

RESEARCH ARTICLE

Open Access

Second generation physical and linkage maps of yellowtail (*Seriola quinqueradiata*) and comparison of synteny with four model fish

Jun-ya Aoki^{1*}, Wataru Kai¹, Yumi Kawabata¹, Akiyuki Ozaki², Kazunori Yoshida³, Takashi Koyama⁴, Takashi Sakamoto⁴ and Kazuo Araki^{1*}

Abstract

Background: Physical and linkage maps are important aids for the assembly of genome sequences, comparative analyses of synteny, and to search for candidate genes by quantitative trait locus analysis. Yellowtail, *Seriola quinqueradiata*, is an economically important species in Japanese aquaculture, and genetic information will be useful for DNA-assisted breeding. We report the construction of a second generation radiation hybrid map, its synteny analysis, and a second generation linkage map containing SNPs (single nucleotide polymorphisms) in yellowtail.

Results: Approximately 1.4 million reads were obtained from transcriptome sequence analysis derived from 11 tissues of one individual. To identify SNPs, cDNA libraries were generated from a pool of 500 whole juveniles, and the gills and kidneys of 100 adults. 9,356 putative SNPs were detected in 6,025 contigs, with a minor allele frequency $\geq 25\%$. The linkage and radiation hybrid maps were constructed based on these contig sequences. 2,081 markers, including 601 SNPs markers, were mapped onto the linkage map, and 1,532 markers were mapped in the radiation hybrid map.

Conclusions: The second generation linkage and physical maps were constructed using 6,025 contigs having SNP markers. These maps will aid the *de novo* assembly of sequencing reads, linkage studies and the identification of candidate genes related to important traits. The comparison of marker contigs in the radiation hybrid map indicated that yellowtail is evolutionarily closer to medaka than to green-spotted pufferfish, three-spined stickleback or zebrafish. The synteny analysis may aid studies of chromosomal evolution in yellowtail compared with model fish.

Keywords: Yellowtail, Radiation hybrid map, Linkage map, Synteny analysis

Background

Teleostei is the most numerous and diversified group of vertebrates. Among them, the order Perciformes is widely distributed throughout the world and is the most diversified of all fish orders. Perciformes are also evolutionarily interesting, because the order includes fish species with various features and forms. However, there is insufficient genomic information for the fishes in this order to allow analyses of chromosomal evolution and the identification of important trait loci for breeding. Yellowtail (*Seriola quinqueradiata*), a member of the order Perciformes, is a popular and important species in

Japanese aquaculture. In this study, we report the results of transcriptome analysis, including single nucleotide polymorphism (SNP) identification.

Sequencing of expressed sequence tags (ESTs) derived from cDNA libraries is considered a useful method to identify transcripts in species that lack a sequenced genome [1]. However, obtaining a high-quality EST database using Sanger's sequencing method is time consuming and expensive. Consequently, this method has been restricted to a few aquaculture fishes, such as Atlantic salmon [2] rainbow trout [3], and catfish [4]. Recently, next generation sequencing, e.g. Illumina sequencing, has been developed and has allowed transcriptome analysis in aquaculture fish species, such as catfish [5], silver carp [6], and large yellow croaker [7], and has permitted the identification of gene-associated markers in

* Correspondence: junaoki@affrc.go.jp; arakin@affrc.go.jp

¹National Research Institute of Aquaculture, Fisheries Research Agency, 224-1 Hiruta, Tamaki-cho, Watarai-gun, Mie 519-0423, Japan

Full list of author information is available at the end of the article

catfish [8], carp [9] and rainbow trout [10]. The parallel pyrosequencing technology, commercialized by 454 pyrosequencing, is another next generation sequencing technology and has been used to identify gene-associated SNPs derived from ESTs, e.g. in the blunt snout bream [11], lake whitefish [12], cutthroat trout [13], Atlantic herring [14] and Atlantic cod [15]. We developed this pyrosequencing technology to find SNPs in the transcripts of yellowtail. SNPs in coding regions are one of the most important DNA variants for quantitative trait locus (QTL) mapping, because some Mendelian and genetically complex traits are caused by SNPs in coding regions [16].

Physical and linkage maps are important aids in the assembly of genome sequences, for comparative analysis of synteny and for identifying candidate genes by QTL analysis. High-resolution linkage maps have been constructed for various fishes, such as medaka, zebrafish, Atlantic salmon [17], rainbow trout [18], Asian seabass [19], catfish [20], Atlantic cod [21], pufferfish (fugu) [22] and carp [23]. In genetic studies of *Seriola*, the first linkage map between *S. quinqueradiata* and *S. lalandi* was produced by Ohara et al. [24]. Recently, the linkage map has been improved by the addition of more simple sequence repeat (SSR) markers, derived from bacterial artificial chromosome (BAC) end sequences [25]. In this study, a second generation linkage map was produced, represented by 24 linkage groups and having 601 SNP markers.

Moreover, we report the construction of a second generation radiation hybrid (RH) map. An RH map is a powerful tool for building a physical map of the whole genome, and can be used to aid the construction of gene orthology relationships through conserved synteny analysis. Seven RH maps have been reported in six teleost fishes, including two maps of zebrafish [26,27], gilthead sea bream [28], medaka [29], European sea bass [30], Nile tilapia [31] and yellowtail [32]. The second generation RH map of yellowtail has more than 1,500 markers and will be useful for genome sequence assembly and comparative analysis of synteny. The release of the complete genome information of the model fish species medaka (*Oryzias latipes*), zebrafish (*Danio rerio*), green-spotted pufferfish (*Tetraodon nigroviridis*), and three-spined stickleback (*Gasterosteus aculeatus*) has accelerated the evolutionary studies. Finally, we performed synteny analysis to compare yellowtail with model fish species.

Results and discussion

De novo transcriptome assembly

A cDNA library was generated from pooled RNA samples extracted from 11 tissues from a single individual. Sequencing on the Roche/454 GS FLX Titanium platform generated 1,353,405 reads. The CLC Genomic Workbench carried out the *de novo* assembly. After trimming the adapters and filtering out the low-quality and short reads, 1,345,753 high-

quality reads were assembled into 56,449 contigs, with 276,945 reads remaining as singletons. The average length of the contigs was 782 bp, and the N_{50} size was 959 bp.

Gene ontology analysis

Of 56,449 contigs, 24,035 (43%) had a significant hit and matched 15,280 unique protein records in the nr protein database. Gene ontology (GO) analysis was conducted on these 24,035 contigs. 17,076 sequences were assigned to at least one GO term describing three functional groups: biological process, molecular function and cellular component. Summaries of the level 2 GO assignments are shown in Figure 1. Among the 17,076 sequences, the molecular function ontology comprised the majority of GO assignments (83%), followed by biological processes (79%) and cellular components (74%). In the molecular function category, binding and catalytic activity represented about 80% of the total. For biological processes, sequences involved in cellular processes (19%), metabolic processes (18%), and biological regulation (13%) were highly represented. Finally, the cell and organelle term represented about 70% of the cellular component. Transcriptome assembly using next generation DNA sequencing and GO analysis has been reported in other fish species [33-35]. Our GO analysis of yellowtail transcriptome revealed similar results to those of other fishes.

