ACTA BIOLÓGICA COLOMBIANA
http://www.revistas.unal.edu.co/index.php/actabiol

Facultad de Ciencias
Departamento de Biología
Sede Bogotá

UNIVERSIDAD NACIONAL DE COLOMBIA

ARTÍCULO DE INVESTIGACIÓN / RESEARCH ARTICLE

ECOLOGÍA

MODELLING OF 3D-STRUCTURES OF THE RARE MELANOCORTIN-1-RECEPTOR MUTATIONS ASSOCIATED TO MELANISM IN THE BANANAQUIT

Modelado de estructuras-3D de mutaciones raras del receptor-1-melanocortina asociadas al melanismo en el mielero

Raúl SEDANO-CRUZ[1]*, Daniel OSORIO[2]

[1]Facultad de Ciencias Naturales y Exactas, Departamento de Biología (Grupos de Investigación en Ecofisiología, Biogeografía y Evolución; y Ecología Animal), Universidad del Valle, Calle 13 No 100-00, Oficina 3120, Cali, Colombia.

[2]Escuela de Biología, Universidad Industrial de Santander, Cra. 27 Calle 9, Bucaramanga, Colombia.

For correspondence: raul.sedano@correounivalle.edu.co

Received: 15th August 2019, **Returned for revision**: 28th Novemeber 2019, **Accepted**: 23th December 2020.
Associate Editor: Edna Matta Camacho.

Citation/Citar este artículo como: Sedano-Cruz R, Osorio D. Modelado de estructuras-3D de mutaciones raras del receptor-1-melanocortina asociadas al melanismo en el mielero. Acta Biol Colomb. 2021;26(1):30-41. Doi: http://dx.doi.org/10.15446/abc.v26n1.81432

ABSTRACT

Melanism in plumage color is often associated to the single nucleotide polymorphism of the *melanocortin-1-receptor*. Despite the striking association between the substitution of a Glutamic-acid by for a Lysine at position 92 on the MC1R protein and a completely black plumage, an in-depth understanding of the effect of missense mutations on the conformational change and behavior of the MC1R in the lipid bilayer caused by the absence of a crystal structure is lacking. We examine the structural basis for receptor activation using DNA sequences from the GenBank to perform *in silico* protein homology-based modeling. Our tridimensional model shows that the Alanine for a 179-Threonine substitution is a structural complement of the charge-reversing effect associated to the substitution of a Glutamic-acid by for a Lysine at position 92 on the MC1R. We proposed the possibility of gradual evolution in stability and electrostatic properties of the MC1R by the sequential accumulation of these two rare substitutions. These two rare substitutions further perturb physical-chemical properties that may be necessary folding requirements of the constitutively active MC1R forms without altering of ligand binding affinity. The computational coarse-grained molecular dynamics of the MC1R binding affinities to the melanocyte-stimulating hormone predicted the disparity in ligand binding *among* alleles. We speculate that the disparity in structural constraints and ligand binding among the alleles within heterozygous individuals may contribute as a mechanism to the plumage color variation in the *Coereba flaveola*.

Keywords: birds, eumelanin, E92K, protein evolution, 3D-folding.

RESUMEN

El melanismo en el color del plumaje se asocia frecuentemente al polimorfismo del *receptor melanocortina-1*. La ausencia de una estructura cristalográfica de la asociación entre la sustitución del Glutamato por Lisina en la posición 92 de la proteína MC1R y el plumaje completamente negro, no ha permitido tener un mejor entendimiento del efecto de mutaciones no sinónimas en la conformación y en el comportamiento en la membrana del MC1R. Examinamos la estructura asociada a la activación del receptor usando secuencias de ADN obtenidas del GenBank, para un modelamiento *in silico* de formas homólogas de la proteína. El modelo tridimensional muestra que la sustitución de Alanina por la Treonina en la posición 179 es un complemento estructural al efecto de reversión de carga asociado a la sustitución del Glutamato por Lisina en la posición 92 del MC1R. Proponemos la posibilidad de evolución gradual de la estabilidad y de propiedades electrostáticas del MC1R por la acumulación de estas substituciones. Estas perturban las propiedades fisicoquímicas que podrían ser necesarias para el plegamiento de las formas constitutivamente activas del MC1R sin alterar la afinidad de empalme con el ligando. La modelación computacional de la dinámica molecular de la afinidad de empalme del MC1R a la hormona estimulante de meloncitos predice la disparidad de la unión con el ligando entre alelos. Consideramos que posiblemente la disparidad entre alelos en heterocigotos en cuanto a restricciones estructurales y la unión con el ligando podría contribuir a la variación en el color del plumaje en *Coereba flaveola*.

Palabras clave: aves, eumelanina, evolución de proteínas, E92K, proteína en torción-3D.

INTRODUCTION

Studies on protein structure using computational-based models are advancing our understanding of the complex acquisition of pigmentation on integumentary tissue in vertebrates. Single nucleotide polymorphism of the *melanocortin-1 receptor* (*MC1R*) has been identified in different vertebrates that are associated to plumage and hair melanism (Robbins *et al.,* 1993; Takeuchi *et al.,* 1996; Lu *et al.,* 1998; Kerje *et al.,* 2003; Ling *et al.,* 2003; Hoekstra *et al.,* 2006; Nadeau *et al.,* 2006; Beaumont *et al.,* 2007). In some cases, the identical-by-type amino-acid substitution is associated with pigmentation in multiple taxa (Theron *et al.,* 2001; Nadeau *et al.,* 2006; Cibois *et al.,* 2012), which suggests convergence of constraints in receptor functional structure on essential residues (Lu *et al.*, 1998). Many disease-like phenotypes are caused by a single amino-acid substitution on key genes because of the loss of stability of the proteins, protein misfolding and changes in the protein-ligand docking requirements (Honig and Nicholls, 1995; Wang and Moult, 2001; Goldberg, 2003).

A variety of functional roles have been documented for the amino acid position 92 of the MC1R, including darkening pigmentation of plumage associated to E92K (Takeuchi *et al.*, 1996; Theron *et al.*, 2001; Nadeau *et al.*, 2006; Cibois *et al.*, 2012), D92K and M92K substitutions (Jackson, 1997) in a number of Galliformes. But also E92K mutation is associated with the dark coat color of the *somber* mice (Robbins *et al.*, 1993) and V92K substitution is associated darker skin color and melasma in humans (Lu *et al.*, 1998; Suryaningsih *et al.*, 2019). Although there is evidence of differential gene regulation MC1R-independent for melanic plumage pigmentation (Abolins-Abols *et al.*, 2018; Krishnan and Cryberg, 2019), the parallel evolution of amino acid position 92 shows its significance in the regulatory mechanism of MC1R function for birds and also mammals. Thus, to address questions of the dependency of plumage pigmentation on the MC1R gene single nucleotide polymorphism (SNP), an understanding of the underlying protein structure is required to reveal possible mechanisms for the dysfunctions.

