nature genetics

Resurrecting ancestral alcohol dehydrogenases from yeast

J Michael Thomson[1,4], Eric A Gaucher[2], Michelle F Burgan[3,4], Danny W De Kee[2], Tang Li[2], John P Aris[1] & Steven A Benner[1,3]

Modern yeast living in fleshy fruits rapidly convert sugars into bulk ethanol through pyruvate. Pyruvate loses carbon dioxide to produce acetaldehyde, which is reduced by alcohol dehydrogenase 1 (Adh1) to ethanol, which accumulates. Yeast later consumes the accumulated ethanol, exploiting Adh2, an Adh1 homolog differing by 24 (of 348) amino acids. As many microorganisms cannot grow in ethanol, accumulated ethanol may help yeast defend resources in the fruit[1]. We report here the resurrection of the last common ancestor[2] of Adh1 and Adh2, called Adh_A. The kinetic behavior of Adh_A suggests that the ancestor was optimized to make (not consume) ethanol. This is consistent with the hypothesis that before the Adh1-Adh2 duplication, yeast did not accumulate ethanol for later consumption but rather used Adh_A to recycle NADH generated in the glycolytic pathway. Silent nucleotide dating suggests that the Adh1-Adh2 duplication occurred near the time of duplication of several other proteins involved in the accumulation of ethanol, possibly in the Cretaceous age when fleshy fruits arose. These results help to connect the chemical behavior of these enzymes through systems analysis to a time of global ecosystem change, a small but useful step towards a planetary systems biology.

Generating ethanol from glucose in the presence of dioxygen, only to then reoxidize the ethanol, is energetically expensive (**Fig. 1**). For each molecule of ethanol converted to acetyl-coenzyme A, a molecule of ATP is used. This ATP would not be 'wasted' if pyruvate made initially from glucose were delivered directly to the citric acid cycle. This implies that yeast has a reason, transcending simple energetic efficiency, for rapidly converting available sugar in fruit to bulk ethanol in the presence of dioxygen.

One hypothesis to explain this 'inefficiency' holds that yeast, which are relatively resistant to ethanol toxicity, may accumulate ethanol to defend resources in the fruit from competing microorganisms[3]. Although the ecology of wine yeast is certainly more complex than this simple hypothesis implies[4], fleshy fruits offer a large reservoir of carbohydrate, and this resource must have value to competing organisms as well as to yeast. For example, humans have exploited the preservative value of ethanol since prehistory[5].

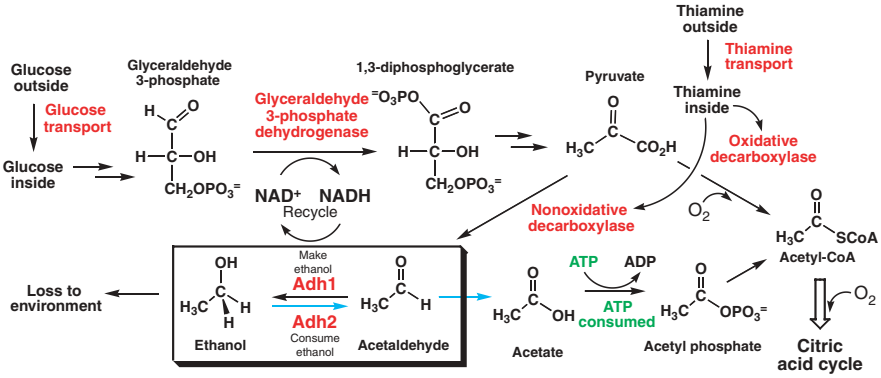

Figure 1 The pathway by which yeast make, accumulate and then consume ethanol. Enzymes in red are associated with gene duplications that, according to the transition redundant exchange clock[18], arose nearly contemporaneously. The make-accumulate-consume pathway is boxed. The shunting of the carbon atoms from pyruvate into (and then out of, blue arrows) ethanol is energy-expensive, consuming a molecule of ATP (green) for every molecule of ethanol generated. This ATP is not consumed if pyruvate is oxidatively decarboxylated directly to acetyl-coenzyme A to enter the citric acid cycle directly (dashed arrow to the right). If dioxygen is available, the recycling of NADH does not need the acetaldehyde-to-ethanol reduction.

[1]Department of Anatomy and Cell Biology, University of Florida, Gainesville, Florida 32610, USA. [2]Foundation for Applied Molecular Evolution, Gainesville, Florida 32601, USA. [3]Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA. [4]Present addresses: Department of Cell & Developmental Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA (J.M.T.); Department of Pharmacology and Cancer Biology, Duke University, Durham, North Carolina 27710, USA (M.F.B.). Correspondence should be addressed to S.A.B. (benner@chem.ufl.edu).

