

Genome Scanning Reveals the Genetic Basis of a Color Pattern Morphotype in an Island Population of the European Adder (*Vipera berus*)

Nina Casillas ¹, Bent Petersen ^{2,3}, Sarah L.F. Martin ¹, Kjeld Henrik Ophus⁴, Henrik Bringsøe ⁵, Jaelle C. Brealey ^{1,6}, Vanessa C. Bieker ^{1,7}, Jaime Morin-Lagos ¹, Dag Dolmen ¹, Thomas Madsen ⁸, Rasmus Nielsen ⁹, Morten E. Allentoft ^{10,11}, José Cerca ^{1,12,13,14,†}, Michael D. Martin ^{1,*,†}

¹Department of Natural History, University Museum, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

²Center for Evolutionary Hologenomics, Globe Institute, University of Copenhagen, Copenhagen, Denmark

³Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), AIMST University, Bedong, Kedah 08100, Malaysia

⁴Steinrøysa 16, Torp 1658, Norway

⁵Irisvej 8, Køge DK-4600, Denmark

⁶Land and Biodiversity, Norwegian Institute of Nature Research (NINA), Trondheim 7034, Norway

⁷Royal Botanic Gardens, Kew, Richmond, Surrey, UK

⁸Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Waurn Ponds, Victoria, Australia

⁹Center for Theoretical Evolutionary Genomics, Department of Integrative Biology, University of California, Berkeley, CA, USA

¹⁰TrEnD Laboratory, School of Molecular and Life Science, Curtin University, Perth, Australia

¹¹GeoGenetics Section, Globe Institute, University of Copenhagen, Copenhagen, Denmark

¹²Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway

¹³Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden

¹⁴SciLifeLab, Karolinska Institutet Science Park, Solna, Sweden

†Jointly contributed to research supervision.

*Corresponding author: E-mail: mike.martin@ntnu.no.

Accepted: April 27, 2026

Abstract

Color polymorphism is an important trait due to its role in defense strategies, feeding habits, environmental responses such as temperature regulation, and overall fitness. While the genetic basis of color polymorphism is well understood in mammals, it remains relatively understudied in nonavian reptiles. *Vipera berus*, the European adder, is the most widely distributed and northernmost terrestrial snake in the world. Similar to other European vipers, *V. berus* exhibits a range of dorsal color patterns spanning from zig-zag to complete melanism. However, a unique longitudinal dorsal stripe pattern was recently discovered in the isolated Gossa Island (~45 km²) population of Norway, where ~5% of snakes exhibit this distinct pattern instead of the common dorsal zig-zag pattern. In this study, we investigated the genetic basis of the unique phenotype observed on Gossa by combining low-depth whole-genome shotgun sequencing and pairwise F_{ST} genome scanning. We discovered a moderately sized, significantly differentiated genomic region containing five associated genes, including the premelanosome gene (*PMEL*) that encodes a well-known transmembrane glycoprotein that

© The Author(s) 2026. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

is a key component of melanosome organelles. We implicate a functionally relevant variant in the encoded premelanosome protein. Our results provide insight into the evolution and genetic basis of pigmentation of squamate reptiles using the unbiased approach of genome scanning, a method that has rarely been used in this context.

Key words: *Vipera berus*, aposematism, crypsis, dorsal pattern, *PMEL*, pigmentation.

Significance

We provide strong evidence from population genomics that the genetic control of stripe pigmentation in the European adder (*Vipera berus*) is the *PMEL* gene that encodes the premelanosome protein. This is only the second time this gene has been implicated in stripe pigmentation in a snake, so the work suggests *PMEL* may play a universal role in determining squamate skin stripes. The *PMEL* gene is evolutionarily conserved and important in vertebrate development, but its entire range of functions is still unknown. Thus, this research provides evidence for further work toward *PMEL*'s full functional characterization.

Introduction

The investigation of intraspecific melanin-based polymorphism has captured the interest of researchers across various taxa, as it offers insights into sexual selection, predator–prey interaction, thermoregulation, ecological niches, pattern formation, and mimicry (Murray and Myerscough 1991; Roulin 2004; McKinnon and Pierotti 2010; Murray 2011). Snakes, in particular, have garnered interest due to their remarkable diversity of pattern and color variations that are unparalleled among vertebrates (Allen et al. 2013). While color polymorphism in mammals and birds has been the focus of numerous studies—largely centered on melanocyte biology—our understanding remains incomplete and is well resolved only in a handful of model species. Investigations into the genetic basis of color polymorphism in reptiles have been comparatively limited.

The primary protein targets implicated in studies of nonavian reptile color polymorphism have been pro-opiomelanocortin (*POMC*), agouti-signaling protein (*ASIP*), premelanosome protein (*SILV/PMEL*), the tyrosinase gene family, and the melanocortin 1 receptor (*MC1R*) (Rosenblum et al. 2004; Corso et al. 2012; Cox et al. 2013; Ducrest et al. 2014; Brown et al. 2022). Although melanocytes do contribute to pattern formation, in reptiles color patterning is more complex due to the presence within the dermis of additional pigmentation cells called chromatophores (Alibardi 2013; Ullate-Agote et al. 2020; Tang et al. 2023). In many reptiles (including most snakes and lizards), chromatophore stratification begins with melanophores at the base of the dermis, producing black pigment patterns (Kuriyama et al. 2020). Ascending from the melanophores are the iridophores, erythrophores, and xanthophores, which are located in the upper layers of the dermis and produce light-scattering effects and red/yellow pigments. These chromatophores are derived from

neural crest cells, which migrate in developing embryos and give rise to a vast range of cells in addition to chromatophores (Tang et al. 2023). Reptile pattern diversity is the result of the distribution and layering of chromatophores and melanocytes in the dermis and epidermis (Alibardi 2013; Saenko et al. 2013; Ullate-Agote et al. 2020). While the vertebrate model organism zebrafish has played a significant role in advancing understanding of chromatophores and pattern formation, there has been only limited research on the complex color and pattern morphology of reptiles (Kuriyama et al. 2020).

The European adder (*Vipera berus* Linnaeus, 1758; family Viperidae) is the most widely distributed



Fig. 1. Both *V. berus* morphs can be observed on the Norwegian island of Gossa. a) A male (individual M2) and female (individual F1) with the zig-zag dorsal pattern. b) A male (individual M4) and female (individual F4) with the comparatively rare dorsal stripe phenotype. Photos © 2017 by K. H. Ophus.

terrestrial snake in the world, with its range stretching from Sakhalin Island in the Far East of Russia to Scotland (Lourdais et al. 2013; Horreo and Fitze 2018) and from Northern Greece to above the Arctic Circle (Andersson 2003; Ursenbacher et al. 2006). Due to its northern distribution, in most parts of its range the adder lives at low altitudes, but in, eg southeastern Europe with warmer climatic conditions, *V. berus* is restricted to high altitudes (1,000 to 2,600 m above sea level; Ursenbacher et al. 2006). Despite the European adder's wide range, by far the most common skin pattern is a dorsal zig-zag (Fig. 1a). In some parts of its range (such as Scandinavia), however, melanism is common, with a population frequency ranging from 10% to >80% (Madsen et al. 2022). The zig-zag pattern had been suggested to be an example of aposematic coloration (Wüster et al. 2004), disruptive coloration (Valkonen et al. 2011; Bauwens and Claus 2019), and resulting in a flicker-fusion effect (Shine and Madsen 1994). *Vipera berus* is one of the few snakes that show sexual dichromatism from birth, and this dichromatism is maintained in adults, where males have a black zig-zag pattern with white, gray, or light blue coloration while females often have a dark brown zig-zag pattern and light brown body color (Andrén and Nilson 1981; Cui et al. 2016; Madsen et al. 2022).

Apart from melanism, deviations from the zig-zag pattern morphology are considered rare for the European adder, and a study conducted in south-east England between 2005 and 2020 observed only three adders to have a straight dorsal line (Hills 2020). Although this striped phenotype is rare for *V. berus*, it is common in *V. seoanei* (Lataste, 1879; 4 Mya divergence from *V. berus*; Martínez-Freiría et al. 2017; Šmíd and Tolley 2019) and in *Vipera aspis* (Linnaeus, 1758; 15 Mya divergence from *V. berus*; Ducrest et al. 2014; Šmíd and Tolley 2019). Similar to Swedish *V. berus*, melanism in *V. aspis* is quite common (Ducrest et al. 2014). In an analysis of the *POMC* and the melanocortin-1-receptor (*MC1R*) genes, a single-nucleotide polymorphism (SNP) within the 3'-untranslated region (3'UTR) of the *POMC* gene was discovered to be correlated with the dorsal color polymorphism in *V. aspis* (Ducrest et al. 2014). Since *POMC* is a precursor to several hormones, and consequently responsible for several vital functions in the body, it is highly conserved across multicellular animals (Cai and Hruby 2016; Dores et al. 2016).

On the island of Gossa off the west coast of Central Norway (Møre og Romsdal county), a population of *V. berus* was recorded displaying the striped phenotype (Fig. 1b) in ~5% of individuals during preliminary field observations in 2017 (2 striped of ca. 42 live individuals) and 2020 (5 striped of 92 live individuals). Gossa, also called Gossen, is ~45 km² in area and covered in boreal

heathland, moorland, and marsh (Thorsnæs 2024). Several locations have rocky outcrops with crevices that go deep into the earth (K. H. Ophus, personal report). The landscape is speckled with numerous ponds and streams providing habitat for frogs (*Rana temporaria*), which constitute the main diet of the Gossa adder population (K. H. Ophus, personal report).

