An Expressed Sequence Tag (EST)-enriched genetic map of turbot (*Scophthalmus maximus*): a useful framework for comparative genomics across model and farmed teleosts

Carmen Bouza1, Miguel Hermida1, Belén G Pardo1, Manuel Vera1, Carlos Fernández1, Roberto de la Herrán2, Rafael Navajas-Pérez2, José Antonio Álvarez-Dios3, Antonio Gómez-Tato4 and Paulino Martínez1*

**Abstract**

**Background:** The turbot (*Scophthalmus maximus*) is a relevant species in European aquaculture. The small turbot genome provides a source for genomics strategies to use in order to understand the genetic basis of productive traits, particularly those related to sex, growth and pathogen resistance. Genetic maps represent essential genomic screening tools allowing to localize quantitative trait loci (QTL) and to identify candidate genes through comparative mapping. This information is the backbone to develop marker-assisted selection (MAS) programs in aquaculture. Expressed sequenced tag (EST) resources have largely increased in turbot, thus supplying numerous type I markers suitable for extending the previous linkage map, which was mostly based on anonymous loci. The aim of this study was to construct a higher-resolution turbot genetic map using EST-linked markers, which will turn out to be useful for comparative mapping studies.

**Results:** A consensus gene-enriched genetic map of the turbot was constructed using 463 SNP and microsatellite markers in nine reference families. This map contains 438 markers, 180 EST-linked, clustered at 24 linkage groups. Linkage and comparative genomics evidences suggested additional linkage group fusions toward the consolidation of turbot map according to karyotype information. The linkage map showed a total length of 1402.7 cM with low average intermarker distance (3.7 cM; ~2 Mb). A global 1.6:1 female-to-male recombination frequency (RF) ratio was observed, although largely variable among linkage groups and chromosome regions. Comparative sequence analysis revealed large macrosyntenic patterns against model teleost genomes, significant hits decreasing from stickleback (54%) to zebrafish (20%). Comparative mapping supported particular chromosome rearrangements within Acanthopterygii and aided to assign unallocated markers to specific turbot linkage groups.

**Conclusions:** The new gene-enriched high-resolution turbot map represents a useful genomic tool for QTL identification, positional cloning strategies, and future genome assembling. This map showed large synteny conservation against model teleost genomes. Comparative genomics and data mining from landmarks will provide straightforward access to candidate genes, which will be the basis for genetic breeding programs and evolutionary studies in this species.

**Keywords:** Turbot, Scophthalmus maximus, Pleuronectiformes, Genetic map, Recombination frequency, Comparative mapping
Background

The turbot (Scophthalmus maximus) is a flatfish of great commercial value, which represents one of the most promising marine species of European aquaculture. Production reached 9,142 t in 2009 [1], and it is predicted to double up in size in 2014. Turbot has also become very popular in the Chinese market, and production in this country has been reported around 50,000 t in 2006 [2]. Genetic breeding programs are being carried out by several turbot companies supported by microsatellite parentage tools [3]. Increasing growth rate, controlling sex ratio (females largely outgrow males) and enhancing disease resistance currently constitute the main goals of genetic breeding programs in this species.

The small turbot genome (C value: 0.86 pg; http://www.genomeseize.com/fish.htm) is organized in 2n = 44 chromosomes with no sex-associated chromosome heteromorphism [4, 5]. An important investment effort has been devoted in the recent years to increase genomic resources in this species to provide new molecular tools to support genetic breeding programs. An Expressed Sequence Tag (EST) database constructed using cDNA libraries from immune tissues [6, 7] has been recently enriched using new generation sequencing (NGS) technologies and currently contains 35,000 contigs and 65,000 singletons. This database was used to design the first turbot oligo-microarray [8], which enabled to identify differentially expressed (DE) genes for pathogen resistance [9, 10]. Co-localization of DE genes through comparative mapping with disease-resistance QTL constitutes a primary goal to identify candidate genes for resistance to pathogens [11].

EST databases are essential not only for functional annotation, but also for the identification of gene-associated markers (type I [6]). New microsatellites and single nucleotide polymorphisms (SNP) originating from the EST database have recently been developed in turbot [7, 12, 13]. These markers were used to identify candidate genes subjected to divergent selection [14], and to begin constructing an EST-linked genetic map in this species [12]. Finally, a 5X BAC genomic library containing ~46,000 clones of ~125 kb on average has been constructed and it is being exploited for physical mapping of specific genomic regions (B. Pardo, unpublished data).

Genetic maps are essential tools to locate genomic regions associated with productive characters, which can eventually be applied in marker-assisted selection programs or used to identify genes related to specific traits through fine mapping and/or positional cloning strategies [15-17]. Additionally, they provide the support to study genome organization and evolution through comparative mapping, and provide useful landmarks for genome assembly [18-24]. A first generation turbot consensus map (242 anonymous microsatellites; 26 linkage groups (LG)) was reported by Bouza et al. [25]. It has been used to identify QTL for sex determination [26], growth rate [27] and resistance to pathogens [28, 29]. Recently, a new microsatellite genetic map has been reported by Ruan et al. [2] using 158 anonymous markers.

