Adapting genetic algorithms for artificial evolution of visual patterns under selection from wild predators

Short title: Camouflage evolution in the wild

Emmanuelle S. Briolat[1], George R. A. Hancock[1] & Jolyon Troscianko[1]

[1]: Centre for Ecology and Conservation, Faculty of Environment, Science and Economy, University of Exeter, Penryn, TR10 8ER, UK

Corresponding author:

E. S. Briolat, esb204@exeter.ac.uk

Centre for Ecology and Conservation, University of Exeter, Penryn, TR10 8ER, UK

Abstract

Camouflage is a widespread and well-studied anti-predator strategy, yet identifying which patterns provide optimal protection in any given scenario remains challenging. Besides the virtually limitless combinations of colours and patterns available to prey, selection for camouflage strategies will depend on complex interactions between prey appearance, background properties and predator traits, across repeated encounters between co-evolving predators and prey. Experiments in artificial evolution, pairing psychophysics detection tasks with genetic algorithms, offer a promising way to tackle this complexity, but sophisticated genetic algorithms have so far been restricted to screen-based experiments. Here, we present methods to test the evolution of colour patterns on physical prey items, under selection from wild predators in the field. Our techniques expand on a recently-developed open-access pattern generation and genetic algorithm framework, modified to operate alongside artificial predation experiments. In this system, predators freely interact with prey, and the order of attack determines the survival and reproduction of prey patterns into future generations. We demonstrate the feasibility of these methods with a case study, in which free-flying birds feed on artificial prey deployed in semi-natural conditions, against backgrounds differing in three-dimensional complexity. Wild predators reliably participated in this experiment, foraging for 11 to 16 generations of artificial prey and encountering a total of 1,296 evolved prey items. Changes in prey pattern across generations indicated improvements in several metrics of similarity to the background, and greater edge disruption, although effect sizes were relatively small. Computer-based replicates of these trials, with human volunteers, highlighted the importance of starting population parameters

for subsequent evolution, a key consideration when applying these methods. Ultimately, these methods provide pathways for integrating complex genetic algorithms into more naturalistic predation trials. Customisable open-access tools should facilitate application of these tools to investigate a wide range of visual pattern types in more ecologically-relevant contexts.

Keywords

camouflage; evolution; genetic algorithms; predation; selection; virtual ecology

Introduction

Camouflage is a common anti-predator defence strategy, with a long history as a case study for evolution [1]. A vast body of work has investigated how visual patterns help conceal prey from predators, identifying a range of strategies from disruptive coloration and countershading to masquerade and iridescence [1,2]. Understanding which patterns produce the most effective protection is an important goal for camouflage research, particularly with respect to practical applications, yet determining which combinations of traits provide optimal camouflage is not straightforward. The typical approach of comparing the detectability of a set of phenotypes, varying in specific traits, has been hugely successful in determining relevant features for crypsis, but is poorly-suited to exploring camouflage optimisation: the vast phenotypic space available for colours and patterns, along with interacting effects of viewing conditions, backgrounds and receiver vision, make it unfeasible to test all possible combinations in this way [3]. Moreover, prey patterns are shaped by long-term interactions between communities of predators and prey, learning across repeated encounters and co-evolving over evolutionary time. In particular, search

image formation, whereby predators learn to more efficiently detect common prey types, is thought to lead to apostatic selection, preventing the evolution of a stable optimum and favouring polymorphisms in prey patterns [4,5].

Inspired by the processes of natural selection, artificial evolution of visual patterns using genetic algorithms presents an alternative approach, allowing for exploration of a wide phenotypic space and accounting for the effects of repeated interactions between predators and prey. Originally developed as a general solution to optimisation problems with vast parameter spaces [6], genetic algorithms are an increasingly popular technique in biological research, ideally suited to exploring visual pattern space. So far, genetic algorithms have been used to investigate the evolution of several aspects of visual signalling, primarily camouflage patterns [4,7–12], but also aposematism [13], motion dazzle [14] and even nectar guides in flowers [15]. In brief, genetic algorithms set out both virtual genetic structures underpinning phenotypes and rules governing survival and reproduction into the next generation, based on the responses of an observer; across multiple encounters with the observer in successive generations, optimal solutions to the task should arise. When the observer is a foraging predator, prey items that are easiest to detect are selected against, and so lost from the population, while those that are most difficult to find will survive and breed, ultimately leading to the evolution of better-camouflaged patterns.

Following similar principles, early experiments in virtual ecology tracked populations of different-coloured pastry baits under selection from wild birds, with the proportion of different bait types in each generation determined by the ratio of survivors in the previous round [16,17]. Building on this simple paradigm, genetic algorithms incorporate processes of

mutation and sexual reproduction, allowing for more complex variation in prey phenotype, and for novel traits to arise. In addition to the algorithms determining how patterns are retained, lost or modified between generations, the ways in which different pattern traits are encoded is also increasingly sophisticated. While no models claim to replicate the actual genetic architecture of a particular species, advances in our understanding of how animal patterns are genetically determined and physically produced during development [18] have inspired systems better able to produce a range of realistic natural patterns [12].

So far, these more complex genetic algorithms have generally been applied only in computer-based experiments, where they are most easily implemented, but which fail to capture the impacts of naturalistic viewing conditions and relevant receiver vision. Techniques such as projecting the search area on larger screens, using natural background images with occlusion layers, or simulating the visual system of dichromats can increase the realism of the search task [11], but still fall short of replicating the effects of real predator behaviour. To our knowledge, genetic algorithms have only been combined with selection by non-human predators in one scenario: seminal experiments on frequency-dependent selection in which blue jays (*Cyanocitta cristata*) searched for artificial moths with evolving patterns on screens [4,7]. Besides restricting prey patterns to greyscale targets, these experiments required highly-motivated captive predators - undergoing up to eight months of pre-training then 100 generations' worth of trials [7] - so are difficult to expand to other species and contexts.

Here, we describe a more generalisable process that combines the benefits of field trials involving natural predators with state-of-the-art tools for artificial evolution, expanding on a

recently-released genetic algorithm and pattern generation framework, the CamoEvo toolbox [12]. Prey patterns are generated by CamoEvo, with modifications facilitating implementation in the field, and deployed on physical prey items, concealing a food reward. Wild predators then interact with these items in semi-controlled arenas, and the order in which prey are attacked guides the evolution of subsequent generations of patterns. As a proof-of-principle, we present a case study in which we tested for effects of background complexity on the evolution of camouflage patterns, in artificial prey under selection from wild birds. We also replicated the field trials in computer-based experiments with human volunteers, to explore how field experiments could be rapidly prototyped and parametrised in future applications.

