Qualitative and quantitative methods show stability in patterns of *Cepaea nemoralis* shell polymorphism in the Pyrenees over five decades

Daniel Ramos-Gonzalez*[1] and Angus Davison[2]

School of Life Sciences, University Park, University of Nottingham, NG7 2RD, UK.

1. Email: stxdr3@exmail.nottingham.ac.uk, https://orcid.org/0000-0003-4497-6374

2. Email: angus.davison@nottingham.ac.uk, https://orcid.org/0000-0003-2554-8585

Abstract

One of the emerging strengths of working with the land snail genus *Cepaea* is that historical collections can be compared against modern day samples, for instance to understand the impact of changing climate and habitat upon shell morph frequencies. However, one potential limitation is that prior studies scored shell ground colour by eye, usually in the field, into three discrete colours yellow, pink or brown. This incurs both potential error and bias in comparative surveys. In this study, we therefore aimed to use a quantitative method to score shell colour, and evaluated it by comparing patterns of *C. nemoralis* shell colour polymorphism, using both methods on present day samples, and against historical data gathered in the 1960s using the traditional method. The Central Pyrenees were used as an exemplar, because previous intensive surveys sometimes showed sharp discontinuities of morph frequencies within and between valleys. Moreover, selective factors, such as climate or the human impact in the Pyrenees, have significantly changed since 1960s. The main finding was that while quantitative measures of shell colour reduced the possibility of error, and standardised the procedure, the same altitudinal trends were recovered, irrespective of the method. There was remarkable stability in the local shell patterns over five decades, with the exception of one valley that has been subject to increased human activity. Therefore, although subject to potential error, human-scoring of snail colour data remains valuable, especially if persons have appropriate training. In comparison, while there are benefits in taking quantitative measures of colour in the laboratory, there are also several practical disadvantages, mainly in terms of throughput and accessibility. In the future, we anticipate that both methods may be combined, for example, using automated measures of colour taken from photos generated by citizen scientists conducting field surveys.

Introduction

Historically, two of the most important species in studying colour polymorphism have been the west European land snails *Cepaea nemoralis* and *C. hortensis*, because individuals are relatively easy to collect and study, and the colour and banding morphs show straightforward inheritance (Cain & Sheppard, 1950, Cain & Sheppard, 1952, Cain & Sheppard, 1954, Lamotte, 1959, Jones et al., 1977). In the present day, one of the continuing benefits of working with *Cepaea* is an ability to compare the frequencies of shell morphs in historic collections against modern day samples, to infer the potential impact of natural selection and/or drift in changing shell morph frequencies. Of particular use, the "Evolution Megalab" project digitised a large set of 20th century samples (Silvertown et al., 2011, Cameron & Cook, 2012, Worthington et al., 2012). These records, and others deposited in museums, are now being used with modern surveys to produce an increasing number of comparative papers (Cowie & Jones, 1998, Silvertown et al., 2011, Cameron & Cook, 2012, Ożgo & Schilthuizen, 2012, Worthington et al., 2012, Cameron et al., 2013, Cook, 2014).

In nearly all comparative studies of *Cepaea* reported to date, absolute change in frequencies of the main shell morphs, colour and banding, have been reported, but the direction is not always consistent. The conclusions are in part dependent upon the geographic scale and the precision of resampling, whether exact or nearest neighbour. To fully understand changes − or stasis − in shell polymorphism, both global and local surveys are needed (Berjano et al., 2015). For instance, large-scale surveys illustrate the broad picture of the changes in the spatial variation of the polymorphism. In the largest study, a historical dataset of more than six thousand population samples of *C. nemoralis* was compared with new data on nearly three thousand populations (Silvertown et al., 2011). A historic geographic cline among habitats in the frequency of the yellow shells was shown to have persisted into the present day. However, there was also an unexpected decrease in the frequency of unbanded shells, and a corresponding increase in frequency of banded and mid-banded morph particularly (Silvertown et al., 2011). A UK-wide study also used Evolution Megalab data, but reported a somewhat different pattern of change. Yellow and mid-banded morphs had increased in woodland, whereas unbanded and mid-banded increased in hedgerow habitats (Cook, 2014).

In comparison to these large surveys, the majority of comparative studies have been at a more local scale. The benefit of these is that resampling is often precise (Cowie & Jones, 1998, Cook et al., 1999, Ożgo & Schilthuizen, 2012, Cameron et al., 2013, Ożgo et al., 2017), and it is also possible to take local factors into account. Most of the original historic studies took place in the UK. Following resampling, modern comparative surveys have tended to find an increase in yellow and mid-banded shells (as above) (Silvertown et al., 2011, Ożgo & Schilthuizen, 2012, Cameron et al., 2013, Ożgo et al., 2017), but with exceptions (Cowie & Jones, 1998, Cook et al., 1999, Cameron & Cook, 2012), depending upon the precise scale of comparison. Moreover, patterns of change are not always consistent within the same study.

One potential limitation of all of these works is that shell ground colour was scored by eye, usually in the field, into three discrete colours yellow, pink or brown. Even if persons are trained, there is still bias and error, and potential for dispute over what defines each colour. In practise, it is frequently difficult to distinguish the colours, and define different shades of the same colour. Therefore, to understand whether colour variation is in reality continuous, and to investigate how the variation may be perceived by an avian predator, psychophysical models of colour vision were applied to shell reflectance measures, finding that both achromatic and chromatic variation are continuously distributed over many perceptual units in indiscrete in *Cepaea nemoralis* (Davison et al., 2019). Nonetheless, clustering analysis based on the density of the distribution did reveal three groups, roughly corresponding to human-perceived yellow, pink and brown shells.

This prior work raises the possibility that reproducible, quantitative shell colour measures, based on spectrophotometry in the laboratory, can be used to compare and test regular shell colour data, avoiding the requirement to bin measures into colour categories. In this study, we therefore aimed 1) to use the quantitative method to score shell colour, and 2) evaluated it by comparing patterns of *C. nemoralis* shell colour polymorphism using both methods on present day samples, and against historical data gathered using the traditional method.

To achieve this aim, the Central Pyrenees were used as an exemplar location, because they were intensively surveyed during the 1960s and 70s (Figure 1), sometimes showing sharp discontinuities of frequencies of morphs (Arnold, 1968, Arnold, 1969, Cameron et al., 1973, Jones & Irving, 1975) and genotypes (Ochman et al., 1983) within and between. They are also particularly interesting for their geographic and ecological variation, including a diverse range of different microclimates, within and among the valleys, due to the interaction of three main climates, Atlantic, Mediterranean and Alpine, as well as a large altitudinal differences and incidence of precipitation. Moreover, selective factors, such as climate or the human impact in the Pyrenees, have significantly changed since the 1960s (García-Ruiz, 2015).

The main finding was that while quantitative measures of shell colour reduced the possibility of error, and standardised the procedure, the same altitudinal trends were recovered, irrespective of the method. There was remarkable a stability in the local shell patterns over five decades. Overall, while there are key benefits in taking quantitative measures of colour in the laboratory, there are also several practical disadvantages. In the future, with the increasing use of digital cameras to capture and record species presence, there is the potential that colour and banding data may be extracted from the images uploaded to public databases and apps such as iRecord, iNaturalist and SnailSnap (Harvey, 2018, Horn et al., 2018, Kerstes et al., 2019). For the moment, the fact remains that human-scoring of snail colour data is valuable, especially with appropriate training.

Materials and Methods

Shell samples and human-scoring of shell phenotype

The Valle de Vielha, Valle de Jueu, Valle Noguera de Tort and Valle Noguera Ribagorzana, hereafter abbreviated as “Vielha”, “Jueu”, “Tort” and “Riba”, were selected for sampling (Figure 1). This is because they had been previously sampled in 1962 by Arnold (1968), and in 1966 and 1969 by Cameron *et al.* (1973), with the colour and banding data made available via the Evolution Megalab database. New samples were collected in October 2017 and June 2018. By choice, we aimed to sample in the same location as described in past surveys, using the coordinates recorded in the Megalab database; when this was not possible, samples were collected from the nearest adjacent site with suitable habitat for snails.

Snail shell colour was qualitatively scored in the laboratory as either yellow, pink or brown, by DRG. Similarly, following previous convention, shells were scored as “unbanded” (00000), “mid-banded” (00300) or “banded” (all banding versions except mid-banded). These three categories were used in all subsequent analyses. As *C. nemoralis* in the Pyrenees is polymorphic for other characters, we also scored the lip colour, as either pale (usually white) or any other colour (usually black or dark brown), and measured the shell height (H) and width (W) using a Vernier calliper with 0.05 mm precision, then calculating the shape as H/W.

Quantification of shell colour

The ground colour of adult snail shells from Vielha and Jueu valley was measured using an Ocean Optics spectrometer (model USB2000+UV-VIS-ES) and a Xenon light source (DT-MINI-2-GS UV-VIS-NIR), as described previously (Davison et al., 2019). Briefly, the shell underside was used because it is generally unbanded and the least damaged/exposed to sunlight, holding the probe at a 45° incident angle, ~2 mm from the shell. Each sample was quantified three times, non-consecutively, recalibrating using light (WS-1) and dark standards after 2 to 5 quantifications, software was recalibrated by using light standards (Davison et al., 2019). Data was collected using Ocean Optics SpectraSuite 2.0.162, using an integration time of 750 msec, boscar width of 5, and scans to average 10. Reflectance spectra were analysed following a modified protocol described below (Delhey et al., 2014, Davison et al., 2019), using Pavo 2.2.0 R package to bin raw reflectance spectra (1 nm) (Maia et al., 2013, Maia et al., 2018), and then R version 3.4.1 (2017-06-30) for further analyses (Delhey et al., 2014).

