Untargeted Metabolomics of the Cavefish *Astyanax mexicanus* Reveals the Basis of Metabolic Strategies in Adaptation to Extreme Conditions

J Kyle Medley[1,†,✉], **Jenna Persons**[1,2,†], **Robert Peuß**[1,3], **Luke Olsen**[1,4], **Shaolei Xiong**[1], **and Nicolas Rohner**[1,4,✉]

[1]Stowers Institute for Medical Research, Kansas City, Missouri, U.S.A.
[2]Current Address: Newman University, Wichita, Kansas, Missouri, U.S.A.
[3]Current Address: University of Münster, Schlossplatz 2, 48149 Münster, Germany
[4]Department of Molecular and Integrative Physiology, University of Kansas Medical Center, Kansas City, Kansas, U.S.A.

[†]These authors contributed equally to this work.

The Mexican tetra, *Astyanax mexicanus*, has undergone remarkable physiological and behavioral changes in order to colonize a number of subterranean caves in the Sierra de El Abra region of Mexico. A hallmark of cave-adapted populations is enhanced survival under low-nutrient conditions coupled with hyperglycemia, increased body fat, and insulin resistance, but cavefish appear to avoid the progression of the respective pathologies associated with these conditions and do not exhibit reduced longevity. The metabolic strategies underlying these adaptations are not fully understood. Here, we provide an untargeted metabolomics study of long- and short-term fasting in two *A. mexicanus* cave populations and one surface population. We find that, although cavefish share many similarities with metabolic syndrome normally associated with the human state of obesity, important differences emerge, including cholesterol esters, urate, intermediates of protein glycation, metabolites associated with hypoxia and longevity, and unexpectedly elevated levels of ascorbate (vitamin C). This work highlights the fact that certain metabolic features associated with human pathologies are not intrinsically harmful in all organisms, and suggests promising avenues for future investigation into the role of certain metabolites in evolutionary adaptation and health. We provide a transparent pipeline for reproducing our analysis and a Shiny app for other researchers to explore and visualize our dataset.

Obesity | Fasting resistance | Metabolomics | Extreme adaptation | Evolution
Correspondence: ***nro@stowers.org (Nicolas Rohner), jmedley@stowers.org (Kyle Medley)***

Introduction

Metabolites are the intermediates and end products of biochemical reactions that serve essential roles as energy, signaling, and regulatory molecules (1, 2). The complete set of metabolites, or 'metabolome', is a functional read-out of interactions between genes, gene products, and the environment (3). Metabolomics techniques aim to capture a comprehensive view of the metabolome of cells, tissues, and organisms in normal and pathogenic states (4).

The enzymes and pathways involved in core processes for the metabolism of carbohydrates, lipids, amino acids and nucleotides are remarkably well conserved across taxa (5). Thus, metabolomics studies in model organisms may have translative capacity (6). Recent work in invertebrates, rodents, and zebrafish have increased our mechanistic understanding of the roles of metabolism in cell cycle, tissue growth, and aging (7–10). Other studies have characterized metabolic responses to environmental and intrinsic challenges, such as diet and disease (11, 12). Given that cellular metabolism is highly conserved, profiles in these studies have confirmed signatures of disease progression, characterized drug responses, and identified novel metabolites and pathways for future exploration (13). Yet, few studies have the advantage of examining naturally evolved adaptations that enable animals to overcome environmental challenges (14).

The Mexican tetra, *Astyanax mexicanus* is an ideal species to study evolutionary adaptations to environmental challenges (15). The two morphotypes of *Astyanax* diverged approximately 200,000 years ago as flooding forced river fish from sunny, abundant rivers in Mexico into caves with complete darkness, and food scarcity (16). Cavefish have evolved a suite of metabolic phenotypes to cope with the cave environment, including lower metabolic rate, increased appetite, fat storage, and starvation resistance (17–19). Cavefish are also insulin resistant, hyperglycemic, and exhibit increased caloric intake (20), a feature often associated with decreased longevity. Interestingly, cavefish live long, healthy lives without ill-effects of metabolic disease (20). *Astyanax* may provide natural solutions to overcome the challenges associated with modern diseases, like diabetes (21), and are also an intriguing model for heart regeneration (22). However, the extent of metabolic adaptations associated with cave-life remains unknown, as neither the cavefish metabolome or their response to dietary challenges have been comprehensively described (23).

In this study, we compared the metabolome of surface fish, and two independently-evolved cave populations using untargeted mass spectrometry (MS) of primary metabolites and lipids. Using these exploratory, hypothesis-generating techniques, we characterized the response of energetically important tissues (the liver, muscle, and brain) of each population

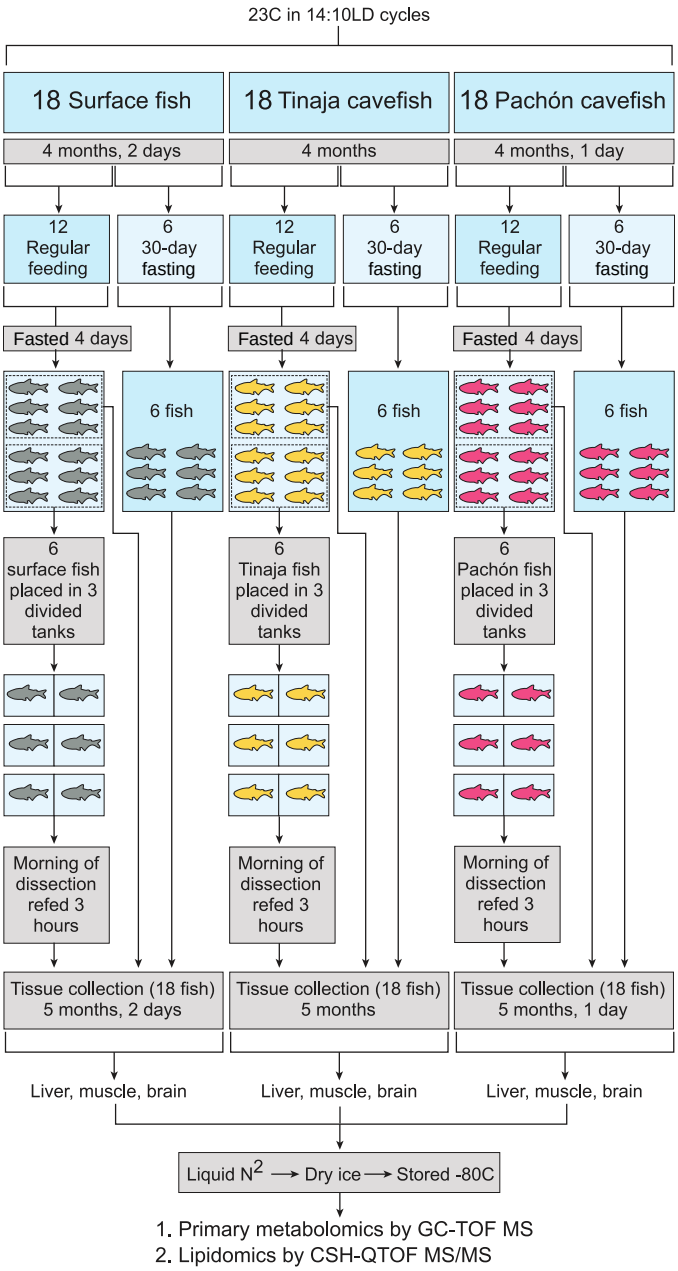


Fig. 1. Experimental setup (A) and PCA for each tissue and metabolite category. Pachón, Tinaja, and surface *A. mexicanus* fry were raised for 4 months and then separated evenly into fasted (30–day) and non–fasted groups. At 4 days prior to collection, non–fasted fish were again divided into two groups (6 fish each) and either fasted for the remaining 4 days (first group) or fasted for 4 days and refed 3 hours prior to collection (second group). Thus, 6 fish were obtained for each of the following conditions: 30–day fasting, 4–day fasting, and 4–day fasting followed by re–feeding.

in short/long-term fasted, and fed conditions. We demonstrate that metabolite profiles in Pachón and Tinaja cavefish are more similar to each other than surface fish in each of the three feeding states, within tissues. We identify metabolic signatures of each tissue, population, and feeding state using O-PLS and a Bayesian generalized linear model (GLM). We constructed inter-population and inter-feeding comparisons and fit separate GLMs to each case (Fig S1) Cavefish exhibit many similarities with human models of obesity and metabolic syndrome, but also differ from these conditions in terms of redox metabolites / by-products and certain classes of lipids. Our results lay the groundwork to explore the mechanistic roles of metabolites and pathways in the adaptation of cavefish and highlight potential roles for certain molecules in resistance to nutrient deprivation and maintenance of robust health under physiological conditions that would be characterized as pathological in humans.

We also provide a shiny app at `https://cavefin.shinyapps.io/shiny` to allow others to explore and visualize our dataset.

Results

Our experimental design aimed to a) characterize the response of the *A. mexicanus* metabolome to different feeding states in energetically expensive tissues and b) utilize comparisons across populations and feeding states to identify

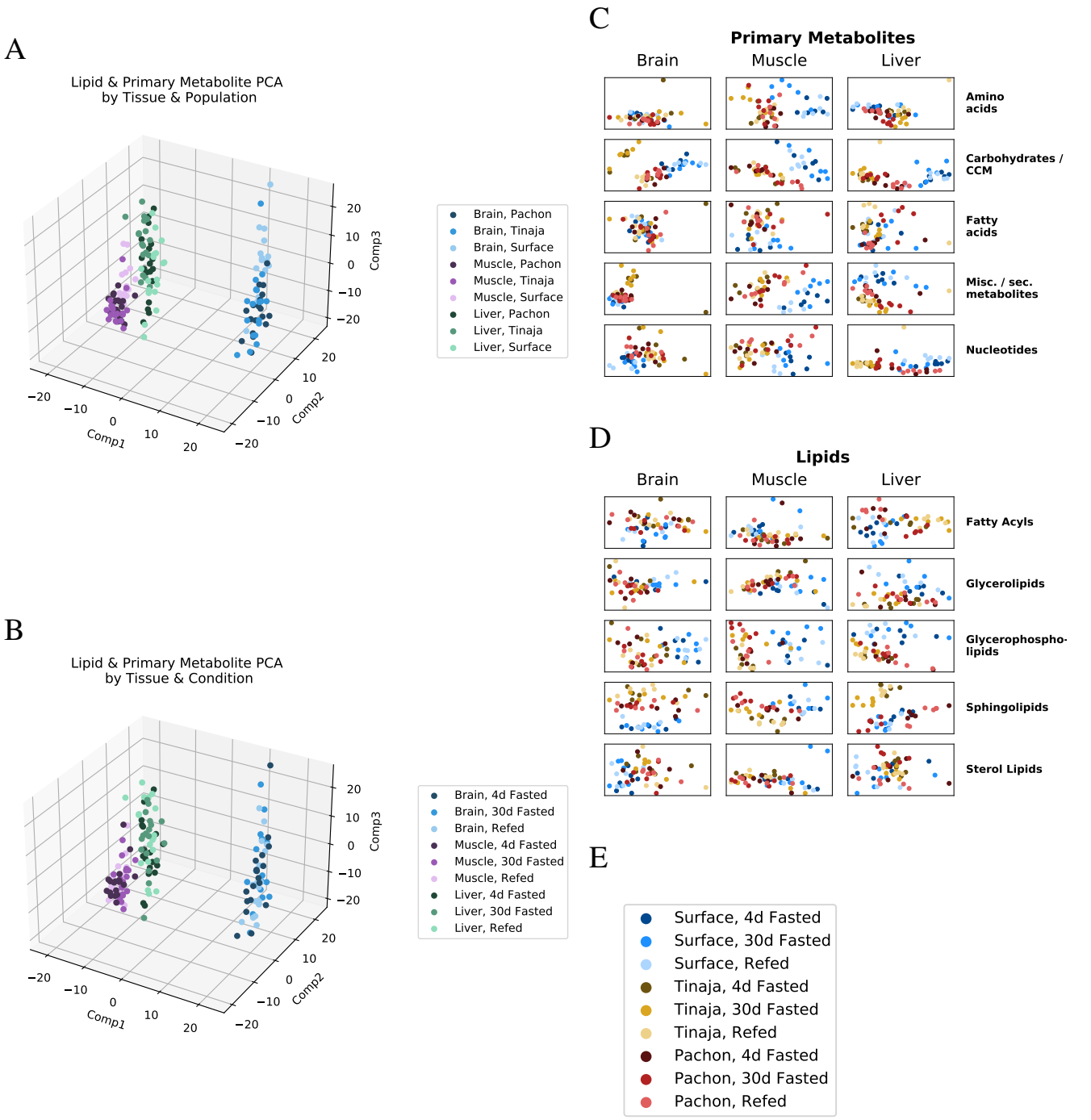


Fig. 2. Global trends in lipid and primary metabolite data. To visualize overall patterns in the metabolome of different experimental groups, we performed principal component analysis (PCA) first on all lipids and primary metabolites (A,B), then on individual categories thereof (C,D). Specifically, PCA of mTIC–normalized log–peak intensities of combined lipid and primary metabolite data (filtered by peaks matching either a KEGG compound or a valid LIPIDMAPS id): (A) Samples were colored according to tissue (blue, purple, green) and population (shading of the primary color). (B) Shading by feeding state instead. (C) First (X-axis) and second (Y-axis) PCA components of primary metabolite data subset by five main categories (see Supplementary Methods) for brain, muscle, and liver of all three populations. (D) PCA of lipid data subset by LIPIDMAPS categorical designation for brain, muscle, and liver of all three populations. (E) Legend for C and D.

metabolites conserved in cavefish populations. Food scarcity is one of the cave's harshest evolutionary pressures. Cavefish have specialized feeding strategies and fat metabolism that helps them thrive in the cave environment (15, 17–19). We raised age-matched offspring of Surface (river) fish, and Pachón and Tinaja cavefish morphs originating from two independent cave colonizations. To understand how the cavefish metabolome adapts to ecologically relevant food challenges, we separated Surface, Pachón and Tinaja populations into three different groups at 4-months: 30-day fasted, 4-day fasted, and "Refed" (fed at 3 hours prior to collection after 4-days without food) (Fig 1). We chose 4-months so that tissue collection would take place prior to sexual maturity (which occurs at 6–9 months). We made all efforts to equalize the

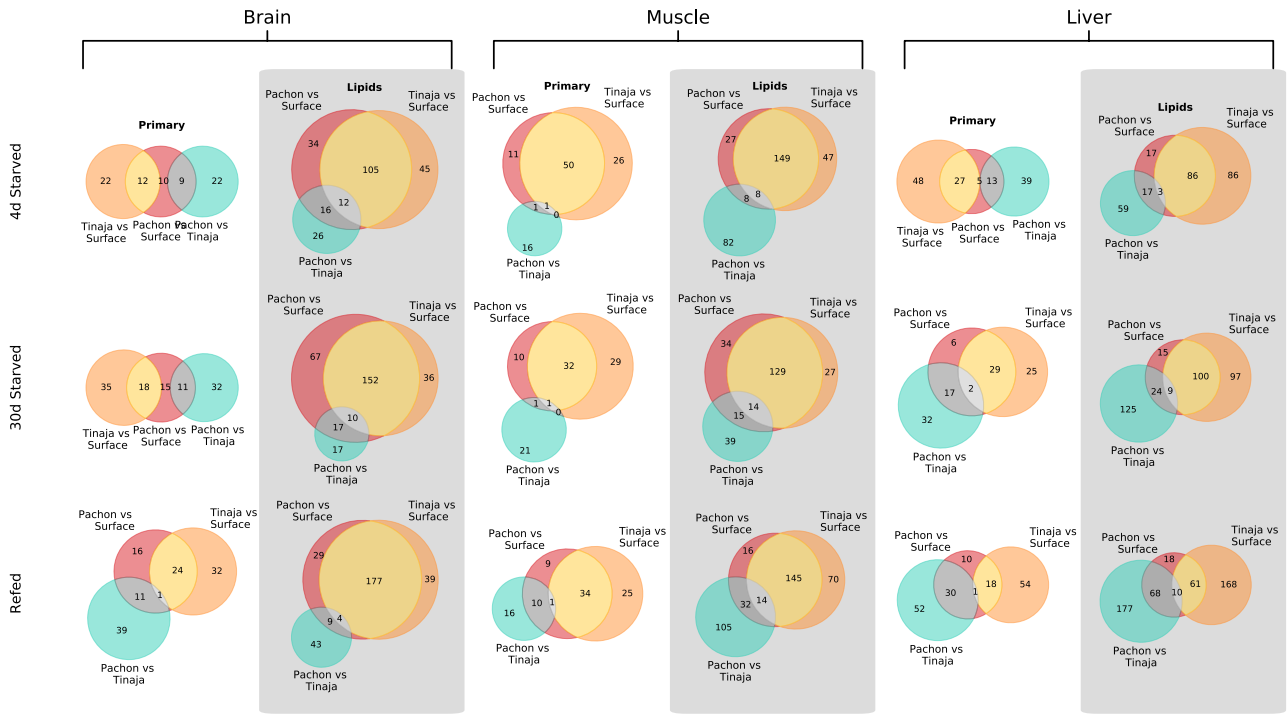


Fig. 3. Common signatures of primary metabolites and lipids in cave populations. In order to determine the degree of agreement between differentially abundant metabolites in Pachón and Tinaja cave populations, we fit an O-PLS / Bayesian GLM statistical model to filtered peak intensity data. The number of metabolites (primary metabolites and lipids respectively) significant at the $p = 0.05$ level are shown. Metabolites in the intersection of two or more sets are constrained to agree on a directional basis (i.e. the differential abundance of a given metabolite in *both* cave populations must be either greater or lesser than surface, not a mixture).

mass and standard length (SL) distributions of fish in each group prior to separation (Fig S2, Table S1, Initial mass and Initial SL).

