Original article

Multilocus structure of the smooth newt (*Triturus vulgaris*, Caudata) natural populations

J Crnobrnja [1], ML Kalezić [2*], N Tucić [2]

[1] *Institute for Biological Research "Siniša Stanković", 29 Novembra 142, 11000 Belgrade;*
[2] *Institute of Zoology, Faculty of Biology, Studentski trg 16, 11000 Belgrade, Yugoslavia*

Summary – The linkage disequilibrium between pairs of 7 polymorphic loci in 27 natural populations of the smooth newt (*Triturus vulgaris*) was examined. Pairwise linkage disequilibrium parameters were estimated from zygotic frequencies using Burrow's method. The average rate of significance of the linkage disequilibrium parameter in 27 populations was about 8%. The variation of linkage disequilibrium among populations was studied by analysis of variance of the correlation coefficients between different loci in zygotes. This analysis did not reveal systematic associations between alleles at different loci over 27 populations. In only one case, Me (malic enzyme) × Pgm (phosphoglucomutase), were correlation coefficients of the same sign and magnitude in a number of populations.

linkage disequilibrium / multilocus association / allozyme / *Triturus vulgaris*

Résumé – Structure multilocus de populations naturelles de triton vulgaire (*Triturus vulgaris*, Caudata). *Dans 27 populations naturelles de triton vulgaire (*Triturus vulgaris*) on a examiné les déséquilibres de liaison entre 7 locus polymorphes pris 2 à 2. Les déséquilibres de liaison entre locus ont été estimés à partir des fréquences zygotiques en utilisant la méthode de Burrows. Le pourcentage moyen de déséquilibres de liaison significatifs dans les 27 populations est proche de 8 %. La variation des déséquilibres de liaison entre les populations a été étudiée par analyse de variance des coefficients de corrélation des gènes non alléliques dans les zygotes. Cette analyse ne montre pas d'associations systématiques entre les gènes non alléliques dans les 27 populations. Dans un seul cas, Me (enzyme malique) × Pgm (phosphoglucomutase), les coefficients de corrélation ont le même signe et la même ampleur dans plusieurs populations.*

déséquilibre de liaison / association multilocus / allozyme / *Triturus vulgaris*

---

* Correspondence and reprints

INTRODUCTION

It has been suggested that association between alleles at different loci (linkage disequilibrium, LD) might be a useful indicator of the action of natural selection (Lewontin, 1974). However, several workers have pointed out that such associations could arise by genetic drift, non-random matings, founder effects and hitchhiking (see Hedrick, 1982 for review). Thus, the mere presence of allozyme LD, for example, does not appear to be critical in the evaluation of the adaptive significance of allozyme variation.

Although it is an extremely difficult task to attribute any pattern of LD to a particular cause, one approach has been to consider evidence from more than one population. Lewontin (1974), for example, suggested that if significant LD is observed which is consistent in the magnitude and sign in many populations, then this pattern can be attributed to selection.

There have been several experimental studies designed to detect LD among allozymes in animal and plant populations (reviewed by Barker, 1979; Brown, 1979). Disequilibrium has been found in several animal species, such as salamanders (Webster, 1973; Good, 1989), blue mussel (Mitton and Koehn, 1973), and the fish (Mitton and Koehn, 1975). Most studies, however, have been with several species of *Drosophila*, and particularly *D melanogaster* (*eg* Langley *et al*, 1978; Laurie-Ahlberg and Weir, 1979 and references therein). Although *Drosophila* studies have frequently shown LD between allozymes and inversions, there is little evidence for stable LD among allozymes in *Drosophila* as in other animal populations.

We have previously (Kalezić and Tucić, 1984; Gonzales-Candelas *et al*, in press) described allele frequencies as well as different environmental and geographical variables which influenced the genetic structure of *Triturus vulgaris* populations. Here we report a survey of LD among 7 polymorphic allozymes in the same 27 natural populations of the common newt *Triturus vulgaris*. The main objective of this study was to seek for consistency in the magnitude and direction of LD among *T vulgaris* populations.

