

QUALITY OF LIFE AND MANAGEMENT OF LIVING
RESOURCES PROGRAMME (1998-2002)

**Genetic Catalogue, Biological Reference Collections and
Online Database of European Marine Fishes**

FINAL REPORT

ANNEXES

Contract number: **QLRI-CT-2002-02755**

Project acronym: **FishTrace**

QoL action line: **Area 14.1, Infrastructures**

Reporting period: **01/01/03 - 30/06/06**

Annexes

Annex I: 1st FishTrace Plenary Meeting. Madrid, 12-14th February, 2003.

Annex II: FishTrace Database Structuration Meeting. Paris, 13-14th May, 2003.

Annex III: FishTrace Standardization Meeting. Stockholm, 4-7th June, 2003.

Annex IV: 2nd FishTrace Plenary Meeting. Las Palmas de Gran Canaria (Spain), 19-21st November, 2003.

Annex V: FishTrace Meeting on Data Validation. Ispra, Italy. 15-16th April, 2004.

Annex VI: 3rd FishTrace Plenary Meeting. Paris, 25-27th November, 2004.

Annex VII: FishTrace Database & Web Interface Meeting. Madrid, 11th March, 2005.

Annex VIII: 4th FishTrace Plenary Meeting. Funchal (Portugal), 24-26th October, 2005.

Annex IX: FishTrace Concluding Meeting. Kavala (Greece), May 10-12th, 2006.

Annex X: Sampling and Taxonomy (WP2) Protocol.

Annex XI: Results from Molecular Genetic Procedures Standardisation.

Annex XII: Molecular identification and DNA barcoding (WP3). Protocol and PCR conditions.

Annex XIII: Phylogenetic Validation of Sequences. Guidelines.

Annex XIV: Preparing sequence files for the Sequin tool.

Annex XV: Reference Collections (WP5) Protocol, Loan Request Form and Invoice of Specimens.

Annex XVI: Guidelines for Validation purposes (WP7) including a protocol defining format for the Validation tasks.

Annex XVII: Concise Protocol for the online validation process.

Annex XVIII: The Bibliography Module in the FishTrace Database.

Annex XIX: Results from the genetic population structure analysis.

Annex XX: Results from haplotype connectivity network analysis.

Annex XXI: Final Report on Validation and Missing Data.

Annex XXII: Control Data Validation Flow Document.

Annex XXIII: Leaflet presenting the FishTrace project.

Annex I: 1st FishTrace Plenary Meeting. Madrid, 12-14 February, 2003.

1st FISHTRACE MEETING

12-14 February 2003

Madrid, Spain

Held at NH Prisma Hotel, and Department of Biochemistry and Molecular Biology IV, UCM

Meeting Participants

Participant 1 (UCM): José M. Bautista, Amalia Diez, Gema González, Susana Schönhuth, Rafael González, Faye Taylor; **Participant 2 (EC-JRC-IPSC):** Philippe Carreau; **Participant 3 (NRM):** Sven Kullander, Michael Noren; **Participant 4 (ICCM):** Jose Antonio González, Ignacio Lozano, Fernanda Marreo, Montserrat Gimeno; **Participant 5 (IFREMER):** Monique Etienne, Veronique Verrez-Bagnis, Marc Jerome; **Participant 6 (RIVO):** Hilde van Pelt; **Participant 7 (IMAR):** Mafalda Freitas, Rosa Pestana; **Participant 9 (NAGREF):** Grigorius Krey; Alexis Tsangridis; **Participant 10 (MNHN):** Guy Duhamel, Patrice Pruvost.

Meeting Agenda:

1st DAY, February 12 – Wednesday

1. Introduction to FishTrace
2. Presentation of participant teams
3. Workpackages structure and information flow.
4. Schedules: General, Workpackages, Participant teams, Milestones and deliverables
5. Visit to the molecular biology laboratories at UCM and demonstrations on sampling for taxonomy, genetics and museums.

2nd DAY, February 13 - Thursday

6. WP1: Standardization of Methods and Legal Aspects
7. WP2: Sampling and Taxonomy.
8. WP3: Molecular Genetic Identification.
9. WP4: Biogeographical Molecular Haplotyping.
10. WP5: Biological Reference Collections.
11. WP6: Database Development.
12. WP7: Validation in Database.
13. WP 8: Implementation of WWW Technologies.

3rd DAY, February 14 - Friday

14. WP 9: Dissemination and Applicability.
15. Organization of practical work and information flow (Taxonomic, genetic and Collections Data Blocks: Research teams involved and flow)
16. Information flow at each geographical location and to the database
17. Implementation of the overviewing performance and periodical progress reports
- 18.. Financial and administrative matters.
19. Next Meetings.

MEETING MINUTES

1.- Introduction to FishTrace:

The meeting started at 09:30 hours at the NH Prisma Hotel.

The coordinator for the project, Dr José Bautista, welcomed the participants and presented the origins and general structure of the project, as well as the European dimension in terms of coverage of the geographic areas.

2.- Presentation of participant teams:

10:00 - 11:15: All participants did a presentation of the scientific groups participating in the project. They included details on their institutions; experience and other projects in which they participate; personnel assigned to this project, their performance and responsibilities, including the geographical area of their analysis (sampling, genetics and/or collections), and a briefing of the Workpackages in which the institution participates. Following are the names of those who presented each institution:

Participant 1 (UCM): Amalia Diez

Participant 2 (EC-JRC-IPSC): Philippe Carreau

Participant 3 (NRM): Sven Kullander

Participant 4 (ICCM): Jose Antonio González

Participant 5 (IFREMER): Monique Etienne

Participant 6 (RIVO): Hilde V Pelt

Participant 7 (IMAR): Mafalda Freitas

Participant 9 (NAGREF): Grigorius Krey

Participant 10 (MNHN): Guy Duhamel

Ignacio Lozano did the presentation on behalf of Participant 8 (TFMC) and informed the participants that Dr. Sebastian Jimenez, the scientist in charge of this group, could not attend the meeting.

3. Workpackages structure and information flows.

At 11:45: Dr Bautista presented the general objectives of the project drawn over each Workpackage structure and how the success on the project depends on the maintenance of the information flow among Workpackages.

At the general discussion on this topic, it was pointed out the urgent need for a quick standardization of all methods to start the tasks for Workpackages 2 and 3, and a detailed definition of the database fields in order to make them compatible with other cross-linked databases of interest for FishTrace.

Keeping the right of use on photographs and sequences for specific purposes (books, publications, tools for genetic identification) was also discussed with the understanding that the project is founded to make it available to the public. In this respect it was also observed the need to start Workpackage 9, dissemination, as soon as possible with the coordination activity of IFREMER to detect possible interested parties in our complete set of information (sequences, photographs, etc).

4. Schedules: General, Workpackages, Participant teams, Milestones and deliverables

At 12:30 : there was a definition of the project schedules indicating the timetables to be covered. The context of the project planning including each Workpackage, the contribution of each participant institution, the milestones to be reached and the deliverables during the time course of the project.

5. Visit to molecular biology laboratories at UCM and demonstrations on sampling for taxonomy, genetics and museums.

At 15:30: there was a visit to the coordinating group's laboratory at the Universidad Complutense de Madrid. There was a demonstration on taxonomy sampling, for genetics and museums. Tentative protocols for sampling and molecular genetics were given to each participant.

2nd Day

6. WP1: Standardization of Methods and Legal Aspects

At 09:30 Dr Van Pelt presented the tasks covered by Workpackage 1 (standardization) for which RIVO is the leader. In the presentation a list of subjects to standardize was shown, making special emphasis in those aspects of sampling, DNA, PCR and sequencing, data compilation, images and collection. Particular attention was given to legal aspects on public access to the database; patents and exchange flow of biological samples. The list of standardized methods of FishTrace will be available through the FishTrace web site, so all groups are able to provide identical data quality.

A "legal committee" formed by Hilde van Pelt (RIVO), Guy Duhamel (MNHN), Monique Etienne (IFREMER) and Philippe Carreau (EJC) will help and research to advise on legal questions related to data releasing from FishTrace.

During the discussion it was agreed that the protocols for standardization must circulate from the Workpackage leader of each standardized tasks to the other contributors in the same Workpackage. Finally, RIVO will act as a referee for all method standardization and produce the definitive standardized protocol.

7. WP2: Sampling and Taxonomy.

At 10:15 Dr Kullander presented the tasks covered by Workpackage 2 (Sampling + Taxonomy) for which NRM is the Workpackage leader.

Within the objectives of this Workpackage it was described the sampling procedure and the corresponding data and material to be obtained from each sample. Particular emphasis was given to the systematic work, and the relationship with possible different haplotypes detected in Workpackage 4 from the different geographical locations.

The need for a standardization and agreement on the systematic by all groups contributing was clear during the discussion, in particular related to the fields in the database. Issues related to regional data compilation and literature was also raised. This is expected to maintain a certain degree of originality in terms of discovering the most local particularities and not to download from other databases.

It was also mentioned that fish photos must retain the copyright of the photographer but they will be freely released among the participants of the project.

The need for agreements on what to do with excess of samples and the interchange of them among institutions was also foreseen at the discussion.

8. WP3: Molecular Genetic Identification.

At 11:20 Dr. Bautista presented the tasks covered by Workpackage 3 (Molecular Genetics Identification) for which UCM is the Workpackage Leader.

Within the scope of the essential information of this Workpackage it was defined the requirements for a sound molecular genetic identification of the samples received by the groups contributing to this Workpackage. Topics covered included the methodology for DNA, extraction, PCR and sequencing. Tentative protocols will be released to the groups participating and a review based on previous experience from all groups will help to define a definitive protocol.

Emphasis was given to the quality control of the sequences obtained through a dichotomy scheme that has been working in previous regional projects.

Exceptions to the general protocol must be considered since it was pointed out that experimental difficulties could arise in some samples (10-15% of the samples) due to DNA quality, mislabelling, misidentified, technical sequencing error, etc...

A brief discussion helped to explain the use of sequencing both genes (Rhodopsin and cytochrome b) for a precise genetic characterization, and how sequences from both genes act as an internal quality control for taxonomic and genetic procedures.

9. WP4: Biogeographical Molecular Haplotyping.

At 12:05 Dr. Krey presented the tasks covered by Workpackage 4 (Biogeographical Molecular Haplotyping) for which NAGREF is the Workpackage leader. Within the pursued objectives it was pointed out the need to establish if cyt b and rhodopsin would be efficient markers for polymorphisms studies. Emphasis was given to intra-species polymorphism detection as a result of the mutation rate, the sample size and the biology of the species examined.

An estimation of the number of the species represented in one or more than one geographical location, suggested trying to balance the number of species in which it would be possible to do haplotyping at the cyt b gene using a high number of specimens at each geographical location. Also the life history of each species would be considered, in terms of genetics variability, established populations (pelagic/migratory verses demersal/benthic). It was also mentioned the methodology to be employed and the data analysis for individual sequence and consensus as well as for haplotype diversity annotation.

10. WP5: Biological Reference Collections.

At 15:30 Dr. Freitas presented the methods and procedures included in WP 5 (Biological reference collections) for which IMAR is the Workpackage leader.

It was examined within the objectives the way to coordinate the collections of biological material obtained in the project, at each institution (because of the heterogeneity interests and possibilities) and among participating institutions, including exchange, loan and access to the biological collections.

Dr Freitas also anticipated some difficulties regarding critical revision of species lists, size and intra-specific morphological variations.

An issue was raised during the discussion on what the most adequate number of vouchers would be for systematic, taken into account the high number of species to be collected and sampled. It was also mentioned the great scientific opportunity with FishTrace to afford a wide characterization of species from taxonomy to genetics compiled with the storage of reference biological samples, all compiled in a new database. For this reason it was found convenient to try to get a significant number of specimens for future analysis.

11. WP6: Database Development.

16:30: Philippe Carreau presented the JRC group since it was not done at the day of the presentations. After that Mr Carreau presented Workpackage 6 (Database development) for which JRC is the Workpackage leader. Architecture of the FishTrace database will be based on the present Fishgen database. The database will be extended for many new species from FishTrace and will contain a new data structure with new tables related to the new data provided and new user specifications. Improvements will be done on:

- a) Taxonomy database structure through the participants input,
- b) Data quality by describing more precise sample acquisitions process, and analysis procedure.

- c) Quality field will be included regarding method reliability, applicability.
- d) Methods to transform genetic information in computer understandable language

Database will be implemented with data from samples and common fields to groups of samples that will be defined as such.

Computing FishTrace to Fish Base dataset will help to join genetic information to fish morphology and more generalist data.

In addition GIS (geographical information system) capabilities will be added to the system.

The discussion raised items on the needs to standardize the methods as soon as possible since it implies fields definition for the database.

During the discussion, the need for quality data and for the previous expected definition on the data was raised.

It was also called the requirements to agree on releasing the data online according to the database, as well as the need to interchange and make compatible FishTrace to other databases.

12. WP7: Validation in Database.

At 17:45 Dr Gonzalez presented Workpackage 7 (Validation in database) for which ICCM is the Workpackage leader. Three types of data sets collected for the offline database must be validated previously for release on the online database.

