MONASH University

Environmental, genetic and fitness correlates of male seasonal plumage ornaments in a genetically monogamous bird

Marie Fan

MSc Biological Sciences

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in 2018

School of Biological Sciences

Copyright notice

Table of contents

Abstract

Elaborate ornamentation is often assumed to improve reproductive success of the bearer, by attracting mates or deterring rivals. In particular, bright, ornate plumage in birds may signal the quality of an individual, provided that its production and maintenance are costly. In some species, such costliness may account for the development of seasonal plumages, when individuals alternate between dull non-breeding and colourful breeding plumage. Seasonal plumages can signal aspects of individual quality through multiple components, such as the timing of plumage production or variation in breeding colour quality. Hence, they offer an opportunity for unique insights into the evolution of multiple sexual signals and seasonal adaptations, which are particularly puzzling in species that form long-term, genetically monogamous partnerships. Here I investigate the function of male seasonal plumages in the cooperatively breeding purple-crowned fairy-wren, *Malurus coronatus*. Conspicuous seasonal male plumages of fairy-wrens are classically viewed as female choice-driven, since early acquisition of the breeding plumage is critical for obtaining extra-pair paternity (EPP), which is extremely high in this genus. However, *M. coronatus* shows nearly no EPP, and yet, males annually develop an unusual and striking, purple-and-black breeding plumage. Using microscopy, I first investigate the proximate mechanisms underlying the production of the purple colour, unique in the genus, in comparison with other structural colours of male fairy-wrens. I show that feather nanomorphology is overall similar and that the large colour diversity in the genus arises from variation in multiple structural components (Chapter 1). Then, using thirteen years of data and a detailed population pedigree, I examine the environmental, genetic and life-history correlates of (i) the timing of breeding plumage acquisition, (ii) breeding plumage completeness, and (iii) the quality (reflectance) of the breeding colouration. I establish that, despite some potential to act as a sexual signal of male quality, early pre-breeding moult does not provide any fitness benefits, and may constitute a vestigial trait, following the loss of extreme EPP levels and female extra-pair mate choice (Chapter 2). However, breeding plumage completeness plays a crucial role in male-male interactions, primarily in competition among subordinates for breeding positions. Possibly, male breeding plumage represents a sexual trait that shifted functions – from female choice to male-male competition, or a dual-function trait that lost its function in female choice (Chapter 3). Moreover, I demonstrate that the quality (reflectance) of the purple breeding colouration is highly variable, substantially heritable and predictive of male breeding success (Chapter 4). Overall, my findings show that the

maintenance of male seasonal plumages in *M. coronatus* results from the dynamic interplay between female choice and male-male competition, illustrating the remarkable flexibility in function of sexual ornaments. They also support the multi-signalling potential of seasonal plumages, with different components correlating with different aspects of individual quality and fitness. More broadly, my thesis significantly contributes to our understanding of how complex fluctuations in sexual selection may shape the evolutionary trajectory of male ornaments, and highlights the need to continue testing and re-evaluating some of the fundamental assumptions of sexual selection, which can yield new research perspectives.

Publications during enrolment

Fan, M., Hall, M. L., Kingma, S. A., Mandeltort, L. M., Hidalgo Aranzamendi, N., Delhey, K., & Peters, A. (2017). No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy? *Behavioral Ecology*, **28**, 1055-1067.

Fan, M., Teunissen, N., Hall, M. L., Hidalgo Aranzamendi, N., Kingma, S. A., Roast, M., Delhey, K., & Peters, A. (2018). From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird. *Journal of Animal Ecology*, **87**, 1274-1285.

Eastwood. J., Hall, M. L., Teunissen, N., Kingma, S. A., Hidalgo Aranzamendi, N., Fan, M., Roast, M., Verhulst, S., & Peters, A. *In press*. Early life telomere length predicts lifespan and lifetime reproductive success in a wild bird. *Molecular Ecology*, MEC-18-0689, accepted on 7-Dec-18.

Roast, M., Aulsebrook, A., Fan, M., Hidalgo Aranzamendi, N., Teunissen, N., & Peters, A. Temperature and rainfall as short-term drivers of individual variation in constitutive innate immunity in a tropical bird? *Physiological and Biochemical Zoology*, PBZ-18062R1, accepted on 10-Dec-18 pending minor revisions.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer-reviewed journals. The core theme of the thesis is behavioural and evolutionary ecology. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the School of Biological Sciences under the supervision of A/Prof Anne Peters and Dr Kaspar Delhey.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of Chapters 2 and 3, my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution	Co-author(s), Monash student Y/N
2	No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy?	Published	60%. Concept, data analysis, data collection and writing of manuscript	15% A. Peters: concept, design, data analysis and edits to manuscript 7.5% K. Delhey: concept, data analysis and edits to manuscript 7.5% M. L. Hall: data collection and edits to manuscript 5% S. A. Kingma: data collection and edits to manuscript 2.5% L. M. Mandeltort: data analysis and comment on manuscript 2.5% N. Hidalgo Aranzamendi: data collection and edits to manuscript	No No No No Yes Yes
3	From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird	Published	60%. Concept, design, data collection, data analysis and writing of manuscript	15% A. Peters: concept, design, data analysis and edits to manuscript 10% K. Delhey: concept, design, data analysis and edits to manuscript 5% N. Teunissen: design, data collection and edits to manuscript	No No Yes

				2.5% M. L. Hall: data collection and edits to manuscript 2.5% N. Hidalgo Aranzamendi: data collection and edits to manuscript 2.5% S. A. Kingma: data collection and edits to manuscript 2.5% M. Roast: data collection and edits to manuscript	No Yes No Yes

I have renumbered sections of published papers in order to generate a consistent presentation within the thesis. Table and figure numbers however have not been changed and reflect chronological order within each chapter. Chapter structure and the formatting of tables, figures and references are in accordance with the requirements of the journal to which manuscripts have been submitted.

Student signature: [redacted] **Date:** 17.12.2018

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: [redacted] **Date:** 17.12.2018

Acknowledgements

First, and foremost, I would like to warmly thank Anne Peters and Kaspar Delhey for being absolutely fantastic supervisors! Working with them for the past three and half years has been a real pleasure and I could not have asked for better supervision. The excellent relationship we built over the years enabled me to have a very smooth and enjoyable PhD journey. They always provided the guidance, support and feedback that I needed, and I have learned a lot from their valuable expertise and experience in research. I thank them for all the time and effort they invested in me, which hopefully made me a better scientist. Special thanks to Anne for believing in my abilities to do a good job even though I had almost no previous knowledge on how to work with birds! It was a steep learning curve and much hard work, but I absolutely enjoyed every moment that I spent in the field and will never forget this incredible experience.

This study would not have been possible without a remarkable group of co-authors. First, I would like to acknowledge Michelle L. Hall and Sjouke A. Kingma who, thanks to their constant enthusiasm, extensive knowledge of fairy-wrens, constructive criticism and positive feedback, substantially improved the quality of my work. Many thanks to Nataly Hidalgo Aranzamendi for taking me under her wing and teaching me everything I needed to know about bird fieldwork during my first field season. I also want to thank Niki Teunissen and Michael Roast for sharing my PhD journey, both in the field where we had some epic adventures, and in the office where we had stimulating conversations about fairy-wrens and statistical analyses, but also in my personal life as amazing friends. I am also grateful to Liliana D'Alba and Matthew Shawkey who welcomed me at the University of Ghent in Belgium and took the time to train me to conduct microscopy analyses of feather micro- and nanostructure.

Having the opportunity to do fieldwork in Mornington Wildlife Sanctuary has been a great privilege. I thank the staff of Australian Wildlife Conservancy for providing this opportunity and for their constant help and support during each of my field seasons. The success of my fieldwork would not have been possible without the valuable assistance and hard work of several other lab members and volunteers: my gratitude to Chris Smith, Megan Milligan, Alessandro Turchi, Zachary Emery, Síofra Sealy, Sergio Nolazco and Laurent Lermusiaux, with whom I had a truly great time in the field!

Many thanks to all the members of Anne's research group for being great lab and office mates, in particular Alex McQueen for sharing with me the stress and joy of doing a PhD, and Justin Eastwood for the many hours spent in the lab extracting DNA for the rest of the group! I would also like to thank some friends I made before or during my PhD and who contributed to making my time in Australia a wonderful experience: among others are Clémentine Lasne, Laure Dumont, Sébastien Menant, Sonia Sanchez, Cathy Cavallo and Nancy van Nieuwenhove.

I am very grateful to Bob Wong, Tim Connallon, Martin Burd and Rohan Clarke for their support and valuable feedback on my research throughout my PhD. Receiving some external input from researchers working in different fields was particularly useful.

This work would not have been possible without the financial support of numerous institutions: my gratitude to the School of Biological Sciences at Monash University, the Australian Research Council and the Australian Wildlife Conservancy. Additional support was provided by the Holsworth Wildlife Endowment and the Rosemary Grant Research Award. Moreover, my PhD was supported by a Monash Graduate Scholarship (MGS) and a Monash International Postgraduate Research Scholarship (MIPRS).

Lastly, and most importantly, I could not have done this PhD without the support of my family: my mum Lan and my sister Céline, whose constant love always encouraged me to give my best, and of course, my best friend / teammate / life partner, Laurent. Thank you for your understanding and patience every time I had to go away for fieldwork, for the many hours spent discussing fairy-wren science, and for always believing in me and pushing me to surpass myself. Finally, thank you for giving me the two most beautiful gifts of my life, who have been as understanding and well-behaved as they could be (!) during the completion of this work.

General introduction

Background

The evolution and maintenance of elaborate, conspicuous male ornamentation – that appears detrimental to survival – is generally attributed to sexual selection (Darwin, 1871): males with more elaborate ornamentation are more successful in competition for mates or breeding resources. This could come about by two basic means, either by being preferred in female mate choice or by being more successful at male-male competition for access to mates. It is now well established that females of many species choose their mates, and that they generally prefer more elaborate or more colourful males. Likewise, when males compete directly for access to females, or indirectly for breeding resources such as territories, such contests are often settled on the basis of ornamental traits, with more elaborate males winning (reviewed in Andersson, 1994; Senar, 2006; Santos et al., 2011). Although both mechanisms of sexual selection are credited with driving the evolution of elaborate male characters, most research has focused on female mate choice (Jones & Ratterman, 2009; McCullough et al., 2016). Moreover, the two mechanisms are often viewed as distinct and operating in isolation, characterised by different underlying processes and costs, and selecting for different types of traits (Lachmann et al., 2001; Hurd & Enquist, 2005; Jones & Ratterman, 2009; McCullough et al., 2016).

Nevertheless, some ornamental traits may have a dual function, being selected by both mechanisms that interact in a synergistic or conflicting way (Berglund et al., 1996; Qvarnström & Forsgren, 1998; Wong & Candolin, 2005; Hunt et al., 2009). Dual function-traits are thought to arise through male-male competition, and subsequently be co-opted for use in female choice, as proposed by the 'armament-ornament model' (Berglund et al., 1996). According to this model, females should exploit signals used in male-male aggressive interactions because their reliability is constantly tested in these interactions and cannot be faked without incurring high

costs (Berglund et al., 1996). However, the reverse process – female choice cues co-opted for use in male contests – may also occur, although there is very little evidence for it and it is less clear by which mechanisms such co-option would occur (but see Morris et al., 2007). Determining the context in which a dual function-trait initially evolved can be challenging as it requires studying both mechanisms in closely related species and phylogenetic information; hence, very few studies have been able to do so (Borgia & Coleman, 2000; Morris et al., 2007). Regardless of the direction of co-option, owing to potential differences between either mechanism (e.g., timing, strength and form of selection), as well as environmental heterogeneity, selection through both mechanisms may fluctuate and as a result, dual-function traits may be subject to complex sexual selection dynamics (Hunt et al., 2009; Miller & Svensson, 2014).

Whether females evaluating prospective mates, or males evaluating prospective rivals, can make informed decisions based on ornamental traits rests upon the potential of ornaments to signal individual quality. How well an ornament is developed may inform about the bearer's overall phenotypic and genetic constitution (Zahavi, 1975; Hamilton & Zuk, 1982; Andersson, 2006), such that individuals with more elaborate ornaments are expected to be better in a range of aspects than those with less elaborate ornaments (e.g., physical condition, parental care ability, immune function, etc.; Andersson, 1994). Such quality indicators require condition-dependence of their expression, as when their production and maintenance are costly (Zahavi, 1975; Johnstone, 1995; Cotton et al., 2004). Such costliness may arise, for example, in bright colours that make individuals more susceptible to predation (Endler, 1980; Kodric-Brown & Brown, 1984; Andersson, 1994; Huhta et al., 2003), in traits that contain rare and important nutrients (e.g., carotenoids; McGraw, 2006), require sustained care (Griggio et al., 2010) or generally depend on the overall condition of the individual (Andersson, 1994). Alternatively, female preference for more ornamented males may be explained by the Fisherian model (Fisher,

1930). According to this model, such female preferences can spread in populations because the genes responsible for preferences in females are transmitted to male progeny that also inherit the preferred ornaments from their fathers. Such ornamented males will achieve higher mating success as they will be more successful in attracting choosy females. As a result, a genetic correlation develops between the ornament and the preference, leading to further exaggeration of both. This process will continue until the trait is constrained by natural selection.

Strong directional selection operating on ornamental traits should drive beneficial alleles to fixation and as a result, lead to the erosion of underlying genetic variation and cause reduced phenotypic variability (Taylor & Williams, 1982; Merilä & Sheldon, 1999). The maintenance of genetic variance in sexually selected traits in the face of directional selection, generally from female preferences, is an ongoing evolutionary conundrum, known as the lek paradox, and has been debated for decades (Taylor & Williams, 1982; Pomianskowski & Møller, 1995; Kotiaho et al., 2001, 2008). Variability may be maintained by temporally or spatially variable patterns of selection (Chaine & Lyon, 2008; Fargevieille et al., 2017) or frequency-dependent selection (Dale, 2006), but it is unclear how common these types of selection are. A hypothesis that is broadly applicable is the genic capture model which suggests that (1) the expression of sexually selected traits is condition-dependent, and (2) condition itself depends on genes at multiple loci and therefore shows high genetic variance (Rowe & Houle, 1996; Kotiaho et al., 2001; Tomkins et al., 2004). Although the expression of many traits may reflect an individual's condition to some extent, ornamental traits are expected to show heightened condition-dependence compared to non-ornamental traits, because of stronger sexual selection, which should lead to greater exaggeration and higher sensitivity to condition (Rowe & Houle, 1996; Cotton et al., 2004; Bonduriansky & Rowe, 2005; but see Johnstone et al., 2009). Evidence for this prediction remains however scarce as a small number of studies have used suitable non-ornamental 'controls' as a comparison (Cotton et al., 2004; Bonduriansky & Rowe, 2005; Gosden &

Chenoweth, 2011). Furthermore, although several resolutions of the lek paradox have been proposed, the issue of the heritability of ornamental expression remains contentious. Indeed, heritability is defined as the ratio of additive genetic variance to total phenotypic variance, which also includes non-additive genetic and environmental sources of variation (Merilä & Sheldon, 1999). Consequently, traits exhibiting large additive genetic variance may not necessarily be highly heritable if, for instance, they also show substantial environmental variance (Price & Schluter, 1991; Tibbetts, 2010). So far, empirical work on signal heritability has been equivocal, as a large range of heritability estimates – from low to high – have been found for various sexually selected traits, but also because the data are relatively sparse and full of uncertainties (Pomiankowski & Møller, 1995; Merilä & Sheldon, 1999; Hadfield et al., 2007; Tibbetts, 2010; Charmantier et al., 2017). Nonetheless, a general expectation is that if a trait signals genetic benefits to the intended receivers, then its expression should be heritable.

The evolution of male ornaments in species where male variance in reproductive success is low appears more perplexing. In polygynous and lekking species, where males show larger variance in mating success, more elaborate males generally obtain more mates (Andersson, 1994; Dunn et al., 2001). Why sexual ornaments are also common in socially monogamous species requires a different explanation (Kraaijeveld et al., 2007). Darwin (1871), and later Fisher (1930), suggested that the presence of male sexual ornaments in monogamous species was associated with condition-dependent timing of breeding in females, a hypothesis deemed plausible based on various genetic models (O'Donald, 1980; Kirkpatrick et al., 1990). In genetically monogamous species that form long-term pair bonds, males might have fewer opportunities to advertise for mates, and such species are often characterised by biparental care coupled with cooperative defence of territories (Tobias et al., 2012). Mutual choice and social competition for non-sexual resources (territory, food) might therefore provide an explanation for the development of ornamentation (often, but not necessarily, mutual) in monogamous

species, but little research has been done to explore this possibility (Tobias et al., 2011, 2012). An alternative, that has received far more attention, is that many socially monogamous species are not genetically monogamous, and that extra-pair (EP) mating increases skew in male success, and female choice for more ornamented males accounts for the prevalence of ornamental traits among males (Owens & Hartley, 1998; Dunn et al., 2001; Forstmeier et al., 2014; Arct et al., 2015).

Fairy-wrens (*Malurus* spp.) provide excellent examples of female (EP) mate choice, strong male bias in EP reproductive success, signalling and condition-dependence of seasonal plumages. Male plumage varies considerably both across and within species, including in hue and extent (i.e. proportion of the body covered) of ornamental colouration (Fig. 1), and levels of dichromatism differ across species as well (Johnson et al., 2013). More particularly, almost all Australian fairy-wrens display seasonal plumages, with males alternating between dull non-breeding plumage and colourful breeding plumage (Peters et al., 2013). Possibly, seasonal plumages have evolved to reduce the costs of producing and maintaining a colourful plumage (e.g., reduced predation risk), although the additional annual process of moulting is likely to entail substantial costs in itself (Lindström et al., 1993; Peters, 2000; Peters et al., 2001). Seasonal plumages encompass several potential signalling possibilities: the timing of plumage production, the extent of coloured plumage, and the quality (hue, saturation) of the colour itself (Mulder & Magrath, 1994; Peters et al., 2013). Each of these plumage components may vary with some intrinsic (sex, age, genetic quality, etc.) and ecological factors (climate, habitat, etc.), and convey specific information about individual quality. They may operate independently, each signalling a single aspect of quality, or all together to provide a better overall indication of quality (Møller & Pomiankowski, 1993). Therefore, as multi-component signals, seasonal plumages constitute complex traits that may inform about various aspects of individual quality, possibly to the same or different receivers, and impact individual fitness in different ways.

A defining feature of the mating system of fairy-wrens is extra-pair paternity (EPP) that can reach very high levels (> 50% of brood in several species; Kingma et al., 2009; Peters et al., 2013; Brouwer et al., 2017; Fig. 1). In several fairy-wren species, it has been established that the timing of pre-breeding moult – typically highly variable between males, and often age- or climate-dependent, or both – is a signal used by females to assess the quality of EP mates (Mulder & Magrath, 1994; Karubian, 2002; Cockburn et al., 2008; Webster et al., 2008; van de Pol et al., 2012; Peters et al., 2013). The most detailed information comes from studies in superb fairy-wrens for which early moult has been shown to be a strong predictor of male EP success, with males acquiring their breeding plumage before or during winter being the most successful in obtaining EP fertilisations (Dunn & Cockburn, 1999; Green et al., 2000; Double & Cockburn, 2003). However, the strength of directional selection varies considerably between years, depending on the climatic conditions (Cockburn et al., 2008). Furthermore, a similar but weaker pattern has been found in red-winged fairy-wrens: males that moult earlier tend to achieve higher EP success, although variation in success is not substantial, possibly because of the importance of EP mating for inbreeding avoidance in this species (Brouwer et al., 2011). In addition, studies in the red-backed fairy-wren demonstrated that bright males sire significantly more EP young than dull males (Karubian, 2002; Webster et al., 2008), and among mostly bright males (> 66% of the body covered in breeding plumage), increased extent of breeding plumage is associated with increased EP success (Karubian et al., 2009). Nevertheless, aside from these knowledge on moult timing, the potential for seasonal plumages to act as multi-component signals of individual quality has only been explored to a limited extent in the genus, and we still know little about the evolutionary significance of other signalling components, including the quality of the plumage colour, or how the different colours are produced (Peters et al., 2013).

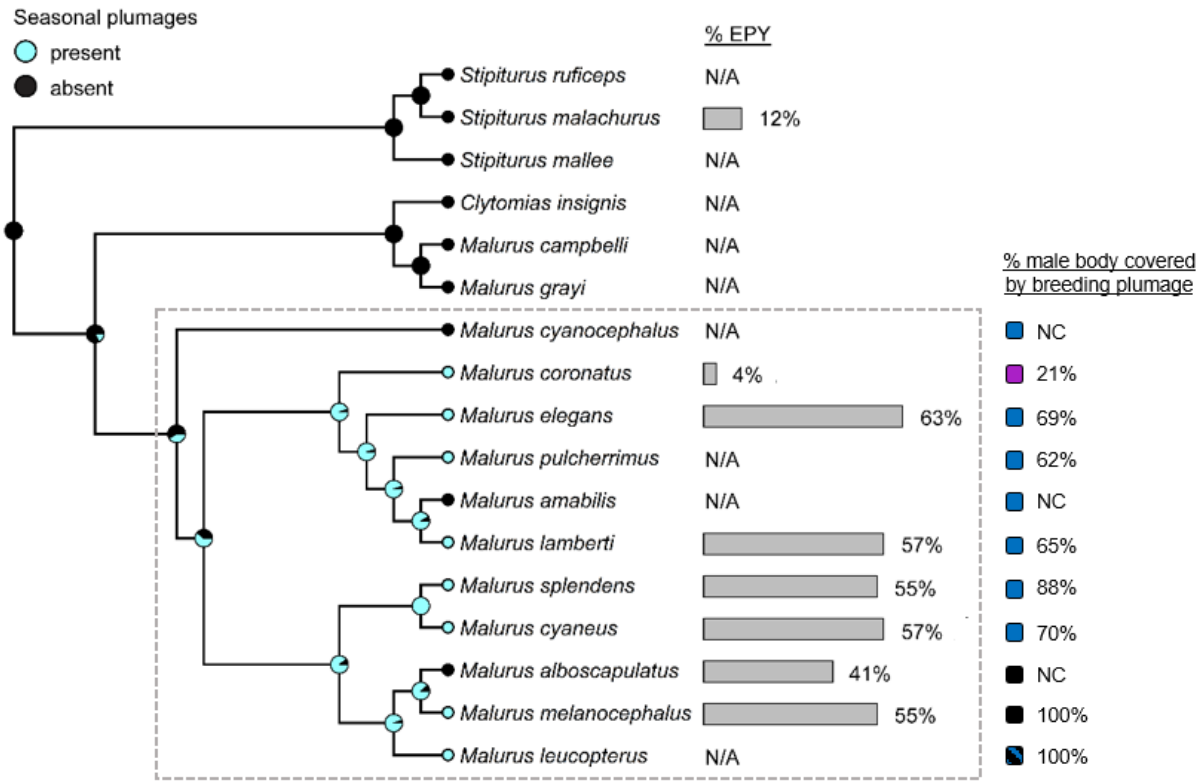


Figure 1. Seasonal plumage and extra-pair paternity (EPP) in fairy-wrens and emu-wrens. Shown is the ancestral state reconstruction of changes in seasonal plumages for 17 Malurid species, using stochastic mapping (for details see Chapter 2) and based on the supermatrix phylogeny of Marki et al. (2017). Absence of seasonal plumages in *M. amabilis* is based on Schodde (1982). % EPY = levels of EPP (proportion of offspring sired by extra-pair males) are shown for each species (Brouwer et al., 2017; N/A: no data available). For fairy-wrens (inside the grey dotted box) are also shown the proportion of male body covered by the breeding plumage (NC = no seasonal change in appearance, with males in colourful plumage year-round; Peters et al., 2013), as well as the dominant hue of the breeding plumage (blue, black or purple; *M. leucopterus* has both blue and black subspecies).

Thesis aim

In this thesis, I first compare the mechanisms of male plumage colour production (i.e. feather morphology) among multiple species of fairy-wren. Additionally, I use the purple-crowned

fairy-wren, *Malurus coronatus*, as a model species to investigate the function of male seasonal plumage ornaments independent of female choice for EPP. More specifically, I aim to determine the genetic and environmental causes of variations in the expression of male seasonal breeding plumage, as well as the life-history consequences of such variations. I examine three potential signalling components: (i) the timing of acquisition of the breeding plumage, (ii) breeding plumage completeness (or the extent of breeding plumage), and (iii) the colour quality (reflectance) of the breeding plumage.

Study species: the puzzling case of the purple-crowned fairy-wren

The purple-crowned fairy-wren *Malurus coronatus* is a small Australian songbird. Two subspecies have been distinguished, of which I studied *M. coronatus coronatus*, and I will refer to this subspecies for the remainder of this thesis.

M. coronatus coronatus is endemic to the monsoonal tropics of north-west Australia. It is restricted to patchy riparian vegetation of *Pandanus aquaticus* and maintains all-purpose territories year-round, linearly arranged along rivers and creeks (Rowley & Russell, 1997; Fig. 2). Like all fairy-wrens, it is highly philopatric and breeds cooperatively, with dominant breeders often (40-70% of pairs) assisted by a number of non-breeding male and female subordinates (mostly offspring from previous broods; Rowley & Russell, 1997; Kingma et al., 2010, 2011 a, b). Subordinate individuals may acquire a breeder position either by taking over part of the natal territory or establishing a new territory, or by filling a vacancy left by a deceased same-sex breeder (either inheritance of the home territory or dispersal to another, mostly limited to neighbouring territories – Kingma et al., 2011b; Hidalgo Aranzamendi et al., 2016). The species can breed year-round (in response to rainfall; Hidalgo Aranzamendi, 2017), with a peak in breeding activity at the study site during the wet season (December-March), and

a smaller peak in the late dry season (August-September) in some years (Hall & Peters, 2009; Peters et al., 2013).

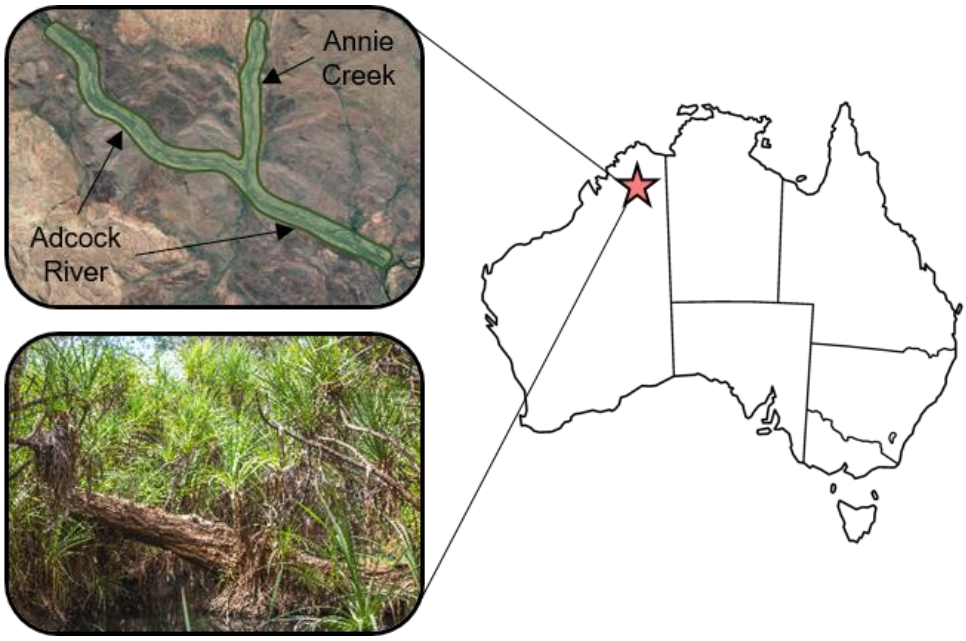


Figure 2. Field study site. Shown are the location of Australian Wildlife Conservancy's Mornington Wildlife Sanctuary (pink star) and photographs of the study site – top (Google maps): the green polygon depicts the study area composed of Annie Creek and the Adcock River along which fairy-wren territories are linearly arranged; bottom (L. Lermusiaux): *Pandanus aquaticus* vegetation.

Although the life-history and ecology of *M. coronatus* is very similar to other fairy-wrens, the species appears to be different in several aspects. Most prominently, it shows very low rates of EP mating (in only 5% of broods; Kingma et al., 2009, 2013; Hidalgo Aranzamendi et al., 2016; Fig. 1) and partners form long-term genetically monogamous partnerships, characterised by close pair cohesion and collaboration (Hall & Peters, 2008, 2009). Nevertheless, all males (dominant breeders and subordinates) undergo a pre-breeding moult once per year, where the dull brown non-breeding head plumage is replaced by purple and black feathers (Peters et al., 2013; Fig. 3). Other plumage patches do not change noticeably in colouration over the year,

including the black cheek patches and the blue tail, and the rest of the plumage which is mainly brown above and buff-white below (Rowley & Russell, 1997; Delhey et al., 2013; Fig. 3). Most males initiate pre-breeding moult in July-September before breeding starts, but in some cases this moult overlaps temporally with breeding (Rowley & Russell, 1997; Peters et al., 2013) and males can breed in brown plumage (6% of cases; Chapter 2).

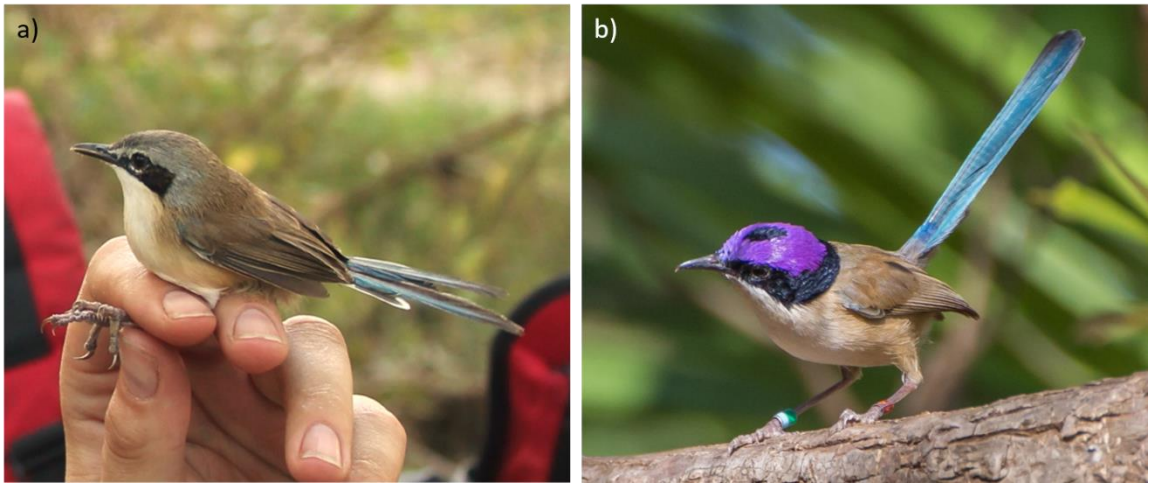


Figure 3. Non-breeding and breeding plumages in male purple-crowned fairy-wrens. Photographs show a male in (a) non-breeding plumage and (b) breeding plumage. Photos: (a) N. Hidalgo Aranzamendi, and (b) L. Lermusiaux.

Given that extreme levels of EPP appear to drive variation in the timing of pre-breeding moult in other species, it is important to ask what the function of the seasonal moult in male purple-crowned fairy-wrens could be. Was the trait co-opted to signal some aspects of male quality in another context, either to the social female or to potential rivals? Or do other components of the plumage have a signalling function that would explain the persistence of this ornamental plumage? Moreover, male ornamentation appears to be reduced compared to other fairy-wrens: is this the result of weaker sexual selection or other selective forces or constraints preventing the development of ornamental traits in this species? The presence of a seasonal breeding plumage in this genetically monogamous bird makes this species a puzzling and

unusual case among other studied fairy-wrens, and challenges current sexual selection theories (Peters et al., 2013). The study of such a model species in comparison with other members of the same genus, provides an excellent opportunity to test various evolutionary theories on multiple sexual traits and assess the potential for seasonal plumages to act as complex multi-component signals of individual quality. It will also help to further our understanding of the mechanisms and evolutionary forces underlying the functional flexibility (i.e. gain, loss or co-option of function) of sexual ornaments within and between species, offering possible new explanations for the diversity of sexual ornaments and functions.

Thesis outline

The main aim of this thesis is to determine the function of male seasonal plumages in the purple-crowned fairy-wren *M. coronatus*. First, I investigate the proximate mechanisms (feather microstructure) underlying the unusual purple colour of male purple-crowned fairy-wrens – an evolutionary innovation in the genus, that is otherwise dominated by structural blue and black colours (Figs 1 and 4). To assess how unique the purple colour is at the nanometre scale, I analyse more broadly how feather microstructure relates to male breeding plumage colouration in multiple fairy-wren species (Chapter 1; Fig. 4). Then, I focus on three plumage components – namely the timing of breeding plumage acquisition (Chapter 2), breeding plumage completeness (Chapter 3), and the colour quality (reflectance) of the breeding plumage (Chapter 4) – to assess how and why seasonal breeding plumage as a multi-component signal varies between and within individuals, and how these variations relate to individual fitness (Fig. 4). Finally, I discuss the main findings of all chapters to summarise our current understanding of the function of male seasonal plumages in this species and provide more general implications regarding the evolution of sexual ornaments.

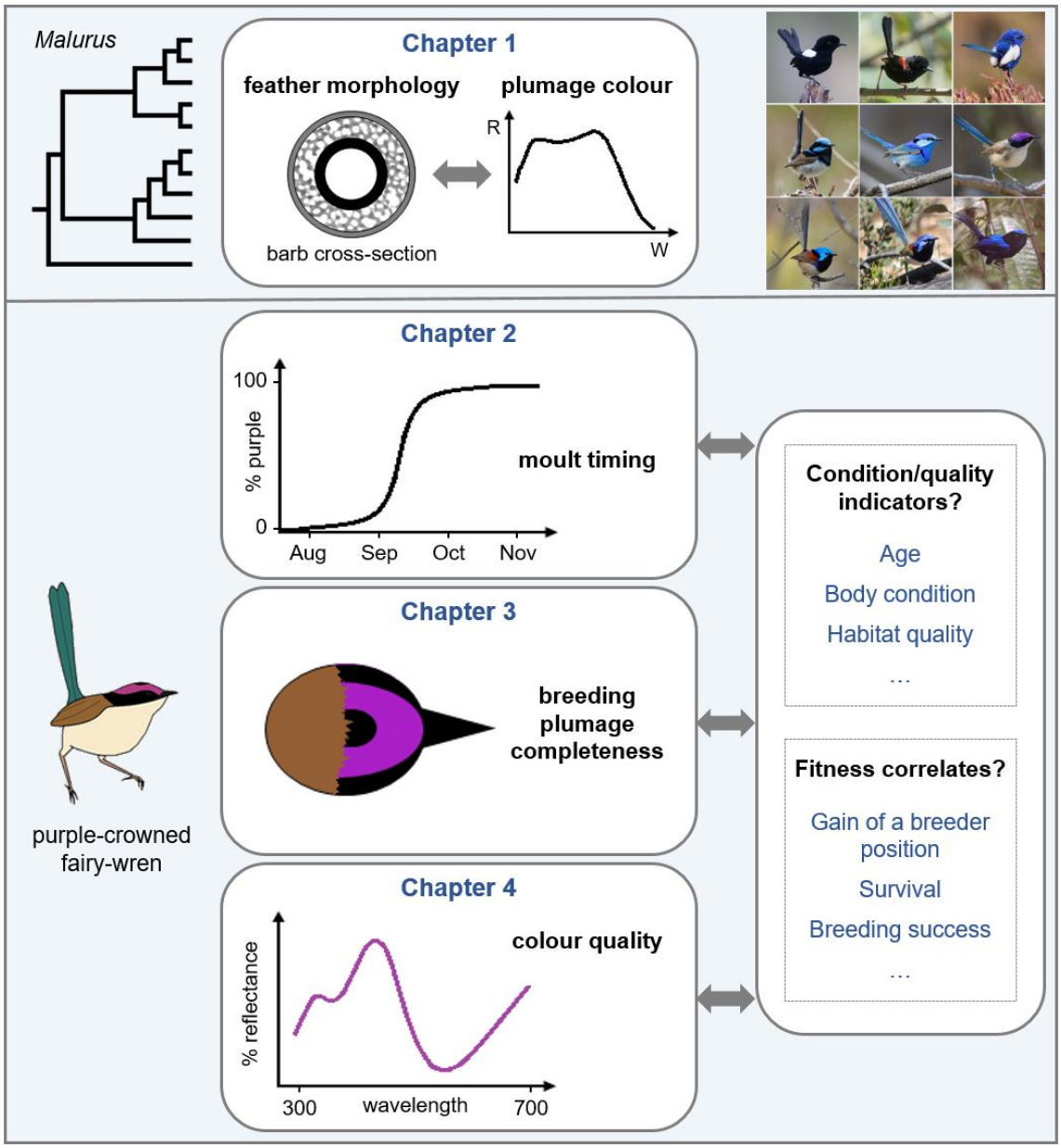


Figure 4. Thesis outline. Chapter 1 focuses on how variation in feather barb morphology at the nanoscale relates to male plumage colour in multiple fairy-wren species. Chapters 2 to 4 focus on how variation in respectively the timing of breeding plumage acquisition, breeding plumage completeness and colour quality, relates to male quality and fitness.

In **Chapter 1**, I focus on the mechanisms of production of the unique purple colour of male *M. coronatus*, as well as the other structural colours displayed in the genus for comparison. Closely related species often differ dramatically in the colour they display, and understanding

the mechanistic changes behind these interspecific differences can provide important insights into the evolution of colour divergence. The breeding plumages of male fairy-wrens are characterised by various non-iridescent blue, indigo and almost pure ultraviolet colours (Fig. 4). Non-iridescent structural plumage colouration is known to primarily arise from the interaction of light with the medullary 'spongy layer', a highly organised nanostructure found within feather barbs (Prum, 2006). Nevertheless, feather barbs contain several other components that may vary and interact with the spongy layer. Using n = 30 feather samples from structurally coloured plumage patches of multiple male fairy-wrens, I examine the relationship between barb morphology and the observed interspecific structural colour variation. This study reveals that the feather barbs producing the purple and other non-iridescent structural colours have a similar structure overall, and that the large diversity of structural colours results from variation in multiple microstructural components, rather than a single key component. This plurality of mechanisms most likely facilitates a greater variability in plumage colouration and may account for the rapid divergence of male ornamentation in this genus.

For the second part of my thesis, I focus on the function of the multiple components of male seasonal plumage ornaments in *M. coronatus* in relation to its unusual (for a fairy-wren) monogamous mating system.

In **Chapter 2**, I assess whether and how the timing of breeding plumage acquisition varies within and between males in relation to various intrinsic and environmental factors in order to test for condition-dependence. I also examine how these variations relate to multiple aspects of male fitness, systematically testing all adaptive hypotheses reflecting how intra- and intersexual selection can operate in monogamous mating systems (Fig. 4). Although there is virtually no opportunity for female EP mate choice in *M. coronatus*, all males annually produce a bright purple-and-black breeding plumage, and like other fairy-wrens, the timing of pre-breeding moult shows substantial variation between and within males (Rowley & Russell, 1997; Peters

et al., 2013). Using six years of data on timing of moult (n = 279 from 137 individuals), I investigate the adaptive significance of pre-breeding moult timing as a sexual signal under (near) genetic monogamy. In stark contrast with other members of the genus, there is no evidence for fitness benefits or costs associated with early moult. Based on phylogenetic information (see Marki et al., 2017; Fig. 1), I discuss the hypothesis of a trait loss scenario due to relaxed sexual selection in *M. coronatus*. This chapter contributes to our understanding of the evolutionary loss of sexually selected traits, an important, yet relatively unexplored area of sexual selection.

In Chapter 2, I also establish that breeder males moult earlier than subordinate males, which suggests that male breeding plumage might function as a social signal of dominance and competitiveness. Hence, in **Chapter 3**, I examine the role of the extent of breeding plumage in male-male social competition and male territoriality (Fig. 4). Because acquiring and retaining a breeder position is critical for male reproductive success in this species (subordinates very rarely reproduce until they gain a breeder position; Kingma et al., 2009), male-male competition may drive the persistence of the breeding plumage. Using detailed records of plumage completeness (n = 279 from 108 males) over six years, combined with an experimental study using realistically coloured 3D-printed models, I test (i) whether variation in breeding plumage of subordinate males predicts success in obtaining a breeder position, and (ii) whether plumage state of male intruders affects the strength of territorial defence by breeder males. My findings, in agreement with these hypotheses, highlight the crucial role played by male-male competition in a genus renowned for strong selection through female choice. Based on phylogenetic information, I propose two novel evolutionary scenarios with important implications regarding the gains and losses of function in sexually selected traits, providing further insights into the complex dynamics of sexual selection.

To address the question of whether the quality of the breeding colouration has a signalling function, I assess its variability, condition-dependence, heritability and fitness correlates in **Chapter 4** (Fig. 4). Since male *M. coronatus* display a multidimensional colour phenotype that includes several putative ornamental and non-ornamental colours (Fig. 3), I also assess these aspects for the other colours to test specific predictions regarding the adaptive significance of multiple ornamental colours in comparison with multiple non-ornamental colours (Bonduriansky & Rowe, 2005). In addition to the breeding colours, other plumage patches include a tail of a conspicuous blue colouration that is possibly ornamental, as well as buff-white and brown patches that are much more cryptic and most likely non-ornamental (Rowley & Russell, 1997; Fig. 3). Using eight years of plumage reflectance data (> 1000 spectra from over 140 males per colour), I quantify colour variation using psychophysical models of avian colour vision (Vorobyev et al., 1998) to take into account the perceptual ability of conspecifics. These data are used to test the predictions that the expression of ornamental colours is (i) more condition-dependent, (ii) more variable, (iii) more heritable, and (iv) more strongly related to male fitness compared to non-ornamental colours. This integrated assessment provides only partial support for these predictions, illustrating the complexity of visual signals which, in the future, should be investigated using more comprehensive approaches.

Fieldwork

To investigate the questions outlined above, I use long-term data from a colour-banded population of *Malurus coronatus coronatus* resident on Annie Creek and the Adcock River at the Australian Wildlife Conservancy's Mornington Wildlife Sanctuary (17º31'S, 126º6'E; north Western Australia; Fig. 2).

The long-term study of this population started in July 2005, and until November 2017, all birds were banded with a unique combination of coloured leg bands. Basic morphological

measurements and blood samples were taken from all birds at banding and any subsequent captures. From July 2005 to March 2011, weekly population censuses were conducted year-round to document individual survival, changes in group composition, dispersal events and individual plumage colouration (completeness of the breeding plumage). During these years, breeding attempts were also closely monitored (Kingma, 2011), and from 2005 to 2009, plumage colour (reflectance) measurements were collected year-round from captured birds.

From October 2011 to November 2017, biannual population censuses were conducted in October-November and May-June (respectively before and after the main breeding season) to document individual survival, changes in group composition, dispersal events and individual plumage colouration. In October-November 2015, 2016 and 2017, plumage reflectance measurements were collected from captured birds. Moreover, in May-June and October-November 2016, model presentation experiments were conducted.

Because of adult philopatry and delayed young dispersal, reproductive success during those years could be accurately assigned through year-round surveys between 2005 and early 2011, and biannual surveys – before and after the main breeding season – between late 2011 and 2017. Moreover, to minimise errors in survival estimates due to emigration, intensive yearly censuses were conducted from 2007 onwards to find birds that had dispersed outside the core area. Because purple-crowned fairy-wrens are habitat specialists, long distance dispersers could be found by highly efficient search techniques (> 90% detection) restricted to suitable riparian habitat (Hidalgo Aranzamendi et al., 2016). The estimates of the number of immigrants and emigrants for those years are similar (Eastwood et al., in press), which also suggests that the surveys outside the core study area did identify the majority of emigrants, and our estimates of survival and lifetime reproductive success for those years are highly accurate.

In summary, between 2005 and 2017, over two hundreds individuals were followed from birth until death, and detailed information on life-history strategies, neonatal conditions, age,

body condition, plumage reflectance and moult patterns were recorded. My PhD study is therefore based on existing data and newly collected data, combined with a detailed pedigree of the study population that currently spans seven generations.

Thesis organisation

This thesis is presented as a "thesis including published works" consisting of a general introduction, two peer-reviewed and published papers (Chapters 2 and 3), two chapters to be revised for submission to peer-reviewed journals (Chapters 1 and 4) and a general discussion. Chapter 2 is published in *Behavioral Ecology* and Chapter 3 in *Journal of Animal Ecology*. Although I was responsible for the planning, experimental design, data collection and analysis, and manuscript preparation for each chapter, the first-person plural is used in subsequent chapters to reflect the collaborative nature of my research.

References

Andersson, M. (1994). *Sexual selection.* Princeton University Press.

Andersson, M. (2006). Condition-dependent indicators in sexual selection: development of theory and tests. In J. R. Lucas & L. W. Simmons (Eds.), *Essays in Animal Behaviour: Celebrating 50 Years of Animal Behaviour* (pp. 255-269). Elsevier Academic Press.

Arct, A., Drobniak, S. M., & Cichoń, M. (2015). Genetic similarity between mates predicts extrapair paternity—a meta-analysis of bird studies. *Behavioral Ecology*, **26**, 959-968.

Berglund, A., Bisazza, A., & Pilastro, A. (1996). Armaments and ornaments: an evolutionary explanation of traits of dual utility. *Biological Journal of the Linnean Society*, **58**, 385-399.

Bonduriansky, R., & Rowe, L. (2005). Sexual selection, genetic architecture, and the condition dependence of body shape in the sexually dimorphic fly *Prochyliza xanthostoma* (Piophilidae). *Evolution*, **59**, 138-151.

Borgia, G., & Coleman, S. W. (2000). Co-option of male courtship signals from aggressive display in bowerbirds. *Proceedings of the Royal Society of London B: Biological Sciences*, **267**, 1735-1740.

Brouwer, L., van de Pol, M., Atema, E. L. S., & Cockburn, A. (2011). Strategic promiscuity helps avoid inbreeding at multiple levels in a cooperative breeder where both sexes are philopatric. *Molecular Ecology*, **20**, 4796-4807.

Brouwer, L., van de Pol, M., Aranzamendi, N. H., Bain, G., Baldassarre, D. T., Brooker, L. C., ... & Hall, M. L. (2017). Multiple hypotheses explain variation in extra-pair paternity at different levels in a single bird family. *Molecular Ecology*, **26**, 6717-6729.

Chaine, A. S., & Lyon, B. E. (2008). Adaptive plasticity in female mate choice dampens sexual selection on male ornaments in the lark bunting. *Science*, **319**, 459-462.

Charmantier, A., Wolak, M. E., Grégoire, A., Fargevieille, A., & Doutrelant, C. (2017). Colour ornamentation in the blue tit: quantitative genetic (co) variances across sexes. *Heredity*, **118**, 125.

Cockburn, A., Osmond, H. L., & Double, M. C. (2008). Swingin' in the rain: condition dependence and sexual selection in a capricious world. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 605-612.

Cotton, S., Fowler, K., & Pomiankowski, A. (2004). Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? *Proceedings of the Royal Society of London B: Biological Sciences*, **271**, 771-783.

Dale, J. (2006). Intraspecific variation in bird coloration. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1*. (pp. 36-86). Harvard University Press.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex.* John Murray.

Delhey, K., Hall, M., Kingma, S. A., & Peters, A. (2013). Increased conspicuousness can explain the match between visual sensitivities and blue plumage colours in fairy-wrens. *Proceedings of the Royal Society of London B: Biological Sciences*, **284**, 20121771.

Double, M. C., & Cockburn, A. (2003). Superb Fairy-wrens (*Malurus cyaneus*) parasitize the reproductive success of attractive dominant males. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 379-384.

Dunn, P. O., & Cockburn, A. (1999). Extrapair mate choice and honest signaling in cooperatively breeding Superb Fairy-wrens. *Evolution*, **53**, 938-946.

Dunn, P. O., Whittingham, L. A., & Pitcher, T. E. (2001). Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution*, **55**, 161-175.

Eastwood. J., Hall, M. L., Teunissen, N., Kingma, S. A., Hidalgo Aranzamendi, N., Fan, M., Roast, M., Verhulst, S., & Peters, A. *In press*. Early life telomere length predicts lifespan and lifetime reproductive success in a wild bird. *Molecular Ecology*.

Endler, J. A. (1980). Natural selection on color patterns in *Poecilia reticulata*. *Evolution*, **34**, 76-91.

Fargevieille, A., Grégoire, A., Charmantier, A., del Rey Granado, M., & Doutrelant, C. (2017). Assortative mating by colored ornaments in blue tits: space and time matter. *Ecology and Evolution*, **7**, 2069-2078.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Clarendon Press.

Gosden, T. P., & Chenoweth, S. F. (2011). On the evolution of heightened condition dependence of male sexual displays. *Journal of Evolutionary Biology*, **24**, 685-692.

Green, D. J., Osmond, H. L., Double, M.C., & Cockburn, A. (2000). Display rate by male fairy-wrens (*Malurus cyaneus*) during the fertile period of females has little influence on extra-pair mate choice. *Behavioral Ecology and Sociobiology*, **48**, 438-446.

Griggio, M., Hoi, H., & Pilastro, A. (2010). Plumage maintenance affects ultraviolet colour and female preference in the budgerigar. *Behavioural processes*, **84**, 739-744.

Forstmeier, W., Nakagawa, S., Griffith, S. C., & Kempenaers, B. (2014). Female extra-pair mating: adaptation or genetic constraint? *Trends in Ecology & Evolution*, **29**, 456-464.

Hadfield, J. D., Nutall, A., Osorio, D., & Owens, I. P. F. (2007). Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *Journal of Evolutionary Biology*, **20**, 549-557.

Hall, M. L., & Peters, A. (2008). Coordination between the sexes for territorial defence in a duetting fairy-wren. *Animal Behaviour*, **76**, 65-73.

Hall, M. L., & Peters, A. (2009). Do male paternity guards ensure female fidelity in a duetting fairy-wren? *Behavioral Ecology*, **20**, 222-228.

Hamilton, W. D., & Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites? *Science*, **218**, 384-387.

Hidalgo Aranzamendi, N., Hall, M. L., Kingma, S. A., Sunnucks, P., & Peters, A. (2016). Incest avoidance, extrapair paternity, and territory quality drive divorce in a year-round territorial bird. *Behavioral Ecology*, arw101.

Hidalgo Aranzamendi, N. (2017). Life-history variation in a tropical cooperative bird: Ecological and social effects on productivity (Doctoral dissertation).

Huhta, E., Rytkönen, S., & Solonén, T. (2003). Plumage brightness of prey increases predation risk: An among-species comparison. *Ecology*, **84**, 1793-1799.

Hunt, J., Breuker, C. J., Sadowski, J. A., & Moore, A. J. (2009). Male-male competition, female mate choice and their interaction: determining total sexual selection. *Journal of Evolutionary Biology*, **22**, 13-26.

Hurd, P. L., & Enquist, M. (2005). A strategic taxonomy of biological communication. *Animal Behaviour*, **70**, 1155-1170.

Johnson, A. E., Jordan Price, J., & Pruett-Jones, S. (2013). Different modes of evolution in males and females generate dichromatism in fairy-wrens (Maluridae). *Ecology and Evolution*, **3**, 3030-3046.

Johnstone, R. A. (1995). Sexual selection, honest advertisement and the handicap principle: reviewing the evidence. *Biological Reviews*, **70**, 1-65.

Johnstone, R. A., Rands, S. A., & Evans, M. R. (2009). Sexual selection and condition-dependence. *Journal of Evolutionary Biology*, **22**, 2387-2394.

Jones, A. G., & Ratterman, N. L. (2009). Mate choice and sexual selection: what have we learned since Darwin? *Proceedings of the National Academy of Sciences*, **106**, 10001-10008.

Karubian, J. (2002). Costs and benefits of variable breeding plumage in the Red-backed Fairy-wren. *Evolution*, **56**, 1673-1682.

Karubian, J., Swaddle, J. P., Varian-Ramos, C. W., & Webster, M. S. (2009). The relative importance of male tail length and nuptial plumage on social dominance and mate choice in the

Red-backed Fairy-wren *Malurus melanocephalus*: evidence for the multiple receiver hypothesis. *Journal of Avian Biology*, **40**, 559-568.

Kingma, S. A., Hall, M. L., Segelbacher, G., & Peters, A. (2009). Radical loss of an extreme extra-pair mating system. *BMC Ecology*, **9**, 15.

Kingma, S. A., Hall, M. L., Arriero, E., & Peters, A. (2010). Multiple benefits of cooperative breeding in purple-crowned fairy-wrens: A consequence of fidelity? *Journal of Animal Ecology*, **79**, 757-768.

Kingma SA. 2011. The evolution of the social and genetic mating system of purple-crowned fairy-wrens *Malurus coronatus* (Doctoral dissertation).

Kingma, S. A., Hall, M. L., & Peters, A. (2011a). No evidence for offspring sex-ratio adjustment to social or environmental conditions in cooperatively breeding purple-crowned fairy-wrens. *Behavioral and Ecological Sociobiology*, **65**, 1203-1213.

Kingma, S. A., Hall, M. L., & Peters, A. (2011b). Multiple benefits drive helping behavior in a cooperatively breeding bird: an integrated analysis. *The American Naturalist*, **177**, 486-495.

Kingma, S. A., Hall, M. L., & Peters, A. (2013). Breeding synchronization facilitates extrapair mating for inbreeding avoidance. *Behavioral Ecology*, **24**, 1390-1397.

Kirkpatrick, M., Price, T., & Arnold, S. J. (1990). The Darwin-Fisher theory of sexual selection in monogamous birds. *Evolution*, **44**, 180-193.

Kodric-Brown, A., & Brown, J. H. (1984). Truth in advertising: the kinds of traits favored by sexual selection. *The American Naturalist*, **124**, 309-323.

Kotiaho, J. S., Simmons, L. W., & Tomkins, J. L. (2001). Towards a resolution of the lek paradox. *Nature*, **410**, 684.

Kotiaho, J. S., LeBas, N. R., Puurtinen, M., & Tomkins, J. L. (2008). On the resolution of the lek paradox. *Trends in Ecology & Evolution*, **23**, 1-3.

Kraaijeveld, K., Kraaijeveld-Smit, F. J. L., & Komdeur, J. (2007). The evolution of mutual ornamentation. *Animal Behaviour*, **74**, 657-677.

Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, **98**, 13189-13194.

Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., ... & Foster, S. A. (2009). Relaxed selection in the wild. *Trends in Ecology & Evolution*, **24**, 487-496.

Lindström Å., Visser, G. H., & Daan, S. (1993). The energetic cost of feather synthesis is proportional to basal metabolic rate. *Physiology Zoology*, **66**, 490-510.

Marki, P. Z., Jønsson, K. A., Irestedt, M., Nguyen, J. M., Rahbek, C., & Fjeldså, J. (2017). Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, **107**, 516-529.

McCullough, E. L., Miller, C. W., & Emlen, D. J. (2016). Why sexually selected weapons are not ornaments. *Trends in Ecology & Evolution*, **31**, 742-751.

McGraw, K. J. (2006) Mechanics of melanin coloration in birds. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 243-294). Harvard University Press.

Merilä, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity*, **83**, 103-109.

Miller, C. W., & Svensson, E. I. (2014). Sexual selection in complex environments. *Annual Review of Entomology*, **59**, 427-445.

Møller, A. P., & Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology*, **32**, 167-176.

Morris, M. R., Tudor, M. S., & Dubois, N. S. (2007). Sexually selected signal attracted females before deterring aggression in rival males. *Animal Behaviour*, **74**, 1189-1197.

Mulder, R. A., & Magrath, M. J. L. (1994). Timing of prenuptial molt as a sexually selected indicator of male quality in superb fairy-wrens (*Malurus cyaneus*). *Behavioral Ecology*, **5**, 393-400.

O'Donald, P. (1980). *Genetic models of sexual selection* (Vol. 44). Cambridge University Press.

Owens, I. P. F., & Hartley, I. R. (1998). Sexual dimorphism in birds: why are there so many different forms of dimorphism? *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 397-407.

Peters, A. (2000). Testosterone treatment is immunosuppressive in superb fairy–wrens, yet free–living males with high testosterone are more immunocompetent. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **267**, 883-889.

Peters, A., Astheimer, L. B., & Cockburn, A. (2001). The annual testosterone profile in cooperatively breeding superb fairy-wrens, *Malurus cyaneus*, reflects their extreme infidelity. *Behavioral Ecology and Sociobiology*, **50**, 519-527.

Peters, A., Kingma, S. A., & Delhey, K. (2013). Seasonal male plumage as a multi-component sexual signal: Insights and opportunities. *Emu*, **113**, 232-247.

Pomiankowski, A., & Møller, A. P. (1995). A resolution of the lek paradox. *Proceedings of the Royal Society of London B: Biological Sciences*, **260**, 21-29.

Price, T., & Schluter, D. (1991). On the low heritability of life-history traits. *Evolution*, **45**, 853-861.

Prum, R. O. (2006). Anatomy, physics, and evolution of structural colours. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 295-353). Harvard University Press.

Qvarnström, A., & Forsgren, E. (1998). Should females prefer dominant males? *Trends in Ecology & Evolution*, **13**, 498-501.

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London B: Biological Sciences*, **263**, 1415-1421.

Rowley, I., & Russell, E. (1997). *Fairy-wrens and Grasswrens: Maluridae.* Oxford University Press.

Santos, E. S., Scheck, D., & Nakagawa, S. (2011). Dominance and plumage traits: meta-analysis and metaregression analysis. *Animal Behaviour*, **82**, 3-19.

Schodde, R. (1982). *The Fairy-wrens: a monograph of the Maluridae.* Lansdowne Editions.

Senar, J. (2006). Color displays as intrasexual signals of aggression and dominance. In G. E. Hill & K. J. McGraw (Eds.), *Bird Coloration, vol. 2.* (pp. 87-136). Harvard University Press.

Taylor, P. D., & Williams, G. C. (1982). The lek paradox is not resolved. *Theoretical Population Biology*, **22**, 392-409.

Tibbetts, E. A. (2010). The condition dependence and heritability of signaling and nonsignaling color traits in paper wasps. *The American Naturalist*, **175**, 495-503.

Tobias, J. A., Gamarra-Toledo, V., Garcia-Olaechea, D., Pulgarin, P. C., & Seddon, N. (2011). Year-round resource defence and the evolution of male and female song in suboscine birds: social armaments are mutual ornaments. *Journal of Evolutionary Biology*, **24**, 2118-2138.

Tobias, J. A., Montgomerie, R., & Lyon, B. E. (2012). The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 2274-2293.

Tomkins, J. L., Radwan, J., Kotiaho, J. S., & Tregenza, T. (2004). Genic capture and resolving the lek paradox. *Trends in Ecology & Evolution*, **19**, 323-328.

van de Pol, M., Osmond, H. L., & Cockburn, A. (2012). Fluctuations in population composition dampen the impact of phenotypic plasticity on trait dynamics in superb fairy-wrens. *Journal of Animal Ecology*, **81**, 411-422.

Vorobyev, M., Osorio, D., Bennett, A. T., Marshall, N. J., & Cuthill, I. C. (1998). Tetrachromacy, oil droplets and bird plumage colours. *Journal of Comparative Physiology A*, **183**, 621-633.

Webster, M. S., Varian, C. W., & Karubian, J. (2008). Plumage colour and reproduction in the Red-backed Fairy-wren: why be a dull breeder? *Behavioral Ecology*, **19**, 517-524.

Wiens, J. J. (2001). Widespread loss of sexually selected traits: how the peacock lost its spots. *Trends in Ecology & Evolution*, **16**, 517-523.

Wong, B., & Candolin, U. (2005). How is female mate choice affected by male competition? *Biological Reviews*, **80**, 559-571.

Zahavi, A. (1975). Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, **53**, 205-214.

Chapter 1

Variation in multiple components of feather microstructure produces the highly diverse structural plumage colours of male fairy-wrens

Marie Fan[1], Liliana D'Alba[2], Matthew D. Shawkey[2], Anne Peters[1] & Kaspar Delhey[1,3]

[1]School of Biological Sciences, Monash University, 25 Rainforest Walk, Clayton, Victoria 3800, Australia

[2]Evolution and Optics of Nanostructures Group, Department of Biology, University of Ghent, Ledeganckstraat 35, Ghent 9000, Belgium

[3]Max Planck Institute for Ornithology, Vogelwarte Radolfzell, Am Obstberg 1, D-83175, Radolfzell, Germany

Abstract

In many animal clades, closely related species often differ in colouration. Understanding the mechanistic bases of such differences can provide insights into how the diversity of ornamentation patterns has evolved. Specifically, such studies can determine whether evolutionary changes in colour are driven by single key mechanisms, or are the consequence of changes in multiple underlying variables. Plumage colouration in birds is generally produced by a combination of pigmentary and nanostructural components of feather barbs and barbules. Light scattering by nanostructures is responsible for structural colouration, and non-iridescent structural plumage colour has been shown to be primarily produced by the medullary spongy layer, an array of keratin and air found within feather barbs. Nevertheless, barb morphology is complex, and other structural elements may vary and interact with the spongy layer to generate

interspecific colour variation. Male fairy-wrens (*Malurus* spp., Fam. Maluridae) display a large diversity of ornamental colours, many of them known to be sexually selected, and mostly varying from ultraviolet to blue and purple. Using spectrometry, transmission electron microscopy and Fourier analysis, we examined how different elements of feather barb structure relate to these divergent colour displays, as perceived by conspecifics. Our study revealed that barb morphology is similar among all species, with an outer keratin cortex surrounding a highly organised medullary spongy layer and a basal layer of melanin. Moreover, all these structural elements contribute to the observed interspecific colour variation. A thinner spongy layer, with tighter arrays of keratin and air, on top of thicker melanin layers, is associated with decreased reflectance in the shortwave range (less UV/blue plumage). Moreover, a more regular spongy layer generates colours of greater purity (more chromatic). Additionally, a thinner cortex and thicker spongy layer are responsible for brighter feathers, with greater overall reflectance. These findings indicate that the diversity of structural colours in the nuptial plumages of male fairy-wrens is a result of variation in multiple microstructural elements, rather than consistent changes in a single key mechanism. This plurality of mechanisms is likely to facilitate a greater variability in plumage colouration and may account for the rapid evolution of male ornamentation in this genus. If broadly applicable, this means that the diversity of non-iridescent structural colours – both within and between species – could be independently affected by various aspects of feather microstructure which in turn can be determined by multiple genetic and environmental factors.

Keywords: avian visual space, colour evolution, feather nanostructure, Fourier analysis, *Malurus*, ornamental plumage

Introduction

Ornamental colouration features prominently in animal communication, sexual selection and speciation (Darwin, 1871; Andersson, 1994; Espmark et al., 2000; Price, 2008; Maan & Seehausen, 2011). In many animal clades, closely related species differ considerably in the colour they display, and understanding the mechanistic changes behind these interspecific differences can provide important insights into the evolution of colour divergence. Bird plumages constitute some of the most diverse colour displays and have served as a classic model system to study the evolution of ornaments (Hill & McGraw, 2006). Feather colouration can be produced by two main mechanisms: the deposition of pigments (e.g., melanins and carotenoids) that selectively absorb some wavelengths of light while allowing others to be reflected (McGraw, 2006), or the physical interaction between light and biological tissues that vary periodically in refractive index at the nanometre scale (i.e. structural colouration; Prum, 2006). Structural colours include most blue, violet and ultraviolet (UV) hues, as well as iridescent colours (Auber, 1957; Dyck, 1976; Prum, 2006). Although pigmentary and structural mechanisms are often studied in isolation, many colours are created by the interactions of both components, and variations in either mechanism can cause changes in the resulting colouration (Shawkey & Hill, 2005, 2006; Prum, 2006; Shawkey et al., 2006; Driskell et al., 2010; Shawkey & D'Alba, 2017).

Structural colours are inherently linked to the nanostructural characteristics of the underlying morphology. While iridescent structural colours in birds are generally produced by arrays of melanin granules (Prum, 2006; Shawkey et al., 2009), non-iridescent blue, violet and UV colours primarily result from arrays of keratin and air within feather barbs, that form the medullary 'spongy layer' (Prum, 2006; Shawkey et al., 2009). Many studies on non-iridescent structural plumage colours have therefore focused on the role of the spongy layer, which constitutes the main colour-producing element in the barb and functions both at the intra- and

interspecific level (Prum et al., 1998, 1999; Shawkey et al., 2003, 2005, 2006). Nevertheless, the morphology of feather barb is complex, with multiple other structural elements that may vary and interact with the spongy layer. To date, how this morphological complexity may generate interspecific colour variation, through changes in multiple structural elements, has been rarely investigated, in particular in non-iridescent structural colours (for iridescent structural colours see Eliason et al., 2015). As far as we know, only one study in the eastern bluebird *Sialis sialis* has established a link between UV-blue colouration and the relative amount of keratin cortex in the barb (Shawkey et al., 2005). Moreover, although the presence of melanin beneath the spongy layer is essential for colour production – as it absorbs incoherently scattered white light that would otherwise wash out the coherently scattered colour (Prum, 2006; Shawkey & Hill, 2006; Shawkey & D'Alba, 2017), whether its arrangement or density affects structural colouration remains unclear.

Male fairy-wrens in the genus *Malurus* display a great variety of brightly coloured plumages, with colours covering the entire avian visual sensitivity range (300-700 nm; Cuthill, 2006), but notably characterised by blue, indigo and almost pure UV hues (Fig. 1). Most of these colourful plumages are seasonal (nuptial) and in several species they serve similar, sexually selected functions, playing a crucial role in either female choice or male-male competition (Cockburn et al., 2008; Brouwer et al., 2011; Peters et al., 2013; Fan et al., 2018 / Chapter 3). This large diversity of ornamental colours has evolved rapidly, driven by two recent transitions in fairy-wren visual system from violet-sensitive (V-type) to UV-sensitive (U-type) eyes (Wilkie et al., 2000; Ödeen et al., 2012; Friedman & Remeš, 2015). These shifts have been associated with the evolution of more shortwave-rich colours (UV/blue), most likely due to their increased conspicuousness to visual systems with high UV sensitivity (Ödeen et al., 2012; Delhey et al., 2013; Friedman & Remeš, 2015). Although some studies have confirmed the primary contribution of the medullary spongy layer to the production of bright blue feathers in two

species of fairy-wren (Doucet et al., 2004; Driskell et al., 2010), the mechanistic basis of the many other nuptial colours displayed in the genus, including the unusual purple colouration of male *M. coronatus* (Delhey et al., 2013), remains unknown. Fairy-wrens therefore constitute an ideal group for comparative analyses of non-iridescent structural colour evolution and feather morphology. Understanding the mechanistic causes of the great variability in plumage colouration can further inform on the rapid evolution of male ornamentation in this genus.

Our study aimed to identify the micro- and nanostructures responsible for the large colour diversity in nuptial plumage among male fairy-wrens. We sampled feather barbs from all structurally coloured plumage patches, and tested the hypothesis that colour variation among nine closely related species is caused by variations in feather barb structure. We first used reflectance spectrometry and psychophysical models of avian vision to assess colour variation in the visual space of birds. These visual models provide non-arbitrary descriptors of colour variation that are biologically relevant and allow colours of different hues to be compared in the same currency (Delhey et al., 2015). Then, we used transmission electron microscopy and Fourier analysis to investigate differences in feather barb nano- and microstructure and link these to colour variation in the visual space of birds.

Material and methods

Study species

To investigate how feather barb structure influences the expression of structural colours, we selected all *Malurus* species in which males display structural colours, i.e. blue, indigo and purple plumage patches. Nine of the 11 *Malurus* species were therefore chosen for this purpose: *M. cyanocephalus*, *M. coronatus*, *M. elegans*, *M. pulcherrimus*, *M. amabilis*, *M. lamberti*, *M. cyaneus*, *M. splendens* and *M. leucopterus* (Fig. S1). Some species included multiple subspecies with varying structural colours, and in those cases we included samples for the

different subspecies. For each taxon, we sampled between one and five differently coloured structural plumage patches. In addition, we compared these structurally coloured plumages with two other nuptial colours displayed in the genus, black and rufous, by examining the barb structure of rufous feathers from *M. elegans*, *M. pulcherrimus*, *M. amabilis* and *M. lamberti*, and black feathers from *M. coronatus*, *M. lamberti* and *M. splendens*, as well as *M. alboscapulatus* and *M. melanocephalus* (Fig. S1).

Figure 1. Male fairy-wrens (*Malurus* spp.) display a great diversity of nuptial plumages, including various structural colours. Shown are photographs of a male a) *M. cyaneus*, b) *M. splendens*, c) *M. alboscapulatus*, d) *M. lamberti*, e) *M. coronatus*, and f) *M. cyanocephalus*, in nuptial plumage. Photos: a) K. Delhey, b) and d) Macaulay Library, c) and f) E. Enbody, and e) L. Lermusiaux.

Feather sampling

Blue crown, throat, flank, rump and shoulder feathers were collected from male *M. cyanocephalus bonapartii*. Purple and black crown feathers were collected from male *M.*

coronatus coronatus. Blue crown and cheek and indigo throat feathers were collected from male *M. cyaneus cyaneus*. For male *M. elegans, M. pulcherrimus, M. amabilis,* and *M. lamberti* (including two subspecies *M. l. lamberti* and *M. l. assimilis*), blue crown and cheek feathers were collected, as well as blue throat feathers for *M. elegans* and *M. pulcherrimus,* and black throat feathers for *M. amabilis*, and *M. lamberti*. For male *M. splendens* (including the three subspecies *M. s. splendens*, *M. s. melanotus* and *M. s. callainus*), blue crown, cheek and throat, as well as black nape feathers were collected. Black crown and throat feathers were collected from male *M. alboscapulatus moretoni*. Collection details are provided in Table S1. Feathers were pulled from living birds and museum specimens using tweezers and stored in small manila envelopes at ambient temperature until the time of analysis. For *M. leucopterus leuconotus* and *M. melanocephalus*, we used data on feather structure from Driskell et al. (2010). Hereafter we use the term 'unique patch' to refer to a single plumage patch of a given taxon (e.g., the blue throat of *M. splendens splendens*).

Reflectance spectrometry and visual models

Reflectance of each unique patch was measured on museum specimens and living birds (n = 8-717 for each patch; for details see Table S2) using an AvaSpec-2048 spectrometer connected to an AvaLight-XE xenon pulsed light source (Avantes, Apeldoorn, Netherlands) with a bifurcated fibre optic cable fitted at the end and a cylindrical probe to standardise measuring distance and exclude ambient light. We held the probe perpendicular to the plumage's surface, and collected up to five reflectance spectra per plumage patch. Reflectance spectra between 300 and 700 nm were calculated relative to a WS-2 white standard using the software AVASOFT 7.5 (Avantes; Figs 2, S2 and S3).

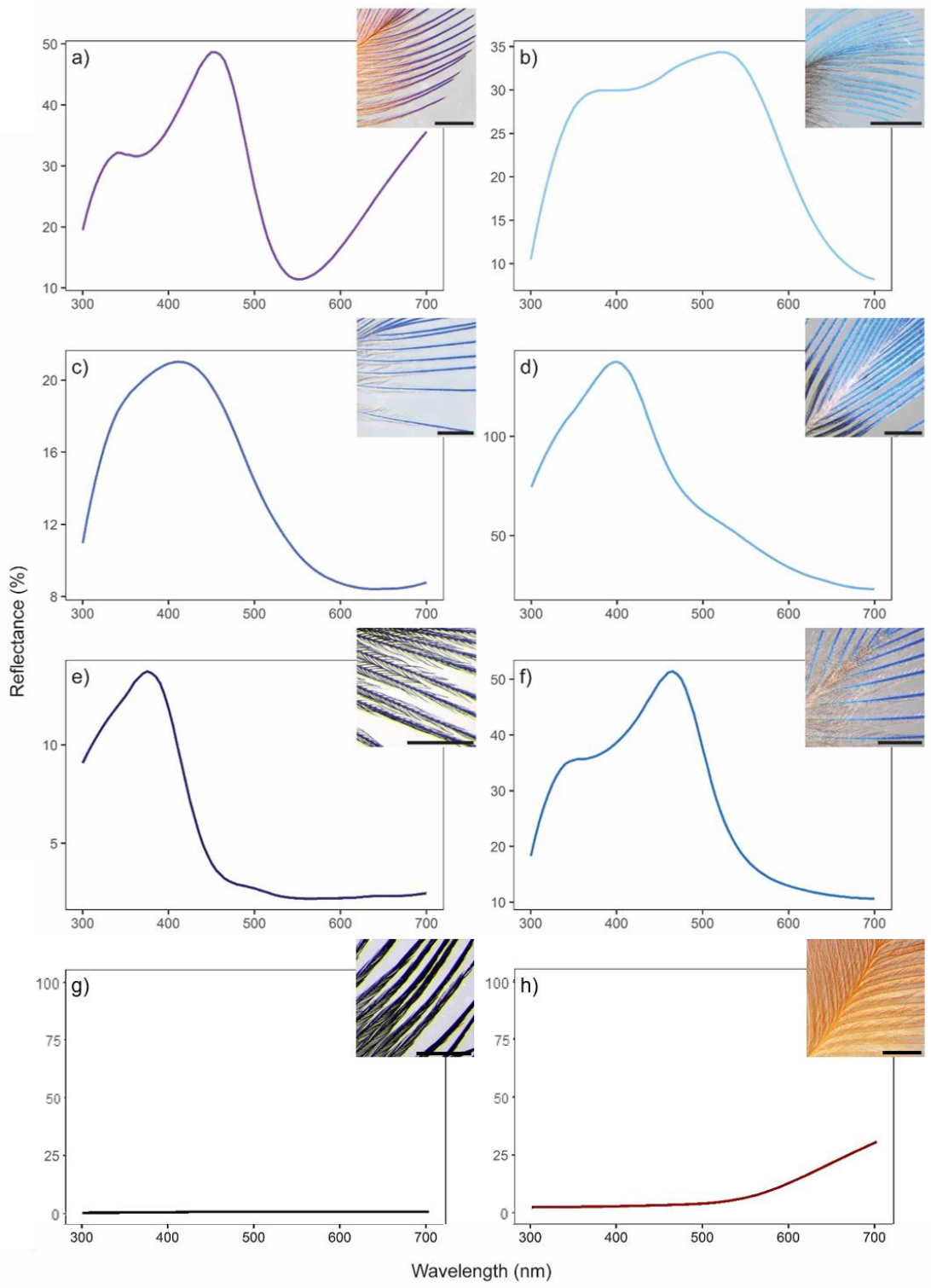


Figure 2. Examples of plumage reflectance spectra and associated feathers. Shown are the average (smoothed) reflectance spectra and detail of the barbs of a) the purple crown of male *M. coronatus coronatus*, b) the silvery blue crown of male *M. elegans*, c) the violet blue crown of male *M. lamberti assimilis*, d) the light blue cheek of male *M. cyaneus cyaneus*, e) the indigo

throat of male *M. cyaneus cyaneus*, f) the cobalt blue throat of male *M. splendens melanotus*, g) the black crown of male *M. alboscapulatus moretoni*, and h) the rufous shoulder of male *M. amabilis*. Line colours correspond approximately to the feather colour perceived by the human eye. Scale bars: a, c, e, f, g, h = 1 mm, b = 2 mm, d = 500 μm.

