Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown

Michael R. McGowen, Lawrence I. Grossman and Derek E. Wildman*

Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, 540 E Canfield St., Detroit, MI 48210, USA

Cetaceans (dolphins and whales) have undergone a radical transformation from the original mammalian body plan. In addition, some cetaceans have evolved large brains and complex cognitive capacities. We compared approximately 10,000 protein-coding genes culled from the bottlenose dolphin genome with nine other genomes to reveal molecular correlates of the remarkable phenotypic features of these aquatic mammals. Evolutionary analyses demonstrated that the overall synonymous substitution rate in dolphins has slowed compared with other studied mammals, and is within the range of primates and elephants. We also discovered 228 genes potentially under positive selection (dN/dS > 1) in the dolphin lineage. Twenty-seven of these genes are associated with the nervous system, including those related to human intellectual disabilities, synaptic plasticity and sleep. In addition, genes expressed in the mitochondrion have a significantly higher mean dN/dS ratio in the dolphin lineage than others examined, indicating evolution in energy metabolism. We encountered selection in other genes potentially related to cetacean adaptations such as glucose and lipid metabolism, dermal and lung development, and the cardiovascular system. This study underlines the parallel molecular trajectory of cetaceans with other mammalian groups possessing large brains.

Keywords: dolphin; Cetacea; brain; genome; evolution; metabolism

1. INTRODUCTION

Cetaceans represent the most successful mammalian colonization of the aquatic environment. Therefore, it is not surprising that cetaceans possess various extremes in mammalian morphology and physiology, including loss of limbs, development of echolocation, and extensive changes in respiratory and cardiovascular anatomy and physiology [1]. One aspect of cetacean biology that is striking is the relative size and complexity of the brain, especially in toothed whales (odontocetes) [2]. Similar to anthropoid primates, odontocetes have evolved brains that are larger than expected for their body size [2,3]. When taking phylogeny into account, relative brain size in some odontocetes, such as the bottlenose dolphin (Tursiops truncatus), is greater than in all non-human primates, and sperm whales (Physeter macrocephalus) possess the largest brains in absolute terms [4,5]. In addition, cetaceans have also evolved many other neurological features convergent with some primates such as a high level of gyriﬁcation (cortical folding), expansion of the insular and cingulate cortices, specialized von Economo neurons, high glia to neuron ratio, increase in synapse number, reduction of the olfactory system and the large relative size of the cerebral cortex [3–6]. Cetaceans offer an opportunity for understanding the convergent evolution of large brains among disparate mammalian groups, as multiple peaks in the evolution of large brains exist within mammals [7].

In addition, tissue of the central nervous system generally requires a larger proportion of energy than other tissues [8], and basal metabolic rate correlates with relative brain size [9]. Therefore, the evolution of large brain size also requires an increase in basal metabolic rate or efficiency of metabolism to accommodate expanded energy demands. Adaptive evolution of genes involved in aerobic metabolism (i.e. the mitochondrial electron transport chain) has been documented within the relatively large-brained anthropoid primates and along the lineage leading to humans [10]. Recent phylogenomic studies of molecular evolution have also discovered parallels in the evolution of aerobic metabolism genes between the large-brained African elephant and primates [11]. This association suggests that the evolution of metabolic genes underlies the evolution of a large brain in other groups with a significant shift in relative brain size and/or cognitive complexity. One study has recovered signatures of molecular adaptation in the cytochrome b gene of cetaceans [12], but other genes expressed in mitochondria have not been evaluated.