Given the limited research on reptile color polymorphism in regards to pigmentation genes outside of melanocytes, this study aimed to determine the genomic basis of the rare striped phenotype that is common in the Gossa population. Our work uses whole-genome sequencing and genome scanning of 76 individuals from the Gossa population, including 19 with a striped phenotype, to further our understanding of pattern development and polymorphism in wild snake populations.

Results

Bioinformatic Processing and Analysis of Population Structure

The 76 combined samples (both field collected and from a natural history collection) were sequenced to a mean genomic depth of 3.4× (minimum: 0.2×, maximum: 12.1×) after mapping against the *V. berus* reference nuclear genome and removing reads with low-quality alignments (Table S1). After processing the dataset with ANGSD and removing repeat regions, we were left with 1,796,192 variant positions used in the subsequent genome-wide association study (GWAS) analyses. Further filtering, by pruning the variants in high linkage disequilibrium (LD), reduced the dataset to 373,558 variants used for inferring the overall population structure via a principal component analysis (PCA). In the PCA, the top four eigenvectors were displayed to visualize the population structure within Gossa (Fig. 2). Notably, the first four principal components explain 18% of the variation, and the structuring of genetic variation is not associated with the phenotypes of interest. The kinship analysis revealed that 19 of the individuals were related with $r_{xy} > 0.3$, including the expected close kinship between mothers and their offspring and siblings within the same brood (Table S2). Selecting a single representative of each kin group resulted in a set of 57 unrelated individuals used in a subset of the subsequent analyses.

F_{ST} Analyses

The pairwise F_{ST} analysis based on the site frequency spectra of the two color morph cohorts including all 76 sequenced individuals resulted in a global weighted F_{ST} value of 0.0020. The sliding window analysis considered 123,909 10 kb windows, 124 of which were

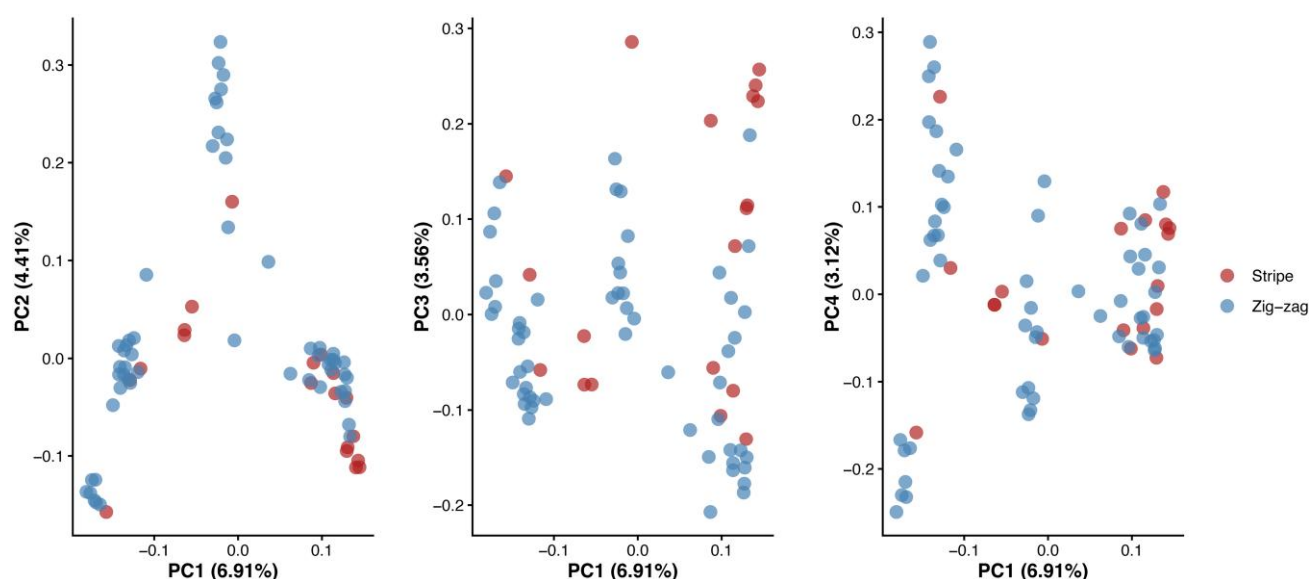


Fig. 2. Population genetic structure of *V. berus* individuals ($n = 76$) on the island of Gossa. Displayed are the first four principal components, which account for ~18% of genetic variation within the Gossa population.

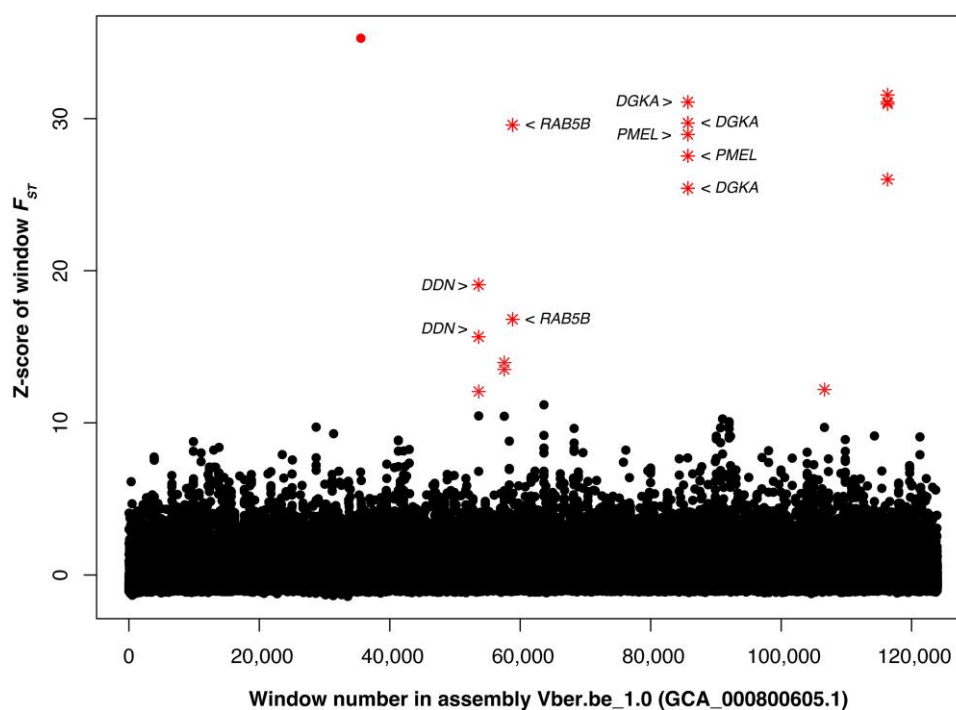


Fig. 3. Genomic windows of extreme differentiation between zig-zag and striped cohorts in a comparison of all 76 sequenced individuals from the Gossa population. The x-axis represents the numbers assigned to nonoverlapping, 10 kb sliding genomic windows in the fragmented genome assembly, and thus do not represent true genomic order. Z-scores are calculated from mean pairwise F_{ST} values summarized in those windows. Outliers with Z-scores >12 are colored red. Labels indicate full-length *V. berus* genes overlapping the outlier windows. Outliers aligned to chromosome 2 in the haplotype-resolved genome assembly are plotted as asterisks.

outliers in the 99.9% quantile (Fig. 3). To identify genomic regions with exceptionally strong genetic differentiation, the outliers were further filtered to include only

those with Z-scores >12, resulting in a list of 19 windows (Table 1). Liftover of these windows to a newly available, improved, haplotype-resolved *V. berus* genome

Table 1 Information about the genomic regions and genes identified in the GWAS and F_{ST} -based genomic scanning analyses

Assembly scaffold	Window range (bp)	F_{ST} outlier scan				GWAS		Gene overlap	Genomic position in improved assembly (scaffold: begin-end)
		All individuals		Unrelated individuals		All individuals	Unrelated individuals		
		Z-score	Mean F_{ST}	Z-score	Mean F_{ST}				
						N	N		
KN620850.1	190,000 to 200,000	-	-	12.50	0.16	-	-		OZ077558.1: 39,865,443–39,885,488 (chr2)
KN620850.1	200,000 to 210,000	-	-	12.52	0.16	-	-		Not resolved
KN623572.1	10,000 to 20,000	35.29	0.59	52.63	0.67	-	-		OZ077561.1: 11,316,985–11,324,071 (chr4)
KN630581.1	110,000 to 120,000	-	-	-	-	-	2		OZ077558.1: 152,896,653–152,950,436 (chr2)
KN631330.1	10,000 to 20,000	19.08	0.32	27.29	0.35	-	-		DDN
KN631330.1	20,000 to 30,000	15.62	0.26	21.32	0.27	-	-		DDN
KN631330.1	30,000 to 40,000	16.22	0.27	24.57	0.31	-	-		
KN631330.1	40,000 to 50,000	12.04	0.21	16.79	0.22	-	-		
KN631330.1	50,000 to 60,000	-	-	12.34	0.16	-	-		
KN631330.1	200,000 to 210,000	-	-	12.51	0.16	-	-	MYL6	OZ077558.1: 152,750,414–152,759,077 (chr2)
KN631588.1	10,000 to 20,000	13.55	0.23	15.95	0.21	-	-		OZ077558.1 :153,672,911–153,691,328 (chr2)
KN631588.1	20,000 to 30,000	13.90	0.24	15.71	0.20	-	-		
KN631677.1	1 to 10,000	-	-	-	-	13	-		OZ077558.1: 153,113,448-153,146,268 (chr2)
KN631677.1	10,000 to 20,000	16.81	0.28	25.39	0.32	3	7	RAB5B	
KN631677.1	20,000 to 30,000	29.58	0.49	35.7	0.45	3	4	RAB5B	
KN631682.1	90,000 to 100,000	-	-	15.4	0.20	-	-		OZ077557.1 :106,638,781–106,648,831 (chr1)
KN632029.1	110,000 to 120,000	-	-	12.36	0.16	-	-		OZ077557.1: 247,590,169–247,597,691 (chr1)
KN632429.1	150,000 to 160,000	-	-	-	-	-	2		OZ077562.1: 87,727,339–87,737,339 (chr5)
KN633259.1	90,000 to 100,000	-	-	-	-	-	1		OZ077559.1: 207,517,687–207,521,595 (chr3)