To summarise spectral information we used psychophysical models of avian colour vision (Vorobyev et al., 1998) following the methods described by Delhey et al. (2015). Visual models require knowledge on the visual sensitivity functions of the four types of cone used by birds in colour vision, the relative abundance of each of these cones in the retina and the spectrum of illuminating light. Colour vision in birds is mediated by four types of single cones sensitive to very short (VS), short (S), medium (M) and long (L) wavelengths of light (Vorobyev et al., 1998). Variation in visual sensitivity between species is mainly restricted to the VS and S cones and birds can be generally classified in two groups: ultraviolet-sensitive (U-type) and violet-sensitive (V-type) species; and U-type species have VS cones with peak sensitivity shifted towards shorter wavelengths (Hart & Hunt, 2007). In our sample, both types of visual sensitivity functions are represented (Ödeen et al., 2012); therefore, we model both to see whether it affects the results (visual sensitivity functions obtained from Endler & Mielke, 2005). As the relative abundance of the four cones in the retina does not differ consistently between U- and V-type species (Hart, 2001), we used average cone proportions as obtained from Hart (2001) (0.38:0.69:1.14:1.00 for VS:S:M:L, respectively) and combined these with behavioural estimates of the Weber fraction (0.1; Vorobyev et al., 1998; Lind et al., 2014; Olsson et al., 2017) using formula (10) in Vorobyev et al. (1998) to obtain the noise-to-signal ratios for each cone type (ν_{VS}=0.162, ν_{S}=0.120, ν_{M}=0.094, ν_{L}=0.1). Finally, we used the spectrum of standard daylight (D65, open habitats) as illuminant (Vorobyev et al., 1998).

Visual models yield a set of quantum catches for the four types of single cones (i.e. how much each cone type is stimulated by a specific combination of reflectance spectrum and irradiance) that can be transformed into three coordinates x, y, z that define the position of each spectrum in the visual space of birds (Figs 3 and S4). This visual space takes the shape of a tetrahedron where each apex represents the sole stimulation of one cone type (Endler & Mielke, 2005). Using the formulae in Cassey et al. (2008), distances between points in visual space are measured in 'just noticeable differences' (jnd), whereby distances > 1 jnd are considered to be discriminable by birds. For each unique structurally coloured patch (i.e. excluding black and rufous patches), we computed the average position, or centroid, of the patch in the visual space (Figs 3 and S4), as well as the standard error (s.e.) around each chromatic coordinate (xyz). Similarly, for each spectrum we computed achromatic variability (i.e. 'brightness' or luminance variation, measured in jnd) as described by Delhey et al. (2015), using formula (7) in Siddiqui et al. (2004) and a noise-to-signal ratio of 0.2 (Olsson et al., 2017) and subsequently the average brightness and standard error for each unique patch.

As an additional analysis, from the reflectance spectra we also computed UV chroma – the summed reflectance of light in the range of 300-400 nm divided by the summed reflectance in the range 300-700 nm – as an index of colour purity (Andersson et al., 1998), using the package 'pavo' in R 3.4.0 (R Core Team, 2017). We did not compute hue and spectral saturation as they did not appear relevant here, given that many of the measured spectra display multiple peaks, with substantial variation in the height and location of the primary and secondary peaks (Figs 2, S2 and S3).

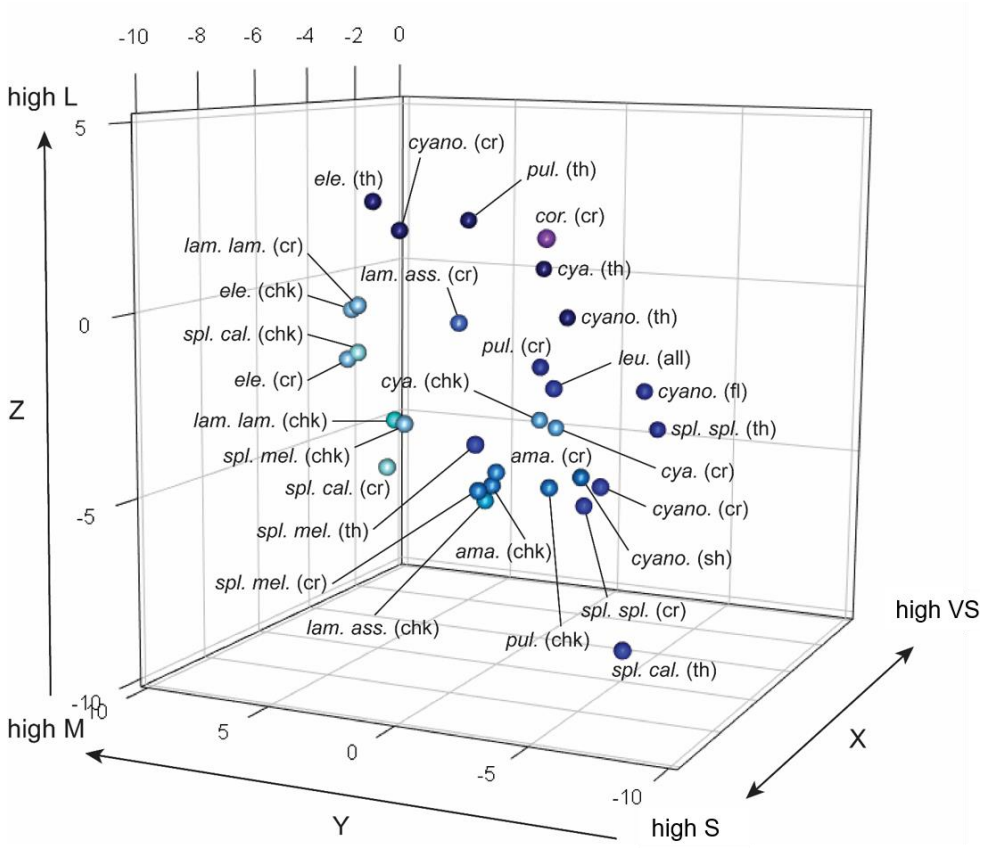


Figure 3. Average positions of the studied structurally coloured patches in the visual space of U-type eyes. The X-axis represents stimulation of the VS cone relative to the S cone, higher values of the Y-axis represents higher stimulation of the M cone relative to VS and S cones, while the Z-axis represents higher relative stimulation of the L cone compared with the other three (units = jnd). Taxon abbreviations: *ama. = M. amabilis*, *cor. = M. coronatus (coronatus)*, *cya. = M. cyaneus (cyaneus)*, *cyano. = M. cyanocephalus (bonapartii)*, *ele. = M. elegans*, *lam. ass. = M. lamberti assimilis*, *lam. lam. = M. lamberti lamberti*, *leu. = M. leucopterus (leuconotus)*, *pul. = M. pulcherrimus*, *spl. cal. = M. splendens callainus*, *spl. mel. = M. splendens melanotus*, *spl. spl. = M. splendens splendens*. Plumage patch abbreviations: cr = crown, ch = cheek, fl = flank, th = throat, ru = rump, sh = shoulder; all* = the whole body except the wings and tail for *M. leucopterus*. Sphere colours correspond approximately to the feather colour perceived by the human eye.

Transmission electron microscopy

To characterise the feather micro- and nanostructural elements responsible for the colours of the different studied patches, we used transmission electron microscopy (TEM). We prepared samples following a protocol similar to Shawkey et al. (2003). Briefly, we cut feather barbs from the upper 5 mm of contour feathers, washed them in 100% Ethanol (twice for 20 min each time) and infiltrated them with Epon in successive concentrations of 15, 50, 70 and 100% (24 h each time). Barbs were then placed into moulds and, after curing the blocks at 60ºC for 16 h in an oven, we trimmed them with a Leica S6 EM Trim2 (Leica Microsystems GmbH, Wetzlar, Germany) and cut 100 nm thin sections using a Leica EM UC6 ultramicrotome (Leica Microsystems) equipped with a MT14878 DIATOME diamond knife (DIATOME, Hatfield, PA, USA). Sections were transferred with a loop to copper grids, stained with uranyl acetate and lead citrate and viewed on a JEOL JEM-1010 transmission electronic microscope (JEOL Ltd., Tokyo, Japan). Micrographs were taken at magnifications between ×800 and ×30,000 (Fig. 4). Our study focuses on the role of the barb structure; therefore we did not investigate the role of barbules, which were only observed in very dark blue and indigo feathers (n = 6 of 30 structurally coloured feathers), as well as black and rufous feathers (Fig. 2). Nevertheless, using microspectrometry we measured the reflectance of single barbs, and the spectra were very similar to those of the corresponding plumage patches (Fig. S5), suggesting a negligible influence of the barbules on the overall plumage colour in those cases.

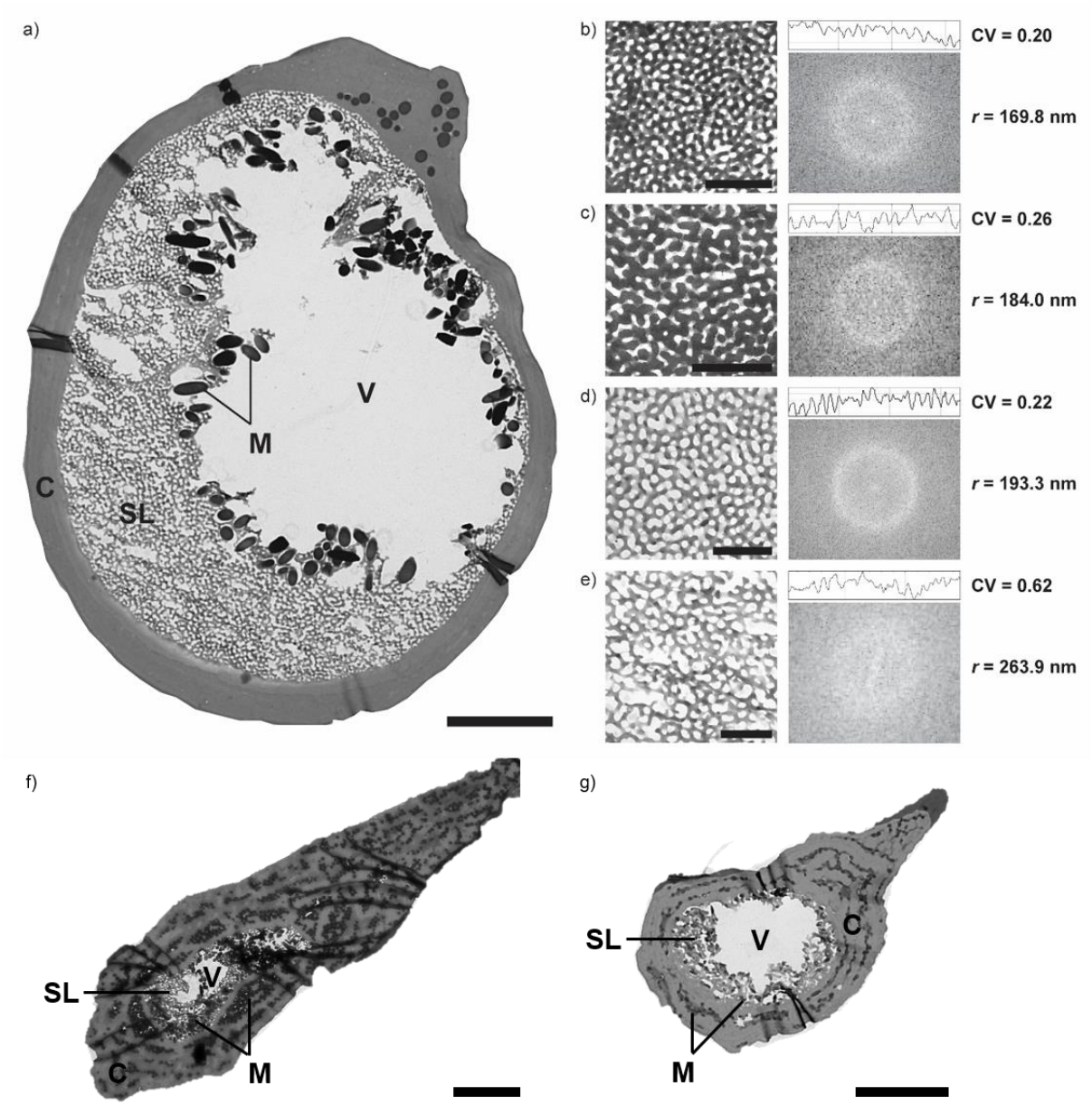


Figure 4. a) Feather barb microstructure from a structurally coloured plumage patch. Shown is a TEM image of a cross-section of a feather barb collected from the blue crown of a male *M. elegans*. Scale bar = 5 μm. b-e) Spongy layer characteristics. Shown are TEM images of the spongy layer found in the feather barbs collected from b) the throat of a male *M. cyaneus cyaneus*, c) the crown of a male *M. coronatus coronatus*, d) the throat of a male *M. splendens melanotus*, and e) the crown of a male *M. elegans*. Scale bars = 1 μm. Also shown are the corresponding 2D profile plots representing spatial fluctuations in dark (keratin) and light (air) areas in the spongy layer and the coefficient of variation (CV) in the spacing between wave

peaks (as an index of nanostructural regularity), as well as the 2D Fourier power spectra and associated spatial frequencies (*r*) given by the internal radius of the rings. Also shown for comparison are TEM images of a cross-section of a feather barb collected from f) the black throat of a male *M. lamberti lamberti* and g) the rufous shoulder of a male *M. amabilis*. Scale bars = 4 μm. C = cortex, SL = spongy layer, M = melanin granule and V = vacuole.

Microstructural variation

We used ImageJ 1.51 to measure the thickness of the keratin cortex and spongy layer (for structurally coloured patches only as black and rufous barbs have little, degraded or no spongy layer; Figs 4f and 4g) at six different evenly spaced points around the barb. Since black and rufous feathers have a pointed shape (and not an oval or round shape), we avoided the pointed area to measure the cortex thickness. Using both ImageJ 1.51 and Matlab, we also measured the cross-sectional area of the spongy layer and of the melanin granules underneath to estimate the ratio of melanin to spongy layer (for structurally coloured patches only). These parameters may affect the amount and/or range of wavelengths of light absorbed by the barb (Shawkey & Hill, 2006; Shawkey et al., 2006; D'Alba et al., 2012) and therefore the reflected colour.

Fourier analysis

For all structurally coloured patches, we performed a two-dimensional Fourier analysis on TEM images of feather barbs using the 2D fast Fourier transform (FFT) tool of ImageJ 1.51. This analysis allows us to determine whether nanostructures are sufficiently organised at an appropriate scale to produce colour by coherent light scattering alone (Prum et al., 1999; Shawkey et al., 2003). For all images, the largest available square portion of keratin and air uninterrupted by melanin granules, cell boundaries or keratin cortex (280-1940 pixels^2) was selected (Fig. 4). We then created 2D Fourier power spectra for each selected portion. All

structurally coloured feather barbs showed discrete rings in the Fourier power spectra (Figs 4b-4e). For each ring, we took five measurements of the internal radius (in nm/cycle) to estimate the characteristic spatial frequency of the spongy layer (with small values of spatial frequency being associated with small scattering elements).

Nanostructural variation

Both the size and regularity of the scattering elements in the spongy layer may affect aspects of the reflected colour (Shawkey et al., 2003, 2005; D'Alba et al., 2012). Therefore, for all structurally coloured patches, we measured the diameter of fifteen keratin rods and fifteen circular air spaces on each TEM image (Figs 4b-4e). The diameters of keratin rods and air spaces were not correlated ($|r| = 0.02$); therefore we summed the average diameters of keratin rods and air spaces to obtain the distance between scatterers (Prum et al., 2003; Shawkey et al., 2005). In addition, we used ImageJ 1.51 to produce profile plots of barb TEM images. We used the same selected square portion of the TEM image as above in the Fourier analysis. The profile plot shows a two-dimensional waveform, representing the spatial density fluctuation in the intensity of pixels (i.e. proxy for spatial variation in refractive index of keratin and air) within the selected area (Figs 4b-4e). We took fifteen measurements of the distance between the peaks in the waveform and then calculated the coefficient of variation (CV) of the inter-peak distance as a measure of nanostructural regularity.

Statistical analyses

All analyses were done in R 3.4.0 (R Core Team, 2017). Since black and rufous feathers have little, degraded or no spongy layer, the analyses described below were performed on structurally coloured patches only. We used a meta-analytical approach to the study of variation (Nakagawa et al., 2015), which accounts for sampling error, multiple measurements per species

and phylogenetic relatedness, implemented using the package "MCMCglmm" (Hadfield, 2010; Hadfield & Nakagawa, 2010). We fitted each colorimetric variable – x, y, z (computed using either U- or V-type visual sensitivity functions) and brightness – as the response variable in separate models. In all models, we fitted the following explanatory variables: the thickness of the keratin cortex, the thickness of the spongy layer, the ratio of melanin to spongy layer, the spatial frequency of the spongy layer provided by the Fourier analysis, the distance between scatterers – calculated as the sum of the diameters of keratin rods and air spaces – and the nanostructural regularity. All continuous explanatory variables were centred and scaled. Random effects included phylogenetic relatedness (as the inverse of the phylogenetic covariance matrix; Hadfield & Nakagawa, 2010; Fig. S1) and species identity, as in some species more than one plumage patch was measured. Sampling error variances (i.e. squared s.e. for each unique patch) for each chromatic coordinate or brightness estimate were also included in the models to account for this uncertainty. We used a relatively uninformative prior for the residuals and Cauchy priors for random effects and normal distributions centred on zero with large variances as fixed effects priors. We ran 101,000 iterations per model, from which we discarded the initial 1,000 (burn-in period). Each chain was sampled at an interval of 100 iterations, so that the effective sample size was 1,000. Model convergence was assessed using trace graphs and autocorrelation plots. For each model, we computed marginal (fixed effects) and conditional (fixed + random effects) R^2 values (Nakagawa & Schielzeth, 2013). We then used the same Bayesian phylogenetic mixed model approach fitting UV chroma as the response variable instead.

Moreover, we assessed the degree of phylogenetic effects on each colorimetric and morphological variable using again a Bayesian phylogenetic mixed model approach. We fitted each colorimetric or morphological variable as the response variable in separate models. In all models, we fitted a simple intercept as the explanatory variable, and the same random effects

as above. We then determined the phylogenetic signal λ (i.e. the proportion of variation accounted by phylogenetic relatedness relative to the total variation accounted by the random effects).

It should be noted that, although we investigated 30 different structural plumage colours from nine different species, our sample size of $n = 1$ feather barb for each colour warrants caution in the interpretation of our results. In particular, there may be some measurement error when quantifying the dimensions of the different structural elements, and potential variation in these dimensions between barbs was not analysed. Nonetheless, the robustness of some effects we observed, as well as the substantial marginal R^2 values of our statistical models (see results), provide strong support for our main conclusions.

Results

Male nuptial plumage in fairy-wrens

We studied the colour and microstructure of feather barbs collected from the nuptial plumages of male fairy-wrens. Structurally coloured patches included various blue (light blue, azure, cobalt blue, navy blue, etc.) and indigo plumages, as well as one case of purple plumage (in male *M. coronatus*; Figs 1 and 2). Only the distal barbs of the feathers displayed the structural colour and had generally few or no barbules (Fig. 2), whereas the proximal barbs displayed either a black (for blue and indigo feathers) or rufous (for purple feathers) colouration. In contrast, black and rufous feathers had both barbs and barbules (Fig. 2), and rufous feathers also displayed a black colouration in their proximal part.

Reflectance spectrometry and avian visual model

Reflectance spectra of all blue and indigo patches were characterised by the presence of one or two discrete peaks in the UV and/or blue range (Figs 2, S2 and S3). The main peak was located

between 360 and 540 nm, i.e. in most cases in the blue range, but in a few cases in the UV or green range (Figs 2, S2 and S3). In contrast, the second peak, when present, was always found between 325 and 375 nm, i.e. in the UV range, and in many cases it appeared as a 'shoulder' (i.e. not a distinct peak) or almost fully merged with the main peak (Figs 2, S2 and S3). In particular, spectra of very dark blue and indigo patches (i.e. the throat of *M. cyaneus*, *M. elegans* and *M. pulcherrimus*, and the rump of *M. cyanocephalus*) peaked in the UV range at a low reflectance value (< 15%; Figs 2, S2 and S3). Similarly, the reflectance spectrum of the purple crown of *M. coronatus* was characterised by two discrete peaks, the smaller one being in the UV range (at 340 nm) and the higher one in the blue range (at 460 nm); however, it also exhibited a gradual and steep increase between 550 and 700 nm, i.e. towards the red range, unlike other spectra (Fig. 2).

Plotting spectra in the bird visual space showed a relatively large colour span in the visual space of both U- and V-type species (Figs 3 and S4). In this representation the main axes of chromatic variation were z (range U-type eyes: 11.5 jnd, V-type eyes: 11.1 jnd) and y (range U-type eyes: 11.3 jnd, V-type eyes: 10.4 jnd), while variation along x was less marked (U-type eyes: 4.5 jnd, V-type eyes: 5.1 jnd). Achromatic variation was also high among the studied colours, with a range of 17.4 jnd.

Reflectance spectra of black and rufous patches displayed no discrete peaks: spectra measured on black patches were flat across the entire avian visual range (very low reflectance), while rufous spectra exhibited low reflectance between 300 and 525 nm, and a gradual and steep increase between 525 and 700 nm, i.e. towards the red range (Fig. 2).

Feather structure

All thirty structurally coloured feather barbs had an oval or round shape and were characterised by the presence of a medullary spongy layer located beneath a keratin cortex and on top of a

layer of melanin granules surrounding large hollow central vacuoles (Fig. 4a). The spongy layer was typically composed of a matrix of irregularly shaped keratin channels and air spaces (Fig. 4b-4e), forming a 'quasi-ordered array' (Prum et al., 2003). The dimensions of these various components showed substantial variation among the barbs examined (Table 1, Fig. 4). In particular, all feather barbs showed discrete rings in the Fourier power spectra (Fig. 4), indicating high levels of nanostructural uniformity and organisation (Prum et al., 1998, 1999), allowing the production of colour by coherent light scattering alone.

In contrast, black and rufous feather barbs exhibited a pointed shape and were characterised by the presence of numerous melanin granules within a thicker keratin cortex (mean ± SE = 3.04 ± 0.22 μm vs. 2.53 ± 0.16 μm for structurally coloured barbs) surrounding little or no spongy layer (n = 3 black samples out of 8 had no spongy layer, and all rufous samples had extremely degraded spongy layers) and hollow central vacuoles (Figs 4f and 4g). In particular, rufous barbs displayed a large amount of relatively small and round melanin granules, suggesting that they mainly contained phaeomelanin (Liu et al., 2003; pers. obs.). Additionally, when a spongy layer was present, it was notably degraded by holes (absence of keratin) and often interrupted by melanin granules (Figs 4f and 4g). As a result, the next section of the results relates to structurally coloured feathers only.

Variation in feather structure and colour

We report results for colour variability as computed using U-type visual sensitivity functions; results using V-type functions were generally similar and are provided in the supplementary material (Tables S8-S11). Analyses were run with all six morphological variables in the same model and for each variable separately.

Table 1. Dimensions of micro- and nanostructural components of feather barbs. For each studied patch are shown the means (± s.e.m.) of the cortex thickness, spongy layer (SL) thickness, SL spatial frequency (from Fourier analysis), diameter of keratin channels, and diameter of circular air spaces, as well as the values of ratio of melanin to SL (areas) and of SL irregularity.

Taxon	Plumage patch	Cortex thickness (nm)	SL thickness (nm)	Ratio melanin/SL	SL spatial frequency (nm)	ø keratin channels (nm)	ø air spaces (nm)	SL irregularity
Malurus cyanocephalus bonapartii	crown	4153.9 ± 480.8	11305.9 ± 1964.0	0.09	242.2 ± 6.9	48.8 ± 2.1	128.3 ± 6.2	0.26
Malurus cyanocephalus bonapartii	throat	2216.2 ± 450.3	2748.3 ± 350.2	0.10	184.4 ± 3.4	52.6 ± 2.5	110.5 ± 6.0	0.34
Malurus cyanocephalus bonapartii	flank	2216.1 ± 750.2	3278.1 ± 684.8	0.14	194.5 ± 5.4	43.9 ± 2.2	105.7 ± 4.6	0.25
Malurus cyanocephalus bonapartii	rump	4341.5 ± 1035.1	4152 ± 508.0	0.25	187.6 ± 4.9	60.1 ± 3.1	115.8 ± 6.4	0.30
Malurus cyanocephalus bonapartii	shoulder	2601.8 ± 713.7	7079.9 ± 549.2	0.14	218.4 ± 4.7	57.3 ± 2.8	112.9 ± 4.4	0.28
Malurus coronatus coronatus	crown	2873.2 ± 182.0	1603.8 ± 220.1	0.52	184.0 ± 5.3	86.9 ± 3.2	86.2 ± 2.6	0.26
Malurus elegans	crown	2032.5 ± 522.3	4786.4 ± 642.6	0.15	263.9 ± 4.4	65.4 ± 4.2	170.5 ± 10.2	0.62
Malurus elegans	cheek	2207.9 ± 258.1	5487.8 ± 655.1	0.11	230.2 ± 15.9	80.9 ± 3.9	121.7 ± 5.5	0.35
Malurus elegans	throat	3094.3 ± 542.0	3021.8 ± 304.2	0.59	176.5 ± 3.1	59.2 ± 3.6	86.2 ± 6.1	0.25
Malurus pulcherrimus	crown	1903 ± 185.6	5477.8 ± 514.2	0.11	167.0 ± 5.9	47.5 ± 1.4	96.9 ± 5.2	0.18
Malurus pulcherrimus	cheek	4048.6 ± 570.8	8834.2 ± 928.4	0.03	182.3 ± 6.7	62.3 ± 2.7	132.6 ± 7.9	0.27
Malurus pulcherrimus	throat	3221.9 ± 467.8	2748.3 ± 350.2	0.39	189.2 ± 0.9	60.0 ± 2.2	111.3 ± 4.8	0.20
Malurus amabilis	crown	2668.5 ± 318.4	6086.6 ± 520.6	0.07	224.3 ± 3.0	68.2 ± 2.7	111.6 ± 5.2	0.42
Malurus amabilis	cheek	2779.5 ± 704.2	12269.7 ± 1550.1	0.02	220.2 ± 5.3	61.5 ± 2.6	121.0 ± 5.8	0.32
Malurus lamberti lamberti	crown	2632.6 ± 452.2	7908.1 ± 671.3	0.13	189.8 ± 6.8	66.0 ± 4.1	136.9 ± 6.3	0.35
Malurus lamberti lamberti	cheek	1408.1 ± 74.4	6834.7 ± 446.9	0.12	208.4 ± 2.2	59.4 ± 2.5	142.5 ± 6.8	0.42
Malurus lamberti assimilis	crown	1605.4 ± 241.0	3250.2 ± 446.1	0.15	186.1 ± 2.7	63.6 ± 4.1	123.6 ± 7.2	0.31
Malurus lamberti assimilis	cheek	3562.5 ± 905.4	5989.0 ± 526.6	0.03	239.4 ± 4.8	74.3 ± 4.6	168.7 ± 10.3	0.39

Malurus cyaneus cyaneus	crown	1877.6 ± 116.1	8854.5 ± 1025.9	0.08	169.3 ± 6.4	71.1 ± 2.5	109.1 ± 4.0	0.25
Malurus cyaneus cyaneus	cheek	1845.3 ± 172.6	8301.8 ± 1578.0	0.03	216.4 ± 2.7	74.9 ± 2.5	107.6 ± 2.4	0.18
Malurus cyaneus cyaneus	throat	3393.1 ± 720.7	2611.8 ± 192.6	0.03	169.8 ± 2.5	58.4 ± 2.1	71.9 ± 2.4	0.20
Malurus splendens splendens	crown	3654.6 ± 559.7	13248.4 ± 653.3	0.03	161.2 ± 5.3	57.7 ± 3.0	107.6 ± 2.6	0.20
Malurus splendens splendens	throat	1439.4 ± 140.0	4130.1 ± 486.4	0.42	189.1 ± 2.6	61.2 ± 2.6	121.9 ± 4.6	0.21
Malurus splendens melanotus	crown	3031.1 ± 423.1	7277 ± 1405.2	0.12	214.9 ± 4.9	70.7 ± 2.0	120.2 ± 5.3	0.20
Malurus splendens melanotus	cheek	1901.7 ± 159.4	11230.0 ± 1885.3	0.01	208.5 ± 2.8	77.1 ± 3.2	128.8 ± 6.0	0.24
Malurus splendens melanotus	throat	1600.5 ± 66.7	6172.9 ± 971.6	0.09	193.3 ± 1.9	64.7 ± 2.3	142.1 ± 4.7	0.22
Malurus splendens callainus	crown	3052.9 ± 297.9	25284.3 ± 2985.9	0.05	213.4 ± 2.3	56.4 ± 2.2	129.4 ± 4.8	0.33
Malurus splendens callainus	cheek	868.7 ± 84.8	3149.8 ± 248.3	0.31	189.1 ± 8.4	66.7 ± 2.6	114.1 ± 4.8	0.38
Malurus splendens callainus	throat	2399.6 ± 323.7	17258.0 ± 2396.8	0.06	190.0 ± 1.9	56.6 ± 3.1	116.1 ± 5.1	0.26
Malurus leucopterus leuconotus	all*	1280.2 ± 211.6	2517.4 ± 259.7	0.27	226.2 ± 1.6	63.0 ± 2.1	151.4 ± 6.2	0.28

Variation along the z axis (i.e. stimulation of the L cone relative to VS+S+M cones) was predicted by both the thickness of the spongy layer and the ratio of melanin to spongy layer. Feathers relatively poor in short- and middle-wavelength reflectance (UV through green) had barbs with a thinner spongy layer sitting above a relatively larger amount of melanin granules (marginal R^2 (U-/V-type) = 0.64/0.64; Tables S3 and S8, Figs 5 and 6).

Variation along the y axis (i.e. stimulation of the M cone relative to VS and S cones) was mainly predicted by nanostructural regularity. More irregular spongy layers were associated with a higher stimulation of the M (green-sensitive) cone relative to the VS and S cones (UV and blue) (marginal R^2 (U-/V-type) = 0.43/0.37; the effect was non-significant when using V-type functions; Tables S3 and S8, Figs 5 and 6). No morphological variable was correlated with variation across the x axis (i.e. stimulation of the VS cone relative to the S cone) based on the full model including all six morphological variables (Tables S3 and S8, Fig. 6).

Achromatic variation (brightness) was associated with the thickness of both the cortex and the spongy layer. Barbs with a thinner cortex and a thicker spongy layer were brighter (although the effect of spongy layer thickness was marginally non-significant; marginal R^2 = 0.65; Table S3, Figs 6 and S6). UV chroma was negatively associated with both the thickness of the spongy layer and the distance between scatterers (although both effects were marginally non-significant; marginal R^2 = 0.30; Table S3, Fig. S6).

We also tested the separate effects of each morphological variable of interest. These analyses indicated that in addition to the effects reported above, the spatial frequency of the spongy layer – as well as the distance between scatterers in some cases – might affect the values of the x- and z-coordinates, as well as brightness: spongy layers with larger spatial frequencies were associated with lower stimulation of the VS cone relative to the S cone and of the L cone relative to the other three, and also with brighter feathers (Tables S4-S7 and S9-S11). These effects were no longer significant when running the full models, which may be in part explained

by some degree of correlation ($r = 0.49$) between the spatial frequency estimated by the Fourier analysis and the distance between scatterers.

Finally, we assessed the degree of phylogenetic constraint on each colorimetric and morphological variable: phylogenetic effects appeared relatively weak for all variables ($\lambda \leq 0.34$; Table S12).

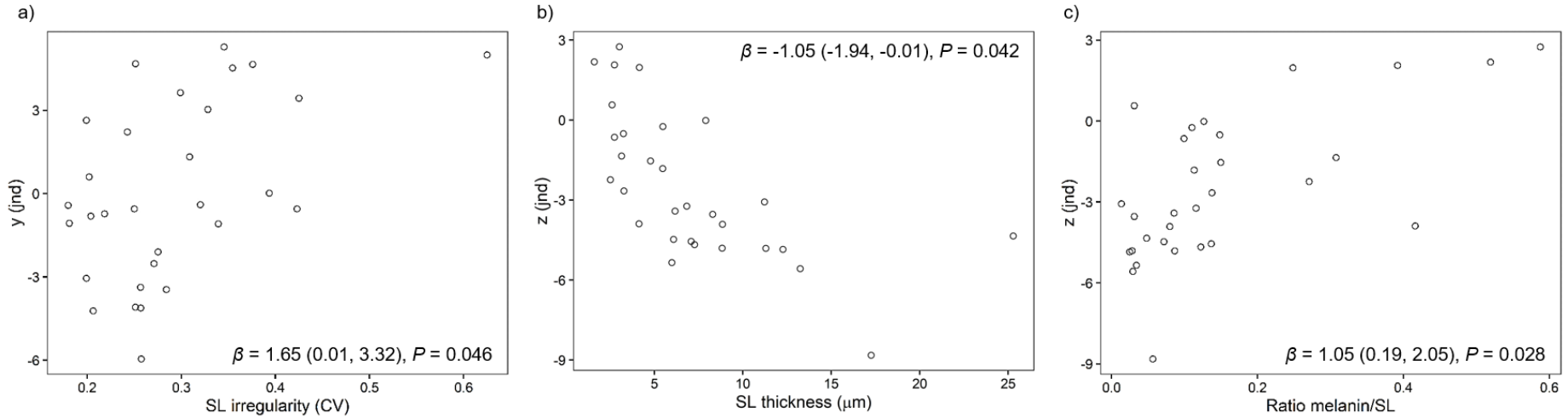


Figure 5. Chromatic variation among structural colours in male fairy-wrens is related to the regularity and thickness of the spongy layer, as well as the ratio of melanin to spongy layer. Shown are a) the y-coordinate as a function of the spongy layer irregularity (CV = index of irregularity, ranging from 0 to 1), and the z-coordinate as a function of b) the thickness of the spongy layer (SL; in μm), and c) the ratio of melanin to spongy layer. Both y and z (in jnd) were computed using the U-type visual sensitivity. For each plot are indicated the posterior mean (β) and 95% credible interval for the depicted correlation and the corresponding *P*-value.

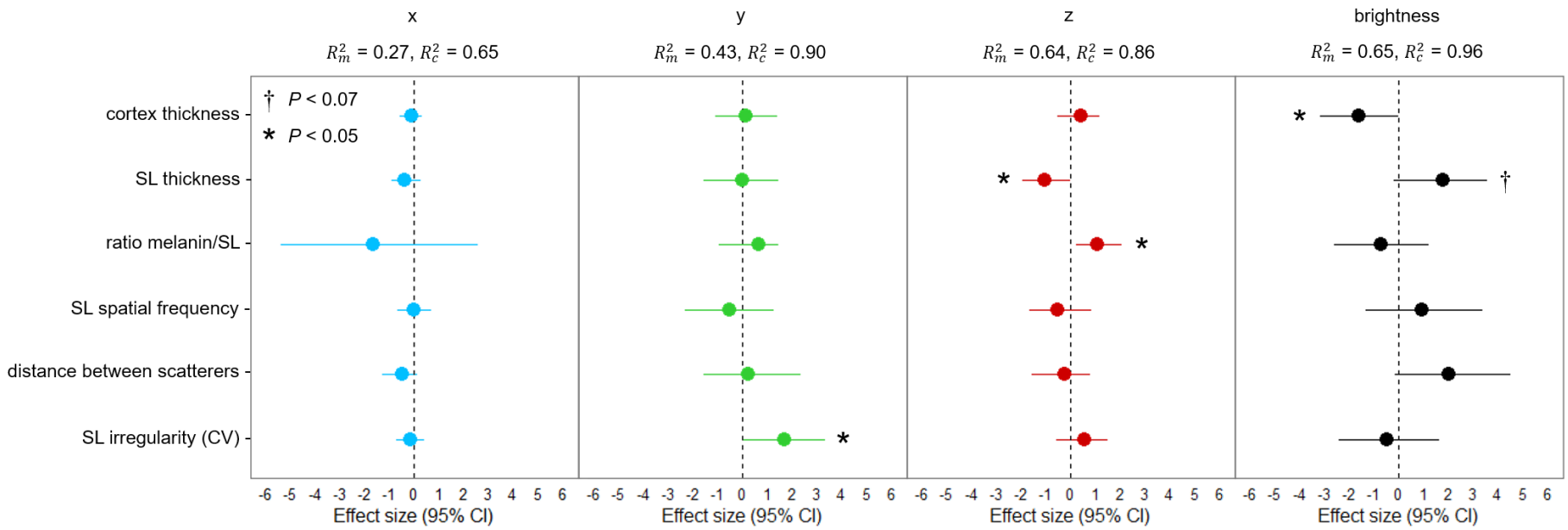


Figure 6. Chromatic and achromatic variation among structural colours in male fairy-wrens is related to different structural elements of feather barb. Shown are the effect size estimates (posterior means and 95% credible intervals (CI); based on the full models presented in Table S3) of correlations between cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity (CV), and chromatic coordinates (xyz) – computed using U-type visual sensitivities – and achromatic brightness (n = 30). The x-coordinate represents stimulation of the VS cone relative to the S cone, higher values of the y-coordinate represents higher stimulation of the M cone relative to VS and S cones, while the z-coordinate represents higher relative stimulation of the L cone compared with the other three. Marginal and conditional R^2 values (R^2_m and R^2_c respectively) are also presented.

Discussion

In this study, we show that differences in 30 patches of male nuptial structural plumage colours across nine species of fairy-wrens are caused by differences in the amount, dimensions and arrangement of multiple structural elements of feather barb. More specifically, we found that (1) chromatic variation is related to the regularity and thickness of the spongy layer, the spacing of its scattering elements, as well as the relative amounts of spongy layer and melanin; and that (2) achromatic brightness varies with the thickness of the keratin cortex and spongy layer. Our results therefore indicate that multiple characteristics of feather microstructure can independently contribute to chromatic and achromatic variation in non-iridescent structural colours as perceived by birds.

Male fairy-wrens display structural colours that vary widely, including ultraviolet (UV), indigo, various blue and purple hues (Figs 1-3 and S2-S4). Transmission electron microscopy revealed that the barb microstructure of all sampled feathers is organised in a very similar way, with an outer keratin cortex above a medullary spongy layer – composed of a matrix of irregularly shaped keratin channels and air spaces – and a basal layer of melanin granules surrounding large hollow central vacuoles (Fig. 4). This similarity suggests that the observed colour variation among the sampled feathers is not caused by the presence or absence of certain structural elements, but primarily by differences in the characteristics (size, density, etc.) of one or more of these elements. The discrete rings observed in the Fourier power spectra of all spongy layers showed that the colours are primarily produced by coherent light scattering due to a nanostructural arrangement of keratin and air – uniform in all directions and highly ordered – in the feather barbs (Prum, 2006). This confirms the critical role played by the medullary spongy layer in the production of structural colours (Prum et al., 1998, 2003, Shawkey et al., 2003, 2005, 2006; Shawkey & Hill, 2006; D'Alba et al., 2012).

Variation in relative reflectance at short and medium wavelengths is generated by changes in the characteristics of the spongy layer. We found some evidence that lower spatial frequencies and smaller distances between scattering elements could account for reflectance spectra richer in shorter wavelengths (Tables S4 and S9). Hence, tighter arrays of keratin and air are responsible for colours richer in UV. Additionally, more regular spongy layers create colours that are richer in shorter wavelengths relative to medium wavelengths (Figs 4-6). This is consistent with previously reported positive associations between nanostructural regularity (variation in size of the keratin channels or air spaces) and spectral saturation (Shawkey et al., 2003, 2005; D'Alba et al., 2012). Indeed, spectra with relatively higher shortwave reflectance and lower mediumwave reflectance are also those exhibiting a narrower curve (Figs 2, S2 and S3) and therefore higher spectral saturation. As suggested by Shawkey et al. (2003), increased uniformity in size of the scattering elements (i.e. increased regularity) is likely to result in a tighter grouping of the reflected light around the wavelength of peak reflectance (i.e. greater purity of the reflected colour), whereas increasing the variation in those elements may spread the reflected light over a broader spectrum, which in our case occurs mostly across the short- and mediumwave ranges.

Variation in relative reflectance at long wavelengths is the result of changes in both the thickness of the spongy layer and the relative amount of melanin granules compared to the amount of spongy layer. Barbs comprising a thinner spongy layer on a relatively larger amount of melanin granules generate colours richer in long-wavelength reflectance (corresponding to darker blue and purple feathers in our sample; Figs 5 and 6). Since the spongy layer is the main colour-producing element, responsible for the UV and blue hues, a decrease in its thickness will be associated with lower reflectance across the shortwave range. Because melanin absorbs light across the entire spectrum, a relatively larger amount of melanin granules will induce a

larger decrease in the reflectance across the shortwave range (where reflectance is the highest) relative to the longwave range (where reflectance is already low).

The importance of melanin to structural colour production is also demonstrated by comparing structural colours with black and rufous colours. Both rufous and purple reflectance spectra exhibit a very similar gradual increase in the red range (Fig. 2). This strongly suggests that purple barbs (of the purple-crowned fairy-wren, *Malurus coronatus*) contain a basal layer consisting mainly of phaeomelanin, as previously proposed by Peters et al. (2013), and that phaeo- instead of eumelanin is primarily responsible for purple instead of blue plumage. Comparison of blue and black plumages confirms the previous findings of Doucet et al. (2004) and Driskell et al. (2010): most sampled black feather barbs (including those from *M. splendens*, *M. lamberti* and *M. coronatus*) contain a spongy layer that could produce a structural blue colour, but the degraded state of the spongy layer, in addition to a higher density of melanin, as well as a thicker and more heavily melanised cortex, make them appear black. Therefore, both the amount and type of melanin are essential aspects of structurally coloured plumages. Yet, as most bird feather barbs contain a mixture of the two types of melanin (eu- vs. phaeomelanin; McGraw, 2006), further research is needed to clarify how variations in the relative proportions of the two contribute to structural colour variation.

Feather brightness (achromatic, dark-to-light variation) is predicted by the thickness of both the keratin cortex and the spongy layer, with thinner cortices and thicker spongy layers producing brighter feathers (Figs 6 and S6). Previous studies of non-iridescent structural plumage colour suggested that the cortex primarily acts to absorb light (Finger, 1995; Shawkey et al., 2005); hence, a thinner cortex will increase the amount of light reflected (total brightness). In addition, both coherent (Benedek, 1971) and incoherent (Kerker et al., 1966; Kerker, 1969; Finger, 1995) scattering models of colour production predict positive associations between the number of scattering elements and brightness. The positive correlation we found between

spongy layer thickness and brightness is in agreement with these theoretical predictions as, for a given spatial frequency, a thicker spongy layer contains a larger number of scattering elements and as a result increases the amount of reflected light. Previous research also proposed that melanin density could affect brightness (Shawkey & Hill, 2006; but see Shawkey et al., 2005), a hypothesis that was partly supported by our results: the ratio of melanin to spongy layer was negatively associated with brightness when tested as a separate effect (but no effect when including all other variables; Table S7, Fig. 6).

In conclusion, feather barb microstructure can vary in several independent ways to generate variation in plumage colouration: either (1) in the thickness of the keratin cortex, or (2) in various characteristics of the spongy layer, including its thickness and regularity, or lastly, (3) in the relative amount of melanin sitting underneath the spongy layer (Fig. 7). Therefore, it is variation in multiple microstructural elements, rather than consistent changes in a single key mechanism, that leads to the diversity of structural colours in the nuptial plumages of male fairy-wrens. The coexistence of several labile parameters is likely to facilitate a greater variability in plumage colouration and may account for the rapid evolution of male ornamentation in this genus. While many studies of structural colours have focused on the role of the spongy layer (Prum et al., 1998, 1999, 2003; Shawkey et al., 2003), our study highlights the complexity of barb morphology that provides potential for a large variety of microstructures, resulting in highly variable structural colouration at the interspecific level. We found relatively small phylogenetic effects on all colorimetric and morphological variables, indicating that feather colour and barb morphology are probably not phylogenetically constrained in this clade of birds. Further studies are needed to improve our understanding of the optical mechanisms taking place between the various structural elements (e.g., how the interaction of the spongy layer and a basal layer of melanin with variable proportions of eu- and phaeomelanin affects the incident light), as well as the genetic and environmental factors influencing these elements.

This will provide further insights into the complex nature of the proximate mechanisms underlying ornamental colours and their role in shaping their evolution.

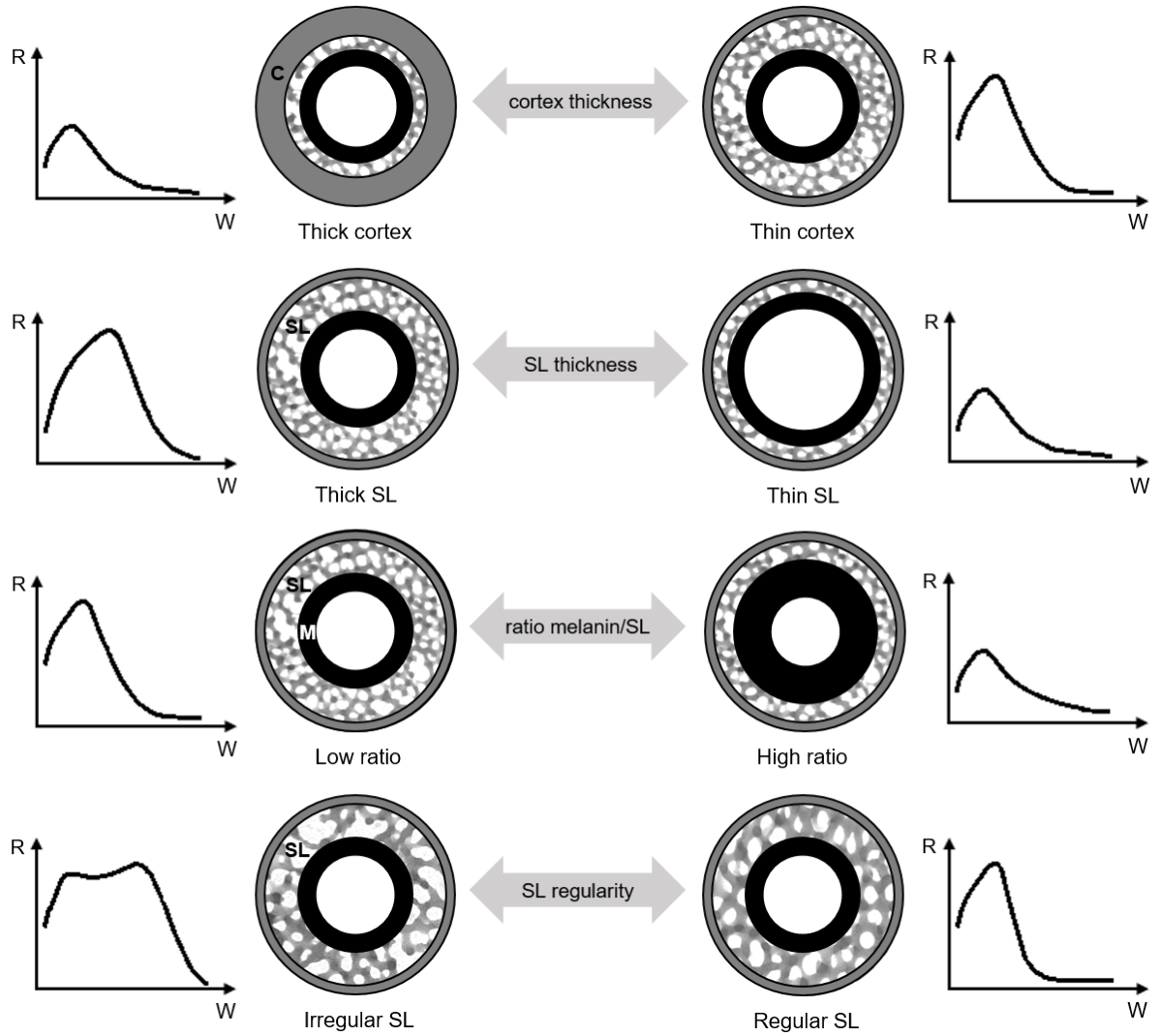


Figure 7. Schematic summarising results of the relationship between feather barb morphology and non-iridescent structural plumage colouration. Shown is a schematic representation of the effects of multiple structural elements of feather barb on plumage reflectance. Barb morphology may affect plumage colour through changes in cortex (C) thickness, spongy layer (SL) thickness, ratio of melanin (M) to SL, and SL regularity. For each depicted effect, the associated variation in the reflectance spectrum (R = reflectance, W = wavelength) is schematised as well.

Acknowledgements

This research was approved by the Australian Bird and Bat Banding Scheme (license 2230); Western Australia Department of Parks and Wildlife; Australian Wildlife Conservancy; and the ethics committees of the School of Biological Sciences at Monash University. We thank the curators at Museum Victoria for access to specimen samples, A. McQueen, E. Enbody and J. A. Jones for feathers sampled from living birds and photographs, and the staff at Australian Wildlife Conservancy's Mornington Wildlife Sanctuary for their continued support and advice. We are grateful for support by the Australian Research Council (FT10100505 & DP150103595 to AP), the Society for the Study of Evolution (Rosemary Grant Research Award to MF), the Holsworth Wildlife Research Endowment – Equity Trustees Charitable Foundation (to MF) and the Australian Wildlife Conservancy and its supporters.

References

Andersson, M. (1994). *Sexual selection.* Princeton University Press.

Andersson, S., Örnborg, J., & Andersson, M. (1998). Ultraviolet sexual dimorphism and assortative mating in blue tits. *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 445-450.

Auber, L. (1957). The structures producing "non-iridescent" blue colour in bird feathers. In *Proceedings of the Zoological Society of London* (Vol. 129, No. 4, pp. 455-486). Blackwell Publishing Ltd.

Benedek, G. B. (1971). Theory of transparency of the eye. *Applied Optics*, **10**, 459-473.

Brouwer, L., van de Pol, M., Atema, E. L. S., & Cockburn, A. (2011). Strategic promiscuity helps avoid inbreeding at multiple levels in a cooperative breeder where both sexes are philopatric. *Molecular Ecology*, **20**, 4796-4807.

Cassey, P., Ewen, J., Blackburn, T., Hauber, M., Vorobyev, M., & Marshall, N. (2008). Eggshell colour does not predict measures of maternal investment in eggs of *Turdus* thrushes. *Naturwissenschaften*, **95**, 713-721.

Cockburn, A., Osmond, H. L., & Double, M. C. (2008). Swingin' in the rain: condition dependence and sexual selection in a capricious world. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 605-612.

Cuthill, I. C. (2006). Color perception. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 3-40). Harvard University Press.

D'Alba, L., Kieffer, L., & Shawkey, M. D. (2012). Relative contributions of pigments and biophotonic nanostructures to natural color production: a case study in budgerigar (*Melopsittacus undulatus*) feathers. *Journal of Experimental Biology*, **215**, 1272-1277.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex.* John Murray.

Delhey, K., Hall, M., Kingma, S. A., & Peters, A. (2013). Increased conspicuousness can explain the match between visual sensitivities and blue plumage colours in fairy-wrens. *Proceedings of the Royal Society of London B: Biological Sciences*, **284**, 20121771.

Delhey, K., Delhey, V., Kempenaers, B., & Peters, A. (2015). A practical framework to analyze variation in animal colors using visual models. *Behavioral Ecology*, **26**, 367-375.

Doucet, S. M., Shawkey, M. D., Rathburn, M. K., Mays, H. L., & Montgomerie, R. (2004). Concordant evolution of plumage colour, feather microstructure and a melanocortin receptor gene between mainland and island populations of a fairy–wren. *Proceedings of the Royal Society of London B: Biological Sciences*, **271**, 1663-1670.

Driskell, A. C., Prum, R. O., & Pruett-Jones, S. (2010). The evolution of black plumage from blue in Australian fairy-wrens (Maluridae): genetic and structural evidence. *Journal of Avian Biology*, **41**, 505-514.

Dyck, J. (1976). Structural colours. *Proceedings of the International Ornithological Congress*, **16**, 426-437.

Eliason, C. M., Maia, R., & Shawkey, M. D. (2015). Modular color evolution facilitated by a complex nanostructure in birds. *Evolution*, **69**, 357-367.

Endler, J. A., & Mielke, P. W. (2005). Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, **86**, 405-431.

Espmark, Y., Amundsen, T., & Rosenqvist, G. (Eds.). (2000). *Animal signals: signalling and signal design in animal communication*. Tapir Academic Press.

Fan, M., Teunissen, N., Hall, M. L., Hidalgo Aranzamendi, N., Kingma, S. A., Roast, M., ... & Peters, A. (2018). From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird. *Journal of Animal Ecology*, **87**, 1274-1285.

Finger, E. (1995). Visible and UV coloration in birds: Mie scattering as the basis of color in many bird feathers. *Naturwissenschaften*, **82**, 570-573.

Friedman, N. R., & Remeš, V. (2015). Rapid evolution of elaborate male coloration is driven by visual system in Australian fairy-wrens (Maluridae). *Journal of Evolutionary Biology*, **28**, 2125-2135.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**, 1-22.

Hadfield, J. D., & Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, **23**, 494-508.

Hart, N. S. (2001). Variations in cone photoreceptor abundance and the visual ecology of birds. *Journal of Comparative Physiology A*, **187**, 685-697.

Hart, N. S., & Hunt, D. M. (2007). Avian visual pigments: characteristics, spectral tuning, and evolution. *The American Naturalist*, **169**(S1), S7-S26.

Hill, G. E., & McGraw, K. J. (Eds.). (2006). *Bird coloration, vol. 1: mechanisms and measurements*. Harvard University Press.

Kerker, M., Farone, W. A., & Jacobsen, R. T. (1966). Color effects in the scattering of white light by micron and submicron spheres. *The Journal of the Optical Society of America*, **56**, 1248-1255.

Kerber, M. (1969). *The scattering of light and other electromagnetic radiation.* Academic Press.

Lind, O., Chavez, J., & Kelber, A. (2014). The contribution of single and double cones to spectral sensitivity in budgerigars during changing light conditions. *Journal of Comparative Physiology A*, **200**, 197-207.

Liu, Y., Kempf, V. R., Brian Nofsinger, J., Weinert, E. E., Rudnicki, M., Wakamatsu, K., ... & Simon, J. D. (2003). Comparison of the structural and physical properties of human hair eumelanin following enzymatic or acid/base extraction. *Pigment Cell Research*, **16**, 355-365.

Maan, M. E., & Seehausen, O. (2011). Ecology, sexual selection and speciation. *Ecology Letters*, **14**, 591-602.

Marki, P. Z., Jønsson, K. A., Irestedt, M., Nguyen, J. M., Rahbek, C., & Fjeldså, J. (2017). Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, **107**, 516-529.

McGraw, K. J. (2006) Mechanics of melanin coloration in birds. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1*. (pp. 243-294). Harvard University Press.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133-142.

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., & Senior, A. M. (2015). Meta-analysis of variation: ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution*, **6**, 143-152.

Ödeen, A., Pruett-Jones, S., Driskell, A. C., Armenta, J. K., & Håstad, O. (2012). Multiple shifts between violet and ultraviolet vision in a family of passerine birds with associated changes in plumage coloration. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 1269-1276.

Olsson, P., Lind, O., & Kelber, A. (2017). Chromatic and achromatic vision: parameter choice and limitations for reliable model predictions. *Behavioral Ecology*, **29**, 273-282.

Peters, A., Kingma, S. A., & Delhey, K. (2013). Seasonal male plumage as a multi-component sexual signal: Insights and opportunities. *Emu*, **113**, 232-247.

Price, T. (2008). *Speciation in birds*. Roberts and Co.

Prum, R. O., Torres, R., Williamson, S., & Dyck, J. (1998). Constructive interference of light by blue feather barbs. *Nature*, **396**, 28-29.

Prum, R. O., Torres, R., Williamson, S., & Dyck, J. (1999). Two-dimensional Fourier analysis of the spongy medullary keratin of structurally coloured feather barbs. *Proceedings of the Royal Society of London B: Biological Sciences*, **266**, 13-22.

Prum, R. O., Andersson, S., & Torres, R. H. (2003). Coherent scattering of ultraviolet light by avian feather barbs. *The Auk*, **120**, 163-170.

Prum, R. O. (2006). Anatomy, physics, and evolution of structural colours. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 295-353). Harvard University Press.

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.

Shawkey, M. D., Estes, A. M., Siefferman, L. M., & Hill, G. E. (2003). Nanostructure predicts intraspecific variation in ultraviolet–blue plumage colour. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 1455-1460.

Shawkey, M. D., Estes, A. M., Siefferman, L., & Hill, G. E. (2005). The anatomical basis of sexual dichromatism in non-iridescent ultraviolet-blue structural coloration of feathers. *Biological Journal of the Linnean Society*, **84**, 259-271.

Shawkey, M. D., & Hill, G. E. (2005). Carotenoids need structural colours to shine. *Biology Letters*, **1**, 121-124.

Shawkey, M. D., & Hill, G. E. (2006). Significance of a basal melanin layer to production of non-iridescent structural plumage color: evidence from an amelanotic Steller's jay (*Cyanocitta stelleri*). *Journal of Experimental Biology*, **209**, 1245-1250.

Shawkey, M. D., Balenger, S. L., Hill, G. E., Johnson, L. S., Keyser, A. J., & Siefferman, L. (2006). Mechanisms of evolutionary change in structural plumage coloration among bluebirds (*Sialia* spp.). *Journal of the Royal Society Interface*, **3**, 527-532.

Shawkey, M. D., Morehouse, N. I., & Vukusic, P. (2009). A protean palette: colour materials and mixing in birds and butterflies. *Journal of the Royal Society Interface*, **6**(S2), S221-S231.

Shawkey, M. D., & D'Alba, L. (2017). Interactions between colour-producing mechanisms and their effects on the integumentary colour palette. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160536.

Siddiqi, A., Cronin, T. W., Loew, E. R., Vorobyev, M., & Summers, K. (2004). Interspecific and intraspecific views of color signals in the strawberry poison frog *Dendrobates pumilio*. *Journal of Experimental Biology*, **207**, 2471-2485.

Vorobyev, M., Osorio, D., Bennett, A. T., Marshall, N. J., & Cuthill, I. C. (1998). Tetrachromacy, oil droplets and bird plumage colours. *Journal of Comparative Physiology A*, **183**, 621-633.

Wilkie, S. E., Robinson, P. R., Cronin, T. W., Poopalasundaram, S., Bowmaker, J. K., & Hunt, D. M. (2000). Spectral tuning of avian violet-and ultraviolet-sensitive visual pigments. *Biochemistry*, **39**, 7895-7901.

Supplementary Material

Table S1. Feathers sampled from (a) living birds and (b) specimens of male *Malurus* provided by Museum Victoria.

Taxon	Collection locality	Latitude	Longitude	Specimen No.	Sampling period
(a) Feathers from living birds					
Malurus cyanocephalus bonapartii	Obo hunting camp, Western Province, Papua New Guinea	7°36'S	141°15'E	NA	April 2017[†]
Malurus coronatus coronatus	Australian Wildlife Conservancy's Mornington Wildlife Sanctuary, Western Australia	17°31'S	126°6'E	NA	November 2015[‡]
Malurus cyaneus cyaneus	Lysterfield Park, Victoria, Australia	37°57'S	145°17'E	NA	August-October 2016[§]
Malurus alboscapulatus moretoni	Garuahi village, Milne Bay Province, Papua New Guinea	10°15'S	150°30'E	NA	May-August 2014[†]
(b) Feathers from museum specimens					
Malurus elegans	Western Australia	NA	NA	R 9952	April 2017
Malurus pulcherrimus	Lake King, Western Australia	32°42'S	120°47'60"E	B 18109	April 2017
Malurus amabilis	Cardwell, Queensland, Australia	18°15'S	146°1'12"E	R 9966	April 2017
Malurus lamberti lamberti	Emu Vale, Queensland, Australia	28°13'48"S	152°15'E	B 1656	April 2017
Malurus lamberti assimilis	Port Curtis, Queensland, Australia	23°23'60"S	150°30'E	R 9957	April 2017
Malurus splendens splendens	Western Australia	NA	NA	22287	April 2017
Malurus splendens melanotus	Ouyen area, Victoria, Australia	35°4'12"S	142°19'12"E	B 10684	April 2017
Malurus splendens callainus	Central Australia, Northern Territory	23°43'12"S	133°49'12"E	R 9928	April 2017

[†]Sampled by E. Enbody and J. A. Jones (Tulane University) under the permit IACUC #0395.

[‡]Sampled by M. Fan (Monash University) under the permit BSCI/2013/10.

[§]Sampled by A. McQueen (Monash University) under the permit BSCI/2015/11.

Table S2. Plumage patches of male *Malurus* measured using reflectance spectrometry. For *M. leucopterus leuconotus*, all blue patches (except the tail) were combined; for *M. melanocephalus*, all black patches were included. For *M. coronatus coronatus* and *M. cyaneus cyaneus*, both museum specimens and living birds were measured.

Taxon	Plumage patch	Colour	Sample size
Malurus cyanocephalus bonapartii	crown	blue	22
Malurus cyanocephalus bonapartii	throat	blue	29
Malurus cyanocephalus bonapartii	flank	blue	18
Malurus cyanocephalus bonapartii	rump	blue	22
Malurus cyanocephalus bonapartii	shoulder	blue	22
Malurus coronatus coronatus	crown	purple	717
Malurus coronatus coronatus	crown centre	black	275
Malurus elegans	crown	blue	14
Malurus elegans	cheek	blue	14
Malurus elegans	throat	blue	8
Malurus elegans	shoulder	rufous	17
Malurus pulcherrimus	crown	blue	25
Malurus pulcherrimus	cheek	blue	18
Malurus pulcherrimus	throat	blue	12
Malurus pulcherrimus	shoulder	rufous	31
Malurus amabilis	crown	blue	49
Malurus amabilis	cheek	blue	25
Malurus amabilis	throat	black	13
Malurus amabilis	shoulder	rufous	49
Malurus lamberti lamberti	crown	blue	16
Malurus lamberti lamberti	cheek	blue	8
Malurus lamberti lamberti	shoulder	rufous	15
Malurus lamberti assimilis	crown	blue	122
Malurus lamberti assimilis	cheek	blue	61
Malurus lamberti assimilis	throat	black	27
Malurus lamberti assimilis	shoulder	rufous	122
Malurus cyaneus cyaneus	crown	blue	258
Malurus cyaneus cyaneus	cheek	blue	228
Malurus cyaneus cyaneus	throat	indigo	226
Malurus splendens splendens	crown	blue	16

Malurus splendens splendens	throat	blue	8
Malurus splendens melanotus	crown	blue	32
Malurus splendens melanotus	cheek	blue	16
Malurus splendens melanotus	throat	blue	14
Malurus splendens callainus	crown	blue	28
Malurus splendens callainus	cheek	blue	14
Malurus splendens callainus	throat	blue	12
Malurus leucopterus leuconotus	all*	blue	223
Malurus melanocephalus	all**	black	446
Malurus alboscapulatus moretoni	crown	black	24
Malurus alboscapulatus moretoni	throat	black	12

Table S3. Effects of multiple, different structural elements of feather barb on chromatic (x, y, z and UV chroma) and achromatic (brightness) variation among structural colours in male fairy-wrens. Shown are the results of Bayesian phylogenetic mixed models examining the effects of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity. Dependent variables were chromatic coordinates in the avian visual space (x, y and z) – computed using U-type visual sensitivities – and UV chroma, as well as achromatic brightness (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values ($^{\dagger}P < 0.07$, $*P < 0.05$ and $***P < 0.001$; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

	x	y	z	UV chroma	Brightness
Fixed effects			Posterior mean (95%CI)		
Intercept	-6.11*** (-7.38, -4.93)	-0.21 (-3.61, 2.87)	-2.44* (-3.96, -0.62)	0.41*** (0.31, 0.54)	72.15*** (67.77, 75.63)
Cortex thickness	-0.12 (-0.60, 0.32)	0.11 (-1.12, 1.37)	0.39 (-0.55, 1.16)	0.01 (-0.03, 0.05)	**-1.63* (-3.16, -0.02)**
SL thickness	-0.40 (-0.93, 0.27)	-0.03 (-1.57, 1.46)	**-1.05* (-1.94, -0.01)**	**-0.04† (-0.08, 0.001)**	**1.77† (-0.19, 3.54)**
Ratio melanin/SL	-1.67 (-5.41, 2.59)	0.62 (-0.95, 1.46)	**1.05* (0.19, 2.05)**	-0.02 (-0.06, 0.03)	-0.71 (-2.61, 1.21)
SL spatial frequency	-0.01 (-0.67, 0.67)	-0.53 (-2.35, 1.27)	-0.52 (-1.66, 0.83)	0.01 (-0.03, 0.06)	0.94 (-1.32, 3.35)
Distance between scatterers	-0.50 (-1.31, 0.13)	0.21 (-1.56, 2.34)	-0.27 (-1.56, 0.77)	2.02 (-0.15, 4.48)	**-0.05† (-0.10, 0.003)**
SL irregularity	-0.15 (-0.73, 0.41)	**1.65* (0.01, 3.32)**	0.52 (-0.59, 1.50)	-0.50 (-2.42, 1.64)	-0.02 (-0.07, 0.04)
Marginal R^2	0.27 (0.03, 0.55)	0.43 (0.04, 0.82)	0.64 (0.29, 0.89)	0.65 (0.27, 0.96)	0.30 (0.05, 0.56)
Conditional R^2	0.65 (0.39, 0.89)	0.90 (0.79, 0.99)	0.86 (0.76, 0.94)	0.96 (0.93, 0.99)	0.67 (0.42, 0.92)
λ	0.23 (0.00, 0.66)	0.15 (0.00, 0.53)	0.12 (0.00, 0.41)	0.13 (0.00, 0.50)	0.18 (0.00, 0.60)

Table S4. Separate effects of feather barb structural elements on the value of the x-coordinate. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the x-coordinate value of plumage structural colours of male *Malurus*, computed using U-type visual sensitivities (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values ($^{\dagger}P < 0.07$, $*P < 0.05$ and $***P < 0.001$; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	-6.06*** (-7.47, -4.28)	-6.13*** (-7.82, -4.33)	-6.12*** (-8.00, -4.49)	-6.08*** (-7.65, -4.52)	-6.09*** (-7.58, -4.77)	-6.07*** (-7.41, -4.69)
Cortex thickness	-0.19 (-0.63, 0.21)	-	-	-	-	-
SL thickness	-	-0.38 (-0.83, 0.05)	-	-	-	-
Ratio melanin/SL	-	-	0.28 (-0.15, 0.79)	-	-	-
SL spatial frequency	-	-	-	**-0.43† (-0.87, 0.02)**	-	-
Distance between scatterers	-	-	-	-	**-0.53* (-0.97, -0.10)**	-
SL irregularity	-	-	-	-	-	-0.38 (-0.84, 0.15)
Marginal R^2	0.03 (0.00, 0.10)	0.06 (0.00, 0.19)	0.03 (0.00, 0.10)	0.08 (0.00, 0.24)	0.12 (0.00, 0.32)	0.08 (0.00, 0.25)
Conditional R^2	0.65 (0.36, 0.94)	0.68 (0.38, 0.94)	0.71 (0.39, 0.97)	0.66 (0.36, 0.94)	0.63 (0.33, 0.93)	0.62 (0.30, 0.93)
λ	0.36 (0.00, 0.76)	0.33 (0.00, 0.74)	0.39 (0.00, 0.81)	0.34 (0.00, 0.75)	0.33 (0.00, 0.73)	0.33 (0.00, 0.73)

Table S5. Separate effects of feather barb structural elements on the value of the y-coordinate. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the y-coordinate value of plumage structural colours of male *Malurus*, computed using U-type visual sensitivities (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values (*P < 0.05; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	-0.48 (-4.15, 2.89)	-0.42 (-3.78, 2.76)	-0.51 (-5.01, 2.58)	-0.22 (-3.59, 3.37)	-0.28 (-3.53, 2.63)	-0.22 (-2.86, 2.75)
Cortex thickness	-0.23 (-1.32, 0.96)	-	-	-	-	-
SL thickness	-	-0.29 (-1.57, 0.87)	-	-	-	-
Ratio melanin/SL	-	-	0.40 (-0.83, 1.61)	-	-	-
SL spatial frequency	-	-	-	0.16 (-1.09, 1.34)	-	-
Distance between scatterers	-	-	-	-	0.48 (-0.81, 1.68)	-
SL irregularity	-	-	-	-	-	**1.29* (0.08, 2.53)**
Marginal R^2	0.05 (0.00, 0.22)	0.06 (0.00, 0.24)	0.06 (0.00, 0.25)	0.06 (0.00, 0.26)	0.09 (0.00, 0.38)	0.27 (0.00, 0.70)
Conditional R^2	0.85 (0.61, 0.99)	0.86 (0.58, 0.99)	0.87 (0.64, 1.00)	0.85 (0.58, 0.99)	0.84 (0.58, 0.99)	0.84 (0.65, 0.99)
λ	0.18 (0.00, 0.53)	0.20 (0.00, 0.62)	0.20 (0.00, 0.61)	0.19 (0.00, 0.57)	0.18 (0.00, 0.58)	0.16 (0.00, 0.55)

Table S6. Separate effects of feather barb structural elements on the value of the z-coordinate. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the z-coordinate value of plumage structural colours of male *Malurus*, computed using U-type visual sensitivities (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values (*P < 0.05, **P < 0.01 and ***P < 0.001; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	-2.14 (-5.14, 0.83)	-2.43* (-4.17, -0.51)	-2.36* (-4.33, -0.48)	-1.93 (-4.86, 1.35)	-2.06 (-4.95, 1.35)	-2.09 (-4.85, 1.43)
Cortex thickness	-0.13 (-1.08, 0.94)	-	-	-	-	-
SL thickness	-	**-1.59** ** **(-2.50, -0.81)**	-	-	-	-
Ratio melanin/SL	-	-	**1.65*** ** **(0.77, 2.49)**	-	-	-
SL spatial frequency	-	-	-	**-1.10*** **(-2.07, -0.15)**	-	-
Distance between scatterers	-	-	-	-	-0.91 (-1.90, 0.08)	-
SL irregularity	-	-	-	-	-	-0.35 (-1.46, 0.70)
Marginal R^2	0.05 (0.00, 0.22)	0.47 (0.08, 0.82)	0.45 (0.06, 0.76)	0.17 (0.00, 0.43)	0.14 (0.00, 0.42)	0.06 (0.00, 0.24)
Conditional R^2	0.79 (0.45, 0.99)	0.81 (0.63, 0.97)	0.83 (0.67, 0.96)	0.86 (0.66, 0.99)	0.84 (0.57, 0.99)	0.81 (0.51, 0.99)
λ	0.18 (0.00, 0.57)	0.11 (0.00, 0.38)	0.18 (0.00, 0.52)	0.20 (0.00, 0.64)	0.20 (0.00, 0.59)	0.17 (0.00, 0.55)

Table S7. Separate effects of feather barb structural elements on brightness. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the brightness of plumage structural colours of male *Malurus* (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values (*P < 0.05, **P < 0.01 and ***P < 0.001; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	72.03*** (67.17, 76.05)	72.12*** (67.88, 75.94)	72.10*** (66.44, 76.61)	71.54*** (66.52, 76.57)	71.75*** (67.39, 75.48)	71.78*** (67.85, 76.97)
Cortex thickness	-1.08 (-2.97, 0.76)	-	-	-	-	-
SL thickness	-	**2.00* (0.23, 4.01)**	-	-	-	-
Ratio melanin/SL	-	-	**-2.15* (-4.01, -0.47)**	-	-	-
SL spatial frequency	-	-	-	**2.76** (1.06, 4.49)**	-	-
Distance between scatterers	-	-	-	-	**2.92*** (1.30, 4.63)**	-
SL irregularity	-	-	-	-	-	1.78 (-0.18, 3.84)
Marginal R^2	0.19 (0.00, 0.65)	0.33 (0.00, 0.77)	0.29 (0.00, 0.66)	0.33 (0.02, 0.68)	0.47 (0.09, 0.89)	0.23 (0.00, 0.62)
Conditional R^2	0.89 (0.68, 0.99)	0.91 (0.78, 0.99)	0.93 (0.82, 1.00)	0.95 (0.89, 1.00)	0.94 (0.87, 0.99)	0.92 (0.79, 1.00)
λ	0.11 (0.00, 0.35)	0.11 (0.00, 0.41)	0.15 (0.00, 0.50)	0.17 (0.00, 0.54)	0.13 (0.00, 0.45)	0.13 (0.00, 0.45)

Table S8. Effects of multiple, different structural elements of feather barb on chromatic (x, y and z) variation for V-type eyes among structural colours in male fairy-wrens. Shown are the results of Bayesian phylogenetic mixed models examining the effects of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on chromatic coordinates (xyz) – computed using V-type visual sensitivities – and achromatic brightness (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values ($^{\dagger}P < 0.07$, $*P < 0.05$ and $**P < 0.01$; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

	x	y	z
Fixed effects		Posterior mean (95%CI)	
Intercept	-1.45 (-2.98, 0.30)	-2.29 (-5.59, 0.86)	-3.42** (-4.99, -1.89)
Cortex thickness	-0.09 (-0.77, 0.64)	0.09 (-0.90, 1.38)	0.31 (-0.46, 1.14)
SL thickness	-0.18 (-0.99, 0.70)	-0.20 (-1.51, 1.17)	**-1.11* (-2.10, -0.17)**
Ratio melanin/SL	-0.26 (-1.15, 0.61)	0.51 (-0.82, 1.84)	**0.96* (0.06, 1.78)**
SL spatial frequency	-0.41 (-1.47, 0.69)	-0.50 (-2.08, 1.17)	-0.50 (-1.61,1.78)
Distance between scatterers	-0.12 (-1.18, 0.93)	0.003 (-1.62, 1.87)	-0.37 (-1.50, 0.70)
SL irregularity	-0.33 (-1.13, 0.67)	**1.40† (-0.02, 2.78)**	0.42 (-0.50, 1.49)
Marginal R^2	0.34 (0.06, 0.64)	0.37 (0.03, 0.76)	0.64 (0.26, 0.88)
Conditional R^2	0.70 (0.44, 0.95)	0.89 (0.76, 0.99)	0.86 (0.78, 0.95)
λ	0.13 (0.00, 0.46)	0.19 (0.00, 0.58)	0.14 (0.00, 0.45)

Table S9. Separate effects of feather barb structural elements on the value of the x-coordinate. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the x-coordinate value of plumage structural colours of male *Malurus*, computed using V-type visual sensitivities (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values ($^{\dagger}P < 0.07$ and $^{*}P < 0.05$; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	-1.44† (-2.92, 0.04)	-1.50* (-2.87, 0.07)	-1.45† (-2.89, 0.13)	-1.42* (-2.74, -0.03)	-1.52† (-3.03, -0.09)	-1.46* (-2.80, 0.15)
Cortex thickness	-0.06 (-0.74, 0.56)	-	-	-	-	-
SL thickness	-	-0.21 (-0.82, 0.54)	-	-	-	-
Ratio melanin/SL	-	-	0.07 (-0.66, 0.69)	-	-	-
SL spatial frequency	-	-	-	**-0.62† (-1.21, 0.03)**	-	-
Distance between scatterers	-	-	-	-	-0.53 (-1.21, 0.14)	-
SL irregularity	-	-	-	-	-	-0.56 (-1.19, 0.10)
Marginal R^2	0.05 (0.00, 0.17)	0.06 (0.00, 0.23)	0.05 (0.00, 0.18)	0.17 (0.00, 0.43)	0.14 (0.00, 0.38)	0.15 (0.00, 0.41)
Conditional R^2	0.55 (0.10, 0.92)	0.56 (0.14, 0.93)	0.56 (0.15, 0.93)	0.59 (0.24, 0.92)	0.57 (0.18, 0.92)	0.57 (0.19, 0.92)
λ	0.15 (0.00, 0.46)	0.14 (0.00, 0.48)	0.15 (0.00, 0.49)	0.14 (0.00, 0.46)	0.14 (0.00, 0.45)	0.14 (0.00, 0.44)

Table S10. Separate effects of feather barb structural elements on the value of the y-coordinate. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the y-coordinate value of plumage structural colours of male *Malurus*, computed using V-type visual sensitivities (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding *P*-values, marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	-2.20 (-4.93, 1.25)	-2.30 (-5.97, 0.21)	-2.32 (-5.30, 0.68)	-2.21 (-5.32, 0.56)	-2.27 (-5.63, 0.56)	-2.14 (-4.89, 0.42)
Cortex thickness	-0.30 (-1.21, 0.77)	-	-	-	-	-
SL thickness	-	-0.45 (-1.47, 0.55)	-	-	-	-
Ratio melanin/SL	-	-	0.53 (-0.55, 1.60)	-	-	-
SL spatial frequency	-	-	-	-0.08 (-1.19, 0.87)	-	-
Distance between scatterers	-	-	-	-	0.16 (-1.00, 1.20)	-
SL irregularity	-	-	-	-	-	0.97 (-0.19, 2.15)
Marginal R^2	0.06 (0.00, 0.25)	0.07 (0.00, 0.28)	0.07 (0.00, 0.24)	0.04 (0.00, 0.19)	0.05 (0.00, 0.22)	0.20 (0.00, 0.58)
Conditional R^2	0.81 (0.52, 0.99)	0.83 (0.54, 0.98)	0.83 (0.53, 1.00)	0.82 (0.51, 0.99)	0.82 (0.50, 0.99)	0.82 (0.57, 0.99)
λ	0.19 (0.00, 0.57)	0.20 (0.00, 0.62)	0.22 (0.00, 0.64)	0.20 (0.00, 0.63)	0.21 (0.00, 0.60)	0.19 (0.00, 0.59)

Table S11. Separate effects of feather barb structural elements on the value of the z-coordinate. Shown are the results of Bayesian phylogenetic mixed models examining the separate effects (i.e. not accounting for other morphological variables) of cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity on the z-coordinate value of plumage structural colours of male *Malurus*, computed using V-type visual sensitivities (n = 30). Explanatory variables were centred and scaled. Shown are posterior means for fixed effect coefficients and their 95% credible intervals (CI) and corresponding P-values ($^{\dagger}P < 0.07$, $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$; in bold), marginal (fixed effects) and conditional (fixed + random effects) R^2 values, and lambda (phylogenetic signal).

