Naturwissenschaften (2013) 100:459–467
DOI 10.1007/s00114-013-1045-2

ORIGINAL PAPER

3-to-1: unraveling structural transitions in ureases

Rodrigo Ligabue-Braun · Fábio Carrer Andreis · Hugo Verli · Célia Regina Carlini

Received: 5 February 2013 / Revised: 5 April 2013 / Accepted: 5 April 2013 / Published online: 26 April 2013

Abstract Ureases are nickel-dependent enzymes which catalyze the hydrolysis of urea to ammonia and carbamate. Despite the apparent wealth of data on ureases, many crucial aspects regarding these enzymes are still unknown, or constitute matter for ongoing debates. One of these is most certainly their structural organization: ureases from plants and fungi have a single unit, while bacterial and archaean ones have three-chained structures. However, the primitive state of these proteins — single- or three-chained — is yet unknown, despite many efforts in the field. Through phylogenetic inference using three different datasets and two different algorithms, we were able to observe chain number transitions displayed in a 3-to-1 fashion. Our results imply that the ancestral state for ureases is the three-chained organization, with single-chained ureases deriving from them. The two-chained variants are not evolutionary intermediates. A fusion process, different from those already studied, may explain this structural transition.

Keywords Urease · Phylogenetic tree · Structural transition · Evolution · Gene fusion · Gene disruption

Background

Ureases (urea amidohydrolases, EC 3.5.1.5) are found in plants, fungi and bacteria. These enzymes catalyze the hydrolysis of urea to ammonia and carbamate. The latter undergoes spontaneous hydrolysis to form carbonic acid and a second ammonia molecule (Mobley et al. 1995). Both urea and urease are hallmarks in the development of natural sciences (as reviewed by Krajewska 2009). After being discovered in human urine in

Communicated by: Sven Thatje

R. Ligabue-Braun and F.C. Andreis contributed equally to this work.

H. Verli and C.R. Carlini share senior authorship.

Electronic supplementary material The online version of this article (doi:10.1007/s00114-013-1045-2) contains supplementary material, which is available to authorized users.

R. Ligabue-Braun · H. Verli · C. R. Carlini
Graduate Program in Cellular and Molecular Biology, Center of Biotechnology, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

R. Ligabue-Braun
e-mail: rodrigobraun@cbiot.ufrgs.br

F. C. Andreis
Biotechnology Undergraduate Program, Center of Biotechnology, UFRGS, Porto Alegre, RS, Brazil
e-mail: fabio.andreis@ufrgs.br

H. Verli
Center of Biotechnology and Faculty of Pharmacy, UFRGS, Porto Alegre, RS, Brazil

C. R. Carlini
Center of Biotechnology and Department of Biophysics-IB, UFRGS, Porto Alegre, RS, Brazil

H. Verli (✉)
Av. Bento Gonçalves, 9500, Campus do Vale, Caixa postal 15005, 91501-970 Porto Alegre, RS, Brazil
e-mail: hverli@cbiot.ufrgs.br

C. R. Carlini (✉)
Av. Bento Gonçalves, 9500, Campus do Vale, Prédio 43431, 91500-970 Porto Alegre, RS, Brazil
e-mail: ccarlini@ufrgs.br

1773, urea became the first organic compound synthesized from inorganic materials, in 1828. The notion that urine-derived ammonia originates from urea dates back to 1798, but only in 1890 the ureolytic enzyme was isolated. Jack bean (*Canavalia ensiformis*) urease was the first enzyme ever to be crystallized (Sumner 1926), proving that enzymes were proteins and that they could be crystallized. Of equal importance, this enzyme was the first shown to posses nickel ions in its active site, which are essential for its activity (Dixon et al. 1975). Despite the apparent wealth of data on ureases, many crucial aspects regarding these enzymes are still unknown, or constitute matter of ongoing debates. Such topics include the very nature of ureolysis (Karplus et al. 1997; Benini et al. 1999; Estiu and Merz 2007) and the mechanisms underlying urease activation (Carter et al. 2009; Zambelli et al. 2011).

One of the most striking aspects of ureases is their structural organization (Fig. 1). Microbial ureases are composed by three or two chains, while plant and fungal ureases are composed by a single subunit. The amino acid sequences of the smaller subunits of microbial ureases are very similar to the corresponding regions in the single units of eukaryotic ureases (Krajewska 2009).

This high similarity observed among ureases from different kingdoms suggest that they all derive from a common ancestral protein, and are likely to have similar tertiary structures and catalytic mechanisms (Jabri et al. 1995; Mobley et al. 1995; Sirko and Brodzik 2000; Carlini and Polacco 2008). These observations, however, do not address a central topic in urease structure, i.e., what was the primitive structural state of these enzymes? To this point, two possibilities arise, as pointed out previously by Hausinger (1993): "Did the gene encoding the single-subunit plant enzyme undergo disruption to yield the multiple genes encoding the two or three bacterial subunits? Or did the bacterial genes fuse to form the gene encoding the plant subunit?" Our work intended to help answering these long-held questions which, despite many efforts in the field, remained unanswered. By means of large-scale phylogenetic analyses, we were able to track the structural transition from three to one subunit in ureases, revealing also that the two-chained variants are not evolutionary intermediates between them. Hypotheses for the genetic fusion in these enzymes are also presented.

