# BMC Evolutionary Biology

Research article

# Molecular evolution of UCP1 and the evolutionary history of mammalian non-shivering thermogenesis

David A Hughes*[1], Martin Jastroch[2], Mark Stoneking[1] and Martin Klingenspor[3]

Address: [1]Max-Planck-Institute for Evolutionary Anthropology, Department of Evolutionary Genetics, Leipzig, Germany, [2]Philipps-Universität Marburg, Faculty of Biology, Department of Animal Physiology, Marburg, Germany and [3]Molecular Nutritional Medicine, Technische Universität München, Else Kröner-Fresenius Center, Freising-Weihenstephan, Germany

Email: David A Hughes* - hughes@eva.mpg.de; Martin Jastroch - jastroch@staff.uni-marburg.de; Mark Stoneking - stoneking@eva.mpg.de; Martin Klingenspor - martin.klingenspor@wzw.tum.de

* Corresponding author

## Abstract

**Background:** Uncoupling protein 1 (UCP1) is a mitochondrial anion carrier, expressed in brown adipose tissue (BAT) of Eutherians. UCP1 is responsible for uncoupling mitochondrial proton transport from the production of ATP, thereby dissipating heat; it is essential for non-shivering thermogenesis (NST) in mammalian BAT. UCP1 orthologs have been identified in non-Eutherian mammals, fish and amphibians. Yet, UCP1 has a unique function in Eutherians in that it is necessary in the production of heat (NST). As such, this study aims to determine the evolutionary mode of UCP1 in Eutherians, where there is clear evidence of UCP1-dependent NST in BAT.

**Results:** Models of adaptive evolution through phylogenetic analysis of amino acid sequences by maximum likelihood were implemented to determine the mode of UCP1 protein evolution in Eutherians. An increase in the rate of amino acid substitutions on the branch leading to Eutherians is observed, but is best explained by relaxed constraints, not positive selection. Further, evidence for branch and site heterogeneity in selection pressures, as well as divergent selection pressures between UCP1 and its paralogs (UCP2-3) is observed.

**Conclusion:** We propose that the unique thermogenic function of UCP1 in Eutherians may be best explained by neutral processes. Along with other evidence, this suggests that the primary biochemical properties of UCP1 may not differ between Eutherians and non-Eutherians.

## Background

Uncoupling protein 1 (UCP1) is a mitochondrial protein carrier which, until recently, was thought to be found only in endothermic placental (Eutherian) mammals [1,2]. In Eutherians, UCP1 is the only gene known to be exclusively expressed in brown adipose tissue (BAT), accounting for up to 5% of the total mitochondrial protein in BAT [3]; UCP1 (also known as thermogenin) provides Euthe-

rians, particularly small mammals, hibernators and newborns, with a unique mechanism of non-shivering thermogenesis (NST) [4]. UCP1-dependent NST is probably a feature of most Eutherian mammals, as it has been found recently in the rock elephant shrew, a member of the Afrotherian mammalian lineage which separated early during the evolution of the Eutherians [5]. NST is produced by increasing the proton conductance in the inner

membrane of brown adipocyte mitochondria. This increased proton conductance uncouples mitochondrial respiration from ATP synthesis and thereby dissipates the proton motive force as heat [6-9]. It is the high oxidative capacity of mitochondria in BAT and the cellular composition of BAT that allows heat dissipation rates at a power of 300 – 400 W/kg [10-12]. It is these properties of BAT, their mitochondria, and the uncoupling activity of UCP1 which provide Eutherians with NST.

UCP1 is a member of the UCP gene family of mitochondrial transport proteins, which contains two closely related paralogs, UCP2 and UCP3. Although UCP2/3 have been intensively investigated, the molecule(s) that these proteins transport and their overall function remain unclear [6,8,13-16]. It is clear that UCP1 in fish and mammals are orthologs, forming a monophyletic clade, and that the tissue specificity of UCP1 differs between fish, where it is expressed in liver, kidneys and brain, and mammals, where it is expressed exclusively in adipose tissue [2]. Although the function of UCP1 in fish is uncertain, it has been shown that UCP1-like uncoupling activity in fish liver mitochondria is mediated by fatty acids and that it is inhibited by purine nucleotides, analogous to its mediation in Eutherians [17]. Interestingly, while UCP1 expression is increased in response to cold in Eutherians, it is down regulated in the liver of fish in response to cold [2].

In this study we investigate how the evolution of UCP1 may have been influenced by the acquisition of BAT-limited expression and the novel role of NST in Eutherians. Previous work has demonstrated an elevated rate of evolution, i.e. substitution rate, on the UCP1 Eutherian lineage [18]. Yet, what remains uncertain is the mode of evolution which acted upon this lineage. Identification of lineage-specific directional selection would indeed be suggestive of the acquisition of a new function by UCP1. In contrast, the lack of support for lineage-specific directional selection would not rule out this possibility, but may be more suggestive of Eutherian and non-Eutherian UCP1 functional similarities and thus an evolution of NST through neutral processes. We implement a series of phylogenetic models by maximum likelihood to estimate selection pressures at non-synonymous or replacement sites and at synonymous sites [19]. Contrasting the rate of evolution at these two sites (dN/dS or ω) provides a means of evaluating the selection pressures which have acted on the protein during the evolution of UCP1. Our goal is to determine the mode of selection occurring on the Eutherian lineage, where novel biochemical function may have evolved in the course of UCP1 becoming a protein obligated to BAT-limited expression and NST.

## Results

Available UCP1, UCP2, and UCP3 sequences were retrieved from 39 species ranging from teleost fishes to mammals (Additional file 1). Nucleotide and protein distance NJ trees, the species tree and a maximum likelihood (ML) tree were generated for the UCP1 locus. No one tree was significantly better than any other as indicated by the RELL method [20] and S-H test [21]. To evaluate the robustness of the results we chose to run all analyses with both the species tree (Figure 1) and the ML tree. As the results were largely concordant between trees and codon substitution models (data not shown), the results using the species tree and the F3X4 codon substitution model are presented in the main text, unless specified otherwise. All analyses for the UCP2 and UCP3 loci used the species tree. Furthermore, all analyses critical for our conclusions were repeated using the alignments from a second alignment software package, namely T-Coffee. Again, this was performed to insure the robustness of our results, as different methods may produce different alignments that then influence the outcome of all subsequent analyses. All results were highly congruent between the two alignment methods; we therefore present parameter estimations derived from alignments using MUSCLE, and present only those derived from T-Coffee which would lead to different conclusions.

The objective here is to determine the mode of selection that has acted on the UCP1 coding sequence in Eutherians, where it is certain that UCP1 is involved in the dissipation of heat essential for non-shivering thermogenesis. We tested for directional selection on the Eutherian branch (Figure 1), and for divergent selection pressure between the Eutherian UCP1 clade and its paralogs UCP2 and UCP3 (Figure 2). This was done by comparing the ω value among lineages and among sites across the UCP1 phylogeny, in order to determine the mode of selection and the magnitude of change in selective constraint (if any) acting across the UCP1 protein as well as within particular functional domains of UCP1.