Here, we compared approximately 10,000 protein-coding sequences (CDS) from the genome of the common bottlenose dolphin (Tursiops truncatus) with nine other amniote genomes, including the closest relative with a sequenced genome, the domestic cow (Bos taurus). Evolutionary analyses showed that the overall protein-coding substitution rate along the dolphin lineage was significantly lower than in other mammals analysed and...
on par with humans and elephants. The dolphin lineage exhibited evidence of positive selection of multiple genes associated with the nervous system, including those involved in human intellectual disorders, sleep, myelin formation, neurite growth and maintenance of synaptic connections, among others. In addition, the dolphin lineage showed a significant increase in comparison with other laurasiatherian genomes in selection on genes that localize to the mitochondria, the source of aerobic metabolism. We also find positive selection in genes that may be linked to specialized adaptations such as glucose and lipid metabolism, diving, fat storage and dermal development.

2. METHODS

CDS of individual genes from the dolphin genome (T. truncatus, turTru1, 2.59× coverage) were downloaded from ENSEMBL v. 61 and used to query the following genomes for homologous sequences: cow (B. taurus, UMD3.1, 9.5×), horse (Equus caballus, EquCab2, 6.8×), dog (Canis familiaris, BROAD2, 7.6×), mouse (Mus musculus, NCBI37), human (Homo sapiens, GRCh37), elephant (Loxodonta africana, losAf3, 7×), opossum (Monodelphis domestica, BROAD05, 6.8×), platypus (Ornithorhynchus anatinus, OANA5, 6×), chicken (Gallus gallus, WASHUC2, 6.6×). The cow was selected because it is the closest-related species with a sequenced genome; both dolphin and cow are included in the mammalian order Cetartiodactyla [13]. All other genomes were selected owing to their placement as successive outgroups to Cetartiodactyla [13] and their relatively deep connections, among others. In addition, the dolphin lineage exhibited evidence of positive selection of multiple genes on par with humans and elephants. The dolphin lineage was dated of 59.1 Myr was used to date the divergence of cow and dolphin [15], and 84.6 was used to date the divergence of cow + dolphin, dog and horse [16]. We used PRISM v. 5 for Mac OS X (GRAPHPAD Software) to test the difference in means for each lineage for dN/dS ratios.

Owing to the inconsistent inclusion of outgroups for all alignments, we confined our analyses to five lineages (dolphin, cow, horse, dog and cow + dolphin stem). Genes from each lineage were divided into non-overlapping bins based on dN/dS as in Goodman et al. [11]. In addition, we constructed bins with dN/dS > 0.5 as well as bins with the top 5 per cent of dN/dS values. Genes were inferred to have undergone positive selection on a particular lineage if dN/dS > 1. We then used the functional annotation clustering tool DAVID [17,18] to discover and group gene ontology (GO) categories into annotation clusters that are enriched for each bin in each lineage. Gene lists for each bin were input separately and used to search against a background of human orthologues of total genes retained for each lineage analysed. DAVID ranks functional annotation clusters by enrichment score (ES) that is equivalent to the geometric mean of the p-values for each GO category included in the annotation cluster. Within each annotation cluster, DAVID lists GO categories that are significantly enriched; corrections for multiple testing are given using the false discovery rate (FDR). All groups with ES > 1.3 were considered for significant enrichment for each list of genes examined; ES > 1.3 corresponds to a p-value of 0.05 [18].

Owing to the importance of the mitochondrion in cellular metabolism, we also measured the degree of selection pressure on genes expressed in the mitochondrion among lineages. Differences in mean dN/dS values of genes belonging to the GO ontology category ‘mitochondrion’ GO:0005739’ were tested for significance using a two-tailed Mann–Whitney U-test as mentioned earlier.

Owing to the imprecise nature of GO categories and the continual stream of new information on specific genes [19], we performed additional literature searches for the top 5 per cent of genes in each lineage to identify genes involved in the central nervous system as well as other potentially relevant physiological systems such as the
circulatory system. In addition, we also searched the Genes to Cognition database that retains data on genes related to cognition [20,21].