For each dataset a different expert will be responsible for data validation. (Data on taxonomy, genetics and collections)

Four steps were suggested for validation: data checking, data compilation, database field validation and data collecting. Deadlines for validation for each data set was also suggested.

In addition the “what” and “how” to validate was given for the taxonomy dataset, and was left undefined for the genetics and collections since is expected that it will be established by the groups responsible in agreement with the other participants involved in these data blocks.

13. WP 8: Implementation of WWW Technologies.

At 16:30, Mr. Carreau presented Workpackage 8 (Implementation of WWW Technologies) for which JRC is the Workpackage leader. The purpose of this WP is to implement the Website/database interconnection.

A web site using technologies based on existing systems to extend web capabilities and database connection will be created for this purpose.

It was mentioned that FihTrace will need to implement the user feedback test methodology. During it's development the interface will be submitted to the users in order to get their input.

Two different interfaces will be defined, one for internal FishTrace use and the other for www use. For the internal use the database will have the capability of giving the number of different entries submitted, validated or already online.

Sequence data will need to be defined in the most standard format (fasta) to provide easy links with Gen-Bank or EMBL databases. In addition taxonomy will be linked with the FishBase web page.

A new server for the project will be provided to quickly inform all participants of news and events on the project with selected channels (or labels) for taxonomy, genetics, collections, database and dissemination.

The domains “Fishtrace.com” “Fishtrace.net” and “Fishtrace.org” have already been obtained by the consortium and a server has been rented so we can soon create the Fishtrace web page.

3rd Day

14. WP 9: Dissemination and Applicability.

At 9:30, Dr. Verrez-Bagnis presented Workpackage 9 (Dissemination and Applicability.) for which IFREMER is the Workpackage leader.

Expected users who could be interested in the FishTrace database are scientists, technical centres, seafood processors, seafood wholesalers and retailers, consumer associations, administration and regulatory bodies, and control laboratories.

Various means of disseminations are foreseen, as articles in scientific, technical and professional journals, brochures to promote the network, booklets presenting a detailed description of the database and the reference collection and genetic catalogue, communications in seminars and specialized symposia (WEFTA: West European Fish technologist association, FASIS: Food Authenticity and Safety International Symposium, etc..., communications at exhibitions like Seafood (Bruxelles), SIAL (Paris), Fisk Seafood International Europe (Bremen)...etc.

The importance of links from web sites of each institution to the FishTrace web site, as well as the links with other databases as FishBase or GenBank was mentioned.

It was also pointed out that a FishTrace logo could be designed as a kind of brand for our project.

The agreed lists of species for the project could be released on the web so to help promote and disseminate the work expectations.

It was also discussed the possibility of designing a brochure to send to a mailing list compiled by each participant. Two formats were suggested, either electronic (in the web site) or paper (to send by post).

15. Organization of practical work and information flow (Taxonomic, genetic and Collections Data Blocks: Research teams involved and flow).

A brain storming session about the organization of the practical work allowed establishing the following general agreements to follow by the FishTrace participants regarding the main data and sampling collection work:

Sampling and taxonomy

The number of specimens to be taken would be five. A muscle tissue sample from all of them must be extracted and all of the individuals will be kept at the museum, except some of those who have storage problems (v.g. IMAR) in which case they will only keep three samples and send the other two to the MNHN and NRM groups who have kindly offered their institutions to keep some vouchers collections (all or part) and any other collection from the project for those groups that are not able to store and maintain them. This will guarantee that all reference collections from FishTrace will be long-term institutionally maintained.

Regarding the morphological characters, the leader for Workpackage 2: sampling and taxonomy (NRM) agreed to establish a protocol to determine those morphological characters that are relevant for taxonomy. NRM will send this protocol to all the participants involved in that task.

The species with big specimens will be taken for sampling and taxonomy but there won't be vouchers at the museums because of the difficulty in storing them. The same rule will apply for the expensive species.

For those species with intraspecific morphological variation, the number of specimens to be taken will be increased, and they will go under high priority to the genetic analysis in order to get a correct validation identifying the specie and the specimen. In case the validation is not successful, more specimens will be taken.

The bibliography will be preferably established at a regional level to avoid including the most generic data already included in other databases like FishBase.

For the distribution of the species, each participant daily sampling and taxonomy will detail the area in which this specie was captured.

In those genus included in the “list of species” with no specification of the specie (marked with *spp* like *Ephinefelus spp*), the sampling at each geographical area will be done in all the species found in this area.

Genetic identification:

All partners involved in this Workpackage agreed to standardize the protocol for DNA extraction. For this task, UCM (leader for WP3) will send muscle tissue samples from several species. Each laboratory will extract that DNA using its own methodology and send it back to UCM to amplify their *cyt b* and rhodopsin genes. All the work will be finished in 6 weeks, after which the results will be sent to Dr. Hilde V. Pelt who will define the protocol for this part.

Regarding the Taq Polymerase, the agreement was to try other enzymes, like high Fidelity-Taq polymerases, and compare them with the ones we are using now.

Haplotyping:

It was suggested to obtain at least 20 specimens of chosen demersal species that are in more than 4 sampling areas. Those chosen species are liable to have more variability than others and would be subjected to specific haplotyping at the *cyt b* only. A tentative list of species for this haplotyping will circulate from the partner responsible for Workpackage 4 (NAGREF) to the groups in charge of Workpackage 3 and from them, once agreed, the list will be send to the groups responsible for sampling and taxonomy (Workpackage 2).

Reference Collections:

IMAR as Workpackage leader for the Biological reference collections will establish a general protocol to organize all collections (vouchers, tissue, otholiths and DNA) based on previously established collections. This protocol will circulate to establish an agreed protocol and they will be sent from the partner responsible for Workpackage 1, standardization (RIVO), to the partners participating in the collections.

Database structure:

A new group was created for defining the data-structure for the database, so it can be defined soon to help Workpackage 6 (database development) to start in time. The MNHN, IMAR, NRM, JRC, NAGREF and ICCM teams form this group. It is expected that the first list of fields for the database could start circulating within four weeks and then a first tentative definitive list of fields two weeks after that.

16. Information flow at each geographical location and to the database

Targeted fish species at regional level involves three groups (taxonomic, museum and genetic). Each 5-specimen species set generate the need for the control of the information flow at regional level.

The objective would be to establish a feedback information control system of identity and location of biological samples, reference collection samples and also early control of the taxonomical identity of samples.

For the control of identity and location of biological samples it was suggested that the taxonomy group creates an Excel file to be sent to the other groups involved on a monthly bases. This will gather updated information on the identity, transference and location of every sample.

For the identity and location of reference collection samples the taxonomic group should create another excel file to be sent to the museum to be filled and returned by

them on a monthly basis. This will help to collect updated information on the identity, transference and location of every sample.

The taxonomy group should create a third Excel file to be sent to the genetics group to be filled out and returned by them on a monthly basis for the early control of the taxonomical identity of samples. This will gather updated information on genetic and taxonomical identity of each specific sample. In addition, it was suggested that the genetics groups must care on detecting possible errors related to the taxonomy identity for both tissue samples from each putative fish species and inform the taxonomy group of any further analysis.

The Excel files mentioned above will be provided by ICCM to all groups in charge of sampling, genetics and collections.

17. Implementation of the overviewing performance and periodical progress report.

It was foreseen the need for an implementation on the overviewing performance through progress indicators.

Within each regional research task, different research groups are responsible for four different sets of data:

- a) Taxonomy + Sampling
- b) Genetics
- d) Haplotyping
- c) Reference Collections

Periodical progress indicators submitted in a simple form will help to overview the geographical task performance and quantify the degree of fulfilling the objectives at any time allowing to detect any possible deviations from the expected work-plans.

At the end of every 4-month period, it was suggested that each group's leader should submit the periodical progress form to the project's co-ordinator so to help the preparing the reports. Below are the progress indicators suggested that will be provided as a form by UCM:

Progress indicators suggested for a Taxonomic group:

- a. Species sampled and taxonomically identified
- b. Vouchers transferred to its curator group
- c. Muscle tissues transferred to its curator group
- d. Pairs of otoliths transferred to its curator group
- e. Muscle tissues transferred to its genetic group
- f. Specific information sheets filled-in (database)
- g. Images obtained from fish
- h. Images obtained from otoliths
- i. Images obtained from other specific fish-products
- j. Meetings/workshops attended
- k. Dissemination activities done (specify)
- l. Periodical progress reports sent

Progress indicators suggested for a Genetic group:

- a. Muscle tissues received
- b. Total DNA extractions
- c. Successful PCRs for Cyt B fragment A
- e. Successful PCRs for Cyt B fragment B
- f. Successful PCRs for Rhodopsin
- g. Sequenced PCR products of Cyt B fragment A
- h. Sequenced PCR products of Cyt B fragment B
- i. Sequenced PCR products of Rhodopsin

- j. Curated sequences of Cyt B fragment A
- k. Curated sequences of Cyt B fragment B
- l. Curated sequences of Rhodopsin
- m. Definitive sequences of Cyt B
- n. Definitive sequences of Rhodopsin
- h. Information sheets filled in (database)
- i. Meetings/workshops attended
- j. Dissemination activities done (specify)
- k. Periodical progress reports sent

Progress indicators suggested for a Reference Collection group:

- a. Vouchers included in collection
- b. Muscle tissues included in collection
- c. Pairs of otoliths included in collection
- d. Total DNA included in collection
- e. Information sheets prepared (database)
- f. Meetings/workshops attended
- g. Dissemination activities done (specify)
- h. Periodical progress reports sent

A starting date to produce the first progress indicator is foreseen at the end of August 2003.

18. Financial and administrative matters.

Mrs. Taylor exposed the main administrative matters concerning the contract which include the following points:

a. Payments:

The contributions from the Commission will be sent to the coordinator in four payments (Advance payment: 40%; 2nd payment: 25%; and 3rd payment: 20%; final payment: 15% which is retained by the Commission for two months after receiving the scientific reports and cost statements)

b. Eligible costs:

They shall fulfill the following conditions:

- be necessary for the project,
- be incurred during the duration of the project,
- be determined in accordance with the accounting principle based on historic costs and the usual internal rules of the principal contractor,
- be recorded in the accounts no later than the contract completion date or in the tax documents,
- exclude any profit margin,
- concern in principle only costs essential for administration, management and coordination of the project.

Eligible costs shall be reimbursed where justified by the participants. To this end, the participants shall maintain on a regular basis the accounts for the projects and appropriate documentation. This documentation must be precise, complete and effective (Ref.: FishTrace).

c. Submission of reports and statements:

It was indicated that the cost statements must be received by the coordinator in the format specified in the General Conditions and scientific reports by 15th of January of each year.

19. Next Meetings.

They were foreseen at least two meetings, one in early June in Stockholm to carry out a workshop meeting on taxonomy and systematics, and where issues arising from the other Workpackages that have already started will be also taken. Another at the end of the year where the report for the first year will be prepared. The location has not been established yet.

It was also mentioned that the group defining the data structure to the database might need a meeting in Paris in case it is needed to accelerate the agreement on the fields of the database. If this is the case the meeting will take place around April.

The sessions finished by February 14th at 18:00 h.

Annex II: FishTrace Database Structuration Meeting. Paris, 13-14 May, 2003.

FISHTRACE Paris MEETING: Database structuration

13-14 May 2003

Paris, France

Held at Muséum national d'histoire naturelle, Paris

Meeting Participants

Participant 1 (UCM): José M. Bautista; **Participant 2 (EC-JRC-IPSC):** Philippe Carreau; Naouma Kourti;

Participant 4 (ICCM): Montserrat Gimeno; **Participant 10 (MNHN):** Guy Duhamel, Patrice Pruvost

FishBase team: N. Bailly and M. L. Palomares (on Wednesday afternoon).

Meeting Agenda:

1st DAY, May 13 – Tuesday

1. Discussion on taxonomy table
2. Discussion on haplotyping concept and structure
3. Visit of the fish collection.
4. Discussion on species information
5. Elaboration on a previous database structure

2nd DAY, May 14 – Wednesday

1. Presentation and discussion on the final database structure
2. FishBase presentation and discussion with FishBase team on link between Fishtrace and FishBase
3. Visit of the bio molecular laboratory
4. Conclusions meeting

MEETING MINUTES

1st DAY, May 13 – Tuesday

The meeting started at 10 am.

1.- Discussion on taxonomy table

We decided to consider two main levels of information in the structure: a) General level and compiled information for each species; b) Specific information from each specimen

The general information would be linked to a main table “Species” or included in the table.

For the taxonomic purpose we won't use the subspecies level. The discussion about the taxonomic structure is reported to the discussion with the FishBase team.

Different fields will be necessary to describe the general information on the species for the Fishtrace project: Biological description specific to the “Fishtrace area” (Classification, Visual description, Diagnosis, Size and meristics parameters, Geolocalisation); Genetics consensus adopted at the issues of the project (Consensual sequence for the CytochromeB, for the Rhodopsine, Polymorphism sequence).

To each field the names of the different scientific contributors will be included.

The bibliography used to describe the species will be cited. A specific table is necessary.

