes.....	281

Liste des sigles et abréviations

- 11 β HSD : 11 β -hydroxystéroïde déhydrogénase
- α ENaC : Gène de la sous-unité α du canal sodium épithélial
- AA : Acide arachidonique
- a.a : Acides aminés
- Abca8a : *ATP-binding cassette, subfamily A, member 8a*
- Ach : Acétylcholine
- ACE : Enzyme de conversion de l'angiotensine
- ACE2 : Enzyme de conversion de l'angiotensine 2
- ACTH : Hormone corticotrope
- ADD1 : Gène de l' α -adducin
- ADH : Hormone antidiurétique
- ADM : Adrénomedulline
- ADN : Acide désoxyribonucléique
- ADRB1 : β 1-adrénorécepteur
- ADRB2 : β 2-adrénorécepteur
- AGT: Angiotensinogène
- Alpk2 : Alpha-kinase 2
- AME : Syndrome d'excès apparent en minéralocorticoïdes
- ANGI : Angiotensine I
- ANGII : Angiotensine II
- ANGIII : Angiotensine III
- ANG(1-7) : Héptapéptide 1-7 de l'angiotensine
- ANP : Facteur natriurétique atrial
- ARN : Acide ribonucléique
- ARNm : Acide ribonucléique messenger
- Art. : Article
- AT1R : Récepteur de l'angiotensine II de type 1
- AT2R : Récepteur de l'angiotensine II de type 2
- BC : Croisement de retour (*back-cross*)

BNP : Facteur natriurétique de type B

bp. : paires de bases

β -ARs : Récepteurs β -adrénergiques

Ca^{2+} : Ion calcium

cAMP : Adénosine monophosphate cyclique

cGMP : Guanosine monophosphate cyclique

Chr. : Chromosome

Cl^- : Ion chlorure

cM : Centimorgan

CmQTL_n : QTL numéro *n* du chromosome C numéro *m*

CmS.Ln : Congénique sur fond DSS avec allèles Lewis numéro *n* du chromosome *m*

CNP : Facteur natriurétique de type C

CNV : *Copy number variation*

CO_2 : Dioxyde de carbone

COX2 : Cyclooxygénase 2

CRLR : *Calcitonin receptor like receptor*

CS : Stéroïde cardiaque

CSK : gène de c-src tyrosine kinase

CUL3 : Cullin 3

CV : Coefficient de variance

CYP11B1 : 11β -hydroxylase

CYP11B2 : Aldostérone synthase

DG : Diacylglycérol

DiFMUP : 6,8-difluoro-4-methylumbelliferyl phosphate

Dr. : Docteur

DSR : Rat *Dahl Salt-Resistant*

DSS : Rat *Dahl Salt-Sensitive*

ENaC : Canal sodium épithélial

EDN3 : Gène de l'endothéline 3

ELISA : *Enzyme-Linked Immunosorbent Assay*

EM : Module épistatique
eQTL : QTL d'expression
ET : Endothéline
ET_A: Récepteur de l'endothéline de type A
ET_B : Récepteur de l'endothéline de type B
Etc. : Et cætera
FHH : Hypertension hyperkaliémique familiale
Gdpd1 : *Glycerophosphodiester phosphodiesterase domaine containing 1*
GP : Peptide gastrique
GPCR : Récepteur couplé aux protéines G
GRA : Aldostéronisme rémissible par les glucocorticoïdes
H⁺ : Ion hydrogène
HCO₃⁻ : Bicarbonate
Hdhd2 : *Haloacid dehalogenase-like hydrolase domain containing 2*
IP₃ : Inositol triphosphate
JAG1: gène de Jagged1
K⁺ : Ion potassium
KLHL3 : Kelch-like 3
K_m : Constante de Michaelis-Menten
Lew : Rat Lewis
LH : Rat *Lyon Hypertensive*
LL : Allèles homozygotes provenant de la souche Lewis
LN : Rat *Lyon Normotensive*
MAF : *Minor allele frequency*
MAP : Pression artérielle moyenne (*mean arterial pressure*)
MAPK : *Mitogen-activated protein kinase*
Mb : Megabases
Mg²⁺ : Ion magnésium
Mn²⁺ : Ion manganèse
mmHg : Millimètre de mercure
MHS : Milan Hypertensive Strain

MNS : Milan Normotensive Strain
MR : Récepteurs des minéralocorticoïdes
Mtmr4 : *Myotbularin related protein 4*
Na⁺ : Ion sodium
NaCl : Chlorure de sodium
NCC : Cotransporteur NaCl sensible au thiazides
NCX1 : Échangeurs Na⁺/Ca²⁺
NEDD4L : *Neural precursor cell expressed, developmentally down-regulated 4-like, E3 ubiquitin protein ligase*
NHE3 : Échangeur sodium-hydrogène
NKCC2 : Cotransporteurs Na⁺/K⁺/2Cl⁻
NO : Monoxyde d'azote
NOS : Synthase du monoxyde d'azote
NPR-A : Récepteur du facteur natriurétique de type A
NPR-B : Récepteur du facteur natriurétique de type B
NPR3 : gène du récepteur du facteur natriurétique 3
O₂ : Oxygène
OMS : Organisation Mondiale de la Santé
³²P : Phosphore 32
PA : Pression artérielle
PAD : Pression artérielle diastolique
PAS : Pression artérielle systolique
PCR : *Polymerase chain reaction*
PGI₂ : Prostacycline
PHAI : Pseudohypoaldostéronisme de type II
PKC : Protéine kinase C
PLA₂ : Phospholipase A₂
PLC : Phospholipase C
PLD : Phospholipase D
pNPP : *p*-nitrophenyl phosphate
Ppm1e : *Protein phosphatase 1E*

Prr11 : *Proline rich 11*

QTL : *Quantitative trait loci*

RAS : Système rénine-angiotensine

RAAS : Système rénine-angiotensine-aldostérone

ROMK : Canal potassique médullaire externe rénal

RT-PCR : *Reverse transcription polymerase chain reaction*

RT-qPCR : *Real Time quantitative polymerase chain reaction*

SGK1 : Sérine/Thréonine kinase 1

SHR : *Spontaneously Hypertensive Rat*

SNA : Système nerveux autonome

SNC : Système nerveux central

SNPs : *Single nucleotide polymorphisms*

SNPS : Système nerveux parasympathique

SNS : Système nerveux sympathique

sQTL : QTL d'épissage

SS : Allèles homozygotes provenant de la souche *Dahl Salt-Sensitive*

TAL : Branche ascendante large

TMEM133 : gène de la protéine transmembranaire 133

VSMC : Cellules musculaires lisses vasculaires

WKY: Rat Wistar Kyoto

WNK : *With no lysine kinase*

WT : type sauvage (*wild type*)

A mi mamá, a mi papá, a Pato, a Sofía, a Laura, a Emilio

Remerciements

Ma, Pa, Pato, Laura, muchísimas gracias por todo, todo, todísimo. Por apoyarme todo el tiempo, por creer en mí, por estar ahí siempre, por sus consejos, por consentirme, por apapacharme, por estar contentos conmigo, por enseñarme tantas cosas. ¡¡Los quiero!!

Sofía, por empezar a existir. Aquí te presento a tu prima mayor.

Litel one, gracias por estar conmigo estos últimos años, por apapacharme, por soportar tanto invierno conmigo, por los empujones y levantadas de ánimo, por chutarte todas mis historias, por tus consejos, por las cenas y las chelas, los juegos, las pláticas intensas. Por ser mi hermano menor. ¡Te quiero Litel!

Gracias a toda mi familia. A mi agüe, a mi abuelita, a mi abuelo, a mi abuelito, a Marta, a Bety, a Sergio, a Monis, a Chimi, a Héctor, a Laura, a François, a Claude, a Nicky; por apoyarme siempre. A mis primos, que quiero mil ocho mil, a Marta Fabienne, a Deby, a Tan, a César, a Panchisca, al Dip, a la Chófora, a Vivit, a Aleus, a Fer, a Dany. A mis sobrinos y ahijados; a Natalia, a Santi, a Fátima, a Federico y a Tiago. A todos ustedes porque, aunque más lejos que otra cosa, siempre me siento acompañada. A Margarita y a Ade, por consentirme siempre, a Monis.

A LeF, por acompañarme todo este tiempo, por todas las pláticas, la ruleta los súper días. ¡Te quiero muchísimo!

A Poupó, por nunca dejarme dudar que iba a estar aquí, por siempre encouragearme, ¡Te quiero mucho!

Elisita, Tasnuf et Soraya, je vous quière mouchissime et je vous extragne!

A Kimberlesa, guapa, gracias por tu amistad, por hacer de los últimos 7 años algo tan agradable. Por toda tu ayuda y por acompañarme en esta etapa.

Vika et à Estelle, pour les fous rires et les moments incroyables au labo (et en dehors). Dudesses, je vous aime et vous allez me manquer énormément!

Sari, pour son soutien, ses conseils, les histoires, les encouragements pre- et post-Ph.D... on se verra derrière un tas de papiers!

Henry, Cécile, Daniel, et tous les collègues du labo pour les moments incroyables.

Yann, pour les encouragements, le soutien (technique et *éléphantaire*), les plans, les nains et les paresseux.

A José, sin tu ayuda y las quedadas tarde en el lab, este proyecto no estaría terminado. Muchísimas gracias por todo.

Suzanne Cossette, pour son aide et ses conseils.

Dr. Richard Bertrand, pour son soutien et son écoute.

Julie, pour toutes les journées au labo, pour tout ce que tu m'as appris, tes conseils, les conversations, les encouragements, pour ton amitié.

Annie, pour les journées au labo, pour tout ce que tu m'as appris aussi, pour les encouragements, les dépannages dernière minute.

Merci à vous deux pour ces 8 années incroyables! Vous allez me manquer!

Dr. Deng, mon directeur de recherche. Merci!

La Fondation Canadienne des Maladies du Cœur, pour leur soutien financier.

Merci à vous tous,

Introduction

1. La pression sanguine

1.1 Définition et mesures

La pression sanguine ou pression artérielle (PA) est définie comme la force hydrostatique exercée par le sang sur la paroi vasculaire et son unité de mesure est le millimètre de mercure (mmHg).

Chez l'humain, chaque contraction cardiaque, appelée systole, propulse environ 70mL de sang vers les organes à travers le réseau artériel qui revient vers le cœur via le réseau veineux. Cette circulation exerce une pression sur la paroi des artères qui atteint son maximum lors du passage du sang, c'est la pression systolique (PAS) dont la valeur optimale se situe autour des 120mmHg. La pression diastolique (PAD) est définie comme la pression entre deux contractions ventriculaires et est à sa valeur optimale aux alentours de 80mmHg. La pression artérielle moyenne (MAP) est, comme son nom l'indique, la moyenne pondérée des PAD et PAS et est l'indicateur de la pression à laquelle sont soumis les vaisseaux continuellement. Ses valeurs devraient se situer entre 70 et 110 mmHg. Elle se calcule comme suit :

$$MAP = [(2 \times PAD) + PAS] / 3$$

La PA dépend principalement du débit cardiaque et de la résistance vasculaire. En effet, le débit cardiaque est dépendant de la fréquence cardiaque et du débit systolique ainsi que du volume sanguin total. La résistance vasculaire dépend du rayon des vaisseaux et de la viscosité du sang [1-3].

1.2 Régulation de la pression artérielle

La régulation de la PA est assurée par un système complexe de contrôle du débit et de la fréquence cardiaques, du volume sanguin ainsi que de la résistance vasculaire. Cette régulation est à la fois neuronale et hormonale; la première agissant surtout à court-terme, la deuxième engagée pour un contrôle à plus long-terme grâce à l'intermédiaire de la régulation du volume sanguin par les reins [1, 2].

1.2.1 Régulation neuronale

Le système nerveux joue un rôle majeur dans le contrôle de la PA. Son action au niveau du système cardiovasculaire requière de l'intégration par le système nerveux central (SNC) des signaux afférents provenant des différents organes ainsi que de différentes parties du cerveau. La réponse à ces signaux sera transmise par les voies efférentes du système nerveux autonome (SNA).

1.2.1.1 Voies afférentes

Le contrôle continu du système cardiovasculaire est assuré par l'hypothalamus ainsi que par deux types de récepteurs capables de détecter des changements au niveau de la PA : les barorécepteurs et les chimiorécepteurs. L'hypothalamus est la région du cerveau qui fait le pont entre le contrôle neuronal et endocrinien de l'homéostasie de l'organisme [4].

Les barorécepteurs, situés notamment au niveau du sinus carotidien et la crosse aortique, sont sensibles aux étirements des parois vasculaires. Les chimiorécepteurs sont situés à proximité des barorécepteurs et sont stimulés par des changements dans les niveaux de CO₂, O₂ et du pH dans le sang.

Les signaux afférents sont traités au niveau du bulbe rachidien et la réponse résulte en l'activation ou inactivation des voies efférentes du SNA, à savoir les systèmes nerveux sympathique (SNS) et parasympathique (SNPS) [5-7].

1.2.1.2 Voies efférentes

Les divisions sympathiques et parasympathiques du SNA sont responsables de la réponse neuronale aux changements de PA. Ces deux systèmes présentent deux types de neurones : un neurone préganglionnaire qui conduit le signal du CNS vers les ganglions et un neurone postganglionnaire chargé du relais du signal vers l'organe cible. Les neurones préganglionnaires des deux divisions sont cholinergiques, relâchant l'acétylcholine (ACh) comme neurotransmetteur, qui se lie aux récepteurs nicotiques des fibres postganglionnaires.

Les fibres postganglionnaires du SNPS, innervent le cœur et les vaisseaux sanguins, et leur contrôle est donc limité à la fonction cardiaque. En contrepartie, le SNS est responsable de l'innervation du cœur, des vaisseaux, des reins et des glandes surrénales.

La stimulation parasympathique a comme résultat prédominant une baisse de la PA, tandis que la stimulation sympathique est le plus souvent liée à une hausse de la pression sanguine [4, 5, 7].

1.2.1.2.1 *Le SNPS*

La stimulation parasympathique via le nerf vague entraîne une diminution de la fréquence cardiaque ainsi qu'une vasodilatation. Les neurones postganglionnaires du SNPS libèrent l'ACh, qui lie les récepteurs muscariniques des organes cibles [4, 7].

1.2.1.2.2 *Le SNS*

Le SNS est composé de neurones postganglionnaires de type adrénérique. Ainsi, la transmission du signal se traduit par la libération de la noradrénaline (ou norépinephrine) qui lie deux types de récepteurs adrénériques, soit les α et β . Le tonus sympathique provoque une augmentation de la fréquence cardiaque via les récepteurs β_1 -adrénériques, une vasoconstriction via les récepteurs du type α_1 et mène à une vasodilatation et à la sécrétion de la rénine via les récepteurs du type β_2 [8, 9].

Le SNS est soumis à son tour à de nombreux mécanismes d'activation ou d'inhibition ce qui permet d'exercer un contrôle rigoureux sur la PA. En effet, certains neurotransmetteurs

comme l'Ach, la sérotonine et la dopamine ainsi que certaines hormones comme les prostaglandines et l'histamine exercent une régulation négative tandis que la bradykinine, l'angiotensine II (ANGII), l'hormone adrénocorticotrope et la leptine sont impliquées dans la régulation positive de la neurotransmission par le SNS [10].

En plus de son action sur la fonction cardiaque et musculaire, le SNS est responsable d'une régulation indirecte et plus prolongée du système cardiovasculaire par l'activation de certaines voies humorales. En effet, l'innervation des glandes surrénales par les fibres du SNS conduit à la libération des catécholamines (épinéphrine et norépinéphrine) dans le sang. De même, les neurones qui innervent les cellules juxtaglomérulaires du rein stimulent l'activation du système rénine-angiotensine (RAS) [4, 5, 7-10].

1.2.2 Régulation hormonale

La régulation de la PA à long terme est assurée par les voies hormonales visant au maintien de l'homéostasie du volume sanguin et du système vasculaire, et ce grâce en grande partie aux systèmes rénaux.

1.2.2.1 Système rénine-angiotensine

Le système humoral régulant la PA le mieux caractérisé est sans doute le système rénine-angiotensine-aldostérone (RAAS). Cependant, la découverte d'autres composantes du RAS, telles que l'enzyme de conversion de l'angiotensine 2 (ACE2), ont permis d'établir l'existence d'une autre branche du RAS qui agit de façon locale au niveau de différents organes tels que le cœur, le rein, le cerveau, les poumons et les glandes surrénales [11, 12]. La Figure 1. (p.8) résume ce système.

1.2.2.1.1 Système rénine-angiotensine-aldostérone

L'activation du RAAS endocrine aboutit à l'augmentation de la PA par l'intermédiaire d'une hausse de la rétention d'eau et de sodium par les reins, ainsi que par la vasoconstriction et ceci grâce à l'action de l'ANGII et de l'aldostérone.

La génération de l'ANGII commence par le clivage l'angiotensinogène (AGT) plasmique par la rénine pour donner l'ANGI, un polypeptide de 10 a.a. L'AGT est une α -globuline constitutivement produite et sécrétée dans le sang par le foie qui, dans des conditions normales, existe sous des concentrations proches de son K_m . La rénine est une protéine synthétisée à partir du clivage de la prorénine, une pré-proenzyme constitutivement sécrétée dans le plasma. En effet, une baisse de PA, de la concentration de Na^+ et/ou du volume sanguin entraînent, via une stimulation sympathique, la sécrétion de la rénine par les cellules juxtaglomérulaires du rein. Il est à noter que la situation opposée, i.e, lors d'une hausse de la PA, la sécrétion de la rénine sera inhibée, entre autres par la présence du produit de la cascade, l'ANGII, permettant ainsi une régulation par rétroaction négative [12].

Lorsque le produit biologiquement inactif du clivage de l'AGT par la rénine, l'ANGI, parvient au niveau de la circulation pulmonaire il est clivé par l'enzyme de conversion de l'angiotensine (ACE) pour donner le polypeptide (8 a.a) actif ANGII. L'ACE est une enzyme sécrétée par les cellules endothéliales faisant partie des métalloprotéinases qui, en plus de son rôle dans le RAAS, est aussi, comme on le verra plus tard, une kininase responsable de l'inhibition de la bradykinine [12, 13].

L'action de l'ANGII se fait par l'intermédiaire de sa liaison à deux types de récepteurs, AT1R et AT2R, tous deux couplés à des protéines G (GPCRs). Chez les rongeurs, deux sous-types de récepteurs AT1R, AT1Ra et AT1Rb, ont été identifiés qui diffèrent au niveau de leur régulation et distribution dans les tissus, contrairement à l'humain où seulement un type est présent [14].

Au niveau moléculaire, lors de sa liaison aux récepteurs de type AT1R, l'ANGII active plusieurs cascades de signalisation comme celle de la phospholipase D (PLD), la phospholipase A2 et la phospholipase C (PLC/ IP_3 /DG) qui permet la mobilisation du Ca^{2+} intracellulaire résultant en la contraction des cellules musculaires lisses, la production d'acide arachidonique (AA), l'activation des MAP kinases (MAPK), ainsi que l'inhibition de l'adénylate cyclase. Au niveau physiologique, l'activation des AT1R au niveau du cerveau provoque la libération de l'hormone antidiurétique (ADH ou vasopressine) ainsi que de l'ACTH et stimule la voie sympathique qui, couplée à l'activation de ces mêmes récepteurs dans le système vasculaire, mène à une vasoconstriction. Au niveau rénal, la stimulation des

AT1R entraîne la rétention de Na^+ et d'eau par activation des canaux ioniques tels que les échangeurs Na^+/H^+ et les canaux sodium épithéliaux (ENaC). Plus encore, l'ANGII déclenche la sécrétion d'aldostérone par les surrénales, qui se lient à des récepteurs minéralocorticoïdes (MR) est responsable de l'activation de gènes encodant certaines protéines régulatrices des ENaC et des pompes Na^+/K^+ ATPase, et module ainsi la réabsorption du sel au niveau rénal [12, 14-18].

1.2.2.1.2 Autres composantes du RAS

Bien qu'il n'y ai pas un effet prouvé sur la PA, dans la plupart des cas, la liaison de l'ANGII aux récepteurs de type AT2R a des effets opposés à ceux de l'activation des AT1R. Ainsi, l'activation de ce type de récepteurs entraîne un effet vasodilatateur, une natriurèse accrue et l'activation la voie bradykinine-monoxyde d'azote (NO). Cependant, il a été récemment prouvé qu'au niveau rénal, ce serait l'activation des AT2R par l'ANGIII, un héptapeptide dérivé de l'ANGII, qui serait responsable de la natriurèse [14, 17].

De plus, il existe une autre voie, l'axe ACE2/ANG-(1-7)/Mas, qui aurait un rôle compensatoire par rapport à l'activation du RAAS classique. L'ANG(1-7) est un peptide résultant du clivage de l'ANGII par ACE2. Il avait été stipulé que la fonction primaire de cette molécule était d'activer la sécrétion de la vasopressine par l'hypophyse. Cependant, il a été récemment démontré que l'activation par ANG(1-7) du récepteur Mas induit une baisse dans la vasoconstriction, de la rétention d'eau et de sel en activant les voies bradykinine-NO, ce qui entre autres prévient les dommages rénaux. Ainsi, l'activation de l'axe ACE2/ANG-(1-7)/Mas contre les effets de la voie RAAS, entre autres par la dégradation de l'ANGII par ACE2, et c'est pour cette raison qu'elle est souvent appelée la branche protectrice du RAS [11, 12, 14-17, 19].

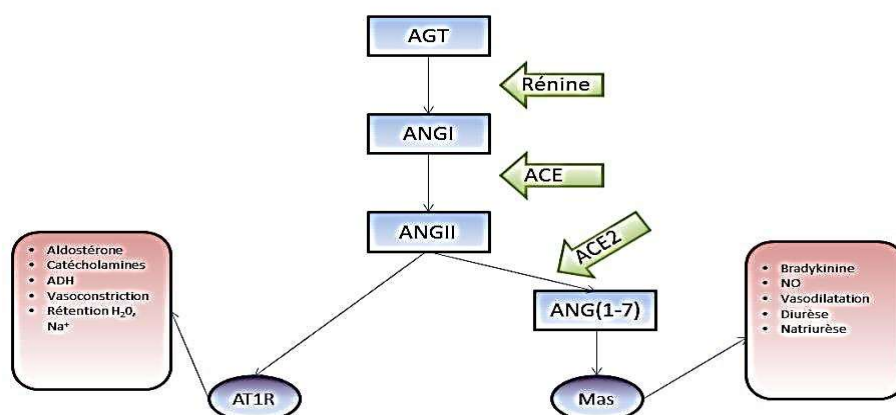


Figure1. Schéma simplifié du système rénine-angiotensine (RAS). L'angiotensinogène (AGT) est clivé par la rénine pour donner l'angiotensine I (ANGI). L'enzyme de conversion de l'angiotensine (ACE) clive l'ANGI pour donner l'angiotensine II (ANGII) qui se lie aux récepteurs AT1R induit une vasoconstriction et la sécrétion d'aldostérone et des catécholamines par les glandes surrénales ainsi que de la vasopressine (ADH) par l'hypophyse provoquant ainsi une augmentation du volume sanguin. L'ANGII est transformée en ANG(1-7) par l'enzyme de conversion de l'angiotensine 2 (ACE2). L'ANG(1-7) lie les récepteurs Mas et provoque une diurèse et natriurèse accrues, et active la voie bradykinine-NO provoquant ainsi une vasodilatation.

1.2.2.2 Autres voies de régulation endocrine

Il existe plusieurs autres facteurs responsables de la régulation de la pression artérielle. Il s'agit de facteurs qui peuvent agir au niveau du système vasculaire ou au niveau du contrôle de la volémie.

1.2.2.2.1 L'endothéline

Lors d'un stress mécanique ou d'une stimulation hormonale dus à une variation de la PA, les cellules endothéliales vont sécréter l'endothéline (ET). Il existe trois types d'endothéline, ET-1, ET-2 et ET-3, ET-1 étant le plus abondant.

Ces peptides de 21 a.a sont tous des vasoconstricteurs puissants et leur action physiologique majeure se traduit par une augmentation du tonus vasculaire et de la PA. ET-1 lie principalement deux types de récepteurs de la famille des GPRCs, soit ET_A et ET_B, qui se différencient par leurs propriétés moléculaires et pharmacologiques ainsi que par leur distribution tissulaire. La liaison de ET-1 aux récepteurs ET_A et ET_B présents au niveau des cellules musculaires lisses vasculaires (VSMC) entraîne une vasoconstriction et la croissance et adhésions cellulaires. Par ailleurs, la liaison de ET-1 aux récepteurs ET_B des cellules endothéliales produit la libération de la prostacycline et du NO ainsi qu'une vasodilatation subséquente [20, 21].

1.2.2.2.2 Le monoxyde d'azote

Les cellules endothéliales produisent le monoxyde d'azote (NO), une substance vasoactive gazeuse via la conversion de la L-arginine par la synthase du monoxyde d'azote (NOS) en L-citrulline. Il existe trois isoformes de NOS, soit NOS1 ou nNOS présente au niveau des cellules neuronales, NOS2 ou iNOS activée par les cytokines et NOS3 ou eNOS présente dans les cellules endothéliales; les isoformes 1 et 3 étant constitutivement exprimées. La diffusion du NO vers les VSMC adjacentes provoque la stimulation de la guanylate cyclase pour augmenter la production intracellulaire du guanosine monophosphate cyclique (cGMP). Parmi les effets de cette cascade, on trouve la vasodilatation au niveau des VSMC, ainsi qu'une inhibition de la réabsorption du NaCl et une action natriurétique au niveau du rein [22, 23].

1.2.2.2.3 L'adrénomedulline

L'adrénomedulline (ADM) est une hormone circulante excrétée par les cellules de la paroi vasculaire suite à une stimulation par l'ANGII, la norépinephrine et la bradykinine. Sa liaison aux récepteurs CRLR (calcitonin receptor like receptor) provoque, par la production de cAMP, une vasodilatation, une stimulation de la production de NO, une inhibition de la sécrétion d'ACTH et d'aldostérone, ainsi qu'une diurèse et natriurèse accrues. En plus de son action au niveau du système cardiovasculaire, l'ADM est aussi impliquée dans la réponse

inflammatoire et l'immunité ainsi que dans la régulation du taux de glucose dans le sang [24-26].

1.2.2.2.4 Les kinines

Les kinines sont des peptides formés à partir du clivage des kininogènes par les kallikréines. Ces peptides sont de puissants vasodilatateurs et ont un rôle protecteur du système vasculaire opposé au RAS. La kallikréine plasmatique est responsable de la production de la bradykinine, peptide responsable de la libération de NO par les cellules endothéliales, ainsi que de la perméabilité capillaire. La bradykinine lie deux types de récepteurs les B₁ et les B₂. Il s'agit de récepteurs couplés aux protéines G, qui sont constitutivement exprimés dans la presque totalité des tissus dans les cas de récepteurs B₂, et transitoirement lors de l'apparition de signaux pro-inflammatoires dans le cas des B₁. Les kinines jouent aussi un rôle important dans le processus inflammatoire et dans la stimulation des récepteurs de la douleur. Ces peptides sont dégradés par les kininases, dont l'ACE fait partie [13, 27].

1.2.2.2.5 Les prostaglandines

Les prostaglandines, dont la prostacycline (PGI₂) fait partie, sont synthétisées à partir d'acides gras insaturés dérivés de l'AA appelés eicosanoïdes par la voie de la cyclooxygénase 2 (COX2). La prostacycline est impliquée dans la vasodilatation, la diminution de l'activité sympathique, l'excrétion d'eau et de sel ainsi que dans la sécrétion de rénine dans le rein [28].

1.2.2.2.6 Les hormones natriurétiques

Les hormones natriurétiques peuvent être divisées en trois groupes : les peptides natriurétiques (NP), les peptides gastriques (GP) et les stéroïdes cardiaques (CS). La famille des NP inclut le facteur natriurétique atrial (ANP), et le facteur de type B (BNP) qui sont tous deux produits au niveau cardiaque. ANP lie les récepteurs NPR-A et NPR-B ayant un domaine guanylate cyclase intracellulaire qui catalyse la formation de cGMP dont les cibles sont les kinases cGMP-dépendantes ainsi que certains canaux ioniques. En plus de son effet

vasodilatateur, ANP stimule la diurèse et natriurèse et contre les actions du RAS en inhibant la sécrétion de rénine par le rein et de l'aldostérone par les glandes surrénales.

Les GP sont produits au niveau de l'intestin à partir de pré-prohormones suite à un apport alimentaire en sodium. Lors de la liaison des GP à leurs récepteurs, la production de cGMP provoque une excrétion de biocarbonate et d'ions chlorure dans l'intestin et inhibe la réabsorption des ions sodium. Dans le rein, elle est responsable d'une augmentation de la diurèse et natriurèse.

Les CS sont des molécules stéroïdiennes, dont la ouabaine fait partie, qui lient et inhibent les pompes Na^+/K^+ ATPase provoquant ainsi une natriurèse accrue. Cependant, en faibles doses continues, la ouabaine peut avoir un effet vasoconstricteur et ce en inhibant la production de NO par l'endothélium [29-31].

1.2.3 Le rôle des ions dans l'homéostasie de la pression artérielle

La régulation à long terme de la PA est étroitement liée au maintien de la volémie, résultant de l'équilibre dans le transport ionique. Il a été stipulé que le rein est l'organe le plus important dans la régulation des électrolytes, et par conséquent de la PA [32].

Lors de la filtration glomérulaire, le sodium diffuse librement et c'est un système d'échangeurs, transporteurs et canaux ioniques qui seront responsables de la réabsorption au long du néphron du 99% du Na^+ filtré. La majorité de cette réabsorption est effectuée par les échangeurs sodium-hydrogène (NHE3) situés dans le tubule proximal du rein. Les pompes Na^+/K^+ ATPases sont aussi importantes dans le maintien de la balance ionique intracellulaire responsable de l'efflux d'autres ions et d'eau au niveau rénal. Au niveau de la branche ascendante large (TAL), le sodium est réabsorbé par les cotransporteurs $\text{Na}^+/\text{K}^+/\text{2Cl}^-$ (NKCC2) qui opère en synergie avec les canaux ROMK. Les cotransporteurs NaCl sensibles aux thiazides (NCC) sont à la base du retour de Na^+ au niveau du tubule distal, tandis que ce sont les ENaC qui régulent le réglage fin de cette réabsorption. En effet, l'activité de ces canaux est assurée notamment par leur taux d'expression, des interactions protéine/protéine et des mécanismes de modifications moléculaires dont les réactions de phosphorylation et

ubiquitination font partie. Plus encore, l'ANGII, l'aldostérone, l'ADH et les catécholamines sont des facteurs déterminants dans la fonction de ces canaux ioniques [1, 31, 33-35].

Lors d'un apport alimentaire riche en sel, le rein ne parvient pas à excréter le Na^+ excédentaire ce qui provoque, dans certains cas, une rétention d'eau et une augmentation subséquente de la PA [31]. Cependant, ce panorama n'est pas toujours vrai et il existe des cas où la pression reste inchangée quelle que soit l'apport exogène en sel. Ces observations ont mené au concept de sensibilité au sel. Les mécanismes de la sensibilité ou résistance au sel sont mal compris, mais des altérations de la fonction rénale et des systèmes endocrinien et sympathique semblent jouer un rôle important. Plus encore, il a été suggéré qu'un apport important en sel induit la vasoconstriction via l'intermédiaire de la ouabaine endogène sur les échangeurs $\text{Na}^+/\text{Ca}^{2+}$ (NCX1) provoquant ainsi une résistance vasculaire accrue [36-39].

1.2.4 Caractère quantitatif et polygénique de la pression artérielle

Comme exposé dans la discussion précédente, la PA est un trait qui varie continuellement et dont la régulation étroite est assurée par plusieurs voies interconnectées. Le maintien de cet équilibre est le résultat de l'interaction et de la fonctionnalité des différentes composantes de ces cascades. Il s'en suit que la PA soit un trait phénotypique quantitatif à caractère polygénique. Par conséquent, la perturbation d'une des composantes de ces voies résulte en un dérèglement de la PA et à l'hypertension.

2. L'hypertension

2.1 Définition et concepts épidémiologiques

L'hypertension est définie comme une élévation persistante de la PA. Les mesures de pression sanguine (PAS/PAD en mmHg) dites optimales se situant en dessous des 120/80 mmHg, les individus qui présentent des mesures supérieures ou égales aux 140/90 mmHg sont considérés comme hypertendus [40].

L'hypertension artérielle est le facteur de risque le plus fréquent (environ 50%) dans le développement des maladies cardiovasculaires. Ainsi, l'impact majeur de cette pathologie se traduit par des lésions aux organes cibles provoquant des accidents vasculaires cérébraux, des anévrismes, des insuffisances cardiaques, des infarctus du myocarde et des maladies rénales. Les statistiques démontrent que les maladies cardiovasculaires expliquent 20% des causes de mortalité à l'échelle de la planète et que l'hypertension touche près d'un milliard de personnes à l'échelle mondiale et affecte de façon non proportionnelle différents groupes démographiques, sociaux et ethniques. Pour les services de santé publique ceci se traduit par un fardeau budgétaire d'environ 3.7 milliards de dollars, selon les données de l'OMS. Cependant, étant donné qu'il s'agit d'une maladie le plus souvent asymptomatique, seulement 70% des patients sont conscients de leur maladie, 60% sont traités et uniquement 35% sont contrôlés. Ainsi, des efforts majeurs sont réalisés pour combattre cette maladie [40, 41].

En effet, il a été démontré que des changements dans les habitudes de vie, tels qu'un mode de vie actif, une réduction dans la consommation de sel, de graisses, d'alcool et de tabac peuvent réduire les PAS et PAD. Il existe aussi plusieurs agents pharmacologiques comme les diurétiques, les vasodilatateurs directs, certains inhibiteurs du RAS ou les bloqueurs des récepteurs β -adrénergiques qui pris seuls ou en combinaison peuvent réduire la PA. Néanmoins, les traitements ne sont pas toujours efficaces et l'étiologie de l'hypertension demeure largement méconnue. Pour cette raison, l'étude de cette maladie est d'une importance majeure [40, 42-45]

2.2 La génétique de l'hypertension

L'hypertension est une maladie complexe résultant de l'interaction de facteurs environnementaux, démographiques et génétiques et ontogéniques. 95% des cas d'hypertension sont d'origine multifactorielle dont l'étiologie est mal établie et cette forme de la maladie est désignée comme hypertension essentielle. Dans le 5% des cas restants, l'hypertension est la résultante d'autres pathologies dont la cause est identifiable [1, 2, 31, 35, 46].

2.2.1 L'hypertension secondaire

Les formes d'hypertension sous-jacentes à d'autres pathologies sont définies comme les formes d'hypertension secondaire. A la base de ce type d'hypertension on peut retrouver des maladies rénales, endocrines, ou encore d'origine toxique [47, 48].

Au niveau rénal, des anomalies dans le maintien de la balance du sodium et de la volémie et mènent à l'hypertension systémique. Parmi ces altérations rénales on retrouve les maladies du parenchyme, la maladie rénale polycystique, la sténose qui peut découler d'une athérosclérose de l'artère rénale ou encore d'anomalies fibro-musculaires [49].

Les causes principales de l'hypertension secondaire endocrine sont l'hyperaldostéronisme primaire, le syndrome de Cushing, les maladies thyroïdiennes et le phéochromocytome. L'hyperaldostéronisme primaire représente jusqu'à 10% des patients hypertendus et est la première cause de l'hypertension secondaire. Il se caractérise par une production excessive d'aldostérone pouvant être causée par un adénome du cortex surrénalien ou par une hyperplasie congénitale [50].

Le syndrome de Cushing est très rare et seuls de 5 à 25 cas par million par an sont diagnostiqués dans la population. Il résulte de la production accrue de cortisol par les surrénales causée soit par une tumeur au niveau de ces organes ou par une sécrétion excessive d'ACTH, résultant d'un adénome au niveau de l'hypophyse [51].

Le phéochromocytome est une tumeur des cellules chromaffines surrenaliennes assez rare (1% des cas d'hypertension) qui entraîne une sécrétion accrue des catécholamines [52].

En outre, certaines substances exogènes peuvent être à la base de l'hypertension secondaire. C'est le cas de certains médicaments comme les anti-inflammatoires non stéroïdiens, des corticostéroïdes, des contraceptifs oraux et de vasoconstricteurs nasaux. On retrouve aussi des substance comme le réglisse, la cocaïne, la caféine et l'alcool [53].

Il existe aussi des formes d'hypertension dues à des mutations au niveau d'un seul gène qui sont transmises de façon mendélienne; il s'agit des formes Mendéliennes d'hypertension.

2.2.1.1 Formes Mendéliennes d'hypertension

Ces formes d'hypertension sont attribuables à la mutation au niveau d'un seul gène et sont transmises de façon mendélienne au sein d'une famille. Il s'agit donc de formes monogéniques d'hypertension affectant presque invariablement les mécanismes de transport d'eau et de sel par les reins. Bien que les formes mendéliennes de l'hypertension soient plus rares et plus sévères que l'hypertension essentielle, leur étiologie a été largement étudiée dans le but de découvrir des pistes qui pourraient mener à la compréhension de celle de l'hypertension essentielle.

2.2.1.1.1 L'aldostéronisme rémissible par les glucocorticoïdes

L'aldostéronisme rémissible par les glucocorticoïdes (GRA) est une maladie autosomique dominante caractérisée par une synthèse élevée d'aldostérone malgré des faibles taux de rénine plasmatique, une hypokaliémie et une alcalose métabolique. Une duplication génique est à la base de cette forme d'hypertension et elle survient lors d'un *cross-over* inégal des gènes encodant l'aldostérone synthase (CYP11B2) et l'enzyme 11 β -hydroxylase (CYP11B1), impliquée dans la synthèse du cortisol. À noter que les gènes *CYP11B2* et *CYP11B1* situés au niveau du chr. 8 chez l'humain, comportent tous les deux 9 exons identiques à 95% dans leur séquence. Il en résulte un gène chimère contenant la région promotrice de *CYP11B1*, sous contrôle de l'ACTH, liée à la partie codante de l'aldostérone synthase. Le produit ainsi généré cause une expression ectopique d'aldostérone, liée avec la

production de cortisol par la zone fasciculée des surrénales. La production excessive d'aldostérone mène à une rétention d'eau, une augmentation du volume plasmatique et à l'hypertension. Ainsi, la suppression de la production des stéroïdes par l'administration de glucocorticoïdes exogènes supprime l'hypertension [1, 35, 54, 55].

2.2.1.1.2 Le syndrome d'excès apparent en minéralocorticoïdes

Les patients atteints du syndrome d'excès apparent en minéralocorticoïdes (AME) présentent une hypertension accompagnée d'hypokaliémie, d'alcalose métabolique, de l'inactivation de la rénine plasmatique et de très faibles taux d'aldostérone circulante. La PA peut être diminuée lors de l'administration d'antagonistes des MR, ce qui suggère qu'un minéralocorticoïde, autre que l'aldostérone est en circulation. En effet, les MR ont une affinité similaire pour le cortisol et l'aldostérone. Étant donné que le cortisol circule avec un ration 1000:1 par rapport à l'aldostérone, et que l'activation des MR est effectuée par cette dernière, l'AME est le résultat de l'inactivation de l'enzyme 11 β -hydroxystéroïde déhydrogénase (11 β HSD) responsable de la conversion de cortisol en cortisone. Cette molécule, contrairement au cortisol, n'active pas les MR, et les patients atteints de AME, maladie autosomique récessive, présentent des mutations dans le gène encodant 11 β HSD qui l'inactivent [1, 35, 54, 55].

2.2.1.1.3 Le syndrome d'hyperkaliémie et d'hypertension de Gordon

Le syndrome d'hyperkaliémie et d'hypertension de Gordon, aussi appelé hypertension hyperkaliémique familiale (FHHT) est essentiellement de type autosomique dominant. Les sujets atteints présentent une hyperkaliémie, hyperchlorémie, une acidose métabolique et sont hypertendus malgré le faible taux de rénine, d'aldostérone plasmatique et l'absence d'insuffisance rénale [56]. Le phénotype est réversible par l'administration de diurétiques thiazidiques et dans certains cas le syndrome découle de l'existence de mutations au niveau des sérine-thréonine kinases *WNK1* et *WNK4* [57]. En effet, les kinases WNK (With no Lysine Kinases) régulent le cotransporteur Na-Cl sensible au thiazides (NCC) exprimé au niveau du néphron distal. A noter que la forme de type sauvage (wt) de *WNK4* (wt $WNK4$) inhibe le cotransporteur NCC en diminuant son expression à la surface cellulaire. De même, cette protéine inhibe le canal ROMK responsable de l'excrétion du potassium. Par ailleurs,

wtWnk1, inhibe Wnk4 relâchant ainsi l'effet inhibiteur de ce dernier sur NCC. Les mutations de Wnk1 associées au PHAII sont notamment des délétions au niveau du premier intron qui entraînent une augmentation de l'expression de cette protéine. Plusieurs mutations ont été trouvées au niveau du gène encodant Wnk4 causant l'hypertension et l'hyperkaliémie. Ces mutations ont pour effet de supprimer l'effet inhibiteur de Wnk4 sur l'expression de NCC au niveau de la membrane et, parallèlement, elles inhibent d'avantage le canal ROMK [34, 54, 56, 57]. Plus récemment, il a été démontré que des mutations au niveau des gènes *KLHL3* et *CUL3* sont aussi à la base de cette maladie. En effet, l'expression de NCC au niveau membranaire est plus importante lorsqu'il y a une perte de fonction de KLHL3 suggérant ainsi une régulation de ces transporteurs dépendante de son ubiquitination par le complexe CUL3-RING-KLHL3 [58, 59].

2.2.1.1.4 Le syndrome de Liddle

Cette forme d'hypertension est de type dominant et se caractérise par une rétention de sodium accrue, des niveaux plasmatiques d'aldostérone bas, l'inactivation de la rénine plasmatique, hypokaliémie et une alcalose métabolique. Les patients atteints sont sensibles à l'amiloride étant donné que cette pathologie est causée par des mutations dans les sous-unités β et/ou γ des ENaC qui entraînent augmentation de l'activité de ces canaux. En effet, les ENaC sont composés de trois sous-unités homologues α , β et γ chacune comprenant un grand domaine extracellulaire, deux domaines transmembranaires et de courts domaines cytoplasmiques. Les motifs PY de ces sous-unités sont mutés ou délétés dans le syndrome de Liddle ce qui empêche l'ubiquitination par Nedd4 et l'internalisation subséquente des canaux. Il existe aussi des mutations aux niveau des ENaC entraînant une perte de fonction, et donc une hyponatrémie, perte d'eau et un hyperaldostéronisme subséquent [1, 54, 60].

2.2.2 L'hypertension essentielle

Contrairement à l'hypertension monogénique, l'hypertension essentielle est d'origine multifactorielle et elle représente 95% d'hypertension. Elle résulte de l'interaction entre

plusieurs gènes, et est ainsi appelée hypertension polygénique et dépend aussi de facteurs environnementaux et démographiques.

2.2.2.1 Les facteurs génétiques

Il est bien connu de nos jours que les facteurs génétiques contribuent de 30% à 60% aux variations de PA. Plusieurs études d'agrégation familiale montrent une corrélation entre la similarité de PA et le lien de parenté. En effet, même s'ils sont séparés, les membres apparentés d'une famille conservent une fréquence similaire de cas d'hypertension. Ceci est attribué non seulement à l'héritabilité génétique, mais aussi à un mode de vie semblable. Cependant, il a aussi été remarqué que les conjoints et les cas d'adoption, qui partagent le même environnement mais dont le patrimoine génétique est différent, n'ont pas une tendance plus grande à développer la maladie. Plus encore, des études sur des jumeaux montrent une très grande corrélation parmi les jumeaux monozygotes, qui est moindre lorsqu'il s'agit de jumeaux dizygotes [1, 2, 61, 62].

2.2.2.2 Les facteurs environnementaux

L'apparition de l'hypertension essentielle est étroitement liée à des facteurs démographiques et des habitudes vie. En effet, il a été démontré que la prévalence de l'hypertension est plus importante dans les pays développés. Ceci est vrai entre les populations urbaines comparées aux populations rurales.

De plus, des facteurs comme l'âge, le niveau d'activité physique, le régime, le stress, la consommation d'alcool et le tabagisme sont aussi reliés à l'apparition de l'hypertension [1, 2, 44, 61, 63].

2.2.2.3 Les interactions gènes-environnement

Comme discuté dans la section 1.2.3, le sodium joue un rôle prépondérant dans l'homéostasie de la PA. En effet, plusieurs études ont montré une corrélation entre l'apport exogène en sodium et l'hypertension. Plus encore, l'importance des facteurs démographiques et ethniques jouent un rôle important dans l'apparition de la pathologie. Ainsi, certains haplotypes reliés au RAS seraient liés à l'hypertension dans les familles de descendance

africaine [46, 63]. Du point de vue évolutif, la sensibilité au sel dériverait de variantes génétiques prédisposant les populations vivant dans les climats chauds et arides à une rétention accrue en eau et en sodium, ce qui expliquerait une fréquence plus importante de ce phénotype chez les individus de descendance africaine [44]. De plus, il a été stipulé que les oestrogènes auraient un effet protecteur sur la sensibilité au sel chez les femmes et ce en augmentant la production de NO et en contrant l'activité du RAAS; plus encore, il a été démontré que la prévalence de l'hypertension sensible au sel chez les femmes est plus importante après la ménopause [64].

Par ailleurs, les apports nutritionnels ont une influence très marquée dans le développement de l'hypertension [38]. À titre d'exemple, une consommation riche en réglisse mime les symptômes du syndrome AME, puisqu'il inhibe l'enzyme 11 β HSD.

L'obésité, la résistance à l'insuline et le syndrome métabolique sont impliqués dans la pathogénèse de l'hypertension. Tous ces facteurs activent le RAAS. À noter qu'une la prévalence de l'hypertension chez les patients diabétiques de type II est de l'ordre de 80%. Plus encore, l'hypertension et le diabète sont les principales causes d'insuffisance rénale [65, 66]. En effet, des régimes riches en glucides (fructose et glucose) induisent une élévation de la PA et des symptômes du syndrome métabolique. Une des bases génétiques de cette observation seraient des variantes d'épissage du gène codant pour le transporteur de l'acide urique et du glucose/fructose GLUT9 associées à un niveau élevé d'acide urique sérique qui inhibe, entre autres la production de NO. De plus, l'insuline et des hauts niveaux de glucose sanguins activent le système sympathique provoquant une activation du RAAS et des récepteurs mineralocorticoïdes par le cortisol qui aboutit à une hausse de la PA [38, 44, 63, 65-67].

D'un autre côté, la théorie du «Thrifty phenotype» [68] propose que la programmation intra-utérine du génome, par l'intermédiaire de modifications épigénétiques, serait aussi à la base d'une prédisposition du nouveau-né à développer ou résister certaines pathologies. À titre d'exemple, une contrainte dans la croissance *in utero* causerait des anomalies dans le nombre et la taille des glomérules et donc dans le mécanisme de pression-natriurèse. Plus encore, il a

été démontré qu'il existe une corrélation inverse entre le poids à la naissance et l'incidence de maladies cardiovasculaires à l'âge adulte.

Les modifications épigénétiques (La méthylation de l'ADN, la modification des histones, les inactivations chromosomiques) ainsi que des facteurs comme les ARN interférence et les microARNs régulent l'expression des gènes sans changer la séquence d'ADN tout en étant stables et héréditaires [63]. Il a récemment été démontré que le niveau de méthylation de la séquence des gènes *11βHSD*, *ACE*, *ADD1*, *ADRB1* et *αENaC*, candidats à l'hypertension, corrèle de façon directe avec l'incidence de la maladie [69-72]. De plus, il a été suggéré qu'un régime riche en sel déclenche la suppression par l'intermédiaire des récepteurs β-adrénergiques de l'expression du gène de *WNK4* au niveau du rein. Cette régulation négative est due à la modulation du promoteur du gène par l'acétylation des histones dépendante de cAMP [73]. Actuellement, l'attention est portée sur la régulation de l'expression génétique par les microARNs, et il s'avère que plusieurs de ces molécules agissent sur des gènes impliqués dans le contrôle de la PA. À titre d'exemple, les microARN miR-584, -31, -181a et -663 sont impliqués dans la régulation de l'expression de certaines composantes du RAS. La voie de maturation des microARN implique les protéines Drosha, Dicer et RNase III; des études sur des modèles murins démontrent que les souris knock-out pour la protéine Dicer spécifique au muscles lisses ont une PAS plus basse que les contrôles qui ont, eux, une contractilité vasculaire diminuée [45, 63, 74].

Le caractère polygénique et l'existence des interactions à la base de l'hypertension essentielle en font une maladie complexe. Ceci, ajouté à l'hétérogénéité génétique et de mode de vie des populations humaines, pose un défi majeur dans l'étude et la compréhension de l'étiologie de l'hypertension. Ainsi, l'utilisation de modèles animaux pour l'identification des composantes génétiques de cette pathologie est une stratégie de choix.

3. Les modèles animaux

3.1 Utilisation de modèles murins pour l'étude de l'hypertension

Étant donné le caractère multifactoriel de l'hypertension essentielle, la complexité de son étude chez l'humain réside, entre autres, dans l'impossibilité de contrôler les facteurs environnementaux sur le trait étudié. Ainsi, pour contrecarrer le problème que pose l'hétérogénéité des populations humaines, des modèles murins, qui imitent certaines caractéristiques des traits complexes chez l'humain comme l'hypertension, ont été développés [75-77].

Les avantages des études génétiques sur les modèles murins consistent principalement à faciliter le contrôle de l'environnement, l'homogénéité génétique et la possibilité d'analyses sur une grande progéniture due à la facilité des croisements et à une courte durée gestationnelle. En ce qui concerne l'étude de l'hypertension, le rat est le modèle de choix; sa taille par rapport à la souris facilitant les études physiologiques [61, 78, 79].

De nos jours, nombreuses souches consanguines de rats pour l'étude de l'hypertension ont été créées. La création de lignées consanguines consiste à faire des croisements sélectifs d'individus portant le trait phénotypique d'intérêt jusqu'à ce qu'il soit fixé et à effectuer ensuite des croisements frère-sœur pendant au moins 20 générations pour obtenir l'hétérogénéité génétique. De nombreux contrôles normotendus ont été créés parallèlement aux modèles hypertendus en suivant cette méthode [75].

Les souches hypertendues peuvent être divisées en deux catégories; d'un côté il existe des modèles dont l'hypertension est spontanée et apparaît sans l'intervention des facteurs environnementaux. Les souches *Spontaneously Hypertensive Rat* (SHR)[80], *Lyon Hypertensive* (LH)[81] et *Milan Hypertensive* (MHS)[82] sont de bons exemples. D'un autre côté, on trouve les souches dont le phénotype est induit par un stimulus environnemental dont la souche *Dahl Salt-Sensitive* (DSS)[36] fait partie. En contrepartie, les contrôles normotendus le plus souvent utilisés sont *Milan Normotensive* (MNS)[82], *Dahl Salt-Resistant* (DSR)[36], *Wistar Kyoto* (WKY)[83], *Lewis* (Lew)[84] et *Lyon Normotensive* (LN)[81].

3.2 Le modèle Dahl *Salt-Sensitive*

Les souches DSS et DSR ont été développées en parallèle à partir de rats Sprague-Dawley par l'équipe du Dr Dahl, suivant les variations de PA observées lorsque les rats étaient soumis à un régime riche en sel. La création de ces souches repose donc sur leur sensibilité ou résistance au sel.

3.2.1 Pression artérielle

Contrairement aux rats DSR, les rats DSS soumis à un régime de 8% NaCl à partir de 3 semaines d'âge développent une hypertension sévère qui atteint les 200 mmHg et provoque la mort de l'animal au cours de 16 semaines. Cependant, si le régime riche en sel est administré à partir des 12 semaines d'âge, l'hypertension apparaît plus lentement et les mesures de PA restent autour des 185 mmHg. Même si les animaux sont soumis à un régime en sel à teneur normale (1% NaCl), ils développent l'hypertension graduellement.

Un régime riche en sel depuis les premières semaines mène à une volémie et un débit cardiaque accrus dont les valeurs retournent à la normale après 8 semaines, et c'est la résistance vasculaire qui fait que l'hypertension soit maintenue. Cependant, lorsque le régime en sel est normal (1% NaCl), la résistance vasculaire augmente sans que le débit et/ou la volémie soient affectés [36, 85, 86].

3.2.2 Fonction rénale

Les études effectuées par Dahl et ses collaborateurs ont démontré que la transplantation de reins de DSS chez DSR était suffisante pour induire une hypertension. L'opération inverse mène à une baisse de PA chez DSS. La souche DSS se caractérise par un mécanisme pression/natriurèse anormal. En effet, l'excrétion d'eau et de sodium chez cette souche est moindre que chez DSR, phénomène qui s'explique par une dérégulation dans la réabsorption des ions au niveau du néphron. Il a été démontré que lorsque le régime est riche en sel, il y a une augmentation de l'expression des ARNm des sous-unités des ENaC au niveau rénal. Plus encore, une augmentation de l'expression de SGK1, kinase responsable de l'activation des ENaC, a été remarquée. La régulation des ENaC et SGK1 induite par l'aldostérone diffère entre les deux souches. En effet, l'administration d'aldostérone diminue l'expression des

ARNm des sous-unités des ENaC chez DSR, contrairement à DSS. Par ailleurs, une activité accrue des cotransporteurs NKCC2 a été détectée chez les rats sensibles au sel. L'activité des pompes Na^+/K^+ ATPases est réduite au niveaux rénal et cardiaque. Cette souche démontre de plus une prédisposition au développement d'une insuffisance rénale [85-89].

3.2.3 Balance des systèmes vasoconstricteur et vasodilatateur

La contraction vasculaire dépend de la concentration intracellulaire de Ca^{2+} ; chez DSS, l'activité des échangeurs NCX1 est augmentée au niveau des vaisseaux. Par ailleurs, bien que les taux de rénine et d'ANGII circulants sont bas, l'activité du RAAS au niveau du cerveau et du rein est accrue. Cette souche démontre une activité sympathique accrue, une atténuation au niveau du baroreflexe et une production de NO diminuée. Plus encore, DSS présente une excrétion et une expression de kallikréine diminuées [22, 86, 90-93].

4. Études génétiques et identification de QTL

4.1 Principes d'identification de QTL influençant la PA:

Le défi majeur dans l'étude de l'étiologie et la pathogenèse de l'hypertension est l'identification des mécanismes qui la causent. Ainsi l'identification des composantes génétiques à la base de la maladie et la compréhension des interactions entre ces dernières sont d'une importance capitale.

4.1.1 Définition de QTL

Comme discuté auparavant, la PA varie continuellement et son homéostasie est sous contrôle de plusieurs gènes; ainsi, il s'agit d'un trait quantitatif et polygénique. Les loci de traits quantitatifs (QTL) sont définis comme unités génomiques responsables de la différence du phénotype observé entre deux souches contrastantes ou entre individus. Ainsi, un QTL influençant la PA serait un gène identifié au niveau d'une région génotypique définie par des analyses génétiques et dont l'expression influence la PA [77, 79, 94].

4.1.2 Approche gène candidat

L'approche gène candidat pour l'hypertension repose sur les bases biochimiques et physiologiques de la régulation de la PA. Ainsi, un gène est considéré comme candidat si sa fonction est liée au phénotype étudié ou s'il est localisé dans une région détectée par une analyse de balayage du génome (étude de liaison ou d'association). Cette approche sous-entend une connaissance préalable de la pathologie du trait en question et/ou de la fonction du gène candidat. Pour cette raison, dans l'étude de l'hypertension, les premières analyses ont

étés effectuées sur des gènes impliqués dans la régulation des systèmes vasculaires, cardiaque, nerveux et rénal.

Bien que cette approche s'est avérée très efficace dans l'étude des formes monogéniques de l'hypertension, elle comporte plusieurs désavantages. Comme discuté précédemment, elle est dépendante des connaissances de la maladie et ne permet pas toujours de découvrir de nouvelles composantes qui n'ont pas été associées au trait. Plus encore, le choix préalable du gène à étudier exclut des composantes pouvant se trouver en amont ou en aval de la cascade qui est analysée [43, 61, 77, 95, 96].

4.1.3 Outils génétiques

Étant donné les limitations de l'approche gène candidat dans l'étude de l'hypertension essentielle et dans le but d'identifier des QTL de PA de façon plus efficace et assertive, plusieurs outils génétiques sont utilisés.

4.1.3.1 Marqueurs génétiques

Le premier pas dans l'analyse génétique d'un trait complexe est l'établissement de marqueurs génétiques reliés au trait en question et dont la position au niveau chromosomique et l'héritabilité peuvent être retracées. L'ADN eucaryote présente plusieurs polymorphismes pouvant servir comme marqueurs génétiques.

Tout d'abord, il existe tout de courtes répétitions en tandem de 1 à 5 paires de bases répandues dans le génome. Ces séquences sont transmises entre générations et le nombre de répétitions peut varier d'un individu à l'autre permettant ainsi le génotypage par une simple réaction de PCR. Ce même type de génotypage est possible grâce à d'autres variantes structurelles de l'ADN parmi lesquelles on retrouve des insertions ou des délétions, ainsi que des duplications géniques.

Les polymorphismes d'un seul nucléotide (SNPs, pour *single nucleotide polymorphisms*) sont des substitutions d'une seule base et se présentent avec une fréquence d'environ 1/1000 bp. Le génotypage peut alors être effectué par séquençage de l'ADN

entourant le SNP, une simple réaction de restriction, l'utilisation de *SNP microarrays* (pouvant interroger < 5 000 000 de variantes à la fois), ou des balises moléculaires [41, 43, 97].

4.1.3.2 Analyses génétiques

La découverte de QTL influençant la PA par des analyses génétiques peut être effectuée au niveau pangénomique en utilisant des marqueurs génétiques étalés sur tout le génome ou par approche gène candidat qui, comme discuté auparavant, requière une connaissance préalable de la fonction du gène et de la pathologie étudiée [77].

4.1.3.2.1 Analyses de liaison génétique

Les analyses de co-ségrégation ont souvent été utilisées pour l'identification de QTL influençant la PA. Cette stratégie est basée sur la fréquence de recombinaison entre deux loci; ainsi, plus la distance chromosomique entre les deux est courte, plus la probabilité qu'il ségrégent ensemble lors de la méiose est grande. En d'autres termes, la probabilité de recombinaison entre un marqueur et le QTL d'intérêt est proportionnelle à la distance qui les sépare.

Les études de liaison génétique visent à établir s'il existe une co-ségrégation entre un marqueur génétique et la maladie étudiée lors de leur transmission d'une génération à la suivante. Elles reposent sur le principe que les allèles d'un QTL influençant un trait donné doivent être liés aux variations de ce dernier au sein de la population étudiée. Ainsi, ces études sont basées sur l'analyse des événements de recombinaison entre le locus causant la maladie dont la position génomique n'est pas établie et un marqueur de position connue, impliquant ainsi un lien physique entre les deux et requière donc d'une cartographie génétique. [77, 79]. De cette façon, les études de liaison reposent sur la probabilité d'avoir une co-ségrégation plus importante que celle pouvant être expliquée par le hasard. Le LOD score est la mesure de la probabilité de liaison entre le marqueur génétique et le locus causant le trait d'intérêt. Un LOD score est le logarithme du ratio de la probabilité que deux loci soient liés sur la probabilité que leur ségrégation soit purement due au hasard. Ainsi, un LOD score ≥ 3 est significatif et indique que deux loci sont 1000 fois plus susceptibles d'être liés que de ne pas l'être. Lorsqu'il s'agit de lignées consanguines, un LOD ≥ 2 est considéré comme suggestif

de la présence de liaison [98]. Ce type d'analyse est utilisé pour l'analyse des variantes génétiques au sein d'une famille et nécessite donc des informations génotypiques et phénotypes de plusieurs générations d'individus. Elle s'est avérée fort utile dans la détection de variantes rares dont l'effet sur le trait étudié est majeur. Bien que son application soit avantageuse dans la découverte de gènes à la base de l'hypertension chez des modèles animaux consanguins, elle comporte certains désavantages vu que la relation entre le génotype et le phénotype qu'elle suggère est de nature corrélative. De plus, cette approche permet d'identifier des régions chromosomiques souvent trop larges pouvant contenir des centaines de gènes. Ainsi, la présence de QTL influençant la PA déterminée par cette approche doit être confirmée et une cartographie fine de ce dernier doit être effectuée [43, 77, 79].

4.1.3.2.2 *Genome-wide association studies (GWAS)*

Les avancées technologiques permettent, de nos jours, de génotyper un nombre considérable de SNPs à des coûts raisonnables. Les GWAS permettent d'associer simultanément un grand nombre de variantes génétiques et un trait donné et visent à établir une corrélation entre ce dernier et les fréquences alléliques dans les populations humaines. Il s'agit donc d'une étude de cas-contrôles au sein d'une population où l'on compare la fréquence allélique des variantes entre les sujets présentant le phénotype d'intérêt et les témoins. Ainsi, si le marqueur étudié est à proximité d'un locus causant la maladie, la fréquence allélique observée devrait être significativement différente entre les cas et les témoins [41, 43, 99].

Étant donné le pouvoir statistique que confère l'association entre une variante génétique et un trait quantitatif dans un grand échantillon de la population, cette approche est devenue l'approche génétique la plus utilisée dans l'étude de l'hypertension essentielle. Jusqu'à date, la contribution la plus grande dans la découverte de nouveaux loci influençant la PAS et la PAD a permis d'identifier 29 SNP dans 28 loci associés à ces phénotypes, parmi lesquels 13 étaient déjà connus [100]. Plus encore, cette stratégie, étant effectuée à l'échelle du génome complet, permet d'identifier des loci pouvant ne pas être préalablement associés au phénotype étudié et dont l'effet est discret. Il est à noter que les variantes communes identifiées ont un impact <1mmHg sur la variance de la PA [43, 96]. Cependant, bien que

plusieurs loci aient été identifiés de cette façon, les désavantages majeurs de cette stratégie reposent sur le fait que les marqueurs génétiques se trouvent souvent au niveau de régions intergéniques et le lien entre ces derniers et le locus responsable du phénotype reste obscure. Néanmoins, les analyses d'association permettent généralement une cartographie plus fine du locus causant le trait que les analyses de liaison (10-100Kb pour les GWAS vs. 2-10Mb pour les études d'association) [99]. De plus, étant donné le seuil élevé de fréquences alléliques ($MAF > 5\%$), certains loci peuvent passer inaperçus en utilisant cette approche. Ceci est particulièrement important étant donné que certaines variantes rares, pouvant expliquer une portion importante des maladies complexes, ne sont pas prises en compte lors des analyses par GWAS [101]. En outre, ce type d'étude est sensible à la taille et à la stratification de la population étudiée et nécessite un nombre considérable de marqueurs à tester [41, 43, 79, 96, 102]. Cependant, cette dernière approche permet de surmonter les contraintes liés au nombre élevé de faux-positifs détectés par GWAS résultant de la stratification des populations [101].

4.1.4 Utilisation du rat dans la découverte de QTL

Comme discuté auparavant, l'utilisation de modèles murins pour l'étude de l'hypertension est une stratégie payante. Chez le rat, la localisation de régions QTL par des études de liaison se fait par l'analyse d'une population significative (en général de 100 individus ou plus) au sein de laquelle les allèles d'un QTL ségrégent de façon mendélienne. Cette population est constituée d'individus de la génération F2 obtenue comme suit : le croisement d'une souche hypertendue (DSS, dans le cas de ce projet) et d'un contrôle normotendu donne naissance à des individus de la génération F1 dont le génotype provient, à parts égales des souches parentales. Les individus F1 ainsi obtenus sont croisés entre eux pour produire une génération F2. En étudiant le phénotype (PA) et le génotype au niveau des marqueurs d'intérêt chez les animaux de la génération F2 on peut établir un patron de co-ségrégation et effectuer ainsi une analyse de liaison génétique [79, 98].

Cependant, l'identification d'une région contenant un QTL qui influence la PA par une étude de liaison génétique nécessite d'une confirmation plus astringente. La localisation de l'intervalle en question résulte, le plus souvent, en une région trop grande pour justifier le gène

d'intérêt et l'analyse des données statistiques présente plusieurs problèmes pouvant donner lieu à des faux positifs [77]. Ainsi, des études physiologiques permettant une cartographie fine du QTL et l'établissement d'une relation de cause à effet doivent être effectués. À cet effet, des souches congéniques de rats sont construites.

4.1.4.1 Souches congéniques de rat

Une souche congénique est obtenue par remplacement du fragment chromosomique d'intérêt chez la souche hypertendue (souche réceptrice) par son homologue provenant de la souche normotendue (souche donneuse). Dans ce cas, si la région chromosomique remplacée contient un QTL influençant la PA, on s'attend à une diminution significative de PA comparativement à la souche hypertendue parentale. La construction d'une souche congénique réciproque est possible; il s'agit alors de remplacer un segment chromosomique sur un fond génétique normotendu par son homologue provenant de la souche hypertendue. Dans ce cas, si la région contient un QTL, on observera une augmentation significative de PA par rapport à la souche normotendue. Les régions QTL ainsi obtenues sont donc susceptibles de contenir des gènes contrôlant la PA [77-79, 103].

La construction d'une souche congénique est resumée dans la Figure 2 (p.29). Le croisement d'une souche normotendue (donneuse) et d'une souche hypertendue (réceptrice) résulte en une génération F1 d'individus hétérozygotes dont le matériel génétique provient des deux parents en proportion égale. Un premier croisement de retour (BC1) est alors effectué entre un individu F1 et la souche parentale réceptrice. Le génotypage par marqueurs génétiques permet de tracer la transmission entre les générations de la région d'intérêt et de la cibler. Ainsi, les individus comportant l'allèle normotendue pour cette région sont recroisés avec la souche parentale pour donner la génération BC2 et ainsi de suite jusqu'à la génération BC8 où le fond génétique est 99% homozygote hypertendu, sauf pour la région ciblée. C'est alors qu'on effectue les croisements frère-sœur en vue d'obtenir un individu homozygote pour la région d'intérêt [76-79, 104].

4.1.4.2 Sous souches congéniques

La construction de sous souches congéniques permet de réduire la région d'intérêt en effectuant des croisements des individus de la souche congénique et de la souche réceptrice.

Ainsi, le fond génétique demeure homozygote hypertendu, et par recombinaison génétique au moment de la méiose, on peut réduire la région d'intérêt jusqu'à 1 ou 2 cM. Des croisements frère-sœur sont alors effectués pour avoir des individus homozygotes.

Suivant cette méthode, plusieurs sous souches peuvent être obtenues à partir de la même congénique. La PA de ces sous souches est mesurée et des cartes de de substitution peuvent ainsi être créées (Figure 3); certaines des souches ainsi obtenues conserveront le même effet sur la PA que la souche parentale, d'autres ne montreront pas de différence de PA

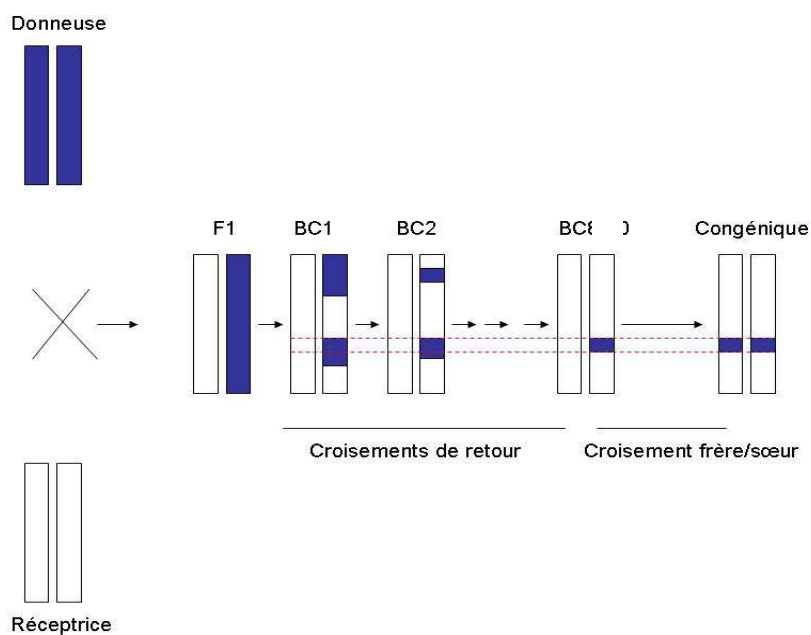


Figure 2. Schéma représentant la construction d'une lignée congénique. La figure ne représente qu'une seule paire de chromosomes. Le segment en pointillés représente la région ciblée lors des croisements de retour.

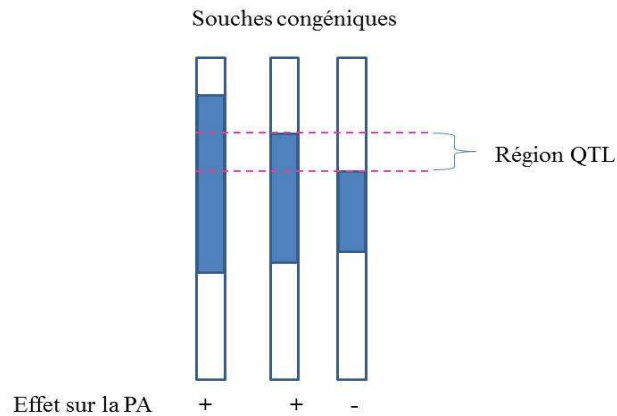


Figure 3. Carte de substitution. En blanc, les régions provenant de la souche receptrice hypertendue, en bleu la région remplacée par l'homologue normotendu. + représente une souche dont la PA est différente de la souche hypertendue, - représente une souche dont la PA n'est pas différente de la souche hypertendue

par rapport à celle-ci. Le QTL est donc localisé dans le segment conservé par les sous souches qui montrent un effet sur la PA et perdu par les sous souche qui ne démontrent plus de différence de PA avec la souche hypertendue [75, 76, 79].*

4.1.4.3. Doubles congéniques

La construction de souches « double congéniques » permet d'étudier les effets combinés de deux QTL distincts. À des fins d'exemplification, on prendra un QTL avec allèles normotendus NN sur le chromosome 1 et un QTL sur le chromosome 2, et ce sur un fond génétique hypertendu de génotype HH. Les congéniques délimitant ces deux QTL sont croisées entre elles et les individus de la génération F1 sont croisés de retour avec la congénique parentale pour le QTL du chr.1. Les individus issus de ce croisement de retour sont génotypés et ceux qui sont NN pour le QTL chr.1 et NH pour le QTL chr.2 sont croisés entre eux. La double congénique est établie lorsque les individus sont NN pour les deux régions QTL [105].

4.1.5 Identification du gène et de la variante de séquence à la base du QTL

Dans l'utilisation de souches congéniques, la recherche au niveau des séquences génomiques entre les souches hyper- et normotendues permet d'identifier le gène responsable du QTL. Le séquençage des régions codantes permet d'établir s'il existe des différences pouvant se traduire en un changement de la séquence protéique de la molécule encodée. Plus encore, le séquençage des régions introniques et intragéniques permet de détecter des différences qui pourraient avoir des effets sur l'expression et/ou l'épissage du gène en question. Les avancées technologiques permettent de séquencer des génomes ou des exomes complets en peu de temps et à des coûts accessibles, obtenant ainsi les SNPs et les variantes structurelles entre ces génomes. Le séquençage de l'ARN des souches étudiées, l'utilisation de *ARN microarrays* ou des analyses de RT-qPCR permettent d'analyser les niveaux d'expression des gènes candidats. Les analyses de *western blot* permettent de détecter des différences dans l'expression ou dans les modifications post-traductionnelles d'une protéine d'intérêt.

Lorsque des différences entre les souches hyper- et normotendues sont observées, il est possible de construire des animaux *knock-in*, *knock-out* (si viables) ou transgéniques en vue de confirmer que le gène en question est responsable de la variation du phénotype observée.

La découverte de la nature moléculaire de QTL peut, certainement, mener à la compréhension de son implication dans la régulation de la PA. Cependant, si à la base du QTL, il s'agit d'un gène de fonction inconnue, l'élucidation de sa contribution au contrôle de la PA devient plus facile si on connaît des éléments avec lesquels il interagit.

5. Interactions des QTL

L'hypertension essentielle étant un trait polygénique, l'identification de la nature moléculaire du QTL seule ne s'avère pas toujours suffisante. Ainsi l'élucidation des modes d'interaction entre les QTL dans le génome est d'une importance capitale.

5.1 Interactions alléliques et effets du fond génétique

Chaque gène peut avoir différentes formes au niveau du même locus, ce sont les allèles. Chaque allèle détermine un phénotype spécifique et un individu comporte 2 allèles pour chaque locus; par conséquent, le phénotype observé est la résultante de l'interaction de ces dernières [79].

Comme pour les formes monogéniques d'hypertension, il existe des QTL influençant la PA qui se comportent suivant une hérédité de type mendélienne [94]. L'analyse de souches congéniques et de des générations F2 issues de croisements entre des souches hyper- et normotendues a permis de démontrer qu'un seul QTL est capable d'entraîner des changements de PA par rapport à la souche parentale; ainsi, ils agissent de façon indépendante d'autres QTL et peuvent être dominants, codominants ou récessifs. À ce sujet, des études effectuées sur des congéniques hétérozygotes sur fond DSS ont montré qu'au niveau du chr.10, huit sur dix QTL démontraient un effet dominant de l'allèle normotendu (Lew). Plus encore, un QTL au niveau du chr.2, montre une dominance complète de l'allèle DSS, étant donné qu'une homozygotie Lew/Lew pour le QTL est nécessaire à la détection d'une variation de PA. Le C17QTL démontre une dominance incomplète étant donné que la magnitude de la différence de PA est dépendante du nombre de copies DSS [106].

Il est important de noter que, dans les études chez le rat, le choix du fond génétique (souche réceptrice) ainsi que celui dont proviennent les allèles introduits jouent un rôle déterminant dans la détection des QTL. En effet, des études de liaison génétique sur des générations F2 issues de croisements entre DSS et les souches MNS, WKY et SHR, ont permis de détecter un QTL uniquement lorsque les croisement étaient DSS x MNS et DSS x

WKY [22, 107]. Plus encore, lors de la construction de souches congéniques et sous congéniques, trois QTL distincts au niveau du chr.18 et un sur le chr.7 ont été détectés lorsque le fond génétique était homozygote DSS et où des allèles Lew/Lew ont été introduits; le résultat des mesures physiologiques se traduisant par une baisse de PA. Cependant, ces QTL n'ont pas été détectés lors de la construction des congéniques réciproques, i.e., allèles DSS pour la même région d'intérêt introduits dans un fond génétique Lewis. D'après ces observations, il a été postulé que le génome Lewis possède, à priori et contrairement à DSS, une capacité à moduler les variations de tension artérielle. [77, 94, 108-110].

5.2 Interactions géniques

5.2.1 Interactions additives

Lorsque la combinaison de deux QTL indépendants et distincts résulte en une variation de PA *grosso modo* équivalente à la somme de leur effet individuel, on peut dire qu'ils interagissent de façon additive. À titre d'exemple, des analyses de QTL au niveau du chr.10 chez le rat ont permis de démontrer l'existence de trois QTL agissant additivement [111]. Le même phénomène a été remarqué au niveau du chr.2 [112]. Lorsque ce type d'interaction se présente, on peut inférer que ces QTL agissent au niveau de différentes voies ou cascades de régulation et que, quand ils sont mis ensemble, leurs effets s'additionnent [94, 110].

5.2.2 Interactions épistatiques

L'épistasie est le phénomène par lequel l'expression des allèles d'un gène masquent les effets phénotypiques d'un autre gène distinct [113]. Les interactions épistatiques entre deux gènes ont été établies pour la première fois lors de l'analyse statistique d'une génération F2 pour laquelle les allèles normotendus du gène *Ace* inhibaient l'expression de celles du gène encodant NPR-A [107]. La preuve fonctionnelle de ce type d'interaction a été obtenue pour la première fois par la construction d'une « double congénique » comportant des QTL pour les

chr.2 et 10 dont les effets phénotypiques combinés étaient significativement moins importants que ceux que prédisait leur additivité mathématique [105]. Plus récemment, par la construction de sous souches congéniques, il a été démontré qu'une région soupçonnée de contenir un seul QTL, contenant en réalité deux QTL distincts dont la somme des effets ne dépasse pas celui de la région QTL originale [108, 114]. L'établissement d'une hiérarchie dans la relation est possible si les allèles de chaque QTL ont des effets opposés. Au niveau du chr.3 chez le rat, deux QTL ont été détectés sur fond génétique DSS avec allèles introduits provenant de Lew; un des QTL porte des allèles qui font baisser la pression (-PA) et l'autre des allèles qui la font augmenter (+PA). L'effet combiné de ces deux QTL est pratiquement égal à l'effet du QTL -PA seul [115]. Le fait que les effets de ces deux QTL ne s'annulent pas mutuellement, et que le phénotype soit semblable à celui observé chez la congénique portant le QTL -PA, indique que ce dernier masque la fonction du QTL +PA. Le QTL -PA est donc épistatique au QTL +PA. Plus encore, ces observations permettent, en partie, de faire le lien entre l'importance du fond génétique choisi et les interactions entre les QTL; en effet, bien que le génome Lewis comporte aussi des allèles menant à une hausse de la tension (QTL +PA), les relations entre les différents QTL de PA de cette souche lui permettent une régulation qui tend vers la normotension

La conclusion logique de l'existence d'interactions de type épistatique entre QTL se traduit comme suit : deux QTL participent à la même cascade si les effets de l'expression des allèles d'un locus masquent les effets des allèles de l'autre. [94, 110, 116].

6. Régions QTL chez le rat chez

De nombreuses études effectuées jusqu'à présent ont permis de détecter un nombre considérable de QTL ayant une influence sur la PA grâce à l'utilisation de modèles animaux [94, 103, 104, 106, 108, 111, 112, 114, 115, 117-128]. En effet, plus d'une soixantaine de QTL ont été identifiés de cette façon; parmi ceux-ci on retrouve *Add1*, *Cyp11b1*, *Rffl* et *Adamts16* [124, 129-131]. Plusieurs des QTL découverts grâce à l'utilisation de modèles animaux sont de nos jours utilisés lors des GWAS visant des gènes candidats, bien que la grande majorité de ces associations n'aient pas encore été confirmées [132, 133].

6.1 Identification de régions QTL par analyses de liaison génétique chez le rat Dahl Salt-Sensitive

Dans les années 1990, des analyses de liaison pangénomiques ont été effectuées sur des générations F2 issues de croisements de rats DSS et des contrôles normotendus montrant l'existence de plusieurs QTL influençant la PA, notamment au niveau des chromosomes 1, 2, 3, 5, 7, 8, 10, 16, 17 et 18 [103, 106, 107, 134]. Le tableau I résume les LOD scores et leurs effets sur la PA. Il est à noter qu'au niveau du chromosome 7, le LOD score est inférieur à 3. Cependant, l'existence d'un QTL de PA détecté dans une population F2 issue du croisement DSS x DSR [135] et la présence du gène candidat *Cyp11b* au niveau de cette région, la rendent intéressante.

Tableau I. Liste des régions QTL influençant la PA identifiés par analyse de liaison génétique dans des générations F2 et leurs effets

Chromosome	Souche normotendue	LOD score	Effet sur la PA (mmHg)	% de variance de PA
1	Lewis	3,0	+30,4	8,9

2	Lewis	2,9	-27,5	8,8
2	MNS	2,6	+13,9	8,0
3	Lewis	3,0	+2,8	8,9
5	Lewis	4,5	+31,4	13,2
7	Lewis	1,6	-18,1	4,8
8	Lewis	2,0	-19,7	6,4
10	Lewis	5,5	+35,4	17,8
16	Lewis	2,2	+25,5	6,7
17	Lewis	2,2	+18,4	6,7
18	Lewis	2,4	+24,8	7,5

Le tableau est adapté Duong *et al.* avec la permission de Macmillan Publishers Ltd: *Heredity*, [106] copyright 2006. Les effets sur la PA correspondent à la moyenne des mesures chez les rats ayant pour génotype les allèles SS au niveau du marqueur étudié moins la moyenne des mesures chez les rats ayant le génotype LL au niveau du même marqueur.

6.2 Construction de congéniques pour cibler les régions QTL

En vue de valider et de réduire les régions QTL ainsi détectées, plusieurs souches et sous souches congéniques ont été construites. De cette façon, le C7QTL a été défini, et, fait intéressant, aucune différence de séquence induisant un changement en acides aminés des protéines encodées par les gènes *Cyp11b1*, *Cyp11b2* et *Cyp11b3* n'a été détectée entre DSS et Lewis [109]. Il est à noter que *Cyp11b1* a été identifié comme étant le gène responsable d'un QTL sur le chromosome 7 entre les souches DSS et DSR [124, 135].

De même, par la construction de souches congéniques visant les régions détectées par analyses de liaison, trois QTL au niveau des chromosomes 1 et 3 [115, 120, 136], deux QTL sur les chromosomes 5, 8 et 17 [121, 137, 138] et le C16QTL [106] ont été identifiés.

Au niveau du chromosome 2, la création de congéniques issues de DSS et Lewis a montré l'existence d'un QTL au niveau de la région détectée par analyse de liaison [139]. Par ailleurs, la construction de souches congéniques issues de DSS et MNS ont permis d'établir 6 régions QTL distinctes s'étalant sur moins de 6Mb chacune. De plus, le séquençage des gènes contenus dans ces intervalles et connus pour avoir un effet sur la PA a montré qu'il n'existait pas de variantes entre les souches parentales se traduisant par des changements au niveau des protéines encodées [112, 140].

En vue de réduire la région QTL du chromosome 10 et de cibler les régions contenant les locus *Nos2* et *Ace*, candidats à l'hypertension, il a été démontré que 3 QTL de PA existaient au niveau de ce chromosome. Ces études ont permis d'exclure *Nos2* et *Ace* comme candidats [75, 111]. Des études ultérieures ont démontré l'existence d'un quatrième QTL grâce à la construction de sous souches congéniques. Au cours de cette analyse, il a été démontré que les C10QTL1, C10QTL3 et C10QTL4 étaient épistatiques et agissaient de façon additive avec le C10QTL2 [126]. La cartographie fine de ces QTL a démontré plus tard la région du C10QTL1 contenait en effet 2 QTL distincts. Ainsi le C10QTL5, interagissant épistatiquement avec les C10QTL1, 3 et 4, a été défini. Cette étude a aussi permis de réduire les régions de ces QTL à moins de 1Mb, rendant ainsi l'analyse des gènes contenus plus simple. À noter que ces QTL trouvent ses homologues au niveau du chr.17 chez l'humain où des analyses de liaison associent la région chromosomique à l'hypertension essentielle [114]. Comme il a été discuté, cinq régions QTL ont été détectées au cours de plusieurs analyses sur le chromosome 10. De par la nature des interactions qui existent entre ces QTL et étant donné la taille réduite de ces régions, les gènes contenus dans les C10QTL1, C10QTL2, C10QTL3 et C10QTL5 ont été séquencés et analysés en vue d'identifier les gènes causant la différence de PA entre DSS et les congéniques. Des gènes n'ayant pas été préalablement associés à l'hypertension ont été identifiés comme candidats pour chaque QTL [141].

6.3 QTL sur le chromosome 18

Comme discuté auparavant, les études de liaison génétique effectuées par Garrett et ses collaborateurs, ont montré l'existence d'un QTL de PA au niveau du chromosome 18. Ces analyses ont été effectuées en utilisant la souche Lewis comme contrôle normotendu étant donné sa grande résistance au sel par rapport à la souche DSR. De plus, les polymorphismes au niveau des marqueurs microsatellites entre DSS et Lewis est supérieur à ceux retrouvés entre DSS et DSR [103]. La construction de souches congéniques délimitant la région détectée par ces études a permis d'identifier l'existence 3 QTL influençant la PA à ce niveau. Plus précisément, la construction des souches C18S.L6, C18S.L7 et C18S.L8 ont permis d'identifier deux régions QTL, soient C18QTL1 et C18QTL2. La souche C18S.L2 délimitait le C18QTL3. Cependant, la taille de ces régions étant trop grande (>30Mb) [108], la construction de sous souches congéniques a permis de les réduire. La cartographie fine du C18QTL3 semblait particulièrement prometteuse étant donné que cette région contient deux gènes associés à l'hypertension; il s'agit du gène codant pour le récepteur β -adrénergique de type II (*ADRB2*) et du gène encodant l'ubiquitine ligase E3 *NEDD4L*. De par leur nature moléculaire et leurs fonctions physiologiques en plus de leur association avec la maladie chez l'humain, ces gènes son de bons candidats.

6.3.1 *ADRB2* et hypertension

Les récepteurs β -adrénergiques (β -ARs) jouent un rôle important dans la régulation de la PA. Le récepteur *ADRB2*, dont le gène ne possède pas d'introns, appartient au sous type β_2 des adrénorécepteurs, qui sont normalement exprimés au niveau des poumons, des reins, des vaisseaux et du foie. Ces récepteurs de type G sont couplés à l'activation de l'adénylate cyclase et à la formation de cAMP. L'activation des récepteurs β_2 par les catécholamines induit la relaxation des muscles lisses et la sécrétion de NO par les cellules endothéliales vasculaires. Chez l'humain, plusieurs polymorphismes de ce gène, parmi lesquels Arg16Gly et Gln27Glu, sont associés à l'hypertension. La substitution Arg16Gly cause une désensibilisation accrue du récepteur puisqu'elle modifie la régulation négative induite par

l'agoniste et affecte la cinétique de son activation; Gln27Glu est associée à une résistance à la désensibilisation [142-144].

6.3.2 NEDD4L et hypertension

NEDD4L est une ubiquitine ligase de type E3 possédant plusieurs domaines WW (motifs tryptophane) responsables de leur liaison aux motifs PY des ENaC et un domaine ubiquitine ligase HECT. L'association de NEDD4L aux ENaC induit leur ubiquitination et leur dégradation subséquente. Des mutations au niveau des motifs PY de ENaC empêchent la liaison de NEDD4L se traduisant par une hyperactivité des canaux et entraînent l'hypertension. NEDD4L est aussi impliqué dans la régulation des NCC. Différents polymorphismes au niveau de *NEDD4L* ont été détectés chez l'humain lors des études de population qui l'associent à l'hypertension. Parmi ces substitutions on retrouve rs4149601(G/A), au niveau de l'exon 1 qui introduit un site d'épissage alternatif résultant en une protéine dont le domaine de ciblage à la membrane est altéré. Plus encore, des études effectuées sur des souris knock-out ont prouvé que la perte de fonction de NEDD4L induit l'hypertension et la sensibilité au sel [145-149].

7. Hypothèse et objectifs

La prévalence élevée de l'hypertension essentielle rend l'étude de cette pathologie prépondérante. Étant donné que seul le 60% des patients atteints est traité et que moins du 35% de ces patients répondent efficacement aux traitements antihypertenseurs, l'identification de nouveaux gènes et mécanismes régulant la PA s'avère d'une importance capitale, et ce, en vue de dispenser des traitements individualisés plus efficaces [45].

La découverte de QTL de PA et de leur mode de fonctionnement au niveau génétique est considérée la façon la plus directe de comprendre les mécanismes à la base de l'hypertension essentielle. L'utilisation de modèles animaux, permettant de contrôler les effets de l'environnement, s'avère une stratégie de choix pour atteindre ce but.

Ainsi, l'identification de la nature moléculaire des QTL influençant la PA est un des principaux objectifs dans l'analyse de l'étiologie de cette maladie. De plus, s'agissant d'un trait polygénique, la découverte et la compréhension des interactions entre les QTL permettraient d'identifier de nouvelles voies de régulation qui pourraient être ciblées lors du choix du traitement à fournir.

Par ailleurs, il a longtemps été pensé qu'un trait quantitatif comme la PA est la résultante de l'action combinée de plusieurs QTL qui contribuent chacun à une variation minime de ce trait. Chez l'humain, aucun QTL à effet majeur n'a été encore identifié. Cependant, comme discuté auparavant, l'observation d'interactions épistatiques et le caractère Mendélien de certains QTL mène à stipuler qu'il existerait une certaine redondance dans les actions des QTL et que ceux-ci seraient organisés dans une architecture génétique hiérarchique permettant de réguler la PA. Ainsi, l'existence d'interactions épistatiques entre deux QTL permettrait de les regrouper dans une même voie de régulation. À l'opposé, l'absence de ce type d'interaction et/ou l'existence d'une interaction additive entre deux QTL révélerait qu'ils font partie de voies distinctes.

En vue de tester l'hypothèse selon laquelle les QTL pourraient être regroupés en différents modules épistatiques, des souches « double congéniques » ont été construites;

l'existence ou absence d'épistasie entre deux QTL dictant s'ils appartiennent au même module.

Bien que la découverte d'une architecture génétique des différents QTL se révèle un moyen efficace dans la compréhension de l'organisation et étiologie de l'hypertension essentielle, l'identification de la nature moléculaire du QTL demeure importante. Ainsi, la détection de variations de séquence ou d'expression du(des) gène(s) responsables du QTL et leur impact sur la fonction physiologique du produit, sont des composantes essentielles à l'étude de la pathologie et à la découverte de nouvelles cibles thérapeutiques et/ou de diagnostic.

Pour cette raison, nous avons décidé non seulement d'analyser les gènes candidats des QTL du chr.18, mais aussi, étant donnée l'épistasie qui existe entre eux, d'élucider leur interaction au niveau moléculaire, apportant ainsi une preuve physiologique à ce phénomène.

Méthodes

8. Cartographie de QTL et mesures de pression artérielle

8.1 Marqueurs génétiques

En vue de réduire les régions QTL du chromosome 18, plusieurs sous-souches congéniques ont été construites. De plus, la construction de souches « double congéniques » a permis d'analyser l'effet de la combinaison de deux QTL distincts. Le génotypage de ces souches a été effectué grâce à des marqueurs microsatellites déjà connus ou de nouveaux marqueurs ont été identifiés afin de délimiter avec plus de précision les nouvelles régions contenant un QTL. Les nouveaux marqueurs microsatellites D18Chm ainsi que plusieurs SNPs permettant le génotypage ont été trouvés à partir de la séquence du génome du rat (<http://www.ncbi.nlm.nih.gov/mapview>) ou par analyse des résultats du séquençage complet des génomes DSS et Lewis.

Les produits de PCR des marqueurs microsatellites ont été migrés sur gel d'agarose 4% pour tester leur polymorphisme. L'amplification par PCR des régions contenant des SNPs entre DSS et Lewis a été suivie d'une migration et purification. Le séquençage du produit ou une restriction enzymatique ont été alors réalisés.

8.2 Séquençage de génomes complets de DSS et Lewis

Le séquençage complet des génomes DSS et Lewis a été effectuée avec les technologies de séquençage Illumina, HiSeq - Illumina Paired-ends 100bp sequencing lane, fourni par le Centre d'innovation Génome Québec et Université McGill.

L'ADN des DSS et Lewis a été extrait et purifié à partir des reins de rats soumis au même protocole que les rats en télémétrie avant l'étape d'implantation des sondes (régime NaCl 0,2% pendant 21 jours suivi de 3 semaines à NaCl 2%). L'extraction et la purification de l'ADN ont été effectuées en utilisant le kit QIAamp DNA Mini Kit® et en suivant les instructions du fabricant.

Les génomes DSS et Lewis sont devenus notre base de données pour identifier des SNPs entre les souches.

Le positionnement des variantes a été possible grâce au développement par Max Chauvet du logiciel de bioinformatique NextiaGen®. Ce logiciel d'analyse massive de données permet de localiser, dans la séquence de chaque gène, les variantes identifiées par l'analyse de SNPs dans un ou plusieurs génomes. En effet, ce logiciel permet de placer et d'assigner une position dans la séquence CDS du transcrit le plus long indexé dans la base de données de *UCSC genome browser* (<https://genome.ucsc.edu/>) aux SNPs se trouvant dans les séquences des exons d'un gène. Avec l'appui des outils d'interprétation de séquences de bases disponibles sur le web, ce logiciel permet de déceler les modifications potentiellement importantes dans les séquences d'acides aminés résultantes. De par sa conception en base de données relationnelle, il peut traiter efficacement, en quelques minutes, des millions d'éléments d'information. Ceci permet la génération rapide de résultats, ainsi que leur mise périodique à partir de sources de référence modifiées en permanence au rythme de l'avancement de la recherche scientifique internationale.

8.3 Mesures de pression artérielle

Les protocoles de manipulation des animaux et de prises de mesures de PA approuvés par le comité de notre institution (CIPA) suivent les Lignes Directrices Canadiennes et sont les mêmes que décrit précédemment [108, 137]. Brièvement, une semaine après la naissance, les rats sont identifiés par le numéro correspondant (tag) et l'ADN est extrait du bout de la queue (clip); le génotypage subséquent est effectué. Les animaux sont maintenus sur un régime faible en sel (0.2% NaCl) jusqu'au sevrage le 21^{ème} jour après la naissance. Le régime à haute teneur en sel (2% NaCl) commence alors et dure trois semaines, jusqu'à l'implantation des sondes de télémétrie.

Les mesures de PA se font par télémétrie via le système Data Sciences Inc. (St-Paul, Minnesota, USA). Il s'agit d'une méthode précise qui permet de prendre des mesures de façon continue et directe à distance, diminuant ainsi les contraintes physiques pour l'animal et par

conséquent la possibilité de fausser la lecture de PA [150]. Le cathéter de la sonde est implanté dans l'artère fémorale. Au moins 10 jours après l'implantation sont alloués avant de débiter de la prise des mesures permettre une récupération post-opératoire. Les mesures sont transmises par la sonde via la radiotélémetrie à une plaque située sous chaque cage. La plaque est reliée à un ordinateur qui collecte pendant dix secondes à toutes les deux minutes les données de PAS, PAD et PAM pour une période de trois semaines. A la fin de cette période, les animaux sont sacrifiés à froid par guillotine et les organes, ainsi que la sonde implantée sont récupérés.

Publications scientifiques

9. Modularization and epistatic hierarchy determine homeostatic actions of multiple blood pressure quantitative trait loci

Authors: Cristina Chauvet¶, Kimberley Crespo¶, Annie Ménard, Julie Roy, Alan Y. Deng*

Centre de recherche, Centre hospitalier de l'Université de Montréal (CRCHUM) and, Département de médecine, Université de Montréal, Montréal, Québec, Canada

¶ These authors contributed equally to the current work

*Corresponding author: Alan Y. Deng

9.1 Abstract

Hypertension, the most frequently diagnosed clinical condition world-wide, predisposes individuals to morbidity and mortality, yet its underlying pathological etiologies are poorly understood. So far, a large number of quantitative trait loci (QTLs) have been identified in both humans and animal models, but how they function together in determining overall blood pressure (BP) in physiological settings is unknown. Here, we systematically and comprehensively performed pair-wise comparisons of individual QTLs to create a global picture of their functionality in an inbred rat model. Rather than each of numerous QTLs contributing to infinitesimal BP increments, a modularized pattern arises: 2 epistatic ‘blocks’ constitute basic functional ‘units’ for nearly all QTLs, designated as epistatic module 1 (EM1) and EM2. This modularization dictates the magnitude and scope of BP effects. Any member can contribute to BP additively to that of EM2, but not to those of the same module. Members of each epistatic module display epistatic hierarchy, which seems to reflect a related functional pathway. Rat homologues of 11 human BP QTLs belong to either EM1 or EM2. Unique insights emerge into the novel genetic mechanism and hierarchy determining BP in the Dahl salt-sensitive SS/Jr (DSS) rat model that implicate a portion of human QTLs. Elucidating the pathways underlying EM1 and EM2 may reveal the genetic regulation of BP.

9.2 Introduction

Essential hypertension, the most prominent of human disorders, leads to stroke and fatal cardiovascular diseases (1,2). Up to now, hypertension is under control in only 34% of patients, and even for them, available treatments are empirical at best as they are not directed at causes of the disease, but rather at symptoms that require life-long medication (3). The greatest challenge in finding primary triggers, not secondary responses, to essential hypertension is to identify quantitative trait loci (QTLs) that underlie the physiological etiologies of blood pressure (BP) determination (2). Genome-wide association studies (GWAS) in humans have localized QTLs for systolic (SBP) and diastolic (DBP) (2,4), mean arterial pressure (MAP) and pulse pressure (5).

Each QTL is assumed to exert a miniscule phenotypic effect, and multiple QTLs work together to incrementally augment BP. Evidently, 28 human QTLs (2) account for less than 1% of total BP changes. By inference, thousands of such QTLs would have to exist to explain the majority of BP variations in general human populations. Although fractionating polygenic risk factors is often adopted in epidemiological studies, the biological impact of such practice in a functional context remains enigmatic.

Coincidentally, experimental models of hypertension have yielded numerous QTLs (6,7). Are they required as quantitative increments cumulatively achieving a BP threshold? In this context, how do Dahl salt-sensitive SS/Jr (DSS) rats, that possess BP-lowering alleles at several QTLs, stay hypertensive (8,9)? Is it because they have more BP-elevating than BP-diminishing QTL alleles that mathematically counterbalance one another?

These issues have to be addressed experimentally, i.e. via the investigation of each QTL individually and in combination with another, while retaining the rest of the genome homogeneously constant. This approach is not feasible in clinical studies, but the insights gained are valuable for understanding the genetic hierarchy and organization of essential hypertension. The current work aims to elucidate the organization that assembles and configures the genetic architecture that determines BP in DSS rats.

9.3 Materials and Methods

N.B. Cette section se trouve à la fin de l'article dans la version publiée. Elle a été déplacée dans cet ouvrage à fin de rendre le texte homogène [151].

9.3.1 Animals

Although inbred DSS rats are known as a model of salt-sensitivity (23), the QTLs governing their BP can act either on low- or high-salt diets (24). Thus, the genetic determination of hypertension in DSS rats can serve as a general model of essential hypertension. Protocols for handling and maintaining animals were approved by our institutional animal ethics committee (CIPA).

9.3.2 Construction of new congenic strains or substrains

DSS and Lewis rats were parents in the production of congenic strains. Breeding and screening procedures were accelerated similarly to those reported previously (8). New strains were validated by polymorphic microsatellite markers with an overall marker density of 10-20 cM/marker across the genome. The genotypes of markers are indicated in Supplemental Figure 1 for the region where crossovers were sought to derive congenic substrains. In the current work, 4 new congenic strains were produced (Supplemental Fig. 1a, 1c) and designated as C1S.L2, C1S.L3, C1S.L4 and C5S.L. C1 refers to Chr 1; S.L denotes congenic substitution of a chromosomal segment from hypertensive DSS rats by their homologue from Lewis rats in the DSS genetic background. A number following S.L indicates # of congenic strains.

One new subline came from C16S.L5 (14), designated as C16S.L7 (Supplemental Fig. 1g). 3 new sublines were derived from C17S.L2 (25), designated as C17S.L9, C17S.L11 and C17S.L12, respectively (Supplemental Fig. 1h).

9.3.3 Systematic assessment of QTL-QTL interactions by congenic combinations

A congenic strain relies on specific replacement of a chromosome segment from the recipient strain by its homologue from a donor, while keeping the remaining genome as that of the recipient (26). In a sense, the congenic strategy is similar to that of ‘knock in’. If BP changes in a congenic strain, a QTL should reside within the chromosome segment replaced.

Supplemental Figure 1 depicts congenic strains entrapping the QTLs under study. Only congenic strains produced in the DSS genetic background were utilized, because those made in the Lewis background did not exhibit any BP effect (18,27) and thus cannot be informative on QTL-QTL interactions.

Each interaction was tested independently in that each time only 3 strains in a comparison were measured for BP along with DSS rats. For example, when analyzing the interaction between *C10QTL1* and *C10QTL2* (Table 1 and Supplemental Fig. 1a), only 3 congenic strains plus DSS were analyzed simultaneously, i.e. DSS, C10S.L30, C10S.L16 and C10S.L30/C10S.L16. To analyze the interaction between *C10QTL2* and *C10QTL4*, only DSS, C10S.L16 (Table 1 and Supplemental Fig. 2b), C10S.L28 and C10S.L16/C10S.L28 were studied, and so forth for the remaining comparisons listed in Table 1.

9.3.4 Animal protocols and BP measurements

Breeding protocols and dietary treatments were the same as reported previously (8). Male rats were weaned at 21 days of age, maintained on a low-salt diet (0.2% NaCl), then fed a high-salt diet (2% NaCl), starting from 35 days of age until the end of the experiment. Telemetry probes were implanted when the rats were 8 weeks old (i.e. after 3 weeks on the 2% NaCl diet). BPs for all strains were measured continuously for 2 weeks, starting from the 10th day of post-operative recuperation.

9.3.5 Statistical analysis

Repeated measures’ analysis of variance (ANOVA) followed by Dunnett’s test (taking multiple comparisons and sample sizes into account) compared DBP, SBP and MAP between

groups, as reported previously(10,28). 2 x2 ANOVA assessed QTL-QTL interactions (or lack thereof) by evaluationg whether the observed effet of a ‘double’ or ‘multiple’ congenic strain combining separate QTLs was significantly different from a predicted sum of effects from each individual QTL (10,28)

9.4 Results

9.4.1 Study design in grouping QTLs

In isolation, the effect magnitude from a single QTL accounts for at least 20% of the BP difference between 2 contrasting parental strains, and no combination with another QTL is necessary to achieve it (Supplemental Fig.1). Epistasis refers to the effect of one gene masking that of another when acting together (6). Epistatic interactions predict that the BP effects of 2 QTLs cannot exceed those of a single QTL alone, (10). Based on this phenomenon,Deng proposed (6) a modularized scheme in that multiple BP QTLs may be assembled on the basis of their epistatic interactions or lack thereof.

To accomplish this task, one QTL can be used as a reference/entry point to which several QTLs can be modularized. A new component in the same epistatic module can be discovered by making a ‘double’ congenic strain, combining a QTL known to be a component of one starting module with another QTL whose modular involvement is not evident. If such a ‘double’ congenic strain has the same BP as each of the 2 ‘single’ congenic strains that constitute it, the 2 QTLs behave epistatically to each other and, consequently, they can be regrouped into the same module. If not, they belong to a different module. By repeating this process with all the known QTLs, more components of the same epistatic module can be added into the existing modular network, and an independent module can be established.

Epistatic hierarchy among QTLs in the same module can be determined if 2 QTLs have discernible BP effects. For example, a BP-decreasing QTL stands clearly higher in epistatic

hierarchy than a BP-increasing QTL on Chromosome (Chr) 3, because the combined effects of the 2 QTLs are the same as those of the BP-decreasing QTL alone (11). Based on this rationale, the following experiments were performed.

As a starting point, combination of 2 congenic strains, C10S.L30 (containing *C10QTL1*) and C10S.L16 (containing *C10QTL2*) (Supplemental Fig.1f), proved that they act additively (Table 1 and Fig. 1a). C10 refers to Chr 10. S.L30 indicates that the congenic strain, #30, was made in the genetic background of DSS rats (see Materials and Methods section for details). QTL2 refers to QTL #2 on Chr 10.

C10QTL1 and *C10QTL2* served as 2 initiating and separate reference/entry points onto which all QTLs (Supplemental Table 1.) could be modularized for their mutual epistatic relationships (6). A new member of the *C10QTL1* module, i.e. epistatic module 1 (EM1), can be found if that QTL acts epistatically to it, such as *C10QTL5* (12). The same principle applies to the module composed of *C10QTL2*, i.e. EM2. A new EM, in addition to EM1 and EM2, can be discovered if a given QTL acts additively with both *C10QTL1* and *C10QTL2*.

9.4.2 Modularization of BP QTLs

An analysis was first performed of *C10QTL4* and *C10QTL2* by combining 2 separate congenic strains, C10S.L28 and C10S.L16 (Supplemental Fig.1f), into 1 single congenic strain, C10S.L28/C10S.L16. The combined effect of *C10QTL4*+*C10QTL2* did not exceed that of *C10QTL4* or *C10QTL2* alone (Table 1 and Fig. 1b), i.e. an epistatic interaction exists between them ($p < 0.002$). Thus, *C10QTL4* and *C10QTL2* belong to EM2 (Table 2).

Next, the interaction between *C10QTL1* and *C16QTL* was evaluated by combining 2 congenic strains, C10S.L30 and C16S.L7 (Supplemental Fig.1f and 1g) into a single congenic strain, C10S.L30/C16S.L7. Simultaneously, the C10S.L16/C16S.L7 combination was produced to assess the interaction between *C10QTL2* and *C16QTL*. *C16QTL* behaves epistatically with *C10QTL1* ($p = 0.049$), but additively with *C10QTL2* ($p = 0.31$) (Table 1 and Fig. 1c, d). Thus, *C16QTL* and *C10QTL1* belong to EM1. Modeled on classifying *C16QTL*, other QTLs were analyzed in ‘double’ congenic combinations. In all, 27 ‘double’ congenic

combinations were created and independently examined (Table 1). All QTLs (Supplemental Fig.1) participated in either EM1 or EM2 except for *C2QTL1* (Table 2). *C2QTL1* acted additively with both *C10QTL1* and *C10QTL2*, and, consequently, belongs to a separate EM: EM3.

EM1, EM2 and EM3 have 16, 8 and 1 members, respectively (Table 2). Three QTLs belonged to either EM1 or EM2, but their precise classification was not established. Genetic hierarchy exists among the QTLs in each module. In EM2, the BP-decreasing *C10QTL2* stands higher in hierarchy than the BP-increasing *C8QTL2*, because the effect of *C10QTL2* masks that of *C8QTL2*, not *vice versa* (Table 1 and Fig.1f).

Next, we addressed the issue of whether or not modularized groupings of BP QTLs had biological meanings or were just statistical rituals devoid of them. We reasoned that, if EM1 and EM2 were fundamental function blocks, BP should stay constant, regardless of how many QTLs are present from each module.

9.4.3 Biological significance of QTL modularization

To prove that EM1 and EM2 constitute 2 essential functional cores in BP regulation, we generated multiple QTL combinations, each of them containing at least 1 member from EM1 and EM2, all carrying BP-lowering QTL alleles (Supplemental Table.1). 3 to 8 BP QTL combinations were similar ($p > 0.15$) to those of 2 QTLs (*C10QTL1* + *C10QTL2*) (Fig. 2), as if EM1 + EM2 exerted a ‘ground floor’ effect. This outcome was not due to bottom limit, because the BP of Lewis rats (95mmHg) seemed still lower ($p < 0.001$) than that of EM1+EM2 combined (110-125 mmHg). Consequently, more QTLs beyond the 2 core members, each from EM1 and EM2, became redundant rather than cumulative in their combined impact on BP.

Next, we examined correspondence between the BP QTLs found in our animal model and those identified in humans.

9.4.4 Implications of modularization for human QTLs

11 out of 42 (26%) human genes or QTL intervals (underlined in Supplemental Table 2) (2,4,5) were included in 8 congenic strains demonstrating BP effects. Their rat homologues belonged to either EM1 or EM2 affecting BP. *MOV10*, *CAPZA1* and *FIGN/GRB14* were contained in large congenic strains, C2S.M (13) and C3S.L2 (11), but were not modularized for lack of appropriate congenic strains that could isolate them from adjacent QTLs.