Materials and Methods

General principles

We developed a methodology for generating and testing artificial prey in the field using genetic algorithms, by substantially adapting an existing screen-based implementation, the CamoEvo toolbox [12], released here as CamoPrint in CamoEvo V2.0 (see supporting code). The properties of the prey and predation scenarios can be customised, and multiple colour and pattern metrics can be used to analyse the outcomes of pattern evolution, providing flexibility to suit a range of research questions (Fig 1). To begin with, an initial population of prey patterns is created by the CamoPrint software, following experimenter inputs for starting parameters, such as prey size, shape, population size and the colour space from which prey colours can be drawn (see [12] and supporting code). This first generation of

patterns is applied to real physical objects and deployed where predators of interest can interact with them. Suitable prey items can take virtually any form required, so long as they can conceal a food reward accessible to predators, or otherwise provide a record of predation attempts. We describe a technique for transferring computer-generated patterns onto plasticine or 3D-printed shapes (see case study), but appropriate two- and three-dimensional targets can also be made more simply by printing on paper (Fig 1). The setting, visual backgrounds and predators recruited will depend on the experimental question; these methods could be implemented in a wide range of scenarios, from relatively controlled predation trials with captive animals to more naturalistic experiments with multiple predators and prey scattered across a large area. As predators interact with prey, the timing or order of predation is recorded, whether through direct observation, filming, or regular checks, as in more traditional artificial predation experiments. Each population of prey can also be split into discrete groups, for example if prey are to be presented in a series of semi-controlled arenas (see case study). In that case, prey can be ranked according to the order of attack within each group; the size and number of groupings, and prey arrangement within arenas, can be customised as needed. Predation outcomes are then fed back into the genetic algorithm software, which uses this information to determine which patterns survive and reproduce into the next generation. To facilitate the integration of physical prey items into the existing CamoEvo workflow, we implemented several modifications to the pattern generation tools – to create the target shapes and shading used in the case study, assign prey to groups and positions in test arenas, and more easily input the results from predation trials – which can be adapted to suit each user's needs (see Supporting Information (SI) for details, and supporting code). Repeating these processes of pattern

generation, prey preparation, predation trials and data entry allows experimenters to follow selection by predators across multiple generations of prey.

Fig 1: Examples of possible target types and general methods workflow. Users can customise a range of experimental parameters, and investigate multiple pattern traits when analysing results. Prey presented here include paper targets, two-dimensional with simple or more specific shapes (top), and three-dimensional rolled paper tubes, as well as 3D-printed shells (middle, as used in case study). Rightmost panels show example prey *in situ*.

Finally, to explore the outcome of selection, experimenters can select a wide range of colour, luminance and pattern metrics to analyse changes in prey appearance across generations. CamoEvo can analyse the computer-generated prey patterns using a range of tools in the widely-used MICA and QCPA toolboxes in ImageJ [19,20], providing a fully open-access workflow, from experimental design to the assessment of pattern evolution. Alternatively, the physical prey items themselves can be photographed and processed using the same open-access tools, so that image analyses can account for the real 3D properties of prey and backgrounds, and relevant visual systems.

Case study - prey pattern evolution against backgrounds of varying three-dimensional complexity

As a proof-of-concept, we present an application of these methods, investigating the evolution of camouflage patterns in response to selection from wild birds. We deployed

prey in arenas with specific visual backgrounds, to test the effect of differences in three-dimensional background complexity on the strength and direction of pattern evolution. Theoretical models [21] and experiments with both human and avian predators [22–24] suggest even relatively poorly-camouflaged targets are more difficult to detect against complex backgrounds, relaxing selection for increased similarity to the background [24]. In addition, three-dimensional backgrounds introduce variation in lighting conditions, between shadowed areas and those under direct lighting, which creates noise in both luminance and colour [25,26]. As a result, we expected to see improved camouflage in our targets across generations, but with a weaker signal of directional change in the population of prey displayed against backgrounds with greater three-dimensional complexity.

Field set-up

Two sets of artificial predation experiments were run in different locations in the grounds of the University of Exeter's Penryn campus (United Kingdom; 50° 10' 12'' N, 5° 7' 22.8'' W), from 17th September to 21st October and 10th November to 11th December 2021. Experiments were approved by the University of Exeter's ethics committee (Application ID: eCORN003896 v3.1).

Our set-up was designed to encourage natural foraging behaviour by wild birds, while allowing us to control the backgrounds prey were viewed against, and accurately track predation events. Prey were placed in six wooden trays (53x53cm,22cm above the ground), set out in areas where birds likely to interact with the set-up were regularly seen. Each tray

183 was lined with a polypropylene ground cover membrane, allowing water to filter through,
184 and a mix of gravel and construction sand, forming the visual background (Fig 2). Many
185 alternative backgrounds could be used here; this gravel mix was selected as a relatively
186 natural and heterogeneous option, that was also stable enough to remain in place in poor
187 weather conditions and could be sculpted to vary the complexity of the scene. To test for an
188 effect of background complexity, each experimental run monitored the survival of two
189 populations of prey, presented in trays in which the gravel was either smoothed flat or given
190 3D structure, with wavy furrows creating peaks and troughs in various orientations
191 (henceforth, smooth and furrowed treatments; Fig 3).

192

193 **Fig 2: Experimental set-up for field trials.** Tray array for the first run (a), with the main
194 predators involved, *C. corone* (b) and *E. rubecula* (c), and a side view of a tray and
195 mechanical timing gate (d). Here, the first two items attacked were associated with the
196 green and red marbles (leftmost prey).

197

198 **Fig 3: Schematic of the process for selection and evolution of prey patterns in each**
199 **generation.** Inset: Diagram representing prey positioning in the trays, with strings
200 connecting to the mechanical timing gate.

201

202 Each population consisted of 24 prey, divided into six groups of four per tray. Every
203 morning, half of each population was placed out into three trays with smooth or furrowed
204 gravel respectively, so each generation of both populations was fully exposed to predators

205 across two consecutive days (Fig 3). The location of the smooth and furrowed backgrounds
206 among the trays was alternated daily, and the prey positions in each tray were pseudo-
207 randomised using a custom script in ImageJ [27], designed to reduce spatial clustering. In
208 CamoPrint, custom additions to the CamoEvo toolbox automated the allocation of prey
209 patterns to each tray and labelled them for easier identification (see SI and supporting
210 code). The artificial prey items themselves consisted of 3D-printed oval shells (18x14 mm)
211 which birds had to flip over to access a concealed food reward (dried mealworm and suet).
212 Prey pattern was manipulated by covering the shells in a thin (~1mm) layer of white tack
213 (UHU, Bühl, Germany), then applying custom patterns printed on waterslide transfer paper
214 (Rolurious) with a laser printer (Colour Laser Jet Enterprise M553, HP, USA). Prior to the
215 experimental trials, wild birds were gradually attracted to the trays and trained to search for
216 the artificial prey, starting with a single tray with visible baits, moving on to food in
217 upturned prey shells, then concealed rewards, and finally increasing the number of trays
218 and prey. During training, prey patterns consisted of checkerboards with colours sampled
219 from the colour space of the experimental patterns. Trials began once visiting birds had
220 consumed four prey in all six trays for at least three consecutive days. In the first run, the
221 trays were visited by carrion crows (*Corvus corone*); in the second, a robin (*Erithacus*
222 *rubecula*) was the main predator, with occasional visits from blackbirds (*Turdus merula*) and
223 great tits (*Parus major*) (Fig 2; S1 Fig; S1 Video).

224

225 During experimental trials, the order in which prey items were attacked in each tray served
226 as our measure of prey detectability. To automatically record this, the prey shells were
227 tethered from the underside to coloured strings running under the trays, connecting them

228 to a mechanical timing gate. The four prey items in each tray were always assigned string
229 colours in the same order – red, green, blue and yellow from left to right– with marbles of
230 matching colours in the timing gate; when an item was attacked, its string was pulled,
231 releasing a marble of the corresponding colour to create a visible record of attack order
232 (Figs 2,3). This system was not always sufficiently sensitive to record attacks from smaller
233 birds, so as a fail-safe, each tray was also filmed continuously with a small spy camera
234 (Coulax).

235

236 Prey were set out each morning (8 - 9am), and trays were checked every two to three hours;
237 each trial ended once at least three items had been attacked per tray. If no prey were
238 attacked over the course of the day, the trays were covered up overnight, and the same
239 prey presented again the following day. Once each generation had been fully exposed to
240 predation, the order of prey attacks in every tray was entered into the genetic algorithm
241 software, and a new set of prey patterns was produced for the next generation. See Fig 3 for
242 a full schematic of the experimental procedure.