2.- Discussion on genetics table

Jose Bautista explained differences between polymorphism and haplotyping. A special table will be necessary to describe haplotyping.

GenBank accessions numbers for the different sequences will be included at the “species” level.

The amplification methods for each PCR will be defined.

3.- Visit to the fishes collection.

4.- Discussion on sampling table

All the information on sampling will be included in this table: measurement, weight, count.

The data on otoliths and DNA will be mentioned in this table. The data from the tissues extract from the sampling will be integrated in another table (more than one tissue per specimen)

The reference to the collection number where the specimen will be stored will also be included idem for tissue, otoliths and DNA. The link between Fishtrace and the other collection databases for tissues, vouchers or otolith will be established through the reference number of the different collection.

The collecting data will be integrated in another table links to Specimen. (One data for one or more sampling).

5.- Discussion on pictures table

The discussion on pictures didn't produce a final decision. The information on the pictures will probably be included in one table, this table will be linked to the other through specific table link this one to the other one.

The scale in the picture must be integrated directly. (Suggestion to show a coin of 1 Euro).

6.- Discussion on the general database structure

We established the structure of the data base (Annexe 1)

2nd DAY, May 13 – Wednesday

The meeting started at 9.30

7- Discussion on the general database structure

Philippe Carreau presented the database structure established at the issues of the first day. Some comments and fields were added (i.e.: names of collaborators responsible for the data or identification).

8- Presentation of the FishBase structure and discussion of the link between Fishtrace and FishBase .

Nicolas Bailly introduced us to the FishBase data structure.

We established that the Fishtrace structure for taxonomy must be compatible with FishBase . The data from FishBase could be downloaded for the Fishtrace's species including the data from different table: Species, Synonyms, Country (for the distribution) (cf. Annexe 2 : FishBase taxonomic structure)

Nicolas Bailly will send to Philippe Carreau the specific data.

FishBase could be linked to Fishtrace for the specific data or pictures.

9- visit to the bio molecular section

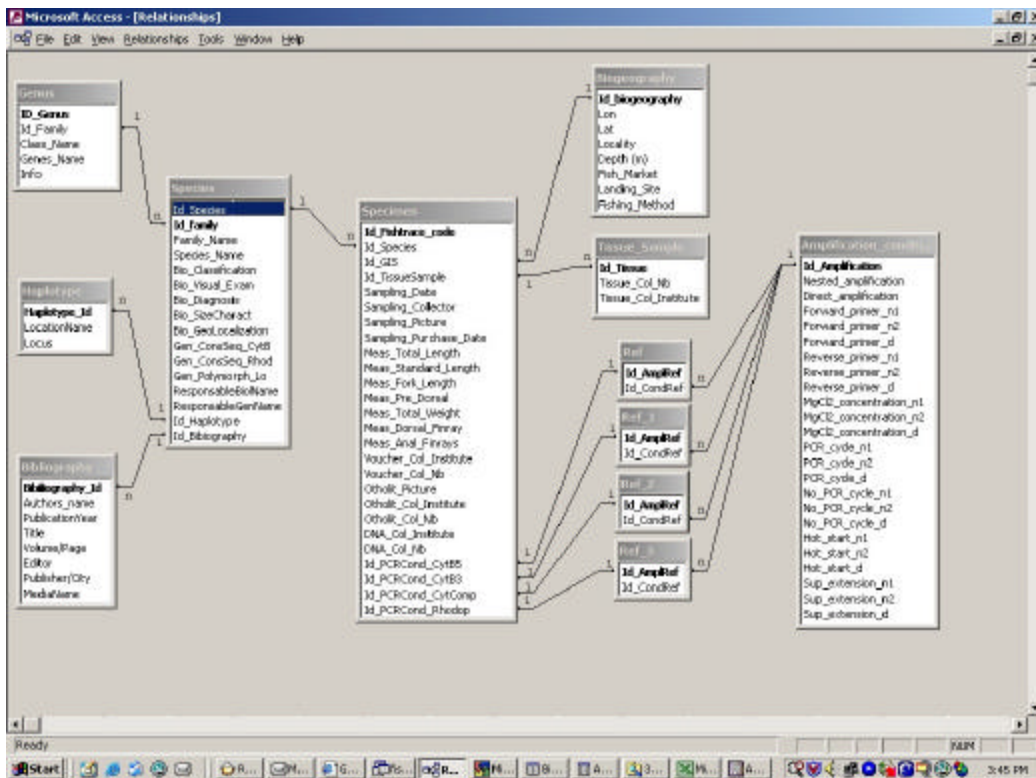
10- Conclusion of the meeting

Apart from the general structure of the database included in Philippe's schema, the following agreements for the database structure and web interface has been reached:

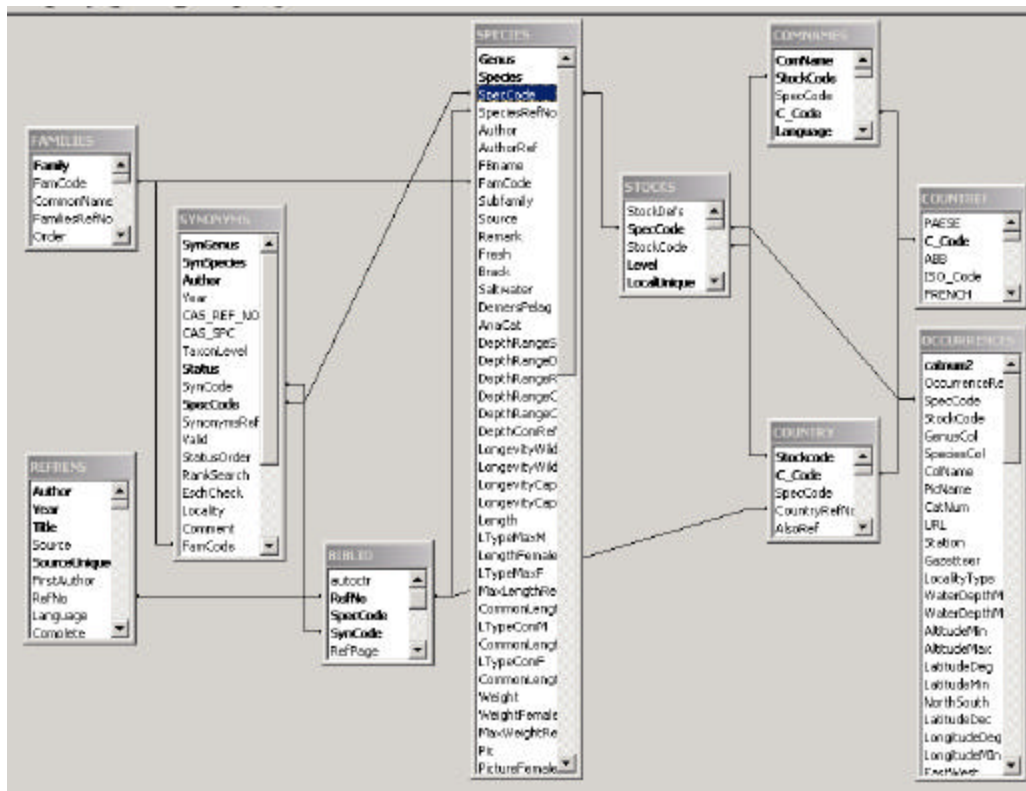
1. Two main levels of information in the structure: a) General and compiled information for each species (SPECIES). b) Specific information from each specimen analyzed including sampling, PCR, and collection information (SPECIMEN).
2. Not use the level subspecies, but add the level Order and Class
3. Species included in the same field in the form of "Genus + Species".
4. Include the Family code for identification at the Eschmeyer's Catalogue of fishes.
5. GenBank accessions numbers for Rhodopsin and Cyt b will be included at SPECIES level
6. Not common names, only FAO names in English, French and Spanish, plus the "FishBase " name. Link to FishBase for other names and languages. For those common and vernacular names not yet included in FishBase submit them to the FishBase team for inclusion..
7. Indicate in the text (v.g. visual identification) the correspondence between the data mentioned and the bibliography cited with a link to it.
8. Special emphasis to local european literature on the FishTrace species.
9. Include the scale in the fish pictures. A euro in the picture could do the job.
10. Include the methodology and protocol employed for obtaining data (measures, pictures, sequencing, bibliography, etc.) so anybody in the future can keep submitting species to the FishTrace database.
11. The general information of each species will be distributed to the partners participating in sampling so as not to duplicate efforts and facilitate the compiling of regional information.
12. ICCM (M. Gimeno) will translate into English 2 species from the Spanish database (Pescabase) to test initial structure of the FishTrace database.
13. ICCM (M. Gimeno) will define the fields for "Visual Identification" and "Diagnosis" comprised for the level of SPECIES.

14. World map and layers with coordinates for distribution: GSP capability
15. The JRC could decide the technological aspect for the database no special specification had been done during the meeting. The long term storage for the database must be discussed after
16. Phase 1: A first database structure, with the most important field, will be presented in Stockholm meeting for discussion. Development will be down after. Phase 2: Other development like GIS facilities will be down after.
17. The queries and interface will be taken at Stockholm meeting.

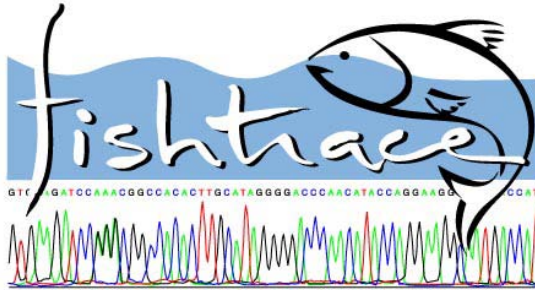
Annexe 1 : Fishtrace database structure



Annexe 2 : FishBase taxonomics structures



Annex III: FishTrace Standardization Meeting. Stockholm, 4-7 June, 2003.



FISHTRACE MEETING
4 - 7 June 2003
Stockholm, Sweden
Held at the Swedish Museum of
Natural History (NRM) and the
Simpnäs Field Station.

Report from the second FishTrace meeting, Stockholm 4-7 June 2003

Objectives

To finalize WP2 and WP5 protocols and to conclude work on the database format.

Venue

Swedish Museum of Natural History and Simpnäs Field Station

Dates

4 - 7 June 2003

Participant list

Hilde van Pelt (Nederlands Instituut voor Visserij Onderzoek)
Ignacio Lozano (The Canarian Institute of Marine Science)
João Delgado (Museu Municipal do Funchal)
José M. Bautista (Universidad Complutense de Madrid)
Jose-Antonio González (The Canarian Institute of Marine Science)
Mafalda Freitas (Instituto do Mar Portugal)
Manuel Biscoito (Museu Municipal do Funchal)
Michael Norén (Swedish Museum of Natural History)
Panos Leoutarakis (National Agricultural Research Foundation)
Patrice Pruvost (Museum National d' Histoire Naturelle)
Philippe Carreau (Joint Research Center Italy, Ispra)
Rosa Pestana (Instituto do Mar Portugal)
Sebastián Jiménez (Museo de Ciencias Naturales de Tenerife)
Sven Kullander (Swedish Museum of Natural History)

Hosts: Sven O Kullander (sven.kullander@nrm.se), Michael Norén (michael.noren@nrm.se), Peter Mortensen (peter.mortensen@nrm.se) and Alia Jaensson (has since left NRM).



Picture courtesy of Panos Leoutarakis

AGENDA

Day 1 Wednesday:

- 09:00 Introduction to NRM, Recapitulation of progress. Individual presentations.
- 10:30 Group work: WP2. Carried on to day 2
- 12:00-13:30 Lunch
- 13:30-14:00 Group work: WP5.
- 19:00 Dinner

Day 2 Thursday:

- 09:30 Transfer to Simpnäs
- 11:00 Lunch
- 13:00 Arrival, setup, nets set
- 13:30 Group work on Field Forms and tissue sampling
- 18:00 First catch recovery
- 18:10 Practical exercise
- 19:00 Dinner

Day 3 Friday:

- 09:00 Second catch recovery (practical exercise)
- 09:10 Practical exercise
- 11:00 Discussion on procedures and possible modifications
- 14:00 Final protocol draft including modifications from discussions Day1-3
- 15:00 Return to Stockholm

Day 4 Saturday:

09:00 Database issues, from Day 1.

Definition of the data fields for database, submission of data to the database, final definition of the species for haplotyping and its methodology. Use of the FishTrace website for communications.

Day 5 Sunday:

Departure day

The goals of this meeting were to define the final Sampling protocol, Sampling form, and to agree on the structure and fields of the Database.