MATERIALS AND METHODS

The study of linkage disequilibrium was carried out on 16 populations of the nominotypical subspecies, 5 populations of *T v meridionalis*, 3 populations of *T v dalmaticus*, and 3 populations of *T v graecus*. For population localities, number of individuals surveyed, and loci abbreviations see Kalezić (1983). Among the 22 loci studied, the following loci were moderately to highly polymorphic in most of the analyzed populations: acid phosphatase (Acph-2), esterase (Est-4), α-glycerophosphatase (α-Gpdh-), malate dehydrogenase (Mdh-2), malic enzyme (Me), octanol dehydrogenase (Odh), and phosphoglucomutase (Pgm). Several loci were polymorphic for more than 2 electromorphs. In these cases, the least common alleles were pooled so that there were just 2 allelic classes for the estimation of LD. The samples, with about 40 individuals, were assayed at each pair of loci. The estimates of LD were made from zygotic frequencies. The method of estimation, proposed by PM Burrows (see Cockerham and Weir, 1977; Weir, 1979), incorporates

the departures from Hardy–Weinberg equilibrium for the sample frequencies at each locus and does not require the assumption of random mating.

An unbiased estimate of Burrows' coefficient of LD is:

$$D = 1/2[4f(11/11) + 2f(11/01) + 2f(11/10) + f(11/00) + f(10/01)] - 2p_1p_2$$

where p_1 and p_2 are the allele frequencies of the "1" alleles at 2 loci, respectively, $f(11/11)$ is the frequency of double homozygotes for the "1" alleles, and $f(11/01)$ is the frequency of zygotes heterozygous at the first locus and homozygous at the second locus for the "1" allele.

The statistical significance of Burrows' coefficient can be tested as follows (Langley et al, 1978):

$$ND^2/p_1p_2(1 - p_1)(1 - p_2) = 4NR^2 = \chi^2$$

where χ^2 has an approximate chi-square distribution with one degree of freedom (Cockerham and Weir, 1977), N is the number of individuals in the sample and R a correlation coefficient defined by Langley et al (1978):

$$R = D/2[p_1p_2(1 - p_1)(1 - p_2)]^{1/2}.$$

This coefficient corresponds to the total correlation between genes, including within gametes as well as between uniting gametes correlations. R generally lies between -1 and $+1$, but between -0.5 and $+0.5$ only when there is no correlation between uniting gametes.

The variation of LD among populations was evaluated by analysis of variance of the correlation coefficients of different genes in zygotes. To analyze variation of R_k (k indexes the kth population) attributable to the differences between populations of different subspecies (d_A^2) or differences between populations of the same subspecies (d_{AB}^2), a weighted analysis of variance (where weights, $4N_{ik}$, are the reciprocals of sampling variances) was performed (for the analysis of variance scheme and more details see Langley et al, 1978, pp 217–220). A test of $d_A^2 = 0$ is the usual F ratio test with $(m_2 - 1)$ and $(m_1 - m_2)$ degrees of freedom (m_2 is the number of subspecies). But, if $d_{AB}^2 = 0$ then the weighted sum of squares at AB level is approximately equal to a chi-square distribution with $(m_1 - m_2)$ degrees of freedom.

RESULTS

Burrows' disequilibrium parameters were calculated for 318 combinations of pairs of polymorphic loci in 27 smooth newt populations. Since pairwise values of D for analyzed loci are too numerous to report here, the results are summarized. Table I gives only these locus pairs in each population with significant D values. There were 24 significant D values among polymorphic loci in these populations. If the disequilibrium parameters were independent, then 16 ($= 0.05 \times 318$) significant values would be expected due to chance alone at the 0.05 probability level. More importantly, the significant values of D were nonrandomly distributed over populations. The percent significant pairs of loci per population varies from zero

Table I. Locus pairs with significant D in 27 populations of 4 subspecies of *T vulgaris*. Percent of significant D out of total combinations of loci pairs for each population is given in parentheses.

