Étude expérimentale de l'influence des protéodies sur le développement foliaire chez *Arabidopsis thaliana* (L.) Heynh. et sur la synthèse d'α-amylase chez *Hordeum vulgare* L.

Rémi Tournebize, projet de statistique, première année ingénieur agronome

OBJECTIF

La présente étude vise à apprécier la validité de l'influence des protéodies aux cours d'expériences contrôlées sur *Arabidopsis thaliana* écotype Columbia et *Hordeum vulgare*. Il s'agit en particulier de mettre à l'épreuve la théorie d'une régulation épigénétique telle que formulée par M. Joël Sternheimer dans un brevet européen déposé le 2 juin 1993. Il ne sera donc pas question ici de traiter en détail des questions fondamentales relatives à cette théorie, nous renvoyons pour cela aux travaux en physique quantique menés par l'auteur (Sternheimer, 2007). En quelques mots néanmoins, le mécanisme allégué reposerait sur une régulation de l'expression génétique au stade co-traductionnel grâce à une « mise en résonance » d'un phénomène bio-physique de type vibratoire par une séquence de fréquences sonores obtenues par transposition dans la gamme tempérée de la fréquence propre des acides aminés constitutifs de la protéine ciblée. Le brevet invoque en effet l'émergence de phénomènes quantiques au niveau de l'acide aminé conjugué à son ARNt du fait d'une stabilisation vis à vis de l'agitation thermique, stabilisation qui confèrerait à l'aa-ARNt des propriétés ondulatoires – ce point reste pour nous qui n'avons aucune connaissance en microphysique tout à fait obscur. Les interactions entre les acides aminés déjà intégrés dans la chaîne polypeptidique et l'aa-ARNt présenteraient ainsi une nature d'*interférences* ondulatoires plutôt que d'interactions corpusculaires et obéiraient dès lors à des contraintes « musicales » avec des fréquences de résonance disposées selon la gamme au tempérament égal. M. Sternheimer entend ainsi établir une règle de correspondance[1] entre les notes de la gamme tempérée et les des vingt-deux acides aminés du vivant, correspondance à caractère « stimulatrice » si la séquence sonore est « en phase avec l'élongation » (nous citons) et à caractère « inhibitrice » dans le cas inverse ; « protéodie » (terme dérivé de protéine et du grec *ôdê*, chant) étant la dénomination de cette séquence sonore qu'elle soit stimulatrice ou inhibitrice.

MATÉRIEL ET MÉTHODES

Influence des protéodies sur le développement foliaire d'*Arabidopsis thaliana* écotype Columbia

Enregistrement des protéodies

Les protéodies sont synthétisées sur logiciel *Bio2MIDI* (développé par la société *Algorithmic Arts*) selon les règles de transposition établies par M. Sternheimer et stipulées dans son brevet[1]. Les notes sont jouées en legato à un tempo de 110. Les séquences révisées d'acides aminés de six protéines d'*A. Thaliana* sont obtenues sous format FASTA[2] sur la banque de données Swiss-Prot, il est question :

(*a*) de la protéine d'actine qui joue un rôle fondamental dans la formation du cytosquelette, très abondante dans la cellule, elle est aussi parmi les protéines les plus conservées du vivant,

(*b*) d'une sous-unité du complexe cytochrome b6-f chloroplastique, complexe essentiel au photo-autotrophisme puisqu'il est intervient dans les deux voies du transfert électronique entre les deux photosystèmes PSII et PSI,

(*c*) de la sous-unité γ du complexe F0-F1 ATP-synthase mitochondrial, complexe essentiel au métabolisme aérobie,

(*d*) de l'enzyme nitrilase 1, impliquée dans l'une des voies de synthèse de l'auxine, hormone végétale qui stimule l'auxèse au niveau des parties aériennes,

(*e*) de la grosse sous-unité de l'enzyme RubisCo, protéine catalysant la fixation du CO2 sur les substrats initiateurs du cycle de Krebs

(*f*) de l'acide indole 3-acétique amido-synthétase chargé de l'inactivation de la molécule d'auxine (acide indole 3-acétique) par conjugaison avec un acide aminé (Böttcher, 2010).

Ces séquences protéiques sont ensuite transposées en séquences de fréquences sonores, avec répétition des séquences (a), (b), (c), (d) de courte durée, puis assemblées sur le logiciel *Audacity* pour aboutir à une séquence musicale globale d'une trentaine de minutes.

Séquence globale :	**inhibitrice**	**stimulatrice**
Séquences unitaires ***a***, ***b***, ***c***, ***d***, ***e*** :	inhibitrices	stimulatrices
Séquence unitaire ***f*** :	stimulatrice	inhibitrice

1 *cf*. tableau en annexe

2 Codes d'accession Uniprot : ***(a)*** P53492-1 (***b)*** Q9ZR03-1 (***c)*** Q96250-1 (***d)*** O03042-1 (***e)*** P32961-1 (***f)*** Q9SZT9-1

Conditions de culture
Les graines sont préalablement stratifiées 72h à 4°C (synchronisation) puis semées au pinceau en surface de trois planches de seize cubes (4x4) de laine de roche humectée à raison d'une graine par cube. Les graines sont mises à germer en terrarium maintenu à température de 22°C ± 0,7°C, à hygrométrie proche 75% et disposé sous une fenêtre à l'abri des rayons directs du soleil. En complément de l'éclairement naturel, un éclairage artificiel est placé 40 cm au-dessus du terrarium ; l'ampoule utilisée est une Envirolite type croissance allongée, branchée sur minuteur pour une durée de fonctionnement de 8h30. On veille à ce que le flux lumineux reçu entre chaque lot soit le plus constant possible, en outre, une permutation des planches est effectuée quotidiennement afin de s'assurer qu'aucun biais d'éclairement hétérogène soit en cause des variabilités éventuelles. Le taux de germination moyen est de 90%. Dès le stade deux cotylédons, le terrarium est tenu ouvert. Les planches sont bassinées tous les deux jours en eau de ville, hebdomadairement en solution fertilisée. Tous les soirs, les planches stimulées et inhibées sont successivement mises en caisson insonorisé où sont jouées les séquences musicales idoines à volume « ambiant pour l'oreille humaine ».

