Genet. Res., Camb. (1982), **40**, *pp.* 65–72
Printed in Great Britain

Association of the cyanogenic loci in white clover

By RICHARD A. ENNOS

Department of Genetics, University of Newcastle, Newcastle upon Tyne NE1 7RU

(*Received 19 November 1981, and in revised form 18 January 1982*)

SUMMARY

The inheritance of cyanogenesis in *Trifolium repens* is governed by two unlinked genes, Ac/ac and Li/li which determine the presence/absence of cyanogenic glucosides and their hydrolysing enzyme linamarase respectively. Plants possessing dominant alleles at both loci release hydrogen cyanide when their leaves are damaged. Analysis of phenotype frequencies in ten natural populations showed a significant positive association in phenotypic state between these two loci within populations. Taking into account the results of a subsidiary study which showed no significant deviation from random mating in a natural population of *T. repens*, the most likely explanation for the observed association in phenotypic state is selection favouring the cyanogenic phenotype.

1. INTRODUCTION

It has long been appreciated that interactions between alleles at different loci in the genome affect the fitness of individuals. If certain combinations of alleles are favoured, this can lead to the development of non-random association between alleles at interacting loci (Lewontin & Kojima, 1960). The extent of these non-random associations depends chiefly on the strength of selection for the favoured allelic combination, the recombination rate between the loci, and the breeding system (Karlin & Feldman, 1970; Allard, 1975; Hendrick, Jain & Holden, 1978).

Only rarely has it been possible to study association between loci whose biochemical functions are understood and between which an epistatic interaction affecting fitness is demonstrable. Such a study can be made, however, in natural populations of the self-incompatible perennial herb, white clover (*Trifolium repens* L.), which are polymorphic for the two loci concerned with cyanogenesis (the ability to release hydrogen cyanide (HCN) from damaged leaf tissue).

The inheritance of cyanogenesis is governed by two unlinked loci (Corkill 1943; Atwood & Sullivan, 1943). Ac/ac determines the presence/absence of cyanogenic glucosides in the leaves, and Li/li the presence/absence of their hydrolysing enzyme linamarase. Presence is dominant to absence in both cases. A complementary epistatic interaction occurs between these two loci such that only those individuals possessing dominant alleles at both loci (phenotype *AcLi*) are capable of releasing HCN when damaged.

It has been proposed that cyanogenic individuals are at a selective advantage because cyanogenesis provides protection against the grazing activities of small herbivores (Jones, 1966). Support for this hypothesis has come from both laboratory and field studies (Angseesing, 1974; Angseesing & Angseesing, 1973; Dritschilo *et al.* 1979; Ennos 1981*a*), though selective grazing of acyanogenic individuals has not always been found (Bishop & Korn, 1969; Miller *et al.* 1975). In order to provide an additional test of the selective grazing hypothesis it was decided to investigate the degree of non-random association in phenotypic state between the Ac and Li loci within natural populations of *T. repens*. As understanding of the mating system is of crucial importance when inferences are to be drawn from data relating to phenotypic association at two loci, an estimate of outcrossing rate in a natural population of *T. repens* was made using a convenient, electrophoretic, marker locus.

2. MATERIALS AND METHODS

(i) *Mating System*

Outcrossing rate was estimated using a polymorphic esterase marker gene scored in the following way. Crude leaf extracts were absorbed onto filter paper wicks and run on horizontal starch gels at a constant current of 40 mA for 5 h. The buffers used were: 28 mM-LiOH with 0·19 M boric acid (bridge), and 4 mM citric acid, 46 mM-tris, 3 mM-LiOH, 22 mM boric acid (gel). The staining solution was 0·03 % α,β naphthyl acetate, 0·1 % fast garnet, 0·05 M-tris-HCl buffer pH 7·1. To ascertain the inheritance of the esterase bands, controlled crosses were conducted using the emasculation technique of Williams (1954).

Forty seed heads were collected at least 5 m apart from a natural population of *T. repens* on Tooting Common, London. Seeds were scarified, germinated in petri dishes, and sown by family in seed trays. At the trifoliate leaf stage these progeny were scored for their genotype at the polymorphic esterase locus. Mating system parameters were estimated from progeny genotype arrays using the maximum likelihood programme of Clegg, Kahler & Allard (1978), which is based on the mixed mating model (see Clegg, 1980).

