

# Duplications and Functional Convergence of Intestinal Carbohydrate-Digesting Enzymes

Antonio Brun,<sup>†,1,2</sup> Daniel Mendez-Aranda,<sup>†,3</sup> Melisa E. Magallanes,<sup>†,2</sup> William H. Karasov,<sup>1</sup> Carlos Martínez del Río,<sup>4</sup> Maude W. Baldwin,<sup>\*,†,3</sup> and Enrique Caviedes-Vidal<sup>\*,†,2,5</sup>

<sup>1</sup>Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI

<sup>2</sup>Instituto Multidisciplinario de Investigaciones Biológicas de San Luis, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

<sup>3</sup>Max Planck Institute for Ornithology, Seewiesen, Germany

<sup>4</sup>Department of Zoology and Physiology, University of Wyoming, Laramie, WY

<sup>5</sup>Departamento de Bioquímica y Ciencias Biológicas, Universidad Nacional de San Luis, San Luis, Argentina

<sup>†</sup>These authors contributed equally to this work as the first authors.

<sup>\*</sup>These authors contributed equally to this work as the last authors.

**\*Corresponding authors:** E-mails: mbaldwin@orn.mpg.de; enrique.caviedes@gmail.com.

**Associate editor:** Meredith Yeager

## Abstract

**Vertebrate diets and digestive physiologies vary tremendously. Although the contribution of ecological and behavioral features to such diversity is well documented, the roles and identities of individual intestinal enzymes shaping digestive traits remain largely unexplored. Here, we show that the sucrase-isomaltase (SI)/maltase-glucoamylase (MGAM) dual enzyme system long assumed to be the conserved disaccharide and starch digestion framework in all vertebrates is absent in many lineages. Our analyses indicate that independent duplications of an ancestral SI gave rise to the mammalian-specific MGAM, as well as to other duplicates in fish and birds. Strikingly, the duplicated avian enzyme exhibits similar activities to MGAM, revealing an unexpected case of functional convergence. Our results highlight digestive enzyme variation as a key uncharacterized component of dietary diversity in vertebrates.**

**Key words:** evolution, physiology, functional convergence, gene duplications, carbohydrate digestion, sucrase..

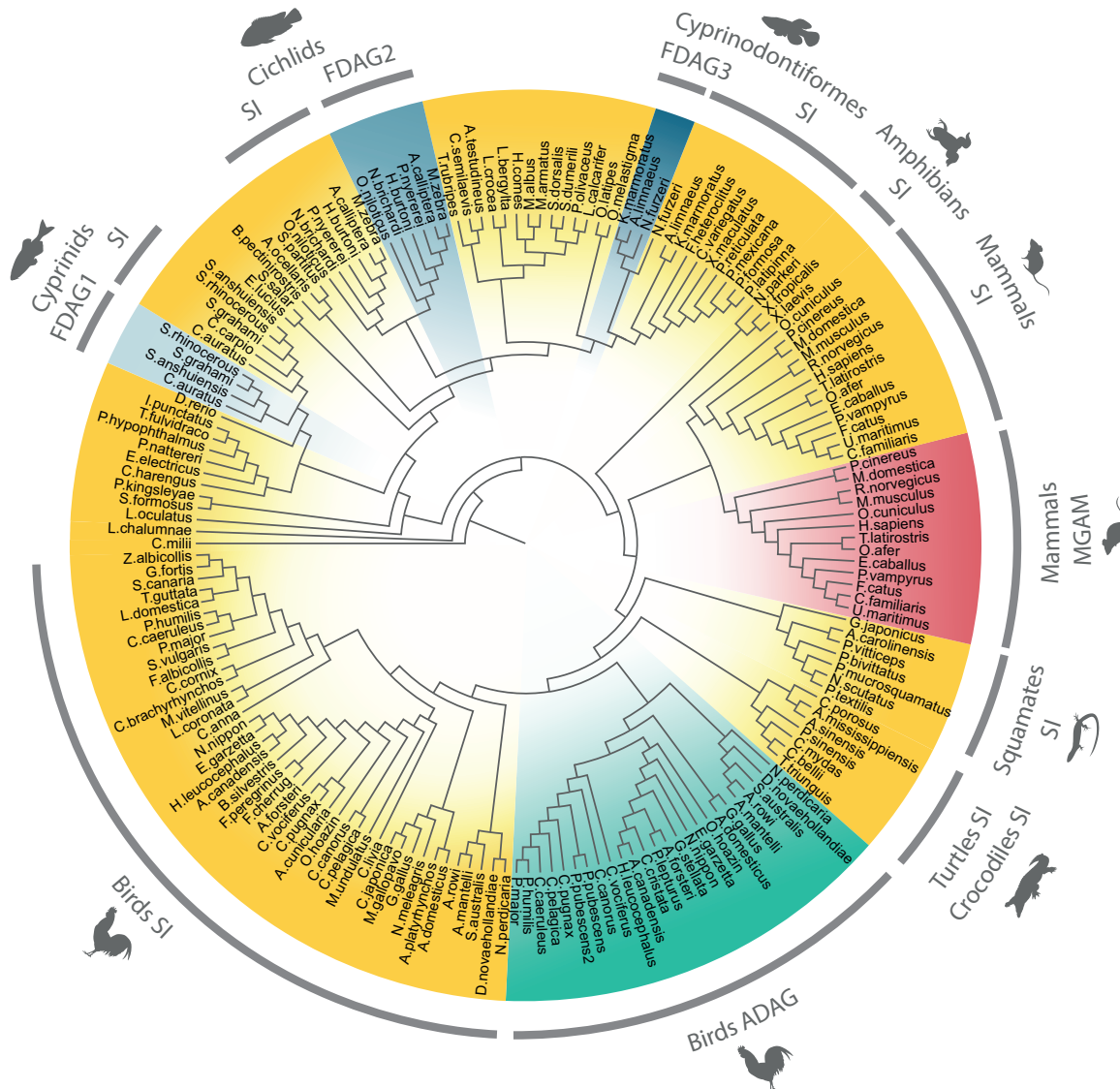
## Introduction

The efficient absorption of nutrients requires that complex molecules are broken down into simpler biochemical components. This process relies on the activity of specific enzymes in the intestinal tract, and the differential catalytic activities of these enzymes likely impact the evolution of dietary diversity (Jiao et al. 2019) and the roles that animals play in their ecosystems (Karasov et al. 2011). In vertebrates, the intestinal membrane-bound  $\alpha$ -glucosidases (AGs) are the primary enzymes involved in carbohydrate digestion, enabling the uptake of glucose and fructose by breaking the bonds in dietary disaccharides, as well as in the oligosaccharides produced by amylase breakdown of starch and glycogen (Swallow 2003). These enzymes are essential for animals with diets rich in simple sugars and in starches, such as nectarivores, frugivores, and granivores (Karasov et al. 2011). However, as most studies in non-model systems have measured hydrolytic activity using intestinal homogenates (McWhorter et al. 2009; Karasov et al. 2011), the identities and functions of specific enzymes in other species are currently unknown.

