In Vitro Cell.Dev.Biol.-Plant (2008) 44:229–232
DOI 10.1007/s11627-007-9100-3

MICROPROPAGATION

Use of regression analysis in plant cell, tissue, and organ culture experiments

J. C. Lorenzo · M. Garcia-Borroto

Received: 21 March 2007 / Accepted: 6 November 2007 / Published online: 24 January 2008 / Editor: M. Kane

Abstract Several authors have suggested that plant biotechnologists perform regression or trend analysis to compare means of related quantitative treatments (e.g., doses of inositol). The present paper compares two statistical strategies to determine the effect of inositol (0–400 mg l^{-1}) on proteolytic activity in the culture medium during pineapple growth in temporary immersion bioreactors. Strategy 1 involved one-way analysis of variance (ANOVA) followed by Tukey's Honestly Significant Difference (HSD). Strategy 2 consisted in the development of different regression analyses to determine the best fit equations to describe the experimental results. Curvefit software (version 2.10-0, May 15, 1987, Thomas S. Cox) was used. Cauchy, Normal, Parabola, and Hoerl equations were the best fitted according to their determination coefficients (R^2). The optimal inositol concentrations to increase proteolytic activity were determined from the equations. Quite different results were obtained following strategy 2: 126.76 mg l^{-1} inositol from Cauchy, 131.29 mg l^{-1} from Normal, 145.06 mg l^{-1} from Parabola, and 14.05 mg l^{-1} from Hoerl equations. In contrast, experimental data identified 200 mg l^{-1} inositol as the most adequate concentration to increase proteolytic activity in the culture medium. The statistical strategy 1, one-way ANOVA followed by Tukey HSD clearly supported this biological observation. In this paper, regression analysis was not useful to describe our experimental results.

J. C. Lorenzo (✉) · M. Garcia-Borroto
Bioplant Center, University of Ciego de Avila,
Ciego de Avila 69450, Cuba
e-mail: jclorenzo@bioplantas.cu
URL: www.bioplantas.cu

Keywords Statistical analysis · Biotechnological experiments · One-way ANOVA · Tukey HSD · Curvefit · Cauchy · Normal · Parabola · Hoerl

Introduction

Mize and Chun (1988), Compton (1994), and Ibanez et al. (2003) have suggested that plant biotechnologists perform regression analysis to compare means of related quantitative treatments such as explant age, hormone concentration, and time course. A mathematical model is constructed to describe the relationship between the treatment levels and the response variable. To identify the most appropriate model, a forward step procedure is commonly used. The simplest model (linear) is tested first and more complicated models (quadratic, cubic, etc.) are tested after the rejection of lower order models. Determination coefficients (R^2) are valuable indicators of the appropriateness of a model to describe the relation between variables. A high R^2 coefficient suggests that much of the variability is described by the model (Petersen 1977; Compton 1994). After the equation is obtained, a numerical procedure is applied to find the curve maximum value. This is regarded as the optimal treatment level.

To this point, regression analysis seems to be an excellent way to process monofactorial experiments where several quantitative treatment levels are compared. However, the most important limiting factor for this sort of statistical management is that mathematical functions rarely describe biological measurements properly. The present paper compares two statistical strategies, one

based on analysis of variance (ANOVA) and one using regression analyses, to determine the effect of inositol on proteolytic activity in the culture medium during pineapple growth in temporary immersion bioreactors. To our knowledge, this kind of statistical comparison of a plant biotechnology experiment has not been published to date.

