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

Adam M. M. Stuckert\textsuperscript{1,2,3} | Mathieu Chouteau\textsuperscript{4} | Melanie McClure\textsuperscript{4} | Troy M. LaPolice\textsuperscript{2,5} | Tyler Linderoth\textsuperscript{6} | Rasmus Nielsen\textsuperscript{6} | Kyle Summers\textsuperscript{3} | Matthew D. MacManes\textsuperscript{2}

\textsuperscript{1}Department of Biology and Biochemistry, University of Houston, Houston, Texas, USA
\textsuperscript{2}Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, New Hampshire, USA
\textsuperscript{3}Department of Biology, East Carolina University, Greenville, North Carolina, USA
\textsuperscript{4}Laboratoire Écologie, Évolution, Interactions Des Systèmes Amazoniens (LEEISA), CNRS, IFREMER, Université de Guyane, Cayenne, France
\textsuperscript{5}Department of Biology, Pennsylvania State University, University Park, Pennsylvania, USA
\textsuperscript{6}Department of Integrative Biology, University of California, Berkeley, California, USA

\textbf{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 defence, are excellent models to understand the link between vivid colour 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 colour 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 310Kbp, a scaffold N50 of 390Kbp and 85% of expected tetrapod genes. We leveraged this genome to conduct gene expression analyses throughout development of four colour morphs of \textit{Ranitomeya imitator} and two colour morphs from both \textit{R. fantastica} and \textit{R. variabilis} which \textit{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 \textit{qdpr} and \textit{xdh}) as a key driver of the variation in colour between morphs of these species, and identify several plausible candidates for colouration in vertebrates (e.g. \textit{cd36}, \textit{ep-cadherin} and...
1 | INTRODUCTION

The diversity of animal colouration in the natural world has long been a focus of investigation in evolutionary biology (Beddard, 1892; Fox, 1936; Gray & McKinnon, 2007; Longley, 1917). Colour phenotypes can be profoundly shaped by natural selection, sexual selection, or both. Further, these colour phenotypes are often under selection from multiple biotic (e.g. competition and predation) and abiotic (e.g. temperature and salinity) factors (Rudh & Qvarnström, 2013). The mechanisms underlying colour 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 defence), provide examples of the effects of strong selection on such phenotypes (Kang et al., 2017; Ruxton et al., 2004; Sherratt, 2006, 2008). In these biological systems, geographically heterogeneous predation produces rapid diversification of colour and pattern within a species or group of species. This produces a diversity of polytypic phenotypes (defined as distinct defensive warning colour signals in distinct localities) that are maintained geographically, with each population characterised by a unique phenotype that deters predators. This spatial mosaic of local adaptations maintained by the strong stabilising 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 & Barton, 1989), velvet ants (Wilson et al., 2015), millipedes (Marek & Bond, 2009) and poison frogs (Stuckert, Saporito, et al., 2014; Stuckert, Venegas, & Summers, 2014; Symula et al., 2001, 2003). 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 colour patterns, as the different R. imitator colour 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 colour production in amphibians in general (Burgon et al., 2020), and poison frogs in particular (Rodríguez et al., 2020; Stuckert et al., 2019; Twomey, Johnson, et al., 2020), 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 to 11 Gb, with an interquartile range of 3–5 Gb; 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 et al., 2018; Rogers et al., 2018). As a result, there is a relative dearth of publicly-available amphibian genomes (14 anuran species as of 1 September 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 et al., 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, R. 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 utilised this high quality R. 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 colour throughout development both between colour morphs and between species. We aimed to identify the genes responsible for colour patterning that are convergent between model and mimic, as well as those genes whose role in colour and pattern may be species-specific or even population-specific. Colour 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 colour patterns and mimicry in poison frogs.

2 | METHODS

2.1 | Permits

2.1.1 | 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).

2.1.2 | 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 authorisation number 232-2016- SERFOR/ DGGSPFFS and export permit No 17PE 001718 and the authorisation from the French Direction de l’Environnement, de l’Agriculture, de l’Alimentation et de la forêt en Guyane number 973-ND0073/ SP2000116-13.