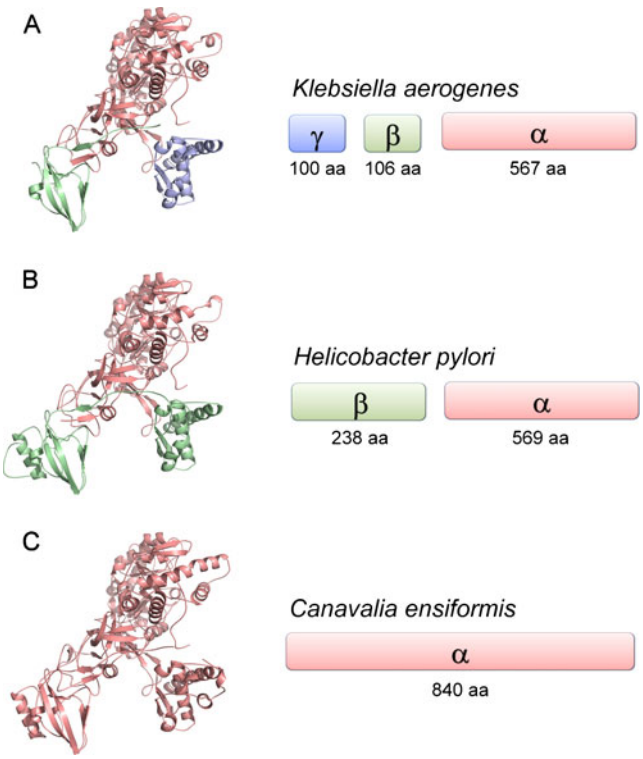

Fig. 1 Structural organization of ureases. Three-dimensional and schematic representations of the subunit organization of ureases. **a** Typical microbial urease composed by three chains (crystallographic structure from PDB ID 1FWJ). **b** Helicobacteraceae urease composed by two chains (crystallographic structure from PDB ID 1E9Z). **c** Typical eukaryotic urease composed by a single chain (crystallographic structure from PDB ID 3LA4)

Methods

All urease amino acid sequences (resulting from the "urease" keyword search) were retrieved from the National Center for Biotechnology Information (Sayers et al. 2012) on July 5, 2010. In order to obtain only complete urease sequences, manual filtering was carried out. Sequences that were incomplete, mistakenly labeled or related to urease accessory proteins were removed. All sequences related to Bacteria and Archaea in the resulting data were cross-compared to find similarities among them. The ClustalW algorithm (Larkin et al. 2007) was employed on all alignments, clustering together sequences with identity greater than 95 %. For practical reasons, these sequences were grouped by source organism genus. From the original 14,221 sequences data set, 32 sequences were from eukaryotic sources and did not require further filtering. Through sequence separation by genus, the number was reduced to 162 microbial representative sequences. These microbial ureases had their subunits' sequences aligned to ureases from plants and fungi, and these subunits were then joined together in a single sequence based on these alignments (forming a single $\gamma\beta\alpha$ sequence, or equivalent). The number of microbial sequences was further decreased to reduce the computational load. Based on the ClustalW alignment guide tree, highly similar sequences (80 % identical or higher) were considered as identical, and only one representative sequence was chosen. The final microbial data set was put together and aligned with the eukaryotic sequences, totaling 124 sequences in the final urease list (Table S1). Alignment sections with long gaps were removed.

To inspect the aligned urease sequences for highly variable regions, SimPlot software (Lole et al. 1999) was employed. To do so, the edited dataset was branched into three distinct subsets: complete urease sequences (without

long gaps), highly variable regions and conserved regions. Amino acid sequences evaluation by the ProtTest software (Abascal et al. 2005) and MEGA5 software (Tamura et al. 2011) suggested the Whelan and Goldman (WAG) model with discrete Gamma distribution to account for evolutionary rate differences among sites (four categories, +G), considering some sites to be evolutionarily invariable (+I) (Whelan and Goldman 2001). Maximum likelihood (ML) phylogenetic trees were calculated with MEGA5 with 1000 bootstrap replicates, and Bayesian inference (BI) trees with MrBayes software (Huelsenbeck and Ronquist 2001) for 7.5×10^6 generations, sampled every tenth generation. All trees were rooted in the branch containing the most archaean urease sequences. The obtained trees were visualized and edited with FigTree software (Rambaut 2012).

Results

Using ML and BI upon our alignments (Fig. S1), we were able to generate two phylogenetic trees for the complete urease sequences dataset (Figs. 2 and 3) and two for the conserved regions dataset (Figs. 4 and 5). For the highly variable regions dataset, one tree was generated through ML (Fig. S1), with very low bootstrap values. Convergence could not be reached throughout BI calculations, meaning that no considerable result could be obtained with these variable regions.

All trees displayed similar general convergence regarding their branching, with minor differences. We observed that, in the majority of cases, sequences belonging to organisms within the same phylum grouped together. All inferences displayed a similar distribution trend regarding ureases from five microbial groups composed of Euryarchaeota, Firmicutes, Actinobacteria, Proteobacteria and Cyanobacteria. Also, all phylogenetic inferences suggest that the number of urease chains evolved in a 3-to-1 fashion: three-chained ureases (those of most microbes) were of earlier existence, with a later structural unification originating single-chained ureases (those of plants and fungi). Two-chained ureases, belonging to Helicobacteraceae, are displayed as special situations among the three-chained enzymes, opposing the hypothesis considering them as intermediates between single- and three-chained ureases.