Published online 1 May 2005; doi:10.1038/ng1553

Both the timing of Adh expression in *Saccharomyces cerevisiae* and the properties of the expressed proteins are consistent with this hypothesis. The yeast genome encodes two alcohol dehydrogenases (Adhs) that interconvert ethanol and acetaldehyde (**Fig. 1**)[6]. The first (Adh1) is expressed at high levels constitutively. Its kinetic properties optimize it as a catalyst to make ethanol from acetaldehyde[7,8]. The Michaelis constant (K_M) for ethanol for Adh1 is high (17,000–20,000 μM), consistent with ethanol being a product of the reaction. After the sugar concentration drops, the second dehydrogenase (Adh2) is derepressed. This paralog oxidizes ethanol to acetaldehyde with kinetic parameters suited for this role. The K_M for ethanol for Adh2 is low (600–800 μM), consistent with ethanol being its substrate.

Adh1 and Adh2 are homologs[8] differing by 24 of 348 amino acids. Their common ancestor, called Adh_A, had an unknown role. If Adh_A existed in a yeast that made, but did not accumulate, ethanol, its physiological role would presumably have been the same as that of lactate dehydrogenase in mammals during anaerobic glycolysis: to recycle NADH generated by the oxidation of glyceraldehyde-3-phosphate (**Fig. 1**)[9]. Lactate in human muscle is removed by the bloodstream; ethanol would be lost by the yeast to the environment. If so, Adh_A should be optimized for ethanol synthesis, as is modern Adh1. The kinetic behaviors of Adh_A should resemble those of modern Adh1 more than those of modern Adh2, with a high K_M for ethanol.

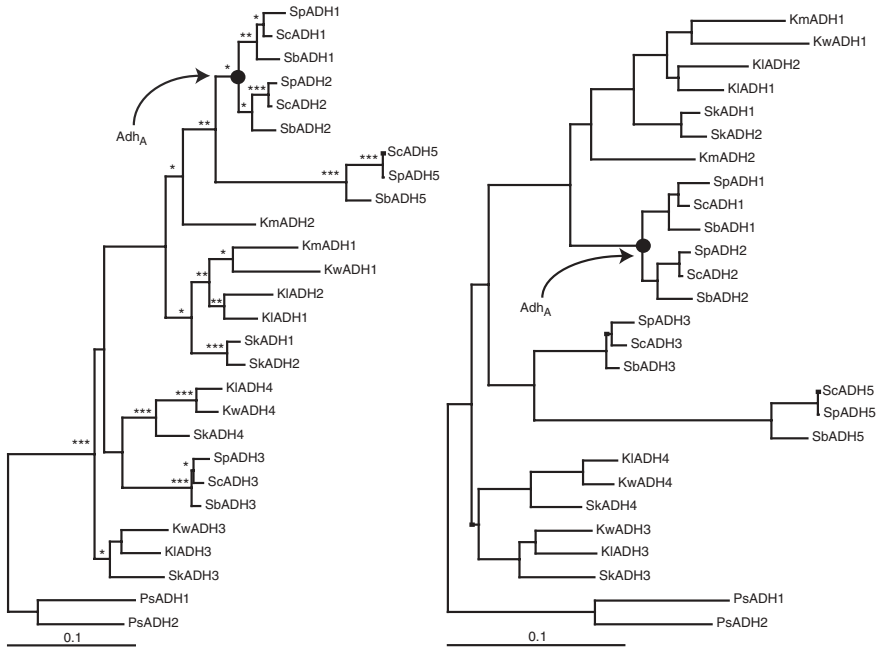

Figure 2 Maximum likelihood trees interrelating sequences determined in this work with sequences in the publicly available database. Shown are the two trees with the best (and nearly equal) ML scores using the following parameters estimated from the data: substitutions A↔C, A↔T, C↔G and G↔T = 1.00; A↔G = 2.92; and C↔T = 5.89; empirical base frequencies; and proportion of invariable sites and the shape parameter of the gamma distribution set to 0.33 and 1.31, respectively. The scale bar represents the number of substitutions per codon per unit evolutionary time. Single, double and triple asterisks represent bootstrap values greater than 50%, 70% and 90%, respectively.

