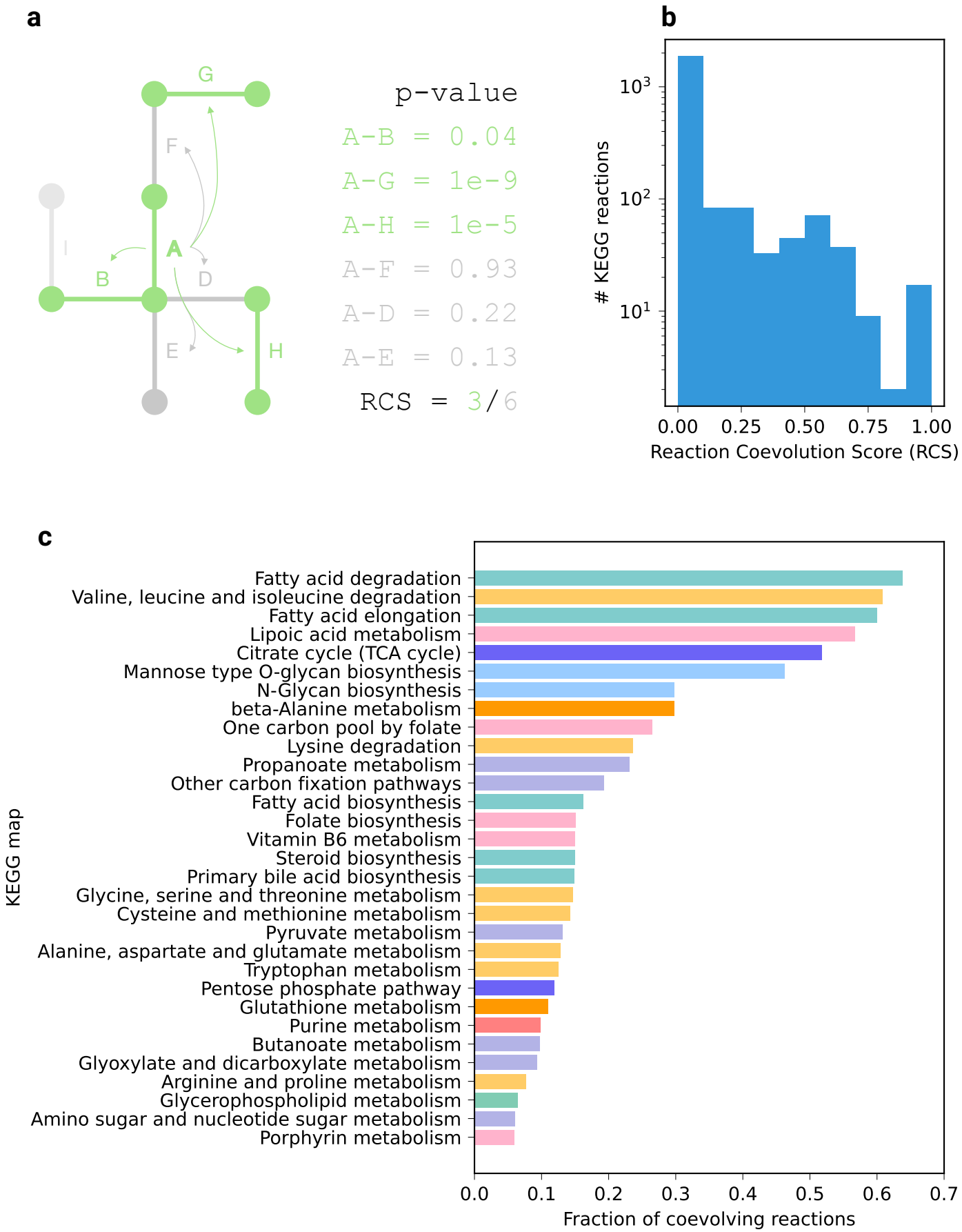


Figure S1. Reaction Coevolution Score (RCS).

a, Example of comparison between orthogroups mapped to the focus reaction (bold) and orthogroups mapped to neighborhood reaction (shortest path <=2). Three out six comparisons have a significant cotr score (green curves), producing an RCS of 0.5. Reactions without a mapped human gene (light gray) are not considered in the calculation. **b**, Distribution of RCS values in 2256 reactions with mapped human genes the 125 KEGG maps. **c** Ranking of KEGG maps according to the fraction of reactions with RCS>0; color code is as in Figure 1.

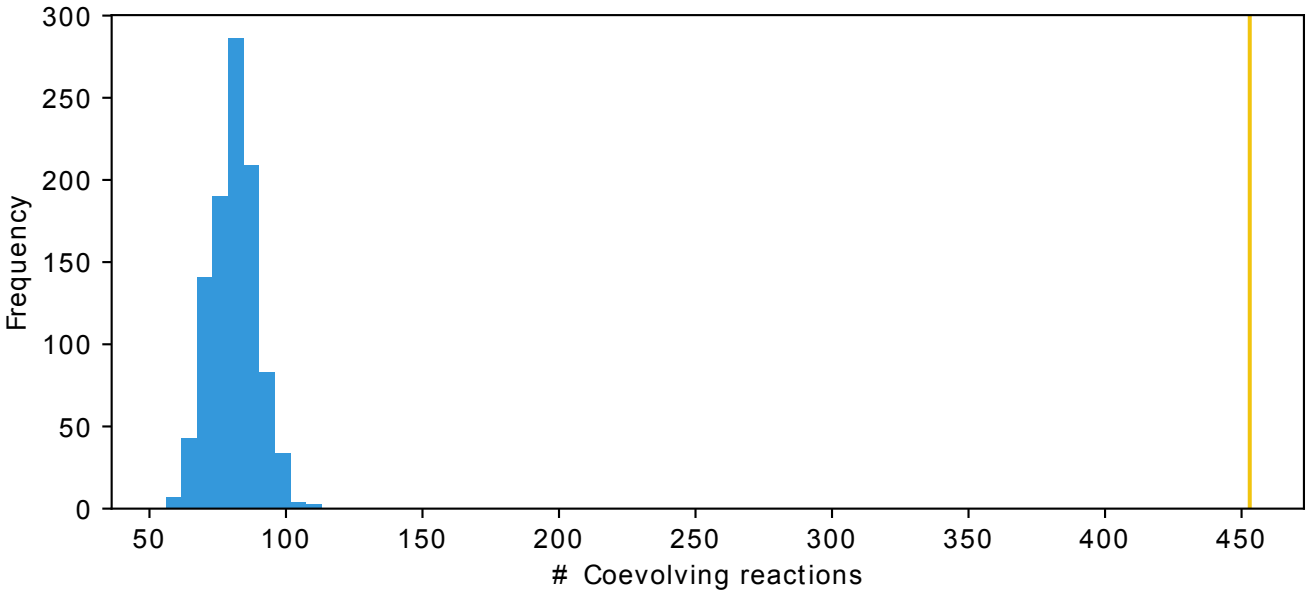


Figure S2. RCS obtained by random permutations

Distribution of the number of reactions with RCS>0 obtained by permutations (n=1000) of reaction/gene assignments across different KEGG maps compared with the number of reactions with RCS>0 (n=453) obtained in the original dataset.

