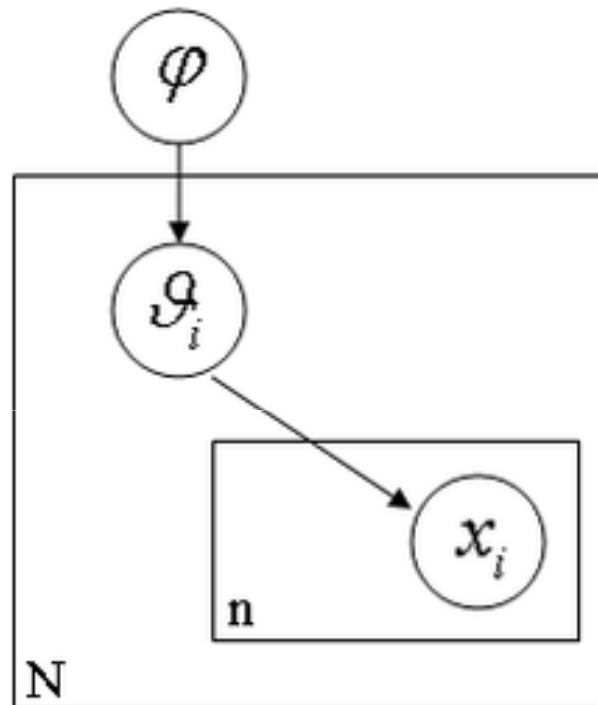


Modèles bayésiens hiérarchiques/bases et applications en biostatistique



Jean-Louis Foulley
jean-louis.foulley@jouy.inra.fr

Sommaire

- ❑ Introduction
- ❑ Inférence bayésienne
- ❑ Modèle en grappe
- ❑ Modèles linéaires
 - ❑ Formulation
 - ❑ Ex Données longitudinales
 - ❑ Implantation en Winbugs
- ❑ Modèles non linéaires
 - ❑ Ex données de croissance
 - ❑ Différenciation génétique
- ❑ Conclusion

Construction d'un modèle

Combinaison de différents types d'informations

- Objectifs de l'analyse statistique
- Structures des données
 - Plan d'expérience
 - Unités expérimentales
 - Traitements, blocs
 - Populations, individus
 - Mode d'échantillonnage
 - Nature des variables
 - Réponses, covariables, échelle
 - Structures longitudinales, transversales, temporelles, spatiales
- Statut et valeurs des paramètres
 - Effets fixes vs effets aléatoires
 - Connaissances préalables
 - Position, dispersion, quantiles
 - Bibliographie, historique
- Lois de probabilité

Construction d'un modèle

Modélisation

- **Structure hiérarchique** , dite aussi **multi-niveaux** «emboîtés »
 - ❑ Mesures intra patients intra cliniques intra régions
 - ❑ Cellules intra organes intra individus intra familles intra groupes socio-professionnels
 - ❑ Gènes intra locus intra individus (gamètes) intra lignées, intra populations
 - ❑ Cf Goldstein H (1999), Raudenbush SW & Brick AS (2002)
- **Paradigme bayésien**
 - ❑ Informations structurelles
 - ❑ Informations sur les paramètres

Bayes/Théorème

Th (ou formule) de Bayes (version discrète)

Formule des probabilités composées

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$$

Soit une partition de Ω ("les causes")

$$\Omega = \left(\cup B_j; j = 1, \dots, n \right), B_j \cap B_k = \emptyset, \forall j \neq k$$

$$P(B_j | A) = P(B_j)P(A|B_j) / P(A)$$

$A = \left(\cup_{k=1, \dots, n} (A \cap B_k) \right)$ d'où formule des probabilités totales

$$P(A) = \sum_{k=1}^n P(B_k)P(A|B_k)$$

$$P(B_j | A) = \frac{P(B_j)P(A|B_j)}{\sum_{k=1}^n P(B_k)P(A|B_k)} \Rightarrow \frac{P(B_j | A)}{P(\bar{B}_j | A)} = \frac{P(B_j)P(A|B_j)}{P(\bar{B}_j)P(A|\bar{B}_j)}$$

Forme due à Laplace (1774) et non pas à Bayes (1763)

cf loi de l'éponymie de S Stiegler (1980) "No scientific discovery is named after his original discoverer"

Bayes/Théorème/Exemple

On applique un test pour détecter une maladie.

Le test est positif (+)

-98% des cas quand il est administré à une personne malade (M)

-6% des cas quand il l'est à une personne bien portante ($\bar{M}$)

Un patient se fait tester. **Le résultat est positif.**

Quelle est la pté (p) qu'il soit atteint de la maladie.

On sait par des enquetes que la maladie touche 4% de la population

$$\frac{p}{1-p} = \frac{P(M | +)}{P(\bar{M} | +)} = \frac{P(M) P(+|M)}{P(\bar{M}) P(+|\bar{M})} = \frac{4}{96} \frac{98}{6} = 0.6805$$

$$p = 0.6805 / (1 + 0.6805) = \mathbf{0.40}$$

Bayes/Théorème/Forme générale

Th de Bayes (forme générale)

y : données θ : paramètres

$$[y, \theta] = [y | \theta][\theta] = [\theta | y][y]$$

où $[.]$ désigne une densité ou une probabilité

$$\pi(\theta | y_{obs}) = \frac{f(y_{obs} | \theta) \pi(\theta)}{m(y_{obs})} : \text{densité a posteriori dite "posterior"}$$

$$\pi(\theta) : \text{densité a priori des paramètres } \theta \text{ ou "prior"}$$

$$f(y_{obs} | \theta) : \text{vraisemblance (en tant que fonction des paramètres)}$$

$$m(y_{obs}) = \int f(y_{obs} | \theta) \pi(\theta) d\theta : \text{densité marginale (données, modèle)}$$

Bayes/Modèle en grappes

Structure en grappes ("clusters") $C_1, C_2, \dots, C_i, \dots, C_I$

Ex: descendants intra famille; patients intra cliniques

étudiants intra universités; chercheurs intra laboratoires

-Décomposition de la **variabilité entre classes et intra - classes**

-Prise en compte de la **corrélation intra - classe**

Bayes/Modèle en grappes

Pour $i = 1, \dots, p; j = 1, \dots, n_i$

$$1) y_{ij} | \theta_i \sim_{\text{id}} \mathcal{N} \left(\theta_i, \underbrace{\sigma^2}_{\text{Variance intra C}} \right) \Leftrightarrow y_{ij} = \theta_i + e_{ij}$$

$$2) \theta_i \sim_{\text{id}} \mathcal{N} \left(\mu, \underbrace{\tau^2}_{\text{Variance entre C}} \right) \Leftrightarrow \theta_i = \mu + u_i$$

$$y_{ij} \sim \mathcal{N}(\mu, \tau^2 + \sigma^2) \Leftrightarrow y_{ij} = \mu + u_i + e_{ij}$$

$$\text{Cov}(y_{ij}, y_{ij'}) = \tau^2 \text{ (covariance intraclass)}$$

$$\theta_i | \mathbf{y} \sim \mathcal{N}(\hat{\theta}_i, v_i) \text{ où}$$

$$\frac{\hat{\theta}_i}{v_i} = \frac{\bar{y}_i}{\sigma^2 / n_i} + \frac{\mu}{\tau^2} \quad \frac{1}{v_i} = \frac{1}{\sigma^2 / n_i} + \frac{1}{\tau^2}$$

Bayes/Modèle en grappes

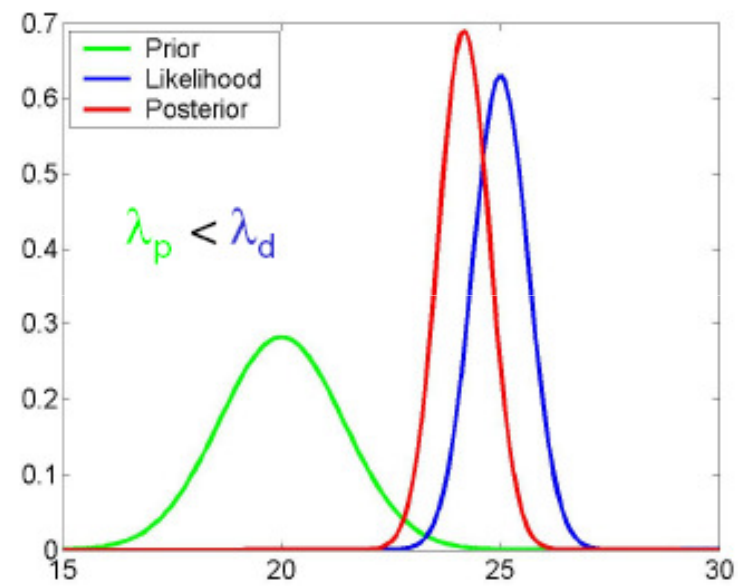
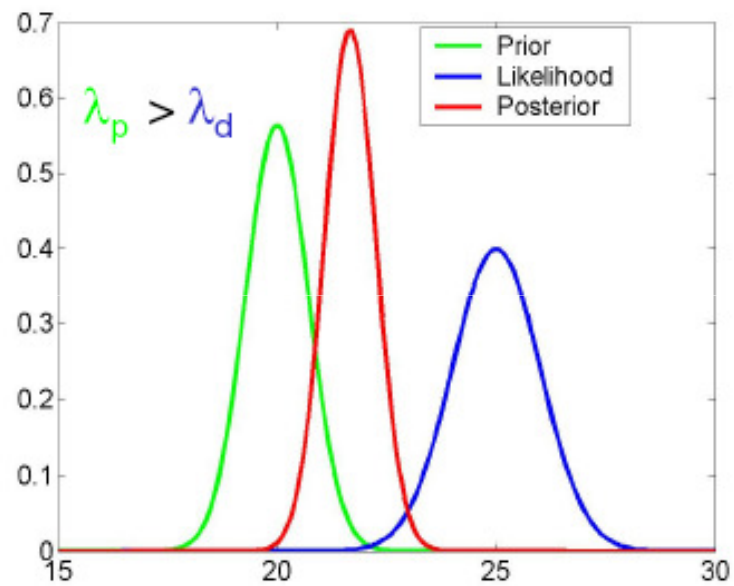
$$\hat{\theta}_i = \mu + b_i (\bar{y}_i - \mu) \quad \text{vs} \quad \hat{\theta}_{i,ML} = \bar{y}_i$$

$$\text{où } 0 \leq b_i = \underbrace{\frac{n_i \tau^2}{n_i \tau^2 + \sigma^2}}_{\text{Shrinkage}} \leq 1 \quad \bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$$