In a previous analysis, the aim was to understand how an avian predator might perceive the shell colours, so the tetrachromatic colorimetric standards of a blackbird (*Turdus merula*) were used (Davison et al., 2019). In this new analysis, the main aim was to compare human qualitative scores of shell colour against quantitative scores, so as to better understand any biases. Reflectance spectra analysis were therefore analysed using human CIE colour trichromatic coordinates (Smith & Guild, 1931, Westland et al., 2012), as follows.

CIE standards are based on the stimulation of the different photoreceptors’ cells (cones) of the retina. In humans, three main groups of cones are found, L (long wavelength, peaking at 560 nm), M (medium wavelength, peaking at 530 nm), and S (short wavelength, peaking at 420 nm) (Hunt, 2004). The visual colour spectra (300-700 nm) were converted using the three chromatic coordinates of the visual space, xyz, where Euclidean distances between points reflect perceptual differences, generated from quantum catches for each photoreceptor (Cassey et al., 2008). The human trichromatic coordinates (xyz), determined from the tristimulus values (XYZ), were calculated by Pavo 2.2.0 R package, a colour spectral and spatial perceptual analysis, organization and visualization package, and the “standard daylight” (d65) irradiance spectrum (Smith & Guild, 1931, Maia et al., 2018). Then, a principal component analysis (PCA) was undertaken as described previously (Delhey et al., 2014, Scrucca et al., 2016, Davison et al., 2019).

Analysis of phenotype frequencies and correlation

To compare past and present-day datasets, the change in the frequencies of colour and banding traits for each sample site were calculated. To detect any overall trends in each valley, any differences were evaluated using independent paired T-student (parametric) or paired rank Wilcoxon Test (non-parametric), selected according to normality (Shapiro-Wilk normality test) and homogeneity (F-test).

Linear mixed regression models were conducted for colour and banding from past and present datasets. Outliers were removed following the interquartile range method, using a Shapiro-Wilk normality test to test for deviations from normality. The Pearson correlation (parametric) or Kendall rank correlation test (non-parametric) were performed to evaluate correlation and any significance with altitude. Kendall rank correlation coefficient "Tau" were transformed into Pearson "r" coefficient to evaluate correlation and to conduct Fishers' Z-transformation (Walker, 2003, Fisher, 1921). The correlation breached the assumption of normality required in the standard comparative test. Therefore, Fishers' Z-transformation was applied to calculate the significance of the difference between the past and current correlation coefficients against altitude.

Maps, plots and statistical tests were made using R version 3.4.1 (2017-06-30), the ggplot2 3.2.1 package for data visualization, and the ggmap 3.0.0 R package, to generate maps. Maps were acquired from the Geo-location APIs platform in Google maps source (https://console.cloud.google.com/apis/dashboard).

Results

Past and present-day geographic distribution of colour and banding morphs

Snails were mainly found in open areas such as hedgerows, scrubs, meadows and grass, and rare in woodlands. In high altitude areas, snails were discovered mostly on meadows or screes. In total, snails were collected from 138 sample sites ranging from 823 m to 1921 m above sea level. However, only 108 sites and 2633 individuals were used for the analysis, as we only considered sites with ten or more individuals collected (Table 1). Of the filtered 108 sites, 87 were judged to be the same as a previous study, based on previous coordinates, or up to 50 m distance away. In comparison, in the previous surveys, Arnold (in 1962) collected 5006 snails from 123 sites in the Vielha and Jueu valleys (Arnold, 1968). Cameron (in 1966 and 1969) sampled 2177 and 2145 snails from 48 and 55 sites located in Jueu, Ribagorzana and Tort respectively (Cameron et al., 1973). Therefore, a total of 226 historical sample sites and 9328 individuals were available for comparison (Table 1). Full details of all sample sites are in the supporting information (Tables S1, S2).

As in previous studies from the Pyrenees, the new survey showed that the pattern of shell morph distribution depends upon the specific valley, frequently showing associations with altitude (Figures 2, 3). Yellow and unbanded shells tended to predominate in the higher regions of the Vielha and Jueu valleys. In the intermediate or lower sites (below ~1200m), pink and yellow shells had similar frequencies, with most shells also having bands. In Ribagorzana yellow shells were commonly distributed in all sites, whereas pinks were usually found in the upper valley and brown morphs in the intermediate and lower valley. Brown populations were only found in the Ribagorzana and Tort valleys. In addition, unbanded morphs prevailed in Ribagorzana. In contrast, in the adjoining Tort valley, yellow predominated in all sites, with banded morphs predominant in almost the entire valley.

Spatial patterns of variation in morph frequencies were largely the same as recorded in the past, including colour and banding (Figures 2, 3) as well as lip-colour (Figure S1). To formally test this, directional changes in the frequencies of shell types at each location between the 1960s and the present-day were tested using independent paired T-student or paired rank Wilcoxon tests (Table 3; Table S3). This confirmed little overall change in the distribution of the main colour and banding types in Vielha, Jueu, and Tort (Table 3; Table S3; and Figure 4). The exception was in Ribagorzana valley, where the proportion of banded shells has risen from ~3% to 14%, with substantially fewer brown shells recorded and more yellow shells (Table 3).

The present-day relationship between altitude and frequency of colour and banding morphs was plotted (Figure 5). Jueu and Tort valleys showed a significant positive correlation between altitude and the frequency of yellows, with the former also showing a positive significant altitude-unbanded association (Figure 5; Table S4). As expected, pink and banded shells showed the reverse trend, but with non-significant altitudinal correlations; mid-banded shells did not show any correlation with altitudes. Tort showed a significant positive (but shallow) relationship between yellow-altitude and banded-altitude (Table S4, Figure 5, $r = 0.27$, 0.34 respectively and $p < 0.05$). There was also significant positive association of the white-lip morph with altitude in three valleys (Figure S2), in addition to associations of higher altitude with larger shell size (H + W), and relatively tall spires (H/W) (Figure S3).

Fishers' Z-transformation was used to test the significance of the difference between the past and present altitudinal correlation coefficients. There were no significant changes in Jueu, Ribagorzana and Tort (Table 2). In comparison, in the past sample from Vielha valley, both colour (Table S4, yellow shells $r = 0.48$, $p < 0.001$) and banding (Table S4, unbanded shells, $r = 0.51$ and banded shells, $r = -0.48$, $p < 0.001$) showed a moderate association with altitude. In the present-day, colour and banding did not show a significant correlation with altitude.

Unfortunately, it was not possible to make the same comparisons with lip-colour and shell measurements,because the former data was not uploaded to the Evolution Megalab database, and the size measures were not recorded in the original studies.

Quantitative measures of shell colour and banding and associations with altitude

The reflectance spectra of 813 shells from Vielha and Jueu valleys was measured, a subset of the total collected (2633; Table 1), because some shells were too damaged to record quantitative colour. A PCA on the xyz coordinates showed three axes which together explained 99% of the chromatic variation, PC1 51%, PC2 44%, and PC3 4%. As previously reported (Davison et al., 2019), the third axis, PC3, tended to separate pink and yellow shells (Figure S4). Therefore, to visualize the present-day relationship between altitude and quantitative chromatic variation, PC3 was used because all the individuals in Vielha and Jueu were yellow or pink (Figure 6). In Vielha, there was weak negative, but non-significant association, of altitude and PC3, whereas Jueu showed a moderate positive correlation. These indicates that in Vielha there was no association of shell colour with altitude, whereas in Jueu yellow shells were more common at high altitude.

Past and present-day associations, using qualitative and quantitative methods

We compared altitude-colour associations between historical and present-day samples from Vielha and Jueu, using the different methods.

For Jueu valley (Figure 7), the same significant altitudinal associations were recovered whether using historical data (n= 1862), the present-day data with human-scoring of colour (n=637), or quantitative measures of colour or pattern as manual scoring (n=206; Figure 7). Fishers' Z-transformation test showed no significant changes among the altitudinal correlations for each of these four graphs (Table 2).

For Vielha valley (Figure 7), there was a significant altitudinal association with colour only in the historical dataset (n = 4756, r = 0.48, p = 0.0001), compared with a non-significant positive relationship using the present-day data with human-scored colour (n = 942, r = 0.14, p = 0.355), and a non-significant negative relationship using quantitative measures of colour (n = 607, r = -0.09, p = 0.056). To further explore these differences, we also tested for a correlation using the present-day data with human-scored colour, but just using the subset of shells which were considered sufficiently undamaged for spectrophotometry (Figure 7 inset graph). This showed a negative relationship (r = -0.08, p = 0.588), likely indicating that some (old) pink shells were mistakenly scored as yellow in the qualitative analysis.

Discussion

Quantitative versus qualitative methods to score shell phenotype

In prior studies, the shell ground colour was scored by eye, sorting individuals into three discrete categories, either yellow, pink or brown. In this study, in addition to the human-scoring of shell colour, we evaluated a quantitative method, based on spectrophotometry in the laboratory, by comparing patterns of *C. nemoralis* shell colour polymorphism from the past and the present day. The main finding was that while spectrophotometry of shell colour has the benefit of being quantitative and is objective, the same trends were recovered. In fact, there was a remarkable stability in the local shell patterns in most valleys over five decades.