Low-coverage genomes, such as the dolphin genome (2.59×), are adversely affected by sequencing error in the form of an estimated 1–4 errors per kilobase [22]. Although this has little effect on overall rates of substitution, it can modestly affect dN/dS ratios by up to 10 per cent [22]. We have implemented several procedures to document and ameliorate the potential effects of error on our data. Firstly, we used gene builds from Ensembl release 61 [23]. Ensembl infers putative protein-coding genes in 2× genomes using a whole genome alignment procedure to construct gene scaffolds and project protein-coding transcripts onto the 2× genome from a more complete reference genome. Putatively, erroneous indels or stop codons have been excised by removing them from the coding sequence [23].

Secondly, to examine the extent of base misspecification and erroneous indels in the remaining CDS regions of the Ensembl 2× dolphin genome, CDS sequences of 152 genes from T. truncatus and two closely related species (Tursiops aduncus or Stenella coeruleoalba) were downloaded from GenBank. All T. truncatus gene segments represent individuals distinct from that used to construct the dolphin genome. In the case of 14 genes from other species, CDS in these genes were actually more similar to the T. truncatus genome than those labelled in GenBank as T. truncatus. These 152 genes (126 690 bp) were compared with homologous sequences in the dolphin genome, and both base differences and numbers of indels tabulated.

Finally, to confirm dolphin genome sequence identity, we directly sequenced five genes with evidence of dN/dS > 0.9 (CYCS, DBL, PLN, PRR9 and TTR) from T. truncatus. DNA was extracted from the villous trophoblast of the placenta using the DNeasy Tissue Kit (Qiagen, Valencia, CA, USA). We amplified sequences using standard PCR protocols and primers listed in the electronic supplementary material, table S1. All sequences were deposited in GenBank.

### 3. RESULTS

Our scan resulted in 10 025 dolphin orthologues that met our criteria (see §2) for comparison with other genomes. Length of these multiple sequence alignments ranged from 26 355 bp in SYNE1 to 150 bp in G20orf196 and FDP5. Among laurasiatherian lineages (dolphin, cow, horse, dog), means for dN/dS, rN and rS were all significantly different from one another (Mann–Whitney U-test, all tests p < 0.05; table 1 and the electronic supplementary material, figures S2–S7). Mean rates of synonymous substitution (rS) along the dolphin lineage are the slowest among all four lineages measured in this study (table 1). Rates of synonymous substitution in dolphin are comparable to estimates of the human lineage, as described in Goodman et al. [11] (table 1). Values of mean rS for other lineages examined (cow, horse, dog) are greater than human and elephant but less than mouse and tenrec. Differences between rN and rS range from a 5.35-fold difference in dolphin to a more than sevenfold difference in dog. This is reflected in the mean (dN/dS) of each lineage in which the highest is in dolphin and the lowest in dog (table 1). When compared with Goodman et al. [11], dN and dS are closer to human and elephant than tenrec and mouse, and mean dN/dS is greater in dolphin than in elephant or human (see the electronic supplementary material, table S2).

Within dolphin, we discovered 228 genes (2.26%) with a dN/dS > 1.0, signifying them as candidates for positive selection in this lineage. The dolphin–cow stem lineage recovered 153 positively selected genes (1.54%). These were in stark contrast to the other lineages analysed (horse, cow and dog) that possessed 48 (0.51%), 32 (0.32%) and 11 (0.12%) positively selected genes, respectively. If we consider genes with evidence of relaxed selection pressure (defined here as 0.5 < dN/dS < 1.0) as well as those exhibiting positive selection dN/dS > 1.0, the dolphin lineage had 1029 (approx. 10%) genes, approximately twice as many or more as the cow–dolphin stem (566; 5.69%), horse (446; 4.77%), cow (374; 3.76%) and dog (238; 2.54%). At least 46 per cent of genes in all lineages except dolphin are under purifying selection (dN/dS < 0.1); in dolphin, this number is only 36.4 per cent or 3646 genes.

