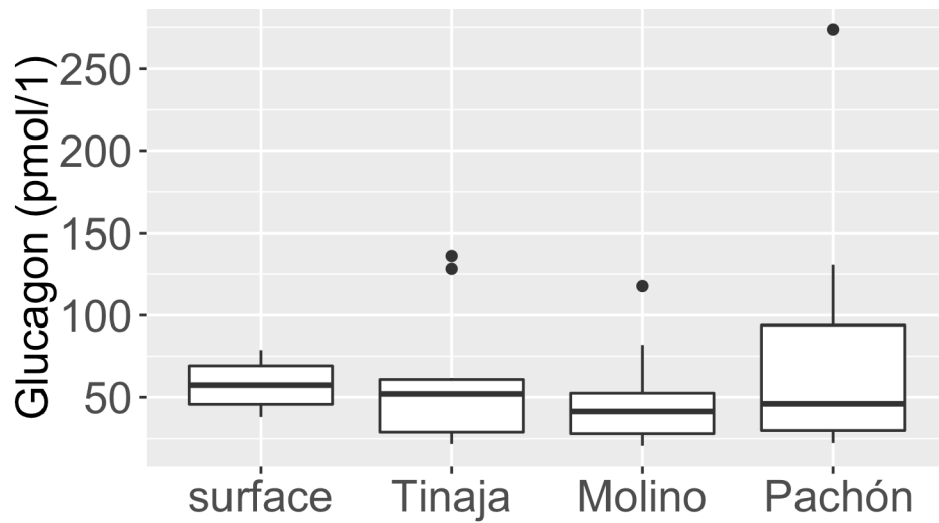


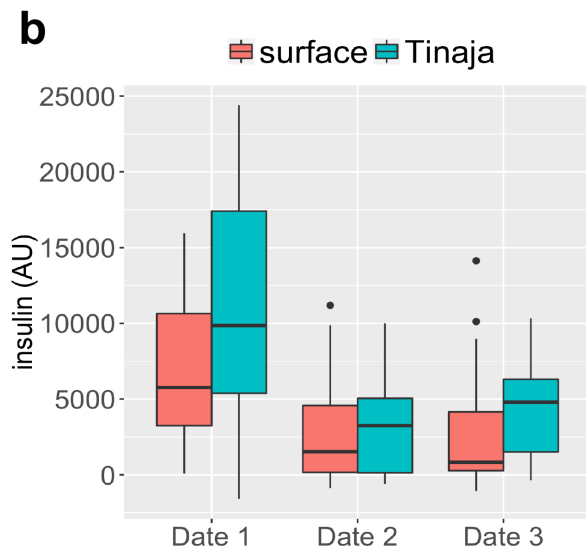
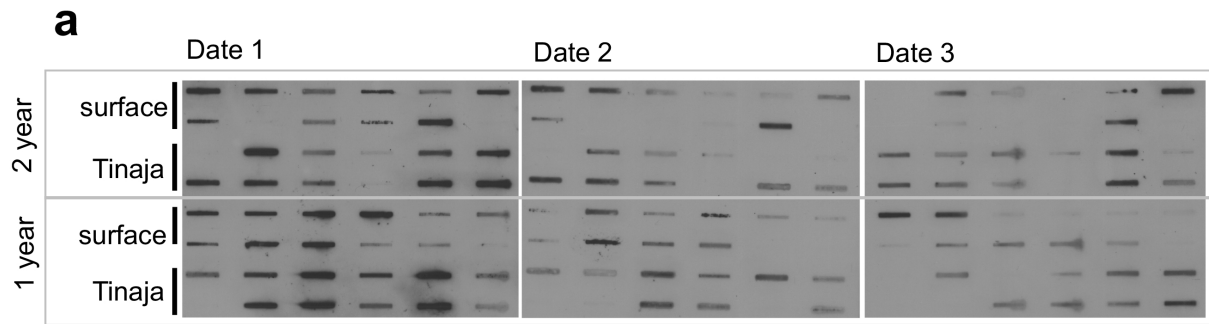
**Extended Data Figure 1. Tinaja and surface larvae have same number of glucagon and insulin positive cells at 10-11 days post fertilization (dpf).**

**a**, Wholemount immunohistochemical detection of insulin and glucagon positive cells in 10dpf Tinaja larvae. **b**, Number of glucagon and insulin positive cells in 10-11dpf surface and Tinaja larvae. **c**, Average number of glucagon and insulin positive cells, fish length, ratio of insulin to glucagon positive cells, and p-value comparing the Surface and Tinaja values (determined using students t-test).