Population	Number of polymorphic loci	Locus pair with significant D (in parentheses)
T v vulgaris		
1 Jankovac	6 (6.67%)	Mdh-2 × Me (0.032 9)
2 Šuljam	6 (6.67%	Acph-2 × Me (−0.081 7)
3 Zmijinje Lake	3 (33.33%)	Est-4 × Me (0.050 0)
4 Zminičko Lake	4 (16.67%)	Mdh-2 × Odh (0.043 3)
5 Tutin	6 (0.00)	None
6 Radevo	4 (0.00)	None
7 "Jezero"	4 (0.00)	None
8 Belo Polje	5 (10.00%)	Mdh-2 × Odh (0.030 8)
9 Rtanj	6 (20.00%)	Acph-2 × Odh (0.035 9), Est-4 × α-Gpdh (0.079 2) α-Gpdh × Me (0.087 5)
10 Sisevac	7 (0.00)	None
11 Smilčići	5 (20.00%)	Mdh-2 × Me (0.044 8), Mdh-2 × Odh (0.012 1)
12 Orle	7 (9.52%)	Acph-2 × Odh (0.101 8), Odh × Pgm (0.030 5)
13 Jelah	7 (9.52%)	Acph-2 × Me (−0.081 4), α-Gpdh × Odh (0.030 0)
14 Barn Pet Sel	7 (9.52%)	Acph-2 × α-Gpdh (0.017 8), Acph-2 × Odh (0.016 2)
15 Ravenica	7 (4.76%)	Est-4 × Me (−0.065 0)
16 Bukovac	7 (4.76%)	α-Gpdh × Odh (−0.057 5)
T v meridionalis		
17 Podstrmec	7 (14.29%)	Est-4 × Mdh-2 (0.036 9), α-Gpdh × Pgm (0.053 9), Mdh-2 × Me (−0.045 9)
18 Salakovac	6 (0.00)	None
19 Krk	4 (0.00)	None
20 Švica	6 (0.00)	None
21 Lički Osik	5 (0.00)	None
T v dalmaticus		
22 Domanovići	3 (0.00)	None
23 Kruševica	3 (33.33%)	α-Gpdh × Me (0.010 4), Me × Odh (−0.051 9)
24 Balj Lok	1 (0.00)	None
T v graecus		
25 Streževo	1 (0.00)	None
26 Visoka Čuka	5 (0.00)	None
27 Bitolj	4 (16.67%)	Mdh-2 × Me (0.1006)

(in 12 populations) to 33.3% (in 2 populations). The average percent of significant D values over all 27 populations is 7.99%.

In 2 populations, one of *T v vulgaris* and one of *T v meridionalis*, 3 combinations of locus pairs exhibited significant D values (table I). Four populations of *T v vulgaris* and 1 population of *T v dalmaticus* had statistically significant D values

for 2 combinations of loci. The Mdh-2 × Me combination showed the most frequent significant value of D (in 4 populations), and in 3 populations significant D's for Acph-2 × Odh, a-Gpdh × Odh and Mdh-2 × Odh were detected. The Odh locus shows significant disequilibrium parameters with Acph-2, a-Gpdh and Mdh-2 in 3 populations, and with Pgm and Me loci in and 1 population (table I).

Table II shows the results of the analysis of variance of the correlation coefficients based on Burrows' disequilibrium parameters over 27 populations divided into 4 smooth newt subspecies. In this table d_A and d_{AB} are reported rather than their squared values (variance component attributable to subspecies differentiation in R_k, and variance component attributable to population variation in R_k, respectively), so that the scale is the same as that for R. Often d_A and d_{AB} have actual numerical estimates which are negative, in which case 0.0 is reported.

Table II. Analyses of variance of the correlation coefficient between alleles at different loci.

