Genetics of plumage colour polymorphism in the Gouldian finch (*Erythrura gouldiae*)

by

Hanyuan Zhang

A thesis submitted for the degree of Doctor of Philosophy

Department of Animal and Plant Sciences

The University of Sheffield

March 2015

ABSTRACT

External pigmentation plays an important role in the evolution, ecology and behaviour of birds. However, our understanding of the genetic basis of variation in plumage colouration is as yet very limited, and especially for loci that control carotenoid pigmentation. Therefore, this thesis details the exploration of the molecular genetic basis of head colour polymorphism that occurs in natural populations of the Gouldian finch *(Erythrura gouldiae)*.

In order to identify the genomic regions encoding variation in plumage traits, a microsatellite-based genetic linkage map was constructed. There was a high degree of synteny throughout the microsatellite map between Gouldian finch and the zebra finch genome assembly. Restriction site associated DNA (RAD) sequencing was then used in a genome-wide association study (GWAS) in an attempt to identify the genomic location of the *Yellow* locus responsible for determining the yellow-headed/red-headed colour polymorphism. A candidate region on chromosome 8 was discovered to contain significant allele frequency differences. Subsequently, using the candidate gene approach, haplotypes co-segregating with different phenotypes were identified in the *CYP2J2-2* gene on an unassembled section of zebra finch chromosome 8. This result suggests that the *CYP2J2-2* gene is associated with this carotenoid-based polymorphism. From the linkage map of chromosome 8, a 2.9-Mb region also showed evidence of association with the yellow/red polymorphism. The available sample size was limited so these results need to be confirmed by testing more samples. Additionally, the hypothesis, based on a previous study, was tested and rejected that the melanin-based polymorphism determined by the Gouldian finch *Red* locus was caused by an inversion. In summary, this thesis identified a candidate region for the *Yellow* locus and provides an empirical foundation for the future study of the genetic mechanism and molecular evolution of plumage colour polymorphism in the Gouldian finch.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere thanks to my supervisor, Professor Terry Burke, for his encouragement and support during my four years' PhD process. I thank him for his professional guidance and the generous time for giving me feedback throughout my research. Without his constant guidance, this thesis could not have been completed. His wisdom with endless knowledge and valuable experience makes him a role model for me.

Thanks to Sarah Pryke and Simon Griffith for all the samples they provided. Thanks to Professor Mike Siva-Jothy, an understanding and helpful supporter who gave me suggestions on my first-year upgrade report and also encouraged me to fight for my PhD as an overseas student. Thanks also to Dr Thomas Webb for suggestions at my upgrade meeting.

Thanks to Karim Gharbi, Cathlene Eland and other staff in the Genepool genomics facility at the University of Edinburgh for the RAD sequencing service they have provided.

I wish to thank Dr. Kang-Wook Kim for the molecular training in the lab, guidance and experimental and data analysis discussions throughout. I really appreciate the help of Andy Krupa and Stephen Dorward, who helped me to test the microsatellites in chapter 4. Andy Krupa and Gavin Horsburgh also offered me great support when I was working in the lab. Deborah Dawson gave me valuable suggestions for microsatellite map construction. Celine Pagnier helped me a lot in arranging the meetings with Terry Burke. A lot of my PhD and postdoc friends, Haslina Razali, Claudius Kerth, Mauricio Montano Rendon, Steeves Buckland and Caitriona McInerney, made my PhD life more pleasurable and colourful. Thanks to all my friends in Sheffield and in my home country for companionship and love.

Thanks for the financial support from China Scholarship Council (CSC), Department for Business Innovation & Skills (BIS), and the University of Sheffield.

Finally, my mum and dad, thanks to you two for the endless love and support. I dedicate this thesis to my beloved parents. I love you!

CONTENTS

ABSTRACT **i**

ACKNOWLEDGEMENTS **ii**

CONTENTS **iv**

CHAPTER ONE: General introduction **1**

1.1. Pigments in birds **1**

1.1.1. Carotenoid-based colouration 1

1.1.2. Melanin-based colouration 3

1.1.3. Structural and uncommon colours 4

1.2. Plumage colouration polymorphism hypotheses **5**

1.2.1. Apostatic selection or frequency-dependent selection hypothesis 6

1.2.2. Disruptive selection hypothesis 6

1.2.3. Non-random mating hypothesis 6

1.3. Genetic basis of colour variation **7**

1.3.1. Single-locus control of colour polymorphism 7

1.3.2. Colour polymorphisms 8

1.3.3. Genetic basis of plumage colour polymorphisms 8

1.3.4. The ecological and physiological functions of colouration 9

1.4. The mechanism of carotenoid pigmentation **10**

1.4.1. The classification of carotenoids 10

1.4.2. The biological function of carotenoids 10

1.4.3. Carotenoid related functions, synthesis, and genetics in birds 11

1.5. Genetics of carotenoid colouration **12**

1.6. Study species: the Gouldian finch **13**

1.6.1. Distribution, history, threat, and recovery 13

1.6.2. Gouldian finch plumage colour polymorphism 15

1.7. Outline of this thesis **16**

CHAPTER TWO: Linkage mapping on the autosomal chromosomes **19**

Abstract **19**

2.1. Introduction **20**

2.2. Materials and methods **24**

2.2.1. Mapping population and genomic DNA extraction 24

2.2.2. Microsatellite marker development 25

2.2.3. Marker amplification and polymorphism 25

2.2.4. Multiplexing microsatellite markers 26

2.2.5. Marker genotyping 31

2.2.6. Three-step error checking 31

2.2.7. Construction of autosomal chromosome genetic map 32

2.2.8. Genome coverage estimate 32

2.3. Results **33**

2.4. Discussion **42**

CHAPTER THREE: A genome-wide association study of head colour polymorphism **44**

Abstract **44**

3.1. Introduction **45**

3.1.1. Restriction-site associated DNA 45

3.1.2. Genome-wide association studies 46

3.1.3. *Stacks* 47

3.2. Methods **50**

3.3. Results **58**

3.4. Discussion **61**

CHAPTER FOUR: A genetic mapping and candidate gene approach of the yellow-headed colour trait **63**

Abstract **63**

4.1. Introduction **65**

4.1.1. Candidate genes for carotenoid colouration in the Gouldian finch 65

4.1.2. The candidate gene approach 66

4.2. Materials and methods **67**

4.2.1. Chromosome 8 mapping and linkage analysis 67

4.2.2. Candidate gene approach: the *CYP2J2-2* gene 70

4.3. Results **74**

4.3.1. Construction of chromosome map and linkage analysis 74

4.3.2. Construction of haplotypes and genetic analysis of *CYP2J2-2* gene 83

4.4. Discussion **87**

CHAPTER FIVE: An extended study of melanin-based head colour plumage polymorphism **89**

Abstract **89**

5.1. Introduction **90**

5.1.1. Melanin-based colouration 92

5.1.2. Melanin-based head colour polymorphism in the Gouldian finch 92

5.1.3. Recent research progress 93

5.2. Materials and Methods **95**

5.2.1. Samples 95

5.2.2. Amplification of the *Red* core region and extended region 95

5.2.3. Confirmation of the identity and size of amplified products 98

5.3. Results 98

5.4. Discussion 102

CHAPTER SIX: General discussion 103

6.1. Carotenoid-based colouration 103

6.2. Melanin-based colouration 105

REFERENCES 107

CHAPTER ONE: General introduction

1.1. Pigments in birds

Pigmentation in birds is controlled via a limited number of routes, being due primarily to the deposition of carotenoids and melanins, and to the development of structurally-based colour (Hill and McGraw, 2006). Here, I discuss carotenoids first, since carotenoids are the main pigment focused on in this study.

1.1.1. Carotenoid-based colouration

Carotenoids generate a range of red, orange and yellow hues in avian feathers, and are the second most-common class of pigment in the avian integument. Carotenoids occur extensively in vertebrate integument tissues, including feathers, cheeks, wattles, combs, beaks, eyes and tarsal skin (Fox and Vevers, 1960). The pigments typically produce a red, orange or yellow colouration. Carotenoids are lipid-soluble hydrocarbons with light-absorbing properties. Carotenoids are divided into carotenes and xanthophylls, which can be again subdivided into hydroxyl-carotenoids or ketocarotenoids, depending on their molecular structure (Goodwin, 1984).

Recently, carotenoid-based colouration became the focus of interest in evolutionary, functional and behavioural studies of morphology (Hill, 2002; McGraw, 2006). In birds, carotenoids have functions in providing resistance to oxidation, and providing an honest signal of individual quality via ornamental traits (Hill, 1990; Krinsky, 1993; Lozano, 1994; Vershinin, 1999).

Identity, occurrence, and distribution of carotenoids

Birds lack any enzyme for manufacturing carotenoids and, therefore, the only source of carotenoids for birds is from their food (Goodwin, 1984). Carotenoids are sequestered from both plant and animal components of the diet. The carotenoid pigmentation in the plumage of captive birds has been shown to be increased by feeding them on a wild variety of foods with a deep orange or red colour (Beebe, 1906; McGraw and Hill, 2001; McGraw *et al.,* 2001). The greater flamingo, *Phoenicopterus ruber*, is a typical example in obtaining the carotenoid pigments for its red feathers from pigmented crustaceans (Fox, 1962). In the Gouldian finch (*Erythrura gouldiae*), the focal species in this study, three different types of carotenoids have been identified from yellow and red feathers. Lutein, zeaxanthin, and 3'-dehydrolutein occur in yellow plumage, and astaxanthin, a-doradexanthin, and papilioerythrinone occur in red plumage (Stradi, 1999).

Carotenoid metabolism

There are two alternative biochemical strategies for carotenoid colouration: either accumulating just one or a few specific carotenoid forms, or accumulating many different types of carotenoids. In birds, carotenoids are metabolically transformed in the follicles that produce pigmented feathers (del Val *et al.*, 2009). The most basic and conserved metabolic pathway for producing carotenoids is cleaving carotenes into vitamin A (Bauernfeind, 1981). In carotenoid metabolism, yellow carotenoids (e.g. the canary xanthophylls, anhydrolutein, dehydrolutein) are made by dehydrogenation. Red carotenoids (e.g. astaxanthin and canthaxanthin), in contrast, are generated by the 4-oxidation process (Stradi, 1998). Hill and McGraw (2004) first demonstrated that carotenoid metabolites are deposited in the avian integument by detecting lipid fractions that contained derived carotenoids in male American goldfinches growing yellow feathers. However, this is not the only way that carotenoids are

produced and deposited in the integument. Different types of carotenoids can also be produced in the duodenum, and then enriched in the liver, and finally deposited into feathers by circulation in the blood plasma (Khachik *et al.,* 2002; McGraw *et al.,* 2002; McGraw and Schuetz, 2004). So far, there is no evidence in birds to show that red ketocarotenoids are formed in tissues other than the integument.

Physiological actions of carotenoids

Carotenoids have extensive roles in an animal's physiology and biochemistry. For example, carotenoids were reported to have an antioxidant role in protecting the immune system by binding to free radicals during the activation and proliferation of immune cells (Chew, 1993; Hughes, 2001). Carotenoids have also been speculated to have potent effects on embryonic development and health protection (Blount *et al.*, 2000; Hammershøj *et al.*, 2010). The physiological functions of carotenoids, requiring them in tissues other than the integument, potentially create a trade-off with being colourful. Lozano (1994) hypothesised that only the highest-quality individuals will have the capacity to allocate carotenoids to the integument for sexual ornamentation, rather than using them in body maintenance, when they face a carotenoid allocation trade-off. This hypothesis has been supported by a series of studies (Brawner *et al.,* 2000; Lindström and Lundström, 2000; McGraw and Hill, 2000; Saks *et al.,* 2003): more colourful birds are healthier. Carotenoid colours might therefore have evolved as an honest signal for individual quality.

1.1.2. Melanin-based colouration

Melanins are the most abundant pigment in animals that can be synthesized *in vivo*, and provide a base pigment for most feathers. Melanin was first isolated from feathers more than a century ago (Verne, 1926), but only in recent decades has attention been paid to the

significance of two main forms of melanin – phaeomelanins and eumelanins – in the colouration of birds (Hill and McGraw, 2006). Phaeomelanin tends to be brown or buff; conversely, eumelanin presents black and grey colours. As the base pigment in feathers, melanins produce a variety of tones and patterns of black, grey and brown colours. However, the intensity and the absorbance of these colours can be modified by the tissues' microstructural arrangements (Finger *et al.*, 1992), so producing a wide range of different structural colours (Zi *et al.*, 2003). Melanins can also mask other types of brightly coloured pigments (Moreau, 1958). Melanin synthesis during feather growth and keratinization is controlled by melanocytes.

Melanins can be synthesized endogenously by birds in peripheral tissues, like the skin, especially the epidermis where the melanocytes are held (Mason and Frost-Mason, 2000). A variety of factors regulates the synthesis and accumulation of melanins, including season, sex, and integumentary form. Tyrosinase activity, and thus melanin synthesis, are determined by strong genetic control (e.g. *melanocortin-1 receptor*, *MC1R*) (Takeuchi *et al.,* 1996) as well as the stimulation of hormones (such as sex steroids, LH, thyroxin, and MSH) (Teshigawara *et al.*, 2001) in birds. There might also be other elements that affect the production of melanin, such as substrate availability or mineral concentration.

Melanins play many roles in biochemistry and physiology, including as antioxidants that protect cells from oxidation (Lozano, 1994), cation chelators that avoid damage from toxic heavy metal accumulation (Riley, 1997), tissue strengtheners (Riley, 1992) and antimicrobial deterrents (Shawkey *et al.*, 2003), as well as photoprotectants (Ortonne, 2002).

1.1.3. Structural and uncommon colours

There are two distinct ways in which colour is formed in feathers: one is by pigment, the

other via the refraction of light caused by feather structure. Feathers can be coloured by either one or both routes. Structural colours are significant in ornamental plumage and skin (Prum and Torres, 2003). These colours are produced by physical interactions between light waves and the nano-structures within the organism. They display abundant hues that cannot be produced by chemically synthesized pigments. All of the known structural colours in birds are produced by coherent scattering; well-known examples of such colours include butterfly wing scales and peacock tail covert feathers (Fox, 1976; Durrer, 1986; Parker, 1999).

Birds chemically synthesize classes of colorants other than carotenoids and melanins. At least four other, rarer, classes of pigmentation have been identified: pterins (Fox, 1976), psittacofulvins (Forshaw, 1989), porphyrins (Marks, 1969) and flavins (Gliszczynska-Swiglo and Koziolowa, 2000). As yet, there is no full understanding about the function or evolution of these colorants, and of how these different rare colorants are used.

1.2. Plumage colouration polymorphism hypotheses

Plumage colour patterning is closely related to adaptive benefits in birds. Different colour pigmentation produces striking phenotypic differences. These morphological characteristics enable birds to hide from or to, alternatively, warn potential predators, to hunt prey, attract mates, and for other behavioural benefits (Hill and McGraw, 2006). Numerous studies of the evolution, genetics and maintenance of plumage or pelage colour polymorphism have been completed over the past few decades, e.g. adaptive melanism in mice (Hoekstra and Nachman, 2003) and plumage traits involved in mate choice (Mundy *et al.*, 2004). In the Gouldian finch, the head colour polymorphism is potentially an honest signal that reflects behavioural and physiological differences (e.g. pre- and post-zygotic incompatibilities; King *et al.*, 2015). Several main hypotheses have been presented to explain the evolutionary

mechanisms, maintenance, functions and fitness consequences of plumage colour polymorphisms in birds (Galeotti *et al.,* 2003).

1.2.1. Apostatic selection or frequency-dependent selection hypothesis

Apostatic selection, or negative frequency-dependent selection, is a form of selection that favours individuals of the rarer morph in the local vicinity, for example because the unusual morph is less familiar to potential prey species (Clarke, 1962; Rohwer, 1983; Rohwer and Paulson, 1987). In other words, prey will react less readily to the unusual morph. Under this hypothesis, the prey species are viewed as the apostatic selective agent, and any consequent slight advantage will spread in the population, leading to a balanced polymorphism.

1.2.2. Disruptive selection hypothesis

Mather (1955) first proposed that disruptive selection would favour the morphologically extreme individuals in a population. Heterogeneity in space and time (Galeotti *et al.,* 2003) is crucial for the evolution of morphological traits. Environmental factors like habitat and climatic differences may select for colour polymorphisms and also for changes in the physiological functions of colour pigmentation (Hoekstra and Nachman, 2003). Pigmentation might affect an individual's thermal balance. For instance, birds in cold and dry climates tend to have pale plumage compared with dark coloured plumage in hot humid areas due to different degrees of melanin deposition (Bartle and Sagar, 1987; Galeotti and Cesaris, 1996).

1.2.3. Non-random mating hypothesis

The maintenance of colour polymorphism may be caused by mate-choice preference (Mundy

et al., 2004). In this case, colouration is a key component of communication within species (Butcher and Rohwer, 1988). There are two classes of the non-random mating hypothesis: (1) Sexual selection: female preference for striking males may select for and maintain colour polymorphism by a trade-off with natural selection (Johnston *et al.*, 2013). The intensity of pigmentation, considered to be a costly trait, should reflect fitness indicators (e.g. the nurture and health) of individuals. (2) Disassortative mating: colour polymorphism may result from a preference for mating with a partner of a different morph. As the colour is genetic this behaviour will result in more heterozygous offspring.

In addition, Galeotti *et al.* (2003) also indicate that colour polymorphism may be neutral, primordial and non-adaptive. The above mechanisms are not mutually exclusive and may co-exist in the wild at the same time.

1.3. Genetic basis of colour variation

1.3.1. Single-locus control of colour polymorphism

Many colour traits have simple, Mendelian inheritance patterns and are controlled by a single locus (e.g. plumage colour polymorphism in the Gouldian finch, *Erythrura gouldiae;* white / blue plumage colour in the snow goose, *Chen caerulescens*). In a population, the segregation of alleles for different colour morphs is an important clue about the likely genetic basis of colouration. Other evidence comes from the genetic transmission of colour variation among subspecies bred in captivity or between species in natural hybrids. The genetic basis of Mendelian colouration has been described in many domesticated avian species, such as in the domestic chicken (*Gallus gallus*), rock pigeon (*Columba livia*), and zebra finch (*Taeniopygia guttata*; Landry, 1997).

1.3.2. Colour polymorphisms

Colour polymorphisms are generally defined by Buckley (1987) as the presence of different plumage morphs or phases – discrete categories of plumage colouration in individuals of the same sex and age in an interbreeding population. The polymorphism rate among avian species has been elucidated by Galeotti *et al.* (2003) to be about 3.5% of 334 species. Most polymorphisms affect both sexes in adults, yet polymorphisms also exist that affect only a single sex in adult plumage or only immature individuals (Mulder *et al.,* 2002). The most common polymorphisms affect only the distribution of melanin. But polymorphism is also caused by pigments other than melanin. An autosomal locus controls a red/yellow carotenoid-based cheek colour polymorphism in the Gouldian finch (*Erythrura gouldiae*; Southern, 1945). Polymorphisms that involve changes in the distribution of both melanin and carotenoids simultaneously are extremely rare. The only clearly known case is the Gouldian finch, in which one sex-linked locus controls variation between black- and red-faced morphs, and a separate autosomal locus determines the presence of red or yellow faces. How are such polymorphisms maintained in the face of inevitable allelic loss under genetic drift? This may be due to gene flow, assortative or disassortative mating, overdominant selection, opposing directional selection in different environments, and frequency-dependent selection (Hartl and Clark, 1997). Not surprisingly, different mechanisms are involved in different species (Lank, 2002).

1.3.3. Genetic basis of plumage colour polymorphisms

Thus far, only a few avian pigmentation or patterning genes have been isolated, including the *MC1R* gene associated with melanic polymorphisms (Takeuchi *et al.*, 1996; Mundy, 2005), the *TYRP1* gene involved in the biosynthesis of melanin in pied flycatchers *(Ficedula hypoleuca)* (Leskinen *et al.*, 2012), the *melanophilin* gene that produces colour dilution in

the chicken (Vaez *et al.*, 2008) and the *agouti* gene in Japanese quail *(Coturnix japonica)* (Nadeau *et al.*, 2008). Only *MC1R* and *TYRP1* have been shown to account for polymorphisms in natural populations. There is little known about the molecular genetic control of carotenoid colouration in birds. The lack of information on the control of carotenoid deposition in feathers is perhaps the most conspicuous gap in current studies of polymorphism. This gap is due to a deficiency of knowledge of the basic biochemical and physiological pathways of carotenoid deposition in birds. In view of the significance of carotenoids in plumage traits involved in sexual selection and in signalling individual quality, the identification of such loci should be a priority.

A knowledge of the genetic basis of Gouldian finch plumage polymorphism would be valuable to understanding the evolutionary dynamics of this colour trait and for identifying what might be homologous genetic mechanisms of colour polymorphism among different species. In the Gouldian finch, there is an epistatic interaction between the two naturally occurring polymorphic loci: the *Yellow* (red or yellow face) locus is epistatic to the *Red* (red or black face) locus (Southern, 1945; Murray, 1963).

1.3.4. The ecological and physiological functions of colouration

Genes and the environment both influence colour expression. The environmental parameters include age, nutrition, season, parasites, territory quality, and the social environment (Hill and McGraw, 2006). Birds use colour to exchange information for many different purposes, including individual quality, genetic compatibility, individual identity, attractiveness, kinship, and to signal their presence. Colour functions as an intrasexual signal of aggression and dominance, such as fighting ability, signals social status, correlates of dominance, and physiological costs (Rohwer, 1975), and affects intrasexual selection – direct competition between individuals of the same sex for mates. Additionally, in intersexual selection females

prefer mates with ornamental colouration. One hypothesis is that social hierarchies are driven by mixed evolutionarily stable strategies (Maynard, 1982; Számadó, 2000). There are many benefits to females of assessing colour displays. Direct benefits signalled by colour displays include dominance-related benefits, courtship feeding, paternal care of offspring, nest defence, and parasite avoidance (Reynolds and Gross, 1990). Indirect benefits that have been described to be signalled by colour include immunocompetence, good-genes benefits, and genetic compatibility (Mays Jr and Hill, 2004).

In captive populations of Gouldian finches, the three head colour morphs mate assortatively (Fox *et al.,* 2002); no evidence has been found for a general female colour morph preference. The maintenance of the colour polymorphism in Gouldian finches has not yet been explained.

1.4. The mechanism of carotenoid pigmentation

1.4.1. The classification of carotenoids

Carotenoids are organic pigments that occur naturally in many plants, bacteria and fungi. There is no clear evidence that carotenoids can be manufactured by animals (Goodwin, 1986), except in one species of aphid (Moran and Jarvik, 2010). In birds, carotenoids can be obtained in the diet. Carotenoids are split into two classes: (1) xanthophylls: carotenoids with molecules containing oxygen, e.g. lutein and zeaxanthin; (2) carotenes: unoxygenated carotenoids, such as β-carotene and lycopene. Carotenoids are responsible for colours ranging from deep red through bright orange to pale yellow, depending on their structure.

1.4.2. The biological function of carotenoids

Carotenoid colouration is an important biological signal in birds that can be perceived, identified and interpreted by other individuals within and between species (Blount and McGraw, 2008). In the Gouldian finch, carotenoid colouration is mainly used to signal to conspecifics. There are several signal functions of carotenoid colouration: (1) Sexual signalling. Colours, carotenoid-based in particular, are among many mating-choice related traits, from song, dances, pheromones to body size and symmetry (Hill, 1999). (2) Social status signalling, including intra-sexual competition and individual recognition. (3) Parent–offspring signalling. For example, the reddening mouths of nestling barn swallows (*Hirundo rustica*) are caused by lutein from dietary supplementation (Saino *et al.,* 2000). (4) Other conspecific signal functions. All the above research enhances our understanding of the physiological control of avian pigmentation.

1.4.3. Carotenoid related functions, synthesis, and genetics in birds

Brush and Power (1976) first reported the carotenoids responsible for colourful plumage in the house finch and pointed out that the different constituents of carotenoids in feathers resulted in plumage colour variation. Subsequent studies by Hill (1991) and other researchers (Hill *et al.,* 1994) found that plumage redness in the male house finch is strongly correlated with female mate choice, with overwinter survival, nutritional condition and the extent to which males provision females during incubation. The distribution of carotenoid pigments in plumage shows genetic differences among subspecies (Inouye *et al.*, 2001). Moreover, the type and amount of carotenoid pigment may also have differ between subspecies. Age significantly affects the plumage colour, which means hatch-year males are generally less colourful than adults (Gill and Lanyon, 1965; Hill, 1992; Hill, 1993) and, due to the same carotenoid composition in nestlings and adults, suggests that the same plumage colouration may result from different combination of carotenoids. In addition, male house finches must maintain complicated physiological systems for carotenoid absorption, transport and

deposition in order to convert the carotenoids in their diet into red colouration in their feathers (Allen, 1987; Erdman *et al.,* 1993).

1.5. Genetics of carotenoid colouration

Colour variation due to carotenoids provides some of the best examples of condition-dependent sexual selection (Hill, 1991; Blount *et al.,* 2003). A key reason why these examples are so noticeable is that birds are unable to manufacture carotenoids *in vivo* themselves, but have to obtain them from their diet before depositing them in feathers, sometimes following metabolic conversion (McGraw and Parker, 2006). In house finches (*Haemorhous mexicanus*), the male plumage colour ranges from pale yellow to dark red due to differences in carotenoid concentration and composition. Females prefer to pair with males exhibiting the most red (with rich carotenoids). This trait benefits the female's reproductive success by indicating which males can provide her with a territory that is apparently rich in resources, and possibly also by providing her offspring with the best genes for displaying the carotenoids that are available (Hill, 1991). As in other examples, though the sexual selection preference of females is well tested, the mechanistic basis underlying the selected trait and the actual reasons for the female preference are not.

As mentioned above, there are a variety of factors that limit carotenoid deposition into the integument in animals. Genetic mechanisms are one of the factors in carotenoid colouration. Birds preferentially accumulate certain type of carotenoids in different parts of the integument. The physiological control of carotenoid colouration, such that dietary carotenoids can be converted into more colourful derived carotenoids by specific enzymes, suggests there are proteins with carotenoid-binding and transport functions, and corresponding genes encoding them (Brush, 1990). The lack of knowledge of the genetic mechanisms of carotenoid colour variation limits our understanding of the trade-offs related

to carotenoid signalling and the molecular mechanisms underlying sexual selection.

A study on the hybrid origins of domestic chicken by Eriksson *et al.* (2008) demonstrated that a carotenoid-related gene – *Yellow skin*, which is controlled by a recessive allele (W*Y) in domestic chicken – is caused by inhibiting the expression of *BCDO2* in skin. *BCDO2* is expected to have a critical function in the metabolism of Vitamin A in vertebrates (Kiefer *et al.,* 2001). This is the only gene that has been identified to be directly involved in the deposition of carotenoids into the avian integument. The yellow-skinned mutant (yellow, due to loss of function in *BCO2*, formerly *BCDO2*) deposits carotenoids obtained in the diet into its skin.

Walsh *et al.* (2011) reported 11 candidate genes related to carotenoid colouration in the red-billed quelea (*Quelea quelea*). Ten of the 11 candidates were expressed in both developing feathers and bill tissue. However, there was no evidence for an association between candidate-gene expression level and carotenoid colour variation of the male morph and sexually dimorphic bill in queleas. Thus, the mechanism of the yellow/white skin polymorphism in chicken and the carotenoid breast feather colour polymorphism is controlled by different pathways. No locus involved in feather deposition has been identified to date, and no locus has been identified to be involved in the metabolic conversion of dietary yellow to derived red ketocarotenoids.

1.6. Study species: the Gouldian finch

1.6.1. Distribution, history, threat, and recovery

The Gouldian finch, *Erythrura gouldiae* was one of the most widespread finches in northern Australia from the Kimberley region to Queensland. The local declines started to occur by

the 1940s. It has declined dramatically since the 1970s and it is now absent from many former sites (Franklin, 1999; Goodfellow, 2001). The adult Gouldian finch population size was reported to be about 2,500 or less at the end of the last century (Dostine, 1998; Garnett and Crowley, 2000). Even though the actual population size may be greater than estimated due to the huge area that might potentially contain suitable habitat in western and northern Australia (O'Malley, 2006), the exact trend in Gouldian finch population size is unclear. The decline of Gouldian finches is a combined consequence of the avicultural trade, mortality rate, resource competition with other species, parasites, and habitat degradation (Bell, 1996; Tidemann, 1996; Franklin *et al.,* 1999; Franklin *et al.,* 2005). Actions to reverse the decline of the Gouldian finch have been developed and implemented gradually since the late 1990s (Dostine, 1998)

The Gouldian finch is generally considered to be a resident, rather than migratory, species but birds do often move over significant distances. They rely on preferred grass seeds as their main food resource, and also feed on insects as supplements (Immelmann, 1982; Dostine *et al.*, 2001; Dostine and Franklin, 2002). The mean length and weight of an adult Gouldian finch are 12-15 cm and 15-30 g, respectively (Higgins *et al.,* 2006). The generation time of the Gouldian finch is about one year and the average life span of birds that reach adulthood is five or more years in the wild. The adult plumage is predominantly green on the dorsal surface and yellow on the ventral surface, with a blue upper tail, purple breast, cream under the tail, and a bright blue-bordered band behind the large face mask, which is red, yellow or black, depending on the morph (Figure 1.1). The body colour is similar in both sexes; however, the female is paler and has a shorter tail.

Figure 1.1. Three different head colour morphs in the Gouldian finch (*Erythrura gouldiae*). Phenotypes from left to right: red, yellow and black (Photograph by Sarah R. Pryke.)

1.6.2. Gouldian finch plumage colour polymorphism

The Gouldian finch is one of the small numbers of species that has a plumage polymorphism in the wild. The polymorphism at the *Yellow* locus in the Gouldian finch involves differences in carotenoid composition in both the plumage and beak (red-tipped beak or yellow-tipped beak, Table 2.1). The causative locus is probably involved in the conversion of dietary yellow to derived red ketocarotenoids (McGraw and Schuetz, 2004).

The red allele at the autosomal *Yellow* locus is dominant to the yellow one, and this locus is epistatic to the sex-linked *Red* locus, which controls the recessive black morph (Pryke and Griffith, 2006). The different colour morphs are associated with specific behaviours that affect their life history. Pryke and Griffith (2007) indicated that both male and female red Gouldian finches are much more aggressive than black and yellow birds (King *et al.*, 2015); meanwhile, red ones provide less help to their mates to raise their chicks. Both male and female birds with red heads are shorter-lived than individuals of the other

colour morphs (Pryke and Griffith, 2009b). Black-headed Gouldian finches, comprising approximately 70 per cent of the wild population, are less aggressive and have a longer lifespan. The yellow morphs, which are very rare in the wild – less than one per thousand – avoid all aggressive interactions with both red and black morphs.