(continued)

Table 1 Continued

Assembly scaffold	Window range (bp)	F_{ST} outlier scan				GWAS		Gene overlap	Genomic position in improved assembly (scaffold: begin-end)
		All individuals		Unrelated individuals		All individuals	Unrelated individuals		
		Z-score	Mean F_{ST}	Z-score	Mean F_{ST}				
KN633923.1	1 to 10,000	-	-	-	-	-	2		OZ077558.1: 153,015,275–153,085,309 (chr2)
KN633923.1	10,000 to 20,000	31.09	0.52	45.41	0.58	-	-	DGKA	
KN633923.1	20,000 to 30,000	29.71	0.50	44.78	0.57	10	4	DGKA	
KN633923.1	30,000 to 40,000	25.42	0.43	39.15	0.5	3	3	DGKA	
KN633923.1	40,000 to 50,000					1	1		
KN633923.1	50,000 to 60,000	28.96	0.48	45.09	0.57	10	4	PMEL	
KN633923.1	60,000 to 70,000	27.55	0.46	37.82	0.48	13	4	PMEL	
KN636360.1	30,000 to 40,000	12.17	0.21	16.45	0.21	-	-		OZ077558.1: 143,676,703–143,686,820 (chr2)
KN636872.1	10,000 to 20,000	-	-	-	-	1	-		OZ077558.1: 188,348,181–188,358,181 (chr2)
KN637755.1	200,000 to 210,000	-	-	-	-	1	-		OZ077558.1: 144,041,451–144,052,658 (chr2)
KN637802.1	10,000 to 20,000	31.09	0.51	46.04	0.58	2	-		OZ077558.1: 152,969,566–153,006,056 (chr2)
KN637802.1	20,000 to 30,000	26.00	0.43	38.55	0.49	2	5		
KN637802.1	30,000 to 40,000	31.56	0.53	46.91	0.59	10	12		
KN637802.1	40,000 to 50,000	30.97	0.52	43.03	0.55	4	3		
KN637802.1	50,000 to 60,000	-	-	-	-	-	1		
KN638099.1	230,000 to 240,000	-	-	-	-	-	2		OZ077559.1: 205,645,872–205,653,643 (chr3)
KN639025.1	10,000 to 20,000	-	-	-	-	-	1		OZ077557.1: 130,513,389–130,523,066 (chr1)

N indicates the number of highly significant ($P < 5e-8$) SNP associations identified in the GWAS within each genomic window. The “-” character indicates when a genomic window did not pass filtering thresholds for a particular analysis. A blank in the rightmost column indicates there was no overlapping gene annotated within the window. Final column indicates the position of the outlier window(s) in the recently released, improved assembly rVip8er3.hap1.1

Table 2 Information about missense variants within the predicted *PMEL* and *DDN* genes segregating in the Gossa population

Gene ID	Scaffold	Scaffold position (bp)	Casual nucleotide variant	Missense amino acid variant	Freq. in striped cohort (%)	Freq. in zig-zag cohort (%)	Amino acid genotype freqs. in striped cohort (%)			Amino acid genotype freqs. in zig-zag cohort (%)		
							A/A	A/R	R/R	A/A	A/R	R/R
<i>PMEL</i>	KN633923.1	55,190	G > A	Gly206Ser	0	59.5	0	0	100	17.4	82.6	0.0
<i>DDN</i>	KN631330.1	31,193	A > G	Ser100Gly	100	50.0	100	0	0	16.0	68.0	16.0
<i>DDN</i>	KN631330.1	31,058	A > T	Arg145Trp	100	45.5	100	0	0	13.6	63.7	22.7
<i>DDN</i>	KN631330.1	29,939	A > C	Thr518Pro	100	56.8	100	0	0	31.8	50.0	18.2
<i>DDN</i>	KN631330.1	29,545	G > C	Ser649Thr	100	55.0	100	0	0	20.0	70.0	10.0

No missense variants were detected within the genes *DGKA*, *RAB5B*, or *MYL6*. R indicates the reference allele, and A indicates the alternate allele.

assembly (rVipBer3.hap1.1; Owens et al. 2025) revealed that 90% (17 of 19) of these regions align to a single 250 kb locus of chromosome 2 (Table 1). A second F_{ST} analysis of the two color morph cohorts using genomes from the 57 unrelated individuals resulted in largely similar results, but with a smaller global weighted F_{ST} value of 0.0017, yielding 25 outlier windows (Table 1). These align to a 1 Mb region of chromosome 2 that completely encompasses the 250 kb region identified in the analysis with all the individuals, as well as smaller, 7 to 20 kb regions on chromosomes 1 and 2. There was considerable overlap between the outlier windows identified in the two F_{ST} analyses (Table 1).

GWAS Analysis

The GWAS analysis including related individuals yielded an overall low inflation factor of 1.092 (Fig. S1) and revealed 76 SNPs with highly significant ($P < 5e-8$) associated with the striped phenotype (Fig. S2). The most significant of these mostly fall within the same genomic windows revealed as extreme outliers in the F_{ST} analysis (Table S3). Another GWAS considering only unrelated individuals also yielded a low inflation factor of 1.087 (Fig. S3) and revealed 58 SNPs with highly significant ($P < 5e-8$) associations with the striped phenotype (Table S4, Fig. S4).

Ab Initio Genome Annotation

After an initial run of *RepeatMasker* and *RepeatModeler*, a second run of *RepeatMasker* and *RepeatModeler* was performed using the generated model as a custom library input. This second run resulted in masking 491,372,682 bp (32.1%) of the genome assembly. This custom repeat library, as well as the genome assembly sequence, a *Boa constrictor* transcriptome, protein homology evidence from Squamata, and several run iterations using *MAKER2*, *SNAP*, *AUGUSTUS*, and *BUSCO*, resulted in a final *ab initio* annotation containing 7,948 gene models. Owing to the highly fragmented

nature of the genome assembly, the *BUSCO* analysis of 3,354 total *BUSCO* groups revealed 930 (27.7%) complete *BUSCO*s, 913 complete and single-copy *BUSCO*s (27.2%), 17 complete and duplicated *BUSCO*s (0.5%), 521 fragmented *BUSCO*s (15.5%), and 1,903 missing *BUSCO*s (56.8%).

Homology-Based Gene Prediction and Annotation in Outlier Windows

The homology-based prediction of *V. berus* genes in identified highly differentiated ($Z > 12$) outlier genomic windows produced the full-length position, sequence, and exon structure of *PMEL* (premelanosome: 713 aa, 12 exons), *DGKA* (diacylglycerol kinase alpha: 745 aa, 23 exons), *RAB5B* (Ras-related protein Rab-5B: 214 aa, 5 exons), *DDN* (dendrin: 659 aa, 2 exons), *MYL6* (myosin light polypeptide 6: 151 aa, 5 exons). Highly differentiated outlier windows from both F_{ST} analyses (all sequenced individuals or only unrelated individuals) and containing highly significant SNPs from both GWAS analyses overlapped the genes *RAB5B*, *PMEL*, and *DGKA*. The gene *DDN* was contained in only those outlier windows from the F_{ST} analysis with all sequenced individuals, and *MYL6* was contained in only those outlier windows from the F_{ST} analysis with only unrelated individuals (Table 1). Notably, all four of these genes were determined to fall within the same ~1 Mb locus of chromosome 2. Together with the clustering of highly differentiated windows, this indicates that the genomic signal underlying the striped phenotype is concentrated within a single locus, rather than distributed across multiple independent regions. The remaining highly differentiated ($Z > 12$) outlier genomic windows contained no annotated genes.