The MC1R protein is homologous to the large family of G protein-coupled receptors (GPCRs) (Palczewski *et al.*, 2000; Cherezov *et al.*, 2007). Similar to these GPCRs, the MC1R protein is composed of seven transmembrane helices, intracellular and extracellular loops, an N-terminus outside the cell, and an intracellular C-terminus (Ling *et al.*, 2003). The MC1R protein is expressed on the surface of melanocytes and controls the melanogenesis by binding with the *melanocyte-stimulating hormone* (α-MSH) (Jackson, 1997; Holder and Haskell-Luevano, 2004) which results in a conformational change and receptor activation for eumelanin deposition to darken the pigmentation (Lu *et al.*, 1998; Ling *et al.*, 2003). The *MC1R* gene loss-of-function mutations are associated with an increase in pheomelanin production, which results in red/yellow hair and plumage (Jackson, 1997; Kerje *et al.*, 2003). To date, the work on the adaptive nature of pigmentation has focused on the DNA coding region of the MC1R gene (Holder and Haskell-Luevano, 2004; Beaumont *et al.*, 2007), with little progress in understanding the effects of missense mutations on conformational change and receptor activation (Guruprasad *et al.*, 1990; Mundy, 2005; Janin and Sternberg, 2013), particularly in non-domesticated bird species.

The strongest evidence that SNP could increase the MC1R protein activity has been documented from the *Coereba flaveola* (Linnaeus, 1758) (Bananaquit) 21 years ago (Mundy, 2005). This neotropical passerine has two discrete morphs that are segregated by two rare autosomal E92K and A179T mutations that are associated with melanism, particularly in the birds on the islands of Grenada and St. Vincent in the Caribbean (Theron *et al.*, 2001). Thus, single amino-acid substitutions presumably locked the MC1R protein into an active state and hence the eumelanin synthesis to produce a completely melanic plumage (Theron *et al.*, 2001). Such phenotype that arose from missense mutations are presumed to be caused by defects in the structure of the receptor, ligand-binding affinities or intracellular signaling transmission (Guruprasad *et al.*, 1990; Wang and Moult, 2001; Janin and Sternberg, 2013). However, absent is an in-depth understanding of the effect of missense mutations on the conformational change and behavior of the MC1R protein with a ligand in the lipid bilayer.

The *in silico* three-dimensional protein homology-based modeling is an approach that is used to find structural information as an alternative to experimental techniques. The computational modeling of amino-acid sequences is an instrument to examine the proteins variants from SNP (Wang and Moult, 2001). The three-dimensional computational protein modeling uses the accessible data on homologous proteins and can provide a firm basis to attain reliable testimony of the structural information for GPCRs (Tan *et al.*, 2009; Shahlaei and Mousavi, 2014). In this study, we examined the *in silico* three-dimensional SNP predictive model of the MC1R forms of the Bananaquit to further describe the implications of rare mutations to the protein structure (Ibarrola-Villava *et al.*, 2014).

Our study integrated homology modeling and coarse-grained computational molecular dynamics to examine the MC1R binding affinities to the α-MSH. We first test the hypothesis of the importance of A179T substitution as a genetic complement to the E92K mutation on the structural characteristics of the MC1R (Theron *et al.*, 2001). Because the E92K mutation probably induces the constitutive activity of the MC1R in the Bananaquit (Theron *et al.*, 2001), and further abrogates the activation by the α-MSH (Lu *et al.*, 1998; Ling *et al.*, 2003), the constitutively active receptor do not need to bind α-MSH (Holder and Haskell-Luevano, 2004). Thus, we predicted that the melanic MC1R protein

forms had fewer three-dimensional requirements for the binding affinity to the α-MSH. Our 3D-model suggested an alternative mechanism in which regional eumelanin production on plumage might be related to the competitive affinity among MC1R alleles for the α-MSH. We introduce *in silico*-based evidence of the physical-chemical requirements of the MC1R protein, and we stress the evolutionary implications of the functional structure of the protein by the effects of rare substitutions on the transmembrane receptor MC1R of *C. flaveola*.

MATERIALS AND METHODS

Homology analysis

The thirty-one nucleotide sequences for the MC1R gene of the *Coereba flaveola* (GenBank accession numbers: AF362575 - AF362606) all of which correspond to 31 haplotypes previously defined by a set of SNPs in the open reading frame (Theron *et al.,* 2001) were translated into protein sequences with the seqinr R-package (Charif and Lobry, 2007). After duplicated amino acid sequences were removed, the seven unique proteins (Y1, Y2, Y4, Y10, M1, M4 and M5) (Appendix 1) were used to model three-dimensional (3D) structures. The geographic variation in 3D-structures and protein-ligand binding affinities among the populations in Panama (Central America), Puerto Rico (The Greater Antilles) and Grenada and St. Vincent (GSV, The Lesser Antilles) was compared using the published MC1R phylogeography for *C. flaveola* (Theron *et al.,* 2001).

In silico modeling

Because the experimental structure for each MC1R in *C. flaveola* was undetermined, the 3D-structure for each unique amino acid sequence (Y1, Y2, Y4, Y10, M5, M4 and M1) was generated. First, we performed a proteins search with high homology using the HHSearch at HHPredServer (Söding *et al.,* 2005), using the Protein Data Bank as a reference (Berman *et al.,* 2000). From the members of the family A of the GPCRs, three experimentally-determined structures defined as the templates to the MC1R showed 27-28 % amino-acid sequence identity and full coverage of protein regions included the Adenosine A2A Receptor (PDB ID: 3eml), the β-2-adrenergic receptor (PDB ID: 2rh1_A) and the Serotonin 5HT1B receptor (PDB ID: 4iaq). The MC1R and MSH structures were generated by I-TASSER 3D-protein modeler (Zhang, 2008) and the 3D-models were energy minimized with the Amber94 force field using the GROMACS package (Lindahl *et al.,* 2001). The quality of the energy minimized 3D-models was evaluated with the Z-score using the QMEAN server (Benkert *et al.,* 2008). For the *molecular dynamics* analysis with MSH as the ligand, we used the minimum fragment sequence ('MEHFRW') required for equipotency to the MSH (Holder and Haskell-Luevano, 2004).

Physical-chemical properties of 3D protein structures

The secondary and tertiary structures were aligned among allele forms (Y1, Y2, Y4, Y10, M1, M4 and M5) to examine the variation in 3D-folding. The comparative 3D-protein structural analysis was conducted using the Multiseq extension (Roberts *et al.,* 2006), as implemented in the Visual Molecular Dynamics package (Humphrey *et al.,* 1996). The physical and chemical properties were computed for all haplotypes sequences, which were assigned into the melanic 'M' and yellow 'Y' haplotypes groups (Theron *et al.,* 2001). Thus, for the Bananaquit's melanic and yellow haplotype groups we compared the average molecular weight, amino-acid length, electric charge and isoelectric point (Moore, 1985), GRAVY index (Kyte and Doolittle, 1982), and the instability index (Guruprasad *et al.,* 1990).