Relation avec l'estimateur de James-Stein

$$\Rightarrow \left| \begin{array}{l} \tau^2 = 0 \Rightarrow b_i = 0 \Rightarrow \hat{\theta}_i = \mu \\ \sigma^2 \rightarrow 0 \text{ ou } n_i \rightarrow \infty \Rightarrow b_i = 1 \Rightarrow \hat{\theta}_i = \bar{y}_i \end{array} \right.$$

Bayes/Shrinkage/Ex



Bayes/Modèle en grappes

$$\begin{array}{l} \text{Str. populations} \left\{ \begin{array}{l} 1) y_{ij} \mid \theta_i \sim_{\text{id}} \mathcal{N}(\theta_i, \sigma^2) \text{ "intra-sujets"} \\ 2a) \theta_i \sim_{\text{id}} \mathcal{N}(\mu, \tau^2) \text{ "entre-sujets"} \end{array} \right. \\ \text{Info. paramètres} \left\{ \begin{array}{l} 2b) \sigma \sim \text{U}(0, \Delta_\sigma) \text{ par exemple} \\ 3) \pi(\mu) = 1 \text{ (prior plat)} \quad \tau \sim \text{U}(0, \Delta_\tau) \end{array} \right. \end{array}$$

Modèle linéaire / Lindley & Smith (1972)

Approche de Lindley-Smith du modèle linéaire
(cad avec covariables discrètes et/ou continues)

$$1) \mathbf{y}_{N \times 1} | \boldsymbol{\theta} \sim \mathcal{N} \left(\underbrace{\mathbf{Z}}_{N \times q} \underbrace{\boldsymbol{\theta}}_{q \times 1}, \mathbf{R} \right) \quad \text{Ex } \mathbf{R} = \sigma^2 \mathbf{I}_N$$

$$2) \boldsymbol{\theta}_{p \times 1} | \boldsymbol{\beta} \sim \mathcal{N} \left(\underbrace{\mathbf{A}}_{q \times p} \underbrace{\boldsymbol{\beta}}_{p \times 1}, \mathbf{G} \right)$$

$$\pi(\sigma^2) \propto 1/\sigma^2 : \text{Jeffreys}$$

$$\text{ou } \sigma \sim \text{U}(0, \Delta_\sigma) \quad \text{ou } \sigma^2 \sim \text{IG}(1/2\nu, 1/2\nu\sigma^{*2})$$

$$3) \pi(\boldsymbol{\beta}) = 1 \quad (\boldsymbol{\beta} \text{ traité comme effet fixe})$$

$$\text{ou } \beta_k \sim \text{U}(\beta_{k,\text{inf}}, \beta_{k,\text{sup}})$$

$$\mathbf{G} \sim \text{IW} \left[(\rho \mathbf{G}^*)^{-1}, \rho \right]$$

Modèle bayésien hiérarchique/schéma type

$$1^{\text{er}} \text{ niveau} : [\mathbf{y} | \boldsymbol{\theta}_1] \sim f(\mathbf{x} | \boldsymbol{\theta}_1)$$

$$2^{\text{ème}} \text{ niveau: } [\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2] \sim \pi_1(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2)$$

$$3^{\text{ème}} \text{ niveau: } [\boldsymbol{\theta}_2 | \boldsymbol{\theta}_3] \sim \pi_2(\boldsymbol{\theta}_2 | \boldsymbol{\theta}_3)$$

$$4^{\text{ème}} \text{ niveau: } [\boldsymbol{\theta}_3] \sim \pi_3(\boldsymbol{\theta}_3)$$

$$[\boldsymbol{\theta}] = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3] = [\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2, \boldsymbol{\theta}_3][\boldsymbol{\theta}_2, \boldsymbol{\theta}_3]$$

$$[\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2, \boldsymbol{\theta}_3] = \pi_1(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2)$$

$$[\boldsymbol{\theta}_2, \boldsymbol{\theta}_3] = \pi_2(\boldsymbol{\theta}_2 | \boldsymbol{\theta}_3) \pi_3(\boldsymbol{\theta}_3)$$

$$\pi(\boldsymbol{\theta}) = \pi_1(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2) \pi_2(\boldsymbol{\theta}_2 | \boldsymbol{\theta}_3) \pi_3(\boldsymbol{\theta}_3)$$

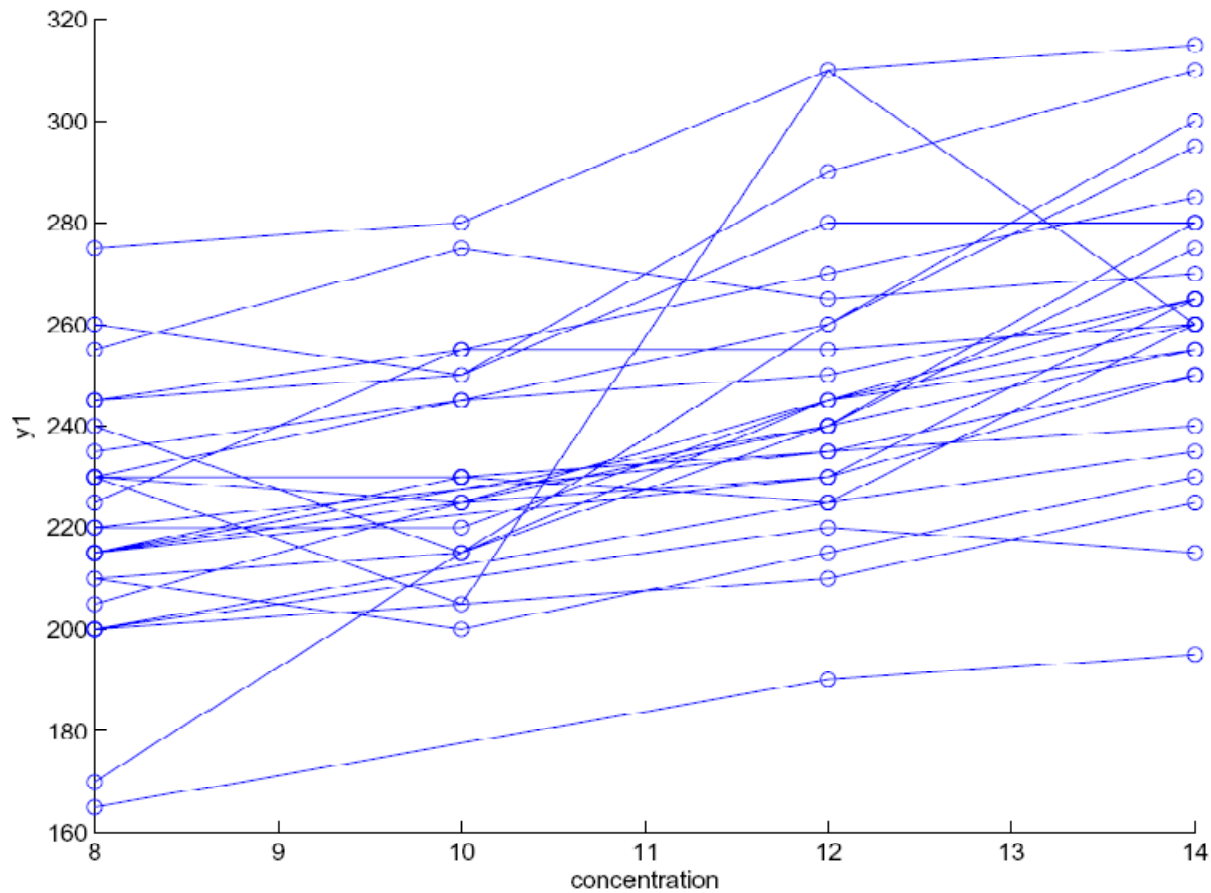
Exemple/Données Pothoff

Growth measurements in 11 girls and 16 boys: Pothoff and Roy, 1964; Little and Rubin, 1987

Girl	Age (years)				Boy	Age (years)			
	8	10	12	14		8	10	12	14
1	210	200	215	230	1	260	250	290	310
2	210	215	240	255	2	215		230	265
3	205		245	260	3	230	225	240	275
4	235	245	250	265	4	255	275	265	270
5	215	230	225	235	5	200		225	260
6	200		210	225	6	245	255	270	285
7	215	225	230	250	7	220	220	245	265
8	230	230	235	240	8	240	215	245	255
9	200		220	215	9	230	205	310	260
10	165		190	195	10	275	280	310	315
11	245	250	280	280	11	230	230	235	250
					12	215		240	280
					13	170		260	295
					14	225	255	255	260
					15	230	245	260	300
					16	220		235	250

distance from the centre of the pituitary to the pterygomaxillary fissure (unit 10^{-4} m)

Exemple/Données Pothoff



Modèle/Ecriture

i : indice de l'individu $i = 1, \dots, I = 25$ (11F+16G-2outliers)

j : indice de la mesure à l'age t_j (8,10,12,14ans)

Formulation: modèle hiérarchique

$$1) y_{ij} = \underbrace{\theta_{i1} + \theta_{i2}(t_j - 8)}_{\eta_{ij}} + e_{ij}, \quad e_{ij} \sim_{\text{iid}} \mathcal{N}(0, \sigma_e^2)$$

$$2) \boldsymbol{\theta}_i = \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \underbrace{\alpha_0}_{\text{Garçon}} + \underbrace{\alpha}_{\text{Fille-Garçon}} x_i \\ \underbrace{\beta_0}_{\text{Garçon}} + \underbrace{\beta}_{\text{Fille-Garçon}} x_i \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \right]$$

$$x_i = I_{[i=\text{"Fille"}]}$$

Formulation: modèle "mixte classique"

$$y_{ij} = (\alpha_0 + \alpha x_i + \textcolor{red}{a}_i) + (\beta_0 + \beta x_i + \textcolor{red}{b}_i)(t_j - 8) + e_{ij}$$

$$y_{ij} = \underbrace{\alpha_0 + \beta_0(t_j - 8)}_{\text{profil moyen "garçon) \textbf{Fixe}}} + \underbrace{(\alpha + \beta(t_j - 8)) x_i}_{\text{écart "fille-garçon" \textbf{Fixe}}} + \underbrace{\textcolor{red}{a}_i + \textcolor{red}{b}_i(t_j - 8)}_{\text{profil individuel ajusté \textbf{Aléatoire}}} + e_{ij}$$

Mise en œuvre de l'analyse bayésienne

Echantillonnage de Gibbs = cas particulier de
Métropolis-Hastings avec acceptation certaine.