Variant Effect Analysis and Inheritance Patterns

The analysis of gene variant effects within the predicted *PMEL*, *DGKA*, *RAB5B*, *MYL6*, and *DDN* genes revealed

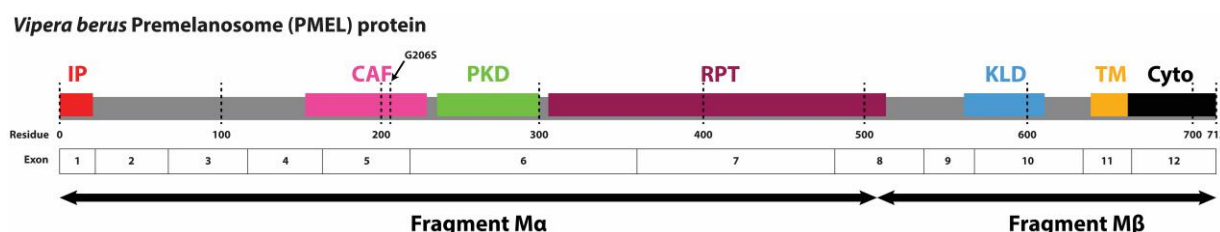


Fig. 4. Premelanosome protein domains predicted from the translated coding regions of the *V. berus* *PMEL* gene. The location of the Gly206Ser missense variant is indicated by the arrow. $M\alpha$ and $M\beta$ indicate the two fragments formed by proteolytic cleavage of the full-length *PMEL* protein (Berson et al. 2003). IP, signal internal propetide targeting the nascent *PMEL* protein to the endoplasmic reticulum; CAF, core amyloid fragment that drives fibril formation in developing melanosomes; PKD, polycystic kidney disease domain; RPT, repeat domain; CS, proprotein convertase cleavage site responsible for separating the luminal $M\alpha$ fragment from the membrane-associated $M\beta$ fragment; KLD, Kringle-like domain; TM, single-pass transmembrane helix anchoring *PMEL* in the melanosomal membrane; Cyto, cytoplasmic domain involved in intracellular trafficking.

that the Gossa population of *V. berus* carries no variants causing nonsynonymous substitutions (missense variants) within predicted exons of *DGKA*, *RAB5B*, or *MYL6*. However, one missense variant (Gly206Ser) was detected in *PMEL* exon 5, which is within the core amyloid fragment domain (Table 2, Fig. 4). In addition, four missense variants were detected within *DDN* exons at residues 100, 145, 518, and 649 (Table 2).

The identified *PMEL* and *DDN* missense alleles have substantially different frequencies in the striped and zig-zag cohorts (Table 2). For the 39 individuals with sequencing depth sufficient to assess genotype at the relevant site, all 16 striped individuals were homozygous for the reference Gly206Ser missense variant. Of the 23 zig-zag individuals, 19 were heterozygous for the missense variant, while 4 were homozygous for the alternate allele.

Discussion

Reptile pattern polymorphism and the mechanisms driving pattern variation are under-studied topics, as the majority of pigmentation research in this group is limited to genes specific to melanocytes (Rosenblum et al. 2004; Corso et al. 2012; Cox et al. 2013; Kuriyama et al. 2020). With the pattern variation of adders in the Gossa population being so rare outside of Gossa, determining the genomic basis is significant, not only for the Gossa population, but also for all reptiles that exhibit polymorphism in color and pattern. Unlike most previous studies that targeted known pigmentation genes, we took an unbiased approach by scanning the whole genome for associations in a population that displays no overall genetic structure associated with pigmentation phenotype, finding a single genomic region with significant genetic differentiation containing several genes that are strong candidates for playing roles in the Gossa stripe morph phenotype. Our discovery that *PMEL* likely plays a role in *V. berus* pattern polymorphism

broadens our understanding of the genetic components at the basis of these polymorphisms in snakes.

Most noteworthy is the identification of *PMEL* in our genome scanning analyses. This gene encodes a type I transmembrane glycoprotein called the "Pre-melanosomal protein," which is localized to the melanosome and plays a role in eumelanin production, although the gene is also expressed in the outermost layer of the retina and the basal ganglia (Hodges et al. 2023) and may also have other functions (Brunberg et al. 2006; Fowler et al. 2006; Chrystal et al. 2021). Because an aspect of pattern formation depends on the ability of pigment to bind to vesicles in the melanocytes and chromatophores, *PMEL* plays a vital role in the development of melanosomes, and therefore color pattern morphology, by creating tight parallel sheets of fibrils for melanin to synthesize and adhere to in stages III and IV of melanosome biogenesis (Watt et al. 2013; Bissig et al. 2016). These tight fibrils give the melanosomes their ellipsoidal shape, characteristic of the darkly pigmented eumelanin. Pheomelanin-producing melanocytes (responsible for red and yellow pigments in skin) lack this ellipsoidal shape and are instead spherical, as was observed in *PMEL* knock-out mice (Hellström et al. 2011).

PMEL is involved in numerous pigmentation disorders in a variety of species such as chickens, horses, cattle, dogs, mice, and zebrafish (Drögemüller et al. 2010; Watt et al. 2011; Claerhout et al. 2018), caused by both loss-of-function and gain-of-function mutations (Chrystal et al. 2021). Our association of the *V. berus* stripe morph phenotype with a single amino acid substitution at the border between the *PMEL* transmembrane region and the cytoplasmic domain is interesting, since mutations in or near the transmembrane domain have already been identified in pigmentation disorders in both chickens (Dominant White; Karlsson et al. 2009) and horses (Silver; Brunberg et al. 2006). Previous work has suggested that mutations in *PMEL*'s well-conserved transmembrane domain (Brunberg et al.

2006) produce divergent forms of amyloid fibrils that reduce melanocyte pigmentation and melanocytes that are depleted of melanosomes (Brumbaugh 1971) and may even be toxic to the melanocytes (Watt et al. 2011).

PMEL was very recently reported as the gene underlying the Terrazzo morph phenotype in the corn snake (*Pantherophis guttatus*) system (Tzika et al. 2024). In that work, a 3 bp polymorphism in the *PMEL* gene sequence was shown to be responsible for the Terrazzo skin pattern, which is characterized by stripes, in contrast to the blotched pattern of the wild type. This multi-nucleotide polymorphism introduces a premature stop codon that produces a truncated premelanosome protein product lacking the Kringle-like domain, the transmembrane helix, and the cytoplasmic domain at carboxyl-terminal end, which ultimately results in the protein's reduced expression. That work also used gene-editing and single-cell transcriptomics to provide strong evidence that *PMEL* is involved in both chromatophore patterning during corn snake embryonic development and melanophore melanogenesis (Tzika et al. 2024).

We determined that the *V. berus* *PMEL* protein product is 713 aa in length, as in the *P. guttatus* genome. We hypothesize that underlying the *V. berus* striped phenotype is a missense variant within the 713-aa *PMEL* protein. This Gly206Ser amino acid substitution falls specifically within *PMEL*'s core amyloid-forming (CAF) domain. Such amyloids play a critical role in melanosomes by providing a scaffold for melanin deposition, which is essential for pigmentation (Yanagisawa et al. 2025). A previous study of the human *PMEL* protein identified a different A Gly>Ser variant, 27 residues downstream and also in the CAF domain, which was found to alter amyloid formation by changing the protein's β -sheet organization and thereby enhancing polymerization capacity (Yanagisawa et al. 2025).

Beyond *PMEL*, several additional genes (*DGKA*, *RAB5B*, *DDN*, and *MYL6*) were identified within highly differentiated outlier windows. However, our results strongly suggest that these genes do not represent independent loci underlying the striped phenotype. Instead, leftover of outlier windows to the improved, haplotype-resolved genome assembly revealed that the vast majority of differentiated regions, including all candidate genes, fall within a single ~250 kb to 1 Mb locus on chromosome 2. This physical clustering indicates that the observed signals likely reflect a single underlying genomic region rather than multiple independent genomic targets. In this context, the most parsimonious explanation is that the missense variant identified in *PMEL* represents the primary causal mutation underlying the striped phenotype, while the signals observed in *DGKA*, *RAB5B*, *DDN*, and *MYL6* are due to physical linkage and genetic hitchhiking

within the same genomic region. Such clustering of functionally related genes within a genomic region is not unexpected, particularly for genes involved in pigmentation or cellular processes associated with melanogenesis.

DGKA encodes an enzyme that phosphorylates diacylglycerol to phosphatidic acid and is involved in lipid-mediated signaling pathways that can influence melanogenesis (Kawaguchi et al. 2012; Sakane et al. 2021). *RAB5B* is an early endosomal Rab GTPase and regulator of endocytic trafficking, its silencing impairs melanosome uptake by keratinocytes (Correia et al. 2018). *MYL6* encodes a component of the myosin motor complex involved in muscle and nonmuscle cytoskeletal function. *MYL6* is expressed in smooth muscle and various nonmuscle tissues, where it contributes to actin-based contractility and cellular motion (Heissler and Sellers 2014). Similarly, *DDN* encodes a cytoskeletal/scaffold protein predominantly expressed in neurons and in the kidney (Herb et al. 1997). Although actin–myosin machinery is important for pigment cell processes (Ripoll et al. 2018), we are not aware of any experimental evidence specifically linking *MYL6* or *DDN* to melanocyte function, pigmentation processes, or skin patterning in snakes or other vertebrates.

Although these additional genes have functions that could plausibly relate to pigmentation biology or chromatophore dynamics, we find no direct evidence from our data supporting a causal role for them in the striped phenotype. However, no missense variants were detected in *DGKA*, *RAB5B*, or *MYL6*, and although several missense variants were identified in *DDN*, there is currently no clear evidence linking this gene to pigmentation or pattern formation. Taken together, these observations support a model in which these genes represent linked passengers within the same genomic region rather than independent contributors to phenotype determination.

We cannot exclude the possibility that regulatory variants within this locus influence the expression of multiple genes; however, the convergence of F_{ST} outliers, GWAS signals, and a clear functional candidate mutation within *PMEL* strongly supports a model in which a single major-effect locus underlies the striped phenotype. Although *DGKA* occurs nearby *PMEL* in the corn snake genome, the amino acid variants in the corn snake Terrazzo morph individuals were not predicted to produce changes in the domain structure of the *DGKA* protein product (Tzika et al. 2024). Similarly, the *V. berus* genomes sequenced in this study contained no missense variants in the *DGKA* gene, nor in the *MYL6*, or *RAB5B* genes. Still there may be noncoding variants outside the gene regions that affect the expression of these genes and thus play some meaningful role in the striped phenotype. Future studies incorporating transcriptomic data and improved genome annotation will be valuable

for disentangling potential regulatory interactions within this region and confirming the extent to which linked genes contribute, if at all, to phenotypic variation.

