In Vitro Cell. Dev. Biol.—Plant 37:539–542, September–October 2001
DOI: 10.1079/IVP2001221

1054-5476/01 $10.00+0.00

SOMACLONAL VARIATION RATE EVOLUTION IN PLANT TISSUE CULTURE: CONTRIBUTION TO UNDERSTANDING THROUGH A STATISTICAL APPROACH

F. X. CÔTE*, C. TEISSON, AND X. PERRIER

CIRAD, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, CIRAD-Flhor & CIRAD-Biotrop, TA 50/04 Avenue Agropolis, 34398 Montpellier Cedex 5, France

(Received 11 August 2000; accepted 11 May 2001; editor T. A. Thorpe)

Summary

In order to better understand somaclonal variant rate evolution in plant tissue culture, a statistical approach has been adopted. According to this approach, the variant percentage could be calculated by: $\%V = [1 - (1 - p)^n] \times 100$, where %V is the percentage of variant, p the probability of variation and n the number of multiplication cycles. A numerical estimation was performed to characterize the variance of this function. It has been demonstrated that a wide scale of variance is associated with '%V', due to the occurrence of variations after a variable number of multiplication cycles in the different lines of culture. Two main conclusions can be drawn from this model: (1) a variant rate increase can be expected as an exponential function of the number of multiplication cycles; (2) after a given number of multiplication cycles, variable off-types percentages can be expected. Due to the complexity of biological systems, this statistical approach could obviously not be applied directly for the calculation and forecasting of variant rates in tissue culture. However, this approach results in a better understanding of two apparently confusing experimental features often reported in tissue culture: the increase of the variant rate as a function of the length of the culture period on the one hand, and, on the other hand, the observations of different variant rates among lines cultured for the same lengths of time under strictly identical culture conditions. This approach also underlined that the comparison of somaclonal variant percentage between batches of plants from different *in vitro* treatments could be, in some cases, insufficient for ascertaining a difference of variability generated by tissue culture.

Key words: genetic stability; micropropagation; somaclonal variation frequency; somaclonal variation probability.

Introduction

In many species micropropagated by tissue culture, some of the regenerated plants do not conform to the source plant material. The term 'somaclonal variation' has been proposed to describe the variability produced by *in vitro* multiplication (Larkin and Scowcroft, 1981). For Meins (1983), it applies to stable and transmissible modifications, i.e., changes which persist in the absence of the event that induced them and which are transmitted during mitosis.

The possible origins of variations have been discussed in several reviews (Swartz, 1990; Karp, 1991, 1995). In numerous *in vitro* multiplication systems, it has been observed that variant rates increase with the number of multiplication cycles or with the length of time in culture (Lörz and Scowcroft, 1983; Sree Ramulu et al., 1984; Cassells and Morrish, 1987; Lee and Phillips, 1987; Benzion and Phillips, 1988; Gaponenko et al., 1988; Hartmann et al., 1989; Gözükirmizi et al., 1990; Morrish et al., 1990; Müller et al., 1990; Wang et al., 1992; Gavidia et al., 1996; Yang et al., 1999). However, in several *in vitro* regeneration systems, it has also been observed that progenies of plants derived from the same initial source material and multiplied following a strictly identical culture protocol, including the same length of time in culture, exhibit variable percentages of off-types (Fukui, 1983; Sree Ramulu et al., 1983, 1984; Karp and Maddock, 1984; Benzion and Phillips, 1988; Wang et al., 1992). In order to better understand variant rate evolution in tissue culture, a statistical approach to somaclonal variation has been adopted. A mathematical model has been developed to describe the theoretical variant rate evolution in tissue culture. The possibility of interpreting part of the complex situation of biological systems by applying this model is discussed.

Hypotheses of Calculation

Four hypotheses have been retained for the variant percentage calculation: (1) the variations considered in this calculation are those which are stable and transmissible; (2) a micropropagation system with the occurrence of only one variation type has been considered; (3) the probability of variation has been considered to be the same during the successive multiplication cycles; and (4) the multiplication rate of the variant and that of true-to-type plants have been considered to be identical.

It is obvious that such hypotheses could not reflect the complexity of the biological systems. However, we chose such simple hypotheses deliberately to facilitate the theoretical calculation. The validity and limits of such hypotheses and, consequently, those of the statistical approach will be discussed later.

*Author to whom correspondence should be addressed: Email cote@messmpl.cirad.fr

The statistical approach developed here could be applied indifferently to plant micropropagation or to cell multiplication. However, in order to facilitate the description of the model, we chose to use only terms referring to plant micropropagation in this paper.

