

Supplementary information

Supplementary Data 1– All CRE coordinates (Provided as a separate excel file)

This Excel spreadsheet provides a list of cis-regulatory elements mapped in this study. The column headers are as below:

Header	Description
name	peak name
seqnames	chromosome
start	start
end	end
width	peak width
peaks	peak name
TS	log2(Tinaja/Surface)
TS.q	adj P-Value
PS	log2(Pachon/Surface)
PS.q	adj P-Value
PT	log2(Pachon/Tinaja)
PT.q	adj P-Value
A	geometric mean of signal
nearest.gene	Nearest Gene
	distance to TSS of nearest
dist.tss	gene
s	1 = peak in Surface, 0 = not
p	1 = peak in Pachon, 0 = not
t	1 = peak in Tinaja, 0 = not
h3k4me3.s	H3K4me3 overlap in Surface
h3k4me3.p	H3K4me3 overlap in Pachon
h3k4me3.t	H3K4me3 overlap in Tinaja
h3k27ac.s	H3K27Ac overlap in Surface
h3k27ac.p	H3K27Ac overlap in Pachon
h3k27ac.t	H3K27Ac overlap in Tinaja
h3k27me3.s	H3K27me3 overlap in Surface
h3k27me3.p	H3K27me3 overlap in Pachon
h3k27me3.t	H3K27me3 overlap in Tinaja

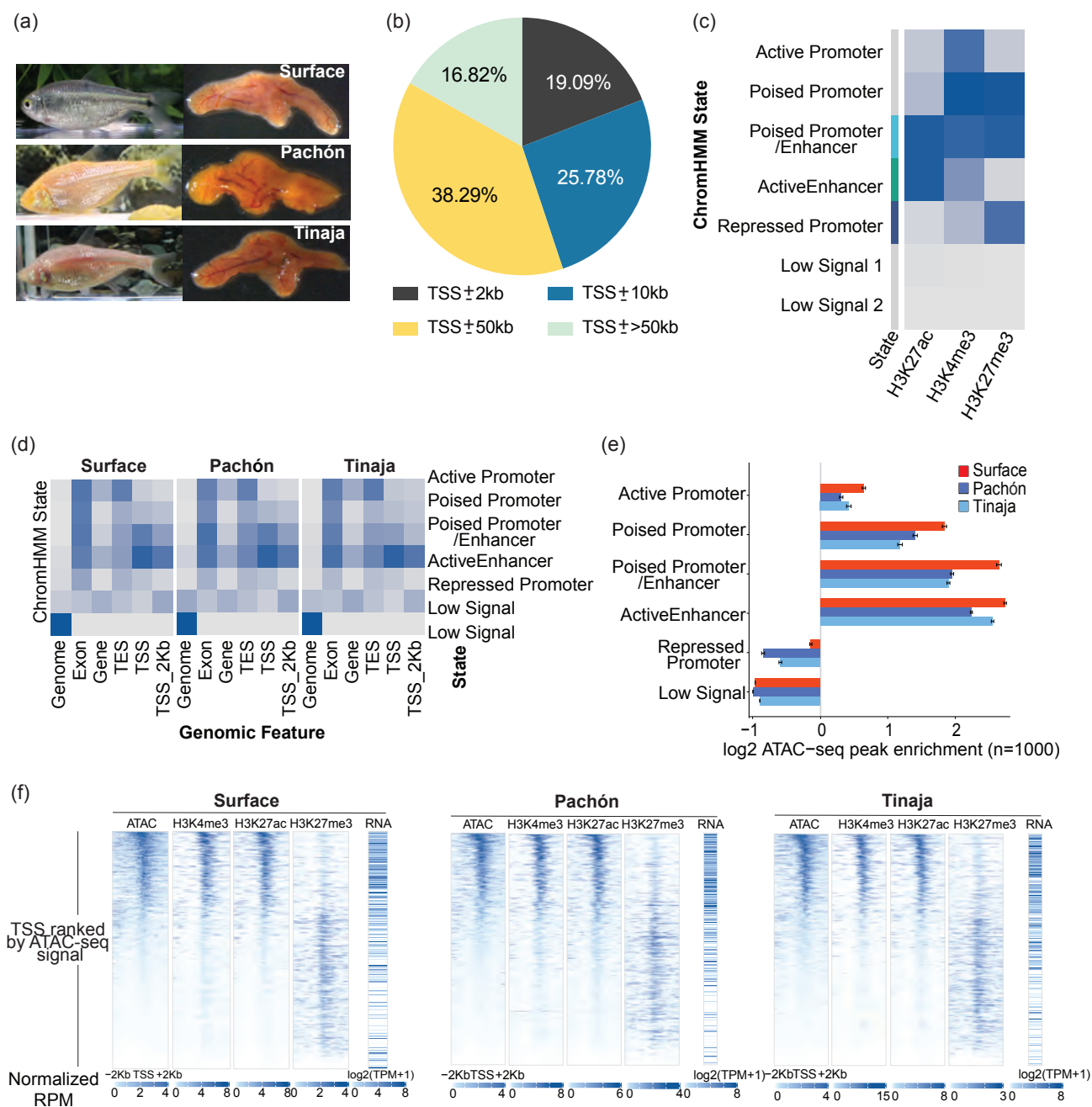
Supplementary Data 2– RNA-seq_AstMex2 (Provided as a separate excel file)

This excel spreadsheet provides the transcriptome data. The column headers are as below:

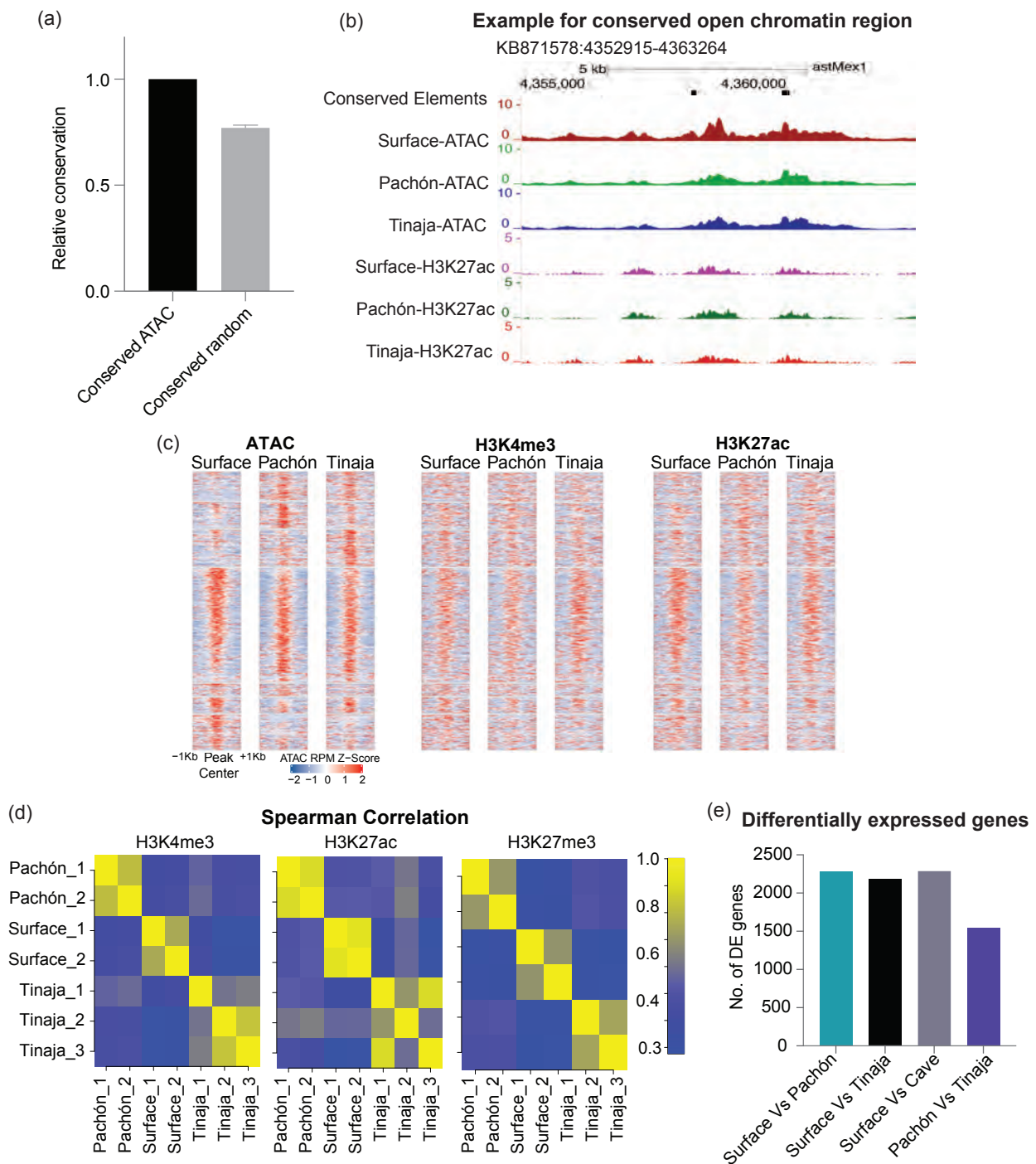
Header	Description
ens_id	ensembl id
gene	gene name
chr	Chromosome no.
start	
end	
Strand	
Biotype	
s_pachon_atac_1	CPM Pachón 1
s_pachon_atac_2	CPM Pachón 2
s_surface_atac_1	CPM Surface 1
s_surface_atac_2	CPM Surface 2
s_tinaja_atac_1	CPM Tinaja 1
s_tinaja_atac_2	CPM Tinaja 1
pachon_vs_surface.log2fc	Log fold change
pachon_vs_surface.log2cpm	Log CPM ratio
pachon_vs_surface.pval	P value
pachon_vs_surface.padj	Adjusted P value
pachon_vs_tinaja.log2fc	Log fold change
pachon_vs_tinaja.log2cpm	Log CPM ratio
pachon_vs_tinaja.pval	P value
pachon_vs_tinaja.padj	Adjusted P value
tinaja_vs_surface.log2fc	Log fold change
tinaja_vs_surface.log2cpm	Log CPM ratio
tinaja_vs_surface.pval	P value
tinaja_vs_surface.padj	Adjusted P value