Fixed effects	Posterior mean (95%CI)					
Intercept	-3.20* (-5.98, -0.26)	-3.42** (-4.98, -1.61)	-3.37** (-5.15, -1.32)	-2.88† (-5.47, 0.26)	-3.03† (-5.78, -0.02)	-3.09† (-5.83, 0.49)
Cortex thickness	-0.13 (-1.08, 0.94)	-	-	-	-	-
SL thickness	-	**-1.63** (-2.50, -0.86)**	-	-	-	-
Ratio melanin/SL	-	-	**1.67*** (0.78, 2.39)**	-	-	-
SL spatial frequency	-	-	-	**-1.20** (-2.06, -0.22)**	-	-
Distance between scatterers	-	-	-	-	**-1.04* (-1.98, -0.08)**	-
SL irregularity	-	-	-	-	-	-0.50 (-1.61, 0.70)
Marginal R^2	0.05 (0.00, 0.19)	0.49 (0.12, 0.84)	0.44 (0.07, 0.75)	0.19 (0.00, 0.45)	0.17 (0.00, 0.45)	0.08 (0.00, 0.28)
Conditional R^2	0.80 (0.48, 0.99)	0.81 (0.66, 0.95)	0.84 (0.70, 0.97)	0.86 (0.68, 0.99)	0.84 (0.60, 0.99)	0.81 (0.52, 0.99)
λ	0.20 (0.00, 0.59)	0.12 (0.00, 0.47)	0.22 (0.00, 0.58)	0.21 (0.00, 0.62)	0.22 (0.00, 0.63)	0.19 (0.00, 0.60)

Table S12. Phylogenetic effects on feather colour and morphological variables. Shown are the results of Bayesian phylogenetic mixed models examining the degree of phylogenetic constraint on a) the x-, y, and z-coordinates computed using U- and V-type visual sensitivities and brightness, and on b) cortex thickness, spongy layer (SL) thickness, ratio of melanin to SL, spatial frequency of the SL (from Fourier analysis), distance between scatterers (sum of the diameters of keratin channels and air spaces) and SL irregularity (n = 30). Shown are posterior means for lambda (phylogenetic signal) and the 95% credible intervals.

a)

	U-type			V-type			-
	x	y	z	x	y	z	Brightness
λ	0.34 (0.00, 0.73)	0.19 (0.00, 0.60)	0.19 (0.00, 0.57)	0.15 (0.00, 0.49)	0.20 (0.00, 0.59)	0.19 (0.00, 0.57)	0.13 (0.00, 0.44)

b)

	Cortex thickness	SL thickness	Ratio melanin/SL	SL spatial frequency	Distance between scatterers	SL irregularity
λ	0.006 (0.00, 0.02)	0.002 (0.00, 0.005)	0.17 (0.00, 0.60)	0.12 (0.00, 0.41)	0.17 (0.00, 0.50)	0.22 (0.00, 0.63)

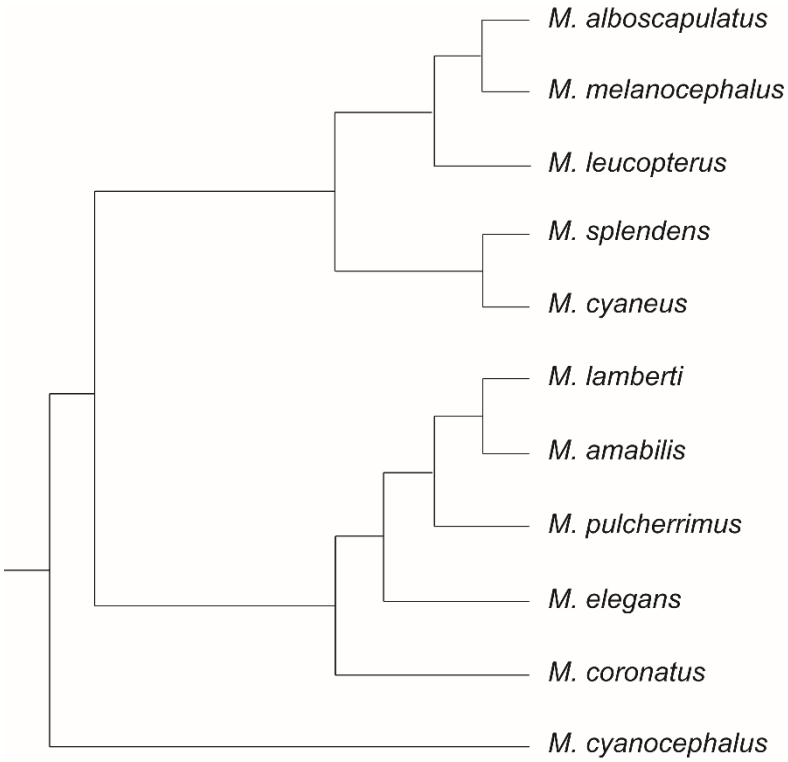


Figure S1. Relationships among *Malurus* species. Diagram based on the supermatrix phylogeny of Marki et al. (2017)**.**

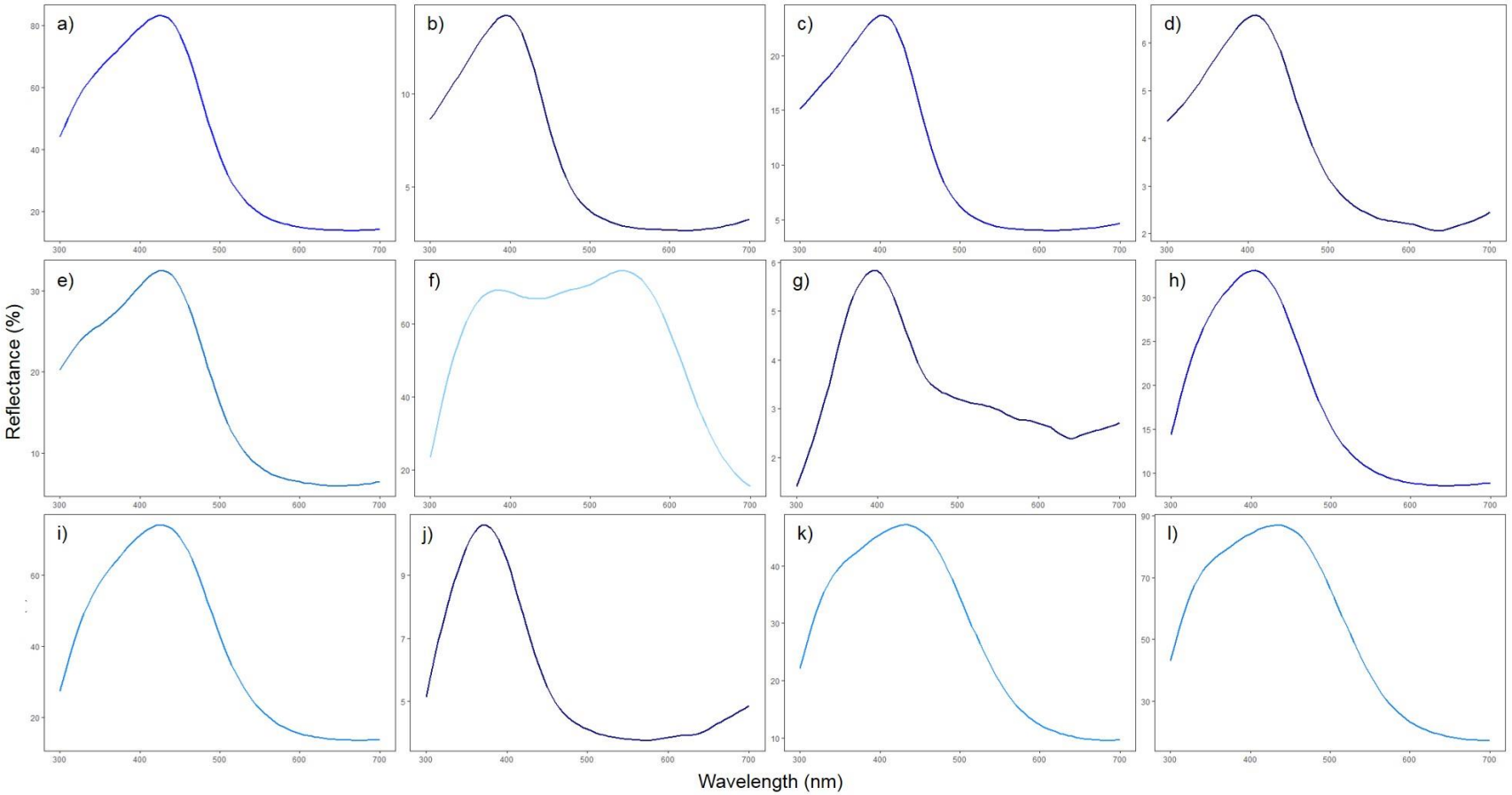


Figure S2. Average (smoothed) reflectance spectra of plumage patches of male *Malurus*. a) crown, b) throat, c) flank, d) rump and e) shoulder of male *M. cyanocephalus bonapartii*; f) cheek and g) throat of male *M. elegans*; h) crown, i) cheek and j) throat of male *M. pulcherrimus*; k) crown and l) cheek of male *M. amabilis*. Line colours correspond approximately to the feather colour perceived by the human eye.

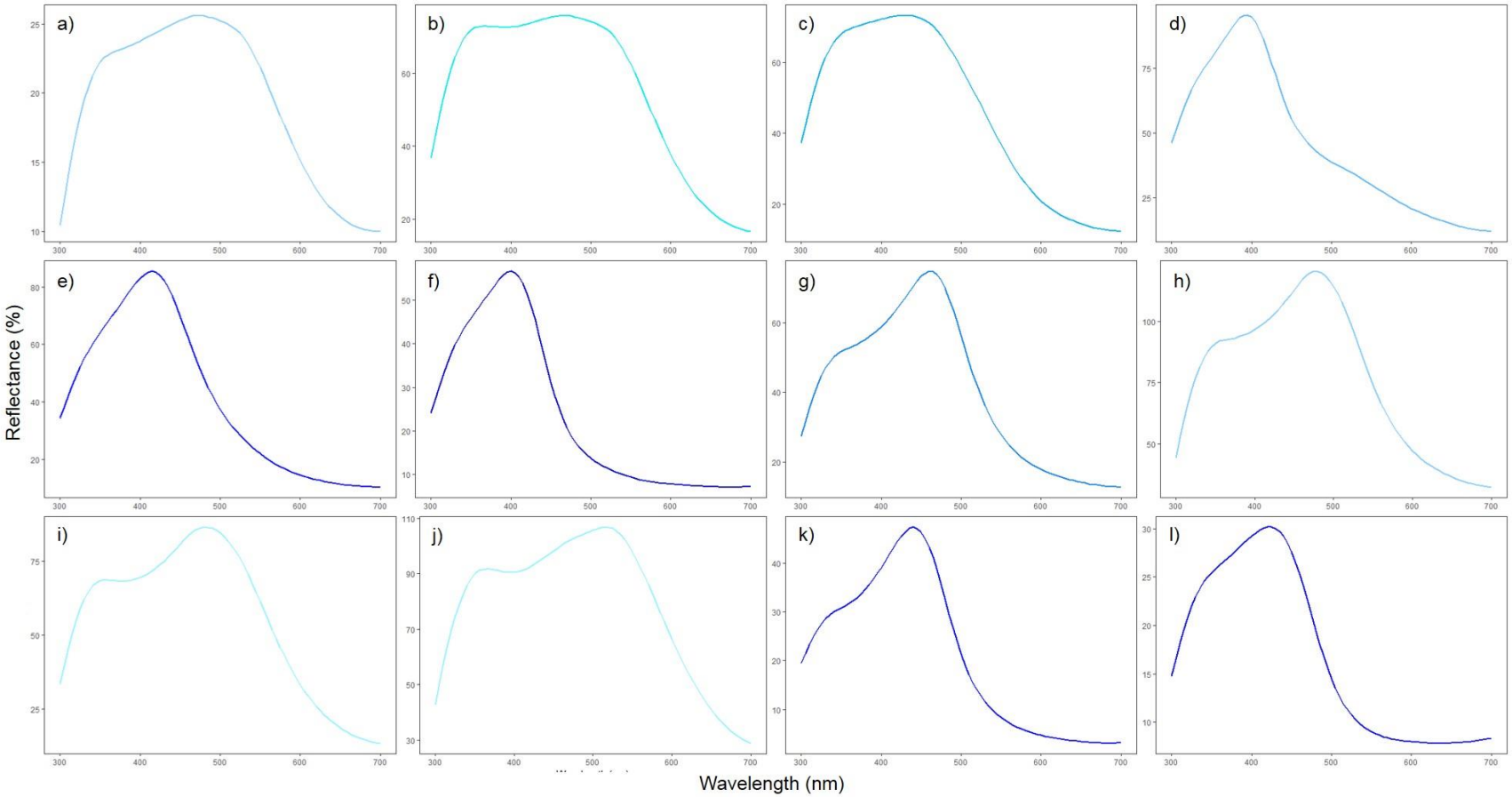


Figure S3. Average (smoothed) reflectance spectra of plumage patches of male *Malurus*. a) crown and b) cheek of male *M. lamberti lamberti*; c) cheek of male *M. lamberti assimilis*; d) crown of male *M. cyaneus cyaneus*; e) crown and f) throat of male *M. splendens splendens*; g) crown and h) cheek of male *M. splendens melanotus*; i) crown, j) cheek and k) throat of male *M. splendens callainus*; l) the whole body except the wings and tail of male *M. leucopterus leuconotus*. Line colours correspond approximately to the feather colour perceived by the human eye.

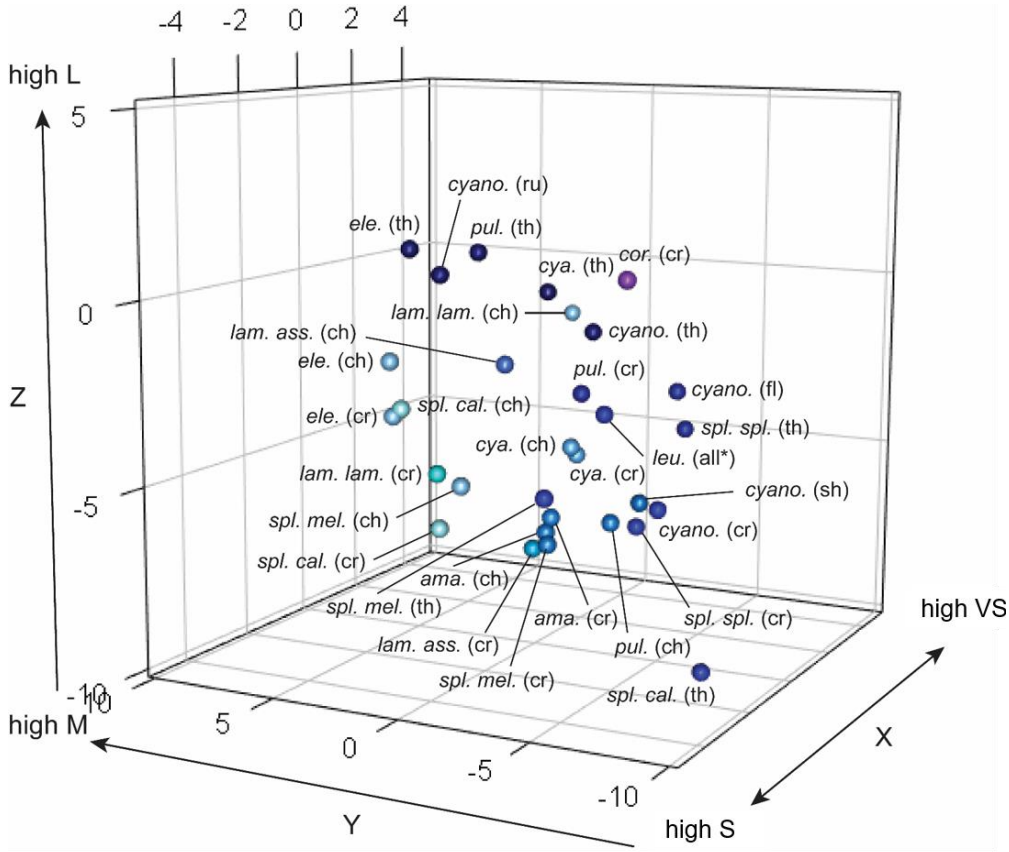


Figure S4. Average positions of the studied colour patches in the visual space of V-type species. The X-axis represents stimulation of the VS cone relative to the S cone, higher values of the Y-axis represents higher stimulation of the M cone relative to VS and S cones, while the Z-axis represents higher relative stimulation of the L cone compared with the other three (units = jnd). Taxon abbreviations: *ama. = M. amabilis*, *cor. = M. coronatus (coronatus)*, *cya. = M. cyaneus (cyaneus)*, *cyano. = M. cyanocephalus (bonapartii)*, *ele. = M. elegans*, *lam. ass. = M. lamberti assimilis*, *lam. lam. = M. lamberti lamberti*, *leu. = M. leucopterus (leuconotus)*, *pul. = M. pulcherrimus*, *spl. cal. = M. splendens callainus*, *spl. mel. = M. splendens melanotus*, *spl. spl. = M. splendens splendens*. Plumage patch abbreviations: cr = crown, ch = cheek, fl = flank, th = throat, ru = rump, sh = shoulder; all* = the whole body except the wings and tail for *M. leucopterus*. Sphere colours correspond approximately to the feather colour perceived by the human eye.

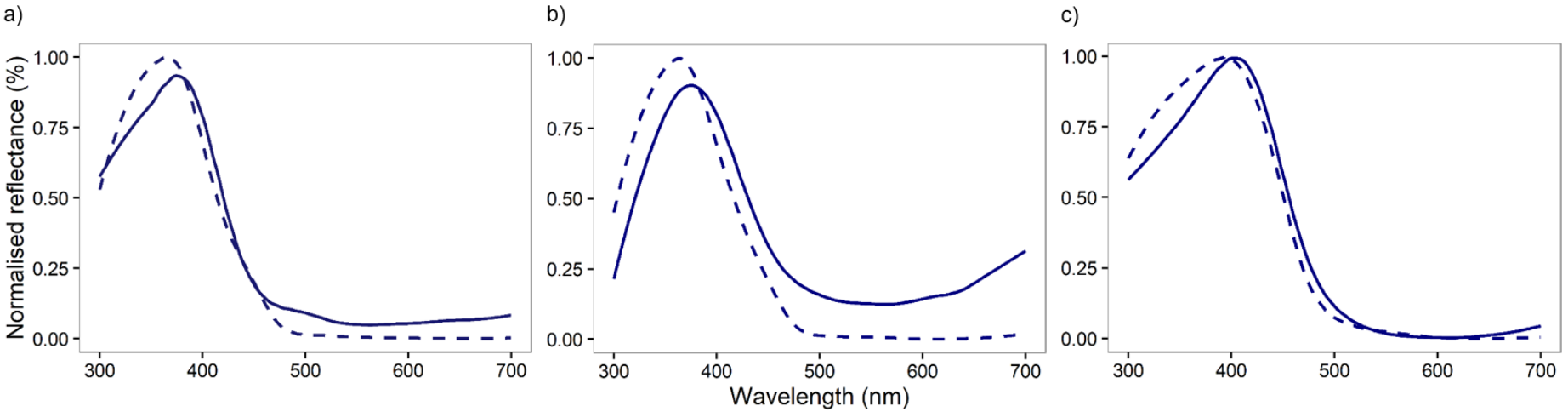


Figure S5. Barbules have a negligible influence on feather colour in the studied plumages. Shown are examples of (normalised and smoothed) reflectance spectra measured on the whole plumage patch (solid line) and on a single barb (dotted line): a) throat of male *M. cyaneus cyaneus*, b) throat of male *M. pulcherrimus*, and c) flank of male *M. cyanocephalus bonapartii* (all three plumage patches are formed by feathers comprising barbules).

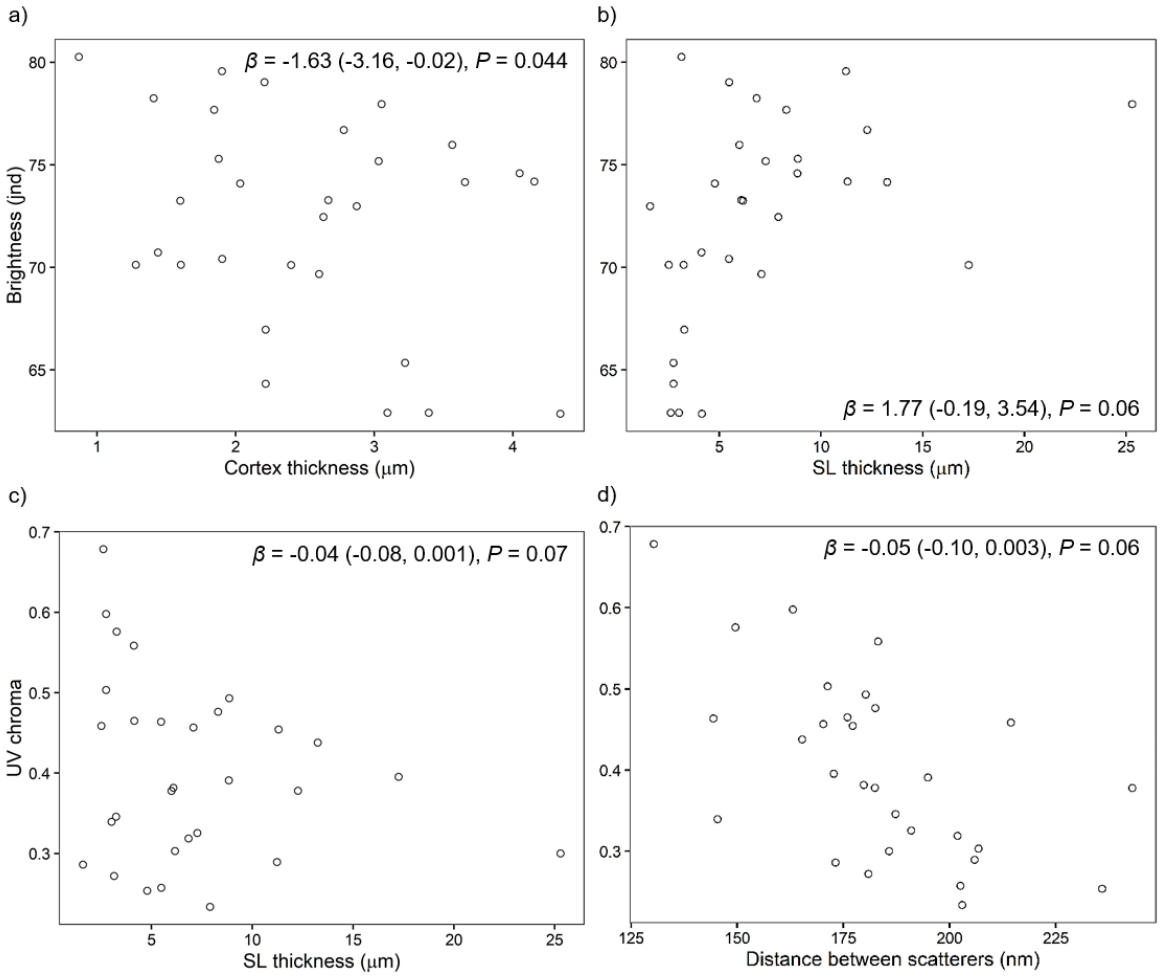


Figure S6. Brightness and UV chroma of structural colours in male fairy-wrens are related to the thickness of the cortex and spongy layer, as well as the distance between scatterers. Shown are brightness (in jnd) as a function of a) the thickness of the cortex (in μm) and b) the thickness of the spongy layer (SL; in μm), and UV chroma as a function of c) the thickness of the spongy layer (in μm) and d) the distance between scatterers (in nm). For each plot are indicated the posterior mean (β) and 95% credible interval for the depicted correlation and the corresponding *P*-value.

Reference

Marki, P. Z., Jønsson, K. A., Irestedt, M., Nguyen, J. M., Rahbek, C., & Fjeldså, J. (2017). Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, **107**, 516-529.

Chapter 2

No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy?

This thesis chapter is in the same form as the final manuscript published in the peer-reviewed journal *Behavioral Ecology*. The full reference for the published paper is:

Fan, M., Hall, M. L., Kingma, S. A., Mandeltort, L. M., Hidalgo Aranzamendi, N., Delhey, K., & Peters, A. (2017). No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy? *Behavioral Ecology*, **28**, 1055-1067.

Behavioral Ecology

The official journal of the
ISBE
International Society for Behavioral Ecology

Behavioral Ecology (2017), 28(4), 1055–1067. doi:10.1093/beheco/arx065

Original Article

No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy?

Marie Fan,[a] Michelle L. Hall,[b,c] Sjouke A. Kingma,[c,d] Lisa M. Mandeltort,[a] Nataly Hidalgo Aranzamendi,[a] Kaspar Delhey,[a,c] and Anne Peters[a,c]

[a]School of Biological Sciences, Monash University, 25 Rainforest Walk, Clayton, Victoria 3800, Australia, [b]School of BioSciences, University of Melbourne, Melbourne, Parkville, Victoria 3010, Australia, [c]Max Planck Institute for Ornithology, Vogelwarte Radolfzell, Schlossallee 2, D-78315 Radolfzell, Germany, and [d]Behavioural and Physiological Ecology, Groningen Institute for Evolutionary Life Sciences, University of Groningen, PO Box 11103, 9700CC Groningen, The Netherlands

Received 1 December 2016; revised 26 February 2017; editorial decision 27 February 2017; accepted 6 April 2017; Advance Access publication 27 April 2017.

The evolution of male ornamentation has long been the focus of sexual selection studies. However, evidence is accumulating that sexually selected traits can also be lost, although the process is ill-understood. In male fairy-wrens (*Malurus* spp.), early molt into the seasonal breeding plumage is critical for obtaining extra-pair paternity (EPP), which reaches very high levels in these socially monogamous songbirds. A notable exception is the purple-crowned fairy-wren, *Malurus coronatus*, which, like its congeners, breeds cooperatively, but where EPP is very rare. Nevertheless, males develop a conspicuous seasonal breeding plumage at highly variable times. Based on 6 years of molt data collected for 137 individuals, we investigated the adaptive significance of pre-breeding molt timing as a sexual signal under (near) genetic monogamy. Molt timing varied between and within individuals with age and climate: molt was completed earlier in older males and after wetter years. Despite its potential to act as a sexual signal of male quality, fitness benefits and costs of early molt appear limited: molt timing did not correlate with 1) the likelihood of gaining a breeding position; 2) female mate preference (EPP/cuckoldry, divorce); 3) female reproductive investment (breeding timing, clutch size, number of clutches); 4) breeding performance (hatching success, fledging success, fledgling survival, annual reproductive success); and 5) male survival. However, although molt timing did not predict which subordinates would become breeders, breeders molted earlier than subordinates. The lack of EPP in this species might imply relaxed sexual selection on early molt with potential to lead to trait disappearance.

Key words: **evolutionary trait loss, extra-pair paternity, monogamy, pre-breeding molt timing, relaxed sexual selection, seasonal breeding plumage.**

INTRODUCTION

A long-standing goal in evolutionary biology is to understand why novel traits arise and how these traits contribute to an individual's fitness. In particular, the evolution of elaborate, conspicuous ornamentation of males has been the focus of many studies in various taxonomic groups (Darwin 1871; Andersson 1994; Andersson and Simmons 2006). Such ornamental traits may signal different types of information related to the bearer's overall body condition and genetic constitution (Zahavi 1975; Hamilton and Zuk 1982; Cotton et al. 2004; Andersson 2006). Males with more elaborate ornamentation are often of higher quality, being preferred as mates by females or more successful when competing for access to females or resources such as territories (Darwin 1871). As a result, such males are generally thought to achieve higher mating success, a benefit that presumably offsets the apparent survival cost of ornamental expression (Lozano 1994; Hill 2002).

In contrast, far less is known about the evolutionary loss of traits, in particular sexually selected traits, although recent phylogenetic studies have shown it to be taxonomically widespread (Wiens 2001; e.g. insects—Emlen et al. 2005; fish—Basolo 1996; amphibians—Emerson 1996; lizards—Wiens 1999; Quinn and Hews 2000; Ord and Stuart-Fox 2006; birds—Ödeen and Björklund 2003; de Kort and ten Cate 2004). Understanding why and how sexually selected traits are lost is important as it can provide valuable insights into the

Address correspondence to M. Fan. E-mail: marie.fan@monash.edu.

nature and role of various sources of selection and how such sources might interact and affect individual fitness (Wiens 2001; Lahti et al. 2009). However, determining causes and mechanisms of trait loss is challenging, which might explain why this area of research remains relatively unexplored (Wiens 2001; Porter and Crandall 2003; Lahti et al. 2009). Evolutionary loss of traits is assumed to be the result of a weakening or removal of one or several sources of selection maintaining the trait ("relaxed selection"; Lahti et al. 2009). For example, sexually selected male traits may wane when the strength of female choice is overcome by random or environmental effects, such as genetic drift or ecological shifts, resulting in female preferences being reduced, lost or reversed (Basolo 1996; Wiens 1999, 2001; Ödeen and Björklund 2003; Ord and Stuart-Fox 2006; Wong and Rosenthal 2006). Trait loss can occur in various gradations, ranging from complete loss to vestigialization, when the trait is still present in a degraded form. Alternatively, traits can persist due to other remaining sources of selection or positive correlations with other functional traits, although such processes are not always easy to identify (Lahti et al. 2009; Ellers et al. 2012).

Avian species have frequently been used as model systems to investigate the evolution of sexually selected traits as males of many species display colorful sexual ornaments (Owens and Hartley 1998; Dunn et al. 2001; McGraw et al. 2002). In particular, extra-pair paternity (EPP) has been suggested to be an important driver of sexual selection for the evolution of male plumage ornamentation in birds, especially in socially monogamous species that have low apparent variance in male mating success (Møller and Birkhead 1994; Owens and Hartley 1998; Dunn et al. 2001). The ability of pair-bonded males to fertilize females other than their social mate increases this variance (Dunn et al. 2001) and females might receive indirect genetic benefits from these extra-pair fertilizations (although evidence for such benefits remains limited; Griffith et al. 2002; Akçay and Roughgarden 2007; Forstmeier et al. 2014). In some species, females may assess potential extra-pair partners on the basis of their elaborate plumage features, which can therefore be subject to strong directional female preference (Møller and Birkhead 1994; Owens and Hartley 1998; Dunn et al. 2001). If such a selective force is weakened or removed, we expect the balance of selection pressures maintaining conspicuous ornamentation to change, affecting its evolution and maintenance. Several studies based on phylogenetic reconstruction have provided evidence that the loss of ornamental plumage coloration is common across taxa and may occur in either or both sexes within a species, and in one or multiple species within a genus, typically because of a weakening of sexual selection due to ecological constraints (Irwin 1994; Omland 1997; Burns 1998; Omland and Lanyon 2000; Schroeder et al. 2009). As a consequence, we might hypothesize that when female extra-pair mate choice is no longer operating in a monogamous system, evolutionary loss of male ornamentation is likely to occur unless other selective forces are still acting to maintain such ornamentation.

A classic example of male sexual ornaments associated with EPP is found among fairy-wrens (*Malurus* spp.), an Australo-Papuan genus in which males display a conspicuous seasonal ornamental plumage (Rowley and Russell 1997). The fairy-wren mating system is dominated by EPP that has been reported to reach exceptionally high levels in several species (>50% of broods, up to 95% in *M. cyaneus*; Kingma et al. 2009; Peters et al. 2013). Typically, the timing of breeding plumage acquisition is highly variable between males, with some males developing the breeding plumage many months prior to the start of breeding (reviewed in Peters et al. 2013). In several species, including the superb (*M. cyaneus*) and the red-winged (*M. elegans*) fairy-wrens, the timing of breeding plumage acquisition determines the likelihood that a male obtains EPP, and there appears to be strong directional sexual selection for early molt into conspicuous ornamental plumage (at least in some years; Cockburn et al. 2008; Brouwer et al. 2011; van de Pol et al. 2012; Peters et al. 2013). In contrast, very little is known about how molt timing might relate to EPP and more generally to mating success in other groups of sexually dimorphic bird species (e.g. whydahs, *Vidua* spp.—Barnard 1995; mallard, *Anas platyrhynchos*—Omland 1996).

A notable exception to the well-known high EPP levels characterizing Australian fairy-wrens is the purple-crowned fairy-wren *M. coronatus*, in which only 5% of broods have EPP (Kingma et al. 2009; Hidalgo Aranzamendi et al. 2016), despite a broadly similar ecology and social system (Kingma et al. 2009). Similar to other fairy-wrens, the purple-crowned fairy-wren is a cooperative breeder with long-term social partnerships and seasonal plumages, as once a year both (dominant) breeder and subordinate males molt from a dull brown non-breeding plumage into a bright purple-and-black breeding plumage (Rowley and Russell 1997). Interestingly, like other fairy-wrens, the timing of pre-breeding molt shows large variation between and within males (Peters et al. 2013). Because there is virtually no opportunity for female extra-pair mate choice in this species, male pre-breeding molt timing may therefore represent a case of a vestigial sexually selected trait and constitutes a very good candidate for the study of relaxed selection and evolutionary trait loss. Alternatively, molt timing may have been co-opted to signal male quality in another context either to the social female or to potential rivals and this could be the reason for its presence despite low levels of EPP.

Here, we assess this possibility by examining: 1) the nature of the intrinsic and environmental factors underlying variability in pre-breeding molt timing (to test for condition-dependence) and 2) the consequences of this variability on different aspects of male fitness. As sexually selected traits may evolve through inter- and/or intrasexual selection, we systematically test the adaptive hypotheses reflecting how both types of selection can operate in monogamous mating systems and show that none of them is supported in our study species, despite substantial variation in molt timing. Based on a phylogenetic analysis of the genus *Malurus*, we discuss the possibility of a trait loss scenario in *M. coronatus*, which could explain the occurrence of such a variable trait in a genetically monogamous fairy-wren.

METHODS

Study species

We studied a color-banded population of *Malurus coronatus coronatus* resident along Annie Creek and the Adcock River at Australian Wildlife Conservancy's Mornington Wildlife Sanctuary in northwest Australia (17°31'S, 126°6'E) from July 2005 to November 2015. The species is restricted to patchy riparian vegetation of *Pandanus aquaticus* and groups maintain all-purpose territories year-round, linearly arranged along creeks and rivers (Rowley and Russell 1993, 1997). Like most other species of fairy-wren, *M. coronatus* breeds cooperatively. Only the dominant male and female reproduce, and they are often (40–70% of pairs) accompanied by a number of non-breeding male and female subordinates (mostly offspring from previous broods), of which most contribute to nestling feeding (Kingma et al. 2010, 2011a, 2011b). *M. coronatus* can breed year-round with a distinct peak in breeding activity during

the wet season (December–March), and a smaller peak in the late dry season (August–September) in some years (Rowley and Russell 1993, 1997; Hall and Peters 2009; Peters et al. 2013). Unlike other fairy-wrens where EPP rates are very high, *M. coronatus* engages in very limited extra-pair mating (in 5% of broods, and only in 3% of broods from non-incestuously mated pairs; Kingma et al. 2009, 2013; Hidalgo Aranzamendi et al. 2016).

Once per year, males undergo a pre-breeding molt where the brown non-breeding head plumage is replaced by purple and black feathers (Rowley and Russell 1997; Peters et al. 2013; Figure 1). Other plumage patches do not change noticeably in coloration over the year, including the black cheek patches and the blue tail, and the rest of the plumage which is mainly brown above and buff-white below (Rowley and Russell 1993, 1997; Delhey et al. 2013; Figure 1). Most males initiate pre-breeding molt in July–September before breeding starts, but in a proportion of males molt overlaps with the start of the breeding season (Rowley and Russell 1997; Peters et al. 2013; see also results section). It may be noted that females also undertake a pre-breeding molt of the head once per year, where the brown non-breeding head plumage is replaced by slate-grey feathers (Rowley and Russell 1997), although this is not the focus of the present study.

Field data collection

From July 2005 to March 2011, weekly population censuses were conducted year-round to document plumage coloration, group size and social status of each uniquely color-banded male. Each bird could be unambiguously assigned breeder (dominant) or subordinate status from behavioral cues (the most obvious being that only the dominant pair sings duets; Hall and Peters 2008, 2009). At each sighting, plumage coloration of each individual was scored as a percentage of the complete breeding plumage. Throughout the study, birds were routinely captured and tarsus length was measured using caliper to the nearest 0.1 mm. Tarsus length could be an important indicator of male quality as body size correlates with song frequency (pitch) in male *M. coronatus* and male songs are thought to be sexually selected in the genus *Malurus* (Hall et al. 2013). We used average tarsus length in mm for all available adult captures, adjusted for handler (i.e. predicted values of a model including the identity of the handler as a random intercept).

Nesting activity was monitored by closely following the female, and once found, nests were checked regularly to determine laying date, clutch size (1–4 eggs), hatching and fledging success (for details see Kingma et al. 2011a; Hidalgo Aranzamendi et al. 2016). All nestlings were individually banded on approximately day 7 and parentage of all offspring was determined by genotyping all individuals in our population using 6 or 9 microsatellite loci (for details see Kingma et al. 2009, 2013). After the nestling period, the nest area was monitored to locate the fledglings and determine their survival. Birds captured as adults at the start of the study were classified as "age unknown" with a minimum age based on the presence or absence of offspring (of known age) or the completeness of the breeding plumage. As birds were followed throughout their life, instances of divorce (dissolution of a pair bond where both individuals remain alive) were also documented (for details see Hidalgo Aranzamendi et al. 2016).

From October 2011 to November 2015, biannual population censuses were conducted in November and May–June (for details see Hidalgo Aranzamendi et al. 2016), documenting plumage coloration, group size and social status of all individuals. All new unbanded birds (fledglings, subordinates, or immigrants) were banded, aged by age-specific development of appearance and behavioral cues (tail length, begging behavior, plumage color, bill color) and their parentage was determined using 9 microsatellite loci (Kingma et al. 2013; Hidalgo Aranzamendi et al. 2016). Newly banded birds for which parentage could not be assigned were considered as immigrants and classified as "age unknown", being at least 3 months old (for details see Hidalgo Aranzamendi et al. 2016).

Figure 1
Non-breeding and breeding plumages in male purple-crowned fairy-wrens. Photographs show a male in non-breeding plumage viewed from the side (a) and from above (b), and a male in breeding plumage viewed from the side (c) and from above (d).

From 2007 onwards, intensive yearly censuses covering almost all suitable habitat along the tributaries that join the study site were conducted to find birds that had dispersed outside the core area (emigrants), providing reliable information on the survival (presence or absence) of all individuals (for details see Hidalgo Aranzamendi et al. 2016). Birds were declared dead on the basis of failure to sight them in regular surveys and assigned a death date estimate with a given error (mean = 59.6 ± 5.6 days, range: 3–354.5 days; for details see Supplementary Appendix S1).

Territories are stable year-round, and most boundaries remain stable through the years, being easily determined from movement patterns of birds in each group and locations of agonistic interactions between groups. Occasional changes in boundaries (shifts, territory splitting, or establishment of new territories) were recorded throughout the study. Territory quality was assessed based on the proportion of the territory covered by *Pandanus aquaticus* following methods described in Kingma et al. (2011a). *M. coronatus* does not occupy habitat without *Pandanus* and the distribution of *Pandanus* varies considerably between territories. Territories with greater *Pandanus* cover have lower nest failure (Hidalgo Aranzamendi et al., unpublished data) and more subordinates (Kingma et al. 2011b), indicating that these territories are more productive or attractive for subordinates. Moreover, dominant females target territories with greater *Pandanus* cover during breeder dispersal (i.e. divorce; Hidalgo Aranzamendi et al. 2016).

Daily records of rainfall were obtained from a local weather station at Mornington Wildlife Sanctuary from October 2004 to December 2015 (Australian Bureau of Meteorology weather station 002076).

Individual molt profiles

Using the scores of plumage coloration collected from July 2005 to March 2011, molt profiles were derived for each male. A year was defined as starting on 01 July (mid-winter), and ending on 30 June of the following year. The date of pre-breeding molt completion was determined as the date (number of days from 01 July) halfway between 1) the last date an individual was observed with an incomplete breeding plumage (<100% purple) and 2) the first date it was observed with a complete breeding plumage (100% purple). Half the duration between these 2 dates constituted the error of the pre-breeding molt completion date estimate (mean = 14.3 ± 1.1 days, range: 1–170 days). When birds were not observed frequently enough, the obtained estimates were unreliable (error > 21 days, half the average duration of pre-breeding molt) and excluded from further analysis (n = 85 of 390 excluded). When birds did not develop a complete breeding plumage (i.e. final percentage of purple <100%), the maximum score was used in a similar manner to determine the pre-breeding molt completion date. Some birds temporarily interrupted the pre-breeding molt at an intermediate coloration score before reaching a complete breeding plumage (mean = 56 ± 4% of coverage, range: 15–90%), before resuming it subsequently. When this interruption exceeded the average pre-breeding molt duration of 6 weeks, these males were classified as having "interrupted pre-breeding molt" and the profiles were excluded from overall analysis (n = 9 first-year males, 14 second-year and older males, and 5 males of unknown age; i.e. a total of 28).

Statistical analyses

We first analyzed (A) whether the timing of molt completion depended on intrinsic and environmental factors. Then, (B) in order to identify potential benefits and costs associated with early molt, we investigated correlations between molt completion timing and various fitness variables related to both inter- and intrasexual selection. Hereafter we will simply use the term "molt date" to refer to the pre-breeding molt completion date. Details of all statistical models follow below.

(A) Intrinsic and environmental effects on timing of pre-breeding molt

We built a linear mixed model (LMM) with molt date as a response variable, and age, social status (dominant or subordinate), tarsus length, territory quality and group size as fixed effects. Bird identity, territory identity and year were included as random intercepts to account for non-independence in the data. We first restricted this analysis to birds whose age was accurately known (n = 130) and showed that the molt date only significantly varied between the first and second years of life, and not beyond (Tables 1 and Supplementary Table S1; for details see Supplementary Appendix S2 (A)). As a consequence, this analysis and subsequent ones were performed using 2 age classes, "1" and "2+", and birds of unknown age but known to be at least in their second year of life (n = 120) were included in the "2+" class. Adjusted repeatability of individual molt behaviour was calculated from the last full model (with 2 age levels, n = 247) and the 95% confidence interval inferred from a parametric bootstrap, following Nakagawa and Schielzeth (2010).

As we found that dominant (breeder) males molted earlier than subordinate males, we tested whether this difference was due to within-individual change. To do this, we extracted the molt dates in 2 consecutive years for 1) subordinates who remained subordinate in the second year (n = 19) and 2) subordinates who became dominant in the second year (n = 16; i.e. a total sample size of n = 35 observations), and standardized them relative to the subordinate population for each year (i.e. for a given year we calculated the average date of subordinates and subtracted it from the actual date of each individual). We then built an LMM with the difference between 2 consecutive standardized molt dates as a response variable, and the interaction between the status change (either "from subordinate to subordinate" or "from subordinate to dominant") and the age class change (either "from 1 to 2+" or "from

Table 1

Older and dominant males produce the annual breeding plumage earlier

Parameter	β	SE	df	t value	P
Fixed effects					
Intercept	173.44	8.58	35.76	20.23	<0.001
Age†	**−63.55**	**7.12**	**235.76**	**−8.92**	**<0.001**
Status*	**−28.81**	**6.11**	**174.14**	**−4.71**	**<0.001**
Tarsus length	−3.50	3.95	77.48	−0.89	0.38
Territory quality	1.08	0.60	52.61	1.80	0.08
Group size	0.69	1.36	223.20	0.50	0.62
Random effects					
$\sigma^2_{individual}$	249.45				
$\sigma^2_{territory}$	141.21				
σ^2_{year}	126.60				
$\sigma^2_{residual}$	680.38				

Shown are results from an LMM examining the effects of age (levels 1, 2+), social status, tarsus length, territory quality, and group size on the date of pre-breeding molt completion in individuals of known age and individuals of unknown age at least in their second year of life (n = 247). Significant values (P < 0.05) are in bold.
Reference categories are†age 1, *subordinate.

2+ to 2+") as a fixed effect. Bird identity was included as a random intercept.

To determine whether population-level molt completion timing depended on rainfall, we used all the scores of plumage coloration collected from July 2005 to November 2015 and determined for each individual whether it had reached a complete breeding plumage (≥95%) by the end of November for each year. Therefore, this approach combined both molt timing and final percentage of purple, which allowed us to use 11 years of data. Specifically, we investigated whether the likelihood of being fully purple (≥95%) by the end of November depended on the cumulative rainfall over the past 12 months (from November of the preceding calendar year to October). We built a generalized linear mixed model with penalized quasi-likelihood (GLMMPQL) with the plumage coloration status (fully purple = 1, partially purple = 0) as a binomial response variable, and age, social status and the cumulative rainfall over the past 12 months as fixed effects. Bird identity nested in territory identity was included as a random intercept. Using the cumulative rainfall of the previous wet season (from November of the preceding calendar year to March) gave very similar results (past 12 months *vs* previous wet season: marginal $R^2 = 0.658/0.655$, conditional $R^2 = 0.981/0.980$).

(B) Fitness benefits and costs of early molt

We systematically evaluated all hypothesized sexually selected fitness benefits and costs of early molt that could apply to males in a socially monogamous year-round pair bond.

Gain of a breeder position. We tested whether the timing of molt completion predicted the likelihood of gaining a dominant (breeder) position among subordinates as this is a critical determinant of male reproductive success (subordinates do not participate in reproduction, except for 2 cases recorded in our study population; Kingma et al. 2009). To do so, we built a GLMMPQL with the annual status change (obtained a breeder position within the year = 1, remained subordinate within the year = 0) as a binomial response variable, and molt date (relative to the subordinate population for each year), age and tarsus length as fixed effects. Bird identity nested in territory identity nested in year was included as a random intercept. Cases where such a position was gained by inheritance ($n = 6$) or splitting of the natal territory ($n = 9$) were excluded.

As the opportunity of gaining a breeder position only arises when a dominant male of a territory dies or moves, and because males usually do not disperse far from their natal territory, the opportunity to compete for a vacancy is probably not equal among subordinates. For these reasons, we also performed a case-by-case analysis for each record of a subordinate gaining a breeder position by dispersing in a given year and whose molt date was known for this particular year ($n = 12$). For each case, by using information on territories (group composition and geographical location), we assessed which other males located within the same distance to the vacancy as the "winner" could also compete. Then, we compared the molt date between the competitors using Welch's *t*-test.

Timing of breeding. We tested whether early molt was related to the quality of the female partner. Because high quality females in other species often start breeding earlier (McGraw et al. 2001; Sheldon et al. 2003; Garant et al. 2007), we tested whether timing of molt completion was correlated to the dominant female's breeding timing and whether such a correlation differed between dominants and subordinates. To do this, for territories where the date of the first egg laid by the dominant female was known ($n = 161$), we built an LMM with molt date as a response variable, and age (levels 1, 2+) as well as the interaction between social status and the date of the first egg laid on a territory as fixed effects. Bird identity, territory identity and year were included as random intercepts.

Social female reproductive investment and social reproductive success. Because high quality females in other species also often have higher annual reproductive output (McGraw et al. 2001; Sheldon et al. 2003; Garant et al. 2007), we tested whether male molt timing and female reproductive investment (clutch size and number of clutches) were correlated. Subsequently, we compared male social reproductive success. We also investigated the risk of a male being divorced (females initiate divorce; Hidalgo Aranzamendi et al. 2016) as a proxy for female preference for early molting males.

For all the models described below, we excluded $n = 5$ incestuous pairs of 186 pairs as those have a higher probability to end in divorce for inbreeding avoidance (Hidalgo Aranzamendi et al. 2016) and hatching success in incestuous pairs is known to be much lower (>30% of reduction compared to non-incestuous pairs; Kingma et al. 2013). We standardized the molt dates relative to the dominant population for each year. In addition, in all models, bird identity, territory identity and year were included as random intercepts (nested in each other if a GLMMPQL was used).

To test whether reproductive investment of a female varied with her partner's molt date, we built a generalized linear mixed model (GLMM) with the annual number of attempts (0–10) as a count response variable and a zero-truncated Poisson model with clutch size (1–4) as a count response variable. In both models, we fitted molt date, age, tarsus length, territory quality, and group size as fixed effects. Additionally, to determine whether the timing of molt completion predicted the likelihood of experiencing divorce, we built a GLMM with the occurrence of divorce within a year (experienced divorce within the year = 1, did not = 0) as a binomial response variable, and molt date, tarsus length, territory quality and group size as fixed effects. The variance in molt date did not differ between males that divorced and those that did not (Levene's test for equality of variances, $F_{1,179} = 0.002$, $P = 0.97$). Furthermore, for each record of divorce for which the molt dates of both the previous partner and the new one were known ($n = 18$), we tested whether the new partner completed his molt significantly earlier than the previous partner using Welch's *t*-test.

To test whether molt date predicted success of breeding attempts—1) hatching success, 2) fledging success, and 3) fledgling survival, we built a GLMMPQL and used the "cbind" function to create a binomial response variable, binding together respectively 1) the number of hatched eggs and the number of unhatched eggs, 2) the number of fledged nestlings and the number of unfledged nestlings, and 3) the number of fledglings that survived for at least 6 weeks and the number of fledglings that did not. Molt date, age, tarsus length, territory quality, and group size were fitted as fixed effects. For the fledging success model, we excluded clutches in which all eggs did not hatch. For the fledgling survival model, we excluded broods with no fledgling. We also tested whether molt date predicted annual reproductive success, assessed as the number of 6-week-old fledglings produced. We built a zero-inflated Poisson model with annual reproductive success as a count response variable. The fixed and random effects listed above were fitted in the

count part of the model and an intercept was fitted for the binomial part of the model (i.e. the zero-inflation did not depend on any explanatory variable).

Extra-pair paternity (EPP) and cuckoldry rates. To determine whether the timing of molt completion predicted the likelihood of obtaining EPP among dominants (as it does in other *Malurus*; Dunn and Cockburn 1999; Brouwer et al. 2011), we built a GLMM with whether a male obtained EPP within a year (obtained EPP within the year = 1, did not = 0) as a binomial response variable, and fitted the same fixed and random effects as in the divorce model above (molt date relative to the dominant population for each year). Although similar results were obtained when including incestuous pairs, we excluded these (n = 4) as they have higher EPP for inbreeding avoidance (Kingma et al. 2013). The variance in molt date did not differ between males that obtained EPP and those that did not (Levene's test for equality of variances, $F_{1,166}$ = 0.65, P = 0.42).

To test whether the timing of molt completion predicted the likelihood of losing paternity among dominants we used a similar GLMM as above, but with the occurrence of paternity loss within a year (lost paternity within the year = 1, did not = 0) instead as a binomial response variable.

Annual survival and remaining lifespan. To assess potential survival costs of early molt (hypothesized to be important in superb fairy-wrens; Peters et al. 2000; Cockburn et al. 2008), we related molt timing to both annual survival and remaining lifespan. Because both response variables could be affected by age, we first restricted the analyses described below to birds whose age was accurately known and showed that neither variable significantly differed between each age class (Supplementary Tables S8 and S9; for details see Supplementary Appendix S2 (B)). Therefore, we expanded these analyses to the age classes "1" and "2+" instead.

To test whether the timing of molt completion correlated with individual annual survival, we built a GLMM with annual survival (survived until the end of the year = 1, died before the end of the year = 0) as a binomial response variable, and molt date (relative to the whole population for each year), age (levels "1" and "2+"), social status, tarsus length, territory quality, and group size as fixed effects. Bird identity, territory identity, and year were included as random intercepts. Birds for which the year of death could not be unambiguously established (i.e. death date estimates with error >183 days or when the error interval included 01 July of a given year) were excluded (n = 5 of 205).

To test whether the timing of molt completion correlated with the remaining lifespan among birds that had died by the end of the study, we built a negative binomial model with remaining lifespan (in months) as a count response variable, and fitted the exact same fixed and random effects as in the annual survival model above. Remaining lifespan for a given year was calculated as the number of months between the molt date in the considered year and the death date. Birds whose death date estimate had an error >15 days were excluded (n = 85 of 205).

All analyses were done in R 3.2.0 (R Core Team 2015). LMMs were built using the packages "lme4" (Bates et al. 2011) and "lmerTest" (Kuznetsova et al. 2015; for details see Supplementary Appendix S2 (C)). GLMMs were first fitted without random term to estimate dispersion. If data were not over- or underdispersed, GLMMs were built using "lme4"; otherwise we either built 1) GLMMPQLs using the package "MASS" for GLMMs with binomial response variable, or 2) negative binomial or zero-inflated Poisson models using the package "glmmADMB" for overdispersed GLMMs with count data, or zero-truncated Poisson models using the same package for underdispersed GLMMs with count data (for details see Supplementary Appendix S2 (C)). All continuous variables were centered and scaled. Estimates of model coefficients (β) and their standard error (±SE) are presented.

RESULTS

Intrinsic effects on timing of pre-breeding molt

In total, 279 seasonal molt profiles with reliable information on the date of pre-breeding molt completion were obtained from 137 males over a 6-year period. At the population level, pre-breeding molt in male purple-crowned fairy-wrens was on average completed by early October, although the timing of completion was highly variable within and between individuals and within and between years—occurring from May at the earliest to February of the following year at the latest. Only 16% of males (4 of 25) developed a complete breeding plumage in their first year, whereas most (92%, 206 of 224) males in their second year or older reached a complete breeding plumage (all 3-year-old and older males developed more than 98% of the complete breeding plumage; Supplementary Figure S1).

Timing of pre-breeding molt was related to age. A quadratic relationship was observed between age (in months) and the pre-breeding molt completion date (Supplementary Table S1a), with males completing their molt earlier as they aged. This effect appeared to flatten out above 2 years of age, as was confirmed when comparing age categories (Supplementary Table S1b): 1-year-old males completed their molt significantly later than males in older age classes, but timing of molt completion did not differ between 2-year-old males and older males (Supplementary Table S2, Supplementary Figure S2). When using only 2 age classes, "1" and "2+", the strong effect of age on pre-breeding molt completion date remained (Table 1, Figure 2). Timing of pre-breeding

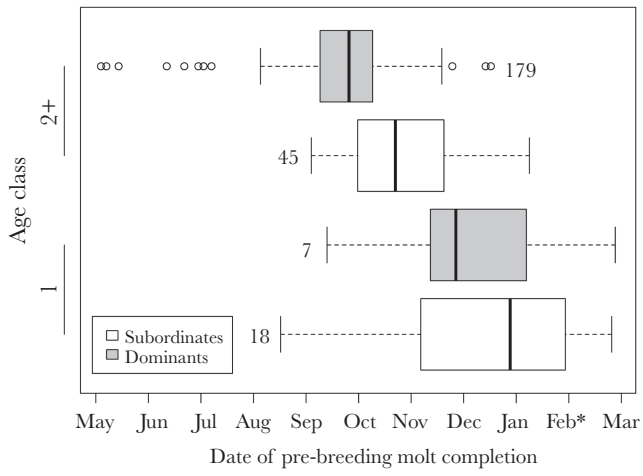


Figure 2
Older and dominant males produce the annual breeding plumage earlier. Shown are the effects of age and social status on the timing of pre-breeding molt completion in male purple-crowned fairy-wrens (n = 249). Boxes depict the 25th, 50th, and 75th percentiles, and whiskers the 10th and 90th percentiles. The sample size is indicated for each category. Month of peak breeding initiation is February indicated with an asterisk (*).

molt was strongly related to social status (Table 1), with subordinate males completing their molt significantly later than dominant males in both age classes (Figure 2). Based on the full model with significant fixed terms, adjusted repeatability of individual molt completion timing was 0.21 (95% confidence interval: 0.21–0.33; Table 1), indicating moderately consistent among-individual differences across the years. Individual tarsus length had no effect on the pre-breeding molt completion date (Table 1).

The comparison of within-individual changes in molt timing between subordinates who remained subordinate and subordinates who became dominants confirmed the significant advancement with age, but showed no difference due to the change of social status (Supplementary Table S3b).

Environmental effects on timing of pre-breeding molt

Higher cumulative rainfall in the past 12 months significantly increased the likelihood of males being fully purple (≥95%) by the end of November (Table 2, Figure 3). In addition to the effect of rainfall, the likelihood of being fully purple by the end of November was also age- and status-dependent (Table 2, Figure 3), being consistent with the age- and status-dependence of the timing of molt completion reported above. In these analyses, territory quality and group size had no effect on the pre-breeding molt completion date (Table 1).

Fitness benefits and costs of early molt

At the population level, the timing of pre-breeding molt completion in subordinates did not predict the likelihood of gaining a breeder position via dispersal (Supplementary Table S3a, Figure 4a). When we performed a case-by-case analysis, comparing the molt timing of a "winner" of a breeding vacancy and potential competitors located within the same distance to the vacancy, timing of molt completion did not predict success (Welch's *t*-test, $t = 0.05$, df = 35, $P = 0.48$, $n = 19$ pairwise comparisons, including 9 cases of earlier molt completion by the winner and 10 cases of later molt completion).

Over the course of the study, 162 individual molt profiles associated with a reproductive event were recorded in 70 breeder males.

Table 2

By the start of the peak breeding initiation, male breeding plumage is more complete in older and dominant males, and following wetter years

Parameter	β	SE	df	*t* value	*P*
Fixed effects					
Intercept	−5.72	0.64	383	−8.94	<0.001
Age†	**4.87**	**0.34**	**383**	**14.43**	**<0.001**
Status*	**2.99**	**0.32**	**383**	**9.19**	**<0.001**
Rainfall	**1.86×10^{-3}**	**0.58×10^{-3}**	**383**	**3.21**	**0.002**
Random effects					
$\sigma^2_{territory}$	0.77				
$\sigma^2_{individual\ in\ territory}$	2.12				
$\sigma^2_{residual}$	0.55				

Shown are results from a GLMMPQL examining the effects of age (levels 1, 2+), social status, and cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) on the likelihood of being fully purple by the end of November ($n = 778$ observations for 300 individuals over 11 years).
Significant values ($P < 0.05$) are in bold
Reference categories are †age 1, *subordinate.

The majority of males (71%) were displaying a complete breeding plumage when the first eggs were laid, while some males (23%) were still undergoing pre-breeding molt, and a few others (6%) were still in non-breeding plumage. No correlation between pre-breeding molt timing of males and timing of breeding initiation by the breeder female (date of first egg) was found (Supplementary Table S4), neither for dominants nor for subordinates.

Among dominants, timing of pre-breeding molt completion was not associated with any component of reproductive success. Specifically, female investment, including the annual number of breeding attempts (Supplementary Table S5, Figure 4b) and clutch size (Supplementary Table S5, Figure 4c) did not vary with molt timing of the breeding partner. Timing of male pre-breeding molt completion was not correlated with hatching success, fledging success and fledgling survival (Supplementary Table S6, Supplementary Figure S3). Finally, the annual number of fledglings produced that survived for at least 6 weeks was not predicted by molt timing (Supplementary Table S5, Figure 4d).

Dominant males who completed their molt earlier were not more likely to gain EPP (Supplementary Table S7, Figure 4e) and those that molted later were not more likely to lose paternity (Supplementary Table S7) or to be divorced (Supplementary Table S7, Figure 4f). In addition, divorcing females did not re-pair with a male that had molted earlier than their previous partner (Welch's *t*-test, $t = -1.33$, df = 33, $P = 0.90$, $n = 18$ pairwise comparisons, including 7 cases of earlier molt completion by the new partner, 10 cases of later molt completion, and 1 case of completion on the same day). Those males that were still undergoing pre-breeding molt or in non-breeding plumage when their partner initiated a clutch, sired 100% of those offspring (32 broods for males in partial breeding plumage, 4 broods for males in non-breeding plumage).

No association between annual survival and timing of pre-breeding molt was found (Supplementary Table S10, Figure 4g). Furthermore, no association between remaining lifespan and timing of pre-breeding molt was found among birds of known lifespan (Supplementary Table S10, Figure 4h).

DISCUSSION

Based on our findings, the timing of acquisition of the purple-and-black breeding plumage in male *M. coronatus* appears to show limited correlations with male quality and environmental conditions: its expression varies with both age and climate, but the age effect is restricted to early life, and the climate effect appears relatively small. In addition, dominants complete their pre-breeding molt earlier than subordinates, suggesting the potential to signal social status and competitive ability among males, although molt timing did not relate to the chances of subordinates to become dominant. Finally, although we investigated numerous reproductive benefits and costs potentially associated with early molt completion, we could not identify any, suggesting weak or even absent female choice acting on this trait in male purple-crowned fairy-wrens.

Pre-breeding molt timing as a signal of male quality? Effects of age and rainfall

The timing of breeding plumage acquisition in male *M. coronatus* is highly variable both between and within individuals. Sexually selected ornaments are often assumed to show high variability, maintained through high condition-dependence (Delhey and Peters 2017), although the latter has proved difficult to test (Cotton et al. 2004; Delhey and Peters 2008; Delhey et al. 2017). Substantial

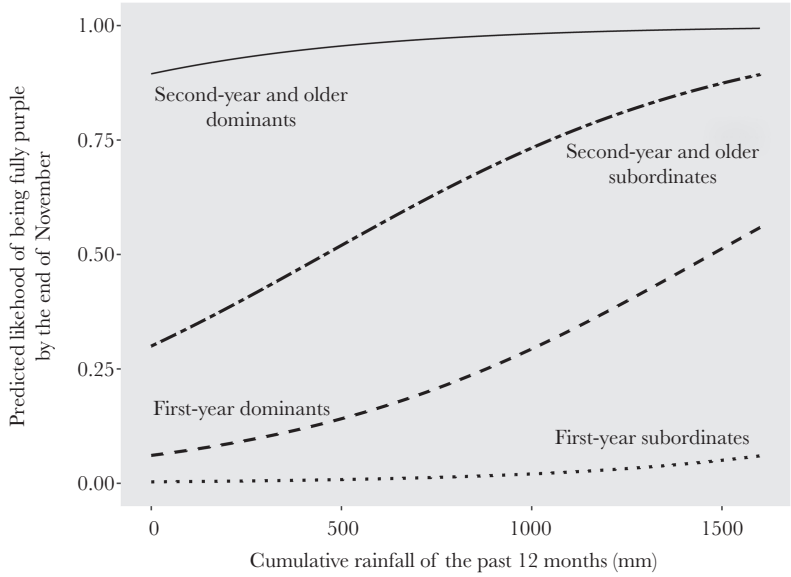


Figure 3
By the start of the peak breeding initiation, male breeding plumage is more complete in older and dominant males, and following wetter years. Shown are the predicted correlations between the likelihood of being fully purple by the end of November and the cumulative rainfall over the past 12 months (n = 778 observations) for 1-year-old subordinates (dotted line), 1-year-old dominants (dashed line), 2-year-old and older subordinates (dotted and dashed line), and 2-year-old and older dominants (solid line).

variation in molt timing (range of up to 10 months vs. 3–9 months in other studied fairy-wren species; Peters et al. 2013) therefore provides support for some potential to act as a male quality indicator, either for females in mate assessment or for males when gauging rivals. Moreover, pre-breeding molt timing is moderately repeatable (r = 0.21; Table 1), suggesting that it could serve as a reliable signal of male genetic quality (Lynch and Walsh 1998; Bell et al. 2009) and consequently be subject to female preferences (Andersson 1986; Iwasa and Pomiankowski 1994).

The timing of pre-breeding molt shows some degree of age-dependence as a significant advancement occurs within the first 2 years of life (Table 1, Figure 2). Males might benefit from age-dependent investment in their signals, reserving the production of costly traits until older ages (Kokko 1997) and honest signaling should favor delaying development of costly signals and therefore preferences for older males (Brooks and Kemp 2001; Proulx et al. 2002). Age-dependence of molt timing could reflect honest signaling of male quality in *M. coronatus*. However, the observed improvement is restricted to early life, which contrasts with patterns in male superb fairy-wrens that show life-long improvement, with the proportion of early molters increasing linearly with age (van de Pol et al. 2012). In superb fairy-wrens, early acquisition of breeding plumage, months prior to breeding, is critical for obtaining extra-pair fertilizations that dominate the mating system of this species (Dunn and Cockburn 1999). Similarly, male red-winged fairy-wrens acquire their breeding plumage earlier as they get older (Russell et al. 1991) and early molters are preferentially chosen as extra-pair mates by females (Brouwer et al. 2011). That male purple-crowned fairy-wrens do not show further improvement with age is in line with their much lower levels of EPP (Kingma et al. 2009).

The timing of pre-breeding molt in male *M. coronatus* is also sensitive to climate, similar to several other fairy-wrens: higher cumulative rainfall correlates with higher likelihood of displaying a complete breeding plumage by the end of November (Table 2, Figure 3). This indicates that males complete their molt earlier following wetter years, and this effect appears particularly noticeable in years with extreme climate, especially during drought (e.g. 2005). Molting is energetically and nutritionally demanding in small birds (Payne 1972; Lindström et al. 1993; Murphy 1996; Cornelius et al. 2011). Higher rainfall is likely to result in prolonged vegetation growth and arthropod abundance (Lowman 1982), therefore providing more resources to these insectivorous birds and increasing their ability to undertake their pre-breeding molt early. Such dependence on environmental variability could be interpreted as evidence for condition-dependence. In comparison, in superb fairy-wrens, molt timing shows a strong sensitivity to summer rainfall, and more particularly old males can strongly advance their molt after high summer rainfall (Cockburn et al. 2008; van de Pol et al. 2012).

Both annual survival (reflecting short-term survival) and remaining lifespan (reflecting long-term survival) show no correlation with the timing of pre-breeding molt (Supplementary Table S10, Figures 4g and h). This might suggest that early molt does not entail substantial costs. Cockburn et al. (2008) hypothesized that cold conditions during winter might have detrimental effects on the survival of early molting superb fairy-wrens in temperate SE Australia. In comparison, individuals from our study population experience much milder tropical climatic conditions where temperatures rarely go below 4 °C (vs. −8 °C in Cockburn et al. 2008), which may account for the lack of noticeable survival costs in early molters. However, without experimental data it is hard to interpret this absence of correlation, since only males able to undergo the pre-breeding molt early (and that can absorb the potential costs) might do so (Peters 2000), masking potential survival costs associated with early molt. In addition, it is possible that the selective disappearance of certain phenotypes makes some patterns of survival selection difficult to detect, especially based on the data that were available to us.

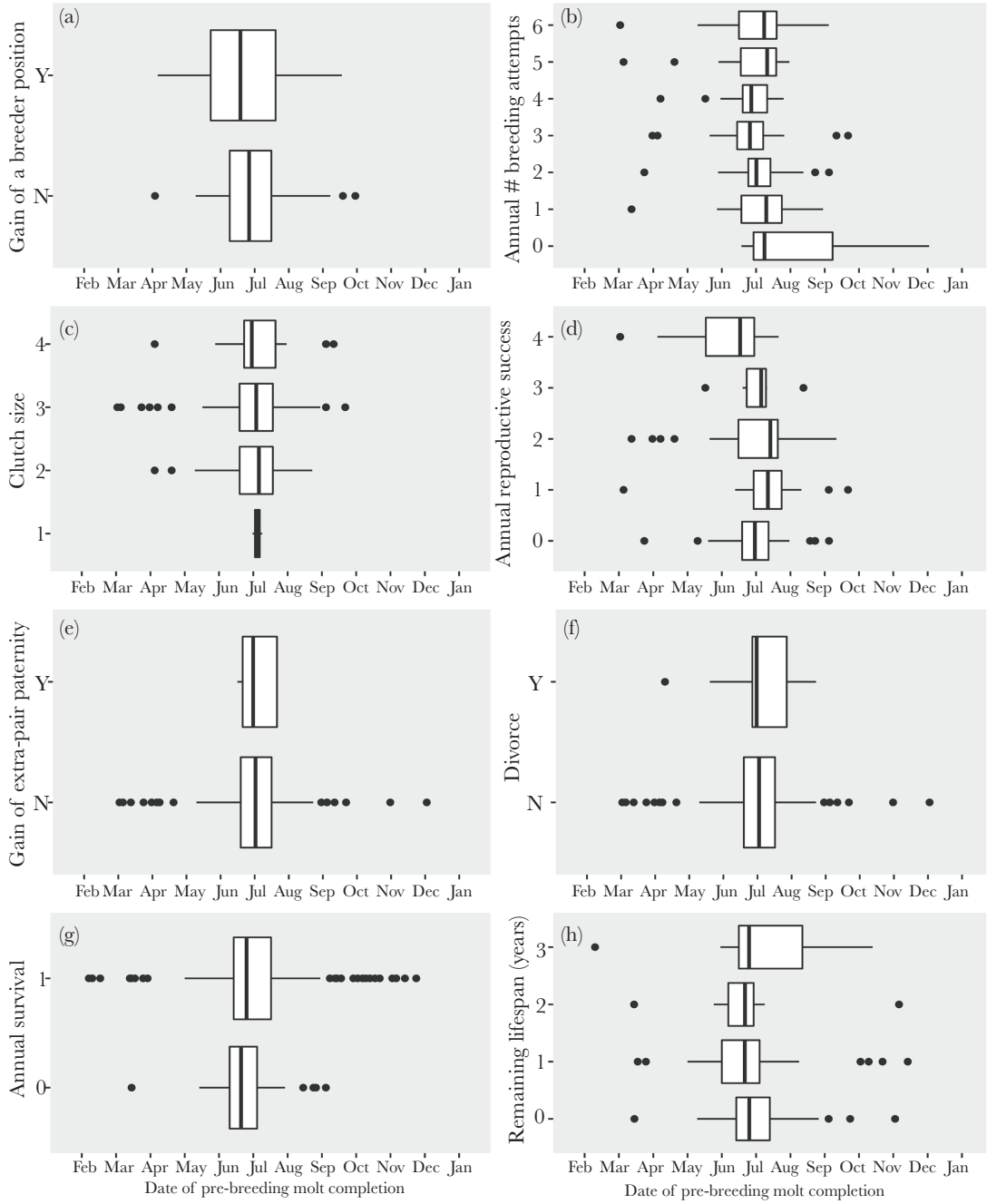


Figure 4

Molt performance is not related to fitness indicators. Shown is the absence of correlation between the timing of pre-breeding molt completion and (a) the likelihood of gaining a breeder position (n = 47), (b) the annual number of breeding attempts (6 and above combined into one class; n = 166), (c) clutch size (n = 259), (d) annual reproductive success (i.e. number of 6-week-old fledglings produced; 4 and above combined into one class; n = 151), (e) the likelihood of gaining EPP (n = 168), (f) the likelihood of experiencing divorce (n = 181), (g) annual survival (n = 242), and (h) remaining lifespan (3 years and above combined into one class; n = 120).

Early molt: no fitness benefits through female choice

When investigating an extensive panel of male fitness proxies linked to female choice, we consistently found no evidence of any association with pre-breeding molt timing. Although it is possible that we missed some fitness indicators that are not immediately obvious or easily measured, our data quite strongly suggest weak or absent intersexual selection favoring early molt in male purple-crowned fairy-wrens.