Hypotheses regarding the evolution of AGs in vertebrates have been shaped by work on mammals. In primates, rodents, and most other mammals that have been investigated (Pontremoli et al. 2015; Nichols et al. 2017), two enzymes

are expressed in the intestinal “brush-border” membrane. Each enzyme has two similar subunits each with differing but overlapping catalytic activities (Nichols et al. 2017). The active sites in maltase-glucoamylase (MGAM) hydrolyze the  $\alpha$ -1,4 glycosidic bonds in maltose and in the small linear chains of glucose (amylose) that result from amylase’s action on starch. Sucrase-isomaltase (SI) has a C-terminal subunit that hydrolyses the  $\alpha$ -1,2 bond between glucose and fructose in sucrose, and an N-terminal subunit that breaks the  $\alpha$ -1,6 bond of isomaltose (Chaudet et al. 2019). Both subunits of sucrase-isomaltase have maltase activity. In mammals, the overlapping, but not identical, substrate specificities, kinetics, and expression patterns (Jones et al. 2011) of SI and MGAM expand the range of substrates hydrolyzed (Nichols et al. 2003).

In many birds, including chickens, ostriches, and pigeons, two AGs have also been described (Prakash et al. 1983; Hu et al. 1987; Oosthuizen et al. 1998). The fact that these avian enzymes exhibit properties similar to the better-known mammalian enzymes led to the hypothesis that the duplication of an ancestral  $\alpha$ -glucosidase gene happened before the split of mammals and birds (Hu et al. 1987), an assumption that has influenced many later studies (Karasov 2011; Karasov and Douglas 2013; Tamaoki et al. 2016). However, our phylogenetic analysis of 175 annotated AGs in the genomes of 129



**Fig. 1.** A single ancestral enzyme duplicated independently in vertebrates. ML gene tree of  $\alpha$ -glucosidases annotated from 129 vertebrate genomes (see [supplementary fig. S1, Supplementary Material](#) online and [supplementary table S1, Supplementary Material](#) online). A single ancestral *SI* (yellow) present in all jawed vertebrates has duplicated independently in mammals (red) and birds (green), as well as three times in fish (blue).

vertebrates ([fig. 1](#) and [supplementary table S1, Supplementary Material](#) online) disproves the assumption of an ancient two-enzyme system and reveals several unexpected findings.

**Results**

Phylogenetic analysis indicates that surprisingly, there appears to be a single gene encoding one  $\alpha$ -glucosidase in all jawed vertebrates ([fig. 1](#) and [supplementary fig. S1, Supplementary Material](#) online), orthologous to the mammalian sucrase-isomaltase gene (*SI*). Although the function of this single early glucosidase is not known, previous studies have hypothesized that the ancestral AG enzyme was a single-subunit protein that hydrolyzed isomaltose and maltose (Hu et al. 1987), which was duplicated in tandem early in vertebrate history and that later mutations in the active site of the C-terminal subunit resulted in sucrase activity (Naumoff 2007). Although

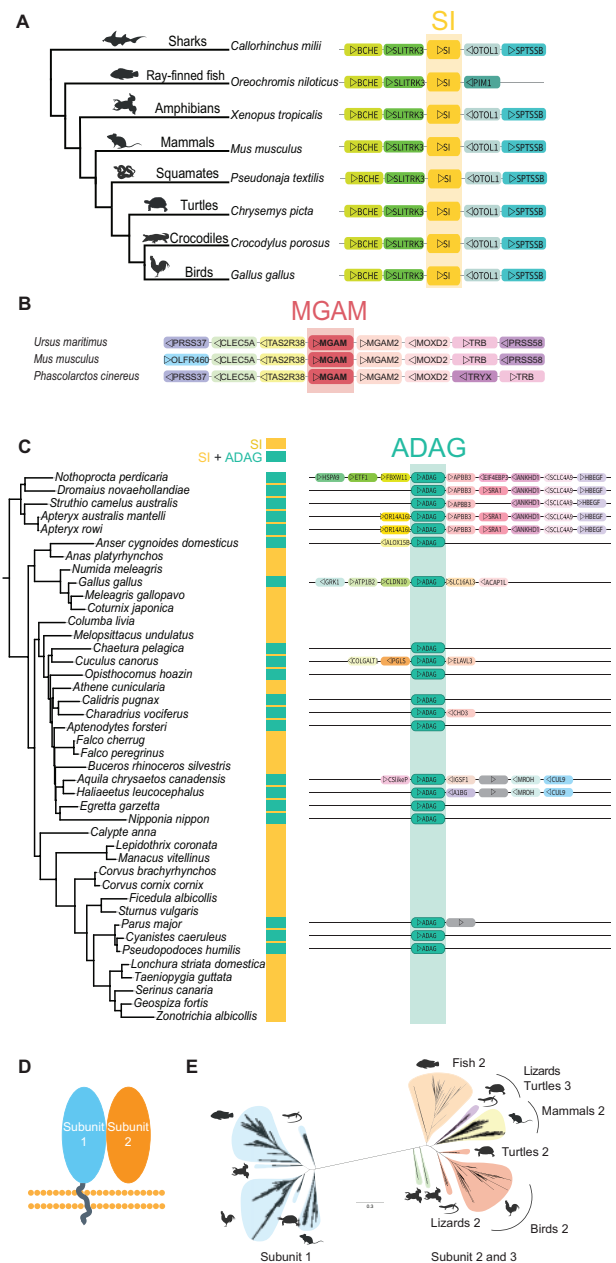
lamprey *SI* predictions are absent, examination of contigs from the sea lamprey genome containing vertebrate flanking loci revealed an  $\alpha$ -glucosidase sequence. Additionally, tunicates and amphioxus as well as other invertebrates have predictions from the same larger GH31 protein family that are annotated as *SI* or *MGAM*. In invertebrates, the syntenic genes flanking vertebrate *SI* are absent, and although many predictions are shorter or appear to be composed of a single subunit, others are long with multiple active sites. Therefore, *SI* appears to be present in the ancestor of jawless vertebrates and may have an earlier origin.

In addition, we document multiple independent duplications of this ancestral vertebrate enzyme: three in different teleost fish families, one in the ancestor of birds, and one early in the evolutionary history of mammals ([fig. 1](#)). In each case, the duplication results in two daughter copies (co-orthologs); here, we retain the term *SI* for the copy that remains in the ancestral genomic location. The mammalian duplicate

(MGAM) is present in all mammalian genomes surveyed and appears to have subsequently duplicated in many mammals giving rise to additional MGAM copies (not shown on the phylogeny), although the functionality of this second duplicate has not been assessed. In birds and fish, the gene predictions of the *SI* duplicate are also called MGAM in genomic databases, but as they are not true orthologs, we here refer to them as avian-derived  $\alpha$ -glucosidase (ADAG) and fish-derived  $\alpha$ -glucosidase (*FDAG1-3*) for clarity. Fish duplicates are found in only three lineages and do not appear to be the result of the whole-genome duplication in teleosts. *FDAG1* is present in the Cyprinidae (perhaps due to the whole genome duplication in this group; Li et al. 2015); *FDAG2* is found in all cichlid genomes examined, and *FDAG3* is present in the Aplocheilidae family of the Cyprinodontiformes (fig. 1, supplementary figs. S1 and S2, Supplementary Material online). In birds, ADAG appears to be the result of an early duplication: ADAG forms a monophyletic clade that includes members of early branching lineages such as the paleognaths (including ostriches, emus, and tinamous), as well as representatives across the entire bird phylogeny (fig. 1).

