

Current Biology

Nectar-feeding bats and birds show parallel molecular adaptations in sugar metabolism enzymes

Highlights

- Nectar bats and birds show parallel positive selection in sugar metabolism genes
- Metabolomic profiling supports elevated respiratory flux in nectar bats
- Glycolytic enzymes of nectar bats show elevated catalytic activity
- Adaptations in sugar metabolism may enable hovering and avoidance of glucose toxicity

Authors

Joshua H.T. Potter, Rosie Drinkwater, Kalina T.J. Davies, ..., Dunja Aksentijevic, Liliana M. Dávalos, Stephen J. Rossiter

Correspondence

jhtpotter@gmail.com (J.H.T.P.),
s.j.rossiter@qmul.ac.uk (S.J.R.)

In brief

Potter et al. show that nectar-feeding bats and hummingbirds have undergone parallel positive selection in enzymes controlling sugar metabolism. Metabolite and enzyme data also support high respiratory flux in nectar-feeding bats and help explain how these specialists can access and survive on a unique energy-rich food source among vertebrates.

Report

Nectar-feeding bats and birds show parallel molecular adaptations in sugar metabolism enzymes

Joshua H.T. Potter,^{1,*} Rosie Drinkwater,¹ Kalina T.J. Davies,¹ Nicolas Nesi,¹ Marisa C.W. Lim,² Laurel R. Yohe,^{2,3} Hai Chi,⁴ Xiaoqing Zhang,^{4,5} Ilya Levantis,¹ Burton K. Lim,⁶ Christopher C. Witt,⁷ Georgia Tsagkogeorga,¹ Mario dos Reis,¹ Yang Liu,⁴ William Furey,⁸ Matthew J. Whitley,⁸ Dunja Aksentijevic,⁹ Liliana M. Dávalos,^{2,10} and Stephen J. Rossiter^{1,11,*}

¹School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, UK

²Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794, USA

³Department of Geology & Geophysics, Yale University, 210 Whitney Avenue, New Haven, CT 06511, USA

⁴College of Life Sciences, Shaanxi Normal University, Xi'an 710119, China

⁵College of Animal Science and Veterinary Medicine, Shenyang Agricultural University, Shenyang 110866, China

⁶Department of Natural History, Royal Ontario Museum, Toronto, ON, Canada

⁷Museum of Southwestern Biology and Department of Biology, University of New Mexico, Albuquerque, NM, USA

⁸Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA

⁹William Harvey Research Institute, Centre for Biochemical Pharmacology, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK

¹⁰Consortium for Inter-Disciplinary Environmental Research, Stony Brook University, Stony Brook, NY 11794, USA

¹¹Lead contact

*Correspondence: jhtpotter@gmail.com (J.H.T.P.), s.j.rossiter@qmul.ac.uk (S.J.R.)

<https://doi.org/10.1016/j.cub.2021.08.018>

SUMMARY

In most vertebrates, the demand for glucose as the primary substrate for cellular respiration is met by the breakdown of complex carbohydrates, or energy is obtained by protein and lipid catabolism. In contrast, a few bat and bird species have convergently evolved to subsist on nectar, a sugar-rich mixture of glucose, fructose, and sucrose.^{1–4} How these nectar-feeders have adapted to cope with life-long high sugar intake while avoiding the onset of metabolic syndrome and diabetes^{5–7} is not understood. We analyzed gene sequences obtained from 127 taxa, including 22 nectar-feeding bat and bird genera that collectively encompass four independent origins of nectarivory. We show these divergent taxa have undergone pervasive molecular adaptation in sugar catabolism pathways, including parallel selection in key glycolytic and fructolytic enzymes. We also uncover convergent amino acid substitutions in the otherwise evolutionarily conserved aldolase B (ALDOB), which catalyzes rate-limiting steps in fructolysis and glycolysis, and the mitochondrial gatekeeper pyruvate dehydrogenase (PDH), which links glycolysis and the tricarboxylic acid cycle. Metabolic profile and enzyme functional assays are consistent with increased respiratory flux in nectar-feeding bats and help explain how these taxa can both sustain hovering flight and efficiently clear simple sugars. Taken together, our results indicate that nectar-feeding bats and birds have undergone metabolic adaptations that have enabled them to exploit a unique energy-rich dietary niche among vertebrates.

RESULTS AND DISCUSSION

To determine whether the transition to nectar feeding in vertebrates has involved molecular adaptations for processing and regulating large quantities of simple sugars, we analyzed sequence data for over 13,500 orthologous coding genes across 22 genera of nectar-feeding bats and birds, representing four independent origins of this dietary specialization. All of these taxa share behavioral and phenotypic traits for accessing the nectar of tubular flowers, including hovering flight, elongated rostra, and hyper-extensible tongues.^{8–10} To expand the taxonomic background for comparative analyses, we combined these with sequence data for 32 non-nectarivorous bat genera and supplemented these with published orthologs from a further six bat genera and 28 bird genera. Our final dataset for bats consisted of 49 genera (69 species), including 10 nectar-feeding

genera from the divergent neotropical bat subfamilies Glossophaginae and Lonchophyllinae (family Phyllostomidae) and one (*Eonycteris*) from the Pteropodidae (Old World fruit bats) (for full species list, see [Data S1A](#)). Our bird dataset of 39 genera (41 species) contained 11 genera of hummingbirds (Trochilidae) ([Data S1B](#)).

We screened for episodic positive selection at each of the four origins of nectar feeding by performing branch-site and clade models on all orthologous gene alignments using published species topologies ([Figure S1](#); [STAR Methods](#)). Analyses of bats and birds were conducted separately with bat- or bird-only alignments, in each case examining one focal lineage while removing other nectar feeders from the background set of taxa. Our results yielded 219 candidate positively selected genes (PSGs; $p < 0.05$) in Glossophaginae (branch-site, 144; clade, 77; [Data S2A](#) and [S2B](#)), which included branch-site loci reported previously;¹¹ 269

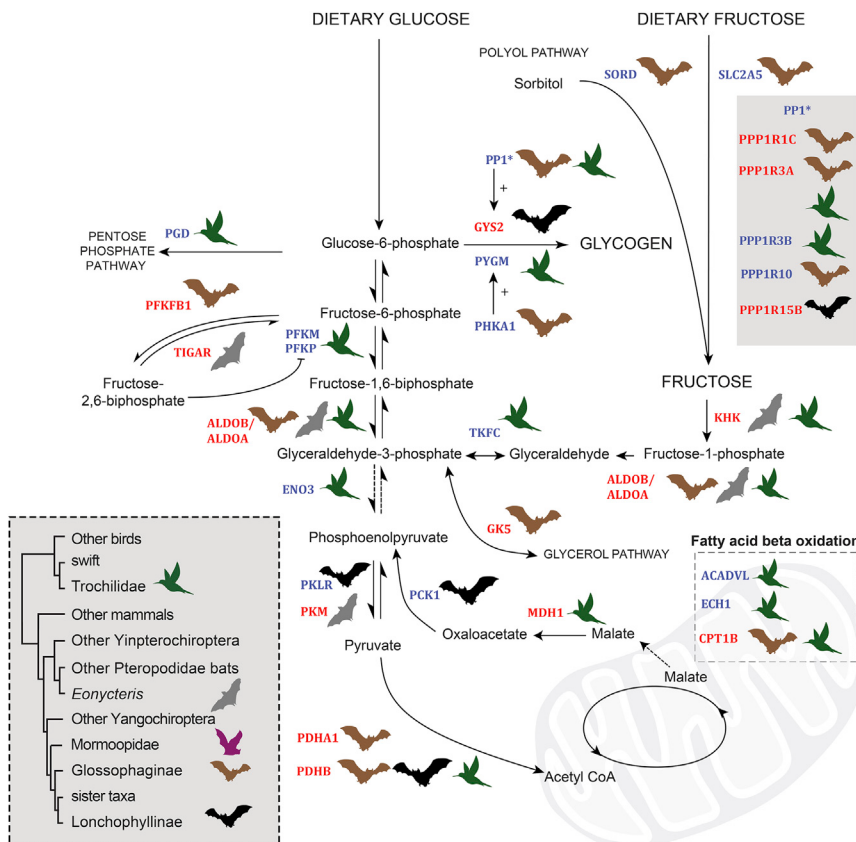


Figure 1. Glycolysis and associated pathways showing the positions of genes under positive selection in four lineages of nectar-feeders

(A) Thirty-three major glycolysis, fructolysis, glycogenesis, TCA, or peripheral metabolism genes showed positive selection in at least one lineage of nectar-feeding bat or bird, with five genes showing selection in more than one lineage. Icons refer to lineages shown in (B). Gene names in blue were significant at the standard threshold ($p < 0.05$) and those in red are significant after robust FDR ($q < 0.05$). None of these key metabolic PSGs was under selection in the control lineage.

(B) Simplified phylogeny including three focal lineages of nectar-feeding bats (Glossophaginae, Lonchophyllinae, and *Eonycteris*), the nectar-feeding hummingbird lineage (Trochilidae), and a control lineage of insectivorous bats (Mormoopidae). For details of species, see [Data S1A](#) and [S1B](#).

in Lonchophyllinae (branch-site, 239; clade, 33; [Data S2C](#) and [S2D](#)); 141 in *Eonycteris* (branch-site only; [Data S2E](#)); and 661 in Trochilidae (branch-site, 622; clade, 45; [Data S2G](#) and [S2H](#)).