We tested this hypothesis using a paleobiochemical experiment[2]. We cloned and sequenced 15 additional homologs of Adh from yeasts related to *S. cerevisiae* (**Supplementary Table 1** online). We constructed a maximum likelihood evolutionary tree using PAUP*4.0 (**Fig. 2**)[10] to combine these with sequences already in the database. We then reconstructed maximum likelihood sequences for Adh_A using both codon and amino acid models in PAML[11] (**Table 1**). When the posterior probability that a particular amino acid occupied a particular site was >80%, we assigned that amino acid at that site in Adh_A.

When the posterior probability was <80% or the most probabilistic ancestral states estimated using the codon and amino acid models were not in agreement, we considered the site to be ambiguous and constructed alternative ancestral sequences. For example, the posterior probabilities for two amino acids (methionine and arginine) were nearly equal at site 168 in Adh_A, three amino acids (lysine, arginine and threonine) were plausibly present at site 211, and two (aspartic acid and asparagine) were plausible for site 236. To handle these ambiguities, we generated twelve (all 2 × 2 × 3 combinations) candidate Adh_A proteins by constructing genes that encoded them, using these genes to transform a strain of *S. cerevisiae*[12] from which both Adh1 and Adh2 had been deleted, and then expressing them from the Adh1 promoter. After showing that the ancestral sequences could rescue the double deletion phenotype (**Supplementary Fig. 1** online), we isolated the candidate ancestral proteins, purified them to homogeneity on a Cibracon-blue agarose column[13] and then analyzed their kinetic behaviors[14] (**Table 1**).

To assess the quality of the data, we calculated Haldane values ($K_{eq} = V_f K_{iq} K_p / V_r K_{ia} K_b$)[15], where V_f and V_r are forward and reverse maximal velocities; K_{ia} and K_{iq} are disassociation constants for NAD^+ and NADH, respectively; and K_b and K_p are Michaelis constants for ethanol and acetaldehyde, respectively) from the experimental data. These reproduced the literature equilibrium constant for the reaction to within a factor of two. One variant, called MTN, had very low catalytic activity in both directions. We inferred that this particular candidate ancestor was not present in the ancient yeast.

Notably, the kinetic properties of the remaining ancestral Adh_A candidates resembled those of Adh1 more than those of Adh2 (**Table 1**). From this, we inferred that the ancestral yeast did not have an Adh specialized for the consumption of ethanol, similar to modern Adh2, but rather had an Adh specialized for making ethanol, similar to modern Adh1. This suggests that before the Adh1-Adh2 duplication, the ancestral yeast did not consume ethanol. This implies that the ancestral yeast also did not accumulate ethanol under aerobic conditions for future consumption and that the make-accumulate-consume strategy emerged after Adh1 and Adh2 diverged. These interpretations are robust with respect to the ambiguities in the reconstructions.

For modern Adh1, reported K_M values range from 17,000 to 24,000 μM for ethanol (our value is 20,000 μM), from 170 to 240 μM for NAD^+ (ours is 220 μM), from 1,100 to 3,400 μM for acetaldehyde (ours is 1,500 μM) and from 110 to 140 μM for NADH (ours is 164 μM)[14]. These comparisons, together with the Haldane analysis, provide a view of the experimental error in the kinetic parameters reported here. Our interpretations are based on differences well outside of this error.

Table 1 Kinetic properties of Adh1, Adh2 and candidate ancestral Adh_A

	K_M (μM)			
Sample[a]	Ethanol	NAD^+	Acetaldehyde	NADH
Adh1	20,060	218	1,492	164
MKD	17,280	511	1,019	144
MKN	13,750	814	1,067	1,106
MRD	11,590	734	1,265	287
MRN	10,960	554	1,163	894
MTD	10,740	467	959	190
MTN	NA	NA	NA	NA
RKD	8,497	449	1,066	142
RKN	7,238	407	1,085	735
RRD	7,784	400	1,074	203
RRN	8,403	172	1,156	1,142
RTD	6,639	254	1,083	316
RTN	7,757	564	1,158	477
Adh1[b]	24,000	240	3,400	140
Adh1[c]	17,000	170	1,100	110
Adh2[b]	2,700	140	45	28
Adh2[c]	810	110	90	50
Adh3[c]	12,000	240	440	70
Adh1[c] (*S. pombe*)	14,000	160	1,600	100
Adh1 (M270L)[c]	19,000	630	1,000	80
KIP20369[d]	27,000	2,800	1,200	110
KIX64397[d]	23,000	2,200	1,700	180
KIX62766[d]	2,570	310	100	20
KIX62767[d]	1,560	200	3,100	30

[a]The three letters designate the amino acids at positions 168, 211 and 236 (*e.g.*, MKD = Met168 Lys211 Asp236). The remaining residues were the same as in Adh1, except for the following changes (using sequence numbering of Adh1 from *S. cerevisiae*): Asn15 Pro30 Thr58 Ala74 Glu147 Leu213 Ile232 Cys259 Val265 Leu270 Ser277 Asn324. [b]From ref. 8. [c]From ref. 16. [d]From ref. 30. KI, *Kluyveromyces lactis*.

