

Nombre optimal de ramets par clone dans deux tests clonaux

H Muranty, F Santi, LE Pâques, J Dufour

Station d'amélioration des arbres forestiers, Inra-Orléans, 45160 Ardon, France

(Reçu le 24 mai 1994 ; accepté le 13 avril 1995)

Résumé — Le nombre optimal de ramets par clone dans des dispositifs de tests clonaux a été étudié à l'aide de tirages aléatoires dans des données réelles de vigueur, de forme et de qualité du bois. Ces données provenaient d'un essai de 32 clones de merisier (*Prunus avium*) répliqué sur trois sites et âgé de 7 ans et d'un essai monostationnel de 96 clones de mélèze hybride (*Larix x eurolepis*) âgé de 8 ans. La précision d'estimation des valeurs génétiques et des paramètres génétiques de variance-covariance approche son maximum vers six à huit ramets par clone. Pour estimer les gains génétiques correspondant aux différentes situations pour le nombre de ramets, on a considéré que le nombre d'arbres en test était constant et que le nombre de génotypes étudiés augmentait quand le nombre de ramets par clone diminuait. Dans ce cadre, avec six ramets par clone, l'espérance de gain pour des caractères d'héritabilités comprises entre 0,20 et 0,70 est de 20 % plus importante que celle obtenue avec 18 ramets par clone, nombre de ramets par clone utilisé fréquemment dans les tests clonaux. Les résultats obtenus pour un index multilocal sont similaires à ceux obtenus pour des caractères étudiés dans un seul site.

sélection clonale / optimisation / progrès génétique / prédiction des valeurs génétiques / estimation des paramètres génétiques

Summary — **Optimal number of ramets per clone in two clonal tests.** Many forest tree breeding programmes are now based on clonal tests with the purpose of selecting improved genotypes for clonal forestry. Of course, manpower, budget and space for tests are limited and it is important to obtain the best selection efficiency within these constraints. The present paper reports a study of optimization of statistical efficiency in clonal tests. The goal is to maximize genetic gain. We examined the consequences of trade-offs between the number of ramets per clone (ie, accuracy of evaluating genotype means) and the number of clones tested (ie, selection intensity), when the total number of trees tested is held constant. The data originated from a test of 32 clones of wild cherry (*Prunus avium*) aged 7 years and repeated on three sites, and from a test of 96 clones of hybrid larch (*Larix x eurolepis*), aged 8 years and planted on only one site. Characteristics studied were height increments, girth, branching parameters, wood quality and form notations. Clonal variability was significant at a 0.1% level for all traits, and broad sense heritabilities varied between 0.2 and 0.7 (table 1). n (2 to 16) ramets per clone were randomly extracted from the whole data sets to obtain simulated data. The simulations

showed that accuracy of genotype evaluation, as measured by R^2 (or CD: average square of correlation between evaluated genotype mean and true genotype mean), increases with the number of ramets per clone and tends to a maximum which is the R^2 obtained with the whole data sets (figs 1 and 2). But with six or eight ramets, the accuracy is only 3 to 10% less than the maximal accuracy (fig 2). This results in a few changes in the ranking of clones (fig 2), but the correlation between the ranks obtained with the whole data and those obtained with each simulation remains high (table III). Variations of genetic gain with the number of ramets per clone was investigated. As an example, figure 4 shows this variation for height increment of wild cherry observed in one site: its maximum is at four ramets per clone. For many characteristics, the maximum estimated genetic gain is obtained with two ramets per clone (table III). With six to eight ramets per clone, accuracy of genotype means evaluation is sufficient and genetic gain is predicted to be 20 to 30% higher than genetic gain obtained with 18 ramets per clone, the number of ramets currently used in clonal tests. Precision of genetic variances and covariances evaluation was also investigated: it increases with the number of ramets and tends to the precision obtained with the whole data sets (fig 5). With six ramets per clone or more, this precision is still quite good.

clonal selection / optimization / genetic gain / prediction of clonal value / genetic parameters estimation

INTRODUCTION

Les tests clonaux sont utilisés dans l'amélioration d'espèces végétales, en vue d'une sortie variétale clonale. Des arbres forestiers se trouvent parmi les espèces concernées, en particulier l'eucalyptus, le peuplier et le pin radiata qui sont des espèces cultivées à grande échelle, mais également l'épicéa, le douglas, le cryptomérida, le mélèze hybride, le merisier. La sélection par comparaison de clones en vue d'une sortie variétale clonale apparaît comme une des meilleures méthodes de sélection. En effet, elle permet d'utiliser toute la variabilité génétique, additive et non-additive, contrairement à la sélection sur l'aptitude générale à la combinaison (AGC) suivie de propagation par graines des arbres sélectionnés dans des vergers à graines (Libby et Rauter, 1984). De plus, le clonage permet de réaliser des copies d'un génotype pour le tester sur plusieurs sites. Il peut aussi être utilisé dans des tests de descendance pour obtenir des estimations plus fines des paramètres génétiques.

Les moyens humains et financiers, de même que l'espace disponible pour les tests, sont limités. Il faut donc rechercher la méthode de sélection qui réalise le meilleur compromis entre l'efficacité de

sélection et les contraintes de moyens. Dans ce cadre, un des paramètres sur lequel on peut agir est le nombre de ramets par clone, c'est-à-dire le nombre de copies clonales de chaque génotype utilisé dans les tests. Intuitivement, plus ce nombre est grand, plus l'évaluation des génotypes est précise. Les surfaces disponibles et les densités de plantation fixent le nombre d'arbres observables et donc, si le nombre de ramets par clone augmente, le nombre de clones évalués diminue. En général, le nombre de clones qui doivent être sélectionnés est fixé a priori. Il est assez important pour que la variabilité génétique restant dans le sous-ensemble de clones sélectionnés et qui seront plantés à grande échelle assure une bonne stabilité de la variété dans l'espace et dans le temps. Par conséquent, la diminution du nombre de génotypes testés qui est consécutive à l'augmentation du nombre de ramets par clone entraîne une diminution de l'intensité de sélection. Un compromis doit être recherché entre la précision d'estimation des génotypes et l'intensité de sélection pour obtenir un gain maximum par unité de temps et de moyens.

Quelques auteurs ont étudié l'optimisation du nombre de ramets par clone dans les tests clonaux d'arbres forestiers. Ils se sont basés surtout sur les formules de pré-

diction de gain génétique (Lindgren, 1985 ; Shaw et Hood, 1985 ; Russel et Libby, 1986) ou sur des données simulées par ordinateur et suivant bien les hypothèses de l'analyse de variance (Russel et Loo-Dinkins, 1993). Russel et Libby (1986) montrent que le nombre optimal de ramets par clone qui permet d'obtenir le meilleur gain génétique pour une variété clonale est inférieur ou égal à 6 quand l'héritabilité est comprise entre 0,1 et 0,5 et la pression de sélection assez faible (2 à 5 % pour un ramet). Shaw et Hood (1985) montrent que le clonage dans des tests de descendance permet d'augmenter le gain génétique sur la valeur additive et que le nombre optimal de ramets par clone dans ce cas est compris entre (1) 2 et 6. Russel et Loo-Dinkins (1993) montrent qu'en présence d'interaction génotype-milieu, pour un nombre total d'arbres évalués constant, tous sites confondus, le nombre optimal de ramets par clone et par site est de 1 ou 2, et le nombre optimal de sites entre 4 et 6. Lindgren (1985) propose une formule pour évaluer les nombres de sites et de ramets par clone et par site optimaux en fonction de l'importance de l'interaction génotype-milieu et d'un paramètre économique évaluant le coût d'un site par rapport au coût d'un ramet.