To investigate selection on aerobic metabolism genes, we compared dN/dS scores of the 548 total genes classified in the mitochondrion cellular component GO category (GO:0005739) (figure 1). Seven mitochondrial expressed genes were each found to be under positive selection both in the dolphin terminal lineage and in the dolphin–cow stem lineage, compared with one gene in horse and zero in cow and dog. The percentage of total mitochondrial genes with dN/dS > 0.5 was approximately 8.5 (48 out of 548) in dolphin compared with 1.5–5 in the other lineages. The somatic isoform of the major electron transport chain gene cytochrome c (CYCS) has a dN/dS = 0.9731 and four amino acid differences from cow. The proportion of genes with dN/dS < 0.1 was only 31 per cent in dolphin, but 44–51 per cent in all other lineages. Mann–Whitney U-tests on pairs of lineages identified the mean dN/dS in both the dolphin lineage (0.2136) and the horse lineage (0.1543) as significantly different from all others and one another (all p < 0.0001); the mean dN/dS of the dolphin lineage was greater than other lineages. These results imply either relaxed purifying selection or greater selective pressure on mitochondrial expressed genes in the dolphin lineage.

### Table 1. Mean dN, dS, dN/dS, rN and rS from four terminal branches using the free-ratio model of branch-specific evolution in Yang [14].

<table>
<thead>
<tr>
<th>branch</th>
<th>mean dN</th>
<th>mean dS</th>
<th>mean dN/dS</th>
<th>mean rN (× 10⁻¹⁰)</th>
<th>mean rS (× 10⁻⁸)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dolphin</td>
<td>0.016</td>
<td>0.083</td>
<td>0.238</td>
<td>2.67</td>
<td>1.40</td>
</tr>
<tr>
<td>cow</td>
<td>0.023</td>
<td>0.163</td>
<td>0.146</td>
<td>3.82</td>
<td>2.75</td>
</tr>
<tr>
<td>horse</td>
<td>0.023</td>
<td>0.144</td>
<td>0.163</td>
<td>2.66</td>
<td>1.71</td>
</tr>
<tr>
<td>dog</td>
<td>0.027</td>
<td>0.204</td>
<td>0.134</td>
<td>3.26</td>
<td>2.42</td>
</tr>
</tbody>
</table>
Annotation clustering analyses using bins with dN/dS > 1 in some lineages, we compared GO categories of annotation clusters with ES > 1.3 generated from functional annotation clustering analyses using bins with dN/dS > 0.5 and the top 5 per cent of dN/dS scores per lineage (see the electronic supplementary material, table S3). Complete lists with ES and FDR are shown in the electronic supplementary material, supplemental dataset S1. In general, annotation clusters associated with the immune system (GO:0006952), those involved with transcription and translation (GO:0006412), and those involved with transcription and translation (GO:0006350) show a greater proportion of mitochondrion-expressed genes in bins with greater dN/dS values as well as larger relative brain size.

Owing to the small number of genes with dN/dS > 1 in some lineages, we compared GO categories of annotation clusters with ES > 1.3 generated from functional annotation clustering analyses using bins with dN/dS > 0.5 and the top 5 per cent of dN/dS scores per lineage (see the electronic supplementary material, table S3). Complete lists with ES and FDR are shown in the electronic supplementary material, supplemental dataset S1. In general, annotation clusters associated with the immune system (GO:0006952 ~ defence response; GO:0006954 ~ inflammatory response; GO:0009617 ~ response to bacterium; GO:002684 ~ positive regulation of immune system process, among others), extracellular region (GO:0005576 ~ extracellular region; GO:0005525 ~ extracellular space) and the plasma membrane (GO:0031224 ~ intrinsic to membrane; GO:0016021 ~ integral to membrane) figured prominently in the gene clusters with the highest ES (greater than 1.3) of all lineages. Representation of most of the individual GO terms within the enriched clusters listed earlier were significant based on FDR in all lineages. Immunity genes are common among positively selected genes in many lineages of mammals examined [11,24]. Genes involved in male reproduction (GO:0048232 ~ male gamete generation, GO:0007276 ~ spermatogenesis) were also enriched in the dN/dS > 0.5 bin of dolphin (ES = 1.72) to a greater extent than other lineages (ES = 0.06–0.93); however, individual GO terms were not considered significant based on FDR. Extremely conserved genes (dN/dS < 0.025) in all lineages included GO categories such as GTP binding (GO:0005525), those involved with transcription and translation (GO:0006412 ~ transcription, GO:0006350 ~ transcription, GO:0006396 ~ RNA processing) and chromosome organization (GO:0051276).

