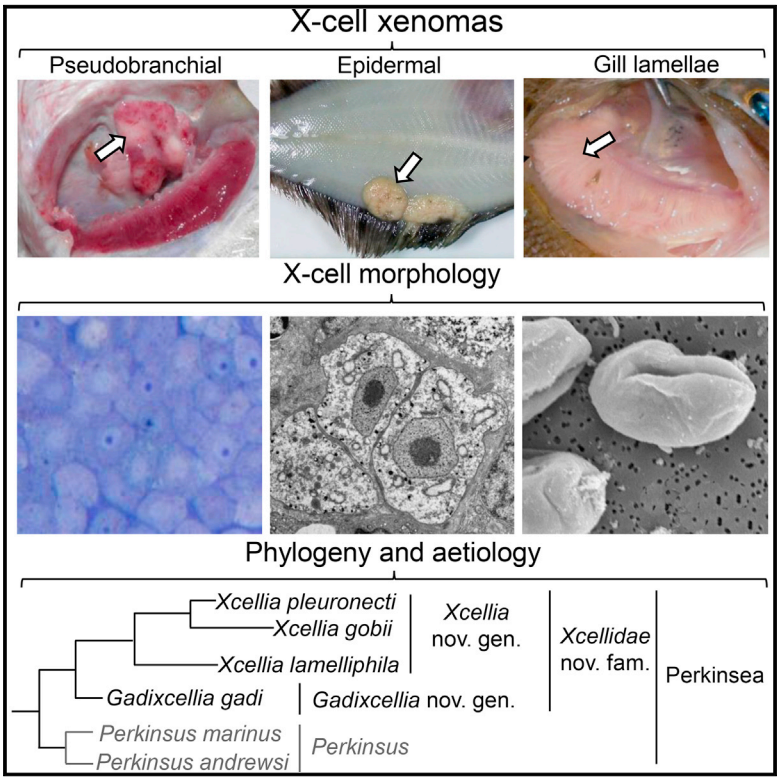
Report

Current Biology

X-Cells Are Globally Distributed, Genetically Divergent Fish Parasites Related to Perkinsids and Dinoflagellates

Graphical Abstract



Highlights

- X-cells group in two highly divergent clades, the Xcellidae, sister to *Perkinsus*
- Creation of the genera *Gadixcellia* and *Xcellia* to accommodate these fish parasites
- Xcellidae nov. fam. are the first perkinsids known to cause pathology in fish
- Evidence that these parasites are important juvenile fish pathogens

Authors

Mark A. Freeman, Janina Fuss, Árni Kristmundsson, ..., Patrick J. Keeling, Kamran Shalchian-Tabrizi, David Bass

Correspondence

mafreeman@rossvet.edu.kn

In Brief

Freeman et al. show that "X-cells" causing disease and parasitic xenomas in fish are related to the alveolate protist *Perkinsus*, a serious pathogen of shellfish. X-cells are found globally in many teleost fish families, including cod and flatfish, and may particularly affect juveniles. Two new genera are created: *Gadixcellia* and *Xcellia*.

Freeman et al., 2017, Current Biology *27*, 1645–1651
June 5, 2017 © 2017 Elsevier Ltd.
http://dx.doi.org/10.1016/j.cub.2017.04.045

Current Biology
Report

X-Cells Are Globally Distributed, Genetically Divergent Fish Parasites Related to Perkinsids and Dinoflagellates

Mark A. Freeman,[1,8,*] Janina Fuss,[2] Árni Kristmundsson,[3] Marit F.M. Bjorbækmo,[2] Jean-François Mangot,[2] Javier del Campo,[4] Patrick J. Keeling,[5] Kamran Shalchian-Tabrizi,[2] and David Bass[6,7]
[1]Center for Conservation Medicine and Ecosystem Health, Ross University School of Veterinary Medicine, PO Box 334, Basseterre, St. Kitts, West Indies
[2]Section for Genetics and Evolutionary Biology (EVOGENE), Department of Biosciences, University of Oslo, 0316 Oslo, Norway
[3]Institute for Experimental Pathology at Keldur, University of Iceland, 112 Reykjavik, Iceland
[4]Botany Department, University of British Columbia, 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
[5]Canadian Institute for Advanced Research, Botany Department, University of British Columbia, 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
[6]Centre for Environment, Fisheries and Aquaculture Science (Cefas), Barrack Road, The Nothe, Weymouth, Dorset DT4 8UB, UK
[7]Department of Life Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, UK
[8]Lead Contact
*Correspondence: mafreeman@rossvet.edu.kn
http://dx.doi.org/10.1016/j.cub.2017.04.045

SUMMARY

"X-cells" have long been associated with tumor-like formations (xenomas) in marine fish, including many of commercial interest. The name was first used to refer to the large polygonal cells that were found in epidermal xenomas from flatfish from the Pacific Northwest [1]. Similar looking cells from pseudobranchial xenomas had previously been reported from cod in the Atlantic [2] and Pacific Oceans [3]. X-cell pathologies have been reported from five teleost orders: Pleuronectiformes (flatfish), Perciformes (perch-like fish), Gadiformes (cods), Siluriformes (catfish), and Salmoniformes (salmonids). Various explanations have been elicited for their etiology, including being adenomas or adenocarcinomas [4, 5], virally transformed fish cells [6–8], or products of coastal pollution [9, 10]. It was hypothesized that X-cells were protozoan parasites [1, 11–13], and although recent molecular analyses have confirmed this, they have failed to place them in any phylum [14–18], demonstrating weak phylogenetic associations with the haplosporidians [16] or the alveolates [15]. Here, we sequenced rRNA genes from European and Japanese fish that are known to develop X-cell xenomas. We also generated a metagenomic sequence library from X-cell xenomas of blue whiting and Atlantic cod and assembled 63 X-cell protein-coding genes for a eukaryote-wide phylogenomic analysis. We show that X-cells group in two highly divergent clades, robustly sister to the bivalve parasite *Perkinsus*. We formally describe these as *Gadixcellia* and *Xcellia* and provide a phylogenetic context to catalyze future research. We also screened Atlantic cod populations for xenomas and residual pathologies and show that X-cell infections are more prevalent and widespread than previously known.

RESULTS

The Phylogenetic Position of X-Cells

Sixty-three protein coding genes (Table S1) were mined from a metagenomic dataset generated from X-cell pseudobranchial xenomas from cod, *Gadus morhua*, sampled from Iceland, and from blue whiting, *Micromesistius poutassou*, from Norway. A phylogenomic analysis showed that these X-cell lineages were closely related to each other and were, together, sister to *Perkinsus marinus* (Alveolata) with maximal support (Figure 1). Concatenated small- and large-subunit (SSU and LSU, respectively) rRNA gene phylogenies also showed this relationship and that there are at least two highly distinct X-cell clades, one containing pseudobranchial parasites of Gadiformes (cod and blue whiting), and the other, much longer branched, containing gill and epidermal X-cells from Perciforms and Pleuronectiformes (Figure S1A). These clades are robustly sisters to each other, and we have created a new family to contain them: Xcellidae nov. fam. (see the Supplemental Information). As 18S rRNA is the most comprehensively sampled taxonomic marker gene, we searched GenBank for all related sequence types to see whether X-cells have previously been detected in organismal and environmental samples and whether they and perkinsids have any mutual relatives, but none were found (Figure S2B). To investigate the diversity and distribution of X-cells in environmental samples, we screened VAMPS [19], BioMarKs [20], and Tara Oceans [21], representing $>10^9$ high-throughput sequencing (HTS) reads from many marine studies, for X-cell sequences: only a single example was found, a 119 bp SSU V9 region OTU (operational taxonomic unit) 98.3% similar (100% match coverage) to an X-cell specimen sequenced from cod (GU296508) from an