2.2 | Generating a de novo genome for Ranitomeya imitator

2.2.1 | Genome sequencing approach

Because all sequencing technologies are known to be biased in both known and unknown ways, we utilised a variety of sequencing technologies to assemble this complex and large genome. At the time of genome construction, both sequencing technologies and assembly algorithms were undergoing rapid change, and as such the generation of an optimal assembly required substantial trial and error. For instance, several of the authors of this paper were involved in the strawberry poison frog genome assembly Rogers et al. (2018) and the difficulties encountered in the attempt to assemble that genome made it clear that short read data alone are insufficient if the goal is to assemble a highly-contiguous and complete genome. To overcome the limitations associated with short read data, we collected linked and long read data from technologies thought to be complementary to one another (Illumina 10X, Oxford Nanopore, and Pacific Biosystems).

2.2.2 | 10X Chromium

A single, likely male, subadult R. imitator of the ‘intermedius’ morph from the amphibian pet trade was used to produce a 10X library. This individual had no visible ovaries, and sex chromosomes in poison frogs are currently uncharacterised and therefore are not a mechanism used to identify sex. This frog was euthanised and high molecular weight DNA was extracted from liver tissue using the QIAGEN Blood & Cell Culture DNA Kit. 10X Genomics Chromium Genome library (Weisenfeld et al., 2017) was prepared by the DNA Technologies and Expression Analysis Cores at the University of California Davis Genome Center and sequenced on an Illumina HiSeq X by Novogene Corporation (Mudd et al. in prep).

2.2.3 | Long read (Oxford nanopore and Pacific BioSciences) sequencing

Captive bred subadult frogs from the pet trade that originated from the region near Tarapoto, Peru (green-spotted morph, Figure 1) were euthanised and the skin and gastrointestinal tract

FIGURE 1 Normative photographs of frogs from the populations used in this study. The centre box represents the four mimetic colour 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 colour 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. [Colour figure can be viewed at wileyonlinelibrary.com]

[Image of four frogs: R. variabilis, R. imitator, R. fantastica, and R. variabilis]
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 8 M cells on a Pacific Biosciences Sequel II.

2.2.4 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 & 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 & Durbin, 2009) before running P RNA_Scaffold with default settings. We then used Arcs version 3.82 (Coome 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/arc5/blob/master/Examples/arscs-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_Scaffold 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 (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.

2.2.5 Repeats and genome annotation

We modelled genomic repeats with Repeat Modeller version 2 (Smit and Hubley, 2008–2015) using RepBase database 20170127. Repeat families modelled by Repeat Modeller 2 were annotated using Transposon Classifier ‘RFSB’ (Riehl et al., 2022). We extracted the vertebrate specific repeats from RepBase and merged this with the classified consensus output from Repat Modeller 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’). We produced this transcript evidence from gene expression data of various R. imitator tissues and populations which we assembled.

FIGURE 2 Flowchart of genome assembly approach. Colours of boxes represent the type of data used in that step (see internal legend) and italicised font indicate the program(s) used. [Colour figure can be viewed at wileyonlinelibrary.com]
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 Data S1. 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 to a second round of Maker annotation with the flag ‘snaphm1’ 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 three total rounds of annotation with gene prediction in this manner, then visualised 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 visualising 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.

2.3 | Identification of colour pattern candidate genes

2.3.1 | 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 analysed in a manner concordant with these differences.

2.3.2 | 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 (red-headed) 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 (Table S1).

Tadpoles were anaesthetised 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 standardised Trizol protocol. RNA was extracted from the whole skin using a standardised 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 50bp paired end reads on three lanes of the Illumina HiSeq 2500 at the New York Genome Center.

2.3.3 | Ranitomeya fantastica and R. variabilis

We set up a captive colony in Peru (see Data S1) consisting of between six 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 five times a week. Tadpoles were raised individually in 21 oz 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 Table S1) were fixed in an RNAlater (Ambion) solution. To do so, tadpoles were first euthanised in a 250mg/L benzocaine hydrochloride bath, then rinsed with distilled water before the whole tadpole was placed in RNAlater and stored at 4°C for 6 h 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 standardised 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.

2.3.4 | Differential gene expression

We indexed our new R. 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 five 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 et al., 2015).
Differential expression analyses were conducted in R version 4.2.2 (Team, 2019) using the package DESeq2 version 1.38.3 (Love et al., 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 htsesq-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. We tested the effect of colour morph and the other which tested the effect of developmental stage. Both models included sequencing lane, tadpole age, and colour morph as fixed effects. We used a 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 colour genes that play a role in colour or pattern production in other taxa. This a priori list was originally used in Stuckert et al. (2019), but was updated by searching for genes that have been implicated in colouration in genomics studies from the last 3 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 colour 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 colour morph as fixed effects. We then identified the set of genes that are differentially expressed between colour 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 colour morphs in R. imitator (via the Likelihood Ratio Test described above) for genes that are differentially expressed between colour morphs in R. fantastica or R. variabilis (via the Walds tests we conducted).

