Marine ecosystems are responsible for virtually all production of omega-3 (ω3) long-chain polyunsaturated fatty acids (PUFA), which are essential nutrients for vertebrates. Current consensus is that marine microbes account for this production, given their possession of key enzymes including methyl-end (or “ωx”) desaturases. ωx desaturases have also been described in a small number of invertebrate animals, but their precise distribution has not been systematically explored. This study identifies 121 ωx desaturase sequences from 80 species within the Cnidaria, Rotifera, Mollusca, Annelida, and Arthropoda. Horizontal gene transfer has contributed to this hitherto unknown widespread distribution. Functional characterization of animal ωx desaturases provides evidence that multiple invertebrates have the ability to produce ω3 PUFA de novo and further biosynthesize ω3 long-chain PUFA. This finding represents a fundamental revision in our understanding of ω3 long-chain PUFA production in global food webs, by revealing that numerous widespread and abundant invertebrates have the endogenous capacity to make significant contributions beyond that coming from marine microbes.

INTRODUCTION

Marine ecosystems are responsible for virtually all natural production of omega-3 (ω3) long-chain (≥C20) polyunsaturated fatty acids (PUFA), which are essential dietary components for vertebrates that are required for neural development and health promotion while also having beneficial roles across a variety of human pathologies and disorders (1). With an expected increased demand for ω3 oil supplements prescribed for prevalent cardiovascular and inflammatory diseases in humans (1), there is intense interest in exploring novel sources of ω3 oils (2, 3). The current dogma is that photosynthetic marine microalgae, heterotrophic protists, and bacteria account for natural ω3 long-chain PUFA production (4–6), because they have all the enzymatic components necessary for de novo synthesis, unlike higher trophic organisms such as fish. Plants have the ability for de novo synthesis of PUFA of up to 18 carbons but limited elongation capacity to produce long-chain PUFA (7). Although some bacteria and lower eukaryotes can produce PUFA de novo via a polyketide synthase pathway (6), PUFA biosynthesis in most aquatic microbes requires fatty acid (FA) desaturases that catalyze the insertion of new double bonds (unsaturations) (8, 9). Δ9 desaturases mediate the insertion of the first unsaturation in stearic acid (18:0) to produce oleic acid (18:1ω9). Insertion of further unsaturations is achieved through sequential reactions driven by methyl-end (or “ωx”) desaturases that introduce a double bond between an existing unsaturation and the methyl end of the fatty acid (8). Thus, the action of an ω6 (or Δ12) desaturase enables the conversion of 18:1ω9 to 18:2ω6ω9 (linoleic acid [LA]), the latter acting as an intermediate metabolite in de novo ω3 PUFA biosynthesis because it is subsequently converted to 18:3ω3 [α-linolenic acid (ALA)] by an ω3 (or Δ15) desaturase (4, 5). Although Δ9 desaturases are present in virtually all organisms (8), de novo ω3 PUFA production is only possible if both Δ12 and Δ15 desaturases coexist within an organism.

Vertebrates lack the ωx desaturases necessary for PUFA biosynthesis and thus require dietary LA and ALA to satisfy essential FA requirements (9, 10). The same condition has been largely accepted for all animals, with few exceptions reported. Two pulmonate molluscs (11) and some insects (12, 13) showed the ability to biosynthesize LA, but not ALA. Recently, the deep-sea tube worm Riftia pachyptila (Annelida) has been shown to have an ω3 desaturase enabling the biosynthesis of ALA but not LA (14). Currently, only two terrestrial metazoans, the nematode Caenorhabditis elegans (15–17) and the arthropod Bemisia argentifolii (18), have been reported to have ωx desaturases with both Δ12 and Δ15 activities. If, contrary to current consensus, these few instances of metazoan ωx desaturases, in fact, reflect a prevalent occurrence of these enzymes across animals, then a major revision of our understanding of ω3 PUFA production in global food webs is required. The present study provides a comprehensive search of available sequence databases and analysis of the distribution of ωx desaturase genes across the Metazoa. We establish that, contrary to common perception, ωx desaturases are widely distributed across animal phyla. Furthermore, selected ωx desaturase genes from key lineages with widespread occurrence in aquatic ecosystems, namely, Cnidaria (Acropora millepora), Rotifera (Adineta vaga), Mollusca (Patella vulgata), Annelida (Platyneris dumerilii), and Arthropoda (Lepeophtheirus salmonis), were functionally characterized to clarify their role in PUFA biosynthesis. Our results demonstrate that multiple invertebrates can actively produce ω3 PUFA de novo. Moreover, we also establish additional desaturase capabilities in animal ωx desaturases that enable the biosynthesis of physiologically essential ω3 long-chain PUFA.

