THE ROYAL SOCIETY
PUBLISHING

**Evolutionary biology**

# Complete loss of the MHC II pathway in an anglerfish, *Lophius piscatorius*

Arseny Dubin, Tor Erik Jørgensen, Truls Moum, Steinar Daae Johansen and Lars Martin Jakt

Genomics group, Faculty of Biosciences and Aquaculture, Nord University, 8049 Bodø, Norway

AD, 0000-0001-8360-5475; TEJ, 0000-0002-6401-0713; LMJ, 0000-0002-3787-8138

Genome studies in fish provide evidence for the adaptability of the vertebrate immune system, revealing alternative immune strategies. The reported absence of the major compatibility complex (MHC) class II pathway components in certain species of pipefish (genus *Syngnathus*) and cod-like fishes (order Gadiformes) is of particular interest. The MHC II pathway is responsible for immunization and defence against extracellular threats through the presentation of exogenous peptides to T helper cells. Here, we demonstrate the absence of all genes encoding MHC II components (CD4, CD74 A/B, and both classical and non-classical MHC II $\alpha/\beta$) in the genome of an anglerfish, *Lophius piscatorius*, indicating loss of the MHC II pathway. By contrast, it has previously been reported that another anglerfish, *Antennarius striatus*, retains all MHC II genes, placing the loss of MHC II in the *Lophius* clade to their most recent common ancestor. In the three taxa where MHC II loss has occurred, the gene loss has been restricted to four or five core MHC II components, suggesting that, in teleosts, only these genes have functions that are restricted to the MHC II pathway.

## 1. Introduction

The vertebrate adaptive immune system that generates diversity through genetic recombination appears to have evolved in the common ancestor of the jawed vertebrates (gnathostomes) [1]. Although this system increased in complexity with gnathostome evolution, it is thought that the acquisition of all required cellular processes, tissues and genes happened relatively quickly as most components are present across all jawed vertebrates [1]. T-cell receptors (TCR), B-cell receptors (BCR) and the major histocompatibility complex (MHC) classes I and II are all present throughout the gnathostome lineages, from the Chondrichthyes to terrestrial vertebrates [1]. Although the specific sites of haematopoiesis vary, homologous tissues and organs including the thymus and spleen are also present across the gnathostomes [1–4].

An intact adaptive immune system has been found in almost all vertebrate species that have had their genomes sequenced, but recent work has demonstrated the loss of components of the adaptive immune system in the elephant shark, pipefish, coelacanth and the entire Gadiformes order [5–10]. These observations demonstrate an unexpected plasticity of adaptive immunity.

