Molecular Biology and Evolution

No Rosetta Stone for a Sense–Antisense Origin of Aminoacyl tRNA Synthetase Classes

Tom A. Williams, Kenneth H. Wolfe, and Mario A. Fares

Department of Genetics, Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin, Ireland

Aminoacyl tRNA synthetases (aaRS) are crucial enzymes that join amino acids to their cognate tRNAs, thereby implementing the genetic code. These enzymes fall into two unrelated structural classes whose evolution has not been explained. The leading hypothesis, proposed by Rodin and Ohno, is that the two classes originated as a pair of sense–antisense genes encoded on opposite strands of a single DNA molecule. This unusual idea obtained its main support from reports of a "Rosetta stone": a locus where genes for heat shock protein 70 (HSP70) and an Nicotinamide adenine dinulecotide-specific glutamate dehydrogenase (NAD-GDH), which are structurally homologous to the two classes of aaRS, overlap extensively on complementary DNA strands. This remarkable locus was first characterized in the oomycete *Achlya klebsiana* and has since been reported in many other species. Here we present evidence that the open reading frames on the antisense strand of *HSP70* genes are spurious, and we identify a more probable candidate for the gene encoding the oomycete NAD-GDH enzyme. These results cast extensive doubt on the Rosetta Stone argument.

Introduction

Aminoacyl tRNA synthetases (aaRS) are the enzymes that specifically join amino acids to their cognate tRNAs prior to translation, thereby implementing the genetic code. AaRS can be divided into two very different structural classes, with 10 members each. The evolutionary origin of these enzymes, and therefore the system of translation as we know it, is a tantalizing mystery. Rodin and Ohno (1995) made the dramatic proposal that the two structural classes of aaRS arose on the opposite strands of the same DNA molecule. This proposal was based on the observation that two conserved motifs of Class I aaRS (the HIGH and KMSKS motifs) could potentially be encoded by the complementary strands of DNA sequences coding for two conserved motifs present in Class II aaRS (Motifs 2 and 1, respectively).

The Rodin–Ohno hypothesis received important support from Carter and Duax (2002), who identified a gene from the oomycete *Achyla klebsiana* as a possible "Rosetta stone" for sense–antisense coding of proteins related to Class I and Class II aaRS. Their work was based on an earlier report that heat shock protein 70 (HSP70) and an Nicotinamide adenine dinulecotide-specific glutamate dehydrogenase (NAD-GDH) are encoded as a sense–antisense pair by a single DNA sequence in *A. klebsiana* (LeJohn, Cameron, Yang, and Rennie 1994). The reported overlap between the genes was extensive, with over 1,800 bp of the proposed *NAD-GDH* gene being located on the reverse complement of the *HSP70* gene (fig. 1). Because canonical dehydrogenases are structurally similar to Class I aaRS, and HSP70 has structural homology to Class II aaRS, Carter and Duax (2002) concluded that the *A. klebsiana* gene proved that such structurally divergent proteins could be encoded on opposite strands of the same DNA molecule. Antisense open reading frames (ORFs) were later reported opposite *HSP70* genes in *Drosophila auraria* (Konstantopoulou et al. 1995) and a variety of other organisms (Rother et al. 1997; Silke 1997) and opposite the *HSP70*-related gene *GRP78* of *Neurospora crassa* (Monnerjahn et al. 2000). The concept of two genes evolving as a completely overlapping sense–antisense pair is surprising, given the heightened selective constraints that would act on both sequences and is unprecedented outside of virus genomes.

Although LeJohn and colleagues performed a thorough biochemical characterization of the NAD-GDH activity in *A. klebsiana*, including purification of the enzyme (LeJohn, Cameron, Yang, MacBeath, et al. 1994; LeJohn, Cameron, Yang, and Rennie 1994; Yang and LeJohn 1994), their evidence that the NAD-GDH enzyme is encoded by the ORF opposite *HSP70* is questionable (see Results and Discussion). Here we present evidence that the antisense ORF is spurious, even though it is present in many species, and we identify a different oomycete gene as a more probable candidate for the locus encoding the NAD-GDH enzyme.