Genomic resources have greatly increased in aquaculture species especially after the arrival of NGS, and several genome projects are underway in several fish species (http://www.genomesonline.org/cgi-bin/GOLD/index.cgi). However, most comparative genomic studies still rely on model species. Genome sequences with high coverage are available in zebrafish (Danio rerio), fugu (Fugu rubripes), Tetraodon (Tetraodon nigroviridis), medaka (Oryzias latipes) and stickleback (Gasterosteus aculeatus) (http://www.ensembl.org). Since gene-associated markers are much more conserved than anonymous ones, they constitute the preference target to go further on comparative mapping and evolutionary genomics [24, 30, 31]. Comparative mapping also represents the best strategy to capture candidate genes at genomic regions associated with productive characters in aquaculture species [32-35].

The aim of this study was to enrich the turbot genetic map using EST-linked markers to create a more powerful tool for comparative genomic and evolutionary studies in turbot. This second-generation genetic map will be useful for identifying candidate genes associated to productive traits and for marker-assisted selection in genetic breeding programs for turbot industry.

Results and discussion

Genetic markers and segregation analysis

The existence of a three-generation pedigree facilitated the consistent detection of null alleles. Among the 463 informative mapping markers (Additional file 1: Table S1), 20 loci (4.3%), 18 microsatellites (4.6%) and two SNP (2.7%) showed null alleles in any of the eight diploid families (DF and QF1-7), in accordance with previous data [3, 7]. Deviations from Mendelian expectations were detected at 27.5% loci (P < 0.05) mostly due to SNP (24.7% over 91 tests, P < 0.05) than to microsatellites (10.8% over 916 tests, P < 0.05) as previously reported in turbot [7, 25]. As suggested [36], the existence of paralogous genes due to the teleost gene duplication probably interferes with SNP genotyping, hence the higher proportion of Mendelian deviations observed. However, this fact did not determine a lower mapping success at deviated loci, showing a very similar proportion of framework markers in the turbot map as the non-deviated ones (74.8% vs 72.4%).

The turbot consensus map

The use of several mapping families has the advantage of increasing the number of informative meiosis, especially useful for low polymorphic markers, and also enables the comparison between genetic maps of...
different families or sexes. New genetic markers (mostly EST-linked), in addition to those previously reported [12, 25], were used to construct nine family maps to be integrated in a new consensus map. A large set of common informative markers were used to anchor the different family maps in order to integrate them into a single consensus map (Additional file 2: Table S2). This map consisted of 24 linkage groups named LG1 to LG24 (Figure 1). Markers in homologous linkage groups were compared among family maps (Additional file 3: Figure S1), full collinearity being observed at 13 linkage groups and very minor discrepancies at the 11 remaining ones, always involving closely linked markers (mostly < 3 cM).

The resulting consensus map (Figure 1) contained 438 out of 463 informative markers (94.4%), 180 EST-linked (41.1%) and 258 anonymous (58.9%) (Table 1). Among them, 336 were framework (72.4%), 63 mapped at LOD < 3 (13.6%), 39 accessory (8.4%) and 26 remained unlinked (5.6%). The 24 linkage groups of the consensus map represent a reduction from the previous 26 ones [25] in the way towards the expected 22 linkage groups according to turbot karyotype (n = 22; [4, 5]). Thus, groups LG4 and LG25 and groups LG10 and LG26, respectively, merged into single groups named LG4 and LG10 in the new consensus map (Figure 1). These two fusions had been suggested only based on paternal segregation data by Bouza et al. [25]. Additionally, some markers shared by different linkage groups suggested two additional fusions between LG8 and LG18 and between LG16 and LG19. If these fusions were confirmed, it would represent the final convergence to the expected 22 linkage groups.

The total map length (1402.7 cm) was very similar to that previously reported [25], but intermarker distance substantially decreased from 6.5 to 3.7 cm, thus the map being among the most dense maps within non-model teleosts [23, 31, 38-40]. Only four terminal regions involving non-framework markers at LG2, LG5, LG6 and LG12 showed distances higher than 20 cm, a threshold considered relevant for QTL identification [41]. The framework map covered 1193.4 cM, thus approaching the total length

![Figure 1 Consensus turbot map. Framework markers in bold characters; accessory markers indicated by parentheses beside the closest marker and listed at the end of Additional file 5: Figure S2; LOD < 3 markers in normal type.](image-url)
estimate. Considering the estimated genome size of the turbot between 600–800 Mb [5, 42], the present map would have on average a marker every ~2 Mb, thus representing a very useful tool for QTL identification and positional cloning strategies. Besides, this map will be valuable for physical mapping starting from the available BAC library and for genome assembling in future turbot genome projects.

| Table 1 Genetic markers and map and genome length in the turbot maps |
|------------------|-----------------|------------------|
|                  | Consensus | Paternal | Maternal |
| Total markers    | 463       | –        | –        |
| Mapped markers   | 438       | 221      | 241      |
| EST-linked       | 180       | 55       | 60       |
| Anonymous        | 258       | 165      | 181      |
| Framework markers| 336       | 215      | 225      |
| EST-linked       | 115       | 55       | 49       |
| Anonymous        | 211       | 159      | 176      |
| LOD < 3 markers  | 62        | 6        | 16       |
| Accessory        | 39        | –        | –        |
| Unlinked         | 26        | –        | –        |
| Total length     | 1402.7    | 854.2    | 1369.1   |
| Max. distance^b | 30.5      | 24.5     | 30.4     |
| Mean distance^c  | 3.7       | 4.4      | 6.3      |
| Framework length | 1193.4    | 781.7    | 1274     |
| Max. distance    | 28.1      | 23       | 30.4     |
| Mean distance    | 3.8       | 4.1      | 6.3      |
| Gen. length^a (total) | 1530.8 | 1063.4   | 1596.1   |
| Gen. length^a (framework) | 1375.8 | 979.2    | 1576.4   |

LG24 length was added to obtain total and estimated lengths of paternal and maternal maps for comparison with consensus map. *Genome length was estimated according to Hubert and Hedgecock [37]. Max. dist.: Maximum intermarker distance (cM) in each map. Average intermarker distance.

