



The fishes of Genome 10K

Giacomo Bernardi^{a,*}, Edward O. Wiley^b, Hicham Mansour^c, Michael R. Miller^d, Guillermo Orti^e, David Haussler^f, Stephen J. O'Brien^g, Oliver A. Ryder^h, Byrappa Venkatesh^{i,*}

^a Department of Ecology and Evolutionary Biology, University of California Santa Cruz, United States

^b Department of Ecology and Evolutionary Biology, University of Kansas, Natural History Museum and Biodiversity Research Center, Lawrence, KS, United States

^c Bioscience and Bioengineering Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal 23955–6900, Saudi Arabia

^d Institute of Molecular Biology, University of Oregon, United States

^e The George Washington University, Department of Biological Sciences, United States

^f Howard Hughes Medical Institute, UCSC, CBSE/ITI E2501, University of California, Santa Cruz, Santa Cruz, CA, United States

^g National Cancer Institute, Laboratory of Genomic Diversity, Frederick, MD, United States

^h San Diego Zoo Institute for Conservation Research, Escondido, CA, United States

ⁱ Institute of Molecular and Cell Biology, Agency for Science, Technology and Research, Biopolis, Singapore, Republic of Singapore

ARTICLE INFO

Article history:

Received 23 February 2012

Received in revised form 27 February 2012

Accepted 27 February 2012

Keywords:

Genome 10K

Fish genome

Vertebrate genome

ABSTRACT

The Genome 10K project aims to sequence the genomes of 10,000 vertebrates, representing approximately one genome for each vertebrate genus. Since fishes (cartilaginous fishes, ray-finned fishes and lobe-finned fishes) represent more than 50% of extant vertebrates, it is planned to target 4,000 fish genomes. At present, nearly 60 fish genomes are being sequenced at various public funded labs, and under a Genome 10K and BGI pilot project. An additional 100 fishes have been identified for sequencing in the next phase of Genome 10K project.

© 2012 Elsevier B.V. All rights reserved.

In 2009, a group of scientists met in Santa Cruz, California, to launch a project to sequence the genomes of ten thousand vertebrates, the Genome 10K project (G10KCOS, 2009). With recent reductions in the cost of DNA sequencing, the project, which would have been inconceivable just a few years earlier, is now considered very accessible. The main objective of the project was to build genomic resources that would help in understanding the structure, function and evolution of vertebrate genomes. These resources, and the deeper understanding of vertebrates enabled by them, were expected to be used in managing stocks of commercially important species to ensure their sustained yield, in developing effective strategies for maintaining biodiversity, and in the conservation and propagation of threatened and endangered species. Since fishes represent approximately 50% of living vertebrates, about 4000 fish genomes will be targeted for sequencing under the Genome 10K project. Thus far, 4 fish species have been completely sequenced and published by various publicly funded research groups (Aparicio et al., 2002; Jaillon et al., 2004; Kasahara et al., 2007; Star et al., 2011) and an additional 24 are currently being sequenced by various groups (Table 1). In November 2010, the Genome 10K project took its first major step by launching

a pilot project to sequence the genomes of 101 vertebrates in collaboration with BGI, Shenzhen. This list of 101 vertebrates includes 28 fishes (Table 2). Sequencing of these species is now ongoing at BGI and is expected to be completed by the end of 2012.

In a recent meeting at UCSC, Genome 10K scientists met as a full group, to resolve details of approaches, protocols, assembly, annotation, and analyses pipelines. Scientists were divided into smaller working units, mostly taxon-based, and the goals for these smaller groups were two fold. First, taxon groups needed to identify a unified protocol to gather samples and their full information. Second and more importantly, groups were asked to identify priorities to determine which genomes would be prime candidates for being in the next sequencing batch. We were asked to identify 100 fish species for this next tier.