Because the well-documented dependency of a protein 3D-folding structure on voltaic charge and hydrophobicity of an amino acid sequence (Moore, 1985), we computed an index that take into account these parameters to further provide specifications of the 3D-folding propensity of an amino acid sequence as implemented in FoldIndex© (Uversky *et al.,* 2000). We implemented a Mann-Whitney's U test to evaluate whether or not our M and Y haplotypes groups were different in terms of 3D-folding propensity. We also computed the propensity in FoldIndex© for the first 50 residues of the MC1R, which correspond to the first extracellular element of the protein because this segment accumulated most helicoid and alpha helix structural variation within *C. flaveola*.

Analysis of mc1r mutants

To further examine the change by the point substitutions on the physical-chemical properties, we conducted one-by-one E92K and A179T substitutions on the consensus amino acid sequence for the yellow haplotypes (identity among Y alleles ranged 98.8-99.7 %, Appendix 2). For these mutants, we computed box-plots for the average amino-acid length, electric charge and isoelectric point (Moore, 1985), GRAVY index (Kyte and Doolittle, 1982), molecular weight and the instability index (Guruprasad *et al.,* 1990). The box-plots were useful to compare between these synthetic mutants and the haplotype variation in *C. flaveola,* and therefore we could propose possible mechanisms for the mutational effects on the MC1R constitutive activation.

Binding energy calculations

Through the HADDOCK web server using default parameters (De Vries *et al.,* 2010), a peptide-protein docking

analysis was conducted to establish the initial configuration for each of the selected MC1R 3D-structures (best models) and the MSH 3D-fragment. This preliminary docking configuration constituted a starting point to further simulate the *molecular dynamics*. The computationally demanding simulations of the *molecular dynamics* were performed using the Groningen Machine for Chemical Simulations (GROMACS 4.6.5) (Lindahl *et al.,* 2001), which was deployed and compiled for parallel processing on a node with 8 Graphic Processing Units TESLA M2050 in the GUANE cluster of the Universidad Industrial de Santander with the optimized parameters for the Amber94 force field (Ponder and Case, 2003). The addition of the appropriate number of counter ions of chloride neutralized the simulated molecular system. The energy minimization procedures and a positional restraint phase of 100 picoseconds (ps) were performed to relieve unfavorable interactions; a simulation of 100 ps was subsequently equilibrated with a canonical ensemble using a constant number and further regulation of either volume (V), pressure (P) or temperature (NVT and NPT), with P = 1 bar and T = 296 K, for 10 nanoseconds (ns), with a time step of 2 femtoseconds. The long-range electrostatics were treated with the Particle Mesh Ewald method (Darden *et al.,* 1993), using a 1.4 nanometers cut-off for the Lennard-Jones interactions (Lennard-Jones and Strachan, 1935). This configuration further allowed the calculations of full energy interaction for the protein (Berezovsky *et al.,* 1997). The pair list in the computational process was updated every ten steps during each simulation. Three independent 10 ns long *molecular dynamics* runs (30 ns production run) were performed to allow for improved conformational sampling and to have an average of the binding estimates of free energies. The free energy calculations for the peptide-protein docking for each independent simulation were performed with the post-processing GMXAPBS tool (Spiliotopoulos *et al.,* 2012).

RESULTS

Helicoid and α-helix structural variation

3D-model structure is highly conserved among the MC1R alleles (Fig. 1), with a wide range of interaction energy comparing the primary structure of the proteins and their spatial structure (Table 1). Furthermore, of the 314 amino-acid MC1R sequence, most helicoid and α-helix structural variation was within the first 50 amino acid residues of the MC1R (Fig. 2). But, the melanic haplotype group in *C. flaveola* was poorly predicted by the structural variation or the folding propensity of this first extracellular portion of the MC1R protein (data not shown). Thus, we examined if the variation between the haplotype groups may be associated to other physical and chemical properties of the entire MC1R protein, or related to energetic requirements for the binding affinity between the MSH and the MC1R protein.

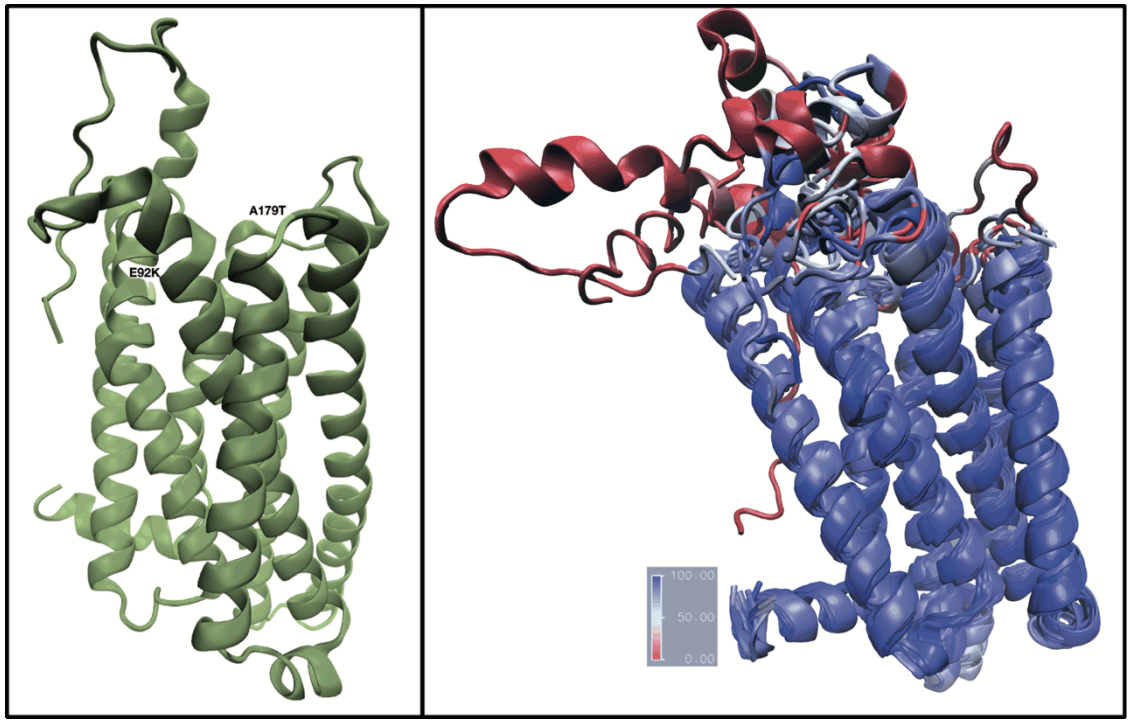

Figure 1. Structure of the MC1R. Left panel: sites subject to the E92K and A179T substitutions on the 3D-structural visualization of the MC1R in the Bananaquit. Right panel: intraspecific overlap of 3D-structural variation of the MC1R (see Table 1); the scale shows conserved and highly variable protein portions from blue (100 %) to red (0 %) respectively

