Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE

Divergence in Regulatory Regions and Gene Duplications May Underlie Chronobiological Adaptation in Desert Tortoises

N. Jade Mellor[1] | Timothy H. Webster[2] | Hazel Byrne[2] | Avery S. Williams[1] | Taylor Edwards[3] | Dale F. DeNardo[4] | Melissa A. Wilson[4] | Kenro Kusumi[4] | Greer A. Dolby[1]

[1]Department of Biology, University of Alabama at Birmingham, Birmingham, Alabama, USA | [2]Department of Anthropology, University of Utah, Salt Lake City, Utah, USA | [3]Arizona Molecular Clinical Core, University of Arizona, Tucson, Arizona, USA | [4]School of Life Sciences, Arizona State University, Tempe, Arizona, USA

Correspondence: Greer A. Dolby (gdolby@uab.edu)

Received: 13 March 2024 | **Revised:** 5 November 2024 | **Accepted:** 11 November 2024

Handling Editor: Holly Bik

Funding: This study was supported by the Arizona State University (ASU) Fostering Postdoctoral Research in the Life Sciences award to THW and GAD, ASU's College of Liberal Arts and Sciences, ASU Foundation for Desert Reptile Genomes Fund to KK, Arizona Game and Fish Department Heritage Grant #I17001 to MAW; and NIGMS R35GM124827 to MAW.

Keywords: circadian | circannual | mucin | promoter | reinforcement | speciation

ABSTRACT

Many cellular processes and organismal behaviours are time-dependent, and asynchrony of these phenomena can facilitate speciation through reinforcement mechanisms. The Mojave and Sonoran desert tortoises (*Gopherus agassizii* and *G. morafkai* respectively) reside in adjoining deserts with distinct seasonal rainfall patterns and they exhibit asynchronous winter brumation and reproductive behaviours. We used whole genome sequencing of 21 individuals from the two tortoise species and an outgroup to understand genes potentially underlying these characteristics. Genes within the most diverged 1% of the genome ($F_{ST} \geq 0.63$) with putatively functional variation showed extensive divergence in regulatory elements, particularly promoter regions. Such genes related to UV nucleotide excision repair, mitonuclear and homeostasis functions. Genes mediating chronobiological (cell cycle, circadian and circannual) processes were also among the most highly diverged regions (e.g., *XPA* and *ZFHX3*). Putative promoter variants had significant enrichment of genes related to regulatory machinery (ARC-Mediator complex), suggesting that transcriptional cascades driven by regulatory divergence may underlie the behavioural differences between these species, leading to asynchrony-based prezygotic isolation. Further investigation revealed extensive expansion of respiratory and intestinal mucins (MUC5B and MUC5AC) within *Gopherus*, particularly *G. morafkai*. This expansion could be a xeric-adaptation to water retention and/or contribute to differential *Mycoplasma agassizii* infection rates between the two species, as mucins help clear inhaled dust and bacterial. Overall, results highlight the diverse array of genetic changes underlying divergence, adaptation and reinforcement during speciation.

1 | Introduction

Adaptive divergence underlies the origin and maintenance of habitat and niche differences among speciating lineages (Sobel et al. 2010; Weissing, Edelaar, and Van Doorn 2011). While protein coding changes are often emphasised in analyses across phylogenies (i.e., *dN*/*dS*), changes in transcriptional and translational regulatory sequences, such as enhancers, promoters

and untranslated regions (UTRs), can also cause meaningful phenotypic differences that facilitate ecological differentiation (Franchini et al. 2019; Santos et al. 2007). For example, the deletion of a tissue-specific *Pitx1* enhancer in stickleback fish results in the loss of pelvic spines in some lineages (Chan et al. 2010). And in the Midas cichlid fish radiation, novel microRNAs (miRNAs) evolved between recently diverged species that regulate differential expression of genes associated with macromolecule modification and protein synthesis (Franchini et al. 2019). In such cases, it is the regulation of protein products that underlies phenotypic divergence.

Changes in regulatory sequences generally do not affect the resultant amino acid sequence, but they can affect: (i) the timing and signalling mode of transcription including during development, (ii) the magnitude of transcription (Cong, Liu, and Tanksley 2002), (iii) the tissue or cell specificity of transcription (Chan et al. 2010), (iv) the residence time of the transcript in the cell and amount of translational product (Franchini et al. 2019), (v) localisation of the transcript within the cell (Ahmed et al. 2010) and (vi) which protein isoform is produced (Matlin, Clark, and Smith 2005) (reviewed in Figure S2). Importantly, some ecological adaptations require a suite of coordinated phenotypic changes that may be more efficiently met by regulatory changes that cascade to shape expression of a suite of related genes. For example, time-dependent (i.e., chronobiological) processes that depend on internal pace-making, such as cell cycle regulation, circadian (daily) rhythms and circannual (annual) rhythms, are based on a set of core pace-maker genes that communicate an internal temporal oscillator to a suite of genes and in doing so underlie core biological functions such as sleep and digestion (Tahara and Shibata 2014). It seems reasonable then that regulatory changes may be important to the evolution of such genes. It is often regulatory modifications that coordinate the transcription of pace-maker genes to external cues (called zeitgebers; Page et al. 2020). Therefore, regulatory changes may play an important role in adaptive divergence between speciating lineages who have chronobiological differences.

Beyond functional changes related to the ecologies of diverging species, several other types of genes are expected to diverge early in speciation and play a role in genomic reinforcement (the genetic mechanisms that impede cross-lineage reproduction). Among these are genes related to immune function because the host immune system is in an arms-race with the rapid evolution of local pathogens (Loker 2012). Immune gene copy number (Malmstrøm et al. 2016) and immune gene diversification (Loker 2012) are associated with speciation. Proteins involved in sperm–egg recognition and other reproductive incompatibilities, such as chromosome pairing, are also expected to accumulate during reinforcement to form pre- and post-zygotic isolation mechanisms (e.g., Baty et al. 2024; Coyne and Orr 2004). Prezygotic isolation mechanisms can arise through factors like divergent sexual selection (Boughman 2001), different mating times (Taylor and Friesen 2017) or when sperm cannot fertilise an egg. Postzygotic incompatibilities can cause lower fitness of hybrids in the environment and can occur via misregulation of cell cycle or mitonuclear interactions in hybrids.

Agassiz's and Morafka's desert tortoises (referenced here as the Mojave and Sonoran desert tortoises respectively), *Gopherus*

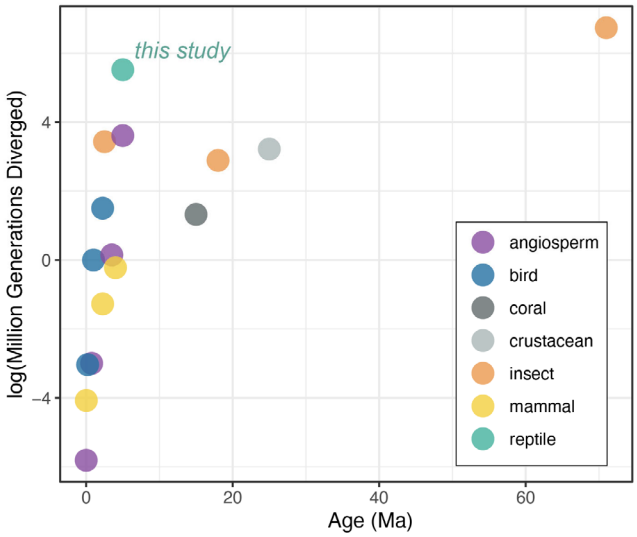

FIGURE 1 | Published studies that used whole genome sequencing to study speciation. Estimated age of divergence and number of generations since divergence from other studied speciation events. Tortoises (this study) are thought to have been diverged for a greater number of generations than other well-studied speciation events. For a list of data sources see Methods.

agassizii and *G. morafkai*, are a good system to study regulatory changes in the context of ecological adaptation. Divergence began about 5 Mya and they are presumed to be in the later stages of speciation (Figure 1, Lamb and Lydeard 1994), offering a good comparison with other more recent speciation systems. Evidence shows they differ in aspects of ecology, life history, reproduction and morphology; many of these differences are time-dependent processes that we hypothesise could be underpinned by regulatory changes. They inhabit the Mojave and Sonoran deserts respectively (Murphy et al. 2011), and likely diverged in parapatry or allopatry with reduced gene flow from formation of the Colorado River as well as adaptation to regional climate disparities formed by the North American Monsoon (Dolby, Dorsey, and Graham 2019; Edwards, Tollis, et al. 2016; Edwards, Vaughn, et al. 2016). The deserts differ in the amount and timing of rainfall, and this results in different vegetation communities which germinate at different times (Germano 1994; Germano et al. 1994). The tortoises differ in growth rate (Curtin, Zug, and Spotila 2009), size (Curtin, Zug, and Spotila 2009), mating and egg laying season(s) (Averill-Murray 2002; Averill-Murray, Christopher, and Henen 2018; Ruby and Niblick 1994; Wallis, Henen, and Nagy 1999) and clutch size (Averill-Murray, Christopher, and Henen 2018; Wallis, Henen, and Nagy 1999). The species also present different susceptibility to the *Mycoplasma agassizii* infection that causes Upper Respiratory Tract Disease (URTD; Dickinson et al. 2005; Gov 2015). URTD has contributed to population declines of the Mojave Desert tortoise, which is Threatened under the US Endangered Species Act (ESA) (US Fish and Wildlife Service 1994; Sandmeier et al. 2009). The two species exhibit a narrow zone of hybridisation at the ecotone of the Mojave and Sonoran deserts in western Arizona and populations of both species are expected to be impacted by the effects of global climate change (Edwards et al. 2015; McLuckie et al. 1999; Smith et al. 2023; Weiss and Overpeck 2005).

In this study we sequenced 10 genomes from the central range of each of the two species along with an outgroup Texas tortoise (*G. berlandieri*) individual. We estimate historical effective population sizes, genome diversity and analyse the genomic regions of differentiation that might underpin the ecological divergence of these lineages. We find strong divergence in regulatory regions (primarily in promoters) of genes related to immune, osmoregulatory and chronobiological processes. We also find a lineage-specific expansion of Mucin 5 genes in the Sonoran Desert tortoise.

2 | Methods and Materials

2.1 | Speciation Literature

To put this speciation event in context of other speciation studies we collected species divergence and generation time estimates from published studies that used genome-wide sequence data to study speciation. Generation time and divergence data were collected for angiosperms (Choi, Purugganan, and Stacy 2020; Guo et al. 2018; Koch and Matschinger 2007; Koornneef and Scheres 2001; Wang et al. 2020), arthropods (Becking et al. 2019; Becking, Gilbert, and Cordaux 2020; Chen et al. 2015; Maebe et al. 2016; Sadd et al. 2015; Sowilem, Kamal, and Khater 2013), birds (Cornetti et al. 2015; Ericson et al. 2019; Joseph et al. 2009; Leroy et al. 2021), corals (Császár et al. 2010; Mao, Economo, and Satoh 2018), mammals (Figueiró et al. 2017; Kumar et al. 2017; K. Li et al. 2020; Shanas et al. 1995) and snails (Chueca, Schell, and Pfenninger 2021). We recorded the generation time (in Mya) and author-reported divergence age for each study (Table S1).

2.2 | Sampling and Sequencing

We downloaded sequence data for 10 individuals representing the geographical centre of the *G. agassizii* range from Scott et al. (2020) (collected from the centre of the Mojave Desert, Figure S1). We sequenced 10 additional *G. morafkai* individuals located 50 km north of Tuscon, AZ from prior collections that followed University of Arizona IACUC protocols (Taylor IACUC 00-084) and state collecting permits (Edwards SCP # SP606722). These samples comprised populations from Picacho Mountain and West Silverbell Mountain also in the geographic middle of the Sonoran Desert tortoise range. An outgroup sample of the Texas tortoise (*G. berlandieri*) was provided by the Amphibian and Reptile Diversity Research Center at the University of Texas at Arlington. Samples were shipped to Yale Center for Genomic Analysis for genomic DNA extractions, Illumina library preparations (150 PE) and sequencing on the Illumina HiSeq 2500.

2.3 | Data Processing

To assess quality of raw sequencing reads we used FastQC v0.11.9 (Andrews 2010) and MultiQC v1.8 (Ewels et al. 2016), after which we trimmed samples for quality and to remove adaptors using BBDuk 'ktrim=r k=21 mink=11 hdist=2 tpe tbo' from BBMap v38.79 (Bushnell 2014). We mapped reads to the gopAga2.0 reference genome (Dolby et al. 2020) using BWA *mem* v0.7.17 (Li and Durbin 2009) and used Samtools v1.10 (Li et al. 2009) to fix mate pairs and sort BAM files, in both cases using default parameters. We marked duplicates with Picard v2.22.8 (Broad Institute n.d.) before conducting INDEL realignment with GATK v3.7 (McKenna et al. 2010).

2.4 | Population Statistics and Demography

For analyses requiring called genotypes, which included principal components analysis (PCA) and historical demographic analysis, we first calculated genome-wide sequencing coverage with Mosdepth (Pedersen and Quinlan 2018) to use in downstream filtering. We jointly called variants per scaffold across samples in the 21-individual dataset using a three-step process in GATK v4.1.7 (McKenna et al. 2010). First, we used HaplotypeCaller to preliminarily call variants in each sample, outputting a gVCF file. We then combined gVCFs using GenomicsDBImport and finally jointly genotyped samples with GenotypeGVCFs. To generate a PCA for the 21-individual dataset, we combined the VCFs for all scaffolds larger than 10 Mbp using BCFtools v1.10.2 *concat* module (Li et al. 2009). We used Plink v1.90b6.12 (Purcell et al. 2007) for linkage disequilibrium-based SNP pruning with the options '--indep-pairwise 50 10 0.1'.

To calculate sliding window statistics along the genome for the 20-individual dataset (no *G. berlandieri* outgroup), we used genotype likelihoods in ANGSD v0.921 which handles low coverage sequencing (Korneliussen, Albrechtsen, and Nielsen 2014). With these likelihoods, we calculated F_{ST} and genetic diversity (θ) in sliding windows using a window size of 50,000 bp, a step size of 10,000 bp and parameters: '-dosaf 1 -gl 1'. To calculate genetic diversity, we ran '-doThetas 1 -doSaf 1 -GL 1 -fold 1' assuming gopAga2.0 as the ancestral state. Finally, we generated a PCA for the 20-individual dataset based on genotype likelihoods using ANGSD and ngsTools (Fumagalli et al. 2013, 2014). We calculated fixed major and minor allele frequences from the likelihoods. Posterior Genotype probabilities were calculated using the frequency as the prior. We only used biallelic sites with a minimum mapping quality of 20, a minimum base quality of 20 and with data in at least 11 individuals (to avoid species-specific missing data).