The teleost order Lophiiformes (Anglerfishes) harbours at least 321 living species, approximately half of which express some degree of sexual parasitism [11]. In these species, males attach to the females either temporarily or permanently. In extreme cases, this leads to fusion of male and female circulatory systems [12]. Why this fusion does not result in tissue rejection is unknown, but suggests a specialized adaptive immune system. Phylogenetic inference based on sequencing data and morphology has concluded that male sexual
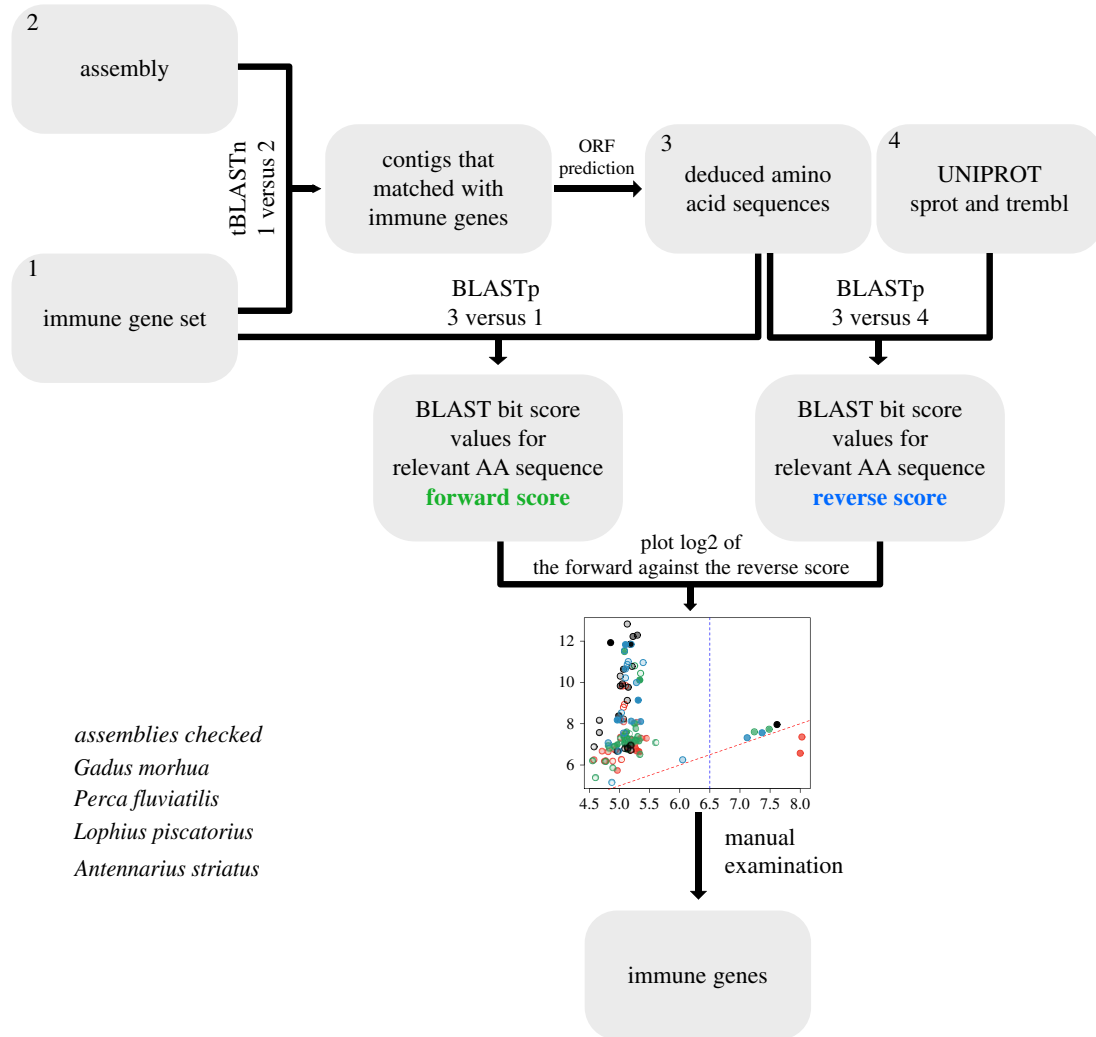
**Figure 1.** Outline of the gene mining process. Sequence inputs and outputs are shown as boxes with processes indicated by connecting arrows. BLAST inputs are numbered to indicate what was used as a query and subject (denoted as *query(number)* versus *subject(number)*). (Online version in colour.)

parasitism within Lophiiformes must have multiple origins [11,13,14], suggesting a common selective pressure or a shared genetic predisposition.

Here, we present two independently obtained draft genomes of an anglerfish, *Lophius piscatorius*, and show that it has lost all components of the MHC II arm of the adaptive immune system. The MHC II pathway is known to be involved in allogenic rejection [15] and our observations suggest that loss of MHC II may have contributed to the immune tolerance observed in sexually parasitic anglerfish species.

## 2. Material and methods

### (a) Sample collection, DNA isolation and sequencing

Samples from two *L. piscatorius* individuals (referred to as BF1 and BF2) were collected in the Bodø coastal waters, Nordland County in collaboration with local fishermen. BF1 skeletal muscle and BF2 kidney were used for subsequent total DNA isolation, library preparation and sequencing. BF1 total DNA sequencing was performed using Illumina MiSeq and SOLiD 5500 technologies (sequence depth: 24×). BF2 total DNA was sequenced by Dovetail Genomics, USA on an Illumina HiSeq X instrument (sequence depth: 150×) as a service. The Illumina libraries were 300 bp paired-end reads with 600 bp insert size for the MiSeq, and 150 bp paired-end reads with 350 bp insert size for the HiSeq.