(Algo de Monte Carlo générant une chaîne de Markov
dont la distribution stationnaire est la distribution cible)

$\mathbf{X} = (X_1, \dots, X_i, \dots, X_I)$ générées ainsi

$$X_1^{(1)} \sim X_1 \mid x_2^{(0)}, \dots, x_i^{(0)}, \dots, x_I^{(0)}$$

$$X_2^{(1)} \sim X_2 \mid x_1^{(1)}, x_3^{(0)}, \dots, x_i^{(0)}, \dots, x_I^{(0)}$$

$$X_i^{(1)} \sim X_i \mid x_1^{(1)}, x_2^{(1)}, \dots, x_{i-1}^{(1)}, x_{i+1}^{(0)}, \dots, x_I^{(0)}$$

...

De façon générale

$$X_i^{(t)} \sim X_i \mid x_1^{(t)}, x_2^{(t)}, \dots, x_{i-1}^{(t)}, x_{i+1}^{(t-1)}, \dots, x_I^{(t-1)}$$

Winbugs/Openbugs

BUGS stands for Bayes Using Gibbs Sampling.

Software for performing Bayesian analyses via MCMC methods

WinBugs is the Windows version of BUGS

Developers

- MRC Biostatistics, Cambridge

- Imperial College, School of Medicine, London

Can be downloaded **freely** (registration for key)

Modèle/Code OpenBugs/

```
model {  
#Modèle hiérarchique  
# 1)Vraisemblance  
for (i in 1:I) {  
  for (j in 1:n){  
    Y[i, j] ~ dnorm(eta[i, j], tauR)  
    eta[i,j]<-theta[i,1]+theta[i,2]*(age[j]-8)  
  }  
#2a) Loi des theta  
theta[ i,1:2]~ dmnorm(mu[i,1:2], tau[1:2,1:2])  
mu[i,1]<-alphaG+alphaF_G*sex[i]  
mu[i,2]<-betaG+betaF_G*sex[i]  
}
```

Modèle/Code OpenBugs(suite)

#2b) Précision résiduelle

tauR <- pow(sigmaR, -2)

sigmaR ~ dunif(0, 50)

varR <- 1/tauR #sortie variance

#3a) Prior position : loi plate

alphaG ~ dflat() **alphaF_G** ~ dflat()

betaG ~ dflat() **betaF_G** ~ dflat()

#3b) prior dispersion (entre sujets)

tau[1:2, 1:2] ~ dwish(R[1:2, 1:2], 3)

var[1:2, 1:2] <- inverse(tau[1:2, 1:2]) #sortie

rho <- var[1, 2] / sqrt(var[1, 1] * var[2, 2]) #sortie

sigma_inter <- sqrt(var[1, 1]) #sortie

sigma_pente <- sqrt(var[2, 2]) #sortie

}

Data/Code OpenBugs(suite)

Data

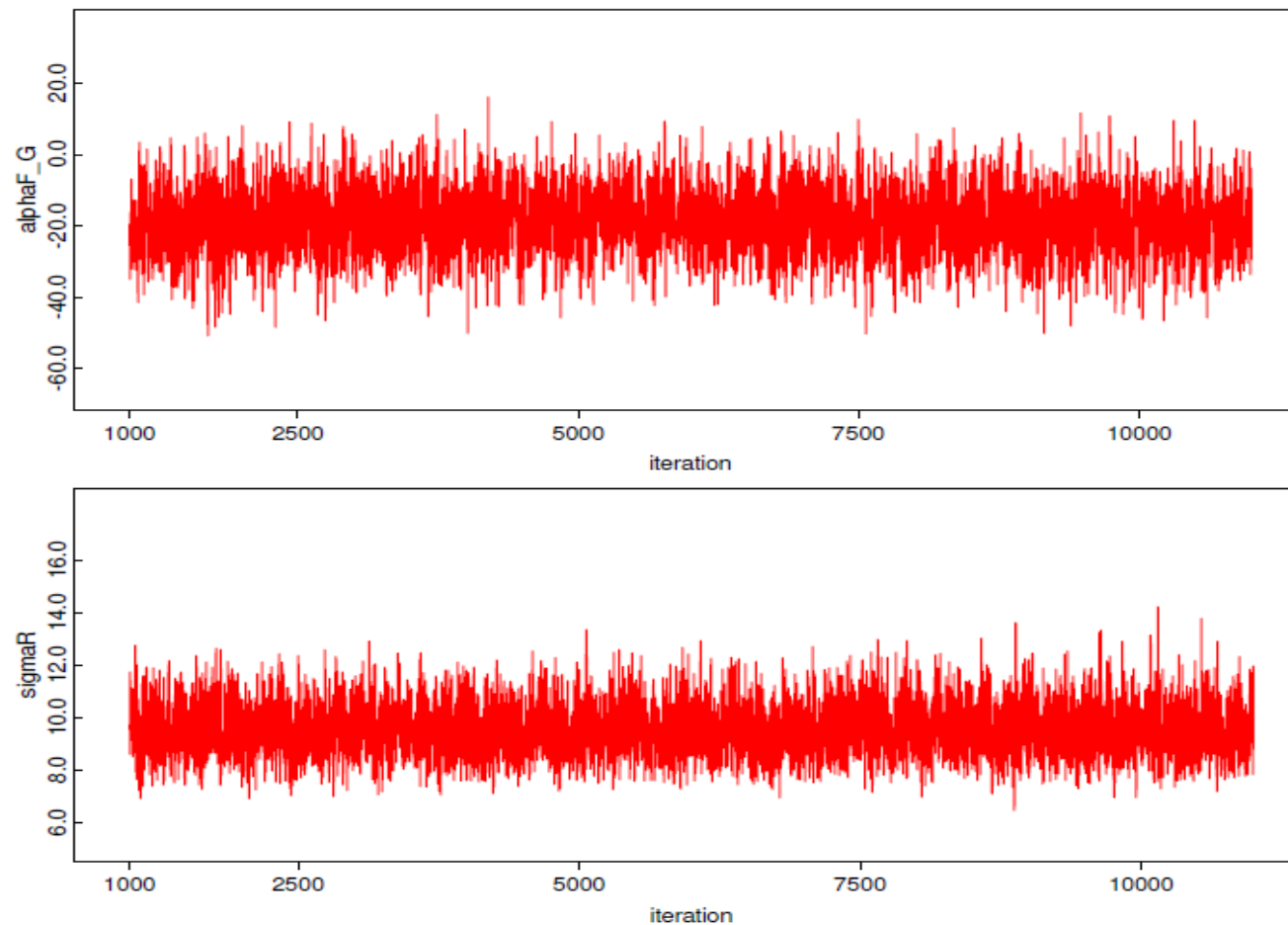
```
list(l=25,n=4,  
Y = structure(.Data = c(  
210,200,215,230,  
210,215,240,255,  
205,NA,245,260,  
235,245,250,265,  
215,230,225,235,  
200,NA,210,225,  
215,225,230,250,  
230,230,235,240,  
200,NA,220,215,  
165,NA,190,195,  
245,250,280,280,  
  
260,250,290,310,  
215,NA,230,265,  
230,225,240,275,  
255,275,265,270,  
200,NA,225,260,  
245,255,270,285,  
220,220,245,265,  
240,215,245,255,  
#230,205,310,260,  
275,280,310,315,  
230,230,235,250,  
215,NA,240,280,  
#170,NA,260,295,  
225,255,255,260,  
230,245,260,300,  
220,NA,235,250),  
.Dim = c(25, 4)),  
age=c(8, 10, 12,14),  
sex=c(  
1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),  
R=structure(.Data=c(800,0,0,10),.Dim=c(2,2)))  
⇐
```

Inits/Code OpenBugs(suite)

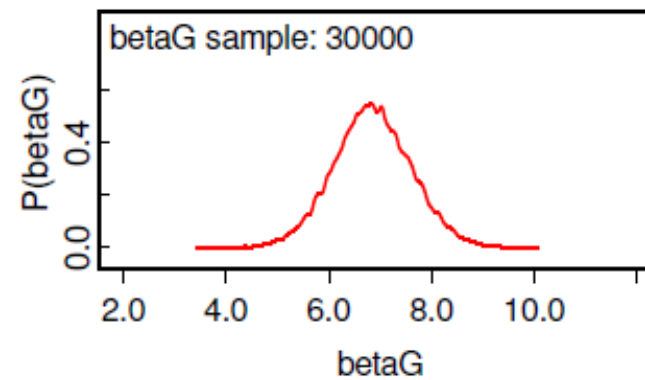
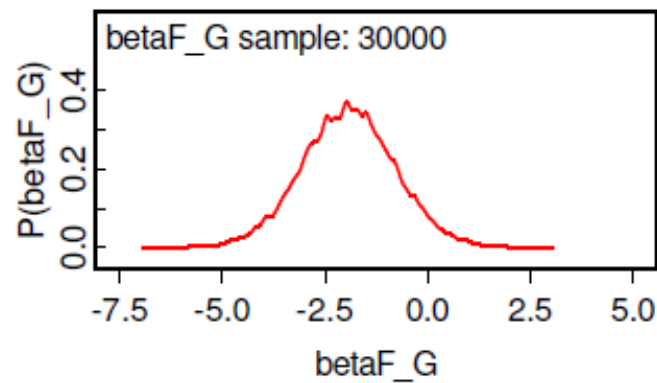
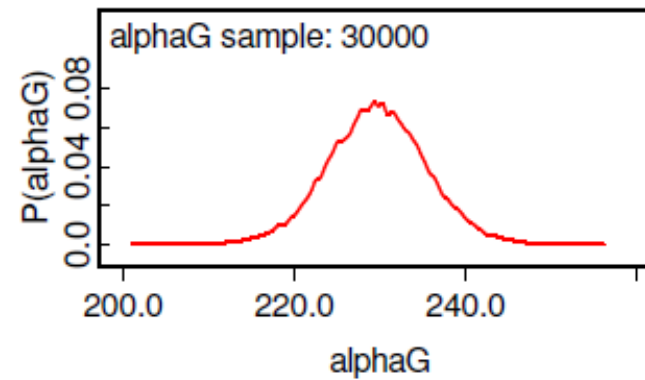
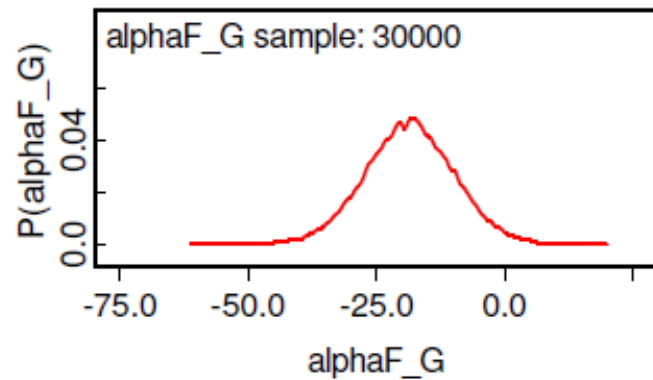
Inits

```
list(alphaG = 200, alphaF_G=0, betaG=10,  
betaF_G=0, sigmaR=20,  
tau = structure(.Data=c(2.5E-3,0,0,0.2),.Dim=c(2,2)))
```

