The genomics of mimicry: Gene expression throughout development provides insights into convergent and divergent phenotypes in a Müllerian mimicry system

Adam M M Stuckert, Mathieu Chouteau, Melanie Mcclure, Troy M Lapolice, Tyler Linderoth, Rasmus Nielsen, Kyle Summers, Matthew D Macmanes

The genomics of mimicry: gene expression throughout development provides insights into convergent and divergent phenotypes in a Müllerian mimicry system

Running title: The genomics of mimicry

Adam M M Stuckert[1,2,3]*, Mathieu Chouteau[4], Melanie McClure[4], Troy M LaPolice[2,5], Tyler Linderoth[6], Rasmus Nielsen[6], Kyle Summers[3], Matthew D MacManes[2]

[1]Department of Biology and Biochemistry, University of Houston

[2]Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire

[3]Department of Biology, East Carolina University

[4]Laboratoire Écologie, Évolution, Interactions des Systèmes Amazoniens (LEEISA), Université de Guyane, CNRS, IFREMER, 97300 Cayenne, France

[5]Department of Biology, Pennsylvania State University

[6]Department of Integrative Biology, University of California, Berkeley

*corresponding author: astuckert@uh.edu

Rationale:

We have updated this manuscript in light of the discovery that we received a combination of *Ranitomeya imitator* genomic reads and rat genomic reads. As a result, we have completely redone the entirety of the analyses in our previously published paper (Stuckert et al. 2021). We have re-assembled and re-annotated the genome, re-analyzed repeat content, and rerun our gene expression analyses. We have modified the content of the manuscript where appropriate (e.g., methods, results, discussion), but not areas of the manuscript that remain unchanged (e.g., the motivation and the introduction). We have tried to remain true to the initial manuscript and provide merely a correction to the record instead of additional analyses or updates. The results broadly hold, with small changes throughout. The main differences are that our newly assembled genome is shorter than the previous chimeric assembly (~6 Gbp vs 6.8 Gbp). One puzzling finding in our initial manuscript was the presence of a significant proportion of duplicated genes (~20%) that were supported with read data. We are now able to explain this result: we had genes present from 2 species. We also implemented a new pipeline to model and identify repeat content in the genome, and as a result we identified ~77% of the genome as repeats and we are able to annotate the vast majority of repeats. The results from the gene expression analyses are largely consistent between this version and the previously published version.

Abstract:

A common goal in evolutionary biology is to discern the mechanisms that produce the astounding diversity of morphologies seen across the tree of life. Aposematic species, those with a conspicuous phenotype coupled with some form of defense, are excellent models to understand

the link between vivid color pattern variations, the natural selection shaping it, and the underlying genetic mechanisms underpinning this variation. Mimicry systems in which species share a conspicuous phenotype can provide an even better model for understanding the mechanisms of color production in aposematic species, especially if comimics have divergent evolutionary histories. Here we investigate the genetic mechanisms by which mimicry is produced in poison frogs. We assembled a 6.02 Gbp genome with a contig N50 of 310 Kbp, a scaffold N50 of 390 Kbp, and 85% of expected tetrapod genes. We leveraged this genome to conduct gene expression analyses throughout development of four color morphs of *R. imitator* and two color morphs from both *R. fantastica* and *R. variabilis* which *R. imitator* mimics. We identified a large number of pigmentation and patterning genes differentially expressed throughout development, many of them related to melanophores/melanin, iridophore development, and guanine synthesis. We also identify the pteridine synthesis pathway (including genes such as *qdpr* and *xdh*) as a key driver of the variation in color between morphs of these species, and identify several plausible candidates for coloration in vertebrates (e.g., *cd36, ep-cadherin, perlwapin*). Finally, we hypothesize that keratin genes (e.g., *krt8*) are important for producing different structural colors within these frogs.

Keywords:

Amphibians, aposematism, color pattern, color production, Dendrobatidae, Ranitomeya

Introduction:

The diversity of animal coloration in the natural world has long been a focus of investigation in evolutionary biology (Gray and McKinnon 2007; Beddard 1892; Longley 1917; Fox 1936). Color phenotypes can be profoundly shaped by natural selection, sexual selection, or both.

Further, these color phenotypes are often under selection from multiple biotic (e.g. competition, predation) and abiotic (e.g. temperature, salinity) factors (Rudh and Qvarnström 2013). The mechanisms underlying color and pattern phenotypes are of general interest because they can help explain the occurrence of specific evolutionary patterns, particularly in systems where these phenotypes embody key adaptations driving biological diversification.

Adaptive radiations in aposematic species (those species which couple conspicuous phenotypes with a defense), provide examples of the effects of strong selection on such phenotypes (Sherratt 2008, 2006; Ruxton, Sherratt, and Speed 2004; Kang et al. 2017). In these biological systems, geographically heterogeneous predation produces rapid diversification of color and pattern within a species or group of species. This produces a diversity of polytypic phenotypes (defined as distinct defensive warning color signals in distinct localities) that are maintained geographically, with each population characterized by a unique phenotype that deters predators. This spatial mosaic of local adaptations maintained by the strong stabilizing selection exerted by predators also results in convergence of local warning signals in unrelated species. Examples of impressive diversification within species and mimetic convergence between species have been documented in many biological systems, including *Heliconius* butterflies (Mallet and Barton 1989), velvet ants (Wilson et al. 2015), millipedes (Marek and Bond 2009), and poison frogs (Symula, Schulte, and Summers 2001; Stuckert, Venegas, and Summers 2014; Stuckert et al. 2014; Twomey et al. 2013), to name only a few of the documented examples of diverse aposematic phenotypes (Briolat et al. 2019).

Understanding the genomic architecture underpinning these diversification events has been of substantial importance to the field of evolutionary biology (Hodges and Derieg 2009). Hence, it is essential to examine these diverse color pattern phenotypes and determine the

mechanisms by which both divergence and convergence occurs at the molecular level. The majority of our knowledge of the genomics of warning signals and mimicry comes from butterflies of the genus *Heliconius*. In these insects, a small number of key genetic loci of large phenotypic effect controls the totality of phenotypic variation observed within populations of the same species. These also provide the genetic underpinnings for mimetic convergence between distinct species (e.g., *WntA (Martin et al. 2012)* and *optix (Reed et al. 2011; Supple et al. 2013)*), though there are many others likely involved as well (Kronforst and Papa 2015). While *Heliconius* butterflies are excellent subjects for the study of mimicry, characterizing the genetics of mimicry in a phylogenetically distant system is critically important to determining whether adaptive radiations in warning signals and mimetic relationships are generally driven by a handful of key loci. Furthermore, given the dramatic differences in life history and social behavior between these clades, studying aposematism and mimicry in *Ranitomeya* is likely to provide general and important insights into convergent mimetic phenotypes and the evolutionary processes that produce them.

Here we investigate the genetics of convergent color phenotypes in *Ranitomeya* from Northern Peru. In this system, the mimic poison frog *(Ranitomeya imitator,* Dendrobatidae; (Schulte 1986) underwent a rapid adaptive radiation, such that it mimics established congenerics (*R. fantastica, R. summersi*, and *R. variabilis*), thereby "sharing" the cost of educating predators—a classic example of Müllerian mimicry (Symula, Schulte, and Summers 2003, 2001; Stuckert, Venegas, and Summers 2014; Stuckert et al. 2014). Previous analyses have indicated that the model species were already present prior to the ancestral *R. imitator* arriving in Peru (Symula et al. 2001). This is a powerful system for the evolutionary study of color patterns, as the different *R. imitator* color morphs have undergone an adaptive radiation to converge on shared phenotypes with the species that *R. imitator* mimics.

Furthermore, despite the arrival of new genomic data that provides critical insights into the mechanisms of color production in amphibians in general (Burgon et al. 2020), and poison frogs in particular (Rodríguez et al. 2020; Twomey, Johnson, et al. 2020; Stuckert et al. 2019), our knowledge of amphibian genomics is far behind that of other tetrapods, both in terms of genomic resources as well as in our ability to make inferences from these resources. This is largely due to the challenging nature of these genomes, most of which are extremely large (frog genomes range from 1-11Gb, with an interquartile range of 3-5Gb; Funk et al. 2018) and rich with repeat elements that have proliferated throughout the genome (Rogers et al. 2018). This makes many amphibians nearly-intractable systems for in-depth genomic analyses (Funk, Zamudio, and Crawford 2018; Rogers et al. 2018). As a result, there is a relative dearth of publicly-available amphibian genomes (14 anuran species as of September 1st, 2020). In fact, many of the available genomes are from a single group of frogs with a genome size of less than 1 Gbp, which is on the lower bound of known amphibian genome sizes (Funk, Zamudio, and Crawford 2018).

We investigated the genetics of Müllerian mimicry by first generating a high quality 6.8 Gbp *de novo* genome assembly for the mimic poison frog, *Ranitomeya imitator* (Figure 1). This is an important new resource for amphibian biologists, as it fills a substantial gap in the phylogenetic distribution of available amphibian genomes and enables more detailed comparative work. A comparison between our *R. imitator* genome and the *Oophaga pumilio* genome provides insights into genome evolution within the family Dendrobatidae, particularly the proliferation of repeat content. We highlight these results in this manuscript. We then utilized this high quality *Ranitomeya imitator* genome to examine gene expression patterns using RNA sequencing of skin tissue from early tadpole development all the way through to the end of metamorphosis in both the mimic (*R. imitator*) and model species (*R. fantastica* and *R. variabilis*). As such, we were able to

keep track of patterns of expression in genes responsible for color throughout development both between color morphs, and between species. We aimed to identify the genes responsible for color patterning that are convergent between model and mimic, as well as those genes whose role in color and pattern may be species-specific or even population-specific. Color patterns within these species begin to appear early during development when individuals are still tadpoles, which is consistent with observations that chromatophores (the structural elements in the integument that contain pigments) develop from the neural crest early during embryonic development (DuShane 1935). This comparative genomic approach allowed us to carefully examine the genes and gene networks responsible for diversification of color patterns and mimicry in poison frogs.

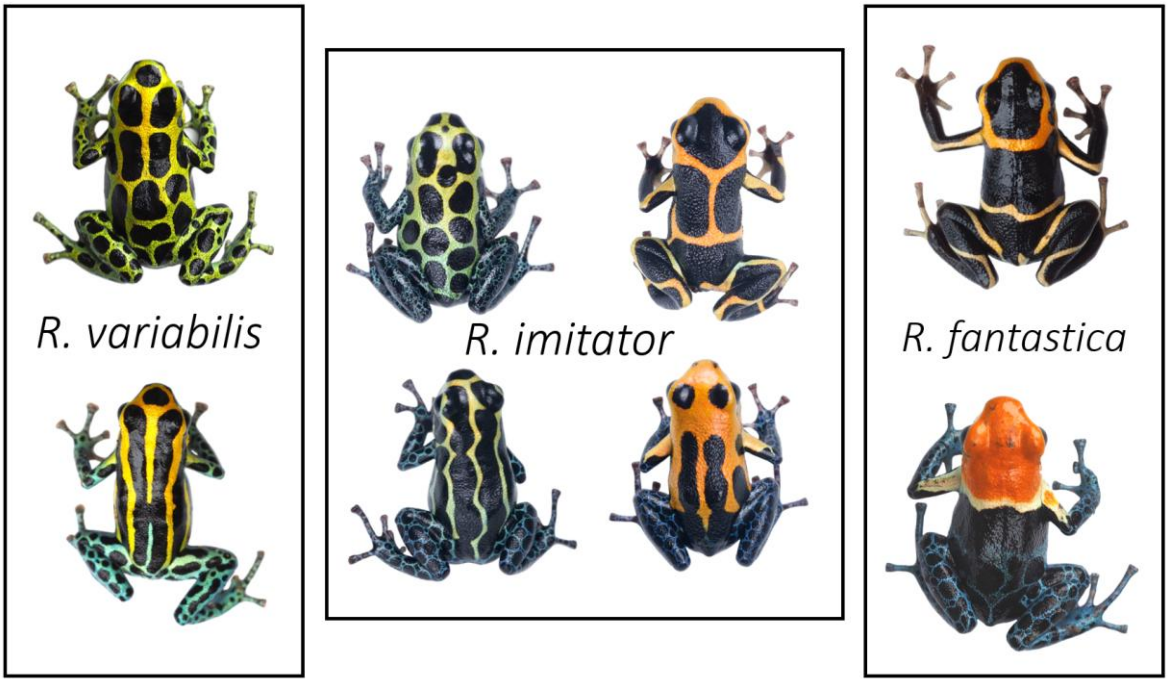


Figure 1. Normative photographs of frogs from the populations used in this study. The center box represents the four mimetic color morphs of *Ranitomeya imitator.* The left exterior box represents the two morphs of *R. variabilis* that *R. imitator* mimics, and the rightmost box represents two color morphs of *R. fantastica* that *R. imitator* has a convergent phenotype with. *R. imitator* photos by AS, *R. fantastica* and *R. variabilis* photos by MC.

Methods:

Permits:

Ranitomeya imitator:

Animal use and research comply with East Carolina University's IACUC (AUP #D281) and the University of New Hampshire's IACUC (AUP #180705).

Ranitomeya fantastica and *R. variabilis:*

The protocol for *R. fantastica* and *R. variabilis* sample collection was approved by the Peruvian Servicio Forestal y de Fauna Silvestre through the authorization number 232-2016-SERFOR/DGGSPFFS and export permit N° 17PE 001718 and the authorization from the French Direction de l'Environnement, de l'Agriculture, de l'Alimentation et de la forêt en Guyane number 973-ND0073/SP2000116-13.

Data accessibility:

All read data, our *de novo* genome assembly, and our annotations are archived with the European Nucleotide Archive (accession number PRJEB28312; https://www.ebi.ac.uk/ena/browser/view/PRJEB28312). The genome assembly is available at the European Nucleotide Archive. Additional transcript evidence from unpublished studies as well as the gff file used in analyses, and all code will be made available on DataDryad upon acceptance.

Generating a *de novo* genome for *Ranitomeya imitator:*

Genome Sequencing Approach:

Because all sequencing technologies are known to be biased in both known and unknown ways, we utilized a variety of sequencing technologies to assemble this complex and large genome.

172 At the time of genome construction, both sequencing technologies and assembly algorithms were
173 undergoing rapid change, and as such the generation of an optimal assembly required substantial
174 trial and error. For instance, several of the authors of this paper were involved in the strawberry
175 poison frog genome assembly Rogers et al. (2018) and the difficulties encountered in the attempt
176 to assemble that genome made it clear that short read data alone are insufficient if the goal is to
177 assemble a highly-contiguous and complete genome. To overcome the limitations associated with
178 short read data, we collected linked and long read data from technologies thought to be
179 complementary to one another (Illumina 10X, Oxford Nanopore, and Pacific Biosystems).
180
181 *10X Chromium:*
182 A single, likely male, subadult *R. imitator* of the 'intermedius' morph from the amphibian
183 pet trade was used to produce a 10X library. This individual had no visible ovaries, and sex
184 chromosomes in poison frogs are currently uncharacterized and therefore are not a mechanism
185 used to identify sex. This frog was euthanized and high molecular weight DNA was extracted from
186 liver tissue using the QIAGEN Blood & Cell Culture DNA Kit. 10X Genomics Chromium Genome
187 library (Weisenfeld et al. 2017) was prepared by the DNA Technologies and Expression Analysis
188 Cores at the University of California Davis Genome Center and sequenced on an Illumina HiSeq X by
189 Novogene Corporation (Mudd et al. in prep).
190
191 *Long read (Oxford Nanopore and Pacific BioSciences) sequencing:*
192 Captive bred subadult frogs from the pet trade that originated from the region near
193 Tarapoto, Peru (green-spotted morph, Figure 1) were euthanized and the skin and gastrointestinal
194 tract was removed in order to reduce potential contamination from skin and gut microbial

communities. To obtain the recommended mass of tissue for genomic DNA extraction, each frog was dissected into eight approximately equal chunks of tissue from the remaining portions of the whole-body and DNA was extracted using a Qiagen Genomic Tip extraction kit. DNA concentration was quantified with a Qubit 3.0 and fragment length was assessed with a TapeStation using a D1000 kit.

