

Genetic and ecological differentiation between the butterfly sisterspecies *Colias alfacariensis* and *Colias hyale*

Daniel F.R. Cleary¹, Henri Descimon² and Steph B.J. Menken¹

¹Institute for Biodiversity and Ecosystem Dynamics (IBED), Faculty of Science, University of Amsterdam, P.O. Box 94766, 1090 GT Amsterdam, The Netherlands, cleary@science.uva.nl;

²Laboratoire de Systématique Évolutive, Université de Provence, Marseille, France

Keywords: Lepidoptera, Pieridae, allozymes, population structure, gene flow

Abstract

Clear habitat separation between the sister species *Colias alfacariensis* and *C. hyale* is shown when occurring sympatrically. *Colias hyale* is found more often in moist cultivated pastures while *Colias alfacariensis* is more abundant in dry uncultivated habitat. Out of a total of 16 loci, no diagnostic loci were found between *C. alfacariensis* and *C. hyale*, and both species shared most major polymorphisms. Exceptions were the marked differences in allele frequencies at the HK locus and only *C. hyale*, but not *C. alfacariensis* was further invariable at the GOT2 locus, which is usually highly polymorphic in the Pieridae. *Colias hyale* has a significantly lower level of heterozygosity than its sister species *C. alfacariensis*. In *Colias alfacariensis* heterozygosity is highest in the Alps and lowest in the low-lying region of Northern France. Both species show high levels of gene flow over a large geographic area. Within *C. alfacariensis*, but not in *C. hyale*, the F_{ST} value of the PGI locus is significantly different from zero effectively separating the species into populations with high levels of the 'b' allele to the west and North, and low levels of the allele in the Alps and Italy. This could point to selection within the PGI locus in line with the well established pattern of selection at the PGI locus in other species of *Colias*. Glaciations have been an important force in shaping the evolutionary history of European biota, leading to extinction, but also allowing new species to evolve into the newly available land as the ice sheets retreated. The genetic and distributional pattern found between both *Colias* species suggests that habitat shifts and subsequent adaptation during glaciations could have played an important role in their speciation.

Contents

Introduction	131
Material and methods	132
Habitat preference	132
Population analysis	132
Results	134

Habitat preference	134
Population analysis	134
Discussion	136
Acknowledgements	138
References	138

Introduction

Coexistence of species requires the existence of evolutionarily continual interspecific trade-offs in the abilities of species to deal with factors that limit their fitness and abundance. These include trade-offs between competitive abilities and dispersal abilities, susceptibility to disease, herbivory or predation, between abilities to live off average conditions or abilities to exploit resource pulses, and between abilities to compete for alternative resources in a heterogeneous landscape (Tilman, 2000). Feeding ecology, reproductive behaviour, and physiology are characteristics of a species niche. Niches tend to be conservative over time scales associated with sister taxon pairs in, for example, birds, mammals, and butterflies but this conservatism breaks down at the level of families (Peterson et al., 1999). By analysing the present day difference in genetic makeup and the often subtle differences in ecology between species within a pair of sister species we can get an idea of the evolutionary forces that led to their differentiation into separate species.

The present study concerns two species of the genus *Colias* [Lepidoptera: Pieridae] namely *C. alfacariensis* and *C. hyale* (Brunton, 1998). *Colias* is one of the most widespread butterfly genera in the world. It is present on all continents except

Australia, and ranges from the Antarctic (e.g., *Colias antarctica*) to the Arctic (e.g., *Colias boothi*), across Asia, Africa, North and South America, from the lowlands to the highest mountains.

Colias alfacariensis and *C. hyale* are sister species (Brunton, 1998), and are difficult to tell apart due to the large variation in wing patterns and other morphological characters of the butterflies. In general, *C. hyale* has a more northerly and *C. alfacariensis* a more southerly distribution with a large region of overlap in central Europe. In recent years *C. hyale* has become increasingly rare in northern Europe due to abandonment in agricultural areas of its larval food plant *Medicago sativa* (Descimon, (1994(5)).

Colias alfacariensis flies in May-June and August-September in two, and in southern Europe sometimes in three generations. Food plants are *Hippocrepis comosa* and *Coronilla varia*. *Colias hyale* has similar flight periods and the food plants are *Medicago sativa*, *Vicia cracca*, *Coronilla* spp., *Lotus* spp., *Cytisus* spp., *Trifolium* spp., and other Fabaceae (Higgins and Riley, 1970). Both species are thus oligophagous on Fabaceae although *C. hyale* has a larger range of potential food plants.