Pour obtenir des grumes de merisier de bonne qualité, le sylviculteur doit tailler et élaguer les arbres dès le plus jeune âge. Ceci représente un coût de suivi par arbre important qui est compensé par une plantation à grand espacement. En conséquence peu d'éclaircies sont pratiquées entre la plantation et la récolte. L'amélioration génétique doit donc fournir un matériel de grande qualité et homogène, ce qui justifie la voie clonale pour le merisier. Les graines de mélèze hybride sont difficiles à obtenir, aussi les meilleurs individus seront propagés par voie clonale. De plus, le mélèze hybride peut être planté à grand espacement en reboisement de terres agricoles. Des tests clonaux de mélèze hybride

sont donc réalisés. Pour utiliser au mieux les moyens humains et financiers dont disposent les programmes d'amélioration du merisier et du mélèze de l'Inra, on doit donc rechercher le nombre de ramets par clone qui permette d'obtenir le plus grand gain génétique par unité de temps et de moyens. Nous avons utilisé des données réelles de ces deux espèces pour simuler l'effet de la diminution du nombre de ramets sur la précision d'analyse d'un test clonal. Nous nous sommes intéressés à la variation de la précision d'estimation des valeurs génétiques des clones, la variation de leur classement, et celle de l'espérance de gain génétique en fonction du nombre de ramets. La précision d'estimation des paramètres génétiques des populations dont les clones sont issus a aussi été observée.

MATÉRIEL ET MÉTHODES

Les données utilisées pour cette étude provenaient d'une part d'un test clonal monostationnel de mélèze hybride et d'autre part d'un test clonal multilocal de merisier.

Le dispositif expérimental de mélèze hybride (*Larix eurolepis* = *L. decidua* x *L. kaempferi*) a été planté en décembre 1984 sur un seul site : Éclache (Puy-de-Dôme), à environ 1 000 m d'altitude et recevant des précipitations abondantes. Au total, 96 clones y ont été testés. Les plants ont été produits par bouturage au printemps 1982. Le nombre de ramets par clone varie de deux à 18 (en moyenne de 12 à 14 ramets par clone). Les arbres ont été plantés à l'écartement de 3 m x 3 m. Ils ont été disposés en blocs incomplets à composition aléatoire (Pâques, 1992).

Les mesures obtenues pour chaque arbre concernent les hauteurs totales en fin de saison de végétation de 1986 à 1992, soit de 2 à 8 ans après la plantation (en cm), de la circonférence à 1,30 m en 1992 (en mm), et deux mesures de la dureté du bois estimée par le pilodyn, réalisées en 1992 sur écorce dans deux orientations perpendiculaires (en 10^e de mm). Une note de flexuosité du tronc a été attribuée en 1992, selon une échelle de 1 (mauvais) à 5 (bon). À partir des mesures de hauteur, un accroissement total de 1986 à 1992 a été calculé.

Le dispositif de comparaison de clones de merisier (*Prunus avium* L.) a été planté pendant l'hiver 1984-1985 avec des plants issus de culture in vitro. Il comprend 32 clones et est répété sur trois sites : Neufchatel-en-Bray (Seine-Maritime), Sarrazac (Dordogne) et Saveuse (Somme). Les sites de Neufchatel et de Saveuse sont des sites forestiers moyennement favorables au merisier, avec une croissance moyenne annuelle de 60 et 48 cm par an respectivement. À Sarrazac, l'essai est implanté sur un ancien terrain agricole, et les conditions y sont très favorables au merisier, ce qui se traduit par une croissance de 110 cm par an. Le nombre de ramets par clone et par site varie de six à 18. Les arbres ont été plantés avec un écartement de 7 m x 3 m. Ils ont été disposés en blocs incomplets à composition aléatoire.

Pour chaque arbre, les hauteurs totales à la plantation et en fin de saison de végétation (en cm) de 1985 à 1991, c'est-à-dire de 1 à 7 ans après la plantation, et la circonférence à 1,30 m en 1991 (en mm) ont été mesurées. Le nombre de branches apparues en 1986 (2^e année) sur la pousse de 1985, les angles et diamètres des trois branches terminales sur la pousse de 1985 ont été observés à Neufchatel et Sarrazac. La vigueur a été évaluée par l'accroissement en hauteur de la 2^e à la 7^e année, qui correspond à une croissance normale après la reprise, hors crise de transplantation et avant que les arbres entrent en compétition entre eux.

Les données ont été ajustées des effets de terrain par la méthode de Papadakis itérée (Papadakis, 1937 ; Bartlett, 1978 ; Pichot, 1993). Cette méthode est une analyse spatiale des effets de terrain, qui utilise les performances des proches voisins d'un plant pour estimer les effets du terrain appliqués à ce plant. Les données ajustées de chaque site ont été analysées avec le modèle $X_{ij} = \mu + G_i + \varepsilon_{ij}$ où X_{ij} est une mesure du caractère considéré sur un ramet du clone i ; μ , la moyenne générale ; G_i , l'effet du clone i ; ε_{ij} , l'erreur résiduelle du modèle. Les composantes de la variance ont été estimées et l'héritabilité au sens large des caractères a été calculée en utilisant la formule :

$$h_{sl}^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_e^2}$$

où σ_G^2 est la variance clonale ; et σ_e^2 , la variance résiduelle.

Des analyses multistationnelles ont été effectuées avec les données concernant le merisier. D'une part, une analyse de variance classique a

été réalisée avec le modèle $X_{ijk} = \mu + G_i + E_j + (G \times E)_{ij} + \varepsilon_{ijk}$ où X_{ijk} est une mesure du caractère considéré sur un ramet du clone i dans le site j ; μ , la moyenne générale ; G_i , l'effet du clone i ; E_j , l'effet du site j ; $(G \times E)_{ij}$, le terme d'interaction pour le clone i et le site j ; et ε_{ijk} , l'erreur résiduelle du modèle. Les composantes de la variance ont été calculées ainsi que l'héritabilité au sens large, estimée selon la formule

$$h_{sl}^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_{G \times E}^2 + \sigma_e^2}$$

où σ_G^2 est la variance clonale ; $\sigma_{G \times E}^2$, la variance d'interaction clone x site ; et σ_e^2 , la variance résiduelle. Les rapports $\sigma_{G \times E}^2 / \sigma_G^2$ et $\sigma_e^2 / \sigma_{G \times E}^2$ ont également été calculés. Ils constituent 2 façons de mesurer l'importance de l'interaction génotype-milieu. Le rapport $\sigma_{G \times E}^2 / \sigma_G^2$ est un des paramètres dont l'influence sur le nombre optimal de ramets par clone a été étudié par Russel et Loo-Dinkins (1993). Le rapport $\sigma_e^2 / \sigma_{G \times E}^2$ est utilisé par Lindgren (1985) pour trouver le nombre de sites et le nombre de ramets par clone optimaux en fonction des contraintes économiques et biologiques.

D'autre part, un index multilocal a été estimé pour chaque clone pour le caractère de croissance après reprise en attribuant un poids arbitraire de un aux mesures de Neufchatel et Saveuse (sites forestiers) et un poids arbitraire de deux à celles de Sarrazac (site agricole). On peut ainsi valoriser les corrélations génétiques entre sites pour ce caractère. Les poids choisis peuvent se justifier car la sélection du merisier vise à égalité les sites forestiers et les terres agricoles.