When paralogs are generated by duplication, the duplicate acquiring the new functional role is believed to evolve more rapidly than the one retaining the primitive role[16]. If this were generally true, one might identify the functionally innovative duplicate by a bioinformatics analysis. This may be true for many genes, but chemical principles do not obligate this outcome, and it is not manifest with these Adh paralogs. Here, the rate of evolution is not markedly faster in the lineage leading to Adh2 (having the derived function) than in the lineage leading to Adh1 (having the primitive function). Thus, a paleobiochemistry experiment was necessary to assign the primitive behavior.

The Haldane ratio relates various kinetic parameters (k_{cat}, K_M, K_{diss}) that can change as a result of changing the amino acid sequence to the overall equilibrium constant, which the enzyme (being a catalyst) cannot change. Thus, if a lower K_M for ethanol is selected, other terms in the Haldane equation must change to keep the ratio the same. This is observed in data for the ancestral proteins prepared here and the natural enzymes.

The assignment of a primitive function to Adh_A raises a broader historical question: did the Adh1-Adh2 duplication, and the accumulate-consume strategy that it presumably enabled, become fixed in response to a particular selective pressure? Connecting molecular change to organismic fitness is always difficult[17] but is necessary if reductionist biology is to move through systems biology to a planetary biology that determines why as well as how changes occurred[18]. Hypothetically, the emergence of a make-accumulate-consume strategy may have been driven by the domestication of yeast by humans selecting for yeast that accumulate ethanol. Alternatively, the strategy might have been driven by the emergence of fleshy fruits that offered a resource worth defending using ethanol accumulation. We might distinguish between the two by estimating the date of the Adh1-Adh2 duplication. Even with large errors in the estimate, a distinction should be possible, as human domestication occurred in the past million years, whereas fleshy fruits arose in the Cretaceous age, after the first angiosperms appeared in the fossil record 125 million years ago[19] but before the extinction of the dinosaurs 65 million years ago[20].

The topology of the evolutionary tree (**Fig. 2**) suggests that the Adh1-Adh2 duplication occurred before the divergence of the *sensu strictu* species of *Saccharomyces*[21] but after the divergence of *Saccharomyces* and *Kluyveromyces*. The date of divergence of *Saccharomyces* and *Kluyveromyces* is unknown but is estimated to have occurred 80 ± 15 million years ago[22]. This date is consistent with a transition redundant exchange clock[23], which exploits the fractional identity (f_2) of silent sites in conserved twofold redundant codon systems to estimate the time since the divergence of two genes (**Supplementary Note** online). Between pairs of presumed orthologs from *Saccharomyces* and *Kluyveromyces*, f_2 is typically 0.82, not much lower than the f_2 value separating Adh1 and Adh2 (0.85)[18] but much lower than paralog pairs within the *Saccharomyces* genome that seem to have arisen by more recent duplication (~0.98)[24].

Notably, Adh1 and Adh2 are not the only pair of paralogs with an f_2 value between 0.80 and 0.86 (ref. 18). We analyzed ~350 pairs of paralogs in the yeast genome that shared at least 100 silent sites and diverged by less than 120 point-accepted replacements per 100 sites, and we identified 15 pairs with f_2 values between 0.80 and 0.86. These represent eight duplications that occurred near the time of the Adh1-Adh2 duplication, if f_2 values are assumed to support a clock.

These duplications are not randomly distributed in the yeast genome. Rather, six of the eight duplications involve proteins that participate in the conversion of glucose to ethanol (**Table 2**). Furthermore, the enzymes arising from the duplicates are those that seem, from expression analysis, to control flux from hexose to ethanol[1,25]. These include proteins that import glucose, pyruvate decarboxylases that generate acetaldehyde from pyruvate, the transporter that imports thiamine for these decarboxylases and the Adhs (**Fig. 1**). If the f_2 clock (within its expected variance) is assumed to date paralogs in yeast, this cluster suggests that several proteins other than Adh duplicated as part of the emergence of the new make-accumulate-consume strategy, near the time that fleshy fruit arose.