SNP identification

Sequencing produced 570,846 raw reads derived from the full-length library and 456,482 raw reads derived from the 3'-anchored library. Quality-based variant calling using the CLC Genomics Workbench detected 9,356 biallelic putative SNPs in 6,025 contigs, with a minor allele frequency (MAF) $\geq 25\%$. SNPs with a high allele frequency are more suited for constructing a linkage map efficiently using genotyping of one family because polymorphisms of SNPs decrease in one family. These contigs were registered to DDBJ/EMBL/GenBank as accession number FX884179–FX890203.

Mapping of SNP markers to the linkage map

Direct sequencing identified 143 informative SNPs that were heterozygous in either one of the parents, and these heterozygous SNPs were used for linkage analysis using the F_1 mapping progeny. In the SNPtype assay, 458 SNPs were mapped to the linkage map (Tables 1 and 2, Figure 2, Additional file 1). 2081 markers containing 601 SNPs, which were polymorphic in the F_1 mapping progeny, were mapped in the linkage map, and 606 markers were common for both sexes. In this study, SNPs were mapped for the first time in a yellowtail linkage map.

In our genotyping analysis, many polymorphisms in wild yellowtail were observed; however, using one family decreased the SNP frequency to about 10%. Moreover,

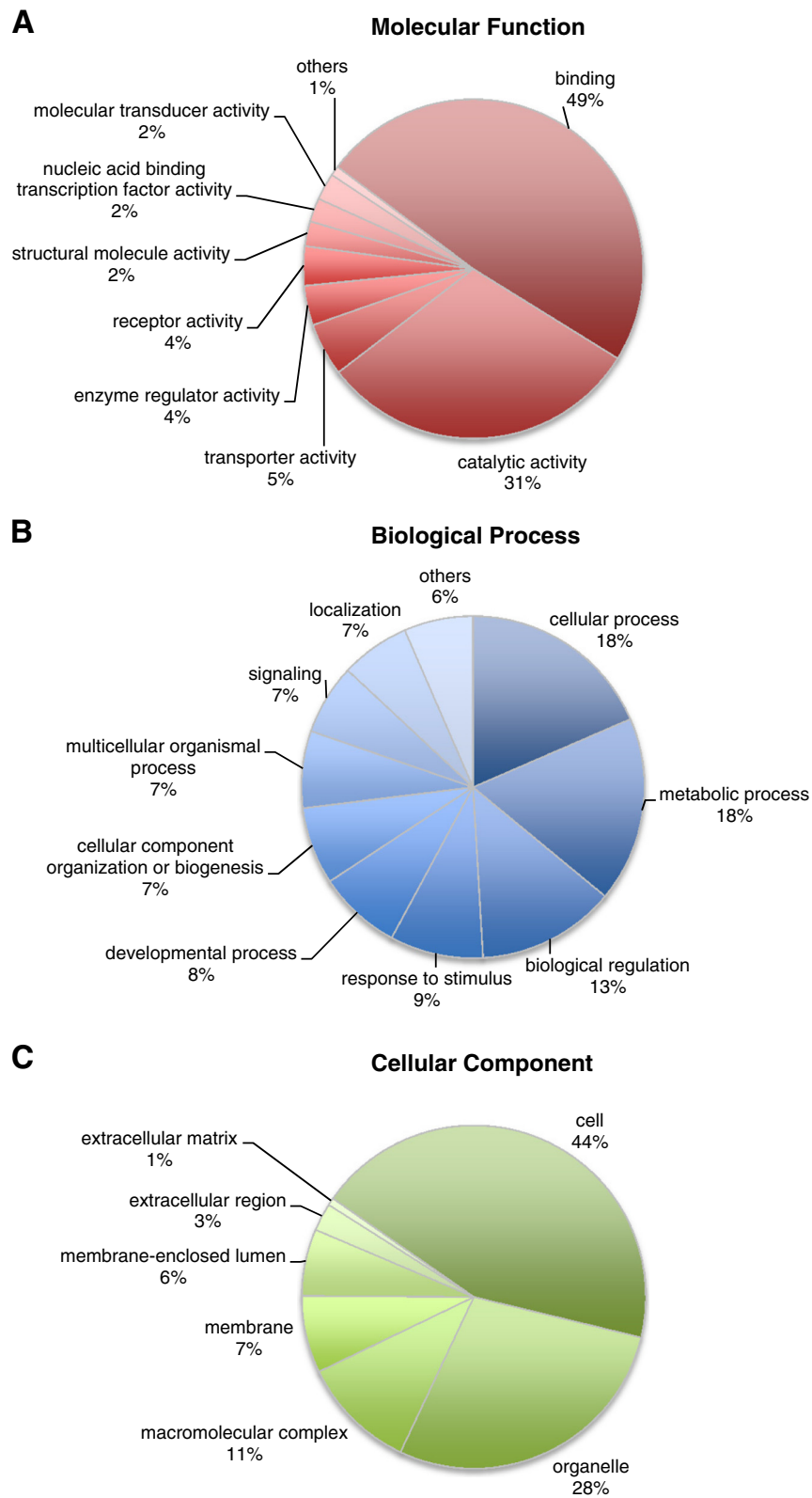


Figure 1 Gene Ontology assignment for assembled contigs. **(A)** Molecular function, **(B)** Biological process and **(C)** Cellular component assignment.

Table 1 Summary of the yellowtail genetic linkage map

Linkage group	Female			Male				
	Length (cM)	Number of all markers	Number of SNP markers	Number of SSR markers	Length (cM)	Number of all markers	Number of SNP markers	Number of SSR markers
Squ1	47.49	53	20	33	49.28	48	14	34
Squ2	49.16	82	11	71	44.53	77	11	66
Squ3	40.18	38	9	29	52.71	40	11	29
Squ4	47.35	35	7	28	63.16	36	11	25
Squ5	59.55	40	2	38	50.41	52	13	39
Squ6	51.40	44	14	30	42.62	37	14	23
Squ7	50.60	44	15	29	38.24	48	16	32
Squ8	55.14	41	15	26	51.87	37	14	23
Squ9	30.20	47	9	38	52.93	49	13	36
Squ10	40.39	53	16	37	49.61	54	18	36
Squ11	0	18	8	10	54.82	25	13	12
Squ12	43.99	45	18	27	50.96	43	15	28
Squ13	16.71	29	4	25	47.11	35	10	25
Squ13'	7.84	2	1	1	-	-	-	-
Squ14	24.84	41	8	33	45.10	46	12	34
Squ14'	2.22	2	0	2	-	-	-	-
Squ15	36.98	52	17	35	52.74	50	14	36
Squ16	36.17	36	11	25	51.52	39	12	27
Squ17	38.02	28	7	21	50.47	34	11	23
Squ18	45.56	48	12	36	65.96	42	9	33
Squ19	40.56	45	15	30	59.43	41	15	26
Squ20	51.19	47	16	31	44.66	45	14	31
Squ21	44.88	39	10	29	54.78	36	9	27
Squ22	45.91	33	8	25	53.08	37	12	25
Squ23	41.05	38	13	25	44.05	41	15	26
Squ24	81.86	59	22	37	57.91	50	17	33
Total	1029.24	1039	288	751	1227.95	1042	313	729

we used a nanofluidic dynamic array to perform high-throughput genotyping against targeted SNPs. We considered that the nanofluidic dynamic array was useful to genotype SNPs in one family, as well as sequencing analysis by Sanger's method.