Although our functional annotation analyses did not recover enriched GO categories involved in neurological function (see the electronic supplementary material, table S3), our literature and database searches recovered at least 27 genes with evidence of expression and/or function in the nervous system and dN/dS > 1 on the dolphin lineage (table 2). Many of these genes come from a diverse array of functional types and are not indexed with nervous system function GO categories, even though evidence supporting a role in the nervous system exists to the contrary (table 2; see the electronic supplementary material, table S4). Seven genes are identified as being involved in intellectual disabilities and/or microcephaly (ERCC8, AP4S1, MCPH1, TTR; [25–27]), schizophrenia (MAL; [28]) or Alzheimer’s susceptibility (AGER, RNFL82; [29,30]). Five genes are involved in neuroendocrine function, neuropeptide hormonal activity, or function as hormones that affect the brain (AGRP, C2orf40, EDN2, NMU, TTR). Other genes function in the development of myelin (MAL), neuronal or brain development (CNPY1, ZNF597, PCP4L1, MCPH1), neural potassium channel function (KCNK18), neurite growth (CD47, LRFN1, CNPY2) and synaptic function (BAALC, DBI, SYPL1, AP4S1, LRFN1) (see the electronic supplementary material, table S4). Some genes have primary functions in other systems, but are also expressed in the synaptosome of mouse (table 2) [20]. To determine whether the excess of nervous system genes is specific to the dolphin genome, we also looked at nervous system genes in the three other laurasiatherian genomes (cow, dog and
A general error rate of 1.57 bases per 1000 bp and 0.34 genome. We found 201 bp and 44 indels differ between sequences with homologous regions in the dolphin (126,690 bp) from GenBank to compare these lineages, see the electronic supplementary material, for a complete list of all positively selected genes in all lineages. Five of the 152 genes also had a dN/dS > 1 in our genome scan (CSN2, MCPH1, KCNJ2, BAG2, IL1A); of these, only two differed from the genome (MCPH1, IL1A) with one amino acid difference each. In addition, we sequenced five genes, and discovered no difference in CDS in four genes (CYS2, DBI, PLN, PRR9). In CYS2, we discovered a polymorphism at two adjacent nucleotides resulting in a potential amino acid change in at least one allele; it is more likely that this difference is due to individual variation than error in sequencing.

4. DISCUSSION
(a) Slow molecular rate in cetaceans
Our results compiled from approximately 10,000 protein-coding genes confirm the assessment that rates of synonymous substitution along the cetacean lineage are slower on average than other mammals, as established by several analyses of smaller datasets [31–35]. Mean rS in dolphins (1.40 × 10−9 substitutions per site yr−1) was lower than the mammalian average mutation rate of 2.2 × 10−9 substitutions per site yr−1 as measured by Kumar & Subramanian [36]. Mean rS was on par with elephants and humans, groups that also have been well-documented as undergoing a rate slowdown [11,37,38]. Multiple studies have shown that the rate of synonymous substitution in both mitochondrial and nuclear genes is negatively correlated and dN/dS is positively correlated with mass, longevity, population size and sometimes generation time [31,37,39–41]. Welch et al. [40] detected that dN is correlated with body mass as well. The common bottlenose dolphin (T. truncatus) can weigh up
to 500 kg and has a lifespan of at least approximately 40–50 years, with age at first reproduction approximately 9.5 years and a generation time of approximately 21.5 years.