Les analyses ont été réalisées avec les logiciels Splus (Statistical Sciences, Inc, 1988), pour les statistiques de base et les graphes, et Sélect (Mangin, 1992), pour interpréter les données en modèle à effets aléatoires. Les analyses site par site ont été réalisées simultanément pour tous les caractères étudiés. En revanche l'analyse multilocale a concerné le caractère de croissance seul, car l'analyse de tous les caractères dans tous les sites simultanément n'était pas possible. Les Blup (*best linear unbiased predictor*) des valeurs génétiques et d'index ont été calculés, ainsi que les coefficients de détermination (CD) des estimations des valeurs génétiques qui sont les carrés des corrélations entre les valeurs génétiques vraies et leurs estimations. La moyenne de ces coefficients de détermination (c'est-à-dire le CD moyen) constitue une mesure globale de la précision d'évaluation des génotypes. Elle peut

être utilisée pour prédire l'espérance de gain génétique selon la formule $\Delta G = i\sigma_G\sqrt{CD}_{moyen}$ (Gallais, 1991).

À partir des fichiers comportant l'ensemble des données ajustées des effets de terrain par la méthode de Papadakis, des fichiers "simulés" où les clones ont n ramets en essai ont été créés : les n ramets ont été choisis par tirage aléatoire. Les clones d'effectifs insuffisants (c'est-à-dire ayant moins de n ramets) étaient représentés par tous leurs ramets. Pour chaque nombre de ramets considéré, dix (analyse sur plusieurs sites) ou 20 (un seul site) fichiers ont été analysés.

RÉSULTATS PRÉLIMINAIRES

Pour tous les caractères étudiés, chez le mélèze et chez le merisier, l'analyse de variance montre que l'effet clone est significatif au seuil de 0,001. Les valeurs de l'héritabilité au sens large quand l'estimation est réalisée en considérant un seul site sont données dans le tableau I. Selon le caractère considéré et selon le site dans lequel elle est évaluée, l'héritabilité au sens large peut varier fortement, de 0,2 à 0,7 ici. Le tableau II donne les valeurs des compo-

santes de la variance pour l'analyse multilocal des données de merisier. Ces composantes représentent une faible part de la variation totale car les effets site sont très importants. Pour tous les caractères, les valeurs des rapports $\sigma_{GxE}^2 / \sigma_G^2$ et $\sigma_e^2 / \sigma_{GxE}^2$ nous montrent que la variance due à l'interaction site-clone représente la plus petite part de la variance totale : le rapport $\sigma_{GxE}^2 / \sigma_G^2$ est toujours inférieur à 1 (de 0,248 à 0,884) et le rapport $\sigma_e^2 / \sigma_{GxE}^2$ est toujours compris entre 1 et 10 (de 2,5 à 9,3). L'héritabilité des caractères est cependant plus faible dans l'analyse multilocale que dans les analyses site par site, à cause de l'interaction site-clone.

RÉSULTATS DES SIMULATIONS

Précision d'estimation des valeurs génétiques

Le coefficient de détermination moyen (CD) permet d'observer comment varie la préci-

Tableau I. Héritabilités au sens large des caractères étudiés, analyse site par site.

	<i>Merisier</i>			
	<i>Croissance</i>	<i>Circonférence</i>	<i>Nombre de branches</i>	<i>Angle de branchaison</i>
Neufchatel	0,32 (0,064)	0,38 (0,069)	0,20 (0,054)	0,45 (0,070)
Sarrazac	0,56 (0,065)	0,70 (0,054)	0,66 (0,058)	0,57 (0,065)
Saveuse	0,25 (0,059)	0,33 (0,065)		
	<i>Mélèze</i>			
	<i>Accroissement 86-92</i>	<i>Circonférence</i>	<i>Pilodyn moyen</i>	<i>Flexuosité</i>
Éclache	0,27 (0,037)	0,26 (0,037)	0,56 (0,039)	0,51 (0,041)

Entre parenthèses, écart type d'estimation de l'héritabilité.

Tableau II. Composantes de la variance pour l'analyse multilocale des caractères observés chez le merisier.

Caractères	Variance clonale	Variance d'interaction	Héritabilité	$\frac{\sigma_{G \times E}^2}{\sigma_G^2}$	$\frac{\sigma_e^2}{\sigma_{G \times E}^2}$
	σ_G^2 (%)	$\sigma_{G \times E}^2$ (%)			
Croissance	746 (2,4)	266 (0,86)	0,21	0,356	9,3
Circonférence	896 (6,5)	435 (3,2)	0,31	0,485	3,6
Nombre de branches	2,06 (10,4)	1,82 (9,2)	0,24	0,884	2,5
Angle de branchaison	23,4 (38,5)	5,80 (9,5)	0,41	0,248	4,7

Entre parenthèses, pourcentage de la variance totale représentée par les variances clonales et d'interaction ; ^a l'écart type de l'héritabilité n'a pu être calculé.

sion d'estimation des valeurs génétiques avec le nombre de ramets. En figure 1, on constate que le coefficient de détermination moyen augmente avec le nombre de ramets et tend vers un CD moyen maximum qui correspond à celui de l'analyse utilisant toutes les données : ce phénomène est observé par tous les caractères étudiés (fig 2). En revanche, pour les faibles nombres de ramets, le coefficient de détermination est d'autant plus faible que l'héritabilité est faible (fig 2). De plus, pour un même nombre de ramets, la variation, selon les simulations, du CD moyen est de plus en plus faible quand le nombre de ramets augmente (fig 1). En général, pour six ou huit ramets, la précision est inférieure de 3 à 10 % à la précision maximale seulement.

Conséquences sur les classements

Cette diminution de précision d'estimation des valeurs génétiques se traduit par de faibles changements de classement ; la corrélation du classement des clones obtenu après simulation avec le classement de référence obtenu avec le fichier complet (corrélation de Pearson sur les rangs) augmente avec le nombre de ramets mais elle reste généralement élevée. Ainsi, pour un nombre

de ramets supérieur ou égal à six, elle descend rarement en-dessous de 0,8 et elle atteint 0,9 pour un nombre de ramets assez faible, variant de trois à sept selon les sites et les caractères (tableau III).

Les changements de classement qui ont des conséquences graves sont ceux qui concernent les clones qui vont être sélectionnés. La figure 3 donne un exemple de comparaison graphique des rangs des neuf premiers clones selon l'analyse des fichiers simulés avec le rang qu'ils ont selon l'analyse du fichier complet. Pour l'ensemble des caractères étudiés, ces clones se trouvent en général au début du classement. Pour les caractères les plus héréditaires ($h^2 \geq 0,5$), on observe que les neuf premiers clones sont toujours dans la première moitié, voire le premier quart du classement de référence. En revanche, pour les caractères moins héréditaires (par exemple la croissance à Saveuse, $h^2 = 0,25$), ce n'est qu'à partir de huit à dix ramets qu'ils sont présents dans la première moitié du classement de référence. Cependant, si l'on diminue le nombre de ramets, on va pouvoir augmenter le nombre de clones étudiés. On doit donc s'intéresser surtout aux quelques premiers clones, qui en moyenne seraient sélectionnés, alors que les suivants ont toutes chances d'être moins bons que certains des génotypes étudiés en