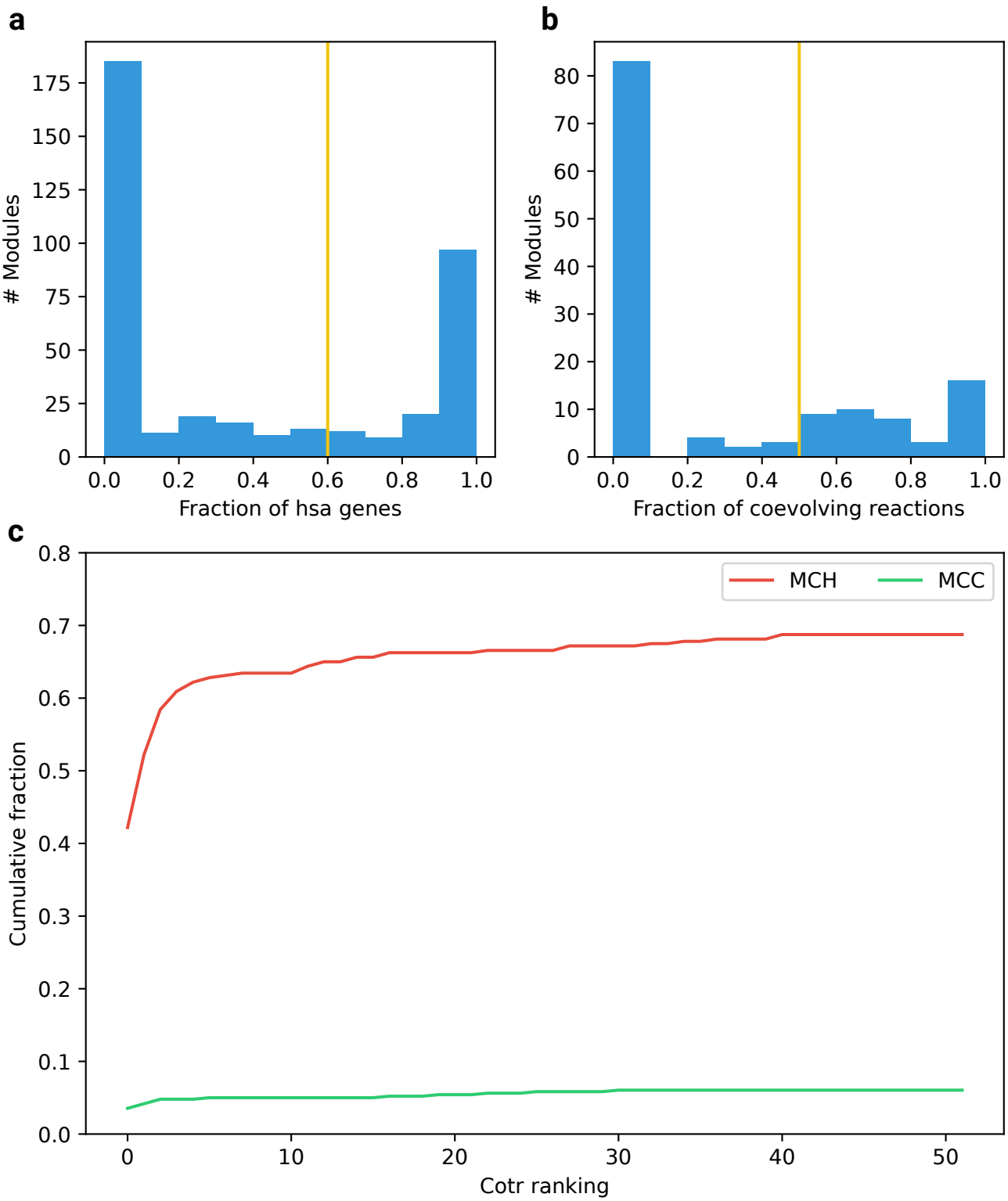


Figure S3. Coevolution in human metabolic pathways (KEGG modules)

a, Distribution of the fraction of reactions mapped to human genes in KEGG modules (n=392); modules with f > 0.6 (orange line) were considered present in humans (n=138). **b**, Distribution of the module coevolution score (MCS) in 138 human pathways; modules with MCS ≥ 0.5 (orange line) were considered coevolution hotspots. **c**, Fraction of reactions in modules with correctly predicted genes as a function of the cotr ranking (according to the gene p-value) in metabolic coevolution hotspots (MCH) and coldspots (MCC).

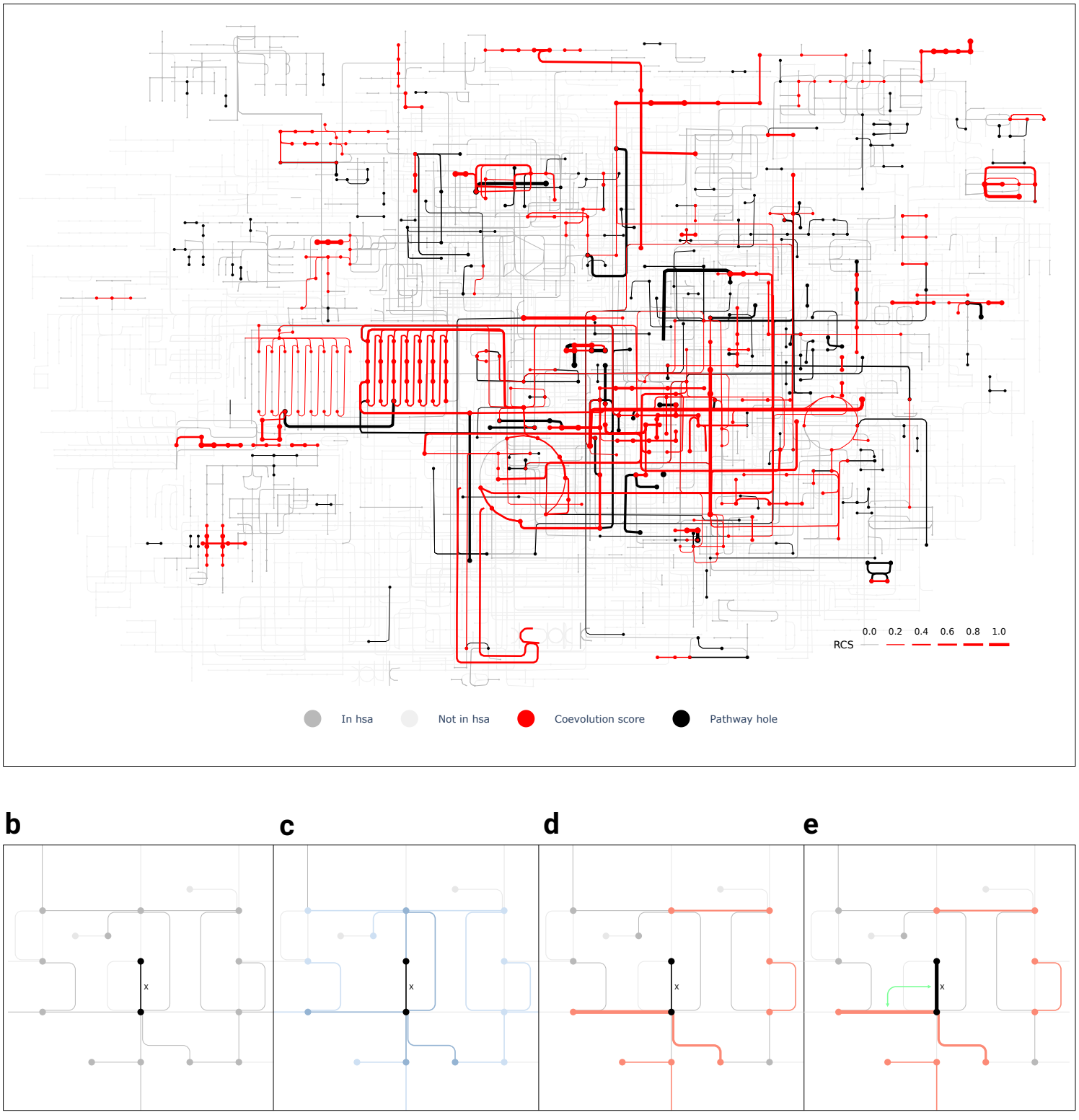


Figure S4 Pathway holes in human metabolism

a, Metabolic Pathways map showing unmapped reactions (black lines) connected to reactions mapped to human genes with (red lines) or without (dark gray lines) a positive reaction coevolution score (RCS); the thickness of the lines is proportional to the RCS value. B-E. Schematic example of a portion of a KEGG map illustrating how pathway holes are scored. Lines represent reactions, circles represent compounds. **b**, An unassigned reaction (X, the pathway hole) contoured by reactions mapped to human genes (dark gray). **c**, RCS values of reactions within two steps of X (blue) are considered. **d**, Reactions with positive RCS (red) are shown with a line thickness proportional to their RCS values. **e**, The pathway hole (X) inherits the highest RCS (green arrow) among the nearby values.