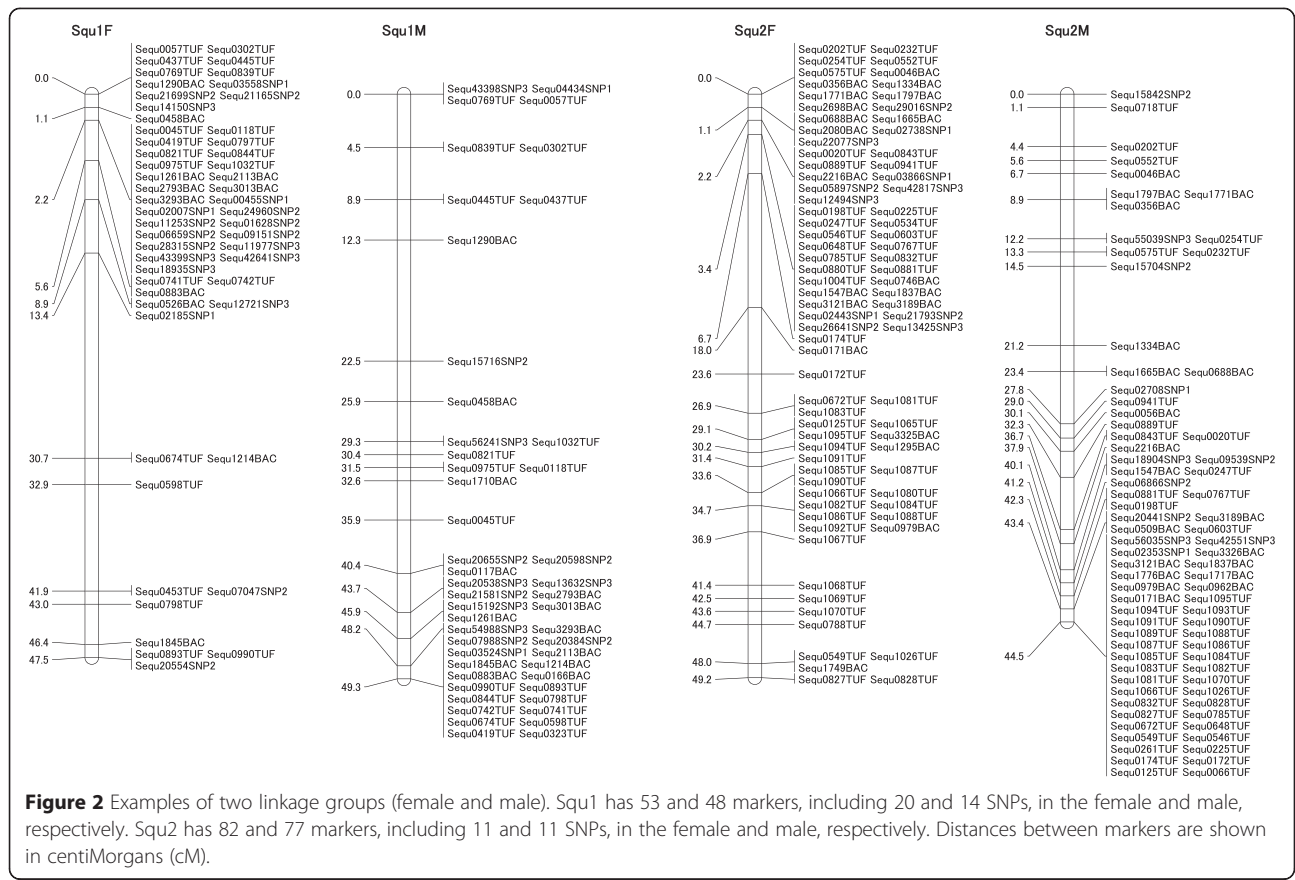
Construction of the RH map

PCR on a dynamic array produces high-throughput gene expression data that are essentially identical in quality to

Table 2 Summary of the markers of the linkage map

Total genetic markers	2081
Total mapped SSR markers	1480
Total mapped SNP markers	601
SNPs identified by direct sequencing	143
SNPs identified by SNPtype assay	458

conventional microliter qRT-PCR and are superior to publicly available array data from the same tissue type [36]. In our previous study, 580 markers were mapped in the first RH map [32]. 1,563 markers, containing the previous 580 markers, were used to construct the RH map. The two-point analysis, performed at a LOD score of 4.0 and a distance threshold of 50, resulted in 75 groups using CarthaGene software [37,38]. Furthermore, with reference to the locations of several markers on the constructed genetic linkage map, 1,532 markers (1433 EST markers containing SNPs and 99SSR markers) were distributed to 24 linkage groups (Tables 3 and 4, Figure 3, Additional file 2). Thirty-one markers in eight groups were not distributed among the genetic linkage groups because of their LOD score or a distance threshold error. The RH map was constructed with a final set of 1,532 markers. In each group, the RH map ranged from



640.7 to 1,343.3 centiRays (cR), with an average of approximately 1,096 cR. The combined size of all RH groups was 26,293.5 cR. The estimated size of the yellowtail genome is 800 Mbp [24], which inferred a value of 1 cR = 30 kbp (800 Mbp/26,293.5 cR). We have obtained a large quantity of data in a short time using the BioMark™ HD system since the construction of the first RH map.

The RH map was compared with the linkage map to confirm the accuracy of the local order of markers (Additional file 3). In linkage group 11 of mapped to 0 cM, because chromosome recombination is unlikely to occur during meiosis. It is possible to map these genes accurately on the physical map using RH. The position of some markers was different between the RH and linkage maps, because the physical lengths of the RH map are different from the genetic lengths of the linkage map. In addition, we supposed which region of the RH map is not recombined at meiosis by comparing the RH and the linkage map. The accuracy of the local order of markers will be confirmed when the whole genome sequence becomes available. Currently, we are trying to map genome contigs of yellowtail onto the physical map.

