Toxicogenomic dose-response models for DNA chips data from rats treated by flutamide

Olivier Colomban [1,2], Bénédicte Naudet [3], Délphine Maucort-Boulch [3], Pascal Roy [3], Michel Tod [2], Pascal Girard [2]

[1] Université Claude Bernard Lyon 1, Lyon, France; [2] EA3738 CTO, Faculté de Médecine Lyon-Sud, F-69600 Oullins, France; [3] Hospices Civils de Lyon, Centre Hospitalier Lyon Sud, Service de Biostatistique, F-69310, Pierre-Bénite, France.

Background and Objectives

- The use of genomic technology for assessing health risks associated which chemical exposure has great potential.
- **Objectives**: To fully characterize testicular toxicity in adult Wistar rats induced by flutamide (FLU), to estimate the benchmark doses (BMD), and to estimate the BMD lower confidence limit (BMDL) modifying gene expression [1].
- To achieve this objective, changes in toxicogenomic responses (gene behavior) in the testes, will be investigated on 43,379 genes (full-genome analysis) in rats exposed to FLU at different dose levels by oral gavage for 28 consecutive days.

Methods

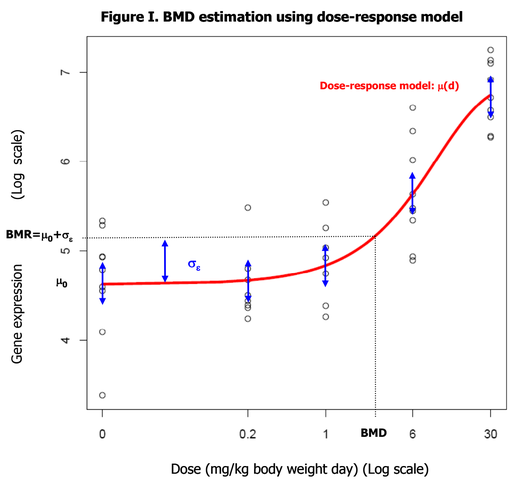
- **42** rats were randomized between 5 arms: **9** rats in vehicule (control group), **8** rats at 0.2, **7** rats at 1, **9** rats at 6 and **9** rats at 30 mg/kg body weight per day.
- For each rat, microarray (Agilent 4x44k) was made.
- **Linear model (M0) was applied to detect** significant change of log(gene-expression) from baseline. False discovery rate was controlled during this step for the slope of linear model (α=20%). [2]
- When significant, non-linear models were tested: **stimulation** or **inhibition** of gene expression (exp., Emax, logistic; see Table I).[3]

Table I. NON-LINEAR MODELS (D=dose)

MODEL	MODEL DESCRIPTION: $\mu_i(D)$	BEHAVIOR	NAME
Exponential	$a_i \times (b_i + (1-b_i) \times e^{(-k_i \times D)}) + \varepsilon_i$	Stimulation or inhibition	M1
Emax	$a_i + \frac{G_{max_i} \times D}{\sqrt{D50_i^2 + D}} + \varepsilon_i$	Stimulation	M2
	$a_i - \frac{G_{max_i} \times D}{\sqrt{D50_i^2 + D}} + \varepsilon_i$	Inhibition	M3
Logistic	$\frac{G_{max_i}}{1 + e^{-\frac{(D-c_i)}{k_i}}} + \varepsilon_i$	Stimulation	M4
	$\frac{a_i}{1 + b_i \times e^{-k_i \times D}} + \varepsilon_i$	Inhibition	M5

- **Model selection**: For each gene, choice between linear and non-linear models was based on Schwarz criterion (BIC).
- **BMD estimation**: defined as the dose level leading to a change in predicted baseline $\pm$ 1 residual SD: **BMR=$\mu_0 \pm \sigma$ ➔ BMD=μ^{-1}(BMR)**[1]. When normality of residual was rejected (Kolmogorov-Smirnov or Shapiro-Wilk test, α=5%), a bootstrap method was used to estimate the quantile of the BMR distribution.

Figure I. BMD estimation using dose-response model



- **BMDL** : intersection with the **BMR** and the upper (or lower for inhibition) 90% CI on the model. BMDL was estimate by **BMD-$z_{95\%}$xSD** with **z ~ N(0,1)**[4]; we used the *delta-method* to compute SD and to evaluate CI on the model.
- All analysis were performed using R 2.10® software, for non-linear models with *nls( )* function; for the delta-method we used *deltamethod( )* function.

Results

- Genes data were log normalized (quantile normalisation[5]) and after a QC control 32,944 genes were retained.

Linear dose-response relationship

- A significant linear change from baseline in gene behavior was detected for **6,343** genes (19.2%): **3,304** stimulations (52%) and **3,039** inhibitions (48%).