As mentioned above, the black-headed type is the most common phenotype in wild Gouldian finch populations (more than 70%), then red-headed type; the rarest, yellow type, which is recessive to both red and black, is controlled by a gene on an autosomal chromosome (Murray, 1963, cited in Brush and Seifried, 1968). Brush and Seifried (1968) also concluded that there was an epistatic interaction between the red and yellow loci, but under different and independent genetic control. Bennett (1966) reported that the black and red phenotypes are controlled by a pair of alleles at a sex-linked locus, and that the red phase expressed carotenoid pigment and was dominant.

1.7. Outline of this thesis

Molecular ecology has provided novel insights into natural populations by using molecular genetic approaches to measure relationships among individuals, populations and species. Recently, it has become feasible to achieve the next level study of the evolution in natural populations – the molecular genetic basis of adaptation. Many of the previous studies in this area have focused on the genetic basis of morphological variation (Hubbard *et al.*, 2010). An animal's external morphology is especially susceptible to both natural and sexual selection. A better understanding of morphological variation is expected by to follow from elucidating the underlying genetic mechanisms of the specific trait. Birds have been ecological model species for studies of both natural and sexual selection. Variation in plumage colouration has in particular been linked to sexual selection and reproductive behaviour. So far, only a few genes have been identified to account for the plumage colour variation seen in nature.

However, the genetic basis of the examples of variation that have most convincingly been associated with components of fitness (Sheldon *et al*., 1997) has not been identified.

This thesis principally examines the molecular genetic basis of the plumage colour variation in the Gouldian finch that is associated with carotenoid pigmentation patterning (Red/Yellow plumage) of the head colour morphs. The main purpose is to demonstrate the genetic mechanism underlying plumage polymorphism in Gouldian finches. This polymorphic system is known to be under simple Mendelian genetic control, and is probably under selection to enable adaptive pre-copulatory mate choice of a compatible partner in Gouldian finches (Pryke, 2007; Pryke and Griffith, 2009a). Although this trait is not, therefore, itself a condition-dependent indicator, it is probably closely related to fitness. The relevant pathways are likely to be conserved among species, and once they have been identified it will become feasible to investigate the mechanistic links between genetics, environment and the signal in relevant species.

In Chapter 2, a preliminary genetic map of Gouldian finch is constructed by genotyping pedigreed captive birds with microsatellite markers, and aims primarily to confirm the genetic associations among different species using pedigree analysis. With the benefit of high-throughput genomic technologies, Chapter 3 mainly uses RAD sequencing and long-PCR sequencing to attempt to identify the carotenoid locus that is associated with the Red/Yellow colour trait. In Chapter 4, the candidate gene approach is performed using long-PCR and sequencing to attempt to locate the causative region, which is also expected to be coincident with the potential markers detected in the GWAS analysis. Combining all this knowledge with available genome information increases the confidence of confirming the likely candidate locus. Chapter 5 demonstrates the genetic control of the Black/Red head in the Gouldian finch. The initial study to locate the causal gene was reported by Kim (2011).

To conclude, Chapter 6 provides a general discussion that describes the broader potential of this research and possibilities for further research on Gouldian finch plumage genetics. After identifying the causative loci, a follow-up study will be able to confirm gene functions, test for convergent genetic mechanisms of adaptation in colouration among multiple passerine species, and use molecular genetic analyses to test for evolutionary signals of selection at the identified locus. This research should also enhance our currently limited understanding of plumage colour evolution in birds.

CHAPTER TWO: Linkage mapping on the Gouldian finch autosomal chromosomes

Abstract

A microsatellite-based linkage map for an avian species with an interesting natural plumage polymorphism, the Gouldian finch *(Erythrura gouldiae)*, was constructed using 33 microsatellite markers derived from the zebra finch (*Taeniopygia guttata*). The pedigree included 672 individuals from 80 captive families. All the linkage groups were homologous to zebra finch. The linkage map consisted of six linkage groups covering a total distance of 344.4 cM, with an average marker spacing of 28.7 cM between loci. The length of each linkage group varied from 3.1 cM to 178.3 cM. This map length will be an underestimate, as many genotyped segments remained unassigned at a LOD threshold of 1.0. Six linkage groups (Ego2, Ego5, Ego6, Ego7, Ego10, and Ego11) have been assigned to six chromosomes, respectively. When comparing the homologues of these microsatellites between zebra finch and Gouldian finch, no inter- or intra-chromosomal rearrangement was identified. In other words, all of the positions of markers included in this map appeared to be conserved, which is consistent with the high degree of synteny that has been observed among other bird species.

This initial linkage map for Gouldian finch, and the microsatellite markers developed in this study, will serve as a genomic resource for the Gouldian finch as well as for other passerine species. This map is also the first step towards the mapping of traits associated with phenotype and behaviour (e.g. plumage colour polymorphism), and for improving the breeding and recovery programmes of this endangered finch.

2.1. Introduction

Passerine birds, which comprise more than half of all bird species, are important model species in behavioural ecological and evolutionary research, especially in studies of natural and sexual selection (Richman and Price, 1992; Norris, 1993), mate choice (Komdeur, 1992), hybridization (Veen *et al.,* 2001), speciation (Irwin *et al.,* 2001), and individual trade-offs (Gustafsson and Sutherland, 1988). There are over 5000 identified species in the Passeriformes. However, genetic studies have been showed by a lack of genomic information on these diverse species.

Until the late 20th Century, the taxonomy of passerines mostly relied on morphological similarity. So far, the only genomic resource for passerine birds was from the zebra finch (Warren *et al.,* 2010). Genetic (linkage) maps are only available for a few species, such as the great reed warbler (Hansson *et al.,* 2005), zebra finch (Stapley *et al.,* 2008), collared flycatcher (Backström *et al.,* 2008) and Siberian jay (Jaari *et al.*, 2009). Therefore, there is an urgent need to improve our knowledge of the genetics of passerines in order to improve the understanding of passerine evolution and diversity, the relationships among passerine families, and the taxonomic methods constrained by the lack of genomic resources.

In avian species, the rate of rearrangement of chromosomes has been shown to be very low (Burt *et al.,* 1999). Chromosome painting and *in situ* hybridization probes have also provided evidence of a remarkably stable karyotype (with slight interchromosomal rearrangements) in birds (Guttenbach *et al.,* 2003; Shibusawa *et al.,* 2004; Fillon *et al.,* 2007; Griffin *et al.,* 2007). Moreover, in both cytogenetic and linkage mapping studies (Shetty *et al.,* 1999; Derjusheva *et al.,* 2004; Hansson *et al.,* 2005; Backström *et al.,* 2006; Dawson *et al.,* 2007; Hansson *et al.,* 2010), birds show highly conserved synteny. Thus, chicken (*Gallus gallus*), which provided the first avian genome sequence (International Chicken Genome

Sequencing Consortium, 2004), although belonging to the Galliformes, has also been used as a model species for genetic and genomic studies in passerines (Backström *et al.,* 2008; Stapley *et al.,* 2008; Jaari *et al.*, 2009). There are only two exceptions between chicken and passerine karyotypic homology: the groups Gga1 and Gga4 in chicken are each split into two linkage groups in passerine species (Derjusheva *et al.,* 2004).

A genetic linkage map is an important tool for understanding the genetic basis of the association between genotype and phenotype, for QTL mapping, for genome sequence assembly, and for many genetic applications such as identifying DNA markers for DNA profiling, map-based gene cloning, and so on. Linkage mapping has great significance of providing evidence for the locus detected by other methods (RAD sequencing, candidate gene analysis, etc.), which locus is indeed co-segregating with the target trait. The linkage mapping approach also allows candidate genes to be identified even if they are not detected by other methods. This method has been frequently used in evolutionary research in recent years.

Microsatellites (also referred to as SSRs or simple sequence repeats) are versatile, co-dominant, highly polymorphic and locus-specific molecular markers. Genetic linkage maps that are mainly based on microsatellite markers have been successfully constructed for several avian species (Groenen *et al.*, 1998; Kayang *et al.*, 2004; Backström *et al.*, 2008; Jaari *et al.*, 2009; Hansson *et al.,* 2010).

The Gouldian finch (*Erythrura gouldiae*) is now endangered in its native range in Australia (Woinarski *et al.*, 1992; Tidemann *et al.*, 1999; O'Malley *et al.*, 2006; Brazill-Boast *et al.*, 2011); meanwhile, due to its striking colouration, is a popular cage bird internationally. Rare among birds, the Gouldian finch possesses three distinct alternative head colour morphs: red-headed, black-headed and yellow-headed. Previous research

demonstrated that the three morphs co-existed at relatively stable ratios for more than 100 years after skins were first collected (Franklin and Dostine, 2000). The black-headed morph, which is the most common morph, is almost three times (70-80%) as frequent as the red form (20-30%) in the wild (Gilby *et al.,* 2009); given that black is recessive, this means that the black allele is much more common than the red. The yellow-headed birds are very rare in the wild, with a low but constant frequency for almost one and half centuries, at less than 0.1% of the population (Gilby *et al.,* 2009).

These three morphs have corresponding associated behavioural charactistics, pigment colouration pathways and feather structure (Brush and Seifried, 1968; Griffith and Pryke, 2006; Pryke, 2007). The Gouldian finch cheek feathers contain two large categories of pigments: melanins (eumelanin and phaeomelanin) and carotenoids (lutein epoxide and canthaxanthin) (Brush and Seifried, 1968). Eumelanin is present in black-coloured feathers, while canthaxanthin is present in red-headed and lutein in yellow-headed birds. Phaeomelanin which reflects brown-red colour, is distributed in all three morphs. The three morphs have been shown to be controlled by two genes with recessive epistatic interaction: a sex-linked (Z-linked) genetic polymorphism at the *Red* locus and an autosomal genetic polymorphism at the *Yellow* locus (Southern, 1946; Murray, 1963). The red allele at the *Red* locus is dominant, and the black allele recessive. *Yellow* is a modifier of the red morph, making red-headed birds yellow, with yellow being recessive. Black-headed adults only reveal their *Yellow* phenotype in their bill, which can be tipped with red or yellow. The head colour inheritance in the Gouldian finch is therefore an example of genetic interaction – recessive epistasis – between two genes. The Gouldian finch phenotypes and corresponding genotypes are summarized in Table 2.1.

Table 2.1. Gouldian finch head colour morphs and genotypes.

Sex	Head colour	Beak-tip colour	Genotype
Males (ZZ)	Red	Red	Z^RZ^RYY
	Red	Red	Z^RZ^rYY
	Red	Red	Z^RZ^RYy
	Red	Red	Z^RZ^rYy
	Black	Red	Z^rZ^rYY
	Black	Red	Z^rZ^rYy
	Yellow	Yellow	Z^RZ^Ryy
	Yellow	Yellow	Z^RZ^ryy
	Black	Yellow	Z^rZ^ryy
Females (ZW)	Red	Red	Z^RWYY
	Red	Red	Z^RWYy
	Black	Red	Z^rWYY
	Black	Red	Z^rWYy
	Yellow	Yellow	Z^RWyy
	Black	Yellow	Z^rWyy

Nothing is exactly known about the evolutionary origins of the red, black and yellow-headed morphs. It is possible that they arose in geographically separate populations; subsequently, they would have merged together into one single population and coexisted, not having developed a strong reproductive barrier (Pryke and Griffith, 2006, 2007; Pryke, 2010). This hypothesis might help to explain the significant behavioural and physiological differences that have been detected among the different finch morphs (Brush and Seifried, 1968; Pryke and Griffith, 2007). These other differences must either be pleiotropic effects of the colour-determining loci, or controlled by additional genes that are in linkage disequilibria with colour.

All in all, the development of genomic resources for Gouldian finch is an essential step towards understanding the behavioural and physiological differences among the three subspecies. Thus far, there are almost no genetic and genomic resources for the Gouldian finch. The main impediment to constructing the linkage map is the insufficient number of genetic markers. Only 35 genetic markers, principally for the Gouldian finch Z chromosome, were until recently available (Kim, 2011). Therefore, in this chapter, I developed and tested 59 autosomal microsatellite markers in the Gouldian finch, of which 42 were polymorphic. I used 33 of these markers to construct the first partial genomic linkage map in 672 captive birds, which included all three different morphs, belonging to 80 families.

2.2. Materials and methods

2.2.1. Mapping population and genomic DNA extraction

In total, 672 pedigreed Gouldian finch blood samples were collected from an Australian captive population, which was established using wild birds from across Australia (Pryke and

Griffith, 2006). All the samples were stored in 100% ethanol and kept at room temperature. Genomic DNA samples were extracted from blood using the ammonium-acetate precipitation method (Bruford *et al.*, 1998). The quality and quantity of DNA were assessed on 0.8% agarose gels and NanoDrop ND8000, respectively. The concentration of the originally extracted DNA was between 100–2200 ng/μl, except for 6 samples of very low concentration. The stocks were diluted gradually, and the normalized concentration of extracted DNA was ~25 ng/μl.

2.2.2. Microsatellite marker development

Microsatellite markers have been developed in previous research (Kim, 2011) using traditional cloning followed by Sanger or Roche 454 sequencing. To select microsatellite markers for primer design, SciRoKo33 (Kofler *et al.,* 2007) was used to assess microsatellite quality for microsatellite with 2- to 6-base-pair repeat units (normally >10 repeats) and relatively homogeneous repeats. To ensure an appropriate distance between markers on each chromosome, all the Gouldian finch microsatellite markers were selected and assigned to chromosomal locations in the zebra finch *genome in silico*, based on high sequence similarity between the two species (Dawson *et al.,* 2006, 2007). The location of cloned markers in the zebra finch genome was identified via standalone BLAST (http://www.ncbi.nlm.nih.gov/blast/download.shtml) with an e value $< 1 \times e^{-10}$ as the filter threshold. Fifty-nine of 581 autosomal chromosome microsatellite markers from enriched libraries were selected, distributed across 23 autosomal chromosomes. A preliminary microsatellite map chart was drawn using MapChart 2.2 (Voorrips, 2002).

2.2.3. Marker amplification and polymorphism

Fifty-nine microsatellite markers were first tested for polymorphism in 8–16 individuals.

Primers were designed based on zebra finch sequence alignments using online Primer 3.0 (Rozen and Skaletsky, 2000) with default parameters. PCR reactions were completed using a DNA Engine Tetrad PTC-225 Peltier thermal cycler in a 2-µl Qiagen multiplex system consisting of 2 µl template DNA at ~10 ng/µl and dried by evaporation at room temperature for 30 minutes, 1 µl of Qiagen Mastermix (Qiagen Inc., Kenta *et al.,* 2008), 1 µl of primer mix (all primers are at 0.2 µM with fluorescent label) with a drop of oil to prevent evaporation. The PCR program used for all the successful amplification was: 15 min of initial denaturation at 95°C, then 30 cycles of 30 s at 94°C, 90 s at 60°C and 60 s with annealing temperature at 72°C, followed by a 30-min extension phase at 60°C. The PCR reaction was terminated with a final minute at 15°C. PCR amplification was checked by electrophoresing PCR product on a 1.5% agarose gel stained with SYBR safe and photographed using GeneSnap. Four DNA samples were used to test amplification; 8–16 samples were used to test polymorphism. After testing and analysing all 59 microsatellite loci, 47 markers were found to amplify, 42 of which had relatively high polymorphism (number of alleles ≥ 4, PIC ≥ 0.387). Sixteen markers which were failed to amplify or monomorphic/ low-polymorphic were excluded from map construction.

2.2.4. Multiplexing microsatellite markers

To maximise efficiency, seven to nine markers were pooled into each multiplex set. Five multiplex sets for linkage mapping including 40 single markers were created using Multiplex Manager (Holleley and Geerts, 2009). The allele size range of each marker with 5-HEX, 6-FAM or NED fluorescent labels (Operon) of the 5’ ends of the forward primers is shown in Table 2.2 and Table 2.3. All five multiplex microsatellite sets were found to amplify well.

Table 2.2. Characterisation and predicted genomic locations of 33 microsatellite loci of Gouldian finch. All the 33 characterised polymorphic loci were assigned to the zebra finch genome and included in the construction of the Gouldian finch linkage maps and further analysis. ID: name of locus. Clone: clone name. Chr: chromosome. Chr Loc: the location of each microsatellite sequence was assigned to the zebra finch as the high sequence similarity (December 2011, ENSEMBL Release 65; Dawson *et al.,* 2006, 2007). Repeat motif: the motif and repeat number of each motif (subscripted). Fluoro: fluorescent label information. P: primers concentration (µM). S: multiplex set. ES: expected allele size (bp). OS: observed allele size range (bp). N: number of individuals amplified. A: number of alleles observed. PIC: polymorphic information content. H_O: observed heterozygosity. H_E: expected heterozygosity.

ID	Clone	Chr	Chr Loc	Repeat motif	E-value	Fluoro	Primers sequence (5'-3')	P	S	ES	OS	N	A	PIC	Ho	He
EG015	GFL 52G04	1	6828225	$(TG)_{25}$	3.00E-116	[6-FAM]	F: TGACAGTCCAGTACTTCAAACCA R: TCGATTCCCTTTGTGAACCT	0.8	1	210	181-212	8	7	0.806	0.71	0.89
EG016	GFL 51B10	1A	55925730	$(AATCTA)_{12}$	0.00E+00	[HEX]	F: GTGCCTGCTTGTCCTTTCTC R: CATGCAGGTAGCGTGGTTC	0.05	3	231	230-278	8	8	0.736	0.75	0.81
EG017	GFL 58A01	2	27572704	$(ATCT)_{27}$	4.00E-126	[HEX]	F: TCAGCAGTATGCTCCACACAG R: TTGCAACACCTTCCTTCCTC	0.2	5	283	216-246	8	8	0.843	0.75	0.92
EG018	GFL 50H10	2	38238559	$(TAGA)_{22}$	2.00E-143	[HEX]	F: AAGGCAACAGTGAATGACCAG R: CACTTGATTCATGGGTGGTG	0.1	1	321	257-320	11	8	0.889	1.00	0.96
EG019	GFL 53F10	2	84249843	$(GT)_{21}$	0.00E+00	[HEX]	F: CTCAACTGCTGCTCTCTATTGC R: TTTGTCATACTGTAGCTGGGTCTC	0.4	3	336	320-344	7	8	0.815	0.88	0.89
EG020	GFL 50E06	2	112371989	$(TCTA)_{15}$	0.00E+00	[NED]	F: GCTAGATGTGCCACTGGTTG R: TCTAATTGGAGATGACAGGATAATATG	0.1	4	260	217-279	10	8	0.845	0.88	0.92
EG021	GFL 57C07	2	150206821	$(GT)_{24}$	5.00E-155	[6-FAM]	F: AGCAATCGTGGGATAACCTG R: CTTTCCTCCCACCAAGGAG	0.4	1	311	293-326	10	8	0.871	0.88	0.94
EG022	GFL 56B02	3	45845379	$(TAGA)_{16}$	2.00E-153	[6-FAM]	F: CAAAGGGTGAACAGGAAACTG R: CTTGTTGACCTCTGGCTTCC	0.2	3	304	288-358	13	8	0.899	0.75	0.97
EG023	GFL 60B06	4	11251339	$(TATC)_{18}$	1.00E-151	[6-FAM]	F: ACAGCAGCTGACTTTCTCCAG R: TGCCTTGGAATGTGTAAGAGG	0.2	4	283	266-295	8	8	0.799	0.88	0.88
EG024	GFL 51G12	4A	1237082	$(TG)_{32}$	1.00E-55	[HEX]	F: GATCAAAGGGAAAGACATGGAC R: GATCTGTACTTATGTTTCAATACATTTTCC	0.1	2	268	231-251	5	7	0.689	0.43	0.79
EG025	GFL 52F10	5	5291621	$(GT)_{19}$	4.00E-75	[6-FAM]	F: CAGGTCTCAGGGAGCCTTTAC R: ATCTGTGGATTTGGGCATTC	1.0	5	227	214-230	6	7	0.731	1.00	0.82
EG026	GFL 56G12	5	17210892	$(TG)_{20}$	0.00E+00	[HEX]	F: TCTGCAGGGACAGCAATAAC R: CATTGGCAAACAAATTCAGC	0.4	3	188	176-203	9	8	0.862	0.88	0.93
EG027	GFL 51D03	5	32440401	$(TG)_{19}$	2.00E-90	[6-FAM]	F: GCCATCAGTTCTGGCTTTG R: CCATGTCACAGGTCCCATC	0.6	3	182	174-208	11	8	0.864	0.75	0.93

EG028	GFL 57F07	5	51035472	$(GT)_{23}$	1.00E-134	[6-FAM]	F: CTGGCTCACAGAAAAGTCCTG R: ATAATCCCCTTCCCAACTGC	0.2	4	233	206-235	7	8	0.755	0.63	0.83
EG029	GFL 57A03	6	1358746	$(GT)_{24}$	0.00E+00	[HEX]	F: GATCATTTGGCTCACAATCC R: CCAATAATTTGATGCTGAACG	1.0	5	184	166-183	6	8	0.785	0.88	0.87
EG030	GFL 61D01	6	20602611	$(GT)_{10}$	0.00E+00	[6-FAM]	F: GAAATGTGAGAAGGCAGTTGG R: TGAGAAATACTGCAGCATGAAAC	0.2	3	221	219-232	5	8	0.708	0.75	0.80
EG031	GFL 61A01	6	27381241	$(TG)_{22}$	7.00E-169	[HEX]	F: TTTATAATGCTCCCATTATGTGTATTC R: TGACACCAGTGTTGCAGATG	0.2	1	185	155-185	6	8	0.702	0.88	0.78
EG032	GFL 57G11	7	9804927	$(ATAG)_{18}$	0.00E+00	[HEX]	F: GAAGACATCTGCAGAGATGGTG R: CAGTTGTCTGTTTGCTGGAATC	0.2	5	342	333-376	5	8	0.679	0.88	0.78
EG033	GFL 58B12	7	24615928	$(AC)_{24}$	5.00E-125	[6-FAM]	F: ATGGTCCTAAGAGCCCAAGG R: TTGCCCAGCTTGATTGTAGAG	0.8	5	281	254-281	6	7	0.605	0.57	0.68
EG034	GFL 54F10	8	3584502	$(TG)_{12}$	5.00E-109	[HEX]	F: TTTCCTCATCACATTGATTGC R: TTGCTTCAGTTTGGAGTTGC	0.4	4	175	175-187	5	8	0.737	1.00	0.83
EG035	GFL 51B11	10	1312335	$(AC)_{29}$	4.00E-112	[6-FAM]	F: TTTGAAGGGAGAGAATCAATGTG R: TGCAGCAAACGCTAACAAAG	0.4	2	318	278-320	9	8	0.791	0.88	0.87
EG036	GFL 53D09	10	7625348	$(GT)_{11}$	0.00E+00	[6-FAM]	F: CAGCAGCTCTCATGTACTCAGG R: TTGCAGTCAACAAGGCAAAC	0.2	2	350	338-353	8	8	0.736	0.63	0.81
EG037	GFL 56H03	10	14121335	$(GGAT)_{18}$	5.00E-54	[NED]	F: CATCCAGATTTGAGGAGGAGAG R: GAGAGGCATCATCAAACTAGGG	0.4	2	253	140-233	8	8	0.824	0.63	0.90
EG038	GFL 52G03	11	3135544	$(TA)_{8}$	6.00E-159	[HEX]	F: TCAGATAGGCTCTGTGGAGATG R: TCCAATTCATTGCCTGTTAGG	0.2	1	212	210-226	6	6	0.760	0.83	0.86
EG039	GFL 61B03	11	18029936	$(TG)_{14}$	0.00E+00	[HEX]	F: GAAAGCTTGCTCGTTGTCTG R: TAACGTGACAGGAGCAGAGG	0.2	2	292	293-321	7	8	0.770	1.00	0.85
EG040	GFL 56D02	12	916788	$(CATT)_{12}$	6.00E-131	[6-FAM]	F: AGGTTAAGGTGGCTGCACTG R: TGTCAGGTTGCATTGGTTTG	1.0	5	338	326-350	7	8	0.770	0.75	0.85

EG041	GFL 56E09	13	4571665	(GAT)$_{31}$	2.00E-114	[NED]	F: GGGTATTTCACTGGCAGAAGC R: CCTGGGATTCATCTGGAAGTC	0.2	5	255	212-269	9	8	0.853	0.88	0.93
EG042	GFL 50F06	14	13274951	(TG)$_{23}$	3.00E-106	[6-FAM]	F: GATCCTTCTGCAGGTTTGG R: TGAAAGTCGATTTCCCTTTG	0.2	2	243	217-262	10	8	0.871	1.00	0.94
EG043	GFL 51F02	18	10243236	(AGAT)$_{12}$	9.00E-127	[6-FAM]	F: CCTGGCTGAGATGAAGATTG R: AAATTGGGTGCTACATAGGAGTAG	0.1	2	335	314-336	4	8	0.387	0.25	0.44
EG044	GFL 50D01	22	1546346	(TCAA)$_{10}$	2.00E-154	[6-FAM]	F: GGAGGGAGGGCTTAGCTTT R: TGAGGCCACATTTGTTGTGT	0.2	1	347	339-351	4	8	0.387	0.25	0.44
EG045	GFL 51F05	24	3017845	(TG)$_{8}$	0.00E+00	[HEX]	F: GGGCTGGTGTCAGTAGATGG R: GCAAATTATCTCCCTGCTTCC	0.2	2	191	191-202	5	8	0.701	0.88	0.79
EG046	GFL 51C09	26	282394	(TTGA)$_{9}$	1.00E-149	[6-FAM]	F: TACAGCAATGGCTCCAAATG R: ACGCAGAAGCAGGGACTG	0.2	2	190	188-200	5	8	0.708	0.50	0.80
EG047	GFL 53G06	27	2547462	(TG)$_{14}$	1.00E-89	[HEX]	F: CCTGCTACACTTGATAATTTCCTTC R: ACTCCCATCCCTCCCATTAC	0.2	4	207	215-234	7	5	0.820	1.00	0.93

2.2.5. Marker genotyping

Two-microlitre Qiagen Multiplex PCR reactions were used to test the multiplexes and generate all genotyping data. Genotyping raw data were generated on the ABI 3730 48-well capillary sequencer (Applied Biosystems) using ROX500 size standards. The data were analysed, alleles scored, genotyped and error checked using GeneMapper v.3.7 software (Applied Biosystems). CERVUS 3.0 (Kalinowski *et al.,* 2007) was used to check polymorphism, estimate null allele frequencies and calculate the observed and expected heterozygosities.

2.2.6. Three-step error checking

Due to random multiplex-PCR error, genotype scoring error, possible occasional mutations and pedigree information error, there is the potential for error in the raw genotyping data. Therefore, error checking was conducted in three steps: (1) General error checking. The most common genotyping error was a 1-bp binning mismatch, which could be caused by multiplex interactions. Missing scores due to overlapping allele size range bins in multiplexes also caused errors. Other errors, such as inherited null alleles and rare mutations could be detected and corrected manually by comparing pedigree information and genotype data. Several loci, for which null alleles could not be eliminated entirely, were excluded from further analysis. Only clearly segregated loci were included. (2) Error checking by CriMaker1.0 – a Perl program provided by K. W. Kim (2011). After minimising the general error, the GeneMapper export files were converted to CERVUS import files. Then, CriMaker1.0 was used to convert the genotype data into CRIMAP format and identify any outliers from expectation under the Mendelian inheritance of markers. During this step, the initial number of mismatches between parents and offspring was reduced from 4401 to 1639. Cases where the offspring from one family collectively had more than 4 alleles were reduced

from 729 to 290 by eliminating the markers with the highest error rate and individuals with high mismatch rates (implying that they had been misidentified). (3) Error checking in CRIMAP. Remaining random genotyping mismatches that were apparent from a single individual were corrected by reviewing raw data on GeneMapper. On the other hand, where an error was detected repeatedly in the same family, then it was likely that the family information was incorrect and the affected individuals were excluded from further analysis.

2.2.7. Construction of autosomal chromosome genetic map

After error-checking, the first Gouldian finch linkage map was constructed using CRIMAP (http://biobase.dk/Embnetut/Crimap/analyse1.html; Lander and Green, 1987; Green *et al.,* 1990), for 672 pedigreed individual DNA samples. The CRIMAP calculates two-point recombination fractions, provides logarithmic odds (LOD) scores for recombination estimates, and tests marker order. The 'twopoint' command was used to calculate the recombination fractions and LOD score between all pairs of markers. The 'build' command was used to generate the order and the interval of markers for each linkage group. Then, the most likely marker order of significantly linked loci was determined by the options 'flipsn' command and 'fixed' command. The initial threshold between linked markers was LOD = 3.0. The groups of pairwise markers with LOD > 3.0 were considered to be linked. To improve the map, markers were considered to be linked if the LOD score was higher than 1.0, where there was also *a priori* support for the localisation, such as an expectation concerning the chromosomal location and linkage based on a BLAST search in zebra finch (Slate *et al.,* 2002; Dawson *et al.,* 2006, 2007). All map distances are given in Kosambi centiMorgans (cM).

2.2.8. Genome coverage estimate

The mean marker spacing among linkage groups was calculated by dividing the total length of all linkage groups on the map by the number of intervals within these linkage groups. The coverage of this linkage map was estimated by calculating the sum of the position differences of the first and last interval in each linkage group in the zebra finch, then dividing this by the total length of the zebra finch genome (~1.2 Gb; Ensembl database http://www.ensembl.org/Taeniopygia_guttata/Info/Index)

2.3. Results

In total, 59 genomic microsatellite primer pairs were tested for polymorphism (Figure 2.1). One locus was monomorphic, four had low to moderate polymorphism ($0.11 \leq PIC \leq 0.34$, $A \leq 3$) and 12 primer pairs failed to amplify products or amplified a faint/fuzzy product (Table 2.3). A total of 42 high polymorphic microsatellite markers were identified (Table 2.2, 2.3). Forty markers were arranged to construct five multiplex microsatellite sets. For later linkage mapping and analysis, 672 DNA samples from 80 families were used. This genetic map includes 33 microsatellites distributed among six linkage groups. For each linkage group, two or more microsatellites were included to identify the genetic location and proceed to the linkage analysis. Eighteen of the 33 autosomal loci (54.5%) were linked to at least one other locus with a LOD value higher than 1.0. These linked autosomal loci were placed on six zebra finch chromosomes (Figure 2.2). This linkage map covered a total distance of 344.4 cM, with an average marker spacing of 28.7 cM between loci (Figure 2.3). The length of each linkage group varied from 3.1 cM to 178.3 cM. After building the map, linkage analysis was performed for yellow-headed informative individuals within pedigreed samples to detect the probable genomic region of the *Yellow* gene. However, there was no positive association detected with 33 specific markers ($LOD \leq 1$; Table 2.4).