### Branch and site ω rate variation

The one-ratio model (M0), which constrains ω to be equal across all lineages of the phylogeny, was used to evaluate the global ω value across the UCP1 phylogeny. The UCP1 species phylogeny (Figure 1) was used in this analysis. Global ω was estimated to be 0.124, indicative of strong purifying selection. Omega estimates for UCP2 and UCP3 phylogenies, also using species trees appropriate for the UCP2 and UCP3 data, were 0.072 and 0.083, respectively. Allowing ω to vary among branches across the UCP1 phylogeny (M1 model) results in a significantly improved fit to the data (likelihood ratio test (LRT), $2\Delta\ell = 217$, p < 0.001, df = 44). Thus, there is lineage-specific heterogeneity in ω across the UCP1 phylogeny.
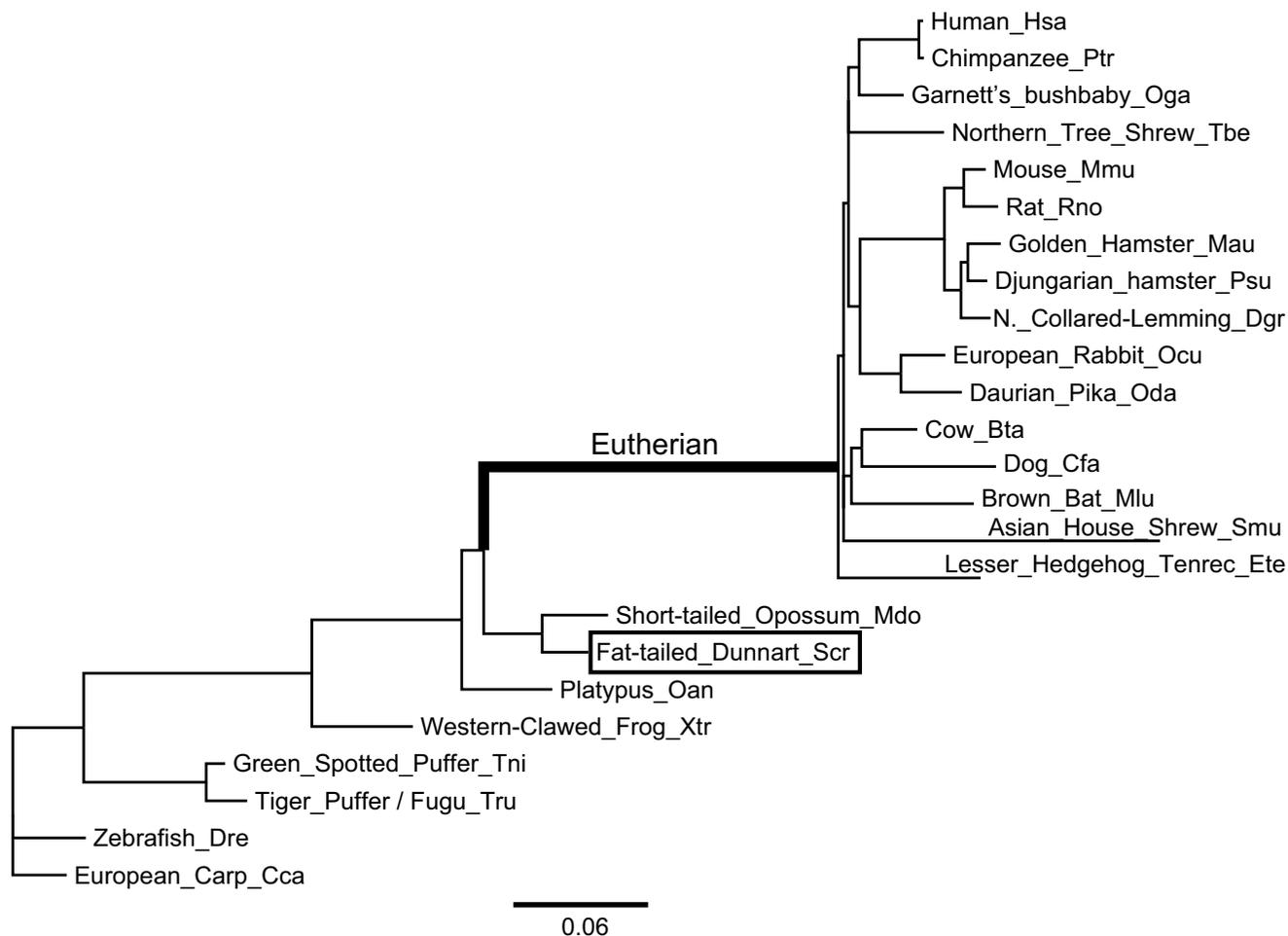
**Figure 1**
**Branch lengths are estimates of the number of substitutions per codon as performed in the M2 model with the Eutherian branch set as the foreground branch**. Each lineage is defined with the common name of each species and the three letter species code.

Along with lineage heterogeneity, variation in ω across sites can also occur. Theoretically, different protein domains with different functions may experience different selection pressures, which can be tested by fitting the data to a model comprising different site classes. The M3(discrete) model partitions the sites among three site categories, based on a discrete distribution. Each site category has an estimated proportion ($p_0$, $p_1$, $p_2$) and their ω values ($\omega_0$, $\omega_1$, $\omega_2$) are estimated from the data. This model results in a significantly better fit to the data than the null M0 model (LRT: $2\Delta\ell = 468$, df = 4, p-value < 0.001). This observation is also true for the UCP2 and UCP3 loci (Table 1). The ω estimates vary among sites for each gene; however, in general the fraction of sites with low ω estimates, indicative of strong selective constraints, is reduced for UCP1 relative to UCP2 and UCP3. Only 46% of UCP1 sites are under strong purifying selection

(the lowest ω category), compared to 60–61% for UCP2 and UCP3 (Table 1). This is consistent with the previous observation that the M0 ω estimate is higher for UCP1 (0.124) than for UCP2 (0.072) or UCP3 (0.083).

***Relaxed constraint on the Eutherian lineage***
To determine if the Eutherian UCP1 lineage (Figure 1) has evolved at a different rate, as compared to the rest of the phylogeny, a two ratio model (M2) was implemented. In this model there are two branch categories, and thus two independent estimates of ω: category 0, for the background branches, i.e. all lineages in the phylogeny excluding the Eutherian lineage; and category 1 for the foreground branch, i.e. the Eutherian lineage. In this model the estimated $\omega_1$ is 0.582 for the Eutherian branch, and 0.113 for the background branches. A LR test indicates that the M2 model provides a significantly better fit
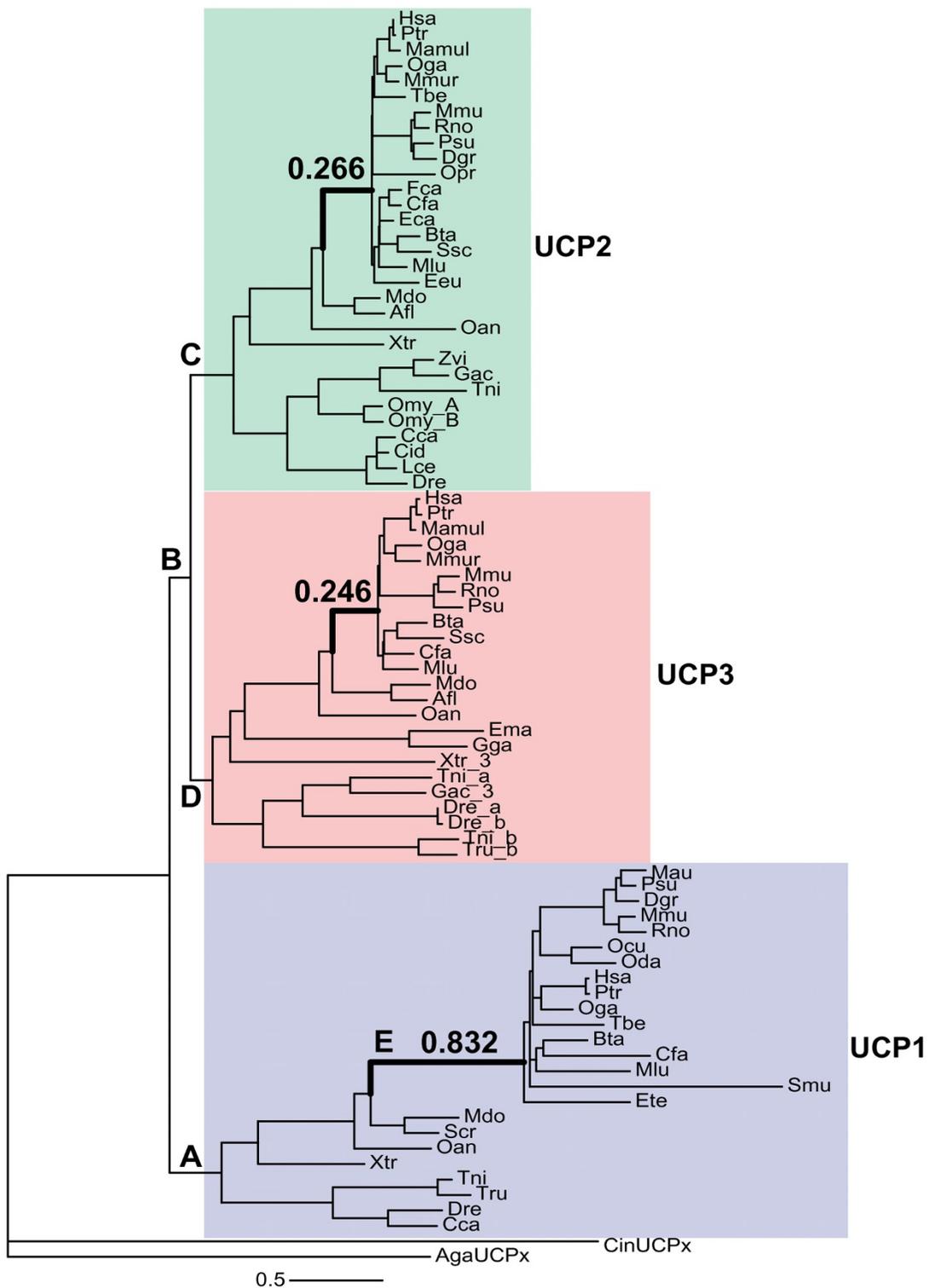
**Figure 2**
**UCP species tree**. Thick branches are Eutherian lineages for each respective gene. The values above each Eutherian branch are branch length estimates from the M0 model. Three letter species codes are given at the tip of each lineage. Species names can be found in Supplementary Table 1.