Materials and Methods

Analysis of *Aphanomyces euteiches* EST data

Aphanomyces euteiches expressed sequence tags (ESTs; sequenced by Gaulin et al. 2008) homologous to the *A. klebsiana HSP70*/antisense-ORF (AS-ORF) genomic locus were identified by BlastN searches at National Centre for Biotechnology Information (NCBI), using as a query the sequence of the whole *A. klebsiana* genomic locus (6,575 bp made by merging GenBank accession numbers U02504 and U02505; LeJohn, Cameron, Yang, and Rennie 1994). The top-scoring 280 ESTs (BlastN $E \leq 3e-15$) were retrieved and assembled into contigs using CAP3 (Huang and Madan 1999). Two contigs corresponding to probable *A. euteiches HSP70* genes were identified. Contigs 1 and 2 had 83% DNA sequence identity to each other and 87% and 79% identity, respectively, to the *A. klebsiana* genomic sequence. The proteins encoded by contigs 1 and 2 had 92% amino acid sequence identity to each other and 97% and 92% amino acid sequence identity, respectively, to the *A. klebsiana* protein. We reasoned that contig 1 is the ortholog of the *A. klebsiana HSP70* gene sequenced by LeJohn et al. and show the locations of the 54 ESTs making up this contig in figure 1.

A contig of *A. euteiches* ESTs coding for the proposed NAD-GDH (fig. 2) was identified by TblastN searches at AphanoDB (Madoui et al. 2007) using the *N. crassa* protein as a query. The sequence shown in figure 2 was assembled by merging GenBank accession numbers CU354392 and CU354866, AphanoDB contig Ae_15AL7142, and AphanoDB sequence trace file NX0AINT6YK19CM1.SCF.

Key words: Rosetta Stone, Aminoacyl tRNA Synthetase, Hsp70.

E-mail: faresm@tcd.ie.

Mol. Biol. Evol. 26(2):445–450. 2009
doi:10.1093/molbev/msn267
Advance Access publication November 26, 2008

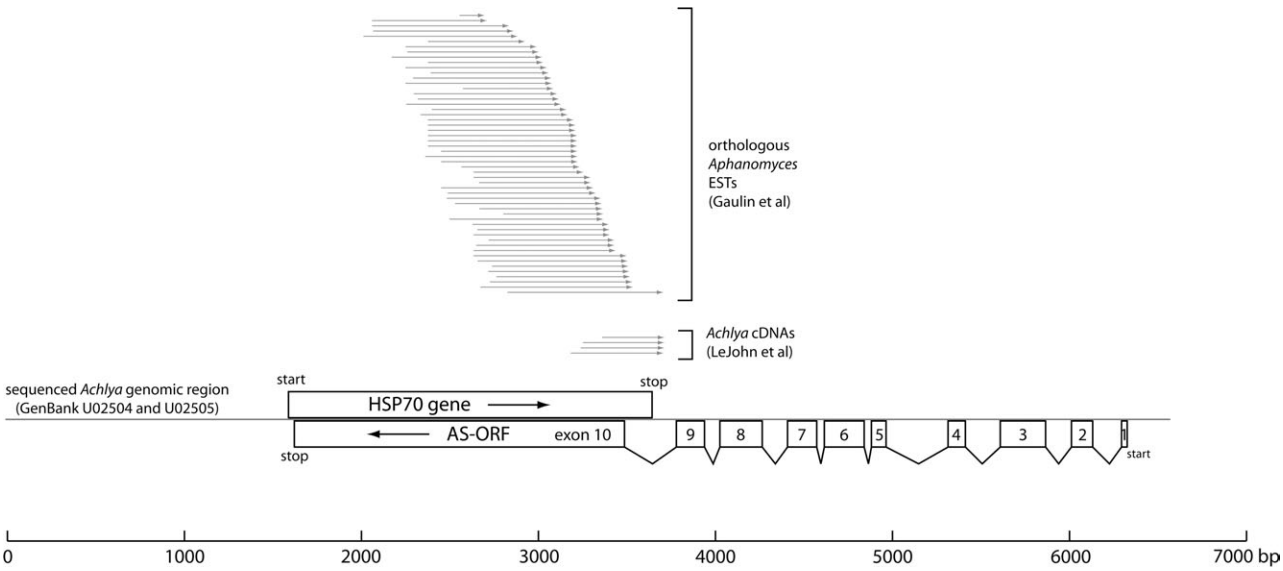

FIG. 1.—Map of the *Achlya klebsiana* genomic region as annotated by LeJohn, Cameron, Yang, and Rennie(1994). The *HSP70* gene is oriented from left to right and contains no introns. AS-ORF is a putative gene on the opposite DNA strand, consisting of 10 exons and 9 introns. Exon 10 of AS-ORF overlaps with the *HSP70* gene. LeJohn, Cameron, Yang, and Rennie (1994) proposed that AS-ORF is the gene coding for NAD-GDH. Gray arrows show the positions and transcriptional orientations of the four *Achlya* cDNAs sequenced by LeJohn, Cameron, Yang, MacBeath, et al. (1994) and of 54 ESTs from the orthologous locus in *Aphanomyces*. There are no ESTs or cDNAs corresponding to transcription from right to left in either species.