Recombination frequency (RF) between sexes and families

RF is a species-specific parameter, but also variable within species according to sex, family, chromosome, and genomic region [43]. These differences constitute an important factor to be considered when constructing genetic maps and when maps are applied for QTL identification and marker assisted selection (MAS) programs. RF differences between sexes have been described in most fish species when constructing genetic maps [20, 30, 31, 39, 44-48], including Pleuronectiformes [40]. Recombination differences between families have also been reported, especially in humans and in domestic species [49, 50], but few studies have been focused on this variation in fish and other aquaculture species [37, 45, 51, 52]. In turbot, we observed a 1.6:1 female: male (F:M) RF ratio from a limited sample of common female/male marker pairs in the previous turbot map [25], but no significant RF differences between the two female maps constructed. Ruan et al. [2] also reported a higher F:M ratio (1.3:1) in this species.

In the present study, the availability of a large number of homogeneously distributed common markers in nine mapping families (Additional file 2: Table S2) offered the opportunity for a detailed study on RF among families within sex, and between sexes. A global F:M ratio of 1.6:1 was observed (Figure 2A), thus corroborating our previous estimate [25]. The F:M ratio was largely variable among linkage groups (Additional file 4: Table S3; Additional file 5: Figure S2), ranging between 0.93 at LG15, the only linkage group with higher male RF, and 23.22 at LG21, where a suggestive sex-determining QTL was previously reported [26]. These results support our previous observation related to the differential crossing-over patterns among turbot chromosomes when estimating gene-centromere distances [53]. RF differences
among families within females (Figure 2B) were much lower than within males (Figure 2C). Accordingly, RF comparisons between females showed no significant differences, while they were significant between males at some cases. Inter-family RF differences have also been documented in other aquaculture species [37, 52].

**Comparative mapping**

Similarity of turbot sequences against stickleback (Gac), *Tetraodon* (Tni), medaka (Ola), fugu (Tru) and zebrafish (Dre) genomes revealed large macrosyntenic patterns (Figure 3; Additional file 6: Figure S3 and Additional file 7: Figure S4). Total significant hits decreased from stickleback (~50%) to *Tetraodon*, fugu, medaka (~40%) and zebrafish (~20%) genomes (Table 2), mostly in agreement with the closer phylogenetic relationship of turbot to stickleback, *Tetraodon*, medaka and fugu (Acanthopterygii) than to zebrafish (Ostariophysi) [54]. The stickleback (Gasterosteiformes) genome was the most informative one in our study, despite that medaka (Beloniformes) has been found to be more closely related to turbot (Pleuronectiformes) genome was the most informative one in our study, despite that medaka (Beloniformes) has been found to be more closely related to turbot (Pleuronectiformes) using mitogenome data [55]. This may reflect phylogenetic discordances between mitochondrial and nuclear genomes [56].

Gene-derived markers have demonstrated better performance than anonymous ones for comparative mapping [24, 30]. Accordingly, more EST-linked than anonymous turbot markers matched against model genomes (Table 2). Most unique hits were included in the turbot map, thus being relevant to identify syntenic regions. Matches showed high average identity (~90%), the length similarity and identity increasing from zebrafish to stickleback and being higher for EST-linked than for anonymous markers (Table 2), as reported in teleosts [38, 47].

**Macrosynteny between the turbot map and model teleost genomes**

Mapping of 180 gene-derived markers to the turbot map has substantially improved previous comparative analysis based on anonymous loci [25], allowing the assessment of large syntenies between the turbot and the model fish genomes (Figure 3; Additional file 6: Figure S3 and Additional file 7: Figure S4). As expected, conserved syntenies (multiple significant hits regardless of their order) were higher against Acanthopterygii (20 to 25 conserved syntenies with four or more hits) than against zebrafish (only 14 small syntenies; Additional file 7: Figure S4) genomes. A remarkable one to one correspondence between the turbot linkage groups and the Acanthopterygii chromosomes was observed (Figure 3), in agreement with previous comparative mapping among model teleosts [21, 57]. Synteny conservation was particularly extensive between the turbot and stickleback genomes (Table 2; Figure 3; Additional file 8: Table S4), aiding to establish a predicted location for most unlinked turbot markers from unique stickleback chromosomes. However, gene order appeared less conserved for most macrosyntenies (Additional file 8: Table S4 and Additional file 9: Table S5) reflecting linkage mapping limitations and/or chromosome rearrangements over evolutionary time [22, 30]. Collinearity appeared to be particularly conserved at macrosyntenic scale (Additional file 9: Table S5), as reported for other teleosts [30].

Comparative mapping provided additional support to the new LG4 and LG10, as well as to the fusion between LG8 and LG18, since they were syntenic to single chromosomes in all model Acanthopterygii (Figure 3; Additional file 6: Figure S3). By contrast, the independent syntenic relationship observed for LG16 and LG19 against model genomes (Figure 3) do not support the weak linkage signal observed between them. Further work will be required to establish the final merging on 22 linkage groups, both focusing on these linkage groups and on the smallest ones, particularly LG24. To achieve this goal, i) we are including new markers from Ruan et al. [2] in the turbot map; ii) we are performing two-color *in situ* hybridization with BAC probes associated to putative merging groups; and iii) we expect a draft of the turbot genome to be completed in the near future.