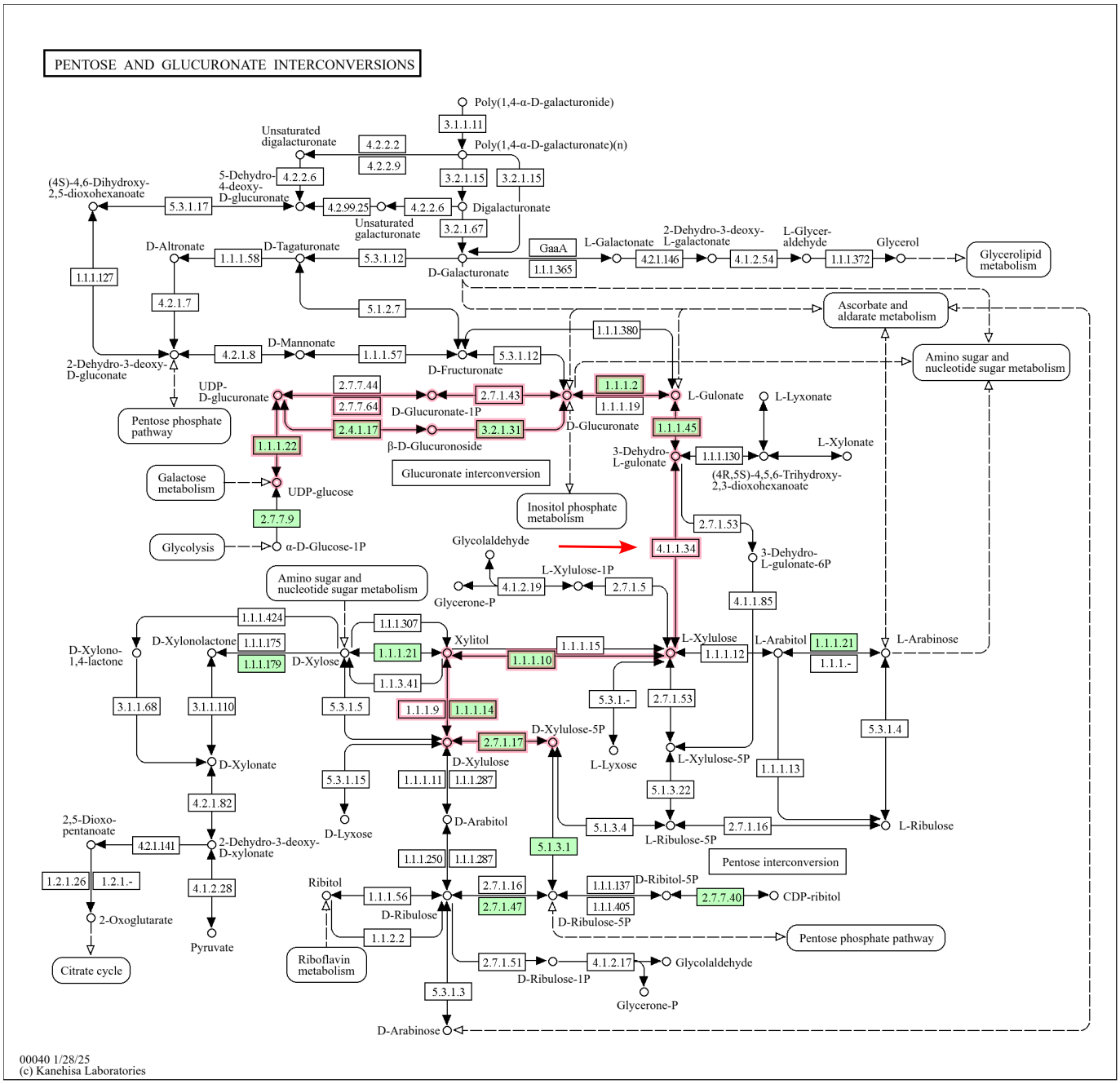


Figure S5 KEGG map of the unassigned 4.1.1.34 reaction

KEGG Orthology (KO) view of the "Pentose and glucuronate interconversion" map (https://www.genome.jp/pathway/ko00040) of *Homo sapiens*, with the Glucuronate pathway module (M00014) highlighted in red. The unassigned 4.1.1.34 reaction (red arrow) is represented by a white box.

Co-evolved Genes in All Clades for gene C11ORF54

This table presents the top 100 co-evolved genes with the query gene in each of 17 clades. A gene's per clade score indicates how co-evolved the gene is with the query from 1 to 100 (lower is better). The CladeOScope score indicates how co-evolved the gene is across all the clades. Lower CladeOScope score = stronger co-evolution.

Results Table

Show 10 entries Search:

Gene	CladeOScope Score	Eukaryota	Chordata	Mammalia	Archelosauria	Ecdysozoa	Nematoda	Arthropoda	Platyhelminthes	Alveolata
CRYL1	1	-	-	-	-	001	001	090	-	NA
TOP3A	1	-	-	-	-	-	-	-	001	NA
DPP3	2	-	-	-	-	081	002	-	-	NA
TTLL12	2	-	004	018	046	-	-	-	-	NA
DPY30	2	-	020	001	-	-	-	-	-	NA
SLC25A29	2.24	-	-	-	-	-	005	-	-	NA
TAF5	3.46	-	-	-	006	-	-	-	-	NA
LSS	3.46	-	026	003	-	-	-	-	-	NA
ADAR	3.46	-	-	-	-	-	004	-	-	NA
ALDH3A2	3.74	-	-	-	-	-	-	001	-	NA

Showing 1 to 10 of 817 entries Previous 1 2 3 4 5 … 82 Next

Figure S6 C11orf54 co-evolutionary associations in CladeOScope

Output of a CladOScope (https://tabachlab.shinyapps.io/CladeOScope/) database search using C11orf54 as a query, showing the top-ranked position of CRYL1 (arrow).

a

Your Input:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
C11orf54	*Ester hydrolase C11orf54; Exhibits ester hydrolase activity on the substrate p- nitrophenyl acetate. (315 aa)*									

Predicted Functional Partners:

CEP295	*Centrosomal protein of 295 kDa; Centriole-enriched microtubule-binding protein involved in centriole biogenesis. Essential f...*									0.622
VSTM5	*V-set and transmembrane domain-containing protein 5; Cell adhesion-like membrane protein of the central nervous system ...*									0.585
CCDC82	*Coiled-coil domain containing 82.*									0.506
MED17	*Mediator of RNA polymerase II transcription subunit 17; Component of the Mediator complex, a coactivator involved in the r...*									0.498
SMCO4	*Single-pass membrane protein with coiled-coil domains 4; Belongs to the SMCO4 family.*									0.492
TBC1D22B	*TBC1 domain family member 22B; May act as a GTPase-activating protein for Rab family protein(s).*									0.490
AKR1C2	*Aldo-keto reductase family 1 member C2; Works in concert with the 5-alpha/5-beta-steroid reductases to convert steroid ho...*									0.483
URB2	*URB2 ribosome biogenesis homolog.*									0.480
SERPINB5	*Serpin B5; Tumor suppressor. It blocks the growth, invasion, and metastatic properties of mammary tumors. As it does not ...*									0.475
GNPDA2	*Glucosamine-6-phosphate deaminase 2.*									0.469

b

Your Input:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
C11orf54	*Ester hydrolase C11orf54; Exhibits ester hydrolase activity on the substrate p- nitrophenyl acetate. (315 aa)*									

Predicted Functional Partners:

TRIP13	*Pachytene checkpoint protein 2 homolog; Plays a key role in chromosome recombination and chromosome structure develop...*		X	X		X	X	X	X	0.041
TTLL12	*Tubulin--tyrosine ligase-like protein 12; Negatively regulates post-translational modifications of tubulin, including detyrosinati...*		X	X		X	X	X	X	0.041
CORO1A	*Coronin-1A; May be a crucial component of the cytoskeleton of highly motile cells, functioning both in the invagination of larg...*		X	X		X	X	X	X	0.041
DECR2	*Peroxisomal 2,4-dienoyl-CoA reductase; Auxiliary enzyme of beta-oxidation. Participates in the degradation of unsaturated fat...*		X	X		X	X	X	X	0.041
DECR1	*2,4-dienoyl-CoA reductase, mitochondrial; Auxiliary enzyme of beta-oxidation. It participates in the metabolism of unsaturate...*		X	X		X	X	X	X	0.041
HNRNPL	*Heterogeneous nuclear ribonucleoprotein L; Splicing factor binding to exonic or intronic sites and acting as either an activato...*		X	X		X	X	X	X	0.041
CEP41	*Centrosomal protein of 41 kDa; Required during ciliogenesis for tubulin glutamylation in cilium. Probably acts by participating...*		X	X		X	X	X	X	0.041
PSMA2	*Proteasome subunit alpha type-2; Component of the 20S core proteasome complex involved in the proteolytic degradation of...*		X	X		X	X	X	X	0.041
YKT6	*Synaptobrevin homolog YKT6; Vesicular soluble NSF attachment protein receptor (v-SNARE) mediating vesicle docking and f...*		X	X		X	X	X	X	0.041
PPIF	*Peptidyl-prolyl cis-trans isomerase F, mitochondrial; PPIase that catalyzes the cis-trans isomerization of proline imidic peptid...*		X	X		X	X	X	X	0.041

Figure S7 C11orf54 associations in String

Output of a String (https://string-db.org/) database search using C11orf54 as a query showing the 10 highest score interactors according to all available (**a**) or co-occurrence only (**b**) evidence sources.