The task was not easy as it came with much responsibility. With input from colleagues, we had to balance pragmatic approaches (which samples are actually going to be available), with purely scientific arguments. The main scientific arguments were both theoretical (species that would fill important gaps in the phylogenetic tree), and applied (importance to commercial fisheries or endangered). These 100 species will be considered as “gold standards”, and efforts will be made to sequence their transcriptomes in addition to sequencing whole genomes, and to generate stable cell lines. This means that besides the expected voucher specimen, high molecular weight DNA and good quality RNA from various tissues will be extracted, which will allow the construction of different types of libraries for *de novo*

* Corresponding authors.

E-mail addresses: bernardi@ucsc.edu (G. Bernardi), mcbbv@imcb.a-star.edu.sg (B. Venkatesh).

Table 1

Fish genomes sequenced or being sequenced at various public funded laboratories (sample number in parentheses). The Genome Center at Washington University (1, 2, 6, 14), IMCB, Singapore (2, 26), Mount Desert Island Biological Lab (3), Broad Institute (4,12, 15–19, 27,28), Leif Andersson's Lab, Sweden (5), Sanger Institute (6), USDA, Auburn University (7), Atlantic salmon Consortium (Norway, Canada and Chile) (8), USDA-Agricultural Research Service; University of Oregon (9), Genoscope (9, 27), Centre for Ecological and Evolutionary Synthesis, University of Oslo (10), National Institute of Genetics, Japan (11), Leibniz Institute for Age Research, Jena, Germany (13), DOE Joint Genome Institute (17, 20–23, 26), European Consortium (24), Fisheries Research Agency (Japan), University of Tokyo, Kyushu University (25). Asterisks (*) indicate species where assembly and analysis of whole genome sequences have been published. Provided are sample number (column 1) and family number (column 2) per Nelson, 2006 (Nelson, 2006).

S #	F #	Species	Common name	Family	Order
1	2	<i>Petromyzon marinus</i>	Sea lamprey	Petromyzontidae	Petromyzontiformes
2	5	<i>Callorhynchus milii</i>	Elephant shark	Callorhynchidae	Chimaeriformes
3	48	<i>Leucoraja erinacea</i>	Little skate	Rajidae	Rajiformes
4	62	<i>Lepisosteus oculatus</i>	Spotted gar	Lepisosteidae	Lepisosteiformes
5	97	<i>Clupea harengus</i>	Atlantic herring	Clupeidae	Clupeiformes
6	102	<i>Danio rerio</i>	Zebrafish	Cyprinidae	Cypriniformes
7	143	<i>Ictalurus punctatus</i>	Channel catfish	Ictaluridae	Siluriformes
8	175	<i>Salmo salar</i>	Atlantic salmon	Salmonidae	Salmoniformes
9	175	<i>Oncorhynchus mykiss</i>	Rainbow trout	Salmonidae	Salmoniformes
10	220	<i>Gadus morhua*</i>	Atlantic cod	Gadidae	Gadiformes
11	252	<i>Oryzias latipes*</i>	Medaka	Adrianichthyidae	Beloniformes
12	258	<i>Nothobranchius furzeri</i>	Turquoise killifish	Notobranchiidae	Cyprinodontiformes
13	266	<i>Xiphophorus maculatus</i>	Platyfish	Poeciliidae	Cyprinodontiformes
14	291	<i>Gasterosteus aculeatus</i>	Three-spined stickleback	Gasterosteidae	Gasterosteiformes
15	409	<i>Oreochromis niloticus</i>	Nile tilapia	Cichlidae	Perciformes
16	409	<i>Astatotilapia burtoni</i>	Burton's mouthbrooder	Cichlidae	Perciformes
17	409	<i>Maylandia zebra</i>	Zebra mbuna	Cichlidae	Perciformes
18	409	<i>Pundamilia nyererei</i>	Nyererei cichlid	Cichlidae	Perciformes
19	409	<i>Neolamprologus brichardi</i>	Fairy cichlid	Cichlidae	Perciformes
20	409	<i>Rhamphochromis esox</i>	Tiger cichlid	Cichlidae	Perciformes
21	409	<i>Melanochromis auratus</i>	Golden mbuna	Cichlidae	Perciformes
22	409	<i>Mchenga conophoros</i>	Happy cichlid	Cichlidae	Perciformes
23	409	<i>Labeotropheus fuelleborni</i>	Blue mbuna	Cichlidae	Perciformes
24	332	<i>Dicentrarchus labrax</i>	European Seabass	Moronidae	Perciformes
25	475	<i>Thunnus orientalis</i>	Pacific bluefin tuna	Scombridae	Scombriformes
26	509	<i>Takifugu rubripes*</i>	Fugu	Tetraodontidae	Tetraodontiformes
27	509	<i>Tetraodon nigroviridis*</i>	Freshwater pufferfish	Tetraodontidae	Tetraodontiformes
28	512	<i>Latimeria chalumnae</i>	South African coelacanth	Latimeriidae	Coelacanthiformes

