Convergent evolution and structural adaptation to the deep ocean in the eukaryotic chaperonin CCTα

Alexandra A.-T. Weber[1,2,3*], Andrew F. Hugall[1], Timothy D. O'Hara[1]

[1]Sciences, Museums Victoria, GPO Box 666, Melbourne VIC 3001, Australia

[2]Centre de Bretagne, REM/EEP, Ifremer, Laboratoire Environnement Profond, 29280 Plouzané, France

[3]Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

Email addresses: aweber@museum.vic.gov.au; ahugall@museum.vic.gov.au; tohara@museum.vic.gov.au

*Corresponding author: Alexandra A.-T. Weber

ORCID ID:

A.A.-T. Weber: https://orcid.org/0000-0002-7980-388X

T.D. O'Hara: https://orcid.org/0000-0003-0885-6578

Abstract

The deep ocean is the largest biome on Earth and yet it is among the least studied environments of our planet. Life at great depths requires several specific adaptations, however their molecular mechanisms remain understudied. We examined patterns of positive selection in 416 genes from four ophiuroid families (216 species) displaying independent events of deep-sea colonization. We found consistent signatures of molecular convergence in 5 genes, including the CCTα gene (Chaperonin Containing TCP-1 subunit α), which is essential for protein folding. CCTα protein stability profiles across the ophiuroid tree of life (725 species) revealed that depth-adapted proteins display higher stability within and next to the substrate-binding region, an expectation for high-pressure adapted proteins. As CCT has previously been categorized as a 'cold-shock' protein, we propose that adaptation mechanisms to cold and deep-sea environments may be linked and highlight that efficient protein folding is a key metabolic deep-sea adaptation.

The deep ocean (>200m) covers approximately two-thirds of the global sea floor area, yet it is among the least studied environments of our planet in terms of biodiversity, habitats and ecosystem functioning[1]. It harbors specific environmental conditions such as high pressure, low temperatures (0-4°C), absence of light and scarcity of food. Life at great depths requires multiple metabolic adaptations resulting in a physiological bottleneck[2], limiting the vertical distribution of species[3,4]. Enzymatic processes, protein folding, assembly of multi-subunit proteins and lipoprotein membranes are influenced by pressure and temperature at the cellular level[4–7]. Thus, as an adaptation to deep-sea environments, high-pressure adapted proteins (i.e. barophilic proteins) have been shown to be more stable (i.e. more resistant to denaturation) than their shallow-water counterparts[2,4,8,9]. However, this has been measured in only a handful of proteins and taxa[10–17]. Interestingly, patterns of protein adaptation to temperature show higher flexibility (i.e. decreased stability) with decreasing temperature[18,19]. As pressure and temperature strongly co-vary in the deep sea - temperature decreases as pressure increases - it therefore can be difficult to disentangle the respective combined or opposing effects of these factors on protein stability evolution.

Patterns of positive selection have been investigated to uncover genes underlying adaptation to specific environments, including the deep-sea, in non-model species[20–25]. Although valuable, these studies typically focused on a single or few shallow-deep transitions in a limited number of species, and thus lack the comparative power to separate confounding effects. With almost 2100 species, brittle stars (Ophiuroidea) are a large and ancient class of echinoderms[26,27]. These diverse marine invertebrates have colonized every marine habitat, highlighting their strong adaptive abilities. Furthermore, their phylogeny is well-resolved[28,29] and they represent a major component of the deep-sea fauna, making them important models for marine biogeography[30,31]. It is usually assumed that deep-sea organisms colonized the deep-sea from shallow waters; however, colonization from deep to shallow waters has also been reported[3,32]. Four large independent ophiuroid families (Amphiuridae, Ophiodermatidae, Ophiomyxidae and Ophiotrichidae) have a common ancestor from shallow water with extant species occurring in the deep-sea[32]. Due to these repeated and independent colonization events, these four brittle star families provide an ideal framework to test for convergent molecular evolution to the deep sea.

Results

Five genes involved in protein biogenesis are recurrently positively selected in deep-sea brittle stars

We used 416 single-copy orthologs from 216 species (288 individuals) of four brittle star families (Figure 1A) to examine patterns of positive selection in deep-sea species (>200m). For each gene of each family, we used four different positive selection methods and one method detecting relaxation of selection (Figure 1B-C). To minimize false positive detection, we only kept candidate genes that had significant signature of positive selection in at least three methods and which also did not show relaxation of selection. We found 36 candidate genes in Amphiuridae, 9 in Ophiodermatidae, 6 in Ophiomyxidae and none in Ophiotrichidae (Table S2; Figure S1). Five genes were positively selected in at least two families, among which one (CCTα) was selected in all three families and significant in each one of the selection detection methods (Table 1). To confirm that positive selection was detected only in shallow-deep transition, we performed positive selection analyses on the five common candidate genes but this time labelling shallow-water lineages as 'Foreground'. Most of the genes did not display signatures of positive selection in shallow-water environments, except PFD3 in Amphiuridae and tkt in Ophiotrichidae, which were significant in the BUSTED, aBSREL and MEME methods (Table S3). This suggests that these two genes are evolving faster not only in response to deep-sea adaptation but also in response to different environmental conditions.

Table 1: common positively selected candidate genes in three families and their characteristics (3 of 4 methods, not displaying relaxation of selection). *Positively selected in 4 of 4 methods, not displaying relaxation of selection. Bold: common Biological Process annotation.

Gene name	Description	Blast Reference sequence	GO terms: Biological Process	Positively selected in
CCTα *	chaperonin containing TCP1 complex subunit α	XP_780270.1	**protein folding**	Amphiuridae, Ophiodermatidae, Ophiomyxidae
PFD3	prefoldin subunit 3	XP_797937.1	macromolecular complex assembly; protein complex assembly; **protein folding**	Amphiuridae, Ophiodermatidae
tkt*	transketolase 2 isoform X2	NP_1229589.1	biological process	Amphiuridae, Ophiodermatidae
rpl34	subunit ribosomal	XP_797232.1	ribosome biogenesis; **translation**	Amphiuridae, Ophiomyxidae
rpl8	60S ribosomal L8	XP_796001.1	**Translation**	Amphiuridae, Ophiomyxidae