In an exploratory analysis of color polymorphism in the closely related asp viper (*V. aspis*), two genes (*POMC* and *MC1R*) were screened to test for genotype–phenotype correlation (Ducrest et al. 2014). Both genes are related to melanin-specific pigmentation; however, *POMC* also contributes to several other functions and hormones (Cai and Hruby 2016). Although there is evidence that the *POMC* gene is involved in pattern variation in *V. aspis* (Ducrest et al. 2014), unfortunately the location/sequence for the *POMC* gene could not be recovered from the genome annotation of *V. berus* and, therefore, could not be tested for associations with the striped phenotype in the Gossa population. The annotation used in this study resulted in only 7,948 gene models, which falls well short of the estimated 21,194 predicted protein-coding genes found in *Deinagkistrodon acutus* (Günther, 1888), another member of the family Viperidae (Yin et al. 2016). Future work is needed to determine whether *POMC* also plays a role in color polymorphism in *V. berus*.

In many snake species, longitudinal stripes assist in a rapid-escape strategy by disrupting the predator's ability to maintain a focal point on the body when the snake escapes (Jackson et al. 1976; Madsen et al. 2022). Though the striped phenotype is perhaps more conspicuous in closed habitats, it is more likely to have a benefit in open and wetland habitats (Allen et al. 2013), such as are common on Gossa. This is supported by striped individuals of the harmless viperine snake (*Natrix maura*, Linnaeus, 1758) having been shown to flee from approaching predators at further distances than zig-zag individuals (Santos et al. 2018). In that same study of *N. maura*, striped individuals were almost exclusively found in more open habitats, while the zig-zag counterparts were found in all habitats (Santos et al. 2018). Thus, it is possible that the *V. berus* population on Gossa face similar selective pressure since Gossa is largely covered in heathland, moorland and marsh. Therefore, it would be interesting for future studies to search for further evidence of selective sweeps in the aforementioned genomic regions, especially those containing *PMEL*, *DGKA*, and *RAB5B*.

However, a recent study (Madsen et al. 2022) suggested that when zig-zag adders are moving, the pattern creates a flicker-fusion effect (Shine and Madsen 1994). This effect makes it harder for a visual hunting predator to focus on a reference point or track the moving adder because the dorsal zig-zag blurs into a dorsal stripe that appears to stay stationary while the snake moves away. From this view, the striped pattern on some of Gossa adders most likely provides just as effective an antipredator

effect as does the zig-zag morph. It should be noted that the geographical origin of the striped morph of *V. berus* on the island is not known. There exists a photo of an approximately striped adder specimen from the mainland ~2 km away, and an observation has been made of a striped adder on a neighboring island further away (D. Dolmen, personal report).

Although closely related snakes have evolved the striped and melanistic phenotypes in addition to the zig-zag phenotype, they appear to differ in the specific genetic mechanisms used to achieve this variation. In this study, we discovered a single, moderately sized, gene-containing genomic region with significant association with the striped phenotype on Gossa. The ability of *V. berus* to adapt color patterns to local environments is highly beneficial and as a result, the species has evolved phenotypes similar to other species within *Vipera*. We provide the second report of *PMEL* as a molecular control of pattern variation in snakes, here determining the dorsal pattern of the unique stripe morphs observed at high frequency in Gossa's *V. berus* population. Further genomic resources for this organism will be highly beneficial for further contextualizing our findings and for exploring the extreme outlier windows we identify with an improved genome annotation based on mRNA evidence from the taxon of interest.

Materials and Methods

Sample Acquisition

During the field seasons of 2017 to 2021, *V. berus* snakes were captured on the island of Gossa, Norway under permits 2017/8183, 2,638 from the Norwegian Environment Agency, granted to KHO. Ideal locations on the island were targeted, which included south-facing rocky outcrops, where *V. berus* naturally hibernates. After capture, each snake was photographed (Figs. S5 to S16), individually identified using their head scalation pattern, marked with water-based, nontoxic marker to avoid recaptures, then released. When possible, the individual's sex and capture location were recorded (Table S1). DNA samples were collected using buccal swabs (Boca Scientific) according to the manufacturer's protocol, which were then placed into a collection tube with a silica capsule. Additionally, ventral scale clippings were collected from each individual and stored in 95% ethanol. Further skin tissue samples were taken from four individuals collected in 2016 to 2017 and stored in ethanol in the vertebrate collections at the NTNU University Museum (Hårsaker 2025; Table S1). To further explore the basis of the striped morph, during the 2017 fieldwork, two pregnant striped adders (F1S and F2S) were collected and maintained under observation and care in captivity (terraria) from August 2017 to April 2018. A total of 14 neonates were born in captivity

in August 2017 (7 in each brood). Within the brood of individual F1S, 5 out of 7 (71%) neonates were striped (Fig. S11). In contrast, within the brood of individual F2S all 7 individuals were zig-zag morphs (Fig. S12). Shed skin was collected from each of the neonates and then subjected to DNA extraction and genomic sequencing. A total of 76 snakes were sequenced, 57 having the zig-zag phenotype (36 males and 21 females) and 19 having the striped phenotype (10 males and 9 females; Table S1).

DNA Extraction, Library Preparation, and Sequencing

Genomic DNA was extracted from buccal swabs, scale clippings, and museum skin samples with the Qiagen DNeasy Blood and Tissue extraction kit following the manufacturer's protocol. The extracted DNA was sheared using the Covaris ME220 or Bioruptor Pico sonication instruments with a target fragmentation length of 300 bp. Double-stranded DNA genomic libraries were constructed using blunt-end adapters (Carøe et al. 2018). High-throughput sequencing was performed initially on the Illumina MiniSeq platform in 75 bp paired-end format, then at the Norwegian National Sequencing Centre on the Illumina HiSeq× platform in 150-bp paired-end format, and at the Novogene Europe sequencing facility on the Illumina NovaSeq 6000 platform in 150 bp paired-end format.

Ab initio Genome Annotation and Repeat Masking

Because no gene annotation is available for the *V. berus* reference genome (Vber.be_1.0), we used an *ab initio* approach to obtain an annotation of the masked genome using *MAKER* v3.01.03, *SNAP* and *Augustus* (Korf 2004; Stanke et al. 2006; Cantarel et al. 2008). Prior to the *ab initio* genome annotation, repeats and low-complexity regions were masked from the genome assembly using *RepeatMasker* v4.1.2-p1 (Smit et al. 2013–2015) to apply the taxon-specific “common viper” subset of the curated Dfam/Repbase repeat database. About 3.2% of the genome assembly (49,031,054 bp) was masked in this initial process. Specific repetitive elements remaining in the masked genome were predicted de novo using *RepeatModeler* v2.0.2a (Flynn et al. 2020). Subsequently, a second round of *RepeatMasker* was performed on the previously masked genome, using the model generated from *RepeatModeler* as a custom library input. This custom repeat library and the annotated repeat regions of the *V. berus* genome assembly were extracted for later use.

Genome annotation was performed using the genome annotation pipeline *MAKER2* v2.31.9 (Holt and Yandell 2011) in an iterative process. For the first round of annotation, the genome assembly file, the custom repeat library from the previously described repeat masking

process, a *B. constrictor* transcriptome (Card et al. 2019) used as alternative expressed-sequence tag (EST) evidence, and protein homology evidence from 2,689 unique experimentally verified protein sequences from Squamata (lizards and snakes) [TaxID:8509] were used as the input files to the annotation pipeline. The custom repeat library was used to soft mask the repetitive regions. Additional regions with low complexity were soft masked using *RepeatMasker*. Gene predictors *SNAP* v2013-11-29 (Korf 2004) and *AUGUSTUS* v3.3.1 (Stanke et al. 2006) were trained by performing iterative runs of *MAKER2* as recommended by the authors (Cantarel et al. 2008). As the first round of annotation was based on the alignment of the EST evidence to the genome, the *est2genome* option in the *MAKER2* control file was set to allow *MAKER2* to infer gene models directly from the EST evidence. The *protein2genome* option was also enabled to allow *MAKER2* to infer predictions from protein homology. After the completion of the first round of annotations, gene models were retained and used to train *SNAP* to obtain a *SNAP* HMM file. We then trained *AUGUSTUS* using *BUSCO* v5.1.3 (Manni et al. 2021). First, training sequences were identified using the gene models predicted by *MAKER2* from the first run by excising regions with mRNA annotations and 1,000 bp on either side. These were used to run *BUSCO* using the “tetrapoda_odb10” set of conserved genes. After training both *SNAP* and *Augustus*, *MAKER2* was run again, with *SNAP* hmm and *Augustus* files. A total of three rounds of *BUSCO*/*Augustus* retrainings and four *MAKER2* iterations were performed.

Bioinformatic Data Processing, Read Mapping, and Variant Discovery

The pipeline *Paleomix* v1.3.6 was selected for bioinformatic processing of the raw genomic sequencing data due to its efficiency in handling datasets that originate from a combination of modern and degraded samples (Schubert et al. 2014). Initially, *AdapterRemoval* v2.3.2 (Schubert et al. 2016) was used to trim adapters, as low-quality bases, and ambiguous bases from the 3' end of sequence reads, as well as to merge overlapping read pairs. The trimmed sequences were then mapped against an available highly fragmented *V. berus* reference genome assembly (Vber.be_1.0; NCBI accession code GCA_000800605.1) using the *Burrows-Wheeler Aligner* (BWA) v0.7.16a (Li and Durbin 2010) with the *backtrack* algorithm, allowing for a minimum map quality of 30. Polymerase chain reaction (PCR) duplicates were removed using *Picard Tools* v2.9.1 (Broad Institute 2020). *Genome Analysis Tool Kit* (GATK) v3.7 (McKenna et al. 2010) was used to perform local sequence alignment around indels. Mean sequencing

depth was assessed at all sites in the reference genome using *samtools* v1.16.1 (Li et al. 2009).