assembly and annotation. In practical terms, this means that the 100 species need to be caught fresh and carefully sampled, DNA barcoded and stored. The 100 species to be sequenced (Table 3) were therefore

identified considering those specific constraints. The collection of these species has begun in earnest with help from affiliates of Genome 10K project spread around the globe.

Table 2

Fishes that are currently being sequenced under the Genome 10K pilot project. Provided are sample number (column 1), and family number (column 2) per Nelson, 2006 (Nelson, 2006).

S #	F #	Species	Common name	Family	Order
1	20	<i>Carcharodon carcharias</i>	Great white shark	Alopiidae	Lamniformes
2	59	<i>Polypterus senegalus</i>	Bichir	Polypteridae	Polypteriformes
3	60	<i>Acipenser sinensis</i>	Chinese sturgeon	Acipenseridae	Acipenseriformes
4	63	<i>Amia calva</i>	Bowfin	Amiidae	Amiiformes
5	65	<i>Scleropages formosus</i>	Golden arowana	Osteoglossidae	Osteoglossiformes
6	74	<i>Anguilla anguilla</i>	European freshwater eel	Anguillidae	Anguilliformes
7	102	<i>Aristichthys nobilis</i>	Bighead carp	Cyprinidae	Cypriniformes
8	102	<i>Megalobrama amblycephala</i>	Wuchang bream	Cyprinidae	Cypriniformes
9	102	<i>Hypophthalmichthys molitrix</i>	Silver carp	Cyprinidae	Cypriniformes
10	102	<i>Gobiocypris rarus</i>	Rare gudgeon	Cyprinidae	Cypriniformes
11	119	<i>Astyanax mexicanus</i>	Blind cave fish	Characidae	Characiformes
12	200	<i>Diaphus dumerilii</i>	Lanternfish	Myctophidae	Myctophiformes
13	280	<i>Hoplostethus atlanticus</i>	Orange roughy	Trachichthyidae	Beryciformes
14	295	<i>Hippocampus comes</i>	Tiger tail seahorse	Syngnathidae	Gasterosteiformes
15	300	<i>Monopterus albus</i>	Finless eel	Synbranchidae	Synbranchiformes
16	332	<i>Lateolabrax japonicus</i>	Japanese seabass	Moronidae	Perciformes
17	338	<i>Epinephelus coioides</i>	Grouper	Serranidae	Perciformes
18	378	<i>Sparus aurata</i>	Gilthead sea bream	Sparidae	Perciformes
19	381	<i>Pseudosciaena crocea</i>	Large yellow croaker	Sciaenidae	Perciformes
20	426	<i>Eleginops maclovinus</i>	Patagonian blenny	Eleginopidae	Perciformes
21	427	<i>Dissostichus mawsoni</i>	Antarctic toothfish	Nototheniidae	Perciformes
22	431	<i>Chanecephalus aceratus</i>	Blackfin icefish	Channichthyidae	Perciformes
23	460	<i>Periophthalmodon schlosseri</i>	Giant mudskipper	Gobiidae	Perciformes
24	475	<i>Thunnus albacares</i>	Yellowfin tuna	Scombridae	Perciformes
25	483	<i>Pampus argenteus</i>	Pomfret	Stromateidae	Perciformes
26	492	<i>Paralichthys olivaceus</i>	Bastard halibut	Paralichthyidae	Pleuronectiformes
27	502	<i>Cynoglossus semilaevis</i>	Tongue sole	Cynoglossidae	Pleuronectiformes
28	511	<i>Mola mola</i>	Sunfish	Molidae	Tetraodontiformes