(ii) *Association between Loci*

Ten natural populations of white clover were chosen for analysis from the area around Liverpool and Chester, England. Nine were located in permanent pasture, and one (Ainsdale) in sand-dune slacks. None of the populations had been ploughed, or deliberately sown, within the last 50 years. Approximately 100 stolons were sampled from each population, the distance between each sample being at least 5 m to ensure that the same clones were not repeatedly sampled. Each stolon was rooted in potting compost and grown in a warm greenhouse for two months. The method of de Araujo (1976) was used to determine the phenotype of each sample at the Ac and Li loci.

3. RESULTS

(i) *Mating System*

Plants could readily be scored at a single, codominant, polymorphic, esterase locus with two alleles, Est^F and Est^S. Results of the controlled crosses (Table 1) demonstrate that the locus shows regular, diploid, Mendelian inheritance.

Table 1. *Results of controlled crosses involving Esterase locus, χ^2 values test for deviation from expected Mendelian ratios*

Cross	Progeny Genotype Est^F/Est^F	Est^F/Est^S	Est^S/Est^S	Total	d.f.	χ^2
$Est^F/Est^S \times Est^F/Est^S$	36	70	33	139	2	0·13
$Est^F/Est^S \times Est^S/Est^S$	0	43	43	86	1	0·00

Table 2(*a*). *Expected and observed numbers of progeny estimated using the maximum likelihood programme of Clegg* et al. (1978), *χ^2 for goodness of fit of model, and estimates of frequency of allele* Est^F *in pollen pool* ($\hat{p}$), *and outcrossing rate* ($\hat{t}$) *together with their standard deviations*

Maternal genotype		Progeny genotype Est^F/Est^F	Est^F/Est^S	Est^S/Est^S
Est^F/Est^F	EXP	133	·99	0
Est^F/Est^F	OBS	150	114	0
Est^F/Est^S	EXP	119	199	80
Est^F/Est^S	OBS	122	165	79
Est^S/Est^S	EXP	0	51	31
Est^S/Est^S	OBS	0	50	32

$\chi^2_{(1)} = 10{\cdot}8$ $P < 0{\cdot}001$; $\hat{t} = 1{\cdot}0491$, SD $\hat{t} = 0{\cdot}0591$, $\hat{p} = 0{\cdot}5928$, SD $\hat{p} = 0{\cdot}0209$

Table 2(*b*). *Outcrossing estimates for two homozygous maternal genotypes*

Maternal genotype	$\hat{t}$	S.D. $\hat{t}$	$\hat{p}$
E^F/E^F	1·1161	0·0948	0·6131
E^S/E^S	0·9945	0·0927	

From the progeny genotype arrays the maximum likelihood programme of Clegg *et al.* (1978) was used to infer the genotype of each maternal parent, and to estimate both the frequency of esterase alleles in the pollen pool and the outcrossing rate (Table 2*a*). There is no significant deviation from random outcrossing ($\hat{t} = 1{\cdot}0491$, S.D. $\hat{t} = 0{\cdot}0591$). However the data do not fit the mixed mating model very well ($\chi^2_{(1)} = 10{\cdot}8$, $P < 0{\cdot}001$). There is a marked deficiency of heterozygotes among the progeny of heterozygous parents. For a diallelic locus, equal numbers of heterozygous and homozygous progeny are expected from the heterozygous maternal parents, whatever the outcrossing rate. The deficiency of heterozygous

progeny may indicate selection at the esterase locus (or at loci in tight linkage disequilibrium) during the fertilisation process.

Outcrossing rate may also be calculated using the maximum likelihood estimator $\hat{t} = H/p$ where H = proportion of A_1A_2 genotypes among progeny of A_2A_2 maternal parents and p = frequency of allele A_1 in the pollen pool. Table 2*b* gives outcrossing estimates obtained when homozygous maternal parents are identified using the maximum likelihood programme, and the frequency of allele Est^F is calculated using all progeny scored. Estimates of outcrossing rate are reasonably consistent for the two homozygous genotypes and neither estimate differs significantly from $t = 1{\cdot}0$. Again there is no evidence for any significant level of inbreeding in this population of *T. repens*, either through selfing or population substructuring.