Our untargeted metabolomics study yielded a total of 174 identified metabolites linked to KEGG (24) / HMDB (25) IDS and 483 identified lipids linked to LIPIDMAPS IDS. We examined the effect of normalizing identified peak values by the total sum of peaks (mTIC) and by sample weight (Fig S3 and found that mTIC is more robust to variations in sample weight. Hence, we employed mTIC-normalized data for the remainder of the analysis.

Figure 2A/B shows that metabolites cluster primarily by tissue, in line with previous studies in mammals (26). Figure 2C/D shows how clustering patterns depend strongly on the chemical classification of identified metabolites. Some lipid and primary metabolite categories show a clear separation between different populations (e.g. carbohydrates, Fig 2C, and glycerophospholipids, Fig 2D, across most tissues), whereas other categories have a less pronounced change (amino acids in the brain, Fig 2C, and fatty acyls in most tissues). In order to quantify separation of feeding states as a function of population and metabolite category, we used a supervised machine learning method based on orthogonal projection of latent structures (O–PLS, Fig S4). We then used O–PLS to remove "orthogonal" variation (27) from each metabolite category and fit a Bayesian logistic regression model to the de–noised data (Supplemental Methods).

Common Metabolite Signatures and Adaptive Response. In general, the metabolome of all three populations shows a large degree of similarity within a given tissue (Fig S5), highlighting the influence of genetic ancestry, even in subpopulations that show markedly different phenotypes.

To quantify the degree to which the metabolomic signature of Pachón and Tinaja cave populations is conserved, we computed the directional intersection of metabolites either increased or decreased (but not a mixture of both) in Pachón and Tinaja with respect to surface according to the GLM (Fig 3). In other words, metabolites in the yellow wedge represent the intersection of metabolites that have significant differential abundance in Pachón (red circle) and Tinaja (gold circle) *and* are increased in both Pachón and Tinaja *or* decreased in both Pachón and Tinaja.

A large convergent signal was displayed between Pachón and Tinaja cave populations, with the differences between cave populations typically being fewer than the differences between each respective cave population and surface. Muscle displayed the highest degree of common primary metabolites, and both muscle and brain display a large overlap of differentially abundant lipids. Given that Pachón and Tinaja represent independent cave populations, adaptive shaping of the metabolome thus appears to have the strongest effect in muscle, the site of energy expenditure during locomotion, and the brain, where lipids may play an important role in signaling. Thus, adaptation appears to play at least a partial role in shaping the metabolome of cave-dwelling *A. mexicanus*. Adaptation is particularly highlighted in certain classes of metabolites, which display extreme changes in both cave populations, as described in the following subsections.

Sugar Phosphate Metabolism. Given the overall similarity at the tissue level for most classes of metabolites (Fig

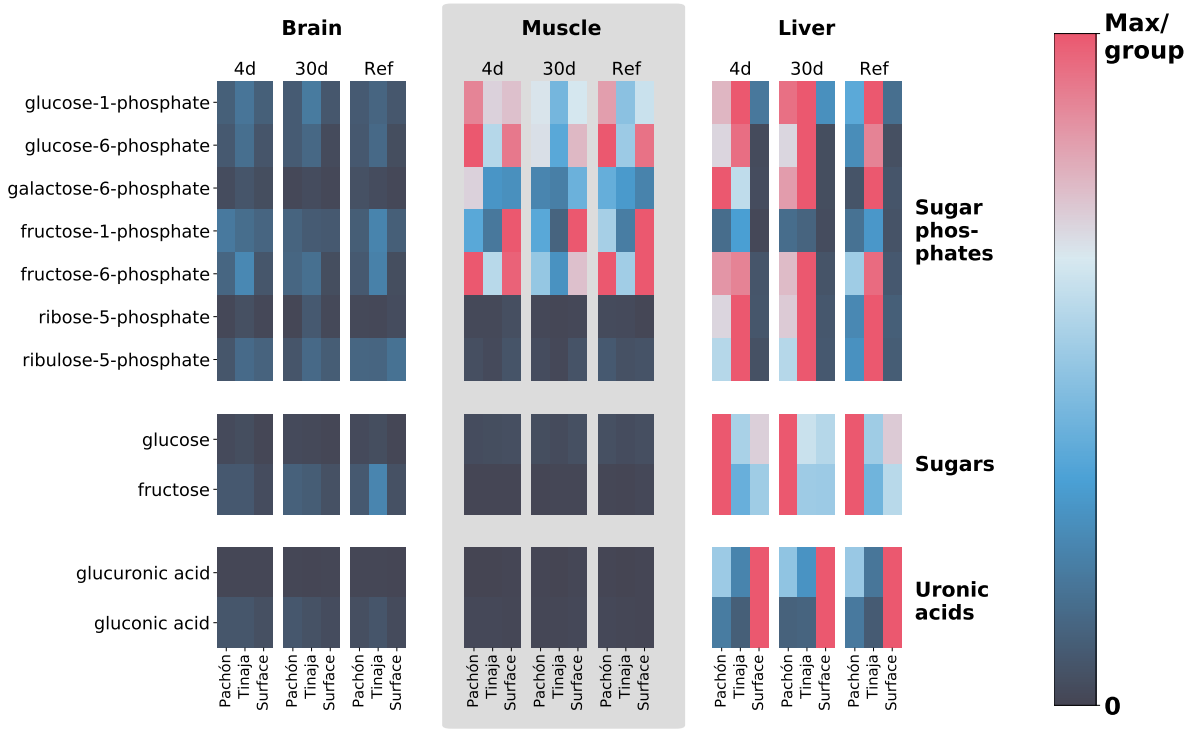


Fig. 4. Extreme alterations to sugar metabolites in cave vs. surface populations. To show differences in cavefish sugar metabolism, we selected 5- and 6-carbon sugars, sugar phosphates, and uronic acids (oxidized forms of sugars that form building blocks of proteoglycans). The row axis shows metabolites within each of these classes, and the column axis shows different populations (short range, labels at bottom), feeding states (medium-range, labels at top) and tissues (long-range, labels at topmost point). Color indicates the mTIC-normalized peak intensity for the average of 6 biological replicates. Red indicates the maximum value for a given row (i.e. across all populations, tissues, and feeding states), whereas navy blue (the bottom of the color bar) corresponds to a peak intensity of zero, not to the minimum value. Thus, dark cells correspond to very little / zero intensity, as opposed to simply corresponding to the minimum intensity within the row. Given that mTIC intensities are "semi–quantitative", red values thus correspond to the most abundant group for a given metabolite, and navy blue corresponds to lack of abundance.

S5), phenotypic differences are likely to be linked to a relatively small subset of the metabolome. We sought to identify metabolites that could be responsible for the drastic change in phenotype of cave populations. Sugars and sugar phosphates are important energy metabolites and hence candidates for adaptations related to resistance to nutrient deprivation. This class of metabolites displays a dramatic change during short- and long-term fasting, particularly in the liver (Fig 4). Hepatic glucose production is derived from gluconeogenesis and glycogenolysis, the latter relying on stored glycogen, which is quickly exhausted during fasting (28), indicating that hepatic gluconeogenesis likely plays a role in sustaining survival under long-term nutrient deprivation in *A. mexicanus*. Surprisingly, surface fish also show stable (albeit generally lower) sugar levels in the liver under different feeding states (Fig 4), indicating that sugar production in the liver may be driven by overall demand rather than supply. This may point to a shift from oxidative to sugar-based metabolism as an energy source in energetically expensive tissues. Cavefish possess a larger amount of body fat (17, 19), and hence have a larger pool of glycerol to serve as a substrate for gluconeogenesis. We find that both cave populations exhibit decreased levels of glycerol in the 30-day fasted state, particularly in Tinaja (Table S2), indicating increased consumption of this intermediate as a substrate for gluconeogenesis may be the source of increased sugar / sugar phosphate abundance in cave populations.

Regardless of the substrates leading to sugar metabolite accumulation, it is clear that a large difference exists between cave and surface populations within this class of metabolites. However, the specifics of this alteration to sugar metabolism appear to be population-specific, with Tinaja showing a large increase in sugar phosphates and Pachón showing an increase in unphosphorylated sugars respectively in short/long-term fasted states in comparison to surface (Fig 4).

Other tissues show a mixed response, with muscle displaying increased levels of most sugar phosphates in Pachón but decreased levels of fructose-1-phosphate in both cave populations with respect to surface. The brain displays low levels of sugar / sugar phosphate metabolites overall but possesses increased sugar metabolite abundance in cave populations for certain metabolites and feeding states (Table S2). In particular, fructose and fructose phosphates tend to be upregulated in the brain, suggesting that reliance on glycolytic metabolites may be a strategy used by cavefish to survive in low nutrient or hypoxic environments, an adaptation that is also present in naked mole rats (29).

While the levels of most simple sugars and sugar phosphates are increased in cavefish with respect to surface, the levels of gluconic acid and glucoronic acid show the opposite pattern (Fig 4). Gluconic acid and glucoronic acid belong to uronic acids, a class of sugar acids that are major building blocks of proteoglycans. It has previously been observed that *A. mexicanus* cave morphs lack advanced glycation end products (20), which are a defining feature of diabetes and are normally associated with chronic hyperglycemia in humans. Al-

tered metabolism of sugar acids in cavefish may play a role in inhibiting excessive protein glycation and the adverse health effects thereof.

Ascorbate. A highly unexpected and unexplained feature of our analysis is the abundance of vitamin C, particularly in its oxidized form dehydroascorbic acid (DHAA), across all tissues in cave populations (Fig 5). Ascorbic acid (AA), the reduced, active form, is also more prevalent in muscle tissue. DHAA can be recycled back to AA using reducing cofactors such as NADH and NADPH, which can in turn be regenerated from the pentose phosphate pathway and TCA cycle using simple sugars, which cavefish have in abundance. For this reason, vitamin C content in food labeling is usually reported as the sum of AA and DHAA (32). Thus, cavefish seem to have a larger total 'pool' of vitamin C (including interconvertible oxidized and reduced forms, Fig 5).

Many cavefish populations exhibit increased appetite and carry an allele of the melanocortin 4 receptor that predisposes them to hyperphagia (17). The increased appetite could cause cavefish to consume more overall food, which could be responsible for the AA/DHAA increase in the refed state. However, this does not pertain to the 30-day fasted state, where AA/DHAA levels are also elevated across all tissues. There is widespread consensus that teleosts, like humans, lack the ability to produce AA endogenously due to the absence of gulonolactone oxidase, which catalyzes the final step in AA biosynthesis (33). In humans, this enzyme is a pseudogene, whereas in teleosts the gene is absent entirely, thought to be lost in the distant evolutionary past. Thus, the additional AA/DHAA supply likely comes from selective reuptake in the kidney, a process that also occurs in humans to conserve AA/DHAA, or it may be produced by commensal microbiota in cavefish. Trace amounts of AA/DHAA in the feed used in the aquatics facility used to house the fish in this experiment may recirculate throughout the water filtration system and be redistributed to all tanks, including those housing fish in the fasted groups. Nevertheless, it remains that even in the case of circulating trace amounts of AA/DHAA, cavefish appear to exhibit selective retainment of AA/DHAA in larger quantities.

The advantages of AA conservation in adaptation to an environment where prolonged starvation is common are self-evident. AA is involved in collagen formation, and its deficiency leads major loss of integrity of connective tissue. Thus, the ability to retain what little ascorbate is present in underground cave environments would confer an enormous survival advantage to fish.

Another factor that could influence the AA/DHAA ratio is the effect of insulin resistance and hyperglycemia on the GLUT family of transporters, particularly GLUT4 in adipose / muscle tissue (32, 34). DHAA competes with glucose for transport across the membrane by GLUT4, whereas AA is taken up by Na^+ transporters. GLUT4 activity is dependent on membrane translocation and this process is dysregulated in diabetes (35). This combination of elevated blood sugar and insulin resistance suggests that GLUT4 could be less active in cavefish and cause DHAA to accumulate in the extracellular space. Finally, Pachón cavefish possess a reduction in neutrophils, one cell type which are normally involved in the uptake of AA and reduction to DHAA, compared to surface (36).

Adaptation to Hypoxic Conditions. Energy metabolism in most organisms can be viewed as a balance between oxidative processes (cellular respiration via oxidative phosphorylation and the electron transport chain) and sugar metabolism, and the relative contributions of these two processes can have important physiological consequences, as in the well-known Warburg effect in cancer. *A. mexicanus* cave morphs have considerably upregulated sugar metabolism (Fig 4), and also display decreased levels of several products of oxidative metabolism. One important metabolite in this category that displays differences in cave populations is α-ketoglutarate (α-KG), which has increased abundance in the liver in all feeding states and in the brain in certain feeding states in both cave populations. α-KG supplementation has been linked to lifespan extension in *C. elegans* (30) and mice (31). Furthermore, uronic acids, the oxidative products of simple sugars, are significantly reduced in the liver of both cave populations (Fig 4, Table S2), suggesting that troglomorphs are characterized by decreased reliance on oxidative metabolism and increased reliance on sugar metabolism. This would enable cavefish to survive in a hypoxic subterranean environment, similar to naked mole rats (29).

Obesity and Inflammation-Related Metabolites. In humans, chronic inflammation of adipose tissue is a common feature of obesity and can often lead to insulin resistance and eventually type 2 diabetes (37). Cave populations of *A. mexicanus* have been previously reported to exhibit pronounced insulin resistance (20), but do not accumulate advanced glycation end products and do not appear to have diminished longevity.