The combined set of 1,246 PSGs contained 33 loci strongly associated with the derivation of energy from sugar, either directly in glycolysis, fructolysis, or peripheral pathways, or from fatty acid metabolism ([Figure 1](#)). Of these PSGs, approximately one-third were also significant with robust false discovery rate (FDR) correction.¹² In contrast, tests performed on the control lineage of insectivorous bats, Mormoopidae, revealed that none of these 33 loci were under positive selection ([Data S2F](#)). Finally, because the phylogeny of birds has undergone revision in recent years, we repeated branch-site and clade models of selection for an alternative bird topology in which hummingbirds were placed in a basal position within the Neoaves (see [STAR Methods](#) for details) and found highly similar numbers of PSGs (branch-site, 591; clade, 54; [Data S2I](#) and [S2J](#)). Under this earlier proposed topology, the set of sugar-metabolism genes recovered was very similar, although *ALDOA*, *ENOL3*, *MDH2*, and *PDHB* were not recovered ([Figure 1](#); [Data S2I](#) and [S2J](#)).

Although a number of other dietary genes have been identified as under positive selection in bats and other mammals, pervasive molecular adaptation in sugar metabolism genes has not previously been described, likely reflecting the more limited taxon sampling of earlier studies and the lack of species representing multiple independent origins of nectar feeding.^{13,14}

For each of the four sets of PSGs from the nectarivorous lineages, we conducted functional enrichment based on biological

process gene ontology (GO) and found strong evidence of parallel enrichment associated with carbohydrate metabolism ([Data S3A–S3E](#)). The same term, “fructose catabolic process to hydroxyacetone phosphate and glyceraldehyde-3-phosphate” (GO:0061624), was ranked third of 10 significantly enriched ($p < 0.01$) GO terms in *Eonycteris* and second of 16 significantly enriched GO terms in hummingbirds.

In Lonchophyllinae, the second most significant GO term of 18 was “detection of chemical stimulus involved in sensory perception of sweet taste” (GO:0001582), due to selection in the genes *TAS1R2* and *TAS1R3*, which encode taste receptors that detect fructose and sucrose.

In total, 49 PSGs were shared between more than one nectarivorous lineage. A functional enrichment test of these loci revealed fructose catabolism (GO:0061624) to be most significantly enriched ($p < 0.001$), pertaining to the genes aldolase B (*ALDOB*) and fructokinase (*KHK*). In addition to these, a further two multi-lineage PSGs are also implicated in sugar metabolism (*PDHB* and *PPP1R3A*), and two in fatty acid beta-oxidation (*CPT1B* and *ABCD3*). *PDHB*, under selection in divergent nectar-feeding phyllostomid bats and hummingbirds, and *PDHA1*, selected in Glossophaginae, encode the beta and alpha subunits of the enzyme pyruvate dehydrogenase (PDH). In aerobic conditions, the endpoint of glycolysis is pyruvate. By converting pyruvate to acetyl CoA, PDH links glycolysis and oxidative phosphorylation via the TCA (tricarboxylic acid) cycle and catalyzes a rate-limiting step in the aerobic production of ATP from glucose.¹⁵ Aldolase B, encoded by *ALDOB*, was found to be under positive selection in both Glossophaginae and *Eonycteris* and is a rate-limiting enzyme that reversibly catalyzes the central hydrolysis stage of both glycolysis and fructolysis.¹⁶ Similarly, fructokinase (*KHK*) was under positive selection in *Eonycteris* and hummingbirds and is responsible for the first step of ingested fructose metabolism. Together, *ALDOB* and *KHK* are critical for metabolizing fructose in the liver, which in

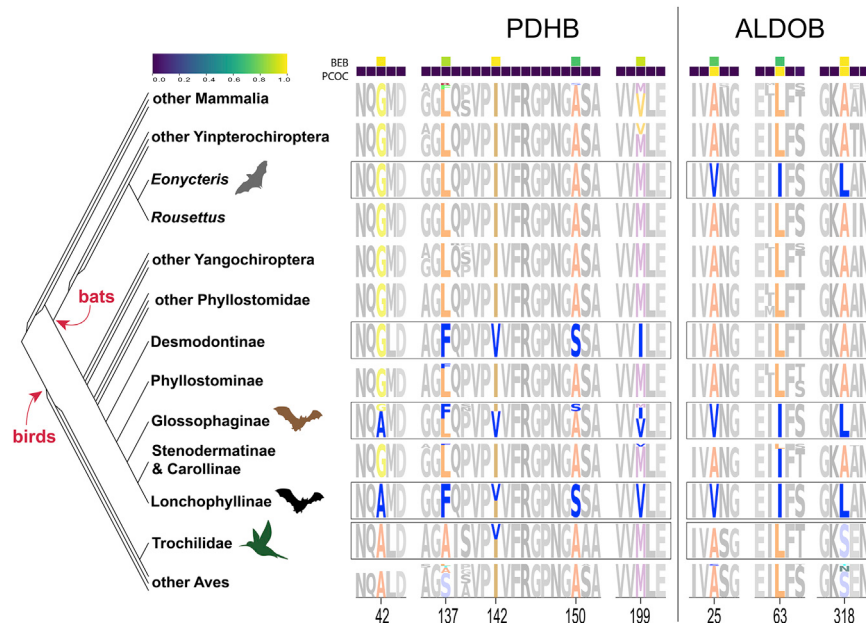


Figure 2. Examples of two glycolytic genes (PDHB and ALDOB) under positive selection in multiple lineages of nectar-feeders

Manually curated set of sites with Bayes empirical Bayes (BEB) posterior probabilities >0.5 in at least one lineage of nectar-feeders, overlaid on PCOC posterior probabilities of convergence based on a four-way comparison of nectar-feeding clades, shown for two key glycolytic enzymes. For each site, the height of the amino acid residue code is proportional to its frequency in the group. Residues inferred to have undergone convergence since the corresponding ancestral state are shown in blue. Examples of convergent sites in nectar-feeding bats include L137F, I142V, and M199V in PDHB and A25V, L63I, and A318L in ALDOB.

(A320T and S505N A516V), a gene implicated in glucose homeostasis; *NT5C1A* (A150S); and *PDHB* (G42A, I142V, and M199V). In this latter gene, one replacement (G42A) was not observed in any other mammalian lineage (Figure 2), while

nectar-feeders will accumulate both through direct dietary input and from the conversion of excess nectar-derived glucose. Indeed, nectar comprises approximately equal proportions of glucose and fructose.¹⁷ As such, the detected parallel positive selection in these key enzymes across multiple nectar feeding lineages reveals how these taxa have evolved to subsist on sugar-rich diets while avoiding glucose- and fructose-induced metabolic syndrome.¹⁸ Furthermore, our finding that the sweet taste receptor gene *TAS1R3* showed positive selection in three nectar-feeding lineages (the two phyllostomid bat lineages, Glossophaginae and Lonchophyllinae, and hummingbirds) highlights a possible adaptive significance of sweet taste perception and contrasts strongly with findings from the common vampire bat (also a member of Phyllostomidae), which has been shown to have lost this gene.¹⁹ Further losses have also been documented in other taste receptor genes in blood and fruit-feeding phyllostomids, including the bitter taste receptors, *TAS2R*, highlighting the evolution of sensory perception linked to dietary diversification.^{20,21}

Since parallel molecular adaptation might arise from selection acting on the same or different residues, we examined the inferred sites under positive selection in our PSGs, highlighted by Bayes empirical Bayes (BEB) posterior probabilities. Of the 49 loci under positive selection in two or more nectar-feeding lineages, we found that eight loci shared between one and eight positively selected BEB amino acid positions (*ABCD3*, *ALDOB*, *CHD1L*, *GLIPR1*, *GPR162*, *MAP6D1*, *NT5C1A*, and *PDHB*). We checked these gene alignments and noted remarkable convergence between *Eonycteris* and at least one member of the Glossophaginae in both *ABCD3* (I256V) and *ALDOB* (A25V and A318L), with the latter case involving identical substitutions in all three nectarivorous bat lineages against a fully conserved ancestral state in other mammals (Figure 2). We also confirmed that, in relation to their common ancestor, identical convergent replacements have occurred between at least one member of each of the Lonchophyllinae and Glossophaginae in *GPR162*

two substitutions (I142V and A150S) were also seen in only one other species, the common vampire bat, *Desmodus rotundus*, noteworthy in the context of sugar metabolism because this taxon is an obligate blood-feeder and therefore ingests carbohydrates solely in the form of glucose.

Positive selection with shared nonsynonymous sites and substitutions in multiple nectar-feeding clades is strong evidence for parallel molecular adaptation. We also ran explicit tests of molecular convergence using PCOC, which tests whether substitutions in *a priori* convergent lineages converge on similar physicochemical properties.²² For the enzymes ALDOB and PDHB, we tested all nectarivorous lineages simultaneously (four-way) and pairwise against an expanded background of birds and mammals (results from other PSGs not shown). In *ALDOB*, PCOC confirmed that the observed parallel positively selected substitutions in nectar bats are convergent (posterior probability [PP] > 0.99) and also identified three additional putative convergent sites (Figure 2). In *PDHB*, four-way PCOC analyses did not detect high PP of convergence for any residues by nectar-feeder; however, pairwise comparisons of Lonchophyllinae and Glossophaginae revealed strong convergence for A135G and M199V (PP > 0.9).