RESULTS AND DISCUSSION

Our search for ωx desaturase-like sequences showed the presence of these genes in both prokaryotes (Cyanobacteria) and eukaryotes including Archaeplastida (red and green algae, Viridiplantae), Excavata, SAR (stramenopiles, alveolates, and rhizaria), Amoebozoa, Haptophyceae, Apusozoa, Heterolobosea, Cryptophyta, and Opisthokonta (divided into Fungi, Metazoa, and other Opisthokonta) (Fig. 1 and fig. S1). In
Metazoa, we found hitherto unknown ox desaturase diversity. Sequences with homology to ox desaturases were retrieved from the phyla Cnidaria, Rotifera, Mollusca, Annelida, Arthropoda, and Nematoda. Overall, 121 ox desaturase sequences from 80 animal species were identified (Fig. 1). A list of all genomes and transcriptomes searched, including those in which no ox desaturase genes were found, can be accessed as described in the Data and materials availability section, below.

Phylogenetic analysis using all recovered metazoan sequences, along with a representative subset of sequences from the remaining nonmetazoan groups, revealed at least three distinct ox desaturase gene lineages (clades 1 to 3 in Fig. 1) in animals; most likely, these have independent evolutionary origins, because they do not form a single monophyletic animal clade. Clade 1 was composed exclusively of sequences from nematodes, including two C. elegans ox desaturases (15–17), with an outgroup of nonmetazoan genes (for example, microalgae, a fungus, and an alveolate) (Fig. 1). Clade 2 comprised two distinct gene lineages in Cnidaria (Fig. 1), with sequences from stony corals, coral anemones, sea anemones, and zoanthids. The cnidianian sequences (clade 2) were nested within a broad and well-supported clade that included most of the eukaryotic nonmetazoan sequences of the data set, but the specific sister group to clade 2 was not well resolved (Fig. 1). Clade 3 included sequences from Lophotrochozoa (namely, Rotifera, Mollusca, and Annelida) and nonnematode Edyssozoa (copepod arthropods). Oomycota sequences formed a well-resolved sister group to the clade 3 animal genes (Fig. 1). As in clade 2, there were frequently two genes in most clade 3 species. Beyond clades 1 to 3, three smaller sets of animal sequences from the arthropods Bemisia tabaci (Hemiptera), Locusta migratoria (Orthoptera), and Smittius viridis (Colembola) appeared scattered in the tree (Fig. 1). The division of most metazoan sequences into three main clades is also consistent with the amino acid compositions of the three histidine box motifs (H-box1–3) typical of desaturases (fig. S2) (19).