We then analyzed rat homologues of human genes implicated in BP regulation or harboring structural mutations. C1S.L2 traps 4 human QTLs around ADM (Supplemental Fig.1a) that interact epistatically with *C10QTL1* + *C10QTL2* (Table 1), i.e. they belong to either EM1 or EM2 (Table 2). C3S.L7 (14) carries the segment harboring *Jag1/Gnas/EDN3*, corresponding to *C3QTL3* (EM1) (Table 2 and Supplemental Fig.1b). The coding regions and exon-intron junctions for *Adm* (15) and *Edn3* (16) are identical between DSS and Lewis [and DSS rats (17) for *Edn3*] (Supplemental Table. 3a, b).

Although a non-synonymous *HFE* SNP is associated with human BP (2), and *Hfe* is situated in the *C17QTL2*-lodging interval (Supplemental Fig.1h), no structural and splicing *Hfe* mutations were detected (Supplemental Table.3c). Barring variants impacting on their gene expressions, *Adm*, *Edn3* and *Hfe* themselves are not supported as BP QTLs. Nevertheless, *C17QTL2* marked by *Hfe* exists and belongs to EM2 (Table 2).

11 out of 42 human QTLs, such as *GUCY1A3-GUCY1B3*, *SLC39A8*, and *NPR3-C5orf23*, were excluded as BP QTLs contrasting DSS rats with a specific normotensive strain (Table 2 and Supplemental Fig.1e). They may still influence BP in other strain comparisons, because different normotensive, and probable hypertensive, strains may have distinctive genetic determinants (18). Thus, the specific gene for these human QTLs await functional identification.

9.5 Discussion

Significant revelations from this study are: **(a)** 28 QTLs that can provoke hypertension regulate BP via 3 functional modules, namely, EM1, EM2 and EM3. Modularity in QTL actions gives rise to a novel genetic concept. This unique mechanism may help shed light on the functionality of certain human BP QTLs. **(b)** EM1 and EM2 are fundamental modules that provide novel insights into the genetic mechanisms and hierarchy globally governing BP homeostasis.

9.5.1 Epistatic modularization, not quantitative accumulation, is the *modus operandi* in biological BP determination by QTLs.

The prevailing view is that a complex and quantitative trait, such as BP, must be realized by arithmetic accumulation of numerous QTLs, each of which has only a miniscule influence on overall phenotype. Counter-intuitively, the effects of DSS QTLs individually and in assembly do not produce a spectrum of continuous BP variations, but rather impact BP as ‘units’ in relatively clear-cut ‘leaps’ (Fig. 1). This genetic insight suggests that the aggregating outcome of multiple QTLs in DSS rats is achieved via a modularized mechanism, not progressive increments. All QTLs in the same EM may be interpreted to be components up- or down-stream in a common pathway. Accordingly, EM1 and EM2 seem to be multi-component pathways.

The modularized modality of QTLs in controlling BP in the DSS inbred model has added a new dimension to, but does not nullify or replace, the established norm governing those in outbred populations at large. BP appears as a continuum in a distribution curve in general human populations (2) and outbred animals. The gene actions determining this continuum should now include the modularized mechanism in addition to the small-effect-but-cumulative paradigm (2).

Although the modularity of BP QTLs does exist in an inbred model, it is not evident whether or not, and to what extent, it can be generalized to an outbred population, such as in humans. An outbred population consists of myriad individuals, and each of them, or monozygotic twins (or multiple births), can be viewed as an inbred strain. Conceivably, QTLs

may be modularized in their influence on BP in some individuals, but not in others. It is possible that certain QTLs may be modularized across heterogeneous individuals. The only way to establish modularity between 2 QTLs is to isolate a QTL and then combine both of them. This approach is not achievable in outbred populations, but individual QTLs can be isolated in inbred rodent models and tested, as in the current study.

In this context, it is necessary to distinguish the actual biological impact of a QTL from the statistically-estimated magnitude of phenotypic effect and the statistically-fractionated contribution of each QTL to the whole phenotype. For example, after typical genetic analysis in a F₂(DSS x Lewis) population (19), the magnitude of BP effect for each QTL was calculated and the contribution to global BP calculated statistically as a percentage, accounting for total variance. While these estimates are widely accepted as the standard genetic interpretation of a QTL action (20), the true biological effects of QTLs individually and in combination only become clear in isolation and in pair-wise combinations. It turns out that all QTLs contrasting DSS and Lewis rats are modularized in their effects on BP (Table 2). Considering that individuals in a F₂ population are genetically heterogeneous, they approximately resemble an outbred population. Thus, the statistical treatment of QTLs in an outbred or F₂ population follows the principles of population genetics, whereas the modularity of QTLs is governed by their biological impact on BP. They represent 2 different aspects of QTL characterization, namely, statistical vs. biological, depending on the context of analysis and the rigor of functional validation.

Discovering QTL modularity is important in that it paves the way for identifying and understanding a common pathway to BP homeostasis and the hierarchical relationship among pathway components.

Once again, the limited application of the modularized mechanism is obvious, as it appears in inbred strains. This mechanism is yet to be tested in outbred populations. Whether or not QTLs would function cumulatively or otherwise in a physiological setting in outbred human populations awaits further functional proves beyond simple statistics. Nevertheless, the modularity exists as shown in our current work, and is a possibility that warrants attention in the future research of human polygenic hypertension.

9.5.2 Pathway of epistatic modules

Candidate genes have been identified for several QTLs in EM1. They are *Loc100363423* for *C10QTL3*, *Proline-rich 11 (Prr11)* for *C10QTL5*, *Alpha kinase 2 (Alpk2)* for *C18QTL3*, and 5 candidates for *C10QTL1* (Table 2 and Supplemental Fig. 1). Since all these genes are not known to regulate BP biologically and even their cellular functions are not identified, the physiological mechanisms by which they influence BP might fall in the same pathway, because they belong to the same module. This framework may pave the way for identifying one QTL based on a known mechanism of another QTL in the same epistatic module. For example, assuming that *Alpk2* may encode a kinase that might be involved in BP control, *PRR11* could somehow participate in the *ALPK2* pathway.

One candidate gene was found for *C10QTL2* in EM2 to be *ATP-binding cassette, subfamily A (ABC1), member 8a (Abca8a)*. *ABCA8A* belongs to a family of transporters that are implicated in the high-density lipoprotein metabolism, atherogenesis and coronary heart disease via control of cholesterol efflux (21). The mechanism of its involvement in BP homeostasis is not known, but based on *Abca8a* being in EM2, the mode of its action may be independent of those of *Alpk2* and *Prr11*, which belong to EM1.

Nevertheless, these QTLs in EM1 and EM2 are novel and may influence BP either directly or indirectly via regulation of other genes or their products known to affect BP.

In contrast, genes encoding well-known molecules are either excluded or unsupported as etiologically BP-impacting agents, such as *Ace*, *Nos2*, *Adm*, *Edn3*, *Npr1*, *Npr3*, *Gucyl1a3*, *Gucyl1b3*, *Atp1a1*, *Adrb2*, and *Nedd4l* (Table 2 and Supplemental Fig. 1). How can they be important in BP regulation? One possibility is that they may not be primary triggers, but secondary responders to triggers such as *Abca8a* and *Prr11* in BP homeostasis.

Given the power of resolution in modern GWAS of BP QTLs, only 11 human QTLs corresponding to those found in the Dahl model were detected. The majority of rat QTLs were ‘missed’, especially those on Chr 2, 10 and 18, in human studies (Supplemental Fig. 1).

A few factors may explain this discrepancy. First, certain QTLs found in rats may be rodent-specific and they may not function as triggers to initiate hypertension in humans.

Second, since a high-salt diet was included in the BP protocol, certain QTLs found under this condition may mostly apply to salt-accelerated hypertension. Third, in human studies, both women and men were assumed to be recruited and the data on them were pooled, whereas in our current animal model, only males were investigated. There is evidence that some rat QTLs affect BP only in males, not in females (22). Thus, certain rat QTLs may be sex-specific. Finally, some rat QTLs were detected only when the genetic background was homogeneous, not when it was heterogeneous. For example, a QTL on Chr 7 was detected only in a congenic strain (18), but was missed in a F₂ (DSS x Lewis) population (19). Human populations are heterogeneous, whereas our congenic strains are homogeneous.

9.6 Conclusion

In summary, multiple QTLs function via specific modules in the inbred DSS model. In designing anti-hypertensive drugs, simultaneously targeting mediators in 2 separate modules seems more effective in lowering BP of a hypertensive individual than targeting those in the same module. Functional hierarchy exists among QTLs in the same EM, which is demonstrated by the interplay between BP-increasing and BP-decreasing QTL alleles. 26% of human QTLs have rat homologues that belong to EM1 and EM2 combined, implying that certain human QTLs might function via a modularized mechanism. The QTLs identified in animal models (Supplemental Table. 1) may facilitate the discovery of further genetic architecture in humans that is hidden from the most powerful GWAS, but is pertinent to essential hypertension, since 116 human QTLs are estimated to exist (2) and only 28 have been found.

Acknowledgements : This work was supported by grants from the Canadian Institutes of Health Research and a joint China-Canada Health Research Initiative to AYD. CC is supported by a doctoral fellowship from the Heart and Stroke Foundation of Canada.

Conflict of interest statement : The authors have no conflict of interest to declare.

9.7 References

1. Kearney,P.M., Whelton,M., Reynolds,K., Muntner,P., Whelton,P.K., and He,J. (2005) Global burden of hypertension: analysis of worldwide data. *Lancet* 365, 217-223.
2. International Consortium for Blood Pressure (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478, 103-109.
3. Chobanian,A.V., Bakris,G.L., Black,H.R., Cushman,W.C., Green,L.A., Izzo,J.L., Jones,D.W., Materson,B.J., Oparil,S., Wright,J.T., et al. (2003) Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension* 42, 1206-1252.
4. Kato,N., Takeuchi,F., Tabara,Y., Kelly,T.N., Go,M.J., Sim,X., Tay,W.T., Chen,C.H., Zhang,Y., Yamamoto,K., et al. (2011) Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat Genet* 43, 531-538.
5. Wain,L.V., Verwoert,G.C., O'Reilly,P.F., Shi,G., Johnson,T., Johnson,A.D., Bochud,M., Rice,K.M., Henneman,P., Smith,A.V., et al. (2011) Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet* 43, 1005-1011.
6. Deng,A.Y. (2007) Genetic basis of polygenic hypertension. *Hum Mol Genet* 16, R195-R202.
7. Rapp,J.P. (2000) Genetic analysis of inherited hypertension in the rat. *Physiol Rev* 80, 135-172.
8. Ariyarajah,A., Palijan,A., Dutil,J., Prithiviraj,K., Deng,Y., and Deng,A.Y. (2004) Dissecting quantitative trait loci into opposite blood pressure effects on Dahl rat chromosome 8 by congenic strains. *J Hypertens* 22, 1495-1502.

9. Eliopoulos,V., Dutil,J., Deng,Y., Grondin,M., and Deng,A.Y. (2005) Severe hypertension caused by alleles from normotensive Lewis for a quantitative trait locus on chromosome 2. *Physiol Genomics* 22, 70-75.
10. Charron,S., Duong,C., Menard,A., Roy,J., Eliopoulos,V., Lambert,R., and Deng,A.Y. (2005) Epistasis, not numbers, regulates functions of clustered Dahl rat quantitative trait loci applicable to human hypertension. *Hypertension* 46, 1300-1308.
11. Palijan,A., Dutil,J., and Deng,A.Y. (2003) Quantitative trait loci with opposing blood pressure effects demonstrating epistasis on Dahl rat chromosome 3. *Physiol Genomics* 15, 1-8.
12. Chauvet,C., Charron,S., Menard,A., Xiao,C., Roy,J., and Deng,A.Y. (2008) Submegabase resolution of epistatically interacting quantitative trait loci for blood pressure applicable for essential hypertension. *J Hypertens* 26, 893-901.
13. Dutil,J., Eliopoulos,V., Tremblay,J., Hamet,P., Charron,S., and Deng,A.Y. (2005) Multiple quantitative trait loci for blood pressure interacting epistatically and additively on dahl rat chromosome 2. *Hypertension* 45, 557-564.
14. Duong,C., Charron,S., Xiao,C., Hamet,P., Menard,A., Roy,J., and Deng,A.Y. (2006) Distinct quantitative trait loci for kidney, cardiac, and aortic mass dissociated from and associated with blood pressure in Dahl congenic rats. *Mamm Genome* 17, 1147-1161.
15. Nicholls,M.G., Lainchbury,J.G., Lewis,L.K., McGregor,D.O., Richards,A.M., Troughton,R.W., and Yandle,T.G. (2001) Bioactivity of adrenomedullin and proadrenomedullin N-terminal 20 peptide in man. *Peptides* 22, 1745-1752.
16. Inoue,A., Yanagisawa,M., Kimura,S., Kasuya,Y., Miyauchi,T., Goto,K., and Masaki,T. (1989) The human endothelin family: three structurally and pharmacologically distinct isopeptides predicted by three separate genes. *Proc Nat Acad Sci USA* 86, 2863-2867.

17. Lee,S.J., Liu,J., Westcott,A.M., Vieth,J.A., DeRaedt,S.J., Yang,S., Joe,B., and Cicila,G.T. (2006) Substitution mapping in dahl rats identifies two distinct blood pressure quantitative trait loci within 1.12- and 1.25-mb intervals on chromosome 3. *Genetics* 174, 2203-2213.
18. Crespo,K., Chauvet,C., Blain,M., Menard,A., Roy,J., and Deng,A.Y. (2011) Normotension in Lewis and Dahl salt-resistant rats is governed by different genes. *J Hypertens* 29, 460-465.
19. Garrett,M.R., Dene,H., Walder,R., Zhang,Q., Cicila,G.T., Assadnia,S., Deng,A.Y., and Rapp,J.P. (1998) Genomic scan and congenic strains for blood pressure quantitative trait loci using Dahl salt-sensitive rats. *Genome Res* 8, 711-723.
20. Lander,E., Kruglyak,L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11, 241-247.
21. Jensen,M.K., Pai,J.K., Mukamal,K.J., Overvad,K., and Rimm,E.B. (2007) Common genetic variation in the ATP-binding cassette transporter A1, plasma lipids, and risk of coronary heart disease. *Atherosclerosis* 195, e172-e180.
22. Deng,A.Y., Menard,A., Xiao,C., and Roy,J. (2008) Sexual dimorphism on hypertension of quantitative trait loci entrapped in dahl congenic rats. *Clin Exp Hypertens* 30, 511-519.
23. Dahl,L.K., Heine,M., and Tassinari,L. (1962) Role of genetic factors in susceptibility to experimental hypertension due to chronic excess salt ingestion. *Nature* 194, 480-482.
24. Huang,B.S., Ahmad,M., Deng,A.Y., and Leenen,F.H. (2007) Neuronal responsiveness to central Na⁺ in 2 congenic strains of Dahl salt-sensitive rats. *Hypertension* 49, 1315-1320.
25. Grondin,M., Eliopoulos,V., Lambert,R., Deng,Y., Ariyaratnam,A., Moujahidine,M., Dutil,J., Charron,S., and Deng,A.Y. (2005) Complete and overlapping congenics proving the existence of a quantitative trait locus for blood pressure on Dahl rat chromosome 17. *Physiol Genomics* 21, 112-116.

26. Deng,A.Y. (2007) Positional cloning of quantitative trait loci for blood pressure: how close are we?: a critical perspective. *Hypertension* 49, 740-747.
27. Charron,S., Lambert,R., Eliopoulos,V., Duong,C., Menard,A., Roy,J., and Deng,A.Y. (2005) A loss of genome buffering capacity of Dahl salt-sensitive model to modulate blood pressure as a cause of hypertension. *Hum Mol Genet* 14, 3877-3884.
28. Rapp,J.P., Garret,M.R., and Deng,A.Y. (1998) Construction of a double congenic strain to prove an epistatic interaction on blood pressure between rat chromosomes 2 and 10. *J Clin Invest* 101, 1591-1595.
29. Chauvet,C., Menard,A., Tremblay,J., Xiao,C., Shi,Y., L'heureux,N., Cardin,S., Tardif,J.C., Nattel,S., and Deng,A.Y. (2009) Cardiac pathways distinguish two epistatic modules enacting BP quantitative trait loci and candidate gene analysis. *Hypertens Res* 32, 631-637.
30. Chauvet,C., Crespo,K., Menard,A., Wu,Y., Xiao,C., Blain,M., Roy,J., and Deng,A.Y. (2011) alpha-Kinase 2 is a novel candidate gene for inherited hypertension in Dahl rats. *J Hypertens* 29, 1320-1326.
31. Chauvet,C., Menard,A., Xiao,C., Aguila,B., Blain,M., Roy,J., and Deng,A.Y. (2012) Novel genes as primary triggers for polygenic hypertension. *J Hypertens* 30, 81-86.

9.8 Tables, figures and legends

Figure 1. Analysis of interactions among representative QTLs

The left three columns in each panel represent differences in mean arterial pressure (MAP) between a congenic and DSS strains, i.e. BP lowered by a congenic strain. The last column to the right indicates a ‘predicted’ value of MAP lowered by combining 2 different QTLs. The BP effect of C10QTL1 is represented by congenic strain C10S.L30; the BP effect of C10QTL2 is represented by congenic strain C10S.L16; the combined BP effect of [C10QTL1 and C10QTL2] is represented by a ‘double’ congenic strain C10S.L30+C10S.L16. *p* indicates the results on the interaction as performed by 2 x 2 ANOVA. The actually ‘observed’ BP effect of the ‘double’ congenic strain combining C10QTL1+C10QTL2 is not different (*p* interaction <0.067) from a ‘predicted’ sum of BP effects of C10QTL1+C10QTL2, indicating additivity. In contrast, the actually ‘observed’ BP effect of another ‘double’ congenic strain combining C10QTL2+C10QTL4 is different (*p* interaction <0.002) from a ‘predicted’ sum of BP effects of C10QTL2+C10QTL4, indicating epistasis. (a), additivity between C10QTL1 and C10QTL4; (b), epistasis between C10QTL2 and C10QTL4; (c), additivity between C10QTL2 and C16QTL; (d), epistasis between C10QTL1 and C16QTL; (e), additivity between C8QTL2 and C10QTL1; (f) epistasis between C8QTL2 and C10QTL2.

Figure 2. BPs of single and multiple congenic strains producing 1-8 QTL combinations

DBP, diastolic blood pressure; MAP, mean arterial pressure; SBP, systolic blood pressure. Error bars represent SEM. *n* of strains ranges from 5 to 12. BP response patterns, such as diurnal variations among all congenic strains, were not different. For simplicity of presentation and comparison, a 24-hour BP average is considered as 1 data point on the graph for each strain. C16QTL and C17QTL1 are defined by single congenic strains C16S.L7 and C17S.L11 respectively (Supplemental Fig. 1g, 1h). A 2-QTL combination was achieved by combining 2 congenic strains carrying them, then 3, 4, and so on. 2 (or multiple) closely-linked QTLs are combined into a congenic strain containing both of them (or multiples). 2 QTLs are C10QTL1 (epitomizing EM1) + C10QTL2 (exemplifying EM2); 3 QTLs are C10QTL4 (EM2) + C18QTL3 (EM1) + C2QTL6 (EM1); 4 QTLs are C10QTL1 (EM1) + C10QTL2 (EM2) + C2QTL4M (EM1) + C2QTL5 (EM1); 5 QTLs are C2QTL5 (EM1) +

C10QTL4 (EM2) + C16QTL (EM1) + C18QTL3 (EM1) + C18QTL4 (EM1); 6 QTLs are C10QTL4 (EM2) + C10QTL5 (EM1) + C10QTL1 (EM1) + C10QTL3 (EM1) + C10QTL2 (EM2) + C3QTL1 (EM1); 7 QTLs are C10QTL4 (EM2) + C10QTL5 (EM1) + C10QTL1 (EM1) + C10QTL3 (EM1) + C10QTL2 (EM2) + C18QTL1 (EM1) + C18QTL2 (EM1); and 8 QTLs are C10QTL2 (EM2) + C10QTL3 (EM1) + C3QTL3 (EM1) + C2QTL3 (EM2) + C2QTL2 (EM1) + C2QTL4M (EM1) + C2QTL5 (EM1) + C2QTL6 (EM1). EM, epistatic module (Table 2). ANOVA + Dunnett values for all BP components are $p < 0.03$ (most conservative) for DSS vs congenic strains trapping QTLs; $p < 0.01$ for congenic strains harboring C16QTL or C17QTL1 vs congenic strains containing 2-8 QTLs and Lewis.

Figure 1

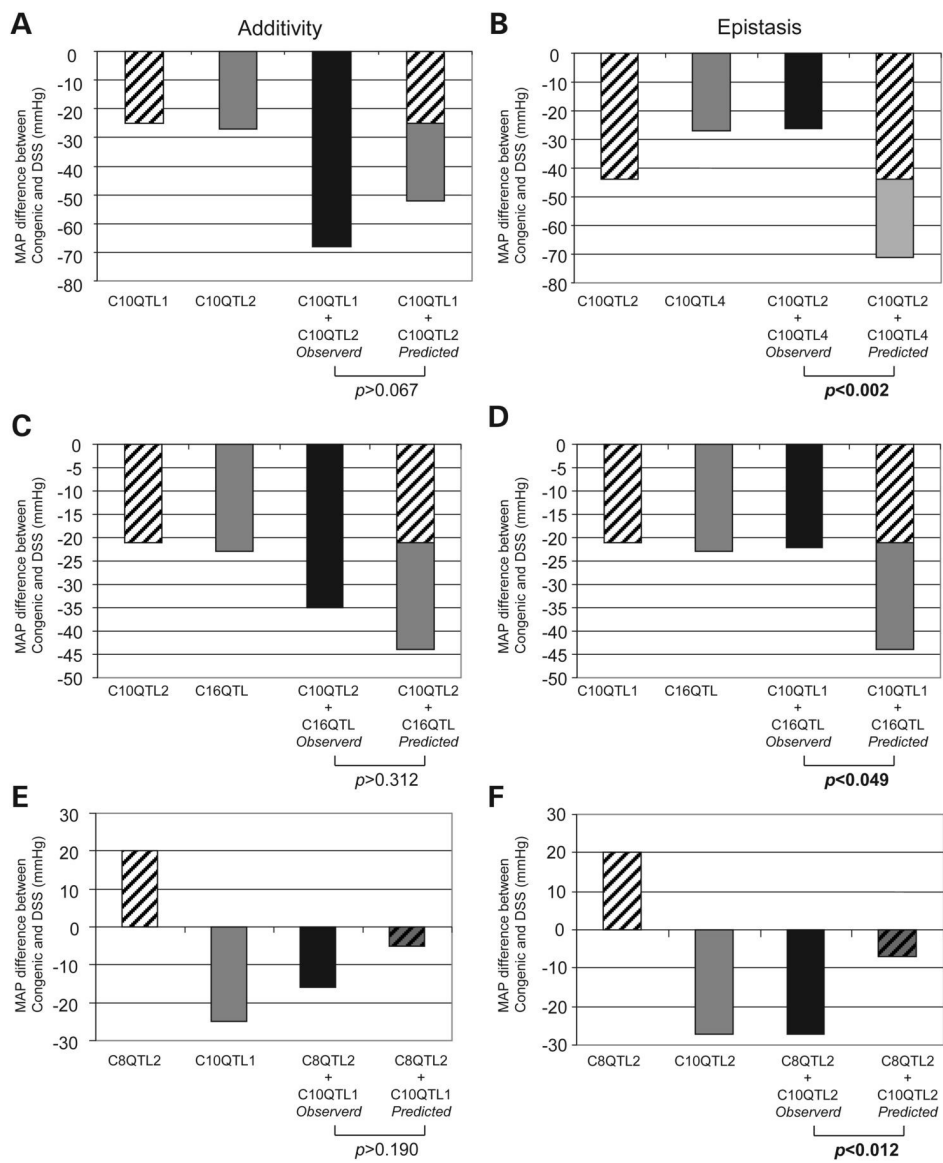


Figure 2

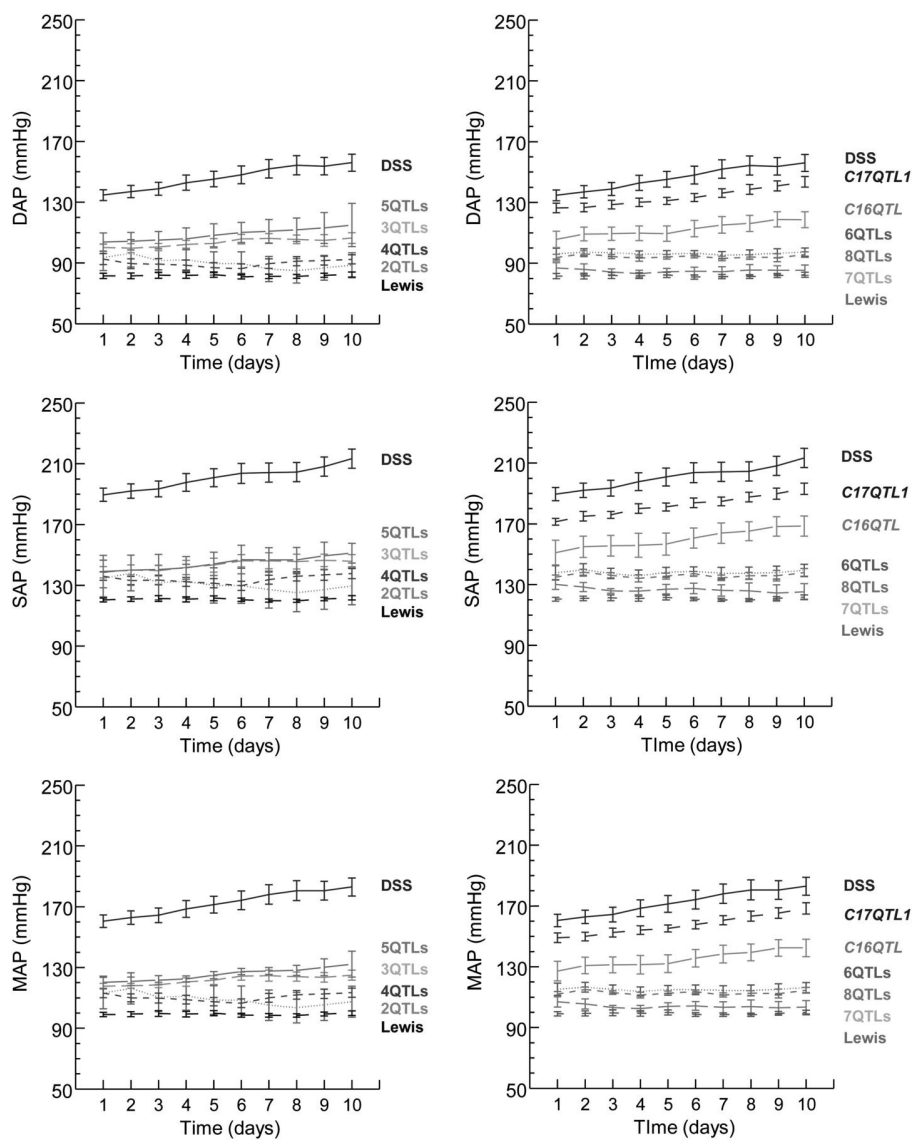


Table 1. Assessment of QTL-QTL interactions by congenic combinations

Congenic strains	QTL designation	n	MAP (mmHg)	MAP differences		Mechanism
				MAP (mmHg)	MAP (mmHg) minus DSS	
C10S.L30	<i>C10QTL1</i>	10	151 ± 3	-26		
C10S.L16	<i>C10QTL2</i>	7	150 ± 3	-27		
C10S.L30/C10S.L16	<i>C10QTL1+C10QTL2</i> observed	5	109 ± 8	-68	<i>p</i> =0.067	Additive
	<i>C10QTL1+C10QTL2</i> predicted			-53		
C10S.L16	<i>C10QTL2</i>	7	150 ± 3	-34		
C10SL28	<i>C10QTL4</i>	9	133 ± 5	-27		
C10S.L16/C10S.L28	<i>C10QTL2+C10QTL4</i> observed	5	151 ± 4	-26	<i>p</i> <0.002	Epistasis
	<i>C10QTL2+C10QTL4</i> predicted			-71		
C10S.L30	<i>C10QTL1</i>	10	151 ± 3	-21		
C16S.L7	<i>C16QTL</i>	7	149 ± 2	-23		
C10S.L30/C16S.L7	<i>C10QTL1+C16QTL</i> observed	8	150 ± 9	-22	<i>p</i> =0.049	Epistasis
	<i>C10QTL1+C16QTL</i> predicted			-44		
C10S.L16	<i>C10QTL2</i>	6	151 ± 3	-21		
C16S.L7	<i>C16QTL</i>	7	149 ± 2	-23		
C10S.L16/C16S.L7	<i>C10QTL2+C16QTL</i> observed	8	137 ± 3	-35	<i>p</i> =0.31	Additive
	<i>C10QTL2+C16QTL</i> predicted			-44		
C8S.L2	<i>C8QTL2</i>	9	197 ± 5	20		
C10S.L16	<i>C10QTL2</i>	7	150 ± 3	-27		
C8S.L2/C10S.L16	<i>C8QTL2+C10QTL2</i> observed	8	150 ± 4	-27	<i>p</i> <0.01	Epistasis
	<i>C8QTL2+C10QTL2</i> predicted			-7		
C8S.L2	<i>C8QTL2</i>	9	197 ± 5	20		
C10S.L30	<i>C10QTL1</i>	10	151 ± 3	-26		
C8S.L2/C10S.L30	<i>C8QTL2+C10QTL1</i> observed	8	161 ± 1	-16	<i>p</i> =0.19	Additive
	<i>C8QTL2+C10QTL1</i> predicted			-6		

C8S.L1	C8QTL1	7	138 ± 5	-39		
C10S.L30	C10QTL1	10	151 ± 3	-26		
C8S.L1/C10S.L30	C8QTL1+C10QTL1 observed	5	134 ± 2	-43	p<0.01	Epistasis
	C8QTL1+C10QTL1 predicted			-65		
C3S.L3	C3QTL1	7	144 ± 5	-33		
C10S.L30	C10QTL1	10	151 ± 3	-26		
C3S.L3/C10S.L30	C3QTL1+C10QTL1 observed	12	136 ± 3	-41	p<0.03	Epistasis
	C3QTL1+C10QTL1 predicted			-59		
C3S.L3	C3QTL1	7	144 ± 5	-33		
C10S.L16	C10QTL2	7	150 ± 3	-27		
C3S.L3/C10S.L16	C3QTL1+C10QTL2 observed	5	133 ± 5	-44	p=0.18	Additive
	C3QTL1+C10QTL2 predicted			-60		
C17S.L11	C17QTL1	5	129 ± 8	-48		
C10S.L16	C10QTL2	7	150 ± 3	-27		
C17S.L11/C10S.L16	C17QTL1+C10QTL2 observed	5	146 ± 3	-31	p<0.001	Epistasis
	C17QTL1+C10QTL2 predicted			-75		
C17S.L11	C17QTL1	5	137 ± 2	-46		
C17S.L9	C17QTL2	8	151 ± 7	-32		
C17S.L11/C17S.L9	C17QTL1+C17QTL2 observed	7	144 ± 6	-39	p<0.003	Epistasis
	C17QTL1+C17QTL2 predicted			-70		
C2S.L4	C2QTL4L	8	199 ± 4	22		
C10S.L16	C10QTL2	10	151 ± 3	-27		
C2S.L4 /C10S.L16	C2QTL4L+C10QTL2 observed	5	157 ± 7	-20	p<0.05	Epistasis
	C2QTL4L+C10QTL2 predicted			-5		
C2S.L4	C2QTL4L	8	199 ± 4	22		
C10S.L30	C10QTL1	10	151 ± 3	-26		
C2S.L4 /C10S.L30	C2QTL4L+C10QTL1 observed	6	183 ± 9	6	p=0.28	Additive
	C2QTL4L+C10QTL1 predicted			-4		
C16S.L7	C16QTL	7	134 ± 5	-43		
C18S.L3	C18QTL2	5	124 ± 4	-53		

C16S.L7/C18S.L3	C16QTL+C18QTL2 observed	5	141 ± 4	-36	<i>p</i><0.001	Epistasis
	C16QTL+C18QTL2 predicted			-96		
C3S.L7	C3QTL3	5	123 ± 2	-54		
C10S.L30	C10QTL1	11	144 ± 2	-33		
C3S.L7/C10S.L30	C3QTL3+C10QTL1 observed	5	116 ± 3	-61	<i>p</i><0.02	Epistasis
	C3QTL3+C10QTL1 predicted			-87		
C1S.L3	C1QTL2	10	134 ± 7	-43		
C10S.L16	C10QTL2	7	150 ± 3	-27		
C1S.L3/C10S.L16	C1QTL2+C10QTL2 observed	6	138 ± 5	-39	<i>p</i><0.008	Epistasis
	C1QTL2+C10QTL2 predicted			-70		
C1S.L2	C1QTL3	8	125 ± 4	-52		
C10S.L30/C10S.L16	C10QTL1+C10QTL2	10	107 ± 6	-70		
	C1QTL3+C10QTL1+C10QTL2					
C1S.L2/C10S.L30/C10S.L16	observed	6	109 ± 2	-68	<i>p</i><0.001	Epistasis
	C1QTL3+C10QTL1+C10QTL2 predicted			-122		
C1S.L4	C1QTL1	16	143 ± 3	-34		
C10S.L30/C10S.L16	C10QTL1+C10QTL2	5	109 ± 8	-69		
	C1QTL1+C10QTL1+C10QTL2					
C1S.L4/C10S.L30/C10S.L16	observed	5	118 ± 3	-59	<i>p</i><0.001	Epistasis
	C1QTL1+C10QTL1+C10QTL2 predicted			-103		
C7S.L	C7QTL	6	138 ± 3	-39		
C10S.L30/C10S.L16	C10QTL1+C10QTL2	10	107 ± 6	-70		
	C7QTL+C10QTL1+C10QTL2					
C7S.L/C10S.L30/C10S.L16	observed	5	125 ± 3	-52	<i>p</i><0.001	Epistasis
	C7QTL+C10QTL1+C10QTL2 predicted			-109		
C2S.M15	C2QTL3	8	144 ± 4	-33		

C10S.L16	C10QTL2	7	150 ± 3	-27		
C2S.M15/C10S.L16	C2QTL3+C10QTL2 observed	5	133 ± 2	-44	p=0.018	Epistasis
	C2QTL3+C10QTL2 predicted			-60		
C2S.M15	C2QTL3	8	144 ± 4	-33		
C10S.L25	C10QTL3	7	150 ± 2	-27		
C2S.M15/C10S.L25	C2QTL3+C10QTL3 observed	6	129 ± 4	-48	p=0.13	Additive
	C2QTL3+C10QTL3 predicted			-60		
C2S.M19	C2QTL2	9	124 ± 6	-53		
C10S.L30	C10QTL1	10	151 ± 3	-26		
C2S.M19/C10S.L30	C2QTL2+C10QTL1 observed	5	135 ± 5	-42	p<0.001	Epistasis
	C2QTL2+C10QTL1 predicted			-81		
C2S.M19	C2QTL2	5	138 ± 8	-39		
C10S.L16	C10QTL2	7	150 ± 3	-27		
C2S.M19/C10S.L16	C2QTL2+C10QTL2 observed	5	125 ± 2	-52	p=0.20	Additive
	C2QTL2+C10QTL2 predicted			-66		
C2S.M21	C2QTL5	16	149 ± 4	-28		
C10S.L30/C10S.L16	C10QTL1+C10QTL2	5	109 ± 8	-68		
C2S.M21/C10S.L30/C10S.L16	C2QTL5+C10QTL1+C10QTL2					
	observed	5	148 ± 9	-29	p<0.001	Epistasis
	C2QTL5+C10QTL1+C10QTL2 predicted			-96		
C2S.M7	C2QTL1	14	145 ± 3	-32		
C10S.L16	C10QTL2	7	150 ± 3	-27		
C2S.M7/C10S.L16	C2QTL1+C10QTL2 observed	5	133 ± 8	-44	p=0.12	Additive
	C2QTL1+C10QTL2 predicted			-59		
C2S.M7	C2QTL1	14	145 ± 3	-32		
C10S.L30	C10QTL1	10	151 ± 3	-26		
C2S.M7/C10S.L30	C2QTL1+C10QTL1 observed	7	131 ± 3	-46	p=0.12	Additive
	C2QTL1+C10QTL1 predicted			-58		
C2S.M7	C2QTL1	14	145 ± 3	-32		

C3S.L3	C3QTL1	7	144 ± 5	-33		
C2S.M7/C3S.L3	C2QTL1+C3QTL1 observed	9	117 ± 2	-60	<i>p</i> =0.56	Additive
	C2QTL1+C3QTL1 predicted			-65		

Table 1 legend: MAP, mean arterial pressure; DSS, Dahl salt-sensitive rats. QTL-QTL interactions were analyzed by 2 x 2 ANOVA as presented in Fig. 1 and analyzed as described in the Fig. 1 legend. No multiple comparisons were made between studies concerning epistatic interactions or a lack of it. Congenic strains defining QTLs are reported in Supplemental Fig.1

Table 2. Modularization of BP QTLs by epistasis

EM #	QTLs included	Model QTL	Candidate genes for model QTLs
EM1	<i>C10QTL1</i> , <i>C10QTL3</i> , <i>C10QTL5</i> , <i>C3QTL1</i> , <i>C3QTL2</i> , <i>C3QTL3</i> , <i>C8QTL1</i> , <i>C16QTL</i> , <i>C18QTL1</i> , <i>C18QTL2</i> , <i>C18QTL3</i> , <i>C18QTL4</i> , <i>C2QTL2</i> , <i>C2QTL4M</i> , <i>C2QTL5</i> , <i>C2QTL6</i> , (<i>C7QTL</i> , <i>C1QTL1</i> and <i>C1QTL3</i>)	<i>C10QTL5</i> , <i>C18QTL3</i>	<i>Proline-rich 11 (Prr11)</i> <i>Alpha kinase 2 (Alpk2)</i>
EM2	<i>C10QTL2</i> , <i>C10QTL4</i> , <i>C2QTL4L</i> , <i>C8QTL2</i> , <i>C1QTL2</i> , <i>C17QTL1</i> , <i>C17QTL2</i> , <i>C2QTL3</i> , (<i>C7QTL</i> , <i>C1QTL1</i> and <i>C1QTL3</i>)	<i>C10QTL2</i>	<i>ATP-binding cassette, subfamily A (ABC1), member 8a (Abca8a)</i>
EM3	<i>C2QTL1</i>	<i>C2QTL1</i>	Unidentified

Table 2 legend: QTLs in parentheses belong to either EM1 or EM2. All QTL-QTL interactions were assessed in the current work except for those closely linked on the same chromosome, such as C2QTL2, C2QTL4M, C2QTL5 and C2QTL6, being in the same epistatic module, EM1 (29), C3QTL1 and C3QTL2 in the same module (11), C10QTL1, C10QTL3 and C10QTL5 in the same module (10,12), and C18QTL1, C18QTL2, C18QTL3

and C18QTL4 in the same module (27,30). *Abca8a* and *Ppr11* are identified from (31), and *Alpk2* from (30).

Abbreviations

BP, blood pressure; *C1QTL1*, QTL1 on Chr 1; *C2QTL2*, QTL2 on Chr 2, and so on; Chr, chromosome; DBP, diastolic blood pressure; DSS, Dahl salt-sensitive rats; EM, epistatic module; GWAS, genome-wide association studies; MAP, mean arterial pressure; QTL, quantitative trait locus; SBP, systolic blood pressure

9.10 Supplementary data

Supplemental data Table 1: Summary of QTLs and their blood pressure effects defined by congenic strains

Chromosome	QTL names	Congenic strains	DSS allele effect	References
1	C1QTL1 C1QTL2 C1QTL3	C1S.L4 C1S.L3 C1S.L2	BP-decreasing BP-decreasing BP-decreasing	Fig. 1, current work
2	C2QTL4L C2QTL1 C2QTL2 C2QTL3 C2QTL4M C2QTL5 C2QTL6	C2S.L4 C2S.M7 C2S.M19 C2S.M15 C2S.M18 C2S.M21 C2S.M24	BP-increasing BP-decreasing BP-decreasing BP-decreasing BP-decreasing BP-decreasing BP-decreasing	(1-3)
3	C3QTL1/- BPQTL C3QTL2/+ BPQTL C3QTL3	C3S.L3/S.L3 C3S.L4/S.L4 C3S.L6	BP-decreasing BP-increasing BP-decreasing	(4) (5)
5		C5S.L1	No effect	Fig. 1, current work
7	C7QTL	C7S.L1	BP-decreasing	(6)
8	C8QTL1/- BP QTL C8QTL2/+ BP QTL	C8S.L1 C8S.L2	BP-decreasing BP-increasing	(7)
10	C10QTL1 C10QTL2 C10QTL3 C10QTL4 C10QTL5	C10S.L30 C10S.L16 C10S.L25 C10S.L28 C10S.L29	BP-decreasing BP-decreasing BP-decreasing BP-decreasing BP-decreasing	(8)
16	C16QTL	C16S.L7	BP-decreasing	(5) and current work
17	C17QTL1 C17QTL2	C17S.L11 C17S.L9	BP-decreasing BP-decreasing	(9) and current work

18	C18QTL1 C18QTL2 C18QTL3	C18S.L4 C18S.L3 C18S.L2	BP-decreasing BP-decreasing BP-decreasing	(10)
----	-------------------------------	-------------------------------	---	------

Legend to Supplemental data Table 1: The extent of their BP effects is shown in Figure 1 in the text. Others QTLs in the DSS genome were not tested because they did not appear in the genetic contrast between DSS and Lewis [or the Milan normotensive (MNS)] rats from which current congenic strains were made. These additional QTLs are summarized in Tables 1 and 2 (11) and located on Chromosomes 9, 12, 13 and 15.

Supplemental data Table 2. Correspondence of human blood pressure quantitative trait loci (QTLs) and their rat homologues

Human QTL interval	Human SNPs	Rat Chr/ Congenics	Comments References	Epistatic module (EM) #
<u><i>FURIN-FES</i></u>	rs2521501	1 C1S.L2	Included in Fig. 1(a), (12)	EM1 or EM2
<u><i>ADM</i></u>	rs7129220	1 C1S.L2	Included in Fig. 1(a), (12)	EM1 or EM2
<u><i>PLEKHA7</i></u>	rs381815	1 C1S.L2	Included in Fig. 1(a), (12)	EM1 or EM2
<u><i>PLCE1</i></u>	rs932764	1 C1S.L2	Included in Fig. 1(a), (12)	EM1 or EM2
<i>MOV10</i>	rs2932538	2 C2S.M	Included in Fig. 1(e), (1,12)	undefined
<i>CAPZA1</i>	rs17030613	2 C2S.M	Included in Fig. 1(e), (1,13)	undefined
<u><i>JAG1</i></u>	rs1327235	3 C3S.L7	Included in Fig. 1(b), (5,12)	EM1
<u><i>GNAS-EDN3</i></u>	rs6015450	3 C3S.L7	Included in Fig. 1(b), (5,12)	EM1
<i>FIGN/GRB14</i> (MAP)	rs1446468 rs16849225	3 C3S.L2	Included in Fig. 1(b), (4,13,14)	undefined
<u><i>NOV</i></u> (PP)	rs2071518	7 C7S.L	Included in Fig. 1(d), (6,14)	EM1 or EM2
<u><i>CYP1A2</i></u>	rs1378942	8	Included in Fig. 1(i), (7,12)	EM1
<u><i>CSK-ULK3</i></u>	rs6495122	C8S.L1		
<u><i>FLJ32810-</i></u> <u><i>TMEM133</i></u>	rs633185	8 C8S.L2	Included in Fig. 1(i), (7,12)	EM2

<u>ADAMTS8</u>	rs11222084	8	Included in Fig. 1(i), (7,14)	EM2
(PP)		C8S.L2		
		17	Included in Fig. 1h (9,12)	EM2
<u>HFE</u>	rs1799945	C17S.L9		
<i>GUCY1A3-</i>	rs13139571	2	Excluded, Fig. 1(e), (2,12)	
<i>GUCY1B3</i>		C2S.M12		
<i>MECOM</i>	rs419076	2	Excluded, Fig. 1(e), (2,12)	
		S.M2		
<i>ENPEP</i>	rs6825911	2	Excluded, Fig. 1(e), (2,13)	
		C2S.M10		
<i>SLC39A8</i>	rs13107325	2	Excluded, Fig. 1(e), (2,12)	
		C2S.M10		
<i>NPR3-C5orf23</i>	rs1173771	2	Excluded, Fig. 1(e), (3,12,13)	
	rs1173766	C2S.L3		
<i>MTHFR</i>	rs17367504	5	Excluded, Fig. 1(c), (12)	
		C5S.L		
<i>CASZ1</i>	rs880315	5	Excluded, Fig. 1(c), (13)	
		C5S.L		
<i>ZNF652</i>	rs16948048	10	Excluded, Fig. 1(f), (12,15)	
		C10S.L12		
<i>PLCD3</i>	rs12946454	10	Excluded, Fig. 1(f), (12,15)	
		C10S.L8		
<i>GOSR2</i>	rs17608766	10	Excluded, Fig. 1(f), (12,15)	
		C10S.L8		
<i>CACNB2</i>	rs11014166	17	Excluded, Fig. 1(h), (9,12)	
	rs4373814	C17S.L3		
<i>CYP17A1</i>	rs1119154,	1	(12)	Untested
	rs1004467	Non		
	rs11191548			

<i>ADRB1</i> (MAP)	rs2782980	1 Non	(14)	Untested
<i>CNNM2/</i> <i>NT5C2</i>	rs11191548	1 Non	(13)	Untested
<i>PIK3CG</i> (PP)	rs17477177	6 Non	(14)	Untested
<i>ATP2B1</i>	rs2681492/ rs2681472 rs17249754	7 Non	(12)	Untested
<i>ULK4</i>	rs9815354	8 Non	(12)	Untested
<i>MAP4</i> (MAP)	rs319690	8 Non	(14)	Untested
<i>EBF1</i>	rs11953630	10 Non	(12)	Untested
<i>STL7</i>	rs17030613	11 None	(13)	Untested
<i>RPL6/PTPN11/</i> <i>ALDH2</i>	rs11066280	12 Non	(13)	Untested
<i>SH2B3</i>	rs653178, rs3184504	12 Non	(12)	Untested
<i>TBX3-TBX5</i>	rs2384550 rs35444	12 Non	(12)	Untested
<i>FGF5</i>	rs16998073	14 Non	(12)	Untested
<i>CHIC2</i> (PP)	rs871606	14 Non	(14)	Untested
<i>SLC4A7</i>	rs13082711	15 Non	(12)	Untested
<i>BAT2-BAT5</i>	rs805303	20 Non	(12)	Untested
<i>c10orf107</i>	rs1530440	No rat homologue	(12)	Untested

Legend to supplemental datat Table 2: The underlined intervals in bold emphasize the inclusion (i.e. included) of a human homologue in a congenic strain exhibiting a BP effect (Fig. 1 in the text). ‘Excluded’ refers to the human homologue in the rat that have been

excluded as a candidate gene because a congenic strain carrying it does not exhibit a BP effect. Epistatic modules (EM) are defined in Table 1 and summarized in Table 2 in the text. *ADAMTS8*, metallopeptidase with thrombospondin type 1 motif, 8; *ADM*, adrenomedullin; *ADRB1*, adrenergic, beta-1-, receptor; *ALDH2*, aldehyde dehydrogenase 2 family (mitochondrial); *ATP2B1*, ATPase, Ca⁺⁺ transporting, plasma membrane 1; *BAT2*(*PRRc2b*), proline-rich coiled-coil 2B; *BAT5* (*ABHD16A*), ab hydrolase domain containing 16A; *C5orf23*, chromosome 5 open reading frame 23; *C10orf107*, chromosome 10 open reading frame 107; *CACNB2*, calcium channel, voltage-dependent, beta 2 subunit; *CAPZA1*, capping protein (actin filament) muscle Z-line, alpha 1; *CASZ1*, castor zinc finger 1; *CHIC2*, cysteine-rich hydrophobic domain 2; *CNNM2*, cyclin M2; *CSK*, c-src tyrosine kinase; *CYP1A2*, cytochrome P450, family 1, subfamily a, polypeptide 2; *CYP17A*, cytochrome P450, family 17, subfamily a, polypeptide 1; *EBF1*, early B-cell factor 1; *EDN3*, endothelin 3; *ENPEP*, glutamyl aminopeptidase (aminopeptidase A); *FES*, feline sarcoma oncogene; *FGF5*, fibroblast growth factor 5; *FIGN*, fidgetin; *FLJ32810*(*ARHGAP42*), Rho GTPase activating protein 42 ; *FURIN*, paired basic amino acid cleaving enzyme; *GNAS*, GNAS (guanine nucleotide binding protein, alpha stimulating) complex locus; *GOSR2*, golgi SNAP receptor complex member 2; *GRB14*, growth factor receptor-bound protein 14; *GUCY1A3*, guanylate cyclase 1, soluble, alpha 3; *GUCY1B3*, guanylate cyclase 1, soluble, beta 3; *HFE*, hemochromatosis; *JAG1*, jagged 1; *MAP4*, microtubule-associated protein 4; *MECOM* , MDS1 and EVI1 complex locus; *MTHFR*, methylenetetrahydrofolate reductase (NAD(P)H); *MOV10*, Moloney leukemia virus 10, homolog (mouse); *NOV*, nephroblastoma overexpressed gene; *NPR3*, natriuretic peptide receptor C/guanylate cyclase C (atriuretic peptide receptor C); *NT5C2*, 5'-nucleotidase, cytosolic II; *PIK3CG*, phosphoinositide-3-kinase, catalytic, gamma polypeptide; *PLCD3*, phospholipase C, delta 3; *PLCE1*, phospholipase C, epsilon 1; *PLEKHA7*, pleckstrin homology domain containing, family A member 7; *PTPN11*, protein tyrosine phosphatase, non-receptor type 11; *RPL6*, ribosomal protein L6; *SH2B3*, SH2B adaptor protein 3; *SLC4A7*, solute carrier family 4, sodium bicarbonate cotransporter, member 7; *SLC39A8*, solute carrier family 39 (metal ion transporter), member 8; *STL7*, Serum triglyceride level QTL 7; *TBX3*, T-box 3; *TBX5*, T-box 5; *TMEM133*, transmembrane protein 133; *ULK3*, unc-51-like kinase 3; *ULK4*, unc-51-like kinase 4 (C. elegans); *ZNF652*, zinc finger protein 652. Chr, chromosome; PP, pulse pressure; MAP, mean arterial pressure.

Supplemental data Table 3 (a). Sequence alignment of *Adrenomedulin (Adm)* between Dahl salt-sensitive (DSS) and Lewis rats

```

DSS      GTTTCCTGGTGACACTAGGCAGACACTCCAGCCTTTACCGCTCCTGGTTTCTCGGCTTCT 60
Lewis    GTTTCCTGGTGACACTAGGCAGACACTCCAGCCTTTACCGCTCCTGGTTTCTCGGCTTCT 60
*****

DSS      CATCGCAGTCAGTCTTGGACTTTGCGGGTTTTGCCGCTGTCAGAAGGACGTCTCGGACTT 120
Lewis    CATCGCAGTCAGTCTTGGACTTTGCGGGTTTTGCCGCTGTCAGAAGGACGTCTCGGACTT 120
*****

          ↓

DSS      TCTGCTTCAAGTGCTTGACAACCTCACCTTTTCCAGCAGGtaggaaccg...tcctttccagGGT 180
Lewis    TCTGCTTCAAGTGCTTGACAACCTCACCTTTTCCAGCAGGtaggaaccg...tcctttccagGGT 180
*****

DSS      ATCGGAGCATCGCTACAGAAATCAAGCTGGTTTCCATCGCCCTGATGTTATTGGGTTTCGCT 240
Lewis    ATCGGAGCATCGCTACAGAAATCAAGCTGGTTTCCATCGCCCTGATGTTATTGGGTTTCGCT 240
*****

          ↓

DSS      CGCCGTTCTCGGCGCGGACACCGCACGGCTCGACACTTCCTCGCAGTTCGAAAGAAgtg 300
Lewis    CGCCGTTCTCGGCGCGGACACCGCACGGCTCGACACTTCCTCGCAGTTCGAAAGAAgtg 300
*****

DSS      agtttg...tgttttctagGTGGAATAAGTGGGCGCTAAGTCGTGGGAAGAGGGAAGTACAA 360
Lewis    agtttg...tgttttctagGTGGAATAAGTGGGCGCTAAGTCGTGGGAAGAGGGAAGTACAA 360
*****

DSS      GCGTCCAGCAGCTACCTACGGGCTCGTTGATGAGAAGACAGTCCCGACCCAGACTCTT 420
Lewis    GCGTCCAGCAGCTACCTACGGGCTCGTTGATGAGAAGACAGTCCCGACCCAGACTCTT 420
*****

          ↓

DSS      GGGCTCCAGGACAAGCAGAGCAGCTTAGCACCCCAAGCCAGgtaactaact...ttctcc 480
Lewis    GGGCTCCAGGACAAGCAGAGCAGCTTAGCACCCCAAGCCAGgtaactaact...ttctcc 480
*****

DSS      gcagCACTCAGAGCACAGCCACATTCGAGTCAAACGCTACCGCCAGAGCATGAACCAGG 540
Lewis    gcagCACTCAGAGCACAGCCACATTCGAGTCAAACGCTACCGCCAGAGCATGAACCAGG 540
*****

DSS      GGTCCCGCAGCACTGGATGCCGCTTTGGGACCTGCACAATGCAGAACTGGCTCACCAGA 600
Lewis    GGTCCCGCAGCACTGGATGCCGCTTTGGGACCTGCACAATGCAGAACTGGCTCACCAGA 600
*****

```

DSS TCTACCAGTTTACAGACAAAGACAAGGACGGCATGGCCCCAGAAACAAGATCAGCCCTC 660
 Lewis TCTACCAGTTTACAGACAAAGACAAGGACGGCATGGCCCCAGAAACAAGATCAGCCCTC 660

DSS AAGGCTATGGCCGCCGGCGCCGGCGTTCCCTGCCAGAGGTCTCCGAGCCCGGACTGTGG 720
 Lewis AAGGCTATGGCCGCCGGCGCCGGCGTTCCCTGCCAGAGGTCTCCGAGCCCGGACTGTGG 720

DSS AGTCTCCAGGAGCAGACACACTCAGCTCCAGCCTCCCCGGCGCACCAAGACATCTCCA 780
 Lewis AGTCTCCAGGAGCAGACACACTCAGCTCCAGCCTCCCCGGCGCACCAAGACATCTCCA 780

DSS GAGTCTCTAGGTTA **TAC**GTGCGGGTGGCAGCATTGAACAGTCGGGCGAGTATCCCATTGG 840
 Lewis GAGTCTCTAGGTTA **TAC**GTGCGGGTGGCAGCATTGAACAGTCGGGCGAGTATCCCATTGG 840

DSS CGCCTGCGGAATCAGAGAGCTTCGCACCCTG**A**GCGGACTGAGACAATCTTGCAAGATCT 900
 Lewis CGCCTGCGGAATCAGAGAGCTTCGCACCCTG**G**GCGGACTGAGACAATCTTGCAAGATCT 900

DSS GCCTGGCTGCCCTAGGGGAGGCAGAGGAACCCAAGATCAAGCCAGGCTCACGTCAGAAA 960
 Lewis GCCTGGCTGCCCTAGGGGAGGCAGAGGAACCCAAGATCAAGCCAGGCTCACGTCAGAAA 960

DSS CCGAGAATTACAGGCTGATACTCTCTCCGGGCAGGGGTCTGAGCCACTGCCTTGCCCCT 1020
 Lewis CCGAGAATTACAGGCTGATACTCTCTCCGGGCAGGGGTCTGAGCCACTGCCTTGCCCCT 1020

DSS CATAAACTGGTTTTCTCACGGGGCATAAGCCTCATTACTACTTGAACCTTCCAAAACCTA 1080
 Lewis CATAAACTGGTTTTCTCACGGGGCATAAGCCTCATTACTACTTGAACCTTCCAAAACCTA 1080

DSS GCGAGGAAAAGTGCAATGCTTGTTATACAGCCAAAGGTAACATATATTTAAGTTTGTT 1140
 Lewis GCGAGGAAAAGTGCAATGCTTGTTATACAGCCAAAGGTAACATATATTTAAGTTTGTT 1140

DSS GATGTCAAGAGGTTTTTTTTT 1162
 Lewis GATGTCAAGAGGTTTTTTTTT 1162

Supplemental data Table 3(b). Sequence alignment of *Endothelin3 (Edn3)* among Dahl salt-sensitive (DSS), Lewis and Dahl salt-resistant (DSR) rats

```

DSS      GGGCAGCTCGCTCTGAAAGTTCGTGACCGCCTCAGCCAACCTAACTCTGAGCCCTGGGACG 60
Lewis    GGGCAGCTCGCTCTGAAAGTTCGTGACCGCCTCAGCCAACCTAACTCTGAGCCCTGGGACG 60
DSR      GGGCAGCTCGCTCTGAAAGTTCGTGACCGCCTCAGCCAACCTAACTCTGAGCCCTGGGACG 60
*****

DSS      CCCAGTTCAGGCAGCGGCAGGACTCGAAAGCTGTACCCAGTCTCACTACCCTTTTGCGGT 120
Lewis    CCCAGTTCAGGCAGCGGCAGGACTCGAAAGCTGTACCCAGTCTCACTACCCTTTTGCGGT 120
DSR      CCCAGTTCAGGCAGCGGCAGGACTCGAAAGCTGTACCCAGTCTCACTACCCTTTTGCGGT 120
*****

DSS      CACAAGCGGCCACCCTCCAGGCCCGGTGCTCCCGCGCTGATCGGGGTTCAATGGAGCTGG 180
Lewis    CACAAGCGGCCACCCTCCAGGCCCGGTGCTCCCGCGCTGATCGGGGTTCAATGGAGCTGG 180
DSR      CACAAGCGGCCACCCTCCAGGCCCGGTGCTCCCGCGCTGATCGGGGTTCAATGGAGCTGG 180
*****

                                     ↓

DSS      GGCTGTGGCTCCTTCTCGGGCTCACAGTGACCTCCGCTGCAGgtaagcagga...tccccaac 240
Lewis    GGCTGTGGCTCCTTCTCGGGCTCACAGTGACCTCCGCTGCAGgtaagcagga...tccccaac 240
DSR      GGCTGTGGCTCCTTCTCGGGCTCACAGTGACCTCCGCTGCAGgtaagcagga...tccccaac 240
*****

DSS      agCTGCACTGCCTGCACAGCCTGAAATGCTGGGCAGGAGCGGGGACCAGGCAGGTCTGG 300
Lewis    agCTGCACTGCCTGCACAGCCTGAAATGCTGGGCAGGAGCGGGGACCAGGCAGGTCTGG 300
DSR      agCTGCACTGCCTGCACAGCCTGAAATGCTGGGCAGGAGCGGGGACCAGGCAGGTCTGG 300
*****

DSS      GGACCAGGAGGAAAAGAGGGTGCCTGCACACCACCGACCTCGGCGCTGCACGTGCTTCAC 360
Lewis    GGACCAGGAGGAAAAGAGGGTGCCTGCACACCACCGACCTCGGCGCTGCACGTGCTTCAC 360
DSR      GGACCAGGAGGAAAAGAGGGTGCCTGCACACCACCGACCTCGGCGCTGCACGTGCTTCAC 360
*****

DSS      TTATAAGGACAAGGAGTGTGTCTACTACTGCCACCTGGACATCATCTGGATCAACTCC 420
Lewis    TTATAAGGACAAGGAGTGTGTCTACTACTGCCACCTGGACATCATCTGGATCAACTCC 420
DSR      TTATAAGGACAAGGAGTGTGTCTACTACTGCCACCTGGACATCATCTGGATCAACTCC 420
*****

                                     ↓

DSS      Tgagtgagttgg...tactctgcagACAGACTGTGCCCTATGGACTGTCCAACCACAGAGGAA 480
Lewis    Tgagtgagttgg...tactctgcagACAGACTGTGCCCTATGGACTGTCCAACCACAGAGGAA 480
DSR      Tgagtgagttgg...tactctgcagACAGACTGTGCCCTATGGACTGTCCAACCACAGAGGAA 480
*****

```

```

DSS      GCCTCCGGGGAAAGAGGTCTCGGGGCCAGTTCAGAAAAGTCCCAGTCTTCTCCACAGA 540
Lewis    GCCTCCGGGGAAAGAGGTCTCGGGGCCAGTTCAGAAAAGTCCCAGTCTTCTCCACAGA 540
DSR      GCCTCCGGGGAAAGAGGTCTCGGGGCCAGTTCAGAAAAGTCCCAGTCTTCTCCACAGA 540
*****

DSS      CACGCTTACGTTGTGCCTGTGCAGGGGTGGATGACAAGGCCTGCGCATACTTCTGTGCAC 600
Lewis    CACGCTTACGTTGTGCCTGTGCAGGGGTGGATGACAAGGCCTGCGCATACTTCTGTGCAC 600
DSR      CACGCTTACGTTGTGCCTGTGCAGGGGTGGATGACAAGGCCTGCGCATACTTCTGTGCAC 600
*****

          ↓

DSS      ACGTCACCAGgtatgacaca...gtccccacagTTATTCCAGGAGAGCAGAAAAGGCAGCTGC 660
Lewis    ACGTCACCAGgtatgacaca...gtccccacagTTATTCCAGGAGAGCAGAAAAGGCAGCTGC 660
DSR      ACGTCACCAGgtatgacaca...gtccccacagTTATTCCAGGAGAGCAGAAAAGGCAGCTGC 660
*****

DSS      AGAAGAGAAGCAGGAGACTGGAGGTCCACGTCAAAGgtgagggcag...cttgaatagGCTG 720
Lewis    AGAAGAGAAGCAGGAGACTGGAGGTCCACGTCAAAGgtgagggcag...cttgaatagGCTG 720
DSR      AGAAGAGAAGCAGGAGACTGGAGGTCCACGTCAAAGgtgagggcag...cttgaatagGCTG 720
*****

DSS      AAGTCAAGGACAGATAAAGTCCACCAGCCTTAGCTGGCTTTATCGGACCACAACCTGATGC 780
Lewis    AAGTCAAGGACAGATAAAGTCCACCAGCCTTAGCTGGCTTTATCGGACCACAACCTGATGC 780
DSR      AAGTCAAGGACAGATAAAGTCCACCAGCCTTAGCTGGCTTTATCGGACCACAACCTGATGC 780
*****

          ↓

DSS      TTCTTGCTTCCTGCGGTGGATTTCCTCCGGCTCTCCctgacctgccc 826
Lewis    TTCTTGCTTCCTGCGGTGGATTTCCTCCGGCTCTCCctgacctgccc 826
DSR      TTCTTGCTTCCTGCGGTGGATTTCCTCCGGCTCTCCctgacctgccc 826
*****

```

Supplemental data Table 3(c). Sequence alignment of *Hemochromatosis (Hfe)* between Dahl salt-sensitive (DSS) and Lewis rats

```

DSS      gtttgaagccTCAGCAATGGCTACAGGGTGAAGTCTTGGATCCTCCACGTTTCCAGGTCC 60
Lewis    gtttgaagccTCAGCAATGGCTACAGGGTGAAGTCTTGGATCCTCCACGTTTCCAGGTCC 60
*****

DSS      TAGTGAAAACCGGTGGACCCAGCTGGAGGCATGACCGATCAGCTGGGCTCCCTGTGCGG 120
Lewis    TAGTGAAAACCGGTGGACCCAGCTGGAGGCATGACCGATCAGCTGGGCTCCCTGTGCGG 120
*****

          ↓

DSS      CTGCTATTGCTGCTGCTGTTGTTGCTGCTGTGGTCCGTGGCCCCGAGGCGCTGCGGCC 180
Lewis    CTGCTATTGCTGCTGCTGTTGTTGCTGCTGTGGTCCGTGGCCCCGAGGCGCTGCGGCC 180
*****

```

DSS ggtgagtgcg...gacttccccagGTTACATTCTCTACGATATCTCTTCATGGGTGCCTCAA 240
Lewis ggtgagtgcg...gacttccccagGTTACATTCTCTACGATATCTCTTCATGGGTGCCTCAA 240

DSS AGCCAGACCTCGGGCTGCCTTTCTTTGAGGCTCTGGGTTATGTGGATGACCAGCTCTTTG 300
Lewis AGCCAGACCTCGGGCTGCCTTTCTTTGAGGCTCTGGGTTATGTGGATGACCAGCTCTTTG 300

DSS TATCCTACAATCACGAGAGTCGCCGTGCTGAGCCCAGGGCCCCATGGATCTTGGGGCAGA 360
Lewis TATCCTACAATCACGAGAGTCGCCGTGCTGAGCCCAGGGCCCCATGGATCTTGGGGCAGA 360

DSS CATCAAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAAAGGGTGGGATTACATGTTCA 420
Lewis CATCAAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAAAGGGTGGGATTACATGTTCA 420

↓

DSS TAGTGGACTTCTGGACCATCATGGGCAACTATAACCACAGTAAGGgtatgtagag...tcatt 480
Lewis TAGTGGACTTCTGGACCATCATGGGCAACTATAACCACAGTAAGGgtatgtagag...tcatt 480

DSS tccagTCACGAAGTTGAGAGTGGTGCCTGAGTCTCACATCCTGCAGGTGATCCTAGGATG 540
Lewis tccagTCACGAAGTTGAGAGTGGTGCCTGAGTCTCACATCCTGCAGGTGATCCTAGGATG 540

DSS TGAGGTGCATGAAGACAACAGTACCAGTGGCTTCTGGAAATATGGCTACGATGGGCAAGA 600
Lewis TGAGGTGCATGAAGACAACAGTACCAGTGGCTTCTGGAAATATGGCTACGATGGGCAAGA 600

DSS TCACCTTGAATTCTGCCCCAAGACACTGAACTGGAGTGCAGCCGAGCCAAGGGCCTGGGC 660
Lewis TCACCTTGAATTCTGCCCCAAGACACTGAACTGGAGTGCAGCCGAGCCAAGGGCCTGGGC 660

DSS CACCAAGATGGAGTGGGAAGAGCACAGGATCCGTGCCAGACAGAGCAGGGACTACCTGCA 720
Lewis CACCAAGATGGAGTGGGAAGAGCACAGGATCCGTGCCAGACAGAGCAGGGACTACCTGCA 720

DSS GAGGGACTGCCCCAGCAGCTGAAGCAGGTCTGGAGCTCCAGAGAGGGGTCTGGGACA 780
Lewis GAGGGACTGCCCCAGCAGCTGAAGCAGGTCTGGAGCTCCAGAGAGGGGTCTGGGACA 780

↓

DSS GCAAGgtacacggaa...ttccttcaagTGCCTACTTTGGTAAAAGTGACTCGCCACTGGGCC 840
Lewis GCAAGgtacacggaa...ttccttcaagTGCCTACTTTGGTAAAAGTGACTCGCCACTGGGCC 840

DSS TCTACAGGGACCTCTCTAAGGTGTCAGGCTCTGAATTTCTTCCCCCAGAACATCACTATG 900
 Lewis TCTACAGGGACCTCTCTAAGGTGTCAGGCTCTGAATTTCTTCCCCCAGAACATCACTATG 900

DSS AGGTGGTTGAAGGACAGCCAGCCCCTAGATGCCAAGGATGTCAACCCTGAGAACGTGCTG 960
 Lewis AGGTGGTTGAAGGACAGCCAGCCCCTAGATGCCAAGGATGTCAACCCTGAGAACGTGCTG 960

DSS CCAAATGGGGATGGGACCTATCAGGGCTGGCTGACCTTGGCTGTGGCCCTGGAGAAGAG 1020
 Lewis CCAAATGGGGATGGGACCTATCAGGGCTGGCTGACCTTGGCTGTGGCCCTGGAGAAGAG 1020

DSS ACAAGGTTCAAGTGGAGCACCAGGCCTGGATCAGCCTCTCACTGCCACTTGG 1080
 Lewis ACAAGGTTCAAGTGGAGCACCAGGCCTGGATCAGCCTCTCACTGCCACTTGG 1080



DSS Ggtgagacgg...ttgttttagAGCCCTCACGGTCTCAGGACATGATTATTGGAATCATAA 1140
 Lewis Ggtgagacgg...ttgttttagAGCCCTCACGGTCTCAGGACATGATTATTGGAATCATAA 1140

DSS GTGGGATCACCATTTGTGCCATCTTCTTTGTTGGAATTCTGATCCTAGTCTTAAGGAAAA 1200
 Lewis GTGGGATCACCATTTGTGCCATCTTCTTTGTTGGAATTCTGATCCTAGTCTTAAGGAAAA 1200

DSS GGAAGTTTCAGgtgagtggga...atcttcccagGAGGAACCATGGGTGACTATGTCTTAAC 1260
 Lewis GGAAGTTTCAGgtgagtggga...atcttcccagGAGGAACCATGGGTGACTATGTCTTAAC 1260

DSS AGAGTGTGAGTGA CCTGCAGCATGCAGAAGCACAGAAGAGAGAAGACTCAGCCAAAGACT 1320
 Lewis AGAGTGTGAGTGA CCTGCAGCATGCAGAAGCACAGAAGAGAGAAGACTCAGCCAAAGACT 1320

DSS TGGAGGGGACACACTTGCTCCATTCTAGAACACAGCTGGACCTAACACACAGAACTGCC 1380
 Lewis TGGAGGGGACACACTTGCTCCATTCTAGAACACAGCTGGACCTAACACACAGAACTGCC 1380

DSS TGAGGACTCTGCCCTTAGCTTTCCTGTTTGCTTTCTTAAGGTGTTTTCTCCAGTTAAGT 1440
 Lewis TGAGGACTCTGCCCTTAGCTTTCCTGTTTGCTTTCTTAAGGTGTTTTCTCCAGTTAAGT 1440

DSS TCAGTTCCTGAATAATAGTGACTGCCCCAGCTGCAACCTCTCCCTTCAGAACCAGTCTCA 1500
 Lewis TCAGTTCCTGAATAATAGTGACTGCCCCAGCTGCAACCTCTCCCTTCAGAACCAGTCTCA 1500

```

DSS      TGATCTTTAAGCTGCTACTTGCAGGCATCCTTCGTTTTCTGCATCCACCTAGACTTCGTA 1560
Lewis    TGATCTTTAAGCTGCTACTTGCAGGCATCCTTCGTTTTCTGCATCCACCTAGACTTCGTA 1560
          *****

DSS      TGTCTACTTAAAAAGCCCCACTAAATTTGGGGGACACATGATTCATTTCCACATCTGAAG 1620
Lewis    TGTCTACTTAAAAAGCCCCACTAAATTTGGGGGACACATGATTCATTTCCACATCTGAAG 1620
          *****

DSS      AAGTTATGAACCTTCATCCTGGGATGCACACATTCTTGTGCCAGAATTTTTCATACATAT 1680
Lewis    AAGTTATGAACCTTCATCCTGGGATGCACACATTCTTGTGCCAGAATTTTTCATACATAT 1680
          *****

DSS      CCTAGGACCCATTCAATTGTCATTTGAGCCTCTCTATCTGTAGTACTACTCTGACTTC 1740
Lewis    CCTAGGACCCATTCAATTGTCATTTGAGCCTCTCTATCTGTAGTACTACTCTGACTTC 1740
          *****

          ↓

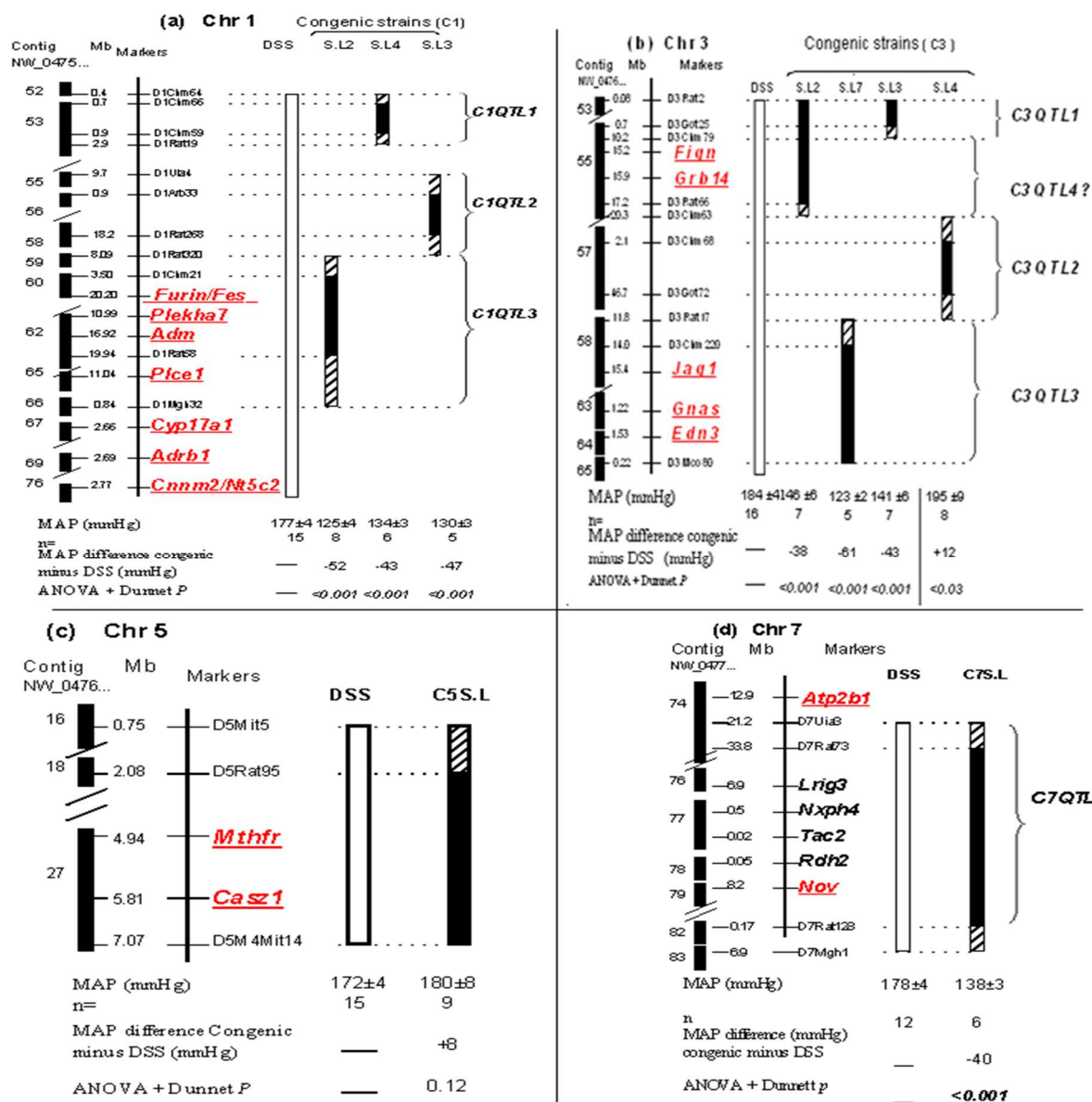
DSS      TCTGCCATTGGAGTGTATGGCAATAAAGCTATGAACGTTacacactgtg 1790
Lewis    TCTGCCATTGGAGTGTATGGCAATAAAGCTATGAACGTTacacactgtg 1790
          *****

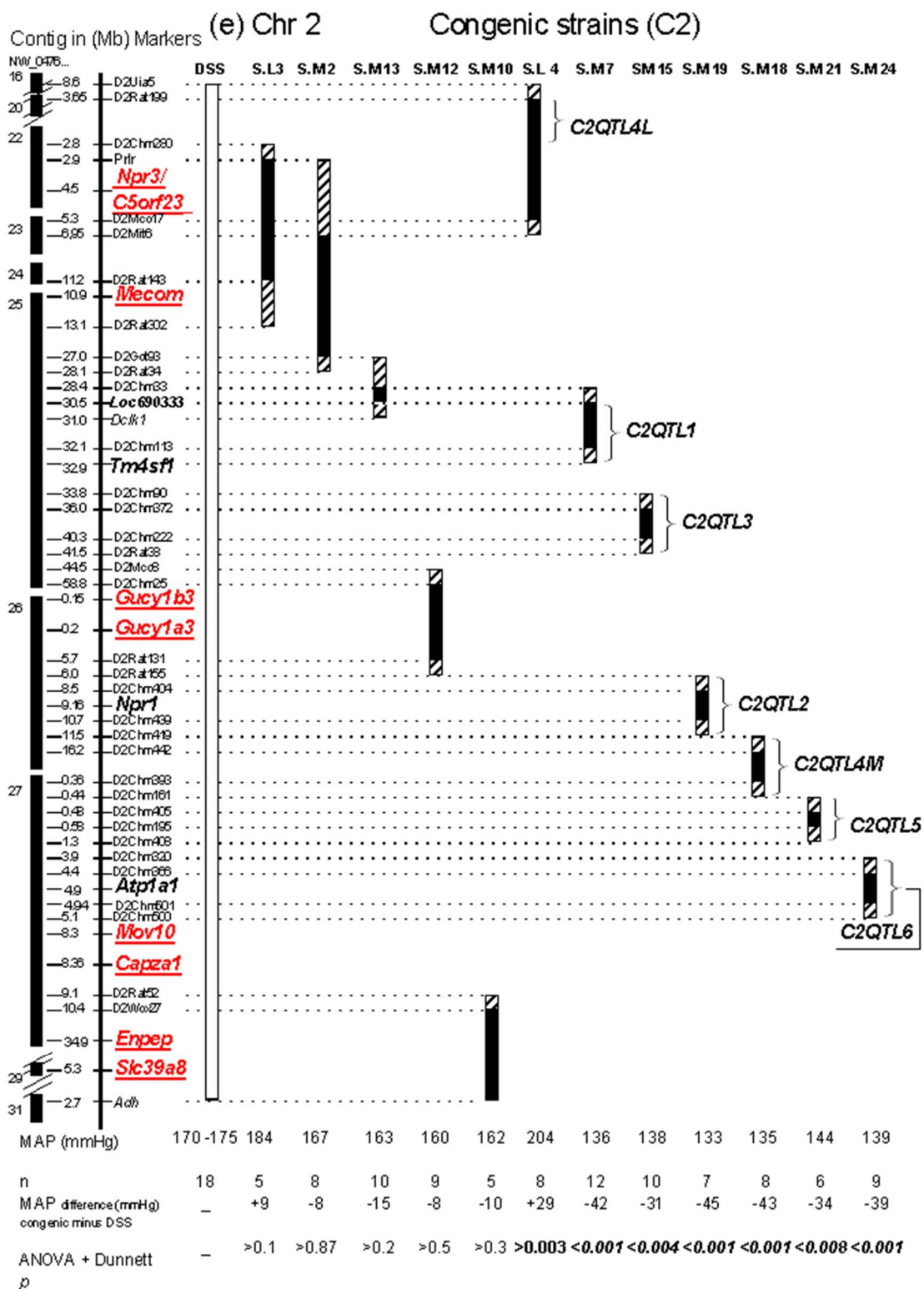
```

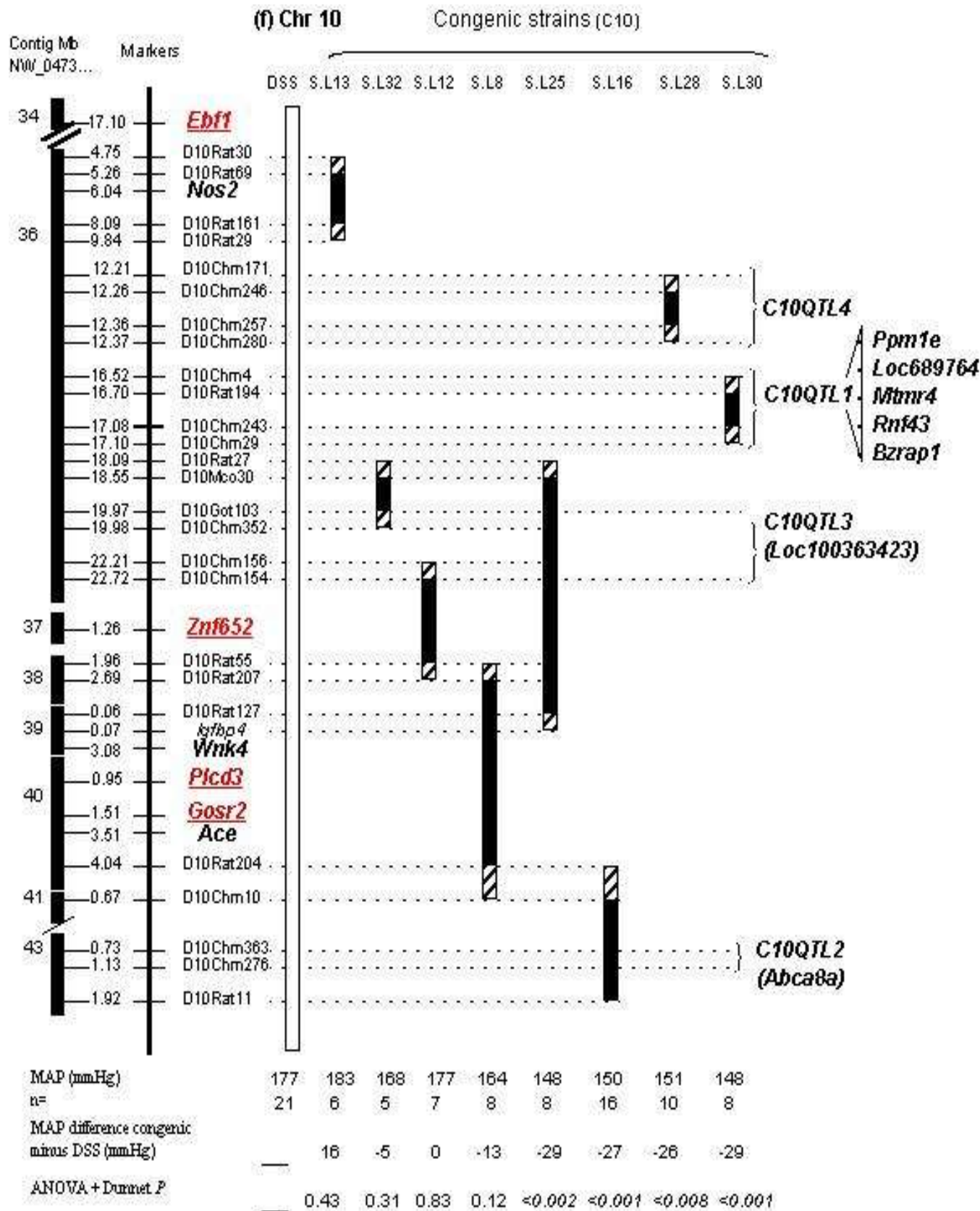
Legend to Supplemental data Table 3: * indicates nucleotide identity. gDNAs were first amplified by PCR and then sequenced. Nucleotides of the mRNA are in capital letters. The intro/exon junctions (10 pb for each side and separated by“...” for each exon) are given in minuscules. ↓ marks the last nucleotide of each exon. Initiation and stop codons are marked in blocks.

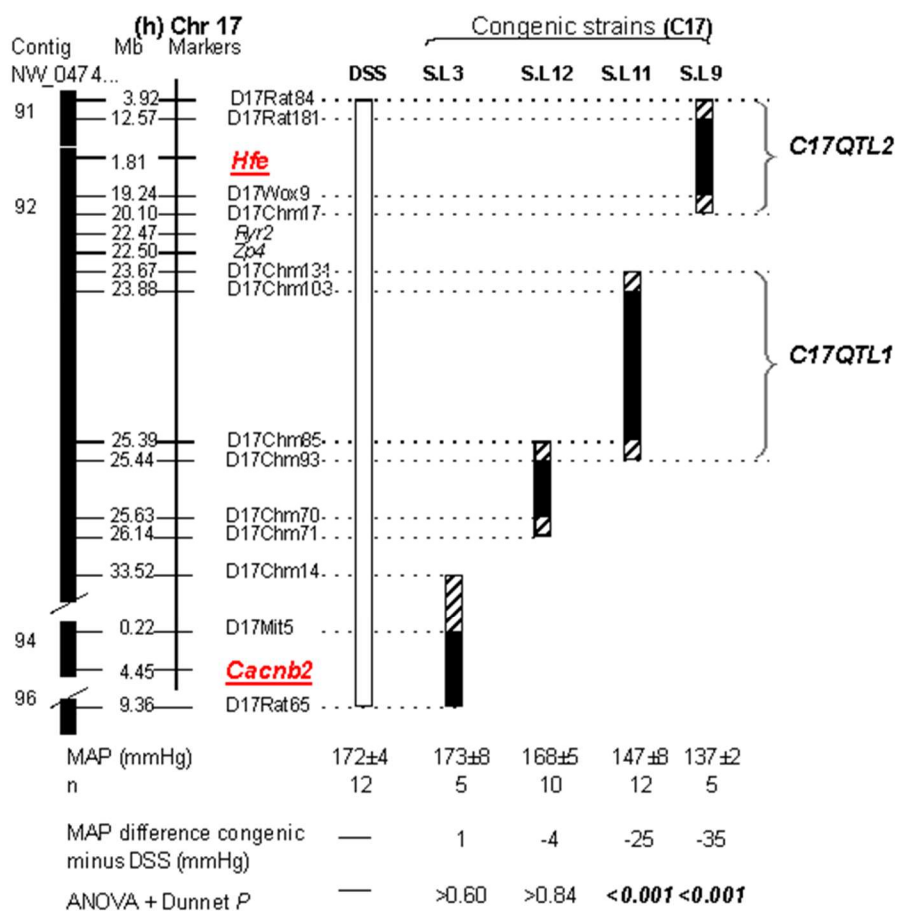
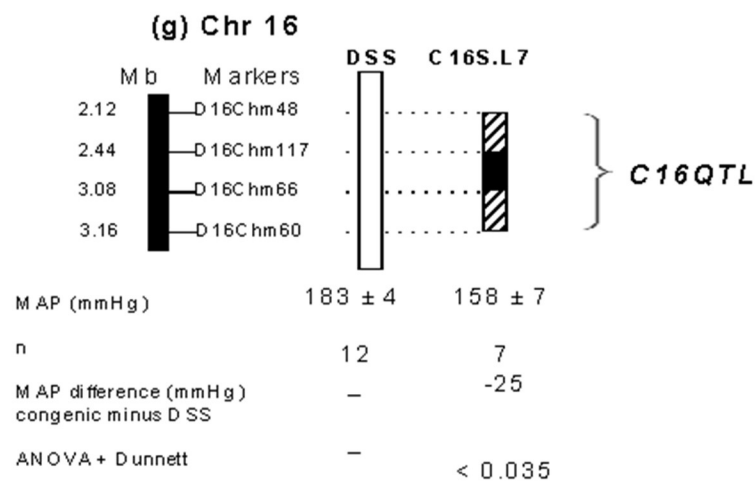
Supplemental data Figure 1. Definition of QTLs for blood pressure by congenic strains. A solid bar under congenic strains symbolizes the DSS fragment (a white bar) that has been replaced by that of Lewis (S.L) or MNS (S.M). Striped bars on ends of the solid bars denote the ambiguity of crossover breakpoints between markers. Mean arterial pressures (MAPs) for all the strains are averages for the duration of the measurement. \pm SEM of MAPs for the congenic strains on Chromosomes 2 and 10 ranges from 2 to 5 and there are not included in the figure for a lack of space. Systolic and diastolic arterial pressures are consistent with their MAPs of all the strains (data not shown). Genes **underlined** denote the positions for the rat homologues of human BP QTLs (12-14), and their full gene names are given in the legend for Supplemental Table 2. C3QTL4? is only implicated from a conserved homology between the rat and human QTLs. Rat genes of biological interest are named in the legend for Supplemental Tables 1 and 2 except for *Ace*, angiotensin I converting enzyme; *Adrb2*, adrenergic, beta-2-, receptor, surface; *Alpk2*, alpha-kinase 2; *Atp1a1*, ATPase, Na⁺/K⁺ transporting, alpha 1 polypeptide; *Nedd4l*, neural precursor cell expressed, developmentally down-regulated 4-like; *Nos2*, inducible nitric oxide synthase2; *Nxph4*, neurexophilin 4; *Npr1*, natriuretic peptide receptor A/guanylate cyclase A; *Rdh2*, retinol dehydrogenase 2; *Tac2*, tachykinin 2; *Wnk4*, WNK lysine deficient protein kinase 4. Newly-produced congenic strains are C1S.L2, C1S.L3, C1S.L4, C5S.L, C16S.L7, C17S.L9, C17S.L11, and C17S.L12. The rest of the congenic strains are as reported previously (1,2,4-8,10,15-17). DSS, Dahl salt-sensitive rat; MNS, Milan normotensive rat; n, number of animals used. Chr, chromosome.

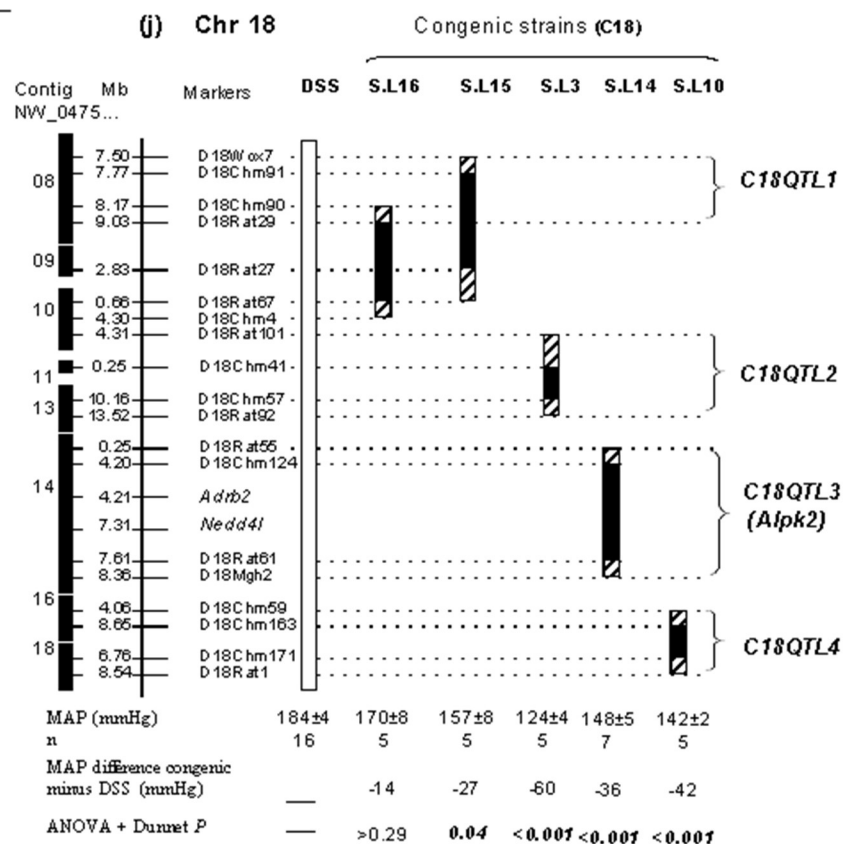
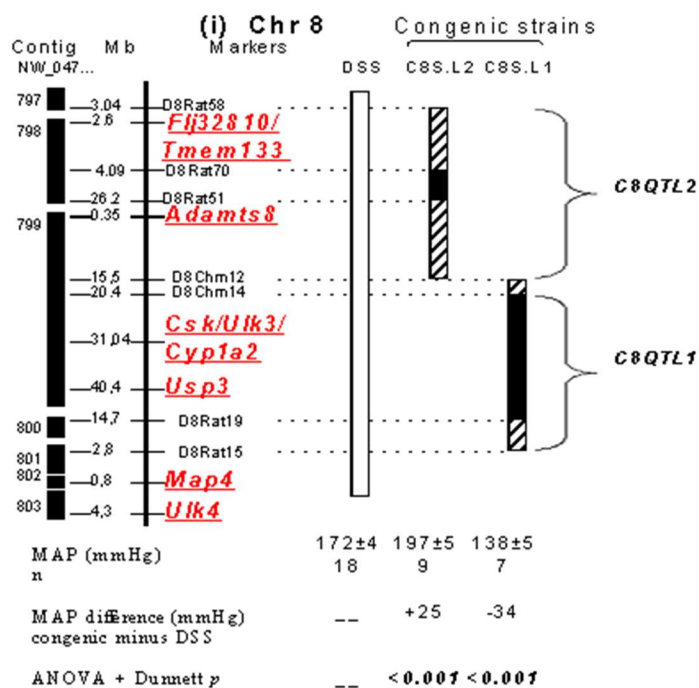
Figure 1











Reference List

1. Chauvet,C., Menard,A., Tremblay,J., Xiao,C., Shi,Y., L'heureux,N., Cardin,S., Tardif,J.C., Nattel,S., and Deng,A.Y. (2009) Cardiac pathways distinguish two epistatic modules enacting BP quantitative trait loci and candidate gene analysis. *Hypertens.Res.*, **32**, 631-637.
2. Dutil,J., Eliopoulos,V., Tremblay,J., Hamet,P., Charron,S., and Deng,A.Y. (2005) Multiple quantitative trait loci for blood pressure interacting epistatically and additively on dahl rat chromosome 2. *Hypertension*, **45**, 557-564.
3. Eliopoulos,V., Dutil,J., Deng,Y., Grondin,M., and Deng,A.Y. (2005) Severe hypertension caused by alleles from normotensive Lewis for a quantitative trait locus on chromosome 2. *Physiol Genomics*, **22**, 70-75.
4. Palijan,A., Dutil,J., and Deng,A.Y. (2003) Quantitative trait loci with opposing blood pressure effects demonstrating epistasis on Dahl rat chromosome 3. *Physiol. Genomics*, **15**, 1-8.
5. Duong,C., Charron,S., Xiao,C., Hamet,P., Menard,A., Roy,J., and Deng,A.Y. (2006) Distinct quantitative trait loci for kidney, cardiac, and aortic mass dissociated from and associated with blood pressure in Dahl congenic rats. *Mamm. Genome*, **17**, 1147-1161.
6. Crespo,K., Chauvet,C., Blain,M., Menard,A., Roy,J., and Deng,A.Y. (2011) Normotension in Lewis and Dahl salt-resistant rats is governed by different genes. *J.Hypertens.*, **29**, 460-465.
7. Ariyarajah,A., Palijan,A., Dutil,J., Prithiviraj,K., Deng,Y., and Deng,A.Y. (2004) Dissecting quantitative trait loci into opposite blood pressure effects on Dahl rat chromosome 8 by congenic strains. *J.Hypertens*, **22**, 1495-1502.
8. Chauvet,C., Charron,S., Menard,A., Xiao,C., Roy,J., and Deng,A.Y. (2008) Submegabase resolution of epistatically interacting quantitative trait loci for blood pressure applicable for essential hypertension. *J. Hypertens*, **26**, 893-901.
9. Grondin,M., Eliopoulos,V., Lambert,R., Deng,Y., Ariyarajah,A., Moujahidine,M., Dutil,J., Charron,S., and Deng,A.Y. (2005) Complete and overlapping congenics proving the existence of a quantitative trait locus for blood pressure on Dahl rat chromosome 17. *Physiol. Genomics*, **21**, 112-116.
10. Charron,S., Lambert,R., Eliopoulos,V., Duong,C., Menard,A., Roy,J., and Deng,A.Y. (2005) A loss of genome buffering capacity of Dahl salt-sensitive model to modulate blood pressure as a cause of hypertension. *Hum.Mol.Genet.*, **14**, 3877-3884.

11. Deng,A.Y. (2007) Genetics of polygenic hypertension from animal models to humans. *Curr. Hyp. Rev*, **3**, 284-297.
12. The International Consortium for blood pressure (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*, **478**, 103-109.
13. Kato,N., Takeuchi,F., Tabara,Y., Kelly,T.N., Go,M.J., Sim,X., Tay,W.T., Chen,C.H., Zhang,Y., Yamamoto,K., et al. (2011) Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat Genet*, **43**, 531-538.
14. Wain,L.V., Verwoert,G.C., O'Reilly,P.F., Shi,G., Johnson,T., Johnson,A.D., Bochud,M., Rice,K.M., Henneman,P., Smith,A.V., et al. (2011) Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet*, **43**, 1005-1011.
15. Palijan,A., Lambert,R., Dutil,J., Sivo,Z., and Deng,A.Y. (2003) Comprehensive congenic coverage revealing multiple blood pressure quantitative trait loci on Dahl rat chromosome 10. *Hypertension*, **42**, 515-522.
16. Chauvet,C., Crespo,K., Menard,A., Wu,Y., Xiao,C., Blain,M., Roy,J., and Deng,A.Y. (2011) alpha-Kinase 2 is a novel candidate gene for inherited hypertension in Dahl rats. *J.Hypertens.*, **29**, 1320-1326.
17. Chauvet,C., Menard,A., Xiao,C., Aguila,B., Blain,M., Roy,J., and Deng,A.Y. (2012) Novel genes as primary triggers for polygenic hypertension. *J.Hypertens*, **30**, 81-86.

9.11 Addendum

Errata:

L'affirmation «less than 1% of BP changes » dans le deuxième paragraphe de la section 9.2 (p.50) devrait être changée par « less than 1% of BP variance », qui aurait été plus acceptable.

L'affirmation « *If such a 'double' congenic strain has the same BP as each of the 2 'single' congenic strains (...)* » (deuxième paragraphe, section 9.3.1, p.51) n'est pas précise. Une définition plus générale et adéquate serait de dire que lorsque l'effet phénotype observé chez une « double congénique » est significativement moins important que la somme prédite de l'additivité mathématique des deux QTL, on peut parler de l'existence d'interactions épistatiques.

Limitations de l'étude :

En vue de tester l'existence d'interactions épistatiques entre les QTL, des analyses de variance à deux critères de classification permettent d'évaluer si l'effet phénotypique observé chez une souche « double-congénique » est significativement différent de celui que prédit la somme arithmétique de leurs effets individuels. Ainsi, l'hypothèse nulle décrit que la différence entre l'effet observé et la somme prédite des effets est égale à 0, et peut être rejetée sous le modèle additif avec un seuil de signification de 5%.

On conclue l'absence d'interactions épistatiques lorsque $p > 0,05$. Cependant, sans avoir défini la puissance statistique du test, quand $p > 0,05$, on ne peut pas rejeter l'hypothèse nulle sans pour autant pouvoir tirer d'autres conclusions sur l'hypothèse alternative [152].

En définissant puis en augmentant la puissance statique du test on augmente les probabilités de détecter un effet d'ampleur minimale mais biologiquement significatif; cependant, ceci est fait au détriment du seuil de signification. Pour surmonter ces contraintes, il faudrait augmenter le nombre n d'individus de chaque souche étudiée [153].

10. Alpha kinase 2 is a novel candidate gene for inherited hypertension in Dahl rats

Authors: Cristina Chauvet^{1*}, Kimberley Crespo^{1*}, Annie Ménard¹, Yanrui Wu², Chunjie Xiao², Marilyne Blain¹, Julie Roy¹, Alan Y. Deng¹

Running head: Gene discovery for hypertension

Institutes: ¹Research Centre, CRCHUM (Centre hospitalier de l'Université de Montréal), Department of Medicine, Université de Montréal, Montréal, Québec, Canada

²Biology Department, Yunnan University, Kunming, Yunnan, China

Grant supports: Canadian Institutes for Health Research (CIHR) to AYD, and a China-Canada Joint Health Research Initiative from CIHR (CCI 102928) to AYD and Natural Science Foundation of China (NSFC 30911120481) to CX.

Conflict of interest: None

Corresponding author: Alan Y. Den

* These authors contributed equally to the current work.

Word count: 4,671

10.1 Abstract

Objectives: The interval harboring a quantitative trait locus (QTL) for blood pressure (BP), *C18QTL3*, contains beta-2 adrenergic receptor (*Adrb2*) and neural precursor cell expressed, developmentally down-regulated 4-like (*Nedd4l*) genes. None of the other genes in the *C18QTL3*-residing interval is known to affect BP. The identification of *C18QTL3* might uncover a brand-new gene that could prosper into a novel diagnostic and/or therapeutic target for essential hypertension, if neither *Adrb2* nor *Nedd4l* could be upheld as candidate genes.

Methods: Congenic fine resolution was combined with gene analyses.

Results: The gene encoding alpha kinase 2 (*Alpk2*) contains a 3 base-pair deletion and multiple non-conserved mutations in its coding region in Dahl salt-sensitive (DSS) rats. In contrast, the gastrin-releasing peptide gene (*Grp*) possesses 2 non-conserved mutations, designated as single nucleotide polymorphisms 1 and 2 (i.e. SNP1 and SNP2), but could not be supported as a candidate gene, because the C18S.L14 congenic strain displayed a homozygous DSS genotype at both SNP1 and SNP2. Furthermore, *Adrb2* and *Nedd4l* could not account for the BP-diminishing effect of Lewis alleles in C18S.L14, since their DSS alleles bear functionally-identical domains as those of Lewis, and no evidence of differential expression and splicing was evident. No significant nucleotide variations were found in 13 other genes closely-linked to *Alpk2*.

Conclusion: *Alpk2* emerged as a strong candidate gene for *C18QTL3*. The present study is the first to implicate *Alpk2* in the genetics of polygenic hypertension and paves the way for novel gene discovery.

Key words: fine congenic resolution, *C18QTL3*, *Alpk2*, *Adrb2*, *Nedd4l*,

10.2 Introduction

An important criterion in establishing a gene in essential hypertension is whether or not variations in that gene can alter blood pressure (BP) *in vivo* [1]. Previously, we found an association of coding variants in the β -2 adrenergic receptor gene (*ADRB2*) with essential hypertension in a population-specific manner [2], although this evidence by itself is inadequate to support its biological significance [3]. Coincidentally, *Adrb2* is located in the region harboring a quantitative trait locus (QTL), *C18QTL3*, for which a replacement of hypertensive DSS alleles by those of normotensive Lewis rats lowered BP [4]. Thus, the candidacy of *Adrb2* as a *C18QTL3* may be tested through fine congenic resolution and gene analysis in our animal models.

It so happens that both *Adrb2* and the neural precursor cell expressed, developmentally down-regulated 4-like gene (*Nedd4l*) reside in the *C18QTL3* interval [5] and are separated from each other by roughly 3.29 megabases (Mb) (Figure. 1). Several association studies indicated that *NEDD4L* may constitute a susceptibility gene for essential hypertension [6-8]. Thus, *Adrb2* and *Nedd4l* appear to be logical gene candidates for *C18QTL3*, even though their human homologues, *ADRB2* and *NEDD4L*, are located on CHRs 5 and 18 [2,8] separately.

The current investigation had 3 objectives. The first goal was to conduct a comprehensive fine congenic resolution of BP QTLs on Chr 18, focusing on *Adrb2* and *Nedd4l*. The second goal was to assess the candidacy of *Adrb2* and *Nedd4l* as a BP QTL. Finally, if neither *Adrb2* nor *Nedd4l* could be supported as a BP QTL, what gene(s) may be the BP QTL.

10.3 Methods

10.3.1 Animals

Protocols for handling as well as maintaining animals were approved by our institutional animal committee (CIPA).

10.3.2 Construction of new congenic sub-strains

Genetically, congenic strains bear different Chr 18 segments of DSS that are replaced by the homologues of Lewis rats. The earlier congenic strains [5] are the basis for deriving new congenic sub-strains used in the current work. The breeding and screening procedure in this process were similar to those reported previously [5]. Emphasis was placed on producing congenic sub-strains with no chromosome overlaps. Among them a congenic sub-strain was highly valued that specifically captured *Adrb2* and/or *Nedd4l* while retaining as fewer other genes surround them as possible.

For the present work, 5 informative new congenic sub-strains were produced (Figure 1), designated:

DSS.Lewis-(D18Chm162-D18Chm172)/(abbreviated as C18S.L10)

DSS.Lewis-(D18Chm124-D18Chm126)/Lt(C18S.L13)

DSS.Lewis-(D18Chm124-D18Rat61)/Lt (C18S.L14)

DSS.Lewis-(D18Chm91-D18Rat27)/Lt(C18S.L15)

DSS.Lewis-(D18Rat29-D18Rat67)/Lt(C18S.L16).

10.3.3 Animal protocols, BP phenotyping and statistical analyses

Breeding protocols, dietary treatments, implantation, postoperative care and BP measurement schedules were essentially the same as documented previously [5]. Briefly, male rats (n = at least 5) were weaned at 21 days of age, maintained on a low salt diet (0.2% NaCl) and followed by a high salt diet (2% NaCl) starting from 35 days of age until the end of the experiment. Telemetry probes were implanted at 56 days of age (namely 3 weeks from the time of the high salt diet). BPs of all rats were measured continuously during 12 days. BP of DSS was pooled from 4 separate measurements as they were not different. In the BP presentation, averaged readings for the duration of measurement of 12 days were provided for each strain. Since systolic and diastolic pressures were consistent (data not shown) with mean arterial pressures (MAP) of all the strains, only their MAPs are presented.

Repeated measures' analysis of variance (ANOVA) followed by Dunnett's test, which corrects for multiple comparisons and unequal sample sizes, was used to compare a parameter in MAP between 2 groups as reported previously [5]. During the BP comparison, ANOVA was first applied to assess the inter-group differences. If it was significant, the Dunnett's test ensued to identify which group was different from DSS, and the level of significance. 2 x 2 ANOVA determines a QTL-QTL interaction (or lack thereof) by evaluating whether the observed effect of a congenic strain combining 2 separate QTLs is significantly different from a predicted sum of effects from each individual QTL [5].

10.3.4 Mutation screening in *Adrb2* and *Nedd4l* genes

The entire *Adrb2* coding [9] and regulatory [10] domains were first amplified separately by the polymerase chain reaction (PCR) from genome DNAs of DSS and Lewis strains, and then sequenced.

The entire *Nedd4l* coding region except for the first 18 nucleotides was amplified by PCR from kidney cDNAs with nested and overlapping primer pairs. The remaining 18 nucleotides from the ATG start codon plus sequences unable to be accurately determined by direct sequencing are contained within exons 1 and 2. The coding regions, exon-intron junctions for exons 1, 2 and 3 as well as a putative promoter section were sequenced for DSS and Lewis.

10.3.5 Analysis of gene expressions of *Adrb2* and *Nedd4l*

Nedd4l is known to be functional in the kidneys that can potentially impact on BP homeostasis [11]. Thus its expressions in the kidneys were compared. For consistency, the same organ was used to analyze *Adrb2*. Quantitative real-time reverse transcription PCR (qRT-PCR) was utilized to determine and compare renal expressions of the 2 genes.

10.3.6 Sequencing of additional candidate genes

Several genes (Figure 1) in the *C18QTL3*-residing interval (Supplement 1) were chosen for sequencing because they are known genes in the rat, mouse and humans. Their coding regions from both DSS and Lewis were first amplified by PCR from genomic DNAs and/or appropriate cDNAs, and then sequenced. When a mutation was detected, the same segment was independently PCR-amplified and re-sequenced. Thus, all mutations were verified.