Comparative mapping also suggested fusion events in the stickleback (LG7 and LG16 merge into Gac4) and *Tetraodon* (LG5 and LG7 merge into Tni1) lineages as the most parsimonious hypothesis considering the ancestral n = 24 teleost karyotype [58] (Figure 3; Additional file 6: Figure S3). In accordance with the low rate of interchromosomal rearrangements in teleosts [57], only one turbot translocation between LG1 and LG22 was suggested from comparison with model species.

Overall incidence of multiple matches against the five model teleost genomes was low, although higher from EST-linked (~10–16%) than from anonymous (<4%) markers (Table 2). This is likely related to the higher retention of duplication events on coding sequences along vertebrate evolution [59], and particularly, to the fish-specific (3R) whole genome duplication validated by comparative studies [19, 21, 24]. Close to 40% of the duplicated hits detected across model genomes in this study (Additional file 10: Table S6) were congruent with the sets of orthologous and paralogous chromosomes identified between the *Tetraodon* and medaka genomes, which have been essential to reconstruct the vertebrate protokaryotype [57, 60]. This information could capacitate to predict positions for unallocated duplicated genes on the turbot map.

**Anchoring the turbot map onto model and farm teleost genomes**

Our study confirmed the findings of previous comparative mapping for farmed teleosts. Conserved synteny against
closely related model genomes has been shown, either within Acanthopterygii (*Tetraodon*, medaka, fugu or stickleback), such as in the halibut, tilapia, Japanese flounder, European seabream or seabass [32, 38, 40, 61], or within Ostariophysi (zebrafish), such as in the catfish or grass carp [23, 30]. Anchoring of several farm fish maps against model teleost genomes is highly relevant to boost in aquaculture technologies, providing straightforward access to the gene content within specific syntenic regions from model species. For instance, linking the advances in the genomic analysis of commercially important pleuronectiform and perciform species will be possible using the stickleback as common anchoring genome given its informativeness for comparative mapping in turbot, tilapia, European seabream and seabass [32]. Also, the conservation of microsyntenies in the turbot map will be valuable to search for candidate genes of productive traits around QTL by data mining on the model fish genomes [33, 34]. For this task, although the stickleback genome has demonstrated to be the most informative one, other model

Figure 3 Macrosynteny analysis between the turbot linkage map and model Acanthopterygii genomes. In gray background syntenies with two or more significant hits. UL: unlinked markers in the turbot map; UN: unrandom genomic regions of the model fish species.
species within Acanthopterygii will also provide essential information, particularly medaka, a closely related model species to Pleuronectiformes [55].

### Conclusions
A gene-enriched turbot consensus map has been constructed with a marker density in the range of those described in farm fish species with large genomic resources. The availability of multiple reference families enabled us to obtain detailed data on RF between and within sexes. The higher evolutionary conservation of EST-linked markers allowed the detection of large macrosynteny patterns with model fish species. This map provides essential information to identify genomic variation and candidate genes associated to productive traits for further application in MAS programs. The turbot map also provides useful landmarks for future turbot genome assembling and for evolutionary studies within pleuronectiforms and teleosts.

#### Methods

**Mapping families**

The haploid (HF) and diploid (DF) families from our previous studies [12, 25], and seven additional families used for QTL identification (QF1-7) [26-28] were used to construct the new turbot map. DF was the main reference because of its higher marker density. HF was maintained in our analysis because a large set of anonymous markers had only been mapped in this family [25], but no new markers were added to this family. QF families were used when markers were non-informative in DF family. The seven QF families had been used for QTL screening on sex determination, growth and resistance to pathogens and thus, they were anchored by a common set of markers [26]. QF families were obtained from the genetic breeding programs of the companies Stolt Sea Farm SA and Insuïna SA, where a three-generation pedigree was available for all of them. Grandparents, parents and around 100 offspring (between 91 and 113) were analyzed in each QF family.

#### Microsatellite and SNP markers

The following 463 markers (388 microsatellites and 75 SNP) were informative in the nine mapping families (HF, DF and QF1-7) (Additional file 1: Table S1): i) 261 mostly anonymous microsatellites obtained from partial genomic libraries (Sma-USC codes) or RAPD markers (TUR codes) including; 248 from the previous map [25], 7 from Pardo et al. [62], 3 RAPD-derived from Liu et al. [63], and 3 novel markers characterized in the present work; and ii) 202 EST-linked markers, including 127 microsatellites: 43 from Bouza et al. [12], 75 from Navajas-Pérez et al. [13] (SmaUSC-E and Sma-E codes, respectively), and 9 from Chen et al. [64]; SMAC codes; and 75 SNPs from Vera et al. [17]; SmaSNP codes. For simplicity, hereinafter we shall refer to those microsatellites derived from enriched-genomic libraries or RAPD as anonymous microsatellites (despite some of them being annotated), and to the other group as EST-derived markers. Microsatellite and SNP genotyping was carried out on an ABI 3730 DNA Sequencer. Primers and PCR conditions for three new microsatellites were described for the first time in this work (Sma-USC286, Sma-USC287 and Sma-USC288; Additional file 1: Table S1). Chi-square tests were applied to check for deviations from Mendelian expectations (1:1, 1:2:1 and 1:1:1:1) at each locus and within each family analyzed.