Supplementary Data 3 – Aligned S-P-CRE-*hpdb*

DNA sequence alignment for surface and Pachón CRE-Hpdb highlighting the sites that were mutated for reporter assay.

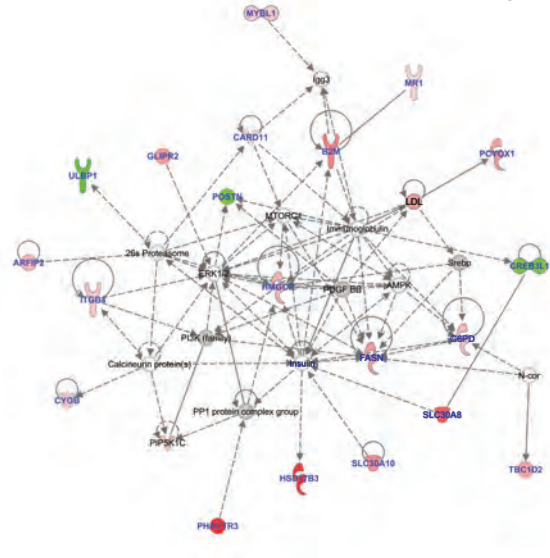


Supplementary Figure 1: (a) Left panel shows representative images of surface, Pachón and Tinaja fishes and right panel shows their corresponding livers. (b) Pie chart showing genomic distribution for the CREs mapped. We defined promoters as CREs within $\pm 2\text{kb}$ of the transcription start site (TSS); Proximal CREs within $\pm 10\text{kb}$ of TSS; Distal elements within $\pm 50\text{kb}$ of TSS and the rest are categorized as intergenic. (c) Heatmap showing the chromatin features comprising each of the seven chromatin states into which the CREs were categorized using ChromHMM. The two low signal states belong to different genomic regions as depicted in Fig. S1d. (d) ChromHMM emission heatmap showing the distribution of various chromatin states across various features in the genome of surface, Pachón and Tinaja. (e) Enrichment of ATAC-seq peaks in the ChromHMM states. The two low signal states were combined into a single low signal state for better depiction. The x-axis shows the average enrichment ratio \pm standard deviation ($n=1000$) for each ChromHMM state. (f) Heatmaps for ATAC-seq, H3K4me3, H3K27ac and H3K27me3 signal at all the promoters (TSS $\pm 2\text{Kb}$) along with the corresponding RNA levels for the surface fish. Plotted are reads per million (RPM) normalized to ChIP-seq signals at promoters.

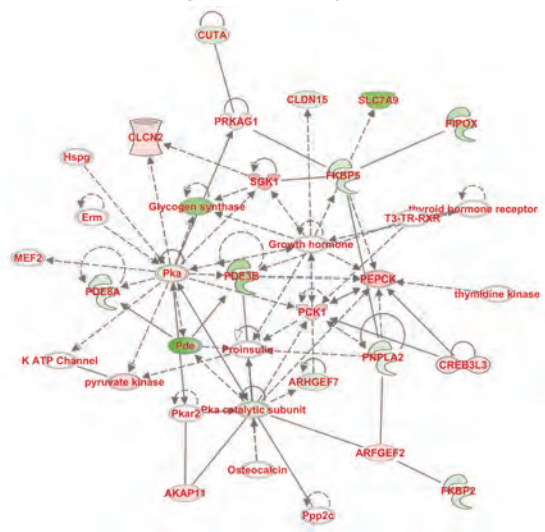


Supplementary Figure 2: (a) Overlap of ATAC peaks with evolutionarily conserved regions using 11-fish GERP scores normalized to 1 (Black bar). The grey bar represents average relative conservation of same number of randomly picked, non-exonic and size matched genomic regions. The standard deviation from 1000 iterations of picking random set of regions (b) Browser shot of one of the conserved candidate-CREs. (c) Heatmaps showing high correlation between ATAC-seq, H3K4me3 and H3K27ac signals in surface, Pachón and Tinaja. (d) Heatmaps for spearman correlations between different samples for each chromatin feature measured in the study – H3K4me3, H3K27ac, H3K27me3. (e) Number of differentially expressed genes between different morphotypes.

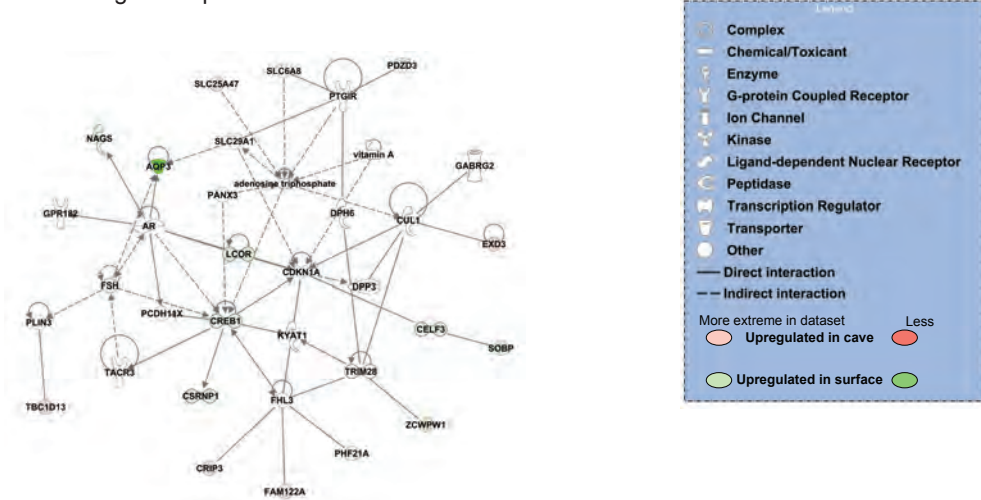
(a) Lipid metabolism network in cave-biased genes