3 | RESULTS

3.1 | Ranitomeya imitator genome

3.1.1 | Genome assembly

Our final genome assembly was 6.02 Gbp in length and consisted of 79,800 scaffolds ranging from 387 to 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 to 4,395,411 bp). A total of 6641 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 orthologues and 1.8% duplicated orthologues. An additional 4.9% were fragmented and 10.3% were missing.

3.1.2 | Repetitive elements

Our analyses indicate a high proportion of repeats in the R. imitator genome. Repeat Modeller masked 75.69% of total bases in the genome, of which 74.63% consisted of repeat elements (Table S2). 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 >1 Gbp in total genomic content: hAT (1.13 Gbp) and Gypsy (1.2 Gbp; see Figure 3). 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 (Table 1). Summary statistics on the number of instances, range, average length, and standard deviation of range can be found in Table S2.

3.1.3 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 artefact 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, 2018); additional details on this protocol can be found in the Data S1. 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

| Element type         | Number of elements | Total length of elements (bp) | Percentage of genomic sequence |
|----------------------|--------------------|--------------------------------|-------------------------------|
| Retroelements        | 2,940,777          | 1,853,517,364                  | 30.77                         |
| SINEs                | 9140               | 1,074,368                      | 0.02                          |
| Penelope             | 155,182            | 74,628,225                     | 1.24                          |
| LINEs                | 604,697            | 430,550,850                    | 7.15                          |
| CRE/SLACS            | 0                  | 0                              | 0                             |
| L2/CR1/Rex           | 355,946            | 294,429,256                    | 4.89                          |
| R1/LOA/Jockey        | 268                | 27,552                         | 0.00                          |
| R2/R4/NeSL           | 10,535             | 2,162,443                      | 0.04                          |
| RTE/Bov-B            | 16,036             | 8,907,551                      | 0.15                          |
| L1/CIN4              | 64,839             | 46,701,023                     | 0.78                          |
| LTR elements         | 2,326,940          | 1,421,892,146                  | 23.61                         |
| BEL/Pao              | 14,889             | 16,156,219                     | 0.27                          |
| Ty1/Copia            | 335,291            | 90,264,134                     | 1.50                          |
| Gypsy/DIR51          | 1,755,793          | 1,231,339,651                  | 20.44                         |
| Retroviral           | 122,545            | 50,623,347                     | 0.84                          |
| DNA transposons       | 8,305,215          | 2,638,017,678                  | 43.79                         |
| Hobo-Activator       | 3,302,857          | 1,409,227,847                  | 23.39                         |
| Tc1-Is630-Pogo       | 894,738            | 463,756,460                    | 7.70                          |
| En-Spm               | 0                  | 0                              | 0.0                           |
| MULE-MuDR            | 3713               | 183,543                        | 0.00                          |
| PiggyBac             | 30,726             | 19,118,673                     | 0.32                          |
| Tourist/Harbinger    | 68,929             | 21,248,842                     | 0.35                          |
| Other (Mirage, P-element, Transib) | 365,851 | 61,666,857                     | 1.02                          |
| Rolling-circles      | 12,894             | 5,925,694                      | 0.10                          |
| Unclassified         | 16,324             | 3,669,707                      | 0.06                          |
| Total interspersed repeats | 4,495,204,749 | 74.63                          |
| Small RNA            | 6867               | 775,734                        | 0.01                          |
| Satellites           | 7136               | 3,040,805                      | 0.05                          |
| Simple repeats       | 669,788            | 47,935,402                     | 0.80                          |
| Low complexity       | 81,259             | 7,023,995                      | 0.12                          |

Note: 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 modelled repeats from *Ranitomeya imitator*. 

**TABLE 1** Repeat elements classified by Repeat Masker.
reads and mapping rates in each sample please see (Table S3). 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 4). 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 colour morph and developmental stage for each species independently. For a list of all differentially expressed colour genes see Table S4. For a list of all differentially expressed genes at alpha <0.01 see Table S5.