10.4. Results

10.4.1 QTL placements and assessment of QTL interactions

The region originally-thought to contain *C18QTL3* [5] turns out to lodge 2 QTLs, *C18QTL3* and *C18QTL4*, since the 2 congenic sub-strains delimiting them, namely C18S.L14 and C18S.L10, have no genomic overlaps between them (Figure 1). C18S.L14 and C18S.L10 defining the QTL-lodging intervals both autonomously and directly exhibit BP effects, each accounting for at least 41% of the MAP difference between DSS (184 mmHg) and Lewis (96mmHg).

In comparison, *C18QTL1* is placed in the non-overlapping segment between C18S.L15 exhibiting a BP effect and C18S.L16 lacking it. This ‘subtractive’ comparative mapping can not rule out the possibility that the congenic strain lacking a BP effect could harbor a QTL, which would require a combination with another putative QTL in the non-overlapping region carried by the strain displaying an effect. Although this imaginary combination has never been demonstrated in experimental hypertension [12], for the sake of argument, the placement of *C18QTL1* is conservatively tentative. No chromosome crossovers occurred in the segment harboring *C18QTL2* [5], and consequently, *C18QTL2* could not be finely resolved.

The C18S.L2 congenic strain carries 2 QTLs, namely *C18QTL3* caught by C18S.L14, and *C18QTL4* by C18S.L10 (Figure 1). A 2 x 2 ANOVA [5] demonstrates an epistatic interaction ($p < 0.001$) between *C18QTL3* and *C18QTL4*.

10.4.2 Evaluation of candidate genes *Adrb2* and *Nedd4l*

C18S.L14 bears a small chromosome fragment containing *C18QTL3* (Figure 1) and still lodges *Adrb2* and *Nedd4l*. This outcome prompted us to directly analyze these 2 genes, because of our prior knowledge on their functional roles in BP regulation [11,13] and on their association with essential hypertension [2,3,6-8]. Thus, relevant insight could be translated into BP homeostasis via their signaling or regulatory mechanisms, once a functional mutation was detected in either 1 of these 2 genes.

10.4.2.1 *Adrb2*

Since the *Adrb2* 5'UTR that regulates its expression has been characterized to reside within 1.1 kb [10,14,15], 1481 base pairs (Bp) of the genome fragment 5' upstream of its initiation codon have been sequenced. DSS and Lewis carry the identical coding region as well as the known regulatory 5'UTR (GenBank accession number GQ160814 and Supplement 2).

10.4.2.2 *Nedd4l* coding regions

From kidney cDNAs, except for the first 6 amino acids, the entire coding *Nedd4l* region including 455 bp of 3'UTR, was determined by PCR-amplifications followed by sequencing (Supplement 3). Exon #1 [16] carries the first 8 amino acids and yields the necessary sequence (Supplement 4).

10.4.2.3 *Nedd4l* exon-intron junctions

The last nucleotide (Supplement 4) in Exon #1 of the rat *Nedd4l* gene corresponds to that of the human *NEDD4L* [17]. This nucleotide belongs to amino acid #8 of the rat and #16 of humans and can lead to the formation of a splice variant in humans believed to contribute to essential hypertension in a population-dependent fashion [17-19]. DSS and Lewis, however, bear identical sequences in Exon #1 and at the exon-intron junction (Supplement 4). Exons 2

and 3 were reported to be differentially-spliced in DSS [16]. No nucleotide differences were found within 19 bps of intronic regions (Supplement 4).

10.4.2.4 *Nedd4l* 5' upstream domain

Although the promoter location for *Nedd4l* is not known in humans, mice or rats, based on the information on other promoters such as for *Adrb2* [14,15], it seemed informative to determine a reasonable genomic tract 5' from the ATG initiation codon. No nucleotide differences between DSS and Lewis were detected in this region (GenBank accession number GQ160817) (Supplement 4).

10.4.2.5 Analysis of *Adrb2* and *Nedd4l* gene expressions

Quantitative polymerase chain reactions (qPCRs) were carried out to evaluate renal gene expressions of the 2 genes. No differences for either of them were found between DSS and Lewis rats (Supplement 5).

10.4.3 Alpha kinase 2 (*Alpk2*), not gastrin-releasing peptide (*Grp*) gene, as a leading candidate for *C18QTL3*

Since *Adrb2* and *Nedd4l* were not upheld as candidate genes for *C18QTL3*, another gene(s) that can should be present. Indeed, among those in the *C18QTL3*-residing interval (Supplement 1), only *Alpk2* emerged as a prominent candidate because of the following reasons.

First, in the coding region of DSS rats, *Alpk2* contains a deletion of 3 base pairs that eliminates an alanine amino acid at the # 637 position as well as 5 non-synonymous mutations (Table 1 and Supplement 6). The *Alpk2* alleles of the C18S.L14 congenic strain are the same as that of Lewis, i.e. LL homozygotes at the positions of the deletion and 5 mutated nucleotides causing non-synonymous amino acid changes (data not shown). Second, although both *Alpk2* and *Grp* possess non-conserved mutations (Table 1), C18S.L14 possesses DSS homozygotes (i.e. SS) (data not shown) at the 2 mutated nucleotides in *Grp* (Table 1 and Supplement 7). Therefore, *Alpk2*, not *Grp*, is required for the BP-lowering effect of *C18QTL3*

from Lewis. In contrast, the remaining genes sequenced (Table 1 and Figure 1) were not supported as candidates for a lack of non-synonymous mutations (Supplements 8-19). Finally, *Alpk2* seems to be a functional gene because it is expressed in all the organs examined (Figure 2).

10.5 Discussion

Major findings in the current work are (a) 2 unique QTLs, *C18QTL3* and *C18QTL4*, were resolved that interact with each other epistatically in a chromosome segment. Each of these QTLs can affect BP independently of each other. (b) Two eminent genes, *Adrb2* and *Nedd4l*, were comprehensively analyzed and could not be supported as *C18QTL3*. (c) *Alpk2* became a leading candidate gene for *C18QTL3*. Thus, *C18QTL3* seems represented by a novel gene previously unknown for BP control.

10.5.1 Fine congenic resolution of BP QTLs

The use of congenic strains with no chromosome overlaps to one another is essential in affirmatively delimiting the fragments lodging QTLs predicated on a cause-effect relationship. From our systematic analyses, several BP QTLs on Chr 18 were defined to the distinct genome segments (Figure 1), paving the way for their positional cloning.

C18QTL3 is conclusively demarcated. The number of 67 genes dwelling in the chromosome interval, however, could not be further reduced at this time, because no polymorphic markers were found after screening 72 microsatellites and no single nucleotide polymorphisms (SNPs) were detected so far in the ambivalent region between D18Rat55 and D18Chm124 markers.

C18QTL1 and *C18QTL4* (Figure 1) as well as *C18QTL2* [5] are either ambiguously defined or contain too many genes to be sequenced. A further reduction in the QTL-residing intervals is necessary. Possibly, more QTLs may emerge from fine congenic resolution of

these QTLs. Evidently, if the QTL-residing interval can be restricted to contain only 1 gene, the gene has to be responsible for the QTL.

10.5.2 Alpha kinase 2 (*Alpk2*) gene is a standout candidate to be *C18QTL3*

Of all the genes sequenced (Table 1), only *Alpk2* harbors mutations resulting in non-conserved changes in amino acids comparing DSS with both Lewis and the C18S.L14 congenic strain. The 3-base-pair deletion, although in frame with the rest of the coding domain, removes an amino acid in DSS. Thus, *Alpk2* poised to be a candidate gene for *C18QTL3*. The emergence of *Alpk2* lays the ground work for the identification of the gene responsible for *C18QTL3*.

Alpha kinase 2 [20,21] belongs to a class of protein kinases that modulate protein functions via phosphorylation and often affect signal transduction as a consequence [22]. *ALPK2* was found to be differentially expressed in human gastric cancer cell lines comparing to non-malignant cells [23].

No functional studies on *Alpk2*, however, have been documented [20,21]. Because its role in BP modulation is completely unknown, the emergence of *Alpk2* as a candidate for *C18QTL3* will likely lead to a novel mechanistic insight into the biological process, biochemical pathway or signaling cascade steering towards BP homeostasis.

Alpk2 is expressed ubiquitously in all the organs of DSS and Lewis rats of 11 weeks of age after 6 weeks on a high salt diet (Figure 2). *Alpk2* is a predicted gene in the rat and no report on its gene and/or cDNA structure is available :

(http://www.ncbi.nlm.nih.gov/nucore/XM_574162.3)

Its homologues in the mouse and humans are merely described from collections of full-length cDNA clones (<http://www.ncbi.nlm.nih.gov/gene/225638>). No data on its enzymatic activity has been chronicled beyond a structural comparison with other protein kinases [20,21]. Thus, its candidate status in hypertension provides an impetus for thoroughly characterizing *Alpk2*, such as its true initiation codon and length of the protein etc. As a result, an analysis of

its biochemical functions can benefit. Despite a lack of functional proof, it is reasonable to predict that the deletion plus the 5 significant mutations in *Alpk2* would likely affect the enzymatic activity of alpha kinase 2, i.e. Δ CCG1910-1912, T376C changing C126R, C475T changing P159S, A2918T changing Q973L, A3754G changing K1252E and A6217G changing I2073V (Table 1 and Supplement 6).

Another noteworthy feature from the present work is that congenic resolution is powerful in separating 2 closely-linked and equally-valid genes that harbor non-conserved mutations, *Alpk2* and *Grp*. A chromosome crossover is between 2 nucleotides. A congenic strain can precisely define the inclusion or exclusion of a gene because of it. In the present case, *Alpk2* is included as, whereas *Grp* is excluded from, a candidate for *C18QTL3*. This resolution is in sharp contrast to the linkage result where only 1 QTL with a broad chromosome coverage was detected [24].

From our current screening work, a possibility cannot be excluded that an additional candidate gene for *C18QTL3* might exist among those not yet sequenced (Supplement 1 and Table 1). A further congenic resolution minimizing the presence of other genes surrounding *Alpk2* may be needed to resolve this issue. In our previous attempt at reducing the *C18QTL3*-containing interval, more than 250 F₂ rats have been screened, but no crossovers were detected (data not shown). With *Alpk2* as the central target, screening a larger number of F₂ rats seems fully justified in order to further narrow down the *C18QTL3*-containing interval. Nevertheless, the appearance of *Alpk2* provides a sufficient and necessary focal point for further genetic, molecular, cellular, biochemical and physiological analyses of *C18QTL3*.

In contrast, *Adrb2* and *Nedd4l* cannot be sustained as candidate genes for *C18QTL3*. Evidently, DSS and Lewis strains contrast sharply in BP (Figure 1), but not in the *Adrb2* and *Nedd4l* genes themselves (Supplements 2-5). Consequently, the BP effect exerted by *C18QTL3* cannot be attributed to them. A BP disparity was found in a Chr 18 congenic strain made from a cross of spontaneously hypertensive and diabetic BB rats [25]. It seems worthwhile comparing *Adrb2* genes in these strains.

10.6 Conclusion

A systematic fine congenic resolution yielded multiple QTLs and each of them can independently influence BP. Neither of the prominent candidate genes, *Adrb2* and *Nedd4l*, could be bolstered to be *C18QTL3*, despite their distinguished roles in BP modulation and their highly-visible associations with essential hypertension. The discovery of *C18QTL3* and other QTLs on Chr 18 appeared likely to unmask novel genes and their underlying physiological mechanisms that may engender innovative diagnostic tools and therapeutic targets for essential hypertension. Indeed, *Alpk2* emerged to be a credible candidate gene for *C18QTL3*. The current work has provided the first evidence that *Alpk2* is a novel gene potentially responsible for hypertension of DSS rats, and now has become a plausible target for the genetic research of essential hypertension.

10.7 Reference List

1. Cowley AW, Jr. The genetic dissection of essential hypertension. *Nat Rev Genet* 2006; **7**: 829-840.
2. Wu H, Tang W, Li H, Zhou X, Yang Y, Yu H, Li K, Xiao C, Deng AY. Association of the beta2-adrenergic receptor gene with essential hypertension in the non-Han Chinese Yi minority human population. *J Hypertens* 2006; **24**: 1041-1047.
3. Hahntow IN, Koopmans RP, Michel MC. The beta2-adrenoceptor gene and hypertension: is it the promoter or the coding region or neither? *J Hypertens* 2006; **24**: 1003-1007.
4. Wu H, Tang W, Li H, Zhou X, Yang Y, Yu H, Li K, Xiao C, Deng AY. Functional significance of single nucleotide polymorphisms within the 5'-flanking region of beta2-adrenergic receptor gene. *J Hypertens* 2006; **24**: 2474-2476.
5. Charron S, Lambert R, Eliopoulos V, Duong C, Menard A, Roy J, Deng AY. A loss of genome buffering capacity of Dahl salt-sensitive model to modulate blood pressure as a cause of hypertension. *Hum Mol Genet* 2005; **14**: 3877-3884.
6. Fava C, von Wöhrn F, Berglund G, Carlson J, Hedblad B, Rosberg L, Burri P, Almgren P, Melander O. 24-h ambulatory blood pressure is linked to chromosome 18q21-22 and

- genetic variation of NEDD4L associates with cross-sectional and longitudinal blood pressure in Swedes. *Kidney Int* 2006; **70**: 562-569.
7. Pankow JS, Dunn DM, Hunt SC, Leppert MF, Miller MB, Rao DC, Heiss G, Oberman A, Lalouel JM, Weiss RB. Further evidence of a quantitative trait locus on chromosome 18 influencing postural change in systolic blood pressure: the Hypertension Genetic Epidemiology Network (HyperGEN) Study. *Am J Hypertens* 2005; **18**: 672-678.
 8. Russo CJ, Melista E, Cui J, DeStefano AL, Bakris GL, Manolis AJ, Gavras H, Baldwin CT. Association of NEDD4L ubiquitin ligase with essential hypertension. *Hypertension* 2005; **46**: 488-491.
 9. Gocayne J, Robinson DA, FitzGerald MG, Chung FZ, Kerlavage AR, Lentjes KU, Lai J, Wang CD, Fraser CM, Venter JC. Primary structure of rat cardiac beta-adrenergic and muscarinic cholinergic receptors obtained by automated DNA sequence analysis: further evidence for a multigene family. *Proc Natl Acad Sci U S A* 1987; **84**: 8296-8300.
 10. Jiang L, Kunos G. Sequence of the 5' regulatory domain of the gene encoding the rat beta 2-adrenergic receptor. *Gene* 1995; **163**: 331-332.
 11. Shi PP, Cao XR, Sweezer EM, Kinney TS, Williams NR, Husted RF, Nair R, Weiss RM, Williamson RA, Sigmund CD, Snyder PM, Staub O, Stokes JB, Yang B. Salt-sensitive hypertension and cardiac hypertrophy in mice deficient in the ubiquitin ligase Nedd4-2. *Am J Physiol Renal Physiol* 2008; **295**: F462-F470.
 12. Deng AY. Positional Cloning of Quantitative Trait Loci for Blood Pressure: How Close Are We?: A Critical Perspective. *Hypertension* 2007; **49**: 740-747.
 13. Brodde OE, Leineweber K. Beta2-adrenoceptor gene polymorphisms. *Pharmacogenet Genomics* 2005; **15**: 267-275.
 14. Drysdale CM, McGraw DW, Stack CB, Stephens JC, Judson RS, Nandabalan K, Arnold K, Ruano G, Liggett SB. Complex promoter and coding region beta 2-adrenergic receptor haplotypes alter receptor expression and predict in vivo responsiveness. *Proc Natl Acad Sci U S A* 2000; **97**: 10483-10488.
 15. Johnatty SE, Abdellatif M, Shimmin L, Clark RB, Boerwinkle E. Beta 2 adrenergic receptor 5' haplotypes influence promoter activity. *Br J Pharmacol* 2002; **137**: 1213-1216.
 16. Umemura M, Ishigami T, Tamura K, Sakai M, Miyagi Y, Nagahama K, Aoki I, Uchino K, Rohrwasser A, Lalouel JM, Umemura S. Transcriptional diversity and expression of NEDD4L gene in distal nephron. *Biochem Biophys Res Commun* 2006; **339**: 1129-1137.

17. Dunn DM, Ishigami T, Pankow J, von Niederhausern A, Alder J, Hunt SC, Leppert MF, Lalouel JM, Weiss RB. Common variant of human NEDD4L activates a cryptic splice site to form a frameshifted transcript. *J Hum Genet* 2002; **47**: 665-676.
18. Araki N, Umemura M, Miyagi Y, Yabana M, Miki Y, Tamura K, Uchino K, Aoki R, Goshima Y, Umemura S, Ishigami T. Expression, Transcription, and Possible Antagonistic Interaction of the Human Nedd4L Gene Variant: Implications for Essential Hypertension. *Hypertension* 2008; **51**: 773-777.
19. Wen H, Lin R, Jiao Y, Wang F, Wang S, Lu D, Qian J, Jin L, Wang X. Two Polymorphisms in NEDD4L Gene and Essential Hypertension in Chinese HansΓÇöA Population-Based Case-Control Study. *Clinical and Experimental Hypertension* 2008; **30**: 87-94.
20. Middelbeek J, Clark K, Venselaar H, Huynen MA, van Leeuwen FN. The alpha-kinase family: an exceptional branch on the protein kinase tree. *Cell Mol Life Sci* 2010; **67**: 875-890.
21. Ryazanov AG, Pavur KS, Dorovkov MV. Alpha-kinases: a new class of protein kinases with a novel catalytic domain. *Current Biology* 1999; **9**: R43-R45.
22. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; **298**: 1912-1934.
23. Junnila S, Kokkola A, Karjalainen-Lindsberg ML, Puolakkainen P, Monni O. Genome-wide gene copy number and expression analysis of primary gastric tumors and gastric cancer cell lines. *BMC Cancer* 2010; **10**: 73.
24. Garrett MR, Dene H, Walder R, Zhang Q, Cicila GT, Assadnia S, Deng AY, Rapp JP. Genomic scan and congenic strains for blood pressure quantitative trait loci using Dahl salt-sensitive rats. *Genome Research* 1998; **8**: 711-723.
25. Kloting I, Voigt B, Kovacs P. Metabolic features of newly established congenic diabetes-prone BB.SHR rat strains. *Life Sci* 1998; **62**: 973-979.

10.8 Tables, figures and legends

Table 1. Mutation screening of candidate genes in the C18QTL3-containing interval

Gene	Genomic start position	Genomic end position	Number of exons	Size of codons (bp)	Mutation detected Lewis/ DSS	Change in amino acid (AA) Lewis/DSS
<i>Nid67</i>	2 460 479	2 460 661	2	183	No	No
<i>Rbm22</i>	2 542 718	2 553 093	11	1 263	No	No
<i>Arsi</i>	2 903 564	2 904 977	2	1 722	No	No
<i>Slc26a2</i>	3 204 489	3 208 486	5	2 220	No	No
<i>Csnk1a1</i>	3 575 198	3 598 085	9	978	No	No
<i>Adrb2</i>	4 210 345	4 208 483	1	1 257	No	No
<i>Ecg2</i>	4 640 157	4 643 629	4	225	No	No
<i>Apcdd1</i>	5 122 234	5 124 665	2	870	No	No
<i>Wdr7</i>	5 905 753	6 176 665	27	4 467	No	No
<i>St8sia3</i>	6 501 348	6 507 560	4	1 143	No	No
<i>Fech</i>	6 708 742	6 710 340	11	1 269	No	No
<i>Nars</i>	6 753 801	6 769 584	12	1 677	No	No
<i>Nedd4l</i>	7 167 842	7 509 317	31	2 892	T321A, C1995T G2139A	No
<i>mir122a</i>	7 546 271	7 546 356	1	85	No	No
<i>Alpk2</i> predicted	7 596 221	7 691 046	10	6 303	T376C C475T C618T ΔCCG1910 A2918T	C126R P159S No A637/- Q973L

					A3754G	K1252E
					C5610G	No
					A6217G	I2073V
					T283G	S95A
<i>Grp</i>	8 196 564	8 209 539	3	444	A286G	R96G

Footnote to Table 1: Gene locations of contig NW_047514 on Chr 18 are indicated on the map in Figure 1. The position of a mutation corresponds to the designation from the ATG start codon of that gene. The amino acid position begins from the first methionine. Δ symbolizes a deletion. *Adrb2*, adrenergic, beta-2-, receptor, surface; *Alpk2*, alpha-kinase 2; *Apcdd1*, adenomatosis polyposis coli down-regulated 1; *Arsi*, arylsulfatase family member I; *Csnk1a1*, casein kinase 1, alpha 1; *Ecg2*, esophagus cancer-related protein 2; *Fech*, ferrochelatase; *Grp*, gastrin-releasing peptide; *mir-122a*, microRNA 122a; *Nars*, asparaginyl-tRNA synthetase; *Nedd4l*, neural precursor cell expressed, developmentally down-regulated 4-like; *Nid67*, putative small membrane protein NID67; *Rbm22*, RNA binding motif protein 22; *Slc26a2*, solute carrier family 26 (sulfate transporter), member 2; *St8sia3*, ST8 alpha-N-acetylneuraminide alpha-2,8-sialyltransferase 3; and *Wdr7*, WD repeat domain 7.

Figure 1. Fine congenic resolution of BP QTLs on Chr 18.

The contigs and marker positions in megabases (Mb) within them are taken from (<http://www.ncbi.nlm.nih.gov/mapview>). Solid bars under congenic sub-strains symbolize the Dahl salt-sensitive (DSS) chromosome fragments that have been replaced by those of Lewis. Open bars on ends of solid bars indicate the ambiguities of crossover breakpoints between markers. Please consult the legend in Table 1 for gene nomenclature. D18Chm markers are produced from the rat genome sequence. Mean arterial pressures (MAPs) for DSS and congenic strains are averaged for the period of measurement and are given at the bottom of the map to facilitate strain comparisons. Significant p values are emphasized in bold and italics. \pm indicates SEM. C18S.L10, C18S.L13, C18S.L14, C18S.L15 and

C18S.L16 are newly-generated from this study. Their full names are listed in the method section under Constructions of new congenic sub-strains. C18S.L2 was produced previously [5]. The placements of the BP QTLs are indicated to the right by brackets. Below a given QTL designation is the size of the QTL-residing interval in Mbs, and the number of genes and undefined loci.

Figure 2. Organ expressions of *Alpk2* and *Grp* assayed by reverse transcriptase

polymerase chain reaction (RT-PCR). The organs are from Dahl salt-sensitive (DSS) on the left panel and Lewis rats on the right. Numbers to the left indicate the size of the fragment in base pairs.

Primers for RT-PCR are forward 5'TCCCGAGCACTCTGGAAATA 3' and reverse 5'AAGTGAAGCTCCTCCGTGAA 3' for *Alpk2*,

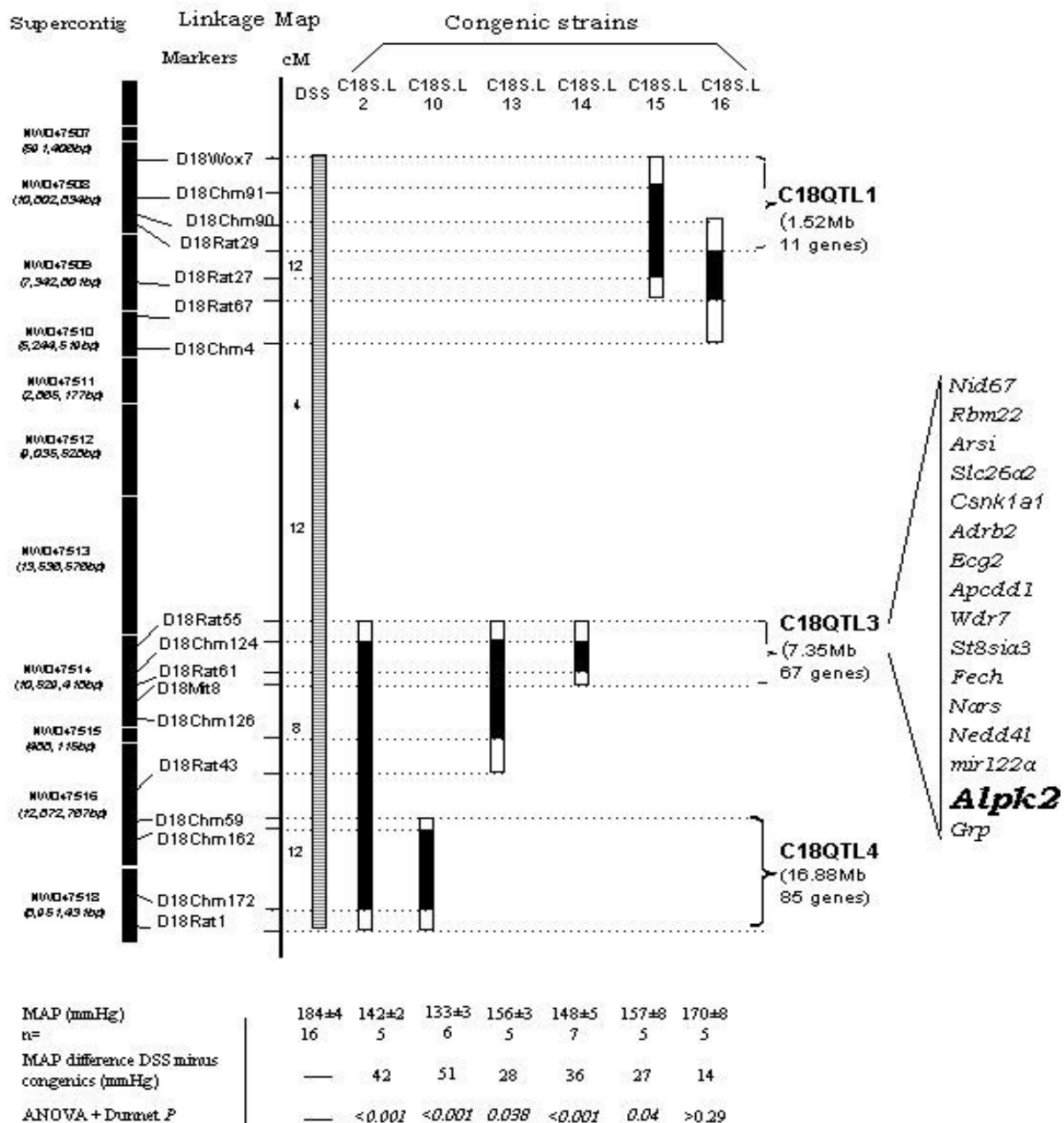
5' GAGCTCTCGCTCTTGCTGTT 3' and reverse 5' CTGGATCCCAAGTAGGCTGG 3'for *Grp*,

5' ACTGCCGCATCCTCTTCCTC 3' and reverse 5' CCGCTCGTTGCCAATAGTGA 3'for *β -actin*.

Two primers for each gene are located in 2 different exons to avoid amplifying genomic DNAs contaminated in RNA preparations, since no products were seen when genomic DNAs were amplified (data not shown). Results shown are from 1 rat of each strain and they have been replicated with multiple rats of the same strain (data not shown). All rats were males, 11 weeks of age and fed a high salt diet for 6 weeks.

Figure 1

Rat Chromosome 18



10.9 Supplementary data

Supplement 1. Genes residing in the chromosome interval harboring *C18QTL3* as defined by microsatellite markers D18Rat55 and D18Mit8.

<i>RGD1561768_predicted</i>	similar to Rpl7a protein
<i>LOC364882</i>	similar to ribosome L13
<i>RGD1309362</i>	similar to interferon-inducible GTPase
<i>RGD1559715_predicted</i>	similar to MGC108823 protein
<i>MGC108823</i>	similar to interferon-inducible GTPase
<i>LOC502172</i>	hypothetical LOC502172
<i>RGD1305184_predicted</i>	similar to CDNA sequence BC023105
<i>MGC105567</i>	similar to cDNA sequence BC023105
<i>LOC689233</i>	similar to immunity-related GTPase family, cinema 1
<i>Nid67</i>	putative small membrane protein NID67 (role in neurogenesis)
<i>Dctn4</i>	dynactin 4
<i>Rbm22</i>	RNA binding motif protein 22
<i>LOC689274 (Myoz3)</i>	myozenin 3
<i>Synpo</i>	synaptopodin
<i>Ndst1</i>	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 1
<i>Rps14</i>	ribosomal protein S14
<i>Cd74</i>	Cd74 molecule, major histocompatibility complex, class II invariant chain
<i>Tcof1_predicted</i>	Treacher Collins Franceschetti syndrome 1, homolog
<i>RGD1310242 (Arsi)</i>	arylsulfatase family, member 1
<i>Camk2a</i>	calcium/calmodulin-dependent protein kinase II α
<i>Slc6a7</i>	solute carrier family 6 (neurotransmitter transporter, L-proline), member 7 Na(+)-dependent transporter
<i>LOC689355</i>	hypothetical protein LOC689355
<i>Cdx1_predicted</i>	caudal type homeo box 1
<i>Pdgfrb</i>	platelet derived growth factor receptor β polypeptide
<i>Csf1r</i>	colony stimulating factor 1 receptor
<i>Slc26a2</i>	solute carrier family 26 (sulfate transporter), member 2
<i>Pde6a_predicted</i>	phosphodiesterase 6A, cGMP-specific, rod, alpha
<i>Ppargc1b</i>	peroxisome proliferator-activated receptor gamma, coactivator 1 β
<i>RGD1560471_predicted</i>	similar to hypothetical protein 4933429F08
<i>Csnk1a1</i>	casein kinase 1, alpha 1
<i>rno-mir-145</i>	MicroRNA
<i>Il17b</i>	interleukin 17B
<i>RGD1563290_predicted</i>	prenylcysteine oxidase 1 like
<i>LOC688777 (Grpel2)</i>	GrpE-like 2, mitochondrial
<i>RGD1311580_predicted</i>	actin filament associated protein 1-like 1 (<i>Afap111</i>)
<i>RGD1565118_predicted</i>	actin binding LIM protein family, member 3, (<i>Ablim3</i>)
<i>RGD1309038 (Sh3tc2)</i>	SH3 domain and tetratricopeptide repeats 2
<i>RGD1561455</i>	similar to Ras GTPase-activating-like protein IQGAP2
<i>Adrb2</i>	beta 2 adrenergic receptor
<i>Htr4</i>	5-hydroxytryptamine (serotonin) receptor 4
<i>Rbxo38-pred</i>	F-box protein 38
<i>Ecg2</i>	esophagus cancer-related protein 2
<i>Loc689570</i>	similar to Kazal type serine protease inhibitor 4
<i>Loc689616 (Apccd1)</i>	adenomatosis polyposis coli down-regulated 1

<i>Napg</i>	N-ethylmaleimide-sensitive factor attachment protein, gamma
<i>Loc688934 (Fam38b)</i>	family with sequence similarity 38, member B
<i>RGD1306866-pred</i>	similar to Hypothetical protein KIAA0233
<i>Txn11</i>	thioredoxin-like 1 mouse homolog
<i>Wdr7</i>	WD repeat domain 7
<i>RGD1565894-pred</i>	similar to ribosomal protein L31
<i>Loc502174</i>	similar to 40S ribosomal protein S7
<i>Loc502175</i>	similar to peroxiredoxin 1
<i>St8sia3</i>	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3
<i>RGD1562699</i>	hypothetical protein LOC291558
<i>RGD1564677</i>	similar to transcription factor ONECUT2
<i>Fech-pred</i>	ferrochelatase
<i>Nars</i>	asparaginyl-tRNA synthetase
<i>Atp8b1-pred</i>	ATPase, Class I, type 8B, member 1
<i>Loc498874</i>	hypothetical LOC498874
<i>Nedd4l</i>	neural precursor cell expressed, developmentally down-regulated 4-like
<i>Loc502176</i>	hypothetical protein LOC502176
<i>rno-mir-122a</i>	MicroRNA
<i>Alpk2_predicted</i>	alpha-kinase 2
<i>Malt1_predicted</i>	mucosa associated lymphoid tissue lymphoma translocation gene 1
<i>RGD1560815_predicted</i>	similar to acidic ribosomal phosphoprotein P1
<i>Znf532_predicted</i>	zinc finger protein 532
<i>RGD1304731_predicted</i>	similar to RIKEN cDNA 5330437102 gene
<i>Sec11b</i>	SEC11 homolog C (<i>S. cerevisiae</i>), human homolog
<i>Grp</i>	gastrin-releasing peptide
67 genes + 2 microRNAs	

Supplement 2. Aligned sequence comparison of the β -2 adrenergic receptor gene (*Adrb2*) between Dahl salt-sensitive (DSS) and Lewis strains

```

DSS      TACTTAAGAAATTGTTTGAAGAAGATCATGCAGACTGGAGAGACAGCTCAGTGGTTAA -1361
Lewis    TACTTAAGAAATTGTTTGAAGAAGATCATGCAGACTGGAGAGACAGCTCAGTGGTTAA -1361
*****
DSS      GAGCATTGGCTGACTTCCAGAGGACCAGGGTTCGATTGCCAGCACTCACAAAGGCAGCTT -1301
Lewis    GAGCATTGGCTGACTTCCAGAGGACCAGGGTTCGATTGCCAGCACTCACAAAGGCAGCTT -1301
*****
DSS      GTGTGTTGATACTATCTGTATATTCAGTCCTAGGGAATCTGATGCTCTCTTTCTCCTGAC -1241
Lewis    GTGTGTTGATACTATCTGTATATTCAGTCCTAGGGAATCTGATGCTCTCTTTCTCCTGAC -1241
*****
DSS      CTCAAAACATGCACAAACGTGGTGCACAGACAAACACAGAGATAAGACTCCCATGGCGCA -1181
Lewis    CTCAAAACATGCACAAACGTGGTGCACAGACAAACACAGAGATAAGACTCCCATGGCGCA -1181
*****
DSS      TAACATAAGTAATGTAGTTTATTTTTTTTATTTTCAAAAAGAAAATCTTGCATTTTCG -1121
Lewis    TAACATAAGTAATGTAGTTTATTTTTTTTATTTTCAAAAAGAAAATCTTGCATTTTCG -1121
*****
DSS      TTTTCAATGCATCCCTGGATATCCACCATATAACAAGAAAATGATGCCACACTCCCAAG -1061
Lewis    TTTTCAATGCATCCCTGGATATCCACCATATAACAAGAAAATGATGCCACACTCCCAAG -1061
*****
DSS      AATTAAGAATAACACCCCCACCCCAACTCTCCCCAAGTGGGTACACATCAATTCCTAAAG -1001
Lewis    AATTAAGAATAACACCCCCACCCCAACTCTCCCCAAGTGGGTACACATCAATTCCTAAAG -1001
*****
DSS      AAATGCAAAGGGCGAAGTTACTGCCTTGGTGGCGTTGCCGTGCATGTTCTCAAGGAGGGAG -961
Lewis    AAATGCAAAGGGCGAAGTTACTGCCTTGGTGGCGTTGCCGTGCATGTTCTCAAGGAGGGAG -961
*****
DSS      GACGCCAAAAAATATTTTTACCAAGTGTGCTTAGAAACTCTGTAGCTGTTGGAGTCAC -901
Lewis    GACGCCAAAAAATATTTTTACCAAGTGTGCTTAGAAACTCTGTAGCTGTTGGAGTCAC -901
*****
DSS      CTCCTGTGTTCTTGCCAAATAACTGAAAGGGAATACACTGGGTGCTGCGAATGGGGTGT -841
Lewis    CTCCTGTGTTCTTGCCAAATAACTGAAAGGGAATACACTGGGTGCTGCGAATGGGGTGT -841
*****
DSS      CATTATGAGTGCTAGAGGAGGGTAGTACCTGGATGTGAACACCGGTGTCTAAGTGATCC -781
Lewis    CATTATGAGTGCTAGAGGAGGGTAGTACCTGGATGTGAACACCGGTGTCTAAGTGATCC -781
*****
DSS      GCACTCCATGTCTCCACCGGAGCCCTTGTGTGTGTTTCTTTCAGCTTGCCTCTGGGTGTT -721
Lewis    GCACTCCATGTCTCCACCGGAGCCCTTGTGTGTGTTTCTTTCAGCTTGCCTCTGGGTGTT -721
*****
DSS      CCTGAGTTTGTATAGCCACCTCTCTGTGTTGGACCAGGAGACAGCTGTGCTCACCAGC -701
Lewis    CCTGAGTTTGTATAGCCACCTCTCTGTGTTGGACCAGGAGACAGCTGTGCTCACCAGC -701
*****
DSS      ATGAGCGGGCGGGGTGTGCTCGGTGCTGTGGCTGTGGTTCGGCATAAGTCTGAGCAC -641
Lewis    ATGAGCGGGCGGGGTGTGCTCGGTGCTGTGGCTGTGGTTCGGCATAAGTCTGAGCAC -641
*****
DSS      GTCCCAGTCAGCAGGGTGCCTTTGTGCCCTGTGTGCGGTGCCGGGTGAGCTTGTCTTT -581
Lewis    GTCCCAGTCAGCAGGGTGCCTTTGTGCCCTGTGTGCGGTGCCGGGTGAGCTTGTCTTT -581
*****
DSS      CTGTTCTAAAAGAAAATGCACCCACTGCCACTGGCGTCCAGAGACCTTCCAACCCGCAC -521
Lewis    CTGTTCTAAAAGAAAATGCACCCACTGCCACTGGCGTCCAGAGACCTTCCAACCCGCAC -521
*****
DSS      AGGTGACCTACATGAACAGGTAGCACTTCCAGACAGCGCCTGGGTGGTGGCGGGTCCGCA -461
Lewis    AGGTGACCTACATGAACAGGTAGCACTTCCAGACAGCGCCTGGGTGGTGGCGGGTCCGCA -461
*****
DSS      CAGCAGTCCAGATTTCTCTCTCGGGAGCCCCAGCTAGGGTAGCTGGAAGGGAGTGGTGG -401
Lewis    CAGCAGTCCAGATTTCTCTCTCGGGAGCCCCAGCTAGGGTAGCTGGAAGGGAGTGGTGG -401
*****

```

DSS CCTGGCCTCCGGGGAGCAGCTTGGGCCCGCCCGGGCCAAACCCAGGAGGAGGGGGCTA -341
Lewis CCTGGCCTCCGGGGAGCAGCTTGGGCCCGCCCGGGCCAAACCCAGGAGGAGGGGGCTA -341

DSS GGCAAGGAGGGTTGGCCCTTAGTGCTCGCCCTCCGGGCATGCGACGGTGCCATTGGCT -281
Lewis GGCAAGGAGGGTTGGCCCTTAGTGCTCGCCCTCCGGGCATGCGACGGTGCCATTGGCT -281

DSS CAAAGTTGCTGCACGTCACCTGGGCAAGTCCCTAAAGTCTAGTGCACATAACGGGCAGGG -221
Lewis CAAAGTTGCTGCACGTCACCTGGGCAAGTCCCTAAAGTCTAGTGCACATAACGGGCAGGG -221

DSS TGCCTGCAAGGCTGCTTCTCCCGCGTTCAGGCTGCAGCTGGCAGGCATCGCGAGCCCGG -161
Lewis TGCCTGCAAGGCTGCTTCTCCCGCGTTCAGGCTGCAGCTGGCAGGCATCGCGAGCCCGG -161

DSS AGCACCCACGAGCTCAGTGTGCAGGACGCGCCCCAGCACAGCCACTACGGTCTCTGAA -101
Lewis AGCACCCACGAGCTCAGTGTGCAGGACGCGCCCCAGCACAGCCACTACGGTCTCTGAA -101

DSS TGAAGCTTCCAGGAGTCCGCCCCGGACGGCTGCGCCCCATCGGAGGTGCACCCGCTGAG -41
Lewis TGAAGCTTCCAGGAGTCCGCCCCGGACGGCTGCGCCCCATCGGAGGTGCACCCGCTGAG -41

DSS AGCGTCTGGGCACTGAAAGCCGGTGGCTCACCTGCCGGCC -1
Lewis AGCGTCTGGGCACTGAAAGCCGGTGGCTCACCTGCCGGCC -1

M E P H G N D S D F L L A P N G S R A P
DSS **ATG**GAGCCACACGGGAATGACAGCGACTTCTTGCTGGCACCCCAATGGAAGCCGAGCGCCA 60
Lewis **ATG**GAGCCACACGGGAATGACAGCGACTTCTTGCTGGCACCCCAATGGAAGCCGAGCGCCA 60

G H D I T Q E R D E A W V V G M A I L M
DSS GGCCACGACATCACTCAGGAACGGGACGAAGCGTGGTGGTGGGCATGGCCATCCTCATG 120
Lewis GGCCACGACATCACTCAGGAACGGGACGAAGCGTGGTGGTGGGCATGGCCATCCTCATG 120

S V I V L A I V F G N V L V I T A I A K
DSS TCGGTTATCGTCTGGCCATCGTGTGTTGGCAACGTGCTGGTGCATCACAGCCATTGCCAAG 180
Lewis TCGGTTATCGTCTGGCCATCGTGTGTTGGCAACGTGCTGGTGCATCACAGCCATTGCCAAG 180

F E R L Q T V T N Y F I T S L A C A D L
DSS TTCGAGCGACTACAAACCGTCACCAACTACTTCATAACCTCCTTGGCGTGTGCTGATCTA 240
Lewis TTCGAGCGACTACAAACCGTCACCAACTACTTCATAACCTCCTTGGCGTGTGCTGATCTA 240

V M G L A V V P F G A S H I L M K M W N
DSS GTCATGGGCCTAGCGGTGGTGCCTTTTGGGGCCAGTCATATCCTTATGAAAATGTGGAAT 300
Lewis GTCATGGGCCTAGCGGTGGTGCCTTTTGGGGCCAGTCATATCCTTATGAAAATGTGGAAT 300

F G N F W C E F W T S I D V L C V T A S
DSS TTTGGCAATTTCTGGTGGCAGTCTGGACTTCCATTGATGTGTTGTGCGTCACAGCCAGC 360
Lewis TTTGGCAATTTCTGGTGGCAGTCTGGACTTCCATTGATGTGTTGTGCGTCACAGCCAGC 360

I E T L C V I A V D R Y V A I T S P F K
DSS ATCGAGACCCTGTGCGTGATTGCAGTGGATCGCTATGTTGCTATCACATCGCCCTTCAAG 420
Lewis ATCGAGACCCTGTGCGTGATTGCAGTGGATCGCTATGTTGCTATCACATCGCCCTTCAAG 420

Y Q S L L T K N K A R V V I L M V W I V
DSS TACCAGAGCCTGCTGACCAAGAATAAGGCCCGAGTGGTGCATCCTAATGGTGTGGATTGTG 480
Lewis TACCAGAGCCTGCTGACCAAGAATAAGGCCCGAGTGGTGCATCCTAATGGTGTGGATTGTG 480

S G L T S F L P I Q M H W Y R A T H K Q
DSS TCGGGCCTTACCTCCTTCTTGCTATCCAGATGCCTGGTACCGTGCCACCCACAAGCAA 540
Lewis TCGGGCCTTACCTCCTTCTTGCTATCCAGATGCCTGGTACCGTGCCACCCACAAGCAA 540

```

A I D C Y A K E T C C D F F T N O A Y A
DSS GCCATCGACTGTTATGCCAAGGAGACTTGCTGTGACTTCTTCACGAACCAGGCCTATGCT 600
Lewis GCCATCGACTGTTATGCCAAGGAGACTTGCTGTGACTTCTTCACGAACCAGGCCTATGCT 600
*****
I A S S I V S F Y V P L V V M V F V Y
DSS ATCGCTTCTCTATCGTATCTTTCTACGTGCCCTGGTGGTCATGGTCTTTGTCTATTCC 660
Lewis ATCGCTTCTCTATCGTATCTTTCTACGTGCCCTGGTGGTCATGGTCTTTGTCTATTCC 660
*****
R V F Q V A K R Q L Q K I D K S E G R F
DSS AGGGTCTTCCAGGTGGCCAAAAGGCAGCTGCAGAAGATAGACAAATCCGAGGGCAGATTT 720
Lewis AGGGTCTTCCAGGTGGCCAAAAGGCAGCTGCAGAAGATAGACAAATCCGAGGGCAGATTT 720
*****
H A Q N L S Q V E Q D G R S E H G L R R
DSS CATGCCAAAACCTCAGCCAGTGGAGCAGGATGGGAGGAGCGGGCAGGACTCCGAAGG 780
Lewis CATGCCAAAACCTCAGCCAGTGGAGCAGGATGGGAGGAGCGGGCAGGACTCCGAAGG 780
*****
S S K F C L K E H K A L K T L G I I M G
DSS TCCTCCAAGTTCTGCTTGAAGAGCACAAAGCCCTCAAGACTTTAGGCATCATCATGGGC 840
Lewis TCCTCCAAGTTCTGCTTGAAGAGCACAAAGCCCTCAAGACTTTAGGCATCATCATGGGC 840
*****
T F T L C W L P F F I V N I V H V I R A
DSS ACCTTCACCCCTCGCTGGCTGCCCTTCTTCATTGTCAATATTGTCCACGTATCCGGGGC 900
Lewis ACCTTCACCCCTCGCTGGCTGCCCTTCTTCATTGTCAATATTGTCCACGTATCCGGGGC 900
*****
N L I P K E V Y I L L N W L E Y V N S A
DSS AACCTCATCCCTAAGGAAGTTTACATCCTCCTTAACTGGTTGGGCTATGTCAACTCTGCC 960
Lewis AACCTCATCCCTAAGGAAGTTTACATCCTCCTTAACTGGTTGGGCTATGTCAACTCTGCC 960
*****
F N P L I Y C R S P D F R I A F Q E L L
DSS TTCAATCCTTTATCTACTGTCCGAGTCCAGATTTCCAGGATTGCCTTCCAGGAGCTTCTG 1020
Lewis TTCAATCCTTTATCTACTGTCCGAGTCCAGATTTCCAGGATTGCCTTCCAGGAGCTTCTG 1020
*****
C L R R S S S K T Y G N G Y S S N S N G
DSS TGCCTTCGCCGGTCTTCTTCGAAAACCTATGGGAACGGCTACTCTAGCAACAGCAACGGT 1080
Lewis TGCCTTCGCCGGTCTTCTTCGAAAACCTATGGGAACGGCTACTCTAGCAACAGCAACGGT 1080
*****
R T D Y T G E Q S A Y Q L G Q E K E N E
DSS AGGACAGACTACACAGGGAGCAGAGCGCATATCAGCTGGGGCAGGAGAAAGAAAATGAA 1140
Lewis AGGACAGACTACACAGGGAGCAGAGCGCATATCAGCTGGGGCAGGAGAAAGAAAATGAA 1140
*****
L L C E E A P G M E G F V N C Q G T V P
DSS CTGCTTTGTGAGGAAGCCCTGGCATGGAAGGCTTTGTGAACTGTCAAGGTACTGTGCCT 1200
Lewis CTGCTTTGTGAGGAAGCCCTGGCATGGAAGGCTTTGTGAACTGTCAAGGTACTGTGCCT 1200
*****
S L S I D S Q G R N C N T N D S P L -
DSS AGCCTTAGCATTGACTCTCAAGGGAGGAAGTGAACACAAACGACTCGCCACTG TAG 1257
Lewis AGCCTTAGCATTGACTCTCAAGGGAGGAAGTGAACACAAACGACTCGCCACTG TAG 1257
*****
TGTAGGCTTTCTACTCTTTAAGACGGCCCTCCCTGACAGGACACTAACCAGACTATTTT 1317
DSS TGTAGGCTTTCTACTCTTTAAGACGGCCCTCCCTGACAGGACACTAACCAGACTATTTT 1317
Lewis TGTAGGCTTTCTACTCTTTAAGACGGCCCTCCCTGACAGGACACTAACCAGACTATTTT 1317
*****
ACTTGAGTGAATAACTTTAGAATAAAATTTGATAGAGATTTTGCAGAAGGGGCACATCC 1377
DSS ACTTGAGTGAATAACTTTAGAATAAAATTTGATAGAGATTTTGCAGAAGGGGCACATCC 1377
Lewis ACTTGAGTGAATAACTTTAGAATAAAATTTGATAGAGATTTTGCAGAAGGGGCACATCC 1377
*****
TTCTGCCCTTTTTATTTTATTTTAAAGCTGCAAGTGAGAGAGAGAGGGGAGAGAGAG 1437
DSS TTCTGCCCTTTTTATTTTATTTTAAAGCTGCAAGTGAGAGAGAGAGGGGAGAGAGAG 1437
Lewis TTCTGCCCTTTTTATTTTATTTTAAAGCTGCAAGTGAGAGAGAGAGGGGAGAGAGAG 1437
*****

```

```

DSS      AGAGAGAGAGAGAGAAAGAGAGAGAGAGAGACTGTATTTCAAGTGCCTTAATGTTTTCCCTTG 1497
Lewis    AGAGAGAGAGAGAGAAAGAGAGAGAGAGAGACTGTATTTCAAGTGCCTTAATGTTTTCCCTTG 1497
*****
DSS      TACAGTTCAGTTCCTCTTTGCATGGAACCTAAAAGTTTCTGTCTGAAGTGTGGTGGTCT 1557
Lewis    TACAGTTCAGTTCCTCTTTGCATGGAACCTAAAAGTTTCTGTCTGAAGTGTGGTGGTCT 1557
*****
DSS      GTCGTGTTGGATGGTGTTCATGTATCTACCTCACTGGTCAAGTATTAAGGATAATATA 1617
Lewis    GTCGTGTTGGATGGTGTTCATGTATCTACCTCACTGGTCAAGTATTAAGGATAATATA 1617
*****
DSS      TATATATATTGCTGCTGGAATTCATATGTAAGGAGAGAGTTTTCTTCCCTGTACCCTTG 1677
Lewis    TATATATATTGCTGCTGGAATTCATATGTAAGGAGAGAGTTTTCTTCCCTGTACCCTTG 1677
*****
DSS      GACTTGAAATATCCTGTGCTTTGGACCTTTCTGCTGTGAATGTGGACTCTCTCACTCC 1737
Lewis    GACTTGAAATATCCTGTGCTTTGGACCTTTCTGCTGTGAATGTGGACTCTCTCACTCC 1737
*****
DSS      ACTTATTCGCTCAAATGGAGTGTGTAGGCAGGGATTTGAGGGACAACATCAGTTGTTTT 1797
Lewis    ACTTATTCGCTCAAATGGAGTGTGTAGGCAGGGATTTGAGGGACAACATCAGTTGTTTT 1797
*****
DSS      TCTGAGCAAAGTCTAAAGTTTACAGTAAATAAATTGTTTGACCATGA 1844
Lewis    TCTGAGCAAAGTCTAAAGTTTACAGTAAATAAATTGTTTGACCATGA 1844
*****

```

Footnote to Supplement 2: * denotes nucleotide identity. Initiation and termination codons are marked by bold and italic letters. Sequences for DSS and Lewis are deposited to GenBank (#GQ160814). 5' promoter and 3' untranslated regions are given before the ATG initiation codon and after the TAG stop codon respectively.

Supplement 3. Aligned coding sequence comparison of neural precursor cell expressed, developmentally down-regulated 4-like gene (*Nedd4l*) between Dahl salt-sensitive (DSS) and Lewis strains

```

                                     ↓
DSS      M S E S P E R R G E S R I L R V K V V S
Lewis    ATBTCTGAGTCTCCCGAGAGGAGGGGAGAATCCCGTATTCTCAGAGTAAAAGTTGTTTCT 60
          ATBTCTGAGTCTCCCGAGAGGAGGGGAGAATCCCGTATTCTCAGAGTAAAAGTTGTTTCT 60
          *****

                                     ↓
DSS      G I D L A K K D I F G A S D P Y V K L S
Lewis    GSGATTGACCTCGCCAAAAAGGACATATTTGGAGCCAGTGACCATACTGAAGCTCTCG 120
          GSGATTGACCTCGCCAAAAAGGACATATTTGGAGCCAGTGACCATACTGAAGCTCTCG 120
          *****

                                     ↓
DSS      L Y V A D E N R E L A L V Q T E T I K K
Lewis    TTGTACGTAGCTGACGAGAACAGAGAAGTGGCTTGCTTCAGACTAAGACAATTA AAAAG 180
          TTGTACGTAGCTGACGAGAACAGAGAAGTGGCTTGCTTCAGACTAAGACAATTA AAAAG 180
          *****

                                     ↓
DSS      T L N P K W N E E F Y F R V M P S H H R
Lewis    ACGCTGAACCCAAAGTGGAAATGAAGAGTTTTATTTTAGAGTAAATCCATCTAATCACAGG 240
          ACGCTGAACCCAAAGTGGAAATGAAGAGTTTTATTTTAGAGTAAATCCATCTAATCACAGG 240
          *****

                                     ↓
DSS      L L F E V F D E N R L T R D D F L G Q V
Lewis    CTCTTATTTGAAGTATTTGACGAGAACAGATTGACACGAGACGACTTTCTGGGCCAGGTG 300
          CTCTTATTTGAAGTATTTGACGAGAACAGATTGACACGAGACGACTTTCTGGGCCAGGTG 300
          *****

                                     ↓
DSS      D V P L S H L P T E D P T W E R F Y T F
Lewis    GACGTGCCTCTTAGTCACCTCCGACAGAAATCCAACCATGGAGAGACCTATAACATT 360
          GACGTGCCTCTTAGTCACCTCCGACAGAAATCCAACCATGGAGAGACCTATAACATT 360
          *****

                                     ↓
DSS      K D F L L R F R S H K S R V K G F L R L
Lewis    AAGGATTTTCTCCTGCGACCTAGAAGTCATAAATCACGCGTCAAGGGGTTTTGAGGTTG 420
          AAGGATTTTCTCCTGCGACCTAGAAGTCATAAATCACGCGTCAAGGGGTTTTGAGGTTG 420
          *****

DSS      K M A Y M P R N G G Q D E E N S E Q R D
Lewis    AAAATGGCCTACATGCCGAAAAACGGAGGTCAGGATGAAGAAAAACAGCGAGCAGAGGAC 480
          AAAATGGCCTACATGCCGAAAAACGGAGGTCAGGATGAAGAAAAACAGCGAGCAGAGGAC 480
          *****

                                     ↓
DSS      D M E H G W E V V D S N D S A S Q H Q E
Lewis    GACATGGAGCATGGATGGGAAGTTGTTGACTCAAACGACTCAGCTTCCAGCACCAGGAG 540
          GACATGGAGCATGGATGGGAAGTTGTTGACTCAAACGACTCAGCTTCCAGCACCAGGAG 540
          *****

DSS      E L P P P P L P P G W E E K V D N L G R
Lewis    GAGCTCCCTCCTCCTCCCTGCCACCAGGATGGGAAGAGAAAAGTGACAAATTTAGGCCGA 600
          GAGCTCCCTCCTCCTCCCTGCCACCAGGATGGGAAGAGAAAAGTGACAAATTTAGGCCGA 600
          *****

                                     ↓
DSS      T Y Y V N H N N R S T Q W H R F S L M D
Lewis    ACTTACTATGTCAACCACAACAACAGGAGTACGCAAGTGGCACCAGCTAGCTTGTATGGAT 660
          ACTTACTATGTCAACCACAACAACAGGAGTACGCAAGTGGCACCAGCTAGCTTGTATGGAT 660
          *****

```

```

V S S E S D N H I R Q I N Q E A A H R R
DSS GTATCGTCAGAGTCAGACAATAACATCAGGCAGATCAATCAGGAGGCCGCACACCCGGCT 720
Lewis GTATCGTCAGAGTCAGACAATAACATCAGGCAGATCAATCAGGAGGCCGCACACCCGGCT 720
*****

F R S R R H I S E D L E P E A S E G G G
DSS TTCCGCTCTCGGAGGCACATTAGTGAAGACTTGGAGCCTGAGGCCTCTGAAGGCGGTGGA 780
Lewis TTCCGCTCTCGGAGGCACATTAGTGAAGACTTGGAGCCTGAGGCCTCTGAAGGCGGTGGA 780
*****

↓

E G F E P W E T I S E E V N M A G D S L
DSS GAAGGCCCTGAGCCTTGGGAGACCATTTAGAGGAAGTGAACATGGCGGGAGATTCTCTC 840
Lewis GAAGGCCCTGAGCCTTGGGAGACCATTTAGAGGAAGTGAACATGGCGGGAGATTCTCTC 840
*****

S L A L P P F F A S P V S R T S F Q E L
DSS AGCCTGGCTTACCCCACTCTGCTCCCACTGTCAGGACCCAGCCCAAGSAGCTG 900
Lewis AGCCTGGCTTACCCCACTCTGCTCCCACTGTCAGGACCCAGCCCAAGSAGCTG 900
*****

S E E L S R R L Q I T F D S N G E Q F S
DSS TCGGAAGACTGAGCAGAAGTTGCAGATCACTCCGGACTCCAATGGGGAACAGTTCAGT 960
Lewis TCGGAAGACTGAGCAGAAGTTGCAGATCACTCCGGACTCCAATGGGGAACAGTTCAGT 960
*****

↓

A L I Q R E P S S R L R S C S V T D T V
DSS GCTCTGATCCAGAGAGAGCCCTCGTCTAGGCTCCGGTCTGCAAGTGTACCGACAGGTT 1020
Lewis GCTCTGATCCAGAGAGAGCCCTCGTCTAGGCTCCGGTCTGCAAGTGTACCGACAGGTT 1020
*****

↓

A E Q A H L P P P S T P T R R A R S S T
DSS GCTGAACAAGCTCACCTTCCACCGCCAGCACCCCAACTAGGCGAGCCCGTTGCTCAACT 1080
Lewis GCTGAACAAGCTCACCTTCCACCGCCAGCACCCCAACTAGGCGAGCCCGTTGCTCAACT 1080
*****

↓

V T G G E E S T P S V A Y V H T T P G L
DSS GTCACGGTGGTGGAGAAATCAACGCCATCAGTGGCCTATGTACATACCAGCCGGGCTA 1140
Lewis GTCACGGTGGTGGAGAAATCAACGCCATCAGTGGCCTATGTACATACCAGCCGGGCTA 1140
*****

P S S W E E R K D A K G R T Y Y V N H N
DSS CCTTCAGGCTGGGAGAAAAGAAAAGATGCTAAGGGACGCACATACTATGTCAATCATAAC 1200
Lewis CCTTCAGGCTGGGAGAAAAGAAAAGATGCTAAGGGACGCACATACTATGTCAATCATAAC 1200
*****

↓

N R T T T W T R F I M Q L A E D G A S G
DSS AATCGAACCAACTTGGACTCGACCAATCATGCAAGTTCAGAAAGACGGTSCCTCTGGA 1260
Lewis AATCGAACCAACTTGGACTCGACCAATCATGCAAGTTCAGAAAGACGGTSCCTCTGGA 1260
*****

S A T N S N N H L V E P Q I R R P R S L
DSS TCAGCCACAACAGTAACAACCACTAGTCAACCCAGATCCGCGGGCTCGTAGCCTC 1320
Lewis TCAGCCACAACAGTAACAACCACTAGTCAACCCAGATCCGCGGGCTCGTAGCCTC 1320
*****

↓

S S P T V T L S A F L E G A K D S F I R
DSS AGCTCGCCAACAGTAACCTTATCTGCCCACTGGAGGTTGCCAAGGATTCAACCCATACGC 1380
Lewis AGCTCGCCAACAGTAACCTTATCTGCCCACTGGAGGTTGCCAAGGATTCAACCCATACGC 1380
*****

R A V K D T L S N P Q S F Q F S F Y N S
DSS CSTGCTGTGAAGATACCCCTTCCCAATCCACAGTCCCTCAGCCATCACTTACAACCTC 1440
Lewis CSTGCTGTGAAGATACCCCTTCCCAATCCACAGTCCCTCAGCCATCACTTACAACCTC 1440
*****

P K P Q H K V T Q S F L P F G W E M R I
DSS CCCAAACCACAACAAAGTCACACAGAGCTTCTGCCACCAGGCTGGGAGATGAGGATA 1500
Lewis CCCAAACCACAACAAAGTCACACAGAGCTTCTGCCACCAGGCTGGGAGATGAGGATA 1500
*****

```

```

          ↓
    A P H G R P F F I D E N T K T T T W E D
DSS      GCGCCAAACGGCCGACCCCTTCTTCATTGACCATAACACAAAGACGACAACCTGGGAAGAT 1560
Lewis    GCGCCAAACGGCCGACCCCTTCTTCATTGACCATAACACAAAGACGACAACCTGGGAAGAT 1560
          *
          *
          *
    F R L K F P V H M R S K A S L N P N D L
DSS      CCACGGCTGAAATTTCCAGTACATATGCGSTCAAAAAGCATCTTAAACCCCAATGACCTG 1620
Lewis    CCACGGCTGAAATTTCCAGTACATATGCGSTCAAAAAGCATCTTAAACCCCAATGACCTG 1620
          *
          *
          *
          ↓
    G P L F P G W E E R I H L D G R T F Y I
DSS      GSCCCTCTTCTCCTGGCTGGGAAGAAAAGSATCCACCTGGACGGCCGCACATTTTACATT 1680
Lewis    GSCCCTCTTCTCCTGGCTGGGAAGAAAAGSATCCACCTGGACGGCCGCACATTTTACATT 1680
          *
          *
          *
          ↓
    D H N S K I T Q W E D F R L Q N P A I T
DSS      GACCATAATAGCAAATTACTCAGTGGGAAGATCCAAGACTACAGAATCCAGCCATAACT 1740
Lewis    GACCATAATAGCAAATTACTCAGTGGGAAGATCCAAGACTACAGAATCCAGCCATAACT 1740
          *
          *
          *
          ↓
    G P A V P Y S R E F K Q K Y D Y F R K
DSS      GGTCCGGCTGTGCCSTACTCCAGAGAATTTAAGCAGAAATACGACTACTTTAGGAAGAAA 1800
Lewis    GGTCCGGCTGTGCCSTACTCCAGAGAATTTAAGCAGAAATACGACTACTTTAGGAAGAAA 1800
          *
          *
          *
          ↓
    L K E P A D I P N R F E M K L H R N N I
DSS      TTAAGAAACCTGCTGATATCCAAACAGSTTTGAAATGAAACTCCACAGAAACAACATA 1860
Lewis    TTAAGAAACCTGCTGATATCCAAACAGSTTTGAAATGAAACTCCACAGAAACAACATA 1860
          *
          *
          *
    F E E S Y R R I N E V K R P D V L K A R
DSS      TTTGAAAGAAATCCTATCGGAGGATCATGTCTGTAAAGAGACTGACGTCTTAAAGGCTAGA 1920
Lewis    TTTGAAAGAAATCCTATCGGAGGATCATGTCTGTAAAGAGACTGACGTCTTAAAGGCTAGA 1920
          *
          *
          *
    L W I E F E S E K G L D Y G G V A R E W
DSS      CTGTGGATCGAGTTTGAATCAGAGAAAAGGCTGGACTATGGGGCGTGGCCAGAGAGTGG 1980
Lewis    CTGTGGATCGAGTTTGAATCAGAGAAAAGGCTGGACTATGGGGCGTGGCCAGAGAGTGG 1980
          *
          *
          *
    F F L L S K E M F M P Y Y G L F E Y S A
DSS      TTCTTTACTGTCTAAAGAGATGTTCAATCCCTACTATGGCCTCTTCGAGTACTCTGCC 2040
Lewis    TTCTTTACTGTCTAAAGAGATGTTCAATCCCTACTATGGCCTCTTCGAGTACTCTGCC 2040
          *
          *
          *
          ↓
    T D N Y T L Q I N P N S G L C N E D H L
DSS      ACGGACAACCTACACACTTCAGATCAATCCCAACTCAGGCTGTGTAAACGAAGATCACTTA 2100
Lewis    ACGGACAACCTACACACTTCAGATCAATCCCAACTCAGGCTGTGTAAACGAAGATCACTTA 2100
          *
          *
          *
    S Y F T F I G R V A G L A V F R G K L L
DSS      TCCTATTTCACTTTCATTGGGAGAGTTGCTGGCCTGGCCTGTGTTTCATGGGAAACTCTTA 2160
Lewis    TCCTATTTCACTTTCATTGGGAGAGTTGCTGGCCTGGCCTGTGTTTCATGGGAAACTCTTA 2160
          *
          *
          *
          ↓
    D G F F I R F F Y K M M L G K Q I T L N
DSS      GATGGATTCTTCATTGACCTTTCTACAAGATGATGCTGGGSAAGCAAATAACTCTGAAT 2220
Lewis    GATGGATTCTTCATTGACCTTTCTACAAGATGATGCTGGGSAAGCAAATAACTCTGAAT 2220
          *
          *
          *
          ↓
    D M E S V D S E Y Y N S L K W I L E N D
DSS      GACATGGAGTCTGTGGACAGCGAGTACTACAACCTTTTGAAGTGGATCTTGGAAAACGAC 2280
Lewis    GACATGGAGTCTGTGGACAGCGAGTACTACAACCTTTTGAAGTGGATCTTGGAAAACGAC 2280
          *
          *
          *

```

```

          P T E L D L M F C I D E E N F G Q T Y Q
DSS      CCCACGGAACTTGACCTCATGTTCTGCATAGATGAAGAGAACTTTGGGCAGACTTACCAA 2340
Lewis    CCCACGGAACTTGACCTCATGTTCTGCATAGATGAAGAGAACTTTGGGCAGACTTACCAA 2340
          *
          V D L K F N G E E I M V T N E N K R E Y
DSS      GTGGACCTGAAGCCCAATGGGTGAGAAATCATGGTAACCAATGAGAACAAAGCGAGAATAT 2400
Lewis    GTGGACCTGAAGCCCAATGGGTGAGAAATCATGGTAACCAATGAGAACAAAGCGAGAATAT 2400
          *
          I D L V I Q W R F V N R V Q K Q M N A F
DSS      ATTGACTTGGTCATCCASTGGAGATTTGTGAACAGGGTCCAGAAGCAAAATGAATGCCCTTC 2460
Lewis    ATTGACTTGGTCATCCASTGGAGATTTGTGAACAGGGTCCAGAAGCAAAATGAATGCCCTTC 2460
          *
          L E G F T E L L F I D L I K I F D E N E
DSS      TTGGAGGGATTACAGAACTTCTTCCAATCGACTTGATTAATAATTTTGGATGAAAATGAG 2520
Lewis    TTGGAGGGATTACAGAACTTCTTCCAATCGACTTGATTAATAATTTTGGATGAAAATGAG 2520
          *
          L E L L M C G L G D V D V N D W R Q H S
DSS      CTGGAGTTGCTGATGTCCGGCCTTGGTGATGTGGACGTGAATGATGGAGACAGCACTCC 2580
Lewis    CTGGAGTTGCTGATGTCCGGCCTTGGTGATGTGGACGTGAATGATGGAGACAGCACTCC 2580
          *
          I Y K N G Y C P N H P V I Q W F W E A V
DSS      ATTTACAAGAACGGCTACTGCCCAACCACTGTCCAGTGGTTCGGAAAGGCCGTG 2640
Lewis    ATTTACAAGAACGGCTACTGCCCAACCACTGTCCAGTGGTTCGGAAAGGCCGTG 2640
          *
          L L M D A E K R I R L L Q F V T G T S R
DSS      CTGCTGATGGATGCTGAGAAGCGGATCCGGTACTACAGTTTGTACAGGCACCTCCAGA 2700
Lewis    CTGCTGATGGATGCTGAGAAGCGGATCCGGTACTACAGTTTGTACAGGCACCTCCAGA 2700
          *
          V P M N G F A E L Y G S N G P Q L F T I
DSS      GTACCCATGAATGGATTGCGGAACTTTATGGTTCCAAATGGTCTCAGCTGTTTACAATA 2760
Lewis    GTACCCATGAATGGATTGCGGAACTTTATGGTTCCAAATGGTCTCAGCTGTTTACAATA 2760
          *
          E Q N G S P E K L P R A H T C F H R L D
DSS      GAGCAATGGGGTAGTCTGAAAAACTGCCAGAGCTCATACTGCTTTAACCGCCTTGAT 2820
Lewis    GAGCAATGGGGTAGTCTGAAAAACTGCCAGAGCTCATACTGCTTTAACCGCCTTGAT 2820
          *
          L P P Y E T F E D L R E K L L M A V E N
DSS      TTACCTCCGTATGAAACCTTTGAAGATCTACGGGAGAACTTCTCATGGCCGTGAGAAAT 2880
Lewis    TTACCTCCGTATGAAACCTTTGAAGATCTACGGGAGAACTTCTCATGGCCGTGAGAAAT 2880
          *
          A Q G F E G V D -
DSS      GCTCAAGGCTTCGAAGGTGTGGATTAA 2907
Lewis    GCTCAAGGCTTCGAAGGTGTGGATTAA 2907
          *

```

Footnote to Supplement 3: * denotes amino acid identity. ↓ indicates the last nucleotide of each exon. Exon and intron junctions are not different (data not shown). Genomic DNAs and/or cDNA from kidneys from DSS and Lewis were first amplified by PCR and then sequenced. When a nucleotide difference was seen (indicated by shade), it was re-PCR'd and sequence-confirmed.

Supplement 4. Aligned comparison of putative promoter, exons #1, 2 and 3 and their junctions of neural precursor cell expressed, developmentally down-regulated 4-like gene (*Nedd4l*) between Dahl salt-sensitive (DSS) and Lewis rats.

Promoter and exon 1:

DSS	AGATCCTATCCTTTAGGACTGAGTGAGCAAAGTCTCTCACTCTCTACACATTATCTAGCT	-1606
Lewis	AGATCCTATCCTTTAGGACTGAGTGAGCAAAGTCTCTCACTCTCTACACATTATCTAGCT	-1606

DSS	GTGAGACTCATGTTAACTCCCATCTGTATAAGAAAGGCTCTGTGATAAAGGTTGGTCAA	-1546
Lewis	GTGAGACTCATGTTAACTCCCATCTGTATAAGAAAGGCTCTGTGATAAAGGTTGGTCAA	-1546

DSS	CAAGCTCTGCTTGAACCTCTGGGTTCAAGGCAAGTCACATCAGAGGGGCAAGTGGT	-1486
Lewis	CAAGCTCTGCTTGAACCTCTGGGTTCAAGGCAAGTCACATCAGAGGGGCAAGTGGT	-1486

DSS	AGGGGTGGGTGGTAGGTAAAGCAAGGCAAGGCGAGGGTGCCTGTGGTCTTGTCCCAGGA	-1426
Lewis	AGGGGTGGGTGGTAGGTAAAGCAAGGCAAGGCGAGGGTGCCTGTGGTCTTGTCCCAGGA	-1426

DSS	CTCAGTGGTGGTGGTAGCTGACTTTGGGAATCCTTACTTCAAGGCTCTGGTGGAAATCAA	-1366
Lewis	CTCAGTGGTGGTGGTAGCTGACTTTGGGAATCCTTACTTCAAGGCTCTGGTGGAAATCAA	-1366

DSS	CTGCAAGCAGTCATTATCCATCTTTATATGTAAGTACACATGTGTGTGCAGGTACATGT	-1306
Lewis	CTGCAAGCAGTCATTATCCATCTTTATATGTAAGTACACATGTGTGTGCAGGTACATGT	-1306

DSS	GTGGAGGGCCGAGCTCCATGCAAGGCTCTCTTCTTCAATAACTTTGAACCTTATCTTTGAG	-1246
Lewis	GTGGAGGGCCGAGCTCCATGCAAGGCTCTCTTCTTCAATAACTTTGAACCTTATCTTTGAG	-1246

DSS	ACTGGGTCCTCACTGAATGGGTTGACTAGGATCCCCAGGAACCATCTGATCTCCCT	-1186
Lewis	ACTGGGTCCTCACTGAATGGGTTGACTAGGATCCCCAGGAACCATCTGATCTCCCT	-1186

DSS	TCAACATCACTAGGATGGCAGGCACTTGTACCCAGAGGGTTTGGCTTCATTTGTTTTAC	-1126
Lewis	TCAACATCACTAGGATGGCAGGCACTTGTACCCAGAGGGTTTGGCTTCATTTGTTTTAC	-1126

DSS	ATATGTGTTGGATACTAGGGATCCCAACTCAGGTTTTCTGTGTTAOCCTGAACCATCTCTC	-1066
Lewis	ATATGTGTTGGATACTAGGGATCCCAACTCAGGTTTTCTGTGTTAOCCTGAACCATCTCTC	-1066

DSS	CAGTCCCTTATACCTATTATAACGATATTATACCAITATATGATAATTACATCATATTGC	-1006
Lewis	CAGTCCCTTATACCTATTATAACGATATTATACCAITATATGATAATTACATCATATTGC	-1006

DSS	TACGATATAGCCTATTATACCATTAGAATGATAATATTGTAAGTTCATTGTTGTAGAGGC	-946
Lewis	TACGATATAGCCTATTATACCATTAGAATGATAATATTGTAAGTTCATTGTTGTAGAGGC	-946

DSS	TTTTAGATTTTTCTCAGATGAAAATGAGTCTTGACCTCCAACCTCTCAAGATACCTTTTAT	-886
Lewis	TTTTAGATTTTTCTCAGATGAAAATGAGTCTTGACCTCCAACCTCTCAAGATACCTTTTAT	-886

DSS	ACTEATCTATGACTTAAAAGATCTTTCAAGTGTGTGTATGTGTATGATGATGATGAGT	-826
Lewis	ACTEATCTATGACTTAAAAGATCTTTCAAGTGTGTGTATGTGTATGATGATGATGAGT	-826

DSS	GCATGTATGTGTGAGTGTACGTGTGCATGTGTGTATATGCAGAAAAAGGAGTGTATGT	-766
Lewis	GCATGTATGTGTGAGTGTACGTGTGCATGTGTGTATATGCAGAAAAAGGAGTGTATGT	-766

DSS	GTATTTGTATATGAGGGGGAACATGCATATGTGTACCACACATAGTACACATGTGAAGAT	-706
Lewis	GTATTTGTATATGAGGGGGAACATGCATATGTGTACCACACATAGTACACATGTGAAGAT	-706

DSS	CAGAGAACAGCCTTTGAAAAGTGGCTTTGAAAACACCGCATGCACGTGCTAGTTCCTCTGG	-646
Lewis	CAGAGAACAGCCTTTGAAAAGTGGCTTTGAAAACACCGCATGCACGTGCTAGTTCCTCTGG	-646

DSS	AGCTCCTGGGATCTGTGTCCACCTTCCATGCTTCGTAGGAGCACTGTGGTTACAGACTC	-586
Lewis	AGCTCCTGGGATCTGTGTCCACCTTCCATGCTTCGTAGGAGCACTGTGGTTACAGACTC	-586

DSS	CGTGCCTGGGCTCCGAGGGTCTGAACTGAGGCTCTCGCAITTTGATATGTGTGGCAAGCGC	-526
Lewis	CGTGCCTGGGCTCCGAGGGTCTGAACTGAGGCTCTCGCAITTTGATATGTGTGGCAAGCGC	-526

```

*****
DSS      TGCTCACTGTGCTTTCTCCCCATCCTCCTAAGATTTGCTTTTTTAGSACAGATTCATAA -466
Lewis    TGCTCACTGTGCTTTCTCCCCATCCTCCTAAGATTTGCTTTTTTAGSACAGATTCATAA -466
*****
DSS      ATGAGATTCGTGTGAGTAGATCATATGAATGACATTTAGGGGATTTAGAGACACTAATC -406
Lewis    ATGAGATTCGTGTGAGTAGATCATATGAATGACATTTAGGGGATTTAGAGACACTAATC -406
*****
DSS      CTTTATGCAAGCGTCTGCTGGGCTGTGACTTCCCTGTATTATGAGGGCTCATGGGAGGTGC -346
Lewis    CTTTATGCAAGCGTCTGCTGGGCTGTGACTTCCCTGTATTATGAGGGCTCATGGGAGGTGC -346
*****
DSS      TCCAGAAGCAGGSAAGGCCGGCTGCAGGGCAGGTTGACCTCAGTATCTTCTTGTAGA -286
Lewis    TCCAGAAGCAGGSAAGGCCGGCTGCAGGGCAGGTTGACCTCAGTATCTTCTTGTAGA -286
*****
DSS      TTTCAAGTCAAGTGTCTACATTCACCTAAGGGGAGGACCTGATTTCTCCAGAAAAGTCTAG -226
Lewis    TTTCAAGTCAAGTGTCTACATTCACCTAAGGGGAGGACCTGATTTCTCCAGAAAAGTCTAG -226
*****
DSS      CTTTCTCTTATTTCAGTTTATTTTTCATCATTACCTATATGTGAGTGTTCCTGATT -166
Lewis    CTTTCTCTTATTTCAGTTTATTTTTCATCATTACCTATATGTGAGTGTTCCTGATT -166
*****
DSS      TTGCTTAATTTGGSAATGGGCCACCCAGSTAAGACCATTGCTCGGTGGSAACGCAAA -106
Lewis    TTGCTTAATTTGGSAATGGGCCACCCAGSTAAGACCATTGCTCGGTGGSAACGCAAA -106
*****
DSS      ACTTGGAAACCGGTTGTGAACCTAGCCTCATCTGTTTGTCTTCCAGACCCGATAACAGT -46
Lewis    ACTTGGAAACCGGTTGTGAACCTAGCCTCATCTGTTTGTCTTCCAGACCCGATAACAGT -46
*****
DSS      TTGTCGAGGTCCTAACAAATTATCTTTGCTGGGCAGGACTCGGG -1
Lewis    TTGTCGAGGTCCTAACAAATTATCTTTGCTGGGCAGGACTCGGG -1
*****
DSS (Exon #1)  ATGTCTGAGTCTCCCGAGAGGAGGTTGAGCAGCTTGTCTTCTACCCCGTGGCGTGTGACC 60
Lewis         ATGTCTGAGTCTCCCGAGAGGAGGTTGAGCAGCTTGTCTTCTACCCCGTGGCGTGTGACC 60
*****
DSS          GTTAGTGCCTTGGCTCCATCATAACTGTTATCCCTTCCCTCCTCCC 106
Lewis         GTTAGTGCCTTGGCTCCATCATAACTGTTATCCCTTCCCTCCTCCC 106
*****

```

Exon #2:

```

DSS (Exon #2)  ACGGATCGGTTTTAACCTGTGCTTTTTCTCCCTTTTCTAACAGGGAGAATCC 57
Lewis         ACGGATCGGTTTTAACCTGTGCTTTTTCTCCCTTTTCTAACAGGGAGAATCC 57
*****
DSS (Exon #2)  CGTATTCTCAGAGTAAAAGTTGTTCTGGGATTGACCTCGCCAAAAGGACATATTGGA 117
Lewis         CGTATTCTCAGAGTAAAAGTTGTTCTGGGATTGACCTCGCCAAAAGGACATATTGGA 117
*****
DSS (Exon #2)  GCCAGTACGTTTGCCTTTTNGTGTGTGTAGGGGGTGGGAGTGTCTCTGGAAAAGTAT 177
Lewis         GCCAGTACGTTTGCCTTTTNGTGTGTGTAGGGGGTGGGAGTGTCTCTGGAAAAGTAT 177
*****
DSS          AAGCCTTGGTGTGA 192
Lewis         AAGCCTTGGTGTGA 192
*****

```

Exon #3 :

```

DSS (Intron 3)          TTTACATTTTAAATCCATTTTGCTGTTTTTGTGCT 37
Lewis                  TTTACATTTTAAATCCATTTTGCTGTTTTTGTGCT 60
*****

DSS                    TGTTTGTTGCTAATGAAAAAAAAAAGTATAGCACCATTTAACGTGAGCCGTTTCCTTGCA 96
Lewis                  TGTTTGTTGGTAATGAAAAAAAAAAGTATAGCACCATTTAAATGTGAGCCGTTTCCTTGCA 96
*****;*****

DSS (Exon #3)         GTGACCCATACGTGAAGCTCTCGTTGTACGTAGCTGACGAGAACAGAGAACTCGCTCTGG 156
Lewis                  GTGACCCATACGTGAAGCTCTCGTTGTACGTAGCTGACGAGAACAGAGAACTCGCTCTGG 156
*****

DSS (Exon #3)         TTCAGACTAAGACAATTAAAAAGGTAGGTCGGCCAGCTTCTATGAAATTCATGTAAGTCC 216
Lewis                  TTCAGACTAAGACAATTAAAAAGGTAGGTCGGCCAGCTTCTATGAAATTCATGTAAGTCC 216
*****

DSS                    CAATTTGACATTTCTGATTTTTCCTTGTAGCAGTTTGCCTTTAAGSATAAGATAGCAA 276
Lewis                  CAATTTGACATTTCTGATTTTTCCTTGTAGCAGTTTGCCTTTAAGSATAAGATAGCAA 276
*****

DSS                    TTTTTTTTTAGTATTTCTCTGAAAATCATCC 309
Lewis                  TTTTTTTTTAGTATTTCTCTGAAAATCATCC 309
*****

```

Footnote to Supplement 4: * denotes nucleotide identity and: a difference. Exons are highlighted in bold letters. Genomic DNAs from DSS and Lewis were first amplified by PCR and then sequenced.

Supplement 5. Comparison of renal gene expressions of β -2 adrenergic receptor (*Adrb2*) and neural precursor cell expressed, developmentally down-regulated 4-like (*Nedd4l*) by qRT-PCR

Genes	Mean R	Ratio of mean R Lewis/ mean R DSS	<i>p</i>
<i>Adrb2</i>	0.005 ± 0.001 (Lewis)/ 0.009 ± 0.006 (DSS)	0.568	0.37
<i>Nedd4l</i>	0.009 ± 0.001 (Lewis)/ 0.012 ± 0.004 (DSS)	0.752	0.60

Footnote to Supplement 5: R represents the ratio of the expression of the gene target relative to that of the *Gapdh* reference. R is the average value of triplicates from 1 rat. Mean R is the average value of R from 3 rats of the same strain. DSS, Dahl salt-sensitive rats. ± refers to SEM. *P*, *t*-test. Rats were 11 week old males fed a 2% NaCl diet starting from 5 weeks of age. Kidneys from 1 rat of DSS and Lewis respectively were homogenized together and used in triplicate for qRT-PCR. In total, kidneys from 3 rats were used and each in triplicate for each strain. qPCR primers for *Nedd4l* are 5'-CTCTGCCACGGACAACACTACA-3' and 5'-TTA TTT GCT TCC CCA GCA TC-3'); for *Adrb2* are 5'-CGTGCCCTGGTGGTCATGG-3' and 5'-CCCGCTCCTCCCATCCTGCT-3; and for the *Gapdh* standard are 5'ATGGGAAGCTGG TCATCAAC-3' and 5'GTGGTTCACACCCATCACAA-3'. qRT-PCR reactions were performed using a Corbett RotorGene 3000 (Corbett Life Sciences, Mississauga, Ontario, Canada) and the Qiagen HotStarTaq Master Mix Kit (Qiagen, Mississauga, Ontario, Canada) with SYBR Green chemistry.

Supplement 6. Coding sequence alignment of *alpha*-kinase 2 (*Alpk2*) between Dahl salt-sensitive (DSS) and Lewis rats

```

M A D F G C P Q R H T L R F L S T L L S
DSS ATGGCAGATCCAGGCTGCCCCAGAGGCCACACACTACGTTTTTTATCAACGTTACTTTCC 60
LEWIS ATGGCAGATCCAGGCTGCCCCAGAGGCCACACACTACGTTTTTTATCAACGTTACTTTCC 60
*****

Q K V F E K S D V V L R C M I A G Q F K
DSS CAGAAGTTCGGGAGAAGTCTGACGTCGCTCCGATGCATGATAGCCGGTCAGCCCAAG 120
LEWIS CAGAAGTTCGGGAGAAGTCTGACGTCGCTCCGATGCATGATAGCCGGTCAGCCCAAG 120
*****

F E V T W Y K H G Q A I D S D N T V S S
DSS CCAGAGTGCACCTGGTATAAGAATGGTCAAGCCATAGATTCGGAAACACCCTTCCAGT 180
LEWIS CCAGAGTGCACCTGGTATAAGAATGGTCAAGCCATAGATTCGGAAACACCCTTCCAGT 180
*****

Y E F F E N Q Y I R L L L H L P G C A Q S
DSS TACGAGTTCITGAGAACCAGTATATTCATTGGTACATCTCCAGGCTGTGCCAGAGC 240
LEWIS TACGAGTTCITGAGAACCAGTATATTCATTGGTACATCTCCAGGCTGTGCCAGAGC 240
*****

D A A V Y Q V S A R S C V G M I C C S A
DSS GACGCTGCGGTTTATCAGGTCCTCAGCGAAGGCTGCGTGGGGATGATCTGCTGTCAGCC 300
LEWIS GACGCTGCGGTTTATCAGGTCCTCAGCGAAGGCTGCGTGGGGATGATCTGCTGTCAGCC 300
*****

E L E V Q C L Q D P Q V S F D P G G G Q
DSS TCDCGAGGTCACGTCCTGACAGACCCACAGGTCCTCTCTGATCCAGAGGTGGCGSS 360
LEWIS TCDCGAGGTCACGTCCTGACAGACCCACAGGTCCTCTCTGATCCAGAGGTGGCGSS 360
*****

D V A G E R E T E I R E E D S T R R T D
DSS GACGTGGCTGGGAGTGTGAAACAGAGATACGTGAAGAAGACAGCACACCCGACACAG 420
LEWIS GACGTGGCTGGGAGTGTGAAACAGAGATACGTGAAGAAGACAGCACACCCGACACAG 420
*****

C

E K N N P C E T E E S M A A G A Y N S T
DSS GAGAAGTGGAAACCCCTGTAAACAGAGAAAGTATGGCCGACGGGCTTACATGCAACT 480
LEWIS GAGAAGTGGAAACCCCTGTAAACAGAGAAAGTATGGCCGACGGGCTTACATGCAACT 480
*****

P

H S S F D K F D F E C S P Q I V V S R D
DSS CATTCTCCCCAGACAAATTCGACCCCTCGTGTCAACCCAGATAGTGGTCAGTCATGAC 540
LEWIS CATTCTCCCCAGACAAATTCGACCCCTCGTGTCAACCCAGATAGTGGTCAGTCATGAC 540
*****

S A V S D S E N F Q Y V K E A R Q R T E
DSS TCTGCTGTTCCGATTCGAGAATCCCAATACGTTAAAGAGGCAAGACAAAGCACAGS 600
LEWIS TCTGCTGTTCCGATTCGAGAATCCCAATACGTTAAAGAGGCAAGACAAAGCACAGS 600
*****

F F N S D N M Q A S S F N S N I E K Q D
DSS CCTTTCAATTCAGATAAATGCAAGCAAGTTCCTTTAATTCAAATATTGAGAAGCAAGAT 660
LEWIS CCTTTCAATTCAGATAAATGCAAGCAAGTTCCTTTAATTCAAATATTGAGAAGCAAGAT 660
*****

V S Q L R T A H A T V S G L T S D G L G
DSS GTATCCAGCTCAGGACGGCCCATGCCACTGTGTCAAGGTTAACTGGTGATGGCCTAGSA 720
LEWIS GTATCCAGCTCAGGACGGCCCATGCCACTGTGTCAAGGTTAACTGGTGATGGCCTAGSA 720
*****

H E E F N E S V T P S R Q T P K A Q K Y
DSS CATGAAGAACCTAATGAGAGTGTCAACCCCTAGTCACCAACCCCAAGCACAGAAATAT 780
LEWIS CATGAAGAACCTAATGAGAGTGTCAACCCCTAGTCACCAACCCCAAGCACAGAAATAT 780
*****

I S F S L P L P E A T L Y P T P S D S N
DSS ATCAGCTTCAGTCTTCCACTGCCGAGGCAACCCCTGTATCCTTACCCACGACAGCAAC 840
LEWIS ATCAGCTTCAGTCTTCCACTGCCGAGGCAACCCCTGTATCCTTACCCACGACAGCAAC 840
*****

```



```

S I N M Q P B P Q V F S E E S D S D Y E
DSS TCCATCAACATGCASC CGGACCCAGSTTCCAGCGAAGACTCTGCAGTGACTACGAA 900
LEWIS TCCATCAACATGCASC CGGACCCAGSTTCCAGCGAAGACTCTGCAGTGACTACGAA 900
*****
L C P E I T L T Y T E E F S D D D I E Y
DSS CTTTGCCAGAGATAACCCCTAACGTACACGGAAGSTTTTCGGATGATGACTGGAGTAT 960
LEWIS CTTTGCCAGAGATAACCCCTAACGTACACGGAAGSTTTTCGGATGATGACTGGAGTAT 960
*****
L E C S D V M T O Y S N A V W Q R N L Q
DSS CTGGAAATGTTCCGATGTTATGACGSAATTACTCTAATGOSSTTTGGCAAAGGAACCTGCAG 1020
LEWIS CTGGAAATGTTCCGATGTTATGACGSAATTACTCTAATGOSSTTTGGCAAAGGAACCTGCAG 1020
*****
G T D R V F L L E S D D E M E F N E C
DSS GGGACTGACCAGSTTTTTTATTAGAAGCGATGACGAAGAGATGGAATTCATGAGTGT 1080
LEWIS GGGACTGACCAGSTTTTTTATTAGAAGCGATGACGAAGAGATGGAATTCATGAGTGT 1080
*****
G P G G C E H F F S E M G C O P Q V S D
DSS GGTCCCGGTGGGTGTGAGCATTTCCTCAGTGAAGTGGTGTGGGCTCAGGTGTCGGGT 1140
LEWIS GGTCCCGGTGGGTGTGAGCATTTCCTCAGTGAAGTGGTGTGGGCTCAGGTGTCGGGT 1140
*****
G M W E M N V A T O F C S Y R S O P Q E
DSS GGCATGTGGTCTATGAATGTCCGCACTGGCTTCTGTAGCTATGCTCACAAACCCAGGAA 1200
LEWIS GGCATGTGGTCTATGAATGTCCGCACTGGCTTCTGTAGCTATGCTCACAAACCCAGGAA 1200
*****
V G D R S E O T E R H S P L P L S E M
DSS GTGGGGACAGGAGCAGGSAACCTCCAGSCATAGCCCAITACCCCTCATTAGAAATG 1260
LEWIS GTGGGGACAGGAGCAGGSAACCTCCAGSCATAGCCCAITACCCCTCATTAGAAATG 1260
*****
T L Y L G P R Q D G I A T M T E F G R A
DSS ACTCTACCCCTGGGACTCACAGGACGSAATAGCTACGATGACAGAACCGGGAGGGCT 1320
LEWIS ACTCTACCCCTGGGACTCACAGGACGSAATAGCTACGATGACAGAACCGGGAGGGCT 1320
*****
F L P T A S E A V E N D C S G I R G E T
DSS CCACTCCCCTGCTCCGAGGCTGTGAAATGATTTGTCAGGAAATCGGGAGAAACC 1380
LEWIS CCACTCCCCTGCTCCGAGGCTGTGAAATGATTTGTCAGGAAATCGGGAGAAACC 1380
*****
R D K P E A B E E F S E N L Q T H D E
DSS AGAGACAAGCCTGAAGCAGGAGGSAATTCCTCAGTGACAACTACAGACCATGGATAAG 1440
LEWIS AGAGACAAGCCTGAAGCAGGAGGSAATTCCTCAGTGACAACTACAGACCATGGATAAG 1440
*****
A E T E A S A K F L S G S D K D E A K
DSS GCAGAGACAGGACATCGGGAAGCCCTTCTCGGGGCTCAGATAAGTCAGAGCCGAAG 1500
LEWIS GCAGAGACAGGACATCGGGAAGCCCTTCTCGGGGCTCAGATAAGTCAGAGCCGAAG 1500
*****
Q D L E S L A G E R T E E K Y P G S R E
DSS CAGGTTTTGAAAAGCCTGGCTGGGSAAGCACGGAAGAAAAATACCCAGGCAGCAGGAAG 1560
LEWIS CAGGTTTTGAAAAGCCTGGCTGGGSAAGCACGGAAGAAAAATACCCAGGCAGCAGGAAG 1560
*****
T A L R P T R A R N P G H K A N V E R Q
DSS ACAGCCCTGAGACCCACAGGGCAGGTGGCCGSAATGAAGCAATGTCAAGAAAGCAG 1620
LEWIS ACAGCCCTGAGACCCACAGGGCAGGTGGCCGSAATGAAGCAATGTCAAGAAAGCAG 1620
*****
L L E D G P P K G T L D P L P K E P T E
DSS CTGCTGAAGACGGTCCCCAAAGGCACTCTTGATCCCTTCCCAAGGAACCAACAAAG 1680
LEWIS CTGCTGAAGACGGTCCCCAAAGGCACTCTTGATCCCTTCCCAAGGAACCAACAAAG 1680
*****
Q P L T R S Y G Q E P A H T E T G A P G
DSS CAGCCCTAACCGASGCTATGGACAAGAACCCTGCACACACAGAGACGGGAGCTCCTGGC 1740
LEWIS CAGCCCTAACCGASGCTATGGACAAGAACCCTGCACACACAGAGACGGGAGCTCCTGGC 1740
*****
W M S R F R E V C I F L P A E Q D S K T
DSS TGGAAATCCCACTTCCGAGAAGTGTGATTCCTCTTCTGCAGAGCAAGACTCCAAAACC 1800
LEWIS TGGAAATCCCACTTCCGAGAAGTGTGATTCCTCTTCTGCAGAGCAAGACTCCAAAACC 1800
*****
F R P E A D P L S K E G D S S F E G S G
DSS CCTGSAACCCAGCAGACCCACTTCAAGGAAGGGATACAGCTTCGAGGGAGGAGSA 1860
LEWIS CCTGSAACCCAGCAGACCCACTTCAAGGAAGGGATACAGCTTCGAGGGAGGAGSA 1860
*****

```

```

A L L S T L S E A S Q I P Q T - G R L
DSS GCAC T G C T T A G T A C G T T A T C T G A A G C A A G C C A G A T T C C A G A C C A G A C C G --- O C C G T C T T 1917
LEWIS GCAC T G C T T A G T A C G T T A T C T G A A G C A A G C C A G A T T C C A G A C C A G A C C G C C G C G T C T T 1920
*****

A

Q E T I G E N S Y L D Q T F A F S V F A
DSS C A G G A A C C A T T G G G G A G A C A B T T A T C T G A C C A G A G G C C G G C C T T T T C C G T G C C C G C T 1977
LEWIS C A G G A A C C A T T G G G G A G A C A B T T A T C T G A C C A G A G A G C C G G C C T T T T C C G T G C C C G C T 1980
*****

E E E S T F T G V Q T R F V S N L T E I
DSS G A G G A G A G T C T A C G T T C A C G G G T C C A A A C A C A T T T C G T C T C C A A C T T G A C T G A G A T C 2037
LEWIS G A G G A G A G T C T A C G T T C A C G G G T C C A A A C A C A T T T C G T C T C C A A C T T G A C T G A G A T C 2040
*****

D R G N S S L A Q Y L D L E S C F Q S P
DSS G A C A G G G A A A C T C A T C A T T G G C C A G T A C C T G G G G C T G G A G A G C T G T C C C C A A A G C C C C 2097
LEWIS G A C A G G G A A A C T C A T C A T T G G C C A G T A C C T G G G G C T G G A G A G C T G T C C C C A A A G C C C C 2100
*****

Q Q E S R Q N R E G D R F G A L W A E S
DSS C A G C A G A A A G C A G A C A A A C A G A A G G T G A C A G G C C T G G C C T C T C T G G G C A G A A T C A 2157
LEWIS C A G C A G A A A G C A G A C A A A C A G A A G G T G A C A G G C C T G G C C T C T C T G G G C A G A A T C A 2160
*****

A Q E L R P L E D H D E E V S Q A P A S
DSS G C A C A A G A C T G A G A C C C C T A G A A G A C A A T G A T G A A G A A G T G T C C C A G G C T C C A G C C T C A 2217
LEWIS G C A C A A G A C T G A G A C C C C T A G A A G A C A A T G A T G A A G A A G T G T C C C A G G C T C C A G C C T C A 2220
*****

V A L F Q G D G T H F R E S E A L S D A
DSS G T G G C T C C C C T C A G G G T G A T G G T A C C C A C T T C A G G A G T C A G A G G C T C T C T G A T G C T 2277
LEWIS G T G G C T C C C C T C A G G G T G A T G G T A C C C A C T T C A G G A G T C A G A G G C T C T C T G A T G C T 2280
*****

F S Q P T A P S L P L E H V D S G R G
DSS T T C T C C C A G C C T A C T G C T C C C C C C C C C C T T G G A A A T G T G G A C A G T G G G T C A A G G G G C 2337
LEWIS T T C T C C C A G C C T A C T G C T C C C C C C C C C C T T G G A A A T G T G G A C A G T G G G T C A A G G G G C 2340
*****

R E A A C V M G C F E A G D Q E T C Y A
DSS A G A G A G C T S C G T G T G A T G G G G T T T T G A A G C T G G T G A T C A G A A A C A T G T T A T G C T 2397
LEWIS A G A G A G C T S C G T G T G A T G G G G T T T T G A A G C T G G T G A T C A G A A A C A T G T T A T G C T 2400
*****

T M D L F V G A P V D E Y L P E E I C P
DSS A C C A T G G A T C C C C T G T T G G A G C A C C A S T T G A T A A A T A T T T G C C T G A A G A A A T T T G C C C C 2457
LEWIS A C C A T G G A T C C C C T G T T G G A G C A C C A S T T G A T A A A T A T T T G C C T G A A G A A A T T T G C C C C 2460
*****

M D L E L T E G Q R E V C D L C E P D K
DSS A T G G A C T T G S A G C T G A C A G A A G G T C A A G A G A A G T A T G T A T T A T G T T C C T G A C A A G 2517
LEWIS A T G G A C T T G S A G C T G A C A G A A G G T C A A G A G A A G T A T G T A T T A T G T T C C T G A C A A G 2520
*****

I L A V L Q A Q G S E S P Q A T Y K R S
DSS A T A T T G G C T G T T C A C A A G C A C A A G G T T C T G A G T C T C C A C A G G C C A C G T A C A A G C A C A G C 2577
LEWIS A T A T T G G C T G T T C A C A A G C A C A A G G T T C T G A G T C T C C A C A G G C C A C G T A C A A G C A C A G C 2580
*****

E D G K S A E G A L F H S T L A M D T S
DSS G A G S A T G G G A A G T C A G C C G A G G G C C T C T T T T C A T A G T A C C T T A G C C T G G G A C A C A T C A 2637
LEWIS G A G S A T G G G A A G T C A G C C G A G G G C C T C T T T T C A T A G T A C C T T A G C C T G G G A C A C A T C A 2640
*****

F E A R E D A T G G T A A D T G H S F S
DSS C C G G A G G C C A G A G A G A T G C T A C G G A G G S A C A G C A G C T G A T A C G G G A A T T C C C C T C C 2697
LEWIS C C G G A G G C C A G A G A G A T G C T A C G G A G G S A C A G C A G C T G A T A C G G G A A T T C C C C T C C 2700
*****

I E S S T L L Y N V G G F S E I Q A P C
DSS A T C T T C T T C T A C A C T A C T C T A C A A T G T A G G A G G T T T G G G A A T A C A G G C C C T T G T 2757
LEWIS A T C T T C T T C T A C A C T A C T C T A C A A T G T A G G A G G T T T G G G A A T A C A G G C C C T T G T 2760
*****

S E N I S F V H S N E G G Y G S S N L S
DSS T C T G A A A A T A T C C C T T T G T A A A C A G T A A C G A A G A G G C T A T G G A A G C T C C A A T C T C A G C 2817
LEWIS T C T G A A A A T A T C C C T T T G T A A A C A G T A A C G A A G A G G C T A T G G A A G C T C C A A T C T C A G C 2820
*****

```

DSS V P I A I D T L A S Y G S I R E C S E E 2877
 LEWIS GTTCCGATTGCCATAGACAGCTTCCAGCTACGGTCCATCAGGGAGTCCTCAGAAGAG 2880

DSS Q S T E P A A H V D C H L V T R E T E G 2937
 LEWIS CAGTCAACAGAACAGCTGCTAATGTTGACTGTGATCGTGGTACCAGAGAGACAGAGGGC 2940
 CAGTCAACAGAACAGCTGCTAATGTTGACTGTGATCGTGGTACCAGAGAGACAGAGGGC

Q

DSS I L T D A T E V R K I K C R T V S V F H 2997
 LEWIS ATACTAACTGATGCCACCGAAGTCCACAAAATCAAATGCCACCGTTTCTGTGCCCCAC 3000
 ATACTAACTGATGCCACCGAAGTCCACAAAATCAAATGCCACCGTTTCTGTGCCCCAC

DSS I N D F V D S A D Q V S C E A Q D E E N 3057
 LEWIS ATCAACGACTTGTGATGGTGGTACCAGGTCTCGTGTGAGGCACAAGATGAAGAGAAT 3060
 ATCAACGACTTGTGATGGTGGTACCAGGTCTCGTGTGAGGCACAAGATGAAGAGAAT

DSS S F S L P D D F L S S F T G E A T R E T 3117
 LEWIS TCDDCATCGCTCCDDGATGACCGGTTAAGTAGTTCACAGGTGAAGCAACCAGGGAGACT 3120
 TCDDCATCGCTCCDDGATGACCGGTTAAGTAGTTCACAGGTGAAGCAACCAGGGAGACT

DSS L V F A P S D A G T H G H F L L P E S Q 3177
 LEWIS CTTGTCCAGGCCCCASTGATGCAAGAACTCATGGCCACTTCTTACTGCCCGAGGGACAG 3180
 CTTGTCCAGGCCCCASTGATGCAAGAACTCATGGCCACTTCTTACTGCCCGAGGGACAG

DSS G L Y L R F L Q I D T Q P G Y E S Q T V 3237
 LEWIS GGTTTGTACTCAGGCCCTTTCAGATTGATACCCAGCCTGGGTATGAGGCCAGACTGTG 3240
 GGTTTGTACTCAGGCCCTTTCAGATTGATACCCAGCCTGGGTATGAGGCCAGACTGTG

DSS E G A R S G G L E E D F Q E K G M G T E 3297
 LEWIS GAGGGAGCCACAGCGGAGSCCTTGAAGAGGACTTCCAAGAAAAGGAAATGGGACAAA 3300
 GAGGGAGCCACAGCGGAGSCCTTGAAGAGGACTTCCAAGAAAAGGAAATGGGACAAA

DSS Q C I R P Q S T S H Q G S L S A N D F Q 3357
 LEWIS CAATSCATCCGGCCACAGACACATCCCATCAGGTTCTCTTTCTGCAAAATGATTTCCAA 3360
 CAATSCATCCGGCCACAGACACATCCCATCAGGTTCTCTTTCTGCAAAATGATTTCCAA

DSS E E L F S I P A M Q Q E V N V E P F E H 3417
 LEWIS GAAAGTTTGCCTCCATACCTGCCATGCAACAGGAGTCAACGTGGAAACCTTCGAGCAC 3420
 GAAAGTTTGCCTCCATACCTGCCATGCAACAGGAGTCAACGTGGAAACCTTCGAGCAC

DSS S F A D S G E E T E C S S D Q R T S V S 3477
 LEWIS TCCCCAGCAGATCCGGGGAAGAACTGAGTGTAGCTCAGACCAGAGGACCAATGTTTCT 3480
 TCCCCAGCAGATCCGGGGAAGAACTGAGTGTAGCTCAGACCAGAGGACCAATGTTTCT

DSS V L A E K T M G E G S P L V S S V P A L 3537
 LEWIS GTGTTGGCTGAGAGACCATGGGAGAAGGCASTCCTTTGGTAAGCAGTGTCCAGCTCTC 3540
 GTGTTGGCTGAGAGACCATGGGAGAAGGCASTCCTTTGGTAAGCAGTGTCCAGCTCTC

DSS F D I L L G E E D G I G L G S W S V S S 3597
 LEWIS CCTGACATCCTCTTGGAGAGAAGATGGCATTGGACTAGGAAGTGGTCTGTGGCCAGC 3600
 CCTGACATCCTCTTGGAGAGAAGATGGCATTGGACTAGGAAGTGGTCTGTGGCCAGC

DSS K V K I I T L E A P V F E I M P F E L V 3657
 LEWIS AAAGTGAAGATCATAACTCTAGAAGCTCCTGTTTTGAAATCTGGCCCCAGAACTAGTG 3660
 AAAGTGAAGATCATAACTCTAGAAGCTCCTGTTTTGAAATCTGGCCCCAGAACTAGTG

DSS T H S G Y K E A E V G L T A P G R S W A 3717
 LEWIS ACBCATTCTGGGTACAAGGAGGAGAAAGTGGTCTCAGGCCACTGGAAGGAGCTGGGCT 3720
 ACBCATTCTGGGTACAAGGAGGAGAAAGTGGTCTCAGGCCACTGGAAGGAGCTGGGCT

DSS L E D I L R A G T R F E P G A L E V T T 3777
 LEWIS CTGTCTGACATCCTCAGAGCAGGCACCGGCCAGCCAGGAGGCTTGGAGTAACAACA 3780
 CTGTCTGACATCCTCAGAGCAGGCACCGGCCAGCCAGGAGGCTTGGAGTAACAACA

K


```

W V F S P K A D A I M A I G N N R D I C
DSS 756GTTCCACGCCCAAGCCGATGCCATTATGSCCTTGGAAATAACAGGGACATCTGC 3837
LEWIS 756GTTCCACGCCCAAGCCGATGCCATTATGSCCTTGGAAATAACAGGGACATCTGC 3840
*****
E D A A P D R Q A Y C N E Q T S Q C L G
DSS GAAGATGCTGCACAGACAGACAAAGCTACTGCAACASTCAGACTTCCAGTGTCTGGGC 3897
LEWIS GAAGATGCTGCACAGACAGACAAAGCTACTGCAACASTCAGACTTCCAGTGTCTGGGC 3900
*****
Q P R L L E S S V D F V E E K E L N V T
DSS CAACCCCGACTCCTGGAGTCATCTGTTGACCCGTGTGAGSAGAAGGATTAATGTCAAG 3957
LEWIS CAACCCCGACTCCTGGAGTCATCTGTTGACCCGTGTGAGSAGAAGGATTAATGTCAAG 3960
*****
D S P E E T S R T G E V E M A E T T D E
DSS GACTCTCCATCAGAGACTTCCAGAACTGGAGAGGTGGAAATGGCTGAGACAAGGATGAG 4017
LEWIS GACTCTCCATCAGAGACTTCCAGAACTGGAGAGGTGGAAATGGCTGAGACAAGGATGAG 4020
*****
E Q D G R Q R K L P H P T V A N Q S V N
DSS GAACAGGGGGAGGSCAGCACAAGCTGCCSCACCCACCCGTGCTAACAGTCTGTGAAC 4077
LEWIS GAACAGGGGGAGGSCAGCACAAGCTGCCSCACCCACCCGTGCTAACAGTCTGTGAAC 4080
*****
F P R I L E S S V D F I D D R G D K F E
DSS TTCCCTAGGATCCTGGAATCCTCTGTGGACCCCATTGATGACAGGGGTGGGAAGCCAGAG 4137
LEWIS TTCCCTAGGATCCTGGAATCCTCTGTGGACCCCATTGATGACAGGGGTGGGAAGCCAGAG 4140
*****
F S D E N I E A S E E T T G N M C Q R V
DSS CCTCTGACTCCAACATAGAAGCAAGTGAATCCACCACTGGAAACATGTGTGAGGGGTA 4197
LEWIS CCTCTGACTCCAACATAGAAGCAAGTGAATCCACCACTGGAAACATGTGTGAGGGGTA 4200
*****
D I Q T A H L R V P H P Q D N G E I I P
DSS GACATCCAACTGCTCACCTACGGGTCCACATCCCAGGACAATGGGSAATCATTCCA 4257
LEWIS GACATCCAACTGCTCACCTACGGGTCCACATCCCAGGACAATGGGSAATCATTCCA 4260
*****
N E H T T N Q T H V D R E R A D A K A S
DSS AACGAAAACACAACCAAACTCATGTAGACAGAGAGGAGCAGATGCCAAAGCCAGT 4317
LEWIS AACGAAAACACAACCAAACTCATGTAGACAGAGAGGAGCAGATGCCAAAGCCAGT 4320
*****
Q R H A A K D A I M Q G Q C P G E E R Q
DSS CAGCATAATGCAGCAAAGAGCGCTATTGGCAAGGCCAGTGCCTCCGGTGAAGAGAGACAA 4377
LEWIS CAGCATAATGCAGCAAAGAGCGCTATTGGCAAGGCCAGTGCCTCCGGTGAAGAGAGACAA 4380
*****
S I P S V C T V S P T Q D G G D R S L S
DSS GGGATCCAAGTGTGTGACCCGTGAGCCCAACAAAGATGGTGTGACAGAAGCCTAGSA 4437
LEWIS GGGATCCAAGTGTGTGACCCGTGAGCCCAACAAAGATGGTGTGACAGAAGCCTAGSA 4440
*****
E A G Q R G K D E T E V T S F M S P L S
DSS GAAGCTGGGCAAGGGGAAAGGACGAGACTGAGGTCACTTCCCCCATGTCTCCCTTTCT 4497
LEWIS GAAGCTGGGCAAGGGGAAAGGACGAGACTGAGGTCACTTCCCCCATGTCTCCCTTTCT 4500
*****
N C P A G M T Y T S V T A E T S N S T G
DSS AACTGTCTSCAGGAATGACGTACACATCTGTGACGGCTGAGACCAGTAACCTCCACAGGC 4557
LEWIS AACTGTCTSCAGGAATGACGTACACATCTGTGACGGCTGAGACCAGTAACCTCCACAGGC 4560
*****
H I Y G G S E P R T H Q R V I P V K E
DSS CACATTTATGGCGGATCTGAGCCCAAGAACCCATCAACGTGTAAATCCTGTGAAGAGGGAA 4617
LEWIS CACATTTATGGCGGATCTGAGCCCAAGAACCCATCAACGTGTAAATCCTGTGAAGAGGGAA 4620
*****
K G T I E N E C G K H V P S S N D L T D
DSS AAAGSAACCATCGAGAACAAAGTGGGAAACATGTGCCCTCTCAAATGATCTCACGGAC 4677
LEWIS AAAGSAACCATCGAGAACAAAGTGGGAAACATGTGCCCTCTCAAATGATCTCACGGAC 4680
*****
T L C T S S P K G N V T R S P T S P R A
DSS ACACCTGACACTTCATCTCCCAAGGAAATGTCAACGCTCGGCCAAGGAGCCCTCGGCA 4737
LEWIS ACACCTGACACTTCATCTCCCAAGGAAATGTCAACGCTCGGCCAAGGAGCCCTCGGCA 4740
*****
E E L K S E E L Q I A E T K P L N S S D
DSS GAGGAACGAAATCAGAGGACTTCAAATTCGGAACCAACCCCTAAACTCATCTGAC 4797
LEWIS GAGGAACGAAATCAGAGGACTTCAAATTCGGAACCAACCCCTAAACTCATCTGAC 4800
*****

```

```

FRYMTLAFISGEHESEEDFE
DSS CCCCACAAATGACCTTGGCTTTCATTTACGGAGAACATGAGTCAGAGAAAGACCTGAA 4857
LEWIS CCCCACAAATGACCTTGGCTTTCATTTACGGAGAACATGAGTCAGAGAAAGACCTGAA 4860
*****
SLLLEKDLCKQKGSSTLESSEKKS
DSS AGCTTGTACTTAAGGACCTGTGTCAAAAAGGGCTTACCCCTGGAGAGGGGAAAAAGTCC 4917
LEWIS AGCTTGTACTTAAGGACCTGTGTCAAAAAGGGCTTACCCCTGGAGAGGGGAAAAAGTCC 4920
*****
REEQQRFPVVANISEKAPDAQS
DSS AGAGAGGAACAGCAGAGGCTGTGGTGGCCAAACATCAGCAAGGCACCCGGGGCCCAATCA 4977
LEWIS AGAGAGGAACAGCAGAGGCTGTGGTGGCCAAACATCAGCAAGGCACCCGGGGCCCAATCA 4980
*****
AIAQSEESGKKEAESESSHLA
DSS GCRAATAGCTGGGTGAGAGGAGGGCAAAAAACAAGAGGCTTCAGGGAGTGGACACTTGGCT 5037
LEWIS GCRAATAGCTGGGTGAGAGGAGGGCAAAAAACAAGAGGCTTCAGGGAGTGGACACTTGGCT 5040
*****
AGIKKKKILERVAAALRLLEEE
DSS GCAGGGATAAAGAAAGAAAATCTATCCAGGGTGGCAGCCCTGAGGCTGAGGCTAGAGGAA 5097
LEWIS GCAGGGATAAAGAAAGAAAATCTATCCAGGGTGGCAGCCCTGAGGCTGAGGCTAGAGGAA 5100
*****
KERVENSTLRKAPKFERBLS
DSS AAGGAACGTGTGAAGAACTCCACTCTGAGGAAGGCACCTAAGTTGAACGGTCTTGTGCC 5157
LEWIS AAGGAACGTGTGAAGAACTCCACTCTGAGGAAGGCACCTAAGTTGAACGGTCTTGTGCC 5160
*****
RTDEEKRDPRRAPCKAESKAP
DSS CGCACTGATGAGAAAAGAGACCCCGAAGGGCCCTTGCAAAGCTGAAGGGAAAAGCTCCA 5217
LEWIS CGCACTGATGAGAAAAGAGACCCCGAAGGGCCCTTGCAAAGCTGAAGGGAAAAGCTCCA 5220
*****
VLLKKKIQAEPAPERESGNIML
DSS GTATTGCTGAAGAAAGATCCAGGCTGAGCCGGCTCCCGAGCAGCTTGGAAATATAATGCTG 5277
LEWIS GTATTGCTGAAGAAAGATCCAGGCTGAGCCGGCTCCCGAGCAGCTTGGAAATATAATGCTG 5280
*****
TCQFSEIEHEDSTVCMTEKDEK
DSS ACCTGTCACTTTTCAGAAATCCATGAAAGACTCCACCCTTGTGGACAAAAGATTCCGAA 5337
LEWIS ACCTGTCACTTTTCAGAAATCCATGAAAGACTCCACCCTTGTGGACAAAAGATTCCGAA 5340
*****
SIAQLKKKSAQDSSESVSLAIA
DSS TCGATAGCCGACTCAAGAAAAGCCAGGGGACAGCTCCAGTGTTCCTTGGCCATCGCC 5397
LEWIS TCGATAGCCGACTCAAGAAAAGCCAGGGGACAGCTCCAGTGTTCCTTGGCCATCGCC 5400
*****
QAGQKDKDQLLYYCCVKNSYGW
DSS CAAGCTGGTCAGAAAGACAGGGCTCTATTACTGCTGCGTCAAGAACAGCTACGGAAAA 5457
LEWIS CAAGCTGGTCAGAAAGACAGGGCTCTATTACTGCTGCGTCAAGAACAGCTACGGAAAA 5460
*****
VTAEFNLTAEVLKQLSDHTE
DSS GTCACCTGCTGAGTTTAACTTCACAGCCGAAGTTCACAAACAGCTTCCAGTCACACAGAA 5517
LEWIS GTCACCTGCTGAGTTTAACTTCACAGCCGAAGTTCACAAACAGCTTCCAGTCACACAGAA 5520
*****
YRCCEIEEFSQLIFKEDVFN
DSS TACAGAGGATGTGAAGAGATTGAATTCAGCCAGCTCATCTTCAAAGAAGATGTTTTCAAT 5577
LEWIS TACAGAGGATGTGAAGAGATTGAATTCAGCCAGCTCATCTTCAAAGAAGATGTTTTCAAT 5580
*****
DSYEGDHLRGQIEFTEELHFG
DSS GACAGTACTTCGGGGACCACTACCGGGCAGATCTTACGGAGGAGCTTCACTTCGGC 5637
LEWIS GACAGTACTTCGGGGACCACTACCGGGCAGATCTTACGGAGGAGCTTCACTTCGGC 5640
*****
EGVRRKAFRSTVKNQQLMFPVF
DSS GAAGGGGTGCACCGCAAAGCTTTCGGCAGCACAGTGTGCAAGGGCTCATGCCCCGTCTTC 5697
LEWIS GAAGGGGTGCACCGCAAAGCTTTCGGCAGCACAGTGTGCAAGGGCTCATGCCCCGTCTTC 5700
*****
QPEHACVLEKVNNAVAVHSTRN
DSS CAGCCCGCCACGCATGCTACTCAAGGTGCACAACGCTGTGCCCCATGGACCCAGAAAC 5757
LEWIS CAGCCCGCCACGCATGCTACTCAAGGTGCACAACGCTGTGCCCCATGGACCCAGAAAC 5760
*****

```

```

                                ↓
DSS      N D E L V Q R N Y K L A A Q E C Y V Q N
LEWIS   AATGATGAACCTGTSCAGAGGAAGTACAAGCTGGCTGCCAGGAATGCTACGTTCAGAAT 5817
        AATGATGAACCTGTSCAGAGGAAGTACAAGCTGGCTGCCAGGAATGCTACGTTCAGAAT 5820
        *****
        T A R Y Y A E I Y A A E A Q P L E G F E
DSS      ACTGCCAGATACTATGCCAAGATCTATGCCGCTGAAGCAGACGCTCTGGAAGGCTTCGGA 5877
LEWIS   ACTGCCAGATACTATGCCAAGATCTATGCCGCTGAAGCAGACGCTCTGGAAGGCTTCGGA 5880
        *****

                                ↓
DSS      E V P E I I P I F L I R R F E N N I F Y
LEWIS   GAGGTGCCGAGATCATTCTATTTCCTTATCCGTCCGCCGAGAACAAACATCCCTTAC 5937
        GAGGTGCCGAGATCATTCTATTTCCTTATCCGTCCGCCGAGAACAAACATCCCTTAC 5940
        *****
        A T V E E E L I G E F V R Y S I R D G K
DSS      GCCACAGTGGAGAGGAGCTGATTGGAGAATTCTGTGAAATATCCATCCGAGACGGGAAG 5997
LEWIS   GCCACAGTGGAGAGGAGCTGATTGGAGAATTCTGTGAAATATCCATCCGAGACGGGAAG 6000
        *****
        E I N F L R R D S E A D Q K C C T F Q H
DSS      GAAATCAACTTCTCAGACGAGATTCGGAGGCTGGTTCAGAAAGTGTTCACCTTCCAGCAC 6057
LEWIS   GAAATCAACTTCTCAGACGAGATTCGGAGGCTGGTTCAGAAAGTGTTCACCTTCCAGCAC 6060
        *****
        W V Y Q K T S G C L L V T D M Q G E Q D
DSS      TGGGTATACCAGAAAACAAGTGGCTGTCTCCTGGTCAGGACATGCAGGGTGGAGCAGGAC 6117
LEWIS   TGGGTATACCAGAAAACAAGTGGCTGTCTCCTGGTCAGGACATGCAGGGTGGAGCAGGAC 6120
        *****
        R P T A L Q G L W S R G V P K V S F L R
DSS      AGACCGACAGCTCTGCAGGGGCTGTGGAGCCGGGGTGTCTCTAAGGTTTCTTCTTGAGS 6177
LEWIS   AGACCGACAGCTCTGCAGGGGCTGTGGAGCCGGGGTGTCTCTAAGGTTTCTTCTTGAGS 6180
        *****

        K E N R G S E E E E R S V
DSS      AAGGAATGAGAGGAAGTGAAGAGSAGSAGSAGTCTGTTGGGTACCAACAGCATCTCCT 6237
LEWIS   AAGGAATGAGAGGAAGTGAAGAGSAGSAGSAGTCTGTTGGGTACCAACAGCATCTCCT 6240
        *****

                                I
DSS      F Q D V T H V P L R T M L G K T D V R K
LEWIS   CCCCAGAGSTTACTCAGTCCCATTCGGGACGATGCTTGGAAAGACTGACGTTAGGAAA 6297
        CCCCAGAGSTTACTCAGTCCCATTCGGGACGATGCTTGGAAAGACTGACGTTAGGAAA 6300
        *****

DSS      -
LEWIS   TGA 6300
        TGA 6303
        ***

```

Footnote to Supplement 6: * indicates nucleotide identity. The amino acid sequence is given on top. The cDNA length of *Alpk2* is based on those of the mouse (http://www.ncbi.nlm.nih.gov/nuccore/NM_001037294.1) and humans (http://www.ncbi.nlm.nih.gov/nuccore/NM_052947.3). Genomic DNAs and/or cDNA from the heart from DSS and Lewis were first amplified by PCR and then sequenced. When a nucleotide difference was seen (indicated by shade), it was re-PCRed and sequence-confirmed. The deleted nucleotides are indicated by ---. The 5 amino acid changes caused by 5 individual mutations are indicated by bold and large lettering. ↓ marks the last nucleotide of each exon.