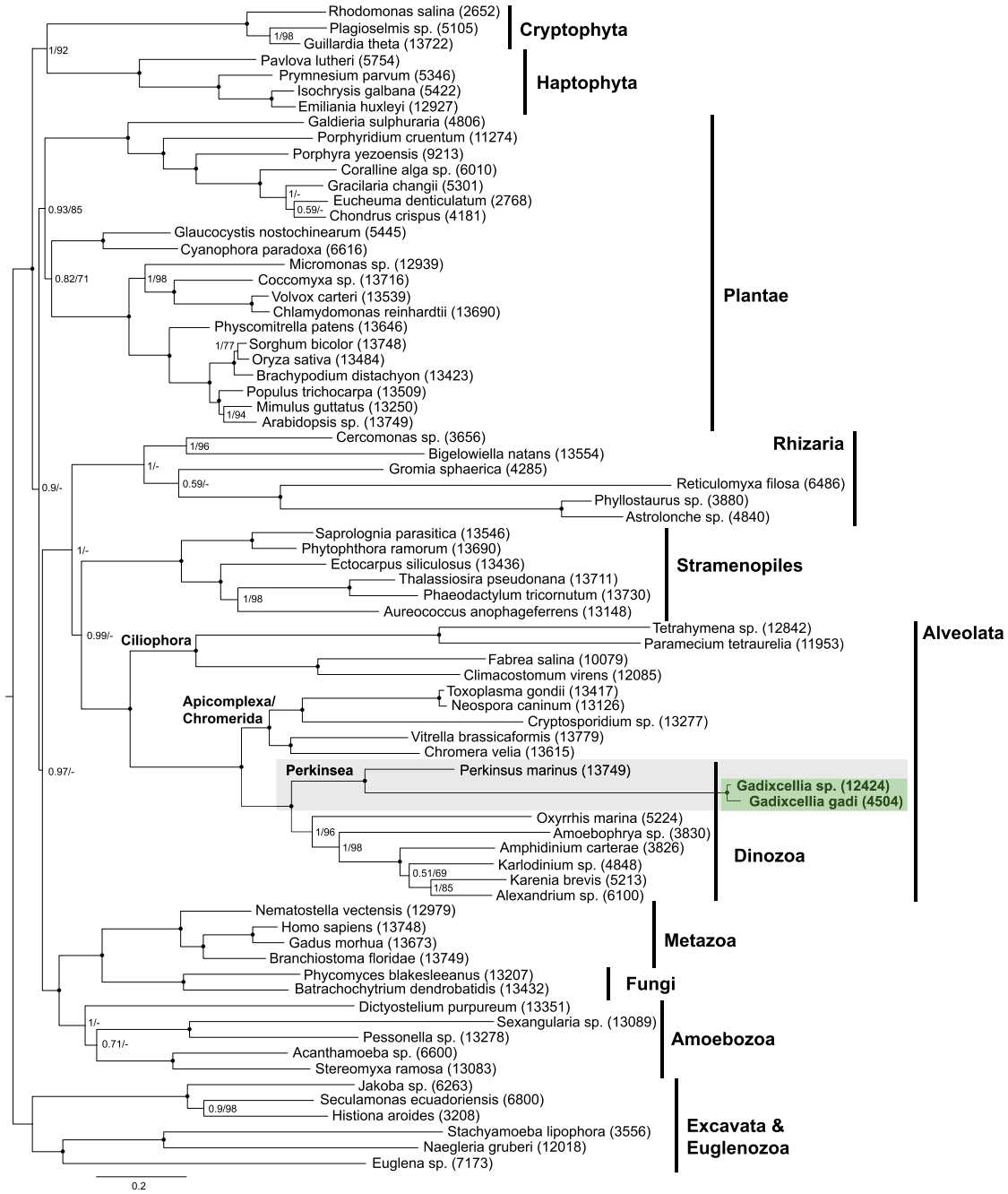


Figure 1. Phylogenomic Analysis of *Gadixcellia*

A global eukaryote tree based on 63 proteins and from 74 taxa is shown (Bayesian analysis with PhyloBayes). *Gadixcellia* spp. (green box) are robustly grouped with *Perkinsus* (Perkinsea) (gray box) within the Dinozoa. Black dots indicate maximal support for both posterior probabilities (1.0) and maximum-likelihood bootstraps (100%) at the respective nodes. On nodes with lower support values, they are indicated in the following order: PhyloBayes/RAxML. Total number of amino-acid positions present in the alignment for each species is indicated after the species name in parentheses.

See also Figure S1 and Table S1.

oceanic bathypelagic zone (2,007 m in depth) in the Gulf of California [22]. Now that they are phylogenetically characterized, diagnostics may be developed to confirm that fish pathologies observed in many species around the world are caused by X-cells.

X-Cells in North Atlantic Gadoids

Sequence data for the shorter branched lineage, to which we give the name *Gadixcellia* nov. gen. (taxonomic treatment is given in the Supplemental Information), are, so far, restricted to members of the cod family in the Northern Atlantic (cod and

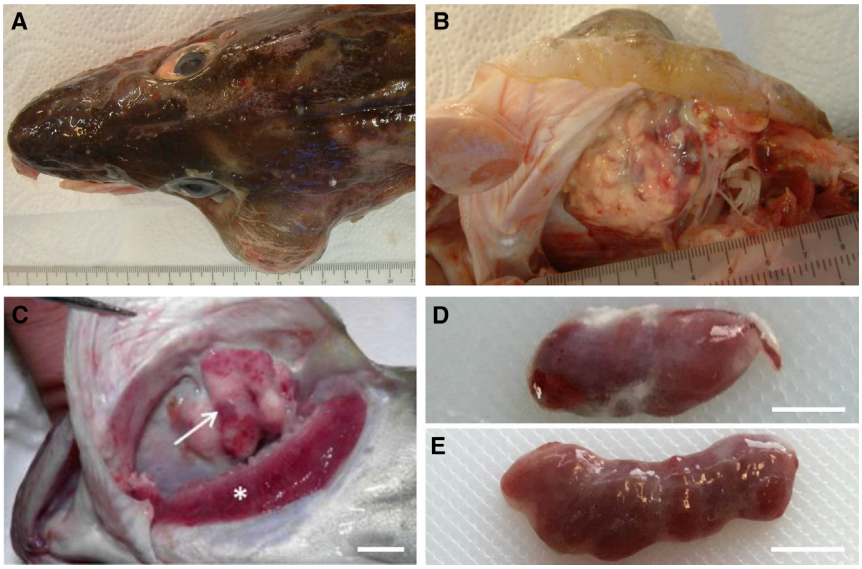


Figure 2. Gross Appearance of Pseudobranchial Xenomas in Atlantic Cod, from Iceland, Caused by *Gadixcellia gadi*

(A) Large unilateral xenoma about 40 mm in diameter.

(B) Xenoma viewed from under the operculum with no visible pseudobranchial tissue remaining.

(C) Significant xenoma causing disruption of the pseudobranch (white arrow), while the gill arches remain uninfected (white asterisk). Scale bar, 10 mm.