MEETING MINUTES**Day 1**

1. The meeting started at 9, with an introduction by Sven Kullander, and a short presentation of the project progress and objectives by the co-ordinator, José Bautista.
2. The proposed sampling protocol, already sent to participants was introduced by Michael Norén.
3. Participants grouped in four discussion groups to discuss separate portions of the WP2 protocol.
4. Lunch
5. Continued group discussion and communal summary. Each group reported on distinct problems, clarifications, and additions, and the protocol was modified according to the suggestions. A major problem was to define a realistic level of morphological information. It is important to avoid blank fields in the database, and include parameters relatively straightforward to record, still to include morphological information relevant to geographic variability and species diagnoses. The scheme of deposition of vouchers also changed slightly, so that in principle all vouchers are saved. The system of numbering the samples was simplified to exclude redundant information. These changes are reflected in the modified protocol which can serve as standard for this kind of sampling.
6. WP5, Biological collections, was discussed in the large group and accepted without major changes. Issues regarding voucher deposition were already covered in the WP2 discussion. It was also clarified that specimens have to be fixed in formalin before final storage in alcohol, and that DNA tissue samples preserved in 95% ethanol and stored cool (in a fridge or freezer) will be sufficient for long term DNA storage. Discussion continued till after 1700h.
7. The group then made a tour of the Vertebrate Department, in particular the Fish Collection and facilities.
8. Dinner 19.00

Day 2

1. Travel by bus to Simpnäs between 9.00 – 13.00.
2. Sven Kullander took the participants on a tour of the Simpnäs facility, after which the participants were given some time to install themselves in the four different houses of the Simpnäs facility.

3. At 15.00, one group of participants joined Peter Mortensen and set gill nets, while another collected fish using a beach seine, after which Michael Noren gave a demonstration of field DNA sampling.
4. Dinner at 19.00

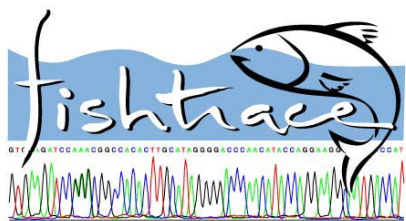
Day 3

1. Starting 9.00 the participants tested the proposed sampling protocol by taking samples from the collected fish.
2. Discussion and views on the sampling protocol, which resulted in the addition of a number of new fields, and the deletion of a number of old. Also, many fields were re-worded for clarity. It was also suggested that the protocol should be distributed in MS Word Form format, to make it easier to fill in on computer.
3. Starting 13.00, there was a discussion on the structure and necessary fields of the database.
4. At 14.00 everyone joined in and cleaned the houses and packed the equipment.
5. Bus back to Stockholm at 15.00.

Day 4

1. Starting 9.00, Philippe Carreau held a presentation on the present status of the FishTrace website and database, and also explained what changes will be made in the near future.
2. 10.00 Concluding discussion on the structure and content of the database, and how it should interface to the FishTrace research groups and to external databases such as FishBase and GenBank. It was decided that keeping the names and synonyms list updated was best achieved by linking to FishBase. Also, a number of fields in the database should be deleted, a number added, and several modified.
3. It was decided that the next FishTrace meeting be held in the Canary Islands towards the end of November.
4. 16.00 The meeting ended with concluding remarks by Sven Kullander and José M. Bautista.

Annex IV: 2nd FishTrace Plenary Meeting. Las Palmas de Gran Canaria (Spain), 19-21 November, 2003.



FISHTRACE MEETING

NOVEMBER 19-21, 2003
Las Palmas de Gran Canarias
Canary Islands, Spain.

Report from the second FishTrace Progress Meeting.

Objectives:

To show and reflect the progress of all work-packages and tasks during the first period of the project. To find agreement positions on the definitive structure and fields of the Database. To define the legal aspects. To solve problems encountered at different stages of the different workpackages and connections.

Venue:

Held at Las Palmas at the Reina Isabel Hotel and the ICCM conference hall.

Dates:

19 - 21 November 2003.

Participant list:

ICCM:

José Antonio González (**Host**)
Ignacio Lozano
Montse Gimeno

IFREMER:

Monique Etienne
Véronique Verrez-Bagnis

IMAR:

Manuel Biscoito
João Delgado
Rosa Pestana

JRC:

Philippe Carreau
Romas Statkus

MNHN:

Patrice Pruvost
Samuel Iglesias

NAGREF:

Grigorios Krey
Laurence Favre-Krey
Alexis Tsangridis
Panos Leoutarakis

NRM:

Michael Norén

RIVO:

Hilde van Pelt

TFMC:

Fátima Hernández
Sebastián Jiménez

UCM:

José M. Bautista
Gema González
Rafael G. Sevilla

Agenda:

1st DAY

- | | |
|---|---------------|
| 1. Welcome to fishtracers | J.A. González |
| 2. Meeting introduction (on behalf of J.M. Bautista) | R.G. Sevilla |
| 3. Workpackages progress info | |
| 3.1 WP1. Standardisation of Methods and Legal Aspects | H. van Pelt |
| 3.2 WP2. Sampling and taxonomy | M. Norén |
| 3.3 WP3. Molecular Genetic Identification | R.G. Sevilla |
| 3.4 WP4. Biogeographical Molecular Haplotyping | G. Krey |
| 3.5 WP5. Biological Reference Collections | M. Biscoito |
| 3.6 Thematic discussion on WP1 & 2, on WP3 & 4, and on WP5 | |
| 3.7 WP7. Validation in Database | J.A. González |
| 3.8 Including Information flow at regional level | J.A. González |
| 3.9 General discussion on WP7 | |
| 3.10 WP9. Dissemination and Applicability | M. Etienne |
| 3.11 General discussion on WP9 | |
| 4. Workplan progress info by partner: | |
| 4.1 Partner no. 5 (IFREMER) | V. Verrez |
| 4.2 Partner no. 10 (MNHN) | S. Iglesias |
| 4.3 Partner no. 1 (UCM) | G. González |
| 4.4 Partner no. 2 (EC-JRC-IPSC) | P. Carreau |
| 4.5 Partner no. 3 (NRM) | M. Norén |
| 4.6 Partner no. 4 (ICCM) | J.A. González |
| 4.7 Partner no. 8 (TFMC) | S. Jiménez |
| 4.8 Partner no. 6 (RIVO) | H. van Pelt |
| 4.9 Partner no. 7 (IMAR) | R. Pestana |
| 4.10 Partner no. 9 (NAGREF) | G. Krey |

2nd DAY

- | | |
|--|---------------|
| 1. Introduction to PescaBase database/web | J.A. González |
| 2. Workpackages progress info (continued) | |
| 2.1 WP6. Database Development | P. Carreau |
| 2.2 Testing off line database | |
| 2.3 WP8. Implementation of www Technologies | P. Carreau |
| 2.4 General discussion on WP6 & 8 | |
| 3. A review of fish genetics implications on fisheries management | R. Statkus |
| 4. Troubleshooting on Taxonomy, Collections, Database, | |

3rd DAY

- | | |
|--|---------------|
| 1. Visit to Canarian Institute of Marine Sciences | D. Schuchart |
| 2. Workshop on FishTrace Legal Aspects | H. van Pelt |
| 3. Workshop on Project's prospects | J.M. Bautista |
| 4. Troubleshooting on Genetics | J.M. Bautista |
| 5. First Scientific and Financial Reports | J.M. Bautista |
| 6. Next FishTrace meeting | J.M. Bautista |

Meeting Minutes:

1st DAY

1. Welcome to fishtracers

9:00 The meeting started with a welcoming introduction by José A. González.

2. Meeting introduction

A short presentation about the progress on each work-package of the project during this first period, by Rafael G. Sevilla (on behalf of José M. Bautista, co-ordinator of the project), based on data from Progress Report, September 2003.

3. Workpackages progress info

3.1. Hilde Van Pelt introduced a detailed presentation about the standardisation of methods (Sampling and identification protocols are available online) and the importance of the quick establishment of Legal Aspects of the project (Patents: for the identification method is possible. Availability, exchange, loans and gifts during and post-Fishtrace was discussed at the end of the meeting (2. Workshop on FishTrace Legal Aspects).

3.2. Michael Norén continued with WP2 progress during this first year, concluding that almost half of the total number of targeted species for each geographical areas are sampled and taxonomically identified. SB: 26/57, NS: 21/46, CB: 21/53, CS: 27/62, MA:24/48, CI: 16/49, WM: 73/80, EM: 29/54, EE: 0/43 (Results explained on Meeting Introduction presentation).

3.3. Rafael G. Sevilla did the same for WP3 explaining that the first task of this workpackage was the standardisation of methods: Extraction (Tissue samples from six different species were provided by UCM, and sent to all groups involved in this workpackage, where DNA was extracted using its own methodology, as described at the presentation). Optimised amplification protocols (Primers and PCR conditions), were provided too. Sequencing and correcting sequences were explained, detailing the checking points that the correction need to pass before including a sequence on the Database.

At present there are 279 DNA extractions from tissue samples by all partners. The extraction of DNA from around 1000 tissue samples are necessary to complete the expected number of sequences for the Genetic Catalogue, so a third of this task is already done.

3.4. Greg Krey continued with WP4 progress info: For two of the species selected for haplotyping (*M. merluccius* and *M. surmuletus*), 20 specimens are presently available for analysis. The rest of species are not complete and the definitive list should be discussed soon.

3.5. Manuel Bischoff exposed the progress on WP5: The total number of vouchers, tissues and otolith samples included in collections were: 152 species, 356 specimens, 219 tissue samples and 54 pairs of otoliths.

3.6. Thematic discussion on WP1, 2, 3, 4, and WP5.

WP1: Postponed to last day.

WP2: The groups involved in this workpackage did not find any trouble except for the identification of some species as it was discussed on the next point (WP4).

WP3: Among Genetist involved in this project some troubles were found with the amplification of Rhodopsin gene. To solve these, it was suggested a revision of the protocols for the amplification of both genes.

WP4: There was a discussion about species (e.g. *Solea solea*) complicated to identify taxonomically. A list was put forward with the kind of species that should receive a special treatment: (e.g. *Solea*, *Epinephelus*, *Serranus*,...). It was established that taxonomist-collectors had to check this conscientiously. Greg Krey suggested an exchange of sequences between genetic groups to construct phylogenetic trees for this aim (alignments, comparisons, etc.), and exposed the problem of hybrids.

WP5: Some fields should be changed into Reference Collections Progress Reports and into Database to give more and better quality information.

3.7. José A. González continued with WP7. We must validate the data set on taxonomy, biology, bibliography, molecular genetics and reference collections. The curated data will be included at the on-line database. For this, a three/four-expert group will be proposed by project's co-ordinator as co-responsible for data validation in database:

- responsible for taxonomic-biological-bibliographic data validation
- responsible for collections data validation
- molecular genetic data validation
- responsible for photos and otoliths images.

Four steps of validation:

1. Specific data checking (errors, missing data,....)
2. Specific data compilation
3. Specific database field validation
4. Data collating and arranging

The methodology to follow for validation on each task is described in the presentation.

3.8. Including Information flow at regional level

José A.González proposed a system for the information flow based on the exchange of three EXCEL files to establish a feedback information control system of:

- Identity and location of every removed biological sample
- Identity and location of every reference collection sample
- Early control of the taxonomical identity of every specific sample

It was shown an example of each EXCEL file. Moreover, it was explained how the information flow works between the Institutions involved.

3.9. General discussion on WP7. A question about validation on taxonomic results and definitive sequences was raised. It was agreed a referee committee composed at least by three experts, one for each task: Sampling-Taxonomy, Collections and Genetics, to validate data.

As well as, the information flow at regional level system was agreed. In brief these files will be ready and sent to all partners in due course.

3.10. WP9. Dissemination and Applicability: Objectives: disseminate the results & provide free-access to the deliverables developed in the network. Monique Etienne explained that this workpackage will start on January 2004. Some questions must be answered before then: There must be someone responsible for writing the results. These results could be disseminated by a scientific, technical or professional journal (e.g. Seafood, Eurofish, etc). In our project we put the sequence on Internet and we keep samples in collection so concerning the Bio-Mérieux request for DNA chip, some questions were raised: To sell or not to sell the results & the samples?
Do we have this possibility in a EU contract?

3.11. General discussion on WP9. It was agreed to print a first draft of a brochure with a small CD containing extended information about the Project that could be useful to show on the next seminars/symposiums to come: WEFTA'05 (Holland), FASIS'05 (Nantes), SeaFood Exhibition (Brussels), etc. First draft: February 2004. First draft of a book compiling results and information of the Project could be ready by May 2004 and printed by the end of December 2004. Language: It will be printed in English and subsequently translated at the relevant language for each country. It was discussed the chance to link our website to others like: FishBase, GenBank, MAPA, FROM, all groups' websites, etc. It was agreed not to sell samples and the free acces for the public to the database but with a previous registration. It was agreed to be very careful with articles that could be send to scientific magazines during the project, so everything must be consulted with the coordinator before sending it to any journal, magazine, etc.

4. Workplan progress info by partner

4.1. Véronique Verrez-Bagnis continued explaining IFREMER's workplan progress:

WP1: They have participated in the standardisation of DNA extraction from muscle tissue samples sent by UCM. They have used the phenol / chlorophorm / proteinase K method for the DNA extraction and sent the extracts to UCM.

WP2 and WP4: They (MNHN and IFREMER) have to sample about 57 fish species from the CB zone (Bay of Biscay and English Channel), and they have already sampled 37 species, 31 of them are complete. They took all these specimens during research cruises on IFREMER ships. Most of these specimens have been brought in ice from ship to lab., but when the cruise was far from the coast and without any call, they have been obliged to freeze them (a picture was taken before freezing in this case). They were able to complete the environmental data table precisely (geographical coordinates and conditions of the catch in particular). Colleagues from the MNHN will check the identity of the preserved specimens.

