The Colour Gene *MuPKS* in Budgerigars

John Mulley AO PhD, December 2021

Introduction

This describes (in language simplified as much as possible without losing meaning) the discovery of the budgerigar Colour gene. The identity of the Colour gene was arrived at by identifying its Blue mutation. The original references are given for anyone wishing to dive deeper into the technical detail but as breeders all we need to know for breeding budgerigars is that the various mutations of the Colour gene obey Mendelian inheritance. However, there have always been those among us who wonder how a blue bird can have a yellow top? That being the case, some have wondered can we have the reverse, a green bird with a white top? Since feathers in other parrots can be red some wonder why budgerigars do not have red feathers?

Melanised budgerigar feathers reflect blue light. The addition of yellow pigment gives green feathers. The Blue mutation blocks the synthesis of yellow pigment leaving the bird with blue feathers. Knowing the identity of the Colour gene and the enzyme it encodes enables us to infer mechanisms for the expression of Golden Faced Blue, Yellow Faced Blue and White Cap which are variations of the amount and distribution of the yellow pigment. Self-coloured non green non blue varieties lack melanin but can express varying amounts of yellow pigmentation from zero to 100% depending on which variant of the Colour gene they express.

The genome sequence for the budgerigar was published in 2014. This was a collaborative effort among genome scientists from the USA and China and provided the DNA blueprint for discovery of genes responsible for heritable variations the budgerigar: Ganapathy, G., Howard, J.T., Ward, J.M., Li, J., Li, B., Li, Y., Xiong, Y., Zhang, Y., Zhou, S., Schwartz, D.C., Schatz, M., Aboukhalil, R., Fedrigo, O. Bukovnik, L., Wang, T., Wray, G., Rasolonjatovo, I., Winer, R., Knight, J. R., Koren, S., Warren, W.C., Zhang, G., Phillippy, A.M., Jarvis, E.D. (2014). High-coverage sequencing and annotated assemblies of the budgerigar genome. Gigascience 3, 11-19.

Discovery of the Colour Gene

The colour gene *MuPKS* is the first budgerigar gene to be described at the molecular level, in 2017. This is the gene that encodes the enzyme that synthesises the yellow pigment psittacofulvin. Discovery was achieved by a team of molecular geneticists, genome scientists, biochemists and cell biologists from the USA and Taiwan: Cooke, T.F., Fischer, C.R., Wu, P., Jiang, T-X,. Xie, K.T., Kuo, J., Doctorov, E., Zehnder, A., Khosla, C., Chuong, C-M., Bustamante, C.D. (2017). Genetic mapping and biochemical basis of yellow pigmentation in budgerigars. Cell 171: 427-439.

Mutations of the colour gene *MuPKS* are responsible in budgerigars for the Green-Golden Faced Blue-Yellow Faced Blue-White Cap-Blue allelic series. *MuPKS* is that part of the

budgerigar DNA sequence that codes for the enzyme polyketide synthase that produces the yellow pigment responsible for green budgerigars. The recessive blue mutation of *MuPKS* is a loss of function mutation that abolishes production of yellow pigment leaving the melanised feather structure to reflect blue light. *MuPKS* is the first gene to be identified for a Mendelian trait in budgerigars.

Psittacofulvin in the red to yellow pigment spectrum is the mechanism for feather colouration in parrots. The red component is missing from budgerigars, a mystery yet to be solved. Mapping and identifying genes for the other Mendelian traits in budgerigars could theoretically be carried out in the same way that succeeded for the discovery of the *MuPKS* gene. Steps that led to the discovery of the colour gene can be summarised as follows:

Physical map of the budgerigar genome

The project to discover the colour gene was based on the premise that on an evolutionary timescale relatively few generations have elapsed since the time the first blue budgerigar appeared in an aviary. Although there may have been multiple independent origins of a recessive blue mutation circulating among wild green budgerigars split for blue, hidden by the dominant green allele, most likely only one of these became the founding blue mutation in our domesticated budgerigars. The founding blue bird needed to carry a pair of blue mutations and likely these were identical by descent from a single ancestral mutational event being finally exposed some generations later through inbreeding. This domesticated blue bird survived because it was protected in an aviary environment.