Figure 2.1. Fifty-nine Gouldian finch microsatellite loci mapped onto 23 autosomal chromosomes in the zebra finch genome. Thirty-three loci distributed among 20 linkage groups were used for linkage mapping and further analysis (Underlined). An additional 10 markers showed a relatively high degree of polymorphism, but were excluded from linkage mapping and further analysis (**Bold**). Four loci were moderately polymorphic or monomorphic (*Italics*) and 12 loci failed to amplify or amplified a faint/fizzy product (Plain text). Clone name are shown in the map.

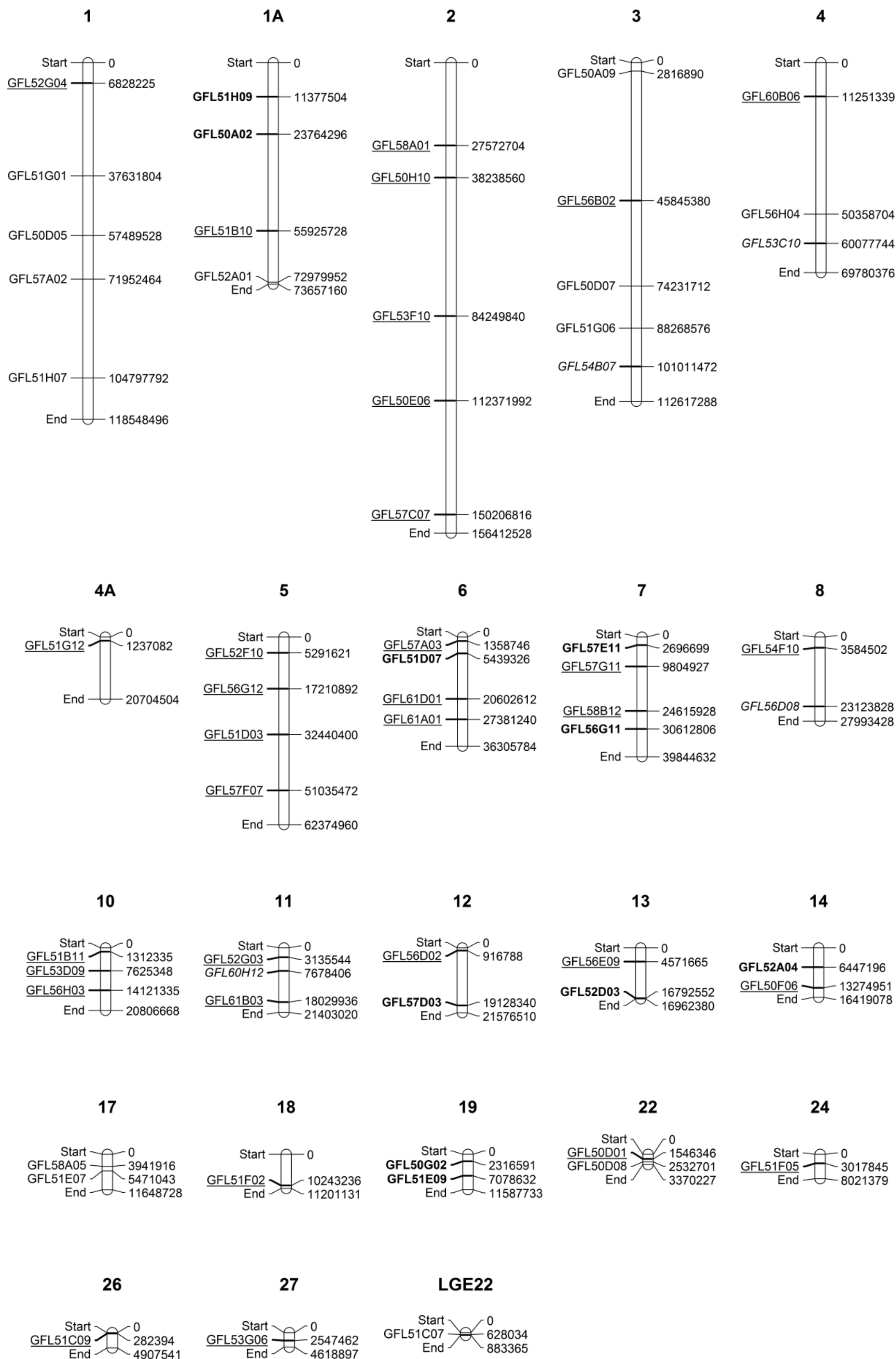

1
Start 0
GFL52G04 6828225
GFL51G01 37631804
GFL50D05 57489528
GFL57A02 71952464
GFL51H07 104797792
End 118548496
1A
Start 0
GFL51H09 11377504
GFL50A02 23764296
GFL51B10 55925728
GFL52A01 72979952
End 73657160
2
Start 0
GFL58A01 27572704
GFL50H10 38238560
GFL53F10 84249840
GFL50E06 112371992
GFL57C07 150206816
End 156412528
3
Start 0
GFL50A09 2816890
GFL56B02 45845380
GFL50D07 74231712
GFL51G06 88268576
GFL54B07 101011472
End 112617288
4
Start 0
GFL60B06 11251339
GFL56H04 50358704
GFL53C10 60077744
End 69780376
4A
Start 0
GFL51G12 1237082
End 20704504
5
Start 0
GFL52F10 5291621
GFL56G12 17210892
GFL51D03 32440400
GFL57F07 51035472
End 62374960
6
Start 0
GFL57A03 1358746
GFL51D07 5439326
GFL61D01 20602612
GFL61A01 27381240
End 36305784
7
Start 0
GFL57E11 2696699
GFL57G11 9804927
GFL58B12 24615928
GFL56G11 30612806
End 39844632
8
Start 0
GFL54F10 3584502
GFL56D08 23123828
End 27993428
10
Start 0
GFL51B11 1312335
GFL53D09 7625348
GFL56H03 14121335
End 20806668
11
Start 0
GFL52G03 3135544
GFL60H12 7678406
GFL61B03 18029936
End 21403020
12
Start 0
GFL56D02 916788
GFL57D03 19128340
End 21576510
13
Start 0
GFL56E09 4571665
GFL52D03 16792552
End 16962380
14
Start 0
GFL52A04 6447196
GFL50F06 13274951
End 16419078
17
Start 0
GFL58A05 3941916
GFL51E07 5471043
End 11648728
18
Start 0
GFL51F02 10243236
End 11201131
19
Start 0
GFL50G02 2316591
GFL51E09 7078632
End 11587733
22
Start 0
GFL50D01 1546346
GFL50D08 2532701
End 3370227
24
Start 0
GFL51F05 3017845
End 8021379
26
Start 0
GFL51C09 282394
End 4907541
27
Start 0
GFL53G06 2547462
End 4618897
LGE22
Start 0
GFL51C07 628034
End 883365

Table 2.3. Twenty-six additional microsatellites with predicted chromosome locations tested but not include in linkage mapping and further analysis. Clone: clone name. Chr: chromosome. Chr Loc: The location of each microsatellite sequence was assigned to the zebra finch as the high sequence similarity (December 2011, ENSEMBL Release 65; Dawson *et al.,* 2006, 2007). Repeat motif: the motif and repeat number of each motif (subscripted). Fluoro: fluorescent label information. P: primers concentration (μM). S: multiplex set. ES: expected allele size (bp). OS: observed allele size range (bp). A: number of alleles observed. N: number of individuals amplified. PIC: polymorphic information content. H_O: observed heterozygosity. H_E: expected heterozygosity. Status: the marker amplification and polymorphism status (High poly=high polymorphic, Moderate=low to moderate polymorphic, Mono=monomorphic, Failed= failed to amplify or amplified a faint/fuzzy product).

Clone	Chr	Chr Loc	Repeat motif	E-value	Fluoro	Primers sequence (5'-3')	P	S	ES	OS	A	N	PIC	Ho	He	Status
GFL 51D07	6	5439326	$(CTAT)_{15}$	0.00E+00	NED	F: TCAGAATTGAGGAAAACCTTAGA R: TGCATCTTGTCCATTAGCTG	0.2	1	207	202-265	8	8	0.83	0.63	0.9	High poly
GFL 57E11	7	2696699	$(GATA)_{24}$	1.00E-146	HEX	F: GATCATCAAGTCACACCTTATATATTTT R: GTTGTTTCAGTGCCAGGTCT	0.04	5	275	262-284	9	8	0.79	0.5	0.86	High poly
GFL 52D03	13	16792553	$(TTG)_{8}$	2.00E-63	NED	F: ATTGACAGGTGTGGGAGGAG R: AGCAAGAAGGGCTTCTACAGG	0.1	3	247	171-258	7	7	0.77	1	0.86	High poly
GFL 52A04	14	6447196	$(AC)_{19}$	8.00E-89	HEX	F: CAGCAGCTCTCATGTACTCAGG R: TTGCAGTCAACAAGGCAAAC	0.6	1	350	345-360	9	8	0.85	0.5	0.93	High poly
GFL 51E09	19	7078632	$(AAAG)_{5}$	3.00E-172	6-FAM	F: TGGTTGAAATGGAAAGCAGA R: CAGCTGTAAAACCCCTCTGG	0.2	1	253	249-257	6	8	0.75	0.5	0.83	High poly
GFL 51H09	1A	11377504	$(GGAA)_{15}$	3.00E-71	HEX	F: CATCCAGATTTGAGGAGGAGAG R: GAGAGGCATCATCAAACTAGGG	0.2	4	253	247-337	10	8	0.87	0.63	0.94	High poly
GFL 50A02	1A	23764297	$(AGAT)_{12}$	8.00E-163	6-FAM	F: AGGCAATGCCAGTGTATGC R: TGCGACTGGTATTTGGACAG	0.05	3	270	257-269	4	6	0.67	0	0.79	High poly
GFL 56G11	7	30612806	$(TG)_{23}$	0.00E+00	-	F: AGGGAGAGAGAGAGAAATCGAAG R: CCCTCCGTACACAAACCTACTC	0.2	-	242	232-242	5	7	0.69	0.57	0.78	High poly
GFL 50G02	19	2316591	$(CA)_{14}$	2.00E-117	-	F: ATCCAGGGAGATAAAGGAGGAG R: CGCTAGAAGGAGGGAGCTG	0.2	-	234	231-246	7	8	0.81	1	0.88	High poly
GFL 57D03	12	19128340	$(TG)_{11}$	4.00E-150	-	F: CCAGTCTGGGATTTCTGGTC R: TCCTCAGAACTGTAATTCTTCTGC	0.2	-	296	295-297	2	8	0.34	0.38	0.46	Moderate
GFL 53C10	4	60077744	$(AC)_{14}$	0.00E+00	-	F: CTCCACCTTGACAGCTGATG R: CAAGAGGCATTGGTGAGAGTC	0.2	-	225	227-230	2	8	0.2	0.25	0.23	Moderate
GFL 56D08	8	23123828	$(TAGA)_{10}$	0.00E+00	-	F: TTTGAAGGGAGAGAATCAATGTG R: TGCAGCAAACGCTAACAAAG	0.2	-	318	301-322	3	8	0.29	0.38	0.34	Moderate
GFL 60H12	11	7678406	$(AC)_{12}$	0.00E+00	-	F: AACCGTTGGGTCTGAATTTG R: CCTGAAGATGAAAGCAACCAG	0.2	-	187	179-181	2	8	0.11	0.13	0.13	Moderate

GFL 54B07	3	101011470	$(ATCT)_{18}$	4.00E-105	-	F: CAAGAGGCCACCATCACTTC R: CCCTCAGGGCAGATAAGAGTG	0.2	-	310	-	1	8	0	0	0	Mono
GFL 50D07	3	74231714	$(TG)_{18}$	2.00E-111	-	F: ATCTAAGAGGTATCTGGAATGTATCTG R: AAGTGCCAGGCCTATGAAG	0.2	-	185	-	0	0	0	0	0	Failed
GFL 50D05	1	57489527	$(TG)_{22}$	1.00E-116	-	F: CTGCATTGCAAACTCTTTCG R: GAAATTGCTACGGCAAGAGC	0.2	-	321	-	-	-	-	-	-	Failed
GFL 51G01	1	37631804	$(TCTT)_{26}$	3.00E-139	-	F: TTCAAGAACATCTAATGTACAAAATGAA R: TGAATGAGTTACTTTATTTACAGGCATT	0.2	-	343	-	-	-	-	-	-	Failed
GFL 51H07	1	104797792	$(TCTA)_{18}$	0.00E+00	-	F: TTGTTGAAGCCAAACACTGC R: AGGGTTCAGTTTCCTTCATCAG	0.2	-	343	-	-	-	-	-	-	Failed
GFL 57A02	1	71952466	$(GAAAG)_{23}$	9.00E-125	-	F: AGGGATAACTAAGAAATAAGTCACTGAAG R: GGCACATATCCTTCTATCCATTATC	0.2	-	349	-	-	-	-	-	-	Failed
GFL 50A09	3	2816890	$(AC)_{16}$	2.00E-134	-	F: GAGCCTCCTTTGAGGAAGAC R: TCCCATTTGAATATCAGCTTG	0.2	-	334	-	-	-	-	-	-	Failed
GFL 51G06	3	88268573	$(TTCCC)_{13}$	4.00E-78	-	F: TGGAATGAAAGAATGAAATAACAA R: AGGGAACTGTTTCAAAATAGAGATT	0.2	-	349	-	-	-	-	-	-	Failed
GFL 56H04	4	50358705	$(CTAT)_{29}$	0.00E+00	-	F: CTGCATGGAAAGTCTCAGAGG R: GAAACTTACCTGCCCTGGTG	0.2	-	238	-	-	-	-	-	-	Failed
GFL 58A05	17	3941916	$(GAGAA)_{32}$	2.00E-84	-	F: ACCCTGCACCAAAATCTGAC R: AGAAAATGACCGAAAAGATAGCTG	0.2	-	317	-	-	-	-	-	-	Failed
GFL 51E08	17	5471043	$(AT)_{8}$	1.00E-130	-	F: AACAGCCTCTGGATGTTTCC R: GATCCAGCATCCTATCTGAGC	0.2	-	324	-	-	-	-	-	-	Failed
GFL 52A01	1A	72979953	$(AC)_{19}$	1.00E-68	-	F: CCCTGTTCCAAGAGTGGTTTAG R: TTTCCCTCCCTTCCTCTCC	0.2	-	165	-	-	-	-	-	-	Failed
GFL 51C07	LGE22	628034	$(GGATT)_{18}$	2.00E-80	-	F: AAGGAACTGCCTGTGGAAAG R: TTTCTTTGCCTCCAAACTCC	0.2	-	294	-	-	-	-	-	-	Failed

Figure 2.2. First-generation sex-averaged linkage map of the Gouldian finch (*Erythrura gouldiae*) consisting of six linkage groups representing six chromosomes in the zebra finch. All map distances in the Gouldian finch genome, with positions, are given in Kosambi centiMorgans (cM). Loci assigned locations on zebra finch chromosomes with marker order supported by a LOD > 1.0 are presented. The yellow-headed morph locus was not found to be linked with any of these markers using 'twopoint' linkage.

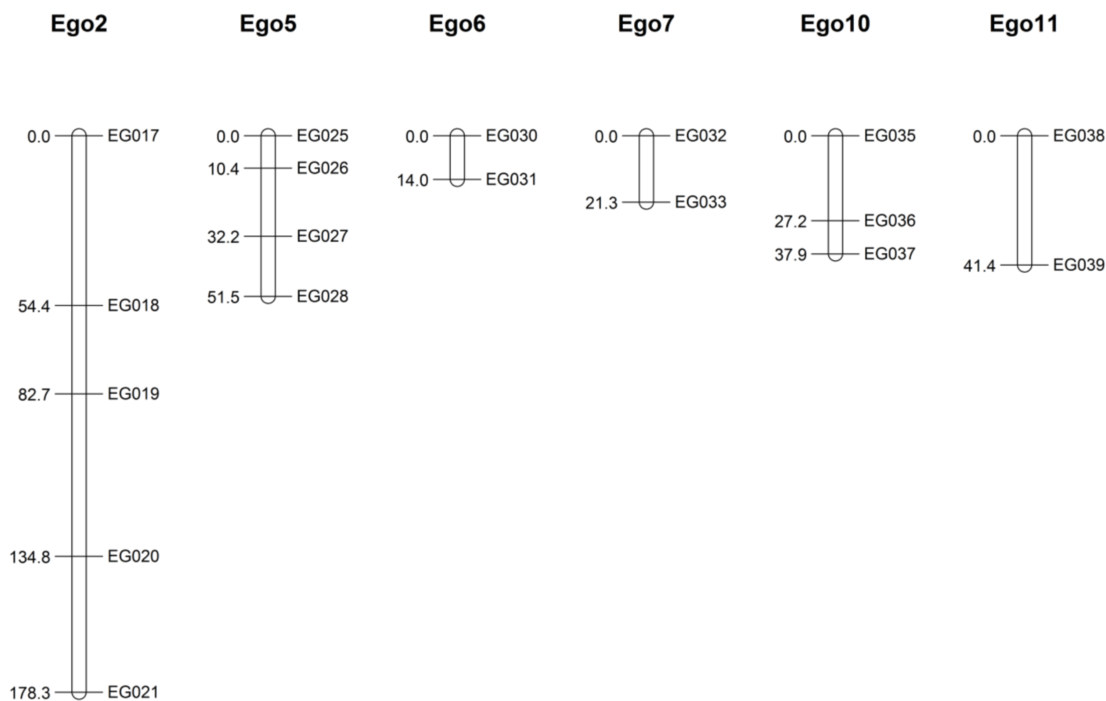


Figure 2.3. Comparison of linkage groups in Gouldian finch and zebra finch. This linkage map included 18 microsatellite loci distributed among six linkage groups, which were found to be homologous to zebra finch chromosomes. The order of markers perfectly matches the physical map. The units are cM in the Gouldian finch and Mbp in the zebra finch.

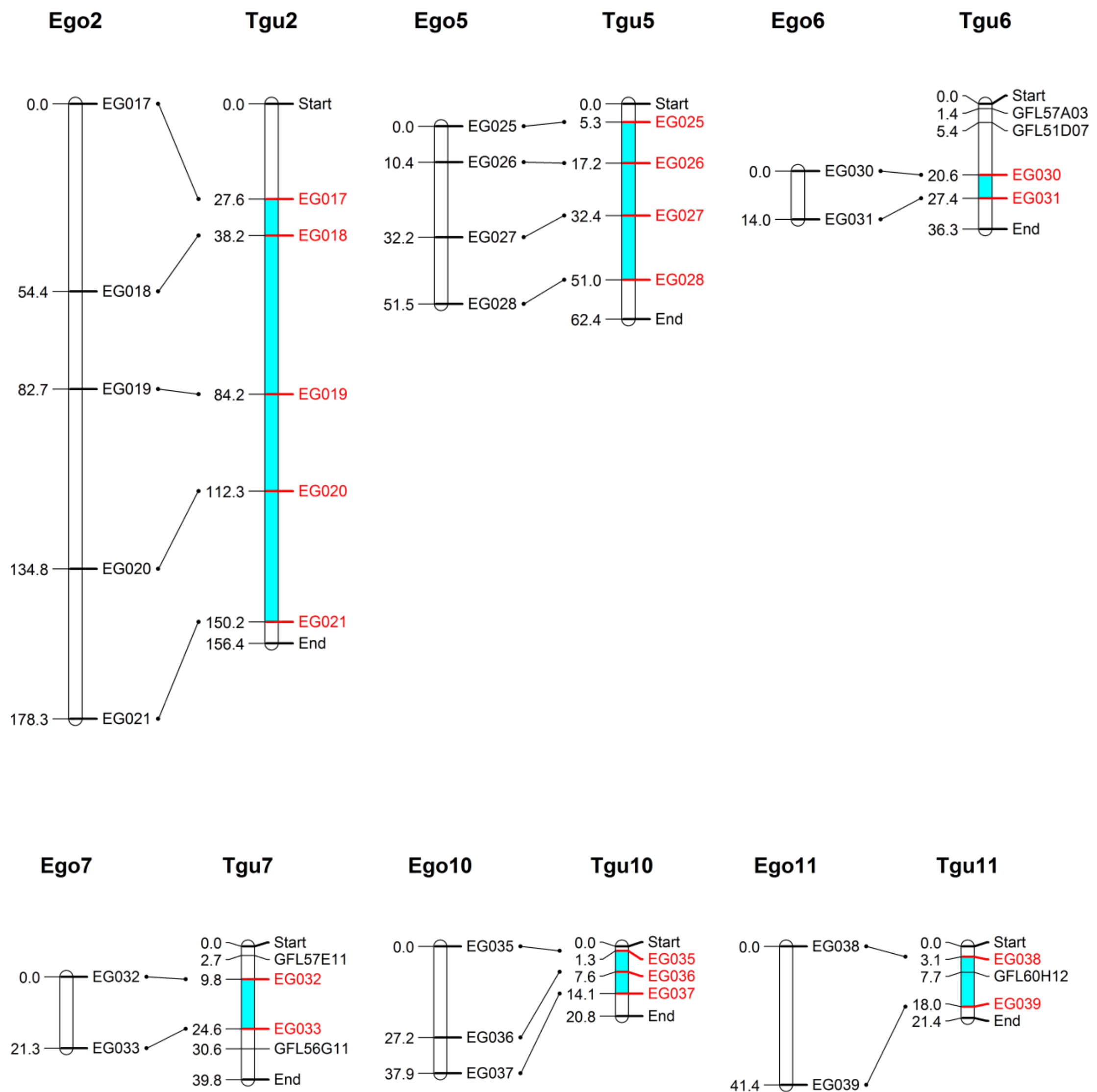


Table 2.4. Association of genomic microsatellite loci with the yellow-headed colour trait. LOD ≥ 3 is the general threshold value considered to demonstrate a specific genotype–phenotype association. Ten of nineteen markers with LOD > 0 are shown.

Marker ID	Morph	Recombination fraction	LOD
EG017	Yellow	0.13	0.72
EG036	Yellow	0	0.6
EG016	Yellow	0.2	0.54
EG039	Yellow	0.22	0.27
EG028	Yellow	0.21	0.27
EG037	Yellow	0.15	0.23
EG024	Yellow	0.15	0.23
EG038	Yellow	0.26	0.13
EG042	Yellow	0.2	0.07
EG041	Yellow	0.39	0.02

2.4. Discussion

In this chapter, 59 new Gouldian finch microsatellite markers were developed and tested based on the zebra finch genome. Of the 59 markers, 47 amplified and 42 markers were usefully polymorphic and readily scored. Forty polymorphic markers were organized into five multiplex microsatellite sets for autosomal chromosome linkage mapping. The first partial microsatellite-based sex-averaged linkage map of the Gouldian finch was constructed using 672 pedigreed captive individuals from 80 families. This linkage map included 18 microsatellites distributed among six linkage groups (Ego2, Ego5, Ego6, Ego7, Ego10, and Ego11), which are homologous to zebra finch chromosomes. The order of markers perfectly matched the physical map predicted from zebra finch, though the marker interval was relatively large and further study is required (Figure 2.3). The comparative microsatellite-based map demonstrated a high degree of synteny and marker order conservation between zebra finch and Gouldian finch. The stability between the genomes of the two species is typical of birds but also reflects the close evolutionary relationship between the two species.

The proportion of the genetic length of the Gouldian finch genome covered by the map is more difficult to estimate. The genetic coverage should be less than the physical coverage estimated above due to the absence of markers for some chromosomes and the high recombination rate typically observed towards the telomeres. To estimate the coverage of the Gouldian finch genome, the physical information from the zebra finch genome has been used. The zebra finch and the Gouldian finch share 217.6 Mb (or 18.1% of the total, as the zebra finch autosomal genome is about 1.2 Gb). More sequence could be considered to be encompassed by the Gouldian finch map, considering the exclusion of the 15 highly polymorphic but unlinked single markers and the flanking sequence associated with the end markers in each linkage group. Assuming the estimated genome size of Gouldian finch is

similar to the zebra finch, the physical map described in this chapter is expected to cover 20-25% of the Gouldian finch genome.

To detect linkage between marker loci and the phenotype at the yellow head colour locus, informative individuals within the pedigreed samples were used. However, no locus was significantly associated with the yellow-headed trait. This might be either because the *Yellow* gene was located outside of these linkage groups or because of the relatively low density of microsatellite coverage. The Gouldian finch is expected to have the typical passerine chromosome set of 2n ~ 78-80 (Shetty *et al.,* 1999), yet we only defined seven linkage groups on six chromosomes with LOD score higher than 3.0. Two linkage groups (corresponding to chromosome 2) could be integrated into one linkage group by reducing the required LOD score threshold from 3.0 to 1.0, to anchor with shared linked microsatellite loci; each of the six linkage groups then corresponded to a different chromosome. Further research will be required to expand the linkage map. Polymorphic markers can be used in comparative analysis on passerine or other avian genomes. This preliminary linkage map and future high-resolution maps will contribute to different aspects of evolutionary biology research, such as comparative genomic studies, parentage analysis, adaptive population divergence, and identifying the genetic mechanisms underlying fitness traits.

In brief, the linkage map developed here includes six linkage groups covering 344.4 cM of the Gouldian finch genome. This is the first autosomal genomic linkage map of the Gouldian finch. Even though the density of the map is relatively low, it is still potentially valuable for mapping phenotypic loci, as demonstrated in previous mapping research (Miwa *et al.,* 2006). This map is also an initial step for further mapping and is a potential resource for locating genes underlying important functional fitness traits in the Gouldian finch, as well as quantitative trait loci.

CHAPTER THREE: A genome-wide association study of head colour polymorphism in Gouldian finches (*Erythrura gouldiae*)

Abstract

The Gouldian finch (*Erythrura gouldiae*) is considered to be a species with one of the most fascinating plumage polymorphisms, with three distinctive colour morphs controlled by two separate loci. Research to identify the genetic basis of the *Red* locus has achieved definite progress in recent years (Kim, 2011); however, the genetic basis of the *Yellow* locus is still undetermined. To identify the genetic variants that might contribute to the red- / yellow-headed polymorphism controlled by this locus, we used a genome-wide association study. Sixty-three Gouldian finch samples, including both red and yellow morphs, were genotyped by RAD-sequencing, in which 63,833 single-nucleotide polymorphisms (SNPs) were identified. We attempted to identify markers with a significant difference in genotype frequency between the red and yellow morphs using the *Stacks* program. We found several candidate SNPs, and in particular, a broad peak in F_{ST} between the morphs on chromosome 8 that coincided with multiple SNPs that showed significant allele frequency differences. The region on chromosome 8 should be considered as a candidate region until further investigations are done. These RAD data represent a valuable resource in this species with a relatively unknown genome. By sequencing more Gouldian finches in further studies, combined with other analyses, a more definitive association might be identified.

3.1. Introduction

3.1.1. Restriction-site associated DNA

The development of genetic tools and technologies in the last decade, such as microarrays and next-generation sequencing (NGS), has greatly improved studies on the genetic basis of important traits (Ekblom and Galindo, 2011; Gheyas and Burt, 2013). Sequencing of the entire genome has provided the opportunity to identify abundant genetic variants.

Restriction-site associated DNA (RAD) tags are short flanking DNA sequences adjacent to restriction sites, after the DNA has been digested by the respective restriction enzyme (Miller *et al.,* 2007b; Baird *et al.,* 2008). DNA sequence polymorphisms (e.g. single nucleotide polymorphisms) can be identified by comparing the same set of RAD tags among individuals. The polymorphisms so identified are called RAD markers, and can be useful for association mapping, and in ecological genetics. The core mechanism of RAD-tag isolation includes digesting DNA with a restriction enzyme, ligating adapters to the overhangs, shearing the DNA into small fragments, and isolating the small fragments or amplifying fragments containing adapters using PCR. The method used to isolate RAD tags was initially by hybridising to microarrays (Lewis *et al.*, 2007; Miller *et al.,* 2007a), but then adapted for ultra-high-throughput DNA sequencing (Baird *et al.,* 2008; Hohenlohe *et al.*, 2010; RAD tag sequencing, RAD-sequencing, or RAD-seq). For example, the latest Illumina HiSeq 2500 system can be used to sequence entire genomes of vertebrates (i.e. gigabases) within a few days (http://www.illumina.com/systems/hiseq_2500_1500.ilmn). Whole-genome sequencing of large sample sizes is currently prohibitively expensive. RAD sequencing is an example of reduced representational sequencing, which is designed to obtain the same subset of genome sequence, usually randomly distributed, from many locations distributed across the whole genome. RAD sequencing can generate a high density of genetic markers, at a lower error

rate, and be made more useful still by using paired-end sequencing (Shendure and Ji, 2008; Davey *et al.*, 2011). Using RAD sequencing, a pooled library of tens or hundreds of samples enables the cost of sequencing to be reduced, and massively parallelised sequencing significantly increases efficiency by producing millions of short DNA fragment reads (up to 125-bp single-end and 250-bp paired-end using the HiSeq) (Metzker, 2010).

SNPs (Single Nucleotide Polymorphisms) are the most common type of DNA sequence variation between related sequences (Suh and Vijg, 2005). SNP markers, with their properties of massive quantity, stable inheritance through generations and enabling high-resolution mapping to linked traits (Thomas *et al.,* 2011), are important genetic markers in molecular genetic studies (Berger *et al.*, 2001; Stickney *et al.*, 2002). Therefore, SNP markers have been applied in many aspects of genomics (Subodh *et al.*, 2014). However, discovering high-density SNPs was, previously, costly and time-consuming using classic genotyping platforms. The rapid developments of high-throughput next-generation sequencing, which is a prerequisite for whole-genome association tests (Schena *et al.*, 1995), makes discovering massive numbers of SNPs in a shorter time for a lower cost a reality, so enabling the genome-wide analysis of polymorphisms.

3.1.2. Genome-wide association studies

Genome wide association analysis among individuals within a species is a crucial tool for understanding the genetic basis of phenotypic polymorphisms. A genome-wide association study (GWAS) is used to examine whether specific genetic variants (e.g. SNPs) between different individuals are associated with a specific trait (Klein *et al.*, 2005). The GWAS normally includes two groups with different traits or phenotypes (cases and controls). The theoretical basis is: if an allele of a SNP is statistically more frequent in the case group than in the control group, than this SNP is considered to be associated with the trait of interest

(Clarke *et al.*, 2011). This provides the location of a genomic region that may influence the trait or phenotype (Ikram *et al.*, 2010). One of the benefits of this approach is that GWAS does not rely on candidate genes or chromosomes but investigates the entire genome. This approach is statistically more powerful than linkage studies for detecting genetic effects (Risch and Merikangas, 1996). It should be noted here that an associated variant does not necessarily correspond to the causal loci (e.g. false positives) (Pearson and Manolio, 2008; Manolio *et al.*, 2010).