Both qualitative and quantitative methods have benefits and also disadvantages. Spectrophotometry produces a quantitative output for an individual shell, which better reflects the non-discrete nature of variation in snail shell colour, and is reproducible. However, it is only accessible to a few persons, requires expensive equipment, and ideally, that the reflectance measures are taken in the laboratory. All of these latter factors together reduce throughput. In comparison, field-based methods do not require the snails to be taken to a laboratory, are rapid and accessible to a wide range of persons, including citizen scientists. The disadvantage is that the shell colour phenotype must be binned into one of three subjective categories, with the snails from a sometimes ill-defined single location making a single data point. Moreover, the data that is collected must be carefully filtered (e.g. Silvertown et al., 2011) to remove misidentified species (especially confusion with *C. hortensis*, juvenile *Cornu aspersum* and *Arianta arbustorum*), a difficult task because the specimen is not preserved. Nonetheless, human-scoring of snail colour data remains valuable, especially with appropriate training.

In the future, we anticipate that a model that takes the best of both methods may be used instead. Websites and apps such as SnailSnap, iNaturalist and iRecord (Harvey, 2018, Horn et al., 2018, Kerstes et al., 2019) are already being used extensively by the general public to capture records and images of snails, which are then identified using a combination of machine-learning methods and input from persons with various degrees of expertise. For example, iNaturalist has over 9000 observations, including photos, of *C. nemoralis* at “research grade” quality (including >1000 in the UK, but only 29 in the Pyrenean region). One suggestion is that it would be relatively straightforward to extend the use of a machine-learning based method to inspect individual images, and then record the colour and the band category. A more sophisticated (but difficult to implement) alternative would be to extract quantitative colour data from the images, but this would have to be robust to the wide variety of circumstances under which the photos were taken; likely including some sort of colour control (e.g. a card; van den Berg et al., 2020) would limit the number of participants.

Past and present-day geographic distribution of colour and banding morphs

By analysing the geographical and altitudinal distribution of colour and banding attributes in the Central Pyrenees and comparing with previous studies, we aimed to understand how local factors, human impact and the rapid climate change acted upon the variation of *C. nemoralis* shell polymorphism.

Broadly, we found a remarkable stability in the local shell patterns in most valleys over five decades, despite large changes in habitat, human impact and a rapid climate changes over five decades. Most valleys still showed visibly similar patterns of shell types, whether colour, banding, lip colour or shell-shape (Figures 2, 3, S1), concordant with another study over the wider Pyrenean region (Ellis, 2004).

There were just a few exceptions to the general pattern. For instance, the altitudinal cline in the frequency of yellows that was present in both Vielha and Jueu valleys is now only present in the latter valley. The present-day absence of a clinal relationship is striking, and contrasts with the paired comparisons at each location, which did not show any overall change in the frequency of yellow or pink in Vielha over the decades (Figure 4). The explanation for the discrepancy (Table S3) is that while pinks have increased in frequency at higher altitudes in Vielha, they have also decreased in frequency at lower altitudes. Vielha is interesting because the establishment of Baqueira-Beret ski resort (now the largest in Spain) has led to an increase of human activity and the construction of infrastructure such as dams, tunnels or mines, with a corresponding growth of urban areas in the adjoining tributary valleys. In comparison, the Jueu valley has remained largely intact, perhaps because it is a protected reserve. The loss of altitudinal-colour variation in this valley is therefore likely explained by the accidental movement of individuals and changing local habitat.

The only other location that showed change was in the Ribagorzana valley, where the proportion of banded shells has risen from ~3% to 14%, with substantially fewer brown shells recorded and more yellow shells. The explanation for changes in this valley are less clear. One possibility is that we were more likely to score an intermediate shell as pink rather than brown compared with previous workers. However, this can probably be discounted because the lower proportion of recorded brown shells in our samples from Ribagorzana is matched by an increased proportion of yellow rather than pink shells. The general finding of reduced browns is perhaps in line with other studies. Cowie and Jones (1998) and Cook et al. (1999) documented an overall decrease in the frequency of the brown shells, Ożgo and Schilthuizen (2012) identified that brown shells decreased in expenses of yellow shells, Cameron et al. (2013) reported a general increase of yellows and Cook (2014) found an increase of yellows in woodland habitats.

From phenotype to genotype

One limitation of comparative studies on *Cepaea* is that there is a risk that we ascribe "just-so" explanations to changes in the frequencies of a particular phenotype over time. For example, in this study, we have concluded that the changes that we observed in Vielha valley are due to immigration of new individuals (because of construction), but of course it is not possible to discount natural selection, especially because of changed habitat associated with the construction industry. The corollary is that we also lack understanding or explanation for circumstances when phenotype frequencies remain stable. These questions are perhaps best-addressed with manipulative experiments.

Recent progress in genomic technologies will certainly offer a solution, including the availability of a first draft *C. nemoralis* genome (Saenko et al., 2020). For example, it should be possible to use genomics to understand the relative roles of migration/founder effect and selection in determining the population structure of *Cepaea* populations. In particular, genomics may be used to understand the history of a population e.g. is there evidence for recent immigration to the high altitude regions of the Vielha valley, from snails that perhaps originate from elsewhere? Alternatively, is there evidence for a selective sweep at the loci that control the shell phenotype, perhaps indicative of a local response to a change in the selective regime?

Some of the other remaining issues, that we have only touched upon here, are the correlations between altitude and multiple phenotypic traits (banding, colour, lip colour, size, shape), as well as both linkage and linkage disequilibrium between the genes involved (Gonzalez et al., 2019, Cook, 2013). Given that lip colour is ordinarily a dark colour in *C. nemoralis* across most of Europe (with some exceptions), and that this is the main character that distinguishes this species from *C. hortensis*, the wide variation in this character in the

Pyrenees is particular mysterious. In the future, we hope to understand the genetic basis for these characters; it is hoped that this will bring forth an era in which we are better able to understand the impact of the multiple factors (Jones et al., 1977), including natural selection and random genetic drift, that determine the patterns of shell types that are present in nature.

Supplementary materials

Supplementary Figures S1 to S5 and Tables S1 to S5.

Acknowledgments: This work was supported by the University of Nottingham; the Biotechnology and Biological Sciences Research Council [grant number BB/M008770/1], via a studentship to Daniel Ramos Gonzalez. Thanks to Hannah Jackson and Alejandro Garcia Alvarez for helping in the sampling collection in the Pyrenees; to Jonathan Silvertown and the Evolution Megalab team who provided the historical data used in this research, MaŁgorzata Ożgo and Robert Cameron who provided helpful comments on the manuscript, and Sophie Poole who helped with some of the shell colour measurements.

Author Contributions: The work was jointly conceived and carried out by DRG and AD. The data were mainly analysed by DRG, with input from AD. Both authors drafted and wrote the manuscript.

Conflict of Interest statement: The authors have no conflict of interest to declare.

Figure 1. Overview of sampling locations in the Pyrenees, including this work, and previous work by others in the 1960s (Arnold, 1968, Cameron et al., 1973).

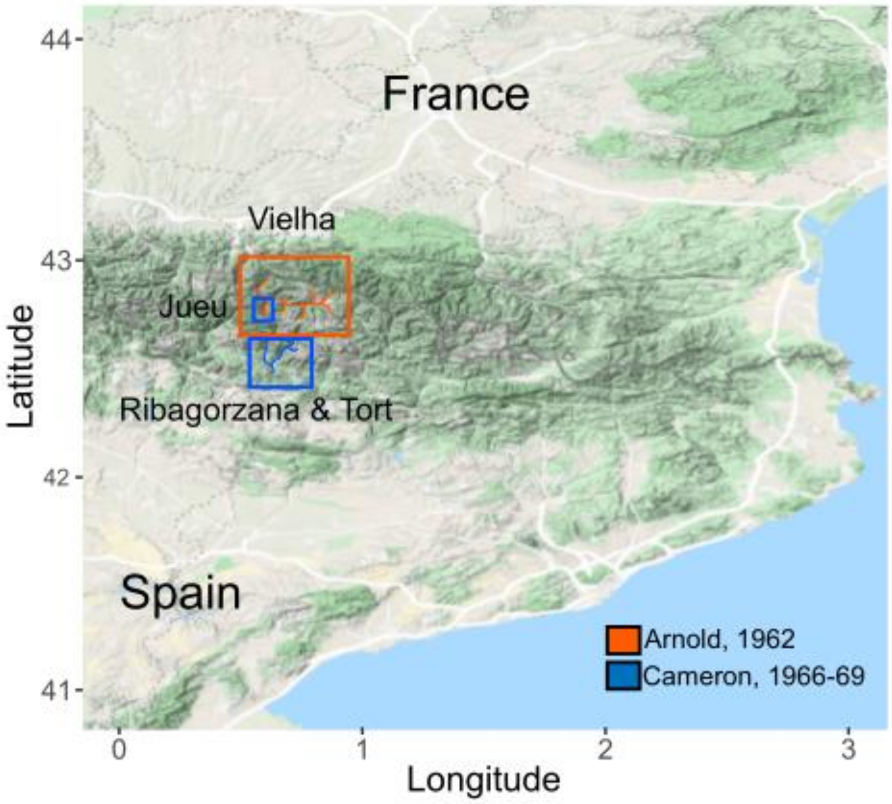

Figure 2. Past and present distribution of yellow, pink and brown shell morphs in Pyrenean valleys, based on sampling in the 1960s and 2017/18. Pie charts show frequencies of yellow (yellow), pink (pink) and brown (brown) morphs in each location. Valle Noguera de Tort is the left valley and Valle Noguera Ribagorzana is the right valley.