### 3.1.3.1 | Between developmental stage comparisons

In our comparison of developmental stages, we found many differentially expressed genes (q value <.01) in each species (R. imitator = 2264, R. fantastica = 2350, R. variabilis = 3016; Table 2). Most of these are unlikely to be related to colour 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 colour 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 and scarb1; Figure 5), the synthesis of pteridines (e.g. ghchfr and qdpr; Figure 5), genes related to melanophore development and melanin synthesis (dct, kit, mif, mlp, notch1, notch2, sfxn1, sox9, sox10, tyr and tyrp1; Figure 6), genes putatively related to the production of iridophores and their guanine platelets (e.g. gart, gas1, paics, pax3-a, rab27a, rab27b and rab7b; Figure 7), and genes related to patterning (notch1 and notch2).

#### 3.1.3.2 | Between morph comparisons

In our comparison of colour morph we found many significantly differentially expressed genes in each species (R. imitator = 1558, R. fantastica = 1266, R. variabilis = 1168; Table 2). Most of these are unlikely to be related to colour and patterning, although a small fraction of differentially expressed genes (average 2.5%) are related to the generation of colour or patterning in other taxa (R. imitator = 33, R. fantastica = 36, R. variabilis = 29). Amongst genes that were differentially expressed between colour morphs we identified genes related to carotenoid metabolism or xanthophore production (e.g. aldh1a1, pax7 and scarb1; Figure 5), the synthesis of pteridines (e.g. ghchfr and qdpr; Figure 5), genes related to melanophore development and melanin synthesis (kit, mlp, sfxn1 and sox9), and genes putatively related to the production of iridophores and their guanine platelets (e.g. atic, dock7, gart, paics, pax2, pax3-a, rab27a, rab27b, rab7a and rabggtt; Figure 7).

### 3.1.4 | 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 versus spotted comparison, we identified 345 differentially expressed genes in *R. imitator* and 1407 in *R. variabilis*. Of these genes, 55 were shared between the two species and 10 were in our a priori colour gene list (*cd36*, *crabp2*, *edaradd* and *mvd*). In the banded versus redheaded comparison, we identified 1025 differentially expressed genes in *R. imitator* and 1488 in *R. fantastica*. Of these genes, 141 were shared between the two species and one was in our a priori colour 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 amongst 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], *vat11*, *olfactory* protein, *piwi1*, *saa3*, *pdc*, *ddb_g0268948*, *cd36*, *oval1* and *perlwapin*). We briefly discuss these genes and their potential import in the Data S1.

4 | DISCUSSION

The genetic, biochemical, cellular, physiological and morphological mechanisms that control colouration in mimetic systems are of interest because of their substantial impacts on survival and reproductive success. Despite this, these mechanisms are poorly characterised. 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 colour production within species, illuminating the mechanisms underlying Müllerian mimicry, and providing a rich foundation upon which future research may be built.

4.1 | 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 contiguity and genic content indicates that long read technologies are, unsurprisingly, critically important and can produce genomes with contiguity spanning large regions, even for species with large genomes containing many long, repetitive regions. Further, the relatively high error rate of long read technologies (even prior to the recent advent of
higher per-base accuracy in long read technologies) does not preclude the ability to assemble and identify genes, as evidenced by our high BUSCO score. This genome is a valuable resource and is well-suited to a variety of future work, especially RNA sequencing analyses like those we present below.

4.2 | Repeat elements in the *Ranitomeya imitator* genome

Over 77% of our genome assembly consists of repeats. This is a larger proportion of the genome than that found in the strawberry poison frog (*O. pumilio*), which was estimated to consist of ~70% repeats (Rogers et al., 2018). However, this is almost certainly due to the availability of long-reads in our study. Given that the *O. pumilio* assembly is both larger, and built from short-reads, their identification of repeat elements is likely an under-estimate of total repeat content. In comparison to these two poison frog genomes, the genome of *Xenopus tropicalis* is a little over one-third repeat elements and *Nanorana parkeri* is ~48% repeats, which are both comparable to a mammalian genome (Hellsten et al., 2010; Sun et al., 2015). Scattered throughout the *R. imitator* genome are a large number of repeat elements which represent over half of the assembled genome. We found that some classes of repeat elements contained similar abundance in the *R. imitator* and *O. pumilio* assemblies (e.g. Gypsy: 1.2Gb in *R. imitator* and 1Gb in *O. pumilio*), whereas several other repeat classes had strikingly different abundance (e.g. hAT: 1.1Gb in *R. imitator* and 255Mbp in *O. pumilio*; Copia: 97Mbp in *R. imitator* and 298Mbp in *O. pumilio*). This indicates that repeat elements have undergone different patterns of expansion or contraction 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 analysing repeats (RepDeNovo in *O. pumilio* and RepeatMasker in *R. imitator*) which are sources of potential bias.