An ancestral blue mutation should under those circumstances exhibit linkage disequilibrium with nearby genetic markers. Linkage disequilibrium in populations is where the distribution of alleles at closely spaced adjacent positions (loci) is non-random, retaining the pattern surrounding the mutation in the ancestral blue bird. The rate of decay of linkage disequilibrium over time by recombination between homologous chromosomes is proportional to the distance between loci.

If linkage disequilibrium remains for markers close to the blue mutation in the budgerigar DNA sequence, then the location of the blue mutation within the budgerigar genome can in theory be determined by the technique of association mapping. But before that, many polymorphic genetic markers needed to be characterised and mapped to linkage groups corresponding to each of the budgerigar chromosomes.

SNP discovery and Mapping the Blue Locus to a Chromosome

Ideal genetic markers are SNPs (single-nucleotide polymorphisms). These are stable single base changes within the DNA sequence discovered in this instance in the budgerigar by an advanced DNA sequencing method. SNPs are commonly used as research tools in association mapping to pinpoint the multiple genes responsible for production traits in commercial species of livestock.

DNA was sequenced from 234 budgerigars of which 105 were blue as well as 15 museum specimens of wild budgerigars from Australian museum specimens. This provided 69,855 SNP markers at polymorphic frequency suitable for association mapping. Polymorphic markers are positions in the genome that segregate alternative nucleotide bases in the DNA sequence. SNP variation within the DNA sequence of any organism, including us, generally has little if any detectable effect on the biological fitness or appearance of the carrier. SNP positions within the genome sequence can be precisely determined and several from a single region on chromosome 1 (the largest chromosome in budgerigars) were consistent with complete linkage disequilibrium but only among budgerigars with the blue mutation. This was the entrée to discovery of the Blue mutation, and hence the colour gene among about 20,000 other genes that may exist in the budgerigar genome.

The Single Ancestral Haplotype Associated with the Blue Trait

A haplotype is a run of alleles at adjacent SNP positions "closely" located with each other along the chromosome. All blue budgerigars shared a single haplotype in complete linkage disequilibrium within a 400,000 base pair DNA sequence located between SNPs at base pair coordinates 21,019,187 and 21,445,705. Within this region of the chromosome there were 11 distinct DNA sequences predicted from computer software analysis to represent genes.

The sequence of one of these genes which they named *MuPKS* was predicted to encode a previously uncharacterised enzyme. The predicted amino acid sequence for this enzyme closely aligned with polyketide synthases across species. This was the most promising of the 11 putative genes from the region of interest for production of psittacofulvin because related polyketide synthases in bacteria and fungi were known to synthesise yellow pigments with similar molecular structures to psittacofulvin.

***MuPKS* Expression in Feathers**

None of the 11 putative genes within the blue associated haplotype showed any differential gene expression between blue and green budgerigars. This included *MuPKS* that was found to be highly expressed in feathers of both blue and green budgerigars. Thus, the recessive blue phenotype is not caused by changes in the expression of any of these 11 genes in budgerigars, including *MuPKS*. It then became necessary to look downstream of gene expression.

It was interesting to discover that the *MuPKS* gene was expressed hundreds to thousands of times higher in budgerigar feathers than its gene homologs in chicken and crow feathers that do not carry yellow psittacofulvin pigment. So, there was still something peculiar about *MuPKS* in budgerigar feathers or its gene product that required further study.

This line of enquiry was developed further after the activity of polyketide synthase was examined in the feathers of green and blue budgerigars. For an enzyme to work properly, firstly the gene encoding that enzyme must be expressed at normal levels, and that is the case for

MuPKS. Secondly, the transcribed mRNA must be stable. Thirdly, the active site of the translated enzyme must function normally.