**Table 1: M3(discrete) sites model for the UCP1, UCP2 andUCP3 phylogenies**

|  | $p_0$ | $\omega_0$ | $p_1$ | $\omega_1$ | $p_2$ | $\omega_2$ | p-value |
|---|---|---|---|---|---|---|---|
| **UCP1** | 0.46 | 0.01 | 0.43 | 0.18 | 0.11 | 0.48 | <0.001 |
| **UCP2** | 0.61 | 0.004 | 0.31 | 0.12 | 0.07 | 0.65 | <0.001 |
| **UCP3** | 0.6 | 0.01 | 0.3 | 0.14 | 0.11 | 0.5 | <0.001 |

$p_0$, $p_1$, and $p_2$ are the estimated proportion of sites, and $\omega_0$, $\omega_1$, and $\omega_2$ are the estimated $\omega$ values, for site class 0 (strong purifying selection), site class 1 (purifying selection), and site class 2 (weak purifying selection), respectively.

to the data that the M0 model ($2\Delta L = 47.14$, $p < 0.001$, df = 1). Thus, there is a significantly higher $\omega$ value for UCP1 on the Eutherian lineage relative to the rest of the phylogeny. However, this estimate is not indicative of positive selection, it is rather indicative of relaxed constraint. By contrast, the UCP2 and UCP3 $\omega$ estimates for the Eutherian lineage do not differ from the rest of the phylogeny, as the LR tests do no indicate a better fit for the M2 model over the M0 model for these two genes (UCP2: p = 0.11; UCP3: p = 0.95). This contrast between the Eutherian UCP1 lineage, vs. the Eutherian UCP2 and UCP3 lineages, is further illustrated by the contrast in branch lengths (number of substitutions per codon) in the UCP phylogeny (Figure 2).

In principle, the increased $\omega$ estimate on the Eutherian UCP1 lineage could be the result of either an increase in dN or a decrease in dS. To determine the best explanation for this observation we compared the observed dN and dS values for the Eutherian lineage to the distribution of these estimates across the UCP1 phylogeny, given the estimates from the M2 model. The Eutherian UCP1 lineage dS is 0.27, which falls in the 70th percentile of dS values for all lineages in the phylogeny, whereas dN for the Eutherian lineage is 0.16, which is the largest dN value in the UCP1 phylogeny. Hence, it is dN, rather than dS, that appears responsible for the increased $\omega$ estimate on the Eutherian UCP1 lineage, again indicative of relaxed constraint.

### Branch-sites model
Branch models have little power to detect directional selection which has acted on just a few codon sites across a lineage, as any signature of selection would be lost in the abundance of negative selection acting across the remaining codons. For this reason, a branch-sites method has been developed that allows variation in $\omega$ across individual codons on a specific lineage [22,23]. This model (MA) designates two categories of branches, again foreground and background, where positive selection is modeled only on the foreground branch. Further, the model includes four site categories for both foreground and background (0, 1, 2a, 2b). Class 0 sites are sites under purifying selec-

tion, where $\omega_0$ is estimated to be between 0 and 1. Class 1 sites are neutral sites, where $\omega_1$ is set to 1. Site classes 2a and 2b are sites which are modeled to be under purifying selection ($\omega_0$) or neutral evolution ($\omega_1$), respectively, on the background branches. On the foreground branch, these sites are under positive selection, where $\omega_2$ is allowed to be greater than 1. We have implemented the MA (branch-sites) model, designating the Eutherian branch of the UCP1 phylogeny (Figure 1) as the foreground branch. The results of this analysis are robust across the different codon substitution models and trees. As done for all analyses, the MA model was tested with both the species tree and the ML tree in combination with the F3X4, F61 and FMutSel models of codon substitution. The MA model ($\omega_2 = 1.45$, $p_2 = 0.33$, $\ell = -10101.60$) did not provide a significantly better fit than the null model, where the null model is neutrality (site class 2 is fixed to equal 1, and thus tests if the MA $\omega_2$ estimate is significantly greater than 1), in any of the analyses ($2\Delta\ell = 0.84$, df = 1, p = 0.36). Thus, these analyses do not provide any evidence for directional selection on the Eutherian UCP1 lineage.

Because it is theoretically possible for novel functions or functional divergence to evolve without positive selection [24,25], it may remain informative to functionally investigate the sites identified in these analyses. Each of the branch-sites models do identify a few robust sites which have greater than a 95% posterior probability of being correctly assigned to site class 2, i.e. those which may have experienced selection, where $\omega$ is greater than one (but not significantly so as demonstrated in the LRT), by the Bayes empirical Bayes (BEB) method [26]; these sites are: 13M, 19S, 26L, 73K, 121L, 150I, 167T, 221V, 256M. It must be emphasized that this analysis is only considered to be informative when the null model of neutrality is rejected and a model of selection provides a better fit to the data. Thus, there is no support for these sites to have been the subject of selection; yet these sites could prove to be informative for future studies on UCP1 function.

### Reduction of UCP1 evolution in Eutherians
To further investigate the possibility of selection we tested for a relative reduction in dN across the Eutherian UCP1 phylogeny. Following the acquisition of a new function by a protein via positive selection at replacement sites, it is predicted that protein evolution should slow down and be dominated by purifying selection [27,28]. To determine if there was a reduction in the rate of non-synonymous substitutions across the UCP1 phylogeny, after the demonstrated relative increase of dN on the Eutherian lineage, we used the UCP2 and UCP3 phylogenies as comparative loci. In the branch (M2) and branch-sites (MA) analysis neither UCP2 nor UCP3 exhibit evidence for either directional selection or relaxed constraints on the

Eutherian branch. The evolutionary history of UCP1 is unique in this respect. Thus we assumed that the estimated evolutionary constraints across the Eutherian phylogenies of the UCP1 paralogs are a good proxy for what may be expected across the UCP1 phylogeny. Estimated dN values, from the M1(free ratio) analysis, for all congruent branches on the UCP1, UCP2 and UCP3 phylogenies were used to perform a one-sided paired Wilcoxon test. If the Eutherian branch had experienced selection at replacement sites we would predict that dN across Eutherians in the UCP1 phylogeny should, at the very least, not be greater than that observed in its paralogs. Contrary to this prediction we find that the rate of substitution at replacement sites remains greater (median dN = 0.01) than that observed across the UCP2 (median dN = 0.003, p-value < 0.001) and UCP3 (median dN = 0.004, p-value = 0.0365) phylogenies. We also found that the rate of replacement substitutions across the UCP3 phylogeny is greater than that estimated in the UCP2 phylogeny (p-value < 0.001). Therefore, congruent lineages across UCP phylogenies do differ from each other in the rate of dN, yet the UCP1 phylogeny has the largest dN rate. There is thus no evidence that UCP1 evolution has slowed in Eutherians, relative to UCP2 or UCP3, as may be expected following a selective event for novel function. Thus these observations refute a conclusion of directional selection and support a model of relaxed constraint on the Eutherian UCP1 lineage.