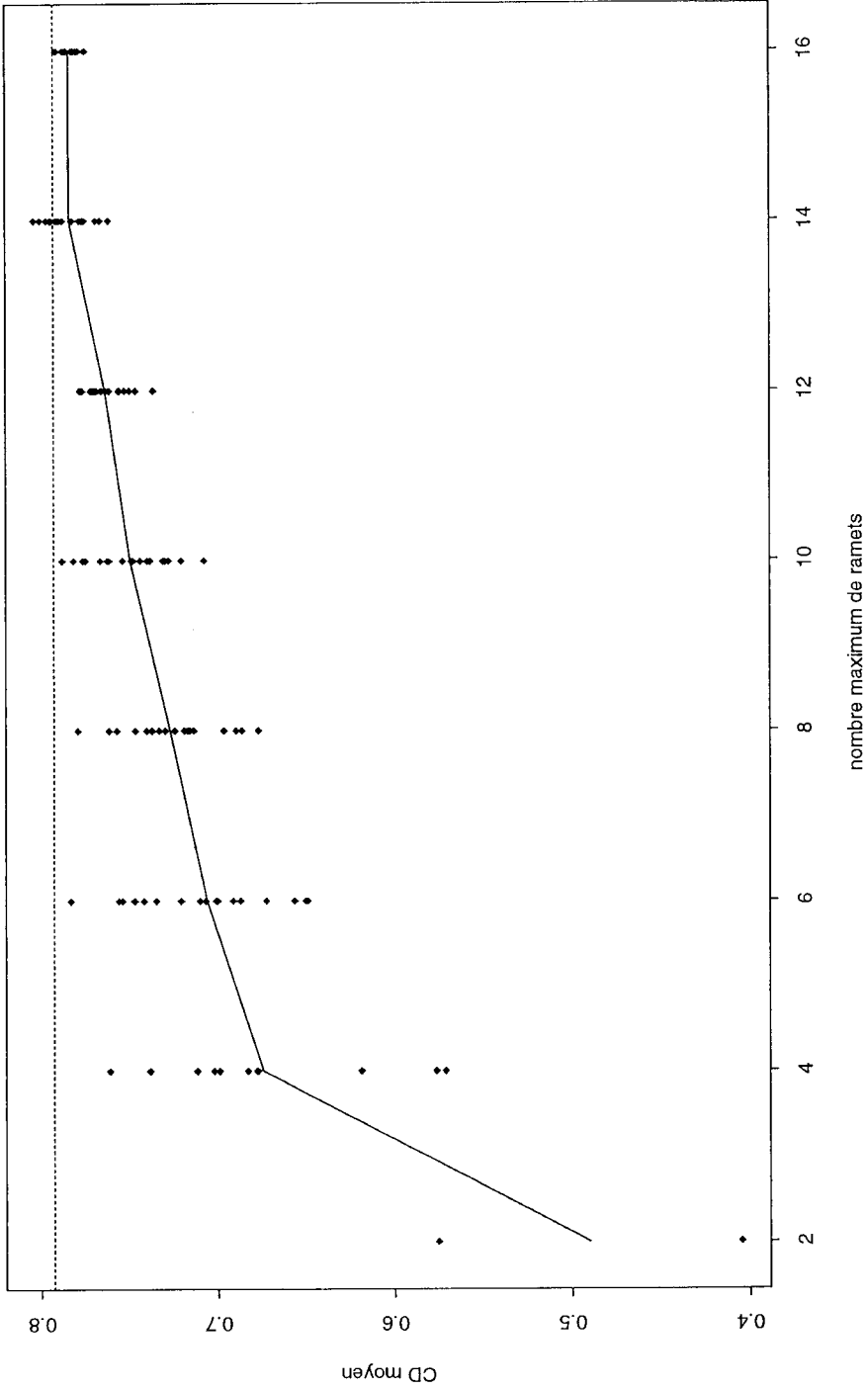


Fig 1. Évolution des coefficients de détermination moyens en fonction du nombre de ramets par clone : exemple de l'accroissement en hauteur du merisier à Saveuse. Chaque point correspond à une simulation, la ligne pleine à la moyenne des simulations pour chaque nombre de ramets, la ligne en pointillé au CD moyen de l'analyse du fichier complet.

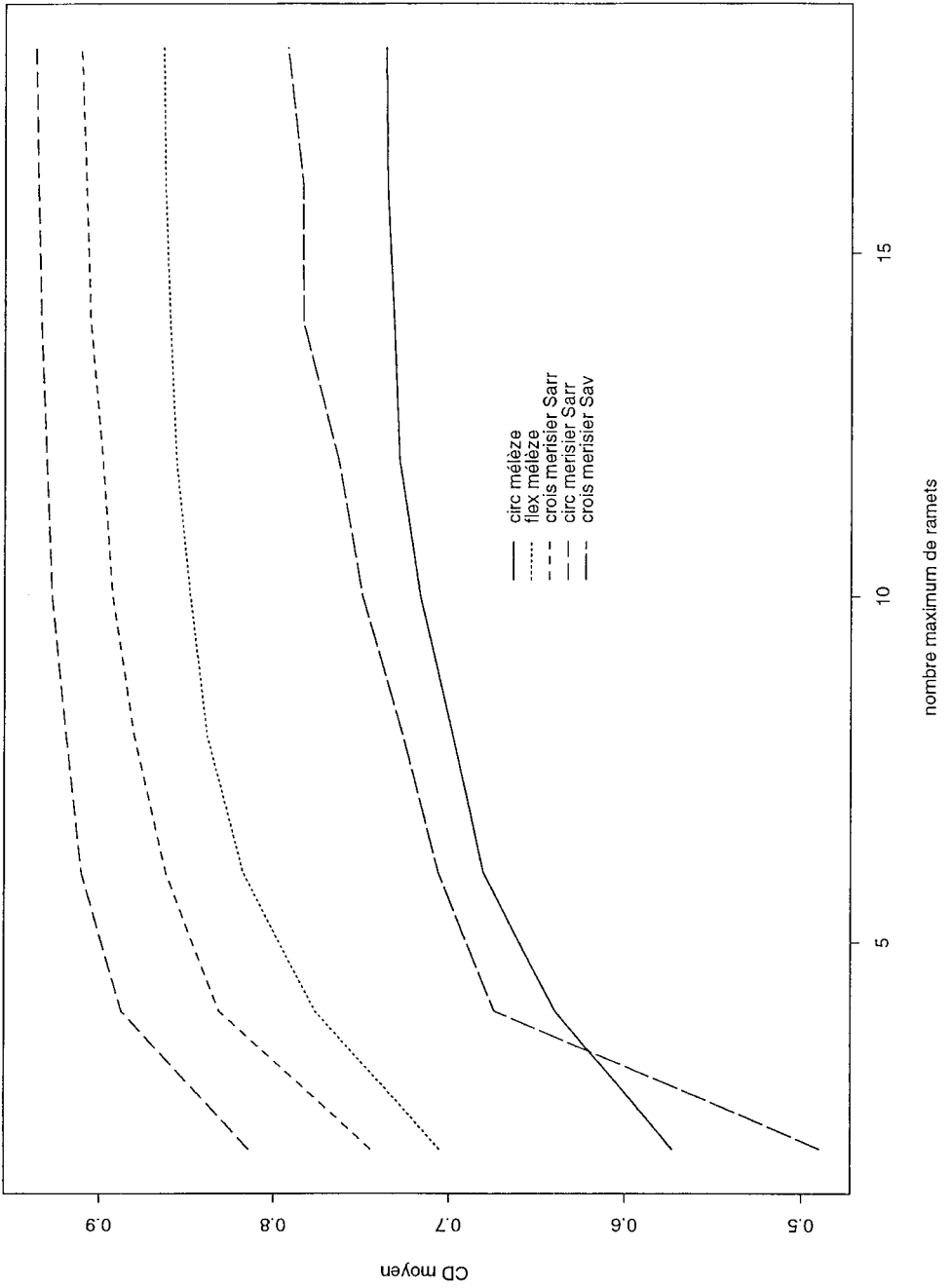


Fig 2. Évolution des coefficients de détermination moyens en fonction du nombre de ramets par clone pour quelques caractères, moyenne sur les simulations. ——— circ méléze, flex méléze, - - - - - crois merisier Sarr, - · - · - circ merisier Sarr, - - - - - crois merisier Sav.