(D and E) Excised pseudobranchs from Atlantic cod; normal looking (D) and deformed and nodular (E). Scale bars, 10 mm.

See also Figure S2.

blue whiting). The cod-derived lineage (*G. gadi* nov. sp.) causes pseudobranchial xenomas that sometimes extend into the gill or operculum (Figures 2A–2C). Pseudobranchial xenomas caused by *G. gadi* were found in 2%–3% of Atlantic cod ranging from 30 to 75 cm (n = 300), caught in one day's research expedition north of Iceland, compared to only 0.03% in cod ranging from 60 to 120 cm (n = 3,200) caught off West Iceland. During the Norwegian cruise from Kirkenes to Tromsø, there were 566 Atlantic cod caught, of which four were visibly infected with X-cells (0.71%). Evidence of historical/recovered *G. gadi* infections, characterized by abnormally shaped and nodular pseudobranchs (Figures 3D and 3E), was seen in numerous cod individuals, ranging from 4.5% in younger (30–75 cm) fish from the north of Iceland to 6% in larger ones (60–120 cm) caught west of Iceland.

Blue whiting were trawled from six locations off the west coast of Iceland (n = 212). Four of 70 fish from only one of these locations had pseudobranchial xenomas. From Norway, blue whiting were caught at 6 of 22 trawling locations during a research expedition from Kirkenes to Tromsø. Only eight blue whiting were caught in total during the cruise. At one of the locations (71.295′ N, 26.323″ W) three blue whiting were caught, two of which had pseudobranchial xenomas. Two distinct *Gadixcellia* genotypes were found in blue whiting from Iceland, one of which, type I, was 98% similar, with respect to SSU rDNA, to *G. gadi* from Atlantic cod, whereas type II, possibly a different species, was only 91.2% similar (Figure S2B; Table S2). The Norwegian blue whiting pseudobranchial xenoma used in the phylogenomic study was type I.

X-Cells from Flatfish, Eelpout, Icefish, and Gobies

Members of the very long-branched clade of X-cells (*Xcellia* nov. gen.; taxonomic treatment is given in the Supplemental Information) cause either X-cell masses to form between the secondary gill lamellae, which can lead to gill fusion and loss of function, or epidermal xenomas, often on the fins and operculum (Figure S2). The former pathology is seen in eelpout, common dab, and icefish. Of the six species of eelpout, *Lycodes* spp. (Zoarcidae), sampled (total number, 220; 4–85 fish per species) north of Iceland, only one species, *L. seminudus* (4 out of 20 individuals infected), had visible X-cell gill masses/lesions. Three of these fish were trawled from a single location. The other five species of *Lycodes* did not have noticeable X-cell infections. The eelpout X-cell SSU sequences were almost identical (>99.5%) to other gill lesion X-cells from common dab, *Limanda limanda*, and Antarctic cod icefishes, *Trematomus* spp., and are assigned to the new species *Xcellia lamelliphila* nov. sp. (Supplemental Information).

A second *Xcellia* clade, sister to the eelpout-dab-icefish X-cells but with only 86%–89% SSU sequence similarity to them (Figure S1B; Table S2), was represented by parasites causing epidermal xenomas in the flathead flounder (*Hippoglossoides dubius*), the northern black flounder (*Pseudopleuronectes obscurus*), and the Japanese goby (*Acanthogobius flavimanus*), all sampled from Japan, for which we create the two new species *X. pleuronecti* nov. sp. and *X. gobii* nov. sp. (taxonomic treatment is given in the Supplemental Information).

X-Cell Morphology, Histopathology, and Ultrastructure

Histological examination and scanning electron microscopy showed parasite cells in very high numbers in the xenomas (Figures 3A and 3B). Normal gadoid pseudobranchial tissue has a continuous lamellar appearance (Figure 3C), but when infected or in recovery, the lamellar organization breaks down and is separated by fibrotic tissue (Figure 3D). Some visibly deformed and nodular pseudobranchs showed indications of previous X-cell infection and a subsequent recovery process (Figure 3D). Occasionally, degenerate X-cells were observed within these areas, among fibrotic tissue and recovering pseudobrancial tissues (data not shown). *Xcellia* cells in epidermal xenomas were limited to the epidermal layer of the skin and did not penetrate into the dermal layer or musculature. However, *Xcellia* xenomas often extended to the edge of the fins, where they appeared as sac-like structures and where masses of X-cells were found in compartments surrounded by connective tissue [15]. Remarkably, no additional developmental stages (flagellated stages, spores/environmental stages) were seen in any X-cell preparations.

Transmission electron microscopy revealed numerous similarities in X-cells from all hosts. All X-cells had a relatively large nucleus with fine granular chromatin and a prominent electron-dense nucleolus (Figure 4). X-cells were often tightly packed together, suggesting recent division (Figure 4B); however, actively

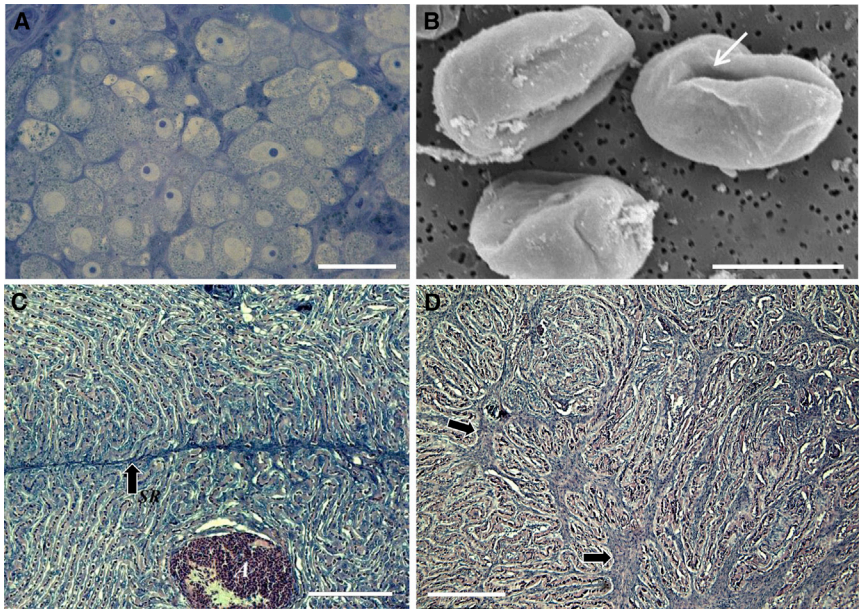


Figure 3. Histology and SEM of *Gadixcellia gadi*

(A) A semithin section from a xenoma from Atlantic cod showing numerous X-cells. Scale bar, 10 μm.

(B) SEM of X-cells reveals no discernible external features except for a groove (white arrow), possibly caused by shrinkage during sample preparation, as no other features are associated with them. Scale bar, 10 μm.

(C) Giemsa-stained normal pseudobranchial tissue with a continuous lamellar structure and a support rod (*SR*) and artery (*A*). Scale bar, 300 μm.