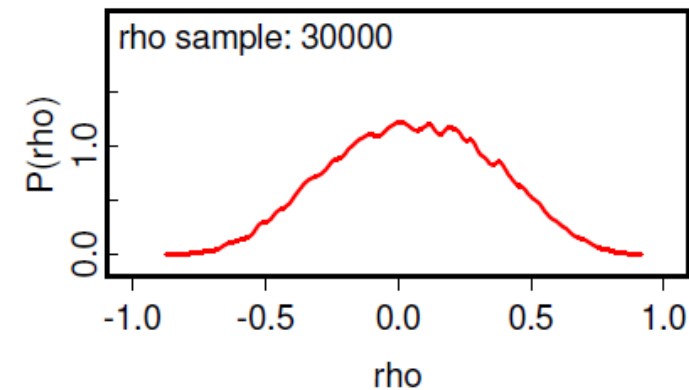
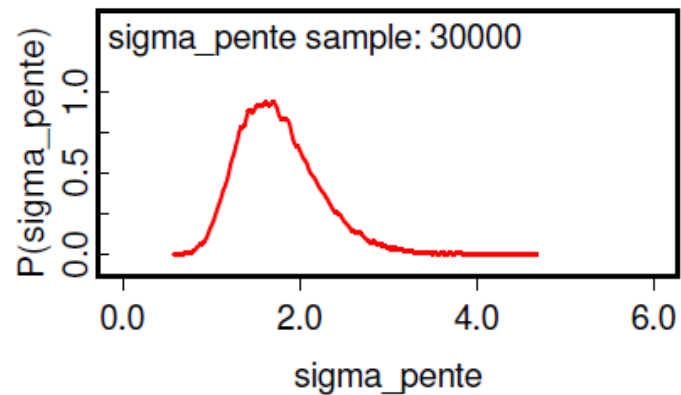
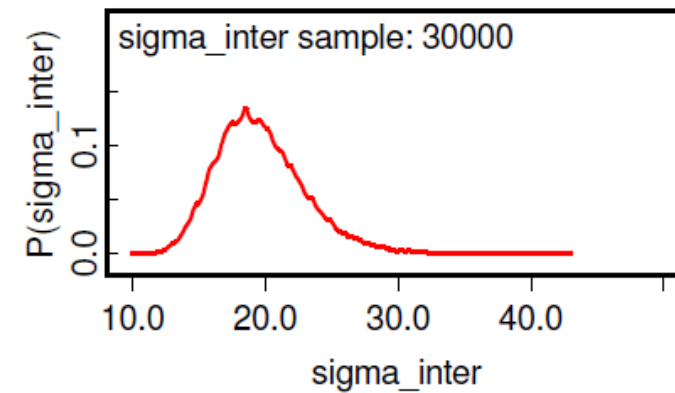
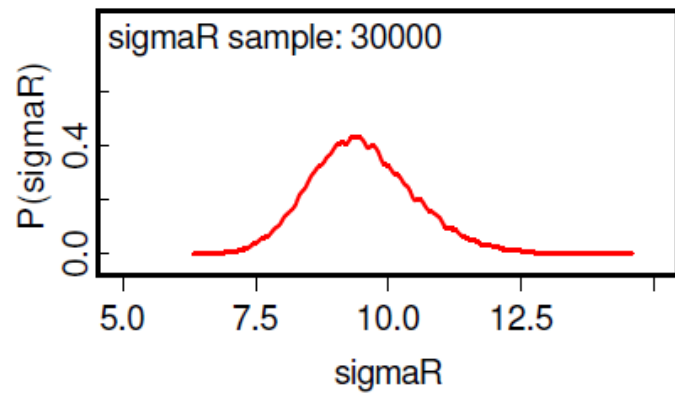
OpenBugs/Résultats/Historique



Code OpenBugs/Résultats/Densité



Code OpenBugs/Résultats/Densité (suite)



Code OpenBugs/Résultats/Statistiques

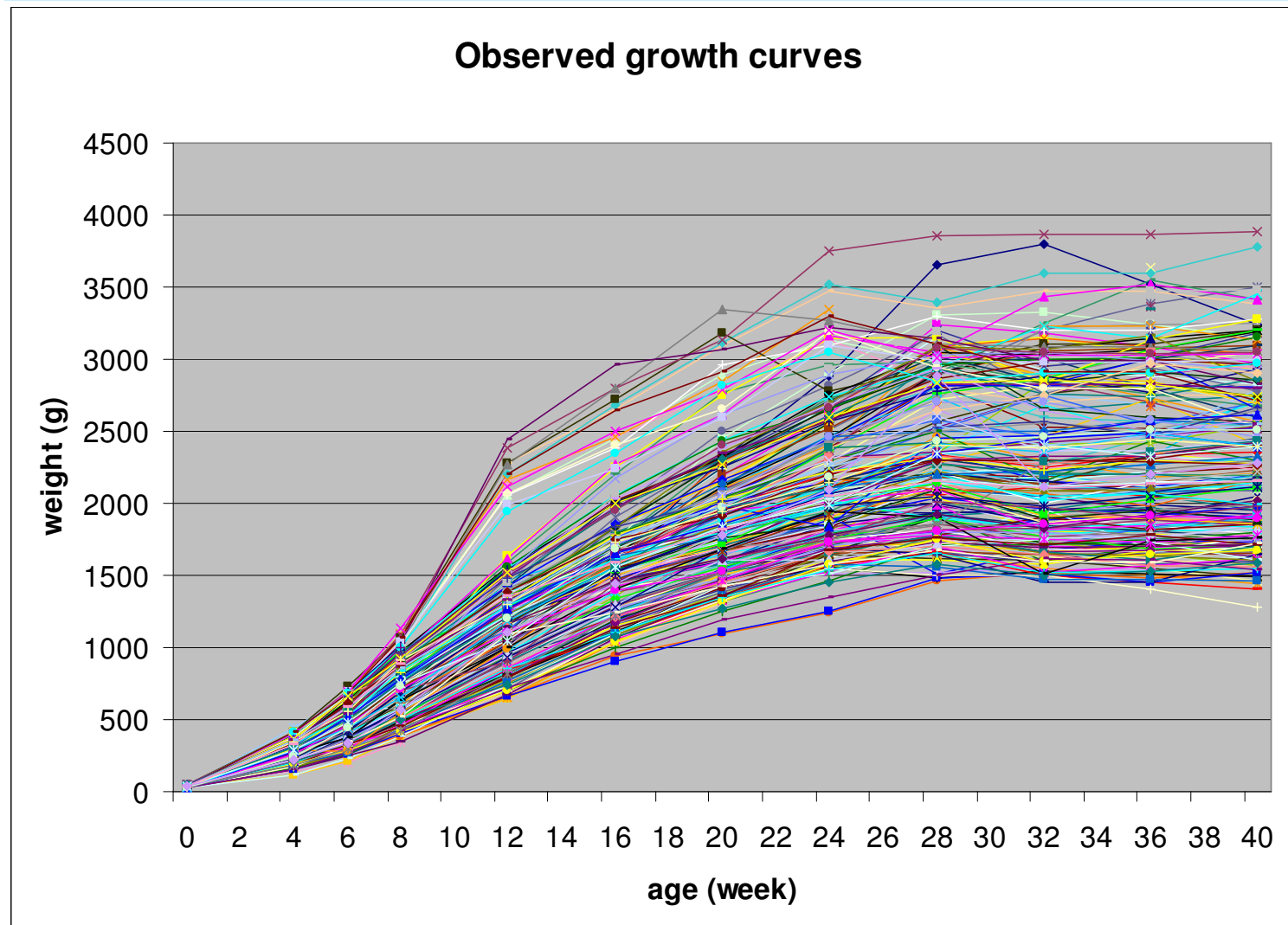
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
alphaF_G	-18.49	8.784	0.08733	-35.87	-18.43	-1.176	1001	40000
alphaG	229.6	5.807	0.05766	218.1	229.6	241.1	1001	40000
betaF_G	-1.92	1.143	0.01724	-4.175	-1.922	0.3412	1001	40000
betaG	6.848	0.7599	0.01205	5.359	6.839	8.364	1001	40000
rho	0.05471	0.3006	0.003281	-0.5162	0.05652	0.6147	1001	40000
sigmaR	9.529	0.9778	0.008937	7.794	9.46	11.64	1001	40000
sigma_inter	19.65	3.351	0.02534	14.15	19.3	27.26	1001	40000
sigma_pente	1.763	0.456	0.005156	1.043	1.709	2.806	1001	40000
tauR	0.01136	0.00231	2.067E-5	0.007375	0.01117	0.01647	1001	40000