The goal of this study is to analyse the ecology and genetical population structure of both species and to try to gain insight into the differentiation process that produced both present-day taxa. Lepidoptera are an important source for understanding the genetics of speciation (Menken et al., 1992; Ritchie and Phillips, 1998). The analysis of spatial pattern (e.g., relative densities and habitat preferences), direct measures of migration (e.g., Mark-release-recapture), and indirect measures (e.g., genetic differentiation) are all important for understanding how populations are structured. Lewis et al. (1997; and refs therein) suggest that at least two of these methods should be applied in order to provide a good indication of population structure.

The specific questions we address are the following. 1) Do both species differ in their habitat preference? 2) What is the spatial, genetical population structure of both species?

Materials and methods

Habitat preference

Individuals of both *C. alfacariensis* and *C. hyale* were sampled for an allozyme study in France and Italy in the summer of 1992. Species determination was based on the studies of Berger and Fontaine (1947), Jarvis (1955), and Dutreix (1980). Habitat preference between *C. alfacariensis* and *C. hyale* was studied by means of field observations in the French Alps where both species occur sympatrically.

A total of 920 adult *C. alfacariensis* and 1155 adult *C. hyale* were observed from the 14th of August until the 8th of September, 1974 by HD. The location and geographical position of all sites is presented in Table 1. During the field observations the following environmental variables were noted at each locality: 1. Aspect, 2. Vegetation (cultivated – uncultivated), 3. Humidity (dry – irrigated), 4. Altitude 5. Date, 6. Weather (warm – cold), in addition to the frequency of occurrence of adults of *C. alfacariensis* and *C. hyale*.

The relative abundances of both species were regressed onto the first axis of a principal components analysis (PCA based on the environmental variables and using Statistica for Windows, version 5, StatSoft Inc.) to see if both species varied significantly (positively or negatively) in their response to this synthetic environmental gradient. The relative abundance here is the absolute abundance of species A at a given site divided by the absolute abundances of both species at the same site. We use relative abundance here because of the high variance in the absolute abundance among sites.

Population analysis

For the allozyme electrophoresis analyses, populations of both *Colias* species were sampled in the French Alps and other localities, including the Provence, Northern France and Italy (Table 1). After capture the butterflies were stored in liquid nitrogen. The head and thorax were subsequently homogenised in the laboratory in 180 ml of a 0.1M Tris-HCl buffer (pH 8.0) and the extracts were cen-

Table 1. Location and geographical position of all sampling sites. Allozyme indicates that the population sampled was used in allozyme electrophoretic analysis. Field indicates that the site was used in the habitat preference study. The two populations from northern France were collected by a colleague in the surroundings of Paris and Metz. Population refers to the populations used for electrophoresis. A – J: *C. alfacariensis*, K – M: *C. hyale*.

Population	Site	Location	Lat	Long	Allozyme	Field
A	Poggio del Commune	Italy	44 57.64	6 41.01	x	
B+M	Val de Prés	French Alps	44 58.47	6 30.33	x	x
C+L	Le Monétier les Bains	French Alps	44 53.52	6 37.02	x	x
D	Puy Saint Pierre	French Alps	44 53.52	6 37.02	x	
E	Saint Zachery	French Alps	43 23.07	5 42.22	x	
F	Pontèves	Provence	43 33.55	6 02.53	x	
G	Artigues	Provence	43 36.05	5 47.54	x	
H	Mimet	Provence	43 24.54	5 30.19	x	
I	Seine River outside Paris	Provence			x	
J+K	Surroundings of Metz	Northern France			x	
	Gero	Northern France	44 59.03	6 28.56		x
	les Guibertes	French Alps	44 58.01	6 31.57		x
	le Cassset	French Alps	44 59.03	6 28.56		x
	Saint Chaffery	French Alps	44 55.32	6 36.11		x
	le Rosier	French Alps	44 56.02	6 40.58		x
	Villard Saint Pancrace	French Alps	44 52.09	6 38.05		x
	la Vachette	French Alps	44 54.58	6 39.55		x
	Laus de Cervières	French Alps	44 51.40	6 43.35		x
	Granon	French Alps	44 58.04	6 39.01		x
	Cervièrès	French Alps	44 51.56	6 43.21		x

trifuged for 6 minutes at 9000 x g. The supernatant from the extract was then frozen until used for electrophoresis.

Two gel matrices were used, namely cellulose acetate and starch gels. The starch gels were prepared according to Harris and Hopkinson (1977) and the exact preparation of buffers and gels follows Raijmann et al. (1997). Two buffer solutions were used in the starch gels, namely a Tris-Citric acid buffer (TC 8.0) and a Tris-Malic acid buffer (TM 7.4). Agar overlays were applied to all gel slices.