3.1.3 *Stacks*

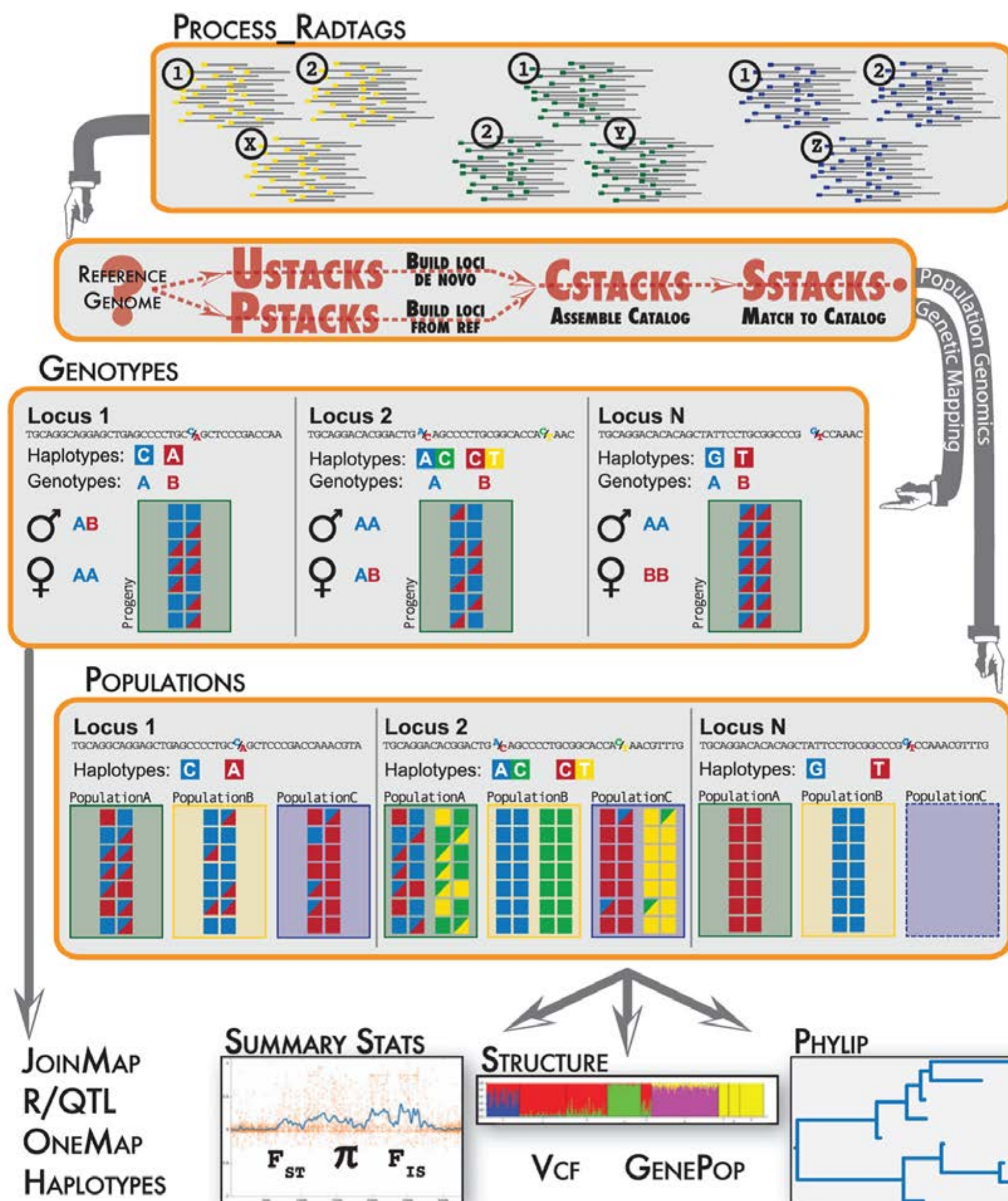
A massive number of polymorphisms can be identified by RAD-seq, which can be used to generate ultra-dense genetic linkage maps, providing a powerful genomic resource in otherwise little-known species (Amores *et al.*, 2011). A set of efficient and easy-to-use software tools is then essential for exploiting the millions of genotypic data points in genetic mapping or population genomic studies. *Stacks* is a useful software platform for the analysis of such short-read sequencing data and genotyping a large number of loci in each different individual, either with respect to a reference genome or *de novo*. This program uses a maximum-likelihood approach to distinguish sequence polymorphisms from sequencing errors (Catchen *et al.,* 2013). *Stacks* can be widely used to contribute to most aspects of genomic analysis, such as high-density linkage mapping, reference genome assembly, population genomics and phylogeography.

Stacks is a C++ and Perl based modular program operating in a Unix-like environment and relies on OpenMP libraries with parallelised algorithms. The core component programs of *Stacks* include five main parts: process_radtags.pl (demultiplexing and cleaning up Illumina raw reads), ustacks/pstacks (using cleaned FASTA/FASTQ files to identify loci and detect haplotypes), cstacks (merging loci and creating a catalogue), sstacks (matching to the

catalogue), and populations/genotypes (population genomics analysis or genetic mapping, depending on the aim of the study). The easiest way to run *Stacks* is via packaged programs: (1) using the denovo_map.pl program without a reference genome, combining ustacks, cstacks, and sstacks, calling markers with markers.pl, and indexing the database with the index_radtags.pl program; (2) using the ref_map.pl program with a reference genome, combining the pstacks, cstacks, sstacks, marker.pl and index_radtags.pl programs. The workflow of *Stacks* has been summarised by Catchen *et al.* (2013) and restated in Figure 3.1.

As mentioned in chapter one, the red- / yellow-headed colour trait in the Gouldian finch is believed to be controlled mainly by a single Mendelian locus, the *Yellow* allele, which is recessive. The objective of this chapter is to use RAD markers in GWAS to identify candidate RAD-tags which might be associated with the red- / yellow-headed polymorphism. In order to map the Gouldian finch *Yellow* locus, genome-wide association analysis was undertaken using the Illumina Hi-seq.

Figure 3.1. Flowchart for the *Stacks* program. There are five main steps in the *Stacks* pipeline. The process_radtags program demultiplexes the data, checks the average quality of the reads, and cleans up the Illumina raw data. The ustacks/pstacks program identifies unique loci in each individual from cleaned short sequencing reads. The cstacks program merges loci from different individuals and creates a catalogue. The sstacks program matches loci in the catalogue. The populations or genotypes program analyses the loci using different input files. The populations program analyses the loci within and among populations and gives an output of calculated population genetic statistics with different formats. The genotypes program maps haplotypes, records genotypes, and formats genotypes for different uses. (Catchen *et al.,* 2013)



3.2. Methods

Sixty-three individually labelled genomic DNA samples were analysed. The sample set involved yellow- and red-headed birds, comprising two captive families segregating for yellow, in addition to 11 yellow-headed captive birds from different families, and 12 putative yellow-headed wild birds from Wyndham, Australia. The genomic DNA samples were prepared using the phenol/chloroform method (Bruford *et al.*, 1998). The quality and quantity of DNA samples was assessed by gel electrophoresis and the NanoDrop spectrophotometer (Figure 3.2 and Table 3.1). The minimum quantity and appropriate concentration of the genomic DNA samples was 2 μg and 50-100 ng/μl, respectively. The concentrated original DNA samples (>100 ng/μl) were diluted to 50-100 ng/μl. All the samples were of good quality with high molecular weight, with very limited degradation and only a trace of RNA in few samples, which would not affect the sequencing.

Figure 3.2. The gel electrophoresis of 63 DNA samples of the Gouldian finch that were used in RAD-Seq (same order with sample data in Table 3.1 from first row left to second row right). The mass of the 1-kb standard is 125 ng/μl for the brightest band.

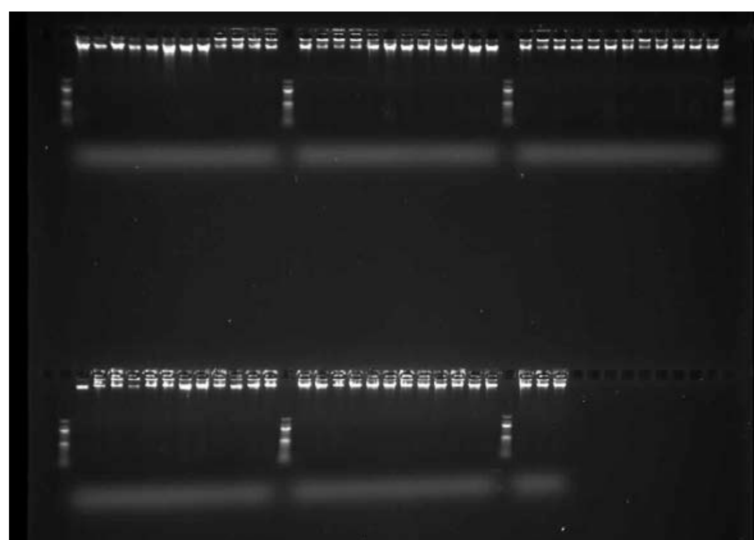


Table 3.1. The RAD-sequencing quantity and quality control, and RAD library information of Gouldian finches. The OD260:280 ratio of pure DNA is around1.8. The appropriate concentration of the genomic DNA is 50-100 ng/µl. The minimum quantity of the total genomic DNA amount is 2µg.

Sample ID	Father	Mother	Head	Beak	Library	MID	OD 260:280	Concentration(ng/µl)	Volume(µl)	Total amount(µg)	Buffer
735			Red		1	ACACTGAC	1.82	75	51	3.8	TE
1216			Red		1	ACGTAGCA	1.94	79	29	2.3	TE
902	638	322	Black	yellow	1	CACACAGT	1.85	75	44	3.3	TE
990	638	322	Black		1	CAGTCTCA	1.93	70	29	2.0	TE
991	638	322	Black		1	GTACTCGT	1.83	75	32	2.4	TE
992	638	322	Black		1	GTCATGTG	1.73	75	52	3.9	TE
1015	638	735	Yellow		1	TGCATCGT	1.86	75	37	2.8	TE
1092	638	735	Black	red	1	TGTGACTG	1.82	75	53	4.0	TE
2122	638	1216	Black		1	ACACGACA	1.66	47	44	2.1	TE
2123	638	1216			1	ACGTCTAC	1.73	36	59	2.1	TE
2124	638	1216	Black		1	CAGTGTGT	1.84	67	34	2.3	TE
2167	638	1216	Yellow		1	CATGATCA	1.94	25	84	2.1	TE
2168	638	1216	Red		1	GTACGCTG	1.82	52	39	2.0	TE
2169	638	1216	Black	yellow	1	GTCAGTGT	1.89	31	69	2.1	TE
2170	638	1216	Yellow		1	TGTGTGAC	2.05	30	69	2.0	TE
2171	638	1216	Black	yellow	1	TGTGCAGT	1.91	51	44	2.2	TE
817			Black	yellow	2	ACGTGCTG	1.83	75	51	3.8	TE
898	407	789	Black	red	2	ACTGATAC	1.86	75	51	3.8	TE
862	310	25	Red		2	CATGCGAC	1.85	75	50	3.8	TE
853	664	202	Red		2	CGATACTA	2.04	75	42	3.1	TE
957	664	202	Black		2	GTACATCA	1.83	75	51	3.9	TE

1151	817	898	Black	yellow	2	GTGTACTG	1.82	75	50	3.8	TE
1128	862	853	Red		2	TATATGCG	1.87	75	39	2.9	TE
1129	862	853	Red		2	TGCACTAC	2.01	75	37	2.8	TE
1493	862	957	Red		2	ACTGCTCA	1.83	75	41	3.1	TE
1494	862	957	Red		2	ACTGTGTG	1.84	75	34	2.5	TE
1495	862	957	Red		2	CATGTCGT	1.84	75	37	2.8	TE
1496	862	957	Yellow		2	CGATCAGC	1.82	75	30	2.2	TE
1531	862	1151	Red		2	GTCACTCA	1.85	75	36	2.7	TE
1532	862	1151	Yellow		2	GTGTCAGT	1.85	75	48	3.6	TE
1533	862	1151	Red		2	TATAGCAT	1.81	75	51	3.8	TE
2203	862	1151			2	TGCAGCTG	1.89	65	34	2.2	TE
2204	862	1151			3	AGAGCTAG	1.88	81	29	2.4	TE
2205	862	1151	Yellow		3	AGCTGTGA	1.84	75	31	2.4	TE
2206	862	1151	Yellow		3	CTCTGCTC	1.90	75	29	2.2	TE
1460	116	735	Yellow		3	CTGACATC	1.96	75	51	3.8	TE
1204	657	739	Yellow		3	GATCTGTC	1.81	78	29	2.3	TE
1299	793	738	Yellow		3	GATCATAG	1.82	75	50	3.8	TE
2187	865	1087	Yellow		3	TCAGCATC	1.81	78	29	2.3	TE
1487	902	1032	Yellow		3	TCGATGTC	1.98	47	44	2.1	TE
2294	950	1256	Yellow		3	AGAGTCGA	1.79	47	44	2.1	TE
1529	1043	939	Yellow		3	AGCTATAG	1.84	75	31	2.3	TE
1277	1077	936	Yellow		3	CTCTCTAG	1.85	75	34	2.6	TE
2141	1085	1080	Yellow		3	CTGAGCAG	1.70	54	39	2.1	TE
2212	1085	1080	Yellow		3	GAGAGCTC	1.84	32	69	2.2	TE
1352	1122	857	Yellow		3	GATCGTGA	1.87	75	31	2.3	TE
2258	1234	818	Yellow		3	TCAGTGCT	1.82	43	49	2.1	TE
2271	1244	1266	Yellow		3	TCGACTCT	2.00	50	44	2.2	TE

1364	Yellow		4	AGCTCTCT	1.87	75	41	3.1	TE
1366	Yellow		4	AGTCGCAG	1.83	75	35	2.6	TE
2233	Yellow		4	CTAGTGTC	1.85	75	29	2.2	TE
61046	Black	yellow	4	CTAGATAG	1.84	75	47	3.6	TE
61045	Black	yellow	4	GACTACGA	1.87	75	40	3.0	TE
61079	Black	yellow	4	GCTACAGC	1.85	75	34	2.5	TE
61211	Black	yellow	4	TAGCGACG	1.88	75	28	2.1	TE
61210	Black	yellow	4	TCGAGTGA	1.85	75	31	2.3	TE
61280	Black	yellow	4	AGTCACGA	1.84	75	37	2.8	TE
61428	Black	yellow	4	AGTCTGCT	1.90	75	29	2.2	TE
61245	Black	yellow	4	CTAGCTCT	1.84	75	39	3.0	TE
61499	Black	yellow	4	CTGATGCT	1.79	75	29	2.2	TE
61212	Black	yellow	4	GACTCATC	1.85	75	33	2.5	TE
61273	Black	yellow	4	GCTAGACG	1.81	75	43	3.2	TE
61278	Black	yellow	4	TAGCTGAT	1.82	75	39	2.9	TE

The RAD-sequencing was completed by the GenePool in NBAF-Edinburgh on the Illumina HiSeq platform (Project 2012135). The Illumina HiSeq 2000 platform was able to provide 150M 100-base paired-end sequences per lane. Assessing the required marker density, the sample size and the budget, the *Sbf*I restriction enzyme was used with the respective linkers, aiming to have 10x coverage. Paired-end 100-bp sequences were obtained to improve the efficiency of sequence alignment. Fifty-one red and yellow morphs from captive populations and twelve wild yellow birds were pooled together to make four libraries. The zebra finch genome was used as a reference genome for locating the RAD sequencing reads. The sequencing reads could be identified as orthologues to the zebra finch genome (which is ~1.2 Gb) (Warren *et al.*, 2010). In a previous study, ten individually labelled samples each of red- and black-headed; wild-caught Gouldian finches were RAD-sequenced to narrow down the location of the *Red* candidate region (Kim, 2011). Two RAD-tags were statistically associated ($p<10^{-10}$) with the *Red* locus. Using the Illumina HiSeq at NBAF-Edinburgh, each individual was expected to generate approximately 69,000 SNP-containing sequences, and thus we generated raw Illumina reads to a depth of 17x coverage per locus based on this estimate. Each forward sequence contained a 5-basepair MID (molecular identifier) that allowed reads to be assigned to a specific individual, followed by 8 bases of the *Sbf*I restriction enzyme footpoint (CCTGCAGG), leaving 87 locus-specific bases per read. The interval between markers was expected to average approximately 18 kb (57 sites/megabase). The RAD-seq parameters and expected restriction site count are shown in Table 3.2.

Table 3.2. RAD-Seq parameters and restriction site count in the zebra finch reference genome used to predict the expected number of RAD tags in the Gouldian finch.

Read length	100-bp
Read type	Paired end
Enzyme	*Sbf*I
Overhang	TGCA
Motif	CCTGCAGG
No. of sites in genome	69,041
Sites/Mb	57
No. of tags	138,082
No. sequences for coverage	2,347,394
Genome Size (bp)	1,221,622,102
GC contents	0.46
Coverage	17
Plexity	63
No. reads per lane on HiSeq	150,000,000
Sequences per pool	147,885,822
Does the pool fit in one lane?	Yes

Note: Sites/Mb: no. of restriction sites per megabase. No. of tags: no. of sites in the genome in double strands. No. sequences for coverage: no. of tags in genome multiplied by coverage. Coverage: how many times each site is expected to be read for inclusion in the double-stranded contig. Plexity: how many samples can be pooled for sequencing. Sequence per pool: no. sequences for coverage multiplied by plexity.

Basic bioinformatics analysis, such as quality control and base calling were processed by GenePool. *Stacks* was then used to analyse the Illumina reads and extract RAD-tags. Within *Stacks*, the process_radtags program was called first to check the Illumina raw reads, which were converted to FASTQ format with the correct barcodes and RAD cut sites, and to demultiplex the pooled data (Kelley *et al.*, 2010). The process_radtags.pl discarded the problematic reads caused by various sequencing errors, such as a prolonged drop in sequence quality, wrong or non-existent barcodes, insufficient restriction enzyme cutting sites, and skipped nucleotides. However, the program can correct isolated errors in reads or occasional errors in barcodes (Catchen *et al.,* 2011). The data were called as gzipped, paired-end, Illumina HiSeq data by specifying the input file with different parameters. The output of the process_radtags gave two reads, one single-end and paired-end per barcode, in a FASTQ file.

Quality-checked sequence reads in FASTQ format were aligned to the zebra finch genome using SeqMan NGen 4.0 to produce a BAM file then processed with SAMtools (Li *et al.*, 2009). SAMtools is a software used to process short sequence alignments in different formats, such as SAM, BAM, and CRAM. The BAM (Binary Alignment/Map) format is a binary alignment file used to store the read alignments based on reference sequences and for use in downstream analysis.

The ref_map.pl pipeline used the BAM file to build loci (pstacks), create a catalogue (cstacks) and match each sample against the catalogue (ustacks). The output files of ref_map were in TSV format. Three files were generated by pstacks for each individual on each chromosome: xxx.tags.tsv, xxx.snps.tsv, and xxx.alleles.tsv. The xxx.tags.tsv files mainly included the aligned reference chromosome and location, each primary read and a consensus sequence, locus information and the raw sequence reads. The xxx.snps.tsv files included the positions of the SNPs within the locus and nucleotide types. The xxx.alleles.tsv files were composed of the haplotypes of the called SNPs at each locus together with the percentages

and raw numbers of reads for each haplotype. The cstacks generated three similar files for each chromosome: batch_x.catalog.tags.tsv, batch_x.catalog.snps.tsv, and batch_x.catalog.alleles.tsv. These files also contained information such as consensus sequence, identified SNPs and observed alleles for each locus, respectively. The sstacks program outputted the xxx.matches.tsv files for each individual within each chromosome. The xxx.matches.tsv files included the locus matched to the catalogue – the matching haplotype.

A Perl program 'population.pl' was employed to do population genomic analysis, and to compute the population genetic statistics, such as expected heterozygosity, observed heterozygosity, nucleotide diversity estimation and population differentiation (F_{ST}). To prepare the population map, the sample set was split into two populations: red- and yellow-headed morphs. The F_{ST} per site was calculated for pairwise red–yellow populations. The output files of the populations program were in TSV format, and included SNP locations and population genetic statistics. The batch_x.sumstats.tsv files included a set of SNP information and population genetic statistics (e.g. observed heterozygosity, observed homozygosity, expected heterozygosity, expected homozygosity). The batch_x.sumstats-summary.tsv files compared the summary of population genetic statistics between two populations. The batch_x.fst_1-2.tsv files calculated the F_{ST}, P values for Fisher's exact test and LOD score. The population program also exported the batch_x.haplotypes.tsv file and batch_x.markers.tsv file, which contained the haplotype information for each individual and marker information. DNASTAR's SeqMan NGen program (NGen) was used for contig assembly (Madison, http://www.dnastar.com/t-products-seqman-ngen.aspx). A repeat masked version of the zebra finch reference assembly was used for mapping. Plots from the STACKS output to compare the red vs. yellow populations were generated from the RAD data using modifying scripts in R (K-W Kim, unpublished).

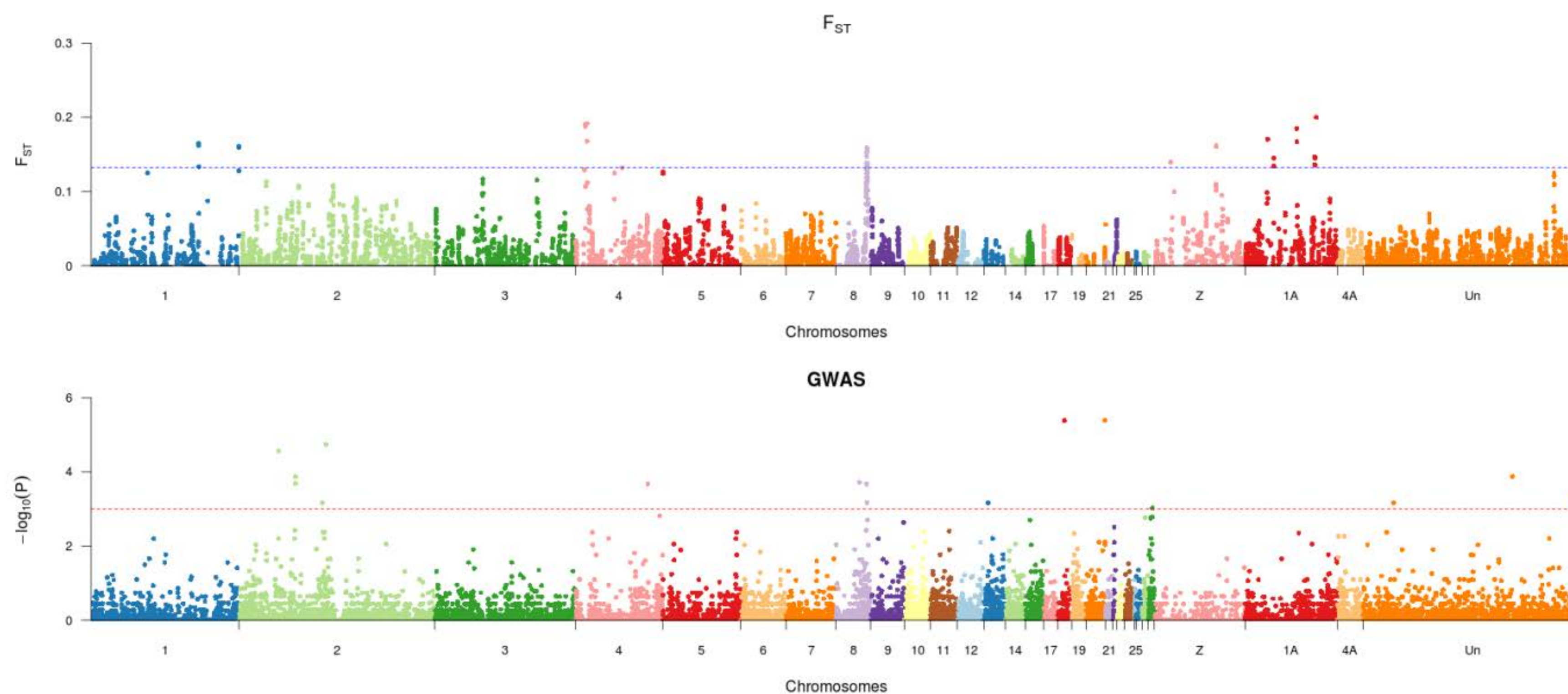
3.3. Results

Four RAD libraries were sequenced in a single Illumina HiSeq lane to produce 150M 100-bp paired-end sequences. A total of 201,303,018 bp in Illumina raw reads were generated for the Gouldian finch samples, of which 172,268,208 raw reads were retained and used in downstream analysis after data filtering. The population program was used to compare two populations: the captive red with the captive yellow individuals. The analysis included 23 captive yellow and 12 captive red individuals.

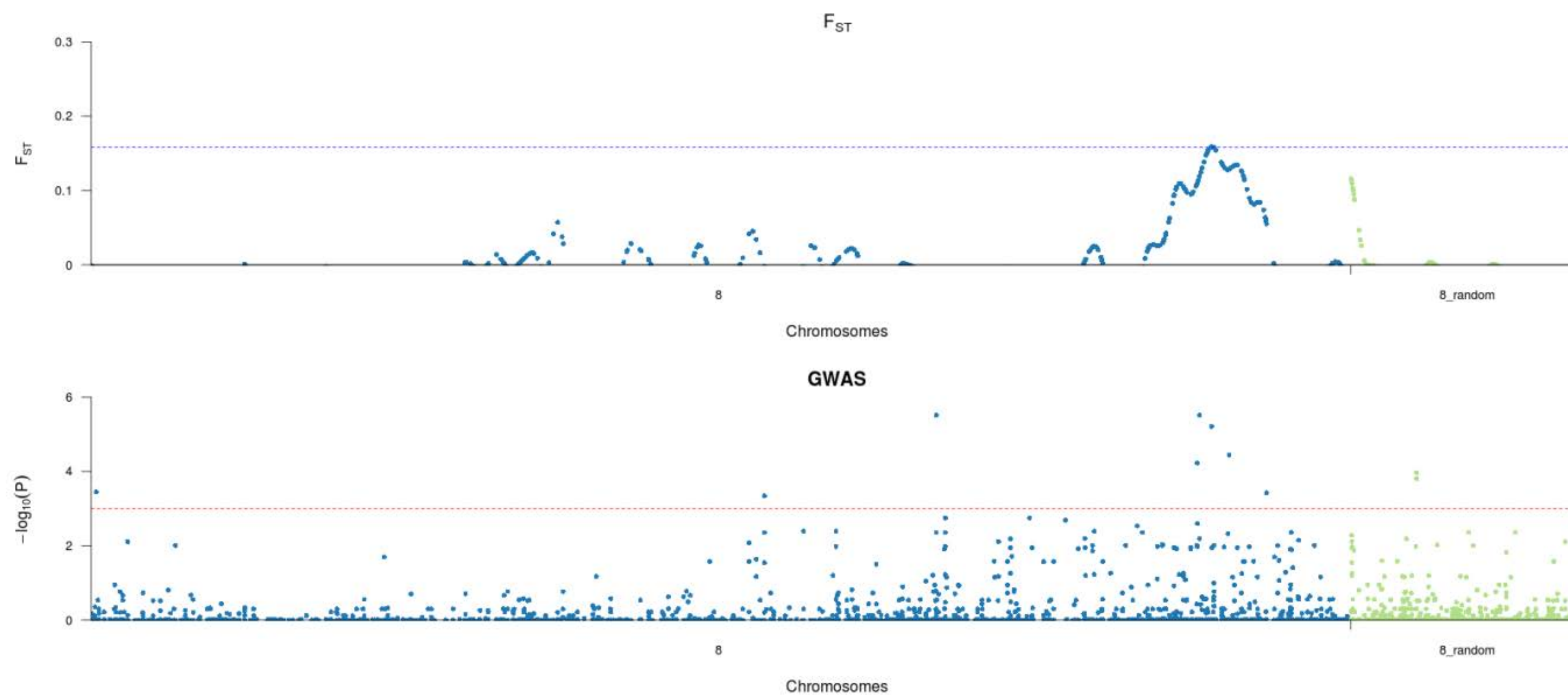
The genome-wide mean F_{ST} value, including all SNPs in each analysis, was 0.017 in the yellow vs. red individuals within the same captive population. This is a moderate level of genetic differentiation between the two subpopulations being compared. F_{ST} values higher than the 99.9th percentile of all F_{ST} values were considered to represent extreme differentiation (Figure 3.3). Figures 3.3 (a) and (b) show F_{ST} value and P-values for Fisher's exact tests for SNPs samples across the genome and within chromosome 8, respectively. Raw F_{ST} values were smoothed using a kernel smoothing algorithm. P-values for Fisher's exact test were corrected for false discovery rate using Benjamini & Hochberg's (1995) method. Several contigs were found to have SNPs showing significantly high F_{ST} (chromosomes 1, 1A, 4, 8 and Z) or P-values (chromosomes 2, 4, 8, 13, 18 and 20). However, only SNPs on chromosome 8 showed both high F_{ST} and P-values that corresponded to the same region. The position of the maximally significant point within these RAD contigs is at around the 25 Mb position.

Figure 3.3. F_{ST} and (b) P-values for Fisher's exact test (a) across the genome and (b) on chromosome 8, respectively. F_{ST}: raw F_{ST} values were smoothed using a kernel smoothing algorithm. The blue horizontal line shows the 99.9th percentile of all F_{ST} values. P-values for Fisher's exact test were corrected for false discovery rate by using Benjamini & Hochberg's (1995) method. The red horizontal line indicates P = 0.001. (Figure prepared by K-W Kim)

(a)



(b)



3.4. Discussion

In an attempt to identify the genetic variants that contribute to the *Yellow* polymorphism, we undertook a genome-wide association study in 63 Gouldian finch samples including red and yellow morphs, genotyped by RAD-sequencing, and identified 63,833 single-nucleotide polymorphisms. The probability of each observed difference in genotype frequencies between the red and yellow morphs was calculated in a *Stacks* analysis.

The threshold level is subjective as we used 0.001 in this study, there would be more significant SNPs at a less stringent threshold if we use 0.05 instead. We used a significance level of $P = 0.001$ to indicate SNPs with significant frequency differences between the two morphs and that were therefore potentially in linkage disequilibrium (LD) with the causative locus. Figure 3.3 shows that much broader F_{ST} peak on chromosome 8 than observed on other chromosomes. This is the only peak where the allelic differentiation detected by F_{ST} is apparently coincident with significant P values calculated using Fisher's exact test. As the samples in this study are from small families in a restricted captive population, in which individuals are typically separated by just a few generations, the implied high extent of the LD is not unexpected.