Materials and Methods

Experimental data used in this paper are from Perez et al. (2004, Fig. 1*a*). The effect of inositol (0, 100, 200, 300, 400 mg l^{-1}) on proteolytic activity in the culture medium during pineapple growth in temporary immersion bioreactors was evaluated. Each treatment was replicated with

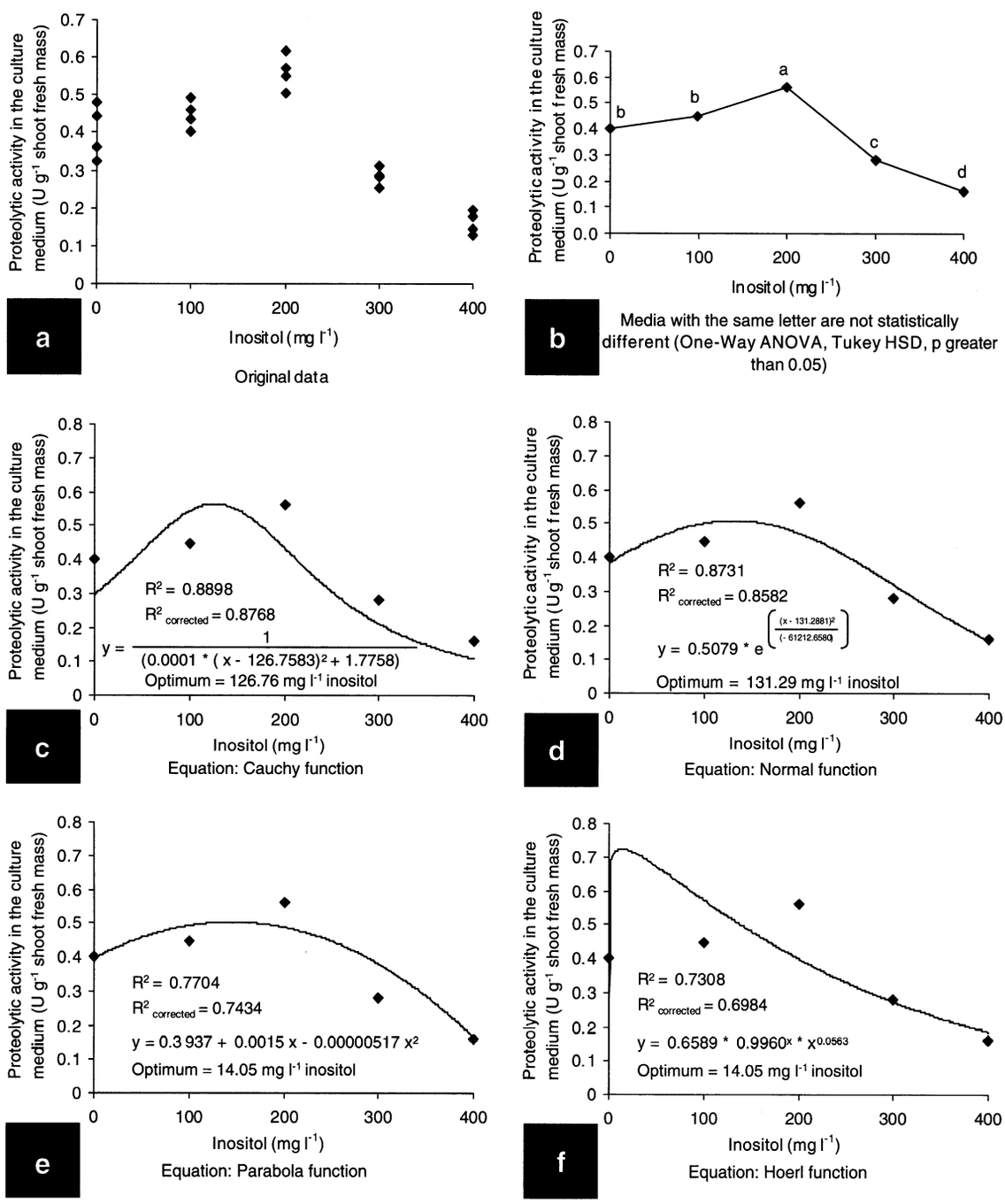

Figure 1. Statistical analyses of the effect of inositol on proteolytic activity in the culture medium during pineapple growth in temporary immersion bioreactors (Data from Perez et al. 2004). (*a*) Original data used for statistical analyses (four measurements per treatment). (*b*) Strategy 1, one-way ANOVA followed by Tukey HSD. (*c-f*) Strategy 2, best fits of regression analyses obtained by curvefit (version 2.10-0, May 15, 1987, Thomas S. Cox; only treatment averages are shown).