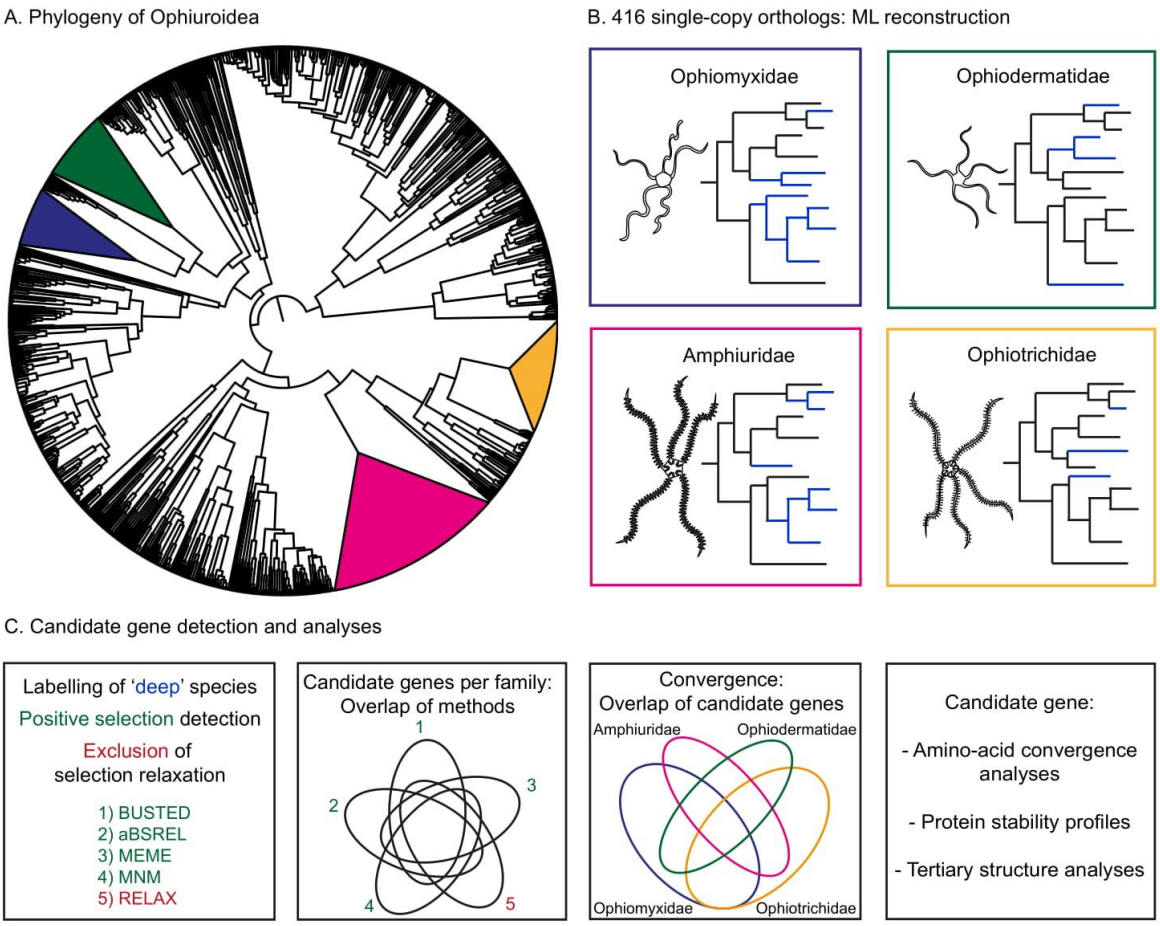

Figure 1: Workflow used in this study. A: Schematic representation of the phylogeny of Ophiuroidea (redrawn from[27]). Four families (288 individuals from 216 species) with a shallow-water common ancestor and extant species in shallow (0-200m) and deep (>200m) environments are highlighted in different colors. The width of each triangle is proportional to the number of species in each family. Ophiomyxidae (blue), Ophiodermatidae (green), Amphiuridae (pink) and Ophiotrichidae (yellow). B: For each family and each one of the 416 single-copy orthologs, Maximum Likelihood (ML) reconstructions were performed. C: For each resulting ML tree, deep (>200m) species were labeled as foreground branches (colored blue) and four positive selection detection methods were used (BUSTED, aBSREL, MEME, MNM). To detect and exclude candidate genes displaying relaxation of selection, i.e. accumulation of substitutions not due to increased selection pressure, the method RELAX was used. The final set of candidate genes for each family encompassed genes positively selected in at least 3 methods and not displaying relaxation of selection. Convergent evolution was examined by overlapping candidate genes per family. For the most interesting candidate gene, amino-acid convergence analyses, protein stability profiles and tertiary structure analyses were performed.