### (b) Bioinformatic analysis

The raw reads were trimmed from adapters and low-quality bases using Cutadapt [16]. Only Illumina data were used for the assemblies. Prior to assembly, overlapping read pairs were merged using FLASH (v. 1.2.11) [17]. Final assemblies were constructed with SPAdes (v. 3.10.0) [18]. Basic assembly statistics were calculated with QUAST (v. 4.4.1) [19] and gene-space completeness assessed using BUSCO (v. 2.0) [20] with the actinopterygii dataset (odb9).

MHC genes were identified using methods similar to those used in [10] (figure 1). Briefly, a set of adaptive immune system-related protein sequences (bait-sequences) were used to identify contigs containing potential orthologues. Genes and open reading frames (ORFs) were predicted from these contigs and aligned both to the bait-set and sequences within the UniProt database to separate orthologues from non-orthologous genes containing homologous sequences. The resulting alignment scores were visualized and identities of candidate orthologues manually confirmed by inspection of alignments and annotations.

We performed this analysis on our *L. piscatorius* assemblies as well as on assemblies of *Antennarius striatus*, *Gadus morhua* and *Perca fluviatilis* [10]. If a gene was not identified in *L. piscatorius*, the unassembled reads (SOLiD and Illumina) were searched using tBLASTn. Matching reads were reassembled (CLC GW v. 11, QIAGEN, Aarhus, Denmark) and verified by reciprocal BLASTn against NCBI nr.

More detailed descriptions are provided in the electronic supplementary material.

# 3. Results

## (a) Genome assembly

The resulting *L. piscatorius* assemblies contained 664 (BF1) and 724 (BF2) megabases with N50 values of 6.9 kb and 108 kb, respectively. We used the BUSCO [20] actinopterygii set of 4584 conserved genes to estimate the gene-space completeness of these assemblies. We could detect at least 75% of these genes in both our assemblies (complete and fragmented), with 91.5% of complete genes identified in the BF2 assembly (electronic supplementary material, figure S1).

The gene space completeness of our assemblies is thus similar to that obtained for the *A. striatus* assembly (66.5% complete and 15.8% fragmented, electronic supplementary material, figure S1). Hence, our assemblies are comparable to or better than assemblies in [10] in terms of continuity, coverage and gene-space completeness (electronic supplementary material, figure S1).

## (b) Adaptive immune system genes in *L. piscatorius*

We used tBLASTn with a set of adaptive immune system genes to identify orthologous genes in *L. piscatorius* as well as in species previously reported to either have (*A. striatus*, *P. fluviatilis*) or lack (*G. morhua*) genes coding for MHC II components. Candidate orthologues were readily observed for all MHC I genes in all species. By contrast, we were unable to identify genes coding for CD4, CD74 A/B, MHC II $\alpha/\beta$ in either *L. piscatorius* or *G. morhua* assemblies (figure 2 and table 1). We repeated this analysis using an extended bait set including the non-classical MHC II $\alpha/\beta$ lineages [21]; this too failed to find any candidate orthologues in *L. piscatorius*. Similarly to the MHC I components we were also able to clearly identify orthologues of 22 out of 23 additional genes that have functions in the adaptive immune system (electronic supplementary material, figure S6, ST2).

To confirm the absence of MHC II orthologues in *L. piscatorius* we also searched for short sequences in the unassembled reads that could be aligned with the missing genes. Using tBLASTn, we identified 18 and 62 reads from BF1 and BF2, respectively, which aligned with an MHC II $\beta$ subunit. To locate the position of these potential MHC II sequences, the matching reads were assembled into contigs and mapped back to the original assemblies. This identified a region of approximately 300–480 bp in length present in both assemblies. When translated, the predicted reading frame was interrupted by multiple stop codons (electronic supplementary material, figure S2), indicating that the fragment represents a remnant of an MHC II $\beta$ chain gene. Hence, we conclude that the MHC II pathway is absent in *L. piscatorius*.

To confirm the presence of genes syntenic to CD4 and CD74 in *L. piscatorius*, we identified contigs containing these genes [5] and aligned them to the stickleback (*Gasterosteus aculeatus*) loci (electronic supplementary material, figure S7). All highly conserved genes were identified in either a single (CD74) or three (CD4) contigs and the gene predictions lying within these contigs aligned both in terms of direction and order. CD74 in *L. piscatorius* seems to have been lost through a deletion of a region lying between ndst1a and SCL35A4 that has removed both CD74 and almost all intergenic space. For CD4, we were unable to identify a contig spanning the expected CD4 position; nevertheless, our analysis confirms the presence of the expected syntenic genes within our assembly.