2.5 | Genomic Diversity

We calculated genome-wide mean genetic diversity (θ) per species using ANGSD ('realSFS -fold 1' and 'thetaStat do_stat -win 50000 -step 10000') in 50-kb windows with a 10-kb step size. From these values we calculated present-day effective population size (N_e) using $\theta = 4N_e\mu$ using a mutation rate of $1.1E^{-9}$ mutations/site/year (Tollis et al. 2018) and a generation time of 20 years to convert the mutation rate to generations.

2.6 | Effective Population Size

For each species, diploid consensus sequences were obtained for one individual from the mapped BAM files for all scaffolds over 10 Mbp in length using BCFtools v1.10.2 mpileup and call modules, and vcfutils vcf2fq module (Li et al. 2009). We only retained sites with a minimum mapping quality, base quality

and root-mean-square mapping quality of 30. We applied the following depth cut-off filters with vcfutils: a minimum depth of four and maximum depth of 21 (approximately five times the genome-wide average) for *G. agassizii* and *G. morafkai*, and a minimum depth of eight and maximum depth of 37 for *G. berlandieri*. We then used filtered consensus sequences in Pairwise Sequential Markovian Coalescent (PSMC) analysis (Li and Durbin 2011) with parameters '-N30 -t6 -r5 -p "4+30*2+4+6+10"' (Nadachowska-Brzyska et al. 2015) to assess changes in effective population size through time. We applied a generation time of 25 years (US Fish and Wildlife Service 1994) and a mutation rate of $2.675e^{-8}$ substitutions per site per generation (Tollis et al. 2017) to scale the PSMC results. We assessed the variance of effective population size (N_e) estimates with 50 bootstraps in randomly sampled segments with replacement.

In a second approach, we estimated changes in effective population size through time for *G. agassizii* and *G. morafkai* using SMC++ (Terhorst, Kamm, and Song 2017). SMC++ calculates the site frequency spectrum (SFS) relative to a 'distinguished individual' selected from the pool of samples against which all other samples are compared. We generated 10 SMC++ formatted files per scaffold greater than 10 Mbp in length for each species, with a different sample serving as the 'distinguished individual' in each. We converted input data from a VCF file filtered for a minimum depth of four and genotype quality of 30 to SMC++ formatted input files using the *vcf2smc* script masking contiguous stretches of homozygosity greater than 30 kbp. We calculated a single composite likelihood estimate for each species from the product of estimates across each scaffold and 'distinguished individual'. We conducted runs assuming a mutation rate of $2.675e^{-8}$ substitutions per site per generation, 20 knots (inflection points), 50 EM iterations and a polarisation error of 0.5. We limited the time interval of inference to between 300 and 500,000 generations before present. We applied a generation time of 25 years to scale the SMC++ results.

2.7 | Variants per Genic Element

Based on the F_{ST} results from ANGSD, we conservatively chose the 1% of windows with the highest F_{ST} values between the two species to constitute highly diverged regions. These regions were intersected using BEDTools v2.24.0 (Quinlan and Hall 2010) with the genome annotation file for *G. agassizii* (Dolby et al. 2020) to determine which genes or genic elements were in highly diverged regions (hereafter called 'diverged genes').

We removed duplicates from the overlapping windows of the intersected dataset and characterised polymorphisms based on which genic element they occurred in (see Data Accessibility for scripts). Because promoters were not annotated for these taxa, polymorphisms within 1000 bp upstream of the annotated start site of the gene were assumed to be in the promoter of that gene (estimate based on Sur and Taipale 2016). However, promoter lengths vary across the genome and this approach has the potential to both include non-promoter sequence and exclude some portions of promoters (see Figure S5A), depending on the gene. We also note that 5′ UTRs were not annotated for all genes, so this region includes 5′ UTRs for some genes (for reference, the average length of annotated 5′ UTRs among highly diverged genes in the *G. agassizii* genome was 181 bp). This could potentially underestimate 5′ UTR divergence, but these are limitations of genome annotations for non-model species. Polymorphisms not part of the gene or potential promoter region were considered intergenic.

2.8 | Validation of Diverged Genes

For a handful of genes, the gene ID of a high F_{ST} gene occurred more than once in the annotation. We validated these to check for duplications and for signs of pseudogenisation, which could present as high F_{ST}. To do so, we blasted the human orthologue for these genes to the *G. agassizii* proteins and retained the top blast hit as the true orthologue. In three situations (*PRDM10*, *TRDC* and *ZNF236*) the e-values were identical, so we viewed the matches in IGV (Freese, Norris, and Loraine 2016); if a tandem duplication seemed plausible (i.e., genes were next to each other) we retained them (*PRD10* and *ZNF236*). Finally, we validated the most diverged genes and those of particular biological interest (*VCX3a*, *XPA*, *MED12*, *MUC5AC* and *MUC5B*) for signs of pseudogenisation by checking the annotated genome transcripts for incomplete reading frames, internal stop codons and missing final stop codons. We labelled any transcripts that did not pass as a pseudogene. Data visualisation was done in R with tidyverse (Wickham et al. 2019) and ggplot2 (Wickham 2011).

2.9 | Functional Enrichment

We determined biologically enriched processes within the 112 highly diverged genes that contained putatively functional variants (defined here as variants in the promoter, UTR or exons) with two approaches. First, we ran that list of de-duplicated functional diverged genes through g:Profiler v.e109_eg56_p17_773ec798 to perform statistical enrichment analysis. We then re-ran g:Profiler by genic feature (Raudvere et al. 2019) to determine if there was element-specific GO term enrichment, using the g:SCS method for adjusted *p*-values in both cases. Because each high F_{ST} gene was effectively tested twice for enrichment (once as part of the whole list and once as part of a gene element list), we manually corrected for multiple tests using the Bonferroni method which gave a revised alpha significance threshold of 0.025 (instead of 0.05).

We took a complementary approach to functional interpretation by examining thematic clusters of genes whose products are known to interact within the cell or organism by running the final functional diverged gene list through the STRING Interaction Database (Szklarczyk et al. 2023). We then manually curated the network to look for biologically meaningful clusters (subnetworks). We used human as the reference because of its extensive functional annotation, did not allow the addition of genes and set the minimum required interaction score to medium confidence which is the default setting (0.4).

2.10 | Mucin 5 Re-Annotation and Synteny

During the validation process (see Validation of diverged genes), we discovered that Mucin 5 returned several hits, many of which contained complete ORFs. To determine the best orthology of the *MUC5B* and *MUC5AC* genes on the diverged list, we analysed the synteny and relatedness of all genes in the MUC-2-5-6

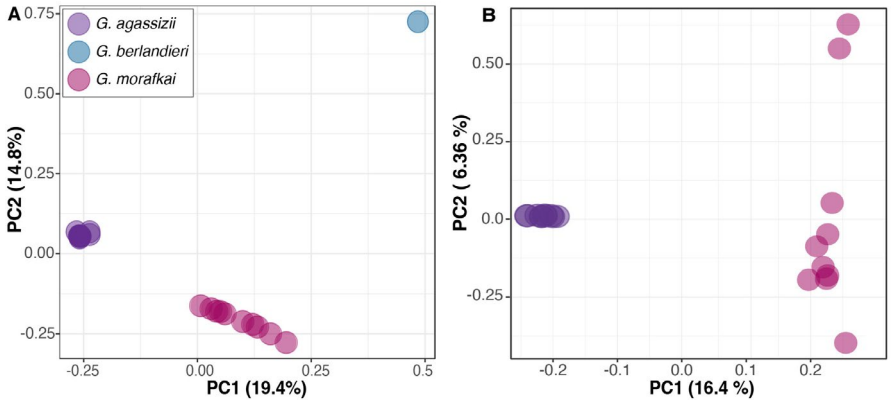

FIGURE 2 | Principal components analysis of whole-genome variation for (A) Mojave Desert tortoise (*Gopherus agassizii*), Sonoran Desert tortoise (*Gopherus morafkai*) and outgroup Texas tortoise (*Gopherus berlandieri*), and (B) solely within *G. agassizii* and *G. morafkai*.

TABLE 1 | Statistics on data generated in this study including sample size, sequence coverage, nucleotide diversity, effective population size, number of polymorphic loci and number of genes found in the 1% of the genome most diverged between *Gopherus agassizii* and *Gopherus morafkai*.

	G. agassizii	*G. morafkai*	*G. berlandieri*
Sample size (N)	10	10	1
Sequence coverage per sample (range)	2.9× (2.4–4.2×)	4.5× (3.4–5.8×)	7.5×
Genetic diversity (θ)	0.003	0.005	—
Effective population size (N_e)	34,000	57,000	—
Number genes in diverged regions (genes with putatively functional variants)	458 (127)		—

mucin array, defined as the genomic region between *TOLLIP* and *APA2A*, for six additional species downloaded from NCBI (*Gallus gallus*, *Mus musculus*, *Malaclemys terrapin*, *Trachemys scripta elegans*, *Chelonoidis niger abingdonii* and *G. evgoodei*), along with the long-read reference genome of *G. morafkai* (Baty et al. 2024). We checked predicted transcripts of genes in this array for ORFs and used annotations to manually map synteny based on visualisations in IGV. We examined the mouse, terrapin and tortoise annotations for unannotated mucin genes by blasting *G. agassizii* transcripts to the genome within the TOLLIP to APA2A syntenic region (Altschul et al. 1990). Finally, to estimate whether variants in the diverged mucin genes were potentially functional, we identified protein domains using analysis of the predicted transcripts through SMART, which searched for PFAM domains, signal peptides and internal repeats (Letunic, Khedkar, and Bork 2021). We recorded when SNPs and INDELs from our highly diverged gene list co-located within these protein domains.

2.11 | Mucin 5 Gene Tree Reconstruction

To determine relatedness of genes in the re-annotated mucin array, we used gene sequences from the synteny analysis to make a mucin phylogeny for *G. agassizii*, *G. morafkai*, *G. evgoodei*, *Chelonoides niger abingdonii*, *Malaclemys terrapin* with *Mus musculus* as an outgroup. We imported sequences into Geneious 2023.2.1 (https://genious.com), translated and aligned using Clustal Omega (Sievers et al. 2011). The protein alignment was fed into ModelTest-NG (Darriba et al. 2020) where a VT+I+G4 model was selected based on lowest AICc score. The alignment was then transformed into a NEXUS file for MrBayes v.3.2.7 (Ronquist et al. 2012) which we ran for 20,000,000 generations each over three runs with four chains per run and a 20% burn-in. Tracer v.1.7.2 was used to assess convergence of the MCMC chains (Rambaut et al. 2018).

2.12 | 3′UTR Differentiation

We evaluated whether the variants in 3′ UTRs of diverged genes mapped to known miRNA-binding sites and therefore could affect post-transcriptional and translational regulation. We downloaded conserved miRNA sequences from TargetScan 8.0 (Agarwal et al. 2015; McGeary et al. 2019) and used the reverse complement of the 6-8mer seed region of the miRNAs as the potential-binding site in the mRNA. We then pulled the 3′ UTR sequences of *G. agassizii* genes in high F_{ST} regions that had 3′ UTR changes and blasted the miRNA sequences to these mRNAs, retaining coordinates of blast hits. Finally, we cross-referenced the blast hit coordinates with the locations of the SNP and INDEL variants, accounting for strandedness.

3 | Results

3.1 | Pairwise F_{ST}, Genetic Diversity, N_e

Gopherus agassizii and *G. morafkai* are strongly differentiated in the PCA, though they fall closer to each other than the outgroup *berlandieri* (Figure 2A) and remain separate with the

outgroup removed (Figure 2B). Mean genome-wide pairwise F_{ST} between *G. agassizii* and *G. morafkai* was 0.34 and the 1% cut-off gave a highly diverged F_{ST} threshold of ≥0.63. After filtering there were 22,145 total variants in the top 1% of diverged regions (Table 1). Genetic diversity (θ) in *G. morafkai* was nearly double that of *G. agassizii* (0.005 vs. 0.003), equating to estimates of higher N_e (57,000 vs. 34,000). These are corroborated by reconstructions of historical N_e, which remained higher for *G. morafkai* throughout the last million or so years (Figure 3, Figure S3). Their historical N_e patterns were similar with a decline ~50 kya, an increase ~12 kya and additional decline ~9 kya for *G. morafkai* which led these trends ahead of *G. agassizii*. Sampling covered a similar geographic area in the centre of each range and diversity differences are therefore unlikely to be a sampling artefact; in fact, the sampling range for *G. morafkai* was slightly smaller but yielded much higher diversity. We note that differences in sequencing coverage could impact diversity estimates (*G. agassizii* samples had 2.9× average coverage, whereas *G. morafkai* samples had 4.5× average coverage), although ANGSD is relatively robust to low sequence depth (Lou et al. 2021).

3.2 | Variants per Genomic Element

The high F_{ST} regions included 458 unique genes based on an intersection with the *G. agassizii* reference genome (Table 1). About 80% of the high F_{ST} variants mapped to the flanking intergenic regions within these windows for which we have no evidence for functional importance (Figure 4a). INDEL and SNP variants showed identical distributions (Figure 4a). In total there were 3637 genic (defined here as promoter, UTR, exonic and intronic) SNPs and 15,393 intergenic SNPs (to see the difference between flanking variation versus promoter variation, see Figure S5). Removing genes with only flanking intergenic or intron changes left 127 genes with promoter, UTR or exon changes (Figure 4b; Table 1). The number of 5′ UTR variants is likely undercounted as not all genes had annotated 5′ UTRs, and these would be considered under promoter divergence in our classification scheme. Hereafter, we refer to promoter, UTR, or exon variants as 'functional' or 'putatively functional'. Of those variants that occurred within the 3′ UTR, none localised to conserved 3′ UTR miRNA-binding sites, but this does not consider species-specific miRNA-binding sites, which are unknown for these taxa.

