



**HAL**  
open science

## Détection de QTL en interactions

Céline Delmas, Charles-Elie Rabier

► **To cite this version:**

Céline Delmas, Charles-Elie Rabier. Détection de QTL en interactions. 42èmes Journées de Statistique, 2010, Marseille, France, France. inria-00494742

**HAL Id: inria-00494742**

**<https://inria.hal.science/inria-00494742>**

Submitted on 24 Jun 2010

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

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

# DÉTECTION DE QTL EN INTERACTIONS

Céline Delmas<sup>1</sup> & Charles-Elie Rabier<sup>1,2</sup>

<sup>1</sup> *Station d'Amélioration Génétique des Animaux  
INRA, Auzeville B.P. 52627, 31326 Castanet Tolosan  
celine.delmas@toulouse.inra.fr*

<sup>2</sup> *Université de Toulouse,  
Institut de Mathématiques de Toulouse, U.P.S. et I.N.S.A,  
F-31062 Toulouse Cedex 9  
charles-elie.rabier@insa-toulouse.fr*

**Résumé :** Le chromosome est modélisé par un segment  $[0, T]$ . Chaque locus correspond à une position  $t \in [0, T]$ . Nous construisons une statistique de test en tout point  $(s, t) \in [0, T]^2$  pour tester la significativité de l'interaction entre les deux loci sur un caractère. Nous étudions la loi de cette statistique de test sur  $[0, T]^2$ . Nous proposons alors des méthodes pour la détection et la localisation de loci et d'interactions entre loci ayant un effet significatif sur un caractère.

**Abstract :** A chromosome is represented by a segment  $[0, T]$ . Each locus is represented by a position  $t \in [0, T]$ . We give a test statistics at each point  $(s, t) \in [0, T]^2$  to test the significance of the interaction between the two loci  $s$  and  $t$  on a trait. We give the asymptotic distribution of this test statistics on  $[0, T]^2$ . Then we propose methods to detect loci and interactions between loci whose effect is significant.

**Mots-clés :** Détection de QTL, Interactions, Processus et champs gaussiens, Modèles de mélange.

**Keywords :** QTL detection, Interactions, Gaussian processes and fields, Mixture models.

## 1 Introduction et modélisation

On s'intéresse au problème de la détection et de la localisation de zones du génome (ensemble des chromosomes) ayant un effet sur un caractère quantitatif. Ces zones sont appelées QTL (Quantitative Trait Loci). Chaque chromosome est représenté par un segment  $[0, T]$ . On teste la présence d'un QTL en chaque point  $t \in [0, T]$ . Notons  $Y_i$  la valeur du caractère quantitatif mesuré sur l'individu  $i$ ,  $i$  varie de 1 jusqu'à  $n$ . Supposons qu'il existe  $m$  QTL positionnés en  $t_1^*, \dots, t_m^*$  agissant de manière additive sur le caractère. Notons  $q_1, \dots, q_m$  les effets de ces QTL sur le caractère. Notons  $X_i(t)$  le génotype de l'individu  $i$  au point  $t$ . Dans le cas particulier des populations que nous étudions, la variable aléatoire  $X_i(t)$  est à valeurs dans  $\{-1, 1\}$ . La loi du processus  $X_i$  est obtenue à partir du modèle

de Haldane (1919) (voir Rabier *et al.*, 2009). Nous pouvons alors poser le modèle suivant sur le caractère :

$$Y_i = \mu + \sum_{j=1}^m X_i(t_j^*)q_j + \epsilon_i \quad (1)$$

$\epsilon_i$  est une erreur de mesure supposée gaussienne de moyenne nulle et de variance  $\sigma^2$ ,  $\mu$  est un paramètre inconnu. Lorsque nous rajoutons les possibles interactions entre QTL, le modèle devient :

$$Y_i = \mu + \sum_{j=1}^m X_i(t_j^*)q_j + \sum_{j=1}^m \sum_{k>j}^m X_i(t_j^*)X_i(t_k^*)\gamma_{jk} + \epsilon_i \quad (2)$$

Cette modélisation du processus  $X_i$  réalisée dans le cadre d'une analyse de liaison peut être étendue à une analyse par déséquilibre de liaison en utilisant par exemple la modélisation proposée par Scheet et Stephens (2006) puis Servin et Stephens (2007).

## 2 Résultats

Dans le cadre du modèle (1) une statistique de test  $S_n(t)$  peut être construite en tout point  $t \in [0, T]$  pour tester la significativité de l'effet QTL au point  $t$ . Nous avons montré que  $S_n$  est asymptotiquement un processus gaussien de fonction de covariance connue et dont la fonction moyenne dépend des effets QTL.

Dans le cadre du modèle (2) une statistique de test  $S_n(s, t)$  peut être construite en tout point  $(s, t) \in [0, T]^2$  pour tester la significativité de l'interaction entre les loci  $s$  et  $t$ . Nous donnons la loi asymptotique du champ aléatoire  $S_n$ .

Nous montrons que le problème de l'estimation des interactions et des effets QTL significatifs se ramène au problème de l'estimation des zéros dans la moyenne d'un vecteur gaussien dont la matrice de variance est connue. Nous utilisons alors des techniques proposées dans la Littérature pour résoudre un tel problème (Huet, 2006 ; Gendre, 2009). Nos résultats sont validés sur des simulations.

## Bibliographie

- [1] Gendre X. (2009), Estimation par sélection de modèle en régression hétéroscédastique *Thèse de l'Université de Nice Sophia Antipolis - UFR Sciences*.
- [2] Haldane J.B.S. (1919), The combination of linkage values and the calculation of distance between the loci of linked factors *Journal of genetics* 8, 299-309.
- [2] Huet S. (2006), Model Selection for estimating the non zero components of a Gaussian vector *ESAIM : P&S*, 10, 164–183.
- [3] Rabier C.E., Azaïs J.M., Delmas C. (2009), Likelihood Ratio Test Process for Quantitative Trait Loci Detection <http://hal.archives-ouvertes.fr/hal-00421215/fr/>.

- [4] Scheet P., Stephens M. (2006), A Fast and Flexible Statistical Model for Large-Scale Population Genotype Data : Applications to Inferring Missing Genotypes and Haplotypic Phase *American Journal of Human Genetics* 78, 629–644.
- [5] Servin B., Stephens M. (2007), Imputation-based analysis of association studies : Candidate regions and quantitative traits *PLoS Genet* 3, 1296–1308.