Other PSGs that we identified in single focal lineages of nectar-feeders encode proteins with additional roles in carbohydrate metabolism, including glycolysis, glucose homeostasis, and gluconeogenesis. These include enzymes that regulate or catalyze rate-limiting steps in glycolysis (e.g., phosphofructokinase 1, PKLR) or gluconeogenesis (e.g., phosphoenolpyruvate carboxykinase 1, PCK1), as well as the hydrolysis of dietary sucrose by sucrase isomaltase (SI) in the intestine. To avoid the risks of long-term glucose toxicity, nectar-feeders must either rapidly process ingested sugars or otherwise be able to tolerate elevated blood sugar. We found that key proteins underlying glycogen synthesis and regulation were under positive selection in nectar-feeding phyllostomids, including glycogen synthase 2 (GYS2) that catalyzes glycogen synthesis in the liver. Glycogen

may be especially important for these taxa as a sink for excess sugar when feeding, as well as a fuel source when fasting between feeding bouts.²³ Our results from GYS2 advance previous reports of parallel substitutions between fruit-eating species.^{24,25} Finally, in glossophagine bats, we found molecular adaptation in *SORD*, which encodes sorbitol dehydrogenase (SDH). This protein functions to clear excess dietary glucose by conversion to sorbitol and then fructose via the polyol pathway. It is thus possible that nonsynonymous changes in *SORD* increase SDH activity, thereby leading to clearance of sorbitol, which when left to accumulate causes nerve damage in diabetic humans.^{26,27}

We assessed the potential impact of the observed convergent changes in ALDOB on the protein structure by mapping substitutions onto a homology model based on the human enzyme. We found that the three replacements in aldolase B, which are shared by all three focal lineages of nectar-feeding bats (A25V, L63I, and A318L), are not located directly at this enzyme's active site, nor in the area of protein-protein interface within the homotetramer. Thus, any impact of these sites might either be subtle or the result of interplay between these residues and the other derived changes, although conservative substitutions (e.g., L→I) have been shown to have major structural and functional consequences.^{28,29} Alternatively, the homology model for bat aldolase B may not accurately reflect the true structure of the enzyme, concealing the effects of the identified mutations. We repeated this approach for the alpha and beta subunits of PDH (*PDHA1* and *PDHB*) and found that all positively selected substitutions in the heterotetramer are located away from the coenzyme thiamin diphosphate (ThDP) and magnesium ion, including three residues on or near the protein surface (α A136T, β L137F, and β A150S), implying no direct impact on the main catalytic centers. Three sites, however, were located near the potassium ion (β G42A, β L137F, and β I142V), with one of these in the same position as a known pathological substitution (β I142M) linked to PDH deficiency in humans.³⁰ An alternative explanation is that the molecular adaptations seen in nectar-feeders affect enzymatic activity by altering the interactions between PDH and its regulatory kinases and phosphatases, perhaps shifting the overall equilibrium to a higher catalytic activity.

To test whether the putative adaptive substitutions in ALDOB and PDH in nectar-feeding bats are associated with altered carbohydrate metabolism, we performed assays of enzyme activity. For the rate-limiting homotetramer ALDOB, because activity is measured indirectly by recording NADH consumption of downstream enzymes, we used a bacterial expression system. We compared the activity of purified *Anoura* ALDOB containing the three convergent substitutions to a triple mutant protein in which these were replaced by their respective ancestral states (V25A, I63L, and L318A). We recorded a modest but significant higher activity level for the wild-type (*Anoura*) enzyme ($p < 0.001$) (Figure 3A). Further work is needed to reproduce this result in an *in vivo* system in which metabolic demand can also be manipulated. For PDH, for which activity is directly measured, and for which *in vitro* reconstitution of the multi-enzyme complex is not trivial for recombinant proteins, comparisons of nectarivorous and other bats were performed on endogenous pectoral muscle. We recorded on average higher activity in the nectar-feeders than in other feeding groups, representing ancestral forms, again

suggestive of enhanced oxidative glucose metabolism (Figure 3B), although further work is needed to rule out differences in protein expression. We also modeled PDH activity as a function of diet while correcting for phylogenetic non-independence and found that the phylogenetic mean of the nectarivores was again highest, followed by that of frugivores and then insectivores, although these differences were not significant, likely reflecting the inherent lack of statistical power in the dual origin of nectarivory in phyllostomids (Figure S2A).

We also performed unbiased metabolomic profiling in muscle and liver tissue collected from wild-caught bats. We found that muscle was characterized by markedly reduced (19-fold) lactate in nectar-feeders compared to other groups, consistent with decreased anaerobic glycolysis. At the same time, a low succinate/fumarate ratio in muscle indicates high TCA cycle/oxidative phosphorylation flux, while high glycogen is consistent with efficient uptake and deposition of glucose (Figure 3C). These trends, which were also supported by comparisons of phylogenetic means (Figure S2B), suggest rapid shunting of intracellular free glucose into glycogen, possibly as a means of protecting cellular proteins from glycosylation and the effects of glucotoxicity, comparable to the shunting of fatty acid metabolites into triglycerides for protection against the effects of lipotoxicity.³¹ Despite these changes in metabolic profile, we found no evidence of energetic deficit in any group, with comparable phosphocreatine/ATP ratios (Figure 3D).

Further insights come from the liver metabolomes, in which glycogen and lactate concentrations are suggestive of enhanced glucose availability, deposition, and utilization in nectar-feeding bats (Figures 3E and S2C). Of particular note, these bats also showed elevated fructose concentration in the liver, adding weight to the likely adaptive significance of the observed parallel amino acid substitutions in both ALDOB and KHK, enzymes that together convert fructose in the liver to glyceraldehyde for entry into glycolysis. Finally, in both muscle and liver we also detected elevated acetate (Figures 3C, 3E, S2B, and S2C), likely reflecting the conversion of glucose-derived pyruvate under conditions of nutritional excess and thus also supporting a state of hyperactive glucose metabolism.³² Taken together, therefore, these metabolome profiles are suggestive of increased carbohydrate metabolic load in both muscle and liver tissues, which play key roles in energy utilization and regulation, respectively.

Although the high energetic demands of nectar feeding are long established, the specific cellular mechanisms through which these are met are largely unknown. Stable isotope studies have shown that, during foraging bouts, nectarivorous bats and birds initially obtain energy by oxidizing their fat reserves before progressively shifting to newly obtained dietary sugar,^{2–4} and it is suggested that hovering may help to regulate blood sugar levels.^{2,3,33} Our results help to explain how these organisms are able to both rapidly deplete surges in simple sugars circulating after feeding and release energy necessary to maintain costly hovering flight.³⁴ Indeed, since hovering is essential for accessing flowers, some of the adaptations relating to glucose metabolism uncovered in this study may have their roots in the evolution of hovering, and these changes subsequently provided the metabolic conditions necessary for a full transition to an obligate sugar-rich diet.

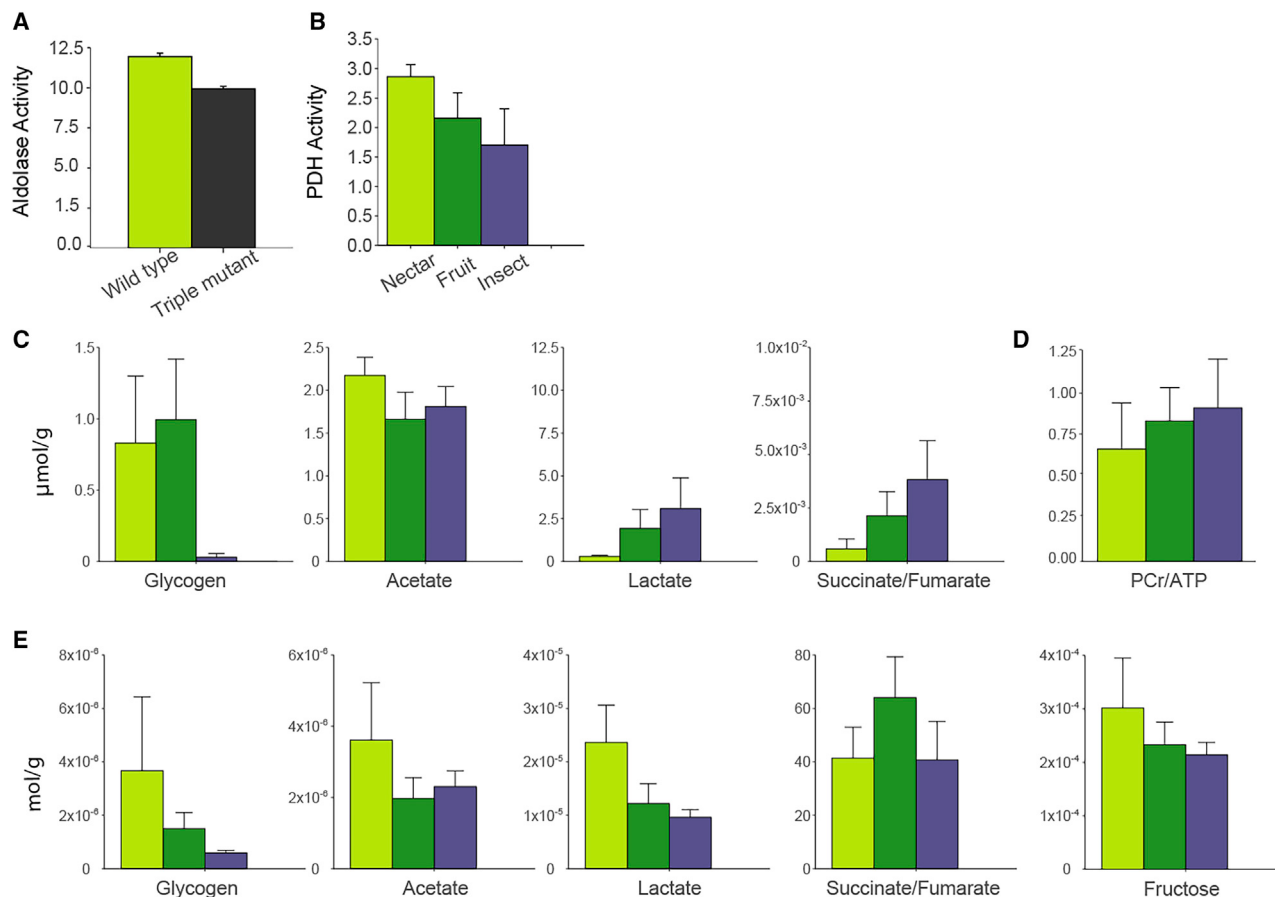


Figure 3. Indicators of glycolytic metabolism

(A) Aldolase activity ($\mu\text{mol min}^{-1} \text{mg}^{-1}$) of *Anoura* wild-type (nectar-feeding) and triple-mutant (ancestral) enzymes. Error bars are standard errors. (B) Pyruvate dehydrogenase activity ($\mu\text{mol min}^{-1} \mu\text{L}^{-1}$) from pectoral muscle for three groups of divergent dietary specialists (light green, nectarivory; dark green, frugivory; purple, insectivory). Error bars are standard errors. (C and E) Metabolite concentrations recorded in (C) pectoral muscle and (E) liver tissue for each of the bat feeding guilds. Error bars are standard errors. (D) Phosphocreatine to ATP ratio (PCr/ATP) calculated from metabolomic profiles for the same dietary groups with error bars showing standard errors. Comparable PCr/ATP ratios between the groups confirm that there is no impaired energetics in the muscle tissue samples. Posterior distributions of phylogenetic means are provided in [Figure S2](#).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
 - Lead contact
 - Materials availability
 - Data and code availability
- [EXPERIMENTAL MODEL AND SUBJECT DETAILS](#)
 - Focal species
 - Culturing of bacteria
- [METHOD DETAILS](#)
 - Sample collection
 - RNA extraction and sequencing
 - Transcriptome assemblies
 - Ortholog identification and alignment
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)

- Identification of genes under positive selection
- Correcting for multiple tests
- Functional GO enrichment
- Tests for convergence
- Protein modeling
- Quantification of enzyme activity
- Quantification of metabolites
- Comparisons of PDH activity levels and metabolomic profiles across species

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2021.08.018>.