In order to compare the metabolome of *A. mexicanus* cave populations to the known metabolic signatures of obesity (38), we calculated changes in lipid categories (the coarsest abstraction used in LipidMaps; Table S3), classes (a more detailed partitioning scheme used in LipidMaps; Table S4), and, within free fatty acids specifically, the degree of saturation (Table S5). Additionally, we also examined a subset of the lipid data corresponding to individual lipid species with common names (Table S6). The metabolome displays a remarkable overlap with the proinflammatory signature associated with obesity that, in humans, leads to insulin resistance. This signature consists of (1) the elevation of saturated fatty acids (SFAs) in muscle in most feeding states (Table S5), which have a direct and pronounced proinflammatory effect in mammals through the recruitment of macrophages (39) (although the importance of fatty acid release in insulin resistance is disputed (40, 41)), (2) abundance of ceramides in muscle in all feeding states (Table S4), which are known direct mediators of insulin signaling (39). Indeed, the only feeding state for which skeletal muscle did not display increased SFA abundance was 30-day fasting, which could simply indicate the exhaustion of free SFA pools. Additionally,

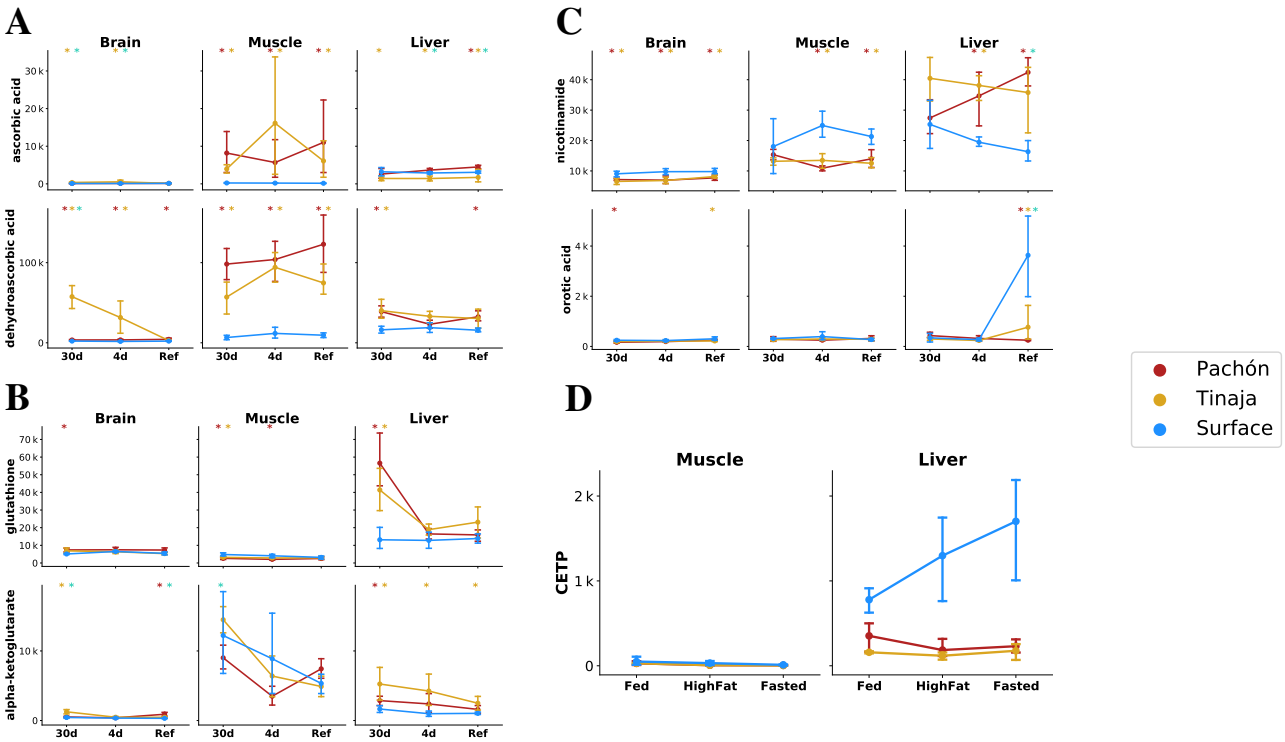


Fig. 5. The cave and surface fish metabolome highlights differences in essential nutrients, redox metabolism, and other biological processes. To analyze adaptive changes in redox metabolism in cave populations, we compared the abundance of several metabolites in our dataset known to influence or be cofactors in various redox reactions. Metabolites exhibiting distinctly altered patterns in cave and surface populations are plotted as mTIC-normalized peak intensities (A,B,C, error bars indicate 2.5 and 97.5 percentiles), whereas gene expression data of CETP, a protein involved in cholesterol ester transport, as expressed in counts per million (CPM,D). The various metabolites shown here and discussed in the text have important roles in many physiological processes, particularly redox metabolism. Asterisks indicate significance at the 0.05 level according to an O-PLS / Bayesian GLM model (Supplementary Methods) for Pachón vs surface (*), Tinaja vs surface (*), and Pachón vs Tinaja (*). (A) Ascorbate and dehydroascorbate (AA/DHAA) are essential vitamins with important antioxidant roles. (B) Glutathione, another antioxidant, is significantly increased in the liver and brain under 30-day fasting. Alpha-ketoglutarate is a longevity-associated tricarboxylic acid cycle (TCA) intermediate (30, 31). (C) Nicotinamide is a precursor to NAD^+ synthesis via a salvage pathway, and exhibits preferential localization to the liver in cavefish. Orotic acid is a metabolite that causes fatty liver disease in rats when added to a chow diet. (D) RNA-Seq data (unpublished) of CETP, a carrier protein for cholesterol esters and triglycerides between lipoproteins. Values on the abscissa (y–axis) are counts per million (CPM) for muscle (left) and liver (right) for older fish under normal feeding ("Fed"), 2-month fasting ("Fasted"), and a high-fat diet ("HighFat").

palmitate, a precursor of ceramide biosynthesis (39), is elevated in muscle in all feeding states (Table S6). Sphingoid bases are significantly more abundant in muscle in all feeding states (Table S3), suggesting generally upregulated sphingolipid biosynthesis in cave populations.

In contrast to proinflammatory metabolites, omega–3 fatty acids (ω-3 FAs) such as DHA and EPA have protective effects against inflammation (37, 39, 42). These molecules bind to the GRP120 receptor on macrophages and adipocytes, and the activated receptor then modulates the activity of PPARγ and ERK (37, 42). ω-3 FAs are less abundant in the liver under 4–day fasting (Table S6) and are generally not upregulated in most feeding states and tissues (Table S6). Thus, ω-3 FAs do not appear to offset for the proinflammatory signature of cavefish SFA and ceramide signatures, suggesting that cavefish possess an alternate compensatory mechanism to prevent chronic tissue inflammation. Overall, cavefish appear to exhibit many metabolic similarities with obesity and health conditions associated with it.

However, this is not a universal trend. Cirulli et al. report a strong association between urate levels and BMI, likely due to insulin resistance interfering with uric acid secretion in the kidney (38). In contrast, cavefish appear to have significantly reduced levels of uric acid in muscle, and in other tissues levels are comparable with surface (except for a small but significant increase in Pachón liver during fasted states, Table S2). Uric acid is increased in humans during prolonged fasting (43), but this trend again does not hold for cavefish, particularly Tinaja. Mannose, which is associated with obesity and insulin resistance in humans (38), was abundant in the Pachón liver in all feeding states, but was reduced in Tinaja compared to surface fish.

Finally, cholesterol esters (Table S6), but not cholesterol (Table S2), were less abundant in cave populations. We investigated factors that might influence levels of cholesterol esters using a separate RNA-Seq dataset (unpublished) containing samples from muscle and liver. Cholesteryl ester transfer protein (CETP), which packages cholesterol into lipoproteins, is downregulated in both Pachón and Tinaja compared to surface fish (Fig 5), suggesting that the decreased abundance of cholesterol esters in cave populations may be due to decreased levels of this important carrier protein. A human variant of CETP associated with decreased serum levels of the protein and larger low-density lipoprotein (LDL) and high-density lipoprotein (HDL) particle sizes has been linked to exceptional longevity (44). The LDL / HDL cholesterol ratio, mediated in part by CETP (45), is a major contributor to risk of atherosclerosis and coronary heart disease (46).

Given the similarities between *A. mexicanus* cave population metabolic signatures and metabolic syndrome, this raises the question of whether cholesterol ester content and CETP may contribute to longevity under obesity-like conditions in these populations.

The presence of free saturated fatty acids, ceramides, and sphingolipids suggests that the link between these metabolites and pathological conditions is not universal, and that organism physiology plays a major role in the progression of metabolic syndrome and diabetes-like pathologies. *A. mexicanus* cavefish do not exhibit shortened lifespan or accumulation of the advanced glycation endproducts (20) normally associated with diabetes.

Studies in other vertebrates demonstrate that high glucose exposure results in an increased cytokine response of macrophages upon lipopolysaccharide exposure but reduces the overall phagocytic rate (47, 48). Our data provides evidence that there is a high glucose concentration in the central organs (brain and liver) of Pachón cavefish in all treatment groups (Fig 4), which leads us to the hypothesis that high glucose concentration is the main driver of the hypersensitivity and reduced phagocytic rate of innate immune cells including macrophages in Pachón cavefish (36). We found further indication that changes in the metabolism are a possible driver of immunological differences between cavefish and surface fish. We found decreased amounts of arachidonic acid in the brain of Pachón cavefish upon 30-day fasting. Arachidonic acid (AraA) provides the basis for the production of membrane phospholipids in the brain (49). AraA, however, is also the substrate of the prostaglandin-endoperoxide synthase 2 (cox-2) that converts AraA to Prostaglandin H2 (PGH2), which in turn is used to produce important mediators of inflammation and a main driver of fever initiation in mammals (50). Cox-2 inhibitors are used to prevent the conversion of AraA to PGH2 and thereby inhibit the progression of inflammation and confer analgesia (51). Cavefish show a decreased expression of cox-2 in the granulocytes and macrophages upon LPS stimulation (36), which might be indicative for a general depletion of the AraA pool in cavefish. In addition, while fish do not develop a fever they show behavioral markers for fever in warmer water for a prolonged time period upon parasite infection (52). Tabin and colleagues (53) showed that cavefish do not show a behavioral fever upon parasite infection, while surface fish show the classic behavioral fever response upon parasite infection. Since conversion of AraA to PGH2 by cox-2 is driving this fever response in fish (52) we hypothesize that the low food abundance in combination with the low parasite diversity in the natural habitat of cavefish would enable the fish reduce the production of AraA, which in turn leads to the inhibition of inflammatory phenotypes and reduced behavioral fever.

Resistance to Nutrient Deprivation. In order to determine the basis of cavefish adaptation to low-nutrient environments, we sought a statistical test that would be sensitive to metabolites that change significantly between refed and short/long-term fasted states and insensitive to metabolites that remain relatively stable across feeding states. We further hypothesized that certain metabolites may have an important role in cave adaptation. Pachón and Tinaja represent independently evolved populations, and we reasoned that a test for parallel adaptation should be selective for metabolites that show the same differential feeding state response pattern across cave populations (e.g. differentially increased in both Pachón and Tinaja fasted states relative to surface). In order to construct this test, we fitted a Bayesian GLM to a linear combination $(P+T)/2-S$ of O-PLS-filtered z-score values (see Supplementary Methods), where P stands for Pachón, T stands for Tinaja, and S stands for surface. We used this test to identify metabolites that might have a role in the fasting response of cavefish, i.e. metabolites that are differentially abundant in cave populations in the fasted versus refed state (Pachón and Tinaja are assigned equal weight), and generally show the opposite pattern in surface. Figure 6 shows the results of this test for 30-day fasting vs. refeeding (A, which corresponds to the most extreme experimental groups), and the two other possible comparisons between feeding states (B/C).

Using this adaptive criterion, we examined which energy metabolites seem to be specifically correlated with different fasting states (Tables S7, S8, and S9 show the results of different fasting state comparisons). Sugar metabolites do not appear to exhibit a strong differential feeding state response. However, long-chain fatty acids such as palmitate and stearate (Fig 6C) do show differential abundance between long- and short-term fasting, suggesting that cavefish may rely on increased usage of fat stores in long-term fasting. Furthermore, analysis of the 30-day fasting response in cavefish liver highlights orotic acid (OA, Fig 6A), an intermediate in pyrimidine synthesis that has been implicated in fatty liver condition (54). OA is suppressed in all feeding states in cave populations, but exhibits a sharp spike in refed surface fish (Fig 5).

Steatosis, or non-alcoholic fatty liver condition, can be caused by reduced FA oxidation / secretion, or increased FA update or lipogenesis (55), with the majority of fat accumulation in the liver coming from uptake of non-esterified fatty acids in humans (56). Lipidomics data indicate that free fatty acid content in the 30-day fasted state is lower for both cave populations with respect to surface (Table S3). The higher body fat of cave populations may be due to increased fatty acid synthesis in the liver under conditions of nutrient availability. Increased lipid accumulation from this process could render cave populations vulnerable to steatosis / fatty liver condition. Given the role of OA in fatty liver condition, OA suppression may therefore have a potential role in preventing excessive lipid accumulation in the liver of cave populations.

Starvation has detrimental effects on an organism in many ways. One detrimental effect is the depletion of antioxidant substances and the resulting oxidative stress through increasing levels of reactive oxygen species (ROS) (57). Studies that focus on the impact of food deprivation on oxidative stress in fish show that prolonged starvation decreases the capacity of fish to ameliorate oxidative stress (57). Glutathione is a major antioxidant that detoxifies ROS and thereby prevents cellular damage from oxidative stress (58). Cavefish

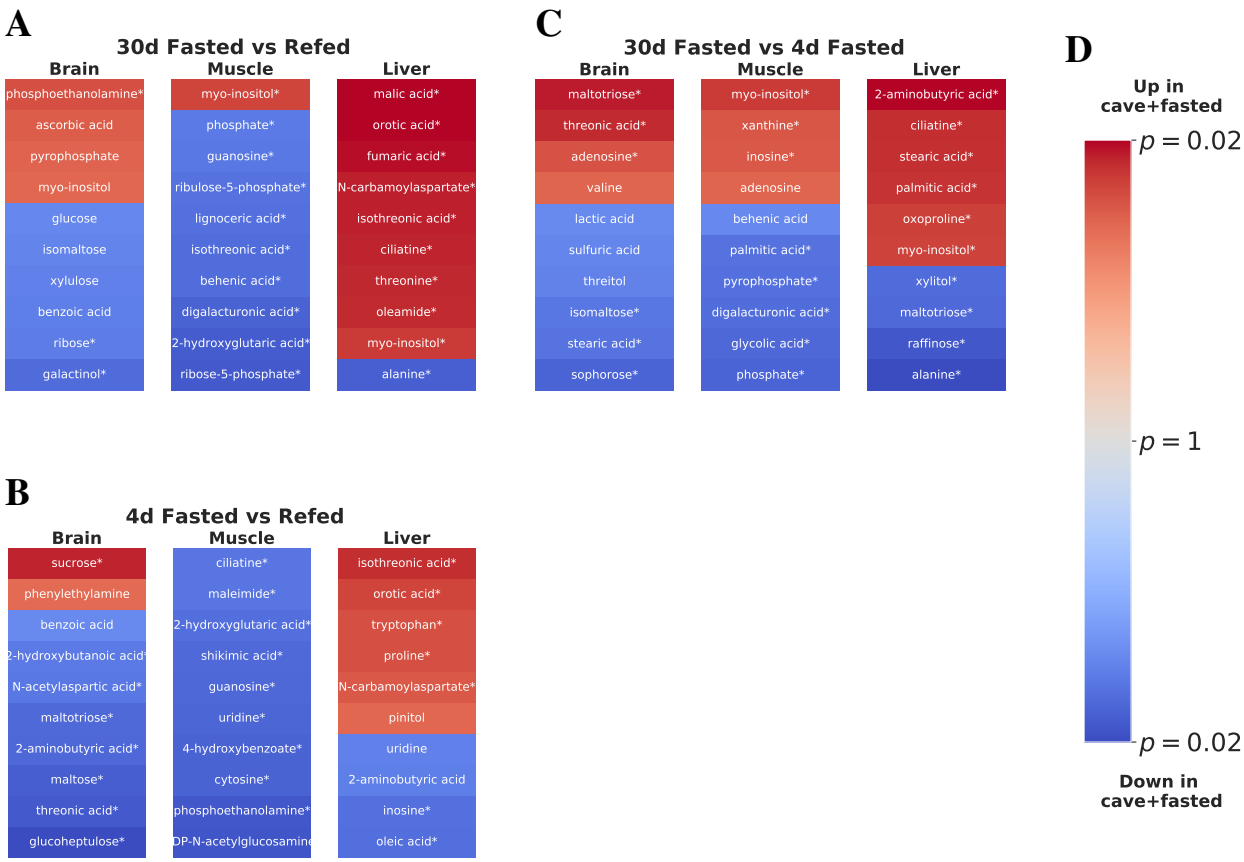


Fig. 6. Adaptive metabolic signature in response to food deprivation in cave populations. To identify metabolites linked to adaptations promoting survival in a nutrient-limited environment shared between cave populations, we fit an O–PLS / GLM statistical model to the response to fasting, i.e. the difference between 30-day fasting and refeeding for Pachón, Tinaja, and surface populations. Specifically, we fitted a Bayesian GLM (see Supplementary Methods) to the linear combination $(P+T)/2-S$, with P, T, and S referring to the normalized z-scores for each Pachón, Tinaja, and surface sample respectively. Coloring in the figure indicates metabolites increased (red) or decreased (blue) in both cave populations in 30–day fasting (i.e. "up" refers to metabolites that are increased in Pachón and Tinaja in the 30–day fasted state with respect to surface), and color intensity corresponds to $\log_{10}$ p-value, with lighter colors indicating less significant p-values and darker colors indicating more significance. The most significant 20 differentially abundant metabolites (regardless of direction) in each tissue for both cave populations with respect to surface are displayed. An asterisk (*) indicates significance at the 0.05 level.

face prolonged periods of nutrient deprivation in their natural environment (17). Adaptation to the cave environment may have led to changes in glutathione metabolism in cavefish to protect against oxidative stress under prolonged fasting. Indeed, in an earlier study we were able to demonstrate that cavefish show an increased expression of genes that are involved in the metabolism of glutathione, which is indicative of an increased stress level compared to surface fish in their natural habitat (59). Here we can confirm the ability of cavefish to respond towards fasting mediated oxidative stress by elevation of reduced glutathione in the liver and brain (Fig 5, Table S2). We did not observe a significant increase of glutathione in the surface fish in the fasted states (Fig 5, Table S2). Given the remarkable differences in glutathione regulation between cavefish and surface fish, *A. mexicanus* is an ideal model to study the oxidative stress response upon states of low nutrient availability.