243

244 *Pattern generation and algorithm set-up*

245 Prey patterns were generated using the CamoEvo toolbox's animal pattern generator, with a
246 modified script for the 3D targets used in this experiment (See SI). Targets were oval, with
247 colourations occupying a CIELAB space with ranges 0 – 100 in luminance, -50 – 50 in A
248 (green-red) and -10 – 70 in B (blue-yellow). Every population began with an identical set of
249 patterns, randomly-generated using these parameters with a uniform distribution of genes

250 (Generation 0; S2 Fig). Details of the pattern generation and genetic algorithm framework
251 can be found in [12].

252

253 CamoEvo is primarily designed for computer-based experiments, so we implemented
254 several modifications to facilitate its application to a field experiment with physical prey
255 items, where prey survival must be manually reported (described in full in SI). These
256 modifications form the basis of the CamoPrint tools featured in CamoEvo V2.0. Once the
257 trials for each generation were complete, prey were assigned a fitness ranking from 1 to 4 in
258 every tray, based on the order in which they were attacked. The genetic algorithm then
259 selected the last two patterns attacked in each tray (ranks 3 & 4) to survive and reproduce,
260 creating the next generation of patterns. However, to reduce the likelihood of relatively
261 well-camouflaged prey being eliminated by chance, we applied CamoEvo's built-in rescue
262 system [12]: if a prey pattern was ranked 1 or 2 in generation N, but ranked 4 (the highest
263 possible rank) in generation N-1, the pattern also survived into the next generation (N+1). In
264 the first run, trials stopped after generation 10, when crows stopped visiting the site; in the
265 second, trials ended as planned after 15 generations. To account for the final set of predator
266 decisions, we used CamoPrint to produce an additional generation of patterns (generation
267 11 or 16 respectively), included in subsequent analyses.

268

269 Prey camouflage was assessed with digital photography and image analysis, using well-
270 established methods implemented in the Multispectral Image Calibration and Analysis
271 (MICA) and QCPA plugins for ImageJ [19,20]. The following commonly-used metrics of
272 camouflage efficacy were analysed, accounting for perception by relevant avian visual

systems: chromatic (ΔS) and achromatic differences (ΔL) between mean colours in the Receptor Noise-Limited space [20,28,29], pattern difference based on granularity analyses [30], and target edge disruption, measured using the GabRat method [31]. To account for prey pattern, where component colours may better match the background than the mean prey colour, we also calculated a frequency-weighted average measure of colour difference between all colours in each prey pattern and the most similar colours in the background. Statistical analyses were carried out in R version 3.5.2 [32]. See SI for details.

Screen-based search tasks

The outcomes of the field trials suggested our methods could be improved by refining the experimental parameters, to adjust task difficulty. To investigate this, we replicated the field trials with matching screen-based search tasks with human volunteers, allowing us to more rapidly test the impacts of different starting conditions; compared to a month-long run of field trials, each replicate of the screen-based task took roughly only half an hour for a participant to complete. These search tasks were also designed using the CamoEvo toolbox, and run using its built-in user interface for screen-based psychophysics experiments. All parameters were selected to mimic the field experiment as closely as possible, with the exception of starting colour space: target patterns occupied either the same colour space as prey in the field (henceforth the full colour space), or a narrower colour space, corresponding to 2 standard deviations around the mean colour of the background images. In the initial generation 0, targets from the full colour space were thus identical to generation 0 targets in the field, while those from the narrow colour space featured the same patterns but with a restricted range of colours (S5 Fig). Each participant viewed a

single population of targets over 16 generations, shown on photographs of the same smooth or furrowed gravel backgrounds as in the field. Every image contained four targets to find, and participants were tasked with clicking on them as quickly as possible, within a maximum of 15 seconds per slide. Survival of targets into the next generation and pattern evolution were governed by the same rules as in the field, based on the order in which targets were clicked. Changes in the detectability of targets over successive generations were quantified using capture time, the time taken for participants to click on each target, as recorded by the CamoEvo toolbox. Measures of prey phenotype used to assess camouflage efficacy, and statistical analyses, were also selected to match those used for prey in the field trials as closely as possible (see SI for details of the experimental design and analyses). In addition, we used CamoEvo to produce 12 replicate populations, with identical starting conditions to the experimental populations, but in which prey patterns were selected at random to survive and reproduce into the next generations. These provide control trends, against which experimental populations can be compared to demonstrate the effect of selection by human volunteers, as opposed to neutral drift.

A total of 12 volunteers completed these screen-based tasks, between 23rd March and 13th April 2022 (three per combination of background type [smooth or furrowed] and colour space [full and narrow]). Experiments took place in a dark room, with images displayed on a 27” monitor (iiyama, Japan). Participants were recruited among staff and students at the University of Exeter, aged 18-70 with normal or corrected-to-normal vision and not knowingly colour-blind. The experiments complied with the declaration of Helsinki and received ethical approval from the University of Exeter’s CLES Cornwall Ethics committee

(application ID: 512348, date approved: 22/03/22). Participants received an information sheet presenting the rationale and rules of the search task, and signed a consent form to approve the use of their anonymised data in this study.

Results

Case study - pattern evolution in field trials

Prey populations showed evidence of improved camouflage across generations, though results varied between background types and experimental runs (Fig 4). In the first run, colour and luminance differences between prey patterns and the background decreased, as expected, for the population shown on smooth gravel, but actually increased slightly for the population on furrowed gravel (linear model (lm), generation:background, $F_{2,570} = 7.625$, $F_{2,570} = 17.425$, $p < 0.001$, for colour difference between means and the weighted average colour difference respectively; $F_{2,570} = 18.997$, $p < 0.001$ for luminance; Fig 4a(i-iii)). Conversely, pattern differences decreased across generations in the more complex furrowed treatment, but increased in the smooth (lm, generation:background, $F_{1,572} = 16.282$, $p < 0.001$; Fig 4a(iv)), suggesting that the relative importance of pattern and colour matching varied between backgrounds. In the second run, luminance difference decreased in both populations, but most strongly on the furrowed backgrounds, (lm, generation:background, $F_{1,812} = 6.869$, $p = 0.009$; Fig 4b(iii)). However, there was little improvement in colour-matching: differences between mean target and background colours declined across generations only on the furrowed background (lm, generation:background, $F_{2,810} = 10.761$, $p < 0.001$; Fig 4b(i)) while the weighted average measure found little change

in that population, and an increase in colour difference across generations on the smooth background (lm, generation:background, $F_{1,812} = 28.069$, $p < 0.001$; Fig 4b(ii)). Unlike in the first run, there were no significant trends in pattern difference (lm, generation:background, $F_{2,810} = 1.328$, $p = 0.266$; background, $F_{1,812} = 1.524$, $p = 0.217$; generation, $F_{2,813} = 0.975$, $p = 0.378$; Fig 4b(iv)). When improvements in background-matching were observed, effect sizes were small, and targets remained conspicuous, with colour and luminance differences well above accepted detection thresholds of 1 to 3 ΔS or ΔL. For example, in the smooth background population in the first run, colour differences between means declined from 10.28 ΔS in generation 0 to 9.32 ΔS in generation 11 (Fig 4a(i)). Luminance differences reached lower levels, suggesting that luminance matching was more critical for camouflage, but this could be due to lower starting values; for example, in the same population, mean luminance difference started at 7.07 ΔL in generation 0, reaching 3.47 ΔL by generation 11 (Fig 4a(iii)).

