

Disequilibrium in some *Cepaea* populations

LM Cook

The Manchester Museum, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Phenotypic disequilibrium is a measure available, along with morph frequency, in many published sets of sample data recording the polymorphism of shell colour and pattern in the snail *Cepaea nemoralis*. The relation of disequilibrium to morph frequency for the colour (pink/yellow) and banding (unbanded/banded) loci has been examined for a large and widespread set of data. The direction of disequilibrium is a function of frequency at the two loci in a way that suggests that selection favours combinations of common morphs,

whichever they are. This could indicate that such combinations are common because they have been selected. The data are therefore consistent with the proposal that populations are generally subject to selection of varying directions at different places and times, acting on the phenotype. In combination with migration, such selection could lead to prolonged polymorphism.

Heredity (2005) 94, 497–500. doi:10.1038/sj.hdy.6800645
Published online 2 March 2005

Keywords: *Cepaea*; polymorphism; directional selection; disequilibrium

Introduction

The snail *Cepaea nemoralis* (L.) has shell colour and pattern polymorphism determined by a few genetic loci. The species is widespread and often abundant in a broad range of western European habitats. There are general morph frequency changes between different climates and habitats, but almost every colony shows some polymorphism. The origin and maintenance of this variability is still unclear. Reviews of different aspects of ecology and population genetics are given by Lamotte (1951), Cain and Sheppard (1954), Jones *et al* (1977), Clarke *et al* (1978), Cain (1983), Goodhart (1987) and Cook (1998). Differences between colonies have been ascribed to random processes (early authors) or primarily to selection (later authors), but the almost universal polymorphism seems to imply a polymorphism-protecting mechanism. The main proposals are that it is (a) intrinsic heterozygote advantage that does not rely on local selective conditions, (b) frequency-dependent selection by predators favouring rare forms, (c) balance of mutation and random drift, and (d) interplay of randomly fluctuating selection and migration.

None of these hypotheses has won universal acceptance or is clearly the preferred explanation. There is circumstantial evidence for selection, but few definite instances. It is therefore worth continuing to look for evidence which might favour one or other interpretation. A large body of published data on morph frequencies in natural populations is available for further analysis. The degree of phenotype linkage disequilibrium in *Cepaea* has been discussed in particular contexts (eg by Cain and Sheppard, 1954; Arnold, 1968; Cain and Currey, 1968), but not in general over a range of conditions. It is

interesting to examine what patterns of disequilibrium exist, and what light they may be able to shed on the polymorphism.

Genetics of the polymorphism

Shell colour in *Cepaea nemoralis* is determined by five linked and several independent loci, each having 2–4 alleles exhibiting dominance (Murray, 1975). The most important in the linked group are the ground colour and banding loci. Banding segregates for unbanded and five banded, unbanded being dominant. Ground colour may be brown (top dominant), pink, or yellow (bottom recessive). Distinct alleles within each of these categories produce shades of colour, but they are rarely distinguished. Brown individuals, if they are dark brown, are rarely banded, and there is usually assumed to be metabolic incompatibility between the two conditions that prevents expression of this phenotype. Pink and yellow, on the other hand, exhibit the full range of banding frequency in different populations. Linkage is usually close, although there are two reports of a 20% crossover value. Excluding them, the best estimate of recombination rate is 0.2%, with an upper 95% confidence limit of 1% (Cook and King, 1966). Mid-banded is a common phenotype in which only the central band of the possible five is present. It is produced by a locus unlinked to the first two, with a dominant allele that suppresses four of the bands normally present in five-banded animals.

Almost all natural populations of *Cepaea nemoralis* are polymorphic for visible differences in phenotype. They often show disequilibria, the most common being between colour and the presence or absence of bands. The frequencies of the phenotypes pink unbanded (PU), yellow unbanded (yU), pink banded (Pb), and yellow banded (yb) may be represented as *a*, *b*, *c*, and *d*. The phenotypic disequilibrium is then measured by the determinant $D = ad - bc$, the four combinations being present proportionally when $D = 0$. The range of D is ± 0.25 when both loci have equal morph frequencies, dropping towards zero as morph frequencies become

Correspondence: LM Cook, The Manchester Museum, University of Manchester, Oxford Road, Manchester M13 9PL, UK.