## (c) MHC II in *A. striatus*

*Antennarius striatus* has been reported to contain both MHC I and MHC II pathway genes [10]. Since both *A. striatus* and *L. piscatorius* are members of the Lophiiformes order, we considered the possibility that the identification of *A. striatus* MHC II orthologues could have resulted from cross-contamination of the sample. Although we did observe the presence of cross-contaminating mitochondrial sequences from distantly related teleost taxa, the relative sequencing depth of contaminant and *A. striatus* sequences, combined with the unimodal depth distribution preclude the possibility that the MHC II sequences were derived from contaminant DNA fragments (electronic supplementary material, figure S4). Our observations thus confirm the presence of MHC II, while at the same time highlighting the potential of cross-contamination leading to confounding results.

# 4. Evolutionary considerations

Of the 81 teleosts examined so far, only members of the Gadiformes order (27 species sequenced), pipefishes (*Syngnathus typhle* and *S. scovelli*) and now *L. piscatorius* lack a functional MHC II pathway [5,6,9,10]. Notably, all these species lack CD4 and both MHC II $\alpha$ and $\beta$ chains (classical and non-classical molecules). In addition, CD74 A/B has been completely lost in *L. piscatorius* and the Gadiformes order. Nevertheless, all contain a complete set of MHC I pathway components [6,9,10] (figure 2; electronic supplementary material, figure S6). In contrast to terrestrial vertebrates, teleost MHC II $\alpha$ and $\beta$ often occupy multiple loci on different chromosomes [22]. This means that loss of the complete MHC II pathway requires multiple independent gene deletions; e.g. the loss of MHC II in *A. striatus* would require around 10 deletions (table 1). This suggests that loss of any critical MHC II pathway component leads to the loss of remaining core parts and argues that genes lost in these species are unlikely to have functions outside of MHC II in teleosts.

By contrast, genes such as CIITA, which is conventionally thought to be a specific activator of MHC II gene expression [23], is likely to have roles outside the MHC II system because in the absence of neo-functionalization, it would have no function after MHC II loss. Hence, our results are consistent with reports indicating an additional role of CIITA in the regulation of MHC I expression [24–26].

Sexual parasitism in anglerfish suggests some form of specialized immune system allowing for allogenic tolerance between fused individuals. Although CD8+ cytotoxic lymphocytes are thought to be the primary effectors of allogenic rejection, it is clear that MHC II components have both enabling and effector functions [15], and a long line of publications show that repression of MHC II components can contribute to immune tolerance or alleviate immune rejection [27–30]. Hence, it is tempting to speculate that the loss of MHC II in the Lophiodei suborder is not restricted to *Lophius* species and has played an enabling role in the development of sexual parasitism.