Discussion

A. mexicanus has been advanced as a model of resilience under ostensibly pathological conditions including hyperglycemia, diabetes (21), and insulin resistance (20). Here, we have provided a large, untargeted study of the metabolome of *A. mexicanus* surface fish and two cave populations in order to investigate the molecular underpinnings of these adaptations.

We were particularly interested in the role of metabolism in cave adaptation of the two *A. mexicanus* cavefish populations in this study: Pachón and Tinaja. We found evidence of cave-mediated positive selection in the form of overlap of changes in metabolite levels (particularly lipids) in certain tissues under all feeding conditions (Fig 3). This suggests that parallel adaptation to cave environments requires satisfying certain common metabolic needs that are an inherent part of the niche. The obvious candidate for this evolutionary conflux is adaptation to a low-nutrient environment. However, metabolic strategies for survival in such environments are not currently well-understood. We found that drastic alterations in energy metabolism, together with shifts in mediators of redox metabolism and ascorbate, an essential vitamin which is lacking in the cave environment, constitute a major feature of cave adaptation in these populations.

Cavefish appear to have substantially altered sugar metabolism, and exhibit higher levels of sugars and sugar phosphates. However, the opposite trend occurs for uronic acids, which are the oxidized forms of simple reducing sug-

ars and can be formed enzymatically or non-enzymatically. This incongruency can be resolved by noting the overall trend to decreased reliance on oxidative metabolism (and enzymes that catalyze oxidative processes) and increased reliance on sugar metabolism. This trend stems from sugars and sugar phosphates, antioxidants such as ascorbate and glutathione, and α-KG (which has been shown to inhibit the electron transport chain in *C. elegans*). Due to drastic fluctuation in oxygen level in the subterranean niche, cavefish may rely on a shift from oxidative to predominantly sugar-derived energy metabolism, as compared to their surface-dwelling cousins. Reduction in uronic acids, which are derived from sugars using oxidative processes, can thus be seen as part of this trend. However, the specific reduction of uronic acids in particular may have an additional survival benefit for cavefish by inhibiting protein glycation and thus preventing accumulation of advanced glycation end products. Further investigation is required to fully understand the evolutionary and physiological implications of these metabolic changes.

The altered sugar metabolism of cavefish may be indicative of a shift from oxidative to sugar-based energy metabolism, similar to certain metabolic adaptations in naked mole rats (29). Further work is required to establish the extent of hypoxic conditions an *A. mexicanus* evolution. However, we also find that certain redox-related metabolites, including α-KG, glutathione, and ascorbate, all exhibit distinctive abundance patterns in cavefish. These patterns may be in response to hypoxia, poor nutrient conditions, differences in metabolic rate, or some other aspect of the cave niche.

Our data indicate that upregulation of glucose and long-chain fatty acid production is a common feature shared by Pachón and Tinaja cave populations, suggesting that certain cave habitats do require considerable changes in energy metabolism. Pachón and Tinaja likely have a greater reliance on fat stores for locomotion, as evidenced by increased SFA content in muscle in fasting and refeeding. The decrease of ω-3 FAs during fasting (Table S6) coupled with the increase of palmitate (Table S9) in long- vs short-term fasting suggests that cavefish metabolism may be preferentially biased towards storing caloric intake as energy-rich saturated fatty acids. Whether cavefish possess adaptations to counteract the deleterious effects of high body fat, such as suppression of orotic acid (a metabolite implicated in steatosis in the liver), requires further investigation.

In summary, *A. mexicanus* troglomorphic populations share many metabolic similarities with human obesity and diabetes mellitus, but also display important differences which may help to explain their resistance to diabetes-like pathologies. We found considerable overlap between the human metabolic signature of obesity and cavefish metabolism. However, we also found important differences in ascorbate, which is known to serve diverse physiological roles, nicotinamide, which is a precursor to NAD^+ synthesis and hence is related to oxidative metabolism, α-ketoglutarate, which has been implicated in longevity in *C. elegans* (30). Unexpectedly, ω-3 FAs and niacin, compounds with anti-inflammatory properties, were not increased in cave populations. The elevated presence of α-ketoglutarate and redistribution of nicotinamide, as well as overall trends toward increased sugar metabolism, suggests decreased overall reliance on oxidative metabolism, which could extend lifespan and reduce the cellular damage associated with inflammation.

Conclusion

Our goals for this study were (1) to provide a comprehensive untargeted study of primary metabolites and lipids in *A. mexicanus*, an extreme-adapted organism with important connections to human health, (2) examine the molecular basis for low-nutrient adaptation in cave-dwelling subpopulations, and (3) identify metabolic changes that might explain *A. mexicanus* longevity in the face of a phenotype with properties linked to obesity and diabetes.

Our findings show that the adaptation to a low nutrient environment in *A. mexicanus* is linked to extreme changes in sugar and fat metabolism, and that increased reliance on these energy sources in the liver insulates other tissues from catabolism under long-term fasting.

All in all, our results highlight the role of *A. mexicanus* as an evolutionary example of extreme metabolism and a model of human obesity, and suggest important roles for certain metabolites in fish and other species.

Acknowledgments

JKM, JP, LO, SX, and NR were supported by institutional funding from the Stowers Institute for Medical Research. Additionally, NR is supported by the Edward Mallinckrodt Foundation, NIH Grant R01 GM127872, DP2AG071466, NSF IOS-1933428, and EDGE award 1923372.

We are deeply grateful to Zachary Zakibe, Andrew Ingalls, Alba Aparicio Fernandez, David Jewell, and Molly Miller for their tireless efforts in maintaining the large fish fish population of the SIMR Cavefish Facility and providing invaluable assistance for this study and others. We would also like to thank Elizabeth Evans, Diana Baumann, M. Shane Merryman, and the SIMR Reptile and Aquatics facility for coordinating and supporting animal research at Stowers and in the Rohner lab in particular. We are grateful to the West Coast Metabolomics Center for providing data acquisition and analysis for this study. Finally, we would like to thank Jaya Krishnan for helpful discussions relating to the manuscript. This pre-print manuscript is based on a LaTeXtemplate kindly provided by Ricardo Henriques.

Methods

Experimental Model and Subject Details

Surface morphs of *Astyanax mexicanus* were reared from offspring of Mexican surface fish collected in the Río Choy. Pachón and Tinaja morphs were reared from fish originating from the Pachón and Tinaja caves. A total of 18 fish from each population were used in experiments. Sex was not determined due to difficulties in determining sex in juvenile *A. mexicanus* fish. This study was approved by the In-

stitutional Animal Care and Use Committee (IACUC) of the Stowers Institute for Medical Research under protocol 2019-084. Animals were euthanized according to an IACUC–approved euthanasia protocols based on American Veterinary Medical Association (AVMA) guidelines using Tricaine mesylate. The method currently in use has been updated to reflect 2020 AVMA guidelines and uses 30 minutes of opercular movement cessation unless a secondary method is employed.

Method Details

Fish husbandry. All *Astyanax* are housed in glass fish tanks on recirculating aquaculture racks (Pentair, Apopka, FL) with a 14:10 LD photoperiod. Each rack system is equipped with mechanical, chemical and biological filtration and UV disinfection. Water quality parameters are monitored and maintained daily as described in previous studies (19, 36). Fish were fed once per day with mysis shrimp and twice per day with Gemma diet. Gemma feed is Protein 59%; Lipids 14%; Fiber 0.2%; Ash 14%; Phosphorus 1.3%; Calcium 1.5%; Sodium 0.7%; Vitamin A 23000 IU/kg; Vitamin D3 2800 IU/kg; Vitamin C 1000 mg/kg; Vitamin E 400 mg/kg. Health examinations of all fish were conducted by aquatics staff twice daily. *Astyanax* colonies are screened biannually for ectoparasites and endoparasites and no pathogens were present at the time of this study. Fish treatment and care was approved by the Institutional Animal Care and Use Committee (IACUC) of the Stowers Institute for Medical Research. NR's institutional authorization for use of *Astyanax mexicanus* in research is 2019-084.

Feeding regimen and tissue collection. Age-matched offspring of Surface, Pachón, and Tinaja populations were reared in similar densities at 23°C in 14:10LD cycles as described previously. Fish were the result of a group mating event within populations. Fish were housed only with members of their population for their entire lives. At 4-months (Tinaja), 4 months and 1 day (Pachón), 4 months and 2 days (Surface), fish of each population were separated into two tanks. 12 fish were separated and starved for 30-days until tissue collection. The 12 fish were maintained on regular feeding schedules until 4-days prior to tissue collection when food was withheld from each population's regular feeding tank. The mass (g) and length (cm) of each fish was recorded at separation. All efforts were made to equalize mass and length distributions in each cohort. On the evening before tissue collection, 6 fish from the 4-day starved tank were separated and placed into three, 3L-tanks. Tanks were divided down the middle such that all 6 fish (2 in each tank) were housed individually. Singly housed fish were refed for 3-hours with 10mg of Gemma 500 on the morning of the dissection day for each population. Dissection occurred at 5 months (October 5th, 2019; Tinaja), 5 months and 1 day (October 6th, 2019; Pachón), 5 months and 2 days (October 7th, 2019; Surface). Fish were re-fed in intervals between 8:30am-12:00pm. At each 3-hour time point, a re-fed fish, a 4-day starved fish, and a 30-day starved fish was euthanized in MS-222. To reduce variability between populations dissected on subsequent days, all dissections took place between 11:30-3pm and were handled identically. Prior to dissection, the final mass and length were recorded for each fish. The liver, muscle, and brain were dissected and placed in 1.5mL plastic tubes. Tissues were flash frozen on liquid nitrogen, transferred to dry ice and stored at -80C. Samples were shipped to West Coast Metabolomics Center on dry ice overnight for analysis.

Sample Preparation. Samples were prepared using the Matyash protocol (60). This procedure allows efficient extraction of lipids in a non-polar methyl tert-butyl ether (MTBE) layer, and extraction of primary metabolites in the polar water/methanol layer (61, 62). From each sample, 4.1 mg of frozen liver or brain tissue (+/- 0.3 mg) or 10.1 mg of muscle tissues (+/- 0.3mg) was weighed and placed into 1.5 mL Eppendorf tubes. Samples were ground prior to extraction using beads with a Spex Sample Prep GenoGrinder with stainless steel 2–3 mm beads for 30 s. 975 μL of ice cold, 3:10 (v/v) MeOH/MTBE + QC mix/CE (22:1) extraction solvent was added to each homogenized sample. Samples were vortexed for 10 sec and shaken for 5 min at 4°C. 188uL room temperature LC/MS water was added and samples vortexed for 20 sec, then centrifuged for 2 min at 14,000 rcf. The upper organic phase was transferred to two separate tubes (350 μL each) for lipidomics (CSH) analysis. The bottom aqueous phase was transferred to two additional tubes (110 μL each) for primary metabolism (GC-TOF) analysis. One tube from each phase was reserved as a backup, the other tube was dried down completely using centrivap. Both were kept at -20°C until ready for analysis. As an additional step prior to GC-TOF analysis, samples were resuspended in 500 μL of degassed, -20°C mixture of acetonitrile (ACN): isopropanol (IPA): water (H2O) (3:3:2, v/v/v). Samples were vortexed for 10 sec and then centrifuged at 14,000 rcf for 2 min. 450 μL of supernatant was transferred to a new tube and concentrated to complete dryness using a Labconco Centruvap cold concentrator.

Primary Metabolite Data Acquisition. Metabolite abundances were quantified by gas-chromatography, time-of-flight mass spectrometry (GC-TOF/MS) using previously established methods (63). Briefly, an Agilent 6890 gas chromatograph (Santa Clara, CA) equipped with a Gerstel automatic linear exchange systems (ALEX) which included a multipurpose sample dual rail and a Gerstel cold injection system (CIS) was used with a Leco Pagasus IV time-of-flight mass spectrometer running Leco ChromaTOF software. The injection temperature was ramped from 50°C to a final temperature of 275°C at a rate of 12°C/s and held for 3 minutes. Injection volume was 0.5 µl with 10 µl/s injection speed on a splitless injector with a purge time of 25 seconds. The liner (Gerstel # 011711-010-00) was changed automatically every 10 samples to reduce sample carryover. The injection syringe was washed three times with 10 μL ethyl acetate before and after each injection. For gas chromatography, a Rtx-5Sil MS column (30 m long, 0.25 mm i.d.) with 0.25 µm 95%

dimethyl 5% diphenyl polysiloxane film was used (Restek, Bellefonte PA). The GC column was equipped with an additional 10 m integrated guard column. 99.9999% pure Helium with a built-in purifier was set at a flow rate of 1mL/minute. The oven temperature was held constant at 50°C for 1minute, ramped at 20°C/minute to 330°C, and then held constant for 5 minutes. The transfer line temperature between gas chromatograph and mass spectrometer was set to 280°C. The mass spectra were acquired at a rate of 17 spectra/second, with a scan mass range of 80–500 Da at an ionization energy of -70 eV, 1800 V detector voltage, and 250°C ion source.

Primary Metabolite Data Processing. Raw GC-TOF MS data files were preprocessed using ChromaTOF version 4.0 without smoothing, a 3 s peak width, baseline subtraction just above the noise level, and automatic mass spectral deconvolution and peak detection at signal/noise (s/n) levels of 5:1 throughout the chromatogram. Results were exported with absolute spectra intensities and further processed by a filtering algorithm implemented in the metabolomics BinBase database (64). The BinBase algorithm (rtx5) used the following settings: validity of chromatogram (107 counts/s), unbiased retention index marker detection (MS similarity > 800, validity of intensity range for high m/z marker ions), retention index calculation by 5th order polynomial regression. Spectra were cut to 5% base peak abundance and matched to database entries from most to least abundant spectra using the following matching filters: retention index window ±2000 units (equivalent to about ±2 s retention time), validation of unique ions and apex masses (unique ion must be included in apexing masses and present at >3% of base peak abundance), mass spectrum similarity must fit criteria dependent on peak purity and signal/noise ratios and a final isomer filter. Failed spectra were automatically entered as new database entries if signal/noise ratios were larger than 25 and mass spectral purity better than 80%. Data was reported as peak height using the unique quantification ion at the specific retention index, unless a different quantification ion was manually set in the BinBase administration software BinView.

Lipid Data Acquisition. Lipid abundances were determined by charged-surface hybrid column-electrospray ionization quadrupole time-of-flight tandem mass spectrometry (CSH-ESI QTOF MS/MS). For positively charged lipids, an Agilent 6530 QTOF mass spectrometer with resolution 10,000 was used and for negatively charged lipids, an Agilent 6550 QTOF mass spectrometer with resolution 20,000 was used. Electrospray ionization was used to ionize column elutants in both positive and negative modes. Compounds were separated using a Waters Acquity ultra-high-pressure, liquid-chromatography charged surface hybrid column (UPLC CSH) C18 (100 mm length × 2.1 mm internal diameter; 1.7 um particles). The conditions in positive mode were as follows: mobile phase A (60:40 acetonitrile:water + 10 mM ammonium formiate + 0.1% formic acid, mobile phase B (90 : 10 isopropanol:acetonitrile + 10 mM ammonium formiate + 0.1% formic acid). The conditions in negative mode were as follows: mobile phase A (60:40 acetonitrile:water + 10 mM ammonium acetate), mobile phase B (90:10 isopropanol:acetonitrile + 10 mM ammonium acetate). 5 μL of each brain, liver and muscle sample was injected in negative mode. 0.5 uL of each brain and liver, and 0.25 uL of muscle samples was injected in positive mode. In both modes, the column temperature was 65°C, at a flow rate of 0.6 mL/minute, an injection temperature of 4 °C, and a gradient of 0 minutes 15%, 0–2 minutes 30%, 2–2.5 minutes 48%, 2.5–11 minutes 82%, 11–11.5 minutes 99%, 11.5–12 minutes 99%, 12–12.1 minutes 15%, and 12.1–15 minutes 15%. The ESI capillary voltage was set to + 3.5 and -3.5 kV, and the collision energy to 25 for positive and negative modes. Mass–to–charge ratios (m/z) were scanned from 60 to 1200 Da and spectra acquired every 2 seconds. Automatic valve switching was used after each injection to reduce sample carryover for highly lipophilic compounds.

A. Lipid Data Processing. Raw lipidomic data were processed using MS-DIAL (65) followed by blank subtractions in Microsoft Excel and data cleanup using MS-FLO (66). Briefly, data were converted to files using Abf Converter. All default parameters were used for processing of MS-DIAL data, except for minimum peak height and width which were adjusted to the instrument. Results are exported from MS-DIAL and a blank reduction is performed for all features which are found in at least one sample. Blank reduction takes the maximum peak height relative to the blank average height and the average of all non-zero peak heights for samples. Duplicates and isotopes are examined using MS-FLO and deleted if confirmed. Peaks were annotated by manually comparing the MS/MS spectra and the accurate masses of precursor ions to spectra in the Fiehn laboratory LipidBlast spectral library (67). Additional peaks are manually curated from sample chromatograms. Manually curation was confirmed by using MassHunter Quant software to verify peak candidates based on peak shape and height reproducibility, and retention time reproducibility in replicate samples. The data were reported as peak heights for the specific quantification ion at the specific retention time.