The following loci (locus abbreviations in parentheses) gave consistently good results and were used for the analyses: Aldolase (*ald*), Creatine Kinase 1 (*ck1*) and 2 (*ck2*), Glutamate-oxaloacetate transaminase 1 (*got1*) and 2 (*got2*), Glycerol-3-phosphate dehydrogenase (*gpd*), Hexokinase (*hk*), Isocitrate dehydrogenase 1 (*idh1*) and 2 (*idh2*), Malate dehydrogenase 1 (*mdh1*) and 2 (*mdh2*), Malic enzyme (*me*), Mannose phosphate isomerase (*mpi*), Phosphogluconate dehydrogenase (*6pgdh*), Glucose phosphate isomerase (*pgi*), and Phosphoglucomutase (*pgm*). The loci *ck1*, *ck2*, *gpd*, *got2*, and *pgm*

gave the best resolution with TM; *got1*, *hk*, *mdh1*, *mdh2*, *me*, *mpi*, *idh1*, *idh2*, and *pgi* gave the best resolution with TC; *6pgdh* gave the best resolution with a Tris-Glycine buffer (pH 8.5) on cellulose acetate. Locus number 1 refers to the least anodally migrating locus. Allelic designations use 'a' as the least anodally migrating allele, then 'b' and so forth.

Individual genotypes were used as input data for BIOSYS-1 (Swofford and Selander, 1981). The allelic frequencies and mean heterozygosity were derived directly from the BIOSYS output. Variation in heterozygosity between *C. alfacariensis* and *C. hyale* was tested for significance with a one-way Anova (using Statistica for Windows, version 5, StatSoft Inc.).

The inbreeding coefficients (F_{ST} , F_{IT} , and F_{IS} ; Wright, 1951) were calculated according to Weir and Cockerham's (1984) F , f , and Θ respectively. F_{ST} values were tested for departure from zero by the method of Workman and Niswander (1970), and inbreeding coefficients (F_{IS}) were tested by the method of Li and Horvitz (1953).

The amount of gene flow ($N_e m$) was estimated

by the F statistic (F_{ST}) method of Weir and Cockerham (1984), which uses Theta as an estimate of F_{ST} . Theta is assumed to be a better estimate of F_{ST} than Wright (1978) estimate and does not make assumptions about number of populations, sample sizes, or heterozygote frequencies so that it is more appropriate for small data sets (Weir and Cockerham, 1984). F_{ST} values were calculated using the computer programme THETA (Ellis, 1994), which uses the individual genotypes as input. F_{ST} is a useful estimator of gene flow according to the equation $N_e m \approx \frac{1}{4} \frac{1}{F_{ST}} - 1$ (Wright, 1951). Since there is an inverse relationship between F_{ST} and $N_e m$ we consider F_{ST} values that do not differ significantly from zero to indicate unlimited gene flow, i.e., a panmictic populations structure.

Results

Habitat preference

Along the synthetic environmental axis (PCA-1) *C. australis* has a highly significant ($R^2 = 0.599$, $F_{1,33} = 49.20$, $P < 0.001$) positive relation to PCA-1 whereas *C. hyale* has a highly significant ($R^2 = 0.599$, $F_{1,33} = 49.24$, $P < 0.001$) negative relation (Fig. 1). Of the six environmental variables, vegetation and altitude were positively correlated with PCA-1, aspect and humidity negatively correlated, and date and weather uncorrelated. Humidity is thus probably the most important environmental determinant of both species since there is an obvious

relation (e.g., exposed ridges tend to be drier than valley bottoms) between humidity and the other three variables that had high loadings for PCA-1 (vegetation, altitude, and aspect).

Population analysis

Four of the loci (*ald*, *mdh1*, *ck2*, and *got1*) are monomorphically fixed for the same allele in the populations of both species. The remaining set of 12 allozyme loci did not contain loci that were diagnostic between *C. alfacariensis* and *C. hyale* (Appendix I). Four loci (*got2*, *mpi*, *pgi*, and *pgm*) show a very high degree of genetic polymorphism, but *got2* is monomorphic in all populations of *C. hyale*. An intermediate level of genetic polymorphism is found among the loci *hk*, *idh1*, and *me*. Low levels of genetic variability are present in the remaining loci (Appendix 1).