Most Lophiiformes phylogenies place the Lophioidei suborder at the most basal position in the anglerfish taxonomy,
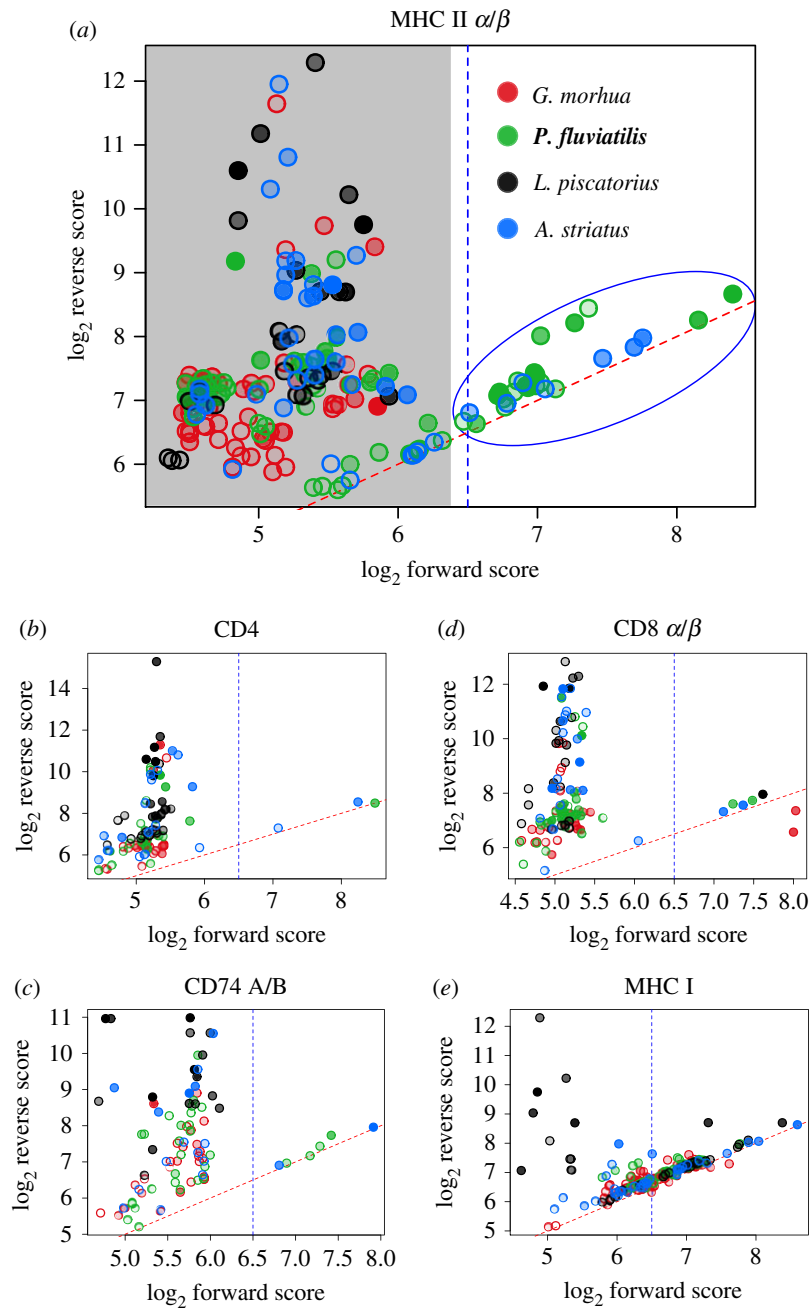
**Figure 2.** Identification of MHC I and MHC II pathway orthologues. Illustration of identification criteria (*a*). The alignment scores of putative orthologues against the initial bait set (forward score, *X*-axis) plotted versus scores against the UniProt database (reverse score, *Y*-axis). Grey shading indicates hits that were dismissed from further analysis, as they align better to genes not belonging to the MHC pathways. Alignments of short but highly conserved gene fragments can have high alignment scores without indicating functional orthologues; such points are shown with a high transparency by scaling the fill transparency by the ratio of alignment to subject (UniProt) length. Orthologues should have values closer to $Y = X$ indicated by the dashed red line. Candidate orthologues shown within the blue ellipse appear as outliers, and can be identified by a forward score threshold indicated by the dashed blue line. Plots for the MHC II components (*a–c*) lack candidate orthologues for *Lophius piscatorius* (black) and *Gadus morhua* (red), whereas candidate orthologues are evident in all species for the MHC I genes (*d,e*). Results for the $\alpha/\beta$ chains of CD8 and MHC II, and the CD74 A/B genes are shown combined.

**Table 1.** Number of candidate orthologues identified after forward/reverse screening (see §2) and manual inspection of the plots (figure 2). Numbers in brackets indicate individual hits after the forward score threshold was applied, but before manual examination of UniProt IDs identified unrelated genes.

| gene | *Gadus morhua* | *Perca fluviatilis* | *Lophius piscatorius* | *Antennarius striatus* |
|---|---|---|---|---|
| CD4 | **0** | 1 | **0** | 2 |
| CD74 A/B | **0** | 4 | **0** | 2 |
| MHC II $\alpha/\beta$ | **0** | 21 (22) | **0** | 6 (7) |
| CD8 $\alpha/\beta$ | 2 | 2 | 1* | 2 |
| MHC I | 49 | 34 (35) | 18 (19) | 12 (13) |

*Predicted sequence appears as a fusion protein of $\alpha$ and $\beta$ chains.

followed by Antennarioidei [11,14,31,32]. If Lophioidei is basal, then MHC II loss is likely to be specific to the Lophioidei suborder since *A. striatus* clearly possesses all MHC II components. That would also mean that our observations are unlikely to be relevant to sexual parasitism in other anglerfish clades. However, inferences of higher-order taxonomies are still fraught with difficulty, exemplified by the fact that the phylogenetic topology of the taxa involved based on mitochondrial DNA is subject to change depending on the choice of outgroup [11,13] (electronic supplementary material, figure S5). Exploring the presence and absence of MHC II genes in other anglerfish species can thus provide a test of the conventional phylogeny, as well as the likelihood of MHC II loss being one of the enabling adaptations preventing intra-species tissue rejection.