A Coding SNP within the *MuPKS* Gene associated with Yellow Pigmentation

Most SNPs occur in non-coding DNA sequence between genes and between the coding elements within genes. Mutant base changes that occur within coding sequences of genes are detrimental if they occur within crucial functional domains. Since variable transcription for *MuPKS* was eliminated, coding changes were sought within the sequences of the 11 genes within the critical 400,000 base pairs that may affect activity of the translated protein. The only protein coding change within a gene was found to cause a substitution of the amino acid arginine with the amino acid tryptophan (abbreviated R → W) at residue 644 in the active site of the enzyme polyketide synthase encoded by *MuPKS*. All blue budgerigars were W/W and green birds were either R/R or R/W, indicating that some of the green birds were split blue.

Position 644 of polyketide synthase is highly conserved throughout evolution with an arginine (R) across numerous species from bacteria to vertebrates. That suggests that an amino acid substitution at 644 would be detrimental and carriers weeded out by natural selection if the associated phenotype occurred in wild populations. A single blue budgerigar within a flock of greens would quickly attract a predator's attention. However, protection within an aviary of domesticated budgerigars negates natural selection. To cut a long story short, it was shown experimentally that although the blue mutation does not affect gene expression the enzyme it encodes with the R → W amino acid substitution is not active. This loss-of-function mutation kills the active site of this enzyme so the enzyme does not produce psittacofulvin. Without the yellow pigment the feathers of the budgerigar reflect blue light.

Reconstitution of Psittacofulvin Synthesis in a Yeast Host

Strong circumstantial evidence as outlined above supported the R644W substitution in polyketide synthase as the reason why no yellow pigment is produced. This enables us to see reflected blue light from the budgerigar feather without the yellow pigment. Definitive proof came after cloning the budgerigar *MuPKS* gene and expressing it in a yeast host. Organic extracts from the yeast expressing the wildtype budgerigar *MuPKS* gene in yeast were yellow but organic extracts from yeast expressing the blue sequence of the *MuPKS* gene were clear. An experiment mutating the wildtype *MuPKS* gene that had been inserted into the yeast to the blue form of *MuPKS* also produced clear extracts. Comparison between the yellow pigment from the yeast extracts with the yellow pigment from budgerigar feathers by mass spectrometry yielded identical long chain carbon compounds. *MuPKS* is therefore causative beyond doubt for the yellow psittacofulvin pigments found in green budgerigar feathers.

Diverse Functions of Polyketide Synthases

Parrots are the only birds to exhibit psittacofulvin pigmentation, but *MuPKS* homologs are widespread across phyla. This suggests that this family of enzymes have a variety of essential

functions in addition to pigmentation in budgerigar feathers. How this gene has been harnessed by the parrots to produce the yellow to red psittacofulvins remains a mystery requiring further investigation. Why budgerigars produce only the yellow form of psittacofulvin is also a mystery, other than knowing that one of the long chain carbon components of psittacofulvin was found to be missing in the budgerigar. This adaptation combining yellow pigment with the blue reflected light from the underlying feather structure is clearly advantageous for the evolutionary success of the native wild green budgerigar.

Yellow Faced Blues, Golden Faced Blues and White Caps

Yellow Faced Blues, Golden Faced Blues and White Caps are not "varieties" superimposed on Blue. They are alternative colour morphs, distinct from Blue, but controlled by different allelic mutations of the same colour gene *MuPKS* that is responsible for green and blue budgerigars. A Green split Yellow Faced Blue can never pass on Blue because its genotype is Green/Yellow Faced Blue. A Golden Faced Blue/Yellow Faced Blue hybrid does not carry the more recessive Blue mutation so mated to Blue it cannot produce a Blue, only Single Factor Golden Faced Blue or Yellow Faced Blue which are both dominant to Blue. Similarly, a Yellow Faced Blue/White Cap hybrid does not carry a Blue mutation but mated to Blue can be predicted to breed only Yellow Faced Blues and Single Factor Whites Caps, both of which are split for Blue.

Whilst the molecular basis for green and blue budgerigars is now fully understood following the research published in the journal Cell, DNA sequencing to determine the *MuPKS* mutations responsible for Yellow Faced Blues, Golden Faced Blues and White Caps has not yet been reported. Until then, the cellular mechanisms responsible for these phenotypes can only be speculated. However, armed as we now are with knowledge of *MuPKS* and its gene product polyketide synthase some educated hypotheses can be proposed.