Through focused sequencing or higher-density SNP analysis in this region, ideally in a larger sample of wild-caught Gouldian finches, the association might be confirmed and more precisely localised. A RAD protocol could be used in which a more frequent restriction site (e.g. *Pst*I) is used to generate a higher density of RAD tags, or indeed whole genomes could be compared.

In conclusion, we have used GWAS via RAD-seq to discover a region of chromosome 8 that is potentially associated with the *Yellow* locus that causes the red- / yellow-headed

polymorphism in Gouldian finches. The results are a reminder of the potentially limiting effects of sample size and LD on our ability to use GWAS to discover the genetic variants determining a specific trait. There is no firm evidence for implicating chromosome 8 or rejecting other regions. However, we should treat the region on chromosome 8 around the 25-MB position as a potential candidate region until further investigations are done. Overall, the sample size of SNPs and individuals used here was too small to enable the gene to be pinpointed.

CHAPTER FOUR: A genetic mapping and candidate gene approach of the yellow-headed colour trait in the Gouldian finch, *Erythrura gouldiae*

Abstract

Carotenoids are a major class of pigment colouration in birds and are responsible for the bright yellow, orange or red appearance of the plumage in many avian species. However, little is known about the molecular genetic basis of carotenoid-based plumage polymorphisms in birds. This chapter aims to identify the candidate genomic region responsible for carotenoid-based plumage differences in the Gouldian finch by using a combination of mapping and candidate gene approaches.

The linkage map of Gouldian finch chromosome 8 was constructed using 13 microsatellite markers isolated from the Gouldian finch but assigned to zebra finch chromosome 8 based on sequence similarity. The pedigree included 672 individuals from 80 captive families. The linkage group is homologous to the zebra finch (*Taeniopygia guttata*) chromosome Tgu8 / Tgu 8_random. This linkage group covered a distance of 102.8 cM, with an average marker spacing of 12.85 cM, which increased the length of mapped Gouldian finch genome to 447.2 cM when combined with the map of chapter 2. This map length could be an underestimate, as four genotyped segments remain unassigned at a LOD threshold of 3.0. No intra-chromosomal rearrangement was apparent when comparing the homologous loci between the zebra finch and Gouldian finch. This result further confirms the high degree of synteny between these two Estrilid species. Two loci with relatively high LOD values for an association with *Yellow* are GF08_21 (LOD=1.63) and GF08_08 (LOD=1.20), which are

in the region 20.2-23.1 Mb on zebra finch chromosome 8. This suggests that the *Yellow* gene might be close to this region.

A *CYP2J2* gene, which is close to this potential region (24.6 Mb in zebra finch), has been implicated in carotenoid colouration in the zebra finch (N Mundy, pers comm). Therefore, based on the position of F_{ST} peak found in chapter 3 and candidate causal gene in zebra finch, we tested for an association between the *CYP2J2* gene and the *Yellow* locus involved in carotenoid colouration in the Gouldian finch. The *CYP2J2-2* gene intron 2 was amplified and sequenced in captive, pedigreed Gouldian finches. Ten SNPs were identified in 627 bp of intron 2, and five different haplotype blocks were constructed. After analysing the SNPs, *CYP2J2-2* haplotypes were found to be associated with different phenotypes (LOD=2.71). Of five haplotypes, one was consistently associated with the red allele and the other four haplotypes co-segregated consistently with the yellow allele, except for 2 SNPs in one individual. This co-segregation suggests that the *CYP2J2-2* gene is likely to be associated with *Yellow*. However, due to the limited available family size, insufficient data were available to provide strong statistical support for this association. More phenotyped birds need to be genotyped as a priority.

4.1. Introduction

4.1.1. Candidate genes for carotenoid colouration in the Gouldian finch

The Gouldian finch head colour polymorphism is a stable, genetically determined feature of natural populations of the species (Franklin and Dostine, 2000). The different head colours, whether red, yellow or black, are associated with their behaviour, personality, and the choice of mates and sex allocation in offspring (Pryke and Griffith, 2009; Williams *et al.*, 2012). As discussed in chapter one, birds cannot synthesise carotenoids *in vivo* (Goodwin, 1984; Schiedt, 1990). Carotenoids are taken up in food and deposited in different tissues (such as skin, plumage, beak, claw and viscera) via alternative pathways (Stradi, 1998; Andersson *et al.*, 2007; Prager *et al.*, 2009). The yellow-headed Gouldian finch (with a yellow-tipped bill) probably lacks the capacity to convert the dietary yellow into a derived red ketocarotenoid (Brush and Seifried, 1968). In another example, in the house finch (*Haemorhous mexicanus*), which varies in colour from yellow to red, it has been shown that the pigmentary basis of plumage colouration between red and yellow is that the carotenoid in the red plumage is 3-hydroxy-echinenone and in yellow feathers is lutein (Inouye *et al.*, 2001).

The *CYP2J2* gene belongs to cytochrome P450, family 2, subfamily J, polypeptide 2. This gene is a protein-coding gene that encodes an enzyme from the cytochrome P450 superfamily. The cytochrome P450 (CYP) proteins are haemoproteins and monooxygenases, which use various kinds of substrates in enzymatic reactions to catalyse reactions such as drug metabolism, synthesis of cholesterol, steroids and other lipids (Capdevila *et al.*, 1995). The CYP2J2 protein is an epoxygenase that catalyses the formation of epoxide with a carbon-carbon double bond (Berlin *et al.,* 2010). In birds, four CYP candidate genes have been discovered on chromosome 8 so far. The *CYP2J2* gene includes 9 exons and 8 introns, like other CYP2 genes (King *et al.*, 2002). The latest RNA-seq data from zebra finch

demonstrated that three of the four genes were expressed in red bill tissue (N. Mundy, pers. comm.). *CYP2J2-1* and *CYP2J2-2* are arranged in tandem in the zebra finch genome. *CYP2J2-1* and *CYP4X* were strongly expressed in both wild type (red-billed) and yellow-billed zebra finches in beak, liver and retina tissues. *CYP2J2-2* was strongly expressed in the wild type beak, but was not (or was only weakly) expressed in yellow bills. These expression results suggest that a regulatory mutation affects the tissue-specific expression of *CYP2J2-2*. *CYP2J2-2* was localised to chromosome 8 random 34,138–43,936 bp in zebra finch.

To date, there is only one example of a genomic sequence being associated with a specific phenotypic variant in the Gouldian finch – the genetic basis of the melanin-based polymorphism at the *Red* locus (Kim, 2011). Here, we tested the hypothesis that *CYP2J2* was the causative gene for the carotenoid-based polymorphism in the Gouldian finch. We included the mapping of markers close to the *CYP2J2* gene region and tested these for cosegregation with *Yellow*, as this could provide further support for an association.

4.1.2. The candidate gene approach

The candidate gene approach is widely used to identify genetic factors for specific traits (Hubbard *et al.*, 2010). This research technique effectively detects genetic variants in a potentially causal gene that is known to contribute to a specific function or phenotypic trait in another species (usually a laboratory model). If the frequencies of the single nucleotide polymorphisms (SNPs)/haplotypes in the candidate region differ significantly between morphs, then this SNP/haplotype is considered to be associated with the specific trait. The haplotype is a set of tightly linked alleles on one chromosome from either parent and inherited as a unit (Clark, 2004). As the rapid development of genome-wide association studies (GWAS) over the past few years, a massive amount of genes have been detected and

identified by the high-throughput genotyping platforms (Wilkening *et al.*, 2009). However, the coverage of the genome, the available study population, or the budget could restrict the performance of GWAS. Therefore, classic candidate gene approaches are still valuable, particularly for detecting polymorphisms with low allele frequencies, incomplete GWAS datasets, and post-GWAS to identify the actual causative variants.

In our case, given the low allele frequency of the yellow head phase, the low density of genomic markers, and the limited extent of the available pedigree data, the whole-genome association study in chapter three did not discover a genetic difference associated with the different morphs. Therefore, this study aimed to explore the molecular basis of plumage colour polymorphism in the Gouldian finch, using the candidate gene approach via chromosome linkage mapping and sequencing of a specific gene. The genomic region of interest was the *CYP2J2* gene. If the locus proves to be associated with the yellow head colour in the Gouldian finch, then the candidate gene approach could be further utilized to test whether this gene also accounts for carotenoid-based phenotypes in other avian species.

4.2. Materials and methods

4.2.1. Chromosome 8 mapping and linkage analysis

Mapping population and genomic DNA extraction

The study population was the same as the population used in chapter 2. Six hundred and seventy-two blood samples of pedigreed Gouldian finch obtained from an Australian captive population (Pryke and Griffith, 2006) were included in this study. The samples were blood samples stored in absolute ethanol, DNA extracted using ammonium acetate and quantified

and diluted for PCR following the methods described in Chapter 2.

Microsatellite marker development

Microsatellite markers had previously been developed using traditional cloning followed by Sanger sequencing (Kim *et al.*, 2015). The SciRoKo33 software was used to assess microsatellite quality (Kofler *et al.,* 2007). All the Gouldian finch microsatellite markers were assigned to chromosome locations on the zebra finch genome based on their sequence similarity (Dawson *et al.,* 2006) and were selected to ensure adequate coverage of markers on chromosome 8. The locations of the Gouldian finch microsatellite sequences in the zebra finch genome were identified using a standalone BLAST (http://www.ncbi.nlm.nih.gov/blast/download.shtml) based on an e value $< 1 \times e^{-10}$ as the filter threshold. Twenty-eight microsatellite sequences assigned to zebra finch chromosome 8 were selected and tested for use in linkage mapping.

Marker amplification and polymorphism

Primer sets were designed based on zebra finch sequence alignments using online Primer3 v0.4.0 (Rozen and Skaletsky, 2000) with default parameters. PCR reactions were completed using a DNA Engine Tetrad PTC-225 Peltier thermal cycler in a 2-µl Qiagen multiplex system consisting of 2 µl template DNA at ~10 ng/µl, and dried by evaporation at room temperature for 30 minutes, followed by the addition of 1 µl of Qiagen HotStar PCR Master Mix (Qiagen Inc., Kenta *et al.,* 2008), 1 µl of primer mix (all primers are at 0.2 µM with the forward fluorescently labelled) and a drop of oil to prevent evaporation. The PCR program used for all the successful amplification was: 15 min of initial denaturation at 95°C, then 30 cycles of 30 s at 94°C, 90 s at an annealing temperature at 60°C and 60 s at 72°C, followed by a 30-min extension phase at 60°C. The PCR reaction ended with a final minute at 15°C.

PCR amplification was checked by electrophoresing PCR product on a 1.5% agarose gel stained with SYBR safe and photographed using GeneSnap. Two Gouldian finch individuals were used initially to test amplification, then eight individuals were used to test for polymorphism in the loci that amplified successfully. These individuals were the same as those used for testing the microsatellite markers in chapter 2. To increase efficiency, four to five markers were pooled into each multiplex set. The multiplex sets used for linkage mapping were created using Multiplex Manager (Holleley and Geerts, 2009). Two-microlitre Qiagen Multiplex PCR reactions were used to test the multiplexes.

Genotyping and error checking

Raw genotypic data were generated on the ABI 3730 48-well Capillary Sequencer (Applied Biosystems) using ROX500 size standards. Genotypic data were analysed, alleles scored, and error checked using GeneMapper v.3.7 software (Applied Biosystems). CERVUS 3.0 (Kalinowski *et al.,* 2007) was used to check polymorphism, estimate null allele frequencies and calculate the observed and expected heterozygosities.

Error checking was conducted by: (1) General error checking. Single-basepair binning mismatches caused by multiplex interactions, inherited null alleles and rare mutations could be detected and corrected manually by comparing pedigree information and genotypic data. (2) Error checking by CriMaker1.0 – a Perl program provided by K. W. Kim (2011). The GeneMapper export files were converted to CERVUS import files. Then, CriMaker1.0 was used to convert CERVUS format file into CRIMAP format. Any outliers from expectation under the Mendelian inheritance of markers, such as the offspring from one family having more than 4 alleles were identified and corrected by running the CRIMAP Prepare program.

Construction of chromosome genetic map

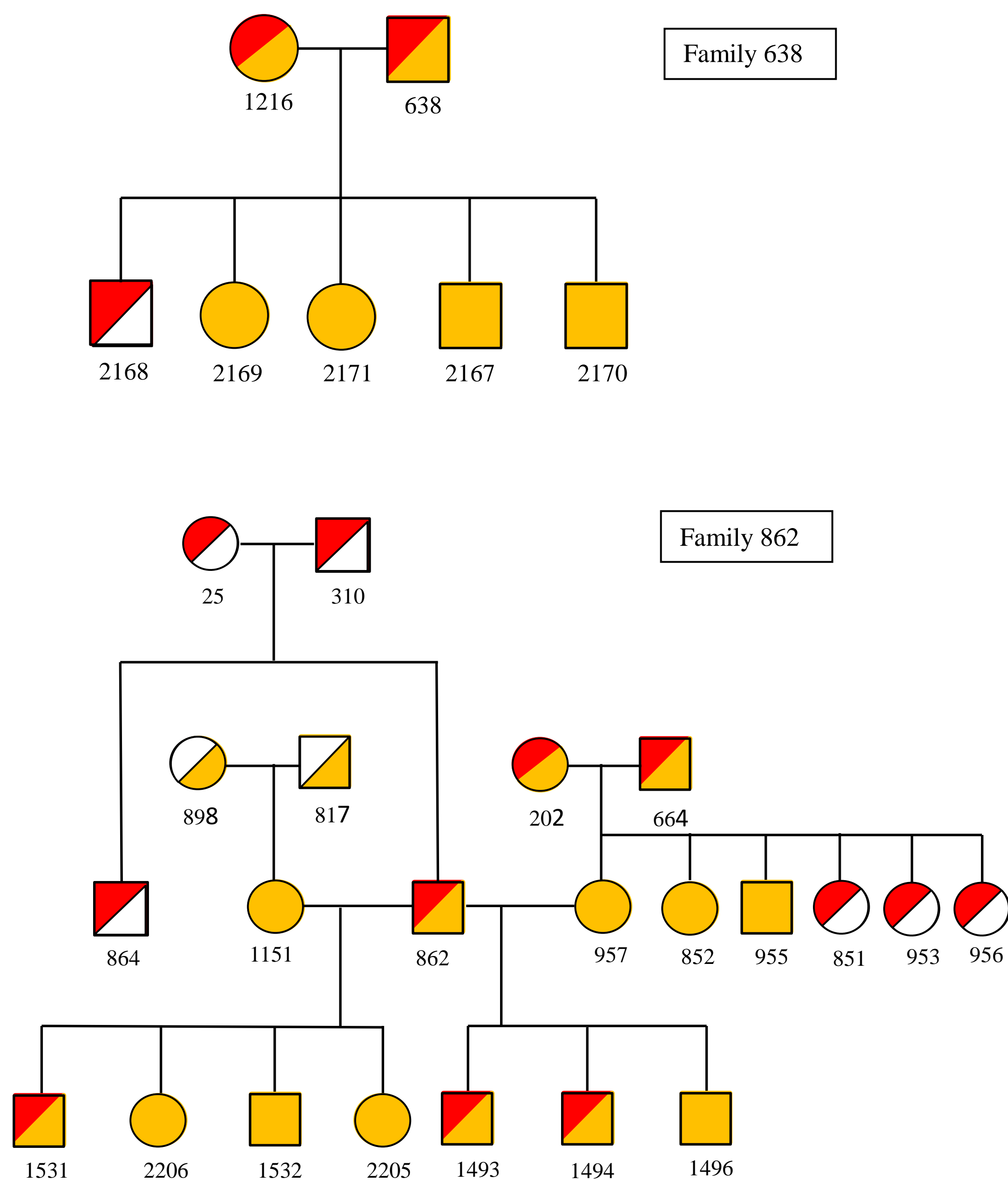
The Gouldian finch chromosome 8 genetic map was constructed using CRIMAP (http://biobase.dk/Embnetut/Crimap/analyse1.html; Lander and Green, 1987; Green *et al.*, 1990). The CRIMAP 'twopoint' program was used to calculate the LOD score between each pair of markers. The 'build' command was used to calculate the order and the interval of markers. The threshold for linkage was a LOD value of 3.0. Groups of pairwise markers with LOD scores higher than 3.0 were considered to be linked. All map distances are given in Kosambi centiMorgans (cM).

4.2.2. Candidate gene approach: the *CYP2J2-2* gene

Sample population and pedigrees of families

Sixty-three pedigree individuals, which were the same as the individuals analysed using RAD sequencing (chapter three), were Sanger-sequenced with a single read to identify any SNPs. These individuals included two captive families segregating for yellow (including red and yellow birds) and 23 yellow-headed birds. The pedigrees are shown in Figure 4.1. These two families included 29 individuals with known phenotypes and zygosity (heterozygote or homozygote), or in which phenotypes and zygosity could be identified from parent-offspring information.

Figure 4.1. Pedigree of the families involved in the *CYP2J2-2* intron 2 haplotype analyses. Solid red symbols indicate homozygous red individuals, based on phenotype and parent-offspring analysis where available. Yellow symbols indicated homozygous yellow individuals (all yellow phenotypes are homozygotes). The half red and half yellow symbols represent heterozygous red individuals, while half blank indicates unknown heterozygosity (if the other half is red) or unknown phenotype (if the other half is yellow). Below each symbol is the individual ID. Males are represented by square symbols, females with circular symbols.



Primer design and testing

Owing to the more rapid evolution of introns than exons in most genes, we focused initially on variation in the candidate gene intron sequences. The respective sequences were downloaded from the ENSEMBL zebra finch chromosome 8 random, on the assumption that there would be high homology between zebra finch and Gouldian finch. The *CYP2J2-2* gene is 9.8 kb in length, and consists of 9 exons and 8 introns. The length of selected region *CYP2J2-2* intron 2 was 627 bp, which ensured the adequate single-end sequencing of the PCR products and that there would be sufficient sequence to enable the PCR of a fragment of useful length after primer design. Primer3 v0.4.0 online was used to design two pairs of primers for *CYP2J2-2* intron 2. Two primer pairs were designed for intron 2 (627 bp) of *CYP2J2-2* gene were designed. To test primer amplification and specificity, two individuals (one black, one red) were tested with four primer pair combinations. The concentration of the DNA templates used in amplification testing was 20 ng/µl. The primers were diluted to 5 µM using 1x low TE buffer. The PCR amplification reaction was performed in a 20-µl PCR mix using a DNA Engine Tetrad PTC-225 Peltier thermal cycler. The PCR mix consisted of 2 µl of 10x PCR buffer (Bioline), 0.8 µl of 50 mM $MgCl_2$, 2 µl of 2 mM each dNTP, 0.1 µl of Biotaq *Taq* DNA polymerase (Bioline), 2 µl of 5 µM forward and reverse primers each, 1 µl of 20 ng/µl DNA template, and 10.1 µl ultrapure H_2O. The PCR profile was initially denatured at 94°C for 3 min, followed by 35 cycles of 94°C for 30 s, an annealing temperature of 60°C for 30 s, 72°C for 60 s, and then a final extension at 72°C for 5 min. The PCR reaction was terminated with a final minute at 15°C. The quantity of PCR product were checked by electrophoresis using a 1% agarose gel which was run at 130 volts for 30 min and DNA visualised using SYBR® Safe. The concentration of DNA in PCR products was checked using a NanoDrop spectrophotometer. The primer combination that produced a specific product of the expected size and a high DNA concentration was selected

after primer testing.

Amplification and sequencing

The high-concentration products were diluted before sequencing. For products that included non-specific amplification fragments, the QIAquick Gel Extraction Kit (QIAGEN) was used to obtain purified DNA product. The amplified products were cleaned up with ExoSAP-IT (USB). Sanger sequencing was completed using the BigDye Terminator v 3.1 cycle sequencing kit (Applied Biosystems). The sequencing program was 96°C for 60 s, then 94°C for 10 s, 50°C for 5 s, 60°C for 4 min for 30 cycles, followed by 15°C for 5 min. The BigDye system included 1 µl of BigDye v 3.1, 1.5 µl of Seq-buffer, and 0.32 µl of 5 µM forward primer, 2 µl of PCR product and 3.7µl of H_2O. The precipitation used the ethanol/EDTA method and the products were separated on an ABI 3730 DNA Analyser.

SNP identification and analysis

CodonCode Aligner v. 2.0.6 (CodonCode Corporation) and BioEdit were used to call the raw sequence bases and align to consensus contigs. High-quality sequences were obtained for most samples, with adequate polymorphic nucleotide sites. Low-quality single-read sequences were excluded from analysis. Sequence variations from each aligned contig were evaluated by manual inspection. Haplotypes were constructed using the SNP genotypes of individuals in pedigreed families. The CRIMAP 'twopoint' program was used to calculate the LOD score between each marker with individuals' genotype. The threshold for linkage was a LOD value of 3.0.

4.3. Results

4.3.1. Construction of chromosome map and linkage analysis

In this chapter, 28 microsatellite loci were tested for linkage mapping (Figure 4.2). After testing and analysing, 23 of 28 markers amplified, 13 of which were polymorphic (k=2–12; PIC=0.164–0.778) (Table 4.1). The mean number of alleles per locus was 5.62. The average polymorphic information content (PIC) was 0.41. Nine loci were monomorphic and six primers failed to amplify products (Table 4.2). These fifteen markers were excluded from map construction. Thirteen polymorphic markers were arranged into three multiplex microsatellite sets. All markers in the three multiplex microsatellite sets amplified well. For subsequent linkage mapping and analysis, 672 individuals from 80 families were typed. This genetic map includes nine microsatellites distributed across chromosome 8 (Figure 4.3). Marker GF08_15, which physically located on chromosome 8 random in zebra finch, was aligned to telomere of chromosome 8 in the Gouldian finch by linkage analysis using CRIMAP. This chromosomal map covered a total length of 102.8 cM, with an average marker spacing of 12.85 cM between loci. After building the map, linkage analysis was performed for yellow-headed informative individuals within pedigreed individuals to detect the probable genomic region of the *Yellow* gene. No statistically significant association with colour was detected with any of the 13 microsatellite markers. However, the loci with a highest LOD score were GF08_21 and GF08_08, in the region 20.2-23.1 Mb on chromosome 8 (Figure 4.4). The highest LOD was 1.63 with locus GF08_21, for both the head and beak trait information included, followed by 1.20 with GF08_08, and 1.16 with GF08_21 for the head trait only. This suggests that the *Yellow* candidate region is probably close to this region.

Figure 4.2. Twenty-eight Gouldian finch microsatellite loci mapped onto chromosomes 8 and six on 8_random in the zebra finch genome. Thirteen loci were used for linkage mapping and further analysis (underlined). Nine loci were monomorphic (**Bold**) and six loci failed to amplify (*Italics*). Marker names are shown in the map.

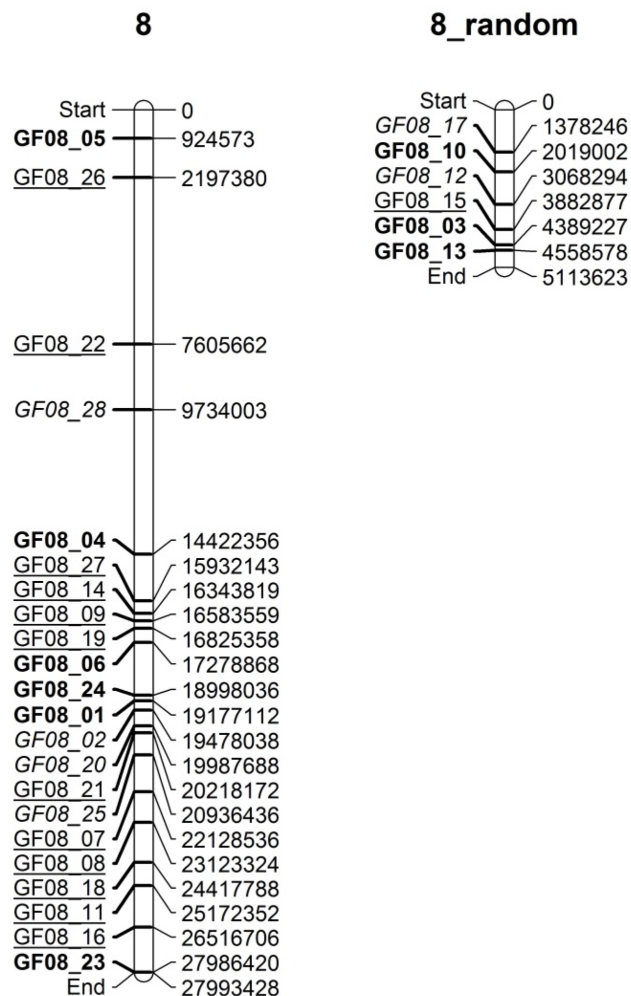


Table 4.1. Characterization and predicted genomic locations of 13 microsatellite loci of Gouldian finch *Erythrura gouldiae* on chromosome 8. All the 13 characterized polymorphic loci were assigned to the zebra finch genome and included in the construction of the Gouldian finch linkage maps and further analysis. Locus: name of locus. Chr: zebra finch chromosome. Chr Loc: The location of each microsatellite sequence was assigned to the zebra finch as the high sequence similarity (December 2011, ENSEMBL Release 65; Dawson *et al.,* 2006, 2007). Repeat motif: the motif and repeat number of each motif (subscripted). Fluoro-label: fluorescent label information. A: number of alleles observed. N: number of individuals amplified. PIC: polymorphic information content. H_O: observed heterozygosity. H_E: expected heterozygosity.

Marker	Chr	Chr Loc	Repeat motif	E-value	Fluoro label	Primers sequence (5'-3')	Primers conc. (μM)	Multiplex set	Exp allele size (bp)	Obs allele size (bp)	A	PIC	H_O	H_E
GF08_07	8	22128537	$(CA)_9$	2.00E-129	[6-FAM]	F: CTGACACCCCATCCACTCAA	0.2	1	192	186-206	9	0.778	0.712	0.805
						R: TGGGTAGTATTGTTGCACTTGG	0.2							
GF08_08	8	23123325	$(TAGA)_9$	0.00E+00	[HEX]	F: CAGGTCACCAGATGATGCAG	0.2	2	193	174-266	12	0.364	0.384	0.379
						R: CATAGTGGGCCCTCTGGTC	0.2							
GF08_09	8	16583559	$(AT)_5$	6.00E-83	[6-FAM]	F: CAGCCCATCCCAAGATTCTA	0.2	2	202	198-223	4	0.259	0.1	0.279
						R: AAGCTGGTGAAAGGAGAGGA	0.2							
GF08_11	8	25172353	$(CA)_5$	1.00E-111	[6-FAM]	F: GAGCCTTGGAAGGGCTTT	0.2	3	207	204-213	4	0.02	0.02	0.02
						R: AGCAGCGAGCAGACTAATCC	0.2							
GF08_14	8	16343819	$(TG)_{12}$	8.00E-58	[HEX]	F: AGTGGTTGGATGGCTTGTCT	0.2	3	243	229-248	8	0.41	0.314	0.447
						R: CCCCTTTTATTCCCAGCTCT	0.2							
GF08_15	8_random	3882877	$(CA)_6$	2.00E-114	[6-FAM]	F: GCATCATTCCAAAAGGCAGT	0.2	1	261	256-262	4	0.164	0.158	0.17
						R: GTGATGTGCAGTTTGGCTGT	0.2							
GF08_16	8	26516706	$(TG)_5$	2.00E-80	[HEX]	F: GGAGCACCTGGGACAGAG	0.2	3	280	280-286	4	0.437	0.262	0.508
						R: TGCTCCATCTCCACTTGGTT	0.2							
GF08_18	8	24417788	$(AG)_8$	9.00E-135	[HEX]	F: AGAAATCCCTCATTTTCCCTTC	0.2	2	282	272-282	5	0.676	0.232	0.725
						R: AACGTTCCCAACTGCTATATTCA	0.2							
GF08_19	8	16825358	$(AC)_8$	3.00E-116	[6-FAM]	F: CCAACAGCCTCTGAAAAGGA	0.2	3	293	287-291	3	0.257	0.243	0.302
						R: AGGCAGGAGTCCCATATGAA	0.2							
GF08_21	8	20218173	$(AC)_{10}$	0.00E+00	[6-FAM]	F: GTTTGGCCTCTCAAATCTTCC	0.2	1	329	322-335	8	0.576	0.611	0.643
						R: TGATGGATGTAGCCTGAAAGG	0.2							
GF08_22	8	7605662	$(GA)_6$	2.00E-167	[HEX]	F: TGGGAATTCTCAGGAAAGGA	0.2	2	345	344-346	2	0.362	0.437	0.475
						R: TTTTCTAGCAGCCCATCACC	0.2							
GF08_26	8	2197380	(GA)5	1.00E-128	[HEX]	F: AGCTGGCATCACAAGGAAGT	0.2	1	365	364-367	4	0.404	0.492	0.492
						R: TGGCCCCTCTAAGCAGATAA	0.2							
GF08_27	8	15932143	$(CAA)_6$	7.00E-81	[6-FAM]	F: ATCTTGGCAATGATGACCTG	0.2	3	369	366-376	6	0.612	0.546	0.639
						R: AGCTCTGGCACGAGACAAAC	0.2							

Table 4.2. Fifteen additional microsatellites with predicted chromosome locations tested but not include in linkage mapping and further analysis. Locus: name of locus. Chr: chromosome. Chr Loc: The location of each microsatellite sequence was assigned to the zebra finch as the high sequence similarity (December 2011, ENSEMBL Release 65; Dawson *et al.,* 2006, 2007). Repeat motif: the motif and repeat number of each motif (subscripted). Fluoro-label: fluorescent label information. A: number of alleles observed. N: number of individuals amplified. PIC: polymorphic information content. H_O: observed heterozygosity. H_E: expected heterozygosity. Status: the marker amplification and polymorphism status (Mono=monomorphic, Failed= failed to amplify or amplified a faint/fuzzy product).