## 5. Conclusion

The classical MHC II components are responsible for the presentation of exogenous peptides to T helper cells and constitute an important part of the gnathostome adaptive immune system. Here, we report two draft genome assemblies of *L. piscatorius* and demonstrate a complete loss of the classical MHC II pathway in this species. The finding of a third taxon that lacks MHC II function corroborates the dispensability of MHC II in teleosts, and suggests that the genes lost in all three clades have no function outside of the MHC II.

## References

1. Flajnik MF. 2018 A cold-blooded view of adaptive immunity. *Nat. Rev. Immunol.* **18**, 438–453. (doi:10.1038/s41577-018-0003-9)

2. Orkin SH, Zon LI. 2008 Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644. (doi:10.1016/j.cell.2008.01.025)

3. Press CM, Evensen Ø. 1999 The morphology of the immune system in teleost fishes. *Fish Shellfish Immun.* **9**, 309–318. (doi:10.1006/fsim.1998.0181)

4. Neely HR, Flajnik MF. 2016 Emergence and evolution of secondary lymphoid organs. *Annu. Rev. Cell Dev. Biol.* **32**, 693–711. (doi:10.1146/annurev-cellbio-111315-125306)

5. Star B et al. 2011 The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207–210. (doi:10.1038/nature10342)

6. Haase D, Roth O, Kalbe M, Schmiedeskamp G, Scharsack JP, Rosenstiel P, Reusch TBH. 2013 Absence of major histocompatibility complex class II mediated immunity in pipefish, *Syngnathus typhle*: evidence from deep transcriptome sequencing. *Biol. Lett.* **9**, 20130044. (doi:10.1098/rsbl.2013.0044)

7. Amemiya CT et al. 2013 The African coelacanth genome provides insights into tetrapod evolution. *Nature* **496**, 311–316. (doi:10.1038/nature12027)

8. Venkatesh B et al. 2014 Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**, 174–179. (doi:10.1038/nature12826)

9. Small CM, Bassham S, Catchen J, Amores A, Fuiten AM, Brown RS, Jones AG, Cresko WA. 2016 The genome of the Gulf pipefish enables understanding of evolutionary innovations. *Genome Biol.* **17**, 258–280. (doi:10.1186/s13059-016-1126-6)

10. Malmstrøm M et al. 2016 Evolution of the immune system influences speciation rates in teleost fishes. *Nat. Genet.* **48**, 1204–1210. (doi:10.1038/ng.3645)

11. Miya M et al. 2010 Evolutionary history of anglerfishes (Teleostei: Lophiiformes): a mitogenomic perspective. *BMC Evol. Biol.* **10**, 58–84. (doi:10.1186/1471-2148-10-58)

12. Pietsch TW. 2005 Dimorphism, parasitism, and sex revisited: modes of reproduction among deep-sea ceratioid anglerfishes (Teleostei: Lophiiformes). *Ichthyol. Res.* **52**, 207–236. (doi:10.1007/s10228-005-0286-2)

13. Shedlock AM, Pietsch TW, Haygood MG, Bentzen P, Hasegawa M. 2004 Molecular systematics and life history evolution of anglerfishes (Teleostei: Lophiiformes): evidence from mitochondrial DNA. *Steenstrupia* **28**, 129–144.