One potential sexually selected benefit of preferred, more ornamented males, is to pair with females of higher quality, those that start breeding earlier or lay more eggs (McGraw et al. 2001; Sheldon et al. 2003; Garant et al. 2007). We found no evidence for such a process in *M. coronatus*: females paired to an earlier-molting male do not start egg-laying earlier, increase the number

of breeding attempts or lay larger clutches (Supplementary Tables S4 and S5, Figures 4b and c). In agreement with this, our findings did not support any association between molt timing and reproductive success. Early molting males do not achieve greater success of breeding attempts (hatching success, fledging success, fledgling survival) or greater annual reproductive success (Supplementary Tables S5 and S6, Figure 4d and Supplementary Figure S3).

There were no indications that females preferred early molting males as social or EP mates. There was no link between male molt timing and divorce (Supplementary Table S7, Figure 4f). Most divorces are initiated by females that leave their partner to take a breeding position elsewhere (Hidalgo Aranzamendi et al. 2016). This provides an opportunity for females to choose their partner, but pre-breeding molt timing appears not to play a role in such partner assessment by prospecting females. Additionally, late molters do not suffer increased paternity loss and early molters do not increase their chances of gaining EPP (based on the few recorded cases) (Supplementary Table S7, Figure 4e), in contrast to 2 other *Malurus* (Dunn and Cockburn 1999; Brouwer et al. 2011). This is in agreement with the absence in *M. coronatus* of male adaptations that play a central role in the EPP-based mating system of other Australian fairy-wrens (Figure 5; Kingma et al. 2009).

Early molt: a signal of social status or competitive ability?

Independently of the effect of age, the timing of pre-breeding molt in male *M. coronatus* varies according to social status: the dominant male in the group is generally the first to acquire breeding plumage (Table 1, Figure 2). It is possible that, by delaying their molt compared to dominant males, subordinate males attempt to avoid costly aggressive interactions (Senar 2006; Santos et al. 2011). Alternatively, molt timing might serve as a signal of dominance and competitiveness among males. However, subordinates that molt earlier do not increase their chances to acquire a breeder position (Supplementary Table S3a). Because subordinates do not participate in reproduction (aside from 2 cases of a subordinate male gaining EPP; Kingma et al. 2009), gaining a breeding position is a critical determinant of male success. This can occur by splitting the natal territory, by inheritance (through an orderly, age-based queue) or by dispersal (Kingma 2011; Kingma et al. 2011b; Hidalgo Aranzamendi et al. 2016). In this third case, because breeding vacancies are scarce and male dispersal mostly limited to neighboring territories, the opportunity to compete is rare, which may account for the absence of role of molt timing in the acquisition of such a vacancy.

Timing of pre-breeding molt: the ghost of female EP mate choice past?

We found no association between pre-breeding molt timing and any of the numerous proxies for male fitness we used in this study, suggesting that this trait is currently not under intersexual selection in male purple-crowned fairy-wrens. However, substantial variation exists, partially linked to age and environmental conditions, and it has a potential genetic component (inferred through its repeatability), attributes which are common to sexually selected traits (Cotton et al. 2004; Delhey and Peters 2008). Therefore, the early development of the breeding plumage could be used as a sexual signal, similar to how it functions in other fairy-wrens. Moreover, early molt may function as a social signal of dominance and competitiveness among males, although it is not intuitively clear how this could

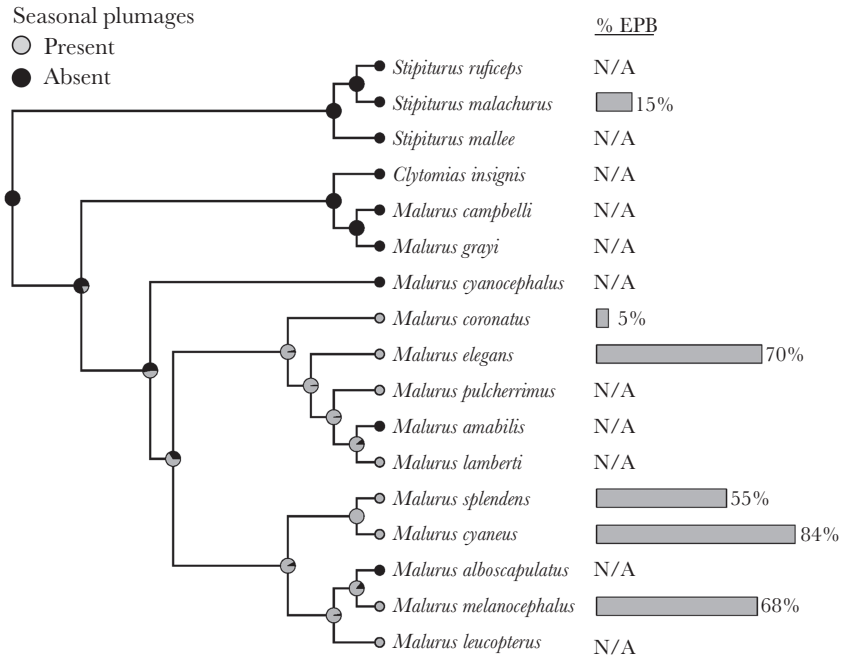


Figure 5
Evolution of seasonal plumages in fairy-wrens and emu-wrens. Shown is the ancestral state reconstruction of changes in seasonal plumages for 17 species of Malurid, using stochastic mapping and based on the supermatrix phylogeny of Marki et al. (2017). Levels of extra-pair paternity expressed as the proportion of broods containing at least one nestling resulting from an extra-pair fertilization (% EPB) are shown for each species (Rowe and Pruett-Jones 2013; N/A: no data available). Reconstruction of ancestral presence/absence of seasonal plumages was done using stochastic mapping using the function "make.simmap" from the R package "phytools" (Revell 2012). We ran 100 stochastic mappings and pie charts in nodes summarize the posterior probability of being one state or the other. For this analysis, we set the root stage as absence of seasonal plumages and we allowed different transition rates between states (model = "ARD").

work, or could be tested, since males are territorial year-round, whether they are in breeding plumage or not. This raises the question of what is the function of this variation in the timing of breeding plumage acquisition in male purple-crowned fairy-wrens.

Ancestral state reconstruction (Figure 5; based on the recent supermatrix phylogeny by Marki et al. 2017) suggests that variation in pre-breeding molt in male purple-crowned fairy-wrens represents a vestigial trait, on its way to disappearing. Kingma et al. (2009) suggested that *M. coronatus* is derived from an ancestor with high EPP, but subsequently lost the extreme levels of EPP and associated behavioral adaptations that characterize the other *Malurus* (i.e. extra-territorial intrusions, extra-pair courtship displays). According to such a scenario, the highly variable timing of pre-breeding molt is a formerly sexually selected trait that lost its function as sexual selection became weaker ("relaxed" sexual selection; Lahti et al. 2009). This is in agreement with the reduction in the extent of the body covered by breeding plumage in purple-crowned fairy-wrens compared to other species in the genus (Kingma et al. 2009). We hypothesize that the loss of extreme levels of EPP removed the main evolutionary force driving variation in timing of seasonal acquisition of bright male plumage in fairy-wrens, resulting in the loss of its main function. Therefore, this trait could represent an evolutionary vestige that may undergo reduction or even complete loss, unless it acquires a novel function, or is correlated with other traits currently under selection (Lahti et al. 2009). Although traits can persist and continue to be expressed for long periods of time after a source of selection is removed (Lahti et al. 2009), variation in molt timing in male *M. coronatus* appears not to have undergone noticeable reduction, which hints that it may still be under some selection pressures, perhaps difficult to detect if it is only noticeable in specific conditions, or serves other purposes. In particular, molt timing may function as a social signal of dominance and competitiveness among males, which could explain why variation in this trait persists.

CONCLUSION

We found no evidence of sexual selection currently acting on the timing of breeding plumage acquisition, based on the investigation of the male fitness variables presented in this study. However, its strong status-dependence warrants future studies to focus on a potential role in mediating male-male social interactions. Additionally, although the timing of pre-breeding molt is not related to fitness indicators for male purple-crowned fairy-wrens, further work is needed to assess the importance of other signal components, such as variation in the quality of the purple color which represents a unique and highly conspicuous evolutionary innovation in this species (Delhey et al. 2013).

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Behavioral Ecology* online.

FUNDING

This work was supported by a Minerva Fellowship of the Max Planck Society (to A.P.); the Australian Research Council (FT 110100505 to A.P. and DE 120102323 to K.D.) and the Holsworth Wildlife Research Endowment – Equity Trustees Charitable Foundation (to M.F.).

Research was approved by the Australian Bird and Bat Banding Scheme (ABBBS license 2230 to A.P.); the Western Australian Department of Conservation and Land Management / Environment and Conservation / Parks and Wildlife; the Australian Wildlife Conservancy; and the ethics committees of the School of Biological Sciences at Monash University (BSCI/2011/28 and BSCI/2015/11) and the Max Planck Institute for Ornithology, in accordance with national and international legislation for animal use in research. Our thanks to numerous volunteers during fieldwork and the support of staff at Australian Wildlife Conservancy's Mornington Wildlife Sanctuary.

Data accessibility: Analyses reported in this article can be reproduced using the data provided by Fan et al. (2017). The data are embargoed for a 12-month period as they are currently being used by the authors in other analyses, and will be made available in the repository at the end of this period.

Handling editor: Francisco Garcia-Gonzalez

REFERENCES

Akçay E, Roughgarden J. 2007. Extra-pair paternity in birds: review of the genetic benefits. Evol Ecol Res. 9:855.

Andersson M. 1986. Evolution of condition-dependent sex ornaments and mating preferences: sexual selection based on viability differences. Evolution. 40:804–816.

Andersson M. 1994. Sexual selection. Princeton (NJ): Princeton University Press.

Andersson M. 2006. Condition-dependent indicators in sexual selection: development of theory and tests. In: Lucas JR, Simmons LW, editors. Essays in Animal Behaviour: Celebrating 50 years of Animal Behaviour. London (UK): Elsevier Academic Press. p. 255–269.

Andersson M, Simmons LW. 2006. Sexual selection and mate choice. Trends Ecol Evol. 21:296–302.

Barnard P. 1995. Timing of ornament growth, phenotypic variation, and size dimorphism in two promiscuous African whydahs (Ploceidae: Vidua). Biol J Linnean Soc. 55:129–141.

Basolo AL. 1996. The phylogenetic distribution of a female preference. Syst Biol. 45:290–307.

Bates D, Maechler M, Bolker B. 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375–42 [cited 2017 April 19]. Available from: https://cran.r-project.org/web/packages/lme4/index.html.

Bell AM, Hankison SJ, Laskowski KL. 2009. The repeatability of behaviour: a meta-analysis. Anim Behav. 77:771–783.

Brooks R, Kemp DJ. 2001. Can older males deliver the good genes? Trends Ecol Evol. 16:308–313.

Brouwer L, van de Pol M, Atema ELS, Cockburn A. 2011. Strategic promiscuity helps avoid inbreeding at multiple levels in a cooperative breeder where both sexes are philopatric. Mol Ecol. 20:4796–4807.

Burns KJ. 1998. A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (Thraupidae): the role of female versus male plumage. Evolution. 52:1219–1224.

Cockburn A, Osmond HL, Double MC. 2008. Swingin' in the rain: condition dependence and sexual selection in a capricious world. Proc R Soc B. 275:605–612.

Cornelius JM, Perfito N, Zann R, Breuner CW, Hahn TP. 2011. Physiological trade-offs in self-maintenance: plumage molt and stress physiology in birds. J Exp Biol. 214:2768–2777.

Cotton S, Fowler K, Pomiankowski A. 2004. Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? Proc R Soc B. 271:771–783.

Darwin C. 1871. The descent of man and selection in relation to sex. London (UK): John Murray.

Delhey K, Peters A. 2008. Quantifying variability of avian colours: are signalling traits more variable? PLoS One. 3:e1689.

Delhey K, Peters A. 2017. The effect of colour producing mechanisms on plumage sexual dichromatism in passerines and parrots. Funct Ecol. 31:903–914.

Delhey K, Hall M, Kingma SA, Peters A. 2013. Increased conspicuousness can explain the match between visual sensitivities and blue plumage colours in fairy-wrens. Proc R Soc B. 284:20121771.

Delhey K, Szecsenyi B, Nakagawa S, Peters A. 2017. Conspicuous plumage colours are highly variable. Proc R Soc B. 284:20162593.
Dunn PO, Cockburn A. 1999. Extrapair mate choice and honest signaling in cooperatively breeding superb fairy-wrens. Evolution. 53:938–946.
Dunn PO, Whittingham LA, Pitcher TE. 2001. Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. Evolution. 55:161–175.
Ellers J, Toby Kiers E, Currie CR, McDonald BR, Visser B. 2012. Ecological interactions drive evolutionary loss of traits. Ecol Lett. 15:1071–1082.
Emerson SB. 1996. Phylogenies and physiological processes—the evolution of sexual dimorphism in Southeast Asian frogs. Syst Biol. 45:278–289.
Emlen DJ, Marangelo J, Ball B, Cunningham CW. 2005. Diversity in the weapons of sexual selection: horn evolution in the beetle genus *Onthophagus* (Coleoptera: Scarabaeidae). Evolution. 59:1060–1084.
Fan M, Hall ML, Kingma SA, Mandeltort LM, Hidalgo Aranzamendi N, Delhey K, Peters A. 2017. Data from: no fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy? Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.pn974.
Forstmeier W, Nakagawa S, Griffith SC, Kempenaers B. 2014. Female extra-pair mating: adaptation or genetic constraint? Trends Ecol Evol. 29:456–464.
Garant D, Kruuk LE, McCleery RH, Sheldon BC. 2007. The effects of environmental heterogeneity on multivariate selection on reproductive traits in female great tits. Evolution. 61:1546–1559.
Griffith SC, Owens IP, Thuman KA. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol Ecol. 11:2195–2212.
Hall ML, Peters A. 2008. Coordination between the sexes for territorial defence in a duetting fairy-wren. Anim Behav. 76:65–73.
Hall ML, Peters A. 2009. Do male paternity guards ensure female fidelity in a duetting fairy-wren? Behav Ecol. 20:222–228.
Hall ML, Kingma SA, Peters A. 2013. Male songbird indicates body size with low-pitched advertising songs. PLoS One. 8:e56717.
Hamilton WD, Zuk M. 1982. Heritable true fitness and bright birds: a role for parasites? Science. 218:384–387.
Hidalgo Aranzamendi N, Hall ML, Kingma SA, Sunnucks P, Peters A. 2016. Incest avoidance, extrapair paternity, and territory quality drive divorce in a year-round territorial bird. Behav Ecol. 27:1808–1819.
Hill GE. 2002. A red bird in a brown bag: the function and evolution of colorful plumage in the house finch. Oxford (UK): Oxford University Press.
Irwin RE. 1994. The evolution of plumage dichromatism in the New World blackbirds: social selection on female brightness. Am Nat. 144:890–907.
Iwasa Y, Pomiankowski A. 1994. The evolution of mate preferences for multiple sexual ornaments. Evolution. 48:853–867.
Kingma SA, Hall ML, Segelbacher G, Peters A. 2009. Radical loss of an extreme extra-pair mating system. BMC Ecology. 9:15.
Kingma SA, Hall ML, Arriero E, Peters A. 2010. Multiple benefits of cooperative breeding in purple-crowned fairy-wrens: A consequence of fidelity? J Anim Ecol. 79:757–768.
Kingma SA. 2011. The evolution of the social and genetic mating system of purple-crowned fairy-wrens *Malurus coronatus* [Doctoral dissertation]. [Konstanz (Germany)]: University of Konstanz.
Kingma SA, Hall ML, Peters A. 2011a. No evidence for offspring sex-ratio adjustment to social or environmental conditions in cooperatively breeding purple-crowned fairy-wrens. Behav Ecol Sociobiol. 65:1203–1213.
Kingma SA, Hall ML, Peters A. 2011b. Multiple benefits drive helping behavior in a cooperatively breeding bird: an integrated analysis. Am Nat. 177:486–495.
Kingma SA, Hall ML, Peters A. 2013. Breeding synchronization facilitates extrapair mating for inbreeding avoidance. Behav Ecol. 24:1390–1397.
Kokko H. 1997. Evolutionarily stable strategies of age-dependent sexual advertisement. Behav Ecol Sociobiol. 41:99–107.
de Kort SR, ten Cate C. 2004. Repeated decrease in vocal repertoire size in *Streptopelia* doves. Anim Behav. 67:549–557.
Kuznetsova A, Brockhoff PB, Christensen RH. 2015. Package 'lmerTest'. R package version. 2015:2-0 [cited 2017 April 19]. Available from: https://cran.r-project.org/web/packages/lmerTest/index.html.
Lahti DC, Johnson NA, Ajie BC, Otto SP, Hendry AP, Blumstein DT, Coss RG, Donohue K, Foster SA. 2009. Relaxed selection in the wild. Trends Ecol Evol. 24:487–496.
Lindström Å, Visser GH, Daan S. 1993. The energetic cost of feather synthesis is proportional to basal metabolic rate. Physiol Zool. 66:490–510.
Lozano GA. 1994. Carotenoids, parasites, and sexual selection. Oikos. 70:309–311.
Lowman MD. 1982. Seasonal variation in insect abundance among three Australian rain forests, with particular reference to phytophagous types. Aust J Ecol. 7:353–361.
Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Sunderland (MA): Sinauer.
Marki PZ, Jønsson KA, Irestedt M, Nguyen JMT, Rahbek C, Fjeldså J. 2017. Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). Mol Phylogenet Evol. 107:516–529.
McGraw KJ, Stoehr AM, Nolan PM, Hill GE. 2001. Plumage redness predicts breeding onset and reproductive success in the house finch: a validation of Darwin's theory. J Avian Biol. 32:90–94.
McGraw KJ, Mackillop EA, Dale J, Hauber ME. 2002. Different colors reveal different information: how nutritional stress affects the expression of melanin-and structurally based ornamental plumage. J Exp Biol. 205:3747–3755.
Møller AP, Birkhead TR. 1994. The evolution of plumage brightness in birds is related to extrapair paternity. Evolution. 48:1089–1100.
Murphy ME. 1996. Energetics and nutrition of molt. In: Carey C, editor. Avian energetics and nutritional ecology. New York (NY): Springer US. p. 158–198.
Nakagawa S, Schielzeth H. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev. 85:935–956.
Ödeen A, Björklund M. 2003. Dynamics in the evolution of sexual traits: losses and gains, radiation and convergence in yellow wagtails (*Motacilla flava*). Mol Ecol. 12:2113–2130.
Omland KE. 1996. Female Mallard mating preferences for multiple male ornaments. I. Natural variation. Behav Ecol Sociobiol. 39:353–360.
Omland KE. 1997. Examining two standard assumptions of ancestral reconstructions: repeated loss of dichromatism in dabbling ducks (Anatini). Evolution. 51:1636–1646.
Omland KE, Lanyon SM. 2000. Reconstructing plumage evolution in orioles (Icterus): repeated convergence and reversal in patterns. Evolution. 54:2119–2133.
Ord TJ, Stuart-Fox D. 2006. Ornament evolution in dragon lizards: multiple gains and widespread losses reveal a complex history of evolutionary change. J Evol Biol. 19:797–808.
Owens IPF, Hartley IR. 1998. Sexual dimorphism in birds: why are there so many different forms of dimorphism? Proc R Soc B. 265:397–407.
Payne RB. 1972. Mechanisms and control of molt. Avian Biol. 2:103–155.
Peters A. 2000. Testosterone treatment is immunosuppressive in superb fairy-wrens, yet free-living males with high testosterone are more immunocompetent. Proc R Soc B. 267:883–889.
Peters A, Astheimer LB, Boland CR, Cockburn A. 2000. Testosterone is involved in acquisition and maintenance of sexually selected male plumage in superb fairy-wrens, *Malurus cyaneus*. Behav Ecol Sociobiol. 47:438–445.
Peters A, Kingma SA, Delhey K. 2013. Seasonal male plumage as a multicomponent sexual signal: Insights and opportunities. Emu. 113:232–247.
Porter ML, Crandall KA. 2003. Lost along the way: the significance of evolution in reverse. Trends Ecol Evol. 18:541–547.
Proulx SR, Day T, Rowe L. 2002. Older males signal more reliably. Proc R Soc B. 269:2291–2299.
Quinn VS, Hews DK. 2000. Signals and behavioural responses are not coupled in males: aggression affected by replacement of an evolutionarily lost colour signal. Proc R Soc B. 267:755–758.
R Core Team. 2015. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing [cited 2017 April 19]. Available from: https://www.R-project.org/.
Revell LJ. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 3:217–223.
Rowe M, Pruett-Jones S. 2013. Extra-pair paternity, sperm competition and their evolutionary consequences in the Maluridae. Emu. 113:218–231.
Rowley I, Russell E. 1993. The purple-crowned fairy-wren *Malurus coronatus*. II. Breeding biology, social organization, demography and management. Emu. 93:235–250.
Rowley I, Russell E. 1997. Fairy-wrens and Grasswrens. Oxford (UK): Oxford University Press.
Russell EM, Rowley I, Brown RJ, Brown MN. 1991. Acquisition of nuptial plumage in the Red-winged Fairy-wren *Malurus elegans*. Corella. 15:125–133.

Santos ES, Scheck D, Nakagawa S. 2011. Dominance and plumage traits: meta-analysis and metaregression analysis. Anim Behav. 82:3–19.
Schroeder J, Lourenço PM, Hooijmeijer JC, Both C, Piersma T. 2009. A possible case of contemporary selection leading to a decrease in sexual plumage dimorphism in a grassland-breeding shorebird. Behav Ecol. 20:797–807.
Senar J. 2006. Color displays as intrasexual signals of aggression and dominance. In: Hill GE, McGraw KJ, editors. Bird Coloration. Vol. II. Cambridge (MA): Harvard University Press. p. 87–136.
Sheldon BC, Kruuk LE, Merila J. 2003. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. Evolution. 57:406–420.
van de Pol M, Osmond HL, Cockburn A. 2012. Fluctuations in population composition dampen the impact of phenotypic plasticity on trait dynamics in superb fairy-wrens. J Anim Ecol. 81:411–422.
Wiens JJ. 1999. Phylogenetic evidence for multiple losses of a sexually selected character in phrynosomatid lizards. Proc R Soc B. 266:1529–1535.
Wiens JJ. 2001. Widespread loss of sexually selected traits: how the peacock lost its spots. Trends Ecol Evol. 16:517–523.
Wong BB, Rosenthal GG. 2006. Female disdain for swords in a swordtail fish. Am Nat. 167:136–140.
Zahavi A. 1975. Mate selection—a selection for a handicap. J Theor Biol. 53:205–214.

Supplementary Material

Appendix S1: Death date estimation

Between 2005 and 2010 – Birds were declared dead on the basis of failure to sight them in regular intensive surveys, and death dates could be assigned with a maximal error of 19.5 days.

Between 2011 and 2015 – As two censuses were conducted every year, the death date was determined as the date halfway between (i) the last date an individual was sighted during a census period and (ii) the first date it was not sighted during a subsequent census period when the territory where it was last seen was surveyed again (the death being subsequently confirmed as the individual was not sighted anywhere else). Half the duration between these two dates constituted the error of the death date estimate (mean=131.9±15.1 days, range: 3-354.5 days).

Appendix S2: Statistical analyses

(A) Early effect of age on timing of pre-breeding molt

We built a linear mixed model (LMM) with molt date as a response variable, age, social status, tarsus length, territory quality and group size as fixed effects, and bird identity, territory identity and year as random intercepts. Only birds whose age was accurately known were included in the model (n=130), thus excluding adult birds born prior to the start of the study (founder population). We first included (i) age in months both as a continuous linear and quadratic variable, and then (ii) age in years as a fixed factor (levels 1, 2, 3, 4 and 5; there were only two six-year-old individuals who were excluded from the analysis). Using Tukey's Honest Significant Difference (HSD) test, we performed post hoc comparisons between each

age class, which indicated that the molt date only significantly varied between the first and second years of life (Tables 1 and S1).

(B) Absence of age effect on both annual survival and remaining lifespan

To test whether the timing of molt completion correlated with individual annual survival, we built a generalized linear mixed model with penalized quasi-likelihood (GLMMPQL) with annual survival (survived until the end of the year = 1, died before the end of the year = 0) as a binomial response variable, and molt date (relative to the whole population for each year), age, social status, tarsus length, territory quality and group size as fixed effects. Bird identity, territory identity and year were included as random intercepts. Similarly, we built a negative binomial model with remaining lifespan (in months) as a count response variable, and fitted the exact same fixed and random effects as in the annual survival model above. Both models were restricted to birds whose age was accurately known and age was included in years as a fixed factor (levels 1, 2, 3, 4 and 5). In addition, birds for which the year of death could not be unambiguously established (i.e. death date estimates with error>183 days or when the error interval included July 1st of a given year) were excluded from the annual survival analysis, whereas birds whose death date estimate had an error>15 days were excluded from the remaining lifespan analysis. Tukey's HSD test indicated that neither variable significantly differed between each age class (Tables S8 and S9). Including age in years both as a continuous linear and quadratic variable confirmed the absence of age effect (not shown).

(C) Further details on statistical models

LMMs – Significance of fixed terms was assessed by *P*-values obtained using Satterthwaite approximations of degrees of freedom.

Overdispersed GLMMs with count data – If overdispersion was due to an excess of zeros, zero-inflated poisson models were built; otherwise negative binomial models were used.

Underdispersed GLMMs with count data – Zero-truncated poisson models were built after preliminarily identifying the absence of zeros as the source of underdispersion.

Table S1. Males advance the production of their annual breeding plumage within the first two years of life, but not beyond. Shown are results from LMMs examining the effects of age, social status, tarsus length, territory quality and group size on the date of pre-breeding molt completion in individuals of known age only, considering (a) age in months as a covariate (n=128) and (b) age in years as a fixed factor (levels 1, 2, 3, 4, 5; n=24, 41, 32, 18, 10, respectively).

(a)

Parameter	β	SE	df	t value	P
Fixed effects					
Intercept	106.26	11.87	9.10	8.96	<0.001
Age	**-1.59**	**0.32**	**100.48**	**-4.99**	**<0.001**
Age^2	**0.07**	**0.01**	**103.47**	**4.86**	**<0.001**
Status*	-11.65	9.21	105.62	-1.27	0.21
Tarsus length	-5.26	4.67	90.37	-1.13	0.26
Territory quality	0.33	0.87	22.78	0.38	0.71
Group size	3.33	2.08	118.29	1.60	0.11
Random effects					
$\sigma^2_{individual}$	6.39×10^{-10}				
$\sigma^2_{territory}$	413.26				
σ^2_{year}	411.70				
$\sigma^2_{residual}$	822.48				

*The reference category is subordinate.

(b)

Parameter	β	SE	df	t value	P
Fixed effects					
Intercept	161.28	10.95	12.78	14.72	<0.001
Age†					
β_{age2}	**-45.83**	**8.15**	**99.78**	**-5.63**	**<0.001**

β_{age3}	**-60.49**	**9.38**	**95.12**	**-6.45**	**<0.001**
β_{age4}	**-67.31**	**11.66**	**101.08**	**-5.77**	**<0.001**
β_{age5}	**-64.66**	**14.08**	**103.58**	**-4.59**	**<0.001**
Status*	-9.98	8.83	107.67	-1.13	0.26
Tarsus length	-0.72	4.56	95.62	-0.16	0.88
Territory quality	0.52	0.86	21.40	0.60	0.56
Group size	2.85	1.99	111.27	1.44	0.15
Random effects					
$\sigma^2_{individual}$	0.00				
$\sigma^2_{territory}$	428.57				
σ^2_{year}	304.05				
$\sigma^2_{residual}$	714.91				

Reference categories are as follows: †age 1, *subordinate.

Table S2. First-year males produce their annual breeding plumage later than males in older age classes. Shown are results from Tukey's Honest Significant Difference (HSD) test for post hoc comparisons between each age class (1-5 years old).

Pairwise comparison (n=125)	β	SE	z value	P
2-1	**-45.83**	**8.15**	**-5.63**	**<0.001**
3-1	**-60.49**	**9.38**	**-6.45**	**<0.001**
4-1	**-67.31**	**11.66**	**-5.77**	**<0.001**
5-1	**-64.66**	**14.08**	**-4.59**	**<0.001**
3-2	-14.66	7.14	-2.05	0.23
4-2	-21.48	9.62	-2.23	0.16
5-2	-18.83	12.31	-1.53	0.53
4-3	-6.82	8.84	-0.77	0.93
5-3	-4.17	11.50	-0.36	1.00
5-4	2.65	11.28	0.24	1.00

Table S3. Subordinate males that molt earlier do not increase their chances to gain a breeder position via dispersal and social improvement is not associated with advancement of molt timing. Shown are results from statistical models examining (a) the effects of pre-breeding molt completion date, age (levels 1, 2+) and tarsus length on the likelihood of gaining a breeder position among subordinates (GLMMPQL, *n*=47), and (b) the effect of the interaction between change of age class (levels 1/2+, 2+/2+) and change of social status (levels 'subordinate/subordinate', 'subordinate/dominant') on the within-individual difference between dates of pre-breeding molt completion in consecutive years (LMM, *n*=35).

(a)

Parameter	β	SE	df	*t* value	*P*
Fixed effects					
Intercept	-10.47	2.60	29	-4.03	<0.001
Molt date	0.01	0.04	10	0.35	0.73
Age†	5.35	3.21	10	1.67	0.13
Tarsus length	1.81	1.96	10	0.92	0.38
Random effects					
σ^2_{year}	3.38×10^{-9}				
$\sigma^2_{territory\ in\ year}$	2.95×10^{-8}				
$\sigma^2_{individual\ in\ territory\ in\ year}$	8.51				
$\sigma^2_{residual}$	9.69×10^{-9}				

†The reference category is age 1.

(b)

Parameter	β	SE	df	*t* value	*P*
Fixed effects					
Intercept	-63.56	16.36	31.00	-3.89	<0.001
Age change†	**62.62**	**19.05**	**31.00**	**3.29**	**0.003**
Status change*	-13.21	26.71	31.00	-0.50	0.62

Age change x status change	6.40	30.19	31.00	0.21	0.83
Random effects					
$\sigma^2_{individual}$	3.08×10^{-12}				
$\sigma^2_{residual}$	1337.40				

Reference categories are as follows: †1/2+, *subordinate/subordinate.

Table S4. Dominant males that molt earlier are not paired with females that start to lay earlier. Shown are results from an LMM examining the effects of age (levels 1, 2+) and the interaction between social status and date of the first egg laid on a given territory, on the date of pre-breeding molt completion (*n*=161).

Parameter	β	SE	df	*t* value	*P*
Fixed effects					
Intercept	165.29	13.21	45.02	12.51	<0.001
Age†	**-48.64**	**9.76**	**151.33**	**-4.98**	**<0.001**
Status*	**-38.96**	**11.40**	**152.45**	**-3.42**	**<0.001**
1st egg date	-0.02	0.07	153.99	-0.31	0.76
1st egg date x status	0.02	0.08	145.71	0.30	0.76
Random effects					
$\sigma^2_{individual}$	323.75				
$\sigma^2_{territory}$	188.19				
σ^2_{year}	204.02				
$\sigma^2_{residual}$	641.10				

Reference categories are as follows: †age 1, *subordinate.

Table S5. Dominant males that molt earlier do not achieve greater seasonal reproductive success. Shown are results from statistical models examining the effects of pre-breeding molt completion date, age (levels 1, 2+), tarsus length, territory quality and group size on reproductive success among dominants: annual number of breeding attempts (GLMM, *n*=166), clutch size (zero-truncated poisson model, *n*=259) and annual number of six-week-old fledglings produced (zero-inflated poisson model, *n*=151).

	Annual number of breeding attempts *n*=166				**Clutch size** *n*=259				**Annual number of 6-week-old fledglings** *n*=151			
Parameter	β	SE	*z* value	*P*	β	SE	*z* value	*P*	β	SE	*z* value	*P*
Fixed effects												
Intercept	0.34	0.38	0.88	0.38	1.10	0.26	4.29	<0.001	-0.81	1.06	-0.76	0.45
Molt date	-0.06	0.05	-1.21	0.23	8.62×10^{-3}	0.04	0.20	0.84	-4.25×10^{-3}	2.61×10^{-3}	-1.63	0.10
Age†	**0.80**	**0.37**	**2.18**	**0.03**	-0.08	0.26	-0.31	0.76	1.01	1.06	0.96	0.34
Tarsus length	-0.04	0.05	-0.78	0.43	0.01	0.04	0.31	0.75	0.05	0.14	0.38	0.71
Territory quality	-0.06	0.05	-1.35	0.18	-0.02	0.04	-0.36	0.72	0.03	0.02	1.20	0.23
Group size	0.01	0.04	0.25	0.81	5.51×10^{-3}	0.04	0.13	0.89	**0.17**	**0.06**	**2.59**	**0.01**
Random effects												
$\sigma^2_{individual}$	0.00				4.08×10^{-9}				7.50×10^{-8}			
$\sigma^2_{territory}$	4.93×10^{-10}				5.15×10^{-9}				3.09×10^{-8}			
σ^2_{year}	0.06				5.21×10^{-8}				0.08			

†The reference category is age 1.

Table S6. Early molt does not predict greater success of breeding attempts, including hatching success, fledging success and fledgling survival. Shown are results from GLMMPQLs examining the effects of pre-breeding molt completion date, age (levels 1, 2+), tarsus length, territory quality and group size on hatching success (n=145), fledging success (n=111) and fledgling survival (n=76).

	Hatching success n=145					**Fledging success** n=111					**Fledgling survival** n=76				
Parameter	β	SE	df	t value	P	β	SE	df	t value	P	β	SE	df	t value	P
Fixed effects															
Intercept	0.46	0.31	104	1.45	0.15	-0.80	0.34	83	-2.39	0.02	-0.46	1.69	60	-0.27	0.79
Molt date	6.11×10^{-4}	1.70×10^{-3}	104	0.36	0.72	-1.98×10^{-3}	1.80×10^{-3}	83	-1.10	0.27	-1.68×10^{-3}	5.94×10^{-3}	60	-0.28	0.78
Age[†]	-0.35	0.32	104	-1.12	0.27	0.46	0.34	83	1.35	0.18	0.92	1.69	60	0.55	0.59
Tarsus length	-0.02	0.09	104	-0.17	0.86	0.17	0.09	83	1.89	0.06	7.03×10^{-3}	0.35	60	0.02	0.98
Territory quality	-0.03	0.06	104	-0.54	0.59	-1.67×10^{-3}	0.06	83	-0.03	0.98	0.09	0.23	60	0.41	0.68
Group size	0.06	0.05	31	1.20	0.24	0.03	0.05	18	0.61	0.55	-0.12	0.18	6	-0.65	0.54
Random effects															
σ^2_{year}	0.08					1.43×10^{-4}					0.57				
$\sigma^2_{territory\ in\ year}$	0.12					0.10					1.17×10^{-4}				
$\sigma^2_{individual\ in\ territory\ in\ year}$	0.12					0.10					1.17×10^{-4}				
$\sigma^2_{residual}$	0.62					0.73					1.10				

[†]The reference category is age 1.

Table S7. Dominant males that molt earlier do not achieve greater extra-pair success and those that molt later are not more likely to lose paternity or divorce. Shown are results from GLMMs examining the effects of pre-breeding molt completion date, tarsus length, territory quality and group size on the likelihood of getting EPP in a given year (*n*=168), the likelihood of losing paternity in a given year (*n*=168) and on the likelihood of experiencing divorce in a given year (*n*=181).

	Obtained EPP within the year (Y/N)				**Lost paternity within the year (Y/N)**				**Divorced within the year (Y/N)**			
	n=5 gains, 163 non-gains				*n*=14 losses, 159 non-losses				*n*=15 divorces, 166 non-divorces			
Parameter	β	SE	*z* value	*P*	β	SE	*z* value	*P*	β	SE	*z* value	*P*
Fixed effects												
Intercept	-4.49	0.93	-4.81	<0.001	-2.56	0.31	-8.18	<0.001	-2.47	0.32	-7.66	<0.001
Molt date	-5.65×10^{-3}	0.02	-0.34	0.74	-2.58×10^{-3}	8.67×10^{-3}	-0.30	0.77	3.01×10^{-3}	8.19×10^{-3}	0.37	0.71
Tarsus length	1.69	1.19	1.42	0.16	0.31	0.47	0.66	0.51	0.21	0.44	0.48	0.63
Territory quality	**0.30**	**0.15**	**1.95**	**0.05**	3.22×10^{-3}	0.07	0.05	0.96	-0.02	0.06	-0.38	0.70
Group size	-0.08	0.40	-0.21	0.84	**0.38**	**0.17**	**2.26**	**0.02**	0.10	0.17	0.61	0.54
Random effects												
$\sigma^2_{individual}$	0.00				0.00				1.27×10^{-10}			
$\sigma^2_{territory}$	0.00				0.00				0.00			
σ^2_{year}	0.00				0.00				0.10			

Table S8. Age does not affect survival or life expectancy and there is no relationship between male molt timing and survival or life expectancy. Shown are results from statistical models examining the effects of pre-breeding molt completion date, age (levels 1, 2, 3, 4, 5), social status, tarsus length, territory quality and group size on (a) annual survival (GLMMPQL, *n*=122) and (b) remaining lifespan (in months) in a given year (negative binomial model, *n*=51), in individuals of known age only.

(a)

	Annual survival (Y/N) *n*=122				
Parameter	β	SE	df	*t* value	*P*
Fixed effects					
Intercept	0.17	0.77	97	0.22	0.83
Molt date	0.01	8.83×10^{-4}	10	1.55	0.15
Age†					
β_{age2}	1.58	0.93	10	1.71	0.12
β_{age3}	-0.08	0.84	10	-0.09	0.93
β_{age4}	1.63	1.39	10	1.17	0.27
β_{age5}	0.14	1.19	10	0.12	0.91
Status*	1.44	0.83	10	1.74	0.11
Tarsus length	0.14	0.37	10	0.38	0.71
Territory quality	5.06×10^{-4}	0.06	97	8.68×10^{-3}	0.99
Group size	0.11	0.20	97	0.55	0.59
Random effects					
σ^2_{year}	1.41×10^{-7}				
$\sigma^2_{territory\ in\ year}$	7.77×10^{-6}				
$\sigma^2_{individual\ in\ territory\ in\ year}$	5.02×10^{-4}				
$\sigma^2_{residual}$	1.01				

Reference categories are as follows: †age 1, *subordinate.

(b)

	Remaining lifespan (months)			
	n=51			
Parameter	β	SE	*z* value	*P*
Fixed effects				
Intercept	1.66	0.74	2.23	0.03
Molt date	1.12×10^{-3}	3.18×10^{-3}	0.35	0.73
Age†				
β_{age2}	-0.06	0.41	-0.14	0.89
β_{age3}	-0.45	0.76	-0.59	0.55
β_{age4}	-0.33	1.11	-0.30	0.76
β_{age5}	-0.45	1.49	-0.30	0.76
Status*	0.25	0.18	1.44	0.15
Tarsus length	0.15	0.36	0.43	0.67
Territory quality	-0.03	0.04	-0.83	0.41
Group size	-0.05	0.04	-1.30	0.19
Random effects				
$\sigma^2_{individual}$	0.72			
$\sigma^2_{territory}$	4.38×10^{-9}			
σ^2_{year}	1.43			

Reference categories are as follows: †age 1, *subordinate.

Table S9. There is no difference of survival or life expectancy between each age class. Shown are results from Tukey's HSD test for post hoc comparisons between each age class (1-5 years old) for both annual survival and remaining lifespan (in months) in a given year.

	Annual survival (Y/N)				**Remaining lifespan (months)**			
	n=122				*n*=51			
Pairwise comparison	β	SE	*z* value	*P*	β	SE	*z* value	*P*
2-1	1.58	0.89	1.78	0.37	-0.06	0.41	-0.14	1.00
3-1	-0.08	0.81	-0.10	1.00	-0.45	0.76	-0.59	0.95
4-1	1.63	1.33	1.23	0.72	-0.33	1.11	-0.30	1.00
5-1	0.14	1.14	0.13	1.00	-0.45	1.49	-0.30	1.00
3-2	-1.66	0.77	-2.17	0.18	-0.40	0.98	-0.40	0.99
4-2	0.05	1.27	0.04	1.00	-0.28	1.39	-0.20	1.00
5-2	-1.44	1.09	-1.33	0.66	-0.39	1.82	-0.22	1.00
4-3	1.71	1.14	1.50	0.55	0.12	0.83	0.14	1.00
5-3	0.22	0.93	0.24	1.00	0.003	1.25	0.002	1.00
5-4	-1.49	1.33	-1.12	0.79	-0.11	0.83	-0.14	1.00

Table S10. There is no relationship between male molt timing and survival or life expectancy. Shown are results from statistical models examining the effects of pre-breeding molt completion date, age, social status, tarsus length, territory quality and group size on annual survival (GLMM, *n*=242) and remaining lifespan (in months) in a given year (negative binomial model, *n*=120).

	Annual survival (Y/N) *n*=242				**Remaining lifespan (months)** *n*=120			
Parameter	β	SE	*z* value	*P*	β	SE	*z* value	*P*
Fixed effects								
Intercept	0.19	0.74	0.26	0.80	1.68	0.61	2.75	0.006
Molt date	9.87×10^{-3}	5.40×10^{-3}	1.83	0.07	-1.23×10^{-5}	1.62×10^{-3}	-0.01	0.99
Age†	0.71	0.69	1.03	0.30	-0.04	0.17	-0.23	0.82
Status*	1.00	0.54	1.86	0.06	0.19	0.15	1.23	0.22
Tarsus length	0.04	0.26	0.15	0.88	0.06	0.18	0.35	0.72
Territory quality	0.01	0.04	0.30	0.77	0.01	0.03	0.40	0.69
Group size	0.23	0.13	1.72	0.09	-9.26×10^{-3}	0.02	-0.40	0.69
Random effects								
$\sigma^2_{individual}$	0.00				0.73			
$\sigma^2_{territory}$	3.61×10^{-10}				0.10			
σ^2_{year}	0.32				1.94			

Reference categories are as follows: †age 1, *subordinate.

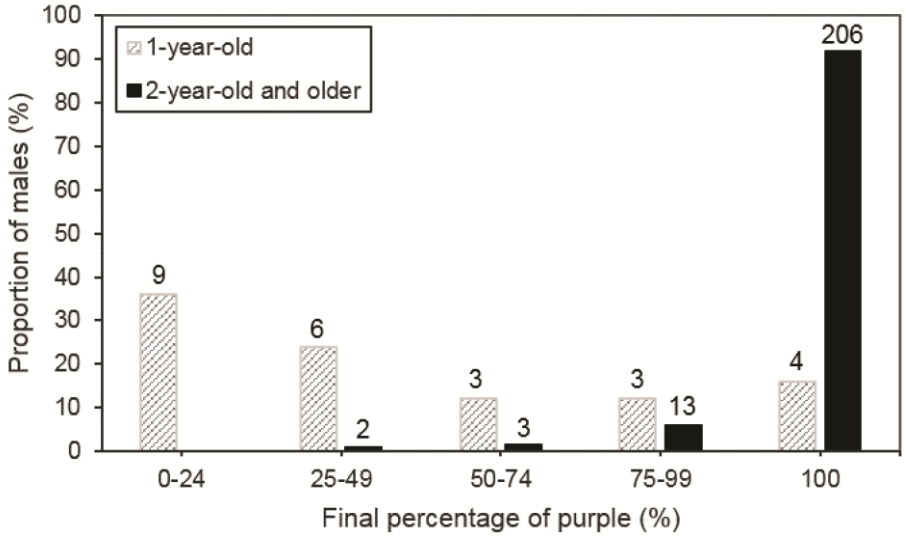


Figure S1. Most second-year and older males are able to produce a complete breeding plumage, whereas most first-year males fail to do so. Shown is the proportion (%) of one-year-old (striped) and older (black) males reaching different ranges of final percentage of purple (%). Numbers represent total number of males in each category.

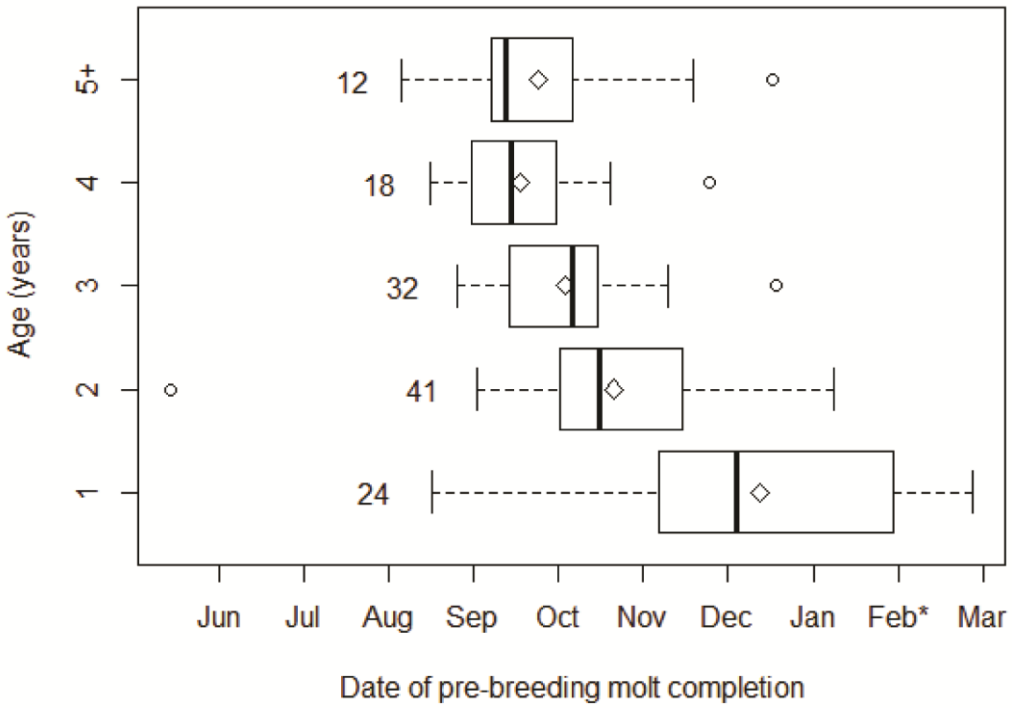


Figure S2. First-year males produce their annual breeding plumage later than males in older age classes. Shown is the effect of age on the timing of pre-breeding molt completion in male purple-crowned fairy-wrens, including individuals of known age only (*n*=130). Diamonds represent the mean, boxes depict the 25th, 50th and 75th percentiles, and whiskers the 10th and 90th percentiles. The sample size is indicated for each age class. Birds of five years old and older have been combined into one class. Month of peak breeding initiation is February indicated with an asterisk (*).

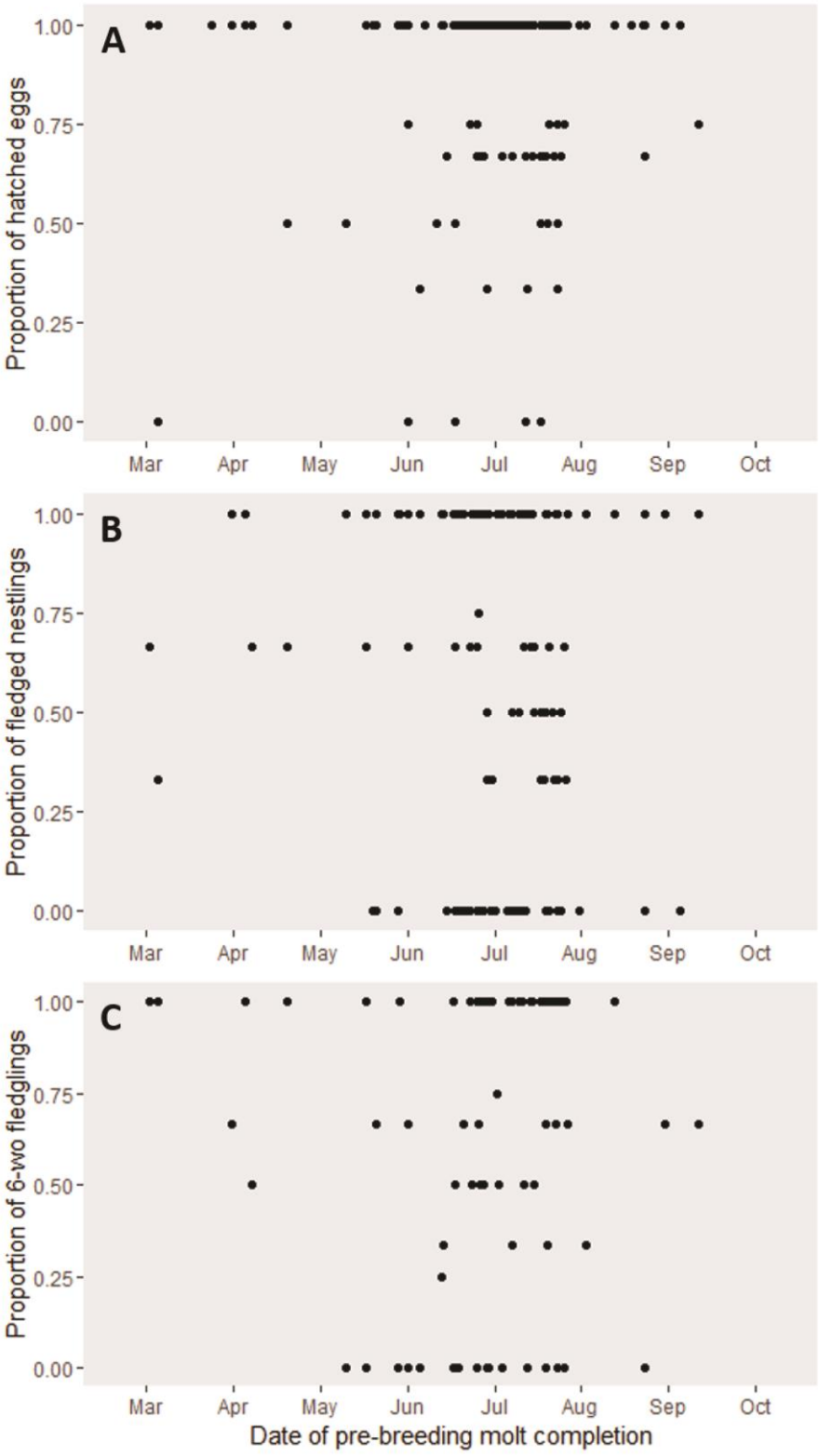


Figure S3. Molt performance is not related to success of breeding attempts. Shown is the absence of correlation between the timing of pre-breeding molt completion and (A) hatching success (*n*=145), (B) fledging success (*n*=111) and (C) fledgling survival (*n*=76). (A)

Hatching success is represented as the proportion of hatched eggs out of the total clutch size; (B) fledging success is represented as the proportion of fledged nestlings out of the total number of nestlings; and (C) fledgling survival is represented as the proportion of fledglings that survived for at least six weeks out of the total number of fledglings.

Chapter 3

From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird

This thesis chapter is in the same form as the final manuscript published in the peer-reviewed journal *Journal of Animal Ecology*. The full reference for the published paper is:

Fan, M., Teunissen, N., Hall, M. L., Hidalgo Aranzamendi, N., Kingma, S. A., Roast, M., Delhey, K., & Peters, A. (2018). From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird. *Journal of Animal Ecology*, **87**, 1274-1285.

Received: 19 December 2017 | Accepted: 13 May 2018

DOI: 10.1111/1365-2656.12855

RESEARCH ARTICLE

Journal of Animal Ecology | BRITISH ECOLOGICAL SOCIETY

From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird

Marie Fan[1] | Niki Teunissen[1] | Michelle L. Hall[2,3] | Nataly Hidalgo Aranzamendi[1] | Sjouke A. Kingma[3,4] | Michael Roast[1] | Kaspar Delhey[1,3] | Anne Peters[1,3]

[1]School of Biological Sciences, Monash University, Clayton, Vic., Australia

[2]School of BioSciences, University of Melbourne, Melbourne, Vic., Australia

[3]Max Planck Institute for Ornithology, Vogelwarte Radolfzell, Radolfzell, Germany

[4]Behavioural and Physiological Ecology, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

Correspondence
Marie Fan, School of Biological Sciences, Monash University, Clayton, Vic., Australia.
Email: marie.fan@monash.edu

Funding information
Max-Planck-Gesellschaft; Australian Research Council, Grant/Award Number: DP150103595 and FT10100505; Holsworth Wildlife Research Endowment

Handling Editor: Elizabeth Derryberry

Abstract

1. The evolution of conspicuous male traits is thought to be driven by female mate choice or male–male competition. These two mechanisms are often viewed as distinct processes, with most studies focusing on female choice.
2. However, both mechanisms of sexual selection can act simultaneously on the same trait (i.e., dual function) and/or interact in a synergistic or conflicting way. Dual-function traits are commonly assumed to originate through male–male competition before being used in female choice; yet, most studies focusing on such traits could not determine the direction of change, lacking phylogenetic information.
3. We investigated the role of conspicuous male seasonal plumage in male–male competitive interactions in the purple-crowned fairy-wren *Malurus coronatus*, a cooperatively breeding bird. Male breeding plumage in most *Malurus* species is selected by female choice through extra-pair mate choice, but unlike its congeners, *M. coronatus* is genetically monogamous, and females do not seem to choose males based on breeding plumage acquisition.
4. Our study shows that, within groups, subordinate males that were older, and therefore higher-ranked in the queue for breeder position inheritance, produced a more complete breeding plumage. In line with this, subordinate males that were older and/or displayed a more complete breeding plumage were more successful in competitively acquiring a breeder position.
5. A role as a signal of competitive ability was experimentally confirmed by presenting models of males: in breeding colours, these received more aggression from resident breeder males than in nonbreeding colours, but elicited limited response from females, consistent with competitors in breeding plumage being perceived as a bigger threat to the breeder male.
6. The role of the conspicuous breeding plumage in mediating male–male interactions might account for its presence in this genetically monogamous species. As phylogenetic reconstructions suggest a past female choice function in *M. coronatus*, this could represent a sexual trait that shifted functions, or a dual-function trait that lost one function. These evolutionary scenarios imply that intra- and

intersexual functions of ornaments may be gained or lost independently and offer new perspectives in understanding the complex dynamics of sexual selection.

KEYWORDS
evolutionary trait loss, extra-pair paternity, functional shift, male–male competition, monogamy, seasonal breeding plumage, social dominance, trait co-option

1 | INTRODUCTION

The evolution and maintenance of elaborate male traits have traditionally been attributed to sexual selection, which operates through two mechanisms: mate choice and same-sex competition (Andersson, 1994; Darwin, 1871). Indeed, numerous male secondary sexual characters are assumed to serve to attract breeding partners and/or to repel opponents in contests over breeding resources and opportunities (Andersson, 1994; Clutton-Brock, 2007; Darwin, 1871). Such characters may act as signals of male quality, providing choosy females and/or rival males with information related to individual body condition, fighting ability and genetic constitution (Andersson, 2006; Hamilton & Zuk, 1982; Zahavi, 1975). Although both mechanisms of sexual selection are credited with the evolution of extravagant male characters, most research has focused on female mate choice (Jones & Ratterman, 2009; McCullough, Miller & Emlen, 2016). Moreover, many studies have depicted a dichotomous view of these mechanisms, with different underlying genetic processes at play and costs involved (Hurd & Enquist, 2005; Jones & Ratterman, 2009; Lachmann, Szamado & Bergstrom, 2001; McCullough et al., 2016).

Nevertheless, various studies have now established that some traits could be selected by male–male competition as well as female choice (i.e., dual function; Berglund, Bisazza & Pilastro, 1996; Hoi & Griggio, 2008; Tarof, Dunn & Whittingham, 2005). Hereby, both mechanisms can interact in a synergistic or conflicting way through mutual reinforcement or opposing selection (Hunt, Breuker, Sadowski & Moore, 2009; Qvarnström & Forsgren, 1998; Wong & Candolin, 2005). Choice by females and competition between males may operate simultaneously or sequentially within populations; as both mechanisms can differ in strength and form, total sexual selection operating on a trait may differ considerably from that imposed by either choice or competition in isolation (Hunt et al., 2009). Furthermore, due to spatial and temporal environmental heterogeneity, selection through both mechanisms may fluctuate and therefore generate complex dynamics in overall sexual selection (Miller & Svensson, 2014). The vast majority of dual-function traits are assumed to originate through male–male competition, and subsequently be co-opted for use in female choice (Berglund et al., 1996). This assumption rests upon the idea that females exploit signals used in male–male aggressive interactions because the honesty of such traits is constantly tested in these interactions and cannot be faked without incurring substantial costs (Berglund et al., 1996). However, the reverse process—female choice cues co-opted for use in male contests—may also occur, although very little evidence for this phenomenon exists (but see Morris, Tudor & Dubois, 2007). Because determining the context in which a dual-function trait initially evolved requires studies of both intra- and intersexual selection in closely related species, as well as phylogenetic information, so far only few studies have been able to test this scenario (Borgia & Coleman, 2000; Morris et al., 2007).

Here, we use the purple-crowned fairy-wren *Malurus coronatus* to investigate the respective roles of female choice and male–male competition in shaping the evolutionary trajectory of a conspicuous male trait. Like most other *Malurus*, the species breeds cooperatively, forms long-term social partnerships and displays seasonal plumages, as both breeder and subordinate males moult annually from a dull nonbreeding plumage into a conspicuous breeding plumage (Peters, Kingma & Delhey, 2013). The timing of moult into breeding plumage by male fairy-wrens is viewed as a classic female choice-driven trait: early moult is strongly selected by female choice for extra-pair (EP) mates who dominate fertilisations (Peters et al., 2013; Brouwer et al., 2017; Figure 1). However, EP mating is very limited in *M. coronatus* (<5% of broods) and mostly driven by incest avoidance (Hidalgo Aranzamendi, Hall, Kingma, Sunnucks & Peters, 2016; Kingma, Hall & Peters, 2013; Kingma, Hall, Segelbacher & Peters, 2009), suggesting no role of female EP mate choice in the evolution of male breeding plumage in this species (Figure 1). In addition, there is no apparent reproductive benefit of early moult (Fan et al., 2017) which, based on the most recent phylogeny (Marki et al., 2017), may thus constitute a vestigial sexual trait (Figure 1). Nevertheless, Fan et al. (2017) reported that breeder males moult earlier than subordinates. Because acquiring and retaining a breeder position is critical for male reproductive success, we hypothesise that male–male competition might drive the persistence of the breeding plumage. We use 6 years of data to test whether variation in breeding plumage of subordinate males predicts success in obtaining a breeder position. Second, we performed model presentations to experimentally test whether plumage state of simulated male intruders affects the strength of territorial defence by breeder males.

2 | MATERIALS AND METHODS

2.1 | Study species

We studied a colour-banded population of *M. coronatus* at the Australian Wildlife Conservancy's Mornington Wildlife Sanctuary (17°31′S, 126°6′E; north Western Australia) from July 2005 to

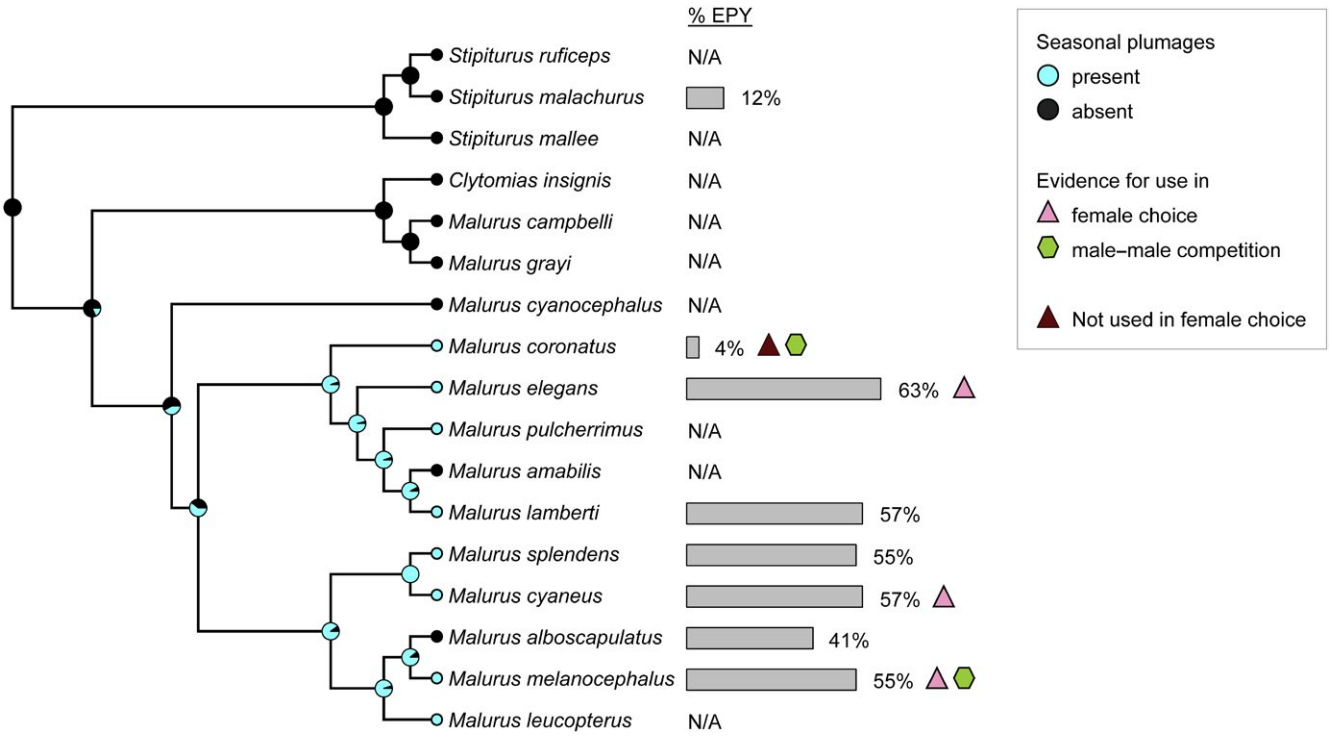


FIGURE 1 Evolution of seasonal plumage in fairy-wrens and emu-wrens. Shown is the ancestral state reconstruction of changes in seasonal plumages for 17 Malurid species, using stochastic mapping (for details see Fan et al., 2017) and based on the supermatrix phylogeny of Marki et al. (2017). Absence of seasonal plumages in *M. amabilis* is based on Schodde (1982). % EPY = levels of extra-pair paternity (proportion of offspring sired by extra-pair males; Brouwer et al., 2017; N/A: no data available) are shown for each species, as well as whether there is evidence that seasonal plumages function in female choice and/or male–male competition (this study; Karubian et al., 2008; Peters et al., 2013; Fan et al., 2017)

November 2016. These birds are restricted to patchy riparian vegetation of *Pandanus aquaticus* and maintain all-purpose territories year-round, linearly arranged along Annie Creek and the Adcock River (Kingma et al., 2009). *Malurus coronatus* can breed year-round with a peak in breeding activity at this site during the wet season (December–March), and a smaller peak in the late dry season (August-September) in some years (Hall & Peters, 2009; Peters et al., 2013).

Malurus coronatus breeds cooperatively, whereby 40%–70% of dominant breeding pairs (distinguished by duet singing; Hall & Peters, 2008, 2009), are accompanied by a number of nonbreeding male and female subordinates (Kingma, Hall, Arriero & Peters, 2010; Kingma, Hall & Peters, 2011a,b).

Males replace the dull brown nonbreeding head plumage annually with purple-and-black feathers (Peters et al., 2013; Figure 2a,c), and this moult overlaps temporally with breeding in some cases (29% of breeder males; Fan et al., 2017). First-year males and subordinate males complete their moult later than older males and breeder males, respectively (Fan et al., 2017). Moreover, although they usually moult to some degree, only 16% of first-year males develop a complete breeding plumage, whereas most older males do so (Fan et al., 2017; see also results).

Subordinate individuals may acquire a breeder position either by taking over part of the natal territory or establishing a new territory (n = 65), or by filling a vacancy left by a deceased breeder (either inheritance of the home territory or dispersal to another, mostly limited to neighbouring territories—Kingma et al., 2011b; Hidalgo Aranzamendi et al., 2016; n = 96). Males less commonly (n = 23) take over a territory of a breeder male that has dispersed after divorce or death of the breeder female, and we have no evidence for eviction of the former breeder in such cases. Dispersal to settle as a subordinate elsewhere is relatively uncommon (n = 24 of 163 records of subordinate male dispersal).

2.2 | Field methods

From July 2005 to March 2011, weekly population censuses were conducted year-round (01 July = start of austral year) to document group size and social status of each uniquely colour-banded male. From October 2011 to November 2016, this information was recorded biannually in population censuses in October-November and May-June (for details see Hidalgo Aranzamendi et al., 2016; Fan et al., 2017). At each sighting, each observer scored the extent of breeding plumage on a scale between 0% and 100% in 5% increments; finer scores could be assigned when birds were captured. Parentage of local birds was determined using six or nine microsatellite loci (for details see Kingma et al., 2009; Hidalgo Aranzamendi et al., 2016). In addition, throughout the study, birds were routinely

FIGURE 2 Nonbreeding and breeding plumages in male purple-crowned fairy-wrens and 3D-printed models used in simulated territorial intrusions. Photographs show the three model types used in the model presentation experiments (b, d, f) and the birds they respectively represent (a, c, e): male purple-crowned fairy-wren in nonbreeding plumage (a, b) and breeding plumage (c, d), and northern fantail (control; e, f). Photos: (a) N. Teunissen; (b-e) L. Lermusiaux; (f) M. Fan

captured to measure tarsus length (a measure of body size). Body size is an important predictor of success in male–male competition generally (Hunt et al., 2009). Tarsus length could be an indicator of male quality in *M. coronatus* as it correlates with song frequency (pitch) in certain male songs (Hall, Kingma & Peters, 2013).

Territory quality was assessed (yearly between 2005 and 2008, and once in 2015) based on the *Pandanus* cover following Kingma et al. (2011a). *Malurus coronatus* generally does not occupy habitat without *Pandanus* (wherein 51% of daytime is spent and 95% of nests are built; Kingma et al., 2011a) and the distribution of *Pandanus* varies considerably within the population.

2.3 | Model presentation experiments

To test whether intruder males in conspicuous plumage are perceived by resident breeder males as a greater threat than males in nonbreeding plumage, we conducted two series of 3D-printed model presentation experiments in 2016. We also tested whether resident breeder males are more aggressive when in breeding plumage themselves and recorded the responses of breeder females to assess their level of interest and aggressiveness. Using 3D-printed models, we avoided any interactive behaviour between live models and residents that could override any potential signal function of coloration (Senar, 2006). Furthermore, because we only manipulated coloration, we could dissociate its role from the effects of potential confounding factors (e.g., age, body condition or size).

Experiments were conducted in 44 territories (21 with no subordinates, 23 with 1–4 subordinates, mean = 1.4) in May-June (start of dry season, when males were completing the moult out of breeding plumage) and 41 territories (19 with no subordinates, 22 with 1–6 subordinates, mean = 1.2) in October-November (end of dry season, when males were completing the moult into breeding plumage); 38 breeder males and 37 breeder females were tested in both seasons.

Models of male purple-crowned fairy-wrens were printed based on a 3D scan of a taxidermic mount and painted in colours that replicated the natural colours, as assessed by avian visual models (for details see Supporting Information Appendix S1a). We made 19 exemplars of a "brown" male (0% breeding plumage), 19 of a "purple-and-black" male (100% breeding plumage) and two of a northern fantail (*Rhipidura rufiventris*, a bird of similar size; also 3D-printed and painted) used as a control (Figure 2). In each territory, the experiment was replicated with each model type in a randomised sequence, with randomly chosen exemplars. There was a minimum of 3 days between replicates to minimise habituation. For each replicate, we placed the model in a relatively open spot well within the territory (identical for all three replicates when possible) and broadcast a standardised

playback of conspecific contact calls and male solo songs to draw the attention of the territory occupants to the model (see Supporting Information Appendix S1b). One observer (MF) continuously recorded the response of the male and female breeders for a period of 15 min from the start of the playback (see Supporting Information Movie S1). We considered that a focal individual started to respond when it either a) started to sing or b) entered a 3-m radius around the model for the first time, and no longer responded when it a) did not sing and b) was > 3 m from the model, for 3 consecutive minutes. We measured (a) latency of response—time between start of the playback and start of the focal individual's response; (b) duration of response; (c) closest approach; (d) time spent within 3 m of the model; (e) time spent within 1 m; and (f) number of songs (solos and duets); and recorded other aggressive behaviours (swooping, pecking). All presentations in which models were not detected (too distant or not sufficiently visible to be seen by the focal individual; see Supporting Information Appendix S1b) were excluded from further analysis (n = 38 of 254). We also recorded the date and time of the day (hourly), substrate type (soil, gravel, stone, rock, wood, grass or leaf litter) and sun strength (from 0 = overcast to 3 = sunny, clear sky). Experiments started at least 30 min after sunrise and ended early afternoon (except for one replicate; see Supporting Information Appendix S1b). No experiment was conducted on rainy and/or windy days.

2.4 | Statistical analyses

We investigated the role of the extent of breeding plumage in male–male interactions by analysing whether it (a) depends on intrinsic and environmental factors (to test for condition-dependence) and (b) predicts the likelihood of subordinates gaining a breeder position. Furthermore, we tested (c) whether breeding plumage of simulated male intruders affects the strength of territorial defence by resident breeder males (and females), and whether this varies with resident male plumage state, using 3D-printed models displaying either extreme of the plumage state range ("purple-and-black" vs. "brown" plumage).

All analyses were carried out in R 3.4.0 (R Core Team, 2017). Linear mixed models (LMM) were built using the packages LME4 (Bates, Maechler, Bolker & Walker, 2015) and LMERTEST (Kuznetsova, Brockhoff & Christensen, 2015). Generalised LMMs (GLMM) were first fitted as generalised linear models without random term to estimate dispersion. If data were under- or overdispersed, appropriate models were selected (see below). All continuous explanatory variables were centred.

2.4.1 | Intrinsic and environmental effects on the extent of breeding plumage

We obtained 279 records of maximum extent of breeding plumage from 108 males over a 6-year period. We built an LMM with the maximum % breeding plumage in a given year as the response variable, and age, within-group rank, tarsus length, territory quality and group size as fixed effects. Bird identity, territory identity and year were included as random intercepts to account for non-independence in the data. Within-group rank was assessed on 01 October, when prebreeding moult is on average completed at the population level (Fan et al., 2017). All breeders were assigned a rank of 1 and subordinates a rank that ranged from 2 to 7 depending on how many other subordinate males were present in the territory and their relative age (the oldest male had a rank of 2, the second oldest a rank of 3, etc.; subordinates of the same age had the same rank). For subordinates, this reflects their rank in the queue for inheriting a breeder position (Kingma et al., 2011b). This analysis was first restricted to birds whose age was accurately known (n = 129 banded as nestlings), but as the maximum % breeding plumage only varied significantly between the first and second year (Supporting Information Tables S1 and S2; see Supporting Information Appendix S2a), it was repeated using two age classes, "1" and "2+", with all birds of unknown age but known to be at least 2 years old (n = 120) included in the "2+" class. When using age classes, we included the age*within-group rank interaction.

2.4.2 | Competitive acquisition of a breeder position

We tested whether the maximum extent of breeding plumage predicted the likelihood of gaining a breeder position elsewhere among subordinate males at the population level (n = 53). We therefore excluded cases of inheritance (n = 12) or splitting of the natal territory (n = 13), as well as when prebreeding moult was temporarily interrupted for > 6 weeks (n = 22; for details see Fan et al., 2017). We used a GLMM with penalised quasi-likelihood (GLMMPQL; underdispersion because many individuals did not obtain a breeder position) using the package "MASS" with the annual status change (became a breeder within the year = 1, did not = 0) as a binomial response variable, and maximum % breeding plumage, age and tarsus length as fixed effects. Due to high correlation between age (levels "1", "2+"; n = 15 and 38, respectively) and maximum % breeding plumage ($|r|$ > 0.7; Dormann et al., 2013), we fitted these predictors in two separate models. Bird identity nested in territory identity nested in year was included as a random intercept. Although GLMMPQLs may at times yield problematic estimates (see Bolker et al., 2009), the results of the models above appear to be robust as they were quantitatively similar to those obtained when using age in months instead (n = 45), as well as when rerunning the models without individuals that died as subordinates before the end of the considered year (n = 39)—to account for potential selective disappearance associated with particular extents of breeding plumage (Supporting Information Table S5).

As breeding vacancies primarily arise when a breeder male of a territory dies or, less commonly, moves away, and because subordinate males usually do not disperse far from their natal territory, the opportunity to compete for a vacancy is probably unequal among subordinates, depending on the distance to a vacancy. Therefore, we performed a case-by-case analysis comparing the percentage of breeding plumage of the "winner" of a vacancy at the time it

appeared and other potential competitors at the same time (referred as "losers"). Losers could be either 1) other within-group subordinate males in cases of inheritance or 2) subordinate males located within the same distance to the vacancy in cases of dispersal (n = 7 inheritance and 22 dispersal cases, each case involving one or more losers; see Supporting Information Appendix S2b). We built an LMM with the difference in % breeding plumage (calculated for all pairs of winners and losers, n = 61) as the response variable, the route used to gain dominance (inheritance or dispersal) as a fixed effect, and winner identity as a random intercept (to control for the fact that winners were compared to multiple losers). A similar LMM was built with the difference in age (in months; calculated for all pairs of winners and losers of known age, n = 34) instead as the response variable.

2.4.3 | Model presentation experiments

Focusing on breeder males only, we tested whether their aggressiveness varied with the 3D-printed model type by investigating five variables (four were fitted in LMMs and transformed to ensure normal distribution of residuals): (a) duration of response (*sqrt*-transformed), (b) closest approach (*sqrt*-transformed), (c) time spent within 3 m (*log*-transformed), (d) time spent within 1 m *(log*-transformed) and (e) number of songs—fitted in a negative binomial model using the package GLMMADMB (non-zero-inflated overdispersion). Physical aggression towards the model occurred in only seven instances (see Supporting Information Appendix S2g); this behaviour was therefore not analysed statistically. For all analyses, we fitted 3D model type (levels "control", "brown", "purple-and-black"), territory quality, group size, presence of fledglings in the group (yes/no), season (start/end of dry season), replicate number (1–3), time of the day and sun strength (0–3, which may affect colour perception; Romero, Hernández-Andrés, Nieves & García, 2002) as fixed effects. Latency of response (independent of model type: Kruskal–Wallis test, $\chi_2^2 = 1.69$, p = 0.43) was also included as variation in the duration of playback heard might affect aggressiveness. Age in years was also included; because it was highly correlated with dominance tenure (i.e., time spent as a breeder; $|r|$ = 0.84) and fitting either variable gave similar results (no effect of either), we only reported the results for age. We also included the presence of within-group subordinate males unrelated to the breeder male (yes/no) and to the breeder female (yes/no) as the presence of reproductive competitors (unrelated males) could affect aggressiveness of the breeder male. Model exemplar, date and substrate type were fitted as random intercepts, whereas bird identity was fitted as both a random intercept to allow for individual variation in baseline responsiveness and a random slope, varying with model type to allow for individual variation in the degree of escalation. We compared the full statistical model including 3D model type with the reduced model excluding it to check for the overall effect of 3D model type using a likelihood-ratio test. Post hoc comparisons between 3D model types were performed using Tukey's Honest Significant Difference test. In addition, we carried out a principal component analysis on the five response variables to summarise variation in aggressiveness (detailed in Supporting Information Appendix S2c), which provided similar results.