### Selection for UCP1 on the Therian lineage

Recently, it was discovered that a single marsupial species, the fat-tailed dunnart (*Sminthopsis crassicaudata*), also contains BAT which is recruited in response to cold acclimatization, and that UCP1 expression increases two-fold under cold conditions [1]. Furthermore, the metabolic rate of this species increased in response to norepinephrine treatment [29], consistent with regulation by UCP1. We therefore tested the Therian branch (which includes marsupials but excludes monotremes) for branch-site directional selection on UCP1. We found robust support across all analyses (species tree, ML tree, each codon substitution model and alignments, p < 0.001) for selection acting on a few sites on the Therian lineage. In these analyses, 90% of the sites on the Therian branch are under strong purifying selection with a ω of 0.08, 8.2% of sites are neutrally evolving, and 1.5% of sites have an estimated ω of 8 (i.e., dS was estimated to be equal to 0.0). Two sites, 79P and 178T, were identified with >75% probability by the BEB method of having been the target of directional selection. While it currently remains unknown if UCP1 in the fat-tailed dunnart truly has a thermogenic function, these two identified sites may prove informative in further studies of UCP1 function.

### Functional divergence of UCP1

Directional selection is not a necessary requirement for functional divergence to occur in the evolution of gene families, as a new function for a gene can arise through neutral processes via genetic co-option [30]. To address this possibility of divergent evolution of UCP1 and its paralogs via neutral processes, we used the MC model [31], which allows for divergent selection pressures at one site class between different clades of a phylogeny or between paralogs. A prediction of this model is that one gene, in a gene family, can experience reduced functional constraint, via genetic redundancy, allowing the accumulation of mutations that will alter protein function, particularly at a time when the environment changes. First, the M0 model was used to obtain the global constrained estimate of ω of 0.096 for the entire UCP phylogeny (Figure 2). We then ran the M2 branch model using each UCP Eutherian lineage as the foreground branch. The ω estimate for the UCP1 Eutherian branch was 0.680, and was a significantly better fit to the data than the M0 model (p-value < 0.001), as was found previously when only the UCP1 phylogeny was analyzed (Figure 1, Relaxed Constraint on the Eutherian Lineage). For the UCP2 locus, the ω estimate for the Eutherian lineage was 0.044, which is significantly smaller than the global ω estimate (p-value = 0.03). The ω estimate for the UCP3 Eutherian lineage was 0.081, which is not significantly different from the global ω estimate (p = 0.68). Thus, UCP1 (but not UCP2 or UCP3) exhibits a significantly higher ω value on the Eutherian lineage in the context of the entire UCP phylogeny, consistent with our previous observations and relaxed functional constraint.

We then fit our data to the MC model with five different clades, A-E, set as the foreground branches (Figure 2). The MC model (clade model C) has five parameters ($p_0$, $p_1$, $\omega_0$, $\omega_2$ and $\omega_3$) estimated. There are three site classes; site class 0 are defined sites of purifying selection where $\omega_0$ is estimated to be between 0 and 1 ($p_0$ = proportion of class 0 sites), site class 1 are neutral sites with $\omega_1$ set equal to 1 ($p_1$ = proportion of class 1 sites), and site class 2 is the divergent site class where there is an independent estimate of ω for the background ($\omega_2$) and foreground ($\omega_3$) branches (proportion of sites in class 2: $p_2 = 1 - p_0 - p_1$). The LR test for each of these analyses is the comparison of the MC model to the M1a(neutral) sites model, in which there are two site classes across the entire UCP phylogeny (site class 0 is under purifying selection and ω is estimated from the data; site class 1 is neutral and ω is set equal to 1). The M1a(neutral) model estimates that 90% of the sites are under strong purifying selection with a ω estimate of 0.08. Each MC analysis provided a significantly better fit to the data than the M1a(neutral) model (Table 2). The parameter estimates (Table 2) are very similar across analyses, suggesting that the model is indeed robust to altering

**Table 2: MC clade model parameter estimates, LRT, and p-value**

| Clade | $p_0$ | $\omega_0$ | $p_2$ | Background $\omega$ | Foreground $\omega$ | $2\Delta\ell$ | p |
|---|---|---|---|---|---|---|---|
| A_UCP1 | 0.57 | 0.02 | 0.39 | 0.17 | 0.24 | 1095 | <0.001 |
| B_UCP2/3 | 0.58 | 0.02 | 0.39 | 0.22 | 0.18 | 1087 | <0.001 |
| C_UCP2 | 0.58 | 0.02 | 0.38 | 0.22 | 0.14 | 1099 | <0.001 |
| D_UCP3 | 0.59 | 0.02 | 0.37 | 0.19 | 0.22 | 1084 | <0.001 |
| E_Eutherian_UCP1 | 0.56 | 0.02 | 0.4 | 0.17 | 0.29 | 1119 | <0.001 |

$p_0$ is the estimated proportion of sites in site class 0 (purifying selection), $\omega_0$ is the estimated $\omega$ value for site class 0; $p_2$ is the estimated proportion of sites in site class 2 (divergent selection), Background $\omega$ ($\omega_2$) is the estimated $\omega$ value for divergent sites on the background branches; Foreground $\omega$ ($\omega_3$) is the estimated $\omega$ value for divergent sites on the foreground branches; $2\Delta\ell$ is the LR test statistic for comparing the MC and M1a(neutral) models; p is the p-value of the LR test.

foreground and background branches. To summarize, overall about 60% of the sites in the UCP phylogeny are under strong purifying selection, about 1% of the sites are evolving neutrally, and about 39% of the sites are under divergent selection pressures between UCP1 and its paralogs. The results for clades B, C, and D (Table 2, foreground $\omega$) suggest that there has been more purifying selection across the UCP2 phylogeny than that of the UCP3.

The model with the highest log likelihood is with the Eutherian UCP1 clade (clade E) set as the foreground branch. Similar results are obtained when the entire UCP1 clade (clade A) is set as the foreground branch. However, the log likelihood is lower and the $\omega_3$ estimate for the foreground divergent class is larger when only considering the Eutherian UCP1 clade (clade E). The $\omega_3$ observed (0.24) for the entire UCP1 clade (clade A) may largely reflect the divergent selection pressures in the Eutherian clade ($\omega_3 = 0.29$), with the marsupial and fish lineages reducing the $\omega_3$ estimate for the foreground branch in the entire UCP1

phylogeny (clade A). Furthermore, it is the Eutherian UCP1 lineage itself that is driving the functional divergence of the Eutherian UCP1 clade. Repeating the MC model but setting the Eutherian UCP1 lineage (E) (i.e. just branch E, not the clade) as the foreground branch, an $\omega$ of 0.87 is estimated for divergent sites, while background lineages have an $\omega$ of 0.18, where $p_2 = 0.390$ ($2\Delta\ell = 1119$, $p < 0.001$; identical $\ell$ to Clade E (Table 2)). In conclusion, these analyses indicate that the UCP1 phylogeny exhibits large divergent selection pressures, which may principally involve the Eutherian UCP1 lineage, consistent with the possible acquisition of the new function of UCP1 through neutrally evolving, relaxed constraints.