Syntenic relationship with model fish

The 1,433 yellowtail EST marker sequences of the RH map were compared with the cDNA sequences of four fish species: medaka (*Oryzias latipes*), zebrafish (*Danio rerio*), three-spined stickleback (*Gasterosteus aculeatus*) and green-spotted pufferfish (*Tetraodon nigroviridis*) using the TBLASTX algorithm. Among the 1,433 yellowtail marker sequences, 1,036 genes (72.3%) had homologs in medaka, 1,064 genes (74.2%) had homologs in zebrafish, 1,073 genes (74.9%) had homologs in three-spined stickleback, and 1,032 genes (72.0%) had homologs in green-spotted pufferfish (Additional file 4). These values were not significantly different among the four fishes. Oxford grids between yellowtail and the four fish species are shown in Figure 4. The modal number of chromosomes in yellowtail is 48 [32]. This number is the same as that of medaka; all chromosomes could be paired one-to-one between yellowtail and medaka. The number of chromosomes in the two fishes is the same and their chromosomal structures are similar. These results suggested that we would observe conserved synteny between yellowtail and medaka, and that the syntenic relationship between yellowtail and zebrafish would be rather low. However, they

Table 3 Summary of the yellowtail radiation hybrid (RH) map

Group	Size (cR)	No. of markers		
		Genes	SSRs	Total
1	1327.1	75	4	79
2	1337.0	64	6	70
3	1151.0	55	5	60
4	1187.1	65	4	69
5	720.4	42	4	46
6	1139.4	61	4	65
7	982.3	56	3	59
8	968.0	57	4	61
9	1161.1	65	3	68
10	1132.1	60	4	64
11	640.7	33	3	36
12	1233.3	92	9	101
13	1221.8	56	4	60
14	1152.7	51	3	54
15	1343.3	71	4	75
16	1283.6	56	4	60
17	1047.6	63	5	68
18	897.4	52	3	55
19	1287.6	64	4	68
20	1126.1	59	3	62
21	909.8	51	4	55
22	764.0	48	4	52
23	968.1	58	4	62
24	1312.0	79	4	83
Subtotal	26293.5	1433	99	1532
Unlinked		31		31
Total	26293.5	1464	99	1563

cR: centiRay.

might not be evolutionarily closer than the relationships between yellowtail and medaka. Medakas are a member of the order *Beloniformes*, which includes freshwater and marine fish, such as Pacific saury and flying fish. Zebrafish are a member of *Cypriniformes*, which consists exclusively of freshwater fish. Our synteny results reflected the known taxonomic relationships of these fishes. In

Table 4 Summary of the markers of the RH map

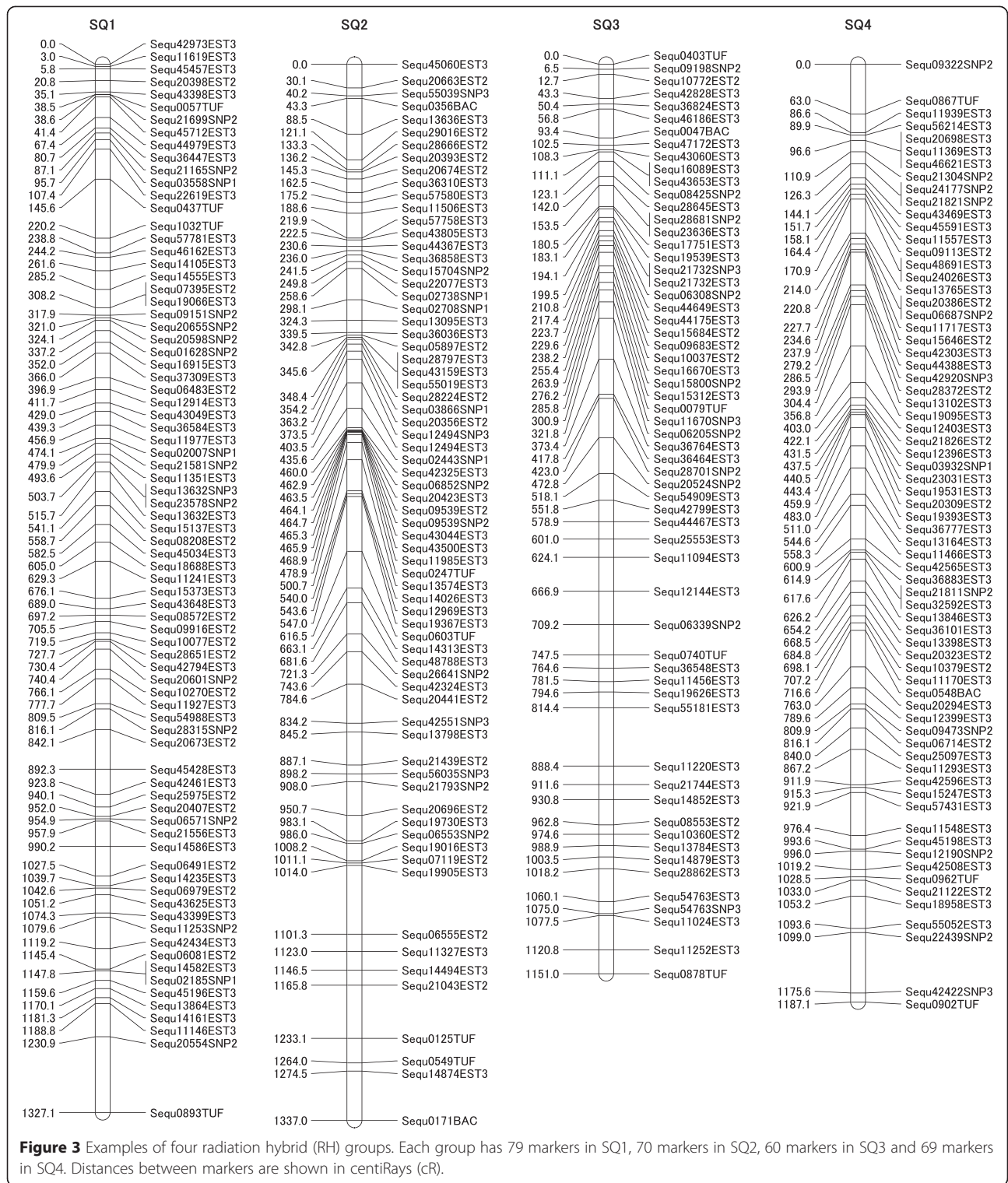
Total genotyped markers	1563
Total mapped markers	1532
Mapped EST markers containing SNPs	1433
Mapped SSR	99
Mapped markers in the previous study [32]	580

Teleostei, it is thought that a whole genome duplication and eight subsequent major rearrangements occurred about 314–404 million years ago [39]. Moreover, medaka and zebrafish are thought to have diverged after the eight major rearrangement events, after which medaka and green-spotted pufferfish diverged. The chromosome number in three-spined sticklebacks is the same as that of green-spotted pufferfish. Furthermore, the Oxford grid between yellowtail and green-spotted pufferfish is similar to that between yellowtail and three-spined sticklebacks (Figure 4). However, in the Oxford grid between green-spotted pufferfish and three-spined sticklebacks, the chromosome groups were not paired one-to-one (Additional file 5). Yellowtail SQ11 and 14 correspond to linkage group 1 of three-spined sticklebacks, SQ5 and 18 correspond to linkage group 4, and SQ19 and 22 correspond to linkage group 7, respectively. Then, yellowtail SQ4 and 18 correspond to chromosome 1 of green-spotted pufferfish, SQ6 and 7 correspond to chromosome 2, and SQ11 and 12 correspond to chromosome 3, respectively. These chromosomes in three-spined sticklebacks and green-spotted pufferfish could be paired one-to-two with those of yellowtail. The number of chromosomes in three-spined sticklebacks and green-spotted pufferfish is $N=21$ and that of yellowtail is $N=24$. Thus, three-spined sticklebacks and green-spotted pufferfish have three fewer chromosomes than yellowtail; each chromosome is thought to have merged after the divergence from medaka about 191.8 M years ago. By analysis of the whole genome sequence, conserved segments and/or conserved segment orders will be distinguished using the RH map, and will aid studies of the dynamics of chromosome evolution between yellowtail and model fish species.