Unlike *SI* and MGAM (fig. 2A and B), ADAG is not always flanked by members of the same block of syntenic genes (fig. 2C). In birds, ADAG is present on very small contigs, and we found shared syntenic blocks only in two cases of closely related groups: the paleognaths (tinamous, *Nothoprocta perdicaria*; ostrich, *Struthio camelus*; emu, *Dromaius novaehollandiae*; and kiwis, *Apteryx rowi* and *A. australis*) and two species of eagles (*Haliaeetus leucocephalus* and *Aquila chrysaetos*). Notably, ADAG appears to be absent from the genomes of many lineages including, surprisingly, all songbirds, except the three members of the titmouse family (Paridae: *Cyanistes caeruleus*, *Parus major*, and *Pseudopodoces humilis*; fig. 1). This puzzling distribution, together with the short contig size and lack of information regarding syntenic genes, suggests that ADAG may lie in a region of the genome difficult to assemble. However, as we demonstrate below, it appears that in some songbird species, absence of ADAG may reflect an actual gene loss. The observation that ADAG is absent from the most recent high-quality assembly of the zebra finch (*Taeniopygia guttata*) genome (accession number PRJNA510143), even though the genes that flank this gene in paleognaths are present, is consistent with this hypothesis (supplementary fig. S3, Supplementary Material online).

Examination of *SI* of nonavian reptiles revealed another unexpected result. In the genomes of amphibians, as well as of lizards and turtles, *SI* appears to have three subunits (caused by duplications of domains in *SI*), rather than the two found in other vertebrates. The first subunit domain is orthologous across all vertebrates, but the second subunit duplicated in multiple lineages. Amphibians appear to have duplicated this subunit domain independently; another duplication occurred either in the ancestor of reptiles (followed by subsequent loss in archosaurs) or in the amniote ancestor (followed by loss in both archosaurs and mammals; fig. 2D–E and supplementary fig. S4A, Supplementary Material online). The second subunit domain of birds appears to be



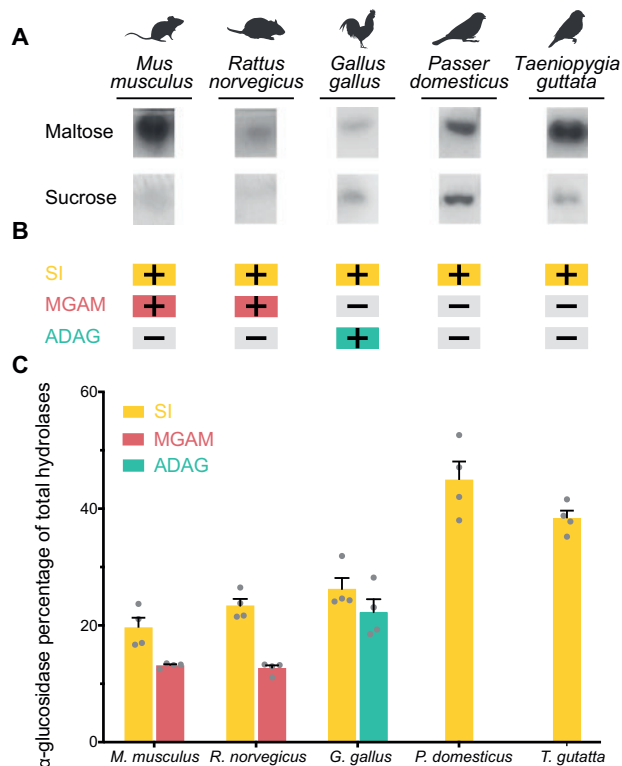
**Fig. 2.** Differences in conserved synteny and subunit domain composition of amniote  $\alpha$ -glucosidases. (A–C) Syntenic region surrounding *SI*, *MGAM*, and *ADAG* in mammals and birds; (D–E) phylogenetic relationships of  $\alpha$ -glucosidase subunits in vertebrates. The flanking genes surrounding *SI* are conserved across all vertebrates (A), and whereas *MGAM* also exhibits conserved synteny (B), *ADAG* is missing from many genomes and has shared synteny only in the paleognaths (in an emu, ostrich, tinamou, and in kiwis) and in eagles (C). Tree in (C) is an extract of figure 1, showing species where both enzymes co-occur; additionally, *ADAG* (but not *SI*) was found in woodpeckers (where it duplicated again), as well in the tropicbird, loon, and in a penguin (supplementary fig. S1, Supplementary Material online). (D) Schematic of the two subunit domains of the mammalian *SI*: both subunit domains are GH31 protein family members and possess an active site, but only the first subunit domain is associated with the membrane. (E) Phylogenetic analysis of individual subunit domains. Subunit 1 (blue) is orthologous across vertebrates, but the second subunit underwent a separate duplication in amphibians (green), and again either in the ancestor of reptiles followed by subsequent loss in

orthologous to the middle subunit of lizards and turtles (supplementary fig. S4B, Supplementary Material online). In addition, duplications are also seen in the second subunit of mammalian MGAM (supplementary fig. S4A, Supplementary Material online).

Given the lack of orthology both between the bird and the mammalian duplicated enzymes (MGAM and ADAG) and the surprising apparent absence of ADAG from many bird genomes, the functionality and activity of the avian enzymes were unclear. To address this, we used proteomics as well as a cell-based assay to functionally characterize avian SI and ADAG. To verify if the enzymes observed in the genomes were expressed, we used a proteomic method to detect the presence and relative abundance of the combined AGs in the intestinal brush border of three bird species (chicken, house sparrow, and zebra finch) as well as of mouse and rat. We isolated the intestinal brush-border membranes of the intestinal villi (Mac Donal et al. 2008) and proteins from this extract were run on a gel and tested for hydrolytic activity against sucrose or maltose (Finlayson et al. 1990); bands showing activity were subsequently extracted and analyzed by nanoscale liquid chromatography (fig. 3A and B and supplementary fig. S5, Supplementary Material online) coupled to tandem mass spectrometry (nano-LC-MS/MS; McConnell et al. 2011).

Consistent with the genomic data, the species that were expected to have SI, as well as either MGAM or ADAG (rat, mouse and chicken), expressed two enzymes (fig. 3B and C), whereas the two songbird species examined (zebra finch and house sparrow), expressed only SI. As has been shown in other studies (Gericke et al. 2016), AGs are highly expressed: among the five vertebrate species we assayed, SI was the most highly expressed (20–45%) of all hydrolases detected in the brush-border membrane (fig. 3C and supplementary table S3, Supplementary Material online) (McConnell et al. 2011; Gericke et al. 2016). In species with either MGAM or ADAG, these duplicated enzymes also represent a very large fraction (13–22%) of the membrane-bound hydrolases as well (fig. 3C and supplementary table S3, Supplementary Material online).