4.3 | Gene expression

While the colours and patterns of *Ranitomeya* poison frogs are extremely variable, they always consist of vivid colour patches overlaying a background that is largely black. Recent evidence indicates that much of the differences in colour 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 colouration in the yellow to red ranges of the visible spectrum. Black and brown colouration is produced by melanophores and the melanin pigments found within the chromatophores (Bagnara et al. 1968; Duellman & 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 melamin synthesis pathway (Posso-Terranova & Andrés, 2017; Stuckert et al., 2019), pteridine synthesis pathway (Rodríguez et al., 2020; Stuckert et al., 2019) and carotenoid processing pathways (Twomey, Johnson, et al., 2020) for their roles in producing different colour 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 colour morphs in both the model and mimic. Of those genes that were shared, only five were in our a priori colour gene list (*cd36*, *crabp2*, *edaradd*, *mvd* and *adpr*). This seems to indicate: (1) the pattern of differential gene expression affecting convergent colour morph development is largely species-specific, (2) the expression patterns of a small number of genes have a very large effect on colour morph divergence, and/or (3) we have insufficient power to identify these convergent genes. Overall, these results suggest that the convergent colour patterns of these species are likely to have evolved via the expression of distinct underlying genetic and biochemical pathways. In sum, colour 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 colouration per se. Nevertheless, we identified a number of very promising candidate colour genes that are likely to play a role in the production of mimetic phenotypes in this system. In the rest of this 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 Data S1.

4.4 | Yellow, orange and red colouration

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 & 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 (Croucher et al., 2013; Grether et al., 2001; Mcgraw et al., 2006; McLean et al., 2017; Obika & Bagnara, 1964). In this study we found a number of key pteridine synthesis genes that were differentially expressed between colour morphs (Figure 5). Prominent amongst these are the aforementioned xanthine dehydrogenase (*xdh*) and quinoid dihydropteridine reductase (*adpr*). Xanthine dehydrogenase appears to be highly conserved and its expression plays a role in the production of pterin-based colouration in a variety of taxa such as spiders (Croucher et al., 2013), fish (Parichy et al., 2000; Salis et al., 2019), salamanders (Frost & Bagnara, 1979; Thorsteinsdottir & Frost, 1986),
and the dendrobatid frogs *D. auratus* and *O. pumilio* (Rodríguez et al., 2020; Stuckert et al., 2019). 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 (Ponzon 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 5), notably paired box 7 (pax7) and xanthine dehydrogenase (*xdh*), both of which have been linked to xanthophore differentiation (Epperlein & Löfberg, 1990; Nord et al., 2016; Parichy et al., 2000; Reaume et al., 1991).

### 4.5 | Melanophore genes

Melanin-based colouration is the best understood aspect of colouration, 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 colouration 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 colour 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 colour and pattern, both throughout development and between colour 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*, *mph*, *notch1*, *notch2*, *sox9*, *sox10*, *tyr* and *tyrpt1*, 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* and *notch2*; Figure 6) 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.

### 4.6 | Iridophore genes

Iridophores are largely responsible for white, blue and green colouration, which is mainly determined by the reflection of light from iridophores (Bagnara et al. 2007). Recently, Twomey, Johnson, et al. (2020); Twomey, Kain, et al. (2020) found that variation in colouration 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*, *dct*, *edn3*, *erbb3*, *impdh2*, *paics*, *rab27a* and *rab27b*) or colour morphs (arf6, *dock7*, *dct*, *erbb3*, *gart*, *gne*, *paics*, *rab27a*, *rab27b*, *rab7a* and *rabbgta*) in our study (Figure 7). We also found differential expression of a number of genes that are known to impact guanine or purine synthesis throughout development (*adsl*, *gart*, *gas1* and *qdpr*) and between colour morphs (*atic* and *qdpr*). A number of these genes (*adsl*, *dct*, *dock7*, *gart*, *qdpr* and *rabbgta*) 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 colours (e.g. *krt1* and *krt2*), although more evidence is needed to verify their role in colouration (Burgon et al., 2020; Cui et al., 2016; McGowan et al., 2006; Stuckert et al., 2019). We identified a number of these that are differentially expressed between colour morphs (e.g., *krt1*, *krt2* and *krt8*). Genes that influence keratin, and organisation of the epidermis generally, are good candidates for the production of different colours, as they may produce structural influences on colour (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 colouration in several taxa (Stubbs 2017; Stuckert et al., 2019). In their in-depth analysis of genes controlling colouration, Linderoth et al. (2023) identified *kr8.2* as a putative causative agent in leg colouration (*kr8.2* is the *Xenopus* ortholog of the human *kr8*), and thus naming differences likely primarily stem from databases used in annotation between studies). Thus, colour and pattern may be in part driven by both variants and differential expression of *krt8*.