The software package *ANGSD* v0.934 (Korneliussen et al. 2014) was used to parse the BAM files for variant discovery and genotype likelihood (GL) estimation after filtering the mapped genomic sequences, requiring a minimum phred-scaled base quality 20 (*-minQ 20*) and a sequence mapping quality of 25 (*-minMapQ 25*). Although typically related individuals should be removed from population genetic analyses, we took a less conservative approach, electing to include both the neonates and their mothers so as to increase the number of striped individuals in our analysis. PCR duplicates and mapped sequences with multiple best hits were removed (*-remove_bads -uniqueOnly*). *ANGSD* was used to calculate GLs using the algorithm from GATK (McKenna et al. 2010), infer and output the major and minor alleles at each variant site (*-doMajorMinor 1 -doGeno 1*), require a minimum SNP *P*-value of $1e^{-6}$ (*-SNP_pval*), and calculate genotype posterior probability using the allele frequency as a prior (*-doPost 1*). For each sample, a minimum of two high-quality mapped reads was required to estimate the GL at a particular site (*-setMinDepthInd 2*). Only sites that were covered by at least two high-quality mapping sequences in at least half the individuals (*-minInd 38*) and only genotypes with a 95% posterior probability (*-postCutoff 0.95*) were included in downstream analyses. Finally, *ANGSD* was used to generate a plink-compatible output file (*-doPlink*). Previously identified *V. berus* genome regions containing transposable elements and repeats were also removed from all downstream analyses. This process resulted in the GLs at what we hereafter refer to as the “final dataset of filtered SNP sites.”

Principal Components Analysis of Population Structure

PLINK v1.90 (Purcell et al. 2007) was used to calculate pairwise LD within the final dataset of filtered SNP sites with the (*-r2*) command. To allow unrecognized chromosome codes (*-allow-extra-chr*) was included, as well as (*-allow-no-sex*), which keeps the phenotype values of missing-sex samples. A threshold of 0 was used for the analysis (*-ld-window-r2 0*) to visualize the decay of r^2 in LD. The average LD distances were plotted in a 100 kb window (*-ld-window-kb 100*) with a window limit of 100 (*-ld-window 100*) to limit the variant count. From this plot, the threshold for pruning was determined (Fig. S17). *PLINK* was then used to establish an independent pairwise cut-off at $r^2 = 0.4$, using 25 kb windows and a step size of 5 kb. To estimate the covariance matrix and allele frequencies, the variants remaining after pruning were analyzed by the software

PCAngsd v1.03 (Meisner and Albrechtsen 2018, 2019). *PCAngsd* uses the GLs from heterogeneous populations to infer population structure using PCA of low-depth genomic sequencing data. The covariance matrix was then imported into *R* v4.0.5 to calculate the eigenvectors and eigenvalues using the *eigen* function included in the base *R*. A PCA plot was created with *ggplot2* v3.3.5 (Wickham 2016) using the eigenvectors.

Kinship Analysis

To estimate kinship between individuals, *angsd* was run exactly as before except that only sites with a minor allele frequency of at least 5% (*-maf 1 -minmaf 0.05*) were utilized. The software *ngsRelate* v2 (Korneliussen and Moltke 2015) was used to estimate pairwise relatedness r_{xy} (Hedrick and Lacy 2015) among the 76 individuals. For sets of individuals with $r_{xy} > 0.3$, only the individual with the highest mean sequencing depth was selected, resulting in a set of 57 genomes from unrelated individuals, of which 14 were striped and 43 were zig-zag.

Genome Scanning Analysis

An F_{ST} -based genome scanning approach was selected in order to identify genomic regions that are highly differentiated in a comparison of striped versus zig-zag individuals. The analysis was first conducted using the entire set of all sequenced individuals and then a second time using a subset of only unrelated individuals. To estimate the folded site frequency spectrum (SFS) from the striped and unstriped cohorts, first the site allele frequency (saf) likelihoods were calculated using the GLs from the final dataset of filtered SNP sites using *realSFS* in *ANGSD* v0.934 (Korneliussen et al. 2014) with the flag *-doSaf 1*. The *V. berus* reference genome sequence was used in place of the ancestral state (assigned using the flag *-anc*), thus using the reference allele as a proxy to define a consistent orientation of variants in the folded SFS. The GATK model was implemented using the *-GL 2* option and the quality threshold was set using *-minMapQ 25* to ignore reads with a mapping quality score below 25. Further filtering was applied, requiring a minimum quality score of 20 using *-minQ 20*, removing duplicate reads (*-remove_bads 1*) and discarding sites with multiple best hits (*-uniqueOnly 1*). Then, using the striped and unstriped site allele frequencies, the 2D SFS was calculated. The *-fold* option was included to estimate the folded SFS since no ancestral state was used. The F_{ST} index was calculated from the allele frequencies from both striped and unstriped sub-populations as well as the folded 2D SFS using the Hudson estimation method (*-whichFST 1*) in *realSFS*. The *realSFS* F_{ST} statistics were

then used to calculate weighted F_{ST} values for each genomic window. Using the *realSFS stat2* function, F_{ST} was calculated in nonoverlapping sliding windows of length 10 kb (*-win 10000* and *-step 10000*). The results from the F_{ST} analysis were refined by correcting negative F_{ST} values to zero and removing windows with fewer than 100 variants. To identify outliers and potential genes of interest, Z-scores were calculated and ranked using the quantile function in base R. The extreme outlier windows (Z-scores >12) were extracted from the *V. berus* reference genome. The genome annotation was used to identify genes positioned within outlier regions.

GWAS Analysis

To complement the F_{ST} -based genome scanning, a GWAS analysis approach was also implemented, again comparing the striped cohort versus the zig-zag cohort within the set of all sequenced individuals. For this, GLs from the final dataset of filtered SNP sites were utilized, including both LD-pruned and unpruned SNP data. Genotype probabilities were imputed from the likelihoods with *Beagle* v.3.3.2 (Browning et al. 2018) and converted to genotype dosages. GWAS was performed with *GEMMA* (Zhou and Stephens 2012). A centered genotype relatedness matrix was included in the *GEMMA* model to account for the high relatedness among individuals in the population. The genomic inflation factor (λ) was calculated and results were visualized with Q-Q and Manhattan plots, generated with the *qqman* package in R (Turner 2018). The GWAS analysis was performed again with the set of unrelated individuals.

Homology-Based Gene Prediction and Annotation in Outlier Windows

As particular genes within outlier windows became of interest, we determined many of these to be only partially annotated by the *BUSCO/Augustus/MAKER2* annotations. Thus *Exonerate* v2.4.0 *protein2genome* algorithm was used to predict the location and full extent of protein-coding genes within the identified outliers windows of interest, using homology evidence from all RefSeq protein sequences from *P. guttatus* (NCBI RefSeq accession GCF_029531705.1). Within these outlier windows, *V. berus* gene annotations from *Exonerate* were further considered if they matched with >85% identity and >90% similarity along the entire length of a *P. guttatus* gene. The exon structures and amino acid sequences of the genes by *Exonerate* were confirmed by extracting the nucleotide sequence of each exon using the *getfasta* algorithm of *BEDTools* (Quinlan and Hall 2010) and then using the *translate* algorithm of *SeqKit* (Shen et al. 2016) to obtain the

amino acid sequence. In each case, the amino acid sequence was then compared to that of the corresponding *P. guttatus* RefSeq protein.

Variant Effect Analysis

To explore the potential effects of variants within the genes of interest identified in the extreme outlier windows, the *HaplotypeCaller* algorithm within the *Genome Analysis Tool Kit* (GATK) v4.3 (McKenna et al. 2010) was used to create sample-specific GVCF files. Then GATK *GenotypeGVCFs* was run using default settings (eg *-sample-ploidy 2*, *-heterozygosity 0.001*, *-indel-heterozygosity 1.25e-4*, *-heterozygosity-stdev 0.01*, *-standard-min-confidence-threshold-for-calling 30.0*, *-max-alternate-alleles 6*, and *-max-genotype-count 1024*) to produce a multisample VCF file containing genomic variants and the genotypes for each sample. The software *snpeff* v5.2c (Cingolani et al. 2012) was used to annotate genomic variants within the outlier regions annotated as genes. Only the three predicted full-length genes (*PMEL*, *DGKA*, *DDN*) predicted to overlap with these windows were considered, and only missense variants were analyzed. *VCFtools* v0.1.17 (Danecek et al. 2011) was used to filter out individual genotypes supported by fewer than 2 sequencing reads or a Phred-scaled genotype quality <15. The InterProScan (Jones et al. 2014) website was used to identify and assign functional domains within the predicted amino acid sequence of the *V. berus* *PMEL* gene.

