this viroid is a new variant of HSVd and was tentatively given the name yellow corky vein strain of HSVd (HSVd–ycv, accession no. AJ490824). The present investigation constitutes the first record of detection of a HSVd variant in citrus in India and also molecular characterization of a viroid infecting citrus. This HSVd variant, named HSVd-ycv, is a new viroid variant, which merits investigation in terms of its pathogenic ability to other hosts.

1. Flores, R., Di-Serio, T. and Hernandez, C., *Semin. Virol.*, 1997, **8**, 65–73.
2. Diener, T. O., *Virology*, 1971, **45**, 411–428.
3. Hataya, T., *Recent Res. Dev. Viroid*, 1999, **1**, 789–815.
4. Singh, R. P., In *Proceedings of the Golden Jubilee International Conference on*, New Delhi, 2000, pp. 95–96.
5. Ramachandran, P., in *Diseases of Horticultural Crops—Fruits* (eds Varma, L. K. and Sharma, R. C.), Indus Publishing House, New Delhi, 1999, pp. 563–581.
6. Mishra, M. D., Hammond, R. W., Owens, R. A., Smith, D. R. and Diener, T. O., *J. Gen. Virol.*, 1991, **72**, 1–5.
7. Reddy, G. S., Dakshinamurty, V. and Reddy, V. R. K., *Indian Phytopathol.*, 1974, **27**, 82–84.
8. Reddy, M. R. S. and Naidu, P. H., *Indian Phytopathol.*, 1989, **42**, 455–456.
9. Azad, P., In Proceedings of the 12th Conference of International Organization Citrus Virol., IOCV, Riverside, 1993, p. 458.
10. Rustem, F. A. and Ahlawat, Y. S., *Indian J. Agric. Sci.*, 1999, **69**, 789–794.
11. Rustem, F. A., Ramachandran, P. and Ahlawat, Y. S., In Proceedings of the Golden Jubilee International Conference, New Delhi, 2000, pp. 778–779.
12. Semancik, J. S., Szychowski, J. A., Rakowski, A. G. and Symons, R. H., *J. Gen. Virol.*, 1993, **74**, 2427–2436.
13. Singh, R. P. and Boucher, A., *Phytopathology*, 1987, **77**, 1588–1591.
14. Sambrook, J. and Russell, D. W., *Molecular Cloning – A Laboratory Manual*, CSHL Press, Cold Spring Harbor, New York, 2001.
15. Sanger, F., Nicklen, S. and Caulson, A. R., *Proc. Natl. Acad. Sci. USA*, 1977, **74**, 5463–5467.
16. Ole, M. and Anders, W., *CABIOS*, 1996, **12**, 247–249.
17. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G., *Nucleic Acids Res.*, 1997, **24**, 4876–4882.
18. Flores, R., Randles, J. W., Bar-Joseph, M. and Diener, T. O., *Arch. Virol.*, 1998, **143**, 623–629.
19. Reanwarakorn, K. and Semancik, J. S., *Phytopathology*, 1999, **89**, 568–574.

ACKNOWLEDGEMENTS. We thank Prof. R. F. Lee, Citrus Research Centre, Lake Alfred, FL, USA for the design and synthesis of primers and Prof. Y. S. Ahlawat, Head, Plant Virology Unit, Division of Plant Pathology, IARI, New Delhi for facilities and constant encouragement and Dr V. G. Malathi for sequence analysis. AR thanks the Dean, P.G. School, IARI for financial assistance rendered through IARI Merit Fellowship.

Received 25 October 2002; revised accepted 15 July 2003

Prediction of seed longevity in the genebank: How reliable are the estimates?