---

**Table 2** Similarity of turbot EST-linked and anonymous sequences against model teleost genomes

| Model species | Gac | Tni | Ola | Tru | Dre |
|---------------|-----|-----|-----|-----|-----|
| Total hits    | 246 (54.4) | 179 (39.4) | 186 (41.0) | 194 (42.7) | 91 (20.5) |
| Unique hits E < 10^{-15} (%) | 210 (48.0) | 151 (33.3) | 146 (32.9) | 162 (35.7) | 69 (15.5) |
| Multiple hits E < 10^{-15} (%) | 36 (7.9) | 28 (6.3) | 38 (8.4) | 32 (7.2) | 22 (5.0) |
| EST-linked sequences (E) | | | | | |
| Unique hits E < 10^{-15} (%) | 121 (60.8) | 98 (49.2) | 88 (44.2) | 93 (46.7) | 50 (25.1) |
| Unique, L-L (%) | 98 (49.2) | 60 (30.2) | 70 (35.2) | 66 (33.2) | 44 (22.1) |
| Multiple hits E < 10^{-15} (%) | 25 (12.6) | 21 (10.6) | 32 (16.1) | 25 (12.6) | 19 (9.6) |
| Anonymous sequences (A) | | | | | |
| Unique hits A (%) | 89 (36.3) | 53 (26.6) | 57 (22.4) | 69 (27.1) | 19 (7.8) |
| Unique, L-L (%) | 78 (31.8) | 40 (16.3) | 53 (21.6) | 66 (26.9) | 17 (6.9) |
| Multiple hits A (%) | 11 (4.3) | 7 (2.9) | 7 (2.7) | 7 (2.9) | 3 (1.2) |
| Unique hits summary | | | | | |
| Mean size alignment bp | 119.0 | 106.4 | 110.3 | 109.0 | 107.1 |
| Mean size A-E bp | 133.0 | 115.0 | 119.4 | 119.8 | 114.0 |
| Maximum size A-E bp | 257–504 | 252–401 | 215–422 | 222–438 | 181–295 |
| Mean E-value | 1.4E-07 | 4.2E-07 | 3.6E-07 | 4.0E-07 | 6.1E-07 |
| Minimum E-value | 1.0E-140 | 1.0E-116 | 1.0E-120 | 1.0E-119 | 2.0E-62 |
| Identity % | 89.8 | 89.9 | 89.0 | 89.5 | 88.2 |
| Identity Range | 79.4–100 | 80.4–100 | 79.9–96.7 | 79.4–100 | 80.3–96.2 |
| Retained at ≤ E 10^{-15} (%) | 182 (87%) | 122 (81%) | 110 (75%) | 129 (80%) | 50 (72%) |

BLASTn matches of 454 turbot microsatellite and SNP flanking sequences (199 EST-linked to 255 anonymous ones) against the stickleback (Gasterosteus aculeatus), spotted green pufferfish (Tetrodon nigroviridis), medaka (Ola: Oryzias latipes), fugu (Tru: Takifugu rubripes) and zebrafish (Dre: Danio rerio) genomes at E < 10^{-5} threshold, most of which retained at E ≤ 10^-15. *Unique or ≥ 2 significant matches, respectively, against model species genomes. “Unique, L-L”-linked markers at the turbot genetic map having unique significant matches with chromosome assignment on the model species-genomes. A-E: size alignment figures in base pairs (bp) for Anonymous (A) to EST-linked (E) markers.

---
Map construction

Linkage analysis in mapping populations

A consensus genetic map was constructed using the nine reference family maps. Also, female and male genetic maps were constructed averaging via female and via male segregation, respectively, with the same diploid reference families. HF family was only used to build the female map. The software JOINMAP 3.0 [65] was used for map construction starting from all haploid and diploid mapping populations (HF, DF and QF1-7). The genotypes of the haploid gynogenetic progeny were coded as JOINMAP type CP population and analyzed within aJOINMAP type CP population and analyzed within a reference family. Also, female and male genetic maps were also coded in HAP configuration with known linkage phase to construct female and male maps. Diploid family data (DF and Q1-7) were coded as JOINMAP type CP population and analyzed within a known-phase model. Clustering and order of markers, as well as integrated linkage analysis to construct consensus, female and male maps were carried out using JOINMAP 3.0 with a LOD threshold > 3.0 for framework mapping, as previously reported [25]. The graphic maps were generated using MAPCHART 2.2 [66].

Comparison of recombination frequency (RF) between sexes and families

Only RF between framework markers (LOD > 3.0) was considered for comparisons. Common marker pairs were identified at each linkage group in the different mapping families to compare RF between families within each sex (i.e. segregating in the male or in the female) and between sexes. Comparisons between families within sex were performed both for all family pairs and globally using information of all families. Comparison between sexes was performed by averaging RF of common marker pairs across families within each sex. The mean and standard error of RF differences (between-families within males, between-families within females and between sexes) were obtained. Comparison was performed for the whole genetic map, but also for each linkage group. For these analyses, a minimum of 10 common marker pairs between the evaluated families was considered. The significance of RF differences for each pair of families was estimated using t-tests. Normality of RF distributions was checked using Kolmogorov-Smirnov tests. Non-parametric Mann-Whitney rank-order test was applied to evaluate the significance of RF differences between sexes.