Variant Percentage Calculation

The number of true-to-type plants after n multiplication cycles is equivalent to the number of true-to-type plants after $n-1$ cycles multiplied by the expected probability of getting a true-to-type plant and by the multiplication rate. This is equivalent to:

$$T_n = T_{n-1}(1-p)t \quad (1)$$

where T_n is the number of true-to-type plants after n multiplication cycles, T_{n-1} the number of true-to-type plants after $n-1$ multiplication cycles, p the expected variation probability (the value '$1-p$' being the expected probability to get a true-to-type plant), and t the multiplication rate.

The recurrence on T_n using equation 1 makes it possible to write:

$$T_{n-1} = T_{n-2}(1-p)t, \ldots, T_1 = T_0(1-p)t$$

where T_{n-1} is the number of true-to-type plants after $n-1$ multiplication cycles, T_{n-2} the number after $n-2$ multiplication cycles, T_1 the number after 1 cycle and T_0 the initial source plant at the origin of the multiplication. This initial plant being a true-to-type plant, we find $T_0 = 1$; it can thus be deduced from equation 1 that:

$$T_n = (1-p)^n t^n \quad (2)$$

The percentage of variants after n multiplication cycles is equal to:

$$\%V_n = [(t^n - T_n) \times 100]/t^n \quad (3)$$

where $\%V_n$ is the percentage of variants after n multiplication cycles, T_n the number of true-to-type plants after n multiplication cycles, t the multiplication rate, and n the number of multiplication cycles (t^n being the total number of plants produced after n multiplication cycles).

Combining equations 2 and 3 makes it possible to write:

$$\%V_n = [t^n - (1-p)^n t^n] \times 100/t^n$$

and finally:

$$\%V_n = [1 - (1-p)^n] \times 100 \quad (4)$$

The evolution of the theoretical percentage of variants during the successive multiplication cycles is thus an exponential function which tends to 100% as n tends to infinity. Moreover, the percentage of somaclonal variants appears independent of t, the multiplication rate.

Estimation of the Variance of the Variant Percentage

The percentage of variants ($\%V_n$) calculated with equation 4 represents the theoretical average percentage of off-types of an infinity of lines (in this paper, a line or a progeny designates the total of plants derived from the same initial source plant). This calculated average percentage is the mean of the variant percentages in each line. The analytical estimation of the variance of this calculated average percentage is no longer easily obtainable when n is larger than 3. We thus chose a numerical estimation to characterize this variance. The estimation used a simple computer program which simulated the multiplication of the initial source plant of a line over n multiplication cycles with a given t value of multiplication rate. The program also simulated the variation occurrence with an expected probability p. This has been obtained by computer selection at random of a uniform variable between 0 and 1 for each plant produced. If the variable was equal or inferior to the chosen value of p, then this plant and all plants derived from this plant were considered as variant plants. The computerized simulation of a large number of progenies (500) was carried out in order to obtain a distribution of the variation percentages between lines with an average close to the expected value of $\%V_n$ calculated in equation 4. For these simulations, the value of t was chosen as 2, the value of n was chosen as 4, 8, or 12 and the value of p was chosen as 10^{-2}, respectively. These values were chosen according to bibliographical references. The values $t = 2$ and $n = 4$, 8, or 12 correspond, respectively, to plausible values of the multiplication rate and multiplication cycle numbers in micropropagation through *in vitro* budding or *in vitro* cutting. The value $p = 10^{-2}$ is comparable to the variation frequency corresponding to the substitution of a nucleotide of the allele *Adh1* in maize multiplied *in vitro* (Dennis et al., 1987). A variation frequency with a comparable value can also be deduced from the studies by Reuveni and Israeli (1990) on banana micropropagation through *in vitro* budding. Lower frequencies ranging from 2×10^{-7} to 8.2×10^{-3} have been reported by Meins (1983) for cell cultures of *Nicotiana tabacum*.