### 4.7 | 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 colouration.

STUCKERT ET AL.
based on their function in other organisms. Amongst these are ep-cadherin, vat11, saa3, seleno, ovol1, piwi1, pdc, ddb_g0268948, cd36 and perhwapin.

5 | CONCLUSION

In this study we examined the molecular mechanisms by which mimetic phenotypes are produced in a Müllerian mimicry system. Through our efforts, we have produced the first high-quality poison frog genome, a 6.02 Gbp, contiguous genome assembly with good genic coverage. We leveraged this to examine gene expression in the skin throughout development of four comimetic morphs from three species of Ranitomeya. We identified a large number of genes related to melanophores, melanin production, iridophore production and guanine synthesis that were differentially expressed throughout development, indicating that many of these are important in the production of pigmentation, albeit not colour morph specific colouration. Genes related to xanthophore production, carotenoid pathways, melanin production and melanophore production were rarely differentially expressed between colour morphs, however those genes that were differentially expressed may be critically important in producing polytypic differences within species that drive mimetic phenotypes. Our results indicate that divergence between colour morphs seems to involve differences in expression and/or timing of expression, but that convergence for the same colour 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 colour morphs across species. Thus, production of these colours 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 colours and patterns. Further, we highlight the potential importance of genes in the keratin family for producing differential colour via structural mechanisms.

AUTHOR CONTRIBUTIONS

Designed research: AMMS, MC, MM, RN, KS, MDM. Performed research: AMMS, MC, MM, KS, TML. Analysed data: AMMS, TML. Contributed to writing: AMMS, MC, MM, TML, TL, RN, KS, MDM.

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

FUNDING INFORMATION

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, Extreme Science and Engineering Discovery Environment (XSEDE) XSEDE Resource Allocation (XRAc) MCB110134 to MDM and AMMS, and University of Houston startup to AMMS. MC received support from Agence Nationale de la Recherche (ANR) grants RANAPAOSA (ref. ANR-20-CE02-0003), from an ‘Investissement d’Avenir’ grant CEBA (ref. ANR-10-LABX-25-01) and from a Marie Skłodowska-Curie fellowship (FITINV, N 655857).

CONFLICT OF INTEREST STATEMENT

The authors confirm that they have no conflict of interest.

OPEN RESEARCH BADGES

This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available on Data Dryad (https://doi.org/10.5061/dryad.gtt76hvp).

DATA AVAILABILITY STATEMENT

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). Gene counts are available on Data Dryad (https://doi.org/10.5061/dryad.gtt76hvp). Code is available at Zenodo (https://zenodo.org/doi/10.5281/zenodo.4758403).

ORCID

Adam M. M. Stuckert https://orcid.org/0000-0003-2250-2223

REFERENCES

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A python framework to work with high-throughput sequencing data. Bioinformatics, 31(2), 166–169.

Bagnara, J. T., Fernandez, P. J., & Fujiij, R. (2007). On the blue coloration of vertebrates. Pigment Cell Research, 20(1), 14–26.

Bagnara, J. T., Taylor, J. D., & Hadley, M. E. (1968). The dermal chromatophore unit. The Journal of Cell Biology, 38(1), 67–79.

Beddard, F. E. (1892). Animal coloration: An account of the principal facts and theories relating to the colours and markings of animals. S. Sonnenschein & Company.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Briolat, E. S., Burdfield-Steel, E. R., Paul, S. C., Rönkä, K. H., Seymour, B. M., Stankovich, T., & Stuckert, A. M. M. (2019). Diversity in warning coloration: Selective paradox or the norm? Biological Reviews of the Cambridge Philosophical Society, 94(2), 388–414.

Brown, J. L., Twomey, E., Amezquita, A., DeSouza, M. B., Caldwell, J., Lötters, S., Von May, R., Melo-Sampaio, P. R., Mejia-Vargas, D.,...
Park, H. Y., Kosmadaki, M., Yaar, M., & Gilchrest, B. A. (2009). Cellular mechanisms regulating human melanogenesis. *Cellular and Molecular Life Sciences, 66*(9), 1493–1506.