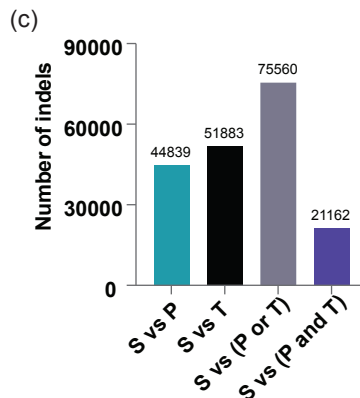
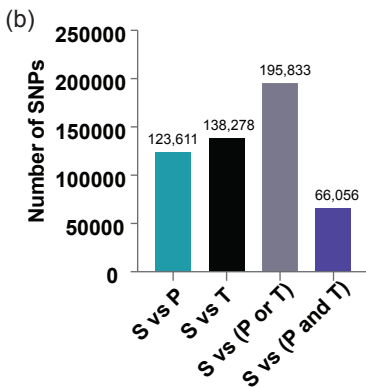
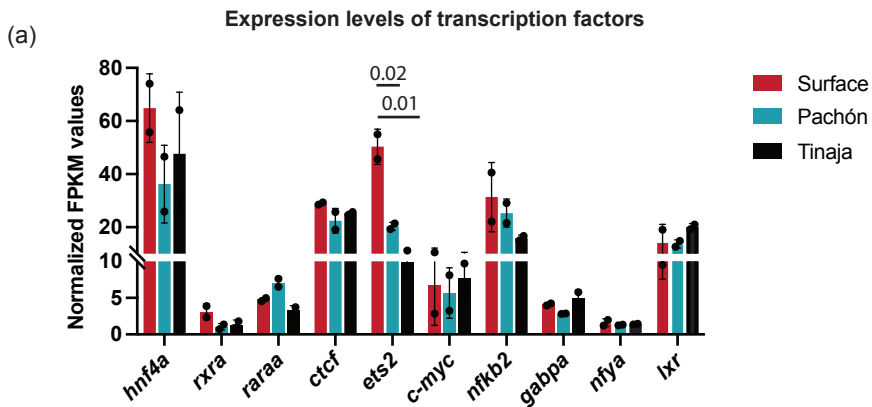
(b) HNF4a targets in carbohydrate metabolism network



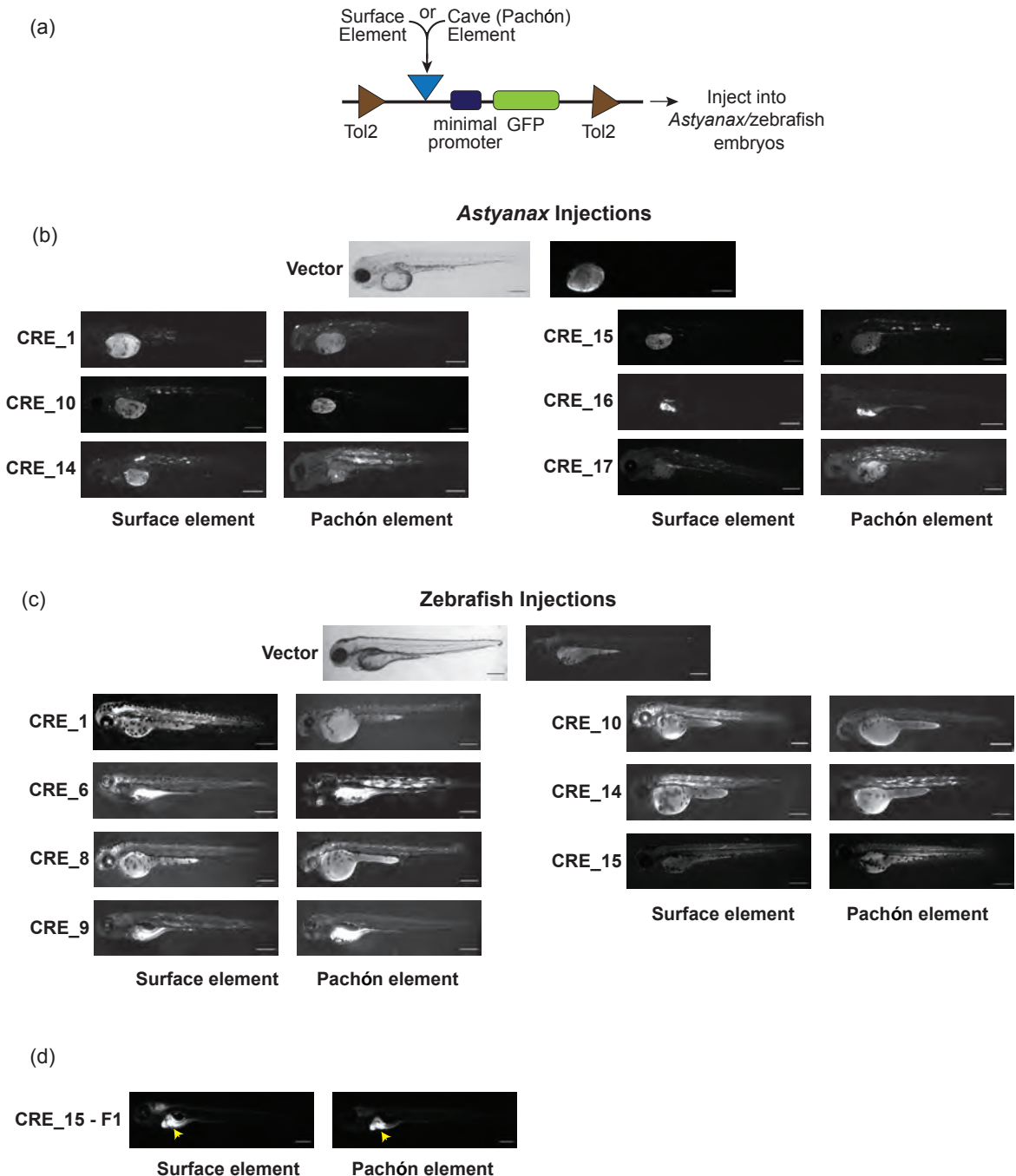
(c) HNF4a targets in lipid metabolism network