We also tested whether the overall response differed between male and female breeders using the data from both sexes and similar statistical models, but fitting 3D model type, latency of response, sex, territory quality, group size, presence of fledglings, season, replicate number, time of the day and sun strength as fixed effects, and bird identity, territory identity, model exemplar, date and type of substrate as random intercepts. As we could not fit a "sex*3D-printed model type" interaction to investigate sex differences in degree of escalation, we then focused on females only and used models similar to those for males (for details see Supporting Information Appendix S2d).

To test whether aggressiveness of breeder males varied with their extent of breeding plumage, we included the percentage of breeding plumage of focal males at the time of the experiments in all the analyses described above. This variable was highly correlated with "season" ($|r|$ = 0.91; 90% of males had ≥90% breeding plumage in October–November, and 86% had <50% in May–June); therefore, we only assessed this effect within each season in separate models, in which % breeding plumage of the focal male was included as a fixed effect and "season" excluded.

3 | RESULTS

3.1 | Intrinsic and environmental effects on the extent of breeding plumage

The maximum extent of breeding plumage produced in a given year was related to age: males produced a more complete breeding plumage as they aged (quadratic relationship between age in months and maximum % breeding plumage achieved; Supporting Information Table S1a). This effect flattened out after 2 years of age, with 1-year-old males developing a less complete breeding plumage (mean = 45%, range: 5%–100%) than males in older age classes, and no significant difference between 2-year-old and older males (respective means (ranges): 93% (35%–100%) vs. 100% (98%–100%); Supporting Information Tables S1b and S2). When using only two age classes, "1" and "2+", the strong effect of age on the maximum percentage of breeding plumage remained (Supporting Information Table S3, Figure 3). Only 16% (4 of 25) of first-year males developed a complete breeding plumage, compared to 92% (206 of 224) of males in their second year or older (all 3-year-old and older males achieved ≥98%).

The maximum extent of breeding plumage was also related to social status as breeder males produced a more complete breeding plumage overall than subordinate males in their group (Supporting Information Table S3, Figure 3). Individual tarsus length, territory quality and group size had no effect on the extent of breeding plumage attained (Supporting Information Table S3).

3.2 | Competitive acquisition of a breeder position

Within social groups, the maximum extent of breeding plumage achieved by subordinate males was related to age relative to other within-group subordinate males (i.e., rank, in the queue for breeder

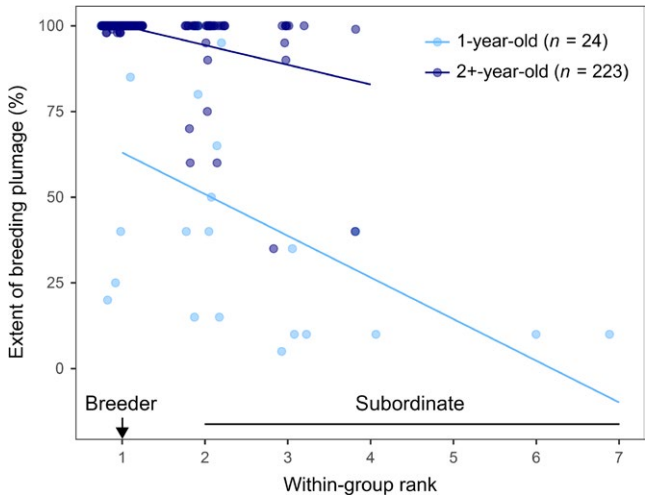


FIGURE 3 Within social groups, breeder males produce a more complete breeding plumage than subordinate males, and among subordinate males, those that are older, and therefore higher-ranked in the social hierarchy, produce a more complete breeding plumage. Moreover, breeding plumage in ≥ 2-year-old males is less affected by within-group rank than in 1-year-old males. Shown is the effect of the interaction of age and within-group rank on % breeding plumage achieved (age: $\beta = 38.77 \pm 2.82$, $t_{225} = 13.73$, $p < 0.001$; rank: $\beta = -12.19 \pm 1.64$, $t_{232} = -7.45$, $p < 0.001$; age × rank: $\beta = 5.93 \pm 1.93$, $t_{200} = 3.08$, $p = 0.002$). Dots depict the (horizontally jittered) raw data, and lines the linear regression lines

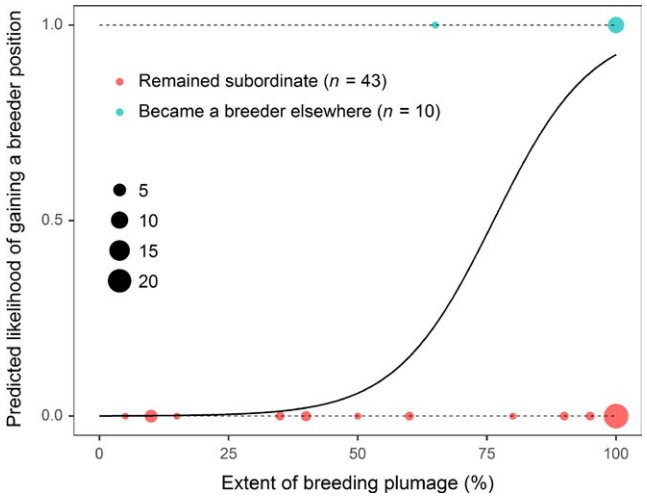


FIGURE 4 At the population level, subordinate males that produce a more complete breeding plumage increase their chances to gain a breeder position by dispersal within that year. Shown is the effect of maximum % breeding plumage on the likelihood of gaining a breeder position elsewhere. Dots depict the raw data for subordinates that did and did not become a breeder (area indicating the number of observations), the dashed lines the upper and lower limits, and the solid line the predicted correlation ($\beta = 0.11 \pm 0.03$, $t_{14} = 3.07$, $p = 0.008$)

position inheritance): subordinate males that were older produced a more complete breeding plumage than their younger group-mates, an effect that was greatest among comparisons with first-year males (Supporting Information Table S3, Figure 3).

At the population level, the maximum extent of breeding plumage achieved by subordinate males in a given year, but not their age, predicted their probability to gain a breeder position by dispersal in that year: subordinate males producing a more complete breeding plumage increased their chances of becoming a breeder elsewhere (Supporting Information Table S5, Figure 4). Tarsus length did not predict whether males obtained a breeder position (Supporting Information Table S5). Similar results were obtained when excluding individuals that died as subordinates before the end of the considered year (Supporting Information Table S5).

Consistent with this, the case-by-case analysis indicated that the extent of breeding plumage predicted success in obtaining a breeder position: males with more breeding plumage than nearby competitors at the time the vacancy appeared were more likely to fill the vacancy (Supporting Information Table S6, Figure 5a). Age also predicted success: older males were more likely to fill the vacancy than younger males (Supporting Information Table S6, Figure 5b; for further analysis see Supporting Information Appendix S2b).

3.3 | Model presentation experiments

The likelihood that individuals detected the model was high (>80%): 46 breeder males detected the 3D-printed model in 216 (of 255) experimental replicates (72 control, 74 brown, 70 purple-and-black) and 47 females did so in 210 replicates (69 control, 72 brown, 69 purple-and-black). For males, model type significantly affected the duration of response (likelihood-ratio test, $\chi_2^2 = 12.98$, $p = 0.002$), the closest approach ($\chi_2^2 = 8.52$, $p = 0.01$) and the time spent within 3 m of the model ($\chi_2^2 = 13.49$, $p = 0.001$). Post hoc pairwise comparisons of 3D model types showed that males responded for longer, approached closer and spent more time within 3 m of purple-and-black models, compared to both brown and control models (the difference in closest approach to purple-and-black and brown models was marginally nonsignificant; Supporting Information Table S7c, Figure 6a–c). However, none of these variables differed significantly between the brown and control models (Supporting Information Table S7c, Figure 6a–c). In contrast, 3D model type had no effect on the time spent within 1 m of the model (likelihood-ratio test, $\chi_2^2 = 3.16$, $p = 0.21$) and the number of songs performed ($D_2 = 2.43$, $p = 0.30$; see also Supporting Information Table S8c, Figure 6d). Females always responded less strongly than males (Supporting Information Table S10) and we found no evidence that their response was affected by the type of model used (likelihood-ratio test, duration: $\chi_2^2 = 3.84$, $p = 0.15$; closest approach: $\chi_2^2 = 3.54$, $p = 0.17$; number of songs: $D_2 = 0.71$, $p = 0.70$; see also Supporting Information Table S11). There was no indication that male response was affected by the behaviour of other group members (Supporting Information Table S12; see Supporting Information Appendix S2e).

Aggressiveness was also related to territory quality: breeder males occupying lower quality territories were more aggressive as they

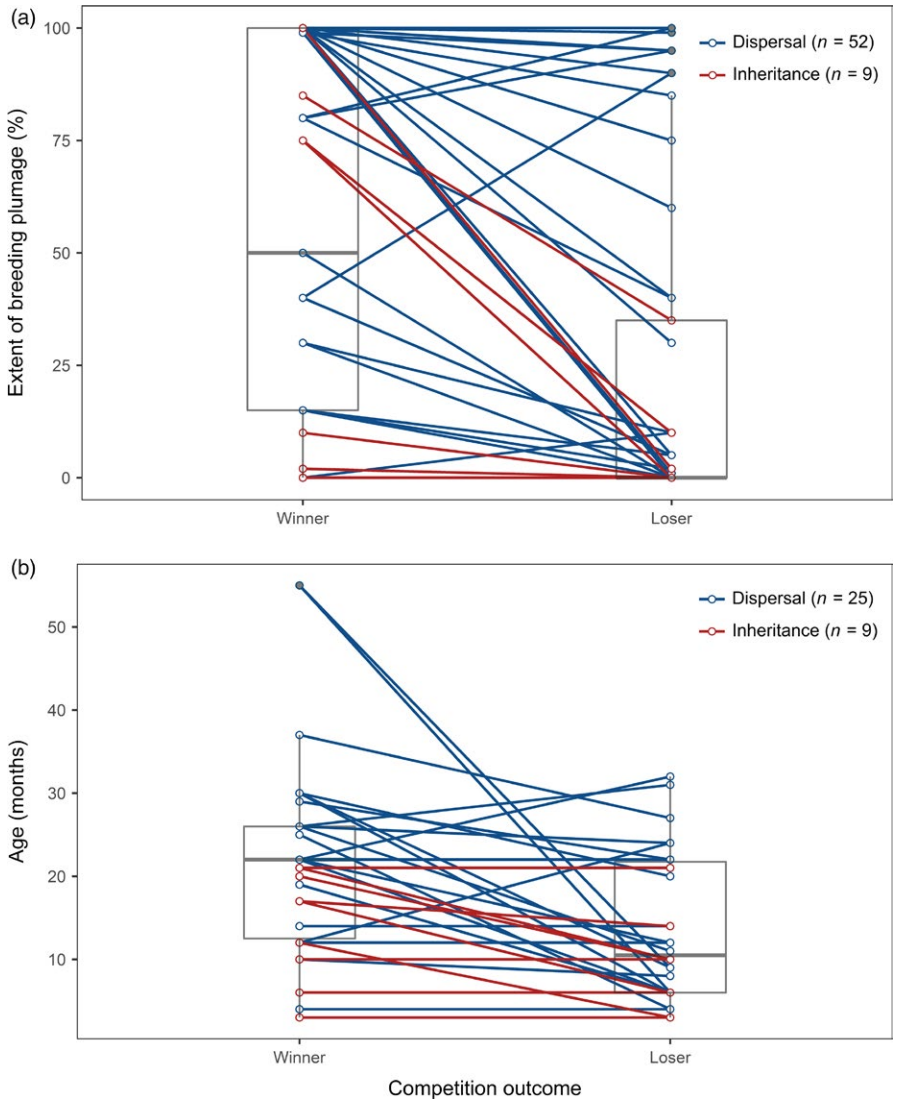


FIGURE 5 Subordinate males that are older and/or display a more complete breeding plumage have higher chances to inherit a breeder vacancy or acquire it elsewhere (when a breeder without a successor dies or moves away). Shown are (a) % breeding plumage and (b) age of winners and losers of competitions for breeder vacancies. Lines depict the pairwise comparisons (difference in % breeding plumage: $\beta = 26.17 \pm 7.66$, $t_{24} = 3.42$, $p = 0.002$; in age: $\beta = 9.43 \pm 2.97$, $t_{22} = 3.18$, $p = 0.004$). Grey boxplots depict the interquartile range (box), medians (dark line) and 2.5% and 97.5% quantiles (whiskers) for winners vs. losers. In inheritance cases, the winner always displayed a similar or larger extent of breeding plumage than the loser, and for brothers of the same age, the one with more breeding plumage inherited the vacancy most of the time

responded for longer, approached closer, spent more time close and performed more songs (Supporting Information Tables S7a and S8a), independently of the model type (i.e., no significant interaction between the effects of territory quality and 3D model type; Supporting Information Table S13; see Supporting Information Appendix S2f). Moreover, males were significantly more aggressive during the period when they usually moult out of breeding plumage (May–June) than when they usually moult into breeding plumage (October–November; Supporting Information Tables S7a and S8a). However, because the two experimental periods differed in the percentage of breeding plumage of breeder males (most males were mostly brown in May-June and fully purple in October–November), the biological significance of this result must be interpreted with caution. When examining each season separately, we found no indication that the percentage of breeding plumage of breeder males affected their aggressiveness (May–June: $\beta = 0.02 \pm 0.03$, $t_{47} = 0.70$, $p = 0.49$; October–November: $\beta = -0.005 \pm 0.03$, $t_{37} = -0.17$, $p = 0.87$). In all these analyses, age, group size, presence of subordinate males unrelated to either breeder, replicate number and time of the day had no effect on the level of aggression (Supporting Information Tables S7a and S8a).

4 | DISCUSSION

Our findings show that the extent of breeding plumage achieved by males increased both with their age and within-group rank. In addition, it appeared to be a strong predictor of success in male–male contests for the acquisition of a breeding territory, with subordinate males in more complete breeding plumage being more likely to win a breeder position. Simulated male intruders in breeding plumage received more aggression from resident breeder males than those in nonbreeding plumage. Taken together, our results strongly suggest that the conspicuous breeding plumage of male *M. coronatus* functions as an intrasexual signal of dominance and competitiveness.

4.1 | Obtaining a breeder position

In *M. coronatus*, subordinate males very rarely (<1%) reproduce until they gain a breeder position (Kingma et al., 2009). This is achieved by inheriting a position in the natal territory or by dispersing, usually to a neighbouring territory, and may occur at any time of the year depending on the time of disappearance of a breeder (Supporting

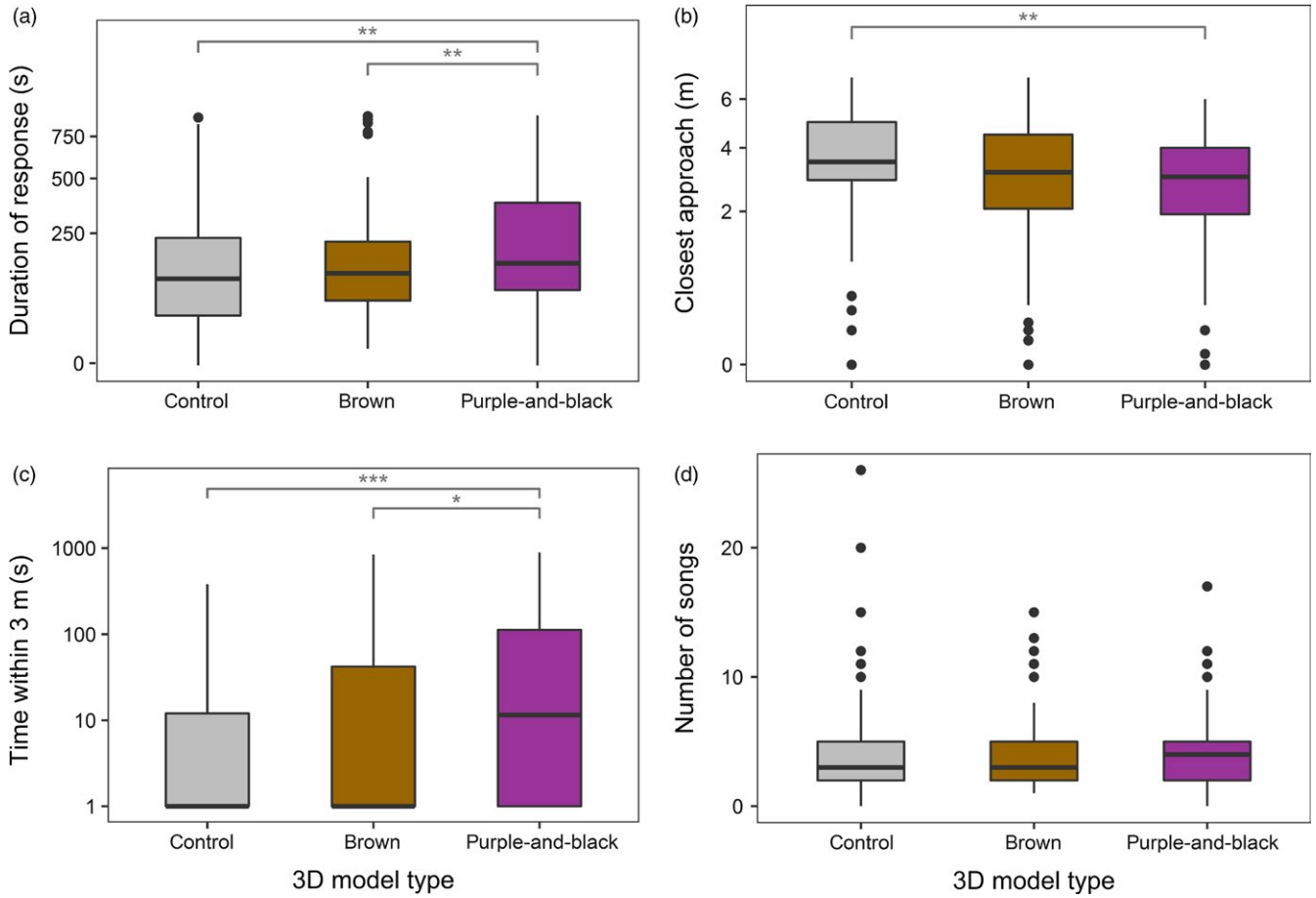


FIGURE 6 Simulated intruders in breeding plumage are perceived as a greater threat to breeder males. Breeder males (a) respond for longer, (b) approach closer and (c) spend more time within 3 m of purple-and-black models compared to brown and control models; however, (d) the number of songs they perform is independent of 3D model type (n = 216). Boxplots depict the interquartile range (box), medians (dark line), 2.5% and 97.5% quantiles (whiskers) and outliers (black dots). Significant pairwise differences are indicated as follows: *p < 0.05, **p < 0.01, ***p < 0.001 (see Supporting Information Tables S7c and S8c for tests statistics). The y-axis scale is *sqrt*-transformed in (a, b) and *log*-transformed in (c)

Information Table S4; see Supporting Information Appendix S2b). Inheritance involves no overt aggression between subordinates within groups and appears mostly to occur through an orderly, age-based queue (Kingma et al., 2011b). When ranking each subordinate male by its relative age in the social group, we found that higher-ranked subordinates produce a more complete breeding plumage (Figure 3), and the extent of breeding plumage of subordinate males correlates with their likelihood of inheriting a breeder position. It is possible that higher-ranked (older) males invest more in acquisition of the breeding plumage, and lower-ranked males refrain from doing so as this reflects their relative prospects for social improvement in the near future. Lower-ranked males may thereby also reduce the aggression received from higher-ranked males (Doucet, McDonald, Foster & Clay, 2007; Karubian, Sillett & Webster, 2008), as the results of the simulated territorial intrusions suggest that males in breeding plumage present a greater perceived threat (see below). Alternatively, lower-ranked subordinates may be physiologically suppressed by higher-ranked individuals ("stress of subordination" or "psychological castration"; Reyer, Dittami & Hall, 1986; Creel, 2001; Brouwer et al., 2009), which constrains the development of their breeding plumage. Either way, less complete breeding plumage of subordinate males in groups with higher-ranked males reduces chances for such subordinates to competitively acquire a breeder position elsewhere, indicating an additional unexpected cost for such subordinates.

The extent of breeding plumage predicts the likelihood of gaining a breeder position. At the population level, subordinates with more complete breeding plumage are more likely to competitively take over a vacant breeder position outside of their natal territory that year (Figure 4). In addition, on a case-by-case basis, males with more complete breeding plumage have greater chances to win a vacancy regardless whether this is achieved by inheritance or dispersal (Figure 5a). Because the extent of breeding plumage is highly correlated with age, and age also predicts competitive outcomes (Figure 5b), the former may serve as a visual signal of age among competitors. However, breeding plumage appears a stronger predictor than age at the population level (Supporting Information Table S5), possibly because it may be easier to assess than male age,

especially by unfamiliar individuals. Taken together, our findings indicate that a larger extent of breeding coloration is associated with increased competitive ability to access a breeder position, similar to observations made in several other bird species (reviewed in Senar, 2006 and Santos, Scheck & Nakagawa, 2011).

4.2 | Territoriality: defence of the breeder position

Simulated territorial intrusions using 3D-printed models of males supported a role for breeding plumage in male–male competitive social interactions. Models in breeding colours elicited stronger aggressive responses from resident breeder males, compared to both brown and control models (Figure 6). This suggests that male intruders in complete breeding plumage are perceived as a greater threat to the resident male, as breeder males, and not females, led the defensive response and displayed higher aggressiveness (Supporting Information Table S10). On the other hand, simulated intrusions by pairs (using playback of duets; Hall & Peters, 2008) have been shown to elicit highly coordinated responses from resident breeding pairs, indicative of cooperative territorial defence. To resident breeder males, an unknown single male in breeding plumage presumably represents a risk of territorial usurpation (which might occasionally occur: we observed 23 cases of breeder dispersal, although we do not know if these were voluntary or evictions by the successor). This finding is also consistent with our observation that subordinate males in a territory generally display less complete breeding plumage than the breeder male, which could be an attempt to signal subordination and avoid within-group agonistic interactions. Similar to that, Karubian et al. (2008) reported that in *M. melanocephalus*, dull brown males are socially subordinate to bright males, and bright caged stimulus males receive higher levels of aggression than dull ones.

Independently of intruders' plumage state, and against expectations, we found that males residing in higher-quality, therefore more valuable, territories displayed less aggression (Supporting Information Tables S7a and S8a). Such males generally live in denser areas but we found no evidence for lower aggression in areas with higher population density (Supporting Information Table S14; see Supporting Information Appendix S2h). Alternatively, if defeated males from lower quality territories are of lower quality or in lower condition, they may be less able to competitively acquire another territory and therefore act more desperately and be more risk-prone (Cain & Langmore, 2016; Grafen, 1987; Wolf, Van Doorn, Leimar & Weissing, 2007).

4.3 | From ornament to armament or loss of function?

Our correlational and experimental results suggest that male–male competition is the selective mechanism responsible for the persistence of male conspicuous breeding plumage, as evident from its importance for obtaining and defending a breeder position. Although we cannot completely rule out the possibility that the extent of breeding plumage is used by resident females to assess potential mates when they attempt to settle as the new breeder, none of our previous and current findings support this. Our model presentation experiment strongly points to a role of the breeding plumage in male–male competition only, with females displaying limited interest in the simulated intruder, and no apparent discrimination of breeding plumage vs. brown models (Supporting Information Tables S10 and S11), and there is no evidence for female choice or reproductive benefits of early acquisition of the breeding plumage (Fan et al., 2017). This is in strong contrast with closely related *Malurus*, where seasonal timing of acquisition of the male ornamental plumage is critical for female EP mate choice, and thereby for male reproductive success, due to very high EPP levels (Figure 1). Because *M. coronatus*' phylogenetic position is nested within this clade (Marki et al., 2017; Figure 1), male breeding plumage in this species could represent a sexual trait formerly selected by female EP mate choice that underwent a shift in function, being subsequently selected by male–male competition. Alternatively, it is possible that breeding plumage elaboration ancestrally had a dual function, being used in both female choice and male–male competition (as may currently be the case in at least one species; Figure 1), whereby the former function was lost as extreme EPP levels and female EP mate choice disappeared (see Kingma et al., 2009). Until more studies investigate the role of *Malurus* breeding plumage in male–male competition (Peters et al., 2013), we cannot confirm which evolutionary scenario—loss or shift of function—is the most likely, but either way our results highlight greater flexibility in function of sexual ornaments than widely appreciated.

Darwin (1859) already appreciated that when a function of a trait is lost, the trait will disappear unless it has multiple functions or switches functions (Lahti et al., 2009). This is also true for sexual ornaments, but not often considered, with studies generally focusing on co-option of an ornament. For example, a substantial number of studies have shown that ornaments used for intrasexual competition in males and females can also be preferred in mate choice (Berglund et al., 1996; Stern & Servedio, 2017). This is generally explained by the armament-ornament hypothesis, which states that traits used in intrasexual competition become co-opted for intersexual mate choice, because their ability to signal individual qualities is also useful for assessment of potential mates (Berglund et al., 1996). Our study shows that the alternative scenarios, traits used in intersexual mate choice switching function to intrasexual competition, or a dual-function trait losing one function, should also be considered. Such scenarios, and our results, imply that intra- and intersexual functions of ornaments can be gained or lost independently and offer possible explanations for the diversity of sexual ornaments and functions. Because these two mechanisms may differ in the selection they impose on sexual traits (Hunt et al., 2009), a shift or loss of function may be characterised by changes in trait expression or selection switching among multiple signalling components. A growing number of studies investigating sexual traits across different populations, at different times, have demonstrated that sexual selection is subject to fluctuations, which may generate very complex evolutionary dynamics (Miller & Svensson, 2014). Our study of the role of male–male competition in a genus renowned for strong selection through

female mate choice further illustrates the complexity of the interplay between the mechanisms of sexual selection and that our view of how sexual selection works may still be incomplete. More generally, it shows that the integrated studies of both mechanisms of sexual selection and all signal components of an ornamental trait in closely related species, with detailed phylogenetic information, can help to uncover new—or rediscover old—evolutionary scenarios and provide further insights into the complex dynamics of sexual selection.

ACKNOWLEDGEMENTS

This research was approved by the Australian Bird and Bat Banding Scheme (license 2230); Western Australia Department of Parks and Wildlife; Australian Wildlife Conservancy; and the ethics committees of the School of Biological Sciences at Monash University; and the Max Planck Institute for Ornithology. Our thanks to many volunteers during fieldwork, especially to L. Lermusiaux for help with the experiments and for photographs, and the support of staff at Australian Wildlife Conservancy's Mornington Wildlife Sanctuary. We are grateful for support by the Max Planck Society (to AP), the Australian Research Council (FT10100505 & DP150103595 to AP), the Holsworth Wildlife Research Endowment—Equity Trustees Charitable Foundation (to MF) and the Australian Wildlife Conservancy and its supporters.

AUTHORS' CONTRIBUTIONS

M.F., A.P. and K.D. conceived the ideas and designed methodology, analysed the data, and led the writing of the manuscript; all authors collected data, contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.t46100m (Fan et al., 2018).

ORCID

Marie Fan http://orcid.org/0000-0002-9332-0227
Michelle L. Hall http://orcid.org/0000-0002-1263-8314
Nataly Hidalgo Aranzamendi http://orcid.org/0000-0001-6967-0780
Anne Peters http://orcid.org/0000-0001-8071-0560

REFERENCES

Andersson, M. (1994). *Sexual selection*. Princeton, NJ: Princeton University Press.

Andersson, M. (2006). Condition-dependent indicators in sexual selection: Development of theory and tests. In J. R. Lucas, & L. W. Simmons (Eds.), *Essays in Animal Behaviour: Celebrating 50 Years of Animal Behaviour* (pp. 255–269). London, UK: Elsevier Academic Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Berglund, A., Bisazza, A., & Pilastro, A. (1996). Armaments and ornaments: An evolutionary explanation of traits of dual utility. *Biological Journal of the Linnean Society*, *58*, 385–399. https://doi.org/10.1111/j.1095-8312.1996.tb01442.x

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*, 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Borgia, G., & Coleman, S. W. (2000). Co-option of male courtship signals from aggressive display in bowerbirds. *Proceedings of the Royal Society of London B: Biological Sciences*, *267*, 1735–1740. https://doi.org/10.1098/rspb.2000.1203

Brouwer, L., Groothuis, T. G., Vedder, O., Eikenaar, C., Richardson, D. S., & Komdeur, J. (2009). Do primary males physiologically suppress subordinate males? An experiment in a cooperatively breeding passerine. *Ethology*, *115*, 576–587. https://doi.org/10.1111/j.1439-0310.2009.01643.x

Brouwer, L., van de Pol, M., Aranzamendi, N. H., Bain, G., Baldassarre, D. T., Brooker, L. C., … Hall, M. L. (2017). Multiple hypotheses explain variation in extra-pair paternity at different levels in a single bird family. *Molecular Ecology*, *26*, 6717–6729. https://doi.org/10.1111/mec.14385

Cain, K. E., & Langmore, N. E. (2016). Female song and aggression show contrasting relationships to reproductive success when habitat quality differs. *Behavioral Ecology and Sociobiology*, *70*, 1867–1877. https://doi.org/10.1007/s00265-016-2192-1

Clutton-Brock, T. (2007). Sexual selection in males and females. *Science*, *318*, 1882–1885. https://doi.org/10.1126/science.1133311

Creel, S. (2001). Social dominance and stress hormones. *Trends in Ecology & Evolution*, *16*, 491–497. https://doi.org/10.1016/S0169-5347(01)02227-3

Darwin, C. (1859). *On the origin of species*. London: John Murray.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London, UK: John Murray.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Münkemüller, T. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Doucet, S. M., McDonald, D. B., Foster, M. S., & Clay, R. P. (2007). Plumage development and molt in Long-tailed Manakins (*Chiroxiphia linearis*): Variation according to sex and age. *The Auk*, *124*, 29–43. https://doi.org/10.1642/0004-8038(2007)124[29:PDAMIL]2.0.CO;2

Fan, M., Hall, M. L., Kingma, S. A., Mandeltort, L. M., Hidalgo Aranzamendi, N., Delhey, K., & Peters, A. (2017). No fitness benefits of early molt in a fairy-wren: Relaxed sexual selection under genetic monogamy? *Behavioral Ecology*, *28*, 1055–1067. https://doi.org/10.1093/beheco/arx065

Fan, M., Teunissen, N., Hall, M. L., Hidalgo Aranzamendi, N., Kingma, S. A., Roast, M., … Peters, A. (2018). Data from: From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.t46100m

Grafen, A. (1987). The logic of divisively asymmetric contests: Respect for ownership and the desperado effect. *Animal Behaviour*, *35*, 462–467. https://doi.org/10.1016/S0003-3472(87)80271-3

Hall, M. L., Kingma, S. A., & Peters, A. (2013). Male songbird indicates body size with low-pitched advertising songs. *PLoS ONE*, *8*, e56717. https://doi.org/10.1371/journal.pone.0056717

Hall, M. L., & Peters, A. (2008). Coordination between the sexes for territorial defence in a duetting fairy-wren. *Animal Behaviour*, *76*, 65–73. https://doi.org/10.1016/j.anbehav.2008.01.010

Hall, M. L., & Peters, A. (2009). Do male paternity guards ensure female fidelity in a duetting fairy-wren? *Behavioral Ecology*, *20*, 222–228. https://doi.org/10.1093/beheco/arn139

Hamilton, W. D., & Zuk, M. (1982). Heritable true fitness and bright birds: A role for parasites? *Science*, *218*, 384–387. https://doi.org/10.1126/science.7123238

Hidalgo Aranzamendi, N., Hall, M. L., Kingma, S. A., Sunnucks, P., & Peters, A. (2016). Incest avoidance, extrapair paternity, and territory quality drive divorce in a year-round territorial bird. *Behavioral Ecology*, *27*, 1808–1819. https://doi.org/10.1093/beheco/arw101

Hoi, H., & Griggio, M. (2008). Dual utility of a melanin-based ornament in bearded tits. *Ethology*, *114*, 1094–1100. https://doi.org/10.1111/j.1439-0310.2008.01566.x

Hunt, J., Breuker, C. J., Sadowski, J. A., & Moore, A. J. (2009). Male–male competition, female mate choice and their interaction: Determining total sexual selection. *Journal of Evolutionary Biology*, *22*, 13–26. https://doi.org/10.1111/j.1420-9101.2008.01633.x

Hurd, P. L., & Enquist, M. (2005). A strategic taxonomy of biological communication. *Animal Behaviour*, *70*, 1155–1170. https://doi.org/10.1016/j.anbehav.2005.02.014

Jones, A. G., & Ratterman, N. L. (2009). Mate choice and sexual selection: What have we learned since Darwin? *Proceedings of the National Academy of Sciences*, *106*, 10001–10008.

Karubian, J., Sillett, T. S., & Webster, M. S. (2008). The effects of delayed plumage maturation on aggression and survival in male red-backed fairy-wrens. *Behavioral Ecology*, *19*, 508–516. https://doi.org/10.1093/beheco/arm159

Kingma, S. A., Hall, M. L., Arriero, E., & Peters, A. (2010). Multiple benefits of cooperative breeding in purple-crowned fairy-wrens: A consequence of fidelity? *Journal of Animal Ecology*, *79*, 757–768.

Kingma, S. A., Hall, M. L., & Peters, A. (2011a). No evidence for offspring sex-ratio adjustment to social or environmental conditions in cooperatively breeding purple-crowned fairy-wrens. *Behavioral and Ecological Sociobiology*, *65*, 1203–1213. https://doi.org/10.1007/s00265-010-1133-7

Kingma, S. A., Hall, M. L., & Peters, A. (2011b). Multiple benefits drive helping behavior in a cooperatively breeding bird: An integrated analysis. *The American Naturalist*, *177*, 486–495. https://doi.org/10.1086/658989

Kingma, S. A., Hall, M. L., & Peters, A. (2013). Breeding synchronization facilitates extrapair mating for inbreeding avoidance. *Behavioral Ecology*, *24*, 1390–1397. https://doi.org/10.1093/beheco/art078

Kingma, S. A., Hall, M. L., Segelbacher, G., & Peters, A. (2009). Radical loss of an extreme extra-pair mating system. *BMC Ecology*, *9*, 15. https://doi.org/10.1186/1472-6785-9-15

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *Package 'lmerTest'*. R version 2.0-33.

Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, *98*, 13189–13194. https://doi.org/10.1073/pnas.231216498

Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., … Foster, S. A. (2009). Relaxed selection in the wild. *Trends in Ecology & Evolution*, *24*, 487–496. https://doi.org/10.1016/j.tree.2009.03.010

Marki, P. Z., Jønsson, K. A., Irestedt, M., Nguyen, J. M., Rahbek, C., & Fjeldså, J. (2017). Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, *107*, 516–529. https://doi.org/10.1016/j.ympev.2016.12.021

McCullough, E. L., Miller, C. W., & Emlen, D. J. (2016). Why sexually selected weapons are not ornaments. *Trends in Ecology & Evolution*, *31*, 742–751. https://doi.org/10.1016/j.tree.2016.07.004

Miller, C. W., & Svensson, E. I. (2014). Sexual selection in complex environments. *Annual Review of Entomology*, *59*, 427–445. https://doi.org/10.1146/annurev-ento-011613-162044

Morris, M. R., Tudor, M. S., & Dubois, N. S. (2007). Sexually selected signal attracted females before deterring aggression in rival males. *Animal Behaviour*, *74*, 1189–1197. https://doi.org/10.1016/j.anbehav.2007.01.019

Peters, A., Kingma, S. A., & Delhey, K. (2013). Seasonal male plumage as a multi-component sexual signal: Insights and opportunities. *Emu*, *113*, 232–247. https://doi.org/10.1071/MU12083

Qvarnström, A., & Forsgren, E. (1998). Should females prefer dominant males? *Trends in Ecology & Evolution*, *13*, 498–501. https://doi.org/10.1016/S0169-5347(98)01513-4

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reyer, H. U., Dittami, J. P., & Hall, M. R. (1986). Avian helpers at the nest: Are they psychologically castrated? *Ethology*, *71*, 216–228.

Romero, J., Hernández-Andrés, J., Nieves, J. L., & García, J. A. (2002). Color coordinates of objects with daylight changes. *Color Research and Application*, *28*, 25–35.

Santos, E. S., Scheck, D., & Nakagawa, S. (2011). Dominance and plumage traits: Meta-analysis and metaregression analysis. *Animal Behaviour*, *82*, 3–19. https://doi.org/10.1016/j.anbehav.2011.03.022

Schodde, R. (1982). *The fairy-wrens: A monograph of the Maluridae*. Melbourne, Vic.: Lansdowne Editions.

Senar, J. (2006). Color displays as intrasexual signals of aggression and dominance. In G. E. Hill, & K. J. McGraw (Eds.), *Bird Coloration*, Vol. *2* (pp. 87–136). Cambridge, MA: Harvard University Press.

Stern, C. A., & Servedio, M. R. (2017). Evolution of a mating preference for a dual-utility trait used in intrasexual competition in genetically monogamous populations. *Ecology & Evolution*, *7*, 8008–8016. https://doi.org/10.1002/ece3.3145

Tarof, S. A., Dunn, P. O., & Whittingham, L. A. (2005). Dual functions of a melanin-based ornament in the common yellowthroat. *Proceedings of the Royal Society of London B: Biological Sciences*, *272*, 1121–1127. https://doi.org/10.1098/rspb.2005.3053

Wolf, M., Van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature*, *447*, 581–584. https://doi.org/10.1038/nature05835

Wong, B., & Candolin, U. (2005). How is female mate choice affected by male competition? *Biological Reviews*, *80*, 559–571. https://doi.org/10.1017/S1464793105006809

Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology*, *53*, 205–214. https://doi.org/10.1016/0022-5193(75)90111-3

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Fan M, Teunissen N, Hall ML, et al. From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird. *J Anim Ecol*. 2018;00:1–13. https://doi.org/10.1111/1365-2656.12855

Supplementary Material

Appendix S1: Model presentation experiments

(a) 3D model development

Printing & infiltration – 3D models of male purple-crowned fairy-wrens were printed using a Projet 660 3D printer and a gypsum-based powder (Visijet PXL Core) as the core material. The models were then soaked with a 3DS ColorBond superglue to be strengthened and sealed, and left to dry. All materials listed above were purchased from 3D Systems (Rock Hill, USA). Painting – In October and November 2015, we measured plumage reflectance in 36 males either in non-breeding plumage or in partial or complete breeding plumage. We used an AvaSpec-2048 spectrometer connected to an AvaLight-XE xenon pulsed light source using a bifurcated fibre optic cable fitted at the end with a cylindrical probe to standardise measuring distance and exclude ambient light. For each male we took five measurements of six plumage patches at a perpendicular probe angle as follows: 1) purple head crown, 2) black head centre (both head patches were measured on males displaying a sufficiently large extent of breeding plumage only), 3) black cheek, 4) buff-white throat, 5) brown back and 6) blue tail; i.e. a maximum of 180 measurements per plumage patch. Reflectance spectra between 300 and 700 nm (the visual sensitivity range of birds; Cuthill, 2006) were calculated relative to a WS-2 white standard using the software AVASOFT 7.5. All spectrometry materials and software were purchased from Avantes (Apeldoorn, Netherlands). Each set of (max.) 180 reflectance spectra associated with a plumage patch was subsequently transformed into a set of xyz coordinates that defined the position of these spectra in the 3D visual space of birds, following the methods described by Delhey et al. (2014). Visual models require knowledge on the visual sensitivity functions of the four types of cone used by birds in colour vision, the relative abundance of each of these cones in the retina and the spectrum of illuminating light. Colour

vision in birds is mediated by four types of single cone sensitive to very short (VS), short (S), medium (M) and long (L) wavelengths of light (Vorobyev et al., 1998). Variation in visual sensitivity between species is mainly restricted to the VS and S cones and birds can be generally classified in two groups: ultraviolet-sensitive (U-type) and violet-sensitive (V-type) species; and U-type species have VS cones with peak sensitivity shifted towards shorter wavelengths (Hart & Hunt, 2007). As *Malurus coronatus* is a V-type species (Ödeen et al., 2012), we used the associated visual sensitivity function obtained from Endler and Mielke (2005). We used average cone proportions as obtained from Hart (2001) (0.38:0.69:1.14:1.00 for VS:S:M:L, respectively) and combined these with behavioural estimates of the Weber fraction (0.1; Vorobyev et al., 1998; Lind, Chavez & Kelber 2014) using formula (10) in Vorobyev et al. (1998) to obtain the noise-to-signal ratios for each cone type (ν_{VS}=0.162, ν_S=0.120, ν_M=0.094, ν_L=0.1). We used the spectrum of standard daylight (D65, open habitats) as illuminant (Vorobyev et al., 1998).

In order to replicate the colour of each plumage patch, we tested different mixtures of paints on spare 3D models, by measuring their reflectance using the same method as above. We used different types of 'Acrylic paints' (J. burrows®, Rockland, USA) containing barium sulphate precipitate (purple crown: 'Violet' and 'Light Blue'; black head centre and cheek: 'Black Night'; white throat and brown back: 'Snow White' and 'Expresso'; and blue tail: 'Ocean Blue', 'Bright Aqua Green' and 'Black Night'), as well as an 'Interference Violet (Fine)' paint (GOLDEN Fluid Acrylics®, New Berlin, USA) containing titanium dioxide-coated mica particles for the purple crown only. The reflectance spectra of the painted colours were subsequently transformed into coordinates of avian visual space (as described above) to be compared with the natural colours. Differences in colouration between two reflectance spectra are represented by Euclidean distances and their unit is the 'just noticeable difference' (JND), whereby a difference below 1 JND is deemed not to be discriminable (Vorobyev et al.,

1998). For each pair of colours (e.g., natural purple *vs* painted purple), we computed the minimum Euclidean distance between the average position of the painted colour and the position of the spectra measured for the natural colour to assess whether the painted colour sufficiently resembled the natural colour (Fig. S1). Minimum Euclidean distances for all pairs of colours were below 1 JND: purple: d = 0.96, black: d = 0.13, buff-white: d = 0.18, brown: d = 0.80, blue: d = 0.32 (Fig. S1).

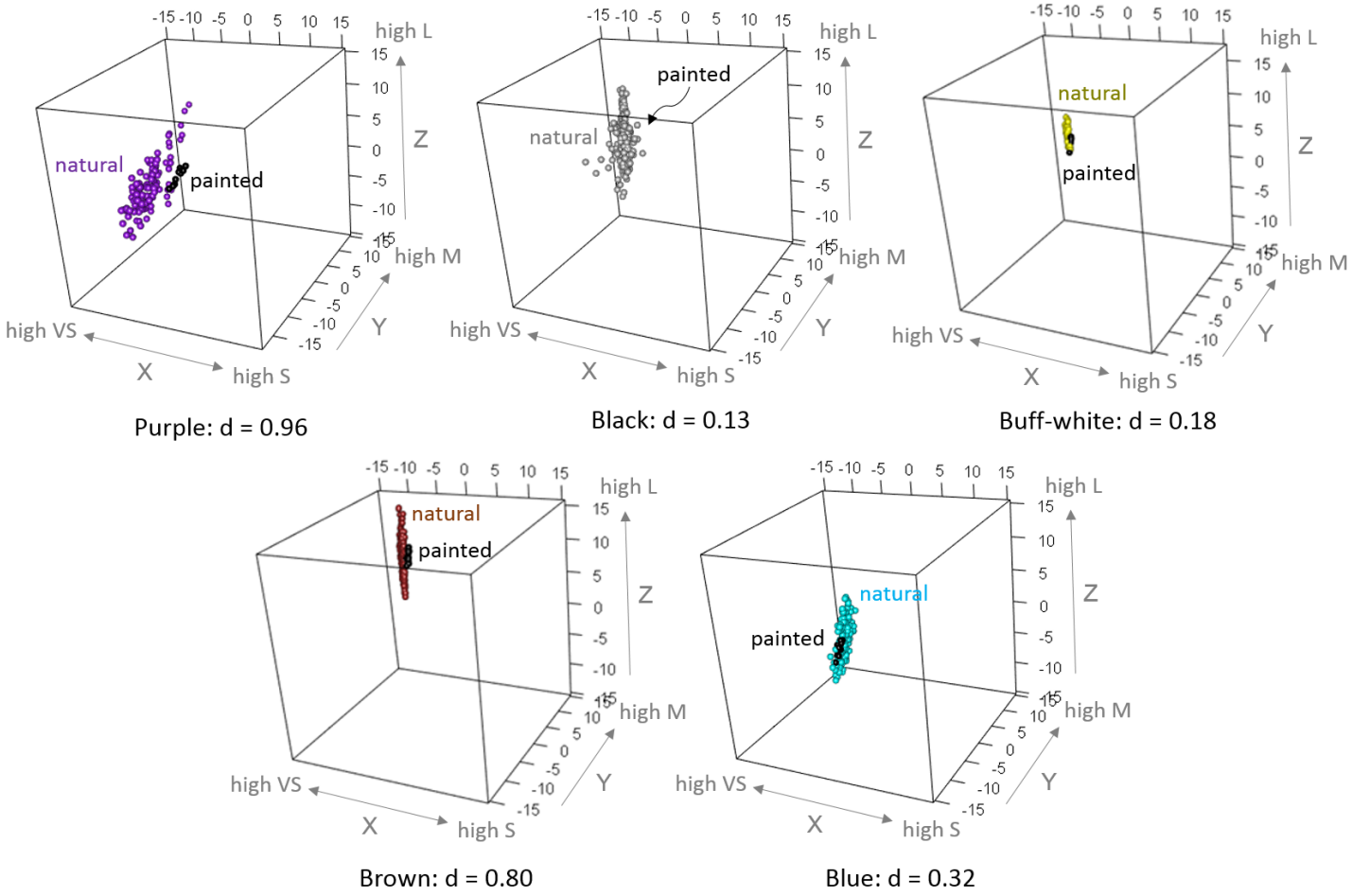


Figure S1. Graphic representation of the methods used to compare natural and painted colours in the avian visual space. In this representation, the X-axis represents stimulation of the VS cone relative to the S cone, higher values of the Y-axis represents higher stimulation of the M cone relative to VS and S cones, while the Z-axis represents higher relative stimulation of the L cone compared with the other three. Axis units are JNDs.

Minimum Euclidean distances (d) in JNDs are indicated for each pair of natural and painted colours and are all below 1: purple (*n* = 114 and 10 measurements, respectively), black (*n* = 270 and 10, respectively), buff-white (*n* = 180 and 15, respectively), brown (*n* = 180 and 20, respectively) and blue (*n* = 174 and 15, respectively).

(b) Details on the experimental protocol

3D model location – The 3D model was visible from anywhere within a 3-m radius from the model, and located > 10 m of any territory boundary to avoid eliciting responses from neighbouring territory occupants. We tried as much as possible to place the models in the same location for the three replicates of each experiment, but it was not always possible because of the location of the birds prior to the experiment (we needed to make sure that they could hear the playback from their initial location and tried to place the model ≤ 7 m of the birds).

Playback – The standardised playback was composed of contact calls and male solo songs recorded from individuals of our population between July 2005 and March 2011 (for details see Hall & Peters, 2008). The complete sequence was as follows: 2-min sequence of intermittent contact calls – male solo song – 2 brief contact calls – male solo song – 2 brief contact calls – male solo song – 2 brief contact calls; lasting for 2 min 38 s in total. The playback was broadcast via a Moshi^{TM} BassBurger speaker (Moki International Ltd., Braeside, Australia) connected to an Apple iPod (Apple Inc., Cupertino, USA) and concealed behind the model.

Non-detection of the 3D model – We considered that a focal individual did not detect the presence of the model when it remained > 7 m from the model, or closer but remaining behind dense vegetation that obviously obstructed the view. Non-detection of the model was independent of model type (Pearson's χ^2 test, $\chi_2^2 = 0.78$, $P = 0.68$).

One experimental replicate conducted in the afternoon – In only one instance the focal individual was not found in the morning (no sign of its presence in the territory) and a replicate was conducted in the afternoon on the same day in similar weather conditions.

Movie S1. Resident breeder male purple-crowned fairy-wren displaying aggressive behaviour towards a simulated male intruder. The video depicts a male responding to the presence of a

3D model of a male purple-crowned fairy-wren in non-breeding plumage (brown model) in his territory – accompanied by playback of conspecific contact calls and songs – by approaching and pecking the model, and performing solo songs.

Appendix S2: Statistical analyses

(a) Early effect of age on the extent of breeding plumage

We built an LMM with the maximum % breeding plumage in a given year as the response variable, age, within-group rank, tarsus length, territory quality and group size as fixed effects, and bird identity, territory identity and year as random intercepts. Only birds whose age was accurately known at the time of pre-breeding moult completion were included in the model (*n* = 129), thus excluding adult birds born prior to the start of the study (founder population). We first included 1) age in months both as a continuous linear and quadratic variable, and then 2) age in years as a cofactor (levels 1, 2, 3, 4 and 5; the only two 6-year-old individuals were excluded from the analysis). Using Tukey's Honest Significant Difference (HSD) test, we performed post hoc comparisons between age classes, which indicated that the maximum % breeding plumage only significantly varied between the first and later years of life (Tables S1-S2).

(b) Acquisition of a breeder position

Absence of pattern in time of breeder position acquisition in the year – Using all the records of a subordinate obtaining a breeder position (either by splitting the natal territory, inheritance or dispersal) between January 2006 and December 2010 (*n* = 81), we examined whether these acquisitions occurred year-round or at any particular time(s) of the year. To do this, we listed for each year the number of acquisitions for each month and built a GLMM with the number of acquisitions as the count response variable, and month as a fixed effect and year as a random

intercept. We then used Tukey's HSD test to perform post hoc comparisons between months, which indicated that the acquisition of a breeder position can occur year-round (Table S4).

Case-by-case analysis – We used all records of a subordinate obtaining a breeder position either by 1) inheritance (while other local subordinate males were present in the territory) or 2) dispersal, and whose percentage of breeding plumage was known when the breeding vacancy became available. For cases of inheritance, we considered that other within-group subordinate male(s) could also compete. For cases of dispersal, by using information on territories (group composition and geographical location), we assessed which other males located within the same distance to the vacancy as the 'winner' could also compete. For both inheritance and dispersal cases, we calculated the difference in % breeding plumage (at the time the breeding vacancy appeared) for all pairs of winners and losers (e.g., for a competition involving 3 individuals A, B and C and where A was the winner, we determined the difference in % breeding plumage both between A and B, and A and C). Similarly, we calculated the difference in age (in months) for all pairs of winners and losers of known age.

As a further analysis, for all cases we determined which loser was the most competitive in terms of % breeding plumage and age (i.e. the most purple-and-black loser and the oldest loser, which were the same individual in competitions with a single loser, but could be different individuals in competitions involving multiple losers), and performed a comparison with the winner, using a paired *t*-test. Cases including individuals of unknown age were entirely excluded from the age comparisons. We found that the winner was significantly more purple-and-black (respectively older) than the most purple-and-black loser (respectively oldest loser) for both inheritance and dispersal cases (winner *vs* most purple-and-black loser: $t_{28} = 2.78$, $P = 0.005$ – including 18 cases where the winner was more purple-and-black, eight cases where the winner and loser were equally purple-and-black and three cases where the winner was less purple-and-black; winner *vs* oldest loser: $t_{20} = 2.36$, $P = 0.01$ – including 12 cases where the

winner was older, six cases where the winner and loser were of the same age and three cases where the winner was younger). Finally, for all cases, we calculated the average % breeding plumage and age of the losers (again excluding entire cases with individuals of unknown age) – defining an 'average loser' in each competition (identical to the loser in competitions with a single loser) – and performed a comparison with the winner, using a paired *t*-test. We found that the winner was significantly more purple-and-black and older than the average loser for both inheritance and dispersal cases (winner *vs* average loser: % breeding plumage, $t_{28} = 4.15$, $P < 0.001$; age, $t_{20} = 3.07$, $P = 0.003$), being consistent with the results reported above and in the results section.

(c) Model presentation experiments: principal component analysis (PCA)

We performed a PCA on the five investigated variables (preliminary centred) using a correlation matrix: 1) duration of response, 2) closest approach, 3) time spent within 3 m of the model, 4) time spent within 1 m, and 5) number of songs. Component loadings, as well as standard deviations and variances, are as follows:

	PC_1	**PC_2**	PC_3	PC_4	PC_5
Duration of response	**0.50**	**0.22**	-0.19	0.72	0.38
Closest approach	**-0.44**	**-0.21**	-0.85	0.19	-0.09
Time ≤ 3 m	**0.47**	**-0.43**	-0.01	0.19	-0.74
Time ≤ 1 m	**0.43**	**-0.57**	-0.24	-0.45	0.49
Number of songs	**0.39**	**0.63**	-0.43	-0.46	-0.23
Standard deviation	**1.84**	**0.97**	0.64	0.39	0.32
Proportion of variance	**0.68**	**0.19**	0.08	0.03	0.02
Cumulative proportion	**0.68**	**0.87**	0.95	0.98	1.00

PC_1 and PC_2 (*sqrt*-transformed) were then each fitted as the response variable in two separate LMMs, with the same fixed and random effects as in the statistical models fitting each of the five investigated variables separately (see material and methods section). The model outputs indicated that PC_1, which has a medium positive loading for duration of response, depended

on 3D model type (likelihood-ratio test, $\chi_2^2 = 12.63$, $P = 0.002$) and was significantly higher when purple-and-black models were used compared to both brown and control models (Table S9c). In contrast, PC_2, which has a quite strong positive loading for number of songs, was not related to 3D model type (likelihood-ratio test, $\chi_2^2 = 1.79$, $P = 0.41$; see also Table S9c), being consistent with the results reported in the results section.

(d) Comparison of aggressiveness of breeder males *vs* females

When testing whether the overall response differed between male and female breeders using the data from both sexes, we found that the overall response of female breeders was much weaker than the males' response (see Table S10). As a result, we had limited power to detect sex differences in baseline aggressiveness and degree of escalation by fitting a sex*3D-printed model type interaction. Consequently, we subsequently focused on females only and used the same models as we did for males (except for time spent within 1 m of the model as in 192 of 203 cases females did not approach that close), but excluding the presence of subordinate males unrelated to either breeder (results provided in Table S11).

(e) Comparison of aggressiveness of breeder males *vs* subordinates

In May-June only, for territories including one or more subordinates, the behaviour of one or two (male or female) subordinates was recorded. We tested whether the overall response differed between breeder males and subordinates by investigating the following variables: 1) duration of response (*sqrt*-transformed), 2) closest approach (*sqrt*-transformed), 3) time spent within 3 m of the model (*log*-transformed), and 4) time spent within 1 m *(log*-transformed). In all models we fitted model type, social status, territory quality, group size, presence of fledglings, replicate number, time of the day and sun strength as fixed effects. Bird identity, territory identity, model exemplar, date and type of substrate were fitted as random intercepts.

LMMs were used for all response variables except number of songs for which we used a negative binomial model. Latency of response was not included as it cannot be defined in a consistent way for both breeders and subordinates since the latter sing much less and do not duet. We did not run the model including number of songs as the response variable since all the values (except for two of 83 observations) for subordinates were equal to zero. The model outputs indicated that the response of subordinates was generally similar or weaker compared to breeder males (results provided in Table S12). The data available on subordinates were insufficient to test for an effect of 3D model type on their aggressiveness.

(f) No variation in the effect of territory quality on aggressiveness with 3D model type

As we found a significant effect of territory quality in all the aggressiveness-related analyses, we tested whether this varied with the 3D model type used by including the 3D model type*territory quality interaction in all statistical models (except number of songs as the model could not converge). The model outputs indicated that the effect of territory quality was independent of the 3D model type used as the interaction term was non-significant (Table S13).

(g) Physical aggression towards the 3D model

Three of 46 males physically attacked the 3D models: 1) a male with 0% breeding plumage in one season and 100% in the other attacked both the control and purple-and-black models in both seasons; 2) a male with 2% breeding plumage attacked the purple-and-black model; and 3) a male, respectively with 25% and then 35% breeding plumage (i.e. moulting into the breeding plumage), successively attacked the control and brown models. These observations provide further support for a lack of effect of the extent of breeding plumage of resident breeder males on the level of aggression they display.

(h) Absence of effect of population density on aggressiveness

We tested whether the significant effect of territory quality was related to population density. Indeed, territories of relatively high quality (average territory quality > 10 out of 20, based on *Pandanus* cover) are spatially grouped along Annie creek where the average population density is 45 individuals.km^{-1}, whereas many low-quality territories (territory quality < 10) are grouped along the Adcock river where the population density is substantially lower, being on average 7 individuals.km^{-1} with in multiple cases gaps between territories with (suitable) vegetation. To do this, we determined for each focal individual in each season the number of direct neighbours as a proxy for population density (range: 0-16), and added it as a fixed effect in all the analyses (except number of songs as the model could not converge; Table S14). The model outputs indicated that (i) the number of direct neighbours had no effect on aggressiveness whereas the effect of territory quality remained (and all other effects were quantitatively similar; Table S14), (ii) territory quality and number of direct neighbours were not highly correlated ($|r|$ = 0.50), and the comparison of the models with and without the number of direct neighbours indicated that (iii) they had similar support (likelihood-ratio test, duration: $\chi_1^2 = 1.08$, $P = 0.30$; closest approach: $\chi_1^2 = 1.22$, $P = 0.27$; time within 3 m: $\chi_1^2 = 0.95$, $P = 0.33$; time within 1 m: $\chi_1^2 = 1.98$, $P = 0.16$).

References

Cuthill, I. C. (2006). Color perception. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 3-40). Cambridge, MA: Harvard University Press.

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.

Endler, J. A., & Mielke, P. W. (2005). Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, **86**, 405-431.

Hall, M. L., & Peters, A. (2008). Coordination between the sexes for territorial defence in a duetting fairy-wren. *Animal Behaviour*, **76**, 65-73.

Hart, N. S. (2001). Variations in cone photoreceptor abundance and the visual ecology of birds. *Journal of Comparative Physiology A*, **187**, 685-697.

Hart, N. S., & Hunt, D. M. (2007). Avian visual pigments: characteristics, spectral tuning, and evolution. *The American Naturalist*, **169**, S7-S26.

Lind, O., Chavez, J., & Kelber, A. (2014). The contribution of single and double cones to spectral sensitivity in budgerigars during changing light conditions. *Journal of Comparative Physiology A*, **200**, 197-207.

Ödeen, A., Pruett-Jones, S., Driskell, A. C., Armenta, J. K., & Håstad, O. (2012). Multiple shifts between violet and ultraviolet vision in a family of passerine birds with associated changes in plumage coloration. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 1269-1276.

Vorobyev, M., Osorio, D., Bennett, A. T., Marshall, N. J., & Cuthill, I. C. (1998). Tetrachromacy, oil droplets and bird plumage colours. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, **183**, 621-633.

Table S1. Males that are older both at the population and social group levels produce a more complete breeding plumage. Shown are results from LMMs examining the effects of age, within-group rank, tarsus length (in mm), territory quality and group size on the maximum % breeding plumage in individuals of known age only, considering (a) age in months as a covariate (n = 127) and (b) age in years as a cofactor (levels 1, 2, 3, 4, 5; n = 24, 41, 32, 18, 10, respectively; reference: *age 1). Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2): (a) / (b) individual = 4.63 / 5.26, territory = 13.49 / 23.76, year = 2.48 / 0.00, residual = 291.14 / 225.30.

(a)

Fixed effects	β	SE	df	t	P
Intercept	96.48	2.32	4	41.53	<0.001
Age	**1.05**	**0.14**	**40**	**7.44**	**<0.001**
Age^2	**-0.05**	**0.01**	**78**	**-6.41**	**<0.001**
Within-group rank	**-10.63**	**2.57**	**81**	**-4.13**	**<0.001**
Tarsus length	0.01	2.24	22	0.004	1.00
Territory quality	0.58	0.37	24	1.56	0.13
Group size	-0.21	1.25	81	-0.17	0.87
Marginal / conditional R^2 = 0.58 / 0.61					

(b)

Fixed effects	β	SE	df	t	P
Intercept	53.57	3.52	107	12.22	<0.001
Age*					
β_{age2}	**41.44**	**4.18**	**104**	**9.91**	**<0.001**
β_{age3}	**43.55**	**4.55**	**109**	**9.57**	**<0.001**
β_{age4}	**40.36**	**5.38**	**112**	**7.50**	**<0.001**
β_{age5}	**42.31**	**6.38**	**110**	**6.63**	**<0.001**
Within-group rank	**-9.14**	**2.39**	**84**	**-3.81**	**<0.001**
Tarsus length	-0.49	2.08	15	-0.24	0.82
Territory quality	0.68	0.36	21	1.89	0.07
Group size	-0.26	1.13	113	-0.23	0.82
Marginal / conditional R^2 = 0.67 / 0.70					

Table S2. 2-year-old and older males produce a more complete breeding plumage than 1-year-old males. Shown are results from Tukey's HSD test for post hoc comparisons between age classes (1-5 years old; n = 125 comparisons). Significant values ($P < 0.05$) are in bold.

Pairwise comparison	β	SE	z	P
2-1	**41.44**	**4.18**	**9.91**	**<0.001**
3-1	**43.55**	**4.55**	**9.57**	**<0.001**
4-1	**40.36**	**5.38**	**7.50**	**<0.001**
5-1	**42.31**	**6.38**	**6.63**	**<0.001**
3-2	2.12	3.68	0.57	0.98
4-2	-1.08	4.54	-0.24	1.00
5-2	0.87	5.66	0.15	1.00
4-3	-3.20	4.58	-0.70	0.96
5-3	-1.24	5.70	-0.22	1.00
5-4	1.95	6.06	0.32	1.00

Table S3. Within social groups, breeder males produce a more complete plumage than subordinate males, and among subordinate males, those that are relatively older, and therefore higher-ranked in the social hierarchy, produce a more complete breeding plumage. Moreover, breeding plumage in older males (≥ 2 years old) is less affected by within-group rank than in one-year-old males. Shown are results from an LMM examining the effects of the interaction between age (levels 1, 2+; reference: *age 1) and within-group rank, tarsus length (in mm), territory quality and group size on the maximum % breeding plumage in individuals of known age and individuals of unknown age at least in their second year of life (n = 247). Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2): individual = 29.48, territory = 8.81, year = 0.00, residual = 94.59.

Fixed effects	β	SE	df	t	P
Intercept	59.32	2.77	235	21.39	<0.001
Age*	**38.77**	**2.82**	**225**	**13.73**	**<0.001**
Within-group rank	**-12.19**	**1.64**	**232**	**-7.45**	**<0.001**
Age* × Within-group rank	**5.93**	**1.93**	**200**	**3.08**	**0.002**
Tarsus length	0.92	1.36	27	0.68	0.50
Territory quality	0.30	0.20	19	1.48	0.16
Group size	-0.05	0.48	208	-0.10	0.92
Marginal / conditional R^2 = 0.69 / 0.78					

Table S4. The acquisition of a breeder position can occur year-round. Shown are results from (a) a GLMM examining the effect of month on the number of acquisitions (reference: *January), and (b) Tukey's HSD test for post hoc comparisons between months. Significant values ($P < 0.05$) are in bold. Variance of the random effect 'year': $\sigma^2 = 0.01$.

(a)

Fixed effects	β	SE	z	P
Intercept	0.78	0.31	2.56	0.01
Month*				
β_{Feb}	-1.01	0.58	-1.73	0.08
β_{Mar}	-0.20	0.45	-0.45	0.66
β_{Apr}	-0.20	0.45	-0.45	0.66
β_{May}	-0.45	0.48	-0.94	0.35
β_{Jun}	**-1.30**	**0.65**	**-2.00**	**0.05**
β_{Jul}	-0.45	0.48	-0.94	0.35
β_{Aug}	-0.45	0.48	-0.94	0.35
β_{Sep}	-0.32	0.46	-0.69	0.49
β_{Oct}	-1.01	0.58	-1.73	0.08
β_{Nov}	-1.01	0.58	-1.73	0.08
β_{Dec}	-0.45	0.48	-0.94	0.35
Marginal / conditional $R^2 = 0.20 / 0.22$				

(b)

Pairwise comparison	β	SE	z	P
Feb-Jan	-1.01	0.58	-1.73	0.85
Mar-Jan	-0.20	0.45	-0.45	1.00
Apr-Jan	-0.20	0.45	-0.45	1.00
May-Jan	-0.45	0.48	-0.94	1.00
Jun-Jan	-1.30	0.65	-2.00	0.69
Jul-Jan	-0.45	0.48	-0.94	1.00
Aug-Jan	-0.45	0.48	-0.94	1.00
Sep-Jan	-0.32	0.46	-0.69	1.00
Oct-Jan	-1.01	0.58	-1.73	0.85
Nov-Jan	-1.01	0.58	-1.73	0.85
Dec-Jan	-0.45	0.48	-0.94	1.00
Mar-Feb	0.81	0.60	1.35	0.97
Apr-Feb	0.81	0.60	1.35	0.97

May-Feb	0.56	0.63	0.89	1.00
Jun-Feb	-0.29	0.76	-0.38	1.00
Jul-Feb	0.56	0.63	0.89	1.00
Aug-Feb	0.56	0.63	0.89	0.99
Sep-Feb	0.69	0.61	1.13	1.00
Oct-Feb	3.20×10^{-5}	0.71	0.00	1.00
Nov-Feb	7.52×10^{-5}	0.71	0.00	0.89
Dec-Feb	0.56	0.63	0.89	1.00
Apr-Mar	-4.30×10^{-5}	0.47	0.00	1.00
May-Mar	-0.25	0.50	-0.50	1.00
Jun-Mar	-1.10	0.67	-1.65	0.97
Jul-Mar	-0.25	0.50	-0.50	0.97
Aug-Mar	-0.25	0.50	-0.50	1.00
Sep-Mar	-0.12	0.49	-0.24	1.00
Oct-Mar	-0.81	0.60	-1.35	0.89
Nov-Mar	-0.81	0.60	-1.35	1.00
Dec-Mar	-0.25	0.50	-0.50	1.00
May-Apr	-0.25	0.50	-0.50	1.00
Jun-Apr	-1.10	0.67	-1.65	0.89
Jul-Apr	-0.25	0.50	-0.50	1.00
Aug-Apr	-0.25	0.50	-0.50	1.00
Sep-Apr	-.1.18	0.49	-0.24	1.00
Oct-Apr	-0.81	0.60	-1.35	0.97
Nov-Apr	-0.81	0.60	-1.35	0.97
Dec-Apr	-0.25	0.50	-0.50	1.00
Jun-May	-0.85	-.69	-1.23	0.99
Jul-May	-2.58×10^{-5}	-.53	0.00	1.00
Aug-May	-6.42×10^{-5}	0.53	0.00	1.00
Sep-May	0.13	0.52	0.26	1.00
Oct-May	-0.56	0.63	-0.89	1.00
Nov-May	-0.56	0.63	-0.89	1.00
Dec-May	-6.41×10^{-5}	0.53	0.00	1.00
Jul-Jun	0.85	0.69	1.23	0.99
Aug-Jun	0.85	0.69	1.23	0.99
Sep-Jun	0.98	0.76	1.45	0.95
Oct-Jun	0.29	0.76	0.38	1.00
Nov-Jun	0.29	0.69	0.38	1.00

Dec-Jun	0.85	0.69	1.23	0.99
Aug-Jul	-3.85×10^{-5}	0.53	0.00	1.00
Sep-Jul	0.13	0.52	0.26	1.00
Oct-Jul	-0.56	0.63	-0.89	1.00
Nov-Jul	-0.56	0.63	-0.89	1.00
Dec-Jul	-3.84×10^{-5}	0.53	0.00	1.00
Sep-Aug	0.13	0.52	0.26	1.00
Oct-Aug	-0.56	0.63	-0.89	1.00
Nov-Aug	-0.56	0.63	-0.89	1.00
Dec-Aug	1.14×10^{-7}	0.53	0.00	1.00
Oct-Sep	-0.69	0.61	-1.13	0.99
Nov-Sep	-0.69	0.61	-1.13	0.99
Dec-Sep	-0.13	0.52	-0.26	1.00
Nov-Oct	4.32×10^{-5}	0.71	0.00	1.00
Dec-Oct	0.56	0.63	0.89	1.00
Dec-Nov	0.56	0.63	0.89	1.00

Table S5. At the population level, subordinate males that produce a more complete breeding plumage increase their chances to gain a breeder position by dispersal within that year. Shown are results from GLMMPQLs examining the effects of maximum % breeding plumage, age (reference: *age 1) and tarsus length (in mm) on the likelihood of gaining a breeder position among subordinates, including or excluding individuals that died as subordinates before the end of the considered year. Due to high correlation between age and maximum % breeding plumage, two models were built to fit (a) maximum % breeding plumage and (b) age separately. Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2): (a) all males / males still alive at the end of the year only: year = 4.08×10^{-5} / 1.19×10^{-15}, territory in year = 1.30×10^{-6} / 6.00×10^{-15}, individual in territory in year = 7.96 / 9.00, residual = 5.86×10^{-5} / 3.24×10^{-15}; (b) all males / males still alive at the end of the year only: year = 1.18×10^{-5} / 1.28×10^{-12}, territory in year = 7.60×10^{-5} / 8.33×10^{-12}, individual in territory in year = 7.70 / 8.51, residual = 4.01×10^{-5} / 4.17×10^{-12}.

(a)

	All males (n = 53)					Males still alive at the end of the year only (n = 39)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P
Intercept	-8.14	1.14	32	-7.13	<0.001	-6.47	1.50	22	-4.31	<0.001
Maximum % breeding plumage	**0.11**	**0.03**	**14**	**3.07**	**0.008**	**0.14**	**0.04**	**10**	**3.22**	**0.009**
Tarsus length	1.42	1.70	14	0.83	0.42	1.44	2.31	10	0.63	0.55
Marginal / conditional R^2	0.19 / 1.00					0.19 / 1.00				

(b)

	All males (n = 53)					Males still alive at the end of the year only (n = 39)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P
Intercept	-9.52	2.06	32	-4.61	<0.001	-8.96	2.84	22	-3.16	0.005
Age*	3.90	2.45	14	1.59	0.13	4.81	3.31	10	1.45	0.18

Tarsus length	1.37	1.64	14	0.83	0.42	1.56	2.20	10	0.71	0.49
Marginal / conditional R^2	0.07 / 1.00					0.08 / 1.00				

Table S6. Subordinate males that are older and/or display a more complete breeding plumage have higher chances to inherit a breeder vacancy or acquire it elsewhere (when a breeder without a successor dies or moves away). Shown are results from LMMs examining whether subordinate males that gain a breeder position display a more complete breeding plumage and/or are older than their competitors, controlling for the route used to gain the position (inheritance or dispersal; reference: *dispersal). Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2): difference in % breeding plumage / age: winner = 873.63 / 123.45, residual = 692.26 / 23.58.

Winner *vs* loser	Difference in % breeding plumage ($n = 61$)					Difference in age (in months, $n = 34$)				
Fixed effects	β	SE	df	T	P	β	SE	df	t	P
Intercept	**26.17**	**7.66**	**24**	**3.42**	**0.002**	**9.43**	**2.97**	**22**	**3.18**	**0.004**
Route*	6.85	16.30	30	0.42	0.68	-4.89	5.41	22	-0.90	0.38
Marginal / conditional R^2	0.004 / 0.56					0.03 / 0.84				

Table S7. Breeder males respond for longer, approach closer and spend more time within 3 m of purple-and-black models compared to brown and control models. Shown are results for (a) fixed effects and (b) random effects from LMMs examining the effects of 3D model type (reference: *control), latency of response (in s), age (in years), territory quality, group size, presence of fledglings, presence of subordinate males unrelated to the breeder male, presence of subordinate males unrelated to the breeder female, season (reference: †start of the dry season), replicate number, time of the day and sun strength on duration of response, closest approach and time spent within 3 m of the model (n = 216) , and (c) from Tukey's HSD test for post hoc pairwise comparisons between 3D model types (as P-values in (a) were obtained using Satterthwaite approximations of degrees of freedom and not a multiple comparison procedure). Significant values ($P < 0.05$) are in bold. Response variables were either [a]*sqrt*- or [b]*log*-transformed.