In general, *C. alfacariensis* tends to have both a higher mean number of alleles and percentage of polymorphic loci (Table 2). Consequently, *Colias alfacariensis* also has a significantly higher heterozygosity (range: 0.148–0.194) than *C. hyale* (range: 0.084–0.094) ($F_{1,11} = 72.70$, $P < 0.001$). Heterozygosity is highest for *C. alfacariensis* populations in the geographically complex Alps and drops off towards the lower-lying regions to the South (Provence and Italy) and the North (Metz and Seine), with the lowest values occurring in Northern Europe. Within *C. alfacariensis*, the Alpine/Italian on the one hand and the Provence/Northern populations on the other hand consistently differ at the PGI locus (Fig. 2).

In *C. alfacariensis*, estimated single locus F_{IS} significantly differs from zero at *me* (Table 3). In *C. hyale* the PGM locus significantly differs from zero. F_{IS} values were generally slightly positive in both species indicating a small deficiency of heterozygotes. In both species total genetic diversity (F_{IT}) was mainly the result of inbreeding (F_{IS}), among population variation (F_{ST}) being less important. Within populations of both species the most common allele at a locus is usually the same (Appendix 1). Within *C. hyale* none of the loci have a F_{ST} value significantly different from zero indicating that, at least until recently, there was a panmictic

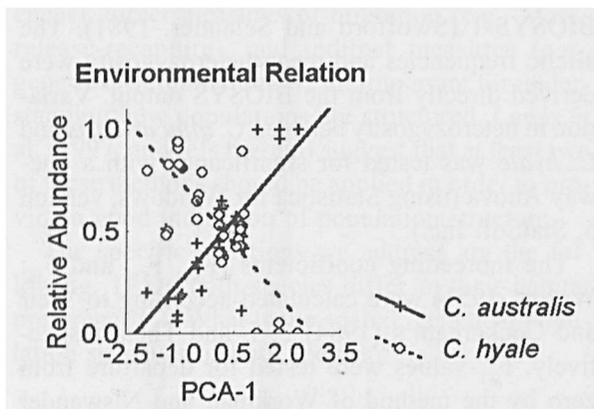


Fig. 1. Regression of *C. alfacariensis* and *C. hyale* onto the first Principal component axis (PCA-1). The slope of both lines is highly significant ($P < 0.001$).

Appendix 1: Allele frequencies for all populations (coded by letters A to M as in Table 1) are given at 12 polymorphic loci. Letters A to J are populations of *C. alfacariensis*, letters K to M are populations of *C. hyale* as in Table 1.