Fig 4: Metrics of camouflage efficacy across generations, in targets from the first (a) and second run (b) of field trials, alongside expected trends for improved camouflage (c). All plots display targets shown against smooth backgrounds in the left panel, and furrowed backgrounds on the right. Trend lines and shaded ribbons represent final model predictions, with 95% confidence intervals. Generation was always initially fitted as a second-degree polynomial; non-linear trends are shown if the polynomial term was significant in final models.

363 Most consistently, targets in all populations displayed increased edge disruption across
364 generations (Fig 4a,b(v)); this improvement appeared to level off in the population shown
365 on furrowed backgrounds in the first run (lm, generation:background type, $F_{2,570} = 4.754$, $p =$
366 0.009), but not in the second (lm, generation:background, $F_{2,810} = 1.976$, $p = 0.139$;
367 background, $F_{1,812} = 0.531$, $p = 0.467$; generation, $F_{1,814} = 72.172$, $p < 0.001$).

368

369 *Evolution in screen-based replicates*

370 In the screen-based experiments, camouflage metrics reveal clear differences between
371 populations of prey utilising the full and narrow colour spaces. For populations using the
372 narrow colour space, patterns improved in every camouflage metric across generations (Fig
373 5a). There were significant decreases in luminance, colour and pattern differences between
374 targets and the backgrounds (lme, for luminance: generation, $\chi^2_1 = 114.07$, $p < 0.001$; for
375 colour: generation, $\chi^2_2 = 191.31$, $p < 0.001$; for pattern: generation, $\chi^2_1 = 8.466$, $p = 0.004$),
376 and an increase in edge disruption in target patterns (lme, for GabRat in L: generation, $\chi^2_1 =$
377 8.043, $p = 0.005$), all with no significant differences between populations shown against
378 smooth or furrowed backgrounds (see S1 Table). For the weighted average colour difference
379 metric, there was a significant effect of the interaction between generation and background
380 type, describing different patterns of overall decrease in colour differences
381 (generation:background type, $\chi^2_2 = 7.434$, $p = 0.024$; Fig 5a(iii)). These improvements in
382 target camouflage made it harder for participants to find them, as suggested by a significant
383 increase in the time taken to capture the first target on each slide over successive
384 generations for the narrow colour space, regardless of background type (lme, generation:

385 background, $\chi^2_2 = 4.762$, $p = 0.093$, background: $\chi^2_1 = 3.826$, $p = 0.050$, generation: $\chi^2_2 = 6.340$,
386 $p = 0.012$; Fig 5a(vi)).

387 By contrast, for the full colour space, there was substantial variation in the trajectory of
388 pattern evolution between replicate populations, making it difficult to see consistent
389 evidence of selection for improved camouflage (Fig 5b). There were no significant trends in
390 luminance or pattern difference between the targets and the background across
391 generations and no evidence of increased edge disruption in prey patterns (Fig 5b(i), (iv-v),
392 S2 Table). There was however some evidence for improved colour-matching overall:
393 average colour difference between targets and backgrounds decreased slightly across
394 generations (lme, generation, $\chi^2_2 = 7.256$, $p = 0.027$), and using the weighted average colour
395 difference metric suggests that this improvement was greater against the smooth
396 backgrounds (lme, generation:background, $\chi^2_2 = 6.965$, $p = 0.031$, Fig 5b(ii-iii), S2 Table).
397 Capture time was shaped by the interaction between generation and background type, but
398 did not increase overall (lme, generation: background, $\chi^2_2 = 10.832$, $p = 0.004$; Fig 5b(vi)).

399

400 **Fig 5: Changes in camouflage metrics in the screen-based experiments, for prey**
401 **populations using the narrow (a) and full (b) colour spaces.** Plots show the mean and
402 standard error of each metric per generation, along with predictions from the best statistical
403 models. Shades of each colour represent different populations, with filled dots for those
404 shown on smooth backgrounds, and open dots for furrowed backgrounds. Dotted lines
405 represent population-level predictions, and thicker lines overall trends for fixed effects only.

406

Results from the control populations, simulating the effects of neutral drift, confirm that consistent selection for improved camouflage was only seen in experimental populations using the narrow colour space. Under those conditions, there were significant differences between the trajectories of experimental populations, selected by volunteers, and control replicates across generations: average colour and luminance differences between target patterns and their backgrounds decreased in the selected populations, while they increased slightly in the controls (lme, generation:scenario, for luminance $\chi^2_1 = 9.443$, $p = 0.002$, for colour, $\chi^2_1 = 15.557$, $p < 0.001$; S6 Fig). Similar trends for the weighted average colour difference metrics and edge disruption were not significant, suggesting that volunteers in our experiment were selecting primarily for improved colour and luminance match (see S6 Fig, S3 Table). However, for the full colour space, analysing all replicates together revealed no significant interaction between scenarios (control versus experimental) and generation, for any metrics (see S6 Fig, S4 Table).

Discussion

This study presents methods to support investigations of camouflage evolution using genetic algorithms in the field, with prey items exposed to wild predators. Our detailed case study describes how to implement these tools, from modifications of the genetic algorithm software to techniques for making artificial prey and recording predation events, and demonstrates the feasibility of this approach. Free-flying birds reliably interacted with multiple generations of prey, and prey patterns improved in several metrics of camouflage efficacy, suggesting the potential for adaptive evolution.

430 However, this case study also highlights important factors to consider when designing this
431 type of evolutionary experiment. Trends for improved camouflage in the field were
432 relatively weak, and inconsistent among runs with different predators and background
433 types. We tested whether these findings were linked to the starting colour space available
434 for prey patterns, using screen-based search tasks with human volunteers, which closely
435 replicated the field trials and were much quicker to run. These experiments revealed
436 substantial variation in outcomes produced by individual participants, when target patterns
437 utilised the same colour space as in the field: while camouflage improved across generations
438 in some populations of prey, others became more conspicuous, such that few overall trends
439 for improved camouflage were detected, and populations were not significantly different
440 from control replicates, simulating neutral drift. These results were unexpected, as similar
441 psychophysics experiments in artificial evolution typically find much clearer trends towards
442 improved camouflage [12]. However, the number of prey and the way they are presented
443 strongly affect learning in this type of experiment [33]. Here, the presence of multiple prey
444 per image increases the influence of random factors, such as the position of prey relative to
445 one another. Under these circumstances, participant learning outstripped target evolution,
446 highlighting the unexplored power of predator learning in holding back the evolution of
447 camouflage from a conspicuous evolutionary start point, a scenario that may be relevant to
448 aposematic species, for example. Predator cognition, encompassing learning, neophobia,
449 forgetting, recognition errors and adaptive decision-making, is a key consideration in the
450 evolution and maintenance of aposematism as a signalling strategy [34,35]. Theoretical
451 models suggest that elements of both predator psychology and prey ecology, such as prey
452 density and aggregation, can contribute to the initial evolution of aposematism [36], and
453 our set-up presents opportunities to investigate these effects, by varying the profitability

454 and arrangement of prey in both the field and online systems. Predator behaviour and
455 cognition are similarly important for the evolution of camouflage yet have received less
456 attention in this context [37], so further experiments with real predators could also shed
457 more light on the conditions required for crypsis to evolve, given different predator
458 communities. In our experiments, the screen-based trials suggest the trajectory of evolution
459 is essentially random when prey utilise the full colour space, so any specific results from the
460 field case study should be interpreted with caution.
461
462 By contrast, the screen-based experiments showed that populations of prey utilising a
463 narrower colour space, based on colours found in the backgrounds, were not only more
464 cryptic from the start, but also consistently evolved to be better camouflaged over
465 successive generations, and clearly behaved differently from control replicates. We
466 therefore predict that field trials would similarly yield greater improvement in camouflage if
467 the prey patterns started with a more restricted colour space, based on the colours of the
468 background. In addition, very few metrics of camouflage efficacy revealed differences in
469 prey pattern evolution between populations displayed against the smooth and furrowed
470 backgrounds, even in the screen-based experiments with the narrow colour space. In the
471 field runs, trends did vary between background types, but there was little consistency across
472 metrics, or between repeats, and thus no clear sense that selection was stronger for
473 populations on the simpler smooth background. This suggests that the variation in three-
474 dimensional complexity between these backgrounds was also insufficient to select for
475 different evolutionary trajectories in this scenario. As such, these experiments represent a
476 proof-of-principle for virtual evolution in the wild, but future applications should carefully
477 consider starting parameters, to better calibrate the difficulty of the task.