FUGUE Searches

FUGUE (Shi et al. 2001) searches were performed using the *A. klebsiana* AS-ORF protein sequence (accession number AAA17563; LeJohn, Cameron, Yang, and Rennie 1994) and the protein sequence encoded by the *A. euteiches* NAD-GDH contig described above. The program was run on the FUGUE web server at http://tardis.nibio.go.jp/fugue/prfsearch.html, which searches the Homologous Structure Alignment Database (HOMSTRAD) database (Mizuguchi et al. 1998).

Phylogenetic Tree Construction

Bacterial homologs of the protein encoded by AS-ORF were identified by a BlastP search at NCBI, using the AS-ORF sequence as a query, and the *HSP70* genes complementary to these AS-ORFs were retrieved. To assemble a set of dnaK sequences without an antisense ORF, we retrieved the protein and nucleotide sequences of the 1,000 BlastP hits to *Escherichia coli* O157:H7 *dnaK* (accession number NP_285706.1). We removed non-*dnaK* genes from this set and then counted the number of stop codons in the reverse complement of their coding sequences. We chose the sequences with the highest number of stop codons but avoided including numerous closely related sequences to increase the phylogenetic coverage of the "without AS-ORF" gene set. This resulted in 27 *dnaK* genes with at least 10 stop codons in the reverse complement of their coding sequences. The accession numbers of these *dnaK* proteins are provided in supplementary table 1 of the Supplementary Material online. The *dnaK* sequences were aligned with MUSCLE (Edgar 2004) using the default parameters. We used PROTTEST (Abascal et al. 2005) to pick the appropriate model of protein evolution (RtREV + I + G + F), and the trees were built using phyml (Guindon and Gascuel 2003), with 100 bootstraps. A consensus tree was produced using CONSENSE, which is part of the PHYLIP package (Felsenstein 1989). The Majority Rule (Extended) method was used to construct the tree and assign bootstrap values to branches.

Simulations

We identified eight *dnaK* genes used in the tree above that had more than 100 nucleotides of C-terminal sequence, which did not overlap with the corresponding AS-ORF. We aligned the protein sequences of these genes with MUSCLE (Edgar 2004) and used this alignment to build a codon-based nucleotide alignment. The codon alignment was split into two subalignments, corresponding to the regions of the *dnaK* genes that did and did not overlap the AS-ORF. To assess sequence conservation within each alignment, we used DNADIST, which is part of the PHYLIP package (Felsenstein 1989). DNADIST generated a table of pairwise Jukes–Cantor distances (Jukes and Cantor 1969) between the eight sequences, and we took the mean pairwise distance as a measure of conservation. We performed 100-fold bootstrapping on the shorter nonoverlapping alignment. Significance was tested using a *t*-test with one degree of freedom, based on the value ((Conservation(nonoverlapping) − Conservation(overlapping))/standard error(Conservation(nonoverlapping)). The *t*-value was 3.738, resulting in $P = 0.083$ for a one-tailed *t*-test.

Results and Discussion

Although many *HSP70* genes contain an AS-ORF on their opposite strand (Rother et al. 1997), the only reports