R. L. Sapra*,†, Prem Narain, S. R. Bhat‡, S. K. Lal§ and S. K. Jain||**

*Bioinformatics Centre, Library Building, Indian Agricultural Research Institute, New Delhi 110 012, India
§Division of Genetics, Indian Agricultural Research Institute, New Delhi 110 012, India
‡National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, New Delhi 110 012, India
||Division of Seed Technology, Indian Agricultural Research Institute, New Delhi 110 012, India
**B–3/27A, SFS Flats, Lawrence Road, New Delhi, India

Germplasms of crop plants are stored as seeds in genebanks at low temperature with low seed moisture content where they remain viable for several decades. The longevity of seeds stored in genebanks is predicted using viability equations developed by subjecting seeds to accelerated ageing under controlled laboratory conditions. We discuss here the serious consequences of provisionally or unreliably developed estimates of seed longevity parameters in making such predictions. A slight under or over estimation to the tune of 0.01 of the linear temperature parameter may result in a difference ranging from 46 to 74 years in the expected longevity. Whereas in case of the quadratic temperature parameter, a minor estimation difference (0.0001) may cause a difference of 11 to 12 years. A nonlinear estimation method based on Levenberg–Marquardt iterative convergence algorithm was applied for the reliable estimation of viability parameters for *Lupinus polyphyllus* seeds. The said estimation procedure resulted into comparatively narrow confidence intervals; and almost four to five times gain in precision over the conventional linear estimation in estimating potential longevity and moisture sensitivity parameters.

SEEDS are stored in the genebank under low moisture and temperature conditions to enhance their longevity. Prediction of storability of samples is essential to plan periodic regeneration and replacement. Seed longevity is mainly influenced by the environmental conditions such as storage temperature and moisture content of the seeds. Since seeds remain viable for several decades under practical storage conditions, conducting real time experiments to know storability of seeds are not feasible. Instead, seed longevity is determined under accelerated ageing conditions (i.e. high temperature and high moisture) and these results are extrapolated to predict longevity under genebank storage conditions. During the last three decades several attempts have been made to quantify the relationship between seed longevity and storage environment. Such relationships have been described by the viability/

†For correspondence. (e-mail: saprarl@rediffmail.com)

longevity equations that predict the viability of seeds at any point of time for varying combinations of temperature and moisture.

Estimation of longevity parameters requires utmost precision as inaccurate estimates would lead to unreliable predictions whose consequences could be disastrous. For example, an underestimation would put a burden on maintenance by recommending frequent regeneration while an overestimation of longevity would result in loss of germplasm. These inaccurate predictions may go undetected under the accelerated seed ageing conditions that are commonly used to test these parameters. However, when predicted for favourable storage conditions (cooler and drier), the deviations from the true values become acutely magnified. Thus, a choice of proper estimation procedure or model is equally important in addition to the careful planning, designing and conducting of the experiment for achieving reliability in the results.

For the last three decades, the linear model fitting approaches[1,2] are being extensively used for quantifying the relationship between seed viability and environmental conditions. Though the relationship between viability and environmental conditions is a nonlinear one, the longevity parameters for most of the species have been estimated using the linear model approach[2]. Nonlinear model estimation has become comparatively easier now and is being applied to a wide range of situations, even to finite populations. Wilson and co-workers[3] suggested a single-step nonlinear regression for computational convenience for estimating the viability parameters of field bean seeds (*Phaseolus vulgaris* L.). Stahl and Steiner[4] also used SAS nonlinear procedure (NLIN) for estimating species-specific constants for wheat seeds without stating any specific advantage for doing so.

Here, we re-examine the data and results of Dickie *et al.*[5] for *Lupinus polyphyllus* and suggest the use of nonlinear regression procedure. We compare our results based on the nonlinear model with those of Dickie *et al.*[5] linear model and show that the use of nonlinear model yields more precise estimates that are free from transformation bias. The issue of reliability has also been emphasized in relation to accurate predictions.