3.3 | Functional Enrichment

We ran G:profiler on the entire functional gene list and on each functional element individually (defined here as the promoter, 5′ UTR, exon and 3′ UTR; see Table 2). Processes enriched from the entire list were *transferase activity* and *carbohydrate derivative binding*. For promoter variants, *transferase activity* (p value = 0.027) and *nucleotide excision repair in xeroderma pigmentosum* (p value = 0.025) were not significant at the Bonferroni corrected alpha value ($\alpha = 0.025$). Exon variants were enriched for the *DRIP complex*, the *ARC92-Mediator*

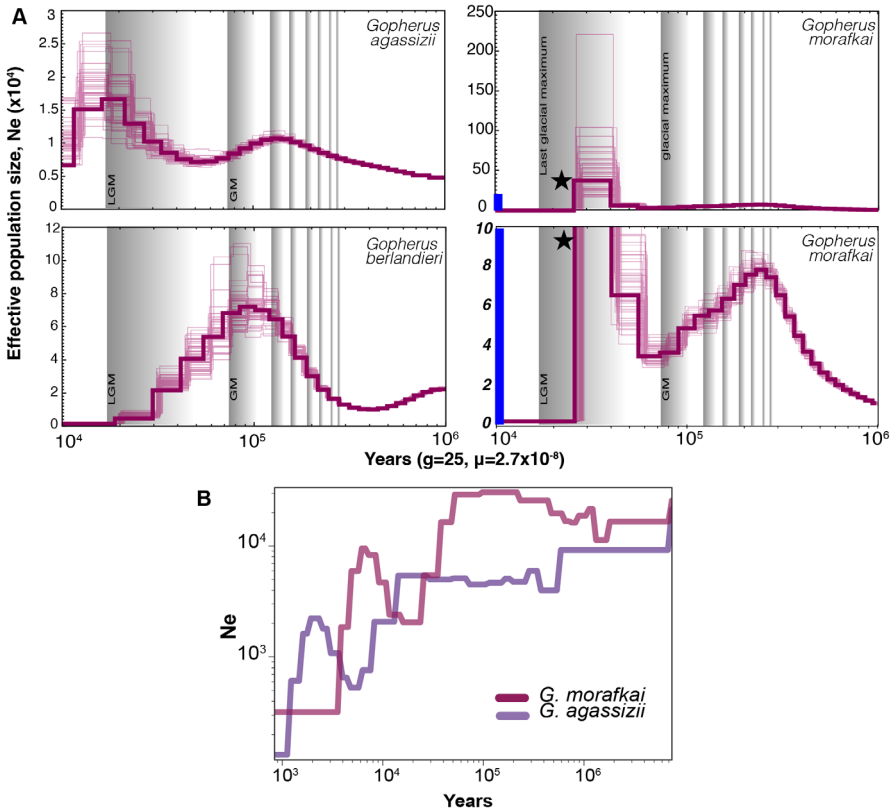

FIGURE 3 | Historical demography for the three *Gopherus* lineages (A) using PSMC. Note difference in scale of y-axis between panels, *G. morafkai* is shown twice (on right) for clarity, and x-axes are log scale. Analysis assumes a 25-year generation time and mutation rate of 2.7×10^{-8}. Pleistocene glacial cycles (peak glacial periods) are depicted in greyscale. Effective population size is much higher for *G. morafkai* (middle). (B) Historical demography of *G. agassizii* and *G. morafkai* based on 10 low coverage genomes per species in SMC++.

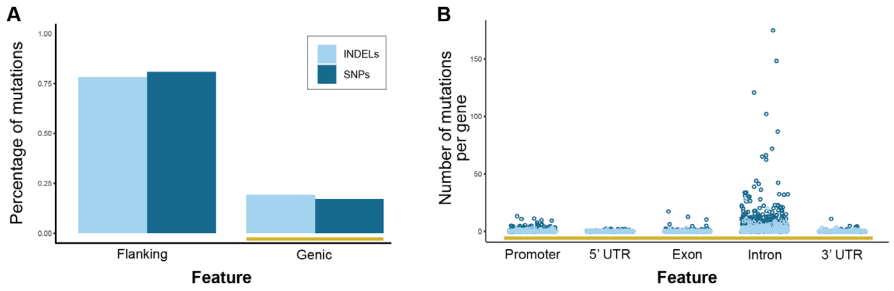

FIGURE 4 | Distribution of how variants in the 1% most-diverged genomic regions map to genomic elements. (A) Most variants map to flanking intergenic regions as expected; (B) of variants that mapped to genic regions in (a), these are the total variants mapping to genic elements according to the annotation for *Gopherus agassizii*. Variants mapped to promoter, UTRs or exons are considered 'functional' in downstream analyses. Promoter is defined as 1 kb immediately upstream of the annotated start site of the gene.

TABLE 2 | Statistically enriched gene ontology (GO) categories for genes most diverged between *Gopherus agassizii* and *Gopherus morafkai* based SNPs and INDELs in promoter, UTR, and exon regions.

GO name	GO ID	Adjusted *p*	Associated genes
Promoter			
Transferase activity, transferring phosphorus-containing groups^	GO:0016772	2.75E-02	*OAS1*, *FCSK*, *MST1R*, *TRPM6*, *ACVR1*, *PRKAR2A*, *PNPT1*, *PARP3*, *POLQ*, *POLD4*, *MYO3B*, *HGS*, *TK1*
Nucleotide excision repair in xeroderma pigmentosum[a]	WP:WP5114	2.53E-02	*XPA**, *POLD4*, *XAB2*, *GPS1*
Exon			
Transcription coregulator activity[a]	GO:0003712	2.59E-02	*MED12**, *DYRK1A*, *SIN3B*, *MED26**, *ZNF541*, *BRD8*
DRIP complex	CORUM:548	5.70E-03	*MED12**, *MED26**
DRIP complex	CORUM:549	5.70E-03	
ARC92-Mediator complex	CORUM:909	5.70E-03	
ARC complex	CORUM:288	6.64E-03	
ARC complex	CORUM:232	7.65E-03	
Entire dataset			
Transferase activity, transferring phosphorus-containing groups	GO:0016772	2.40E-04	*PRKRA*, *PAPSS2*, *CLK1*, *STK17B*, *OAS1*, *FCSK*, *DYRK1A*, *MST1R*, *TGFBR3*, *TRPM6*, *ACVR1*, *PRKAR2A*, *PNPT1*, *PARP3*, *POLQ*, *POLD4*, *MYO3B*, *HGS*, *TK1*
Carbohydrate derivative binding	GO:0097367	5.18E-03	*TGFBR3*, *LAYN**, *CLEC4G*, *CYR61*, *MST1R*, *FCSK*, *RAB2B*, *RAC3*, *TK1*, *ACVR1*, *PRKAR2A*, *TTLL12*

Note: Asterisks indicate the 13 genes with the highest number of variants.
[a]Processes not significant after Bonferroni correction.

complex as well as the *ARC complex*. *Transcription coregulator activity* was not significant at the corrected alpha value (p-value = 0.026).

The STRING interaction analysis produced a network with 112 nodes and 54 edges. The average node degree was 0.96 and the average local clustering coefficient was 0.28. The genes interacted significantly more than expected by chance with a protein–protein interaction (PPI) enrichment p-value of 0.017.

3.4 | Mitonuclear Divergence Genes

Nine diverged genes formed a STRING cluster relating to mitochondrial function that included the mitochondrial ribosome (mitoribosome) and the cytosolic ribosome (pink cluster, Figure 5). Nuclear-mitochondrial dysregulation is a noted source of incompatibility and cytonuclear discordance in hybrids (Burton, Pereira, and Barreto 2013). Genes in this cluster included Mitochondrial Translation Release Factor 1-like

(*MTRF1L*, 3′ UTR +2 bp INDEL) which plays a role in termination of mitochondrial translation; 3′ UTR INDELs may lead to divergent post-transcriptional modification, perhaps to change the transcript residence time or localisation within the cell. *SLC25A38* similarly had a 3′ UTR INDEL (+5) as well as promoter SNP and INDEL variants. *SLC25A38* transports glycine into the mitochondrial matrix while PNPT1 (one SNP) localises to the intermembrane of the mitochondrion and plays a role in the transport of RNA into the mitochondrion and its degradation (Liu et al. 2018), suggesting changes to mitochondria-RNA interactions. Two additional genes (*MSTO1* and *ARMCX3*) had several changes each and both function in the distribution of mitochondria within the cell (Donkervoort et al. 2019; López-Doménech et al. 2012).

Also in this cluster were genes involved in the processing of ribosomal RNA transcripts and structure of the 40S subunit of the cytosolic ribosome (*TSR1*, *RRP9* and *RPSA*). *TSR1* (four 3′ UTR and two exon SNPs) and *RRP9* (one exonic SNP and two exonic INDELs) variants suggest differences in both the protein and post-translational modification of these pre-ribosomal proteins. They interacted with *ZPR1* and *WDR6*, which are implicated in positive and negative cell cycle regulation respectively, and *TK1*, which performs cell-cycle dependent salvaging of dTTP—a nucleoside necessary for DNA replication and repair. Together, these genes can affect mitochondrial and non-mitochondrial translation as well as nuclear–mitochondrial interactions. However, addition of mitochondrial data could better resolve potential mito-nuclear incompatibilities.

3.5 | Chronobiological Genes

3.5.1 | UV DNA Repair, Cell Cycle, Circadian Rhythm

UV causes damage to DNA that must be repaired for progression through the cell cycle, and circadian rhythms of the cell are coupled to the cell cycle through shared regulatory proteins. We found many diverged genes at the intersection of these three processes (yellow cluster, Figure 5). Transcription coupled (TC) and Global Genome (GG) Nucleotide Excision Repair (NER) pathways are critical for repairing UV DNA damage, and we find NER pathways significantly enriched among these genes (REAC:R-HSA-5696398, WP:WP5114). At the hub of this cluster was *XPA* (XPA, DNA damage recognition and repair factor; two 5′ UTR and 11 promoter SNPs, Table 3) which is a core gene in the DNA repair response to UV damage. It accumulates at damaged DNA, acting as a scaffold for proteins of the NER excision complex to bind (Borszéková Pulzová, Ward, and Chovanec 2020). XAB2 (XPA-Binding Protein 2, one promoter SNP) co-regulates transcription of *XPA* and plays a role in the TC-NER response to UV DNA damage (Borszéková Pulzová, Ward, and Chovanec 2020; Kuraoka et al. 2008; Nakatsu et al. 2000) and interacts with RNA polymerase II (Mirastschijski et al. 2019). A core circadian pacemaker gene controls *XPA*, which exhibits a 24-h cycling of its concentration between the cytoplasm and nucleus (Dakup and Gaddameedhi 2017). This circadian process is interrupted at the occurrence of UV DNA damage (Dakup and Gaddameedhi 2017) and could therefore be related to the significant differences in UV DNA exposure between the deserts (Baty et al. 2024) and/or associated with circadian behavioural differences, which are not well documented.

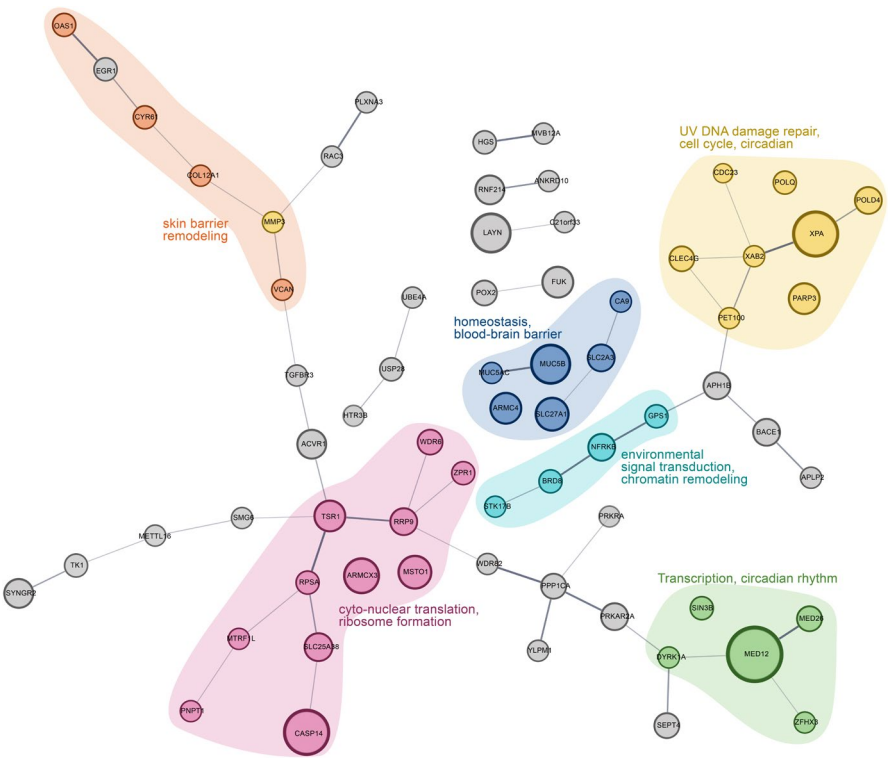

FIGURE 5 | STRING interaction network of genes within highly diverged F_{ST} windows that interact based on available evidence. Clusters with biologically meaningful themes are colour-coded (see Section 4). Each node is a gene, edges are weighted to reflect the degree of confidence of node–node interactions. All genes shown were in the highly diverged regions with putatively functional variants; nodes are scaled by the number of variants in putatively functional regions (promoter, exons or UTR). Non-interacting genes are not shown.

TABLE 3 | Genes with the greatest number of potentially functional variants (SNPs or INDELs) in promoter (pr), 5′ UTR, exon, 3′ UTR regions where promoter is defined as 1 kb immediately upstream of each annotated gene. Nonzero values are bolded.

Gene name	Gene function	Number of changes			
		Pr.	5′ UTR	Exon	3′ UTR
IGHV3-11	Antigen recognition	**6**	0	**18**	**4**
IGHV3	Adaptive immunity	**10**	0	**13**	0
MED12	RNA polymerase II transcription mediator	**16**	**1**	**1**	0
CLEC1A	Cell signalling, adhesion; immune response	**1**	**2**	**2**	**13**
XPA	DNA damage repair	**11**	**2**	0	0
CASP14	Apoptosis	**13**	0	0	0
TRDV3	T-cell receptor	**5**	0	**6**	0
LAYN	Hyaluronic acid binding	**10**	0	0	0
MUC5B	Gel-forming mucin production	0	0	**10**	0
TRAV18	T-cell receptor	**3**	0	**6**	0
GTF2IRD2	Transcription regulation	**9**	0	0	0
DER	Oxidoreductase	**9**	0	0	0
ARMCX3	Tumour suppression	0	0	0	**8**

UV radiation also causes DNA strand breaks. The gene *POLD4* (DNA polymerase Delta 4 Accessory Subunit) was in the same STRING cluster. *POLD4* is central to genome replication and repair, plays a role in NER (Blank, Kim, and Loeb 1994) and response to UV damage (Kojima et al. 2021). In this cluster, *POLQ* (DNA Polymerase Theta; two promoter SNPs) is responsible for the error-prone translesion repair synthesis that is necessary for replicating a genome with strand breaks and is thought to release stalled DNA replication forks (Costantino et al. 2014). *PARP3* (five promoter SNPs and one INDEL) is also involved in DNA strand break recognition and genome stability. Another cluster of genes (orange, Figure 5) were involved in the skin barrier and skin remodelling related to UV damage (*COL12A1*, *EGR1*, *VCAN* and *OAS1*). Skin is the main target of UV exposure, and the hub of this cluster was Matrix Metalloprotease 3 (*MMP3*), which cleaves collagen for degradation in response to UV damage (Mirastschijski et al. 2019). Overall, there are strong results related to differences in skin and DNA damage resulting UV exposure, which differs between the deserts (Baty et al. 2024).