Fig. 1. Distribution of ox desaturases. (A) A condensed phylogenetic tree depicting relationships among ox desaturases from Cyanobacteria (8) and Eukaryota (279). Metazoan phyla (represented with 121 sequences) are named in the tree and assigned with colors for the following taxa: Nematoda (green), Cnidaria (blue), Arthropoda (purple), and Lophotrochozoa (orange). Nonmetazoan clades were collapsed and named according to their general taxonomic composition. “Miscellaneous Clade A” contains representatives of Amoeboza, Apusozoa, Haptophyceae, Heterolobosea, Nucleariidae, SAR, Chlorophyta, Rhodophyta, Cryptophyta, and the insect L. migratoria. Representatives of Excavata, Amoebozoa, Fungi, SAR, and the hexapod S. viridis compose “Miscellaneous Clade B.” The phylogenetic tree presented corresponds to the 80% rule consensus tree estimated in MrBayes. Colored dots on the nodes represent supporting values of 80% or above for MrBayes and PhyloBayes analyses (blue) or for all three (MrBayes, PhyloBayes, and RAxML) analyses (red). (B) Summary of taxonomic groups highlighting metazoans in which ox desaturases have been identified in this study. Number of species is indicated in brackets. A list of all genomes and transcriptomes searched, including those in which no ox desaturase genes were found, can be accessed as described in the Data and materials availability section.
Horizontal gene transfer (HGT) can be inferred in some animal species and may explain the general phylogenetic pattern of metazoan ox desaturase genes. For example, several ox desaturase–like sequences from the silverleaf whitefly, *B. tabaci*, are most closely related to one of the Viridiplantae ox desaturase clades in the phylogeny (Fig. 1 and fig. S1), which includes FAD2 plant desaturases functionally characterized as Δ12 desaturases (20). To exclude the possibility that these sequences might represent plant sequence contamination in *B. tabaci*, we examined the genomic neighborhoods of *B. tabaci* ox desaturases. The *B. tabaci* sequences are genome-anchored and distributed across five different scaffolds (Fig. 2A). *B. tabaci* ox desaturase sequences are flanked by genes that have clear orthologs in other metazoans (Fig. 2A). This genomic arrangement implies that the *B. tabaci* ox desaturase sequences are bona fide components of the whitefly genome. Thus, HGT likely accounts for the occurrence of ox desaturases in *B. tabaci*, with the genes possibly transferring from plants on which this whitefly feeds. Further cases of HGT appear to exist in a locust and a collembolan. Collembola have ox desaturase sequences that are most closely related to sequences from Oomycota species that are plant pathogens, which are presumably prevalent in the soil habitat occupied by collembolans. Our data thus indicate that the described ability of some collembolans to synthesize LA is due to an endogenous capacity derived from the presence of ox desaturases in this hexapod and not to any associated microbes (21). Furthermore, HGT has been suggested for the *Fat-1* and *Fat-2* genes.
ω-3 desaturases in *C. elegans* (15–17), with protists postulated as probable “donor taxa” (22). More generally, HGT between closely associated organisms has been established previously (23, 24), whereas other desaturase genes have been demonstrated to have arisen via HGT in arthropods, such as the two-spotted spider mite, *Tetranychus urticae* (25). Therefore, although additional evolutionary processes are also likely to be affecting metazoan ω-3 desaturase genes, it is reasonable to suppose that HGT could account for multiple ω-3 desaturases in metazoans. HGT is a particularly plausible scenario for aquatic invertebrates (26).

Having established that ω-3 desaturase sequences are phylogenetically widespread across the Metazoa, it was a central aim to confirm that these genes are capable of de novo ω-3 PUFA biosynthesis. Using a yeast heterologous expression system, we investigated the functions of two distinct ω-3 desaturases for each of five representative species of selected metazoan lineages including Cnidaria (*A. millepora*), Rotifera (*A. vaga*), Mollusca (*P. vulgata*), Annelida (*P. dumerilii*), and Arthropoda (*L. salmonis*) (Fig. 3 and Tables 1 and 2). With the exception of *A. vaga*, a fairly consistent pattern of desaturase capabilities was found among representatives of Cnidaria, Mollusca, Annelida, and Arthropoda. Thus, one gene from the pair of ω-3 desaturases in each species exhibited Δ12 desaturation ability, enabling the biosynthesis of 18:2ω6 (Fig. 3B and Table 1). The second ω-3 desaturase of each pair exhibited Δ15 desaturation ability, being able to convert 18:2ω6 to 18:3ω3 (ALA) (Fig. 3 and Table 2) and