A putative sex determination locus of yellowtail was located in the Squ12 linkage group [40]; however, its related gene has not been identified. The sex determination gene in fish varies according to fish species: *Dmy* in medaka [41], *Amhy* in Patagonian pejerrey [42], *Amhr2* in fugu [43] and *sdY* in the rainbow trout [44]. Therefore, the identification of sex determination genes is difficult, especially when the conservation of these loci is not high. Detailed chromosome information is provided by analysis of the RH map and whole genome sequences. We are interested in whether yellowtail appeared earlier in evolution than medaka, and when yellowtail diverged from medaka after the major rearrangements of the chromosomes. Currently, we are studying SNPs at the genome level and performing DNA chip analyses using SNPs in ESTs.

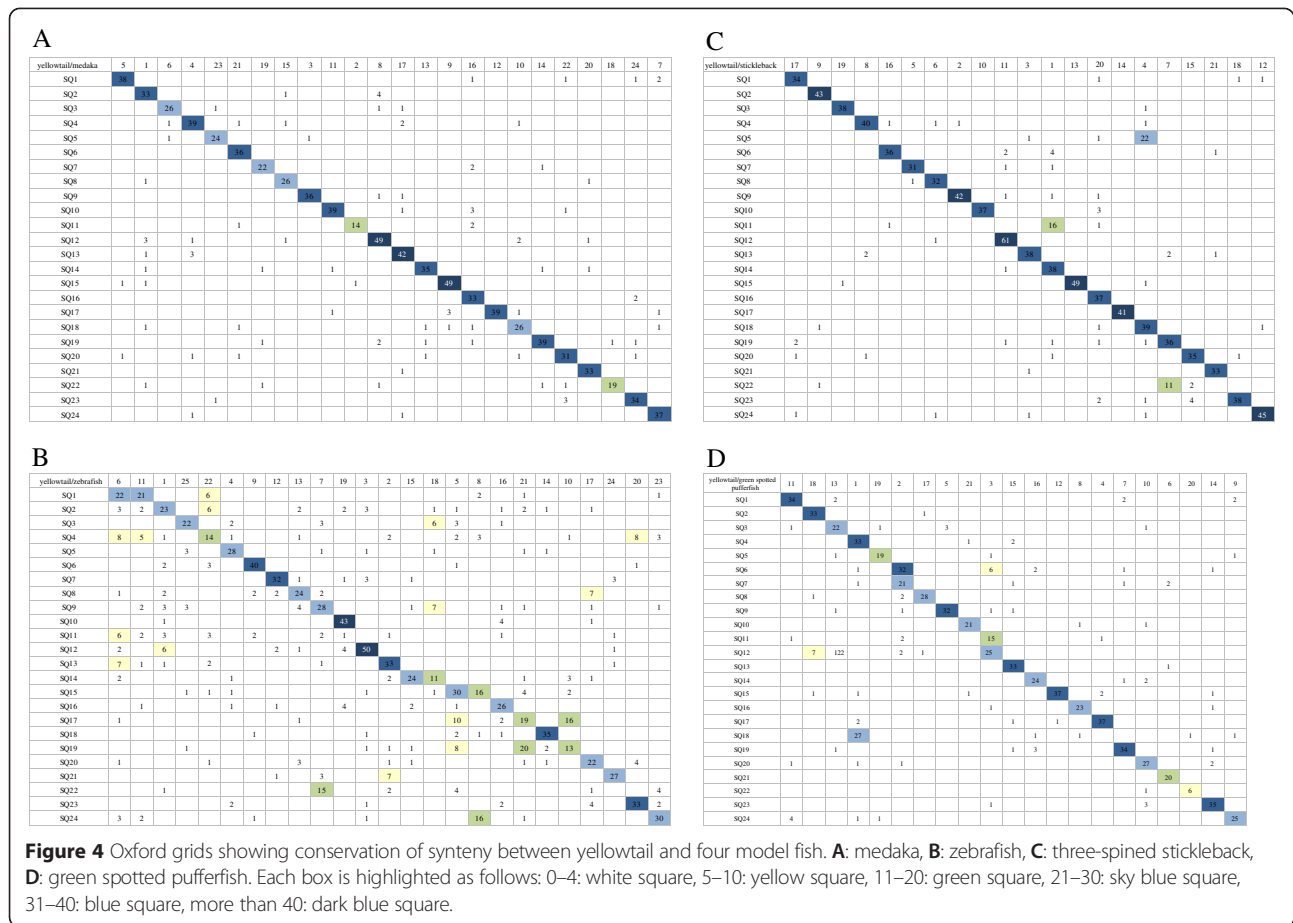
Conclusions

In the transcriptome analysis, 9,356 SNPs were identified in 6,025 contigs. A linkage map for yellowtail was then



constructed, which comprised 24 linkage groups including 601 SNP markers, and an RH map was constructed with 1,532 markers. Our synteny results indicated conserved synteny between yellowtail and medaka, and that the syntenic relationship between yellowtail and zebrafish would

be rather low: these results reflect the known taxonomic relationships of these fish. The high-density maps will help the assembly of the genome sequence in the comparative analysis of synteny and will aid the search for candidate genes using quantitative trait locus (QTL) analysis.



Methods

Ethics statement

Field permits are not required for this species in Japan. The Institute Animal Care and Use Committee of the National Research Institute of Aquaculture (IACUC-NRIA No. 3) approved the fish handling, husbandry and sampling methods.

Preparation of a cDNA library for transcriptome *de novo* assembly of a single individual

A single adult yellowtail was captured in the Eastern China Sea off Nagasaki prefecture. Total RNAs were extracted from 11 tissues (muscle, brain, heart, liver, intestine, kidney, spleen, gonad, gill, fin and bladder) using RNAiso Plus (Takara, Shiga, Japan). Eurofins MWG Operon (Hamburg, Germany) performed the cDNA library construction and 454-pyrosequencing. After purification of poly (A)⁺ RNA, first-strand cDNA synthesis was primed with an N6 randomized primer. The 454 adapters A and B were ligated to the 5' and 3' ends of the cDNA. These fragments were finally amplified with 11 cycles of PCR using a proofreading enzyme. Normalization was carried out by one cycle of denaturation and re-association of the cDNA. The re-associated ds-cDNA was separated from the remaining

ss-cDNA by passing the mixture over a hydroxyl apatite column. After hydroxyl apatite chromatography, the ss-cDNA was amplified by seven cycles of PCR. One sequencing run was performed on the Roche/454 GS FLX Titanium platform. All reagents and protocols used were from Roche 454 Life Sciences, USA.