Non-Linear dose-response relationship

- Non linear model for stimulation behavior:

Table IIa. NON-LINEAR STIMULATION MODELS

MODEL	NAME	CONVERGENCE	# STIMULATION
Exponential	M1	2,352 (71.2%)	3,304
Emax	M2	2,706 (81.9%)	
Logistic	M4	2,697 (81.6%)	

For 2,325 genes, it was possible to achieve convergence for the 3 non-linear models and at least one model for 2,732 genes.

- Non linear model for inhibition behavior:

Table IIb. NON-LINEAR INHIBITION MODELS

MODEL	NAME	CONVERGENCE	# INHIBITION
Exponential	M1	2,336 (76.9%)	3,039
Emax	M3	2,451 (80.7%)	
Logistic	M5	2,440 (80.3%)	

For 2,180 genes, it was possible to achieve convergence for the 3 non-linear model and at least one model for 2,598 genes.

Model selection:

Table III. CHOICE OF THE BEST MODEL (Shwarz criterion)

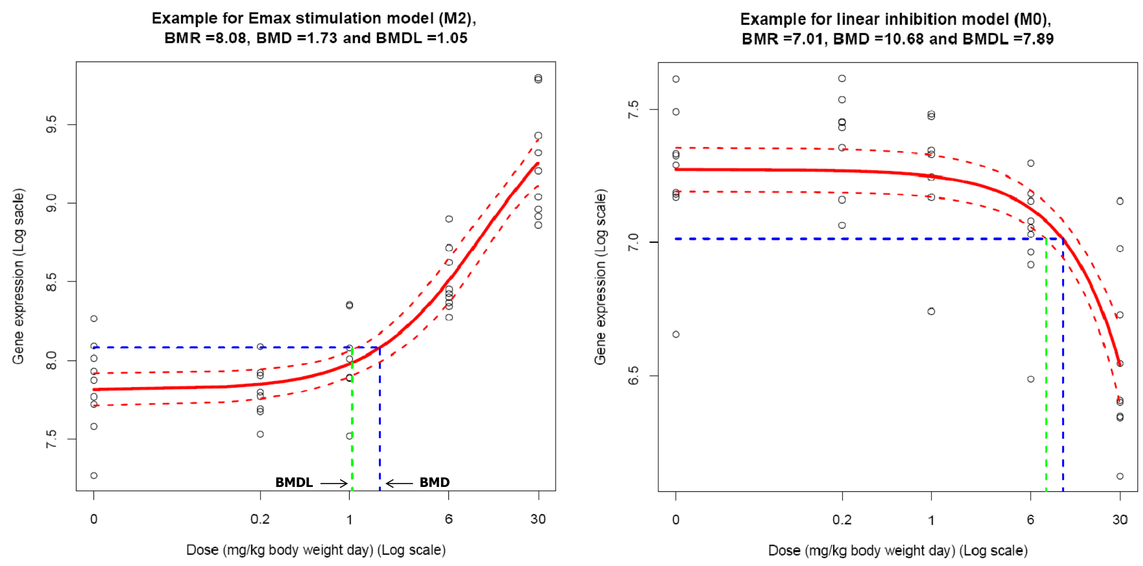
	Linear	Exponential	Emax		Logistic		TOTAL
	M0	M1	M2	M3	M4	M5	
STIMULATION	2,159	80	532	-	533	-	3,304
	65.4%	*2.4%*	*16.1%*	-	*16.1%*		
INHIBITION	2,484	182	-	303	-	70	3,039
	81.7%	*6%*	-	*10%*	-	*2.3%*	
TOTAL	**4,643**	**262**	**835**		**603**		6,343
	73.2%	*4.1%*	*13.2%*		*9.5%*		

- Linear and Emax models were the main preferred significant models, over exponential and logistic shape.

BMD and BMDL estimation:

- For 96.8% of estimated models, residuals were normal and for 3.2% bootstrap was used to estimate the quantile.
- Example for 2 genes:

Figure II. BMD analysis for stimulation and inhibition behavior



- Good estimation of the BMD at the intersection but not all the time.

Conclusion and perspectives

- An algorithm has been created to model dose-effect relationship toxicity expressed by DNA chips.
- This algorithm allows to characterizing the benchmark doses for a large set of genes.
- Running time : ~ 6 hours.

References

(1) Russel S. Toxicological Sciences, 98: 240-248; 2007; (2) Benjamini Y, Hochberg Y: Crit Rev Toxicol 57: 289-300; 2003. (3) Benchmark Dose Software (BMDS) Version 2.1, User's Manuel; 2009. (4) Moerbeek M. Risk Analysis, 24: 2004. (5) B. M. Bolstad, BIOINFORMATICS; 2003.