WP3: They have amplified the rhodopsin and the cyt.b genes using the UCM protocol and the 1.1. conditions as recommended (without any problem). Since June,

they have got a DNA sequencer (CEQ 8000 Beckman Coulter), so now they are able to sequence DNA by themselves. They have sequenced the Rhodopsin gene for these 6 species and the sequencing of Cyt-B gene is ongoing.

WP7: They have, in the beginning, made a meeting in Nantes to agree for the sampling. They are using a common Excel file to follow the sampling and the pictures of the specimens are all sent to MNHN.

4.2. Samuel Iglesias talked about the advances, problems and solutions in sampling and collection in the Bay of Biscay and English Channel areas where MNHN is involved: 34 species with 160 specimens are included in the collection; about 100 specimens received from RIVO; $\frac{3}{4}$ parts of species are completed for haplotyping. 153 Tissues samples and 21 pairs of otoliths are included in the collection too. Some problems and its proposed solutions about quality of voucher, photographs and difficult to found species were exposed and can be consulted on his presentation.

4.3. Gema González continued with UCM's workplan progress. This group is in charge of 4 geographical areas (CS, WM, MA and EE), and the total results obtained up to september 2003 were: 221 DNA extractions from 273 muscle tissues received.

12 definitive sequences of CytB and 31 of Rhodopsin were obtained. For this work, they have used a new 6100 Nucleic Acid PrepStation (Applied) to carry out the extractions.

4.4. Philippe Carreau explained developments on the database concluding that the first step is to load the data because it determines data quality. The second step is to create a web public interface that could be similar to Pescabase. Finally, the third step is to exploit the database : Extract statistics, cross query results, etc. Another utility could be to create maps & GIS (geospatial information). The work plan of JRC for the different workpackages is as follows :

WP1 : Methodological Standardization and establishment of data structure.

WP6 : Database development (Database implemented).

WP8 : Implementation of WWW technologies (Web data loading interface implemented).

WP9 : Dissemination: user-oriented applicability of database (General web site has been delivered).

At the end of the meeting, some items were clarified (4. Troubleshooting on Taxonomy, Collections and Database):

Train and make remarks on the loading interface.

Definitely fix the table architectures related to species (as "regional info") in continuation with Stockholm.

Decide the basic public interface design.

Make propositions for database exploitation.

4.5. Michael Norén continued with the progress raised by NRM during this first year: Sampling: 30 (of 59) species sampled, 130 vouchers (entire fish in formalin), 300 tissue samples (some are whole, small, fish in alcohol), 150 photos. About 2/3 entered into

museum database. Sequencing: 55 DNA extractions, 200 PCR reactions, 92 ABI-lanes. Only 5' fragment sequenced for 8 taxa. Complete Cytochrom B sequences from 14 taxa.

4.6. José A. González reported the progress of his group in terms of the researchers task organisation, mayor team, fish and samples displacements around the mainland of Spain and total number of species sampled and taxonomically identified until September, 2003 for each area they are encharged: 27 Spp for Cantabric Sea, 76 for Western Mediterranean, 37 for Canary Islands and 8 for Extra- European Seas. As well as the species sampled for haplotyping in each area were related: *M. merluccius* (CS), *M. merluccius*, *M. surmuletus*, *P. erythrinus*, *P. pagrus* and *S. solea* (WM) and *P. erythrinus*, *P. pagrus* and *M. surmuletus* (CI).

4.7. Sebastián Jiménez explained the progress of TFMC, Institution that works quite close to ICCM, so as it was told, he participated at the same displacements than ICCM team working in biometric sampling and taxonomic identification at Cantabric Sea and Western Mediterranean areas. Total results until September, 2003 were: 51 (CS) and 42 (WM) voucher in the TFMC collection; 38 (CS) and 115 (WM) pairs of otoliths in collection; 130 (CS) and 328 (WM) tissue samples in TFMC collection.

He also shown the way he works photographs with the Adobe Photoshop.

4.8. Hilde van Pelt (postponed to last day: Workshop on FishTrace Legal Aspects)

4.9. Rosa Pestana exposed the progress and the workplan of her group (IMAR): A total of 48 targeted species are included in the list for IMAR. 32 of them are sampled and 24 are completed at present. Concernin WP1, they did not find any problem and the methodological standardisation and establishment of data structure was carry out properly. About WP2, they have started sampling and identification of the specimens as well as the transference of tissue samples to genetics team. Some species were replaced (described in the presentation) from the first list of targeted species of their geographical area (MA), and they have completed the sampling of two species for Haplotyping. The protocols to organize collections were done and the reference collections were started (WP5).

4.10. Greg Krey explained the workplan for his group, detailing some tasks to do: (described in the presentation)

- Establish genetic identity of sampled species.
- Construct phylogenetic trees.
- Population genetic structure of selected species.

The most common problems found after all interventions were: the list of problematic species, definitive format of the photos and other difficulties on Sampling and Taxonomy.

2nd DAY

1. Introduction to PescaBase database/web

José A. González introduced PescaBase Database system and interface to all partners. He explained the parallelism between FishTrace and PescaBase objectives, and tested on-line the results obtained and the information filled at PescaBase at present.

2. Workplan progress info (continuation)

2.1. Philippe Carreau continued with progress on Database: All tables were build except "Bibliography" and "Haplotype", and "Regional_info" should be clarified. These items were discussed on Troubleshooting on Database. Search tools have been developed on "Specimen", "Species" and "Collecting_Environment".

2.2 Testing off line database. Concluding the discussion on the structure and content of the database, and how it should interface in the FishTrace research groups and users.

Panos Leoutarakis volunteered to test filling the database, followed by a group discussion on the definitive structure and content of database. The list of targeted species was discussed bearing the problems that appeared with the identification of some individuals belonging to these "problematic" species. It was suggested more flexibility and the possibility to include easier or exclude more difficult species to identify. Concluding this theme, it was agreed that the team leader of the institute in this situation would write a Memo to JRC group to change the list would it be necessary to include/exclude any new/old species. For those difficult to identify species, it was proposed by Patrice and Michael to send a picture with the name of the specie to some expert taxonomy group to validate it. All participants discussed about Bibliography on Database.

3. A review of fish genetics implications on fisheries management

Romas Statkus informed all partners about fish genetic implications on fisheries management. For this purpose Romas introduced to the group six European Projects working in related studies about fish populations, migrations, etc., using a brief but very informative presentation with examples: Cods, Tuna, etc.

4. Troubleshooting on Taxonomy, Collections and Database

Michael Norén presented the most common problems that appeared on WP2: species for which they were unable to obtain five specimens, broken otoliths, non-applicable or missing biometric data, very small fishes, etc.

It was suggested again as the best solution an opened list of targeted species.

The list of species for Haplotyping was changed, bearing the difficulty to find some old species of the candidate list for some groups. *Mugil cephalus*, was changed for *Micromesistius poutassou*.

Another issue discussed was the photos final format. Detailed, JPEG and named by FishTrace protocol: FishTrace code is used to name the photo and details must come out as "remarks". The optimal size is: 1024x768 pixels with a Photoshop compression of 6-8 (1,3 Mb). To conclude the best of all photos taken of each specimen, and the more informative one will be chosen for the database.

When contributing data each group will include the ones exclusive to its geographical area. For common species it is discussed between the groups that have them in their targeted lists before sending the information to the database.

Several fields in the database should be added (Official European Name), and modified (gillrakers, finrays, etc.), but the main discussion was centered on Regional Info fields. It was agreed to add two new fields on Regional Info sheet: Common missing Identification and Common substitutions of species on markets.

PescaBase Database system and its interface was chosen as definitive option for FishTrace's webpage. José A. Gonzalez allowed to Philippe Carreau to get in contact with the PescaBase webmaster. This will be working by May-June 2004.

3rd DAY

1. Visit to the Canarian Institute of Marine Sciences

9:00 a.m. Travel by bus to Taliarte, where the ICCM is located. Once there, Dominique Schuchart toured the group through the Institute and explained the different tasks for each line of research developed there and how they work, as well as other services offered by the ICCM (e.g. Injured Wild Turtles Rehabilitation). After this tour the group attended Manuel Biscoito's lecture about "Deep-sea Hydrothermal Vents".

Once finished, the group met again at ICCM lecture hall and continued with the meeting program.

2. Workshop on FishTrace Legal Aspects

Hilde Van Pelt started the discussion about legal rights on results obtained in this project (including photos, etc.) and possible patents. It was agreed that sequences cannot be patented but it is feasible to patent methods as well as Oligos, Primers, etc. The suggested sale of samples was discarded. Tissue samples (preserved as reference collections) are appropriate tools to check the availability of methods by the analysis of those samples in the future. It was agreed not to exchange or give away any information without previously consulting the consortium. Free access to FishTrace Database for the public but with a previous registration was discussed too.

A discussion arose on using this information commercially or just as exclusive research guidance for the project.

3. Workshop on Project's prospects

José M. Bautista informed about the necessity of having the database in full working order for January 2004, publishing an article compiling the standardisation of the methods as well as the first results obtained. The sources will include scientific articles on genetics, research and fishery biology. We agreed on writing reports with the problems found and circulate them between the team leaders via the web.

4. Troubleshooting on Genetics

José M. Bautista exposed the inconveniences found when dealing with the genetic molecular identification and the protocol to obtain better results and to be able to optimise on time and costs. The phylogenetics studies and the microarrays development for identification are powerful tools to validate the data obtained hence we should investigate more on this as from today. Once the presentation was finished a debate was opened on haplotypes and the exchanges of sequences between the teams in charge on this task. Greg Krey suggested using fresh water fishes for studies on hybrids, José A. González suggested *serranids* and Veronique just the rodopsine gen. José A. Gonzalez warned about the problems he's had with the Cytb amplification in the Conger, conger samples. José M. Bautista suggested a design on alternative primers when analysis this species.

5. First Scientific and Financial Reports

José M. Bautista insisted on the importance in publishing. For this reason it is very urgent that we establish and validate all aspects of the project for good (database, collection, genetics) In order to work properly as from January 2004, it was agreed to exchange all the information (Sampling, Taxonomy, Molecular Genetics) on the 5 species targeted on the list (5 for each geographical area) from which we would start filling the validated database.

6. Next FishTrace meeting

It was decided that the next FishTrace meeting be held in Ispra towards April-May of 2004, with the team leaders participating so the final structure of the database would be agreed as definitive there. Other meeting must be fixed until end 2004 to summarize what is developed during the first two years of the project.

The goals of this meeting were to agree on the structure and fields of the Database as well as to define the legal aspects. A short term publication of Methods and first Results is the next task to complete by starting 2004.

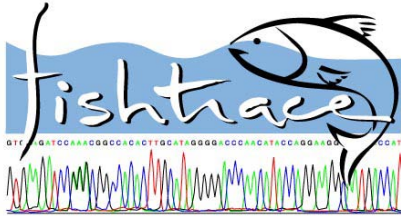
Urgent to send:

- 1st scientific report.**
- 1st financial report.**
- 5 completed species to fill on Database.**
- New PCR conditions.**
- Sequences to make phylogenetic trees.**



All participants at the FishTrace Meeting (M.Norén not present), at *El Gabinete Literario*.
20 November 2003. Las Palmas de Gran Canaria, Canary Islands, Spain.

Annex V: FishTrace Meeting on Data Validation. Ispra, Italy. 15-16 April, 2004.



FISHTRACE MEETING

APRIL 15 - 16, 2004

Ispra, Italy.

Report from the FishTrace Meeting on Data Validation, Database Check and Database Result Presentation.

Objectives:

To define data validation and database checking procedures according to type of data. Subjects were database check, data validation and database result presentation. To show and reflect the progress of all partners/workpackages during this period of the project.

Venue:

Held at room 2 building 36, JRC Ispra site.

Dates:

15 - 16 April, 2004.

Participant list:

JRC: (Host)

Naouma Kourti
Philippe Carreau

ICCM:

José Antonio González
Ignacio J. Lozano

IFREMER:

Véronique Verrez-Bagnis

IMAR:

Manuel Biscoito

MNHN:

Patrice Pruvost

NAGREF:

Grigorios Krey

NRM:

Michael Norén

RIVO:

Hilde van Pelt

TFMC:

Sebastián Jiménez

UCM:

José M. Bautista

Agenda:

1st DAY

1. Welcome to Fishtracers

N. Kourti & P. Carreau

2. Meeting introduction

J.M. Bautista

3. Status presentations

3.1. Methodological Standardization status

H. von Pelt

3.2. Partner no. 1 (UCM)

J.M. Bautista

3.3. Partner no. 3 (NRM)

M. Norén

3.4. Partner no. 4 (ICCM)

J.A. González

3.5. Partner no. 8 (TFMC)

S. Jiménez

3.6. Partner no. 5 (IFREMÉR)

V. Verrez-Bagnis

3.7. Partner no. 6 (RIVO)

H. van Pelt

3.8. Partner no. 7 (IMAR)

M. Biscoito

3.9. Partner no. 9 (NAGREF)

G. Krey

3.10. Partner no. 10 (MNHN)

P. Pruvost

3.11. Database status

P. Carreau

4. Data validation & database check

4.1. Data validation

J.A. González

4.2. Genetic Data Validation

J.M. Bautista

4.3. Quality check exercise results

P. Carreau

5. Creation of interim working groups according to type of data.

2nd DAY

6. Communal Summary from the working groups

6.1. Genetic

J.M. Bautista

6.2. Taxonomy

M. Biscoito

6.3. Regional Info

J.A. González

7. Data result presentation

7.1. GIS capability

P. Carreau

7.2. Web site user needs

J.M. Bautista

8. General conclusion

8.1. Next Fishtrace meeting

J.M. Bautista

8.2. Conclusions and Project Schedules

J.M. Bautista

Meeting Minutes:

1st DAY

1. Welcome to fishtracers

9:00 The meeting started with a welcoming introduction by N. Kourti and P. Carreau.