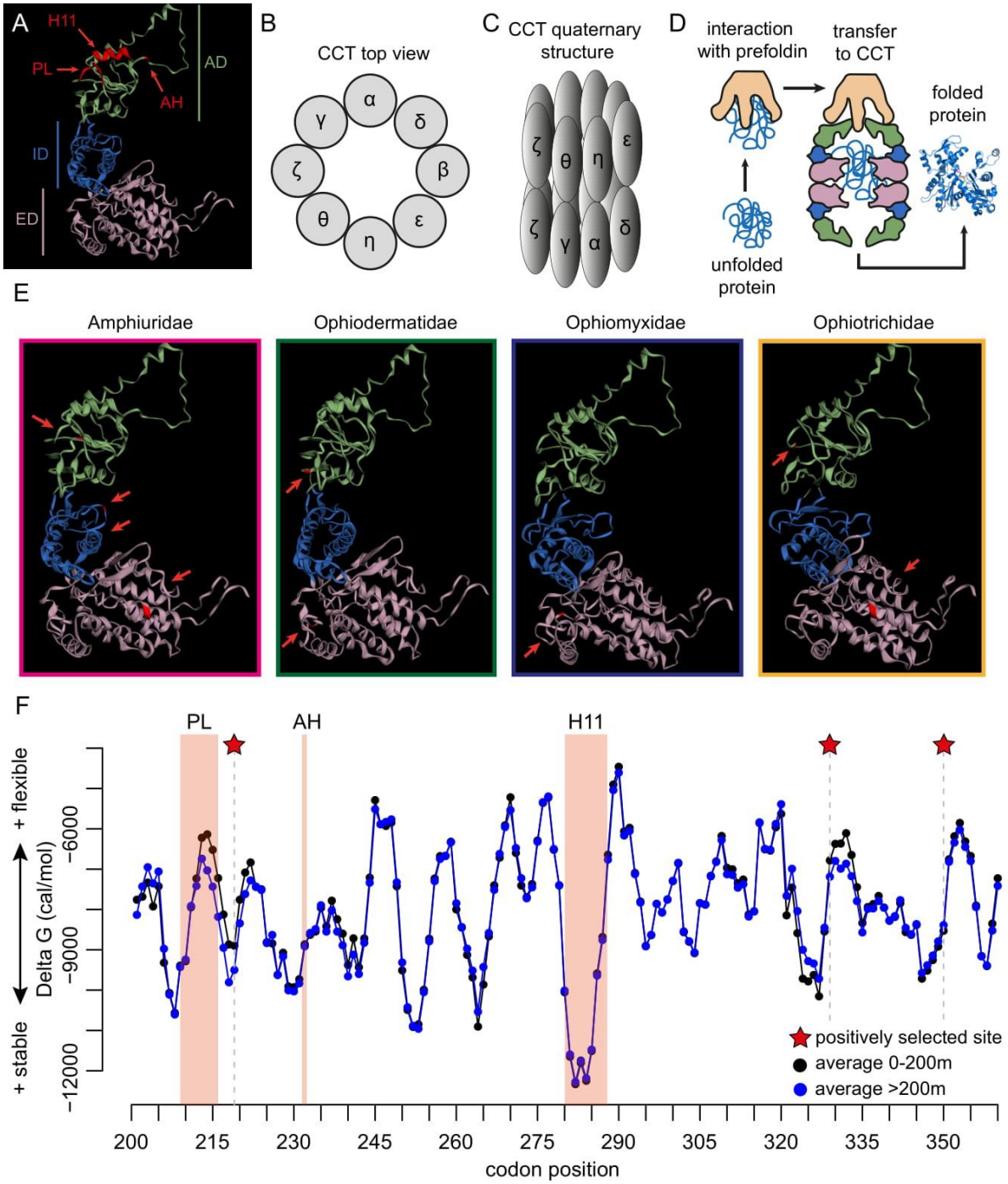

Figure 2: Structure and function of the CCT complex, selection analyses on CCTα and comparison of stability values from CCTα apical domain between shallow and deep species. A: Model of tertiary structure of the CCTα subunit. Each subunit is composed of an apical domain (AD; green) containing the substrate binding regions (PL: Proximal Loop; H11: Helix 11; AH: Apical Hinge), an intermediate domain (ID; blue) and an equatorial domain (ED; pink) containing the nucleotide binding site and where hydrolysis takes place. B: Model of the top view of the CCT complex, encompassing 8 paralogous subunits. C: Quaternary structure model of the CCT complex encompassing a double ring of 8 paralogous subunits. D: Simplified model of Prefoldin (PFD)-CCT interaction in the folding of newly synthesized actin or tubulin. A-D: Adapted from Bueno-Carrasco & Cuellar, 2018, "Mechanism and Function of the Eukaryotic

Chaperonin CCT". E: Localization of the positively selected sites on the tertiary structure of CCTα in the four ophiuroid families investigated. F: Average protein stability profiles for each codon of the CCTα apical domain in 324 species (424 individuals) from shallow water (0-200m) and 401 species (543 individuals) from deep water (>200m) representative of the whole ophiuroid class. A smaller (i.e. more negative) value of delta G is indicative of substitutions increasing stability. The substrate binding regions PL, AH and H11 are highlighted as well as the positively selected sites.

CCTα, the sole candidate gene detected as significantly positively selected in three families, is a subunit of the octameric Chaperonin Containing TCP1 (CCT) complex, a cytosolic eukaryotic chaperonin having a central role in protein folding (Figure 2A-D)[33,34]. CCT is estimated to fold ~10% of newly synthesized proteins, including actin and tubulin, and is involved in numerous core cellular processes such as cytoskeleton formation, cell signaling, cell recognition and protein degradation. Interestingly, PFD3, a subunit of the hexameric co-chaperone prefoldin interacting with CCT[35,36] was positively selected in two families (Amphiuridae: MNM, MEME, BUSTED; Ophiodermatidae: MNM, aBSREL, MEME, BUSTED) (Table 1; Tables S2, S6; Figure 2D), although it was also positively selected in shallow-water Amphiuridae (Table S3). The two other prefoldin subunits present in our dataset (PFD1 and PFD5) did not show consistent signature of positive selection (Table S4), suggesting that sub-units of this co-chaperone can evolve relatively independently from each other. Finally, two ribosomal proteins (Rpl8 and Rpl34) were positively selected in two families, suggesting that protein biogenesis (protein synthesis and folding) may have a central role in deep-sea adaptation.

CCTα and deep-sea adaptation: convergence at the gene but not amino-acid level

Although positive selection was detected at the gene level in three families, positive selection at the site level (MEME) was detected in all four families. Interestingly, the sites displaying positive selection in CCTα were not the same among the four families (Figure 2E; Table S5). Four sites were found in the equatorial domain, i.e. the ATP binding region, while three sites were found in the apical domain, i.e. the substrate binding region [33]. In addition, convergent evolution at the site level was not detected when examining amino-acid profiles (PCOC posterior probabilities at all sites < 0.9). Thus, convergent patterns of positive selection were detected at the pathway and gene levels but convergent evolution was not detected at the amino-acid level. It has been shown that rates of molecular convergence decrease with time[37] and the last common ancestor of Amphiuridae, Ophiodermatidae and Ophiomyxidae is estimated to be approximately 250 million years old[28]. Furthermore, convergence at the amino-acid level is often the least common compared to convergence at higher levels of biological hierarchy (e.g. gene, pathway or species levels)[38–40]. While we tested four subunits of the octameric CCT complex, CCTα was the only one to be detected as showing significant signal of

positive selection (Table S6). This might be due to the different degrees of subunit specialization, as CCTα has intermediate binding properties (i.e. neither high ATP affinity nor high substrate affinity) compared to the other subunits [33]. Thus, CCTα might be functionally less constrained to evolve rapidly. Interestingly, it was shown that CCTα, CCTγ and CCTζ evolved under positive selection after duplication events which led to sub-functionalization in eukaryotes, most likely in response to folding increasingly complex cytosolic proteins[41].