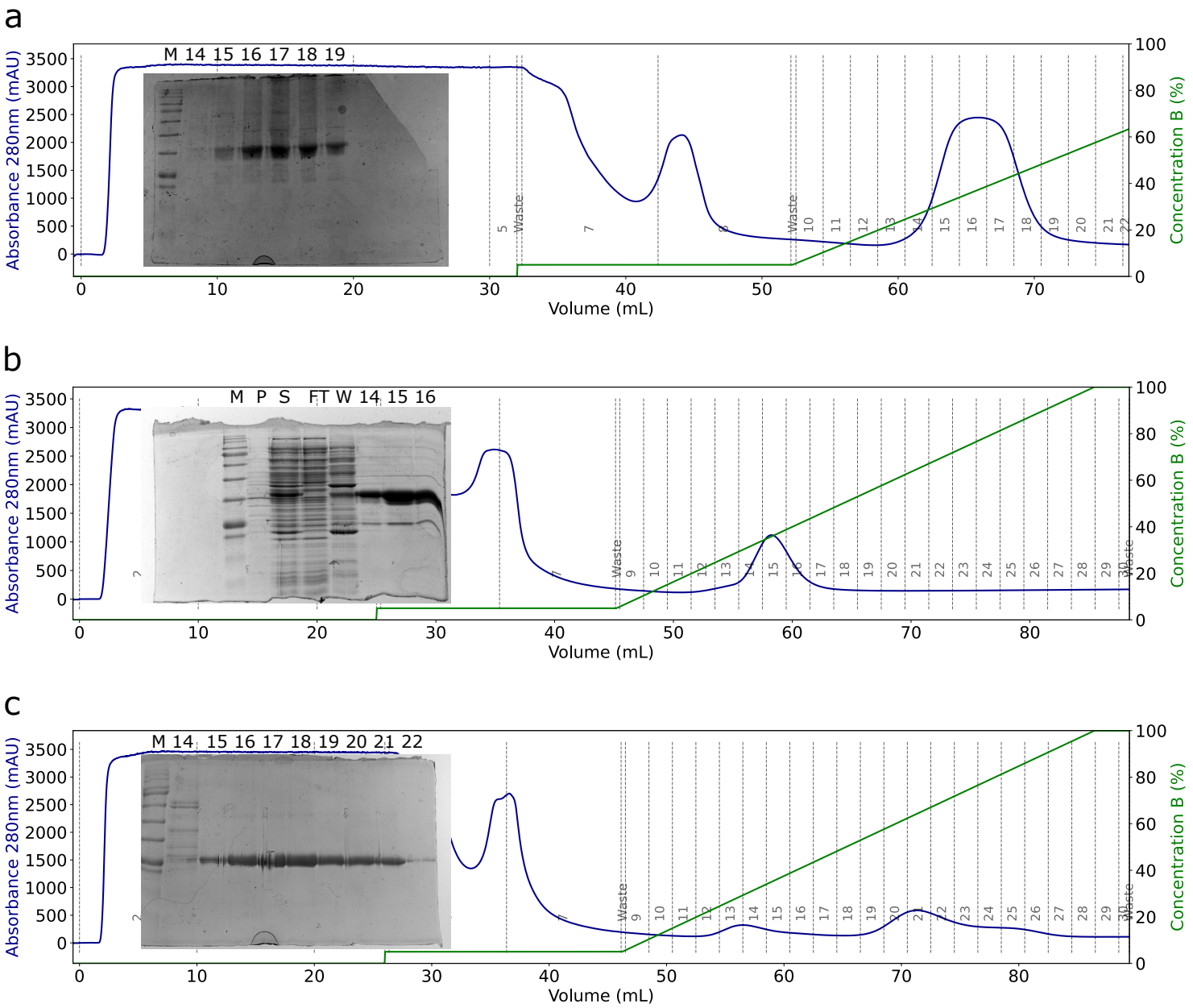


Figure S8: Expression and purification of recombinant C11orf54, CRYL1, and DCXR

Chromatogram of affinity chromatography for CRYL1 (**a**), C11orf54 (**b**), and DCXR (**c**) using a 5 mL HisTrap FF column on an ÄKTA pure system. Insets display SDS-PAGE gels of the fractions indicated at the top of each gel.

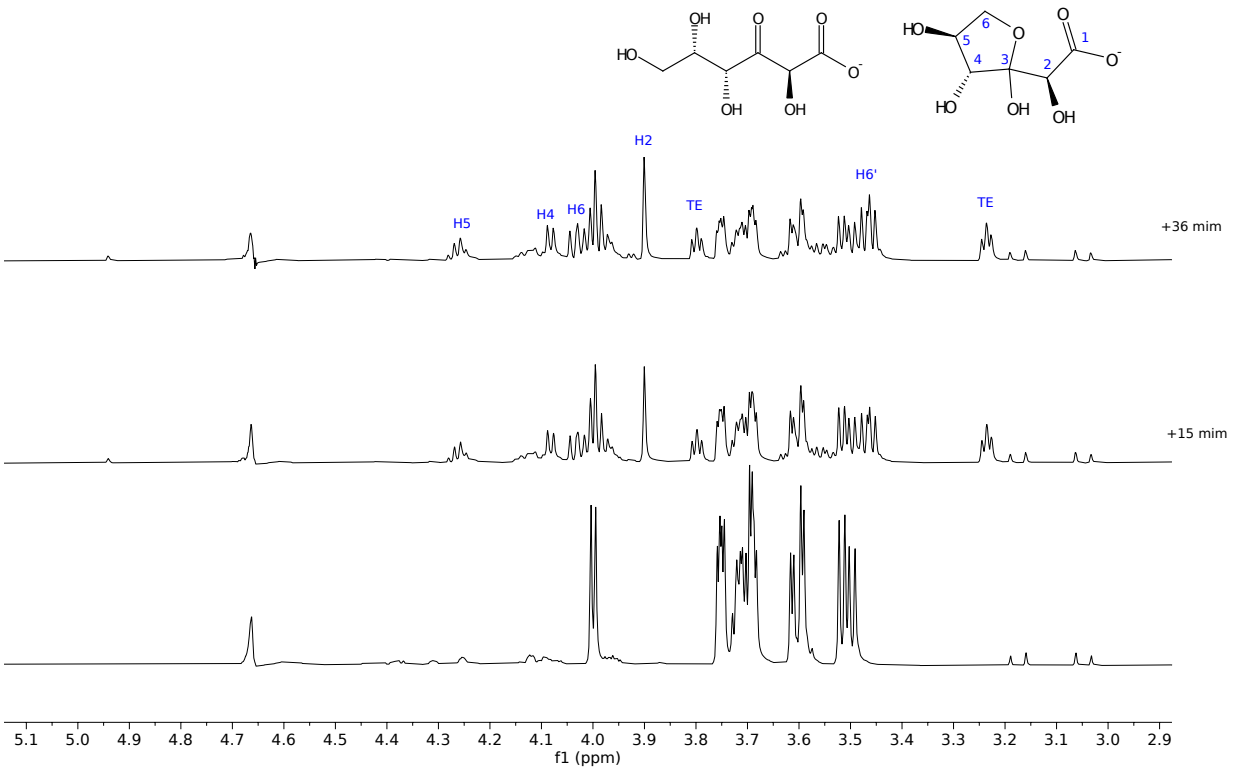


Figure S9: BKG proton assignment in the presence of excess CRYL1

Partial ^{1}H NMR spectra of the conversion of L-gulonate (5 mM) into 3-dehydro-L-gulonate (BKG) catalyzed by an excess (40 uM) of CRYL1. Signals corresponding to the enzyme buffer component Triethanolamine (TE) are indicated.