Table 3

One hundred fish species currently being considered as “gold standard” for the Genome 10K project. Provided are sample number (column 1), family number (column 2) per Nelson, 2006 (Nelson, 2006), and genome size in pg, with one asterisk for values of species in the same genus, two asterisks for values of species within the same family (right column) (Hinegardner and Rosen, 1972; Smith and Gregory, 2009; Gregory, 2011).

S #	F #	Species	Common name	Family	Order	G. size
1	7	<i>Hydrolagus collieri</i>	Ratfish	Chimaeridae	Chimaeriformes	1.51
2	30	<i>Sphyrna</i> sp.	Hammerhead shark	Sphyrnidae	Carcharhiniformes	3.5*
3	44	<i>Pristis</i> sp.	Sawfish	Pristidae	Pristiformes	2.8*
4	42	<i>Torpedo californica</i>	Pacific electric ray	Torpedinidae	Torpediniformes	7.00
5	59	<i>Erpetoichthys calabaricus</i>	Ropefish	Polypteridae	Polypteriformes	4.70
6	61	<i>Polyodon spatula</i>	Paddlefish	Polyodontidae	Acipenseridae	1.60
7	64	<i>Hiodon alosoides</i>	Goldeye	Hiodontidae	Hiodontiformes	
8	65	<i>Pantodon buchholzi</i>	Freshwater butterflyfish	Osteoglossidae	Osteoglossiformes	0.77
9	66	<i>Notopterus notopterus</i>	Knife fish	Notopteridae	Osteoglossiformes	1.3**
10	70	<i>Megalops atlanticus</i>	Tarpon	Megalopidae	Elopiformes	1.2**
11	71	<i>Albula vulpes</i>	Bonefish	Albulidae	Elopiformes	
12	79	<i>Gymnothorax mordax</i>	California moray	Muraenidae	Anguilliformes	2.45*
13	95	<i>Anchoa</i> sp.	Anchovy	Engraulidae	Clupeiformes	1.9*
14	97	<i>Sardinella</i> sp.	Sardine	Clupeidae	Clupeiformes	1.3**
15	98	<i>Chanos chanos</i>	Milkfish	Chanidae	Gonorhynchiformes	
16	102	<i>Paedocypris progenetica</i>	Paedocypris	Paedocyprinida	Cypriniformes	
17	102	<i>Cyprinus carpio</i>	Carp	Cyprinidae	Cypriniformes	1.70
18	102	<i>Catla catla</i>	Indian Carp	Cyprinidae	Cypriniformes	1.7**
19	106	<i>Botia kubotai</i>	Burmese border loach	Cobitidae	Cypriniformes	1.3**
20	119	<i>Serrasalmus rhombeus</i>	Piranha	Characidae	Characiformes	1.6*
21	131	<i>Corydoras julii</i>	Leopard catfish	Callichthyidae	Siluriformes	4.20
22	151	<i>Plotosus lineatus</i>	Striped catfish	Plotosidae	Siluriformes	1.75**
23	159	<i>Pangasius gigas</i>	Pangasius	Bagridae	Siluriformes	2.5**
24	161	<i>Electrophorus electricus</i>	Electric eel	Gymnotidae	Gymnotiformes	0.99*
25	166	<i>Argentina</i> sp.	Herring smelt	Argentinidae	Argentiniformes	0.85*
26	171	<i>Alepocephalus</i> sp.	Slickhead	Alepocephalidae	Argentiniformes	
27	172	<i>Thaleichthys pacificus</i>	Eulachon	Osmeridae	Osmeriformes	
28	174	<i>Galaxias maculatus</i>	Inanga	Galaxiidae	Osmeriformes	1.05*
29	175	<i>Oncorhynchus tshawytscha</i>	Chinook salmon	Salmonidae	Salmoniformes	3.30
30	175	<i>Coregonus clupeaformis</i>	Lake whitefish	Salmonidae	Salmoniformes	3.3**
31	176	<i>Esox lucius</i>	Pike	Esocidae	Esociformes	1.00
32	182	<i>Chauliodus</i> sp.	Viperfish	Stomiidae	Stomiiformes	
33	187	<i>Synodus</i>	Lizard fish	Synodontidae	Aulopiformes	1.2**
34	200	<i>Benthoosema pterotum</i>	Skinnycheek lanternfish	Myctophidae	Myctophiformes	2.0**