E-mail: lcook@manchester.ac.uk

Received and accepted 21 December 2004; published online 2 March 2005

more extreme. The calculated values are often standardized to ± 1 by dividing by marginal totals. One alternative measure is the logarithm of the cross-product ratio, or $\log(ad) - \log(bc)$, and another is the correlation coefficient. All these measures have drawbacks under some circumstances, but are equally good if disequilibria are simply to be separated into positive and negative categories. For the purpose of discussion here, morph combinations in samples in which P and U are both greater than 0.5 or both less than 0.5 in frequency will be referred to as associated, and those in which P and b or y and U are greater than 0.5 in frequency will be referred to as disassociated. The set of available data for a sample can then be represented as the frequency of banded plotted on the frequency of yellow, divided into pairs of associated and disassociated quadrants, while a third axis represents the estimate of D .

Observations from natural populations

Published surveys allow us to examine the relation of disequilibrium to frequency in natural populations. Results have been used which give numbers of PU , yU , Pb , and yb . Those in which one or other class was missing were not used. This excludes smaller samples and ones that may have been founded with less than four classes. No other restriction of sample size has been imposed, on the assumption that sufficient data points will be obtained to make this unnecessary. Brown shells have been excluded, except for one study of Irish shells (Clarke *et al.*, 1968), which have very pale, and frequently banded, brown shells included with the pinks (pale browns are locally distributed and express banding). In one survey (Arnold, 1968), animals segregated for white lip as well as colour and banding. Lip colour is determined by a locus linked to ground colour and banding and shows disequilibrium with respect to them. The lip colour classes have been combined here. Numerous samples from the same homogeneous area could introduce bias, so repeat samples made at different times at the same site have been excluded.

Samples were grouped into the four equally sized morph frequency quadrants and the diagonal sets summed into associated and disassociated classes. The number showing positive and negative disequilibria in each class has then been determined. The signed square root of the corrected heterogeneity chi squared for the number of individuals of each morph was then calculated for each sample. This normal deviate is positive when $ad > bc$, otherwise negative. It indicates the tendency for the sign of the disequilibrium to change between associated and disassociated quadrants. In order to examine the degree of deviation within the associated and disassociated classes, these values have then been summed separately for each class, and an overall normal deviate calculated for each group by dividing the sum by the square root of the number of contributory samples.

A wide range of material has been examined from different habitats in Britain, Ireland, and France, some of the surveys covering wide areas. The most extensive are from sand dunes throughout the British Isles (Cain, 1968) and from various parts of France (Lamotte, 1951, 1959). Even where surveys are more intensive, they always involve areas much larger than the panmictic unit. Individual sets of data are therefore heterogeneous, so that it is reasonable to sum the data to provide a single ensemble.

Results

There were 512 samples in the 13 data sets examined. The mean sample size is 109.7, and 87% of samples have 30 or more individuals. Results are given in Tables 1 and 2. In Table 1, 10 of the 12 normal deviates are positive, indicating that positive disequilibrium is more common in associated and negative in disassociated combinations. This effect is significant when the separate sets are combined and for the overall totals (both $P < 0.001$). Likewise, in Table 2 the calculated normal deviates vary from one data set to another both in size and direction,

Table 1 Number of samples with frequencies in associated and disassociated phases classified by direction of disequilibrium