Supplement 7. Coding sequence alignment of *Gastrin releasing peptide gene (Grp)* between Dahl salt-sensitive (DSS) and Lewis rats.

```

DSS      N R G S E L S L L L L L A L V L C Q A P R 20
Lewis   ATGCGGGGCTCCGAGCTCTCGCTCTTGGCTGTGGCTCTGGTCCCTCGCCAGGCTCCCCGG 60
Lewis   ATGGCGGGCTCCGAGCTCTCGCTCTTGGCTGTGGCTCTGGTCCCTCGCCAGGCTCCCCGG
*****

DSS      G F A A F V E T G A G S G T V L A K M W 40
Lewis   GGGCAGCTGCCCCGTTGTGACAGGCGCAGGTGAGSCACAGTCTGGCCAAGATGTAT 120
Lewis   GGGCAGCTGCCCCGTTGTGACAGGCGCAGGTGAGSCACAGTCTGGCCAAGATGTAT
*****

DSS      P R G S H W A V G H L M G K K S T D E L 60
Lewis   CCGCCGSCAGCCACTGGGCTGTAGGACACTTAATGGGSAAGANGACACAGATGAGTTE 180
Lewis   CCGCCGSCAGCCACTGGGCTGTAGGACACTTAATGGGSAAGANGACACAGATGAGTTE
*****

DSS      P P L Y A A D R D G L K E Q L R S Y I R 80
Lewis   CCGCCCTGTATGCAAGCTGACAGAGATGGCCTGAAAGGACAACTGAGGGGATACATCCGC 240
Lewis   CCGCCCTGTATGCAAGCTGACAGAGATGGCCTGAAAGGACAACTGAGGGGATACATCCGC
*****

DSS      W E E A A R N L L G L L E A A G H R S H 100
Lewis   TGGGAAGAAGCTGCAAGSAATTTGCTTGGCCTCCTCGAAGCCGCGGGAAACAGAAGCCAC 300
Lewis   TGGGAAGAAGCTGCAAGSAATTTGCTTGGCCTCCTCGAAGCCGCGGGAAACAGAAGCCAC
*****
                S R

DSS      Q F P Q D Q F L G S L Q P T W D F E D G 120
Lewis   CAGCCACTCAGSATCAGCCTCTGGGAGTCTCCAGCCTACTTGGGATCCAGAGSATGGC 360
Lewis   CAGCCACTCAGSATCAGCCTCTGGGAGTCTCCAGCCTACTTGGGATCCAGAGSATGGC
*****

DSS      S Y F S D A Q N A K L V D S L L O V L K 140
Lewis   AGCTACTTCAGTGTGCTCAAATGCTAAATGGTAGACTCTCTGCTCCAGGTTCTCAAG 420
Lewis   AGCTACTTCAGTGTGCTCAAATGCTAAATGGTAGACTCTCTGCTCCAGGTTCTCAAG
*****

DSS      G K E G T A S - 147
Lewis   GGGAAAGGAAGGAACTGCCAGCTGA 444
Lewis   GGGAAAGGAAGGAACTGCCAGCTGA
*****

```

Footnote to Supplement 7: * indicates nucleotide identity. The amino acid sequence is given on top. A change in amino acid is indicated by bold and larger lettering. The amino acid region underlined represents the mature peptide GRP. Genomic DNAs and brain cDNAs from DSS and Lewis were first amplified by PCR and then sequenced. When a nucleotide difference was seen (indicated by shade), it was re-PCR'd and sequence confirmed. ↓ indicates the last nucleotide of each exon. The exon-intron junctions are not different (data not shown).

Supplement 9. Coding sequence alignment of the RNA binding motif protein 22 gene (*Rbm22*) between Dahl salt-sensitive (DSS) and Lewis rats

```

M A T S E L G S N T Y N R Q N W E D A D F
DSS ATGGCGACTTCTCTGGGTTCCAACACTTACAACAGSCAGAACTGGGAGSACGGGACTTC 60
Lewis ATGGCGACTTCTCTGGGTTCCAACACTTACAACAGSCAGAACTGGGAGSACGGGACTTC 60
*****

F I L C Q T C L G E N F Y I R M T K E K
DSS CCCATTCTGTGTCAAGACTGTCTTGGAGAAAACCCCTATATAAGAAATGACCAAAAGAAAA 120
Lewis CCCATTCTGTGTCAAGACTGTCTTGGAGAAAACCCCTATATAAGAAATGACCAAAAGAAAA 120
*****

Y G K E C K I C A R F F T V F R W C F G
DSS TATGSAAGAAGATGCAAAATCTGTGCCAGGCCATTCCAGTATTTCGATGGTGCCTGGG 180
Lewis TATGSAAGAAGATGCAAAATCTGTGCCAGGCCATTCCAGTATTTCGATGGTGCCTGGG 180
*****

V R W R F K K T E V C Q T C S K L K N V
DSS GTCCGCATGGGCTTCAAGAAGACTGAAGTGTGCCAGACCTGCAGTAAATGAAAGAAATGTT 240
Lewis GTCCGCATGGGCTTCAAGAAGACTGAAGTGTGCCAGACCTGCAGTAAATGAAAGAAATGTT 240
*****

C Q T C L L D L E Y G L P I Q V R D A G
DSS TGTCCAGACTGCCCTGTAGACTAGAGTACGGCTACCCATCCAGGTCCGTGATGCAGGG 300
Lewis TGTCCAGACTGCCCTGTAGACTAGAGTACGGCTACCCATCCAGGTCCGTGATGCAGGG 300
*****

L S F K D D M P K S D V N K E Y Y T Q N
DSS TTATCATTAAAGATGACATGCCATAATCAGATGTCAATAAAGAAATATTACACCCAAAAT 360
Lewis TTATCATTAAAGATGACATGCCATAATCAGATGTCAATAAAGAAATATTACACCCAAAAT 360
*****

M E R E I S N S D G T R P V G M L G K A
DSS ATGGAAGAGAAATTTCTAACTCTGATGGAACACGGCCAGTCCGAAATGCTGGGAAAAGCA 420
Lewis ATGGAAGAGAAATTTCTAACTCTGATGGAACACGGCCAGTCCGAAATGCTGGGAAAAGCA 420
*****

T S T E D M L L K L A R T T P Y Y K R N
DSS ACATCCACTAGTGACATGCTACTCAAATGGCCCCGACACACCTTACTACAAAAGGAAT 480
Lewis ACATCCACTAGTGACATGCTACTCAAATGGCCCCGACACACCTTACTACAAAAGGAAT 480
*****

R P H I C S E N V E G E C K R G E E C F
DSS CGGCCCCACATTTGTCCTTCTGGGTGAAAGGAGAGTGTAAAGSGGGAGAGGATGTGCC 540
Lewis CGGCCCCACATTTGTCCTTCTGGGTGAAAGGAGAGTGTAAAGSGGGAGAGGATGTGCC 540
*****

Y R H E K P T D P D D E I A D Q H I K D
DSS TACAGACATGAGAAGCCTACCGATCCAGATGACCCCTCTTCTGATCAGAATATAAAGAC 600
Lewis TACAGACATGAGAAGCCTACCGATCCAGATGACCCCTCTTCTGATCAGAATATAAAGAC 600
*****

R Y Y G I N D F V A D K L L K R A S T M
DSS CGGTACTATGGAATTAACGACCCCTGTGGTGTATAAGCTTCTCAAGCGGGCTTCAACCATG 660
Lewis CGGTACTATGGAATTAACGACCCCTGTGGTGTATAAGCTTCTCAAGCGGGCTTCAACCATG 660
*****

P R L D P P E D E T I T T L Y V G G L G
DSS CCTGCTTGAACCCACCGGAGGATAAGACGATCACACACTGTATGTGGTGGTCTGGGA 720
Lewis CCTGCTTGAACCCACCGGAGGATAAGACGATCACACACTGTATGTGGTGGTCTGGGA 720
*****

D T I T E T D L E N H E Y Q F G E I E T
DSS GACACCATCACTGAAACAGACCTCAGGAATCATTCTACCACTTCGGAGAGATTCCGACA 780
Lewis GACACCATCACTGAAACAGACCTCAGGAATCATTCTACCACTTCGGAGAGATTCCGACA 780
*****

V T V V Q R Q Q C A F I Q F A T R Q A A
DSS GTCACTGTTGTGASAGCAGCAGTGTGCTTCAATCCAGTTCGCCACCGGCAAGCTGCC 840
Lewis GTCACTGTTGTGASAGCAGCAGTGTGCTTCAATCCAGTTCGCCACCGGCAAGCTGCC 840
*****

```

```

DSS      E V A A E K E F N K L I V N G R R L N V
Lewis    GAGTGGCTGCCGAGAAGTCTTTCAACAAGCTGATTGTGAACGGCCGACGACTCAATGTC 900
          GAGTGGCTGCCGAGAAGTCTTTCAACAAGCTGATTGTGAACGGCCGACGACTCAATGTC 900
          *
          ↓
DSS      K W G R S Q A A R G K E K E K D D T T D
Lewis    AAATGGGGACGATCCCAAGCAGCCAGAGGAAAAGAAAAGAGAAGGATGGAACACAGAC 960
          AAATGGGGACGATCCCAAGCAGCCAGAGGAAAAGAAAAGAGAAGGATGGAACACAGAC 960
          *
          ↓
DSS      S G I K L E P V F D L P G A L P P P F A
Lewis    TCTGGGATCAAGTTAGAACCCTGTACCAGGACTGCCTGGAGCTCTGCCTCCACCTCCTGCC 1020
          TCTGGGATCAAGTTAGAACCCTGTACCAGGACTGCCTGGAGCTCTGCCTCCACCTCCTGCC 1020
          *
DSS      A E E E A S A H Y F N L P F S G P P A V
Lewis    GCAGAAGAGGAAGGCTGTGCCAACTACTTCAAACCTGCCCCCAAGTGGGCTCCAGCCGTA 1080
          GCAGAAGAGGAAGGCTGTGCCAACTACTTCAAACCTGCCCCCAAGTGGGCTCCAGCCGTA 1080
          *
          ↓
DSS      V N I A L P P P P G I A P P P P P G F G
Lewis    GTGAACATTGCCCCGCCCCACCCCTGGGATGCCCCACCCACCCCCAGGTTTGGG 1140
          GTGAACATTGCCCCGCCCCACCCCTGGGATGCCCCACCCACCCCCAGGTTTGGG 1140
          *
DSS      F H L F H F M G F P P P F M R A E G F I
Lewis    CCACACTTGTCCACCCAATGGGACCGCCCCCTCTTTCATGAGAGCTCCAGGACCAATC 1200
          CCACACTTGTCCACCCAATGGGACCGCCCCCTCTTTCATGAGAGCTCCAGGACCAATC 1200
          *
DSS      H Y F S Q D P Q R M G A H A G K E E S P
Lewis    CACTACCCGTCTCAGGACCTCAGAGAATGGGAGCTCATGCTGGAAAGCACAGCAGCCCC 1260
          CACTACCCGTCTCAGGACCTCAGAGAATGGGAGCTCATGCTGGAAAGCACAGCAGCCCC 1260
          *
DSS      TAG 1263
Lewis    TAG 1263
          ***

```

Footnote to Supplement 9: * indicates nucleotide identity. Amino acid sequence is given on top. Genomic DNAs and adrenal cDNAs from DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of each exon. Exon-intron junctions are the same (data not shown).

Supplement 10. Coding sequence alignment of the *arylsulfatase family, member 1* gene (*Ars1*) between Dahl salt-sensitive (DSS) and Lewis rats

```

M H A L S G F S L V S L L S L G Y L S W
DSS ATGCATGCCCTCTCAGGATTCTCTCTTGTGAGCCTACTCAGTTTGGGCTACCTGTCTCTGG 60
Lewis ATGCATGCCCTCTCAGGATTCTCTCTTGTGAGCCTACTCAGTTTGGGCTACCTGTCTCTGG 60
*****
D W A K F G L V A D G P A E A E D Q F S
DSS GATTGGGCCAAGCCCGGCTTGTGGCTGATGGGCTGCGGAGSCTGAGSATCAGCCTTCG 120
Lewis GATTGGGCCAAGCCCGGCTTGTGGCTGATGGGCTGCGGAGSCTGAGSATCAGCCTTCG 120
*****
A A P F Q F F H I I F I L T D D Q G Y H
DSS GCGGCTCCACCGCAACCTCCACACATTATCTTCATTCTCACCSATGACCAAGGGTACCAC 180
Lewis GCGGCTCCACCGCAACCTCCACACATTATCTTCATTCTCACCSATGACCAAGGGTACCAC 180
*****
D V G Y H G S D I E T P T L D R L A A E
DSS GACGTGGGATACCATGGCTCAGATATTGAGACCCCAACGCTGGACCGGCTGGCAGCTGAG 240
Lewis GACGTGGGATACCATGGCTCAGATATTGAGACCCCAACGCTGGACCGGCTGGCAGCTGAG 240
*****
G V K L E N Y Y I Q P I C T P S R S Q L
DSS GCGCTCAAACCTGGAGAATTATTACATCCAAACCCATATGTACGCGCTTCGAGGAGCCAACTC 300
Lewis GCGCTCAAACCTGGAGAATTATTACATCCAAACCCATATGTACGCGCTTCGAGGAGCCAACTC 300
*****
      ↓
L T G R Y Q I R T G L Q H E I I R P R Q
DSS CTCACTGGCAGGTACCAGATCCATACAGGACTACAGCACTCCATCATCCGCCACGGCAG 360
Lewis CTCACTGGCAGGTACCAGATCCATACAGGACTACAGCACTCCATCATCCGCCACGGCAG 360
*****
F N C L F L D Q V T L P Q K L Q E A G Y
DSS CCCAACTGTCTGCTCTGGACCASSTGACGCTGCCCCAGAAGCTACAGGAAGCAGGCTAC 420
Lewis CCCAACTGTCTGCTCTGGACCASSTGACGCTGCCCCAGAAGCTACAGGAAGCAGGCTAC 420
*****
S T H M V G E W H L G F Y R K E C L E T
DSS TCCACCCACATGGTGGCAASTGGCATCTGGGTTTCTACCGGAAGSAGTGTCTGCCCTACC 480
Lewis TCCACCCACATGGTGGCAASTGGCATCTGGGTTTCTACCGGAAGSAGTGTCTGCCCTACC 480
*****
R R G F D T F L G S L T G H V D Y Y T Y
DSS CGCAGGGGCTTTGACACCTTCCTGGGTTCCCTCACAGGCAATGTGGACTATTACACCTAT 540
Lewis CGCAGGGGCTTTGACACCTTCCTGGGTTCCCTCACAGGCAATGTGGACTATTACACCTAT 540
*****
D W C D G F G V C G F D L H E G E S V A
DSS GACAACCTGTGATGGCCAGGGGTTTGTGGTTTTGATCTGCATGAGGGTGAAGTGTGGCC 600
Lewis GACAACCTGTGATGGCCAGGGGTTTGTGGTTTTGATCTGCATGAGGGTGAAGTGTGGCC 600
*****
W G L S G Q Y S T M L Y A Q R A S H I L
DSS TGGGGGCTCAGCGCCAGTACTCCACTATGCTCTATGCTCAGCGTGAAGCCACATCCTG 660
Lewis TGGGGGCTCAGCGCCAGTACTCCACTATGCTCTATGCTCAGCGTGAAGCCACATCCTG 660
*****
A S H S P Q K P L F L Y V A F Q A V H T
DSS GCCAGCCACAGTCCCAGAAAGCCCTCTTCTCTATGTTGGCCTTCAGGCAGTACACACA 720
Lewis GCCAGCCACAGTCCCAGAAAGCCCTCTTCTCTATGTTGGCCTTCAGGCAGTACACACA 720
*****
P L Q S P R E Y L Y R Y R T M G N V A R
DSS CCCCTACAGTCACTCGAGASTACCTCTACCGCTACCGCACAATGGGCAACGTAGCGAGG 780
Lewis CCCCTACAGTCACTCGAGASTACCTCTACCGCTACCGCACAATGGGCAACGTAGCGAGG 780
*****
R K Y A A N V T C M D E A V R K I T W A
DSS CGCAAGTACGCAGCCATGGTGAACCTGCATGGATGAAGCTGTGGCAACATCACTGGGCT 840
Lewis CGCAAGTACGCAGCCATGGTGAACCTGCATGGATGAAGCTGTGGCAACATCACTGGGCT 840
*****
L K R Y G F Y H N S V I I F S S D N G G
DSS CTCAAGCGCTATGGTTTCTATAACAACAGTGTCAATTATCTTCTCCAGTGAACAACGGTGG 900
Lewis CTCAAGCGCTATGGTTTCTATAACAACAGTGTCAATTATCTTCTCCAGTGAACAACGGTGG 900
*****

```



```

          Q T F S G G S E H W F L R G R K G T Y W E
DSS      CAAACCTTCTCAGGGGGTAGTAACTGGCCCTTCGAGGACGAAAGGGTACTTATTGGGAA 960
Lewis    CAAACCTTCTCAGGGGGTAGTAACTGGCCCTTCGAGGACGAAAGGGTACTTATTGGGAA 960
          *
          G G V R G L G F V H S P L L K K K R R T
DSS      GGAGGTGTGAGGGGGCTGGGTTTTGTCCACAGCCCACTGCTCAAGAAAAAAGACGGACC 1020
Lewis    GGAGGTGTGAGGGGGCTGGGTTTTGTCCACAGCCCACTGCTCAAGAAAAAAGACGGACC 1020
          *
          S R A L V H I T D W Y F T L V G L A G G
DSS      AGTCGGCCCTTGGTCCATATCACAGACTGGTACC CAACACTGGTGGGTCTGGCAGGTGGC 1080
Lewis    AGTCGGCCCTTGGTCCATATCACAGACTGGTACC CAACACTGGTGGGTCTGGCAGGTGGC 1080
          *
          T T S A A D G L D G Y D V W P A I S E G
DSS      ACTACATCCGCTGCTGATGGACTTGATGGCTATGACGTGTGGCCAGCCATTAGTGGAGGC 1140
Lewis    ACTACATCCGCTGCTGATGGACTTGATGGCTATGACGTGTGGCCAGCCATTAGTGGAGGC 1140
          *
          R A S P R T E I L H H I D P L Y N H A R
DSS      CGSGCCTCACCACGACAGAGATCCTACACAACATTGACCCCTCTACAACCATGCCCCG 1200
Lewis    CGSGCCTCACCACGACAGAGATCCTACACAACATTGACCCCTCTACAACCATGCCCCG 1200
          *
          H G E L E G G F G I W N T A V Q A A I R
DSS      CATGGCTCCTTGGAGGGTGGCTTTGGCATCTGGAATACAGCAGTGCAGSCTGCTATCCGA 1260
Lewis    CATGGCTCCTTGGAGGGTGGCTTTGGCATCTGGAATACAGCAGTGCAGSCTGCTATCCGA 1260
          *
          V G E W K L L T G D F G Y G D W I P P Q
DSS      GTCGGAGAGTGGAAAGCTCCTCACTGGAGACCCAGGTTATGGTGACTGGATTCCACCTCAG 1320
Lewis    GTCGGAGAGTGGAAAGCTCCTCACTGGAGACCCAGGTTATGGTGACTGGATTCCACCTCAG 1320
          *
          T L A S F F G S W W H L E R M A S I R Q
DSS      ACAC TGGCCCTCCTTCCCTGGCAGTTGGTGGAAACCTAGAGCGGATGGCCAGCATCCGCCAG 1380
Lewis    ACAC TGGCCCTCCTTCCCTGGCAGTTGGTGGAAACCTAGAGCGGATGGCCAGCATCCGCCAG 1380
          *
          A V W L F N H I S A D P Y E R E D L A D Q
DSS      GCTGTGGCTCCTTTAACATCAGTGTGACCCCTTATGAACGAGAGGACCTAGCTGACCAG 1440
Lewis    GCTGTGGCTCCTTTAACATCAGTGTGACCCCTTATGAACGAGAGGACCTAGCTGACCAG 1440
          *
          R P D V V R T L L A R L A D Y N R T A I
DSS      CGACCTGATGTAGTCCGACCCCTGCTGGCTCGCCCTGGCTGATTATAACCGTACTGCCATC 1500
Lewis    CGACCTGATGTAGTCCGACCCCTGCTGGCTCGCCCTGGCTGATTATAACCGTACTGCCATC 1500
          *
          F V R Y P A A N P R A R P D F N G G A W
DSS      CCTGTGCGTTACCCAGCTSCGAACCTCGGGCCCATCCTGACTTAAATGGGGGTGCTTGG 1560
Lewis    CCTGTGCGTTACCCAGCTSCGAACCTCGGGCCCATCCTGACTTAAATGGGGGTGCTTGG 1560
          *
          G F W A S D E D E E E E D E E E E G R A
DSS      GGACCCCTGGSCCAGTATGAGSATGAAGAAGAGSAGSATGAAGAGSAGSAAGCAGAGCT 1620
Lewis    GGACCCCTGGSCCAGTATGAGSATGAAGAAGAGSAGSATGAAGAGSAGSAAGCAGAGCT 1620
          *
          R S F P R G R R K K E C K I C K L R S F
DSS      AGAAGTTTCCCCGGGGTGCAGCAAGAAGAAATGCAAGATTTGCAAGCTTCGATCTTTT 1680
Lewis    AGAAGTTTCCCCGGGGTGCAGCAAGAAGAAATGCAAGATTTGCAAGCTTCGATCTTTT 1680
          *
          F R K L N T R L N S H R I -
DSS      TTCGGTAAACTCAATACCAGSCTGATGTCCACCGGATCTGA 1722
Lewis    TTCGGTAAACTCAATACCAGSCTGATGTCCACCGGATCTGA 1722
          *

```

Footnote to supplement 10: * indicates nucleotide identity. Amino acid sequence is given on top. Genomic DNAs from DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of exon 1. The exon-intron junction is the same (data not shown).

Supplement 11. Coding sequence alignment of the *esophagus cancer-related protein 2* gene (*Ecg2*) between Dahl salt-sensitive (DSS) and Lewis rats.

```

DSS      M K L L D G L L L L F T A T C L C H S F
ATGAAGTCCTTGGTGGTCTCTGCTGCTCTTCACAGCAACCTGTTTGGCAACAGCTTC 60
Lewis   ATGAAGTCCTTGGTGGTCTCTGCTGCTCTTCACAGCAACCTGTTTGGCAACAGCTTC 60
*****

          ↓                               ↓
DSS      S E V T S H F S F T V D C D I Y K K Y P
TCCGAAGTTACCAGCCACCCTTCACCCACAGTGGACTGTGACATATACAAGAAGTACCCA 120
Lewis   TCCGAAGTTACCAGCCACCCTTCACCCACAGTGGACTGTGACATATACAAGAAGTACCCA 120
*****

DSS      V V A I P C P I E H I P V C G S D Y I T
GTAGTGGCCATCCCTTGCCCCATTGAAACATGCCAGTTTGTGGATCTGACTATATCACT 180
Lewis   GTAGTGGCCATCCCTTGCCCCATTGAAACATGCCAGTTTGTGGATCTGACTATATCACT 180
*****

          ↓
DSS      Y G H K C K L C T E I L R S N G E I Q F
TACGGGAATAAATGCAAGTTGTGTACAGAGATCTTGAGAAGTAATGGGAAAATTCAGTTT 240
Lewis   TACGGGAATAAATGCAAGTTGTGTACAGAGATCTTGAGAAGTAATGGGAAAATTCAGTTT 240
*****

DSS      L H E E H C -
CTTCATGAAGGGCACTGCTAA 261
Lewis   CTTCATGAAGGGCACTGCTAA 261
*****

```

Footnote to supplement 11: * indicates nucleotide identity. Amino acid sequence is given on top. Genomic DNAs from DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of each exon. The exon-intron junctions are the same (data not shown).

Supplement 12. Coding sequence alignment of the *adenomatosis polyposis coli* down-regulated 1 gene (*Apcdd1*) between Dahl salt-sensitive (DSS) and Lewis rats

```

DSS      N V N P F Y T P L S G I H M H T E L M R
Lewis    ATGGTGAACCCCCCTACCCCCACTTTCAGGGATTACATGCACACTGAGCTCATGAGA 60
          ATGGTGAACCCCCCTACCCCCACTTTCAGGGATTACATGCACACTGAGCTCATGAGA 60
          *
          R L K S L N V F V S L F Q N H N R A C I
DSS      AGACTAAAGTCTCTGAACGTTTTTGTTCCTTTGTTTCAGAAATCACAACCATGCGTGCATA 120
Lewis    AGACTAAAGTCTCTGAACGTTTTTGTTCCTTTGTTTCAGAAATCACAACCATGCGTGCATA 120
          *
          A C R I I Y R S D E H H P P I L F P K A
DSS      GCCTGCCGCATCATTACCGGTCCGATGAACACCACCCCTCCCATCCTGCCCCCCAAGGCT 180
Lewis    GCCTGCCGCATCATTACCGGTCCGATGAACACCACCCCTCCCATCCTGCCCCCCAAGGCT 180
          *
          D L T I G L H G E W V S Q R C E V R P E
DSS      GACCTGACCATTTGGCCTCCATGGGAATGGGTGAGCCAGCGCTGTGAGGTACGCCCGSAA 240
Lewis    GACCTGACCATTTGGCCTCCATGGGAATGGGTGAGCCAGCGCTGTGAGGTACGCCCGSAA 240
          *
          V L F L T R H F I F H D M N N T W E G H
DSS      GTCCTCTTCCCTCACCCGCCACTTCATCTTCCATGACAACAACAACCTGGGAAGGGCAT 300
Lewis    GTCCTCTTCCCTCACCCGCCACTTCATCTTCCATGACAACAACAACCTGGGAAGGGCAT 300
          *
          Y Y H Y E D P V C E H P T F T I Y A R G
DSS      TACTACCACTACTCGGACCCCGTCTGCAAGCACCCACATTACCATCTACGCTCGAGGC 360
Lewis    TACTACCACTACTCGGACCCCGTCTGCAAGCACCCACATTACCATCTACGCTCGAGGC 360
          *
          R Y S E R G V L S S R V M G G T E F V F K
DSS      CGCTACAGCCCGGTGTGCTGTCTATCTAGGGTCAATGGGTGGCACAGAATTTGTGTTCAA 420
Lewis    CGCTACAGCCCGGTGTGCTGTCTATCTAGGGTCAATGGGTGGCACAGAATTTGTGTTCAA 420
          *
          ↓
          V N H M K V T P M D A A T A S L L N V F
DSS      GTGAATCACATGAAGGTCACTCCCATGGATGCAGCCACAGCCTCCCTCCTCAATGTCTTC 480
Lewis    GTGAATCACATGAAGGTCACTCCCATGGATGCAGCCACAGCCTCCCTCCTCAATGTCTTC 480
          *
          E G N E C G A E G S W Q V G I Q Q D V T
DSS      AGTGGGAATGAGTGTGGGCTGAAGGCTCCTGGCAGGTGGGTATCCAGCAGGATGTGACA 540
Lewis    AGTGGGAATGAGTGTGGGCTGAAGGCTCCTGGCAGGTGGGTATCCAGCAGGATGTGACA 540
          *
          H T N G C V A L G I K L P H T E Y E I F
DSS      CACACCAATGGCTGTGTGGCTCTGGGCATCAAACCTACCTCACACAGATATGAGATCTTC 600
Lewis    CACACCAATGGCTGTGTGGCTCTGGGCATCAAACCTACCTCACACAGATATGAGATCTTC 600
          *
          K N E Q D A R G R Y L L F N G Q R P S D
DSS      AAAATGGAGCAAGACGCCCGGGCCGCTACCTGCTATTCAATGGCCAGAGGCCAGCGAT 660
Lewis    AAAATGGAGCAAGACGCCCGGGCCGCTACCTGCTATTCAATGGCCAGAGGCCAGCGAT 660
          *
          G S S F D R P E K R A T S Y Q M P L V Q
DSS      GGTTCAGCCCCAGACAGACCCGAGAAGAGAGCCACATCCTACCAGATGCCCTTGGTCCAG 720
Lewis    GGTTCAGCCCCAGACAGACCCGAGAAGAGAGCCACATCCTACCAGATGCCCTTGGTCCAG 720
          *
          C A S S S P R A E D I L E D S R A H L Y
DSS      TGTGCCCTTCTCCTCACCGAGAGCTGAAGACTTGTGGAAGACAGTCCGAGCCCATCTGTAT 780
Lewis    TGTGCCCTTCTCCTCACCGAGAGCTGAAGACTTGTGGAAGACAGTCCGAGCCCATCTGTAT 780
          *
          G R A A G R T A G S L L L F A F V G L W
DSS      GGCCGAGCAGCAGGAGGACAGCTGGGTCCCTGTTGCTTCTGCTTCTGCTTCTGCTGCTTGG 840
Lewis    GGCCGAGCAGCAGGAGGACAGCTGGGTCCCTGTTGCTTCTGCTTCTGCTTCTGCTGCTTGG 840
          *

```

```
          T L P H W R V L R -
DSS      A C A C T A C C A C A T T G G C G C G T C C T C A G A T A G 870
Lewis    A C A C T A C C A C A T T G G C G C G T C C T C A G A T A G 870
          *****
```

Footnote to Supplement 12: * indicates nucleotide identity. Amino acid sequence is given on top. Genomic DNAs from DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of exon 1. The exon-intron junctions are the same (data not shown).

Supplement 13. Coding sequence alignment of the *WD repeat domain 7* gene (*Wdr7*) between Dahl salt-sensitive (DSS) and Lewis rats.

```

M A G N S L V L P I V L W G R K A P T H
DSS  ATGGCAGGAAACAGTCTAGTCTGCCCATTGTTCTCTGGGGCCGCAAAGCACCCACACAC 60
Lewis ATGGCAGGAAACAGTCTAGTCTGCCCATTGTTCTCTGGGGCCGCAAAGCACCCACACAC 60
*****
C I S S I L L T D D G G T I V T G C H D
DSS  TGTATTTCTCAACTACTGCTGACAGATGATGGGGCCACGATTGTGACTGGATGCCACGAT 120
Lewis TGTATTTCTCAACTACTGCTGACAGATGATGGGGCCACGATTGTGACTGGATGCCACGAT 120
*****
                                     ↓
G Q I C L W D L S E E L E V N P R A L L
DSS  GGACAGATATGTCTCTGGGATCTCTCAGAAAGAGCTGSAAGTTAATCCCCGAGCACTATTA 180
Lewis GGACAGATATGTCTCTGGGATCTCTCAGAAAGAGCTGSAAGTTAATCCCCGAGCACTATTA 180
*****
F G H T A A I T C L S K A C A S G D K Q
DSS  TTTGGTCACACAGCAGCCATCACTTGTGTTGTCAAAGCCTGTGCTTCTGGAGACAAACAG 240
Lewis TTTGGTCACACAGCAGCCATCACTTGTGTTGTCAAAGCCTGTGCTTCTGGAGACAAACAG 240
*****
                                     ↓
Y T V S A S A N G E M C L W D V N D G R
DSS  TACACCGTGAGTSCGTCCSCAAATGGAGAGATGTGCCTCTGGGATGTGAATGACGGCAGA 300
Lewis TACACCGTGAGTSCGTCCSCAAATGGAGAGATGTGCCTCTGGGATGTGAATGACGGCAGA 300
*****
                                     ↓
C I E F T K L A C T H T G I O F Y Q E S
DSS  TGTATTGAATTTACAAAATTAGCCTGCACACATACTGGCATAACAGTTCCTACCGATTCTCT 360
Lewis TGTATTGAATTTACAAAATTAGCCTGCACACATACTGGCATAACAGTTCCTACCGATTCTCT 360
*****
V G N Q R E G R L L C H G H Y P E I L V
DSS  GTTGGAAATCAGCGAGAGGGAAGGCTTCTGTGCCATGGGCATTACCCCTGAAATCCCTCGTT 420
Lewis GTTGGAAATCAGCGAGAGGGAAGGCTTCTGTGCCATGGGCATTACCCCTGAAATCCCTCGTT 420
*****
V D A T S L E V L Y S L V S K I S P D W
DSS  GTGGATGCCACCGCCTTGAGGTGTGATTTCTTTGGTATCAAAGATATCACCAGACTGG 480
Lewis GTGGATGCCACCGCCTTGAGGTGTGATTTCTTTGGTATCAAAGATATCACCAGACTGG 480
*****
                                     ↓
I S S M S I I R S H R T Q E D T V V A L
DSS  ATTAGCTCCATGAGTATTATCCGCTCTCACCGGACACAAGAGSACACTGTGGTAGCCGTG 540
Lewis ATTAGCTCCATGAGTATTATCCGCTCTCACCGGACACAAGAGSACACTGTGGTAGCCGTG 540
*****
                                     ↓
S V T G I L K V W I V T S E I S G L Q D
DSS  TCCGTGACAGGTATTCTGAAGGTGTGGATTGTGACCTCTGAAATTAGTGGATTGCAGGAC 600
Lewis TCCGTGACAGGTATTCTGAAGGTGTGGATTGTGACCTCTGAAATTAGTGGATTGCAGGAC 600
*****
T E P I F E E E S K P I Y C Q N C Q S L
DSS  ACTGAGCCAATATTTGAGGAGGAATCCAAACCAATTTATTGTCAGAATTGCCAAAGCCCTC 660
Lewis ACTGAGCCAATATTTGAGGAGGAATCCAAACCAATTTATTGTCAGAATTGCCAAAGCCCTC 660
*****
                                     ↓
S F C A F T Q R S L L V V C S K Y W R V
DSS  TCTTTTGTGCACTTACACAGAGGTCCCTCTTGGTTGTATGCTCCAAGTACTGGAGGGTG 720
Lewis TCTTTTGTGCACTTACACAGAGGTCCCTCTTGGTTGTATGCTCCAAGTACTGGAGGGTG 720
*****
F D A G D Y S L L C S G P S E D G Q T N
DSS  TTCGATGCTGGCGACTACTCCCTGTGTGTTCCAGGTCTAGTGAAGATGGACAGACATGG 780
Lewis TTCGATGCTGGCGACTACTCCCTGTGTGTTCCAGGTCTAGTGAAGATGGACAGACATGG 780
*****

```



```

T G G D F V S A D E V I I W T E N G Q S
DSS      ACTGGAGGGGACITTTGTCTGCGAGACAAAGTCATTATTTGGACTGAAAACGGGCAGAGT 840
Lewis    ACTGGAGGGGACITTTGTCTGCGAGACAAAGTCATTATTTGGACTGAAAACGGGCAGAGT 840
*****

          ↓
Y I Y K L P A S C L P A S D S F R S D V
DSS      TACATTTACAAACTCCCTGCCAGTTGCCCTCCAGCTAGTGATTCAATCCGCAGTGACGTG 900
Lewis    TACATTTACAAACTCCCTGCCAGTTGCCCTCCAGCTAGTGATTCAATCCGCAGTGACGTG 900
*****

G K A V E N L I P P V Q H S L L D Q K D
DSS      GGGAAAGCASTTAAAAATCTGATTCTCCCGTGACGATAGCCTCTTGGATCAGAAAGAT 960
Lewis    GGGAAAGCASTTAAAAATCTGATTCTCCCGTGACGATAGCCTCTTGGATCAGAAAGAT 960
*****

          ↓
R E L V I C P P V T R F F Y G C K E Y L
DSS      AGAGAGTTGGTAATTTGTCTCCTGTACTCGGTTTTTCTATGGATGCAAGGAATATTTG 1020
Lewis    AGAGAGTTGGTAATTTGTCTCCTGTACTCGGTTTTTCTATGGATGCAAGGAATATTTG 1020
*****

H K L L I Q G D S S G R L S I W N I A D
DSS      CATAAGCTACTAATTCAGSGTGATTCTTCTGGAGGTTAAGTATTGGAAACATAGCAGAC 1080
Lewis    CATAAGCTACTAATTCAGSGTGATTCTTCTGGAGGTTAAGTATTGGAAACATAGCAGAC 1080
*****

          ↓
I A D K Q E A N E G L E T T T C I S L Q
DSS      ATAGCAGACAACAGGAAGCCAATGAAGGGCTAAAAACGACAACCTTGTATTAGTTTGCAG 1140
Lewis    ATAGCAGACAACAGGAAGCCAATGAAGGGCTAAAAACGACAACCTTGTATTAGTTTGCAG 1140
*****

D A F D K L E P C P A G I I D Q L S V I
DSS      GATGCATTTGACAAACTGAAGCCTTGCTCTGCGAATTATCGATCAGCTGAGTGTGATT 1200
Lewis    GATGCATTTGACAAACTGAAGCCTTGCTCTGCGAATTATCGATCAGCTGAGTGTGATT 1200
*****

F W E N E P L E V T A S V Y I P A H G R
DSS      CCAACAGCAACGAACCACTTAAAGTAACGCGAGTGTCTACATACCAGCACACGGGGGC 1260
Lewis    CCAACAGCAACGAACCACTTAAAGTAACGCGAGTGTCTACATACCAGCACACGGGGGC 1260
*****

L V C G R E D G S I I I V P A T Q T A I
DSS      CTTGTTTGGCGCCGGGAAGACGGAAGCATCATTATTTGCCCTGCCACCCAGACGGGCATA 1320
Lewis    CTTGTTTGGCGCCGGGAAGACGGAAGCATCATTATTTGCCCTGCCACCCAGACGGGCATA 1320
*****

          ↓
V Q L L Q G E H M L R R G W F F H R T L
DSS      GTTCAGCTGCTCAGGGAGAACACATGCTCAGACGAGGTTGGCCCCGACAGAACCCCTC 1380
Lewis    GTTCAGCTGCTCAGGGAGAACACATGCTCAGACGAGGTTGGCCCCGACAGAACCCCTC 1380
*****

R G H R N E V T C L L L Y P H Q V E A R Y
DSS      CGTGGCCACCGGAACAAAGTCACATGTTTGGCTGTATCCTCATCAGGTCTCAGCTCGGTAT 1440
Lewis    CGTGGCCACCGGAACAAAGTCACATGTTTGGCTGTATCCTCATCAGGTCTCAGCTCGGTAT 1440
*****

D Q R Y L I S G G V D F S V I I W D I F
DSS      GATCAAAGATACCTGATATCCGGAGGTGTGGATTTTCCGTCATCATTGGGACATTTT 1500
Lewis    GATCAAAGATACCTGATATCCGGAGGTGTGGATTTTCCGTCATCATTGGGACATTTT 1500
*****

S G E M K H I F C V H G G E I T Q L L V
DSS      TCCGGAGAAATGAAACATATCTTCTGTGTTTCATGGTGGTGGATCACACAACCTCTGGTC 1560
Lewis    TCCGGAGAAATGAAACATATCTTCTGTGTTTCATGGTGGTGGATCACACAACCTCTGGTC 1560
*****

          ↓
F P E N C S A R V Q H C V C S V A S D H
DSS      CCGCCAGAAAACGTGAGTSCAAGAGTTCAACACTGCCTCTGTTCTGTGGCCAGTGACCAC 1620
Lewis    CCGCCAGAAAACGTGAGTSCAAGAGTTCAACACTGCCTCTGTTCTGTGGCCAGTGACCAC 1620
*****

S V G L L S L R E E R C I M L A E R H L
DSS      TCTGTAGGCTGCTAAGTCTGGGAGAGAAAAAATGCATCATGCTGGCGTCTCGTCACCTG 1680
Lewis    TCTGTAGGCTGCTAAGTCTGGGAGAGAAAAAATGCATCATGCTGGCGTCTCGTCACCTG 1680
*****

```

```

F P I Q V I E W R P S D D Y L V V G C T
DSS TTCCCTATTGAGGTGATCAAGTGGAGGCTTCTGACGACTACCTGGTGGTGGGGTGCACG 1740
Lewis TTCCCTATTGAGGTGATCAAGTGGAGGCTTCTGACGACTACCTGGTGGTGGGGTGCACG 1740
*****

          ↓
D G S V C V W Q M D T G A L D R C A M G
DSS GACGGCTCTGTGTGTCTGGCAGATGGACACTGGTGCCTGGACCGCTGTGCAATGGGG 1800
Lewis GACGGCTCTGTGTGTCTGGCAGATGGACACTGGTGCCTGGACCGCTGTGCAATGGGG 1800
*****

I T A V E I L N A C D E A V F A A V D S
DSS ATAACAGCCGTGGAGATTCTCAATGCTTGTGACGAAGCTGTCCCTGCAGCASTGGACTCA 1860
Lewis ATAACAGCCGTGGAGATTCTCAATGCTTGTGACGAAGCTGTCCCTGCAGCASTGGACTCA 1860
*****

L E E F A V N L K Q A M T R R S L A A L
DSS CTCAGTCACCCAGCACTCAACCTGAAGCAAGCCATGACACGGCGGAGTCTCGCCGCCCTT 1920
Lewis CTCAGTCACCCAGCACTCAACCTGAAGCAAGCCATGACACGGCGGAGTCTCGCCGCCCTT 1920
*****

K N M A H H K L Q T L A T N L L A S E A
DSS AAAAACATGGCCACCACAAAGCTGCAAAACCTTGCACCTAACCTTTTGGCTTCTGAGGCC 1980
Lewis AAAAACATGGCCACCACAAAGCTGCAAAACCTTGCACCTAACCTTTTGGCTTCTGAGGCC 1980
*****

          ↓
S D K G N L F R Y S H N S L M V Q A I F
DSS YCTGACAAAGGGGATTTACCTAAATATCTCATAACTCCCTGATGGTTCAAGCAATAAAG 2040
Lewis YCTGACAAAGGGGATTTACCTAAATATCTCATAACTCCCTGATGGTTCAAGCAATAAAG 2040
*****

T N L T D F D I R V L F F D V E A L I I
DSS ACAAACTTAACAGACCCGGATATCCATGTGCTTTCTTTGATGTGGAAGCTTTGATTATT 2100
Lewis ACAAACTTAACAGACCCGGATATCCATGTGCTTTCTTTGATGTGGAAGCTTTGATTATT 2100
*****

Q L L T E E A S R P N T A L I S F E N L
DSS CAACTCCTGACTGAAGAAGCCTTAGGCCGAATACTGCACCTATTTCCTCCAGAGAATCTG 2160
Lewis CAACTCCTGACTGAAGAAGCCTTAGGCCGAATACTGCACCTATTTCCTCCAGAGAATCTG 2160
*****

Q K A S G S S D K G S S F L T G K R A A
DSS CAGAAAGCATCTGGCAGTTCCAGACAAAGGGGGCTCTTTCTGACTGGAAAACGAGCGGCA 2220
Lewis CAGAAAGCATCTGGCAGTTCCAGACAAAGGGGGCTCTTTCTGACTGGAAAACGAGCGGCA 2220
*****

V L F Q Q V E E T I R E H I K E H L L D
DSS GTTCTTTCCAGCAAGTGAAGAAACTATCAAAGAGAACATAAAGGAGCACCTCCCTTGAT 2280
Lewis GTTCTTTCCAGCAAGTGAAGAAACTATCAAAGAGAACATAAAGGAGCACCTCCCTTGAT 2280
*****

E E E D E E E V M R Q R R E E S D P E Y
DSS GAGSAGSAGSACGAGSAGSAGGTGATGAGGCAGAGGAGGSAAGAAAGTACCCCTGAGTAC 2340
Lewis GAGSAGSAGSACGAGSAGSAGGTGATGAGGCAGAGGAGGSAAGAAAGTACCCCTGAGTAC 2340
*****

R A S K S K F L T L L E Y N L T M D T A
DSS CGGGCCAGCAAGTCCAAACCACTTACCCTACTAGAATACAACCTTACTATGGATACCSCA 2400
Lewis CGGGCCAGCAAGTCCAAACCACTTACCCTACTAGAATACAACCTTACTATGGATACCSCA 2400
*****

K L F M S C L H A W G L N E V L D E V C
DSS AAATTATTGATGCTCTGCTCCACGCTGGGGTTTGAATGAAGTCTGSAATGAAGTTTGC 2460
Lewis AAATTATTGATGCTCTGCTCCACGCTGGGGTTTGAATGAAGTCTGSAATGAAGTTTGC 2460
*****

L D R L G M L K P H C T V S F G L L S R
DSS CTCGATCGCCCTCGGCATGCTGAAAACCACTGCACAGTGTCTTTGGTCTCCTATCCAGA 2520
Lewis CTCGATCGCCCTCGGCATGCTGAAAACCACTGCACAGTGTCTTTGGTCTCCTATCCAGA 2520
*****

G G H M S L M L P G Y H Q A A G K L L Q
DSS GGAGGTGATGCTCTGATGCTTCTGTTATAATCAGGCTGCTGGAAAAGCTACTGCAG 2580
Lewis GGAGGTGATGCTCTGATGCTTCTGTTATAATCAGGCTGCTGGAAAAGCTACTGCAG 2580
*****

```

```

A K A E A G R E G F A T E S V G K G T Y
DSS 0CGAAAGCAGAAGCAGGACGGAAGGGGCCAGCAACGGAGAGCCTAGGCAAGGGGACCTAC 2640
Lewis 0CGAAAGCAGAAGCAGGACGGAAGGGGCCAGCAACGGAGAGCCTAGGCAAGGGGACCTAC 2640
*****
T V S R A V T T Q H L L S I I S L A N T
DSS ACAGTGTCCCGAGCGGTCAACCAGCAACATCTGTGTGCCATCATATCTCTGGCAAAATACT 2700
Lewis ACAGTGTCCCGAGCGGTCAACCAGCAACATCTGTGTGCCATCATATCTCTGGCAAAATACT 2700
*****
↓
L M S M T N A T F I G D H M K K G F T R
DSS TTAATGAGCATGACCAATGCGACGTTTCATTGGAGATCACATGAAGAAGGGCCCCACCAGG 2760
Lewis TTAATGAGCATGACCAATGCGACGTTTCATTGGAGATCACATGAAGAAGGGCCCCACCAGG 2760
*****
F P R F G T F D L S K A R D S P P A S S
DSS CCGCCTAGACCAGGACCCCCAGACCTTTCTAAGSCAAGGGATTCCCCCCCAGCCTCCAGT 2820
Lewis CCGCCTAGACCAGGACCCCCAGACCTTTCTAAGSCAAGGGATTCCCCCCCAGCCTCCAGT 2820
*****
↓
N I V Q G Q I K Q A A A P V S A R S A A
DSS AACATTGTGCAAGGACAGATTA AACCAAGCTGCTGGCCCTGTCTCTGCTCGGTCTGCCGCC 2880
Lewis AACATTGTGCAAGGACAGATTA AACCAAGCTGCTGGCCCTGTCTCTGCTCGGTCTGCCGCC 2880
*****
D H S S S A S A S F A L R T C F L V H E
DSS GACCACTCTGGCTCTGCCTCTGCCTCTCCTGCTTTACGTACCTGCTTCTTAGTGAATGAA 2940
Lewis GACCACTCTGGCTCTGCCTCTGCCTCTCCTGCTTTACGTACCTGCTTCTTAGTGAATGAA 2940
*****
↓
G W S Q L A A M H C V M L P D L L G L G
DSS GGATGGAGCCAACTAGCTGCCATGCACTGTGTCTATGCTGCCCGACCTGCTGGGCTGGGT 3000
Lewis GGATGGAGCCAACTAGCTGCCATGCACTGTGTCTATGCTGCCCGACCTGCTGGGCTGGGT 3000
*****
↓
K F R F F L L E M L A R R W Q D R C L E
DSS AAATTCAGSCTCCTCTTTCTGGAGATGCTAGCTCGAAGATGGCAAGATCGGTGCTTGGAG 3060
Lewis AAATTCAGSCTCCTCTTTCTGGAGATGCTAGCTCGAAGATGGCAAGATCGGTGCTTGGAG 3060
*****
V R E A A Q A L L L A E L R R I E Q A G
DSS GTGAGAGAGSCTSCACAGGCCCTTCTTCTAGCAGAGCTGAGAAGAATTGAGCAGGCAGGA 3120
Lewis GTGAGAGAGSCTSCACAGGCCCTTCTTCTAGCAGAGCTGAGAAGAATTGAGCAGGCAGGA 3120
*****
R K E T I D T W A P Y L F Q Y M D H V I
DSS CGGAAGGAGACTATTGATACTGGGCTCCTTACTTACCTCAGTACATGGACCATGTGCATA 3180
Lewis CGGAAGGAGACTATTGATACTGGGCTCCTTACTTACCTCAGTACATGGACCATGTGCATA 3180
*****
↓
S F G V T A E A M Q T M A A A P D A S G
DSS TCACCTGGAGTCACGGCGGAAGCCATGCAGACTATGGCAGCTGCTCCAGATGCCTCGGGG 3240
Lewis TCACCTGGAGTCACGGCGGAAGCCATGCAGACTATGGCAGCTGCTCCAGATGCCTCGGGG 3240
*****
↓
F E A K V Q E E E H D L V D D D I T T G
DSS CCAGAAGCCAAAGTCCAGGAAGAAGGATGACCTCGTGGACGATGACATCACCACCTGGT 3300
Lewis CCAGAAGCCAAAGTCCAGGAAGAAGGATGACCTCGTGGACGATGACATCACCACCTGGT 3300
*****
C L S S V P Q M K K M S T S Y E E R R K
DSS TGCTTATCAAGTGTCCACAAATGAAAAAGATGTCCACATCTTACGGAAGAAAGAAAGAA 3360
Lewis TGCTTATCAAGTGTCCACAAATGAAAAAGATGTCCACATCTTACGGAAGAAAGAAAGAA 3360
*****
Q A T A I V L E G V I G A E F G A E I E
DSS CAGGCCACTGCTATTGTTCTCCTGGGAGTGATAGGAGCAGAGTTTGGAGCTGAAAATTGAA 3420
Lewis CAGGCCACTGCTATTGTTCTCCTGGGAGTGATAGGAGCAGAGTTTGGAGCTGAAAATTGAA 3420
*****

```



```
F P K L L T R P R S S S Q I P E G F G L  
DSS CCACCAAACTGTGACCAGACCTCGGAGCTCTAGTCAAATTCCTGAAGGATTTGGTTTG 3480  
Lewis CCACCAAACTGTGACCAGACCTCGGAGCTCTAGTCAAATTCCTGAAGGATTTGGTTTG 3480  
*****  
T S G G S N Y S L A R H T C K A L T F L  
DSS ACAAGTGGAGGTTCCAACTACTCTCTGGCCAGACATACTGCAAGGCCACTGACATTTCTT 3540  
Lewis ACAAGTGGAGGTTCCAACTACTCTCTGGCCAGACATACTGCAAGGCCACTGACATTTCTT 3540  
*****  
L L Q P P S P K L P P H S T I R R T A I  
DSS CTGCTACAGCCACCAAGTCCCAAATTCCTCCTCATAGCACCATCCGGAGAAGTGGCATT 3600  
Lewis CTGCTACAGCCACCAAGTCCCAAATTCCTCCTCATAGCACCATCCGGAGAAGTGGCATT 3600  
*****  
D L I G R G F T V W E P Y M D V S A V L  
DSS GACCTGATCGGGCGAGGTTTACCCTGTGTGGAGCCTTACATGGACGTGTCTGTGTCCTG 3660  
Lewis GACCTGATCGGGCGAGGTTTACCCTGTGTGGAGCCTTACATGGACGTGTCTGTGTCCTG 3660  
*****  
M G L L E L C A D A E K Q L A N I T H G  
DSS ATGGGGCTGCTGGAGCTGTGTGCAAGTCTGAGAAGCAGCTGGCCAAACATCACAATGGGG 3720  
Lewis ATGGGGCTGCTGGAGCTGTGTGCAAGTCTGAGAAGCAGCTGGCCAAACATCACAATGGGG 3720  
*****  
L P L S P A A D S A R S A R H A L S L I  
DSS CTGCCCTCTGAGCCCTGCAGCTGACTCTGCCCGATCCGCAAGACACGCCCTTTCTCTCATA 3780  
Lewis CTGCCCTCTGAGCCCTGCAGCTGACTCTGCCCGATCCGCAAGACACGCCCTTTCTCTCATA 3780  
*****  
A T A R F F A F I T T I A K E V H R H T  
DSS GCCACCGCCAGACCACCCGCCCTTTCATCACCACCATAGCTAAGGAGGTGCACAGACACAG 3840  
Lewis GCCACCGCCAGACCACCCGCCCTTTCATCACCACCATAGCTAAGGAGGTGCACAGACACAG 3840  
*****  
A L A A N T Q S Q Q S I H T T T L A R A  
DSS GCCCTTGCAGCAAAATACCCAGTCCAGCAGAGTATCCACACCACACACTGGCAAGGGCT 3900  
Lewis GCCCTTGCAGCAAAATACCCAGTCCAGCAGAGTATCCACACCACACACTGGCAAGGGCT 3900  
*****  
K G E I L R V I E I L I E K M P T D V V  
DSS AAAGGCGAAATCCTGAGAGTCAATTGAAATTCCTTATCGAAAAGATGCTACGGATGTTGTG 3960  
Lewis AAAGGCGAAATCCTGAGAGTCAATTGAAATTCCTTATCGAAAAGATGCTACGGATGTTGTG 3960  
*****  
D L L V E V M D I I M Y C L E G S L V K  
DSS GATCTTCTTGTGGAGTTCATGGACATCATCATGTACTGCCCTGSAAGGATCTTTAGTTAAG 4020  
Lewis GATCTTCTTGTGGAGTTCATGGACATCATCATGTACTGCCCTGSAAGGATCTTTAGTTAAG 4020  
*****  
K K G L Q E C F F A I C R F Y M V S Y Y  
DSS AAGAAGGGTCTTCAGGAGTGTTCOCAGCCATCTGCAGSTTCACATGSTCAGCTATTAT 4080  
Lewis AAGAAGGGTCTTCAGGAGTGTTCOCAGCCATCTGCAGSTTCACATGSTCAGCTATTAT 4080  
*****  
E R S H R I A V G A R H G S V A L Y D I  
DSS GAGCGAGTCACAGAAATTCAGTGGAGCAGCCATGGCTCAGTGGCCCTGTATGACATC 4140  
Lewis GAGCGAGTCACAGAAATTCAGTGGAGCAGCCATGGCTCAGTGGCCCTGTATGACATC 4140  
*****  
R T G K C Q T I H G H K G P I T A V S F  
DSS CGSACTGGGAAATGTCAGCAATCCACGGACACAAGGGACCTATCACTGCAAGTGCCTTT 4200  
Lewis CGSACTGGGAAATGTCAGCAATCCACGGACACAAGGGACCTATCACTGCAAGTGCCTTT 4200  
*****  
A P D G R Y L A T Y S H T D S H I S F W  
DSS GCTCCTGATGGGCTTACCTTGCCACCTACTCAAACACTGACAGCCACATTTCTTCTGG 4260  
Lewis GCTCCTGATGGGCTTACCTTGCCACCTACTCAAACACTGACAGCCACATTTCTTCTGG 4260  
*****
```

```

      ↓
      Q M N T S L L G S I G M L N S A P Q L R
DSS   CAGATGAACACCTCACTTCTGGGAAGCATTGGCATGCTGAACTCAGCACCTCAGCTGGC 4320
Lewis CAGATGAACACCTCACTTCTGGGAAGCATTGGCATGCTGAACTCAGCACCTCAGCTGGC 4320
      *
      C I K T Y Q V F F V Q P A S P G S H N A
DSS   TGCATCAAGACCTACCAGGTACCTCCAGTGCAGCCCGCATCCCCTGGCTCGCACAACGCC 4380
Lewis TGCATCAAGACCTACCAGGTACCTCCAGTGCAGCCCGCATCCCCTGGCTCGCACAACGCC 4380
      *
      L R L A R L I W T S N R N V I L M A H D
DSS   CTCAGGTTGGCCCGGCTCATCTGGACTTCCAACCGSAATGTTATCCTCATGGCCACGAT 4440
Lewis CTCAGGTTGGCCCGGCTCATCTGGACTTCCAACCGSAATGTTATCCTCATGGCCACGAT 4440
      *
      G K E H R F M V -
DSS   GGGAAGGAGCACCGCTTCATGGTCTGA 4467
Lewis GGGAAGGAGCACCGCTTCATGGTCTGA 4467
      *

```

Footnote to Supplement 13: * denotes nucleotide identity. The amino acid sequence is given on top. Genomic and cDNAs from the brain of DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of each exon. Intron-exon junctions are the same (data not shown).

Supplement 14. Coding sequence alignment of the *ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3 gene (St8Sia3)* between Dahl salt-sensitive (DSS) and Lewis rats.

```

DSS      M R N C K E M A R V A S I L E S L V M L S V
Lewis    ATGAGAAATTGC AAAATGGCCGAGTCCCGAGTACTAGGGCTGGTCATGCTCAGCGTG 60
          ATGAGAAATTGC AAAATGGCCGAGTCCCGAGTACTAGGGCTGGTCATGCTCAGCGTG 60
          *****
          A L L I L S L I S Y V S L K K E M I F T
DSS      GCCCTGCTGATTTATCGCTTATCAGCTACGTGTCCCTGAAAAAGGAGAACATCTTCACC 120
Lewis    GCCCTGCTGATTTATCGCTTATCAGCTACGTGTCCCTGAAAAAGGAGAACATCTTCACC 120
          *****
          ↓
DSS      T F K Y A S F G A P R M Y M F H A G F R
Lewis    ACTCCCAAGTACCCAGCCCGGGGGCGCCCGAATGTACATGTTCACCGGGGATTCGGG 180
          ACTCCCAAGTACCCAGCCCGGGGGCGCCCGAATGTACATGTTCACCGGGGATTCGGG 180
          *****
          S O F A L K E F L D F S E V F I T M E L T
DSS      TCACAGTTTGCACCTGAAGTTTCTAGACCCGTCATTTGTGCCATTACGAATTCCTCACC 240
Lewis    TCACAGTTTGCACCTGAAGTTTCTAGACCCGTCATTTGTGCCATTACGAATTCCTCACC 240
          *****
          H E L Q E K F S E K W T F N R T A F L H Q
DSS      CATGAACCCAGAGAAACCTTCTAAATGGACATTTAATCGSACAGCGTTTTTACATCAA 300
Lewis    CATGAACCCAGAGAAACCTTCTAAATGGACATTTAATCGSACAGCGTTTTTACATCAA 300
          *****
          ↓
DSS      R Q E I L Q H V D V I K H F E L T E N E
Lewis    AGSCAAGAAATCTTCAGCATGTCCGATGTAATAAAAAATTTTCTTTAACCAAGAAATG 360
          AGSCAAGAAATCTTCAGCATGTCCGATGTAATAAAAAATTTTCTTTAACCAAGAAATG 360
          *****
          V R I S Q L I H Y D Y S S H K Y V F S I
DSS      GTTCGGATCGGACAACCTGATACATTACGATTTATCCAGCCATAAAATATGTTTCTCTATT 420
Lewis    GTTCGGATCGGACAACCTGATACATTACGATTTATCCAGCCATAAAATATGTTTCTCTATT 420
          *****
          S N H F R S L L P D V S F I L N E R Y N
DSS      AGCAATAACTTCGGTCCCTGCTCCCAAGTGTGTCCGCCATTCTGAAATAAGCGTTATAAC 480
Lewis    AGCAATAACTTCGGTCCCTGCTCCCAAGTGTGTCCGCCATTCTGAAATAAGCGTTATAAC 480
          *****
          I C A V V G N S G I L T G S Q C G Q E I
DSS      ATTTGTGCTGTGGTCCGAAACAGTGAATCTTGACAGGAGTCAGTGTGGACAAGAAATA 540
Lewis    ATTTGTGCTGTGGTCCGAAACAGTGAATCTTGACAGGAGTCAGTGTGGACAAGAAATA 540
          *****
          D E S D F V F R C N F A P T E A F H K D
DSS      GATAAATCGGATTTTGTTTTTCGGTGCATTTTGGCCCGACAGAGGCTTTCACAAAAGAT 600
Lewis    GATAAATCGGATTTTGTTTTTCGGTGCATTTTGGCCCGACAGAGGCTTTCACAAAAGAT 600
          *****
          V D K K T N L T T F N P S I L E K Y Y N
DSS      GTTGGGAAGAAAACCAACTCAACTTTCAATCCTAGCATCTTGGAAAAATATTACAAC 660
Lewis    GTTGGGAAGAAAACCAACTCAACTTTCAATCCTAGCATCTTGGAAAAATATTACAAC 660
          *****
          N L L T I Q D R N H F F L S L K K L D G
DSS      AACCTTTAACCATCCAGGACCGTAACAACCTTCTCCTCAGTTTAAAAAAGCTCGAGCGG 720
Lewis    AACCTTTAACCATCCAGGACCGTAACAACCTTCTCCTCAGTTTAAAAAAGCTCGAGCGG 720
          *****
          A I L W I F A F F F H T S A T V T R T L
DSS      GCCATTCTGATCCCGCATTTTCTTCCACACTTCTGCGACTGTAACGAGAACCCTTA 780
Lewis    GCCATTCTGATCCCGCATTTTCTTCCACACTTCTGCGACTGTAACGAGAACCCTTA 780
          *****
          V D F F V E R R S Q L K V Q L A W P G H
DSS      GTGGATTTTTCGTTGAGCACAGAGTCAGTTAAAGGTCCAGTTGGCTTGGCCTGGAAAT 840
Lewis    GTGGATTTTTCGTTGAGCACAGAGTCAGTTAAAGGTCCAGTTGGCTTGGCCTGGAAAT 840
          *****
          ↓
DSS      I M Q H V N R Y W K N K H L S F K R L S
Lewis    ATCATGCACATGTCACAGGTACTGSAAAAACAACACTTGTCCACCAAAACGACTSAGC 900
          ATCATGCACATGTCACAGGTACTGSAAAAACAACACTTGTCCACCAAAACGACTSAGC 900
          *****

```

```

          T S I L M Y T L A E A I C E E I H L Y S
DSS      ACAGSTATCCTTATGTACACCCCTTGCATCTGCAATATGTGAAGAGATCCACTTGTACGGT 960
Lewis   ACAGSTATCCTTATGTACACCCCTTGCATCTGCAATATGTGAAGAGATCCACTTGTACGGT 960
          *
          F W P F S F D P N T R E D L F Y H Y Y D
DSS      TTCTGGCCCTTTCGGATTGACCCCAACACCAGGGAGGATCTGGCCACCACTACTATGAC 1020
Lewis   TTCTGGCCCTTTCGGATTGACCCCAACACCAGGGAGGATCTGGCCACCACTACTATGAC 1020
          *
          K E G T K F T T I R W Q E S H Q L P A E F
DSS      AAAAAAGGAACCAAATTTACCACCAAGTGGCAGGAGTCTCACCAGCTGCCCTGCTGAGTTT 1080
Lewis   AAAAAAGGAACCAAATTTACCACCAAGTGGCAGGAGTCTCACCAGCTGCCCTGCTGAGTTT 1080
          *
          Q L L Y R M H G E S L T K L T L S R C A
DSS      CAGCTGCTCTATCGAATGCATGGGGAAAGGCTCACAAAGCTCACTCTGTACACTGTGCC 1140
Lewis   CAGCTGCTCTATCGAATGCATGGGGAAAGGCTCACAAAGCTCACTCTGTACACTGTGCC 1140
          *
DSS      -
Lewis   TAA 1143
          ***

```

Footnote to Supplement 14: * denotes nucleotide identity. The amino acid sequence is given on top. Genomic DNAs from DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of each exon. The exon-intron junctions are the same (data not shown).

Supplement 15. Coding sequence alignment of the *Ferrochelatase* gene (*Fech*) between Dahl salt-sensitive (DSS) and Lewis rats.

```

DSS      M L S A S A N H A A A L R A A G A I F R
Lewis    ATGCTTTGCGCCGCGCCCAACATGCGCTGCGGCCCTGCGCGGCTGCGGGGCTCTGTGTCGCG 60
          *
          ↓
DSS      E F L V Y G E S R A C Q P W R C Q S G P
Lewis    GAGCCGCTGGTGTATGGCAGCTCGAGGGCCTGTCCAGCCATGGAGGTGCAATCGGGTCCA 120
          *
          *
DSS      A V V A T T E E A H H A K T T K P Q A Q
Lewis    GCAGTGGTGGCCACTACGGAGAAAGCGCATCATGCCAAAACCAAAAACCAAAAGCTCAA 180
          *
          *
          ↓
DSS      F E R R K F K T G I L M L N M G G P E T
Lewis    CCAGAAAGSAGSAGCCCAAACGGGCATATTAATGTTAAACATGGGAGGCCCGGAAACC 240
          *
          *
DSS      L G E V Q D F L Q R L F L D R D L M T L
Lewis    CTTGGAGAGTTCAAGACTTCCTTCAGAGGCTCTTCCTGGACCGAGACTCATGACCCCT 300
          *
          *
          ↓
DSS      F I Q N E L A P F F I A E R R T P E I Q E
Lewis    CCCATTCAAATAAGCTGGCACCATTTCATCGCCAAAACCGGAACCCCAAATTCAGAG 360
          *
          *
DSS      Q Y R I G G G S P I K N W T S K Q S E
Lewis    CAGTATCCAGAAATCGAGGTGGATCCCCATCAAGATGTGGACTTCCAAAGCAAGAGAA 420
          *
          *
          ↓
DSS      G M V K L L D E L S P D T A F H K Y Y I
Lewis    GGCATGGTGAAGCTGCTGGATGAGCTGTCCCTGACACAGCACTCCAAATACTATATT 480
          *
          *
DSS      G F R Y V H F L T E E A I E E M E R D G
Lewis    GGATTCGGTACGTCCTCCCTTGACAGAGAGGCAATTGAAGAGATGGAGAGATGSA 540
          *
          *
          ↓
DSS      L E R A I A E T Q Y P Q Y S C S T E G S
Lewis    CTAGAAAGGGCCATTGCTTTCCACAGATATCCACAGTACAGCTGCTCCACACAGGCAGC 600
          *
          *
DSS      S L N A I Y R Y Y N E V G R K P T K K W
Lewis    AGCTTAAATGCCATTACAGATACTATAACGAGGTGGGGCGSAGCCCACTGAAGTGG 660
          *
          *
          ↓
DSS      S T I D R W P T H F L L I Q C F A D H I
Lewis    ASCACTATCGACAGGTGGCCACGGCACCCCTCCTCATCCAGTGTCTTTCAGACCACATT 720
          *
          *
DSS      L R E L D H F P E E K R S E V V I L F S
Lewis    CTGAAAGAGCTGGACCAATTTCCAGAGGAGAGAGAAATGAGGTGGTCAATCGTTTTCT 780
          *
          *
          ↓
DSS      A H S L P M S V V N R G D F Y F Q E V S
Lewis    GCCCACTCCCTGCGGATGTCTGTTGTCAACAGAGGGGACCCCTATCCCAAGAGGTAGGA 840
          *
          *

```

```

DSS      A T V H R V M E Q L G Y P N F Y R L V W
Lewis    GCCACTGTCCACAGAGTCATGGAGCAGCTGGGTTATCCGAAACCCCTACCGACTGGTCTGG 900
          *****
          ↓
DSS      Q E E V S P V P W L G P Q T D E A I K S
Lewis    CAGTCCAAGGTTGGTCCAGTACCTGGTTGGGOCCTCAAACAGATGAGGCTATCAAAGG 960
          CAGTCCAAGGTTGGTCCAGTACCTGGTTGGGOCCTCAAACAGATGAGGCTATCAAAGG 960
          *****
DSS      L C E R G R E N I L L V P I A F T S D H
Lewis    CTTTGTGAGCGGGGAGGAAGAATATCCTCTTGGTTCCAATAGCATTACCAAGTATCAC 1020
          CTTTGTGAGCGGGGAGGAAGAATATCCTCTTGGTTCCAATAGCATTACCAAGTATCAC 1020
          *****
DSS      I E T L Y E L D I E Y S Q V L A Q K C S
Lewis    ATTGAGACCCCTTATGAACCTGGACATTGAATACTCTCAAAGTGTAGCTCAGAAATGTGGA 1080
          ATTGAGACCCCTTATGAACCTGGACATTGAATACTCTCAAAGTGTAGCTCAGAAATGTGGA 1080
          *****
DSS      A E N I R R A E S L N D N F L F S K A L
Lewis    GCTGAAAACATCAGAAGAGCAGAGTCTCTTAATGGAAATCCATTGTTCTCCAAGGCCCTT 1140
          GCTGAAAACATCAGAAGAGCAGAGTCTCTTAATGGAAATCCATTGTTCTCCAAGGCCCTT 1140
          *****
DSS      A D L V H S H I Q S N K L C S T Q L S L
Lewis    GCTGACCTGGTGCACCTCCACATCCAGTCAAACAAGCTGTGCTCCACGAGTGGAGTCTG 1200
          GCTGACCTGGTGCACCTCCACATCCAGTCAAACAAGCTGTGCTCCACGAGTGGAGTCTG 1200
          *****
DSS      E C P L C V N P V C R E T K E F F T S Q
Lewis    AGCTGTCCGCTCTGTGTAATCCTGTCTGCAGGAAGACTAAATCCTTCCACCAGCCAA 1260
          AGCTGTCCGCTCTGTGTAATCCTGTCTGCAGGAAGACTAAATCCTTCCACCAGCCAA 1260
          *****
DSS      Q L -
Lewis    CAGCTGTGAC 1270
          CAGCTGTGAC 1270
          *****

```

Footnote to supplement 15: * denotes nucleotide identity. The amino acid sequence is given on top. Genomic and cDNAs from the kidneys of DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of each exon. The exon-intron junctions are the same (data not shown).

Supplement 16. Coding sequence alignment of the *asparaginyl-tRNA synthetase* gene (*Nars*) between Dahl salt-sensitive (DSS) and Lewis rats

```

                                     ↓
DSS      N S E E V I R G T A D M V L E L Y V S D
Lewis    ATGTCTCAGAGGTGATCAGGGGGACTGCAGACATGGTTCTAGAGCTCTATGTCTCTGAC 60
          ATGTCTCAGAGGTGATCAGGGGGACTGCAGACATGGTTCTAGAGCTCTATGTCTCTGAC 60
          *****

DSS      R E G N D A T G D G T K E K P F K T G L
Lewis    CGAGAAGGAAATGACBCAACGGGGGATGGAAACCAAGGAGAAACCATTTAAACAGGCCTA 120
          CGAGAAGGAAATGACBCAACGGGGGATGGAAACCAAGGAGAAACCATTTAAACAGGCCTA 120
          *****

                                     ↓
DSS      K A L M T V G K E P F P T I Y V D S Q K
Lewis    AAGGCTTTGATGACAGTTGGAAAAGAGCCGTTTCCACCATTATGTGGATTACAGAGAG 180
          AAGGCTTTGATGACAGTTGGAAAAGAGCCGTTTCCACCATTATGTGGATTACAGAGAG 180
          *****

                                     ↓
DSS      E N E R W D V I F K S Q M K N I K K M W
Lewis    GAAAATGAGAGGTGGGACGTTATTTTAAAGTCACAGATGAAGAACATCAAAAAGATGTGG 240
          GAAAATGAGAGGTGGGACGTTATTTTAAAGTCACAGATGAAGAACATCAAAAAGATGTGG 240
          *****

                                     ↓
DSS      H R E Q M K S D S R E K K E A E D N L R
Lewis    CACCAGAACAGATGAAGAGTGATTCCCGGGAGAAAGAGGCAGAGATAACTTACGA 300
          CACCAGAACAGATGAAGAGTGATTCCCGGGAGAAAGAGGCAGAGATAACTTACGA 300
          *****

DSS      R E K N L E E A K K I I I E N D P S L F
Lewis    AGAGAAAAGAACCTGGAGGAAGCAAAAAAATTATCATAAAAACGACCCGAGCCTGGCC 360
          AGAGAAAAGAACCTGGAGGAAGCAAAAAAATTATCATAAAAACGACCCGAGCCTGGCC 360
          *****

                                     ↓
DSS      E P A C V K I S A L E G Y R G Q R V K V
Lewis    GAGCCAGCATGTGTAAGATTAGTGCATTAGAAGGATACAGAGCCAGAGAGTGAAGGTG 420
          GAGCCAGCATGTGTAAGATTAGTGCATTAGAAGGATACAGAGCCAGAGAGTGAAGGTG 420
          *****

                                     ↓
DSS      F G W V H R L R R Q G K H L M F L V L R
Lewis    TTTGGCTGGGTCCACAGSTTACGAAGGCAAGGAAAGAATCTGATGTTTTGGTGTGGGA 480
          TTTGGCTGGGTCCACAGSTTACGAAGGCAAGGAAAGAATCTGATGTTTTGGTGTGGGA 480
          *****

                                     ↓
DSS      D G T G Y L Q C V L S D D L C Q C Y N G
Lewis    GATGGTACGGGTTATCTTCAGTGTGCTTGTGTCAGATGACTTGTGTGAGTGTACAAATGA 540
          GATGGTACGGGTTATCTTCAGTGTGCTTGTGTCAGATGACTTGTGTGAGTGTACAAATGA 540
          *****

DSS      V V L S T E S E V A V Y G T L N L T P K
Lewis    GTAGTCTTGTCCACTGAGAGTAGCGTGGCGGTATACGGGACACTGAACCTAACTCCAAA 600
          GTAGTCTTGTCCACTGAGAGTAGCGTGGCGGTATACGGGACACTGAACCTAACTCCAAA 600
          *****

                                     ↓
DSS      G K Q A F G G R E L S C D F W E L V G L
Lewis    GGCAAACAGGCTCCGGGAGGCCATGAGCTGAGCTGTGACTTCTGGGAACCTGGTGGGGCTG 660
          GGCAAACAGGCTCCGGGAGGCCATGAGCTGAGCTGTGACTTCTGGGAACCTGGTGGGGCTG 660
          *****

DSS      A P A G G A D N L I N E E S D V D V Q L
Lewis    GCCCCAGCTGGAGGACTGACAACCTGATCAATGAGGAGTCTGATGTGGACGTCACAGCTC 720
          GCCCCAGCTGGAGGACTGACAACCTGATCAATGAGGAGTCTGATGTGGACGTCACAGCTC 720
          *****

DSS      N M R H M I R G E N M S K I L K A R S
Lewis    AACCAACCGGCATATGATGATCCGGGGAGAGAACATGTCCAAAATCTGAAAAGCACGGTCC 780
          AACCAACCGGCATATGATGATCCGGGGAGAGAACATGTCCAAAATCTGAAAAGCACGGTCC 780
          *****

```

```

M I T R C F R D H F F D R G Y C E V T T
DSS ATGATCACCAGGTGCTTTAGGGACCACTCTTCGACAGAGGCTACTGTGAAGTAAACCACT 840
Lewis ATGATCACCAGGTGCTTTAGGGACCACTCTTCGACAGAGGCTACTGTGAAGTAAACCACT 840
*****
F T L V Q T Q V E G G A T L F K L D Y F
DSS CCAACACTGGTGCAGACACAGGTGGAAGGTGGGGCCACACTCTTCAAGCTTGACTATTTT 900
Lewis CCAACACTGGTGCAGACACAGGTGGAAGGTGGGGCCACACTCTTCAAGCTTGACTATTTT 900
*****
G E E A F L T Q S S Q L Y L E T C L P A
DSS GGGGAAGAABCAATTTTGAACCCAGTCCCTCACAGCTGTACCTGGAGACCTGCCTTCCAGCC 960
Lewis GGGGAAGAABCAATTTTGAACCCAGTCCCTCACAGCTGTACCTGGAGACCTGCCTTCCAGCC 960
*****
L G D V F C I A Q S Y R A E Q S R T R R
DSS CTGGGAGATGTTTTTGCATAGCCAGTCTTACAGGGCTSAACAGTCCAGGACACGAAGG 1020
Lewis CTGGGAGATGTTTTTGCATAGCCAGTCTTACAGGGCTSAACAGTCCAGGACACGAAGG 1020
*****
H L A E F T H V E A E C P F L T F E D L
DSS CATCTGGCTGAGTTCACCTCACGTGGAAGCCGAGTGTCTTTTCCCTGACCTTCGAGGACCTC 1080
Lewis CATCTGGCTGAGTTCACCTCACGTGGAAGCCGAGTGTCTTTTCCCTGACCTTCGAGGACCTC 1080
*****
L S R L E D L V C D V V D R V L K S F V
DSS CTGAGCCGTCTAGAGGACCTGGTGTGTGACGTGGTGGACAGAGTCTTGAAGTCAACAGTG 1140
Lewis CTGAGCCGTCTAGAGGACCTGGTGTGTGACGTGGTGGACAGAGTCTTGAAGTCAACAGTG 1140
*****
A S I V Y D L N E N F K A P K R P F R R
DSS GCGAGCATAGTGTATGACCTTAACCCGAACCTTAAAGCCCCAAAACGGCCTTTCCGGGGG 1200
Lewis GCGAGCATAGTGTATGACCTTAACCCGAACCTTAAAGCCCCAAAACGGCCTTTCCGGGGG 1200
*****
M N Y S D A I E W L R E H D V K K E D G
DSS ATGAACATTCAGATGCCATTTGAGTGGCTGAGGGAGCAGGATGTAAGAAAGAAGACGGG 1260
Lewis ATGAACATTCAGATGCCATTTGAGTGGCTGAGGGAGCAGGATGTAAGAAAGAAGACGGG 1260
*****
T F Y E F G D D I F E A P E R L M T D T
DSS ACSTTCTACGAGTTCGGAGACGATATCCCGAAGCGCCTGAGAGACTGATGACAGACACC 1320
Lewis ACSTTCTACGAGTTCGGAGACGATATCCCGAAGCGCCTGAGAGACTGATGACAGACACC 1320
*****
I N E E I L L C R F P V E I E S F Y M Q
DSS ATTAATGAACCAATCCTGCTGTGTGGTTTTCTGTGGAGATCAAGTCTCTACATGCAG 1380
Lewis ATTAATGAACCAATCCTGCTGTGTGGTTTTCTGTGGAGATCAAGTCTCTACATGCAG 1380
*****
R C P E D P R L T E S V D V L M P N V G
DSS CGCTGTCTGAGGATCCTCGGCTTACCGAATCTGTGGATGTGTGATGCCCAATGTGCGT 1440
Lewis CGCTGTCTGAGGATCCTCGGCTTACCGAATCTGTGGATGTGTGATGCCCAATGTGCGT 1440
*****
E I V G G S M R S W D S E E I L E G Y K
DSS GAGATTGSGGAGGCTCAATGCGCTCCTGGSACAGTGAGGAGATTCTAGAAGGCTATAAA 1500
Lewis GAGATTGSGGAGGCTCAATGCGCTCCTGGSACAGTGAGGAGATTCTAGAAGGCTATAAA 1500
*****
R E G I D P A P Y Y W Y T D Q R K Y G T
DSS AGGGAAGGATGACCCCGCTCCTTATTACTGGTATACAGATCAGAGAAAATATGGTACA 1560
Lewis AGGGAAGGATGACCCCGCTCCTTATTACTGGTATACAGATCAGAGAAAATATGGTACA 1560
*****
C P H G G Y G L G L E R P L S W I L N R
DSS TGTCTCACGGAGGTATGGCTTGGGCTTGGAAACGATTTCTAAGCTGGATTCTGAACAGG 1620
Lewis TGTCTCACGGAGGTATGGCTTGGGCTTGGAAACGATTTCTAAGCTGGATTCTGAACAGG 1620
*****

```



```

Y H I R D V C L Y P R F L Q R C R P -
DSS:  TATCACATCCGAGACGTGTGCTTGTACCCGAGATTTCTCCAGCGCTGCAGGCCATAA 1677
Lewis: TATCACATCCGAGACGTGTGCTTGTACCCGAGATTTCTCCAGCGCTGCAGGCCATAA 1677
*****

```

Footnote to supplement 16: * denotes nucleotide identity. The amino acid sequence is given on top. Genomic and cDNAs from the brain of DSS and Lewis were first amplified by PCR and then sequenced. ↓ indicates the last nucleotide of each exon. Intron-exon junctions are the same (data not shown).

Supplement 17: Sequence alignment of the microRNA mlr-122a between Dahl salt-sensitive (DSS) and Lewis rats

```

DSS      CCTTAGCAGAGCTCTGGAGTGTGACAATGGTGTGGTCCAAAACATCAAACGCCATCA 60
Lewis    CCTTAGCAGAGCTCTGGAGTGTGACAATGGTGTGGTCCAAAACATCAAACGCCATCA 60
*****

DSS      TCACACTAAACAGCTACTGCTAGGC 85
Lewis    TCACACTAAACAGCTACTGCTAGGC 85
*****

```

Footnote to Supplement 17: * indicates nucleotide identity. Genomic DNA from DSS and Lewis were first amplified by PCR and then sequenced.

Supplement 18: Coding sequence alignment of the solute carrier family 26 (sulfate transporter), member 2 (Slc26a2) between Dahl salt-sensitive (DSS) and Lewis rats

```

M S E E N K E Q H N L E P R D L P E E A 30
Slc26a2-DSS ATGTCTTCAGAAAATAAAGAGCAGCATAATCTTTCACCAAGGGACTTACCTGAAGAAGCC 60
Slc26a2-Lewis ATGTCTTCAGAAAATAAAGAGCAGCATAATCTTTCACCAAGGGACTTACCTGAAGAAGCC 60
*****
Y G F P F F E L F L G A Q R G S S T D L R 40
Slc26a2-DSS TATGGCTTCCACCTGAGCTCCCTGGGGCTCAAAGAGSATCAAGCACTGACTAAGS 120
Slc26a2-Lewis TATGGCTTCCACCTGAGCTCCCTGGGGCTCAAAGAGSATCAAGCACTGACTAAGS 120
*****
Q F E F S D R R E R A F R R I H M E L H E 60
Slc26a2-DSS CAGTTTGAAGCCAGTGTATCGGAGAAGAGCTTTTCGTAGGATCCACATGGAACTCCATGAG 180
Slc26a2-Lewis CAGTTTGAAGCCAGTGTATCGGAGAAGAGCTTTTCGTAGGATCCACATGGAACTCCATGAG 180
*****
K F D T N I R Q L V M R K L Q K E C Q C 80
Slc26a2-DSS AAACCACTACTAACATCAGACAGCTTGTCTAGAGAAGCTTCAGAAGAGTGGCCAGTGT 240
Slc26a2-Lewis AAACCACTACTAACATCAGACAGCTTGTCTAGAGAAGCTTCAGAAGAGTGGCCAGTGT 240
*****
N A T E I R N R I F D F F P V L R M L P 100
Slc26a2-DSS AATGCAACCAAAATCAGAAATAGGATTTTGGATTTTTCCTGTTTTGAGGTGGCTCCCA 300
Slc26a2-Lewis AATGCAACCAAAATCAGAAATAGGATTTTGGATTTTTCCTGTTTTGAGGTGGCTCCCA 300
*****
K Y D L K K N I L G D M H S B L I V G I 120
Slc26a2-DSS AAATATGATCTGAAGAAAACATTTAGGTGACATGATGCTGGGCTGATTTGGGTATA 360
Slc26a2-Lewis AAATATGATCTGAAGAAAACATTTAGGTGACATGATGCTGGGCTGATTTGGGTATA 360
*****
L L V P Q S I A Y S L L A G Q E P I Y G 140
Slc26a2-DSS TTGTTGGTGGCCAGTCTATTGCTTACTCCCTGTTGGCTGGCCAGAACCTATCTATGGT 420
Slc26a2-Lewis TTGTTGGTGGCCAGTCTATTGCTTACTCCCTGTTGGCTGGCCAGAACCTATCTATGGT 420
*****
L Y T S F F A S I I Y F L F G T S R H I 160
Slc26a2-DSS CTGTACACATCATTTTTGGCAGCATTATTACTTCTGTTGGTACCTCCGCCACATC 480
Slc26a2-Lewis CTGTACACATCATTTTTGGCAGCATTATTACTTCTGTTGGTACCTCCGCCACATC 480
*****
S V G I F G I L C L M I G E V V D S E L 180
Slc26a2-DSS TCTGGGGCATTTTTGGAACTGTGGCTTATGATTGGTGAGGTAGTTGACCGAGAAGCTA 540
Slc26a2-Lewis TCTGGGGCATTTTTGGAACTGTGGCTTATGATTGGTGAGGTAGTTGACCGAGAAGCTA 540
*****
H N A C F D I D T T S S I A M F S N G 200
Slc26a2-DSS CATAAAGCTGGCCGACATGATACTACATCATCTTCCATAGCAATGTTTCAAATGGA 600
Slc26a2-Lewis CATAAAGCTGGCCGACATGATACTACATCATCTTCCATAGCAATGTTTCAAATGGA 600
*****
C V V V N H T I D G L C D K S C Y A I E 220
Slc26a2-DSS TGTGTGTGTAAACCATACATTAGACGGACTCTGTGACAAAAGCTGTATGCAATTAAG 660
Slc26a2-Lewis TGTGTGTGTAAACCATACATTAGACGGACTCTGTGACAAAAGCTGTATGCAATTAAG 660
*****
I G S E T V T F N A G V Y Q V A M G F F Q 240
Slc26a2-DSS ATTGGCAGCACTGTGACATTCATGGCTGGAGITTTATCAGGTAGCCATGGGCTTCTTTCAA 720
Slc26a2-Lewis ATTGGCAGCACTGTGACATTCATGGCTGGAGITTTATCAGGTAGCCATGGGCTTCTTTCAA 720
*****
V G F V S V Y L S D A L L S G F V T G A 260
Slc26a2-DSS GTGGGCTTGTGTCTGTCTACTCTCAGATGCTTGTGTGAGCGGTTTGTCACTGGTGCC 780
Slc26a2-Lewis GTGGGCTTGTGTCTGTCTACTCTCAGATGCTTGTGTGAGCGGTTTGTCACTGGTGCC 780
*****
E F T I L T S Q A K Y L L G L S L P R S 280
Slc26a2-DSS TCCTTCACCATCTCTCAGCTCTCAGGCCAAGTACCTCTGGGGCTGAGCCTTCTCTGGAGC 840
Slc26a2-Lewis TCCTTCACCATCTCTCAGCTCTCAGGCCAAGTACCTCTGGGGCTGAGCCTTCTCTGGAGC 840
*****
N G V G S V I T T M I H I F R N T H K T 300
Slc26a2-DSS AATGGTGTAGGCTCAGTCACTACTACCTGGATCCACATCTTCAGAAAATATTCATAAGACC 900
Slc26a2-Lewis AATGGTGTAGGCTCAGTCACTACTACCTGGATCCACATCTTCAGAAAATATTCATAAGACC 900
*****
H I C D L I T S L L C L I V L V P T K E 320
Slc26a2-DSS AACATCTGTGACCTCATCACCAGCCTTTTGTGTCTCTGTTCTTGTGTCACCAACCAAGAG 960
Slc26a2-Lewis AACATCTGTGACCTCATCACCAGCCTTTTGTGTCTCTGTTCTTGTGTCACCAACCAAGAG 960
*****

```

```

L H E Y F K S K L F A F I P T E L I V V 340
Slc26a2-DSS CTCAACGAATACTTCAAGTCCAAAGCTCCCGCACCAATCCAACTGAGCTCATTGTGGTT 1020
Slc26a2-Lewis CTCAACGAATACTTCAAGTCCAAAGCTCCCGCACCAATCCAACTGAGCTCATTGTGGTT 1020
*****
V A A T L A S S F S K L N E N Y H S S I 360
Slc26a2-DSS GTGGCAGCCACATTGGCTTCTCATTTTGGAAAACCAATGAGAATTACAATTCAGTATT 1080
Slc26a2-Lewis GTGGCAGCCACATTGGCTTCTCATTTTGGAAAACCAATGAGAATTACAATTCAGTATT 1080
*****
A G G I F T G F M F P Q A P D W S L I P 380
Slc26a2-DSS GCCGGGCAAAATCCCAACGGGTTTATGCCACCCCAAGCGCCAGACTGGAGCCTCATTCTT 1140
Slc26a2-Lewis GCCGGGCAAAATCCCAACGGGTTTATGCCACCCCAAGCGCCAGACTGGAGCCTCATTCTT 1140
*****
M V A V D A I A I S I I G F A I T V S L 400
Slc26a2-DSS AATGTGGCTGTGTAGCCATAGCTATTTCTATCATTGGTTTTGTATCACTGTATCACTT 1200
Slc26a2-Lewis AATGTGGCTGTGTAGCCATAGCTATTTCTATCATTGGTTTTGTATCACTGTATCACTT 1200
*****
E E H F A K K H G Y T V K A N Q E H Y A 420
Slc26a2-DSS TCTGAGATGTTTGGCAAGAAACATGGCTACAGGTTGAAAGCCAAATCAGGAAATGTATGCC 1260
Slc26a2-Lewis TCTGAGATGTTTGGCAAGAAACATGGCTACAGGTTGAAAGCCAAATCAGGAAATGTATGCC 1260
*****
I G F C H I I P S F F H C I T T E A A L 440
Slc26a2-DSS ATTGGCTTTTGGCAATATCATACCTCTCTTCCACTGCATAACTACTAGTCCGCGCTT 1320
Slc26a2-Lewis ATTGGCTTTTGGCAATATCATACCTCTCTTCCACTGCATAACTACTAGTCCGCGCTT 1320
*****
A K T L V K E E T G C Q T Q L S A I V T 460
Slc26a2-DSS GCAAAGACACTGGTTAAAGAGTCAACAGGTTGCCAGACCCCACTGTCCGCTATAGTACA 1380
Slc26a2-Lewis GCAAAGACACTGGTTAAAGAGTCAACAGGTTGCCAGACCCCACTGTCCGCTATAGTACA 1380
*****
E L V L L L V L L L I A P L F Y E L Q K 480
Slc26a2-DSS TCCTGTCTCTCTGTGTGGTCTCTGCTAATAGTCCCTTATTCTATTCCCTCCAAAAA 1440
Slc26a2-Lewis TCCTGTCTCTCTGTGTGGTCTCTGCTAATAGTCCCTTATTCTATTCCCTCCAAAAA 1440
*****
C V L G V I T I V N L R G A L L E F R D 500
Slc26a2-DSS TGTGTCTTGGTGTGATCACCATTGTAAATCTCCGGGTTGCACTTCTGAAAATTAGAGAC 1500
Slc26a2-Lewis TGTGTCTTGGTGTGATCACCATTGTAAATCTCCGGGTTGCACTTCTGAAAATTAGAGAC 1500
*****
L P K M N R L S B M D T V I N F V T M L 520
Slc26a2-DSS CTGCCAAAGATGGGAGGCTCAGCAGAAATGGACACAGTATCTGGTTTGTGACGATGCTG 1560
Slc26a2-Lewis CTGCCAAAGATGGGAGGCTCAGCAGAAATGGACACAGTATCTGGTTTGTGACGATGCTG 1560
*****
E S A L L L S T E I B L L V G V C F S N F 540
Slc26a2-DSS TCCTCTGCTCTGTAAAGCACTGAAATAGGCTGCTGCTGGGTTTGTTTTCAATGTTT 1620
Slc26a2-Lewis TCCTCTGCTCTGTAAAGCACTGAAATAGGCTGCTGCTGGGTTTGTTTTCAATGTTT 1620
*****
C V I L R T Q M P K I S L L G L E E S S 560
Slc26a2-DSS TGTGTGATTCTCCGTACTCAGATGCCAAGATTTCACTGCTTGGTTTGGAAAGAAGATCT 1680
Slc26a2-Lewis TGTGTGATTCTCCGTACTCAGATGCCAAGATTTCACTGCTTGGTTTGGAAAGAAGATCT 1680
*****
E I F K S I S T Y E N L R S R S G I K V 580
Slc26a2-DSS GAAATCTTCGAGTCCATCTCCACTTACAAGAACCTTCGGAGTAAGTCAGGCATCAAGGTT 1740
Slc26a2-Lewis GAAATCTTCGAGTCCATCTCCACTTACAAGAACCTTCGGAGTAAGTCAGGCATCAAGGTT 1740
*****
F R F I A P L Y Y I N K E C F K S A L Y 600
Slc26a2-DSS TTCGGCTCATAGCCCTCTCTACTACATAAACAAAGAATGCTTTAAATCAGCTTTGTAC 1800
Slc26a2-Lewis TTCGGCTCATAGCCCTCTCTACTACATAAACAAAGAATGCTTTAAATCAGCTTTGTAC 1800
*****
E E T L N P V L V K A A W K E R A A K E R 620
Slc26a2-DSS AAGAAAACGCTAAACCCAGTCTTGGTAAAGCAGCCTGGAAAAAGGCAAGAAAAGSAAA 1860
Slc26a2-Lewis AAGAAAACGCTAAACCCAGTCTTGGTAAAGCAGCCTGGAAAAAGGCAAGAAAAGSAAA 1860
*****
L K E E T V Y F R G D P D E V S M Q L S 630
Slc26a2-DSS CTGAAAGAGGAAACAGTGACTTTCCATGGGGACCCGGATGAAGTCTCAATGCAGCTTCC 1920
Slc26a2-Lewis CTGAAAGAGGAAACAGTGACTTTCCATGGGGACCCGGATGAAGTCTCAATGCAGCTTCC 1920
*****
H D E L E L H T V V I D C S A I Q F L D 660
Slc26a2-DSS CATGATCCCTTGGAGCTGCACACCTTGTGATGACTGTAGTCAATACAGTTTTGGAT 1980
Slc26a2-Lewis CATGATCCCTTGGAGCTGCACACCTTGTGATGACTGTAGTCAATACAGTTTTGGAT 1980
*****

```

```

          T A G I H T L E E V R R E Y E A I G I Q 680
Slc26a2-DSS ACAGCAGGGATCCACACACTGAAAGAAGTTCCGCCGGATTATGAAGCCATTGGTATCCAG 2040
Slc26a2-Lewis ACAGCAGGGATCCACACACTGAAAGAAGTTCCGCCGGATTATGAAGCCATTGGTATCCAG 2040
          *
          V L L A Q C H F S V R D S L A K G E Y C 700
Slc26a2-DSS GTTTTACTGGCTCAGTSCAATCCTTCTGTGAGGGATTCCCTTAGCCAAAGGAGAATATTGC 2100
Slc26a2-Lewis GTTTTACTGGCTCAGTSCAATCCTTCTGTGAGGGATTCCCTTAGCCAAAGGAGAATATTGC 2100
          *
          K K E E E N L L F Y S L S E A V A F A E 720
Slc26a2-DSS AAAAAGSAAAGAAAATCTTCTCTTCTACAGTTTGCTGAAGCAGTAGCTTTTGCAGAA 2160
Slc26a2-Lewis AAAAAGSAAAGAAAATCTTCTCTTCTACAGTTTGCTGAAGCAGTAGCTTTTGCAGAA 2160
          *
          E S Q K E K G V C V V H G L S L E G D - 739
Slc26a2-DSS GAGTCTCAGAAGGAGAAAGGAGTGTGTGTGTCAATGGTTTGAGTCTTCTGGTGACTGA 2220
Slc26a2-Lewis GAGTCTCAGAAGGAGAAAGGAGTGTGTGTGTCAATGGTTTGAGTCTTCTGGTGACTGA 2220
          *

```

Footnote to Supplement 18: * indicates nucleotide identity. Amino acid sequence is given on top. Genomic DNAs from DSS and Lewis were first amplified by PCR and then sequenced since the *Slc26a2* coding region resides within two exons. ↓ indicates the end of exon 1.