(a)

	Duration of response[a] (s)					Closest approach[a] (m)					Time spent ≤ 3 m[b] (s)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P	β	SE	df	t	P
Intercept	15.00	1.22	51	12.34	<0.001	1.73	0.13	8	13.27	<0.001	1.30	0.46	75	2.83	<0.001
3D model type*															
β_{brown}	0.44	0.61	63	0.72	0.48	-0.10	0.08	1	-1.26	0.44	**0.55**	**0.26**	**58**	**2.15**	**0.04**
$\beta_{purple\text{-}and\text{-}black}$	**2.49**	**0.74**	**39**	**3.38**	**0.002**	-0.22	0.07	1	-3.02	0.23	**1.14**	**0.25**	**48**	**4.63**	**<0.001**
Latency of response	**-0.06**	**6.98×10^{-3}**	**183**	**-8.40**	**<0.001**	**1.99×10^{-3}**	**6.53×10^{-4}**	**169**	**3.05**	**0.003**	**-7.50×10^{-3}**	**2.64×10^{-3}**	**174**	**-2.84**	**0.005**
Age	0.14	0.22	39	0.61	0.54	1.33×10^{-2}	0.02	40	0.54	0.59	-0.07	0.08	42	-0.89	0.38
Territory quality	**-0.49**	**0.11**	**45**	**-4.36**	**<0.001**	**0.05**	**0.01**	**45**	**3.86**	**<0.001**	**-0.12**	**0.04**	**47**	**-2.82**	**0.007**
Group size	-0.20	0.47	95	-0.44	0.66	-0.04	0.05	107	-0.81	0.42	-0.03	0.17	91	-0.20	0.84
Presence of fledglings	1.20	0.91	171	1.31	0.19	**-0.23**	**0.09**	**175**	**-2.52**	**0.01**	0.60	0.36	176	1.69	0.09

Presence of subordinate ♂ unrelated to breeder ♂	-0.93	1.82	71	-0.51	0.61	0.13	0.19	88	0.67	0.51	-0.42	0.65	67	-0.64	0.53
Presence of subordinate ♂ unrelated to breeder ♀	-0.55	1.28	110	-0.43	0.67	0.05	0.13	122	0.41	0.68	-0.03	0.48	104	-0.07	0.95
Season[†]	**-2.45**	**0.61**	**154**	**-4.03**	**<0.001**	**0.17**	**0.06**	**151**	**2.78**	**0.006**	-0.11	0.24	171	-0.48	0.64
Replicate number	-0.45	0.31	160	-1.47	0.14	-9.27×10^{-3}	0.03	159	-0.33	0.74	0.03	0.12	163	0.26	0.79
Time of the day	-0.12	0.20	162	-0.62	0.53	-0.01	0.02	167	-0.62	0.54	0.13	0.08	175	1.74	0.08
Sun strength	**-0.76**	**0.31**	**160**	**-2.46**	**0.02**	7.90×10^{-3}	0.03	163	0.27	0.79	-0.02	0.12	165	-0.17	0.87

(b)

Random effects	Duration of response[a]	Closest approach[a]	Time spent ≤ 3 m[b]
$\sigma^2_{individual}$			
Intercept	11.41	0.14	1.03
Brown	1.76	0.03	0.55
Purple-and-black	8.56	0.02	0.32
$\sigma^2_{model\ exemplar}$	0.00	3.78×10^{-3}	0.00
$\sigma^2_{substrate\ type}$	1.01	0.01	0.14
σ^2_{date}	0.00	0.00	0.00
$\sigma^2_{residual}$	11.34	0.10	1.81
Marginal / conditional R^2	0.50 / 0.78	0.31 / 0.75	0.24 / 0.64

(c)

	Duration of response[a]			Closest approach[a]			Time spent ≤ 3 m[b]		
Pairwise comparison	$\beta\pm$SE	*z*	*P*	$\beta\pm$SE	*z*	*P*	$\beta\pm$SE	*z*	*P*
brown *vs* control	0.44±0.61	0.72	0.75	-0.10±0.08	-1.26	0.42	0.55±0.26	2.15	0.08
purple-and-black *vs* control	**2.49±0.74**	**3.38**	**0.002**	**-0.22±0.07**	**-3.02**	**0.007**	**1.14±0.25**	**4.63**	**<0.001**
purple-and-black *vs* brown	**2.05±0.63**	**3.26**	**0.003**	-0.13±0.06	-2.17	0.07	**0.59±0.23**	**2.53**	**0.03**

Table S8. Breeder males neither spend more time within 1 m nor sing more in presence of purple-and-black models compared to brown and control models. Shown are results for (a) fixed effects and (b) random effects from statistical models examining the effects of 3D model type (reference: *control), latency of response (in s), age (in years), territory quality, group size, presence of fledglings, presence of subordinate males unrelated to the breeder male, presence of subordinate males unrelated to the breeder female, season (reference: †start of the dry season), replicate number, time of the day and sun strength on (*log*-transformed) time spent within 1 m of the model and number of songs (n = 216) , and (c) from Tukey's HSD test for post hoc pairwise comparisons between 3D model types (as P-values in (a) were obtained using Satterthwaite approximations of degrees of freedom and not a multiple comparison procedure). Significant values ($P < 0.05$) are in bold.

(a)

	Time spent ≤ 1 m (s) (LMM)					Number of songs (negative binomial model)			
Fixed effects	β	SE	df	t	P	β	SE	z	P
Intercept	0.68	0.38	17	1.81	0.09	1.51	0.17	9.12	<0.001
3D model type*									
β_{brown}	0.12	0.28	4	0.42	0.70	-0.13	0.09	-1.54	0.12
$\beta_{purple\text{-}and\text{-}black}$	0.37	0.28	5	1.32	0.25	-0.05	0.09	-0.52	0.60
Latency of response	-1.15×10^{-3}	1.70×10^{-3}	179	-0.68	0.50	$\mathbf{-6.07\times10^{-3}}$	$\mathbf{1.15\times10^{-3}}$	**-5.27**	**<0.001**
Age	0.04	0.06	43	0.57	0.57	0.02	0.02	0.99	0.32
Territory quality	**-0.12**	**0.03**	**48**	**-3.92**	**<0.001**	**-0.05**	**0.01**	**-4.03**	**<0.001**
Group size	0.19	0.13	112	1.49	0.14	-0.02	0.06	-0.42	0.68
Presence of fledglings	0.41	0.24	189	1.76	0.08	0.23	0.12	1.92	0.05
Presence of subordinate ♂ unrelated to breeder ♂	-0.38	0.47	88	-0.79	0.43	-0.02	0.22	-0.09	0.93

Presence of subordinate ♂ unrelated to breeder ♀	-0.38	0.34	128	-1.14	0.26	-0.24	0.16	-1.46	0.14
Season[†]	**-0.31**	**0.15**	**155**	**-2.01**	**0.05**	-0.11	0.09	-1.30	0.19
Replicate number	1.14×10^{-4}	0.07	160	0.002	1.00	0.06	0.05	1.19	0.23
Time of the day	0.05	0.05	177	1.05	0.30	-0.04	0.03	-1.54	0.12
Sun strength	-0.02	0.08	166	-0.21	0.83	**-0.12**	**0.05**	**-2.73**	**0.006**

(b)

Random effects	Time spent ≤ 1 m	Number of songs
$\sigma^2_{individual}$		
Intercept	0.80	0.07
Brown	0.03	2.31×10^{-5}
Purple-and-black	0.07	1.13×10^{-7}
$\sigma^2_{model\ exemplar}$	0.10	1.15×10^{-7}
$\sigma^2_{substrate\ type}$	0.03	0.02
σ^2_{date}	0.00	4.45×10^{-3}
$\sigma^2_{residual}$	0.69	-
Marginal / conditional R^2	0.23 / 0.72	-

(c)

	Time spent ≤ 1 m			Number of songs		
Pairwise comparison	$\beta\pm$SE	z	P	$\beta\pm$SE	z	P
brown *vs* control	0.12±0.28	0.42	0.91	-0.13±0.09	-1.54	0.26
purple-and-black *vs* control	0.37±0.28	1.32	0.38	-0.05±0.09	-0.52	0.85
purple-and-black *vs* brown	0.25±0.18	1.41	0.33	0.09±0.12	0.72	0.74

Table S9. The higher level of aggression displayed by breeder males towards purple-and-black models compared to brown and control models is expressed through the duration of response, but not through the number of songs performed. Shown are results for (a) fixed effects and (b) random effects from LMMs examining the effects of 3D model type (reference: *control), latency of response (in s), age (in years), territory quality, group size, presence of fledglings, presence of subordinate males unrelated to the breeder male, presence of subordinate males unrelated to the breeder female, season (reference: †start of the dry season), replicate number, time of the day and sun strength on the (*sqrt*-transformed) principal components PC_1 and PC_2 (n = 216), and (c) Tukey's HSD test for post hoc pairwise comparisons between 3D model types. Significant values (P < 0.05) are in bold.

(a)

	PC_1					PC_2				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P
Intercept	1.71	0.08	41	22.33	<0.001	2.41	0.08	10	29.74	<0.001
3D model type*										
β_{brown}	0.04	0.04	75	1.18	0.24	-0.05	0.07	4	-0.70	0.52
$\beta_{purple\text{-}and\text{-}black}$	**0.16**	**0.04**	**45**	**4.21**	**<0.001**	-0.04	0.07	5	-0.60	0.57
Latency of response	**-1.93×10^{-3}**	**4.08×10^{-4}**	**169**	**-4.72**	**<0.001**	**-1.30×10^{-3}**	**3.50×10^{-4}**	**169**	**-3.72**	**<0.001**
Age	0.02	0.02	42	1.03	0.31	-2.84×10^{-3}	7.40×10^{-3}	55	-0.38	0.70
Territory quality	**-0.03**	**7.53×10^{-3}**	**46**	**-4.56**	**<0.001**	-3.47×10^{-3}	3.70×10^{-3}	67	-0.94	0.35
Group size	0.01	0.03	108	0.36	0.72	-3.64×10^{-4}	0.02	93	-0.02	0.98
Presence of fledglings	**0.13**	**0.06**	**180**	**2.31**	**0.02**	0.05	0.04	107	1.25	0.22
Presence of subordinate ♂ unrelated to breeder ♂	-0.08	0.11	78	-0.73	0.47	0.03	0.06	67	0.50	0.62

Presence of subordinate ♂ unrelated to breeder ♀	-0.03	0.08	125	-0.43	0.67	-0.02	0.05	93	-0.37	0.72
Season†	**-0.15**	**0.04**	**160**	**-4.17**	**<0.001**	6.16×10^{-3}	0.03	37	0.20	0.85
Replicate number	1.96×10^{-3}	0.02	159	0.11	0.91	5.35×10^{-3}	0.02	33	0.29	0.77
Time of the day	6.19×10^{-3}	0.01	165	0.52	0.60	-6.73×10^{-3}	0.01	159	-0.67	0.50
Sun strength	-0.02	0.02	164	-1.15	0.25	**-0.05**	**0.02**	**76**	**-3.07**	**0.003**

(b)

Random effects	PC_1	PC_2
$\sigma^2_{individual}$		
Intercept	0.04	2.89×10^{-3}
Brown	7.00×10^{-3}	0.01
Purple-and-black	0.01	0.03
$\sigma^2_{model\ exemplar}$	0.00	5.98×10^{-3}
$\sigma^2_{substrate\ type}$	6.37×10^{-3}	0.00
σ^2_{date}	0.00	1.87×10^{-3}
$\sigma^2_{residual}$	0.04	0.03
Marginal / conditional R^2	0.40 / 0.78	0.13 / 0.53

(c)

	PC_1			PC_2		
Pairwise comparison	$\beta\pm$SE	*z*	*P*	$\beta\pm$SE	*z*	*P*
brown *vs* control	0.04±0.04	1.18	0.46	-0.05±0.07	-0.70	0.76
purple-and-black *vs* control	**0.16±0.04**	**4.21**	**<0.001**	-0.04±0.07	-0.60	0.82
purple-and-black *vs* brown	**0.12±0.04**	**3.42**	**0.002**	0.005±0.05	0.11	0.99

Table S10. Breeder females display a weaker response to simulated intrusions compared to breeder males. Shown are results from statistical models examining the effects of 3D model type (reference: *control), latency of response (in s), sex (reference: •female), territory quality, group size, presence of fledglings, season (reference: †start of the dry season), replicate number, time of the day and sun strength on (a) duration of response, closest approach and time spent within 3 m of the model, and (b) time spent within 1 m of the model and number of songs. Variance of random effects and marginal / conditional R^2 for all models are presented in (c). Significant values ($P < 0.05$) are in bold. Response variables were either [a]*sqrt*- or [b]*log*-transformed (except number of songs).

(a)

	Duration of response[a] (s) (LMM, n = 418)					Closest approach[a] (m) (LMM, n = 420)					Time spent ≤ 3 m[b] (s) (LMM, n = 420)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P	β	SE	df	t	P
Intercept	12.77	1.72	33	7.43	<0.001	1.86	0.13	28	13.79	<0.001	0.90	0.50	22	1.80	0.09
3D model type*															
β_{brown}	0.59	1.26	13	0.47	0.64	-0.07	0.10	9	-0.77	0.46	0.30	0.33	5	0.91	0.40
$\beta_{purple\text{-}and\text{-}black}$	2.31	1.26	13	1.83	0.09	-0.16	0.10	9	-1.70	0.12	0.79	0.33	5	2.38	0.07
Latency of response	**-0.05**	**4.34×10^{-3}**	**375**	**-10.83**	**<0.001**	**1.22×10^{-3}**	**4.03×10^{-4}**	**344**	**3.04**	**0.003**	**-5.36×10^{-3}**	**1.71×10^{-3}**	**324**	**-3.13**	**0.002**
Sex•	**0.89**	**0.32**	**277**	**2.80**	**0.006**	**-0.18**	**0.06**	**45**	**-3.19**	**0.003**	**0.45**	**0.18**	**45**	**2.56**	**0.01**
Territory quality	**-0.54**	**0.12**	**49**	**-4.67**	**<0.001**	**0.03**	**8.91×10^{-3}**	**49**	**3.41**	**0.001**	**-0.08**	**0.04**	**49**	**-2.18**	**0.03**
Group size	-0.05	0.39	147	-0.14	0.89	2.91×10^{-3}	0.03	117	0.09	0.92	-0.13	0.14	111	-0.99	0.33
Presence of fledglings	1.24	0.72	355	1.71	0.09	**-0.15**	**0.06**	**317**	**-2.41**	**0.02**	0.39	0.27	318	1.45	0.15
Season†	**-2.40**	**0.59**	**28**	**-4.09**	**<0.001**	**0.20**	**0.05**	**28**	**3.83**	**<0.001**	-0.26	0.20	34	-1.26	0.22
Replicate number	-0.57	0.31	25	-1.86	0.08	0.02	0.03	24	0.77	0.45	0.08	0.11	26	0.74	0.46
Time of the day	0.01	0.16	258	0.08	0.94	-0.02	0.01	239	-1.65	0.10	**0.17**	**0.06**	**233**	**2.87**	**0.004**

Sun strength	-0.44	0.29	54	-1.54	0.13	-3.60×10^{-3}	0.03	52	-0.14	0.89	-0.03	0.10	56	-0.32	0.75

(b)

	Time spent ≤ 1 m[b] (s) (LMM, *n* = 420)					Number of songs (negative binomial model, *n* = 420)			
Fixed effects	*β*	SE	df	*t*	*P*	*β*	SE	*z*	*P*
Intercept	0.04	0.42	32	0.09	0.93	1.10	0.14	7.85	<0.001
3D model type*									
β_{brown}	0.07	0.37	18	0.20	0.84	-0.07	0.07	-0.97	0.33
$\beta_{purple\text{-}and\text{-}black}$	0.25	0.37	18	0.67	0.51	0.02	0.07	0.28	0.78
Latency of response	4.18×10^{-4}	1.00×10^{-3}	358	0.42	0.68	**-5.89×10^{-3}**	**7.93×10^{-3}**	**-7.43**	**<0.001**
Sex•	**0.43**	**0.15**	**50**	**2.82**	**0.007**	**0.33**	**0.05**	**6.08**	**<0.001**
Territory quality	**-0.09**	**0.02**	**47**	**-3.92**	**<0.001**	**-0.04**	**0.01**	**-3.61**	**<0.001**
Group size	0.05	0.08	117	0.61	0.54	-0.06	0.05	-1.25	0.21
Presence of fledglings	0.06	0.16	315	0.37	0.71	**0.19**	**0.09**	**2.07**	**0.04**
Season[†]	-0.19	0.12	32	-1.57	0.13	**-0.20**	**0.07**	**-2.65**	**0.008**
Replicate number	-0.01	0.06	28	-0.21	0.83	3.94×10^{-3}	0.04	0.10	0.92
Time of the day	0.02	0.03	25	0.69	0.49	-0.04	0.02	-1.82	0.07
Sun strength	0.05	0.06	60	0.76	0.45	-0.07	0.04	-1.92	0.06

(c)

Random effects	Duration of response[a]	Closest approach[a]	Time spent ≤ 3 m[b]	Time spent ≤ 1 m[b]	Number of songs
$\sigma^2_{individual}$	0.00	0.05	0.31	0.42	6.95×10^{-7}
$\sigma^2_{model\ exemplar}$	2.55	0.01	0.15	0.24	1.13×10^{-7}
$\sigma^2_{substrate\ type}$	2.86	0.01	0.12	0.03	7.05×10^{-3}
σ^2_{date}	1.42	9.63×10^{-3}	0.12	0.05	0.01
$\sigma^2_{territory}$	14.47	0.06	1.25	0.29	0.08

$\sigma^2_{residual}$	10.05	0.10	1.76	0.51	-
Marginal / conditional R^2	0.43 / 0.82	0.24 / 0.70	0.18 / 0.61	0.16 / 0.72	-

Table S11. Breeder females' response to the presence of simulated male intruders is independent of intruders' plumage state. Shown are results from statistical models examining effects of 3D model type (reference: *control), latency of response (in s), age (in years), territory quality, group size, presence of fledglings, season (reference: †start of the dry season), replicate number, time of the day and sun strength on (a) duration of response and closest approach, and (b) time spent within 3 m of the model and number of songs (and variance of random effects and marginal / conditional R^2 for all models are presented in (c); n =203), and (d) Tukey's HSD test for post hoc pairwise comparisons between 3D model types. Significant values ($P < 0.05$) are in bold. Response variables were either [a]*sqrt*- or [b]*log*-transformed (except number of songs).

(a)

	Duration of response[a] (s) (LMM)					Closest approach[a] (m) (LMM)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P
Intercept	12.99	1.42	19	9.16	<0.001	1.76	0.10	73	17.69	<0.001
3D model type*										
β_{brown}	0.86	0.98	4	0.88	0.43	-0.06	0.06	63	-1.11	0.27
$\beta_{purple\text{-}and\text{-}black}$	1.97	1.10	6	1.79	0.12	-0.12	0.06	53	-1.90	0.06
Latency of response	**-0.05**	**5.88×10^{-3}**	**145**	**-7.68**	**<0.001**	7.23×10^{-4}	4.56×10^{-4}	159	1.59	0.11
Age	-0.07	0.30	42	-0.23	0.82	0.02	0.02	46	0.91	0.37
Territory quality	**-0.63**	**0.11**	**66**	**-5.86**	**<0.001**	**0.02**	**7.98×10^{-3}**	**68**	**2.37**	**0.02**
Group size	-0.35	0.45	81	-0.77	0.44	8.74×10^{-3}	0.03	82	0.26	0.80
Presence of fledglings	1.11	0.97	156	1.15	0.25	**-0.15**	**0.08**	**162**	**-1.99**	**0.05**
Season†	**-2.01**	**0.67**	**154**	**-3.01**	**0.003**	**0.17**	**0.05**	**145**	**3.13**	**0.002**
Replicate number	-0.33	0.35	148	-0.93	0.35	0.04	0.03	152	1.64	0.10
Time of the day	0.11	0.22	160	0.48	0.64	**-0.04**	**0.02**	**160**	**-2.23**	**0.03**

Sun strength	**-0.72**	**0.35**	**149**	**-2.06**	**0.04**	0.01	0.03	161	0.39	0.70

(b)

	Time spent ≤ 3 m[b] (s) (LMM)					Number of songs (negative binomial model)			
Fixed effects	*β*	SE	df	*t*	*P*	*β*	SE	*z*	*P*
Intercept	1.23	0.44	92	2.79	0.006	1.17	0.17	6.77	<0.001
3D model type*									
β_{brown}	0.09	0.25	78	0.38	0.71	-5.36×10^{-3}	0.11	-0.05	0.96
$\beta_{purple\text{-}and\text{-}black}$	0.50	0.30	45	1.66	0.10	0.07	0.11	0.69	0.49
Latency of response	-3.04×10^{-3}	2.11×10^{-3}	157	-1.44	0.15	**-6.14×10^{-3}**	**1.02×10^{-3}**	**-6.00**	**<0.001**
Age	-0.09	0.10	46	-0.89	0.38	-0.01	0.03	-0.50	0.62
Territory quality	-0.05	0.04	59	-1.50	0.14	**-0.03**	**0.01**	**-2.43**	**0.02**
Group size	-0.08	0.15	75	-0.52	0.60	**-0.11**	**0.05**	**-2.08**	**0.04**
Presence of fledglings	0.21	0.34	151	0.61	0.54	0.16	0.13	1.23	0.22
Season[†]	-0.23	0.24	132	-0.97	0.34	**-0.32**	**0.10**	**-3.20**	**0.001**
Replicate number	0.10	0.13	150	0.83	0.41	-0.04	0.05	-0.80	0.42
Time of the day	**0.18**	**0.08**	**169**	**2.27**	**0.02**	-0.06	0.03	-1.82	0.07
Sun strength	-0.11	0.13	157	-0.90	0.37	-0.03	0.05	-0.59	0.56

(c)

Random effects	Duration of response[a]	Closest approach[a]	Time spent ≤ 3 m[b]	Number of songs
$\sigma^2_{individual}$				
Intercept	14.53	0.07	1.01	0.04
Brown	3.58	0.03	0.33	1.31×10^{-5}
Purple-and-black	13.21	0.06	1.44	1.35×10^{-6}
$\sigma^2_{model\ exemplar}$	0.87	0.00	0.00	3.77×10^{-6}

$\sigma^2_{\text{substrate type}}$	0.00	2.52×10^{-3}	0.02	1.20×10^{-7}
σ^2_{date}	0.00	0.00	0.00	8.78×10^{-7}
$\sigma^2_{\text{residual}}$	12.84	0.09	1.80	-
Marginal / conditional R^2	0.48 / 0.77	0.21 / 0.59	0.12 / 0.55	-

(d)

	Duration of response[a]			Closest approach[a]			Time spent ≤ 3 m[b]			Number of songs		
Pairwise comparison	$\beta\pm$SE	z	P	$\beta\pm$SE	z	P	$\beta\pm$SE	z	P	$\beta\pm$SE	z	P
brown *vs* control	0.86±0.98	0.88	0.65	-0.06±0.06	-1.11	0.51	0.09±0.25	0.38	0.92	-0.005±0.11	-0.05	1.00
purple-and-black *vs* control	1.97±1.10	1.79	0.17	-0.12±0.06	-1.90	0.14	0.50±0.30	1.66	0.22	0.07±0.11	-0.69	0.76
purple-and-black *vs* brown	1.11±0.77	1.44	0.31	-0.06±0.06	-1.13	0.50	0.41±0.27	1.54	0.27	0.08±0.15	0.53	0.85

Table S12. Subordinates display a similar or weaker response to simulated intrusions compared to breeder males. Shown are (a) results from LMMs examining the effects of 3D model type (reference: *control), social status (reference: •subordinate), territory quality, group size, presence of fledglings, replicate number, time of the day and sun strength on (a) duration of response and closest approach, and (b) time spent within 3 m and 1 m of the model. Variance of random effects and marginal / conditional R^2 for all models are presented in (c). Significant values ($P < 0.05$) are in bold. Response variables were either [a]*sqrt*- or [b]*log*-transformed.

(a)

	Duration of response[a] (s) (n = 176)					Closest approach[a] (m) (n = 184)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P
Intercept	5.96	2.41	17	2.47	0.03	1.76	0.24	20	7.42	<0.001
3D model type*										
β_{brown}	0.97	1.88	6	0.52	0.63	-0.16	0.22	13	-0.75	0.46
$\beta_{purple\text{-}and\text{-}black}$	1.50	1.88	5	0.80	0.46	-0.13	0.22	13	-0.59	0.56
Social status•	**6.94**	**0.98**	**30**	**7.06**	**<0.001**	0.07	0.06	39	1.30	0.20
Territory quality	**-0.63**	**0.13**	**35**	**-4.65**	**<0.001**	**0.05**	**0.01**	**37**	**3.73**	**<0.001**
Group size	**-1.68**	**0.55**	**32**	**-3.04**	**0.005**	0.10	0.05	46	1.90	0.06
Presence of fledglings	1.97	1.54	48	1.28	0.21	-0.11	0.14	69	-0.79	0.43
Replicate number	-1.04	0.64	27	-1.64	0.11	-0.02	0.04	16	-0.49	0.63
Time of the day	-0.40	0.35	104	-1.16	0.25	6.91×10^{-3}	0.03	105	0.25	0.81
Sun strength	-0.19	0.53	38	-0.36	0.72	3.19×10^{-3}	0.04	27	0.08	0.94

(b)

	Time spent ≤ 3 m[b] (s) (n = 179)					Time spent ≤ 1 m[b] (s) (n = 184)				
Fixed effects	β	SE	df	t	P	β	SE	df	t	P

Intercept	1.71	0.63	14	2.72	0.02	0.24	0.72	19	0.34	0.74
3D model type*										
β_{brown}	0.44	0.47	4	0.93	0.41	0.44	0.67	14	0.66	0.52
$\beta_{purple\text{-}and\text{-}black}$	0.29	0.47	4	0.60	0.58	0.55	0.67	14	0.82	0.43
Social status•	-0.27	0.27	35	-0.98	0.33	-6.88×10^{-3}	0.14	73	-0.05	0.96
Territory quality	**-0.13**	**0.05**	**33**	**-2.62**	**0.01**	**-0.14**	**0.04**	**35**	**-3.93**	**<0.001**
Group size	**-0.60**	**0.20**	**47**	**-3.02**	**0.004**	0.15	0.15	45	-1.02	0.31
Presence of fledglings	0.34	0.54	67	0.62	0.54	-0.26	0.39	68	-0.66	0.51
Replicate number	-0.09	0.17	18	-0.54	0.60	-0.07	0.13	21	-0.52	0.61
Time of the day	0.06	0.11	113	0.52	0.60	0.03	0.08	113	0.37	0.71
Sun strength	0.01	0.15	31	0.08	0.93	0.09	0.12	30	0.75	0.46

(c)

Random effects	Duration of response[a]	Closest approach[a]	Time spent ≤ 3 m[b]	Time spent ≤ 1 m[b]
$\sigma^2_{individual}$	6.27	4.42×10^{-3}	0.21	1.31×10^{-15}
$\sigma^2_{model\ exemplar}$	5.06	0.08	0.27	0.75
$\sigma^2_{substrate\ type}$	0.98	0.00	0.01	0.00
σ^2_{date}	2.80	7.26×10^{-3}	0.07	0.10
$\sigma^2_{territory}$	6.12	0.12	1.61	1.02
$\sigma^2_{residual}$	15.64	0.08	1.78	0.58
Marginal / conditional R^2	0.55 / 0.81	0.23 / 0.78	0.24 / 0.66	0.23 / 0.82

Table S13. The effect of territory quality on the aggressiveness of breeder males is independent of the 3D model type used. Shown are results from statistical models examining the effects of the interaction between 3D model type (reference: *control) and territory quality ('TQ' below), latency of response (in s), age (in years), group size, presence of fledglings, presence of subordinate males unrelated to the breeder male, presence of subordinate males unrelated to the breeder female, season (reference: †start of the dry season), replicate number, time of the day and sun strength on (a) duration of response and closest approach, and (b) time spent within 3 m and 1 m of the model. Variance of random effects and marginal / conditional R^2 for all models are presented in (c). Significant values ($P < 0.05$) are in bold. Response variables were either [a]*sqrt*- or [b]*log*-transformed.

(a)

	Duration of response[a] (s) (LMM)					Closest approach[a] (m) (LMM)				
Fixed effects	β	SE	df	*t*	*P*	β	SE	df	*t*	*P*
Intercept	14.79	1.23	47	12.06	<0.001	1.75	0.12	72	14.38	<0.001
3D model type*										
β_{brown}	0.45	0.62	38	0.72	0.48	-0.10	0.06	45	-1.62	0.11
$\beta_{purple\text{-}and\text{-}black}$	**2.47**	**0.73**	**37**	**3.38**	**0.002**	**-0.22**	**0.06**	**68**	**-3.79**	**<0.001**
TQ	**-0.46**	**0.13**	**44**	**-3.65**	**<0.001**	**0.05**	**0.01**	**45**	**3.85**	**<0.001**
3D model type × TQ										
β_{brown} × TQ	0.03	0.11	37	0.26	0.80	-0.01	0.01	44	-1.26	0.22
$\beta_{purple\text{-}and\text{-}black}$ × TQ	-0.16	0.13	35	-1.20	0.24	2.90×10^{-6}	0.01	65	0.00	1.00
Latency of response	**-0.06**	**6.99×10^{-3}**	**181**	**-8.38**	**<0.001**	**1.91×10^{-3}**	**6.52×10^{-4}**	**172**	**2.93**	**0.004**
Age	0.14	0.22	39	0.62	0.54	0.01	0.02	41	0.52	0.61
Group size	-0.19	0.47	95	-0.40	0.69	-0.04	0.05	114	-0.77	0.45

Presence of fledglings	1.09	0.92	153	1.19	0.24	**-0.21**	**0.09**	**185**	**-2.34**	**0.02**
Presence of subordinate ♂ unrelated to breeder ♂	-0.97	1.82	71	-0.53	0.60	0.11	0.19	87	0.58	0.56
Presence of subordinate ♂ unrelated to breeder ♀	-0.55	1.28	108	-0.43	0.67	0.06	0.13	130	0.44	0.66
Season[†]	**-2.39**	**0.61**	**144**	**-3.93**	**<0.001**	**0.15**	**0.06**	**170**	**2.58**	**0.01**
Replicate number	-0.41	0.31	154	-1.35	0.18	-0.01	0.03	158	-0.48	0.63
Time of the day	-0.12	0.20	160	-0.61	0.54	-0.01	0.02	165	-0.77	0.44
Sun strength	**-0.70**	**0.31**	**155**	**-2.24**	**0.03**	2.68×10^{-3}	0.03	163	0.09	0.93

(b)

	Time spent ≤ 3 m^b (s) (LMM)					Time spent ≤ 1 m^b (s) (LMM)				
Fixed effects	β	SE	df	*t*	*P*	β	SE	df	*t*	*P*
Intercept	1.19	0.46	76	2.61	0.01	0.63	0.39	18	1.63	0.12
3D model type*										
β_{brown}	**0.56**	**0.26**	**49**	**2.17**	**0.03**	0.13	0.30	5	0.41	0.70
$\beta_{purple\text{-}and\text{-}black}$	**1.13**	**0.24**	**58**	**4.70**	**<0.001**	0.37	0.30	5	1.23	0.27
TQ	**-0.09**	**0.04**	**42**	**-2.04**	**0.05**	**-0.11**	**0.03**	**49**	**-3.44**	**0.001**
3D model type × TQ										
β_{brown} × TQ	-0.02	0.05	48	-0.42	0.68	-0.03	0.03	143	-1.06	0.29
$\beta_{purple\text{-}and\text{-}black}$ × TQ	-0.08	0.04	55	-1.90	0.06	-0.06	0.03	117	-1.80	0.08
Latency of response	**-7.78×10^{-3}**	**2.62×10^{-3}**	**172**	**-2.97**	**0.003**	-1.40×10^{-3}	1.69×10^{-3}	177	-0.83	0.41
Age	-0.07	0.08	42	-0.89	0.38	0.04	0.06	43	0.57	0.57
Group size	-0.02	0.17	91	-0.14	0.89	0.19	0.12	113	1.54	0.13
Presence of fledglings	0.58	0.36	178	1.62	0.11	0.40	0.24	191	1.70	0.09
Presence of subordinate ♂ unrelated to breeder ♂	-0.45	0.65	66	-0.70	0.49	-0.41	0.48	89	-0.85	0.40

Presence of subordinate ♂ unrelated to breeder ♀	-0.02	0.48	104	-0.04	0.97	-0.39	0.34	129	-1.14	0.26
Season[†]	-0.10	0.24	170	-0.40	0.69	**-0.30**	**0.15**	**148**	**-1.96**	**0.05**
Replicate number	0.05	0.12	161	0.43	0.67	0.01	0.07	157	0.19	0.85
Time of the day	0.13	0.08	174	1.69	0.09	0.04	0.05	176	0.89	0.38
Sun strength	0.01	0.12	165	0.10	0.92	-3.18×10^{-4}	0.08	161	-0.004	1.00

(c)

Random effects	Duration of response[a]	Closest approach[a]	Time spent ≤ 3 m[b]	Time spent ≤ 1 m[b]
$\sigma^2_{individual}$				
Intercept	11.50	0.14	1.00	0.82
Brown	2.33	0.03	0.58	0.02
Purple-and-black	8.40	0.02	0.23	0.05
$\sigma^2_{model\ exemplar}$	0.00	0.00	0.00	0.12
$\sigma^2_{substrate\ type}$	1.04	0.01	0.13	0.03
σ^2_{date}	0.00	0.00	0.00	0.00
$\sigma^2_{residual}$	11.21	0.10	1.79	0.67
Marginal / conditional R^2	0.50 / 0.78	0.30 / 0.74	0.26 / 0.65	0.28 / 0.74

Table S14. The number of direct neighbours has no effect on the aggressiveness of breeder males, and including it in the statistical models does not affect the effects of other explanatory variables. Shown are results from statistical models examining the effects of 3D model type (reference: *control), latency of response (in s), age (in years), territory quality, number of direct neighbours, group size, presence of fledglings, presence of subordinate males unrelated to the breeder male, presence of subordinate males unrelated to the breeder female, season (reference: †start of the dry season), replicate number, time of the day and sun strength on (a) duration of response and closest approach and (b) time spent within 3 m and 1 m of the model and number of songs (n = 216). Variance of random effects and marginal / conditional R^2 for all models are presented in (c). Significant values ($P < 0.05$) are in bold. Response variables were either [a]*sqrt*- or [b]*log*-transformed (except number of songs).

(a)

	Duration of response[a] (s) (LMM)					Closest approach[a] (m) (LMM)				
Fixed effects	β	SE	Df	t	P	β	SE	Df	t	P
Intercept	15.01	1.22	48	12.33	<0.001	1.73	0.13	9	13.22	<0.001
3D model type*										
β_{brown}	0.44	0.61	62	0.72	0.47	-0.10	0.08	1	-1.27	0.42
$\beta_{purple\text{-}and\text{-}black}$	**2.49**	**0.74**	**39**	**3.38**	**0.002**	-0.22	0.07	1	-3.04	0.20
Latency of response	**-0.06**	**6.98×10^{-3}**	**182**	**-8.40**	**<0.001**	**2.00×10^{-3}**	**6.54×10^{-4}**	**170**	**3.05**	**0.003**
Age	0.12	0.22	39	0.56	0.58	0.01	0.02	40	0.58	0.56
Territory quality	**-0.43**	**0.13**	**44**	**-3.28**	**0.002**	**0.04**	**0.01**	**45**	**2.77**	**0.008**
# direct neighbours	-0.23	0.23	52	-0.98	0.33	0.03	0.03	56	1.09	0.28
Group size	-0.07	0.49	100	-0.14	0.89	-0.05	0.05	115	-1.05	0.30
Presence of fledglings	1.33	0.93	169	1.44	0.15	**-0.24**	**0.09**	**178**	**-2.63**	**0.009**

Presence of subordinate ♂ unrelated to breeder ♂	-1.31	1.85	75	-0.70	0.48	0.16	0.19	94	0.85	0.40
Presence of subordinate ♂ unrelated to breeder ♀	-0.39	1.29	108	-0.30	0.77	0.03	0.13	120	0.27	0.79
Season†	**-2.52**	**0.61**	**153**	**-4.12**	**<0.001**	**0.17**	**0.06**	**152**	**2.89**	**0.004**
Replicate number	-0.44	0.31	159	-1.44	0.15	-9.47×10^{-3}	0.03	160	-0.33	0.74
Time of the day	-0.15	0.20	160	-0.75	0.46	-8.86×10^{-3}	0.02	167	-0.47	0.64
Sun strength	**-0.77**	**0.31**	**158**	**-2.49**	**0.01**	8.46×10^{-3}	0.03	162	0.29	0.77

(b)

	Time spent ≤ 3 m^b (s) (LMM)					Time spent ≤ 1 m^b (s) (LMM)				
Fixed effects	*β*	SE	df	*t*	*P*	*β*	SE	df	*T*	*P*
Intercept	1.31	0.46	72	2.86	0.006	0.69	0.37	15	1.84	0.09
3D model type*										
β_{brown}	**0.55**	**0.25**	**56**	**2.16**	**0.03**	0.12	0.27	4	0.43	0.69
$\beta_{purple\text{-}and\text{-}black}$	**1.14**	**0.25**	**50**	**4.62**	**<0.001**	0.37	0.28	4	1.34	0.25
Latency of response	**-7.47×10^{-3}**	**2.64×10^{-3}**	**173**	**-2.83**	**0.005**	-1.16×10^{-3}	1.70×10^{-3}	168	-0.68	0.50
Age	-0.08	0.08	41	-0.95	0.35	0.03	0.06	42	0.49	0.63
Territory quality	-0.09	0.05	44	-1.93	0.06	**-0.10**	**0.04**	**46**	**-2.81**	**0.007**
# direct neighbours	-0.08	0.09	52	-0.93	0.36	-0.09	0.06	58	-1.32	0.19
Group size	0.01	0.18	94	0.06	0.96	0.23	0.13	117	1.77	0.08
Presence of fledglings	0.64	0.36	176	1.78	0.08	0.46	0.24	178	1.92	0.06
Presence of subordinate ♂ unrelated to breeder ♂	-0.57	0.67	68	-0.85	0.40	-0.50	0.48	89	-1.04	0.30
Presence of subordinate ♂ unrelated to breeder ♀	0.01	0.48	100	0.03	0.98	-0.34	0.34	120	-1.00	0.32
Season†	-0.14	0.24	170	-0.58	0.56	**-0.34**	**0.15**	**147**	**-2.16**	**0.03**

Replicate number	0.03	0.12	162	0.29	0.77	1.32×10^{-3}	0.07	155	0.02	0.99
Time of the day	0.12	0.08	174	1.58	0.12	0.04	0.05	165	0.83	0.41
Sun strength	-0.03	0.12	163	-0.21	0.83	-0.02	0.08	163	-0.26	0.80

(c)

Random effects	Duration of response[a]	Closest approach[a]	Time spent ≤ 3 m[b]	Time spent ≤ 1 m[b]	Number of songs
$\sigma^2_{individual}$					
Intercept	11.21	0.14	0.99	0.78	0.06
Brown	1.82	0.03	0.52	0.02	5.45×10^{-8}
Purple-and-black	8.58	0.02	0.31	0.07	7.37×10^{-9}
$\sigma^2_{model\ exemplar}$	0.00	3.62×10^{-3}	0.00	0.10	2.11×10^{-9}
$\sigma^2_{substrate\ type}$	1.05	0.01	0.15	0.04	0.02
σ^2_{date}	0.00	0.00	0.00	0.00	4.17×10^{-3}
$\sigma^2_{residual}$	11.35	0.10	1.83	0.70	-
Marginal / conditional R^2	0.50 / 0.78	0.31 / 0.75	0.25 / 0.63	0.26 / 0.72	-

Chapter 4

An integral assessment of variability, heritability, condition dependence and fitness correlates of the multidimensional colour phenotype in a passerine bird

Abstract

Elaborate ornamental traits are commonly assumed to be honest signals of individual quality, owing to the presumed costs involved in their production and/or maintenance. Individuals with more elaborate traits are thus expected to achieve higher mating success, through increased success in mate choice or same-sex competition. Ornamental traits are often highly variable, presumably because of high underlying genetic variation, and it has been suggested that their expression should be more heritable than non-ornamental traits. In many bird species, males display colourful plumages with multiple distinct patches of different developmental origins, forming complex colour phenotypes. Despite this complexity, colourful ornaments are often studied in isolation, with little information on the species' perceptual ability and no suitable non-ornamental controls to be compared with. Based on 8 years of plumage reflectance data, we assessed the adaptive significance of the multidimensional male colour phenotype in the purple-crowned fairy-wren *M. coronatus*, a cooperatively breeding bird. We tested the predictions that the expression of ornamental colours (purple, blue and black) is (1) more condition-dependent, (2) more variable and more heritable, and (3) more strongly related to male fitness compared to non-ornamental colours (buff-white and brown). Our results show that, contrary to predictions of heightened condition-dependence in ornaments, only brightness of the buff-white and brown colouration increased with body condition. Nevertheless, ornamental colours exhibited greater levels of variability, while chromatic variation in the

purple colouration displayed substantial heritability and predicted male annual reproductive success: more purple males sired more offspring. Despite this partial support for predictions, the lack of consistent patterns illustrates the complexity of visual signals and highlights the need to study colour phenotypes in their entirety and using suitable non-signalling controls.

Keywords: additive genetic variance, avian visual space, condition-dependence, *Malurus*, ornamental plumage, sexual selection

Introduction

Many animals use elaborate or conspicuous traits as quality signals to choose mates and/or assess rivals (Darwin, 1871; Andersson, 1994). Indeed, such ornamental traits may convey various information regarding the bearer's phenotypic condition, fighting ability and genetic constitution, and this in a reliable way if trait production and/or maintenance entail substantial costs (Zahavi, 1975; Hamilton & Zuk, 1982; Cotton et al., 2004; Andersson, 2006). As only individuals of high quality can afford the costs of extreme versions of such ornaments, they are generally assumed to be more successful in mate choice or same-sex competition, therefore achieving higher mating success – a benefit that presumably offsets the apparent survival cost of signal elaboration (Darwin, 1871; Lozano, 1994; Hill, 2002). Alternatively, under Fisher's (1930) runaway selection model, a genetic correlation develops between the ornamental trait and the mating preference for that trait. This process will continue, resulting in simultaneous enhanced mate choice and exaggeration of the trait, until the trait is constrained by natural selection, when it becomes so large that it is a major impediment to survival.

Avian species have frequently been used as model systems to investigate the evolution of conspicuous traits as males of many species display colourful sexual ornaments (Owens & Hartley, 1998; Dunn et al., 2001; McGraw et al., 2002). Yet, despite the fact that many bird species display complex phenotypes, including multiple distinct colour patches or patterns, the study of adaptive functions of colourful ornaments is often performed on single traits, and not informed by the perceptual ability of the species involved. Quantifying and comparing colour expression can be a challenging task, but recent methodological advances, such as psychophysical models of avian colour vision (Vorobyev et al., 1998), allow us to do so in a robust way. Moreover, although ornamental colours are implicitly assumed to be more costly and more strongly correlated with individual fitness than non-ornamental colours, studies using non-ornamental 'controls' are largely missing (Cotton et al., 2004; Bonduriansky & Rowe,

2005; Gosden & Chenoweth, 2011; but see Kemp & Rutowski, 2007; Tibbetts, 2010). Nonetheless, in order to fully understand why colourful ornaments evolve and how they contribute to an individual's fitness, it is crucial to (1) use an integrated approach by considering the complete colour phenotype, and (2) contrast the information content and fitness correlates of ornamental colours with those of non-ornamental controls.

Ornamental colours are often highly variable within species. Delhey et al. (2017) previously demonstrated that intraspecific levels of chromatic variability in bird plumage colouration are strongly correlated with conspicuousness, with more conspicuous colours displaying greater variability. This pattern may be explained by a higher sensitivity to resource availability and/or developmental conditions; but it could also be linked to the need for colours that are very different from natural backgrounds to have greater levels of discriminability (Delhey et al., 2017). Ornamental colours are thus predicted to be more variable than non-ornamental, more cryptic, colours. However, ornamental traits are generally under strong directional selection, which should lead to depletion of the underlying genetic variation and reduced phenotypic variability (Taylor & Williams, 1982; Merilä & Sheldon, 1999). The maintenance of genetic variance in sexually selected traits despite persistent directional selection is an ongoing evolutionary conundrum, known as the lek paradox, and has been debated for decades (Taylor & Williams, 1982; Pomianskowski & Møller, 1995; Kotiaho et al., 2001, 2008). Several resolutions have been proposed, with the most popular theory being based on the ideas of condition-dependence and genic capture: it was suggested that (1) the expression of ornamental traits depends on condition, and (2) condition itself depends on genes at multiple loci that therefore generate high mutational variance (Rowe & Houle, 1996; Kotiaho et al., 2001; Tomkins et al., 2004). Although the expression of many traits may reflect an individual's condition to some extent, ornamental colours are expected to show heightened condition-dependence, being subject to stronger sexual selection – leading to greater

exaggeration and higher sensitivity to condition – compared to non-ornamental traits (Rowe & Houle, 1996; Cotton et al., 2004; Bonduriansky & Rowe, 2005). Yet, Johnstone et al. (2009) argued that it may not necessarily be the case depending on the nature of the costs of expressing such traits and how they affect individuals of differing condition. Moreover, the question of the heritability of ornaments remains contentious as heritability is the ratio of additive genetic variance to total phenotypic variance, which includes environmental effects (Merilä & Sheldon, 1999). Consequently, high levels of additive genetic variance may not necessarily result in high heritability (Price & Schluter, 1991), and so far empirical work on signal heritability has been equivocal (Pomiankowski & Møller, 1995; Merilä & Sheldon, 1999; Hadfield et al., 2007; Tibbetts, 2010; Charmantier et al., 2017). If an ornamental trait is indicative of 'good genes', we may expect its expression to be highly heritable; conversely, the trait could be an effective signal of the individual's environmental experiences, due to strong environmental effects, and therefore show low heritability.

Bird plumage colours can be produced by two main mechanisms: the deposition of pigments (e.g., melanins and carotenoids), or the interaction of light with ordered or semi-ordered nanostructures (i.e. structural colouration). Melanins, which generate hues ranging from black to grey and brown to rufous, are endogenously synthesised (McGraw, 2006). As a result, their expression is generally thought to be strongly genetically determined (Roulin & Ducrest, 2013) and weakly sensitive to individual condition or quality (Badyaev & Hill, 2000; Hill, 2006; McGraw, 2006), although support for this second assumption is very mixed (Hill & Brawner, 1998; McGraw et al., 2002; Roulin, 2016). On the other hand, highly ordered nanostructures composed of keratin and air, responsible for non-iridescent UV, violet and blue colours, need to be assembled and arranged in three-dimensional arrays during feather development (Prum, 2006). Whether structural colouration is costly to produce and/or related to individual quality appears, however, to vary among studies (Keyser & Hill, 2000; McGraw

et al., 2002; Prum, 2006; Peters et al., 2011). Nevertheless, given that the optical and developmental mechanisms that produce structural colouration differ from those that produce pigment-based colouration, we may expect ornaments of structural vs. pigmentary origin to differ in their information content, heritability and how they correlate with individual fitness. To date, only a few studies have simultaneously examined these aspects for multiple plumage ornaments of different developmental origins within a same species (Hadfield et al., 2007; Hegyi et al., 2007).

Here we aim to compare the condition-dependence, heritability and fitness correlates of ornamental vs. non-ornamental colours using multiple colour traits within a single species, the purple-crowned fairy-wren *Malurus coronatus*. The plumage of male purple-crowned fairy-wrens includes multiple colour patches, produced through different mechanisms and with variable signalling potential. Most prominently, all males annually produce a purple (structural and melanin-based) and black (melanin-based) breeding plumage, largely restricted to the head (Rowley & Russell, 1997; Delhey et al., 2013; Fig. 1). Fan et al. (2018 / Chapter 3) previously established that the presence (vs. absence), as well as the extent of male breeding colouration signal male competitive ability, primarily for access to breeding vacancies that rarely arise in this cooperatively breeding species. However, whether variations in the quality of the purple and black breeding colours themselves relate to any aspect of male quality and/or fitness is not known. Additionally, male *M. coronatus* also have a conspicuous blue tail (structural) colouration that is sexually dichromatic, being less conspicuous and less blue in females and juveniles (Rowley & Russell, 1997; Delhey et al., 2013; unpublished data; Fig. 1), constituting a putative ornament. In contrast, the rest of the plumage, which includes buff-white and brown (melanin-based) patches, is much more cryptic and displayed by both sexes at all ages (Rowley & Russell, 1997; Delhey et al., 2013; Fig. 1); therefore, these patches are unlikely to carry any signalling function and may be considered as non-ornamental colours.

The co-occurrence of these multiple conspicuous and cryptic plumage patches in the purple-crowned fairy-wren offers the opportunity to investigate the adaptive significance of multiple ornamental colours in comparison with multiple non-ornamental colours. Moreover, using a wild study system allows us to test the above predictions in the range of naturally experienced conditions and relate the colours to fitness. Finally, because we use psychophysical models of avian colour vision (Vorobyev et al., 1998), we can model variation in the chromatic (colour) and achromatic (brightness) components of each patch colour in a way that reflects the perceptual ability of the intended receivers (conspecifics). Here we use 8 years of data to test whether putative ornamental and non-ornamental colours of male purple-crowned fairy-wrens differ in (1) their condition-dependence – by testing what intrinsic and environmental factors underlie variability in colouration, (2) their heritability – using animal models based on a detailed pedigree of the population, and (3) how they correlate with male reproduction and survival. Although we found no support for heightened condition-dependence of the ornamental colours, there is some evidence for greater levels of heritability in comparison with non-ornamental colours, and some associations with male reproductive success.

Materials and methods

Study species

We studied a colour-banded population of *Malurus coronatus coronatus* at the Australian Wildlife Conservancy's Mornington Wildlife Sanctuary in northwest Australia (17º31'S, 126º6'E) from July 2005 to November 2017. These birds are restricted to patchy riparian vegetation of *Pandanus aquaticus* and maintain all-purpose territories year-round, linearly arranged along Annie Creek and the Adcock River (Kingma et al., 2009). *M. coronatus* can breed year-round with a peak in breeding activity at this site during the wet season (December-

March), and a smaller peak in the late dry season (August-September) in some years (Hall & Peters, 2009; Peters et al., 2013).

M. coronatus breeds cooperatively, whereby 40-70% of dominant breeding pairs (distinguished by duet singing; Hall & Peters, 2008, 2009) are accompanied by a number of non-breeding male and female subordinates (mostly offspring from previous broods), of which most contribute to nestling feeding (Kingma et al., 2010, 2011a, b). Subordinate males may acquire a breeder position either by filling a vacancy left by a deceased breeder (either inheritance of the home territory or dispersal to another – Kingma et al., 2011b; Hidalgo Aranzamendi et al., 2016), or less commonly by taking over part of the natal territory or establishing a new territory.

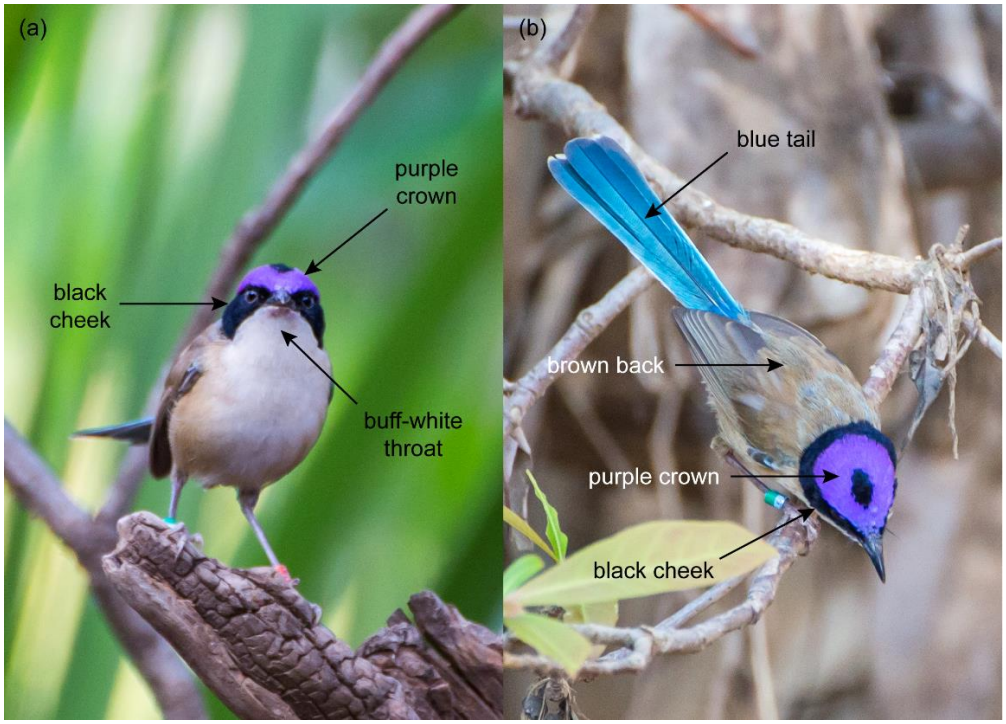


Figure 1. Breeding plumage in male purple-crowned fairy-wrens. Photographs show a male in purple-and-black breeding plumage viewed (a) from the front and (b) from above. Arrows indicate the five plumage patches whose reflectance was measured in this study: purple crown (a, b), blue tail (b), buff-white throat (a), brown back (b) and black cheek (a, b). Photos: L. Lermusiaux.

Males replace their dull brown non-breeding head plumage annually with purple and black feathers: they acquire a purple crown surrounding a black round central patch, as well as a black nape (Peters et al., 2013; Fig. 1). Adult males display black cheeks year-round but presumably go through a moult of the cheek feathers when they develop their purple-and-black breeding plumage (pers. obs.). Most males initiate pre-breeding moult in July-September before breeding starts, but in some cases it overlaps temporally with breeding (29% of breeder males; Fan et al., 2017 / Chapter 2). First-year males and subordinate males complete their moult later than older males and breeder males respectively (Fan et al., 2017 / Chapter 2). Moreover, although they usually moult to some degree, only 16% of first-year males develop a complete breeding plumage, whereas most older males do so (Fan et al., 2017, 2018 / Chapters 2 and 3). Previous work has shown that subordinate males displaying a more complete breeding plumage are more likely to acquire a breeder position by inheritance or dispersal (Fan et al., 2018 / Chapter 3). Other plumage patches do not change noticeably in colouration over the year, including the blue tail, and the rest of the plumage that is mainly brown above (back and wings) and buff-white below (throat, breast and belly; Rowley & Russell, 1997; Delhey et al., 2013; Fig. 1). It may be noted that females also annually undertake a pre-breeding moult of the head, but replace the brown non-breeding head plumage with slate-grey feathers, and possess rufous cheek patches and a turquoise tail (as juveniles do; Rowley & Russell, 1997).

Field methods

From July 2005 to March 2011, weekly population censuses were conducted year-round (01 July = start of year) to document group size and social status of each uniquely colour-banded male. Breeding activity was intensively monitored (for details see Kingma et al., 2009; Hidalgo Aranzamendi et al., 2016), and all nestlings banded and assigned an accurate hatch date. Birds captured as adults at the start of the study were classified as ‘age unknown’ with a minimum

age based on the presence or absence of offspring (of known age) or the completeness of the breeding plumage. From October 2011 to November 2017, biannual population censuses were conducted in October-November and May-June, documenting group size and social status of all individuals (for details see Hidalgo Aranzamendi et al., 2016; Fan et al., 2017 / Chapter 2). All new unbanded birds (fledglings, subordinates or immigrants) were banded, aged by age-specific development of appearance and behavioural cues (tail length, begging behaviour, plumage colour, bill colour). Parentage of all local birds was determined using six or nine microsatellite loci (for details see Kingma et al., 2009; Hidalgo Aranzamendi et al., 2016).

From 2005 to 2009 and from 2015 to 2017, we measured plumage reflectance of all birds that were captured using mist-nets, with some individuals being measured multiple times within a year or in different years (for details see 'Colour analysis' section below). For each captured individual, we also measured tarsus length (a measure of body size) and body mass – together indicative of body condition at the time of colour measurement. Tarsus length could be an indicator of male quality in *M. coronatus* as it correlates with song frequency (pitch) in certain male songs (Hall et al., 2013).

From 2007 onwards, intensive yearly censuses covering almost all suitable habitat along the tributaries that join the study site were conducted to find birds that had dispersed outside the core area (emigrants), providing reliable information on the survival (presence or absence) of all individuals (for details see Hidalgo Aranzamendi et al. 2016). Birds were declared dead on the basis of failure to sight them in regular surveys and assigned a death date estimate (for details see Fan et al., 2017 / Chapter 2).

Territories are stable year-round, and most boundaries remain stable through the years. Occasional changes in boundaries (shifts, territory splitting or establishment of new territories) were recorded throughout the study. Territory quality was assessed (yearly between 2005 and 2008, and once in 2013, 2015 and 2017) based on the *Pandanus* cover following Hidalgo

Aranzamendi et al. (2016). For sampling seasons between surveys, data were interpolated at even increments or decrements. *Malurus coronatus* generally does not occupy habitat without *Pandanus* (wherein 51% of daytime is spent and 95% of nests are built; Kingma et al., 2011a) and the distribution of *Pandanus* varies considerably within the population.

Daily records of rainfall were obtained from a local weather station at Mornington Wildlife Sanctuary from October 2004 to December 2017 (Australian Bureau of Meteorology weather station 002076).

Colour analysis

From 2005 to 2009 (throughout the year) and from 2015 to 2017 (in October-November only), a total of 195 males were captured using mist-nets and their plumage reflectance measured, with some males being measured multiple times within a year or in different years. Plumage reflectance measurements were collected in the shade using an AvaSpec-2048 spectrometer connected to an AvaLight-XE xenon pulsed light source using a bifurcated fibre optic cable fitted at the end with a cylindrical probe to standardise measuring distance and exclude ambient light. For each captured male we collected five reflectance spectra – at different predefined and standardised spots – of five plumage patches at a perpendicular probe angle as follows: (1) purple crown (for males in partial or complete breeding plumage), (2) black cheek, (3) buff-white throat, (4) brown back and (5) blue tail (Fig. 1). Reflectance spectra between 300 and 700 nm (the visual sensitivity range of birds; Cuthill, 2006) were calculated relative to a WS-2 white standard using the software AVASOFT 7.5 (Figs 2a and 2b). All spectrometry materials and software were purchased from Avantes (Apeldoorn, Netherlands). We also scored the extent of breeding plumage of each captured male (0-100% purple; for details see Fan et al., 2017, 2018 / Chapters 2 and 3).

To compute chromatic variability, we used psychophysical models of avian colour vision (Vorobyev et al., 1998) following the methods described by Delhey et al. (2015). Visual models require knowledge on the visual sensitivity functions of the four types of cone used by birds in colour vision, the relative abundance of each of these cones in the retina and the spectrum of illuminating light. Colour vision in birds is mediated by four types of single cone sensitive to very short (*VS*), short (*S*), medium (*M*) and long (*L*) wavelengths of light (Vorobyev et al., 1998). Variation in visual sensitivity between species is mainly restricted to the *VS* and *S* cones and birds can be generally classified in two groups: ultraviolet-sensitive (U-type) and violet-sensitive (V-type) species; and U-type species have *VS* cones with peak sensitivity shifted towards shorter wavelengths (Hart & Hunt, 2007). As *M. coronatus* is a V-type species (Ödeen et al., 2012), we used the associated visual sensitivity function obtained from Endler and Mielke (2005). We used average cone proportions as obtained from Hart (2001) (0.38:0.69:1.14:1.00 for *VS*:*S*:*M*:*L*, respectively) and combined these with behavioural estimates of the Weber fraction (0.1; Vorobyev et al., 1998; Lind et al., 2014) using formula (10) in Vorobyev et al. (1998) to obtain the noise-to-signal ratios for each cone type (v_{VS} = 0.162, v_S = 0.120, v_M = 0.094, v_L = 0.1). We used the spectrum of standard daylight (D65, open habitats) as illuminant (Vorobyev et al., 1998).

Visual models yield a set of quantum catches for the four types of single cones (i.e. how much each cone type is stimulated by a specific combination of reflectance spectrum and irradiance) that can be transformed into three coordinates x, y, z that define the position of each spectrum in the visual space of birds (Figs 2c and 2d). This visual space takes the shape of a tetrahedron where each apex represents the sole stimulation of one cone type (Endler & Mielke, 2005). Using the formulae in Cassey et al. (2008), distances between points in visual space are measured in 'just noticeable differences' (jnd), whereby distances > 1 jnd are considered to be

discriminable by birds. We computed chromatic variability of all colours except black that contains very little discriminable chromatic variation.

For each patch colour (except black), we then performed a Principal Component Analysis (PCA) on the three xyz coordinates, using a covariance matrix (Table S1). The computed principal component (PC) scores are measured in the same unit as the original variables (jnd) and can be used as independent chromatic variables (Figs 2c and 2d). PCs that explain low amounts of chromatic variability (< 10% of variation) with a range < 4 jnd are less likely to carry meaningful biological information. This was the case for purple PC3, blue PC3, buff-white PC2 and PC3, and brown PC2 and PC3 (Table S1), which were therefore excluded from further analysis.

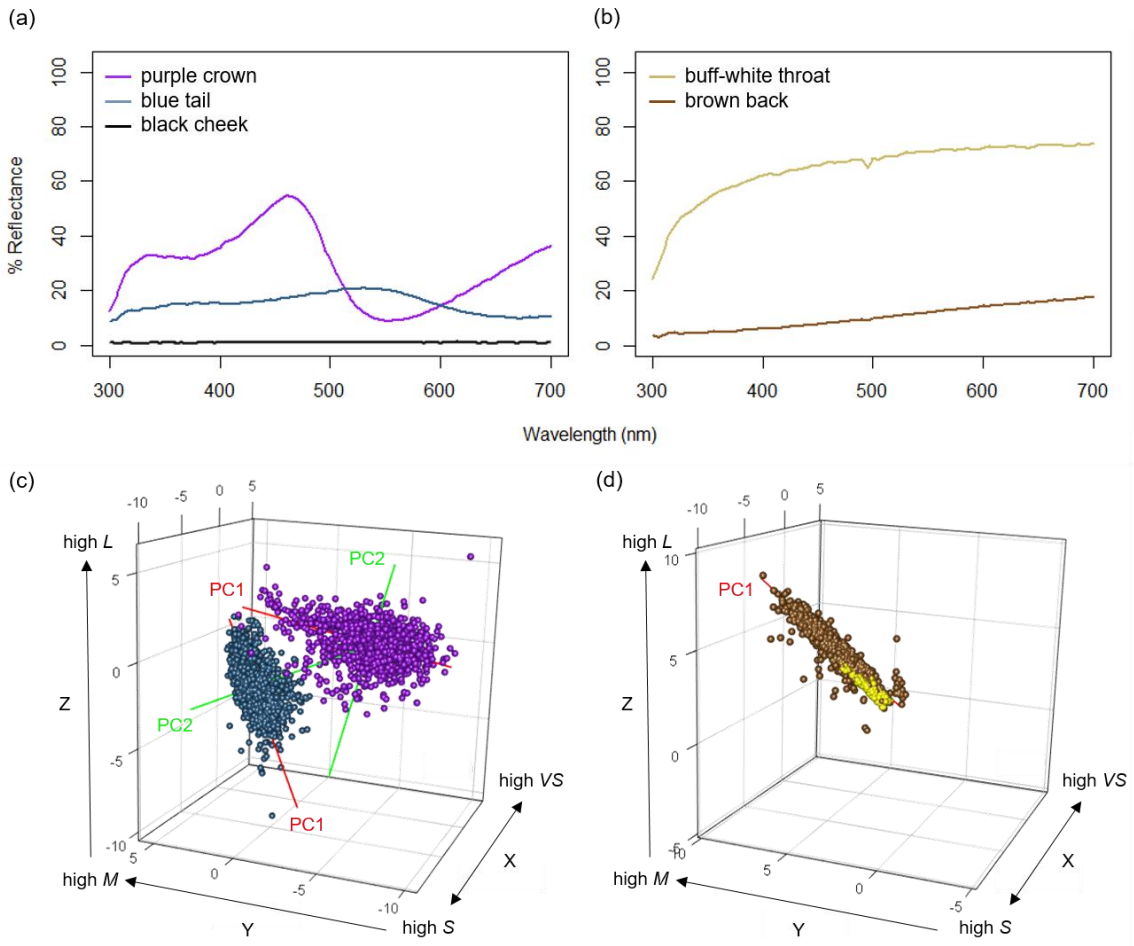


Figure 2. Plumage reflectance spectra measured in male purple-crowned fairy-wrens and represented in their visual space. Shown are examples of reflectance spectrum measured in a

male in breeding plumage for (a) the purple crown, black cheek and blue tail colouration, and (b) the buff-white throat and brown back colouration. Chromatic variability of each patch (except the black cheek) computed in the avian visual space is represented for (c) the purple crown and blue tail colouration, and (d) the buff-white throat (in yellow) and brown back colouration. PCs explaining ≥ 95% of chromatic variation for each colour are depicted, in red for PC1 axes and in green for PC2 axes; PC1 axes are very similar for the buff-white and brown colouration and therefore indistinguishable in (d). The X-axis represents stimulation of the *VS* cone relative to the *S* cone, higher values of the Y-axis represents higher stimulation of the *M* cone relative to *VS* and *S* cones, while the Z-axis represents higher relative stimulation of the *L* cone compared with the other three (units = jnd).

Similarly, we computed achromatic variability (i.e. 'brightness' or luminance variation) as described by Delhey et al. (2015) for all plumage patches. We used a Weber fraction of 0.2, following Olsson et al. (2017). For each spectra, the computed value of brightness was noted L (lightness) and measured in jnd.

As multiple (up to five) measurements were taken per plumage patch per capture, we averaged the values for each colorimetric variable (PC1, PC2 and L; the differences between measurements were small and mostly below the discrimination threshold; see Appendix S1). We used these average PC and L values in the statistical analyses below (purple crown: n = 233; black cheek: n = 248; blue tail: n = 350; buff-white throat: n = 208; brown back: n = 349).

Statistical analyses

For each colour, we first analysed (A) the degree of variability of the PCs and L and (B) whether they depended on intrinsic and environmental factors to test for condition-dependence. Then, (C) to model additive genetic effects, we estimated individual heritability of PCs and L

using a detailed pedigree of our study population. Finally, (D) we investigated correlations between PCs and L and various fitness variables in order to identify potential fitness benefits and costs associated with plumage colour elaboration. To control for potential false positives due to multiple testing when investigating condition-dependence and fitness correlates, we also used the Benjamini-Hochberg procedure to decrease the false discovery rate (Benjamini & Hochberg, 1995). Because our hypotheses state that ornamental colours should differ from non-ornamental colours, we divided the examined variables into two groups: the first group included purple PC1, PC2, L, blue PC1, PC2, L, and black L (i.e. a total of 7 tests), while the second group included buff-white PC1 and L, and brown PC1 and L (i.e. a total of 4 tests). The false discovery rate was set at 0.05.

All analyses were done in R 3.4.0 (R Core Team, 2017). Linear mixed models (LMM) were built using the packages "lme4" (Bates et al., 2015) and "lmerTest" (Kuznetsova et al., 2015). Generalised LMMs (GLMM) were first fitted as generalised linear models without random term to estimate dispersion. If data were under- or overdispersed, appropriate models were selected (see below). All continuous explanatory variables were centred.

(A) Chromatic and achromatic variability of plumage reflectance

We tested whether the ornamental colours display higher degree of variability compared to non-ornamental colours. To assess chromatic variability, for each plumage colour (except black) we computed the distance between each reflectance spectrum and the centroid (i.e. average position) of that plumage colour in the bird visual space. To assess achromatic variability, for each plumage colour we computed the distance to the average L value of that plumage colour. Then, to test differences among patch colours in chromatic and achromatic variability respectively, we built a linear model (LM) with distance to the centroid (average L value respectively; both *log*-transformed to ensure normal distribution of residuals) as the

response variable, and patch colour as a fixed effect. Using Tukey's Honest Significant Difference (HSD) test, we performed post hoc comparisons between patch colours. In addition, we also examined differences between ornamental vs. non-ornamental colours by running a similar model as above but including 'colour type' (i.e. ornamental or non-ornamental) as a fixed effect instead.

(B) Intrinsic and environmental effects on plumage reflectance

The analysis described below was performed on the biologically meaningful colorimetric variable of each patch colour (PC1, PC2 and L for purple and blue; PC1 and L for buff-white and brown; L only for black). Each of these chromatic and achromatic variables was fitted as the response variable in separate LMMs. In each LMM, we fitted age, social status (subordinate or dominant), body condition, group size and territory quality as fixed effects. To determine whether population-level plumage reflectance depended on rainfall, we also included the cumulative rainfall over the past year (from November of the preceding calendar year to October). For LMMs associated with the purple crown reflectance, we also included the extent of breeding plumage (0-100% purple) of each male to control for potential variation in males in partial breeding plumage due to the presence of brown feathers (all colorimetric variables associated with purple were indeed positively correlated with breeding plumage completeness; Table S3). Although different plumage patches have different moult schedules (see Appendix S2a), for consistency we also included the extent of breeding plumage (as a proxy for moulting stage) in the LMMs associated with the other plumage patches. In addition, for LMMs associated with the purple crown and black cheek reflectance, we included the time interval since pre-breeding moult completion to account for potential within-year variation in colour due to fading, soiling and/or abrasion (Delhey et al., 2010); however, as the model outputs indicated that including or excluding this variable did not affect the results, we only report the

results for the models excluding it (see Appendix S2b). Values of body condition were obtained by extracting the residuals of a linear regression of body mass against tarsus length and time of the day (hourly). Bird identity and year were included as random intercepts to account for non-independence in the data. This analysis was first restricted to birds whose age was accurately known (n = 81-155 for the different patch colours), but as age did not affect any chromatic or achromatic variables (except for buff-white PC1, but the variations with age were below the discrimination threshold; Tables S3-S7; see Appendix S2c), it was repeated using two age classes, "1" and "2+", with all birds of unknown age but known to be at least 2 years old included in the "2+" class (n = 169-235).

(C) Heritability of plumage reflectance

In order to estimate heritability of the chromatic and achromatic variables for each patch colour, we used the genetic pedigree information available for our population (Fig. S1) to transform our mixed effects models into animal models (Wilson et al., 2010). Using a Bayesian framework implemented with the package "MCMCglmm" (Hadfield, 2010), we fitted the LMMs detailed above, but only kept the fixed effects that were both statistically significant (P < 0.05) and biologically important, i.e. when the variation of the colorimetric variable across the range of realistic values of the fixed effect (e.g., from 9 to 13 g for body mass) was larger than 1 jnd. This approach allowed us to use simpler models with larger sample sizes. Random effects included bird identity twice (to model additive genetic effects and permanent environmental effects given that we had repeated measurements on multiple males) and year. We ran between 2 005 000 and 30 005 000 iterations per model, from which we discarded the initial 5000 (burn-in period). Each chain was sampled so that the effective sample size was 1000 (i.e. at an interval between 2000 and 30 000 iterations according to the total number of iterations). The total number of iterations was chosen to ensure low autocorrelation among

thinned samples. Fixed effect priors were normally distributed and centred on zero with large variances. Different priors were used for the residuals and random effects; mean values of the posterior distributions were robust to different relatively un-informative prior settings. The detailed structure of each model is summarised in Table S10. Posterior means and 95% credible intervals were estimated across the thinned samples for the mean effects (fixed effects), variances and variance ratios (i.e. repeatabilities and heritabilities).

In addition, we ran the exact same animal models using ASReml-R (Butler, 2009) in order to check their robustness. Random effects were tested for significance using likelihood ratio tests of models with and without each effect, assuming a chi-squared distribution with 1 d.f. Heritability was calculated as the additive genetic variance to total phenotypic variance ratio.

(D) Fitness benefits and costs associated with plumage reflectance

We evaluated several hypothesised fitness benefits and costs of colour elaboration that could apply to male *M. coronatus*, including the acquisition of a breeder position (for subordinates), reproductive success and survival. We did not investigate associations with EP success or the likelihood of losing paternity since we had a few observations only of such events (2 cases of EPP and 6 cases of cuckoldry in our dataset).

All the models described below were run separately for each patch colour. In all models we fitted the colorimetric variables (PC1, PC2 and L for purple and blue; PC1 and L for buff-white and brown; L only for black) as fixed effects, and bird identity and year as random intercepts (unless specified otherwise). In addition, for models associated with the purple crown colouration, because we found that PC1 and PC2 varied with the extent of breeding plumage – most likely due to the presence of brown feathers in the crown (see Table S3), we adjusted the values of these two PCs to account for this (for details see Appendix S2d). Hereafter we will simply use ‘PC1’ and ‘PC2’ to refer to these adjusted values.

(a) Acquisition of a breeder position

We tested whether plumage reflectance predicted the likelihood of acquiring a breeder position elsewhere among subordinate males at the population level (n = 36-82 for the different patch colours) as this is a critical determinant of male reproductive success (subordinates do not participate in reproduction, except for two cases recorded in our study population; Kingma et al., 2009). Cases of inheritance or splitting of the natal territory were excluded. We built a GLMM (or in case of under- or overdispersion a GLMM with template model builder – GLMMTMB – using the package "glmmTMB"; Magnusson et al., 2017) with the annual status change (became a breeder within the year = 1, did not = 0) as a binomial response variable, the colorimetric variable(s), age (levels "1" and "2+") and tarsus length as fixed effects, and territory identity as an additional random intercept. In the case of the brown back colouration, the model could not converge when including both territory identity and year as random effects; therefore, we fitted each factor in separate models, which provided quantitatively similar results and showed that the variance of both was close to zero.

(b) Social reproductive success

We tested whether plumage reflectance correlated with male social reproductive success. We built a zero-inflated poisson model using the package 'glmmADMB' (Skaug et al., 2015) with annual reproductive success (assessed as the number of six-month-old fledglings produced in a given year) as a count response variable (n = 90-111), and the colorimetric variable(s), tarsus length, group size, territory quality and the cumulative rainfall over the past year as fixed effects. Fixed and random effects were fitted in the count part of the model, and an intercept was fitted for the binomial part of the model (i.e. the zero-inflation did not depend on any explanatory variable). Incestuous pairs (n = 7) were excluded as hatching success in those is known to be much lower (> 30% of reduction compared to non-incestuous pairs; Kingma et al., 2013).

(c) Annual survival and remaining lifespan

To assess potential survival costs of colour elaboration, we related plumage reflectance to both annual survival and remaining lifespan. For both models, we fitted the colorimetric variable(s), age, social status, tarsus length, group size, territory quality and cumulative rainfall over the past year as fixed effects. Year was not included as a random effect in the annual survival model associated with the black cheek as it generated convergence issues.

To test for correlations with individual annual survival, we built GLMMTMBs (overdispersion due to the relatively small proportion of individuals who died) with annual survival (survived until the end of the year = 1, died before the end of the year = 0) as a binomial response variable. For both the purple crown and black cheek models, explanatory variables were scaled to allow the model to converge. To investigate correlations with remaining lifespan among birds that had died by the end of the study, we built a zero-inflated poisson model with remaining lifespan in years as a count response variable.

Results

Chromatic variability

Male purple breeding colouration is characterised by a multi-peaked spectrum, as observed to a lesser extent for the blue tail colouration (Fig. 2a). For both colours, the PCA indicates a complex pattern of chromatic variation with two main principal components: purple PC1 (range of 13 jnd) and purple PC2 (range of 10 jnd) explain 76% and 20% of variation respectively, while blue PC1 (range of 12 jnd) and blue PC2 (range of 5 jnd) explain 81% and 18% of variation respectively (Fig. 2c, Table S1). The loadings of each component indicate that, compared to purple PC1 and PC2, blue PC1 and PC2 are oriented differently in the visual space, i.e. they stimulate the bird vision cones differently (Fig. 2c, Table S1). Higher values of purple PC1 are associated with spectra with higher reflectance in the shorter wavelengths (UV,

blue) relative to longer wavelengths (green, red), and higher values of purple PC2 with spectra with higher red reflectance relative to shorter wavelengths (green) and simultaneously higher blue reflectance relative to UV, i.e. overall more purple spectra (Fig. S2). In contrast, higher values of blue tail PC1 correspond to spectra with lower red reflectance relative to shorter wavelengths (UV, blue and green), and higher values of PC2 to spectra with higher reflectance in the UV, blue and red ranges relative to green and yellow reflectance (Fig. S3).