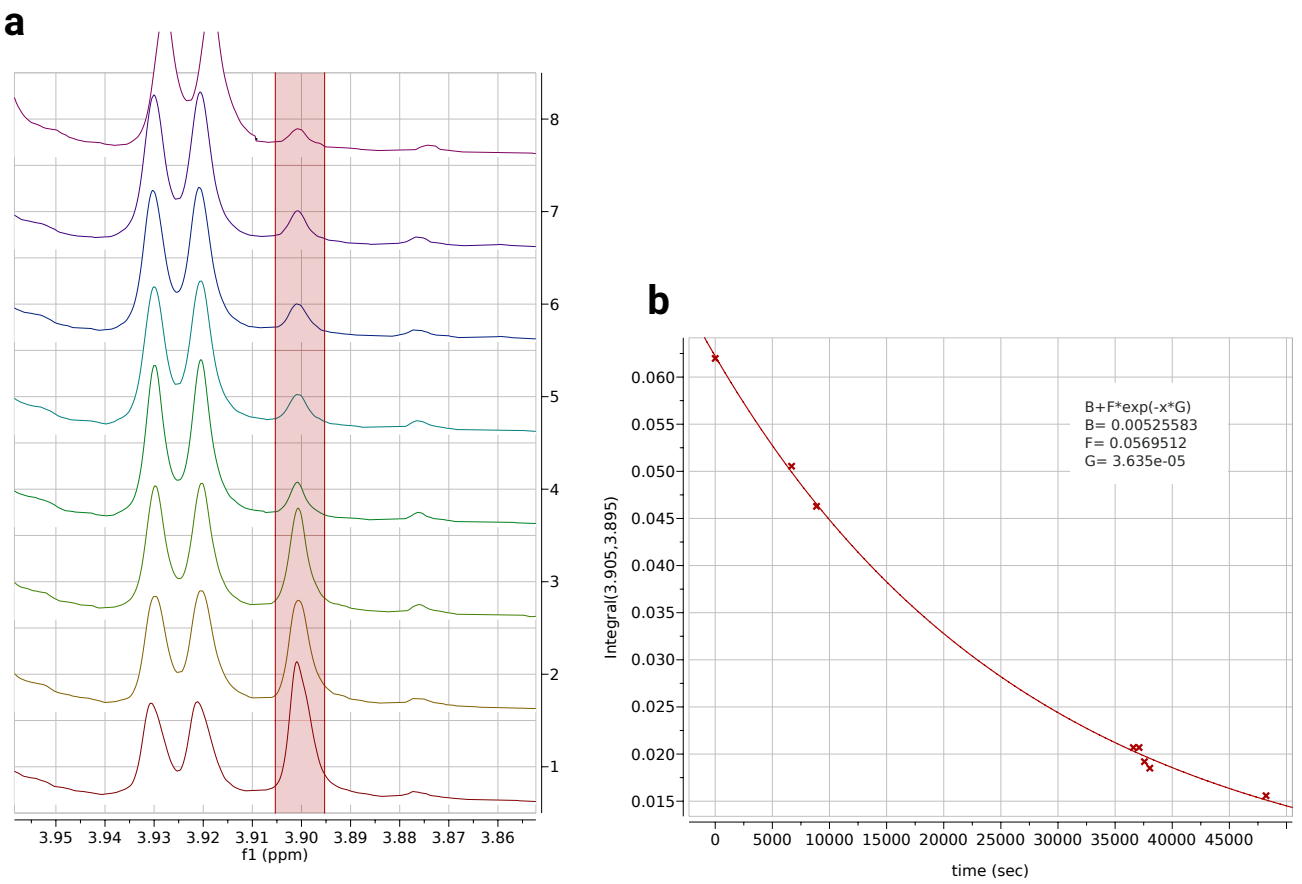
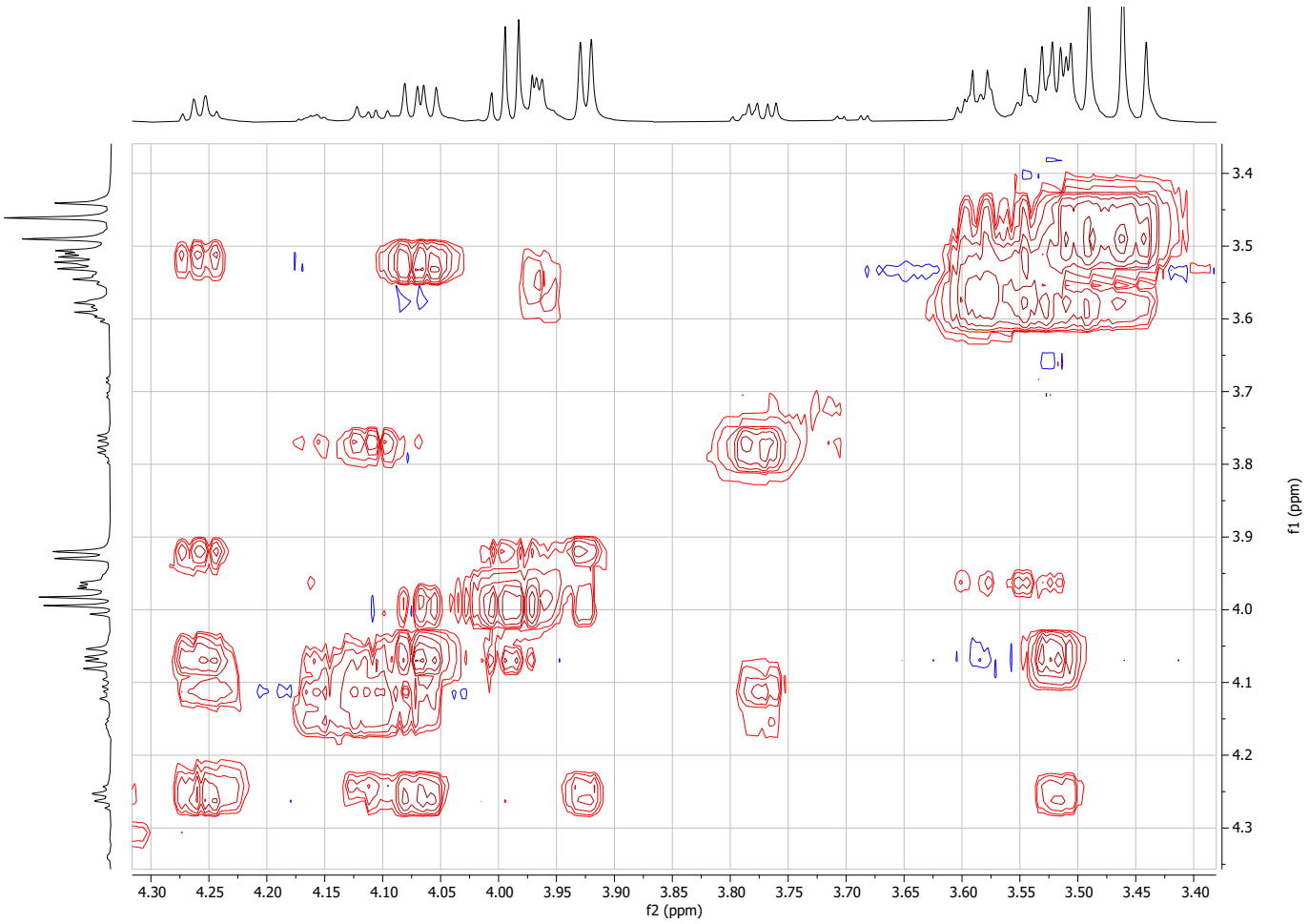


Figure S10: Non-enzymatic decay of β-keto-L-gulonate (BKG)

Proton NMR spectra of the CRYL1 reaction product were recorded at various time points, starting 5 hours after the addition of CRYL1 (40 uM) to L-gulonate (5mM). The non-enzymatic decay of beta-keto-gulonate (BKG) was monitored by measuring the integral of the singlet signal at 3.90 ppm. The decay was fitted to an exponential decay model (red curve), obtaining a decay constant (G, inset) of 3.6 x 10^{-5} s^{-1} ($t_{½}$ ≈ 5.3 h).



Extended Data Fig. 11: Correlation spectroscopy (COSY) of the C11orf54 reaction product

The 2D ^{1}H-^{1}H NMR Correlation Spectroscopy map (400 MHz, 10% D_2O, 25 °C) of the product obtained from L-gulonate (5 mM) with the addition of CRYL1 (5 µM) C11orf54 (5 µM).