Liftover to Improved Genome Assembly

To better understand the proximity of the genomic regions identified in this study, scaffolds of the highly fragmented *V. berus* reference genome assembly (*Vber.be_1.0*) were aligned to haplotype 1 of an improved, haplotype-resolved assembly (*rVipBer3.hap1.1*; Owens et al. 2025) that became available late in the life of the present project. This improved assembly lacks a gene annotation at present. To achieve the liftover, *MUMmer* v4.0.1 (Marçais et al. 2018) was used with the options *-maxmatch -l 100 -c 500* to identify all maximal exact matches between the reference and query sequences with a minimum seed length of 100 bp and requiring matches to form alignment clusters of at least 500 bp. The *delta-filter* algorithm *-1* option was used to retain only the best mutually consistent set of alignments such that each region participates in at most one alignment.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

The authors are grateful to Karstein Hårsaker of the NTNU University Museum for facilitating the sampling of the vertebrate collections. Some support for the study was provided by Norwegian Research Council awards 287327 and 325589 to M.D.M. Some computations were performed on resources provided by Sigma2, the National Infrastructure for High-Performance Computing and Data Storage in Norway, under project NS9819K. The authors are grateful to the Norwegian National Sequencing Centre (Oslo) and Novogene UK (Cambridge) for their Illumina sequencing services.

Author Contributions

M.E.A., H.B., R.N., and M.D.M. conceived the study. Funding was provided by NTNU via an Onsager Fellowship to M.D.M. M.D.M. and J.C. supervised the research. K.H.O. conducted the fieldwork, in periods together with N.C. and D.D. K.H.O. also performed registration, photographing, sampling and systematization of the records and material, and collected and kept the pregnant females and their young. N.C. and M.D.M. performed sampling from museum collections. N.C., S.L.F.M., J.M.L., and M.D.M. performed the laboratory work. N.C., M.D.M., J.C., B.P., V.C.B., and J.C.B. performed the analyses. All photographs were provided by K.H.O. N.C. and M.D.M. prepared the figures and wrote the manuscript with contributions from all authors.

Data Availability

All genomic data generated for this study are available on the European Nucleotide Archive under accession code PRJEB84000. The genome annotation files generated in this study and complete amino acid sequences of *V. berus* PMEL, DDN, and DGKA are available on the Dryad digital repository at <https://doi.org/10.5061/dryad.s1rn8pknn>.

Literature Cited

- Alibardi L. Observations on the ultrastructure and distribution of chromatophores in the skin of chelonians. *Acta Zool.* 2013;94:222–232. <https://doi.org/10.1111/j.1463-6395.2011.00546.x>.
- Allen WL, Baddeley R, Scott-Samuel NE, Cuthill IC. The evolution and function of pattern diversity in snakes. *Behav Ecol.* 2013;24:1237–1250. <https://doi.org/10.1093/beheco/art058>.
- Andersson S. Hibernation, habitat and seasonal activity in the adder, *Vipera berus*, north of the Arctic Circle in Sweden. *Amphib-Reptila.* 2003;24:449–457. <https://doi.org/10.1163/156853803322763918>.
- Andrén C, Nilson G. Reproductive success and risk of predation in normal and melanistic colour morphs of the adder, *Vipera berus*. *Biol J Linn Soc Lond.* 1981;15:235–246. <https://doi.org/10.1111/j.1095-8312.1981.tb00761.x>.
- Bauwens D, Claus K. Seasonal variation of mortality, detectability, and body condition in a population of the adder (*Vipera berus*). *Ecol Evol.* 2019;9:5821–5834. <https://doi.org/10.1002/ece3.5166>.
- Berson JF, et al. Proprotein convertase cleavage liberates a fibrillogenic fragment of a resident glycoprotein to initiate melanosome biogenesis. *J Cell Biol.* 2003;161:521–533. <https://doi.org/10.1083/jcb.200302072>.
- Bissig C, Rochin L, van Niel G. PMEL amyloid fibril formation: the bright steps of pigmentation. *Int J Mol Sci.* 2016;17:1438. <https://doi.org/10.3390/ijms17091438>.
- Broad Institute. 2020. Picard tools. Broad Institute, GitHub repository. <http://broadinstitute.github.io/picard/>.
- Brown AR et al. A community-science approach identifies genetic variants associated with three color morphs in ball pythons (*Python regius*). *PLoS One.* 2022;17:e0276376. <https://doi.org/10.1371/journal.pone.0276376>.
- Browning BL, Zhou Y, Browning SR. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet.* 2018;103:338–348. <https://doi.org/10.1016/j.ajhg.2018.07.015>.
- Brumbaugh JA. The ultrastructural effects of the I and S loci upon black-red melanin differentiation in the fowl. *Dev Biol.* 1971;24:392–412. [https://doi.org/10.1016/0012-1606\(71\)90088-1](https://doi.org/10.1016/0012-1606(71)90088-1).
- Brunberg E et al. A missense mutation in PMEL17 is associated with the silver coat color in the horse. *BMC Genet.* 2006;7:46. <https://doi.org/10.1186/1471-2156-7-46>.
- Cai M, Hruby VJ. The melanocortin receptor system: a target for multiple degenerative diseases. *Curr Protein Pept Sci.* 2016;17:488–496. <https://doi.org/10.2174/1389203717666160226145330>.
- Cantarel BL et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 2008;18:188–196. <https://doi.org/10.1101/gr.6743907>.
- Card DC et al. Genomic basis of convergent island phenotypes in *Boa constrictors*. *Genome Biol Evol.* 2019;11:3123–3143. <https://doi.org/10.1093/gbe/evz226>.
- Carøe C et al. Single-tube library preparation for degraded DNA. *Methods Ecol Evol.* 2018;9:410–419. <https://doi.org/10.1111/2041-210X.12871>.
- Chrystal PW et al. Functional domains and evolutionary history of the PMEL and GPNMB family proteins. *Molecules.* 2021;26:3529. <https://doi.org/10.3390/molecules26123529>.
- Cingolani P et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012;6:80–92. <https://doi.org/10.4161/fly.19695>.
- Claerhout H et al. Isolated sulfite oxidase deficiency. *J Inherit Metab Dis.* 2018;41:101–108. <https://doi.org/10.1007/s10545-017-0089-4>.
- Correia MS et al. Melanin transferred to keratinocytes resides in nondegradative endocytic compartments. *J Invest Dermatol.* 2018;138:637–646. <https://doi.org/10.1016/j.jid.2017.09.042>.
- Corso J, Gonçalves GL, de Freitas TRO. Sequence variation in the melanocortin-1 receptor (MC1R) pigmentation gene and its role in the cryptic coloration of two South American sand lizards. *Genet Mol Biol.* 2012;35:81–87. <https://doi.org/10.1590/S1415-47572012005000015>.
- Cox CL, Rabosky ARD, Chippindale PT. Sequence variation in the Mc1r gene for a group of polymorphic snakes. *Gene.* 2013;513:282–286. <https://doi.org/10.1016/j.gene.2012.10.065>.
- Cui S et al. The adder (*Vipera berus*) in Southern Altay mountains: population characteristics, distribution, morphology and