Because our proteomic analyses do not distinguish between the activity of the bird SI and ADAG, we used a cell-based assay (Rodríguez et al. 2013) to study each enzyme in isolation. We amplified chicken SI and ADAG from chicken intestinal cDNA, generated expression vectors, and transfected COS-7 cell cultures with plasmids containing either of the two bird enzymes or synthesized human SI as a control (Rodríguez et al. 2013). Transfected cells were then incubated with either maltose or sucrose and the activity against these substrates was measured. Cells transfected with either chicken or human SI hydrolyzed both maltose and sucrose (fig. 4). In contrast, chicken ADAG hydrolyzed maltose, but



**FIG. 3.** SI, MGAM, and ADAG are expressed and active against maltose and sucrose. Extracted and solubilized native proteins from the intestinal brush-border membrane of five species ( $n = 2$  individuals per species) were run on nondenaturing gels; activity against maltose and sucrose is shown as dark bands (A). Extracted bands were analyzed by mass spectrometry revealing presence (+) or absence (-) of SI, MGAM, and ADAG (B). Abundance (spectral counts) of SI (yellow), MGAM (red), and ADAG (green) in relation to total brush-border membrane hydrolases (mean  $\pm$  SE,  $n = 4$  individuals per species) (C).

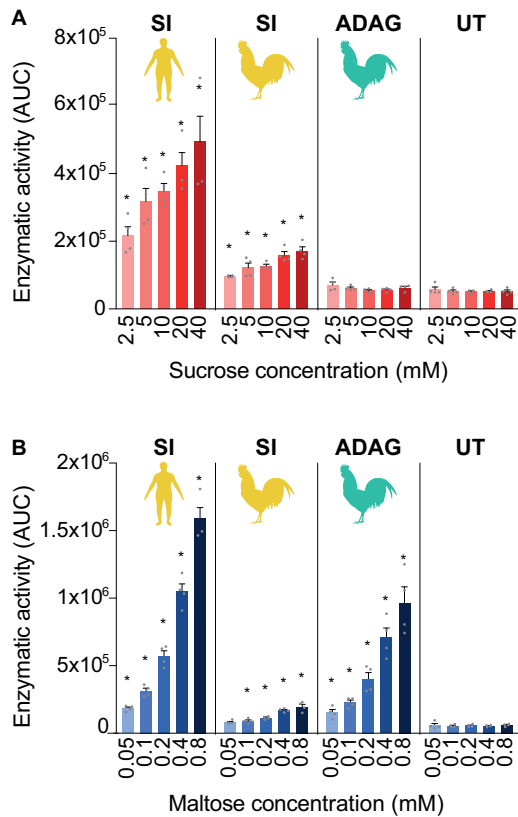
activity against sucrose was indistinguishable to that of mock-transfected cells (fig. 4; Welch's two-tailed  $t$ -test,  $*P \leq 0.05$ ). In the mammals that have been studied, MGAM has strong activity against maltose and very low (in laboratory rats; Lee et al. 2016) or no (in pigs and humans; Sørensen et al. 1982; Semenza et al. 2014) activity against sucrose. Our results are consistent with previous proteomic characterization of two enzymes with different functions in chickens (Hu et al. 1987) and support the hypothesis of functional convergence, after independent duplications, between bird ADAG and mammalian MGAM.

## Discussion

The two-enzyme (sucrase/maltase) system that completes the digestion of carbohydrates such as sucrose and starch in mammals is considered widespread among vertebrates and assumed to be the result of a single ancestral gene duplication. Here, unexpectedly, we reveal that most vertebrates have a single common gene, which in some fishes, as well as in birds and mammals, duplicated independently to produce a two-enzyme system. Strikingly, this enzyme expressed by the duplicated gene converged in function in chickens and

**FIG. 2.** Continued

archosaurs, or in the ancestor of amniotes, followed by loss in archosaurs and mammals. The second subunits of mammals and reptiles are shown in yellow and red; a third subunit (purple) is retained in lizards and turtles.



**Fig. 4.** Chicken ADAG and SI exhibit functional complementarity. Functional assays detecting free glucose due to sucrase or maltase activity from cells expressing human SI, chicken SI, or chicken ADAG. After incubation with 2.5–40 mM sucrose (A) and 0.05–0.8 mM maltose (B), enzymatic activity was compared with mock-transfected cells (UT) ( $n = 4$  replicates, mean  $\pm$  SE; Welch’s two-tailed  $t$ -test,  $*P \leq 0.05$ , Holm correction). Like human MGAM, chicken ADAG hydrolyzes maltose and has undetectable sucrase activity.

mammals. Genomic and proteomic comparative analyses can upend long-held hypotheses and open new questions with direct implications for understanding how animals process food.

In mammals, both SI and MGAM not only hydrolyze maltose but also hydrolyze starch and the oligosaccharides that result from the action of pancreatic and salivary amylases (Diaz-Sotomayor et al. 2013). Because MGAM seems to have higher activity against oligosaccharides with higher numbers of residues than SI and has higher specificity for maltose (Lee et al. 2016), the joint action of SI and MGAM has been hypothesized to be indispensable for small intestinal digestion of plant-derived  $\alpha$ -D-glucose oligomers to glucose (Nichols et al. 2003). However, a large number of songbird species seem to retain only a single intestinal  $\alpha$ -glucosidase gene (fig. 1), and seemingly paradoxically, many of these species are grain- or seedeaters that subsist on high-starch diets (Karasov 1990). The physiological consequences of a dual- or single-enzyme system across birds remain to be explored.

The unexpected diversity of vertebrate glucosidases opens up many questions. Phylogenetically guided comparative functional assays can determine the function of the ancestral

SI and the catalytic abilities of the distinct subunits, as well as the relative roles of the two enzymes in the digestion of different substrates and their relationship to distinct dietary types. Within birds, the genomic location and true phylogenetic distribution of ADAG, including the pattern of apparent loss among many songbirds may be resolved by sequencing of additional high-quality genomes. In addition, functional studies of  $\alpha$ -glucosidases can be used to examine the evolutionary forces shaping the evolution of gene duplications. Duplication often results in subfunctionalization or neofunctionalization of the duplicated gene (Zhang 2003), and although a large number of models regarding the origin and consequences of gene duplicates have been proposed, distinguishing between the role of selection and drift in the fixation and maintenance of duplicate copies can be challenging (Innan and Kondrashov 2010; Levasseur and Pontarotti 2011). The multiplicity of origins of vertebrate  $\alpha$ -glucosidases presents an opportunity to identify the mechanisms involved in the evolution of gene duplications, as well as to explore their evolutionary and functional ramifications (Crow and Wagner 2006; Innan and Kondrashov 2010). The functional consequences of duplications of SI (together with the subsequent duplication of ADAG in woodpeckers and MGAM in mammals and the additional subunit duplication in some reptiles) as well as the selective forces shaping these duplication events remain unknown. Digestive differences between species (such as the lack of sucrase activity in pinnipeds [Wacker et al. 1984] and some bird clades [Martínez del Río 1990]) may reflect underlying changes in the catalytic abilities of  $\alpha$ -glucosidases. The extent to which the immense variability in vertebrate diets and digestive abilities is related to functional changes of key digestive enzymes, as well as the extent of convergent changes in enzymatic properties in species with similar diets, will be exciting to explore.