Ponzone, A., Spada, M., Ferraris, S., Dianzani, I., & De Sanctis, L. (2004). Dihydropyrrteridine reductase deficiency in man: From biology to treatment. *Medicinal Research Reviews, 24*(2), 127–150.

Posso-Terranova, A., & Andrés, J. A. (2017). Diversification and convergence of apomictic phenotypes: Truncated receptors and cellular arrangements mediate rapid evolution of coloration in harlequin poison frogs. *Evolution, 71*(11), 2677–2692.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (Version 3.6.0) [Computer Software]. R Foundation for Statistical Computing.

Reaume, A. G., Knecht, D. A., & Chovnick, A. (1991). The rosy locus in *Drosophila melanogaster*: Xanthine dehydrogenase and eye pigments. *Genetics, 129*(4), 1099–1109.

Reed, R. D., Papa, R., Martin, A., Hines, H. M., Kronforst, M. R., Chen, R., Georg Haider, H., Nijhout, F., & Owen Mcmillan, W. (2011). Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science, 333*(August), 1137–1141.

Riehl, K., Riccio, C., Miska, E. A., & Hemberg, M. (2022). TransposonUltimate: Software for transposon classification, annotation and detection. *Nucleic Acids Research, 50*(11), e64–e64.

Rodriguez, A., Mundy, N. I., Ibáñez, R., & Pröhl, H. (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, R. L., Zhou, L., Chu, C., Márquez, R., Corl, A., Linderoth, T., Freeborn, L., MacManes, M. D., Xiong, Z., Zheng, J., Guo, C., Xun, X., Kronforst, M. R., Summers, K., Wu, Y., Yang, H., Richards-Zawacki, C. L., Zhang, G., & Nielson, R. (2018). Genomic takeover by transposable elements in the strawberry poison frog. *Molecular Biology and Evolution, 35*(12), 2913–2927.

Ruan, J., & Li, H. (2019). Fast and accurate Long-read assembly with wtdbg2. *Nature Methods, 17*(2), 155–158. https://doi.org/10.1038/s41592-019-0669-3

Rudh, A., & Qvarnström, A. (2013). Adaptive colouration in amphibians. *Seminars in Cell and Developmental Biology, 24*(6-7), 553–561.

Ruxton, G. D., Sherratt, T. N., & Speed, M. P. (2004). Avoiding attack: The evolutionary ecology of crypsis, warning signals and mimicry (Vol. 17). Oxford University Press.

Salis, P., Lorin, T., Lewis, V., Rey, C., Marcionetti, A., Escande, M.-L., Roux, N., Besseau, L., Salamin, N., Sémon, M., Parichy, D., Voll, J.-N., & Laudet, V. (2019). Developmental and comparative transcriptomic identification of iridophore contribution to white barring in clownfish. *Pigment Cell and Melanoma Research, 32*(3), 391–402.

Schulte, R. (1986). Eine Neue Dendrobatas—art Aus Ostperu (Amphibia: Salienta: Dendrobatidae). *Sauria, 8*, 11–20.

Sherratt, T. N. (2006). Spatial mosaic formation through frequency-dependent selection in Mullerian mimicry complexes. *Journal of Theoretical Biology, 240*(2), 165–174.

Sherratt, T. N. (2008). The evolution of Müllerian mimicry. *Die Naturwissenschaften, 95*(8), 681–695.

Simão, F. A., Waterhouse, R. M., Ioanidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics, 31*(19), 3210–3212.

Sk öld, H. N., Aspengren, S., Cheney, K. L., & Wallin, M. (2016). Fish chromatophores—from molecular motors to animal behavior. *International Review of Cell and Molecular Biology, 321*, 171–219.

Smit, A. F. A., & Hubley, R. (2008–2015). RepeatModeler Open-1.0. [http://www.repeatmasker.org]

Smit, A. F. A., Hubley, R., & Green, P. (2013–2015). RepeatMasker Open-4.0. [http://www.repeatmasker.org] Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics-Oxford, 19*(2), 215–225.

Stubbs, W. K. (2017). Forensic applications of associating human scalp hair morphology and pigmentation analysis at the microscopic and molecular level. PhD diss., Purdue University.

Stuckert, A. M. M., Moore, E., Coyle, K. P., Davison, I., MacManes, M. D., Roberts, R., & Summers, K. (2019). Variation in pigmentation gene expression is associated with distinct apomictic color morphs in the poison frog, *Dendrobates Auratus*. *BMC Evolutionary Biology, 19*(1), 85.