Energetic landscapes reveal structural adaptation within and next to the proximal loop binding region

Next we calculated site-specific protein stability profiles of CCTα in 967 individuals of 725 species representative of the whole Ophiuroidea class, to test the hypothesis that deep-sea adapted proteins are more stable than their shallow-water counterparts. For each site, we compared the average stability measure of 324 shallow-water species (0-200m depth) vs. 401 deep-water species (>200m depth), where lower delta G values correspond to higher stability (Figures 2F, S2A). We focused on the apical domain as it encompasses the substrate binding region, whose position and structure are highly conserved across eukaryotes [42]. This region is composed of the proximal loop (PL), the apical hinge (AH) and Helix 11 (H11) (Figure 2A). While AH and H11 are almost invariant across all ophiuroids, the stability measure was lower (i.e. more stable) in deep compared to shallow species within the PL and in two sites following the PL (codons 214-217), close to a positively selected site (codon 219) (Figures 2F, S2A). In contrast, three codons displayed significantly higher flexibility in deep compared to shallow species (Figure S2), suggesting that increased flexibility may play a role in deep-sea adaptation outside the ligand binding region, possibly related to low temperature. Nevertheless, when averaging delta G values across 10 codons, only the signal close to PL remained significant in the phylogenetically-corrected ANOVA contrasting stability values of shallow and deep species (Figure S3). This indicates that substitutions towards a more stable PL occurred independently in the ophiuroid tree of life. It has previously been reported that the shallow groove created by the conserved H11 and the flexible PL allows the binding of a variety of substrates[42,43]. Thus, our results suggest that substitutions in the PL and in adjacent amino acids allow efficient substrate binding in deep-sea species. Similarly, in a study on metabolic enzymes from 37 ctenophores, numerous sites associated with adaptation to depth, temperature or both were located close to the ligand binding region [44].

Discussion

We have shown that over deep evolutionary timescales, CCTα, a sub-unit of the most complex eukaryotic chaperonin CCT, displays recurrent signatures of accelerated evolution and

structural adaptation in transition from shallow to deep-sea habitats across 725 brittle star species. This was not the case in a study including nine sea urchin species, which found that only the CCTε subunit was positively selected, but not in the two deep-sea species investigated[20]. There are numerous studies on the role of CCT at shorter evolutionary timescales, which revealed its role in cold-stress response. Notably, CCT has been characterized as a 'cold-shock' protein in several eukaryotes due to the overexpression of the investigated subunits when organisms were exposed to cold stress[45–48]. Furthermore, CCT has been shown to display specific structural[49] and functional[50] adaptations to cold environment in Antarctic fish, in addition to being overexpressed in Antarctic fish exposed to heat stress[51]. There is also evidence for a link between cold-stress response and high-pressure stress response in bacteria[52,53]. Moreover, cold-inducible protein families are expanded in a hadal amphipod[54], and several proteins involved in cold shock have been shown to evolve under positive selection in deep-sea amphipod and fish[21]. Taken together, our findings support the hypothesis that cold shock proteins play an important role in deep-sea adaptation[3].

While we acknowledge that our study lacks functional validation to demonstrate that the changes are truly adaptive (which would be experimentally demanding as CCT folds ~10% of newly synthesized proteins), we minimized false inferences by applying stringent positive selection detection criteria. Furthermore, we used a proxy of functional validation by investigating *in silico* protein stability profiles in a dataset with great comparative power, both in terms of phylogenetic and environmental diversity. Finally, experimental testing on deep-sea organisms remains technically challenging, so we made use of the power of molecular data to reveal new insights in deep-sea adaptation. Further studies should include whole genomes to obtain a more complete view of deep-sea adaptation mechanisms[55]. Also, while we focused on intrinsic adaptations, mechanisms of extrinsic adaptations through osmolyte concentration should not be overlooked[56], but they were beyond the scope of this study. With increasing interests in deep-sea biodiversity, ecosystems and resources in the last decades[57–59], these are exciting times for diving deeper into mechanisms of deep-sea adaptation.

Methods

Phylogenomic data generation and processing

The data used in this study is an extension (436 additional samples) of a previously published exon-capture phylogenomic datamatrix of 1484 exons in 416 genes for 708 individual ophiuroid samples representative of the whole class Ophiuroidea[30]. The full dataset used here encompassed 1144 individual ophiuroid samples accounting for 826 species. Details on specimen collection, environmental parameters and list of species are available in Table S1. The

set of 416 single-copy genes were first determined in a transcriptome analysis[29] and the subsequent exon-capture system laboratory, bioinformatic and phylogenetic procedures are described in[28,60] and dryad packages https://doi.org/10.5061/dryad.db339/10 and https://datadryad.org/stash/dataset/doi:10.5061/dryad.rb334. Briefly, base-calling used a minimum read coverage of five. Exon boundaries were initially based on the *Strongylocentrotus purpuratus* and *Danio rerio* genomes, and then revised using the exon-capture read mapping information. For all selection analyses, codons immediately adjacent to exon boundaries were ignored. The primary data had IUPAC-coded heterozygous sites, which were then randomly resolved. However, these sites had little influence as both ambiguity-coded and randomly resolved datasets returned the same positive selection test results. A global phylogenetic tree of all 1144 samples for 416 genes (273kb sites) was generated via RAxML v.8. (Stamatakis 2014) using a codon position partition model. First a fully resolved all compatible consensus topology was generated from 200 RAxML fast bootstrap samples (the -f –d command), onto which branch lengths were then optimized using a codon position GTR-Γ model (the -f –e command). The tree was then rooted according to[28] defining the sister superorders Ophintegrida and Euryophiurida.

Four brittle star families were investigated that included species displaying independent events of deep-sea colonization from shallow-water [32]: Amphiuridae (111 individuals from 95 species, depth range: -0.5m to -5193m; temperature range: -1.6°C to 28.8°C), Ophiodermatidae (60 individuals from 38 species, depth range: -0.5m to -1668m; temperature range: 2.6°C to 28.3°C), Ophiomyxidae (41 individuals from 29 species, depth range: -1.5m to -792m; temperature range: 4.6°C to 28.7°C) and Ophiotrichidae (78 individuals from 62 species, depth range: -1m to -405m; temperature range: 10.5°C to 29.5°C). Positive selection analyses were conducted separately per family. 1664 alignments were generated, representing each gene (416) in each family (4). In each alignment, a maximum of 30% missing data per sequence was allowed. Alignments that lacked deep species (>200m) after filtering were not used. As all these four families belong to the superorder Ophintegrida, sequences of *Asteronyx loveni* belonging to the sister superorder Euryophiurida (Asteronychidae) were used as outgroups. After filtering, 1649 alignments were available for further analyses.