## Materials and Methods

### Phylogenetic Analyses

We searched for protein sequences of  $\alpha$ -glucosidase genes from the NCBI database using BlastP (Altschul et al. 1990) with human MGAM (accession number: NP\_004659.2) and SI (NP\_001032.2) as query sequences. Sequences that did not contain a transmembrane domain (tested using the TMHMM server v2.0) were excluded. As platypus SI appears to be composed of a single subunit, we did not include either platypus SI or MGAM sequences in the phylogeny. MGAM2 was also omitted as many sequences contained long repetitive regions. We recovered a data set that includes 175  $\alpha$ -glucosidase sequences from 129 vertebrates. Alignments, model-testing, and phylogenetic inference were all run on the CIPRES Science Gateway (Miller et al. 2010). First, a multiple sequence alignment was generated by MUSCLE ver.3.7 (Edgar 2004). We used the program ModelFinder (Kalyaanamoorthy et al. 2017), implemented in IQ-Tree v.1.6.10 (Nguyen et al. 2015) to select the most appropriate model of amino acid substitution, and phylogenies were inferred using both maximum likelihood (ML) and Bayesian methods. ML trees were inferred with RAXML v.8.2.10

(Stamatakis 2015), using the Jones, Taylor, and Thornton matrix substitution model with the CAT model of rate of heterogeneity (PROTCATIJTT) and Lewis ascertainment bias correction; support was assessed using 1,000 nonparametric bootstrap replicates. After trimming poorly aligned regions using the program trimAl (Capella-Gutiérrez et al. 2009), the Bayesian trees were inferred using ExaBayes v.1.5 (Aberer et al. 2014) using the JTTDCMUT model. For Bayesian analyses, we ran two ExaBayes replicates until convergence criteria (sdsfConvergence = 0.01) were met (after 11,525,000 generations), sampling every 1,000 generations. Convergence of parameters and branch lengths was assessed by checking that effective sample sizes were >200 in the Tracer v1.7.1 software (Rambaut et al. 2018), and that the average standard deviation of split frequencies neared zero (0.009996) using ExaBayes (“sdsf” tool). We calculated a majority rule consensus tree, discarding 25% of initial samples as burn-in (ExaBayes, “consense”). The topology of the best-supported trees built by ML or Bayesian methods was congruent (supplementary fig. S1, Supplementary Material online).

### Phylogeny of $\alpha$ -Glucosidase Subunits

The phylogeny of the separated subunit domains of the  $\alpha$ -glucosidase sequences from figure 1 and supplementary table S1, Supplementary Material online (SI and derived  $\alpha$ -glucosidase, FDAG1-3, MGAM, or ADAG) was inferred using RAXML v.8.2.10 on the CIPRES Science Gateway after alignment of subunits in MUSCLE ver.3.7. Individual subunits were generated by dividing each species' protein sequence into two fragments, one homologous to the N-terminal region (referred to as subunit 1 and including up to residue 954, which is the cleavage point between subunits, Ren et al. 2011) and the other to the C-terminal region (subunit 2, including all residues from residue 960 to the protein's end) of human MGAM. In some species, a third subunit was identified and extracted (~residue 1857 of human MGAM to the protein's end). RAXML analysis was conducted as described above; support was assessed using 1,000 rapid bootstrap replicates. We visualized phylogenetic trees with iTOL (Interactive Tree of Life) v4.2.3 (Letunic and Bork 2016) and FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). To distinguish between subunit domain duplication in the ancestor of amniotes (followed by loss in archosaurs and in mammals) or duplication in the ancestor of reptiles (followed by loss in archosaurs), given the low support at this node, the approximately unbiased topology test (Shimodaira and Hasegawa 1999; Shimodaira 2002) was conducted in IQ-Tree (Nguyen et al. 2015) but neither of these scenarios was able to be excluded.

### Analysis of Conservation of Synteny

We generated the syntenic blocks depicted in figure 2 and supplementary figure S2, Supplementary Material online, by searching for up- and downstream genes in the 10 kb windows flanking the  $\alpha$ -glucosidases (SI, MGAM, ADAG, or FDAG1-3) using NCBI (<http://www.ncbi.nlm.nih.gov/gene/>). Uncharacterized loci are shown in gray. If uncertainties in orthology or gene names were apparent, the identity of the

flanking genes was confirmed using sequence similarity searches by BlastP. To examine changes in gene order and to verify absence of genomic regions containing ADAG, we searched for genes flanking paleognath ADAG in other bird groups (supplementary fig. S2A, Supplementary Material online). We used a similar approach and searched for genes flanking chicken ADAG in songbirds (supplementary fig. S2B, Supplementary Material online).

### Animal Use

All animals were held and euthanized following protocols approved by the Institutional Animal Care and Use Committee at the University of Wisconsin (UW)-Madison or following the regulations of the European Union and the German Animal Welfare Regulation Governing Experimental Animals (TierSchVersV). Rodents were provided by the UW-Madison Research Animal Resource Center. Adult chickens for proteomic analysis were supplied by the Poultry Research Laboratory at UW-Madison. Zebra finches were obtained from a commercial vendor and house sparrows were captured with mist nets on the campus of the UW-Madison. Chicken tissue for cDNA synthesis was obtained from hatching chickens from the lab of Prof. Luksch (Technische Universität München).

### Proteomic Approaches (Overview)

Briefly, our approach involved first separating enterocytes from small intestinal tissue and isolating their brush-border membranes. We analyzed isolated brush-border proteins in two ways. First, we assessed the activity of  $\alpha$ -glucosidases against sucrose and maltose in native gels (“in-gel” hydrolysis, described below). After identifying the bands in the gels where activity was present, we cut the gel bands, solubilized their proteins, and used nanoliquid chromatography followed by mass spectrometry (nano-LC-MS/MS) to determine the presence or absence of SI, MGAM, and ADAG in the bands. We also quantified the relative abundance of  $\alpha$ -glucosidases in solubilized protein from entire isolated brush-border membranes.

### Preparation and Solubilization of Brush-Border Membranes, and In-Gel Enzymatic Hydrolysis

The separation of enterocytes and the preparation of brush-border membranes from the full length of the small intestine followed the methods described by Mac Donal et al. (2008). Isolated brush-border membranes were solubilized in buffered detergent (1% Triton X100, PH = 6.8) for 10 min at 40 °C. In-gel hydrolysis methods followed those described in references (Maestracci 1976; Manchenko 2003; Gaspari and Cuda 2011). Protein ( $\approx 10 \mu\text{g}$ ) was loaded in a polyacrylamide gel (run at 100 V and  $\approx 350 \text{ mA}$ ) using Tris-glycine (pH 8.3) as a running buffer. Subsequently, gels were incubated in a 56 mM maltose or sucrose buffered solution (maleate/NaOH, pH = 6.5). After 30- to 60-min incubations, the substrate solution was removed and the gels were developed in a glucose peroxidase/*o*-dianisidine solution (GO Assay Kit, Sigma GAGO-20) until distinct orange bands developed (1 h for maltose and 5 h for sucrose). The reaction was

stopped by rinsing with deionized water. The entire section of the gel in which stain was detectable was cut and the proteins within them were extracted and processed for nano-LC-MS/MS (see below).