Locus		A	B	C	D	E	F	G	H	I	J	K	L	M	
<i>ck1</i>	<i>a</i>	-	-	0.026	0.056	-	-	0.036	-	-	-	0.022	-	-	
	<i>b</i>	1.000	1.000	0.974	0.944	1.000	1.000	0.964	1.000	1.000	1.000	0.978	1.000	1.000	
<i>got2</i>	<i>a</i>	-	-	-	-	0.023	-	-	-	-	-	-	-	-	
	<i>b</i>	0.079	0.063	-	-	0.045	0.024	0.036	-	-	-	-	-	-	
	<i>c</i>	0.263	0.250	0.211	0.333	0.318	0.429	0.357	0.500	0.547	0.292	-	-	-	
	<i>d</i>	0.632	0.688	0.711	0.667	0.545	0.548	0.607	0.500	0.453	0.583	1.000	1.000	1.000	
	<i>e</i>	0.026	-	0.079	-	0.068	-	-	-	-	-	0.125	-	-	-
<i>gpd</i>	<i>a</i>	0.026	-	-	-	-	-	-	0.071	0.016	-	0.022	-	-	
	<i>b</i>	0.974	1.000	1.000	1.000	1.000	1.000	1.000	0.929	0.984	1.000	0.978	1.000	1.000	
<i>hk</i>	<i>a</i>	0.921	0.688	0.895	0.889	0.977	1.000	0.893	0.929	0.828	0.917	-	-	-	
	<i>b</i>	0.079	0.313	0.105	0.111	0.023	-	0.107	0.071	0.172	0.083	1.000	1.000	1.000	
<i>idh1</i>	<i>a</i>	-	-	0.079	0.056	0.068	-	0.107	0.000	0.078	0.042	0.022	-	-	
	<i>b</i>	1.000	1.000	0.921	0.944	0.909	1.000	0.857	1.000	0.906	0.958	0.978	1.000	0.975	
	<i>c</i>	-	-	-	-	0.023	-	0.036	-	0.016	-	-	-	0.025	
<i>idh2</i>	<i>a</i>	-	0.063	-	-	0.023	-	0.036	-	-	-	-	-	-	
	<i>b</i>	1.000	0.938	0.974	1.000	0.955	1.000	0.929	1.000	1.000	1.000	1.000	0.958	0.950	
	<i>c</i>	-	-	0.026	-	0.023	-	0.036	-	-	-	-	0.042	0.050	
<i>mdh2</i>	<i>a</i>	-	-	-	-	-	0.024	-	-	-	-	0.022	-	-	
	<i>b</i>	0.947	1.000	0.974	0.889	1.000	0.929	1.000	1.000	0.953	0.958	0.978	1.000	0.975	
	<i>c</i>	0.053	-	0.026	0.111	-	0.048	-	-	0.047	0.042	-	-	0.025	
<i>me</i>	<i>a</i>	-	-	0.053	-	-	-	-	-	-	-	-	0.042	0.025	
	<i>b</i>	0.921	0.938	0.789	0.889	0.841	0.905	0.857	0.929	0.953	0.958	1.000	0.958	0.975	
	<i>c</i>	0.079	0.063	0.158	0.111	0.159	0.095	0.143	0.071	0.047	0.042	-	-	-	
<i>mpi</i>	<i>a</i>	-	-	0.079	0.111	0.023	0.048	0.036	-	-	-	-	-	-	
	<i>b</i>	0.289	0.375	0.342	0.389	0.455	0.190	0.429	0.214	0.328	0.542	0.109	0.167	0.050	
	<i>c</i>	0.026	-	-	-	-	-	-	0.143	-	-	-	-	-	
	<i>d</i>	0.605	0.563	0.500	0.389	0.409	0.357	0.429	0.571	0.672	0.333	0.826	0.750	0.750	
	<i>e</i>	0.053	0.063	0.079	0.056	0.114	0.286	0.071	0.071	-	0.083	0.065	0.083	0.150	
	<i>f</i>	0.026	-	-	0.056	-	0.119	0.036	-	-	0.042	-	-	-	0.050
	<i>g</i>	0.026	-	-	-	-	-	0.036	-	-	-	-	-	-	
<i>6pgdh</i>	<i>a</i>	0.026	-	-	-	-	-	0.036	-	-	-	-	-	-	
	<i>b</i>	0.974	1.000	1.000	1.000	1.000	1.000	0.964	1.000	1.000	1.000	0.978	1.000	1.000	
	<i>c</i>	-	-	-	-	-	-	-	-	-	-	0.022	-	-	
<i>pgi</i>	<i>a</i>	0.132	0.063	0.000	0.000	0.091	0.143	0.036	0.143	0.109	0.125	0.022	0.000	0.025	
	<i>b</i>	0.263	0.125	0.105	0.222	0.591	0.643	0.536	0.571	0.656	0.708	0.457	0.417	0.475	
	<i>c</i>	0.316	0.375	0.553	0.444	0.318	0.190	0.393	0.286	0.234	0.167	0.500	0.458	0.450	
	<i>d</i>	0.289	0.375	0.342	0.333	-	0.024	0.036	-	-	-	0.022	0.125	0.050	
	<i>e</i>	-	0.063	-	-	-	-	-	-	-	-	-	-	-	
<i>pgm</i>	<i>a</i>	0.053	-	0.026	-	0.023	0.024	-	0.071	0.016	-	0.022	-	0.025	
	<i>b</i>	-	-	-	-	-	-	-	-	0.016	-	-	-	-	
	<i>c</i>	0.053	0.125	0.079	0.056	0.045	-	0.036	0.071	-	0.042	-	0.042	0.100	
	<i>d</i>	0.421	0.188	0.395	0.333	0.386	0.238	0.321	0.143	0.313	0.458	0.217	0.208	0.275	
	<i>e</i>	-	-	-	-	0.045	-	-	-	0.047	-	-	-	-	
	<i>f</i>	0.421	0.688	0.500	0.611	0.477	0.667	0.643	0.714	0.547	0.500	0.717	0.708	0.575	
	<i>g</i>	0.053	-	-	-	0.023	0.071	-	-	0.063	-	0.043	0.042	0.025	

population structure with high levels of gene flow. Over the same geographic range, *C. alfacariensis* has an estimated $N_e m$ of less than 10. The difference here though is solely due to the PGI locus, which effectively divides the species into a North-

ern-Provence and Alpen-Italian component. Excluding this locus gives a panmictic population structure, since none of the other loci differ significantly from zero, quite similar to its sister species.

Table 2. Population genetics summary statistics at 12 loci in all populations (standard errors in parentheses).