Phylogenetic reconstruction and positive selection analyses

For each of the 1649 alignments, a Maximum-Likelihood phylogeny was reconstructed using RAxML v.8.2.11 with the following parameters: -x 12345 -# 100 -f a -m GTRGAMMA -p 12345 (Figure 1B). Deep (>200m) species (tips) and monophyletic groups of deep species (nodes) were labeled as “Foreground” branches for positive selection analyses (Amphiuridae: 46 species; 11 independent events of deep-sea colonization; Ophiodermatidae: 7 species; 4 independent events; Ophiomyxidae: 10 species; 4 independent events; Ophiotrichidae: 4

species; 4 independent events). Then, the package HyPhy was used to conduct several positive selection analyses (Figure 1C): 1) BUSTED (Branch-site Unrestricted Statistical Test for Episodic Diversification) [61] to test for gene-wide positive selection (at least one site on at least one branch); 2) aBSREL (adaptive Branch-Site Random Effects Likelihood) [62] to detect specific branches evolving under episodic positive selection; 3) MEME (Mixed Effects Model of Evolution) [63] to find sites evolving under episodic positive selection. 4) MNM method; it has been shown recently that mutations at adjacent sites often occur as a result of the same mutational event (i.e. multinucleotide mutations, MNMs) and therefore may bias classical branch-site tests for positive selection [64]. The authors of that study developed a new model of positive selection detection incorporating MNMs, which we also used to detect positive selection on deep lineages (>200m). For each gene, p-values were corrected for multiple testing using the Holm method [65]. The p-value significance level used for all the positive selection detection methods was 0.05. Finally, we used: 5) RELAX [66] to test for relaxation of selection, and exclude potential candidate genes displaying relaxation of selection. For each of the four families, positively selected candidate genes of each method were overlapped on a Venn diagram (Figures 1C and S1). To be considered as a candidate gene for positive selection in one family and to minimize the risk of false positives, a gene had to display a significant signal in at least three out of four methods including MEME and MNM (BUSTED, MEME, MNM or MEME, MNM, aBSREL) and not display relaxation of selection (RELAX). This set of candidates was used for functional annotation. Final sets of positively selected genes per family were then compared among each other to test for convergent evolution. To confirm that positive selection was detected only in deep-sea lineages, positive selection was also tested in shallow-water lineages for the five genes displaying convergent positive selection signatures (CCTα, PFD3, tkt, rpl34, rpl8; see Results, Table 1). For each family, the same number of shallow-water species as was used for deep-water species was randomly labeled as 'Foreground' in each gene tree and positive selection tests were performed for the five methods as described above.

Gene Ontology annotations and amino-acid convergence analyses

To explore which functions may be involved in deep-sea adaptation, the representative sequence of each of the 416 genes was extracted from the sea urchin *Strongylocentrotus purpuratus* genome and blasted against the nr database from NCBI using BLAST+. We used *S. purpuratus* as reference because sequence annotation for this species is of high quality (no high-quality brittle star reference genome is currently available) and to use a single complete representative sequence for each gene. The top 50 hits were extracted and loaded in BLAST2GO v.4.1. for annotation [67]. Mapping, annotation and slim ontology (i.e. GO subsets of broader categories) were performed with BLAST2GO using default parameters, except for the

annotation cut-off parameter that was set to 45. GO categories were described using the level 3 of slim ontology.

CCTα, the only candidate gene displaying positive selection signal in three families (see Results) was further analyzed for signatures of convergent evolution. Specifically, amino-acid profiles were investigated for convergent shifts using PCOC[68]. This method, which has been shown to display high sensitivity and specificity, detects convergent shifts in amino-acid preferences rather than convergent substitutions. The CCTα amino-acid alignment encompassing the four families and outgroups was used to generate a maximum-likelihood phylogeny as previously described but this time using the PROTGAMMAWAG protein model of sequence evolution (Figure S4). For each family, positively selected branches resulting from aBSREL analyses were labeled as foreground branches (i.e. the branches with the convergent phenotype in the nucleotide topology) in four different scenarios: i) Amphiuridae, Ophiodermatidae, Ophiomyxidae; ii) Amphiuridae, Ophiodermatidae; iii) Amphiuridae, Ophiomyxidae; iv) Ophiodermatidae, Ophiomyxidae. Detection of amino-acid convergence in these four scenarios was then performed using PCOC and a detection threshold of 0.9 [68].

Protein structure modeling and protein stability profile

To infer the position of positively selected mutations on CCTα, the corresponding amino-acid sequence of the individual Amphiura_constricta_MVF214041 was used to obtain the secondary and tertiary protein structures of this gene. This species was chosen because its CCTα sequence had no missing data. The secondary structure was modeled using InterPro 72.0 web browser (https://www.ebi.ac.uk/interpro/). The protein model was generated using the normal mode of the online Phyre^2 server [69]. The online server EzMol 1.22 was used for image visualization and production [70].

We then examined the protein stability profiles of CCTα across the whole ophiuroid class (967 sequences with less than 30% missing sites, representing 725 species) using eScape v2.1 [71,72]. This algorithm calculates a per-site estimate of Gibbs free energy of stabilization based on a sliding window of 20 residues. More specifically, it models the contribution of each residue to the stability constant, a metric that represents the equilibrium of the natively folded and the multiple unfolded states of a protein [73]. Sites adapted to elevated pressure (or high temperature at atmospheric pressure) are expected to display stabilizing mutations (i.e. more negative delta G values), whereas sites adapted to low temperatures at atmospheric pressure are expected to display mutations increasing flexibility (i.e. decreasing stability, thus more positive delta G values) [18,19]. For each site of the apical domain (codons 200-361), we calculated the average delta G value for all 324 shallow-water species (424 individuals) (0-200m) and 401 deep-water

317 species (543 individuals) (>200m). To test the difference between these average values in a
318 phylogenetic context, we used phylogenetically-corrected ANOVA (R function phylANOVA of the
319 phytools v.0.6-60 R package; 10,000 simulations). To correct for relatedness among species, we
320 used the global RAxML phylogenetic tree pruned to the 967 tips. To investigate regions rather
321 than individual codons, we contrasted shallow vs. deep species along the whole gene, averaging
322 delta G values across 10 residues and performing a phylogenetically-corrected ANOVA as
323 previously described.