four bioreactors. The experiment was completely randomized. Two statistical strategies were compared. Strategy 1 involved one-way ANOVA followed by Tukey Honestly Significant Difference (HSD) tests. Strategy 2 consisted in the development of different regression analyses to determine the best fit equations to describe the experimental results. The Statistical Package for Social Sciences (Version 8.0 for Windows, SPSS Inc.) was used to process the experiment according to Strategy 1. Curvefit software (version 2.10-0, May 15, 1987, Thomas S. Cox) was used to develop Strategy 2. This software was used to perform the following regression analyses: Straight line, Line through origin, Reciprocal straight line, Line and reciprocal, Hyperbola, Reciprocal hyperbola, Second-order hyperbola, Parabola, Parabola at origin, Power, Mod power, Root, Supergeometric, Mod geometric, Exponential, Mod exponential, Logarithmic, Reciprocal logarithmic, Hoerl, Mod Hoerl, Normal, Logarithmic normal, Beta, Gamma, and Cauchy functions. MatLab 7.0 was used to determine optimal inositol concentrations.

Results and Discussion

Statistical results obtained following Strategy 1 are shown in Fig. 1*b*. Two hundred mg l^{-1} inositol significantly increased the proteolytic activity in the culture medium. Fig. 1*c–f* shows the results obtained with strategy 2. Cauchy (Fig. 1*c*), Normal (Fig. 1*d*), Parabola (Fig. 1*e*), and Hoerl (Fig. 1*f*) equations were the best fitted according to their determination coefficients (R^2).

Optimal inositol concentrations that significantly increased proteolytic activity were determined from the equations. Quite different results were obtained: 126.76 mg l^{-1} inositol from Cauchy, 131.29 mg l^{-1} from Normal, 145.06 mg l^{-1} from Parabola, and 14.05 mg l^{-1} from Hoerl equations.

In contrast, the biological response of replicates indicated 200 mg l^{-1} inositol to be the most effective concentration to increase proteolytic activity in the culture medium (Fig. 1*a*). Strategy 1 clearly supported this biological observation (Fig. 1*b*). Figure 1*b* shows homogeneous subsets described by Tukey HSD test, but similar results were obtained with the following tests: Student–Newman–Keuls, Tukey B, Duncan, Scheffe, Gabriel, Ryan–Einot–Gabriel–Welsch F, Ryan–Einot–Gabriel–Welsch Range, Hochberg, and Waller–Duncan. Figure 1 shows that the regression methods left biological meaning unclear.

Some recent references have applied regression analyses to explain biological responses (Aynalem et al. 2006; Mokhtassi et al. 2006; Onyeka et al. 2006; Sharma et al. 2006; Zhu et al. 2006; Browne 2007). In our opinion, only Onyeka et al. (2006) and Zhu et al. (2006) demonstrated that regression contributed to clarity of their biological results. Onyeka et al. (2006) established the relationships between anthracnose disease parameters on inoculated tissue culture-derived whole plants of *Dioscorea alata*. Figures with straight line functions and high R^2 (0.648–0.869) were provided. Zhu et al. (2006) studied phosphate uptake by two wheat cultivars and their doubled haploid lines. They found that an exponential regression plot fitted the data closely (R^2 = 0.921). On the contrary, Mokhtassi et al. (2006) and Sharma et al. (2006) published values of simple correlations coefficients, but plots were not provided. Therefore, data dispersion could not be observed. Aynalem et al. (2006) estimated straight line functions, but some R^2 were as low as 0.40 and consequently the regression functions did not fit the data properly. Browne (2007) also estimated straight line functions but determination coefficients were extremely low (0.12 and 0.27).