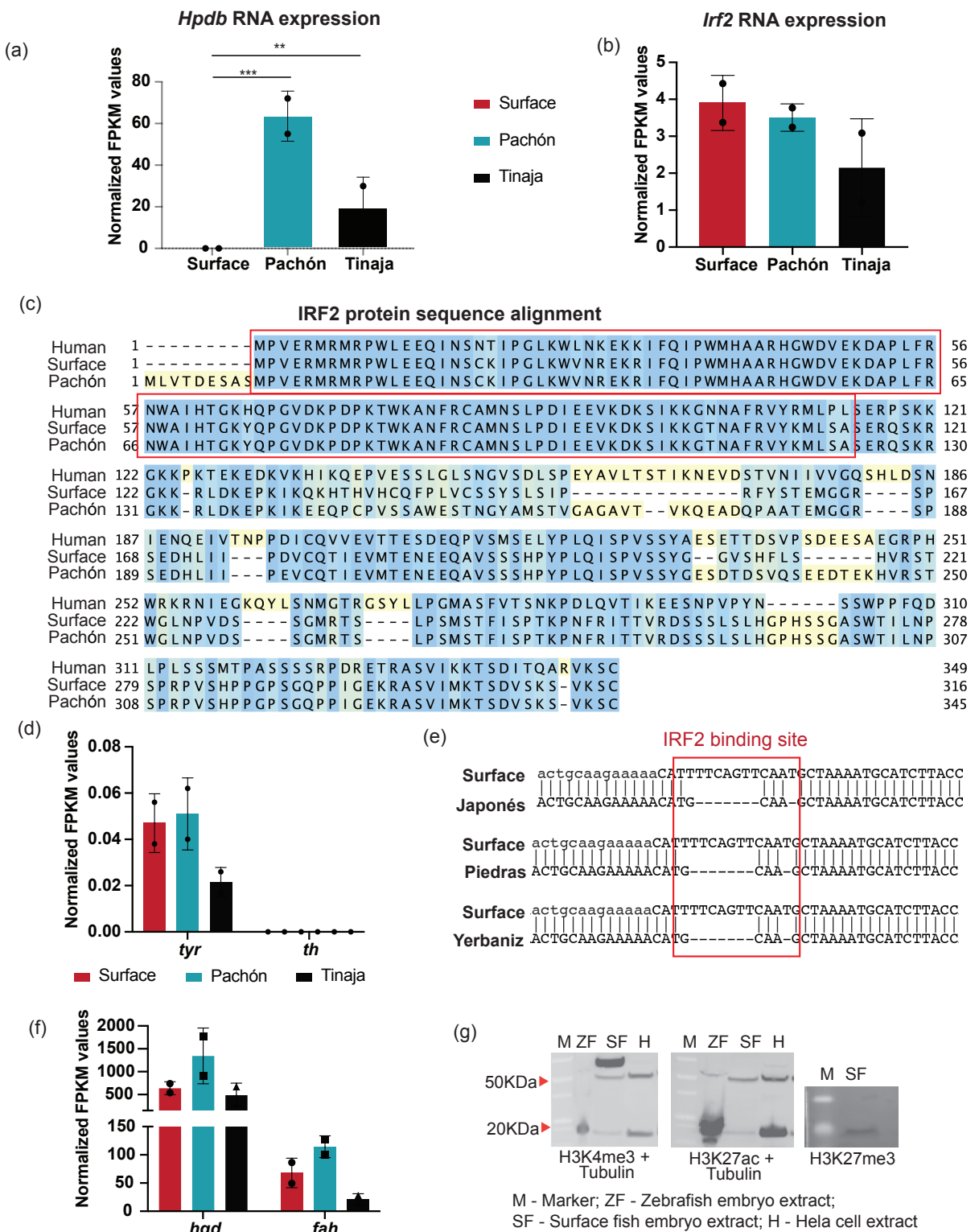
Supplementary Figure 3: (a) Lipid metabolism network genes are overrepresented among genes associated with c-CREs. The figure shows lipid metabolism network centered around insulin highlights multiple genes from our dataset (colored molecules) that are differentially regulated between surface and Pachón. (b) HNF4a target genes are involved in carbohydrate metabolism network. Many differentially expressed genes between surface and Pachón are direct/indirect targets of HNF4a as highlighted (colored molecules) in the network. (c) HNF4a target genes are involved in lipid metabolism network. Many differentially expressed genes between surface and Pachón are direct/indirect targets of HNF4a as highlighted (colored molecules) in the network.



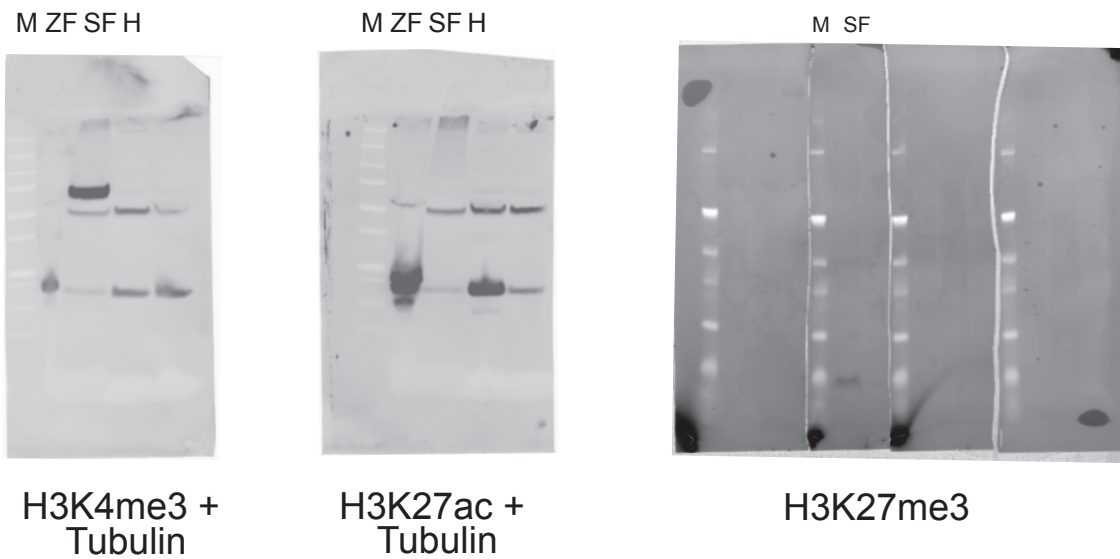
Supplementary Figure 4: (a) Expression levels of transcription factors represented as normalized FPKM values in each morphotype. The graph represents average values +/- s.d. from n=2 biologically independent RNA-seq experiments. p-values were calculated using two-tailed student's t-test. (b, c) No. of SNPs (b) and Indels (c) between surface and Pachón (S vs P), surface and Tinaja (S vs T), surface and either of the two cave populations (S vs (P or T)) and between surface and both the cave populations (S vs (P and T)).



Supplementary Figure 5: (a) Schematic of the reporter construct used for assaying enhancer activity in *A. mexicanus* and zebrafish larvae. (b) Representative bright field image of uninjected 4 days post fertilization *A. mexicanus* surface fish larvae, injected with vector alone (Vector) and injected with surface and the Pachón alleles of candidate CREs (see methods for more details). Scale bar = 0.5 mm. (c) Representative bright field image of uninjected and injected 4 dpf zebrafish larvae with vector alone (Top panel) and 4 dpf zebrafish larvae injected with surface and the Pachón alleles of candidate CREs. (d) F1 stable zebrafish line expressing surface and Pachón CRE_15, respectively. Reporter expression is seen in the liver/gut region (yellow arrowheads). Scale bars = 0.5 mm.



Supplementary Figure 6: (a) Expression levels of *hpdb* gene in surface, Pachón and Tinaja livers from RNA-seq experiment. (b) Expression levels of *Irf2* gene in the livers of adult surface, Pachón and Tinaja. (c) Alignment for human and Pachón cavefish IRF2 protein. Amino acids within the red box comprises the highly conserved DNA-binding domain of IRF2. (d) Expression levels of genes belonging to tyrosine metabolism that make melanin. *Tyr*: Tyrosinase, *th*: Tyrosine hydroxylase. (e) DNA sequence alignments for the wild-caught Japonés, Piedras and Yerbaniz fishes showing the deletion of IRF2 binding site. (f) Expression levels of genes belonging to tyrosine metabolism downstream to *Hpdb*, measured in the livers of adult surface, Pachón and Tinaja. *Hgd*: Homogentisate 1,2- dioxygenase, *Fah*: fumarylacetoacetate hydrolase. All bar graphs represent average values \pm s.d. from $n=2$ biologically independent RNA-seq experiments. * indicates p value <0.05 and ** indicates p value <0.005 using two-tailed student's t -test. (g) Western blots to validate antibodies against H3K4me3, H3K27ac, and H3K27me3. Anti-tubulin was used as loading control. Original uncropped blots are presented in Supplementary figure 7.



M - Marker; ZF - Zebrafish embryo extract;
 SF - Surface fish embryo extract; H - Hela cell extract

Supplementary Figure 7: Original uncropped images for Western blots showed in Supplementary Figure 6g.

Supplementary Table 1– *Astyanax* CREs overlap with human liver enhancers

This Excel sheet provides gerp loci that are orthologous to mapped CREs in the study. The coordinates are in Astmex 1 genome.

gerp = Genomic Evolutionary Rate Profiling

gerp coordinates are loci that are conserved between human and *Astyanax*.

Atac.width – width of the ATAC-seq peak.

Supplementary Table 2 – Candidates for functional assay

Coordinates and nearest gene descriptions of the CREs tested in the cell lines.