The BEB method is implemented in each analysis to determine the probability [26] of each site belonging to site class 0 (purifying), site class 1 (neutral), or site class 2 (divergent). Site classifications are consistent across each analysis. Figure 3 is a plot of site class probability, i.e. the probability of each codon belonging to one of the three site categories, namely purifying, neutral or divergent.
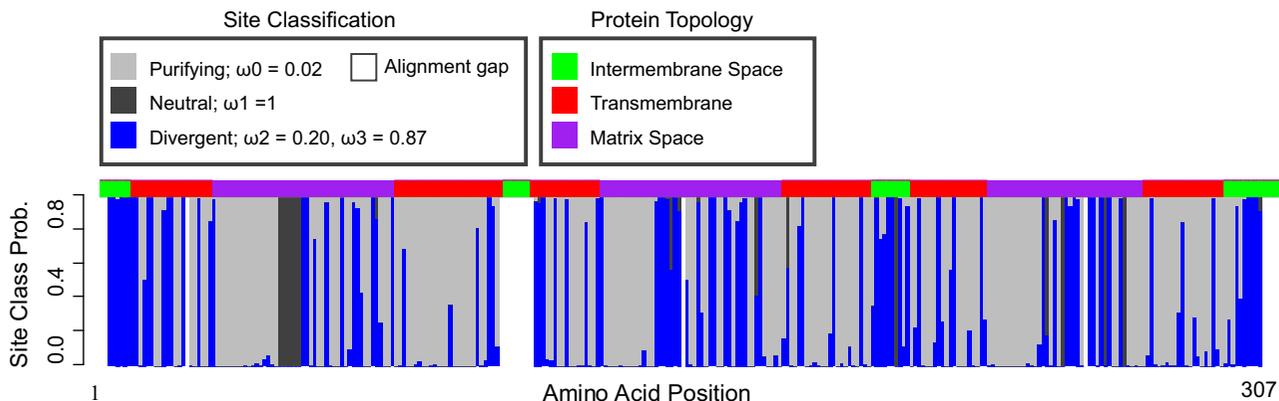


**Figure 3**
**A bar plot of site class probabilities for each codon, represented by amino acids in the figure**. Site classifications and their omega estimate are defined in the figure. The protein topology, defined as intermembrane, transmembrane and matrix space protein portions was placed above the barplot. Topology was taken from [56].

Approximately 40% of all sites are under divergent selection pressure, with nonsynonymous substitutions accumulating at a rate ~1.7× greater in the UCP1 Eutherian clade (clade E) relative to its paralogs, as compared to synonymous substitutions. Sites categorizations are nearly identical when the MC model is run with either the Eutherian UCP1 clade or the Eutherian UCP1 lineage set as the foreground branch (5% discrepancy in site classification between analyses). It is only the ω estimations that are largely different for the divergent class between the two analyses (0.29 vs. 0.87). In contrast to the Eutherian clade analysis, the Eutherian lineage analysis has a nonsynonymous substitution rate that is ~5× greater than that of all other lineages in the entire phylogeny. Divergent sites are found across the entire protein and in all functional domains, assigned by their topology to regions located in the intermembrane, transmembrane, and matrix space. Notably, the protein regions exposed to the intermembrane space are significantly over-represented with divergent sites ($\chi^2$ = 26.32, p-value = 0.0005, p computed by Monte Carlo simulation). This result suggests that functional constraints are more relaxed for the regions of the protein exposed to the intermembrane space in Eutherians. This may have enabled the recruitment of novel interactions with proteins which potentially regulate the activity of UCP1.

### Positive selection following duplication

In addition to any selection on UCP1 for NST in Eutherians, there may have been positive selection on other sites in any of the UCP loci following the gene duplications that created them. We therefore used the UCP phylogeny (Figure 2) and the branch-sites model (MA) to test lineages A (UCP1), B (UCP2/3), C (UCP2), and D (UCP3) for positive selection acting on only a few sites. Again, LR tests were used to compare the MA model to the MA1 model, in which the ω estimate for site class 2 is set to 1, to assess if the MA model with positive selection is significantly better than the MA1 model of neutrality. No robust statistical evidence for selection was discovered for any basal UCP lineage. Statistical evidence for selection was identified for UCP1 (A), UCP2/3 (B) and UCP2 (C) when the analysis was carried out using the MUSCLE alignment (Table 3), but these results were not supported by the T-Coffee alignments, and hence any conclusions

about positive selection on any of the UCP loci following gene duplication depend on the alignment.

### Codon usage bias

A recently developed codon substitution model, FMutSel, explicitly models mutational biases and selection at silent codon sites [32]. This model was compared to the null model, FMutSel0, which only models mutational biases, and does not include the additional codon fitness parameter in FmutSel which models selection on silent sites. The LRT between these two models has 41 degrees of freedom, resulting from the parameter rich FMutSel model (60 parameters). It has been suggested that codon usage bias could have a drastic effect on estimation of dN and dS [22,33]. The species tree for each UCP locus was used in these analyses, not the full UCP phylogeny found in Figure 2. We find that each of the UCP loci has statistically significant evidence for selection at silent sites (Table 4). This result is not surprising, as about 90% of 5,369 genes tested reject the null model, indicating that selection on silent sites is a common phenomenon across the mammalian genome [32]. Thus, codon usage bias is widespread and silent sites are not neutrally evolving. This conclusion, however, does not discredit the use of ω (the comparison of dN and dS) as a means of identifying sites of selection. Selection on the protein alters the rate of evolution at replacement sites, and it is this rate of change before and after selection that is contrasted. Thus, it does not matter if the rate of substitution at silent sites is driven by mutation or selection [32].

Interestingly, even though the FmutSel model was a much better fit to the data than the null model for all three loci, the statistical significance of the difference between the models for the UCP2 and UCP3 loci are orders of magnitude larger than for the UCP1 locus (Table 4). This may suggest that selection on silent sites has been reduced across the UCP1 phylogeny. The proportion of advantageous mutations (P+) estimated in the UCP phylogenies is about 0.3 – 0.4, similar to the average estimate of 0.39 across 5,639 genes [32]. In contrast, the estimated average selection coefficient for advantageous (S+) and disadvantageous (S-) mutations (Table 4) for each of these UCP loci are greater than three times the average selection coefficient for other genes [32]. Thus, mutations at silent sites

**Table 3: MA branch-sites parameter estimates following gene duplication**

| Lineage | p0 | ω0 | p2 | ω2 | Selected Sites | 2Δℓ | p |
|---|---|---|---|---|---|---|---|
| **A_UCP1** | 0.887 | 0.078 | 0.022 | 468.880 | >90% 256M | 6.803 | 0.009 |
| **B_UCP2/3** | 0.875 | 0.076 | 0.034 | 999.000 | >80% 148A, 274M, >90% 120L | 7.709 | 0.005 |
| **C_UCP2** | 0.886 | 0.076 | 0.021 | 14.950 | >95% 207A, 300A | 4.780 | 0.028 |
| **D_UCP3** | 0.878 | 0.076 | 0.029 | 2.870 | >70% 130C | 0.028 | 0.868 |

$p_0$ and $p_2$ are the estimated proportion of sites, and $\omega_0$ and $\omega_2$ are the estimated ω values for site class 0 (purifying selection) and site class 2 (positive selection), respectively; Selected sites are those with a large posterior probability of belonging to the selected site category.

**Table 4: FMutSel LRT and parameter estimates with the M0 model**

| Locus | FMutSel $\ell$ | FMutSel0 $\ell$ | 2$\Delta\ell$ | p-value | $\omega$ | P+ | S+ | S- |
|-------|----------|-----------|--------|---------|----------|------|------|-------|
| UCP1 | -10029.62 | -10118.34 | 177.44 | <0.001 | 0.129 | 0.40 | 1.42 | -1.58 |
| UCP2 | -8657.37 | -8777.49 | 240.24 | <0.001 | 0.074 | 0.31 | 1.90 | -2.44 |
| UCP3 | -8857.95 | -9004.26 | 292.62 | <0.001 | 0.099 | 0.41 | 1.57 | -1.93 |

FMutSel$\ell$ is the log likelihood of the FMutSel substitution model; FMutSel0$\ell$ the log likelihood of the FMutSel0 substitution model; 2$\Delta\ell$ is the LR test statistic for comparing these two models; p is the p-value of the LR test; $\omega$ is the estimated $\omega$ value for the designated gene phylogeny using the global M0 model; P+ is the proportion of advantageous mutations that could occur across the gene; S+ is the selection coefficient for advantageous mutations; S- is the selection coefficient for disadvantageous mutations.

in these three loci would seem to be either highly preferential or highly deleterious, as compared to the genome average.