Quantification and Statistical Analysis

B. Weight change and K-factor calculations. Percent weight change for each fish was calculated using formula 1. Mass and length measurements were recorded at the beginning and end of feeding regimens.

$$\Delta Wt(\%) = (m_{\text{final}} - m_{\text{initial}})/m_{\text{final}} \times 100, \qquad \textbf{(1)}$$

where m is mass. K-factor is a metric that represents both the mass and length of individuals and is frequently used in aquaculture research to assess an animal's physical condition (68). K-factor for each fish was calculated at the beginning of feeding regimens (app. 4 months) and on the day of dissection (30-days later) using the formula (c) below. Percent K-factor change was calculated using formula (d).

$$K = (m/x^3) \times 100, \qquad \textbf{(2)}$$

where x is the standard length.

$$\Delta K(\%) = (K_{\text{final}} - K_{\text{initial}})/K_{\text{final}} \times 100 \quad \textbf{(3)}$$

Data in Supplemental Fig S2 and Table S1 were analyzed and graphically represented using Prism software (GraphPad, Prism version 8.3.0 for Mac, GraphPad Software, San Diego, California USA). Data was first analyzed for normality using four independent methods: D'Agostino-Pearson, Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling. When comparing between more than two groups, data that passed three of four normality tests were analyzed using One-way ANOVA with Tukey correction for multiple comparisons between all groups. Data which failed more than one normality test, was analyzed with Kruskal-Wallis test using Dunn's for multiple comparison correction. The tests used in each figure are reported in the figure legends. p-values less than 0.05 are reported and the level of significance is indicated using the * system (ns, $p > 0.05$; *, $p \leqslant 0.05$; **, $p \leqslant 0.01$; ***, $p \leqslant 0.001$; ****, $p \leqslant 0.0001$).

Further Data Processing. Processed primary metabolite data were vector normalized using mTIC. First, the sum of all peak heights for all identified metabolites, excluding the unknowns, for each sample was calculated. Such peak sums are called "mTIC" and represent the sum of genuine metabolites (identified compounds) in each sample. This method avoids unidentified peaks that could represent potentially non-biological artifacts such as column bleed, contaminants, or routine machine maintenance. mTIC averages for each sample were compared to determine if the variance between samples was significantly different (p < 0.05). Samples were then normalized to the average mTIC "mTICaverage" within populations (Surface, Pachón, or Tinaja) and within organs (brain, muscle, or liver). For example, each biological replicate of the Tinaja brain group was normalized to the average mTIC of all Tinaja brain replicates regardless of feeding state. The equation (a) below was then used to normalize each metabolite (i) of a sample (j). After normalization, data are reported as 'relative semi-quantifications' or normalized peak heights.

$$y_{ij}\,(\text{normalized}) = (x_{ij,raw}/\text{mTIC}_j) \times \overline{\text{mTIC}}, \quad \textbf{(4)}$$

where x_{ij} is the raw peak intensity for metabolite i in sample j, mTIC_j is the average identified peak intensity in sample j, $\overline{\text{mTIC}}$ is the global average identified peak intensity, and y_{ij} is the mTIC-normalized intensity of metabolite i in sample j.

Metabolite Categorization. Metabolites were categorized according to their respective subclass classification in the human metabolite database (25) (if the subclass was absent, we instead used the superclass of the respective metabolite). Metabolite classes with low membership were manually reassigned to arrive at five broad metabolite categories:

- **Carbohydrates and central carbon metabolites (CCM)**. Simple sugars such as glucose, fructose, and various phosphates thereof, as well as core metabolites in glycolysis, gluconeogenesis, the TCA cycle, and the pentose phosphate pathway.
- **Amino acids**. All amino acids and intermediates in amino acid biosynthesis and degradation.
- **Fatty acids**. All free fatty acids, intermediates, and metabolites involved in lipogenesis and β-oxidation.
- **Miscellaneous / secondary metabolites**. Metabolites that do not fall in any of the other categories.
- **Nucleotides**. All nucleotides, nucleosides, nucleobases, and byproducts / intermediates of nucleotide metabolism.

Within each of metabolite category, we further normalized log_{10} peak intensities using z–score normalization prior to performing PCA, O–PLS (described below) or any supervised classification or statistical modeling.

O–PLS. In order to remove sources of variation not useful in discriminating the feeding state of different samples, we used O–PLS (27), a technique commonly used in spectroscopy to correct for systematic variation (27). O–PLS is often applied to raw spectra in order to eliminate the influence of background signals, but here we apply it instead to mTIC normalized peak intensities. Our main use of O–PLS is to remove biological noise that is uncorrelated with feeding state, such as baseline differences or trends among different populations. While z–score normalization already removes many of these artifacts, we observed that O–PLS generally enhanced the predictive accuracy of our PLS classifier. Given an input matrix X of n samples and m spectral features (metabolite peak intensities in our study), and a target matrix Y of classes or measured values (here, the feeding state), the final output of O–PLS (referred to here as X') is again an m by n matrix consisting of X with the systematic variation orthogonal to Y removed.

Characterization of Feeding State Responses. In order to determine which tissues and metabolite categories are most strongly implicated in (1) the starvation response within a given population, (2) differences in metabolite levels between different populations for a given feeding state, we used a simple 1–component PLS classifier trained on the output X' of O–PLS.
The discriminant Q^2 value is a metric of PLS model accuracy and is given by

$$\text{Q}^2 = 1 - \frac{\sum_k (y_k - \hat{y}_k)^2}{\sum_k (y_k - \hat{y}_k)^2}$$

However, we use a truncated version DQ^2 (69), where y_k is replaced by $y'_k = \max(min(y_k, 1), -1)$ and y'_k is used in place of y_k. This metric does not penalize the PLS model for correct predictions that overshoot the target class label.
Using this framework, we employed a two–step model comprised of a O–PLS model followed by a single–component PLS model to discriminate refed versus long term-starved

samples. We trained this combined model on z score-normalized log-transformed data for primary metabolites subdivided into categories. The output of the initial-stage O–PLS model consists of the original data with a PLS component representing 'orthogonal' noise removed. This denoised data was then used to train a one-component PLS classifier on labels representing feeding state. This results in a DQ^2 value for the ability to discriminate refed versus starved states. Finally, an iterative scheme was used to randomly permute the label indices of the input data, resulting in a distribution of DQ^2 values. The significance level of the original predictive DQ^2 value was calculated using a two–tailed survival function of a normal distribution fitted to the DQ^2 values.

Identification of Significant Metabolites. We employed a logistic regression model to identify important features (metabolites, lipids and classes thereof). We were specifically interested in marginal p-values of each individual metabolite, hence we constructed separate single-covariate models for each metabolite or lipid. Models were further based on different types of comparisons: (1) we compared different feeding states within a given population and (2) different populations within a given feeding state. Logistic regression models (and GLMs in general) tend to suffer from complete separation of observed covariates (70). This renders maximum–likelihood estimates of the model parameters impossible. We therefore used the `bayesglm` function of the `arm` R package (71) to obtain estimates for model coefficients, even in the case of complete separation. The `bayesglm` requires specifying a prior distribution. We found that the highly conservative default prior (corresponding to a an assumption that the response to a change in input should typically not exceed roughly ± 5 on the logistic scale, or, equivalently, no typical change in input should cause a shift in probability from 0.01 to 0.50, or 0.50 to 0.99(71)) was sufficient to identify important metabolic changes in our comparisons. However, given the conservative nature of this prior distribution, we did not perform FDR correction.
Fig S1 shows the procedure for training the GLM. We first split the input dataset into two matrices: one containing populations as category labels (bottom left), and one containing feeding states as category labels (upper right). We then subset each of these into the three possible pairwise comparisons from each group, compute z-score-normalized values within the comparison, filter the resulting matrix using O-PLS to remove orthogonal noise, and use the `bayesglm` function to fit a model to the respective comparison for discriminating either populations (within a given feeding state) or feeding states (within a given population). In each case, the GLM consists of a single covariant corresponding to metabolite / lipid peak heights (for individual metabolites) or classes of metabolites / lipids (for Table S4 and S3).

Differences between Feeding States and Shared Metabolites. To identify metabolites that might play a role in cave adaptation, we sought to fit logistic regression to an input capturing the difference between refed and and starved samples and differences between cave and surface populations simultaneous. We implemented this using the following formula:

$$x = (P + T)/2 - S, \tag{5}$$

where P, T, and S are z-score normalized mTIC peak intensities for starved vs. refed samples. In general, these vectors have length 12 (6 refed and 6 starved samples). A Bayesian logistic regression model was then fitted to the x vector for each metabolite as before, with each element of the response vector labelled accordingly (starved or refed).

Additional Resources

The Shiny App for this study may be accessed at https://cavefin.shinyapps.io/shiny.

Bibliography

1. Caroline H Johnson, Julijana Ivanisevic, and Gary Siuzdak. Metabolomics: beyond biomarkers and towards mechanisms. *Nature reviews Molecular cell biology*, 17(7):451–459, 2016.
2. Markus M Rinschen, Julijana Ivanisevic, Martin Giera, and Gary Siuzdak. Identification of bioactive metabolites using activity metabolomics. *Nature Reviews Molecular Cell Biology*, 20(6):353–367, 2019.
3. Thomas Nägele. Linking metabolomics data to underlying metabolic regulation. *Frontiers in Molecular Biosciences*, 1:22, 2014.
4. Carlos Guijas, J Rafael Montenegro-Burke, Benedikt Warth, Mary E Spilker, and Gary Siuzdak. Metabolomics activity screening for identifying metabolites that modulate phenotype. *Nature biotechnology*, 36(4):316–320, 2018.
5. José M Peregrín-Alvarez, Chris Sanford, and John Parkinson. The conservation and evolutionary modularity of metabolism. *Genome biology*, 10(6):1–17, 2009.
6. Akhila Rajan and Norbert Perrimon. Of flies and men: insights on organismal metabolism from fruit flies. *BMC biology*, 11(1):38, 2013.
7. Pernille Sarup, Simon Metz Mariendal Pedersen, Niels Chr Nielsen, Anders Malmendal, and Volker Loeschcke. The metabolic profile of long-lived drosophila melanogaster. *PLoS One*, 7(10):e47461, 2012.
8. Janna Hastings, Abraham Mains, Bhupinder Virk, Nicolas Rodriguez, Sharlene Murdoch, Juliette Pearce, Sven Bergmann, Nicolas Le Novère, and Olivia Casanueva. Multi-omics and genome-scale modeling reveal a metabolic shift during c. elegans aging. *Frontiers in molecular biosciences*, 6:2, 2019.
9. Phan Nguyen Thuy An, Masamitsu Yamaguchi, Takeshi Bamba, and Eiichiro Fukusaki. Metabolome analysis of drosophila melanogaster during embryogenesis. *PloS one*, 9(8): e99519, 2014.
10. Masayuki Oginuma, Philippe Moncuquet, Fengzhu Xiong, Edward Karoly, Jérome Chal, Karine Guevorkian, and Olivier Pourquié. A gradient of glycolytic activity coordinates fgf and wnt signaling during elongation of the body axis in amniote embryos. *Developmental cell*, 40(4):342–353, 2017.
11. Matthew J Laye, ViLinh Tran, Dean P Jones, Pankaj Kapahi, and Daniel EL Promislow. The effects of age and dietary restriction on the tissue-specific metabolome of drosophila. *Aging cell*, 14(5):797–808, 2015.
12. Daniel Wilinski, Jasmine Winzeler, William Duren, Jenna L Persons, Kristina J Holme, Johan Mosquera, Morteza Khabiri, Jason M Kinchen, Peter L Freddolino, Alla Karnovsky, et al. Rapid metabolic shifts occur during the transition between hunger and satiety in drosophila melanogaster. *Nature communications*, 10, 2019.
13. Laura Palanker Musselman, Jill L Fink, and Thomas J Baranski. Coa protects against the deleterious effects of caloric overload in drosophila. *Journal of lipid research*, 57(3):380–387, 2016.
14. Mélanie Viltard, Sylvère Durand, Maria Pérez-Lanzón, Fanny Aprahamian, Deborah Lefevre, Christine Leroy, Frank Madeo, Guido Kroemer, and Gérard Friedlander. The metabolomic signature of extreme longevity: Naked mole rats versus mice. *Aging (Albany NY)*, 11(14):4783, 2019.
15. William R Jeffery. Astyanax surface and cave fish morphs. *EvoDevo*, 11(1):1–10, 2020.
16. Adam Herman, Yaniv Brandvain, James Weagley, William R Jeffery, Alex C Keene, Thomas JY Kono, Helena Bilandžija, Richard Borowsky, Luis Espinasa, Kelly O'Quin, et al. The role of gene flow in rapid and repeated evolution of cave-related traits in mexican tetra, astyanax mexicanus. *Molecular ecology*, 27(22):4397–4416, 2018.
17. Ariel C Aspiras, Nicolas Rohner, Brian Martineau, Richard L Borowsky, and Clifford J Tabin. Melanocortin 4 receptor mutations contribute to the adaptation of cavefish to nutrient-poor conditions. *Proceedings of the National Academy of Sciences*, 112(31):9668–9673, 2015.
18. Kathrin Hüppop. Oxygen consumption of astyanax fasciatus (characidae, pisces): a comparison of epigean and hypogean populations. *Environmental Biology of Fishes*, 17(4): 299–308, 1986.
19. Shaolei Xiong, Jaya Krishnan, Robert Peuß, and Nicolas Rohner. Early adipogenesis contributes to excess fat accumulation in cave populations of astyanax mexicanus. *Developmental biology*, 441(2):297–304, 2018.