For Nanopore sequencing we prepared libraries for direct sequencing via Oxford Nanopore using a LSK-109 kit. Samples were loaded onto either R9 or R10 flowcells, which yielded minimum throughput. We basecalled raw fast5 files from Nanopore sequencing using the "read_fast5_basecaller.py" script in the ONT Albacore Sequencing Pipeline Software version 2.3.4.

For Pacific Biosystems (PacBio) sequencing we used a Circulomics short-read eliminator (Circulomics Inc, Baltimore, MD, USA) kit to size select extracted DNA from 10 Kb progressively up to 25 Kb. After this, we sent ~15 µg of high molecular weight DNA to the Genomics Core Facility in the Icahn School of Medicine at Mt. Sinai (New York, USA) for library preparations and sequencing. Here libraries were prepared with 20-25 Kb inserts and were sequenced on three SMRTcell 8M cells on a Pacific Biosciences Sequel II.

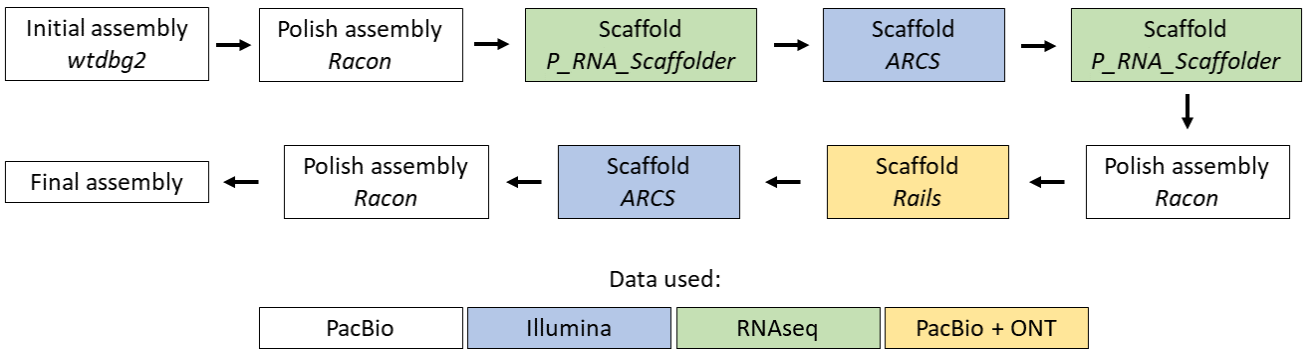


Figure 2. Flowchart of genome assembly approach. Colors of boxes represent the type of data used in that step (see internal legend) and italicized font indicate the program(s) used.

Genome assembly:

We took a multifaceted approach to constructing the *Ranitomeya imitator* genome, which contained iterative scaffolding steps. We detail our general approach here, and have a graphical depiction of this in Figure 2. PacBio has an algorithm to classify subreads that pass their minimum quality expectations for subreads as "good" subreads, and all subsequent analyses used only those reads passing this threshold. We used the contig assembler wtdbg2 version 2.5 *(Ruan and Li 2019)* to create our initial assembly and create consensus contigs using only the PacBio data. Wtdbg2 uses a fuzzy *de bruijn* graph approach to assemble reads into contigs. We corrected individual base errors in our assembly using Racon v1.4.19 (Vaser et al. 2017) with PacBio data aligned with Minimap2 (Li 2018) as the input.

Prior to conducting downstream gene expression analyses, we scaffolded our genome assembly with gene expression data in order to recover fragmented genic regions. To do so, we used the program P_RNA_Scaffolder (Zhu et al. 2018) with the RNA seq data used to assemble the transcript evidence as input data. This data was mapped to the reference assembly using BWA-MEM version 0.7.17-r1188 (Li and Durbin 2009) before running P_RNA_Scaffolder with default settings. We then used Arcs version 3.82 (Coombe et al. 2018) to scaffold our existing assembly using Illumina 10X data. Prior to doing this, we used the "basic" function within the program Longranger (Marks et al. 2019) to trim, error correct, and to identify barcodes in the 10X data. We then used the "arks.mk" makefile provided in the Arcs GitHub page (https://github.com/bcgsc/arcs/blob/master/Examples/arcs-make), to run the Arcs software. This makefile aligns the barcoded 10X data from Longranger using BWA, then runs Arcs to scaffold our

assembly. Following this, we ran an additional round of P_RNA_Scaffolder and then Racon with the same settings.

We then ran RAILS version 5.26.1 (Warren 2016) to scaffold again, with the same input data as for Cobbler. We aligned the long-read data with Minimap2 because it maps a higher proportion of long reads than other aligners and we used BWA on the short reads because it is more accurate for short-read data (Heng Li 2018). After this, we scaffolded the assembly again using Arcs, before completing one final round of polishing with Racon.

We examined genome quality in two main ways. First we examined the presence of genic content in our genome using Benchmarking Universal Single-Copy Orthologs v5.2.2 (BUSCO; (Simão et al. 2015) using the tetrapod database (tetrapoda_odb10; 2021-02-19). Our second method of genome examination was genome contiguity. We used the assemblathon perl script (https://github.com/KorfLab/Assemblathon/blob/master/assemblathon_stats.pl) to calculate overall numbers of scaffolds and contigs as well as contiguity metrics such as N50 for both.

Repeats and genome annotation:

We modeled genomic repeats with Repeat Modeler version 2 (Smit and Hubley 2008-2015) using RepBase database 20170127. Repeat families modeled by Repeat Modeler 2 were annotated using Transposon Classifier "RFSB" (Riehl et al. 2021). We extracted the vertebrate specific repeats from RepBase and merged this with the classified consensus output from Repat Modeler as input to Repeat Masker version 4.1.2-p1 (Smit, Hubley, and Green 2013-2015).

We annotated our genome using Maker version 3.01.02 (Campbell et al. 2014). We used transcript evidence that we produced from *R. imitator* to aid in assembly (“est2genome=1”l). We produced this transcript evidence from gene expression data of various *R. imitator* tissues and

populations which we assembled. We then used TransDecoder(https://github.com/TransDecoder/TransDecoder) and the Pfam database (Mistry et al. 2021; downloaded August 2021) to predict coding regions. For additional details on transcript evidence used to annotate this genome please see the Appendix. We also annotated the genome with protein evidence from UniProt (UniProt Consortium 2019) and the *Xenopus laevis* protein set ("protein2genome=1"). Following this first round of annotation, we did several rounds of gene prediction using SNAP (Korf 2004) and Augustus (Stanke and Waack 2003). We ran SNAP after our first round of annotation, and used the predictions as input into a second round of Maker annotation with the flag "snaphmm" in the control file. Within this second round of Maker, we ran Augustus gene predictions with species set as "human". Predictions using transcript and protein evidence were turned off. We ran 3 total rounds of annotation with gene prediction in this manner, then visualized both predicted gene length and Annotation Edit Distance (AED), a score which indicates perfect support for a predicted gene (0.0) to a complete lack of support (1.0). Predicted gene lengths dramatically increased using gene prediction, with a subsequent decrease in AED. After visualizing predicted gene length and AED scores we chose to use three total rounds of annotation, as the proportion of genes with "good" AED scores dropped precipitously after this. We also filtered out gene products with poor support (AED > 0.7) from our analyses.

IDENTIFICATION OF COLOR PATTERN CANDIDATE GENES

Gene expression sample preparation:

Samples were prepared differently for the mimic (*R. imitator*) and the model species (*R. fantastica* and *R. variabilis*). During the course of our work, we discovered that there were multiple

groups approaching the same questions using collected samples from different species, but at slightly different timepoints. In light of this, we chose to combine our efforts into a single manuscript in an attempt at making broader inferences. We acknowledge this, and as a result, the data in this manuscript are analyzed in a manner concordant with these differences.

Ranitomeya imitator:

The initial breeding stock of *Ranitomeya imitator* was purchased from Understory Enterprises, LLC (Chatham, Canada). Frogs used in this project represent captive-bred individuals sourced from the following wild populations going clockwise from top left in Figure 1: Tarapoto (green-spotted), Sauce (orange-banded), Varadero (redheaded), and Baja Huallaga (yellow-striped). Tadpoles were sacrificed for analyses at 2, 4, 7, and 8 weeks of age. We sequenced RNA from a minimum of three individuals at each time point from the Sauce, Tarapoto, and Varadero populations (except for Tarapoto at 8 weeks), and two individuals per time point from the Huallaga population. Individuals within the same time points were sampled from different family groups (Appendix Table 1).

Tadpoles were anesthetized with 20% benzocaine (Orajel), then sacrificed via pithing. Whole skin was removed and stored in RNA later (Ambion) at -20°C until RNA extraction. Whole skin was lysed using a BeadBug (Benchmark Scientific, Sayreville, NJ, USA), and RNA was then extracted using a standardized Trizol protocol. RNA was extracted from the whole skin using a standardized Trizol protocol, cleaned with DNAse and RNAsin, and purified using a Qiagen RNEasy mini kit. RNA Libraries were prepared using standard poly-A tail purification with Illumina primers, and individually barcoded using a New England Biolabs Ultra Directional kit as per the

manufacturer's protocol. Individually barcoded samples were pooled and sequenced using 50 bp paired end reads on three lanes of the Illumina HiSeq 2500 at the New York Genome Center.

Ranitomeya fantastica and *Ranitomeya variabilis*:

We set up a captive colony in Peru (see Appendix) consisting of between 6 and 10 wild collected individuals per locality. We raised the tadpoles on a diet consisting of a 50/50 mix of powdered spirulina and nettle, which they received 5 times a week. Tadpoles were raised individually in 21oz plastic containers, within outside insectaries covered with 50% shading cloth, and water change was performed with rainwater. Three tadpoles per stage (1, 2, 5, 7, and 8 weeks after hatching; see Appendix Table 1) were fixed in an RNAlater (Ambion) solution. To do so, tadpoles were first euthanized in a 250 mg/L benzocaine hydrochloride bath, then rinsed with distilled water before the whole tadpole was placed in RNAlater and stored at 4°C for 6h before being frozen at -20°C for long-term storage. Before RNA extraction, tadpoles were removed from RNA later and the skin was dissected off. Whole skin was lysed using a Bead Bug, and RNA was then extracted using a standardized Trizol protocol. RNA libraries were prepared using standard poly-A tail purification, prepared using Illumina primers, and individually dual-barcoded using a New England Biolabs Ultra Directional kit. Individually barcoded samples were pooled and sequenced on four lanes of an Illumina HiSeq X at NovoGene (California, USA). Reads were paired end and 150 base pairs in length.

Differential gene expression:

We indexed our new *Ranitomeya imitator* genome using STAR version2.7.10a_alpha_220601 (Dobin et al. 2013). We aligned our reads to our genome using STAR

version2.7.10a_alpha_220601 (Dobin et al. 2013), allowing 10 base mismatches (--outFilterMismatchNmax 10), a maximum of 5 multiple alignments per read (--outFilterMultimapNmax 5), and discarding reads that mapped at less than 50% of the read length (--outFilterScoreMinOverLread 0.5). We then counted aligned reads using htseq-count version 2.0.2 (Anders, Pyl, and Huber 2015).

Differential expression analyses were conducted in R version 4.2.2 (Team 2019) using the package DESeq2 version 1.38.3 (Love, Anders, and Huber 2014). Some genes in our annotated genome are represented multiple times, and thus the alignment is nearly to gene level with some exceptions. As a result, when we imported data into R we corrected for this by merging counts from htseq-count into a gene-level count. We filtered out low expression genes by removing any gene with a total experiment-wide expression level ≤ 50 total counts. cDNA libraries for *R. imitator* were sequenced at a different core facility than those of *R. fantastica* and *R. variabilis*, so in order to statistically account for batch effects we analyzed the data from each species independently (combining all species and batch effects in our dataset produces rank deficient models). For each species we compared two models using Likelihood Ratio Tests, one which tested the effect of color morph and the other which tested the effect of developmental stage. Both models included sequencing lane, developmental stage, and color morph as fixed effects. We used a Benjamini and Hochberg (Benjamini and Hochberg 1995) correction for multiple comparisons and used an alpha value of 0.01 for significance. We then extracted data from our models for particular *a priori* color genes that play a role in color or pattern production in other taxa. This *a priori* list was originally used in Stuckert et al. (Stuckert et al. 2019), but was updated by searching for genes that have been implicated in coloration in genomics studies from the last three years. Plots in this manuscript were produced using ggplot2 (Wickham 2011).

Finally, we ran an analysis with the specific intent of identifying genes involved in the production of different color morphs that are convergent between model (*R. fantastica* or *R. variabilis*) and mimic (*R. imitator*). To do this we conducted a Walds test within species between the spotted and striped morph of *R. imitator* and *R. variabilis* that incorporated sequencing lane, tadpole age, and color morph as fixed effects. We then identified the set of genes that are differentially expressed between color morphs in both species, as well as those that showed species-specific patterns. We did this same within species comparison using the banded and redheaded morphs of *R. imitator* and *R. fantastica*. In order to further elucidate potential genes that may influence convergent and divergent phenotypes in multiple species, we examined the list of all differentially expressed genes between color morphs in *R. imitator* (via the Likelihood Ratio Test described above) for genes that are differentially expressed between color morphs in *R. fantastica* or *R. variabilis* (via the Walds tests we conducted).

Results:

Ranitomeya imitator Genome:

Genome assembly:

Our final genome assembly was 6.02 Gbp in length and consisted of 79,800 scaffolds ranging from 387-5,545,867 bp in length with a scaffold N50 of 389,727 bp (83,475 total contigs with an N50 of 310,614 bp, ranging from 387-4,395,411 bp). A total of 6,641 contigs were placed into scaffolds by our iterative scaffolding and gap-filling. Based on our BUSCO analysis, the final genome contained 84.8% of expected tetrapod genes. We assembled 83.0% single copy orthologs and 1.8% duplicated orthologs. An additional 4.9% were fragmented and 10.3% were missing.