Quantification de la croissance
On estime la croissance en mesurant à l'aide d'un pied-à-coulisse les longueurs des feuilles avec leur pétiole pour les deux premiers verticilles, ainsi que la longueur de l'hypocotyle. Ces mesures sont consignées dans un fichier *csv.*

Influence des protéodies sur la synthèse d'α-amylase par la couche à aleurone de grains d'orge, *Hordeum vulgare* L.

Enregistrement des protéodies
Méthode similaire à la précédente. Les protéines considérées (séquences révisées) sont deux isozymes de l'α-amylase, type A et type B – d'accessions Uniprot respectives : P00693 et P04063.

Conditions de pré-germination
Trois lots de dix graines d'orge sont constitués et mis à pré-germer pendant 72 heures dans des pots de verre remplis de 2 mL d'eau de ville chacun. Les graines proviennent d'un magasin d'alimentation biologique et sont destinées à la consommation humaine de germes d'orge.

Les pots sont placés dans les cavités de moellons en béton (isolation thermique et phonique) chacun avec un haut-parleur de petite taille, puis couverts par des briques d'argile. Ces trois lots sont répartis dans trois pièces tenues obscures et à température voisine de 20°C ± 0,5°C. Les séquences protéodiques sont émises automatiquement toutes les cinq heures grâce au *crontab* de Linux, à un volume « ambiant pour l'oreille humaine ». L'insonorisation est sans reproche pour une oreille humaine, le seul biais que nous envisageons relevant des micro-variations thermiques, mais l'isolation *a priori* permise par les caissons en béton limite leur amplitude.

Gels de révélation
Six boîtes de Pétri sont utilisées pour contenir le gel révélateur de l'activité amylolytique des semences germées. Le milieu de révélation contient pour 100 mL d'eau du robinet : 5 mL d'agar-agar et 2,5 mL d'amidon de maïs. Après ébullition, il est versé 5 mL de milieu à l'aide d'une seringue dans chaque boîte de Pétri. La gélification se produit au cours du refroidissement. Il importe que la hauteur de gel soit identique entre chaque boîte (influence sur la superficie de l'auréole de décoloration).

Après 72 heures (correspondant physiologiquement à la percée de la radicule ou à la formation des premières racines selon les lots ; le choix de cette durée d'imbibition s'est fait en accord avec le moment d'apparition du pic de sécrétion d'α-amylase chez le riz, *cf.* Kaneko, 2002) les graines sont coupées en deux au niveau du sillon central, puis les moitiés mises côte à côte albumen retourné sur le milieu de révélation. On répartit cinq couples de demi-graines par boîte de Pétri. Les boîtes sont fermées ; l'incubation s'effectue dans les mêmes conditions que la pré-germination, toujours sous l'influence des protéodies pour les lots inhibés et stimulés.

Révélation *cf.* Fig. 1 p. 8
Au bout de 24 heures d'incubation, les graines sont retirées des milieux, on révèle la présence d'amidon à l'aide d'une teinture d'iode de pharmacie diluée à 4% et répartie superficiellement pendant 20 secondes. Les milieux sont rincés et immédiatement photographiés et scannés accompagnés d'un repère millimétré.

Quantification
La mesure de la superficie des auréoles de décoloration s'effectue à l'aide du logiciel gimp. Auparavant, les clichés subissent une correction du contraste avec réglages constants. La sélection des aires décolorées se fait au moyen de l'outil de sélection contiguë paramétré avec une tolérance de teinte (seuil) de 15 sur 255. On relève le nombre de pixels contenu dans cette sélection avec l'outil *Histogramme*, qui sera ensuite converti en aire mm² par une règle de trois, connaissant l'échelle grâce aux repères millimétrés (les pixels étant carrés, $A_{pix}=L_{pix}^2$). Les mesures de superficie sont consignées dans un fichier *csv.*

Analyse des observations sur *Hordeum vulgare* L.

TRAITEMENT STATISTIQUE

■ **Modèle statistique**

Le test ANOVA s'envisage sur la nature des variables en jeu : il s'agit d'expliquer une variable quantitative par une variable (facteur) qualitative, en supposant une distribution gaussienne des observations et sachant que le nombre de facteurs (3) ne justifie pas de recourir à un test classique de Fisher. On prendra donc un modèle de relation linéaire entre la surface de décoloration et le facteur *i*, avec pour paramétrisation $S_{ij}=\mu+\alpha_i+\varepsilon_{ij}$ où $S_{ij}\sim\mathcal{N}(m_i,\sigma)$ avec S_{ij} superficie de l'auréole de décoloration de la graine *j* soumise à la protéodie *i* où *i* prend les valeurs -1, 0 et 1 correspondant respectivement à la protéodie inhibitrice, au témoin sans protéodie et à la protéodie stimulatrice ; *j* un entier compris entre 1 et 10.

▫ μ : superficie moyenne des auréoles.

$$\mu = \frac{1}{30}\sum_{i=1}^{3}\sum_{j=1}^{10}(Y_{ij})$$

▫ α_i : terme factoriel qui rend compte de l'effet de la protéodie *i* relativement au groupe témoin, d'où $\alpha_1=0$.

▫ $\varepsilon_{ij}\sim^{iid}\mathcal{N}(0,\sigma)$ terme d'erreur qui rend compte de la variabilité résiduelle (variabilité intra-groupe non expliquée par le facteur sous l'hypothèse d'indépendance).

▫ m_i : moyenne des superficies d'auréoles pour le groupe soumis à la protéodie *i*.

$$m_i = \mu + \alpha_i = \sum_{j=1}^{10}\frac{Y_{ij}}{10}$$

■ **Hypothèses structurales du test**

Normalité de la distribution des termes d'erreur résiduelle ; indépendance des termes d'erreur résiduelle ; homoscédasticité (identicité des variances entre catégories : σ ne dépend pas de *i*).