### Protein Extraction from Gels for Submission of Peptides for Spectrometry

The entire stained region of the native gel containing activity staining was excised, destained in MeOH/H<sub>2</sub>O/NH<sub>4</sub>HCO<sub>3</sub> (50%:50%:25 mM), and the in-gel proteins were denatured for 10 min in an SDS/DTT/Tris-HCl solution (2%/1 mM/50 mM, pH 7) with subsequent 2 × 5 min washes in MeOH/H<sub>2</sub>O/NH<sub>4</sub>HCO<sub>3</sub> (50%:50%:25 mM). Gel fragments were washed twice for 5 min in MeOH/H<sub>2</sub>O/NH<sub>4</sub>HCO<sub>3</sub> (50%:50%:100 mM), dehydrated for 5 min in acetonitrile (ACN)/H<sub>2</sub>O/NH<sub>4</sub>HCO<sub>3</sub> (50%:50%:25 mM) then once more for 1 min in 100% ACN, dried in a vacuum concentrator (Speed-Vac, Thermo Fisher Scientific) for 2 min, reduced in 25 mM DTT (dithiothreitol in 25 mM NH<sub>4</sub>HCO<sub>3</sub>) for 30 min at 52 °C, alkylated with 55 mM IAA (iodoacetamide in 25 mM NH<sub>4</sub>HCO<sub>3</sub>) in darkness at room temperature for 30 min, washed twice in H<sub>2</sub>O for 30 s, equilibrated in 25 mM NH<sub>4</sub>HCO<sub>3</sub> for 1 min, dehydrated for 5 min in ACN/H<sub>2</sub>O/NH<sub>4</sub>HCO<sub>3</sub> (50%:50%:25 mM) then once more for 30 s in 100% ACN, dried again and rehydrated with 20 μl of trypsin solution (10 ng/μl trypsin [Promega] in 25 mM NH<sub>4</sub>HCO<sub>3</sub>/0.01% Protease MAX w/v [Promega]). An additional 30 μl of digestion solution (25 mM NH<sub>4</sub>HCO<sub>3</sub>/0.01% Protease MAX w/v) was added to facilitate complete rehydration and excess supernatant needed for peptide extraction. The digestion was conducted for 3 h at 42 °C. The supernatant was removed, and gel pieces were extracted for peptides with 2 × gel volume of ACN/H<sub>2</sub>O/trifluoroacetic acid (TFA) solution (70/30/0.75%). Extracted peptides were combined with the supernatant. Degraded Protease MAX was removed via centrifugation (max speed 16,000 × g, 10 min) and the solid phase extracted (ZipTipC<sub>18</sub> pipette tips [Millipore]). The extracted peptides were subsequently submitted for nano-LC-MS/MS analysis.

### Protein Extraction from Whole Brush-Border Membranes for Submission of Peptides for Spectrometry

Fifteen-microgram aliquots of total proteins from each brush-border membrane preparation were heated to 95 °C for 5 min. Samples were dried in a vacuum concentrator (Speed-Vac) and resolubilized in 50 μl 8 M CO(NH<sub>2</sub>)<sub>2</sub> in 50 mM NH<sub>4</sub>HCO<sub>3</sub> supplemented with 5 mM Tris, pH 8. 17 μl MeOH were added followed by 200 μl 50 mM NH<sub>4</sub>HCO<sub>3</sub>. Samples were aspirated to resuspend membranes and 2 μl DTT was added to a final concentration of 2 mM. Samples were then incubated at 57 °C for 30 min, after which 2 μl of C<sub>2</sub>H<sub>4</sub>INO (Sigma) was added to a final concentration of 5 mM followed by incubation at room temperature in the dark for 30 min. A final 2-μl aliquot of DTT was added to quench excess of C<sub>2</sub>H<sub>4</sub>INO. Trypsin (Promega) and Lys-C (Wako) were dissolved in 50 mM NH<sub>4</sub>HCO<sub>3</sub> to reach a concentration

of 0.1 μg/μl and added to yield a final enzyme to substrate (E:S) ratio of 1:50 each. Samples were incubated at 42 °C for 1 h, then an additional 3 μl aliquot of each protease was added and samples were incubated overnight at 37 °C with shaking. Following digestion, MeOH was removed by vacuum concentration and TFA was added to achieve a pH of 2. Solid-phase extraction was done using OMIX tips (Agilent) according to the manufacturer's instructions, eluting peptides with 100 μl of 75% ACN, 0.5% CH<sub>2</sub>O<sub>2</sub> in H<sub>2</sub>O. Eluates were dried by vacuum concentration and resolubilized in 20 μl 0.1% CH<sub>2</sub>O<sub>2</sub> in H<sub>2</sub>O for nano-LC-MS/MS analysis.

### Nano-LC-MS/MS and Data Analysis

Peptides from entire brush-border membrane and enzymatic in-gel hydrolysis were analyzed by nano-LC-MS/MS using the Agilent 1100 nanoflow system (Agilent) connected to a next-generation hybrid linear ion trap-orbitrap mass spectrometer (LTQ-Orbitrap Elite, Thermo Fisher Scientific) equipped with an EASY-Spray electrospray source using standard procedures (Gaspari and Cuda 2011). Chromatography of peptides prior to mass spectral analysis was accomplished with a capillary emitter column (PepMap C18, 3 μM, 100 Å, 150 × 0.075 mm, Thermo Fisher Scientific) onto which 3 μl of extracted peptides were automatically loaded. Raw MS/MS data were converted to mascot generic format (mgf) files using MSConvert (ProteoWizard: Open Source Software for Rapid Proteomics Tools Development). The resulting mgf files were used to search against Uniprot databases (UniProt Consortium 2018) from *Mus musculus* and *Rattus norvegicus*. Each of these databases includes 87,156 and 57,407 entries, respectively. RefSeq databases (O'Leary et al. 2016) were used for *T. guttata* and *Gallus gallus*. Each of these databases has 47,054 and 75,106 entries, respectively. Since no protein databases are currently available for *Passer domesticus*, we used the *T. guttata* RefSeq database. We added to the amino acid sequence databases a list of common contaminants and decoy sequences to establish false discovery rates using Mascot search engine 2.2.07 (<http://www.matrixscience.com/>) with variable methionine oxidation with asparagine and glutamine deamidation plus fixed cysteine carbamidomethylation. Peptide mass tolerance was set at 15 ppm and fragment mass at 0.6 Da. Protein annotations, significance of identification, and spectral-based quantification were done using the Scaffold software (version 4.3.2, Proteome Software Inc.). Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003, <http://proteinprophet.sourceforge.net/>), and protein identifications were accepted if they could be established at >99.0% probability with a 1% false discovery rate and containing at least 5 identified peptides. Proteins that contain similar peptides and that could not be differentiated based on MS/MS analysis alone were grouped based on parsimony.

### Comparative Proteomic Analyses

The α-glucosidases represented a very large fraction of the hydrolases (62 ± 6%, n = 5 species) found in the bands excised from zymography gels (supplementary table S2,

Supplementary Material online), consistent with a previous study in mice (McConnell et al. 2011). They represented  $40.1 \pm 3.2\%$ , ( $n = 5$  species) of the hydrolases in whole brush-border membranes (supplementary table S3, Supplementary Material online and fig. 3C). As predicted by the genomic data, we found SI and MGAM in *M. musculus* and *R. norvegicus*, and SI and ADAG in *G. gallus*. We only detected SI, not ADAG, in the extracted gel slices from the songbirds *T. guttata* and *P. domesticus*, even when we queried the databases using the sequence for *G. gallus* ADAG and the ADAGs of 11 other avian species (supplementary table S4, Supplementary Material online). SI, MGAM, and ADAG were the most abundant hydrolases in the extracted bands (supplementary table S2, Supplementary Material online). In the analysis of the whole membranes, the relative abundance of SI in the songbirds ( $38.4 \pm 1.5\%$  and  $44.9 \pm 3.6\%$  in *T. guttata* and *P. domesticus*, respectively) was roughly twice as high as that measured in *M. musculus* ( $19.6 \pm 2.0\%$ ), *R. norvegicus* ( $23.4 \pm 1.3\%$ ), and *G. gallus* ( $26.2 \pm 2.2\%$ ); ( $n = 4$  individuals/species, fig. 3 and supplementary table S3, Supplementary Material online).