Tableau III. Résultats des simulations.

Espèce Sites Caractères	Corrélation des rangs (r^2)		Nombre de ramets donnant un gain maximum	ΔG_6 <hr/> ΔG_{16}
	Minimum	Nombre de ramets tel que $r^2 \geq 0,9$		
Merisier				
Saveuse				
Croissance	0,65	7	4	1,25 (8,6 10 ⁻³)
Circonférence	0,74	6	2	1,23 (9,5 10 ⁻³)
Sarrazac				
Croissance	0,86	4	2	1,28 (2,1 10 ⁻³)
Circonférence	0,89	4	2	1,29 (9,1 10 ⁻⁴)
Nb de branches	0,89	4	2	1,29 (2,1 10 ⁻³)
Angle de branchaison	0,74	6	2	1,27 (4,1 10 ⁻³)
Neufchatel				
Sarrazac				
Saveuse				
Index pour la croissance	0,85	3	2	/ (*)
Mélèze				
Éclache				
Circonférence	0,77	6	2	1,26 (4,8 10 ⁻³)
Pilodyn moyen	0,89	3	2	1,29 (1,2 10 ⁻³)
Flexuosité	0,85	3	2	1,28 (1,4 10 ⁻³)

* Pas de simulations pour 16 ramets. Entre parenthèses, écart type.

plus, et ne seraient donc pas sélectionnés. Ainsi, si le nombre de ramets passe de 18 à six, et si huit clones doivent être sélectionnés, «8/3» clones à sélectionner sont dans l'échantillon étudié avec 18 ramets. On observe que les deux ou trois premiers clones, selon les analyses des fichiers simulés, sont proches du début du classement de référence, c'est-à-dire dans les dix premiers, au moins quand le nombre de ramets par clone est supérieur ou égal à six.

Gains génétiques prédits

Les coefficients de détermination moyens calculés après analyse des fichiers simu-

lés ont été utilisés pour calculer les espérances de gain génétique en faisant varier l'intensité de sélection i avec le nombre de ramets sous l'hypothèse que le nombre total d'arbres observés est constant, et en utilisant l'estimation la plus précise de la variance génétique, celle obtenue avec tous les ramets. À titre d'exemple, la figure 4 montre que, pour l'accroissement en hauteur observé à Saveuse, le gain génétique prédit maximum est obtenu pour quatre ramets. Pour certains caractères plus héréditaires, et pour l'index multilocal de croissance du merisier, le gain prédit est maximum pour deux ramets par site (tableau III). Cela signifie que le gain prédit est maximal pour un nombre de ramets par clone pour lequel la précision n'avait pas été jugée suffisante,

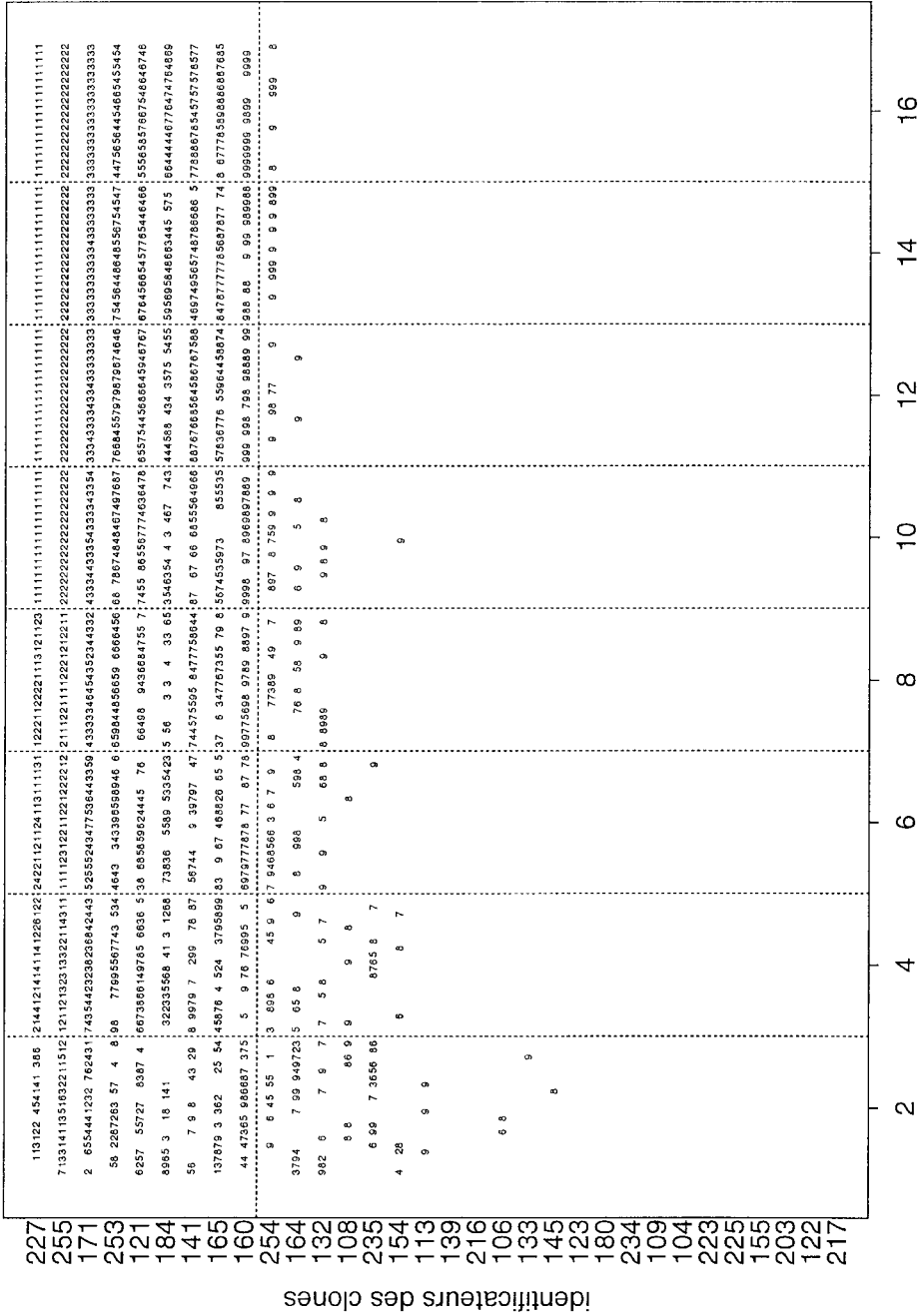


Fig 3. Comparaison graphique des rangs des premiers clones dans les simulations avec leur rang de référence : exemple de l'accroissement en hauteur du merisier à Sarrazac. En ordonnée, les identificateurs sont placés selon les rangs des clones obtenus dans l'analyse du fichier complet, le premier étant en haut. Les chiffres indiquent les rangs des neuf premiers clones de chaque simulation.