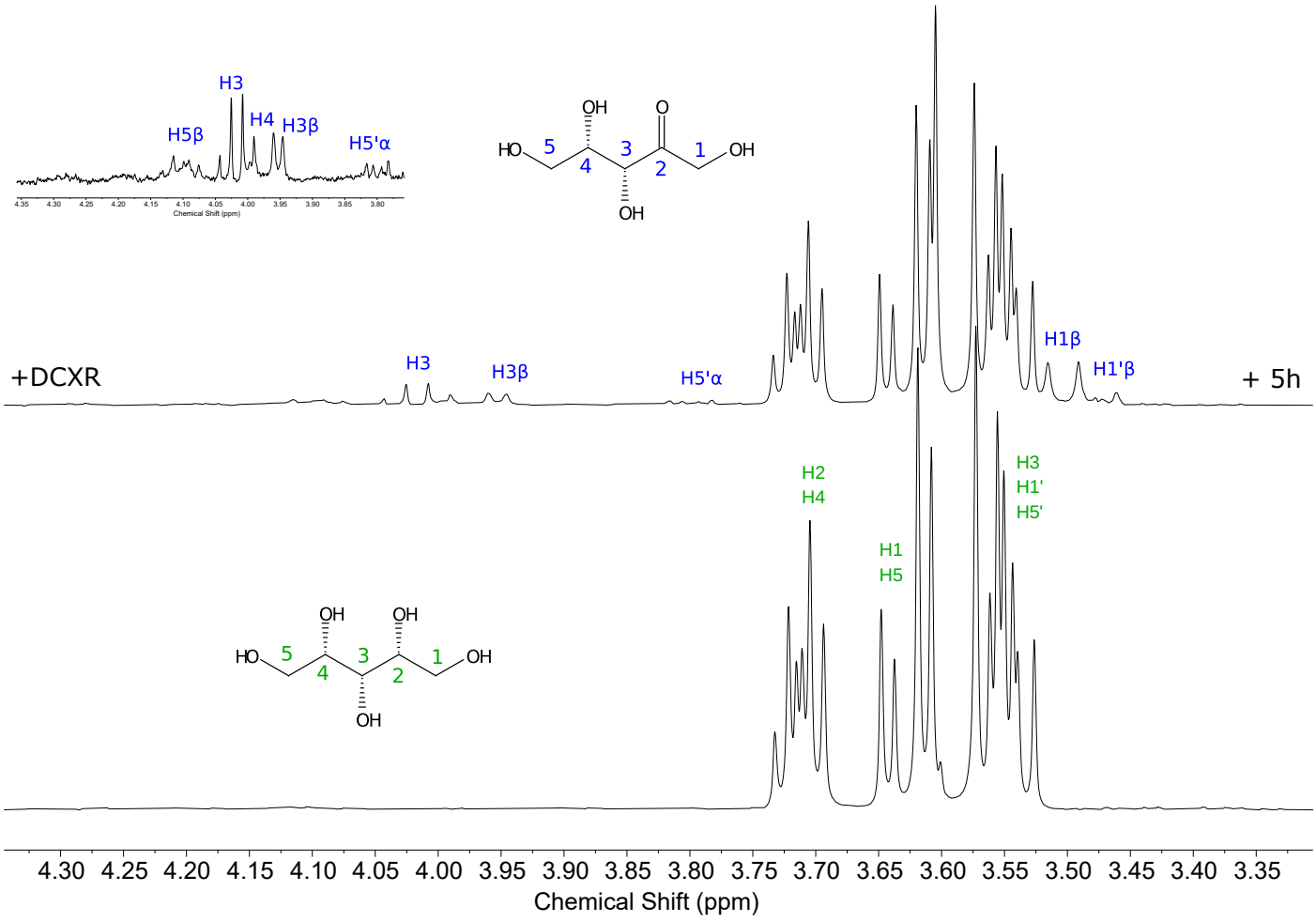
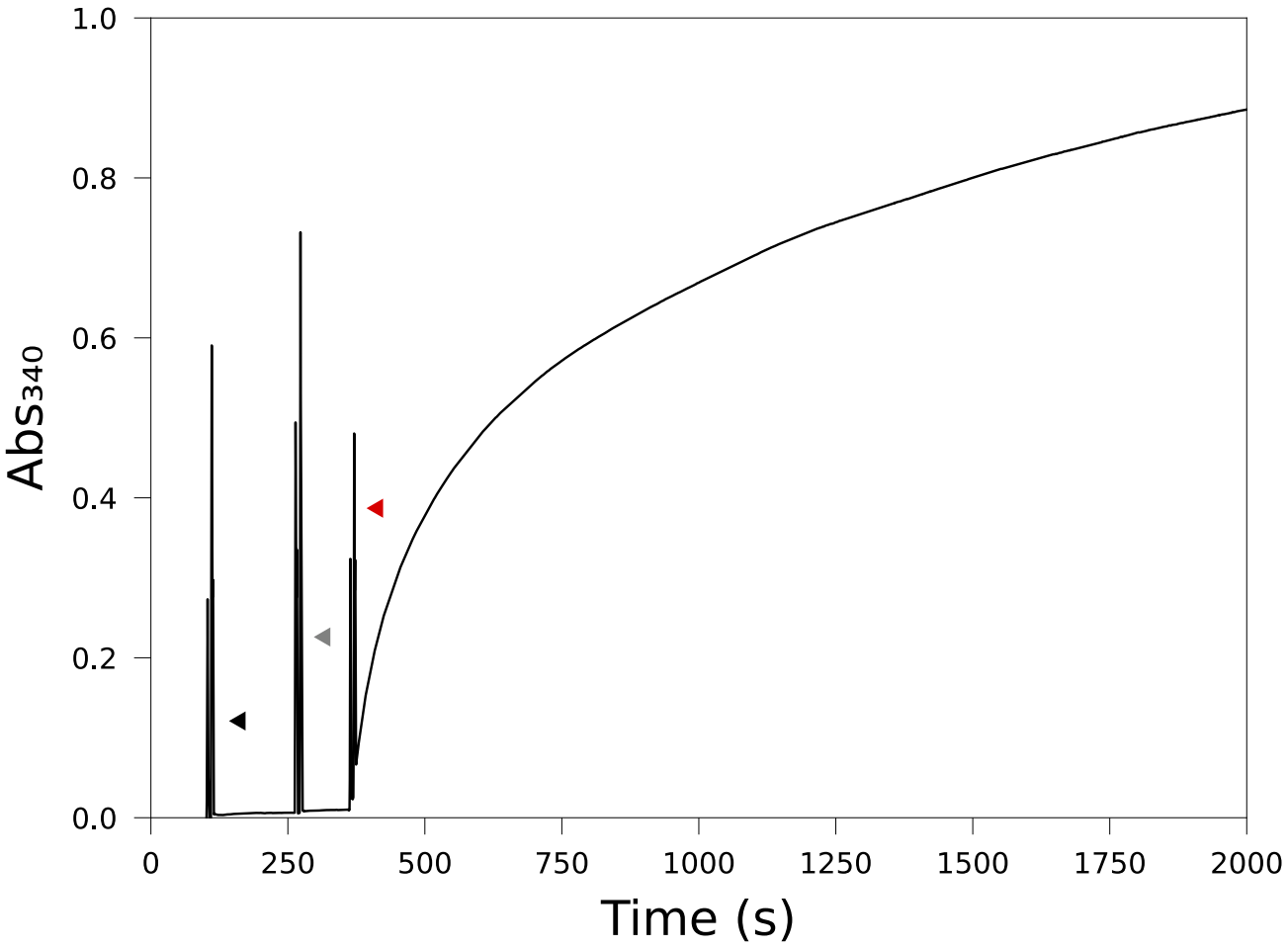


Figure S12: DCXR-catalyzed reverse reaction supports a functional link with C11orf54

Time-resolved ^{1}H NMR spectra showing the conversion of xylitol into L-xylulose by DCXR. The proton assignment of xylitol (5 mM) is shown at time 0' in a reaction buffer (20 mM sodium phosphate, 10% D_2O) containing $NADP^+$ (0.5 mM), pyruvate (20 mM) and lactate dehydrogenase (LDH) for $NADP^+$ recycling. Upon addition of DCXR (5 µM), the production of both the open-chain and furanose forms of L-xylulose becomes evident, as supported by proton assignment.



Supplementary Fig. 13: Coupled assay confirms NADH but not NADPH production via C11orf54

Time-course trace of reaction with L-gulonate (1 mM) and NADP+ (0.4 mM) after the addition of CRYL1 (black arrow), C11orf54 (grey arrow) and NAD+ (red arrow) monitored via spectrophotometry by tracking NADPH/NADH production at 340 nm.

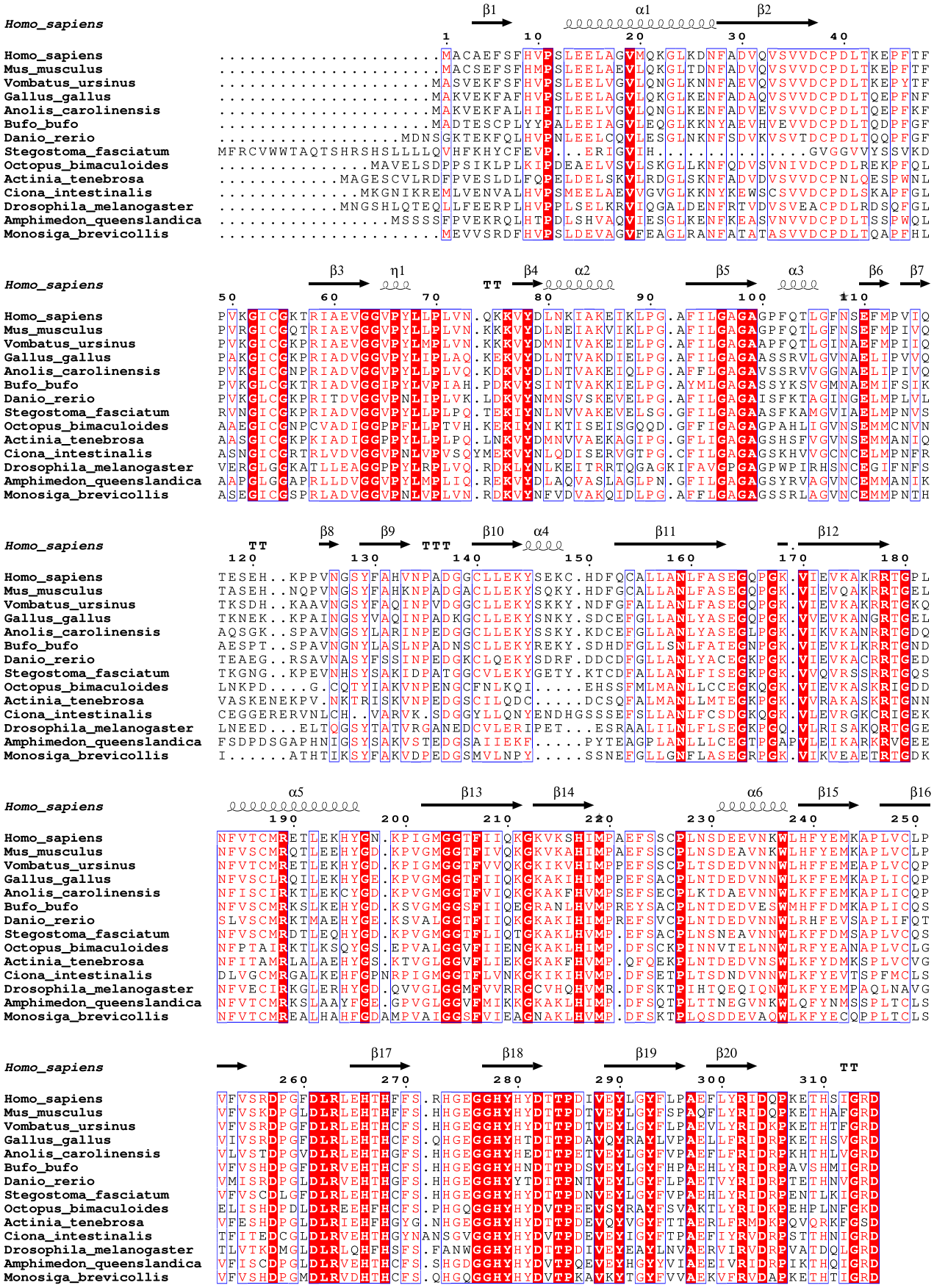
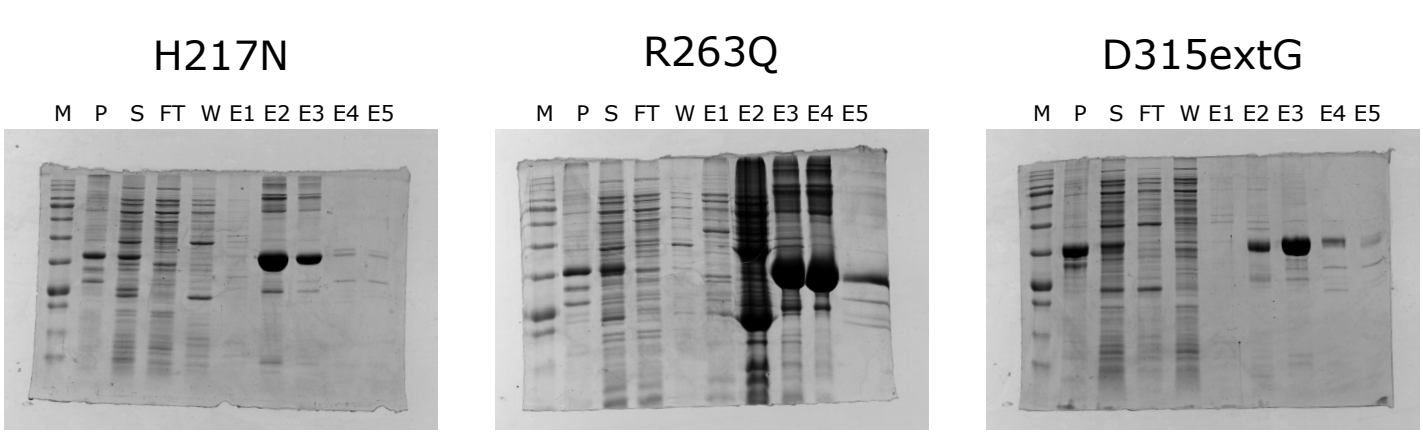