Species	Population	Sample size per Locus	Mean no. of alleles per locus	Percentage of loci polymorphic*	Mean heterozygosity Direct-count
<i>C. australis</i>	A	19	2.2 ± 0.4	56.3	0.17 ± 0.062
<i>C. australis</i>	B	8	1.8 ± 0.3	43.8	0.19 ± 0.073
<i>C. australis</i>	C	19	2.1 ± 0.3	62.5	0.19 ± 0.061
<i>C. australis</i>	D	9	1.9 ± 0.3	56.3	0.19 ± 0.059
<i>C. australis</i>	E	22	2.3 ± 0.4	50.0	0.17 ± 0.057
<i>C. australis</i>	F	21	1.9 ± 0.3	37.5	0.16 ± 0.066
<i>C. australis</i>	G	14	2.2 ± 0.3	62.5	0.17 ± 0.052
<i>C. australis</i>	H	7	1.8 ± 0.3	43.8	0.2 ± 0.077
<i>C. australis</i>	I	32	1.9 ± 0.3	56.3	0.15 ± 0.049
<i>C. australis</i>	J	12	1.8 ± 0.2	50.0	0.16 ± 0.058
<i>C. australis</i>	Mean		2.0	51.9	0.175
<i>C. hyale</i>	K	23	1.8 ± 0.3	50.0	0.08 ± 0.04
<i>C. hyale</i>	L	12	1.6 ± 0.2	31.3	0.09 ± 0.046
<i>C. hyale</i>	M	20	1.9 ± 0.3	43.8	0.08 ± 0.039
<i>C. hyale</i>	Mean		1.8	41.7	0.087

* A locus is considered polymorphic if the frequency of the most common allele does not exceed 0.990

Table 3. Estimates of Wrights F-statistics (F_{IS} , F_{ST} , and F_{IT}) at 12 allozyme loci for *C. alfacariensis* and *C. hyale*.

Locus	<i>Colias australis</i>			<i>Colias hyale</i>		
	F_{ST}	F_{IT}	F_{IS}	F_{ST}	F_{IT}	F_{IS}
<i>ck1</i>	0.003	-0.006	-0.009	-0.009	-0.003	0.006
<i>got2</i>	0.018	0.015	-0.003	-	-	-
<i>gpd</i>	0.002	-0.006	-0.008	-0.009	-0.003	0.006
<i>hk</i>	0.035	0.175	0.145	-	-	-
<i>idh1</i>	0.010	-0.049	-0.059	-0.013	-0.009	0.004
<i>idh2</i>	0.004	-0.010	-0.014	0.004	-0.017	-0.022
<i>mdh2</i>	0.007	0.140	0.146	-0.013	-0.009	0.004
<i>me</i>	0.007	0.219	0.225***	-0.004	-0.011	-0.006
<i>mpi</i>	0.027	0.176	0.153	-0.014	0.265	0.275
<i>6pgdh</i>	0.003	-0.003	0.000	-0.009	-0.003	0.006
<i>pgi</i>	0.120***	0.122	0.003	-0.022	0.007	0.029
<i>pgm</i>	0.000	0.088	0.088	-0.016	0.350	0.360**
Global	0.037	0.107	0.073	-0.016	0.167	0.180
Average	0.029 (0.011)	0.098 (0.022)	0.071	-0.023 (0.008)	0.158 (0.050)	0.177 (3.862)

Significant departures from zero: **; $P < 0.01$; *** $P < 0.001$. Average and (standard deviation) obtained by a bootstrap procedure over samples (1000 repeats).

Discussion

Colias alfacariensis and *C. hyale* both have a pronounced habitat preference, with *C. hyale* found on moist cultivated areas and *C. alfacariensis* found on dry uncultivated areas. *Colias hyale* has a consistently lower level of heterozygosity than *C. alfa-*

cariensis, has fewer alleles per locus, and most populations are monomorphic at loci that Geiger (1982) has determined as being usually highly polymorphic in the Pieridae, such as *got2*. The fact that no diagnostic loci were found between *C. alfacariensis* and *C. hyale* is not uncommon in Lepidoptera. For instance, Aubert et al. (1997) also found

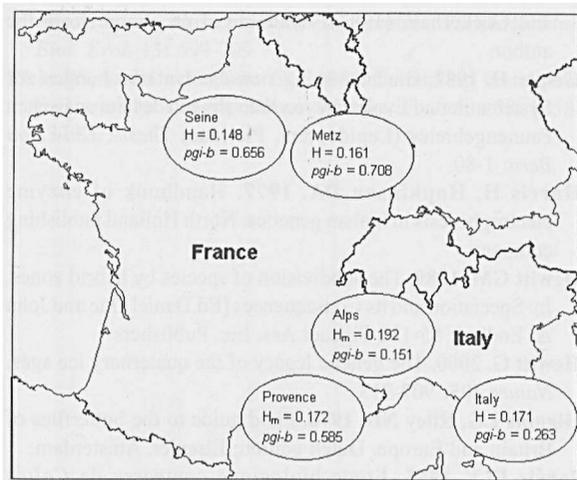


Fig. 2. Map of the localities where the populations of *C. alfacariensis* were sampled. Five main localities are shown, representing a single population with the exception of the Alps and Provence where three and four populations, respectively, were sampled. Within each circle average heterozygosity of 12 loci and the frequency of the *pgi-b* allele are given. For the Provence and Alps the values are means of the three populations. H = heterozygosity of a single population, H_m = mean heterozygosity from more than one population.

no diagnostic loci between the sister species *Papilio machaon* and *P. hospiton*.