Preparation of cDNAs for SNP discovery

Full-length and 3'-anchored cDNA libraries were prepared from a pool of 500 whole juveniles, and the gills and kidneys from 100 adult yellowtails. All fish were captured in the East China Sea off Nagasaki prefecture. The full-length cDNA library was constructed using the SMART technology (Takara), with slight modifications. First-strand cDNA synthesis was primed with the SMART-Sfi IA oligonucleotide and the CDS-Sfi IB primer using ReverTra Ace[®] (Toyobo, Osaka, Japan). The first-strand cDNA was initially amplified for 20 cycles by Long-Distance PCR using Advantage 2 polymerase (Takara). The products were purified using the NucleoSpin purification kit (Macherey-Nagel, Duren, Germany). The cDNA was then normalized with the TRIMMER-DIRECT Kit (Evrogen, Moscow Russia), which uses duplex-specific nuclease treatment. The normalized

cDNA samples were diluted and used for PCR amplification with 18 cycles using Advantage 2 polymerase (Clontech, Shiga, Japan). The library was finally purified using a NucleoSpin Extract II (Macherey-Nagel). The 3'-anchored cDNA library was constructed according to the method of Eveland et al. [45]. First-strand cDNA synthesis was primed with 6 pmol of biotinylated (T₁₂) B-adaptor modified from Margulies et al. [46]. After purification using a DNAClear Purification kit (Invitrogen, CA, USA), the cDNA was bound to Dynabeads M-270Streptavidin (Invitrogen) and digested with *MspI* (Promega, WI, USA). An adaptor modified from Margulies et al. [46] was ligated to the restriction site of the cDNA. The template strand was eluted with 10 mM Tris-HCl (pH 8.5) and purified using the QIAquick column (Qiagen, Hilden, Germany). Each cDNA library was sequenced on the Roche/454 GS FLX Titanium platform by Takara Dragon Genomics Center in Japan.

De novo assembly

The raw pyrosequencing data obtained from a cDNA library of a single individual were processed using CLC Genomics Workbench (CLC Bio, Denmark). Adaptor sequences were trimmed, and low-quality reads were filtered out using the following parameters: ambiguous limit of 2; quality limit of 0.05; minimum read length of 50 nt. *De novo* assembly of the processed reads was performed using the following parameters: word size of 22 (automated calculation); bubble size of 445 (automated calculation); mismatch cost of 2; insertion cost of 3; deletion cost of 3; length fraction of 0.8; and similarity of 0.8.

Gene ontology analysis

Functional annotation of the assembled contigs was performed by BLASTx searching against the NCBI non-redundant (nr) protein database, using the Blast2GO suite [47] with a cutoff *E*-value of e^{-6} . Gene ontology (GO) analysis of the assembled contigs provides an ontology of defined terms representing gene product properties. The ontology covers three functional groups: biological process, molecular function, and cellular component. To examine broad-level classifications of gene functions, we mapped the GO terms to the Generic GO-Slim database using the Blast2GO suite.

SNPs discovery

The CLC Genomics Workbench (CLC Bio) was used to analyze the sequencing results of the cDNA libraries. CLC bio's *de novo* assembly algorithm uses de Bruijn graphs to represent overlapping reads, which is a common approach for short read *de novo* assembly that allows a large number of reads to be handled efficiently [48-50]. Adaptor sequences were trimmed, and low-quality reads were filtered out using the following parameters: ambiguous limit

of 2 and a quality limit of 0.05. The processed reads were mapped to the reference sequence generated from the assembled contigs of a single individual yellowtail, using the following parameters: mismatch cost of 2; insertion cost of 3; deletion cost of 3; length fraction of 0.8; similarity of 0.8; and non-specific match handling=ignore.

SNP detection based on the neighborhood quality standard algorithm [51,52] using the CLC Genomics Workbench was performed using the following parameters: neighborhood radius of 5; maximum gap and mismatch count of 0; minimum neighborhood base quality of 15; minimum central base quality of 20; non-specific match handling = ignore; minimum coverage of 10; minimum allele frequency (MAF) of 5%; and filter homopolymer indels = yes. Only biallelic SNPs were considered.

Construction of a linkage map

To construct the second generation linkage map, a mapping population was obtained by artificial crossing, as described previously [25]. Direct sequencing of the PCR products and the SNPtype assay of BioMark™ HD system in Fluidigm (USA) were used to perform SNP genotyping.

For direct sequencing, Primer 3 software [53] was used to design primers flanking the SNP sites. PCR amplification was generated using Hot-Start Gene *Taq* (Nippon gene, Tokyo, Japan), and performed over 40 cycles: 15 sec at 95°C, 10 sec at 59°C, 40 sec at 72°C, in addition to a 5-min predenaturation at 95°C and a 3-min postcycle incubation at 72°C. The obtained PCR products of the parents and two offspring of one yellowtail family were sequenced directly using the ABI3130xl Genetic Analyzer (Applied Biosystems, CA, USA) or ABI3730xl DNA Analyzer (Applied Biosystems).

In the SNPtype assay, SNP genotyping of the parents and 94 offspring of one yellowtail family was carried out according to the protocol of the manufacturer. Specific-Target Amplification (STA) primers and Locus-Specific Primers (LSP) for each SNP locus were pooled, and a pre-amplification was performed using 2× Multiplex PCR Master Mix (Qiagen) with the following PCR protocol: 95°C for 15 min, followed by 14 cycles of 95°C for 15 sec and 60°C for 4 min. Pre-amplified DNA was then diluted 1:100 in DNA suspension buffer and 2.5 μL combined with 2× Fast Probe Master Mix (Biotium, CA, USA), 20× Sample Loading Reagent (Fluidigm), and the reference dye ROX (Invitrogen). In parallel, each SNPtype assay mix included allele-specific primers (ASP1 and 2) and LSP was mixed with 2× Assay Loading Reagent. These were then loaded onto 96.96 Dynamic Arrays IFC (Fluidigm), which uses nanofluidic circuitry to combine 96 loci with 96 samples in 9,216 reaction chambers. Thermo-cycling and fluorescence detection were conducted on the BioMark™ HD system (Fluidigm). Finally, linkage analysis was performed as described previously [25]. Linkage analysis was performed using

genotype data converted to a backcross format. As the grandparent genotypes were unknown, pairwise analyses were performed, and markers were sorted into linkage groups at a minimum LOD score of 4.0. A goodness-of-fit for Mendelian segregation distortion was tested for all alleles using the chi-square test ($p < 0.05$, d.f. = 1). Finally, the marker order was determined and double recombination events were checked with MapManagerQTX version 2.0 [54]. The resultant genetic map was visualized using MapChart version 2.2 [55].