ACKNOWLEDGMENTS

We thank J. Day, C. Economou, P. Howard, R. Rose, M. Struebig, P. Ungerer, K. Warren, and J. Crowe for technical support. For help with permits and field support in the Dominican Republic, we thank J. Flanders, J. Almonthe, M.E.

Lauterbur, Y.M. León, M.S. Nuñez, and J. Salazar; in Peru, F. Cornejo, J. Pacheco, E. Paliza, H. Portocarrero, M.K. Ramos, E. Rengifo, J.N. Ruiz, and C. Tello; in Costa Rica, B. Matarrita, M. Porras, L.B. Miller, A. Kaliszewska, S. Santana, and the La Selva Biological Station staff. We are grateful to A. Atkinson (King's College London) and H. Toms (QMUL) for access to NMR facilities. This work was funded by a European Research Council starting grant (310482) awarded to S.J.R., a Fundamental Research Funds for the Central Universities grant (GK202102006) to Y.L., Barts Charity grant (MRC 0215), British Heart Accelerator Award (AA/18/5/34222) to D.A., and National Science Foundation grants 1442142 to L.M.D. and S.J.R., 1601477 to M.C.W.L. and L.M.D., and 1701414 to L.R.Y. and L.M.D.

AUTHOR CONTRIBUTIONS

S.J.R. and J.H.T.P. conceived the project. S.J.R., L.M.D., L.R.Y., J.H.T.P., B.K.L., K.T.J.D., and C.C.W. undertook fieldwork. RNA extraction and library preparation were performed by J.H.T.P., K.T.J.D., and M.C.W.L. Transcriptome assemblies, alignment, and annotation were undertaken by J.H.T.P., K.T.J.D., and G.T. for bats, and by M.C.W.L., J.H.T.P., and N.N. for hummingbirds. J.H.T.P. and N.N. performed the molecular evolution and convergence analyses, with input from S.J.R., I.L., and K.T.J.D. Enzyme assays were performed by R.D., J.H.T.P., and D.A. Protein structure modeling was performed by W.F. and M.J.W. D.A. interpreted the metabolome data; H.C., X.Z., and Y.L. performed functional assays on impacts of mutations; J.H.T.P., I.L., M.d.R., and L.M.D. conducted statistical analyses; and J.H.T.P., S.J.R., and R.D. wrote the paper with input from all co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 12, 2020

Revised: April 21, 2021

Accepted: August 5, 2021

Published: September 2, 2021

REFERENCES

- Baker, H.G., Baker, I., and Hodges, S.A. (1998). Sugar composition of nectars and fruits consumed by birds and bats in the tropics and subtropics. *Biotropica* 30, 559–586.
- Kelm, D.H., Simon, R., Kuhlow, D., Voigt, C.C., and Ristow, M. (2011). High activity enables life on a high-sugar diet: blood glucose regulation in nectar-feeding bats. *Proc. Biol. Sci.* 278, 3490–3496.
- Suarez, R.K., and Welch, K.C. (2017). Sugar metabolism in hummingbirds and nectar bats. *Nutrients* 9, 743.
- Voigt, C.C., Kelm, D.H., and Visser, G.H. (2006). Field metabolic rates of phytophagous bats: do pollination strategies of plants make life of nectar-feeders spin faster? *J. Comp. Physiol. B* 176, 213–222.
- Brunner, Y., Schvartz, D., Priego-Capote, F., Couté, Y., and Sanchez, J.-C. (2009). Glucotoxicity and pancreatic proteomics. *J. Proteomics* 71, 576–591.
- Rolo, A.P., and Palmeira, C.M. (2006). Diabetes and mitochondrial function: role of hyperglycemia and oxidative stress. *Toxicol. Appl. Pharmacol.* 212, 167–178.
- Johnson, R.J., Nakagawa, T., Sanchez-Lozada, L.G., Shafiu, M., Sundaram, S., Le, M., Ishimoto, T., Sautin, Y.Y., and Lanaspa, M.A. (2013). Sugar, uric acid, and the etiology of diabetes and obesity. *Diabetes* 62, 3307–3315.
- Winter, Y., and von Helversen, O. (2003). Operational tongue length in phyllostomid nectar-feeding bats. *J. Mammal.* 84, 886–896.
- Arbour, J.H., Curtis, A.A., and Santana, S.E. (2019). Signatures of echolocation and dietary ecology in the adaptive evolution of skull shape in bats. *Nat. Commun.* 10, 2036.
- Voigt, C.C., and Winter, Y. (1999). Energetic cost of hovering flight in nectar-feeding bats (Phyllostomidae: Glossophaginae) and its scaling in moths, birds and bats. *J. Comp. Physiol. B* 169, 38–48.
- Potter, J.H.T., Davies, K.T.J., Yohe, L.R., Sanchez, M.K.R., Rengifo, E.M., Struebig, M., et al. (2021). Dietary diversification and specialisation in new world bats facilitated by early molecular evolution. *Mol. Biol. Evol.* 38, 3864–3883.
- Pounds, S., and Cheng, C. (2006). Robust estimation of the false discovery rate. *Bioinformatics* 22, 1979–1987.
- Gutiérrez-Guerrero, Y.T., Ibarra-Laclette, E., Martínez Del Río, C., Barrera-Redondo, J., Rebollar, E.A., Ortega, J., León-Paniagua, L., Urrutia, A., Aguirre-Planter, E., and Eguiarte, L.E. (2020). Genomic consequences of dietary diversification and parallel evolution due to nectarivory in leaf-nosed bats. *Gigascience* 9, gaa059.
- Vandeweghe, M.W., Sotero-Caio, C.G., and Phillips, C.D. (2020). Positive selection and gene expression analyses from salivary glands reveal discrete adaptations within the ecologically diverse bat family Phyllostomidae. *Genome Biol. Evol.* 12, 1419–1428.
- Patel, M.S., Nemeria, N.S., Furey, W., and Jordan, F. (2014). The pyruvate dehydrogenase complexes: structure-based function and regulation. *J. Biol. Chem.* 289, 16615–16623.
- Santer, R., Rischewski, J., von Weihe, M., Niederhaus, M., Schneppenheim, S., Baerlocher, K., Kohlschütter, A., Muntau, A., Posselt, H.-G., Steinmann, B., and Schneppenheim, R. (2005). The spectrum of aldolase B (ALDOB) mutations and the prevalence of hereditary fructose intolerance in Central Europe. *Hum. Mutat.* 25, 594.
- Chalcoff, V.R., Aizen, M.A., and Galetto, L. (2006). Nectar concentration and composition of 26 species from the temperate forest of South America. *Ann. Bot.* 97, 413–421.
- Vos, M.B., and Lavine, J.E. (2013). Dietary fructose in nonalcoholic fatty liver disease. *Hepatology* 57, 2525–2531.
- Zhao, H., Xu, D., Zhang, S., and Zhang, J. (2012). Genomic and genetic evidence for the loss of umami taste in bats. *Genome Biol. Evol.* 4, 73–79.
- Hong, W., and Zhao, H. (2014). Vampire bats exhibit evolutionary reduction of bitter taste receptor genes common to other bats. *Proc. Biol. Sci.* 281, 20141079.
- Wang, K., Tian, S., Galindo-González, J., Dávalos, L.M., Zhang, Y., and Zhao, H. (2020). Molecular adaptation and convergent evolution of frugivory in Old World and neotropical fruit bats. *Mol. Ecol.* 29, 4366–4381.
- Rey, C., Guéguen, L., Sémon, M., and Boussau, B. (2018). Accurate detection of convergent amino-acid evolution with PCOC. *Mol. Biol. Evol.* 35, 2296–2306.
- Fang, L., Shen, B., Irwin, D.M., and Zhang, S. (2014). Parallel evolution of the glycogen synthase 1 (muscle) gene *Gys1* between Old World and New World fruit bats (Order: Chiroptera). *Biochem. Genet.* 52, 443–458.
- Qian, Y., Fang, T., Shen, B., and Zhang, S. (2014). The glycogen synthase 2 gene (*Gys2*) displays parallel evolution between Old World and New World fruit bats. *J. Mol. Evol.* 78, 66–74.
- Zhu, L., Yin, Q., Irwin, D.M., and Zhang, S. (2015). Phosphoenolpyruvate carboxykinase 1 gene (*Pck1*) displays parallel evolution between Old World and New World fruit bats. *PLoS ONE* 10, e0118666.
- Yagihashi, S., Mizukami, H., and Sugimoto, K. (2011). Mechanism of diabetic neuropathy: where are we now and where to go? *J. Diabetes Investig.* 2, 18–32.
- Greene, D.A., Lattimer, S.A., and Sima, A.A.F. (1987). Sorbitol, phosphoinositides, and sodium-potassium-ATPase in the pathogenesis of diabetic complications. *N. Engl. J. Med.* 316, 599–606.
- Wu, E.Y., Walsh, A.R., Materne, E.C., Hiltner, E.P., Zielinski, B., Miller, B.R., 3rd, Mawby, L., Modeste, E., Parish, C.A., Barnes, W.M., and Kermekchiev, M.B. (2015). A conservative isoleucine to leucine mutation causes major rearrangements and cold sensitivity in KlenTaq1 DNA polymerase. *Biochemistry* 54, 881–889.
- Mackenzie, A.E., Quon, T., Lin, L.C., Hauser, A.S., Jenkins, L., Inoue, A., Tobin, A.B., Gloriam, D.E., Hudson, B.D., and Milligan, G. (2019). Receptor selectivity between the G proteins $G_{\alpha_{12}}$ and $G_{\alpha_{13}}$ is defined by a single leucine-to-isoleucine variation. *FASEB J.* 33, 5005–5017.