Both *Colias* species, but especially *C. hyale* are known to be migrants (Higgins and Riley, 1970). This explains why both species maintain a panmictic population structure over a large geographical area with little genetic differentiation between the populations. This is also the case with the Monarch butterfly (*Danaus plexippus*) (Eanes and Koehn, 1978) famed for its migratory ability. Other species of North American *Colias* also show a high level of movements between patches of their habitat (Watt et al., 1977).

The loss of genetic variation in *C. hyale* and its preference for high humidity conditions suggest that it may represent one of the many examples of northern split-off species that evolved from species that were able to survive the periods of glaciation in ice-free areas of the Mediterranean (Hewitt, 2000). Such populations were subsequently able to colonise the area left behind after the retreating ice sheet.

Split-offs usually result in reduced genetic variation coupled with independent evolution to sub-

specific or specific status (Marchi et al., 1996). Populations at the northern edge of their range would be able to expand, usually via long-distance dispersal and rapidly colonize the newly available habitat allowing the best-adapted hybrid genomes to spread. These repeated colonising and founding events would eventually lead to a loss of alleles and thus to homozygosity. Many European and North American species show a lower genetic diversity in their northern populations (Hewitt, 2000). The same holds for other northern species, for example, the Dall sheep, *Ovis dalli*, only has a heterozygosity of 0.015 compared to an average of 0.041 for 184 mammalian taxa Hewitt (1989). Exactly such a pattern in levels of genetic variability can be found between *C. alfacariensis* and *C. hyale*. Not only does *C. hyale* have a lower genetic diversity than *C. alfacariensis*, but also Northern populations of *C. alfacariensis* have lower levels of heterozygosity than southern populations. The loss of heterozygosity in this and other Northern species is argued to be due to repeated fluctuation of ice fronts, which will cause a series of shifts in the periglacial range of the species involved. Range changes and consequently changes in population size provide evidence for a general loss of heterozygosity in northern species compared to southern species. This can further be predicted for subdivided species with extensive North-South ranges, depending on the life history and population dynamics of the species involved.

For the southern populations, the temporal predictability of a suitable habitat patch should decrease the rate of extinction, making the colonisation of other suitable habitat patches unnecessary for the continued survival of the species in a particular area. Such decreases in extinction probably would be associated with lower gene flow (Liebherr, 1988).

Selection at *pgi* has been demonstrated in a number of *Colias* species and PGI allozyme polymorphisms occur in parallel among many *Colias* species (Watt et al., 1996). In different American *Colias* complexes lowland and alpine groups have genotypes that display major functional differences, related to energy, allowing predictions about flight performance and fitness consequences which have been tested and confirmed (Watt et al., 1996; Watt,

1992). PGI polymorphs of *Colias* taxa may have diverged from an ancestral polymorphism and have been maintained till today (Watt et al., 1996).

Males of kinetically favoured *pgi* genotypes show greater flight capacity, survivorship, and overall mating success. There is also active female choice in selecting these favoured genotypes, where favoured males have an even greater advantage in mating older, more discriminating females than they do in mating previously unmated less discriminating females (Watt et al., 1986). A similar selection might be occurring in *C. alfacariensis*, but interestingly enough not in its sister species *C. hyale*. As in the American *Colias* species the main difference in the PGI locus among *C. alfacariensis* populations is between the alpine populations of France and the high elevation population in Italy and the low-lying populations in the Provence and Northern France.

Acknowledgements

We wish to thank Leon Rajimann and Wil van Ginkel for help in the lab. Barbra Schütt and Filippo Fabiani provided help with field collections.