Supplementary Table 3 – Sequencing QC

The Excel sheet provides quality control data for all the genomics experiments

Supplementary Table 4 – List of primers

List of primers and oligos used in the study

Supplementary Table 1

gerp.chr	gerp.start	gerp.end	atac.chr	atac.start	atac.end	atac.peak.na	atac.width
KB882187	313632	313634	KB882187	313052	313837	ATAC_69725	786
KB882187	313631	313632	KB882187	313052	313837	ATAC_69725	786
KB882187	313519	313631	KB882187	313052	313837	ATAC_69725	786
KB882179	2033362	2033571	KB882179	2032949	2033987	ATAC_68604	1039
KB882179	2033576	2033587	KB882179	2032949	2033987	ATAC_68604	1039
KB882179	2038158	2038175	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038175	2038179	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038182	2038189	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038189	2038195	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038211	2038212	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038212	2038223	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038223	2038282	KB882179	2038077	2038467	ATAC_68606	391
KB882179	2038282	2038348	KB882179	2038077	2038467	ATAC_68606	391
KB872048	43374	43466	KB872048	42981	44063	ATAC_28460	1083
KB872048	43360	43363	KB872048	42981	44063	ATAC_28460	1083
KB872048	43292	43316	KB872048	42981	44063	ATAC_28460	1083
KB872048	43234	43288	KB872048	42981	44063	ATAC_28460	1083
KB872048	43186	43231	KB872048	42981	44063	ATAC_28460	1083
KB871615	752068	752069	KB871615	751613	752512	ATAC_5049	900
KB882276	471611	471614	KB882276	471124	472444	ATAC_82490	1321
KB882279	1103427	1103475	KB882279	1103382	1104071	ATAC_82891	690
KB872295	340588	340603	KB872295	339441	340861	ATAC_34119	1421
KB882146	1860708	1860709	KB882146	1859547	1861894	ATAC_62181	2348
KB882173	1560183	1560184	KB882173	1560067	1560772	ATAC_67477	706
KB872043	47412	47415	KB872043	46789	47491	ATAC_28357	703
KB871633	1305786	1305800	KB871633	1305630	1306221	ATAC_6879	592
KB871601	499967	499973	KB871601	499828	500316	ATAC_3747	489
KB871586	685	687	KB871586	365	993	ATAC_2282	629
KB872367	13027	13036	KB872367	12401	13906	ATAC_35721	1506
KB882267	1090697	1090708	KB882267	1089722	1090923	ATAC_81144	1202
KB882230	1341923	1341924	KB882230	1341693	1343132	ATAC_76424	1440
KB872081	2322129	2322148	KB872081	2321050	2322357	ATAC_29500	1308
KB871615	691535	691543	KB871615	691158	692379	ATAC_5041	1222
KB882177	653613	653614	KB882177	653371	655098	ATAC_68132	1728
KB882193	1501882	1501887	KB882193	1501407	1503644	ATAC_70718	2238
KB871938	4185456	4185460	KB871938	4184948	4185880	ATAC_24857	933
KB871668	12493	12507	KB871668	12350	13316	ATAC_9428	967
KB882184	1149742	1149769	KB882184	1148858	1149784	ATAC_69306	927
KB882227	496453	496454	KB882227	495978	498188	ATAC_75929	2211
KB882224	961535	961649	KB882224	960807	961682	ATAC_75567	876
KB882083	1161942	1161947	KB882083	1160582	1162333	ATAC_44963	1752
KB882231	1569399	1569403	KB882231	1568376	1569795	ATAC_76569	1420
KB882202	143136	143137	KB882202	141406	143509	ATAC_71980	2104
KB882209	1761892	1761893	KB882209	1761393	1762393	ATAC_73509	1001

Supplementary Table 1 continued

KB871668	12439	12440	KB871668	12350	13316	ATAC_9428	967
KB882196	940187	940188	KB882196	938697	940382	ATAC_71132	1686
KB882089	2009829	2009839	KB882089	2009175	2010194	ATAC_47048	1020
KB882118	2670309	2670337	KB882118	2669855	2671426	ATAC_55559	1572
KB872059	29027	29028	KB872059	28271	29834	ATAC_28759	1564
KB882105	3040799	3040804	KB882105	3039975	3041007	ATAC_51855	1033
KB871747	600651	600656	KB871747	599622	601264	ATAC_14729	1643
KB872013	377173	377293	KB872013	376828	379207	ATAC_27548	2380
KB871663	614254	614259	KB871663	614208	614993	ATAC_9156	786
KB882105	3040572	3040573	KB882105	3039975	3041007	ATAC_51855	1033
KB882143	1681123	1681171	KB882143	1680961	1682502	ATAC_61224	1542
KB872026	73840	73974	KB872026	73714	74276	ATAC_27876	563
KB882312	105772	105857	KB882312	105518	106441	ATAC_86107	924
KB882084	3846069	3846082	KB882084	3845242	3846812	ATAC_45517	1571
KB872498	47664	47665	KB872498	47049	47877	ATAC_37042	829
KB882179	2038348	2038351	KB882179	2038077	2038467	ATAC_68606	391
KB882251	1492967	1492973	KB882251	1492743	1493331	ATAC_79288	589
KB882297	1028781	1028791	KB882297	1028464	1029811	ATAC_84755	1348
KB871630	5460928	5460936	KB871630	5460558	5461241	ATAC_6588	684
KB882090	3841535	3841560	KB882090	3840822	3841997	ATAC_47466	1176
KB882273	73107	73109	KB882273	71950	73268	ATAC_82060	1319
KB882163	2426620	2426629	KB882163	2426028	2426742	ATAC_65513	715
KB871864	38321	38322	KB871864	37848	39277	ATAC_21607	1430
KB882187	1271242	1271248	KB882187	1270372	1271253	ATAC_69779	882
KB872048	158399	158400	KB872048	158318	159081	ATAC_28467	764
KB882297	1028781	1028791	KB882297	1028464	1029811	ATAC_84755	1348
KB882088	3752756	3752843	KB882088	3752561	3752848	ATAC_46855	288
KB871592	56235	56264	KB871592	54711	56558	ATAC_2846	1848
KB882244	1087369	1087376	KB882244	1087075	1087655	ATAC_78458	581
KB872029	225035	225047	KB872029	223050	225047	ATAC_28005	1998
KB882124	2392205	2392213	KB882124	2391697	2393457	ATAC_57043	1761
KB871793	430097	430112	KB871793	429344	430376	ATAC_17877	1033
KB882179	2018181	2018183	KB882179	2017387	2018390	ATAC_68603	1004
KB882214	880297	880306	KB882214	879843	880966	ATAC_74237	1124
KB882147	1666068	1666251	KB882147	1666036	1666332	ATAC_62409	297
KB882086	948618	948620	KB882086	948571	950339	ATAC_45956	1769
KB882145	7044550	7044644	KB882145	7043575	7044662	ATAC_61964	1088
KB882144	2411717	2411721	KB882144	2411006	2411828	ATAC_61485	823
KB882199	920653	920724	KB882199	918710	921288	ATAC_71605	2579
KB872164	77906	77908	KB872164	77064	78604	ATAC_31956	1541
KB872273	83780	83781	KB872273	83017	84229	ATAC_33768	1213
KB872082	893945	894052	KB872082	892896	894968	ATAC_29747	2073
KB882106	4096705	4096706	KB882106	4096326	4096864	ATAC_52251	539
KB882103	526032	526038	KB882103	525864	527052	ATAC_50800	1189
KB878679	1206	1215	KB878679	345	1287	ATAC_43261	943