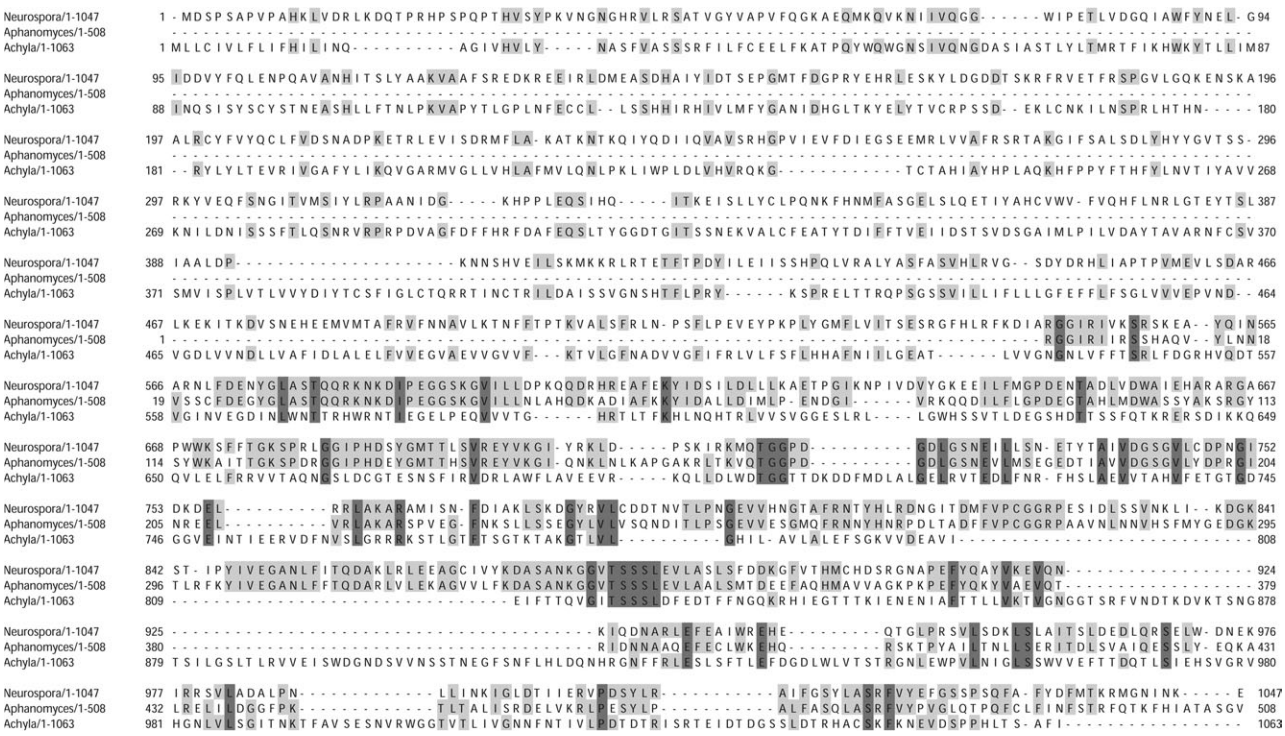

FIG. 2.—Protein sequence alignment of *Neurospora crassa* NAD-GDH with the partial putative NAD-GDH inferred from *Aphanomyces euteiches* EST data and the translation of the *Achlya klebsiana* AS-ORF. The *N. crassa* NAD-GDH shows 52% amino acid identity to the *A. euteiches* sequence but only 15.1% identity to the *A. klebsiana* sequence over the region where all three sequences could be aligned. Shading indicates residue identity. Alignment generated by T-Coffee (Notredame et al. 2000) using the default parameters and visualized in Jalview (Clamp et al. 2004).

that this AS-ORF codes for a GDH are three consecutive papers published by LeJohn and colleagues about the *A. klebsiana* locus (LeJohn, Cameron, Yang, MacBeath, et al. 1994; LeJohn, Cameron, Yang, and Rennie. 1994; Yang and LeJohn 1994). The experimental evidence that the *A. klebsiana* AS-ORF codes for NAD-GDH hinges on the specificity of the polyclonal antibody used in these studies. This antibody was raised against purified *A. klebsiana* NAD-GDH protein, but it was subsequently found to have dual specificity against both NAD-GDH and HSP70 in cell extracts (LeJohn, Cameron, Yang, MacBeath, et al. 1994; Yang and LeJohn 1994). When the antibody was used to screen an *A. klebsiana* cDNA expression library in λgt11, four cDNA clones were isolated, and all these appeared to be transcripts of the 3′ end of *HSP70* (LeJohn, Cameron, Yang, MacBeath, et al. 1994; fig. 1); two of the four cDNAs had poly(A) tails. Even though the recovery of *HSP70* cDNAs from the library is consistent with the antibody's anti-HSP70 activity (which was demonstrated and commented upon by LeJohn, Cameron, Yang, MacBeath, et al. 1994), LeJohn, Cameron, Yang, and Rennie (1994) pursued the hypothesis that the large AS-ORF on the opposite strand might code for NAD-GDH. Although LeJohn et al. did demonstrate transcription of both strands of the locus (fig.9 of LeJohn, Cameron, Yang, MacBeath, et al. 1994), we cannot find any experimental evidence in their papers that the AS-ORF actually codes for the observed NAD-GDH enzyme. The immunological reaction between the antibody and the cDNAs cloned in λgt11 cannot be interpreted as proof that the AS-ORF codes for NAD-GDH, because the antibody has been shown (LeJohn, Cameron, Yang, MacBeath, et al. 1994; Yang and LeJohn 1994) to have specificity for another protein (HSP70) that is made by the same cDNAs. Moreover, the genomic structure of 10 exons and 9 introns that was proposed for the complete AS-ORF gene (LeJohn, Cameron, Yang, and Rennie 1994) is not supported by any cDNA or EST evidence, and the structures of the four cDNAs cloned by LeJohn, Cameron, Yang, MacBeath, et al. (1994) is consistent only with them being derived from HSP70 mRNA (not spliced AS-ORF mRNAs; fig. 1).