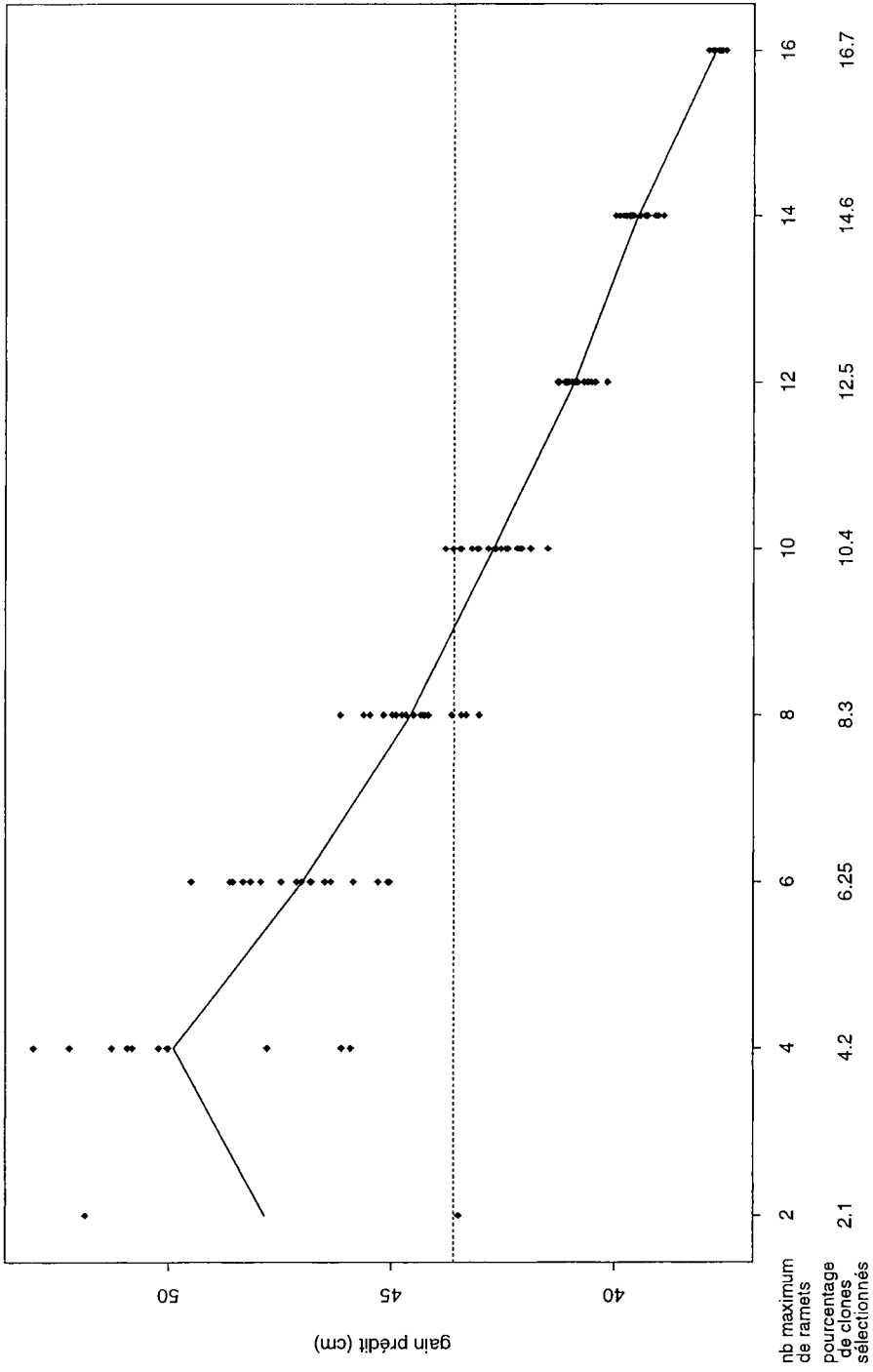


Fig 4. Évolution des gains génétiques en fonction du nombre de ramets par clone testés : exemple de l'accroissement en hauteur du merisier à Saveuse. Chaque point correspond à une simulation, la ligne pleine à la moyenne des simulations pour chaque nombre de ramets, la ligne en pointillé à 1,2 fois le gain prédit pour le test complet.

sur la base des CD moyens ou de la coïncidence des classements des premiers clones. Le rapport des espérances de gain génétique à six et 16 ramets est donné au tableau III : il est toujours compris entre 1,2 et 1,3. La précision d'estimation de ce rapport a été calculée à partir des variances sur les simulations des coefficients de détermination moyens : l'écart type est toujours inférieur à 0,01. De même, avec une diminution du nombre de ramets par clone de 18, nombre de ramets par clone utilisé jusqu'à présent, à six ou huit, nombre pour lequel la précision d'estimation des valeurs génétiques est suffisante, l'espérance de gain génétique serait augmentée de 20 à 30 %.

Précision d'estimation des paramètres de variances et de covariances génétiques

Les tests clonaux peuvent servir à évaluer les paramètres de variances et de covariances génétiques de la population dont sont issus les clones. Le logiciel Sélect donne une estimation de ces paramètres ainsi que de leur variance d'estimation, qui est une mesure de la précision d'estimation de ces paramètres. La figure 5 montre que ces variances diminuent rapidement quand le nombre de ramets augmente, c'est-à-dire que l'estimation des paramètres génétiques devient plus précise. De plus elles tendent vers un minimum, qui correspond à la variance d'estimation obtenue avec le fichier complet. En général pour un nombre de ramets supérieur ou égal à six, le plus souvent pour un nombre de ramets supérieur ou égal à quatre seulement, les variances d'estimation des variances et des covariances sont comprises entre ce minimum et deux fois ce minimum.

DISCUSSION

Comme indiqué précédemment, les données utilisées ont été ajustées des effets

de terrain avant les simulations et ces effets de terrain ont été estimés à partir du nombre total de ramets en test. Cela pose un problème car l'efficacité des méthodes d'analyse des effets de terrain peut dépendre du nombre de répétitions par génotype pour un nombre total d'arbres (ou une surface de test) constant(e). Ici, les fichiers simulés ont été obtenus en choisissant aléatoirement un nombre donné de ramets pour chaque génotype, sans tenir compte de leur disposition sur le terrain. Pour utiliser des données non ajustées, il aurait fallu considérer les arbres non retenus comme manquants, ce qui rend l'estimation des effets de terrain moins performante, voire impossible. Une simulation complète des données par ordinateur peut permettre de tester l'efficacité des méthodes d'estimation des effets de terrain pour un nombre de ramets par clone variable. Pichot (1993) donne un exemple de ce type de simulation et montre que le méthode de Papadakis itérée reste performante même avec seulement deux, quatre ou six ramets par clone.

Pour réaliser des simulations à «moyens constants», c'est-à-dire avec un nombre total d'arbres constant, il faudrait faire varier simultanément le nombre de clones (c) et le nombre de ramets par clone (r) de façon à garder $c \times r$ constant. Mais pour cela, il aurait fallu disposer d'un nombre initial de clones plus important, pour ne pas avoir un nombre de clones trop faible (< 20 , modèle aléatoire inapplicable) quand le nombre de ramets est grand. Or pour le nombre de clones de merisier étudiés (32), pour comparer six ramets à 18 ramets par clone, il faut considérer dix clones seulement à 18 ramets pour 32 clones à six ramets, ce qui est vraiment faible. Ce type de simulation a déjà été réalisé par Russel et Libby (1986) pour un test sur un seul site. De plus, si le nombre de clones est suffisamment important pour que l'échantillon de clones soit représentatif de la population, il n'a pas d'influence sur la précision d'estimation des