I contacted Professor Carlos Bustamante from Stanford University, the senior and corresponding author on the research paper published in Cell (2017) 171: 427-439 describing the discovery of the *MuPKS* colour gene and its blue mutation. I wondered if they had taken the research further, to include sequencing the other allelomorphs to detect any of the other *MuPKS* mutations. He referred me on to the lead investigator, his former graduate student (PhD student in Australian terminology) Dr Thomas Cooke. Tom remains extremely interested in the topic although has now moved on to unrelated post-doctoral research.

Thomas speculated that temperature sensitive mutations of *MuPKS* might explain the Golden Faced Blues and Yellow Faced Blues and a regulatory mutation in the promoter region of the *MuPKS* gene or a mutation affecting the stability of its mRNA might explain the White Caps. He referred to a known temperature sensitive mutation in Siamese cats.

Abbreviations that follow are: B for Blue, YF for Yellow Faced Blue, GF for Golden Faced Blue, WC for White Cap, SF for Single Factor and DF for Double Factor.

Yellow Faced Blues

Siamese cats give us a clue. An established temperature sensitive mutation of the gene encoding the tyrosine kinase enzyme restricts the distribution of melanin to the coolest part of the Siamese cat: its face, ears, feet and tail.

Thomas Cooke suggested a similar mechanism may account for the restricted distribution of yellow pigment to the crown, frontal and mask of the Yellow Faced Blues. He mentioned that the budgerigar polyketide synthase probably functions as a dimer (by analogy to fatty acid synthase, to which it is closely related in terms of amino acid sequence). Dimers are the functional form of an enzyme commonly seen in enzymology consisting of two amino acid sequences encoded by the same gene bonded end to end to enable them to function enzymatically.

The phenotype YF blue is a hybrid between YF and B (blue). These hybrids are predicted to have YF-YF, YF-B and B-B polyketide synthase dimer configurations. Subunits would randomly associate in the dimer ratios of 1 : 2 : 1, respectively. The proposal is that yellow pigment is distributed in budgerigar feathers as follows, based on the hypothesis of temperature sensitivity for the YF mutation:

- If the YF-YF homodimer is temperature unstable and therefore non-functional it produces no yellow pigment. That would explain the white face phenotype in the DF YF blues.
- We already know that the B-B homodimer is not functional as proven by the research describing the arginine to tryptophan substitution in polyketide synthase. There is no need to suggest it might be temperature sensitive.
- Assuming that B is temperature stable then it may partially stabilise the YF-B heterodimer allowing some functional enzymatic activity derived from the YF subunit. This dimer configuration might biochemically rescue yellow pigment production to a limited extent. That might explain why yellow pigment is only deposited in the cooler feathers of the SF YF budgerigar, in the feathers of the crown, frontal and mask.

Meaningful laboratory experiments could be designed to test this hypothesis. For example, the YF sequence could be cloned into the above-mentioned yeast host and pigment levels assessed at different temperatures with and without co-expression with B. I am not aware of any alternative hypotheses to explain the phenotypes associated with the YF mutation.

Golden Faced Blues

The proposal is that the GF-B heterodimer is slightly more temperature stable than the YF-B heterodimer, with the GF dimer component more active than the YF dimer component.

- That explains the brighter yellow colour on the crown, frontal and mask with more efficient rescue of yellow pigment production by the GF-B heterodimer relative to the YF-B heterodimer.
- Slightly enhanced temperature stability of GF relative to YF also explains why in the SF GF the GF-B heterodimer creates some spillage of yellow pigment throughout the warmer body feathers as the bird ages.
- If the GF-GF homodimer in the DF GF has some degree of temperature stability then that may be sufficient to deposit yellow pigment within the feathers on the cooler crown, frontal and mask, as observed, but not enough to suffuse through the warmer body feathers as seen in the SF GF.