Comparative mapping

Given the high percentage of anonymous sequences used for linkage mapping in the turbot, NCBI-BLASTn was used to compare turbot containing-marker sequences against updated versions of model fish genomes downloaded from ftp://ftp.ensembl.org: Tetraodon nigroviridis v.8.61, Takifugu rubripes v.5, Danio rerio Zv9.6, Oryzias latipes v.1.61 and Gasterosteus aculeatus v.1.61. BLAST searching was performed by using a minimum alignment length of 40 bp with a score > 80 as recommended for EST mapping across species and two E-value thresholds (E ≤ 10⁻¹⁰ and E < 10⁻⁵) [26, 32, 67]. Turbot marker sequences were also used as queries for analysis against the stickleback cDNA database, as the most informative model genome in this study (see Results). We wrote a BioPerl BLAST parser to extract the desired hits including sequence similarity figures and genome location information for each model species.

Additional files

- Additional file 1: Table S1. Characteristics of the genetic markers included in the turbot map.
- Additional file 2: Table S2. Informative markers used to construct the turbot consensus map.
- Additional file 3: Figure S1. Correspondence between the nine family maps used to construct the turbot consensus map.
- Additional file 4: Table S3. Number of markers and map length for each linkage group (LG) of turbot.
- Additional file 5: Figure S2. Consensus, female and male genetic maps of turbot.
- Additional file 6: Figure S3. Comparative mapping between the turbot map and the five model teleost genomes.
- Additional file 7: Figure S4. Oxford grid showing syntenies between the turbot linkage map and the zebrafish genome.
- Additional file 8: Table S4. Putative syntenic markers (210) between the turbot genetic map and the stickleback genome.
- Additional file 9: Table S5. Putative syntenic markers (60) between the turbot genetic map and four model Acanthopterygii genomes.
- Additional file 10: Table S6. Conservation of multiple similarity hits between the turbot and four model Acanthopterygii genomes.

Competing interests

Some of the authors have received research funds in the past five years from private turbot companies related to genetic breeding programs. However, funding of the present manuscript has completely come from public institutions.

Authors’ contributions

BGP was in charge of libraries construction to develop EST linked markers. MV was responsible of SNP development and genotyping in the reference families, and RH and RN did the same task with EST linked microsatellites. JAAD and AGT constructed EST databases and developed bioinformatic tools for marker development and comparative mapping. CF contributed to the revision and management of sequence data for marker development. MH was responsible of genetic map construction using Joinmap. CB performed the comparative genomics analysis and together with PM wrote the manuscript. PM performed RF analyses and supervised and coordinated all the work. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by the projects: Consolider Ingenio Aquagenomics (CSD200700002), Spanish Ministerio de Ciencia e Innovación (AGL2009-13273), and Xunta de Galicia local Government (09MMA011261PR). We are indebted to Lucía Insua, María Portela, Susana Sánchez, María López, Mónica Otero and Sonia Gómez for technical assistance. B.G. Pardo was supported by an Isidro Parga Pondal research fellowship from Xunta de Galicia (Spain).
2. Ruan X, Wang W, Kong J, Yu F, Huang X: Genetic linkage mapping of turbot (Scophthalmus maximus L) using microsatellite markers and its application in QTL analysis. Aquaculture 2010, 308:89–100.

3. Castro J, Bouza C, Presa P, Pino-Queirido A, Raza A, Ferreiro I, Sánchez L, Martínez P: Potential sources of error in parentage assessment of turbot (Scophthalmus maximus) using microsatellite loci. Mar Biotechnol 2010, 12:19–35.

4. Bouza C, Sánchez L, Martínez P: Karyotypic characterization of turbot (Scophthalmus maximus) with conventional, fluorochrome, and restriction endonuclease banding techniques. Mar Biol 1994, 120:609–613.

5. Cuñado N, Terrones J, Sánchez L, Martínez P, Santos JL: Synaptonemal complex analysis in spermatocytes and oocytes of turbot, Scophthalmus maximus (Pisces, Scophthalmidae). Genome 2001, 44:1143–1147.

6. Pardo BG, Fernández C, Millán A, Bouza C, Vera M, Álvarez-Dios JA, Calaza M, Gómez-Tato A, Vázquez M, Cabaleiro S, Magaríños B, Lemos ML, Leiro JM, Martínez P: Designed expression sequence tags (ESTs) from immune tissues of turbot (Scophthalmus maximus) challenged with pathogens. BMC Vet Res 2008, 4:37.

7. Vera M, Álvarez-Dios JA, Millán A, Pardo BG, Bouza C, Hermeda M, Fernández C, de la Herrán R, Molina-Luzón MJ, Martínez P: Validation of single nucleotide polymorphism (SNP) markers from an immuneExpressed Sequence Tag (EST) turbot, Scophthalmus maximus, database. Aquaculture 2011, 313:31–41.

8. Millán A, Gómez-Tato A, Fernández C, Pardo BG, Álvarez-Dios JA, Calaza M, Bouza C, Vázquez M, Cabaleiro S, Martínez P: Design and performance of a turbot (Scophthalmus maximus) oligo-microarray based on ESTs from immune tissues. Mar Biotechnol 2010, 12:452–465.

9. Millán A, Gómez-Tato A, Pardo BG, Fernández C, Bouza C, Vera M, Álvarez-Dios JA, Cabaleiro S, Lamas J, Lemos ML, Martínez P: Gene expression profiles of the spleen, liver, and head kidney in turbot (Scophthalmus maximus) along the infection process with Aeromonas salmonicida using an immune-enriched oligo-microarray. Mar Biotechnol 2011, doi:10.1007/s10126-011-9374-7.