Repetitive *elements:*

Our analyses indicate a high proportion of repeats in the *R. imitator* genome. Repeat Modeler masked 75.69% of total bases in the genome, of which 74.63% consisted of repeat elements (see Table 1). Many of the repeats were retroelements (30.77% of the genome), 7.15% of which were LINEs. Over 23% of the genome is LTR elements, including 20.44% of the genome in GYPSY/DIRS1 repeats. We were able to classify the vast majority of repeats, with only 0.06% of repeats remaining unclassified. Two repeat types represent >1Gbp in total genomic content: hAT (1.13 Gbp) and Gypsy (1.2 Gbp). Given the quality of repeat databases and the scarcity of amphibian genomic resources in these databases, our results likely represent an underrepresentation of repeats in the genome as a whole. An unknown proportion of the genome's repeat elements are likely to be unassembled and missing in the genome, contributing to our underestimation of repeat content. Summary statistics on the number of instances, range, average length, and standard deviation of range can be found in Supplemental Table 2.

Element type			Number of elements	Total length of elements (bp)	Percentage of genomic sequence
Retroelements			2940777	1853517364	30.77
	SINEs		9140	1074368	0.02
	Penelope		155182	74628225	1.24
	LINEs		604697	430550850	7.15
		CRE/SLACS	0	0	00
		L2/CR1/Rex	355946	294429256	4.89
		R1/LOA/Jockey	268	27552	0.00
		R2/R4/NeSL	10535	2162443	0.04
		RTE/Bov-B	16036	8907551	0.15
		L1/CIN4	64839	46701023	0.78
	LTR elements		2326940	1421892146	23.61

		BEL/Pao	14889	16156219	0.27
		Ty1/Copia	335291	90264134	1.50
		Gypsy/DIRS1	1755793	1231339651	20.44
		Retroviral	122545	50623347	0.84
DNA transposons			8305215	2638017678	43.79
	hobo-Activator		3302857	1409227847	23.39
	Tc1-IS630-Pogo		894738	463756460	7.70
	En-Spm		0	0	0.0
	MULE-MuDR		3713	183543	0.00
	PiggyBac		30726	19118673	0.32
	Tourist/Harbinger		68929	21248842	0.35
	Other (Mirage, P-element, Transib)		365851	61666857	1.02
Rolling-circles			12894	5925694	0.10
Unclassified			16324	3669707	0.06
Total interspersed repeats				4495204749	74.63
Small RNA			6867	775734	0.01
Satellites			7136	3040805	0.05
Simple repeats			669788	47935402	0.80
Low complexity			81259	7023995	0.12

Table 1. Repeat elements classified by Repeat Masker. Most repeats fragmented by insertions or deletions were counted as a single element. Classification of repeats was done with RepeatMasker version 4.1.4 using rmblastn version 2.13.0+ with a custom database of known vertebrate repeats and modeled repeats from *R. imitator.*

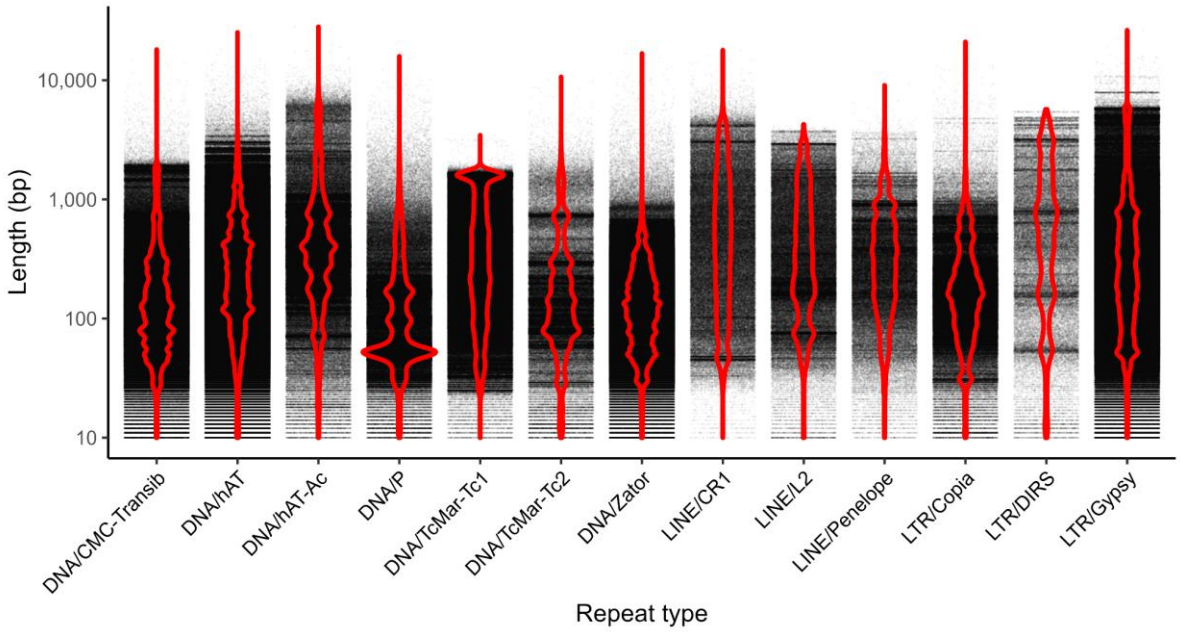


Figure 4. A violin plot of the length of individual repeats in the *R. imitator* genome. Each point represents a single instance of the repeat. Only repeats with >50 Mbp of total content are plotted.

Gene expression:

We aligned an average of 23.6 million reads (± 6.5 sd) per sample. On average, 72.78% of reads were uniquely mapped (± 5.39% sd). Mapping rates were slightly higher in *R. imitator* than in *R. fantastica* or *R. variabilis*, because the libraries were of slightly higher quality in *R. imitator.* To further test if this was an artifact derived from mapping reads from other species to the *R. imitator* genome, we also mapped these reads to species-specific transcriptome assemblies. We assembled these using data for each species in this study using the Oyster River Protocol (MacManes 2017); additional details on this protocol can be found in the appendix. After mapping our read data to these species-specific transcriptome assemblies we found similar results to that of our genome-guided mapping. Thus, our mapping rates are driven primarily by the slightly lower quality of the *R. fantastica* and *R. variabilis* cDNA libraries (as exhibited by slightly lower Phred scores towards the 3' end of reads) and not by species-specific differences in coding regions. For data on the number of

reads and mapping rates in each sample please see (Supplemental Table 3). All gene expression count data can be found on DataDryad. Patterns of gene expression are largely driven by developmental stage (principal component 1; 42% of variation) and species (principal component 2; 13% of variance; see Figure 6). We note that differences in sample preparation and sequencing localities between *R. imitator and R. fantastica/variabilis* may be driving a portion of the variation. However, the pattern found in principal component 2 closely parallels phylogeny, as *R. fantastica* and *R. variabilis* are more closely related to each other than either are to *R. imitator* (Brown et al. 2011). We then conducted a test for the effect of color morph and developmental stage for each species independently. For a list of all differentially expressed color genes see Supplemental Table 4. For a list of all differentially expressed genes at alpha < 0.01 see Supplemental Table 5.

Between developmental stage comparisons:

In our comparison of developmental stages we found many differentially expressed genes (q value < 0.01) in each species (*R. imitator* = 2,264, *R. fantastica* = 2,350, *R. variabilis* = 3,016; Table 2). Most of these are unlikely to be related to color and patterning, although a small fraction of differentially expressed genes (average 3.7%) are found in our *a priori* list of genes that influence the generation of color or patterning in other taxa (*R. imitator* = 91, *R. fantastica* = 95, *R. variabilis* = 90). Amongst genes that were significantly differentially expressed between developmental stages we identified genes related to carotenoid metabolism (e.g., *bco1, retsat, scarb1,* Figure 6), the synthesis of pteridines (e.g., *gchfr, qdpr, xdh;* Figure 6), genes related to melanophore development and melanin synthesis (*dct, kit, mitf, mlph, notch1, notch2, sfxn1, sox9, sox10, tyr*, and *tyrp1;* Figure 7), genes putatively related to the production of iridophores and their guanine

platelets (e.g., *gart, gas1, paics, pax3-a, rab27a, rab27b, rab7b,* Figure 8), and genes related to patterning (*notch1, notch2*).

Between morph comparisons:

In our comparison of color morph we found many significantly differentially expressed genes in each species (*R. imitator* = 1,558, *R. fantastica* = 1,266, *R. variabilis* = 1,168; Table 2). Most of these are unlikely to be related to color and patterning, although a small fraction of differentially expressed genes (average 2.5%) are related to the generation of color or patterning in other taxa (*R. imitator* = 33, *R. fantastica* = 36, *R. variabilis* = 29). Amongst genes that were differentially expressed between color morphs we identified genes related to carotenoid metabolism or xanthophore production (e.g., *aldh1a1, pax7, scarb1;* Figure 6), the synthesis of pteridines (e.g., *gchfr, qdpr;* Figure 6), genes related to melanophore development and melanin synthesis (*kit, mlph, sfxn1, sox9*), and genes putatively related to the production of iridophores and their guanine platelets (e.g., *atic, dock7, gart, paics, pacx2, pax3-a, rab27a, rab27b, rab7a, rabggta;* Figure 8).

Convergent gene expression patterns between model and mimic:

We compared gene expression between morphs for which we could make a direct comparison of expression between species with shared convergent phenotypes (i.e., striped vs spotted *R. imitator* and *R. variabilis*, banded vs redheaded *R. imitator* and *R. fantastica*). In the striped vs spotted comparison, we identified 345 differentially expressed genes in *R. imitator* and 1,407 in *R. variabilis*. Of these genes, 55 were shared between the two species and 10 were in our *a priori* color gene list (*cd36, crabp2, edaradd, mvd*). In the banded vs redheaded comparison, we identified 1,025 differentially expressed genes in *R. imitator* and 1,488 in *R. fantastica*. Of these

genes, 141 were shared between the two species and one was in our a priori color gene list (*qdpr*). The genes that are differentially expressed between morphs in both model and mimic are good candidates for future investigation, particularly if they fall out as important in other studies. To identify additional good candidates, we examined the top 10 genes by log fold change in each mimetic pair. There are several genes among these that seem promising candidates based on their log-fold changes between morphs and roles in other organisms (e.g., ep-*cadherin*, several genes in the cyp2a family [*cyp2f5, cyp2j5, cyp2c13*], *vat1l, selenoi, olfactory protein, piwil1, saa3, pdc, ddb_g0268948, cd36, ovol1,* and *perlwapin*. We briefly discuss these genes and their potential import in the Supplemental Discussion.

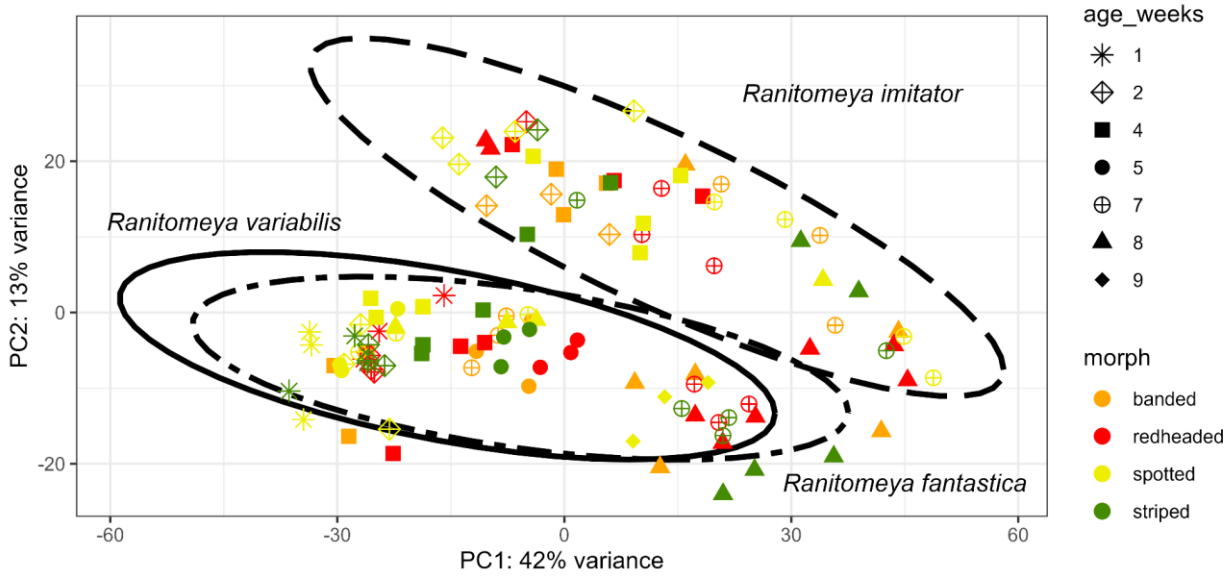


Figure 5. Principal component analysis of gene expression. The axes are labeled with the proportion of the data explained by principal components 1 and 2.