![Diagram](http://advances.sciencemag.org/)

**Fig. 3.** Functional characterization of metazoan ω-3 desaturases. (A) The de novo production of ω-3 PUFA requires both ω-6 (Δ12) (blue arrow) and ω-3 (Δ15) desaturases (red arrow). Both LA (18:2ω6) and ALA (18:3ω3) can be subsequently modified through the ω-6 and ω-3 long-chain PUFA biosynthesis pathways that proceed separately, at least in vertebrates (9), or can be interconnected by ω-3 desaturases with Δ15, Δ17, or Δ19 activities (green arrows). (B) Chromatograms of FA methyl esters (FAME) from yeast expressing ω-3 desaturases from *A. vaga* (MF448339) (top) and *A. millepora* (KY658237) (middle) show Δ12/Δ15 and Δ12 activities, respectively. Further ω-3 desaturases with Δ12 activity were characterized from *P. vulgata*, *P. dumerilii*, and *L. salmonis* (Table 1). (C) Chromatograms of FAME from yeast expressing the ω-3 desaturase from *P. vulgata* (KY658241) and grown with supplemented ω-6 substrates including 18:2ω6 (left), 20:4ω6 (middle), and 22:4ω6 (right) show the resulting ω-3 products 18:3ω3, 20:5ω3, and 22:5ω3, respectively. Similar functions obtained for ω-3 desaturases from *A. millepora*, *A. vaga*, *P. dumerilii*, and *L. salmonis* are shown in Table 2. Molecular structures and diagnostic mass ions of 4,4-dimethyloxazoline derivatives from desaturation products are shown in figs. S3 and S4.
thus completing all the desaturation steps required for de novo biosynthesis of ω3 PUFA. The capability for de novo biosynthesis of ω3 PUFA through the combined action of two distinct enzymes in representatives of Cnidaria, Mollusca, Annelida, and Arthropoda was also contained within each of the two ox desaturases characterized from the rotifer *A. vaga* (Fig. 3 and Table 2). Our results contradict the currently accepted paradigm (6, 9, 27) because they show unequivocally that both Δ12 and Δ15 desaturase activities are present in multiple metazoans, enabling them to endogenously produce ω3 PUFA de novo. The yeast system used in this study thus confirms that the enzyme activities reported in vivo experiments are due to endogenous metazoan genes and not microbes (28, 29). One of the ωx desaturases described in cnidarians, molluscs, annelids, and crustaceans was able to produce a variety of ω3 long-chain PUFA from their corresponding ω6 precursors (Table 2), in addition to 18:3ω3, confirming that these enzymes are true ω3 desaturases with coexisting Δ15, Δ17, and Δ19 activities. This represents long-chain PUFA-synthesizing capabilities previously unknown outside microbes (5, 14, 15).

The present study shows that rather than merely converting dietary PUFA [like the “trophic upgrading” done by fish (30)], aquatic invertebrates can now be recognized as net producers of ω3 long-chain PUFA in agreement with studies (29), providing much more widespread biosynthesis of these compounds in the marine environment than previously appreciated. Consequently, models that estimate global production of ω3 long-chain PUFA, which are largely constructed on the assumption that single-cell microorganisms are the sole primary producers, need extensive revision (27, 31). ω3 long-chain PUFA production by metazoans will undoubtedly be significant due to the abundance of the animals that are now known to have ωx desaturases have in global ecosystems.

### Table 1. Characterization of Δ12/Δ15 activities of metazoan ωx desaturases.

Table 1. Characterization of Δ12/Δ15 activities of metazoan ωx desaturases. FA profiles of yeast transformed with empty pYES2 (control) were compared with those of yeast transformed with the corresponding ωx desaturase (*A. millepora* (KY658237), *A. vaga* (ex1, MF448337; ex2, MF448339), *P. vulgata* (KY658242), *P. dumerilii* (KY658238), and *L. salmonis* (KY658240)). FA compositions (% of total FA) are given as means ± SEM of replicate samples (*n* = 6 for control and *n* = 4 for transgenic yeast). An asterisk indicates significant differences between control and yeast expressing the corresponding desaturase (Dunnett’s test; *P* < 0.05). nd, not detected.