#### Cloning of ADAG, Avian SI, and Human SI

We obtained the two chicken enzymes by polymerase chain reaction amplification from cDNA (SMARTScribe reverse-transcriptase, Clontech) synthesized from RNA extracted from oral and intestinal tissue using the RNeasy Fibrous Tissue Kit (Qiagen). Primers were designed in the software Geneious (Biomatters) using the predicted sequences XM\_015291762.2 (for SI) and XM\_015273018.2 (for ADAG) from NCBI, and PCR was performed using Advantage GC 2 Polymerase (Clontech). The human SI NM\_001041.4 was synthesized by Genewiz. These three sequences were cloned into the pEAK10 expression vector (Toda et al. 2011; Baldwin et al. 2014).

#### Cell Culture and Functional Assays

COS-7 cells were obtained from the Leibniz Institute DSMZ and maintained in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% fetal bovine serum. Cell culture and transfection were performed as described in Rodríguez et al. (2013) with modifications following Toda et al. (2011) and Baldwin et al. (2014). Cells were seeded into 12-well plates in low-glucose DMEM and transiently transfected in duplicate with 700 ng of plasmid per transfection using Lipofectamine 2000; untransfected cells were mock-transfected (no addition of plasmid). After transfection of 6.5 h, cells were moved to six-well plates for 24 h in DMEM supplemented with GlutaMAX (Gibco), then transferred to 96-well black-walled CellBIND surface-treated plates (Corning, NY). After overnight culture, cells were washed seven times with phosphate-buffered saline (PBS) and incubated at 37 °C for 2 h in PBS with either 2.5, 5, 10, 20, or 40 mM sucrose (Sigma–Aldrich), or 0.05, 0.1, 0.2, 0.4, or 0.8 mM maltose (Sigma–Aldrich). After incubation, enzymatic activity was assessed by detecting glucose in the supernatant, using the Glucose and Sucrose Assay Kit (Sigma–Aldrich) following manufacturer's recommendations.

Fluorescence was monitored using a FlexStation 3 microplate reader (Molecular Devices Co., Sunnyvale, CA), using excitation and emission wavelengths of 530 and 590 nm and measuring every 30 s for 20 min. Four replicates of two independent transfections were analyzed by calculating the mean of the area under the curve (AUC) for the first 4 min; data shown are mean values with error bars representing standard error of the mean. The mean AUC for each substrate and concentration for each of the three transfections was compared with the untransfected controls using Welch's two-tailed *t*-test followed by the Holm adjustment for multiple comparisons.

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

This work was supported by the Max Planck Society (to M.W.B.), the National Science Foundation (IOS-1354893) and the Department of Forest and Wildlife Ecology, UW-Madison (to W.H.K.), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) (PIP 834 to E.C.-V.), and the Universidad Nacional de San Luis (PROICO 2-0814 to E.C.-V.). We thank Jimena Manzur for her help during a preliminary stage of this research, Gregory A. Barrett-Wilt and Greg Sabat for their assistance with proteomic analysis, Cristián Gutiérrez-Ibáñez and Harald Luksch for providing chicken tissue, Chad D. Brock for advice on bioinformatics, and Fernando Aversa for providing additional computing resources and assistance. We thank Talia Karasov, Pablo Oteiza, Ian Baldwin, and two anonymous reviewers for comments on the manuscript. Most animal silhouettes are from PhyloPic (<http://phylopic.org/>) under Creative Commons Attribution (<http://creativecommons.org/licenses/by/3.0/>) and Attribution-ShareAlike (<http://creativecommons.org/licenses/by-sa/3.0/>) or in the public domain. Artist credit: Jim Bendon, T. Michael Keesey, Andrew Butko, Andrew A. Farke, Ryan Cupo, Milton Tan, Steven Traver, Daniel Jaron, Oliver Griffith, Pedro de Siracusa, Matt Martyniuk, Tony Ayling and Tracy A. Heath.

## References

- Aberer AJ, Kobert K, Stamatakis A. 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. *Mol Biol Evol.* 31(10):2553–2556.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215(3):403–410.
- Baldwin MW, Toda Y, Nakagita T, O'Connell MJ, Klasing KC, Misaka T, Edwards SV, Liberles SD. 2014. Evolution of sweet taste perception in hummingbirds by transformation of the ancestral umami receptor. *Science.* 345(6199):929–933.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 25(15):1972–1973.
- Chaudet MM, Amiri M, Marth N, Naim HY, Rose DR. 2019. Phylogenetic analysis reveals key residues in substrate hydrolysis in the isomaltase