In eukaryotes, the G_2/M checkpoint is a critical cell cycle checkpoint that prevents a damaged nuclear genome from continuing to mitosis by stalling and allowing time for DNA repair (Löbrich and Jeggo 2007). While *POLD4* function is necessary for this progression (Huang et al. 2010), the same UV cluster (yellow, Figure 5) contained Cell Division Cycle 23 (*CDC23*)—which encodes a protein also necessary for progression past the G_2/M checkpoint and which is normally highly conserved in eukaryotes (Zhang et al. 2011). (It is also key for spindle assembly during mitosis and meiosis (Harper, Burton, and Solomon 2002; Zhou et al. 2020)). *POLD4* had several promoter SNPs and INDELs and *CDC23* had a promoter INDEL (+2 bp), changes that could alter binding specificity, timing or magnitude of transcription of these DNA repair/cell cycle checkpoint genes. They could together coordinate to regulate the cell-cycle-UV damage response due to UV differences between the deserts in a time (circadian)-dependent way. All of these genes occurred in the most diverged genomic regions.

3.5.2 | Signal Transduction and Chromatin Remodelling

There was significant divergence in a cluster of genes related to transcription signal processing and perhaps environmental cues (teal, Figure 5). GPS1 is suppressive to immune-response signal transduction, and STK17B is also involved in intracellular signal transduction as well as positive regulation of apoptosis. *NFRKB* (four promoter SNPs) is thought to be a regulator of the INO80 chromatin remodelling complex that displaces nucleosomes to improve access to DNA for transcription, repair and replication (Eustermann et al. 2018). Similarly, at least one isoform of BRD8 alters nucleosome structure within chromatin, increasing accessibility of DNA for positive regulation of transcription. Some human BRD8 isoforms interact with thyroid hormone receptors (THRs) to increase thyroid-related transcription activity, such as metabolism. Thyroid hormone and its regulation are proposed to play a key role in circannual pace-making (Hazlerigg and Lincoln 2011; Lomet et al. 2018). This cluster of interacting genes documents divergence in the regulatory apparatus of genes involved in nucleosome remodelling and signal transduction that affect the thyroid-mediated physiological behaviours (which can be seasonal). If the transcriptional of nucleosome remodellers is altered, this could affect the transcription of functionally unrelated neighbouring genes, while changes in transcription related to THR could have cascading effects along signalling pathways.

3.5.3 | Transcriptional Regulation and Circadian/Circannual Rhythms

We observed divergence in regulatory regions within a cluster of genes relating to transcription initiation and circadian rhythm (green, Figure 5), along with significant divergence in the ARC/Mediator transcriptional complexes based on exon changes (CORUM288, CORUM:909 and others, Table 2). MED12 and MED26 are components of the Mediator, which facilitates the start of transcription by interacting with transcription factors and binding to RNA pol II, helping it localise to the core promoter. Surprisingly, *MED12* had extensive changes (one 5′ UTR SNP, 13 promoter SNPs, 3 promoter INDELs and 1 exon SNP) despite being a highly evolutionarily conserved subunit (Fondell 2013; Figure 6); *MED26* had exon and promoter variants.

Interacting with the Mediator genes was *ZFHX3* (one exon SNP, three promoter variants), a transcription factor of, and highly expressed in, the Suprachiasmatic Nucleus (SCN) of the hypothalamus. The SCN plays the critical and primary role in regulation of circadian rhythms based on zeitgebers (external cues, often photoperiod). Variants in *ZFHX3* affect circadian rhythm clocks (Parsons et al. 2015), circadian-related response to light (Hughes et al. 2021) and sleep cycles (Balzani et al. 2016) in mice. Though much less well understood, the hypothalamus–pituitary axis may also shape circannual seasonal rhythms (Hazlerigg and Lincoln 2011; Wood and Loudon 2014), such as periods of mating or torpor. These are controlled at least in part through changes in thyroid hormone and THRs; BRD8 (prior section) interacted with these receptors in the network and could play a role in the strong seasonal behavioural differences observed between these species. Also in this cluster was *SIN3B*, which negatively regulates RNA polymerase II, transcription and plays a role in circadian clock-dependent heterochromatin formation (hub gene, yellow cluster; Zhu and Belden 2020).

3.6 | Homeostasis: Water and Energy Balance Genes

Five genes associated with the lung mucosa and/or blood–brain barrier (BBB; blue, Figure 5). Of these, two members of the Solute Carrier Family 2 (*SLC27a1*, *SLC2a3* [*3′ UTR changes*]) are integral parts of the facultative transport of glucose across the blood–brain barrier (Deng et al. 2015; Hoskin et al. 2003). *SLC27a1* (five promoter SNPs and two 5′ UTR SNPs) is involved in the transport of lipids across the blood–brain barrier (Mitchell et al. 2011). The BBB is selectively permeable, allowing nutrients to reach the brain while blocking pathogens (Zhao et al. 2022). The abundance of 3′ UTR and promoter changes in these genes could play a role in transcriptional or post-transcriptional regulation of energy balance and differences in the metabolism and transport of energy molecules across the BBB; they could also relate to blocking of pathogens across the barrier. Also in this cluster was Carbonic Anydrase 9 (*CA9*), a zinc metalloenzyme that forms bicarbonate. Related genes in the CA family are integral to forming cerebrospinal fluid, and while CA9 primarily localises to the gastrointestinal mucosa in humans, it is present in the cerebellum and may play a role in pH regulation of cerebrospinal fluid, potentially linking the BBB to pathogen defence in the lung and gut mucosa.

Genes related to lung mucosal properties and host innate immunological defence were MUC5AC (one promoter SNP) and MUC5B, which are gel-forming secreted mucin proteins specific to the lung and nasal epithelia. *MUC5B* had 10 exon SNPs, six of which were nonsynonymous and all 10 located in functional protein domains. They were not identified as interacting in the STRING network with *CA9*, but data indicates *CA9* and *MUC5AC* are highly co-expressed (Spearman correlation of 0.98; proteinatlas.org). In the lung, *CA9* could form the bicarbonate needed to change calcium ion concentrations necessary to stimulate the release of both mucins (Hansson 2019). Upon secretion the mucin proteins expand in volume over 1000-fold (Hansson 2019) and bind and sequester water to form a diffusion barrier (Hansson 2019; Schade, Flemström, and Holm 1994), potentially playing a significant role in xeric-adapted water retention and pathogen defence.

3.7 | Structural Variation in Mucin Gene Family Arrays

Based on reference genome analysis, we identified major structural differences in the gene family arrays of lung-specific Mucins across taxa. Of the MUC5-MUC2-MUC6 array, the outgroup (*C. abingdonii*) had five MUC genes, *G. evgoodei* had six, *G. agassizii* had nine, of which five had complete ORFs and four were pseudogenised (Figure 6D), and *G. morafkai* had 16 mucin genes—15 with complete ORFs and only one pseudogenised. Therefore, an expansion of MUC5-2-6 genes appear specific to desert tortoises, but *G. morafkai* shows a large, lineage-specific expansion while *G. agassizii* appears to have lineage-specific pseudogenisation of its somewhat expanded array. Two of those functional genes (one *MUC5AC* and one *MUC5B*) were on the list of highly diverged genes. *MUC6* in *Gopherus agassizii* and *C. n. abingdonii* did not have complete ORFs. For both *G. agassizii* and *G. morafkai* duplications occurred across the array, instead of at the ends of the array like in other chelonians.

The phylogenetic reconstruction for mucin genes converged with an effective sample size of 35,177 according to Tracer. The mucin phylogeny had five clades (Figure 6C) with MUC6 and MUC2 forming clear, distinct groups. In the MUC6 group there were two genes annotated as MUC2 (from *G. evgoodei* and *M. terrapin*), but they locate to the end of the mucin array next to MUC6 and based on synteny are probably annotation errors. The MUC5 clade included a MUC5B group and a MUC5AC group. The MUC5AC group had two subgroups, a 'core' MUC5AC group and a more conglomeratic MUC5AC/MUC5B group which we infer to be duplicated from a MUC5AC ancestor. The homology of the mixed MUC5AC group is difficult to resolve and could have arisen through or been complicated by gene conversion during DNA repair-based recombination (Chen et al. 2007). The 'core' MUC5AC subgroup contained only annotated MUC5ACs except for two that were labelled as MUC5B. The conglomeratic subgroup contained genes that were annotated as MUC5B, MUC5AC and MUC2.

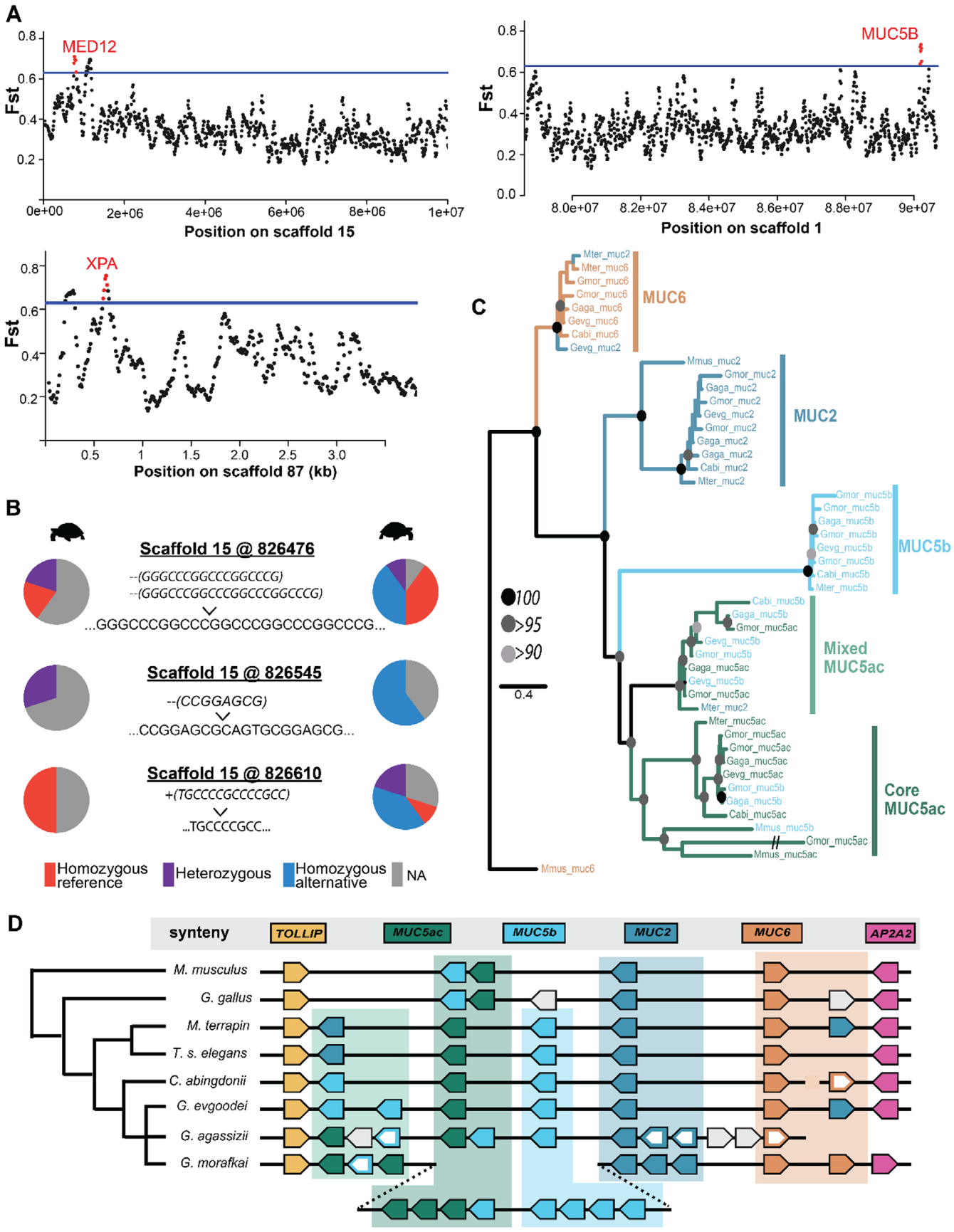

FIGURE 6 | Legend on next page.

FIGURE 6 | Characterisation and summary of diverged genes. (A) Manhattan plots showing F_{ST} of highly diverged genes discussed in the text (red): Mediation Complex Subunit 12 (*MED12*), DNA Damage Recognition and Repair Factor (*XPA*) and Mucin 5B (*MUC5B*). Outliers in black contained 'non-functional' variants and are not discussed. (B) Characterisation of promoter INDEL variants of MED12 as well as the variant proportion in *Gopherus agassizii* (left) and *Gopherus morafkai* (right) relative to the *G. agassizii* reference allele. (C) Phylogenetic reconstruction of the MUC2-5-6 gene array based on predicted proteins from mucin genes identified in this study and (D) synteny of the mucin gene array based showing the expansion of MUC5 genes in *G. morafkai*. Gene colours are based on original annotations and coloured boxes are homology estimations based on the phylogeny in C. Genes with white centers are pseudogenized.

4 | Discussion

In this study we examined 20 genomes from sister desert tortoise lineages, along with an outgroup individual, to understand genomic diversity and identify regions of potential adaptation differentiating the two lineages. These lineages have been diverged for more generations than many other speciation genomic study systems (Figure 1); this, combined with the many phenotypic differences between species, suggests they might be considered to be in the later stages of speciation. Consistent with this, we found a mean genome-wide pairwise F_{ST} of 0.34 (Figure S4), reflecting species-level differentiation (Araya-Donoso et al. 2022).

We found Sonoran Desert tortoises have much higher genetic diversity and current and historical N_e than Mojave Desert tortoises, which is consistent with their current conservation statuses under the US Endangered Species Act (USFWS 2010, 2022). A previous study found higher N_e estimates using only transcriptomic data, but with broader geographic sampling (Edwards, Tollis, et al. 2016). In the top 1% of diverged genomic regions, 80.5% of SNP and INDEL variants mapped to intergenic regions flanking genes, and 17% mapped to introns, leaving only 2% falling in putatively functional regions (promoters, UTRs and exons). Of the 458 genes identified in the diverged regions, 331 had only intergenic or intronic variation. This distribution is consistent with these regions having low selective constraint, but we note that these regions can contain many regulatory and other functional elements (Christmas et al. 2023; Nobrega et al. 2003; The ENCODE Project Consortium 2012). However, like many other non-model species, these regions have not been annotated for our species so we do not interpret them here. Among putatively functional regions, we find extensive divergence in promoters of genes related to chronobiological processes, such as circadian rhythm, circannual rhythm and cell cycle progression. Promoter divergence could affect transcription of pace-maker genes that regulate chronobiological functions, implying a meta-transcriptional contribution to reinforcement and speciation through potential differential regulation of key behavioural phenotypes.