The putative AS-ORF protein has ~20% amino acid sequence identity to known dehydrogenases when an alignment is forced (LeJohn, Cameron, Yang, and Rennie 1994; Carter and Duax 2002). Although this level of sequence identity does not preclude a distant but valid homology, protein–protein Blast searches of the AS-ORF against the NCBI nonredundant database result in no significant hits to any member of the canonical NAD-GDH family. In fact, there are no significant full-length (1063 aa) hits of any kind to the AS-ORF protein but only to that portion of the sequence that overlaps with *HSP70* on the opposite strand (601 aa at the C-terminus of AS-ORF). To investigate the relationship between the AS-ORF and typical dehydrogenases, we used FUGUE (Shi et al. 2001), a program that searches for distant but biologically relevant homologies by fitting a query sequence against a structure database that contains archaeal, eukaryotic, and bacterial representatives of the GDH family. Although submitting the NAD-GDH sequence of *N. crassa* (Kapoor et al. 1993) to FUGUE results in a highly significant

hit to the dehydrogenase structure (Z score = 8.44), submitting the AS-ORF sequence results in no significant hits (the best hit has a Z score of 2.85). FUGUE also provides the option to search using an alignment of Position Specific Iterative-Blast-derived homologs of the query sequence. Using this option, the AS-ORF sequence returned a hit ($Z = 4.67$), but this was not to a dehydrogenase structure, and the score was below the recommended cutoff provided by the program ($Z = 6$). Therefore, it seems that the AS-ORF is not homologous to previously described dehydrogenases in either sequence or structure.

We then investigated whether the NAD-GDH activity of *A. klebsiana* could be encoded by a locus other than that characterized by LeJohn, Cameron, Yang, and Rennie (1994). Our analysis made use of recent EST data (Gaulin et al. 2008) from *A. euteiches*, a closely related oomycete in the same family (Saprolegniaceae) as *A. klebsiana*. Using the *N. crassa* NAD-GDH protein sequence (Kapoor et al. 1993) as the query in a TblastN search, we obtained an *A. euteiches* contig encoding a protein fragment (508 aa). Over the 614-position region where all three sequences (*N. crassa* NAD-GDH, *A. euteiches* contig, and *A. klebsiana* AS-ORF) could be aligned, the *N. crassa* sequence displayed 52% amino acid identity to the *A. euteiches* contig and 15.1% identity to the translation of the *A. klebsiana* AS-ORF (fig. 2). The level of identity observed between the *N. crassa* and *A. euteiches* sequences is typical of that observed between members of the NAD-GDH protein family (Kersten et al. 1999). In contrast to the AS-ORF, this protein returned highly significant hits to canonical GDHs in both BlastP and FUGUE searches (Z score = 6.61). We suggest that this gene encodes the NAD-GDH enzyme of *A. euteiches*. We were also able to find an *A. euteiches* contig orthologous to the *HSP70*/AS-ORF locus of *A. klebsiana*. Although there was abundant evidence of transcription of the *HSP70* gene at this *A. euteiches* locus (54 ESTs), there were no ESTs corresponding to transcription of the complementary strand. Further, there were no ESTs corresponding to exons 1–9 of the AS-ORF, which do not overlap the *HSP70* sequence (fig. 1). These data suggest that *A. euteiches* has an NAD-GDH enzyme that is a typical member of the GDH family and that this enzyme is not encoded by the *A. euteiches* counterpart of the AS-ORF but at a different locus. We propose that the observed NAD-GDH biochemical activity of *A. klebsiana* (Yang and LeJohn 1994) is encoded by an *A. klebsiana* gene orthologous to the candidate oomycete NAD-GDH gene we identified in *A. euteiches* (fig. 2) but that this *A. klebsiana* gene has not yet been sequenced.