30. Okajima, K., Korotchkina, L.G., Prasad, C., Rupar, T., Phillips, J.A., 3rd, Ficicioglu, C., Hertecant, J., Patel, M.S., and Kerr, D.S. (2008). Mutations of the E1 β subunit gene (PDHB) in four families with pyruvate dehydrogenase deficiency. *Mol. Genet. Metab.* 93, 371–380.
31. Listenberger, L.L., Han, X., Lewis, S.E., Cases, S., Farese, R.V., Jr., Ory, D.S., and Schaffer, J.E. (2003). Triglyceride accumulation protects against fatty acid-induced lipotoxicity. *Proc. Natl. Acad. Sci. USA* 100, 3077–3082.
32. Liu, X., Cooper, D.E., Cluntun, A.A., Warmoes, M.O., Zhao, S., Reid, M.A., Liu, J., Lund, P.J., Lopes, M., Garcia, B.A., et al. (2018). Acetate production from glucose and coupling to mitochondrial metabolism in mammals. *Cell* 175, 502–513.e13.
33. Peng, X., He, X., Liu, Q., Sun, Y., Liu, H., Zhang, Q., Liang, J., Peng, Z., Liu, Z., and Zhang, L. (2017). Flight is the key to postprandial blood glucose balance in the fruit bats *Eonycteris spelaea* and *Cynopterus sphinx*. *Ecol. Evol.* 7, 8804–8811.
34. McWhorter, T.J., Bakken, B.H., Karasov, W.H., and del Rio, C.M. (2006). Hummingbirds rely on both paracellular and carrier-mediated intestinal glucose absorption to fuel high metabolism. *Biol. Lett.* 2, 131–134.
35. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
36. Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512.
37. Löytynoja, A. (2014). Phylogeny-aware alignment with PRANK. *Methods Mol. Biol.* 1079, 155–170.
38. Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738.
39. Yohe, L.R., Devanna, P., Davies, K.T.J., Potter, J.H.T., Rossiter, S.J., Teeling, E.C., et al. (2019). Tissue collection of bats for -omics analyses and primary cell culture. *J. Vis. Exp.* 152, e59505.
40. Lim, M.C.W., Witt, C.C., Graham, C.H., and Dávalos, L.M. (2019). Parallel molecular evolution in pathways, genes, and sites in high-elevation hummingbirds revealed by comparative transcriptomics. *Genome Biol. Evol.* 11, 1552–1572.
41. Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K.A., Fang, X., Wynne, J.W., Xiong, Z., Baker, M.L., Zhao, W., et al. (2013). Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* 339, 456–460.
42. Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., et al.; Broad Institute Sequencing Platform and Whole Genome Assembly Team; Baylor College of Medicine Human Genome Sequencing Center Sequencing Team; Genome Institute at Washington University (2011). A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478, 476–482.
43. Pavlovich, S.S., Lovett, S.P., Koroleva, G., Guito, J.C., Arnold, C.E., Nagle, E.R., Kulcsar, K., Lee, A., Thibaud-Nissen, F., Hume, A.J., et al. (2018). The Egyptian rousette genome reveals unexpected features of bat antiviral immunity. *Cell* 173, 1098–1110.e18.
44. Dong, D., Lei, M., Hua, P., Pan, Y.-H., Mu, S., Zheng, G., Pang, E., Lin, K., and Zhang, S. (2017). The genomes of two bat species with long constant frequency echolocation calls. *Mol. Biol. Evol.* 34, 20–34.
45. Zepeda Mendoza, M.L., Xiong, Z., Escalera-Zamudio, M., Runge, A.K., Thézé, J., Streicker, D., Frank, H.K., Loza-Rubio, E., Liu, S., Ryder, O.A., et al. (2018). Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat. *Nat. Ecol. Evol.* 2, 659–668.
46. Eckalbar, W.L., Schlebusch, S.A., Mason, M.K., Gill, Z., Parker, A.V., Booker, B.M., Nishizaki, S., Muswamba-Nday, C., Terhune, E., Nevenon, K.A., et al. (2016). Transcriptomic and epigenomic characterization of the developing bat wing. *Nat. Genet.* 48, 528–536.
47. Seim, I., Fang, X., Xiong, Z., Lobanov, A.V., Huang, Z., Ma, S., Feng, Y., Turanov, A.A., Zhu, Y., Lenz, T.L., et al. (2013). Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*. *Nat. Commun.* 4, 2212.
48. Choo, S.W., Rayko, M., Tan, T.K., Hari, R., Komissarov, A., Wee, W.Y., Yurchenko, A.A., Kliver, S., Tamazian, G., Antunes, A., et al. (2016). Pangolin genomes and the evolution of mammalian scales and immunity. *Genome Res.* 26, 1312–1322.
49. Francischetti, I.M.B., Assumpção, T.C.F., Ma, D., Li, Y., Vicente, E.C., Uieda, W., and Ribeiro, J.M.C. (2013). The “Vampirome”: Transcriptome and proteome analysis of the principal and accessory submaxillary glands of the vampire bat *Desmodus rotundus*, a vector of human rabies. *J. Proteomics* 82, 288–319.
50. Shaw, T.I., Srivastava, A., Chou, W.-C., Liu, L., Hawkinson, A., Glenn, T.C., Adams, R., and Schountz, T. (2012). Transcriptome sequencing and annotation for the Jamaican fruit bat (*Artibeus jamaicensis*). *PLoS ONE* 7, e48472.
51. Davies, K.T.J., Tsagkogeorga, G., Bennett, N.C., Dávalos, L.M., Faulkes, C.G., and Rossiter, S.J. (2014). Molecular evolution of growth hormone and insulin-like growth factor 1 receptors in long-lived, small-bodied mammals. *Genes* 549, 228–236.
52. Aken, B.L., Achuthan, P., Akanni, W., Amode, M.R., Bernsdorff, F., Bhai, J., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., et al. (2017). Ensembl 2017. *Nucleic Acids Res.* 45 (D1), D635–D642.
53. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.
54. King, T., Butcher, S., and Zalewski, L. (2017). Apocrita - high performance computing cluster for Queen Mary University of London.
55. Zhang, J., Nielsen, R., and Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* 22, 2472–2479.
56. Junier, T., and Zdobnov, E.M. (2010). The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26, 1669–1670.
57. Bielawski, J.P., and Yang, Z. (2004). A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J. Mol. Evol.* 59, 121–132.
58. Weadick, C.J., and Chang, B.S.W. (2012). Complex patterns of divergence among green-sensitive (RH2a) African cichlid opsins revealed by Clade model analyses. *BMC Evol. Biol.* 12, 206.
59. Yang, Z., and dos Reis, M. (2011). Statistical properties of the branch-site test of positive selection. *Mol. Biol. Evol.* 28, 1217–1228.
60. Yang, Z., Wong, W.S.W., and Nielsen, R. (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22, 1107–1118.
61. Tsagkogeorga, G., McGowen, M.R., Davies, K.T.J., Jarman, S., Polanowski, A., Bertelsen, M.F., and Rossiter, S.J. (2015). A phylogenomic analysis of the role and timing of molecular adaptation in the aquatic transition of cetartiodactyl mammals. *R. Soc. Open Sci.* 2, 150156.
62. Davies, K.T.J., Bennett, N.C., Faulkes, C.G., and Rossiter, S.J. (2018). Limited evidence for parallel molecular adaptations associated with the subterranean niche in mammals: a comparative study of three superorders. *Mol. Biol. Evol.* 35, 2544–2559.
63. Alexa, A., and Rahnenfuhrer, J. (2010). topGO: enrichment analysis for gene ontology. R package version 2.
64. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
65. Meredith, R.W., Janečka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simão, T.L.L., Stadler, T., et al. (2011). Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* 334, 521–524.
66. Kuhl, H., Frank-Vilches, C., Bakker, A., Mayr, G., Nikolaus, G., Boerno, S.T., Klages, S., Timmermann, B., and Gahr, M. (2021). An unbiased molecular approach using 3'-UTRs resolves the avian family-level tree of life. *Mol. Biol. Evol.* 38, 108–127.