The six duplications proposed to be part of the emergence of the make-accumulate-consume strategy (with f_2 values between 0.80 and 0.86) are not associated with one of the documented blocks of genes duplicated in ancient fungi, possibly as part of a whole-genome duplication[16,26]. But two duplications in genes that are not associated with fermentation that have f_2 values between 0.80 and 0.86 are part of a duplication block (**Table 2** and **Supplementary Note** online). The silent sites for most gene pairs associated with blocks are nearly equilibrated (with the prominent exception of ribosomal proteins), suggesting that most blocks arose by duplications more ancient than those with f_2 values between 0.80 and 0.86. Therefore, the hypothesis that a set of six time-correlated duplication (**Fig. 1** and **Table 2**) generated the make-accumulate-consume strategy in yeast near the time when fermentable fruit emerged is not inconsistent with the whole-genome duplication hypothesis.

The ecology of fermenting fruit is complex. In rotting fruits, *S. cerevisiae* becomes dominant after fermentation begins, whereas

Table 2 Duplication in the *Saccharomyces cerevisiae* genome where $0.80 < f_2 < 0.86$

SGD name	gi number	Trivial name	Annotation and comments
Inosine-5′-monophosphate dehydrogenase family (3 paralogs, 3 pairs, 2 duplications)[a]			
$f_2 = 0.803$[b]	Pair associated with Wolfe duplication blocks 1 and 44		
YAR073W	gil456156	IMD1	Nonfunctional homolog, near telomere, not expressed
YLR432W	gil665971	IMD3	Inosine-5′-monophosphate dehydrogenase
$f_2 = 0.825$[b]			
YLR432W	gil665971	IMD3	Inosine-5′-monophosphate dehydrogenase
YHR216W	gil458916	IMD2	Inosine-5′-monophosphate dehydrogenase
Subfamily pair: YHR216W and YAR073W	$f_2 = 0.93$ (proposed recent duplication creating a pseudogene)		
Sugar transporter family A (4 paralogs, 4 pairs, 3 duplications)[c]			
$f_2 = 0.805$[b]	Pair not associated with any duplication block		
YJR158W	gil1015917	HXT16	Sugar transporter repressed by high glucose levels
YNR072W	gil1302608	HXT17	Sugar transporter repressed by high glucose levels
$f_2 = 0.806$[b]	Pair not associated with any duplication block		
YDL245C	gil1431418	HXT15	Sugar transporter induced by low glucose, repressed by high glucose
YNR072W	gil1302608	HXT17	Sugar transporter repressed by high glucose levels
$f_2 = 0.809$[b]	Pair not associated with any duplication block		
YJR158W	gil1015917	HXT16	Sugar transporter repressed by high glucose levels
YEL069C	gil603249	HXT13	Sugar transporter induced by low glucose, repressed by high glucose
$f_2 = 0.810$[b]	Pair not associated with any duplication block		
YEL069C	gil603249	HXT13	Sugar transporter induced by low glucose, repressed by high glucose
YDL245C	gil1431418	HXT15	Sugar transporter
Subfamily pair: YEL069C and YNR072W	$f_2 = 0.932$ (proposed recent duplication)		
Subfamily pair: YJR158W and YDL245C	$f_2 = 1.000$ (proposed very recent duplication)		
Chaperone family A (2 paralogs, 1 pair, 1 duplication)[a]			
$f_2 = 0.81$	Pair associated with Wolfe duplication block 48		
YMR186W	gil854456	HSC82	Cytoplasmic chaperone induced 2-3 fold by heat shock
YPL240C	gil1370495	HSP82	Cytoplasmic chaperone, pheromone signaling, Hsf1p regulation
Phosphatase/thiamine transport family A (2 paralogs, 1 pair, 1 duplication)[c]			
$f_2 = 0.818$	Pair not associated with any duplication block		
YBR092C	gil536363	PHO3	Acid phosphatase implicated in thiamine transport
YBR093C	gil536365	PHO5	Acid phosphatase, one of three repressible phosphatases
Pyruvate decarboxylase family A (2 paralogs, 1 pair, 1 duplication)[c]			
$f_2 = 0.835$	Pair not associated with any duplication block		
YLR044C	gil1360375	PDC1	Pyruvate decarboxylase, major isoform
YLR134W	gil1360549	PDC5	Pyruvate decarboxylase, minor isoform
By ortholog analysis, *Saccharomyces bayanus* (gil515236) diverged from *S. cerevisiae* after the $f_2 = 0.835$ duplication; *Kluyveromyces* diverged before.			
Glyceraldehyde-3-phosphate dehydrogenase family (3 paralogs, 3 pairs, 2 duplications)[c]			
$f_2 = 0.845$[b]	Pair not associated with any duplication block		
YJL052W	gil1008189	TDH1	Glyceraldehyde-3-phosphate dehydrogenase
YGR192C	gil1323341	TDH3	Glyceraldehyde-3-phosphate dehydrogenase
$f_2 = 0.845$[b]	Pair not associated with any duplication block		
YJL052W	gil1008189	TDH1	Glyceraldehyde-3-phosphate dehydrogenase
YJR009C	gil1015636	TDH2	Glyceraldehyde-3-phosphate dehydrogenase
Subfamily pair: YJR009C and YGR192C	$f_2 = 0.991$ (proposed very recent duplication)		
Alcohol dehydrogenase family (2 paralogs, 1 pair, 1 duplication)[c]			
$f_2 = 0.848$	Pair not associated with any duplication block		
YMR303C	gil798945	Adh2	Alcohol dehydrogenase, glucose-repressible
YOL086C	gil1419926	Adh1	Alcohol dehydrogenase, constitutive
Spermine transporter family (2 paralogs, 1 pair, 1 duplication)[a]			
$f_2 = 0.86$	Pair associated with Wolfe duplication block 34		
YGR138C	gil1323230	TPO2	Spermine transporter activity
YPR156C	gil849164	TPO3	Spermine transporter activity