<table>
<thead>
<tr>
<th>FA</th>
<th>Control</th>
<th><em>A. millepora</em></th>
<th><em>A. vaga</em></th>
<th><em>P. vulgata</em></th>
<th><em>P. dumerilii</em></th>
<th><em>L. salmonis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>16:0</td>
<td>29.0 ± 1.0</td>
<td>31.7 ± 1.0</td>
<td>25.7 ± 0.6</td>
<td>24.4 ± 0.7*</td>
<td>24.8 ± 0.4*</td>
<td>28.5 ± 1.1</td>
</tr>
<tr>
<td>16:1ω7</td>
<td>35.0 ± 1.2</td>
<td>23.3 ± 2.4*</td>
<td>22.4 ± 0.3*</td>
<td>19.5 ± 0.5*</td>
<td>36.4 ± 0.6</td>
<td>34.5 ± 1.0</td>
</tr>
<tr>
<td>16:2ω4</td>
<td>nd</td>
<td>2.8 ± 0.3*</td>
<td>6.9 ± 0.3*</td>
<td>10.4 ± 0.3*</td>
<td>0.4 ± 0.1</td>
<td>0.2 ± 0.0</td>
</tr>
<tr>
<td>16:3ω1</td>
<td>nd</td>
<td>nd</td>
<td>0.5 ± 0.0*</td>
<td>1.5 ± 0.1*</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>18:0</td>
<td>10.1 ± 0.9</td>
<td>15.8 ± 0.6*</td>
<td>15.8 ± 1.2*</td>
<td>17.0 ± 1.0*</td>
<td>10.6 ± 0.3</td>
<td>11.1 ± 0.4</td>
</tr>
<tr>
<td>18:1ω9</td>
<td>25.6 ± 0.6</td>
<td>13.9 ± 2.2*</td>
<td>13.0 ± 1.4*</td>
<td>8.6 ± 0.4*</td>
<td>24.1 ± 1.0</td>
<td>21.2 ± 0.4</td>
</tr>
<tr>
<td>18:2ω6</td>
<td>0.3 ± 0.1</td>
<td>12.5 ± 0.7*</td>
<td>13.8 ± 0.4*</td>
<td>15.1 ± 0.5*</td>
<td>3.8 ± 0.6*</td>
<td>3.3 ± 0.5*</td>
</tr>
<tr>
<td>18:3ω3</td>
<td>nd</td>
<td>nd</td>
<td>1.8 ± 0.1*</td>
<td>3.4 ± 0.2*</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>

†Mostly 16:1ω7, although it also includes some 16:1ω9.

### Table 2. Functional characterization of metazoan ω3 desaturases.

Table 2. Functional characterization of metazoan ω3 desaturases. Transgenic yeast expressing the corresponding ω3 desaturases (*A. millepora* (KY658236), *A. vaga* (ex1, MF448337; ex2, MF448339), *P. vulgata* (KY658241), *P. dumerilii* (KY707886), and *L. salmonis* (KY658239)) were grown in the presence of exogenously added substrates 18:2ω6, 18:3ω6, 20:2ω6, 20:3ω6, 20:4ω6, and 22:4ω6. Conversions are expressed as a percentage of total FA substrate converted to desaturated product, with the corresponding activity (Δ) detected also shown. nd, not detected.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Product</th>
<th>Conversion (%)</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>A. millepora</em></td>
<td><em>A. vaga</em></td>
<td><em>P. vulgata</em></td>
</tr>
<tr>
<td>18:2ω6</td>
<td>18:3ω3</td>
<td>70.5</td>
<td>10.6</td>
</tr>
<tr>
<td>18:3ω6</td>
<td>18:4ω3</td>
<td>64.8</td>
<td>14.3</td>
</tr>
<tr>
<td>20:2ω6</td>
<td>20:3ω3</td>
<td>15.1</td>
<td>1.1</td>
</tr>
<tr>
<td>20:3ω6</td>
<td>20:4ω3</td>
<td>25.7</td>
<td>4.2</td>
</tr>
<tr>
<td>20:4ω6</td>
<td>20:5ω3</td>
<td>32.0</td>
<td>1.0</td>
</tr>
<tr>
<td>22:4ω6</td>
<td>22:5ω3</td>
<td>2.8</td>
<td>nd</td>
</tr>
</tbody>
</table>
**MATERIALS AND METHODS**