324 **References**

325 1. Ramirez-Llodra, E. Z. *et al.* Deep, diverse and definitely different: unique attributes of the world's
326 largest ecosystem. *Biogeosciences* **7**, 2851–2899 (2010).
327 2. Gross, M. & Jaenicke, R. Proteins under pressure: the influence of high hydrostatic pressure on
328 structure, function and assembly of proteins and protein complexes. *Eur. J. Biochem.* **221**, 617–630
329 (1994).
330 3. Brown, A. & Thatje, S. Explaining bathymetric diversity patterns in marine benthic invertebrates and
331 demersal fishes: physiological contributions to adaptation of life at depth. *Biol. Rev.* **89**, 406–426
332 (2014).
333 4. Somero, G. N. Adaptations to high hydrostatic pressure. *Annu. Rev. Physiol.* **54**, 557–577 (1992).
334 5. Carney, R. S. Zonation of deep biota on continental margins. in *Oceanography and Marine Biology*
335 221–288 (CRC Press, 2005).
336 6. Jaenicke, R. Protein stability and molecular adaptation to extreme conditions. in *EJB Reviews 1991*
337 291–304 (Springer, 1991).
338 7. Pradillon, F. & Gaill, F. Adaptation to deep-sea hydrothermal vents: some molecular and
339 developmental aspects. *J. Mar. Sci. Technol.* 37–53 (2007).
340 8. Siebenaller, J. F. Effects of the deep-sea environment on invertebrates. *Comp. High Press. Biol.* 319–
341 341 (2010).
342 9. Somero, G. N. Life at low volume change: hydrostatic pressure as a selective factor in the aquatic
343 environment. *Am. Zool.* **30**, 123–135 (1990).
344 10. DAHLHOFF, E. & SOMERO, G. N. Pressure and temperature adaptation of cytosolic malate
345 dehydrogenases of shallowand deep-living marine invertebrates: evidence for high body
346 temperatures in hydrothermal vent animals. *J. Exp. Biol.* **159**, 473–487 (1991).
347 11. Lemaire, B. *et al.* Molecular adaptation to high pressure in cytochrome P450 1A and aryl hydrocarbon
348 receptor systems of the deep-sea fish Coryphaenoides armatus. *Biochim. Biophys. Acta BBA-Proteins*
349 *Proteomics* **1866**, 155–165 (2018).
350 12. Morita, T. Comparative sequence analysis of myosin heavy chain proteins from congeneric shallow-
351 and deep-living rattail fish (genus Coryphaenoides). *J. Exp. Biol.* **211**, 1362–1367 (2008).
352 13. Morita, T. Structure-based analysis of high pressure adaptation of α-actin. *J. Biol. Chem.* **278**, 28060–
353 28066 (2003).
354 14. Siebenaller, J. F. & Somero, G. N. Pressure-adaptive differences in the binding and catalytic properties
355 of muscle-type (M 4) lactate dehydrogenases of shallow-and deep-living marine fishes. *J. Comp.*
356 *Physiol.* **129**, 295–300 (1979).
357 15. Siebenaller, J. & Somero, G. N. Pressure-adaptive differences in lactate dehydrogenases of congeneric
358 fishes living at different depths. *Science* **201**, 255–257 (1978).
359 16. Suka, A. *et al.* Stability of cytochromes c′ from psychrophilic and piezophilic Shewanella species:
360 implications for complex multiple adaptation to low temperature and high hydrostatic pressure.
361 *Extremophiles* **23**, 239–248 (2019).
362 17. Wakai, N., Takemura, K., Morita, T. & Kitao, A. Mechanism of deep-sea fish α-actin pressure tolerance
363 investigated by molecular dynamics simulations. *PloS One* **9**, e85852 (2014).
364 18. Fields, P. A., Dong, Y., Meng, X. & Somero, G. N. Adaptations of protein structure and function to
365 temperature: there is more than one way to 'skin a cat'. *J. Exp. Biol.* **218**, 1801–1811 (2015).
366 19. Saarman, N. P., Kober, K. M., Simison, W. B. & Pogson, G. H. Sequence-based analysis of thermal
367 adaptation and protein energy landscapes in an invasive blue mussel (Mytilus galloprovincialis).
368 *Genome Biol. Evol.* **9**, 2739–2751 (2017).