20. Misty R Riddle, Ariel C Aspiras, Karin Gaudenz, Robert Peuß, Jenny Y Sung, Brian Martineau, Megan Peavey, Andrew C Box, Julius A Tabin, Suzanne McGaugh, et al. Insulin resistance in cavefish as an adaptation to a nutrient-limited environment. *Nature*, 555(7698): 647–651, 2018.
21. Jaya Krishnan and Nicolas Rohner. Sweet fish: Fish models for the study of hyperglycemia and diabetes. *Journal of diabetes*, 11(3):193–203, 2019.
22. William T Stockdale, Madeleine E Lemieux, Abigail C Killen, Juanjuan Zhao, Zhilian Hu, Joey Riepsaame, Noémie Hamilton, Tetsuhiro Kudoh, Paul R Riley, Ronny van Aerle, et al. Heart regeneration in the mexican cavefish. *Cell reports*, 25(8):1997–2007, 2018.
23. Lucie Marandel, Elisabeth Plagnes-Juan, Michael Marchand, Therese Callet, Karine Dias, Frederic Terrier, Stéphane Père, Louise Vernier, Stephane Panserat, and Sylvie Rétaux. Nutritional regulation of glucose metabolism-related genes in the emerging teleost model mexican tetra surface fish: a first exploration. *Royal Society open science*, 7(2):191853, 2020.
24. Minoru Kanehisa and Susumu Goto. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30, 2000.
25. David S Wishart, Yannick Djoumbou Feunang, Ana Marcu, An Chi Guo, Kevin Liang, Rosa Vázquez-Fresno, Tanvir Sajed, Daniel Johnson, Carin Li, Naama Karu, et al. Hmdb 4.0: the human metabolome database for 2018. *Nucleic acids research*, 46(D1):D608–D617, 2018.
26. Siming Ma, Sun Hee Yim, Sang-Goo Lee, Eun Bae Kim, Sang-Rae Lee, Kyu-Tae Chang, Rochelle Buffenstein, Kaitlyn N Lewis, Thomas J Park, Richard A Miller, et al. Organization of the mammalian metabolome according to organ function, lineage specialization, and longevity. *Cell metabolism*, 22(2):332–343, 2015.
27. Johan Trygg and Svante Wold. Orthogonal projections to latent structures (o-pls). *Journal of Chemometrics: A Journal of the Chemometrics Society*, 16(3):119–128, 2002.
28. Hye-Sook Han, Geon Kang, Jun Seok Kim, Byeong Hoon Choi, and Seung-Hoi Koo. Regulation of glucose metabolism from a liver-centric perspective. *Experimental & molecular medicine*, 48(3):e218–e218, 2016.
29. Thomas J Park, Jane Reznick, Bethany L Peterson, Gregory Blass, Damir Omerbašić, Nigel C Bennett, P Henning JL Kuich, Christin Zasada, Brigitte M Browe, Wiebke Hamann, et al. Fructose-driven glycolysis supports anoxia resistance in the naked mole-rat. *Science*, 356(6335):307–311, 2017.
30. Randall M Chin, Xudong Fu, Melody Y Pai, Laurent Vergnes, Heejun Hwang, Gang Deng, Simon Diep, Brett Lomenick, Vijaykumar S Meli, Gabriela C Monsalve, et al. The metabolite α-ketoglutarate extends lifespan by inhibiting atp synthase and tor. *Nature*, 510(7505): 397–401, 2014.
31. Azar Asadi Shahmirzadi, Daniel Edgar, Chen-Yu Liao, Yueh-Mei Hsu, Mark Lucanic, Arash Asadi Shahmirzadi, Christopher D Wiley, Garbo Gan, Dong Eun Kim, Herbert G Kasler, et al. Alpha-ketoglutarate, an endogenous metabolite, extends lifespan and compresses morbidity in aging mice. *Cell Metabolism*, 32(3):447–456, 2020.
32. John X Wilson. The physiological role of dehydroascorbic acid. *FEBS letters*, 527(1-3):5–9, 2002.
33. Biyun Ching, Shit F Chew, and Yuen K Ip. Ascorbate synthesis in fishes: A review. *IUBMB life*, 67(2):69–76, 2015.
34. Shaohui Huang and Michael P Czech. The glut4 glucose transporter. *Cell metabolism*, 5 (4):237–252, 2007.
35. Javier R Jaldin-Fincati, Martin Pavarotti, Scott Frendo-Cumbo, Philip J Bilan, and Amira Klip. Update on glut4 vesicle traffic: a cornerstone of insulin action. *Trends in Endocrinology & Metabolism*, 28(8):597–611, 2017.
36. Robert Peuß, Andrew C Box, Shiyuan Chen, Yongfu Wang, Dai Tsuchiya, Jenna L Persons, Alexander Kenzior, Ernesto Maldonado, Jaya Krishnan, Jörn P Scharsack, et al. Adaptation to low parasite abundance affects immune investment and immunopathological responses of cavefish. *Nature Ecology & Evolution*, pages 1–15, 2020.
37. Vivian A Paschoal, Evelyn Walenta, Saswata Talukdar, Ariane R Pessentheiner, Olivia Osborn, Nasun Hah, Tyler J Chi, George L Tye, Aaron M Armando, Ronald M Evans, et al. Positive reinforcing mechanisms between gpr120 and pparγ modulate insulin sensitivity. *Cell Metabolism*, 2020.
38. Elizabeth T Cirulli, Lining Guo, Christine Leon Swisher, Naisha Shah, Lei Huang, Lori A Napier, Ewen F Kirkness, Tim D Spector, C Thomas Caskey, Bernard Thorens, et al. Profound perturbation of the metabolome in obesity is associated with health risk. *Cell metabolism*, 29(2):488–500, 2019.
39. Christopher K Glass and Jerrold M Olefsky. Inflammation and lipid signaling in the etiology of insulin resistance. *Cell metabolism*, 15(5):635–645, 2012.
40. Pauline Morigny, Marianne Houssier, Aline Mairal, Claire Ghilain, Etienne Mouisel, Fadila Benhamed, Bernard Masri, Emeline Recazens, Pierre-Damien Denechaud, Geneviève Tavernier, et al. Interaction between hormone-sensitive lipase and chrebp in fat cells controls insulin sensitivity. *Nature Metabolism*, 1(1):133–146, 2019.
41. Fredrik Karpe, Julian R Dickmann, and Keith N Frayn. Fatty acids, obesity, and insulin resistance: time for a reevaluation. *Diabetes*, 60(10):2441–2449, 2011.
42. David Young Oh, Saswata Talukdar, Eun Ju Bae, Takeshi Imamura, Hidetaka Morinaga, WuQiang Fan, Pingping Li, Wendell J Lu, Steven M Watkins, Jerrold M Olefsky, et al. Gpr120 is an omega-3 fatty acid receptor mediating potent anti-inflammatory and insulin-sensitizing effects. *Cell*, 142(5):687–698, 2010.
43. Takayuki Teruya, Romanas Chaleckis, Junko Takada, Mitsuhiro Yanagida, and Hiroshi Kondoh. Diverse metabolic reactions activated during 58-hr fasting are revealed by non-targeted metabolomic analysis of human blood. *Scientific reports*, 9(1):1–11, 2019.
44. Nir Barzilai, Gil Atzmon, Clyde Schechter, Ernst J Schaefer, Adrienne L Cupples, Richard Lipton, Suzanne Cheng, and Alan R Shuldiner. Unique lipoprotein phenotype and genotype associated with exceptional longevity. *Jama*, 290(15):2030–2040, 2003.
45. JK Christison, Kerry-Anne Rye, and Roland Stocker. Exchange of oxidized cholesteryl linoleate between ldl and hdl mediated by cholesteryl ester transfer protein. *Journal of lipid research*, 36(9):2017–2026, 1995.
46. Margaret E Brousseau, Ernst J Schaefer, Megan L Wolfe, LeAnne T Bloedon, Andres G Digenio, Ronald W Clark, James P Mancuso, and Daniel J Rader. Effects of an inhibitor of cholesteryl ester transfer protein on hdl cholesterol. *New England Journal of Medicine*, 350 (15):1505–1515, 2004.
47. Thais Soprani Ayala, Fernando Henrique Galvão Tessaro, Grasielle Pereira Jannuzzi, Leonardo Mendes Bella, Karen Spadari Ferreira, and Joilson O Martins. High glucose environments interfere with bone marrow-derived macrophage inflammatory mediator release, the tlr4 pathway and glucose metabolism. *Scientific reports*, 9(1):1–15, 2019.
48. Sofia Pavlou, Jaime Lindsay, Rebecca Ingram, Heping Xu, and Mei Chen. Sustained high glucose exposure sensitizes macrophage responses to cytokine stimuli but reduces their phagocytic activity. *BMC immunology*, 19(1):24, 2018.
49. Francesca Bosetti. Arachidonic acid metabolism in brain physiology and pathology: lessons from genetically altered mouse models. *Journal of neurochemistry*, 102(3):577–586, 2007.
50. Alexandre A Steiner, Andrei I Ivanov, Jordi Serrats, Hiroshi Hosokawa, Allison N Phayre, Jared R Robbins, Jennifer L Roberts, Shigeo Kobayashi, Kiyoshi Matsumura, Paul E Sawchenko, et al. Cellular and molecular bases of the initiation of fever. *PLoS Biol*, 4 (9):e284, 2006.
51. Aristides G Eliopoulos, Calin D Dumitru, Chun-Chi Wang, Jeonghee Cho, and Philip N Tsichlis. Induction of cox-2 by lps in macrophages is regulated by tpl2-dependent creb activation signals. *The EMBO journal*, 21(18):4831–4840, 2002.
52. Krzysztof Rakus, Maygane Ronsmans, and Alain Vanderplasschen. Behavioral fever in ectothermic vertebrates. *Developmental & Comparative Immunology*, 66:84–91, 2017.
53. Julius A Tabin, Ariel Aspiras, Brian Martineau, Misty Riddle, Johanna Kowalko, Richard Borowsky, Nicolas Rohner, and Clifford J Tabin. Temperature preference of cave and surface populations of astyanax mexicanus. *Developmental biology*, 441(2):338–344, 2018.
54. Susan B Standerfer and Philip Handler. Fatty liver induced by orotic acid feeding. *Proceedings of the Society for Experimental Biology and Medicine*, 90(1):270–271, 1955.
55. Seung-Hoi Koo. Nonalcoholic fatty liver disease: molecular mechanisms for the hepatic steatosis. *Clinical and molecular hepatology*, 19(3):210, 2013.
56. Kerry L Donnelly, Coleman I Smith, Sarah J Schwarzenberg, Jose Jessurun, Mark D Boldt, Elizabeth J Parks, et al. Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *The Journal of clinical investigation*, 115(5): 1343–1351, 2005.
57. Miriam Furne and Ana Sanz. Starvation in fish – sturgeon and rainbow trout as examples. In V. R. Preedy and V. B. Patel, editors, *Handbook of Famine, Starvation, and Nutrient Deprivation*. Springer International Publishing AG, 2017.
58. Alfonso Pompella, Athanase Visvikis, Aldo Paolicchi, Vincenzo De Tata, and Alessandro F Casini. The changing faces of glutathione, a cellular protagonist. *Biochemical pharmacology*, 66(8):1499–1503, 2003.
59. Jaya Krishnan, Jenna L Persons, Robert Peuß, Huzaifa Hassan, Alexander Kenzior, Shaolei Xiong, Luke Olsen, Ernesto Maldonado, Johanna E Kowalko, and Nicolas Rohner. Comparative transcriptome analysis of wild and lab populations of astyanax mexicanus uncovers differential effects of environment and morphotype on gene expression. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 2020.
60. Vitali Matyash, Gerhard Liebisch, Teymuras V Kurzchalia, Andrej Shevchenko, and Dominik Schwudke. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *Journal of lipid research*, 49(5):1137–1146, 2008.
61. Oliver Fiehn. Metabolomics by gas chromatography–mass spectrometry: Combined targeted and untargeted profiling. *Current protocols in molecular biology*, 114(1):30–4, 2016.
62. Tomas Cajka and Oliver Fiehn. Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. *TrAC Trends in Analytical Chemistry*, 61:192–206, 2014.
63. Oliver Fiehn, Gert Wohlgemuth, Martin Scholz, Tobias Kind, Do Yup Lee, Yun Lu, Stephanie Moon, and Basil Nikolau. Quality control for plant metabolomics: reporting msi-compliant studies. *The Plant Journal*, 53(4):691–704, 2008.
64. Kirsten Skogerson, Gert Wohlgemuth, Dinesh K Barupal, and Oliver Fiehn. The volatile compound binbase mass spectral database. *BMC bioinformatics*, 12(1):321, 2011.
65. Hiroshi Tsugawa, Tomas Cajka, Tobias Kind, Yan Ma, Brendan Higgins, Kazutaka Ikeda, Mitsuhiro Kanazawa, Jean VanderGheynst, Oliver Fiehn, and Masanori Arita. Ms-dial: data-independent ms/ms deconvolution for comprehensive metabolome analysis. *Nature methods*, 12(6):523–526, 2015.
66. Brian C DeFelice, Sajjan Singh Mehta, Stephanie Samra, Tomas Cajka, Benjamin Wancewicz, Johannes F Fahrmann, and Oliver Fiehn. Mass spectral feature list optimizer (ms-flo): a tool to minimize false positive peak reports in untargeted liquid chromatography–mass spectroscopy (lc-ms) data processing. *Analytical chemistry*, 89(6):3250–3255, 2017.
67. Tobias Kind, Kwang-Hyeon Liu, Do Yup Lee, Brian DeFelice, John K Meissen, and Oliver Fiehn. Lipidblast in silico tandem mass spectrometry database for lipid identification. *Nature methods*, 10(8):755–758, 2013.
68. TS Imam, U Bala, ML Balarabe, TI Oyeyi, et al. Length-weight relationship and condition factor of four fish species from wasai reservoir in kano, nigeria. *African Journal of General Agriculture*, 6(3):125–130, 2010.
69. Johan A Westerhuis, Ewoud JJ van Velzen, Huub CJ Hoefsloot, and Age K Smilde. Discriminant q 2 (dq 2) for improved discrimination in plsda models. *Metabolomics*, 4(4):293–296, 2008.
70. Dongming Huang, Nathan Stein, Donald B Rubin, and SC Kou. Catalytic prior distributions with application to generalized linear models. *Proceedings of the National Academy of Sciences*, 117(22):12004–12010, 2020.
71. Andrew Gelman, Aleks Jakulin, Maria Grazia Pittau, Yu-Sung Su, et al. A weakly informative default prior distribution for logistic and other regression models. *The annals of applied statistics*, 2(4):1360–1383, 2008.

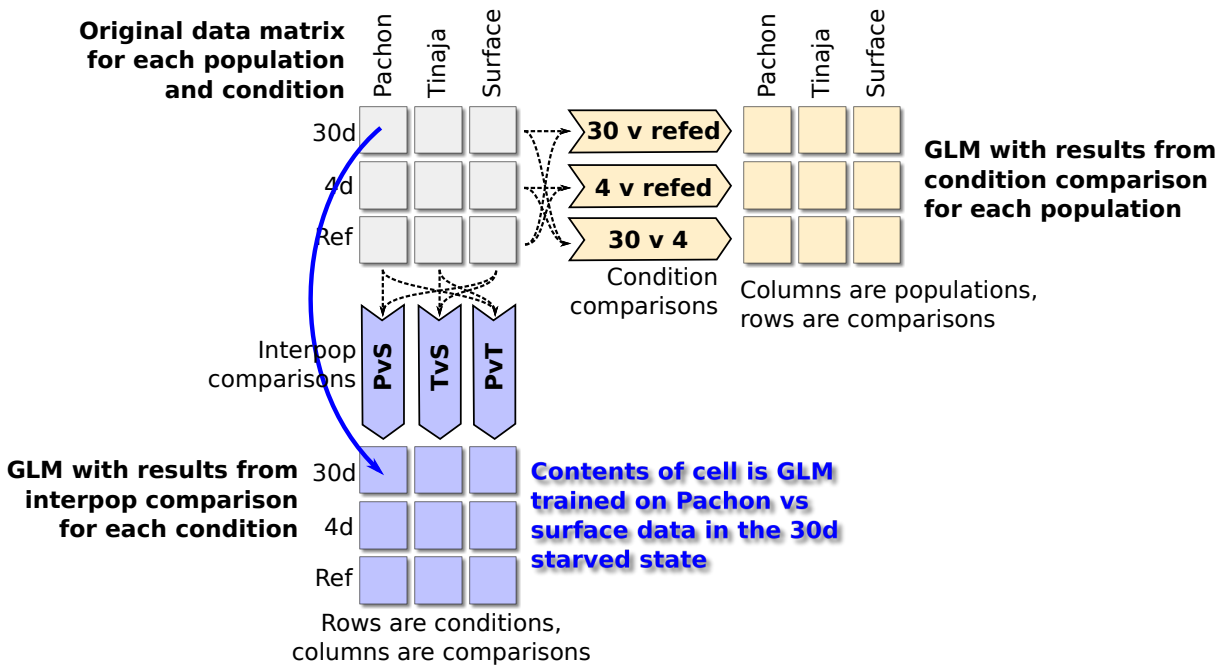


Fig. S1. Schematic depiction of comparisons used for fitting GLM parameters. A Bayesian logistic regression model was used to identify significant features (metabolites and categories) among different groups. To train the model, we first split the entire dataset (upper-left matrix) into two components of pairwise comparisons: a matrix containing category labels for every pairwise comparison of the three populations in our study (lower-left, for example the PvS comparison contains category labels for Pachón and surface and excludes data from Tinaja), and a separate matrix containing category labels for every pairwise comparison of feeding conditions. We then trained a Bayesian logistic regression model using the `bayesglm` R package using the appropriate category labels for each comparison. Predictors were either O–PLS–filtered metabolite peaks or metabolite classes (e.g. fatty acids, nucleotides, etc.).