10. Pardo BG, Millán A, Gómez-Tato A, Fernández C, Bouza C, Álvarez-Dios A, Cabaleiro S, Lamas J, Leiro JM, Martínez P: Gene expression profiles of spleen, liver and head kidney in turbot (Scophthalmus maximus) along the infection process with Phililisterides dicentrarchi using an immune-enriched oligo-microarray. Mar Biotechnol 2012, doi:10.1007/s10126-012-9449-9.

11. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop L, Minelli A, Cardon LR, Abecasis GR: Tag (EST) database for microsatellite screening and validation. Mol Ecol Res 2012, doi:10.1111/j.1755-0998.2012.03216.x.

12. Vilas R, Bouza C, Vera M, Millán A, Martínez P: Variation in anonymous and EST-microsatellite genes across adaptive population divergence in turbot. Mar Ecol Prog Ser 2010, 420:231–239.

13. Danzmann RG, Gharbi K: Linkage mapping in aquaculture species. In Aquaculture Genome Technologies. Edited by Li ZJ. Oxford: Blackwell Publishing; 2007:139–167.

14. Canario A, Bargelloni L, Volckaert F, Houston RD, Massault C, Guignen Y: Genomics toolbox for farmed fish. Rev Fish Sci 2008, 16:1–13.

15. Moen T, Barsanki M, Sonesson AK, Kjellum S: Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar): population-level association between markers and trait. BMC Genomics 2009, 10:368.

16. Shimoda N, Knapik EW, Zimt J, Sim C, Yamada E, Kaplan S, Jackson D, de Sauvage F, Jacob H, Fishman MC: Zebralis genetic map with 2000 microsatellite markers. Genomics 1999, 58:219–232.

17. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, Aittamaa J, Zdobnov EM, De Geyter I, Delouche N, Billoud L, Duret L, Weissenbach J, Bouneau L, Leclerc M, Lescot L, Caboche M, Schutte V, Saurin W, Wincker P, Quétier F, Soubre O, Vincent A, Soubrier M: A consensus linkage map of Aequorea victoria reveals the early vertebrate proto-karyotype. Nature 2004, 431:946–957.

18. Kai W, Kikuchi K, Kato C, Takechi T, Kinouchi K, Shimizu J, Itoh H, Ootake M, Venkatesh B, Miyaki K, Suzuki Y: A genetic linkage map for the tiger pufferfish, Takifugu rubripes. Genetics 2005, 171:227–238.

19. Kai W, Kikuchi K, Tohari S, Cheak A, Tan A, Fujwa H, Asoya S, Sutoke H, Naruse K, Bremner S, Suzuki Y, Venkatesh B: Integration of the genetic map and genome assembly of fugu facilitates insights into distinct features of genome evolution in Tetuloids and mammals. Genome Biol Evol 2011, 3:442–449.

20. Gross JB, Protas M, Conrad M, Schied PE, Vidal O, Jeffery WR, Borowsky R, Tabin CJ: Synteny and candidate gene prediction using an anchored linkage map of Astyanax mexicanus. Proc Natl Acad Sci U S A 2008, 105:20166–20171.

21. Xia JH, Liu F, Zhu ZY, Fu JJ, Jing YB, Yue GH: A consensus linkage map of the grass carp (Ctenopharyngodon idella) based on microsatellites and SNPs. BMC Genomics 2010, 11:135.

22. Sarropoulou E, Nousdili D, Magoulas A, Kotoulas G, Vougiouklakis T, Vozikis E, Loutsis D, Mileitis J, Toumbas A, Verikouropoulos E: The genetic map of Catarrhineus weissiennbs:54 Page 9 of 10

23. Canario A, Bargelloni L, Volckaert F, Houston RD, Massault C, Guignen Y: Genomics toolbox for farmed fish. Rev Fish Sci 2008, 16:1–13.
35. Greenwood AK, Jones FC, Chan YF, Brady SD, Asher DM, Grimwood J, Schmutz J, Myers RM, Kingsley DM, Peichel CL: The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. Heredity 2011, 107:155–166.

36. Dakin EE, Avise JC: Microsatellite null alleles in parentage analysis. Heredity 2004, 93:504–509.

37. Hubert S, Hedegerock D: Linkage maps of microsatellite DNA markers for the pacific oyster Crassostrea gigas. Genetics 2004, 168:351–362.

38. Chistiakov DA, Tsigopoulos CS, Lagnel J, Guo YM, Hellems B, Hales CS, Volckaer FA, Kotoulas G: A combined AFLP and microsatellite linkage map and pilot comparative genomic analysis of European sea bass Dicentrarchus labrax L. Ann Genet 2006, 39:623–634.

39. Ruxrod CE, Pals Y, Garth SA, Vallejo RL: A second generation genetic map of rainbow trout (Oncorhyncus mykiss). BMC Genet 2008, 9:74.

40. Castaño-Sánchez C, Fuji K, Ozaki A, Hasegawa O, Sakamoto T, Morishima K, Bouza CE, Goutti A, Danzmann RG, Gharbi S, Sakamoto T, Høyheim B, Franch R, Louro B, Tsalavouta M, Chatziplis D, Tsigenopoulos CS, Stermer LE, Guyomard R: A second generation genetic linkage map based on microsatellite DNA markers. A genetic linkage map for Arctic char (Salvelinus alpinus): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 2004, 47:304–315.

41. Lee BY, Lee WJ, Streelman JT, Carleton KL, Howe AE, Liao XL, Meng L, Guo BL, Riquet J, Huang L, Milan Bouza C, Martínez P: Independent evolution of the Japanese flounder (Paralichthys olivaceus). BMC Genomics 2010, 11:554.