20. Kober, K. M. & Pogson, G. H. Genome-wide signals of positive selection in strongylocentrotid sea urchins. *BMC Genomics* **18**, 555 (2017).
21. Lan, Y. *et al.* De novo transcriptome assembly and positive selection analysis of an individual deep-sea fish. *BMC Genomics* **19**, 394 (2018).
22. Oliver, T. A. *et al.* Whole-genome positive selection and habitat-driven evolution in a shallow and a deep-sea urchin. *Genome Biol. Evol.* **2**, 800–814 (2010).
23. Sun, J. *et al.* Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nat. Ecol. Evol.* **1**, 0121 (2017).
24. Zhang, Y. *et al.* Adaptation and evolution of deep-sea scale worms (Annelida: Polynoidae): insights from transcriptome comparison with a shallow-water species. *Sci. Rep.* **7**, 46205 (2017).
25. Weber, A. A.-T. *et al.* Positive selection on sperm ion channels in a brooding brittle star: consequence of life-history traits evolution. *Mol. Ecol.* **26**, 3744–3759 (2017).
26. Stöhr, S., O'Hara, T. D. & Thuy, B. World Ophiuroidea Database. (2017).
27. Stöhr, S., O'Hara, T. D. & Thuy, B. Global diversity of brittle stars (Echinodermata: Ophiuroidea). *PLoS ONE* **7**, e31940 (2012).
28. O'Hara, T. D., Hugall, A. F., Thuy, B., Stöhr, S. & Martynov, A. V. Restructuring higher taxonomy using broad-scale phylogenomics: The living Ophiuroidea. *Mol. Phylogenet. Evol.* **107**, 415–430 (2017).
29. O'Hara, T. D., Hugall, A. F., Thuy, B. & Moussalli, A. Phylogenomic resolution of the class Ophiuroidea unlocks a global microfossil record. *Curr. Biol.* **24**, 1874–1879 (2014).
30. O'Hara, T. D., Hugall, A. F., Woolley, S. N., Bribiesca-Contreras, G. & Bax, N. J. Contrasting processes drive ophiuroid phylodiversity across shallow and deep seafloors. *Nature* **565**, 636 (2019).
31. Woolley, S. N. *et al.* Deep-sea diversity patterns are shaped by energy availability. *Nature* **533**, 393–410 (2016).
32. Bribiesca-Contreras, G., Verbruggen, H., Hugall, A. F. & O'Hara, T. D. The importance of offshore origination revealed through ophiuroid phylogenomics. in *Proc. R. Soc. B* vol. 284 20170160 (The Royal Society, 2017).
33. Bueno-Carrasco, M. T. & Cuéllar, J. Mechanism and Function of the Eukaryotic Chaperonin CCT. *ELS John Wiley Sons Ltd Ed* 1–9 (2019) doi:10.1002/9780470015902.a0028208.
34. Valpuesta, J. M., Carrascosa, J. L. & Willison, K. R. Structure and function of the cytosolic chaperonin CCT. *Protein Fold. Handb.* 725–755 (2005).
35. Gestaut, D. *et al.* The Chaperonin TRiC/CCT Associates with Prefoldin through a Conserved Electrostatic Interface Essential for Cellular Proteostasis. *Cell* **177**, 751–765 (2019).
36. Martín-Benito, J. *et al.* Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. *EMBO J.* **21**, 6377–6386 (2002).
37. Storz, J. F. Causes of molecular convergence and parallelism in protein evolution. *Nat. Rev. Genet.* **17**, 239–250 (2016).
38. Bolnick, D. I., Barrett, R. D., Oke, K. B., Rennison, D. J. & Stuart, Y. E. (Non) parallel evolution. *Annu. Rev. Ecol. Evol. Syst.* **49**, 303–330 (2018).
39. Tenaillon, O. *et al.* Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature* **536**, 165 (2016).
40. Tenaillon, O. *et al.* The molecular diversity of adaptive convergence. *Science* **335**, 457–461 (2012).
41. Fares, M. A. & Wolfe, K. H. Positive selection and subfunctionalization of duplicated CCT chaperonin subunits. *Mol. Biol. Evol.* **20**, 1588–1597 (2003).
42. Joachimiak, L. A., Walzthoeni, T., Liu, C. W., Aebersold, R. & Frydman, J. The structural basis of substrate recognition by the eukaryotic chaperonin TRiC/CCT. *Cell* **159**, 1042–1055 (2014).
43. Yam, A. Y. *et al.* Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat. Struct. Mol. Biol.* **15**, 1255 (2008).
44. Winnikoff, J. R., Francis, W. R., Thuesen, E. V. & Haddock, S. H. D. Combing transcriptomes for secrets of deep-sea survival: Environmental diversity drives patterns of protein evolution. *Integr. Comp. Biol.* (2019).
45. He, Y., Wang, L., Zhu, W., Dong, Z. & Liu, N. Effects of salinity on cold tolerance of Malaysian red tilapia. *Aquac. Int.* **25**, 777–792 (2017).
46. Kayukawa, T. *et al.* Expression of mRNA for the t; chcomplex polypeptide–1, a subunit of chaperonin CCT, is upregulated in association with increased cold hardiness in Delia antiqua. *Cell Stress Chaperones* **10**, 204 (2005).
47. Somer, L., Shmulman, O., Dror, T., Hashmueli, S. & Kashi, Y. The eukaryote chaperonin CCT is a cold shock protein in Saccharomyces cerevisiae. *Cell Stress Chaperones* **7**, 47 (2002).
48. Yin, Q. *et al.* Molecular cloning of Litopenaeus vannamei TCP-1-eta gene and analysis on its relationship with cold tolerance. *Yi Chuan Hered.* **33**, 168–174 (2011).

49. Pucciarelli, S., Parker, S. K., Detrich, H. W. & Melki, R. Characterization of the cytoplasmic chaperonin containing TCP-1 from the Antarctic fish Notothenia coriiceps. *Extremophiles* **10**, 537–549 (2006).
50. Cuellar, J. *et al.* Assisted protein folding at low temperature: evolutionary adaptation of the Antarctic fish chaperonin CCT and its client proteins. *Biol. Open* **3**, 261–270 (2014).
51. Buckley, B. A. & Somero, G. N. cDNA microarray analysis reveals the capacity of the cold-adapted Antarctic fish Trematomus bernacchii to alter gene expression in response to heat stress. *Polar Biol.* **32**, 403–415 (2009).
52. Welch, T. J., Farewell, A., Neidhardt, F. C. & Bartlett, D. H. Stress response of Escherichia coli to elevated hydrostatic pressure. *J. Bacteriol.* **175**, 7170–7177 (1993).
53. Wemekamp-Kamphuis, H. H., Karatzas, A. K., Wouters, J. A. & Abee, T. Enhanced levels of cold shock proteins in Listeria monocytogenes LO28 upon exposure to low temperature and high hydrostatic pressure. *Appl Env. Microbiol* **68**, 456–463 (2002).
54. Lan, Y. *et al.* Molecular adaptation in the world's deepest-living animal: Insights from transcriptome sequencing of the hadal amphipod Hirondellea gigas. *Mol. Ecol.* **26**, 3732–3743 (2017).
55. Gaither, M. R. *et al.* Genomics of habitat choice and adaptive evolution in a deep-sea fish. *Nat. Ecol. Evol.* **2**, 680 (2018).
56. Yancey, P. H. & Siebenaller, J. F. Co-evolution of proteins and solutions: protein adaptation versus cytoprotective micromolecules and their roles in marine organisms. *J. Exp. Biol.* **218**, 1880–1896 (2015).
57. Danovaro, R., Corinaldesi, C., Dell'Anno, A. & Snelgrove, P. V. The deep-sea under global change. *Curr. Biol.* **27**, R461–R465 (2017).
58. Danovaro, R., Snelgrove, P. V. & Tyler, P. Challenging the paradigms of deep-sea ecology. *Trends Ecol. Evol.* **29**, 465–475 (2014).
59. Glover, A. G., Wiklund, H., Chen, C. & Dahlgren, T. G. Point of View: Managing a sustainable deep-sea 'blue economy'requires knowledge of what actually lives there. *eLife* **7**, e41319 (2018).
60. Hugall, A. F., O'Hara, T. D., Hunjan, S., Nilsen, R. & Moussalli, A. An exon-capture system for the entire class Ophiuroidea. *Mol. Biol. Evol.* **33**, 281–294 (2015).
61. Murrell, B. *et al.* Gene-wide identification of episodic selection. *Mol. Biol. Evol.* **32**, 1365–1371 (2015).
62. Smith, M. D. *et al.* Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
63. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764 (2012).
64. Venkat, A., Hahn, M. W. & Thornton, J. W. Multinucleotide mutations cause false inferences of lineage-specific positive selection. *Nat. Ecol. Evol.* **2**, 1280 (2018).
65. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 65–70 (1979).
66. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **32**, 820–832 (2014).
67. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
68. Rey, C., Guéguen, L., Sémon, M. & Boussau, B. Accurate detection of convergent amino-acid evolution with PCOC. *Mol. Biol. Evol.* **35**, 2296–2306 (2018).
69. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845 (2015).
70. Reynolds, C. R., Islam, S. A. & Sternberg, M. J. EzMol: A web server wizard for the rapid visualization and image production of protein and nucleic acid structures. *J. Mol. Biol.* **430**, 2244–2248 (2018).
71. Gu, J. & Hilser, V. J. Sequence-based analysis of protein energy landscapes reveals nonuniform thermal adaptation within the proteome. *Mol. Biol. Evol.* **26**, 2217–2227 (2009).
72. Gu, J. & Hilser, V. J. Predicting the energetics of conformational fluctuations in proteins from sequence: a strategy for profiling the proteome. *Structure* **16**, 1627–1637 (2008).
73. D'Aquino, J. A. *et al.* The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct. Funct. Bioinforma.* **25**, 143–156 (1996).