Our analysis of the *A. euteiches* EST data revealed two *HSP70* genes in this species, one orthologous (*AeHSP70-1*) and the other paralogous (*AeHSP70-2*) to the *HSP70*/AS-ORF locus of *A. klebsiana*. We noticed that the *AeHSP70-1* and *AeHSP70-2* sequences both also contain long antisense ORFs. Indeed, BlastP searching the *A. klebsiana* AS-ORF predicted protein sequence against the NCBI protein sequence database reveals a family of proteins with several phylogenetically scattered eukaryotic members and about 30 bacterial members (Rother et al. 1997; Carter and Duax 2002). Many of these bacterial sequences have been annotated as NAD-GDH enzymes on the basis of their similarity to the *A. klebsiana* AS-ORF. The presence of database homologs of AS-ORF might suggest that it is a functional gene even if, as argued above, it does not code for NAD-GDH. However, the observation that all the Blast hits to the AS-ORF sequence are within the region that overlaps *HSP70* raised the possibility that the apparent conservation of this family might be an artifact due to the presence on the opposite strand of sequences coding for HSP70, one of the most conserved proteins yet described (Gupta and Golding 1993). Indeed, every one of the apparent AS-ORF homologs in the database contains an intact *HSP70* gene on its opposite strand. We used bacterial sequences to test this hypothesis in two ways.

First, if there exist two kinds of *HSP70* genes—those with a functional AS-ORF and those without—then the two groups must have very different evolutionary histories, which should be reflected in their phylogeny. If the AS-ORFs are functional, there should be phylogenetic separation of *HSP70* sequences with and without the antisense gene. But if the AS-ORFs are artifacts, the *HSP70* sequences should cluster generally according to the species phylogeny. We retrieved 29 *dnaK* (bacterial *HSP70*) sequences from GenBank that were encoded by the complements of the AS-ORF homologs identified by Blast. We also retrieved 27 *dnaK* sequences that do not have an intact AS-ORF, based on the presence of at least 10 stop codons in the reverse complement of their coding sequences. A maximum likelihood, bootstrapped protein phylogeny of these sequences revealed extensive mixing between the *dnaK* sequences with and without AS-ORFs, a result that is difficult to explain if the antisense sequences are real genes (fig. 3). Furthermore, there is a clear correlation between the presence of an intact AS-ORF and the G + C content of the *HSP70* gene (fig. 3), as expected on purely statistical grounds (Merino et al. 1994; Silke 1997).

Second, eight of the *dnaK* sequences with AS-ORFs that were used to build the phylogeny contained a region of more than 100 nucleotides at the 3′ end of *dnaK* that was not overlapped by the AS-ORF, which permitted a further test of the hypothesis. The AS-ORFs are in the same frame as the *dnaK* genes, such that the third codon positions in one are base paired with the first codon positions in the other. We were therefore able to divide a nucleotide alignment of these eight sequences into two subalignments containing the third codon positions of *dnaK* from the regions that 1) overlapped and 2) did not overlap the antisense ORF. If the AS-ORF sequences are a real gene, the additional selective constraint provided by the first codon positions of this gene should result in higher conservation of codon third positions in the overlapping but not the nonoverlapping, regions of *dnaK*. We compared sequence conservation between the two subalignments using bootstrapping of the nonoverlapping region to increase robustness and taking mean pairwise Jukes–Cantor differences as a measure of conservation. We find that the difference in conservation between these regions is not statistically significant ($P = 0.083$, one-tailed t-test). This suggests that the apparent conservation of the bacterial AS-ORF homologs is due to the high conservation of *HSP70* on the other strand.

Our results point to two conclusions: First, the family of ORFs that has been identified antisense to *HSP70* genes in many species is unlikely to be a real gene family because