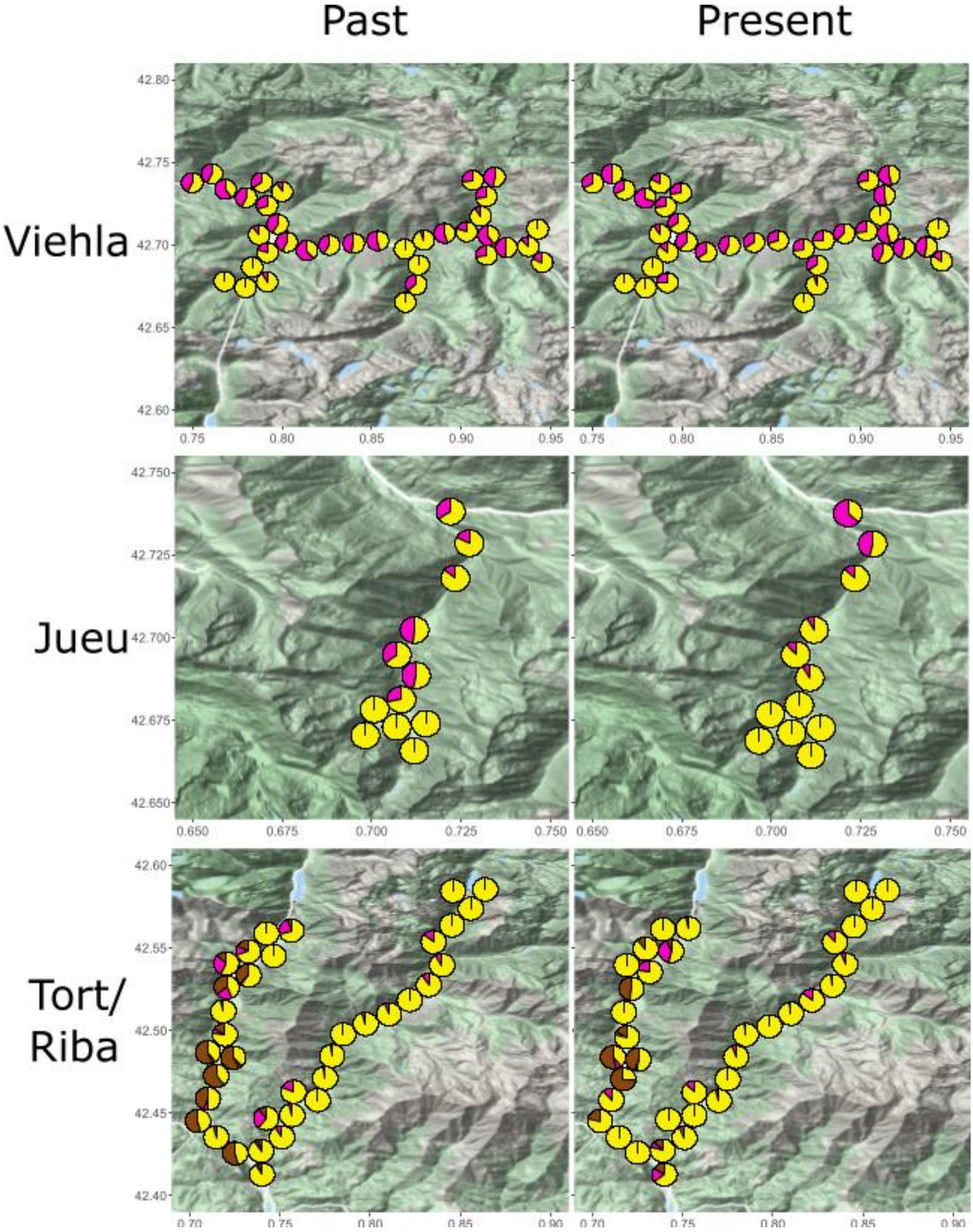

Figure 3. Past and present distribution of banded, mid-banded and unbanded shell morphs in four Pyrenean valleys, based on sampling in the 1960s and 2017/18. Pie charts show frequencies of banded (green), mid-banded (red) and unbanded (yellow) morphs in each location. Valle Noguera de Tort is the left valley and Valle Noguera Ribagorzana is the right valley.

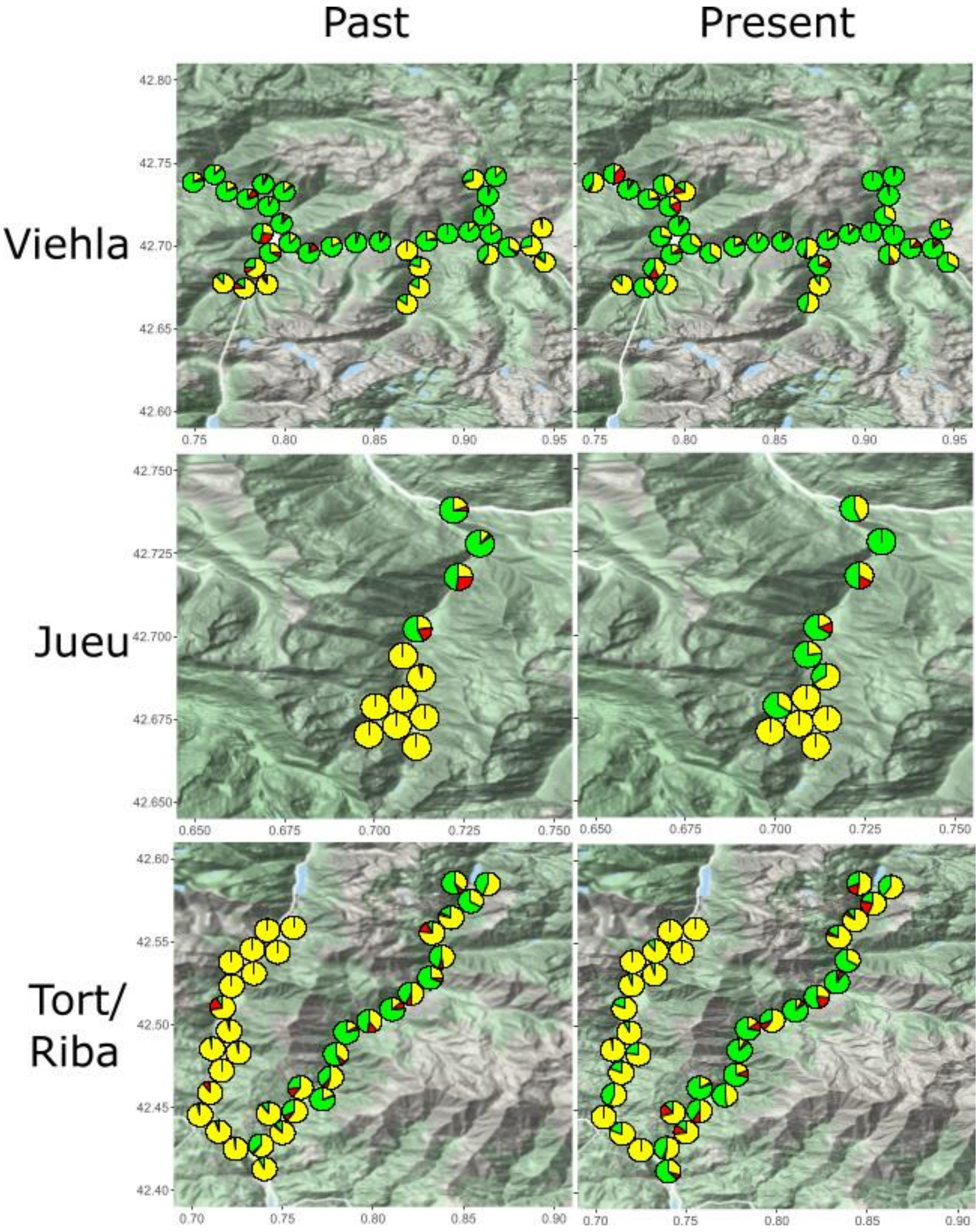

Figure 4. Changes in frequency of colour and banding types between paired sites (same location, or within 50 m) in four Pyrenean valleys over five decades, tested using paired T-test or Wilcoxon signed-rank test. Ribagorzana is the only valley that showed significant changes, with the frequency of brown ($p < 0.05$) and unbanded ($p < 0.01$) shells decreasing, and the proportion of banded shells increasing ($p < 0.01$).