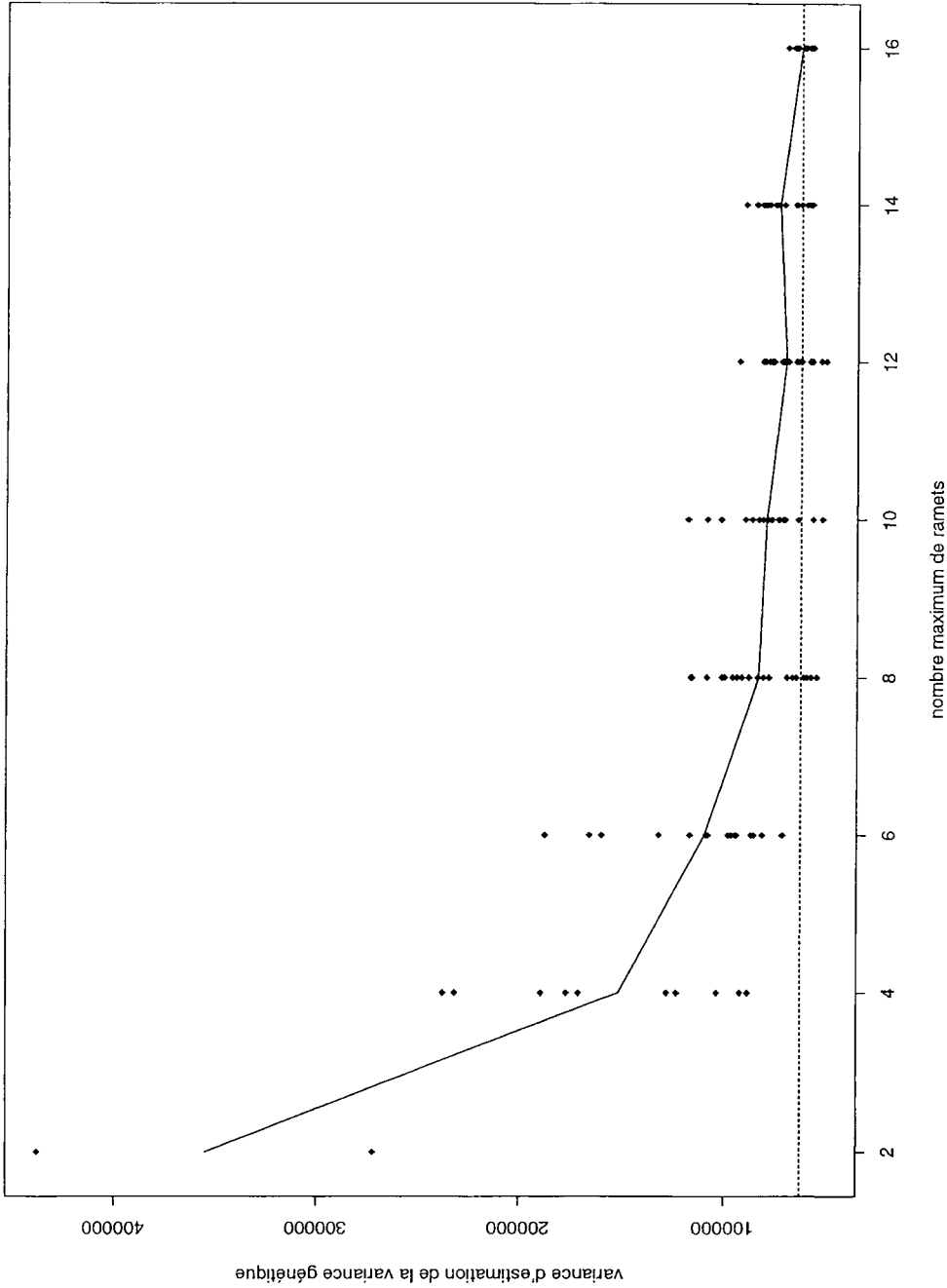


Fig 5. Évolution de la précision d'estimation de la variance génétique : exemple de l'accroissement en hauteur du merisier à Saveuse. Chaque point correspond à une simulation, la ligne pleine à la moyenne des simulations pour chaque nombre de ramets, la ligne en pointillé à la précision obtenue dans l'analyse du fichier complet.

valeurs génétiques et des paramètres génétiques.

L'héritabilité au sens large d'un caractère est un des paramètres qui détermine le nombre de répétitions par clone à planter en test pour obtenir le plus grand gain génétique (Russel et Libby, 1986). Celles des caractères étudiés dans cette étude, situées entre 0,20 et 0,70, se trouvent dans un intervalle suffisamment large pour recouvrir la gamme des héritabilités observées chez d'autres espèces pour de nombreux caractères. Les analyses des fichiers simulés montrent qu'avec de telles héritabilités les valeurs génétiques sont estimées avec une précision proche du maximum pour six à huit ramets. De même la variance d'estimation des paramètres de variances-covariances génétiques atteint son minimum vers six à huit ramets. Or beaucoup des tests clonaux ont été mis en place avec plus de six ramets par clone (Russel et Libby, 1986) ; en général 15 à 20 ramets par clone sont utilisés. Notre étude montre que planter un si grand nombre de ramets par clone est un luxe qui ne permet pas de gagner beaucoup sur la précision des estimations. Comme il a été dit en introduction, pour un nombre total d'arbres testés constant et un nombre de clones sélectionnés constant, la diminution du nombre de ramets par clone permet une augmentation du nombre de clones testés et par conséquent une augmentation de l'intensité de sélection, obtenue par une pression de sélection plus forte. La comparaison des figures 1 et 4 montre que cette augmentation de l'intensité de sélection compense bien la perte de précision de l'estimation des valeurs génétiques. Ainsi, en n'utilisant que six ramets par clone, on peut espérer des gains génétiques de 20 % supérieurs à ceux obtenus avec 18 ramets par clone. Donc, pour obtenir le plus grand gain génétique, le nombre de ramets par clone doit être faible. Ces résultats sont en accord avec ceux, théoriques, de Russel et Libby (1986), qui montrent que, pour des

héritabilités au sens large comprises entre 0,1 et 0,5, et pour des pressions de sélection faibles (de 2 à 5 % pour un ramet), le nombre optimal de ramets par clone est compris entre (1) 2 et 6.

Le problème de l'interaction génotype–milieu

L'index multilocal de croissance du merisier permet de constater que les résultats obtenus par les simulations pour un caractère évalué dans un seul site sont aussi valables pour un index multilocal ou multivariable. Cette bonne concordance est sans doute due au fait que peu de clones présentent de grand changement de performance relative d'un site à l'autre (Muranty et al, en prép). L'interaction génotype–milieu représente d'ailleurs la plus petite part de la variabilité génotypique. En effet, le rapport $\sigma_{G \times E}^2 / \sigma_G^2$ est toujours inférieur à 1, ce qui correspond à peu près à la gamme étudiée par Russel et Loo-Dinkins (1993) pour ce rapport, et le rapport $\sigma_e^2 / \sigma_{G \times E}^2$ est toujours compris entre 1 et 10 alors que Lindgren (1985) montre que ce rapport est généralement compris entre 10 et 100 pour des tests des descendances. Mais, à cause de la variabilité génétique à l'intérieur des familles, dans ces tests, la variance résiduelle est attendue plus forte et l'interaction génotype–milieu plus faible par rapport aux tests clonaux ; ceci peut expliquer le décalage observé sur le rapport $\sigma_e^2 / \sigma_{G \times E}^2$. L'importance de l'interaction génotype–milieu a une incidence sur la stratégie d'amélioration. En effet quand elle est très forte, il peut être intéressant de définir des zones plus homogènes et de sélectionner des clones pour chaque zone si des clones spécialisés sont identifiables. En revanche si elle est assez faible, ce qui est observé ici pour le merisier, il peut être suffisant de réaliser des tests multistationnels pour identifier des clones performants et stables. Cette étude