(ii) *Association between Loci*

Numbers of individuals of each phenotype at the ten sites are given in Table 3*a*. For each site, gametic phase disequilibrium D has been calculated on the assumption of random mating, using the expression (Turner, 1968)

$$\hat{D} = \frac{\sqrt{(a_4 n)} - \sqrt{[(a_3 + a_4)(a_2 + a_4)]}}{n},$$

a_2, a_3 and a_4 are the number of individuals of phenotypes *Acli*, *acLi* and *acli* respectively, and n is the total number of individuals scored. The significance of D has been tested using the statistic

$$Q = 4n\hat{D}^2 / \hat{P}_1(2 - \hat{P}_1)\,\hat{q}_1\,(2 - \hat{q}_1),$$

where $\hat{P}_1$ is the frequency of allele Ac and $\hat{q}_1$ the frequency of allele Li in the population. Q is asymptotically χ^2 distributed with one degree of freedom (Hill, 1974).

Eight of the ten populations show positive values of $\hat{D}$, and two of these values are significantly different from $D = 0$. In neither of the two populations showing a negative value of $\hat{D}$ is the deviation from $D = 0$ significant. Taken over all populations, there tends to be a positive association in allelic state between the Ac and Li loci.

A second way of analysing the data is to estimate phenotypic association between the Ac and Li loci, rather than allelic association. This analysis requires no estimation of genotypic frequencies. A multidimensional χ^2 test has been performed on the phenotype frequency data, to look for significant interactions between the factors population, phenotype at the Ac locus, and phenotype at the Li locus (Table 3*b*). The Ac × Li interaction term is highly significant ($P < 0{\cdot}001$) indicating a non-random association in phenotypic state at the Ac and Li loci. Within the limits of experimental error this phenotypic association is consistent over populations (Ac × Li × population interaction term is non-significant). We may conclude that the data provide strong evidence for positive phenotypic association between the Ac and Li loci in natural populations of *T. repens*.

Table 3(*a*). *Numbers of individuals of each phenotype at each site, estimate of gametic disequilibrium* $\hat{D}$, *and test statistic* Q *distributed as* χ^2 *with one degree of freedom* (*Hill*, 1974)

Site	*AcLi*	*Acli*	*acLi*	*acli*	Total	$\hat{D}$	*Q*
Ainsdale	2	23	2	69	96	0·006	1·2714
Brimstage	25	33	10	33	101	0·044	3·9310*
Chester	12	47	10	29	98	−0·012	0·4175
Thornton-le-Moors	14	45	7	30	96	0·010	0·2857
Churton	27	41	7	25	100	0·040	2·7682
Doddleston North	20	30	7	41	98	0·051	7·2487**
Doddleston South	23	21	3	6	53	0·042	0·9181
Chapel House	16	57	7	17	97	−0·016	0·5456
Belgrave Moat	18	55	4	27	104	0·025	1·7518
Grange Farm	6	39	3	51	99	0·014	1·8795

* $P < 0·05$, ** $P < 0·01$.

Table 3(*b*). χ^2 *analysis of data in Table* 3(*a*) *with respect to interaction of the factors population, Ac phenotype, and Li phenotype*

	d.f.	χ^2	
Ac × Population	9	87·7***	—
Li × Population	9	57·96***	—
Ac × Li	1	25·5***	—
Ac × Li × Population	9	10·74	*** $P < 0·001$

4. DISCUSSION

White clover possesses a gametophytic self-incompatability system, governed by a multiple allelic series at a single locus (Atwood, 1942*a*). Although self-compatible and pseudo-self-compatible individuals have been reported, the frequency with which they occur is apparently very low (Atwood, 1942*b*, *c*). In a large population with little or no population subdivision, random mating is therefore expected. In this study no significant difference from random mating was found. Though this result may not seem surprising, a number of other self-incompatible species have been found to have much lower levels of heterozygosity than expected on the assumption of random mating (Brown, 1979). Possible reasons for deficiency of heterozygotes have been reviewed by Brown (1979), who observes that localised inbreeding, due to substructuring of the population coupled with limited pollen dispersal, may be involved. *T. repens* apparently shows no localised inbreeding for the esterase locus in the population studied. It may be significant in this context that *T. repens* clones are relatively mobile, able to grow through the sward by means of rooting stolons (Harper, 1977). Mobility of adult plants will reduce the likelihood of substructuring in the population, and the localised inbreeding which may accompany it if pollen dispersal by bees is low.