4.1 | Sonoran Desert Tortoises Have Higher Genetic Diversity

The census sizes (N) of Mojave Desert and Sonoran Desert tortoises were previously estimated at ~336,000 (Allison and McLuckie 2018) and 470,000–970,000 (Arizona Ecological Services Field Office 2015), respectively. Based on genetic diversity, our estimates of effective population size (N_e; 34,000 and 57,000 respectively) are proportional to these census sizes. Consistent with the ESA listing decisions for both species, the Mojave Desert tortoise has much lower (almost half) the genetic diversity of the Sonoran Desert tortoise (Table 1, Figure 2). Both species were sampled within similarly small geographic extents within their ranges, so this is unlikely to be a consequence of sampling bias. Lower sequencing coverage of the Mojave individuals could affect results, though we picked analysis tools generally robust to this. Historical N_e was consistently higher for Sonoran Desert tortoises (Figure 3), with fluctuations that may be attributed to range changes during glaciations. The terrain of the two species' ranges differs, with the Sonoran Desert having greater topographic relief and mountain chains that could offer suitable microclimates during periods of climate change, whereas Mojave Desert tortoises mainly live in bajadas and creosote flats. Sonoran Desert tortoises are bounded by the Colorado River, Colorado Plateau and Grand Canyon, limiting northward migration in the face of climate change. Montane microclimates could provide suitable habitat in lieu of latitudinal shifts. We note that these habitat features can shape both census population size and population structure, the latter of which can strongly impact estimates of effective population size (Charlesworth 2009; Mazet et al. 2016).

4.2 | Putatively Non-Functional vs. Functional Variants

The mean genome-wide pairwise F_{ST} was high (0.34) (Figure S4), reflecting species-level differentiation (Araya-Donoso et al. 2022). The vast majority (80%) of variants found in highly diverged regions ($F_{ST} \geq 0.63$) were in intergenic flanking regions and are not associated with any annotated functional regions. In total, 331 of 458 genes in these highly diverged regions had only linked (flanking) or intronic variation. While some flanking variants probably occur in enhancers and could affect gene transcription, it seems unlikely that is true of most. Introns can be retained in final protein products and play regulatory roles in some genes, though to what degree and in what genes introns have this role in *Gopherus* is unknown (Rose 2019; Schmitz et al. 2017). Annotating regulatory RNA genes and regulatory elements, particularly enhancers, in the genomes of non-model organisms remains a challenge and is a current limit of such studies, including this one. Genome-wide F_{ST} analysis is a common way to study diverging lineages and can reveal regions of potential evolutionary importance beyond coding changes, including promoter and UTR changes. However, we note that transcriptomic or proteomic data is necessary to test whether changes in such candidate regions actually affect expression of those genes. Our results advise caution that when interpreting genes in high-F_{ST} regions as putatively adaptive, mapping the variation to genic elements is necessary for making functional interpretations and avoiding false positives (Amato et al. 2009; Hebert, Renaut, and Bernatchez 2013); adding expression or protein data is also needed to validate predictions.

4.3 | Divergence in Regulation of Chronobiological Processes

We found extensive divergence in core chronobiological processes, namely circadian rhythm, cell cycle regulation and circannual rhythms (yellow, teal, green clusters, Figure 5). The core hub gene in our analysis, *XPA*, is positively regulated by *BMAL1*—a key circadian clock gene (Dakup et al. 2018). In fact, cell cycle and cellular circadian biorhythms are strongly coupled through shared regulator genes: circadian clock gene *BMAL1* regulates key members of the G_2/M transition (and UV repair, XPA) meanwhile cell cycle genes (*CDK1* and *NR1d1*) feedback to regulate *BMAL1* (Gaucher, Montellier, and Sassone-Corsi 2018; Yan and Goldbeter 2019). From an evolutionary view, divergence in regulatory regions of a cohort of UV-cell-cycle and circadian rhythm genes could provide an important reinforcement mechanism whereby the regulation of biorhythms at the cellular level are mis-timed in hybrids, leading to cellular instability (summarised in Figure 7, Figure S2). This mechanism would be a classic example of Bateson–Dobzhansky–Mueller incompatibility (reviewed in Mack and Nachman 2017; Payne et al. 2022).

Sonoran Desert tortoises will bask in the sun and emerge to drink during winter while Mojave Desert tortoises brumate throughout winter (Sullivan et al. 2014). These differences in winter precipitation and temperature have been hypothesised to contribute to niche separation between the two tortoise species (Edwards et al. 2015). Timing of mating and egg-laying differ: Mojave Desert tortoises mate in spring/summer and lay April–July while Sonoran Desert tortoises mate in early fall following the summer monsoon and lay in summer (Averill-Murray 2002; Averill-Murray, Christopher, and Henen 2018; Ruby and Niblick 1994; Wallis, Henen, and Nagy 1999). Sonoran Desert tortoises have at most one clutch per year while Mojave Desert tortoises may adopt a bet-hedging strategy and have additional clutches (up to three)—a behaviour thought to be tied to the unpredictability of rainfall over their distribution (Averill-Murray, Christopher, and Henen 2018; Germano 1993).

A common zeitgeber is photoperiod, mediated by melatonin and light cues to the hypothalamus. The Mojave and Sonoran deserts are at roughly the same latitude and thus have similar photoperiods. The zeitgeber in this case could be precipitation, where summer monsoon/non-monsoon rainfall asynchrony cues offset seasonal behaviours, forming prezygotic isolation barriers (Hau et al. 2017; Quintero et al. 2014). Studies done in the Sonoran Desert have shown that concentrations of testosterone and luteinising hormone increase during summer monsoons, when breeding occurs (Deviche et al. 2006; Small, Sharp, and Deviche 2007). Precession-driven rainfall differences have been the zeitgeber for divergence among African giraffes (Coimbra et al. 2021; Thomassen et al. 2013). Divergence in regulatory elements underlying these chronobiological molecular processes may encode such behavioural differences, if those differences manifest at the transcriptional level.

Some studies pose that cell cycle, circadian and circannual rhythms are coupled (Abrieux et al. 2020; Gaucher, Montellier, and Sassone-Corsi 2018; Hazlerigg and Lincoln 2011; Wood and Loudon 2014). Divergence in regulatory regions of *MED12*, *MED26*, *ZFHX3* and *SIN3B* (which regulate gene transcription by different mechanisms) could manifest meta-transcriptional effects through the divergence of genes that control downstream regulation of suites of genes. Regulatory changes in chromatin remodelling complexes and hormone pathways may bring cascading transcriptional effects. It makes intuitive sense that regulatory changes may mediate the molecular processes that control chronobiological behaviours, where it is likely to be the timing or signalling of transcription, rather than changes in the specific proteins, that matter. This is highlighted by the regulatory divergence in transcription factor *ZFHX3*, which coordinates transcription to zeitgebers, in this case perhaps rainfall seasonality (Deviche et al. 2006; Small, Sharp, and Deviche 2007), which is thought to underlie behavioural reproductive differences of these lineages (Figure 7B).

4.4 | Evolution of Cytonuclear Discordance

Cytonuclear discordance is proposed to play an important role in speciation via selection against hybrids (reviewed in Burton, Pereira, and Barreto 2013). While mitochondria are maternally inherited, the mitochondria-related genes diverged here are encoded by the nuclear genome. Given the documented mitochondrial divergence between the species (Edwards, Vaughn, et al. 2016), hybridisation of these lineages would allow for mito-nuclear misalignment where nuclear-encoded genes of core mitochondrial functions interact poorly with the mito-genome of the other lineage. In this case, mito-nuclear interactions would be a natural target of selection. Mitochondrial genomes evolve faster than nuclear genomes; however, nuclear genes encoding core mitochondrial functions could be under associated selection pressure to maintain functional parity with the mitogenome of its own nascent species (Burton, Pereira, and Barreto 2013). Our results suggest divergence in the localisation of mitochondria within the cell (*MSTO1* and *ARMCX3*), transportation of proteins related to mitochondrial functions (*SLC25A38* and *PNPT1*) and mitoribosomal translation (*MTRF1L*). *MTRF1L* and *SLC25A38* had 3′ UTR INDELs, implying potential post-transcriptional differences between the species.

4.5 | Expansion of Lung Mucins in *Gopherus*

We find extensive expansion of the MUC5 group in desert tortoises compared with less xeric tortoises and turtles (Figure 6D). The Mojave Desert tortoise has three more mucin copies in this array than *G. evgoodei* (nine total, four of which are pseudogenised; Figure 6D). The Sonoran tortoise has 16 total with only one pseudogenised. The transcript lengths of MUC5s vary greatly, indicating that if they are transcribed, they may have different functions. In humans, this region of the genome frequently recombines (Rousseau et al. 2004) and duplications in other species have resulted in tissue-specific expression (Sveen et al. 2017), potentiating greater pathogen specialisation (Padra et al. 2014). In desert tortoises, evolution of the lung mucosal barrier could aid in sequestration of water to limit loss, and this could relate to the difference in drought duration between the deserts. The addition of transcription data would help determine which mucin genes are being expressed, and whether expression level is proportional to copy number or if dosage compensation has evolved to equalise transcript levels.

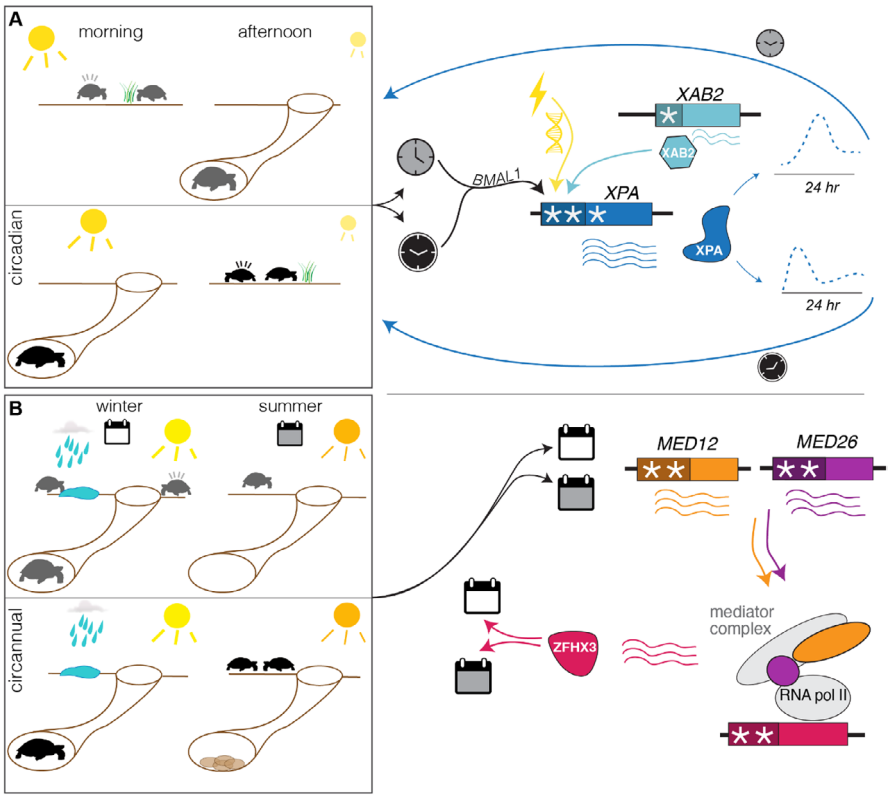

FIGURE 7 | Schematic showing evolutionary divergence in the regulatory regions of chronobiological processes. (A) *XPA* plays a key role in nucleotide excision repair of UV DNA damage and exhibits a circadian rhythm regulated by the clock gene *BMAL1*. *XPA* and its regulator *XAB2* showed extensive regulatory divergence which may relate to differences in circadian rhythm and DNA repair. Offsetting of internal clocks between the two species could results in hybrid incompatibilities. Circadian differences of tortoises shown here are only suggestions. (B) *MED12* and *MED26* (part of the Mediator Complex, grey) and *ZFHX3* all have promoter changes; *ZFHX3* is a transcription factor expressed in the Suprachiasmatic Nucleus of the Hypothalamus that is directly involved in the regulation of sleep–wake cycles and circadian rhythms. Regulatory changes can impact the timing, specificity, or magnitude of transcription that underlie known seasonal differences in the timing of mating, laying and brumation behaviours of the tortoise species. Cell cycle, circadian and circannual rhythms are proposed to be coupled (see Section 4.22).

Mucin 5 polymers also play a critical role in host defence by binding and trapping inhaled pathogens and could, therefore, play a role in adaptation to different pathogen challenges. A related diverged gene was *ARMC4* (four exon SNPs and two promoter SNPs), which localises to the ciliary of respiratory epithelium and contributes to motor function of the cilia. Cilia transport mucus and foreign bodies (including pathogens) out of the lungs (Howell et al. 2023). Additionally, both deserts experience dust storms and MUC5B has been shown to be upregulated in response to dust storms in humans (but not MUC5AC; Kim, Ye, and Shin 2011).

Taken together, there may be adaptive differences in brain–lung homeostasis (see results) related to dietary energy and osmoregulation through lung and nasal mucosa that may be motivated by water and pathogen differences between the deserts. The species may be evolving differently in response to *Mycoplasma agassizii* (Brown et al. 1994), the causative agent of URTD, which presents with nasal discharge and lesions of the upper respiratory tract mucosa (Brown et al. 1999). URTD disproportionately affects Mojave Desert tortoises and contributes to their population declines, many of which are below viability (Gov 2015; USFWS 2010). While the prevalence of *M. agassizii* among Sonoran Desert tortoises is also high, reports of URTD in wild Sonoran Desert tortoises are rare (Berry et al. 2015; Gov 2015; Jones n.d.). Given the differences in structural variation, genetic diversity and effective sizes (Table 1), Mojave Desert tortoises may be less able to adapt to novel pathogens or respond to greater drought challenges. Conversely, mucus formation requires water, so the lower MUC expansion in *G. agassizii* could be a response to lower water availability.

5 | Conclusion

Analysis of genome-wide divergence of speciating desert tortoises revealed extensive differentiation in promoter regions as well as an expansion of mucin genes. Parsing the distribution of SNPs and INDELs across genic elements (promotor, UTRs and exons) indicate divergence among these lineages in energy and water homeostasis, UV-related DNA damage repair and cell cycle checkpoints, and environmental signal transduction coupled to chromatin conformation, transcription initiation and circadian rhythm. Cell cycle checkpoint and circadian and circannual processes are biorhythms critical to organism function which underlie key behavioural and reproductive differences between these species. It would make sense that some

adaptations, particularly temporal adaptations, be mediated through divergence in the timing and conditions under which genes are expressed. Results here suggest desert tortoises have speciated in part through divergence of regulatory elements that may control the expression timing of time-keeping genes and/or hormone signalling, which can in turn have knock-on effects at the cellular and organismal level. Additional transcriptomic and proteomic data will be critical for testing this hypothesis in the future.

Author Contributions

N. Jade Mellor: led analyses, original writing, vizualizations, methodology, validation. **Timothy H. Webster:** contributed writing, analyses, and resources. **Hazel Byrne:** contributed writing, analyses, and resources. **Avery S. Williams:** contributed analyses. **Taylor Edwards:** contributed writing and resources. **Dale F. DeNardo:** contributed to writing, resources, and investigation. **Melissa A. Wilson:** contributed to writing, analyses, investigation and led resources. **Kenro Kusumi:** contributed to writing, resources, and project administration. **Greer A. Dolby:** contributed to analyses, visualization, methodology, resources, led investigation, project administration, supervision, writing.