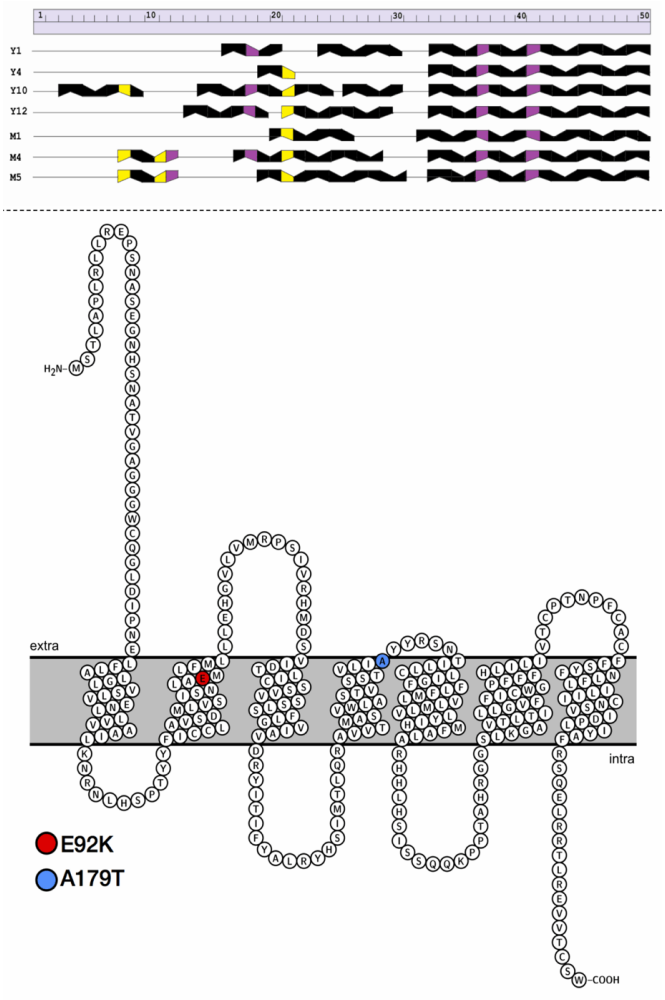

Figure 2. 2D-structural transmembrane location of the sites subject to the E92K and A179T substitutions on the MC1R (Bottom panel). Helicoid and alpha helix structural variation among haplotypes for the first 50 amino acids of the extra-cellular string of the MC1R in *C. flaveola* (Top panel).

Physical-chemical and folding structural properties

We found that molecular weight is not dependent on the E92K or the A179T mutations (Fig. 3); however, E92K and A179T mutations (Figs. 1 and 2) were estimated to be important for the variation of the MC1R structure by *in silico* analysis. An analysis of the dipeptide composition from the DNA sequences showed that the protein stability was highly dependent on the E92K and A179T substitutions (Fig. 3). These two residues likely contribute to a net gain in instability of the MC1R protein that is perfectly associated with melanic plumage. If the entire melanic morph were solely dependent on the constitutive activation caused by the E92K substitution as the current model suggests (Theron *et al.,* 2001), we also expected that a single substitution would explain most of the three-dimensional variation in protein structure. Although not pharmacologically characterized, the E92K mutation was estimated to be of structural importance by our *in silico* analyses. However, we also found evidence that the A179T substitution constituted an intragenic complement to E92K. Our 3D-model provided evidence of the charge-reversing effect of the E92K substitution (Fig. 3). The protein charge and isoelectric point are dependent on E92K, but not on A179T substitution, this further supports that the E92K residue participates in important electrostatic interactions. By contrast, a significant dependency of MC1R hydropathicity (GRAVY index) on A179T, but not on E92K substitution, suggested that a single substitution might explain some but not all the structural changes in the MC1R (Fig. 3). Before this study, the structural significance of the A179T substitution was undetermined because of the segregation among *ca* 5.7 % of yellow and 100 % of melanic morphs (Theron *et al.,* 2001).

As compared to the yellow MC1R forms, the melanic alleles are characterized by (i) higher protein instability, (ii) net charge and isoelectric point and (iii) lower hydropathicity (Fig. 3). Our 3D-folding structural analysis, as a combined measure of charge and hydrophobicity calculated in FoldIndex©, showed that the median folding propensity for melanic and yellow haplotypes groups were 0.485 and 0.492, respectively. We found a significant effect of haplotype group, the mean ranks of melanic and yellow haplotype groups were 3.0 and 12.5, respectively; $U = 0$, $Z = 3.1941$, $p < 0.05$. Thus, the high net charge suggested charge-charge repulsion combined with low GRAVY index minimally means a less compact structure in the melanic protein forms as compared to the yellow alleles (Table 1). The combination of high net charge and low mean hydrophobicity, as seen in the melanic protein forms in the Bananaquit, may represent a necessary prerequisite for the conformational changes in the structure of proteins under physiological conditions (Uversky *et al.,* 2000).

Binding affinity to α-MSH

The binding affinities to the MSH between MC1R alleles show a disparity range of -0.16 to -2.07 kcal/mol and minimally mean the competitive binding affinity within diplotypes. Although the disparity between alleles to bind to the MSH-peptide is equivalent between the melanic and yellow diplotype groups, a geographic pattern emerged. The heterozygous genotyped in Puerto Rico had the lowest binding disparity between alleles (mean = -0.10 Kcal/mol, n = 9) than those diplotypes in GSV (mean = -0.90 Kcal/mol, n = 17) and Panama (mean = -0.72 Kcal/mol, n = 6). The allele disparity in the affinity to bind to the MSH in Panama was primarily caused by the A179T substitution also present at GSV. Although the 179T haplotypes (Y4, Y9 and Y10) did not occur in Puerto Rico (Theron *et al.,* 2001), these haplotypes did contribute to the disparity in binding affinity within the diplotypes in GSV and Panama. This implies that

Table 1. Structural variation of the MC1R in terms of the 3D-folding propensity. The Folding index (-1 to 1) is based on the average residue hydrophobicity and the net protein charge, the interaction energy and the free energy available for the binding of the MC1R protein to the MSH-peptide are averaged from three independent computational simulations.

Alleles	Haplotypes	Residue E92K	Residue A179T	Folding Index	Full Interaction Energy (kcal/mol)	Average ΔG (kcal/mol)
Y1	Y2, Y3, Y5, Y6, Y7, Y8, Y11, Y13, Y14, Y15, Y16, Y17, Y18, Y19, Y20, Y21, Y22, Y23, Y24, Y25, Y26	E	A	0.492	-1033.8	-12.343
Y12	Y12	E	A	0.492	-1089.0	-13.273
Y4	Y9	E	T	0.489	-632.6	-12.183
Y10	Y10	E	T	0.493	-994.7	-14.417
M5	M2	K	T	0.485	-1098.1	-13.883
M4	M4	K	T	0.488	-933.7	-14.073
M1	M3	K	T	0.483	-1165.5	-14.237
TC	TC	E	A	0.493	-1131.7	-15.830

the disparity of affinities between alleles to bind to the MSH-peptide is constraint by the phylogeography of the MC1R gene of *C. flaveola*.