	Dbar	Dhat	DIC	pD
Y	673.9	638.9	708.8	34.93
total	673.9	638.9	708.8	34.93

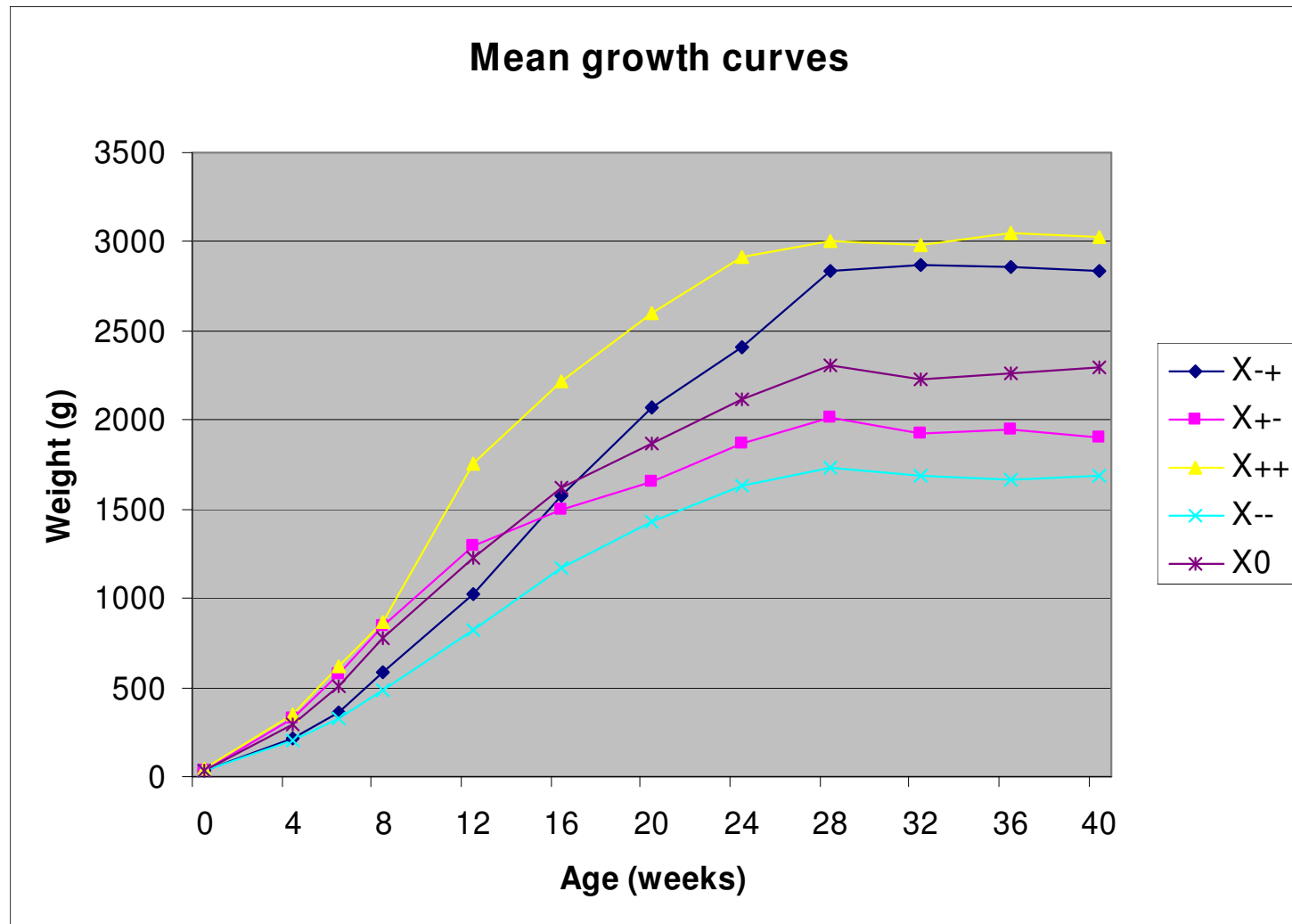
Exemple/Données de croissance

- Expérience de Ricard à l'INRA de sélection divergente sur le poids à 2 âges: 8 et 36 semaines
- 5 lignées:
 - $+-(\text{HL}); -+(\text{LH})$
 - $++(\text{HH}); --(\text{LL})$
 - Témoin (C)
- Données=3058 contrôles effectués sur 265 femelles à la dernière génération

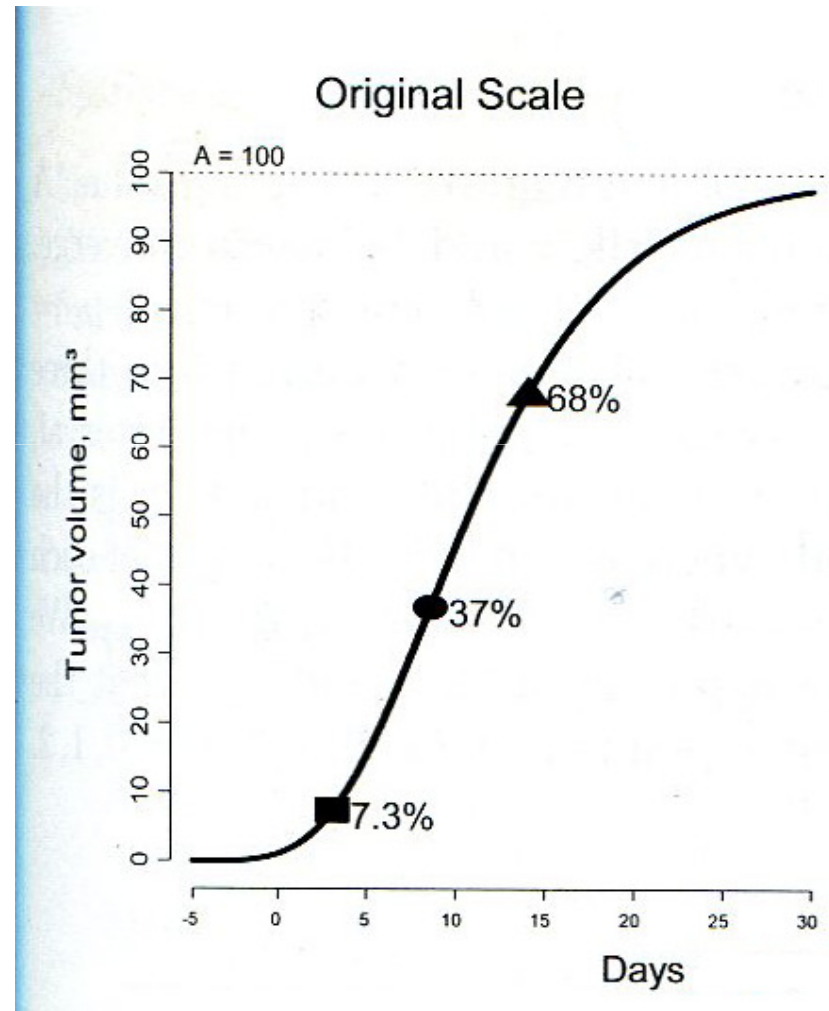
Croissance/Courbes individuelles



Croissance/Courbes par souche



Croissance/Fonction de Gompertz



Fonction de croissance

Gompertz $f(t) = A \exp[-B \exp(-Ct)]$

$$A = \lim_{t \rightarrow \infty} f(t) \quad B = -\ln(f(0)/A)$$

$$C = \underbrace{f'(t_I) / f(t_I)}_{\text{vit. croissance relative}} \quad t_I = \text{pt d'inflexion}$$

$$d^2 f(t) / dt^2 = 0 \Rightarrow t_I = (\ln B) / C$$

$$f(t_I) = \boxed{A/e} = 0.37 A$$

$$\ln \left[\ln \left(\underbrace{f(t)/A}_{\text{degré de maturité}} \right) \right] = \ln(-B) - Ct$$

Croissance/Sources de variation

- Entre souches (h)
- Entre individus intra souches (i)
- Entre mesures intra individus (j)

Croissance/Modèle hiérarchique non linéaire

$$1) y_{hij} | \boldsymbol{\theta}_{hi} \sim_{id} \mathcal{N} \left[f(t_j, \boldsymbol{\theta}_{hi}), \sigma_{e,hij}^2 \right]$$

h = souche; i = individu intra souche, j = mesure intra individu

$$2) \boldsymbol{\theta}_{hi} \sim_{id} \mathcal{N}(\mathbf{A}_{hi}\boldsymbol{\beta}, \boldsymbol{\Gamma}) \text{ où } \boldsymbol{\theta}_{hi} = (A_{hi}, B_{hi}, C_{hi})'$$

soit $k = hij$

$$\ln(\sigma_{e_k}^2) \sim \mathcal{N}(\mathbf{p}_k' \boldsymbol{\delta}, \gamma_k) \begin{cases} a) \gamma_k = 0 & \text{Effets fixes seuls} \\ b) \gamma_k = \mathbf{q}_k' \boldsymbol{\Omega} \mathbf{q}_k & \text{App. structurale mixte} \end{cases}$$

$$\ln(\sigma_{e_k}^2) = \mathbf{p}_k' \boldsymbol{\delta} + \mathbf{q}_k' \mathbf{v} \text{ avec } \mathbf{v} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$$

3) $\boldsymbol{\beta}, \boldsymbol{\Gamma}$ comme précédemment

$$a) \pi(\boldsymbol{\delta}) = 1$$

$$b) \boldsymbol{\Omega} \sim \text{Inv-Wishart}$$

Croissance/Modélisation des variances

Modèles

$$y_{hij} = A_{hi} \exp \left[-B_{hi} \exp(-C_{hi} T_j) \right] + e_{hij}$$

avec $T_j = t_j / 100$ (multiplication du coef C par 100)

$$A_{hi} = \underbrace{\alpha_h}_{\substack{\text{strain} \\ \text{(fixed)}}} + \underbrace{a_{hi}}_{\substack{\text{animal within strain} \\ \text{(aléatoire)}}} \quad a_{hi} \sim_{iid} \mathcal{N}(0, \sigma_a^2) \Leftrightarrow A_{hi} \sim_{iid} \mathcal{N}(\alpha_h, \sigma_a^2)$$

$$B_{hi} = \beta_h + b_{hi} \quad b_{hi} \sim_{iid} \mathcal{N}(0, \sigma_b^2) \Leftrightarrow B_{hi} \sim_{iid} \mathcal{N}(\beta_h, \sigma_b^2)$$

$$C_{hi} = \gamma_h + c_{hi} \quad c_{hi} \sim_{iid} \mathcal{N}(0, \sigma_c^2) \Leftrightarrow C_{hi} \sim_{iid} \mathcal{N}(\gamma_h, \sigma_c^2)$$

$$\sigma_{ab} = \text{Cov}(a_{hi}, b_{hi}) \quad \sigma_{ac} = \text{Cov}(a_{hi}, c_{hi}) \quad \sigma_{bc} = \text{Cov}(b_{hi}, c_{hi}) ;$$