(D) An abnormal/nodular pseudobranch with severe interlamellar fibrosis suggestive of recovery from prior infection with *G. gadi*. Scale bar, 200 μm.

dividing cells were not encountered, and there was an absence of centriole formation, suggesting that division is amitotic. The cytoplasm contained numerous large mitochondria, usually with limited inner membrane folding, but cristae were tubular or ampulliform (Figure 4; Figure S3A). Other cytoplasmic organelles/vesicles were present, often in large numbers, which were sometimes associated with the plasma membrane and had the appearance of pseudopodia-like appendages (Figure 4; Figure S3C). Cortical alveolae were sometimes seen (Figure 4A) but were not always a prominent feature, perhaps reduced or secondarily lost in some lineages. Multinucleate plasmodial X-cells (*Gadixcellia*) were only found in Atlantic cod xenomas (Figure 4C) and not in other fish species (unconfirmed in blue whiting).

DISCUSSION

Members of the new family Xcellidae share a range of phenotypic characteristics, such as the formation of xenomas in fish and some remarkable similarities in cell morphology and content in histological sections. However, the genetic variation between the genera *Gadixcellia* and *Xcellia* is unusually high, with similarity in the SSU rDNA as low as 74.9% (Table S2). The even greater genetic distances between X-cells and other protist taxa have made this unusual group of fish parasites hard to definitively place in previous, more limited, phylogenetic studies. In the present study, the phylogenetic position of X-cells is resolved for the first time using a phylogenomic approach, as a long-branched sister lineage to *Perkinsus*. The perkinsids comprise a large diversity of SSU sequence types, many of which are detectable in environmental samples [23, 24]. However, even in a recent large-scale, perkinsid-focused coastal marine habitat sequencing survey of the V4 SSU gene region, X-cell sequences were not detected [24], possibly because their divergent SSU genes are not amplified with "general" eukaryotic primers, although V9 primers [21] should amplify them. An alternative/additional reason why X-cell sequences do not occur in environmental sequence datasets could be that X-cells are tightly host associated and/or transmitted directly from fish to fish or via sediment reservoirs [15, 25]. Therefore, knowledge of X-cells is currently limited to direct sampling of individual fish, infection rates in which are generally low but apparently very patchy in spatial occurrence. This patchy distribution of X-cell-infected fish—reported here for blue whiting and eelpout and previously reported for common dab in the Atlantic [26] and for numerous flatfish from Hokkaido, Japan [27]—again supports the theory that transmission to fish may be dependent upon certain conditions, e.g., substrate types. There are parallels here with other highly divergent parasites; in particular, the recently described *Paramikrocytos* [28], which, despite being locally extremely prevalent, causing intense infections in crabs, was not detectable in the surrounding water using either highly specific or broadly targeted PCR primers.

Some perkinsid lineages are known to be parasitic: *Perkinsus* spp. in bivalve mollusks (e.g., X75762; Figure S1B), *Parvilucifera* in dinoflagellates (KF395485), a pathogen of the southern leopard frog *Rana sphenocephala* (EF675616) [29], and *Rastrimonas* [30] (no SSU data were available). The lifestyles of the many other lineages, so far detected only in environmental sequencing studies, are unknown, although they are suspected to be parasitic. For example, SSU types closely related to an abundant diversity of environmental lineages and the *R. sphenocephala* pathogen were recently also detected in frog liver tissue [31]. Resolving the phylogenetic position of X-cells shows that the two now-known perkinsid parasites of vertebrates are not specifically related to each other, X-cells being more closely related to parasites of bivalves than to the frog parasite. It is likely that other—perhaps most—perkinsid lineages will be shown to have at least a symbiotic stage in their life cycles, collectively with a potentially very wide host range including other vertebrates. X-cells do not show any particular morphological similarities with perkinsids, which have flagellate stages, release hypnospores on killing the host bivalve, and exhibit local cell clusters rather than the cell masses seen in X-cells. Even when mature X-cell xenomas are observed detached from the fish, no further developmental stages are seen (data not shown). The massive X-cell nucleolus is actually more similar to those seen in apicomplexan macrogamonts, e.g., *Pseudoklossia*/*Margolisiella* spp. infecting mollusks [32, 33].

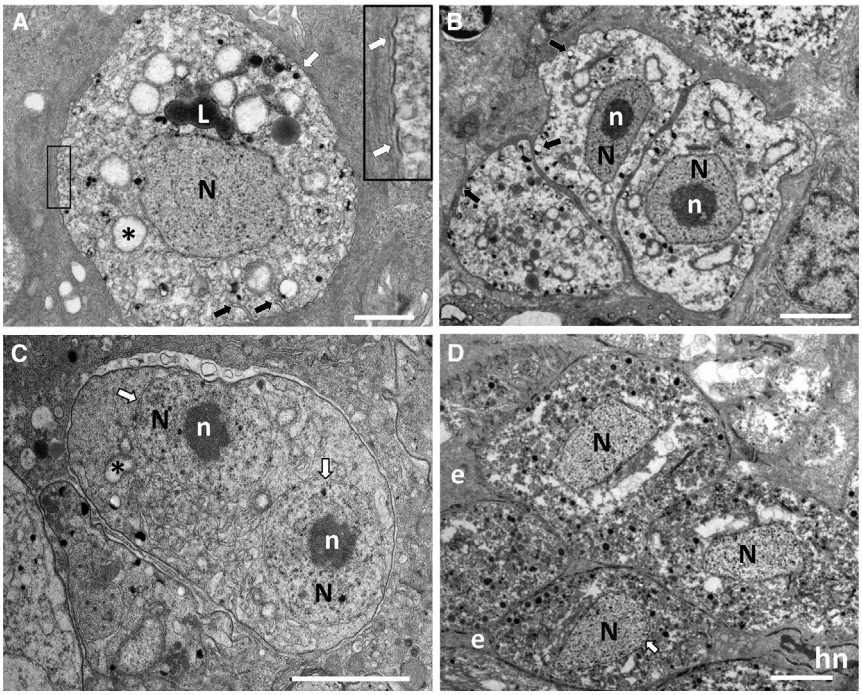


Figure 4. TEM of *Xcellia* spp. and *Gadixellia gadi*

(A) *Xcellia pleuronecti* from northern black flounder surrounded by a complex of host envelope cells, microtubules, and desmosomes. The parasite has a large central nucleus (N) and numerous mitochondria (black asterisk) and lipid droplets (L) in the cytoplasm. Note the cortical alveoli supporting the plasma membrane (white arrows and inset), that in places forms large micropore-like structures (black arrows). Scale bar, 1 μm.

(B) A group of *Xcellia pleuronecti* cells, probably products of division, note the prominent nucleoli (n) and numerous small vesicles associated with the plasma membrane (black arrows). Scale bars, 2 μm. N, nucleus.

(C) *Gadixellia* cell with two nuclei (N) enclosed with by nuclear membranes (arrows), both with relatively large nucleoli (n) and mitochondria in the cytoplasm (black asterisk). Scale bar, 2 μm.

(D) A cluster of *Xcellia lamelliphila* from common dab with large nuclei (N), surrounded by host envelope cells (e). Note the host nucleus (hn) and high numbers of cytoplasmic vesicles in the X-cell cytoplasm. Scale bar, 3 μm.

See also Figure S3.