Both buff-white throat and brown back colouration are characterised by a reflectance spectrum with no peak (Fig. 2b). In both cases, chromatic variation is essentially along one principal component axis that explains ≥ 97% of variation (Fig. 2d, Table S1). Visualisation of the spectra in the bird visual space shows that the chromatic variation of the brown colouration encompasses the chromatic variation of the buff-white colouration, but exhibits greater variation in one direction (range of 4 jnd and 10 jnd for buff-white PC1 and brown PC1 respectively; Fig. 2d). In line with this, the loadings of PC1 are very similar for both colours (Table S1), and higher values of PC1 are associated with higher UV reflectance and lower red reflectance (Fig. S4).

Comparison of chromatic variability among the four patch colours showed that variability significantly decreases in the following order: purple > blue > brown > buff-white (Table S2b). Additionally, the ornamental colours (purple and blue) display higher chromatic variability compared to the non-ornamental colours (buff-white and brown; Table S2c).

Achromatic (brightness) variability

Purple, blue, buff-white and brown L show variations in the population close to 10 jnd. In contrast, black L show a substantially larger variation of 40 jnd. Comparison of achromatic variability among the five patch colours showed that variability decreases in the following order: black > blue > purple, brown and buff-white (the last three colours are similarly variable;

Table S2b). Additionally, the ornamental colours (purple, black and blue) display higher achromatic variability compared to the non-ornamental colours (buff-white and brown; Table S2c).

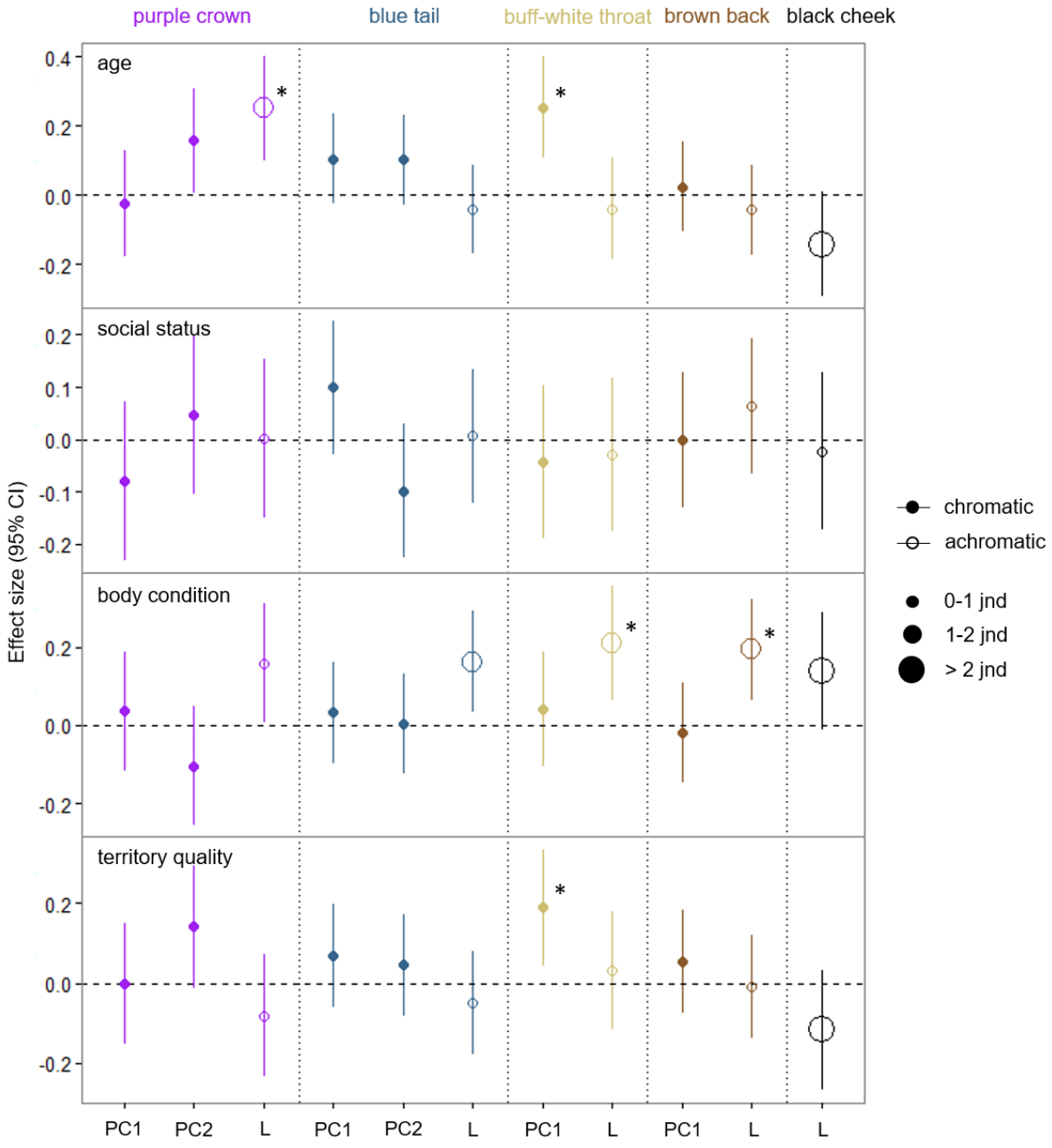


Figure 3. Male plumage brightness varies with age and body condition, while plumage chromatic variation is little or not affected by intrinsic and environmental factors. Shown are the effect size estimates of correlations between age class (1- vs. 2+-year-old), social status (subordinate vs. dominant), body condition and territory quality and the chromatic (PC1, PC2)

and achromatic (L) components of the purple crown, blue tail, buff-white throat, brown back and black cheek colouration in male purple-crowned fairy-wrens. Circle area depicts whether the variation in PCs and L across the range of realistic values of the predictors is between 0 and 1 jnd (i.e. not discriminable by conspecifics), between 1 and 2 jnd, or larger than 2 jnd, and error bars the 95% confidence intervals. Asterisks depict significant correlations after running the FDR procedure. Purple: higher PC1 = higher UV/blue reflectance relative to green/red, higher PC2 = higher red reflectance relative to green and higher blue reflectance relative to UV, i.e. more purple overall; blue: higher PC1 = lower red reflectance relative to UV, blue and green, higher PC2 = higher UV, blue and red reflectance relative to green/yellow; buff-white and brown: higher PC1 = higher UV and lower red reflectance.

Intrinsic and environmental effects on plumage reflectance

No aspects of chromatic variation of the purple crown, blue tail, buff-white throat and brown back colouration show significant association with any of the intrinsic and environmental variables that we investigated, namely age class (1- or 2-year-old and older), social status (subordinate or dominant), body condition, group size, territory quality and rainfall accumulated over the previous year (Fig. 3, Tables S3-S7). Some correlations were detected for a few chromatic variables, but only two of them remain statistically significant after running the false discovery rate (FDR) procedure – between buff-white PC1 and age class, as well as territory quality – and the associated biological effects are all small (variation below the discrimination threshold of 1 jnd; Fig. 3, Tables S3-S7 and S15). The random effect 'year' explains up to 16% of the phenotypic variance of purple and blue PC1 and PC2, while it explains 35% and 44% of the phenotypic variance of buff-white PC1 and brown PC1 respectively (Table S8), indicating substantially larger annual variations in the non-ornamental colours.

In contrast, achromatic brightness (L) of the different patch colours is related to age and/or body condition. Purple L is positively correlated with age, with 2-year-old and older males producing brighter crowns than 1-year-old males (Fig. 3, Table S3). This difference is quite small (~ 1 jnd), but the effect appears robust as it holds after running the FDR procedure (Table S15). Conversely, black L shows a negative trend with age, with 2-year-old and older males producing darker cheeks than 1-year-old males; this difference is large but marginally non-significant and does not remain after running the FDR procedure (Fig. 3, Tables S7 and S15). In addition, the brightness of all colours exhibits a positive association with body condition (marginally non-significant for the black cheek; Fig. 3, Tables S3-S7): males in better condition display a brighter plumage overall. While these differences are large for the black colouration, they are much smaller and only discriminable between individuals with the lowest and highest values of body condition for the other colours (just below the discrimination threshold for purple; Fig. 3; Tables S3-S7). However, only the associations with the non-ornamental colours (buff-white and brown L) remain after running the FDR procedure (Table S15). For all colours, the 'year' effect explains between 15% and 28% of the phenotypic variance of L (Table S8), indicating moderate annual variation in overall plumage brightness.

Heritability of plumage reflectance

Purple PC2, blue PC1 and blue PC2 show moderate to substantial heritability estimates (between 0.19 and 0.30 using MCMCglmm, and between 0.27 and 0.37 using ASReml-R; Fig. 4, Table S11), with moderate to substantial estimates of additive genetic variance for purple PC2 and blue PC1 (0.18 and 0.55 respectively; Table S11). In contrast, the heritability estimates of all the other chromatic and achromatic variables are rather low (< 0.12, except for purple L with an estimate of 0.19 when using ASReml-R; Fig. 4. Table S11). We note however

that the credible intervals (and standard errors) for some estimates are very broad (Fig. 4, Table S11) and therefore it is difficult to make confident inferences about their exact values.

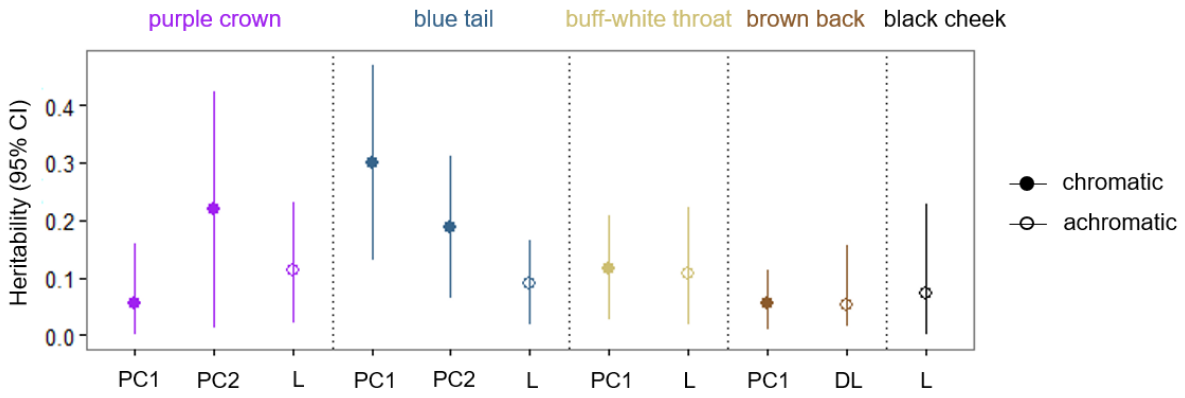


Figure 4. Some chromatic components of the purple crown and blue tail colouration of male purple-crowned fairy-wrens show moderate to substantial heritability. Shown are the posterior mean values of heritability of the chromatic (PC1, PC2) and achromatic (L) components of the purple crown, blue tail, buff-white throat, brown back and black cheek colouration in male purple-crowned fairy-wrens. Error bars depict the 95% credible intervals.

Fitness benefits and costs associated with plumage reflectance

We found a positive association between purple PC2 and annual reproductive success, with males displaying more purple crowns siring more 6-month-old offspring in a given year – and this effect remains after running the FDR procedure (Figs 5 and 6, Tables S13 and S15). Black L is also positively related to annual reproductive success, with males displaying brighter black cheeks producing more 6-month-old offspring in a given year, but this effect does not hold using the FDR procedure (Fig. 5, Tables S13 and S15). In addition, we found a statistically significant positive association between blue L and the probability of acquiring a breeder position elsewhere, as well as a negative association between blue PC2 and annual survival (Fig. 5, Tables S12 and S14). Only the former remains after running the FDR procedure (Table S15) and the effect appears to be weak (i.e. males with low and high values of blue L have very

similar predicted probabilities of becoming a breeder). No significant association was detected between any other colorimetric variable and the probability of acquiring a breeder position, annual reproductive success, annual survival and remaining lifespan (Tables S12-S14).

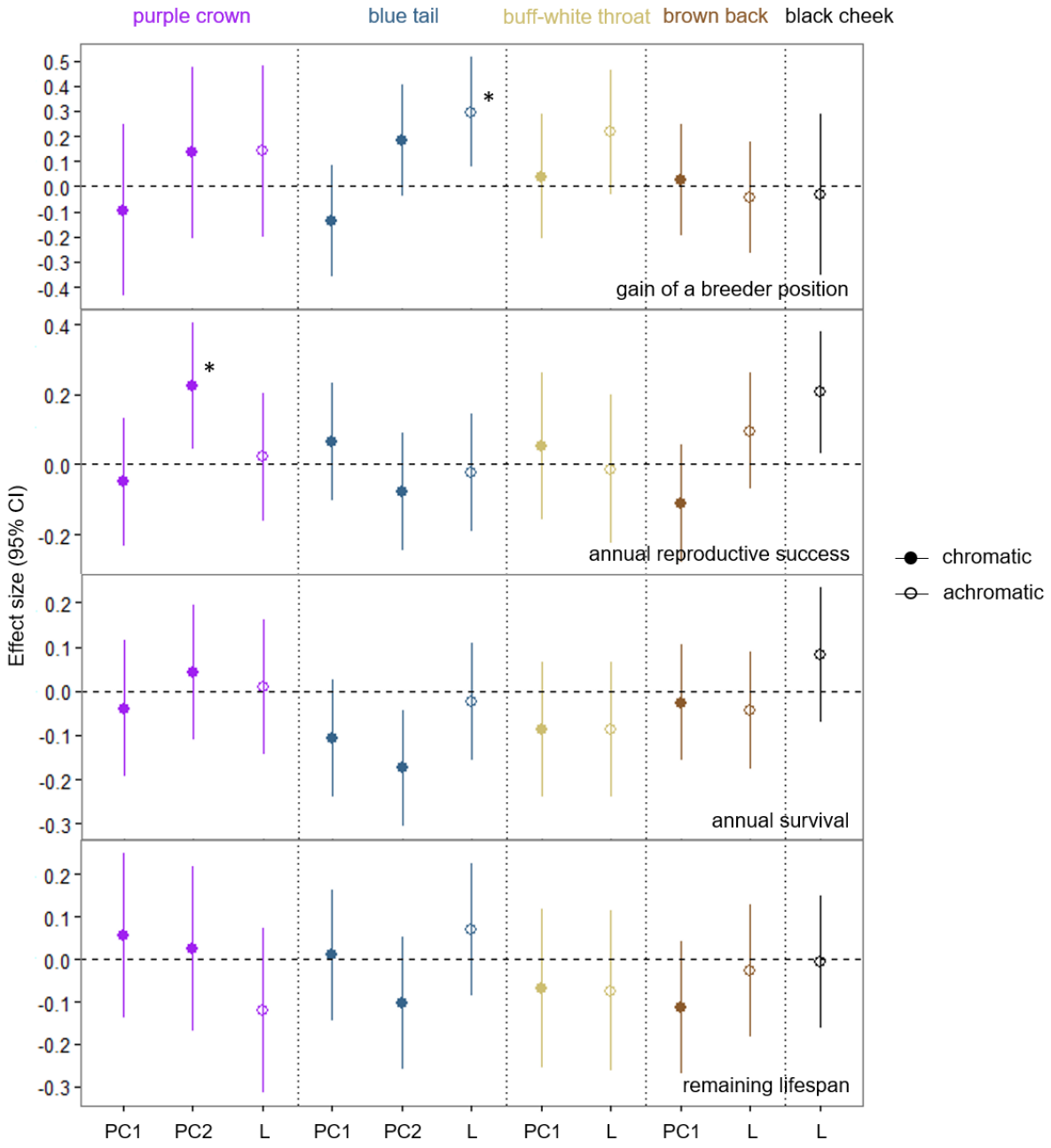


Figure 5. The quality of the purple breeding plumage is related to male annual reproductive success, but not to other fitness indicators. Shown are the effect size estimates of correlations between the probability of acquiring a breeder position elsewhere, annual reproductive success (i.e. number of 6-month-old offspring produced), annual survival and remaining lifespan and

the chromatic (PC1 and PC2) and achromatic (L) components of the purple crown, blue tail, buff-white throat, brown back and black cheek colouration in male purple-crowned fairy-wrens. Error bars depict the 95% confidence intervals. Asterisks depict significant correlations after running the FDR procedure.

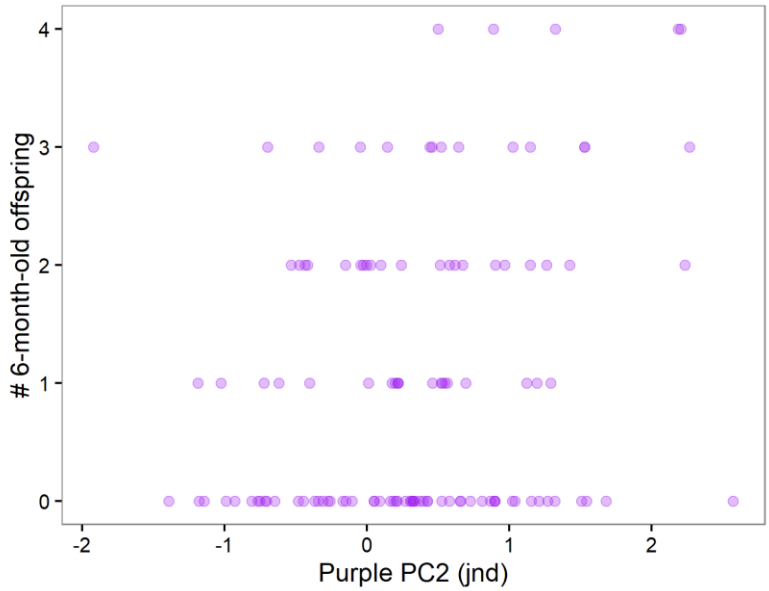


Figure 6. Dominant breeder males that display more purple crowns achieve greater annual reproductive success. Shown is the number of 6-month-old offspring produced in a given year (raw data) as a function of purple PC2 ($\beta = 0.32 \pm 0.13$, $z = 2.44$, $P = 0.02$).

Discussion

In this study, we examined the condition-dependence, heritability and fitness correlates of the multidimensional colour phenotype of male *M. coronatus*, which includes both ornamental and non-ornamental colours. Our results indicate that (1) although there is limited support for heightened condition-dependence of ornamental colours compared to non-ornamental colours, we found some evidence for (2) higher levels of heritability and (3) associations with male fitness for some ornamental colours. Specifically, seasonal male purple breeding colouration

appeared to show high phenotypic variation, heritable variation, and an association with male reproductive success.

Condition-dependence

Ornamental traits subject to directional selection often show high variability (but see Evans & Barnard, 1995; Reinhold, 2011), an important feature that can inform about differences in quality among signallers. Strong directional selection should, however, rapidly deplete the underlying genetic variation and lead to reduced phenotypic variability (Taylor & Williams, 1982; Merilä & Sheldon, 1999). Fluctuations in selection over time and space – affecting its strength, form and direction – possibly account for some of the persistent variability (Dale, 2006; Chaine & Lyon, 2008); however, the best supported hypothesis is heightened sensitivity to condition (Rowe & Houle, 1996). Our data provide support for greater levels of variability in the putative ornamental colours of male *M. coronatus* compared to the non-ornamental colours: the purple and blue plumage displayed higher chromatic variability, while the black plumage showed higher achromatic variability than any other colour (Table S2, Fig. 2). This is in line with the study of Delhey et al. (2017) which, based on more than 100 plumage colours across 55 bird species, demonstrated that more conspicuous colours display greater variability. Furthermore, the purple and blue colouration are characterised by complex patterns with two axes of chromatic variation (Fig. 2), contrasting with most plumage colours that generally show a single axis of chromatic variation (Endler et al., 2005; Delhey et al., 2010, 2013, 2014).

However, there was no clear indication that condition-dependence was stronger in ornamental colours: only the brightness of the buff-white and brown plumage was correlated with male body condition, while none of the other colours were related to condition (Fig. 3). Males in better body condition display brighter buff-white and brown patches than those in poorer condition, and the strength of this effect was similar for both colours (although only

discriminable between individuals in very different condition; Fig. 3). A similar positive association between individual condition and brightness of melanin-based patches was previously reported in several passerine species, and linked to changes in feather microstructure (barbule density and barb thickness; D'Alba et al., 2014). Possibly, in male *M. coronatus*, body condition also influences the growth of feather microstructure in buff-white and brown patches (both melanin-based), which affects their brightness in a similar way. Nevertheless, we must interpret this result with caution as our test of condition-dependence presents some weaknesses: to compute body condition we measured body mass at the time of colour measurement, when males were possibly at different stages of moult (during the pre- or post-breeding moult, or in between). As the moulting process may entail substantial energetic costs (Lindström et al., 1993), this might have impacted our assessment of body condition.

Although there was no evidence for a link between ornamental colouration and condition, the brightness of the purple breeding colouration appeared to increase with male age. That 2-year-old and older males produce brighter purple crowns compared to first-year males, may represent an age-dependent investment in signals, where the production of costly traits is delayed until older ages to indicate higher viability at that point (Kokko, 1997). Age may function as a signal of (genetic) quality since an older age necessarily indicates the ability of an individual to survive up to that age at least (Trivers, 1972; Manning, 1985). Honest signalling should favour an increase in the level of sexual advertisement over time and therefore preferences for older males (Brooks & Kemp, 2001; Proulx et al., 2002). Although the average difference in brightness between first-year males and older males is relatively small (~1 jnd; Fig. 3), age-dependence of the purple breeding colouration could reflect honest signalling of male quality in *M. coronatus*.

Decomposition of the total phenotypic variance of all colour components showed significant variation between years that appeared generally larger for the buff-white and brown

colouration, especially for their chromatic components. Such an annual effect may be driven by large-scale environmental influences (Garant et al., 2004; Masello et al., 2008; Evans & Sheldon, 2011) during the period of moult, such as wet spells and drought. The stronger effect observed on the non-ornamental colours suggests that their expression may be less robust against environmental variation, possibly because they are less or not important for fitness, contrary to ornamental colours, as proposed by the 'canalisation hypothesis' (Waddington, 1942, 1959). According to this hypothesis, the effect of an environmental perturbation is stronger on traits where deviations from the optimal trait value are less costly in fitness terms, i.e. when they are less well canalised. Robustness to perturbations may be achieved by preferentially allocating resources to better canalised traits at the expense of less canalised traits (Stearns et al., 1995; Liefting et al., 2015).

Overall, our data suggest that in male *M. coronatus*, changes in the brightness of the purple ornamental plumage may serve to distinguish older males from first-year males, and brightness of the non-ornamental, dull plumage uniformly reflects male condition. Hence, there is limited support for the prediction that ornamental colours exhibit greater condition-dependence in comparison with non-ornamental colours.

Heritability

When assessing the additive genetic contributions to the different colours, we found that some chromatic components of the ornamental seasonal purple crown and sexually dichromatic blue tail colouration (purple PC2 and blue PC1) exhibited greater levels of additive genetic variance (Table S11) and were also more heritable than any other colour components, albeit not highly heritable (Fig. 4). This suggests that some level of genetic variance underlies these ornamental colours and as a result, their expression may convey some information about the bearer's genetic quality that can be used by conspecifics (see also discussion of the link to fitness

below). Further work is however required to determine more specifically the nature of the genetic aspects that may be signalled. The observed differences in heritability could also be driven by differences in the underlying mechanisms of colour production: structural colours (i.e. purple and blue) showed higher heritable variation than melanin-based colours (i.e. black and brown). However, this is in disagreement with previous studies that support a tight genetic control of melanin-based colouration (heritability estimates between 0.53 and 1.0, based on studies focusing mostly on polymorphic species; Roulin & Ducrest, 2013), which has been linked to the endogenous production of melanin. In contrast, knowledge about the quantitative genetics of bird structural colouration is quasi non-existent, but a study by Hadfield et al. (2007) in blue tits showed little heritability of the structural blue patches (0.12 and 0.05 for the cap and primary coverts respectively), while Charmantier et al. (2017) observed in the same species variable heritability of the chromatic and achromatic parts of the blue crown colouration (between 0.06 and 0.23, depending on the subspecies and sex). Therefore, our findings clearly highlight the need for more studies on the genetic basis of feather structural colouration. Even so, it should be noted that, unfortunately, we found much uncertainty in our estimates of additive genetic variance and heritability and hence, we cannot draw firm conclusions about the exact magnitude of heritable variation for the different colour traits.

Fitness benefits: quality of the purple breeding colouration and male reproductive success

While the extent of breeding colouration is a strong predictor of success when subordinate males compete for breeding vacancies (Fan et al., 2018 / Chapter 3), our present findings indicate that the quality of the breeding colours is not important for obtaining a breeding position. However, it seems that the purple plumage reflectance correlates with male reproductive success once males have acquired a breeder position (Figs 5 and 6). Specifically, males with more purple crowns (higher purple PC2) achieve greater annual reproductive

success (assessed as the number of 6-month-old offspring produced; Figs 5 and 6). Such a pattern may reflect different reproductive tactics by differently coloured males (Duckworth et al., 2003) or differential investment by the social partner (Horváthová et al., 2012). Purple PC2 shows some heritable variation (Fig. 4) and possibly acts as an indicator of male genetic quality, potentially influencing parental investment from either or both partners and/or offspring viability (Wolf et al., 1997; Horváthová et al., 2012). Possibly, more purple males are able to provide better parental care or produce offspring of higher fitness (Wolf et al., 1997). Alternatively, females paired with more purple males may increase their reproductive effort, e.g. through larger clutch sizes, increased incubation effort or feeding rates (Horváthová et al., 2012). As we have little knowledge about how male plumage colour relates to reproductive and parental investment, such hypotheses – that are not mutually exclusive – remain to be tested. Nonetheless, our data suggest that the purple breeding colouration of male *M. coronatus* may have an additional signalling function in dominant breeder males, which, through mechanisms that are yet to be uncovered, impacts male reproductive output.

Conclusion

Using psychophysical models of colour perception, we performed a comprehensive study of the information content, heritability and fitness correlates of all ornamental and non-ornamental colour traits displayed in a single species. Overall, we found limited support for the hypothesis that ornamental colours are more costly or more strongly correlated with individual quality and fitness. There was some evidence for greater levels of heritability for some ornamental colours, although the degree of uncertainty in our estimates warrants some caution in the interpretation of these results. Nevertheless, the main ornamental colour, i.e. the purple breeding colouration, appeared to meet most of our expectations, and in particular correlate with male reproductive success. Combined with previous studies on the adaptive function of

male breeding plumage in *M. coronatus*, our data illustrate the complexity of such a seasonal plumage, as the timing of acquisition and extent of breeding colouration constitute other dimensions to be considered in addition to the quality of the colour itself. More generally, through the integral assessment of multiple ornamental colours, in comparison with multiple non-ornamental colours, our study demonstrates that some traits may not follow any clear pattern, therefore challenging the existing theories regarding the evolution of ornamental traits, and calling for more studies using an integrated approach. This will help to provide a more complete picture, although not always easy to interpret, of why and how complex animal signals might evolve.

Acknowledgements

This research was approved by the Australian Bird and Bat Banding Scheme (license 2230); Western Australia Department of Parks and Wildlife; Australian Wildlife Conservancy; and the ethics committees of the School of Biological Sciences at Monash University and the Max Planck Institute for Ornithology. Our thanks to numerous volunteers during fieldwork and the support of staff at Australian Wildlife Conservancy's Mornington Wildlife Sanctuary. We are grateful for support by the Max Planck Society, the Australian Research Council (FT10100505 & DP150103595), the Holsworth Wildlife Research Endowment – Equity Trustees Charitable Foundation / Ecological Society of Australia, and the Australian Wildlife Conservancy and its supporters.

References

Andersson, M. (1994). *Sexual selection. Princeton*. Princeton University Press.

Andersson, M. (2006). Condition-dependent indicators in sexual selection: development of theory and tests. In J. R. Lucas & L. W. Simmons (Eds.), *Essays in Animal Behaviour: Celebrating 50 Years of Animal Behaviour* (pp. 255-269). Elsevier Academic Press.

Badyaev, A. V. & Hill, G. E. (2000). Evolution of sexual dichromatism: contribution of carotenoid- versus melanin-based colouration. *Biological Journal of the Linnean Society*, **69**, 153-172.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1-48.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 289-300.

Bonduriansky, R., & Rowe, L. (2005). Sexual selection, genetic architecture, and the condition dependence of body shape in the sexually dimorphic fly *Prochyliza xanthostoma* (Piophilidae). *Evolution*, **59**, 138-151.

Brooks, R., & Kemp, D. J. (2001). Can older males deliver the good genes? *Trends in Ecology & Evolution*, **16**, 308-313.

Burley, N. (1988). The differential-allocation hypothesis: an experimental test. *The American Naturalist*, **132**, 611-628.

Butler, D. (2009). ASReml R package version 3.0. VSN International Ltd. Hemel Hempstead, UK.

Cassey, P., Ewen, J., Blackburn, T., Hauber, M., Vorobyev, M., & Marshall, N. (2008). Eggshell colour does not predict measures of maternal investment in eggs of *Turdus* thrushes. *Naturwissenschaften*, **95**, 713-721.

Chaine, A. S., & Lyon, B. E. (2008). Adaptive plasticity in female mate choice dampens sexual selection on male ornaments in the lark bunting. *Science*, **319**, 459-462.

Charmantier, A., Wolak, M. E., Grégoire, A., Fargevieille, A., & Doutrelant, C. (2017). Colour ornamentation in the blue tit: quantitative genetic (co) variances across sexes. *Heredity*, **118**, 125.

Cotton, S., Fowler, K., & Pomiankowski, A. (2004). Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? *Proceedings of the Royal Society of London B: Biological Sciences*, **271**, 771-783.

Cuthill, I. C. (2006). Color perception. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1*. (pp. 3-40). Harvard University Press.

D'Alba, L., Van Hemert, C., Spencer, K. A., Heidinger, B. J., Gill, L., Evans, N. P., ... & Shawkey, M. D. (2014). Melanin-based color of plumage: role of condition and of feathers' microstructure. *Integrative and Comparative Biology*, **54**, 633-644.

Dale, J. (2006). Intraspecific variation in bird coloration. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1*. (pp. 36-86). Harvard University Press.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. John Murray.

Delhey, K., Burger, C., Fiedler, W., & Peters, A. (2010). Seasonal changes in colour: a comparison of structural, melanin-and carotenoid-based plumage colours. *PLoS One*, **5**, e11582.

Delhey, K., Hall, M., Kingma, S. A., & Peters, A. (2013). Increased conspicuousness can explain the match between visual sensitivities and blue plumage colours in fairy-wrens. *Proceedings of the Royal Society of London B: Biological Sciences*, **284**, 20121771.

Delhey, K., Delhey, V., Kempenaers, B., & Peters, A. (2015). A practical framework to analyze variation in animal colors using visual models. *Behavioral Ecology*, **26**, 367-375.

Delhey, K., Szecsenyi, B., Nakagawa, S., & Peters, A. (2017). Conspicuous plumage colours are highly variable. *Proceedings of the Royal Society of London B: Biological Sciences*, **284**, 20162593.

Duckworth, R. A., Badyaev, A. V., & Parlow, A. F. (2003). Elaborately ornamented males avoid costly parental care in the house finch (*Carpodacus mexicanus*): a proximate perspective. *Behavioral Ecology and Sociobiology*, **55**, 176-183.

Dunn, P. O., Whittingham, L. A., & Pitcher, T. E. (2001). Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution*, **55**, 161-175.

Endler, J. A., & Mielke, P. W. (2005). Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, **86**, 405-431.

Endler, J. A., Westcott, D. A., Madden, J. R., & Robson, T. (2005). Animal visual systems and the evolution of color patterns: sensory processing illuminates signal evolution. *Evolution*, **59**, 1795-1818.

Evans, M. R., & Barnard, P. (1995). Variable sexual ornaments in scarlet-tufted malachite sunbirds (*Nectarinia johnstoni*) on Mount Kenya. *Biological Journal of the Linnean Society*, **54**, 371-381.

Evans, S. R., & Sheldon, B. C. (2011). Quantitative genetics of a carotenoid-based color: heritability and persistent natal environmental effects in the great tit. *The American Naturalist*, **179**, 79-94.

Fan, M., Hall, M. L., Kingma, S. A., Mandeltort, L. M., Hidalgo Aranzamendi, N., Delhey, K., & Peters, A. (2017). No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy? *Behavioral Ecology*, **28**, 1055-1067.

Fan, M., Teunissen, N., Hall, M. L., Hidalgo Aranzamendi, N., Kingma, S. A., Roast, M., ... & Peters, A. (2018). From ornament to armament or loss of function? Breeding plumage acquisition in a genetically monogamous bird. *Journal of Animal Ecology*, **87**, 1274-1285.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Clarendon Press.

Garant, D., B. C. Sheldon, & L. Gustafsson. (2004). Climatic and temporal effects on the expression of secondary sexual characters: genetic and environmental components. *Evolution*, **58**, 634-644.

Gosden, T. P., & Chenoweth, S. F. (2011). On the evolution of heightened condition dependence of male sexual displays. *Journal of Evolutionary Biology*, **24**, 685-692.

Hadfield, J. D., Nutall, A., Osorio, D., & Owens, I. P. F. (2007). Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *Journal of Evolutionary Biology*, **20**, 549-557.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**, 1-22.

Hall, M. L., & Peters, A. (2008). Coordination between the sexes for territorial defence in a duetting fairy-wren. *Animal Behaviour*, **76**, 65-73.

Hall, M. L., & Peters, A. (2009). Do male paternity guards ensure female fidelity in a duetting fairy-wren? *Behavioral Ecology*, **20**, 222-228.

Hall, M. L., Kingma, S. A., & Peters, A. (2013). Male songbird indicates body size with low-pitched advertising songs. *PLoS One*, **8**, e56717.

Hamilton, W. D., & Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites? *Science*, **218**, 384-387.

Hart, N. S. (2001). Variations in cone photoreceptor abundance and the visual ecology of birds. *Journal of Comparative Physiology A*, **187**, 685-697.

Hart, N. S., & Hunt, D. M. (2007). Avian visual pigments: characteristics, spectral tuning, and evolution. *The American Naturalist*, **169**(S1), S7-S26.

Hegyi, G., Szigeti, B., Török, J., & Eens, M. (2007). Melanin, carotenoid and structural plumage ornaments: information content and role in great tits *Parus major*. *Journal of Avian Biology*, **38**, 698-708.

Hidalgo Aranzamendi, N., Hall, M. L., Kingma, S. A., Sunnucks, P., & Peters, A. (2016). Incest avoidance, extrapair paternity, and territory quality drive divorce in a year-round territorial bird. *Behavioral Ecology*, arw101.

Hill, G. E., & Brawner, W. R. (1998). Melanin-based plumage colouration in the house finch is unaffected by coccidial infection. *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 1105-1109.

Hill, G. E. (2002). *A red bird in a brown bag: the function and evolution of colorful plumage in the house finch*. Oxford University Press.

Hill, G. E. (2006). Environmental regulation of ornamental coloration. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1*. (pp. 507-560). Harvard University Press.

Horváthová, T., Nakagawa, S., & Uller, T. (2012). Strategic female reproductive investment in response to male attractiveness in birds. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 163-170.

Johnstone, R. A., Rands, S. A., & Evans, M. R. (2009). Sexual selection and condition-dependence. *Journal of Evolutionary Biology*, **22**, 2387-2394.

Kemp, D. J., & Rutowski, R. L. (2007). Condition dependence, quantitative genetics, and the potential signal content of iridescent ultraviolet butterfly coloration. *Evolution*, **61**, 168-183.

Keyser, A. J., & Hill, G. E. (2000). Structurally based plumage coloration is an honest signal of quality in male blue grosbeaks. *Behavioral Ecology*, **11**, 202-209.

Kingma, S. A., Hall, M. L., Segelbacher, G., & Peters, A. (2009). Radical loss of an extreme extra-pair mating system. *BMC Ecology*, **9**, 15.

Kingma, S. A., Hall, M. L., Arriero, E., & Peters, A. (2010). Multiple benefits of cooperative breeding in purple-crowned fairy-wrens: A consequence of fidelity? *Journal of Animal Ecology*, **79**, 757-768.

Kingma, S. A., Hall, M. L., & Peters, A. (2011a). No evidence for offspring sex-ratio adjustment to social or environmental conditions in cooperatively breeding purple-crowned fairy-wrens. *Behavioral and Ecological Sociobiology*, **65**, 1203-1213.

Kingma, S. A., Hall, M. L., & Peters, A. (2011b). Multiple benefits drive helping behavior in a cooperatively breeding bird: an integrated analysis. *The American Naturalist*, **177**, 486-495.

Kingma, S. A., Hall, M. L., & Peters, A. (2013). Breeding synchronization facilitates extrapair mating for inbreeding avoidance. *Behavioral Ecology*, **24**, 1390-1397.

Kokko H. 1997. Evolutionarily stable strategies of age-dependent sexual advertisement. *Behavioral and Ecological Sociobiology*, **41,** 99-107.

Kotiaho, J. S., Simmons, L. W., & Tomkins, J. L. (2001). Towards a resolution of the lek paradox. *Nature*, **410**, 684.

Kotiaho, J. S., LeBas, N. R., Puurtinen, M., & Tomkins, J. L. (2008). On the resolution of the lek paradox. *Trends in Ecology & Evolution*, **23**, 1-3.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmerTest'. R version 2.0-33.

Liefting, M., van Grunsven, R. H., Morrissey, M. B., Timmermans, M. J., & Ellers, J. (2015). Interplay of robustness and plasticity of life-history traits drives ecotypic differentiation in thermally distinct habitats. *Journal of Evolutionary Biology*, **28**, 1057-1066.

Lind, O., Chavez, J., & Kelber, A. (2014). The contribution of single and double cones to spectral sensitivity in budgerigars during changing light conditions. *Journal of Comparative Physiology A*, **200**, 197-207.

Lindström Å., Visser, G. H., & Daan, S. (1993). The energetic cost of feather synthesis is proportional to basal metabolic rate. *Physiology Zoology*, **66**, 490-510.

Lozano, G. A. (1994). Carotenoids, parasites, and sexual selection. *Oikos*, 309-311.

Magnusson, A., Skaug, H. J., Nielsen, A., Berg, C. W., Kristensen, K., Maechler, M., ... & Brooks, M. E. (2017). glmmTMB: Generalized linear mixed models using a template model builder. R package version 0.1, 3.

Manning, J.T. (1985). Choosy females and correlates of male age. *Journal of Theoretical Biology*, **116**, 349-354.

Masello, J. F., T. Lubjuhn, & P. Quillfeldt. (2008). Is the structural and psittacofulvin-based coloration of wild burrowing parrots *Cyanoliseus patagonus* condition dependent? *Journal of Avian Biology*, **39**, 653-662.

McGraw, K. J., Mackillop, E. A., Dale, J., & Hauber, M. E. (2002). Different colors reveal different information: how nutritional stress affects the expression of melanin-and structurally based ornamental plumage. *Journal of Experimental Biology*, **205**, 3747-3755.

McGraw, K. J. (2006) Mechanics of melanin coloration in birds. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 243-294). Harvard University Press.

Merilä, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity*, **83**, 103-109.

Ödeen, A., Pruett-Jones, S., Driskell, A. C., Armenta, J. K., & Håstad, O. (2012). Multiple shifts between violet and ultraviolet vision in a family of passerine birds with associated changes in plumage coloration. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 1269-1276.

Olsson, P., Lind, O., & Kelber, A. (2017). Chromatic and achromatic vision: parameter choice and limitations for reliable model predictions. *Behavioral Ecology*, **29**, 273-282.

Owens, I. P. F., & Hartley, I. R. (1998). Sexual dimorphism in birds: why are there so many different forms of dimorphism? *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 397-407.

Peters, A., Kurvers, R. H. J. M., Roberts, M. L., & Delhey, K. (2011). No evidence for general condition-dependence of structural plumage colour in blue tits: an experiment. *Journal of Evolutionary Biology*, **24**, 976-987.

Peters, A., Kingma, S. A., & Delhey, K. (2013). Seasonal male plumage as a multi-component sexual signal: Insights and opportunities. *Emu*, **113**, 232-247.

Pomiankowski, A., & Møller, A. P. (1995). A resolution of the lek paradox. *Proceedings of the Royal Society of London B: Biological Sciences*, **260**, 21-29.

Price, T., & Schluter, D. (1991). On the low heritability of life-history traits. *Evolution*, **45**, 853-861.

Proulx, S. R., Day, T., & Rowe, L. (2002). Older males signal more reliably. *Proceedings of the Royal Society of London B: Biological Sciences*, **269**, 2291-2299.

Prum, R. O. (2006). Anatomy, physics, and evolution of structural colours. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 295-353). Harvard University Press.

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.

Reinhold, K. (2011). Variation in acoustic signalling traits exhibits footprints of sexual selection. *Evolution: International Journal of Organic Evolution*, **65**, 738-745.

Roulin, A., & Ducrest, A. L. (2013). Genetics of colouration in birds. In *Seminars in cell & developmental biology, vol. 24, no. 6-7.* (pp. 594-608). Academic Press.

Roulin, A. (2016). Condition-dependence, pleiotropy and the handicap principle of sexual selection in melanin-based colouration. *Biological Reviews*, **91**, 328-348.

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London B: Biological Sciences*, **263**, 1415-1421.

Rowley, I., & Russell, E. (1997). *Fairy-wrens and Grasswrens: Maluridae.* Oxford University Press.

Taylor, P. D., & Williams, G. C. (1982). The lek paradox is not resolved. *Theoretical Population Biology*, **22**, 392-409.

Skaug, H., Fournier, D., Nielsen, A., Magnusson, A., & Bolker, B. M. (2015). glmmADMB: Generalized Linear Mixed Models Using AD Model Builder. R package version 0.7. 2.1.

Stearns, S. C., Kaiser, M., & Kawecki, T. J. (1995). The differential genetic and environmental canalization of fitness components in *Drosophila melanogaster*. *Journal of Evolutionary Biology*, **8**, 539-557.

Tibbetts, E. A. (2010). The condition dependence and heritability of signaling and nonsignaling color traits in paper wasps. *The American Naturalist*, **175**, 495-503.

Tomkins, J. L., Radwan, J., Kotiaho, J. S., & Tregenza, T. (2004). Genic capture and resolving the lek paradox. *Trends in Ecology & Evolution*, **19**, 323-328.

Trivers, R.L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual Selection and the Descent of Man 1871–1971*. (pp. 136-179). Heinemann.

Vorobyev, M., Osorio, D., Bennett, A. T., Marshall, N. J., & Cuthill, I. C. (1998). Tetrachromacy, oil droplets and bird plumage colours. *Journal of Comparative Physiology A*, **183**, 621-633.

Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, **150**, 563-565.

Waddington, C. H. (1959). Canalization of development and genetic assimilation of acquired characters. *Nature*, **183**, 1654-1655.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., ... & Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, **79**, 13-26.

Wolf, J. B., Moore, A. J., & Brodie III, E. D. (1997). The evolution of indicator traits for parental quality: the role of maternal and paternal effects. *The American Naturalist*, **150**, 639-649.

Zahavi, A. (1975). Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, **53**, 205-214.

Supplementary Material

Appendix S1: Colour analysis

Before computing average values for each colorimetric variable for each patch colour, we tested whether the multiple measurements taken for each patch at each capture differed among each other. To do this, we built an LMM fitting the colorimetric variable of interest (PC1, PC2 or L of each colour) as the response variable, the measurement number (1-5) as a fixed effect and the unique capture identity as a random intercept. Subsequently, using Tukey's Honest Significant Difference (HSD) test, we performed post hoc comparisons between measurements, which indicated that the differences between measurements (absolute difference noted |d| below) were small and mostly below the discrimination threshold: purple – PC1: $|d| < 0.98$, PC2: $|d| < 1.33$, L: $|d| < 0.48$; blue – PC1: $|d| < 0.28$, PC2: $|d| < 0.36$, L: $|d| < 0.70$; buff-white – PC1: $|d| < 0.39$, L: $|d| < 0.40$; brown – PC1: $|d| < 0.63$, L: $|d| < 0.49$; black – L: $|d| < 0.81$.

Appendix S2: Statistical analyses

(a) Timing of moult of plumage patches other than the head

Body moult, which includes the throat, breast, belly, back and presumably the cheeks, peaks during April and May, while the tail is moulted constantly over the year without a clear peak (Schodde, 1982; Rowley & Russell, 1993; unpublished data).

(b) Absence of effect of time interval since pre-breeding moult completion on the purple-and-black breeding plumage reflectance

We tested whether the time interval since pre-breeding moult completion (i.e. time interval between date of moult completion and date of reflectance measurement) affected the purple

crown and black cheek reflectance. Because the date of pre-breeding moult completion was not known for all individuals in all years, we determined the average date of moult completion at the population level based on pre-breeding moult timing data collected between 2005 and 2010 (for details see Fan et al., 2017 / Chapter 2). We did so for dominants and subordinates separately as the former complete their moult earlier than subordinates (Fan et al., 2017 / Chapter 2). Then, for each male, we calculated the time interval (in days) between the average date of moult completion (corresponding to their current social status) and the date of reflectance measurement.

When including the time since moult completion in the models, the outputs indicated that (i) the time since moult completion had no discriminable effect on purple reflectance (given that the average period of breeding plumage maintenance and post-breeding moult is seven months; unpublished data) whereas the effect of extent of breeding plumage remained (and all other effects were quantitatively similar; Table S9), (ii) the extent of breeding plumage and time since moult completion were not highly correlated ($|r| < 0.30$), and the comparison of the models with and without time since moult completion indicated that (iii) they had similar support (except for purple L, but in both models the statistically significant effects are not biologically discriminable; likelihood-ratio test, purple: PC1, $\chi_1^2 = 1.61$, $P = 0.20$; PC2, $\chi_1^2 = 0.24$, $P = 0.62$; L, $\chi_1^2 = 14.56$, $P < 0.001$; black L: $\chi_1^2 = 0.67$, $P = 0.41$).

(c) Absence of linear effect of age on plumage reflectance

For each patch colour and each chromatic or achromatic variable (L only for the black cheek), we built a linear mixed model (LMM) with the chromatic or achromatic variable as the response variable, age, social status, body condition, group size, territory quality, the cumulative rainfall over the past year and the extent of breeding plumage as fixed effects, and bird identity and year as random intercepts. Only birds whose age was accurately known were

included in the model (n = 81-155 for the different patch colours), thus excluding adult birds born prior to the start of the study (founder population). We first included (1) age in months as a covariate, and then (2) age in years as a fixed factor (levels 1, 2, 3, 4 and 5; the few older individuals were excluded from the analysis). Using Tukey's HSD test, we performed post hoc comparisons between age classes, which indicated that age did not affect any chromatic or achromatic variables (except for buff-white PC1, but the variations with age were below the discrimination threshold; Tables S3-S7).

(d) Calculation of adjusted values of purple PC1 and PC2

For models associated with the purple crown reflectance, because we found that PC1 and PC2 varied with the extent of breeding plumage (most likely due to the presence of brown feathers; see Table S3), we adjusted their values to account for this. Specifically, we (1) built an LMM with average PC1 (respectively PC2) as the response variable, the extent of breeding plumage (i.e. % purple) as a fixed effect, and bird identity and year as random effects; (2) extracted the coefficient estimate β associated with % purple; and (3) adjusted the value of average PC1 (respectively PC2) as follows:

$$\text{adjusted PC1 (or PC2)} = \beta_{\%\ \text{purple}} \times (100 - \%\ \text{purple}) + \text{average PC1 (or PC2)}$$

References

Fan, M., Hall, M. L., Kingma, S. A., Mandeltort, L. M., Hidalgo Aranzamendi, N., Delhey, K., & Peters, A. (2017). No fitness benefits of early molt in a fairy-wren: relaxed sexual selection under genetic monogamy? *Behavioral Ecology*, **28**, 1055-1067.

Rowley, I., & Russell, E. (1993). The purple-crowned fairy-wren *Malurus coronatus*. II. Breeding biology, social organization, demography and management. *Emu*. **93**, 235-250.

Schodde, R. (1982). *The Fairy-wrens: a monograph of the Maluridae.* Lansdowne Editions.

Table S1. Different plumage colours show different patterns of chromatic variation in male *M. coronatus*. Shown are the results of principal component analyses (component loadings, standard deviations and variances) performed on all the reflectance measurements of the purple crown, blue tail, buff-white throat and brown back respectively, collected in male purple-crowned fairy-wrens over eight years. The x-coordinate represents stimulation of the *VS* cone relative to the *S* cone, the y-coordinate represents stimulation of the *M* cone relative to *VS* and *S* cones, and the z-coordinate represents stimulation of the *L* cone compared with the other three (units = jnd).

	Purple crown (*n* = 1061)			Blue tail (*n* = 1803)			Buff-white throat (*n* = 1096)			Brown back (*n* = 1827)		
	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
x	0.11	0.40	0.91	-0.07	0.14	0.99	0.41	0.65	-0.64	0.33	0.62	-0.71
y	-0.98	-0.14	0.17	-0.42	0.90	0.10	-0.67	-0.26	-0.70	-0.66	-0.38	-0.65
z	-0.19	0.91	-0.38	-0.90	0.41	-0.12	-0.62	0.71	0.33	-0.67	0.68	0.29
Standard deviation	1.09	0.56	0.24	0.74	0.36	0.07	0.63	0.08	0.04	1.67	0.30	0.09
Proportion of variance	0.76	0.20	0.04	0.81	0.18	0.01	0.98	0.02	0.00	0.97	0.03	0.00
Cumulative proportion	0.76	0.96	1.00	0.81	0.99	1.00	0.98	1.00	1.00	0.97	1.00	1.00

Table S2. Ornamental colours of male *M. coronatus* show higher chromatic and achromatic variability than non-ornamental colours. Shown are results from LMs examining the effects of (a, b) patch colour (reference: *purple) and (c) type of colour (reference: †non-ornamental) on chromatic variability (calculated as the *log*-transformed distance to the colour centroid in the bird visual space; black excluded; n = 5787) and achromatic variability (calculated as the *log*-transformed distance to the colour average brightness; n = 6987). Levels of significance are specified as follows: *P < 0.05, and ***P < 0.001 (in bold). NA = not applicable.

(a)

	Chromatic variability			Achromatic variability		
Fixed effects	β	SE	t	β	SE	t
Intercept	0.57	0.02	24.84***	-0.43	0.04	-12.72***
Colour*						
Black	NA	NA	NA	**1.54**	**0.05**	**32.02*****
Blue	**-0.44**	**0.03**	**-15.21*****	**0.18**	**0.04**	**4.19*****
Buff-white	**-1.54**	**0.03**	**-47.70*****	-0.05	0.05	0.32
Brown	**-0.52**	**0.03**	**-17.90*****	**-0.10**	**0.04**	**-2.25***
Adjusted R^2		0.30			0.21	

(b)

	Chromatic variability			Achromatic variability		
Pairwise comparisons	β	SE	t	β	SE	t
blue vs. purple	**-0.44**	**0.03**	**-15.21*****	**0.18**	**0.04**	**4.19*****
buff-white vs. purple	**-1.54**	**0.03**	**-47.70*****	-0.05	0.05	-0.99
brown vs. purple	**-0.52**	**0.03**	**-17.90*****	-0.10	0.04	-2.25
buff-white vs. blue	**-1.10**	**0.03**	**-38.26*****	**-0.23**	**0.04**	**-5.34*****
brown vs. blue	**-0.08**	**0.02**	**-3.09***	**-0.28**	**0.04**	**-7.50*****
brown vs. buff-white	**1.02**	**0.03**	**35.68*****	-0.05	0.04	-1.16
black vs. purple	NA	NA	NA	**1.54**	**0.05**	**32.02*****
blue vs. black	NA	NA	NA	**-1.35**	**0.04**	**-31.87*****
buff-white vs. black	NA	NA	NA	**-1.59**	**0.05**	**-33.31*****
brown vs. black	NA	NA	NA	**-1.64**	**0.04**	**-38.65*****

(c)

	Chromatic variability			Achromatic variability		
Fixed effects	β	SE	t	β	SE	t
Intercept	0.32	0.02	24.84***	-0.51	0.02	-21.99***
Type†	**0.62**	**0.02**	**28.18*****	**0.62**	**0.03**	**20.42*****
Adjusted R^2		0.12			0.06	

Table S3. Older males produce a brighter purple crown colouration compared to first-year males. Additionally, males in better body condition produce a brighter purple crown colouration than males in worse condition, but this effect is not discriminable by conspecifics. Shown are results from LMMs examining the effects of age, social status (reference: [†]subordinate), body condition, group size, territory quality, cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) and % purple on average PC1, PC2 and L of the purple crown colouration in (a, b) individuals of known age only, considering age (a) in months as a covariate (n = 81) and (b) in years as a cofactor (levels 1, 2, 3, 4, 5; n = 13, 23, 21, 11, 10, respectively; reference: [#]age 1; Tukey's HSD test for post hoc comparisons between age classes shown in (c)), and in (d) individuals of known age and individuals of unknown age at least in their second year of life (n = 169; reference: [#]age 1). Body condition = residuals of a linear regression of body mass (in g) against tarsus length (in mm) and time of the day. In bold are values for which $P < 0.07$ and levels of significance are specified as follows: •$P < 0.07$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; effects that are biologically important (i.e. > 1 jnd across the range of realistic values) are underlined. See Table S8 for variance of random effects.

(a)

	Purple PC1				Purple PC2				Purple L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	-0.05	0.26	69	-0.18	-0.23	0.16	14	-1.48	72.24	0.29	7	250.70
Age	4.02×10^{-3}	6.81×10^{-3}	54	0.59	-5.07×10^{-3}	3.99×10^{-3}	24	-1.27	5.07×10^{-3}	4.51×10^{-3}	61	1.12
Social status[†]	-0.10	0.40	72	-0.25	-9.17×10^{-3}	0.24	62	-0.04	-2.82×10^{-4}	0.26	69	0.00
Body condition	-0.02	0.21	73	-0.10	-0.02	0.13	68	-0.17	**0.28**	**0.14**	**68**	**2.04***
Group size	0.06	0.10	73	0.67	-0.04	0.06	71	-0.63	-0.07	0.06	69	-1.10
Territory quality	0.02	0.04	56	0.44	8.22×10^{-4}	0.02	33	0.04	2.77×10^{-3}	0.02	60	0.12
Rainfall	5.63×10^{-4}	8.18×10^{-4}	73	0.69	-4.43×10^{-4}	5.17×10^{-4}	4	-0.86	-1.57×10^{-4}	1.35×10^{-3}	4	-0.12

% purple	**0.04**	**5.19×10^{-3}**	**72**	**7.21*****	**0.02**	**3.23×10^{-3}**	**53**	**7.56*****	**8.96×10^{-3}**	**3.69×10^{-3}**	**71**	**2.43***
Marginal / conditional R^2	0.49 / 0.67				0.50 / 0.55				0.16 / 0.51			

(b)

	Purple PC1				Purple PC2				Purple L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	-0.17	0.43	60	-0.39	-0.27	0.29	20	-0.94	71.59	0.37	16	191.73
Age[#]												
Age 2	-0.07	0.48	66	-0.15	0.14	0.28	66	0.50	0.70	0.31	65	2.26
Age 3	-0.36	0.55	65	-0.65	-0.01	0.33	63	-0.04	0.93	0.36	66	2.58
Age 4	-0.43	0.65	64	-0.66	-0.37	0.40	54	-0.92	1.01	0.44	63	2.27
Age 5	-0.21	0.64	66	-0.32	-0.35	0.39	56	-0.89	0.76	0.44	65	1.76
Social status[†]	0.16	0.44	67	0.37	0.13	0.24	62	0.55	-0.14	0.26	63	-0.53
Body condition	0.22	0.23	67	0.93	-0.01	0.13	65	-0.10	**0.36**	**0.14**	**64**	**2.56****
Group size	0.06	0.11	67	0.57	-0.05	0.06	65	-0.75	-0.06	0.07	64	-0.92
Territory quality	0.04	0.04	51	0.97	0.02	0.02	42	0.98	7.74×10^{-3}	0.03	56	0.30
Rainfall	7.24×10^{-4}	9.96×10^{-4}	67	0.73	-4.18×10^{-4}	8.58×10^{-4}	4	-0.49	-1.06×10^{-5}	1.41×10^{-3}	4	-0.01
% purple	**0.04**	**5.85×10^{-3}**	**66**	**7.10*****	**0.02**	**3.51×10^{-3}**	**62**	**6.28*****	**8.12×10^{-3}**	**3.82×10^{-3}**	**65**	**2.12***
Marginal / conditional R^2	0.52 / 0.63				0.46 / 0.67				0.27 / 0.62			

(c)

	Purple PC1			Purple PC2			Purple L		
Pairwise comparisons	β	SE	z	β	SE	z	β	SE	z
2-1	-0.07	0.48	-0.15	0.14	0.28	0.50	0.70	0.31	2.26
3-1	-0.36	0.55	-0.65	-0.01	0.33	-0.04	0.93	0.36	2.58
4-1	-0.43	0.65	-0.66	-0.37	0.40	-0.92	1.01	0.44	2.27
5-1	-0.21	0.64	-0.32	-0.35	0.39	-0.89	0.76	0.44	1.75

3-2	-0.28	0.42	-0.68	-0.16	0.24	-0.64	0.23	0.27	0.86
4-2	-0.36	0.55	-0.66	-0.51	0.32	-1.60	0.30	0.35	0.87
5-2	-0.13	0.56	-0.24	-0.49	0.32	-1.55	0.06	0.35	0.17
4-3	-0.08	0.51	-0.15	-0.36	0.28	-1.26	0.07	0.31	0.24
5-3	0.15	0.54	0.28	-0.34	0.30	-1.11	-0.17	0.33	-0.51
5-4	0.23	0.60	0.38	0.02	0.34	0.06	-0.24	0.36	-0.67

(d)

	Purple PC1				Purple PC2				Purple L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	0.08	0.38	83	0.22	-0.61	0.27	59	-2.21	71.57	0.32	58	226.34
Age[#]	-0.14	0.39	145	-0.35	**0.49**	**0.25**	**159**	**1.95**•	**1.07**	**0.29**	**157**	**3.68*****
Social status[†]	-0.26	0.25	154	-1.04	0.10	0.16	158	0.64	-0.01	0.18	157	-0.05
Body condition	0.07	0.15	160	0.46	-0.13	0.09	158	-1.37	**0.23**	**0.11**	**158**	**2.06***
Group size	0.07	0.07	157	1.11	-0.07	0.04	157	-1.56	-0.05	0.05	155	-0.94
Territory quality	-2.21×10^{-5}	0.02	111	0.00	**0.03**	**0.01**	**123**	**2.07***	-0.02	0.02	132	-1.44
Rainfall	-7.69×10^{-4}	6.16×10^{-4}	8	-1.25	1.74×10^{-6}	6.47×10^{-4}	6	0.00	4.15×10^{-4}	7.44×10^{-4}	6	-0.56
% purple	**0.05**	**3.24×10^{-3}**	**108**	**15.28*****	**0.01**	**2.09×10^{-3}**	**143**	**5.58*****	**7.84×10^{-3}**	**2.49×10^{-3}**	**157**	**3.14****
Marginal / conditional R^2	0.62 / 0.70				0.26 / 0.59				0.22 / 0.46			

Table S4. Males in better body condition produce a brighter blue tail colouration than males in worse condition, but this effect is only discriminable between males at the extremes of the body condition range. Other intrinsic and environmental factors explain little of the chromatic and achromatic variations of the blue colouration. Shown are results from LMMs examining the effects of age, social status (reference: †subordinate), body condition, group size, territory quality, cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) and % purple on average PC1, PC2 and L of the blue tail colouration in (a, b) individuals of known age only, considering age (a) in months as a covariate (n = 155) and (b) in years as a cofactor (levels 1, 2, 3, 4, 5; n = 64, 27, 21, 12, 10, respectively; reference: #age 1; Tukey's HSD test for post hoc comparisons between age classes shown in (c)), and in (d) individuals of known age and individuals of unknown age at least in their second year of life (n = 235; reference: #age 1). Body condition = residuals of a linear regression of body mass (in g) against tarsus length (in mm) and time of the day. In bold are values for which $P < 0.07$ and levels of significance are specified as follows: •$P < 0.07$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; effects that are biologically important (i.e. > 1 jnd across the range of realistic values) are underlined. See Table S8 for variance of random effects.

(a)

	Blue PC1				Blue PC2				Blue L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	-0.37	0.17	9	-2.13	0.01	0.11	10	0.12	70.77	0.25	8	287.94
Age	-5.92×10^{-3}	6.05×10^{-3}	128	-0.98	**-6.07×10^{-3}**	**3.00×10^{-3}**	**130**	**-2.02***	9.62×10^{-3}	6.94×10^{-3}	133	1.39
Social status†	0.31	0.33	146	0.95	0.06	0.16	144	0.38	-0.26	0.37	144	-0.71
Body condition	**0.32**	**0.13**	**146**	**2.48***	-0.02	0.06	143	-0.39	0.20	0.15	143	1.34
Group size	-0.05	0.06	138	-0.87	0.02	0.03	139	0.62	-0.12	0.07	139	-1.82

Territory quality	0.03	0.02	113	1.09	0.01	0.01	98	1.25	**-0.06**	**0.03**	**108**	**-2.33***
Rainfall	-5.17×10^{-4}	7.78×10^{-4}	6	-0.67	8.66×10^{-4}	5.02×10^{-4}	10	1.73	-1.61×10^{-3}	1.13×10^{-3}	7	-1.42
% purple	$\mathbf{7.62\times10^{-3}}$	$\mathbf{3.48\times10^{-3}}$	**99**	**2.19***	$\mathbf{3.95\times10^{-3}}$	$\mathbf{1.78\times10^{-3}}$	**145**	**2.22***	$\mathbf{-7.69\times10^{-3}}$	$\mathbf{4.12\times10^{-3}}$	**143**	**-1.87•**
Marginal / conditional R^2	0.14 / 0.39				0.10 / 0.30				0.12 / 0.30			

(b)

	Blue PC1				Blue PC2				Blue L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	-0.32	0.23	15	-1.37	-0.01	0.13	15	-0.09	70.64	0.30	10	235.64
Age[#]												
Age 2	0.32	0.30	105	1.09	0.18	0.15	123	1.22	-0.33	0.29	122	-1.13
Age 3	0.05	0.40	93	0.13	0.23	0.20	122	1.15	0.33	0.39	123	0.85
Age 4	-0.14	0.51	91	-0.28	-0.30	0.25	117	-1.20	0.62	0.49	121	1.26
Age 5	-0.39	0.50	101	-0.78	0.01	0.24	122	0.05	0.40	0.47	120	0.85
Social status[†]	0.35	0.30	98	1.16	8.07×10^{-3}	0.15	118	0.05	-0.19	0.29	115	-0.67
Body condition	0.06	0.15	102	0.42	0.01	0.07	122	0.16	0.19	0.15	119	1.30
Group size	6.84×10^{-3}	0.06	118	0.11	-2.99×10^{-4}	0.03	119	-0.10	0.02	0.06	117	0.30
Territory quality	0.03	0.03	93	0.98	7.66×10^{-3}	0.01	67	0.58	-0.03	0.03	83	-1.13
Rainfall	-4.61×10^{-4}	8.37×10^{-4}	5	-0.55	4.15×10^{-4}	5.29×10^{-4}	9	0.79	-1.45×10^{-3}	1.27×10^{-3}	7	-1.15
% purple	5.57×10^{-3}	3.73×10^{-3}	110	1.49	2.54×10^{-3}	1.85×10^{-3}	122	1.37	-6.28×10^{-3}	3.59×10^{-3}	119	-1.75
Marginal / conditional R^2	0.09 / 0.65				0.13 / 0.42				0.11 / 0.56			

(c)

	Blue PC1			Blue PC2			Blue L		
Pairwise comparisons	β	SE	z	β	SE	z	β	SE	z
2-1	0.32	0.30	1.09	0.18	0.15	1.22	-0.33	0.29	-1.13
3-1	0.05	0.40	0.13	0.23	0.20	1.15	0.33	0.39	0.85

4-1	-0.14	0.51	-0.28	-0.30	0.25	-1.20	0.62	0.49	1.26
5-1	-0.39	0.50	-0.78	0.01	0.24	0.05	0.40	0.47	0.85
3-2	-0.27	0.32	-0.84	0.05	0.17	0.28	0.66	0.33	2.03
4-2	-0.47	0.41	-1.14	-0.48	0.21	-2.26	0.94	0.41	2.30
5-2	-0.71	0.43	-1.66	-0.17	0.21	-0.79	0.73	0.41	1.78
4-3	-0.20	0.38	-0.52	**-0.53**	**0.19**	**-2.73***	0.28	0.37	0.76
5-3	-0.44	0.42	-1.04	-0.22	0.21	-1.03	0.07	0.41	0.17
5-4	-0.24	0.45	-0.55	0.31	0.23	1.37	-0.21	0.44	-0.49

(d)

	Blue PC1				Blue PC2				Blue L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	-0.37	0.23	19	-1.64	0.04	0.12	22	0.33	70.75	0.28	13	249.52
Age[#]	0.34	0.22	219	1.58	0.15	0.12	225	1.31	-0.13	0.22	223	-0.57
Social status[†]	0.32	0.20	225	1.57	-0.15	0.11	221	-1.41	0.02	0.20	221	0.10
Body condition	0.05	0.11	219	0.48	3.68×10^{-3}	0.06	226	0.06	**0.29**	**0.11**	**223**	**2.52***
Group size	-4.63×10^{-3}	0.05	220	-0.10	-0.04	0.02	218	-1.41	0.02	0.05	219	0.39
Territory quality	0.02	0.02	164	1.16	9.68×10^{-3}	9.82×10^{-3}	142	0.99	-0.02	0.02	171	-0.82
Rainfall	-2.89×10^{-5}	7.74×10^{-4}	5	-0.04	4.36×10^{-4}	4.24×10^{-4}	7	1.03	-1.63×10^{-3}	1.11×10^{-3}	6	-1.46
% purple	1.41×10^{-3}	2.30×10^{-3}	177	0.62	$\mathbf{3.12\times10^{-3}}$	$\mathbf{1.26\times10^{-3}}$	**194**	**2.47***	-3.53×10^{-3}	2.42×10^{-3}	211	-1.46
Marginal / conditional R^2	0.08 / 0.56				0.10 / 0.43				0.09 / 0.49			

Table S5. Males in better body condition produce a brighter buff-white throat colouration than males in worse condition, but this effect is only discriminable between males at the extremes of the body condition range. Other intrinsic and environmental factors explain little of the chromatic and achromatic variations of the buff-white colouration. Shown are results from LMMs examining the effects of age, social status (reference: [†]subordinate), body condition, group size, territory quality, cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) and % purple on average PC1 and L of the buff-white throat colouration in (a, b) individuals of known age only, considering age (a) in months as a covariate (n = 129) and (b) in years as a cofactor (levels 1, 2, 3, 4, 5; n = 46, 24, 22, 12, 10, respectively; reference: [#]age 1; Tukey's HSD test for post hoc comparisons between age classes shown in (c)), and in (d) individuals of known age and individuals of unknown age at least in their second year of life (n = 182). Body condition = residuals of a linear regression of body mass (in g) against tarsus length (in mm) and time of the day. In bold are values for which $P < 0.07$ and levels of statistical significance are specified as follows: $^{\bullet}P < 0.07$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; effects that are biologically important (i.e. > 1 jnd across the range of realistic values) are underlined. See Table S8 for variance of random effects.

(a)

	Buff-white PC1				Buff-white L			
Fixed effects	β	SE	df	t	β	SE	df	t
Intercept	0.02	0.13	5	0.13	78.27	0.22	7	361.88
Age	2.16×10^{-3}	1.74×10^{-3}	99	1.24	-5.53×10^{-3}	5.47×10^{-3}	110	-1.01
Social status[†]	-0.12	0.09	115	-1.23	0.19	0.30	119	0.65
Body condition	0.05	0.04	112	1.01	-0.02	0.14	116	-0.16
Group size	-0.02	0.02	115	-1.22	0.01	0.05	116	0.21
Territory quality	4.79×10^{-3}	6.93×10^{-3}	89	0.69	-0.01	0.02	83	-0.53
Rainfall	7.68×10^{-6}	7.21×10^{-4}	5	0.01	2.23×10^{-4}	1.08×10^{-3}	5	0.21
% purple	$\mathbf{2.96\times10^{-3}}$	$\mathbf{1.11\times10^{-3}}$	**119**	**2.67****	$\mathbf{6.61\times10^{-3}}$	$\mathbf{3.36\times10^{-3}}$	**112**	**1.97$^{\bullet}$**
Marginal / conditional R^2	0.10 / 0.58				0.06 / 0.37			

(b)

	Buff-white PC1				Buff-white L			
Fixed effects	β	SE	df	t	β	SE	df	t
Intercept	-0.15	0.13	7	-1.17	78.68	0.21	10	373.48
Age[#]								
Age 2	0.17	0.08	99	2.11	-0.41	0.24	101	-1.68
Age 3	0.42	0.11	95	3.96	-0.30	0.31	97	-0.96
Age 4	0.53	0.13	85	3.96	-0.06	0.39	80	-0.15
Age 5	0.44	0.13	94	3.41	-0.25	0.38	98	-0.67
Social status[†]	**-0.20**	**0.08**	**99**	**-2.41***	-0.08	0.24	100	-0.35
Body condition	6.50×10^{-3}	0.04	98	0.16	0.09	0.12	98	0.76
Group size	-0.03	0.02	98	-1.65	0.02	0.05	98	0.41
Territory quality	-2.78×10^{-3}	6.97×10^{-3}	74	-0.40	-0.02	0.02	67	-1.15
Rainfall	1.44×10^{-4}	6.29×10^{-4}	5	0.23	-2.35×10^{-4}	8.81×10^{-4}	5	-0.27
% purple	1.15×10^{-3}	1.04×10^{-3}	100	1.10	$\mathbf{7.81\times10^{-3}}$	$\mathbf{3.07\times10^{-3}}$	**99**	**2.54***
Marginal / conditional R^2	0.23 / 0.70				0.11 / 0.24			

(c)

	Buff-white PC1			Buff-white L		
Pairwise comparisons	β	SE	z	β	SE	z
2-1	0.17	0.08	2.11	-0.41	0.24	-1.68
3-1	**0.42**	**0.11**	**3.96*****	-0.30	0.31	-0.96
4-1	**0.53**	**0.13**	**3.96*****	-0.06	0.39	-0.15
5-1	**0.44**	**0.13**	**3.41****	-0.25	0.38	-0.67
3-2	**0.25**	**0.09**	**2.77***	0.11	0.28	0.38
4-2	**0.36**	**0.12**	**3.09***	0.35	0.35	1.01
5-2	0.27	0.11	2.35	0.15	0.34	0.45
4-3	0.10	0.10	1.04	0.24	0.30	0.81
5-3	0.02	0.11	0.14	0.05	0.33	0.15
5-4	-0.09	0.12	-0.74	-0.20	0.36	-0.54

(d)

	Buff-white PC1				Buff-white L			
Fixed effects	β	SE	df	t	β	SE	df	t
Intercept	-0.08	0.10	10	-0.82	78.66	0.26	11	306.27
Age[#]	**0.24**	**0.07**	**171**	**3.38*****	-0.08	0.21	167	-0.38
Social status	-0.04	0.06	172	-0.61	-0.08	0.19	165	-0.41
Body condition	0.02	0.03	170	0.56	**0.31**	**0.11**	**168**	**2.90****
Group size	**-0.03**	**0.01**	**170**	**-2.35***	-0.04	0.04	168	-0.95
Territory quality	**0.01**	$\mathbf{5.22\times10^{-3}}$	**171**	**2.81****	4.75×10^{-3}	0.02	141	0.30
Rainfall	-1.48×10^{-6}	4.71×10^{-4}	5	0.00	1.15×10^{-4}	1.09×10^{-3}	4	0.11
% purple	$\mathbf{2.09\times10^{-3}}$	$\mathbf{8.35\times10^{-4}}$	**174**	**2.51***	3.68×10^{-3}	2.49×10^{-3}	170	1.48
Marginal / conditional R^2	0.22 / 0.49				0.06 / 0.34			

Table S6. Males in better body condition produce a brighter brown back colouration than males in worse condition, but this effect is only discriminable between males at the extremes of the body condition range. Other intrinsic and environmental factors explain little of the chromatic and achromatic variations of the brown colouration. Shown are results from LMMs examining the effects of age, social status (reference: †subordinate), body condition, group size, territory quality, cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) and % purple on average PC1 and L of the brown back colouration in (a, b) individuals of known age only, considering age (a) in months as a covariate (n = 152) and (b) in years as a cofactor (levels 1, 2, 3, 4, 5; n = 63, 27, 21, 11, 10, respectively; reference: #age 1; Tukey's HSD test for post hoc comparisons between age classes shown in (c)), and in (d) individuals of known age and individuals of unknown age at least in their second year of life (n = 233; reference: #age 1). Body condition = residuals of a linear regression of body mass (in g) against tarsus length (in mm) and time of the day. In bold are values for which $P < 0.07$ and levels of significance are specified as follows: •$P < 0.07$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; effects that are biologically important (i.e. > 1 jnd across the range of realistic values) are underlined. See Table S8 for variance of random effects.