Supplementary Table 1 continued

KB882105	2031704	2031730	KB882105	2031704	2032036	ATAC_51778	333
KB882205	1730478	1730486	KB882205	1729904	1732329	ATAC_72843	2426
KB882123	256671	256727	KB882123	255944	257212	ATAC_56589	1269
KB882239	190010	190044	KB882239	188912	190351	ATAC_77708	1440
KB873139	17211	17212	KB873139	16289	18464	ATAC_40593	2176
KB871701	218352	218355	KB871701	218079	218882	ATAC_12045	804
KB882126	212403	212408	KB882126	211085	212603	ATAC_57535	1519
KB882104	2410846	2410860	KB882104	2410420	2412079	ATAC_51514	1660
KB871797	471370	471371	KB871797	470450	471475	ATAC_18074	1026
KB872982	34064	34068	KB872982	33677	34663	ATAC_40202	987
KB882270	3568717	3568719	KB882270	3567048	3568913	ATAC_81656	1866
KB872818	1703244	1703246	KB872818	1702156	1703644	ATAC_39119	1489
KB872164	77906	77908	KB872164	77064	78604	ATAC_31956	1541
KB872498	47664	47665	KB872498	47049	47877	ATAC_37042	829
KB882303	319973	320089	KB882303	319163	320200	ATAC_85296	1038
KB882260	194342	194344	KB882260	193953	194589	ATAC_80223	637
KB882106	207799	207803	KB882106	207578	208449	ATAC_51931	872
KB882150	73565	73566	KB882150	73149	73796	ATAC_62900	648
KB882210	1378232	1378237	KB882210	1377600	1378282	ATAC_73629	683
KB882233	5739030	5739035	KB882233	5738623	5739879	ATAC_77074	1257
KB872615	9820	9821	KB872615	8792	10238	ATAC_37921	1447
KB871578	4360234	4360241	KB871578	4359184	4361413	ATAC_283	2230
KB872081	2322129	2322148	KB872081	2321050	2322357	ATAC_29500	1308
KB872029	225035	225047	KB872029	223050	225047	ATAC_28005	1998
KB871764	63753	63754	KB871764	63089	64448	ATAC_16284	1360
KB882106	282881	282885	KB882106	282563	283630	ATAC_51941	1068
KB882113	1867165	1867170	KB882113	1865837	1867371	ATAC_54073	1535

Supplementary Table 2

CRM number	coordinates	specificity	nearest.gene	gene description
CRM_1	APWO02000126.1:12116-12322	Cave-specific	ENSAMXG00000021129	inositol polyphosphate 5-phosphatase K-like
CRM_2	5:20862951-20863249	Surface-specific	ENSAMXG00000017885	Transducin like enhancer of split 2b
CRM_3	APWO02000046.1:5908919-5909487	Surface-specific	ENSAMXG00000018939	Cytochrome P450, family 2, subfamily K, polypeptide 8
CRM_4	APWO02002143.1:13670-13963	Cave-specific	ENSAMXG00000000698	protein phosphatase 1H-like [Source:NCBI gene;Acc:103035555]
CRM_5	11:11176042-11175510	Surface-specific	ENSAMXG00000004148	Hydroxypyruvate reductase (GRHPR)
CRM_6	APWO02001323.1:49828-50218	Cave-specific	ENSAMXG00000001621	-
CRM_7	APWO02000055.1:404521-404568	Surface-specific	ENSAMXG00000018424	kinesin family member 9
CRM_8	APWO02000139.1:4691964-4692266	Surface-specific	ENSAMXG00000008187	prostaglandin E2 receptor EP4 subtype
CRM_9	APWO02002377.1:5952-6319	Cave-specific	ENSAMXG00000016306	calcium/calmodulin-dependent protein kinase type 1D-like
CRM_10	18:7017570-7017917	Surface-specific	ENSAMXG00000003688	deoxyribodipyrimidine photo-lyase-like
CRM_11	9:36562402-36562882	Surface-specific	ENSAMXG00000018617	iodothyronine deiodinase 2
CRM_12	13:36360727-36361987	Cave-specific	ENSAMXG00000001899	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 4-like
CRM_13	5:8790638-8791408	Surface-specific	ENSAMXG00000008063	Fetuin B
CRM_14	8:5530266-5530674	Cave-specific	ENSAMXG00000018872	glutathione S-transferase pi 2
CRM_15	6:33370826-33370935	Cave-specific	ENSAMXG00000015502	4-hydroxyphenylpyruvate dioxygenase
CRM_16	18:17152529-17153211	Surface-specific	ENSAMXG00000003000	mitogen-activated protein kinase kinase 1
CRM_17	7:9582583-9583260	Cave-specific	ENSAMXG00000008354	slc47a3 Multidrug and toxin extrusion protein 1-like
CRM_18	7:24261100-24262368	Surface-specific	ENSAMXG00000019808	solute carrier family 16 member 7
CRM_19	20:10902951-10903397	Cave-specific	ENSAMXG00000016875	sodium/nucleoside cotransporter 2
CRM_20	APWO02000037.1:3462546-3463098	Cave-specific	ENSAMXG00000016393	squalene epoxidase
CRM_21	25:21808992-21810214	Surface-specific	ENSAMXG00000001966	apoptosis inducing factor, mitochondria associated 2
CRM_22	6:7155483-7156454	Surface-specific	ENSAMXG00000007498	synaptophysin a
CRM_23	17:3671345-3672232	Surface-specific	ENSAMXG00000012893	phosphotriesterase related
CRM_24	19:161341-162334	Surface-specific	ENSAMXG00000009850	nitric oxide synthase trafficking
CRM_25	6:33372486-33373251	Cave-specific	ENSAMXG00000015502	4-hydroxyphenylpyruvate dioxygenase

Supplementary Table 3

QC data for ChIP-seq, ATAC-seq and RNA-seq												
Sample name in Original data	Morphotype	Feature	Total reads	Unmapped reads	Unmapped reads	mapped reads	mapped reads	Multimappe d reads	Multimappe d reads	% Total mapped reads	Replicate	
s_sf3k27ac1	Surface	H3K27ac	28896164	4212313	14.58	19549051	67.65	5134800	17.77	85.42	1	
s_sf3k27ac2	Surface	H3K27ac	29287690	4169598	14.24	19908604	67.98	5209488	17.79	85.76	2	
s_sf3k4me31	Surface	H3K4me3	28284936	4484867	15.86	17396700	61.51	6403369	22.64	84.14	1	
s_sf3k4me32	Surface	H3K4me3	29097834	4530128	15.57	18118761	62.27	6448945	22.16	84.43	2	
s_surface_k27me3_a	Surface	H3K27me3	13313897	2786856	20.93	6895151	51.79	3631890	27.28	79.07	1	
s_surface_k27me3_b	Surface	H3K27me3	13486379	2801383	20.77	7061718	52.36	3623278	26.87	79.23	2	
s_surface_1	Surface	ATAC	39327874	8349780	21.23	28359185	72.11	2618909	6.66	84.81	1	
s_surface_2	Surface	ATAC	60873891	13205999	21.69	43494926	71.45	4172966	6.86	84.68	2	
s_sfinput	Surface	Input	30572883	4952270	16.2	17817175	58.28	7803438	25.52	83.8	1	Input for ChIP samples
s_surface_input	Surface	Input	27355910	4351054	15.91	16049710	58.67	6955146	25.42	84.09	2	Input for H3K27me3 ChIP
s_h3k27ac_1ug	Pachón	H3K27ac	27833853	3116030	11.2	19555813	70.26	5162010	18.55	88.8	1	
s_pachon_h3k27ac_1ug_b	Pachón	H3K27ac	30838536	4748285	15.4	20738671	67.25	5351580	17.35	84.6	2	
s_h3k4me3_1ug	Pachón	H3K4me3	29602992	4482011	15.14	18170864	61.38	6950117	23.48	84.86	1	
s_pachon_h3k4me3_1ug_b	Pachón	H3K4me3	32384913	5065249	15.64	19876454	61.38	7443210	22.98	84.36	2	
s_pachon_k27me3_a	Pachón	H3K27me3	14649941	2825828	19.29	7736917	52.81	4087196	27.9	80.71	1	
s_pachon_k27me3_b	Pachón	H3K27me3	15782262	2994775	18.98	8285425	52.5	4502062	28.53	81.02	2	
s_pachon_1	Pachón	ATAC	52080524	8550858	16.42	40290056	77.36	3239610	6.22	88.52	1	
s_pachon_2	Pachón	ATAC	47791467	8365804	17.5	36915377	77.24	2510286	5.25	88.94	2	
s_pachon_input	Pachón	Input	59420483	7780710	13.09	35856259	60.34	15783514	26.56	86.91	1	Input for ChIP samples
s_pachon_input	Pachón	Input	29918122	4159031	13.9	18022556	60.24	7736535	25.86	86.1	2	Input for H3K27me3 ChIP
s_tinajak27aca	Tinaja	H3K27ac	45553054	7038172	15.45	30352992	66.63	8161890	17.92	84.55	1	
s_tinajak27acb	Tinaja	H3K27ac	43342684	7500452	17.31	28276411	65.24	7565821	17.46	82.69	2	
s_tinaja_h3k27ac_b	Tinaja	H3K27ac	31221992	4193820	13.43	21053814	67.43	5974358	19.14	86.57	3	
s_tinajak4me3a	Tinaja	H3K4me3	49635852	8129866	16.38	31500926	63.46	10005060	20.16	83.62	1	
s_tinajak4me3b	Tinaja	H3K4me3	49633221	8305438	16.73	31648424	63.76	9679359	19.5	83.27	2	
s_tinaja_h3k4me3_a	Tinaja	H3K4me3	35138318	5499463	15.65	21916095	62.37	7722760	21.98	84.35	3	
s_tinaja_k27me3_a	Tinaja	H3K27me3	17427245	3429257	19.68	8960936	51.42	5037052	28.9	80.32	1	
s_tinaja_k27me3_b	Tinaja	H3K27me3	15064770	3072767	20.4	7784172	51.67	4207831	27.93	79.6	2	
s_tinaja_1	Tinaja	ATAC	57278843	12360989	21.58	40782123	71.2	4135731	7.22	84.31	1	
s_tinaja_2	Tinaja	ATAC	53957916	11929952	22.11	38304678	70.99	3723286	6.9	84.35	2	
s_tinaja_input	Tinaja	Input	39893215	5891832	14.77	23582776	59.11	10418607	26.12	85.23	1	Input for ChIP samples
s_tinaja_input	Tinaja	Input	29127120	4450498	15.28	17136677	58.83	7539945	25.89	84.72	2	Input for H3K27me3 ChIP