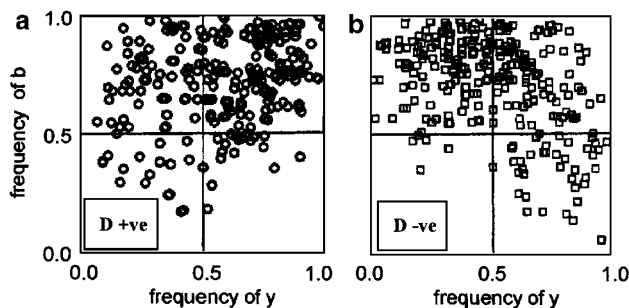
Source	Number of samples	D > 0		D < 0		Normal deviate
		a	d	a	d	
Arnold, 1968	79	25	3	29	22	2.71
Arnold, 1970	12	0	2	4	6	-1.10
Arnold, 1971	6	2	0	2	2	0.31
Cain, 1968	40	9	2	12	17	1.93
Cain and Currey, 1968	86	32	23	14	17	2.95
Cain and Sheppard, 1954	65	20	17	6	22	2.40
Cain <i>et al.</i> , 1990	6	0	0	0	6	—
Cameron, 1992	27	23	2	1	1	0.65
Carter, 1968	79	24	20	18	17	0.05
Clarke <i>et al.</i> , 1968	37	1	5	6	25	-0.15
Currey <i>et al.</i> , 1964	19	3	4	2	10	0.71
Lamotte, 1951, 1959	25	9	6	4	6	0.57
Murray, 1966	31	4	4	3	20	1.66
Trend						3.66
Total	512	152	88	101	171	5.83

a, Number in associated quadrants; d, number in disassociated quadrants. Normal deviates are obtained as the square root of corrected chi squared.

Table 2 Normal deviates measuring deviation of disequilibrium from zero in associated and disassociated quadrants

Source	No. of samples	Associated	No. of samples	Disassociated
Arnold, 1968	54	3.09	25	-5.66
Arnold, 1970	4	-2.65	8	-1.60
Arnold, 1971	4	0.16	2	-0.48
Cain, 1968	21	-0.36	19	-6.22
Cain and Currey, 1968	46	4.83	40	1.01
Cain and Sheppard, 1954	26	6.08	39	-0.17
Cain <i>et al.</i> , 1990	0	—	6	-14.54
Cameron, 1992	24	14.85	3	0.08
Carter, 1968	42	6.79	37	0.98
Clarke <i>et al.</i> , 1968	7	-6.24	30	-8.49
Currey <i>et al.</i> , 1964	5	1.20	14	-1.77
Lamotte, 1951, 1959	13	10.61	12	-1.50
Murray, 1966	7	4.91	24	-6.12
Trend	253	12.49		-12.34
Total data combined		14.71	259	-10.76

$P < 0.05$ for deviates $\geq \pm 1.96$.

**Figure 1** Distribution of disequilibrium values for 512 samples of *Cepaea nemoralis* in relation to morph frequency of banded (b) and yellow (y). (a) Samples with positive determinants, (b) samples with negative determinants.

but with one exception they are greater for samples in associated than in disassociated phase. For the total data, the normal deviate for associated samples is 14.7 and that for disassociated -10.8. Samples in associated phase therefore tend to have phenotypic determinants which are positive, while samples in disassociated phase tend to have negative determinants. The distribution between quadrants of samples with positive and negative determinants is shown in Figure 1. The slope of the reduced major axis has been calculated for the data in the two parts. For the 242 samples with $D \geq 0$ (Figure 1a) it is 0.795, and for the 270 with $D < 0$ (Figure 1b) it is -0.842. These relations are clearly different from each other, and support the contention that samples with positive determinants and samples with negative determinants relate differently to the associated and disassociated axes.

Discussion

Data from a very wide range of locations and habitats have been examined for disequilibrium between yellow and pink morphs at the colour locus and banding and unbanded at the banding locus. The direction of

disequilibrium is clearly related to whether the morph frequency is in the associated phase (high pink, high unbanded or high yellow, high banded) or disassociated phase (high pink, high banded or high yellow, high unbanded).