Previously, we unsuccessfully tried to use regression analyses many times. In some cases, where no biological meanings were obtained, we decided to perform other statistical strategies for publication. For example, we tried to develop mathematical models to describe the effect of the time course on electrolyte leakage and aldehyde content in sugarcane cryopreserved calluses (data in Martinez et al. 2002). Regression analyses of the effects of sodium nitrate, potassium phosphate, magnesium sulfate, potassium chloride, and sucrose on banana leaf lesion area caused by *Fusarium* culture filtrates did not generate reasonable models (data in Companioni et al. 2004). We also failed to find mathematical equations to establish the relationship between the time course and the levels of malondialdehyde, other aldehydes, chlorophyll pigments (a, b, total), phenolics (free and cell-wall linked) and protein in transgenic pineapple plantlets (data in Yabor et al. 2006). Consequently, based on our experience, adequate regression analyses are not always obtained. The comparison shown in Fig. 1 supports that instead of regression analyses, it is better to perform one-way ANOVA followed by Tukey HSD, which will provide a clear biological meaning.

References

Aynalem H, Righetti TL, Reed B. Iron formulation affects *in vitro* storage of hops: an image analysis. In Vitro Cell Dev Biol Plant 42:405–410; 2006.

Browne RA. Components of resistance to fusarium head blight (FHB) in wheat detected in a seed-germination assay with *Microdochium majus* and the relationship to FHB disease development

and mycotoxin accumulation from *Fusarium graminearum* infection. Plant Pathol 56:65–72; 2007.

Companioni B, Mora N, Arzola M, Ventura J, Perez MC, Santos R, Lorenzo JC. Improved technique for rapid and non-destructive *in vitro* differentiation between resistant and susceptible banana clones of *Fusarium oxysporum* f. sp. *cubense*. Biotechnol Lett 26:213–216; 2004.

Compton ME. Statistical methods suitable for the analysis of plant tissue culture data. Plant Cell Tissue Organ Cult 37:217–242; 1994.

Ibanez MA, Martin C, Perez C. Alternative statistical analyses for micropropagation: a practical case of proliferation and rooting phases in *Viburnum opulus*. In Vitro Cell Dev Biol Plant 39:429–436; 2003.

Martinez ME, Mora N, Quiñones J, González-Arnao MT, Engelmann F, Lorenzo JC. Effect of cryopreservation on the structural and functional integrity of cell membranes of sugarcane embryogenic calluses. Cryo-lett 23:237–244; 2002.

Mize CW, Chun YW. Analyzing treatment means in plant tissue culture research. Plant Cell Tissue Organ Cult 13:201–217; 1988.

Mokhtassi A, Akbari G, Mirhadi M, Zand E, Soufizadeh S. Path analysis of the relationships between seed yield and some morphological and phenological traits in safflower (*Carthamus tinctorius* L.). Euphytica 148:261–268; 2006.

Onyeka TJ, Pétro D, Ano G, Etienne S, Rubens S. Resistance in water yam (*Dioscorea alata*) cultivars in the French West Indies to anthracnose disease based on tissue culture-derived whole-plant assay. Plant Pathol. 55:671–678; 2006.

Perez A, Napoles L, Carvajal C, Hernandez M, Lorenzo JC. Effect of sucrose, inorganic salts, inositol and thiamine on protease excretion during pineapple culture in temporary immersion bioreactors. In Vitro Cell Dev Biol Plant 40:311–316; 2004.

Petersen RG. Use and misuse of multiple comparison procedures. Agron. J. 69:205–208; 1977.

Sharma R, Pandey-Chhetri B, Duveiller E. Heritability estimates of spot blotch resistance and its association with other traits in spring wheat crosses. Euphytica 147:317–327; 2006.

Yabor L, Arzola M, Aragón C, Hernández M, Arencibia A, Lorenzo JC. Biochemical side effects of genetic transformation of pineapple. Plant Cell Tissue Organ Cult 86:63–67; 2006.

Zhu Y, Geng C, Tong Y, Smith SE, Smith FA. Phosphate (Pi) and arsenate uptake by two wheat (*Triticum aestivum*) cultivars and their doubled haploid lines. Ann Bot 98:631–636; 2006.