Supplementary Table 3 continued

MACS2 QC		
Morphotype	Chromatin feature/ Experiment	No. of peaks
Surface	H3K27ac	70089
Surface	H3K4me3	56323
Surface	H3K27me3	97948
Surface	ATAC	73675
Pachón	H3K27ac	93527
Pachón	H3K4me3	64432
Pachón	H3K27me3	98198
Pachón	ATAC	72521
Tinaja	H3K27ac	117125
Tinaja	H3K4me3	136219
Tinaja	H3K27me3	99370
Tinaja	ATAC	84653

Supplementary Table 4

S.No.	Primers	Sequence	product size	Use
1	CRM_1F	GGGGAGACAAAAGTCTCCTAGC	847bp	Amplifying candidate regions from <i>Astyanax</i> genome
2	CRM_1R	GCATGACGGCATTGCAAAC		Amplifying candidate regions from <i>Astyanax</i> genome
3	CRM_2F	TTTAAATAAACAGAGACTGAGTGAC	456bp	Amplifying candidate regions from <i>Astyanax</i> genome
4	CRM_2R	GTTTGTAGTGTATGTGTAATGCC		Amplifying candidate regions from <i>Astyanax</i> genome
5	CRM_3F	CTCCCTCTCCCCAGCTAG	1204bp	Amplifying candidate regions from <i>Astyanax</i> genome
6	CRM_3R	GGTTGAAGCTTTTGGGCG		Amplifying candidate regions from <i>Astyanax</i> genome
7	CRM_4F	AGCTAAAGGGGTTTTGATC	353bp	Amplifying candidate regions from <i>Astyanax</i> genome
8	CRM_4R	GGCAATACCTACTGTTTCTG		Amplifying candidate regions from <i>Astyanax</i> genome
9	CRM_5F	CTATGTCTCACAGTTCAGCAG	570bp	Amplifying candidate regions from <i>Astyanax</i> genome
10	CRM_5R	ACCCAGAACAGTATCTCCAAG		Amplifying candidate regions from <i>Astyanax</i> genome
11	CRM_6F	TCTATTTTCATGCCTTCCG	1042bp	Amplifying candidate regions from <i>Astyanax</i> genome
12	CRM_6R	TTTATGCAAGGGTTCACAATTC		Amplifying candidate regions from <i>Astyanax</i> genome
13	CRM_7F	AACATGCTCTCTGCGTAGG	853bp	Amplifying candidate regions from <i>Astyanax</i> genome
14	CRM_7R	CCACAGTTTTCCACTGCAAAG		Amplifying candidate regions from <i>Astyanax</i> genome
15	CRM_8F	TCCTGCAATCTCCAACAAAC	320bp	Amplifying candidate regions from <i>Astyanax</i> genome
16	CRM_8R	TATTGTCCCTGTTCCAGAG		Amplifying candidate regions from <i>Astyanax</i> genome
17	CRM_9F	AGGTGGTGTAACTACTACAAAAG	804bp	Amplifying candidate regions from <i>Astyanax</i> genome
18	CRM_9R	GCTCCACCTACTGTTCCC		Amplifying candidate regions from <i>Astyanax</i> genome
19	CRM_10F	GTGTGTGTTCTTAAGTITTTG	772bp	Amplifying candidate regions from <i>Astyanax</i> genome
20	CRM_10R	TCCATGTCTACTATGCC		Amplifying candidate regions from <i>Astyanax</i> genome
21	CRM_18F	TGCAGGGTGTACAATAGGGC	1331bp	Amplifying candidate regions from <i>Astyanax</i> genome
22	CRM_18R	TACAACAGTGCAACCAGGCA		Amplifying candidate regions from <i>Astyanax</i> genome
23	CRM_12F	ACACAGAGCTCCTGTTGTCG	2043bp	Amplifying candidate regions from <i>Astyanax</i> genome
24	CRM_12R	GCTACTGGTGATCCCACACC		Amplifying candidate regions from <i>Astyanax</i> genome
25	CRM_14F	AACAGACCGTTCATGCTGGA	1216bp	Amplifying candidate regions from <i>Astyanax</i> genome
26	CRM_14R	AAATTACCCAGGACCCCT		Amplifying candidate regions from <i>Astyanax</i> genome
27	CRM_11F	CAGGTCCTGTTTTGAGGGGTT	1668bp	Amplifying candidate regions from <i>Astyanax</i> genome
28	CRM_11R	TGCCCTATCCAAAGGGAAC		Amplifying candidate regions from <i>Astyanax</i> genome
29	CRM_15F	AGTGAGCGGATGGTGTTCCAG	1002bp	Amplifying candidate regions from <i>Astyanax</i> genome
30	CRM_15R	TCCATACGGGAACCTAATGCC		Amplifying candidate regions from <i>Astyanax</i> genome
31	CRM_17F	GCACTTACAGGCCAACAGC	1400bp	Amplifying candidate regions from <i>Astyanax</i> genome
32	CRM_17R	TCACCTGTGATCACCAACAGT		Amplifying candidate regions from <i>Astyanax</i> genome
33	CRM_21F	CCACTGGGAACCTTGCACTGT	2269bp	Amplifying candidate regions from <i>Astyanax</i> genome
34	CRM_21R	GGCACTTTGATTTGGTGGGG		Amplifying candidate regions from <i>Astyanax</i> genome
35	CRM_19F	AAGCTCCACTGCTGATTCGC	1431bp	Amplifying candidate regions from <i>Astyanax</i> genome
36	CRM_19R	TTTGAGGGATAGAGTCATTAACCA		Amplifying candidate regions from <i>Astyanax</i> genome
37	CRM_16F	GCACAAATCTACATCCTGGCA	1565bp	Amplifying candidate regions from <i>Astyanax</i> genome
38	CRM_16R	GAGTGTCTTTGGGCCGTTA		Amplifying candidate regions from <i>Astyanax</i> genome
39	CRM_13F	TACACAGCGCAATGGAACA	2009bp	Amplifying candidate regions from <i>Astyanax</i> genome
40	CRM_13R	GCCAATGGCAACAGACTGAAC		Amplifying candidate regions from <i>Astyanax</i> genome
41	CRM_20F	AGATGGAAGCGTCAGCAGTC	1359bp	Amplifying candidate regions from <i>Astyanax</i> genome
42	CRM_20R	TCAACACGTGTGCGTTTAGC		Amplifying candidate regions from <i>Astyanax</i> genome
43	HLC-PCRtoPGL_gibsonF	agctcgtagcctcgaggatGAGGATATCGAGCTCG		Transferring candidate regions from fish reporter
44	HLC-PCRtoPGL_gibsonR	cgccgagggccagatcttgatAGGGTAATGAGGGCC		vector to luciferase vector
45	cf_rpl13a_2F	CGCAACAAATTGAAGTACCTG		qPCR housekeeping gene
46	cf_rpl13a_2R	GGTTCGTGTTTCATCCTCTTG		qPCR housekeeping gene
47	PGL_SHPD2-del-F	AAAAAACTGCAAGAAAAACAGCAATGCTAAAATGCATCTTAC		Gibson primer for mutating IRF2 binding site in
48	PGL_SHPD2-del-R	TGTTTTTCTTGCAGTTTTTTTTATGTTGTGCAG		S_CRM_15 cloned in reporter vector
49	S_Hpd2_del_F	aaaaaCATTTCAGTTCAATGC		Oligo for EMSA
50	S_Hpd2_del_R	GCATTGAACTGAAAATGTTTT		Oligo for EMSA
53	P_Hpd2_del20_F	AGAAAAACATGCAAGCTAAA		Oligo for EMSA
54	P_Hpd2_del20_R	TTTAGCTTGCATGTTTTTCT		Oligo for EMSA
55	PGL_PHPD2-undel-F	aaaaaCATTTCAGTTCAATGCTAAAATGCATCTTACCAT		Gibson primer for adding IRF2 binding site in
56	PGL_PHPD2-undel-R	ATTGAACTGAAAATGTTTTctgcagttttttatgtgtg		C_CRM_15 cloned in reporter vector
57	Hpdb_mRNA_F1	GAAAGAGGTGCCATCGTTGT		qPCR
58	Hpdb_mRNA_R2	CTCCCTCTCCACCAGAGTG		qPCR
59	Hpdb_mRNA_F2	CACATTGCCATGAACACCTC		qPCR
60	Hpdb_mRNA_R2	CCTCCAGAATGCTGATGTCC		qPCR
61	Hpdb_allele_genRegion_F1	GGAATACAGTGGGCGTGTCC		Genotyping
62	Hpdb_allele_genRegion_R1	AAATCTGAAGGAGCGTGCCA		Genotyping
63	Hpdb_allele_genRegion_F2	TGTCTGGACTGGTTTTGGTG		Genotyping
64	Hpdb_allele_genRegion_R2	GCACATTGCCATGAACACCT		Genotyping
65	pGL4.23_RV3	TAGCAAAATAGGCTGTCCCC		Sequencing of pGL4.23 vectors
66	pUC57-to-pGL-F	gagctcgtagcctcgaggacCTCGCAATGCATCTAGAT		For transferring candidate regions cloned in pUC57 to
67	pUC57-to-pGL-R	gccgagggccagatcttgataCGACGGGCCCCGGGATCCG		luciferase vector