Researchers attempting similar work would benefit from testing their starting parameters in pilot experiments with human volunteers as “predators”, to establish conditions in which robust selection can be expected to take place before running more time-consuming experiments in the field. To that end, the methods described here present a significant advantage, as CamoEvo’s flexibility allows users to set up screen-based experiments that closely mirror their field setups, and rapidly collect data with human volunteers. Equally, the limitations of our field experiments highlight the difficulty of predicting the conditions under which prey patterns can be expected to evolve, which is in of itself an interesting area to explore. Using these field methods to investigate the scenarios that favour the evolution of camouflage in response to real predators would be extremely valuable, to better understand the role of predation as a selective force. In particular, the impact of the starting colour space available to prey has received little attention, yet is likely to be a significant constraint for the subsequent evolution of prey patterns. Beyond prey appearance, initial conditions are an important determinant of evolutionary trajectories more generally. For example, certain traits can promote or restrict the potential for adaptive radiation (eg. when competition between small individuals for a shared resource limits the emergence of larger individuals that could specialise on alternative resources [38]) while more evolutionarily labile traits can facilitate the spread of invasive species, or boost the resilience of native species they encounter [39]. So-called “historical difference experiments”, testing how populations that previously evolved under different conditions perform in a new situation, provide a means to explore the role of starting conditions and historical contingencies in evolution [40]. Our set-up provides opportunities to run this kind of experiment, potentially

501 by combining screen and field experiments for efficiency, and approach these broad
502 evolutionary questions through the study of pattern evolution.
503
504 The system we present here is designed to encourage future research on prey appearance,
505 by providing open-access tools for every step of the process, from pattern generation and
506 the genetic algorithm framework [12], to implementation in the field, and image analysis
507 software for quantifying changes in prey properties [19,20]. Within CamoEvo and
508 CamoPrint, parameters such as target shape, colour space and genetic algorithm rules are
509 fully customisable, while the field set-up can be modified to meet individual needs, for
510 different predators and backgrounds. This should facilitate the application of virtual
511 evolution to investigating a wide range of visual pattern types and functions. In particular,
512 identifying optimal camouflage patterns remains a key question, with many practical
513 applications. Recent studies tackling this problem have gained useful insights by defining
514 optimal patterns based on the statistical properties of the background [41], or using deep
515 learning to extrapolate from human responses to a limited set of phenotypes [3]. Yet
516 genetic algorithms present an especially powerful tool for studying camouflage
517 optimisation, particularly when paired with machine learning techniques [10,11]. Crucially,
518 they account for the effects of repeated interactions between predators and prey, though
519 implementation in screen-based experiments restricts their interpretation. Associating
520 these tools with predation experiments in the field provides opportunities to investigate
521 optimal camouflage while simultaneously considering a wider range of influential factors,
522 from background properties and viewing conditions to predator vision, learning and prey
523 population dynamics. Beyond camouflage, field-based evolutionary approaches supported

by these tools could also provide valuable new insights into other forms of anti-predator colouration, such as aposematism or distant-dependent signalling.

Acknowledgments

We thank Juho Jolkkonen & Andrew Szopa-Comley for fieldwork assistance, all screen experiment volunteers, and Will Allen and an anonymous reviewer for helpful comments on an earlier version. GRAH was funded by a NERC GW4+ studentship (NE/S007504/1), JT & ESB by a NERC Independent Research Fellowship awarded to JT (NE/P018084/1).

Conflict of Interest Statement

The authors declare no conflict of interest.

Author Contributions

JT: Conceptualization, Methodology, Supervision, Writing – review & editing

GRAH: Conceptualization, Methodology, Investigation, Visualization, Writing – review & editing

ESB: Methodology, Investigation, Formal Analysis, Visualization, Writing – original draft, Writing – review & editing

Data Availability

Data and analysis code are available as supplementary material; code for tools used in this paper can be found on GitHub:

https://github.com/GeorgeHancock471/CamoWild_Repository_CE.v1.2

The tools used for creating printed targets, including site randomisation have been condensed and modified for greater ease of use in CamoEvo Version 2.0 and will be included in all future versions. Instructions on using these tools are provided in the CamoEvo 2.0 handbook: https://github.com/GeorgeHancock471/CamoEvo-v2.0-2022_Plugins/tree/main

References

1. Stevens M, Merilaita S. Animal camouflage: current issues and new perspectives. Philos Trans R Soc London B Biol Sci. 2009;364(1516):423–7.
2. Kjernsmo K, Whitney HM, Scott-Samuel NE, Hall JR, Knowles H, Talas L, et al. Iridescence as camouflage. Curr Biol. 2020;30(3):551–5.
3. Fennell JG, Talas L, Baddeley RJ, Cuthill IC. Optimizing colour for camouflage and visibility using deep learning: the effects of the environment and the observer's visual system. J R Soc Interface. 2019;16:20190183.
4. Bond AB, Kamil AC. Visual predators select for crypticity and polymorphism in virtual prey. Nature. 2002;415:609–13.
5. Troscianko J, Nokelainen O, Skelhorn J, Stevens M. Variable crab camouflage patterns defeat search image formation. Commun Biol. 2021;4(1):1–9.
6. Whitley D. A genetic algorithm tutorial. Stat Comput. 1994;4(2):65–85.
7. Bond AB, Kamil AC. Spatial heterogeneity, predator cognition, and the evolution of color polymorphism in virtual prey. Proc Natl Acad Sci. 2006;103(9):3214–9.
8. Sherratt TN, Pollitt D, Wilkinson DM. The evolution of crypsis in replicating populations of web-based prey. Oikos. 2007;116:449–60.
9. Reynolds C. Interactive evolution of camouflage. Artif Life. 2011 Apr 1;17(2):123–36.
10. Talas L, Fennell JG, Kjernsmo K, Cuthill IC, Scott-Samuel NE, Baddeley RJ. CamoGAN: evolving optimum camouflage with Generative Adversarial Networks. Methods Ecol Evol. 2020;11(2):240–7.
11. Fennell JG, Talas L, Baddeley RJ, Cuthill IC, Scott-Samuel NE. The Camouflage Machine : optimizing protective coloration using deep learning with genetic algorithms. Evolution. 2021;1–11.
12. Hancock GRA, Troscianko J. CamoEvo: an open access toolbox for artificial camouflage evolution experiments. Evolution. 2022;870–82.
13. Sherratt TN, Beatty CD. The evolution of warning signals as reliable indicators of prey defense. Am Nat. 2003;162(4):377–89.
14. Hughes AE, Griffiths D, Troscianko J, Kelley LA. The evolution of patterning during movement

in a large-scale citizen science game. Proc R Soc B Biol Sci. 2021;288(1942).

15. Lawson DA, Rands SA. The evolution of floral guides: Using a genetic algorithm to investigate the evolution of floral cue arrangements. Biol J Linn Soc. 2018;123(4):739–53.