■ **Hypothèses statistiques**

Hypothèse nulle $\mathcal{H}_0$: "$m_{-1}=m_0=m_1$"

Alternative $\mathcal{H}_1$: "$m_{-1}\neq m_0 \cup m_{-1}\neq m_1 \cup m_1\neq m_0$" (test bilatéral).

■ **Statistique**

$F=Var_F/Var_R\sim^{\mathcal{H}0}\mathcal{F}(2,27)$

■ **Région critique**

$W=[c;+\infty)$ où *c* est lu sur la table de Fisher de sorte que $Pr(F\in W)=\alpha$ sous $\mathcal{H}_0$.

	Graines « inhibées »	Graines contrôle	Graines « stimulées »
S_{ij} (en mm^2)	49.77 31.50 31.57 72.06 8.75 45.94 36.91 103.95 72.06 8.75	35.96 54.45 44.80 17.33 31.02 17.84 44.23 20.22 44.75 34.41	135.12 80.77 106.15 132.53 111.23 72.35 67.26 94.48 115.94 84.56
Moyenne $\bar{S}$ (mm^2)	46.12	35.4	100.04
Ecart-Type σ (mm^2)	29.83	12.57	24

RÉSULTATS ET INTERPRETATION BIOLOGIQUE

▪ **ANOVA globale**

	∂∂	*SC*	*Variance*	*F*	*Pr(>F)*
Facteur	2	24457	12228	22.482	**$1.788\cdot10^{-6}$**
Résidus	27	14686	543.9		

Signification des abréviations :

▫ *∂∂* : degrés de liberté

▫ *SC* : somme des carrés (variabilité factorielle) où :

$$\underbrace{\sum_{ij}(S_{ij}-\bar{S})^2}_{SC\ Totaux}=\underbrace{\sum_{ij}(S_{ij}-\bar{S}_i)^2}_{SC\ Résiduels}+\underbrace{\sum_{ij}(\bar{S}_i-\bar{S})^2}_{SC\ Factoriels}$$

▫ *Variance* = *SC*/*∂∂*

▫ *F* : la statistique $F=Var_F/Var_R\sim^{\mathcal{H}0}\mathcal{F}(2,27)$

▫ *Pr(>F)* : probabilité critique (*p-value*) de trouver une statistique de valeur supérieure à la statistique des observations, $Pr(F>F_{obs})$, *ie*. borne inférieure des risques α pour lesquels $\mathcal{H}_0$ est rejetée.

Influence hautement significative de la séquence stimulatrice

La valeur de la probabilité critique très nettement inférieure à l'erreur de première espèce α communément fixée à 0.05 nous permet de rejeter avec confiance l'hypothèse nulle, et de valider l'hypothèse alternative. Il existe au moins un i_1 tel que $m_{i1}\neq m_{i2}$: les séquences protéodiques ont une influence hautement significative sur la synthèse d' d'α-amylase dans la graine d'orge. Sous nos conditions expérimentales, la synthèse d'α-amylase apparaît ainsi *doublée* par la seule émission de protéodies stimulatrices.

▪ ANOVA couple par couple

L'analyse de la variance couple par couple (par commodité, nous utilisons un test d'ANOVA quoiqu'un test de Fisher puisse également être envisagé) nous permet de dresser le tableau des probabilités critiques suivant :

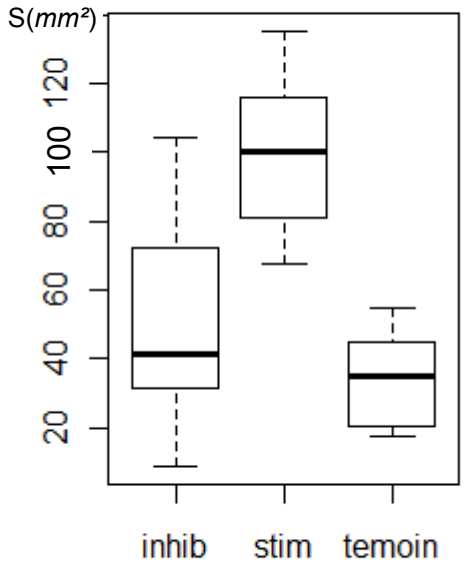
	Stimulation	Inhibition
Témoin	$4.977 \cdot 10^{-7}$	0.2727
Inhibition	$3.072 \cdot 10^{-4}$	

Sous les hypothèses de travail, nos expériences nous permettent de rejeter l'hypothèse $\mathcal{H}_0$. L'ordre de grandeur de la probabilité critique pour le couple stimulation-témoin assure de l'influence de la séquence sonore basée sur la règle de transposition stimulatrice, la différence de moyenne entre m_0 et m_1 étant jugée hautement significative. La qualité de l'ajustement R^2=76% élevée indique que les protéodies expliquent plus des trois quart de la variabilité constatée, indication qui légitime en outre le modèle d'ajustement ici utilisé.

La différence de moyenne entre m_{-1} et m_1 est également significative, quoique les protéodies n'expliquent ici que 52% de la variabilité observée. Cette influence sélective des protéodies suggère que le bruit seul, considéré comme perturbation mécanique, ne peut être à l'origine des écarts aux moyennes. Il existerait donc une information musicale, dont le support est la succession de perturbations mécaniques de fréquences spécifiques, qui influencerait plus ou moins significativement des processus biologiques.

Absence d'influence de la séquence inhibitrice

Il faut noter cependant que la séquence inhibitrice ne joue pas comme symétrique dans ses effets de la séquence stimulatrice, symétrie qu'on pouvait *a priori* envisager.



La probabilité critique pour le couple inhibition-témoin est de 0.27, valeur très largement supérieure au seuil conventionnel de décision de 5% et nous conduit donc à accepter l'hypothèse $\mathcal{H}_0$ c'est-à-dire que l'écart entre les moyennes m_0 et m_{-1} n'est pas déterminé par le facteur : il est purement aléatoire.

La séquence inhibitrice n'a donc pas eu d'influence sur la synthèse d'α-amylase chez *Hordeum vulgare* dans nos conditions expérimentales.