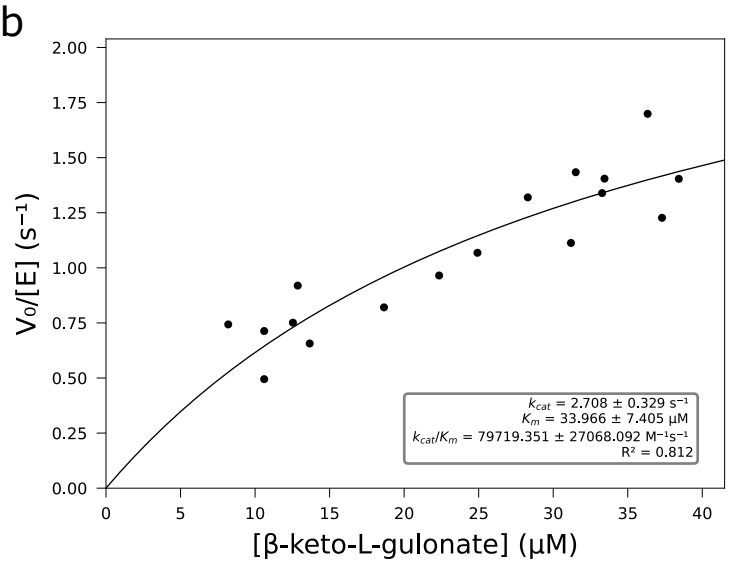
Figure S14 Multiple alignment of BKGD homologs in various species

Multiple sequence alignment of orthologous C11orf54 sequences from OrthoDB v11, with secondary structures based on the human PDB structure (1XCR).