The distribution frequency of the variant percentage in the 500 lines obtained after the computerized simulation is given in Fig. 1a, b, c for $n = 4$, 8, 12, respectively. As expected, the modes of the distributions for the different simulations were related to the number of multiplication cycles. For $n = 4$, $n = 8$, and $n = 12$, these modes were equal to the classes 0–2%, 2–4%, and 8–10%, respectively (Fig. 1a, b, c). The simulations also clearly illustrated the broad distribution of the variant percentage in the different lines for the different values of n studied. For $n = 4$, $n = 8$, and $n = 12$, the expected mean value of the variant percentage calculated with equation 4 was 3.94%, 7.73%, and 11.36%, respectively. However, the simulations demonstrated that variant percentages ranging from a few percent to more than 50% were observed in some lines for the different values of n (Fig. 1a, b). As an example, for $n = 12$ the mean of the distribution, equal to 11.40%, was very close to the expected value of 11.36% calculated by equation 4. However, a mode of 8–10% of variants was observed for approximatively 28% of the lines and values ranging from 4% to 62% of variants were observed around this mode (Fig. 1c). The dispersion was particularly important for the high values, giving an asymmetrical distribution. Clearly identified secondary modes were observed for the values of 32% of variants (1.4% of the lines) and 54% of variants (1.2% of the lines) (Fig. 1c). This distribution is the consequence of the more or less early occurrence of variations in the different lines during the successive multiplication cycles. The peaks at 32% and 54% of variants correspond, for example, to progenies where the first variation appeared, respectively, during the second and first multiplication cycle.

While broad distributions of the variant percentages were observed for the different number of multiplication cycles, these distributions exhibited a more complex pattern with increased values of n. As an example, for $n = 4$ only nine classes of lines were

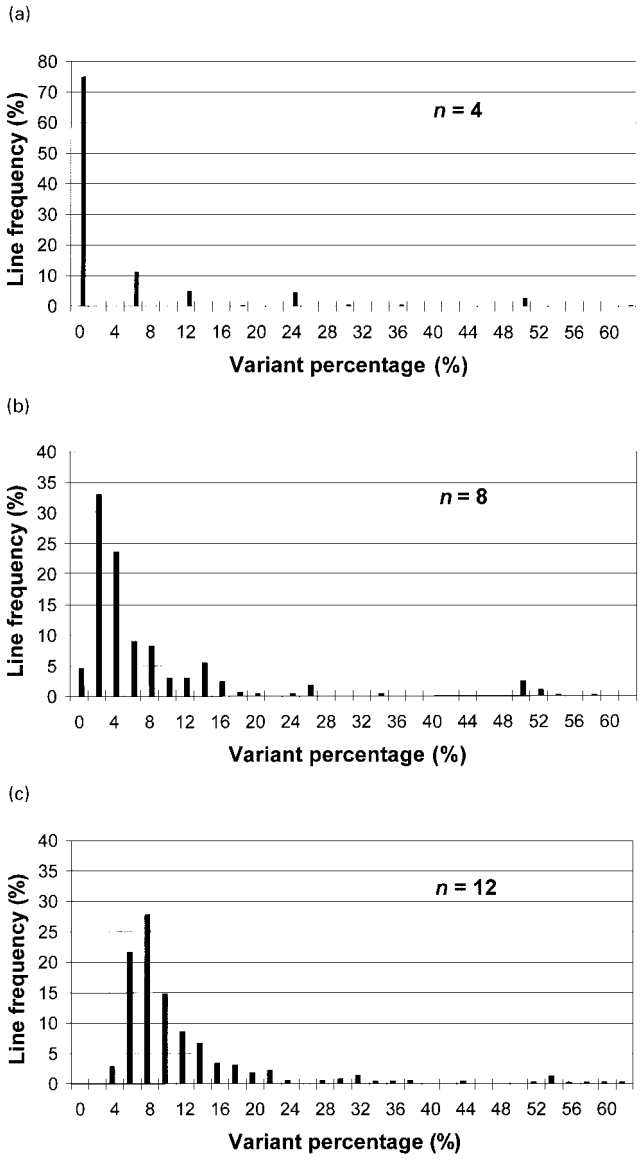

FIG. 1. Distribution frequency of the theoretical variant percentage among different progenies (*lines*) according to a statistical model of somaclonal variation occurrence in plant tissue culture. *a*, *b*, *c*, Distribution frequency after 4, 8, 12 multiplication cycles, respectively.

identified with 75% of the lines concentrated into the mode (Fig. 1a), while 18 and 24 classes were observed for $n = 8$ and $n = 12$, with a maximum of 33% and 28% of the lines concentrated into the mode, respectively (Fig. 1b, c).

DISCUSSION

The first calculation hypothesis retained supposed that the variation was transmitted to the whole *in vitro* progeny of a plant that undergoes a variation. This hypothesis is verified if genetic variations are considered. However, alterations in epigenic characteristics are generally not inherited. In this case the statistical approach proposed would not apply.

The second calculation hypothesis supposed the occurrence of only one variation type. It is obvious that various variation types could affect a regeneration process and that each of these types would exhibit a peculiar probability of variation. However, the arguments developed in the statistical approach are separately applicable for each of the variation types considered. In a regeneration process with different variant types, the global variant percentage will be the sum of the variant percentages calculated for each type of variation.