VALIDATION DES HYPOTHESES

cf. p. 5

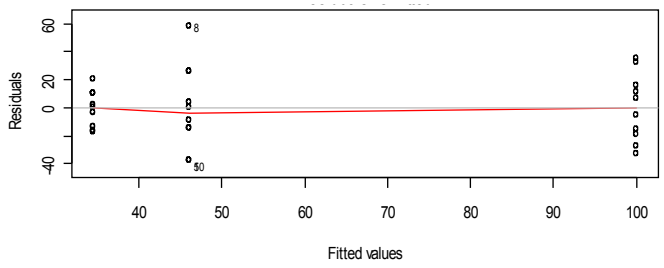
CONCLUSION & CRITIQUE

En l'état de nos connaissances, il n'est pas possible d'expliquer la spécificité de l'effet protéodique, autrement dit le fait qu'une séquence ait un effet significatif sur un paramètre biologique et l'autre non. Nous pourrions suggérer l'existence d'un réseau d'intéractions protéine-protéine de type aléatoire (non "*scale-free*"), où les variations de concentration d'une protéine peuvent générer des effets imprévisibles et non linéaires. Ainsi, la synthèse d'α-amylase pourrait s'intégrer dans un réseau au sein duquel deux actions agonistes n'entraînent pas nécessairement des effets symétriques.

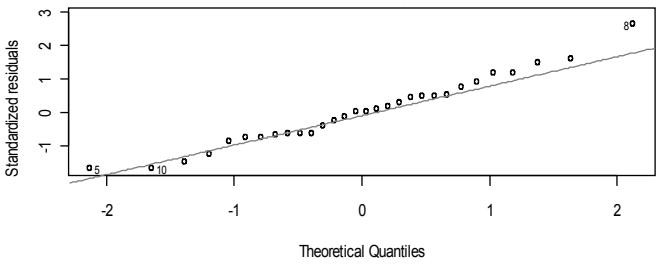
Cependant mentionnons comme biais expérimental pouvant fragiliser nos conclusions et rendre absurde le problème soulevé précédemment : *(1)* le faible effectif de graines soumises à l'expérience et *(2)* le biais de brefs écarts thermiques entre lots de graines, puisqu'il n'a pas été possible avec nos moyens d'associer les lots parfaitement insonorisés dans une même pièce. Il s'agissait de faire un choix entre une excellente insonorisation contre un régime de températures variable, ou une insonorisation médiocre contre un régime de températures identique entre lots. Précisons que le lot *a priori* soumis aux températures en moyenne les plus basses était le lot stimulé. Quoique les graines d'orge puissent présenter une dormance (Donwell, 1980), nous raisonnons ici sur des écarts thermiques en réalité faibles, il ne nous semble pas qu'il faille alors suggérer un débourrement plus précoce du lot stimulé. Par contre, au stade pré-germinatif, et *a fortiori* pour des graines destinées à la consommation de germes d'orge toute l'année, il serait plus vraisemblable qu'une température légèrement plus élevée accélère très légèrement la germination plus qu'elle ne l'inhibe.

(à gauche) *Diagramme en "boîte à moustaches" représentant les grandeurs statistiques des observations pour les trois facteurs : protéodie stimulatrice (stim), protéodie inhibitrice (inhib) et contrôle (témoin).*

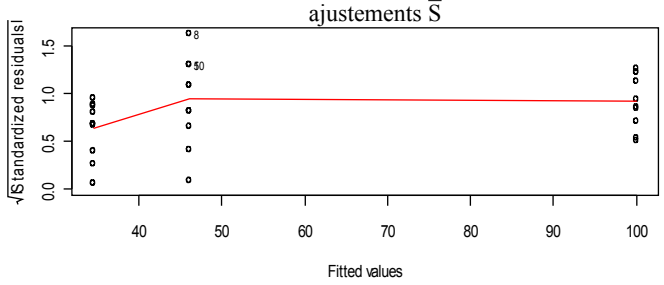
(a) Résidus en fonction des ajustements (=*fitted-values*) $\overline{S}$



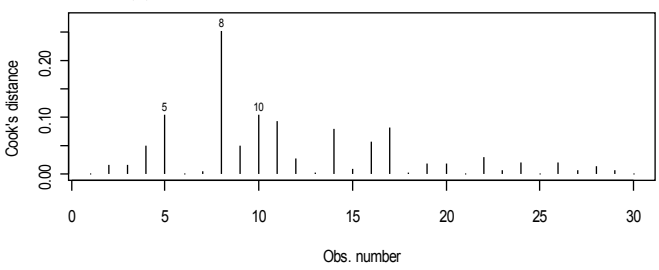
(b) Equivalent droite de Henry (Q-Q plot)



(c) Valeur absolue des résidus standardisés en fonction des ajustements $\overline{S}$



(d) Distance de Cook des différentes observations



NB. Les trois alignements verticaux de points aux graphes *(a)* et *(c)* correspondent respectivement, de gauche à droite, aux résidus des catégories inhibées, témoin et stimulées.

(suite dela page 4) VALIDATION DES HYPOTHESES ET DU MODELE

▪ ANOVA globale

Le diagramme de Cook met en évidence au moins trois mesures marginales qui ont une influence significative sur la statistique, ces trois mesures correspondant à des superficies de décoloration pour le lot inhibé exclusivement. L'alignement des points sur la droite de Henry est correct, hors extrêmes où quelques points accusent un écart sensible par rapport à la loi normale, notamment la mesure n°8 (lot inhibé) particulièrement marginale – le test de Shapiro-Wilk ne permet cependant pas de rejeter l'hypothèse nulle supposant la normalité des observations (*p-value*=8%>α). On vérifie par la commande R `summary(residuals(lm(Yij~i)))` que les écarts résiduels suivent une gaussienne effectivement centrée en 0. L'homoscédasticité sera supposée validée vu la symétrie des points par rapport à l'horizontale en *y=0* du diagramma *(a)* et malgré le fait que la courbe *(c)* présente un étalement des résidus standardisés plus ou moins variables selon les catégories *i* avec une tangente non nulle de la courbe du *scale-location* en $i_{inhibé}$ (l'écart est notamment plus fort entre résidus pour le lot témoin que pour les lots inhibés et stimulés : il y aurait une très légère corrélation avec *i*). Il est intéressant de noter qu'un test de Fisher indique une parfaite identicité entre les écarts-types des lots inhibés et stimulés, ces derniers étant statistiquement légèrement supérieurs à l'écart-type du lot témoin. Pratiquement, il apparaîtrait ainsi des sous-populations réagissant très diversement à l'émission de protéodies : cette variabilité intra-groupe pourrait s'expliquer, pour les lots traités, par ce qu'on transcrit une séquence protéique *spécifique* de l'espèce, quoique cette dernière puisse néanmoins varier légèrement entre individus pour des raisons génétiques.