We suggest that infection by X-cells occurs via contact of fish with the benthos. All known X-cell infections involve fish species with at least one benthic stage in their life cycle. Wild cod are initially pelagic but become benthic when 5–8 cm long [34, 35]. According to Eydal et al. [36], the first signs of macroscopically visible pseudobranchial xenomas occur in fish at 6 months of age (then, 6.5–13 cm long), with a subsequent increase in prevalence, reaching a peak of 25% at age 22 months. Furthermore, Eydal et al. [36] compared 4-month-old cod (at that time, 3.5–7.5 cm and, therefore, with limited benthic contact) placed into tanks with those straight from a hatchery. None of the latter showed X-cell infection, but a low proportion of the former were infected before entering the tanks and subsequently died. No fish taken straight from the hatchery to an equivalent tank were infected. Fish from both tanks were then placed into cages suspended above the benthos, and none succumbed. Therefore, only fish exposed, even transiently, to the benthos became infected. However, most details of the X-cell life cycle remain unknown. A planktonic stage seems unlikely, given that no environmental sequence surveys, including limited screens with X-cell-specific primers, have amplified X-cells, except the single sequence from the Gulf of California. This is consistent with the lack of observation of flagellate or spore-like stages in histological/cytological studies. It is also noteworthy that, with few exceptions, all of the infected, wild-caught juveniles died during rearing in land-based tanks [36]. Samples usually examined for X-cell infection comprise adult fish, in which infection levels are generally low (though locally high) or show moderate (6%) levels of residual pathology. Considering the apparent decrease in visible pseudobranchial xenomas with age of Atlantic cod, it is possible that X-cells cause far more significant juvenile mortality of at least some fish species than is currently realized.

X-cell parasites taken from epidermal xenomas from different flatfish species in northern Japan have the same SSU rDNA sequences as each other [15] but are different from those causing epidermal xenomas in the Japanese goby from similar geographical locations [17]. This indicates that fish X-cells are not host species specific but may be restricted to certain fish families, such as Pleuronectidae and Gobidae. However, SSU rDNA sequences from X-cell lesions from the gills of the common dab in Europe are genetically distinct from those of the epidermal X-cell parasites of other flatfish species [14], implying that tissue tropism is also an important factor [15]. Interestingly, X-cells causing very similar gill pathology in the Antarctic cod icefish from the family Nototheniidae have been shown to have almost identical SSU sequences to those from X-cell gill lesions in the common dab, suggesting that this particular X-cell parasite has a wide distribution and can infect a range of fish species from different family groups [18].

X-cells are known, with DNA confirmation, from the northern Atlantic, Pacific, and Southern (or Antarctic) Oceans. However, their actual distribution is clearly limited by sampling effort, compounded by the difficulty of detecting these parasites independently of infected hosts. Many remain uncharacterized; for example, X-cell-like pathologies have been described in Siluriformes (catfish), Salmoniformes (salmonids), and numerous other Perciformes but have not yet been sequenced. They are also more prevalent than direct observations suggest: the data from juveniles and residual pathology results from adult cod are strongly suggestive of a higher disease incidence earlier in the fishes' life cycle, and our hypothesis that X-cells cause significant undetected mortalities in juvenile/young fish should be tested. Identifying reliable environmental reservoirs for X-cells (for example, alternate hosts, sediment bands in gyre systems) would elucidate X-cell life cycles and greatly facilitate research into the impact and etiology of this new taxon of neglected fish pathogens.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Trawling of fish to screen for xenomas
- METHOD DETAILS
  - Sequence generation
  - Data analysis
  - Microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Phylogenetic analysis of ribosomal RNA gene sequences
  - Phylogenomic analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Results and taxonomic treatments, three figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.04.045.

AUTHOR CONTRIBUTIONS

M.A.F. sampled X-cells in Japan, Scotland, and Iceland; provided histopathology, scanning electron microscopy (SEM)/transmission electron microscopy (TEM) micrographs, and the SSU and LSU data; and drafted the manuscript. A.K. sampled fish in Iceland and performed histological and TEM analyses. P.J.K. helped with the initial stages of the molecular phylogenetic analyses and helped guide the project to completion. J.F. generated sequence data, analyzed the *Gadixcellia* metagenomes, and performed the phylogenomic analyses. M.F.M.B., J.-F.M., and K.S.-T. coordinated sampling and data generated in Norway. J.d.C. mined environmental high-throughput sequence datasets for X-cell sequences. D.B. carried out phylogenetic analyses, guided the project, and co-wrote the manuscript.

ACKNOWLEDGMENTS

This study has been supported financially by the Research Council of Norway (contract Havkyst/216475 "X-cell parasites: an emerging threat to marine fish"). The initial study (M.A.F.) was supported by an HIR grant (no. UMC/625/1/HIR/027) from the University of Malaya, Kuala Lumpur, Malaysia. We thank the Norwegian Institute of Marine Research and especially professor Egil Karlsbakk for letting us join their coastal sampling cruise with the research vessel "Johan Hjort." We thank Arthur A.B. Haraldsen, Torbjørn Gylt, and Estelle Grønneberg for help with sampling. We thank Anders K. Krabberød for his valuable feedback on the multigene phylogenies and The Norwegian Sequencing Centre for Illumina HiSeq sequencing. D.B. was supported by NERC grants NE/H009426/1 and NE/H000887/1. J.d.C. was supported by a Marie Curie International Outgoing Fellowship grant (FP7-PEOPLE-2012-IOF – 331450 CAARL) and a grant from the Tula Foundation to the Centre for Microbial Biodiversity and Evolution at UBC. P.J.K. is a Senior Fellow of the Canadian Institute for Advanced Research. We acknowledge the Marine Research Institute in Iceland for their valuable help in sampling.

Received: January 2, 2017
Revised: March 7, 2017
Accepted: April 21, 2017
Published: May 18, 2017