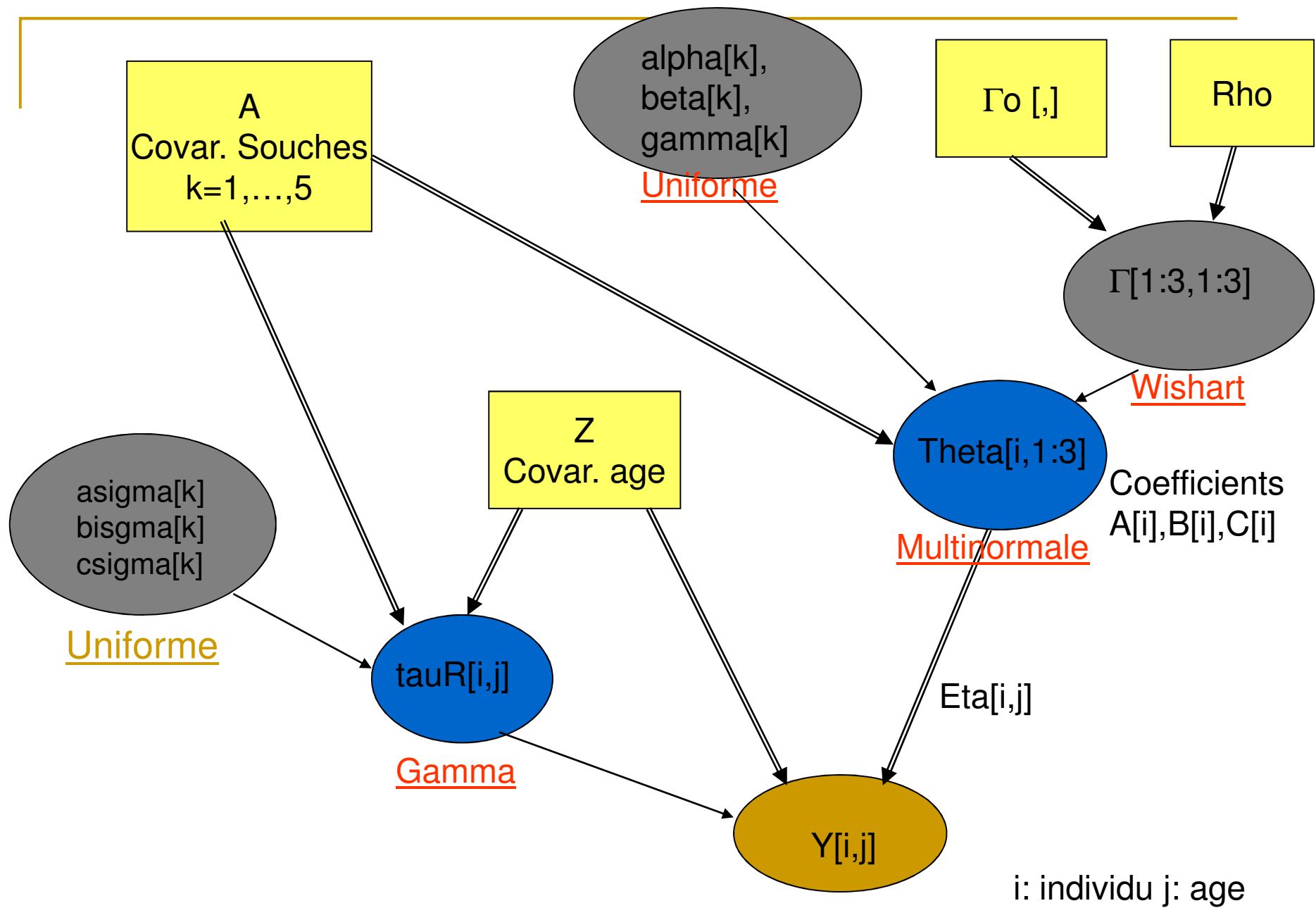
$$e_{hij} \sim_{iid} \mathcal{N}(0, \sigma_{e_{t(j)}}^2)$$

$h = 1(MP), 2(PM), 3(PP), 4(MM), 5(Témoin C)$

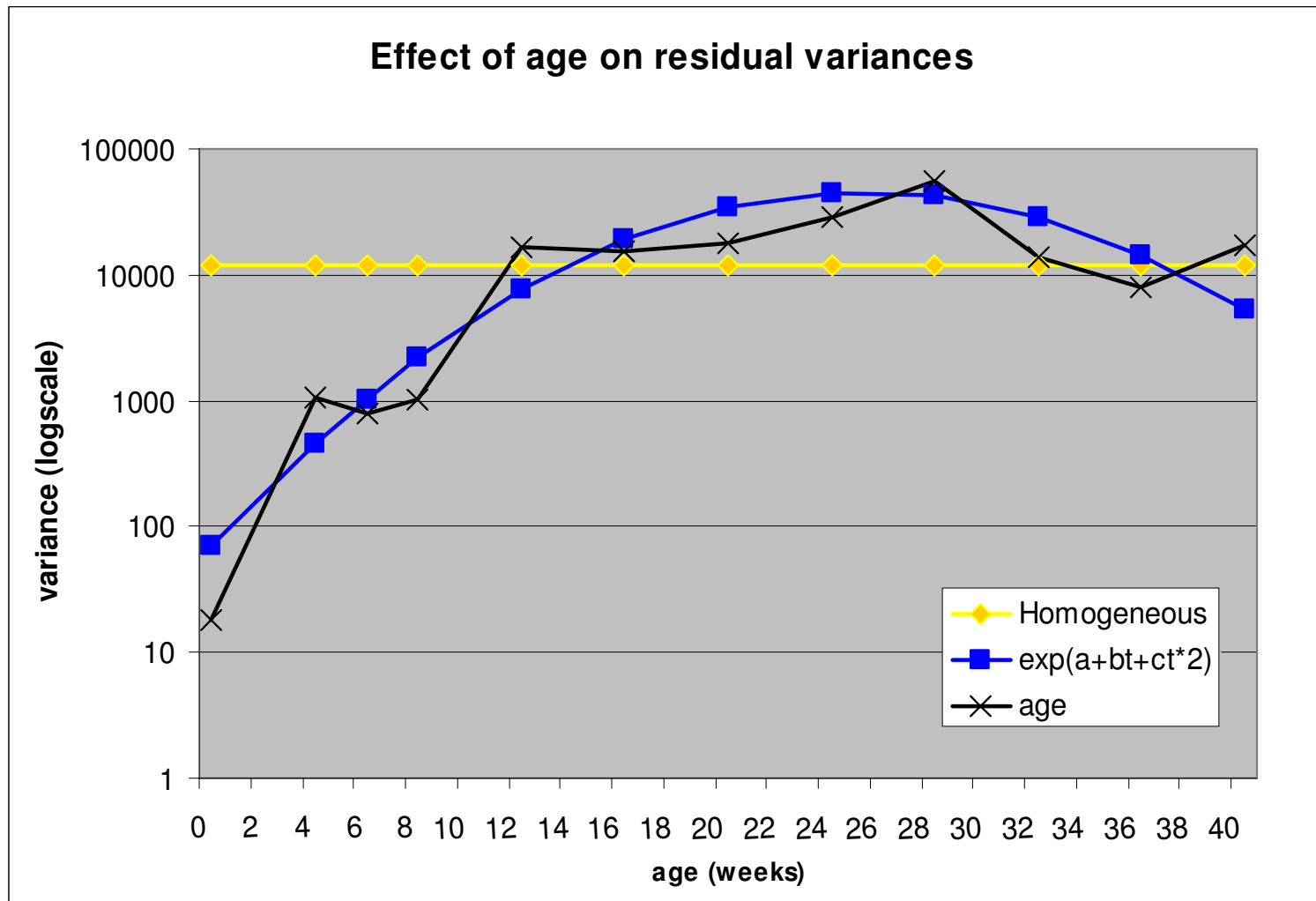
Modèle de variance résiduelle par souche

$$\sigma_{e_{ht}} = \exp \left(\underline{a}_h + \underline{b}_h t^* + \underline{c}_h t^{*2} \right) \quad t^* = (t - 20) / 100 \text{ où } t = \text{age(sem)}$$

Ici $\underline{a} = \text{asigma}$, $\underline{b} = \text{bsigma}$, $\underline{c} = \text{csigma}$



Croissance/Modèle hiérarchique non linéaire/var résiduelle



Comparaison de modèles/DIC

Déviance marginale

$$-D_m = -2 \log [m(\mathbf{y})]; \quad m(\mathbf{y}) = \int f(\mathbf{y} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Difficile à calculer: Foulley, Applibugs, 2010

$$- \boxed{DIC = D(\bar{\boldsymbol{\theta}}) + 2 \underbrace{p_D}_{\text{Complexité}}}$$

$$\boxed{p_D = \bar{D} - D(\bar{\boldsymbol{\theta}})}$$

Spiegelhalter et al (2002), JRSS

Analogie avec $\boxed{AIC = D(\hat{\boldsymbol{\theta}}_{ML}) + 2p}$

$D(\boldsymbol{\theta}) = -2 \ln p(\mathbf{y} | \boldsymbol{\theta}) + cste$: "déviance"

$$\bar{\boldsymbol{\theta}} = E(\boldsymbol{\theta} | \mathbf{y}); \quad \bar{D} = E_{\boldsymbol{\theta} | \mathbf{y}} [D(\boldsymbol{\theta})]$$

Croissance/Modélisation des variances/comparaison

Model	#par	pD	DIC	Gain in DIC
Fixed/Homogeneous	16	16	40710	-4013
Mixed/Homogeneous	22	502	36607	0
Exp(Age Quadra*Strain)	36	547	34227	2380
Constant+Power	24	552	34402	2205
Exp(Age Quadra*Strain+ Random Inter)	37	674	33987	2620 (+240)