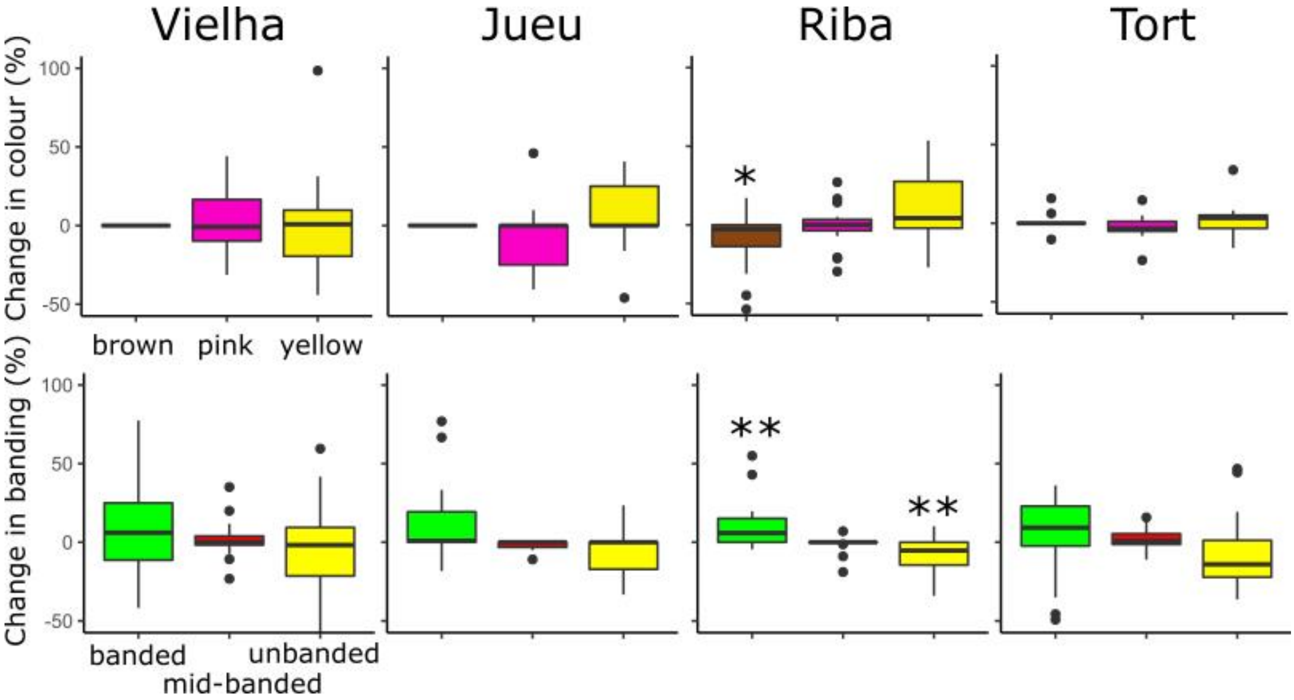

Figure 5. Scatterplots showing the present-day relationship between altitude and frequency of yellow and unbanded morphs in four Pyrenean valleys. Points represent collections of shells from the same location (n ≥ 10). Only samples from Jueu show a significant strong positive relationship between altitude and frequency of yellow and unbanded shells; samples from Tort showed a shallow but significant association for altitude and yellow. Regression line and confidence intervals are shown, alongside the Pearson coefficient and p value.

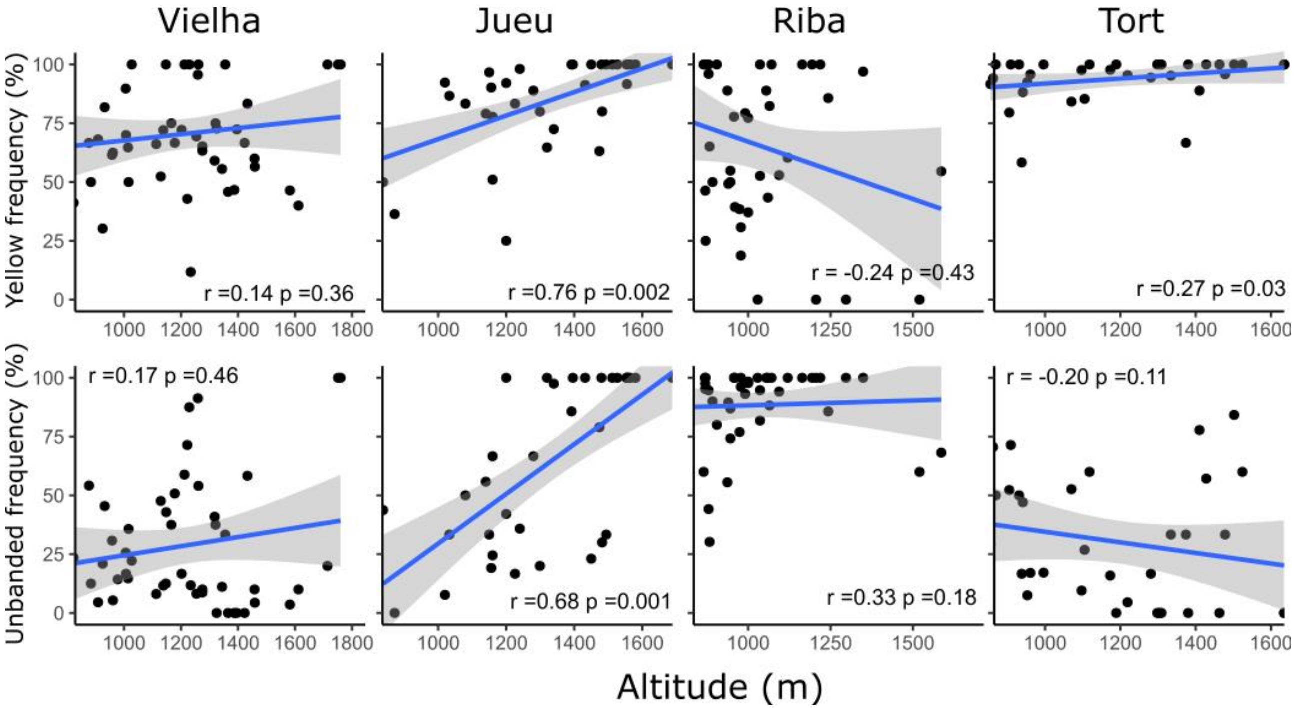

Figure 6. Scatterplots showing the relationship between altitude and chromatic variation (PC3) for individual shells from Vielha and Jueu valleys. Points represent individual shells, coloured according to human-scored colours. There is a strong positive association of PC3 with altitude in shells from Jueu, and a weak non-significant negative association in shells from Vielha. Regression line and confidence intervals are shown, alongside the Pearson coefficient and p value.

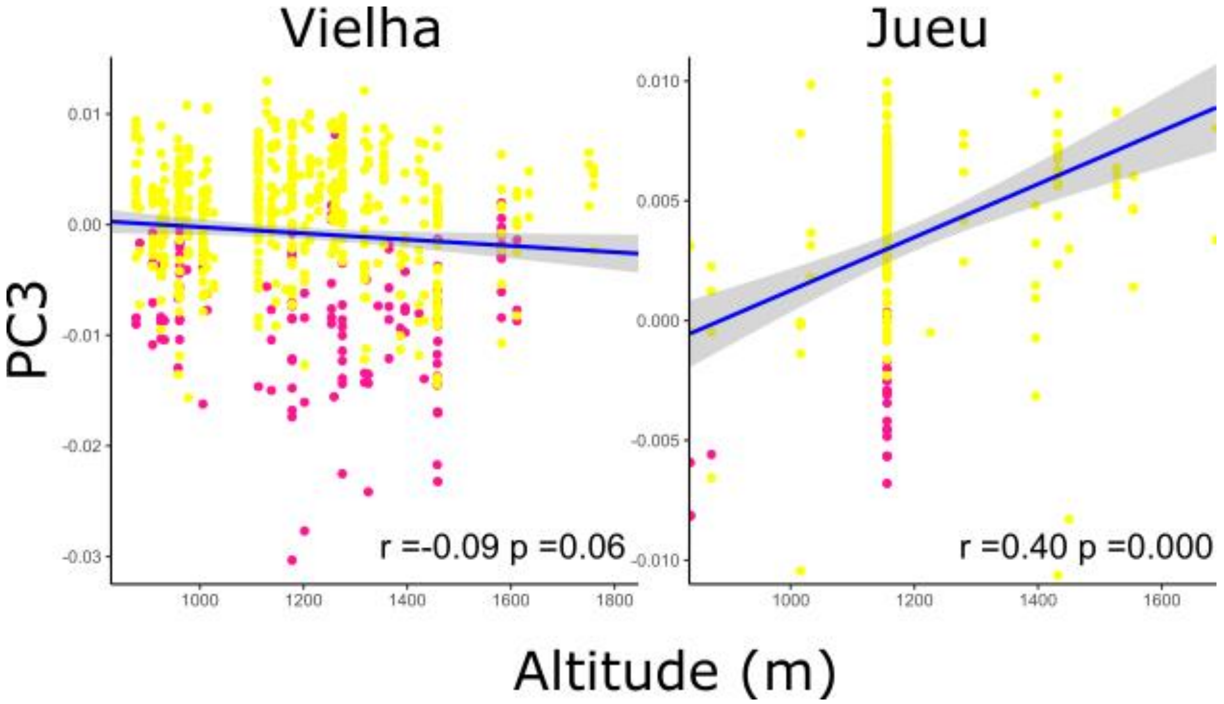

Figure 7. Summary figure showing the relationship between altitude and colour variation for shells from Vielha and Jueu valleys, comparing past and present-day collections, and using qualitative or quantitative methods to score colour. The small inset graph shows the same data, but only using the subset of shells that were considered sufficiently undamaged for spectrophotometry.