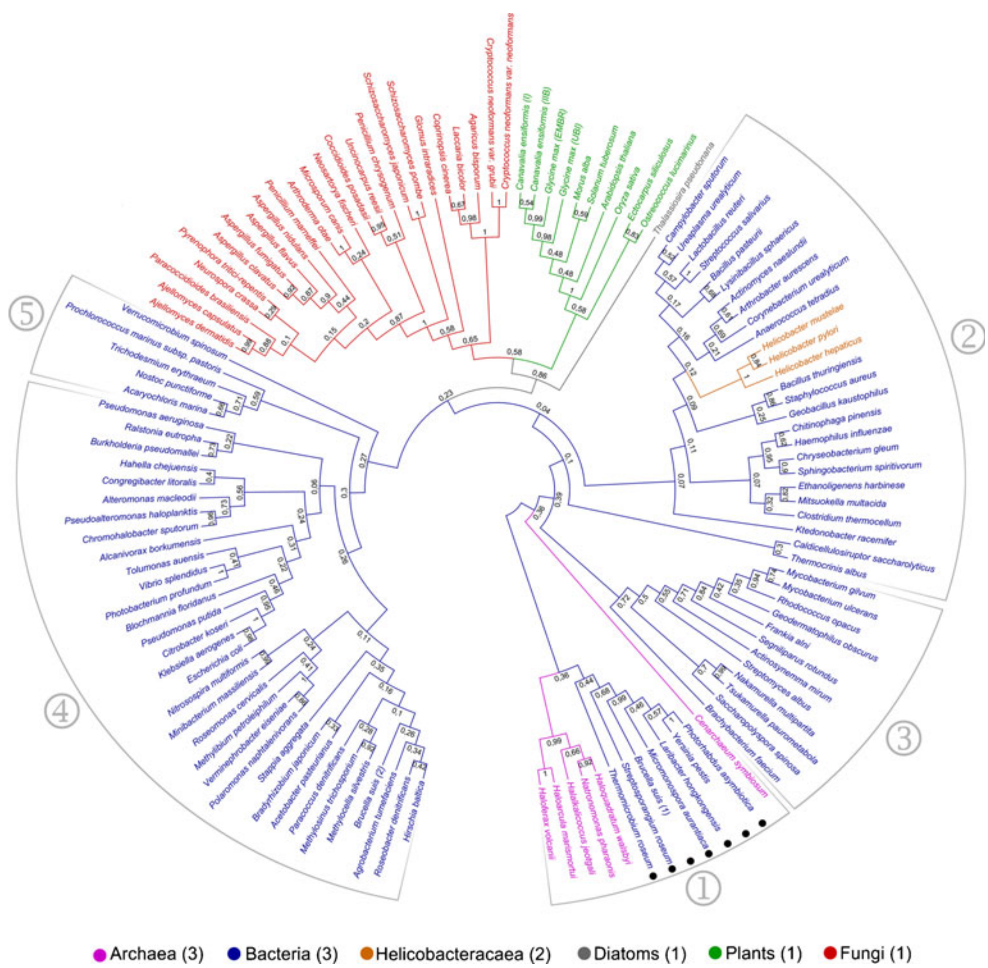

Fig. 2 Molecular phylogenetic analysis of complete urease sequences by maximum likelihood method. The evolutionary history was inferred by using the ML method based on the WAG+G+I model. The number of chains composing ureases from different groups is given in *brackets*. General microbial phyla separations are marked in *grey* (*1* Euryarchaeota, *2* Firmicutes, *3* Actinobacteria, *4* Proteobacteria, *5* Cyanobacteria). Grouping outliers are marked with *black dots*

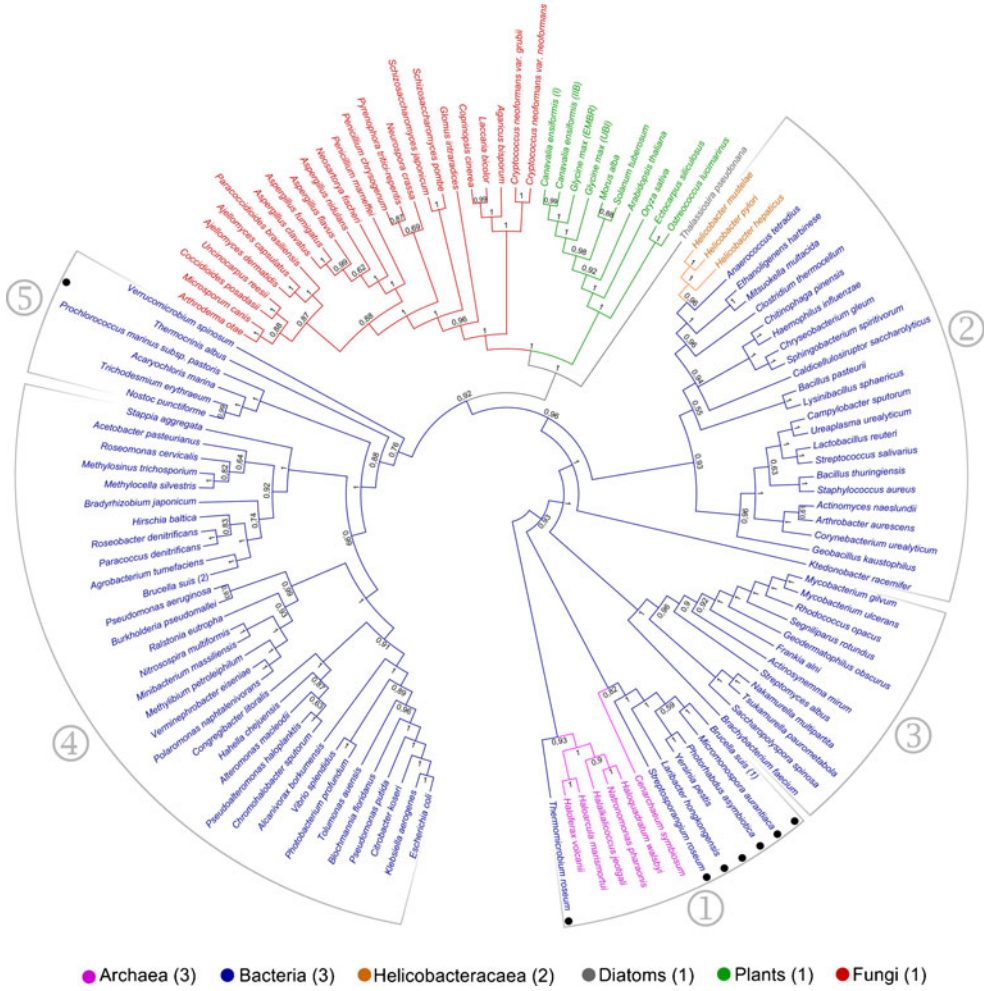

Fig. 3 Molecular phylogenetic analysis of complete urease sequences by Bayesian Inference method. The evolutionary history was inferred by using the Bayesian method based on the WAG+G+I model. The number of chains composing ureases from different groups is given in *brackets*. General microbial phyla separations are marked in *grey* (*1* Euryarchaeota, *2* Firmicutes, *3* Actinobacteria, *4* Proteobacteria, *5* Cyanobacteria). Grouping outliers are marked with *black dots*

Discussion

Urease has been considered an ancient enzyme, related to the putative primordial peptide cycle (Huber et al. 2003). This enzyme is found in the three domains of life, being synthesized in archaea, bacteria, fungi and plants (Carlini and Polacco 2008). There is uncertainty regarding urease presence in animals, and while some findings indicate its presence in some invertebrates (Pedrozo et al. 1996), others indicate that the enzyme comes from exterior sources (Hirayama et al. 2000) and that all animals lost urease in their evolutionary history (Fujiwara and Noguchi 1995). For this reason, ureases from putative animal sources were not included in our datasets.