The review of multi-locus systems by Hedrick *et al.* (1978), has emphasised that there are many ways in which non-random associations between alleles at different loci can arise, and that a non-random association is not necessarily indicative of

an epistatic fitness interaction. Let us consider whether the non-random association in phenotypic state between the Ac and Li loci, detected in this study, could have been generated in the absence of selection favouring the *AcLi* phenotype. Genetic drift in finite populations may generate non-random associations between alleles at different loci. (Hill & Robertson, 1968). The populations in this study were chosen for their stability over the past 50 years. During this time population numbers are unlikely to have fallen below a thousand at any of the sites. The importance of genetic drift during this time in generating non-random association is likely to be small. Moreover genetic drift would, on average, generate equal numbers of positive and negative associations of the same magnitude. The results, however, show a preponderance of positive associations in allelic state (Table 3), which is not consistent with the generation of non-random association through genetic drift.

Nei & Li (1973) have pointed out that non-random association can be generated if subpopulations with different gene frequencies, but in which there is random association of alleles, are unwittingly combined. Again an equal number of positive and negative associations are anticipated under this model. The preponderance of positive associations in natural populations suggests that this is an unlikely explanation for the results obtained.

Finally, positive phenotypic association between loci can be generated if there is some degree of selfing in the population. Bennett & Binet (1956) have shown that with partial selfing, positive association in genotypic state is generated, even when loci are unlinked. In the populations studied, most individuals of phenotype *Ac* and *Li* are likely to be heterozygous at these loci because of the low frequencies of the dominant alleles. Positive association in genotypic state through selfing would therefore generate positive association in phenotypic state. However, as analysis of the mating system here has given no evidence for selfing in natural populations of *T. repens*, the observed positive phenotypic association cannot be ascribed to this cause.

Selection favouring the cyanogenic (*AcLi*) phenotype therefore appears to be the most likely explanation for the observed positive association in phenotypic state between the Ac and Li loci. We may now inquire at what rate such an association would be likely to decay in the absence of selection. The rate of decay is maximum for random mating populations with no linkage between the loci involved. This is the situation which pertains to the Ac and Li loci in the populations observed. It is perhaps surprising, therefore, that any degree of non-random association in phenotypic state is detectable. It is pertinent to remember, however, that *T. repens* is a perennial species, and in these populations shows a very low rate of seedling recruitment. Therefore, if there is powerful selection for the *AcLi* phenotype during initial seedling establishment, as previous work suggests (Ennos, 1981*a*), non-random phenotypic association could be generated which would persist for many years. Indeed if selection for the *AcLi* phenotype continued after seedling establishment, it would merely intensify the positive association in phenotypic state at the Ac and Li loci. Clearly, when considering the rate of decay of non-random associations between loci, we must take into account not only the breeding system

and linkage relationship of the loci concerned, but also the demography of the species.

Evidence for non-random association between alleles at different loci has been sought in random mating populations of many other species. Consistent non-random association has rarely been found unless the loci are tightly linked and functionally related in some way. Supergenes provide the best documented evidence for non-random associations (Hendrick *et al.* 1978). Thus detection of non-random association in a random mating population when the loci are unlinked constitutes very strong evidence for an epistatic fitness interaction between the loci (Zouros & Johnson, 1976). Detection of positive phenotypic association at the unlinked Ac and Li loci in *T. repens* therefore suggests that powerful selection favouring the cyanogenic *AcLi* phenotype is operating in natural populations. This result is consistent with the findings of many selective grazing experiments in the laboratory and in the field (Angseesing, 1974; Angseesing & Angseesing, 1973; Dritschilo *et al.* 1979; Ennos, 1981*a*), and adds weight to their conclusions.

Selection for the cyanogenic phenotype will initially generate non-random association between the Ac and Li loci. However the ultimate result of such selection will be the fixation of the cyanogenic phenotype in the population. There is some evidence that competitive interactions between *Li* and *li* phenotypes are such as to maintain polymorphism at this locus (Ennos, 1981*b*). However much remains to be learned about the maintenance of polymorphism at the Ac and Li loci in *T. repens*.

I would like to thank Dr M. T. Clegg for valuable help and discussion, Dr M. R. Macnair for help with the analysis, Z. J. Nugent for collection of seed material, and the late Dr J. A. Bishop for his enthusiasm and encouragement. Part of this work was conducted while I was in receipt of a Liverpool University Postgraduate Studentship.

REFERENCES

Allard, R. W. (1975). The Mating System and Microevolution. *Genetics*, **79**, 115–126.

Angseesing, J. P. A. (1974). Selective eating of the acyanogenic form of *Trifolium repens*. *Heredity*, **32**, 73–83.