Acknowledgements

We thank Carl Franklin at the Amphibian and Reptile Diversity Research Center at the University of Texas at Arlington for sharing Texas tortoise sample UTA R 59507. We thank Cristina Jones, Tom Jones, Chad Rubke for agency and logistical support, A. Nguyen for SMART analyses, and B. Todd for helpful conversations. We are further grateful to the University of Utah Center for High Performance Computing for computational support and resources.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Scripts used in this study are available https://github.com/njm-data/tortoise_speciation and https://github.com/thw17/Agassizii_morafkai_comparison. The data are available. Sequence data generated for this study are available under PRJNA356763 https://dataview.ncbi.nlm.nih.gov/object/PRJNA356763 (sequence accessions: SRR5100746-55, SRR31204776; sample accessions: SAMN0625105-114, SAMN44560062).

References

Abrieux, A., Y. Xue, Y. Cai, et al. 2020. "EYES ABSENT and TIMELESS Integrate Photoperiodic and Temperature Cues to Regulate Seasonal Physiology in *Drosophila*." *Proceedings of the National Academy of Sciences of the United States of America* 117, no. 26: 15293–15304. https://doi.org/10.1073/pnas.2004262117.

Agarwal, V., G. W. Bell, J.-W. Nam, and D. P. Bartel. 2015. "Predicting Effective microRNA Target Sites in Mammalian mRNAs." *eLife* 4: e05005. https://doi.org/10.7554/eLife.05005.

Ahmed, S., D. G. Brickner, W. H. Light, et al. 2010. "DNA Zip Codes Control an Ancient Mechanism for Gene Targeting to the Nuclear Periphery." *Nature Cell Biology* 12, no. 2: 111–118. https://doi.org/10.1038/ncb2011.

Allison, L. J., and A. M. McLuckie. 2018. "Population Trends in Mojave Desert Tortoises (*Gopherus agassizii*)." *Herpetological Conservation and Biology* 13, no. 2: 433–452.

Altschul, S., W. Gish, W. Miller, E. W. Myers, and D. Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215: 403–410.

Amato, R., M. Pinelli, A. Monticelli, D. Marino, G. Miele, and S. Cocozza. 2009. "Genome-Wide Scan for Signatures of Human Population Differentiation and Their Relationship With Natural Selection, Functional Pathways and Diseases." *PLoS One* 4, no. 11: e7927. https://doi.org/10.1371/journal.pone.0007927.

Andrews, S. 2010. "FastQC: A Quality Control Tool for High Throughput Sequence Data [Computer Software]." http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Araya-Donoso, R., S. M. Baty, P. Alonso-Alonso, et al. 2022. "Implications of Barrier Ephemerality in Geogenomic Research." *Journal of Biogeography* 49, no. 11: 2050–2063. https://doi.org/10.1111/jbi.14487.

Arizona Ecological Services Field Office. 2015. "Endangered and Threatened Wildlife and Plants; 12-Month Finding on a Petition to List Sonoran Desert Tortoise as an Endangered or Threatened Species." *Federal Register* 80, no. 193: 60321–60335.

Averill-Murray, R. C. 2002. "Reproduction of *Gopherus agassizii* in the Sonoran Desert, Arizona." *Chelonian Conservation and Biology* 4, no. 2: 295–301.

Averill-Murray, R. C., T. E. Christopher, and B. T. Henen. 2018. "Reproductive Ecology and Life History of Female Sonoran Desert Tortoises (*Gopherus morafkai*)." *Herpetological Monographs* 32, no. 1: 34–50. https://doi.org/10.1655/HERPMONOGRAPHS-D-18-00003.1.

Balzani, E., G. Lassi, S. Maggi, et al. 2016. "The Zfhx3-Mediated Axis Regulates Sleep and Interval Timing in Mice." *Cell Reports* 16, no. 3: 615–621. https://doi.org/10.1016/j.celrep.2016.06.017.

Baty, S. M., R. Araya-Donoso, A. Paulsen, et al. 2024. "Strong Signatures of Selection on Candidate Genes Underlying Core Speciation Mechanisms in Desert Tortoises." https://doi.org/10.1101/2024.06.06.597788.

Becking, T., M. A. Chebbi, I. Giraud, et al. 2019. "Sex Chromosomes Control Vertical Transmission of Feminizing Wolbachia Symbionts in an Isopod." *PLoS Biology* 17, no. 10: e3000438. https://doi.org/10.1371/journal.pbio.3000438.

Becking, T., C. Gilbert, and R. Cordaux. 2020. "Impact of Transposable Elements on Genome Size Variation Between Two Closely Related Crustacean Species." *Analytical Biochemistry* 600: 113770. https://doi.org/10.1016/j.ab.2020.113770.

Berry, K. H., M. B. Brown, M. Vaughn, et al. 2015. "*Mycoplasma agassizii* in Morafka's Desert Tortoise (*Gopherus morafkai*) in Mexico." *Journal of Wildlife Diseases* 51, no. 1: 89–100. https://doi.org/10.7589/2014-04-083.

Blank, A., B. Kim, and L. A. Loeb. 1994. "DNA Polymerase Delta Is Required for Base Excision Repair of DNA Methylation Damage in *Saccharomyces cerevisiae*." *Proceedings of the National Academy of Sciences of the United States of America* 91, no. 19: 9047–9051. https://doi.org/10.1073/pnas.91.19.9047.

Borszéková Pulzová, L., T. A. Ward, and M. Chovanec. 2020. "XPA: DNA Repair Protein of Significant Clinical Importance." *International Journal of Molecular Sciences* 21, no. 6: 2182. https://doi.org/10.3390/ijms21062182.

Boughman, J. W. 2001. "Divergent Sexual Selection Enhances Reproductive Isolation in Sticklebacks." *Nature* 411, no. 6840: 944–948. https://doi.org/10.1038/35082064.

Broad Institute. n.d. "Picard [Computer Software]." http://broadinstitute.github.io/picard/.

Brown, M. B., G. S. McLaughlin, P. A. Klein, et al. 1999. "Upper Respiratory Tract Disease in the Gopher Tortoise Is Caused by

Mycoplasma agassizii." *Journal of Clinical Microbiology* 37, no. 7: 2262–2269. https://doi.org/10.1128/JCM.37.7.2262-2269.1999.

Brown, M. B., I. M. Schumacher, P. A. Klein, K. Harris, T. Correll, and E. R. Jacobson. 1994. "*Mycoplasma agassizii* Causes Upper Respiratory Tract Disease in the Desert Tortoise." *Infection and Immunity* 62, no. 10: 4580–4586. https://doi.org/10.1128/iai.62.10.4580-4586.1994.

Burton, R. S., R. J. Pereira, and F. S. Barreto. 2013. "Cytonuclear Genomic Interactions and Hybrid Breakdown." *Annual Review of Ecology, Evolution, and Systematics* 44, no. 1: 281–302. https://doi.org/10.1146/annurev-ecolsys-110512-135758.

Bushnell, B. 2014. *BBMap: A Fast, Accurate, Splice-Aware Aligner [Computer Software]*. LBNL-7065E. Lawrence Berkeley National Lab.

Chan, Y. F., M. E. Marks, F. C. Jones, et al. 2010. "Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a *Pitx1* Enhancer." *Science* 327, no. 5963: 302–305. https://doi.org/10.1126/science.1182213.

Charlesworth, B. 2009. "Effective Population Size and Patterns of Molecular Evolution and Variation." *Nature Reviews Genetics* 10, no. 3: 195–205. https://doi.org/10.1038/nrg2526.

Chen, J.-M., D. N. Cooper, N. Chuzhanova, C. Férec, and G. P. Patrinos. 2007. "Gene Conversion: Mechanisms, Evolution and Human Disease." *Nature Reviews Genetics* 8, no. 10: 762–775. https://doi.org/10.1038/nrg2193.

Chen, X.-G., X. Jiang, J. Gu, et al. 2015. "Genome Sequence of the Asian Tiger Mosquito, *Aedes albopictus*, Reveals Insights Into Its Biology, Genetics, and Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 112, no. 44: E5907–E5915. https://doi.org/10.1073/pnas.1516410112.

Choi, J. Y., M. Purugganan, and E. A. Stacy. 2020. "Divergent Selection and Primary Gene Flow Shape Incipient Speciation of a Riparian Tree on Hawaii Island." *Molecular Biology and Evolution* 37, no. 3: 695–710. https://doi.org/10.1093/molbev/msz259.

Christmas, M. J., I. M. Kaplow, D. P. Genereux, et al. 2023. "Evolutionary Constraint and Innovation Across Hundreds of Placental Mammals." *Science* 380, no. 6643: eabn3943. https://doi.org/10.1126/science.abn3943.

Chueca, L. J., T. Schell, and M. Pfenninger. 2021. "Whole-Genome Re-Sequencing Data to Infer Historical Demography and Speciation Processes in Land Snails: The Study of Two *Candidula* Sister Species." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 376, no. 1825: 20200156. https://doi.org/10.1098/rstb.2020.0156.

Coimbra, R. T. F., S. Winter, V. Kumar, et al. 2021. "Whole-Genome Analysis of Giraffe Supports Four Distinct Species." *Current Biology* 31, no. 13: 2929–2938.e5. https://doi.org/10.1016/j.cub.2021.04.033.

Cong, B., J. Liu, and S. D. Tanksley. 2002. "Natural Alleles at a Tomato Fruit Size Quantitative Trait Locus Differ by Heterochronic Regulatory Mutations." *Proceedings of the National Academy of Sciences of the United States of America* 99, no. 21: 13606–13611. https://doi.org/10.1073/pnas.172520999.

Cornetti, L., L. M. Valente, L. T. Dunning, et al. 2015. "The Genome of the "Great Speciator" Provides Insights Into Bird Diversification." *Genome Biology and Evolution* 7, no. 9: 2680–2691. https://doi.org/10.1093/gbe/evv168.

Costantino, L., S. K. Sotiriou, J. K. Rantala, et al. 2014. "Break-Induced Replication Repair of Damaged Forks Induces Genomic Duplications in Human Cells." *Science* 343, no. 6166: 88–91. https://doi.org/10.1126/science.1243211.

Coyne, J., and H. A. Orr. 2004. "Allopatric and Parapatric Speciation." In *Speciation*, 83–124. Sunderland, MA: Oxford University Press.

Császár, N. B. M., P. J. Ralph, R. Frankham, R. Berkelmans, and M. J. H. Van Oppen. 2010. "Estimating the Potential for Adaptation of Corals to Climate Warming." *PLoS One* 5, no. 3: e9751. https://doi.org/10.1371/journal.pone.0009751.

Curtin, A. J., G. R. Zug, and J. R. Spotila. 2009. "Longevity and Growth Strategies of the Desert Tortoise (*Gopherus agassizii*) in Two American Deserts." *Journal of Arid Environments* 73, no. 4–5: 463–471. https://doi.org/10.1016/j.jaridenv.2008.11.011.

Dakup, P., and S. Gaddameedhi. 2017. "Impact of the Circadian Clock on UV-Induced DNA Damage Response and Photocarcinogenesis." *Photochemistry and Photobiology* 93, no. 1: 296–303. https://doi.org/10.1111/php.12662.

Dakup, P. P., K. I. Porter, A. A. Little, et al. 2018. "The Circadian Clock Regulates Cisplatin-Induced Toxicity and Tumor Regression in Melanoma Mouse and Human Models." *Oncotarget* 9, no. 18: 14524–14538. https://doi.org/10.18632/oncotarget.24539.

Darriba, D., D. Posada, A. M. Kozlov, A. Stamatakis, B. Morel, and T. Flouri. 2020. "ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models." *Molecular Biology and Evolution* 37, no. 1: 291–294. https://doi.org/10.1093/molbev/msz189.

Deng, D., P. Sun, C. Yan, et al. 2015. "Molecular Basis of Ligand Recognition and Transport by Glucose Transporters." *Nature* 526, no. 7573: 391–396. https://doi.org/10.1038/nature14655.

Deviche, P., T. Small, P. Sharp, and K. Tsutsui. 2006. "Control of Luteinizing Hormone and Testosterone Secretion in a Flexibly Breeding Male Passerine, the Rufous-Winged Sparrow, *Aimophila carpalis*." *General and Comparative Endocrinology* 149, no. 3: 226–235. https://doi.org/10.1016/j.ygcen.2006.06.004.

Dickinson, V. M., I. M. Schumacher, J. L. Jarchow, T. Duck, and C. R. Schwalbe. 2005. "Mycoplasmosis in Free-Ranging Desert Tortoises in Utah and Arizona." *Journal of Wildlife Diseases* 41, no. 1: 239–243. https://doi.org/10.7589/0090-3558-41.1.239.

Dolby, G. A., R. J. Dorsey, and M. R. Graham. 2019. "A Legacy of Geo-Climatic Complexity and Genetic Divergence Along the Lower Colorado River: Insights From the Geological Record and 33 Desert-Adapted Animals." *Journal of Biogeography* 46, no. 11: 2479–2505. https://doi.org/10.1111/jbi.13685.

Dolby, G. A., M. Morales, T. H. Webster, D. F. DeNardo, M. A. Wilson, and K. Kusumi. 2020. "Discovery of a New TLR Gene and Gene Expansion Event Through Improved Desert Tortoise Genome Assembly With Chromosome-Scale Scaffolds." *Genome Biology and Evolution* 12, no. 2: 3917–3925. https://doi.org/10.1093/gbe/evaa016.

Donkervoort, S., R. Sabouny, P. Yun, et al. 2019. "MSTO1 Mutations Cause mtDNA Depletion, Manifesting as Muscular Dystrophy With Cerebellar Involvement." *Acta Neuropathologica* 138, no. 6: 1013–1031. https://doi.org/10.1007/s00401-019-02059-z.

Edwards, T., K. H. Berry, R. D. Inman, et al. 2015. "Testing Taxon Tenacity of Tortoises: Evidence for a Geographical Selection Gradient at a Secondary Contact Zone." *Ecology and Evolution* 5, no. 10: 2095–2114. https://doi.org/10.1002/ece3.1500.

Edwards, T., M. Tollis, P. Hsieh, et al. 2016. "Assessing Models of Speciation Under Different Biogeographic Scenarios; an Empirical Study Using Multi-Locus and RNA-Seq Analyses." *Ecology and Evolution* 6, no. 2: 379–396. https://doi.org/10.1002/ece3.1865.

Edwards, T., M. Vaughn, P. C. Rosen, et al. 2016. "Shaping Species With Ephemeral Boundaries: The Distribution and Genetic Structure of Desert Tortoise (*Gopherus morafkai*) in the Sonoran Desert Region." *Journal of Biogeography* 43, no. 3: 484–497. https://doi.org/10.1111/jbi.12664.

Ericson, P. G. P., Y. Qu, P. C. Rasmussen, M. P. K. Blom, F. E. Rheindt, and M. Irestedt. 2019. "Genomic Differentiation Tracks Earth-Historic Isolation in an Indo-Australasian Archipelagic Pitta (Pittidae; Aves) Complex." *BMC Evolutionary Biology* 19, no. 1: 151. https://doi.org/10.1186/s12862-019-1481-5.