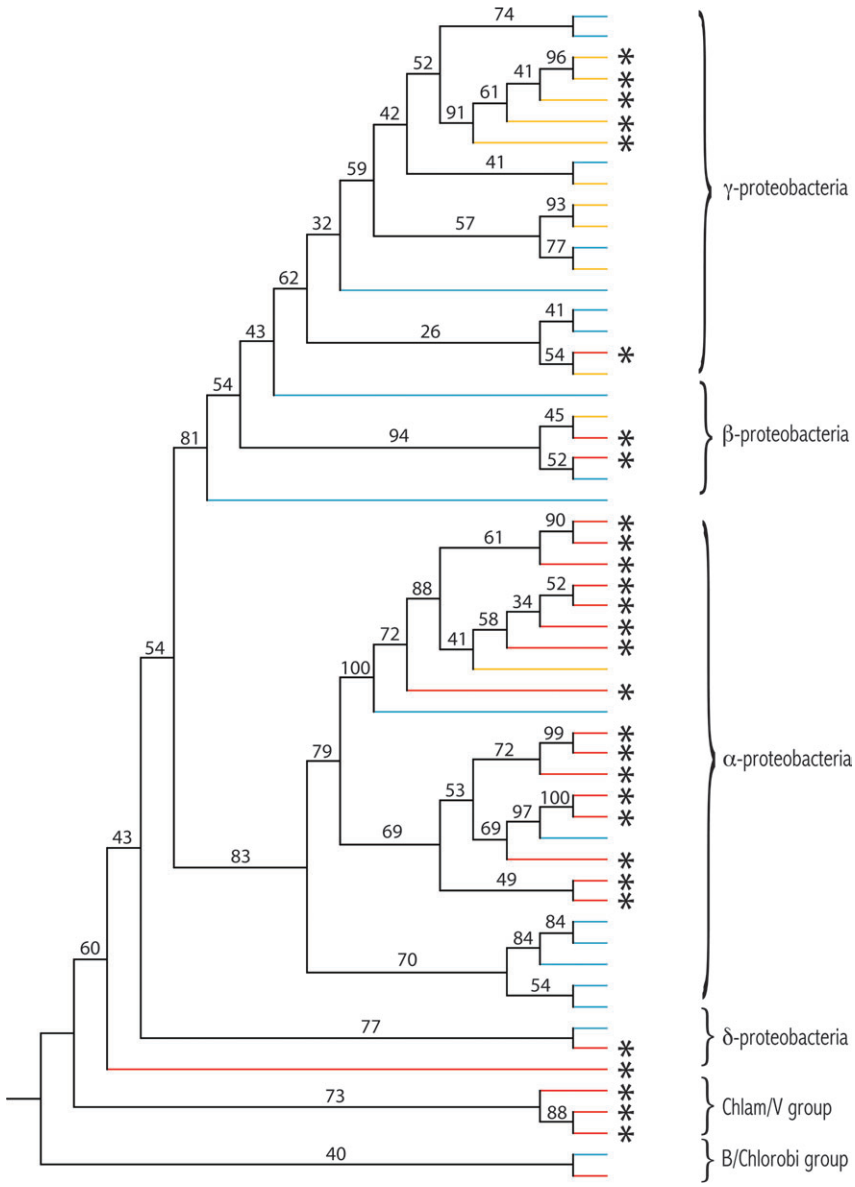

FIG. 3.—Protein phylogeny of *dnaK* sequences with and without an antisense glutamate dehydrogenase ORF (AS-ORF). An asterisk denotes the presence of an AS-ORF. The phylogeny recapitulates the standard view of bacterial relationships and reveals extensive mixing between genes with and without an AS-ORF, as expected if the AS-ORFs are artifactual. Branches are colored by third-position G + C content: high (66.7–100%, red), intermediate (33.4–66.6%, yellow), and low (0–33.3%, blue). The phylogeny is labeled by class. "Chlam/V group" is the Chlamydiae/Verrucomicrobia group. "B/Chlorobi group" is the Bacteriodetes/Chlorobi group. Species names and sequence accession numbers are included in Supplementary Material online.

these ORFs do not show the patterns of phylogenetic distribution, sequence constraint, or even transcription that would be expected if they code for functional proteins. We therefore doubt that any of the ORFs antisense to *HSP70* genes in any species are functional. Second, the evidence that one particular member of this putative family—the AS-ORF of *A. klebsiana*—codes for the NAD-GDH enzyme in this species is tenuous. We do not doubt the biochemical evidence that an NAD-GDH enzyme exists in *A. klebsiana*, but we do question LeJohn et al.'s claim that the gene coding for this enzyme is located antisense to *HSP70* because 1) their papers do not provide any biochemical evidence linking the NAD-GDH enzyme to the *HSP70* locus and 2) we have found a different oomycete gene that seems likely to code for NAD-GDH. Although proof of our proposal would require further biochemical experiments such as the purification and direct amino acid sequencing of an oomycete NAD-GDH enzyme, the results presented here cast substantial doubt on the proposed overlap between the *NAD-GDH* and *HSP70* genes, on which the Rosetta stone idea depends (Carter and Duax 2002). The validity of the Rodin–Ohno hypothesis for the proposed dual-strand origin of the two classes of aaRS therefore rests on the original evidence presented for a statistically significant sense–antisense relationship between the conserved aaRS motif sequences (Rodin and Ohno 1995) and on the

demonstration that an equal spacing between conserved motifs of Class 1 and 2 aminoacyl synthetases is compatible with function (Pham et al. 2007).

Supplementary Material

Supplementary table 1 is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by Science Foundation Ireland to M.A.F. and K.H.W. T.A.W. is supported by a grant from the Irish Council for Science Engineering and Technology.

Literature Cited

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics. 21:2104–2105.

Carter CW, Duax WL. 2002. Did tRNA synthetase classes arise on opposite strands of the same gene? Mol Cell. 10:705–708.

Clamp M, Cuff J, Searle SM, Barton GJ. 2004. The Jalview Java alignment editor. Bioinformatics. 20:426–427.