DISCUSSION

We characterized the 3D-structure of two rare missense mutations in the MC1R that were associated in a previous study with the heavy dark plumage pigmentation of Bananaquits (Theron *et al.,* 2001). Beyond the variation in DNA sequences, the E92K mutation is a polar by polar amino acid substitution in contrast to the A179K substitution that is a change of a charged amino acid by charged amino acid. Thus, we expected significant structural variation between the melanic and the yellow MC1R forms. Our 3D-folding calculations suggested that a less compact structure of the protein is associated to the active state of the melanic MC1R forms. In humans several rare MC1R haplotypes switch to a lower synthesis ratio of eumelanin/pheomelanin, resulting red-haired (Rees, 2000). In the Bananaquit the pattern is reversed, rare melanic haplotypes switch to a higher synthesis ratio of eumelanin/pheomelanin, resulting in a melanic plumage. Our *in silico* model predicts structural parallelism for the MC1R on the melanic morph in *Acrocephalus caffer,* a non-closely related *passerine* to the Bananaquit converging on an identical-by-type E92K substitution (Cibois *et al.,* 2012).

Our *in silico* analysis also predicted gains in the physical-chemical requirements caused by both the E92K and A179T mutations, which may constitute requirements to the gain-of-function for a constitutively active MC1R. Before this study, the coupling effect change of the E92K+A179T substitutions on physical-chemical properties of the MC1R was undetermined (Theron *et al.,* 2001). Because of the epistasis of the E92K and A179T residues for protein instability and low hydropathic change-effect, the implication was that a much larger fraction of the residues is likely constraint to protein stability than any other role of protein structure. The vast majority of known disease-causing missense mutations affect protein stability, through a variety of energy related factors (Wang and Moult, 2001). Here, computational homology modeling has provided a firm framework for the examination of rare amino-acid changes and the MC1R structural variation (Ibarrola-Villava *et al.,* 2014).

Our finding suggested that E92K and A179T carried a functional role at the expenses of the protein stability. Teilum *et al.* (2011) say that a positive correlation may seem a pragmatic expectation for the stability and protein-ligand affinity/function, "the more stable the machine, the better it works". But, in some cases the correlation is reversed; "the more stable the protein the less functional it is" (Teilum *et al.,* 2011). In the case of the Bananaquit, E92K and A179T substitutions perturb the stability of the native state of the MC1R without a significant altering of ligand binding affinities for all melanic and yellow haplotypes. Because the gain-of-function to the constitutively active MC1R presumably do not need to bind the α-MSH (Lu *et al.,* 1998; Ling *et al.,* 2003; Hoekstra 2006; Nadeau *et al.,* 2006; Benned-Jensen *et al.,* 2011), a reduced receptor coupling activity might not be the only contributing factor to the interaction between the MC1R and the phenotype (Beaumont *et al.,* 2005; Benned-Jensen *et al.,* 2011; Krishnan and Cryberg, 2019). For instance, the control of regional eumelanin production in the yellow morph of Bananaquits may be caused by the local variation in the α-MSH and/or in MC1R antagonist (Theron *et al.,* 2001).

However, our *in silico* 3D-model suggested a possible third mechanism in which regional eumelanin production in yellow Bananaquits might be related to the competitive MC1R allele affinity for the α-MSH (Ringholm *et al.,* 2004; Beaumont *et al.,* 2005). The melanic and yellow haplotype forms showed equivalent ranges of binding affinities to the MSH (Table 1). As compared to the melanic haplotypes,

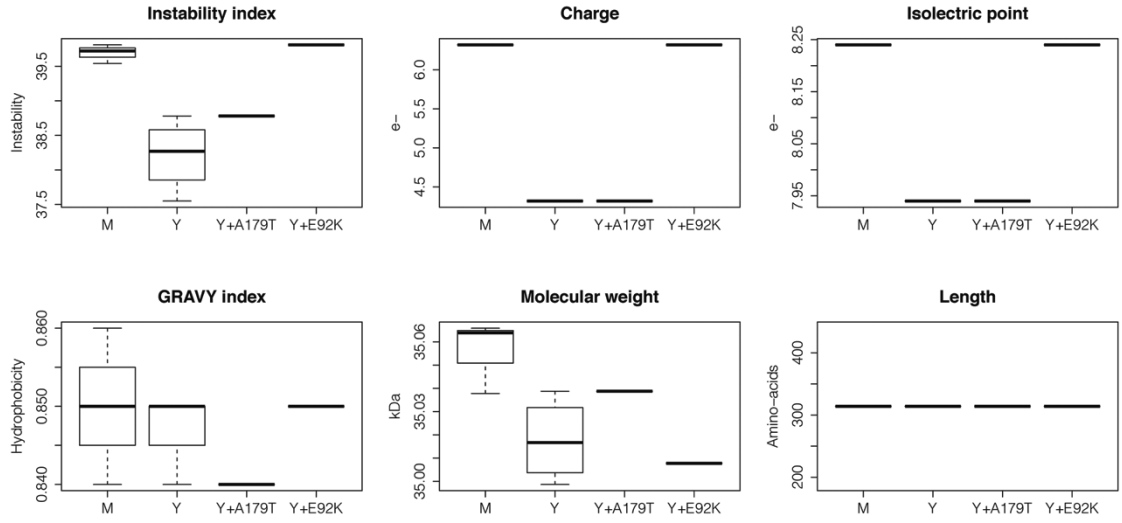

Figure 3. Physical and chemical properties of MC1R protein grouped by Melanic 'M' and Yellow 'Y' phenotypes. From left to right: The Instability index is a measure of how stable a protein might be under standard conditions, Charge is the whole protein charge determination, the Isoelectric point is the condition at which a protein carries no net charge and GRAVY index stands for Grand Average Hydropathy value. The effect change of point substitutions on Yellow haplotype physical-chemical properties was calculated by enforcing Yellow haplotypes with either E92K or A179T mutation.

the yellow 179T-haplotypes show the highest and lowest protein-ligand docking affinities to bind the MSH peptide (Table 1). Thus, the variation in docking affinity to the MSH among MC1R haplotypes is likely dependent on the accumulative effect of overall amino-acid sequence.