Détection de loci sélectionnés

Modèle hiérarchique à plusieurs niveaux

i =locus; j =(sous)population

a_{ij} =Nb de gènes d'un type allélique au locus i dans la pop j

p_{ij} = Fréquence de cet allèle au locus i dans la pop j

$$1) \boxed{a_{ij} \mid p_{ij} \sim B(n_{ij}, p_{ij})}$$

$$2) \boxed{p_{ij} \mid x_i, \lambda_{ij} \sim \text{Beta}(\lambda_{ij} x_i, \lambda_{ij} (1 - x_i))}$$

$E(p_{ij}) = x_i$ (fréquence dans le gene-pool)

$$\text{Var}(p_{ij}) = x_i (1 - x_i) \underbrace{(1 + \lambda_{ij})^{-1}}_{F_{ij}}$$

Marginale après intégration de p_{ij} : $a_{ij} \mid x_i, \lambda_{ij} \sim \text{Beta - Binomiale}$

Détection de loci sélectionnés

3) $x_i \sim \text{Beta}(a_x, b_x)$ par ex $a_x = b_x = 1$

$$\eta_{ij} = \ln F_{ij} / (1 - F_{ij}) = \underbrace{\mu}_{1/\lambda_{ij}} + \underbrace{\alpha_i}_{\text{locus}} + \underbrace{\beta_j}_{\text{population}} + \underbrace{\gamma_{ij}}_{\text{écart additivité}}$$

$$\eta_{ij} \mid a_i, b_j, \sigma_\gamma^2 \sim_{id} \mathcal{N} \left(\underbrace{a_i + b_j}_{\mu + \alpha_i + \beta_j}, \sigma_\gamma^2 \right)$$

avec $a_i = \delta_i a_i^*$ $\delta_i = I(\text{locus } i \text{ sélectionné})$

4) $b_j \mid \mu, \sigma_\beta^2 \sim_{iid} \mathcal{N}(\mu, \sigma_\beta^2)$ $a_i^* \sim_{iid} \mathcal{N}(0, \sigma_{\alpha^*}^2)$

$\delta_i \sim \text{Ber}(P)$ $\sigma_\gamma^2 \sim$ (cf (5ème niveau))

5) $\pi(\mu) = 1$, $\sigma \sim U(0, A)$, $P \sim \text{Beta}(u_1, u_2)$

Application au génome bovin

- 9 populations bovines ouest africaines
- 36320 SNP
- 53 régions sous sélection

BMC Genomics



Research article

Open Access

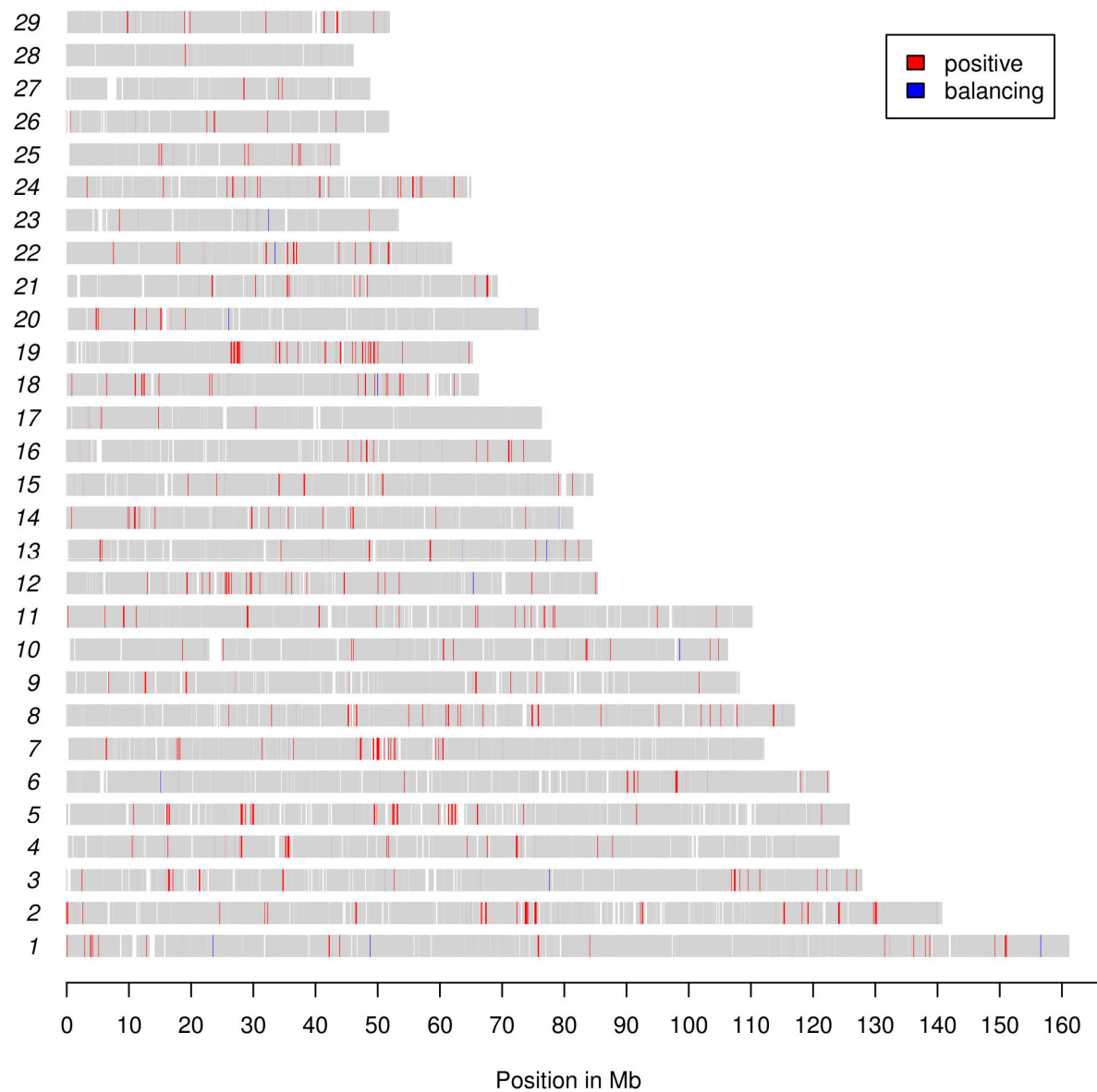
A whole genome Bayesian scan for adaptive genetic divergence in West African cattle

Mathieu Gautier^{*1}, Laurence Flori¹, Andrea Riebler², Florence Jaffrézic¹, Denis Laloé¹, Ivo Gut³, Katayoun Moazami-Goudarzi¹ and Jean-Louis Foulley¹

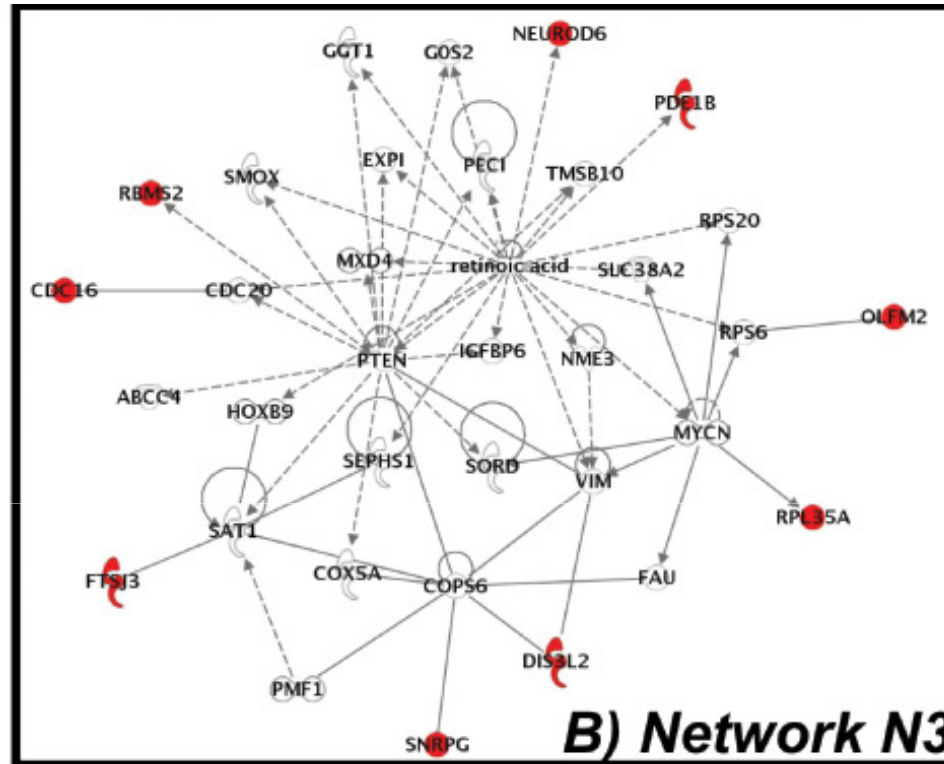
Génome bovin/régions sous sélection

Table 3: Regions under selection identified at the 5% local FDR (q-value) thresholds.