Construction of an RH map

The RH panel, comprising 93 hybrid RH cell lines, and positive and negative controls, was published previously [32]; their 96 DNA samples were extracted. Genotyping reactions were carried out on the Fluidigm platform using a BioMark™ 96.96 nanofluidic dynamic array (Fluidigm) for gene expression analysis, as described previously [32].

The CarthaGene software [37,38] was used to perform two-point linkage analyses and to determine the marker order and inter-marker distances in cR. CarthaGene looks for multiple populations' maximum likelihood consensus maps using a fast EM algorithm for maximum likelihood estimation and powerful ordering algorithms. The group command at a LOD threshold of 4.0 and a distance threshold of 50 determined the linkage groups, and referenced the linkage map of yellowtail.

Syntenic analysis

EST alignments of mapped markers in the RH map were performed using TBLASTX searches against the cDNA database of zebrafish (*Danio rerio*), medaka (*Oryzias latipes*), three-spined stickleback (*Gasterosteus aculeatus*) and green-spotted pufferfish (*Tetraodon nigroviridis*) at Ensembl [56]. The sequences with the lowest E-value ($< 1.0E-5$) were adopted in each EST alignment.

Additional files

Additional file 1: Linkage map of yellowtail. Distances between markers are shown in centiMorgans (cM).

Additional file 2: RH map of yellowtail. Distances between markers are shown in centiRays (cR).

Additional file 3: Comparison of the radiation hybrid (RH) map and the linkage map. Distances between markers are shown in centiRays (cR) on the RH map and in 10^{-1} centiMorgans (cM) on the linkage map.

Additional file 4: Syntenic analysis between the yellowtail radiation hybrid (RH) map and four model fish.

Additional file 5: Oxford grids showing conservation of synteny in green-spotted pufferfish and three-spined stickleback. (BLASTN: E-value $< 1.0E-10$) Each box was highlighted as follows: 0–100: White square, more than 100: Blue square.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JA and KA conceived and designed the study, and wrote the draft manuscript. KA produced the cDNA libraries for the EST analysis. WK and YK carried out transcriptome analysis and identified the SNPs. WK, YK, AO, TK, KY and TS contributed to the construction of the linkage map. JA carried out construction of the RH map, the comparison of RH and linkage maps, and the synteny analysis among model fish. All authors read and approved the final version of the manuscript.

Acknowledgements

This work was supported by the Program for Promotion of Basic and Applied Research for Innovation in Bio-oriented Industry (BRAIN) in Japan.

Author details

¹National Research Institute of Aquaculture, Fisheries Research Agency, 224-1 Hiruta, Tamaki-cho, Watarai-gun, Mie 519-0423, Japan. ²National Research Institute of Aquaculture, Fisheries Research Agency, 422-1 Nakatsuhamaura, Minamiise-cho, Watarai-gun, Mie 516-0193, Japan. ³Goto Laboratory, Seikai National Fisheries Research Institute, Fisheries Research Agency, 122-7, Nunoura, Tamanoura-cho, Goto, Nagasaki 853-0508, Japan. ⁴Faculty of Marine Science, Tokyo University of Marine Science and Technology, 4-5-7 Konan, Minato-ku, Tokyo 108-8477, Japan.

Received: 14 March 2015 Accepted: 29 April 2015

Published online: 24 May 2015

References

- Marra MA, Hillier L, Waterston RH. Expressed sequence tags—ESTablishing bridges between genomes. *Trends Genet.* 1998;14(1):4–7.
- Rise ML, von Schalburg KR, Brown GD, Mawer MA, Devlin RH, Kuipers N, et al. Development and application of a salmonid EST database and cDNA microarray: Data mining and interspecific hybridization characteristics. *Genome Res.* 2004;14:478–90.
- Govoroun M, Gac FL, Guiguen Y. Generation of a large scale repertoire of Expressed Sequence Tags (ESTs) from normalised rainbow trout cDNA libraries. *BMC Genomics.* 2006;7:196.
- Wang S, Peatman E, Abernathy J, Waldbieser G, Lindquist E, Richardson P, et al. Assembly of 500,000 inter-specific catfish expressed sequence tags and large scale gene-associated marker development for whole genome association studies. *Genome Biol.* 2010;11:R8.
- Liu S, Zhang Y, Zhou Z, Waldbieser G, Sun F, Lu J, et al. Efficient assembly and annotation of the transcriptome of catfish by RNA-Seq analysis of a doubled haploid homozygote. *BMC Genomics.* 2012;13:595.
- Fu B, He S. Transcriptome analysis of silver carp (*Hypophthalmichthys molitrix*) by paired-end RNA sequencing. *DNA Res.* 2012;19:131–42.
- Mu Y, Li M, Ding F, Ding Y, Ao J, Hu S, et al. De novo characterization of the spleen transcriptome of the large yellow croaker (*Pseudosciaena crocea*) and analysis of the immune relevant genes and pathways involved in the antiviral response. *PLoS One.* 2014;9:e97471.
- Liu S, Zhou Z, Lu J, Sun F, Wang S, Liu H, et al. Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. *BMC Genomics.* 2011;12:53.
- Xu J, Ji P, Zhao Z, Zhang Y, Feng J, Wang J, et al. Genome-wide SNP discovery from transcriptome of four common carp strains. *PLoS One.* 2012;7:e48140.
- Salem M, Vallejo RL, Leeds TD, Palti Y, Liu S, Sabbagh A, et al. RNA-seq identifies SNP markers for growth traits in rainbow trout. *PLoS One.* 2012;7:e36264.
- Gao Z, Luo W, Liu H, Zeng C, Liu X, Yi S, et al. Transcriptome analysis and SSR/SNP markers information of the blunt snout bream (*Megalobrama amblycephala*). *PLoS One.* 2012;7:e42367.
- Renaut S, Nolte AW, Bernatchez L. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Mol Ecol.* 2010;19 Suppl 1:115–31.
- Houston DD, Elzinga DB, Maughan PJ, Smith SM, Kauwe JSK, Evans RP, et al. Single nucleotide polymorphism discovery in cutthroat trout subspecies using genome reduction, barcoding, and 454 pyro-sequencing. *BMC Genomics.* 2012;13:724.