Acknowledgements

We are grateful to W. Salzburger for providing access to the HPC sciCORE cluster and to J. Sarrazin for comments on a previous version of the manuscript. Calculations were performed at sciCORE (http://scicore.unibas.ch/) scientific computing center at the University of Basel,

Switzerland and at DATARMOR (http://www.ifremer.fr/pcdm) scientific computing center at the Pôle de Calcul et de Données Marines (PCDM), Ifremer, Brest, France. AATW was supported by an Endeavour Postdoctoral Fellowship awarded by the Australian Department of Education and Training (Grant Agreement No 6534_2018) and a Marie Sklodowska-Curie Global Fellowship awarded by the European Union's Horizon 2020 research and innovation programme (Grant Agreement No 797326; Project DEEPADAPT).

Author contributions

TOH collected the samples. TOH and AFH designed and generated the exon-capture data. AFH processed the raw data to generate the phylogenomic dataset. AATW designed the present study, performed positive selection, convergence and stability analyses. AATW drafted the manuscript and it was finalized with input from all co-authors.

Competing interests

None declared

Data availability

Phylogenomic data (including raw reads) and scripts for dataset generation are available in NCBI Bioproject PRJNA311384 and dryad packages: https://doi.org/10.5061/dryad.db339/10 and https://datadryad.org/stash/dataset/doi:10.5061/dryad.rb334. There are no restrictions on data availability.

Code availability

All custom computer codes are available upon request. There are no restrictions on their use.

Supplementary figure captions

Figure S1: Overlap of positively selected candidate genes among three brittle star families. A: Number of candidate genes per family positively selected in at least three positive selection detection methods and not displaying relaxation of selection. B: Number of candidate genes per family positively selected in all four positive selection detection methods and not displaying relaxation of selection. In both conditions of A and B, no gene was positively selected in the family Ophiotrichidae.

Figure S2: Comparison of stability values from CCTα apical domain between shallow and deep species. A: Log transformed p-values of the phylogenetically-corrected ANOVA performed between average delta G values of shallow vs. deep species at each codon of the CCTα apical domain. The substrate binding regions PL, AH and H11 are highlighted in light orange. The positively selected sites are highlighted with a red star. P-value level corresponding to 0.05 is highlighted in red. The most significant codons in a "significance peak" (204, 214, 265 and 324) are highlighted. B: Beanplots of delta G values

between shallow (0-200m) and deep (>200m) species for each one of the most significant codons in the phylogenetically-corrected ANOVA. Horizontal bar represents the average value of the dataset.

Figure S3: Comparison of stability values over 10 codon windows on the complete CCTα gene between shallow and deep species. A: Average protein stability profiles over 10 codon windows for the complete CCTα gene in 324 species (424 individuals) from shallow water (0-200m) and 401 species (543 individuals) from deep water (>200m) representative of the whole ophiuroid class. A smaller (i.e. more negative) value of delta G is indicative of substitutions increasing stability. The substrate binding regions PL, AH and H11 are highlighted as well as the positively selected sites. B: Log transformed p-values of the phylogenetically-corrected ANOVA performed between average delta G values over 10 codon windows of shallow vs. deep species. The positively selected sites are highlighted with a red star. P-value level corresponding to 0.05 is highlighted in red. C: Beanplots of delta G values between shallow (0-200m) and deep (>200m) species for both of the most significant 10 codon windows in the phylogenetically-corrected ANOVA. Horizontal bar represents the average value of the dataset.

Figure S4: Maximum-likelihood reconstruction of CCTα using amino-acid sequences. The four focal families (Amphiuridae, Ophiotrichae, Ophiomyxidae and Ophiodermatidae) and the outgroup (Asteronychidae) are labelled. Positively selected lineages (aBSREL method) of each family are highlighted in red.

Supplementary tables (in separate excel file)

Table S1: List of species used in this study, GPS coordinates and environmental parameters at their sampling locations. Empty cells indicate missing data.

Table S2: Positively selected candidate genes per family and their Gene Ontology annotation (P: Biological Process; F: Molecular Function; C: Cellular Component). Genes positively selected in several families are highlighted in bold.

Table S3: Results of positive selection tests in shallow-water lineages for the five candidate genes for deep-sea adaptation. Significance level: 0.05. Significant results are in bold. NS: not significant

Table S4: Results of positive selection tests for the three prefoldin subunits for each family. Significance level: 0.05. Significant results are in bold. NS: not significant

Table S5: Sites displaying episodic positive selection in CCTα. Method used: MEME. Significance level: 0.05. Significant sites are in bold.

Table S6: Results of positive selection tests for the four CCT subunits for each family. Significant results are in bold. Significance level: 0.05. NS: not significant