## Discussion

It has long been thought that UCP1 had evolved in Eutherian mammals in conjunction with the evolution of brown adipose tissue. Recent discoveries of UCP1 orthologs in fish and non-Eutherian mammals have now discredited this idea [1,2,17], as fish and most non-Eutherian mammals do not contain brown adipose tissue where UCP1 plays a major role in thermogenesis. Furthermore, the recent discovery of UCP1 localized in adipose tissue in a marsupial species, where it is also regulated in response to cold acclimatization [1] has forced a reevaluation of the evolution of non-shivering thermogenesis. In particular, when did the biochemical properties of UCP1 which confer the capability of thermogenesis evolve?

In this study we have used codon substitution models implemented through a maximum likelihood framework to estimate the rate of evolution at silent and replacement sites in UCP1 and its paralogs, UCP2 and UCP3. Different models were used to investigate variation in the rate of evolution between lineages of a phylogeny, between clades of a phylogeny, and to estimate $\omega$ for specific lineages and sites across phylogenies. Our objective was to determine the mode of evolution on the UCP1 Eutherian lineage, to illuminate possible evolutionary consequences of novel functions necessary for NST. We observed a significant, five-fold increase in the rate of UCP1 evolution on the Eutherian lineage, but found no strong statistical evidence for directional selection. Yet, it should be noted that the power of these analyses are dependent on the magnitude of the selection pressure and the proportion of sites affected [34,35]. These models do have underlying assumptions, namely no variation in branch and site $\omega$ values. As we have demonstrated, variation among branch and sites is observed in the UCP1 phylogeny. While large deviations from these assumptions will affect the power of the analysis, the MA model does seem to be robust to moderate deviation in these assumptions [34]. As such, relaxed functional constraint is most consistent with the molecular evolutionary analyses of the UCP data. Further,

our observation of strong purifying selection being the primary mode of evolution throughout the UCP phylogeny is consistent with the findings of a recent study published while this paper was under review [36].

Our conclusion is in contrast to a previous study which implemented a heuristic branch model and concluded that the UCP1 Eutherian lineage had experienced positive selection on replacement sites [18]. That study estimated 82.8 nonsynonymous and 4.3 synonymous substitutions on the Eutherian branch using a least-squares approach, whereas we estimate 88.5 nonsynonymous and 47.5 synonymous substitutions respectively. We repeated the analyses in [18], using their methodology, and obtained essentially the same results they did (data not shown). To determine why these two methods are giving drastically different results, we simulated the evolutionary history of UCP1 using the branch-lengths given by the M0 global model, the global $\omega$ estimate of 0.1134, and the Evolver program in the PAML 4a package (Additional file 2). We then estimated $\omega$ for the Eutherian lineage using both the heuristic method [18] and the ML (M2) model used in this study. The resulting $\omega$ estimate for the ML model was 0.1135, virtually identical to the value used in the simulations, while the $\omega$ estimate for the heuristic approach was 1.239. We repeated the simulation, setting $\omega = 0.5$ for the Eutherian lineage and $\omega = 0.1$ for all other lineages, which mirrors the observed data. The resulting $\omega$ estimate for the Eutherian lineage was 0.47 based on the ML model and 7.75 based on the heuristic approach. In each simulation the heuristic approach greatly under-estimated the number of synonymous substitutions. We conclude that the contradiction between our results and the previous study [18] results from the poor performance of the heuristic least-squares method on trees and lineages with long branch lengths. It appears that least-squares method drastically underestimate the number of synonymous substitutions [37].

An alternative to relaxed functional constraint as an explanation for the increased rate of evolution on the UCP1 Eutherian lineage would be a decrease in the effective population size ($N_e$) of the ancestor of all Eutherian mammals. However, a change in $N_e$ along the Eutherian lineage

would affect all loci in the genome. As such the UCP1 paralogs, UCP2 and UCP3, should also reflect any possible changes in $N_e$. Yet these UCP1 paralogs do not exhibit a rate increase, so a decrease in $N_e$ can be excluded. Therefore it remains that the best explanation for the increase in UCP1 evolution along the Eutherian lineage is relaxation of functional constraints, possibly allowing functional divergence through neutral processes.

Assuming that UCP1 did indeed experience relaxation of functional constraints on the Eutherian lineage, what may have caused this? There are several possible explanations, which are not mutually exclusive. First, it has been documented that the physiological relevance of thermogenic function varies across Eutherians, with NST being most efficient in small mammals and hibernators, particularly those which inhabit environments with long cold seasons [38,39]. If the ancestor of all Eutherians was a large species where NST was not important, or a species which inhabited a moderate environment, functional constraints on UCP1 could have been relaxed. The ancestor of all extant Eutherians lived ~150–100 MYA, during the late Jurassic and early Cretaceous [40]. In general mammals of this period are thought to have been small in size [41] and this era was also generally warmer than today [42,43]. Thus it is possible that constraints on UCP1 were reduced at some time point prior to the Eutherian radiation, some 100 MYA. Second, it could be that some function(s) of UCP1 in Eutherians was taken over by UCP1's paralogs, which would also reduce the level of constraint. It is certain that UCP1 in Eutherians must have had a role in some other tissue(s) prior to becoming an obligate BAT protein. Genes expressed in many tissues tend to be more constrained than those expressed in one or a few tissues [44]. Thus, the release of UCP1 from its role in other tissues would allow UCP1 to evolve tissue specificity to BAT, decrease the selective constraint on UCP1, and allow for genetic co-option (the altering of a genetic system for a new use). For example, as shown in Table 1, the ω estimate ($\omega_0$) assigned to site class 1 (strong purifying selection) suggests that there has been more selective constraint on UCP2, which is expressed in many tissues, than on UCP3, which is expressed in just a few tissues [15,16]. The function(s) of UCP2 and UCP3 are currently uncertain; determining their function(s) would certainly aid in understanding UCP1 evolution.

In the process of genetic co-option or functional gene divergence it is not necessary for directional selection to act on protein evolution in order for divergent/novel functions to evolve. Constraint on a protein can be relaxed when a paralog is available to fulfill the function of that protein. As such, evolution can then act to use existing gene functions for a new purpose. The prediction of heterogeneity in selection pressures between gene paralogs

exemplifies the divergent selection pressures that allow the evolution of divergent functions. We examined the potential of genetic co-option, i.e. functional divergence of UCP1, through the implementation of the MC clade model [31]. We found that about 40% of the sites were under divergent selection pressures, with the UCP1 clade experiencing a greater number of substitutions at replacement sites relative to silent sites (Figure 3). Moreover, intermembrane sites are over-represented in the divergent site category, which may suggest that intermembrane interactions have diverged in Eutherian UCP1. Regardless, it is apparent that there has been tissue-specific functional divergence of UCP1 in BAT.

Cold-induced UCP1 expression was recently detected in adipose tissue of the interscapular region of the Australian fat-tailed dunnart (*Sminthopsis crassicaudata*), a dasyurid marsupial species [1]. Further, the adipose tissue in which it was found was described as being an archetypal brown adipose tissue. As such, these findings challenge the view of NST as an exclusive property of Eutherians. However, this finding is unique to this particular marsupial, as cold-induced UCP1 expression was not found in the dasyurid marsupial *Antechius flavipes* (yellow-footed Antechinus) or the South American didelphid marsupial *Monodelphis domestica* (gray short-tailed opossum). Furthermore, there is currently no evidence for a thermogenic role of UCP1 in the fat-tailed dunnart. Its regulation and expression is suggestive of a thermogenic role, yet no evidence of non-shivering thermogenesis or "adaptive-NST", i.e. an adaptive organismal response to the cold, has been demonstrated. As such, if NST is determined to be present in the fat-tailed dunnart, it would most likely be a case of parallel evolution or multiple lineage loss of NST, as deduced from phylogenetic inference and given that these findings are restricted to a single non-Eutherian species.