References

- Aubert J, Barascud B, Descimon H, Michel F. 1997. Ecology and genetics of interspecific hybridization in the swallowtails, *Papilio hospiton* Gén  and *P. machon* L., in Corsica (Lepidoptera: Papilionidae). *Biological Journal of the Linnean Society* 60: 467-492
- Berger LA, Fontaine M. 1947. Une esp ce meconnue du genre *Colias* F. *Lambillionea* XLVII 11-12: 91-111
- Brunton CFA. 1998. The evolution of ultraviolet patterns in European *Colias* butterflies (Lepidoptera, Pieridae): a phylogeny using mitochondrial DNA. *Heredity* 80: 611-616
- Descimon H. 1994(5). *Colias hyale* (L.): une esp ce en voie de disparition? *Alexanor* 18: 440-442
- Dutreix C. 1980.  tude des deux esp ces affines *Colias hyale* Linn  et *Colias alfacariensis* Verity [Lepidoptera, Pieridae]. *Alexanor* 11: 297-316
- Eanes WF, Koehn RK. 1978. An analysis of genetic structure in the monarch butterfly, *Danaus plexippus* L. *Evolution* 32: 784-797
- Ellis WN 1994. Theta version 2.2. A program for the Macintosh computer for the analysis of population structure and estimation of gene flow levels based on the methods of Weir and Cockerham (1984). Available at no charge from the author.
- Geiger H. 1982. Biochemisch-genetische Untersuchungen zur Systematik und Evolution von Weisslingen des Europ ischen Faunengebietes (Lepidoptera, Pieridae) *Thesis, Univ. von Bern*: 1-80
- Harris H, Hopkinson DA. 1977. Handbook of enzyme electrophoresis in human genetics. North Holland publishing company
- Hewitt GM. 1989. The subdivision of species by hybrid zones. In: Speciation and its Consequences [Ed. Daniel Otte and John A. Endler]: 85-110 Sinauer Ass. Inc. Publishers
- Hewitt G. 2000. The genetic legacy of the quaternary ice ages. *Nature* 405: 907-913
- Higgins LG, Riley ND. 1970. Field guide to the butterflies of Britain and Europe. Dutch edition: Elsevier, Amsterdam.
- Jarvis FLV. 1955. Etude biologique comparee de *Colias alfacariensis* (Verity) et de *Colias hyale* (L.) *Lambillionea* LV, 3-4: 28-34 and 5-6: 37-50
- Lewis OT, Thomas CD, Hill JK, Brookes MI, Crane TPR, Graneau YA, Mallet JLB, Rose OC. 1997. Three ways of assessing metapopulation structure in the butterfly *Plebejus argus*. *Ecological Entomology* 22: 283-293
- Liebherr JK. 1988. Gene flow in ground beetles (Coleoptera: Carabidae) of differing habitat preference and flight-wing development. *Evolution* 42: 129-137
- Marchi A, Addis Exposito Hermosa V, and Crnjar R. 1996. Genetic divergence and evolution of *Polyommatus coridon gennargentii* (Lepidoptera: Lycaenidae) in Sardinia. *Heredity* 77: 16-22
- Menken SBJ, Herrebut WM, Wiebes JT. 1992. Small ermine moths (*Yponomeuta*): their host relations and evolution. *Annu. Rev. Entomol.* 37: 41-66.
- Peterson AT, Sober n J, S nchez-Cordero V. 1999. Conservation of ecological niches in evolutionary time. *Science* 285: 1265-1267
- Rajimann LEL, VanGinkel WE, Heckel DG, Menken SBJ. (1997). Inheritance and linkage of isozymes in *Yponomeuta padellus* (Lepidoptera, yponomeutidae). *Heredity* 78: 645-654
- Ritchie MG, Phillips SDF. 1998. The genetics of sexuals isolation. In: Endless Forms: species and speciation. Oxford University Press. P 470: 291-308
- Statistica for Windows (version 5, StatSoft Inc.).
- Tilman D. 2000 Causes, consequences and ethics of biodiversity. *Nature* 405: 208-211
- Watt WB, Chew FS, Snyder LRG, Watt AG, Rothschild DE. 1977. Population structure of pierid butterflies. I. Numbers and movements of some montane *Colias* species. *Oecologia* 27:1-27.
- Watt WB, Carter PA, Donahue K. 1986. Females' choice of "good genotypes" as mates is promoted by an insect mating system. *Science* 233: 1187-1190
- Watt WB. 1992. Eggs, enzymes, and evolution: natural genetic variants change insect fecundity. *PNAS* 89: 10608-10612
- Watt WB, Donohue K, and Carter PA. 1996. Adaptation at specific loci. VI. Divergence vs. parallelism of polymorphic allozymes in molecular function and fitness-component

effects among *Colias* species (Lepidoptera, Pieridae). *Mol. Biol. Evol.* **13**: 699-709

Weir BS, and Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

Wright S. 1951. The genetical structure of populations. *Annals of Eugenetics* **15**: 323-354

Accepted: 4 September 2002