14. Helyar SJ, Limborg MT, Bekkevold D, Babbucci M, Van Houdt J, Maes GE, et al. SNP discovery using next generation transcriptomic sequencing in Atlantic herring (*Clupea harengus*). *PLoS One*. 2012;7(8):e42089.
15. Carlsson J, Gauthier DT, Carlsson JEL, Coughlan JP, Dillane E, Fitzgerald RD, et al. Rapid, economical single-nucleotide polymorphism and microsatellite discovery based on *de novo* assembly of a reduced representation genome in a non-model organism: a case study of Atlantic cod *Gadus morhua*. *J Fish Biol*. 2013;82:944–58.
16. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. *Science*. 2002;298:2345–9.
17. Lien S, Gidskehaug L, Moen T, Hayes BJ, Berg PR, Davidson WS, et al. A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genomics*. 2011;12:615.
18. Rexroad III CE, Palti Y, Gahr SA, Vallejo RL. A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*). *BMC Genet*. 2008;9:74.
19. Wang CM, Bai ZY, He XP, Lin G, Xia JH, Sun F, et al. A high-resolution linkage map for comparative genome analysis and QTL fine mapping in Asian seabass, *Lates calcarifer*. *BMC Genomics*. 2011;12:174.
20. Ninwichian P, Peatman E, Liu H, Kucuktas H, Somridhivej B, Liu S, et al. Second-generation genetic linkage map of catfish and its integration with the BAC-based physical map. *G3*. 2012;2:1233–41.
21. Hubert S, Higgins B, Borza T, Bowman S. Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC Genomics*. 2010;11:191.
22. Kai W, Kikuchi K, Tohari S, Chew AK, Tay A, Fujiwara A, et al. Integration of the genetic map and genome assembly of fugu facilitates insights into distinct features of genome evolution in teleosts and mammals. *Genome Biol Evol*. 2011;3:424–42.
23. Zhao L, Zhang Y, Ji P, Zhang X, Zhao Z, Hou G, et al. A dense genetic linkage map for common carp and its integration with a BAC-based physical map. *PLoS One*. 2013;8(5):e63928.
24. Ohara E, Nishimura T, Nagakura Y, Sakamoto T, Mushiaki K, Okamoto N. Genetic linkage maps of two yellowtails (*Seriola quinqueradiata* and *Seriola lalandi*). *Aquaculture*. 2005;244:41–8.
25. Fuji K, Koyama T, Kai W, Kubota S, Yoshida K, Ozaki A, et al. Construction of a high-coverage bacterial artificial chromosome library and comprehensive genetic linkage map of yellowtail *Seriola quinqueradiata*. *BMC Research Notes*. 2014;7:200.
26. Geisler R, Rauch GJ, Baier H, van Bebber F, Bross L, Dekens MP, et al. A radiation hybrid map of the zebrafish genome. *Nature Genet*. 1999;23:86–9.
27. Hukriede NA, Joly L, Tsang M, Miles J, Tellis P, Epstein JA, et al. Radiation hybrid mapping of the zebrafish genome. *Proc Natl Acad Sci U S A*. 1999;96:9745–50.
28. Senger F, Priat C, Hitte C, Sarropoulou E, Franch R, Geisler R, et al. The first radiation hybrid map of a perch-like fish: the gilthead seabream (*Sparus aurata* L.). *Genomics*. 2006;87:793–800.
29. Su F, Osada Y, Ekker M, Chevrette M, Shimizu A, Asakawa S, et al. Radiation hybrid maps of medaka chromosome LG 12, 17, and 22. *DNA Res*. 2007;14:135–40.
30. Guyon R, Senger F, Rakotomanga M, Sadequi N, Volckaert FA, Hitte C, et al. A radiation hybrid map of the European sea bass (*Dicentrarchus labrax*) based on 1581 markers: Synteny analysis with model fish genomes. *Genomics*. 2010;96:228–38.
31. Guyon R, Rakotomanga M, Azzouzi N, Coutanceau JP, Bonillo C, D'Cotta H, et al. A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. *BMC Genomics*. 2012;13:222.
32. Aoki J, Kai W, Kawabata Y, Ozaki A, Yoshida K, Tsuzaki T, et al. Construction of a radiation hybrid panel and the first yellowtail (*Seriola quinqueradiata*) radiation hybrid map using a nanofluidic dynamic array. *BMC Genomics*. 2014;15:165.
33. Coppe A, Pujolar JM, Maes GE, Larsen PF, Hansen MH, Bernatchez L, et al. Sequencing, *de novo* annotation and analysis of the first *Anguilla anguilla* transcriptome: EeelBase opens new perspectives for the study of the critically endangered European eel. *BMC Genomics*. 2010;11:635.
34. Salem M, Rexroad III CE, Wang J, Thorgaard GH, Yao J. Characterization of the rainbow trout transcriptome using Sanger and 454-pyrosequencing approaches. *BMC Genomics*. 2010;11:564.
35. Pereiro P, Balseiro P, Romero A, Dios S, Forn-Cuni G, Fuste B, et al. High-throughput sequence analysis of turbot (*Scophthalmus maximus*) transcriptome using 454-pyrosequencing for the discovery of antiviral immune genes. *PLoS One*. 2012;7(5):e35369.
36. Spurgeon SL, Jones RC, Ramakrishnan R. High throughput gene expression measurement with real time PCR in a microfluidic dynamic arra. *PLoS One*. 2008;3(2):e1662.
37. De Givry S, Bouchez M, Chabrier P, Milan D, Schiex T. CARTHA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics (Oxford, England)*. 2005;21:1703–4.
38. Carthagene. Home Page: <http://www7.inra.fr/mia/T/CarthaGene/>.
39. Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, et al. The medaka draft genome and insights into vertebrate genome evolution. *Nature*. 2007;447:714–9.
40. Fujii K, Yoshida K, Hattori K, Ozaki A, Araki K, Okauchi M, et al. Identification of the sex-linked locus in yellowtail, *Seriola quinqueradiata*. *Aquaculture* 2010;308:S51–S55.
41. Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T, et al. DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature*. 2002;417:559–63.
42. Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, Sakamoto T, et al. A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. *Proc Natl Acad Sci U S A*. 2012;109:2955–9.
43. Kamiya T, Kai W, Tasumi S, Oka A, Matsunaga T, Mizuno N, et al. A trans-species missense SNP in *Amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (fugu). *PLoS genet*. 2012;8:7 e1002798.
44. Yano A, Guyomard R, Nicol B, Jouanno E, Quillet E, Klopp C, et al. An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. *Curr. Boil*. 2012;22:1423–1428.
45. Eveland AL, McCarty DR, Koch KE. Transcript profiling by 3'-untranslated region sequencing resolves expression of gene families. *Plant Physiol*. 2008;146:32–44.
46. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bernben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437:376–80.
47. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res*. 2008;36(10):3420.
48. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008;18:821—829.
49. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, et al. *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Res*. 2010;20:265–72.
50. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A*. 2011;108:1513–8.
51. Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, et al. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature*. 2000;407:513–6.
52. Brockman W, Alvarez P, Young S, Garber M, Giannoukos G, Lee WL, et al. Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome Res*. 2008;18:763–70.
53. Primer 3: <http://bioinfo.ut.ee/primer3-0.4.0/>.
54. Manly KF, Olson JM. Overview of QTL mapping software and introduction to Map manager QT. *Mamm Genome*. 1999;10:327–34.
55. Voorrips RE. Mapchart software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002;93:77–8.
56. Ensembl. <http://www.ensembl.org/index.html>.