35	202	<i>Lampris guttatus</i>	Opah	Lampridae	Lampridiformes	
36	209	<i>Aphredoderus sayanus</i>	Pirate perch	Percopsidae	Percopsiformes	
37	222	<i>Chilara taylori</i>	Cusk eel	Ophiidae	Ophiidiformes	
38	226	<i>Porichthys notatus</i>	Midshipman	Batrachoididae	Batrachoidiformes	2.20
39	227	<i>Lophius</i> sp.	Goosefish	Lophiidae	Lophiiformes	1.0*
40	231	<i>Thymichthys</i> sp.	Handfish	Brachionich.	Lophiiformes	
41	242	<i>Ceratias</i> sp.	Sea devil	Ceratiidae	Lophiiformes	
42	245	<i>Mugil cephalus</i>	Grey mullet	Mugilidae	Mugiliformes	0.99
43	246	<i>Odontesthes bonariensis</i>	Pejerrey	Atherinopsidae	Atheriniformes	
44	253	<i>Exocoetus</i> sp.	Flyingfish	Exocoetidae	Beloniformes	1.1*
45	254	<i>Hemiramphus</i> sp.	Halfbeak	Hemiramphidae	Beloniformes	1.1*
46	255	<i>Strongylura marina</i>	Needlefish	Belonidae	Beloniformes	1.1*
47	259	<i>Krytolebias marmoratus</i>	Mangrove rivulus	Rivulidae	Cyprinodontiformes	1.5*
48	262	<i>Fudulus heteroclitus</i>	Mummichog	Fundulidae	Cyprinodontiformes	1.50
49	264	<i>Cyprinodon variegatus</i>	Sheepshead	Cyprinodontid	Cyprinodontiformes	1.60
50	266	<i>Poecilia reticulata</i>	Guppy	Poeciliidae	Cyprinodontiformes	0.96*
51	282	<i>Neoniphon sammara</i>	Sammara squirrelfish	Holocentridae	Beryciformes	0.82
52	288	<i>Zeus faber</i>	John Dory	Zeidae	Zeiformes	1.23**
53	295	<i>Syngnathus</i> sp.	Pipefish	Syngnathidae	Syngnathiformes	0.66**
54	302	<i>Mastacembelus armatus</i>	Tire track eel	Mastacembelid	Synbranchiformes	
55	303	<i>Dactylopterus</i> sp.	Flying gurnard	Dactylopteridae	Dactylopteriformes	
56	304	<i>Sebastes chrysomelas</i>	Black-n-yellow rockfish	Scorpaenidae	Scorpaeniformes	0.96*
57	304	<i>Pterois radiata</i>	Clearfin lionfish	Scorpaenidae	Scorpaeniformes	0.96**
58	313	<i>Platycephalus bassensis</i>	Sand Flathead	Platycephalidae	Scorpaeniformes	0.80*
59	320	<i>Clinocottus analis</i>	Wolly sculpin	Cottidae	Scorpaeniformes	0.93
60	324	<i>Aspidophoroides</i>	Poacher	Agonidae	Scorpaeniformes	0.76**
61	327	<i>Cyclopterus lumpus</i>	Lumpfish	Cyclopteridae	Scorpaeniformes	
62	332	<i>Morone saxatilis</i>	Striped bass	Moronidae	Perciformes	0.89
63	338	<i>Pseudanthias squamipinnis</i>	Lyretail anthias	Serranidae	Perciformes	1.25**
64	349	<i>Micropterus salmoides</i>	Large mouth bass	Centrarchidae	Perciformes	1.00
65	350	<i>Sander vitreus</i>	Walleye	Percidae	Perciformes	1.2**
66	354	<i>Sillago ciliata</i>	Sand whiting	Sillaginidae	Perciformes	0.66*
67	361	<i>Coryphaena hippurus</i>	Mahi mahi	Coryphaenidae	Perciformes	0.61
68	364	<i>Caranx</i> sp.	Jack	Carangidae	Perciformes	0.72*
69	370	<i>Lutjanus campechanus</i>	Red snapper	Lutjanidae	Perciformes	1.40
70	374	<i>Anisotremus virginicus</i>	Porkfish	Haemulidae	Perciformes	0.90**
71	380	<i>Polydactylus</i> sp.	Threadfin	Polynemidae	Perciformes	
72	381	<i>Cynoscion nebulosus</i>	Croaker	Sciaenidae	Perciformes	0.80**
73	391	<i>Girella nigricans</i>	Opaleye	Girellidae	Perciformes	1.10