Supplement 19. Coding sequence alignment for casein kinase 1, alpha 1 (*Csnk1a1*) between Dahl salt-sensitive (DSS) and Lewis rats

```

DSS      N A S S S G S S K A E F I V G G K Y K L V
Lewis    ATGGCGAGCAGCAGCGGCTCCAAGGCCGAATTTATCGTCGGTGGGAAATACAAACTGGTG 60
*****
R K I G S G S F G D I Y L A I N I T N G
DSS      CGGAAGATCGGATCTGGTTCCTTCGGGGACATTTATCTAGCGATCAACATCACCAATGGC 120
Lewis    CGGAAGATCGGATCTGGTTCCTTCGGGGACATTTATCTAGCGATCAACATCACCAATGGC 120
*****
↓
DSS      E E V A V K L E S Q K A R H P Q L L Y E
Lewis    GAGGAAGTGGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
GAGGAAGTGGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
*****
S K L Y E I L Q G G V G I P H I R W Y G
DSS      AGCAAACGTATAAGATTCTTCAAGGTGGGGTTGGCATCCCTCACATACGGTGGTATGGG 240
Lewis    AGCAAACGTATAAGATTCTTCAAGGTGGGGTTGGCATCCCTCACATACGGTGGTATGGG 240
*****
Q E K D Y N V L V M D L L G P S L E D L
DSS      CAAGAAAAGACTATAATGTGCTAGTCATGGATCTTCTGGGACCCAGCCTCGAAGACCTC 300
Lewis    CAAGAAAAGACTATAATGTGCTAGTCATGGATCTTCTGGGACCCAGCCTCGAAGACCTC 300
*****
↓
DSS      F N F C S R R F T M K T V L M L A D Q M
Lewis    TTCAAATTCCTGTTCAAGAAGGTTCACTATGAAGACTGTACTTATGTTAGCTGACCAGATG 360
TTCAAATTCCTGTTCAAGAAGGTTCACTATGAAGACTGTACTTATGTTAGCTGACCAGATG 360
*****
I S R I E Y V H T E N F I H R D I K P D
DSS      ATCAGTAGAATTGAATATGTGCATACAAAGAATTTTATACACAGAGACATTAACAGAT 420
Lewis    ATCAGTAGAATTGAATATGTGCATACAAAGAATTTTATACACAGAGACATTAACAGAT 420
*****
↓
DSS      N F L M G I G R H C N K L F L I D F G L
Lewis    AACTTCCTAATGGGTATTGGGGTCACTGTAATAAGTTATTCCTTATTSATTTTGGTTTG 480
AACTTCCTAATGGGTATTGGGGTCACTGTAATAAGTTATTCCTTATTSATTTTGGTTTG 480
*****
A K K Y R D N R T R Q H I P Y R E D K N
DSS      GCCAAAAAGTACAGAGACAACAGGACAAGGCAACACATACCATAACAGGSAAGATAAAAAAC 540
Lewis    GCCAAAAAGTACAGAGACAACAGGACAAGGCAACACATACCATAACAGGSAAGATAAAAAAC 540
*****
↓
DSS      L T G T A R Y A S I N A H L G I E Q S R
Lewis    CTCCTGGCACTGCCCCGATGCCCAGCATCAATGCACATCTTGGCATTGAGCAGAGTCCG 600
CTCCTGGCACTGCCCCGATGCCCAGCATCAATGCACATCTTGGCATTGAGCAGAGTCCG 600
*****
R D D M E S L G Y V I M Y P N R T S L P
DSS      CGAGATGACATGGAATCTTTAGGATATGTTTTGATGATTTTAAATAGAACCAGTCTGCCA 660
Lewis    CGAGATGACATGGAATCTTTAGGATATGTTTTGATGATTTTAAATAGAACCAGTCTGCCA 660
*****
↓
DSS      W Q G L K A A T E E Q K Y E K I S E E K K
Lewis    TGSCAAGGGCTAAAGSCTGCAACAAGAAACAAAAATATGAAAAGATTAGTGAAGAAG 720
TGSCAAGGGCTAAAGSCTGCAACAAGAAACAAAAATATGAAAAGATTAGTGAAGAAG 720
*****

```

```

                                ↓
N S T P V E V L C K G F P A E F A M Y L
DSS ATGTCCACTCCTGTTGAAGTTTTGTGTAAGGGGTTTCCTGCAGAAATTTGCCATGACTTA 780
Lewis ATGTCCACTCCTGTTGAAGTTTTGTGTAAGGGGTTTCCTGCAGAAATTTGCCATGACTTA 780
*****
N Y C R G L R F E E A F D Y M Y L R Q L
DSS AATTACTGTCGTGGGCTGCGCTTTGAGGAAGCTCCGGATTACATGTATCTGAGGCAGCTG 840
Lewis AATTACTGTCGTGGGCTGCGCTTTGAGGAAGCTCCGGATTACATGTATCTGAGGCAGCTG 840
*****
                                ↓
F R I L F R T L N H Q Y D Y T F D W T M
DSS TTCCGCATCCTCTTCAGSACCCCTGAACCACCAGTATGACTACACGTTTGATTGGACGATG 900
Lewis TTCCGCATCCTCTTCAGSACCCCTGAACCACCAGTATGACTACACGTTTGATTGGACGATG 900
*****
L K Q K A A Q Q A A S S S G Q G Q Q A Q
DSS TTAAGCAGAAAAGCAGCCAGCAGGCAGCCTCTTCCAGTGGGCAGGGTCAGCAGGCCCAA 960
Lewis TTAAGCAGAAAAGCAGCCAGCAGGCAGCCTCTTCCAGTGGGCAGGGTCAGCAGGCCCAA 960
*****
T P T G E -
DSS ACCCCACAGGTTTCTAA 978
Lewis ACCCCACAGGTTTCTAA 978
*****

```

Footnote to Supplement : Amino acid sequence is given on top. Brain cDNAs from DSS and Lewis were first amplified by PCR and then sequenced. ↓ Indicates the last nucleotide of each exon.

10.10 Addendum

Errata :

i) La longueur de l'intervalle chromosomique contenant le C18QTL3 est de 7,35 Mb et contient 67 gènes. Ces données auraient dû être spécifiées dans le résumé

ii) On entend par « variant » tout changement dans la séquence génomique et ce indépendamment de son effet. Le terme « mutation » désigne une variante à effet pathologique et donc causant une maladie [154]. Le terme « *mutation* » n'a pas été correctement employé au long de cette publication, étant donné que l'impact des variantes décrites sur les fonctions des protéines et donc sur le trait étudié n'a pas été analysé en détail.

iii) La nomenclature des variantes recommandée par l'HGVS est la suivante [154]:

C18QTL3

NM_001008300.2(Nedd4l):c.321A>T

NM_001008300.2(Nedd4l):c.1995T>C

NM_001008300.2(Nedd4l):c.2139G>A

XM_006254887.2(Alpk2):c.376C>T, p.(126Arg>Cys)

XM_006254887.2(Alpk2):c.475C>T, p.(159Pro>Ser)

XM_006254887.2(Alpk2):c.618C>T

XM_006254887.2(Alpk2):c.1910_1912delCCG, p.(637delAla)

XM_006254887.2(Alpk2):c.2918C>T, p.(973Gln>Leu)

XM_006254887.2(Alpk2) c.3754G>A, p.(1252Glu>Lys)

XM_006254887.2(Alpk2):c.5610C>G

XM_006254887.2(Alpk2) c.6217A>G, p.(2073Ile>Val)

NM_133570.5(Grp):c.283G>T,p.(95Ala>Ser)

NM_133570.5(Grp):c.286G>A, p.(96Arg>Gly)

11. Two candidate genes for 2 quantitative trait loci epistatically attenuate hypertension in a novel pathway

Authors: Cristina Chauvet, Annie Ménard, Alan Y. Deng

Institution: Research Centre, CRCHUM (Centre hospitalier de l'Université de Montréal),
Department of Medicine, Université de Montréal, Montréal, Québec, Canada

Running title: Genetics of polygenic hypertension

Grant supports: Canadian Institutes for Health Research (CIHR) to AYD and a doctoral
fellowship to CC from the Canadian Heart and Stroke foundation

Conflict of interest: None

Corresponding author: Alan Y. Deng

Word count: 7205

11.1 Abstract

Objectives: Multiple quantitative trait loci (QTLs) for blood pressure (BP) have been detected in rat models of human polygenic hypertension. They influence BP physiologically via epistatic modules. Little is known about the causal genes and virtually nothing is known on modularized mechanisms governing their regulatory connections.

Methods and Results: Two genes responsible for 2 individual BP QTLs on rat Chromosome 18 have been identified that belong to the same epistatic module. Treacher Collins-Franceschetti syndrome 1 (Tcof1) gene is the only function candidate for C18QTL3. Haloacid dehalogenase-like hydrolase domain containing 2 (Hdhd2), although a gene of previously-unknown function, is C18QTL4, and encodes a newly-identified phosphatase. The current work has provided the premier evidence that Hdhd2/C18QTL4 and Tcof1/C18QTL3 may be involved in polygenic hypertension. Hdhd2/C18QTL4 can regulate the function of Tcof1/C18QTL3 via de-phosphorylation, and, for the first time, furbishes a molecular mechanism in support of a genetically-epistatic hierarchy between 2 BP QTLs, and thus authenticates the epistasis-common pathway paradigm.

Conclusions: The pathway initiated by Hdhd2/C18QTL4 upstream of Tcof1/C18QTL3 reveals novel mechanistic insights into BP modulations. Their discovery might yield innovative therapeutic targets and diagnostic tools predicated on a novel BP etiology and mechanism that is determined by a regulatory hierarchy. Optimizing the de-phosphorylation capability and its down-stream target could be anti-hypertensive. The conceptual paradigm of an order and regulatory hierarchy may help unravel genetic and molecular relationships among certain human BP QTLs.

Key words: Genetics of hypertension, QTLs, epistasis, congenic strains, dephosphorylation, posttranslational modification

11.2 Introduction

Uncovering the underlying mechanisms that transduce the sequela of BP variations can facilitate the development of etiology-based and mechanism-derived therapeutic targets. Localization of QTLs has disclosed numerous chromosome regions as well as gene markers that are functionally correlated with BP in both humans [1] and animal models [2,3]. The genetic architecture of BP is composed of numerous QTLs as building blocks. Their overall impact on BP is determined by modularized function cores/pathways that are arranged in epistatic hierarchies [4]. Since there are far fewer pathways than the components composing them [4], deciphering the regulatory relationships among these components in each pathway becomes critical in understanding the regulatory sequence determining the functionality of genetic architecture for BP [5]. Identifying gene variants causal to BP variations is prerequisite in understanding the molecular mechanisms underlying these relationships.

Several BP QTLs closely-linked on Chromosome (Chr) 18 were previously defined by congenic strains cross-bred from hypertensive Dahl salt-sensitive (DSS) and Lewis rats [4,6]. Due to a lack of comprehensive genome sequence data, mutation screenings could not be thoroughly carried out for all the genes present in the QTL-residing intervals; and congenic coverages were left with large ambiguous gaps. In the current work, we aimed at identifying causal genes for certain QTLs on Chr 18 combining [7] congenic resolution as well as total genome sequencing, and elucidating molecular mechanisms on QTL actions and the regulatory relationships between them. We have identified function-altering genetic variants potentially responsible for QTL actions, despite the knockout lethality of individual genes involved.

11.3 Methods

11.3.1 Hypothesis and Study design

NB. Cette section se trouve sous la partie Results dans la publication original[155], elle a été déplacée à fin de rendre le texte plus homogène.

The basic principle [7] in gene discovery based on the genetics of BP QTLs is that, first, a gene candidate has to carry a function-impacting genetic variation between the 2 contrasting strains, e.g. DSS and Lewis rats. In search of such candidates for QTLs on Chr 18, we combined whole genome sequencing with congenic resolution always accompanied by a BP-effect. Second, once a candidate gene for each of 2 epistatically-interacting QTLs is identified, they can be analyzed to reveal potential molecular mechanisms underlying epistasis. The following sections detail these studies.

11.3.2 Animals

Protocols for handling, maintaining and treating animals have been approved by our institutional animal committee (CIPA). All the congenic strains used in the current study are depicted in Fig. 1 and are synthesized from our previous work [4,6,8].

11.3.2 Whole genome sequencing of DSS and Lewis rats

Supplement (Sup.) 1 outlines the methods, procedure, interpretation, sequence calling and data mining along with relative references. The genome sequences of DSS and Lewis became our data base (Sup. 2) for identifying single-nucleotide polymorphisms (SNPs) in genes in a given QTL-containing interval (Fig. 1). Bioinformatics software NextiaGen© developed by Max Chauvet was used to position the variants resulting from the SNP analysis within the mRNA or genomic sequence of each gene. All missense mutations were independently verified and are homozygous for the Lewis alleles, (i.e. LL) in the congenic strain that defines the QTL in question. Copy number variations can also be found.

11.3.3 In vitro function analysis of 2 isoforms for Haloacid dehalogenase-like hydrolase domain containing 2 (Hdhd2)

Cloning of full-length *Hdhd2*: Complete coding regions for the 2 *Hdhd2* isoforms were cloned (Fig. 3) separately following PCR of cDNAs from DSS and Lewis kidneys into the pcDNA™3.1(+) expression vector (Invitrogen) and transformed into DH5 α Competent Cells. The coding identity defining the DSS and Lewis *Hdhd2* isoforms, *Hdhd2*^{DSS}, and *Hdhd2*^{Lewis} was confirmed by sequencing.

Hdhd2 expressions: HEK293 cells were cultured by a standard protocol and then transfected with 4 μ g *Hdhd2*^{DSS}, *Hdhd2*^{Lewis} and GFP-pcDNA vector alone using the lipofectamine protocol (Invitrogen). 48hrs after transfection, cells were harvested following a saline wash. Whole cell lysates were collected by centrifugation. Hdhd2 antibodies [Santa Cruz Biotech, SCB, (C-18:sc-84837)], β -tubulin antibodies (Cellsignal #2146), and goat anti-rabbit IgG-HRP (SCB, c-2301) were used in western blotting according to manufacturer's instructions.

Phosphatase Assay: EnzChek Phosphatase Assay kit (E12020, Invitrogen) was used to perform assays following manufacturer's instructions. Briefly, 250ng of total protein extract was added to varying concentrations of 6,8-Difluoro-4-Methylumbelliferyl Phosphate (DiFMUP) substrate. Fluorescence measures were taken every 10 mins for a period of 120 min at 25°C using excitation at 360nm and emission at 460nm. The kinetic parameters for Hdhd2 as a phosphatase were determined using the same method as we reported previously on analyzing another enzyme [9], and compared between Hdhd2^{DSS} and Hdhd2^{Lewis}. Statistics are done using Systat [4].

Detection of de-phosphorylation of Tcof1 by Hdhd2: 75 μ g of whole cell proteins from Hdhd2^{Lewis}, Hdhd2^{DSS} cultured *in vitro*, or vector alone, were first extracted in the presence of a Sigma protease inhibitor cocktail and were then each incubated for 5 min at 37°C in presence or absence of 5 μ L of Sigma phosphatase inhibitor cocktail 2, followed by the addition of 20 μ g of Jurkat nuclear cell extract (SCB, SC-2132) as a substrate. The treated samples were then loaded onto a 7% SDS gel. After electrophoresis, proteins were transferred onto nitrocellulose membrane at 350mA for 2h using a standard transfer buffer. The

membrane was then washed in 20% methanol/H₂O before blocking with 5% milk/TBST and hybridized with the Treacle antibody (SCB, L-16: sc-49529) by the standard western blotting [10]. The *Tcofl* protein, also known as treacle, and its de-phosphorylation activity were detected based upon its cross reactivity with the SCB polyclonal anti-treacle antibodies. The epitope used to induce antibody productions came from the N-terminus of the human treacle. The designated antibodies cross hybridize with treacles from humans, mouse and rats, because the N-terminus of treacle from these species is highly conserved. Based on this information, we assume that the designated antibodies from the manufacturer are genuinely specific for the rat treacle.

11.4 Results

11.4.2 Restricting chromosome intervals harboring BP QTLs

Congenic strains isolate the genes functionally impacting on BP based on a cause-effect relationship and separate multiple closely-linked BP QTLs (Fig. 1). The genes contained in the intervals harboring *C18QTL3* and *C18QTL4* are listed in Sup. 3. The gap between *C18QTL3* and *C18QTL4* harbors 93 genes. The segment containing *C18QTL2* carries 249 genes and the region containing *C18QTL1* is not well delineated due to uncertainty of which genes are present. Because of these factors, further genetic studies were not pursued on the intervals harboring *C18QTL2* and *C18QTL1*.

11.4.3 *Treacher Collins-Franceschetti syndrome 1* gene (*Tcofl*) is a unique function candidate for *C18QTL3*

The congenic strain, C18S.L14, is different from DSS in the segment of 7.35 megabases (Fig. 1). This segment-‘knock in’ strain lowers ($p < 0.001$) mean arterial pressure (MAP) by 29 mmHg from that of DSS (Fig. 1), indicating that a gene(s) entrapped within is

responsible for *C18QTL3*. 7 genes are present in the *C18QTL3*-residing interval that carry SNPs in the coding regions and/or intron-exon junctions (Table 1).

18 consecutive base pairs (bp) of the *Tcofl* coding domain are deleted in DSS rats comparing to that of C18S.L14 (Sup. 4 and 5) and Lewis (data not shown), which removes 6 sequential amino acids from the conserved treacle domain [11] of the Tcofl protein. The protein product encoded by *Tcofl* is also known as treacle. This internal in-frame deletion results in a shorter and, probably, less efficient but functional, Tcofl product in DSS than that of Lewis (Table 1). DSS is homozygous for the 18-bp *Tcofl* deletion (Fig. 2). The remaining *Tcofl* coding domain is identical between DSS and Lewis alleles (Sup. 4). Thus, this 18-bp deletion alone distinguishes *Tcofl*^{DSS} from *Tcofl*^{Lewis}.

In contrast to *Tcofl*, all the genes located in the *C18QTL3*-residing interval encode identical proteins, except for *Alpha kinase 2 (Alpk2)* which we previously reported [6] and *caudal-type homeobox protein 1 (Cdx1)* (Table 1). *Synonymous SNPs found in the remaining genes (Table 1) are given in Sup. 6*. No DSS/C18S.L4 alternative splicing were detected (data not shown) for the genes carrying intron-exon boundary variants (Table 1). *Cdx1* is uniquely expressed in the intestine, not in the other organs (Fig. 2), consistent with its role in gastrointestinal development [12]. This restricted function and expression pattern down-graded *Cdx1* to be a biological candidate for *C18QTL3*.

Tcofl and *Alpk2* seem equally-valid genetic candidates for *C18QTL3* based on the presence of non-synonymous mutations. However, the predicted impact on the mutations distinguishes them functionally and supports *Tcofl*, not *Alpk2*. First, all the missense mutations in *Alpk2* are predicted to be benign (Table 2). The 3 bp-deletion removing an Ala637 residue in *Alpk2* (Table1) is not located in a predicted functional domain, unlike the 18 bp-deletion in *Tcofl* which is [11]. Second, although function predictions can only be made on missense mutations, not an in-frame deletion, we can gain some insights into the function impact from each deleted amino acid by simulating them as individual missense mutations. The *Tcofl* 18-bp deletion contains 2 important amino acids targeted for phosphorylation, Ser763 and Ser765. Simply mutating each of them would be deleterious (Sup. 7), let alone delete both of them. In contrast, mutating the Ala637 to any of 6 possible other amino acids would be benign. Therefore, *Tcofl* is functionally supported as *C18QTL3*, whereas *Alpk2* is

not. Consequently, *Tcofl* is the only gene that carries function-altering mutations and becomes the sole candidate for *C18QTL3*.

11.4.4 *Hdhd2* is a foremost candidate for *C18QTL4*

Among the mutation-carrying genes in the *C18QTL4*-residing interval (Table 3 and Fig. 1), *Hdhd2* became a leading candidate. It is ubiquitously expressed (Fig. 2, Sup. 5) and contains a single missense mutation at # G337C that causes a Val113Leu substitution (Table 3, Sup. 8).

In addition to *Hdhd2*, 4 other genes also harbor non-synonymous mutations, i.e. *Rbfa*, *Adnp2*, *Loxhd1*, and *Sall3* and, together, constitute the initial pool of genetic candidates for *C18QTL4* (Table 3). Examining expression patterns of these genes dwindled the list to 3, i.e. *Hdhd2*, *Rbfa* and *Adnp2*, because *Loxhd1* and *Sall3* are not detectable in 9 organs examined (Fig. 2). *Loxhd1* is principally expressed in *cochlear* and vestibular hair cells [13]. Missense mutations in *LOXHDI* cause human corneal dystrophy [14] and hearing loss [13]. *Sall3* encodes a zinc finger protein that is restrictively expressed in early development [15].

Synonymous SNPs found in the remaining genes in the *C18QTL4*-residing interval are summarized in Table 3 and their sequence alignments are given in Sup. 6. No alternative splicing of any genes in the *C18QTL4*-residing region has been detected (data not shown).

11.4.5 Function predictions of missense mutations

The single C337G mutation is persistently predicted to be damaging on the *Hdhd2* function, whereas all the missense mutations in *Rbfa* and *Adnp2* [16] are predicted to be benign (Table 2). Based on these results, *Hdhd2* appears the strongest and only function candidate for *C18QTL4*.

11.4.6 Function analysis of *Hdhd2*

First, although *Hdhd2* is an expressed gene (Fig. 2), no function information was previously known and the rat *Hdhd2* was only predicted from annotation. The single

Val113Leu change in Hdhd2, although not synonymous, is relatively-conservative, because both amino acids are aliphatic, non-polar, neutral and hydrophobic. Computer programs predict (Table 2) that this minor change in the size of the aliphatic and hydrophobic side chain is damaging to the Hdhd2 protein function. To validate this prediction, a functional assessment is indispensable. Second, the relationship between *C18QTL3* and *C18QTL4* is epistatic (Fig. 1), implying that they share a common pathway [4]. The function of *Tcof1/C18QTL3* is controlled, among others, by phosphorylation of its conserved treacle domain via a cascade catalyzed by casein kinase [17,18]. Epistasis between *C18QTL3* and *C18QTL4* (Fig. 1) predicts that a gene representing *C18QTL4* may participate in the same cascade of regulations on *Tcof1* including phosphorylation or de-phosphorylation. Hdhd2 might be a phosphatase, since Hdhd1 exhibits a phosphatase activity [19].

11.4.7 The Val113Leu substitution alters the phosphatase function of Hdhd2

We first cloned (Fig. 3a), and then analyzed phosphatase activity of DSS and Lewis Hdhd2 proteins expressed *in vitro* (Fig. 3b). K_m of Hdhd2^{DSS} is significantly higher than that of Hdhd2^{Lewis} (Fig. 3c), despite their V_{max} values are similar (Table 4). Thus, the singular Val113Leu change in Hdhd2^{DSS} decreases the binding affinity of substrate without changing the Hdhd2 activity. The result of enzymatic assays and computer predictions are consistent, and the G337C point mutation in *Hdhd2* indeed alters its function. Thus, *Hdhd2* as *C18QTL4* is functionally strengthened.

11.4.8 Hdhd2 dephosphorylates Tcof1

Based on their epistasis (Fig. 1) implicating the same pathway [4], we next examined whether or not Hdhd2 can modulate *Tcof1* by de-phosphorylation, a mode of post-translational modification. As shown in Fig. 3d, Hdhd2 indeed de-phosphorylates *Tcof1*/treacle *in vitro* as detected with treacle-specific antibodies by Western blotting (i.e. an increased mobility), and this de-phosphorylation was prevented in the presence of phosphatase inhibitors. According to a calculation, the mobility increase by Hdhd2-transformed extracts corresponds to a difference of approximately 15-30 kiloDaltons. No change in mobility was

seen when Jurkat cell extracts was incubated with cultured cells extracts transformed with the vector alone (data not shown).

11.5 Discussion

Major revelations from the current work are (a) 2 genes, *Tcof1* and *Hdhd2*, that are not previously known to participate in BP modulations, have been identified to be candidates for 2 QTLs. (b) *Hdhd2* encodes a newly-identified phosphatase. A single missense G337C mutation causing a Val113Leu substitution results in a change in the substrate binding affinity without affecting the enzymatic activity of Hdhd2. (c) The protein product of *C18QTL4/Hdhd2* regulates the protein product treacle encoded by *C18QTL3/Tcof1* post-translationally, and thus acts up stream of it in an epistatic hierarchy in a novel pathway towards the BP homeostasis. This is the first molecular evidence that 2 BP QTLs may function in a regulatory relationship as predicted by the epistasis/ common pathway hypothesis [4], and the mechanism underlying their epistasis may be a post-translational modification.

11.5.1 Elucidating the regulatory relationship in controlling BP between *Hdhd2* and *Tcof1* by synergizing epistasis and molecular mechanisms

Genetic epistasis (Fig. 1) implicates *Hdhd2* and *Tcof1* in the same pathway by not allowing BP to change accumulatively when combined. The molecular mechanism of Hdhd2 modulating Tcof1 by de-phosphorylation (above) further solidifies the order between them in the pathway with Hdhd2 upstream of Tcof1. Combining these data, a mechanistic pathway incorporating Hdhd2 and Tcof1 emerges that eventually leads to the BP modulation (Fig. 4).

Since only the chromosome segments replaced in C18S.L10, C18S.L14 and C18S.L2 congenic strains are different from DSS (Fig. 1), we can define DSS as the ‘wild type’ and its genes as in an ‘on’ state. Any change from DSS due to a Lewis genome replacement in a congenic strain is defined as a ‘mutant’ and its gene as in an ‘off’ state. When *C18QTL3/Tcof1* is ‘off’ as in C18S.L14 (Fig. 1), BP decreases (marked by a down-pointing arrow in Fig. 4) from the DSS ‘wild type’, indicating that the DSS ‘wild type’ *C18QTL3/Tcof1*

is involved in increasing BP. When the *C18QTL4/Hdhd2* is ‘off’ as in C18S.L10, BP decreased also, indicating that its DSS ‘wild type’ increases BP. When both *C18QTL3/Tcof1* and *C18QTL4/Hdhd2* are ‘off’ as in the congenic strain C18S.L2, BP decreases to the same level as either of the 2 mutants alone, as if *C18QTL3/Tcof1* alone is ‘off’. This is because as long as *C18QTL3/Tcof1* is ‘off’, the downstream regulatory target for *C18QTL4/Hdhd2* is ‘off’, regardless whether *C18QTL4/Hdhd2* itself is ‘on’ or ‘off’. This result molecularly establishes epistasis between them and justifies why the *C18QTL3/Tcof1* ‘mutant’ masks the BP effect of the *C18QTL4/Hdhd2* ‘mutant’, when combined.

An alternative model of a required BP threshold between 2 independent rather than epistatic QTLs is not consistent with experimental data. If *C18QTL3* and *C18QTL4* independently affected BP, their cumulative effects should be twice as much as each of acting alone. The data, however, contradicted this prediction (Fig. 1). Moreover, no BP threshold exists in DSS rats, since it has lost its buffering capacity [20].

Since all 4 adjacent QTLs on Chr. 18 (Fig. 1) influence BP epistatically [6,20], *C18QTL1* and *C18QTL2* are expected to participate in the Hdhd2-Tcof1 pathway (Fig. 4), even though they have not been molecularly identified. Predictably, other QTLs may regulate, or be regulated by, Hdhd2 and/or Tcof1 at any level from its expression, transcription, posttranscription, translation, post-translation to their activity.

Hdhd2 and *Tcof1* co-localize to certain cells potentially relevant to BP regulation such as the aortic smooth muscle cells of the vasculature (data not shown).

11.5.2 *Hdhd2/C18QTL4* as a post-translational regulator of *Tcof1/C18QTL3*

The present work provides the first experimental evidence for any molecular function of Hdhd2 as a true phosphatase, capable of de-phosphorylating Tcof1 (Fig. 3). Added to this molecular impact, their mutual epistatic relationship (Fig. 1) further places them in the same pathway in controlling BP, since Hdhd2 may de-phosphorylate numerous proteins other than Tcof1, conversely Tcof1 can be de-phosphorylated by phosphatases other than Hdhd2 in the organism, but these other relationships may have nothing to do with BP.

Phosphorylation as a molecular basis of epistasis is exemplified by the pathway of mammalian target of rapamycin (mTOR) [21]. mTOR is epistatic to numerous down-stream targets in the same pathway and regulates them via phosphorylation.

Hdhd2 is a unique phosphatase, because a homozygous mouse knockout is embryonic lethal (<http://www.informatics.jax.org/marker/MGI:1924237>), suggesting that its function is irreplaceable. The diminished enzymatic efficiency may appear minor due to the G337C mutation in Hdhd2^{DSS}, but under physiological conditions, the Tcof1 concentration may be highly limited in the cell nucleus in the tissue where the Hdhd2/Tcof1(treacle) pathway plays a vital role in regulating BP. An increase in its affinity to Tcof1 might give Hdhd2^{Lewis} a catalytic advantage, thus favors lower BP. It is well-known that posttranslational modifications by kinases play a role in monogenic hypertension [22].

11.5.3 *Tcof1/C18QTL3* may be a downstream target of *Hdhd2/C10QTL4*

TCOF1 causes the human Treacher Collins syndrome (TCS) via haploinsufficiency [17,18]. The protein product of *TCOF1*, treacle, is a nucleolar phosphoprotein involved in ribosomal processing [23]. Although most TCS-inducing mutations result in truncated or non-functional treacles, an in-frame deletion within *TCOF1* may not necessarily cause outright pathologies, but may lead to a disorder. For example, an in-frame deletion in exon 6A of *TCOF1* does not cause TCS, but has unknown pathogenicity in humans [24]. Coincidentally, a TCS patient carries the 18-bp deletion in treacle at exactly the same segment as that found in the DSS rat (Table 1) along with a 1-bp insertion. The mother of the TCS patient only carries the in-frame 18-bp deletion, but shows no symptom of TCS [25]. Whether she has hypertension or not has not been reported.

Tcof1 is a sole function candidate for *C18QTL3* for the following reasons. First, restoring the 18-bp deletion by C18S.L14 lowers BP (Fig. 1). Second, the 18-bp *Tcof1* deletion has functional consequences, since it reduces the exon containing it by 9.5% (18/188) and entire *Tcof1*/treacle by 0.4% (6/1395). The rat 18-bp-containing exon (#15) corresponds to the human exon 16, which is a *de novo* mutation hot spot for TCS [26]. Finally, the 6 amino acids removed by the deletion (Sup 7) are located in a highly conserved and repeated region involved in the *Tcof1*/treacle phosphorylation [17,18]. The degree of phosphorylation

may affect the efficiency of nucleolar translocation that Tcof1/treacle executes to fulfil its function in ribosomal processing [27]. The Tcof1/treacle de-phosphorylation by Hdhd2 (Fig. 3d) is consistent with the Western blotting results using alkaline phosphatase (AP) of other investigators [28], and thus proves that Hdhd2 functions in the same capacity as AP.

Since the quantity of treacle protein is miniscule in rat organs (data not shown), despite its expression was detected by RT-PCR, *in vivo* measurements of phosphorylation are not achievable in DSS and Lewis rats (data not shown). Whether or not Hdhd2 can dephosphorylate Tcof1/treacle *in vivo* remains to be determined.

11.5.4 Mechanistic connection to blood pressure regulation

The Tcof1/treacle protein is not a known BP-controlling agent itself. Neither is Hdhd2. How can they affect the BP homeostasis? Based on epistasis between them on BP and Hdhd2 capable of de-phosphorylating Tcof1/treacle, Hdhd2 appears to act from upstream in a pathway by regulating Tcof1/treacle possibly at the level of posttranslation (Fig. 4), and consequently, it is not expected to affect BP ‘directly’ and immediately. Since at least 16 components belong to epistatic module 1 [4], ‘indirect’ products of QTLs such as those encoded by *Hdhd2* and *Tcof1* may simply be these whose roles in a pathway are further upstream from a ‘direct’ gene product that is involved in an end-stage physiology leading to BP. Multiple intervening steps bridging them are to be expected [5].

How treacle might influence the BP homeostasis from downstream of the Hdhd2-Tcof1(treacle) pathway can only be speculated at the present. For one thing, inhibiting the p53 function on apoptosis was shown to prevent physiological abnormalities in the *Tcof1* mice [29]. The p53-mediated pathway has also been found to be anti-angiogenic in systolic function [30]. Thus, a possible connection through Tcof1/treacle-p53-angiogenesis-cardiac function-BP could not be excluded.

11.5.5 The BP QTLs are not the prominent genes physiologically known to be involved in BP regulation

In contrast to *Hdhd2/C18QTL4* and *Tcofl1/C18QTL3* whose mechanistic connections to BP are obscure, certain genes with well-recognized roles in BP modulation are present in the QTL-residing interval, but are not supported as a genetic candidate for the QTL in question [6]. Cases in point are the genes encoding adrenergic beta-2 receptor (*Adrb2*) and neural precursor cell expressed developmentally down-regulated 4-like protein (*Nedd4l*), which are well-known BP physiology agents. The *Adrb2* protein is one of antihypertensive therapeutic targets by beta blockers [31]. Both genes exist in the *C18QTL3*-lodging interval (Supplement 3). Knocking out either *Adrb2* [32] or *Nedd4l* [33] alters BP. Physiologically, they appeared the ideal candidates for *C18QTL3*.

However, genetically, *Adrb2* or *Nedd4l* carries neither a function-altering mutation that could impact on the structure, splicing, or expression [6], nor copy number variations (data not shown). Thus, neither of them can be qualified as a genetic candidate for *C18QTL3*, despite their obvious physiological roles in BP regulation. In the absence of a function-changing mutation, *Adrb2* or *Nedd4l* can not cause a difference in BP due to its genetic variation, which is the cornerstone of Mendelian inheritance and the basis of human genome-wide association studies and animal genetics. A QTL has to change BP, but a gene that can affect BP, in spite of its prominent action in the BP physiology, is not necessarily a QTL. In contrast, *Hdhd2* and *Tcofl1* are strong candidate genes to be *C18QTL4* and *C18QTL3* respectively because they bear function-changing mutations, even though their physiological impact on BP seems ‘indirect’.

11.5.6 Implications in human essential hypertension

First, a large number of human QTLs have been localized [1] and how they act together in biologically affecting the overall BP is unknown. Since *C18QTL3* and *C18QTL4* actually function in an epistatic hierarchy via a posttranslational modification in a pathway (Fig. 4), this conceptual framework should facilitate the elucidation of the relationships among some human BP QTLs.

Second, although no significant signals were detected near any of *C18QTL3* or *C18QTL4* in human genome-wide association studies (GWAS) [1], some small-scale studies have shown that QTLs on CHR 18 might exist, albeit in specific human populations [35,36]. This phenomenon can not be overlooked, because SNPs used in a GWAS has a minor-allele frequency of at least 5% [37], those SNPs below that frequency threshold are filtered out of the screening, but may still be relevant in stratified populations. The common disease-common variant hypothesis does not exclude a common disease-rare variant postulate [38].

Finally, from a therapeutic view point, improving the de-phosphorylation capacity and the integrity of a downstream phosphatase substrate lowers BP and thus might become targets against hypertension that are founded on the mechanistic causes of BP changes.

11.6 Conclusion

Current studies have identified 2 genes with function-altering variants, which are qualified to be 2 individual BP QTLs, *Hdhd2/C18QTL4* and *Tcof1/C18QTL3*, that fall into an epistatic hierarchy. A shared pathway constitutes a fundamental BP genetic architecture and reveals, for the first time, a post-translational regulatory mechanism underlying the epistatic relationship between 2 BP QTLs (Fig. 4). The regulatory hierarchy is the first for 2 BP QTLs in any mammal, lays the foundation for a comprehensive pathway construction and for generating novel anti-hypertensive targets based on a BP etiology and driven by the mechanism.

Acknowledgements: This work was supported by an operating grant from the Canadian Institutes for Health Research (CIHR) to AYD. CC is supported by a doctoral fellowship from the Heart and Stroke Foundation of Canada.

Disclosures: None.

11.7 References

1. The International Consortium for blood pressure. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 2011; **478**: 103-109.
2. Cowley AW, Jr. The genetic dissection of essential hypertension. *Nat Rev Genet* 2006; **7**: 829-840.
3. Deng AY. Genetic basis of polygenic hypertension. *Hum Mol Genet* 2007; **16**: R195-R202.
4. Chauvet C, Crespo K, Menard A, Roy J, Deng AY. Modularization and epistatic hierarchy determine homeostatic actions of multiple blood pressure quantitative trait loci. *Hum Mol Genet* 2013; **22**: 4451-4459.
5. Deng AY. Genetic mechanisms of polygenic hypertension: fundamental insights from experimental models. *J Hypertension* 2014; **in press**.
6. Chauvet C, Crespo K, Menard A, Wu Y, Xiao C, Blain M, Roy J, Deng AY. alpha-Kinase 2 is a novel candidate gene for inherited hypertension in Dahl rats. *J Hypertens* 2011; **29**: 1320-1326.
7. Deng AY. Positional Cloning of Quantitative Trait Loci for Blood Pressure: How Close Are We?: A Critical Perspective. *Hypertension* 2007; **49**: 740-747.
8. Duong C, Charron S, Xiao C, Hamet P, Menard A, Roy J, Deng AY. Distinct quantitative trait loci for kidney, cardiac, and aortic mass dissociated from and associated with blood pressure in Dahl congenic rats. *Mamm Genome* 2006; **17**: 1147-1161.
9. Deng AY, Martin LL, Balwierczak JL, Jeng AY. Purification and characterization of an endothelin degradation enzyme from rat kidney. *J Biochem* 1994; **115**: 120-125.
10. Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* 1979; **76**: 4350-4354.
11. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* 2011; **39**: D225-D229.

12. Guo RJ, Suh ER, Lynch JP. The role of Cdx proteins in intestinal development and cancer. *Cancer Biol Ther* 2004; **3**: 593-601.
13. Grillet N, Schwander M, Hildebrand MS, Sczaniecka A, Kolatkar A, Velasco J, Webster JA, Kahrizi K, Najmabadi H, Kimberling WJ, Stephan D, Bahlo M, Wiltshire T, Tarantino LM, Kuhn P, Smith RJH, Muller U. Mutations in LOXHD1, an Evolutionarily Conserved Stereociliary Protein, Disrupt Hair Cell Function in Mice and Cause Progressive Hearing Loss in Humans. *The American Journal of Human Genetics* 2009; **85**: 328-337.
14. Riazuddin S-A, Parker D-S, McGlumphy E-J, Oh E-C, Iliff B-W, Schmedt T, Jurkunas U, Schleif R, Katsanis N, Gottsch J-D. Mutations in LOXHD1, a Recessive-Deafness Locus, Cause Dominant Late-Onset Fuchs Corneal Dystrophy. *The American Journal of Human Genetics* 2012; **90**: 533-539.
15. Shikauchi Y, Saiura A, Kubo T, Niwa Y, Yamamoto J, Murase Y, Yoshikawa H. SALL3 Interacts with DNMT3A and Shows the Ability To Inhibit CpG Island Methylation in Hepatocellular Carcinoma. *Mol Cell Biol* 2009; **29**: 1944-1958.
16. Kushnir M, Dresner E, Mandel S, Gozes I. Silencing of the ADNP-family member, ADNP2, results in changes in cellular viability under oxidative stress. *Journal of Neurochemistry* 2008; **105**: 537-545.
17. Gladwin AJ, Dixon J, Loftus SK, Edwards S, Wasmuth JJ, Hennekam RC, Dixon MJ. Treacher Collins syndrome may result from insertions, deletions or splicing mutations, which introduce a termination codon into the gene. *Hum Mol Genet* 1996; **5**: 1533-1538.
18. The Treacher Collins Syndrome Collaborative Group. Positional cloning of a gene involved in the pathogenesis of Treacher Collins syndrome. The Treacher Collins Syndrome Collaborative Group. *Nat Genet* 1996; **12**: 130-136.
19. Preumont A, Rzem R, Vertommen D, Van-Schaftingen E. HDHD1, which is often deleted in X-linked ichthyosis, encodes a pseudouridine-5-phosphatase. *Biochem J* 2010; **431**: 237-244.
20. Charron S, Lambert R, Eliopoulos V, Duong C, Menard A, Roy J, Deng AY. A loss of genome buffering capacity of Dahl salt-sensitive model to modulate blood pressure as a cause of hypertension. *Hum Mol Genet* 2005; **14**: 3877-3884.
21. Hay N, Sonenberg N. Upstream and downstream of mTOR. *Genes & Development* 2004; **18**: 1926-1945.

22. Wilson FH, Disse-Nicodeme S, Choate KA, Ishikawa K, Nelson-Williams C, Desitter I, Gunel M, Milford DV, Lipkin GW, Achard JM, Feely MP, Dussol B, Berland Y, Unwin RJ, Mayan H, Simon DB, Farfel Z, Jeunemaitre X, Lifton RP. Human hypertension caused by mutations in WNK kinases. *Science* 2001; **293**: 1107-1112.
23. Lin CI, Yeh NH. Treacle recruits RNA polymerase I complex to the nucleolus that is independent of UBF. *Biochemical and Biophysical Research Communications* 2009; **386**: 396-401.
24. Bowman M, Oldridge M, Archer C, O'Rourke A, McParland J, Brekelmans R, Seller A, Lester T. Gross deletions in TCOF1 are a cause of Treacher-Collins-Franceschetti syndrome. *Eur J Hum Genet* 2012; **20**: 769-777.
25. Wang Y, Yin XJ, Han T, Peng W, Wu HL, Liu X, Feng ZC. A novel silent deletion, an insertion mutation and a nonsense mutation in the TCOF1 gene found in two Chinese cases of Treacher Collins syndrome. *Mol Genet Genomics* 2014; 1-4.
26. Splendore A, Silva EO, Alonso LsG, Richieri-Costa An, Alonso N, Rosa A, Carakushanky G, Cavalcanti DP, Brunoni D+, Passos-Bueno MR. High mutation detection rate in TCOF1 among Treacher Collins syndrome patients reveals clustering of mutations and 16 novel pathogenic changes. *Hum Mutat* 2000; **16**: 315-322.
27. Gonzales B, Henning D, So RB, Dixon J, Dixon MJ, Valdez BC. The Treacher Collins syndrome (TCOF1) gene product is involved in pre-rRNA methylation. *Hum Mol Genet* 2005; **14**: 2035-2043.
28. Isaac C, Marsh KL, Paznekas WA, Dixon J, Dixon MJ, Jabs EW, Meier UT. Characterization of the Nucleolar Gene Product, Treacle, in Treacher Collins Syndrome. *Mol Biol Cell* 2000; **11**: 3061-3071.
29. Jones NC, Lynn ML, Gaudenz K, Sakai D, Aoto K, Rey JP, Glynn EF, Ellington L, Du C, Dixon J, Dixon MJ, Trainor PA. Prevention of the neurocristopathy Treacher Collins syndrome through inhibition of p53 function. *Nat Med* 2008; **14**: 125-133.
30. Sano M, Minamino T, Toko H, Miyauchi H, Orimo M, Qin Y, Akazawa H, Tateno K, Kayama Y, Harada M, Shimizu I, Asahara T, Hamada H, Tomita S, Molkentin JD, Zou Y, Komuro I. p53-induced inhibition of Hif-1 causes cardiac dysfunction during pressure overload. *Nature* 2007; **446**: 444-448.
31. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH,

- Lisabeth LD, Magid D, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Nichol G, Paynter NP, Schreiner PJ, Sorlie PD, Stein J, Turan TN, Virani SS, Wong ND, Woo D, Turner MB. Heart Disease and Stroke Statistics 2013 Update: A Report From the American Heart Association. *Circulation* 2013; **127**: e6-e245.
32. Rohrer DK, Bernstein D, Chruscinski A, Desai KH, Schauble E, Kobilka BK. The developmental and physiological consequences of disrupting genes encoding beta 1 and beta 2 adrenoceptors. *Adv Pharmacol* 1998; **42**: 499-501.
 33. Shi PP, Cao XR, Sweezer EM, Kinney TS, Williams NR, Husted RF, Nair R, Weiss RM, Williamson RA, Sigmund CD, Snyder PM, Staub O, Stokes JB, Yang B. Salt-sensitive hypertension and cardiac hypertrophy in mice deficient in the ubiquitin ligase Nedd4-2. *Am J Physiol Renal Physiol* 2008; **295**: F462-F470.
 34. Munroe PB, Barnes MR, Caulfield MJ. Advances in Blood Pressure Genomics. *Circ Res* 2013; **112**: 1365-1379.
 35. Guzman B, Cormand B, Ribases M, Gonzalez-Nunez D, Botey A, Poch E. Implication of chromosome 18 in hypertension by sibling pair and association analyses: putative involvement of the RKHD2 gene. *Hypertension* 2006; **48**: 883-891.
 36. Rutherford S, Johnson MP, Griffiths LR. Sibpair studies implicate chromosome 18 in essential hypertension. *Am J Med Genet* 2004; **126A**: 241-247.
 37. Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *New England Journal of Medicine* 2010; **363**: 166-176.
 38. Kathiresan S, Srivastava D. Genetics of Human Cardiovascular Disease. *Cell* 2012; **148**: 1242-1257.
 39. Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, Mooney SD, Radivojac P. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 2009; **25**: 2744-2750.
 40. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009; **4**: 1073-1081.

11.8 Tables, figures and legends

Table 1. Mutation screening of genes in the *C18QTL3*-containing interval

Gene	Codons			In/Ex junctions		
	Size of codon (bp)	# of exons	Mutation detected Lewis/ DSS	Change in amino acid (AA) Lewis/DSS	Mutation detected in In/Ex junction	Differential Splicing
<i>Tcof1</i>	4188	26	GGAGTCTGACAGT GAAGA 2283Δ2300	ESDSEE 762Δ767	No	
<i>Cdx1</i> (only expressed in intestine)	843	4	G105C	No	No	
			C107T	Thr36Ile	No	
			C117A	No	No	
			C119A	Pro40Gln	No	
			C130A	Pro44Thr	No	
			T163A	Tyr55Asn	No	
			C201A	No	No	
G205A	Ala69Thr	No				
<i>Nedd4l</i>	2892	31	A321T(exon 6)	No	Splice donor	No
			T1995C	No	No	
			A2139G	No	No	
<i>Alpk2</i>	6303	10	T376C	Cys126Arg	No	
			C475T	Pro159Ser	No	
			C618T	No	No	
			CCG1910Δ	Ala637Δ	maybe	No
			A2918T	Gln973Leu	No	
A3754G	Lys1252Glu	No				
C5610G	No	No				

			G90A	No	No	
<i>Zfp532</i>	3999	10	A810G	No	No	
					Ex9(-28)	No
			G1509A	No	No	
					Ex1(+35)	No
<i>Oacyl</i>	2031	15			Ex10(-42)	No
					Ex10(-35)	No
					Ex12(+4)	No
					Ex14(+11)	No
<i>Sec11c</i>	579	6	T171C	No	No	

Footnote to Table 1: Only the genes carrying structural or splice-junction mutations are listed from Supplement 3. The position of a mutation corresponds to the designation from the ATG start codon of that gene. The amino acid position begins from the first methionine. *Alpk2*, alpha kinase 2; *Cdx1*, caudal-type homeobox protein 1; *Nedd4l*, neural precursor cell expressed, developmentally down-regulated 4-like, E3 protein ligase; *Oacyl*, O-acyltransferase like; *Sec11c*, SEC11 homolog C (*S. cerevisiae*); *Tcofl*, Treacher Collins-Franceschetti syndrome 1; *Zfp532*, zinc finger protein 532 (a gene which does not have a human homologue). In: intron, Ex: exon, (+) and (-): nucleotide after before a given exon respectively. No copy number variations were detected for these genes from total genome sequencing.

Table 2. Predictions on a functional impact of a missense mutation occurring in genes residing in *C18QTL3* and *C18QTL4* residing intervals.

Gene	C18QTL	cDNA	Protein	Mutpred	SIFT
<i>Alpk2</i>	3	T376C	Cys126Arg	benign	benign
		C475T	Pro159Ser	benign	benign
		CCG1910Δ	Ala637Δ	N/A	N/A
		A2918T	Gln973Leu	benign	benign
		A3754G	Lys1252Glu	benign	benign
<i>Hdhd2</i>	4	G337C	Val113Leu	Deleterious	Deleterious
<i>Rbfa</i>	4	A487G	Thr163Ala	benign	benign
		A911G	Glu304Gly	benign	benign
<i>Adnp2</i>	4	A353G	Lys118Arg	benign	benign
		A1607C	Gln536Pro	benign	benign
		A2140G	Met714Val	benign	benign
		G2149A	Ala717Thr	benign	N/A
		G3400C	Glu1134Gln	benign	benign
<i>Loxhd1</i>	4	G3346A	Val1116Met	benign	benign
		G3880A	Val1294Ile	benign	benign
		G3991A	Val1331Ile	Benign	Benign
		A5929G	Ile1977Val	benign	benign
<i>Sall3</i>	4	C3460T	Arg1154Cys	benign	Deleterious
		C3483G	Ile1161Met	benign	benign

Footnote to table: MutPred score [39] (<http://mutpred.mutdb.org/>), and SIFT [40] (<http://sift.jcvi.org/>) refer to various prediction programs used for assessing a probable consequence of a missense mutation on the function of a protein it encodes.

Table 3. Mutation screening of genes in the *C18QTL4*-containing interval

Gene	Size of codons (bp)	# of exons	Codons		In/Ex junctions	
			Mutation detected Lewis/ DSS	Change in amino acid (AA) Lewis/DSS	Mutation detected in In/Ex junction	Alternative Splicing
<i>Ctif</i>	1011	5	G189A	No	No	
<i>Pqlc1</i>	744	5			Ex4(-26) Ex5(-28)	
<i>Rbfa</i>	1038	6	G474A	No	No	
			A487G	Thr163Ala	maybe	No
			A911G	Glu304Gly	No	
					Ex4 (+20) Ex5 (+14) Ex6 (+12) Ex7 (-15)	No
<i>Adnp2</i>	3411	3	A353G	Lys118Arg	No	
			A1607C	Gln536Pro	No	
			A1926G	No	No	
			A2140G	Met714Val	No	
			G2149A	Ala717Thr	No	
			A2499G	No	No	
			G3126A	No	No	
			T3144C	No	No	
	G3400C	Glu1134Gln	No			
<i>Pard6g</i>	1149	3	C309T	No	No	
			T537C	No	No	
			G699A	No	No	
<i>Hdhd2</i>	780	6	G337C	Val113Leu	No	

<i>St8sia5</i>	1131	7	C246T	No	No
			T2190C	No	No
			C2859G	No	No
			C2880T	No	No
			G2937A	No	No
			G3346A	Val1116Met	No
			G3880A	Val1294Ile	No
			G3991A	Val1331Ile	No
			G4107A	No	maybe
			C5268T	No	No
			A5929G	Ile1977Val	No
			C6375T	No	No

				Ex2(-35)	Not expressed
				Ex10(-23)	
<i>Loxhd1</i>				Ex16(-32)	
(Expression not detectable Fig. 2)	6651	40		Ex16(-42)	
				Ex18(-49)	
				Ex18(-33)	
				Ex18(-32)	
				Ex19(-49)	
				Ex21(-8)	
				Ex23(+22)	
				Ex23(+42)	
				Ex26(+10)	
				Ex33(-38)	
				Ex33(+5)	
				Ex35(-39)	
				Ex36(+39)	
				Ex37(+9)	

					Ex37(+10)	
					Ex37(+42)	
					Ex39(+27)	
					Ex39(+34)	
			T225C	No	No	
<i>Setbp1</i>	4494	6	G1242A	No	No	
			A1668G	No	No	
					Ex5(-19)	No
<i>Nfatc1</i>	2484	10			Ex5(+41)	No
			T87C	No	No	
			T1003C	No	No	
			A3393G	No	No	
					Ex6(-3)	No
<i>Atp9b</i>	3444	30			Ex10(-36)	No
					Ex12(-2)	No
					Ex14(-6)	No
					Ex19(+45)	No
					Ex21(+25)	No
					Ex22(+33)	No
<i>Sall3</i>			A2865G	No	No	
(Only expressed in early development)	3573	2	C3460T	Arg1154Cys	No	
			C3483G	Ile1161Met	No	

Footnote to Table: Only the genes carrying structural or splice-junction mutations are listed from Supplement 3. The position of a mutation enumerates from the ATG start codon of that gene. The amino acid position begins from the first methionine. *Adnp2*, ADNP homeobox protein 2; *Atp9b*, ATPase, class II, type 9B; *Ctif*, CBP80/20-dependent translation initiation factor; *Hdhd2*, haloacid dehalogenase-like hydrolase domain-containing protein 2; *Loxhd1*, lipoxygenase homology domains 1; *Nfatc1*, nuclear factor of activated T cells, cytoplasmic, calcineurin-dependent; *Pard6g*, par-6 partitioning defective 6 homolog gamma (C. elegans);

Pqlc1, PQ loop repeat containing 1; *Rbfa*, ribosome binding factor A; *Sall3*, sal-like 3 (Drosophila); *Setbp1*, SET binding protein 1; *St8sia5*, ST8 alpha-N-acetyl-neuramidine alpha-2,8-sialyltransferase 5. In: intron, Ex: exon, (+) and (-): nucleotide after and before a given exon. No copy number variations were detected for these genes from total genome sequencing.

Table 4. Kinetic studies of phosphatase properties of Hdh2 *in vitro*

	<i>K</i> m (pM)	<i>V</i> max (pM/sec)
DSS Hdh2	112.39 ±7.61	0.87 ±0.18
Lewis Hdh2	87.57 ±5.54	0.68 ±0.22
<i>p t</i> -test	0.025	0.51

Footnote: *In vitro* assays of phosphatase activities were done in triplicate from cultured cells transfected with Hdh2 clones from Dahl salt sensitive (DSS) and Lewis rat Hdh2 cDNAs respectively. The sole difference of G337C between DSS and Lewis *Hdh2* coding regions were confirmed by sequencing. 3 independent cultures, each in triplicates, were analyzed, i.e. a total of n=9. ±, SEM.

Figure 1. Defining BP QTLs for blood pressure rat Chr 18 by congenic strains. A solid bar under congenic strains represents the DSS fragment (a white bar) that has been replaced by that of Lewis (S.L). Striped bars on ends of the solid bars denote the ambiguity of crossover breakpoints between markers. The congenic regions are not drawn to scale. Mean arterial pressures (MAPs) for all the strains are averages for the duration of the measurement. Systolic and diastolic arterial pressures are consistent with their MAPs of all the strains (data not shown). Rat genes of biological interest are: *Hdh2*, *Haloacid dehalogenase-like hydrolase domain containing 2*; *Tcof1*, *Treacher Collins-Franceschetti syndrome 1*. A 2 x 2 ANOVA on interactions [4] shows epistasis between *C18QTL3* defined by C18S.L14 and *C18QTL4* defined by C18S.L10, because BP lowering (35mmHg) when combining *C18QTL3* and *C18QTL4* in the C18S.L2 congenic strain was highly different ($p < 0.001$) from the predicted sum of BP lowering (73mmHg) by adding the effects of 2 QTLs.

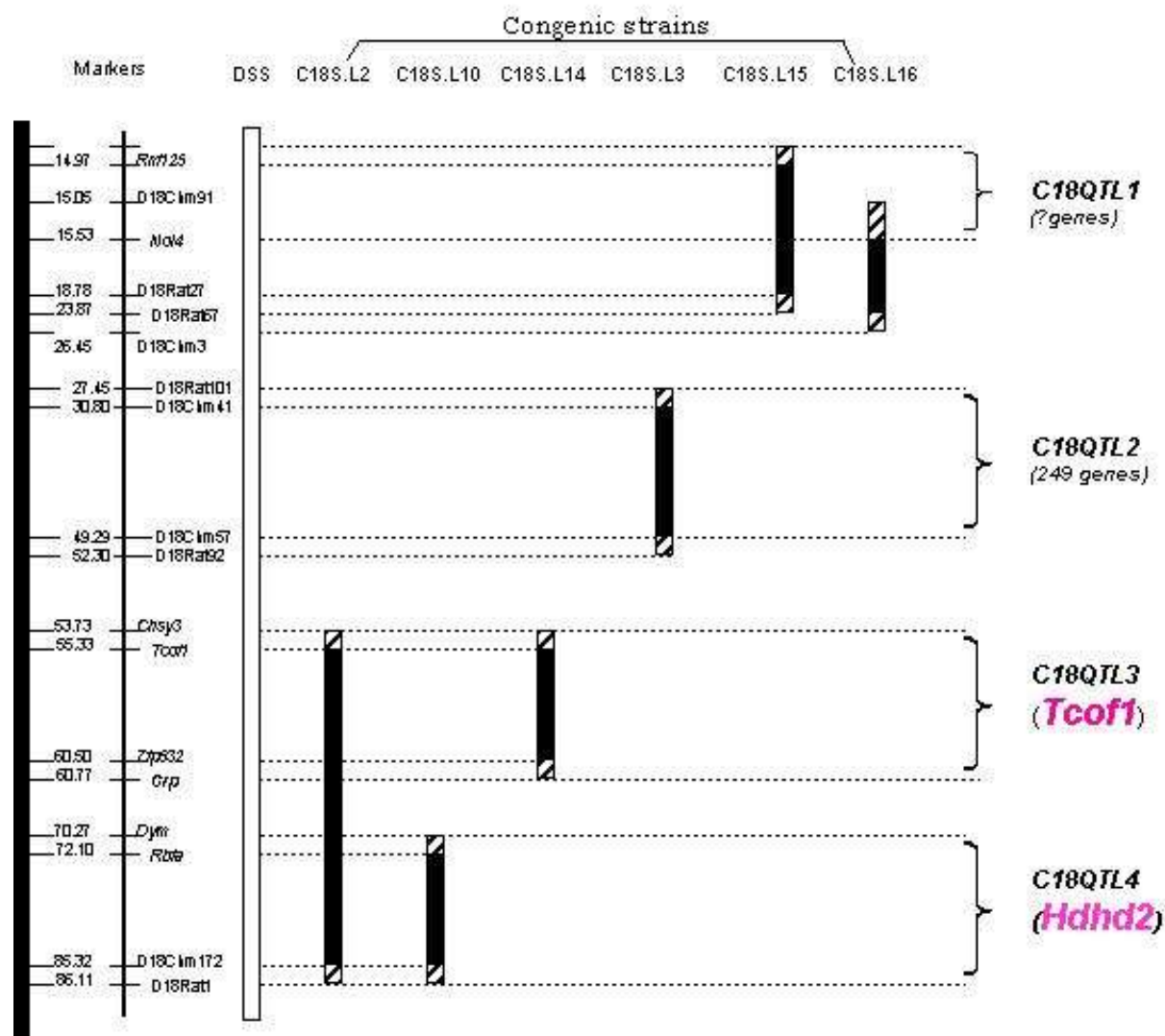
Figure 2. Organ expression pattern of genes assayed by reverse transcriptase polymerase chain reaction (RT-PCR). The organs are from Dahl salt-sensitive (DSS) rats and Lewis rats. Numbers to the right indicate the size of the fragment in base pairs. The full gene names are given in Tables 1&3. Primers for RT-PCRs of genes are listed in Supplement 5 and located in 2 different exons to avoid amplifying genomic DNAs contaminated in RNA preparations. Before RT-PCR, all the mRNA samples were treated with RNase-free DNAs to remove possible traces of genomic DNAs. Results shown are from 1 rat of each strain and they have been replicated with multiple rats of the same strain (data not shown). All rats were males, 11 weeks of age and fed a high salt diet for 6 weeks starting from 5 weeks of age. The *Sall3* expression was analyzed with multiple PCR primer pairs and no products were detected in multiple PCR runs, consistent with its early developmental-specific profile, not in the adults [15].

Figure 3. Functional analysis of 2 Hdh2 isoforms and de-phosphorylation of Tcof1 by Hdh2. a. Scheme in cloning the full-length *Hdh2* cDNA for expressions. b. Confirmation of 2 isoform (DSS and Lewis) proteins expressed *in vitro* by Hdh2 antibodies. c. Lineweaver-Burk plot of the phosphatase activity of Hdh2^{DSS} and Hdh2^{Lewis}. d. Detecting de-phosphorylation of Tcof1 by Hdh2 by Western blotting indicated by 2 parallel left-pointing arrow heads. Tcof1/treacle and its dephosphorylation activities were detected by Tcof1/treacle specific antibodies. J, Jurkat nuclear extract alone; H, Hdh2 cell extract; I, phosphatase inhibitors. Tcof1 was not detected by Western in 150µg (i.e. the up most limit) of crude heart and kidney extracts, presumably its quantity in organs is limited and its location compartmentalized in the nucleus of specialized cells.

Figure 4. The actual order between C18QTL4/Hdh2 and C18QTL3/Tcof1 in one pathway toward BP modulations. The SS genotype (homozygous for DSS) refers to the ‘on’ state of the gene, and the LL genotype (homozygous for Lewis) refers to the ‘off’ state of the gene. A horizontal arrow indicates the general direction leading to the BP control and does not imply an immediate BP-effector. A down pointing arrow indicates a decrease in BP with reference to the DSS ‘wild type’.

Figure 1

Rat Chromosome 18



MAP (mmHg)	177±4	142±2	133±3	148±5	124±4	157±8	170±8
n=	16	5	6	7	5	5	5
MAP difference DSS minus congenics (mmHg)	—	35	44	29	53	20	7
ANOVA + Dunnett <i>P</i>	—	<0.001	<0.001	<0.001	<0.001	0.04	>0.29

Figure 2

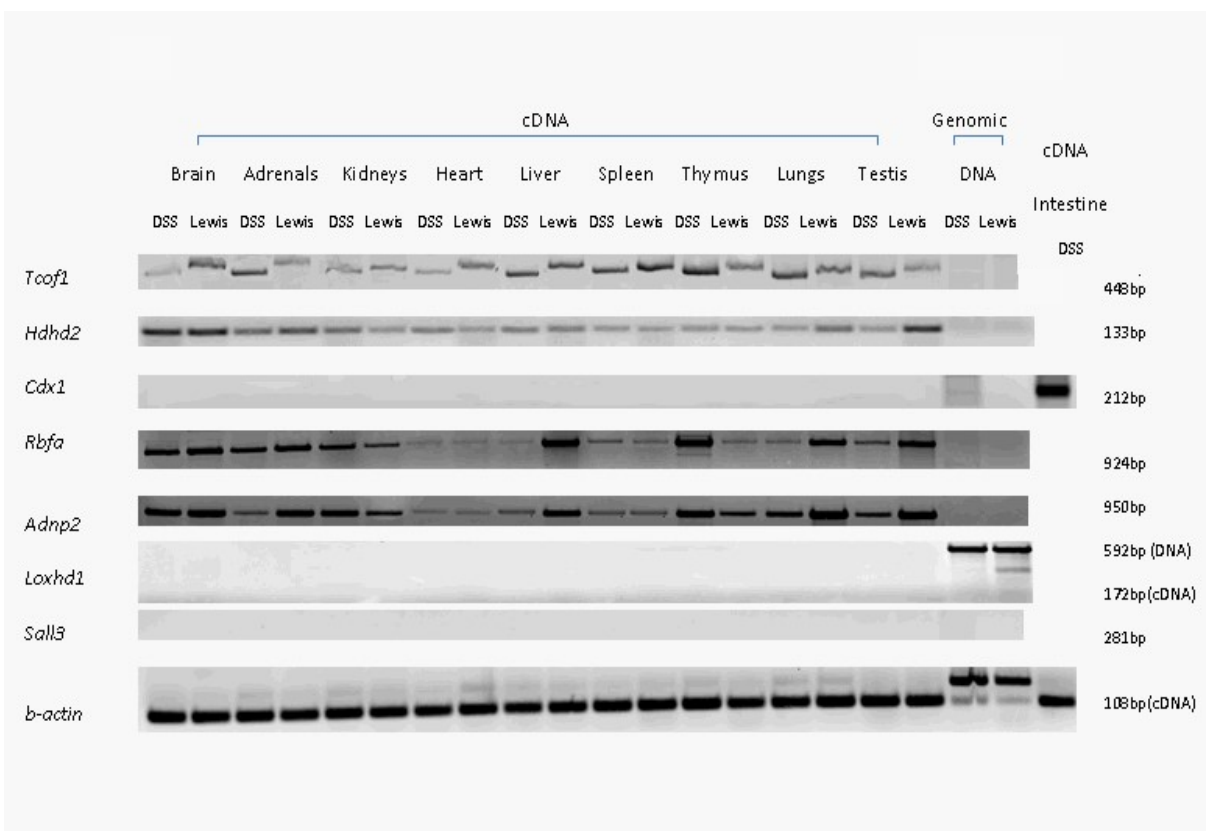


Figure 3

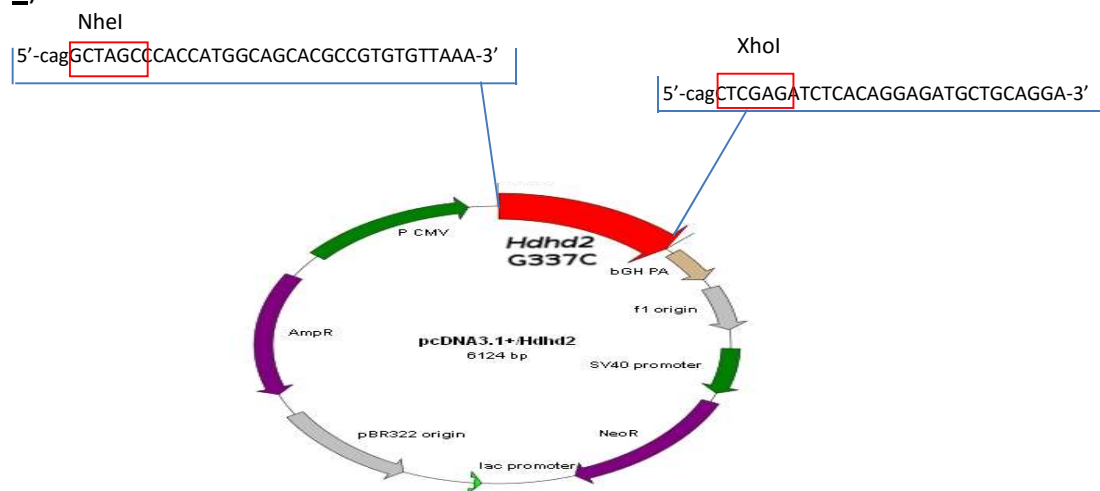
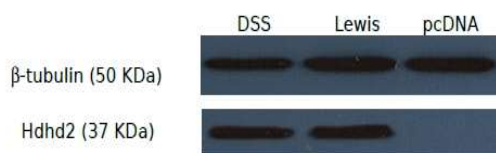
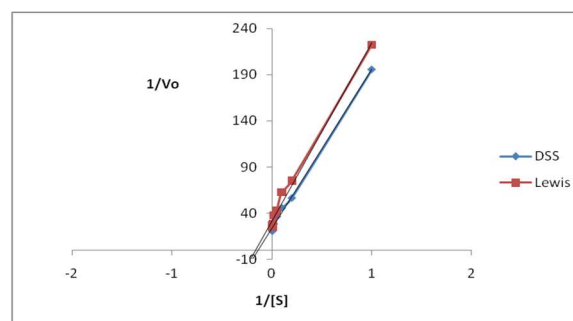
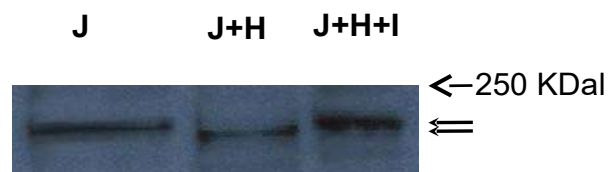
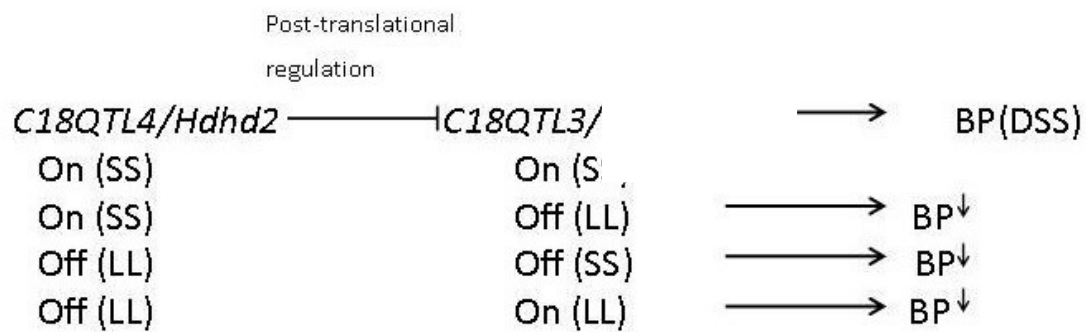
a).**b).****c).****d).**

Figure 4



11.9 Supplementary data

Supplement 1: Total genome sequencing and identification of single nucleotide polymorphisms (SNPs): DNA Sequencing Pipeline Genome Quebec and McGill Innovation Center
October 2013 (bioinformatics.genome@mail.mcgill.ca)

Program	Reference
Genome Analysis Toolkit	http://www.broadinstitute.org/gsa/wiki/index.php/The_Genome_Analysis_Toolkit
BWA	http://bio-bwa.sourceforge.net/
Igtools	http://www.broadinstitute.org/igv/igvtools
samtools	http://samtools.sourceforge.net/
picard	http://picard.sourceforge.net/
snpSift/snpEff	http://snpeff.sourceforge.net/SnpSift.html
Trimmomatic	http://www.usadellab.org/cms/index.php?page=trimmomatic

Deliverables: The following files are delivered to clients: (1) A sample statistics file containing the metrics enumerated in Step 6. (2) A .csv file containing all the variants found in at least one sample, often classified by chromosome because of large file sizes. (3) A .csv file containing all the high impact coding variants found in at least one sample, based on the Step 9 annotations.

Sequencing Reads: Around 50 to 150 million 100 b.p. paired-end reads from the Illumina HiSeq 2000 sequencer. Base quality is encoded in phred 33.

Pipeline steps: The pipeline is executed on Compute Canada clusters via unix bash commands, perl scripts and open source software.

Step 1: Read trimming and clipping of adapters: Reads are trimmed from the 3' end to have a phred score of at least 30. Illumina sequencing adapters are removed from the reads, and all reads are required to have a length of at least 32 b.p. Trimming and clipping are done with the Trimmomatic software [1].

Step 2: Aligning the reads to the genome reference: The filtered reads are aligned to the reference genome. The alignment is done per lane of sequencing, and then merged for a complete Binary Alignment Map file (.bam). The alignment software used is bwa [2], and the merging is done with the picard software [3].

Step 3: Realigning insertions and deletions (INDELS): Insertion and deletion realignment is performed on regions where multiple base mismatches are preferred over indels by the aligner since it can appear to be less costly by the algorithm. Such regions will introduce false positive variant calls which may be filtered out by realigning those regions properly. Realignment is done with the GATK software [4].

Step 4: Fixing the read mates: Once local regions are realigned, the read mate coordinates of the aligned reads need to be recalculated since the reads are realigned at positions that differ from their

original alignment. Fixing the read mate positions is done with picard software [3].

Step 5: Marking duplicates: Aligned reads are duplicates if they have the same 5' alignment positions (for both mates in the case of paired-end reads). All but the best pair (based on alignment score) will be marked as a duplicate in the .bam file. Marking duplicates is done with picard software [3].

Step 6: Compute metrics and generating coverage track: Multiple metrics are computed at this stage and given in the statistics file:

- Number of raw reads
- Number of filtered reads (after Step 1)
- Number of aligned reads (after Step 2)
- Number of duplicate reads (after Step 5)
- Duplicate rate (number of duplicate reads / number of raw reads. Good run max of 25%)
- Median, mean and standard deviation of insert sizes of reads after alignment
- Mean coverage over exons (mean number of reads per base position)
- Percentage of bases covered at X reads (%_bases_above_50 means the % of exons bases which have at least 50 reads. A good run is typically around 50%)

A TDF (.tdf) coverage track is also generated at this step for easy visualization of coverage in the IGV browser [5].

Step 7: Variant calling: Variants (SNPs and INDELS) are called using samtools mpileup and bcftools varfilter [6]. The following options are given to mpileup to filter for low quality variants which could introduce false positive calls: -L 1000 -E -q 1 -u -D -S, where:

```
-L INT max per-sample depth for INDEL calling [250]
-E     extended BAQ for higher sensitivity but lower specificity
-q INT skip alignments with mapQ smaller than INT [0]
-u     generate uncompressed BCF output
-D     output per-sample DP in BCF (require -g/-u)
-S     output per-sample strand bias P-value in BCF (require -g/-u)
```

The output of mpileup is then fed to varfilter, which does an additional filtering of the variants and transforms the output into the VCF (.vcf) format. The arguments used are: -d 2 -D 1200 -Q 15 -l 0.0, where:

```
-d INT minimum read depth [2]
-D INT maximum read depth [10000000]
-Q INT minimum RMS mapping quality for SNPs [10]
-l FLOAT min P-value for strand bias (given PV4) [0.0001]
```

The final .vcf files are filtered for long 'N' INDELS which are sometimes introduced and causing excessive memory usage by downstream tools.

Supplement 2: Selective global evaluation of genome single nucleotide polymorphisms (SNPs) comparing DSS and Lewis rats.

Chromosome (Chr)	Gene number	SNP Count	SNPs in genes Count	SNPs in Exon Count	SNPs in intron-exon junctions count
Chr1	2 643	431 290	111 941	3 574	1 588
Chr2	1 115	506 527	98 039	1 886	1 111
Chr3	1 550	263 475	62 965	1 944	986
Chr4	1 185	270 209	78 309	1 505	897
Chr5	1 060	188 601	46 127	1 153	666
Chr6	719	247 020	46 312	1 088	649
Chr7	1 117	245 827	61 742	1 955	1 019
Chr8	1 057	204 003	51 750	1 448	752
Chr9	666	159 279	42 518	655	480
Chr10	1 507	196 741	61 224	2 582	1 367
Chr11	480	141 888	32 930	934	418
Chr12	508	94 706	29 131	815	626
Chr13	586	151 050	38 025	877	573
Chr14	551	124 193	29 395	714	389
Chr15	548	162 540	38 436	797	373
Chr16	492	109 407	30 684	740	478
Chr17	491	121 618	26 337	532	345
Chr18	414	134 343	36 156	1 092	387
Chr19	454	74 450	20 745	637	432
Chr20	592	112 869	32 476	2 139	936

chrX	595	122 396	14 593	490	241
Total	18 330	4 062 432	989 835	27 557	14 713

Footnote: Gene number refers to the number of known genes. Intron-exon junctions include before and exon sequences up to 100 base pairs.

Supplement 3. Genes residing in the chromosome intervals harboring *C18QTL3* and *C18QTL4* (Fig. 1).

(a) The *C18QTL3*-residing interval: (61 genes, 4 microRNAs, 1 non-coding RNA)

<i>RGD1309362</i>	similar to interferon-inducible GTPase
<i>MGC108823</i>	similar to interferon-inducible GTPase
<i>RGD1305184</i>	similar to CDNA sequence BC023105
<i>MGC105567</i>	similar to cDNA sequence BC023105
<i>Smim3</i>	small integral membrane protein 3
<i>Dctn4</i>	dynactin 4 (p62)
<i>Rbm22</i>	RNA binding motif protein 22
<i>Myoz3</i>	myozenin 3
<i>Synpo</i>	synaptopodin
<i>Ndst1</i>	N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 1
<i>Rps14</i>	ribosomal protein S14
<i>Cd74</i>	Cd74 molecule, major histocompatibility complex, class II invariant chain
<i>Tcof1</i>	Treacher Collins-Franceschetti syndrome 1
<i>Arsi</i>	arylsulfatase family, member 1
<i>Camk2a</i>	calcium/calmodulin-dependent protein kinase II alpha
<i>Slc6a7</i>	solute carrier family 6 (neurotransmitter transporter), member 7
<i>Cdx1</i>	caudal-type homeobox protein 1
<i>Pdgfrb</i>	platelet derived growth factor receptor, beta polypeptide
<i>Csf1r</i>	colony stimulating factor 1 receptor
<i>Hmgxb3</i>	HMG box domain containing 3
<i>Slc26a2</i>	solute carrier family 26 (anion exchanger), member 2
<i>Pde6a</i>	phosphodiesterase 6A, cGMP-specific, rod, alpha
<i>Ppargc1b</i>	peroxisome proliferator-activated receptor gamma, coactivator 1 beta
<i>Mir3557</i>	microRNA mir-3557
<i>Mir378</i>	microRNA mir-378
<i>LOC100360101</i>	homeobox prox 1-like
<i>Arhgef37</i>	Rho guanine nucleotide exchange factor (GEF) 37
<i>Csnk1a1</i>	casein kinase 1, alpha 1
<i>Mir145</i>	microRNA 145
<i>Mir143</i>	microRNA 143
<i>Il17b</i>	interleukin 17B
<i>Pcyox11</i>	prenylcysteine oxidase 1 like
<i>Grpel2</i>	GrpE-like 2, mitochondrial
<i>Afap111</i>	actin filament associated protein 1-like 1
<i>Ablim3</i>	actin binding LIM protein family, member 3
<i>Sh3tc2</i>	SH3 domain and tetratricopeptide repeats 2

<i>Adrb2</i>	adrenoceptor beta 2, surface
<i>Htr4</i>	5-hydroxytryptamine (serotonin) receptor 4, G protein-coupled
<i>Fbxo38</i>	F-box protein 38
<i>Spink7</i>	serine peptidase inhibitor, Kazal type 7
<i>Spink13</i>	serine protease inhibitor Kazal-type 13 serine protease inhibitor Kazal-type 5-like 3
<i>Spink10</i>	serine peptidase inhibitor, Kazal type 10
<i>Spink13</i>	serine protease inhibitor Kazal-type 13 serine protease inhibitor Kazal-type 5-like 3
<i>Apcdd1</i>	adenomatous polyposis coli down-regulated 1
<i>Napg</i>	N-ethylmaleimide-sensitive factor attachment protein, gamma
<i>Piezo2</i>	piezo-type mechanosensitive ion channel component 2
<i>Txn1l</i>	thioredoxin-like 1
<i>Wdr7</i>	WD repeat domain 7
<i>Wdr7</i>	WD repeat domain 7
<i>Si8sia3</i>	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3
<i>RGD1562699</i>	
<i>Onecut2</i>	one cut homeobox 2
<i>Fech</i>	ferrochelataase
<i>Nars</i>	asparaginyI-tRNA synthetase
<i>Atp8b1</i>	ATPase, aminophospholipid transporter, class I, type 8B, member 1
<i>Nedd4l</i>	neural precursor cell expressed, developmentally down-regulated 4-like, E3 ubiquitin protein ligase
<i>LOC502176</i>	hypothetical protein LOC502176
<i>Mir3591</i>	microRNA mir-3591
<i>Mir122</i>	microRNA mir-122
<i>Alpk2</i>	alpha-kinase 2
<i>Malt1</i>	mucosa associated lymphoid tissue lymphoma translocation gene 1
<i>LOC100912140</i>	uncharacterized LOC100912140
<i>RGD1560815</i>	similar to acidic ribosomal phosphoprotein P1
<i>Zfp532</i>	zinc finger protein 532
<i>Oacyl</i>	O-acyltransferase like
<i>Sec11c</i>	SEC11 homolog C (S. cerevisiae)

(b) The *C18QTL4*-residing interval: (69 genes)

<i>Smad7</i>	SMAD family member 7
<i>Ctif</i>	CBP80/20-dependent translation initiation factor
<i>LOC100911356</i>	peroxisomal biogenesis factor 19-like

<i>Zbtb7c</i>	zinc finger and BTB domain containing 7C
<i>Smad2</i>	SMAD family member 2
<i>Ctdp1</i>	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) phosphatase, subunit 1
<i>Kcng2</i>	potassium voltage-gated channel, subfamily G, member 2
<i>Pqlc1</i>	PQ loop repeat containing 1
<i>Hsbp111</i>	heat shock factor binding protein 1-like 1
<i>LOC100911462</i>	thioredoxin-like protein 4A-like
<i>RGD1560212</i>	similar to DNA segment, Chr 18, Wayne State University 98, expressed
<i>Rbfa</i>	ribosome binding factor A
<i>Adnp2</i>	ADNP homeobox 2
<i>Pard6g</i>	par-6 partitioning defective 6 homolog gamma (C. elegans)
<i>LOC100911433</i>	28S ribosomal protein S21, mitochondrial-like
<i>Skor2</i>	SKI family transcriptional corepressor 2
<i>Hdhd2</i>	haloacid dehalogenase-like hydrolase domain containing 2
<i>Katnal2</i>	katanin p60 subunit A-like 2
<i>LOC100911993</i>	ubiquitin carboxyl-terminal hydrolase 48-like
<i>Pias2</i>	protein inhibitor of activated STAT, 2
<i>St8sia5</i>	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 5
<i>Loxhd1</i>	lipoxygenase homology domains 1
<i>Rnf165</i>	ring finger protein 165
<i>RGD1308601</i>	similar to hypothetical protein
<i>Haus1</i>	HAUS augmin-like complex, subunit 1
<i>Atp5a1</i>	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle
<i>LOC100912125</i>	HAUS augmin-like complex subunit 1-like
<i>Pstpip2</i>	proline-serine-threonine phosphatase-interacting protein 2
<i>Epg5</i>	ectopic P-granules autophagy protein 5 homolog (C. elegans)
<i>Siglec15</i>	sialic acid binding Ig-like lectin 15
<i>Slc14a1</i>	solute carrier family 14 (urea transporter), member 1
<i>Slc14a2</i>	solute carrier family 14 (urea transporter), member 2
<i>Sethp1</i>	SET binding protein 1
<i>Mrps211</i>	mitochondrial ribosomal protein S21-like
<i>LOC100909457</i>	uncharacterized LOC100909457
<i>LOC100909499</i>	uncharacterized LOC100909499
<i>Nfate1</i>	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
<i>Atp9b</i>	ATPase, class II, type 9B
<i>Sall3</i>	sal-like 3 (Drosophila)
<i>LOC100909795</i>	colorectal mutant cancer protein-like

<i>LOC100909830</i>	mRNA-decapping enzyme 2-like
<i>LOC100909868</i>	myotilin-like
<i>LOC100909899</i>	myotilin-like
<i>LOC100909761</i>	myotilin-like
<i>Galr1</i>	galanin receptor 1
<i>Mbp</i>	myelin basic protein
<i>Zfp236</i>	zinc finger protein 236
<i>Zfp516</i>	zinc finger protein 516
<i>Tshz1</i>	teashirt zinc finger homeobox 1
<i>Zadh2</i>	zinc binding alcohol dehydrogenase, domain containing 2
<i>Zfp407</i>	zinc finger protein 407
<i>Cndp1</i>	carnosine dipeptidase 1 (metallopeptidase M20 family)
<i>Cndp2</i>	CNDP dipeptidase 2 (metallopeptidase M20 family)
<i>Fam69c</i>	family with sequence similarity 69, member C
<i>LOC100359752</i>	hypothetical protein LOC100359752
<i>Cyb5a</i>	cytochrome b5 type A (microsomal)
<i>Fbxo15</i>	F-box protein 15
<i>Timm21</i>	translocase of inner mitochondrial membrane 21 homolog (yeast)
<i>LOC100365382</i>	hypothetical LOC100365382
<i>Neto1</i>	neuropilin (NRP) and tolloid (TLL)-like 1
<i>Cbln2</i>	cerebellin-like protein
<i>LOC688985</i>	similar to Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (38 kDa BFA-dependent ADP-ribosylation substrate) (BARS-38)
<i>RGD1559726</i>	similar to melanoma antigen family A, 6
<i>RGD1560813</i>	similar to H2A histone family, member V isoform 1
<i>Socs6</i>	suppressor of cytokine signaling 6
<i>Rtn</i>	rotatin
<i>LOC689166</i>	hypothetical protein LOC689166
<i>Cd226</i>	CD226 molecule
<i>Dok6</i>	docking protein 6

Footnote: Designated pseudogenes are not included. Genes in bold bear non-synonymous mutations and are further described in Tables 1-3 in the text.


```

T F G K G L H V R A A S V S A K G M S R
C185.L14 ACCCTGGA AAAAGTCTTCATGTCAGAGCTGCCTCAGTTTCTGCCAAGGGATGTCAGG 900
DSS ACCCTGGA AAAAGTCTTCATGTCAGAGCTGCCTCAGTTTCTGCCAAGGGATGTCAGG 900
*****

K G P I S A T F E K T G P I T P Q A K T
C185.L14 AAAGGGCCATCTCAGCAACCCAGAGAAAGACTGGGCCATAACCCCAAGCCAGACA 960
DSS AAAGGGCCATCTCAGCAACCCAGAGAAAGACTGGGCCATAACCCCAAGCCAGACA 960
*****

G R F Q R D M E T E S E D D E D S E D E
C185.L14 GSAAGGCCACAGAAAGGATATGSAAACCCAGTCAGAGGACGATCTTGACAGTGAAGATGAG 1020
DSS GSAAGGCCACAGAAAGGATATGSAAACCCAGTCAGAGGACGATCTTGACAGTGAAGATGAG 1020
*****

↓
M F V T V T T F Q A R F S G K S F Q V R
C185.L14 ATGCCAGTCACTGTGACCACTCCTCAGGCAAGGCCTTCCGGAAAGAGCCCTCAGGTCAGA 1080
DSS ATGCCAGTCACTGTGACCACTCCTCAGGCAAGGCCTTCCGGAAAGAGCCCTCAGGTCAGA 1080
*****

G T E A P A K E S S Q K G A P A V T F G
C185.L14 GSTACCTCTGCCCTGCCCAAGGAGTCATCCCAAAAAGGGGCTCTGGCAGTCAACCCCTGGG 1140
DSS GSTACCTCTGCCCTGCCCAAGGAGTCATCCCAAAAAGGGGCTCTGGCAGTCAACCCCTGGG 1140
*****

K T R P V A A Q A Q A G E F E T R S S E
C185.L14 AAGACAAGACCTGTGGCAGCCAGGCCAGGCAGGGAAGCCAGAGACCAGGAGCAGTGAG 1200
DSS AAGACAAGACCTGTGGCAGCCAGGCCAGGCAGGGAAGCCAGAGACCAGGAGCAGTGAG 1200
*****

E S E S D S G E T F A A G T L L T T S F A
C185.L14 GAGTCAGAGAGCCAGCAGTGGAGAGACACCAGCAGCTGGGACTCTAACTACAAGTCTGCC 1260
DSS GAGTCAGAGAGCCAGCAGTGGAGAGACACCAGCAGCTGGGACTCTAACTACAAGTCTGCC 1260
*****

↓
K V K P L G K E P Q V R P A S T A N L G
C185.L14 AAGGTGAACCTTTGGGAAAAAGCCCCAGGTCAGACCTGCCCTCCACTGCAAACTGGGG 1320
DSS AAGGTGAACCTTTGGGAAAAAGCCCCAGGTCAGACCTGCCCTCCACTGCAAACTGGGG 1320
*****

S S R K G A N F F C L G E V G S A V L K
C185.L14 TCATCTAIAAAAAGGAGCCACCCACCTGCTTGGGAAGGTGGGGTCAGCAGTTCTCAAG 1380
DSS TCATCTAIAAAAAGGAGCCACCCACCTGCTTGGGAAGGTGGGGTCAGCAGTTCTCAAG 1380
*****

V Q T G K E E E D S E S S E S E S D E D
C185.L14 GTCCAAAAGGGAAAGGAAGAGAGGACTCAGAGAGCAGCAGCGAGGAGTGGACAGTGAC 1440
DSS GTCCAAAAGGGAAAGGAAGAGAGGACTCAGAGAGCAGCAGCGAGGAGTGGACAGTGAC 1440
*****

↓
G A V N A A K A E P P G E K A T P A L F
C185.L14 GGGGCTGTGAACCCAGCCAAAGCCAAAGCCCCAGGGAIAAAAAGCCACTCCAGCACTCCCT 1500
DSS GGGGCTGTGAACCCAGCCAAAGCCAAAGCCCCAGGGAIAAAAAGCCACTCCAGCACTCCCT 1500
*****

R K T G P V A T Q V K T D K G K D H A E
C185.L14 CGGAAGACGGGGCTGTGGCCACCAGGTCAGACTGACAGGGCAAGACCAGCAGAG 1560
DSS CGGAAGACGGGGCTGTGGCCACCAGGTCAGACTGACAGGGCAAGACCAGCAGAG 1560
*****

S E E E S T D S E E E A A P A A S A A Q
C185.L14 AGCAGCGAGGAGTCCAGCCAGCAGGAGGAGGAGGAGCACCAGCAGCTTCTGCTGCTCAG 1620
DSS AGCAGCGAGGAGTCCAGCCAGCAGGAGGAGGAGGAGCACCAGCAGCTTCTGCTGCTCAG 1620
*****

A K P A F I K O H K A S P B E S T A A S
C185.L14 GCTAAACCACTCCGATAAAGCAGATGAAAGCTTCCCTAGGAAAAGGCAGGCTGCATCC 1680
DSS GCTAAACCACTCCGATAAAGCAGATGAAAGCTTCCCTAGGAAAAGGCAGGCTGCATCC 1680
*****

T T G A S A S E P R K A G T K T S S A S
C185.L14 ACCCAGGGGCGAGGGCCCTGCTCCCTGTAAGGCAGGAACAAGACCTCTTCAGCCAGC 1740
DSS ACCCAGGGGCGAGGGCCCTGCTCCCTGTAAGGCAGGAACAAGACCTCTTCAGCCAGC 1740
*****

L S S L A L P K G T Q K P D V D S E E
C185.L14 CTATCATCCCTGGCTCTGCCAAGGGCACCCAGAAAGCAGATGTGGACTCCTCTAGCCAG 1800
DSS CTATCATCCCTGGCTCTGCCAAGGGCACCCAGAAAGCAGATGTGGACTCCTCTAGCCAG 1800
*****

```

W E E E G A A P G T T G V Q G K S G G E
 C188.L14 TGGAGT CAGAAGSAGCTGCTCCTGGCACCACGGGTACAGGGAAAATCTGGGGGAAA 1860
 DSS TGGAGT CAGAAGSAGCTGCTCCTGGCACCACGGGTACAGGGAAAATCTGGGGGAAA 1860
 *
 G L O G R A A S G Q G V A F L R A Q E T
 C188.L14 GGCTCCAAAGGGAGCTGCTCCTCAGGACAAGGGGTGGCCCCACTGCAOCTCAGAAGACA 1920
 DSS GGCTCCAAAGGGAGCTGCTCCTCAGGACAAGGGGTGGCCCCACTGCAOCTCAGAAGACA 1920
 *
 G P S S A Q V E A T A Q E D S E S E S E E
 C188.L14 GGCCCTTCAGTCCCAAGTCAAAGCCACAGCACAGSAAAGCTCAGAGACACGCAAGS 1980
 DSS GGCCCTTCAGTCCCAAGTCAAAGCCACAGCACAGSAAAGCTCAGAGACACGCAAGS 1980
 *
 E S E E E E E D E T P A Q V M A L G R L
 C188.L14 GAATCCAGCAGTGAAGAAGAGATGAGACCCACAGCAGGTCATGGCCCTGGGGAGACTT 2040
 DSS GAATCCAGCAGTGAAGAAGAGATGAGACCCACAGCAGGTCATGGCCCTGGGGAGACTT 2040
 *
 P P A K A N P P P T K T F L A S A S G K
 C188.L14 CCTCCGGCCAAAGCCAAOCCACCTCCCACTAAGACACCTCTGGCATCTGCATCTGGAAAA 2100
 DSS CCTCCGGCCAAAGCCAAOCCACCTCCCACTAAGACACCTCTGGCATCTGCATCTGGAAAA 2100
 *
 A A A V V P P P F E G K A F A S T V Q N S
 C188.L14 GCGGCCGAGTAGTTCCTCCACCCAAAGGGAAAAGCACCTGCGAGTACTGTTTCAGAACAGC 2160
 DSS GCGGCCGAGTAGTTCCTCCACCCAAAGGGAAAAGCACCTGCGAGTACTGTTTCAGAACAGC 2160
 *
 T I E A R G Q R A V P A T G E A G A F A
 C188.L14 ACCATCTCTGCCAGGGGCCAGCGGGCTGTGCCGGCCACGGGAAAAGCAGGGGCCAGCA 2220
 DSS ACCATCTCTGCCAGGGGCCAGCGGGCTGTGCCGGCCACGGGAAAAGCAGGGGCCAGCA 2220
 *
 T Q A Q K G P M A G T G E D S E S E S E E
 C188.L14 ACCCAAGCACAGAAGGGGCCATGGCTGGCACAGGGGAGSACTCAGAGACACGAGTGA 2280
 DSS ACCCAAGCACAGAAGGGGCCATGGCTGGCACAGGGGAGSACTCAGAGACACGAGTGA 2280
 *
 E E S D S E E E T P A Q V K F V G E T S 780
 C188.L14 GA **GGAGTCTGACAGTGAAGA**AGAAAATCCAGCCAGGTCAAGCCCTGTGGGGAAAGACTCT 2322
 DSS GA-----AGAAAATCCAGCCAGGTCAAGCCCTGTGGGGAAAGACTCT 2340
 *
 Q V R A A S A F V K E S F N K G A Y S G
 C188.L14 CAGTCCAGAGTGCCTCGGCCCTGTCAAGGAGTCTCCTAACAAAGSAGCTTATTAGGA 2382
 DSS CAGTCCAGAGTGCCTCGGCCCTGTCAAGGAGTCTCCTAACAAAGSAGCTTATTAGGA 2400
 *
 T S R E T S P A T Q A O T S K T E D S
 C188.L14 ACCTCCAGGAAGACAGGGCCCTCAGCTACCCAGSCTCAGACAGGAAAGACAGAGACTCA 2442
 DSS ACCTCCAGGAAGACAGGGCCCTCAGCTACCCAGSCTCAGACAGGAAAGACAGAGACTCA 2460
 *
 E S S S E E S D S D R E I P P A I T P A
 C188.L14 GAGACAGTAGTGGGAATCTGACAGTGCACAGAGATACCAOCCAGCCATCAOCCAGCC 2502
 DSS GAGACAGTAGTGGGAATCTGACAGTGCACAGAGATACCAOCCAGCCATCAOCCAGCC 2520
 *
 Q D G I S Q S T R E K L S G L A F F E K
 C188.L14 CAGGATGGTATCTCCAGTCTACAGAAAGCAAATCTCAGTCTGGCACCCCCAGAGAAG 2562
 DSS CAGGATGGTATCTCCAGTCTACAGAAAGCAAATCTCAGTCTGGCACCCCCAGAGAAG 2580
 *
 S Y E S S E S S S D E D L F S G Q A I K
 C188.L14 AGCACAGAAGAGTCTCAGAGAGCAGTGTGAGGATCTGCCATCTGGCCAGGCCATTAAA 2622
 DSS AGCACAGAAGAGTCTCAGAGAGCAGTGTGAGGATCTGCCATCTGGCCAGGCCATTAAA 2640
 *
 S P E I S V N E N R G P A A P V P T F E
 C188.L14 TCCCTCCAATTCTGTCAACCCTAATGTGGTCCAGTGCOCAGTCCAACCCAGAA 2682
 DSS TCCCTCCAATTCTGTCAACCCTAATGTGGTCCAGTGCOCAGTCCAACCCAGAA 2700
 *
 Q H Q A V N T R K A Q A S G S T A Q S S
 C188.L14 CAACACCAGGCTGTGAACACAAAGGAAGGCCAGGCCCTCAGGCAGCACTGCCAGAGCTCC 2742
 DSS CAACACCAGGCTGTGAACACAAAGGAAGGCCAGGCCCTCAGGCAGCACTGCCAGAGCTCC 2760
 *

S S E E E D E D M I P A T Q P P T L A I
 C186.L14 TCCTCCGAGAGT GAGGATGAGGACATGATTCCTGCCACACAAACCCCCACCCCTTGCCATC 2802
 DSS TCCTCCGAGAGT GAGGATGAGGACATGATTCCTGCCACACAAACCCCCACCCCTTGCCATC 2820
 R T N V T T P T A L S Q T A A Q P S E S
 C186.L14 AGAACCAATGTGACTACGCCCAAGCCCTCTCACAAAACAGCCGCCAACCCAGCAAGAGT 2862
 DSS AGAACCAATGTGACTACGCCCAAGCCCTCTCACAAAACAGCCGCCAACCCAGCAAGAGT 2880
 E Q S E R M P K G K K P K T A S T Q I S
 C186.L14 GAGCAGTCTAGTCCGATGCCAAAAGGCAAGAAACCAAGACAGCGTCCACTCAGATCAGC 2922
 DSS GAGCAGTCTAGTCCGATGCCAAAAGGCAAGAAACCAAGACAGCGTCCACTCAGATCAGC 2940
 S A N E A L P V T L P Q S T P A Q S K T
 C186.L14 AGTCCCATGGAAGCACTCCCGTGACGCTTCCCGAGAGCACACTGCCAGTCCAAAAC 2982
 DSS AGTCCCATGGAAGCACTCCCGTGACGCTTCCCGAGAGCACACTGCCAGTCCAAAAC 3000
 T N K L G D P K L A E K Q Q L T P G Y P
 C186.L14 ACCAACAGCTGGGGGACCCAAACTTCTGAGAAGCAGCAGCTTACCCAGGCTACCCC 3042
 DSS ACCAACAGCTGGGGGACCCAAACTTCTGAGAAGCAGCAGCTTACCCAGGCTACCCC 3060
 K A F R S S E D S S U T S E S E E D A
 C186.L14 AAAGCCCCAGGAGCTCAGAGGACAGCAGTGCACCTTTCAGAGAGGAGGAGGATGCC 3102
 DSS AAAGCCCCAGGAGCTCAGAGGACAGCAGTGCACCTTTCAGAGAGGAGGAGGATGCC 3120
 K R F Q M S K E S Q R L D P D A S Q K E
 C186.L14 AAGAGACCCAGATGTCCAGTGTCCCAAAGGCTGGATCCAGACGCTTCCAGAGGAA 3162
 DSS AAGAGACCCAGATGTCCAGTGTCCCAAAGGCTGGATCCAGACGCTTCCAGAGGAA 3180
 T V V E E T F T E S S D D E M V A F S Q
 C186.L14 ACTGTGTAGAGGAGACCCCTACAGAAATCCAGGACGACAGAGATGGTGGCCCTCCACAG 3222
 DSS ACTGTGTAGAGGAGACCCCTACAGAAATCCAGGACGACAGAGATGGTGGCCCTCCACAG 3240
 S L L E G Y V T F G L T V A N S Q P E X
 C186.L14 TCCTCCTCTCAGGTTATGTGACTCCCGGCTTAACTGTGGCCAAATTCOCAGCCTCAAAA 3282
 DSS TCCTCCTCTCAGGTTATGTGACTCCCGGCTTAACTGTGGCCAAATTCOCAGCCTCAAAA 3300
 A T F R P D A N F L V S S A P A T K D N
 C186.L14 GCTACTCCTAGGCCAGAGCCCAACCCCTTGGTTCCCTGTCTCCAGCCACCAAGATAAC 3342
 DSS GCTACTCCTAGGCCAGAGCCCAACCCCTTGGTTCCCTGTCTCCAGCCACCAAGATAAC 3360
 P O G K O K S K E Q D S T A D T T L B X
 C186.L14 CCAGATGCCAAGCAGAAATCAAATCCCAAGACTCCACTCCAGACACCCACTCCGTAAA 3402
 DSS CCAGATGCCAAGCAGAAATCAAATCCCAAGACTCCACTCCAGACACCCACTCCGTAAA 3420
 T G R H E A S E G S T P Q E F K K F E X
 C186.L14 ACCGGTAGGAAAGAGGCTCCTCAGGATCCACACTCAGAAAGCCAAAGAGCCCAAGAAAG 3462
 DSS ACCGGTAGGAAAGAGGCTCCTCAGGATCCACACTCAGAAAGCCAAAGAGCCCAAGAAAG 3480
 S T L S S P A P A Q T L F N S I T Q R L
 C186.L14 AGCAGCTTGAGCTCACCAGCCCGSCACAGACTGCGAAGCAGCATCACCCAGCGCTC 3522
 DSS AGCAGCTTGAGCTCACCAGCCCGSCACAGACTGCGAAGCAGCATCACCCAGCGCTC 3540
 L E Q F W F L S E A Q V Q A S V M K V L
 C186.L14 CTGGAGCAGCCCTGGCCCTGAGTGAGGCACAGGTGCAGGCTCTGTGATGAAGTCTCTG 3582
 DSS CTGGAGCAGCCCTGGCCCTGAGTGAGGCACAGGTGCAGGCTCTGTGATGAAGTCTCTG 3600
 T E L L E Q E R Q R A T E A I K E S G R
 C186.L14 ACAGAACTGCTGGAGCAGGAAAGCAGAAAGCCACCCAGGCTATCAAGGAGTGGAAAG 3642
 DSS ACAGAACTGCTGGAGCAGGAAAGCAGAAAGCCACCCAGGCTATCAAGGAGTGGAAAG 3660
 K G G K R K L S G D Q V E A G A P F N K
 C186.L14 AAGGGCCAGAAACGGAAACTATCAGGGGACCAAGTAGAAGCTGGGGCCCAAGAAACAA 3702
 DSS AAGGGCCAGAAACGGAAACTATCAGGGGACCAAGTAGAAGCTGGGGCCCAAGAAACAA 3720
 K K E Q Q L A A G T S A G S P E K A S R
 C186.L14 AAGAGAAAGCAGCAGCTTGGGCTGGGACAAAGTCTGGTTCCOCAGAAAAGGCTCCAGG 3762
 DSS AAGAGAAAGCAGCAGCTTGGGCTGGGACAAAGTCTGGTTCCOCAGAAAAGGCTCCAGG 3780

```

T S K A E S K L N E G E A G G K G F G S
C188.L14 ACTTCCAAGGCCAAAATCAAACATAAACAAAGGGAGTGTGGTGGCAAGGGGAAGGGGTCT 3822
DSB ACTTCCAAGGCCAAAATCAAACATAAACAAAGGGAGTGTGGTGGCAAGGGGAAGGGGTCT 3840
*****
F V F Q G A K E K F E G K L G I K L E S
C188.L14 CCTGTCCCCAAGGAGCCAGGAGAAAGCCGAAAGCAAGCTGGGGATAAAGCTTGAGAGT 3882
DSB CCTGTCCCCAAGGAGCCAGGAGAAAGCCGAAAGCAAGCTGGGGATAAAGCTTGAGAGT 3900
*****
G E Q E D F K S E K E K K E S S K K B K
C188.L14 GGAGAGCAGAGTGCACCCGAGAGCAAGAAAGAAAAGAAATCCAGTAAGAAAAAGAAA 3942
DSB GGAGAGCAGAGTGCACCCGAGAGCAAGAAAGAAAAGAAATCCAGTAAGAAAAAGAAA 3960
*****
K E E R P Q P R T L P H R S R R R E R R
C188.L14 AAGSAAAAGAGACCCCAAGCAAGSACTCTGCCTCACCAGTCCAGAGAGAAAAGAAAGA 4002
DSB AAGSAAAAGAGACCCCAAGCAAGSACTCTGCCTCACCAGTCCAGAGAGAAAAGAAAGA 4020
*****
R R R Q P S L L C E G Q A A H R T E S M
C188.L14 AGAAGAAAGACAGCCGAGCCTGTGTGTGAGGGCCAGGCAGCCATAGAACAGAGTCTTGG 4062
DSB AGAAGAAAGACAGCCGAGCCTGTGTGTGAGGGCCAGGCAGCCATAGAACAGAGTCTTGG 4080
*****
Q S V T I P R F L T S G H C E L G L D F
C188.L14 CAGAGCCTGACCATCCCCAGGCCTCTGACCTCTGGGCACCTGTGAGTTGGGACTGGACTTT 4122
DSB CAGAGCCTGACCATCCCCAGGCCTCTGACCTCTGGGCACCTGTGAGTTGGGACTGGACTTT 4140
*****
L I C F P S A E G D G W R C C -
C188.L14 TTAATTTGCCCTCCCTCGGCAGAAAGGTGATGGCTGGCCTTGTGCTAG 4170
DSB TTAATTTGCCCTCCCTCGGCAGAAAGGTGATGGCTGGCCTTGTGCTAG 4188
*****

```

Footnote to Supplement: * indicates nucleotide identity. Amino acid sequence is given on top. The deleted nucleotides are indicated by --- and in bold. ↓ marks the last nucleotide of each exon.

Supplement 5. Primer sequences used for reverse-transcriptase polymerase chain reactions (RT-PCR) in assessing gene expressions and/or alternative splicings.