(a)

	Brown PC1				Brown L			
Fixed effects	β	SE	df	t	β	SE	df	t
Intercept	0.17	0.33	7	0.51	70.98	0.22	7	325.90
Age	3.11×10^{-3}	5.74×10^{-3}	122	0.54	9.74×10^{-3}	5.66×10^{-3}	130	1.72
Social status†	0.06	0.31	137	0.18	-0.11	0.31	140	-0.36
Body condition	-0.14	0.12	136	-1.11	0.09	0.12	139	0.73
Group size	0.05	0.05	135	1.02	-0.08	0.05	137	-1.48
Territory quality	-0.02	0.02	96	-0.77	-4.61×10^{-3}	0.02	109	-0.21
Rainfall	1.42×10^{-3}	1.53×10^{-3}	7	0.93	-2.09×10^{-3}	1.01×10^{-3}	7	-2.07
% purple	8.30×10^{-4}	3.47×10^{-3}	142	0.24	-2.69×10^{-3}	3.40×10^{-3}	143	-0.79
Marginal / conditional R^2	0.04 / 0.41				0.09 / 0.27			

(b)

	Brown PC1	Brown L

Fixed effects	β	SE	df	t	β	SE	df	t
Intercept	-0.09	0.34	8	-0.27	71.00	0.27	9	262.56
Age[#]								
Age 2	0.29	0.26	115	1.13	-0.21	0.25	119	-0.87
Age 3	0.39	0.35	116	1.09	0.26	0.34	120	0.76
Age 4	0.32	0.44	104	0.73	0.12	0.43	115	0.29
Age 5	0.22	0.43	117	0.51	0.32	0.40	119	0.80
Social status[†]	0.13	0.26	116	0.50	-0.28	0.25	116	-1.11
Body condition	-0.12	0.13	115	-0.88	**0.29**	**0.13**	**117**	**2.28***
Group size	-0.01	0.05	115	-0.26	4.42×10^{-3}	0.05	115	0.09
Territory quality	-4.77×10^{-3}	0.02	85	-0.21	-0.02	0.02	93	-0.98
Rainfall	1.27×10^{-3}	1.50×10^{-3}	7	0.84	-2.40×10^{-3}	1.16×10^{-3}	6	-2.07
% purple	-1.06×10^{-4}	3.27×10^{-3}	117	-0.03	-1.73×10^{-3}	3.09×10^{-3}	118	-0.56
Marginal / conditional R^2	0.07 / 0.51				0.15 / 0.53			

(c)

	Brown PC1			Brown L		
Pairwise comparisons	β	SE	z	β	SE	z
2-1	0.29	0.26	1.13	-0.21	0.25	-0.87
3-1	0.39	0.35	1.09	0.26	0.34	0.76
4-1	0.32	0.44	0.73	0.12	0.43	0.29
5-1	0.22	0.43	0.51	0.32	0.40	0.80
3-2	0.09	0.30	0.30	0.47	0.28	1.67
4-2	0.03	0.38	0.07	0.34	0.36	0.93
5-2	-0.08	0.38	-0.20	0.54	0.35	1.52
4-3	-0.06	0.35	-0.18	-0.13	0.33	-0.40
5-3	-0.17	0.37	-0.45	0.07	0.35	0.19
5-4	-0.10	0.41	-0.25	0.20	0.39	0.52

(d)

	Brown PC1				Brown L			
Fixed effects	β	SE	df	t	β	SE	df	t
Intercept	0.16	0.28	9	0.58	71.05	0.25	11	282.23
Age[#]	0.03	0.17	218	0.17	-0.06	0.19	220	-0.34
Social status	4.60×10^{-3}	0.15	210	0.03	0.15	0.17	217	0.87
Body condition	-0.03	0.09	211	-0.30	**0.30**	**0.10**	**219**	**2.99****
Group size	-0.02	0.04	213	-0.55	0.02	0.04	216	0.52
Territory quality	0.01	0.01	165	1.12	-7.51×10^{-3}	0.02	179	-0.49
Rainfall	5.66×10^{-4}	1.20×10^{-3}	6	0.47	-1.68×10^{-3}	1.00×10^{-3}	6	-1.67
% purple	$\mathbf{5.89\times10^{-3}}$	$\mathbf{1.89\times10^{-3}}$	**219**	**3.12****	-6.73×10^{-4}	2.11×10^{-3}	218	-0.32
Marginal / conditional R^2	0.08 / 0.51				0.08 / 0.43			

Table S7. Older males tend to produce a darker black cheek colouration compared to first-year males. Additionally, males in better body condition tend to produce a brighter black cheek colouration than males in worse condition. Shown are results from LMMs examining the effects of age, social status (reference: †subordinate), body condition, group size, territory quality, cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) and % purple on average L of the black cheek colouration in (a, b) individuals of known age only, considering age (a) in months as a covariate (n = 82) and (b) in years as a cofactor (levels 1, 2, 3, 4, 5; n = 13, 23, 21, 12, 10, respectively; reference: #age 1; Tukey's HSD test for post hoc comparisons between age classes shown in (c)), and in (d) individuals of known age and individuals of unknown age at least in their second year of life (n = 172; reference: #age 1). Body condition = residuals of a linear regression of body mass (in g) against tarsus length (in mm) and time of the day. In bold are values for which $P < 0.07$ and levels of significance are specified as follows: •$P < 0.07$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; effects that are biologically important (i.e. > 1 jnd across the range of realistic values) are underlined. See Table S8 for variance of random effects.

(a)

	Black L			
Fixed effects	β	SE	df	t
Intercept	58.00	1.62	14	35.77
Age	3.45×10^{-3}	0.03	60	0.11
Social status†	0.29	1.86	72	0.16
Body condition	1.40	0.99	71	1.42
Group size	-0.87	0.50	70	-1.74
Territory quality	-0.23	0.19	66	-1.20
Rainfall	4.23×10^{-4}	6.40×10^{-3}	8	0.07
% purple	**-0.10**	**0.02**	**72**	**-4.22*****
Marginal / conditional R^2	0.23 / 0.55			

(b)

	Black L			
Fixed effects	β	SE	df	t
Intercept	60.80	2.10	40	29.01

Age[#]				
Age 2	-3.61	2.19	67	-1.65
Age 3	-3.19	2.51	68	-1.27
Age 4	-3.40	3.01	63	-1.13
Age 5	1.14	2.90	62	0.39
Social status[†]	-0.76	1.92	65	-0.39
Body condition	1.78	1.01	64	1.78
Group size	-0.84	0.54	64	-1.55
Territory quality	**-0.49**	**0.21**	**62**	**-2.30***
Rainfall	-2.75×10^{-3}	5.44×10^{-3}	8	-0.51
% purple	**-0.09**	**0.03**	**65**	**-3.47*****
Marginal / conditional R^2	0.31 / 0.61			

(c)

	Black L		
Pairwise comparisons	β	SE	z
2-1	-3.61	2.19	-1.65
3-1	-3.19	2.51	-1.27
4-1	-3.40	3.01	-1.13
5-1	1.14	2.90	0.39
3-2	0.42	1.87	0.22
4-2	0.21	2.46	0.09
5-2	4.75	2.48	1.92
4-3	-0.21	2.15	-0.10
5-3	4.33	2.34	1.85
5-4	4.54	2.50	1.82

(d)

	Black L			
Fixed effects	β	SE	df	t
Intercept	60.82	1.74	66	34.89
Age[#]	-2.77	1.57	161	-1.76
Social status[†]	-0.32	1.05	152	-0.30
Body condition	1.07	0.59	161	1.80
Group size	-0.49	0.29	158	-1.69
Territory quality	-0.17	0.10	125	-1.72
Rainfall	-6.85×10^{-4}	4.09×10^{-3}	6	-0.17
% purple	**-0.07**	**0.01**	**161**	**-4.77*****
Marginal / conditional R^2	0.20 / 0.45			

Table S8. Variance of random effects fitted in the LMMs presented in Tables S3-S7, including age either in months as a covariate, or in years as a cofactor (levels 1-5), or age class (levels 1, 2+). Sample sizes for these LMMs are respectively *n* = 81, 78 and 69 for the purple crown, *n* = 155, 134 and 235 for the blue tail, *n* = 129, 114 and 182 for the buff-white throat, *n* = 152, 132 and 233 for the brown back, and *n* = 82, 79 and 172 for the black cheek.

LMMs fitting	Random effects	Purple crown			Blue tail			Buff-white throat		Brown back		Black cheek
		PC1	PC2	L	PC1	PC2	L	PC1	L	PC1	L	L
Age in months	Bird ID	0.48	0.04	0.04	0.33	0.03	0.13	0.00	0.21	0.00	0.03	8.34
	Year	0.00	0.00	0.34	0.07	0.05	0.26	0.10	0.17	0.69	0.23	6.98
	Residual	0.93	0.47	0.53	0.97	0.28	1.51	0.09	0.76	1.10	1.06	21.89
Age in years	Bird ID	0.38	0.14	0.09	0.77	0.05	0.28	0.01	0.01	0.00	0.10	12.20
	Year	0.00	0.09	0.36	0.10	0.06	0.43	0.08	0.10	0.69	0.37	2.81
	Residual	1.29	0.37	0.48	0.55	0.21	0.71	0.06	0.61	0.76	0.60	19.77
Age classes	Bird ID	0.26	0.19	0.13	0.53	0.09	0.28	0.00	0.07	0.03	0.11	4.54
	Year	0.03	0.11	0.14	0.15	0.05	0.38	0.04	0.22	0.48	0.31	4.23
	Residual	1.15	0.37	0.62	0.64	0.22	0.81	0.08	0.69	0.58	0.68	18.48

Table S9. There is no discriminable variation in the purple-and-black breeding colouration within years. Shown are results from the same models presented in (a) Table S3 (purple crown reflectance) and b) Table S7 (black cheek reflectance) but including the time interval (in days) since pre-breeding moult completion (i.e. time interval between date of moult completion and date of reflectance measurement) for individuals of known age and individuals of unknown age at least in their second year of life (i.e. considering age as a cofactor with levels 1 and 2+; n = 169 and 172 respectively; reference: #age 1). In bold are values for which $P < 0.07$ and levels of significance are specified as follows: •$P < 0.07$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; effects that are biologically important (i.e. > 1 jnd across the range of realistic values) are underlined. Variance of random effects (σ^2) for (a) PC1 / PC2 / PC3 / L: individual = 0.26 / 0.19 / 0.06 / 0.19, year = 0.01 / 0.11 / 0.00 / 0.14, residual = 1.17 / 0.38 / 0.09 / 0.51; and (b): individual = 4.18, year = 4.41, residual = 18.81.

(a)

	Purple PC1				Purple PC2				Purple L			
Fixed effects	β	SE	df	t	β	SE	df	t	β	SE	df	t
Intercept	0.12	0.38	61	0.31	-0.60	0.27	58	-2.20	71.52	0.31	58	231.72
Age[#]	-0.25	0.39	107	-0.63	**0.51**	**0.26**	**158**	**1.98***	**0.87**	**0.29**	**158**	**3.04*****
Social status[†]	-0.16	0.27	157	-0.61	0.74	0.17	159	0.43	0.27	0.19	159	1.39
Body condition	0.09	0.15	158	0.57	-0.13	0.09	156	-1.41	**0.28**	**0.11**	**156**	**2.62****
Group size	0.07	0.07	156	1.06	-0.07	0.04	156	-1.54	-0.05	0.05	156	-1.14
Territory quality	2.77×10^{-3}	0.02	108	0.12	**0.03**	**0.01**	**124**	**2.03***	-0.02	0.02	132	-0.96
Rainfall	-6.79×10^{-4}	5.64×10^{-4}	62	-1.21	-2.12×10^{-5}	6.48×10^{-4}	6	-0.03	-1.55×10^{-4}	7.43×10^{-4}	6	-0.21
% purple	**0.05**	**3.31×10^{-3}**	**111**	**14.91*****	**0.01**	**2.14×10^{-3}**	**144**	**5.54*****	**6.07×10^{-3}**	**2.42×10^{-3}**	**152**	**2.51***
Time since moult	-1.51	1.35×10^{-3}	36	-1.12	3.95×10^{-4}	9.28×10^{-4}	116	0.43	**-4.16×10^{-3}**	**1.06×10^{-3}**	**133**	**-3.94*****
Marginal / conditional R^2	0.63 / 0.70				0.26 / 0.59				0.29 / 0.57			

(b)

	Black L			
Fixed effects	β	SE	df	t
Intercept	60.70	1.75	65	34.62
Age[#]	**-2.90**	**1.59**	**160**	**-1.83**•
Social status[†]	-6.55×10^{-3}	1.12	154	-0.01
Body condition	**1.14**	**0.60**	**158**	**1.89**•
Group size	-0.49	0.29	157	-1.70
Territory quality	-0.16	0.10	125	-1.64
Rainfall	-3.60×10^{-4}	4.17×10^{-3}	7	-0.09
% purple	**-0.07**	**0.01**	**160**	**-4.82*****
Time since moult	-4.78×10^{-3}	5.74×10^{-3}	150	-0.83
Marginal / conditional R^2	0.20 / 0.45			

Table S10. Specifications for animal models used to estimate the heritability of colorimetric variables for the purple crown, blue tail, buff-white throat, brown back and black cheek. The table indicates the sample size, fitted fixed effects, number of iterations and prior settings used. V_{PC1}, V_{PC2}, V_{PC3} and V_L are the variances of PC1, PC2, PC3 and L respectively for the patch of interest.

Plumage patch	Response variable	Sample size	Fixed effects	# iterations	Residual and random effect priors
Purple crown	PC1	233	% purple	27 505 000	list(G=list(G1=list(V=1, nu=0.02), G2=list(V=1, nu=0.02), G3=list(V=1, nu=0.02)), R=list(V=1, nu=0.02))
	PC2	233	% purple	27 505 000	list(G=list(G1=list(V=matrix(V_{PC2}*0.5), nu=0.5), G2=list(V=matrix(V_{PC2}*0.5), nu=0.5), G3=list(V=matrix(V_{PC2}*0.5), nu=0.5)), R=list(V=matrix(V_{PC2}*0.5), nu=0.5))
	L	233	None (intercept)	27 505 000	list(G=list(G1=list(V=1, nu=0.2), G2=list(V=1, nu=0.2), G3=list(V=1, nu=0.2)), R=list(V=1, nu=0.2))
Blue tail	PC1	350	None (intercept)	20 005 000	list(G=list(G1=list(V=matrix(V_{PC1}*0.5), nu=0.5), G2=list(V=matrix(V_{PC1}*0.5), nu=0.5), G3=list(V=matrix(V_{PC1}*0.5), nu=0.5)), R=list(V=matrix(V_{PC1}*0.5), nu=0.5))
	PC2	350	None (intercept)	27 505 000	list(G=list(G1=list(V=matrix(V_{PC2}*0.5), nu=0.2), G2=list(V=matrix(V_{PC2}*0.5), nu=0.2), G3=list(V=matrix(V_{PC2}*0.5), nu=0.2)), R=list(V=matrix(V_{PC2}*0.5), nu=0.2))
	L	324	Body condition	2 005 000	list(G=list(G1=list(V=matrix(V_L*0.95), nu=0.2), G2=list(V=matrix(V_L*0.95), nu=0.2), G3=list(V=matrix(V_L*0.95), nu=0.2)), R=list(V=matrix(V_L*0.05), nu=0.2))
Buff-white throat	PC1	208	None (intercept)	2 005 000	list(G=list(G1=list(V=1, nu=0.2), G2=list(V=1, nu=0.2), G3=list(V=1, nu=0.2)), R=list(V=1, nu=0.2))
	L	208	Body condition	2 005 000	list(G=list(G1=list(V=matrix(V_L*0.95), nu=0.2), G2=list(V=matrix(V_L*0.95), nu=0.2), G3=list(V=matrix(V_L*0.95), nu=0.2)), R=list(V=matrix(V_L*0.05), nu=0.2))
Brown back	PC1	349	None (intercept)	2 005 000	list(G=list(G1=list(V=1, nu=0.2), G2=list(V=1, nu=0.2), G3=list(V=1, nu=0.2)), R=list(V=1, nu=0.2))
	L	322	Body condition	2 005 000	list(G=list(G1=list(V=matrix(V_L*0.95), nu=0.2), G2=list(V=matrix(V_L*0.95), nu=0.2), G3=list(V=matrix(V_L*0.95), nu=0.2)), R=list(V=matrix(V_L*0.05), nu=0.2))
Black cheek	L	248	% purple	10 005 000	list(G=list(G1=list(V=matrix(V_L*0.95), nu=0.2), G2=list(V=matrix(V_L*0.95), nu=0.2), G3=list(V=matrix(V_L*0.95), nu=0.2)), R=list(V=matrix(V_L*0.05), nu=0.2))

Table S11. Heritability estimates of the colorimetric variables for the purple crown, blue tail, buff-white throat, brown back and black cheek. Shown are results from animal models, including the posterior mean values of the proportion of phenotypic variance in colour components explained by additive genetic variance (i.e. heritability) and their 95% credible intervals (CI), obtained using the package "MCMCglmm", as well as the additive genetic variance, standard error (SE), *P*-values obtained from likelihood ratio tests and heritabilities obtained using ASReml-R.

		MCMCglmm		ASReml			
Plumage patch	Response variable	Heritability	95% CI	V_A	SE (V_A)	*P*	Heritability
Purple crown	PC1	0.06	(0.00, 0.16)	0.06	0.16	0.71	0.04
	PC2	0.22	(0.01, 0.42)	0.18	0.13	0.14	0.27
	L	0.11	(0.02, 0.23)	0.26	0.11	0.14	0.19
Blue tail	PC1	0.30	(0.13, 0.47)	0.55	0.22	<0.001	0.37
	PC2	0.19	(0.06, 0.31)	0.10	0.03	<0.001	0.27
	L	0.09	(0.02, 0.17)	0.22	0.10	0.02	0.11
Buff-white throat	PC1	0.12	(0.03, 0.21)	0.03	0.05	0.04	0.12
	L	0.11	(0.02, 0.22)	0.15	0.17	0.35	0.11
Brown back	PC1	0.06	(0.01, 0.11)	0.07	0.05	<0.001	0.06
	L	0.06	(0.02, 0.16)	0.16	0.11	0.07	0.11
Black cheek	L	0.07	(0.00, 0.23)	2.62	3.01	0.39	0.11

Table S12. At the population level, the quality of the purple crown, black cheek and blue tail colouration of subordinate males does not predict their chances to gain a breeder position by dispersal within that year; neither does the quality of the buff-white throat and brown back colouration. Shown are results from GLMMTMBs examining the effects of PC1, PC2 and L of each colour (PC1 only for buff-white and brown; L only for black), age (levels 1, 2+; reference: #age 1) and tarsus length on the likelihood of acquiring a breeder position by dispersal among subordinates. PC1 and PC2 values were adjusted for the purple crown as described in Appendix S2c. Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2) for purple crown / blue tail / buff-white throat / brown back / black cheek: individual = 7.88×10^{-10} / 1.11×10^{4} / 0.37 / 2.57×10^{3} / 6.03×10^{3}, territory = 1.74 / 8.09×10^{-5} / 1.14×10^{-10} / 0.00 / 1.78×10^{-16}, year = 0.00 / 3.11×10^{-8} / 2.43×10^{-10} / 0.00 / 0.00. For the brown back, territory and year random effects were fitted in separate models which provided almost identical estimates for the fixed effects and individual variance.

	Purple crown (n = 36)			Blue tail (n = 82)			Buff-white throat (n = 66)			Brown back (n = 81)			Black cheek (n = 40)		
Fixed effects	β	SE	z	β	SE	z	β	SE	z	β	SE	z	β	SE	z
Intercept	-1.76	1.81	-0.97	-44.95	15.70	-2.86	-1.67	0.81	-2.06	-23.46	10.79	-2.17	-32.38	22.85	-1.42
PC1	-0.38	0.68	-0.55	-3.84	3.10	-1.24	0.34	1.06	0.32	0.62	2.89	0.22	-	-	-
PC2	1.62	2.04	0.79	6.59	4.04	1.63	-	-	-	-	-	-	-	-	-
L[a]	0.71	0.86	0.82	**12.02**	**4.47**	**2.69****	1.29	0.74	1.75	-1.26	3.21	-0.39	-0.10	0.50	-0.20
Age[#]	-0.45	1.77	-0.26	20.76	11.11	1.87	-0.25	0.81	-0.31	5.96	4.86	1.23	3.97	12.82	0.31
Tarsus length	0.39	1.12	0.35	-1.71	5.10	-0.34	0.25	0.47	0.53	-5.76	5.73	-1.00	-18.10	15.36	-1.18

[a] Although statistically significant, the effect of blue L on the probability of acquiring a breeder position is extremely small (when considering the range of realistic values of blue L).

Table S13. Dominant breeder males that display more purple crowns and brighter black cheeks achieve greater seasonal reproductive success. Shown are results from zero-inflated poisson models examining the effects of PC1, PC2 and L of each colour (PC1 only for buff-white and brown; L only for black), tarsus length (in mm), territory quality, group size and cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) on annual number of six-month-old fledglings produced. PC1 and PC2 values were adjusted for the purple crown as described in Appendix S2c. Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2) for purple crown / blue tail / buff-white throat / brown back / black cheek: individual = 1.16×10^{-7} / 2.54×10^{-5} / 9.81×10^{-6} / 1.35×10^{-6} / 1.28×10^{-7}, year = 3.93×10^{-7} / 0.04 / 0.05 / 1.13×10^{-7} / 4.20×10^{-6}.

	Purple crown (n = 119)			Blue tail (n = 139)			Buff-white throat (n = 90)			Brown back (n = 140)			Black cheek (n = 40)		
Fixed effects	β	SE	z	β	SE	z	β	SE	z	β	SE	z	β	SE	z
Intercept	0.33	0.13	2.54	0.30	0.21	1.41	0.14	0.23	0.62	0.33	0.13	2.57	0.40	0.12	3.27
PC1	-0.05	0.10	-0.54	0.08	0.11	0.76	0.22	0.45	0.48	-0.13	0.10	-1.33	-	-	-
PC2	**0.32**	**0.13**	**2.44***	-0.18	0.20	-0.90	-	-	-	-	-	-	-	-	-
L	0.03	0.11	0.24	-0.03	0.10	-0.29	-0.02	0.15	-0.12	0.12	0.10	1.13	**0.05**	**0.02**	**2.31***
Tarsus length	0.19	0.18	1.01	-8.54×10^{-3}	0.17	-0.05	0.12	0.20	0.59	-0.05	0.15	-0.35	-0.07	0.16	-0.44
Group size	**0.18**	**0.07**	**2.57***	0.14	0.07	1.92	0.14	0.10	1.39	0.11	0.07	1.65	**0.18**	**0.07**	**2.60****
Territory quality	-0.01	0.02	-0.45	-2.95×10^{-3}	0.02	-0.12	0.02	0.04	0.67	-3.14×10^{-3}	0.02	-0.14	-3.96×10^{-3}	0.03	-0.15
Rainfall	-2.62×10^{-4}	6.00×10^{-4}	-0.44	-2.79×10^{-4}	9.00×10^{-4}	0.76	-8.13×10^{-6}	2.04×10^{-3}	0.00	4.47×10^{-4}	6.07×10^{-4}	0.74	3.56×10^{-4}	5.57×10^{-4}	0.64

Table S14. There is no relationship between colour quality of any plumage patch and survival or life expectancy. Shown are results from statistical models examining the effects of PC1, PC2 and L of each colour (PC1 only for buff-white and brown; L only for black), age (levels 1, 2+; reference: [#]age 1), social status (reference: [†]subordinate), tarsus length (in mm), group size, territory quality and cumulative rainfall over the past 12 months (from November of the preceding calendar year to October of the current calendar year) on (a) annual survival (GLMMTMBs) and (b) remaining lifespan (zero-inflated poisson models). PC1 and PC2 values were adjusted for the purple crown as described in Appendix S2c. Explanatory variables were scaled for the purple crown and black cheek to achieve model convergence. Significant values ($P < 0.05$) are in bold. Variance of random effects (σ^2) for purple crown / blue tail / buff-white throat / brown back / black cheek: annual survival: individual = 2556.0 / 4811 / 0.46 / 2186.28 / 1041, year = 124.3 / 178 / 0.78 / 57.21 / not fitted due to model convergence issue; remaining lifespan: individual = 0.48 / 0.75 / 0.44 / 0.81 / 0.63, year = 0.70 / 0.60 / 0.57 / 0.25 / 0.40.

(a)

	Purple crown (n = 167)			Blue tail (n = 223)			Buff-white throat (n = 170)			Brown back (n = 222)			Black cheek (n = 167)		
Fixed effects	β	SE	z	β	SE	z	β	SE	z	β	SE	z	β	SE	z
Intercept	14.21	6.41	2.22	19.46	5.56	3.50	1.21	0.68	1.77	13.84	4.36	3.18	10.08	4.59	2.20
PC1	-0.81	1.61	-0.50	-2.17	1.35	-1.60	-0.95	0.83	-1.14	-0.55	1.43	-0.38	-	-	-
PC2[a]	0.74	1.38	0.54	**-5.85**	**2.24**	**-2.61****	-	-	-	-	-	-	-	-	-
L	0.20	1.95	0.10	-0.38	1.12	-0.34	-0.33	0.29	-1.13	-0.93	1.44	-0.64	0.98	0.93	1.06
Age[#]	7.55	5.32	1.42	6.73	5.00	1.35	0.98	0.77	1.27	5.08	4.04	1.26	4.64	5.27	0.88
Social status[†]	-2.71	3.59	-0.76	-1.74	3.37	-0.52	0.12	0.64	0.19	-1.43	3.87	-0.37	-2.25	3.87	-0.58
Tarsus length	-0.73	1.94	-0.38	-0.45	3.06	-0.15	-0.06	0.36	-0.15	-1.38	2.78	-0.50	-0.78	1.49	-0.53
Group size	0.17	1.60	0.11	-0.48	0.99	-0.49	-0.17	0.16	-1.05	-0.27	1.07	-0.25	0.46	1.75	0.26

Territory quality	-0.96	2.01	-0.48	0.10	0.36	0.27	-0.04	0.06	-0.62	-0.15	0.47	-0.32	-1.44	1.63	-0.88
Rainfall	2.71	4.26	0.64	-1.46	2.60	-0.56	0.44	0.52	0.86	-1.01	2.01	-0.51	-1.03	1.32	-0.78

[a] Although statistically significant, the effects of blue PC2 and PC3 on survival are extremely small (when considering the range of realistic values of blue PC2 and PC3).

(b)

	Purple crown (n = 105)			Blue tail (n = 163)			Buff-white throat (n = 112)			Brown back (n = 162)			Black cheek (n = 115)		
Fixed effects	β	SE	z	β	SE	z	β	SE	z	β	SE	z	β	SE	z
Intercept	-0.87	0.74	-1.18	-0.08	0.40	-0.19	-0.04	0.43	-0.09	0.12	0.34	0.36	-0.06	0.51	-0.11
PC1	0.05	0.09	0.58	9.63×10^{-3}	0.08	0.12	-0.23	0.31	-0.73	-0.14	0.10	-1.43	-	-	-
PC2	0.04	0.15	0.26	-0.20	0.15	-1.32	-	-	-	-	-	-	-	-	-
L	-0.12	0.10	-1.24	0.07	0.07	0.89	-0.07	0.09	-0.78	-0.03	0.08	-0.35	-1.74×10^{-3}	0.02	-0.09
Age[#]	0.66	0.60	1.09	-0.14	0.23	-0.60	-0.26	0.31	-0.83	-0.19	0.23	-0.81	-0.07	0.42	-0.17
Social status[†]	0.04	0.29	0.13	0.05	0.23	0.24	0.35	0.29	1.20	0.04	0.23	0.16	5.35×10^{-3}	0.29	0.02
Tarsus length	-0.05	0.20	-0.24	-0.03	0.16	-0.19	0.01	0.16	0.09	-2.51×10^{-3}	0.16	-0.02	-0.05	0.20	-0.25
Group size	-0.08	0.08	-1.07	-0.10	0.06	-1.70	-0.10	0.07	-1.59	**-0.12**	**0.06**	**-1.98***	-0.07	0.08	-0.93
Territory quality	-0.03	0.03	-0.94	-0.02	0.03	-0.74	-5.22×10^{-3}	0.03	-0.20	-0.02	0.03	-0.71	-0.04	0.03	-1.26
Rainfall	-1.44×10^{-3}	2.09×10^{-3}	-0.69	-1.26×10^{-3}	2.05×10^{-3}	-0.62	1.78×10^{-3}	3.48×10^{-3}	0.51	-1.30×10^{-3}	1.42×10^{-3}	-0.92	-1.66×10^{-3}	1.72×10^{-3}	-0.97

Table S15. Benjamini-Hochberg procedure for correlates between colorimetric variables and intrinsic, environmental and fitness variables. Shown are the results of the Benjamini-Hochberg procedure used to decrease the false discovery rate for correlates between ornamental colours (i.e. purple PC1, PC2 and L, blue PC1, PC2 and L and black L) and (a) intrinsic and environmental variables and (b) fitness variables (ARS = annual reproductive success), and correlates between non-ornamental colours (i.e. buff-white PC1 and L, and brown PC1 and L) and (c) intrinsic and environmental variables and (d) fitness variables. *i* depicts the rank of the *P*-value (from the smallest to the largest), *m* the total number of tests (7 for ornamental colours, 4 for non-ornamental colours) and *Q* the false discovery rate (0.05). The largest *P*-value that verifies $P < (i/m)Q$ is significant and all the *P*-values smaller than it are also significant, even the ones that do not verify $P < (i/m)Q$ (in bold).

		Age		Social status		Body condition		Group size		Territory quality		Rainfall		% purple	
i	$(i/m)Q$	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*
1	0.007	**Purple L**	**<0.001**	Blue PC1	0.118	Blue L	0.013	Black L	0.092	Purple PC2	0.040	Blue L	0.192	**Purple PC1**	**<0.001**
2	0.014	Purple PC2	0.053	Blue PC2	0.160	Purple L	0.041	Purple PC2	0.120	Black L	0.087	Purple PC1	0.248	**Purple PC2**	**<0.001**
3	0.021	Black L	0.080	Purple PC1	0.302	Black L	0.073	Blue PC2	0.159	Purple L	0.152	Blue PC2	0.338	**Black L**	**<0.001**
4	0.029	Blue PC1	0.116	Purple PC2	0.524	Purple PC2	0.173	Purple PC1	0.270	Blue PC1	0.247	Purple L	0.597	**Purple L**	**0.002**
5	0.036	Blue PC2	0.193	Black L	0.765	Blue PC1	0.633	Purple L	0.347	Blue PC2	0.326	Black L	0.872	**Blue PC2**	**0.015**
6	0.043	Blue L	0.566	Blue L	0.924	Purple PC1	0.650	Blue L	0.697	Blue L	0.414	Blue PC1	0.972	Blue L	0.145
7	0.050	Purple PC1	0.725	Purple L	0.957	Blue PC2	0.952	Blue PC1	0.922	Purple PC1	0.999	Purple PC2	0.998	Blue PC1	0.539

(b)

		Becoming a breeder		ARS		Annual survival		Remaining lifespan	
i	$(i/m)Q$	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*
1	0.007	**Blue L**	**0.004**	**Purple PC2**	**0.004**	Blue PC1	0.282	Blue PC2	0.200

2	0.014	Blue PC2	0.034	Black L	0.021	Black L	0.289	Purple L	0.220
3	0.021	Blue PC1	0.117	Blue PC2	0.328	Blue L	0.554	Blue L	0.368
4	0.029	Purple PC2	0.200	Blue PC1	0.391	Purple PC1	0.583	Purple PC1	0.590
5	0.036	Purple L	0.369	Purple L	0.748	Purple PC2	0.598	Purple PC2	0.770
6	0.043	Purple PC1	0.775	Blue L	0.753	Blue PC2	0.621	Blue PC1	0.919
7	0.050	Black L	0.845	Purple PC1	0.856	Purple L	0.917	Black L	0.930

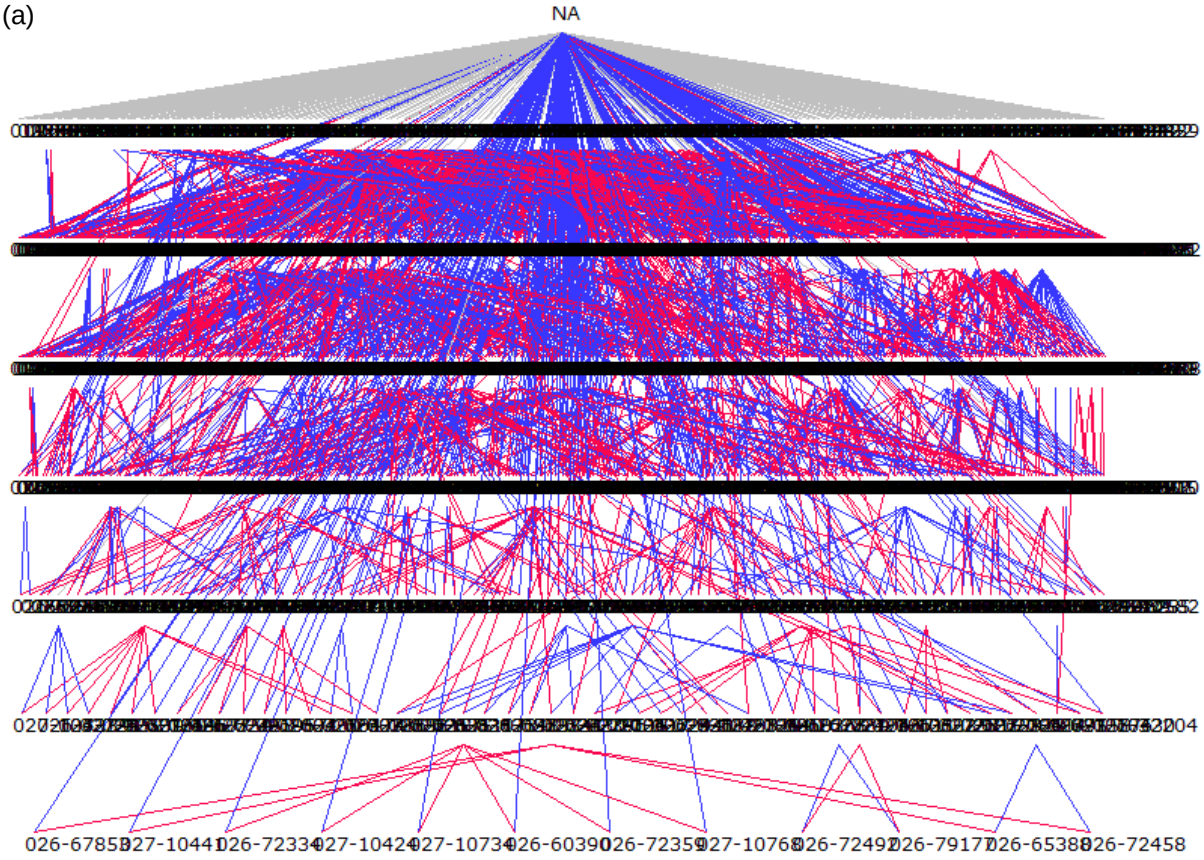
(c)

		Age		Social status		Body condition		Group size		Territory quality		Rainfall		% purple	
i	(*i*/*m*)*Q*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*
1	0.013	**Buff-white PC1**	**<0.001**	Brown L	0.387	**Brown L**	**0.003**	Buff-white PC1	0.020	**Buff-white PC1**	**0.005**	Brown L	0.149	**Brown PC1**	**0.002**
2	0.025	Buff-white L	0.707	Buff-white PC1	0.541	**Buff-white L**	**0.004**	Buff-white L	0.342	Brown PC1	0.265	Brown PC1	0.654	**Buff-white PC1**	**0.013**
3	0.038	Brown L	0.734	Buff-white L	0.684	Buff-white PC1	0.575	Brown PC1	0.583	Brown L	0.628	Buff-white L	0.921	Buff-white L	0.141
4	0.050	Brown PC1	0.866	Brown PC1	0.976	Brown PC1	0.766	Brown L	0.605	Buff-white L	0.766	Buff-white PC1	0.998	Brown L	0.750

(d)

		Becoming a breeder		ARS		Annual survival		Remaining lifespan	
i	(*i*/*m*)*Q*	Ranked	*P*	Ranked	*P*	Ranked	*P*	Ranked	*P*
1	0.013	Buff-white L	0.045	Brown L	0.316	Buff-white PC1	0.247	Brown PC1	0.183
2	0.025	Brown L	0.125	Brown PC1	0.379	Buff-white L	0.369	Buff-white PC1	0.470
3	0.038	Buff-white PC1	0.427	Buff-white PC1	0.500	Brown L	0.559	Buff-white L	0.530
4	0.050	Brown PC1	0.713	Buff-white L	0.910	Brown PC1	0.561	Brown L	0.720

(a)


NA
026-67853
027-10441
026-72334
027-10424
027-10734
026-60390
026-72359
027-10768
026-72492
026-79177
026-65388
026-72458

(b)

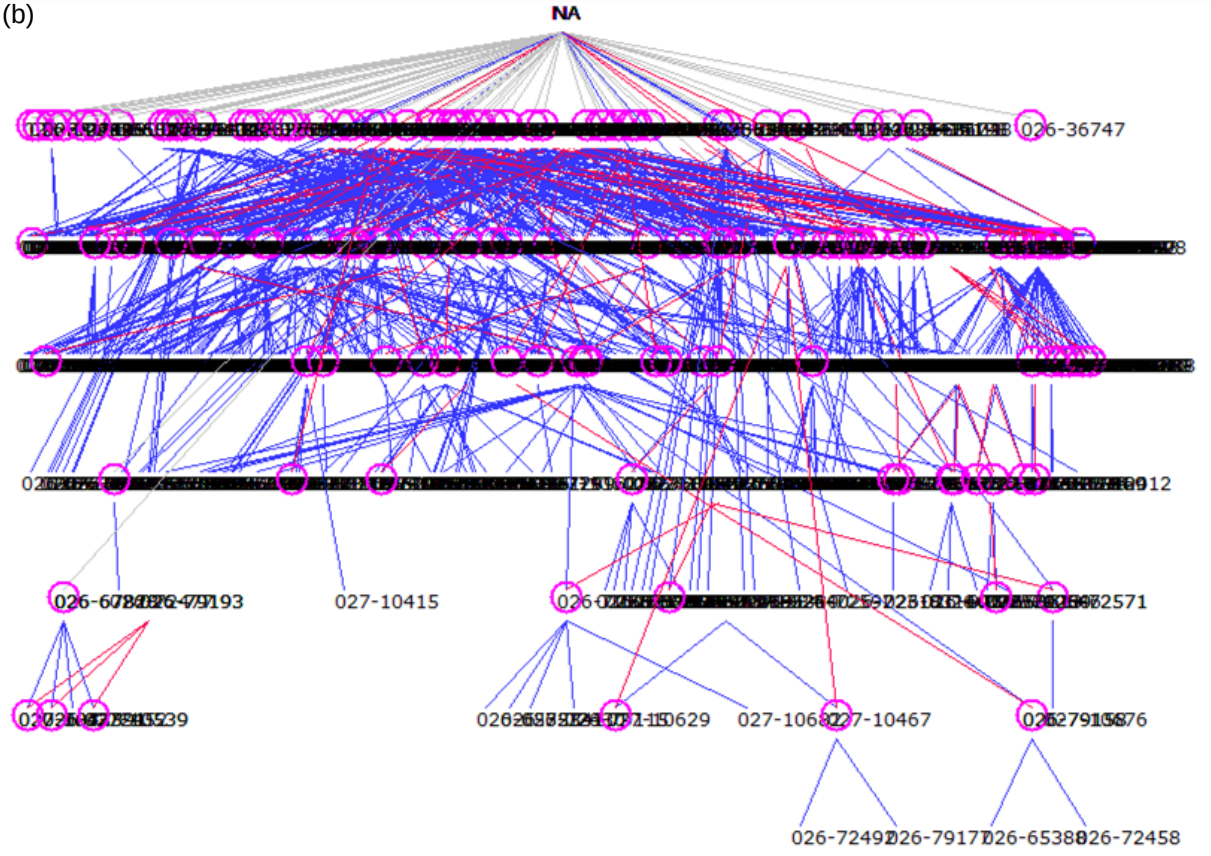

NA
026-36747
027-10415
027-10629
027-10467
026-72492
026-79177
026-65388
026-72458

Figure S1. Pedigree structure. Shown are (a) the full pedigree of our population and (b) the pedigree with the individuals informative for the animal model analyses of purple crown colouration (circled), connected to their parents and offspring. Pink lines depict maternal links and blue lines paternal links.

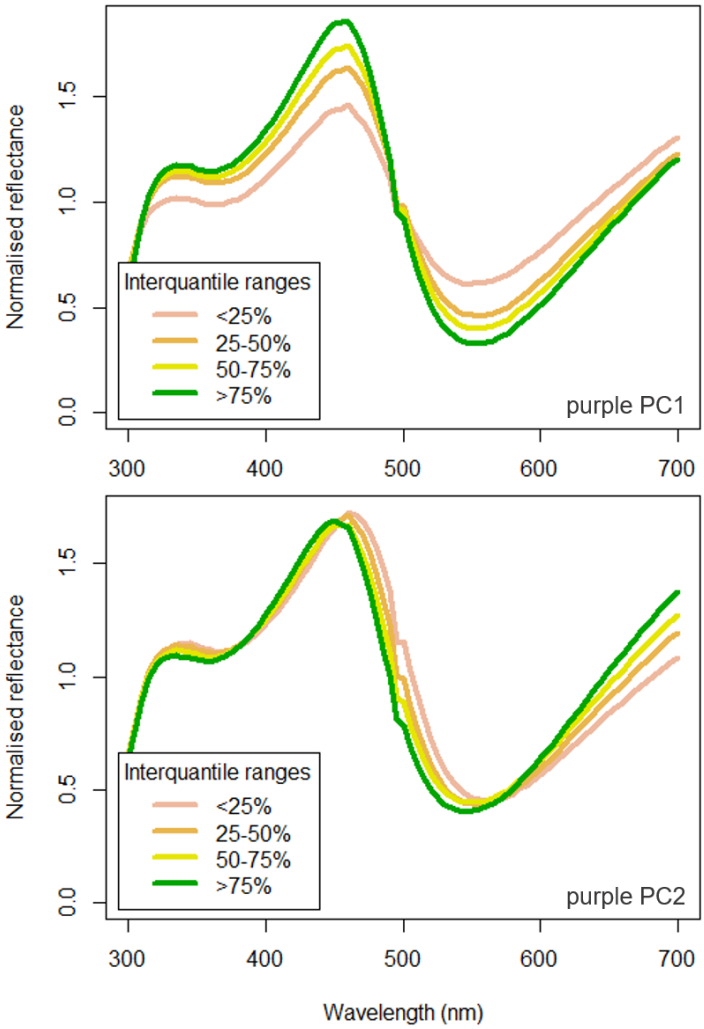


Figure S2. Purple PC1 and PC2 in relation to spectral shape. Shown are the average spectra of the four quantiles belonging to purple PC1 (top) and purple PC2 (bottom).

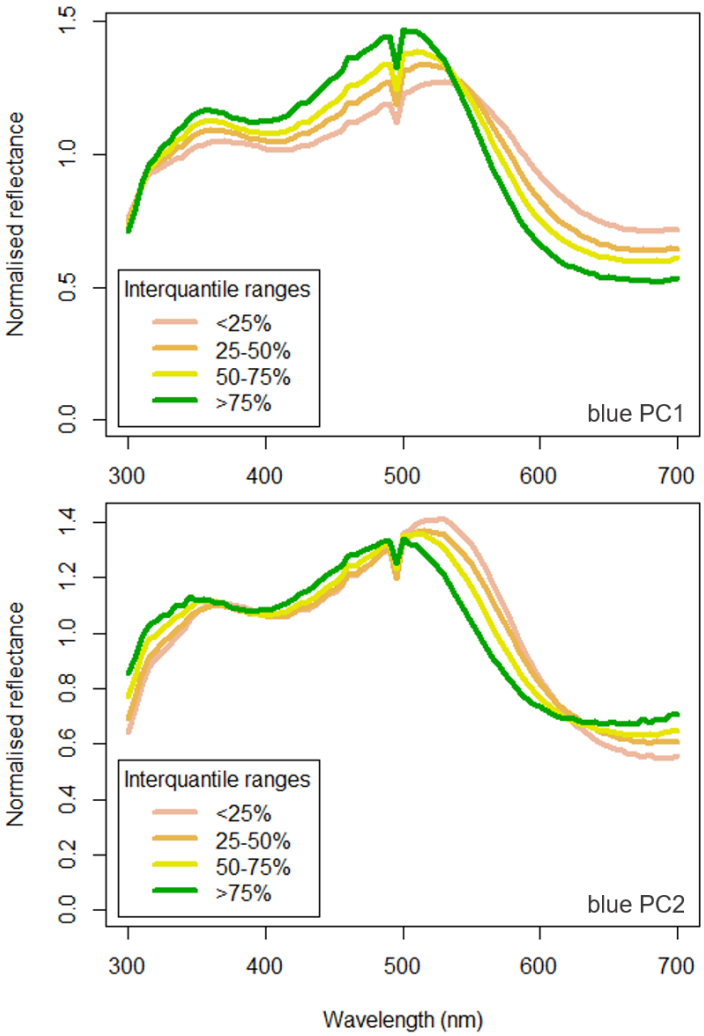


Figure S3. Blue PC1 and PC2 in relation to spectral shape. Shown are the average spectra of the four quantiles belonging to blue PC1 (top) and blue PC2 (bottom).

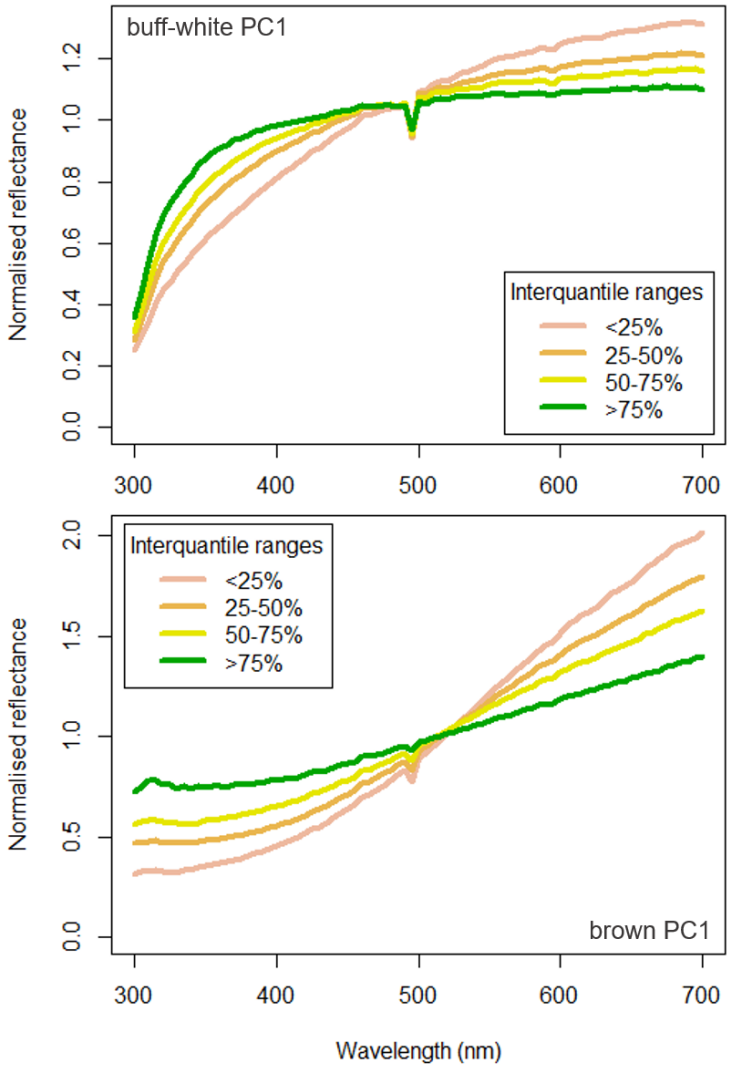


Figure S4. Buff-white PC1 and brown PC1 in relation to spectral shape. Shown are the average spectra of the four quantiles belonging to buff-white PC1 (top) and brown PC1 (bottom).

General discussion

Sexual selection by female choice is deemed responsible for much of the bizarre and bright ornaments that adorn males of many species (Darwin, 1871). Conspicuous seasonal male plumages of fairy-wrens are a classic example of an ornament driven by female choice, as early acquisition of the breeding plumage is critical for obtaining extra-pair paternity (EPP), a defining feature of the fairy-wren mating system (Dunn & Cockburn, 1999; Karubian, 2002; Cockburn et al., 2008; Webster et al., 2008; Kingma et al., 2009; Brouwer et al., 2011; Peters et al., 2013). In my PhD research, I investigated the function of male seasonal plumage ornaments in the purple-crowned fairy-wren, *Malurus coronatus*. Despite a very similar life-history and ecology to other fairy-wrens, EPP is nearly absent in this species (Kingma et al., 2009). Nonetheless, male purple-crowned fairy-wrens annually moult into a striking purple-and-black breeding plumage and there appears to be large variation in the timing of moult (Peters et al., 2013). Therefore, the main aim of my thesis was to understand the nature of the selection pressures driving the presence of a male seasonal plumage in this genetically monogamous species. My study revealed that the evolutionary maintenance of male seasonal plumages in *M. coronatus* results from the dynamic interplay between female choice and male-male competition, illustrating a greater flexibility in function of sexual ornaments than widely appreciated, and contributing to our increasing appreciation of the complexity of sexual selection dynamics. It also provided support for the multi-signalling potential of seasonal plumages, showing that, despite a notable lack of role of the timing of seasonal moult, several other plumage components correlate with different aspects of individual quality and fitness. Below I discuss the main findings of my thesis in detail.

The great diversity of male fairy-wren colours arises from variation in multiple components of feather microstructure

Male *M. coronatus* annually develop an unusual purple crown which represents an evolutionary innovation in the genus since males of other fairy-wren species produce mostly blue and black plumages (Figs 1 and 4 in the Introduction; Fig. 1 in Chapter 1). In Chapter 1, I assessed the proximate mechanisms underlying this large diversity of non-iridescent structural colours. My results highlighted the important contribution of melanin to the production of the purple colour of male *M. coronatus*: the gradual increase in the red range seen in the spectra of the purple crown and various rufous patches (mostly produced by deposition of phaeomelanin; McGraw, 2006; Fig. 2 in Chapter 1) suggests that purple barbs most likely contain a basal layer consisting mainly of phaeomelanin (as hypothesised by Peters et al., 2013). This constitutes the only qualitative difference between purple and the other structural colours, as all other differences I observed were quantitative. All structural colours, including purple, exhibited a similar barb morphology, composed of an outer keratin cortex surrounding a highly organised medullary spongy layer and a basal layer of melanin (Fig. 4 in Chapter 1), as previously reported for other structurally coloured feathers (Prum, 2006). My key finding was that all these structural elements independently contribute to various characteristics of the colour produced (Fig. 7 in Chapter 1 and Fig. 1). The co-occurrence of multiple labile components most likely facilitates a greater variability in structural plumage colouration and explains the rapid evolution of male ornamentation in this genus (Friedman & Remeš, 2015).

Investigating the proximate mechanisms of ornament elaboration allows to identify components that affect ornamental expression and thereby may constitute potential sources of variation in costs. As the costs of both feather microstructure elaboration and melanin (both eu- and phaeomelanin) synthesis are still not well understood (but see Galván & Solano (2015) for melanogenesis-related oxidative stress), whether the production of structural colouration entails

substantial physiological costs remains unclear, especially for the (rather small) purple crown displayed by male *M. coronatus*. To advance our understanding, a first step would be to quantify the melanin composition of the purple plumage, for instance by measuring the concentration of phaeomelanin using chemical techniques (Wakamatsu et al., 2002; Ito et al., 2011). These are technically complex and require advanced equipment, and quantifying melanin was beyond the scope of my PhD. Moreover, additional research could further explore the potential physiological costs of structural plumage elaboration in fairy-wrens, including in *M. coronatus*, by examining how variation in individual condition impacts the different structural elements of barb morphology at the intraspecific level. For instance, given that the production of male breeding plumage in *M. cyaneus* is condition-dependent – with higher summer rainfall providing better conditions (Cockburn et al., 2008; van de Pol et al., 2012) – and that testosterone can be used to experimentally induce the moult into male breeding plumage in this species (Peters et al., 2000), one could assess whether inducing the moult under good vs. poor environmental conditions affects the quality of the barb microstructure.

In the next sections, I discuss my results regarding the function of the multiple components of male seasonal plumage ornaments in *M. coronatus* in relation to its unusual monogamous mating system.

Timing of pre-breeding moult in male *M. coronatus*: a vestigial trait resulting from relaxed female extra-pair mate choice?

Given that extreme levels of EPP appear to drive variation in the timing of pre-breeding moult in other fairy-wrens, I first assessed the variability of pre-breeding moult timing in male *M. coronatus*, and tested whether males derive any fitness benefits from moulting early (Chapter 2). I observed substantial individual variation and found some evidence for a potential to act as a sexual signal of male quality: moult timing shows correlations (although of limited extent)

with male age, social status and climate (Figs 2 and 3 in Chapter 2 and Fig. 1). Such patterns are very similar to what was previously observed in five species of fairy-wren (Rowley, 1981; Russell et al., 1991; Higgins et al., 2001; Cockburn et al., 2008; Webster et al., 2008; van de Pol et al., 2012; Peters et al., 2013). However, despite a thorough investigation of numerous male fitness proxies, I could not identify any fitness benefits or costs associated with early acquisition of the breeding plumage (Fig. 4 in Chapter 2 and Fig. 1). Based on the most recent phylogeny of the genus (Marki et al., 2017), early moult in male *M. coronatus* could be a vestigial trait, on its way to disappearing, owing to the radical loss of extreme levels of EPP that led to relaxed female choice (Lahti et al., 2009; Fig. 5 in Chapter 2). Hence, even though male *M. coronatus* display a seasonal plumage that appears very similar to its congeners in several aspects (except for the inclusion of phaeomelanin instead of eumelanin in feather barbs; Chapter 1), this plumage does not currently serve a similar purpose. Persistence of large variation in moult timing, as well as the observation that dominant breeders moult earlier than subordinates (Fig. 2 in Chapter 2), suggested that this plumage might have another function, possibly as a social signal of dominance and competitiveness among males, which I investigated in Chapter 3.

Although a number of recent phylogenetic studies have demonstrated that the loss of sexually selected traits – due to the weakening or removal of a source of selection – is a widespread evolutionary trend, it remains an understudied area of sexual selection (Wiens, 2001; Lahti et al., 2009). Yet, understanding why and how sexually selected traits are partially or entirely lost, is important as it can provide insights into the nature and role of various sources of selection, as well as how they might interact and affect individual fitness (Wiens, 2001; Lahti et al., 2009). More studies incorporating phylogenetic approaches are needed to identify such losses of sexual traits, and an important future step would be to assess the underlying genetic changes accompanying these phenotypic losses. Furthermore, my findings highlight the

importance of being cautious in the interpretation of comparative analyses and the generalisation of observed patterns. Fairy-wrens form a group of closely related species which have explicitly and implicitly been assumed to all be very similar (Boland & Cockburn, 2002). However, the study of particular species, such as *M. coronatus*, has disproved and continues to disprove this assumption (Cockburn et al., 2013; Buchanan & Cockburn, 2013), revealing more complex selection patterns that could be examined in the rest of the genus.

Male breeding plumage: from ornament to armament or loss-of-function?

Since Chapter 2 suggested a potential role of male breeding plumage in social signalling, I examined the function of breeding plumage completeness (or extent) in male-male competitive interactions in Chapter 3. Becoming a breeder – either by inheriting a position in the natal territory or by dispersing, usually to a neighbouring territory – is a critical step for male reproductive success, although this may be challenging given the scarcity of breeding vacancies in this species. I demonstrated that male breeding plumage plays a crucial role in this context as male subordinates with more extensive breeding plumage have a higher chance to competitively gain a breeder position (Figs 4 and 5 in Chapter 3 and Fig. 1). Moreover, experimentally presented models representing male intruders in purple colour receive more aggression from resident breeder males than models painted in dull non-breeding colour (Fig. 6 in Chapter 3), consistent with competitors in breeding plumage being perceived as a bigger territorial threat to the breeder male. Hence, a role as an indicator of competitive ability seems to drive the persistence of male breeding plumage in this species. Based on the phylogeny of the genus, two novel evolutionary scenarios could be proposed: either (1) male seasonal plumage went through a functional shift, being previously selected by female choice, then by male-male competition, or (2) it was a dual function-trait – being selected by both mechanisms (Berglund et al., 1996) – and lost its function in female choice as extreme EPP levels and female

extra-pair mate choice disappeared (Kingma et al., 2009; Chapter 2). In order to determine which scenario is more likely, further research should examine the potential intrasexual function of male breeding plumage in other fairy-wrens, which to date has been poorly studied. Nonetheless, Karubian et al. (2008) did report that in *M. melanocephalus*, dull males are socially subordinate to bright males, and bright caged stimulus males receive higher levels of aggression than dull ones, suggesting a possible dual function of male breeding plumage in this species. The idea that a trait may persist because of a functional shift is not novel (Lahti et al., 2009), and a substantial number of studies have shown that ornaments used for same-sex competition can also be preferred in mate choice ('armament-ornament' hypothesis; Berglund et al., 1996; Stern & Servedio, 2017). Moreover, spatial and temporal fluctuations in either mechanism of sexual selection may be common (Hunt et al., 2009; Miller & Svensson, 2011). Nevertheless, the alternative scenarios proposed here imply that intra- and intersexual functions of a same ornament can be gained or lost independently, highlighting an under-appreciated flexibility in the functions of sexual ornaments.

Colour quality of the breeding plumage: evidence for multi-component signalling

A potential signalling component that seems obvious, and is most commonly considered for colourful plumages, is the quality of the colouration (Hill & McGraw, 2006a, b). In Chapter 4, I investigated the quality of the breeding colouration of male *M. coronatus* and incorporated all the other plumage patches that together compose their multidimensional colour phenotype. This approach allowed me to test various evolutionary theories – relating to condition-dependence, heritability and fitness – on ornamental colour traits in comparison with non-ornamental colour traits. Contrary to predictions of heightened condition-dependence in ornaments (Rowe & Houle, 1996; Cotton et al., 2004; Bonduriansky & Rowe, 2005), only the expression of non-ornamental colours (buff-white and brown) was associated with body condition (but some

weaknesses of our test of condition-dependence may in part explain this pattern; Fig. 3 in Chapter 4). Nevertheless, the expression of ornamental colours (purple, blue and black) was more variable (Fig. 2 in Chapter 4), and importantly, chromatic variation in the purple breeding colouration displayed substantial heritability (Fig. 4 in Chapter 4) and predicted male annual reproductive success (Figs 5 and 6 in Chapter 4 and Fig. 1). Hence, while breeding plumage completeness is a strong predictor of success when subordinate males compete for breeding vacancies (Chapter 3), the quality of the breeding plumage colouration correlates with male reproductive success once males have acquired a breeder position. Further investigation is needed to determine the potential genetic benefits that are signalled (e.g., 'good genes' contributing to offspring survival, attractiveness or competitive ability, or genetic compatibility), as well as the mechanisms underlying the association with male reproductive success (e.g., differential reproductive investment by differently coloured males or their female partners, or variation in offspring viability with male colour; Duckworth et al., 2003; Horváthová et al., 2012). Nonetheless, these findings illustrate the potential for seasonal plumages to act as multi-component signals directed to different receivers in different contexts (Møller & Pomiankowski, 1993). Moreover, the lack of consistent patterns (i.e. the apparently higher condition-dependence of the non-ornamental colours, combined with the variable heritability and strength of correlation to fitness among the ornamental colours) highlights the complexity of visual signals and the value of using more integrated approaches when studying complex phenotypes. If I had focused exclusively on the purple breeding colouration, as is a very common approach, this would have led to the conclusion that most observed patterns fit with our expectations, whereas incorporating other colours displayed by the species demonstrated that the complete picture might be more complex. Therefore, when assessing the function of a given ornament, using appropriate non-ornamental 'control' traits can broaden our inference: this may be done by comparing several traits of the same type (e.g., metric traits,

colour traits, acoustic traits, etc.) but subject to variable strength of selection, or by randomly selecting several traits of the same type, some among ornamental traits and others among non-ornamental traits.

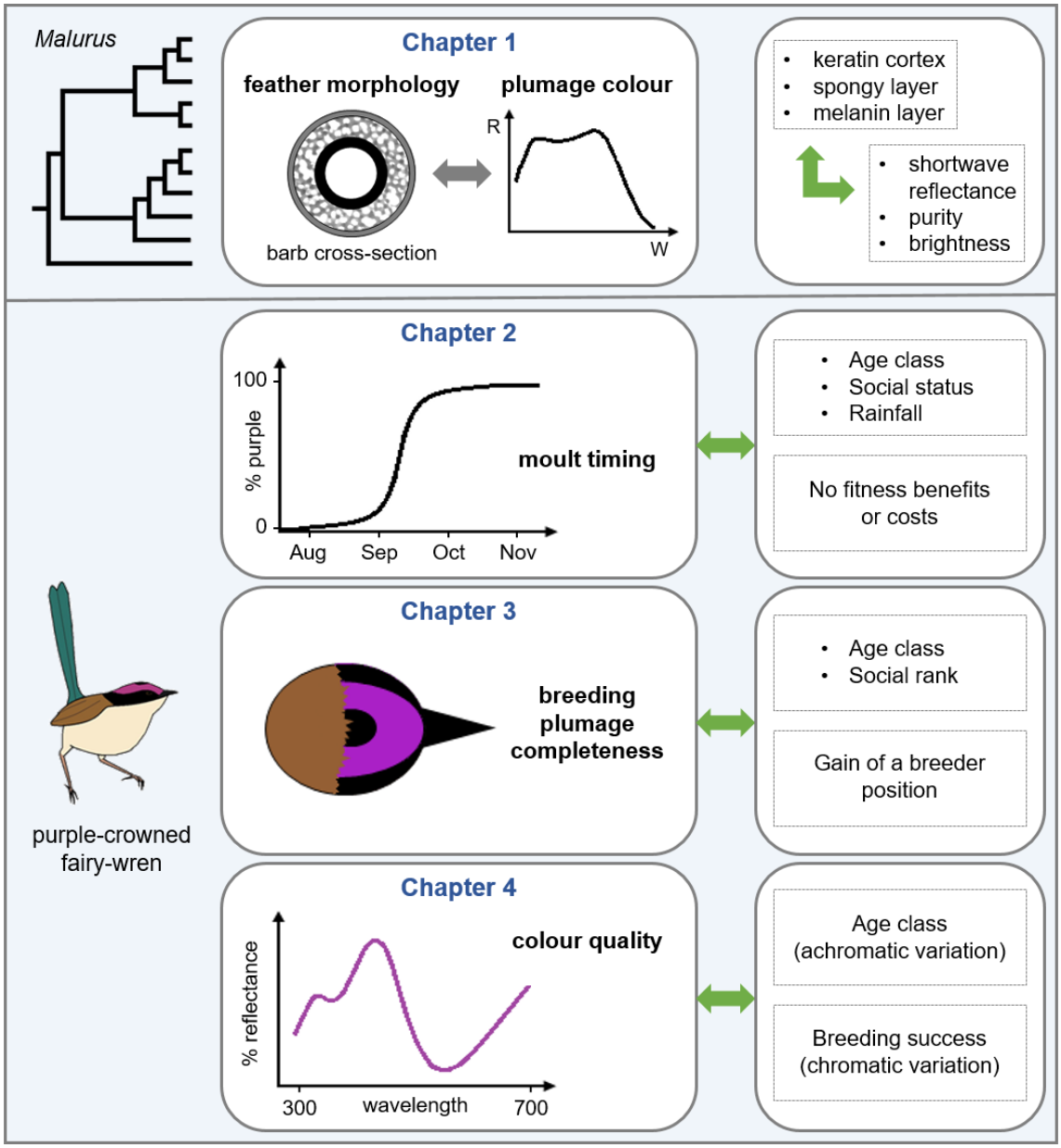


Figure 1. Main findings of my thesis for Chapters 1 to 4.

Concluding remarks and future directions

The main aim of my thesis was to understand the function of male seasonal plumages in *M. coronatus*, a genetically monogamous fairy-wren. The study of several components of male

breeding plumage demonstrated that in this species seasonal plumages provide males with multiple signalling opportunities. Plumage completeness (Chapter 3) signals male competitive ability, primarily for access to breeding vacancies, whereas colour quality (Chapter 4) may signal some genetic benefits and predicts male reproductive success (Fig. 1). Therefore, male seasonal plumage constitutes a complex multi-component signal directed to different receivers for different purposes, depending on the context. On the other hand, despite the timing of moult being critically important in other fairy-wrens, and despite substantial variation in *M. coronatus*, I was unable to identify any selection, or signalling potential, other than social status (Chapter 2). It is possible that the higher susceptibility to predation risk often associated with conspicuous colouration favours a less conspicuous non-breeding plumage (Endler 1980; Götmark & Hohlfält, 1995; Slagsvold et al., 1995; Huhta et al., 2003). For example, male *M. cyaneus* perceive themselves to be at a higher risk of predation while in conspicuous blue plumage and adjust their behaviour accordingly, spending more time in cover and showing greater vigilance (McQueen et al., 2017). Likewise, predation pressure could maintain the annual moult in male *M. coronatus*. Alternatively, maintaining a colourful plumage during the entire dry season when food resources are less abundant may entail important costs.

Although I gained some valuable information about the function of male seasonal plumages in *M. coronatus*, additional experimental studies could be performed to continue this research. For instance, experimental removal of resident breeder males from their territories could provide further insights into male-male competition for breeding vacancies and the behaviour of unpaired females in such a context. Furthermore, the extent or colour of the breeding plumage could be experimentally manipulated on living birds (in a realistic way based on the methods used in Chapter 3) and the consequences of these treatments on various aspects of male fitness assessed. Moreover, similar studies in female *M. coronatus* could help to determine whether the evolution of female seasonal plumages is driven by the same or different selection pressures

(Karubian, 2013) and shed some new light on what I observed in males. Such studies may also inform on the possible role of assortative mating in the evolution of plumage ornaments in both sexes (MacDougall & Montgomerie, 2003; Pryke & Griffith, 2007; Fargevieille et al., 2017), an aspect that has not yet been studied in *Malurus* species given that most female fairy-wrens display a rather dull plumage with no seasonal change (Rowley & Russell, 1997).

In conclusion, the investigation of several plumage components in relation to both mechanisms of sexual selection, combined with phylogenetic information, has improved our understanding of some understudied evolutionary scenarios (trait vestigialisation and loss, co-option of female choice signals for use in male-male competition; Chapters 2 and 3) and allowed to uncover some new ones (loss of function in a dual-function trait; Chapter 3), offering new perspectives in understanding the complex dynamics of sexual selection. Hence, my PhD research emphasises the importance of performing integrated studies of all signal components of an ornamental trait in closely related species and the great value of phylogenetic tools in the assessment of trait evolutionary trajectories. It is undeniable that a tremendous amount of theoretical and empirical work has already been devoted to the evolution and maintenance of elaborate ornaments (Darwin, 1871; Andersson, 1994); nonetheless, my thesis shows that this field remains very complex, with a number of understudied and not well understood areas, which will hopefully motivate further investigation that builds upon my findings.

References

Andersson, M. (1994). *Sexual selection*. Princeton University Press.

Berglund, A., Bisazza, A., & Pilastro, A. (1996). Armaments and ornaments: an evolutionary explanation of traits of dual utility. *Biological Journal of the Linnean Society*, **58**, 385-399.

Boland, C. R., & Cockburn, A. (2002). Short sketches from the long history of cooperative breeding in Australian birds. *Emu*, **102**, 9-17.

Bonduriansky, R., & Rowe, L. (2005). Sexual selection, genetic architecture, and the condition dependence of body shape in the sexually dimorphic fly *Prochyliza xanthostoma* (Piophilidae). *Evolution*, **59**, 138-151.

Brouwer, L., van de Pol, M., Atema, E. L. S., & Cockburn, A. (2011). Strategic promiscuity helps avoid inbreeding at multiple levels in a cooperative breeder where both sexes are philopatric. *Molecular Ecology*, **20**, 4796-4807.

Buchanan, K. L., & Cockburn, A. (2013). Fairy-wrens and their relatives (Maluridae) as model organisms in evolutionary ecology: the scientific legacy of Ian Rowley and Eleanor Russell. *Emu*, **113**, i-vii.

Cockburn, A., Osmond, H. L., & Double, M. C. (2008). Swingin' in the rain: condition dependence and sexual selection in a capricious world. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 605-612.

Cockburn, A., Brouwer, L., Double, M. C., Margraf, N., & van de Pol, M. (2013). Evolutionary origins and persistence of infidelity in *Malurus*: the least faithful birds. *Emu*, **113**, 208-217.

Cotton, S., Fowler, K., & Pomiankowski, A. (2004). Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? *Proceedings of the Royal Society of London B: Biological Sciences*, **271**, 771-783.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. John Murray.

Delhey, K., Hall, M., Kingma, S. A., & Peters, A. (2013). Increased conspicuousness can explain the match between visual sensitivities and blue plumage colours in fairy-wrens. *Proceedings of the Royal Society of London B: Biological Sciences*, **284**, 20121771.

Dunn, P. O., & Cockburn, A. (1999). Extrapair mate choice and honest signaling in cooperatively breeding superb fairy-wrens. *Evolution*, 938-946.

Duckworth, R. A., Badyaev, A. V., & Parlow, A. F. (2003). Elaborately ornamented males avoid costly parental care in the house finch (*Carpodacus mexicanus*): a proximate perspective. *Behavioral Ecology and Sociobiology*, **55**, 176-183.

Endler, J. A. (1980). Natural selection on color patterns in *Poecilia reticulata*. *Evolution*, **34**, 76-91.

Fargevieille, A., Grégoire, A., Charmantier, A., del Rey Granado, M., & Doutrelant, C. (2017). Assortative mating by colored ornaments in blue tits: space and time matter. *Ecology and Evolution*, **7**, 2069-2078.

Friedman, N. R., & Remeš, V. (2015). Rapid evolution of elaborate male coloration is driven by visual system in Australian fairy-wrens (Maluridae). *Journal of Evolutionary Biology*, **28**, 2125-2135.

Galván, I., & Solano, F. (2015). Melanin chemistry and the ecology of stress. *Physiological and Biochemical Zoology*, **88**, 352-355.

Götmark, F., & Hohlfält, A. (1995). Bright male plumage and predation risk in passerine birds: are males easier to detect than females? *Oikos*, 475-484.

Higgins, P. J., Peter, J. M., & Steele, W. K. (Eds.). (2001). *Handbook of Australian, New Zealand and Antarctic Birds, vol. 5: Tyrant-flycatchers to Chats*. Oxford University Press.

Hill, G. E., & McGraw, K. J. (Eds.). (2006a). *Bird coloration, vol. 1: mechanisms and measurements*. Harvard University Press.

Hill, G. E., & McGraw, K. J. (Eds.). (2006b). *Bird coloration, vol. 2: function and evolution*. Harvard University Press.

Horváthová, T., Nakagawa, S., & Uller, T. (2012). Strategic female reproductive investment in response to male attractiveness in birds. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 163-170.

Huhta, E., Rytkönen, S., & Solonén, T. (2003). Plumage brightness of prey increases predation risk: An among-species comparison. *Ecology*, **84**, 1793-1799.

Hunt, J., Breuker, C. J., Sadowski, J. A., & Moore, A. J. (2009). Male–male competition, female mate choice and their interaction: determining total sexual selection. *Journal of Evolutionary Biology*, **22**, 13-26.

Ito, S., Nakanishi, Y., Valenzuela, R. K., Brilliant, M. H., Kolbe, L., & Wakamatsu, K. (2011). Usefulness of alkaline hydrogen peroxide oxidation to analyze eumelanin and pheomelanin in various tissue samples: application to chemical analysis of human hair melanins. *Pigment cell & melanoma research*, **24**, 605-613.

Karubian, J. (2002). Costs and benefits of variable breeding plumage in the Red-backed Fairy-wren. *Evolution*, **56**, 1673-1682.

Karubian, J., Sillett, T. S., & Webster, M. S. (2008). The effects of delayed plumage maturation on aggression and survival in male red-backed fairy-wrens. *Behavioral Ecology*, **19**, 508-516.

Karubian, J. (2013). Female ornamentation in *Malurus* fairy-wrens: a hidden evolutionary gem for understanding female perspectives on social and sexual selection. *Emu-Austral Ornithology*, **113**, 248-258.

Kingma, S. A., Hall, M. L., Segelbacher, G., & Peters, A. (2009). Radical loss of an extreme extra-pair mating system. *BMC Ecology*, **9**, 15.

Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., ... & Foster, S. A. (2009). Relaxed selection in the wild. *Trends in Ecology & Evolution*, **24**, 487-496.

MacDougall, A. K., & Montgomerie, R. (2003). Assortative mating by carotenoid-based plumage colour: a quality indicator in American goldfinches, *Carduelis tristis*. *Naturwissenschaften*, **90**, 464-467.

Marki, P. Z., Jønsson, K. A., Irestedt, M., Nguyen, J. M., Rahbek, C., & Fjeldså, J. (2017). Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, **107**, 516-529.

McGraw, K. J. (2006) Mechanics of melanin coloration in birds. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1*. (pp. 243-294). Harvard University Press.

McQueen, A., Naimo, A. C., Teunissen, N., Magrath, R. D., Delhey, K., & Peters, A. (2017). Bright birds are cautious: seasonally conspicuous plumage prompts risk avoidance by male superb fairy-wrens. *Proceedings of the Royal Society London B: Biological Sciences*, **284**, 20170446.

Miller, C. W., & Svensson, E. I. (2014). Sexual selection in complex environments. *Annual Review of Entomology*, **59**, 427-445.

Møller, A. P., & Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology*, **32**, 167-176.

Ödeen, A., Pruett-Jones, S., Driskell, A. C., Armenta, J. K., & Håstad, O. (2012). Multiple shifts between violet and ultraviolet vision in a family of passerine birds with associated changes in plumage coloration. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 1269-1276.

Peters, A., Astheimer, L. B., Boland, C. R., & Cockburn, A. (2000). Testosterone is involved in acquisition and maintenance of sexually selected male plumage in superb fairy-wrens, *Malurus cyaneus*. *Behavioral Ecology and Sociobiology*, **47**, 438-445.

Peters, A., Kingma, S. A., & Delhey, K. (2013). Seasonal male plumage as a multi-component sexual signal: Insights and opportunities. *Emu*, **113**, 232-247.

Pryke, S. R., & Griffith, S. C. (2007). The relative role of male vs. female mate choice in maintaining assortative pairing among discrete colour morphs. *Journal of Evolutionary Biology*, **20**, 1512-1521.

Prum, R. O. (2006). Anatomy, physics, and evolution of structural colours. In G. E. Hill & K. J. McGraw (Eds.), *Bird coloration, vol. 1.* (pp. 295-353). Harvard University Press.

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London B: Biological Sciences*, **263**, 1415-1421.

Rowley, I. (1981). The communal way of life in the splendid wren, *Malurus splendens*. *Zeitschrift für Tierpsychologie*, **55**, 228-267.

Russell, E. M., Rowley, I., Brown, R. E., & Brown, M. (1991). Acquisition of nuptial plumage in the Red-winged Fairy-wren *Malurus elegans*. *Corella*, **15**, 125-133.

Slagsvold, T., Dale, S., & Kruszewicz, A. (1995). Predation favours cryptic coloration in breeding male pied flycatchers. *Animal Behaviour*, **50**, 1109-1121.

Stern, C. A., & Servedio, M. R. (2017). Evolution of a mating preference for a dual-utility trait used in intrasexual competition in genetically monogamous populations. *Ecology and Evolution*, **7**, 8008-8016.

van de Pol, M., Osmond, H. L., & Cockburn, A. (2012). Fluctuations in population composition dampen the impact of phenotypic plasticity on trait dynamics in superb fairy-wrens. *Journal of Animal Ecology*, **81**, 411-422.

Wakamatsu, K., Ito, S., & Rees, J. L. (2002). The usefulness of 4-amino-3-hydroxyphenylalanine as a specific marker of pheomelanin. *Pigment Cell Research*, **15**, 225-232.

Webster, M. S., Varian, C. W., & Karubian, J. (2008). Plumage colour and reproduction in the Red-backed Fairy-wren: why be a dull breeder? *Behavioral Ecology*, **19**, 517-524.

Wiens, J. J. (2001). Widespread loss of sexually selected traits: how the peacock lost its spots. *Trends in Ecology & Evolution*, **16**, 517-523.