Table 2 continued on following page

Table 2 Continued

SGD name	gi number	Trivial name	Annotation and comments
Sugar transporter family B (3 paralogs, 3 pairs, 2 duplications)[c]			
$f_2 = 0.847$[b]	Pair not associated with any duplication block		
YDR343C	gil1230670	HXT6	Sugar transporter, high-affinity high basal levels
YDR345C	gil1230672	HXT3	Sugar transporter, low-affinity glucose transporter
$f_2 = 0.854$[b]	Pair not associated with any duplication block		
YDR342C	gil1230669	HXT7	Sugar transporter, high-affinity, high basal levels
YDR345C	gil1230672	HXT3	Sugar transporter, low affinity
Subfamily pair: YDR342C and YDR343C	$f_2 = 0.994$ (proposed very recent duplication)		

[a]Not associated with fermentation. These are associated with duplication blocks in the yeast genome[16], where the high value of f_2 (typically equilibrated in block paralog pairs) may reflect either variance or selective pressure to conserve silent sites in individual codons. [b]These pairs represent a family generated with a single duplication with f_2 value between 0.80 and 0.86 and subsequent duplication(s) in the derived lineages. Trees are shown in **Supplementary Figure 2** online. Paralog pairs were considered only if they have with at least 100 aligned silent sites and are not separated by more than 120 point-accepted mutations per 100 aligned amino acid sites. [c]Associated with the pathway to make, accumulate and then consume ethanol. Genes involved in the fermentation pathway that are not rate-limiting[1,25] generally do not have duplicates in the yeast genome (*e.g.*, hexokinase, glucose-6-phosphate isomerase, phosphofructokinase, aldolase, triose phosphate isomerase and phosphoglycerate kinase are all present in one isoform). Enolase has two paralogs (ENO1 and ENO2), where $f_2 = 0.946$. These are distantly related to a homolog known as ERR1, with the silent sites equilibrated. Phosphoglycerate mutase has three paralogs, GM1, GM2 and GM3, with silent sites that are essentially equilibrated (**Supplementary Note** online).

osmotic stress and pH, as well as ethanol, seem to inhibit the growth of competing organisms[1]. Nevertheless, the emergence of bulk ethanol may not be unrelated to other changes in the ecosystem at the end of the Cretaceous age, which include the extinction of the dinosaurs and the emergence of mammals and fruit flies[27,28]. Thus, this paleogenetics experiment is a small but necessary step in connecting the chemical behavior of individual enzymes operating as part of a multienzyme system, through metabolism and physiology, to the ecosystem and the fitness of organisms in it. We expect that this combination of chemistry, enzymology and cell biology with genomics, geology, paleontology and planetary science will be a key activity in addressing the 'why' questions in contemporary biology.

METHODS

Yeast strains and genomic DNA preparation. We purchased strains from ATCC. We used isogenic *S. cerevisiae* strains BY4742 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and BY4741 (MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0) for Adh1 and Adh2 deletions[12]. We used the double deletion strain YMT-1D (MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 adh1::URA3 adh2::URA3) to express Adh_A. We purchased strains for cloning new *ADH* genes from ATCC.