Locus × locus	m_1	m_2	N	$\widehat{R}$	d_A	d_{AB}	Compared subspecies
Acph-2 × Est-4	10	2	40.00	−0.009 3	0.0933	0.00	v,m
Acph-2 × α-Gpdh	9	3	39.56	0.026 5	0.00	0.204 5**	v,m,g
Acph-2 × Mdh-2	12	3	39.08	0.067 5**	0.00	0.337 3***	v,m,g
Acph-2 × Me	11	3	39.00	−0.005 5	0.00	0.149 7***	v,m,g
Acph-2 × Odh	12	3	39.08	0.076 7***	0.00	0.122 3***	v,m,g
Acph-2 × Pgm	10	2	39.30	−0.050 0	0.00	0.086 0*	v,m
Est-4 × α-Gpdh	13	2	39.23	0.006 6	0.108 6**	0.078 1*	v,m
Est-4 × Mdh-2	19	2	38.37	−0.009 2	0.056 6**	0.035 9	v,m
Est-4 × Me	18	2	38.28	0.043 7*	0.00	0.102 9***	v,m
Est-4 × Odh	17	2	39.12	−0.012 2	0.00	0.030 4	v,m
Est-4 × Pgm	13	2	38.23	0.016 2	0.00	0.055 7	v,m
α-Gpdh × Mdh-2	14	3	39.29	−0.015 9	0.00	0.096 1**	v,m,g
α-Gpdh × Me	18	4	39.22	−0.013 6	0.119 2*	0.142 9***	v,m,d,g
α-Gpdh × Odh	17	4	39.18	0.010 0	0.00	0.086 3**	v,m,d,g
α-Gpdh × Pgm	9	2	38.45	0.001 1	0.00	0.122 5**	v,m
Mdh-2 × Me	23	3	38.17	0.026 7	0.162 3***	0.161 2***	v,m,g
Mdh-2 × Odh	21	3	38.76	0.014 5	0.040 7*	0.047 5	v,m,g
Mdh-2 × Pgm	14	2	37.86	−0.050 2	0.015 8	0.00	v,m
Me × Odh	24	4	38.92	−0.006 9	0.025 8	0.043 8	v,m,d,g
Me × Pgm	15	2	38.00	−0.046 6*	0.029 3	0.051 9	v,m
Odh × Pgm	13	2	38.39	−0.042 6	0.092 4**	0.056 8	v,m

m_1 : The number of populations, m_2: the number of subspecies, N: the average number of individuals per population, $\widehat{R}$: the weighted mean of the correlation coefficient, d_A and d_{AB}: square roots of the variance components attributable to subspecies and population variation, respectively. Acronyms v, m, d, and g denote subspecies *T v vulgaris, T v meridionalis, T v dalmaticus* and *T v graecus,* respectively. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

The most noteworthy comparison in table II is between the Me and Pgm loci that show a significant weighted mean of the correlation coefficient ($\widehat{R} = -0.047$,

$P < 0.05$) and nonsignificant variance components at both levels. This indicates a consistency in R_k over the whole set of 15 compared populations of *T v vulgaris* and *T v meridionalis*. The average correlation coefficients were also significant for Acph-2 × Mdh-2 (0.068, $P < 0.01$), Acph-2 × Odh (0.077, $P < 0.001$) and Est-4 × Me (0.044, $P < 0.05$). However, in all these cases the highly significant d^2_{AB} indicate great variation in R_k values among analyzed populations. Although $\widehat{R}$ was not significantly different from zero, the highly significant d^2_A for Est-4 × Mdh-2, Mdh-2 × Odh and Odh × Pgm indicate differences in R_k values between *T v vulgaris, T v meridionalis* and *T v graecus* (second loci pair), and *T v vulgaris* and *T v meridionalis* (first and third comparisons of loci pairs). There are 3 combinations of loci pairs (Est-4 × α-Gpdh, α-Gpdh × Me and Mdh-2 × Me) that show significant d_A and significant d_{AB}. The other comparisons were either insignificant individually for all parameters or showed only significant d_{AB} values.

DISCUSSION

Two studies of LD in natural and laboratory populations of *Drosophila melanogaster* involve the use of genotypic data to calculate Burrows' disequilibrium parameter and are thus comparable to the data reported here. Langley *et al* (1978) studied 8 enzyme loci in some 100 samples from natural populations and in 2 laboratory populations. The frequency of cases of significant LD between pairs of loci in natural populations was 5.1% for linked genes and 6.7% for loci on different chromosomes. In the 2 laboratory populations, the frequency of significant disequilibrium was much greater: 37.5% for linked loci pairs and 10.3% for unlinked loci. Laurie-Ahlberg and Weir (1979) studied 17 enzyme loci in 9 laboratory populations of *Drosophila melanogaster* and found significant associations at frequencies fairly similar to those of Langley and collaborators in laboratory populations: 34.5% and 8.9% for linked and unlinked pairs of loci, respectively.