La validation de l'hypothèse d'indépendance ne pose *a priori* pas de problème, du fait des précautions prises à l'isolation phonique. Graphiquement, la courbe *(a)* ne montre aucune structure particulière des résidus, leurs espérances sont situées sur la droite d'équation *y=0* – l'indépendance des écarts résiduels vis à vis des facteurs est donc bel et bien vérifiée.

▪ ANOVA couple par couple

L'ensemble des hypothèses est validée pour les analyses en couple par couple. Le test Shapiro renvoie systématiquement des probabilités critiques élevées (entre 50% et 75%) donnant pleine confiance en la normalité des distributions des populations particulières.

Analyse des observations sur *Arabidopsis thaliana* (L.) Heynh.

MODELE STATISTIQUE

■ **Modèle statistique**

On considère un estimateur du développement végétal pour tout plant i soumis à la protéodie j : $D_{ij} = M_{ij}^1 + M_{ij}^2 + H_{ij}$ avec M_{ij}^1 la moyenne des longueurs des feuilles du premier verticille, pétiole inclus ; M_{ij}^2 la moyenne des longueurs des feuilles du deuxième verticille, pétiole inclus ; H_{ij} la longueur de l'hypocotyle.

Décomposition ANOVA en modèle linéaire avec une paramétrisation $D_{ij}=\mu+\alpha_i+\varepsilon_{ij}$ où i prend les valeurs -1, 0 et 1 correspondant respectivement à la protéodie inhibitrice, au contrôle sans protéodie, et à la protéodie stimulatrice ; j un entier compris entre 1 et 15.

▫ $\boldsymbol{n_i}$: nombre de plantes du lot i.

▫ $\boldsymbol{\mu}$: estimateur moyen du développement.

$$\mu = \frac{1}{3(n_1+n_0+n_{-1})}\sum_{i=1}^{3}\sum_{j=1}^{n_i}(D_{ij})$$

▫ $\boldsymbol{\alpha_i}$ et $\boldsymbol{\varepsilon_{ij}} \sim^{iid} \mathcal{N}(0,\sigma)$ cf. modèle *Hordeum vulgare*.

▫ $\boldsymbol{m_i}$: moyenne des estimateurs pour le groupe soumis à la protéodie i.

$$m_i = \mu + \alpha_i = \sum_{j=1}^{n_i}\frac{D_{ij}}{n_i}$$

■ **Hypothèses de travail**

Normalité de la distribution des termes d'erreur résiduelle ; indépendance des termes d'erreur résiduelle ; homoscédasticité (identicité des variances entre groupes factoriels).

■ **Hypothèses statistiques**

Hypothèse nulle $\mathcal{H}_0$: "$m_{-1}=m_0=m_1$"

Alternative $\mathcal{H}_1$: "$m_{-1}\neq m_0 \cup m_{-1}\neq m_1 \cup m_1\neq m_0$"

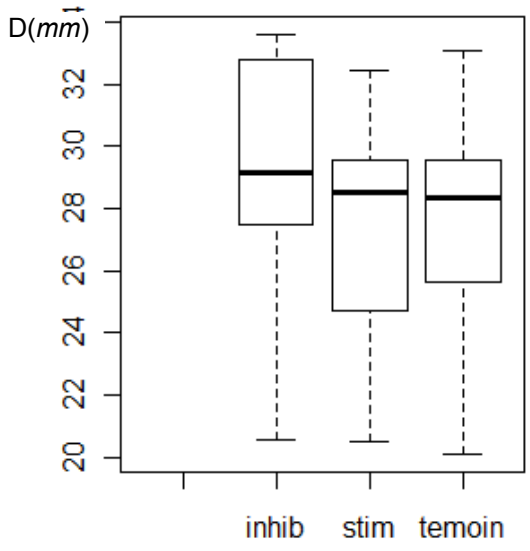
■ **Statistique et région critique** *cf.* p. 3

RESULTATS

	$\partial\partial$	*SC*	*Variance*	*F*	*Pr(>F)*
Facteur	2	8.75	4.37	0.195	**0.8236**
Résidus	41	919.5	22.428		

Signification des abréviations : *cf.* p. 3

L'analyse ANOVA effectuée sous R sur l'ensemble des données des trois lots retourne une probabilité critique de 0.8236 >> α =5%. Nos expériences n'ont donc pas permis de rejeter l'hypothèse nulle : l'écart entre les moyennes des estimateurs du développement végétal ne sont pas statistiquement significatives. Les protéodies n'ont donc pas eu d'influence sur le développement foliaire chez *Arabidopsis thaliana* dans nos conditions de culture. Les tests ANOVA en couple par couple retournent pour le couple stimulé-témoin une *p-value* de 0.4677, pour le couple inhibé-témoin une *p-value* de 0.8463, pour le couple stimulé-inhibé une *p-value* de 0.7074. Chacune de ces probabilités critiques est très largement supérieure au risque de première espèce conventionnellement fixé à 5%.



VALIDATION D'HYPOTHESES

cf. p. 7

INTERPRETATION BIOLOGIQUE ET CRITIQUES

Cette acceptation incontestable de $\mathcal{H}_0$ dans l'expérience effectuée sur *Arabidopsis thaliana* vient en butte avec son net rejet dans l'expérience effectuée sur *Hordeum vulgare*.