***ADH* gene cloning[12].** We designed degenerate primers from an alignment of existing *S. cerevisiae* and *Kluyveromyces lactis* Adh proteins. We obtained partial *ADH* gene sequences using degenerate primers NSDP-1-FOR (positions 127–151 of *S. cerevisiae ADH1*) and NSDP-2-REV (positions 357–381 of *S. cerevisiae ADH1*) and amplification parameters of 96 °C for 2.5 min; followed by 30 cycles of 94 °C for 30 s, 39–50 °C for 30 s and 72 °C for 1 min; followed by a final extension step of 72 °C for 5 min. We pooled PCR products and shotgun-cloned them into TOPO PCR4-TA cloning vector (Invitrogen). We sequenced random clones and used them to generate a consensus sequence for each allele. We then used the consensus sequences to generate gene-specific primers. We used a combination of partial inverse PCR and linker-mediated PCR to clone the remaining flanking DNA.

Ancestral *ADH* genes. We generated ancestral *ADH* genes by site-directed mutagenesis of *S. cerevisiae ADH1* cloned into the pRS415 vector[12]. We constructed the expression plasmid using primers ADH1/-1605-For (corresponding to 1,605 nucleotides upstream of the *ADH1* start site) and ADH1/+378-Rev (corresponding to 378 nucleotides downstream of the *ADH1* stop codon). Amplification parameters were 96 °C for 2.5 min; followed by 20 cycles of 94 °C for 30 s, 58 °C for 45 s and 72 °C for 3.5 min; followed by a final extension step of 72 °C for 5 min. We digested the amplified product with *Xba*I and *Xho*I and cloned it into pRS415. We used the resulting construct pRS415-ADH1 as a template for creating the Adh_A constructs using site-directed mutagenesis (Stratagene).

Expression strain construction[12]. We constructed single deletion strains using PCR. We used isogenic *S. cerevisiae* strains BY4741 (MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and BY4742 (MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0; ATCC) to create deletions of the two primary *ADH* alleles *ADH2* and *ADH1*, respectively. We designed primers to amplify *URA3* with 50 bp of homology to the 5′ and 3′ untranslated regions of *ADH1* or *ADH2*. We used the URA3-containing vector pRS316 as a template and cycled 30 times at 94 °C for 30 s, 55 °C for 45 s and 72 °C for 1 min. We disrupted alleles as described[29]. We created the double deletion strain (YMT-1D) by mating the single deletion strains followed by tetrad dissection.

Protein purification[12,13]. We transformed the double deletion strain YMT-1D with either pRS415-ADH1 or pRS415-Adh_A. We grew cells in 2% glucose yeast minimal medium to midlog phase, with $A_{600} = 0.6$. We applied extracts to a Cibracon blue–coupled agarose column and eluted protein with a gradient of 0–200 μM NADH. We pooled fractions and applied them to a Superdex 200 gel-filtration column to remove excess $NAD^+(H)$ and collect tetrameric Adh.

Kinetic measurements[12,14]. We measured kinetics with NAD^+, ethanol, NADH and acetaldehyde by holding the concentration of one constant and varying the concentrations of the others. All reactions were done in potassium phosphate buffer (83 mM, pH 7.3) with KCl (40 mM) at 25 °C. Ethanol reactions contained semicarbazide (10 mM) to avoid product inhibition by acetaldehyde. We monitored reactions by measure UV absorbance at 340 nm on a Cary Varian spectrophotometer.

GenBank accession numbers. Adh homologs from yeasts related to *S. cerevisiae*, AY216989 through AY217003.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank A. Falcon and T. Barnash for their assistance. Funding was provided by the US National Research Council, the US National Aeronautics and Space Administration's Astrobiology Institute (E.A.G.) and the US National Aeronautics and Space Administration Exobiology program (S.A.B.).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 6 October 2004; accepted 15 March 2005
Published online at http://www.nature.com/naturegenetics/

1. Pretorius, I.S. Tailoring wine yeasts for the new millennium: novel approaches to the ancient art of winemaking. *Yeast* **16**, 675–729 (2000).
2. Thornton, J.W. Resurrecting ancient genes. Experimental analysis of extinct molecules. *Nat. Rev. Genet.* **5**, 366–376 (2004).