**Retrieval of o xo desaturase homologs**

To examine the evolutionary diversity of o xo desaturases among metazoan phyla, we used a combination of strategies: extensive homologous sequence searches, rigorous o xo desaturase sequence selection, and phylogenetic reconstruction. Specifically, we conducted an extensive search of genomic and transcriptomic resources covering species from all domains of life, with a particular focus on Metazoa. Given the vast taxonomic and genetic variability of the “oxo desaturases” (8), homologous sequences were collected by applying similar but complementary methodologies, including HMMER, PSI-BLAST, and HHblits. Query sequences consisted of 113 representative prokaryotic and eukaryotic o xo desaturase protein sequences (19). First, the query sequences were aligned with hmmalign (HMMER software v3.1b2, hmmer.org) to the Pfam Fatty Acid Desaturase Family Profile (FA_desaturase PF00487.22). The resultant multiple sequence alignment (MSA) was then used to search for similar sequences through the HHMER3 web server (https://toolkit.tuebingen.mpg.de/#tools/hmmer) (32) against the nonredundant GenBank protein database (nr, E-value < 1 × 10^-10). Using this approach, around 14,000 distinct hits were selected. Second, the MSA was used to build a position-specific scoring matrix for a PSI-BLAST search (v.2.6.0+; E-value < 1 × 10^-10; 11 iterations) against the nr database. In total, this search retrieved around 56,000 sequences. Third, the query sequences were used as input in a Homology detection by iterative HMM-HMM comparison (HHblits) (33) against the UniProt20 HMM database using the following nondefault options: maximum number of iterations = 8 and realign queries template alignments with Maximum Accuracy alignment algorithm (threshold ≥ 0.1). With HHblits, a total of 520 hits were collected. Finally, all hits were combined and duplicates were removed, resulting in an initial data set of 58,000 sequences.

**Filtering out false-positive and incomplete o xo desaturase gene sequences**

FA desaturase sequences are characterized by the presence of three conserved histidine box motifs (H-box1, H-box2, and H-box3) that typically contain eight histidines between them (19). The distance between histidine boxes (8) and their specific amino acid sequences are characteristic of each subfamily of membrane-bound desaturases including first, o xo, front-end, and sphingolipid desaturases (19). To identify putative o xo desaturase sequences, an automated genomic pipeline was built to filter out non-oxo desaturase sequences. According to the filter thresholds, sequences matching the following criteria were further considered: The FA_desaturase (PF00487.22) Pfam domain must be identified (P < 1 × 10^-10) within the protein sequence, but the Pfam domain Lipid_DES (PF08557.8) should not be present because it is typically associated with sphingolipid desaturases; H-box1 had to be composed of five amino acid residues (HXXXH), which excluded Scd/Δ9 desaturases—HXXXXH; H-box2 had to be composed of five amino acid residues (HXXXXH), which excluded front-end desaturases—HXXXXH; H-box3 had to consist also of five amino acid residues (HXXHH); the number of amino acid residues between the end of H-box1 and the start of H-box2 had to be between 120 and 250. To distinguish o xo from sphingolipid desaturases, the empirical frequency of amino acid occurrence within the histidine boxes of known eukaryotic sequences (19) was used to calculate the probability of a desaturase sequence being assigned to one of these two subfamilies. For example, the third and fourth amino acids in H-box1 of o xo and sphingolipid show very distinctive frequencies (19), with o xo desaturases typically having a cysteine followed by a glycine, whereas sphingolipid desaturases have an isoleucine and a serine. Only putative desaturase sequences with a probability ≥0.7 of being o xo (versus sphingolipid) were retained for further analyses. From the initial putative 58,000 homologous sequences, around 9000 passed all filtering steps (around 15.7%).