**Extended Data Figure 2.**

Boxplot comparing serum glucagon levels between surface, Tinaja, Molino, and Pachón after 24-hour fast, (n=12 fish per population, average 57.87, 59.76, 79.66, 48.89 respectively, p=0.52, one-way ANOVA). Box plots show 25th, 50th, and 75th percentiles (horizontal bars), and 1.5 interquartile ranges (error bars), dots represent outliers.

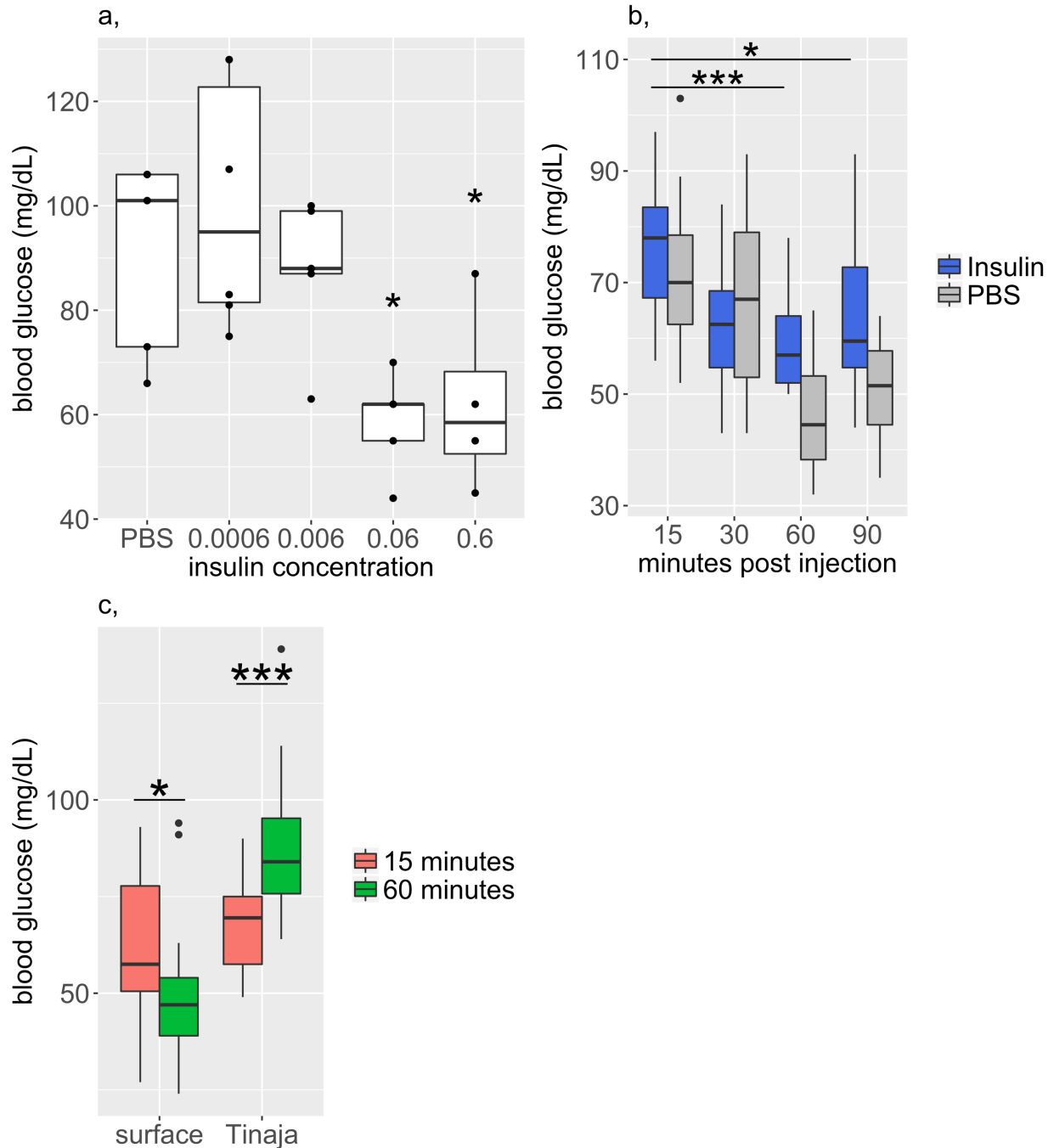


**c**

	surface	Tinaja	p-value
<i>Date 1</i>	7194	10771	0.057
<i>Date 2</i>	3043	3548	0.620
<i>Date 3</i>	2872	4370	0.168