In size, longevity and generation time, cetaceans are some of the most extreme in the mammal world, and on par or greater than that of humans and elephants; therefore, it is unclear which of these are the driving force behind the cetacean slowdown, but similarities between such groups of highly encephalized, large, long-lived mammals are striking.

(b) Adaptive evolution of nervous system genes and the evolution of the brain in cetaceans

Here, we find a large number of nervous system genes putatively under selection compared with other Laurasiatherian lineages. We caution that no direct link has been established between these genes and any morphological or behavioural trait in cetaceans, and changes in amino acids may not correlate with changes in protein function. However, we speculate that the association of large brains and complex social behaviour in cetaceans with the evidence of adaptive evolution in nervous system genes may be linked and briefly discuss some of these genes later in the text.

One aspect of brain evolution and its association with cognition that has received much attention is the evolution of synaptic proteins and the expansion of synaptic connections. Differences in synaptic protein diversity and expression have been hypothesized to correspond with variation in cognition among species. In addition, synaptic plasticity is directly related to memory and learning through interactions at the molecular level. Genes involved in synaptic function have been implicated in the evolution of the brain in humans, especially AMPA and NMDA glutamate receptors. Here, we recovered six proteins with definitive roles in synaptic function, including a membrane protein associated with AMPA glutamate receptors (AP4S1), a synaptic molecule differentially regulated during human development (SYPL1), a synaptic adhesion molecule (LRN1), a component of postsynaptic complexes (RAALC) and a modulator of signal transduction at type A gamma-aminobutyric acid receptors with implications in sleep regulation (DBI), in addition to numerous proteins detected in the postsynaptic density (table 2).

DNA repair genes with primary expression in the central nervous system represent a large number of genes implicated in intellectual disability in humans, as neural progenitor cells seem to be more sensitive to DNA damage. ERCC8 is a gene involved in transcription-coupled DNA damage repair; mutations in this gene result in Cockayne’s syndrome, whose symptoms include microcephaly, neurological defects and premature ageing. ERCC8 shows evidence of positive selection in recent human populations. MCPH1 is another gene with evidence of function in DNA damage repair as well as cell cycle regulation. MCPH1 dysfunction can cause microcephaly and has been documented to be under positive selection along the human lineage. A previous study also documented positive selection in MCPH1 along the dolphin lineage.

Other genes with known expression in the nervous system include those involved in various functions in development and support of neurons. MAL is involved in myelin biogenesis and function and is required for a proper interaction between glia and axons. AGER translocates amyloid-beta peptide across the cell membrane in cortical neurons, leading to mitochondrial dysfunction and has been identified as a major Alzheimer’s disease susceptibility gene. CD47 is involved in axon and dendrite development in the hippocampus, the centre of executive function in the brain. TTR is a thyroid hormone-binding protein that transports thyroxine, which controls metabolic rate, from the bloodstream to the brain. TTR evolution is also accelerated on the human lineage.

(c) Molecular signatures of metabolic evolution

Bottlenose dolphins have a higher metabolic rate than comparable land mammals of the same size and consume oxygen at a rate of approximately 1.1 ml kg⁻¹ or approximately 6.0 ml kg⁻¹; this is greater than both human and elephant. Along with primates, central nervous system metabolism as a function of body metabolism is elevated in dolphins relative to other mammals. Large relative brain size is constrained by the increased demand upon the rate of metabolism itself, and relative brain size is correlated with basal metabolic rate. Therefore, it is not unlikely that elevated rates of evolution in genes expressed in the mitochondrion of dolphins are related to the increased need to provide increasing energy to such a large organ, as argued in Goodman et al. [11]. Here, we found that CYCS, an electron acceptor of the electron transport chain, has also undergone accelerated evolution on the cetacean lineage. Genes expressed in the mitochondria, especially those of the electron transport chain and CYCS specifically, have experienced accelerated evolution in anthropoid primates as well. Examination of the dolphin–anthropoid proteins shows four site differences; half of them are in a region previously suggested to be in the binding site for interaction with complex IV, with one of them being identical to the anthropoid site.