Marker	Chr	Chr Loc	Repeat motif	E-value	Fluoro label	Primers sequence (5'-3')	Primers conc. (μM)	Exp allele size (bp)	Status
GF08_01	8	19177111	$(AG)_5$	1.00E-146	[6-FAM]	F: CAGGAGTGGGACTCATGGTT	0.2	86	Mono
						R: GGGAACTGGAGAGTGTAAAAGC	0.2		
GF08_02	8	19478038	$(CA)_5$	3.00E-119	[HEX]	F: TTTTGCAGCAAGAGCCTACA	0.2	91	Failed
						R: GGAGGTAGTTTGGGTGGGTA	0.2		
GF08_03	8_random	4389227	$(TC)_6$	7.00E-53	[6-FAM]	F: GAGAATGAGTTTCTCTGTGTATCAGTG	0.2	93	Mono
						R: AGCCTCGTATCTCTGGGCTA	0.2		
GF08_04	8	14422356	$(AG)_5$	4.00E-78	[HEX]	F: GCTCAGCTAACCTTGGTTTTG	0.2	99	Mono
						R: TGTTCCTGTGTGCCTACATCA	0.2		
GF08_05	8	924573	$(AT)_5$	4.00E-58	[6-FAM]	F: TAGGCTCCCTGGATTCATTG	0.2	145	Mono
						R: TTATGGATGGCACACCTGAC	0.2		
GF08_06	8	17278868	$(CA)_5$	1.00E-73	[HEX]	F: GGAATGCACTGAACTTCAACC	0.2	163	Mono
						R: CCTTCTCTTCCATCCACTGC	0.2		
GF08_10	8_random	2019002	$(GA)_{10}$	2.00E-95	[HEX]	F: GGTCTTCAGCTGGGACAAAA	0.2	202	Mono
						R: TTACCATATTGGAACCCCAAA	0.2		
GF08_12	8_random	3068294	$(AG)_5$	5.00E-96	[HEX]	F: AAAATATGCAGAGTTTTGCTAGAAT	0.2	209	Failed
						R: AGGATGATTTTGATGTTTGTTTTG	0.2		
GF08_13	8_random	4558578	$(AT)_5$	6.00E-124	[6-FAM]	F: TGGAATAGCAGATGGGGAAC	0.2	238	Mono
						R: TTGTGTTGCCACTTTTGGAA	0.2		
GF08_17	8_random	1378246	$(GA)_6$	6.00E-78	[6-FAM]	F: TTTGTTTGGAAATGTCCCTTG	0.2	281	Failed
						R: GAAACACAAGAAGCGTGCAG	0.2		
GF08_20	8	19987689	$(CT)_6$	5.00E-115	[HEX]	F: TTGGGTTTTCAAGGGGATTA	0.2	308	Failed
						R: CTCTGTGCAGGGAGAGCTG	0.2		
GF08_23	8	27986419	(A)12	6.00E-150	[6-FAM]	F: CCACAGAAACACCCAGGTTC	0.2	348	Mono
						R: TGGGTTTGAGGTAGTGTCCTG	0.2		
GF08_24	8	18998035	$(TG)_5$	5.00E-177	[HEX]	F: GTTTGTGTGCGGTTTGTGAG	0.2	350	Mono
						R: CCTGCTTAATTCTGTGCCTGA	0.2		

GF08_25	8	20936435	$(AC)_6$	0	[6-FAM]	F: GCTGCCAAATAGCCATTCTT	0.2	362	Failed
						R: ACTTTTGAGCAATTTGAGAAGAGA	0.2		
GF08_28	8	9734003	$(TG)_6$	1.00E-171	[HEX]	F: GGTGGGAGGTGTGTATGCTT	0.2	437	Failed
						R: CATTTTTGGTCTGCATCGTG	0.2		

Figure 4.3. Sex-averaged linkage map of Gouldian finch chromosome 8 (left). Nine microsatellite loci with assigned locations on zebra finch chromosomes that were supported by a LOD > 3.0 are presented. Linkage group 8 is compared between Gouldian finch and zebra finch. All the loci were found to be homologous to zebra finch except locus GF08_15, which was on chromosome 8 random in zebra finch. All map distances in the Gouldian finch genome, with positions, are given in Kosambi centiMorgans (cM).

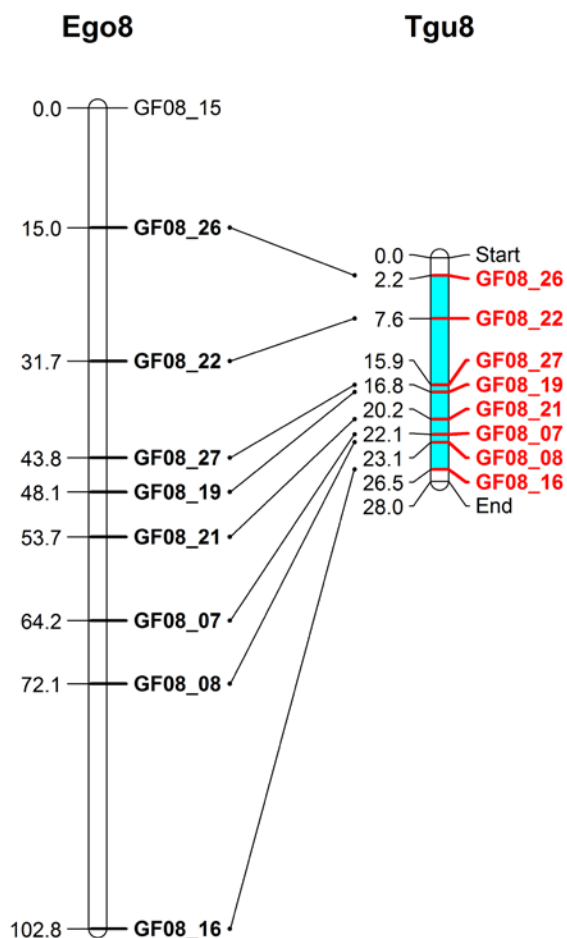
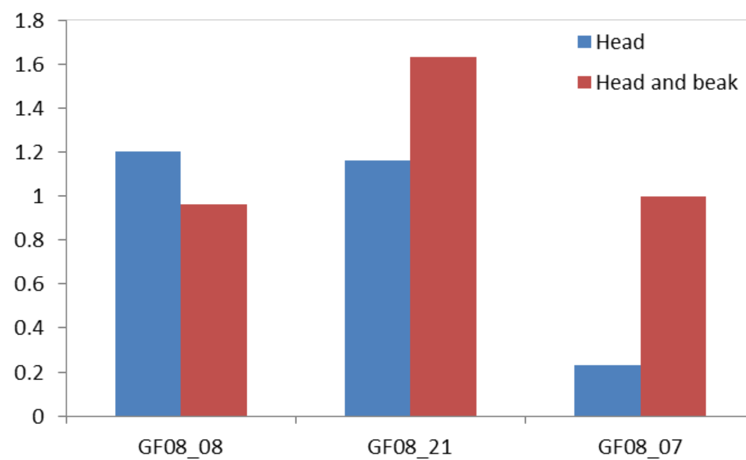


Figure 4.4. Association analysis of chromosome 8 microsatellite loci with the yellow-headed colour trait. Data are shown for the three loci with highest LOD scores. Blue bars show the LOD score with head colour (yellow); red bars show the LOD score using both head and beak colour (yellow head or black head with yellow beak).



4.3.2. Construction of haplotypes and genetic analysis of the *CYP2J2-2* gene

There were 10 genetic variants between sites 93–268 of the *CYP2J2-2* gene intron 2 across the 63 sequenced individuals (Table 4.3). Five haplotypes were constructed using the SNP genotypes of individuals in two pedigreed families. Among the five haplotypes, one haplotype almost perfectly cosegregated with the red allele and the other four haplotypes co-segregated consistently with the yellow allele. The five haplotype blocks in *CYP2J2-2* intron 2 are shown in Table 4.4, which includes 22 individuals from two pedigreed families for which we had sequence data, and four additional sequenced yellow phase individuals. The genotypes of family 638 comprised 3 haplotypes, and family 862 included 5 haplotypes. Only one red haplotype was identified (GTTATGGGCA), while yellow haplotypes included four different types: TGTGTGGCTA, TGCGAACGCG, GGTATGGGCG, and GGTATGGGCA. The genetic haplotypes were perfectly consistent with the phenotypes shown in Figure 4.1, except for two SNPs at sites 93 and 133 in sample 2168 (Table 4.4). For linkage analysis, 29 individuals from 2 families were analysed between 10 SNPs and the Yellow-headed trait using CRIMAP 'twopoint'. Eight SNPs (except sites 93 and 133) showed the same LOD score which was 2.71. The correlated co-segregation indicates that the *CYP2J2-2* gene is likely to be associated with the carotenoid-based head colour polymorphism in the Gouldian finch.

Table 4.3. Single nucleotide polymorphisms identified in the *CYP2J2-2* gene. 1=Allele 1, 2=Allele 2, MAF means minor allele frequency. This frequency estimate based on 26 samples, which included two pedigreed families and 4 additional yellow individuals.

				Genotype frequency			
Gene	SNP (bp)	Allele 1	Allele 2	1/1	1/2	2/2	MAF
CYP2J2-2	93	G	T	0.15	0.50	0.35	0.39 (G)
Intron2	133	G	T	0.65	0.35	0.00	0.18 (T)
	134	C	T	0.00	0.31	0.69	0.16 (C)
	140	A	G	0.15	0.54	0.31	0.42 (A)
	222	A	T	0.00	0.31	0.69	0.16 (A)
	230	A	G	0.00	0.31	0.69	0.16 (A)
	236	C	G	0.00	0.31	0.69	0.16 (C)
	249	C	G	0.15	0.54	0.31	0.42 (C)
	250	C	T	0.31	0.54	0.15	0.42 (T)
	268	A	G	0.42	0.50	0.08	0.33 (G)

Table 4.4. *CYP2J2-2* intron 2 haplotypes correlated with red or yellow alleles. Genetic variants are shown. The different colours represent different haplotype. Red haplotype is GTTATGGGCA (Red), yellow haplotype 1 is TGTGTGGCTA (Orange), yellow haplotype 2 is TGCGAACGCG (Green), and yellow haplotype 3 is GGTATGGGCG (Blue), yellow haplotype 4 is GGTATGGGCA (Purple). The genetic haplotypes correlated almost perfectly with red or yellow alleles; except for two SNPs at sites 93 and 133 in sample 2168 (*Italic*).

ID	Father	Mother	Pheno	Single nucleotide polymorphism genotype position																			
				93		133		134		140		222		230		236		249		250		268	
Family 638																							
1216			red	G	T	T	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
2168	638	1216	red	***T***	T	***G***	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
2169	638	1216	yellow	T	T	G	G	C	T	G	G	A	T	A	G	C	G	G	C	C	T	G	A
2171	638	1216	yellow	T	T	G	G	C	T	G	G	A	T	A	G	C	G	G	C	C	T	G	A
2167	638	1216	yellow	T	T	G	G	C	T	G	G	A	T	A	G	C	G	G	C	C	T	G	A
2170	638	1216	yellow	T	T	G	G	C	T	G	G	A	T	A	G	C	G	G	C	C	T	G	A
Family 862																							
862	310	25	red	G	G	T	G	T	T	A	A	T	T	G	G	G	G	G	G	C	C	A	G
864	310	25	red	G	G	T	G	T	T	A	A	T	T	G	G	G	G	G	G	C	C	A	G
851	664	202	red	G	T	T	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
953	664	202	red	G	T	T	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
956	664	202	red	G	T	T	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
1493	862	957	red	G	T	T	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
1494	862	957	red	G	T	T	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	A	A
1531	862	1151	red	G	T	T	G	T	C	A	G	T	A	G	A	G	C	G	G	C	C	A	G
957	664	202	yellow	T	T	G	G	T	T	G	G	T	T	G	G	G	G	C	C	T	T	A	A
852	664	202	yellow	T	T	G	G	T	T	G	G	T	T	G	G	G	G	C	C	T	T	A	A

955	664	202	yellow	T	T	G	G	T	T	G	G	T	T	G	G	G	G	C	C	T	T	A	A
1151	817	898	yellow	G	T	G	G	T	C	A	G	T	A	G	A	G	C	G	G	C	C	A	G
2206	862	1151	yellow	G	T	G	G	T	C	A	G	T	A	G	A	G	C	G	G	C	C	G	G
1496	862	957	yellow	G	T	G	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	G	A
1532	862	1151	yellow	G	G	G	G	T	T	A	A	T	T	G	G	G	G	G	G	C	C	G	A
2205	862	1151	yellow	G	G	G	G	T	T	A	A	T	T	G	G	G	G	G	G	C	C	G	A
Additional yellows																							
1366			yellow	G	T	G	G	T	C	A	G	T	A	G	A	G	C	G	G	C	C	G	G
1299	793	738	yellow	T	T	G	G	T	T	G	G	T	T	G	G	G	G	C	C	T	T	A	A
2187	865	1087	yellow	G	T	G	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	G	A
2233			yellow	G	T	G	G	T	T	A	G	T	T	G	G	G	G	G	C	C	T	G	A

4.4. Discussion

In chromosomal mapping, 28 new Gouldian finch microsatellite markers were developed and tested, based on the zebra finch genome. Twenty-two of 28 markers were amplified and 13 markers were polymorphic. A microsatellite-based sex-averaged linkage map of the Gouldian finch chromosome 8 was constructed using 672 pedigreed captive individuals from 80 families. This linkage map included 9 loci that are homologous to zebra finch chromosome 8. The linkage order of the markers completely matched the physical map predicted from the zebra finch, except for locus GF08_15, which was on chromosome 8 random in zebra finch (Figure 4.2) and located near to the telomeric region of chromosome 8 in the Gouldian finch. This suggests the location where the section of GF08_15 on chromosome 8 random actually belongs. The comparative microsatellite-based map demonstrated a high degree of synteny between zebra finch and Gouldian finch, which is not unexpected since these two species are members of the same family, but is consistent with the results in chapter 2.

To identify the candidate region, i.e. the location of the *Yellow* gene in the Gouldian finch genome, the linkage analysis was performed using the genotyping data for informative yellow morph individuals within the pedigree. The loci with relatively high LOD values were GF08_21 and GF08_08, which are in the region 20.2-23.1 Mb on chromosome 8. Even though no locus was found to be significantly associated with the yellow-headed trait (LOD $>$ 3) using CRIMAP 'twopoint' analysis within 13 markers, this evidence suggests that the *Yellow* candidate gene might be close to this region. LOD $\geqslant$ 3 is the general threshold value considered to demonstrate a specific genotype–phenotype association. However, the LOD scores were limited by the marker coverage and by having only a few informative families, and in which not all individuals were fully informative (i.e. potential mis-score

yellow beak in black-headed birds). Therefore, LOD > 1 provides suggestive evidence for a candidate region (Slate *et al.,* 2002; Dawson *et al.,* 2006, 2007). The *CYP2J2-1* gene is located at 24.6 Mb on chromosome 8 in zebra finch. In our marker map, GF08_18 is the closest locus to the candidate region if the *CYP2J2-1* gene is located in the same position in the Gouldian finch and zebra finch. However, the quality of the genotyping data for GF08_18 was not very high, and some of the informative data were missed. This might explain why no association could be detected with the GF08_18 locus. Alternatively, there may be interspecific chromosomal differences between zebra finch and Gouldian finch.

Ten SNPs identified in *CYP2J2-2* intron 2 were analysed as a candidate gene SNP set. The single nucleotide diversity frequency was 15.95 SNP/kb in the Gouldian finch *CYP2J2-2* intron 2. Five haplotypes were identified in the pedigreed individuals, three of which were also present in additional individuals screened (Table 4.4), which indicates the stable co-segregation of these haplotypes. The genetic haplotypes were perfectly consistent with the phenotypes, except for two SNPs at sites 93 and 133 in one sample (LOD=2.71). These data suggest that the *CYP2J2-2* gene could be involved in determining the carotenoid-based head colour polymorphism in the Gouldian finch.

To sum up, the *CYP2J2* gene can be considered to be a candidate gene associated with carotenoid-based head colour polymorphism in the Gouldian finch; however, the position of the *Yellow* gene in Gouldian finch remains uncertain. To further test this hypothesis, or to identify an alternative locus, more pedigreed individuals need to be screened for the *CYP2J2* gene and included in the linkage analysis on chromosome 8 in the Gouldian finch.

CHAPTER FIVE: An extended study of melanin-based head colour plumage polymorphism in the Gouldian finch

Abstract

Variation in melanin pigmentation of the plumage in birds has received the attention of many biologists. The biosynthetic pathways and the functions of melanin pigmentation variation have been investigated in several different avian species, although the genetic basis for a number of melanin-based phenotypes has not yet been determined. The aim of this chapter is to review and update the information on the gene controlling the melanin-based head colour polymorphism in the Gouldian finch. A hypothesis has been proposed in a previous study that a polymorphism within a potential inversion, the *Red* core region on the Z chromosome, leads to differentiation between the red- and black-headed morphs in the Gouldian finch. Therefore, in this chapter, 22 loci in the *Red* core region and its flanking region have been amplified in 16 wild Gouldian finches using long-range PCR to investigate further the causal factor of this differentiation between morphs. A 3-kilobase insertion in the *Ego165* locus was identified; however, no break point or inversion within the *Red* core region was discovered. For further study, MiSeq sequencing was planned to discover the cause of the differentiation between the red- and black-headed morphs and the selection mechanism maintaining this polymorphism.

5.1 Introduction

Feathers are a distinguishing feature of birds. The pattern and colour of plumage varies remarkably between species, age classes, sexes and seasons. This is one of the reasons why birds have been extensively used as models in behavioural and ecological studies. The Gouldian finch (*Erythrura gouldiae*) is a colourful songbird with green back, blue and black tails, yellow abdomen, and a purple breast with a black outline and blue border (see Figure 1.1). But what has most interested biologists is the head colour polymorphism in wild Gouldian finch populations: with three different head colour morphs, red, black and yellow. Red and yellow heads are due to carotenoid-based colouration, while the black morph expresses eumelanin pigmentation. The morphs differ in the structure of the head feathers, as well as in pigment composition (Brush and Seifried, 1968). Black-headed feathers are covered with barbules on each barb, as well as containing heavy deposits of eumelanin; in contrast, red and yellow-headed feathers are flattened and lack barbules on the barbs in the areas of carotenoid deposition (Figure 5.1). A similar feather structure was found in the red and yellow morphs. However, the pigment composition in red and yellow is different. The type of carotenoid in red feathers is β-carotene, while the yellow-headed feathers contain α-carotene derived from dietary lutein (Figure 5.2).

Figure 5.1. The structure of cheek feathers from three colour morphs of the Gouldian finch. a, black-headed, barbs with barbules; b and c, barbs of red- and yellow-headed, respectively; no barbules on barbs in the latter two colour morphs (Brush and Seifried, 1968).

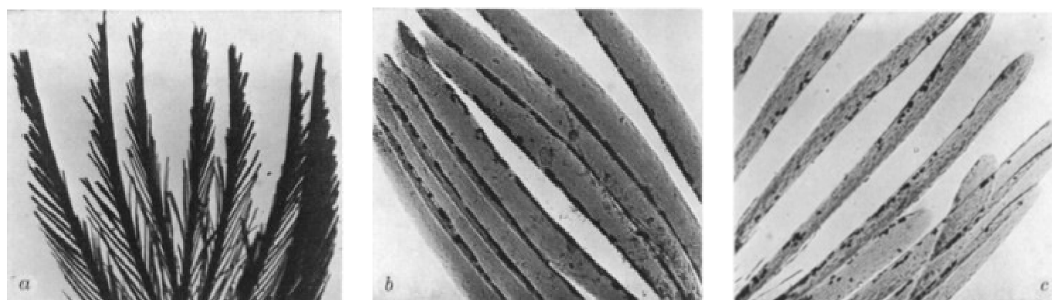
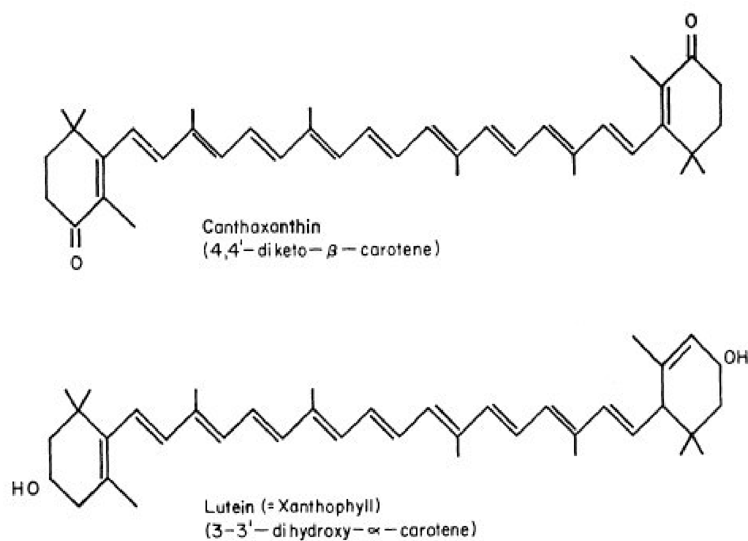


Figure 5.2. Structural formulae of carotenoid pigments partly responsible for colouration in the Gouldian finch. There are separate chemical structures for the α- and β-carotene backbones (Brush and Seifried, 1968).



5.1.1. Melanin-based colouration

Melanin is the most widely used pigment in animal integuments and tissues (Jimbow *et al.*, 1976). Melanin produces a variety of tones and patterns of black, grey and brown colours. There are two different types of melanin in the feathers of birds: phaeomelanin and eumelanin, which produce brown to buff and black to grey colours, respectively. Melanin occurs in all types of integumentary tissues (e.g. skin, scale, feathers) in birds. Melanin also plays many roles in biochemistry and physiology, such as antioxidant (Lozano, 1994), cation chelator (Riley, 1997), tissue strengthener (Riley, 1992), antimicrobial (Shawkey *et al.*, 2003) or photoprotectant (Ortonne, 2002). Melanins can be synthesized endogenously by birds in peripheral tissues. There are a variety of factors regulating the synthesis and accumulation of melanins. Besides season, sex, and integumentary tissue type, a key factor controlling the synthesis of melanin in birds is strong genetic control, such as by *Melanocortin-1 Receptor* (*MC1R*). *MC1R* has been shown to be involved in regulating colour polymorphism in the integument in chicken and other avian species (Takeuchi *et al.*, 1996; Theron *et al.*, 2001; Kerje *et al.*, 2003; Ling *et al.*, 2003).

5.1.2. Melanin-based head colour polymorphism in the Gouldian finch

As described above, the head-colour polymorphism exists in three distinct colour morphs in the Gouldian finch: red, black, and yellow. The *Yellow* locus controls the carotenoid-based head colour polymorphism, whereas, the *Red* locus is involved in the melanin-based polymorphism. These two genes are involved in an epistatic interaction codetermining the head and bill colour of Gouldian finch. The *Red* locus is sex-linked and located on the Z chromosome. Red colour is due to a dominant red allele (Z^R) while the black allele is recessive (Z^r) (Southern, 1945). The sex chromosome is considered to play an important role in the speciation process (Qvarnström and Bailey, 2008). In a series of papers, Pryke and

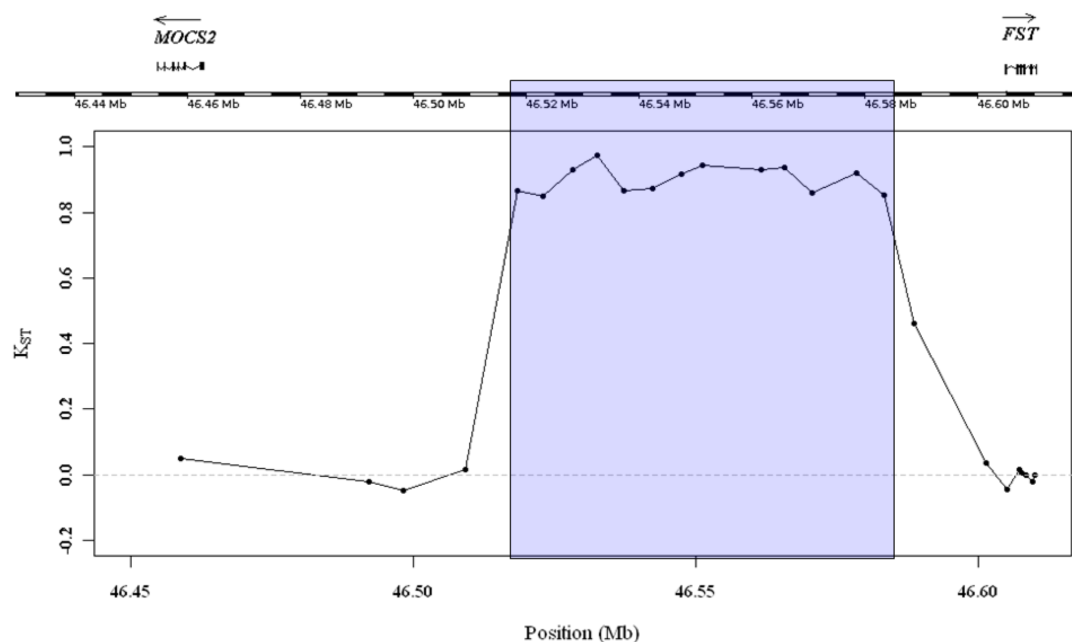
Griffith (2007, 2009a, 2009b) showed that the *Red* locus is associated with nonrandom mating, pre- and post-mating isolation, and incompatibility between the red and black morphs. As these are important fitness effects, restricted recombination may promote the adaptive evolution of the head colour morph because of the interactions between the *Red* gene and genes involved in genetic incompatibility. The position of the *Red* gene on the Z chromosome might have contributed to determining the local recombination rate and to increasing linkage disequilibrium between genes that contribute to the phenotype.

5.1.3. Recent research progress

Previous research identified the genomic region of the *Red* locus and a candidate gene that was significantly associated with melanin-based head colour polymorphism in the Gouldian finch. Kim (2011) demonstrated the region between 34.3 cM to 37.9 cM on the Z chromosome (mapped in the zebra finch genome) was associated with the *Red* locus, showing high LD by linkage mapping. This increased LD was possibly caused by an inversion within this region, providing a possible route of genetic differentiation by reducing the fitness of heterozygotes or gene flow (Rieseberg, 2001). Therefore, in the Gouldian finch, the association between the melanin-based plumage polymorphism and social behaviour and other characteristics seems likely to be due to local LD and, possibly, suppressed recombination (Thomas *et al.*, 2008; Huynh *et al.*, 2011). This candidate region has been narrowed down to 66 kb by RAD-seq, with highly significant LD (K.-W. Kim, unpublished data). This region was close to two genes: the molybdenum cofactor synthesis 2 *MOCS2* and *Follistatin FST,* but closer to the upstream region of *FST.* The *FST* gene is an antagonist of the *TGF-β* superfamily that regulates the development of hair follicles in mammals (McDowall *et al.*, 2008). Therefore, the dynamic expression of *FST* during feather development (Ohyama *et al.*, 2001), and its function in regulating the *TGF-β* superfamily, might both contribute to the head and bill colour polymorphism in the Gouldian finch. In

particular, the *Red* core region is 33.8-42.9 kb upstream of the *FST* gene. Within this region, there was a high level of sequence differentiation between the red and black morphs (Figure 5.3). Most of the loci sequenced within the region showed significant genetic differentiation between morphs. On the other hand, there was no significant difference in the flanking regions and in other genomic regions between the two morphs. Inversion may be an important mechanism for enabling genetic isolation (Rieseberg, 2001; Navarro and Barton, 2003). We therefore hypothesised that the *Red* core region might be inside an inversion. An alternative hypothesis, for which there is no direct evidence, is the introgression of a chromosomal fragment from another species (Kim, 2011).

Figure 5.3. Genetic differentiation between black and red morphs in the *Red* core region (measured by K_{ST}). The shaded region indicated the *Red* core region. The positions of the molybdenum cofactor synthesis 2 (*MOCS2*) and *Follistatin* (*FST*) genes are indicated at the top (Kim, 2011). Position indicates orthologous locations on zebra finch chromosome 8.



However, it remains unclear whether the genetic differentiation within the *Red* locus, and the evolution and maintenance of this plumage polymorphism, was caused by inversion, introgression or some other factor. In this chapter, I amplified the whole *Red* core region to attempt to identify any inversion breakpoints or other characteristics of this locus.

5.2. Materials and Methods

5.2.1. Samples

To test for the presence of an inversion in the *Red* core region, sixteen Gouldian finch blood samples were randomly selected from a wild population collected by Pryke and Griffith (2008) in Kimberley, Australia. These included eight red and eight black-headed females, females were chosen because in birds females are hemizygous for the Z chromosome, on which *Red* locus is located.

5.2.2. Amplification of the *Red* core region and extended region

A 66-kb region upstream of the *FST* gene was found to be associated with the melanin-based colour trait, and showed significant linkage disequilibrium (Kim, 2011). To amplify the whole *Red* core region and extended flanking regions, sixteen loci covering the *Red* core region and six flanking loci were selected for long-range PCR amplification. The long-range PCR primers (Table 5.1) were designed on Primer 3.0 online (Rozen and Skaletsky, 2000). All PCR primers were designed to amplify overlapping region to ensure that the amplified fragments are linearly connected. The PCR fragment size and high specificity required by long-range PCR were considered when adjusting the primer parameters to fit the appropriate PCR product size. The system used in long-range PCR amplification was 20-μl reactions

using QIAGEN LongRange PCR Kit, consisting of 2 µl LongRange PCR buffer (+Mg^{2+}), 1 µl dNTP mix (10 mM each), 1.6 µl each of forward and reverse primer (5 µM), 0.16 µl of Taq, 1 µl of DNA template (100ng/µl) and 12.6 µl of ultrapure water. The PCR reactions were completed using a DNA Engine Tetrad PTC-225 Peltier thermal cycler (MJ Research). The PCR program used for all the successful amplification was 3 min of initial denaturation at 93°C, then 38 cycles of 15 s at 93°C, 30 s at 56°C, and 15 min annealing at 68°C, followed by 60 s at 15°C. PCR amplification was checked by electrophoresing PCR products on a 0.7% agarose gel stained with SYBR safe, photographing using GeneSnap, and quantifying using a NanoDrop spectrophotometer and Fluostar OPTIMA-fluorometer.

Table 5.1. Primers for amplification and sequencing of upstream regions of *FST* (the *Red* core region). (a) Primers designed by K.-W. Kim (unpublished data). (b) Primers designed in this study using Primer 3.0. The positions indicate the distance from the end of PCR product to the first exon of *FST*. The minus number means the locus is located upstream of *FST*.