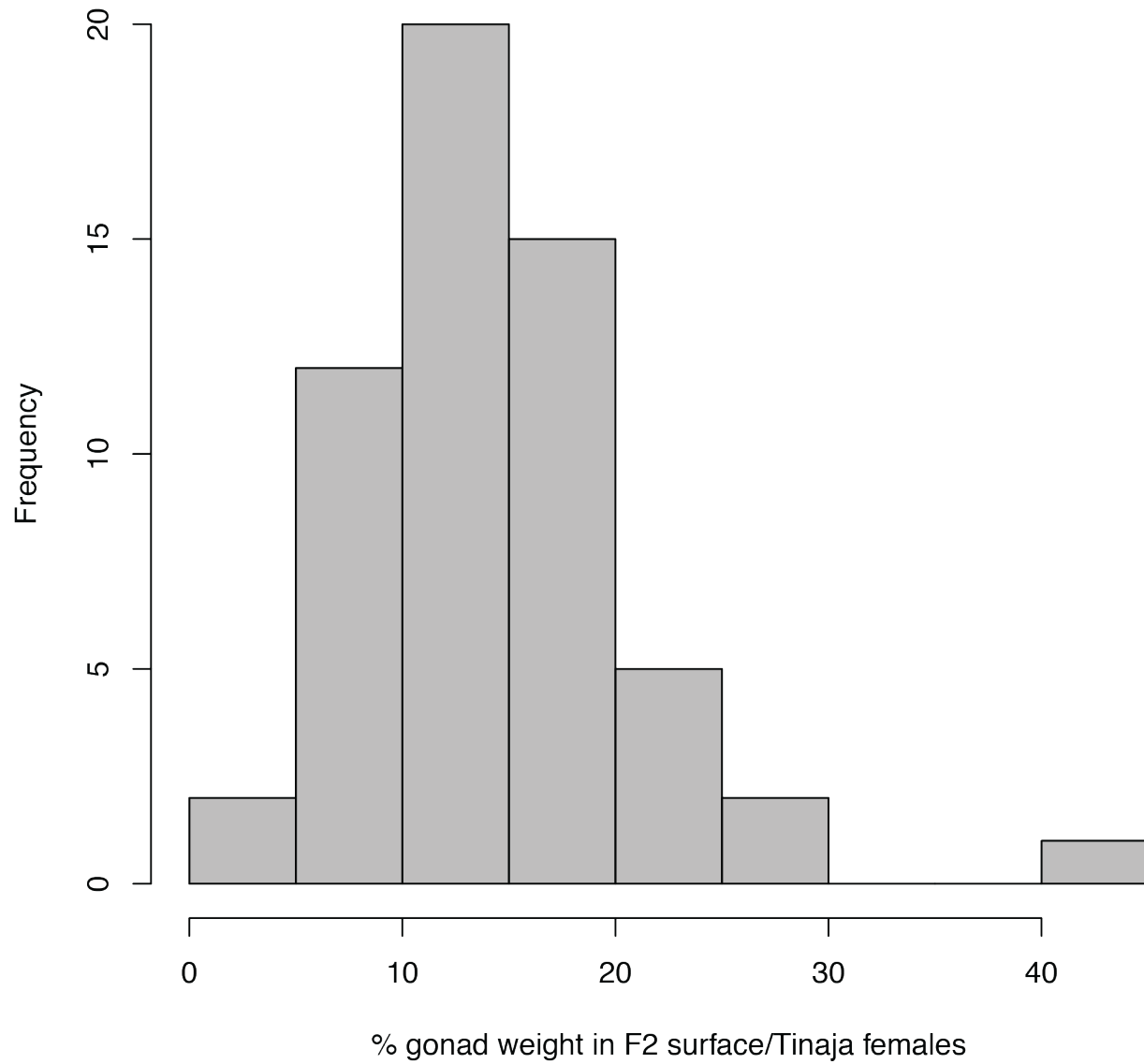
**Extended Data Figure 3.**

**a**, Dot-blot of serum collected from 2-year-old and 1-year-old surface and Tinaja fish at three different dates (9 days and 16 days apart,  $n=12$  for each population and age) probed with anti-insulin antibody. **b**, Quantification of insulin level measured by densitometry of blots (AU, artificial units, median, 25th, 50th, and 75th percentiles (horizontal bars), and error bars at 1.5 interquartile ranges). No significant difference between ages was detected for any date ( $p=0.37, 0.66, 0.89$  respectively) so the values were grouped. **c**, Table showing average insulin value and p-value for each date (students two-sample t-Test).