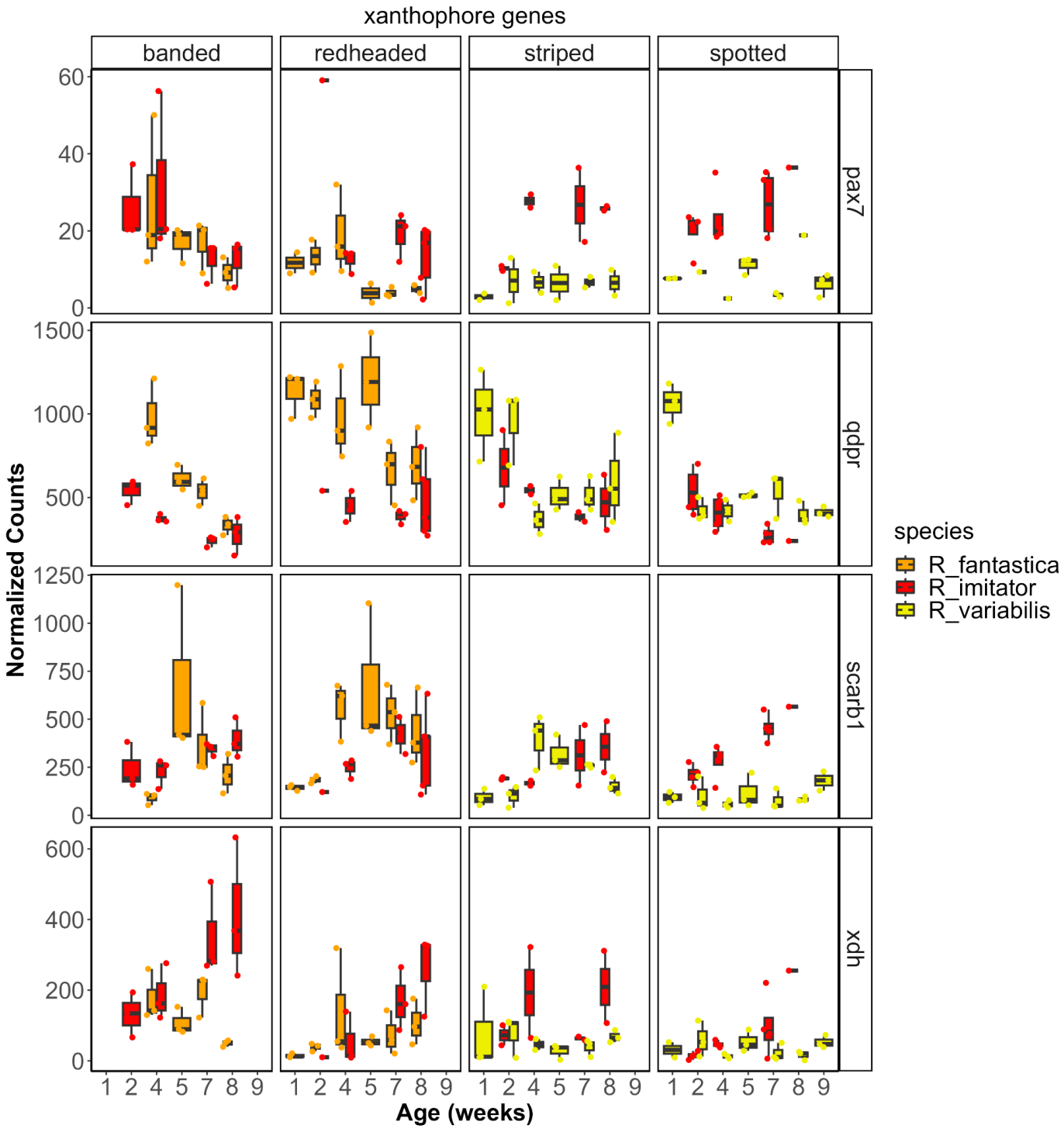
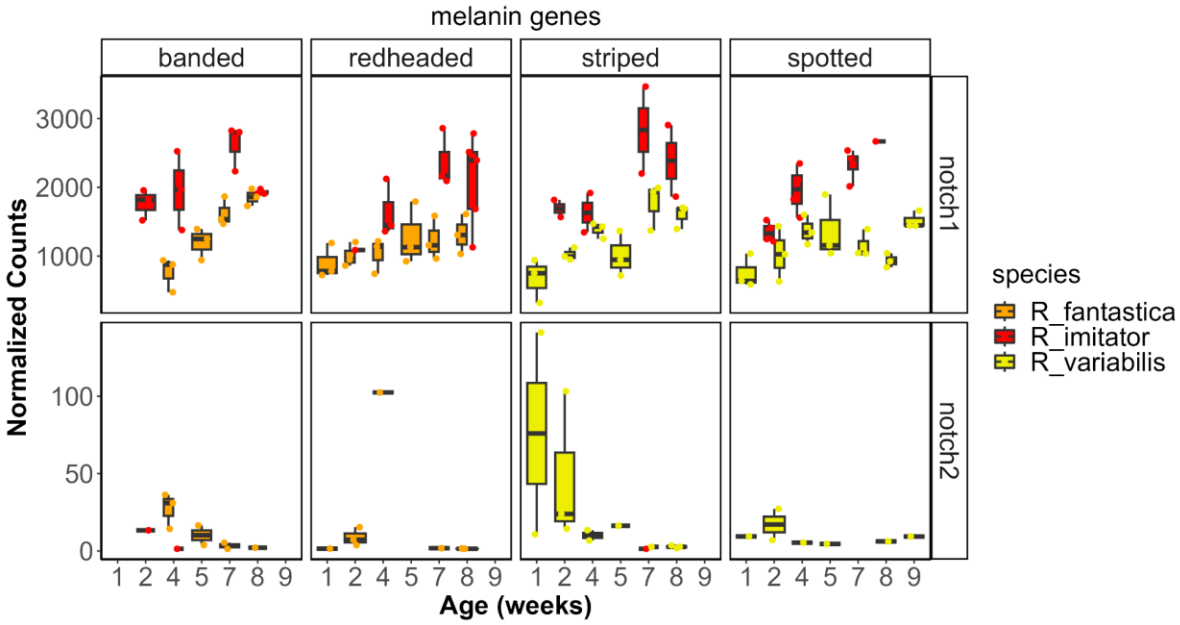


Figure 6. Gene expression for select carotenoid and pteridine genes.



475

476 Figure 7. Gene expression for select melanin genes.

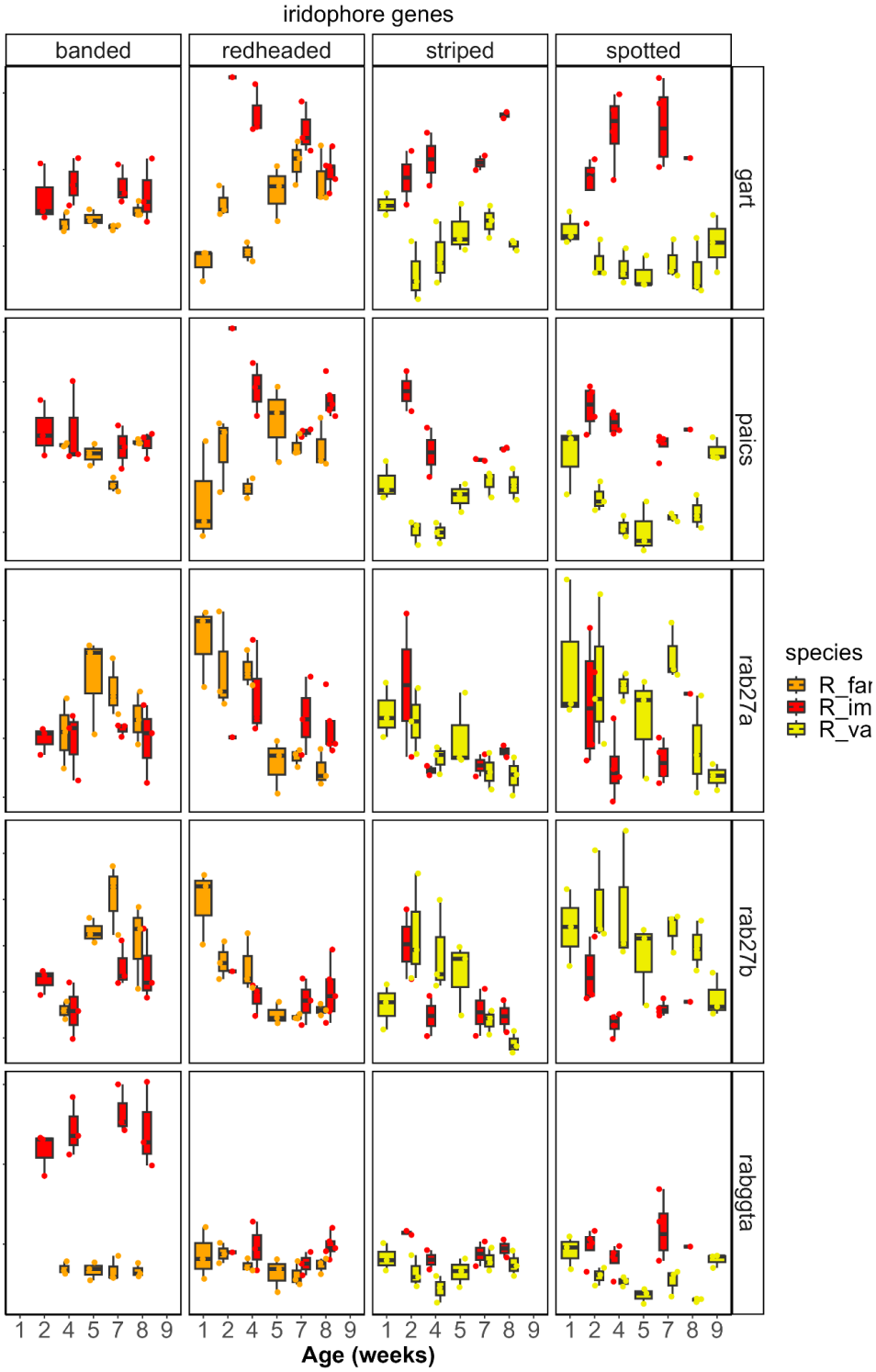


Figure 8. Gene expression for select iridophore genes.

Discussion:

The genetic, biochemical, cellular, physiological and morphological mechanisms that control coloration in mimetic systems are of interest because of their substantial impacts on survival and reproductive success. Despite this, these mechanisms are poorly characterized. Further, genetic mechanisms and genomic resources in amphibians are limited and poorly understood, particularly compared to better known groups like mammals and fish. In this study, we examined how gene expression contributes to differential phenotypes within species in a Müllerian mimicry complex of poison frogs. To do this we assembled a high-quality genome for the mimic poison frog *Ranitomeya imitator*, which we leveraged to conduct gene expression analyses. Here we describe the resulting *R. imitator* genome assembly and highlight key pathways and genes that likely contribute to differential color production within species, illuminating the mechanisms underlying Müllerian mimicry, and providing a rich foundation upon which future research may be built.

Genome:

Our newly assembled *Ranitomeya imitator* genome is a large, high-quality genome. This assembled genome is 6.02 Gbp in length and contains 85% of the expected genes according to our BUSCO results. Further, our genome is relatively contiguous with a contig N50 of over 310 Kbp. For comparison, the *O. pumilio* genome produced with short read technologies had a contig N50 of 385 base pairs and many genic regions were not assembled, presumably because of long intronic regions interspersed with repeat elements (Rogers et al. 2018). This dramatic difference in genome

504 contiguity and genic content indicates that long read technologies are, unsurprisingly, critically
505 important and can produce genomes with contiguity spanning large regions, even for species with
506 large genomes containing many long, repetitive regions. Further, the relatively high error rate of
507 current long read technologies does not preclude the ability to assemble and identify genes, as
508 evidenced by our high BUSCO score. This genome is a valuable resource and is well-suited to a
509 variety of future work, especially RNA sequencing analyses like those we present below.
510
511 *Repeat elements in the* Ranitomeya imitator *genome:*
512 Over 77% of our genome assembly consists of repeats. This is a larger proportion of the
513 genome than that found in the strawberry poison frog (*O. pumilio*), which was estimated to consist
514 of ~70% repeats (Rogers et al. 2018). However, this is almost certainly due to the availability of
515 long-reads in our study. Given that the *O. pumilio* assembly is both larger, and built from short-
516 reads, their identification of repeat elements is likely an under-estimate of total repeat content. In
517 comparison to these two poison frog genomes, the genome of *Xenopus tropicalis* is a little over
518 one-third repeat elements and *Nanorana parkeri* is ~48% repeats, which are both comparable to a
519 mammalian genome (Hellsten et al. 2010; Sun et al. 2015). Scattered throughout the *R. imitator*
520 genome are a large number of repeat elements which represent over half of the assembled
521 genome. We found that some classes of repeat elements contained similar abundance in the *R.*
522 *imitator* and *O. pumilio* assemblies (e.g., Gypsy: 1.2 Gbp in *R. imitator* and 1 Gbp in *O. pumilio*),
523 whereas several other repeat classes had strikingly different abundance (e.g., hAT: 1.1 Gbp in *R.*
524 *imitator* and 255 Mbp in *O. pumilio;* Copia: 97 Mbp in *R. imitator* and 298 Mbp in *O. pumilio*). This
525 indicates that repeat elements have undergone different patterns of expansion or contraction
526 within the family Dendrobatidae, and follow-up investigations into the evolution of repeat

elements may be particularly useful for understanding patterns of genome size within amphibians. However, there is a large caveat to this conclusion: our two studies used very different input data, assembly algorithms, and methods of analyzing repeats (RepDeNovo in *O. pumilio* and RepeatMasker in *R.* imitator) which are sources of potential bias.

Gene expression:

While the colors and patterns of *Ranitomeya* poison frogs are extremely variable, they always consist of vivid color patches overlaying a background that is largely black. Recent evidence indicates that much of the differences in color in poison frogs are derived from the structure (thickness) and orientation of iridophore platelets (Twomey, Kain, et al. 2020). Additionally, specific pigments that are deposited in the xanthophores, such as pteridines and carotenoids, interact with these structural elements to influence integumental coloration in the yellow to red ranges of the visible spectrum. Black and brown coloration is produced by melanophores and the melanin pigments found within the chromatophores (Bagnara et al. 1968; Duellman and Trueb 1986. These data are corroborated by new genomic data that seem to highlight the importance of pigment production and modification genes such as those in the melanin synthesis pathway (Stuckert et al. 2019; Posso-Terranova and Andrés 2017), pteridine synthesis pathway (Stuckert et al. 2019; Rodríguez et al. 2020) and carotenoid processing pathways (Twomey, Johnson, et al. 2020) for their roles in producing different color morphs in poison frogs.

We conducted a targeted analysis of genes which show convergent expression patterns between the model and mimic. Surprisingly, there was minimal overlap in genes that were differentially expressed between convergent color morphs in both the model and mimic. Of those genes that were shared, only five were in our *a priori* color gene list (*cd36, crabp2, edaradd, mvd,*

and *qdpr*). This seems to indicate: 1) the pattern of differential gene expression affecting convergent color morph development is largely species-specific, 2) the expression patterns of a small number of genes have a very large effect on color morph divergence, and/or 3) we have insufficient power to identify these convergent genes. Overall, these results suggest that the convergent color patterns of these species are likely to have evolved via the expression of distinct underlying genetic and biochemical pathways. In sum, color differences are likely to be driven by expression patterns of many genes involved in several pathways.

As for the genes that were differentially expressed throughout development, many are likely related to body restructuring rather than coloration *per se.* Nevertheless, we identified a number of very promising candidate color genes that are likely to play a role in the production of mimetic phenotypes in this system. In the rest of the manuscript we provide a high-level summary of key candidate genes which are differentially expressed. We also include a much more detailed description of these candidate genes and their potential mechanisms in the Supplemental Discussion.

Yellow, orange, and red coloration:

Yellows, oranges, and reds are determined in large part by the presence of pigments deposited within the xanthophores, the outermost layer of chromatophores in the skin (Duellman and Trueb 1986). These pigments are primarily composed of pteridines and carotenoids, and many studies to date have documented that these pigments play a key role in the production of yellows, oranges, and reds (Obika and Bagnara 1964; Grether et al. 2001; Mcgraw, Nolan, and Crino 2006; McLean et al. 2017; Croucher et al. 2013). In this study we found a number of key pteridine synthesis genes that were differentially expressed between color morphs (Figure 7). Prominent

amongst these are the aforementioned xanthine dehydrogenase (*xdh*) and quinoid dihydropteridine reductase (*qdpr*). Xanthine dehydrogenase appears to be highly conserved and its expression plays a role in the production of pterin-based coloration in a variety of taxa such as spiders (Croucher et al. 2013), fish (Parichy et al. 2000; Salis et al. 2019), salamanders (Frost and Bagnara 1979; Thorsteinsdottir and Frost 1986), and the dendrobatid frogs *D. auratus* and *O. pumilio* (Stuckert et al. 2019; Rodríguez et al. 2020). In our study, *xdh* had the highest expression in the orange banded morph of *R. imitator*. Quinoid dihydropteridine reductase (*qdpr*) is another gene involved in the pteridine synthesis pathway and is known to alter patterns of production of the yellow pigment sepiapterin (Ponzone et al. 2004). We found differential expression in this gene across developmental stages in all three species in this study, with the highest expression in the redheaded morph. Further, we identified a number of key genes that are differentially expressed and required for the production of xanthophores (Figure 7), notably paired box 7 (*pax7*) and xanthine dehydrogenase (*xdh*), both of which have been linked to xanthophore differentiation (Epperlein and Löfberg 1990; Reaume, Knecht, and Chovnick 1991; Parichy et al. 2000; Nord et al. 2016).