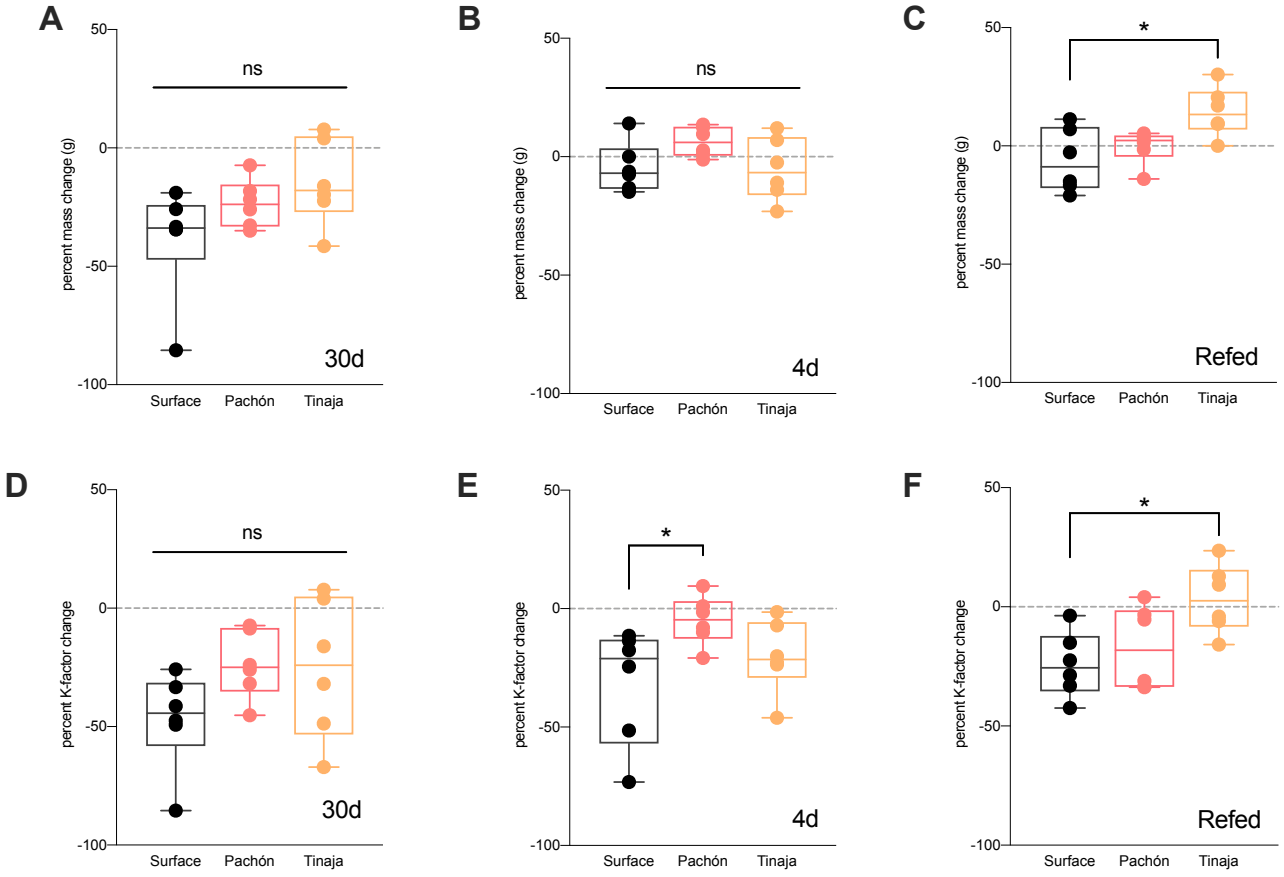
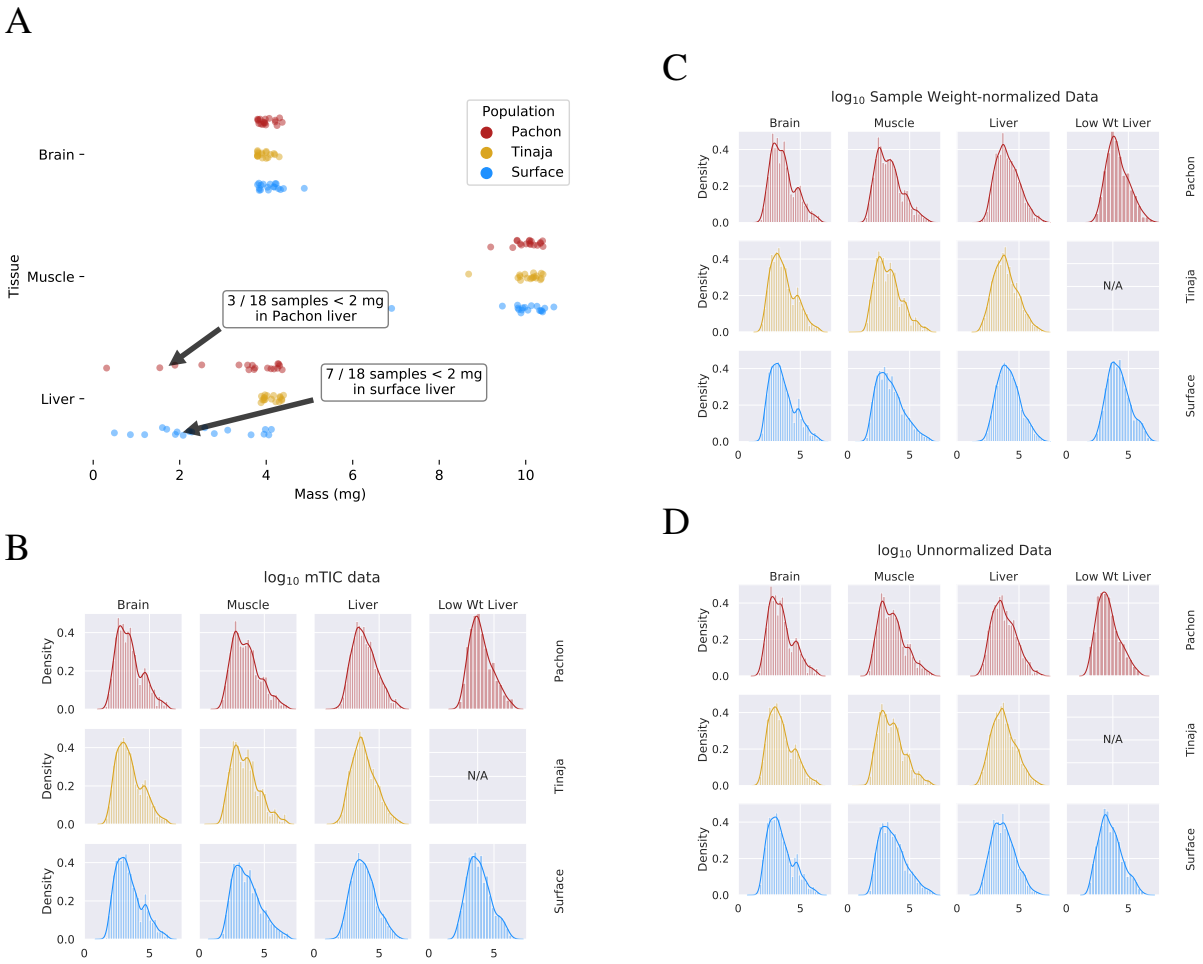


Fig. S2. Percent change of weight and K-factor during fasting regimen.



E

Scheme	Tissue	Population	Median Low Wt. Samples	Median All
mTIC	Liver	Pachon	3.86	3.79
mTIC	Liver	Surface	3.78	3.75
Unnormalized	Liver	Pachon	3.25	3.68
Unnormalized	Liver	Surface	3.45	3.69
Wt. Normalized	Liver	Pachon	4.03	3.97
Wt. Normalized	Liver	Surface	4.14	4.10

Fig. S3. Effect of normalization scheme on peak intensity distribution. In comparing the mass of the different tissue samples used in this study (A), we observed that liver samples exhibited more variability than muscle or brain. In particular, some liver samples of Pachón and surface fish had a mass of less than 2 mg. In order to adjust for this effect, we compared three different normalization schemes: mTIC normalization (B), wherein each peak is normalized to the sum total of all identified peaks in a given sample (see Methods), (C) sample weight-based normalization, where each peak is normalized according to the physical weight of its sample, and (D) unnormalized peaks. By examining the change in median peak intensity for identified compounds in each method (E), we found that mTIC normalization exhibits slightly better correction for low-weight (<2mg) median peak intensities compared to weight-based normalization and both are superior to unnormalized data. We therefore employed mTIC normalization for the remainder of the analysis.

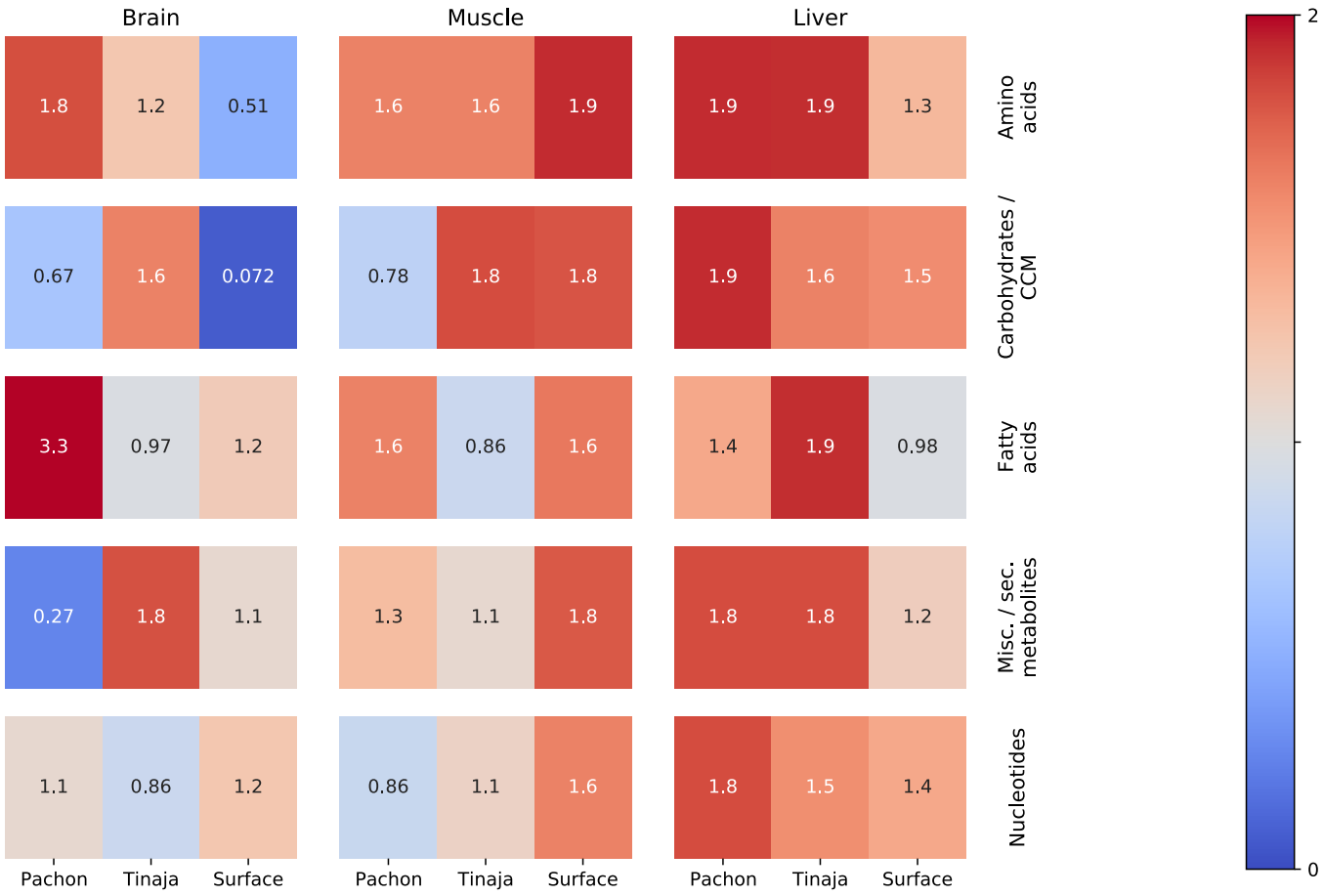


Fig. S4. Classifier performance shows which categories of primary metabolites are most salient in the starvation response. This figure shows $-log_{10}p$ values of an O–PLS / PLS classifier trained to discriminate 30-day fasted vs refed samples for different tissues, populations, and categories. O–PLS is used to produce de–noised data, which was then fitted to a single–component PLS model. The figure shows $-log_{10}p$ values for each classifier trained on different subsets of the data corresponding to different categories of primary metabolites. Red indicates the most robust classifiers and blue indicates the least robust. P–values were obtained by randomly permuting the feeding state indices for 2000 iterations and computing the two–tailed survival function of a normal distribution fitted to the DQ^2 values of the resulting permuted samples. P–values were adjusted for FDR using the BH method.

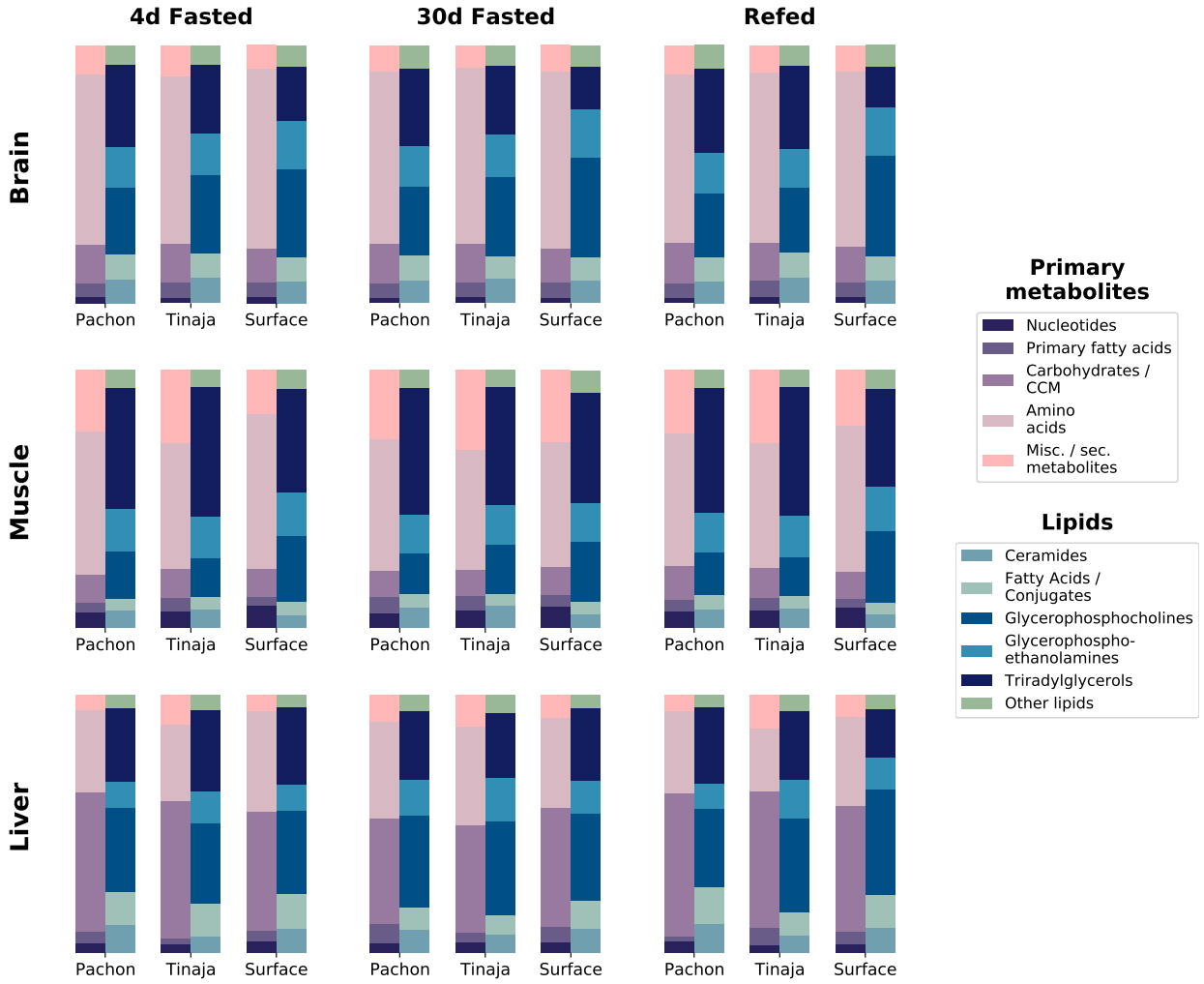


Fig. S5. Relative composition for primary metabolites and lipid classes shows metabolome profile for different tissues, populations, and feeding states. To determine overall composition of the metabolome for different experimental groups, category information for primary metabolites was computed based on five main categories (Supplementary Methods). Lipid data consists of the five most abundant lipid categories: ceramides, fatty acids and conjugates, glycerophosphocholines, and glycerophosphoethanolamines. Some metabolites, such as palmitate, were detected in both primary and lipid data. The peak intensities for all detected metabolites in these respective categories were summed across all replicates and across positive and negative modes to obtain total peak intensities for each population / condition combination. Total peak intensities were then plotted as a fraction of total peak intensity for all identified metabolites within a given experimental group.