**Extended Data Figure 4. a,** We injected different concentrations of human recombinant insulin (sigma, product I9278, stock 9.5-11.5mg/mL) into the intraperitoneal cavity of surface fish to determine the effective dosage for subsequent experiments. Blood glucose levels are significantly less in fish injected with approximately .6 $\mu$ g or .06 $\mu$ g insulin per mg of fish weight compared to .006 $\mu$ g (n=4 fish per dosage, dots represent individual fish, significance codes from one-way ANOVA with HSD post hoc test, \*p<.05). We used .06 $\mu$ g per mg of fish in subsequent experiments. **b,** Blood glucose levels of surface fish over time after

injection of PBS or insulin. Blood glucose levels are significantly lower at 60 and 90 minutes compared to 15 minutes after insulin injection so we chose to focus on the 60 minute timepoint for comparison to Tinaja cavefish (n=10 fish per time point and condition, significance codes from one-way ANOVA with HSD post hoc test, \* $p < .05$ ). **c**, Blood glucose levels at 15 and 60 minutes after insulin injection in surface fish and Tinaja cavefish. Surface fish display a significant decrease in blood glucose levels while cavefish display a significant increase in blood glucose levels (significance codes from two-tailed student's t-test, \* $p < .05$ , \*\*\* $p < 0.0005$ ). Tinaja blood glucose may increase due to the stress of being injected; stress hormones, such as catecholamines, ACTH, and epinephrine, cause transient increases in blood glucose in humans and mice, an effect that cannot be mitigated in the absence of insulin signaling<sup>39,40</sup>. While both cavefish and surface fish likely undergo a stress response upon injection, this is overcome in the surface fish, which have wild type insulin activity, but not in the Tinaja cavefish, which have reduced insulin signaling.



**Extended Data Figure 5.**

Histogram displaying percent gonad weight (gonad weight/weight\*100) of 58 F2 surface/Tinaja females fed on a 6mg diet for 4 months (min = 3.57, 1<sup>st</sup> Qu: 10.64, median = 13.96, mean = 13.96, 3<sup>rd</sup> Qu = 17.57, max = 41.86).

## Extended Data 1

Insulin pathway sequences: [see webarchive file](#)

## Extended Data 2

Locations of Surface Fish sampled:

Rio Sabinos and Rio Sarco (n=12)  
San Rafael Los Castros (n=9)  
Rio Florido (n=1)  
Chamal (n=1)  
Nacimiento del Rio Santa Clara (n=21)  
Rio Subteraneo Valley (n=27)

## Extended Table S1

A. Measures of divergence for each cave population compared to its most closely related-surface population. We also included Río Choy-Pachón and Río Choy-Tinaja as there is some evidence that this surface population interbreeds with these two caves.  $D_{XY}^{41,42}$  is an absolute measure of divergence that is independent of diversity levels,  $F_{ST}^{43}$  is one of the most common ways to measure population differentiation, but is strongly influenced by within-population diversity. hapFLK<sup>44</sup> is a measure of haplotype divergence that takes into account hierarchical population structure. hapFLK is among the most sensitive measures for detecting positive selection<sup>45</sup>. B. Measures of diversity include: H-SCAN<sup>46</sup> which is measured here as the actual physical distance in number base pairs between the two delimiting SNPs that terminate a pairwise homozygosity tract at each end,  $\pi$  is pairwise nucleotide diversity within populations. Tajima's D is the measure of  $\pi$  (observed) -  $\pi$  (expected) based on the number of variable sites in the data and is a measure of deviation from neutrality. Negative values may be caused by selective sweeps or recent population expansion, whereas positive values may be caused by balancing selection or a recent bottleneck<sup>47</sup>.

Total sites refers to total sites with data for six or more individuals per population. Fixed = fixed differences between populations, invariant = invariant sites, percentile in the genome is relative to all other genes in the genome with data available. Data for hapFLK refers to the mean, median, maximum, minimum and standard deviation of p-values for the gene (total of N = 54 sites with p-values in the coding region of the gene).