Melanophore genes:

Melanin-based coloration is the best understood aspect of coloration, in no small part because of a long history of genetic analyses in lab mice (Hoekstra 2006; Hubbard et al. 2010). As a result, there are a large number of genes that are known to influence the production of melanin, melanophores, and melanosomes. In vertebrates, black coloration is caused by light absorption by melanin in melanophores (Sköld et al. 2016). Melanophores (and the other chromatophores) originate from populations of cells in the neural crest early in development (Park et al. 2009). The

four color morphs of *Ranitomeya* used in this study have pattern elements on top of a generally black dorsum and legs, and therefore melanin-related genes are likely to play a key role in color and pattern, both throughout development and between color morphs.

Given that a large portion of pigmentation arises during development when we sampled individuals, we found that many of our differentially expressed candidate genes are in this pathway. Prominent amongst these genes are *dct, kit, mc1r, mitf, mlph, notch1, notch2, sox9, sox10, tyr*, and *tyrp1*, all of which were differentially expressed across development in at least one species. In fact, the well-known patterning genes in the notch pathway (e.g., *notch1, notch2*; Figure 8) seem to be important in all three species, as *notch1* was differentially expressed across development in *R. imitator* and *R. variabilis,* and *notch2* in *R. fantastica* (Hamada et al. 2014). Expression patterns of melanophore and melanin synthesis genes did not follow a consistent pattern overall and instead were variable.

Iridophore genes:

Iridophores are largely responsible for white, blue, and green coloration, which is mainly determined by the reflection of light from iridophores (Bagnara et al. 2007). Recently, Twomey et al. (2020) found that variation in coloration in *Ranitomeya* and related poison frogs is largely driven by a combination of the orientation and thickness of the guanine platelets in iridophores, and their interaction with pigment components.

We found a number of genes that have previously been linked to iridophore production or development to be differentially expressed between developmental stages (*arf6, dct, dgat2, dock7, dst, edn3, erbb3, impdh2, paics, rab27a, rab27b*) or color morphs (*arf6, dock7, dst, erbb3, gart, gne, paics, rab27a, rab27b, rab7a, rabggta*) in our study (Figure 9). We also found differential expression

of a number of genes that are known to impact guanine or purine synthesis throughout development *(adsl, gart, gas1, qdpr*) and between color morphs (*atic, qdpr*). A number of these genes *(adsl, dct, dock7, gart, qdpr, rabggta*) have been implicated in previous work in dendrobatids (Rodríguez et al. 2020; Stuckert et al. 2019) and in other taxa.

Notably, a number of epidermis-structuring genes (such as those in the krt family) have been implicated in the production of structural colors (e.g., *krt1, krt2*), although more evidence is needed to verify their role in coloration (Burgon et al. 2020; Stuckert et al. 2019; McGowan et al. 2006; Cui et al. 2016). We identified a number of these that are differentially expressed between color morphs (e.g., *krt1, krt2, krt8*). Genes that influence keratin, and organization of the epidermis generally, are good candidates for the production of different colors, as they may produce structural influences on color (via reflectance) in a manner that parallels what we see from guanine platelets. Of particular interest is *krt8.* We found differential expression of *krt8* between developmental stages in all three species, as well as between morphs in both *R. fantastica* and *R. variabilis.* Previous work has identified *krt8* as a candidate gene for coloration in several taxa (Stubbs 2017; Stuckert et al. 2019). In their in-depth analysis of genes controlling coloration, Linderoth et al. (Linderoth et al. 2023) identified *krt8.2* as a putative causative agent in leg coloration (*krt8.2* is the *Xenopus* ortholog of the human *krt8*, and thus naming differences likely primarily stem from databases used in annotation between studies). Thus, color and pattern may be in part driven by both variants and differential expression of *krt8*.

New candidate genes:

Many of the genes discussed above are those that have been previously implicated in other taxa. As a result, these are subject to ascertainment bias. One of the novel aspects of this work is

642 the comparative approach across species, and that it is occurring in non-model organisms. We
643 leveraged this to examine several genes which had high log-fold changes and were differentially
644 expressed between morphs in convergent species (i.e., between spotted and striped morphs of *R.*
645 *imitator* and *R. variabilis* or between banded and redheaded morphs of *R. imitator* and *R.*
646 *fantastica*), potentially indicating their importance. There are many more genes that fall into this
647 category than we can discuss. So here we highlight genes in which we were able to decipher
648 plausible roles in coloration based on their function in other organisms. Among these are *ep-*
649 *cadherin*, *vat1l*, *saa3, selenoi*, ov*ol1*, *piwil1, pdc*, *ddb_g0268948*, *cd36*, *perlwapin*.
650
651 *Conclusion:*
652 In this study we examined the molecular mechanisms by which mimetic phenotypes are
653 produced in a Müllerian mimicry system. Through our efforts, we have produced the first high-
654 quality poison frog genome, a 6.02 Gbp, contiguous genome assembly with good genic coverage.
655 We leveraged this to examine gene expression in the skin throughout development of four
656 comimetic morphs from three species of *Ranitomeya.* We identified a large number of genes
657 related to melanophores, melanin production, iridophore production, and guanine synthesis that
658 were differentially expressed throughout development, indicating that many of these are important
659 in the production of pigmentation, albeit not color morph specific coloration. Genes related to
660 xanthophore production, carotenoid pathways, melanin production and melanophore production
661 were rarely differentially expressed between color morphs, however those genes that were
662 differentially expressed may be critically important in producing polytypic differences within
663 species that drive mimetic phenotypes. Our results indicate that divergence between color morphs
664 seems to involve differences in expression and/or timing of expression, but that convergence for

the same color pattern may not be obtained through the same changes in gene expression between species. We identified the importance of the pteridine synthesis pathway in producing these different yellow, orange, and red color morphs across species. Thus, production of these colors are likely strongly driven by differences in gene expression in genes in the pteridine synthesis pathway, and our data indicate that there may be species-specific differences in this pathway used in producing similar colors and patterns. Further, we highlight the potential importance of genes in the keratin family for producing differential color via structural mechanisms.

Funding:

Funding for this project was provided by an East Carolina University Thomas Harriot College of Arts and Sciences Advancement Council Distinguished Professorship and NSF DEB 165536 to KS, NSF DEB 1655585 to MDM, XSEDE XRAC MCB110134 to MDM and AMMS andMC received support from Agence Nationale de la Recherche (ANR) grants RANAPOSA (ref. ANR-20-CE02-0003), from an "Investissement d'Avenir" grant CEBA (ref. ANR-10-LABX-25-01) and from a Marie Sklodowska-Curie fellowship (FITINV, N 655857).

Acknowledgements:

We are grateful to many individuals for their help with frog husbandry in the lab, including but not limited to M Yoshioka, C Meeks, A Sorokin, K Weinfurther, R Sen, N Davison, M Johnson, M Pahl, N Aramburu, M Tuatama, R Mori-Pezo, J Richard and S Gallusser. We are also grateful to Laura Bauza-Davila for her work doing RNA extractions, and Andrew Lang for guidance converting RNA to cDNA and preparing samples for sequencing. This work was completed in part with resources provided by

688 the Research Computing Data Core at the University of Houston. We are grateful to the anonymous
689 reviewers and editors whose comments improved this manuscript.

690

691 **Author contributions:**

692 Designed research: AMMS, MC, MM, RN, KS, MDM

693 Performed research: AMMS, MC, MM, KS, TML

694 Analyzed data: AMMS, TML

695 Contributed to writing: AMMS, MC, MM, TML, TL, RN, KS, MDM

696

References:

Ahi, Ehsan Pashay, Laurène A. Lecaudey, Angelika Ziegelbecker, Oliver Steiner, Ronald Glabonjat, Walter Goessler, Victoria Hois, Carina Wagner, Achim Lass, and Kristina M. Sefc. 2020. "Comparative Transcriptomics Reveals Candidate Carotenoid Color Genes in an East African Cichlid Fish." *BMC Genomics* 21 (1): 54.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2015. "HTSeq—a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31 (2): 166–69.

Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 19 (5): 455–77.

Bao, Weidong, Kenji K. Kojima, and Oleksiy Kohany. 2015. "Repbase Update, a Database of Repetitive Elements in Eukaryotic Genomes." *Mobile DNA* 6 (June): 11.

Beddard, Frank Evers. 1892. *Animal Coloration: An Account of the Principal Facts and Theories Relating to the Colours and Markings of Animals*. S. Sonnenschein & Company.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)*. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20.

Braasch, Ingo, Manfred Schartl, and Jean Nicolas Volff. 2007. "Evolution of Pigment Synthesis Pathways by Gene and Genome Duplication in Fish." *BMC Evolutionary Biology* 7: 1–18.

Briolat, Emmanuelle S., Emily R. Burdfield-Steel, Sarah C. Paul, Katja H. Rönkä, Brett M. Seymoure, Theodore Stankowich, and Adam M. M. Stuckert. 2019. "Diversity in Warning Coloration: Selective Paradox or the Norm?" *Biological Reviews of the Cambridge Philosophical Society* 94 (2): 388–414.

Brown, Jason L., Evan Twomey, Adolfo Amezquita, Moises B. DeSouza, Janalee Caldwell, Stefan Lötters, Rudolf Von May, et al. 2011. "A Taxonomic Revision of the Neotropical Poison Frog Genus *Ranitomeya* (Amphibia: Dendrobatidae)." *Zootaxa* 3083: 1–120.

Burgon, James D., David R. Vieites, Arne Jacobs, Stefan K. Weidt, Helen M. Gunter, Sebastian Steinfartz, Karl Burgess, Barbara K. Mable, and Kathryn R. Elmer. 2020. "Functional Colour Genes and Signals of Selection in Colour-polymorphic Salamanders." *Molecular Ecology* 29 (7): 1284–99.

Campbell, Michael S., Carson Holt, Barry Moore, and Mark Yandell. 2014. "Genome Annotation and Curation Using MAKER and MAKER-P." *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]* 48 (1): 188.

Coombe, Lauren, Jessica Zhang, Benjamin P. Vandervalk, Justin Chu, Shaun D. Jackman, Inanc Birol, and René L. Warren. 2018. "ARKS: Chromosome-Scale Scaffolding of Human Genome Drafts with Linked Read Kmers." *BMC Bioinformatics* 19 (1): 234.

Croucher, Peter J. P., Michael S. Brewer, Christopher J. Winchell, Geoff S. Oxford, and Rosemary G. Gillespie. 2013. "De Novo Characterization of the Gene-Rich Transcriptomes of Two Color-Polymorphic Spiders, Theridion Grallator and T. Californicum (Araneae: Theridiidae), with Special Reference to Pigment Genes." *BMC Genomics* 14 (1): 862.

Cui, Yucong, Yajun Song, Qingling Geng, Zengfeng Ding, Yilong Qin, Ruiwen Fan, Changsheng Dong, and Jianjun Geng. 2016. "The Expression of KRT2 and Its Effect on Melanogenesis in Alpaca

742 Skins." *Acta Histochemica* 118 (5): 505–12.
743 Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe
744 Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq
745 Aligner." *Bioinformatics* 29 (1): 15–21.
746 Duellman, William E., and Linda Trueb. 1986. *Biology of Amphibians*. Baltimore: The John Hopkins
747 University Press.
748 DuShane, Graham P. 1935. "An Experimental Study of the Origin of Pigment Cells in Amphibia." *The*
749 *Journal of Experimental Zoology* 72 (1): 1–31.
750 Eden, Eran, Roy Navon, Israel Steinfeld, Doron Lipson, and Zohar Yakhini. 2009. "GOrilla: A Tool for
751 Discovery and Visualization of Enriched GO Terms in Ranked Gene Lists." *BMC Bioinformatics*
752 10 (February): 48.
753 Emerling, Christopher A. 2018. "Independent Pseudogenization of CYP2J19 in Penguins, Owls and
754 Kiwis Implicates Gene in Red Carotenoid Synthesis." *Molecular Phylogenetics and Evolution*
755 118 (January): 47–53.
756 Epperlein, H. H., and J. Löfberg. 1990. "The Development of the Larval Pigment Patterns in Triturus
757 Alpestris and Ambystoma Mexicanum." *Advances in Anatomy, Embryology, and Cell Biology*
758 118: 1–99.
759 Fox, Denis L. 1936. "Structural and Chemical Aspects of Animal Coloration." *The American*
760 *Naturalist* 70 (730): 477–93.
761 Frost, Sally K. 1978. "Developmental Aspects of Pigmentation in the Mexican Leaf Frog,
762 *Pachymedusa Dacnicolor*."
763 Frost, S. K., and J. T. Bagnara. 1979. "Allopurinol-Induced Melanism In The Tiger Salamander
764 (Ambystoma Iigrinum Nebulosum)." *The Journal of Experimental Zoology* 209 (3): 455–65.
765 Funk, W. Chris, Kelly R. Zamudio, and Andrew J. Crawford. 2018. "Advancing Understanding of
766 Amphibian Evolution, Ecology, Behavior, and Conservation with Massively Parallel
767 Sequencing." *Population Genomics*. https://doi.org/10.1007/13836_2018_61.
768 Grabherr, Manfred G., Brian J. Haas, Moran Yassour, Joshua Z. Levin, Dawn A. Thompson, Ido Amit,
769 Xian Adiconis, et al. 2011. "Full-Length Transcriptome Assembly from RNA-Seq Data without a
770 Reference Genome." *Nature Biotechnology* 29 (7): 644–52.
771 Gray, Suzanne M., and Jeffrey S. McKinnon. 2007. "Linking Color Polymorphism Maintenance and
772 Speciation." *Trends in Ecology & Evolution* 22 (2): 71–79.
773 Grether, Gregory F., David F. Millie, Michael J. Bryant, David N. Reznick, and Wendy Mayea. 2001.
774 "RAIN FOREST CANOPY COVER, RESOURCE AVAILABILITY, AND LIFE HISTORY EVOLUTION IN
775 GUPPIES." *Ecology* 82 (6): 1546–59.
776 Gu, Li-Hong, and Pierre A. Coulombe. 2007a. "Keratin Function in Skin Epithelia: A Broadening
777 Palette with Surprising Shades." *Current Opinion in Cell Biology* 19 (1): 13–23.
778 ———. 2007b. "Keratin Expression Provides Novel Insight into the Morphogenesis and Function of
779 the Companion Layer in Hair Follicles." *The Journal of Investigative Dermatology* 127 (5):
780 1061–73.
781 Hamada, Hiroki, Masakatsu Watanabe, Hiu Eunice Lau, Tomoki Nishida, Toshiaki Hasegawa, David
782 M. Parichy, and Shigeru Kondo. 2014. "Involvement of Delta/Notch Signaling in Zebrafish Adult
783 Pigment Stripe Patterning." *Development* 141 (2): 318–24.
784 Hellsten, Uffe, Richard M. Harland, Michael J. Gilchrist, David Hendrix, Jerzy Jurka, Vladimir
785 Kapitonov, Ivan Ovcharenko et al. "The genome of the Western clawed frog Xenopus
786 tropicalis." *Science* 328, no. 5978 (2010): 633-636.
787 Hodges, Scott A., and Nathan J. Derieg. 2009. "Adaptive Radiations: From Field to Genomic