**Metazoan o xo desaturase homolog search expansion**

To expand the number of distinct metazoan o xo desaturase sequences, query metazoan sequences were searched against the following GenBank databases with BLAST (version 2.6.0+); whole genome shotgun (WGS), transcriptome shotgun assemblies (TSA) sets and tsa_nr. Specifically, the tblastn program was used against 515 WGS and 935 TSA projects from metazoan nonchordate species, and blastp was used to search metazoan o xo desaturase homologs against tsa_nr (E-value < 1 × 10^-10). In addition, species-specific genome/transcriptome projects were also investigated (for example, P. vulgata). Genome WGS contigs containing significantly similar sequences to o xo desaturases were annotated with corresponding transcripts obtained from the TSA and Sequence Read Archive (SRA) except two Radix auricularia sequences (MUCZC01000242 and MUCZC01000057), which were annotated by tblastn using o xo desaturase homologs from other Heterobranchia species as queries. For each transcript, we selected the longest open reading frame (ORF) sequence with a minimum length of 159 amino acids (to guarantee that the three histidine boxes were present in the sequence). All selected TSA and tsa_nr sequences were then evaluated using the filters described above. In total, 41 additional metazoan sequences were found with this step.

**Sequence number reduction: Clustering, isoforms, and identical sequence removal**

To reduce the total number of sequences for the final data set, the software package SiLix v.1.2.10 (34) was applied to the nonmetazoan sequences (based on BLAST “all against all” pair of sequences alignments) for clustering sequences that shared a minimal pair-sequence identity and overlap of 70%. For nonmetazoan taxa, one representative of each taxonomic rank (for example, order or phylum) found in each cluster was selected for further analyses. The total number of sequences after clustering analysis was reduced to 910 and further reduced to 892 after removing identical sequences and splice isoforms (longest one for each species was kept). The sequences selected by IsoSVM (892) were then aligned with MAFFT v.7.309 (reference, parameters: –genafpair –maxiterate 1000), and 100 bootstrap maximum likelihood (ML)–based phylogenetic trees were inferred using ExaML tool v.3.0.18 (35). The ExaML trees were inferred given 100 precomputed starting trees using standard RaxML v.8.2.9 (36) following the ExaML manual instructions (https://github.com/stamatak/ExaML/blob/master/manual/ExaML.pdf). The majority-rule consensus ExaML tree was used to select sequences representing distinct well-supported clades of major taxonomic ranks (for example, keeping only major Dikaria fungal lineages) and to exclude potential contaminations. A last data set reduction (down to 287 sequences) was made by excluding the sequences with pairwise identity below 29% and overlap below 70%.