QTL	Genes	Forward primers (5'- 3')	Reverse primers (5'- 3')	Product size (bp)
<i>C18QTL3</i>	<i>Tcof1</i>	CAACCCACCTCCCCTAAGA	GGTGATGGCTGGTGGTATCT	460
	<i>Cdx1</i>	AGAGCTGGCTGCTAACTTGG	CATTGGTGGGGCATAGACTC	212
	<i>Oacyl</i>	GTTGGTTAAGTGCACGGTCA	CACGCAGTGCCAGAGAAAA	960
	<i>Sec11c</i>	GGACATCTTTGGGGACTTGA	AGCCAGTTCTGGCCTTCTTT	387
<i>C18QTL4</i>	<i>Hdhd2</i>	AGGAGATGATTGCAGGGATG	GCTCTCGCACGTTAGGTAGG	780
	<i>Pqlc1</i>	GACTCAGAGAGCCCAAAGGA	GACGCAGCGCTCAACTTAAT	960
	<i>Rbfa</i>	CTCCTGCGGAAGCAAGAAGT	CATCTCCTGACCCGGAAGTA	924
	<i>Adnp2</i>	CCATCGCCAGTAGTCAGGAT	CACCTTGATTCACAGGGACA	950
	<i>Loxhd1</i>	ATTGACCTGGGGGCTCTTAC	AAGGGACAAATGGCGTGTAG	685
	<i>Atp9b</i>	GCGGACCAGATTCCTCTGTA	TGGACATCCAAGCCATACTG	366
	<i>Sall3</i>	GGAAGTAAACGGTCACAGCAA	CCTCAATGAACCTTGTGAACG	281
Control	<i>β-actin</i>	ACTGCCGCATCCTCTTCCTC	CCGCTCGTTGCCAATAGTGA	118

Footnote: Genes are given in the legends of Tables 1 and 3 and are chosen because they contain structural or intron-exon junction SNPs. Chromosome regions harbouring *C18QTL3* and *C18QTL4* are defined in Fig. 1. bp, base pairs.

Supplement 6. Aligned coding sequence comparisons of genes with single nucleotide polymorphisms (SNPs) in intervals harbouring *C18QTL4* and *C18QTL3* between Dahl salt-sensitive (DSS) and Lewis alleles.

(a). Genes in the *C18QTL4*-residing interval

(1) *CBP80/20-dependent translation initiation factor (Ctlf)*

```

Lewis      M K E P P S D E G E A G S H R H A K E T
DSS        ATGAAGCATCCGCCAGGAGACAAGGGAGAGCGGGCAGCCACCISCAACISGCAAAGAGACC 60
*****
          V T V E N F K L E D G F Q D T G H S G L
Lewis      GTGACCSTTGAGAACCCCAAACCTTGAAGACGGCCACGGGACACCGGACACAGTGGCCTT 120
DSS        GTGACCSTTGAGAACCCCAAACCTTGAAGACGGCCACGGGACACCGGACACAGTGGCCTT 120
*****
          E F P C S P D T L T P T A S E R E P T F Q
Lewis      SAGCCCCCTGCAGCCCTGACACTCTGACCCCAACAGCTTCCGAGAGGCCAACCCCTCAG 180
DSS        SAGCCCCCTGCAGCCCTGACACTCTGACCCCAACAGCTTCCGAGAGGCCAACCCCTCAG 180
*****
          L P G G S E A E T K Q K D T A L P E R L
Lewis      CTGCCAGGGGTTTCAGAGSCTGAGACAAAGCAAAGACACTGCTCTTCCCGAGGCTCTT 240
DSS        CTGCCAGGGGTTTCAGAGSCTGAGACAAAGCAAAGACACTGCTCTTCCCGAGGCTCTT 240
*****

          G E R P K I T L L Q S S K D R L R R R L
Lewis      GGGAGCGGCCCAAATCACCTGCTCCAGTCTTCCAAAGACAGSCTGAGGCGAAGGCTG 300
DSS        GGGAGCGGCCCAAATCACCTGCTCCAGTCTTCCAAAGACAGSCTGAGGCGAAGGCTG 300
*****
          |
          K E K V P D E V A V E A S N P Q P H K M
Lewis      AAAGAAAAGGTACCCGACGAGTAGCAGTAGAGGCCAGCAACCTCAGCCAAACAGATG 360
DSS        AAAGAAAAGGTACCCGACGAGTAGCAGTAGAGGCCAGCAACCTCAGCCAAACAGATG 360
*****

          D R L T E I L N S M R H N S S D V D A K
Lewis      GACAGCTGACGGAGATCCTGAACAGCATGAGGAATAACAGCAGTGAAGTGGACGCCAAG 420
DSS        GACAGCTGACGGAGATCCTGAACAGCATGAGGAATAACAGCAGTGAAGTGGACGCCAAG 420
*****

          L T S F H E E A Q N S T H S E E M L G E
Lewis      CTCACTCCTTCATGGAGGAAGCGCAGAACTCCACCAACTCCGAGGAGATGCTGGGAGAG 480
DSS        CTCACTCCTTCATGGAGGAAGCGCAGAACTCCACCAACTCCGAGGAGATGCTGGGAGAG 480
*****

          I V B T I Y Q E A V S D R S F A P T A A
Lewis      ATGTCGGACCATCTACCGAAGGCCGTGTGACAGCCGACGCTTGGCTTTCAGGCGCGCC 540
DSS        ATGTCGGACCATCTACCGAAGGCCGTGTGACAGCCGACGCTTGGCTTTCAGGCGCGCC 540
*****

          H L C D E M A L F M V E G T E F R S L L
Lewis      AAGCTCTGGACAAAGATGGCCCTTTCATGGTGGAGGGACCAAGTTCGGAGCCTGCTC 600
DSS        AAGCTCTGGACAAAGATGGCCCTTTCATGGTGGAGGGACCAAGTTCGGAGCCTGCTC 600
*****
          |
          L N M L Q K D F T V R E E L Q Q Q D V E
Lewis      CTCAACATGCTCCAGAGSACTTCACTGTGCGTGGAGSACTACAGCAACAGGATGTGGAG 660
DSS        CTCAACATGCTCCAGAGSACTTCACTGTGCGTGGAGSACTACAGCAACAGGATGTGGAG 660
*****

          R W L G F I T F L C E V F G T M R S S T
Lewis      CGCTGGCTGGGCTTTCATCACTTCTGTTGTGAGGTGTTGGCAACATGGGGAGCAGCACT 720
DSS        CGCTGGCTGGGCTTTCATCACTTCTGTTGTGAGGTGTTGGCAACATGGGGAGCAGCACT 720
*****
          |
          G E P F R V L L L Q S Q D V K E D A V L
Lewis      GGGAGCCCTTCCGAGTGTCTGCTTTCAGTCTCAGGATGTGAAGGAAGACGCTGTCTC 780
DSS        GGGAGCCCTTCCGAGTGTCTGCTTTCAGTCTCAGGATGTGAAGGAAGACGCTGTCTC 780
*****

```

```

      C C S N E L Q S T B R L L E E Q L F E M
Lewis  TGCTGCTCCATGGAGCTGCAAAGCACAGGCCGGCTACTAGAGGAGCAGCTGCCBAGATG 840
DSS    TGCTGCTCCATGGAGCTGCAAAGCACAGGCCGGCTACTAGAGGAGCAGCTGCCBAGATG 840
      *
      M T E L L A S A R D E M L C F E E S M L
Lewis  ATGACGGAGCTCCTAGCCACGCGCGGAGATAAGATGCTGTGCCCTCAGAGTCCATGCTA 900
DSS    ATGACGGAGCTCCTAGCCACGCGCGGAGATAAGATGCTGTGCCCTCAGAGTCCATGCTA 900
      *
      T R E L L L E V I E L H A N S W N P L T
Lewis  ACACGCTCCCTGCTCCTGGAGTTCATCGAGCTCCACGCCAATAGCTGGAACCCCTCGAG 960
DSS    ACACGCTCCCTGCTCCTGGAGTTCATCGAGCTCCACGCCAATAGCTGGAACCCCTCGAG 960
      *
      P P I T Q Y Y N R T I Q K L T A -
Lewis  CCCCCATCACGCACTACTACAACAGGACCATCCAGAACTGACAGCTGA 1011
DSS    CCCCCATCACGCACTACTACAACAGGACCATCCAGAACTGACAGCTGA 1011
      *

```

(2) Ribosome binding factor A (Rbfa)

```

      M W A G V A G L R G S C A G L O A L M G
Lewis  ATGTGGCCCGGCTCCGCGGGCTGAGAGGCTCCTCCGCGGGCTGCAGGCTCTGTGGGGC 60
DSS    ATGTGGCCCGGCTCCGCGGGCTGAGAGGCTCCTCCGCGGGCTGCAGGCTCTGTGGGGC 60
      *
      G H R A A L L L G R S F A L H T S V A S
Lewis  GGCCACAGGGCTGCGCTTCTTCGCGCGCTCGCCGCGCCTGCACACCTCGGTGCGCTCC 120
DSS    GGCCACAGGGCTGCGCTTCTTCGCGCGCTCGCCGCGCCTGCACACCTCGGTGCGCTCC 120
      *
      C G E K N L E K F A S E T R K K F M Y
Lewis  TGGGCGCAGCAAACTTGTCAAGAAATTTGCTTGGAAAACAGGAAGAAATTTGGTAT 180
DSS    TGGGCGCAGCAAACTTGTCAAGAAATTTGCTTGGAAAACAGGAAGAAATTTGGTAT 180
      *
      E G P E L G S H L T P R F S E R H E F L T
Lewis  GAAGTCCCTCCCTGGGCTCACITGACCCCCAGGCCATCCAAGCATGAGTTCCTCAAG 240
DSS    GAAGTCCCTCCCTGGGCTCACITGACCCCCAGGCCATCCAAGCATGAGTTCCTCAAG 240
      *
      K N T L K K T R K E E T I R L R V L N G
Lewis  AAGAACAAGTTAAAGAAGACTAGGAAGGAGGACACTATACGCTAAGGTCCTGSAACGGC 300
DSS    AAGAACAAGTTAAAGAAGACTAGGAAGGAGGACACTATACGCTAAGGTCCTGSAACGGC 300
      *
      L L H K S L T E L L C T F E V S Q E V Y
Lewis  CTTCTCCACAAGTCAATGACAGAGCTGCTGTGCACTCCTGAAGTGAAGCAAGAGTGTAT 360
DSS    CTTCTCCACAAGTCAATGACAGAGCTGCTGTGCACTCCTGAAGTGAAGCAAGAGTGTAT 360
      *
      D L N V E L S E V E V T P D F S A C R V
Lewis  GACCTGAATGTCGAGCTCTCCAAGSTCTCTGTGACTCCAGACTTCTCAGCTGCGGAGTG 420
DSS    GACCTGAATGTCGAGCTCTCCAAGSTCTCTGTGACTCCAGACTTCTCAGCTGCGGAGTG 420
      *
      Y W K T G V S A E Q N R H T E A V L Q R
Lewis  TACTGGAAGCAGTGTCTCTGCAGAGCAGAACAGGCACACGGAGGCCCTCTGCAGAGG 480
DSS    TACTGGAAGCAGTGTCTCTGCAGAGCAGAACAGGCACACGGAGGCCCTCTGCAGAGG 480
      *
      S A Y M R H L L I S Q Q T L R N V P F
Lewis  AGGCTCCCTACATGAGGCATCTTCTGATATCTCAGCAAAACCTGAGAAATGTTCCACCC 540
DSS    AGGCTCCCTACATGAGGCATCTTCTGATATCTCAGCAAAACCTGAGAAATGTTCCACCC 540
      *
      I V F V Q D X R D I V L A E V D S L L A
Lewis  ATAGTGTGTTTCAAGACAAGAGACATAGTCTTGGCTGAGGTAGATCGTTACTGGCT 600
DSS    ATAGTGTGTTTCAAGACAAGAGACATAGTCTTGGCTGAGGTAGATCGTTACTGGCT 600
      *

```

```

          V A D F G P F D E E D D L D G L R E P D
Lewis  679GCTGACTTTGGACCCDAGATGAAAGGATGACTTGGATGTTTGAGGAGCCCGAT 660
DSS    679GCTGACTTTGGACCCDAGATGAAAGGATGACTTGGATGTTTGAGGAGCCCGAT 660
          *
          A Q A F H D E P E P T T H F N L C G I D
Lewis  680CAGGCAACCATGACTCCCCAGAACCAACACACCCAAATCTGTGTGGATTGAC 720
DSS    680CAGGCAACCATGACTCCCCAGAACCAACACACCCAAATCTGTGTGGATTGAC 720
          *
          H E A L N K Q I M E Y E B K E E R G L Q
Lewis  681CAGGCACTGAATRAGCAAATAATGGAGTACAAACGCAAGAAGGAGCGAGGCTCCAG 780
DSS    681CAGGCACTGAATRAGCAAATAATGGAGTACAAACGCAAGAAGGAGCGAGGCTCCAG 780
          *
          C V S L A F F S G Q E Q A F D F A H L L
Lewis  682TGTGTGAGCTTGGCCACCATCAGGGCAGGAGCAGCCGCTGATCCCGCTCATTGTCT 840
DSS    682TGTGTGAGCTTGGCCACCATCAGGGCAGGAGCAGCCGCTGATCCCGCTCATTGTCT 840
          *
          R E R E K A R E R Q H Q D A S P R E F L
Lewis  683CGAAAGAAAGAAAGGCCGATCCCGCCAGCACAGGATGCTTCCCCAGAAAGCTTCCT 900
DSS    683CGAAAGAAAGAAAGGCCGATCCCGCCAGCACAGGATGCTTCCCCAGAAAGCTTCCT 900
          *
          L G E E D E D E D S E T E M E C H A Q E
Lewis  684TTGGTGAAGGATGAGSATGAGGACAGCAGCACAGAGTGGAAATGCCACGCCAGGAG 960
DSS    684TTGGTGAAGGATGAGSATGAGGACAGCAGCACAGAGTGGAAATGCCACGCCAGGAG 960
          *
          *
          A E D E W E A E E G G E G V Q Q G L B G
Lewis  685GCTGAGGACGAGTGGGAGCAGAGSSTGGAGTGGAGGCTCCAGAGGCTCGSSTGGC 1020
DSS    685GCTGAGGACGAGTGGGAGCAGAGSSTGGAGTGGAGGCTCCAGAGGCTCGSSTGGC 1020
          *
          K R E Q G -
Lewis  686AAAAGAGAGCAGGGGTAA 1038
DSS    686AAAAGAGAGCAGGGGTAA 1038
          *

```

(3) ADNP homeobox protein 2 (Adnp2)

```

          H F Q I P V Q N L D N I R E Y R K E V K
Lewis  687ATGTTTCAAATTCCTGTGCAAAATCTTGACAACATCAGAAAGTSCGGAGAGGGTGA 60
DSS    687ATGTTTCAAATTCCTGTGCAAAATCTTGACAACATCAGAAAGTSCGGAGAGGGTGA 60
          *
          G I L V D I G L D S C E E L L K D L K G
Lewis  688GGCATCCTTGTGGACATGGACTTGACAGCTGCAAGGAGCTGCTGAAGSATCTTAAAG 120
DSS    688GGCATCCTTGTGGACATGGACTTGACAGCTGCAAGGAGCTGCTGAAGSATCTTAAAG 120
          *
          F D F G E K Y F C H T S W G D V S L W E
Lewis  689TTTATCCGGAGAGAGTACTTTTGTAAATACATGTTGGGAGATGTTTCTTTGGAA 180
DSS    689TTTATCCGGAGAGAGTACTTTTGTAAATACATGTTGGGAGATGTTTCTTTGGAA 180
          *
          P S G K R A K Y R T K F Y C C S L C R Y
Lewis  690CCTTCTGGAAGAGAGCGAAATACAGAACAAAGCCCTACTGCTGTAGTCTCTGCAGGTAC 240
DSS    690CCTTCTGGAAGAGAGCGAAATACAGAACAAAGCCCTACTGCTGTAGTCTCTGCAGGTAC 240
          *
          E T E V L T S I E N H L H R Y H E E E A
Lewis  691TCAACGAAGTGTCTACCTCCCTCAAAAATCACTGCAACCGATACCCGAAAGAGGCT 300
DSS    691TCAACGAAGTGTCTACCTCCCTCAAAAATCACTGCAACCGATACCCGAAAGAGGCT 300
          *
          D Q E L M I E C P N C E F A S Q E V V
Lewis  692GACCAGGACTCATGATCCCTGCCCAACTGCCCCCTTGGCTCTCAGCCCAAGGTTGT 360
DSS    692GACCAGGACTCATGATCCCTGCCCAACTGCCCCCTTGGCTCTCAGCCCAAGGTTGT 360
          *

```


Lewis G K H F R M F H A F A R K V Q S Y T V N
 DSS GGCAAGCACITTCAGAAATGTTCCACGGCCTGCCGGGAAAGTCCAGAGCTACACAGTGAAC 420
 GGCAAGCACITTCAGAAATGTTCCACGGCCTGCCGGGAAAGTCCAGAGCTACACAGTGAAC 420

Lewis I L G E A K T E R S D V I S F T C L K C
 DSS ATCCTGGGTGAGGCGAAGACTTCAAGSAGTGAATGTGATAAGCTTCACATGTTTAAATGT 480
 ATCCTGGGTGAGGCGAAGACTTCAAGSAGTGAATGTGATAAGCTTCACATGTTTAAATGT 480

Lewis H F E N T L Y Y S M K K H V L V A H F N
 DSS AACTTTTCAAACTCTGTACTACAGCATGAAGAAGCATGTGCTGGTGGCCATTTTAAT 540
 AACTTTTCAAACTCTGTACTACAGCATGAAGAAGCATGTGCTGGTGGCCATTTTAAT 540

Lewis Y L I N E Y F G L R T E E T G E Q F K A
 DSS TACTTAATTAACCTCTACTTTGGATTGCGAAGTGGGAAACAGGAGAACAAACCBAAAGCA 600
 TACTTAATTAACCTCTACTTTGGATTGCGAAGTGGGAAACAGGAGAACAAACCBAAAGCA 600

Lewis S D P V S V D E A L P F F E R Y Y C K K C
 DSS AATGATCCAGTTTCTGTGGATAAAGCCCTGCCATTGACAAAGTACTACTGTAAAAAAATGC 660
 AATGATCCAGTTTCTGTGGATAAAGCCCTGCCATTGACAAAGTACTACTGTAAAAAAATGC 660

Lewis S A I A S S Q E A L M Y H I L T S D A H
 DSS AGGSCATGCGCCAGTACTAGGATGCCCTGATGTATCACATCTGCATCAGATGCACAT 720
 AGGSCATGCGCCAGTACTAGGATGCCCTGATGTATCACATCTGCATCAGATGCACAT 720

Lewis R D L E N K L R S V I S E H I K R T G F
 DSS AGGSACTTGGAGAATAAGCTGAGGTCTGTATCTCAGASACATCAAGAGSACCGGGTTT 780
 AGGSACTTGGAGAATAAGCTGAGGTCTGTATCTCAGASACATCAAGAGSACCGGGTTT 780

Lewis L E G M H I A P K P V T H I A L P P N S
 DSS CTGAAGCAAATGCATATGCTCCAAAGCCAGTGAOCCACATAGCTTTACCGCCAAACAGC 840
 CTGAAGCAAATGCATATGCTCCAAAGCCAGTGAOCCACATAGCTTTACCGCCAAACAGC 840

Lewis S A P S I A A P P F C F Q L A L P Q N S
 DSS AGTGGTCCGAGCATTGCAGCCCTCCTCCTTCCAGCTTGGCTTGGCCACAGAACAT 900
 AGTGGTCCGAGCATTGCAGCCCTCCTCCTTCCAGCTTGGCTTGGCCACAGAACAT 900

Lewis Q S F G T V Q S V T V A F G T S G S L T
 DSS CAAAGTCCCGCACTGTGCAGTCAAGTACTGTGSCCCAGGCACTTCTGGAGCCTTACA 960
 CAAAGTCCCGCACTGTGCAGTCAAGTACTGTGSCCCAGGCACTTCTGGAGCCTTACA 960

Lewis H S P P T T A Q S H V A L V S S L P V
 DSS CACTCACCACTACCAGGSCAGTCTCATGTAGCTCTGGTCTCCAGCTCTTGGCTGTG 1020
 CACTCACCACTACCAGGSCAGTCTCATGTAGCTCTGGTCTCCAGCTCTTGGCTGTG 1020

Lewis C Q S S L T L Q Q S A P P P V F L S H S
 DSS TSCCAGAGTACCTCACCTGSCAGTCCGCTCCGCCACCTGTCTTCTCTCACAGT 1080
 TSCCAGAGTACCTCACCTGSCAGTCCGCTCCGCCACCTGTCTTCTCTCACAGT 1080

Lewis V F L S Q P V S T S V L F L T Q P L G F
 DSS GTCCCACTGAGTCAGCCCTCAGTACTTCTGTGCTGCTCCTCACTCAGCCACTTGGGCCT 1140
 GTCCCACTGAGTCAGCCCTCAGTACTTCTGTGCTGCTCCTCACTCAGCCACTTGGGCCT 1140

Lewis V N E S V G T S L L P V N Q A H C S V N
 DSS GTGAATAAGTCTGTTGGAAACAAGCCTCCTCCCTGTGAACCAAGCCATGTGCTCCGTGAAC 1200
 GTGAATAAGTCTGTTGGAAACAAGCCTCCTCCCTGTGAACCAAGCCATGTGCTCCGTGAAC 1200

Lewis Q A V R E G V L P L F K P M G F I N R P
 DSS CAAGCTGTCCGCTGGAGTTTTACCCCTCCCTAAGCCCATGGSSCCATAAACAGACT 1260
 CAAGCTGTCCGCTGGAGTTTTACCCCTCCCTAAGCCCATGGSSCCATAAACAGACT 1260

Lewis V G E G V L P V G P S V N S E V L Q A T 1320
 DSS GTGGCCCTGGTGTCTTGCCCTGTGGGTCCCTCTGTAACTCAGGGGTTCCTCCAGGCTACA 1320

 Lewis E P G V I S V G R A V P S G V L P A G Q 1380
 DSS TCTCCTGGGGTGAATTCGTAGGTGAGCAGTTCATCAGGAGTCTTCTGCGAGGTCAG 1380

 Lewis V T F A S V I P S Q T A T S E V L F T S 1440
 DSS GTGACCCCTGCTGTGTGATCCCTGGGCAGACAGCCACTTCGGGGTCTTACCCACTGSC 1440

 Lewis Q V V Q E S V L P V G Q T A F S R V L F 1500
 DSS CAGTGGTCCAGTCTGTCAGTTCCTCCTGTGGCCAGACAGCCCATCTGAGTTCCTCCT 1500

 Lewis P G Q T V P L E V L P A G Q V V P F G L 1560
 DSS CTTGGCCAGACAGTCCCTTGAGGGTTCCTCCTGCGAGCCAGTGTGACCCCTGGGCTG 1560

 Lewis L S P N Q T V P S G V V P V N Q S V N S 1620
 DSS CTTTCTCAAACCAAACGTTCCTCCTGGGAGTGTCTCCTGTGAATCAAGTGTSAACTCT 1620

 Lewis G V L Q L S Q F V T P G V L P V G F P V 1680
 DSS GGTGTTCTTCAGTCTCAGTCCAGTAAACACAGGAGTCTTCTGTTGGGCCACCCCGT 1680

 Lewis R P G V L Q L S P S V S T S I L P V S Q 1740
 DSS ABGCTTGGTGTCTGACGCTCAGTCCGTCTGTCAACACAGCATCTGCCCCGTGAGCCAG 1740

 Lewis P V R A G T S Q N T T F L T S G S I L R 1800
 DSS CCGGTGAGAGCTGSAACGTCCAAAACACTACTTTCCTTACTCAGGTTCTATTCTCASA 1800

 Lewis Q L I P T G K Q V N G I P T Y T L A P V 1860
 DSS CAGCTCATTCCAACCTGGGAACAGGTGAATGGAATCCCCACTATACGCTGGCCCCAGTG 1860

 Lewis S V T L P V P S A G G L A A V G P F P Q 1920
 DSS TCCCTCAGTCTGCGGGTCCCTCCGCTGAGGCTTSCAGTGTGGACCGCCACCCAG 1920

 Lewis V P V Q F L P S S S G T Q M G S S L P S 1980
 DSS GTGCGTGTGAGTTCCTGCCCTCAAGCTCGGGCACACAGATGGGCAGTCTCTGCCCAGC 1980

 Lewis L P S P Q V L V S P A P S V P V Q A T P 2040
 DSS CTGCCCTGACCACAGGTGCTAGTGAAGCCCTGCCCTAGGGTGTGTTGTCAGGCTACCCCG 2040

 Lewis P V A D A N Q A L R Q A K Q M K T C P V 2100
 DSS CCTGTGGCAGATGCAAATCAGGCACTCAGACAGSCCAASCASTGGAAAACATGCCAGTT 2100

 Lewis C N E L F P A N V Y Q V H M E V A H T Q 2160
 DSS TSCAACGAGCTCTCCCTGCAAAGCTCTACAGSTCCACTGGAAGTGTCTCACACGAG 2160

V E

```

Lewis      E E A K S E E K P E P E R L A A C A P F
DSS        AGCGAGGCCAAGTCCAGTGGAGAAGCCTGAGCCCGAAAGSCTTGCTGCATGCGCCDCCATT 2220
          AGCGAGGCCAAGTCCAGTGGAGAAGCCTGAGCCCGAAAGSCTTGCTGCATGCGCCDCCATT 2220
          ..

Lewis      L E W M R E E T V R C L S C E C L V S Q
DSS        CTGAAGTGGATGAGAGAGAAGACASTGCGCTGCTCTTGTAAAGTGCCTGGTCTCCGAG 2280
          CTGAAGTGGATGAGAGAGAAGACASTGCGCTGCTCTTGTAAAGTGCCTGGTCTCCGAG 2280
          ..

Lewis      E E L M R H L L M H G L G C L F C P C T
DSS        GAGGAGCTGATGACCACTTTGCTCATGSCATGGCTGGGGTGCCTGTCTGTCCATGCACT 2340
          GAGGAGCTGATGACCACTTTGCTCATGSCATGGCTGGGGTGCCTGTCTGTCCATGCACT 2340
          ..

Lewis      F H D V R G L V E H S R T K H L S K E R
DSS        TTTGATGATGTCGCGGGCCTTGTGGAGCACAGCAGGACTAAGCACCTGGGCAAGAGAGA 2400
          TTTGATGATGTCGCGGGCCTTGTGGAGCACAGCAGGACTAAGCACCTGGGCAAGAGAGA 2400
          ..

Lewis      L S M D Y S N R G F Q L D L D A N G N L
DSS        CTGTCTATGGATTACAGTAAACAGAGSTTCCAGCTGGACTTGGATGCTAATGGSAACCTG 2460
          CTGTCTATGGATTACAGTAAACAGAGSTTCCAGCTGGACTTGGATGCTAATGGSAACCTG 2460
          ..

Lewis      L F P H L D F I T I L P R E K L G E R E
DSS        CVGTCCCTCATCTCGATTTCATCACCATACTGCCACGGAGAAACTTGGAGAGGAGAA 2520
          CVGTCCCTCATCTCGATTTCATCACCATACTGCCACGGAGAAACTTGGAGAGGAGAA 2520
          ..

Lewis      V Y L A I L A G I H E E K S L V F V Y V E
DSS        GTGTACCTGGCTATCCTTGTGGAATACACTCCAAGTCTTGGTCCCTGTGTAGSTTAAG 2580
          GTGTACCTGGCTATCCTTGTGGAATACACTCCAAGTCTTGGTCCCTGTGTAGSTTAAG 2580
          ..

Lewis      V R F Q F E V A P K I P N R Q H L T C F
DSS        GTGAGGCTCAGCTGAGSTTGCACCAAGATACCTAACAGACAGAGCTGACCTGCCCG 2640
          GTGAGGCTCAGCTGAGSTTGCACCAAGATACCTAACAGACAGAGCTGACCTGCCCG 2640
          ..

Lewis      F C F G T F M A A D A Y E L H L K E R H
DSS        TTTCTTTTGGCACATTCATGGCTGCTGATGCTACAGCTGSCATCTGAAGGAGAGGCAC 2700
          TTTCTTTTGGCACATTCATGGCTGCTGATGCTACAGCTGSCATCTGAAGGAGAGGCAC 2700
          ..

Lewis      H V M P T V H T M L R S P A F K C I H C
DSS        CATGTGATGCCACAGTCCATACAATGCTCCGGTCTCCGGCCTTTAAGTGCATCCACTGT 2760
          CATGTGATGCCACAGTCCATACAATGCTCCGGTCTCCGGCCTTTAAGTGCATCCACTGT 2760
          ..

Lewis      C G V Y T G N M T L G A I A V H L L R C
DSS        TGTGGGCTTACACTGGAACATGACCTTAGGAGCCATGCTGTCATTGCTCCGCTGT 2820
          TGTGGGCTTACACTGGAACATGACCTTAGGAGCCATGCTGTCATTGCTCCGCTGT 2820
          ..

Lewis      R S A P K D S E S D L Q A Q P D F I E S
DSS        AGAAGTGTCCAAAGGACAGCAGCTCAGACCTGCAAGCCAGCCAGATTTTATCGAGAGC 2880
          AGAAGTGTCCAAAGGACAGCAGCTCAGACCTGCAAGCCAGCCAGATTTTATCGAGAGC 2880
          ..

Lewis      E E L L M V N G E V I P E S T F P L K R
DSS        AGTGAAGTGTGATGGTCAATGGGGAAGTGTATCCCGAGTCCACCTTTCTCTGAGAGA 2940
          AGTGAAGTGTGATGGTCAATGGGGAAGTGTATCCCGAGTCCACCTTTCTCTGAGAGA 2940
          ..

Lewis      K L F E G H L G P E S Q G D G D E F Q L
DSS        AAGTGTCCAGAGGCCATTTAGGGCCAGAGAGCAGGGGACGGGACGAGCCCCAGCTC 3000
          AAGTGTCCAGAGGCCATTTAGGGCCAGAGAGCAGGGGACGGGACGAGCCCCAGCTC 3000
          ..

Lewis      T V D T D A S P S E E K G L S A V P L E
DSS        ACAGTAGACACCGATGCGAGCCAGGTTGAGAGAAAGSCTGAGTCTGTGCTTTGAG 3060
          ACAGTAGACACCGATGCGAGCCAGGTTGAGAGAAAGSCTGAGTCTGTGCTTTGAG 3060
          ..

```

```

Lewis      R Q E N E S R T E G S G A S D D S L Q V
DSS
AGACAGAAGAAATGAGAGCAGGACAGAGGGGTCAGGGGCCAGTGATGACTCCCTGCAAGTG 3120
AGACAGAAGAAATGAGAGCAGGACAGAGGGGTCAGGGGCCAGTGATGACTCCCTGCAAGTG 3120
*****

Lewis      L A L D P S K Y G S R S Y E E K K Q F L
DSS
TTGGCTTAGACCCCAATAAGTATGGAAAGTCCTTCTCTATGAGGAAAAGAAACAGTTCTCT 3180
TTGGCTTAGACCCCAATAAGTATGGAAAGTCCTTCTCTATGAGGAAAAGAAACAGTTCTCT 3180
*****

Lewis      R D Y F H K R P Y P S R K E V E L L S S
DSS
AGAGACTATTTTCCACAGAGACCATATCCTAGTGGAAAAGAAAGTGGAACTACTGTCTCTG 3240
AGAGACTATTTTCCACAGAGACCATATCCTAGTGGAAAAGAAAGTGGAACTACTGTCTCTG 3240
*****

Lewis      L L W V N K I D V A S F F G K R R Y I C
DSS
CTCTTGTGGSTGTGAAAATCGAGCTGGGCTCGTCTCTTTGGGAAAAGGAGGTATATCTGC 3300
CTCTTGTGGSTGTGAAAATCGAGCTGGGCTCGTCTCTTTGGGAAAAGGAGGTATATCTGC 3300
*****

Lewis      M K A I K S H K P S V L L G F D M S E L
DSS
ATGAAGCAATAAAATCCACAAGCCCTCTGTACTTCTGGSTTTTGATATGTCTGAGCTT 3360
ATGAAGCAATAAAATCCACAAGCCCTCTGTACTTCTGGSTTTTGATATGTCTGAGCTT 3360
*****

Lewis      K N V K H R L N F E C E S S H L -
DSS
AAGAATGTCAACACAGGCTGAACCTTGGAGTGTGAGTCAAAAACCTGTAG 3411
AAGAATGTCAACACAGGCTGAACCTTGGAGTGTGAGTCAAAAACCTGTAG 3411
*****

```

Q

(4) Partitioning defective 6 homolog gamma (Pard6g)

```

Lewis      M N R S F H E S Q T L R F Y D C S A V E
DSS
ATGAACGGAGTTTTTCATAAGTCTCAGAACCCCTAAGTTCTACGACTGCAGCCAGTGGAA 60
ATGAACGGAGTTTTTCATAAGTCTCAGAACCCCTAAGTTCTACGACTGCAGCCAGTGGAA 60
*****

Lewis      V K S K F G A E F R R F S L D R H K P S
DSS
GTCAAGAGCAAGTTTTGGGGCCGAGTTCCGAAGGTTCTCTCTTGGACCCACAAAGCCTGG 120
GTCAAGAGCAAGTTTTGGGGCCGAGTTCCGAAGGTTCTCTCTTGGACCCACAAAGCCTGG 120
*****

Lewis      K F E D F Y Q L V V H T H H I S N T E V
DSS
AAGTTTGAAGATTTCTACAGCTGGTGGTGCACACCCACCACATCTCCAACACCGAAGTG 180
AAGTTTGAAGATTTCTACAGCTGGTGGTGCACACCCACCACATCTCCAACACCGAAGTG 180
*****

Lewis      T I G Y A D V H S D L L P I N H D N F
DSS
ACCATTGGCTATGCTGATGTACACGGGGACCTGCTGCCCATCAACAAATGACGACAACTT 240
ACCATTGGCTATGCTGATGTACACGGGGACCTGCTGCCCATCAACAAATGACGACAACTT 240
*****

Lewis      C K A V S S A N P L L R V F I Q E R E E
DSS
TGCAGGGCCGCTCAAGTGGGAACCCCTTGGCTTGGAGTCTTCATCCAGAGCGAGAGGAG 300
TGCAGGGCCGCTCAAGTGGGAACCCCTTGGCTTGGAGTCTTCATCCAGAGCGAGAGGAG 300
*****

Lewis      A D H Y S F G A G T L S R K K K V L V T
DSS
GCAGACCAATACAGCTTCGGAGCAGGCACCTGTGTGAGSRAAGAAGGTGCTGGTGAAC 360
GCAGACCAATACAGCTTCGGAGCAGGCACCTGTGTGAGSRAAGAAGGTGCTGGTGAAC 360
*****

Lewis      L R D E S L R R R A Q L N I S M P H D F
DSS
CTGAGGGATGAGGCTCTGGCTCGGGGTGCCAGCTCAACATCAGCATGCCGACGACTTC 420
CTGAGGGATGAGGCTCTGGCTCGGGGTGCCAGCTCAACATCAGCATGCCGACGACTTC 420
*****

Lewis      R P V S S I I D V D I L F E T H R R V R
DSS
CGCCCGGTGTCCTCCATCATTGACGTGGACATCTCCCTGAGACGACCGCCGGGTGAG 480
CGCCCGGTGTCCTCCATCATTGACGTGGACATCTCCCTGAGACGACCGCCGGGTGAG 480
*****

```

```

Lewis      L Y R H G C E K P L S F Y I R D G T S V
DSS       CTTTACCGGCACGGCTGTGAGAAAGCCGCTGGGCTTCTACATCCGAGATGGCACCAAGGTG 540
          CTTTACCGGCACGGCTGTGAGAAAGCCGCTGGGCTTCTACATCCGAGATGGCACCAAGGTG 540
          *****

Lewis      R V T P H G L E K V P G I F I S R M V P
DSS       CGAGTGACACCCACGGGCTGGAAAAGSTACAGGSAITCTTCATTCCCGAATGGTGCDC 600
          CGAGTGACACCCACGGGCTGGAAAAGSTACAGGSAITCTTCATTCCCGAATGGTGCDC 600
          *****

Lewis      G G L A E S T G L L A V N D E V L E V N
DSS       GCGGGCTTGGCGAGAGCACTGGGCTGCTGGCTGTGAATGACGAAGTCTGGAGGTGAAT 660
          GCGGGCTTGGCGAGAGCACTGGGCTGCTGGCTGTGAATGACGAAGTCTGGAGGTGAAT 660
          *****

Lewis      G I E V A G K T L D Q V T D M M I A N S
DSS       GGGATTGAGSTCGCTGGAAAAGCATTGGATCAAGTCACTGACATGATGATAGCCCAACAGC 720
          GGGATTGAGSTCGCTGGAAAAGCATTGGATCAAGTCACTGACATGATGATAGCCCAACAGC 720
          *****

Lewis      H N L I V T V K P A N Q R N N V V R S S
DSS       CACAACCTCATTGTCACTGTCAAGCCCGCCAAACAAAGSAACAATGTGTCCGACAGCAGC 780
          CACAACCTCATTGTCACTGTCAAGCCCGCCAAACAAAGSAACAATGTGTCCGACAGCAGC 780
          *****

Lewis      R A S G S S V Q S T D S T T S H H S L P
DSS       CCGGCTCCGGCAGCTCTGTCCAGTCCACAGACAGCACCACCAGTCCACCACAGCCTGCCA 840
          CCGGCTCCGGCAGCTCTGTCCAGTCCACAGACAGCACCACCAGTCCACCACAGCCTGCCA 840
          *****

Lewis      G T H A L Q N S E E N E S D E E A D I V
DSS       GGCACCCACGCTCTGAGAAATCTGAGGAAATGGAGAGTGATGAGGAGGCCGACATCGTC 900
          GGCACCCACGCTCTGAGAAATCTGAGGAAATGGAGAGTGATGAGGAGGCCGACATCGTC 900
          *****

Lewis      I E G A L E P H R I P E M Q A V P P G S
DSS       ATCGAGGGGGCTCTAGAGCTCACCACATCCCAAGATGCAGGCTGTGCCTCCAGGCAGC 960
          ATCGAGGGGGCTCTAGAGCTCACCACATCCCAAGATGCAGGCTGTGCCTCCAGGCAGC 960
          *****

Lewis      L S E A N G A S L A H R L H R G D G G L
DSS       CTCTGAGGGCCAAATGGCCAGCCTGGCTCACAGGCTACACAGAGGGGACGGGGGCTG 1020
          CTCTGAGGGCCAAATGGCCAGCCTGGCTCACAGGCTACACAGAGGGGACGGGGGCTG 1020
          *****

Lewis      H S S G R E S N G S I H R F L S S L K P
DSS       CATACTCTGGGAGGGAGAGCAATGGCAGCATCCACAGATTTCTCAGCTCTCTGAAACCA 1080
          CATACTCTGGGAGGGAGAGCAATGGCAGCATCCACAGATTTCTCAGCTCTCTGAAACCA 1080
          *****

Lewis      D P R H S L V L P Q G G V E E H G P A I
DSS       GACCCCGACACAGCCTGTCTCTCCCCAGGAGGGGTGGAGGAACACGGACCAGCCATT 1140
          GACCCCGACACAGCCTGTCTCTCCCCAGGAGGGGTGGAGGAACACGGACCAGCCATT 1140
          *****

Lewis      T L -
DSS       ACCCTCTAG 1149
          ACCCTCTAG 1149
          *****

```

(5) ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 5 (St8sia5)

```

Lewis      M R Y A D P S A N R D L L G N R T L L F
DSS       ATGCGCTACGCAGACCCCTGGCCACAGGATTTGTTGGGAACCGAATTTGCTCTTC 60
          ATGCGCTACGCAGACCCCTGGCCACAGGATTTGTTGGGAACCGAATTTGCTCTTC 60
          *****

Lewis      I F I C A F A L V T L L Q Q I L Y S K S
DSS       ATCTTCATCTGGCGCTTGGCTTGGTGAACCTTGCTCCAGCAGATCTGTACAGCAAGAGC 120
          ATCTTCATCTGGCGCTTGGCTTGGTGAACCTTGCTCCAGCAGATCTGTACAGCAAGAGC 120
          *****

```


Lewis V K E K P G F H A M P S E I F T F L R M
 STCAAGCCCAAGCCCGCTTTCATGCTATGCCCTGGAGATCTTCACCTTCCCTGGCATG 1080
 DSS STCAAGCCCAAGCCCGCTTTCATGCTATGCCCTGGAGATCTTCACCTTCCCTGGCATG 1080

Lewis H S R G I L R V R T G T C N C C -
 CACAGCCCGGCATCCTCCGCTGACACAGCCACTGCAATTGCTGCTGA 1131
 DSS CACAGCCCGGCATCCTCCGCTGACACAGCCACTGCAATTGCTGCTGA 1131

(6) Lipoxygenase homology domains 1 (Loxhd1)

Lewis M M P Q E E R R R E E D I D F L G L Y E
 ATGATGCCCCAGAAAGAAAGCGCGGGAAGAAAGACATCGACTTCCCTGGCTTGTACGAG 60
 DSS ATGATGCCCCAGAAAGAAAGCGCGGGAAGAAAGACATCGACTTCCCTGGCTTGTACGAG 60

Lewis E E L L N Y N S E D D E D E L E H E Y Y
 GAGGAGCTGCTGAACTACAACCTCGGAGGATGATGAGGATGAGTTGGAGCACGAATACTAC 120
 DSS GAGGAGCTGCTGAACTACAACCTCGGAGGATGATGAGGATGAGTTGGAGCACGAATACTAC 120

Lewis K A E V Y E V V T A T G E V R G A G T D
 AAAGCAAAGGTGTACGAGGTGGTCAAGCCACAGGGATGTTCTGGCCACAGGACAGAT 180
 DSS AAAGCAAAGGTGTACGAGGTGGTCAAGCCACAGGGATGTTCTGGCCACAGGACAGAT 180

Lewis A N V F I T I F G E N G L S F H L H L T
 GCCAACGTTTCATCAGGATCTTGGAGAGAACGGGCTCTCTCCCAAGCTCCATCTCAC 240
 DSS GCCAACGTTTCATCAGGATCTTGGAGAGAACGGGCTCTCTCCCAAGCTCCATCTCAC 240

Lewis S K S E S A F E K A N V D V F R V R T N
 AGCAAGCGGAGTCTGCTTTTGAAGAAAGCCACSTTGTATTCGGGTACGGACCAAT 300
 DSS AGCAAGCGGAGTCTGCTTTTGAAGAAAGCCACSTTGTATTCGGGTACGGACCAAT 300

Lewis N V G L I Y K I R I E H E N T G L N A S
 AACSTGGGCTCATCTATAAAATCAGGATCGAGCATGATAACACAGGCTGAATGCCAGC 360
 DSS AACSTGGGCTCATCTATAAAATCAGGATCGAGCATGATAACACAGGCTGAATGCCAGC 360

Lewis M Y L D R V I V T D M K R F H L R Y Y F
 TGGTACCTGGACCGGTTATAGTGACGGACATGAAGAGGCTCACCTCCGCTATTACTTC 420
 DSS TGGTACCTGGACCGGTTATAGTGACGGACATGAAGAGGCTCACCTCCGCTATTACTTC 420

Lewis M C N N W L S E V E S D R Q W C R D L L
 AACTGTAACTAGCTGAGCAAGGTGGAGGTTGACCGTCASTGCTGCTGACCTCCTG 480
 DSS AACTGTAACTAGCTGAGCAAGGTGGAGGTTGACCGTCASTGCTGCTGACCTCCTG 480

Lewis A S E D P M D M P R G N K Y E I K V Y T
 GCCAGCTTCGACCCATGACATGCCAGAGGAAACAATATGAAATCAAGGTATACACT 540
 DSS GCCAGCTTCGACCCATGACATGCCAGAGGAAACAATATGAAATCAAGGTATACACT 540

Lewis G D V I G A G T D A D V F I N I F G E Y
 GGTGATGATGTTGTTGCAGGGACAGATGCTGATGCTTTATCAATATCTTTGGAGAGTAT 600
 DSS GGTGATGATGTTGTTGCAGGGACAGATGCTGATGCTTTATCAATATCTTTGGAGAGTAT 600

Lewis G D T G E R R L E N E K D N F E K G A E
 GGAGACACAGGGGAGCTAGGCTGGAAAATGAGAAAGACAACTTTGAAAAGGGAGCTGAA 660
 DSS GGAGACACAGGGGAGCTAGGCTGGAAAATGAGAAAGACAACTTTGAAAAGGGAGCTGAA 660

Lewis D K F T L D A F D L G Q L M E I N V E H
 GACAAGTTCACACTGGATGCACAGATTTGGGGCAGCTGATGAAGATCAATGTGGCCAC 720
 DSS GACAAGTTCACACTGGATGCACAGATTTGGGGCAGCTGATGAAGATCAATGTGGCCAC 720

```

      H N K E S S A G W F L S K I I I E D I S
Lewis  AACAACAAGGGGGTTCCGGCAGGTTGGTCTCTGCCAAGATCATCATTGAAGATATTGGG 780
DSS    AACAACAAGGGGGTTCCGGCAGGTTGGTCTCTGCCAAGATCATCATTGAAGATATTGGG 780
      *
      H E R K Y D F P L N R W L A L D E D D S
Lewis  AACAAAAGGAAATATGATTTCCCCCTTAAOCGCTGGCTGGCCCTGGATGAAGATGATGGC 840
DSS    AACAAAAGGAAATATGATTTCCCCCTTAAOCGCTGGCTGGCCCTGGATGAAGATGATGGC 840
      *
      H I Q R D I L V G S A E T T A I T Y I V
Lewis  AAAATCCAGAGAGACATCTTAGTGGGCGAGCTGAGACCACAGCCATAACCTACATTGTC 900
DSS    AAAATCCAGAGAGACATCTTAGTGGGCGAGCTGAGACCACAGCCATAACCTACATTGTC 900
      *
      T V F T G D I S G A G T K E K I Y L V M
Lewis  ACTGTCTTCACTGGGATATCCGTGGAGCTGGGACCAAATCTAAAATCTATTGGTCATG 960
DSS    ACTGTCTTCACTGGGATATCCGTGGAGCTGGGACCAAATCTAAAATCTATTGGTCATG 960
      *
      Y G A R Q H R N S Q K I F L E G G V F D
Lewis  TACGAGCCAGAGSAAATAGGAACAGTGSAAAAATCTTTCTGGAAGGTGGTGTGTTCCGAC 1020
DSS    TACGAGCCAGAGSAAATAGGAACAGTGSAAAAATCTTTCTGGAAGGTGGTGTGTTCCGAC 1020
      *
      H G R T D I F H I D L A V L L S P L S R
Lewis  AGGGCCCGCAGACATCTCCACATCGAOCCTGGCTGTCTCTTAGCCACTGAGCCGG 1080
DSS    AGGGCCCGCAGACATCTCCACATCGAOCCTGGCTGTCTCTTAGCCACTGAGCCGG 1080
      *
      V S I G H Q N I S V N R G W Y C E K V V
Lewis  GTCTCCATTGGSCACGGCAACATAGGCGTCAACCGAGGCTGGTACTGTGAGAGSTGGTG 1140
DSS    GTCTCCATTGGSCACGGCAACATAGGCGTCAACCGAGGCTGGTACTGTGAGAGSTGGTG 1140
      *
      V L C P F T G I Q Q T F P C S H W L D E
Lewis  GTACTGTGCCCTTCACTGGCATCCAGCAGACTTTCCCTGCGAGCAACTGGTGGATGAG 1200
DSS    GTACTGTGCCCTTCACTGGCATCCAGCAGACTTTCCCTGCGAGCAACTGGTGGATGAG 1200
      *
      K E A D G L I E R Q L Y E M V S L R K E
Lewis  AAGAAGCAGACGSACTGATTGASCGTCAGCTCTATGAGATGGTGTCTCTGAGSAGAAG 1260
DSS    AAGAAGCAGACGSACTGATTGASCGTCAGCTCTATGAGATGGTGTCTCTGAGSAGAAG 1260
      *
      R L E K F P W S L N V W T T D L K R A G
Lewis  CGGCTGAAAAAATTCCTTGGTCCCTTGGGTCTGGACAAGTGAATGAGAAAGCTGGC 1320
DSS    CGGCTGAAAAAATTCCTTGGTCCCTTGGGTCTGGACAAGTGAATGAGAAAGCTGGC 1320
      *
      T N S P I F I G I Y G K K E R Y D E I L
Lewis  ACCAACTGGCCATCTTCATCCAGATTTATGGGAAGAGGGGCGGACAGAGATTCCT 1380
DSS    ACCAACTGGCCATCTTCATCCAGATTTATGGGAAGAGGGGCGGACAGAGATTCCT 1380
      *
      L N F N H K W F K P G I I E K F R H E L
Lewis  CTGAATCCCAACAACAAATGTTCAAACAGGSCATCATTGAGAAGTTCAGGATGGAGCTC 1440
DSS    CTGAATCCCAACAACAAATGTTCAAACAGGSCATCATTGAGAAGTTCAGGATGGAGCTC 1440
      *
      P D L G R F Y K I R A W H D R Q N P G S
Lewis  CCAGATCTCGGAGGTTTTATAAGATTCGGGCTGGCACGACAGGAGAAATCCTGGCTCT 1500
DSS    CCAGATCTCGGAGGTTTTATAAGATTCGGGCTGGCACGACAGGAGAAATCCTGGCTCT 1500
      *
      G W H L E K M T L M N T L N E D E Y N F
Lewis  GGGTGGCATTAGAAAAGATGACTCTGATGAATACCTTGAATAAGGACAAATATAACTTC 1560
DSS    GGGTGGCATTAGAAAAGATGACTCTGATGAATACCTTGAATAAGGACAAATATAACTTC 1560
      *
      H C H R W L D A N E D D H E I V R E M T
Lewis  AACTGCAACCCCTGGCTGSAATGCAATGAGGATGACAAATGAGATCGTGAGGGAGATGACC 1620
DSS    AACTGCAACCCCTGGCTGSAATGCAATGAGGATGACAAATGAGATCGTGAGGGAGATGACC 1620
      *

```



```

      A E G P T V R R I N G M A R Y R V T V C
Lewis  GCAGANGGCCGACTGTACGGAGGATAATGGGAATGGCCDGGTATCGTGTGACTGTATGC 1680
DSS    GCAGANGGCCGACTGTACGGAGGATAATGGGAATGGCCDGGTATCGTGTGACTGTATGC 1680
      *
      T G E L E S A G T D A N V Y L C L F G D
Lewis  ACAGGGGAAC TTGAAGGTGCAGGAAC TGATGCCAACCTCTACCTCTGCCTTTTGGTGAT 1740
DSS    ACAGGGGAAC TTGAAGGTGCAGGAAC TGATGCCAACCTCTACCTCTGCCTTTTGGTGAT 1740
      *
      V G E T G E R L L Y N C R N N T D L F E
Lewis  GTGGGGACACAGGGGAGAGACTGCTTTACAACCTGCAGAAATAACACCSACCTGTTGAG 1800
DSS    GTGGGGACACAGGGGAGAGACTGCTTTACAACCTGCAGAAATAACACCSACCTGTTGAG 1800
      *
      K G N A D E F T I E S V T M R R V R R V
Lewis  AAAGCAATCCGGATGAATTTACCATTGGAGTCTGTCTACTATGAGGAAGGTGAGGCGAGTG 1860
DSS    AAAGCAATCCGGATGAATTTACCATTGGAGTCTGTCTACTATGAGGAAGGTGAGGCGAGTG 1860
      *
      R V E H D D K G S G S G W Y L D R V L V
Lewis  AGGTCAGGCACGAGGCAAGGCTCGGGCAGCGGCTGACTCGACAGGGTGCTGGTG 1920
DSS    AGGTCAGGCACGAGGCAAGGCTCGGGCAGCGGCTGACTCGACAGGGTGCTGGTG 1920
      *
      R E E G Q F E E D N V E F F C L E W L D
Lewis  ASAGGGAAGGGCAGCCAGAGGTGACAAAGTGGAGTTCCCTGCTCAGGTGGCTGGAC 1980
DSS    ASAGGGAAGGGCAGCCAGAGGTGACAAAGTGGAGTTCCCTGCTCAGGTGGCTGGAC 1980
      *
      K D E D D G Q L V R E L L P S D S N A T
Lewis  AAAGCAAGSATGATGGGAGCTGGTCCGAGAGTTGCTACCCAGTGACAGCAATGCCAGS 2040
DSS    AAAGCAAGSATGATGGGAGCTGGTCCGAGAGTTGCTACCCAGTGACAGCAATGCCAGS 2040
      *
      L K N F R Y H I S V K T G D V S G A S T
Lewis  CTCAAGAACTTCCGCTATCAGCTGAGGAGGAGGAGGATGCTCTGGGGCCAGCACA 2100
DSS    CTCAAGAACTTCCGCTATCAGCTGAGGAGGAGGAGGATGCTCTGGGGCCAGCACA 2100
      *
      D S R V Y I K L Y G E K S D T I K Q V L
Lewis  GATTCGCGGTGTACATCAAGCTCTATGGAGAGAAATCTGACACCATCAAGCAAGTTCTT 2160
DSS    GATTCGCGGTGTACATCAAGCTCTATGGAGAGAAATCTGACACCATCAAGCAAGTTCTT 2160
      *
      L V E D N N L K D Y F E R G R V D E F T
Lewis  CTGCTCTGACAAACCTCAAAGACTAATTTGAACTGGCCGAGTGGATGATTTACT 2220
DSS    CTGCTCTGACAAACCTCAAAGACTAATTTGAACTGGCCGAGTGGATGATTTACT 2220
      *
      L E T L N I G T I I R L V I G H D S T G
Lewis  CTGGAGCCCTGAACATAGGAATATCATCAGGCTGGTATTGGGCACBACAGCACTGGC 2280
DSS    CTGGAGCCCTGAACATAGGAATATCATCAGGCTGGTATTGGGCACBACAGCACTGGC 2280
      *
      M H A G W F L G S V Q I R V F R Q G K Q
Lewis  ATGCATGCAGGCTGGTTCTGGGTAGCGTCCAGATCCCGGTGCCCGCCAAAGCAAGCAG 2340
DSS    ATGCATGCAGGCTGGTTCTGGGTAGCGTCCAGATCCCGGTGCCCGCCAAAGCAAGCAG 2340
      *
      Y T F P A H R N L D K N Q A D G R L E V
Lewis  TACACCTTCCCTGCCAACGGTGGCTGGATAAGAACCAAGCTGATGGGCGTCTGGAGGTG 2400
DSS    TACACCTTCCCTGCCAACGGTGGCTGGATAAGAACCAAGCTGATGGGCGTCTGGAGGTG 2400
      *
      E L Y P S E V V E I Q K L V H Y E I E I
Lewis  GAGCTGTACCCAGCGAGGTGGTGGAGATCCAGAAATGGTCCACTACGAGATTGAGATT 2460
DSS    GAGCTGTACCCAGCGAGGTGGTGGAGATCCAGAAATGGTCCACTACGAGATTGAGATT 2460
      *

```

```

Lewis      M T G D V S G A G T T S E V Y V Q I Y S
DSS        TGGACCGGAGATGTGGGTGGTGGCCGGCACCACCTCCCGAGTCTACGTGCAGATCTACGGG 2520
          TGGACCGGAGATGTGGGTGGTGGCCGGCACCACCTCCCGAGTCTACGTGCAGATCTACGGG 2520
          ..

Lewis      E E G K T E V L F L S S R E K V F E R G
DSS        GAAGAAGGCAGACCGGAAGTACTCTTCCTCTCCAGCCGCTCCAAGTGTITGATCGGGG 2580
          GAAGAAGGCAGACCGGAAGTACTCTTCCTCTCCAGCCGCTCCAAGTGTITGATCGGGG 2580
          ..

Lewis      E K D I F Q L E A A D V G E I Y K I R L
DSS        TCCAAGGACATATTCAGCTGGAAGCGCGGACGTGGGGGAGATCTACAAGATCCGACTG 2640
          TCCAAGGACATATTCAGCTGGAAGCGCGGACGTGGGGGAGATCTACAAGATCCGACTG 2640
          ..

Lewis      G H T G E G F G F S M F V D T V W L R H
DSS        GGACACCGGGCAGGGGCTTTGGGCCAGCTGGTTTGGACACAGTGTGGCTGGGGCAC 2700
          GGACACCGGGCAGGGGCTTTGGGCCAGCTGGTTTGGACACAGTGTGGCTGGGGCAC 2700
          ..

Lewis      L V V Q E E S L T F E E E A R R K K E K
DSS        CTGGTGGTGCAGGAGGAGCCCTGACGCCGGAGSAGGAGGCCCGGAGGAAGSAGAA 2760
          CTGGTGGTGCAGGAGGAGCCCTGACGCCGGAGSAGGAGGCCCGGAGGAAGSAGAA 2760
          ..

Lewis      E K I R Q L L K K E R L K A K L Q R K K
DSS        GAGAAGTTGGCCAGCTGTTGAAGAGGAAAGGCTGAAGGCTAAGCTGCAGCGGAAGA 2820
          GAGAAGTTGGCCAGCTGTTGAAGAGGAAAGGCTGAAGGCTAAGCTGCAGCGGAAGA 2820
          ..

Lewis      K K K K K S S E E E E E G D E E E E S S
DSS        AAGAAAAGSAAAGSAGCCAGCGATGAGGAGGAGGAGGAGATGAGGAAGSAGATCCCTC 2880
          AAGAAAAGSAAAGSAGCCAGCGATGAGGAGGAGGAGGAGATGAGGAAGSAGATCCCTC 2880
          ..

Lewis      S E E S S S E E E E E E E E E E E E E
DSS        TCAGAGAGTCCCTCTCTGAGAGSAGGAGGAGSAGGAGTCAAGAGSAGSAGSAGSAA 2940
          TCAGAGAGTCCCTCTCTGAGAGSAGGAGGAGSAGGAGTCAAGAGSAGSAGSAGSAA 2940
          ..

Lewis      E E Y G P G M Q E V I V Q Y E F V A N R
DSS        GAGGAGTATGGGCCAGGGATGCAGGAGGTGATTGTGCASTACAAGTTCCTCGCTAACCG 3000
          GAGGAGTATGGGCCAGGGATGCAGGAGGTGATTGTGCASTACAAGTTCCTCGCTAACCG 3000
          ..

Lewis      M L A R S R E D N E L V V E L V P A G Q
DSS        TGGCTGGCCCGGSCAAGSAGSACAATGAAGTGTGCTGGAAATGGTTCCAGCTGGTCA 3060
          TGGCTGGCCCGGSCAAGSAGSACAATGAAGTGTGCTGGAAATGGTTCCAGCTGGTCA 3060
          ..

Lewis      P G P E P N T Y E V Q V I T G N V P K A
DSS        CCGGTCCTGAGCCCAATACCTACGAGTCCAGTGTATCAGAGGAACTACCCAGGCT 3120
          CCGGTCCTGAGCCCAATACCTACGAGTCCAGTGTATCAGAGGAACTACCCAGGCT 3120
          ..

Lewis      G T E A N V Y L T I Y G E E Y G D T G E
DSS        GGCATGATGCCAACGTCTACTGACCATCTACGGTGAAGAAACGGGACACAGGGGAG 3180
          GGCATGATGCCAACGTCTACTGACCATCTACGGTGAAGAAACGGGACACAGGGGAG 3180
          ..

Lewis      R F L K K S D K S N K F E Q G Q T D T F
DSS        CGACCCCTGAGAAATCAGACAAAGTCCAACAAGTTTGAAGCAGGGCCAGACAGATACCTC 3240
          CGACCCCTGAGAAATCAGACAAAGTCCAACAAGTTTGAAGCAGGGCCAGACAGATACCTC 3240
          ..

Lewis      T I Y A I D L G A L T K I R I R H D N T
DSS        ACCATCTATGCCATTGACTGGGGGCTTACCAAGATTCGGATCCGTCATGATAACAG 3300
          ACCATCTATGCCATTGACTGGGGGCTTACCAAGATTCGGATCCGTCATGATAACAG 3300
          ..

Lewis      G H R F S M F L D R I D I T D V N H E T
DSS        GSCAACAGCCCGGCTGGTTCCTGGACAGAAATAGACATCACCAGCTGAATAACGAGACC 3360
          GSCAACAGCCCGGCTGGTTCCTGGACAGAAATAGACATCACCAGCTGAATAACGAGACC 3360
          ..

```

```

      ↓
Lewis  T Y Y F P C Q R W L A V E E D D G Q L S
DSS    ACGTACTACTTCCDGTGCCAGCGCTGGTTAGCASTCGAGGAGGATGACGGCCASTTATCC 3420
      *
      *
Lewis  R E L L F V D E S Y V L F S E D E E G G
DSS    AGGAGCTGCTGCCAGTGGATGASTCCTACGTGCTACCTAGTGAAGACBAGAGGGCGGG 3480
      *
      *
Lewis  G Q G D N N F L D N L A L E Q K D K S T
DSS    GGCCAGGGTGACAATAATCCCTCGACAACCTGCTCTGGAGCAGAAAGATAAATCAACC 3540
      *
      *
Lewis  T F S V T I K T E D K E N A B T D A N V
DSS    ACATTTTCAGTGACCATAAAGACTGGGGACAAAAGAACCCAGGCCACAGATGCCAACGTC 3600
      *
      *
Lewis  F I T L E G T K D N H G M T L L K E S K
DSS    TTCATCAGGCTCTTTGGCAGAAAGATAACAACGGGATGACCCTCTGAAAGTCTCCAAA 3660
      *
      *
Lewis  T N S D E F E R D S I E I F T V E T L D
DSS    ACCAACAGTGACAAATTTGAAGGGACAGCATTGAAATCTTCAACCGTGGAGACCTGGAT 3720
      *
      *
Lewis  L G D L W K V R I G H D N T G K A F G W
DSS    CTGGAGACCTTTGGAAAGTTAGSATCGGCCATGACAAACAGGCAAGGCCCCAGGCTGG 3780
      *
      *
Lewis  F V E W V E V E A P E L G K C H T F P C
DSS    TTTGTAGACTGGGTGGAGTGGATGCCCATCTCTTGGAAAGTGCATGACCTTCCCTGC 3840
      *
      *
Lewis  G R W L A K N E D D S S I V R D L F E A
DSS    GGCDDTTGGTTGGCCAGAATGAGGACGAGGGGAGCATCTCAGGGATCTCTTCCACGGG 3900
      *
      *
Lewis  E L Q T R L Y T P F V F Y E I T L Y T S
DSS    GAGCTTCAGACAGGCTCTACAGCCATTGGCCCTTATGAGATCACCTCTACACCAGC 3960
      *
      *
Lewis  D V F A A G T S A N V F I V I Y G C D A
DSS    GATGCTTTGCTGCTGGGACGGACGCCAAGCTCTTCATAGTCACTACGGCTGTGATGCC 4020
      *
      *
Lewis  V C T R Q K F L C T N K R E Q K L L F E
DSS    GTGTGCACCCGGCAGAAGTTCCTGTGCACCAATAAGAGGGAAACAGAAGTGTCTTTCGAG 4080
      *
      *
Lewis  R K S A S R F I V E L E D V G E I I E K
DSS    AGGAAATCAGCCTCCCGCTTTATTGTGAGTTAGAGAGCGTGGGTGAGATCATAGAAAAA 4140
      *
      *
Lewis  I R I G H D N T G I N F G W H C S H V D
DSS    ATTCGGATTGGCCATGACAACACAGGCATAAACCCCTGGGTGGCACTGCTCCACGTGGAC 4200
      *

```

```

      I R R L L P E K D S T E T L T F F C D R
Lewis  ATCCGAGGCTCCTCCCGAGAAAGACGGTACAGAAACCTTGACTTTCDDCTGTGATCSA 4260
DSS    ATCCGAGGCTCCTCCCGAGAAAGACGGTACAGAAACCTTGACTTTCDDCTGTGATCSA 4260
      *
      M L A T S E D D K K E T I R E L V P Y D I
Lewis  TGGCTCGCCACCTCTGAGGATGACAAAGACCATCCGAGAACTGGTCCDCTATGACATC 4320
DSS    TGGCTCGCCACCTCTGAGGATGACAAAGACCATCCGAGAACTGGTCCDCTATGACATC 4320
      *
      F T E K Y M K E G S L R Q V Y K E V E E
Lewis  TTCACTGAGAAATACATGAAAGATGGATCCTTAAGCCAGGCTACAAAGAAATGGAGAA 4380
DSS    TTCACTGAGAAATACATGAAAGATGGATCCTTAAGCCAGGCTACAAAGAAATGGAGAA 4380
      *
      F L S I V L Y S V Q I F T G N V P G A G
Lewis  CCTCTGGACATTTGTGCTGACTCTGTGCAGATCTTCACAGGAAATGTTCTCGGGCCGGG 4440
DSS    CCTCTGGACATTTGTGCTGACTCTGTGCAGATCTTCACAGGAAATGTTCTCGGGCCGGG 4440
      *
      T D A K V Y I T I Y S D L G D T G E R Y
Lewis  ACAGATGCCAAGGTCTACATCACCATCTACGGAGACCTCGGGGACACTGGGGAGCGGTAC 4500
DSS    ACAGATGCCAAGGTCTACATCACCATCTACGGAGACCTCGGGGACACTGGGGAGCGGTAC 4500
      *
      L G K S E H R T N K F E K G T A D T F I
Lewis  CTTGGCAAGTCAGAGAACCCACCCACAAAGTTTGAGAAAGGAAACGGCTGACACCTTCATC 4560
DSS    CTTGGCAAGTCAGAGAACCCACCCACAAAGTTTGAGAAAGGAAACGGCTGACACCTTCATC 4560
      *
      I E A A D L G V I Y K I K L R H D N T K
Lewis  ATCGAGGCCGCTGACCTGGGTGTCTATCAAGATCAAGCTCCGCCATGACAAACCAAG 4620
DSS    ATCGAGGCCGCTGACCTGGGTGTCTATCAAGATCAAGCTCCGCCATGACAAACCAAG 4620
      *
      M C A D N Y V E K Y E I W N D T H E D E
Lewis  TGGTGTCCGACTGGTATGTGGAGAAAGTGGAGATGGAATGACACCAATGAGGACGAG 4680
DSS    TGGTGTCCGACTGGTATGTGGAGAAAGTGGAGATGGAATGACACCAATGAGGACGAG 4680
      *
      F L F L C G R W L S L K K E D G R L E R
Lewis  TTTCTATTCTGTGTGGGCGCTGCTATCCCTGAAGAAAGGAGATGGGCGCTGGAGAG 4740
DSS    TTTCTATTCTGTGTGGGCGCTGCTATCCCTGAAGAAAGGAGATGGGCGCTGGAGAG 4740
      *
      L F Y E K E Y T G D R S S N C S S F A D
Lewis  CTCTCTATGAGAAGGATACACTGGAGACCCGACAGCAACTGCAACAGCCCGCTGAC 4800
DSS    CTCTCTATGAGAAGGATACACTGGAGACCCGACAGCAACTGCAACAGCCCGCTGAC 4800
      *
      F W E I A L S E K M A D V D I D T V T G
Lewis  TTCGGGAGATCGCCCTGAGTTCCAAATGGCAGCCTGGACATTGACACASTGACAGGA 4860
DSS    TTCGGGAGATCGCCCTGAGTTCCAAATGGCAGCCTGGACATTGACACASTGACAGGA 4860
      *
      F M V D Y V Q E G E P V I P Y Y V S V T T
Lewis  CCCATGGTTGACTATGTTCAAGATGGCCCGGTGATCCCTACTATGATCAGTACTACT 4920
DSS    CCCATGGTTGACTATGTTCAAGATGGCCCGGTGATCCCTACTATGATCAGTACTACT 4920
      *
      G K H K E A A T D S R A F V L L I G E D
Lewis  GGGAGCACAAAGGAGGACCCACTGATAGCCGCGCTTGTCTCTGCTCATTGGGGAAGAT 4980
DSS    GGGAGCACAAAGGAGGACCCACTGATAGCCGCGCTTGTCTCTGCTCATTGGGGAAGAT 4980
      *
      D E S T N R I W L D F P G G K R G F S C
Lewis  GATGACGTACCAACCSCATCTGGCTGGACTTTCGCCAGGGAAAGGAGGCTTCAGCTGT 5040
DSS    GATGACGTACCAACCSCATCTGGCTGGACTTTCGCCAGGGAAAGGAGGCTTCAGCTGT 5040
      *
      G S V E E F Y V G G L D V G I I K K I E
Lewis  GGCTCTGTGGAGGAGTTCTACGTTGGTGGCTTGGATGTGGGCATCATCAAGAAAATAGAG 5100
DSS    GGCTCTGTGGAGGAGTTCTACGTTGGTGGCTTGGATGTGGGCATCATCAAGAAAATAGAG 5100

```

```

Lewis      V L Y E M T V N T G D V V G S T D E N
DSS        GTGCTGTATGAAATGACGGTGTGGACTGGCGATGTGGTGGGGTGGAGGCACTGACTCCAAT 5160
          *GTGCTGTATGAAATGACGGTGTGGACTGGCGATGTGGTGGGGTGGAGGCACTGACTCCAAT 5160
          *.....

Lewis      I F M T L Y G I N G S T E E V Q L S K K
DSS        ATCTTCATGACCCCTCATGGCATCAACGGGACACAGAAGAGTGCAGCTGGACAAGAAG 5220
          *ATCTTCATGACCCCTCATGGCATCAACGGGACACAGAAGAGTGCAGCTGGACAAGAAG 5220
          *.....

Lewis      K A S F E R E G N D T F I M E I L D I A
DSS        AAGGCCAGGTTTGAOCGGGAACAAAATGACACCTTCATCATGGAGATCTGGACATTGCT 5280
          *AAGGCCAGGTTTGAOCGGGAACAAAATGACACCTTCATCATGGAGATCTGGACATTGCT 5280
          *.....

Lewis      F F T K M R I R I D D L G S R F E W F L
DSS        CCCTTACGAGATGCGCATCCGTATCGACGGGCTGGGGAGCCGSCCAGAGTGGTTCCTG 5340
          *CCCTTACGAGATGCGCATCCGTATCGACGGGCTGGGGAGCCGSCCAGAGTGGTTCCTG 5340
          *.....

Lewis      E R I L L K N H N T G D L T M E Y Y G D
DSS        GAGAGATCCCTGTGAAGAACATGAACACAGGGGACCTGACCATGTTCTACTACGGAGAC 5400
          *GAGAGATCCCTGTGAAGAACATGAACACAGGGGACCTGACCATGTTCTACTACGGAGAC 5400
          *.....

Lewis      M L S Q E K K G K K T L V C E I C A V I D
DSS        TGGCTGTCCCAAAGAAAGSAAAGAGACACTGGTGTGAAATATGTGGGTCATTGAT 5460
          *TGGCTGTCCCAAAGAAAGSAAAGAGACACTGGTGTGAAATATGTGGGTCATTGAT 5460
          *.....

Lewis      G E E M M E W T S Y T V S V K T S D I L
DSS        GGAGAGGAGATGATGGAGTGGACCTCCTACACAGTCTCGGTGAAGACCACTGACATCTG 5520
          *GGAGAGGAGATGATGGAGTGGACCTCCTACACAGTCTCGGTGAAGACCACTGACATCTG 5520
          *.....

Lewis      G A G T D A N V F I I I F G E H G S S S
DSS        GGGGAGGCACGGACGCTAATGTGTTTCATCATCTCTTGGGGAAAAATGGAGACAGCGGG 5580
          *GGGGAGGCACGGACGCTAATGTGTTTCATCATCTCTTGGGGAAAAATGGAGACAGCGGG 5580
          *.....

Lewis      T L A L E Q S A N W N E F E R N N T D T
DSS        ACCCTGGCCCTGAACAGTCAAGCACTGCAACAAGTTTGAGCAGCAACACAGACACC 5640
          *ACCCTGGCCCTGAACAGTCAAGCACTGCAACAAGTTTGAGCAGCAACACAGACACC 5640
          *.....

Lewis      F N F P D M L S L G H L C K L R V W R D
DSS        TTCAATTTTCTGACATGCTGAGCCTGGGOCACCTCTGCAAACTGAGGCTCTGGCATGAC 5700
          *TTCAATTTTCTGACATGCTGAGCCTGGGOCACCTCTGCAAACTGAGGCTCTGGCATGAC 5700
          *.....

Lewis      N K G I F P G W H L S Y I D V K D N S R
DSS        AACAAAGGATATTTCTGGATGGCATCTGAGCTACATTGACGTGAAGSACAATTCOCST 5760
          *AACAAAGGATATTTCTGGATGGCATCTGAGCTACATTGACGTGAAGSACAATTCOCST 5760
          *.....

Lewis      D E T F R F Q C D C M L S K S E G D R Q
DSS        GACGAGCCTTCGCTTCCAGTGTGACTGCTGGCTGTCTAAGAGTGAAGGTGACAGACAA 5820
          *GACGAGCCTTCGCTTCCAGTGTGACTGCTGGCTGTCTAAGAGTGAAGGTGACAGACAA 5820
          *.....

Lewis      T L E D F A C A N N E I R D E L E E T T
DSS        ACCTCCGAGACTTCGCTGTGCCAACAAATGAGATCCGGGATGAGCTGSAAGSACCACC 5880
          *ACCTCCGAGACTTCGCTGTGCCAACAAATGAGATCCGGGATGAGCTGSAAGSACCACC 5880
          *.....

Lewis      Y E I V I E T G N G D E T R E N I W L I
DSS        TACGAGATTGTCATGAAACAGGCAACGGAGCCGAAACAGAGAGAACTCTGGCTCATC 5940
          *TACGAGATTGTCATGAAACAGGCAACGGAGCCGAAACAGAGAGAACTCTGGCTCATC 5940
          *.....

```

M


```

Lewis      L E G R E N R E K E F L V E N S S R Q R
DSS        TTGGAAGGCAGGAAGAACAGATCCAAGGAATTCCTTGTGSAATAATCTTCCAGACAGCGG 6000
          TTGSAAGGCAGGAAGAACAGATCCAAGGAATTCCTTGTGSAATAATCTTCCAGACAGCGG 6000
          *
          *
          *
Lewis      A E R K G T T D T F E F D S I F L G D I
DSS        GCCTTTCCGAAGGGACACAGACACCTTTGAGTTTGACAGCATCTTCTTGGGTGACATC 6060
          GCCTTTCCGAAGGGACACAGACACCTTTGAGTTTGACAGCATCTTCTTGGGTGACATC 6060
          *
          *
          *
Lewis      A S I C V G H I A R E D S F I P K R E L
DSS        GCCTCCCTCTGTGTGGGCCACCTGGCAAGGGAGGACCGSTTCATCCCCAAGAGAGAGCTA 6120
          GCCTCCCTCTGTGTGGGCCACCTGGCAAGGGAGGACCGSTTCATCCCCAAGAGAGAGCTA 6120
          *
          *
          *
Lewis      V W H V K T I T I T E M E Y G N V Y F F
DSS        GTCTGGCATGTCAAGACCATCACCATCACCGAGATGGAGTATGGCAACSTGTACTTCTTT 6180
          GTCTGGCATGTCAAGACCATCACCATCACCGAGATGGAGTATGGCAACSTGTACTTCTTT 6180
          *
          *
          *
Lewis      H C D C L I P L E R K R E Y F K V F E V
DSS        AACTGTACTGCCTCATCCCCCTCAAGAGGAAGAGGAAGTACTTCAAGSTATTCCAGGTC 6240
          AACTGTACTGCCTCATCCCCCTCAAGAGGAAGAGGAAGTACTTCAAGSTATTCCAGGTC 6240
          *
          *
          *
Lewis      T K T T E S F A S K I Q S L V P V K Y E
DSS        ACTAAAACGACAGAGAGCTTCCCCAGTAAGATCCAGAGCCTGGTCCCAAGTCAAGTACGAG 6300
          ACTAAAACGACAGAGAGCTTCCCCAGTAAGATCCAGAGCCTGGTCCCAAGTCAAGTACGAG 6300
          *
          *
          *
Lewis      V I V T T G Y E F G A G T D A N V F V T
DSS        GTCATCGTGACACGGGCTACGAACCAAGGGCGGGCACTGATGCCAACSTCTTCCGTGACC 6360
          GTCATCGTGACACGGGCTACGAACCAAGGGCGGGCACTGATGCCAACSTCTTCCGTGACC 6360
          *
          *
          *
Lewis      I F G A N S D T G E R E L K Q K M R N L
DSS        ATCTTTGGCCTAAAGGGGACACGGCAAGGCGCAACTGAAACAGAAAATCGGCAACCTC 6420
          ATCTTTGGCCTAAAGGGGACACGGCAAGGCGCAACTGAAACAGAAAATCGGCAACCTC 6420
          *
          *
          *
Lewis      F E R G S T D S F F L E T L E L G E L R
DSS        TTCSAGAGGGCAGTACAGACCCTTCTTTCTGGAGACATTAGAGCTGGGAGAGCTGCSC 6480
          TTCSAGAGGGCAGTACAGACCCTTCTTTCTGGAGACATTAGAGCTGGGAGAGCTGCSC 6480
          *
          *
          *
Lewis      K V R L E R D S S G Y S E M L V E K V
DSS        AAGSTTCGGCTGGAACATGACAGCAGTGGCTACTACTCGGGCTGGCTGGTGGAGAAAGTG 6540
          AAGSTTCGGCTGGAACATGACAGCAGTGGCTACTACTCGGGCTGGCTGGTGGAGAAAGTG 6540
          *
          *
          *
Lewis      E V T N T S T G V A T I F S C G R W L D
DSS        GAGSTCACCAACTAGTACTGGAGTGGCCACCATCTTCTCTCGGCAAGTGGSTTAGAC 6600
          GAGSTCACCAACTAGTACTGGAGTGGCCACCATCTTCTCTCGGCAAGTGGSTTAGAC 6600
          *
          *
          *
Lewis      K S R G D G L T W R E L F P S V -
DSS        AAGTCTCGGGGCGATGGGCTCACCTGGCAGAGCTCTTCCCGTCCGCTCGA 6651
          AAGTCTCGGGGCGATGGGCTCACCTGGCAGAGCTCTTCCCGTCCGCTCGA 6651
          *
          *
          *

```

(7) SET binding protein 1 (Setpb1)

```

Lewis      M E S R E I L S S S R Q R G S S E S E F L
DSS        ATGSAATCCAGGGAAATACTAAGCAGCTCCCGGCAAGAGGGAGGAGTGGGAATTCCTT 60
          ATGSAATCCAGGGAAATACTAAGCAGCTCCCGGCAAGAGGGAGGAGTGGGAATTCCTT 60
          *
          *
          *
Lewis      P G S S E R S P A T P G C S G E F L K G
DSS        CCGGGCTCCTCCTCCAGTCCCAGGCCACTCCTGGGTGTTCCAGSAGAGCCCTAAAAGGC 120
          CCGGGCTCCTCCTCCAGTCCCAGGCCACTCCTGGGTGTTCCAGSAGAGCCCTAAAAGGC 120
          *
          *
          *

```

```

Lewis      I S V G G E R I E F E E E D E L G S G R
DSS        ATCTCTGTGGGTGGAGAGCCATTGAGCCAGAGGAGGAGGACGAACCTGGGCTCAGGACGA
          180
          ATCTCTGTGGGTGGAGAGCCATTGAGCCAGAGGAGGAGGACGAACCTGGGCTCAGGACGA
          180
          *
          *
          *
Lewis      D V E C N S N A D S E E K W V A G D G L E
DSS        GATGTGGATTGCAACTCTAATGCGGACAGTGAAGAAATGGTGGCGGGAGATGGCTTGGAG
          240
          GATGTGGATTGCAACTCTAATGCGGACAGTGAAGAAATGGTGGCGGGAGATGGCTTGGAG
          240
          *
          *
          *
Lewis      E Q E F S I K E A N F S E G S L K L K I
DSS        SAGCAGGAGTTTCCATCAAGAGGCCAACTTCTCCGAGGGGAGTTTAAAGCTGAAGATT
          300
          SAGCAGGAGTTTCCATCAAGAGGCCAACTTCTCCGAGGGGAGTTTAAAGCTGAAGATT
          300
          *
          *
          *
Lewis      Q T T K R A K K P F K N L E N Y I C P F
DSS        CAGACCAGAAACCGGCTAAGAAACCCCCAAAACCTGGAGAACTACATCTGTCCACCT
          360
          CAGACCAGAAACCGGCTAAGAAACCCCCAAAACCTGGAGAACTACATCTGTCCACCT
          360
          *
          *
          *
Lewis      E I K I T I K Q S G D Q K V S R A G K N
DSS        GAGATCAAGATCACCATAAAGCAGTCTGGGACCCAGAAAGGTGTCCTGCTGGGAAAAAT
          420
          GAGATCAAGATCACCATAAAGCAGTCTGGGACCCAGAAAGGTGTCCTGCTGGGAAAAAT
          420
          *
          *
          *
Lewis      G K A T E E E R N H S K K K L L T A S
DSS        GGCARAGCTACAAGGAAGAGAAAGAAACCACTCCAAAAGAGAAGTCTCTCACAGCCAGT
          480
          GGCARAGCTACAAGGAAGAGAAAGAAACCACTCCAAAAGAGAAGTCTCTCACAGCCAGT
          480
          *
          *
          *
Lewis      D F A A S D L K G F Q T G A Y E R F Q K
DSS        GACDCTGCAGCCAGTCACTCAAAGGATTTGAGACACAGGCATACGAGAGGCCCCAGAAA
          540
          GACDCTGCAGCCAGTCACTCAAAGGATTTGAGACACAGGCATACGAGAGGCCCCAGAAA
          540
          *
          *
          *
Lewis      H S T L Q Y D E G L S Q G F T S D T L E
DSS        CACTCAACACTTCAATATGACCCAGGTCTCTCCAGGGCTTACCAGTGACACTTAAAA
          600
          CACTCAACACTTCAATATGACCCAGGTCTCTCCAGGGCTTACCAGTGACACTTAAAA
          600
          *
          *
          *
Lewis      P K R Q Q K S S S Q S R M D W S S N S D
DSS        CCAAAGCACCAGCAAAAAGCAGCAGCCAGAGCCACATGGACTGGTCTCCAACTCTGAC
          660
          CCAAAGCACCAGCAAAAAGCAGCAGCCAGAGCCACATGGACTGGTCTCCAACTCTGAC
          660
          *
          *
          *
Lewis      E G P A T Q N C I I E F E S G R D T T S
DSS        AGTGGACCACTACTCAGAATTGTATCATCAGTCCAGAGTCTGGCAGAGACACTACAAGC
          720
          AGTGGACCACTACTCAGAATTGTATCATCAGTCCAGAGTCTGGCAGAGACACTACAAGC
          720
          *
          *
          *
Lewis      T S K I P A L E P V A S F A K T Q S E K E
DSS        ACCAGCAAGATCCGAGCTCTTGAACCCGGTCTTCTTTCAGAGACCCAGTAAAGAA
          780
          ACCAGCAAGATCCGAGCTCTTGAACCCGGTCTTCTTTCAGAGACCCAGTAAAGAA
          780
          *
          *
          *
Lewis      G S T G S T W S Q L S S S S R D L L L G
DSS        GGCAGTACAGGGAGTACCTGGAGTCAATTGTCCAGCAGCAGCAAGACCTGCTCTTGGGA
          840
          GGCAGTACAGGGAGTACCTGGAGTCAATTGTCCAGCAGCAGCAAGACCTGCTCTTGGGA
          840
          *
          *
          *
Lewis      S V V P S F G S H S S P A T P S N S A E
DSS        AGTGTGGTACCATCCCTGGCAGCCACAGTCTCCAGCAGCCCTAGCAACTCAGCTGAG
          900
          AGTGTGGTACCATCCCTGGCAGCCACAGTCTCCAGCAGCCCTAGCAACTCAGCTGAG
          900
          *
          *
          *
Lewis      C W G L Q P L G D Q S G G G T X D L F E
DSS        TGCAATGGACTCCAGCCATTGGGAGATCAAGAGGGAGGAGTACAAGGATCTCCAGAA
          960
          TGCAATGGACTCCAGCCATTGGGAGATCAAGAGGGAGGAGTACAAGGATCTCCAGAA
          960
          *
          *
          *
Lewis      F P T I S S K K E S S K E D L I S Q T I
DSS        CCACCCTATCAGTAGCAAGAAATCCAGTAAAAAGATCTGATAAGCCAAACCATA
          1020
          CCACCCTATCAGTAGCAAGAAATCCAGTAAAAAGATCTGATAAGCCAAACCATA
          1020
          *
          *
          *

```

Lewis F N S D L D W V K S A Q K A F E T T E S
 DSS CCAAACCTCGSATCTGGATTGGGTCAAGAGTCCCGAGAAAGCATTGAAACCCAGAAAGG 1080
 CCAAACCTCGSATCTGGATTGGGTCAAGAGTCCCGAGAAAGCATTGAAACCCAGAAAGG 1080

Lewis K R E T Y S A E N A Q E T E P T R Q S I
 DSS AAAAGGSAACCTATTCTGCAGACAAATGCTCAAGAGACATCCCGACCCGSCAGAGCATC 1140
 AAAAGGSAACCTATTCTGCAGACAAATGCTCAAGAGACATCCCGACCCGSCAGAGCATC 1140

Lewis S S V S N P E N D S S H V R I T I P I K
 DSS ASTTCCGTCAGTAATCCGSAAAATGATTCCAGTCAATGTCCGGATTACTATCCCCATAAAG 1200
 ASTTCCGTCAGTAATCCGSAAAATGATTCCAGTCAATGTCCGGATTACTATCCCCATAAAG 1200

Lewis T P S L D F T N H K R K R Q S I K A V
 DSS ACACCTTCTCTGGATCCAAACCAACATAAGAGGAAAAGAGCAGTCTATTAAGGCTGTG 1260
 ACACCTTCTCTGGATCCAAACCAACATAAGAGGAAAAGAGCAGTCTATTAAGGCTGTG 1260

Lewis V E K I I P E K A L A S G I T M S E E V
 DSS GTGGAAAAATCATCCCGAGAAAGCACTGGCTTCAGGATTAATATGAGCAGCGAAGTC 1320
 GTGGAAAAATCATCCCGAGAAAGCACTGGCTTCAGGATTAATATGAGCAGCGAAGTC 1320

Lewis V N E I L S N S E G S K K D P R V P K L
 DSS GTTAATAGSATACTTTCCAACTCAGAGGSAAGTAAGAAAGACCTCCGAGTCCCAAAGCTG 1380
 GTTAATAGSATACTTTCCAACTCAGAGGSAAGTAAGAAAGACCTCCGAGTCCCAAAGCTG 1380

Lewis S K M I E N E S P S V G L E T G A N A E
 DSS AGTAAGATGATAGAGAATGAGTCCCCCTCAGTTGGCTTCGAAACAGGTGCGAATGCTGAG 1440
 AGTAAGATGATAGAGAATGAGTCCCCCTCAGTTGGCTTCGAAACAGGTGCGAATGCTGAG 1440

Lewis K I V P G G V S K Q R K P P L V M T S P
 DSS AAAATTGTCCTGGAGGTGTGTCCAACAGAGGAAACCAACCCCTGGTCATGACGTCCAGG 1500
 AAAATTGTCCTGGAGGTGTGTCCAACAGAGGAAACCAACCCCTGGTCATGACGTCCAGG 1500

Lewis T R A E H A P S G K L P E I Q H P K F A
 DSS ACACGCGCAGAGCATGCTCCGTCAGGAAAGCTGCCCTGAGATCCAGCACCCCAAGTTTGTCT 1560
 ACACGCGCAGAGCATGCTCCGTCAGGAAAGCTGCCCTGAGATCCAGCACCCCAAGTTTGTCT 1560

Lewis A K R R C S K A K F P A M L R E A V L A
 DSS GCAAAAAGGAGGTGTAGCAAGSCAAAACCTCCGECTATGCTTCGAGAGGCACTTTAGCC 1620
 GCAAAAAGGAGGTGTAGCAAGSCAAAACCTCCGECTATGCTTCGAGAGGCACTTTAGCC 1620

Lewis T A E K L M V E P P S A Y P I T F S S P
 DSS ACCGCGCAGAACTGATGTTGGAGCCGCGTCTGCTTACCCATTACCCATCCAGCCCT 1680
 ACCGCGCAGAACTGATGTTGGAGCCGCGTCTGCTTACCCATTACCCATCCAGCCCT 1680

Lewis L Y T N T D S L T V I T P V K K K R G R
 DSS CTCTACCAACACCGACAGTCTTACTGTGATCAAGCCGCTCAAAAAGAGCGGGACBA 1740
 CTCTACCAACACCGACAGTCTTACTGTGATCAAGCCGCTCAAAAAGAGCGGGACBA 1740

Lewis P K K Q F L L T V E T I H E G T S T S P
 DSS CCAAAGAAAGCAACCTCTGCTCACCGTGGAGACATTCAAGAAAGTACCTCCACTAGCCCT 1800
 CCAAAGAAAGCAACCTCTGCTCACCGTGGAGACATTCAAGAAAGTACCTCCACTAGCCCT 1800

Lewis V S P I S R E F P G T K K R K R R N L
 DSS GTCAAGCCGATTAGCAGAGAGTTTCCGSAACCAAGAAAGAAAGCGGAGGCGCAATCTG 1860
 GTCAAGCCGATTAGCAGAGAGTTTCCGSAACCAAGAAAGAAAGCGGAGGCGCAATCTG 1860

Lewis A K L A Q L V P G E D K P M S E M E F H
 DSS SCTAAGTTGGCCAGTTAGTCCCGAGAGGACAAACCCATGAGTGAATGAATTTTAC 1920
 SCTAAGTTGGCCAGTTAGTCCCGAGAGGACAAACCCATGAGTGAATGAATTTTAC 1920

Lewis
 DSS
 K K V G E L G V L D K E T I E T I N K M
 AAGAAGGTTGGAAAGCTTGGTGTGTGGATAAGAGACCATCAAACGATCAACAAAGATG 1980
 AAGAAGGTTGGAAAGCTTGGTGTGTGGATAAGAGACCATCAAACGATCAACAAAGATG 1980

Lewis
 DSS
 K T L K R K N I L N Q I L S C B S S V A
 AAGAGCTCAAGAGAAAAATATCTTAAATCAGATCTTGTCTTGTCCAGTAGTGTGGCC 2040
 AAGAGCTCAAGAGAAAAATATCTTAAATCAGATCTTGTCTTGTCCAGTAGTGTGGCC 2040

Lewis
 DSS
 L K A K A P P E T S P G A A S I E S K L
 CTAAGGCCAAAGCTCCCCAGAGACCCAGCCCTGGGGCAGCATCTATCGAAAGCAAACCTG 2100
 CTAAGGCCAAAGCTCCCCAGAGACCCAGCCCTGGGGCAGCATCTATCGAAAGCAAACCTG 2100

Lewis
 DSS
 G E Q I N V S E R G T I Y I G H E R G R
 GGAAAGCAAATTAATGTGAGCAAGAGAGAACCATCTACATTGGCAAGAGAGGGGCAGG 2160
 GGAAAGCAAATTAATGTGAGCAAGAGAGAACCATCTACATTGGCAAGAGAGGGGCAGG 2160

Lewis
 DSS
 K P R T E L F P P E E E F K T A I K H F
 AAGCCGAGAACAGAGCTGCCACCCCATCCGAGSAAACCCAAAACAGCCATCAAGCAACCC 2220
 AAGCCGAGAACAGAGCTGCCACCCCATCCGAGSAAACCCAAAACAGCCATCAAGCAACCC 2220

Lewis
 DSS
 R F V S S Q P E V P A V P S S F Q S P V
 AGGCTGTCTCTAGCCAGCCGGATGTTCCAGCCGTGCTTCCAGCTTCCAGTCACTGTG 2280
 AGGCTGTCTCTAGCCAGCCGGATGTTCCAGCCGTGCTTCCAGCTTCCAGTCACTGTG 2280

Lewis
 DSS
 A S S P A A V H P L A T Q L G G S H G N
 GCTTCTTCAACAGAGCCGTGCACCCACTTGCAACCCAGTTAGGTGGGTCCCATGGCAAC 2340
 GCTTCTTCAACAGAGCCGTGCACCCACTTGCAACCCAGTTAGGTGGGTCCCATGGCAAC 2340

Lewis
 DSS
 L S F A S T E T N F S E L K T M P N L Q
 CTGAGCCCGCCAGCACTGAAACCAATTTTCCGAAATGAAAACATGCGGAACCTCCAG 2400
 CTGAGCCCGCCAGCACTGAAACCAATTTTCCGAAATGAAAACATGCGGAACCTCCAG 2400

Lewis
 DSS
 F I S A L P T K T Q K G I H S G T W K L
 CCCATCAGTGTCTTCCAAACAAAACCCAGAGGGGATCCACAGTGGGACTTGGAAAGCTC 2460
 CCCATCAGTGTCTTCCAAACAAAACCCAGAGGGGATCCACAGTGGGACTTGGAAAGCTC 2460

Lewis
 DSS
 S P P R L M A N S P E H L C E I G S L E
 TCCCCACCCCGGCTGATGGCCAACTCCCTTCCGACCTTTCGAGATTGGCTCGCTGAAG 2520
 TCCCCACCCCGGCTGATGGCCAACTCCCTTCCGACCTTTCGAGATTGGCTCGCTHAAG 2520

Lewis
 DSS
 S I T L S P V S E S H S E E T I P S D S
 GAGATCAGCTATCCCCAGTGAAGTGAAGTCCACAGTGAAGGAGACCATCCCGAGCGACAAT 2580
 GAGATCAGCTATCCCCAGTGAAGTGAAGTCCACAGTGAAGGAGACCATCCCGAGCGACAAT 2580

Lewis
 DSS
 G I G T D N N S T S D Q A E K E S E S R
 GGCATTGGGACCGACAACACAGCACCTCCGACCGAGGACAGAGAGAGGTTCCGAATCCCG 2640
 GGCATTGGGACCGACAACACAGCACCTCCGACCGAGGACAGAGAGAGGTTCCGAATCCCG 2640

Lewis
 DSS
 R R Y S F D F C S L D N P E A I P P D T
 CGGAGTACTCCTTGGACTTCTGTCCCTTGACAACCCGGAGGCCATTCCACTGACACC 2700
 CGGAGTACTCCTTGGACTTCTGTCCCTTGACAACCCGGAGGCCATTCCACTGACACC 2700

Lewis
 DSS
 S T K N R H G H R Q K H L I V D T F L A
 AGCACAAGAGACCGGACCGCCACCGGACAGAGCATCTCATTGTGGACACTTTTCTGGCC 2760
 AGCACAAGAGACCGGACCGCCACCGGACAGAGCATCTCATTGTGGACACTTTTCTGGCC 2760

Lewis
 DSS
 H E S L K K P K H K R K E S L Q N R D
 CACGAAAGCTTCAAGAGCCAAAGCACAAGAGGAAACGGAAAGCCCTGCAGAACCCGGAT 2820
 CACGAAAGCTTCAAGAGCCAAAGCACAAGAGGAAACGGAAAGCCCTGCAGAACCCGGAT 2820

Lewis DSS D L Q F L A E L E E L I T K F Q V F R I 2880
 GATCTCCAGTTTCTGGCTGAGCTGGAAGAGCTCATCACCAGTTTCAAGTGTTCAGGATC 2880
 GATCTCCAGTTTCTGGCTGAGCTGGAAGAGCTCATCACCAGTTTCAAGTGTTCAGGATC 2880

 Lewis DSS S H R G Y T F Y H E N P Y P S I F R I H 2940
 TCCACCGGGCTACACTTTTACACGAGAAATCCGTACCCAGCATCTTCAGSATTAA 2940
 TCCACCGGGCTACACTTTTACACGAGAAATCCGTACCCAGCATCTTCAGSATTAA 2940

 Lewis DSS F D H Y Y P V P Y I Q Y E P L L Y L R R 3000
 TTTGATCACTATTACCCGGTGCCATATATCCAGTACGACCCATTACTCTACCTTCGTAG 3000
 TTTGATCACTATTACCCGGTGCCATATATCCAGTACGACCCATTACTCTACCTTCGTAG 3000

 Lewis DSS T S D L E E E K E R G R P A E T N D T M 3060
 ACTTCAGACTGAAGTCAAGAAGAGCTGGTAGGCCCTGCAAGACCAATGACACCATG 3060
 ACTTCAGACTGAAGTCAAGAAGAGCTGGTAGGCCCTGCAAGACCAATGACACCATG 3060

 Lewis DSS T K V P F L Q G F E Y F I P E S Y Y A 3120
 ACAAGGTGCCCTTTTACAGGGSTTCAECTACCCATTCOCAGTGAAGTACTATGCA 3120
 ACAAGGTGCCCTTTTACAGGGSTTCAECTACCCATTCOCAGTGAAGTACTATGCA 3120

 Lewis DSS F Y G M P Y T S M P M N L S Y Y G Q Y 3180
 CCTATGGAATGCTTACACATCAATGCTTATGATGAACCTTGGTACTATGGTCAGTAT 3180
 CCTATGGAATGCTTACACATCAATGCTTATGATGAACCTTGGTACTATGGTCAGTAT 3180

 Lewis DSS P A F L Y L S H T L G A A S F F M R P T 3240
 CCAGCCCTCTCTATCTGTCCACACACTTGGAGCAGCATCCCCATTCTAGGSCAAC 3240
 CCAGCCCTCTCTATCTGTCCACACACTTGGAGCAGCATCCCCATTCTAGGSCAAC 3240

 Lewis DSS V P F P Q F H T S E H V K M S G A T K H 3300
 GTGCCACGCCCCAGTTCACACAAGTTCOCAGTCAAGATGTCCGGSCAACTAAACAT 3300
 GTGCCACGCCCCAGTTCACACAAGTTCOCAGTCAAGATGTCCGGSCAACTAAACAT 3300

 Lewis DSS K A K H S V H L Q S T V G M S L S D I Q 3360
 AAAGCAAAGCATGAGTCCACTTGCAGGAAACCTTAGGCATGGGCTTGGTGACATCCAG 3360
 AAAGCAAAGCATGAGTCCACTTGCAGGAAACCTTAGGCATGGGCTTGGTGACATCCAG 3360

 Lewis DSS P S L N P F P E V G G A A L S E S R L H E 3420
 CCATCTGAAACCTCCCAAGGTGGGCGGGCCCTCTGTCCAGCAGTAGGCTCCACAAG 3420
 CCATCTGAAACCTCCCAAGGTGGGCGGGCCCTCTGTCCAGCAGTAGGCTCCACAAG 3420

 Lewis DSS R K R K R K R K R K E D R I L G T H D N 3480
 AGGAACACAAACATAAACACAAGCACAAAGAAACCGAATCTTAGGGACCCAGACAAT 3480
 AGGAACACAAACATAAACACAAGCACAAAGAAACCGAATCTTAGGGACCCAGACAAT 3480

 Lewis DSS L S G L F A G E A T G F S S H L L S E R 3540
 CTGAGTGGTCTCTTTGCAGSCAAAGCCACAGGCTTCTCCAGTCACTCTGAGCGAGCG 3540
 CTGAGTGGTCTCTTTGCAGSCAAAGCCACAGGCTTCTCCAGTCACTCTGAGCGAGCG 3540

 Lewis DSS L N S S D R E L P L V S E K S H H K E R 3600
 CTCAACAGTTCGGACAAGAGCTCCCGTTGGTGGTGAAGAGCAAAACAAAGGAGAGA 3600
 CTCAACAGTTCGGACAAGAGCTCCCGTTGGTGGTGAAGAGCAAAACAAAGGAGAGA 3600

 Lewis DSS Q K R Q R G D A S H X V S K N N F E V D 3660
 CAGAAGCACCAGCATGGTGAAGCCAGCCCAAGSTTTCTAAGAACAACITTTAGGTGGAC 3660
 CAGAAGCACCAGCATGGTGAAGCCAGCCCAAGSTTTCTAAGAACAACITTTAGGTGGAC 3660

 Lewis DSS T L S T L S L S D A Q H M T Q A K K K Q 3720
 ACCTGTGACGCTGTCACTTTCGATGCCAGCACTGACACAGGCTAAGSACAGGGT 3720
 ACCTGTGACGCTGTCACTTTCGATGCCAGCACTGACACAGGCTAAGSACAGGGT 3720

```

Lewis      D L E N E P V E S C A R R Y S S S G G D
DSS       GACTTGAGCAACGAGCCTGTGGASTCGTGGCTAAAAAGTACTCTGGCASTGGCGGAGAC 3780
          GACTTGAGCAACGAGCCTGTGGASTCGTGGCTAAAAAGTACTCTGGCASTGGCGGAGAC 3780
          *****
Lewis      S T R S E S L E V F S E M N P S S S K W
DSS       ABCACAAGTTCAGAGAUTCTGGATGTGTTCASTGAAATGAAATCCTTCCAGCGACAAGTGG 3840
          AGCACAAGTTCAGAGAUTCTGGATGTGTTCASTGAAATGAAATCCTTCCAGCGACAAGTGG 3840
          *****
Lewis      D S D M S G S E R R G F E G F G T Y R E
DSS       GCAGTGTACATGAGTGGGASTAAGAGGAGGGGCTTTGAAAGGCTTTGGGACGTACAGGGAG 3900
          GCAGTGTACATGAGTGGGASTAAGAGGAGGGGCTTTGAAAGGCTTTGGGACGTACAGGGAG 3900
          *****
Lewis      K D I Q A F K M N R K E R G S Y E S S H
DSS       AAGSACATCCAGGCCCTTCAAGATGAAATCGCAAGSAGAGAGSTTCTTACBAGTCTTCGATG 3960
          AAGSACATCCAGGCCCTTCAAGATGAAATCGCAAGSAGAGAGSTTCTTACBAGTCTTCGATG 3960
          *****
          ↓
Lewis      T P G M P S P H L K V D Q T T V H S K S
DSS       ACTCCAGGGATGCCAAGTCCCCACTTAAAGTGGACAGACAGTACACAGTAAAGAGC 4020
          ACTCCAGGGATGCCAAGTCCCCACTTAAAGTGGACAGACAGTACACAGTAAAGAGC 4020
          *****
Lewis      E G E V S T M H A R K K E P T A V D E V A
DSS       GAGGGCTCAGTATCCACCATGATGACCAGGAAAGAGCCACCGCAGTGGACAGTGTGCA 4080
          GAGGGCTCAGTATCCACCATGATGACCAGGAAAGAGCCACCGCAGTGGACAGTGTGCA 4080
          *****
          ↓
Lewis      I P E A P V L E L L A A S A A T S D A A
DSS       ATTCCCTCAGCGCCCGTGTATCTCTCCTTGTGTCATCTGCAGCAACGTGGATGCAGCC 4140
          ATTCCCTCAGCGCCCGTGTATCTCTCCTTGTGTCATCTGCAGCAACGTGGATGCAGCC 4140
          *****
Lewis      S S S L K K R F R R R E I E A I Q C E V
DSS       AGCTCCCTCCTGAAAGAGCGATTCAAGAGACGGGAGATCGAAGCCATCCAATGGGAAGTG 4200
          AGCTCCCTCCTGAAAGAGCGATTCAAGAGACGGGAGATCGAAGCCATCCAATGGGAAGTG 4200
          *****
Lewis      R K M C H Y T E L L S T E K N L D H V N
DSS       CGGAAGATGTGCCATTACCAAGCCTTCTGTCCACCAAGAGAAACCTGGACCATGTGAAC 4260
          CGGAAGATGTGCCATTACCAAGCCTTCTGTCCACCAAGAGAAACCTGGACCATGTGAAC 4260
          *****
          ↓
Lewis      K I L K A K R L Q R Q S K T G N H F P P
DSS       AAGATCCTGAAGGCCAAGCGGCTGCAGAGACAATCAAAGACGGGCAACAACCTTCCACCA 4320
          AAGATCCTGAAGGCCAAGCGGCTGCAGAGACAATCAAAGACGGGCAACAACCTTCCACCA 4320
          *****
Lewis      P P P P P L P L K T S R G G K R K H R P
DSS       CCACCGCCCCCGCCACCGCTGCCCAAGADCTCTCGGGGCGGGAAGAGAAAACACAGACCG 4380
          CCACCGCCCCCGCCACCGCTGCCCAAGADCTCTCGGGGCGGGAAGAGAAAACACAGACCG 4380
          *****
Lewis      Q P P A Q P A Q P P P Q P L P Q E E E V
DSS       CAGCCCCCAGCGCAGCCCGCCAGCCCCCGCGAGCCCTTCCCAAGSAGSAGSAGTGG 4440
          CAGCCCCCAGCGCAGCCCGCCAGCCCCCGCGAGCCCTTCCCAAGSAGSAGSAGTGG 4440
          *****
Lewis      X A K R F R K S R A S E S D V L P -
DSS       AAGGCCAAAAGGCCACGGGAAGTCCCGAGCCAGCGAGAGTGTATGTCCTTCCCTAG 4494
          AAGGCCAAAAGGCCACGGGAAGTCCCGAGCCAGCGAGAGTGTATGTCCTTCCCTAG 4494
          *****

```

(8) *ATPase, class II, type 9b (Atp9b)*

```

Lewis      M A D Q I P L Y F V R S A A A A A S H
DSS       ATGGCGGACAGATTCCCTGTACCCCGTGGCAGCGCGGCGGCGGCGGCGGCGGCGGCGG 60
          ATGGCGGACAGATTCCCTGTACCCCGTGGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
          *****
          ↓
Lewis      R R A A Y Y S A V G P G F G A D R R G R
DSS       AGACGTGCGGCTACTACAGCGCCGTGGGCCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
          AGACGTGCGGCTACTACAGCGCCGTGGGCCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
          *****