Nevertheless, since UCP1 in the fat-tailed dunnart does show a response to cold and norepinephrine [29] similar to that found in Eutherians, it may be that any truly novel function of UCP1 started evolving in the Therian (marsupial and Eutherian) lineage. We therefore tested for selection on UCP1 on this lineage and found strong evidence for positive selection, and we identified two sites (79P and 178T) which may be responsible for any novel UCP1 function. These two sites should be the target of further functional investigations. Moreover, it is not clear if UCP1 in marsupials is truly functionally identical to UCP1 in Eutherians; more studies are needed to clarify this issue.

## Conclusion
In agreement with previous studies, we find evidence for acceleration in the evolution of UCP1 on the Eutherian lineage. However, in contrast to previous studies, our analyses indicate that relaxed constraints on UCP1, rather

than directional selection, seem to be responsible. We suggest that this relaxation of constraints on UCP1 may have allowed the acquisition or enhancement of the (potentially) novel role of UCP1 in NST in Eutherians, as supported by the divergent selection pressures identified in the UCP phylogeny. Currently, the most likely hypothesis for the origin of NST hinges upon the evolution of BAT with high mitochondrial oxidative capacities. For this requirement to be meet evolution needs a tissue with many mitochondria which in turn have a high expression level of UCP1 and complexes I-IV of the electron transport chain. This combination can then induce a high mitochondrial oxidative capacity which will lead to heat production (i.e. NST). As such, we hypothesize that UCP1 in Eutherians and non-Eutherians share similar or identical biochemical properties. To test this hypothesis the biochemical properties of UCP1 in non-Eutherians need to be elucidated.

## Methods
### Coding sequence sources
Coding sequences (cds) for UCP1, UCP2 and UCP3 were obtained from the NCBI CoreNucleotide and Ensembl http://www.ensembl.org databases. A table with species names, abbreviations and accession numbers can be found in supplementary materials (Additional file 1).

### Data analysis
Coding sequences were imported into MEGA 3.1 [45] and converted into amino acid sequences, which were then aligned with MUSCLE 3.6 [46] and T-Coffee 3.8 [47] and exported in CLUSTAL W format. The amino acid alignment was subsequently transformed into an aligned cds fasta file with a custom PERL script. The accepted species tree [40,48] was generated by hand and gene trees were generated using various nucleotide and amino acid substitution models in PHYLIP [49]. The gene trees analyzed included: neighbor joining (NJ) trees from nucleotide sequence distances (dnadist; F84, Kimura and Jukes-Cantor substitution models) with a gamma distributed rate variation and a transition/transversion bias ($\kappa$) of 2.0; a DNA maximum likelihood (dnaml) tree with $\kappa$ = 2.0, rate variation among sites, global rearrangements, and randomized input; and lastly a NJ tree from protein distances (protdist) was generated using the JTT and PAM substitution models and a gamma distributed rate variation. Maximum likelihood analyses were performed using the aligned cds and the CODEML program found in the PAML 4a package [19]. The species tree and generated gene trees were analyzed with the CODEML program of the PAML 4a package, which implements the re-sampling estimated log-likelihood (RELL) method, and the likelihood ratio tests K-H and S-H. These are nonparametric likelihood ratio tests, which test if two phylogenetic trees differ significantly in how well they fit the data. As it is not

appropriate to use a maximum likelihood tree with the K-H test [50] we focused on the results from the S-H test and the RELL method. Finally, CODEML of the PAML 4a package was also used to estimate the ratio of non-synonymous or replacement changes per site/synonymous changes per site, also known as the dN/dS ratio or the omega value ($\omega$). CODEML implements a maximum likelihood (ML) method of estimating $\omega$ [51], using a variety of codon substitutions models and estimation of $\kappa$, and accounts for base frequency biases at codon positions. Different site [52,53], branch [54,55], branch-sites [22,23,26], and clade models [31] were implemented in estimating the $\omega$ value. In this study we used the F3X4 and F61 models of codon frequencies [51], as well as the recently developed FMutSel model [32], which models mutational biases and introduces a parameter of codon fitness to model selection at silent sites. The F3X4 and F61 codon models differ in the manner in which they estimate each codon's equilibrium frequency. The F3X4 model derives equilibrium codon frequencies from the frequencies of the three nucleotides at the three codon positions, while the F61 model uses each codon as a free parameter and constrains the sum to one. Analyses such as these allow the determination of the type of evolution (neutrality, negative selection, or positive selection) which predominates across the entire phylogenetic tree, across each site in the sequence, on particular branches, at particular sites on particular branches, or at particular sites across clades. In these analyses an $\omega$ value equal to one is indicative of neutral evolution, $\omega$ greater than one is indicative of positive selection, and $\omega$ less than one is indicative of negative or purifying selection.

## Authors' contributions
MK and MJ initiated the study; DH and MS conceived the study design; DH performed the analyses and analyzed the data; DH, MS and MK wrote the manuscript. All authors approved the final manuscript.

## Additional material

> ### Additional file 1
> *Accession Numbers. Table of accession numbers including: list of species, species abbreviations, and accession numbers for sequences used in the study.*
> Click here for file
> [http://www.biomedcentral.com/content/supplementary/1471-2148-9-4-S1.pdf]

> ### Additional file 2
> *UCP1 Simulations. Description and results of UCP1 simulations and $\omega$ estimations using heuristic and maximum likelihood approaches.*
> Click here for file
> [http://www.biomedcentral.com/content/supplementary/1471-2148-9-4-S2.pdf]