Interval	BT A	Position (size) in Mb	Peak Position in Mb	BF value of the SNP at the peak position (q-value of the smoothed signal)	Gene at or close to the peak
1	1	3.654-12.74 (9.087)	3.923	50 (0.0467)	KRTAP8-1 (3.888-3.888)
2	1	133.1-138.6 (5.465)	138.1	51 (0.0345)	BFSP2 (138.1-138.2)
3	2	0.035-2.687 (2.652)	0.035	51 (0.0001)	NA
4	2	64.04-67.28 (3.235)	64.25	31 (0.0258)	CXCR4 (64.22-64.22)
5	2	71.48-78.32 (6.84)	73.84	51 (0.0001)	SNRPG (73.75-73.75)
6	2	117.7-129.0 (11.35)	124.1	45 (0.0007)	DIS3L2 (123.844-124.15)
7	2	140.0-140.6 (0.662)	140.0	25 (0.0473)	PADI4 (139.9-139.9)
8	3	12.39-19.72 (7.325)	15.96	51 (0.0007)	SEMA4A (15.94-15.96)
9	3	107.9-110.1 (2.219)	108.3	37 (0.0246)	RNF220 (108.2-108.3)
10	4	24.74-29.70 (4.953)	25.47	38 (0.0002)	NA
11	4	32.31-37.17 (4.860)	35.71	36 (0.0009)	NA
12	4	46.84-54.10 (7.259)	51.86	26 (0.0017)	EPDR1 (51.77-51.81)
13	4	66.29-75.27 (8.977)	67.67	50 (0.0065)	NEUROD6 (67.70-67.70)
14	5	15.625-16.745 (1.121)	16.15	26 (0.0342)	NA
15	5	26.81-31.05 (4.245)	28.37	51 (0.0001)	PDE1B (28.34-28.39)



Génome bovin/Analyse de réseaux fonctionnels



S7 in additional file 2). Network N3 contained molecules mainly involved in cancer, cellular function and maintenance, and neurological disease, and among them nine candidate genes (CDC16, DIS3L2, FTSJ3, NEUROD6, OLFM2, PDE1B, RBMS2, RPL3A and SNRPG). Six other

Conclusion

- L'approche bayésienne hiérarchique est bien adaptée au traitement de structures de données et modèles complexes (linéaires, non linéaires & modèles mixtes)
- Elle augmente la robustesse de la loi a priori par intégration sur des paramètres inconnus
- Elle améliore aussi les performances de l'inférence d'un point de vue fréquentiste en réduisant l'arbitraire sur le choix de l'hyperparamètre
 - Estimateurs shrinkés, est. « minimax »
- Elle donne lieu à représentation graphique (DAG)
 - Accessibilité et meilleure compréhension
 - Indispensable pour le Gibbs sampling

Références

Berger JO (1980) Statistical decision theory and Bayesian analysis. 2nd Edition, Springer-Verlag

Carlin BP, Louis TA (2008) Bayesian methods for data analysis; Third edition. CRC Press

Gelman A, Hill J (2006) Data analysis using regression and multilevel (hierarchical) models. Cambridge University Press

Marin JM, Robert CP (2007) Bayesian core. Springer Verlag

Parent E, Bernier J (2007) Le raisonnement bayésien. Springer Verlag

Robert CP, Casella G (2004) Monte Carlo statistical methods, 2nd Edition, Springer-Verlag

Annexes

Modèle/Ecriture bayésienne hiérarchique

i : indice de l'individu $i = 1, \dots, I = 25$ j : indice de la mesure à l'age t_j

$$\text{1er niveau : } \mathbf{y}_i = (y_{ij}) \sim \mathcal{N} \left(\underbrace{\mathbf{Z}_i}_{4(3) \times 2} \underbrace{\boldsymbol{\theta}_i}_{2 \times 1}, \mathbf{R}_i = \sigma^2 \mathbf{I}_{n_i} \right)$$

$$\text{2}^{\text{ème}} \text{ niveau : } \boldsymbol{\theta}_i = \begin{pmatrix} \boldsymbol{\theta}_{i1} \\ \boldsymbol{\theta}_{i2} \end{pmatrix} \sim \mathcal{N}_2 \left(\underbrace{\mathbf{A}_i}_{2 \times 4} \underbrace{\boldsymbol{\beta}}_{4 \times 1}, \underbrace{\mathbf{G}}_{2 \times 2} \right)$$

$$\sigma^2 \sim \mathcal{IG} \left(\frac{1}{2} \nu, \frac{1}{2} \nu \sigma^{*2} \right) \text{ avec } \nu=2 \text{ par ex}$$

3^{ème} niveau : $\pi(\boldsymbol{\beta}) = 1$: $\boldsymbol{\beta}$ traité comme fixe

ou $\beta_k \sim U(\beta_{k,\text{inf}}, \beta_{k,\text{sup}})$

$$\mathbf{G} \sim \mathcal{IW} \left[(\rho \mathbf{G}^*)^{-1}, \rho \right] \text{ (Wishart-inverse)}$$

Rappel Wishart

$$u_i^* \sim_{iid} \mathcal{N}(0,1) \Rightarrow x^* = \sum_{i=1}^n u_i^{*2} \sim \chi_n^2 \sim \mathcal{G}(1/2n, 1/2)$$

$$u_i \sim_{iid} \mathcal{N}(0, \sigma^2) \Rightarrow x = \sum_{i=1}^n u_i^2 \sim \sigma^2 \chi_n^2 \sim \mathcal{G}(1/2n, 1/2\sigma^{-2})$$

$$\mathbf{U}_i \sim_{iid} \mathcal{N}_p(\mathbf{0}, \Sigma) \text{ and } \mathbf{X} = \sum_{i=1}^n \mathbf{U}_i \mathbf{U}_i'$$

$$\mathbf{X} | \Sigma, n \sim \mathcal{W}(\Sigma, n) \Leftrightarrow [\mathbf{X}; \Sigma, n] \propto |\mathbf{X}|^{1/2(n-p-1)} \exp\left[-1/2 \text{tr}(\mathbf{X}\Sigma^{-1})\right]$$

$$\mathbf{E}(\mathbf{X}) = n\Sigma; \text{ } n \text{ degrés de liberté } n \geq p = \dim \Sigma$$

$$\text{Var}(x_{ij}) = n(\sigma_{ij}^2 + \sigma_{ii}\sigma_{jj}); \text{Cov}(x_{ij}, x_{kl}) = n(\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk})$$

Prior Wishart Inverse

$\mathbf{G}_{p \times p}$ non diagonale

Prior classique: Wishart inverse soit si l'on pose $\mathbf{\Omega} = \mathbf{G}^{-1}$

$$\mathbf{\Omega} \sim \mathcal{W}\left[(\nu \mathbf{G}_0)^{-1}, \nu\right] \Leftrightarrow [\mathbf{\Omega} | \mathbf{G}_0, \nu] \propto |\mathbf{\Omega}|^{1/2(\nu-p-1)} \exp\left[-1/2\nu \text{tr}(\mathbf{G}_0 \mathbf{\Omega})\right]$$

$$\mathbb{E}(\mathbf{G}^{-1}) = \mathbf{G}_0^{-1} \quad \mathbb{E}(\mathbf{G}) = \frac{\nu}{\nu - p - 1} \mathbf{G}_0 \quad \text{mode}(\mathbf{G}) = \frac{\nu}{\nu + p + 1} \mathbf{G}_0$$

$\mathbf{G}_0$: par de position; ν : par de dispersion ν petit $\Rightarrow$ variance élevée

Pour $\nu = 0$ $[\mathbf{G}] \propto |\mathbf{G}|^{-1/2(p+1)}$ (Prior de Jeffreys)

Attention: autre notation $\mathbf{G}^{-1} \sim \mathcal{W}\left[\underbrace{(\nu \mathbf{G}_0)}_{\text{"R"}}, \nu\right]$ (Winbugs)

Rappel: Jacobien de la transformation $\mathbf{A}_{p \times p} = \mathbf{B}^{-1}$ est $|\mathbf{B}|^{-(p+1)}$

Mod linéaire/Gibbs/Lois conditionnelles

Rappel: $\boldsymbol{\theta}_i \mid \boldsymbol{\beta}, \mathbf{G}_0 \sim \mathcal{N}(\mathbf{A}_i \boldsymbol{\beta}, \mathbf{G}_0)$

$$\boldsymbol{\theta}_i \mid \mathbf{y}, \boldsymbol{\beta}, \sigma^2, \mathbf{G}_0 \sim \mathcal{N}(\tilde{\boldsymbol{\theta}}_i, \tilde{\mathbf{V}}_{\boldsymbol{\theta}_i})$$

$$\tilde{\boldsymbol{\theta}}_i = \mathbf{A}_i \boldsymbol{\beta} + \mathbf{G}_0 \mathbf{Z}_i' \mathbf{V}_i^{-1} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{A}_i \boldsymbol{\beta})$$

$$\text{avec } \mathbf{V}_i = \sigma^2 \mathbf{I}_{n_i} + \mathbf{Z}_i \mathbf{G}_0 \mathbf{Z}_i'$$

$$\tilde{\mathbf{V}}_{\boldsymbol{\theta}_i} = \mathbf{G}_0 - \mathbf{G}_0 \mathbf{Z}_i' \mathbf{V}_i^{-1} \mathbf{Z}_i \mathbf{G}_0$$

$$\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\theta}_i (i = 1, \dots, I), \sigma^2, \mathbf{G}_0 \sim \mathcal{N}(\tilde{\boldsymbol{\beta}}, \tilde{\mathbf{V}}_{\boldsymbol{\beta}})$$

$\tilde{\boldsymbol{\beta}}$ solution de

$$\left(\sum_{i=1}^I \mathbf{A}_i' \mathbf{G}_0^{-1} \mathbf{A}_i \right) \tilde{\boldsymbol{\beta}} = \sum_{i=1}^I \mathbf{A}_i' \mathbf{G}_0^{-1} \mathbf{y}_i$$

$$\tilde{\mathbf{V}}_{\boldsymbol{\beta}} = \left(\sum_{i=1}^I \mathbf{A}_i' \mathbf{G}_0^{-1} \mathbf{A}_i \right)^{-1}$$

Mod linéaire/Gibbs/Lois conditionnelles

$$\sigma^2 \mid \mathbf{y}, \boldsymbol{\theta}_i (i=1, \dots, I), \boldsymbol{\beta}, \mathbf{G}_0$$

$$\sim \mathcal{IG} \left(\frac{1}{2}(\nu + N), \frac{1}{2} \left[\left(\sum_{i=1}^I \sum_{j=1}^{n_i} (y_{ij} - \mathbf{a}_{ij}' \boldsymbol{\beta})^2 \right) + \nu \sigma_*^2 \right] \right)$$

$$\mathbf{G}_0 \mid \mathbf{y}, \boldsymbol{\theta}_i (i=1, \dots, I), \boldsymbol{\beta}, \sigma^2$$

$$\sim \mathcal{IW} \left[\left(\sum_{i=1}^I (\boldsymbol{\theta}_i - \mathbf{A}_i \boldsymbol{\beta})(\boldsymbol{\theta}_i - \mathbf{A}_i \boldsymbol{\beta})' + \rho \mathbf{G}_0^* \right)^{-1}, I + \rho \right]$$