If our *molecular dynamics* model is correct, the disparity in binding affinity to the α-MSH will be unequal between alleles within A179T heterozygotes. Thus, we speculate that competitive binding affinity to the α-MSH within the yellow heterozygotes might play a role in the variation of regional plumage pigmentation in the *ca*. 41 subspecies of *C. flaveola* (Paynter, 1968). In addition, our model may be also an explanatory framework for recent findings on MC1R SNPs and a diplotype association with regional black and grey plumage traits in pigeons (Ran *et al.*, 2016). We encourage further experimental test on plumage pigmentation as a function of disparity in binding affinity between diplotypes as a future avenue for understanding the mechanism of regional eumelanin production in the Bananaquit.

Polarized genealogies have been used to study a variety of amino acid substitutions for e.g. spectral tuning of the *Opsin* genes (Jacobs, 2012); such cumulative contribution of multiple mutations somewhat supports gradual changes of the absorption spectra of cones photo-pigments among primates (Matsumoto *et al.*, 2014). In our case, Yellow haplotype ancestry for melanic alleles (Theron *et al.,* 2001) allows us to polarize the evolution of 3D-structural requirements of the MC1R protein. In the Bananaquit, the melanic haplotypes on GSV evolved from a first-order mutant A179T yellow, which was followed by a subsequent injection of the E92K substitution (melanic). Otherwise, both substitutions must have to evolve as a single mutational step. This later scenario suggested a punctuated gain in the protein structural requirements, whereas the former suggested a gradual evolution in protein instability. The gradual effect change began with the A179T substitution followed by the E92K mutation, which was further associated with a net gain in protein instability, a reversing change effect in electric charge and low hydropathical properties. Although the stepwise evolutionary model would assume that the A179T haplotypes on GSV had a continental origin (Theron *et al.,* 2001), we acknowledge that waves of yellow haplotypes from the Grater Antilles into GSV cannot be excluded (Bellemain *et al.,* 2012). To us this latter phylogeographic scenario is unlikely to explain the origin of the melanic morph on GSV from the Greater Antilles because it would require either the recent extirpation of 179T haplotypes from Puerto Rico and the Lesser Antilles, or would further require a more recent single-event to the achieve of the 179T + 92K fault in the melanic Bananaquit on GSV.

CONCLUSIONS

In absence of a crystal structure for the MC1R, SNP 3D-computational model is shown instrumental for understanding of the effect of missense mutations on the conformational change and *molecular dynamics* of the MC1R in *C. flaveola*. We provide evidence that the A179T substitution is a significant intragenic complement of the E92K substitution and both contribute to the variation in stability of the MC1R structure. Further, the analysis shows that the melanic MC1R alleles may have less structural-requirements to binding α-MSH as compared to yellow alleles. The *molecular dynamics* model shows disparity in binding affinity to the α-MSH between MC1R alleles within diplotypes associated to plumage melanism. It

stands to reason that this is an interesting mechanism to further examine intraspecific variation in regional melanic plumage pigmentation. Our *in silico* analysis couple with the advancement of the Bananaquit´s phylogeographic history in previous studies led us to suggest that the variation in structural requirements towards melanic MC1R is the result of gradual evolution. This is the first time SNP 3D-computational predictive model was used with an avian phenotype-associated protein. We hope to encourage others to accommodate SNP3d *in silico* predictive models to further examine the genetic and structural parallelism underlying the E92K-SNP interactions, in which there is a dependency of plumage pigmentation on the MC1R missense mutations.

ACKNOWLEDGMENTS

The *in silico* experiments presented in this paper were conducted using the GridUIS-2 platform of the Universidad Industrial de Santander (SC3UIS) High Performance and Scientific Computing Centre, Bucaramanga, Colombia. Thanks to Diana Gil for formatting this manuscript and also to J.F. Ortega at the digital image lab. Departamento de Biología, Universidad del Valle.

DISCLOSURE OF INTEREST

The authors declare that they have no conflicts of interest

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals performed by any of the authors

REFERENCES

Abolins-Abols M, Kornobis E, Ribeca P, Wakamatsu K, Peterson MP, Ketterson ED, *et al*. A role for differential gene regulation in the rapid diversification of melanic plumage coloration in the dark-eyed junco (*Junco hyemalis*). BioRxiv [Internet]. 2018(May):315762. Doi: https://doi.org/10.1101/315762

Goldberg AL. Protein degradation and protection against misfolded or damaged proteins. Nature. 2003;426:895–899. Doi: https://doi.org/10.1038/nature02263

Beaumont KA, Newton RA, Smit DJ, Leonard JH, Stow JL, Sturm RA. Altered cell surface expression of human MC1R variant receptor alleles associated with red hair and skin cancer risk. Hum Mol Genet. 2005;14(15):2145–2154. Doi: https://doi.org/10.1093/hmg/ddi219

Beaumont KA, Shekar SL, Newton RA, James MR, Stow JL, Duffy DL, *et al*., Receptor function, dominant negative activity and phenotype correlations for MC1R variant alleles. Hum Mol Genet. 2007;16(18):2249–2260. Doi: https://doi.org/10.1093/hmg/ddm177

Bellemain E, Gaggiotti OE, Fahey A, Bermingham E, Ricklefs RE. Demographic history and genetic diversity in West Indian *Coereba flaveola* populations. Genetica. 2012;140(4–6):137–148. Doi: https://doi.org/10.1007/s10709-012-9665-6

Benkert P, Tosatto SCE, Schomburg D. QMEAN: A comprehensive scoring function for model quality assessment. Proteins Struct Funct Genet. 2008;71(1):261–277. Doi: https://doi.org/10.1002/prot.21715

Benned-Jensen T, Mokrosinski J, Rosenkilde MM. The E92K melanocortin 1 receptor mutant induces camp production and arrestin recruitment but not ERK activity indicating biased constitutive signaling. PLoS One. 2011;6(9):e24644. Doi: https://doi.org/10.1371/journal.pone.0024644

Berezovsky IN, Tumanyan VG, Esipova NG. Representation of amino acid sequences in terms of interaction energy in protein globules. FEBS Lett. 1997;418(1-2):43–46. Doi: https://doi.org/10.1016/S0014-5793(97)01346-X

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, *et al*. The protein data bank. Nucleic Acids Res. 2000;28(1):235–242. Doi: https://doi.org/10.1093/nar/28.1.235

Charif D, Lobry JR. SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. 2007; p. 1–26. Doi: https://doi.org/10.1007/978-3-540-35306-5_10

Cherezov V, Hanson MA, Stevens RC, Rosenbaum DM, Rasmussen SGF, Foon ST, *et al*. High-resolution crystal structure of an engineered human β2-adrenergic G protein-coupled receptor. Science. 2007; 318(5854):1258–1265. Doi: https://doi.org/10.1126/science.1150577

Cibois A, Thibault JC, Pasquet E. The molecular basis of the plumage colour polymorphism in the Tahiti reed-warbler *Acrocephalus caffer*. J Avian Biol. 2012; 43(1): 3–8. Doi: https://doi.org/10.1111/j.1600-048X.2011.05546.x

Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys. 1993;98(12):10089–10092. Doi:https://doi.org/10.1063/1.464397