(continued on next page)

Table 3 (continued)

S #	F #	Species	Common name	Family	Order	G. size
74	393	<i>Chaetodon auriga</i>	Threadfin butterflyfish	Chaetodontidae	Perciformes	0.87*
75	394	<i>Pygoplites diacanthus</i>	Regal angelfish	Pomacanthidae	Perciformes	0.70**
76	409	<i>Pterophyllum scalare</i>	Freshwater angelfish	Cichlidae	Perciformes	1.2**
77	410	<i>Embiotoca jacksoni</i>	Black surfperch	Embiotocidae	Perciformes	1.00
78	411	<i>Dascyllus trimaculatus</i>	Three-spot damselfish	Pomacentridae	Perciformes	1.05**
79	411	<i>Hypsypops rubicundus</i>	Garibaldi	Pomacentridae	Perciformes	1.05**
80	412	<i>Semicossyphus pulcher</i>	California sheephead	Labridae	Perciformes	0.95**
81	412	<i>Halichoeres hortulanus</i>	Chekerboard wrasse	Labridae	Perciformes	0.95**
82	414	<i>Scarus ghobban</i>	Blue chinned parrotfish	Scaridae	Perciformes	1.9*
83	416	<i>Zoarcetes</i> sp.	Eelpout	Zoarcidae	Perciformes	0.91*
84	416	<i>Thermarces cerberus</i>	Vent eelpout	Zoarcidae	Perciformes	0.91**
85	427	<i>Dissostichus eleginoides</i>	Chilean Seabass	Nototheniidae	Perciformes	1.02*
86	441	<i>Ammodytes americanus</i>	Sand lance	Ammodytidae	Trachiniformes	
87	448	<i>Gibbonsia montereyensis</i>	Crevice kelpfish	Clinidae	Perciformes	1.0**
88	452	<i>Gobiesox maendricus</i>	Northern clingfish	Gobiesocidae	Gobiesociformes	
89	457	<i>Bostrychus sinensis</i>	Four-eyed Sleeper	Eleotridae	Perciformes	1.25
90	457	<i>Dormitator latifrons</i>	Pacific fat sleeper	Eleotridae	Perciformes	1.25
91	460	<i>Gillichthys mirabilis</i>	Longjawed mudsucker	Gobiidae	Perciformes	1.3**
92	467	<i>Siganus luridus</i>	Dusky Rabbitfish	Siganidae	Perciformes	0.65*
93	469	<i>Zanclus cornuta</i>	Moorish idol	Zanclidae	Perciformes	
94	470	<i>Naso elegans</i>	Elegant unicornfish	Acanthuridae	Perciformes	0.72**
95	472	<i>Barracuda barracuda</i>	Gret barracuda	Sphyrnaeidae	Perciformes	0.69
96	475	<i>Rastrelliger kanagurta</i>	Indian mackerel	Scombridae	Perciformes	0.90**
97	476	<i>Xiphias gladius</i>	Swordfish	Xiphiidae	Perciformes	0.88
98	486	<i>Colisa lalia</i>	Dwarf Gourami	Osphronemidae	Perciformes	0.70**
99	486	<i>Betta splendens</i>	Siamese fighting fish	Osphronemidae	Perciformes	0.64
100	510	<i>Diodon histrix</i>	Porcupine fish	Diodontidae	Tetraodontiformes	0.90**