The modified viability equations developed by Ellis and Roberts[2] relate the decline of viability with the period of storage (p), seed moisture content (m) and storage temperature (t), and can be expressed as

$$v = K_i - p/10^{K_E - C_w \log m - C_H t - C_Q t^2}, \quad (1)$$

where v is viability on a probit scale and K_i is the probit of percentage of viability at the beginning of the storage period and is specific to the seed lot. K_E is the potential longevity, C_w is the moisture sensitivity parameter, C_H is the linear temperature sensitivity parameter and C_Q is the quadratic temperature sensitivity parameter. These parameters are estimated in two stages. The first stage estimates the initial viability (K_i) and slopes of seed survival curves ($1/L$) for a number of environments using the probit analysis[6] by fitting the following model:

$$v = K_i - p(1/L). \quad (2)$$

The second stage estimates the species-specific longevity parameters (K_E, C_H, C_W and C_Q) using the following model:

$$L = 10^{K_E - C_w \log m - C_H t - C_Q t^2}. \quad (3)$$

The parameters of the model (3) are estimated after transforming the longevity on a log scale. The model finally fitted through multiple linear regression analysis is:

$$\log L = K_E - C_w \log m - C_H t - C_Q t^2 + e, \quad (4)$$

where e is an error term and assumed to be independently, identically and normally distributed with constant variance. Here, it is worth mentioning that one must be especially careful to check that the above said least square assumptions are not violated after making the desired transformation. In practice, the transformation chosen may not achieve the desired assumptions of independence, constancy and normality of errors. Therefore, the residuals from fit of the transformed data should be checked for these assumptions, as the transformation as such may not be successful. Currie[7] studied the Michaelis–Menton model and found that the best linearizing inverse transformation produced unreliable estimates and concluded that transformation should not be used except in cases where it stabilizes the variance.

Thus, we see here that the longevity model (3) is a nonlinear one and has been linearized (4) for the purpose of estimation of longevity constants. The said transformation has been applied widely for the estimation of longevity parameters of most of the species. Though this works well, it does not give a clear interpretation of the parameters, as an un-transformed longevity model would do so. Besides, simply de-transforming the usual least square prediction equation can lead to severe bias[8]. Thus, when we are led to a model of nonlinear form, we should usually prefer to fit such a model, whenever possible, rather than to fit an alternative, perhaps less realistic, linear model[9].

In case of nonlinear models, the asymptotic theory does not require normally distributed errors. In model (3), we assume the random error in the multiplicative form. A nonlinear model assumes the random error e in the additive form as follows:

$$L = 10^{K_E - C_w \log m - C_H t - C_Q t^2} + e. \quad (5)$$

Table 1. Comparison of various residuals based on linear and nonlinear models using data of Dickie *et al.*[5]

Temp. (°C)	Moisture (%)	Longevity (days)	Log of longevity	Residuals (days) in log scale linear model	De-transformed residuals (days) linear model	Residual (days) nonlinear model
21.00	7.91	526.32	2.72	0.00190	2.30	2.28
21.00	11.73	178.57	2.25	0.00569	2.32	–28.87
21.00	14.11	169.49	2.23	0.20489	63.75	35.15
42.00	7.92	84.75	1.93	0.23107	34.97	13.82
42.00	11.58	12.67	1.10	–0.13788	–4.74	–16.35
42.00	13.35	3.48	0.54	–0.52876	–8.27	–17.29
62.00	8.13	3.30	0.52	–0.17445	–1.63	–6.65
62.00	11.93	2.28	0.36	0.12586	0.57	–1.76
62.00	14.32	1.93	0.29	0.27169	0.90	–0.70

Table 2. Estimates of various parameters based on linear model (6) (Dickie *et al.*[5])

Parameter	Estimate	Std. error	95% confidence interval Lower	Upper
K_E	6.223945	1.026285	3.7126256	8.7352644
C_W	2.765399	0.963299	0.4082063	5.1225916
C_H	0.048609	0.005832	0.0343381	0.0628799

In the present investigation we use the seed survival data of Dickie *et al.*[5] (slopes of curves estimated for nine environments). Statistical analysis was carried out using the Windows version of SPSS package. For nonlinear regression, Marquardt iterative method was chosen as it represents a compromise between the linearization (Gauss–Newton) and the steepest descent method. It almost always converges and appears to work well in many circumstances[9]. Dickie *et al.*[5] obtained estimates of longevity parameters for *Lupinus polyphyllus* species by fitting the following multiple linear regression model (6) to the data shown in Table 1.