67. Rojas, D., Warsi, O.M., and Dávalos, L.M. (2016). Bats (Chiroptera: Noctilionoidea) challenge a recent origin of extant neotropical diversity. *Syst. Biol.* **65**, 432–448.
68. McGuire, J.A., Witt, C.C., Remsen, J.V., Jr., Corl, A., Rabosky, D.L., Altshuler, D.L., and Dudley, R. (2014). Molecular phylogenetics and the diversification of hummingbirds. *Curr. Biol.* **24**, 910–916.
69. Shi, J.J., and Rabosky, D.L. (2015). Speciation dynamics during the global radiation of extant bats. *Evolution* **69**, 1528–1545.
70. Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M., and Lemmon, A.R. (2015). A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569–573.
71. Patel, M.S., and Korotchkina, L.G. (2001). Regulation of mammalian pyruvate dehydrogenase complex by phosphorylation: complexity of multiple phosphorylation sites and kinases. *Exp. Mol. Med.* **33**, 191–197.
72. Santamaria, R., Esposito, G., Vitagliano, L., Race, V., Paglionico, I., Zancan, L., Zagari, A., and Salvatore, F. (2000). Functional and molecular modelling studies of two hereditary fructose intolerance-causing mutations at arginine 303 in human liver aldolase. *Biochem. J.* **350**, 823–828.
73. Faulkes, C.G., Eykyn, T.R., and Aksentijevic, D. (2019). Cardiac metabolomic profile of the naked mole-rat-glycogen to the rescue. *Biol. Lett.* **15**, 20190710.
74. Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**, 1–22.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
RNAlater	Sigma-Aldrich	Cat No: R0901-500ML
Tris/HCl buffer (pH = 7.5)	Sigma-Aldrich	Cat No: B548124
Isopropyl β-D-thiogalactopyranoside	Sangon Biotech	Cat No: A600168
Fructose-1,6-bisphosphate	Acme Biochemical	Cat No: D89910-5g
β-nicotinamide adenine dinucleotide disodium salt	Sangon Biotech	Cat No: A600642
Ethylendinitrilotetraacetic acid	Sigma-Aldrich	Cat No: EDS-100G
α-glycerophosphate dehydrogenase	Sigma-Aldrich	Cat No: G6751-200UN
Triose-phosphate isomerase	Sigma-Aldrich	Cat No: T6258-1MG
Ammonium sulfate (~4.1M) (Qmul)	Sigma-Aldrich	Cat No: A4418-100G
Chloroform-d 99.8 atom % D, contains 0.05% (v/v) TMS	Sigma-Aldrich	Cat No: 612200
Methanol - Fisher Chemical M/4056/15	Fisher Scientific	Cat No: 10365710
Chelex 100 sodium form 50-100 mesh (dry)	Sigma-Aldrich	Cat No: C7901-25G
Honeywell Fluka Universal Indicator Solution Ph	Fisher Scientific	Cat No: 15683370
Deuterium Oxide 99.9 atom % D, contains 0.75 wt. % 3-(trimethylsilyl)propionic-2,2,3,3-d ₄ acid, sodium salt	Sigma-Aldrich	Cat No: 293040
Oxid Phosphate Buffered Saline Tablets (Dulbecco A) containing 8 g/L NaCl, 0.2 g/L KCl, 1.15 g/L Na ₂ HPO ₄ , 0.2 g/L KH ₂ PO ₄	Thermo Scientific Fisher	Cat No: BR0014G
Anoura ALDOB	This study	N/A
Anoura ALDOB triple mutant	This study	N/A
Critical commercial assays		
RNeasy Mini kits	QIAGEN	Cat. No: 74004
Qubit RNA BR Assay Kit	ThermoFisher Scientific	Cat. No: Q10210
Agilent RNA 6000 Nano Kit	Agilent	Part Number: 5067-1511
TruSeq RNA Sample Preparation v2	Illumina	Cat. No: RS-122-2001
NEBNext Ultra RNA Library Prep Kit for Illumina	New England Biolabs	Cat. No: E7530L
PDH Activity Assay Kit	Sigma-Aldrich	MAK 183
BCA Protein Assay Kit	Beyotime Biotechnology	Cat. No: P0010S
Deposited data		
List of studied samples for genetic analysis	This paper	Data S1
Experimental models: Organisms/strains		
<i>E. coli</i> BL21 (DE3) strain	Thermo Fisher	Cat No: EC0114
Recombinant DNA		
Anoura ALDOB gene ligated into pET28a	General Biosystems	http://www.generalbiol.com
Anoura ALDOB mutant ligated into pET28a	General Biosystems	http://www.generalbiol.com
Software and algorithms		
Trimomatic	³⁵	version 0.40
Trinity	³⁶	version 2.4
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi	N/A
Ensembl human genome release GRCh38.p10	Ensembl, http://www.ensembl.org/Homo_sapiens/Info/Index	Release 90, August 2017
PRANK	³⁷	version 170427
PAML	²²	version 4.9e

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Newick Utilities	https://github.com/tjunier/newick_utils/wiki	version 1.6
R	https://cran.r-project.org/src/base/R-3/R-3.6.3.tar.gz	version 3.6.3
topGO	https://bioconductor.org/packages/release/bioc/html/topGO.html	version 2.44
PCOC	https://github.com/CarineRey/pcoc	²²
IQ-TREE	http://www.iqtree.org	version 1.6.12
I-TASSR web server	https://zhanglab.dcmf.med.umich.edu/I-TASSER/	³⁸
Topspin	https://www.bruker.com/en/products-and-solutions/mr/nmr-software/topspin.html	version 4.0.6
Chenomx NMR Profiler	Chenomx, Canada https://www.chenomx.com/products/	version 8.1
MCMCglmm	https://www.rdocumentation.org/packages/MCMCglmm/versions/2.32	version 2.32

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Stephen Rossiter (s.j.rossiter@qmul.ac.uk)

Materials availability

This study did not generate new unique reagents

Data and code availability

- RNA-seq data have been deposited at NCBI GenBank/NCBI Sequence Read Archive (SRA) and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#). Enzyme activity data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Focal species

We studied nectar-feeding bats and hummingbirds, sister taxa, and outgroups. All bats and hummingbirds were collected under site-specific research permits between 2009 and 2017 at several sites in Central and South America ([Data S1C](#)). Bats and hummingbirds were trapped in high and low mist nets of varying length (4–12 m) set up along forest trails, over streams, or near nectar-bearing flowers. Following protocols approved by the Institutional Animal Care and Use Committees (IACUC) of Stony Brook University (2013-2034-NF-4.15.16-BAT and 2014-2090-R1-1.20.17-Bat) and the University of New Mexico (08UNM033-TR-100117), subjects were humanely collected for immediate tissue extraction and specimen preparation.³⁹ Data from outgroups were obtained from GenBank. The genomics level data examined in this project are listed in [Data S1C](#).

Culturing of bacteria

We expressed recombinant proteins using the commercial *E. coli* BL21 (DE3) strain. Cells were cultured at 37°C prior to transformation, and subsequently at 15°C to induce protein expression.

METHOD DETAILS

Sample collection

Bats were collected in the field for genetic, enzyme and metabolome analyses. Tissue samples, either dry or immersed in RNAlater, were flash-frozen in liquid nitrogen in a portable dry-shipper. All bat samples collected by BKL were from specimens subsequently deposited at the Royal Ontario Museum. For hummingbirds, tissue was flash-frozen in liquid nitrogen and subsequently transferred to RNAlater (for details, see Lim et al.⁴⁰). Voucher specimens and additional frozen tissues for hummingbird samples were deposited at

the Museum of Southwestern Biology (University of New Mexico) and at CORBIDI (Lima, Peru). Bats were collected under research permits VAPB-01436 in the Dominican Republic, 0002287 in Peru, and R-018-2013-OT-CONAGEBIO in Costa Rica. Hummingbirds were collected in Peru under research permits 135-2009, 0377-2010, and 0199-2012 (AG-DGFFS-DGEFFS), and 006-2013 (MINAGRI-DGFFS/DGEFFS).

RNA extraction and sequencing

To generate the sequences of protein-coding genes, we performed RNA-Seq. Total RNA was extracted from bat tissue samples with QIAGEN RNeasy Mini kits according to standard protocols, and RNA quality assessed with Agilent 2100 Bioanalyzer and Qubit Fluorometer. For cDNA library construction, we used Illumina TruSeq RNA Sample Preparation v2 for bats, and NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, Ipswich, MA) for birds. Library preparations were performed by BGI, Novogene and TGAC using their standard protocols. The resulting libraries were sequenced on either the NextSeq or HiSeq platforms yielding an average of 25 million paired-reads per sample (Data S1C).

Transcriptome assemblies

Raw RNA-Seq reads were cleaned by adaptor trimming and quality filtering based on phred scores, with Trimmomatic³⁵ using the following parameters: 16-base ‘seed’ alignment, maximum seed mismatch of 2 bases, palindromic clipping threshold score of 30, single-end clipping threshold score of 10 and a four-base sliding window with a cut threshold of 15.

Bat transcriptomes were assembled using default parameters in Trinity v2.4.³⁶ Transcriptome assembly methods for the hummingbird samples are detailed in Lim et al.⁴⁰ In total, we assembled RNA-seq datasets and analyzed coding gene sequence datasets from 22 genera of nectar-feeding vertebrates, encompassing four independent origins: 11 genera of nectar-feeding bats – three lineages comprising eight genera from subfamily Glossophaginae and two from subfamily Lonchophyllinae (family Phyllostomidae), and *Eonycteris spelaea* from family Pteropodidae – and 11 genera of hummingbirds (family Trochilidae).

To increase taxonomic representation we supplemented our *de novo* assemblies with publicly available mRNA refseq transcripts available for 46 outgroup mammal and bird species from NCBI GenBank.^{41–50} We also used additional bat transcriptomes from published SRA data.^{11,41,50,51} Our final dataset comprised 126 mammal and bird species, including 25 nectar-feeders (Data S1C).