Studies." *Proceedings of the National Academy of Sciences of the United States of America* 106 Suppl 1 (June): 9947–54.
Hoekstra, H. E. 2006. "Genetics, Development and Evolution of Adaptive Pigmentation in Vertebrates." *Heredity* 97 (3): 222–34.
Hooper, Daniel M., Simon C. Griffith, and Trevor D. Price. 2019. "Sex Chromosome Inversions Enforce Reproductive Isolation across an Avian Hybrid Zone." *Molecular Ecology* 28 (6): 1246–62.
Hubbard, Joanna K., J. Albert C. Uy, Mark E. Hauber, Hopi E. Hoekstra, and Rebecca J. Safran. 2010. "Vertebrate Pigmentation: From Underlying Genes to Adaptive Function." *Trends in Genetics: TIG* 26 (5): 231–39.
Kang, Changku, Thomas N. Sherratt, Ye Eun Kim, Yujin Shin, Jongyeol Moon, Uhram Song, Jae Yeon Kang, Kyungmin Kim, and Yikweon Jang. 2017. "Differential Predation Drives the Geographical Divergence in Multiple Traits in Aposematic Frogs." *Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology* 00: 1–9.
Kim, Heuijong, Kiyoung Kim, and Jeongbin Yim. 2013. "Biosynthesis of Drosopterins, the Red Eye Pigments of Drosophila Melanogaster." *IUBMB Life* 65 (4): 334–40.
Kronforst, Marcus R., and Riccardo Papa. 2015. "The Functional Basis of Wing Patterning in *Heliconius* Butterflies: The Molecules behind Mimicry." *Genetics* 200 (1): 1–19.
Langfelder, Peter, and Steve Horvath. 2008. "WGCNA: An R Package for Weighted Correlation Network Analysis." *BMC Bioinformatics* 9 (December): 559.
Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform." *Bioinformatics* . https://academic.oup.com/bioinformatics/article-abstract/25/14/1754/225615.
Li, Heng. 2018. "Minimap2: Pairwise Alignment for Nucleotide Sequences." *Bioinformatics* 34 (18): 3094–3100.
Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79.
Longley, W. H. 1917. "Studies Upon the Biological Significance of Animal Coloration. II. A Revised Working Hypothesis of Mimicry." *The American Naturalist* 51 (605): 257–85.
Love, Michael, Simon Anders, and Wolfgang Huber. 2014. "Differential Analysis of Count Data--the DESeq2 Package." *Genome Biology* 15 (550): 10–1186.
MacManes, Matthew D. 2018. "The Oyster River Protocol: A Multi-Assembler and Kmer Approach for de Novo Transcriptome Assembly." *PeerJ* 6: e5428.
Mallet, James, and Nicholas H. Barton. 1989. "Strong Natural Selection in a Warning-Color Hybrid Zone." *Evolution; International Journal of Organic Evolution* 43 (2): 421–31.
Marek, Paul E., and Jason E. Bond. 2009. "A Müllerian Mimicry Ring in Appalachian Millipedes." *Proceedings of the National Academy of Sciences of the United States of America* 106 (24): 9755–60.
Marks, Patrick, Sarah Garcia, Alvaro Martinez Barrio, Kamila Belhocine, Jorge Bernate, Rajiv Bharadwaj, Keith Bjornson, et al. 2019. "Resolving the Full Spectrum of Human Genome Variation Using Linked-Reads." *Genome Research* 29 (4): 635–45.
Martin, A., R. Papa, N. J. Nadeau, R. I. Hill, B. A. Counterman, G. Halder, C. D. Jiggins, et al. 2012. "Diversification of Complex Butterfly Wing Patterns by Repeated Regulatory Evolution of a Wnt Ligand." *Proceedings of the National Academy of Sciences* 109 (31): 12632–37.
McGowan, Kelly A., Swaroop Aradhya, Helmut Fuchs, Martin H. de Angelis, and Gregory S. Barsh.

2006. “A Mouse Keratin 1 Mutation Causes Dark Skin and Epidermolytic Hyperkeratosis.” *The Journal of Investigative Dermatology* 126 (5): 1013–16.

Mcgraw, K. J., P. M. Nolan, and O. L. Crino. 2006. “Carotenoid Accumulation Strategies for Becoming a Colourful House Finch: Analyses of Plasma and Liver Pigments in Wild Moulting Birds.” *Functional Ecology* 20 (4): 678–88.

McLean, Claire A., Adrian Lutz, Katrina J. Rankin, Adam Elliott, Adnan Moussalli, and Devi Stuart-Fox. 2019. “Red Carotenoids and Associated Gene Expression Explain Colour Variation in Frillneck Lizards.” *Proceedings. Biological Sciences / The Royal Society* 286 (1907): 20191172.

McLean, Claire A., Adrian Lutz, Katrina J. Rankin, Devi Stuart-Fox, and Adnan Moussalli. 2017. “Revealing the Biochemical and Genetic Basis of Color Variation in a Polymorphic Lizard.” *Molecular Biology and Evolution* 34 (8): 1924–35.

Nord, Hanna, Nils Dennhag, Joscha Muck, and Jonas von Hofsten. 2016. “Pax7 Is Required for Establishment of the Xanthophore Lineage in Zebrafish Embryos.” *Molecular Biology of the Cell* 27 (11): 1853–62.

Obika, M., and J. T. Bagnara. 1964. “PTERIDINES AS PIGMENTS IN AMPHIBIANS.” *Science* 143 (3605): 485–87.

Parichy, D. M., D. G. Ransom, B. Paw, L. I. Zon, and S. L. Johnson. 2000. “An Orthologue of the Kit-Related Gene Fms Is Required for Development of Neural Crest-Derived Xanthophores and a Subpopulation of Adult Melanocytes in the Zebrafish, Danio Rerio.” *Development* 127 (14): 3031–44.

Park, H. Y., M. Kosmadaki, M. Yaar, and B. A. Gilchrest. 2009. “Cellular Mechanisms Regulating Human Melanogenesis.” *Cellular and Molecular Life Sciences: CMLS* 66 (9): 1493–1506.

Ponzone, Alberto, Marco Spada, Silvio Ferraris, Irma Dianzani, and Luisa De Sanctis. 2004. “Dihydropteridine Reductase Deficiency in Man: From Biology to Treatment.” *Medicinal Research Reviews* 24 (2): 127–50.

Posso-Terranova, Andrés, and José Á. Andrés. 2017. “Diversification and Convergence of Aposematic Phenotypes: Truncated Receptors and Cellular Arrangements Mediate Rapid Evolution of Coloration in Harlequin Poison Frogs.” *Evolution; International Journal of Organic Evolution* 71 (11): 2677–92.

Reaume, A. G., D. A. Knecht, and A. Chovnick. 1991. “The Rosy Locus in Drosophila Melanogaster: Xanthine Dehydrogenase and Eye Pigments.” *Genetics* 129 (4): 1099–1109.

Reed, Robert D., Riccardo Papa, Arnaud Martin, Heather M. Hines, Marcus R. Kronforst, Rui Chen, Georg Halder, H. Frederik Nijhout, and W. Owen Mcmillan. 2011. “Optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry.” *Science* 333 (August): 1137–41.

Robertson, Gordon, Jacqueline Schein, Readman Chiu, Richard Corbett, Matthew Field, Shaun D. Jackman, Karen Mungall, et al. 2010. “De Novo Assembly and Analysis of RNA-Seq Data.” *Nature Methods* 7 (11): 909–12.

Rodríguez, Ariel, Nicholas I. Mundy, Roberto Ibáñez, and Heike Pröhl. 2020. “Being Red, Blue and Green: The Genetic Basis of Coloration Differences in the Strawberry Poison Frog (Oophaga Pumilio).” *BMC Genomics* 21 (1): 301.

Rogers, Rebekah L., Long Zhou, Chong Chu, Roberto Márquez, Ammon Corl, Tyler Linderoth, Layla Freeborn, et al. 2018. “Genomic Takeover by Transposable Elements in the Strawberry Poison Frog.” *Molecular Biology and Evolution* 35 (12): 2913–27.

Ruan, Jue, and Heng Li. 2019. “Fast and Accurate Long-Read Assembly with wtdbg2.” *Nature Methods*, December. https://doi.org/10.1038/s41592-019-0669-3.

Rubio, Andrew O., Adam MM Stuckert, Troy M LaPolice, T. Jeffrey Cole, and Kyle Summers. *In press.*

880 Under pressure: Evidence for selection on color-related genes in poison frogs of the genus
881 *Ranitomeya. Evolutionary Ecology.*
882 Rudh, Andreas, and Anna Qvarnström. 2013. “Adaptive Colouration in Amphibians.” *Seminars in*
883 *Cell & Developmental Biology* 24 (6-7): 553–61.
884 Ruxton, G. D., T. N. Sherratt, and M. P. Speed. 2004. *Avoiding Attack: The Evolutionary Ecology of*
885 *Crypsis, Warning Signals and Mimicry*. Vol. 17.
886 Salis, Pauline, Thibault Lorin, Victor Lewis, Carine Rey, Anna Marcionetti, Marie-Line Escande,
887 Natacha Roux, et al. 2019. “Developmental and Comparative Transcriptomic Identification of
888 Iridophore Contribution to White barring in Clownfish.” *Pigment Cell & Melanoma Research* 32
889 (3): 391–402.
890 Schulte, Rainer. 1986. “Eine Neue Dendrobates—art Aus Ostperu (Amphibia: Salienta:
891 Dendrobatidae).” *Sauria* 8: 11–20.
892 Sherratt, Thomas N. 2006. “Spatial Mosaic Formation through Frequency-Dependent Selection in
893 M??llerian Mimicry Complexes.” *Journal of Theoretical Biology* 240 (2): 165–74.
894 ———. 2008. “The Evolution of Müllerian Mimicry.” *Die Naturwissenschaften* 95 (8): 681–95.
895 Simão, Felipe A., Robert M. Waterhouse, Panagiotis Ioannidis, Evgenia V. Kriventseva, and Evgeny
896 M. Zdobnov. 2015. “BUSCO: Assessing Genome Assembly and Annotation Completeness with
897 Single-Copy Orthologs.” *Bioinformatics* 31 (19): 3210–12.
898 Sköld, Helen Nilsson, Sara Aspengren, Karen L. Cheney, and Margareta Wallin. 2016. “Fish
899 Chromatophores-From Molecular Motors to Animal Behavior.” *International Review of Cell*
900 *and Molecular Biology* 321: 171–219.
901 Smit, A. F. A., and R. Hubley. 2008-2015. *RepeatModeler Open-1.0*. http://www.repeatmasker.org.
902 Smit, A. F. A., R. Hubley, and P. Green. 2013-2015. *RepeatMasker Open-4.0*.
903 http://www.repeatmasker.org.
904 Stuckert, Adam M. M., Emily Moore, Kaitlin P. Coyle, Ian Davison, Matthew D. MacManes, Reade
905 Roberts, and Kyle Summers. 2019. “Variation in Pigmentation Gene Expression Is Associated
906 with Distinct Aposematic Color Morphs in the Poison Frog, Dendrobates Auratus.” *BMC*
907 *Evolutionary Biology* 19 (85): 1–15.
908 Stuckert, Adam M. M., Ralph A. Saporito, Pablo J. Venegas, and Kyle Summers. 2014. “Alkaloid
909 Defenses of Co-Mimics in a Putative Müllerian Mimetic Radiation.” *BMC Evolutionary Biology*
910 14: 1–8.
911 Stuckert, Adam M. M., Pablo J. Venegas, and Kyle Summers. 2014. “Experimental Evidence for
912 Predator Learning and Müllerian Mimicry in Peruvian Poison Frogs (Ranitomeya,
913 Dendrobatidae).” *Evolutionary Ecology* 28 (3): 413–26.
914 Stuckert, Adam M. M. and Troy M. LaPolice. (2021). “AdamStuckert/Ranitomeya_imitator_genome:
915 The genomics of mimicry: gene expression throughout development provides insights into
916 convergent and divergent phenotypes in a Müllerian mimicry system (Version v1.0).” *Zenodo*.
917 http://doi.org/10.5281/zenodo.4758404.
918 Stuckert, Adam M. M. Mathieu Chouteau, Melanie McClure, Troy M LaPolice, Tyler Linderoth,
919 Rasmus Nielsen, … Matthew D MacManes. (2021). “The genomics of mimicry: gene expression
920 throughout development provides insights into convergent and divergent phenotypes in a
921 Müllerian mimicry system.” *Molecular Ecology* 30 (16):4039—4061.
922 Stuckert, Adam M. M. Mathieu Chouteau, Melanie McClure, Troy M LaPolice, Tyler Linderoth,
923 Rasmus Nielsen, … Matthew D MacManes. (2021). “The genomics of mimicry: gene expression
924 throughout development provides insights into convergent and divergent phenotypes in a
925 Müllerian mimicry system.” *Zenodo*. http://doi.org/10.5281/zenodo.4758346.