1. S-CRM_15

2. P-CRM_15

1	1	AGTGAGCGGATGGTGTTCAGTGGCTGAATCTTTGATACCTCCACCCCTCATGAGAACCAATAGTATCCTT	70
2	1	AGTGAGCGGATGGTGTTCAGTGGCTGAATCTTTGATACCTCCACCCCTCATGAGAACCAATAGCATCCTT	70
1	71	TTCTACAGAAACTCCTCTAAGGTTCAATGAGTTCACAGCAAAGATGAAGAACATAGGGATAACACTGTTT	140
2	71	TTCTACAGAAACTCCTCTAAGGTTCAATGAGTTCACAGCAAAGATGAAGAACATAGGGATAACACTGTTT	140
1	141	GCAGAAAACACATTAGAAAATAAGCTGACCTGTTTTGCATTTCCCACCCAGAAGGTGATGTGGTCAAAGC	210
2	141	GCAGAAAACACATTAGAAAATAAGCTGACCTGTTTTGCATTTCCCACCCAGAAGGTGATGTGGTCAAAGC	210
1	211	AGACGAACCTGCCATGTTTCATGCTGATGGAAGAAAATCGGGTAAATCATTTTTATATATAAAAATAAATAT	280
2	211	AGACGAACCTGCCCTGTTTCATGCTGATGGAAGAAAATCGGGTAAATCATTTTTATATATAAAAATAAATAT	280
1	281	ACTGAATACTCACTGCAATTGCTCAGATTTAAATGCAAAAATGTATATATAAAAAAGTTTGTTTTATATA	350
2	281	ACTGAATACTCACTGCAATTGCTCAGATTTAAATGCAAAAATTTATATATAAAAAAGTTTGTTTTATATA	350
1	351	TTTGTGTGTTGGCTTTTTTACATCCTATATTACTCTGTTATCTGTCAAAGGTATAATTATAAAGCATATT	420
2	351	TTTGTGTGTTGGCTTTTTTACATCCTATATTACTCTGTTATCTGTCAAAGGTATAATTATAAAGCATATT	420
1	421	ACTTTTCCTCAAAGACATTTACTACTAAGAGTAAAAGTGGTTCCTGAGTCACATACCTTTTTCACCTTTGT	490
2	421	ACTTTTCCTCAAAGACATTTACTACTAAGAGTAAAAGTGGTTCCTGAGTCACATACCTTTTTCACCTTTGT	490

1 491 CTGTGTAAGTTGTCTAAAAAAAAAACAGAGAAAAACACTGTTTTTGTATTTCTTTTTTATATTTTTGAGTG 560
2 491 CTGTGTAAGTTGTCTAAAAAAAAAACAGAGAAAAACACTGTTTTTGTATTTCTTTTTTATATTTTTGAGTG 560

TT

1 561 GAAGGTCATTAGCTTCAGATGCTTTTTTCATAAAAAAA TAAT AATAATAATAATAATATCTGCACAACATA 630
2 561 GAAGGTCATTAGCTTCAGATGCTTTTTTCATAAAAAAA -AA AATAATAATAATAATATCTGCACAACATA 629

IRF2-del

1 631 AAAAAA ACTGCAAGAAAAACA TTTTCAGTTCAAT GCTAAAATGCATCTTACCATGCTGATGTTCTGCAGG 700
2 630 AAAAAA ACTGCAAGAAAAACA - - - - - TGCAA - GCTAAAATGCATCTTACCATGCTGATGTTCTGCAGG 691

1 701 TCCTGTTTTGAGGGGTTTTGAGGCAGAATGTTGGTCCAACCAGAATTTAAAGAGCAAGAGAAGCTAGGGG 770
2 692 TCCTGTTTTGAGGGGTTTTGAGGCAGAATGTTGGTCCAACCAGAATTTAAAGAGCAAGAGAAGCTAGGGG 761

1 771 GGCGTTTCCCCCTCCACCGGATGCATGCAGTATGCAGAACTTACCTCAAATTAATGCCTGGTCTGGAC 840
2 762 GGCGTTTCCCCCTCCACCGGATGCATGCAGTATGCAGAACTTACCTCAAATTAATGCCTGGTCTGGAC 831

ATA

1 841 TGATGACAATTATCACAAA ACTAATGAGAG - - AGTTA TTGGCACCTAATTTGAAATAATAT - AAGTATAA 907
2 832 TGATGACAATTATCACAAA ACTAATGAGAG ATAGTTT TTGGCACCTAATTTGAAATAATAT A AAGTATAA 901

1 908 TTGATCACAAATGCAATGGAATTTAAAAT C CACACACTAACAGGTTTCTTATGTGCAGCTGCATAATAGAC 977
2 902 TTGATCACAAATGCAATGGAATTTAAAAT T CACACACTAACAGGTTTCTTATGTGCAGCTGCATAATAGAC 971

1 978 AT A ATACAAAAGGCATTAGGTT - CCGTATGGA 1008

2 972 AT - ATACAAAAGGCATTAGGTT C CCGTATGGA 1002