Ortholog identification and alignment

We used a stringent reciprocal BLAST method to define orthologous sequences from transcriptomes. For each transcriptome in turn, transcripts were queried with blastx (E-value threshold 1e-06) against a reference database of the longest protein sequences for each *Homo sapiens* coding gene (Ensembl human genome release GRCh38.p10, Release 90, August 2017).⁵² The reverse tblastn search was also performed. For each search, only the top hit per query sequence was retained. If the top hit was reciprocal in the two searches, the transcript was assumed to be homologous to the human gene. Using a perl script, the BLAST-aligned sequence within the transcript, which represents the coding sequence (CDS), was extracted for multiple species alignments (MSAs) and downstream analysis. Additional filtering was also performed to ensure orthology. For birds, only genes annotated as Ensembl one-to-one orthologs between *H. sapiens* and either the zebra finch (*Taeniopygia guttata*) or chicken (*Gallus gallus*) were retained. For bats, only Ensembl one-to-one orthologs between *H. sapiens* and either of the bat species *Myotis lucifugus* or *Pteropus vampyrus* were retained. If orthology information for the reference human gene was not annotated in either Chiroptera or Aves, we retained the gene if it had one-to-one orthology in the other group. All BLAST-derived CDSs were filtered for length and premature stop codons using custom Perl scripts.

Orthologous CDSs were aligned with PRANK v.170427,³⁷ which provides protein-aware codon based alignments, producing a single MSA per gene with all bird, bat and outgroup taxa. Finished alignments were parsed to a single representative sequence per species; the most complete sequence (fewest gaps) was retained. We subsequently removed any alignment positions (columns) in which a gap had been introduced in the reference *H. sapiens* sequence (i.e., insertions in non-human species were not considered comparable), and we also discarded any sequences comprising more than 50% gaps after alignment. Further quality filtering masked any column where data was missing data from more than 50% of species. In total 13,899 orthologous gene alignments containing nectar-feeding taxa were generated.

QUANTIFICATION AND STATISTICAL ANALYSIS

Identification of genes under positive selection

To identify genes and sites that have undergone molecular adaptation in the evolution of nectar-feeding, we performed high-throughput genewise tests of positive selection on orthologous gene alignments for each of the nectar-feeding lineages in turn: the ancestral branches of Glossophaginae, Lonchophyllinae and Trochilidae, and *Eonycteris spelaea*. We also repeated these tests for a clade of five bat species from the family Mormoopidae, all of which feed on insects. By comparing sets of genes under selection in our focal clades with those in our control group, we were able to disentangle putative cases of molecular adaptation related to sugar-rich diets from those that might be rapidly evolving in general.

All tests were performed in codeml in PAML v4.9e, which has been shown to perform well for identifying episodic selection in vertebrates,⁵³ using Queen Mary’s Apocrita HPC facility.⁵⁴ For a given species tree topology codeml uses maximum likelihood codon models of molecular adaptation to estimate the rates of non-synonymous and synonymous substitutions (dN and dS, respectively).

A dN/dS ratio (ω) > 1 implies positive selection, ~ 1 signifies neutral evolution, and < 1 purifying selection. We first ran branch-site models on the ancestral branch of each clade of nectar feeders (foreground), removing all other nectar-feeders from the background. To assess whether positive selection is a better fit than neutral evolution we implemented branch-site Model A, the preferred standard test for positive selection, and compared this to its respective null model.⁵⁵ Model A classifies sites to one of four ω site categories: $\omega_0 < 1$ (purifying selection); $\omega_1 = 1$ (neutral evolution); $\omega_{2a} > 1$ (positive selection in foreground, purifying in background); and $\omega_{2b} > 1$ (positive selection in foreground, neutral in background). In the null model, ω in site classes 2a and 2b cannot exceed 1 in the foreground branch, such that the null is bounded at neutral evolution. We compared model fit using the likelihood ratio test (LRT), with significance assessed by a χ^2 test under one degree of freedom.

Branch-site models were only run on gene alignments that contained at least one species of the focal nectarivorous lineage, four outgroup species, and a minimum of ten total taxa. When testing nectarivorous bats, we removed non-bat species from each alignment, and, conversely for tests of hummingbirds we removed non-birds from the analyses. Additionally, for bats, when focusing on each lineage in turn, we also removed other nectar-feeders from the alignment background, to avoid masking parallel signatures of adaptation. The tree topologies used for selection tests are detailed below (See [Tests for convergence](#)), and were trimmed to match alignment taxon representation using Newick Utilities v1.6.⁵⁶ Branch lengths were optimized in codeml during model fitting.

As a secondary test of positive selection, we ran clade models on each gene alignment for Glossophaginae, Lonchophyllinae and Trochilidae. Although clade models are typically used to detect diversifying selection between different clades, we designated only a single focal clade and filtered our significant LRTs (χ^2 , 1 df, $p < 0.05$) by cases where the estimated foreground (focal clade) ω was > 1 , and the background (rest of tree) ≤ 1 . We used Clade Model C, which assigns sites to three classes: $\omega_0 < 1$, $\omega_1 = 1$, and $\omega_2 > 0$, where ω_2 is estimated separately for the clade of interest and the background.⁵⁷ This was compared to the null model, Model 2a_rel,⁵⁸ which has the same three classes, but where all are estimated across the tree. Prior to running clade models, we again filtered alignments to be bird- or bat-specific and removed the remaining nectar-feeders from the background. We also ensured that each alignment tested contained at least two genera from the focal lineage; *Eonycteris* was not tested with clade models.

The number of genewise selection tests for each lineage/model after all filtering steps were as follows: Glossophaginae branch-site: 13,487, Glossophaginae clade: 12,571, Lonchophyllinae branch-site: 12,815, Lonchophyllinae clade: 4,689, *Eonycteris* branch-site 6,474, Trochilidae branch-site: 11,220, and Trochilidae clade: 11,220. Additionally, we ran these tests on the control bat lineage Mormoopidae (9525).

Correcting for multiple tests

To reduce the incidence of false positives, which can arise from running many tests, we implemented several filtering steps. Classic Benjamini-Hochberg false discovery rate (FDR) requires that p values are uniformly distributed, a condition that is violated in the case of branch-site models of positive selection, for several reasons. First, the null model, which permits neutral selection in foreground lineages, is nested within the alternative model, which permits neutral evolution or positive selection in the foreground. Due to this shared boundary, when neutral evolution is an equal or better fit than positive selection, the null and alternative models perform equally well and return the same log likelihood. In such cases, the LRT statistic is zero and the p value is 1. Neutral evolution is expected to be as or more likely than positive selection at least 50% of the time, and thus simulations of repeated branch-site tests across multiple genes return a distribution with approximately half of p values equal to 1.⁵⁹ Thus, to account for this condition, it is recommended that p values from the branch-site test should be halved. Second, results from large empirical datasets show that the branch-site test may yield over 80% of p values ~ 1 , so exceeding the expectation of 50%, due to misspecification of the branch-site null model, which does not allow for proteins to be highly-conserved with all sites under purifying selection. Simulations of such distributions with a majority mass of large p values show that application of FDR is overly conservative and discards most or all true positives.¹¹

To account for these issues, we identified putative cases of positive selection based on both raw, and a modified FDR procedure, in each case with additional filtering to improve rigour. First, we applied a robust FDR (rFDR) correction¹² on the cohort of raw p values of < 0.99 . This threshold was used to account for the excess of uninformative p values due to the boundary condition, and also to diagnose putative artifacts of model misspecification, inclusion of which would violate assumptions for FDR.¹¹ The rFDR procedure does not assume a two-sided test, and thus performs more favorably when a poorly specified one-sided null model produces an excess of p values at the right tail, as is our case here because the branch-site test is one-sided. Still, we suspect the inadequacy of the branch-site test to deal with highly conserved proteins makes the rFDR conservative as well.

Second, to be conservative, for all loci for which the null model could be rejected (LRT $p < 0.05$), we undertook post hoc filtering to identify only genes with at least one positively selected site based on a Bayes Empirical Bayes (BEB) of > 0.5 in either site class 2A or 2B.⁶⁰ Additionally, significant genes with clusters of > 5 BEB sites (median interval between selected sites < 10 amino acids) were discarded, to reduce false positives from alignment error.^{61,62} These filtering steps decreased the overall number of significant genes by 17.0% to 49.1% per lineage, which we report as positively selected genes (PSGs).

The clade model test is not susceptible to the same issues as the branch-site test, and thus does not display an extreme excess of p values ≈ 1 . However, since the clade model is a one-tailed test of differential selection, rFDR was applied to all p values from the clade model test. As above, the same BEB filtering steps were applied, along with reporting loci as PSGs only where estimates of ω were > 1 in the clade of interest, and ≤ 1 in the background. For both the branch-site and clade model, rFDR was applied with results from all lineages pooled, thus accounting for testing the same genes in multiple lineages.

Functional GO enrichment

For each nectarivorous lineage, significant results of branch-site and clade models were pooled and gene ontology enrichment for positively selected genes was tested with topGO in R using the ‘weight01’ algorithm and Fisher’s exact test for significance ($p < 0.05$).⁶³ Biological process (BP) GO term annotations were taken from Ensembl. The background for the enrichment test used in each case was the full cohort of genes tested for the specific lineage. The topGO procedure takes the hierarchical structure of GO terms into account such that multiple testing of terms is nonindependent, and thus resulting p values were not corrected.⁶³

Tests for convergence

We screened genes that showed parallel positive selection (i.e., positive selection in more than one lineage of nectar-feeder) for evidence of molecular convergence. For this we used the software PCOC, which has been shown to outperform other methods.²² Here, convergence at a site is characterized as a shift from an ancestral biochemical role to a new biochemical role that is shared by more than one convergent lineage. These biochemical roles are modeled as ‘profiles’ consisting of vectors of amino acid frequencies built from large empirical datasets. PCOC assigns sites to profiles based on observed amino acid frequencies. If a site is inferred, under maximum likelihood, to belong to the same profile in convergent branches, chosen *a priori*, and a different profile in the background, it is considered to show a profile change (PC). Additionally, the site must display at least one substitution, or one change (OC), in each convergent branch. The combination of the two provides a single estimate of posterior probability of PCOC convergence. Model fit was assessed against a null model where all branches belong to the same profile. The complete gene alignments containing birds, bats and outgroup mammals were tested with PCOC against the species tree, with all nectarivorous lineages present labeled as putatively convergent branches, such that PCOC was testing up to a four-way convergent scenario.