Meaningful laboratory experiments could be designed to test these possibilities, as suggested above for YF. I am not aware of any alternative hypotheses to explain the phenotypes associated with the GF mutation.

We know from observation that the intensity of the yellow pigment can vary slightly within the Green, Yellow Faced Blue and Golden Faced Blue Series. That adds additional complexity and all one can do is speculate that additional modifier genes of relatively small effect or other regulatory changes within DNA external to the *MuPKS* gene sequence can enhance or dilute the yellow pigment to some degree. Playing around with Yellow Faced Blue-Golden Faced Blue hybrids can have the same outcome. It is the role of judges to ensure the colour groups remain true to their written Standard for the exhibition budgerigar. Unequivocal deep buttercup yellow expressed in Golden Faced Blue on the frontal, crown and mask must remain its characteristic varietal feature.

White Caps

I described and sent a photograph of a White Cap to Thomas Cooke. He was not aware of this mutation from his interactions with North American budgerigar breeders since the White Cap is currently restricted to Australia. An explanation within a Mendelian framework is straight forward for this mutation and could be explained by any one of three mechanisms. There could be a mutation in the promoter region of *MuPKS* that significantly reduces gene activity but does not completely block its transcription into mRNA, ultimately reducing the amount of polyketide synthase and yellow pigment. Another explanation might be mutation within the *MuPKS* DNA sequence that adversely affects but does not totally kill the active site of the enzyme, reducing yellow pigment. Finally, there could be a mutation that affects the stability of its transcribed mRNA affecting the amount of yellow pigment produced.

Nobody has yet sequenced the *MuPKS* gene in a White Cap to determine which of these mechanisms might be responsible. RNA from pinfeathers from the head of a White Cap is yet to be examined to compare the level of expression of *MuPKS* with pinfeathers in the caps of GF, YF and Green birds. Low level presence of yellow pigment in White Caps does not necessarily mean that it has mutated from YF or GF but genomic sequencing of *MuPKS* in Golden Faces,

Yellow Faces and White Caps will resolve the origin of the White Cap mutation. The White Cap mutation approaches but does not completely present as a green budgerigar with a white top.

Any of the moderate loss-of-function mutational mechanisms proposed for the White Cap might explain the green body colour for the Double Factor White Cap within a Mendelian framework. The amount of yellow pigment produced would simply be additive, depending on whether one dose of the White Cap mutation is carried in the WC-B hybrid (body appearing aqua in the blue form) or two doses of this mutation are carried in the double factor WC-WC (body appearing green). The double dose of WC produces more yellow pigment, hence green body feathers, but not enough yellow pigment is produced to completely cover the cap. Selection of colour modifiers may slightly enhance the extent of the distribution of the white on the cap in the same way that we see subtle variation in intensity of the yellow pigment in the Greens, Yellow Faced Blues and Golden Faced Blues.

I have bred a White Cap/ Yellow Faced Blue hybrid on a grey background. Apart from a slightly paler yellow cap this hen looks very much like a Grey Green or a Single Factor Golden Faced Grey. That observation is consistent with additivity of yellow pigment conferred by the White Cap mutation when combined with the yellow from the Yellow Faced Blue mutation. Test mated to a Normal Sky Blue (the most recessive of the Colour mutations) verified her hybrid Yellow Faced Blue/White Cap genotype. She produced eight Yellow Faced Blues and six White Caps from 14 chicks over two nests, all of which will be split for Blue, consistent with Mendelian expectations.

Conclusion

Identity of the Colour gene and the molecular basis for its Blue mutation is proven beyond doubt. Plausible mechanisms for Yellow Faced Blue, Golden Faced Blue and White Cap remain speculative until their hypothesised molecular basis and cellular mechanisms can be validated in the laboratory or plausible alternative hypotheses emerge to explain those phenotypes.

Acknowledgment: I thank Dr Thomas Cooke for generously sharing his thoughts on possible cellular mechanisms for the Yellow Faced Blue, Golden Faced Blue and White Cap. None of that was in the public domain but provides fascinating insights for students of the budgerigar.