926 Sun, Yan-Bo, Zi-Jun Xiong, Xue-Yan Xiang, Shi-Ping Liu, Wei-Wei Zhou, Xiao-Long Tu, Li Zhong et al.
927 "Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative
928 evolution of tetrapod genomes." Proceedings of the National Academy of Sciences 112, no. 11
929 (2015): E1257-E1262.
930 Supple, Megan a., Heather M. Hines, Kanchon K. Dasmahapatra, James J. Lewis, Dahlia M. Nielsen,
931 Christine Lavoie, David a. Ray, Camilo Salazar, W. Owen Mcmillan, and Brian a. Counterman.
932 2013. "Genomic Architecture of Adaptive Color Pattern Divergence and Convergence in
933 *Heliconius* Butterflies." *Genome Research* 23 (662): 1248–57.
934 Symula, R., Rainer Schulte, and Kyle Summers. 2001. "Molecular Phylogenetic Evidence for a
935 Mimetic Radiation in Peruvian Poison Frogs Supports a Müllerian Mimicry Hypothesis."
936 *Proceedings of the Royal Society B: Biological Sciences* 268 (1484): 2415–21.
937 ———. 2003. "Molecular Systematics and Phylogeography of Amazonian Poison Frogs of the Genus
938 *Dendrobates*." *Molecular Phylogenetics and Evolution* 26 (3): 452–75.
939 Team, R. Core. 2019. "R: A Language and Environment for Statistical computing(Version 3.6.
940 0)[Computer Software]. R Foundation for Statistical Computing, Vienna, Austria."
941 Thorsteinsdottir, Solveig, and Sally K. Frost. 1986. "Pigment Cell Differentiation: The Relationship
942 between Pterin Content, Allopurinol Treatment, and the Melanoid Gene in Axolotls." *Cell*
943 *Differentiation* 19: 161–72.
944 Tsuji Gaku, Takamachi Ito, Takahito Chiba, Chikage Mitoma, Takeshi Nakahara, Hiroshi Uchi, and
945 Masutaka Furue. 2018. The role of the OVOL1-OVOL2 axis in normal and diseased human skin.
946 *Journal of Dermatological Science* 90:227–231.
947 Twomey, Evan, James D. Johnson, Santiago Castroviejo-Fisher, and Ines Van Bocxlaer. 2020. "A
948 Ketocarotenoid-Based Color Polymorphism in the Sira Poison Frog Ranitomeya Sirensis
949 Indicates Novel Gene Interactions Underlying Aposematic Signal Variation." *Molecular Ecology*.
950 https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.15466?casa_token=HfW3FcMDWEsAAA
951 AA:4WvPeH_f3rgYyvO-PmCh18EPVKnr3kW3KSxksgWbdPYhWPyJk20QEbdD3osOTofr6V-
952 F43rVbVH9fA.
953 Twomey, Evan, Morgan Kain, Myriam Claeys, Kyle Summers, Santiago Castroviejo-Fisher, and Ines
954 Van Bocxlaer. 2020. "Mechanisms for Color Convergence in a Mimetic Radiation of Poison
955 Frogs." *The American Naturalist*, January. https://doi.org/10.1086/708157.
956 Twomey, Evan, Justin Yeager, Jason Lee Brown, Victor Morales, Molly Cummings, and Kyle
957 Summers. 2013. "Phenotypic and Genetic Divergence among Poison Frog Populations in a
958 Mimetic Radiation." Edited by Pawel Michalak. *PloS One* 8 (2): e55443.
959 Walker, Bruce J., Thomas Abeel, Terrance Shea, Margaret Priest, Amr Abouelliel, Sharadha
960 Sakthikumar, Christina A. Cuomo, et al. 2014. "Pilon: An Integrated Tool for Comprehensive
961 Microbial Variant Detection and Genome Assembly Improvement." *PloS One* 9 (11): e112963.
962 Warren, Rene L. 2016. "RAILS and Cobbler: Scaffolding and Automated Finishing of Draft Genomes
963 Using Long DNA Sequences." *J. Open Source Software* 1 (7): 116.
964 Weisenfeld, Neil I., Vijay Kumar, Preyas Shah, Deanna M. Church, and David B. Jaffe. 2017. "Direct
965 Determination of Diploid Genome Sequences." *Genome Research* 27 (5): 757–67.
966 Wickham, Hadley. 2011. "ggplot2." *Wiley Interdisciplinary Reviews: Computational Statistics* 3 (2):
967 180–85.
968 Wilson, Joseph S., Joshua P. Jahner, Matthew L. Forister, Erica S. Sheehan, Kevin A. Williams, and
969 James P. Pitts. 2015. "North American Velvet Ants Form One of the World's Largest Known
970 Müllerian Mimicry Complexes." *Current Biology: CB* 25 (16): R704–6.
971 Xu, Gui-Cai, Tian-Jun Xu, Rui Zhu, Yan Zhang, Shang-Qi Li, Hong-Wei Wang, and Jiong-Tang Li. 2019.

"LR_Gapcloser: A Tiling Path-Based Gap Closer That Uses Long Reads to Complete Genome Assembly." *GigaScience* 8 (1). https://doi.org/10.1093/gigascience/giy157.

Ziegler, Irmgard. 2003. "The Pteridine Pathway in Zebrafish: Regulation and Specification during the Determination of Neural Crest Cell-Fate." *Pigment Cell Research / Sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society* 16 (3): 172–82.

Supplemental Materials

SUPPLEMENTAL METHODS:

Genome assembly:

Annotations using Maker:

We annotated our genome using Maker version 3.01.02 (Campbell et al. 2014). We used transcript evidence from *Ranitomeya imitator* to aid in assembly ("est2genome=1"). These data include: 1) a developmental series of tadpole skin across color morphs of captive bred *R. imitator* (this paper, see below), 2) liver, skin, and intestine samples from six different wild *R. imitator* populations (Stuckert et al. unpublished data), and 3) brain samples from captive bred *R. imitator* (Geralds et al. unpublished data). The developmental series data were used in order to accurately annotate genes that are expressed in the skin at different time points in order to target color genes. The addition of data from wild frogs and a variety of other tissue types were used to provide additional transcript evidence in an effort to recover more genes after annotation.

Transcriptome assemblies:

Developmental series:

We used data from the imitator developmental series we analyzed in this paper to make transcriptome across developmental time points in the skin. In order to generate an initial reference transcriptome we assembled 40 M randomly subsampled forward and reverse reads sampled across morphs and time points using seqtk (https://github.com/lh3/seqtk) and used the Oyster River Protocol version 1.1.1 (MacManes 2017) to assemble this dataset. Evidence indicates that there is a substantial diminishment of returns in terms of transcriptome assembly completeness from using over 20-30 million reads (MacManes 2017). Initial error correction was done using RCorrector 1.01, followed by adaptor removal and quality trimming using trimmomatic version 0.36 at a Phred score of ≤ 3 (Bolger et al. 2014) since overly aggressive quality trimming has been shown to reduce assembly completeness (MacManes 2014). The Oyster River Protocol (MacManes 2017) assembles a transcriptome by merging multiple assemblies constructed using a series of different transcriptome assemblers and kmer lengths. We constructed the Independent assemblies with Trinity version 2.4.0 (Grabherr et al. 2011), Shannon version 0.0.2 (Kannan et al. 2016), and SPAdes assembler version 3.11 using 35-mers (Bankevich et al. 2012). This deviates slightly from the Oyster River Protocol specified in MacManes (2017), which specifies kmer lengths of 55 and 75 for SPAdes assemblies, but that exceeds our 50 bp sequencing read length. We then merged these individual assemblies using OrthoFuser (MacManes 2017). Finally, we assessed transcriptome quality using BUSCO version 3.0.1 (Simão et al. 2015) and TransRate 1.0.3 (Smith-Unna et al. 2016).

Transcriptomes from wild frogs of multiple populations:

We *de novo* assembled transcriptomes from six populations along a transition zone from the orange banded morph of *R. imitator* (*R. summersi* mimic) through the yellow striped morph (lowland *R. variabilis* mimic). This includes two pure orange banded populations, two pure yellow striped populations, and two admixed populations with highly variable phenotypes. For each population we randomly chose one individual, concatenated Illumina HiSeq 4000 reads from the liver, intestines, and dorsal skin into a single readset per population and then used the Oyster River Protocol version 2.2.8 (MacManes 2018) to assemble population specific *de novo* transcriptomes. This is similar to the approach above in the section "*Developmental series*", with some minor differences which we detail here. First, we assembled individual assemblies using Trinity version

2.8.5 (Grabherr et al. 2011), two iterations of SPAdes version 3.13.1 (Bankevich et al. 2012) with kmer values of 55 and 75 respectively, and finally Trans-ABySS version 2.0.1 (Robertson et al. 2010). These individually built transcriptomes were then merged together using OrthoFuser (MacManes 2018). Unique contigs which were dropped in Orthofuser were recovered using a reciprocal blast search of the final assembly against the individual assemblies for unique contigs. We removed all contigs with expression lower than one Transcript Per Million (TPM) using the TPM=1 flag in the ORP. Contigs that were dropped due to low expression but which likely represent expressed genes were recovered by blasting these against the UniProt database. Finally, transcriptome quality was assessed using BUSCO version 3.0.1 (Simão et al. 2015) and TransRate 1.0.3 (Smith-Unna et al. 2016).

Brain transcriptome assemblies:

Geralds et al. (unpublished data) conducted an experiment looking at the effects of parental behaviors and tadpole begging on gene expression in the brains of adult *Ranitomeya imitator*. We randomly chose one individual from individuals that were interacting with begging tadpoles, and those that were not. We concatenated these data into a single forward and reverse read file, then used the Oyster River Protocol version 2.2.8 (MacManes 2018) to assemble the brain transcriptome. These details are the same as above in the "*Transcriptomes from wild frogs of multiple populations"* section.

Gene expression:

Ranitomeya imitator:

Frogs from our study populations were collected by Understory and captive bred in Peru prior to shipping captive-bred frogs to a breeding facility in Canada. We purchased individuals that were captive bred in Canada. The frogs used in this study have a similar phenotype to those of individuals found in captivity from their source populations. As with any captive stock sourced from the wild, there is likely an initial bottleneck and corollary reduction in overall heterozygosity, however morphs were not selectively bred to produce desired phenotypes.

Breeding *R. imitator* pairs were placed in 5-gallon terraria containing small (approximately 13 cm) PVC pipes capped on one end and filled halfway with water. We removed tadpoles from the tanks after the male transported them into the pools of water and hand reared them. Although in the wild female *R. imitator* feed their tadpoles unfertilized eggs, tadpoles can survive and thrive on other food items (Brown et al. 2008). We raised experimental tadpoles on a diet of Omega One Marine Flakes fish food mixed with Freeze Dried Argent Cyclop-Eeze, which they received three times a week, with full water changes twice a week until sacrificed for analyses at 2, 4, 7, and 8 weeks of age. At two weeks, tadpoles are limbless, patternless, and a light gray color with two dark black eyeballs. At 4 weeks tadpoles are a slightly darker gray and have back limb buds. Tadpoles had developed their pattern and some coloration as well as reached the onset of metamorphosis at around week 7, and had metamorphosed, were resorbing the tail, and had their froglet patterns at 8 weeks old. Pattern development continues as juveniles and subadults frogs as they grow into the ultimate pattern they possess as adults. Our four sampling periods correspond to roughly Gosner stages 25, 27, 42, and 44 (Gosner 1960). Tadpoles were raised in a homogenous environment, and thus tadpoles were generally at the same developmental stage at the point of sacrifice. Due to inherent variation, some individuals may have been one Gosner stage off the norm. We sequenced a minimum of three individuals at each time point from the Sauce, Tarapoto, and Varadero

populations (except for Tarapoto at 8 weeks), and two individuals per time point from the Huallaga population. Individuals within the same time points were sampled from different family groups (Table 1).Individually barcoded samples were pooled and sequenced using 50 bp paired end reads on three lanes of the Illumina HiSeq 2500 at the New York Genome Center. This yielded on average 24.45M reads per library ± 8.6M sd (range: 10.1-64.M).

Ranitomeya fantastica and *Ranitomeya variabilis*:

We set up a captive colony consisting of between 6 and 10 wild collected individuals per locality, which was maintained at the Tarapoto Research Center (INIBICO. jr. Ventanilla s/n Sector Ventanilla. Banda de Shilcayo. San Martin, Peru). Male-female pairs were placed into individual terrariums which were misted with rainwater and fed with fruit flies daily. Artificial egg deposition sites consisting of short sections of PVC pipe (~10 cm in length) were positioned within each terrarium and we checked terraria for eggs biweekly. When eggs were found, they were transferred into petri dishes to monitor their development. Upon hatching, tadpoles were placed individually into 25 ml plastic cups filled with rain water and fed daily with a pinch of a 50/50 mix of spirulina and nettle leaf powder. Three tadpoles per stage (7, 14, 35, 49 and 56 days after hatching) were fixed in an RNAlater (Ambion) solution. To do so, tadpoles were first euthanized in a 250 mg/L benzocaine hydrochloride bath, then rinsed with distilled water before the whole tadpole was placed in RNAlater in a 2.5ml Eppendorf tube. The Eppendorf tube with sample in RNAlater was then stored at 4°C for 6h before being frozen at -20°C for long-term storage. This protocol was approved by the Peruvian Servicio Forestal y de Fauna Silvestre through the authorization number 232-2016- SERFOR/DGGSPFFS.

Species	Morph	Age (weeks)	n
Ranitomeya imitator	Redheaded	2	2
		4	2
		7	2
		8	2
	White-Banded	2	3
		4	3
		7	3
		8	3
	Spotted	2	4
		4	4
		7	4
		8	1
	Striped	2	3
		4	3
		7	3
		8	3
Ranitomeya fantastica	Redheaded	1	3
		2	3

		4	3
		5	3
		7	3
		8	3
	White-Banded	4	3
		5	3
		7	3
		8	3
Ranitomeya variabilis	Spotted	1	3
		2	3
		4	3
		5	3
		7	3
		8	3
		9	3
	Striped	1	3
		2	3
		4	3
		5	3
		7	3
		8	3

Supplemental Table 1. Sample sizes for gene expression of model species *R. fantastica* and *R. variabilis.*

SUPPLEMENTAL DISCUSSION:

Yellow, orange, and red coloration:

Yellows, oranges, and reds are determined in large part by the presence of pigments deposited within the xanthophores, the outermost layer of chromatophores in the skin (Duellman and Trueb 1986). These pigments are primarily composed of pteridines and carotenoids, and many studies to date have documented that these pigments play a key role in the production of yellows, oranges, and reds (Obika and Bagnara 1964; Grether et al. 2001; Mcgraw, Nolan, and Crino 2006; McLean et al. 2017; Croucher et al. 2013). Given the clear importance of xanthophores, pteridines, and carotenoids in the production of yellows, oranges, and reds, we briefly examine the contribution of genes in all three pathways here, beginning with xanthophore production.

Xanthophores:

We identified a number of key genes that are differentially expressed and required for the production of xanthophores (Figure 7), notably paired box 7 (*pax7*) and xanthine dehydrogenase (*xdh*). *Pax7* is a transcription factor that is required for establishing and differentiating xanthophores in both embryonic and adult zebrafish (Nord et al. 2016). *Pax7* was differentially expressed between morphs in *R. fantastica,* notably with higher expression in the orange-banded morph. *Xdh* is sometimes referred to as a xanthophore differentiation marker, as it is found in xanthoblasts and is required for synthesizing the pteridine pigment xanthopterin (Epperlein and Löfberg 1990; Reaume, Knecht, and Chovnick 1991; Parichy et al. 2000). In our study, this gene exhibited differential expression across development in *R. fantastica.* Additionally, we saw differential expression in this gene between *R. imitator* color morphs, with the highest expression in the orange banded morph (at alpha < 0.05). Given their roles in xanthophore differentiation, *pax7* and *xdh* are excellent candidates for production of xanthophores across all *Ranitomeya,* and

early differences in expression of these genes could lay the groundwork for markedly different colors and patterns.

Pteridine synthesis genes:

Genes and biochemical products in the pteridine pathway are important for pigmentation in the eyes and for vision, as well as for coloration across a wide variety of taxa (e.g., invertebrates, fish, lizards, amphibians). Data from both genetic studies and biochemical assays of pigments point to pteridines as important components of animal coloration. Although a number of genes in the pteridine pathway have been implicated in producing different color patterns, in general the genetic control of pteridine pigmentation is poorly characterized and largely comes from studies of *Drosophila melanogaster (Kim, Kim, and Yim 2013; Braasch, Schartl, and Volff 2007)*. In this study we found a number of key pteridine synthesis genes that were differentially expressed between color morphs (Figure 7). Prominent amongst these are the aforementioned xanthine dehydrogenase (*xdh*) and quinoid dihydropteridine reductase (*qdpr*).

In addition to its role in early xanthophore lineages, *xdh* appears to be highly conserved and its expression plays a role in the production of pterin-based coloration in a variety of taxa such as spiders (Croucher et al. 2013), fish (Parichy et al. 2000; Salis et al. 2019), and the dendrobatid frogs *D. auratus* and *O. pumilio (Stuckert et al. 2019; Rodríguez et al. 2020)*. Experimental inhibition of *xdh* causes a reduction in the quantity of pterins, resulting in an atypical black appearance in axolotls (Sally K. Frost 1978; S. K. Frost and Bagnara 1979; Thorsteinsdottir and Frost 1986). Additionally, typically green frogs with deficiencies in the *xdh* gene appear blue due to the lack of pterins in the xanthophores (Sally K. Frost 1978; S. K. Frost and Bagnara 1979). As discussed above, *xdh* had the highest expression in the orange banded morph of *R. imitator*. Because *xdh* plays a role

in the transformation of pterins into several different yellow pterin pigments such as xanthopterin and isoxanthopterin (Ziegler 2003), differential expression of *xdh* is a plausible mechanism for the production of orange coloration in the banded *R. imitator*. In sum, *xdh* may function in xanthophore production and/or pterin synthesis and is a likely driver of the production of yellow, orange, and green colors in this mimicry system.