$$\log L = K_E - C_w \log m - C_H t + \log e. \quad (6)$$

The fit of the model (6) seems to be reasonably good as the model has explained almost 93 per cent of the variation in longevity ($R^2 = 0.92975$). However, the standard errors for the potential longevity (K_E) and moisture sensitivity (C_W) are unexpectedly very high (Table 2). When the 'Studentized' residuals from fit of the model (6) are plotted against the predicted values we get a pattern of residuals (Figure 1). On examining these residuals, no serious deviations were found from the usual ANOVA assumptions. However, the error corresponding to sixth observation (–0.52876, Table 1) is slightly disturbing and perhaps, this may be the cause of high standard errors associated with the said longevity estimates.

The results of application of nonlinear regression model (7) to the data of Dickie *et al.*[5] are summarized in Table 3.

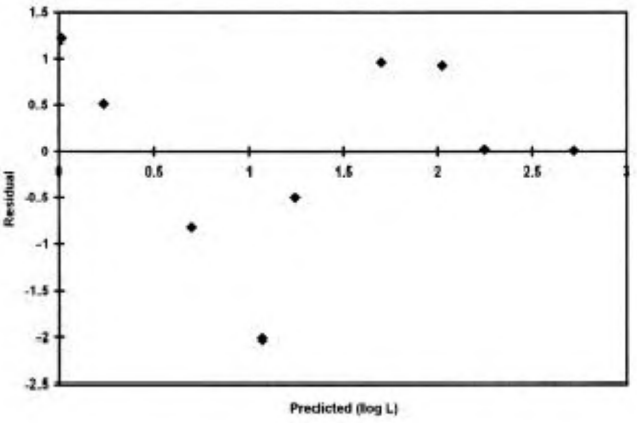

Figure 1. Scatter diagram between predicted seed longevity and residuals (linear model).

$$L = 10^{K_E - C_w \log m - C_H t} + e. \quad (7)$$

On comparing the results of linear model fitting (Table 2) with those of nonlinear model fitting (Table 3), the nonlinear estimation procedure yielded highly precise estimates of the longevity parameters. The standard errors of the estimates for moisture sensitivity (C_W) and potential longevity (K_E) parameters for linear model (Table 2) are almost four to five times those of nonlinear model (Table 3). Confidence intervals generated by the nonlinear model are comparatively narrow for all the parameters. For example, the 95 per cent estimated confidence interval for linear model indicates that the true value of moisture sensitivity parameter lies somewhere between 0.41 and 5.12, whereas the nonlinear model says it lies between 1.83 and 2.87. The same is true for potential longevity parameter. Further, if we compare the maximum absolute deviations (last two columns of Table 1) between the two models, we find it is almost half (35 days vs 65 days). Thus, the nonlinear model gives a better performance when compared to a linear one for the present data set.

Table 3. Estimates, standard errors and confidence intervals of nonlinear regression parameters

Parameter	Estimate	Asymptotic Std. error	Asymptotic 95% confidence interval	
			Lower	Upper
K_E	5.699109257	0.230912814	5.134085955	6.264132558
C_W	2.351936934	0.211937043	1.833345672	2.870528196
C_H	0.041299700	0.005558097	0.027699526	0.054899874

Table 4. Predicted values of longevity for the linear and nonlinear model under interpolated as well as extrapolated conditions

Temperature (°C)	Moisture content (%)	Predicted L (linear model) in days	Predicted L (nonlinear model) in days
–20	2	2310239	656295
–10	2	754339	253572
0	2	246307	97972
10	2	80424	37853
20	2	26260	14625
30	2	8574	5650
40	2	2799	2183
50	2	914	843
–20	5	183313	76062
–10	5	59855	29388
0	5	19544	11354
10	5	6381	4387
20	5	2083	1695
30	5	680	654
40	5	222	253
50	5	72	97
–20	10	26960	14899
–10	10	8803	5756
0	10	2874	2224
10	10	938	859
20	10	306	332
30	10	100	128
40	10	32	49
50	10	10	19
–20	20	3965	2918
–10	20	1294	1127
0	20	422	435
10	20	138	168
20	20	45	65
30	20	14	25
40	20	4	9
50	20	2	4