16. Allen JA, Cooper JM, Hall GJ, McHenry C. 'Evolving pastry': a method for simulating microevolution. J Biol Educ. 1993;27(4):274–82.

17. Cooper JM, Allen JA. Selection by wild birds on artificial dimorphic prey on varied backgrounds. Biol J Linn Soc. 1994 Apr 1;51(4):433–46.

18. Orteu A, Jiggins CD. The genomics of coloration provides insights into adaptive evolution. Nat Rev Genet. 2020;21(8):461–75.

19. Troscianko J, Stevens M. Image calibration and analysis toolbox - a free software suite for objectively measuring reflectance, colour and pattern. Methods Ecol Evol. 2015;6(11):1320–31.

20. van den Berg CP, Troscianko J, Marshall NJ, Cheney KL, Endler JA. Quantitative Colour Pattern Analysis (QCPA): a comprehensive framework for the analysis of colour patterns in nature. Methods Ecol Evol. 2020;11:316–32.

21. Merilaita S. Visual background complexity facilitates the evolution of camouflage. Evolution. 2003;57(6):1248–54.

22. Xiao F, Cuthill IC. Background complexity and the detectability of camouflaged targets by birds and humans. Proc R Soc B Biol Sci. 2016; 283:20161527.

23. Murali G, Mallick S, Kodandaramaiah U. Background complexity and optimal background matching camouflage. Behav Ecol Sociobiol. 2021;75(4).

24. Rowe ZW, Austin DJD, Chippington N, Flynn W, Starkey F, Wightman EJ, et al. Background complexity can mitigate poor camouflage. Proc R Soc B Biol Sci. 2021;288(1963).

25. Lovell PG, Tolhurst DJ, Párraga CA, Troscianko J, Troscianko T, Baddeley R, et al. Stability of the color-opponent signals under changes of illuminant in natural scenes. J Opt Soc Am A. 2005;22(10):2060.

26. Troscianko T, Benton CP, Lovell PG, Tolhurst DJ, Pizlo Z. Camouflage and visual perception. Philos Trans R Soc B Biol Sci. 2009;364(1516):449–61.

27. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671–5.

28. Vorobyev M, Osorio D. Receptor noise as a determinant of colour thresholds. Proc R Soc B Biol Sci. 1998;265:351–8.

29. Renoult JP, Kelber A, Schaefer HM. Colour spaces in ecology and evolutionary biology. 2017;92:292–315.

30. Spottiswoode CN, Stevens M. Visual modeling shows that avian host parents use multiple visual cues in rejecting parasitic eggs. Proc Natl Acad Sci. 2010 May 11;107(19):8672–6.

31. Troscianko J, Skelhorn J, Stevens M. Quantifying camouflage: how to predict detectability from appearance. BMC Evol Biol. 2017;17(1):7.

32. R Development Core Team. R: a language and environment for statistical computing. Team RDC, editor. R Foundation for Statistical Computing. 2015. Available from: http://www.r-

project.org

33. Troscianko J, Lown AE, Hughes AE, Stevens M. Defeating crypsis: detection and learning of camouflage strategies. PLoS One. 2013;8(9):e73733.

34. Servedio MR. The effects of predator learning, forgetting, and recognition errors on the evolution of warning coloration. Evolution. 2000;54(3):751–63.

35. Skelhorn J, Halpin CG, Rowe C. Learning about aposematic prey. Behav Ecol. 2016;27(4):955–64.

36. Speed MP. Can receiver psychology explain the evolution of aposematism? Anim Behav. 2001;61(1):205–16.

37. Skelhorn J, Rowe C. Cognition and the evolution of camouflage. Proc R Soc B Biol Sci. 2016;283(1825):20152890.

38. Ten Brink H, Seehausen O. Competition among small individuals hinders adaptive radiation despite ecological opportunity. Proc R Soc B Biol Sci. 2022;289(1971).

39. Perkins TA. Evolutionarily labile species interactions and spatial spread of invasive species. Am Nat. 2012;179(2):E37–54.

40. Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: replaying life's tape. Science. 2018;362(6415).

41. Michalis C, Scott-Samuel NE, Gibson DP, Cuthill IC, Michalis C. Optimal background matching camouflage. 2017; 284:20170709.

Supporting Information captions

SI1: Additional methodological information for case study field trials

S1 Fig. Evidence of predation from a crow (*C. corone*) in run 1 (a), and a robin (*E. rubecula*), great tit (*P. major*) and blackbird (*T. merula*) in run 2 (b-d).

S2 Fig. Prey patterns in the field trials: generation 0 (left) and the final generation in runs 1 (middle) and 2 (right) for populations shown on smooth (top) and furrowed backgrounds (bottom).

S3 Fig. Altered target shading method, showing the base method (top-down) on the left and the custom method (centre-edge) on the right.

SI 2: Additional information for screen-based search tasks

S4 Fig. Target image modification for screen-based replicate experiments.

S5 Fig. Schematic of the screen-based search tasks, showing the targets in generation 0 and 15, for populations using the full and narrow colour spaces.

S6 Fig. Changes in camouflage metrics in the screen-based experimental populations and control replicates, using the narrow (a) and full (b) colour spaces.

S1 Table. Model simplification tables for full models testing changes in metrics of camouflage efficacy in the screen-based experiment, for the narrow colour space.

S2 Table. Model simplification tables for full models testing changes in metrics of camouflage efficacy in the screen-based experiment, for the full colour space.

S3 Table. Model simplification tables for full models testing changes in metrics of camouflage efficacy between control and experimental populations of the screen experiment, for the narrow colour space.

S4 Table. Model simplification tables for full models testing changes in metrics of camouflage efficacy between control and experimental populations of the screen experiment, for the full colour space.

S1 Video. Example predations in field trials.