## Acknowledgements

## References

1. Jastroch M, Withers KW, Taudien S, Frappell PB, Helwig M, Fromme T, Hirschberg V, Heldmaier G, McAllan BM, Firth BT, Burmester T, Platzer M, Klingenspor M: **Marsupial uncoupling protein 1 sheds light on the evolution of mammalian nonshivering thermogenesis.** *Physiol Genomics* 2008, **32(2):**161-169.
2. Jastroch M, Wuertz S, Kloas W, Klingenspor M: **Uncoupling protein 1 in fish uncovers an ancient evolutionary history of mammalian nonshivering thermogenesis.** *Physiol Genomics* 2005, **22(2):**150-156.
3. Stuart JA, Harper JA, Brindle KM, Jekabsons MB, Brand MD: **A mitochondrial uncoupling artifact can be caused by expression of uncoupling protein 1 in yeast.** *Biochem J* 2001, **356(Pt 3):**779-789.
4. Cannon B, Nedergaard J: **Brown adipose tissue: function and physiological significance.** *Physiol Rev* 2004, **84(1):**277-359.
5. Mzilikazi N, Jastroch M, Meyer CW, Klingenspor M: **The molecular and biochemical basis of nonshivering thermogenesis in an African endemic mammal, Elephantulus myurus.** *Am J Physiol Regul Integr Comp Physiol* 2007, **293(5):**R2120-2127.
6. Krauss S, Zhang CY, Lowell BB: **The mitochondrial uncoupling-protein homologues.** *Nat Rev Mol Cell Biol* 2005, **6(3):**248-261.
7. Nicholls DG, Rial E: **A history of the first uncoupling protein, UCP1.** *J Bioenerg Biomembr* 1999, **31(5):**399-406.
8. Rousset S, Alves-Guerra MC, Mozo J, Miroux B, Cassard-Doulcier AM, Bouillaud F, Ricquier D: **The biology of mitochondrial uncoupling proteins.** *Diabetes* 2004, **53(Suppl 1):**S130-135.
9. Nicholls DG, Locke RM: **Thermogenic mechanisms in brown fat.** *Physiol Rev* 1984, **64(1):**1-64.
10. Foster DO, Frydman ML: **Tissue distribution of cold-induced thermogenesis in conscious warm- or cold-acclimated rats reevaluated from changes in tissue blood flow: the dominant role of brown adipose tissue in the replacement of shivering by nonshivering thermogenesis.** *Can J Physiol Pharmacol* 1979, **57(3):**257-270.
11. Heldmaier G, Buchberger A: **Sources of heat during nonshivering thermogenesis in Djungarian hamsters: a dominant role of brown adipose tissue during cold adaptation.** *J Comp Physiol [B]* 1985, **156(2):**237-245.
12. Puchalski W, Bockler H, Heldmaier G, Langefeld M: **Organ blood flow and brown adipose tissue oxygen consumption during noradrenaline-induced nonshivering thermogenesis in the Djungarian hamster.** *J Exp Zool* 1987, **242(3):**263-271.
13. Bouillaud F, Couplan E, Pecqueur C, Ricquier D: **Homologues of the uncoupling protein from brown adipose tissue (UCP1): UCP2, UCP3, BMCP1 and UCP4.** *Biochim Biophys Acta* 2001, **1504(1):**107-119.
14. Echtay KS: **Mitochondrial uncoupling proteins – what is their physiological role?** *Free Radic Biol Med* 2007, **43(10):**1351-1371.
15. Nedergaard J, Cannon B: **The 'novel' 'uncoupling' proteins UCP2 and UCP3: what do they really do? Pros and cons for suggested functions.** *Exp Physiol* 2003, **88(1):**65-84.
16. Ricquier D, Bouillaud F: **The uncoupling protein homologues: UCP1, UCP2, UCP3, StUCP and AtUCP.** *Biochem J* 2000, **345(Pt 2):**161-179.
17. Jastroch M, Buckingham JA, Helwig M, Klingenspor M, Brand MD: **Functional characterisation of UCP1 in the common carp: uncoupling activity in liver mitochondria and cold-induced expression in the brain.** *J Comp Physiol [B]* 2007, **177(7):**743-752.
18. Saito S, Saito CT, Shingai R: **Adaptive evolution of the uncoupling protein 1 gene contributed to the acquisition of novel nonshivering thermogenesis in ancestral eutherian mammals.** *Gene* 2008, **408(1–2):**37-44.
19. Yang Z: **PAML 4: phylogenetic analysis by maximum likelihood.** *Mol Biol Evol* 2007, **24(8):**1586-1591.
20. Kishino H, Miyata T, Hasegawa M: **Maximum likelihood inference of protein phylogeny and the origin of chloroplasts.** *J Mol Evol* 1990, **31:**151-160.
21. Shimodaira H, Hasegawa M: **Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference.** *Mol Biol Evol* 1999, **16(8):**1114-1116.
22. Yang Z, Nielsen R: **Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages.** *Mol Biol Evol* 2002, **19(6):**908-917.
23. Zhang J, Nielsen R, Yang Z: **Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level.** *Mol Biol Evol* 2005, **22(12):**2472-2479.
24. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J: **Preservation of duplicate genes by complementary, degenerative mutations.** *Genetics* 1999, **151(4):**1531-1545.
25. Zhang J: **Evolution by gene duplication: an update.** *Trends in Ecology & Evolution* 2003, **18(6):**292.
26. Yang Z, Wong WS, Nielsen R: **Bayes empirical bayes inference of amino acid sites under positive selection.** *Mol Biol Evol* 2005, **22(4):**1107-1118.
27. Aguileta G, Bielawski JP, Yang Z: **Gene conversion and functional divergence in the beta-globin gene family.** *J Mol Evol* 2004, **59(2):**177-189.
28. Ohta T: **Pattern of nucleotide substitutions in growth hormone-prolactin gene family: a paradigm for evolution by gene duplication.** *Genetics* 1993, **134(4):**1271-1276.
29. Clements F, Hope P, Daniels C, Chapman I, Wittert G: **Thermogenesis in the marsupial Sminthopsis crassicaudata: effect of catecholamines and diet.** *Aust J Zool* 1998, **46(4):**381-390.
30. True JR, Carroll SB: **Gene co-option in physiological and morphological evolution.** *Annu Rev Cell Dev Biol* 2002, **18:**53-80.
31. Bielawski JP, Yang Z: **A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution.** *J Mol Evol* 2004, **59(1):**121-132.
32. Yang Z, Nielsen R: **Mutation-Selection Models of Codon Substitution and Their Use to Estimate Selective Strengths on Codon Usage.** *Mol Biol Evol* 2008.
33. Yang Z, Bielawski JP: **Statistical methods for detecting molecular adaptation.** *Trends Ecol Evol* 2000, **15(12):**496-503.
34. Anisimova M, Yang Z: **Multiple hypothesis testing to detect lineages under positive selection that affects only a few sites.** *Mol Biol Evol* 2007, **24(5):**1219-1228.
35. Wong WS, Yang Z, Goldman N, Nielsen R: **Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites.** *Genetics* 2004, **168(2):**1041-1051.
36. Hughes J, Criscuolo F: **Evolutionary history of the UCP gene family: gene duplication and selection.** *BMC Evol Biol* 2008, **8:**306.
37. Zhang J, Rosenberg HF, Nei M: **Positive Darwinian selection after gene duplication in primate ribonuclease genes.** *Proc Natl Acad Sci USA* 1998, **95(7):**3708-3713.
38. Sundin U, Moore G, Nedergaard J, Cannon B: **Thermogenin amount and activity in hamster brown fat mitochondria: effect of cold acclimation.** *Am J Physiol* 1987, **252(5 Pt 2):**R822-832.
39. Barger JL, Barnes BM, Boyer BB: **Regulation of UCP1 and UCP3 in arctic ground squirrels and relation with mitochondrial proton leak.** *J Appl Physiol* 2006, **101(1):**339-347.
40. Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A: **The delayed rise of present-day mammals.** *Nature* 2007, **446(7135):**507-512.
41. Luo ZX: **Transformation and diversification in early mammal evolution.** *Nature* 2007, **450(7172):**1011-1019.
42. Steuber T, Rauch M, Masse JP, Graaf J, Malkoc M: **Low-latitude seasonality of Cretaceous temperatures in warm and cold episodes.** *Nature* 2005, **437(7063):**1341-1344.
43. Wilson PA, Norris RD: **Warm tropical ocean surface and global anoxia during the mid-Cretaceous period.** *Nature* 2001, **412(6845):**425-429.
44. Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, Weiss G, Lachmann M, Paabo S: **Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees.** *Science* 2005, **309(5742):**1850-1854.
45. Kumar S, Tamura K, Nei M: **MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.** *Brief Bioinform* 2004, **5(2):**150-163.

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

47. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment.** *J Mol Biol* 2000, **302(1):**205-217.

48. Benton MJ, Donoghue PC: **Paleontological evidence to date the tree of life.** *Mol Biol Evol* 2007, **24(1):**26-53.

49. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6.** *Distributed by the author Department of Genome Sciences, University of Washington, Seattle* 2004.

50. Goldman N, Anderson JP, Rodrigo AG: **Likelihood-based tests of topologies in phylogenetics.** *Syst Biol* 2000, **49(4):**652-670.

51. Goldman N, Yang Z: **A codon-based model of nucleotide substitution for protein-coding DNA sequences.** *Mol Biol Evol* 1994, **11(5):**725-736.

52. Nielsen R, Yang Z: **Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene.** *Genetics* 1998, **148(3):**929-936.

53. Yang Z, Nielsen R: **Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models.** *Mol Biol Evol* 2000, **17(1):**32-43.

54. Yang Z: **Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution.** *Mol Biol Evol* 1998, **15(5):**568-573.

55. Yang Z, Nielsen R: **Synonymous and nonsynonymous rate variation in nuclear genes of mammals.** *J Mol Evol* 1998, **46(4):**409-418.

56. Klingenspor M, Fromme T, Hughes DA Jr, Manzke L, Polymeropoulos E, Riemann T, Trzcionka M, Hirschberg V, Jastroch M: **An ancient look at UCP1.** *Biochim Biophys Acta* 2008.