```

```

Lewis      Y Q L E D E S A H L D E M P L M H E E E
DSS        TACCAGCTGGAGGATGAGTCTGCCCATTTGGATGAAATGGCACTGATGATGTCGGAAGAA 180
          TACCAGCTGGAGGATGAGTCTGCCCATTTGGATGAAATGGCACTGATGATGTCGGAAGAA 180
          *
          *
Lewis      G F E N D E S D Y H T L F R A R I T R R
DSS        GGCTTTGAGAAATGACGAAAGTGATTACCAACATTACTCGAGCTAGGATAACACGGAGA 240
          GGCTTTGAGAAATGACGAAAGTGATTACCAACATTACTCGAGCTAGGATAACACGGAGA 240
          *
          *
Lewis      E R G L E W F V C G G W K F L C T S C C
DSS        GAGAGAGGCTGGAGTGGTTTGTCTCGGAGGATGGAATTCCTCTGTACAAGTTGCTGT 300
          GAGAGAGGCTGGAGTGGTTTGTCTCGGAGGATGGAATTCCTCTGTACAAGTTGCTGT 300
          *
          *
Lewis      D W L I N V C Q R E K E L K A R T V W L
DSS        GATTGGCTGATAAATGTTTGTCAAGGAAAGAAAGAACTAAAAGCTCGAACAGTATGGCTT 360
          GATTGGCTGATAAATGTTTGTCAAGGAAAGAAAGAACTAAAAGCTCGAACAGTATGGCTT 360
          *
          *
Lewis      G C P E E C E E K H P R N S I R H Q K Y
DSS        GGATGCTCCAGAGAAAGTGTGAGAAAAACATCCTAGGAACCTATCAAAAATCAGAAATAC 420
          GGATGCTCCAGAGAAAGTGTGAGAAAAACATCCTAGGAACCTATCAAAAATCAGAAATAC 420
          *
          *
Lewis      N V E T F I P G V L V E Q F E F L N L
DSS        AATGATTTACCTTTATACCTGGGTTTTGTATGAAACAAITCAAGTTTTCTTGAATCTC 480
          AATGATTTACCTTTATACCTGGGTTTTGTATGAAACAAITCAAGTTTTCTTGAATCTC 480
          *
          *
Lewis      Y F L V V S C S Q F V P A L K I G Y L Y
DSS        TATTTTCTGGTAGTGTCTCTGCTCACAGTTTGTACCAGCATTGAAAATGGCTACCTCTAC 540
          TATTTTCTGGTAGTGTCTCTGCTCACAGTTTGTACCAGCATTGAAAATGGCTACCTCTAC 540
          *
          *
Lewis      T Y W A P L G F V M A V T I A R E A I D
DSS        ACCTACTGGGCTCCTCTGGGATTTGTATGTCATGTCCTTACTATTGCACGTGAAGCAATTGAT 600
          ACCTACTGGGCTCCTCTGGGATTTGTATGTCATGTCCTTACTATTGCACGTGAAGCAATTGAT 600
          *
          *
Lewis      E F R R F Q R D K E M H S Q L Y S K L T
DSS        GAATTTCCAGCTTTTCAGCGAGACAAGGAAATGAATTCACAGTTATATAGCAAGCTTACA 660
          GAATTTCCAGCTTTTCAGCGAGACAAGGAAATGAATTCACAGTTATATAGCAAGCTTACA 660
          *
          *
Lewis      V R G K V Q V E S S D I Q V G D L I I V
DSS        GTAAGAGGTAAAGTTCAAGTTAAGAGTTCAGACATACAAGTTGGAGACCTCATCATAGTG 720
          GTAAGAGGTAAAGTTCAAGTTAAGAGTTCAGACATACAAGTTGGAGACCTCATCATAGTG 720
          *
          *
Lewis      E K N Q R I P S D M V F L R T S E K A G
DSS        GAAAAGAACCAAGAAATTCATCAGATATGGTATTTCTTAGAACTTCAGAGAAAGCAGGT 780
          GAAAAGAACCAAGAAATTCATCAGATATGGTATTTCTTAGAACTTCAGAGAAAGCAGGT 780
          *
          *
Lewis      S C F I R T D Q L D G E T D M K L E V A
DSS        TCATGTTTTATTGAAAGGACCCAGCTAGATGGTGGAGCCGACTGGAAAGCTGAAGGTGGG 840
          TCATGTTTTATTGAAAGGACCCAGCTAGATGGTGGAGCCGACTGGAAAGCTGAAGGTGGG 840
          *
          *
Lewis      V S C T Q R L P A L S D L F S I S A Y Y
DSS        GTGAGCTGCACACAACGGCTGCCAGCCCTAGGGGACCTTTTTCTATCAGTGCCTATGTT 900
          GTGAGCTGCACACAACGGCTGCCAGCCCTAGGGGACCTTTTTCTATCAGTGCCTATGTT 900
          *
          *
Lewis      Y A Q K P Q L S I N S F E G T F T R D D
DSS        TATGCTCAGAAACCAACTGGATATTCATAGTTTTGAAAGAACATTTACCAGGGATGAC 960
          TATGCTCAGAAACCAACTGGATATTCATAGTTTTGAAAGAACATTTACCAGGGATGAC 960
          *
          *
Lewis      S D F P I H E S L S I E N T L W A S T I
DSS        AGTGAATCTCCCATTCATGAGAGTCTCAGCATAGAAAACACAATGTGGSCAAGCACCATT 1020
          AGTGAATCTCCCATTCATGAGAGTCTCAGCATAGAAAACACAATGTGGSCAAGCACCATT 1020
          *
          *

```

↓
 Lewis V A S E T V I G V V I Y T E K E T R S V 1080
 DSS GTTGCCTCAGGTACTGTAATAGGTGTTGTCAATTATACTGGAAAAGAGACTCGAAGTGTA 1080

↓
 Lewis M N T S N P K N K V G L L D L E L N Q L 1140
 DSS ATGAACACATCCAATCCAAAAATAAGGTTGGTTTATTGGACCTCGAACTCAATCAGCTG 1140
 ATGAACACATCCAATCCAAAAATAAGGTTGGTTTATTGGACCTCGAACTCAATCAGCTG 1140

Lewis T R A L F L A L V V L S V V M V T L Q G 1200
 DSS ACAAAAGCGCTATTTCTGGCTTTAGTTGTTCTTCTGTGTATGTAACCTTACAGGSA 1200
 ACAAAAGCGCTATTTCTGGCTTTAGTTGTTCTTCTGTGTATGTAACCTTACAGGSA 1200

Lewis F A G P W Y R N L F R F L L L F S Y I I 1260
 DSS ITTGCAGGCCCATGAGTACCGTAATCTTTTTCGGTTCTCTCTCTCTTTCTTCTACATCATT 1260
 ITTGCAGGCCCATGAGTACCGTAATCTTTTTCGGTTCTCTCTCTCTTTCTTCTACATCATT 1260

↓
 Lewis F I S L R V N L D M G K A A Y G W M I M 1320
 DSS CCCATAAGTCTGCGAGTGAACCTGGACATGGCCAAAGCAGCATATGGATGGATGATTATG 1320
 CCCATAAGTCTGCGAGTGAACCTGGACATGGCCAAAGCAGCATATGGATGGATGATTATG 1320

Lewis K D E N I F G T V V R T S T I F E E L G 1380
 DSS AAAGATGAGAAATATCTCGGTACAGTTGTTCCGACCCAGCACTATACCAGAGGAACCTGSA 1380
 AAAGATGAGAAATATCTCGGTACAGTTGTTCCGACCCAGCACTATACCAGAGGAACCTGSA 1380

↓
 Lewis R L V Y L L T D E T G T L T Q N E M V F 1440
 DSS CGCTGGTGTACTTACTGACAGACAAAACAGGAACGCTCACTCAGAACGAAATGTTGTT 1440
 CGCTGGTGTACTTACTGACAGACAAAACAGGAACGCTCACTCAGAACGAAATGTTGTT 1440

Lewis X R L R L G T V S Y G T D T M D E I Q S 1500
 DSS AAGCGCTCCATCTGGGTACTGTCTCTATGGGACAGACACTATGGACGAAATCCAAAGC 1500
 AAGCGCTCCATCTGGGTACTGTCTCTATGGGACAGACACTATGGACGAAATCCAAAGC 1500

↓
 Lewis H V L N S Y L G V H S Q T S G H N P S S 1560
 DSS CAGTCCGTAATTCCTACTTGCAGGTACACTCCCAAACAAGTGGCCACAACCCAAATTTCT 1560
 CAGTCCGTAATTCCTACTTGCAGGTACACTCCCAAACAAGTGGCCACAACCCAAATTTCT 1560

Lewis A P L R R S Q E S T P K V K E S V S S R 1620
 DSS GCTCCACTTGAAGAAGCCAGTCTCAACCCCTAAAGTTAAGAAGAGTGTGAGCAGTABA 1620
 GCTCCACTTGAAGAAGCCAGTCTCAACCCCTAAAGTTAAGAAGAGTGTGAGCAGTABA 1620

Lewis I H E A V K A I A L C H N V T P V Y E A 1680
 DSS ATCCATGAAGCAGTGAAGGCCATCGCCCTTTGTCAATATGTCACTCTGTGTATGAGGCT 1680
 ATCCATGAAGCAGTGAAGGCCATCGCCCTTTGTCAATATGTCACTCTGTGTATGAGGCT 1680

Lewis R T G I T G E T E F A E A D Q D F S D E 1740
 DSS CCGACTGGCATTCACTGGAGGACTGAGTTGCTGAAGCAGACCAGACTTTAGTGATGAG 1740
 CCGACTGGCATTCACTGGAGGACTGAGTTGCTGAAGCAGACCAGACTTTAGTGATGAG 1740

↓
 Lewis N R T Y Q A S E P D E V A L Y R W T E S 1800
 DSS AACCCCACTTACAGGCCTCCAGCCCGATGAGTGGCACTGGTGGATGGACAGAGAT 1800
 AACCCCACTTACAGGCCTCCAGCCCGATGAGTGGCACTGGTGGATGGACAGAGAT 1800

Lewis V G L T L V S R D L A S N Q L R T P S G 1860
 DSS GTGGCCCTCACTCTGGTCAGCAGGACCTTGCTTCCATGCACTGAAGACCCCAAGTGGC 1860
 GTGGCCCTCACTCTGGTCAGCAGGACCTTGCTTCCATGCACTGAAGACCCCAAGTGGC 1860

Lewis Q V L T Y C I L Q H F F F T S E S K R M 1920
 DSS CAGGTCTTGACCTACTGCACTTACAGATGTTCCCTTCACTCGGAGAGCAAGGGATG 1920
 CAGGTCTTGACCTACTGCACTTACAGATGTTCCCTTCACTCGGAGAGCAAGGGATG 1920

```

      ↓
Lewis  I I S T M Q A V F S S V F Y F A S V P L 2880
DSS   ATCATTTCACCATGCAGGCTGTGTTCCTCTCGSTCTTCTACTTTGCATCTGTCCCGCTG 2880
      *
      *
      *
      ↓
Lewis  Y Q G F L M V G Y A T I Y T M F E V F S 2940
DSS   TACCAGGGATTTCCTTATGSTAGSSTATGCAACTATCTACACCATGTTCCDGGTGTCTCC 2940
      *
      *
      *
Lewis  L V L D Q Q D V K P E M A I L Y P E L Y K 3000
DSS   TTAGTGTGGACCAAGATGTAAGSCCAGAAATGGCAATTCTCTACCCAGAGCTGTACAAG 3000
      *
      *
      *
      ↓
Lewis  D L T K G R S L S F K T F L I M V L I S 3060
DSS   GACCTCACCAAGGGAGATCTCTTCCTTCAAGACCTTCTCATCTGGHTTTTGATCAST 3060
      *
      *
      *
      ↓
Lewis  I Y Q G S I L M Y G A L L L F E A E F V 3120
DSS   ATTTACCAAGGAGCATCTCATGTACGGGGCCTGCTGCTCTTGGAGCCGAGTTTGTG 3120
      *
      *
      *
Lewis  H V V A I S F T A L I L T E L L M V A L 3180
DSS   CATGTTGTGGCCATCTCCTTCAACAGCACTGATCCTGACCGAGCTGCTCATGTTGGCCCTG 3180
      *
      *
      *
Lewis  T I R T W H W L M V V A E F L S L G C Y 3240
DSS   ACTATCAGSACTTGGCACTGGCTGATGTTGGTGGGGAGTCTCTCAGCCTGGGCTGCTAT 3240
      *
      *
      *
      ↓
Lewis  V A E L A F L N E Y F G I G R V S F G A 3300
DSS   GTTGCCTCCCTTGCITTTTCTCAATGAGTATTTTGGTATAGGCAGAGTGTCTTTTGGCGCT 3300
      *
      *
      *
      ↓
Lewis  F L D V A F I T T V T F L W K V S A I T 3360
DSS   TTCTTAGATGTTGCCCTTCAACACAAGTGTGACCTTCTTGTGGAAAGTATCAGCTATCACT 3360
      *
      *
      *
Lewis  V V S C L P L Y V L K Y L K R K L S P F 3420
DSS   GTGTCAGTGCCTCCACTCTACGTACTCAAGTACTTGAAGABAAGTTGTCCCTCC 3420
      *
      *
      *
Lewis  E Y S K L S S - 3444
DSS   AGCTACTCCAAGCTGTCTCCTCGA 3444
      *
      *
      *

```

(9) Sal-like 3 (*Drosophila*) (*Sall3*)

```

Lewis  M D V E A S A D Q G P P G P E V P P P P 60
DSS   ATGATGTGGAGGCTTCTGCGGACCAGGGCCCTCCAGGCCCAAGTGTGCCCCACCCGCA 60
      *
      *
      *
Lewis  F A L P P Q P E P A A P S M P S T H V T 120
DSS   CTTGCACTGCCACCACAGCCAGAGCCTGCGGCCTTCAAGATGCTTAGTACCAATGTGACC 120
      *
      *
      *
Lewis  L E T L L S T K V A V A Q F S Q G A R A 180
DSS   CTGGAGACGCTGCTGAGCACCAGSTGGCCGTGGCACAATCTCACAGSSTGCBCGTGCA 180
      *
      *
      *

```

Lewis
 DSS
 G G T T S A G G S V G A V A I P M I L E
 GGCGCACCACAGGCGCTGGTGGCAGCGTGGGTGCTGTGGCCATCCCATGATCCTAGAG 240
 GGCGCACCACAGGCGCTGGTGGCAGCGTGGGTGCTGTGGCCATCCCATGATCCTAGAG 240

Lewis
 DSS
 Q L V A L Q Q Q Q I H Q L L I E Q I R
 CAGCTGGTGGCACTGCAGCAGCAGATCCACCAGCTTCAGCTCATCGAGCAGATCCGC 300
 CAGCTGGTGGCACTGCAGCAGCAGATCCACCAGCTTCAGCTCATCGAGCAGATCCGC 300

Lewis
 DSS
 S Q V A L M S R Q P G P F L K P S A S A
 AGCCAGGTGGCCCTAATGAGCCGGCAGCCTGGGCGCTCCGCTGAAGCCCTCAGCCAGTGC 360
 AGCCAGGTGGCCCTAATGAGCCGGCAGCCTGGGCGCTCCGCTGAAGCCCTCAGCCAGTGC 360

Lewis
 DSS
 F G A A S V Q L P G L T P H A E L Q L E
 CCTGGAGCAGCCTCAGTACAGCTTCGGGCGCTGACTCCGATGCGTCCCTCCAGCTTTC 420
 CCTGGAGCAGCCTCAGTACAGCTTCGGGCGCTGACTCCGATGCGTCCCTCCAGCTTTC 420

Lewis
 DSS
 A G F A T A S A G E S T L S A A F D S
 GCTGGACCAGCCACTGCCTGCTGCTCAGGCTCCACCCTGTGGCAGCCTTCGATGSC 480
 GCTGGACCAGCCACTGCCTGCTGCTCAGGCTCCACCCTGTGGCAGCCTTCGATGSC 480

Lewis
 DSS
 F Q H L S Q P Q P A S S T S T F C S T S
 CCCCAGCAGCTGTGGCAGCGCAGCCCGCTTCGGCACAAGCAACCCCTGTAGCACTAGT 540
 CCCCAGCAGCTGTGGCAGCGCAGCCCGCTTCGGCACAAGCAACCCCTGTAGCACTAGT 540

Lewis
 DSS
 A G F T D S G A H F A C S T G P A F G A
 GCTGGCCCCACAGATTCTGGGGCACACCCAGCCTGTAGCACTGGGCCAGCTCCAGGAGCT 600
 GCTGGCCCCACAGATTCTGGGGCACACCCAGCCTGTAGCACTGGGCCAGCTCCAGGAGCT 600

Lewis
 DSS
 V A A A S S T V G H T V Q P Q N A S T F
 GTGGCCGACAGATCCAGCACTGTAGGTAAACAGGTGCAAGCCCAAAAATGCATCTACGCC 660
 GTGGCCGACAGATCCAGCACTGTAGGTAAACAGGTGCAAGCCCAAAAATGCATCTACGCC 660

Lewis
 DSS
 F A L G P G P L L S S A S S L F S P L L
 CCTGGCCCTGGGTCCCGGACCCCTCCTCAGCTCAGCCTCCAGTCTGCCAAGCCCTCTGCTA 720
 CCTGGCCCTGGGTCCCGGACCCCTCCTCAGCTCAGCCTCCAGTCTGCCAAGCCCTCTGCTA 720

Lewis
 DSS
 F Q T S S S S V I F P H P L V S I A A T
 CCTCAGACTTCATCCAGCAGTGTATCTTCCCAACCCGCTGGTTCAGCATTGTCCACACA 780
 CCTCAGACTTCATCCAGCAGTGTATCTTCCCAACCCGCTGGTTCAGCATTGTCCACACA 780

Lewis
 DSS
 A N A L D P L S A L M K H R R K G K F P H
 GCCAATGCCCTGGACCCCTCTCGGCTCTTATGAAGCACCGCAAGGGCAAGCCCCCTAAC 840
 GCCAATGCCCTGGACCCCTCTCGGCTCTTATGAAGCACCGCAAGGGCAAGCCCCCTAAC 840

Lewis
 DSS
 V S V F E P K A S A E D P F F K H E C R
 GTTTCAGTGTTCGAACCCAGGCCAGTGTGAGGACCCCTTCTTTAAGCACAAATGCAGG 900
 GTTTCAGTGTTCGAACCCAGGCCAGTGTGAGGACCCCTTCTTTAAGCACAAATGCAGG 900

Lewis
 DSS
 F C A K V F G S D E A L Q I H L R S H T
 TTCTGTGCCAAGGTCTTCGGCAGTGCAGCGCGTTGCAGATCCACCTGGGATCCACACA 960
 TTCTGTGCCAAGGTCTTCGGCAGTGCAGCGCGTTGCAGATCCACCTGGGATCCACACA 960

Lewis
 DSS
 G E R F F K C H I C G R R F S T K G N L
 GGGGACGGCCCTTCAAATGTAACATCTGGGGGAACCGCTTTTCCACCAAGGGCAACCTG 1020
 GGGGACGGCCCTTCAAATGTAACATCTGGGGGAACCGCTTTTCCACCAAGGGCAACCTG 1020

Lewis
 DSS
 K V H F Q R H K E K Y P H I Q M N P Y P
 AAGTCCACTTCAGAGGCACAAGSAGAAGTACCCACATTCAGATGAACCCCTACCCCT 1080
 AAGTCCACTTCAGAGGCACAAGSAGAAGTACCCACATTCAGATGAACCCCTACCCCT 1080

Lewis V P E Y L D H V P T C S G I P Y G N S L 1140
 DSS GTCCAGAAATACCTCGACAATGTGCCACCTGCTCTGSGATTCCCTACGGCATGTCCCTA 1140

 Lewis F P E K F V T T W L D S K P V L P T V F 1200
 DSS CCCCAGAAAAGCCGGTGACCACCTGGCTGGACAGCAAGCCAGTGTCTGCCACTGTACCA 1200
 CCCCAGAAAAGCCGGTGACCACCTGGCTGGACAGCAAGCCAGTGTCTGCCACTGTACCA 1200

 Lewis T S V G L Q L P P T V P G T H N Y T D S 1260
 DSS ACATCAGTAGGGCTTCAGCTGCCDCCACTGTCCCTGGCACCCACAATTACACTGACTCC 1260
 ACATCAGTAGGGCTTCAGCTGCCDCCACTGTCCCTGGCACCCACAATTACACTGACTCC 1260

 Lewis P S I T F I S R S F Q R F S F A S S E C 1320
 DSS CCTAGCATCACTCCTATTAGCCGGTCCCACAGAGGCCCTCTCCAGCATCCAGTGAATGC 1320
 CCTAGCATCACTCCTATTAGCCGGTCCCACAGAGGCCCTCTCCAGCATCCAGTGAATGC 1320

 Lewis T S L S F G L N N T E S G I T V R F E S 1380
 DSS ACCTCTGTCTCCAGGCTCAACAATACAGAATCTGGTATCACAGTGAAGCCGTGAGTCA 1380
 ACCTCTGTCTCCAGGCTCAACAATACAGAATCTGGTATCACAGTGAAGCCGTGAGTCA 1380

 Lewis P Q P L L G S P P L T K A E P V S L P C 1440
 DSS CCCCAGCCACTGTGGGTGGGCTCCGCTTACTAAAGCTGAGCCAGTCAAGCCGTGCTTC 1440
 CCCCAGCCACTGTGGGTGGGCTCCGCTTACTAAAGCTGAGCCAGTCAAGCCGTGCTTC 1440

 Lewis T S T R T G D A F V V G G Q V S G L P T 1500
 DSS ACAGTACAAGSACAGSAGACGCTCCAGTGGTGGTGGGAGGTGAGCCAGTCCAGCCCT 1500
 ACAGTACAAGSACAGSAGACGCTCCAGTGGTGGTGGGAGGTGAGCCAGTCCAGCCCT 1500

 Lewis F A A T T V T D S A C T S L G S P G L F 1560
 DSS CCGGCTGCCACCCTGTCCAGACAGCGCCCTGCACAAGCCCTCGGGAGCCCGGCTTCCA 1560
 CCGGCTGCCACCCTGTCCAGACAGCGCCCTGCACAAGCCCTCGGGAGCCCGGCTTCCA 1560

 Lewis A V S D Q F E A Q F F F G S L L D S M Q 1620
 DSS GCTGTCTCCAGCCAGTTCAGGGCCAGTTTCTTTCCGAGGGCTGCTTGAATCTATGCAA 1620
 GCTGTCTCCAGCCAGTTCAGGGCCAGTTTCTTTCCGAGGGCTGCTTGAATCTATGCAA 1620

 Lewis T S E T S H L Q Q L V E N I D H K M T D 1680
 DSS ACCTCAGAGACTCGAAACTGCASCAGCTGGTGGAGAACATCGACAAGAATGACTGAC 1680
 ACCTCAGAGACTCGAAACTGCASCAGCTGGTGGAGAACATCGACAAGAATGACTGAC 1680

 Lewis F N Q C V I C H R V L S C Q S A L K M H 1740
 DSS CCGAACCAGTGTGTCTCTGCCACCGCGTGTGAGCTGCCAGAGCGCACTGAAGATGCAC 1740
 CCGAACCAGTGTGTCTCTGCCACCGCGTGTGAGCTGCCAGAGCGCACTGAAGATGCAC 1740

 Lewis Y R T H T G E R P F K C E I C G R A F T 1800
 DSS TACAGGACGCACACAGGGGAACGGCCCTTCAAATGTAGATCTGTGGCCCGCCTTCAAC 1800
 TACAGGACGCACACAGGGGAACGGCCCTTCAAATGTAGATCTGTGGCCCGCCTTCAAC 1800

 Lewis T E G N L R T H F G V H R A K P P L R V 1860
 DSS ACCAAGGCAACCTGAAGACTCACTTCGGGGTGCACCGTCCCAAGCCCGGCTCCAGTG 1860
 ACCAAGGCAACCTGAAGACTCACTTCGGGGTGCACCGTCCCAAGCCCGGCTCCAGTG 1860

 Lewis Q H S C F I C Q K K F T H A V V L Q Q H 1920
 DSS CAGCATTCTGCCDCCATCTGCCAGAGAAGTTCACAACCGTGTGGTGTGCAACAGCAC 1920
 CAGCATTCTGCCDCCATCTGCCAGAGAAGTTCACAACCGTGTGGTGTGCAACAGCAC 1920

 Lewis I R M H M G G Q I E N T P L P E G L Q E 1980
 DSS ATCCGTATGCACATGGGGGACAGATCCCGAACACGCCACTGCTGAGGGCTTGCAGGAG 1980
 ATCCGTATGCACATGGGGGACAGATCCCGAACACGCCACTGCTGAGGGCTTGCAGGAG 1980

Lewis A M D A E L P F D E K H A E T L E E F D 2040
 DSS GCCATGGATGCCGAACCTGCCCTTTGATGAGAGAATGCGAGAACCCCTCAGCAGCTTTGAC 2040

 Lewis D D V D E N S H E E D S E L K D T A S D 2100
 DSS GAGATGTCGACGAGAATCCATGAGGAGGACTCAGAGCTGAAGGACACAGCCAGCGAC 2100

 Lewis E S K F L L S Y S S S C P P S P F S V I 2160
 DSS TCCTCCAAACCGCTCTTGTCTTACTCGGGCTCCTGTCCACCCCTCACCDCDGTGCGTCACT 2160

 Lewis E S I A A L E N Q M H M I D E V M N C Q 2220
 DSS TCCAGCATGCGCCCTGAGAAATCAGATGAAAATGATGACTGCGTCAAGACTGCCAG 2220

 Lewis Q L A S L K S V E N G S G E S D R L S N 2280
 DSS CAACTGGCCAGCCTGAAGTCGGTGGAAAATGGGTCTGGGAAAAGCGACCGCTTGAGCAAC 2280

 Lewis D S S S A V G S L E E R S A B S P A L S 2340
 DSS SACTCCTCCTCTGCAGTGGGAGACCTGGAGAGCGGAGCGGAGCGCCCTGCCCTGTCT 2340

 Lewis E S S S S Q A L S P A H S N G E S F R S 2400
 DSS GAGTCTCCTCCTCCAGGCTCTGTCACTGCTCAGATGATGATGAGAGCTTCCGCTTC 2400

 Lewis X S F G L S H Q E D P Q E I F L K T E R 2460
 DSS AAGTCGCCAGGGCTTAGCCACCAGGAGATCCGAGGAGATCCCACTGAAGACTGAAAGA 2460

 Lewis L D S P P F S P S N G S A L D L T A G H 2520
 DSS CTGGACAGCCACCCCCCGGCCAGGAAATGGAGTGGCCCTGGACCTGACAGCGGGCCAC 2520

 Lewis F G R P L I K E E A P F S L L F L S R E 2580
 DSS CCTGGTCGGCCACTCATCAAGGAGGAGCCCTTTCAGCCTGCTGTCTTCTGAGCAGAGAA 2580

 Lewis R G K C A S T V C G V C G K F F A C K S 2640
 DSS CGGGSTAAAGTGTGCGAGCACTGTGTGTGTGTCTGTGGCAAGCCCTTTGCTTGCAAAAAT 2640

 Lewis A L E I H Y R S H T K E R P F F V C T V C 2700
 DSS GCCTTGGAAATCCACTACCGCAGCCATACCAAGGAGCGGCGCTTTGTCTGCACAGTCTGC 2700

 Lewis R R G C E T M G N L R Q H L L L Y H K L E 2760
 DSS AGCGGAGGCTGCTCCACTATGGGTAACCTAAAGCAGCACTTACTGACACACAAGTTGAAA 2760

 Lewis E L F S Q V F D P N F T L G F S H S T F 2820
 DSS GAGTGCCTTCTCAGGTGTTTACCCCACTTTACTCTAGGTCCCAGCCACAGCAGCCCT 2820

 Lewis S L A S S P A P T M I K M E V N G H S K 2880
 DSS AGCCTGGCCTCCAGCCCTGGCCCACTGATCAAAATGGAAGTAAACBCTCACAGCAAG 2880

```

Lewis      A I A L S E G P A L P A G V Q V P T G P
DSS        GCCATCGCACTGGGTGAGGGCCCAAGCCCTACCAGCCGGGTCCAGGTACCTACTGGGCC 2940
          SCCATCGCACTGGGTGAGGGCCCAAGCCCTACCAGCCGGGTCCAGGTACCTACTGGGCC 2940
          *****

Lewis      Q T V M S P G L A P M L A P P P R T P
DSS        CAGACAGTGTATGAGCCCTGGCTTGGCCCATGCTGGCACCCCGCCAAGCCGACACCC 3000
          CAGACAGTGTATGAGCCCTGGCTTGGCCCATGCTGGCACCCCGCCAAGCCGACACCC 3000
          *****

Lewis      K Q H N C Q S C G K T F E S A S A L Q I
DSS        AAGCAGCACAACTGTCAGTCACTGTGGGAAGACCTTCTCTCAGCCAGTGGCCCTGAGATC 3060
          AAGCAGCACAACTGTCAGTCACTGTGGGAAGACCTTCTCTCAGCCAGTGGCCCTGAGATC 3060
          *****

Lewis      H E R T H T G E K P F G C T I C G R A F
DSS        CATGAGCCACCCACACTGGGGAGAAGCCCTTGGCTGCACCATCTGTGGCAGGGCCCTTC 3120
          CATGAGCCACCCACACTGGGGAGAAGCCCTTGGCTGCACCATCTGTGGCAGGGCCCTTC 3120
          *****

          ↓

Lewis      T T E G N L E V R H G T H M W N N A P A
DSS        ACTACAAGGGCAATCTCAAGGTACACATGGCCACCCACATGTGGAACAACGGCCCTGGG 3180
          ACTACAAGGGCAATCTCAAGGTACACATGGCCACCCACATGTGGAACAACGGCCCTGGG 3180
          *****

Lewis      R R G R R L S V E N P M A L L G G D A L
DSS        AGGCGTGGCCGCCCTCTGTCTGTGGAAAACCCCATGGCCCTGCTGGGTGGCCAGCCCTTC 3240
          AGGCGTGGCCGCCCTCTGTCTGTGGAAAACCCCATGGCCCTGCTGGGTGGCCAGCCCTTC 3240
          *****

Lewis      K F S E H F Q E D L A A R A H N V S P S
DSS        AAGTTCCTGAAATGTTCCAGAAGGACCTGGCGGCTCGGGCGATGAATGTCCAGCCCGC 3300
          AAGTTCCTGAAATGTTCCAGAAGGACCTGGCGGCTCGGGCGATGAATGTCCAGCCCGC 3300
          *****

Lewis      F W H Q Y A A A I T H G L A H K H N E I
DSS        TTTTGGAACCAAGTATGCCGCTGCCATCACCAATGGGCTCGCCATGAAGAACAATGAGATC 3360
          TTTTGGAACCAAGTATGCCGCTGCCATCACCAATGGGCTCGCCATGAAGAACAATGAGATC 3360
          *****

Lewis      S V I Q N G G I P Q L F V S L G G G A I
DSS        TCAGTCATTGAGAATGGAGGCATCCCTCAGCTCCAGTAAGTCTAGGCGGAGGTGCCATC 3420
          TCAGTCATTGAGAATGGAGGCATCCCTCAGCTCCAGTAAGTCTAGGCGGAGGTGCCATC 3420
          *****

Lewis      P F L G A M A G G V D K T R T G S S P P
DSS        CCGCCTCTGGGTGCCATGGCTGGTGGGTGGACAAGACGGCACTGGCAGTAGTCCACCT 3480
          CCGCCTCTGGGTGCCATGGCTGGTGGGTGGACAAGACGGCACTGGCAGTAGTCCACCT 3480
          *****

          C

Lewis      I V S L D H A S P E T G A S R P F T R F
DSS        ATGTCAGCTTGGACAAAGCAAGCCCGGAGACGGGAGCCAGCCGCGCTTCAAGAGTTC 3540
          ATGTCAGCTTGGACAAAGCAAGCCCGGAGACGGGAGCCAGCCGCGCTTCAAGAGTTC 3540
          *****

          M

Lewis      I E E N E E I G I N -
DSS        ATTGAGGATAACAAAGGAGATTGGAATAAACTAG 3573
          ATTGAGGATAACAAAGGAGATTGGAATAAACTAG 3573
          *****

```

(b). Genes present in the *CI8QTL3*-residing interval(1) *Caudal-type homeobox protein 1 (Cdx1)*

```

Lewis      M G E V P V D Q L S T S S Y H S W L S H
DSS        ATGGGAGAAGTCCCTGTGGACCAGCTCAGCACAAAGCAGCTACCACAGCTGGCTCTCCAC 60
          ATGGGAGAAGTCCCTGTGGACCAGCTCAGCACAAAGCAGCTACCACAGCTGGCTCTCCAC 60
          *
          *
          *
Lewis      A P G A A L K S P G G K G P P T Y A P P
DSS        GCTCCTGGAGCTGCCCTGAAGTCTCCCGGGGAAAGGGCCCTCCAGCTACGCGCCCGG 120
          GCTCCTGGAGCTGCCCTGAAGTCTCCCGGGGAAAGGGCCCTCCAGCTACGCGCCCGG 120
          *
          *
          *
Lewis      G P A P A P E Q Y P D E A G Y T H V E P
DSS        GGCCCGGCGCCCGGCGCCCGGCGAGTACCCCGACTTCGCGGGTACACGCACGTGGAGCG 180
          GGCCCGGCGCCCGGCGCCCGGCGAGTACCCCGACTTCGCGGGTACACGCACGTGGAGCG 180
          *
          *
          *
Lewis      A P A P P P T W A A P F P A P K D D W A
DSS        GCACCCGCGCCCGCTCCGACTGGCTGGCCCTCCCTGCGCCCAAGGACGACTGGCA 240
          GCACCCGCGCCCGCTCCGACTGGCTGGCCCTCCCTGCGCCCAAGGACGACTGGCA 240
          *
          *
          *
Lewis      A A Y G P P P A V P A A S P A P L A F G
DSS        SCTGCATATGGTCCGGGCCCCCGGGTCCCGGCGCCAGCCCGCCCGCTGGCCTTCGGG 300
          SCTGCATATGGTCCGGGCCCCCGGGTCCCGGCGCCAGCCCGCCCGCTGGCCTTCGGG 300
          *
          *
          *
Lewis      P P F D F S P V P A P P G P G P G I L A
DSS        CCCCCCGGACTTTAGCCCAAGTGCCTGGCCCTCCCGGCTCTGGCTGGCATCTGGGG 360
          CCCCCCGGACTTTAGCCCAAGTGCCTGGCCCTCCCGGCTCTGGCTGGCATCTGGGG 360
          *
          *
          *
Lewis      Q S L G S P G A P S S P G V Q R R T P Y
DSS        CAGTCCCTCGGGGTCGCGGAGCACCGTCTCGCCAGGAGTGCAAAAGCGGACGCGCTAC 420
          CAGTCCCTCGGGGTCGCGGAGCACCGTCTCGCCAGGAGTGCAAAAGCGGACGCGCTAC 420
          *
          *
          *
Lewis      E W M R R N V A A A S G G S S S K T R T
DSS        GAATGGATGCGGCCCAACTGGCAGCTGCAGGCGCGGTGGCAGGGTAAAGACTCGGACC 480
          GAATGGATGCGGCCCAACTGGCAGCTGCAGGCGCGGTGGCAGGGTAAAGACTCGGACC 480
          *
          *
          *
Lewis      K D K Y R V V Y T D H Q S L E L E K E F
DSS        AAGSACAAGTACCGCGTGTCTATACAGACCAACAGCTAGAGCTGSAAAAAGGAGTTT 540
          AAGSACAAGTACCGCGTGTCTATACAGACCAACAGCTAGAGCTGSAAAAAGGAGTTT 540
          *
          *
          *
Lewis      H Y S R Y I T I R R K S E L A N L G L
DSS        CACTACAGTCCGTACATCACCATTGCGCGCAAGTCAGAGCTGGCTGCTAACTGGGTCTC 600
          CACTACAGTCCGTACATCACCATTGCGCGCAAGTCAGAGCTGGCTGCTAACTGGGTCTC 600
          *
          *
          *
Lewis      T E R Q Y K I N F Q H R E A K E R E V N
DSS        ACAGAGCGGCAAGTAAAGATCTGTTCCAGAACDGTGCGGCCAAGGAGGTAAGATAAC 660
          ACAGAGCGGCAAGTAAAGATCTGTTCCAGAACDGTGCGGCCAAGGAGGTAAGATAAC 660
          *
          *
          *
Lewis      K K K Q Q Q Q Q Q Q Q Q Q P M P F T Q L
DSS        AAGAAGAAACAGCAGCAGCAGCAGCAACAGCAGCAGCCATGCTCCACACAGTTG 720
          AAGAAGAAACAGCAGCAGCAGCAGCAACAGCAGCAGCCATGCTCCACACAGTTG 720
          *
          *
          *
Lewis      F L P L D G T P T P S G P P L G S L C P
DSS        CCCCTGCCCTGGATGGCACCCACACCATCGGGGCCACCCCTGGGGAGTCTATGCCDC 780
          CCCCTGCCCTGGATGGCACCCACACCATCGGGGCCACCCCTGGGGAGTCTATGCCDC 780
          *
          *
          *

```

```

Lewis      T N A G L L G T P S F V F V K E E F L F
DSS        ACCAATGCTGGTCTTCTGGGCACCCCTCCCAAGTCCCGTCAAGGAGGAGTTTTTACCC 840
           ACCAATGCTGGTCTTCTGGGCACCCCTCCCAAGTCCCGTCAAGGAGGAGTTTTTACCC 840
           *****
Lewis      TAG 843
DSS        TAG 843
           ***

```

(2) Neural precursor cell expressed, developmentally down-regulated 4-like gene (Nedd4l)

```

DSS        M S E S P E R R G E S R I L R V K V V S
Lewis      ATBTCTGAGTCTCCCGAGAGGAGGGGAGAAATCCCGTATTCTCAGAGTAAAAGTTGTTTCT 60
           *****
           |
DSS        G I D L A K K D I F G A S D P Y V K L S
Lewis      GGGATTGACCTCGCCAAAAAGGACATATTTGGAGCCAGTGACCCATACGTGAAGCTCTCG 120
           GGGATTGACCTCGCCAAAAAGGACATATTTGGAGCCAGTGACCCATACGTGAAGCTCTCG 120
           *****
           |
DSS        L Y V A D E N R E L A L V Q T E T I R K
Lewis      TTGTACGTAGCTGACGAGAACAGAGAAGTCCCTCTGGTTCAGACTAAGCAATTA AAAAG 180
           TTGTACGTAGCTGACGAGAACAGAGAAGTCCCTCTGGTTCAGACTAAGCAATTA AAAAG 180
           *****
           |
DSS        T L N P K W N E E F Y F R V N F S N H R
Lewis      ACGCTGAACCCAAAGTGGAAATGAAGAGTTTATTTTAGAGTAAATCCATCTAATCACAGG 240
           ACGCTGAACCCAAAGTGGAAATGAAGAGTTTATTTTAGAGTAAATCCATCTAATCACAGG 240
           *****
           |
DSS        L L F E V F D E N R L L T R D D F L G Q V
Lewis      CTCTTATTTGAAGTATTTGACGAGAACAGATGACACGAGACGACTTTCTGGGCCAGGTG 300
           CTCTTATTTGAAGTATTTGACGAGAACAGATGACACGAGACGACTTTCTGGGCCAGGTG 300
           *****
           |
DSS        D V F L S H L P T E D P T M E R F Y T F
Lewis      GACGTGCCTCTTAGTCACTTCCGACAGAGATCCAACCATGGAGAGACCCTATACATTT 360
           GACGTGCCTCTTAGTCACTTCCGACAGAGATCCAACCATGGAGAGACCCTATACATTT 360
           *****
           |
DSS        K D F L L R P R S H K E R V E G F L R L
Lewis      AAGGATTTTCTCCTGCGACCTAGAAGTCATAAATCACCGGTCAAGGGGTTTTTGAAGTTG 420
           AAGGATTTTCTCCTGCGACCTAGAAGTCATAAATCACCGGTCAAGGGGTTTTTGAAGTTG 420
           *****
           |
DSS        K M A Y M P K N G G Q D E E N S E Q R D
Lewis      AAAATGGCCTACATGCCGAAAAACGGAGGTGAGGATGAAGAAAAACAGCGAGCAGAGGGAC 480
           AAAATGGCCTACATGCCGAAAAACGGAGGTGAGGATGAAGAAAAACAGCGAGCAGAGGGAC 480
           *****
           |
DSS        D M E H G W E V V D E N D S A S Q H Q E
Lewis      GACATGGASCATGGATGGGAAGTTGTTGACTCAAACGACTCAGCTTCCCAGCACCAGGAG 540
           GACATGGASCATGGATGGGAAGTTGTTGACTCAAACGACTCAGCTTCCCAGCACCAGGAG 540
           *****
           |
DSS        E L P P P P L P P G W E E K V D N L G R
Lewis      GAGCTCCCTCCTCCTCCCTGCCACAGGATGGGAAGAGAAAGTGGACAATTTAGGCCGA 600
           GAGCTCCCTCCTCCTCCCTGCCACAGGATGGGAAGAGAAAGTGGACAATTTAGGCCGA 600
           *****
           |
DSS        T Y Y V N H N N R S T Q W H R F S L M D
Lewis      ACTTACTATGTCAACCAACAACAGSAGTACGCAAGTGGCACCAGCCTAGCTTGTATGGAT 660
           ACTTACTATGTCAACCAACAACAGSAGTACGCAAGTGGCACCAGCCTAGCTTGTATGGAT 660
           *****

```



```

V S S E S D N N I R Q I N Q E A A H R R
DSS GTATCGTCAGAGTCAGACAATAACATCAGGCAGATCAATCAGGAGGCCGCACACCGGCGT 720
Lewis GTATCGTCAGAGTCAGACAATAACATCAGGCAGATCAATCAGGAGGCCGCACACCGGCGT 720
*****

F R S R R H I S E D L E P E A S E G G G
DSS TTCCGCTCTCGGAGGCACATTAGTGAAGACTTGGAGCCTGAGGCCTCTGAAGGCGGTGGA 780
Lewis TTCCGCTCTCGGAGGCACATTAGTGAAGACTTGGAGCCTGAGGCCTCTGAAGGCGGTGGA 780
*****

      ↓
E G P E P W E T I S E E V N M A G D S L
DSS GAAGGCCCTGAGCCTTGGGAGACCAATTCAGAGGAAGTGAACATGGCGGGAGATTCTCTC 840
Lewis GAAGGCCCTGAGCCTTGGGAGACCAATTCAGAGGAAGTGAACATGGCGGGAGATTCTCTC 840
*****

S L A L P P P P A S P V S R T S P Q E L
DSS AGCCTGGCTTACCCCACTCTGCTCCCCAGTGTCAAGSACCAGCCCAAGGAGCTG 900
Lewis AGCCTGGCTTACCCCACTCTGCTCCCCAGTGTCAAGSACCAGCCCAAGGAGCTG 900
*****

S E E L S R R L Q I T P D E N G E Q F S
DSS TCGGAAGAAGTGAAGCAGAAGGTTGCAGATCACTCCGGACTCCAATGGGGAACAGTTCAGT 960
Lewis TCGGAAGAAGTGAAGCAGAAGGTTGCAGATCACTCCGGACTCCAATGGGGAACAGTTCAGT 960
*****

      ↓
A L I Q R E P S S R L R S C S V T D T V
DSS GCTCTGATCCAGAGAGAGCCCTCGTCTAGGCTCCGGTCTGCACTGTTACCGACACGGTT 1020
Lewis GCTCTGATCCAGAGAGAGCCCTCGTCTAGGCTCCGGTCTGCACTGTTACCGACACGGTT 1020
*****

      ↓
A E Q A H L P F F S T F T R R A R S S T
DSS GCTGAACAAGCTCACCTTCCACCGCCAGCACCCCACTAGGCGAGCCCGTTCTGCAACT 1080
Lewis GCTGAACAAGCTCACCTTCCACCGCCAGCACCCCACTAGGCGAGCCCGTTCTGCAACT 1080
*****

      ↓
V T G G E E S T F S V A Y V H T T P G L
DSS GTCACGGTGGTGAGGAATCAACGCCATCAGTGGCCTATGTACATACCCAGCCGGGCCTA 1140
Lewis GTCACGGTGGTGAGGAATCAACGCCATCAGTGGCCTATGTACATACCCAGCCGGGCCTA 1140
*****

P S G N E E R K D A K G R T Y Y V N H N
DSS CCTCAGGCTGGGAAGAAAGAAAAGATGCTAAGGAGCCACATACTATGTCAATCATAAC 1200
Lewis CCTCAGGCTGGGAAGAAAGAAAAGATGCTAAGGAGCCACATACTATGTCAATCATAAC 1200
*****

      ↓
N R T T T W T R P I N Q L A E D G A S G
DSS AATCGAACCAACTTGSACTCGACCAATCATGAGCTTGCAGAAGACGGTGCCTCTGGA 1260
Lewis AATCGAACCAACTTGSACTCGACCAATCATGAGCTTGCAGAAGACGGTGCCTCTGGA 1260
*****

S A T N S N N H L V E P Q I R R P R S L
DSS TCAGCCACAACAGTAACAACCACTAGTGAACCCAGATCCGCGGSCCTGTAAGCCTC 1320
Lewis TCAGCCACAACAGTAACAACCACTAGTGAACCCAGATCCGCGGSCCTGTAAGCCTC 1320
*****

      ↓
E E P T V T L E A F L E G G A K D S P I R
DSS AGCTCGCAACAGTAACCTTATCTGCCCACTGGAGGGTGCCAAGGATTACCCATACGC 1380
Lewis AGCTCGCAACAGTAACCTTATCTGCCCACTGGAGGGTGCCAAGGATTACCCATACGC 1380
*****

R A V K D T L E N P Q S P Q P S P Y N S
DSS CGTGCTGTGAAAGATACCCCTTCCAAATCCACAGTCCCTCAGCCATCACCTTACAACCTC 1440
Lewis CGTGCTGTGAAAGATACCCCTTCCAAATCCACAGTCCCTCAGCCATCACCTTACAACCTC 1440
*****

F K P Q H K V T Q S F L P P G W E M R I
DSS CCCAAACCAACACAAAGTCAACAGAGCTTCTGCCACAGGCTGGGAGATGAGGATA 1500
Lewis CCCAAACCAACACAAAGTCAACAGAGCTTCTGCCACAGGCTGGGAGATGAGGATA 1500
*****

```

```

                                     ↓
DSS      A P N G R P F F F I D H M T K T T T W E D
Lewis    GCGCCAAACGGCCGACCCCTTTCATTGACCATAACACAAAGACGACAACTGGGAAGAT 1560
          GCGCCAAACGGCCGACCCCTTTCATTGACCATAACACAAAGACGACAACTGGGAAGAT 1560
          *****

DSS      F R L K F P V H M R S K A S L M P N D L
Lewis    CCACGGCTGAAATTTCCAGTACATATGCGGTCAAAGCATCTTTAAACCCCAATGACCTG 1620
          CCACGGCTGAAATTTCCAGTACATATGCGGTCAAAGCATCTTTAAACCCCAATGACCTG 1620
          *****

                                     ↓
DSS      G P L F P G W E E R I N L D G R T F Y I
Lewis    GGCCTCTTCTCTCTGCTGSGAAGAAAGSATCCACCTGGACGGCCGCACATTTTACATT 1680
          GGCCTCTTCTCTCTGCTGSGAAGAAAGSATCCACCTGGACGGCCGCACATTTTACATT 1680
          *****

                                     ↓
DSS      D H N S K I T Q W E D P R L Q N P A I T
Lewis    GACCATAATAGCAAATTTACTCAGTGGGAAGATCCAAGACTACAGAATCCAGCCATAACT 1740
          GACCATAATAGCAAATTTACTCAGTGGGAAGATCCAAGACTACAGAATCCAGCCATAACT 1740
          *****

                                     ↓
DSS      G P A V P Y S R E F K Q K Y D Y F R K
Lewis    GGTCCGGCTGTGCCGTACTCCAGAGAATTTAAGCAGAAATACGACTACTTTAGGAAGAAA 1800
          GGTCCGGCTGTGCCGTACTCCAGAGAATTTAAGCAGAAATACGACTACTTTAGGAAGAAA 1800
          *****

                                     ↓
DSS      L K K P A D I P N R F E M E L H R H N I
Lewis    TAAAGAAACCTGCTGATATCCAAACAGSTTTGAAATGAAACTCCACAGAAACAACATA 1860
          TAAAGAAACCTGCTGATATCCAAACAGSTTTGAAATGAAACTCCACAGAAACAACATA 1860
          *****

DSS      F E E S Y R R I M S V K R P D V L E A R
Lewis    TTTGAGAATCCATCCGAGGATCATGTCTGTAAGAGACCTGACCTCCTAAAGGCTAGA 1920
          TTTGAGAATCCATCCGAGGATCATGTCTGTAAGAGACCTGACCTCCTAAAGGCTAGA 1920
          *****

DSS      L W I E F E S E K G L D Y G G V A R E W
Lewis    CTGTGGATCGAGTTTGAATCAGAGAAAGGCTGGACTATGGGGCGTGGCCAGAGAGTGG 1980
          CTGTGGATCGAGTTTGAATCAGAGAAAGGCTGGACTATGGGGCGTGGCCAGAGAGTGG 1980
          *****

DSS      F F L L S K E M F N P Y Y G L F E Y S A
Lewis    TTCTTCTTACTGTCCAAAGAGATGTTCAATCCCTACTATGGCCTCTTCGAGTACTCTGCC 2040
          TTCTTCTTACTGTCCAAAGAGATGTTCAATCCCTACTATGGCCTCTTCGAGTACTCTGCC 2040
          *****

                                     ↓
DSS      T D H Y T L Q I N P N S G L C N E D H L
Lewis    ACGGACAACTACACACTTCAGATCAATCCCAACTCAGGCCGTGTGAACGAAGATCACTTA 2100
          ACGGACAACTACACACTTCAGATCAATCCCAACTCAGGCCGTGTGAACGAAGATCACTTA 2100
          *****

DSS      E Y F T F I G R V A G L A V F H G K L L
Lewis    TCCTATTTCACTTTCATTGGGAGAGTTGCTGGCCTGGCGTGTGTTTCATGGGAAACTCTTA 2160
          TCCTATTTCACTTTCATTGGGAGAGTTGCTGGCCTGGCGTGTGTTTCATGGGAAACTCTTA 2160
          *****

                                     ↓
DSS      D G F F I R P F Y K M M L G E Q I T L N
Lewis    GATGGATTCTTCATTGACCTTCTACAAGATGATGCTGGGSAAGCAAATAACTCTGAAT 2220
          GATGGATTCTTCATTGACCTTCTACAAGATGATGCTGGGSAAGCAAATAACTCTGAAT 2220
          *****

                                     ↓
DSS      D M E S V D S E Y Y N S L K W I L E N D
Lewis    GACATGGAGTCTGTGGACAGGAGTACTACAACCTTTTGAAGTGSATCTTGGAAAACGAC 2280
          GACATGGAGTCTGTGGACAGGAGTACTACAACCTTTTGAAGTGSATCTTGGAAAACGAC 2280
          *****

```

```

          ↓
DSS      F T E L D L M F C I D E E N F G Q T Y Q
Lewis    CCCACGGAACTTGACCTCATGTTCTGCATAGATGAAGAGAACTTTGGGCAGACTTACCAA 2340
          *****
          V D L K P N G S E I M V T N E N K R E Y
DSS      GTGGACCTGAAGCCCAATGGGTCAGAAATCATGGTAACCAATGAGAACAAGCGGAGATAT 2400
Lewis    GTGGACCTGAAGCCCAATGGGTCAGAAATCATGGTAACCAATGAGAACAAGCGGAGATAT 2400
          *****
          ↓
DSS      I D L V I Q W R F V N R V Q K Q M N A F
Lewis    ATTGACTTGGTCATCCAGTGGAGATTTGTGAACAGGGTCCAGAAGCAAATGAATGCCTTC 2460
          ATTGACTTGGTCATCCAGTGGAGATTTGTGAACAGGGTCCAGAAGCAAATGAATGCCTTC 2460
          *****
          ↓
DSS      L E G F T E L L P I D L I K I F D E N E
Lewis    TTGGAGGGATTACAGAACTTCTTCCAATCGACTTGATTAATAATTTTGGATGAAAATGAG 2520
          TTGGAGGGATTACAGAACTTCTTCCAATCGACTTGATTAATAATTTTGGATGAAAATGAG 2520
          *****
          ↓
DSS      L E L L M C G L G D V D V N D W R Q H S
Lewis    CTGGAGTTGCTGATGTGCGGCCCTTGGTGTGTGGACGTGAATGATTGGAGACAGCACTCC 2580
          CTGGAGTTGCTGATGTGCGGCCCTTGGTGTGTGGACGTGAATGATTGGAGACAGCACTCC 2580
          *****
          ↓
DSS      I Y K N G Y C P H H P V I Q W F W K A V
Lewis    ATTTACAAGAACGGCTACTGCCCAACCACCTGTCTCCAGTGGTCTGGAAGGCCGTG 2640
          ATTTACAAGAACGGCTACTGCCCAACCACCTGTCTCCAGTGGTCTGGAAGGCCGTG 2640
          *****
          L L M D A E K R I R L L Q E V T G T S R
DSS      CTGCTGATGGATGCTGAGAAGCGGATCCGGTTACTACAGTTTGTACAGGCACCTCCAGA 2700
Lewis    CTGCTGATGGATGCTGAGAAGCGGATCCGGTTACTACAGTTTGTACAGGCACCTCCAGA 2700
          *****
          ↓
DSS      V P M N G F A E L Y G S N G F Q L F T I
Lewis    GTACCCATGAATGGATTGCGCAACTTTATGGTCCCAATGGTCCCTCAGCTGTTACAATA 2760
          GTACCCATGAATGGATTGCGCAACTTTATGGTCCCAATGGTCCCTCAGCTGTTACAATA 2760
          *****
          ↓
DSS      E Q W G S P E K L F R A H T C F N R L D
Lewis    GAGCAATGGGGTAGTCTGAAAAACTGCCAGAGCTCATAATGCTTTAACCGCCTTGAT 2820
          GAGCAATGGGGTAGTCTGAAAAACTGCCAGAGCTCATAATGCTTTAACCGCCTTGAT 2820
          *****
          L P P Y E T F E D L R E K L L M A V E N
DSS      TTACCTCCGATGAAACCTTTGAAGATCTACGGGAGAAGCTTCTCATGGCCGTGGAGAAT 2880
Lewis    TTACCTCCGATGAAACCTTTGAAGATCTACGGGAGAAGCTTCTCATGGCCGTGGAGAAT 2880
          *****
          A Q G F E G V D -
DSS      GCTCAAGGCTTCGAAGGTGTGGATTAA 2907
Lewis    GCTCAAGGCTTCGAAGGTGTGGATTAA 2907
          *****

```

(3) Alpha-kinase 2 (Alpk2)

```

DSS      M T D Y S N A V W Q R N L Q S T D R V F
Lewis    ATGACGGATTACTCTAATGCGGTTTGGCAAAGGAACCTSCAGGGACTGACCGAGTTTTI 60
          ATGACGGATTACTCTAATGCGGTTTGGCAAAGGAACCTSCAGGGACTGACCGAGTTTTI 60
          *****
          L L E S D D E E M E F N E C G F G S C E
DSS      TTATTAGAAAGCGATGACGAGAGATGGAATTCATGASTGTGSTCCCGSTGGSTGTGAG 120
Lewis    TTATTAGAAAGCGATGACGAGAGATGGAATTCATGASTGTGSTCCCGSTGGSTGTGAG 120
          *****

```


H F F E E M G C G P Q V S G M N S M N
 DSS CATTCTTCAGTGAATGGTTGTGGGCTCAGGTGTGGGTGGCATGTGGTCTATGAAT 180
 Lewis CATTCTTCAGTGAATGGTTGTGGGCTCAGGTGTGGGTGGCATGTGGTCTATGAAT 180

 V A T G F C S Y R S Q P Q E V G D R S S
 DSS GTGGCCACTGGCTTCTGTAGCTATCGCTCACAAACCCAGGAAGTGGGGACAGGAGCASC 240
 Lewis GTGGCCACTGGCTTCTGTAGCTATCGCTCACAAACCCAGGAAGTGGGGACAGGAGCASC 240

 G T S R H E P L P L H S E M T L T L G P
 DSS GGAACCTCCAGGCATAGCCATTACCCCTGCAATTCAGAAATGACTCTTACCCCTGGGACCT 300
 Lewis GGAACCTCCAGGCATAGCCATTACCCCTGCAATTCAGAAATGACTCTTACCCCTGGGACCT 300

 H Q D G I A T M T E P G R A P L P T A S
 DSS CACCAGGACGGAAATAGCTACGATGACAGAACCAGGGAGGGCTCCACTCCCACTGCTTCC 360
 Lewis CACCAGGACGGAAATAGCTACGATGACAGAACCAGGGAGGGCTCCACTCCCACTGCTTCC 360

 E A V E N D C S G I R G E T H D E P E A
 DSS GAGGCTGTGAAAATGATTGTTCCAGGAATTCGGGGAGAACCAGAGACAAGCCTGAAGCA 420
 Lewis GAGGCTGTGAAAATGATTGTTCCAGGAATTCGGGGAGAACCAGAGACAAGCCTGAAGCA 420

 S E E F S S D N L Q T N D E A E T E A S
 DSS GGAGAGGAATTCCTCAUTGACAATCTACAGACCTGGATAAAGCAGAGACAGAGGCATCG 480
 Lewis GGAGAGGAATTCCTCAUTGACAATCTACAGACCTGGATAAAGCAGAGACAGAGGCATCG 480

 A K P L S G S D K S E A K Q S L K S L
 DSS GCGAAGCCCTTGTCTGGGGCTCAGATAAGTCAGAGCCGAAGCAGGGTTTGAAGAAGCCTG 540
 Lewis GCGAAGCCCTTGTCTGGGGCTCAGATAAGTCAGAGCCGAAGCAGGGTTTGAAGAAGCCTG 540

 A G E R T E E K Y P G S R E T A L R P T
 DSS GCTGGGSAACGCACGGAAAGAAAATACCCAGSCAGGAGAGACAGCCCTGAGACCCACC 600
 Lewis GCTGGGSAACGCACGGAAAGAAAATACCCAGSCAGGAGAGACAGCCCTGAGACCCACC 600

 R A R W F G M K A N V E K Q L L E D G P
 DSS AGGGGAGSTGGCCGGAAATGAAGSCAAATGTCAAGAAAGCAGTCTGTGAAGAGAGGTCC 660
 Lewis AGGGGAGSTGGCCGGAAATGAAGSCAAATGTCAAGAAAGCAGTCTGTGAAGAGAGGTCC 660

 F K G T L D P L P K E P T K Q F L T R S
 DSS CCAAAAGGCACTCTTGAATCCCTTCCCAAGGAACCAACAAAGCAGCCCTTAACAGGAGC 720
 Lewis CCAAAAGGCACTCTTGAATCCCTTCCCAAGGAACCAACAAAGCAGCCCTTAACAGGAGC 720

 Y G Q E P A H T E T G A F G M N S H F R
 DSS TATGGACAAGAACTGCACACACAGAGACGGGAGCTCCTGGCTGGAAATCCCACTTCCBA 780
 Lewis TATGGACAAGAACTGCACACACAGAGACGGGAGCTCCTGGCTGGAAATCCCACTTCCBA 780

 E V C I P L P A E Q D S K T P R P F A D
 DSS GAAGTGTGATTCCTCTTCTGSCAGAGCAAGACTCCAAAACCCCTCGAACCCAGCAGAC 840
 Lewis GAAGTGTGATTCCTCTTCTGSCAGAGCAAGACTCCAAAACCCCTCGAACCCAGCAGAC 840

 F L E K E S D S E F E G G B A L L E T L
 DSS CCACITTCAAAGGAAGGGATAGCAGCTTCGAGSSAGGAGGACACTGCTTAGTACGTTA 900
 Lewis CCACITTCAAAGGAAGGGATAGCAGCTTCGAGSSAGGAGGACACTGCTTAGTACGTTA 900

 S E A S Q I P D Q T - G R L Q E T I G E
 DSS TCTGAAGCAAGCCAGATCCAGACCAGACCG---GCCGTCTTCAGGAAACATTGGGGAG 957
 Lewis TCTGAAGCAAGCCAGATCCAGACCAGACCGGCCGCCGTCTTCAGGAAACATTGGGGAG 960

A

N S Y L D Q T F A F S V F A E E E S T F
 DSS AACAGTTATCTCGACAGAGCCCGCCCTTTTCCTGTGCCCGCTGAGGAGSAGTCTACGTT 1017
 Lewis AACAGTTATCTCGACAGAGCCCGCCCTTTTCCTGTGCCCGCTGAGGAGSAGTCTACGTT 1020

 T S V Q T H F V E N L T E I D R G N S
 DSS ACCGGGTCCAAACACATTCGCTCCAACTTGACTGAGATCGACAGSSGAAACTCATCA 1077
 Lewis ACCGGGTCCAAACACATTCGCTCCAACTTGACTGAGATCGACAGSSGAAACTCATCA 1080

 L A Q Y L G L E S C F Q S P Q Q E S R Q
 DSS TTGGCCAGTACCTGGGGCTGGAGAGCTGTCCCAAAGGCCCCAGCAGSAAAGCAGACAA 1137
 Lewis TTGGCCAGTACCTGGGGCTGGAGAGCTGTCCCAAAGGCCCCAGCAGSAAAGCAGACAA 1140

N R E E D D R P G A L W A E E A Q E L R P
 DSS AACAGAGAAAGGTGACAGGCTGGGCTCTCTGGGACAGAAATCAGCACAAAGAACTGAGACCC 1197
 Lewis AACAGAGAAAGGTGACAGGCTGGGCTCTCTGGGACAGAAATCAGCACAAAGAACTGAGACCC 1200

 L E D N D E E V S Q A P A S V A L F Q G
 DSS CTAGAGACAAATGATGAAGAAGTGTCCAGGCTCCAGCCTCAGTGGCTCTCCCTCAGGAT 1257
 Lewis CTAGAGACAAATGATGAAGAAGTGTCCAGGCTCCAGCCTCAGTGGCTCTCCCTCAGGAT 1260

 D G T H F R E S E A L S D A F S Q P T A
 DSS GATGGTACCCACTTCAGGAGTCAGAGGCTCTCTCTGATGCTTTCCTCCAGCCTACTGCT 1317
 Lewis GATGGTACCCACTTCAGGAGTCAGAGGCTCTCTCTGATGCTTTCCTCCAGCCTACTGCT 1320

 P S L P L E H V D S G S R S R E A A C V
 DSS CCCTCCCTCCCTTGGAAAAATGTGGACATGGGTCAAGGGGACAGAAAGCTGCCTGTGTG 1377
 Lewis CCCTCCCTCCCTTGGAAAAATGTGGACATGGGTCAAGGGGACAGAAAGCTGCCTGTGTG 1380

 M G C F E A G D Q E T C Y A T M D L P V
 DSS ATGGGGTGTITTTGAAGCTGGTATCAAGAAACATGTTATGCTACCATGGATCTCCCTGTT 1437
 Lewis ATGGGGTGTITTTGAAGCTGGTATCAAGAAACATGTTATGCTACCATGGATCTCCCTGTT 1440

 S A F V D R Y L P E E I C P M D L E L T
 DSS GGAGCACCAATGATAAATATTTGCTGAAAGAAATTTGCCCATGGACTTGGAGCTGACA 1497
 Lewis GGAGCACCAATGATAAATATTTGCTGAAAGAAATTTGCCCATGGACTTGGAGCTGACA 1500

 E G Q R E V C D L C S P D K I L A V L Q
 DSS GAAGGTCAAAGAGAGATATGATTTATGTTCTCCCTGACAAGATATGGCTGTCTACAA 1557
 Lewis GAAGGTCAAAGAGAGATATGATTTATGTTCTCCCTGACAAGATATGGCTGTCTACAA 1560

 A Q G S E S P Q A T Y K R S E D G K S A
 DSS GCACAAAGTTCTGAGTCTCCACAGGCCAAGTACAAGCACAGCGAGGATGGAAAGTCAGCC 1617
 Lewis GCACAAAGTTCTGAGTCTCCACAGGCCAAGTACAAGCACAGCGAGGATGGAAAGTCAGCC 1620

 E G A L F H S T L A W D T S P E A R E D
 DSS GAGGGCTCTTTTTCATAGTACTTGGCTGGGACACATCACGGAGGCGAGAGAGAT 1677
 Lewis GAGGGCTCTTTTTCATAGTACTTGGCTGGGACACATCACGGAGGCGAGAGAGAT 1680

 A T G G T A A D T G N S F S I F S S T L
 DSS GCTACGGGAGGACAGCAGCTGATACGGGAAATCTCCCTCCATCTTCTCTCTACACTA 1737
 Lewis GCTACGGGAGGACAGCAGCTGATACGGGAAATCTCCCTCCATCTTCTCTCTACACTA 1740

 L Y N Y G S F G E I Q A F C S E N I S F
 DSS CTCTACAATGTAGSAGGTTTGGGAGATACAGGCCCTTGTCTGAAAATATCTCCTTT 1797
 Lewis CTCTACAATGTAGSAGGTTTGGGAGATACAGGCCCTTGTCTGAAAATATCTCCTTT 1800

 V N S N E G G Y G S S N L S V F T A I D
 DSS GTAAACAGTAACGAAGGAGGCTATGGAAGCTCCAATCTCAGCGTCCGATTGCCATAGAC 1857
 Lewis GTAAACAGTAACGAAGGAGGCTATGGAAGCTCCAATCTCAGCGTCCGATTGCCATAGAC 1860

 T L A S Y S I R E C S E E Q S T E P A
 DSS ACGCTTGCAGCTAAGGTTCCATCAGGAGTCTCAGAAGSACSTCAACAGAACCGCT 1917
 Lewis ACGCTTGCAGCTAAGGTTCCATCAGGAGTCTCAGAAGSACSTCAACAGAACCGCT 1920

 A N V D C H L V T B E T E G I L T D A T
 DSS GCTAATGTTGACTGTCACTGGTGACCAGAGACAGAGGSCATACTAATGATGCCACC 1977
 Lewis GCTAATGTTGACTGTCACTGGTGACCAGAGACAGAGGSCATACTAATGATGCCACC 1980

 Q
 E V R K I K C R T V S V F H I N D F V D
 DSS GAAGTCCACAAAATCAAATGCCGACCGTTTCTGTTCCCCACATCAACGACTTGTGTTGAT 2037
 Lewis GAAGTCCACAAAATCAAATGCCGACCGTTTCTGTTCCCCACATCAACGACTTGTGTTGAT 2040

 S A D Q V S C E A Q D E E M S F S L F D
 DSS SGTGCTSACCAGTCTCGTGTGAGGACAAAGATGAAGAGATTCCCATGCTCCCGGAT 2097
 Lewis SGTGCTSACCAGTCTCGTGTGAGGACAAAGATGAAGAGATTCCCATGCTCCCGGAT 2100

 D P L S S F T G E A T R E T L V P A P S
 DSS GACCCGTTAAGTAGCTTCACAGGTGAAGCAACCAAGGAGACTCTTGTCCAGCSCCCAGT 2157
 Lewis GACCCGTTAAGTAGCTTCACAGGTGAAGCAACCAAGGAGACTCTTGTCCAGCSCCCAGT 2160

D A S T H S H F L L E E S Q G L Y L R P
 DSS 2217 GATGACAGAACTCATGSCACTTCTTACTGCCCGAGGGACAGGTTTGTACCTCAGGCCT
 Lewis 2220 GATGACAGAACTCATGSCACTTCTTACTGCCCGAGGGACAGGTTTGTACCTCAGGCCT
 L Q I D T Q P G Y E S Q T V E G A H S G
 DSS 2277 CTTGAGATTGATACCCAGCCTGGGTATGAGABCCAGACTGTGGAGGGAGCCACAGCGSA
 Lewis 2280 CTTGAGATTGATACCCAGCCTGGGTATGAGABCCAGACTGTGGAGGGAGCCACAGCGSA
 G L E E D F Q E E G N G T E Q C I R P Q
 DSS 2337 GGCCTTGAAGAGGACTTCCAAGAAAAGGSAATGSGACAAAGCAATGCATCCGSCCAGAG
 Lewis 2340 GGCCTTGAAGAGGACTTCCAAGAAAAGGSAATGSGACAAAGCAATGCATCCGSCCAGAG
 E T S H Q Q S L S A N D F Q S S L P S I
 DSS 2397 AGCACATCCCATCAGGTTCTCTTTCTGCAAAATGATTTCCAAGAAAATTTGCCCTCCATA
 Lewis 2400 AGCACATCCCATCAGGTTCTCTTTCTGCAAAATGATTTCCAAGAAAATTTGCCCTCCATA
 F A N Q Q E V N V E F F E H S P A D S G
 DSS 2457 CTTGCCATGCAACAGGAGSTCAACGTGGAACCTTCGAGCACTCCCCAGCAGATTCCGSS
 Lewis 2460 CTTGCCATGCAACAGGAGSTCAACGTGGAACCTTCGAGCACTCCCCAGCAGATTCCGSS
 E E T E C S S D Q R T S V S V L A E E T
 DSS 2517 GAAGAACTGAGTGTAGCTCAGACCAGAGGACCASTTTTCTGTGTGGCTGAGAAGAAC
 Lewis 2520 GAAGAACTGAGTGTAGCTCAGACCAGAGGACCASTTTTCTGTGTGGCTGAGAAGAAC
 H G E S S E L V S S V P A L P D I L L G
 DSS 2577 ATGGAGAAAGGCASTCTTTGGTAAGCASTGTCCAGCTCTCCCTGACATCCTCCTTGA
 Lewis 2580 ATGGAGAAAGGCASTCTTTGGTAAGCASTGTCCAGCTCTCCCTGACATCCTCCTTGA
 E K D S I G L G S W S V G S K V K I I T
 DSS 2637 GAGAAAGATGGCATTGGACTAGGAAGTTGGTCTGTGGGCAAGAAAGTGAAGATCATAACT
 Lewis 2640 GAGAAAGATGGCATTGGACTAGGAAGTTGGTCTGTGGGCAAGAAAGTGAAGATCATAACT
 L E A P V F E I W P P E L V T H S E Y K
 DSS 2697 CTAGAAGCTCCTGTCTTGAATCTGGCCCCAGAACTAGTGAAGCAATCTGGGTACAAG
 Lewis 2700 CTAGAAGCTCCTGTCTTGAATCTGGCCCCAGAACTAGTGAAGCAATCTGGGTACAAG
 E A E V G L T A P G R S W A L S D I L R
 DSS 2757 GAGGCAGAAATGGTCTCAGGGCACTGSAAGGAGCTGGGCTCTGTCTGACATCCTCAGA
 Lewis 2760 GAGGCAGAAATGGTCTCAGGGCACTGSAAGGAGCTGGGCTCTGTCTGACATCCTCAGA
 A G T R P E P G A L G V T T W V P S F K
 DSS 2817 GCAGSCACAGGCCCTGAGCCAGGAGCCTTGGAGTAACAACATGSGTTCCAGCCCCAAA
 Lewis 2820 GCAGSCACAGGCCCTGAGCCAGGAGCCTTGGAGTAACAACATGSGTTCCAGCCCCAAA
 K
 A D A I N A L G E N N R D I C E D A A P D
 DSS 2877 GCGATGCCATTATGGCCCTTGGAAATAACAGGACATCTGCGAAGATGCTGCACCAGAC
 Lewis 2880 GCGATGCCATTATGGCCCTTGGAAATAACAGGACATCTGCGAAGATGCTGCACCAGAC
 R Q A Y C N S Q T S Q C L G Q F R L L E
 DSS 2937 AGACAGCSTACTGCAACAGTCAGACTTCCAGTGTCTGGGCCAACCCGACTCCTGGAG
 Lewis 2940 AGACAGCSTACTGCAACAGTCAGACTTCCAGTGTCTGGGCCAACCCGACTCCTGGAG
 E S V D P V E E E E L H V T D S P S E T
 DSS 2997 TCATCTGTTGACCTGTGAGGAGAGGAGTTAAATGTCACGGACTCTCCATCAGAGACT
 Lewis 3000 TCATCTGTTGACCTGTGAGGAGAGGAGTTAAATGTCACGGACTCTCCATCAGAGACT
 E R T G E V E M A E T T D E E Q G G R Q
 DSS 3057 TCCAGAACTGGAGAGGTGSAATGGCTGAGACAACGGATGAGGAACAGGGGGAGAGGCAG
 Lewis 3060 TCCAGAACTGGAGAGGTGSAATGGCTGAGACAACGGATGAGGAACAGGGGGAGAGGCAG
 H K L P H P T V A N Q S V N F F R I L E
 DSS 3117 CACAAGCTGCCGCAOCCACCCTTGTAAOCCAGTCTGTGAACTTCCCTAGGATCCTGGAA
 Lewis 3120 CACAAGCTGCCGCAOCCACCCTTGTAAOCCAGTCTGTGAACTTCCCTAGGATCCTGGAA
 S S V D P I D D R G G K F E P S D S M I
 DSS 3177 TCTCTGTGGACCCATTGATGACAGGGSTGGGAAGCCAGAGCCTCTGACTCAACATA
 Lewis 3180 TCTCTGTGGACCCATTGATGACAGGGSTGGGAAGCCAGAGCCTCTGACTCAACATA

```

E A S E S T T G N M C Q R V D I O T A H
DSS      GAAGCAAGTGAATCCACCCTGGAAACATGTGTGAGAGGGTAGACATCCAAACTGCTCAC 3237
Lewis    GAAGCAAGTGAATCCACCCTGGAAACATGTGTGAGAGGGTAGACATCCAAACTGCTCAC 3240
*****
L R V F H P Q D N G E I I P N E H T T N
DSS      CTACGGTCCCACATCCCAGGACAAATGGGAAATCATTCCAAACGAAAACACAACCAAC 3297
Lewis    CTACGGTCCCACATCCCAGGACAAATGGGAAATCATTCCAAACGAAAACACAACCAAC 3300
*****
Q T H V D R E R A D A E A S Q H N A A K
DSS      CAAACTCATGTAGACAGAGAGCGAGCAGATGCCAAAGCCAGTCAGCATAATGACGCAAAA 3357
Lewis    CAAACTCATGTAGACAGAGAGCGAGCAGATGCCAAAGCCAGTCAGCATAATGACGCAAAA 3360
*****
D A I W Q G G C F G G E E R Q G I P S V C
DSS      GAGCCTATTGGCAAGGCCAGTGCOCOCGTGAASAGAGACAAGGGATTCCAAAGTGTGTGC 3417
Lewis    GAGCCTATTGGCAAGGCCAGTGCOCOCGTGAASAGAGACAAGGGATTCCAAAGTGTGTGC 3420
*****
T V S P T Q D G G D R S L G E A G Q R G
DSS      ACCGTGAGCCCAACACAAGATGGTGGTGCAGAAAGCCTAGGAGAAGCTGGGCAAGGGGA 3477
Lewis    ACCGTGAGCCCAACACAAGATGGTGGTGCAGAAAGCCTAGGAGAAGCTGGGCAAGGGGA 3480
*****
K D E T E V T S P M S P L E N C P A G M
DSS      AAGGACGAGACTGAGGTCACTTCCOCCATGTCTCOOCTTTCTAACTGTCTGCAGGAAAG 3537
Lewis    AAGGACGAGACTGAGGTCACTTCCOCCATGTCTCOOCTTTCTAACTGTCTGCAGGAAAG 3540
*****
T Y T S V T A E T S N S T G H I Y G S S
DSS      ACGTACACATCTGTACGGCTGASACCAAGTAACTCCACAGGCCACATTTATGGCGATCT 3597
Lewis    ACGTACACATCTGTACGGCTGASACCAAGTAACTCCACAGGCCACATTTATGGCGATCT 3600
*****
E P R T H Q R V I P V K R E K G T I E N
DSS      GAGCCGAGAACCATCAACGTGTAATTCCTGTGAAGAGGGAAAAGGAACCATGAGAAAC 3657
Lewis    GAGCCGAGAACCATCAACGTGTAATTCCTGTGAAGAGGGAAAAGGAACCATGAGAAAC 3660
*****
K C S K H V P S S N D L T D T L C T S S
DSS      AAGTGTGGAAACATGTGCOCCTCTCAAATGATCTCACGGACACACTGTGCACCTTCATCT 3717
Lewis    AAGTGTGGAAACATGTGCOCCTCTCAAATGATCTCACGGACACACTGTGCACCTTCATCT 3720
*****
P K G N V T R S P T S P R A S E L K E E
DSS      CCCAAGGAAATGTCAACGCTGOCACADGAGCCCTGGGCGAGAGGAAGTGAATCAGAG 3777
Lewis    CCCAAGGAAATGTCAACGCTGOCACADGAGCCCTGGGCGAGAGGAAGTGAATCAGAG 3780
*****
E L Q I A E T K P L N S S D P H T M T L
DSS      GAGCTTCAAATGGGAAACCAAACCCCTAAACTCATCTGACCCCGAACAAATGACCTTG 3837
Lewis    GAGCTTCAAATGGGAAACCAAACCCCTAAACTCATCTGACCCCGAACAAATGACCTTG 3840
*****
A F I S G E H E S E K D F E S L L L K D
DSS      GCTTTCATTTCAGSAGAACATGAGTCAGAGAAAGCCCTGAAAGCTTGTACTTAAGGAC 3897
Lewis    GCTTTCATTTCAGSAGAACATGAGTCAGAGAAAGCCCTGAAAGCTTGTACTTAAGGAC 3900
*****
L C Q K G S T L E S G K K E R E E Q Q R
DSS      CTGTGTCAAAGGGCTCTACCCCTGGAGAGGGGAAAAGTCCAGAGGSAACAGCAGAGG 3957
Lewis    CTGTGTCAAAGGGCTCTACCCCTGGAGAGGGGAAAAGTCCAGAGGSAACAGCAGAGG 3960
*****
F V V A N I S K A P G A Q S A I A G S S
DSS      CCTGTGTGGCCAAATCAGCAAGGCCACCGGGGCCAATCAGCAATAGCTGGGTGAGAG 4017
Lewis    CCTGTGTGGCCAAATCAGCAAGGCCACCGGGGCCAATCAGCAATAGCTGGGTGAGAG 4020
*****
E G K E Q E A S S S G H L A A G I K E K
DSS      GAGGCCAAAAACAAGAGGCTTCAGGGAGTGGACACTTGGCTGCAGGGATAAAGAGAAA 4077
Lewis    GAGGCCAAAAACAAGAGGCTTCAGGGAGTGGACACTTGGCTGCAGGGATAAAGAGAAA 4080
*****
I L S R V A A L R L R L E E K E R V K N
DSS      ATTCTATCCAGGGTGGCAGCCCTGAGGCTGAGGCTAGAGGAAAAGGAADGTGTGAGAAAC 4137
Lewis    ATTCTATCCAGGGTGGCAGCCCTGAGGCTGAGGCTAGAGGAAAAGGAACGTGTGAGAAAC 4140
*****
S T L R E A P K F E R E L S R T D E E R
DSS      TCCACTCTGAGGAAGGCACCTAAGTTTGAACGGTCTTGTCCCGACTGATGAGAAAAGA 4197
Lewis    TCCACTCTGAGGAAGGCACCTAAGTTTGAACGGTCTTGTCCCGACTGATGAGAAAAGA 4200
*****
D P E R A P C K A E S E A P V L L K E I
DSS      GACCCGAGAAGGGCCCTTGCAAAGCTGAAGGGAAAGCTCCAGTATTGCTGAAGAGATC 4257
Lewis    GACCCGAGAAGGGCCCTTGCAAAGCTGAAGGGAAAGCTCCAGTATTGCTGAAGAGATC 4260
*****

```



```

      Q A E P A F E H E G N I M L T C Q F E E
DSS      CAGGCTGAGCCGGCTCCCGAGCACTCTGGAATATAATGCTGACCTGTCACTTTTCAGAA 4317
Lewis    CAGGCTGAGCCGGCTCCCGAGCACTCTGGAATATAATGCTGACCTGTCACTTTTCAGAA 4320
      *
      I H E D S T V C W Y K D S E S I A Q L K
DSS      ATCCATGAAGACTCTACCGTTTGTCTGGACAAAAGATTGAAAGTCGATAGGCCAGCTCAAG 4377
Lewis    ATCCATGAAGACTCTACCGTTTGTCTGGACAAAAGATTGAAAGTCGATAGGCCAGCTCAAG 4380
      *
      ↓
      K S A S D S S S V S L A I A Q A G Q E D
DSS      AAAAGCCGAGGGGACAGCTCCAGTGTTCCTTGGCCATCGCCCAAGCTGGTCAGAAGGAC 4437
Lewis    AAAAGCCGAGGGGACAGCTCCAGTGTTCCTTGGCCATCGCCCAAGCTGGTCAGAAGGAC 4440
      *
      Q G L Y Y C C V K N S Y S K V T A E F N
DSS      CAGGGCTCTATTACTGCTGCTCAAGAACAGCTACGGAAAAGTCACTGCTGAGTTTAA 4497
Lewis    CAGGGCTCTATTACTGCTGCTCAAGAACAGCTACGGAAAAGTCACTGCTGAGTTTAA 4500
      *
      ↓
      L T A E V L E Q L S S H T E Y R G C E E
DSS      CTCACAGCCGAGGTTCTCAAACAGCTCTCCAGTCACACAGAATACAGAGGATGTGAAGAG 4557
Lewis    CTCACAGCCGAGGTTCTCAAACAGCTCTCCAGTCACACAGAATACAGAGGATGTGAAGAG 4560
      *
      I E F E Q L I F K E D V F N D S Y F G D
DSS      ATTGAATCAGCCAGCTCATCTCAAGAAAGATGTTTTCAATGACAGCTACTTGGGGAC 4617
Lewis    ATTGAATCAGCCAGCTCATCTCAAGAAAGATGTTTTCAATGACAGCTACTTGGGGAC 4620
      *
      H L R G Q I F T E E L H F S E G V H R K
DSS      CACTACGGGGCAGATCTCACGAGGAGCTTCACTTGGCCGAAGGGGTGCAACCGCAA 4677
Lewis    CACTACGGGGCAGATCTCACGAGGAGCTTCACTTGGCCGAAGGGGTGCAACCGCAA 4680
      *
      A F R S T V H Q S L M P V F Q F G H A C
DSS      GCITTCGGCAGCACAGTGATGCAAGGCTCATGCCCCCTTCCAGCCCGCCACGSCATG 4737
Lewis    GCITTCGGCAGCACAGTGATGCAAGGCTCATGCCCCCTTCCAGCCCGCCACGSCATG 4740
      *
      V L K V H N A V A H G T R N N D E L V Q
DSS      GTACTCAAGGTGCACAACGCTGTGCCCCATGGGACCAGAAACAATGATGAACCTGTGCG 4797
Lewis    GTACTCAAGGTGCACAACGCTGTGCCCCATGGGACCAGAAACAATGATGAACCTGTGCG 4800
      *
      ↓
      R H Y E L A A Q E C Y V Q H T A R Y Y A
DSS      AGGAATACAGCTCGCTGCCCAGGAATGCTACGTTTCAAGATACTGCCAGATACTATGCC 4857
Lewis    AGGAATACAGCTCGCTGCCCAGGAATGCTACGTTTCAAGATACTGCCAGATACTATGCC 4860
      *
      ↓
      K I Y A A E A Q P L E G F S E V F E I I
DSS      AAGATCTATGCCGCTGAAGCACAGCTCTGGAAGGCTTGGAGAGGTGCCGGAGATCATT 4917
Lewis    AAGATCTATGCCGCTGAAGCACAGCTCTGGAAGGCTTGGAGAGGTGCCGGAGATCATT 4920
      *
      P I F L I R R P E N N I P Y A T V E E S
DSS      CCTATTTTCCTTATCCGTCGGCCCGAGAACAACATCCCTACGCCACAGTGAAGAGGAG 4977
Lewis    CCTATTTTCCTTATCCGTCGGCCCGAGAACAACATCCCTACGCCACAGTGAAGAGGAG 4980
      *
      L I G E F V E Y S I R D G E S I N F L R
DSS      CTGATTGGAGAATTCGTGAATATTCATCCGAGACGGGAAGGAATCAACTTCTCAGA 5037
Lewis    CTGATTGGAGAATTCGTGAATATTCATCCGAGACGGGAAGGAATCAACTTCTCAGA 5040
      *
      R D S E A G Q K C C T F Q H M V Y Q K T
DSS      CGAGATCISAGGCTGGTCAGAAGTGTTCACCTCCAGCACTGGGTATACCAGAAAACA 5097
Lewis    CGAGATCISAGGCTGGTCAGAAGTGTTCACCTCCAGCACTGGGTATACCAGAAAACA 5100
      *
      S G C L L V T D M Q G E Q D R P T A L Q
DSS      AGTGGCTGTCTCCTGGTCAAGSACATGCAAGGTGAGCAAGSACAGCCGACAGCTTGCAG 5157
Lewis    AGTGGCTGTCTCCTGGTCAAGSACATGCAAGGTGAGCAAGSACAGCCGACAGCTTGCAG 5160
      *
      G L W E R S V F E V S F L R K E H R S S
DSS      GGGCTGTGAGCCGGGTGTTCTAAGGTTTCTTCTTGGAGGAAGGAATGAGAGGAAT 5217
Lewis    GGGCTGTGAGCCGGGTGTTCTAAGGTTTCTTCTTGGAGGAAGGAATGAGAGGAAT 5220
      *

```

```

E E E E R S V G Y Q Q A S P P Q D V T H
DSS GAAGAGGAGGAGAGGTCCTTTGGGTACCAACAAGCATCTCCTCCCAAGAGGTTACTCAC 5277
Lewis GAAGAGGAGGAGAGGTCCTTTGGGTACCAACAAGCATCTCCTCCCAAGAGGTTACTCAC 5280
*****
I
V P L R T M L G K T D V R K -
DSS GTCCCATTCGGACGATGCTTGGAAAGACTGACCTTAGSAAATGA 5322
Lewis GTCCCATTCGGACGATGCTTGGAAAGACTGACCTTAGSAAATGA 5325
*****

```

(4) Zinc finger protein 532 (Zfp532)

```

M K A K R E C N R V T V T E R W S E K V F
Lewis ATGAAGGCCAAAAGGAATGCAACCGTGTGACAGTAAGTCTGACACTGGTCCAAAAGTCTTT 60
DSS ATGAAGGCCAAAAGGAATGCAACCGTGTGACAGTAAGTCTGACACTGGTCCAAAAGTCTTT 60
*****
S K G Q G S S Q E H L I K F M T M G D M K
Lewis TCAAAAGGTCAAGSTTCACANGAACATCTATCAAAATTCATGACCATGGGGGATATGAAG 120
DSS TCAAAAGGTCAAGSTTCACANGAACATCTATCAAAATTCATGACCATGGGGGATATGAAG 120
*****
T P D F D D L L A A F D I P D M V D P K
Lewis ACCCCAGACTTGGATGACTCTTGGCAGGTTTGCATACAGATATGTTGATCTCTAAA 180
DSS ACCCCAGACTTGGATGACTCTTGGCAGGTTTGCATACAGATATGTTGATCTCTAAA 180
*****
A A I E S G H D D H E S H I E Q N A H V
Lewis SCAGCCATTGAGTCTGGGCACGATGACCAAGAAAGCCACATTAAGCAGAATGCTCAGGTG 240
DSS SCAGCCATTGAGTCTGGGCACGATGACCAAGAAAGCCACATTAAGCAGAATGCTCAGGTG 240
*****
D D D S H T P S S S D V G V S V I V K N
Lewis GATGACACTCTCACACACCATCATCTCAGATGTTGGGTTAGTGTGATTGTSAGAAT 300
DSS GATGACACTCTCACACACCATCATCTCAGATGTTGGGTTAGTGTGATTGTSAGAAT 300
*****
V R N I D S S E G V E K D G H N P T G N
Lewis GTTCCCAACATCGACTCCTCTGAGGGAGTGGAAAAGATGGCCACAATCCACAGGCAAT 360
DSS GTTCCCAACATCGACTCCTCTGAGGGAGTGGAAAAGATGGCCACAATCCACAGGCAAT 360
*****
G L H N G F L T A S S L D S Y G R D G A
Lewis GSCITGCATAACGGSTTCTCACGSCCTCCTCTCTGACAGCTATGTTAAGGATGGAGCC 420
DSS GSCITGCATAACGGSTTCTCACGSCCTCCTCTCTGACAGCTATGTTAAGGATGGAGCC 420
*****
K S I K G D T P A S E V T L E D P A F S
Lewis AAGTCTTAAAAGGAGACACACTGCTCAGAGGTGACTCTTAAAGACCCAGCCTTCAGC 480
DSS AAGTCTTAAAAGGAGACACACTGCTCAGAGGTGACTCTTAAAGACCCAGCCTTCAGC 480
*****
Q F S P I S S A K E F E D D E H I E V D
Lewis CAGTTCAGCCCATCTCCAGGCGGAGGAGTTTGGAGATGATGAGAAAATAGAGSTGGAT 540
DSS CAGTTCAGCCCATCTCCAGGCGGAGGAGTTTGGAGATGATGAGAAAATAGAGSTGGAT 540
*****
D P F D K E E T R A N F R S N V L T G S
Lewis GACCCACCTGATAAGGAGSAGACGAGGGCCAAATTCAGATCAAATGTGCTGACGGGCTCC 600
DSS GACCCACCTGATAAGGAGSAGACGAGGGCCAAATTCAGATCAAATGTGCTGACGGGCTCC 600
*****
A P Q Q D F D K L E A L G G E N T S E T
Lewis SCTCCCCAGSAGACTTCGACAAACTGAAGGCACCTTGGAGGGGAAAACACAGCAAGACT 660
DSS SCTCCCCAGSAGACTTCGACAAACTGAAGGCACCTTGGAGGGGAAAACACAGCAAGACT 660
*****
G V S A S G H A D K N K V K R E A E S N
Lewis GGAGTCTCTGCATCAGSTCATGAGATAAAAACAAAGTCAAGAGAGAGCAGAAAAGCAAT 720
DSS GGAGTCTCTGCATCAGSTCATGAGATAAAAACAAAGTCAAGAGAGAGCAGAAAAGCAAT 720
*****
S I T L S V Y E P F K V R K A E D K L E
Lewis TCCATAACCTGAGTGTCTATGAGCCGTTTAAAGSTCAGAAAAGCAGAGSATAAGTGAAG 780
DSS TCCATAACCTGAGTGTCTATGAGCCGTTTAAAGSTCAGAAAAGCAGAGSATAAGTGAAG 780
*****
E N S E K M L E S R V L D G R P S S E R
Lewis GAGAACTCTGAGAAGATGCTTGAAGCAGGCTCTGGATGGGAAGCCGAGCTCCGAGAAG 840
DSS GAGAACTCTGAGAAGATGCTTGAAGCAGGCTCTGGATGGGAAGCCGAGCTCCGAGAAG 840
*****

```

```

Lewis      M D F A I A A A T S S E K T K S S S K L S
DSS        AACGACCCCTGCCATTGCTGCTGCCACATCTTCCAAAACAAAGTGTCTCTCCAAAGCTCTG 900
          AACGACCCCTGCCATTGCTGCTGCCACATCTTCCAAAACAAAGTGTCTCTCCAAAGCTCTG 900
          *****
Lewis      S C I A A A I A A L S A K K A A S D S C K
DSS        TCGTGCATAGCCGCCATTGCGGCTCTGAGTGGGAAAAGGCTGCATCTGACTCTGCAAAA 960
          TCGTGCATAGCCGCCATTGCGGCTCTGAGTGGGAAAAGGCTGCATCTGACTCTGCAAAA 960
          *****
Lewis      E P V A H S R E A S P L P K E V H D S P
DSS        GAGCCTGTGGCCAAATCCAGGGAAGCATCTCCACTACCAAAGAAAGTGAATGACAGTCCC 1020
          GAGCCTGTGGCCAAATCCAGGGAAGCATCTCCACTACCAAAGAAAGTGAATGACAGTCCC 1020
          *****
Lewis      K A A D K S P E S Q N L I D G T E K A S
DSS        AAAGCCGCTGACAAGTCTCCGGAATCCAGAAATCTCATCGATGGCACCAGAAAGGCTCT 1080
          AAAGCCGCTGACAAGTCTCCGGAATCCAGAAATCTCATCGATGGCACCAGAAAGGCTCT 1080
          *****
Lewis      L K F S D S P R S V S S E N S S K G S P
DSS        CTGAAGCCGTCAGATAGTCCAGSAGCGTATCCAGTGGAGAACAGCAGCAAAGGSTCCCCA 1140
          CTGAAGCCGTCAGATAGTCCAGSAGCGTATCCAGTGGAGAACAGCAGCAAAGGSTCCCCA 1140
          *****
Lewis      S S F V G S T F A I P K V R I X T I K T
DSS        TCCTCTCTGTGGGTTCTACCCAGCTATCCCAAGTCCGCATCAAGACCATCAAGACA 1200
          TCCTCTCTGTGGGTTCTACCCAGCTATCCCAAGTCCGCATCAAGACCATCAAGACA 1200
          *****
Lewis      S S G E I K R T V T S V L P E V D L D S
DSS        TCATCTGGGAGATCAAGAGGACAGTACCAGAGTACTGCCAGAAAGTGGATCTGGACTCT 1260
          TCATCTGGGAGATCAAGAGGACAGTACCAGAGTACTGCCAGAAAGTGGATCTGGACTCT 1260
          *****
Lewis      G K E P E E Q A A S V M A S V T S L L S
DSS        GGGAGAAAGCCTTCTGAGCAGGCAGCGTCCGATGGCGCTCTGTGACTCCTCTCTGCA 1320
          GGGAGAAAGCCTTCTGAGCAGGCAGCGTCCGATGGCGCTCTGTGACTCCTCTCTGCA 1320
          *****
Lewis      S S A S A A V L S S P P R A P L P T A M
DSS        TCTTACGATCAGCTGGGTCCTCTCTCCACCCCGGAGGACCTCTGCCGACAGCCATG 1380
          TCTTACGATCAGCTGGGTCCTCTCTCCACCCCGGAGGACCTCTGCCGACAGCCATG 1380
          *****
Lewis      V T S A V S S A E L T P K Q V T I K P V
DSS        GTTACAGTGGCGTTTCTCTGSCAGAGCTGACTCCCAAACAGGTAAACATCAAGCCGCG 1440
          GTTACAGTGGCGTTTCTCTGSCAGAGCTGACTCCCAAACAGGTAAACATCAAGCCGCG 1440
          *****
Lewis      A T A F L P V S A V K T A S S Q V I N L
DSS        GCTACAGCTTTTCTCTGTGTCTGCGTCAAGACAGCAGGGTCTCAAGTCACTAACCTG 1500
          GCTACAGCTTTTCTCTGTGTCTGCGTCAAGACAGCAGGGTCTCAAGTCACTAACCTG 1500
          *****
Lewis      K L A N N T T V K A T V I S A A S V Q S
DSS        AAGTGGCTAACAAACAACGSGTAAAGCCAGGTCTCTCGGCTGCTCTGTTCCAGAGC 1560
          AAGTGGCTAACAAACAACGSGTAAAGCCAGGTCTCTCGGCTGCTCTGTTCCAGAGC 1560
          *****
Lewis      A S S A I I K A A N A I G Q Q T V V V P
DSS        GCCAGCAGTGCATCAAGCTGCCAATGCCATCCAGCAGCAAACGTTGTGGTGGCG 1620
          GCCAGCAGTGCATCAAGCTGCCAATGCCATCCAGCAGCAAACGTTGTGGTGGCG 1620
          *****
Lewis      A S S L A N A E L V P E T V H L A N L N
DSS        GCCTCCAGTCTGGCAATGCCAAACTCGGCCAAAGACTGTGCACTTGGCAACCTTAAC 1680
          GCCTCCAGTCTGGCAATGCCAAACTCGGCCAAAGACTGTGCACTTGGCAACCTTAAC 1680
          *****
Lewis      L L P Q G A Q A T S E L R Q V L T K P Q
DSS        CTCTGCTCAGGTTGCCAGGCCACCTCTGAACTCCGCCAAGTGTCAACAACTCAG 1740
          CTCTGCTCAGGTTGCCAGGCCACCTCTGAACTCCGCCAAGTGTCAACAACTCAG 1740
          *****
Lewis      Q Q I K Q A I L N A A A A Q P P E K V S
DSS        CAGCAAATAAAGCAGGCAATACTCAATGCCGCGGCCGCGCAGCCACCAAGAAAGTGTCC 1800
          CAGCAAATAAAGCAGGCAATACTCAATGCCGCGGCCGCGCAGCCACCAAGAAAGTGTCC 1800
          *****
Lewis      R V Q V V S S I Q S S V V E A F N K V L
DSS        CGGTTCCAGTGGTGTCTCTCTTGCAGAGTTCGTGGTGGAAAGCTTTCAACAAGTGTCTG 1860
          CGGTTCCAGTGGTGTCTCTCTTGCAGAGTTCGTGGTGGAAAGCTTTCAACAAGTGTCTG 1860
          *****
Lewis      S S V N P V P V Y I P N L S P P A S A G
DSS        AGCAGCTCAACCCAGTCCCTGTTTACATCCGAACTCAGTCTCCCGCAGTCCCGGG 1920
          AGCAGCTCAACCCAGTCCCTGTTTACATCCGAACTCAGTCTCCCGCAGTCCCGGG 1920
          *****

```

I T L P M R G Y K C L E C G D S F A L E
 Lewis 1980
 DSS ATCACGTTACCGATGCGCGGGTACAAGTGTCTGGAGTGCGGGGACTCCTTCGCGCTGGAA 1980
 ATCACGTTACCGATGCGCGGGTACAAGTGTCTGGAGTGCGGGGACTCCTTCGCGCTGGAA 1980
 * * * * *
 K S L S Q H Y D R R E V S I E V T C N H
 Lewis 2040
 DSS AAGAGCCTGAGCCAGCACTACGACAGGCGCAGCGTGCGCATCGAAGTGAOBTGCAACAC 2040
 AAGAGCCTGAGCCAGCACTACGACAGGCGCAGCGTGCGCATCGAAGTGAOBTGCAACAC 2040
 * * * * *
 C T E N L V F Y N E C S L L E S H A R G H
 Lewis 2100
 DSS TGTACGAAGAACCTTGTITTTACACAAGTGCAGCGCTCTTTCCACGCGCGAGGGCAT 2100
 TGTACGAAGAACCTTGTITTTACACAAGTGCAGCGCTCTTTCCACGCGCGAGGGCAT 2100
 * * * * *
 K E K G V V M G C S H L L L E K F V P A D
 Lewis 2160
 DSS AAGGAGAAAGGCGTGGTATGCAAGTGTCCCACTTAATCTAAAACCASTCCACAGCAGAC 2160
 AAGGAGAAAGGCGTGGTATGCAAGTGTCCCACTTAATCTAAAACCASTCCACAGCAGAC 2160
 * * * * *
 Q M I V E P S S N T A A S T L P S S V G
 Lewis 2220
 DSS CAGATGATAGTGTCTCCGTCCAGCAATACTGCTGCTTCCACTCTGCCGAGCTCGGTGGSA 2220
 CAGATGATAGTGTCTCCGTCCAGCAATACTGCTGCTTCCACTCTGCCGAGCTCGGTGGSA 2220
 * * * * *
 A A T H T V A E V Q P G L T G A V I S A
 Lewis 2280
 DSS GCTGACACACACTGTGSCAAAGTACAGCTTGGCTTAACCGGGCGGTATCTCAGCT 2280
 GCTGACACACACTGTGSCAAAGTACAGCTTGGCTTAACCGGGCGGTATCTCAGCT 2280
 * * * * *
 P P E T F I S P A M P L D E D P E K L C
 Lewis 2340
 DSS CCTCCAGCACACCCATCAGCCAGCCATGCCCTTAGAGCAAGACCCCTCCAGCTCTGT 2340
 CCTCCAGCACACCCATCAGCCAGCCATGCCCTTAGAGCAAGACCCCTCCAGCTCTGT 2340
 * * * * *
 R H S L K C L E C N E V F Q D E T S L A
 Lewis 2400
 DSS AGACACAGTCTCAAGTGTGGAGTGCATGAAGTCTTCCAGGATGAGAGCTCCCTGGCC 2400
 AGACACAGTCTCAAGTGTGGAGTGCATGAAGTCTTCCAGGATGAGAGCTCCCTGGCC 2400
 * * * * *
 ↓
 T H F Q H A A E T C G G Q K T C T V C Q H
 Lewis 2460
 DSS ACACATTTCCAGCATGCCGCGAGACACATGTGGACAAAAGACTTGCATGTCTGCCAGATG 2460
 ACACATTTCCAGCATGCCGCGAGACACATGTGGACAAAAGACTTGCATGTCTGCCAGATG 2460
 * * * * *
 L L E N Q C S Y A S H Q R I H Q H K S P
 Lewis 2520
 DSS CTGCTTCTAACCAAGTGCAGCTATGCTTACACAGAGAAATCCATCAGCACAAATCTCC 2520
 CTGCTTCTAACCAAGTGCAGCTATGCTTACACAGAGAAATCCATCAGCACAAATCTCC 2520
 * * * * *
 Y T C P E C G A I C R S V R F Q S H V T
 Lewis 2580
 DSS TACACTGCCCCGAGTGCGGGGCCATCTGCAGGTGGTGCATTCAGAGCCAGTCAAC 2580
 TACACTGCCCCGAGTGCGGGGCCATCTGCAGGTGGTGCATTCAGAGCCAGTCAAC 2580
 * * * * *
 ↓
 K N C L H Y T R R V G F R C V H C H V V
 Lewis 2640
 DSS AAGAAGTGCCTGCCTACACACGAGAGTGGCTTTCCATGTGTACATTCGCAATGTTGTG 2640
 AAGAAGTGCCTGCCTACACACGAGAGTGGCTTTCCATGTGTACATTCGCAATGTTGTG 2640
 * * * * *
 Y S D V A A L K S H I Q G S H C E V F Y
 Lewis 2700
 DSS TACTCTGATGTGGCGCACTGAAGTCTCACATTCAGGCTCTCACTGTGAAGTCTTCTAC 2700
 TACTCTGATGTGGCGCACTGAAGTCTCACATTCAGGCTCTCACTGTGAAGTCTTCTAC 2700
 * * * * *
 K C F I C P M A F K S A P S T H S H A Y
 Lewis 2760
 DSS AAGTGCCTATCTGTCCAATGGCGTTAAGTCTGCCCCAAGTACACATTCOCCAGCCATC 2760
 AAGTGCCTATCTGTCCAATGGCGTTAAGTCTGCCCCAAGTACACATTCOCCAGCCATC 2760
 * * * * *
 ↓
 T Q H P G V K I G E P K I I Y K C S M C
 Lewis 2820
 DSS ACCGACATCCTGGCGTCAAGATAGGAGAACCAAAAATAATCTATAAGTGTTCATGTGC 2820
 ACCGACATCCTGGCGTCAAGATAGGAGAACCAAAAATAATCTATAAGTGTTCATGTGC 2820
 * * * * *
 D T V F T L Q T L L Y R H F D Q H I D N
 Lewis 2880
 DSS GACACTGTGTCACCCGCAAACCTTGTGTATGCCCACITTGACCAGCACATTGACAAC 2880
 GACACTGTGTCACCCGCAAACCTTGTGTATGCCCACITTGACCAGCACATTGACAAC 2880
 * * * * *
 Q K V S V F K C P D C S L L Y A Q K Q L
 Lewis 2940
 DSS CAGAAGTGTCTGTTTTCAAGTGTCCAGACTGTCTCTTTTATATGCCAGAGCAACTT 2940
 CAGAAGTGTCTGTTTTCAAGTGTCCAGACTGTCTCTTTTATATGCCAGAGCAACTT 2940
 * * * * *

↓
 Lewis M M D H I K S M H B T L K S I E G P P N 3000
 DSS ATGATGGACCATATCAAGTCTATGCAATGGAACATTGAAAAGTATTGAAGGGCTCCAAAC 3000
 *
 Lewis L G I N L F L S S K F A T Q N S A N H S 3060
 DSS TTGGTATAAACTTGCCTTTGAGCAGTAAGCTGCAACTCAGAATTCAGCAAACACAGC 3060
 *
 Lewis R E D A K S M N E K E E L E K K S F S P 3120
 DSS AGAGGGATGCCAAGTCCATGAATGGGAAGAGAAATTGGAAAAGAAGTCCCATCTCCT 3120
 *
 Lewis A K K S T E S E K M A S L G W T C W E C 3180
 DSS GCAAGAAATCCACAGAATCTAAGAGATGGCCAGTCTCGGGTGGACGTGTTGGAGTGT 3180
 *
 Lewis D R L F T Q R D V Y L S H M R K E H G K 3240
 DSS GACCTCTGTTACACAGAGGATGTGTATCTGTCTCACATGAGGAAGGAGCATGGGAAG 3240
 *
 Lewis Q M K K H F C R Q C D K S F S S S H S L 3300
 DSS CAAATGAAGAGCACCCCTGCCGCGAGTGTGACAAAGTCTTCAGCTCCTCCACAGCTG 3300
 *
 Lewis C R H N R I K H K B I R E V Y A C S H C 3360
 DSS TGCCGCCAATCGCATCAAGCACAAAGGCATCAGGAAGSTTTACGCTGCTGCACTGC 3360
 *
 Lewis F D S R R T F T K R L M L E K H I Q L M 3420
 DSS CCAGACTCCCGTCCGACTTCACCAAGCGCTGATGCTGGAGAAGCACATACAGCTGATG 3420
 *
 Lewis H G I K D F D V K E L S D E V T N E E E 3480
 DSS CACGGATCAAGAGCCCTGACGTAAGAGAGCTGAGTGACGAGGTTACCAACGAGGAGAA 3480
 *
 Lewis V E I K E D A K V P S P K R K L E E P V 3540
 DSS GTGGAGATAAAGGAGGACGCCAAGSTTCCAGTCTTAAGGGAAAGTGGAGGAGCCAGTT 3540
 *
 Lewis L E F R F P R G A I T Q F L K K L K I N 3600
 DSS TTAGAGTTCAGGCTCCCGAGGAGCCATCACTCAGCCACTGAAGAACTGAAAATCAAT 3600
 *
 Lewis V F E V R K C A V C G P T T E H L L Q F 3660
 DSS GTCTTTAAGSTCCACAAGTGCSCDGTGTGTGGCTTCAACCCTGAGAACTGCTGCAGTTC 3660
 *
 Lewis H E H I F Q H K S D G S S Y Q C R E C G 3720
 DSS CACGACACATCCCCAGCACAAAGTCGGACGGCTCCTCTACCAGTGCSSGAGTCCGGC 3720
 *
 Lewis L C Y T S H V S L S R H I F I V H K L E 3780
 DSS CTGTGCTACAGTCCCAAGTCTCCTGTCCAGGACCTCTTCATGTTGCACAGCTGAAG 3780
 *
 Lewis E P Q P V S E K Q N G A G E D S Q Q E N K 3840
 DSS GAGCCTCAGCCTGTGTCCAAGCAGAACGGGCTGGGGAAGACAGCCAGCAGGACAAAG 3840
 *
 Lewis F S P E D E A A E G A V S D R K C K V C 3900
 DSS CCCAGCCCGAGGATGAGGCCCGGAGGGGCASTGTGACACAGGAAGTGCAAAAGTGTGC 3900
 *
 Lewis A K T F E T E A A L N T H M R T H G M A 3960
 DSS GCCAAAACCTTCGAAACGSAAGCTGCCTTAAACACACACATGAGGACACAGGCATGGCC 3960
 *

```

Lewis      F I E E E R M S S A E E -
DSS       TTCATCAAGTCCAAAAGAATGAGTTCAGCTGAAAAATAG 3999
          TTCATCAAGTCCAAAAGAATGAGTTCAGCTGAAAAATAG 3999
          *****

```

(5) *O-acyltransferase like (Oacyl)*

```

Lewis      M T A F T L L A C L H A L F P F V S P A
DSS       ATGAGGGGCTTCACITTTGCTAGCTTGCCTGCACGCTCTGTTCCTTTTGTTCCTCTGCA 60
          ATGAGGGGCTTCACITTTGCTAGCTTGCCTGCACGCTCTGTTCCTTTTGTTCCTCTGCA 60
          *****
          R N I S L K C H Q D T D E F L S D L N S
Lewis      AGAAACATTTCTCTCAAGTGCATGCAAGATACATGATGAGTTCCTTTCTGATCTGAATCA 120
DSS       AGAAACATTTCTCTCAAGTGCATGCAAGATACATGATGAGTTCCTTTCTGATCTGAATCA 120
          *****
          |
Lewis      L E F P R E Y A L R M Y D S V G K L E S N
DSS       CTGAAGCCCAAAGAATAACGCTCTGAGAATGTACGACTCGGTGGGGAAGCTTGAAGCAAC 180
          CTGAAGCCCAAAGAATAACGCTCTGAGAATGTACGACTCGGTGGGGAAGCTTGAAGCAAC 180
          *****
          V L N G N V D R L G S Y S E C L S T Q S
Lewis      GTTCTTAACGGCAACGTGACAGGCTGGGCTCTACAGCGAGTGCCTTTCACGCGAGAC 240
DSS       GTTCTTAACGGCAACGTGACAGGCTGGGCTCTACAGCGAGTGCCTTTCACGCGAGAC 240
          *****
          |
Lewis      P E G S F R G Q Y C K L H I L Q D G T D
DSS       CCCBAGGGGAGTTCCCGGGCCACTACTGCAAGCTTCATATTTCTGCGAGSATGGAACAGAC 300
          CCCBAGGGGAGTTCCCGGGCCACTACTGCAAGCTTCATATTTCTGCGAGSATGGAACAGAC 300
          *****
          Y S V G V C V P D S C A E E D V T T M S
Lewis      TACAGTGTGGGTGTGTGCGTCCCTGATTTGTGTGCTGAAGAGGACGTGACCCASATGTCT 360
DSS       TACAGTGTGGGTGTGTGCGTCCCTGATTTGTGTGCTGAAGAGGACGTGACCCASATGTCT 360
          *****
          |
Lewis      Q L E I L R F R N T S F L E P S L S L F
DSS       CAGTGGACATTTCTACGGTTCCAGAAACACTTCGTTCTTGGAACTTCCCTGTCTCTTTT 420
          CAGTGGACATTTCTACGGTTCCAGAAACACTTCGTTCTTGGAACTTCCCTGTCTCTTTT 420
          *****
          T K D S S S S C E I V A R C A A G A L E
Lewis      ACAAAAGACTTCTTTCATCTTGTGAATCGTGGCCAGATGTGCTGCTGGGGCTTGAAG 480
DSS       ACAAAAGACTTCTTTCATCTTGTGAATCGTGGCCAGATGTGCTGCTGGGGCTTGAAG 480
          *****
          |
Lewis      P D M F S S V C L F I T L L G L V L P V
DSS       CCGSACATGTTTGTGTCAGTGTGTCTGTTCACTCACTTGTAGGCTTGTCTTCTCTGTG 540
          CCGSACATGTTTGTGTCAGTGTGTCTGTTCACTCACTTGTAGGCTTGTCTTCTCTGTG 540
          *****
          A G T V Y M V A R D W G L D L K A A S V
Lewis      GCTGGAACTGTTTACATGTTGGCCAGAGACTGGGGCTTGGACTCAAGGCGGCATCTGTG 600
DSS       GCTGGAACTGTTTACATGTTGGCCAGAGACTGGGGCTTGGACTCAAGGCGGCATCTGTG 600
          *****
          C G T P T S L E S L P L G N M E S N R Q
Lewis      TGTGGACTCCTACCACTCTCGAGAGTCTGCTTTGGGAAACATGGAGAGCAACAGGCAA 660
DSS       TGTGGACTCCTACCACTCTCGAGAGTCTGCTTTGGGAAACATGGAGAGCAACAGGCAA 660
          *****
          R N R T S C Q V Q L P F P G P P S R G K
Lewis      AGAAACAGSACCAGTTGCCAAGTTCAGCTCCCTCTCCGGACCCCGAGCAGAGGAAAA 720
DSS       AGAAACAGSACCAGTTGCCAAGTTCAGCTCCCTCTCCGGACCCCGAGCAGAGGAAAA 720
          *****
          |
Lewis      R F L G A L D G V L Q C F S N Q K H M P
DSS       ABGTTTCTAGGAGCCTTGSATGGAGTCTGCACTGCTTTTCTTGGCAGAAGAACATGCCA 780
          ABGTTTCTAGGAGCCTTGSATGGAGTCTGCACTGCTTTTCTTGGCAGAAGAACATGCCA 780
          *****
          A I C N P E L E G S T C Q T L N G I E V
Lewis      GCCATCTGCAACCCAGAGTGCACAGGAGSACCTGCCASACTGAACGSCATCCGGGTC 840
DSS       GCCATCTGCAACCCAGAGTGCACAGGAGSACCTGCCASACTGAACGSCATCCGGGTC 840
          *****

```

```

Lewis      L S L L N I I S G H T S Q M T A W L S L
DSS        CTGAGTCTCCTCTGGATTATCTCGGGSCACACCACTCAGATGACCGCTTGGCTCTCTTTG 900
          CTGAGTCTCCTCTGGATTATCTCGGGSCACACCACTCAGATGACCGCTTGGCTCTCTTTG 900
          *
          ↓
Lewis      D N V L E M K A R V F K N P L Y L Y S R
DSS        GATAATSTGCTTGAATGGAAAGCCAGAGTGCACCAAGAACCCGCTGTACTCTACTCTCGG 960
          GATAATSTGCTTGAATGGAAAGCCAGAGTGCACCAAGAACCCGCTGTACTCTACTCTCGG 960
          *
          ↓
Lewis      S G F F Y L G V D T F F L I S G W L S A
DSS        AGTGGCCCBTTCTACCTCGGGSTGGATACATTTTCTCTGATCASTGGTGGTTAAGTGA 1020
          AGTGGCCCBTTCTACCTCGGGSTGGATACATTTTCTCTGATCASTGGTGGTTAAGTGA 1020
          *
          ↓
Lewis      R S F L K M H Q N T S T G I T P K V I L
DSS        CGGTCAATTCATAAGATGCACCAAGAACCCAGCAGGAATAACCCCTAAAGTCAATTC 1080
          CGGTCAATTCATAAGATGCACCAAGAACCCAGCAGGAATAACCCCTAAAGTCAATTC 1080
          *
          ↓
Lewis      R Y F L K R F T H L Q L L H L Y S V C L
DSS        AGATACTTCTGAAACBCTTTACAAGSTTGCAGCTTCTTCACTGTACTCTGTGTGCTTG 1140
          AGATACTTCTGAAACBCTTTACAAGSTTGCAGCTTCTTCACTGTACTCTGTGTGCTTG 1140
          *
          ↓
Lewis      L V G F F S F V P N G E V W E V A Q F H
DSS        CTGTTTGGTTCTCTCTCTTTTGTCCCTGGGGACCTGTCTGGGAAGTGGCCAGTTCAC 1200
          CTGTTTGGTTCTCTCTCTTTTGTCCCTGGGGACCTGTCTGGGAAGTGGCCAGTTCAC 1200
          *
          ↓
Lewis      W D N C R E V W W T H L L L L N F L S
DSS        TGGGATAACTGCCGGAGSTTTGGTGGACGAATCTGCTGCTGCTGAATAACTTCTCTCT 1260
          TGGGATAACTGCCGGAGSTTTGGTGGACGAATCTGCTGCTGCTGAATAACTTCTCTCT 1260
          *
          ↓
Lewis      V Q E A C H S M T N Y L A S D F Q F H L
DSS        GTGCAAAAACCTGCAATGGCTGSACTTGGTACTTTCARAGCACTTCCAGTTCACCTC 1320
          GTGCAAAAACCTGCAATGGCTGSACTTGGTACTTTCARAGCACTTCCAGTTCACCTC 1320
          *
          ↓
Lewis      H T P L I F L V H G K S K R A L V L L G
DSS        ATGACACCGCTGATCTTCTTAGTCCATGGAAAAAGTAAAGCGCCCTCTCTCTCTCGGG 1380
          ATGACACCGCTGATCTTCTTAGTCCATGGAAAAAGTAAAGCGCCCTCTCTCTCTCGGG 1380
          *
          ↓
Lewis      G L L F L V S S T A T A L L T V A Y K L
DSS        GGCTGCTGTCTCTGGTGTCTTCCACTGCCACTGCTCTGCTCACTGTGCTTATAAACTT 1440
          GGCTGCTGTCTCTGGTGTCTTCCACTGCCACTGCTCTGCTCACTGTGCTTATAAACTT 1440
          *
          ↓
Lewis      P V S S F S A A S E N E T V L Y F S E Y
DSS        CCAGTGTCACTCTCCGTCAGCAGCCAGTGAATAAGAGACTGTACTGTATTCTCAGAGTAT 1500
          CCAGTGTCACTCTCCGTCAGCAGCCAGTGAATAAGAGACTGTACTGTATTCTCAGAGTAT 1500
          *
          ↓
Lewis      Y T K P Y C R C G F F L V G L F L S M F
DSS        TACCAAACCCCTACTGTGCTGTGSCCAATCTTGTGGCCCTCTTCTGAGCATGTC 1560
          TACCAAACCCCTACTGTGCTGTGSCCAATCTTGTGGCCCTCTTCTGAGCATGTC 1560
          *
          ↓
Lewis      M H E D E P T D V L R T K V Q A V L G W
DSS        ATGCATTGACACCACCTACAGAGTCTCAGAACTAAGTGCAGGCTGTGCTGGCTGG 1620
          ATGCATTGACACCACCTACAGAGTCTCAGAACTAAGTGCAGGCTGTGCTGGCTGG 1620
          *
          ↓
Lewis      A F S L T T L F A V V A L A Y V V D D T
DSS        GCTTCTCTCTGACGACTCTCTTTGCACTTGTGGCCCTGGCGTATGTAGTGGATGATA 1680
          GCTTCTCTCTGACGACTCTCTTTGCACTTGTGGCCCTGGCGTATGTAGTGGATGATA 1680
          *
          ↓
Lewis      S P T S E V L A A I Y Q A L H R T L W A
DSS        TCTCCAACTCTCTGTGCTTGTGCTATCTACAGSCCCTGCAACCGACCCCTGTGGCA 1740
          TCTCCAACTCTCTGTGCTTGTGCTATCTACAGSCCCTGCAACCGACCCCTGTGGCA 1740
          *
          ↓
Lewis      A A V G W V I F A C Q E G F G G P V E R
DSS        GCAGCTGTGGCTGGTATCTTTGCTGTGAGSAGGCTTGGAGTCCCGTGAAGC 1800
          GCAGCTGTGGCTGGTATCTTTGCTGTGAGSAGGCTTGGAGTCCCGTGAAGC 1800
          *

```

```

Lewis      L L E C S I W A L P A S I S Y A C Y L V
DSS        TTGCTTTCTGTGGCAICTGGGCTCTCCCTGCTAGCATAAGCTATGCTTGTACTACCTCGTA 1860
           TTGCTTTCTGTGGCAICTGGGCTCTCCCTGCTAGCATAAGCTATGCTTGTACTACCTCGTA 1860
           *
Lewis      H F I V I I L Y N G L Q E T L I H Y T D
DSS        CACCCCATTTGATCATCTCTATAAATGGACTTCAGSAGACCCTTATTCACTACACAGAC 1920
           CACCCCATTTGATCATCTCTATAAATGGACTTCAGSAGACCCTTATTCACTACACAGAC 1920
           *
Lewis      T N N F Y L F S G H C V I T F L C G L A
DSS        ACCAACATGTTCTATCTTTTCTCTGGACACTGCGTCTGACCTTCTCTGTGGGCTGGCC 1980
           ACCAACATGTTCTATCTTTTCTCTGGACACTGCGTCTGACCTTCTCTGTGGGCTGGCC 1980
           *
Lewis      L T L F I E R F W Q E L E W G L -
DSS        CTGACCCCTGTTTCATTGAGGGCCATGGCAAGAACTGAAAGTGGGGCTGTGA 2031
           CTGACCCCTGTTTCATTGAGGGCCATGGCAAGAACTGAAAGTGGGGCTGTGA 2031
           *

```

(6) SEC11 homolog C (*S. cerevisiae*) (Sec11c)

```

Lewis      M V R A G A V G T H L P T S S L D I F G
DSS        ATGTTCTGTGCGGGCCCTGGGAGCGCATCTCCCCAGTCCAGSCTTGGACATCTTTGGG 60
           ATGTTCTGTGCGGGCCCTGGGAGCGCATCTCCCCAGTCCAGSCTTGGACATCTTTGGG 60
           *
Lewis      D L R K M N K R Q L Y Y Q V L N F A M I
DSS        GACTTGAGGAGATGAACAAACGACAGCTTTACTACAGSCTTCAAACCTTGGCATGATT 120
           GACTTGAGGAGATGAACAAACGACAGCTTTACTACAGSCTTCAAACCTTGGCATGATT 120
           *
Lewis      V S A L M I N K S L I V L T G S E S P
DSS        GTGTCTTCTGCGCTCATGATATGAAAGSCTTAATTGTGCTCACGGGCAAGGAGATGCC 180
           GTGTCTTCTGCGCTCATGATATGAAAGSCTTAATTGTGCTCACGGGCAAGGAGATGCC 180
           *
Lewis      I V V V L S G E M E F A F H R G D L L F
DSS        ATGTTGGTGGTACTGAGTGGCAGTATGGAGCCGGCCTTTCACAGGGAGATCTGCTGTC 240
           ATGTTGGTGGTACTGAGTGGCAGTATGGAGCCGGCCTTTCACAGGGAGATCTGCTGTC 240
           *
Lewis      L T N F R E D F I R A G E I V V F K V E
DSS        CTCACAAATTTCCGGGAGSATCCCATCAGAGCTGGTGAATAGTTGTTTTAAAGTTGAA 300
           CTCACAAATTTCCGGGAGSATCCCATCAGAGCTGGTGAATAGTTGTTTTAAAGTTGAA 300
           *
Lewis      G R D I F I V H R V I K V H E K D N G D
DSS        GGAAGAGACATTCGATAGTTTACAGAGTAATCAAGSCTTCAAAAAAGATAATGGAGAC 360
           GGAAGAGACATTCGATAGTTTACAGAGTAATCAAGSCTTCAAAAAAGATAATGGAGAC 360
           *
Lewis      I K F L T K G E N N E V E D R G L Y K E
DSS        ATCAAGTTTCTGACTAAAGGAGATAATAATGAAGTTGATGATCGAGSCTTGTACAAAG 420
           ATCAAGTTTCTGACTAAAGGAGATAATAATGAAGTTGATGATCGAGSCTTGTACAAAG 420
           *
Lewis      G Q N W L E K K D Y V G E A R S F L P Y
DSS        GGCCGAACTGGCTGGAGAAGSAGACGTGGTASGAAGSCTCGAGSCTTCTTACCATAT 480
           GGCCGAACTGGCTGGAGAAGSAGACGTGGTASGAAGSCTCGAGSCTTCTTACCATAT 480
           *
Lewis      V G H V T I I N N D Y P K F E Y A L V A
DSS        GTTGGCATGGTACCATTATAATGAACGACTATCTAAGTCAAGTACBCTCTTGTGGCT 540
           GTTGGCATGGTACCATTATAATGAACGACTATCTAAGTCAAGTACBCTCTTGTGGCT 540
           *
Lewis      V M G A Y V L L K R E S -
DSS        GTGATGGGCGCATACTGTTACTGAAACSTGAATCTAA 579
           GTGATGGGCGCATACTGTTACTGAAACSTGAATCTAA 579
           *

```

Footnote to Supplement: * indicates nucleotide identity. Amino acid sequence is given on top. Genomic DNAs and/or cDNA from an organ from DSS and Lewis were amplified by PCR and then sequenced for

Alpk2, *Nedd4l*, *Rbfa*, and *Adnp2*. Mutations were confirmed for *Loxhd1* by resequencing genomic DNAs. SNPs for the rest of the genes were retrieved from our genome data base of DSS and Lewis rats (Supplement 2). The deleted nucleotides are indicated by ---. Synonymous SNPs are marked in shades. The amino acid changes caused by individual mutations are indicated by bold lettering. ↓ marks the last nucleotide of each exon.