42. Delkens JCA, Hospita F: The use of molecular genetics in the improvement of agricultural populations. Nat Rev Genet 2002, 3:22–32.

43. Hardie DC, Hubert PN: Genome-size evolution in fishes. In Fish Genomes. Oxford: BIOS Scientific Publishers Ltd; 1999.

44. Moen T, Hayes B, Baranski M, Berg PH, Kjelgum S, Koop BF, Davidson WS, Omholt SW, Lien S: A linkage map of the Atlantic salmon (Salmo salar) based on EST-derived SNP marker. BMC Genomics 2008, 9:223.

45. Varanam AA, McGowan C, Stout JA, Gharib K, Ferguson MM, Høyheim B, Davidson EA, Davidson WS, Ruxrod C, Danzmann RG: A genetic linkage map for Arctic char (Salvelinus alpinus): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 2004, 47:304–315.

46. Lee BY, Lee WJ, Streelman JT, Carleton KL, Howe AE, Hulata G, Hasegawa O, Sakamoto T, Young G, Hui F, Titmarsh P, Matsuda K, Hara M, Seeb KE, Streelman JT, Howe AE, Asher DM, Grimwood J, Myers RM, Kingsley DM, Peichel CL: The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. Heredity 2011, 107:155–166.

47. Dakin EE, Avise JC: Microsatellite null alleles in parentage analysis. Heredity 2004, 93:504–509.

48. Ruxrod CE, Pals Y, Garth SA, Vallejo RL: A second generation genetic map of rainbow trout (Oncorhyncus mykiss). BMC Genet 2008, 9:74.

49. Castaño-Sánchez C, Fuji K, Ozaki A, Hasegawa O, Sakamoto T, Morishima K, Bouza CE, Goutti A, Danzmann RG, Gharbi S, Sakamoto T, Høyheim B, Franch R, Louro B, Tsalavouta M, Chatziplis D, Tsigenopoulos CS, Stermer LE, Guyomard R: A second generation genetic linkage map based on microsatellite DNA markers. A genetic linkage map for Arctic char (Salvelinus alpinus): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 2004, 47:304–315.

50. Lee BY, Lee WJ, Streelman JT, Carleton KL, Howe AE, Liao XL, Meng L, Guo BL, Riquet J, Huang L, Milan Bouza C, Martínez P: Independent evolution of the Japanese flounder (Paralichthys olivaceus). BMC Genomics 2010, 11:554.

51. Delkens JCA, Hospita F: The use of molecular genetics in the improvement of agricultural populations. Nat Rev Genet 2002, 3:22–32.

52. Hardie DC, Hubert PN: Genome-size evolution in fishes. In Fish Genomes. Oxford: BIOS Scientific Publishers Ltd; 1999.

53. Moen T, Hayes B, Baranski M, Berg PH, Kjelgum S, Koop BF, Davidson WS, Omholt SW, Lien S: A linkage map of the Atlantic salmon (Salmo salar) based on EST-derived SNP marker. BMC Genomics 2008, 9:223.

54. Varanam AA, McGowan C, Stout JA, Gharib K, Ferguson MM, Høyheim B, Davidson EA, Davidson WS, Ruxrod C, Danzmann RG: A genetic linkage map for Arctic char (Salvelinus alpinus): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 2004, 47:304–315.

55. Lee BY, Lee WJ, Streelman JT, Carleton KL, Howe AE, Hulata G, Hasegawa O, Sakamoto T, Young G, Hui F, Titmarsh P, Matsuda K, Hara M, Seeb KE, Streelman JT, Howe AE, Asher DM, Grimwood J, Myers RM, Kingsley DM, Peichel CL: The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. Heredity 2011, 107:155–166.

56. Dakin EE, Avise JC: Microsatellite null alleles in parentage analysis. Heredity 2004, 93:504–509.

57. Ruxrod CE, Pals Y, Garth SA, Vallejo RL: A second generation genetic map of rainbow trout (Oncorhyncus mykiss). BMC Genet 2008, 9:74.

58. Castaño-Sánchez C, Fuji K, Ozaki A, Hasegawa O, Sakamoto T, Morishima K, Bouza CE, Goutti A, Danzmann RG, Gharbi S, Sakamoto T, Høyheim B, Franch R, Louro B, Tsalavouta M, Chatziplis D, Tsigenopoulos CS, Stermer LE, Guyomard R: A second generation genetic linkage map based on microsatellite DNA markers. A genetic linkage map for Arctic char (Salvelinus alpinus): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 2004, 47:304–315.

59. Lee BY, Lee WJ, Streelman JT, Carleton KL, Howe AE, Liao XL, Meng L, Guo BL, Riquet J, Huang L, Milan Bouza C, Martínez P: Independent evolution of the Japanese flounder (Paralichthys olivaceus). BMC Genomics 2010, 11:554.

60. Delkens JCA, Hospita F: The use of molecular genetics in the improvement of agricultural populations. Nat Rev Genet 2002, 3:22–32.

61. Hardie DC, Hubert PN: Genome-size evolution in fishes. In Fish Genomes. Oxford: BIOS Scientific Publishers Ltd; 1999.

62. Moen T, Hayes B, Baranski M, Berg PH, Kjelgum S, Koop BF, Davidson WS, Omholt SW, Lien S: A linkage map of the Atlantic salmon (Salmo salar) based on EST-derived SNP marker. BMC Genomics 2008, 9:223.