Previous published attempts to establish the evolutionary history of ureases were based only in the catalytic subunit of microbial ureases (Contreras-Rodriguez et al. 2008), the full eukaryotic enzyme (Mulinari et al. 2011), or association of phylogenies at the organism level with enzyme characteristics (Navarathna et al. 2010). In this work we were able to reconstruct the possible evolutionary pathway followed by ureases in their structural transitions from multiple to single chains, supporting the "fusion hypothesis" (Hausinger 1993). It could be argued that this hypothesis was already supported by application of Ockham's razor (Gernert 2007), since the most abundant and primitive ureases appear to be three-chained, and therefore less likely to have all suffered fission. Only this observation, however, is too simplistic to grant validity to this assumption. The convergent results from four phylogenies, built based on two different methods upon two different datasets, nevertheless lend weight to the parsimony-derived conclusion. Additionally, these trees agree with previous analyses of the urease operon organization in fully sequenced microbial genomes (Zambelli et al. 2011). The clades as obtained in this work have *ure* operons with distinct structures: Clade 1 representatives (*H. marismortui*, *N. pharaonis*, *H. walsbyi*) are organized as *UreBCAGDEF*; Clade 2 (*C. thermocellum*, *H. influenzae*, *G. kaaustophilus*, *S. aureus*, *C. urealyticum*, *A. aurescens*, *L. sphaericus*, *U. urealyticum*) are organized as *UreABCEFGD* with the exception of *Helicobacter pylori* and *H. hepaticus* which are organized as *Ure(AB)C*-(unrelated gene)-*UreEFGD*; Clade 3 has only one analyzed representative (*M. gilvum*), organized as *UreABCGDEF*; Clade 4 has mixed

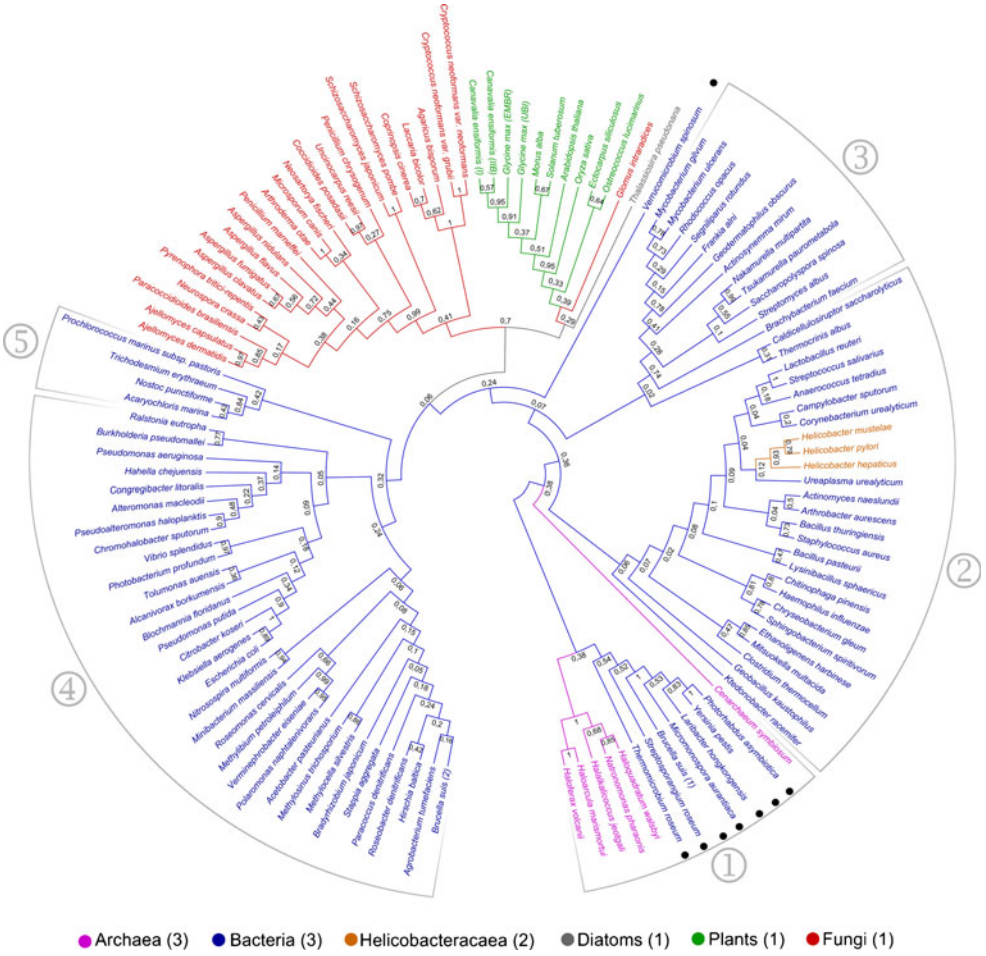

Fig. 4 Molecular phylogenetic analysis of conserved regions of urease sequences by Maximum Likelihood method. The evolutionary history was inferred by using the ML method based on the WAG+G+I model. The number of chains composing ureases from different groups is given in *brackets*. General microbial phyla separations are marked in *grey* (*1* Euryarchaeota, *2* Firmicutes, *3* Actinobacteria, *4* Proteobacteria, *5* Cyanobacteria). Grouping outliers are marked with *black dots*

assembles, including urease genes not organized in recognizable operons (*A. marina*); and Clade 5 (*P. aeruginosa*, *R. eutropha*, *B. pseudomallei*, *H. chejuensis*, *A. macleodii*, *P. haloplanktis*, *A. borkumensis*, *P. putida*, *E. coli*, *N. multiformis*, *P. naphtalivorans*, P. *denitrificans*, *M. silvestris*, *R. denitrificans*) is organized as either *UreDABCEFG* or *UreDA*-(unrelated gene)-*UreBCEFG*.