REFERENCES

1. Brooks, R.E., McArn, G.E., and Wellings, S.R. (1969). Ultrastructural observations on an unidentified cell type found in epidermal tumors of flounders. J. Natl. Cancer Inst. *43*, 97–109.
2. Peyron, A., and Thomas, L. (1929). Contribution a l'étude des tumeurs de revetement branchial chez les poissons. Bull. Assoc. Fr. Etud. Cancer *18*, 825–827.
3. Takahashi, K. (1929). Studie uber die Fishgeschwulste. Z. Krebsforsch. *29*, 1–73.
4. Alpers, C.E., Cain, B.B., Myers, M., Wellings, S.R., Poore, M., Bagshaw, J., and Dawe, C.J. (1977). Pathologic anatomy of pseudobranch tumours in Pacific cod, *Gadus macrocephalus*. J. Natl. Cancer Inst. *59*, 377–398.
5. Lange, E. (1973). Carcinoid-like tumours in the pseudobranch of *Gadus morhua* L. Comp. Biochem. Physiol. A Comp. Physiol. *45*, 477–481.
6. Wellings, S.R., Chuinard, R.G., Gourley, R.T., and Cooper, R.A. (1964). Epidermal papillomas in the flathead sole, *Hippoglossoides elassodon*, with notes on the occurrence of similar neoplasms in other pleuronectids. J. Natl. Cancer Inst. *33*, 991–1004.
7. Peters, N., Peters, G., Stich, H.F., Acton, A.B., and Bresching, G. (1978). On differences in skin tumours of Pacific and Atlantic flatfishes. J. Fish Dis. *1*, 3–25.
8. Peters, N., Stich, H.F., and Kranz, H. (1981). On the relationship between lymphocystis disease and X-cell papillomatosis in flatfish. In Phyletic Approaches to Cancer, C.J. Dawe, J.C. Harshbarger, S. Kondo, T. Sugimura, and S. Takayama, eds. (Japan Scientific Societies Press), p. 111.
9. Stich, H.F., Acton, A.B., and Forrester, C.R. (1976). Fish tumors and sublethal effects of pollutants. J. Fish. Res. Board Can. *33*, 1993–2001.
10. Wellings, S.R., McCain, B.B., and Miller, B.S. (1976). Epidermal papillomas in Pleuronectidae of Puget Sound, Washington; review of the current status of the problem. Prog. Exp. Tumor Res. *20*, 55–74.
11. Dawe, J.C. (1981). Polyoma tumors in mice and X-cell tumors in fish, viewed through telescope and microscope. In Phyletic Approaches to Cancer, C.J. Dawe, J.C. Harshbarger, S. Kondo, T. Sugimura, and S. Takayama, eds. (Scientific Societies Press), pp. 19–49.
12. Kent, M.L., Myers, M.S., Wellings, S.R., and Elston, R.A. (1988). An internal X-cell pseudotumor in a black croaker (*Cheilotrema saturnum*). J. Wildl. Dis. *24*, 142–145.
13. Shinkawa, T., and Yamazaki, F. (1987). Proliferative patterns of X-cells found in the tumorous lesions of Japanese goby. Nippon Suisan Gakkai Shi *53*, 563–568.
14. Freeman, M.A. (2009). X-cell parasites in the European dab *Limanda limanda* are related to other X-cell organisms: a discussion on the potential identity of this new group of parasites. Parasitology *136*, 967–980.
15. Freeman, M.A., Eydal, M., Yoshimizu, M., Watanabe, K., Shinn, A.P., Miura, K., and Ogawa, K. (2011). Molecular identification and transmission studies of X-cell parasites from Atlantic cod *Gadus morhua* (Gadiformes: Gadidae) and the northern black flounder *Pseudopleuronectes obscurus* (Pleuronectiformes: Pleuronectidae). Parasit. Vectors *4*, 15.
16. Miwa, S., Nakayasu, C., Kamaishi, T., and Yoshiura, Y. (2004). X-cells in fish pseudotumors are parasitic protozoans. Dis. Aquat. Organ. *58*, 165–170.
17. Miwa, S., and Kamaishi, T. (2009). X-cells in pseudotumors of yellowfin goby *Acanthogobius flavimanus*: a protistan organism distinct from that in flathead flounder *Hippoglossoides dubius*. Dis. Aquat. Organ. *85*, 53–57.
18. Evans, C., and Tupmongkol, K. (2014). X-cell disease in Antarctic fishes. Polar Biol. *37*, 1261–1269.
19. Huse, S.M., Mark Welch, D.B., Voorhis, A., Shipunova, A., Morrison, H.G., Eren, A.M., and Sogin, M.L. (2014). VAMPS: a website for visualization and analysis of microbial population structures. BMC Bioinformatics *15*, 41.
20. Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., Chambouvet, A., Christen, R., Claverie, J.M., Decelle, J., et al. (2015). Marine protist diversity in European coastal waters and sediments as

revealed by high-throughput sequencing. Environ. Microbiol. *17*, 4035–4049.