- domain of sucrase-isomaltase and its role in starch digestion. *Biochim Biophys Acta*. 1863(9):1410–1416.
- Crow KD, Wagner GP. 2006. What is the role of genome duplication in the evolution of complexity and diversity? *Mol Biol Evol*. 23(5):887–892.
- Diaz-Sotomayor M, Quezada-Calvillo R, Avery SE, Chacko SK, Yan L-K, Lin A-M, Ao Z-H, Hamaker BR, Nichols BL. 2013. Maltase-glucoamylase modulates gluconeogenesis and sucrase-isomaltase dominates starch digestion gluconeogenesis. *J Pediatr Gastroenterol Nutr*. 57(6):704–712.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32(5):1792–1797.
- Finlayson SD, Moore PA, Johnston JR, Berry DR. 1990. Two staining methods for selectively detecting isomaltase and maltase activity in electrophoresis gels. *Anal Biochem*. 186(2):233–235.
- Gaspari M, Cuda G. 2011. Nanoproteomics: methods and protocols. In: Toms SA, Weil RJ, editors. *Nano LC–MS/MS: a robust setup for proteomic analysis*. Totowa (NJ): Humana Press. p. 115–126.
- Gericke B, Amiri M, Naim HY. 2016. The multiple roles of sucrase-isomaltase in the intestinal physiology. *Mol Cell Pediatr*. 3(1):2
- Hu CB, Spiess M, Semenza G. 1987. The mode of anchoring and precursor forms of sucrase-isomaltase and maltase-glucoamylase in chicken intestinal brush-border membrane. Phylogenetic implications. *Biochim Biophys Acta*. 896(2):275–286.
- Innan H, Kondrashov F. 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet*. 11(2):97–108.
- Jiao H, Zhang L, Xie H-W, Simmons NB, Liu H, Zhao H. 2019. Trehalase gene as a molecular signature of dietary diversification in mammals. *Mol Biol Evol*. 36(10):2171–2183.
- Jones K, Sim L, Mohan S, Kumarasamy J, Liu H, Avery S, Naim HY, Quezada-Calvillo R, Nichols BL, Mario Pinto B. 2011. Mapping the intestinal alpha-glucogenic enzyme specificities of starch digesting maltase-glucoamylase and sucrase-isomaltase. *Bioorg Med Chem*. 19:3929–3934.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 14(6):587–589.
- Karasov WH. 1990. Digestion in birds: chemical and physiological determinants and ecological implications. *Stud Avian Biol*. 13:391–415.
- Karasov WH. 2011. Digestive physiology: a view from molecules to ecosystem. *Am J Physiol Regul Integr Comp Physiol*. 301(2):R276–284.
- Karasov WH, Douglas AE. 2013. Comparative digestive physiology. *Compr Physiol*. 3(2):741–783.
- Karasov WH, Martínez del Rio C, Caviedes-Vidal E. 2011. Ecological physiology of diet and digestive systems. *Annu Rev Physiol*. 73(1):69–93.
- Lee B-H, Rose DR, Lin A-M, Quezada-Calvillo R, Nichols BL, Hamaker BR. 2016. Contribution of the individual small intestinal  $\alpha$ -glucosidases to digestion of unusual  $\alpha$ -linked glycemic disaccharides. *J Agric Food Chem*. 64(33):6487–6494.
- Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. 44(W1):W242–245.
- Levasseur A, Pontarotti P. 2011. The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics. *Biol Direct*. 6(1):11.
- Li J-T, Hou G-Y, Kong X-F, Li C-Y, Zeng J-M, Li H-D, Xiao G-B, Li X-M, Sun X-W. 2015. The fate of recent duplicated genes following a fourth-round whole genome duplication in a tetraploid fish, common carp (*Cyprinus carpio*). *Sci Rep*. 5(1):8199.
- Mac Donal O, Chediack JG, Caviedes-Vidal E. 2008. Isolation of epithelial cells, villi and crypts from small intestine of pigeons (*Columba livia*). *Biocell*. 32(3):219–227.
- Maestracci D. 1976. Enzymic solubilization of the human intestinal brush border membrane enzymes. *Biochim Biophys Acta*. 433(3):469–481.
- Manchenko GP. 2003. Handbook of detection of enzymes on electrophoretic gels. Boca Raton (FL): CRC Press LLC.
- Martínez del Rio C. 1990. Dietary phylogenetic and ecological correlates of intestinal sucrase and maltase activity in birds. *Physiol Zool*. 63(5):987–1011.
- McConnell RE, Benesh AE, Mao S, Tabb DL, Tyska MJ. 2011. Proteomic analysis of the enterocyte brush border. *Am J Physiol Gastrointest Liver Physiol*. 300(5):G914–G926.
- McWhorter TJ, Caviedes-Vidal E, Karasov WH. 2009. The integration of digestion and osmoregulation in the avian gut. *Biol Rev*. 84(4):533–565.
- Miller MA, Pfeiffer W, Schwartz T. editors. 2010. Gateway Computing Environments Workshop (GCE). Proceedings of the Gateway Computing Environments Workshop (GCE); 2010 Nov 4; New Orleans, LA.
- Naumoff DG. 2007. Structure and evolution of the mammalian maltase-glucoamylase and sucrase-isomaltase genes. *Mol Biol*. 41(6):962–973.
- Nesvizhskii AI, Keller A, Kolker E, Aebersold R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*. 75(17):4646–4658.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32(1):268–274.
- Nichols BL, Avery S, Sen P, Swallow DM, Hahn D, Sterchi E. 2003. The maltase-glucoamylase gene: common ancestry to sucrase-isomaltase with complementary starch digestion activities. *Proc Natl Acad Sci U S A*. 100(3):1432–1437.
- Nichols BL, Avery SE, Quezada-Calvillo R, Kilani SB, Lin A-M, Burrin DG, Hodges BE, Chacko SK, Opekun AR, Hindawy ME, et al. 2017. Improved starch digestion of sucrase-deficient shrews treated with oral glucoamylase enzyme supplements. *J Pediatr Gastroenterol Nutr*. 65(2):e35–e42.
- O’Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 44(D1):D733–D745.
- Oosthuizen V, Weldrick DP, Naudé RJ, Oelofsen W, Muramoto K, Kamiya H. 1998. Ostrich intestinal glycohydrolases: distribution, purification and partial characterisation. *Int J Biochem Cell Biol*. 30(3):339–352.
- Pontremoli C, Mozzi A, Forni D, Cagliani R, Pozzoli U, Menozzi G, Vertemara J, Bresolin N, Clerici M, Sironi M. 2015. Natural selection at the brush-border: adaptations to carbohydrate diets in humans and other mammals. *Genome Biol Evol*. 7(9):2569–2584.
- Prakash K, Patil SD, Hegde SN. 1983. Studies on the intestinal disaccharidases of the pigeon. III. Separation, purification and properties of sucrase-isomaltase and maltase-glucoamylase. *Arch Int Physiol Biochim*. 91(5):379–390.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol*. 67(5):901–904.
- Ren L, Qin X, Cao X, Wang L, Bai F, Bai G, Shen Y. 2011. Structural insight into substrate specificity of human intestinal maltase-glucoamylase. *Protein Cell*. 2:827–836.
- Rodríguez R, Ramsay AJ, Quesada V, Garabaya C, Campo E, Freije JMP, López-Otín C. 2013. Functional analysis of sucrase–isomaltase mutations from chronic lymphocytic leukemia patients. *Human Mol Genet*. 22(11):2273–2282.
- Semenza G, Auricchio S, Mantei N. 2014. The online metabolic and molecular bases of inherited disease. In: Valle D, Beaudet AL, Vogelstein B, Kinzler KW, Antonarakis SE, Ballabio A, Gibson KM, Mitchell G, editors. *Small-intestinal disaccharidases*. New York: The McGraw-Hill Companies, Inc.
- Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol*. 51(3):492–508.
- Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol*. 16(8):1114–1114.

- Sørensen SH, Norén O, Sjöström H, Danielsen EM. 1982. Amphiphilic pig intestinal microvillus maltase/glucoamylase. Structure and specificity. *Eur J Biochem.* 126:559–568.
- Stamatakis A. 2015. Using RAxML to infer phylogenies. *Curr Protoc Bioinformatics.* 51(1):6.14.11–16.14.14.
- Swallow DM. 2003. Genetic influences on carbohydrate digestion. *Nutr Res Rev.* 16(1):37–43.
- Tamaoki K, Okada R, Ishihara A, Shiojiri N, Mochizuki K, Goda T, Yamauchi K. 2016. Morphological, biochemical, transcriptional and epigenetic responses to fasting and refeeding in intestine of *Xenopus laevis*. *Cell Biosci.* 6:2.
- Toda Y, Okada S, Misaka T. 2011. Establishment of a new cell-based assay to measure the activity of sweeteners in fluorescent food extracts. *J Agric Food Chem.* 59(22):12131–12138.
- UniProt Consortium. 2018. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47:D506–D515.
- Wacker H, Aggeler R, Kretschmer N, O'Neill B, Takesue Y, Semenza G. 1984. A two-active site one-polypeptide enzyme: the isomaltase from sea lion small intestinal brush-border membrane. Its possible phylogenetic relationship with sucrase-isomaltase. *J Biol Chem.* 259(8):4878–4884.
- Zhang J. 2003. Evolution by gene duplication: an update. *Trends Ecol Evol.* 18:292–298.