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 5:113.

Felsenstein J. 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics. 5:164–166.

Gaulin E, Madoui MA, Bottin A, Jacquet C, Mathe C, Couloux A, Wincker P, Dumas B. 2008. Transcriptome of *Aphanomyces euteiches*: new oomycete putative pathogenicity factors and metabolic pathways. PLoS ONE. 3:e1723.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52:696–704.

Gupta RS, Golding GB. 1993. Evolution of *HSP70* gene and its implications regarding relationships between archaebacteria, eubacteria, and eukaryotes. J Mol Evol. 37:573–582.

Huang X, Madan A. 1999. CAP3: a DNA sequence assembly program. Genome Res. 9:868–877.

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Munro HN, editor. Mammalian protein metabolism. New York: Academic Press. p. 21–132.

Kapoor M, Vijayaraghavan Y, Kadonaga R, LaRue KE. 1993. NAD(+)-specific glutamate dehydrogenase of *Neurospora crassa*: cloning, complete nucleotide sequence, and gene mapping. Biochem Cell Biol. 71:205–219.

Kersten MA, Muller Y, Baars JJ, Op den Camp HJ, van der Drift C, Van Griensven LJ, Visser J, Schaap PJ. 1999. NAD+-dependent glutamate dehydrogenase of the edible mushroom *Agaricus bisporus*: biochemical and molecular characterization. Mol Gen Genet. 261:452–462.

Konstantopoulou I, Ouzounis CA, Drosopoulou E, Yiangou M, Sideras P, Sander C, Scouras ZG. 1995. A *Drosophila hsp70* gene contains long, antiparallel, coupled open reading frames (LAC ORFs) conserved in homologous loci. J Mol Evol. 41:414–420.

LeJohn HB, Cameron LE, Yang B, MacBeath G, Barker DS, Williams SA. 1994. Cloning and analysis of a constitutive heat shock (cognate) protein 70 gene inducible by L-glutamine. J Biol Chem. 269:4513–4522.

LeJohn HB, Cameron LE, Yang B, Rennie SL. 1994. Molecular characterization of an NAD-specific glutamate dehydrogenase gene inducible by L-glutamine. Antisense gene pair arrangement with L-glutamine-inducible heat shock 70-like protein gene. J Biol Chem. 269:4523–4531.

Madoui MA, Gaulin E, Mathe C, San Clemente H, Couloux A, Wincker P, Dumas B. 2007. AphanoDB: a genomic resource for *Aphanomyces* pathogens. BMC Genomics. 8:471.

Merino E, Balbas P, Puente JL, Bolivar F. 1994. Antisense overlapping open reading frames in genes from bacteria to humans. Nucleic Acids Res. 22:1903–1908.

Mizuguchi K, Deane CM, Blundell TL, Overington JP. 1998. HOMSTRAD: a database of protein structure alignments for homologous families. Protein Sci. 7:2469–2471.

Monnerjahn C, Techel D, Mohamed SA, Rensing L. 2000. A non-stop antisense reading frame in the grp78 gene of *Neurospora crassa* is homologous to the *Achlya klebsiana* NAD-gdh gene but is not being transcribed. FEMS Microbiol Lett. 183:307–312.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol. 302:205–217.

Pham Y, Li L, Kim A, Erdogan O, Weinreb V, Butterfoss GL, Kuhlman B, Carter CW Jr. 2007. A minimal TrpRS catalytic domain supports sense/antisense ancestry of class I and II aminoacyl-tRNA synthetases. Mol Cell. 25:851–862.

Rodin SN, Ohno S. 1995. Two types of aminoacyl-tRNA synthetases could be originally encoded by complementary strands of the same nucleic acid. Orig Life Evol Biosph. 25:565–589.

Rother KI, Clay OK, Bourquin JP, Silke J, Schaffner W. 1997. Long non-stop reading frames on the antisense strand of heat shock protein 70 genes and prion protein (PrP) genes are conserved between species. Biol Chem. 378:1521–1530.

Shi J, Blundell TL, Mizuguchi K. 2001. FUGUE: sequence–structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J Mol Biol. 310:243–257.

Silke J. 1997. The majority of long non-stop reading frames on the antisense strand can be explained by biased codon usage. Gene. 194:143–155.

Yang B, LeJohn HB. 1994. NADP(+)-activable, NAD(+)-specific glutamate dehydrogenase. Purification and immunological analysis. J Biol Chem. 269:4506–4512.

Michele Vendruscolo, Associate Editor

Accepted November 15, 2008