Angseesing, J. P. A. & Angseesing, W. J. (1973). Field observations on the cyanogenesis polymorphism in *Trifolium repens*. *Heredity*, **31**, 276–282.

de Araujo, A. M. (1976). The relationship between altitude and cyanogenesis in white clover (*Trifolium repens* L.). *Heredity*, **37**, 291–293.

Atwood, S. S. (1942*a*). Oppositional alleles causing cross-incompatability in *Trifolium repens*. *Genetics*, **27**, 333–338.

Atwood, S. S. (1942*b*). Genetics of self-compatability in *Trifolium repens*. *Journal of the American Society of Agronomy*, **34**, 352–364.

Atwood, S. S. (1942*c*). Genetics of pseudo-self-compatability and its relation to cross-incompatability in *Trifolium repens*. *Journal of Agricultural Research*, **64**, 699–709.

Atwood, S. S. & Sullivan, J. T. (1943). Inheritance of a cyanogenetic glucoside and its hydrolysing enzyme in *Trifolium repens*. *Journal of Heredity*, **34**, 311–320.

Bennett, J. H. & Binet, F. E. (1956). Association between mendelian factors with mixed selfing and random mating. *Heredity*, **10**, 51–55.

Bishop, J. A. & Korn, M. E. (1969). Natural selection and cyanogenesis in white clover, *Trifolium repens*. *Heredity*, **24**, 423–430.

BROWN, A. H. D. (1979). Enzyme polymorphism in plant populations. *Theoretical Population Biology*, **15**, 1–42.

CLEGG, M. T. (1980). Measuring plant mating systems. *BioScience*, **30**, 814–818.

CLEGG, M. T., KAHLER, A. L. & ALLARD, R. W. (1978). Estimation of life cycle components of selection in an experimental plant population. *Genetics*, **89**, 765–792.

CORKILL, L. (1943). Cyanogenesis in white clover (*Trifolium repens* L.). V. The inheritance of cyanogensis. *New Zealand Journal of Science and Technology*, **23**B, 178–193.

DRITSCHILO, W., KRUMMEL, J., NAFUS, D. & PIMENTEL, D. (1979). Herbivorous insects colonising cyanogenic and acyanogenic *Trifolium repens*. *Heredity*, **42**, 49–56.

ENNOS, R. A. (1981*a*). Detection of selection in populations of white clover (*Trifolium repens* L.). *Biological Journal of the Linnean Society*, **15**, 75–82.

ENNOS, R. A. (1981*b*). Manifold effects of the cyanogenic loci in white clover. *Heredity*, **46**, 127–132.

HARPER, J. L. (1977). *Population Biology of Plants*. London, Academic Press.

HENDRICK, P., JAIN, S. & HOLDEN, L. (1978). Multilocus systems in evolution. *Evolutionary Biology*, **11**, 101–184.

HILL, W. G. (1974). Estimation of linkage disequilibrium in randomly mating populations. *Heredity*, **33**, 229–239.

HILL, W. G. & ROBERTSON, A. (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, **38**, 226–231.

JONES, D. A. (1966). On the polymorphism of cyanogenesis in *Lotus corniculatus*. Selection by animals. *Canadian Journal of Genetics and Cytology*, **8**, 556–567.

KARLIN, S. & FELDMAN, M. W. (1970). Linkage and selection: Two-locus symmetric viability model. *Theoretical Population Biology*, **1**, 39–71.

LEWONTIN, R. C. & KOJIMA, K. (1960). The evolutionary dynamics of complex polymorphisms. *Evolution*, **14**, 458–472.

MILLER, J. D., GIBSON, P. B., COPE, W. A. & KNIGHT, W. E. (1975). Herbivore feeding on cyanogenic and acyanogenic white clover seedlings. *Crop Science*, **15**, 90–91.

NEI, M. & LI, W. H. (1973). Linkage disequilibrium in subdivided populations. *Genetics*, **75**, 213–219.

TURNER, J. R. G. (1968). On supergenes. II. The estimation of gametic excess in natural populations. *Genetica*, **39**, 82–93.

WILLIAMS, W. (1954). An emasculation technique for certain species of *Trifolium*. *Agronomy Journal*, **46**, 182–184.

ZOUROS, E. & JOHNSON, W. (1976). Linkage disequilibrium between functionally related enzyme loci of *Drosophila mojavensis*. *Canadian Journal of Genetics and Cytology*. **18**, 245–254.