The frequency of significant cases of LD in the smooth newt populations is nearly as large as in the laboratory populations of *Drosophila* for unlinked pairs of loci. The average rate of significant D values in 27 populations of *Triturus vulgaris* amounted to about 8% (table I). Since we have no genetic data on the positions of these loci on chromosomes, it is reasonable to assume that we are dealing with unlinked pairs of loci. It is also important to note that analyses of LD were based on relatively small samples (about 40 individuals per population). But, as has been shown by Brown (1975) and Marinković *et al* (1987), the sample size necessary to detect significant LD must often be quite large. Thus, it is to be expected that more instances of LD would appear in studies of *T vulgaris* populations with larger sample size.

Analysis of variance of the correlation coefficients between alleles at different loci over 27 populations of the 4 smooth newt subspecies did not reveal systematic associations across the range of studied species. In only one case (Me × Pgm) were the R_k values of the same sign and magnitude in a number of populations (which gives rise to significant R and nonsignificant d_A and d_{AB}; table II). This observation could be accounted for by epistatic selection. The lack of consistency in the magnitude and direction of LD among populations in all other cases can be consistent with both neutralist and selectionist hypotheses. In several cases, both regionally consistent selection (within the subspecies range, as in the case of Est-4 ×

Mdh-2, Mdh-2 × Odh and Odh × Pgm which show significant d_A and nonsignificant d_{AB} values) and genetic drift *via* founder effects or population subdivision might be causes of the observed associations. Evidence provided by Gonzales-Candelas *et al* (in press), for the same populations of the smooth newt, that frequent extinction and recolonization were responsible for genetic structure of these populations, also indicate that founder effects might be a primary cause of the observed frequencies of linkage disequilibrium.

ACKNOWLEDGMENTS

REFERENCES

Barker SF (1979) Inter-locus interactions: a review of experimental evidence. *Theor Biol* 16, 323-346

Brown AHD (1975) Sample sizes required to detect linkage disequilibrium between two or three loci. *Theor Popul Biol* 8, 184-201

Brown AHD (1979) Enzyme polymorphism in plant populations. *Theor Popul Biol* 15, 1-42

Cockerham CC, Weir BS (1977) Digenic descent measures for finite populations. *Genet Res* 30, 121-147

Gonzalez-Candelas F, Crnobrnja J, Kalezić ML, Moya A (1992) Gene flow rates in Yugoslavian populations of the smooth newt *Triturus vulgaris*. *J Evol Biol*

Hedrick PW (1982) Genetic hitchhiking: a new factor in evolution? *BioScience* 32, 845-853

Good DA (1989) Hybridization and cryptic species in *Dicamptodon* (Caudata: Dicamptodontidae). *Evolution* 43 728-744

Kalezić ML (1983) Geographical aspects of genetic variability in the smooth newt (*Triturus vulgaris*). *Genetika (Belgrade)* 15, 93-103

Kalezić ML, Tucić N (1984) Genic diversity and population genetic structure of *Triturus vulgaris* (Urodela, Salamandridae). *Evolution* 38, 389-401

Langley CH, Smith DB, Johnson FM (1978) Analysis of linkage disequilibrium between allozyme loci in natural populations of *D melanogaster*. *Genet Res* 32, 215-230

Laurie-Ahlberg CC, Weir BS (1979) Allozymic variation and linkage disequilibrium in some laboratory populations of *D melanogaster*. *Genetics* 92, 1295-1314

Lewontin RC (1974) *The Genetic Basis of Evolutionary Change*. Columbia Univ Press, NY

Marinković D, Tucić N, Moya A, Ayala FJ (1987) Genetic diversity and linkage disequilibrium in *D melanogaster* with different rates of development. *Genetics* 117, 513-520

Mitton JB, Koehn RK (1973) Population genetics of marine pelecypods. III. Epistasis between functionally related isoenzymes of *Mytilus edulis. Genetics* 73, 487-496

Mitton JB, Koehn RK (1975) Genetic organization and adaptive response of allozymes to ecological variables in *Fundulus heteroclitus. Genetics* 79, 97-111

Webster TP (1973) Adaptive linkage disequilibrium between two esterase loci of a salamander. *Proc Natl Acad Sci USA* 70, 1156-1160

Weir BS (1979) Inferences about linkage disequilibrium. *Biometrics* 35, 235-254