The enzyme dihydroorotase is considered the ancestral of all urease-related amidohydrolases (Holm and Sander 1997) and would be the ideal sequence to root the urease phylogenies. Dihydroorotase, however, is much shorter than urease and alignments including its sequence would interfere in the resulting trees. For this reason, a root had to be chosen among the clades in the urease trees. This rooting was performed by manually selecting the Archaean urease clade as a midpoint root, considering the intermediate phylogenetic position of Archaea in the tree of life (Woese et al. 1990), as a way of allowing comparisons among the different trees. When rooted by mathematical midpoint (Hess and Russo 2007), the trees showed no significant change in respect to those rooted in Archaean ureases (Supplementary Figs. S3, S4, S5 and S6). Inclusion in the alignments of regions that are highly variable among species was shown to be a source of "molecular noise" in the phylogenetic analyses. Attempts to obtain phylogenetic data from the highly variable regions were not statistically reliable, while removal of these regions had little impact on the obtained trees. These observations point to advantages in working with less variable regions of alignments, an approach employed for taxonomic phylogenies of fungi (Câmara et al. 2002), arthropods (Arango 2003; Hunt and Vogler 2008), cyanobacteria (Gaylarde et al. 2005), and viruses (Korber et al. 2001).

The occurrence of three-chained ureases in two of the three domains of life (i.e., Bacteria and Archaea) indicates that this structural organization is widespread and may be considered primitive in relation to one-chained ureases. The two-chained enzymes, which could have been taken as intermediates between single- and triple-chained ureases, seem to have arisen from bacterial triple chains in a process unrelated to the origin of single chained enzymes. Bacteria from the genus *Helicobacter* spp. are known to be subject to distinct selective pressures related to surviving in the gastric environment (Gueneau and Loiseaux-De Goër 2002) and the dodecameric macromolecular organization observed for *H. pylori* and *H. mustelae* ureases seem to be an adaptive response to such harsh conditions (Ha et al. 2001; Carter et

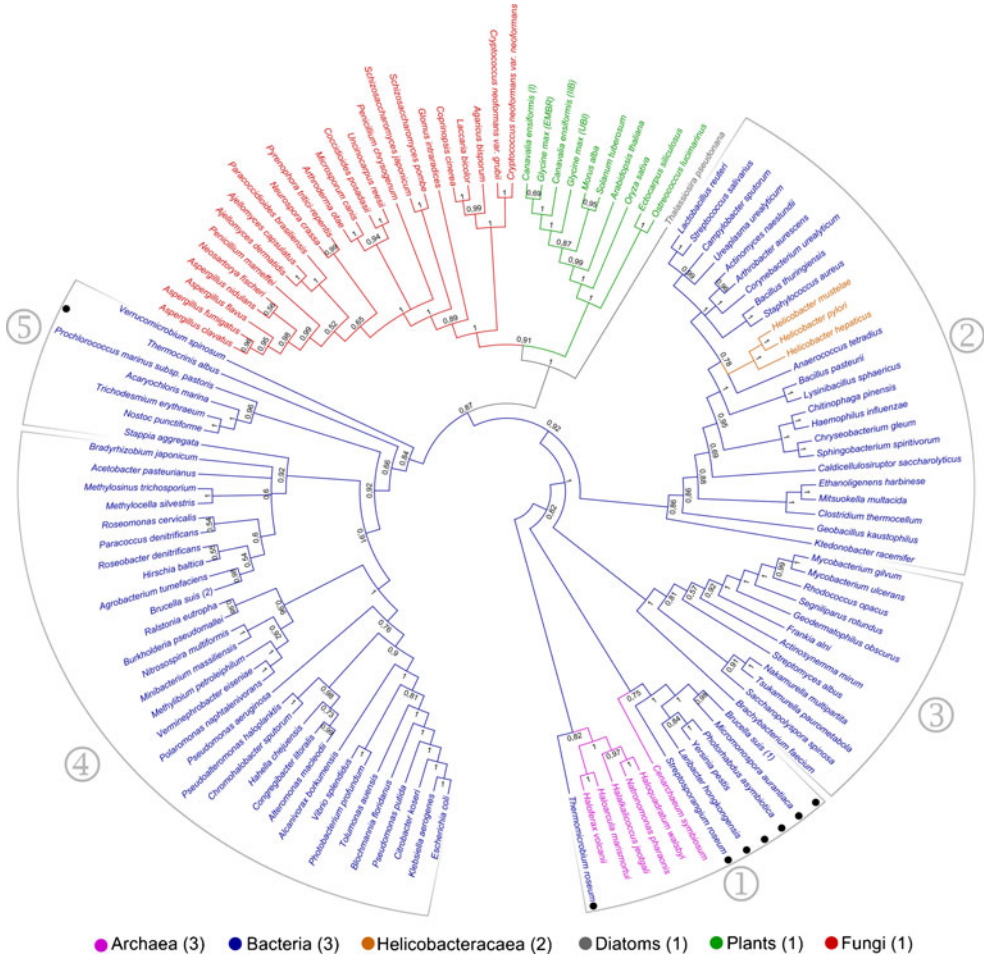

Fig. 5 Molecular phylogenetic analysis of conserved regions of urease sequences by Bayesian Inference method. The evolutionary history was inferred by using the Bayesian method based on the WAG+G+I model. The number of chains composing ureases from different groups is given in *brackets*. General microbial phyla separations are marked in *grey* (*1* Euryarchaeota, *2* Firmicutes, *3* Actinobacteria, *4* Proteobacteria, *5* Cyanobacteria). Grouping outliers are marked with *black dots*

al. 2011). The linking region (connecting the equivalents of subunits γ and β) of *Helicobacter* spp., which is highly divergent in relation to the same region in eukaryotes (Ha et al. 2001), could take part in this differential organization.