Supplement 7. Predictions for a missense mutation on the function of the protein product.

a) Amino acid substitutions within the *Alpk2* Ala637 Δ deletion

Possible Codon	Amino acid position	Amino acid	Mutpred	SIFT
GCC	637	Ala (wild type)	N/A	N/A
TCC	637	Ser	benign	benign
CCC	637	Pro	benign	benign
ACC	637	Thr	benign	benign
GGC	637	Gly	benign	benign
GAC	637	Asp	benign	benign
GTC	637	Val	benign	benign

b) Amino acid substitutions within the *Tcof1* ESDSEE762 Δ 767 deletion

Possible Codon	Amino acid position	Amino acid	Mutpred	SIFT
TCT	763	Ser (wild type)	N/A	N/A
ACT	763	Thr	benign	Deleterious
GCT	763	Ala	Deleterious	Deleterious
CCT	763	Pro	Deleterious	Deleterious
TAT	763	Tyr	benign	Deleterious
IGT	763	Cys	Deleterious	Deleterious
TTT	763	Phe	benign	Deleterious
AGT	765	Ser (wild type)	N/A	N/A
TGT	765	Cys	Deleterious	Deleterious
CGT	765	Arg	Deleterious	Deleterious
GGT	765	Gly	Deleterious	Deleterious
AAT	765	Asn	Deleterious	Deleterious
ATT	765	Ile	Deleterious	Deleterious
ACT	765	Thr	benign	Deleterious

Footnote to supplement: Each of the 3 nucleotide for a codon encoding a given amino acid is hypothetically mutated and assessed for possible functional consequence. MutPred score ¹ (<http://mutpred.mutdb.org/>), and SIFT ² (<http://sift.icvi.org/>) refer to various prediction programs used for assessing a probable consequence of a missense mutation on the function of a protein it encodes.

Supplemental references

- (1) Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, Mooney SD, Radivojac P. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 2009;25:2744-50.
- (2) Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073-81.

11.10 Addendum

Limitations de l'étude :

L'analyse des gènes dans les régions contenant les C18QTL s'est limitée à l'évaluation des variantes codantes et des jonctions intron/exon. Cependant, les variantes non codantes peuvent avoir des conséquences fonctionnelles sur les gènes avoisinants [156]. Ceci représente une des limites de l'étude présentée dans cette publication et est discutée d'avantage dans la section 13.1.2 du présent document.

La nomenclature des variantes recommandée par l'HGVS est la suivante [154]:

C18QTL3

NM_001106143.1(Tcof1):c.2283_2300delGGAGTCTGACAGTGAAGA, p.(762_767delESD SEE)

XM_006254816.2(Cdx1):c.105C>G,

XM_006254816.2(Cdx1):c.107C>T,p. (36Thr>Ile)

XM_006254816.2(Cdx1):c.119A>C,p. (40Gln>Pro)

XM_006254816.2(Cdx1):c.130A>C,p. (44Thr>Pro)

XM_006254816.2(Cdx1):c.163A>T,p. (55Asn>Tyr)

XM_006254816.2(Cdx1):c.205A>G,p. (69Thr>Ala)

NM_001008300.2(Nedd4l):c.321A>T

NM_001008300.2(Nedd4l):c.1995T>C

NM_001008300.2(Nedd4l):c.2139G>A

XM_006254887.2(Alpk2):c.376C>T,p.(126Arg>Cys)

XM_006254887.2(Alpk2):c.475C>T,p.(159Pro>Ser)

XM_006254887.2(Alpk2):c.618C>T

XM_006254887.2(Alpk2):c.1910_1912delCCG, p.(637delAla)

XM_006254887.2(Alpk2):c.2918C>T,p.(973Gln>Leu)

XM_006254887.2(Alpk2):c.3754G>A,p.(1252Glu>Lys)

NM_001107382.1(Zfp532):c.90G>A

NM_001107382.1(Zfp532):c.810G>A

NM_001106139.1(Oacyl):c.1059G>A

NM_153628.2(Sec11c):c.579T>C

C18QTL4

NM_001108891.1(Ctif):c.189A>G

NM_001107373.1(Rbfa):c.474G>A

NM_001107373.1(Rbfa):c.487A>G p.(163Thr>Ala)

NM_001107373.1(Rbfa):c.911A>G,p.(163Glu>Gly)

NM_001014151.2(Hdhd2):c.337G>C,p.(113Val>Leu)

NM_001127373.1(Adnp2):c.353A>G,p.(118Lys>Arg)

NM_001127373.1(Adnp2):c.1607C>G, p.(536Gln>Pro)

NM_001127373.1(Adnp2):c.1926A>G

NM_001127373.1(Adnp2):c.2140A>G,p.(714Met>Val)

NM_001127373.1(Adnp2):c.2149G>A,p.(717Ala>Thr)

NM_001127373.1(Adnp2):c.2499A>G

NM_001127373.1(Adnp2):c.3126G>A

NM_001127373.1(Adnp2):c.3144T>C

NM_001127373.1(Adnp2):c.3400G>C,p.(1134Glu>Gln)

NM_001100973.1(Pard6g):c.309T>C

NM_001100973.1(Pard6g):c.537C>T

NM_001100973.1(Pard6g):c.699A>G

NM_213628.1(St8sia5):c.246C>T

NM_001106132.3(Loxhd1):c.2190T>C

NM_001106132.3(Loxhd1):c.2859C>G

NM_001106132.3(Loxhd1):c.2880C>T

NM_001106132.3(Loxhd1):c.2937G>A

NM_001106132.3(Loxhd1):c.3346G>A,p.(1116Val>Met)

NM_001106132.3(Loxhd1):c.3880G>A,p.(1294Val>Ile)

NM_001106132.3(Loxhd1):c.3991G>A,p.(1331Val>Ile)

NM_001106132.3(Loxhd1):c.4107G>A

NM_001106132.3(Loxhd1):c.5268C>T

NM_001106132.3(Loxhd1):c.5929A>G, p. (1977Ile>Val)

NM_001106132.3(Loxhd1):c.6375C>T

XM_008772236.1(Setbp1):c.225C>T

XM_008772236.1(Setbp1):c.242G>A

XM_008772236.1(Setbp1):c1668A>G

NM_001106130.1(Atp9b):c.7T>C

NM_001106130.1(Atp9b):c.1003T>C

NM_001106130.1(Atp9b):c.3393A>G

NM_001108892.1(Sall3):c.2865A>G

NM_001108892.1(Sall3):c.3460C>T,p.(1154Arg>Cys)

NM_001108892.1(Sall3):c.3483C>G,p.(1161Ile>Met)

11.10.1 Résultats complémentaires

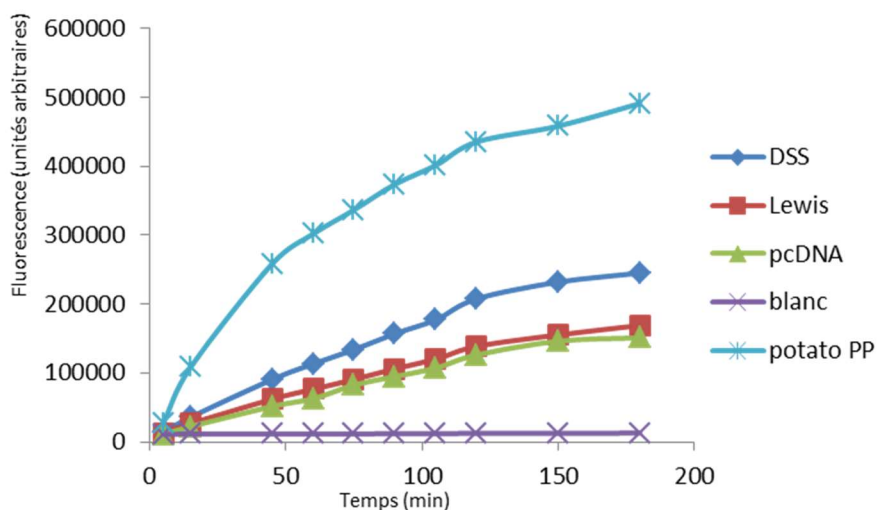


Figure A1. Résultats complémentaires pour l'activité enzymatique de Hdhd2.

Détection d'activité phosphatase sur les lysats cellulaires de HEK293. La fluorescence du produit du DiFMUP est mesurée à différents temps d'incubation (25°C). 250ng de protéines totales extraites des cellules HEK293 transfectées avec les vecteurs pcDNA3.1 seul (pcDNA), pcDNA-Hdhd2^{DSS} (DSS) ou pcDNA-Hdhd2^{Lewis} (Lewis) (cf. section 11.3.3 p.166), ou du contrôle positif fourni, (phosphatase acide de patate, potato PP) (E12020, Invitrogen) sont incubées avec 100µM de DiFMUP.

Erratum : Le coefficient de variation (CV) des expériences de cinétique enzymatique présentées dans la publication est de 8,7%.

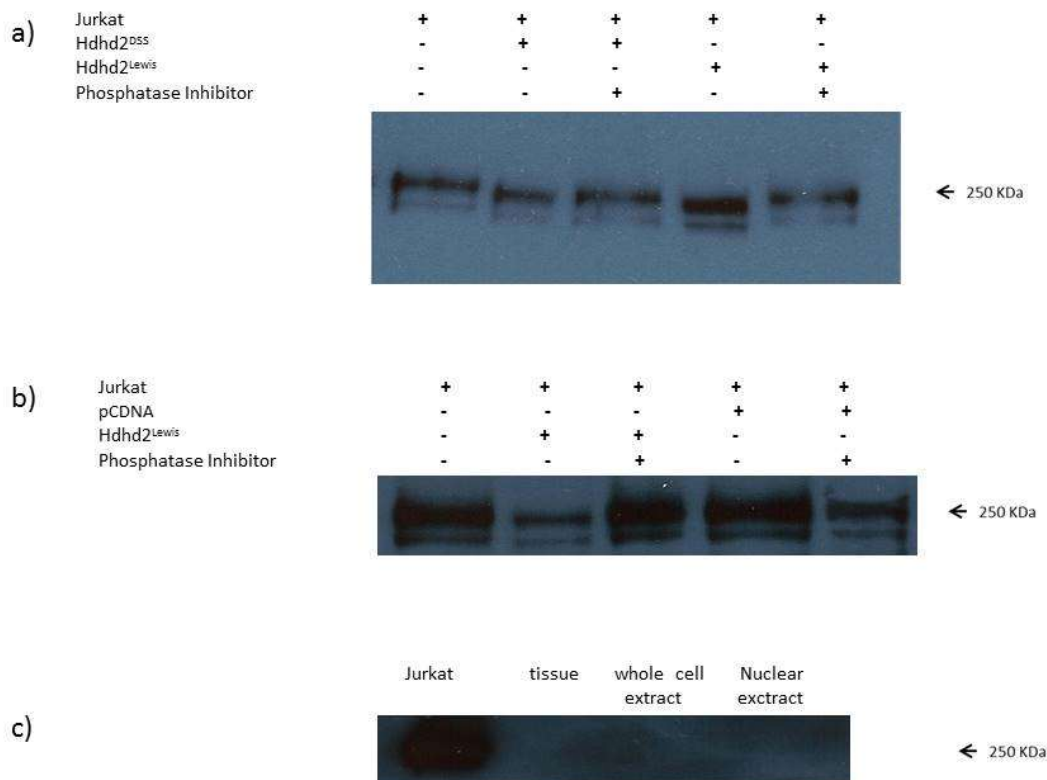


Figure A2. Détection par immunobuvardage de type Western de la déphosphorylation de Tcof1 par les isoformes de Hdhd2 et contrôles.

a) Détection de la déphosphorylation de Tcof par les deux isoformes de Hdhd2. Les conditions expérimentales sont telles que décrites dans la section 11.3.3 du présent document. Jurkat : Jurkat nuclear cell extract (SCB, SC-2132), Hdhd2^{DSS} et Hdhd2^{Lewis} : extraits cellulaires des cellules HEK293 transfectées avec les vecteurs, pcDNA-Hdhd2^{DSS} ou pcDNA-Hdhd2^{Lewis}. **b) Détection de la déphosphorylation de Tcof Hdhd2^{Lewis} et contrôle;** pcDNA : extraits cellulaires des cellules HEK transfectées avec le vecteur vide. **c) Détection de de l'expression de Tcof1.** Jurkat : 50 µg de Jurkat nuclear cell extract (SCB, SC-2132) ont été chargées sur gel; Tissue : 150 µg extraites du rein ont été chargées sur le gel; Whole cell extract : 100 µg de protéines totales du lysat des cellules de l'aorte de DSS ont été chargées. Nuclear extract : 75 µg de protéines totales d'un extrait nucléaire obtenu par lysat des cellules de l'aorte de DSS ont été chargées.

Discussion

12. Nouvelles découvertes dans l'architecture génétique de la pression artérielle chez le rat DSS

D'après les estimations, 1,5 milliards de personnes seront touchées par l'hypertension d'ici l'année 2025 [41]. Ainsi, l'étude et la compréhension des mécanismes à la base de cette pathologie par la découverte de QTL et de leurs interactions est d'une importance capitale.

Les constatations principales du projet exposé sont, tout d'abord, que les 28 QTL influençant la PA identifiés jusqu'à ce jour peuvent être séparés en trois modules qui régulent la pression artérielle. Les modules épistatiques EM1 et EM2 sont les « unités » fondamentales via lesquelles l'ensemble de ces QTL expriment leurs effets. Certaines caractéristiques des QTL, comme l'existence d'allèles ayant des effets opposés, permettent d'établir, en plus, une hiérarchie dans leur organisation. Sur cette même logique, la découverte de la nature moléculaire des QTL ainsi que la mise en évidence d'interactions physiques entre eux permettent d'élargir la compréhension de cette organisation séquentielle. C'est le cas des gènes candidats *Tcofl* et *Hdhd2* dont l'identification pourrait permettre de mieux comprendre le rôle physiologique des QTL et leur effet sur la PA.

12.1 Évidence de l'existence d'interactions épistatiques des QTL

S'agissant d'un trait polygénique, il a longtemps été pensé que la tension artérielle est le résultat de l'addition des effets minuscules des différents QTL. Par cette logique, si chaque QTL exerce un effet minime sur le phénotype, l'accumulation d'un nombre considérable de QTL serait nécessaire pour atteindre un seuil qui permet de détecter une variation de la PA [43].

Cependant, plusieurs observations contredisent ce paradigme. Tout d'abord, les QTL détectés jusqu'à présent chez l'humain par GWAS, ne représentent que <3% de la variance de PA [41]. Cette incapacité à détecter l'ensemble des effets des QTL s'explique par le

phénomène d' « hérabilité manquante » (« missing heritability ») où les variations observées ne représentent pas l'estimation de l'hérabilité du trait étudié [157]. De plus, lorsqu'il s'agit de QTL identifiés par souches congéniques de rats, l'évidence existe que ces derniers agissent de façon monogénique, et que la présence d'un seul QTL est nécessaire à faire varier la PA de façon significative [94, 106, 110]. Le paradoxe réside dans le fait que, si l'on additionne les effets des 28 QTL identifiés jusqu'à présent ($\pm 450\text{mmHg}$), la somme reste largement supérieure à la différence de pression artérielle observée entre les souches parentales DSS et Lewis ($\pm 90\text{mmHg}$).

Plus encore, comme mentionné dans la section 5.2.2 du présent document, la présence d'allèles hypertensifs chez les souches normotendues et d'allèles qui font baisser la tension chez la souche DSS, démontre qu'il existe une organisation structurée dans l'interaction des QTL tout au long du génome permettant à l'organisme de maintenir l'homéostasie de la PA. Il est à noter que les QTL du chromosome 18 discutés dans les résultats du présent document, n'ont pas été repérés lors de la construction de la souche réciproque pour le chromosome 18, c'est-à-dire, lorsque des allèles DSS ont été introduits dans un fond génétique Lewis [108]. Ces observations ne sont pas dues au manque d'allèles hypertensifs chez le contrôle normotendu mais seraient plutôt dues à la capacité que possède le génome récepteur à masquer l'influence des QTL introduits. Ainsi, il existe des interactions des QTL au niveau du génome complet qui, dans le cas de Lewis, empêchent l'expression du QTL détecté chez DSS. La capacité de régulation du génome Lewis est contrastante avec l'incapacité du génome DSS à moduler les impacts qui entraînent des variations de PA.

D'après ces observations, on peut conclure qu'il existe, tout au long du génome, des interactions épistatiques entre les QTL qui expliquent la redondance des QTL.

12.2 Classification des QTL par approche de génétique épistatique

12.2.1 Modularisation des QTL

Dans l'étude de l'hypertension et d'autres maladies complexes, les phénomènes épistatiques ont longtemps été vus comme des obstacles à l'identification des gènes causant la pathologie. Cependant, ce type d'interaction est essentiel et indispensable au maintien de la structure de contrôle de la PA [116].

Lorsqu'il existe des mutations dans deux gènes ou QTL distincts affectant la pression, la première question à poser est si tous deux agissent au niveau de la même cascade de régulation, et le cas échéant, dans quel ordre ils le font.

Au niveau du génome DSS, la classification de QTL par rapport à l'existence ou l'absence d'interactions épistatiques a rendu possible l'établissement de deux modules épistatiques (EM). La presque totalité des QTL identifiés au sein de notre laboratoire appartient à l'un ou à l'autre de ces modules qu'on appelle EM1 et EM2. Les QTL appartenant à un même module sont nécessairement épistatiques entre eux, et donc redondants, ainsi qu'additifs envers les QTL classés dans l'autre module.

Il est à noter qu'au cours de cette étude nous avons identifié un QTL, soit le *C2QTL1*, qui agit de façon additive avec les QTL de EM1 et avec ceux du EM2 et a donc été classifié dans un troisième module, EM3 (Figure 5a.). Plus encore, il existe 3 QTL; plus précisément *C7QTL*, *C1QTL1* et *C1QTL3*, qui présentent des interactions de type épistatique simultanément avec les QTL du EM1 et du EM2. Une explication possible à ce phénomène serait que ces 3 QTL peuvent avoir des effets pléiotropiques [158]. Ainsi, vu qu'ils ne présentent pas d'interactions de type additif envers les membres des EM1 et EM2, ils peuvent être classifiés comme faisant partie de l'un ou l'autre des modules et non d'un module distinct (Figure 5b.).

12.2.2 Établissement d'une hiérarchie dans l'organisation des QTL

L'existence de QTL ayant des effets opposés permet, dans un premier temps, d'établir une hiérarchie entre eux. À titre d'exemple, lors de la construction de la souche congénique C8S.L2, délimitant le *C8QTL2*, on observe que les allèles introduits causent une hausse de la PA par rapport à la souche parentale DSS. Il s'agit donc, comme dans le cas du *C3QTL2*, discuté dans la section 5.2.2, ainsi que dans celui du *C2QTL4L*, de QTL +PA (cf. section 9.8 Table 1. P.68 et section 9.10 Table 1. P.74). Lors de la construction de la souche double congénique C8S.L2/C10S.L16, contenant à la fois le *C8QTL2* et le *C10QTL2*, membres tous deux du EM2, on observe que leur effet combiné est égal à l'effet observé chez la souche C10S.L16; contenant seulement le *C10QTL2* (cf. Fig.1F, section 9.8 p.66). De cette façon, étant donné que *C10QTL2* masque l'effet de *C8QTL2*, il est possible de dire que le premier est épistatique au deuxième et qu'il est placé plus haut dans la hiérarchie régulatrice.

Plus encore, ceci permet d'identifier le type d'interaction existant entre ces deux QTL. Dans ce cas précis, on peut dire que *C8QTL2* exerce un effet inhibiteur sur *C10QTL2* (Figure 5 c., p.265). Pour illustrer cette affirmation, on se référera à la Figure 4, p. 264 qui s'explique comme suit : étant donné que lors de l'identification de QTL de PA, on compare les souches congéniques par rapport à DSS, cette souche sera définie comme étant de type sauvage (WT) et les allèles (SS, homozygotes) de cette souche comme étant « on ». Dans le même ordre d'idées, tout changement chez DSS dû à un remplacement par les allèles provenant de Lewis (LL, homozygotes) dans une congénique, sera considéré comme mutant et en état « off ». Ainsi, quand le *C10QTL2* est en position « off » (souche C10S.L16), la PA diminue par rapport à celle de DSS (flèche pointant vers le bas sur la Figure 4.), indiquant ainsi que la présence d'allèles SS au niveau de ce QTL est nécessaire à l'augmentation de la PA. Si le *C8QTL2* agissait comme activateur du *C10QTL2*, lorsque les deux seraient en position « on », on s'attendrait à voir une augmentation de la PA de base de DSS. Ce phénomène étant aberrant et non-existant, on peut admettre que *C8QTL2* exerce une régulation négative sur *C10QTL2*. En effet, lorsque *C8QTL2* est « off », son effet inhibiteur sur *C10QTL2* est annulé, permettant l'augmentation de PA observée chez la souche C8S.L2. En outre, lorsque les deux QTL sont « off », comme dans la double congénique, la PA diminue au même degré que pour

la souche C10S.L16. Ceci est dû au fait que, lorsque *C10QTL2* est LL, les allèles nécessaires à l'augmentation de la PA pour atteindre la même valeur que DSS, sont « off », et par conséquent, l'état du QTL inhibiteur devient insignifiant. Le même principe peut être appliqué aux *C3QTL1* et *C3QTL2* de EM1, où le premier est épistatique au deuxième et ce dernier agit comme inhibiteur du *C3QTL1*. Il en est de même pour *C2QTL4L* qui exerce aussi une régulation négative sur *C10QTL2*.

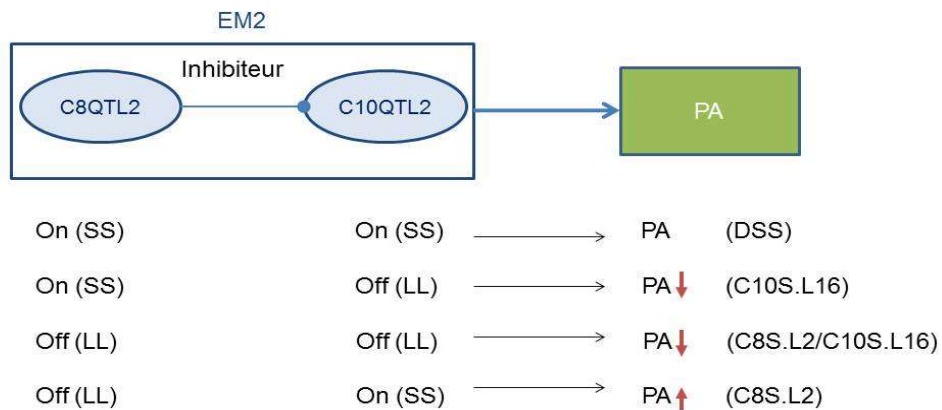


Figure 4. Schéma de la détermination hiérarchique de deux QTL par génétique épistatique. Les flèches verticales démontrent une augmentation (vers le haut) ou une diminution (vers le bas) de la PA par rapport à la souche parentale Lewis.

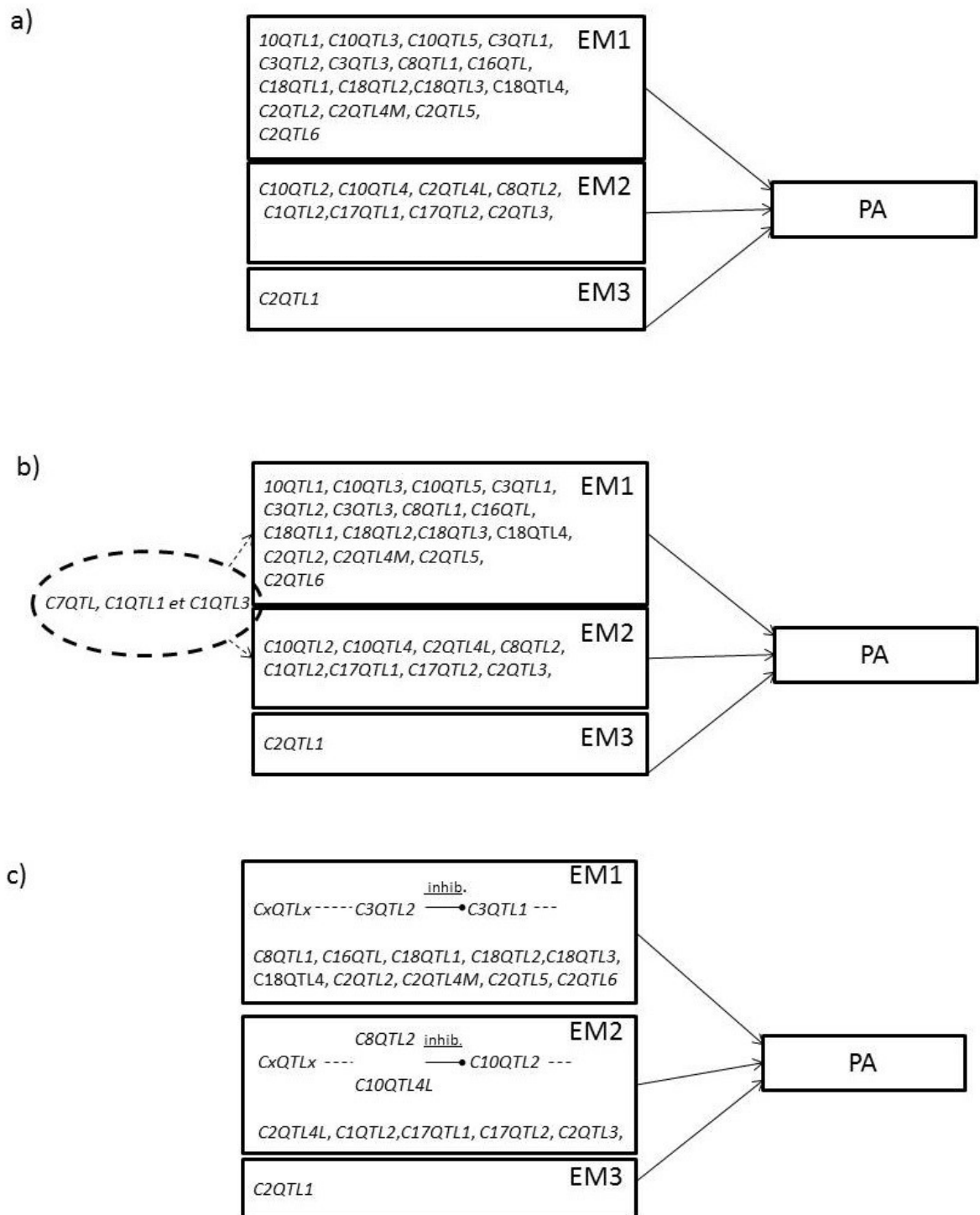


Figure 5. Schéma de la modularisation des QTL. a) Modularisation des QTL en EM. b) Schématisation des effets pléiotropiques (flèches en pointillés) des *C7QTL*, *C1QTL1* et *C1QTL3*, l'ensemble de ces QTL délimité par des pointillés dénote que ces 3QTL font partie à la fois de EM1 et EM2 et non d'un EM distinct. c)

Modularisation des QTL et établissement d'une hiérarchie entre C3QTL1 et C3QTL2, ainsi que C10QTL2 et les C8QTL2, C10QTL4L. CxQTLx peut être n'importe quel QTL du EM en question. —● indique un effet inhibiteur, - - - indique un type de régulation quelconque dont l'existence n'est pas prouvée, → indique une régulation existante.

12.3 Signification biologique de la modularisation des QTL

Comme discuté dans la section 9.4.3, les combinaisons de plus de 2 QTL contenant chacune, à la base, *C10QTL1* et *C10QTL2*, démontrent que la présence de ces deux QTL dans une même congénique exerce un « effet palier » au niveau de la PA. Ceci illustre donc que l'effet d'un autre QTL provenant de l'un ou l'autre des EM devient redondant. De plus, les QTL faisant partie d'un même module épistatique sont supposés d'agir au niveau de la même voie de régulation. Suivant cette logique, le défaut d'un des membres du EM/voie, se traduit par un blocage au niveau de la cascade, la rendant ainsi non-fonctionnelle [77, 94, 110, 159].

De cette façon, le dévoilement d'une architecture génétique expliquant le mode d'action des différents QTL permet de placer provisoirement un QTL au sein d'une voie de régulation de la PA, et ceci, sans que sa nature moléculaire et/ou sa fonction soient connues.

Néanmoins, il est à noter que l'établissement d'une hiérarchie au niveau de ces voies est possible, uniquement, lorsque les deux QTL analysés exercent des effets opposés. Le positionnement séquentiel de deux QTL ayant des effets semblables et de même ampleur demeure impossible sur la base unique d'un changement de PA. Ainsi, il est nécessaire d'analyser des phénotypes intermédiaires et/ou d'établir la nature et fonction des QTL en question.

Pour cette raison, l'identification des gènes à la base des QTL se révèle une stratégie complémentaire à la compréhension des mécanismes sur lesquels les traits complexes sont basés.

13. Identification de QTL au niveau du chromosome 18 du rat.

Depuis les années 1990, l'existence de QTL influençant la PA au niveau du chromosome 18 chez le rat a été établie par analyses de liaison génétique. L'objectif a été, au cours de la présente étude, d'identifier ces QTL et de démasquer les gènes à la base de la différence de PA [103, 108].

13.1 Analyse des régions C18QTL3 et C18QTL4

13.1.1 Cartographie fine des C18QTL

La construction des sous souches congéniques C18S.L10 et C18S.L14 a permis de déceler la présence de deux QTL distincts au niveau de la région délimitée par la souche C18S.L2, contenant l'ancien C18QTL3 (cf. Fig. 1, section 10.8, p.110). C18S.L14 délimite le nouveau C18QTL3, tandis que C18QTL4 est délimité par C18S.L10.

Étant donné les moyens techniques disponibles au moment où l'étude a été entamée et en suivant les principes de l'approche par gène candidat, les premiers gènes analysés ont été *Adrb2* et *Nedd4l*, contenus dans C18QTL3. Aucune variante se traduisant par un changement protéique ou se trouvant au niveau des régions régulatrices n'a été détectée. La région promotrice minimale de *Adrb2* est identifiée chez l'humain à 550pb en amont du codon d'initiation [160]. Cependant, certains SNPs ont été détectés jusqu'à ~1450 pb en amont du ATG [161], pour cette raison, nous avons séquencé 1481 pb dans la région 5' de ce gène.

Étant donné que la séquence promotrice de *Nedd4l* n'était pas connue, nous avons séquencé environ 1650 pb en amont du codon d'initiation, en nous basant sur les régions définies pour *Adrb2*. Plus encore, aucune différence du niveau d'expression dans les reins n'a

été détectée pour ces deux gènes. *Nedd4l* étant impliqué dans la régulation de plusieurs canaux et transporteurs ioniques rénaux comme les ENaC, NCC, NKCC2 et NH3 [147, 149, 162], l'analyse d'expression de ce gène n'a été effectuée qu'au niveau rénal. Pour des raisons de concordance, le même tissu a été utilisé pour l'étude de *Adrb2*. Il serait néanmoins intéressant d'effectuer des analyses sur d'autres organes comme le cerveau, le cœur, le foie et les poumons, étant donné que l'expression de *Nedd4l* est importante au niveau de ces organes et que cette protéine est connue pour réguler des canaux sodiques sensibles au voltage au niveau du cœur et des neurones, des canaux potassiques cardiaques et est impliqué dans la croissance des dendrites [147]. Il serait d'autant plus nécessaire d'effectuer ces analyses d'expression dans des organes autres que le rein étant donné la fonction de *Adrb2* au niveau des cellules vasculaires lisses ainsi que son implication au niveau des systèmes cardiovasculaire, musculaire et immunologique [144, 147, 163, 164].

Ensuite, l'examen détaillé de la séquence codante et des jonctions exon/intron définies à 50 pb en amont et en aval de chaque exon [165] d'autres gènes contenus dans les régions C18QTL3 et C18QTL4 a été poursuivi.

Le développement et la standardisation de techniques de séquençage des génomes complets [166] a rendu possible, postérieurement, l'utilisation des génomes DSS et Lewis comme base de données. De cette façon, non seulement nous avons réussi à identifier toutes les différences de séquence présentes dans les gènes contenus dans les QTL, mais aussi à réduire les régions délimitées par les souches en utilisant des SNPs comme marqueurs génétiques. Par conséquent, nous avons réduit le nombre de gènes contenus au niveau des C18QTL3, C18QTL4 et C18QTL2.

13.1.2 Identification de gènes candidats

En vue d'établir un gène comme candidat à la base d'un QTL celui-ci doit présenter au moins une variante qui affecte sa fonction; ces variantes peuvent affecter la séquence protéique encodée, l'épissage ou le niveau d'expression du gène. Ce dernier peut aussi présenter un nombre variable de copies (CNV) [94, 167].

Étant donné les résultats des prédictions de l'impact des variantes au niveau des gènes contenus dans les régions C18QTL, *Tcofl* et *Hdhd2* ont été retenus comme candidats prédominants pour les *C18QTL3* et *C18QTL4*, respectivement. Bien que le gène *Alpk2* ait été considéré comme candidat prédominant avant d'avoir séquencé tous les gènes contenus dans la région (chapitre 10), les variantes observées, sont prédites comme étant non-délétères. *Tcofl* et *Alpk2* présentent des délétions se traduisant par la perte de un ou plusieurs acides aminés pouvant affecter la fonction des protéines encodées. L'impact de la délétion observée dans la séquence de *Tcofl* serait néanmoins plus important étant donné que les acides aminés manquants correspondent à un des domaines fonctionnels de la protéine [168]. *Hdhd2* est retenu comme candidat prédominant pour le *C18QTL4* étant donné que les résultats présentés dans cette étude suggèrent que la substitution NM_001014151.2(*Hdhd2*):c.337G>C, p.(113Val>Leu) altérerait l'affinité de *Hdhd2* pour le DiFMUP.

Il est à noter qu'au cours des études effectuées, de nombreuses variantes se trouvant au niveau de séquences régulatrices et, par conséquent, pouvant affecter l'expression ou l'épissage des gènes présents dans les régions C18QTL, n'ont pas été analysées. Tout en étudiant ce type de variantes, il serait intéressant d'analyser les profils d'expression des gènes entre DSS et Lewis, ou entre DSS et les souches congéniques au niveau de différents tissus [169-171].

En effet, l'analyse des gènes dans les régions contenant les C18QTL s'est limitée à l'évaluation des variantes codantes et des jonctions intron/exon. Bien que souvent rares, les variantes présentes dans les régions codantes des gènes ont un impact important sur les formes Mendéliennes d'hypertension [54], ainsi que dans des traits complexes.

Cependant, les variantes non codantes, qui sont souvent négligées, peuvent avoir des conséquences fonctionnelles sur les gènes avoisinants [156]. Étant donné que l'on dispose à présent des séquences complètes des génomes de DSS et Lewis il serait intéressant d'analyser ce type de variantes. En effet, on pourrait, grâce à l'existence de bases de données telles que ENCODE [172, 173], GWAVA[174] ou GeneVar, étudier l'implication de cette classe de variantes dans la différence du phénotype observé et ce en recherchant l'existence de: (i) QTL d'expression (eQTL) [175]; (ii) QTL d'épissage (sQTL) [176]; (iii) marqueurs de régulation épigénétiques; pouvant avoir des effets sur l'expression de gènes pertinents pour la régulation

de la PA. Plusieurs techniques sont à présent disponibles pour analyser cette classe de variantes; la méthode de ChIP-seq permet d'établir l'existence d'éléments régulateurs de l'expression génique comme des activateurs ou suppresseurs de l'activité promotrice des gènes [175]. L'analyse des profils d'expression dans des tissus pertinents couplé au séquençage de l'ARN permet aussi de confirmer la présence d'un eQTL ayant un effet sur l'expression des gènes ou des ARN non codants [156]. Plus encore, il existe à présent des techniques permettant de mettre en évidence des interactions régulatrices à l'échelle du génome comme c'est le cas de la capture de la conformation de la chromatine (chromatin conformation capture, 3C) ou de la technique *Chromatin interaction analysis by paired-end-tag sequencing* (ChIA-PET) [174, 177].

A la lumière de la discussion précédente, on ne peut pas exclure la possibilité qu'une composante génétique autre que *Tcof1* et *Hdhd2* soit à la base d'un QTL dans cette région et ait une influence sur la PA. En effet, il serait nécessaire d'isoler chacun des gènes retenus comme candidats et de tester leur effet grâce à des techniques d'édition du génome par des expériences de knock-in ou de transgénèse [110, 178, 179].

13.1.2.1 QTL influençant la PA et gènes liés à sa régulation

Dans les cas monogéniques d'hypertension, ce sont les gènes ayant des fonctions physiologiques directement liées à l'homéostasie de la PA qui sont touchés [54]. De même, des changements dans la fonction de ces derniers par leur inactivation ou encore par leur surexpression, cause des effets directs sur la PA. Cependant, même s'ils résident au niveau d'une région QTL, ils ne sont pas considérés comme gènes candidats s'ils ne possèdent pas des variantes de séquence significatives pouvant aboutir à des changements de fonction. À titre d'exemple, l'inactivation du gène *Adm* contenu dans des régions QTL chez le rat et chez l'humain, se traduit par une variation de la PA [43, 100, 180]. Cependant, comme discuté auparavant (section 9.4.4, p.55), aucune différence de séquence n'a été détectée entre DSS et Lewis et ce gène n'est donc pas retenu comme candidat. Par le même raisonnement, *Adrb2* et *Nedd4l* n'ont pas été retenus comme candidats pour le *C18QTL3*. À noter que, grâce au séquençage des génomes complets de DSS et Lewis, nous avons en plus pu corroborer qu'aucun de ces deux gènes n'est présent en nombre différents de copies (CNV) entre ces

deux souches. Ainsi, un QTL influençant la PA peut être différent d'un gène dont les effets sur la PA sont directs.

Par ailleurs, les candidats des *C18QTL4* et *C18QTL3*, *Hdhd2* et *Tcofl* ne sont pas connus pour avoir une influence directe sur la pression sanguine. Étant donné l'importance des voies de signalisation par phosphorylation/déphosphorylation dans la cellule [181], il est envisageable de penser que ces candidats agissent en amont d'autres facteurs d'une voie qui aboutit la modulation du produit d'un gène directement lié au contrôle de la PA.

13.2 Modularisation de *C18QTL3* et *C18QTL4* et étude de leur interaction

La souche C18S.L2 délimite l'ancien C18QTL3, région contenant le nouveau C18QTL3 et le C18QTL4. Ainsi, C18S.L2 peut être considérée comme la double congénique nécessaire à l'étude de l'interactions des *C18QTL3* et *C18QTL4* (Figure 6.). Les *C18QTL1*, *C18QTL2* et ancien *C18QTL3* sont épistatiques entre eux [108]. Plus encore, les résultats présentés dans le chapitre 9 démontrent que *C18QTL2* appartient à EM1. Par extrapolation de ces données, on peut conclure que *C18QTL3* et *C18QTL4* sont épistatiques et appartiennent à EM1.

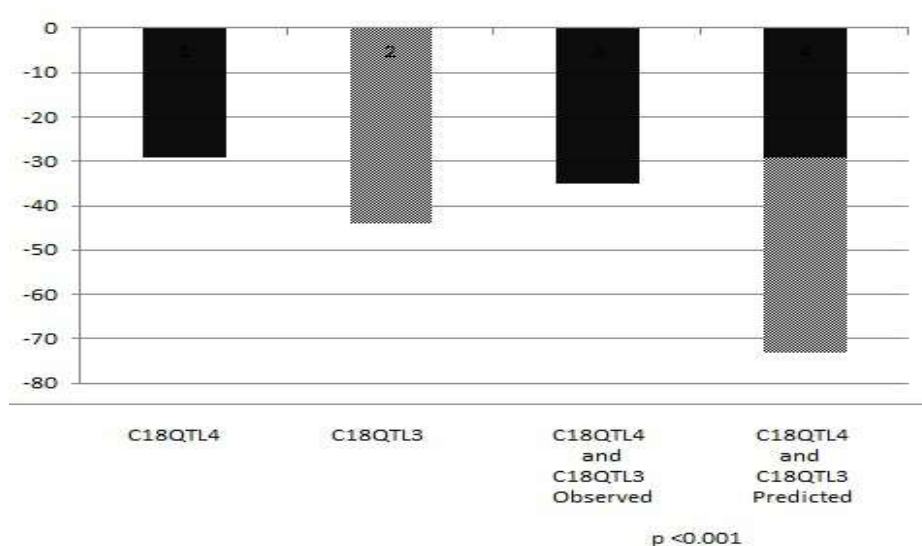


Figure 6. Analyse de l'interaction épistatique entre C18QTL3 et C18QTL4. Les 3 premières barres représentent la différence de MAP observées expérimentalement entre la congénique et DSS (définie à valeur 0). La 4ème barre représente l'effet additif prédit des 2 QTL analysés. C18QTL3 est délimité par C18S.L14, C18QTL3 par C18S.L10 et C18QTL3+C18QTL4 est représenté par la souche C18S.L2.

La suite logique de l'étude est donc d'analyser les bases moléculaires de l'interaction épistatique entre *C18QTL3* et *C18QTL4*. Étant donné que Hdh2 fait partie d'une superfamille de protéines encodant des phosphatases et que Hdh1 est connue pour en être une [182], nous avons testé l'existence de cette activité enzymatique chez Hdh2. Cependant, l'activité de cette protéine nécessite d'une caractérisation plus poussée; certaines méthodes pour atteindre cet objectif sont décrites dans la section Annexes de ce document.

TCOF1, impliquée dans le syndrome de Treacher Collins, est une protéine nucléaire polyphosphorylée impliqué dans la biogénèse ribosomale [183]. Les résultats présentés dans la section 11.4.8, suggèrent une interaction entre Hdh2 et Tcof1.

Cependant, les résultats présentés dans ce document ne représentent qu'une première approche. En effet, il serait nécessaire de prouver l'interaction de ces deux protéines par des expériences plus complètes et concluantes. Tout d'abord, il serait nécessaire d'effectuer les tests avec des protéines purifiées par des techniques d'immunoaffinité; de cette façon, le niveau de phosphorylation de Tcof1 pourrait être analysé grâce à l'utilisation de méthodes

d'immunovubardages de type « *western* » ou ELISA avec des anticorps phospho-spécifiques [184].

Cependant, d'après les expériences réalisées, nous proposons que Hdhd2 déphosphoryle Tcof1 *in vitro*. L'identification de la nature moléculaire de ces QTL rend possible l'analyse de leur interaction. Bien que leur interaction reste à confirmer *in vivo*, les analyses présentées font preuve, pour la première fois, d'une interaction fonctionnelle entre 2 QTL prédite par le modèle de modularisation des QTL chez le rat *Dahl Salt-Sensitive*. Ainsi, étant donné que Hdhd2 agirait en amont de Tcof1 dans la voie dans laquelle ils sont impliqués, une nouvelle hiérarchie dans l'organisation de l'architecture du contrôle de la PA est proposée (Figure 7.).

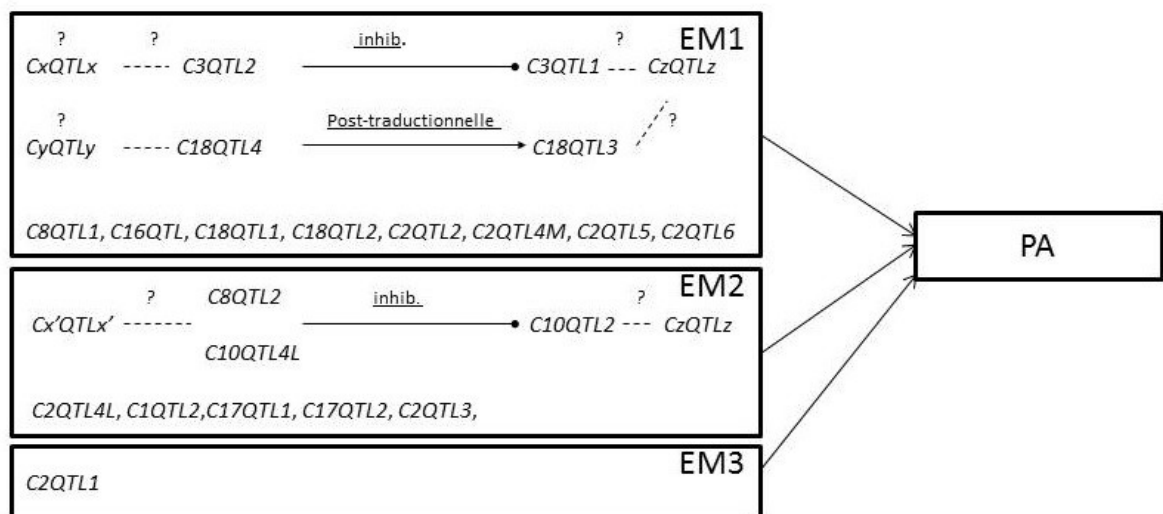


Figure 7. Schéma récapitulatif de modularisation des QTL et des hiérarchies existantes. CxQTLx, CyQTLy, CzQTLz et Cx'QTLx' peuvent être n'importe quel QTL du EM en question. —● indique un effet inhibiteur, - - - indique un type de régulation quelconque dont l'existence n'est pas prouvée (?), → indique une régulation existante.

14. Implications et applications chez l'humain

Bien que plusieurs mécanismes moléculaires à la base des formes monogéniques de l'hypertension soient établis, l'étiologie de l'hypertension essentielle demeure énigmatique. La stratégie de choix pour identifier les variations causant l'apparition de cette pathologie est l'application des *Genome-Wide Association Studies* (GWAS) [1, 41, 57, 58, 62, 100]. Bien que cette approche soit viable dans la découverte de variantes causant l'hypertension, elle comporte quelques bémols. Tout d'abord, les associations ne permettent pas de détecter directement le gène à la base de la variation de PA, étant donné que les SNPs sont souvent dans des régions intergéniques. Ensuite, le seuil de détection est établi à $MAF > 5\%$, ce qui exclut systématiquement des SNPs pouvant être pertinents à la PA dans une sous-population. Plus encore, étant donné qu'il s'agit d'études statistiques, la possibilité d'établir un lien de cause à effet est faible [62, 185]. Ces limitations peuvent être surmontées grâce à l'utilisation de modèles animaux pour l'identification de QTL.

14.1 Effets de l'hétérogénéité génétique sur l'identification de QTL

D'après l'ensemble de l'étude exposée jusqu'à maintenant, on admet que les QTL sont organisés en modules épistatiques chez le modèle DSS. Étant donné qu'il s'agit d'une population consanguine, il est difficile d'extrapoler cette organisation directement chez une population plus hétérogène et d'autant plus chez l'humain. Il faut noter aussi, que s'agissant de conclusions tirées à partir de l'utilisation de modèles animaux, les effets de l'environnement sont modulés et ne sont pas pris en considération. Plus encore, la seule façon de déceler l'existence d'une organisation en modules est d'isoler chaque QTL pour ensuite étudier leurs interactions en paires, tâche qui n'est pas réalisable au sein des populations humaines.

En outre, il est intéressant de remarquer que le caractère monogénique des QTL n'est pas toujours apparent lorsque le fond génétique dans lequel il s'exprime est hétérogène. En effet, le C18QTL1 détecté par congéniques sur fond DSS représente 48% de la différence de

PA entre les souches parentales [108]. Cependant, l'effet observé de ce même QTL dans un fond non homogène ne représente que 7.5% de la variabilité détectée [103].

De cette façon, nous mettons de l'avant le fait que les approches statistiques dans l'étude de traits complexes chez l'humain, couplées aux approches expérimentales dans des modèles animaux peuvent démasquer des QTL de PA qui ne seraient pas identifiables autrement.

14.2 Relation entre les QTL influençant la PA chez l'humain et les QTL de PA chez DSS

Comme il a été discuté dans la section 9.4.4, certains QTL détectés chez l'humain sont retrouvés dans le modèle DSS. C'est le cas, par exemple, de FLJ32810-TMEM133, qui étant découvert par GWAS [100], est retrouvé dans la région délimitée par la souche C8S.L2. Comme exposé auparavant, *C8QTL2* inhibe *C10QTL2* dans une cascade encore inconnue. Le gène candidat prédominant pour *C10QTL2* code pour une protéine de la famille des *ATP-binding cassettes, subfamily A* (ABCA); cette classe de protéines n'étant pas associées à la régulation de la PA, on peut déduire qu'il ne s'agit pas d'un QTL directement lié au trait. Ainsi, au niveau de la cascade de régulation dans laquelle il est impliqué, *C10QTL2* est sûrement en amont d'autres QTL. À noter le QTL défini par *NPR3* chez l'humain est homologue au *C2QTL4L*, régulant lui aussi *C10QTL2*. Ainsi, l'analyse plus approfondie de l'existence d'une relation hiérarchique et régulatrice entre les QTL *FLJ32810-TMEM133* et *NPR3* pourrait se révéler très prometteuse. En suivant ce même ordre d'idées, il serait intéressant d'étudier les interactions entre les QTL *JAG1*, *EDN3* et *CSK*, étant donné que leurs QTL homologues chez DSS font tous partie de EM1.

14.3 Pertinence de la modularisation des QTL

Même si la relation entre les QTL humains discutés dans la section précédente nécessite d'une confirmation expérimentale, la modularisation des QTL chez DSS permet d'ouvrir la voie vers une compréhension plus approfondie et globale de l'architecture à la base de l'hypertension essentielle. La mise en évidence de l'existence de modules épistatiques chez DSS permet, néanmoins, de positionner provisoirement un QTL humain dans une voie de régulation; si et seulement si, il existe un QTL homologue chez la souche murine. Ceci est envisageable même si la nature et fonction du QTL restent à déterminer. Plus encore, la modularisation des QTL permettra peut-être d'améliorer les traitements médicaux étant donné qu'on pourra cibler deux cascades distinctes de façon à contrôler la tension plus efficacement.

Réciproquement, l'identification du gène responsable du QTL chez le rat ainsi que l'élucidation de sa fonction et/ou de son interaction avec un autre QTL, permet d'agrandir le nombre de variantes à tester lors de l'utilisation des GWAS. C'est le cas des *C18QTL3/Tcof1* et *C18QTL4/Hdhd2*; QTL qui pourraient être inclus dans les analyses d'association chez l'humain.

Conclusion

La découverte des bases génétiques de l'hypertension essentielle, facteur majeur de risque des maladies cardiovasculaires, est actuellement un défi très important à relever. Étant donné le caractère polygénique du trait et l'hétérogénéité des populations humaines, l'utilisation des modèles animaux dans l'étude de cette pathologie est une stratégie de choix.

L'identification de QTL ayant une influence sur la PA par construction de souches congéniques se comportant de façon monogénique et l'évidence de l'existence d'interactions épistatiques dans le génome ont été les outils et les moyens par lesquels on a regroupé la presque totalité des QTL en modules épistatiques (EM). Ces EM sont considérés comme les « unités » fondamentales de l'architecture génétique contrôlant la PA. Les QTL faisant partie du même EM sont organisés en cascades séquentielles et hiérarchiques dont la majorité reste inconnue; ceci étant dû, en grande partie, au fait que les gènes candidats pour chaque QTL sont loin d'être des gènes de fonction physiologique proche de la régulation de la PA.

Bien que le dévoilement de cette modularité permette de placer provisoirement un QTL au sein d'une cascade donnée, l'identification de sa nature et fonction moléculaires, ainsi que la façon dans laquelle il interagit avec le reste des QTL, demeure indispensable.

L'identification de la nature moléculaire de *C18QTL4/Hdhd2* et *C18QTL3/Tcofl* faisant partie tous les deux du même module épistatique, dont la présente étude suggère une interaction de Hdhd2 et Tcofl, se révèle être la première évidence expérimentale prédite par la modularisation des QTL.

Bien que l'interaction de ces deux QTL reste à prouver *in vivo*, la découverte de nouvelles composantes dans l'organisation du maintien de la PA rend possible la diversification des cibles des études chez l'humain. La modularisation des QTL chez le rat trace la voie pour la découverte de nouvelles cibles et de nouveaux traitements thérapeutiques plus efficaces, basés sur des voies de régulation hiérarchisées.

Annexes

1. Analyse de l'activité d'une phosphatase et identification du substrat

De nos jours, la technique la plus répandue pour caractériser l'activité phosphatase d'une protéine donnée est l'utilisation du *p*-nitrophenyl phosphate (*p*NPP). Lors de l'hydrolyse de ce substrat artificiel par une phosphatase, un produit chromogénique avec absorbance à 405 nm, le *p*-nitrophenol, est créé, permettant ainsi de mesurer l'activité de la phosphatase étudiée [186]. Le *p*NPP est un substrat de choix étant donné que son taux d'hydrolyse non-enzymatique est très faible et ce, même à des hautes températures. De plus, la majorité des phosphatases présentent une activité significative envers ce substrat à pH neutre. Cette technique comporte plusieurs avantages sur la méthode radioactive, étant donné le danger associé aux manipulations du matériel radioactif et le fait que les concentrations de substrat utilisées peuvent être supérieures au K_m . La majorité des sérine/thréonine phosphatases nécessitent de la présence d'un cation divalent pour leur activité. Ainsi, on peut analyser l'activité de la phosphatase étudiée en présence d'ions métalliques tels que le Mn^{2+} et Mg^{2+} , entre autres. Par ailleurs, il est nécessaire de déterminer aussi la température et le pH optimal pour l'activité enzymatique, paramètres qui peuvent être analysés en effectuant les réactions à différentes températures et en utilisant différentes solutions tampons. Plus encore, cette technique permet d'étudier l'effet des inhibiteurs de phosphatases [187, 188]. Une alternative pour la détermination de l'activité d'une phosphatase est la méthode « in-gel ». Dans cette approche, une molécule marquée au ^{32}P est ajoutée au gel de polyacrylamide et suite la séparation par SDS-PAGE des protéines contenues dans un échantillon, on peut détecter par autoradiographie les sites de perte du radio-isotope. Cette procédure, couplée à des immunobuvardages de type « western » permettent d'identifier l'activité phosphatase de la protéine d'intérêt. Des variantes de cette technique incluent l'utilisation de molécules fluorogéniques comme le DiFMUP qui permettent la coloration du gel après la migration .

En vue d'identifier le substrat d'une protéine phosphatase, plusieurs approches peuvent être utilisées. Tout d'abord, un balayage systématique peut être effectué en utilisant des

molécules comme des sucres et de nucléotides ainsi que des peptides phosphorylés. L'activité de l'enzyme envers les différents substrats peut être détectée grâce à l'addition du vert de malachite [188, 189]. Il existe aussi des puces à peptides phosphorylés qui permettent de détecter leur niveau de déphosphorylation suite à des colorations fluorescentes, comme le « Pro-Q[®] Diamond » (Molecular Probes[®], Invitrogen) [190].

De nos jours, des efforts sont faits pour répertorier les substrats potentiels de certaines phosphatases grâce à des analyses informatiques qui utilisent des techniques de prédiction par d'homologie de séquence, par exemple. Cependant, le manque d'information concernant principalement les sérine/thréonine phosphatases rend la tâche difficile [191]. Lorsqu'une première analyse démontre une activité de la phosphatase envers un ou plusieurs peptides phosphorylés, on peut effectuer un marquage isotopique des protéines dans un lysat cellulaire en présence ou absence de la phosphatase en question, suivi d'une électrophorèse bidimensionnelle sur gel à fin de comparer les profils de phosphorylation. Les protéines ayant été déphosphorylées peuvent être analysées par spectrométrie de masse. De même, on peut effectuer des analyses de spectrométrie de masse en tandem couplée à la chromatographie liquide (LC-MS/MS); de cette façon on peut comparer les phosphoprotéines présentes dans un échantillon traité avec la phosphatase en question vs. l'échantillon non traité [192, 193]. Dans certains cas, la validation du substrat d'une phosphatase peut se faire *in vitro* par la technique de « substrate-trapping ». En effet, cette méthode repose sur la co-précipitation de deux protéines mais étant donné que la formation du complexe entre une enzyme-substrat est transitoire, certains mutants de la phosphatase peuvent être créés en vue d'abolir son activité catalytique tout en gardant intact le site de liaison du substrat. Ce type de technique d'identification de substrat est très répandu dans l'étude des tyrosine-phosphatases [184]. Une approche alternative *in vivo* est l'utilisation de double-hybrides chez la levure où la protéine d'intérêt est fusionnée au domaine de liaison de l'ADN d'un activateur de transcription, comme le GAL4, et le substrat à son domaine d'activation. L'interaction de la phosphatase avec le substrat se traduit alors par la transcription de gènes rapporteurs sous contrôle de GAL4. Cette technique est aussi utilisée pour un balayage de substrats possibles en utilisant des bibliothèques de cDNA et elle requiert souvent de l'activation conditionnelle d'une kinase pour assurer la phosphorylation des protéines encodées [194].

Bibliographie

1. Lifton, R.P., A.G. Gharavi, and D.S. Geller, *Molecular mechanisms of human hypertension*. Cell, 2001. **104**(4): p. 545-56.
2. Cowley, A.W., Jr., *The genetic dissection of essential hypertension*. Nat Rev Genet, 2006. **7**(11): p. 829-40.
3. Sesso, H.D., et al., *Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in Men*. Hypertension, 2000. **36**(5): p. 801-7.
4. Wallin, B.G. and N. Charkoudian, *Sympathetic neural control of integrated cardiovascular function: insights from measurement of human sympathetic nerve activity*. Muscle Nerve, 2007. **36**(5): p. 595-614.
5. Guyenet, P.G., *The sympathetic control of blood pressure*. Nat Rev Neurosci, 2006. **7**(5): p. 335-46.
6. Pilowsky, P.M. and A.K. Goodchild, *Baroreceptor reflex pathways and neurotransmitters: 10 years on*. J Hypertens, 2002. **20**(9): p. 1675-88.
7. Thomas, G.D., *Neural control of the circulation*. Adv Physiol Educ, 2011. **35**(1): p. 28-32.
8. Fisher, J.P. and J.F. Paton, *The sympathetic nervous system and blood pressure in humans: implications for hypertension*. J Hum Hypertens, 2012. **26**(8): p. 463-75.
9. Parati, G. and M. Esler, *The human sympathetic nervous system: its relevance in hypertension and heart failure*. Eur Heart J, 2012. **33**(9): p. 1058-66.
10. Bassi, M., et al., *Control of respiratory and cardiovascular functions by leptin*. Life Sci, 2015. **125C**: p. 25-31.
11. Donoghue, M., et al., *A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9*. Circ Res, 2000. **87**(5): p. E1-9.
12. Fyhrquist, F. and O. Saijonmaa, *Renin-angiotensin system revisited*. J Intern Med, 2008. **264**(3): p. 224-36.
13. Bjorkqvist, J., A. Jamsa, and T. Renne, *Plasma kallikrein: the bradykinin-producing enzyme*. Thromb Haemost, 2013. **110**(3): p. 399-407.
14. Kaschina, E. and T. Unger, *Angiotensin AT1/AT2 receptors: regulation, signalling and function*. Blood Press, 2003. **12**(2): p. 70-88.
15. Mendoza, A. and E. Lazartigues, *The compensatory renin-angiotensin system in the central regulation of arterial pressure: new avenues and new challenges*. Ther Adv Cardiovasc Dis, 2015.
16. Romero, C.A., M. Orias, and M.R. Weir, *Novel RAAS agonists and antagonists: clinical applications and controversies*. Nat Rev Endocrinol, 2015. **11**(4): p. 242-252.
17. Sumners, C., et al., *Angiotensin type 2 receptors: blood pressure regulation and end organ damage*. Curr Opin Pharmacol, 2015. **21**: p. 115-121.
18. Chen, D. and T.M. Coffman, *AT Angiotensin receptors-vascular and renal epithelial pathways for blood pressure regulation*. Curr Opin Pharmacol, 2015. **21**: p. 122-126.
19. Paul, M., A. Poyan Mehr, and R. Kreutz, *Physiology of local renin-angiotensin systems*. Physiol Rev, 2006. **86**(3): p. 747-803.
20. Moorhouse, R.C., et al., *Endothelin antagonism and its role in the treatment of hypertension*. Curr Hypertens Rep, 2013. **15**(5): p. 489-96.

21. Rapoport, R.M., *Nitric oxide inhibition of endothelin-1 release in the vasculature: in vivo relevance of in vitro findings*. Hypertension, 2014. **64**(5): p. 908-14.
22. Deng, A.Y., *Is the nitric oxide system involved in genetic hypertension in Dahl rats?* Kidney Int, 1998. **53**(6): p. 1501-11.
23. Horita, S., et al., *Regulatory roles of nitric oxide and angiotensin II on renal tubular transport*. World J Nephrol, 2014. **3**(4): p. 295-301.
24. Hinson, J.P., S. Kapas, and D.M. Smith, *Adrenomedullin, a multifunctional regulatory peptide*. Endocr Rev, 2000. **21**(2): p. 138-67.
25. Wong, H.K., T.T. Cheung, and B.M. Cheung, *Adrenomedullin and cardiovascular diseases*. JRSM Cardiovasc Dis, 2012. **1**(5).
26. Wong, H.K., et al., *Adrenomedullin and diabetes*. World J Diabetes, 2014. **5**(3): p. 364-71.
27. Regoli, D. and F. Gobeil, Jr., *Critical insights into the beneficial and protective actions of the kallikrein-kinin system*. Vascul Pharmacol, 2015. **64**: p. 1-10.
28. Kawabe, J., F. Ushikubi, and N. Hasebe, *Prostacyclin in vascular diseases. - Recent insights and future perspectives*. Circ J, 2010. **74**(5): p. 836-43.
29. Hamlyn, J.M., *Natriuretic hormones, endogenous ouabain, and related sodium transport inhibitors*. Front Endocrinol (Lausanne), 2014. **5**: p. 199.
30. Hodes, A. and D. Lichtstein, *Natriuretic hormones in brain function*. Front Endocrinol (Lausanne), 2014. **5**: p. 201.
31. Mullins, L.J., M.A. Bailey, and J.J. Mullins, *Hypertension, kidney, and transgenics: a fresh perspective*. Physiol Rev, 2006. **86**(2): p. 709-46.
32. Guyton, A.C., *Blood pressure control--special role of the kidneys and body fluids*. Science, 1991. **252**(5014): p. 1813-6.
33. Castaneda-Bueno, M., J.P. Arroyo, and G. Gamba, *Independent regulation of Na⁺ and K⁺ balance by the kidney*. Med Princ Pract, 2012. **21**(2): p. 101-14.
34. Gamba, G., *Regulation of the renal Na⁺-Cl⁻ cotransporter by phosphorylation and ubiquitylation*. Am J Physiol Renal Physiol, 2012. **303**(12): p. F1573-83.
35. Luft, F.C., *Molecular genetics of human hypertension*. J Hypertens, 1998. **16**(12 Pt 2): p. 1871-8.
36. Dahl, L.K., M. Heine, and L. Tassinari, *Effects of chronic excess salt ingestion. Evidence that genetic factors play an important role in susceptibility to experimental hypertension*. J Exp Med, 1962. **115**: p. 1173-90.
37. Hamlyn, J.M. and M.P. Blaustein, *Salt sensitivity, endogenous ouabain and hypertension*. Curr Opin Nephrol Hypertens, 2013. **22**(1): p. 51-8.
38. Farquhar, W.B., et al., *Dietary Sodium and Health: More Than Just Blood Pressure*. J Am Coll Cardiol, 2015. **65**(10): p. 1042-1050.
39. Hamilton, B.P. and M.P. Blaustein, *Molecular mechanisms linking sodium to hypertension: report of a symposium*. J Investig Med, 2006. **54**(2): p. 86-94.
40. Cloutier, L., et al., *World Health Organization celebrates World Health Day, April 7, 2013--focusing on hypertension*. Can J Cardiovasc Nurs, 2013. **23**(2): p. 9-11.
41. Munroe, P.B., M.R. Barnes, and M.J. Caulfield, *Advances in blood pressure genomics*. Circ Res, 2013. **112**(10): p. 1365-79.
42. Roger, V.L., et al., *Heart disease and stroke statistics--2011 update: a report from the American Heart Association*. Circulation, 2011. **123**(4): p. e18-e209.

43. Ehret, G.B. and M.J. Caulfield, *Genes for blood pressure: an opportunity to understand hypertension*. Eur Heart J, 2013. **34**(13): p. 951-61.
44. Franceschini, N., et al., *Genetics, ancestry, and hypertension: implications for targeted antihypertensive therapies*. Curr Hypertens Rep, 2014. **16**(8): p. 461.
45. Chen, S., *Essential hypertension: perspectives and future directions*. J Hypertens, 2012. **30**(1): p. 42-5.
46. Kelly, T.N. and J. He, *Genomic epidemiology of blood pressure salt sensitivity*. J Hypertens, 2012. **30**(5): p. 861-73.
47. Rimoldi, S.F., U. Scherrer, and F.H. Messerli, *Secondary arterial hypertension: when, who, and how to screen?* Eur Heart J, 2014. **35**(19): p. 1245-54.
48. Vehaskari, V.M., *Heritable forms of hypertension*. Pediatr Nephrol, 2009. **24**(10): p. 1929-37.
49. Wadei, H.M. and S.C. Textor, *The role of the kidney in regulating arterial blood pressure*. Nat Rev Nephrol, 2012. **8**(10): p. 602-9.
50. Sica, D.A., *Endocrine causes of secondary hypertension*. J Clin Hypertens (Greenwich), 2008. **10**(7): p. 534-40.
51. Velez, D.A., M.R. Mayberg, and W.H. Ludlam, *Cyclic Cushing syndrome: definitions and treatment implications*. Neurosurg Focus, 2007. **23**(3): p. E4; discussion E4a.
52. Sukor, N., *Endocrine hypertension--current understanding and comprehensive management review*. Eur J Intern Med, 2011. **22**(5): p. 433-40.
53. Grossman, E. and F.H. Messerli, *Drug-induced hypertension: an unappreciated cause of secondary hypertension*. Am J Med, 2012. **125**(1): p. 14-22.
54. Simonetti, G.D., M.G. Mohaupt, and M.G. Bianchetti, *Monogenic forms of hypertension*. Eur J Pediatr, 2012. **171**(10): p. 1433-9.
55. Moraitis, A.G., W.E. Rainey, and R.J. Auchus, *Gene mutations that promote adrenal aldosterone production, sodium retention, and hypertension*. Appl Clin Genet, 2013. **7**: p. 1-13.
56. Furgeson, S.B. and S. Linas, *Mechanisms of type I and type II pseudohypoaldosteronism*. J Am Soc Nephrol, 2010. **21**(11): p. 1842-5.
57. Wilson, F.H., et al., *Human hypertension caused by mutations in WNK kinases*. Science, 2001. **293**(5532): p. 1107-12.
58. Boyden, L.M., et al., *Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities*. Nature, 2012. **482**(7383): p. 98-102.
59. Louis-Dit-Picard, H., et al., *KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron*. Nat Genet, 2012. **44**(4): p. 456-60, S1-3.
60. Verouti, S.N., et al., *Regulation of blood pressure and renal function by NCC and ENaC: lessons from genetically engineered mice*. Curr Opin Pharmacol, 2015. **21**: p. 60-72.
61. McBride, M.W., et al., *Functional genomics in hypertension*. Curr Opin Nephrol Hypertens, 2006. **15**(2): p. 145-51.
62. Padmanabhan, S., C. Newton-Cheh, and A.F. Dominiczak, *Genetic basis of blood pressure and hypertension*. Trends Genet, 2012. **28**(8): p. 397-408.
63. Leow, M.K., *Environmental origins of hypertension: phylogeny, ontogeny and epigenetics*. Hypertens Res, 2015.

64. Brinson, K.N., O. Rafikova, and J.C. Sullivan, *Female sex hormones protect against salt-sensitive hypertension but not essential hypertension*. *Am J Physiol Regul Integr Comp Physiol*, 2014. **307**(2): p. R149-57.
65. Byrd, J.B. and R.D. Brook, *A critical review of the evidence supporting aldosterone in the etiology and its blockade in the treatment of obesity-associated hypertension*. *J Hum Hypertens*, 2014. **28**(1): p. 3-9.
66. Mule, G., et al., *Metabolic syndrome in hypertensive patients: An unholy alliance*. *World J Cardiol*, 2014. **6**(9): p. 890-907.
67. Mallamaci, F., et al., *A genetic marker of uric acid level, carotid atherosclerosis, and arterial stiffness: a family-based study*. *Am J Kidney Dis*, 2015. **65**(2): p. 294-302.
68. Hales, C.N. and D.J. Barker, *Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis*. *Diabetologia*, 1992. **35**(7): p. 595-601.
69. Alikhani-Koopaei, R., et al., *Epigenetic regulation of 11 beta-hydroxysteroid dehydrogenase type 2 expression*. *J Clin Invest*, 2004. **114**(8): p. 1146-57.
70. Riviere, G., et al., *Epigenetic regulation of somatic angiotensin-converting enzyme by DNA methylation and histone acetylation*. *Epigenetics*, 2011. **6**(4): p. 478-89.
71. Yu, Z., Q. Kong, and B.C. Kone, *Aldosterone reprograms promoter methylation to regulate alphaENaC transcription in the collecting duct*. *Am J Physiol Renal Physiol*, 2013. **305**(7): p. F1006-13.
72. Zhang, L.N., et al., *Lower ADD1 gene promoter DNA methylation increases the risk of essential hypertension*. *PLoS One*, 2013. **8**(5): p. e63455.
73. Mu, S., et al., *Epigenetic modulation of the renal beta-adrenergic-WNK4 pathway in salt-sensitive hypertension*. *Nat Med*, 2011. **17**(5): p. 573-80.
74. Wang, J., et al., *Hypertensive epigenetics: from DNA methylation to microRNAs*. *J Hum Hypertens*, 2015.
75. Sivo, Z., et al., *Accelerated congenics for mapping two blood pressure quantitative trait loci on chromosome 10 of Dahl rats*. *J Hypertens*, 2002. **20**(1): p. 45-53.
76. Nabika, T., Y. Kobayashi, and Y. Yamori, *Congenetic rats for hypertension: how useful are they for the hunting of hypertension genes?* *Clin Exp Pharmacol Physiol*, 2000. **27**(4): p. 251-6.
77. Deng, A.Y., *In search of hypertension genes in Dahl salt-sensitive rats*. *J Hypertens*, 1998. **16**(12 Pt 1): p. 1707-17.
78. McBride, M.W., et al., *Functional genomics in rodent models of hypertension*. *J Physiol*, 2004. **554**(Pt 1): p. 56-63.
79. Rapp, J.P., *Genetic analysis of inherited hypertension in the rat*. *Physiol Rev*, 2000. **80**(1): p. 135-72.
80. Okamoto, K. and K. Aoki, *Development of a strain of spontaneously hypertensive rats*. *Jpn Circ J*, 1963. **27**: p. 282-93.
81. Dupont, J., et al., *Selection of three strains of rats with spontaneously different levels of blood pressure*. *Biomedicine*, 1973. **19**(1): p. 36-41.
82. Bianchi, G., U. Fox, and E. Imbasciati, *The development of a new strain of spontaneously hypertensive rats*. *Life Sci*, 1974. **14**(2): p. 339-47.
83. Belledonne, M., J.M. Preuss, and H.G. Preuss, *Acid excretion in young and adult Wistar Kyoto and spontaneously hypertensive rats*. *Experientia*, 1979. **35**(12): p. 1594-5.

84. Stark, O. and V. Kren, *Five congenic resistant lines of rats differing at the Rt H-1 locus*. Transplantation, 1969. **8**(2): p. 200-3.
85. Rapp, J.P., *Dahl salt-susceptible and salt-resistant rats. A review*. Hypertension, 1982. **4**(6): p. 753-63.
86. Zicha, J., et al., *Age-dependent salt hypertension in Dahl rats: fifty years of research*. Physiol Res, 2012. **61 Suppl 1**: p. S35-87.
87. Alvarez-Guerra, M. and R.P. Garay, *Renal Na-K-Cl cotransporter NKCC2 in Dahl salt-sensitive rats*. J Hypertens, 2002. **20**(4): p. 721-7.
88. Aoi, W., et al., *Aldosterone-induced abnormal regulation of ENaC and SGK1 in Dahl salt-sensitive rat*. Biochem Biophys Res Commun, 2006. **341**(2): p. 376-81.
89. Kato, T., et al., *Decreased sensitivity to renal interstitial hydrostatic pressure in Dahl salt-sensitive rats*. Hypertension, 1994. **23**(6 Pt 2): p. 1082-6.
90. Chao, C., et al., *Differential regulation of kallikrein, kininogen, and kallikrein-binding protein in arterial hypertensive rats*. Am J Physiol, 1996. **271**(1 Pt 2): p. F78-86.
91. St Lezin, E.M., et al., *Effects of renin gene transfer on blood pressure and renin gene expression in a congenic strain of Dahl salt-resistant rats*. J Clin Invest, 1996. **97**(2): p. 522-7.
92. Wang, C., et al., *High-salt diet upregulates kininogen and downregulates tissue kallikrein expression in Dahl-SS and SHR rats*. Am J Physiol, 1996. **271**(4 Pt 2): p. F824-30.
93. Hayakawa, H. and L. Raij, *Nitric oxide synthase activity and renal injury in genetic hypertension*. Hypertension, 1998. **31**(1 Pt 2): p. 266-70.
94. Deng, A.Y., *Genetic basis of polygenic hypertension*. Hum Mol Genet, 2007. **16 Spec No. 2**: p. R195-202.
95. Rapp, J.P. and A.Y. Deng, *Detection and positional cloning of blood pressure quantitative trait loci: is it possible? Identifying the genes for genetic hypertension*. Hypertension, 1995. **25**(6): p. 1121-8.
96. Delles, C. and S. Padmanabhan, *Genetics and hypertension: is it time to change my practice?* Can J Cardiol, 2012. **28**(3): p. 296-304.
97. Syvanen, A.C., *Assessing genetic variation: genotyping single nucleotide polymorphisms*. Nat Rev Genet, 2001. **2**(12): p. 930-42.
98. Lander, E. and L. Kruglyak, *Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results*. Nat Genet, 1995. **11**(3): p. 241-7.
99. Altshuler, D., M.J. Daly, and E.S. Lander, *Genetic mapping in human disease*. Science, 2008. **322**(5903): p. 881-8.
100. International Consortium for Blood Pressure Genome-Wide Association, S., et al., *Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk*. Nature, 2011. **478**(7367): p. 103-9.
101. Ott, J., J. Wang, and S.M. Leal, *Genetic linkage analysis in the age of whole-genome sequencing*. Nat Rev Genet, 2015. **16**(5): p. 275-84.
102. van der Sijde, M.R., A. Ng, and J. Fu, *Systems genetics: From GWAS to disease pathways*. Biochim Biophys Acta, 2014. **1842**(10): p. 1903-1909.
103. Garrett, M.R., et al., *Genome scan and congenic strains for blood pressure QTL using Dahl salt-sensitive rats*. Genome Res, 1998. **8**(7): p. 711-23.

104. Cicila, G.T. and S.J. Lee, *Identifying candidate genes for blood pressure quantitative trait loci using differential gene expression and a panel of congenic strains*. Hypertens Res, 1998. **21**(4): p. 289-96.
105. Rapp, J.P., M.R. Garrett, and A.Y. Deng, *Construction of a double congenic strain to prove an epistatic interaction on blood pressure between rat chromosomes 2 and 10*. J Clin Invest, 1998. **101**(8): p. 1591-5.
106. Duong, C., et al., *Individual QTLs controlling quantitative variation in blood pressure inherited in a Mendelian mode*. Heredity (Edinb), 2007. **98**(3): p. 165-71.
107. Deng, Y. and J.P. Rapp, *Cosegregation of blood pressure with angiotensin converting enzyme and atrial natriuretic peptide receptor genes using Dahl salt-sensitive rats*. Nat Genet, 1992. **1**(4): p. 267-72.
108. Charron, S., et al., *A loss of genome buffering capacity of Dahl salt-sensitive model to modulate blood pressure as a cause of hypertension*. Hum Mol Genet, 2005. **14**(24): p. 3877-84.
109. Crespo, K., et al., *Normotension in Lewis and Dahl salt-resistant rats is governed by different genes*. J Hypertens, 2011. **29**(3): p. 460-5.
110. Deng, A.Y., *Positional cloning of quantitative trait Loci for blood pressure: how close are we?: a critical perspective*. Hypertension, 2007. **49**(4): p. 740-7.
111. Palijan, A., et al., *Comprehensive congenic coverage revealing multiple blood pressure quantitative trait loci on Dahl rat chromosome 10*. Hypertension, 2003. **42**(4): p. 515-22.
112. Dutil, J., et al., *Multiple quantitative trait loci for blood pressure interacting epistatically and additively on Dahl rat chromosome 2*. Hypertension, 2005. **45**(4): p. 557-64.
113. Bateson, W., *Facts Limiting the Theory of Heredity*. Science, 1907. **26**(672): p. 649-60.
114. Chauvet, C., et al., *Submegabase resolution of epistatically interacting quantitative trait loci for blood pressure applicable for essential hypertension*. J Hypertens, 2008. **26**(5): p. 893-901.
115. Palijan, A., J. Dutil, and A.Y. Deng, *Quantitative trait loci with opposing blood pressure effects demonstrating epistasis on Dahl rat chromosome 3*. Physiol Genomics, 2003. **15**(1): p. 1-8.
116. Cordell, H.J., *Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans*. Hum Mol Genet, 2002. **11**(20): p. 2463-8.
117. Kreutz, R., et al., *Dissection of a quantitative trait locus for genetic hypertension on rat chromosome 10*. Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8778-82.
118. Dutil, J. and A.Y. Deng, *Further chromosomal mapping of a blood pressure QTL in Dahl rats on chromosome 2 using congenic strains*. Physiol Genomics, 2001. **6**(1): p. 3-9.
119. Garrett, M.R., et al., *Two linked blood pressure quantitative trait loci on chromosome 10 defined by dahl rat congenic strains*. Hypertension, 2001. **38**(4): p. 779-85.
120. Saad, Y., M.R. Garrett, and J.P. Rapp, *Multiple blood pressure QTL on rat chromosome 1 defined by Dahl rat congenic strains*. Physiol Genomics, 2001. **4**(3): p. 201-14.
121. Garrett, M.R. and J.P. Rapp, *Two closely linked interactive blood pressure QTL on rat chromosome 5 defined using congenic Dahl rats*. Physiol Genomics, 2002. **8**(2): p. 81-6.

122. Garrett, M.R. and J.P. Rapp, *Multiple blood pressure QTL on rat Chromosome 2 defined by congenic Dahl rats*. Mamm Genome, 2002. **13**(1): p. 41-4.
123. Moujahidine, M., et al., *Congenic mapping of a blood pressure QTL on chromosome 16 of Dahl rats*. Mamm Genome, 2002. **13**(3): p. 153-6.
124. Garrett, M.R. and J.P. Rapp, *Defining the blood pressure QTL on chromosome 7 in Dahl rats by a 177-kb congenic segment containing Cyp11b1*. Mamm Genome, 2003. **14**(4): p. 268-73.
125. Meng, H., et al., *Localization of a blood pressure QTL to a 2.4-cM interval on rat chromosome 9 using congenic strains*. Genomics, 2003. **81**(2): p. 210-20.
126. Charron, S., et al., *Epistasis, not numbers, regulates functions of clustered Dahl rat quantitative trait loci applicable to human hypertension*. Hypertension, 2005. **46**(6): p. 1300-8.
127. Deng, A.Y., et al., *Sexual dimorphism on hypertension of quantitative trait loci entrapped in Dahl congenic rats*. Clin Exp Hypertens, 2008. **30**(7): p. 511-9.
128. Herrera, V.L., et al., *Sex-specific effects on spatial learning and memory, and sex-independent effects on blood pressure of a <3.3 Mbp rat chromosome 2 QTL region in Dahl salt-sensitive rats*. PLoS One, 2013. **8**(7): p. e67673.
129. Gopalakrishnan, K., et al., *Increased Expression of Rifiylin in A < 330 Kb Congenic Strain is Linked to Impaired Endosomal Recycling in Proximal Tubules*. Front Genet, 2012. **3**: p. 138.
130. Joe, B., et al., *Positional identification of variants of Adamts16 linked to inherited hypertension*. Hum Mol Genet, 2009. **18**(15): p. 2825-38.
131. Manunta, P. and G. Bianchi, *Pharmacogenomics and pharmacogenetics of hypertension: update and perspectives--the adducin paradigm*. J Am Soc Nephrol, 2006. **17**(4 Suppl 2): p. S30-5.
132. Rafiq, S., S. Anand, and R. Roberts, *Genome-wide association studies of hypertension: have they been fruitful?* J Cardiovasc Transl Res, 2010. **3**(3): p. 189-96.
133. Sober, S., et al., *Targeting 160 candidate genes for blood pressure regulation with a genome-wide genotyping array*. PLoS One, 2009. **4**(6): p. e6034.
134. Deng, A.Y., H. Dene, and J.P. Rapp, *Mapping of a quantitative trait locus for blood pressure on rat chromosome 2*. J Clin Invest, 1994. **94**(1): p. 431-6.
135. Cicila, G.T., et al., *Linkage of 11 beta-hydroxylase mutations with altered steroid biosynthesis and blood pressure in the Dahl rat*. Nat Genet, 1993. **3**(4): p. 346-53.
136. Duong, C., et al., *Distinct quantitative trait loci for kidney, cardiac, and aortic mass dissociated from and associated with blood pressure in Dahl congenic rats*. Mamm Genome, 2006. **17**(12): p. 1147-61.
137. Ariyarajah, A., et al., *Dissecting quantitative trait loci into opposite blood pressure effects on Dahl rat chromosome 8 by congenic strains*. J Hypertens, 2004. **22**(8): p. 1495-502.
138. Grondin, M., et al., *Complete and overlapping congenics proving the existence of a quantitative trait locus for blood pressure on Dahl rat chromosome 17*. Physiol Genomics, 2005. **21**(1): p. 112-6.
139. Eliopoulos, V., et al., *Severe hypertension caused by alleles from normotensive Lewis for a quantitative trait locus on chromosome 2*. Physiol Genomics, 2005. **22**(1): p. 70-5.

140. Chauvet, C., et al., *Cardiac pathways distinguish two epistatic modules enacting BP quantitative trait loci and candidate gene analysis*. Hypertens Res, 2009. **32**(7): p. 631-7.
141. Chauvet, C., et al., *Novel genes as primary triggers for polygenic hypertension*. J Hypertens, 2012. **30**(1): p. 81-6.
142. Leineweber, K. and G. Heusch, *Beta 1- and beta 2-adrenoceptor polymorphisms and cardiovascular diseases*. Br J Pharmacol, 2009. **158**(1): p. 61-9.
143. Insel, P.A., *beta(2)-Adrenergic receptor polymorphisms and signaling: Do variants influence the "memory" of receptor activation?* Sci Signal, 2011. **4**(185): p. pe37.
144. Ferguson, S.S. and R.D. Feldman, *beta-adrenoceptors as molecular targets in the treatment of hypertension*. Can J Cardiol, 2014. **30**(5 Suppl): p. S3-8.
145. Russo, C.J., et al., *Association of NEDD4L ubiquitin ligase with essential hypertension*. Hypertension, 2005. **46**(3): p. 488-91.
146. Manunta, P., et al., *Physiological interaction between alpha-adducin and WNK1-NEDD4L pathways on sodium-related blood pressure regulation*. Hypertension, 2008. **52**(2): p. 366-72.
147. Goel, P., J.A. Manning, and S. Kumar, *NEDD4-2 (NEDD4L): the ubiquitin ligase for multiple membrane proteins*. Gene, 2015. **557**(1): p. 1-10.
148. Svensson-Farbom, P., et al., *A functional variant of the NEDD4L gene is associated with beneficial treatment response with beta-blockers and diuretics in hypertensive patients*. J Hypertens, 2011. **29**(2): p. 388-95.
149. Dahlberg, J., et al., *Genetic variation in NEDD4L, an epithelial sodium channel regulator, is associated with cardiovascular disease and cardiovascular death*. J Hypertens, 2014. **32**(2): p. 294-9.
150. Mills, P.A., et al., *A new method for measurement of blood pressure, heart rate, and activity in the mouse by radiotelemetry*. J Appl Physiol (1985), 2000. **88**(5): p. 1537-44.
151. Chauvet, C., et al., *Modularization and epistatic hierarchy determine homeostatic actions of multiple blood pressure quantitative trait loci*. Hum Mol Genet, 2013. **22**(22): p. 4451-9.
152. Wasserman, L., *All of statistics : a concise course in statistical inference*. Springer texts in statistics. 2004, New York: Springer. xix, 442 p.
153. Fitzner, K. and E. Heckinger, *Sample size calculation and power analysis: a quick review*. Diabetes Educ, 2010. **36**(5): p. 701-7.
154. den Dunnen, J.T. and S.E. Antonarakis, *Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion*. Hum Mutat, 2000. **15**(1): p. 7-12.
155. Chauvet, C., A. Menard, and A.Y. Deng, *Two candidate genes for two quantitative trait loci epistatically attenuate hypertension in a novel pathway*. J Hypertens, 2015. **33**(9): p. 1791-801.
156. Zhang, F. and J.R. Lupski, *Non-coding genetic variants in human disease*. Hum Mol Genet, 2015.
157. Wei, W.H., G. Hemani, and C.S. Haley, *Detecting epistasis in human complex traits*. Nat Rev Genet, 2014. **15**(11): p. 722-33.

158. Gopalakrishnan, K., et al., *Augmented rififylin is a risk factor linked to aberrant cardiomyocyte function, short-QT interval and hypertension*. *Hypertension*, 2011. **57**(4): p. 764-71.
159. Avery, L. and S. Wasserman, *Ordering gene function: the interpretation of epistasis in regulatory hierarchies*. *Trends Genet*, 1992. **8**(9): p. 312-6.
160. Pottier, N., et al., *Promoter polymorphisms in the beta-2 adrenergic receptor are associated with drug-induced gene expression changes and response in acute lymphoblastic leukemia*. *Clin Pharmacol Ther*, 2010. **88**(6): p. 854-61.
161. Wu, H., et al., *Association of the beta2-adrenergic receptor gene with essential hypertension in the non-Han Chinese Yi minority human population*. *J Hypertens*, 2006. **24**(6): p. 1041-7.
162. Rizzo, F. and O. Staub, *NEDD4-2 and salt-sensitive hypertension*. *Curr Opin Nephrol Hypertens*, 2015. **24**(2): p. 111-6.
163. Elenkov, I.J., et al., *The sympathetic nerve--an integrative interface between two supersystems: the brain and the immune system*. *Pharmacol Rev*, 2000. **52**(4): p. 595-638.
164. Gorelik, J., et al., *Spatial control of the betaAR system in heart failure: the transverse tubule and beyond*. *Cardiovasc Res*, 2013. **98**(2): p. 216-24.
165. Pagani, F. and F.E. Baralle, *Genomic variants in exons and introns: identifying the splicing spoilers*. *Nat Rev Genet*, 2004. **5**(5): p. 389-96.
166. Wang, X., et al., *Beyond genome-wide association studies: new strategies for identifying genetic determinants of hypertension*. *Curr Hypertens Rep*, 2011. **13**(6): p. 442-51.
167. Klopocki, E. and S. Mundlos, *Copy-number variations, noncoding sequences, and human phenotypes*. *Annu Rev Genomics Hum Genet*, 2011. **12**: p. 53-72.
168. Marchler-Bauer, A., et al., *CDD: a Conserved Domain Database for the functional annotation of proteins*. *Nucleic Acids Res*, 2011. **39**(Database issue): p. D225-9.
169. Redina, O.E., et al., *Candidate genes in quantitative trait loci associated with absolute and relative kidney weight in rats with Inherited Stress Induced Arterial Hypertension*. *BMC Genet*, 2015. **16 Suppl 1**: p. S1.
170. Rockman, M.V. and L. Kruglyak, *Genetics of global gene expression*. *Nat Rev Genet*, 2006. **7**(11): p. 862-72.
171. Farber, C.R. and A.J. Lusis, *Integrating global gene expression analysis and genetics*. *Adv Genet*, 2008. **60**: p. 571-601.
172. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome*. *Nature*, 2012. **489**(7414): p. 57-74.
173. Mouse, E.C., et al., *An encyclopedia of mouse DNA elements (Mouse ENCODE)*. *Genome Biol*, 2012. **13**(8): p. 418.
174. Ritchie, G.R., et al., *Functional annotation of noncoding sequence variants*. *Nat Methods*, 2014. **11**(3): p. 294-6.
175. Edwards, S.L., et al., *Beyond GWASs: illuminating the dark road from association to function*. *Am J Hum Genet*, 2013. **93**(5): p. 779-97.
176. Zhang, X., et al., *Identification of common genetic variants controlling transcript isoform variation in human whole blood*. *Nat Genet*, 2015. **47**(4): p. 345-52.
177. Dekker, J., et al., *Capturing chromosome conformation*. *Science*, 2002. **295**(5558): p. 1306-11.

178. Deng, A.Y., *Genetics of systolic and diastolic heart failure*. J Hypertens, 2015. **33**(1): p. 3-13.
179. Gilles, A.F. and M. Averof, *Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution*. Evodevo, 2014. **5**: p. 43.
180. Nicholls, M.G., et al., *Bioactivity of adrenomedullin and proadrenomedullin N-terminal 20 peptide in man*. Peptides, 2001. **22**(11): p. 1745-52.
181. Serber, Z. and J.E. Ferrell, Jr., *Tuning bulk electrostatics to regulate protein function*. Cell, 2007. **128**(3): p. 441-4.
182. Preumont, A., et al., *HDHDI, which is often deleted in X-linked ichthyosis, encodes a pseudouridine-5'-phosphatase*. Biochem J, 2010. **431**(2): p. 237-44.
183. Dixon, J., P. Trainor, and M.J. Dixon, *Treacher Collins syndrome*. Orthod Craniofac Res, 2007. **10**(2): p. 88-95.
184. Stanford, S.M., et al., *Cellular biochemistry methods for investigating protein tyrosine phosphatases*. Antioxid Redox Signal, 2014. **20**(14): p. 2160-78.
185. Kathiresan, S. and D. Srivastava, *Genetics of human cardiovascular disease*. Cell, 2012. **148**(6): p. 1242-57.
186. Zhuo, S., et al., *Expression, purification, crystallization, and biochemical characterization of a recombinant protein phosphatase*. J Biol Chem, 1993. **268**(24): p. 17754-61.
187. Wu, Q., et al., *Molecular cloning and characterization of a novel dual-specificity phosphatase 23 gene from human fetal brain*. Int J Biochem Cell Biol, 2004. **36**(8): p. 1542-53.
188. Kuznetsova, E., et al., *Enzyme genomics: Application of general enzymatic screens to discover new enzymes*. FEMS Microbiol Rev, 2005. **29**(2): p. 263-79.
189. Xie, L., Y.L. Zhang, and Z.Y. Zhang, *Design and characterization of an improved protein tyrosine phosphatase substrate-trapping mutant*. Biochemistry, 2002. **41**(12): p. 4032-9.
190. Stasyk, T., et al., *Quantitative detection of phosphoproteins by combination of two-dimensional difference gel electrophoresis and phosphospecific fluorescent staining*. Electrophoresis, 2005. **26**(14): p. 2850-4.
191. Li, X., et al., *Elucidating human phosphatase-substrate networks*. Sci Signal, 2013. **6**(275): p. rs10.
192. Delom, F. and E. Chevet, *Phosphoprotein analysis: from proteins to proteomes*. Proteome Sci, 2006. **4**: p. 15.
193. McLachlin, D.T. and B.T. Chait, *Analysis of phosphorylated proteins and peptides by mass spectrometry*. Curr Opin Chem Biol, 2001. **5**(5): p. 591-602.
194. Bruckner, A., et al., *Yeast two-hybrid, a powerful tool for systems biology*. Int J Mol Sci, 2009. **10**(6): p. 2763-88.