Stuckert, A. M. M., Saporito, R. A., Venegas, P. J., & Summers, K. (2014). Alkaloid defenses of co-mimics in a putative Müllerian mimetic radiation. *BMC Evolutionary Biology, 14*, 76.

Stuckert, A. M. M., Venegas, P. J., & Summers, K. (2014). Experimental evidence for predator learning and Müllerian mimicry in Peruvian poison frogs (*Ranitomeya. Dendrobatidae*). *Evolutionary Ecology, 28*(3), 413–426.

Sun, Y.-B., Xiong, Z.-J., Xiang, X.-Y., Liu, S.-P., Zhou, W.-W., Tu, X. L., Zhong, L., Wang, L., Wu, D. D., Zhang, B. L., Zhu, C. L., Yang, M. M., Chen, H. M., Li, F., Zhou, L., Feng, S. H., Huang, C., Zhang, G. J., Irwin, D., ... Zhang, Y. P. (2015). Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. *Proceedings of the National Academy of Sciences of the United States of America, 112*, E1257–E1262.

Supple, M. A., Hines, H. M., Dasmahapatra, K. K., Lewis, J. J., Nielsen, D. M., Lavoie, C., Ray, D. A., Salazar, C., Mcmillan, W. O., & Counterman, B. A. (2013). Genomic architecture of adaptive color pattern divergence and convergence in *Heliconius* butterflies. *Genome Research, 23*(662), 1248–1257.

Symula, R., Schulte, R., & Summers, K. (2001). Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proceedings of the Royal Society B: Biological Sciences, 268*(1484), 2415–2421.

Symula, R., Schulte, R., & Summers, K. (2003). Molecular systematics and phylogeography of Amazonian poison frogs of the genus *Dendrobates*. *Molecular Phylogenetics and Evolution, 26*(3), 452–475.

Thorsteindottir, S., & Frost, S. K. (1986). Pigment cell differentiation: The relationship between pterin content, allopurinol treatment, and the melanoid gene in axolotls. *Cell Differentiation, 19*, 161–172.

Twomey, E., Johnson, J. D., Castroviejo-Fisher, S., & Van Bocxlaer, I. (2020). A ketocarotenoid-based color polymorphism in the Sira poison frog *Ranitomeya siresis* indicates novel gene interactions underlying aposematic signal variation. *Molecular Ecology, 29*, 2004–2015. https://doi.org/10.1111/mec.15466

Twomey, E., Kain, M., Claeyts, M., Summers, K., Castroviejo-Fisher, S., & Van Bocxlaer, I. (2020). Mechanisms for color convergence in a mimetic radiation of poison frogs. *The American Naturalist, 195*(5), E132–E149. https://doi.org/10.1086/708157

Twomey, E., Yeager, J., Brown, J. L., Morales, V., Cummings, M., & Summers, K. (2013). Phenotypic and genetic divergence among poison frog populations in a mimetic radiation. *PLoS One, 8*(2), e55443.

UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research, 47*(D1), D506–D515.

Vaser, R., Šorić, I., Nagarajan, N., & Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Research, 27*(5), 737–746.

Warren, R. L. (2016). RAILS and Cobbler: Scaffolding and automated finishing of draft genomes using Long DNA sequences. *Journal of Open Source Software, 1*(7), 116.

Weissenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research, 27*(5), 757–767.

Wickham, H. (2011). *ggplot2*. Wiley Interdisciplinary Reviews: Computational Statistics, 3(2), 180–185.
Wilson, J. S., Jahner, J. P., Forister, M. L., Sheehan, E. S., Williams, K. A., & Pitts, J. P. (2015). North American velvet ants form one of the World’s largest known Müllerian mimicry complexes. Current Biology, 25(16), R704–R706.
Zhu, B.-H., Xiao, J., Xue, W., Gui-Cai, X., Sun, M.-Y., & Li, J.-T. (2018). P_RNA_scaffolder: A fast and accurate genome scaffolder using paired-end RNA-sequencing reads. BMC Genomics, 19, 1–13.

SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Stuckert, A. M. M., Chouteau, M., McClure, M., LaPolice, T. M., Linderoth, T., Nielsen, R., Summers, K., & MacManes, M. D. (2024). The genomics of mimicry: Gene expression throughout development provides insights into convergent and divergent phenotypes in a Müllerian mimicry system. Molecular Ecology, 33, e17438. https://doi.org/10.1111/mec.17438