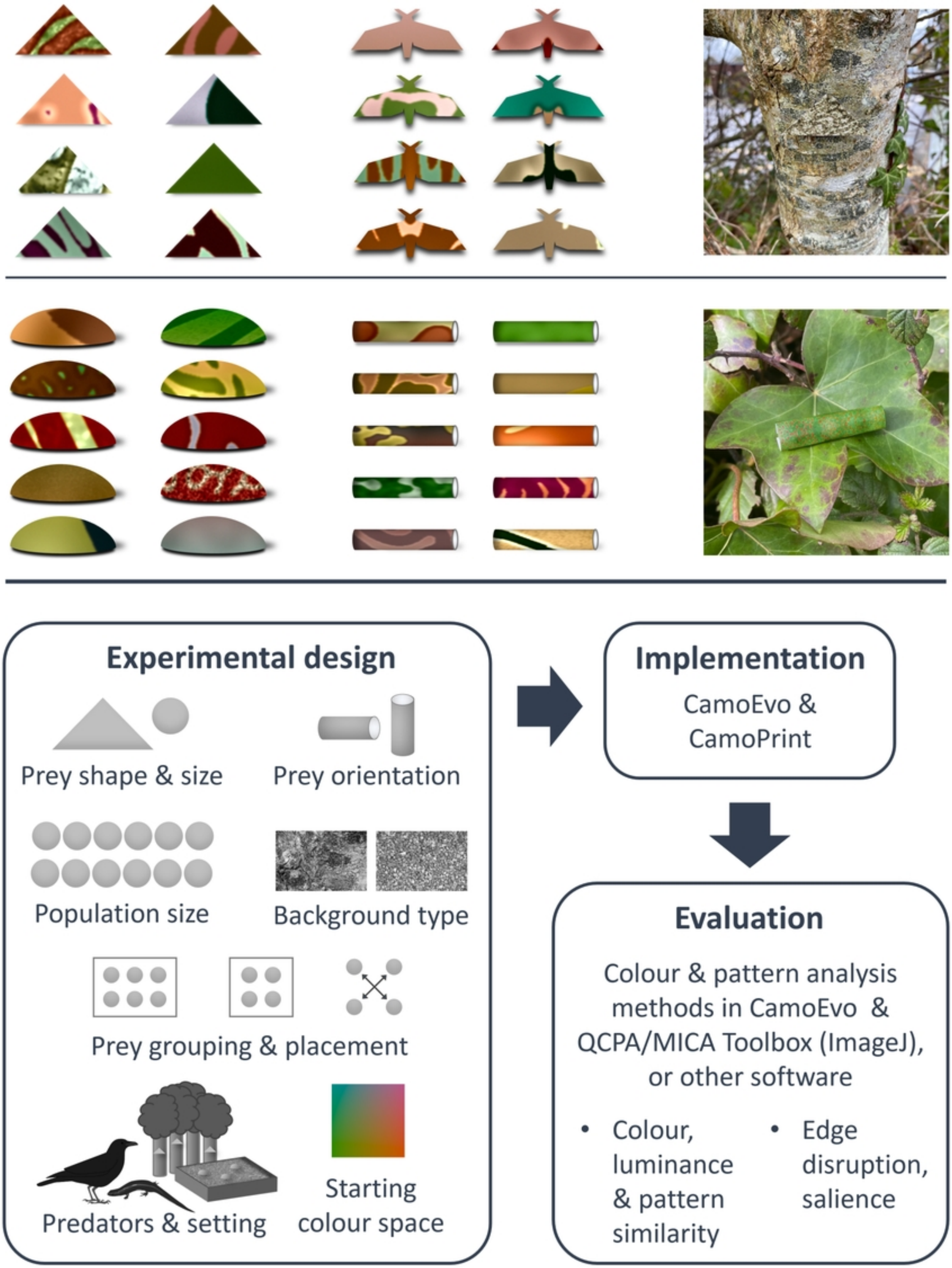
Experimental design
Prey shape & size
Prey orientation
Population size
Background type
Prey grouping & placement
Predators & setting
Starting colour space
Implementation
CamoEvo & CamoPrint
Evaluation
Colour & pattern analysis methods in CamoEvo & QCPA/MICA Toolbox (ImageJ), or other software
• Colour, luminance & pattern similarity
• Edge disruption, salience

Figure 1

Figure 2

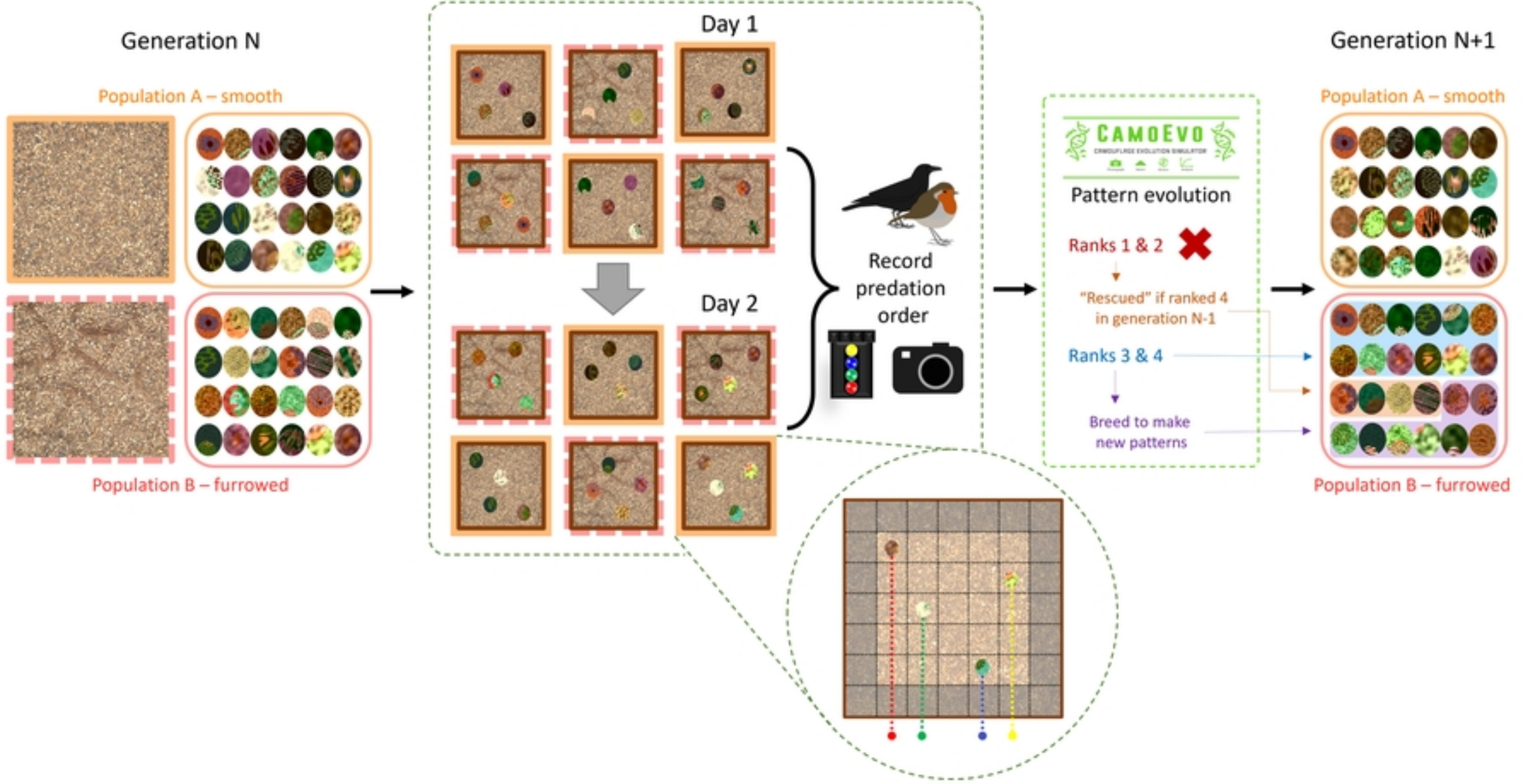
Generation N
Population A – smooth
Population B – furrowed
Day 1
Day 2
Record predation order
CAMOEVO
Pattern evolution
Ranks 1 & 2
"Rescued" if ranked 4 in generation N-1
Ranks 3 & 4
Breed to make new patterns
Generation N+1
Population A – smooth
Population B – furrowed