Table 5. Effect of slight over/underestimation of various parameters on predicted longevity of seeds stored at –20°C and 5% moisture

K_E	C_W	C_H	C_Q	Longevity (years)
5.7	2.4	0.04	0.0004	125.9
5.8	2.4	0.04	0.0004	158.6
5.6	2.4	0.04	0.0004	100.0
5.7	2.5	0.04	0.0004	107.2
5.7	2.3	0.04	0.0004	147.9
5.7	2.4	0.05	0.0004	199.6
5.7	2.4	0.03	0.0004	79.5
5.7	2.4	0.04	0.0005	114.9
5.7	2.4	0.04	0.0003	138.1
5.8	2.5	0.05	0.0005	195.1

The effect of over- or under-estimation of parameters on longevity may not appear great under adverse accelerated ageing conditions (high temperature or high seed moisture) that are commonly employed to test these predictions. However, this may lead to a difference of several decades in predicted longevity, particularly under the favourable storage conditions. Normally, seeds are stored in the genebank at –10 or –20°C with a seed moisture of 5%. We work out the longevity for these two conditions. The longevity predicted at –10°C and 5% moisture content using the linear model is around 60,000 days (Table 4), whereas with the nonlinear model it is around 30,000 days – a difference of 83.47 years. Again, at –20°C and 5% moisture, the difference between the two estimates is over 0.1 million days (273 years).

The seriousness of the problem of slight over- or underestimation, perhaps, can be best appreciated by looking at Table 5. Let us assume that the true values of these species-specific constants for a particular species are 5.7 (K_E), 2.4 (C_w), 0.04 (C_H) and 0.0004 (C_Q); and the corresponding estimated values are 5.8, 2.5, 0.05 and 0.0005. If the seeds are stored at –20°C with 5% moisture content, then according to our calculations, the difference in longevity due to this petty overestimation is approximately 69 years! A slight under- or over-estimation to the tune of 0.01 of the linear temperature parameter (C_H) may result in a difference ranging from 46 to 74 years in the expected longevity. Whereas in case of the quadratic temperature parameter (C_Q), a minor estimation difference (0.0001) may cause a difference of 11 to 12 years. It is certainly more dangerous to have estimates which overestimate the longevity than to those which underestimate it. Overly estimated predictions will send a wrong signal that the seed is alive, whereas in reality it would have perished several years ago. Underestimation would put unnecessary burden on the genebank maintenance by requiring more frequent regeneration of stored seeds (viability above the regeneration standard). Thus, in the light of the above discussions, one should be extra cautious in using the provisional estimates[5,10] or estimates with low precision, developed in the literature over the years for the long-term prediction of seed longevity.

The theory of linear estimation of regression parameters is being extensively used owing to its computational convenience. However, the computational simplicity that comes with the linear model is often lost when one is faced with obtaining point and interval estimates for the original rather than the transformed parameters. For example, say, when the emphasis is on estimating the expected longevity (L) for a given moisture (m) and temperature (t), simply de-transforming the usual least squares prediction equation can lead to severe bias. Linear model estimation gives us the estimate of Log(L). Thus, it is the naive estimate of the longevity that we estimate by taking antilog of estimate of Log(L). If errors of Log(L) are symmetric then errors of L are asymmetric and thus, estimate of L tends to estimate the median rather than the mean of the distribution of L. However, before adopting the nonlinear estimation procedure for other species, it would be worth studying the behaviour of the model under experimental conditions and validating the model for extrapolated conditions. For validating the model one may generate two sets of data. The first set may be used for estimating the parameters and the other for validating the model. Utmost attention should be paid to estimate the parameters with high degree of precision. If the linearizing transformation is not successful, particularly in stabilizing the error variance, one should not go for the ordinary linear least square estimation procedure. Estimates with low precision should not be released as they may prove more damaging. Trusting on such estimates may ultimately lead to a loss of valuable genetic wealth stored for long-term genetic conservation.