During the course of primate evolution, metabolism evolved to provide a constant supply of energy to the increasing demands of a larger brain. Humans have evolved several adaptations, including an increase in visceral fat and the evolution of tissue-specific insulin resistance in the brain that protects nervous tissue from energy shortage; however, malfunctions of this system in humans can cause type 2 diabetes. Interestingly, recent research has proposed that dolphins should be seen as a model for type 2 diabetes, as they possess all the hallmarks of the disease.

Here, we find genomic signatures of adaptive evolution in dolphin genes related to control of food intake such as genes involved in glycerol uptake and/or glucose metabolism (AQP9, OSTN, SOCS6), and a neuropeptide AGRP that is expressed in the hypothalamus and involved in increasing appetite, decreasing metabolism and regulating leptin. In addition, we found several genes under selection related to lipid transport and metabolism (see §3) that may be associated with the large fat reserves found in cetaceans.

(d) Validation of dolphin genome

Here, we discovered little error in the small segment of dolphin genome that we investigated. Changes in nucleotide sequence were predicted to be only approximately 1.5
bases per kilobase, and we could not rule out whether these changes were indeed error or intraspecies variation. These error rates are within the lower range of estimates of sequencing errors of low-coverage mammalian genomes using comparison with ENCODE regions (1–4 bases per kilobase) [21]. Variation within T. truncatus undoubtedly exists as some closely related species had sequences that were more similar or identical to the Tursiops genome sequence when compared with GenBank sequences from T. truncatus. Sequence error had little effect on rates of non-synonymous substitution, as nR of dolphin was on par with horse, a genome with 6.8 × coverage. In addition, rates of substitution (nN, nS) in the dolphin genome were about half that of tenrec, another low-coverage genome analysed using similar protocols [10]. Hubisz et al. [21] discovered that dN/dS ratios could differ up to 10 per cent in low-coverage genomes, which is modest; this could potentially change dN/dS = 1.00–1.11 to dN/dS < 1, or 54 out of 228 genes in our dataset. Therefore, even though we found few differences in the genes we resequenced, we cannot rule out some effect of error on dN/dS. Detailed investigation of individual genes identified here as under adaptive selection in multiple cetacean species are the next step to investigating molecular evolution in this fascinating lineage.

5. CONCLUSIONS

Despite some limitations owing to the low quality of the dolphin genome, this analysis adds to a greater understanding of molecular evolution in cetaceans. Our results provide a baseline for the study of molecular adaptations in cetaceans and some evidence for convergent features in the genomes of dolphins and primates. We documented rates of synonymous substitution in the dolphin lineage that were significantly lower than other mammals and on par with humans and elephants. The dolphin lineage exhibited evidence of positive selection of multiple genes associated with the nervous system, including those involved in synaptic plasticity, as well as genes involved in metabolic processes, both those that putatively supply more energy and those related to glycolytic regulation. In addition, the dolphin lineage shows a significant increase in selection on genes expressed in the mitochondrion in comparison with other mammalian genomes. We also find positive selection in genes possibly linked to cetacean specializations such as deep diving, blubber and fat storage. This study adds to a greater understanding of the molecular landscape surrounding the convergent evolution of larger brains and complex cognition.

We thank M. Islam, Z.-C. Hou and K. Sterner for assistance and advice concerning bioinformatics. W. Gundling provided assistance with amplification of exonic sequence. Three anonymous reviewers provided comments on earlier drafts of this manuscript. Our deceased colleague, M. Goodman, was involved in early stages of planning this project. This research was funded by NSF grants no. BCS-0550209 to L.I.G. and BCS-0827546 to D.E.W.

REFERENCES


Dolphin genome

M. R. McGowen et al.


58 Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. 2006.09.007)