14. Pietsch TW, Orr JW. 2007 Phylogenetic relationships of deep-sea anglerfishes of the suborder Ceratioidei (Teleostei: Lophiiformes) based on morphology. *Copeia* **2007**, 1–34. (doi:10.1643/0045-8511(2007)7[1:PRODAO]2.0.CO;2)

15. Rosenberg AS, Singer A. 1992 Cellular basis of skin allograft rejection: an *in vivo* model of immune-mediated tissue destruction. *Annu. Rev. Immunol.* **10**, 333–360. (doi:10.1146/annurev.iy.10.040192.002001)

16. Martin M. 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12. (doi:10.14806/ej.17.1.200)

17. Magoc T, Salzberg SL. 2011 FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963. (doi:10.1093/bioinformatics/btr507)

18. Bankevich A et al. 2012 SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477. (doi:10.1089/cmb.2012.0021)

19. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013 QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075. (doi:10.1093/bioinformatics/btt086)

20. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018 BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **35**, 543–548. (doi:10.1093/molbev/msx319)

21. Dijkstra JM, Grimholt U, Leong J, Koop BF, Hashimoto K. 2013 Comprehensive analysis of MHC class II genes in teleost fish genomes reveals dispensability of the peptide-loading DM system in a large part of vertebrates. *BMC Evol. Biol.* **13**, 260–273. (doi:10.1186/1471-2148-13-260)

22. Grimholt U. 2016 MHC and evolution in teleosts. *Biology* **5**, 6–25. (doi:10.3390/biology5010006)

23. Neefjes J, Jongsma MLM, Paul P, Bakke O. 2011 Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836. (doi:10.1038/nri3084)

24. Martin BK, Chin K-C, Olsen JC, Skinner CA, Dey A, Ozato K, Ting JP-Y. 1997 Induction of MHC Class I expression by the MHC Class II transactivator CIITA. *Immunity* **6**, 591–600. (doi:10.1016/S1074-7613(00)80347-7)

25. Howcroft TK, Raval A, Weissman JD, Gegonne A, Singer DS. 2003 Distinct transcriptional pathways regulate basal and activated major histocompatibility complex class I expression. *Mol. Cell. Biol.* **23**, 3377–3391. (doi:10.1128/MCB.23.10.3377-3391.2003)

26. Gobin SJP, Peijnenburg A, Keijsers V, van den Elsen PJ. 1997 Site $\alpha$ is crucial for two routes of IFN$\gamma$-induced MHC Class I transactivation: the ISRE-mediated route and a novel pathway involving CIITA. *Immunity* **6**, 601–611. (doi:10.1016/S1074-7613(00)80348-9)

27. Shizuru J, Gregory A, Chao C, Fathman C. 1987 Islet allograft survival after a single course of treatment of recipient with antibody to L3T4. *Science* **237**, 278–280. (doi:10.1126/science.2955518)

28. Sato Y, Bolzenius JK, Eteleeb AM, Su X, Maher CA, Sehn JK, Arora VK. 2018 CD4[+] T cells induce rejection of urothelial tumors after immune checkpoint blockade. *JCI Insight* **3**, e121062. (doi:10.1172/jci.insight.121062)

29. Borges TJ *et al*. 2018 March1-dependent modulation of donor MHC II on CD103[+] dendritic cells mitigates alloimmunity. *Nat. Commun.* **9**, 3482–3497. (doi:10.1038/s41467-018-05572-z)

30. Abrahimi P, Qin L, Chang WG, Bothwell ALM, Tellides G, Saltzman WM, Pober JS. 2016 Blocking MHC class II on human endothelium mitigates acute rejection. *JCI Insight* **1**, e85293. (doi:10.1172/jci.insight.85293)

31. Pietsch TW. 1984 Lophiiformes: Development and relationships. In *Ontogeny and systematics of fishes: based on an international symposium dedicated to the memory of Elbert Halvor Ahlstrom* (eds HG Moser, WJ Richards, DM Cohen, MP Fahay, AW Kendall, SL Richardson), pp. 320–325. New York, NY: American Society of Ichthyologists and Herpetologists.

32. Pietsch TW. 1981 The osteology and relationships of the anglerfish genus *Tetrabrachium* with comments on lophiiform classification. *Fish. Bull.* **79**, 387–419.

33. Dubin A, Jørgensen TE, Moum T, Johansen SD and Jakt LM. 2019 Data from: Complete loss of the MHC II pathway in an anglerfish, *Lophius piscatorius*. Dryad Digital Repository. (doi:10.5061/dryad.4dc007q)

6

royalsocietypublishing.org/journal/rsbl    *Biol. Lett.* **15**: 20190594