One reason for colour/banding disequilibrium is likely to be that a colony was founded by a small number of immigrants in which the four forms were not present proportionally. The generation time in *Cepaea* is 2–3 years and with a recombination rate no higher than 1% centuries could sometimes pass before the proportions were close to equilibrium. Levels of disequilibrium are more extreme between colour and banding than between the unlinked mid-banded and colour. This is consistent with accidental composition. However, for a large number of colonies disequilibria should be random among colonists, other things being equal, so that positive and negative values are equally likely, and if populations exist for a long time with little or no selection they should be near equilibrium.

If any one of the frequencies a , b , c and d becomes small enough to have a high chance of being zero, an extreme disequilibrium value will result. This is most likely in the smallest expected class, so that disequilibrium is most likely to be negative when P and U have high frequency or when both have low frequency. It is most likely to be positive when the frequency of P is high and U low, or vice versa. Similarly, in small samples containing four classes, D tends to rise in the disassociation quadrants and fall in the association quadrants as a result of sampling error. Using only four-class samples may modify this conclusion to some extent, and selection will tend to mask drift effects.

In *Cepaea* an important category of selection to consider is frequency-dependent, resulting from the action of predators which tend to overlook rare forms. If all morphs are equally likely to be attacked, this will protect the polymorphism with an equilibrium at equal frequencies of morphs (differential detectability would modify the equilibrium). Selective values for each phenotype combination are then declining functions of frequency. Disequilibrium will tend to show negative values if the association combinations are at high frequency, because the functions are small, and positive values if disassociation combinations are at high frequency. The result is to deflect the D surface in the same direction as occurs due to sampling error. How easily such an effect could be detected in practice depends, of course, on the strength and pattern of selection. Cook and Pettitt (1998) tested some data on colour and mid-banding for evidence of frequency-dependent selection, with inconclusive results.

Sampling error and centripetal selection therefore tend to generate differences in the direction of disequilibrium between phases, but in a direction opposite to the one observed. The pattern observed would be generated by centrifugal selection (eg negative frequency-dependent selection), moving the samples away from an unstable equilibrium. Apart from the theoretical unlikelihood of such a systematic force, the very wide range of habitats and locations makes it improbable that a common type of selection operates in all of them.

It is possible, however, that the disequilibria indicate selection in the generation sampled or in parental generations, while the frequencies themselves reflect

past history. Certain combinations may be common because they have been favoured in the past. If this is so, it indicates directional selection occurring in different directions in different places. On its own, such a situation would lead to monomorphism for different combinations in different places. In combination with migration, and with medium-term fluctuation in the amount and direction of selection brought about by environmental change, it could retain polymorphism for long periods of time. It could act, for example, on the north-south cline in the frequency of yellow and banded, which has probably been generated by broad latitudinal differences in selection. The evidence is therefore consistent with the argument that the predominant forces determining the patterns we measure in field studies are fluctuating directional selection modified by migration, as proposed by Cook (1998). This is not to say that other factors do not operate, but suggests that they are less readily detectable and may play a smaller role.

Disequilibria tend to occur where there is clinal selection and gene flow (Barton, 1982; Mallet and Barton, 1989; Mallet *et al.*, 1990). In *Cepaea*, there are sometimes well-defined clinal changes between patches of different morph frequencies (eg Cain and Currey, 1968) but that is not true of all the types of situation represented by these data. Cook (1998) suggested that polymorphism might be widespread because selective pressures fluctuate in time and space and there is sufficient long-distance movement to maintain or reintroduce morphs where they have become rare. There is no doubt that *Cepaea* individuals usually migrate very little, so that if there is also a small amount of regular long-distance dispersal the overall pattern of migration visualized is one of extreme leptokurtosis. If this assumption is used, it may be that the Mallet and Barton models are relevant to the processes involved here.