Since PCOC is a phylogeny-based method, we produced a combined mammal and bird tree based on all 13,899 alignments. For this, we used a maximum likelihood concatenation approach in IQ-TREE,⁶⁴ in which we estimated branch-lengths for a fixed tree topology encompassing all 117 species of birds and mammals. The master topology was based on published phylogenies of deeper relationships in birds and mammals^{65,66} and species-level relationships of bats and hummingbirds^{67–69} (Figure S1). We also repeated analyses under alternative phylogenetic hypotheses, in which the hummingbirds were recovered as a basal lineage within the Neoaves.⁷⁰ Each gene in the concatenated supermatrix was assigned an individual partition; we allowed IQ-TREE to choose the appropriate substitution model and apply independent rate estimation for each partition. The final phylogeny estimated branch lengths not shown, see Figure S1) was used for tests of genetic convergence.

Protein modeling

We modeled the potential impact of substitutions shared by nectar-feeding bats in three glycolytic proteins: the alpha (PDHA1) and beta (PDHB) subunits of pyruvate dehydrogenase, and aldolase B (ALDOB). For PDHA1 and PDHB we used the I-TASSR web server³⁸ to generate homology models using the glossophagine bat *Erophylla sezekorni* (Es) amino acid sequence, which shows 98% and 95% sequence identity with the respective human subunits. From the homology modeled alpha and beta subunits, an intact Es homology heterotetramer was created by least-squares superposition of the individual Es subunit models onto their counterparts in the experimentally determined structure of human pyruvate dehydrogenase (PDB: 3EXF). In the Es heterotetramer model, the positions of the coenzyme thiamin diphosphate and the magnesium and potassium ions were retained directly from the structure of the human enzyme. We considered two possibilities for how the identified mutations might alter PDH function. First, the mutations could affect the catalytic centers of the enzyme, the regions near the coenzyme thiamin diphosphate and magnesium ion. We also considered the region near the potassium ion, although the biological role of this ion is not clearly demonstrated. Second, we considered whether the mutations changed the overall throughput of the whole PDH multi-enzyme complex by altering the interaction between PDH and its regulatory proteins. To the best of our knowledge, all mammalian PDH enzymes are regulated via a kinase/phosphatase system, with phosphorylation by kinase leading to decreased PDH catalysis, and removal of the phosphoryl group by the phosphatase leading to increased catalytic activity.⁷¹

For ALDOB, we used I-TASSR³⁸ to generate a homology model of the *Anoura geoffroyi* (Ag) aldolase B enzyme based on the human protein (PDB: 1XDL). The root-mean-square deviation between the *A. geoffroyi* model and the experimental human structure is 0.72 Å. The site of the convergent A25V mutation is distal to the reaction products found in the active site, located more than 18 Å away, and the V25 side chain points outward, away from the protein core. Thus, the minor lengthening of this side chain resulting from the A to V mutation likely has little impact on catalysis. The A318L mutation occurs on the protein surface on a short loop connecting two α -helices, over 23 Å from the closest atoms of the reaction products in the active site, while the L63I mutation is located on the periphery of the active site, approximately 12 Å from closest atoms of the reaction products. This mutation involves conversion of one δ -methyl group of L into a γ -methyl group on I, with the side chain located on an α -helix and pointing into the interface with two other helices in the structure. Although conservative, it is possible that the L $\rightarrow$ I mutation affects the packing of the three helices, and any structural rearrangement of these helices might propagate to the nearby active site.

Quantification of enzyme activity

To assess the functional impact of the three convergent substitutions recorded in the ALDOB of nectar-feeding lineages (A25V, L63I and A318L, see Figure 2), we compared the enzymatic activity of a wildtype protein with that of a triple mutant in which the convergent three residues were replaced by their respective ancestral states (V25A, I63L and L318A). Each protein was expressed in *E. coli* strain BL21 (DE3) by ligating into a pET28a (+) expression vector using the restriction sites *Nco*I and *Xho*I. A His-tag sequence

(CAC-CAC-CAC-CAC-CAC-CAC) was added upstream of the 5' end of the coding region of each gene for subsequent protein purification. Recombinant plasmids were transformed into the *E. coli* and when the OD₆₀₀ reached 0.6–0.8, IPTG was added to induce protein expression (15°C for 16 h). Following collection and ultrasonication of the bacteria, the expressed His-tagged wildtype and mutant proteins were purified by Ni-NTA for 1 h at 4°C. The expressed enzymes were eluted by Tris-HCl buffer and then purified by gel filtration chromatography (Superdex200). SDS-PAGE was used to determine the purity of the two purified samples, and the protein concentration of each was measured using a bicinchoninic acid (BCA) assay (Beyotime Biotechnology).

We quantified activity of the purified proteins following Santamaria et al.⁷² We prepared 200 µl reaction volumes containing 1 µg purified protein, 100 mM Tris/HCl buffer (pH = 7.5), 1 mM fructose-1,6-bisphosphate, 0.2 mM NADH, 0.5 mM EDTA, 2 µg α-glycerophosphate dehydrogenase and 2 µg triose-phosphate isomerase. We recorded absorbance at 340 nm (measuring NADH) at 37°C every minute for up to 30 min using a microplate reader (TECAN). For each of the proteins, we set up six independent reactions, with each reaction divided into two technical replicates.

For pyruvate dehydrogenase (PDH), activity was measured in endogenous tissue. For species sampled in Costa Rica (2017), representing three main dietary guilds, we collected duplicate samples of pectoral muscle (Table S1). Muscle samples from 16 species were immersed in a cryoprotectant (glycerol) and placed in liquid nitrogen. For each taxon, ~1 g pectoral muscle was ground up on liquid nitrogen after washing off glycerol. Using a colorimetric assay kit (cat no: MAK183, Sigma-Aldrich, UK), ground tissue was homogenized on ice with PDH buffer, and 100 µL of the resulting lysate was cleaned by ammonium sulfate (~4.1 M) precipitation following the manufacturer's recommendation. We used 20 µL of clean lysate per reaction with the exception of *Artibeus watsoni* and *Carollia castanea* for which 10 µL of lysate was used due to shortage of material.

For each species, we prepared both positive and negative reactions, where positive reactions contained 20 µL clean lysate, PDH developer and PDH substrate, and negative reactions contained only 20 µL clean lysate and PDH developer. This allowed the final absorbance of value $[(A_{450})_{\text{final}}]$ to be corrected for the species-specific background. The assay was conducted with six standards of known concentrations of NADH (0, 2.5, 5, 7.5, 10 & 12.5 nmol/µL) on the same plate, to generate a standard curve of NADH production, and a positive control was also included on the assay plate. All reactions were prepared in a final volume of 50 µL and conducted in duplicate, including standards and positive controls. A plate reader (BMG labtech, UK) was used to measure the enzymatic reaction which was conducted at 37°C for a total of 135 cycles, measuring absorbance at 450 nm, every 76 s per well. The mean value for each time results for all samples were blank corrected and the average was taken at each time point. PDH activity could then be calculated per species following the assay kit protocol.

Quantification of metabolites

We also collected samples of muscle and liver for metabolomic profiling, again using the same set of individuals collected in Costa Rica, as described above. Frozen tissue was weighed and pulverized (Table S1), and then electromechanically homogenized subject to methanol/water/chloroform dual-phase extraction.⁷³ The upper aqueous phase was separated from the chloroform and protein fractions. 20–30 mg chelex-100 was added to chelate paramagnetic ions, vortexed and centrifuged at 3600 RPM for 10 min at 4°C. The supernatant was added to a fresh Falcon tube containing 10 µL universal pH indicator solution followed by vortexing and lyophilisation. Dual-phase-extracted metabolites were reconstituted in 400 µL deuterium oxide (containing 8 g/L NaCl, 0.2 g/L KCl, 1.15 g/L Na₂HPO₄, 0.2 g/L KH₂PO₄ and 0.0075% w/v trimethylsilyl propanoic acid (TSP)).⁷⁴

Muscle samples were analyzed at 303 K and 800 MHz using a Bruker Avance NEO 800 spectrometer comprising an 18.8 T magnet, NEO console and cryogenically cooled triple-resonance probe. Liver spectra were acquired using a Bruker Avance NEO 600 spectrometer. 1H spectra were acquired in the manufacturer's TopSpin 4.0.6 software using the noesygppr1d pulse sequence with a spectral width of 15.6 ppm, 128 transients of 64 K data points (acquisition time per transient: 2.62 s) and a relaxation delay between transients of 4 s, giving a total duration of 14.7 min per sample. Acquired data were multiplied by an exponential function with a line broadening of 0.3 Hz prior to Fourier transformation, phasing and baseline correction. TopSpin software was used for metabolite quantification. Chemical shifts were normalized by setting the TSP signal to 0 ppm. Peaks of interest were initially integrated automatically using a pre-written integration region text file and then manually adjusted where required. Assignment of metabolites to their respective peaks was carried out based on previously obtained in-house data, confirmed by chemical shift and confirmed using Chemomx NMR Profiler Version 8.1 (Chemomx, Canada). Peak areas were normalized to the TSP peaks and metabolite concentrations quantified per gram tissue wet weight.⁷³

Comparisons of PDH activity levels and metabolomic profiles across species

To evaluate differences in PDH activity and metabolomics profiles while accounting for evolutionary relatedness, we implemented phylogenetic regressions using MCMCglmm.⁷⁴ We modeled each response variable, scaled through logarithmic transformation when applicable, as a function of the predominant diet of the species as a categorical variable. The phylogenetic structure of errors was included as species-specific ('random') effect with a covariance matrix given by a sparse inverse matrix derived from the phylogeny.⁶⁹ Each univariate Bayesian analysis ran for 50,000 iterations with 10% as burnin and thinning every 50 iterations. Convergence was evaluated by examining estimated sample sizes for all model parameters above 200.