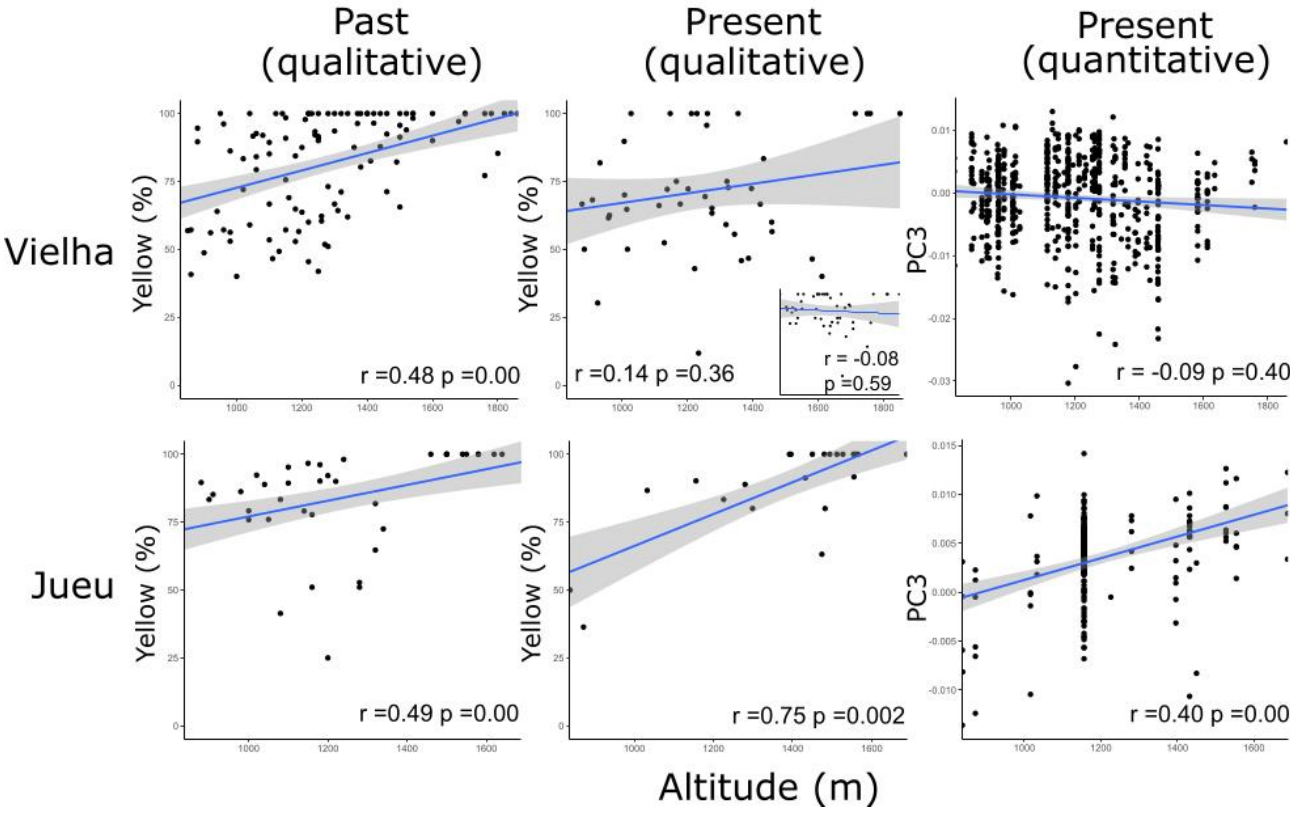

Figure S1. Present-day distribution of pale-lipped shell morphs in four Pyrenean valleys. Pie charts show frequencies of pale-lipped shells (white) versus other forms.

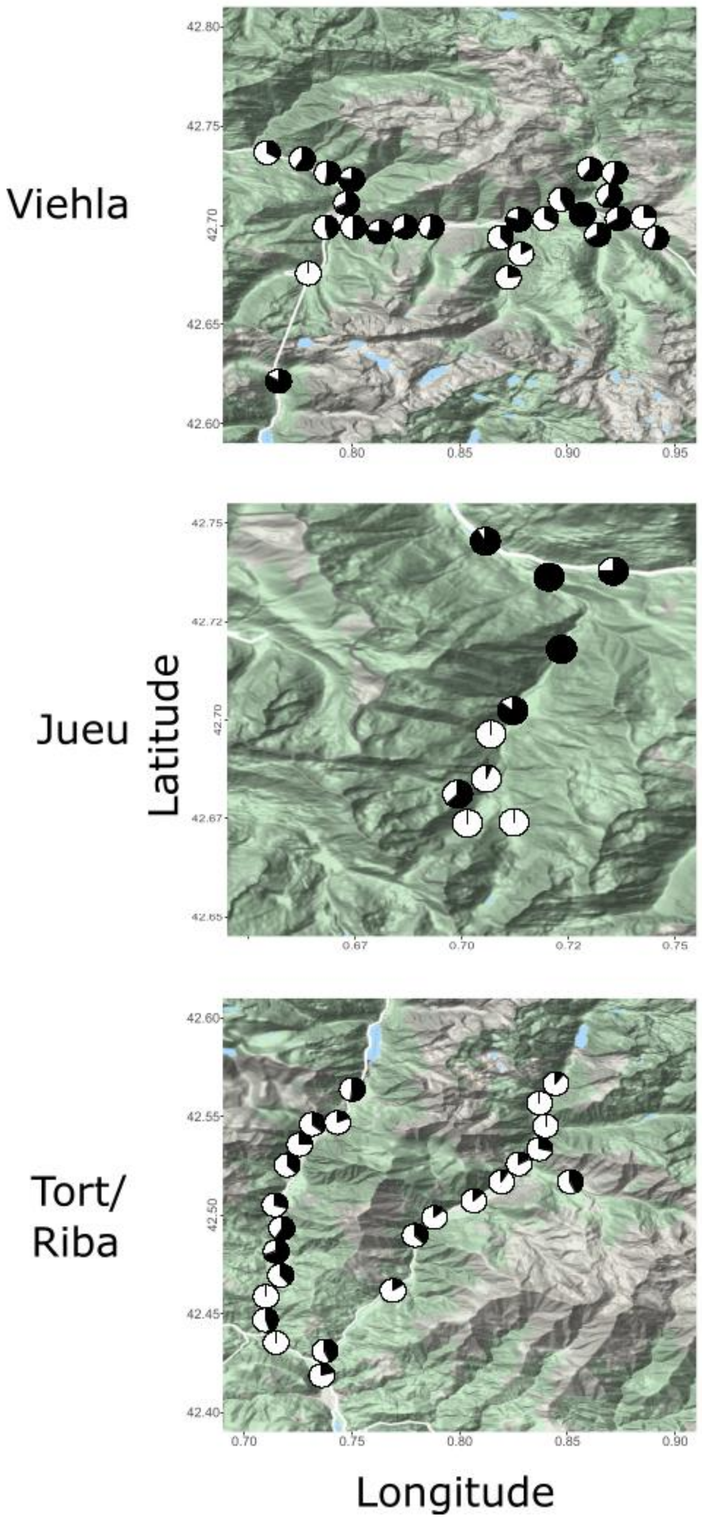

Figure S2. Scatterplots showing the present-day relationship between altitude and frequency of pale-lipped morphs in four Pyrenean valleys. Regression line and confidence intervals are shown, alongside the Pearson coefficient and p value.

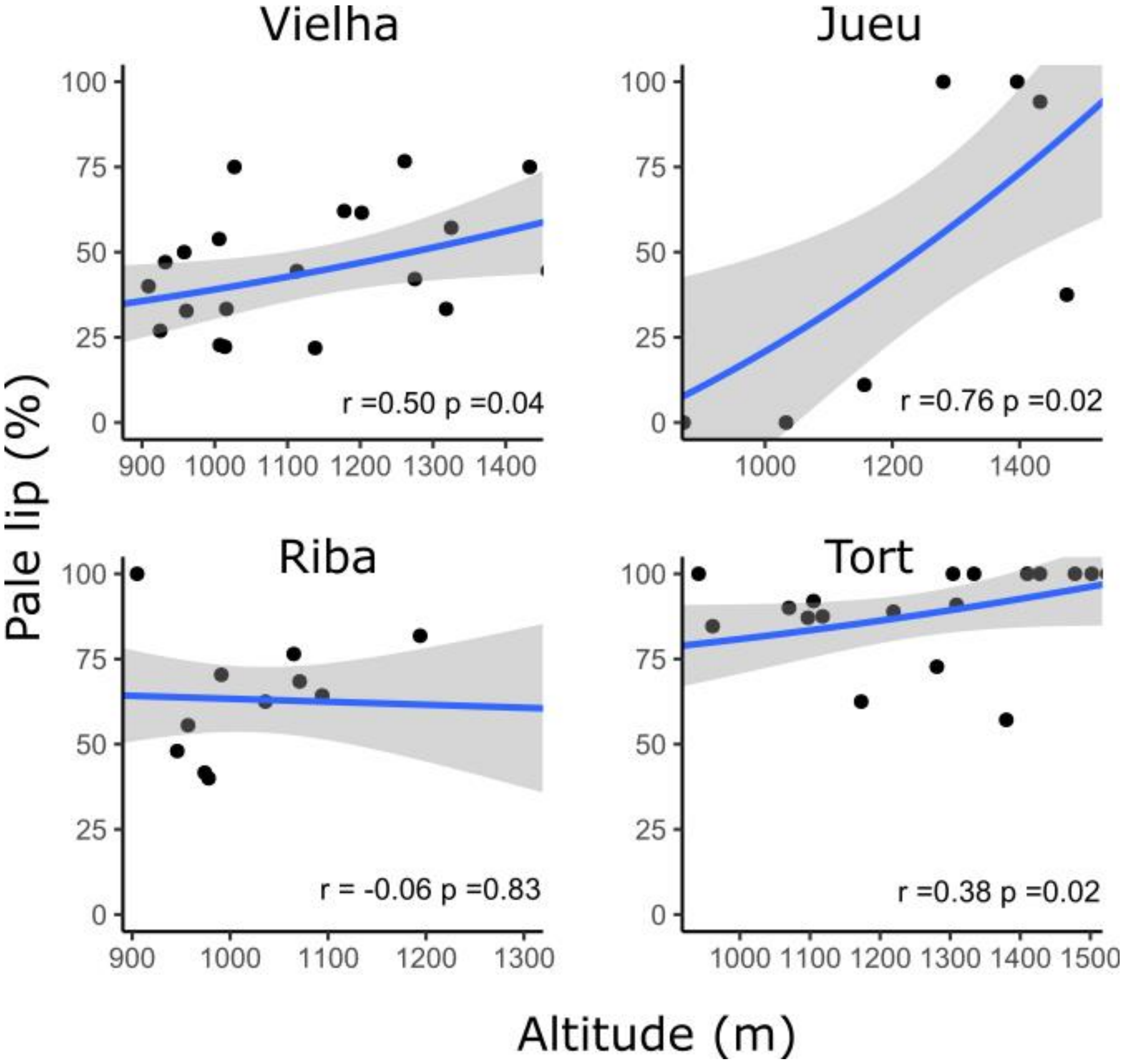

Figure S3. Scatterplots showing the present-day relationship between altitude and size and shape of shells in four Pyrenean valleys. Regression line and confidence intervals are shown, alongside the Pearson coefficient and p value.