**Phylogenetic analyses**

The final sequences were aligned with MAFFT (parameters: –genafpair –maxiterate 1000) and filtered with GUIDANCE v.2.0 (37) with the following parameters: 100 bootstrap replications, sequence, column, and...
site masking cutoff threshold <0.5. The resultant MSA was then filtered to remove columns with gaps in 95% or more of the sequences. The best-fit model of protein evolution was selected using ProtTest v.3.3 (38) based on the Bayesian Information Criterion. The ML phylogenetic inference was performed using RAxML with 100 rapid bootstrap replicates and 20 ML searches. The Bayesian inference was performed using MrBayes v.3.2.1 (39) and PhyloBayes MPI v.1.7a (40). MrBayes was run with four Markov chains incrementally heated with the default values. Each chain started with a randomly generated tree and ran for 1x107 generations with sampling frequency of 1 tree for every 100 generations. The posterior distribution of trees, after discarding the first 25% as burn-in, was summarized in a 50% majority-rule consensus tree. Two independent replicates were conducted and inspected for consistency. The PhyloBayes analysis was executed applying distinct compositional profiles and one substitution rate matrix inferred from the data (CAT+GTR model). Similar to the MrBayes analysis, two independent chains were run in parallel, which terminated reaching moderate convergence after around 16,000 cycles with maximum discrepancy observed across all bipartitions of 0.28. The final protein sequence alignment and the various phylogenetic trees (text and image content) were calculated as \[ \frac{\text{product areas}}{\text{product areas} + \text{substrate area}} \times 100 \] (41). Identity of the desaturation products was determined by preparing 4,4-dimethyloxazoline (DMOX) derivatives of FAME samples followed by gas chromatography–mass spectrometry analysis (41).

**Reagents**

For functional characterization assays, all FA substrates were purchased from Nu-Chek Prep Inc. except 20:4o6, which was from Cayman Chemical. Chemicals used to prepare the *S. cerevisiae* minimal medium uracil were from Sigma-Aldrich, except for the bacteriological agar obtained from Oxoid Ltd. Thin-layer chromatography (20 × 20 cm × 0.25 mm) plates precoated with silica gel 60 (without fluorescent indicator) were purchased from Merck. All solvents were HPLC (high-performance liquid chromatography)–grade and were from Fisher Scientific. DMOX derivatization reagents (2-amino-2-methyl-1-propanol) were purchased from Sigma-Aldrich.

**Statistics**

Comparisons of means of FA composition between control yeast (n = 6) and each transgenic yeast (n = 4) were carried out using Dunnett’s test with P < 0.05 indicating statistical significance. All statistical analyses were performed in R 3.4.1. (www.r-project.org).

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/4/5/eaaar6849/DC1

**REFERENCES AND NOTES**


7 of 8


Acknowledgments: We thank A. Magurran and J. Napier for comments on the manuscript and R. Ruivo for drawings in Figs. 1 and 3. Funding: This work received funding from the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland) funded by the Scottish Funding Council (grant reference HR09011), and their support is gratefully acknowledged. Access to the Institute of Aquaculture laboratories was funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 262336 (AQUAEXCEL), Transnational Access Project Number 0009/06/03/13. Author contributions: M.M.F., D.E.K.F., L.K.B., and L.F.C.C. designed and performed the overall sequence screening strategy. N.K., D.S.F., and D.E.K.F. retrieved the sequences of functionally characterized genes. M.M.F. performed the phylogenetic analyses. O.M., N.K., M.M.F., D.E.K.F., L.F.C.C., and D.R.T. designed the experiments, interpreted the results, and wrote the manuscript. N.K. and Ó.M. performed the cDNA cloning and functional characterization assays. N.K. and J.C.N. analyzed the yeast FAs. All authors contributed to the manuscript with intellectual input and approved the final version. Competing interests: The authors declare that they have no competing interests. Data and materials availability: The list of animal WGS (513) and TSA (935) projects searched with BLAST, the list of final data set sequences used, the protein alignment, and all phylogenetic trees are accessible through figshare.com (https://figshare.com/s/7e518eb36e3496f9281) and the CiMAR website (https://box.cimar.up.pt/owncloud/index.php/s/0dWV3Z37R3mjX/). The sequences functionally characterized in this work have been deposited in GenBank with the accession codes listed in table S1. All other data are available from the corresponding authors upon reasonable request.

Submitted 6 December 2017
Accepted 13 March 2018
Published 2 May 2018
10.1126/sciadv.aar849

Genes for de novo biosynthesis of omega-3 polyunsaturated fatty acids are widespread in animals

DOI: 10.1126/sciadv.aar6849