2. Meeting introduction

A short overview about the progress of the project during this first period by José M. Bautista, based on data from the First Annual Report, February 2004.

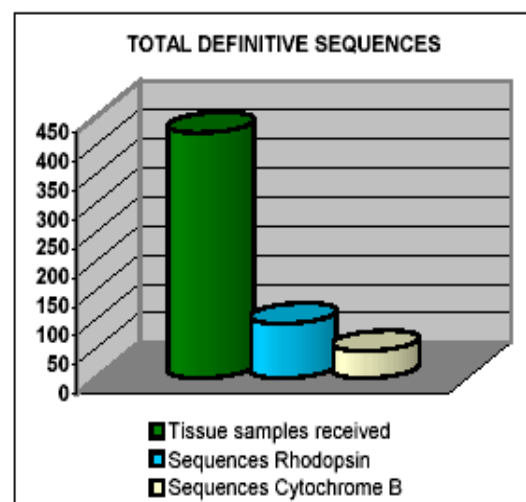
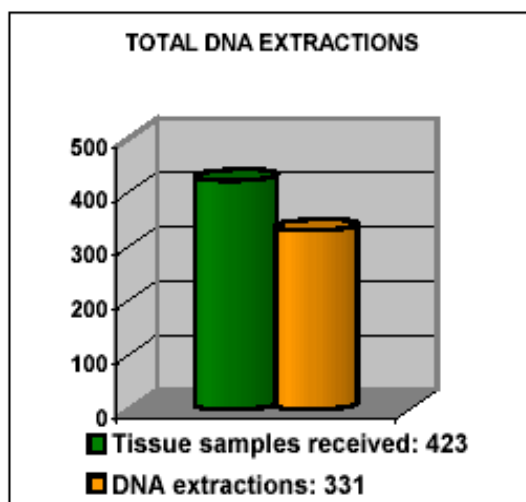
3. Workpackages progress info

3.1. H. van Pelt presented the two milestones corresponding to the standardisation of methods that have been reached, Milestone 1: Standardisation of experimental procedures, which we have reached on schedule and Milestone 2: Legal definition concerning use of the FishTrace online database.

After the intervention of H. van Pelt, it started an overview of the status of the project by each partner, followed by a short discussion on genetic problems.

3.2. J.M. Bautista, leader of UCM, showed the progress on Molecular Genetic Identification (WP3), detailing objectives achieved and results obtained during this period of the project:

- Development of specific protocols.
- Development of specific primers for PCRs and sequencing.
- DNA extraction from 78,25% of received samples.
- 174 succesful PCRs for *cytochrome b* and 88 for *rhodopsin*.
- 47 definitive nucleotide sequences obtained of the complete *cytochrome b*.
- 94 definitive nucleotide sequences of the *rhodopsin* fragment from the target fish species of the sampling areas D, E, F, G and the extra-european species.



Concerning the Biogeographical Molecular Haplotyping (WP4), J.M. Bautista mentioned that 194 tissue samples were received for haplotyping although the specific genetic analyses will start during this second period of the project.

J.M. Bautista showed several pictures of agarose gels (results obtained on the successful amplification of both target genes (Cytochrome B and Rhodopsin). Finally, he presented a list of expected deliverables and achievements by partner 1 at the end of 2004 for the UCM, detached by workpackages:

WP3. To reach 2/3 of the total definitive sequences planned for the project.

WP4. To obtain 1/2 of the total definitive sequences planned for Haplotyping.

WP6. To fill in the corresponding fields of the database with the obtained data.

WP7. To participate in Validation of the Molecular Genetic Identification data.

WP8. To collaborate on the implementation of WWW technologies.

WP9. To collaborate on the Dissemination and Exploitation of the data obtained:

- Presentation at Meeting in Naples (July). International Meeting on Veterinary Morpho-Functional Biotechnologies. "*Species identification by Molecular Genetics: from PCR to Molecular Databases and DNA Microarrays*" (Plenary Lecture)
- Presentation at Meeting in Lisbon (October). European Fisheries and Aquaculture Research Organisations (EFARO). *Title to be determined*
- Manuscript on "*Microgenomic exploration of teleostei: a standard catalogue and biological reference collections for fish identification with 60 newly determined sequences-signatures*".
- Developing of a first Prototype of Microarrays for fish species identification.

Associated dissemination and exploitation actions:

- Presentation at Aquaculture Europe 2004 (Barcelona, October): "*A Multiplex PCR method for species identification of fish fillets from Grouper (Epinephelus spp. and Mycteroperca spp.) and its usual substitution species*"
- Manuscript on "*A Real-Time Multiplex PCR for species identification of fish fillets from Grouper and its usual substitution species*"
- European Patent: DNA Identification Kit in processed foods and animal feeds (Plants, Animals, Birds, Mammals and Fishes)

3.3. M. Norén explained the results obtained by NRM during this period of the project:

Sampling:

- 41 (of 59) species sampled
- 200 vouchers (entire fish in formalin)
- 400 tissue samples
- 270 photos
- About 2/3 entered into museum database.

Sequencing:

- Complete Cytochrome B sequences from 30 taxa (17 taxa in duplicate = completed)
- No Rhodopsin sequences finished.

On schedule:

- Have started cooperation with the Institute of Marine research with sampling.
- Some problems with new sequencing equipment.
- NRM have found some problems amplifying the Rhodopsin gene. The rest of tasks are working well.

3.4. Representing ICCM, J. González detached the work accomplished by this group on Sampling and Taxonomy and the rest of tasks:

Sampling status at ICCM:

Between 80-100% of the species of geographical regions in charge of ICCM are done (except the species from EE: almost 50% done)

- Cantabric Sea & NW Iberian Peninsula (CS):
 - Number of targeted species: 62
 - Number of species sampled & taxonomically ID: 29
 - Species sampled for haplotyping: *M. merluccius*
 - Sampling sites: Basque Institute for Fisheries and Food Technologies (AZTI), Sukarrieta, Bizkaia, Spain. ICCM Labs., Telde, Las Palmas, Spain.

- Western Mediterranean & Bay of Cadiz (WM):
 - Number of targeted species: 86
 - Number of species sampled & taxonomically ID: 86
 - Species sampled for haplotyping: *M.merluccius*, *M.surmuletus*, *P.erythrinus*, *P.pagrus* & *S.solea*
 - Sampling sites: Fishermen Professional Association, Port of Santa Pola, Alicante, Spain. ICCM Labs., Telde, Las Palmas, Spain.

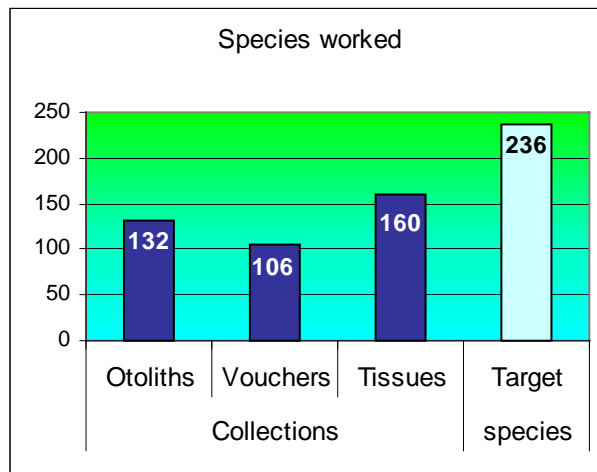
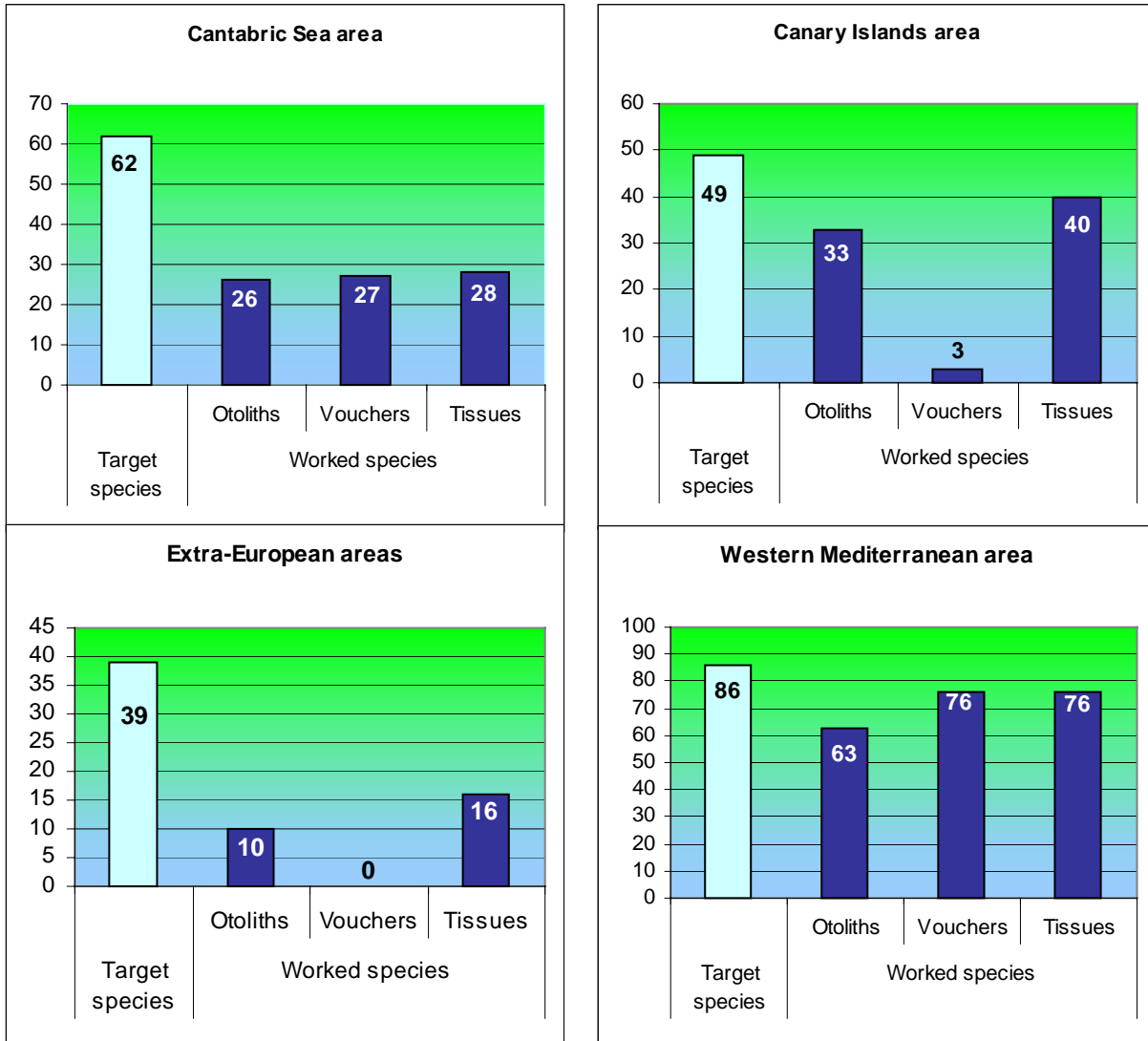
- Canary Islands (CI):
 - Number of targeted species: 49
 - Number of species sampled & taxonomically ID: 44
 - Species sampled for haplotyping: *P.erythrinus*, *P.pagrus*, *M.surmulletus* & *M.merluccius*
 - Sampling sites: ICCM Labs., Telde, Las Palmas, Spain.

- Extra-European Seas (EE):
 - Number of targeted species: 39
 - Number of species sampled & taxonomically ID: 17
 - Geographical origin of specimens: Eastern Atlantic: Morocco-Saharan Bank, Mauritania-Senegal, Guinea Conakry, Namibia-South Africa. North-Western Indian Ocean. South-Eastern Pacific: Chile
 - Sampling site: ICCM Labs., Telde, Las Palmas, Spain.