On fera ici encore une critique d'ordre statistique sur la faiblesse des effectifs d'observation, ainsi que sur l'estimateur que nous avons décidé d'employer. En effet, il est difficile d'estimer rigoureusement le développement végétal dans la mesure où tous les plants ne développent pas le même verticille foliaire en même temps, certains amorçant le stade troisième verticille quand d'autres poursuivent encore l'expansion foliaire du deuxième verticille. Cette variabilité a justifié pour nous de recourir à un estimateur qui permette de comparer des pieds à des stades physiologiques légèrement distincts en prenant en compte plusieurs mesures caractéristiques : les deux verticilles de la base de la rosette, les moins affectés par le stade de croissance de la plante au terme de la période d'expérimentation, ainsi que l'hypocotyle dont on

croit qu'il est le plus sensible aux éventuelles variations de concentration en auxine (hormone jouant sur l'auxèse caulinaire). Il est possible au demeurant que cette estimateur soit biaisé, pour des raisons que nous ne soupçonnons pas ici.

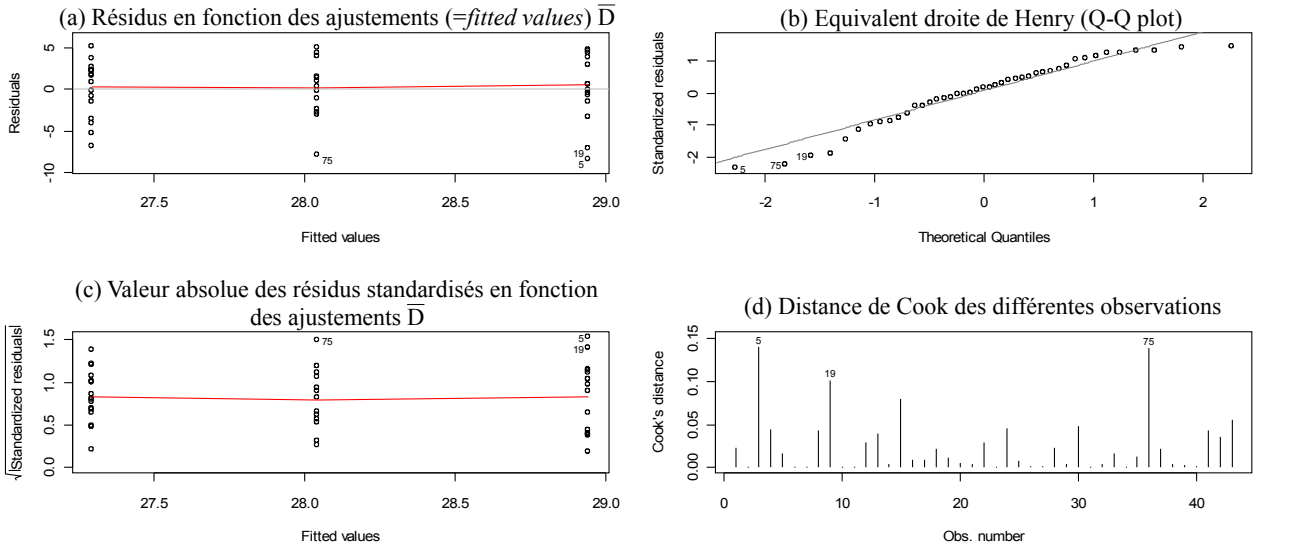
Une remarque d'ordre biologique sur les protéines choisies : certes, dans la cellule, ces protéines peuvent intervenir au sein d'un réseau de réactions de type "*scale-free*" qui, du fait de la plasticité et de la redondance du vivant, rend imperceptible au niveau macro-phénotypique leur variations de concentration. Nous avons néanmoins, pour éviter ce risque, fait le choix de cibler des protéines-clefs ("*hubs*") du métabolisme, de sorte que les manipulations soit clairement décelables.
Toujours au niveau biologique : une stimulation ou une inhibition de trente minutes peut s'avérer trop longue vis à vis d'un mécanisme (lequel ?) éventuellement saturable, ou inversement : des stimulations/inhibitions à raison d'une émission toutes les vingt quatre heures peuvent s'avérer insuffisantes ; enfin, l'action sur un ensemble de protéines peut conduire à des interactions imprévisibles entre ces mêmes protéines stimulés/inhibées qui tempèrent l'effet macro-phénotypique (nous retrouvons l'hypothèse « réseau de réactions » sus-évoquée, mais en considérant l'interaction entre *les* protéines choisies).

Enfin, une considération d'ordre expérimental : l'expérience a été menée sur une durée d'un mois dans un studio d'étudiant de 19 m². Ce paramètre implique deux biais loin d'être négligeables : le fait que le caisson « insonorisé » l'était imparfaitement vu les moyens du bord, et qu'il était impossible de l'isoler davantage en le plaçant dans une pièce séparée, d'où une éventuelle « contamination » des deux planches supposées être isolées de l'influence de la musique en cours d'émission ; le fait par ailleurs qu'il a été joué pendant la durée de l'expérimentation des musiques pour notre loisirs, indépendamment donc de l'expérience. Ces musiques n'étant pas, à proprement parler, des protéodies, nous avons émis l'hypothèse initiale qu'elles ne risquaient pas de perturber le déroulement normal de l'expérience – cette hypothèse ne peut être raisonnablement validée dans la mesure où une « contamination » musicale a bel et bien pu intervenir.

CONCLUSION GENERALE

Le succès des expériences menées sur l'orge ne nous permettent pas, en l'état, de conclure en l'inefficacité des protéodies. L'ensemble de nos expériences en revanche soulève des questions et des incertitudes sur le mécanisme bio-physique à l'oeuvre dans cette « régulation épigénétique », et sur sa reproductibilité dans d'autres conditions et pour d'autres espèces végétales.

(suite de la page 6) VALIDATION DES HYPOTHESES



NB. Pour les remarques annexes, *cf.* p. 5.