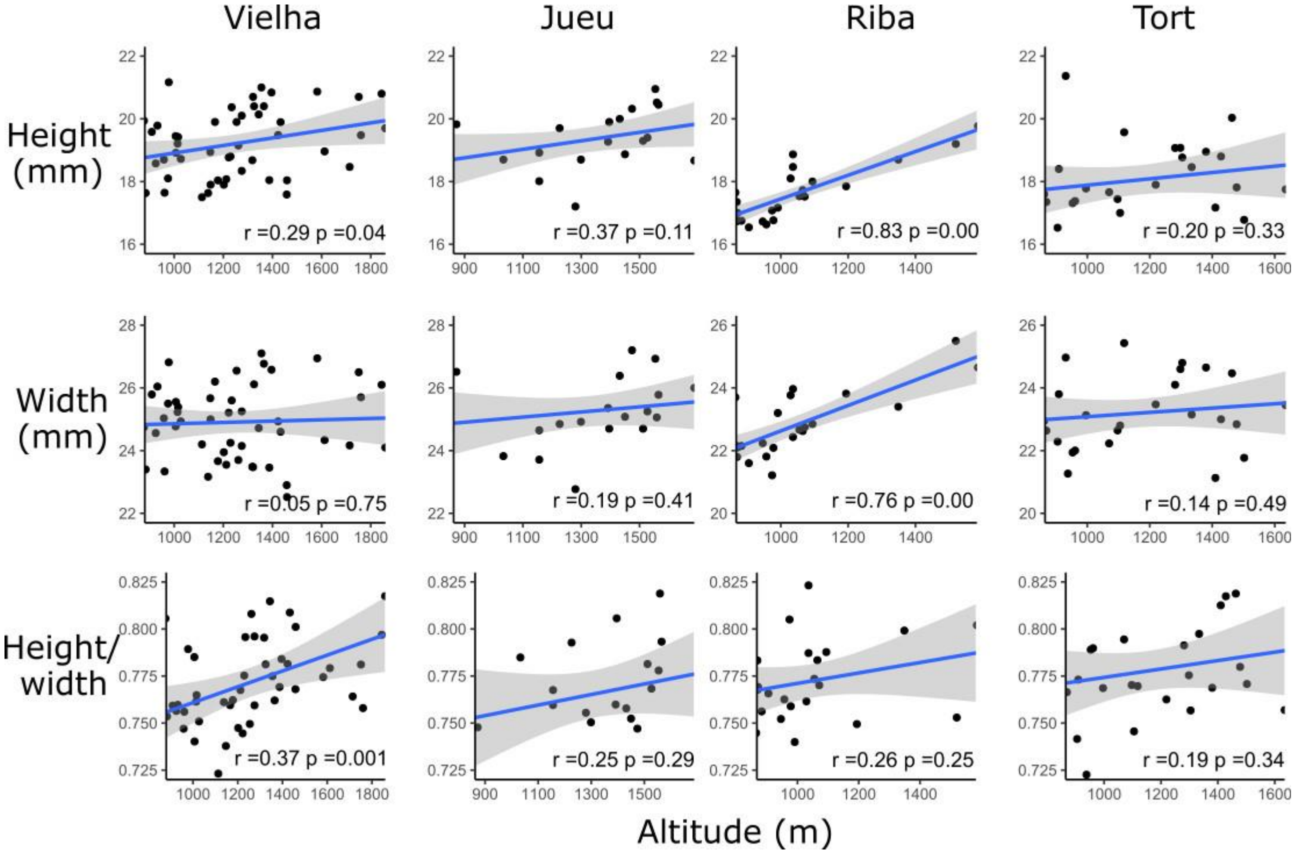

Figure S4. Scatterplot showing variation of visual space coordinates, xyz, on three principal component axes, using shells from Vielha and Jueu valleys in the Pyrenees. Units are in JNDs. Points are coloured according to human-scored classification of the shell, either yellow or pink.

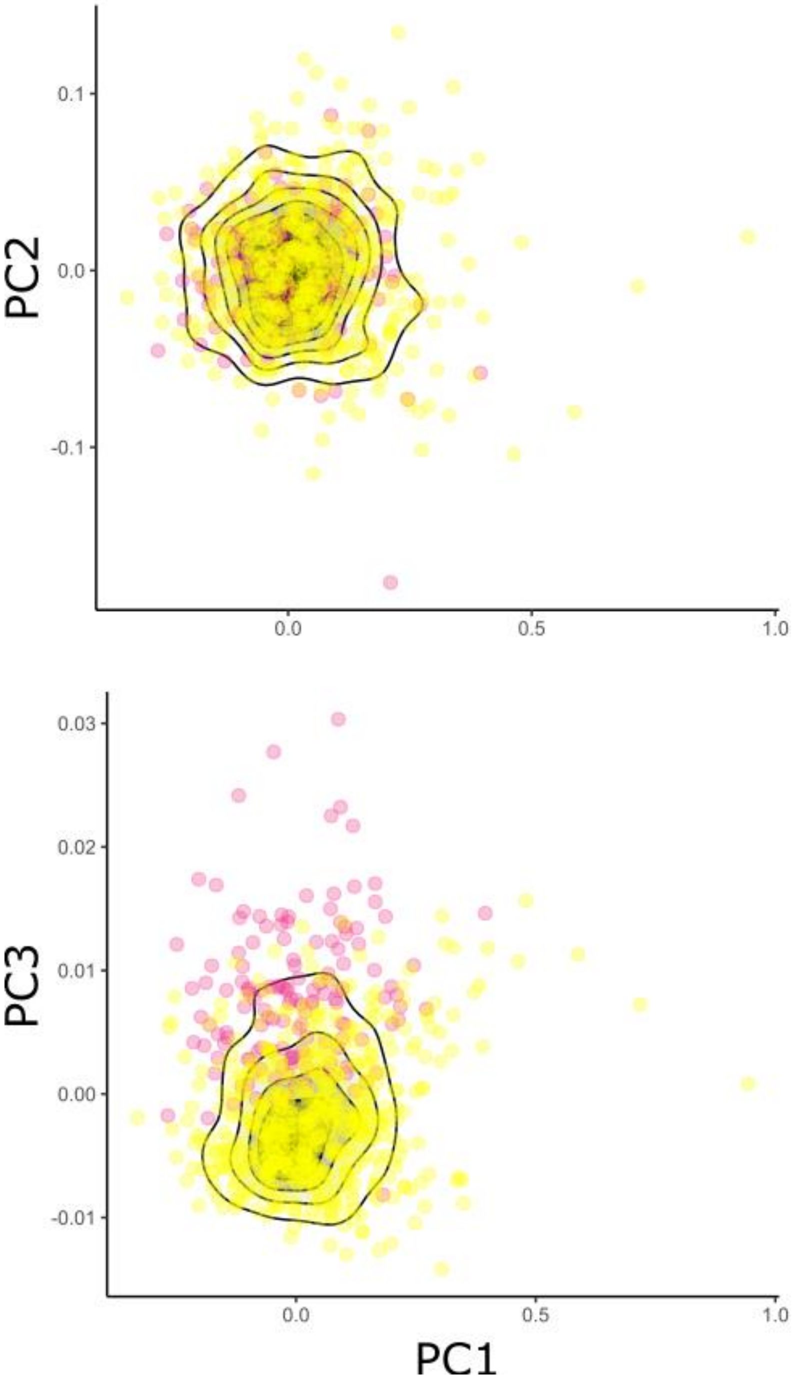

Table 1. Sampling summary; number of sites and snails for each valley.

Valley	Past		Present		Spectrophotometry	
	Sample sites	No. shells	Sample sites	No. shells	Sample sites	No. shells
Vielha	119	4756	43	942	43	607
Jueu	49	1862	17	637	12	206
Riba	34	1545	21	518		
Tort	24	1165	27	536		
Total	226	9328	108	2633	55	813

Table 2. Fisher's r-to-z transformation, significance of the difference between two correlation coefficients

Past vs Present	Vielha		Jueu		Riba		Tort	
	Z-Value	P-Value	Z-Value	P-Value	Z-Value	P-Value	Z-Value	P-Value
Yellow	2.120*	0.017	-1.460	0.072	-0.310	0.378	0.270	0.394
Unbanded	2.210*	0.014	1.510	0.066	-0.310	0.378	0.140	0.444
Banded	-3.16***	0.001	-1.180	0.119	1.120	0.131	-0.590	0.278
Yellow sets								
Yellow subset	3.31***	0.001	-0.740	0.230				
Yellow dataset-subset	0.980	0.164	0.420	0.337				

*p < 0.05. **p <0.01. ***p < 0.001

Table 3. Statistical summary of shell geographical distribution in each valley; independent paired comparison (Student's t-test (parametric),Wilcoxon signed-rank test (non-parametric)).