ne permet pas de décider du nombre optimal de sites pour cette dernière stratégie mais fournit un premier élément de réponse concernant la précision d'évaluation des valeurs génétiques pour le merisier : la comparaison du CD moyen de chaque site avec 18 ramets pour la croissance au CD moyen de l'index multistationnel avec six ramets permet d'envisager le cas où le nombre total de ramets par clone est constant (tableau III). Le CD moyen d'un site est plus important quand le site considéré est très homogène, c'est-à-dire quand l'héritabilité dans ce site est élevée (Sarrazac) et moins important quand l'héritabilité est plus faible (Saveuse, Neufchâtel). Intuitivement, le CD moyen de l'index doit être plus faible que les CD moyens dans chaque site car l'interaction génotype-milieu perturbe la prédiction des valeurs génotypiques. On constate que l'interaction est assez importante pour avoir cet effet vis-à-vis du site où l'héritabilité est forte seulement. Cependant les CD moyens de chaque site ont été obtenus dans une analyse multicaractère alors que l'analyse donnant l'index n'a considéré que le caractère de croissance dans chaque site : si le caractère croissance avait été analysé seul pour chaque site, les CD moyens auraient probablement été plus faibles que ce qu'ils sont dans l'analyse multicaractère, car les caractères étudiés sont corrélés (Muranty et al, en prép). La dispersion de tous les ramets des clones sur plusieurs sites ne diminue donc presque pas la précision (d'évaluation de leurs valeurs génétiques).

Comment décider du nombre optimal de sites ? Pour l'approche classique d'analyse de variance avec interaction, Lindgren (1985) propose une méthode pour trouver les nombres optimaux de sites et de ramets par site en fonction de l'importance de l'interaction génotype-milieu et de paramètres économiques. Mais les données de plusieurs sites sont mieux valorisées quand elles sont utilisées pour évaluer un index

multilocal (Gallais, 1991). Dans ce cas, ce n'est pas tant le nombre de sites qui est important que leur représentativité par rapport à l'ensemble des sites dans lesquels le reboisement se fera avec l'espèce étudiée. Pour en avoir une idée, il faut étudier de nombreux sites ainsi que la structuration de l'interaction génotype-milieu, et éventuellement trouver des variables du milieu expliquant cette structuration. Le nombre «optimal» de sites est alors celui qui permet de représenter au mieux toute la gamme de milieux possible avec les moyens dont on dispose.

Aspects pratiques

Quelques remarques que l'améliorateur devra garder en mémoire au moment de mettre en œuvre un test clonal peuvent être ajoutées aux résultats de cette étude :

- avec un nombre trop faible de ramets, l'évaluation des effets de terrain peut devenir peu performante, si elle utilise un modèle avec des blocs complets. En effet, dans ce cas, l'augmentation du nombre de génotypes qui correspond à la diminution du nombre de ramets par génotype entraîne forcément l'augmentation de la taille des blocs, qui auront ainsi toutes les chances de ne pas être homogènes ;

- il faut également penser au fait que l'héritabilité d'un caractère peut varier avec l'âge de la plantation. Dans cette étude, chez le mélèze, l'héritabilité des mesures de hauteur augmente lentement avec l'âge et passe de 0,174 à 5 ans à 0,233 à 11 ans (données non publiées). Elle reste toujours proche de l'héritabilité de l'accroissement en hauteur qui a été étudié et les conclusions sont donc inchangées. Mais, si l'héritabilité d'un caractère diminuait avec l'âge, il faudrait avoir prévu un nombre de ramets plus important que ce qui est l'optimum au stade jeune pour être encore proche de l'optimum à un stade plus âgé. Dans des tests assez âgés,

en général l'héritabilité augmente jusqu'à un certain âge, puis diminue à cause de la compétition (Franklin, 1979 ; Balocchi et al, 1993 ; Magnussen, 1993). De plus les éclaircies pratiquées peuvent aussi entraîner des diminutions de l'héritabilité des caractères observés. Cependant, il semble peu probable qu'un caractère utilisable en sélection sorte de la large gamme d'héritabilités envisagée dans cette étude (0,2 à 0,7).

CONCLUSION

Les résultats de cette étude permettent d'évaluer la précision d'estimation des valeurs génétiques des clones et les espérances de grains selon le nombre de ramets par clone et de choisir ce nombre pour des tests clonaux selon l'importance accordée à chacun de ces paramètres. Par exemple dans le cadre du travail d'amélioration du merisier, conduit en France par l'Inra, il est nécessaire que la valeur des clones sélectionnés soit estimée avec une bonne précision car ces clones seront rapidement commercialisés sans avoir été testés une seconde fois à large échelle. Au total, six à huit ramets par clone dans un test sont alors un compromis raisonnable : ils permettent d'estimer les valeurs clonales avec une précision presque maximale, même si le gain génétique espéré n'est pas maximal. Pour prendre moins de risque si la commercialisation est moins urgente, une stratégie du type de celle proposée par Libby (1987) peut être utilisée. Elle consiste en plusieurs étapes : d'abord tester un grand nombre de clones avec deux à quatre ramets par clone sur plusieurs sites, puis sélectionner 10 à 20 % de ces clones pour les installer dans un deuxième test afin d'évaluer leurs performances plus précisément, avec un nombre

de ramets plus important. Cette stratégie sera appliquée par l'Inra au merisier après intercroisement des meilleurs clones et au mélèze hybride.

RÉFÉRENCES

- Ballochi CE, Bridgwater FE, Zobel BJ, Jahromi S (1993) Age trends in genetic parameters for tree height in a non selected population of Loblolly pine. *For Sci* 39, 231-251
- Bartlett MS (1978) Nearest neighbour models in the analysis of field experiments (with discussion). *J R Statist Soc B* 40, 147-174
- Franklin EC (1979) Model relating levels of genetic variance to stand development of four North American conifers. *Silvae Genetica* 28, 207-212
- Gallais A (1991) *Théorie de la sélection en amélioration des plantes*. Masson, Paris, 588 p
- Libby WJ (1987) Testing for clonal forestry. *Ann For* 1-2, 69-75
- Libby WJ, Rauter RM (1984) Advantages of clonal forestry. *The Forestry Chronicle*, June 1984, 145-149
- Lindgren D (185) Cost-efficient number of test sites for ranking entries in field trials. *Biometrics* 41, 887-893
- Magnussen S (1993) Growth differentiation in White Spruce crop tree progenies. *Silvae Genetica* 42, 258-265
- Mangin B (1992) SELECT: a program package for assisting in plant selection. *XVth International Biometrics Conference*, Hamilton, New Zealand
- Papadakis J (1937) Méthode statistique pour des expériences en champ. Thessalonique, Institut d'amélioration des plantes. *Bull Sci* 23, 30 p
- Pâques LE (1992) Inheritance and estimated genetics gains in a clonal test of hybrid Larch (*Larix eurolepis*). *Scand J For Res* 7, 355-365
- Pichot C (1993) Analyse de dispositifs par approches itératives prenant en compte les performances des plus proches voisins. *agronomie* 13, 109-119
- Russel JH, Libby WJ (1986) Clonal testing efficiency: the trade-offs between clones tested and ramets per clone. *Can J For Res* 16, 925-930
- Russel JH, Loo-Dinkins JA (1993) Distribution of testing effort in cloned genetic tests. *Silvae Genetica* 42, 98-104
- Shaw DV, JV Hood (1985) Maximising gain per effort by using clonal replicates in genetic tests. *Theor Appl Genet* 71, 392-399