Pop.	Treat-ment	Sample	Initial mass (g)	Final mass (g)	Initial SL (cm)	Final SL (cm)	Initial K-factor	Final K-factor	Weight Chg. (%)	K-factor Chg. (%)
Surface	30d	S30.1	0.47	0.35	2.8	2.9	2.14	1.44	-34.3	-49.2
	30d	S30.2	1.00	0.84	3.4	3.6	2.54	1.80	-19.0	-41.3
	30d	S30.3	0.89	0.48	3.4	3.4	2.26	1.22	-85.4	-85.4
	30d	S30.4	0.86	0.64	3.2	3.3	2.62	1.78	-34.4	-47.4
	30d	S30.5	0.83	0.66	3.3	3.3	2.31	1.84	-25.8	-25.8
	30d	S30.6	1.72	1.29	4.2	4.2	2.32	1.74	-33.3	-33.3
	4d	S4.1	0.43	0.43	2.7	2.8	2.18	1.96	0.0	-11.5
	4d	S4.2	1.01	0.88	3.4	3.9	2.57	1.48	-14.8	-73.2
	4d	S4.3	0.60	0.53	3.1	3.2	2.01	1.62	-13.2	-24.5
	4d	S4.4	0.86	0.8	3.3	3.7	2.39	1.58	-7.5	-51.5
	4d	S4.5	1.71	1.9	4.1	4.5	2.48	2.18	14.1	-13.6
	4d	S4.6	0.51	0.48	2.9	3.0	2.09	1.78	-6.3	-17.6
	Ref	SR.1	0.46	0.4	2.7	2.9	2.34	1.64	-15.0	-42.5
	Ref	SR.2	0.56	0.48	3.0	3.1	2.07	1.61	-16.7	-28.7
	Ref	SR.3	0.40	0.43	2.7	2.8	2.03	1.96	7.0	-3.7
	Ref	SR.4	0.75	0.62	3.1	3.2	2.52	1.89	-21.0	-33.1
	Ref	SR.5	0.77	0.75	3.3	3.5	2.14	1.75	-2.7	-22.5
	Ref	SR.6	0.86	0.97	3.3	3.6	2.39	2.08	11.3	-15.1
Pachón	30d	P30.1	1.03	0.96	3.7	3.7	2.03	1.90	-7.3	-7.3
	30d	P30.2	1.12	0.83	3.6	3.5	2.40	1.94	-34.9	-24.0
	30d	P30.3	1.17	0.99	3.6	3.5	2.51	2.31	-18.2	-8.6
	30d	P30.4	1.29	1.06	3.7	3.8	2.55	1.93	-21.7	-31.8
	30d	P30.5	1.02	0.81	3.5	3.5	2.38	1.89	-25.9	-25.9
	30d	P30.6	0.73	0.55	3.3	3.4	2.03	1.40	-32.7	-45.2
	4d	P4.1	0.82	0.81	3.3	3.5	2.28	1.89	-1.2	-20.8
	4d	P4.2	1.00	1.01	3.5	3.5	2.33	2.36	1.0	1.0
	4d	P4.3	1.34	1.53	3.8	4.1	2.44	2.22	12.4	-10.0
	4d	P4.4	0.32	0.37	2.6	2.8	1.82	1.69	13.5	-8.0
	4d	P4.5	1.05	1.16	3.7	3.7	2.07	2.29	9.5	9.5
	4d	P4.6	1.09	1.12	3.6	3.7	2.24	2.21	2.7	-1.4
	Ref	PR.1	1.06	0.93	3.7	3.9	2.09	1.57	-14.0	-33.5
	Ref	PR.2	0.97	1.00	3.6	3.7	2.08	1.97	3.0	-5.3
	Ref	PR.3	0.89	0.94	3.4	3.5	2.26	2.19	5.3	-3.3
	Ref	PR.4	0.75	0.74	3.1	3.4	2.52	1.88	-1.4	-33.7
	Ref	PR.5	1.18	1.23	3.7	3.7	2.33	2.43	4.1	4.1
	Ref	PR.6	0.66	0.67	3.0	3.3	2.44	1.86	1.5	-31.1
Tinaja	30d	T30.1	0.99	0.70	3.5	3.7	2.31	1.38	-41.4	-67.1
	30d	T30.2	1.37	1.12	3.9	4.0	2.31	1.75	-22.3	-32.0
	30d	T30.3	1.64	1.37	4.0	4.3	2.56	1.72	-19.7	-48.7
	30d	T30.4	1.08	0.93	3.7	3.7	2.13	1.84	-16.1	-16.1
	30d	T30.5	0.94	1.02	3.5	3.5	2.19	2.38	7.8	7.8
	30d	T30.6	0.72	0.75	3.5	3.5	1.68	1.75	4.0	4.0
	4d	T4.1	1.60	1.82	4.1	4.3	2.32	2.29	12.1	-1.4
	4d	T4.2	0.86	0.84	3.2	3.4	2.62	2.14	-2.4	-22.8
	4d	T4.3	0.96	0.78	3.4	3.6	2.44	1.67	-23.1	-46.1
	4d	T4.4	1.22	1.10	3.7	3.8	2.41	2.00	-10.9	-20.1
	4d	T4.5	1.56	1.68	4.1	4.3	2.26	2.11	7.1	-7.1
	4d	T4.6	1.15	1.01	3.6	3.7	2.46	1.99	-13.9	-23.6
	Ref	TR.1	1.23	1.36	3.7	3.9	2.43	2.29	9.6	-5.9
	Ref	TR.2	0.74	1.06	3.3	3.4	2.06	2.70	30.2	23.6
	Ref	TR.3	1.31	1.58	3.8	4.1	2.39	2.29	17.1	-4.1
	Ref	TR.4	0.66	0.83	3.2	3.3	2.01	2.31	20.5	12.8
	Ref	TR.5	0.98	1.08	3.6	3.6	2.10	2.31	9.3	9.3
	Ref	TR.6	1.48	1.48	4.0	4.2	2.31	2.00	0.0	-15.8

Table S1. Weight and K-factor measurements of all samples.

Table S2. An Excel file containing the results of the O–PLS / GLM feature identification pipeline for primary metabolites divided into categories. Columns represent different population and feeding condition combinations.

	Pachon vs. Surface									Tinaja vs. Surface									Pachon vs. Tinaja								
	Brain			Muscle			Liver			Brain			Muscle			Liver			Brain			Muscle			Liver		
	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R
Fatty Acyls								−									−	−									+
Glycerolipids		+	+		+							+			+												
Glycerophospholipids	−	−	−		−							−			−												
Sphingolipids				+	+	+							+	+	+	−	−	−							+	+	+
Sterol Lipids		−	−									−						−									

Table S3. Interpopulation Differences in Abundance of Lipid Categories Peak intensities for all lipids in a given category (determined from the LipidMaps "`CATEGORY`" attribute) were summed to yield a total intensity for each category which is either significantly (at the $p < 0.05$ level) up– (+) or down–regulated (−) in a given cave population with respect to surface (Pachón versus surface and Tinaja versus surface, top row) or the Pachón cave population with respect to the Tinaja cave population (last comparison, top row). The sample set for each tissue / feeding state combination consists of six individuals from each population as shown in Fig 1. P–values were obtained from the OPLS / GLM approach described in Methods. Coloring (■) indicates a class that agrees in significance and directionality between both cave populations and is thus may be related to cave adaptation. Columns corresponding to muscle are highlighted to help distinguish the different tissues. Compare Table S4 based on the "`MAIN_CLASS`" attribute.

	Pachon vs. Surface									Tinaja vs. Surface									Pachon vs. Tinaja								
	Brain			Muscle			Liver			Brain			Muscle			Liver			Brain			Muscle			Liver		
	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R
Ceramides				+	+	+			+			+	+	+	+	−	−	−			−				+	+	+
Diradyl-glycerols							+																				
Docosanoids								−	+								−			+						+	+
Fatty Acids								−									−	−									+
Fatty esters				−	−	−	−		−	−	−	−	−		−	−					+						−
Glycerophospho-cholines		−	−									−	−		−												
Glycerophospho-ethanolamines	−	−	−							−	−	−					+									−	−
Glycerophospho-glycerols				−		−							−	−	−		+	+					+	+			
Glycerophospho-glycerophospho-glycerols					−									−													
Glycerophospho-inositols				−	−	−		+					−	−	−	+	+	+							−	−	−
Glycerophospho-serines									−						−												−
Glycosyldiradyl-glycerols		+																		+							
Monoradylglycerols				+	+	+			−				+	+	+												−
Neutral glycosphingolipids					+	+			−					+													−
Oxidized glycerophospho-lipids		−	−	∅	∅	∅						−	∅	∅	∅							∅	∅	∅			
Phosphosphingo-lipids									−																		−
Sphingoid bases		−		+	+	+		+					+	+	+												−
Sterols		−	−									−						−									
Triradylglycerols		+	+		+							+															

Table S4. Interpopulation Differences in Abundance of Lipid Classes Significant increased (+) or decreased (−) lipid classes based on summed peak intensities for every lipid species belonging to a given LipidMaps "`MAIN_CLASS`" label. ∅ denotes classes which were not detected in a given sample set. LipidMaps also possesses a "`CATEGORY`" attribute that provides a more coarse–grained classification of lipid species, which is used as a basis for a similar analysis shown in Table S3.

	Pachon vs. Surface									Tinaja vs. Surface									Pachon vs. Tinaja								
	Brain			Muscle			Liver			Brain			Muscle			Liver			Brain			Muscle			Liver		
	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R
Monounsat. FAs	+					+						+															
Polyunsat. FAs								−									−	−									+
Saturated FAs				+		+							+		+											+	

Table S5. Intra–population Differences in Fatty Acid Saturation For lipids corresponding to free fatty acids, sauturation was calculated based on the presence of double bonds in LipidMaps structural data and used to classify each LMID as either saturated, monounsaturated, or polyunsaturated. Significance values were again calculated using an OPLS / Bayesian GLM workflow. Coloring and markings as before.

	Pachon vs. Surface									Tinaja vs. Surface									Pachon vs. Tinaja								
	Brain			Muscle			Liver			Brain			Muscle			Liver			Brain			Muscle			Liver		
	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R	4	30	R
13,16-docosadienoic acid							−				−																
20:5 Cholesterol ester	−		−	−	−	−				−		−	−	−	−			−									
22:1 Cholesterol ester																		−									
22:6 Cholesterol ester	−		−	−		−				−	−	−	−		−			−	+								
9,12-Hexadecadienylcarnitine	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−					+	+						
Acylcarnitine C18:0	−	−		−	−				−	−	−	−	−	−	−					+	+						−
Arachidonic acid									+							−		−						+			+
Behenic acid																									+		
Bishomo-alpha-linolenic acid	−	−	−							−	−	−				−		−									+
C17 Sphingosine	−			+	+	+	+						+	+	+			+					+				−
Capric acid					−	−								−			−						−				
Cholesterol	−		−		−	−			−			−		−	−			−					+	+	+		
DHA							−									−		−								+
DPA							−		+							−		−	+						+		+
Dihomolinoleic acid					−		−			−	−					−		−									+
EPA					−		−									−		−									+
Lauric acid					−	−								−				−					−				
Mangold's acid					−		−											−									+
Myristic acid														−				−									
Nervonic acid																−		−							+		
Oleic acid	+	+				+			+			+			+							+					
Palmitic acid				+	+	+				−			+	+	+		+	+									−
Palmitoylcarnitine	−	−		−	−	−	−	−	−	−	−	−	−	−	−	−	−	−		+	+						
Pentadecylic acid	−								−							−		−				−			+		
Physeteric acid				−			−			−						−		−									+
Stearidonic acid					−		−			−						−		−									+
alpha-Linolenic acid					−		−	−	+	−				−	−	−	−	−	+								+
bishomo-gamma-linolenic acid					−		−		+		−				−	−	−	−	+								+
cis-9-palmitoleic acid									+																		+
cis-erucic acid					−		−	−	−							−		−					−				+
cis-gondoic acid					−		−			−						−		−	+								+
tetracosapentaenoic acid	+							+	+										+			+					+

Table S6. Common lipids Lipids with common names were selected from the set of 447 identified lipids and analyzed using the O–PLS / GLM scheme. Many lipids exhibit a strongly conserved pattern between Pachón and Tinaja, and these are highlighted in blue. Of note, omega–3 fatty acids appear to exhibit lower abundance in the 4-day fasted state in liver. Some metabolites, such as palmitate and oleic acid, overlap with primary metabolomics data.

Category	Brain	$\log_2$ fc	Muscle	$\log_2$ fc	Liver	$\log_2$ fc
Amino acids			N-acetylputrescine	-0.53	N-carbamoylaspartate	2.93
					threonine	0.55
					proline	1.66
					isoleucine	0.52
					alanine	-1.01
					2-ketoadipic acid	-1.50
Carbo-hydrates / CCM	galactinol	-2.67	myo-inositol	1.09	malic acid	1.54
	ribose	-0.51	ribose-5-phosphate	-1.21	fumaric acid	1.49
			digalacturonic acid	-1.06	myo-inositol	1.63
			ribulose-5-phosphate	-1.42	lyxitol	-1.95
			glyceraldehyde-3-phosphate	-0.60	1,5-anhydroglucitol	-3.27
					glyceric acid	-3.07
					xylitol	-1.85
					lactic acid	-4.11
Fatty acids			behenic acid	-0.25	oleamide	0.52
			lignoceric acid	-0.02	glutaric acid	0.34
Misc.	phosphoethanolamine	0.72	2-hydroxyglutaric acid	-1.87	isothreonic acid	0.54
			isothreonic acid	-0.74	ciliatine	1.68
			phosphate	-1.43	xanthurenic acid	2.72
Nucleotides			guanosine	-1.02	orotic acid	3.13
			UDP-N-acetylglucosamine	-0.54	adenine	0.91
					uracil	0.63
					hypoxanthine	-0.99
					nicotinamide	-0.86

Table S7. Adaptive Response in 30-day Fasting. Differentially significant metabolites in 30-day fasted states which are similar in both cave populations. This table shows p–values associated with a logistic regression model using OPLS–filtered z–scores for 30-day tasted and refed fish as input. To identify metabolites conserved between both cave populations, we implemented the test $0.5 \cdot (\mathrm{PS} - \mathrm{PR} + \mathrm{TS} - \mathrm{TR}) - (\mathrm{SS} - \mathrm{SR})$ where PS refers to Pachón 30-day fasted, PR refers to Pachón refed, etc. Tables S8 and S9 show conserved metabolites in cave populations versus surface for 4-day fasted versus refed and 30-day fasted versus 4-day fasted respectively. $\log_2$ fc values are also based on this formula. Outliers were excluded for this analysis (Fig **??**).

Category	Brain	$\log_2$ fc	Muscle	$\log_2$ fc	Liver	$\log_2$ fc
Amino acids	N-acetylaspartic acid	-0.23	shikimic acid	-0.75	tryptophan	0.69
			glutamic acid	-0.73	proline	1.01
					N-carbamoylaspartate	3.03
Carbo-hydrates / CCM	sucrose	1.49	ribose-5-phosphate	-2.01		
	glucoheptulose	-1.50				
	threonic acid	-0.50				
	maltose	-1.15				
	maltotriose	-1.40				
Fatty acids	2-hydroxybutanoic acid	-2.35			oleic acid	-1.63
Misc.	2-aminobutyric acid	-0.54	phosphoethanolamine	-1.08	isothreonic acid	0.45
			4-hydroxybenzoate	-0.55		
			2-hydroxyglutaric acid	-2.67		
			maleimide	-0.56		
			ciliatine	-0.81		
Nucleotides			UDP-N-acetylglucosamine	-0.82	orotic acid	3.08
			cytosine	-0.69	inosine	-0.34
			uridine	-1.12		
			guanosine	-1.39		
			uracil	-1.15		

Table S8. Adaptive Response in 4-day Fasting. Differentially significant metabolites in 4-day fasted states which are similar in both cave populations. Cf. Table S7 with the difference that this table compares 4-day fasted versus refed conditions. Thus, metabolites displayed as upregulated in this table are differentially upregulated in 4-day fasted cave fish versus refed cave fish using surface fish as a baseline for comparison. Outliers are not included in this analysis.

Category	Brain	$\log_2$ fc	Muscle	$\log_2$ fc	Liver	$\log_2$ fc
Amino acids					oxoproline	0.79
					glutamic acid	1.45
					threonine	0.67
					alanine	-1.13
Carbo-hydrates / CCM	maltotriose	3.13	myo-inositol	1.26	myo-inositol	1.08
	threonic acid	0.50	digalacturonic acid	-1.49	glucose	0.40
	sophorose	-1.36			raffinose	-1.89
	isomaltose	-1.83			maltotriose	-1.07
					xylitol	-1.50
					1-kestose	-1.82
Fatty acids	stearic acid	-0.25	palmitic acid	-0.24	stearic acid	-0.05
					palmitic acid	0.06
Misc.			phosphate	-2.71	2-aminobutyric acid	1.54
			glycolic acid	-0.99	ciliatine	1.37
			pyrophosphate	-1.28	xanthurenic acid	1.65
Nucleotides	adenosine	1.87	xanthine	0.76		
			inosine	0.92		

Table S9. Adaptive Response in 30- vs 4-day Fasting. Differentially significant metabolites in 30-day vs 4-day fasting which are similar between both cave populations. Cf. Table S7 with the differences that this table compares 30-day fasted versus 4-day fasted conditions. Thus, metabolites displayed as upregulated in this table are differentially upregulated in 30-day fasted cave fish versus 4-day fasted cave fish using surface fish as a baseline for comparison. Outliers are not included in this analysis.