Graphiquement, l'hypothèse de normalité semble validée vu l'alignement des points sur la droite de Henry, ce malgré la présence marginale de quelques points aux extrêmes, signalés par le diagramme des distances de Cook : ce sont les mesures 5, 19 et 75 appartenant respectivement aux lots inhibé, inhibé

et témoin. Un test de Shapiro-Wilk donne une validation statistique de cette hypothèse : la distribution est effectivement gaussienne (*p-value*= 5%), toutefois le test n'est pas d'une haute significativité. Le diagramme en boîte-à-moustaches illustre notamment une dissymétrie de la boîte "inhibée" signalant un léger écart à la normalité. L'alignement de l'espérance des résidus sur la droite *y=0* assure du centrage de la gaussienne et l'écartement entre chaque résidu apparaît visuellement comme constant et indépendant des catégories *i* : l'hypothèse d'homoscédasticité est confirmée. L'indépendance est validée vu la distribution symétrique, sans structure particulière, des points autour de la droite *y=0* au niveau du diagramme (*a).*

REMERCIEMENTS

Nous tenons à remercier M. Bertholomieu pour ses conseils et les graines d'*Arabidopsis* qu'il nous gracieusement offertes.

BIBLIOGRAPHIE

Böttcher Ch. et al. 2010. Sequestration of auxin by the indole-3-acetic acid-amido synthetase GH3-1 in grape berry (*Vitis vinifera* L.) and the proposed role of auxin conjugation during ripening. *Journal of Experimental Botany*. **61** (13): 3615-3625

Dunwell J.M. 1980. Dormancy and germination in Embryos of *Hordeum vulgare* L. *Annals of Botany*, vol. **48**, 2, 203-214

Kaneko M. **et al.** 2002. The α-amylase induction in endosperm during rice seed germination is caused by gibberellin synthesized in epithelium. *Plant Physiol.* **128** (4):1264

Sternheimer J. 29/08/2007. Procédé de régulation épigénétique de la biosynthèse des protéines par résonance d'échelle. Fascicule de brevet européen EP 0 648 275 B1. WO 1993/024645.

Uniprot. http://www.uniprot.org/ Consulté en novembre et décembre 2010

ANNEXE 1

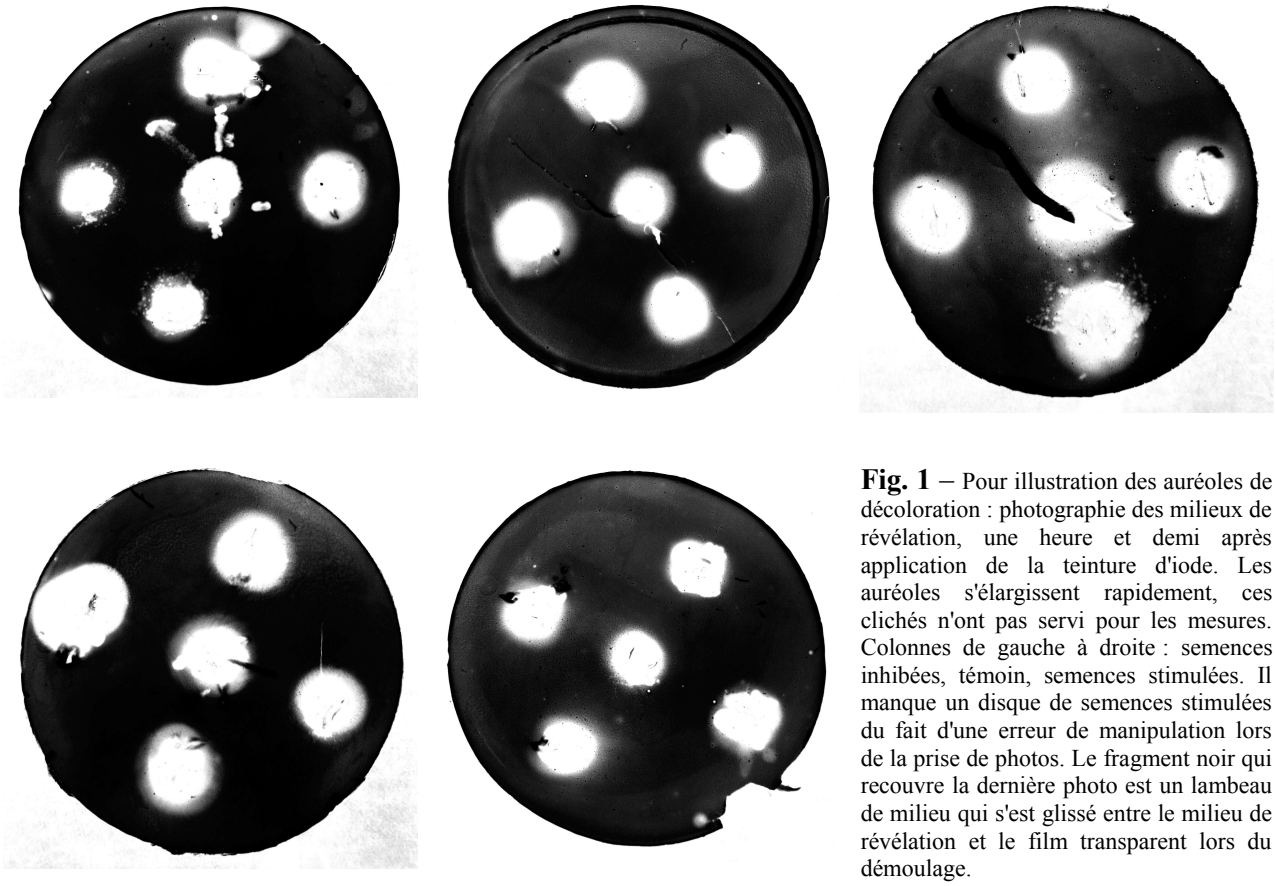


Fig. 1 – Pour illustration des auréoles de décoloration : photographie des milieux de révélation, une heure et demi après application de la teinture d'iode. Les auréoles s'élargissent rapidement, ces clichés n'ont pas servi pour les mesures. Colonnes de gauche à droite : semences inhibées, témoin, semences stimulées. Il manque un disque de semences stimulées du fait d'une erreur de manipulation lors de la prise de photos. Le fragment noir qui recouvre la dernière photo est un lambeau de milieu qui s'est glissé entre le milieu de révélation et le film transparent lors du démoulage.