Eustermann, S., K. Schall, D. Kostrewa, et al. 2018. "Structural Basis for ATP-Dependent Chromatin Remodelling by the INO80 Complex." *Nature* 556, no. 7701: 386–390. https://doi.org/10.1038/s41586-018-0029-y.

Ewels, P., M. Magnusson, S. Lundin, and M. Käller. 2016. "MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report." *Bioinformatics* 32, no. 19: 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

Figueiró, H. V., G. Li, F. J. Trindade, et al. 2017. "Genome-Wide Signatures of Complex Introgression and Adaptive Evolution in the Big Cats." *Science Advances* 3, no. 7: e1700299. https://doi.org/10.1126/sciadv.1700299.

Fondell, J. D. 2013. "The Mediator Complex in Thyroid Hormone Receptor Action." *Biochimica et Biophysica Acta (BBA)—General Subjects* 1830, no. 7: 3867–3875. https://doi.org/10.1016/j.bbagen.2012.02.012.

Franchini, P., P. Xiong, C. Fruciano, et al. 2019. "MicroRNA Gene Regulation in Extremely Young and Parallel Adaptive Radiations of Crater Lake Cichlid Fish." *Molecular Biology and Evolution* 36, no. 11: 2498–2511. https://doi.org/10.1093/molbev/msz168.

Freese, N. H., D. C. Norris, and A. E. Loraine. 2016. "Integrated Genome Browser: Visual Analytics Platform for Genomics." *Bioinformatics* 32, no. 14: 2089–2095. https://doi.org/10.1093/bioinformatics/btw069.

Fumagalli, M., F. G. Vieira, T. S. Korneliussen, et al. 2013. "Quantifying Population Genetic Differentiation From Next-Generation Sequencing Data." *Genetics* 195, no. 3: 979–992. https://doi.org/10.1534/genetics.113.154740.

Fumagalli, M., F. G. Vieira, T. Linderoth, and R. Nielsen. 2014. "ngsTools: Methods for Population Genetics Analyses From Next-Generation Sequencing Data." *Bioinformatics* 30, no. 10: 1486–1487. https://doi.org/10.1093/bioinformatics/btu041.

Gaucher, J., E. Montellier, and P. Sassone-Corsi. 2018. "Molecular Cogs: Interplay Between Circadian Clock and Cell Cycle." *Trends in Cell Biology* 28, no. 5: 368–379. https://doi.org/10.1016/j.tcb.2018.01.006.

Germano, D. J. 1993. "Shell Morphology of North American Tortoises." *American Midland Naturalist* 129, no. 2: 319. https://doi.org/10.2307/2426513.

Germano, D. J. 1994. "Growth and Age at Maturity of North American Tortoises in Relation to Regional Climates." *Canadian Journal of Zoology* 72, no. 5: 918–931. https://doi.org/10.1139/z94-125.

Germano, D. J., B. Bury, T. Esque, T. Fritts, and P. A. Medica. 1994. "Range and Habitats of the Desert Tortoise." *Biology of North American Tortoises* 13: 73–84.

Gov, R. 2015. *Patterns in the Presence of Respiratory Tract Mycoplasmas in Four Species of North American Tortoises*. Reno: University of Nevada.

Guo, X., Q. Hu, G. Hao, et al. 2018. "The Genomes of Two Eutrema Species Provide Insight Into Plant Adaptation to High Altitudes." *DNA Research* 25, no. 3: 307–315. https://doi.org/10.1093/dnares/dsy003.

Hansson, G. C. 2019. "Mucus and Mucins in Diseases of the Intestinal and Respiratory Tracts." *Journal of Internal Medicine* 285, no. 5: 479–490. https://doi.org/10.1111/joim.12910.

Harper, J. W., J. L. Burton, and M. J. Solomon. 2002. "The Anaphase-Promoting Complex: It's Not Just for Mitosis any More." *Genes & Development* 16, no. 17: 2179–2206. https://doi.org/10.1101/gad.1013102.

Hau, M., D. Dominoni, S. Casagrande, et al. 2017. "Timing as a Sexually Selected Trait: The Right Mate at the Right Moment." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 372, no. 1734: 20160249. https://doi.org/10.1098/rstb.2016.0249.

Hazlerigg, D. G., and G. A. Lincoln. 2011. "Hypothesis: Cyclical Histogenesis Is the Basis of Circannual Timing." *Journal of Biological Rhythms* 26, no. 6: 471–485. https://doi.org/10.1177/0748730411420812.

Hebert, F. O., S. Renaut, and L. Bernatchez. 2013. "Targeted Sequence Capture and Resequencing Implies a Predominant Role of Regulatory Regions in the Divergence of a Sympatric Lake Whitefish Species Pair (*Coregonus clupeaformis*)." *Molecular Ecology* 22, no. 19: 4896–4914. https://doi.org/10.1111/mec.12447.

Hoskin, P. J., A. Sibtain, F. M. Daley, and G. D. Wilson. 2003. "GLUT1 and CAIX as Intrinsic Markers of Hypoxia in Bladder Cancer: Relationship With Vascularity and Proliferation as Predictors of Outcome of ARCON." *British Journal of Cancer* 89, no. 7: 1290–1297. https://doi.org/10.1038/sj.bjc.6601260.

Howell, V., J. Huusko, M. Bernal, et al. 2023. "P487: Genetic Findings From Multi-Gene Panel for Primary Ciliary Dyskinesia." *Genetics in Medicine Open* 1, no. 1: 100534. https://doi.org/10.1016/j.gimo.2023.100534.

Huang, Q. M., T. Akashi, Y. Masuda, K. Kamiya, T. Takahashi, and M. Suzuki. 2010. "Roles of POLD4, Smallest Subunit of DNA Polymerase δ, in Nuclear Structures and Genomic Stability of Human Cells." *Biochemical and Biophysical Research Communications* 391, no. 1: 542–546. https://doi.org/10.1016/j.bbrc.2009.11.094.

Hughes, S., J. K. Edwards, A. G. Wilcox, et al. 2021. "*Zfhx3* Modulates Retinal Sensitivity and Circadian Responses to Light." *FASEB Journal* 35, no. 9: e21802. https://doi.org/10.1096/fj.202100563R.

Jones, C. A. n.d. 2008. "*Mycoplasma agassizii* in the Sonoran population of the desert tortoise in Arizona." Tuscon, AZ: The University of Arizona.

Joseph, L., G. J. Adcock, C. Linde, et al. 2009. "A Tangled Tale of Two Teal: Population History of the Grey *Anas gracilis* and Chestnut Teal *A. castanea* of Australia." *Journal of Avian Biology* 40, no. 4: 430–439. https://doi.org/10.1111/j.1600-048X.2008.04652.x.

Kim, S.-T., M.-K. Ye, and S.-H. Shin. 2011. "Effects of Asian Sand Dust on Mucin Gene Expression and Activation of Nasal Polyp Epithelial Cells." *American Journal of Rhinology & Allergy* 25, no. 5: 303–306. https://doi.org/10.2500/ajra.2011.25.3627.

Koch, M. A., and M. Matschinger. 2007. "Evolution and Genetic Differentiation Among Relatives of *Arabidopsis thaliana*." *Proceedings of the National Academy of Sciences of the United States of America* 104, no. 15: 6272–6277. https://doi.org/10.1073/pnas.0701338104.

Kojima, K., M. Ooka, T. Abe, and K. Hirota. 2021. "Pold4, the Fourth Subunit of Replicative Polymerase δ, Suppresses Gene Conversion in the Immunoglobulin-Variable Gene in Avian DT40 Cells." *DNA Repair* 100: 103056. https://doi.org/10.1016/j.dnarep.2021.103056.

Koornneef, M., and B. Scheres. 2001. "*Arabidopsis thaliana* as an Experimental Organism." *Encyclopedia of Life Sciences*: 1–6. https://onlinelibrary.wiley.com/doi/abs/10.1038/npg.els.0002031.

Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. "ANGSD: Analysis of Next Generation Sequencing Data." *BMC Bioinformatics* 15, no. 1: 356. https://doi.org/10.1186/s12859-014-0356-4.

Kumar, V., F. Lammers, T. Bidon, et al. 2017. "The Evolutionary History of Bears Is Characterized by Gene Flow Across Species." *Scientific Reports* 7, no. 1: 46487. https://doi.org/10.1038/srep46487.

Kuraoka, I., S. Ito, T. Wada, et al. 2008. "Isolation of XAB2 Complex Involved in Pre-mRNA Splicing, Transcription, and Transcription-Coupled Repair." *Journal of Biological Chemistry* 283, no. 2: 940–950. https://doi.org/10.1074/jbc.M706647200.

Lamb, T., and C. Lydeard. 1994. "A Molecular Phylogeny of the Gopher Tortoises, With Comments on Familial Relationships Within the Testudinoidea." *Molecular Phylogenetics and Evolution* 3, no. 4: 283–291. https://doi.org/10.1006/mpev.1994.1036.

Leroy, T., Y. Anselmetti, M.-K. Tilak, et al. 2021. "A Bird's White-Eye View on Avian Sex Chromosome Evolution." *Peer Community Journal* 1: e63. https://doi.org/10.24072/pcjournal.70.

Letunic, I., S. Khedkar, and P. Bork. 2021. "SMART: Recent Updates, New Developments and Status in 2020." *Nucleic Acids Research* 49, no. D1: D458–D460. https://doi.org/10.1093/nar/gkaa937.

Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment With Burrows–Wheeler Transform." *Bioinformatics* 25, no. 14: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., and R. Durbin. 2011. "Inference of Human Population History From Individual Whole-Genome Sequences." *Nature* 475, no. 7357: 493–496. https://doi.org/10.1038/nature10231.

Li, H., B. Handsaker, A. Wysoker, et al. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25, no. 16: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Li, K., S. Zhang, X. Song, et al. 2020. "Genome Evolution of Blind Subterranean Mole Rats: Adaptive Peripatric Versus Sympatric Speciation." *Proceedings of the National Academy of Sciences of the United States of America* 117, no. 51: 32499–32508. https://doi.org/10.1073/pnas.2018123117.

Liu, X., R. Fu, Y. Pan, K. F. Meza-Sosa, Z. Zhang, and J. Lieberman. 2018. "PNPT1 Release From Mitochondria during Apoptosis Triggers Decay of Poly(A) RNAs." *Cell* 174, no. 1: 187–201.e12. https://doi.org/10.1016/j.cell.2018.04.017.

Löbrich, M., and P. A. Jeggo. 2007. "The Impact of a Negligent G2/M Checkpoint on Genomic Instability and Cancer Induction." *Nature Reviews Cancer* 7, no. 11: 861–869. https://doi.org/10.1038/nrc2248.

Loker, E. S. 2012. "Macroevolutionary Immunology: A Role for Immunity in the Diversification of Animal Life." *Frontiers in Immunology* 3: 3–25. https://doi.org/10.3389/fimmu.2012.00025.

Lomet, D., J. Cognié, D. Chesneau, E. Dubois, D. Hazlerigg, and H. Dardente. 2018. "The Impact of Thyroid Hormone in Seasonal Breeding Has a Restricted Transcriptional Signature." *Cellular and Molecular Life Sciences* 75, no. 5: 905–919. https://doi.org/10.1007/s00018-017-2667-x.

López-Doménech, G., R. Serrat, S. Mirra, et al. 2012. "The Eutherian Armcx Genes Regulate Mitochondrial Trafficking in Neurons and Interact With Miro and Trak2." *Nature Communications* 3, no. 1: 814. https://doi.org/10.1038/ncomms1829.

Lou, R. N., A. Jacobs, A. P. Wilder, and N. O. Therkildsen. 2021. "A Beginner's Guide to Low-Coverage Whole Genome Sequencing for Population Genomics." *Molecular Ecology* 30, no. 23: 5966–5993. https://doi.org/10.1111/mec.16077.

Mack, K. L., and M. W. Nachman. 2017. "Gene Regulation and Speciation." *Trends in Genetics* 33, no. 1: 68–80. https://doi.org/10.1016/j.tig.2016.11.003.

Maebe, K., I. Meeus, S. Vray, et al. 2016. "A Century of Temporal Stability of Genetic Diversity in Wild Bumblebees." *Scientific Reports* 6, no. 1: 38289. https://doi.org/10.1038/srep38289.

Malmstrøm, M., M. Matschiner, O. K. Tørresen, et al. 2016. "Evolution of the Immune System Influences Speciation Rates in Teleost Fishes." *Nature Genetics* 48, no. 10: 1204–1210. https://doi.org/10.1038/ng.3645.

Mao, Y., E. P. Economo, and N. Satoh. 2018. "The Roles of Introgression and Climate Change in the Rise to Dominance of Acropora Corals." *Current Biology* 28, no. 21: 3373–3382.e5. https://doi.org/10.1016/j.cub.2018.08.061.

Matlin, A. J., F. Clark, and C. W. J. Smith. 2005. "Understanding Alternative Splicing: Towards a Cellular Code." *Nature Reviews Molecular Cell Biology* 6, no. 5: 386–398. https://doi.org/10.1038/nrm1645.

Mazet, O., W. Rodríguez, S. Grusea, S. Boitard, and L. Chikhi. 2016. "On the Importance of Being Structured: Instantaneous Coalescence Rates and Human Evolution—Lessons for Ancestral Population Size Inference?" *Heredity* 116, no. 4: 362–371. https://doi.org/10.1038/hdy.2015.104.

McGeary, S. E., K. S. Lin, C. Y. Shi, et al. 2019. "The Biochemical Basis of MicroRNA Targeting Efficacy." *Science* 366, no. 6472: eaav1741. https://doi.org/10.1126/science.aav1741.

McKenna, A., M. Hanna, E. Banks, et al. 2010. "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data." *Genome Research* 20, no. 9: 1297–1303. https://doi.org/10.1101/gr.107524.110.

McLuckie, A. M., T. Lamb, C. R. Schwalbe, and R. D. McCord. 1999. "Genetic and Morphometric Assessment of an Unusual Tortoise (*Gopherus agassizii*) Population in the Black Mountains of Arizona." *Journal of Herpetology* 33, no. 1: 36. https://doi.org/10.2307/1565541.

Mirastschijski, U., B. Lupše, K. Maedler, et al. 2019. "Matrix Metalloproteinase-3 is Key Effector of TNF-α-Induced Collagen Degradation in Skin." *International Journal of Molecular Sciences* 20, no. 20: 5234. https://doi.org/10.3390/ijms20205234.

Mitchell, R. W., N. H. On, M. R. Del Bigio, D. W. Miller, and G. M. Hatch. 2011. "Fatty Acid Transport Protein Expression in Human Brain and Potential Role in Fatty Acid Transport Across Human Brain Microvessel Endothelial Cells: Fatty Acid Transport Protein Expression in Human Brain." *Journal of Neurochemistry* 117: 735–746. https://doi.org/10.1111/j.1471-4159.2011.07245.x.