21. de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., et al. (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science *348*, 1261605.
22. Lie, A.A.Y., Liu, Z., Hu, S.K., Jones, A.C., Kim, D.Y., Countway, P.D., Amaral-Zettler, L.A., Cary, S.C., Sherr, E.B., Sherr, B.F., et al. (2014). Investigating microbial eukaryotic diversity from a global census: Insights from a comparison of pyrotag and full-length sequences of 18S rRNA genes. Appl. Environ. Microbiol. *80*, 4363–4373.
23. Bråte, J., Logares, R., Berney, C., Ree, D.K., Klaveness, D., Jakobsen, K.S., and Shalchian-Tabrizi, K. (2010). Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. ISME J. *4*, 1144–1153.
24. Chambouvet, A., Berney, C., Romac, S., Audic, S., Maguire, F., De Vargas, C., and Richards, T.A. (2014). Diverse molecular signatures for ribosomally 'active' Perkinsea in marine sediments. BMC Microbiol. *14*, 110.
25. Bass, D., Stentiford, G.D., Littlewood, T.D., and Hartikainen, H. (2015). Diverse applications of environmental DNA methods in parasitology. Trends Parasitol. *31*, 499–513.
26. Diamant, A., and McVicar, A.H. (1990). Distribution of X-cell disease in common dab, *Limanda limanda* L., in the North Sea, and ultrastructural observations of previously undescribed developmental stages. J. Fish Dis. *13*, 25–37.
27. Katsura, K., Yamazaki, F., Hamada, K., Oishi, K., Harada, T., and Shinkawa, T. (1984). Geographic distribution and frequency of tumorous fishes collected from the coastal waters of Hokkaido, Japan. Bull. Japan. Soc. Sci. Fish. *50*, 979–984.
28. Hartikainen, H., Stentiford, G.D., Bateman, K.S., Berney, C., Feist, S.W., Longshaw, M., Okamura, B., Stone, D.M., Ward, G., Wood, C., and Bass, D. (2014). Mikrocytids are a broadly distributed and divergent radiation of parasites in aquatic invertebrates. Curr. Biol. *24*, 807–812.
29. Davis, A.K., Yabsley, M.J., Keel, M.K., and Maerz, J.C. (2007). Discovery of a novel alveolate pathogen affecting southern leopard frogs in Georgia: description of the disease and host effects. EcoHealth *4*, 310–317.
30. Brugerolle, G. (2003). Apicomplexan parasite *Cryptophagus* renamed *Rastrimonas* gen. nov. Eur. J. Protistol. *39*, 101.
31. Chambouvet, A., Gower, D.J., Jirků, M., Yabsley, M.J., Davis, A.K., Leonard, G., Maguire, F., Doherty-Bone, T.M., Bittencourt-Silva, G.B., Wilkinson, M., and Richards, T.A. (2015). Cryptic infection of a broad taxonomic and geographic diversity of tadpoles by Perkinsea protists. Proc. Natl. Acad. Sci. USA *112*, E4743–E4751.
32. Friedman, C.S., Gardner, G.R., Hedrick, R.P., Stephenson, M., Cawthorn, R.J., and Upton, S.J. (1995). *Pseudoklossia haliotis* sp. n. (Apicomplexa) from the kidney of California abalone, *Haliotis* spp. (Mollusca). J. Invertebr. Pathol. *66*, 33–38.
33. Kristmundsson, Á., Helgason, S., Bambir, S.H., Eydal, M., and Freeman, M.A. (2011). *Margolisiella islandica* sp. nov. (Apicomplexa: Eimeridae) infecting Iceland scallop *Chlamys islandica* (Müller, 1776) in Icelandic waters. J. Invertebr. Pathol. *108*, 139–146.
34. Fahay, M.P. (1983). Guide to the early stages of marine fishes occurring in the western North Atlantic Ocean, Cape Hatteras to the southern Scotian Shelf. J. Northwest Atl. Fish. Sci. *4*, 1–423.
35. Lough, R.G., Valentine, P.C., Potter, D.C., Auditore, P.J., Bolz, G.R., Neilson, J.D., and Perry, R.I. (1989). Ecology and distribution of juvenile cod and haddock in relation to sediment type and bottom currents on eastern Georges Bank. Mar. Ecol. Prog. Ser. *56*, 1–12.
36. Eydal, M., Kristmundsson, A., and Bambir, S.H. (2010). Pseudobranchial X-cell pseudotumors in young wild and farmed Atlantic cod *Gadus morhua* in Iceland. Dis. Aquat. Organ. *91*, 83–88.
37. Burki, F., Okamoto, N., Pombert, J.-F., and Keeling, P.J. (2012). The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc. Biol. Sci. *279*, 2246–2254.
38. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.
39. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
40. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. *61*, 539–542.
41. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.
42. Peng, Y., Leung, H.C.M., Yiu, S.M., and Chin, F.Y.L. (2012). IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics *28*, 1420–1428.
43. Kumar, S., Krabberød, A.K., Neumann, R.S., Michalickova, K., Zhao, S., Zhang, X., and Shalchian-Tabrizi, K. (2015). BIR pipeline for preparation of phylogenomic data. Evol. Bioinform. Online *11*, 79–83.
44. Roure, B., Rodriguez-Ezpeleta, N., and Philippe, H. (2007). SCaFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol. Biol. *7*, S2.
45. Lartillot, N., Rodrigue, N., Stubbs, D., and Richer, J. (2013). PhyloBayes MPI: phylogenetic reonstruction with infinite mixtures of profiles in a parallel environment. Syst. Biol. *62*, 611–615.
46. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), pp. 1–8.
47. van der Auwera, G., Chapelle, S., and De Wachter, R. (1994). Structure of the large ribosomal subunit RNA of Phytophthora megasperma, and phylogeny of the oomycetes. FEBS Lett. *338*, 133–136.
48. Bartošová, P., Fiala, I., and Hypša, V. (2009). Concatenated SSU and LSU rDNA data confirm the main evolutionary trends within myxosporeans (Myxozoa: Myxosporea) and provide an effective tool for their molecular phylogenetics. Mol. Phylogenet. Evol. *53*, 81–93.
49. Freeman, M.A., Yokoyama, H., and Ogawa, K. (2008). Description and phylogeny of *Ceratomyxa anko* sp. n. and *Zschokkella lophii* sp. n. from the Japanese anglerfish, *Lophius litulon* (Jordan). J. Fish Dis. *31*, 921–930.
50. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
51. Kristmundsson, Á., and Freeman, M.A. (2013). Sphaeromyxids form part of a diverse group of myxosporeans infecting the hepatic biliary systems of a wide range of host organisms. Parasit. Vectors *6*, 51.
52. Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688–2690.
53. Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. *57*, 758–771.
54. Altekar, G., Dwarkadas, S., Huelsenbeck, J.P., and Ronquist, F. (2004). Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics *20*, 407–415.
55. Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., Armbrust, E.V., Archibald, J.M., Bharti, A.K., Bell, C.J., et al. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biol. *12*, e1001889.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
X-cell xenomas from *Gadus morhua*	Hosts captured along the Icelandic and Norwegian Coast	N/A
X-cell xenomas of *Lycodes* sp.	Hosts captured along the Icelandic coast	N/A
X-cell xenomas of *Pseudopleuronectes obscurus*	[15]	N/A
X-cell xenomas of *Limanda limanda*	[14]	N/A
Critical Commercial Assays		
TruSeq DNA HT Sample Prep Kit	Illumina	FC-121-2003
DNeasy Blood & Tissue Kit	QIAGEN	Cat# 69504
GeneMATRIX Tissue DNA Purification Kit	EURx	Cat# E3550
Deposited Data		
Gadixcellia sp. SSU I-ITS-LSU sequence	This study	GenBank: KY628810
Gadixcellia sp. SSU II	This study	GenBank: KY628814
Xcellia lamelliphila SSU	This study	GenBank: KY628815
Xcellia pleuronecti LSU	This study	GenBank: KY628817
Xcellia lamelliphila LSU	This study	GenBank: KY628818
Gadixcellia gadi LSU	This study	GenBank: KY628819
Additional Marine Transcriptomes	Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP)	http://data.imicrobe.us/project/view/104: MMETSP0795,MMETSP0290, METSP0288,MMETSP1397,MMETSP1345, MMETSP0420,MMETSP0437,MMETSP0439
Reference single protein alignments for phylogenomic analysis	[37]	N/A
Newly added protein data from *Gadixcellia gadi* and *Gadixcellia* sp.	This study	Table S1
Oligonucleotides		
Oligonucleotide sequences for this study can be found in Table S3	This study	Details can be found in Table S3
Software and Algorithms		
Mafft v7.216	[38]	http://mafft.cbrc.jp/alignment/software/
RAxML 8.0.26/8.1.1	[39]	http://sco.h-its.org/exelixis/web/software/raxml/
MrBayes v 3.2.6	[40]	http://mrbayes.sourceforge.net/
blast+/2.2.29	[41]	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastDocs&DOC_TYPE=Download
idba-1.1.1	[42]	http://i.cs.hku.hk/~alse/hkubrg/projects/idba_ud/index.html
BIR	[43]	https://lifeportal.uio.no
SCAFOS 1.2.5	[44]	http://megasun.bch.umontreal.ca/Software/scafos/scafos.html
PhyloBayes MPI 1.5a	[45]	http://megasun.bch.umontreal.ca/People/lartillot/www/downloadmpi.html
Other		
NCBI Nucleotide Database	https://www.ncbi.nlm.nih.gov/nucleotide	NCBI Nucleotide, RRID: SCR_004860
Visualization and Analysis of Microbial Population Structure, VAMPS Project	http://vamps.mbl.edu/overview.php	VAMPS, RRID: SCR_004483

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BioMarKs Data Portal: Biodiversity of Marine Eukaryotes	[20]	http://biomarks.eu/
TaraOceans Expedition 2009-2013	[21]	https://www.embl.de/tara-oceans/start/
CIPRES Science Gateway for inference of large phylogenetic trees	[46]	CIPRES Science Gateway, RRID: SCR_008439

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Mark A. Freeman (mafreeman@rossvet.edu.kn), Ross University School of Veterinary Medicine, Basseterre, St Kitts, West Indies.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Trawling of fish to screen for xenomas

Atlantic cod (*Gadus morhua*) and blue whiting (*Micromesistius poutassou*) were caught off the Icelandic and Norwegian coasts during research expeditions performed by the Icelandic Marine Research Institute and the Norwegian Institute of Marine Research, respectively. Both trawls were approved by the Icelandic Directorate of Fisheries and the Norwegian Food and Safety Authority to comply with fisheries regulations. The samples from Icelandic waters were caught off the west and NW coast, respectively for blue whiting (n = 212; length = 20-40cm) and Atlantic cod (n = 300; length = 30-75cm). In Norway, both Atlantic cod (n = 566) and blue whiting (n = 8) were caught off the North coast. All fish were examined for the presence of pseudobranchial xenomas. In addition, 3,200 cod (length = 60-120cm) caught off West Iceland were examined for xenomas prior to processing in a fisheries factory. Xenomas were either dissected from fresh fish and flash frozen in liquid nitrogen and stored on dry ice aboard the research vessel or removed from fish frozen on the vessel (blue whiting, Iceland). Upon arrival on shore the samples were stored at −80°C until needed for molecular analyses.

Eelpout (*Lycodes* spp.) were sampled by trawling waters north of Iceland during an annual survey performed by the Marine Research Institute in Iceland. These included *Lycodes esmarkii* (length range 43 – 62 cm: n = 10), *L. vahli gracili*s (18 – 27 cm; n = 77), *L. reticulatus* (11 – 23 cm; n = 24), *L. pallidus* (18 – 26 cm; n = 4), *L. seminudus* (16 – 36 cm; n = 20) and *L. eudipleurostictus* (10 – 29 cm; n = 85). Shortly after catching and grading, the fish were frozen aboard the research vessels. After the survey, fish were brought to the laboratory and kept frozen until examined macroscopically for the presence of X-cell gill lesions. Samples from gill xenomas, taken for molecular analysis, were frozen in −80°C until further processed.