We plan to assemble genomes based on 100× to 150× coverage paired-end reads from libraries of various insert lengths and therefore expect to have a good coverage of the genomes. Our goal is to obtain an assembly with a N50 contig size of 50 kb and N50 scaffold size of 5 Mb. However, achieving this goal depends largely on the types and extent of repetitive sequences in the targeted genomes. For gold standard species, where quality RNAs will be collected, we also plan to obtain entire transcriptomes by RNA-seq. As per the Toronto Statement, the maximum time between the end of data collection and submission of the genome paper will be 1 year, after which global analyses and publication by the community would be unimpeded.

Acknowledgement

We would like to thank Michael Berumen, KAUST, for providing samples from the Red Sea.

References

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.-M., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., Gelpke, M.D.S., Roach, J., Oh, T., Ho, I.Y., Wong, M., Detter, C., Verhoeve, F., Predki, P., Tay, A., Lucas, S., Richardson, P., Smith, S.F., Clark, M.S., Edwards, Y.J.K., Doggett, N., Zharkikh, A., Tavtigian, S.V., Pruss, D., Barnstead, M., Evans, C., Baden, H., Powell, J., Glusman, G., Rowen, L., Hood, L., Tan, Y.H., Elgar, G., Hawkins, T., Venkatesh, B., Rokhsar, D., Brenner, S., 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* (New York, N.Y.) 297, 1301–1310.

G10KCOS, 2009. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J. Hered.* 100, 659–674.

Gregory, T.R., 2011. Animal Genome Size Database. <http://www.genomesize.com> 2011.

Hinegardner, R., Rosen, D.E., 1972. Cellular DNA content and the evolution of Teleostean fishes. *Am. Nat.* 106, 621–644.

Jailon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., et al., 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431, 946–957.

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A., Sasaki, T., Shimizu, A., Asakawa, S., Shimizu, N., Hashimoto, S.-I., Yang, J., Lee, Y., Matsushima, K., Sugano, S., Sakaizumi, M., Narita, T., Ohishi, K., Haga, S., Ohta, F., Nomoto, H., Nogata, K., Morishita, T., Endo, T., Shin-I, T., Takeda, H., Morishita, S., Kohara, Y., 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714–719.

Nelson, J.S., 2006. *Fishes of the World*, 4th edn. John Wiley and Sons, Hoboken, New Jersey.

Smith, E.M., Gregory, T.R., 2009. Patterns of genome size diversity in the ray-finned fishes. *Hydrobiologia* 625, 1–25.

Star, B., Nederbragt, A.J., Jentoft, S., Grimholt, U., Malmstrøm, M., Gregers, T.F., Rounge, T.B., Paulsen, J., Solbakken, M.H., Sharma, A., Wetten, O.F., Lanzén, A., Winer, R., Knight, J., Vogel, J.-H., Aken, B., Andersen, Ø., Lagesen, K., Tooming-Klunderud, A., Edvardsen, R.B., Tina, K.G., Espelund, M., Nepal, C., Previti, C., Karlsten, B.O., Møum, T., Skage, M., Berg, P.R., Gjøen, T., Kuhl, H., Thorsen, J., Malde, K., Reinhardt, R., Du, L., Johansen, S.D., Searle, S., Lien, S., Nilsen, F., Jonassen, I., Omholt, S.W., Stenseth, N.C., Jakobsen, K.S., 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 7–10.