(a)

Loci	Target	Position (kb)	Forward primers (5'–3')	Reverse primers (5'–3')
Ego190	Intergenic	-106.3	CAGTGAGGGGACAGCAAATC	GCACTGCAGGAGTGTGGTAA
Ego191	Intergenic	-101.2	GCAGCATGTAAATCAGCATCA	CAGAACAATGCTGGGATCAA
Ego192	Intergenic	-95.2	TGGTTCACAAGCTGCATTTC	GGCACGTAAAAGGGTTTTTG
Ego193	Core	-91.0	CGCCCAATGAGTTATTGACA	CTCTGTGCCAGCCATATGAA
Ego173	Core	-81.6	CAACTTAGGACTGGAACAACTCA	CCCTCCACAAGAATGCAAAT
Ego195	Core	-76.2	CAGGCTGGACAAATCAGGTT	TGCCATGTGGGGTTTTTATT
Ego196	Core	-71.7	GCCTGCTATTGATTTTACATTCC	AATGCCTGAAAGCAGTTTGAA
Ego197	Core	-67.2	GGTGGGTGTCAGGCTGTATT	GCCATGCTCTGTCCCTAAAC
Ego198	Core	-62.1	TCGTGTACCTCCACCATTCA	TGCTGCTAGTTCCACGAGAG
Ego176	Core	-53.4	AACACTCAAGTATTGTGCAAGAAAA	CCCTGGAGGTTCCACAAGTA
Ego165	Core	-42.9	TCCTTCCTGCTGTCCTTGAC	TGGATGAGGCCTTCTACAGG
Ego166	Core	-38.9	ACCACCTTGTTCCCCAAATC	AGCTTCATTCCTCTGGGGTTA
Ego178	Core	-33.8	CGAGGGCAAGTCAGGTTATT	CCCTGATTTGTGGAGACACC
Ego181	Core	-21.2	TGTGATTTGAGGTGTTCCAAA	GCCACAAAAGAAATCTGTCCA
Ego183	Intergenic	-9.7	GCACAGGAACAAACTGCAAA	AAGCCAAATGCATTCAGTCC
Ego184	Intergenic	-3.2	CAGGAATTTGCTTGTTTTTCAC	GGCTGGCCTCTGAATAAAAT

(b)

Loci	Target	Position (kb)	Forward primers (5'–3')	Reverse primers (5'–3')
Ego194	Core	-85.9	TTGGTGGCTTAGTGGTGTATTG	ACTTCCACCAAAACAGATGTCT
Ego175	Core	-57.0	TGCTTTATTCTGGAGTTTGAACC	TCCACTCTTACCTCTGGACTTC
Ego177	Core	-46.9	GGGCTAGAAACATGGCAAAG	GTCATTGCAGCTGTGTTGGA
Ego179	Core	-29.6	AACACCACCCAAGAAAATGC	AAGAAACCTGTTTGATTTCTGGA
Ego180	Core	-25.6	GGAAGACCCTGGAGAAAAACTT	AAGGAGTTACCCAGATGTGCT
Ego182	Intergenic	-15.8	TACCTGGGCTTGTGACAACA	ACATGTACTTGTGGGATAGTTGA

5.2.3. Confirmation of the identity and size of amplified products

In order to confirm that the amplified products were target regions, the long-range PCR products were sequenced by 454 sequencing. The thermal cycling program using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) was 60 s at 96°C, then 30 cycles of 10 s at 94°C, 5 s at 50°C, 4 min at 60°C, followed by 5 min at 15°C. The BigDye v3.1 system included 1 µl BigDye v 3.1, 1.5 µl Seq-buffer, 0.32 µl forward primer (5 µM), 2 µl PCR product and 3.7 µl water for each reaction. Following precipitation using the ethanol/EDTA method, raw sequence data were generated on the ABI 3730 48-well capillary sequencer (Applied Biosystems). Sequence data were analysed by CodonCode Aligner v.2.0.6 (CodonCode Corporation).

5.3. Results

In this test, all 22 loci in the *Red* core region and flanking region were amplified successfully (Table 5.2). Although not every sample amplified at every locus, probably due to the limited available amount of DNA template or the intrinsic difficulty of long-range PCR, the fact that both black and red alleles amplified successfully in each case indicates that neither haplotype included a rearrangement, that is the possibility of an inversion can be excluded. Each black and red sample was sequenced and aligned to the zebra finch reference sequence to confirm the identity of the targeted fragment. A three-kb insertion in the *Ego165* locus (Table 5.3) was identified. But the fragment was found to be present in both haplotypes; therefore, it seems not to be related to the significant differentiation between the alleles. So far, there is no break point within the *Red* core region.

Table 5.2. Long-range PCR amplification of sixteen loci covering the *Red* core region, and six loci in the flanking region, in16 wild-caught Gouldian finch samples. NS: non-specific band.

Sample	Morph	Amplified core region markers							
		Ego193	Ego194	Ego173	Ego195	Ego196	Ego197	Ego198	Ego175
F101	Black	Yes	Yes	--	--	--	--	--	Yes
F102	Black	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F103	Black	Yes	Yes	Yes	Yes	Yes	Yes	Yes NS	Yes
F104	Black	Yes	--	--	Yes	Yes	Yes	Yes	Yes
F105	Black	Yes	Weak	Yes	Yes	Yes	Yes	Yes	Yes
F106	Black	Yes	Yes	Weak	Yes	Yes	Yes	Yes	Yes
F107	Black	Yes	Yes	--	--	--	--	--	Yes
F108	Black	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F121	Red	Yes	--	--	--	--	--	--	Yes
F125	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F140	Red	--	--	Yes	Yes	Yes	Yes	Yes	Yes
F146	Red	Yes	Yes	--	--	--	--	--	Yes
F168	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F169	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F170	Red	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
F171	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Sample	Morph	Amplified core region markers							
		Ego176	Ego177	Ego165	Ego166	Ego178	Ego179	Ego180	Ego181
F101	Black	--	Faint	--	--	--	Yes	--	--
F102	Black	Yes	Yes	Yes	Yes	Weak	Yes	Yes	Yes
F103	Black	Yes	Yes	Yes	Yes	Yes	Yes	--	Yes
F104	Black	Yes	Yes	Weak	No	Weak	--	--	Weak
F105	Black	Yes	Yes	Yes	Yes	Yes	Weak	Yes	Yes
F106	Black	Yes	Yes	Yes	Yes	Yes	Weak	--	Yes
F107	Black	--	Yes	--	--	--	Weak	Yes	--
F108	Black	Yes	Yes	Yes	Yes	Yes	--	Yes	Yes
F121	Red	--	--	--	--	--	--	Yes	--
F125	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F140	Red	Yes	--	Yes	Yes	Yes	Weak	Yes	Yes
F146	Red	--	Yes	--	--	--	Yes	Yes	--
F168	Red	Yes	Yes	Yes	Yes	Yes	--	Yes	Yes
F169	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F170	Red	Yes	Yes	Yes	Yes	Yes	--	Yes	Yes
F171	Red	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Sample	Morph	Amplified flanking region markers					
		Ego190	Ego191	Ego192	Ego182	Ego183	Ego184
F101	Black	--	NS	Yes	Yes NS	--	Weak
F102	Black	Weak	Yes NS	Yes	Yes NS	Yes	--
F103	Black	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F104	Black	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F105	Black	Yes	Yes NS	Yes	Yes NS	Weak	Yes
F106	Black	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F107	Black	Yes	Yes NS	Yes	Yes	Yes	Yes
F108	Black	Yes	Yes	Yes	Yes NS	Yes	Yes
F121	Red	Yes	Weak	Yes	Yes NS	--	Yes
F125	Red	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F140	Red	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F146	Red	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F168	Red	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F169	Red	Yes	Yes NS	Yes	Yes NS	Yes	Yes
F170	Red	Yes	Weak NS	Yes	Yes NS	Weak	Yes
F171	Red	Yes	Yes NS	Yes	Yes NS	Yes	Yes

Table 5.3. Summary of the amplified loci, comparing observed with expected amplicon sizes based on the zebra finch genome. Cases where they did not match are indicated in **bold**.

Locus	Target	Position (kb)	Amplification	Expected size (kb)	Observed size (kb)
Ego190	Intergenic	-106.3	Yes	5.9	5.9
Ego191	Intergenic	-101.2	Yes NS	5.1	5.1
Ego192	Intergenic	-95.2	Yes	6.0	6.0
Ego193	Core	-91.0	Yes	4.2	4.2
Ego194	Core	-85.9	Yes	5.1	5.1
Ego173	Core	-81.6	Yes	5.4	5.4
Ego195	Core	-76.2	Yes	6.6	6.6
Ego196	Core	-71.7	Yes	5.5	5.5
Ego197	Core	-67.2	Yes	5.5	5.5
Ego198	Core	-62.1	Yes	6.0	6.0
Ego175	Core	-57.0	Yes	6.0	6.0
Ego176	Core	-53.4	Yes	4.6	4.6
Ego177	Core	-46.9	Yes	7.9	7.9
Ego165	Core	-42.9	Yes	4.4	4.4
Ego166	Core	-38.9	Yes	**4.9**	**7.9**
Ego178	Core	-33.8	Yes	5.9	5.9
Ego179	Core	-29.6	Yes	4.5	4.5
Ego180	Core	-25.6	Yes	4.0	4.0
Ego181	Core	-21.2	Yes	5.8	5.8
Ego182	Intergenic	-15.8	Yes NS	5.4	5.4
Ego183	Intergenic	-9.7	Yes	6.1	6.1
Ego184	Intergenic	-3.2	Yes	6.5	6.5

5.4. Discussion

Long-range PCR is an efficient and economical method to amplify a relatively long region. However, as the target fragment gets larger, several factors need extra consideration. First, to run a long-range PCR, highly-fidelity DNA polymerases are an important requirement to increase the rate of target amplification. Second, long-range PCR requires a high quality and quantity of DNA template. Here, the concentration of the DNA samples was 5-10 times higher (100 ng/μl) than in normal PCR (10-20 ng/μl). The long extension time (15 min) is also important, as insufficient time for extension could lead to the incomplete amplification of the target fragment.

A previous study showed a high level of fixed polymorphism and linkage disequilibrium in the core region between loci *Ego194* and *Ego181*. Here, we tested for the presence or absence of inversion break points between these markers. Using long-range PCR amplification, we were able to fill all the gaps around the *Red* core region extending to the flanking region for both morphs, to discover the potential inversion that may cause the reduced recombination and the significant linkage disequilibrium. If there had been a paracentric inversion, then the corresponding region could not have been amplified within the affected haplotype using the 5'-3' primers that we designed on the assumption that the sequence was contiguous. However, no inversion was detected in the core region.

Although no inversion was found in this assay, the PCR products could be used for further study, i.e. sequencing, to investigate the origins of the differentiation, and identify the selection mechanism maintaining this polymorphism.

CHAPTER SIX: General discussion

6.1. Carotenoid-based colouration

The carotenoid-based polymorphism in the Gouldian finch is probably controlled by a locus that controls the conversion of dietary yellow to derived red ketocarotenoids, though it might alternatively control a pathway that deposits different carotenoids in the plumage and bill. Previous genetic studies on plumage colour polymorphism in the Gouldian finch have been focused on the control of melanin-based colouration. Because the candidate gene related to melanin-based polymorphism is located on the Z chromosome, before this study only Z chromosome genetic markers and sporadic autosomal markers were available (Kim, 2011).

In this thesis, 81 new microsatellite markers distributed across the autosomal chromosomes, and nearly 64,000 unique SNPs discovered by RAD sequencing, were developed as valuable genetic resources for further studies of population genetics, linkage mapping, and association analysis in the Gouldian finch. These microsatellite markers are also now available for morphological and behavioural studies, such as for mapping the several plumage colour patterning mutations maintained in aviarists' collections of Gouldian finches (see 100 Gouldian finch mutations and modifications on www.gouldianfinch.gr) and for studying the mating system in this and related species.

Additionally, I present a linkage map of the Gouldian finch from which the microsatellite markers were designed, based on sequence similarity between Gouldian finch and zebra finch orthologues. This first autosomal genomic linkage map of Gouldian finch included 42 microsatellite markers in total, representing 7 linkage groups. The linkage map

covered a total distance of 447.2 cM. Even though the density of the map is relatively low, it is still potentially valuable for mapping phenotypic loci, as demonstrated in previous mapping research (Miwa *et al.,* 2006). This map represents an initial step for further mapping and is a potential resource for locating genes underlying important functional fitness traits in the Gouldian finch. In this study, no interchromosomal rearrangements were detected between Gouldian finch and zebra finch microsatellite loci. This result is consistent with the high degree of synteny and relatively conserved marker order among avian species (Stapley *et al.*, 2008).

It has been shown that the genetic control of the head morph variation, consisting of three distinct morphs, is attributable to an epistatic two-locus model (Southern, 1946; Murray, 1963), which involves one melanin-based and one carotenoid-based locus; these genes might be regulatory, e.g. regulating the expression of a gene in the integumentary tissue (Walsh *et al.*, 2011), or structural, e.g. coding for a carotenoid-binding protein, as has been demonstrated in invertebrate species (Keen *et al.*, 1991). I attempted to map the position of the carotenoid-based plumage colour locus in the Gouldian finch through linkage mapping, RAD-sequencing and a candidate-gene approach. There was no statistically significant association between the specific trait and any SNP, probably due to the relatively low density of the map and the small informative population size. However, there was a region on chromosome 8 around the 25-Mb position showed notable F_{ST} and P-value that should be considered as a potential candidate region. Additionally, the SNPs used in the GWAS and discovered using RAD tags are a valuable resource for further investigating plumage and behaviour polymorphisms in the Gouldian finch.

According to the evidence from chapter 3 and candidate *Yellow* gene tested in zebra finch, by chromosome 8 linkage mapping, I identified the possible chromosome location of *Yellow* in the Gouldian finch, a 2.9-cM region from 20.2 cM to 23.1 cM on chromosome 8.

Due to the limited size of the study population and the complexity of the plumage colour variation in this species, it was not possible to define the location of the gene more precisely or to identify the causal gene. However, the study agreed that the carotenoid-based polymorphism is determined by a recessive autosomal allele, distinct from the melanin-determining polymorphic *Red* locus located in the upstream region of *follistatin* (*FST*).

Further study is needed to identify the causative region and gene. This might be achieved by: (1) Candidate region fine-mapping in a larger pedigree, and/or (2) A higher density genome-wide association study in a larger sample of wild birds.

6.2. Melanin-based colouration

Melanins are a major and abundant class of pigments in animals; in this case they are synthesized *in vivo*. Melanins play many roles in biochemistry and physiology, from being antioxidants (Lozano, 1994), cation chelators (Riley, 1997), to affect the strength of feathers (Riley, 1992), antimicrobial deterrents (Shawkey *et al.*, 2003), and photoprotectants (Ortonne, 2002). Melanins can be synthesized endogenously by birds in their peripheral tissues (Mason and Frost-Mason, 2000). A variety of factors regulate the synthesis and accumulation of melanins besides season, sex, and integumentary forms. Tyrosinase activity, and thus melanin synthesis, is under strong genetic control (Takeuchi *et al.,* 1996) in birds.

Recently, the genetic control of melanin pigmentation in avian plumage has received the attention of many biologists (Takeuchi *et al.*, 1996; Ito *et al.*, 2000; Hill and McGraw, 2006). The biosynthetic pathways and functions of melanin have also been demonstrated in birds (McGraw, 2008). Kim (2011) identified a genomic region of the *Red* locus that was significantly associated with the melanin-based head colour polymorphism in the Gouldian

finch. However, it is unknown yet how the genetic differentiation within the *Red* locus arose and how this locus affects the expression of the genes involved in melanin deposition.

Therefore, I amplified and sequenced 22 loci in the *Red* core and flanking regions, using long-range PCR, in order to further characterise the causal basis of the differentiation between the red black morphs as identified in previous research (Kim, 2011). In particular, this work aimed to discover if the high degree of differentiation observed between the red and black haplotypes was due to an inversion, as has been seen following reduced recombination in a 700-kb inversion polymorphism in *Heliconius* butterflies (Merrill *et al.*, 2011). However, no break point or inversion that included the *Red* core region was detected, which suggests that the strong differentiation described previously (Kim, 2011) is maintained by selection, rather than by an inversion. Further study of the origins and selection mechanisms maintaining this highly differentiated polymorphism will require in-depth new-generation sequencing of the region.

REFERENCES

Allen, P. C. (1987) Physiological responses of chicken gut tissue to coccidial infection: comparative effects of *Eimeria acervulina* and *Eimeria mitis* on mucosal mass, carotenoid content, and brush border enzyme activity. Poultry Science **66**: 1306–1315.

Amores, A., J. M. Catchen, A. Ferrara, Q. Fontenot, and J. H. Postlethwait. (2011) Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. Genetics **188**: 799–808.

Andersson, S., M. Prager, and E. I. A. Johansson. (2007) Carotenoid content and reflectance of yellow and red nuptial plumages in widowbirds (*Euplectes spp.*). Functional Ecology **21**: 272–281.

Arnason, E. (1978) Apostatic selection and kleptoparasitism in the parasitic jaeger. The Auk **95**: 377–381.

Backström, N., M. Brandström, L. Gustafsson, A. Qvarnström, H. Cheng, and H. Ellegren. (2006) Genetic mapping in a natural population of collared flycatchers (*Ficedula albicollis*): conserved synteny but gene order rearrangements on the avian Z chromosome. Genetics **174**: 377–386.

Backström, N., N. Karaiskou, E. H. Leder, L. Gustafsson, C. R. Primmer, A. Qvarnström, and H. Ellegren. (2008) A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of avian evolution. Genetics **179**: 1479–1495.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One **3**: e3376.

Bartle, J. A., and P. M. Sagar. (1987) Intraspecific variation in the New Zealand bellbird

Anthornis melanura. Journal of the Ornithological Society of New Zealand **34**: 253–306.

Bauernfeind, J. C. (1981) Carotenoids as colorants and vitamin a precursors: technological and nutritional applications. Academic Press, New York.

Beebe, C. W. (1906) The Bird: Its Form and Function. Henry Holt and Company, New York.

Bell, P. J. (1996) Survey of the nasal mite fauna (Rhinonyssidae and Kytoditidae) of the Gouldian finch, *Erythrura gouldiae*, and some co-occurring birds in the Northern Territory. Wildlife Research **23**: 675–685.

Benjamini, Y., and Y. Hochberg. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B **57**: 289–300.

Bennett, J. H. (1966) Head colour polymorphism in the Gouldian finch. Heredity 534.

Berger, J., T. Suzuki, K. A. Senti, J. Stubbs, G. Schaffner, and B. J. Dickson. (2001) Genetic mapping with SNP markers in Drosophila. Nature Genetics **29**: 475–481.

Berlin, D. S., K. Sangkuhl, T. E. Klein, and R. B. Altman. (2010) PharmGKB summary: cytochrome P450, family 2, subfamily J, polypeptide 2: CYP2J2. Pharmacogenetics and genomics **21**: 308–311.

Blount, J. D., D. C. Houston, and A. P. Møller. (2000) Why egg yolk is yellow. Trends in Ecology and Evolution **15**: 47–49.

Blount, J. D., N. B. Metcalfe, T. R. Birkhead, and P. F. Surai. (2003) Carotenoid modulation of immune function and sexual attractiveness in zebra finches. Science **300**: 125–127.

Blount, J. D., and K. J. McGraw. (2008) Signal functions of carotenoid coloration. In Carotenoids. Volume 4. Edited by Britton G, Liaaen-Jensen S. Basel: Birkhauser, 213–236.

Brawner, W. R., III, G. E. Hill, and C. A. Sundermann. (2000) Effects of coccidial and mycoplasmal infections on carotenoid-based plumage pigmentation in male house

finches. The Auk **117**: 952–963.

Brazill-Boast, J., Dessman, J. K., Davies, G. T. O., Pryke, S. R., and S. C. Griffith. (2011) Selection of breeding habitat by the endangered Gouldian finch (*Erythrura gouldiae*) at two spatial scales. Emu. **111**: 304–311.

Bruford, M. W., O. Hanotte, J. F. Y. Brookfield, and T. Burke (1998) Multi- and single-locus DNA fingerprinting. In: Molecular Analysis of Populations: A Practical Approach, 2nd edn (ed. Hoelzel AR), 287–336. IRL Press, Oxford.

Brush, A. H., and H. Seifried. (1968) Pigmentation and feather structure in genetic variants of the Gouldian finch, *Poephila gouldiae*. The Auk **85**: 416–430.

Brush, A. H., and D. M. Power. (1976) House finch pigmentation: Carotenoid metabolism and the effect of diet. The Auk **93**: 725–739.

Brush, A. H. (1990) Metabolism of carotenoid pigments in birds. The FASEB Journal **4**: 2969–2977.

Burkley, P. A. (1987) Mendelian genes. In F. Cooke and P. A. Buckley, eds., Avian Genetics, 1–44. Academic Press, London.

Burt, D. W., C. Bruley, I. C. Dunn, C. T. Jones, A. Ramage, A. S. Law, D. R. Morrice, I. R. Paton, J. Smith, D. Windsor, A. Sazanov, R. Fries, and D. Waddington. (1999) The dynamics of chromosome evolution in birds and mammals. Nature **402**: 411–413.

Butcher, G. S., and S. Rohwer. (1988) The evolution of conspicuous and distinctive coloration for communication in birds. Current Ornithology **6**: 51–108.

Capdevila, J. H., D. C. Zeldin, K. Makita, A. Karara, and J. R. Falck. (1995) Cytochrome P450 and the metabolism of arachidonic acid and oxygenated eicosanoids. In ‘‘Cytochrome P450: Structure, Mechanism, and Biochemistry’’ (P.R. Ortiz de Montellano, Ed.), 317–339. JAI Press, London.

Catchen, J., A. Amores, P. Hohenlohe, W. Cresko, and J. Postlethwait. (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics **1**:171–182.

Catchen, J., P. Hohenlohe, S. Bassham, A. Amores, and W. Cresko. (2013) Stacks: an analysis tool set for population genomics. Molecular Ecology **22**: 3124–3140.

Chew, B. (1993) Role of carotenoids in the immune response. Journal of Dairy Science **76**: 2804–2811.

Clark, A. G. (2004) The role of haplotypes in candidate gene studies. Genetic Epidemiology. **27**: 321–333.

Clarke, B. (1962) Natural selection in mixed populations of two polymorphic snails. Heredity **17**: 319–345.

Clarke, G. M., C. A. Anderson, F. H. Pettersson, L. R. Cardon, A. P. Morris, and K. T. Zondervan. (2011) Basic statistical analysis in genetic case-control studies. Nature Protocols **6**: 121–133.

Davey, J. W., P. A. Hohenlohe, P. D. Etter, J. Q. Boone, J. M. Catchen, and M. L. Blaxter. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Review Genetics **12**: 499–510.

Dawson, D. A., T. Burke, B. Hansson, J. Pandhal, M. C. Hale, G. H. Hinten, and J. Slate. (2006) A predicted microsatellite map of the passerine genome based on chicken–passerine sequence similarity. Molecular Ecology **15**: 1299–1320.

Dawson, D. A., M. Åkesson, T. Burke, J. M. Pemberton, J. Slate, and B. Hansson. (2007) Gene order and recombination rate in homologous chromosome regions of the chicken and a passerine bird. Molecular Biology and Evolution **24**: 1537–1552.

del Val, E., J. C. Senar, J. Garrido-Fernández, M. Jarén, A. Borràs, J. Cabrera, and J. J. Negro. (2009) The liver but not the skin is the site for conversion of a red carotenoid in a passerine bird. Naturwissenschaften **96**: 797–801.

Derjusheva, S., A. Kurganova, F. Habermann, and E. Gaginskaya. (2004) High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. Chromosome Research **12**: 715–723.

Dostine, P. L. (1998) Gouldian Finch Recovery Plan *Erythrura gouldiae*. Parks and Wildlife

Commission NT and Environment Australia, Darwin.

Dostine, P. L., G. C. Johnson, D. C. Franklin, Y. Zhang, and C. Hempel. (2001) Seasonal use of savanna landscapes by the Gouldian finch, *Erythrura gouldiae*, in the Yinberrie Hills area, Northern Territory. Wildlife Research **28**: 445–458.

Dostine, P. L., and D. C. Franklin. (2002) A comparison of the diet of three finch species in the Yinberrie Hills area, Northern Territory. Emu **102**: 159–164.

Duerr, R. H., K. D. Taylor, S. R. Brant, J. D. Rioux, M. S. Silverberg, M. J. Daly, A. H. Steinhart, C. Abraham, M. Regueiro, A. Griffiths, T. Dassopoulos, A. Bitton, H. Yang, S. Targan, L. W. Datta, E. O. Kistner, L. P. Schumm, A. T. Lee, P. K. Gregersen, M. M. Barmada, J. I. Rotter, D. L. Nicolae, and J. H. Cho. (2006) A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. Science **314**:1461–1463.

Durrer, H. (1986) The skin of birds: colouration. In J. Bereiter-Hahn, A. G. Matoltsky, and K. S. Richards, eds., Biology of the Integument 2: Vertebrates, 239–247. Springer-Verlag, Berlin.

Ekblom, R., and J. Galindo. (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity **107**: 1–15.

Erdman, J. W., JR., T. L. Bierer, and E. T. Gugger. (1993) Absorption and transport of carotenoids. Annals of the New York Academy of Sciences **691**: 76–85.

Eriksson, J., G. Larson, U. Gunnarsson, B. Bed'hom, M. Tixier-Boichard, L. Strömstedt, D. Wright, A. Jungerius, A. Vereijken, E. Randi, P. Jensen, and L. Andersson. (2008) Identification of the *Yellow Skin* gene reveals a hybrid origin of the domestic chicken. PLoS Genetics **4**:e1000010.

Ewing, B., and P. Green (1998) Base-calling of automated sequencer traces using *Phred. II.* Error probabilities. Genome Research **8**: 186–194.

Fillon, V., M. Vignoles, R. P. Crooijmans, M. A. Groenen, R. Zoorob, and A. Vignal. (2007) FISH mapping of 57 BAC clones reveals strong conservation of synteny between

Galliformes and Anseriformes. Animal Genetics **38**: 303–307.

Finger, E., D. Burkhardt, and J. Dyck. (1992) Avian plumage colors: Origin of UV reflection in a black parrot. Naturwissenschaften **79**: 187–188.

Forshaw, J. M. (1989) Parrots of the World. Lansdowne Editions, Melbourne, Australia.

Fox, D. L. (1962) Metabolic fractionation, storage and display of carotenoid pigments by flamingos. Comparative Biochemistry and Physiology **6**: 1–40.

Fox, D. L. (1976) Animal Biochromes and Structural Colours. University of California Press, Berkeley.

Fox, H. M., and G. Vevers. (1960) The Nature of Animal Colours. Sidgwick and Jackson, London.

Fox, S., R. Brooks, M. J. Lewis, and C. N. Johnson. (2002) Polymorphism, mate choice and sexual selection in the Gouldian finch (*Erythrura gouldiae*). Journal of Zoology (London) **50**: 125–134.

Franklin, D. C. (1999) Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement. Biological Conservation **90**: 53–68.

Franklin, D. C., A. H. Burbidge, and P. L. Dostine. (1999) The harvest of wild birds for aviculture: an historical perspective on finch trapping in the Kimberley with special emphasis on the Gouldian finch. Australian Zoologist **31**: 92–109.

Franklin, D. C., and P. L. Dostine. (2000) A note on the frequency and genetics of head colour morphs in the Gouldian finch. Emu **100**: 236–239.

Franklin, D. C., P. J. Whitehead, G. Pardon, J. Matthews, P. McMahon, and D. McIntyre (2005) Geographic patterns and correlates of the decline of granivorous birds in northern Australia. Wildlife Research **32**: 399–408.

Freedman, R. S., E. Wang, S. Voiculescu, R. Patenia, R. L. Bassett, M. Deavers, F. M. Marincola, P. Yang, and R. A. Newman. (2007) Comparative analysis of peritoneum and tumor eicosanoids and pathways in advanced ovarian cancer. Clinical Cancer

Research **13**: 5736–5744.

Galeotti, P., and C. Cesaris. (1996) Rufous and grey colour morphs in the Italian tawny owl: geographical and environmental influences. Journal of Avian Biology **27**: 15–20.

Galeotti, P., D. Rubolini, P. O. Dunn, and M. Fasola. (2003) Colour polymorphism in birds: causes and functions. Journal of Evolutionary Biology **16**: 635–646.

Garnett, S. T. and G. M. Crowley. (2000) The action plan for Australian birds. Environment Australia, Canberra.

Gheyas, A. A., and D. W. Burt. (2013) Microarray resources for genetic and genomic studies in chicken: a review. Genesis **51**: 337–356.

Gilby, A. J., S. R. Pryke, and S. C. Griffith. (2009) The historical frequency of head-colour morphs in the Gouldian finch (*Erythrura gouldiae*). Emu **109**: 222–229.

Gill, D. E., and W. E. Lanyon. (1965) Establishment, growth, and behaviour of an extralimital population of house finches at Huntington, New York. Bird Banding **36**: 1–14.

Gliszczyńska-Świglo, A., and A. Koziolowa. (2000) Chromatographic determination of riboflavin and its derivatives in food. Journal of Chromatography A **881**: 285–297.

Goodfellow, D. L. (2001) Birds of Australia's top end. Reed New Holland, Sydney.

Goodwin, T. W. (1984) The Biochemistry of the Carotenoids. Volume II: Animals. Chapman and Hall, London.

Goodwin, T. W. (1986) Metabolism, nutrition and function of carotenoids. Annual Review of Nutrition **6**: 273–297.

Green, P. H. (1990) Documentation for CRIMAP, version 2.4.

Griffin, D. K., L. B. Robertson, H. G. Tempest, and B. M. Skinner. (2007) The evolution of the avian genome as revealed by comparative molecular cytogenetics. Cytogenetic and Genome Research **117**: 64–77.

Groenen, M. A. M., R. P. M. A. Crooijmans, A. Veenendaal, H. H. Cheng, M. Siwek, and J. J. van der Poel. (1998) A comprehensive microsatellite linkage map of the chicken

genome. Genomics **49**: 265–274.

Gustafsson, L., and W. J. Sutherland. (1988) The cost of reproduction in the collared flycatcher *Ficedula albicollis*. Nature **335**: 813–815.

Guttenbach, M., I. Nanda, W. Feichtinger, J. S. Masabanda, D. K. Griffin, and M. Schmid. (2003) Comparative chromosome painting of chicken autosomal paints 1–9 in nine different bird species. Cytogenetic and Genome Research **103**: 173–184.

Haiman, C. A., G. K. Chen, W. J. Blot, S. S. Strom, S. I. Berndt, R. A. Kittles, B. A. Rybicki, W. B. Isaacs, S. A. Ingles, J. L. Stanford, W. R. Diver, J. S. Witte, A. W. Hsing, B. Nemesure, T. R. Rebbeck, K. A. Cooney, J. Xu, A. S. Kibel, J. J. Hu, E. M. John, S. M. Gueye, S. Watya, L. B. Signorello, R. B. Hayes, Z. Wang, E. Yeboah, Y. Tettey, Q. Cai, S. Kolb, E. A. Ostrander, C. Zeigler-Johnson, Y. Yamamura, C. Neslund-Dudas, J. Haslag-Minoff, W. Wu, V. Thomas, G. O. Allen, A. Murphy, B. Chang, S. L. Zheng, M. C. Leske, S. Wu, A. M. Ray, A. J. M. Hennis, M. J. Thun, J. Carpten, G. Casey, E. N. Carter, E. R. Duarte, L. Y. Xia, X. Sheng, P. Wan, L. C. Pooler, I. Cheng, K. R. Monroe, F. Schumacher, L. L. Marchand, L. N. Kolonel, S. J. Chanock, D. V. D. Berg, D. O. Stram, and B. E. Henderson. (2011) Genome-wide association study of prostate cancer in men of African ancestry identifies a susceptibility locus at 17q21. Nature Genetics **43**: 570–573.