a

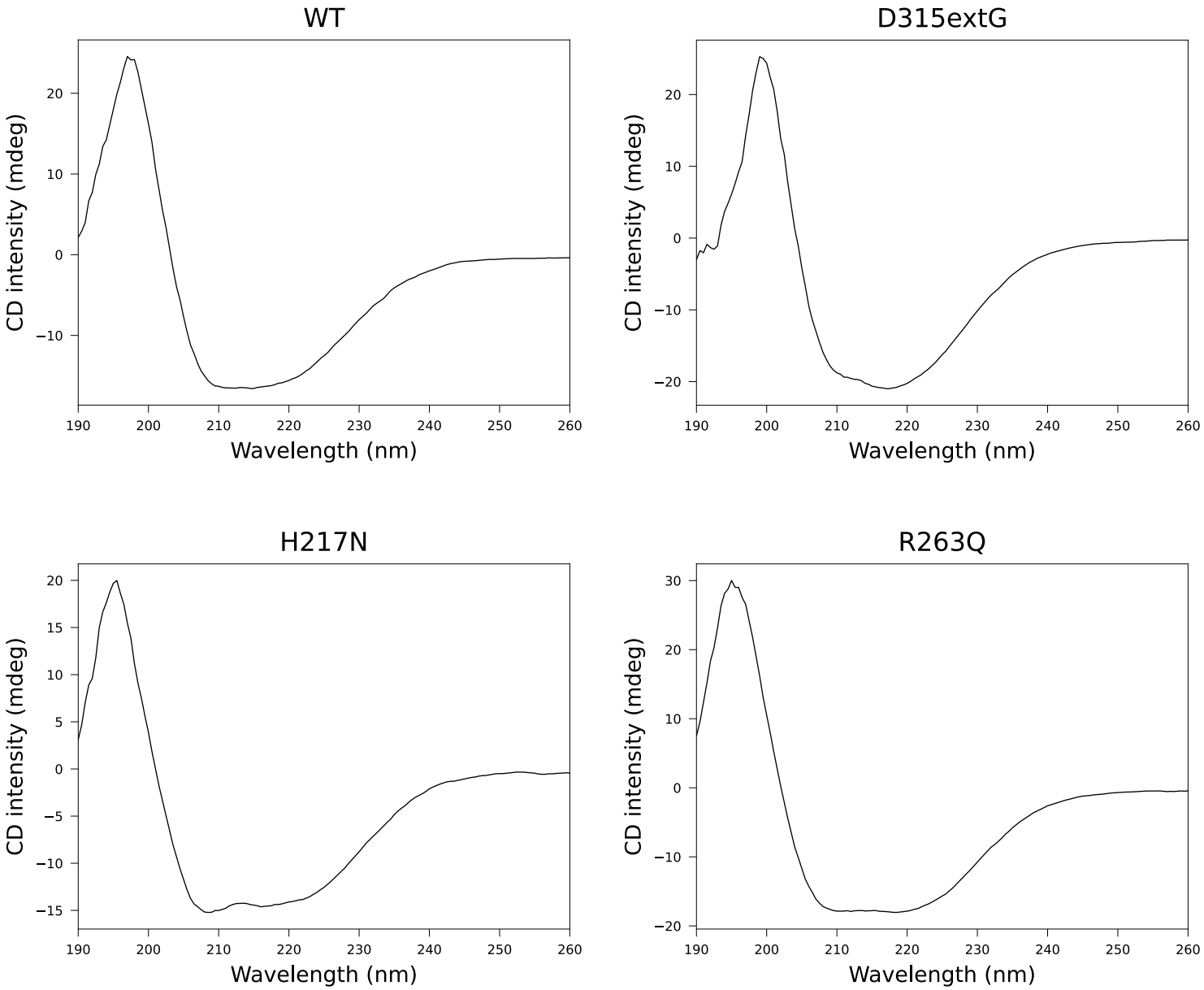


b



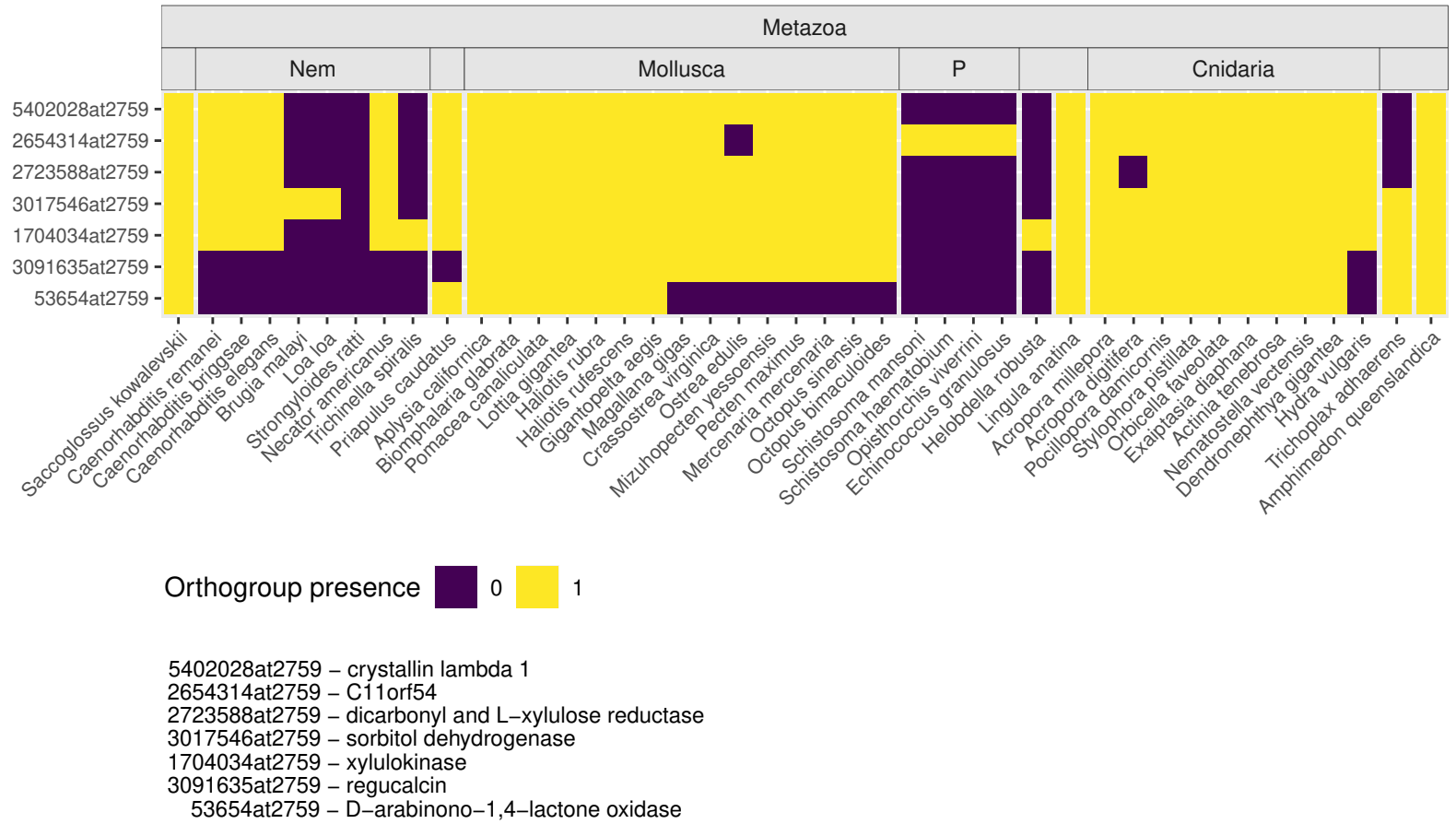
Supplementary Fig. 15 Expression, purification and characterization of C11orf54 variants

a, SDS-PAGE gels of the C11orf54 variants corresponding to the purification fractions (P: pellet; S: supernatant; FT: flowthrough; W: washing; E: elution). **b**, Nonlinear fitting of the initial reaction velocity of the D315extG variant of C11orf54 (0.25 μM) as a function of β-keto-L-gulonate concentration, based on the Michaelis-Menten equation. Data points represent independent experimental replicates.



Supplementary Fig. 16 CD spectroscopy of C11orf54 variants.

CD spectra of C11orf54 and its variants recorded with 1 µM of protein in 20 mM Tris-HCl pH 7.4. The values are means of 3 acquisitions for each spectrum.



5402028at2759 – crystallin lambda 1
2654314at2759 – C11orf54
2723588at2759 – dicarbonyl and L–xylulose reductase
3017546at2759 – sorbitol dehydrogenase
1704034at2759 – xylulokinase
3091635at2759 – regucalcin
53654at2759 – D–arabinono–1,4–lactone oxidase

Figure S17 Detail of the distribution of pentose and ascorbate pathway genes in basal metazoans.

The absence and presence of genes involved in the two alternative pathways is shown in metazoans belonging to Hemichordata (*Saccoglossus kowalevskii*), Nematoda (Nem), Priapulida (*Priapulus caudatus*), Mollusca, Platyhelminthes (P), Annelida (*Helobdella robusta*), Brachiopoda (*Lingula anatina*), Cnidaria, Placozoa (*Trichoplax adhaerens*), Porifera (*Amphimedon queenslandica*). Gene orthogroup classification is according to OrthoDB V.11.

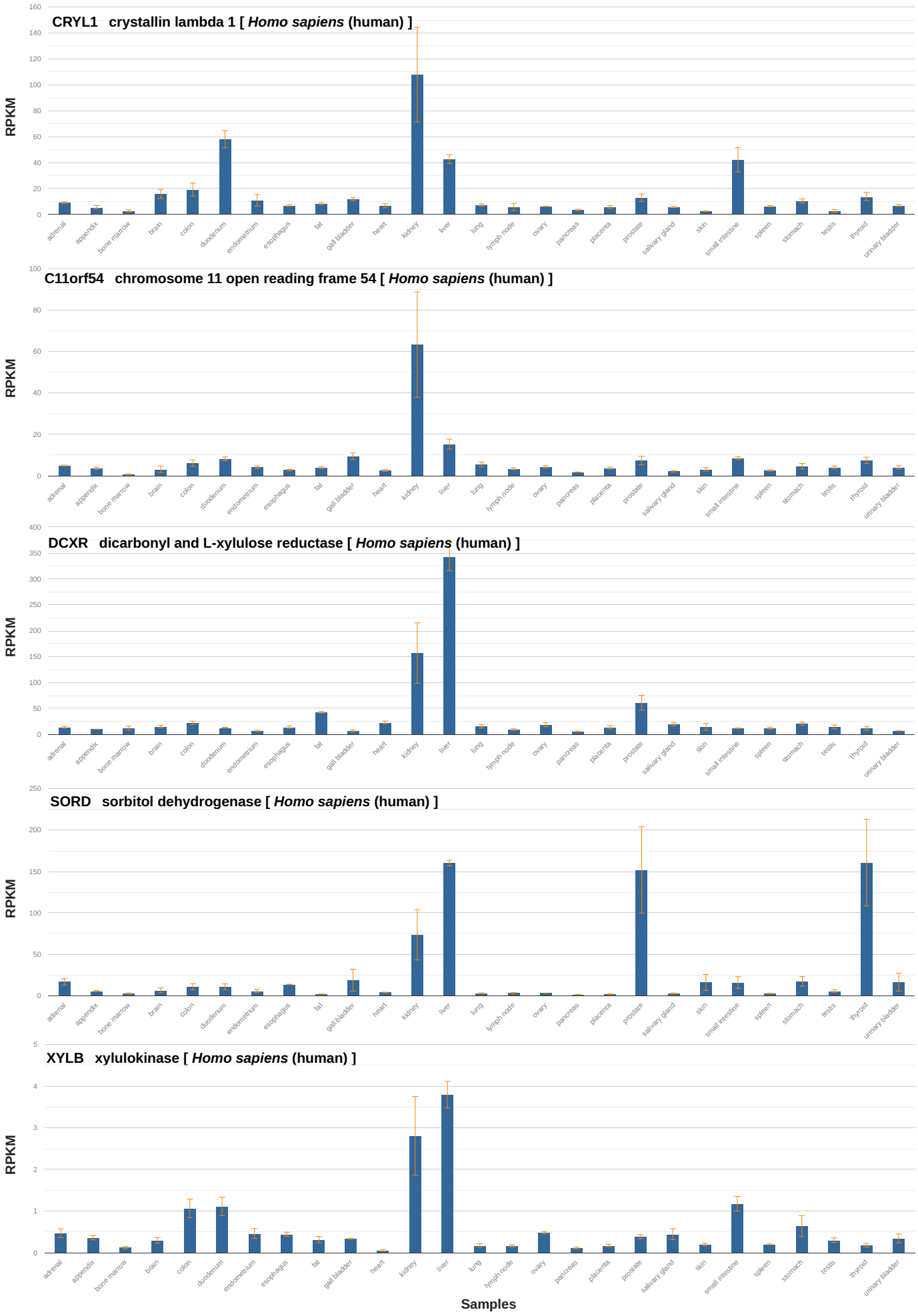


Figure S18 Expression of PP genes in human

Gene expression is according to data provided by the BioProject PRJEB4337 (HPA RNA-seq normal tissues). RPKM: reads per kilobase of exon per million mapped reads.