1. Roberts, E. H., Storage environment and control of viability. In *Viability of Seeds* (ed. Roberts, E. H.), Chapman and Hall, London, 1972, p. 448.
2. Ellis, R. H. and Roberts, E. H., Improved equations for the prediction of seed longevity. *Ann. Bot.*, 1980, **45**, 13–30.
3. Wilson, D. O. Jr., McDonald, M. B. Jr. and Martin, S. K. St., A probit planes method for analyzing seed deterioration data. *Crop. Sci.*, 1989, **29**, 471–476.
4. Stahl, M. and Steiner, A. M., Germination and vigor loss of non-sprouted and sprouted wheat seeds during storage – testing the viability constants. *Seed Sci. Res.*, 1998, **8**, 123–128.
5. Dickie, J. B., McGrath, S. and Linington, S. H., Estimation of provisional seed viability constants for *Lupinus polyphyllus* Lindley. *Ann. Bot.*, 1985, **55**, 147–151.
6. Finney, D. J., In *Probit Analysis*, Cambridge University Press, London, 1962, p. 318.
7. Currie, D. J., Estimating Michaelis–Menten parameters: Bias, variance and experimental design. *Biometrics*, 1982, **38**, 907–919.
8. Seber, G. A. F. and Wild, C. J., In *Nonlinear Regression*, John Wiley & Sons, New York, 1989, p. 753.
9. Draper, N. R. and Smith, H., In *Applied Regression Analysis*, John Wiley & Sons, New York, 1981, p. 709.
10. Dickie, J. B. and Bowyer, J. T., Estimation of provisional seed viability constants for Apple (*Malus domestica* Borkh cv. Greensleevs). *Ann. Bot.*, 1985, **56**, 271–275.

Received 20 February 2003; revised accepted 28 July 2003

Use of image analysis to study the effect of phosphate on honeydew formation and clavine alkaloid synthesis in *in vitro* cultures of *Claviceps fusiformis* Lov.

Asha Jacob* and Paramjeet Jite

Department of Botany, University of Pune, Pune 411 007, India

Image analysis has become an important tool in the study of growth and nutrition in fungi. Here this technique has been used successfully to study the effect of phosphate on honeydew formation and clavine alkaloid synthesis in *in vitro* cultures of *Claviceps fusiformis* Lov. Specific colour reagents for the detection of clavine alkaloids and phosphate are used to demarcate the region of alkaloid production and the pattern of phosphate utilization by the *C. fusiformis* colony. The intensity of colour formed is then measured using image-analysis techniques, and regions of similar intensity are coloured alike using pseudocolours. This enables the division of the colony according to its biochemical make-up, to an accuracy not achieved by the naked eye. Growth of the colony can also be monitored nondestructively on a daily basis. It is seen that as the colony grows, phosphate in the medium is depleted, disrupting growth and causing a decrease in intracellular phosphate. This leads to honeydew formation accompanied by enhanced clavine alkaloid production.

CLAVICEPS fusiformis Lov. causes ergot, a commonly occurring disease on members of Gramineae. Ergot alkaloids cause ergotism in men and cattle on consumption, but have gained importance owing to their pharmacological properties[1]. These clavine alkaloids are also used as a precursor in the synthesis of lysergic acid diethylamide or LSD[2]. Parasitic cultivation of this fungus besides bearing the risk of poisoning, also demands large areas of cultivation and consequent loss of food crop[3]. Therefore, fermentative cultivation of this fungus where the process can be precisely controlled, is a welcome alternative. To optimize nutritional conditions that would maximize biomass and alkaloid production, it is important to raise *in vitro* honeydew stage, as the latter marks the advent of secondary metabolism[3] and the production of alkaloids. *In vivo* honeydew formation requires entry of long filamentous hyphae raised from germinating ascospore of *C. fusiformis* through the stigma and into the ovary[4] of the flower. These come in contact with the epidermal tissue at the base of the ovary and cause plasmolyses. The infected hyphae feeds on the host sucrose and converts it into fructose and glucose by the action of β-D-fructo-

*For correspondence. (e-mail: ashajacob2000@hotmail.com)