Figure 3

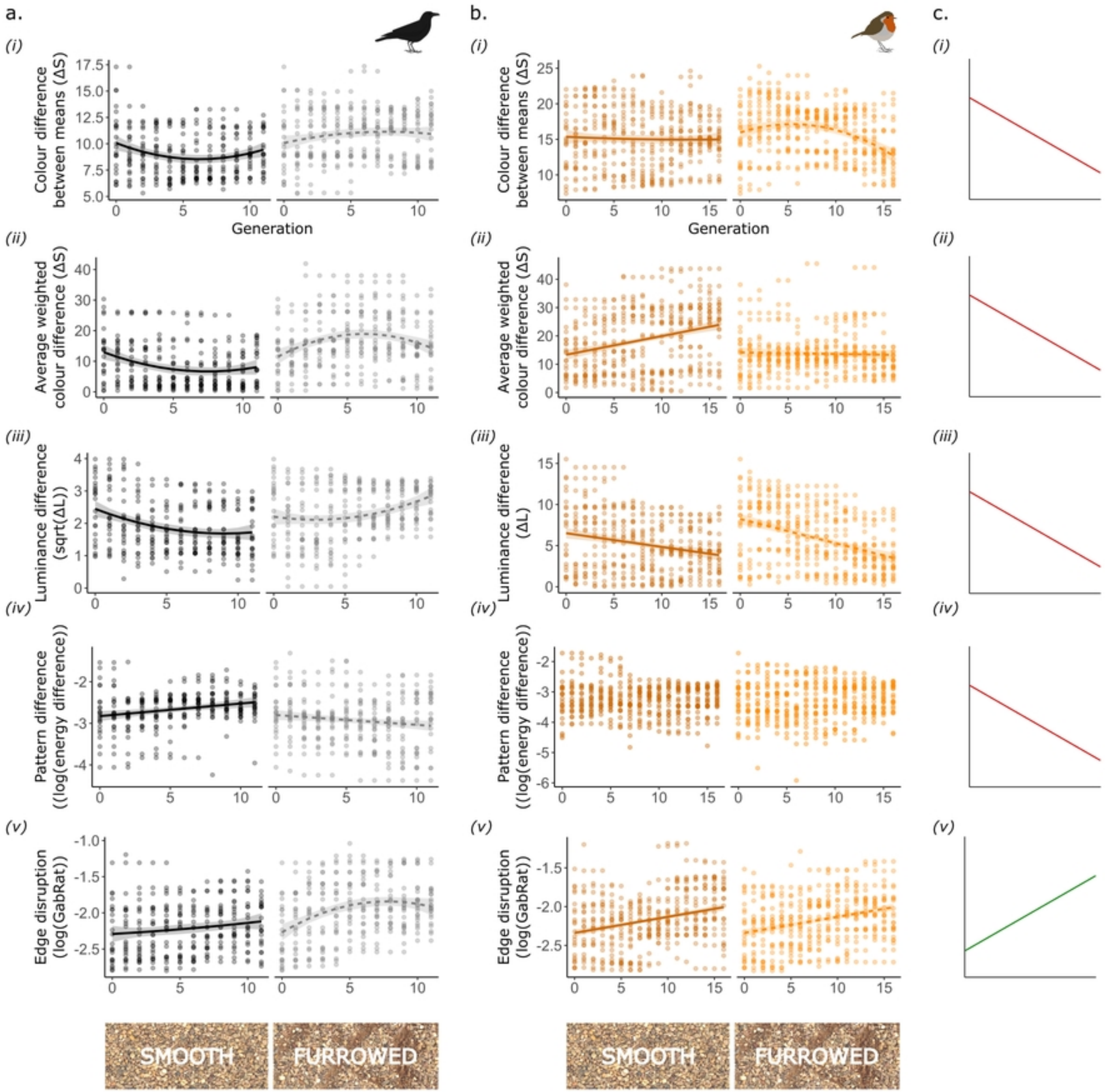
a.
b.
c.
(i)
(ii)
(iii)
(iv)
(v)
Colour difference between means (ΔS)
Average weighted colour difference (ΔS)
Luminance difference (sqrt(ΔL))
Luminance difference (ΔL)
Pattern difference ((log(energy difference))
Edge disruption (log(GabRat))
Generation
SMOOTH
FURROWED

Figure 4

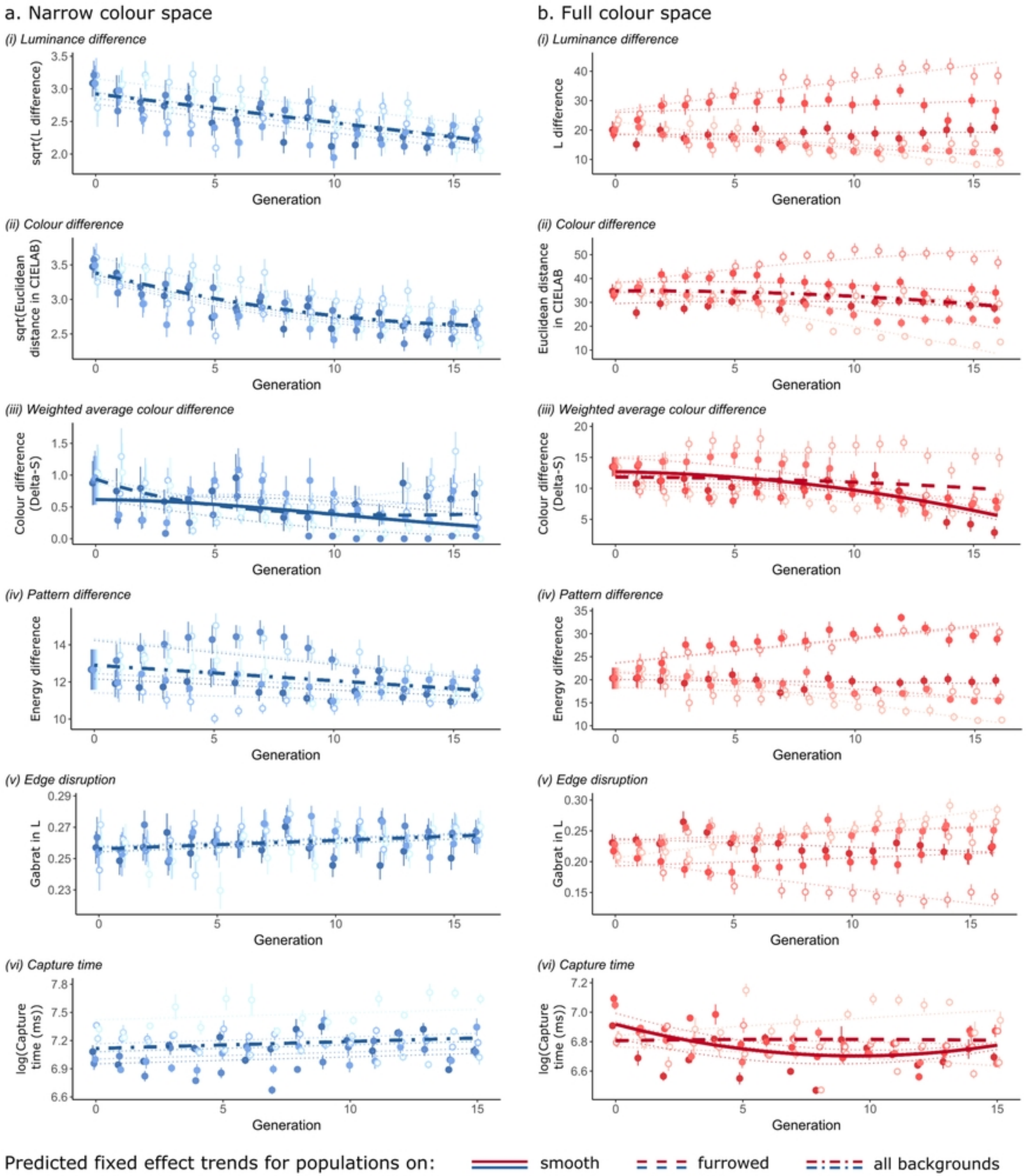
a. Narrow colour space
b. Full colour space
(i) Luminance difference
sqrt(L difference)
L difference
(ii) Colour difference
sqrt(Euclidean distance in CIELAB)
Euclidean distance in CIELAB
(iii) Weighted average colour difference
Colour difference (Delta-S)
(iv) Pattern difference
Energy difference
(v) Edge disruption
Gabrat in L
(vi) Capture time
log(Capture time (ms))
Generation
Predicted fixed effect trends for populations on:
smooth
furrowed
all backgrounds

Figure 5