The third calculation hypothesis supposed that the probability of variation remained the same during the successive multiplication cycles. This is not necessary the case because the *in vitro* multiplication process may consist of distinct stages during which the plant tissue undergoes developmental change. Thus, it cannot be excluded that somaclonal variations different in nature and frequency are possible for each of these stages. However, while various values of p corresponding to the successive stages of the multiplication process will complicate the theoretical calculation, the arguments developed in the statistical approach based on a constant value of p will still be applicable to each of the identified stages of multiplication.

The fourth calculation hypothesis supposed that the multiplication rate of the variants and that of the true-to-type plants were identical. It is, however, likely that these rates are different. The variation could indeed be responsible for changes that reduce the multiplication capacity. In several culture systems it has, for example, been established that the accumulation of chromosomal anomalies in cell lines results in reduced growth rates, sometimes associated with a complete loss of regeneration capability (Swedlund and Vasil, 1985; Singh, 1986; Gaponenko et al., 1988; Franklin et al., 1989; Wang et al., 1992).

In conclusion, this discussion illustrates the possible contribution of a statistical approach to the understanding of variant rate evolution in tissue culture. This approach could not be applied directly for the calculation and forecasting of variant rate. Biological systems do not act in precise mathematical patterns. On the other hand, this statistical approach results in a better understanding of experimental features frequently observed in tissue culture.

A simple mathematical function linked to the transmissibility of variant traits can be one explanation of the often reported increase in the variant rates as a function of subcultures or of the culture period. Benzion and Phillips (1988) have suggested a similar interpretation to explain the increase in the percentage of chromosomal anomalies as a function of the length of time in culture in maize propagated from callus.

There is no incompatibility between the findings of the literature reporting an increase in the percentage of variants along with the number of multiplication cycles and those reporting variable percentage of variants among different lines cultured for the same length of time with strictly identical culture conditions. As the simulation approach emphasizes, the latter observation may indeed be the consequence of the large variance associated with the variant percentage due to the appearance of variations after a variable number of multiplication cycles in the different progenies. The statistical approach proposed helps to better understand experimentally data showing that variations of an identical nature occur after variable numbers of multiplication cycles or time in culture in different systems of *in vitro* multiplication (Sree Ramulu et al., 1984;

Benzion and Phillips, 1988; Sandoval et al., 1991; Vuylsteke et al., 1991).

The conclusions underlined by the statistical approach described here are similar to those of Luria and Delbrück (1943) who proposed a mathematical approach for the distribution of mutation occurrence in *Escherichia coli* cultures. These authors demonstrated that a distribution according to Poisson's law was obtained between lines for a constant frequency of mutation due to the more or less early occurrence of these mutations.

The statistical approach also emphasizes that the comparison of somaclonal variant percentages between batches of plants from different *in vitro* treatments could be insufficient for assertaining a difference of variability between them. For example, while the expected mean of the variant percentage was lower after four multiplication cycles than after 12 multiplication cycles, the computer simulations clearly demonstrated that some lines after four multiplication cycles exhibited more than 60% of variants while some lines after 12 multiplications cycles exhibited less than 5%. In order to characterize the extent of the variability generated by a tissue culture treatment, a powerful experimental design able to take into account the variance of the percentage of variation is indispensable.

Acknowledgment

We are grateful to R. Domergue for his help through the development of a computer program.

References

Benzion, B.; Phillips, R. L. Cytogenetic stability of maize tissue cultures; a cell line pedigree analysis. Genome 30:318–325; 1988.

Cassells, A. C.; Morrish, F. M. Variation in adventitious regenerants of *Begonia rex* Putz. Lucille Closon as a consequence of cell ontogeny, callus ageing and frequency of subculture. Sci. Hortic. 32:135–144; 1987.

Dennis, E. S.; Brettell, R. I. S.; Peacock, W. J. A tissue culture induced *Adh1* mutant of maize results from a single base change. Mol. Gen. Genet. 210:181–183; 1987.

Franklin, C. I.; Mott, R. L.; Vuke, T. M. Stable ploidy levels in long-term callus cultures of loblolly pine. Plant Cell Rep. 8:101–104; 1989.

Fukui, K. Sequential occurrence of mutations in a growing rice callus. Theor. Appl. Genet. 65:225–230; 1983.

Gaponenko, A. K.; Petrova, T. F.; Iskakov, A. R.; Sozinov, A. A. Cytogenetics of *in vitro* cultured somatic cells and regenerated plants of barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 75:905–911; 1988.