3.5. S. Jiménez continued with the work plan progress of TFMC. The results are showed in these figures that were presented at the meeting:

Fishtrace collections status at the Tenerife Museum of Natural Sciences



3.6. V. Vérrez-Bagnis exposed the results obtained by IFREMER in the different tasks that they are in charge:

Sampling and Taxonomy (WP2-WP4):

39 complete species (5/5 specimen)
5 incomplete species (1,3 or 4 specimen/5)
12 non-sampled species (among them *Thunnus alalunga* and *Thunnus thynnus*)
1 species added to the list (*Belone belone*)

Among the Triglidae, 4 species have been sampled:

Chelidonichthys cuculus
Chelidonichthys gurnadus
Chelidonichthys capensis (described in South Africa)
Chelidonichthys lucernus

For the haplotyping task (WP4) these following species are not yet sampled:

Micromesistius poutassou
Pagellus erythrinus (rare in CB zone)
Pagrus pagrus (rare in CB zone)

Remaining sampling will be done during research Ifremer cruises and by MNHN in Concarneau port.

Molecular Genetic Identification (WP3):

For the rhodopsin gene, sequence is done on :

11 species (2 specimen)
4 species (1 specimen)

For the cytochrome b gene, sequence is done on:

5 species (2 specimen)
3 species (1 specimen)

No problem to amplify these 2 genes but some difficulties to sequence the cyt b gene.

Validation in Database and Validation flow at regional level (WP7):

2 specimen of *Engraulis encrasicolus* have been put in the database (amplification condition table and genetic information). The GenBank number for rhodopsin gene and the DNA collection number have not been filled.

3.7. H. van Pelt informed all partners about the progress of RIVO in:

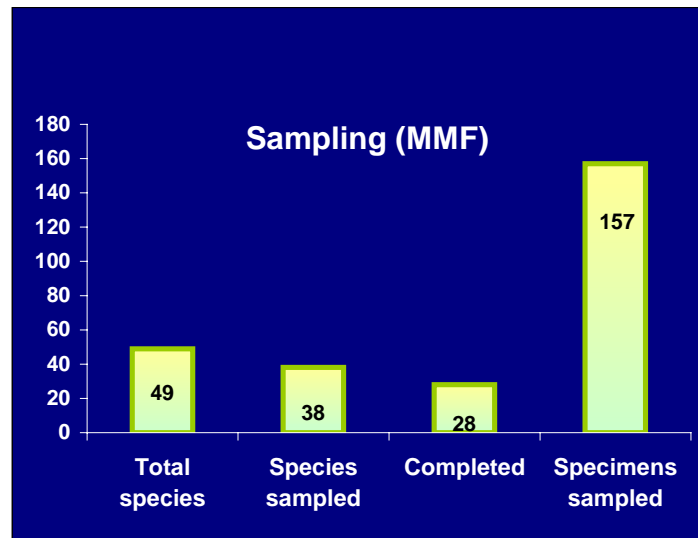
WP2:

38 species collected from the list (total 48)
Last 10 species difficult to collect, due to changes in populations in the Northsea.
Rivo has collected already alternatives, but will continue to find the 10 from the list.

WP3:

CytB sequenced from 18 species
Problems with Rhodopsin PCR, no sequences up to now.

3.8. M. Biscoito continued with the results obtained by his group on WP2 that are going on schedule:



- About 50% of samples are already in collection.
- Samples from landing places have not an exact GIS.
- IMAR has found problems with sampling of some species.
- IMAR has found problems with the interface of Database in the way of placing the data.

3.9. G. Krey showed the current status of NAGREF on WP2 and WP4:

53 species fully collected (5+ specimen from each)

5 incomplete species

2 missing species

Total number of specimens collected: 284

Measurements, counts and otoliths have been taken from 30 species.

Haplotyping (WP4): all species collected: *Merluccius merluccius*, *Micromesistius poutassou*, *Mullus surmuletus*, *Pagrus pagrus*, *Pagrus erythrinus* and *Solea vulgaris*.

Concerning WP3, NAGREF has obtained 11 completed sequences of Cytochrome B of two individuals and 5 of one individual of the same species. In relation to Rhodopsin gene, NAGREF has obtained 20 completed sequences of two individuals and 3 of one individual of the same species to date. They also have completed 6 definitive sequences information of both genes, belonging to these species sampled at geographical area G (EM):

Mic Pou, Pom Sal, Ser Hep, Spi Mae, Spi Sma and Zeu Fab

NAGREF has added four new species to the list of targets. These four species have commercial interest in Greece.

1. *Synapturichthys kleinii*,
2. *Lithognathus mormyrus*,
3. *Diplodus puntazzo*,
4. *Oblada melanura*

They also have removed one specie from the list: *Solea senegalensis*.

On the other hand, NAGREF has found some problems amplifying and sequencing samples.
 G. Krey detached these problems as follows:

Rhodopsin sequence generally ok

ScoRho specimen 01 and 02 differ substantially (10 nucleotide substitutions)

Cytb sequence problematic in certain cases

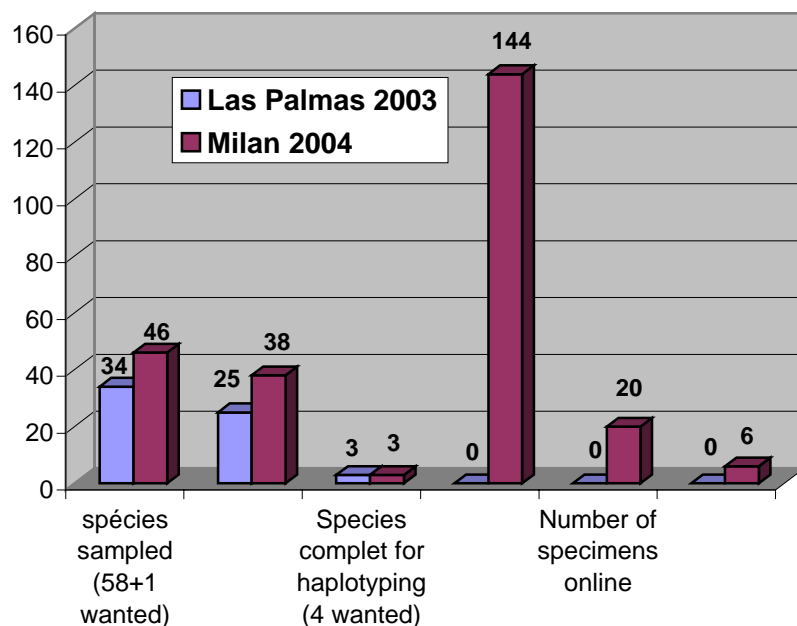
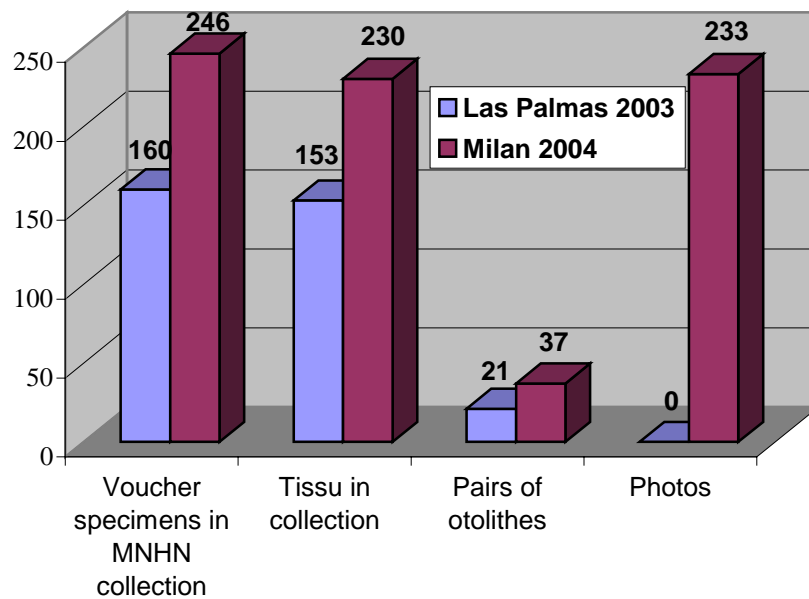
Lack of overlap between 5' and 3' (LepBos, ScoPor, MerMer)

Different 5' sequences among specimen 01 and 02 (LepBos, LopBud, LopPis, PhyBle)

Different 3' sequences among specimen 01 and 02 (MerMer, TriMin, MulBar)

Sequencing problems with MulBar-EM-02 (repeatedly produces poor sequence)

3.10. P. Pruvost exposed the progress of MNHN on WP2. The results are shown on these figures:



P. Pruvost presented a list of species indicating the probability (easy-difficult) to obtain some of them in the geographical areas in charge of MNHN. He continued showing several photos taken to the specimens collected, explaining the way to include pictures on database. They have found problems with collecting some samples. Finally, P. Pruvost enumerated the next tasks that MNHN follows on schedule:

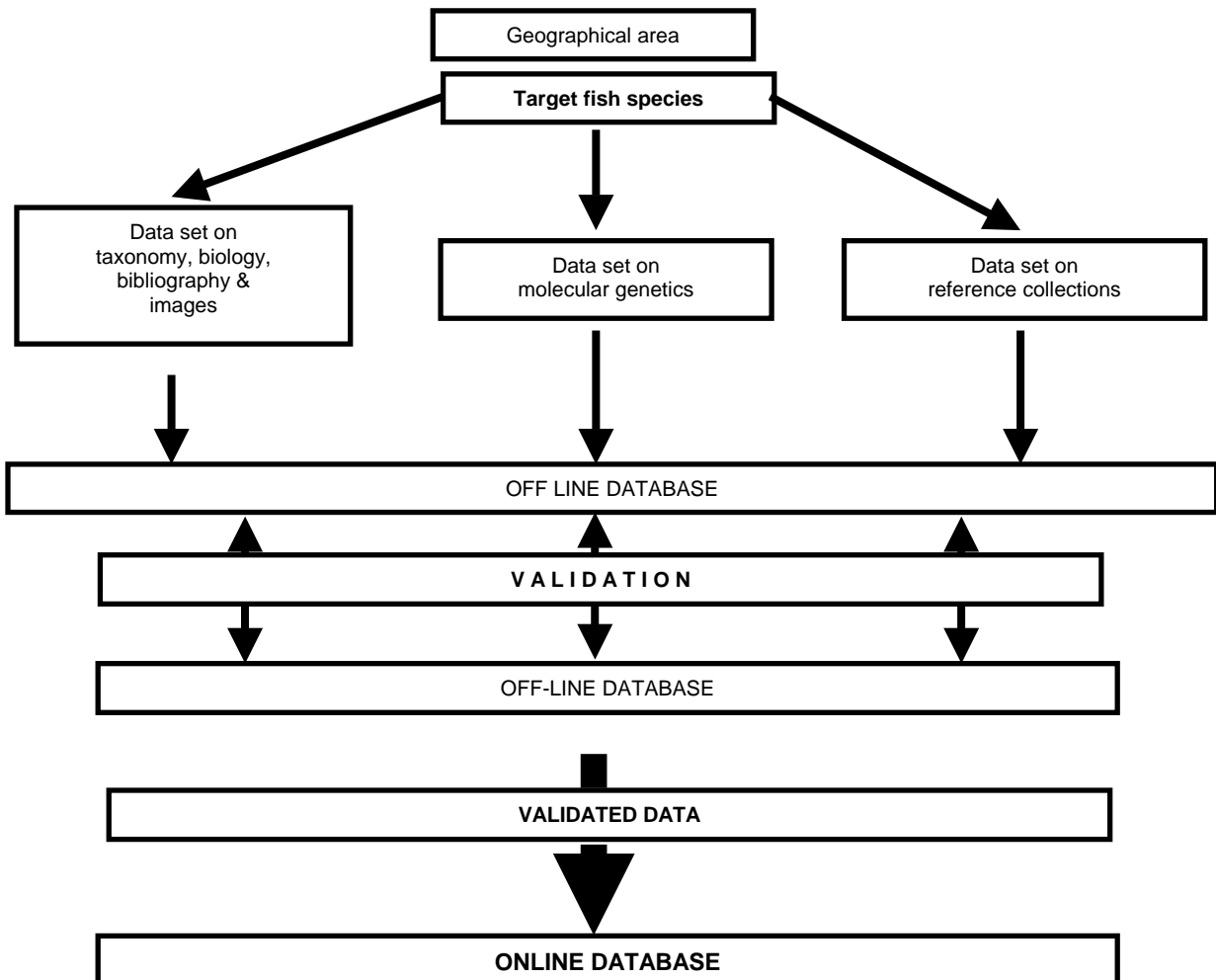
- Finish the voucher sampling
- Obtain otoliths and constitute the collection reference
- Integrate in the MNHN collection the vouchers from the other partner

3.11. Database Status by P. Carreau: Before to start speaking about the results obtained after the Database checking exercise, P. Carreau pointed some aspects of the Database:

- 1) Most of the changes agreed at Las Palmas meeting are made.
- 2) We have produce more information to fill in fields on web page after Las Palmas meeting.
- 3) A new FTP account is working now.