References

- Arnold RW (1968). Studies on *Cepaea* VII. Climatic selection in *Cepaea nemoralis* (L.) in the Pyrenees. *Philos Trans R Soc London B* **253**: 549–593.
- Arnold RW (1970). A comparison of populations of the polymorphic land snail *Cepaea nemoralis* (L.) living in a lowland district in France with those in a similar district in England. *Genetics* **64**: 589–604.
- Arnold RW (1971). *Cepaea nemoralis* in the East Sussex South Downs and the nature of area effects. *Heredity* **26**: 277–298.
- Barton NH (1982). The structure of the hybrid zone in *Uroderma bilobatum* (Chiroptera: Phyllostomatidae). *Evolution* **36**: 863–866.
- Cain AJ (1968). Studies on *Cepaea* V. Sand-dune populations of *Cepaea nemoralis* (L.). *Philos Trans R Soc London B* **253**: 499–517.
- Cain AJ (1983). Ecology and ecogenetics of terrestrial molluscan populations. In: Russell-Hunter WD (ed) *The Mollusca*. 6. *Ecology*. Academic Press: New York. pp 597–647.
- Cain AJ, Cook LM, Currey JD (1990). Population size and morph frequency in a long-term study of *Cepaea nemoralis*. *Proc R Soc London B* **240**: 231–250.
- Cain AJ, Currey JD (1968). Area effects in *Cepaea*. *Philos Trans R Soc London B* **246**: 1–81.
- Cain AJ, Sheppard PM (1954). Natural selection in *Cepaea*. *Genetics* **39**: 89–116.
- Cameron RAD (1992). Change and stability in *Cepaea* populations over 25 years: a case of climatic selection. *Proc R Soc London B* **248**: 181–187.
- Carter MA (1968). Studies on *Cepaea* II. Area effects and visual selection in *Cepaea nemoralis* (L.) and *Cepaea hortensis*. *Philos Trans R Soc London B* **253**: 397–446.
- Clarke BC, Arthur W, Horsley DT, Parkin DT (1978). Genetic variation and natural selection in pulmonate snails. In: Fretter V, Peake J (eds) *The Pulmonates*. 2A. *Systematics, Evolution and Ecology*. Academic Press: New York. pp 220–270.
- Clarke B, Diver C, Murray J (1968). Studies on *Cepaea* VI. The spatial and temporal distributions of phenotypes in a colony of *Cepaea nemoralis* (L.). *Philos Trans R Soc London B* **253**: 519–548.
- Cook LM (1998). A two-stage model for *Cepaea* polymorphism. *Philos Trans R Soc London B* **353**: 1577–1593.
- Cook LM, King JMB (1966). Some data on the genetics of shell-character polymorphism in the snail *Arianta arbustorum*. *Genetics* **53**: 415–425.
- Cook LM, Pettitt CWA (1998). Morph frequencies in the snail *Cepaea nemoralis*: changes with time and their interpretation. *Biol J Linn Soc* **64**: 137–150.
- Currey JD, Arnold RW, Carter MA (1964). Further examples of variation of populations of *Cepaea nemoralis* with habitat. *Evolution* **18**: 111–117.
- Goodhart CB (1987). Why are some snails visibly polymorphic, others not? *Biol J Linn Soc* **31**: 35–57.
- Jones JS, Leith BH, Rawlings P (1977). Polymorphism in *Cepaea*: a problem with too many solutions? *Ann Rev Ecol Syst* **8**: 109–143.
- Lamotte M (1951). Recherches sur la structure génétique des populations naturelles de *Cepaea nemoralis* (L.). *Bull Biol France Suppl* **35**: 1–239.
- Lamotte M (1959). Polymorphism in natural populations of *Cepaea nemoralis*. *Cold Spring Harbor Symp Quant Biol* **24**: 65–84.
- Mallet J, Barton N, Gerardo Lamas M, Jose Santisteban C, Manuel Muerdas M, Eeley H (1990). Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics* **124**: 921–936.
- Mallet J, Barton NH (1989). Inference from clines stabilized by frequency-dependent selection. *Genetics* **122**: 967–976.
- Murray J (1966). *Cepaea nemoralis* in the Isles of Scilly. *Proc Malac Soc Lond* **37**: 167–181.
- Murray JJ (1975). The genetics of Mollusca. In: King RC (ed) *Handbook of Genetics*, Vol 3. Plenum: New York. pp 3–31.