Gavidia, I.; del Castillo Agudo, L.; Perez-Bermudez, P. Selection and long-term cultures of high-yielding *Digitalis obscura* plants: RAPD markers for analysis of genetic stability. Plant. Sci. 121:197–205; 1996.

Gözükirmizi, N.; Ari, S.; Orale, G.; Okatan, Y.; Unsal, N. Callus induction, plant regeneration and chromosomal variations in barley. Acta Bot. Neerl. 39:379–387; 1990.

Hartmann, C.; Henry, Y.; De Buyser, J.; Aubry, C.; Rode, A. Identification of new mitochondrial genome organizations in wheat plants regenerated from somatic tissue cultures. Theor. Appl. Genet. 77:169–175; 1989.

Karp, A. On the current understanding of somaclonal variation. In: Miflin, B. J., ed. Oxford surveys of plant molecular cell biology, vol. 7. London: Oxford University Press; 1991:1–58.

Karp, A. Somaclonal variation as a tool for crop improvement. Euphytica 85:295–302; 1995.

Karp, A.; Maddock, S. E. Chromosome variation in wheat plants regenerated from cultured immature embryos. Theor. Appl. Genet. 67:249–255; 1984.

Larkin, P. J.; Scowcroft, W. R. Somaclonal variation—a novel source of variability from cell cultures for plant improvement. Theor. Appl. Genet. 60:197–214; 1981.

Lee, M.; Phillips, R. L. Genetic variants in progeny of regenerated maize plants. Genome 29:834–838; 1987.

Lörz, H.; Scowcroft, W. R. Variability among plants and their progeny regenerated from protoplasts of Su/su heterozygotes of *Nicotiana tabacum*. Theor. Appl. Genet. 66:67–75; 1983.

Luria, S. E.; Delbrück, M. Mutation of bacteria from virus sensitivity to virus resistance. Genetics 28:491–511; 1943.

Meins, F. Heritable variation in plant cell culture. Annu. Rev. Plant. Physiol. 34:327–346; 1983.

Morrish, F. M.; Hanna, W. W.; Vasil, I. K. The expression and perpetuation of inherent somatic variation in regenerants from embryogenic cultures of *Pennisetum glaucum* (L.) R. Br. (pearl millet). Theor. Appl. Genet. 80:409–416; 1990.

Müller, E.; Brown, P. T. H.; Hartke, S.; Lörz, H. DNA variation in tissue culture-derived rice plants. Theor. Appl. Genet. 80:673–679; 1990.

Reuveni, O.; Israeli, Y. Measure to reduce somaclonal variation in *in vitro* propagated bananas. Acta Hortic. 175:307–313; 1990.

Sandoval, J.; Tapia, A.; Müller, L.; Villalobos, V. Observaciones sobre la variabílidad encontrada en plantas micropropagadas de Musa cv. Falso Cuerno AAB. Fruits 46:533–539; 1991.

Singh, R. J. Chromosomal variation in immature embryo derived calluses of barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 72:710–716; 1986.

Sree Ramulu, K.; Dijkhuis, P.; Roest, S. Phenotypic variation and ploidy level of plants regenerated from protoplasts of tetraploid potato (*Solanum tuberosum* L. cv. 'Bintje'). Theor. Appl. Genet. 65:329–338; 1983.

Sree Ramulu, K.; Dijkhuis, P.; Roest, S.; Bokelman, G. S.; De Groot, B. Early occurrence of genetic instability in protoplasts culture of potato. Plant Sci. Lett. 36:79–86; 1984.

Swartz, H. J. Post culture behavior: genetic and epigenetic effects and related problems. In: Debergh, P. C., Zimmerman, R. H., eds. Micropropagation: technology and application. Dordrecht: Kluwer Academic Publishers; 1990:95–121.

Swedlund, B.; Vasil, I. K. Cytogenetic characterization of embryogenic callus and regenerated plants of *Pennisetum americanum* (L.) K. Schum. Theor. Appl. Genet. 69:575–581; 1985.

Vuylsteke, D.; Swennen, R.; Wilson, G.; De Langhe, E. Somaclonal variation in plantain (*Musa* spp, AAB group) derived from shoot-tip culture. Fruits 46:429–439; 1991.

Wang, X. H.; Lazzeri, P. A.; Lörz, H. Chromosomal variation in dividing protoplasts derived from cell suspensions of barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 85:181–185; 1992.

Yang, H.; Tabei, Y.; Kamada, H.; Kayano, T.; Takaiwa, F. Detection of somaclonal variation in tissue cultured rice cells using digoxigenin-based random amplified polymorphic DNA. Plant. Cell Rep. 18:520–526; 1999.