Quinoid dihydropteridine reductase (*qdpr*) is another gene involved in the pteridine synthesis pathway and is known to alter patterns of production of the yellow pigment sepiapterin (Ponzone et al. 2004). We found differential expression in this gene across developmental stages in all three species in this study. Notably, the expression of *qdpr* showed a stark decline over development. *Qdpr* showed a convergent pattern of differential expression between color morphs in both *R. imitator* and *R. fantastica,* in both cases with the highest expression in the redheaded morph, although expression was also high in the yellow striped morph of *R. imitator*. This indicates that qdpr is likely playing a similar role in color production in both the model and mimic species. The *qdpr* gene was also differentially expressed across populations of another species of poison frog, and was only expressed in light blue or green colored morphs which are likely to derive some of the coloration from pteridines (Stuckert et al. 2019). In combination, these studies indicate that *qdpr* may be playing a role in the production of pteridine pigmentation in this system and in amphibians in general, particularly since *qdpr* alters sepiapterin production (Ponzone et al. 2004).

Carotenoid genes:

Carotenoids are important for both yellow and red coloration across a diversity of life forms (McGraw 2006; Toews et al 2017). While carotenoids are clearly an important class of pigments that broadly influence coloration, few known genes contribute to carotenoid based color

differences (e.g., *bco2*, *scarb1*, *retsat*), although this is likely an underestimation of the actual number of genes playing a role in carotenoid synthesis and processing given that we are continuously discovering new genes that play these roles (Twomey, Johnson, et al. 2020; Emerling 2018). It appears that orange and red colored *Ranitomeya* (outside of *R. imitator)* have slightly higher overall carotenoid concentrations, although these colors are more likely to be derived from the pteridine drosopterin (Twomey et al. 2020). However, different color morphs of *R. imitator* possess similar quantities of carotenoids (Twomey et al. 2020), providing little evidence that carotenoid levels influence coloration in *R. imitator*, but they may be important in other *Ranitomeya spp*. Intriguingly, Crothers et al (2016) found no relationship between coloration and overall carotenoid abundance in the bright orange Solarte morph of *O. pumilio*, but did find a relationship between total dorsal reflectance and two carotenoids (xanthophyll and a canary xanthophyll ester). Thus, the evidence seems to indicate that overall carotenoid concentrations are not particularly important for red, orange, or yellow colorations, and that instead these colors may be driven by a small subset of carotenoids or pteridines.

We found differential expression of a number of carotenoid genes across development (e.g., *aldh1a1*, *aldh1a2*, *bco1*, *retsat*, *scarb1, scarb2*). Although we did not find evidence that many known carotenoid genes are a major contributor to differences between specific color morphs in our study, we did find differential expression of some carotenoid genes between morphs (e.g., *scarb1*). *Scarb1* was differentially expressed during developmental time points in *R. imitator* and *R. fantastica*, as well as between morphs in *R. variabilis*, with higher expression in the yellow striped morph than the green spotted morph. Scarb1 has been identified as responsible for carotenoid uptake in canaries, and mutations in this gene can produce white canaries that lack carotenoids (Toomey et al. 2017). It is quite plausible that it is playing a similar role in *R. variabilis*.

While we found very few putative carotenoid genes identified in other taxa that were differentially expressed between our color morphs, we nevertheless identified several additional putative carotenoid genes. These include retinol dehydrogenase and a variety of cytochrome p450 genes. A recent study by Twomey et al. (2020) identified *CYP3A80* as a novel candidate for carotenoid ketolase that is preferentially expressed in the livers of red *Ranitomeya sirensis.* We found several cytochrome p2 family genes that were differentially expressed between morphs in both model and mimic and expressed high log fold changes (e.g., *cyp2f5, cyp2j5, cyp2c13*), and these would be plausible candidates of importance here.

Melanophore genes:

Many of the genes that were differentially expressed in our dataset were differentially expressed between developmental stages as tadpoles undergo a complete body reorganization as they prepare to metamorphose. In addition to growth, development, and metamorphosis, during this time tadpoles are producing both the structural chromatophores and the pigments that will be deposited within them. In many species, the distribution of chromatophores that will regulate patterns are set early during development. Of these, melanin-based coloration is the best understood aspect of coloration, in no small part because of a long history of genetic analyses in lab mice (Hoekstra 2006; Hubbard et al. 2010). As a result, there are a large number of genes that are known to influence the production of melanin, melanophores, and melanosomes. In vertebrates, black coloration is caused by light absorption by melanin in melanophores (Sköld et al. 2016). Melanophores (and the other chromatophores) originate from populations of cells in the neural crest early in development (Park et al. 2009). The four color morphs of *Ranitomeya* used in this study have pattern elements on top of a generally black dorsum and legs, and therefore melanin-

related genes are likely to play a key role in color and pattern, both throughout development and between color morphs.

Given that a large portion of pigmentation arises during development when we sampled individuals, we found that many of our differentially expressed candidate genes are in this pathway. Prominent amongst these genes are *dct, kit, mc1r, mitf, mlph, notch1, notch2, sox9, sox10, tyr*, and *tyrp1*, all of which were differentially expressed across development in at least one species. It seems likely these genes are contributing to color production across species given the important roles of each of these genes in melanocyte development and melanin synthesis. In fact, the well-known patterning genes in the notch pathway (e.g., *notch1, notch2*; Figure 8) seem to be important in all three species, as *notch1* was differentially expressed across development in *R. imitator* and *R. variabilis,* and *notch2* in *R. fantastica* (Hamada et al. 2014). Expression patterns of melanophore and melanin synthesis genes did not follow a consistent pattern overall and instead were variable.

We found *edaradd* (Ectodysplasin-A Receptor-Associated Death Domain), was differentially expressed between the spotted and striped morphs of *R. imitator* and *R. variabilis*. However, it showed different expression patterns between the two species. This gene is known to play a role in ectodermal dysplasia, a catch-all term for a suite of similar human diseases that produce abnormalities of ectodermal structures. Phenotypically these appear as sparse hair, abnormal teeth and hair, and, most relevant to our study, abnormally light pigmentation (Cluzeau et al. 2011). Another gene in this pathway, *edar* (Ectodysplasin A Receptor) was differentially expressed between color morphs of both *R. fantastica* and *R. variabilis*, potentially indicating the importance of the TNFα-related signaling pathway (Reyes-Reali et al. 2018; Cluzeau et al. 2011). The TNFα-related signaling pathway has not been implicated as a driver of color and pattern in any

context outside of ectodermal dysplasia, and thus this indicates a plausible mechanism for coloration in animals that has not yet been identified.

Iridophore genes:

Iridophores are largely responsible for blue (and to a lesser extent green) coloration, which is mainly determined by the reflection of light from iridophores (Bagnara et al. 2007). This depends on the presence and orientation of guanine platelets, where thicker platelets tend to reflect longer wavelengths of light (Ziegler 2003; Bagnara et al. 2007; Saenko et al. 2013). Iridophores are a key component of white, blue, and green coloration. Recently, Twomey et al. (2020) found that variation in coloration in *Ranitomeya* and related poison frogs is largely driven by a combination of the orientation and thickness of the guanine platelets in iridophores, and their interaction with pigment components. Using a combination of electron microscopy, biochemical pigment analyses, and modeling of the interaction between structural elements in the integument and pigmentation within chromatophores, they found that much more of the variation in coloration (hue) was driven by differences in the guanine platelet thickness of iridophores than expected. We discuss our findings in light of this recent work, with respect to how they can inform future work in conjunction with the results of Twomey et al. (2020).

The precise mechanisms underlying the development of iridophores and the size and orientation of the guanine platelets are unknown. However, previous work has suggested that ADP Ribosylation Factors (ARFs), which are ras-related GTPases that control transport through membranes and organelle structure and Rab GTPases, are likely important in determining the size and orientation of iridophores (Higdon et al. 2013). We found a number of these genes to be differentially expressed between developmental stages (*arf6, dct, dgat2, dock7, dst, edn3, erbb3,*

impdh2, paics, rab27a, rab27b) or color morphs (*arf6, dock7, dst, erbb3, gart, gne, paics, rab27a, rab27b, rab7a, rabggta*) in our study (Figure 9). We also found differential expression of a number of genes that are known to impact guanine or purine synthesis throughout development *(adsl, gart, gas1, qdpr*) and between color morphs (*atic, qdpr*). A number of these genes *(adsl, dct, dock7, gart, qdpr, rabggta*) have been implicated in previous work in dendrobatids (Rodríguez et al. 2020; Stuckert et al. 2019) and in other taxa. The genes that have been implicated in previous studies are clearly good candidates for future study. Additionally, the ARFs and Rab GTPases that we identified and that are upregulated in iridophores relative to other chromatophore types in fish (e.g., Higdon et al. 2013) are also good candidates. Unfortunately, our understanding of how these genes affect the development of iridophores is limited, and thus more targeted examinations of iridophore production and pigment production are needed. Notably, a number of epidermis-structuring genes (such as those in the krt family) have been implicated in the production of structural colors (e.g., *krt1, krt2*), although more evidence is needed to verify their role in coloration (Burgon et al. 2020; Stuckert et al. 2019; McGowan et al. 2006; Cui et al. 2016). We identified a number of these that are differentially expressed between color morphs (e.g., *krt1, krt2, krt8*). Genes that influence keratin, and organization of the epidermis generally, are good candidates for the production of different colors, as they may produce structural influences on color (via reflectance) in a manner that parallels what we see from guanine platelets. Keratins are known to influence the distribution and arrangement of melanosomes, which impact the ultimate color phenotype of animals (Gu and Coulombe 2007a, 2007b). Of particular interest is *krt8.* We found differential expression of *krt8* between developmental stages in all three species, as well as between morphs in both *R. fantastica* and *R. variabilis.* Previous work has identified *krt8* as a candidate gene for coloration in several taxa. For example, gene expression analyses indicate different expression patterns between colors

of monkey pelage which is related to hair thickness (Stubbs 2017) and color morphs of the poison frog *D. auratus* (Stuckert et al. 2019). In their in-depth analysis of genes controlling coloration, Linderoth et al. (Linderoth et al. 2023) identified *krt8.2* as a putative causative agent in leg coloration (*krt8.2* is the *Xenopus* ortholog of the human *krt8*, and thus naming differences likely primarily stem from databases used in annotation between studies). Thus, color and pattern may be in part driven by both variants and differential expression of *krt8*. The combination of differential expression between colors and/or color morphs in multiple expression studies and a plausible mechanism for color modification suggest that genes in the keratin family may well be important for producing color differences, although detailed analyses would need to be conducted to confirm this.

New candidate genes:

Many of the genes discussed above are those that have been previously implicated in other taxa. As a result, these are subject to ascertainment bias. One of the novel aspects of this work is the comparative approach across species, and that it is occurring in non-model organisms. We leveraged this to examine several genes which had high log-fold changes and were differentially expressed between morphs in convergent species (i.e., between spotted and striped morphs of *R. imitator* and *R. variabilis* or between banded and redheaded morphs of *R. imitator* and *R. fantastica*), potentially indicating their importance. There are many more genes that fall into this category than we can discuss. So here we highlight genes in which we were able to decipher plausible roles in coloration based on their function in other organisms. Among these are *ep-cadherin*, *vat1l*, *saa3, selenoi*, *ovol1*, *piwil1, pdc*, *ddb_g0268948*, *cd36*, *perlwapin*.

1307 *Ep-cadherin* was first identified as a novel protein in *Xenopus* that codes for a cadherin
1308 which shares homology to mammalian *E-cadherin* and *P-cadherin* (Ginsberg et al. 1991). In humans,
1309 *E-cadherin* is known to influence melanin transfer in human skin cells (Singh et al. 2017). Cadherins
1310 are key players in tissue organization (Niessen et al. 2011), and tissue organization in the skin is
1311 known to play a role in coloration (Duellman and Trueb 1986; Bagnara and Matsumoto 2006;
1312 Twomey et al. 2020). *Ep-cadherin* has not been implicated in vertebrate coloration, but given the
1313 evidence it is a plausible candidate. *Vat1l* has recently been implicated as a contributor to different
1314 colors in primate pelage via a gene expression analysis (Stubbs 2017), but additional evidence is
1315 lacking. Our results in combination with this previous work suggests that coloration may be an
1316 unappreciated aspect of *vat1l*'s function. To our knowledge, *saa3,* serum amyloid A 3, has not
1317 previously been implicated in pigmentation, but *saa3* is known to bind retinol in vivo and thus is a
1318 good candidate for pigmentation (Derebe et al. 2014). The gene *ddb_g0268948* is a putative
1319 methyltransferase that was recently implicated in color changes associated with ripening in
1320 tomatoes (Xiao et al. 2022) and leaf color in ornamental cabbage (Jin et al. 2018). However, this
1321 gene is poorly characterized and its basic functions remain unknown. Many methyltransferases are
1322 involved in DNA methylation, and can be involved in gene silencing, alternative splicing,
1323 transcriptional activation, and post-transcriptional regulation (Lyko 2018). Therefore the role of
1324 *ddb_g0268948* is likely not in producing pattern or pigmentation *per se,* but rather this gene, if
1325 involved in coloration in vertebrates, is likely driving transcription of genes that are directly
1326 involved in pigment production.
1327 The gene *cd36,* known as CD36 molecule or occasionally SCARB3, is a Scavenger Receptor
1328 Class B member. Two other genes in this class, *scarb1* and *scarb2,* play a role in carotenoid-based
1329 coloration in organisms via carotenoid uptake (e.g., Toomey et al. 2017), and have been implicated

in poison frog coloration (Stuckert et al. 2019; Rodríguez et al. 2020). Further, a related gene, *cameo2*, has been putatively implicated as playing a role in deposition of carotenoids of cocoon silk (Sakudoh et al. 2010). Although *cd36* has not been previously explicitly linked to coloration, it is known to impact carotenoid uptake in the retina (Shyam et al. 2017), and thus *cd36* may play a role in carotenoid uptake in the skin in *Ranitomeya* as well.

We also identified the ovol1 (ovo like transcriptional repressor 1) gene, which is involved in epidermal keratinization (Tsuji et al. 2018). Recently, Rubio et al. (in press) identified this as a gene under selection within the *Ranitomeya* lineage. As a result, *ovol1* may play a role in epidermal organization and the formation of structural colors.

The final gene we highlight is *perlwapin,* a gene that was characterized from molluscs. This gene is thought to play a role in biomineralization of mollusc shells (Marin et al. 2007; Le Luyer et al. 2019) via inhibiting the growth of certain crystallographic planes by binding selectively to edges and thereby preventing further growth (Treccani et al. 2006). There are two plausible crystalline structures which influence coloration that *perlwapin* may act upon: guanine platelets in the iridophores or the pteridine isoxanthopterin.