Guruprasad K, Reddy B, Pandit M. Correlation between the stability of a protein and its dipeptide composition : A novel approach for predicting in vivo stability from its primary structure. Protein Eng. 1990;4(2):155–161. Doi: https://doi.org/10.1093/protein/4.2.155

Hoekstra HE. Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity (Edinb). 2006;97:222–234. Doi: https://doi.org/10.1038/sj.hdy.6800861

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. A Single Amino Acid Mutation Contributes to Adaptive Beach Mouse Color Pattern. Science. 2006;313(5783):101–104. Doi: https://doi.org/10.1126/science.1126121

Holder JR, Haskell-Luevano C. Melanocortin Ligands: 30 Years of Structure-Activity Relationship (SAR) Studies. Med Res Rev. 2004;24(3):325–356. Doi: https://doi.org/10.1002/med.10064

Honig B, Nicholls A. Classical electrostatics in biology and chemistry. Science. 1995;268(5214):1144–1149. Doi: https://doi.org/10.1126/science.7761829

Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. J Mol Graph. 1996;14(1):33–38. Doi: https://doi.org/10.1016/0263-7855(96)00018-5

Ibarrola-Villava M, Peña-Chilet M, Llorca-Cardeñosa MJ, Oltra S, Cadenas CM, Bravo J, *et al*. Modeling MC1R rare variants: A structural evaluation of variants detected in a mediterranean case-control study. J Invest Dermatol. 2014;134(4):1146–1149. Doi: https://doi.org/10.1038/jid.2013.469

Jackson IJ. Homologous pigmentation mutations in human, mouse and other model organisms. Hum Mol Genet. 1997;6(10):1613–1624. Doi: https://doi.org/10.1093/hmg/6.10.1613

Jacobs GH. Losses of functional opsin genes, short-wavelength cone photopigments, and color vision - A significant trend in the evolution of mammalian vision. Vis Neurosci. 2012;30(1–2):39–53. Doi: https://doi.org/10.1017/S0952523812000429

Janin J, Sternberg MJE. Protein flexibility, not disorder, is intrinsic to molecular recognition. F1000 Biol Rep. 2013;5(2):1–7. Doi: https://doi.org/10.3410/b5-2

Kerje S, Lind J, Schutz K, Jensen P, Andersson L. Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken. Anim Genet. 2003;34(4):241–248. Doi: https://doi.org/10.1046/j.1365-2052.2003.00991.x

Krishnan S, Cryberg RL. Effects of mutations in pigeon Mc1r implicate an expanded plumage color patterning regulatory network. BioRxiv. 2019:1–36. Doi: https://doi.org/10.1101/792945

Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105–132. Doi: https://doi.org/10.1016/0022-2836(82)90515-0

Lennard-Jones JE, Strachan C. Material Forces and Configurational Forces in the Interaction of Elastic Singularities. Proc R Soc Lond A Math Phys Sci. 1935;150(870):442–455. Doi: https://doi.org/10.1098/rspa.1935.0114

Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. J Mol Model. 2001;7(8):306–317. Doi: https://doi.org/10.1007/S008940100045

Ling MK, Lagerström MC, Fredriksson R, Okimoto R, Mundy NI, Takeuchi S, *et al*. Association of feather colour with constitutively active melanocortin 1 receptors in chicken. Eur J Biochem. 2003;270(7):1441–1449. Doi: https://doi.org/10.1046/j.1432-1033.2003.03506.x

Lu D, Vage DI, Cone RD. A Ligand-Mimetic Model for Constitutive Activation of the Melanocortin-1 Receptor. Mol Endocrinol. 1998;12(4):592–604. Doi: https://doi.org/10.1210/mend.12.4.0091

Matsumoto Y, Hiramatsu C, Matsushita Y, Ozawa N, Ashino R, Nakata M, *et al*. Evolutionary renovation of L/M opsin polymorphism confers a fruit discrimination advantage to ateline New World monkeys. Mol Ecol. 2014;23(7):1799–1812. Doi: https://doi.org/10.1210/10.1111/mec.12703

Moore DS. Amino acid and peptide net charges: A simple calculational procedure. Biochem Educ. 1985;13(1):10–11. Doi: https://doi.org/10.1016/0307-4412(85)90114-1

Mundy NI. A window on the genetics of evolution: MC1R and plumage colouration in birds. Proc R Soc B Biol Sci. 2005;272(1573):1633–1640. Doi: https://doi.org/10.1098/rspb.2005.3107

Nadeau NJ, Minvielle F, Mundy NI. Association of a Glu92Lys substitution in MC1R with extended brown in Japanese quail (*Coturnixjaponica*). Anim Genet. 2006;37(3):287–289. Doi: https://doi.org/10.1111/j.1365-2052.2006.01442.x

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, *et al*. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. Science. 2000;289(5480):739–745. Doi: https://doi.org/10.1126/science.289.5480.739

Paynter RA. Check-list of birds of the world. [Internet]. vol XIV. Cambridge: Harvard University Press; 1968. p. 1-433. Disponible en: https://www.biodiversitylibrary.org/page/14481194#page/7/mode/1up

Ponder JW, Case DA. Force Fields For Protein Simulations. Adv Protein Chem. 2003; 66: 27–85. Doi: https://doi.org/10.1016/S0065-3233(03)66002-X

Ran J-S, You X-Y, Jin J, Zhou Y-G, Wang Y, Lan D, *et al*. The Relationship between MC1R Mutation and Plumage Color Variation in Pigeons. Biomed Res Int. Hindawi Publishing Corporation. 2016;2016:2–7. Doi: https://doi.org/10.1155/2016/3059756

Rees JL. Review - Pigment Gene Focus The Melanocortin 1 Receptor (MC1R): More Than Just Red Hair. Cell Res. 2000;13:135–140. Doi: https://doi.org/10.1034/j.1600-0749.2000.130303.x

Ringholm A, Klovins J, Rudzish R, Phillips S, Rees JL, Schiöth HB. Pharmacological characterization of loss of function mutations of the human melanocortin 1 receptor that are associated with red hair. J Invest Dermatol. 2004;123(5):917–923. Doi: https://doi.org/10.1111/j.0022-202X.2004.23444.x

Robbins LS, Nadeau JH, Johnson KR, Kelly MA, Roselli-rehfuss L, Baack E, *et al*. Pigmentation Phenotypes of Variant Extension Locus Alleles Result from Point MutationsThat Alter MSH Receptor Function. Cell. 1993;72(6):827–834. Doi: https://doi.org/10.1016/0092-8674(93)90572-8

Roberts E, Eargle J, Wright D, Luthey-Schulten Z. MultiSeq: Unifying sequence and structure data for evolutionary analysis. BMC Bioinformatics. 2006;7:1–11. Doi: https://doi.org/10.1186/1471-2105-7-382

Shahlaei M, Mousavi A. A 3D Model for human melanocortin 4 receptor refined with molecular dynamics simulation. J Reports Pharm Sci. 2014;3(1):42–53. Doi: https://doi.org/10.22110/jrps.v3i1.1677

Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 2005;33((Web Server issue):W244–W248. Doi: https://doi.org/10.1093/nar/gki408

Spiliotopoulos D, Spitaleri A, Musco G. Exploring PHD Fingers and H3K4me0 Interactions with Molecular Dynamics Simulations and Binding Free Energy Calculations: AIRE-PHD1, a Comparative Study. PLoS One. 2012;7(10):e46902. Doi: https://doi.org/10.1371/journal.pone.0046902

Suryaningsih BE, Sadewa AH, Wirohadidjojo YW, Soebono H. Association between heterozygote Val92Met MC1R gene polymorphisms with incidence of melasma: A study of Javanese women population in Yogyakarta. Clin Cosmet Investig Dermatol. 2019;12:489–495. Doi: https://doi.org/10.2147/CCID.S206115

Takeuchi S, Suzuki H, Yabuuchi M, Takahashi S. Possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken. Biochim Biophys Acta - Gene Struct Expr. 1996;1308(2):164–168. Doi: https://doi.org/10.1016/0167-4781(96)00100-5

Tan K, Pogozheva ID, Yeo GSH, Hadaschik D, Keogh JM, Haskell-Leuvano C, *et al*. Functional characterization and structural modeling of obesity associated mutations in the melanocortin 4 receptor. Endocrinology. 2009;150(1):114–125. Doi: https://doi.org/10.1210/en.2008-0721

Teilum K, Olsen JG, Kragelund BB. Protein stability, flexibility and function. Biochim Biophys Acta - Proteins Proteomics. Elsevier B.V.; 2011;1814(8):969–976. Doi: https://doi.org/10.1016/j.bbapap.2010.11.005

Theron E, Hawkins K, Bermingham E, Ricklefs RE, Mundy NI. The molecular basis of an avian plumage polymorphism in the wild. Curr Biol. 2001;11(8):550–557. Doi: https://doi.org/10.1016/s0960-9822(01)00158-0

Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins Struct Funct Genet. 2000;41(3):415–427. Doi: https://doi.org/10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7

De Vries SJ, Van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc. 2010;5:883–897. Doi: https://doi.org/10.1038/nprot.2010.32

Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat. 2001;17(4):263–270. Doi: https://doi.org/10.1002/humu.22

Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9(40):1–8. Doi: https://doi.org/10.1186/1471-2105-9-40

ANEXOS

```
CLUSTAL O(1.2.4) multiple sequence alignment


Y1       MSTLAPLRLLREPSNASEGNHSNATVGAGGGWCQGLDIPNELFLALGLVSLVENLLVVAA
Y4       MSTLAPLRLLREPSNASEGNHSNATVGAGGGWCQGLDIPNELFLALGLVSLVENLLVVAA
Y10      MSTLAALRLLREPSNASEGNHSNATVGAGGGWCQGLDIPNELFLALGLVSLVENLLVVAA
Y12      MSTLAPLRLLREPSNASEGNHSNATVGASGGWCQGLDIPNELFLALGLVSLVENLLVVAA
M1       MSTLAPLRLLREPSNASEGNHSNATVGAGGGWCQGLDIPNELFLALGLVSLVENLLVVAA
M4       MSTLAPLRLLREPSNASEGNHSNATVGAGGGWCQGLDIPNELFLALGLVSLVENLLVVAA
M5       MSTLAPLRLLREPSNASEGNHSNATVGAGGGWCQGLDIPNELFLALGLVSLVENLLVVAA
         ***** ********************.******************************

Y1       ILKNRNLHSPTYYFICCLAVSDMLVSISNLAEMLFMLLLEHGVLVMRPSIVRHMDSVIDT
Y4       VLKNRNLHSPTYYFICCLAVSDMLVSISNLAEMLFMLLLEHGVLVMRPSIVRHMDSVIDT
Y10      VLKNRNLHSPTYYFICCLAVSDMLVSISNLAEMLFMLLLEHGVLVMRPSIVRHMDSVIDT
Y12      ILKNRNLHSPTYYFICCLAVSDMLVSISNLAEMLFMLLLEHGVLVMRPSIVRHMDSVIDT
M1       ILKNRNLHSPTYYFICCLAVSDMLVSISNLAKMLFMLLLEHGVLVMRPSIVRHMDSVIDT
M4       ILKNRNLHSPTYYFICCLAVSDMLVSISNLAKMLFMLLLEHGVLVMRPSIVRHMDSVIDT
M5       ILKNRNLHSPTYYFICCLAVSDMLVSISNLVKMLFMLLLEHGVLVMRPSIVRHMDSVIDT
         :*****************************.:*****************************

Y1       LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMASVWLASTVSSTVLIAY
Y4       LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMASVWLASTVSSTVLITY
Y10      LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMASVWLASTVSSTVLITY
Y12      LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMASVWLASTVSSTVLIAY
M1       LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMASVWLASTVSSTVLITY
M4       LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMAIVWLASTVSSTVLITY
M5       LICSSVVSSLSFLGVIAVDRYITIFYALRYHSIMTLQRAVVTMASVWLASTVSSTVLITY
         ****************************************** ************:**

Y1       YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
Y4       YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
Y10      YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
Y12      YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
M1       YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
M4       YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
M5       YRSNTILLCLIGFFLFMLVLMLVLYIHMFALARHHLHSISSQQKPPTAHRGGSLKGAVTL
         ************************************************************

Y1       TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
Y4       TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
Y10      TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
Y12      TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
M1       TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
M4       TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
M5       TILLGVFFICWGPFFFHLILIVTCPTNPFCACFFSYFNLFLILIICNSVIDPLIYAFRSQ
         ************************************************************

Y1       ELRRTLREVVTCSW
Y4       ELRRTLREVVTCSW
Y10      ELRRTLREVVTCSW
Y12      ELRRTLREVVTCSW
M1       ELRRTLREVVTCSW
M4       ELRRTLREVVTCSW
M5       ELRRTLREVVTCSW
         **************
```

Appendix 1. Alignment of the amino acid sequence for the yellow and melanic MC1R protein forms of the Bananaquit. The asterisks show the consensus sequence by using ≥75% amino-acid identity but residue 92 in which case the 50% majority rule of residue identity is shown. Alignment conducted on the online platform EMBL-EBL using the option Clustal omega for multiple sequence alignment.

Appendix 2. Amino acid sequence identity (%) among MC1R haplotypes in *C. flaveola*.

	Y1	**Y12**	**Y4**	**Y10**	**M4**	**M1**	**M5**
Y1	100						
Y12	99.68	100					
Y4	99.36	99.04	100				
Y10	99.04	98.72	99.68	100			
M4	99.04	98.72	99.04	98.72	100		
M1	99.36	99.04	99.36	99.04	99.68	100	
M5	99.04	98.72	99.04	98.72	99.36	99.68	100