ANNEXE 2

Tableau des correspondances.

Code FASTA	Acide aminé	Note stimulatrice Ecriture anglaise	Fréquence	Note stimulatrice Ecriture française	Note inhibitrice	Fréquence
A	Ala	$_3$C	261,63 Hz	Do, 3° octave	$_4$D	523,25 Hz
B	Asn	$_3$G	392,00 Hz	Sol, 3° octave	$_3$G	392,00 Hz
C	Cys	$_3$F	349,23 Hz	Fa, 3° octave	$_3$A	440,00 Hz
D	Asp	$_3$G	392,00 Hz	Sol, 3° octave	$_3$G	392,00 Hz
E	Glu	$_3$A	440,00 Hz	La, 3° octave	$_3$F	349,23 Hz
F	Phe	$_3$B	493,88 Hz	Si, 3° octave	$_3$D#	311,13 Hz
G	Gly	$_2$A	220,00 Hz	La, 2° octave	$_4$F	698,6 Hz
H	His	$_3$A#	466,16 Hz	La#, 3° octave	$_3$E	329,63 Hz
I	Iso	$_3$G	392,00 Hz	Sol, 3° octave	$_3$G	392,00 Hz
K	Lys	$_3$A	440,00 Hz	La, 3° octave	$_3$F	349,23 Hz
L	Leu	$_3$G	392,00 Hz	Sol, 3° octave	$_3$G	392,00 Hz
M	Met	$_3$A	440,00 Hz	La, 3° octave	$_3$F	349,23 Hz
N	Asn	$_3$G	392,00 Hz	Sol, 3° octave	$_3$G	392,00 Hz
P	Pro	$_3$F	349,23 Hz	Fa, 3° octave	$_3$A	440,00 Hz
Q	Gln	$_3$A	440,00 Hz	La, 3° octave	$_3$F	349,23 Hz
R	Arg	$_3$C	261,63 Hz	Do, 3° octave	$_3$D	293,66 Hz
S	Ser	$_3$E	329,63 Hz	Mi, 3° octave	$_3$A#	466,16 Hz
T	Thr	$_3$F	349,23 Hz	Fa, 3° octave	$_3$A	440,00 Hz
V	Val	$_3$F	349,23 Hz	Fa, 3° octave	$_3$A	440,00 Hz
W	Trp	$_4$D	523,25 Hz	Ré, 4° octave	$_3$C	261,63 Hz
Y	Tyr	$_3$C	261,63 Hz	Do, 3° octave	$_3$D	293,66 Hz

NB. Sous *Bio2Midi*, $_3$X est noté 1X, $_2$X noté 0X, $_4$X noté 2X (où X est une note quelconque).

ANNEXE 3
Copies d'écran R.

```
donnees.Hordeum <-
read.csv("~/Bureau/exp.proteodique/milieux_numerises.csv",sep=";",dec=",",header
=TRUE)

> donnees.Hordeum
    code mesures.px mesures.mm lot
1   inh1       6871       0.00  inhib
2   inh1       4349      31.50  inhib
3   inh1       4358      31.57  inhib
4   inh1       9948      72.06  inhib
5   inh1       1208       8.75  inhib
6   inh2       6343      45.94  inhib
7   inh2       5096      36.91  inhib
8   inh2      14351     103.95  inhib
9   inh2       9948      72.06  inhib
10  inh2       1208       8.75  inhib
```

```
11 stim1      18654     135.12  stim
12 stim1      11151      80.77  stim
13 stim1      14655     106.15  stim
14 stim1      18297     132.53  stim
15 stim1      15357     111.23  stim
16 stim2       9989      72.35  stim
17 stim2       9286      67.26  stim
18 stim2      13044      94.48  stim
19 stim2      16007     115.94  stim
20 stim2      11674      84.56  stim
21  tém1       4964      35.96  temoin
22  tém1       7517      54.45  temoin
23  tém1       6185      44.80  temoin
24  tém1       2392      17.33  temoin
25  tém1       4283      31.02  temoin
26  tém2       2463      17.84  temoin
27  tém2       2792      20.22  temoin
28  tém2       6178      44.75  temoin
29  tém2       6107      44.23  temoin
30  tém2       4750      34.41  temoin
```

▫ **Ajustement au modèle linéaire**

```
> ano <- lm(mesures.mm~lot,data=donnees.Hordeum)
> summary(ano)

Call:
lm(formula = mesures.mm ~ lot, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-37.376 -15.266  -0.139  10.968  57.824
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   46.126      7.375   6.254 1.09e-06 ***
lot.stim       53.913     10.430   5.169 1.93e-05 ***
lot.temoin    -11.625     10.430  -1.115    0.275
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 23.32 on 27 degrees of freedom
Multiple R-squared: 0.6248,     Adjusted R-squared: 0.597
F-statistic: 22.48 on 2 and 27 DF,  p-value: 1.788e-06
```

▫ **Tableau d'analyse de variance globale**

```
> anova(ano)
```

▫ **Graphes : visualisation des résultats et validation des modèles**

```
> boxplot(mesures.mm~lot,data=donnees.Hordeum)
> par(mfrow=c(2,2))
> plot(residuals(ano,which=1)
> plot(residuals(ano,which=2)
> plot(residuals(ano,which=3)
> plot(residuals(ano,which=4)
```

▫ **Tableau d'analyse de variance couple par couple**

```
> anova(lm(mesures.mm~lot,data=data[11:30,]))
> anova(lm(mesures.mm~lot,data=data[1:20,]))
> anova(lm(mesures.mm~lot,data=data[c(1:10,21:30),]))
```