There are advantages associated with linking subunits. By artificial genetic fusion, it has been shown that some oligomeric viral proteins are benefited with enhanced folding rate and structural stability, and increased tolerance to insertions of other segments (Liang et al. 1993; Ma et al. 1993; Peabody 1997). When artificially fused, the genetic subunits of cytochrome ubiquinol oxidase from *E. coli* yielded an active enzyme, similar to the one from *Thermus thermophilus* (Ma et al. 1993). Differently from natural genetic fusions in *Drosophila* (Jones and Begun 2005), which resulted in new genes with new functions, the joining of urease subunits at the genetic level incorporating linking segments kept the original ureolytic activity. Ureases, however, are recognized as multifunctional (or moonlighting) proteins. They have many catalysis-independent effects, including neurotoxicity to mammals, insecticidal activities against Coleopterans and Hemipterans, fungistatic properties, pro-inflammatory roles, glycoconjugate binding properties, platelet activation ability, and inter-specific communication action (reviewed by Carlini and Polacco 2008; Stanisçuaski and Carlini 2012). Most of these properties have not yet been mapped to particular regions of these proteins, the exception being an "entomotoxic domain" containing the insectcidal peptide(s) released by *Canavalia ensiformis* ureases upon insect digestion (Ferreira-DaSilva et al. 2000; Piovesan et al. 2008; Defferrari et al. 2011). This "domain" is located in the β–α intersubunit region, and may be subject to faster divergence rates, since it is not involved in catalysis or subunit association (Mulinari et al. 2007). Also intriguing is the need for non-catalytic subunits/domains in urease, considering that only the TIM-barrel sub-domain from the α subunit is responsible for catalysis (Balasubramanian and Ponnuraj 2010). While the β subunit has been implicated in the urease activation process (Carter et al. 2011), no specific function was ascribed to the γ subunit.

Reviewing the literature we were not able to track other proteins that have undergone a similar process of natural subunit fusion. The events observed for ureases seem unparalleled, and not related to immunoglobulin genetic fusion (Tonegawa 1983) or exon shuffling (Kaessman 2010). The main difference is that for most of the studied microbial ureases, the γ, β, and α subunits are ordered and genetically

adjacent in the same operon (Zambelli et al. 2011), and not dispersed along the genome, as is the case for precursors of many merged genes. It is expected from the different mechanisms that may lead to chimeric genes that both inter and intragenic regions can be equally affected by segment insertions (Kaessman 2010). This does not seem to have occurred for ureases, since only intergenic regions were incorporated.

From the phylogenies, it is inferred that the transition from three subunits to one unit occurred as a single event. One mechanism that could be held responsible for such one-step result would be transcription or translation readthrough, which bypasses stop codons, incorporating intergenic regions as coding sequences. Translation bypass of stop codons has been well documented in yeast, where it takes part in complex regulatory mechanisms (von der Haar and Tuite 2007), while transcriptional readthrough has been implicated in human genetic disorders (Du et al. 2009), plant responses to stress (Hernández-Pinzón et al. 2009), and prostate adenocarcinoma, where many transcription-induced chimeras were found (Nacu et al. 2011). For ureases, however, readthrough events alone would not explain how the fused genes were finally incorporated into the genomes of their source organisms, requiring other subsequent process(s) at transcriptional and translational levels. We speculate that different genetic codes would be responsible for stop codons being unrecognized as such, allowing continuous transcription of the urease $\gamma\beta\alpha$ complex into a single chain. One such candidate would be the Chlorophyceae *Scenedesmus obliquus* mitochondrial code, which takes UCA as a stop codon instead of coding for serine, as occurs in the standard genetic code (Nedelcu et al. 2000). In this scenario, for intergenic codes to be translated it would be required that urease genes were transferred from the mitochondrial genome to the nuclear genome, where stop codons would not be recognized as such. There are, however, some difficulties with this hypothesis. No UCA codon is found in the terminal serine position of region γ from plants and fungi enzymes, and no clear serine position is found in the β chain terminus of the same ureases. Regarding the serine codons in the terminus of γ chain, it could be argued that serine would be beneficial in that position, thus allowing only conservative mutations of that originally misinterpreted codon. Regarding cellular location, urease is generally considered a cytoplasmic protein, but it has also been found in membrane fractions and cell wall from plants (Aguetoni Cambuí et al. 2009). Proteomic studies also indicate that *C. ensiformis* urease is either bound to mitochondria or spatially related to mitochondrial proteins (Demartini et al. 2011). Inter-specific horizontal gene transfer could also be responsible for the hypothetical readthrough. These transfers are now considered a major genome shaping tool when involving transposable elements (Schaack et al. 2010), and many instances of gene transfer from bacteria to eukaryotes have been documented, including organelle-to-nucleus transfers, such as the nuclear mitochondrial insert transferred to the *A. thaliana* chromosome 2 (Dunning Hotopp 2011). Little data is available on inter-specific urease genetic transfers and the only case reported so far is the second *ure* gene cluster from *Brucella suis* (Contreras-Rodriguez et al. 2008).

The genetic code readthrough hypothesis would also require intergenic regions of sufficient length to account for the incorporated segments. Urease operons from some bacteria related to Clade 2, such as *Staphylococcus saprophyticus*, *Streptococcus thermophilus*, and *Corynebacterium glutamicum* (Gene IDs 3616069, 3167116, and 1021080, respectively), have intercistronic regions that would satisfy this requirement. In other cases, exemplified by the *Proteus mirabilis* urease, there are not enough codons in these regions. On the contrary, there is even superposition of the last codon of the β subunit with the start codon of the α subunit (Jones and Mobley 1989). When eukaryotic gene organization is taken into account, the picture becomes even more complex. A preliminary inspection of annotated genomes at the Ensembl database (Flicek et al. 2012), revealed that all urease genes from plants deposited in there, i.e., *Arabidopsis thaliana*, *Brassica rapa*, *Glycine max* (both isoforms), *Oryza sativa*, *Physcomitrella patens*, *Populus trichocarpa*, *Sorghum bicolor*, and *Vitis vinifera*, have their coding sequences arranged in 18 exons. The same number of exons is observed for *Solanum tuberosum*, the first (and so far, only) case of alternative splicing in ureases (Witte et al. 2005). For fungal urease genes deposited at Ensembl, the number of exons is variable. It ranges from a single coding sequence with no introns (*Magnaporthe poae*, *Ustilago maydis*, *Schizosaccharomyces pombe*), to a variable number of exons: three (*Fusarium oxysporum*, *Gaeumannomyces graminis*, *Gibberella moniliformis*, *Magnaporthe oryzae*), four (*Aspergillus fumigatus*, *Neurospora crassa*), five (*A. fumigatus*, *A. niger*, *Nectria haematococca*, *Gibberella zeae*), 6 (*A. flavus*, *A. nidulans*), 7 (*Neosartorya fischeri*, *Phaeosphaeria nodorum*), 8 (*Aspergillus terreus*), 9 (*Fusarium oxysporum*), 13 (*Puccinia graministritici*), 14 (*Gibberella zeae*) or 15 (*Puccinia graministritici*). The difference in exon number between plants and fungi may be a reflection of strict structural conservation in plants. Until more genomes are deciphered, allowing a more in-depth analysis of their urease-coding segments, this interpretation remains speculative.