Murphy, R., K. Berry, T. Edwards, A. Leviton, A. Lathrop, and J. D. Riedle. 2011. "The Dazed and Confused Identity of Agassiz's Land Tortoise, *Gopherus agassizii* (Testudines: Testudinidae) With the Description of a New Species and Its Consequences for Conservation." *ZooKeys* 113: 39–71. https://doi.org/10.3897/zookeys.113.1353.

Nadachowska-Brzyska, K., C. Li, L. Smeds, G. Zhang, and H. Ellegren. 2015. "Temporal Dynamics of Avian Populations During Pleistocene Revealed by Whole-Genome Sequences." *Current Biology* 25, no. 10: 1375–1380. https://doi.org/10.1016/j.cub.2015.03.047.

Nakatsu, Y., H. Asahina, E. Citterio, et al. 2000. "XAB2, a Novel Tetratricopeptide Repeat Protein Involved in Transcription-Coupled DNA Repair and Transcription." *Journal of Biological Chemistry* 275, no. 45: 34931–34937. https://doi.org/10.1074/jbc.M004936200.

Nobrega, M. A., I. Ovcharenko, V. Afzal, and E. M. Rubin. 2003. "Scanning Human Gene Deserts for Long-Range Enhancers." *Science* 302, no. 5644: 413. https://doi.org/10.1126/science.1088328.

Padra, J. T., H. Sundh, C. Jin, N. G. Karlsson, K. Sundell, and S. K. Lindén. 2014. "*Aeromonas salmonicida* Binds Differentially to Mucins Isolated From Skin and Intestinal Regions of Atlantic Salmon in an *N*-Acetylneuraminic Acid-Dependent Manner." *Infection and Immunity* 82, no. 12: 5235–5245. https://doi.org/10.1128/IAI.01931-14.

Page, A. J., S. Christie, E. Symonds, and H. Li. 2020. "Circadian Regulation of Appetite and Time Restricted Feeding." *Physiology & Behavior* 220: 112873. https://doi.org/10.1016/j.physbeh.2020.112873.

Parsons, M. J., M. Brancaccio, S. Sethi, et al. 2015. "The Regulatory Factor ZFHX3 Modifies Circadian Function in SCN via an AT Motif-Driven Axis." *Cell* 162, no. 3: 607–621. https://doi.org/10.1016/j.cell.2015.06.060.

Payne, C., R. Bovio, D. L. Powell, et al. 2022. "Genomic Insights Into Variation in Thermotolerance Between Hybridizing Swordtail Fishes." *Molecular Ecology* 33, no. 22: e16489. https://doi.org/10.1111/mec.16489.

Pedersen, B. S., and A. R. Quinlan. 2018. "Mosdepth: Quick Coverage Calculation for Genomes and Exomes." *Bioinformatics* 34, no. 5: 867–868. https://doi.org/10.1093/bioinformatics/btx699.

Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81, no. 3: 559–575. https://doi.org/10.1086/519795.

Quinlan, A. R., and I. M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26, no. 6: 841–842. https://doi.org/10.1093/bioinformatics/btq033.

Quintero, I., S. González-Caro, P.-C. Zalamea, and C. D. Cadena. 2014. "Asynchrony of Seasons: Genetic Differentiation Associated With Geographic Variation in Climatic Seasonality and Reproductive Phenology." *American Naturalist* 184, no. 3: 352–363. https://doi.org/10.1086/677261.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. "Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7." *Systematic Biology* 67, no. 5: 901–904. https://doi.org/10.1093/sysbio/syy032.

Raudvere, U., L. Kolberg, I. Kuzmin, et al. 2019. "G:Profiler: A Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update)." *Nucleic Acids Research* 47, no. W1: W191–W198. https://doi.org/10.1093/nar/gkz369.

Ronquist, F., M. Teslenko, P. Van Der Mark, et al. 2012. "MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space." *Systematic Biology* 61, no. 3: 539–542. https://doi.org/10.1093/sysbio/sys029.

Rose, A. B. 2019. "Introns as Gene Regulators: A Brick on the Accelerator." *Frontiers in Genetics* 9: 672. https://doi.org/10.3389/fgene.2018.00672.

Rousseau, K., C. Byrne, Y. S. Kim, J. R. Gum, D. M. Swallow, and N. W. Toribara. 2004. "The Complete Genomic Organization of the Human MUC6 and MUC2 Mucin Genes." *Genomics* 83, no. 5: 936–939. https://doi.org/10.1016/j.ygeno.2003.11.003.

Ruby, D. E., and H. A. Niblick. 1994. "A Behavioral Inventory of the Desert Tortoise: Development of an Ethogram." *Herpetological Monographs* 8: 88. https://doi.org/10.2307/1467073.

Sadd, B. M., S. M. Barribeau, G. Bloch, et al. 2015. "The Genomes of Two Key Bumblebee Species With Primitive Eusocial Organization." *Genome Biology* 16, no. 1: 76. https://doi.org/10.1186/s13059-015-0623-3.

Sandmeier, F. C., C. R. Tracy, S. duPré, and K. Hunter. 2009. "Upper Respiratory Tract Disease (URTD) as a Threat to Desert Tortoise Populations: A Reevaluation." *Biological Conservation* 142, no. 7: 1255–1268. https://doi.org/10.1016/j.biocon.2009.02.001.

Santos, H., J. Rousselet, E. Magnoux, M.-R. Paiva, M. Branco, and C. Kerdelhué. 2007. "Genetic Isolation Through Time: Allochronic Differentiation of a Phenologically Atypical Population of the Pine Processionary Moth." *Proceedings of the Royal Society B: Biological Sciences* 274, no. 1612: 935–941. https://doi.org/10.1098/rspb.2006.3767.

Schade, C., G. Flemström, and L. Holm. 1994. "Hydrogen Ion Concentration in the Mucus Layer on Top of Acid-Stimulated and -Inhibited Rat Gastric Mucosa." *Gastroenterology* 107, no. 1: 180–188. https://doi.org/10.1016/0016-5085(94)90075-2.

Schmitz, U., N. Pinello, F. Jia, et al. 2017. "Intron Retention Enhances Gene Regulatory Complexity in Vertebrates." *Genome Biology* 18, no. 1: 216. https://doi.org/10.1186/s13059-017-1339-3.

Scott, P. A., L. J. Allison, K. J. Field, R. C. Averill-Murray, and H. B. Shaffer. 2020. "Individual Heterozygosity Predicts Translocation Success in Threatened Desert Tortoises." *Science* 370, no. 6520: 1086–1089. https://doi.org/10.1126/science.abb0421.

Shanas, U., G. Heth, E. Nevo, R. Shalgi, and J. Terkel. 1995. "Reproductive Behaviour in the Female Blind Mole Rat (*Spalax ehrenbergi*)." *Journal of Zoology* 237, no. 2: 195–210. https://doi.org/10.1111/j.1469-7998.1995.tb02758.x.

Sievers, F., A. Wilm, D. Dineen, et al. 2011. "Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega." *Molecular Systems Biology* 7, no. 1: 539. https://doi.org/10.1038/msb.2011.75.

Small, T. W., P. J. Sharp, and P. Deviche. 2007. "Environmental Regulation of the Reproductive System in a Flexibly Breeding Sonoran Desert Bird, the Rufous-Winged Sparrow, *Aimophila carpalis*." *Hormones and Behavior* 51, no. 4: 483–495. https://doi.org/10.1016/j.yhbeh.2007.01.004.

Smith, C. I., L. C. Sweet, J. Yoder, M. R. McKain, K. Heyduk, and C. Barrows. 2023. "Dust Storms Ahead: Climate Change, Green Energy Development and Endangered Species in the Mojave Desert." *Biological Conservation* 277: 109819. https://doi.org/10.1016/j.biocon.2022.109819.

Sobel, J. M., G. F. Chen, L. R. Watt, and D. W. Schemske. 2010. "The Biology of Speciation." *Evolution* 64, no. 2: 295–315. https://doi.org/10.1111/j.1558-5646.2009.00877.x.

Sowilem, M. M., H. A. Kamal, and E. I. Khater. 2013. "Life Table Characteristics of *Aedes aegypti* (Diptera: Culicidae) From Saudi Arabia." *Tropical Biomedicine* 30, no. 2: 301–314.

Sullivan, B. K., R. Averill-Murray, K. O. Sullivan, J. R. Sullivan, E. A. Sullivan, and J. D. Riedle. 2014. "Winter Activity of the Sonoran Desert Tortoise (*Gopherus morafkai*) in Central Arizona." *Chelonian Conservation and Biology* 13, no. 1: 114–119. https://doi.org/10.2744/CCB-1056.1.

Sur, I., and J. Taipale. 2016. "The Role of Enhancers in Cancer." *Nature Reviews Cancer* 16, no. 8: 483–493. https://doi.org/10.1038/nrc.2016.62.

Sveen, L. R., F. T. Grammes, E. Ytteborg, H. Takle, and S. M. Jørgensen. 2017. "Genome-Wide Analysis of Atlantic Salmon (*Salmo salar*) Mucin Genes and Their Role as Biomarkers." *PLoS One* 12, no. 12: e0189103. https://doi.org/10.1371/journal.pone.0189103.

Szklarczyk, D., R. Kirsch, M. Koutrouli, et al. 2023. "The STRING Database in 2023: Protein–Protein Association Networks and Functional Enrichment Analyses for any Sequenced Genome of Interest." *Nucleic Acids Research* 51, no. D1: D638–D646. https://doi.org/10.1093/nar/gkac1000.

Tahara, Y., and S. Shibata. 2014. "Chrono-Biology, Chrono-Pharmacology, and Chrono-Nutrition." *Journal of Pharmacological Sciences* 124, no. 3: 320–335. https://doi.org/10.1254/jphs.13R06CR.

Taylor, R. S., and V. L. Friesen. 2017. "The Role of Allochrony in Speciation." *Molecular Ecology* 26, no. 13: 3330–3342. https://doi.org/10.1111/mec.14126.

Terhorst, J., J. A. Kamm, and Y. S. Song. 2017. "Robust and Scalable Inference of Population History From Hundreds of Unphased Whole Genomes." *Nature Genetics* 49, no. 2: 303–309. https://doi.org/10.1038/ng.3748.

The ENCODE Project Consortium. 2012. "An Integrated Encyclopedia of DNA Elements in the Human Genome." *Nature* 489, no. 7414: 57–74. https://doi.org/10.1038/nature11247.

Thomassen, H. A., A. H. Freedman, D. M. Brown, W. Buermann, and D. K. Jacobs. 2013. "Regional Differences in Seasonal Timing of Rainfall Discriminate Between Genetically Distinct East African Giraffe Taxa." *PLoS One* 8, no. 10: e77191. https://doi.org/10.1371/journal.pone.0077191.

Tollis, M., D. F. DeNardo, J. A. Cornelius, et al. 2017. "The Agassiz's Desert Tortoise Genome Provides a Resource for the Conservation of a Threatened Species." *PLoS One* 12, no. 5: e0177708. https://doi.org/10.1371/journal.pone.0177708.

Tollis, M., E. D. Hutchins, J. Stapley, et al. 2018. "Comparative Genomics Reveals Accelerated Evolution in Conserved Pathways During the Diversification of Anole Lizards." *Genome Biology and Evolution* 10, no. 2: 489–506. https://doi.org/10.1093/gbe/evy013.

US Fish and Wildlife Service. 1994. *Desert Tortoise (Mojave Population) Recovery Plan*. Portland, Oregon: U.S. Fish and Wildlife Service.

USFWS. 2010. *Mojave Population of the Desert Tortoise (Gopherus agassizii) 5-Year Review: Summary and Evaluation*, 1–123. U.S. Las Vegas, Nevada: Fish and Wildlife Service Desert Tortoise Recovery Office.

USFWS. 2022. "Endangered and Threatened Wildlife and Plants; 12-Month Finding for the Sonoran Desert Tortoise." *Federal Registrar* 87, no. 26: 7077–7079.

Wallis, I. R., B. T. Henen, and K. A. Nagy. 1999. "Egg Size and Annual Egg Production by Female Desert Tortoises (*Gopherus agassizii*): The Importance of Food Abundance, Body Size, and Date of Egg Shelling." *Journal of Herpetology* 33, no. 3: 394. https://doi.org/10.2307/1565636.

Wang, J., N. R. Street, E. Park, J. Liu, and P. K. Ingvarsson. 2020. "Evidence for Widespread Selection in Shaping the Genomic Landscape During Speciation of *Populus*." *Molecular Ecology* 29, no. 6: 1120–1136. https://doi.org/10.1111/mec.15388.

Weiss, J. L., and J. T. Overpeck. 2005. "Is the Sonoran Desert Losing Its Cool?" *Global Change Biology* 11, no. 12: 2065–2077. https://doi.org/10.1111/j.1365-2486.2005.01020.x.

Weissing, F. J., P. Edelaar, and G. S. Van Doorn. 2011. "Adaptive Speciation Theory: A Conceptual Review." *Behavioral Ecology and Sociobiology* 65, no. 3: 461–480. https://doi.org/10.1007/s00265-010-1125-7.

Wickham, H. 2011. "Ggplot2." *WIREs Computational Statistics* 3, no. 2: 180–185. https://doi.org/10.1002/wics.147.

Wickham, H., M. Averick, J. Bryan, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4, no. 43: 1686. https://doi.org/10.21105/joss.01686.

Wood, S., and A. Loudon. 2014. "Clocks for All Seasons: Unwinding the Roles and Mechanisms of Circadian and Interval Timers in the Hypothalamus and Pituitary." *Journal of Endocrinology* 222, no. 2: R39–R59. https://doi.org/10.1530/JOE-14-0141.

Yan, J., and A. Goldbeter. 2019. "Robust Synchronization of the Cell Cycle and the Circadian Clock Through Bidirectional Coupling." *Journal of the Royal Society Interface* 16, no. 158: 20190376. https://doi.org/10.1098/rsif.2019.0376.

Zhang, L., R. Rahbari, M. He, and E. Kebebew. 2011. "CDC23 Regulates Cancer Cell Phenotype and Is Overexpressed in Papillary Thyroid Cancer." *Endocrine-Related Cancer* 18, no. 6: 731–742. https://doi.org/10.1530/ERC-11-0181.

Zhao, Y., L. Gan, L. Ren, Y. Lin, C. Ma, and X. Lin. 2022. "Factors Influencing the Blood-Brain Barrier Permeability." *Brain Research* 1788: 147937. https://doi.org/10.1016/j.brainres.2022.147937.

Zhou, Q., J. Li, W. Yue, et al. 2020. "Cell Division Cycle 23 Is Required for Mouse Oocyte Meiotic Maturation." *FASEB Journal* 34, no. 7: 8990–9002. https://doi.org/10.1096/fj.202000131R.

Zhu, Q., and W. J. Belden. 2020. "Molecular Regulation of Circadian Chromatin." *Journal of Molecular Biology* 432, no. 12: 3466–3482. https://doi.org/10.1016/j.jmb.2020.01.009.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.