3. Boulton, B., Singleton, V.L., Bisson, L.F. & Kunkee, R.E. Yeast and biochemistry of ethanol fermentation. in *Principles and Practices of Winemaking* 139–172 (Chapman and Hall, New York, 1996).
4. Fleet, G.H. & Heard, G.M. Yeast growth during fermentation. in *Wine Microbiology and Biotechnology* 27–54 (Harwood Academic Publishers, Chur, Switzerland, 1993).
5. McGovern, P.E. *et al.* Fermented beverages of pre- and proto-historic China. *Proc. Natl. Acad. Sci. USA* **101**, 17593–17598 (2004).
6. Wills, C. Production of yeast alcohol dehydrogenase isoenzymes by selection. *Nature* **261**, 26 (1976).
7. Fersht, A.R. *Enzyme Structure and Mechanism*. (W.H. Freeman, San Francisco, 1977).
8. Ellington, A.D. & Benner, S.A. Free energy differences between enzyme bound states. *J. Theor. Biol.* **127**, 491–506 (1987).
9. Stryer, L. *Biochemistry* 3rd edn. (Freeman, New York, 1988).
10. Swofford, D.L. *PAUP* Phylogenetic Analysis Using Parsimony Version 4* (Sinauer Associates, Sunderland, Massachusetts, 1998).
11. Yang, Z.H. PAML. A program package for phylogenetic analysis by maximum likelihood. *CABIOS* **15**, 555–556 (1997).
12. Thomson, J.M. *Interpretive Proteomics: Experimental Paleogenetics as a Tool to Analyze Function and Discover Pathways in Yeast*. Ph.D. dissertation. (University of Florida, Gainesville, 2002).
13. Weinhold, E.G., Glasfeld, A., Ellington, A.D. & Benner, S.A. Structural determinants of the stereospecificity of alcohol dehydrogenase from yeast. *Proc. Natl. Acad. Sci. USA* **88**, 8420–8424 (1991).
14. Ganzhorn, A.J., Green, D.W., Hershey, A.D., Gould, R.M. & Plapp, B.V. Kinetic characterization of yeast alcohol dehydrogenases. Amino acid residue 294 and substrate specificity. *J. Biol. Chem.* **262**, 3754–3761 (1987).
15. Segel, I.H. *Enzyme Kinetics* (Wiley, New York, 1975).
16. Kellis, M., Birren, B.W. & Lander, E.S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617–624 (2004).
17. Kreitman, M. & Akashi, H. Molecular evidence for natural selection. *Ann. Rev. Ecol. Systematics* **26**, 403–422 (1995).
18. Benner, S.A., Caraco, M.D., Thomson, J.M. & Gaucher, E.A. Planetary biology. Paleontological, geological, and molecular histories of life. *Science* **293**, 864–868 (2002).
19. Sun, G. *et al. Archaefructaceae*, a new basal angiosperm family. *Science* **296**, 899–904 (2002).
20. Collinson, M.E. & Hooker, J.J. Fossil evidence of interactions between plants and plant-eating mammals. *Philos. Trans. R. Soc. London Ser. B* **333**, 197–208 (1991).
21. Fernandez-Espinar, M.T., Barrio, E. & Querol, A. Analysis of the genetic variability in the species of the *Saccharomyces sensu stricto* complex. *Yeast* **20**, 1213–1226 (2003).
22. Berbee, M.L. & Taylor, J.W. Dating the evolutionary radiations of the true fungi. *Can. J. Bot.* **71**, 1114–1127 (1993).
23. Benner, S.A. Interpretive proteomics. Finding biological meaning in genome and proteome databases. *Adv. Enzyme Regul.* **43**, 271–359 (2003).
24. Lynch, M. & Conery, J.S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
25. Schaaff, I., Heinisch, J. & Zimmerman, F.K. Overproduction of glycolytic enzymes in yeast. *Yeast* **5**, 285–290 (1989).
26. Wolfe, K.H. & Shields, D.C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (2001).
27. Ashburner, M. Speculations on the subject of alcohol dehydrogenase and its properties in *Drosophila* and other flies. *Bioessays* **20**, 949–954 (1998).
28. Barrett, P.M. & Willis, K.J. Did dinosaurs invent flowers? Dinosaur-angiosperm coevolution revisited. *Biological Rev.* **76**, 411–447 (2001).
29. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. & Cullin, C. A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic Acids Res.* **21**, 3329–3330 (1993).
30. Bozzi, A., Saliola, M., Falcone, C., Bossa, F. & Martini, F. Structural and biochemical studies of alcohol dehydrogenase isozymes from *Kluyveromyces lactis*. *Biochim. Biophys. Acta* **1339**, 133–142 (1997).