Conclusions

From the phylogenies presented in this work, we conclude that ureases were originally composed by three chains, and their transition to single-chained enzymes did not involve two-chained intermediates. We also speculate that the 3-to-1 transition took place as a single event, and hypothesize on a mechanism that would

result in the fused urease. Nonetheless, many questions remain unanswered. It is as if the unraveling of urease evolutionary paths (the "what") begins to be established, while the mechanisms underlying the urease structural transitions (the "how") still await further investigation. We expect that the large datasets and multiple approaches employed in this work contribute to enhance the comprehension of the unique case of urease subunits fusion, encouraging further research on the subject.

Acknowledgments The authors thank Charley C. Staats, Cláudia L. Fernandes, Dennis M. Junqueira, and Marilene H. Vainstein for many insightful suggestions and friendly help with the methodology. This work was supported by the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), MCT; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), MEC; Financiadora de Estudos e Projetos (FINEP); and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS).

References

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

Aguetoni Cambuí C, Gaspar M, Mercier H (2009) Detection of urease in the cell wall and membranes from leaf tissues of bromeliad species. Physiol Plant 136:86–93

Arango CP (2003) Molecular approach to the phylogenetics of sea spiders (Arthropoda: Pycnogonida) using partial sequences of nuclear ribosomal DNA. Mol Phylogenet Evol 28:588–600

Balasubramanian A, Ponnuraj K (2010) Crystal structure of the first plant urease from jack bean: 83 years of journey from its first crystal to molecular structure. J Mol Biol 400:274–283

Benini S, Rypniewski WR, Wilson KS, Miletti S, Ciurli S, Mangani S (1999) A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels. Structure 7:205–216

Câmara MP, Palm ME, van Berkum P, O'Neill NR (2002) Molecular phylogeny of Leptosphaeria and Phaeosphaeria. Mycologia 94:630–640

Carlini CR, Polacco JC (2008) Toxic properties of ureases. Crop Sci 48:1665–1672

Carter EL, Flugga N, Boer JL, Mulrooney SB, Hausinger RP (2009) Interplay of metal ions and urease. Metallomics 1:207–221

Carter EL, Boer JL, Farrugia MA, Flugga N, Towns CL, Hausinger RP (2011) Function of UreB in *Klebsiella aerogenes* urease. Biochemistry 50:9296–9308

Contreras-Rodriguez A, Quiroz-Limon J, Martins AM, Peralta H, Avila-Calderon E, Sriranganathan N, Boyle SM, Lopez-Merino A (2008) Enzymatic, immunological and phylogenetic characterization of *Brucella suis* urease. BMC Microbiol 8:121

Defferrari MS, Demartini DR, Marcelino TB, Pinto PM, Carlini CR (2011) Insecticidal effect of *Canavalia ensiformis* major urease on nymphs of the milkweed bug *Oncopeltus fasciatus* and characterization of digestive peptidases. Insect Biochem Mol Biol 41:388–399

Demartini DR, Carlini CR, Thelen JJ (2011) Global and targeted proteomics in developing jack bean (*Canavalia ensiformis*) seedlings: an investigation of urease isoforms mobilization in early stages of development. Plant Mol Biol 75:53–65

Dixon NE, Gazzola C, Blakeley RL, Zerner B (1975) Jack bean urease (EC 3.5.1.5). A metalloenzyme. A simple biological role for nickel? J Am Chem Soc 97:4131–4133

Du L, Damoiseaux R, Nahas S, Gao K, Hu H, Pollard JM, Goldstine J, Jung ME, Henning SM, Bertoni C, Gatti RA (2009) Nonaminoglycoside compounds induce readthrough of nonsense mutations. J Exp Med 206:2285–2297

Dunning Hotopp JC (2011) Horizontal gene transfer between bacteria and animals. Trends Genet 27:157–163

Estiu G, Merz KM Jr (2007) Competitive hydrolytic and elimination mechanisms in the urease catalyzed decomposition of urea. J Phys Chem B 111:10263–10274

Ferreira-DaSilva CT, Gombarovits ME, Masuda H, Oliveira CM, Carlini CR (2000) Proteolytic activation of canatoxin, a plant toxic protein, by insect cathepsin-like enzymes. Arch Insect Biochem Physiol 44:162–171

Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, Gordon L, Hendrix M, Hourlier T, Johnson N, Kähäri AK, Keefe D, Keenan S, Kinsella R, Komorowska M, Koscielny G, Kulesha E, Larsson P, Longden I, McLaren W, Muffato M, Overduin B, Pignatelli M, Pritchard B, Riat HS et al (2012) Ensembl 2012. Nucleic Acids Res 40:D84–D90

Fujiwara S, Noguchi T (1995) Degradation of purines: only ureidoglycollate lyase out of four allantoin-degrading enzymes is present in mammals. Biochem J 312:315–318

Gaylarde PM, Crispim CA, Neilan BA, Gaylarde CC (2005) Cyanobacteria from Brazilian building walls are distant relatives of aquatic genera. Omics 9:30–42

Gernert D (2007) Ockham's razor and its improper use. J Sci Explor 21:135–140

Gueneau P, Loiseaux-De Goër S (2002) Helicobacter: molecular phylogeny and the origin of gastric colonization in the genus. Infect. Genet Evol 1:215–223

Ha NC, Oh ST, Sung JY, Cha KA, Lee MH, Oh BH (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. Nat Struct Biol 8:505–509

Hausinger RP (1993) Urease. In: Hausinger RP (ed) Biochemistry of nickel. Plenum, New York, pp 23–57