All fish used in this study had their internal organs and gonads examined (when possible), there was no relationship to X-cell xenoma presence and the sex of the fish.

METHOD DETAILS

Sequence generation

Genomic DNA (gDNA) for the whole genome sequencing was extracted from frozen xenomas using DNeasy Blood & Tissue Kit (QIAGEN), Animal Tissue Spin-Column Protocol, following the manufacturer's recommendations with a minor modification; for enhanced lysis the sample was mechanically disrupted using sterile tungsten carbide beads for 3 min at 20 Hz/s, and left for lysis at 56°C overnight. The gDNA was subjected to RNase treatment before it was purified using standard isopropanol precipitation. Samples were submitted for sequencing at the Norwegian Sequencing Centre, where the Illumina library was constructed using the Regular TruSeq DNA HT Sample Prep Kit (Illumina, US), before the sample was sequenced on one lane (100 base pair paired end, library size 500 bp) on an Illumina HiSeq2000 platform.

Fresh xenoma tissue or recently thawed material was either fixed in 95% ethanol or placed directly into tissue lysis buffer. DNA was extracted using a GeneMATRIX DNA extraction kit (EURx Poland) following the tissue protocol and used as template for PCRs. Additional X-cell DNA used in this study, was from gill lesions in common dab, and epidermal xenomas in northern black flounder from Japan, which were obtained during previous studies [14, 15]. PCRs targeting SSU and LSU rDNA used primer combinations given in Table S3 with cycling conditions detailed in the original descriptions [47–49].

Data analysis

Apicoplast genes

A blast search of all published apicoplast sequences extracted from GenBank against the metagenome from the blue whiting xenoma did not result in any significant hits. This however does not prove absence of a relic apicoplast/plastid in *Gadixcellia* sp., as the metagenomic coverage is partial and the assembly is incomplete. Further work is required to determine whether such genes occur in X-cells, preferably based on a significantly refined genome assembly or transcriptomic data.

Environmental sequences

We screened 18S V4 and V9 amplicons from three different high-throughput environmental sequences sources: VAMPS [19], BioMarKs [20], and Tara Oceans [21], representing more than 1 million OTUs using blastn [50] for environmental X-cell sequences. Using X-cell 18S reference sequences as query, we retrieved all the HTES OTUs with a similarity higher than 90% and with a coverage of more than 80%.

Microscopy

X-cells were harvested from freshly caught cod by gentle mechanical disruption of the excised xenoma tissues using forceps. Released cells were fixed in 2.5% glutaraldehyde, passed through a 0.4μm Whatman Cyclopore track-etched polycarbonate membrane, and prepared for SEM [51]. In brief, membranes were post-fixed in 1% osmium tetroxide in 100 mM sodium cacodylate buffer pH 7.2 for 2hrs and taken through an ethanol series of 30%, 60%, 90% and 2 × 100% 30 min each, transferred into 50% hexamethyldisilazane (HMDS) in 100% ethanol followed by two changes of 100% HMDS each for 45 min. Excess HMDS was removed and the membranes allowed to air dry overnight. The membranes were then mounted onto aluminum stubs, earthed with silver dag paint, sputter-coated with gold and viewed. Samples for standard wax histology and TEM were taken from fresh pseudobranchial X-cell xenomas from Atlantic cod from Iceland, and from epidermal X-cell xenomas from the northern black flounder, *P. obscurus*, obtained previously from Japan [15] which were prepared and examined following a standard protocol known to be suitable for X-cells [14]. Samples were fixed in 2.5% glutaraldehyde in 0.1M sodium cacodylate buffer (pH 7.2) at 4C for 2h. Specimens were then rinsed in 0.1M sodium cacodylate buffer overnight before being postfixed in 1% osmium tetroxide for 1 hr, dehydrated through an ethanol series, embedded in Spurr's resin and polymerized at 60C for 48 hr. Ultrathin sections (60–80nm) were stained with uranyl acetate and lead citrate prior to viewing.

QUANTIFICATION AND STATISTICAL ANALYSIS

Phylogenetic analysis of ribosomal RNA gene sequences

Ribosomal RNA gene sequence alignments were constructed using the e-ins-i algorithm in MAFFT [38] and refined by eye where necessary. Refined SSU alignments were analyzed in RAxML 8.1.11 [39, 52], BlackBox (GTR+gamma; all parameters estimated from the data); bootstrap values were mapped onto the highest likelihood tree obtained [53]. Bayesian consensus trees were constructed using MrBayes v 3.2.6 [40] in parallel mode [54], on the Cipres Science Gateway [46]. Two separate MC^3 runs with randomly generated starting trees were carried out for 5 million generations each with one cold and three heated chains. The evolutionary model applied included a GTR substitution matrix, a four-category autocorrelated gamma correction and the covarion model. All parameters were estimated from the data. Trees were sampled every 100 generations. 1.2 million generations were discarded as "burn-in" (trees sampled before the likelihood plots reached a plateau) and a consensus tree was constructed from the remaining sample.

Phylogenomic analysis

Candidate sequences for the multigene phylogeny were extracted by tblastn as implemented in blast+ [41] from two IDBA_UD [42] assembled xenoma metagenome blast databases from blue whiting and cod. Sequences from the reference alignments [37] were used as queries. The retrieved sequences were aligned with MAFFT, manually curated and assembled. Then they were re-analyzed together with the reference alignments using the BIR pipeline [43]. In the same run, several transcriptomes from the Marine Microbial Eukaryote Transcriptome Sequencing Project [55] (acc: MMETSP0795, MMETSP0290, MMETSP0288, MMETSP1397, MMETSP1345, MMETSP0420, MMETSP0437, MMETSP0439) were added to extract sequences for a better representation of Alveolates and Amoebozoa. After additional manual selection of candidate genes based on the single gene trees from BIR and manual curation of the alignments, the single gene alignments were concatenated using SCAFOS [44] to determine best candidates. Genera that were represented by more than one species were combined into one. Genera/Species that were afterward represented in less than 10% of the alignments were excluded. Several different setups for alignments were used and tested running RAxML PThreads [39] with rapid bootstrap analysis, with substitution model and matrix set to PROTGAMMAAUTO. Number of replicates for bootstrapping was determined with the option autoMRE. TCA scores were calculated for all trees, and the tree resulting from the alignment containing 74 taxa and 63 different proteins reached the highest relative tree certainty and therefore this alignment was chosen as an input for Phylobayes-MPI [45]. Phylobayes-MPI was run with four chains with the CAT-GTR model and matrix and the removal of constant columns (-dc). Four chains were run for more than 30,000 generations. Two of the four chains were converging respectively to reach a maximum difference of < 0.1. "Burn-in" was set to 400 (trees sampled before the likelihood plots reached a plateau).

DATA AND SOFTWARE AVAILABILITY

Sequences of the ribosomal RNAs are available in GenBank with the accession numbers KY628810, KY628814, KY628815, KY628817, KY628818 and KY628819. The sequences used for the phylogenomic approach are made available in Table S1 (see separate Excel file). Alignments and other resources are made available upon request.