Hammershøj, M., U. Kidmose, and S. Steenfeldt. (2010) Deposition of carotenoids in egg yolk by short-term supplement of coloured carrot (*Daucus carota*) varieties as forage material for egg-laying hens. Journal of the Science of Food and Agriculture **90**: 1163–1171.

Hansson, B., M. Åkesson, J. Slate, and J. M. Pemberton. (2005) Linkage mapping reveals sex-dimorphic map distances in a passerine bird. Proceedings of the Royal Society B **272**: 2289–2298.

Hansson, B., M. Ljungqvist, D. A. Dawson, J. C. Mueller, J. Olano-Marin, H. Ellegren, and J.-A. Nilsson. (2010) Avian genome evolution: insights from a linkage map of the

blue tit (*Cyanistes caeruleus*). Heredity **104**: 67–78.

Hartl, D., and A. Clark. (1997) Principles of Population Genetics. Sinauer, Sunderland, MA.

Higgins, P. J., J. M. Peter, and S. J. Cowling, eds. (2006) Boatbill to Starlings. In: Handbook of Australian, New Zealand and Antarctic Birds. 7. Oxford University Press, Melbourne.

Hill, G. E. (1990) Female house finches prefer colourful males: sexual selection for a condition-dependent trait. Animal Behaviour **40**: 563–572.

Hill, G. E. (1991) Plumage coloration is a sexually selected indicator of male quality. Nature **350**: 337–339.

Hill, G. E. (1992) Proximate basis of variation in carotenoid pigmentation in male house finches. The Auk **109**: 1–12.

Hill, G. E. (1993) Male mate choice and the evolution of female plumage coloration in the house finch. Evolution **47**: 1515–1525.

Hill, G. E., R. Montgomerie, C. Y. Inouye, and J. Dale. (1994) Influence of dietary carotenoids on plasma and plumage colour in the house finch: Intra- and intersexual variation. Functional Ecology **8**: 343–350.

Hill, G. E. (1999) Mate choice, male quality, and carotenoid-based plumage coloration. Proceedings of the International Ornithological Congress **22**: 1654–1668.

Hill, G. E. (2002) A Red Bird in a Brown Bag: The Function and Evolution of Colourful Plumage in the House Finch. Oxford University Press, New York.

Hill, G. E., and K. J. McGraw. (2004) Correlated changes in male plumage colouration and female mate choice in cardueline finches. Animal Behaviour **67**: 27–35.

Hill, G. E., and K. J. McGraw. (2006) Bird Coloration, Volume 1: Mechanisms and Measurements. Harvard University Press, Cambridge, MA.

Hoekstra, H. E., and M. W. Nachman. (2003) Different genes underlie adaptive melanism in different populations of rock pocket mice. Molecular Ecology **12**: 1185–1194.

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko.

(2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genetics. **6**: e1000862.

Holleley, C. E., and P. G. Geerts. (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. BioTechniques **7**: 511–517.

Hubbard, J. K., J. A. C. Uy, M. E. Hauber, H. E. Hoekstra, and R. J. Safran. (2010) Vertebrate pigmentation: from underlying genes to adaptive function. Trends in Genetics **26**: 231–239.

Hughes, D. A. (2001) Dietary carotenoids and human immune function. Nutrition **17**: 823–827.

Huynh, L. Y., D. L. Maney, and L. W. Thomas. (2011) Chromosome-wide linkage disequilibrium caused by an inversion polymorphism in the white-throated sparrow (*Zonotrichia albicollis*). Heredity **106**: 537–546.

Ikram, M. K., S. Xueling, R. A. Jensen, M. F. Cotch, A. W. Hewitt, M. A. Ikram, J. J. Wang, R. Klein, B. E. K. Klein, M. M. B. Breteler, N. Cheung, G. Liew, P. Mitchell, A. G. Uitterlinden, F. Rivadeneira, A. Hofman, P. T. V. M. de Jong, C. M. van Duijn, L. Kao, C. Cheng, A. V. Smith, N. L. Glazer, T. Lumley, B. McKnight, B. M. Psaty, F. Jonasson, G. Eiriksdottir, T. Aspelund, G. B. Consortium, T. B. Harris, L. J. Launer, K. D. Taylor, X. Li, S. K. Iyengar, Q. Xi, T. A. Sivakumaran, D. A. Mackey, S. MacGregor, N. G. Martin, T. L. Young, J. C. Bis, K. L. Wiggins, S. R. Heckbert, C. J. Hammond, T. Andrew, S. Fahy, J. Attia, E. G. Holliday, R. J. Scott, F. M. A. Islam, J. I. Rotter, A. K. McAuley, E. Boerwinkle, E. S. Tai, V. Gudnason, D. S. Siscovick, J. R. Vingerling, and T. Y. Wong. (2010) Four Novel Loci (19q13, 6q24, 12q24, and 5q14) Influence the microcirculation *in vivo*. PLoS Genetics **6**: e1001184.

Immelmann, K. (1982) Australian Finches in Bush and Aviary. Angus and Robertson, Sydney.

Inouye, C. Y., G. E. Hill, R. D. Stradi, and R. Montgomerie. (2001) Carotenoid pigments in male house finch plumage in relation to age, subspecies, and ornamental coloration.

The Auk **118**: 900–915.

International Chicken Genome Sequencing Consortium. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature **432**: 695–716.

Irwin, D. E., S. Bensch, and T. D. Price. (2001) Speciation in a ring. Nature **409**: 333–337.

Jaari, S., M. H. Li, and J. Merilä. (2009) A first-generation microsatellite-based genetic linkage map of the Siberian jay (*Perisoreus infaustus*): insights into avian genome evolution. BMC Genomics **10**: 1.

Jiang, J. G., Y. G. Ning, C. Chen, D. Ma, Z. J. Liu, S. Yang, J. Zhou, X. Xiao, X. A. Zhang, M. L. Edin, J. W. Card, J. Wang, D. C. Zeldin, and D. W. Wang. (2007) Cytochrome p450 epoxygenase promotes human cancer metastasis. Cancer Research **67**: 6665–6674.

Jimbow, K., W. C. Quevedo Jr, T. B. Fitzpatrick, and G. Szabo. (1976) Some aspects of melanin biology: 1950-1975. The Journal of Investigative Dermatology **67**: 72–89.

Johnston, S. E., J. Gratten, C. Berenos, J. G. Pilkington, T. H. Clutton-Brock, J. M. Pemberton, and J. Slate. (2013) Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature **502**: 93–95.

Kalinowski, S. T., M. L. Taper, and T. C. Marshall. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology **16**: 1099–1106.

Kayang, B. B., A. Vignal, M. Inoue-Murayama, M. Miwa, J. L. Monvoisin, S. Ito, and F. Minvielle. (2004) A first-generation microsatellite linkage map of the Japanese quail. Animal Genetics **35**: 195–200.

Keen, J. N., I. Caceres, E. E. Eliopoulos, P. F. Zagalsky, and J. B. C. Findlay. (1991) Complete sequence and model for the A2 subunit of the carotenoid pigment complex, crustacyanin. European Journal of Biochemistry **202**: 31–40.

Kelley, D., M. Schatz, and S. Salzberg (2010) Quake: quality-aware detection and correction

of sequencing errors. Genome Biology **11**: R116.

Kenta, T., J. Gratten, G. Hinten, J. Slate, R. K. Butlin, and T. Burke. (2008) Multiplex SNP-SCALE: a cost-effective medium-throughput SNP genotyping method. Molecular Ecology Resources **8**: 1230–1238.

Kerje, S., J. Lind, K. Schütz, P. Jensen, and L. Andersson. (2003) Melanocortin 1-receptor (*MC1R*) mutations are associated with plumage colour in chicken. Animal Genetics **34**: 241–248.

Khachik, F., F. F. de Moura, D. Y. Zhao, C. P. Aebischer, and P. S. Bernstein. (2002) Transformations of selected carotenoids in plasma, liver, and ocular tissues of humans and in nonprimate animal models. Investigative Ophthalmology and Visual Science **43**: 3383–3392.

Kiefer, C., S. Hessel, J. M. Lampert, K. Vogt, M. O. Lederer, D. E. Breithaupt, and J. von Lintig. (2001) Identification and characterization of a mammalian enzyme catalyzing the asymmetric oxidative cleavage of provitamin A. Journal of Biological Chemistry **276**: 14110–14116.

Kim, K. W. (2011) The molecular genetic basis of a plumage colour polymorphism in the Gouldian finch (*Erythrura gouldiae*). Ph.D. Thesis. University of Sheffield: U.K.

Kim, K.-W., H. Zhang, G. J. Horsburgh, S. Bird, D. A. Dawson, S. C. Griffith, and T. Burke. (2015) Four-way development of microsatellite markers for the Gouldian finch (*Erythrura gouldiae*). Conservation Genetics Resources, In press.

King, L. M., J. Ma, S. Srettabunjong, J. Graves, J. A. Bradbury, L. Li, M. Spiecker, J. K. Liao, H. Mohrenweiser, and D. C. Zeldin. (2002) Cloning of *CYP2J2* gene and identification of functional polymorphisms. Molecular Pharmacology **61**: 840–852.

King, A. J., L. J. Williams, and C. Mettke-Hofmann. (2015) The effects of social conformity on Gouldian finch personality. Animal Behaviour **99**: 25–31.

Klein, R. J., C. Zeiss, E. Y. Chew, J. Y. Tsai, R. S. Sackler, C. Haynes, A. K. Henning, J. P. SanGiovanni, S. M. Mane, S. T. Mayne, M. B. Bracken, F. L. Ferris, J. Ott, C.

Barnstable, and J. Hoh. (2005) Complement Factor H Polymorphism in Age-Related Macular Degeneration. Science **308**: 385–389.

Kofler, R., C. Schlotterer, and T. Lelley. (2007) SciRoKo: a new tool for whole genome microsatellite search and investigation. Bioinformatics **23**: 1683–1685.

Komdeur, J. (1992) Importance of habitat saturation and territory quality for the evolution of cooperative breeding in the Seychelles warbler. Nature **358**: 493–495.

Krinsky, N. I. (1993) Actions of carotenoids in biological systems. Annual Review of Nutrition **13**: 561–587.

Lander, E. S., and P. Green. (1987) Construction of multilocus genetic linkage maps in humans. Proceedings of the National Academy of Sciences of the USA **84**: 2363–2367.

Landry, G. P. (1997) The Varieties and Genetics of the Zebra Finch. Franklin, Louisiana.

Lank, D. B. (2002) Diverse processes maintain plumage polymorphisms in birds. Journal of Avian Biology **33**: 327–330.

Leskinen, P. K., T. Laaksonen, S. Ruuskanen, C. R. Primmer, and E. H. Leder. (2012) The proteomics of feather development in pied flycatchers *(Ficedula hypoleuca)* with different plumage coloration. Molecular Ecology **21**: 5762–5777.

Lewis, Z. A., A. L. Shiver, N. Stiffler, M. R. Miller, E. A. Johnson, and E. U. Selker. (2007) High density detection of restriction site associated DNA (RAD) markers for rapid mapping of mutated loci in Neurospora. Genetics **177**: 1163–1171.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics **25**: 2078–2079.

Lindstrom, K., and J. Lundstrom. (2000) Male greenfinches (*Carduelis chloris*) with brighter ornaments have higher virus infection clearance rate. Behavioral Ecology and Sociobiology **48**: 44–51.

Ling, M. K., M. C. Lagerström, R. Fredriksson, R. Okimoto, N. I. Mundy, S. Takeuchi, and

H. B. Schiöth. (2003) Association of feather colour with constitutively active melanocortin 1 receptors in chicken. European Journal of Biochemistry **270**: 1441–1449.

Lozano, G. A. (1994) Carotenoids, parasites, and sexual selection. Oikos **70**: 309–311.

Manolio, T. A., A. E. Guttmacher, and T. A. Manolio. (2010) Genome wide association studies and assessment of the risk of disease. The New England Journal of Medicine **363**: 166–176.

Marks, G. S. (1969) Heme oxygenase: function, multiplicity, regulatory mechanisms, and clinical applications. The FASEB Journal **2**: 2557–2568.

Mason, K. A., and S. K. Frost-Mason. (2000) Evolution and development of pigment cells: At the crossroads of the discipline. Pigment Cell Research **13**: 150–155.

Mather, K. (1955) Polymorphism as an outcome of disruptive selection. Evolution **9**: 52–61.

Maynard Smith, J. (1982) Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Mays Jr, H. L., and G. E. Hill. (2004) Choosing mates: good genes versus genes that are a good fit. Trends in Ecology and Evolution **19**: 554–559.

McDowall, M., N. M. Edwards, C. A. B. Jahoda, and P. I. Hynd. (2008) The role of activins and follistatins in skin and hair follicle development and function. Cytokine and Growth Factor Reviews **19**: 415–426.

McGraw, K. J., and G. E. Hill. (2000) Differential effects of endoparasitism on the expression of carotenoid- and melanin-based ornamental coloration. Proceedings of the Royal Society B **267**: 1525–1531.

McGraw, K. J., and G. E. Hill. (2001) Carotenoid access and intraspecific variation in plumage pigmentation in male American goldfinches (*Carduelis tristis*) and northern cardinals (*Cardinalis cardinalis*). Functional Ecology **15**: 732–739.

McGraw, K. J., G. E. Hill, R. Stradi, and R. S. Parker. (2001) The influence of carotenoid acquisition and utilization on the maintenance of species-typical plumage

pigmentation in male American goldfinches (*Carduelis tristis*) and northern cardinals (*Cardinalis cardinalis*). Physiological and Biochemical Zoology **74**: 843–852.

McGraw, K. J., E. Adkins-Regan, and R. S. Parker. (2002) Anhydrolutein in the zebra finch: a new, metabolically derived carotenoid in birds. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **132**: 811–818.

McGraw, K. J., and J. G. Schuetz. (2004) The evolution of carotenoid coloration in estrildid finches: a biochemical analysis. Comparative Biochemistry and Physiology B **139**: 45–51.

McGraw, K. J. (2006) Mechanics of carotenoid-based colouration. In G. E. Hill and K. J. McGraw, Bird Coloration, Volume 1: Mechanisms and Measurements, 177. Harvard University Press, Cambridge, MA.

McGraw, K. J., and R. S. Parker. (2006) A novel lipoprotein-mediated mechanism controlling sexual attractiveness in a colorful songbird. Physiology and Behavior **87**: 103–108.

McGraw, K. J. (2008) An update on the honesty of melanin–based colour signals in birds. Pigment Cell & Melanoma Research **21**: 133–138.

Merrill, R. M., B. V. Schooten, J. A. Scott, and C. D. Jiggins. (2011) Pervasive genetic associations between traits causing reproductive isolation in *Heliconius* butterflies. Proceedings of the Royal Society B **278**: 511–518.

Metzker, L. (2010) Sequencing technologies - the next generation. Nature Review Genetics **1**: 31–46.

Miller, M. R., T. S. Atwood, B. F. Eames, J. K. Eberhart, Y. L. Yan, J. H. Postlethwait, and E. A. Johnson. (2007a) RAD marker microarrays enable rapid mapping of zebrafish mutations. Genome Biology **8**: R105.

Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, E. A. Johnson. (2007b) Rapid and cost-effective polymorphism identification and genotyping using restriction site

associated DNA (RAD) markers. Genome Research **17**: 240–248.

Miwa, M., M. Inoue-Murayama, N. Kobayashi, B. B. Kayang, M. Mizutani, H. Takahashi, and S. Ito. (2006) Mapping of panda plumage colour locus on the microsatellite linkage map of the Japanese quail. BMC Genetics **7**: 1471–2156.

Moran, N. A., and T. Jarvik. (2010) Lateral transfer of genes from fungi underlies carotenoid production in aphids. Science **328**: 624–627.

Moreau, R. E. (1958) Some aspects of the Musophagidae. Part 3. Some general features. Ibis **100**: 238–270.

Mulder, R. A., R. Ramiarison, and R. E. Emahalala. (2002) Ontogeny of male plumage dichromatism in Madagascar paradise flycatchers *Terpsiphone mutata*. Journal of Avian Biology **33**: 342–348.

Mundy, N. I., N. S. Badcock, T. Hart, K. Scribner, K. Janssen, and N. J. Nadeau. (2004) Conserved genetic basis of a quantitative plumage trait involved in mate choice. Science **303**: 1870–1873.

Mundy, N. I. (2005) A window on the genetics of evolution: *MC1R* and plumage colouration in birds. Proceedings of the Royal Society B **272**: 1633–1640.

Murray, R. (1963) The genetics of the yellow-masked Gouldian finch. Avicultural Magazine **69**: 108–113.

Nadeau, N. J., F. Minvielle, S. Ito, M. Inoue-Murayama, D. Gourichon, S. A. Follett, T. Burke, and N. I. Mundy. (2008) Characterization of Japanese Quail *yellow* as a Genomic Deletion Upstream of the Avian Homolog of the Mammalian *ASIP* (*agouti*) Gene. Genetics **178**: 777–786.

Navarro, A., and N. H. Barton. (2003) Chromosomal speciation and molecular divergence - accelerated evolution in rearranged chromosomes. Science **300**: 321–324.

Nishimura, M., H. Yaguti, H. Yoshitsugu, S. Naito, and T. Satoh. (2003) Tissue distribution of mRNA expression of human cytochrome P450 isoforms assessed by high-sensitivity real-time reverse transcription PCR. Yakugaku Zasshi. **123**: 369–

375.

Norris, K. (1993) Heritable variation in a plumage indicator of viability in male great tits *Parus major*. Nature **362**: 537–539.

Ohyama, A., F. Saito, H. Ohuchi, and S. Noji. (2001) Differential expression of two BMP antagonists, gremlin and Follistatin, during development of the chick feather bud. Mechanisms of Development **100**: 331–333.

O'Malley, C. (2006) National recovery plan for the Gouldian finch (*Erythrura gouldiae*). National Gouldian Finch Recovery Team, Northern Territory Department of Natural Resources, Environment and the Arts, Palmerston.

Ortonne, J. P. (2002) Photoprotective properties of skin melanin. British Journal of Dermatology **146**: 7–10.

Parker, A. R. (1999) Invertebrate structural colours. In E. Savazzi, ed., Functional Morphology of the Invertebrate Skeleton, 65–90. John Wiley and Sons, London.

Parker, R. S. (1996) Absorption, metabolism, and transport of carotenoids. The FASEB Journal **10**: 542–551.

Pearson, T. A., and T. A. Manolio. (2008) How to interpret a genome-wide association study. The Journal of the American Medical Association **299**: 1335–1344.

Prager, M., and S. Andersson. (2009) Phylogeny and evolution of sexually selected tail ornamentation in widowbirds and bishops (*Euplectes spp.*). Journal of Evolutionary Biology **22**: 2068–2076.

Prum, R., and S. Williamson. (2002) Reaction-diffusion models of within-feather pigmentation patterning. Proceedings of the Royal Society B **269**: 781–792.

Prum, R. O., and R. H. Torres. (2003) Structural colouration of avian skin: convergent evolution of coherently scattering dermal collagen arrays. Journal of Experimental Biology **206**: 2409–2429.

Pryke, S. R., and S. C. Griffith. (2006) Red dominates black: Agonistic signalling among head morphs in the colour polymorphic Gouldian finch. Proceedings of the Royal

Society B **273**: 949–957.

Pryke, S. R. (2007) Fiery red heads: female dominance among head color morphs in the Gouldian finch. Behavioral Ecology **18**: 621–627.

Pryke, S. R., and S. C. Griffith. (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.

Pryke, S. R., and S. C. Griffith. (2009a) Maternal adjustment of parental effort in relation to mate compatibility affects offspring development. Behavioral Ecology **21**: 226–232.

Pryke, S. R., and S. C. Griffith. (2009b) Genetic incompatibility drives sex allocation and maternal investment in a polymorphic finch. Science **323**: 1605–1607.

Pryke, S. R., and S. C. Griffith. (2009c) Postzygotic genetic incompatibility between sympatric color morphs. Evolution **63**: 793–798.

Pryke, S. R. (2010) Sex chromosome linkage of mate preferences and color signal maintains assortative mating between interbreeding finch morphs. Evolution **65**: 1301–1310.

Qvarnström, A., and R. I. Bailey. (2009) Speciation through evolution of sex-linked genes. Heredity **102**: 4–15.

Reynolds, J. D., and M. R. Gross. (1990) Costs and benefits of female mate choice: is there a lek paradox? The American Naturalist **136**: 230–243.

Richman, A. D., and T. Price. (1992) Evolution of ecological differences in the Old World leaf warblers. Nature **355**: 817–821.

Rieseberg, L. H., and J. M. Burke. (2001) A genic view of species integration. Journal of Evolutionary Biology **14**: 883–886.

Riley, P. A. (1992) Materia melancia: Further dark thoughts. Pigment Cell Research **5**: 101–106.

Riley, P. A. (1997) Melanin. International Journal of Biochemistry and Cell Biology **29**: 1235–1239.

Risch, N., and K. Merikangas. (1996) The future of genetic studies of complex human

diseases. Science **273**: 1516–1517.

Rohwer, S. A. (1975) The social significance of avian winter plumage variability. Evolution **29**: 593–610.

Rohwer, S. (1983) Formalizing the avoidance-image hypothesis: critique of an earlier prediction. The Auk **100**: 971–974.

Rohwer, S., and D. R. Paulson. (1987) The avoidance image hypothesis and color polymorphism in Buteo hawks. Ornis Scandinavica **18**: 285–290.

Roulin, A. (2004) The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. Biological Reviews **79**: 815–848.

Rozen, S., and H. J. Skaletsky. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols: methods in molecular biology, 365–386. Humana Press, Totowa, NJ.

Saino, N., P. Ninni, M. Incagli, S. Calza, R. Sacchi, and A. P. Møller. (2000) Begging and parental care in relation to offspring need and condition in the barn swallow (*Hirundo rustica*). The American Naturalist **156**: 637–649.

Saks, L., I. Ots, and P. Hõrak. (2003) Carotenoid-based plumage coloration of male greenfinches reflects health and immunocompetence. Oecologia **134**: 301–307.

Schena, M., D. Shalon, R. W. Davis, and P. O. Brown. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270**: 467–470.

Schiedt, K. (1990) New aspects of carotenoid metabolism in animals. In Carotenoids: Chemistry and Biology (ed. N. I. Krinsky, M. M. Mathews-Roth, and R. F. Taylor), 247–268. Plenum Press, New York.

Shawkey, M. D., S. R. Pillai, and G. E. Hill. (2003) Chemical warfare? Effects of uropygial oil on feather-degrading bacteria. Journal of Avian Biology **34**: 345–349.

Sheldon, B. C., J. Merilä, A. Qvarnström, L. Gustafsson, and H. Ellegren. (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary

sexual character. Proceedings of the Royal Society B **264**: 297–302.

Shendure, J., and H. Ji. (2008) Next-generation DNA-sequencing. Nature biotechnology **26**: 1135–1145.

Shetty, S., D. K. Griffin, and J. A. M. Graves. (1999) Comparative painting reveals strong chromosome homology over 80 million years of bird evolution. Chromosome Research **7**: 289–295.

Shibusawa, M., C. Nishida-Umehara, M. Tsudzuki, J. Masabanda, D. K. Griffin, and Y. Matsuda. (2004) A comparative karyological study of the blue-breasted quail (*Coturnix chinensis*, Phasianidae) and California quail (*Callipepla californica*, Odontophoridae). Cytogenetic and Genome Research **106**: 82–90.

Slate, J., P. M. Visscher, S. MacGregor, D. R. Stevens, M. L. Tate, and J. M. Pemberton. (2002) A genome scan for QTL in a wild population of red deer (*Cervus elaphus*). Genetics **162**: 1863–1873.

Southern, H. N. (1945) Polymorphism in *Poephila Gouldiae*. Journal of Genetics **47**: 51–57.

Southern, H. N. (1946) The inheritance of head colour in the Gouldian finch. Avicultural Magazine **52**: 126–131.

Srivastava, S. K., P. Wolinski, and A. Pereira. (2014) A Strategy for Genome-Wide Identification of Gene Based Polymorphisms in Rice Reveals Non-Synonymous Variation and Functional Genotypic Markers. PLoS ONE **9**: e105335.

Stapley, J., T. R. Birkhead, T. Burke, and J. Slate. (2008) A linkage map of the zebra finch *Taeniopygia guttata* provides new insight into avian genome evolution. Genetics **179**: 651–667.

Stickney, H. L., J. Schmutz, I. G. Woods, C. C. Holtzer, M. C. Dickson, P. D. Kelly, R. M. Myers, and W. S. Talbot. (2002) Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays. Genome Research **12**: 1929–1934.

Stradi, R. (1998) The colour of flight: carotenoids in bird plumage. Solei Gruppo Editoriale Informatico, Milan, Italy.

Stradi, R. (1999) Pigmenti e sistematica degli uccelli. In L. Brambilla, G. Canali, E. Mannucci, *et al.*, ed., Colori in volo: il piumaggio degli ucceli, 117–146. Università degli Stusi di Milano, Milan, Italy.

Suh, Y., and J. Vijg. (2005) SNP discovery in associating genetic variation with human disease phenotypes. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis **573**: 41–53.

Számadó, S. (2000) Cheating as a mixed strategy in a simple model of aggressive communication. Animal Behaviour **59**: 221–230.

Takeuchi, S., H. Suzuki, M. Yabuuchi, and S. Takahashi. (1996) A possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken. Biochimica et Biophysica Acta **1308**: 164–168.

Teshigawara, K., S. Takahashi, T. Boswell, Q. Li, S. Tanaka, and S. Takeuchi. (2001) Identification of avian alpha-melanocyte-stimulating hormone in the eye: temporal and spatial regulation of expression in the developing chicken. Journal of Endocrinology **168**: 527–537.

Theron, E., K. Hawkins, E. Bermingham, R. Ricklefs, and N. I. Mundy. (2001) The molecular basis of an avian plumage polymorphism in the wild: a point mutation in the melanocortin-1 receptor is perfectly associated with melanism in the bananaquit (*Coereba flaveola*). Current Biology **11**: 550–557.

Thomas, J. W., M. Cáceres, J. J. Lowman, C. B. Morehouse, M. E. Short, E. L. Baldwin, D. L. Maney, and C. L. Martin. (2008) The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (*Zonotrichia albicollis*) is a complex rearrangement and suppressor of recombination. Genetics **179**: 1455–1468.

Thomas, P. E., R. Klinger, L. I. Furlong, M. Hofmann-Apitius, and C. M. Friedrich. (2011) Challenges in the association of human single nucleotide polymorphism mentions with unique database identifiers. BMC Bioinformatics **12**: S4.

Tidemann, S. C. (1996) Causes of the decline of the Gouldian finch *Erythrura gouldiae*. Biological Conservation International **6**: 49–61.

Tidemann, S. C., Lawson, C., Elvish, R., Boyden, J., and J. Elvish. (1999) Breeding biology of the Gouldian finch *Erythrura gouldiae*, an endangered finch of northern Australia. Emu **99**: 191–199.

Vaez, M., S. A. Follett, B. Bed'hom, D. Gourichon, M. Tixier-Boichard, and T. Burke. (2008) A single point-mutation within the *melanophilin* gene causes the lavender plumage colour dilution phenotype in the chicken. BMC Genetics **9**:7.

Veen, T., T. Borge, S. C. Griffith, G. P. Sætre, S. Bures, L. Gustafsson, and B. C. Sheldon. (2001) Hybridization and adaptive mate choice in flycatchers. Nature **411**: 45–50.

Verne, J. (1926) Les pigments dans l’Organisme animal, Gaston Doin et Cie Paris, 608.

Vershinin, A. (1999) Biological functions of carotenoids - diversity and evolution. BioFactors **10**: 99–104.

Voorris, R. E. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of Heredity **1**: 77–78.

Walsh, N., J. Dale, K. J. McGraw, M. A. Pointer, and N. I. Mundy. (2012) Candidate genes for carotenoid coloration in vertebrates and their expression profiles in the carotenoid-containing plumage and bill of a wild bird. Proceedings of the Royal Society B **279**: 58–66.

Warren W. C., D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier, A. Kunstner, S. Searle, S. White, A. J. Vilella, S. Fairley, A. Heger, L. Kong, C. P. Ponting, E. D. Jarvis, C. V. Mello, P. Minx, P. Lovell, T. A. F. Velho, M. Ferris, C. N. Balakrishnan, S. Sinha, C. Blatti, S. E. London, Y. Li, Y-C. Lin, J. George, J. Sweedler, B. Southey, P. Gunaratne, M. Watson, K. Nam, N. Backstrom, L. Smeds, B. Nabholz, Y. Itoh, O. Whitney, A. R. Pfenning, J. Howard, M. Volker, B. M. Skinner, D. K. Griffin, L. Ye, W. M. McLaren, P. Flicek, V. Quesada, G. Velasco, C. Lopez-Otin, X. S. Puente, T. Olender, D. Lancet, A. F. A. Smit, R. Hubley, M. K. Konkel, J. A. Walker, M. A.

Batzer, W. Gu, D. D. Pollock, L. Chen, Z. Cheng, E. E. Eichler, J. Stapley, J. Slate, R. Ekblom, T. Birkhead, T. Burke, D. Burt, C. Scharff, I. Adam, H. Richard, M. Sultan, A. Soldatov, H. Lehrach, S. V. Edwards, S-P. Yang, X. Li, T. Graves, L. Fulton, J. Nelson, A. Chinwalla, S. Hou, E. R. Mardis, and R. K. Wilson. (2010) The genome of a songbird. Nature **464**: 757–762.

Wilkening S., B. Chen, J. L. Bermejo, and F. Canzian. (2009) Is there still a need for candidate gene approaches in the era of genome-wide association studies? Genomics **93**: 415–419.

Williams, L. J., A. J. King, and C. Mettke-Hofmann. (2012) Colourful characters: head colour reflects personality in a social bird, the Gouldian finch, *Erythrura gouldiae*. Animal Behaviour **84**: 159–165.

Woinarski, J. C. Z. and S. C. Tidemann. (1992) Survivorship and some population parameters for the endangered Gouldian finch *Erythrura gouldiae* and two other finch species at two sites in tropical northern Australia. Emu **92**: 33–38.

Wu, S., C. R. Moomaw, K. B. Tomer, J. R. Falck, and D. C. Zeldin. (1996) Molecular cloning and expression of *CYP2J2*, a human cytochrome P450 arachidonic acid epoxygenase highly expressed in heart. Journal of Biological Chemistry **271**: 3460–3468.

Zi, J., X. Yu, Y. Li, X. Hu, C. Xu, X. Wang, X. Liu, and R. Fu. (2003) Coloration strategies in peacock feathers. Proceedings of the National Academy of Sciences, USA **100**: 12576–12578.