Hernández-Pinzón I, de Jesús E, Santiago N, Casacuberta JM (2009) The frequent transcriptional readthrough of the tobacco Tnt1 retrotransposon and its possible implications for the control of resistance genes. J Mol Evol 68:269–278

Hess PN, Russo CAM (2007) An empirical test of the midpoint rooting method. Biol J Linn Soc 92:669–674

Hirayama C, Sugimura M, Saito H, Nakamura M (2000) Host plant urease in the hemolymph of the silkworm, *Bombyx mori*. J Insect Physiol 46:1415–1421

Holm L, Sander C (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. Proteins 28:72–82

Huber C, Eisenreich W, Hecht S, Wächtershäuser G (2003) A possible primordial peptide cycle. Science 301:938–940

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Hunt T, Vogler AP (2008) A protocol for large-scale rRNA sequence analysis: towards a detailed phylogeny of Coleoptera. Mol Phylogenet Evol 47:289–301

Jabri E, Carr MB, Hausinger RP, Karplus PA (1995) The crystal structure of urease from *Klebsiella aerogenes*. Science 268:998–1004

Jones CD, Begun DJ (2005) Parallel evolution of chimeric fusion genes. Proc Natl Acad Sci USA 102:11373–11378

Jones BD, Mobley HL (1989) *Proteus mirabilis* urease: nucleotide sequence determination and comparison with jack bean urease. J Bacteriol 171:6414–6422

Kaessman H (2010) Origins, evolution, and phylogenetic impact of new genes. Genome Res 20:1313–1326

Karplus PA, Pearson MA, Hausinger RP (1997) 70 years of crystalline urease: what have we learned? Acc Chem Res 30:330–337

Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V (2001) Evolutionary and immunological implications of contemporary HIV-1 variation. Br Med Bull 58:19–42

Krajewska B (2009) Ureases I. Functional, catalytic and kinetic properties: a review. J Mol Catal B Enzym 59:9–21

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. Bioinformatics 23:2947–2948

Liang H, Sandberg WS, Terwilliger TC (1993) Genetic fusion of subunits of a dimeric protein substantially enhances its stability and rate of folding. Proc Natl Acad Sci USA 90:7010–7014

Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 73:152–160

Ma J, Lemieux L, Gennis RB (1993) Genetic fusion of subunits I, II, and III of the cytochrome bo ubiquinol oxidase from *Escherichia coli* results in a fully assembled and active enzyme. Biochemistry 32:7692–7697

Mobley HL, Island MD, Hausinger RP (1995) Molecular biology of microbial ureases. Microbiol Rev 59:451–480

Mulinari F, Stanisçuaski F, Bertholdo-Vargas LR, Postal M, Oliveira-Neto OB, Rigden DJ, Grossi-de-Sá MF, Carlini CR (2007) Jaburetox-2Ec: an insecticidal peptide derived from an isoform of urease from the plant *Canavalia ensiformis*. Peptides 28:2042–2050

Mulinari F, Becker-Ritt AB, Demartini DR, Ligabue-Braun R, Stanisçuaski F, Verli H, Fragoso RR, Schroeder EK, Carlini CR, Grossi-de-Sá MF (2011) Characterization of JBURE-IIb isoform of *Canavalia ensiformis* (L.) DC urease. Biochim Biophys Acta 1814:1758–1768

Nacu S, Yuan W, Kan Z, Bhatt D, Rivers CS, Stinson J, Peters BA, Modrusan Z, Jung K, Seshagiri S, Wu TD (2011) Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. BMC Med Genomics 4:11

Navarathna DH, Harris SD, Roberts DD, Nickerson KW (2010) Evolutionary aspects of urea utilization by fungi. FEMS Yeast Res 10:209–213

Nedelcu AM, Lee RW, Lemieux C, Gray MW, Burger G (2000) The complete mitochondrial DNA sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. Genome Res 10:819–831

Peabody DS (1997) Subunit fusion confers tolerance to peptide insertions in a virus coat protein. Arch Biochem Biophys 347:85–92

Pedrozo HA, Schwartz Z, Luther M, Dean DD, Boyan BD, Wiederhold ML (1996) A mechanism of adaptation to hypergravity in the statocyst of *Aplysia californica*. Hear Res 102:51–62

Piovesan AR, Stanisçuaski F, Marco-Salvadori J, Real-Guerra R, Defferrari MS, Carlini CR (2008) Stage-specific gut proteinases of the cotton stainer bug *Dysdercus peruvianus*: role in the release of entomotoxic peptides from *Canavalia ensiformis* urease. Insect Biochem Mol Biol 38:1023–1032

Rambaut A (2012) FigTree [http://tree.bio.ed.ac.uk/software/figtree/]

Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Federhen S, Feolo M, Fingerman IM, Geer LY, Helmberg W, Kapustin Y, Krasnov S, Landsman D, Lipman DJ, Lu Z, Madden TL, Madej T, Maglott DR, Marchler-Bauer A, Miller V, Karsch-Mizrachi I, Ostell J, Panchenko A, Phan L, Pruitt KD, Schuler GD et al (2012) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 40:D13–D25

Schaack S, Gilbert C, Feschotte C (2010) Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. Trends Ecol Evol 25:537–546

Sirko A, Brodzik R (2000) Plant ureases: roles and regulation. Acta Bioch Pol 4:1189–1195

Stanisçuaski F, Carlini CR (2012) Plant ureases and related peptides: understanding their entomotoxic properties. Toxins 4:55–67

Sumner JB (1926) The isolation and crystallization of the enzyme urease. J Biol Chem 69:435–441

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302:575–581

von der Haar T, Tuite MF (2007) Regulated translational bypass of stop codons in yeast. Trends Microbiol 15:78–86

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691–699

Witte CP, Tiller S, Isidore E, Davies HV, Taylor MA (2005) Analysis of two alleles of the urease gene from potato: polymorphisms, expression, and extensive alternative splicing of the corresponding mRNA. J Exp Bot 56:91–99

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579

Zambelli B, Musiani F, Benini S, Ciurli S (2011) Chemistry of Ni^{2+} in urease: sensing, trafficking, and catalysis. Acc Chem Res 44:520–530