	Vielha				Jueu				Riba				Tort			
Present (2017/2018)	Mean	S.E.	% char	P-value	Mean	S.E.	% char	P-value	Mean	S.E.	% char	P-value	Mean	S.E.	% char	P-valu
Yellow	70.9	3.1	0.4	0.933	87.1	6.1	5.8	0.427	75.9	5.4	10.4	0.061	94.6	1.5	1.5	0.541
Pink	29.0	3.1	0.3	0.923	12.4	5.9	-6.3	0.379	9.0	3.2	-1.0	0.752	3.9	1.1	-2.3	0.204
Brown									15.4	5.2	-8.8*	0.049	1.4	1.0	0.6	0.570
Unbanded	27.0	4.0	-8.7	0.110	59.9	11.1	-13.2	0.152	85.7**	4.2	-9.6**	0.008	42.9	5.9	-6.6	0.239
Mid-banded	5.2	1.3	1.5	0.327	2.6	1.8	-2.1	0.059	0.4	0.4	-1.2	0.328	8.4	1.5	2.5	0.141
Banded	67.8	4.1	8.5	0.140	37.5	10.6	15.4	0.094	13.9**	3.9	10.8**	0.007	48.7	6.3	4.1	0.499
Past (1962/1969)																
Yellow	70.5	3.8			81.3	5.6			65.5	5.7			93.0	1.4		
Pink	29.3	3.7			18.7	5.6			10.0	2.5			6.1	1.4		
Brown									24.2	5.8			0.8	0.0		
Unbanded	35.7	5.7			73.2	11.3			95.3**	1.7			49.4	5.4		
Mid-banded	3.7	0.9			4.7	2.6			1.5	1.2			6.0	1.1		
Banded	59.3	5.7			22.1	9.8			3.2**	0.7			44.6	5.6		

*p < 0.05. **p <0.01. ***p < 0.001

References

Arnold, R. 1969. The effects of selection by climate on the land-snail *Cepaea nemoralis (L.)*. *Evolution* **23**: 370-378.

Arnold, R. W. 1968. Studies on *Cepaea* VII. Climatic selection in *Cepaea nemoralis* (L) in Pyrenees. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **253**: 549-593.

Berjano, R., Gauthier, P., Parisi, C., Vaudey, V., Pons, V., Renaux, A., Doblas, D. & Thompson, J. D. 2015. Variation of a floral polymorphism at different spatial scales in the Mediterranean geophyte *Narcissus assoanus*. *Journal of Plant Ecology* **9**: 333-345.

Cain, A. J. & Sheppard, P. M. 1950. Selection in the polymorphic land snail *Cepæa nemoralis*. *Heredity* **4**: 275-294.

Cain, A. J. & Sheppard, P. M. 1952. The effects of natural selection on body colour in the land snail *Cepaea-nemoralis*. *Heredity* **6**: 217-231.

Cain, A. J. & Sheppard, P. M. 1954. Natural selection in *Cepaea*. *Genetics* **39**: 89-116.

Cameron, R. A. D., Carter, M. A. & Haynes, F. N. 1973. The variation of *Cepaea nemoralis* in three Pyrenean valleys. *Heredity* **31**: 43-74.

Cameron, R. A. D. & Cook, L. M. 2012. Habitat and the shell polymorphism of *Cepaea nemoralis* (L.): interrogating the Evolution Megalab database. *Journal of Molluscan Studies* **78**: 179-184.

Cameron, R. A. D., Cook, L. M. & Greenwood, J. J. D. 2013. Change and stability in a steep morph-frequency cline in the snail *Cepaea nemoralis* (L.) over 43 years. *Biological Journal of the Linnean Society* **108**: 473-483.

Cassey, P., Honza, M., Grim, T. & Hauber, M. E. 2008. The modelling of avian visual perception predicts behavioural rejection responses to foreign egg colours. *Biology Letters* **4**: 515-517.

Cook, L. M. 2013. Selection and disequilibrium in *Cepaea nemoralis*. *Biological Journal of the Linnean Society* **108**: 484-493.

Cook, L. M. 2014. Morph frequency in British *Cepaea nemoralis*: what has changed in half a century? *Journal of Molluscan Studies* **80**: 43-46.

Cook, L. M., Cowie, R. H. & Jones, J. S. 1999. Change in morph frequency in the snail *Cepaea nemoralis* on the Marlborough Downs. *Heredity* **82**: 336-342.

Cowie, R. H. & Jones, J. S. 1998. Gene frequency changes in *Cepaea* snails on the Marlborough Downs over 25 years. *Biological Journal of the Linnean Society* **65**: 233-255.

Davison, A., Jackson, H. J., Murphy, E. W. & Reader, T. 2019. Discrete or indiscrete? Redefining the colour polymorphism of the land snail *Cepaea nemoralis*. *Heredity* **123**: 162-175.

Delhey, K., Delhey, V., Kempenaers, B. & Peters, A. 2014. A practical framework to analyze variation in animal colors using visual models. *Behavioral Ecology* **26**: 367-375.

Ellis, D. 2004. *Evolutionary genetics of the land snail Cepaea nemoralis in the Central Pyrenees*. University of London, University College London (United Kingdom).

Fisher, R. A. 1921. On the "Probable Error" of a coefficient of correlation deduced from a small sample. *Metron* **1**: 3-32.

García-Ruiz, J. M. (2015) Geo-ecological effects of Global Change in the Central Spanish Pyrenees: A review at different spatial and temporal scales. (J. I. López-Moreno, T. L. M., S. M. Vicente Serrano, P. González-Sampériz, B. L. Valero-Garcés, Y. Sanjuán, S. Beguería, E. Nadal-Romero, N. Lana-Renault, & A., ed.). pp. CSIC - Instituto Pirenaico de Ecología (IPE), CSIC - Instituto Pirenaico de Ecología (IPE).

Gonzalez, D. R., Aramendia, A. C. & Davison, A. 2019. Recombination within the *Cepaea nemoralis* supergene is confounded by incomplete penetrance and epistasis. *Heredity*.

Harvey, M. 2018. Using iRecord to support site-based recording. *Conservation Land Management* **16**: 25-29.

Horn, G., Aodha, O., Song, Y., Cui, Y., Sun, C., Shepard, A., Adam, H., Perona, P. & Belongie, S. (2018) The iNaturalist species classification and detection dataset. pp. pp. 8769-8778. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Hunt, R. W. G. 2004. *The reproduction of color*. Wiley–IS&T Series in Imaging Science and Technology, Chichester UK.

Jones, J. S. & Irving, A. J. 1975. Gene frequencies, genetic background and environment in Pyrenean populations of *Cepaea nemoralis (L)*. *Biological Journal of the Linnean Society* **7**: 249-259.

Jones, J. S., Leith, B. H. & Rawlings, P. 1977. Polymorpism in *Cepaea* - A problem with too many solutions. *Annual Review of Ecology and Systematics* **8**: 109-143.

Kerstes, N. A. G., Breeschoten, T., Kalkman, V. J. & Schilthuizen, M. 2019. Snail shell colour evolution in urban heat islands detected via citizen science. *Communications Biology* **2**: 264.

Lamotte, M. (1959) Polymorphism of natural populations of *Cepaea nemoralis*. In: *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 24. pp. 65-86. Cold Spring Harbor Laboratory Press.

Maia, R., Eliason, C. M., Bitton, P. P., Doucet, S. M. & Shawkey, M. D. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution* **4**: 906-913.

Maia, R., Gruson, H., Endler, J. A. & White, T. E. 2018. pavo 2.0: new tools for the spectral and spatial analysis of colour in R. *bioRxiv*: 427658.

Ochman, H., Jones, J. S. & Selander, R. K. 1983. Molecular area effects in *Cepaea*. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **80**: 4189-4193.

Ożgo, M., Liew, T. S., Webster, N. B. & Schilthuizen, M. 2017. Inferring microevolution from museum collections and resampling: lessons learned from *Cepaea*. *PeerJ* **5**: e3938-e3938.

Ożgo, M. & Schilthuizen, M. 2012. Evolutionary change in *Cepaea nemoralis* shell colour over 43 years. *Global Change Biology* **18**: 74-81.

Saenko, S. V., Groenenberg, D. S. J., Davison, A. & Schilthuizen, M. 2020. The draft genome sequence of the grove snail Cepaea nemoralis. *in review*.

Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. 2016. mclust 5: Clustering, classification and density estimation using Gaussian Finite Mixture models. *R Journal* **8**: 289-317.

Silvertown, J., Cook, L., Cameron, R., Dodd, M., McConway, K., Worthington, J., Skelton, P., Anton, C., Bossdorf, O., Baur, B., Schilthuizen, M., Fontaine, B., Sattmann, H., Bertorelle, G., Correia, M., Oliveira, C., Pokryszko, B., Ozgo, M., Stalazas, A., Gill, E., Rammul, U., Solymos, P., Feher, Z. & Juan, X. 2011. Citizen science reveals unexpected continental-scale evolutionary change in a model organism. *Plos One* **6**.

Smith, T. & Guild, J. 1931. The C.I.E. colorimetric standards and their use. *Transactions of the Optical Society* **33**: 73-134.

van den Berg, C. P., Troscianko, J., Endler, J. A., Marshall, N. J. & Cheney, K. L. 2020. Quantitative Colour Pattern Analysis (QCPA): A comprehensive framework for the analysis of colour patterns in nature. *Methods in Ecology and Evolution* **11**: 316-332.

Walker, D. 2003. JMASM9: Converting Kendall's Tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods* **2**: 525-530.

Westland, W., Ripamonti, C. & Cheung, V. (2012) Computing CIE Tristimulus Values. In: *Computational Colour Science using MATLAB.* pp. 27-47. Wiley.

Worthington, J. P., Silvertown, J., Cook, L., Cameron, R., Dodd, M., Greenwood, R. M., McConway, K. & Skelton, P. 2012. Evolution MegaLab: a case study in citizen science methods. *Methods in Ecology and Evolution* **3**: 303-309.