### A. Measures of divergence

$D_{XY}$	Total sites	Mean	Stdev	Fixed	Invariant	% in genome
Rascon-Tinaja	4033	0.001073	0.023402	0	4016	26.51
Rascon- Pachón	4033	0.001021	0.025561	0	4019	22.71
Río Choy-Pachón	4046	0.002176	0.035289	1	4018	44.97
Río Choy-Tinaja	4046	0.002175	0.032547	0	4014	42.15
Río Choy-Molino	4046	0.001807	0.026271	0	4020	40.59

$F_{ST}$	Total sites	Mean	Stdev	% in genome
----------	-------------	------	-------	-------------

Rascon-Tinaja	12	0.207804	0.291358	38.84
Rascon- Pachón	12	0.269965	0.356356	45.21
Río Choy-Pachón	27	0.245772	0.310325	52.82
Río Choy-Tinaja	25	0.234978	0.271977	52.25
Río Choy-Molino	19	0.164644	0.127403	13.31

hapFLK	Median	Max	Min	Stdev	# of p.values<0.05
hapFLK p-values	0.328636	0.788896	0.086054	0.258214	0

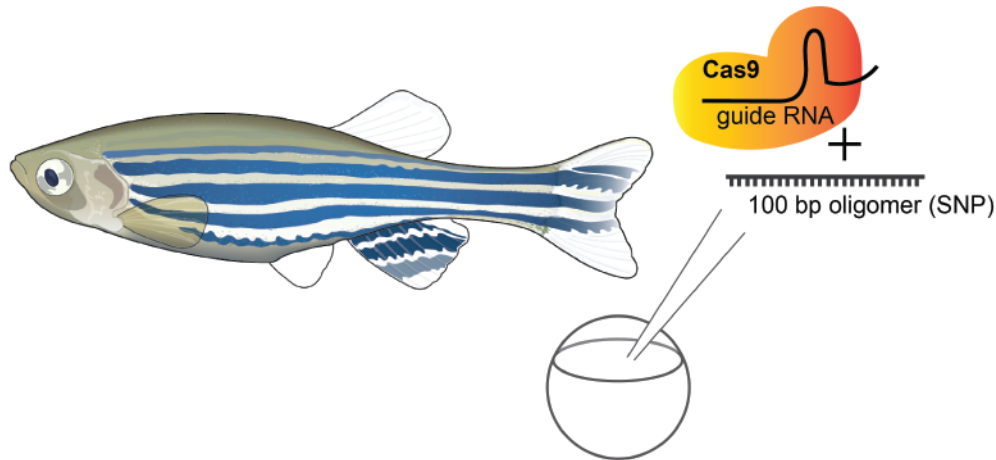
## B. Measures of Diversity

HSCAN	Mean_HSCAN	Percentile in genome (smaller = shorter homozygosity runs)
Tinaja	1617.131611	33.97
Pachón	10877.28727	48.51
Rascon	363.8201	34.83
Molino	141.119421	1.37
Río Choy	328.45057	47.32

$\pi$	Total Sites	Mean	Stdev	% in genome
Tinaja	4047	0.00078	0.01431	53.43
Pachón	4047	0.00031	0.00719	27.22
Rascon	4033	0.00055	0.01149	14.38
Río Choy	4046	0.00182	0.02584	52.33
Molino	4047	0.00127	0.02555	66.81

Tajima'sD	# of SNPS	Tajima'sD	% in genome
Tinaja	15	-0.924334	29.89
Pachón	8	-1.50598	15.32
Rascon	11	-1.40961	11.62
Río Choy	25	0.159596	76.84
Molino	10	2.74051	94.84





### Supplementary Figure 1:

Schematic of CRISPR/Cas9 experiment using homologous directed repair to introduce the P211L mutation in zebrafish *insra*.

### References

- 39 Ziegler, M. G., Elayan, H., Milic, M., Sun, P. & Gharaibeh, M. Epinephrine and the metabolic syndrome. *Curr. Hypertens. Rep.* **14**, 1-7 (2012).
- 40 Surwit, R. S., Schneider, M. S. & Feinglos, M. N. Stress and diabetes mellitus. *Diabetes Care* **15**, 1413-1422 (1992).
- 41 Nei, M. *Molecular Evolutionary Genetics*. (Columbia University Press, 1987).
- 42 Nei, M. *Mathematical model for studying genetic variation in terms of restriction endonucleases*. . Vol. 76 (Proceedings of the National Academy of Sciences, 1979).
- 43 Charlesworth, B. *Measures of divergence between populations and the effect of forces that reduce variability*. Vol. 15 538-543 (1988).
- 44 Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., and Servin, B. in *Genetics* Vol. 193 Ch. 929-941, (2013).
- 45 Schlamp, F., et al. in *Mol. Ecol.* Vol. 25 342-356 (2016).
- 46 Messer, P. W. *H-scan: Detecting hard and soft sweeps in population genomic data*, <<https://messerlab.org/resources/>> (2015).
- 47 Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595 (1989).