- phylogenetic position. *PeerJ*. 2016;4:e2342. <https://doi.org/10.7717/peerj.2342>.
- Danecek P et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Dores RM, Liang L, Davis P, Thomas AL, Petko B. 60 YEARS OF POMC: melanocortin receptors: evolution of ligand selectivity for melanocortin peptides. *J Mol Endocrinol*. 2016;56: T119–T133. <https://doi.org/10.1530/JME-15-0292>.
- Drögemüller C et al. Identification of the bovine arachnomelia mutation by massively parallel sequencing implicates sulfite oxidase (SUOX) in bone development. *PLoS Genet*. 2010;6: e1001079. <https://doi.org/10.1371/journal.pgen.1001079>.
- Ducrest A-L et al. Pro-opiomelanocortin gene and melanin-based colour polymorphism in a reptile. *Biol J Linn Soc Lond*. 2014;111:160–168. <https://doi.org/10.1111/bij.12182>.
- Flynn JM et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A*. 2020;117: 9451–9457. <https://doi.org/10.1073/pnas.1921046117>.
- Fowler DM et al. Functional amyloid formation within mammalian tissue. *PLoS Biol*. 2006;4:e6. <https://doi.org/10.1371/journal.pbio.0040006>.
- Härsaker K, Dolmen D. 2025. Herpetile collection Norwegian University of Science and Technology (NTNU) University Museum. Occurrence dataset. <https://doi.org/10.15468/UDAFNG>.
- Hedrick PW, Lacy RC. Measuring relatedness between inbred individuals. *J Hered*. 2015;106:20–25. <https://doi.org/10.1093/jhered/esu072>.
- Heissler SM, Sellers JR. Myosin light chains: teaching old dogs new tricks. *Bioarchitecture*. 2014;4:169–188. <https://doi.org/10.1080/19490992.2015.1054092>.
- Hellström AR et al. Inactivation of Pmel alters melanosome shape but has only a subtle effect on visible pigmentation. *PLoS Genet*. 2011;7:e1002285. <https://doi.org/10.1371/journal.pgen.1002285>.
- Herb A et al. Prominent dendritic localization in forebrain neurons of a novel mRNA and its product, dendrin. *Mol Cell Neurosci*. 1997;8:367–374. <https://doi.org/10.1006/mcne.1996.0594>.
- Hills L. Dorsal stripe polymorphism of *Vipera berus* in south-east England. *Herpetol Bull*. 2020;153:26–28. <https://doi.org/10.33256/hb153.2628>.
- Hodges ED et al. Disrupting the repeat domain of premelanosome protein (PMEL) produces dysamyloidosis and dystrophic ocular pigment reflective of pigmentary glaucoma. *Int J Mol Sci*. 2023;24:14423. <https://doi.org/10.3390/ijms241914423>.
- Holt C, Yandell M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*. 2011;12:491. <https://doi.org/10.1186/1471-2105-12-491>.
- Horreo JL, Fitze PS. Postglacial colonization of Northern Europe by reptiles. In: Pontarotti P, editor. *Origin and evolution of biodiversity*. Springer International Publishing; 2018. p. 197–214.
- Jackson JF, Ingram W III, Campbell HW. The dorsal pigmentation pattern of snakes as an antipredator strategy: a multivariate approach. *Am Nat*. 1976;110:1029–1053. <https://doi.org/10.1086/283125>.
- Jones P et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
- Karlsson A-C, Kerje S, Hallböök F, Jensen P. The dominant white mutation in the PMEL17 gene does not cause visual impairment in chickens. *Vet Ophthalmol*. 2009;12:292–298. <https://doi.org/10.1111/j.1463-5224.2009.00714.x>.
- Kawaguchi M, Valencia JC, Namiki T, Suzuki T, Hearing VJ. Diacylglycerol kinase regulates tyrosinase expression and function in human melanocytes. *J Invest Dermatol*. 2012;132: 2791–2799. <https://doi.org/10.1038/jid.2012.261>.
- Korf I. Gene finding in novel genomes. *BMC Bioinformatics*. 2004;5:59. <https://doi.org/10.1186/1471-2105-5-59>.
- Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*. 2014;15: 356. <https://doi.org/10.1186/s12859-014-0356-4>.
- Korneliussen TS, Moltke I. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics*. 2015;31:4009–4011. <https://doi.org/10.1093/bioinformatics/btv509>.
- Kuriyama T, Murakami A, Brandley M, Hasegawa M. Blue, black, and stripes: evolution and development of color production and pattern formation in lizards and snakes. *Front Ecol Evol*. 2020;8:232. <https://doi.org/10.3389/fevo.2020.00232>.
- Li H et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26: 589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
- Lourdais O, Guillon M, Denardo D, Blouin-Demers G. Cold climate specialization: adaptive covariation between metabolic rate and thermoregulation in pregnant vipers. *Physiol Behav*. 2013;119:149–155. <https://doi.org/10.1016/j.physbeh.2013.05.041>.
- Madsen T, Stille B, Ujvari B, Bauwens D, Endler JA. Negative frequency-dependent selection on polymorphic color morphs in adders. *Curr Biol*. 2022;32:3385–3388.e3. <https://doi.org/10.1016/j.cub.2022.05.060>.
- Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38: 4647–4654. <https://doi.org/10.1093/molbev/msab199>.
- Marçais G et al. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol*. 2018;14:e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>.
- Martínez-Freiría F, Pérez i de Lanuza G, Pimenta AA, Pinto T, Santos X. Aposematism and crypsis are not enough to explain dorsal polymorphism in the Iberian adder. *Acta Oecol*. (Montrouge). 2017;85:165–173. <https://doi.org/10.1016/j.actao.2017.11.003>.
- McKenna A et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- McKinnon JS, Pierotti MER. Colour polymorphism and correlated characters: genetic mechanisms and evolution. *Mol Ecol*. 2010;19:5101–5125. <https://doi.org/10.1111/j.1365-294X.2010.04846.x>.
- Meisner J, Albrechtsen A. Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*. 2018;210:719–731. <https://doi.org/10.1534/genetics.118.301336>.
- Meisner J, Albrechtsen A. Testing for Hardy-Weinberg equilibrium in structured populations using genotype or low-depth next generation sequencing data. *Mol Ecol Resour*. 2019;19: 1144–1152. <https://doi.org/10.1111/1755-0998.13019>.
- Murray JD. *Mathematical biology II: spatial models and biomedical applications*. Springer Science & Business Media; 2011.

- Murray JD, Myerscough MR. Pigmentation pattern formation on snakes. *J Theor Biol.* 1991;149:339–360. [https://doi.org/10.1016/S0022-5193\(05\)80310-8](https://doi.org/10.1016/S0022-5193(05)80310-8).
- Owens JB et al. The genome sequence of the common adder, *Vipera berus* (Linnaeus, 1758). *Wellcome Open Res.* 2025;10:11. <https://doi.org/10.12688/wellcomeopenres.23470.2>.
- Purcell S et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81:559–575. <https://doi.org/10.1086/519795>.
- Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
- Ripoll L et al. Myosin VI and branched actin filaments mediate membrane constriction and fission of melanosomal tubule carriers. *J Cell Biol.* 2018;217:2709–2726. <https://doi.org/10.1083/jcb.201709055>.
- Rosenblum EB, Hoekstra HE, Nachman MW. Adaptive reptile color variation and the evolution of the *Mcl1r* gene. *Evolution.* 2004;58:1794–1808. <https://doi.org/10.1554/03-741>.
- Roulin A. The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. *Biol Rev Camb Philos Soc.* 2004;79:815–848. <https://doi.org/10.1017/S1464793104006487>.
- Saenko SV, Teyssier J, van der Marel D, Milinkovitch MC. Precise co-localization of interacting structural and pigmentary elements generates extensive color pattern variation in *Phelsuma* lizards. *BMC Biol.* 2013;11:105. <https://doi.org/10.1186/1741-7007-11-105>.
- Sakane F, Hoshino F, Ebina M, Sakai H, Takahashi D. The roles of diacylglycerol kinase α in cancer cell proliferation and apoptosis. *Cancers (Basel).* 2021;13:5190. <https://doi.org/10.3390/cancers13205190>.
- Santos X et al. Ecological significance of dorsal polymorphism in a Batesian mimic snake. *Curr Zool.* 2018;64:745–753. <https://doi.org/10.1093/cz/zox058>.
- Schubert M et al. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat Protoc.* 2014;9:1056–1082. <https://doi.org/10.1038/nprot.2014.063>.
- Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes.* 2016;9:88. <https://doi.org/10.1186/s13104-016-1900-2>.
- Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One.* 2016;11:e0163962. <https://doi.org/10.1371/journal.pone.0163962>.
- Shine R, Madsen T. Sexual dichromatism in snakes of the genus *Vipera*: a review and a new evolutionary hypothesis. *J Herpetol.* 1994;28:114–117. <https://doi.org/10.2307/1564692>.
- Šmíd J, Tolley KA. Calibrating the tree of vipers under the fossilized birth-death model. *Sci Rep.* 2019;9:5510. <https://doi.org/10.1038/s41598-019-41290-2>.
- Smit AFA, Hubley R, Green P. 2013–2015. RepeatMasker Open-4.0. RepeatMasker Open-4.0. <http://www.repeatmasker.org>.
- Stanke M, Schöffmann O, Morgenstern B, Waack S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics.* 2006;7:62. <https://doi.org/10.1186/1471-2105-7-62>.
- Tang C-Y et al. Genetic mapping and molecular mechanism behind color variation in the Asian vine snake. *Genome Biol.* 2023;24:46. <https://doi.org/10.1186/s13059-023-02887-z>.
- Thorsnæs G. 2024. Gossa. *Store Norske Leksikon*. <https://snl.no/Gossa> (Accessed January 12, 2026).
- Turner SD. Qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J Open Source Softw.* 2018;3:731. <https://doi.org/10.21105/joss.00731>.
- Tzika AC, Ullate-Agote A, Helleboid P-Y, Kummrow M. PMEL is involved in snake colour pattern transition from blotches to stripes. *Nat Commun.* 2024;15:7655. <https://doi.org/10.1038/s41467-024-51927-0>.
- Ullate-Agote A et al. Genome mapping of a LYST mutation in corn snakes indicates that vertebrate chromatophore vesicles are lysosome-related organelles. *Proc Natl Acad Sci U S A.* 2020;117:26307–26317. <https://doi.org/10.1073/pnas.2003724117>.
- Ursenbacher S, Carlsson M, Helfer V, Tegelström H, Fumagalli L. Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred from mitochondrial DNA sequence data. *Mol Ecol.* 2006;15:3425–3437. <https://doi.org/10.1111/j.1365-294X.2006.03031.x>.
- Valkonen J, Niskanen M, Björklund M, Mappes J. Disruption or aposematism? Significance of dorsal zigzag pattern of European vipers. *Evol Ecol.* 2011;25:1047–1063. <https://doi.org/10.1007/s10682-011-9463-0>.
- Watt B et al. Mutations in or near the transmembrane domain alter PMEL amyloid formation from functional to pathogenic. *PLoS Genet.* 2011;7:e1002286. <https://doi.org/10.1371/journal.pgen.1002286>.
- Watt B, van Niel G, Raposo G, Marks MS. PMEL: a pigment cell-specific model for functional amyloid formation. *Pigment Cell Melanoma Res.* 2013;26:300–315. <https://doi.org/10.1111/pcmr.12067>.
- Wickham H. Ggplot2: elegant graphics for data analysis. Springer; 2016.
- Wüster W et al. Do aposematism and Batesian mimicry require bright colours? A test, using European viper markings. *Proc Biol Sci.* 2004;271:2495–2499. <https://doi.org/10.1098/rspb.2004.2894>.
- Yanagisawa H et al. Cryo-EM of wild-type and mutant PMEL amyloid cores reveals structural mechanism of pigment dispersion syndrome. *Nat Commun.* 2025;16:5411. <https://doi.org/10.1038/s41467-025-61233-y>.
- Yin W et al. Evolutionary trajectories of snake genes and genomes revealed by comparative analyses of five-pacer viper. *Nat Commun.* 2016;7:13107. <https://doi.org/10.1038/ncomms13107>.
- Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet.* 2012;44:821–824. <https://doi.org/10.1038/ng.2310>.

Associate editor: Andrea Betancourt