4. Data validation & database check

4.1. J.A. González from ICCM, group responsible of WP7, explained in his presentation the procedures to validate the data collected for each task of the project. He started regarding the main objective of WP7: Data compiled in the database are being generated by different teams of taxonomists, geneticists and curators. This important amount of valuable information requires to be checked for reliability and standardisation among the different data sets and the various groups involved. For this task, J.A. González showed the following flow of procedures:



J.A. González defined the main research tasks allocated to each Partner as follows:

Sampling-Taxonomy	NRM, ICCM, RIVO, IMAR, IFREMER, NAGREF & MNHN
Molecular Genetics	UCM, NRM, RIVO, IFREMER & NAGREF
Reference Collections	NRM, IMAR, TFMC & MNHN

Each group is responsible for checking data generated by itself in order to detect possible errors or missing data.

Each **genetic** group is responsible for detecting possible errors related to the species identification:

- a) both samples from each putative species are different,
- b) species pretended to be the same are different, and
- c) samples supposed to be different belongs to a unique species.

Once entered in the off-line DB, sampled species have to be assigned to partners involved in order to proceed to the next step on the validation process. Two questions raised from this matter:

What's a good criterion for the assignment of species to partners?

How do we decide how to share the validation work?

ICCM proposed that the geographic proximity or familiarity with species as two alternative criteria for distribute species for validation. Also it was mentioned that the partners should be more familiarised with their own regional species for validation purposes of data sets on taxonomy, description, biology, images and bibliography. J.A. González proposed that:

- Species sampled from a unique geographic area will be assigned to the partner responsible for sampling in this area.
- Species sampled from two areas will be shared on a fifty-fifty basis between the two partners involved.
- Species sampled from three areas will be shared on a third basis among the three partners involved.
- As a result, each partner involved will deal with around 35 species.

Results of the assignment of species to partners are explained in this table:

	ICCM	IMAR	NAGREF	IFREMER	NRM	All partners
1 Area	58	13	5	5	10	91
2 Areas	11	9	6	5	18	49
3 Areas	0	10	5	14	1	30
4 Areas	0	4	9	0	3	16
5 Areas	0	0	9	3	0	12
6 Areas	0	0	0	5	3	8
7 Areas	0	0	0	2	0	2
TOTALS	69	36	34	34	35	208

Validation tasks to be done with each one of the assigned species proposed by J.A. González:

- Check carefully the information entered by the involved partners in the DB related to the specimen table ("FT Specimen Form"): biometric & morphometric data, specific photos.
- Check carefully the information entered by the involved partners in the DB regional fields: common names, biology, common & legal size, interest to fisheries, forms of utilization, etc.
- Choose representative specific photos & otolith images.
- Fill the appropriate information in the DB common fields: visual examination, diagnosis, maximum size and geographic location.
- A protocol defining format for contents and sequence of information will be provided for standardisation among the different data sets and the various groups involved.

Validation tasks in **Taxonomy and Biology**:

WHAT	HOW
<ol style="list-style-type: none"> 1. Taxonomic classification 2. Biometric & morphometric data 3. Specific photos and otolith images 4. Regional information 	<ol style="list-style-type: none"> 1. According to FishBase 2. Comparing minimum & maximum values with those from main specific references (Unesco & FAO Catalogues and others) 3. Comparing taken images among them and with those contained in FB & literature. Selecting representative specific images 4. Checking data on biology, sizes, interest to fisheries and forms of utilization

Fill Info on **Taxonomy and Biology** in Common Fields:

WHAT	HOW
<ol style="list-style-type: none"> 1. Visual examination 2. Diagnosis 3. Maximum size 4. Biology 5. Geographic location 	<ol style="list-style-type: none"> 1. Entering body shape, mouth, fins and scales characteristics, and colour pattern 2. Entering body shape, mouth char., rays of dorsal, anal, pectoral & pelvic fins, caudal region, lateral line scales, gillrakers, vertebrae and colour 3. Entering the maximum length and maximum weight when relevant 4. Entering sexuality, reproduction size, spawning season and food 5. Entering climate, latitudinal range, world distribution and habitat (environment, substrata and overall depth range).

Information sources: Specimen Forms, Unesco & FAO Catalogues, FishBase & Others

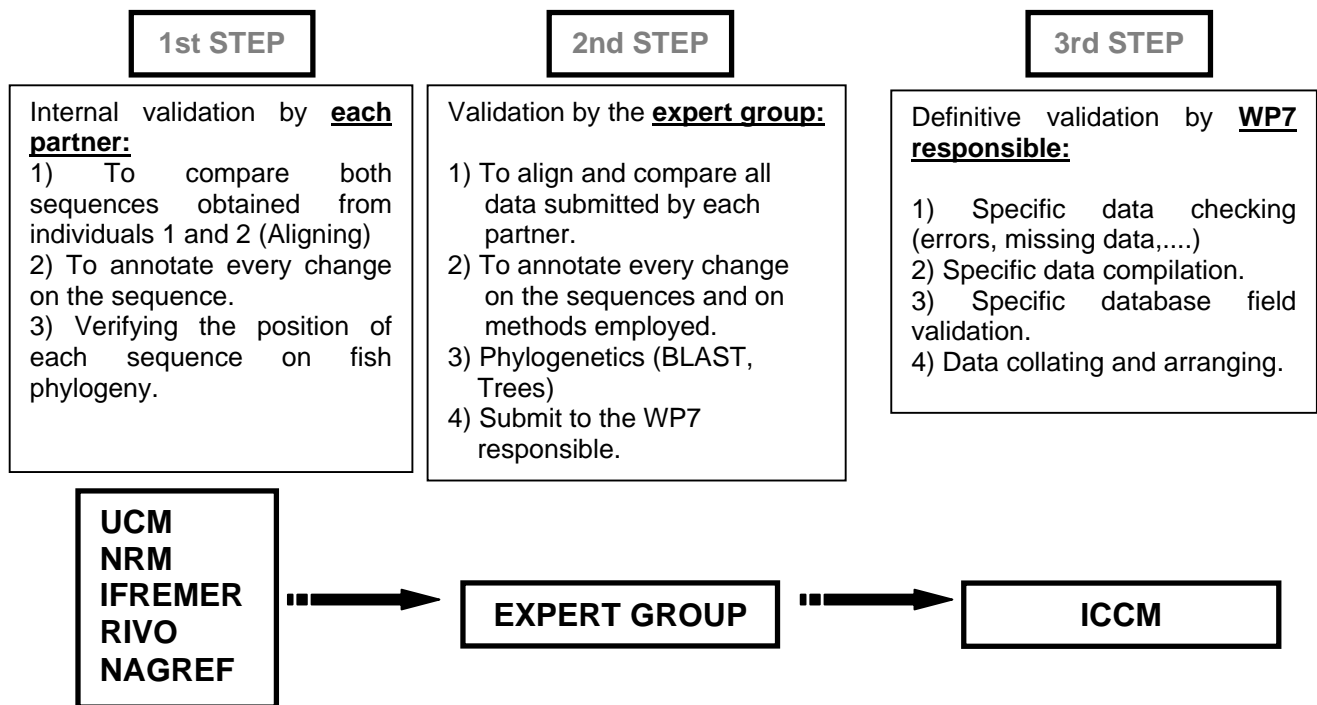
Validation tasks in **Reference Collections:**

Check carefully errors: missing information and registering data entered by the involved partners in the DB for every biological collection (vouchers, muscular tissues, otoliths and DNA)

Validation tasks in **Molecular Genetics:**

WHAT	HOW
To be established by groups involved in molecular genetics and haplotyping	To be established by groups involved in molecular genetics and haplotyping

4.2. Genetic Data Validation. J.M. Bautista, from UCM, proposed the following flow of procedures for the validation of genetic data obtained:



4.3. Quality check and exercise results by P. Carreau:

Why?

- Solve last software issues and database check
- Better to detect problems now than after 200 inputs
- Be certain everybody enter data the same way:
 - Same data format
 - Same meaning
- Be sure everybody understand the instructions !
 - Quite complex exercise, need a little training

- 62 specimens entered
- 12 with genetic completed
- 51 samples including 20 filled with -2
- 34 pictures declaration (with 34 photos loaded on the FTP site)
- Finally -> specimen incl genetic + picture less than 5

JRC pointed out minor errors in the data entered during the quality exercise especially concerning geographic location. IMAR made some comments to improve the data input interface.

A great thank you to those who did “mistakes”:

- Due most of time to misunderstanding -> need clarification
- We won't do them next time

Why didn't each partner manage to complete the 2*5 specimens?

- Lack of information?
- Lack of time?

Read instructions: One of the partner made put 1 and 2 instead of -1 and -2

- 1 for “I will never get the data”
- 2 for “Data will be enter later”

Quality check results (Genetic):

DNA format: One of the partner entered the DNA sequence with carriage return (return to next line) every 80 characters. Recommendation and discussion item. Find an agreement on DNA sequence format. All DNA sequences must be enter the same way.

Make list of DNA quality criteria and find the way to improve the interface to check it. Is it possible and should we had a system to check the DNA sequence length?

Quality check results (GPS position):

- GPS position seem not correct
- Most of them are located on continents

Recommendation and discussion item:

We should do a real effort to get fish location correctly.

-It is a project purpose to establish relation between taxonomy, genetic but also geographical location

-Required to do spatial analysis/statistics (link haplotyping) later
Abandon market position

-Location on continent

-Finally not accurate information for fish location

Use ICES (center location for each ICES area) or other system

Quality check results (Photos):

During the exercise all photos were not entered for all specimens

-Probably required a lot of additional effort

-Some specimen could have been destroyed/degraded (after sampling)

Several photo were entered for reference voucher/otolith

Recommendation and discussion item:

There must be only one photo for reference voucher file and reference otolith

Voucher name should be the id Fishtrace code only

Find agreement to establish rule for otolith photo naming

-Establish rules for otolith names as id_fishtracecode_oto.jpg

-What is a reference otolith (left or right)

During discussion try to review the photo rules (format etc..).

Quality check results (“-2”):

Avoid “-2” (too much -2 during quality check)

Recommendation and discussion item:

Advise: be sure to get all the information before entering the data.

- It is very effort consuming to complete information because you should do 2 additional operations each time: 1) search the file and 2) locate the field.

On this point there are problems with the “collection institute” fields.

- Is it possible to have them before entering the data? In the exercise nobody entered one.
- We should discuss on the way to improve the procedure.

Don't put -1 or -2 without good reason for optional data (you can recognize them by (o))

- I noticed some partner put -2 for the picture comment. What does it mean? Do you want to put a comment later?

Quality check results (Miscellaneous):

Problem with field definition.

- Example of SL_preserved – What is the final decision?
- Don't select “Select below” during a selection.

5. Creation of interim working groups according to type of data

- Genetic Group: formed by G. Krey, H. van Pelt, V. Verrez, J.M. Bautista and M. Norén.
- Specific Common fields in Database (Taxonomy and Biology group): formed by M. Biscoito (IMAR), P. Pruvost (MNHN) and J.A. González (ICCM).
- Specific Regional Fields in Database (Taxonomy and Biology group): formed by J.A. González (ICCM), S. Jiménez (TFMC) and I.J. Lozano (ICCM).

2nd DAY

6. Communal Summary from the working groups

6.1. Validation of Genetic Data. After the discussion, the group raised several conclusions:

- It was agreed to add a new field to the database under “Amplification Condition” form. This field will be called “Amplification Remarks” and will indicate:
 - Type and trademark of Taq DNA polymerase used,
 - MgCl₂ concentration in mM,
 - and any other amplification variable.
- The English checking of the whole database is needed and will be carry out by a person with English as mother language. Faye Taylor (UCM) was proposed for this purpose.
- It is necessary to draw up instructions to fill in the database with examples, and explaining the abbreviations.
- There is a new definition of PCR cycle as follows:
 - Each amplification step is composed by: Temperature (in °C) followed by the required Time (in seconds): *temp – time*. Example: **94-30**
 - Each amplification step must be separated to the next step by a slash: /. Example: **94-30 / 50-30**
 - The steps repeated must be included between parenthesis and followed by the number of cycles. Example: **(94-30/50-30/72-30) x40**: *these three steps must be repeated forty times.*
 - Example of a complete PCR cycle: **95-240 / (94-30/50-30/72-30)x40 / 72-420**

- To submit sequences to GenBank, Sequin (software) will be used with common fields. UCM will circulate a proposal of these fields to all partners.
- The format of the sequences must be "only text" (seq-file.txt) no wrapped text.
- UCM will circulate the sequences for standard length and phylogenetics analysis for validation purposes.
- In the database it is not necessary a control on sequence identity, but on number of characters.
- Delete present data added to database and start from zero now.
- Recommended software for phylogenetics: XAC (faster). UCM will check it for consistent phylogenetics reconstructions with our fish sequences.
- UCM will draw up protocols to validate genetic data and circulate.
- The distribution of species will be defined by ICCM: Important to send soon the final sampling list.

6.2. Validation of Taxonomy and Regional Info Data.

It was delivered a first draft of guidelines for validation purposes by M. Biscoito (IMAR), P. Pruvost (MNHN), S. Jiménez (TFMC), I.J. Lozano and J.A. González (ICCM). See annexed document number 1: *WP7 Data Validation in Database. Guidelines for Validation Purposes*. The final version of this document will be delivered in a few weeks by the group in charge.

The working group composed by M. Biscoito (IMAR), P. Pruvost (MNHN) and J.A. González (ICCM) have created the annexed document number 2: *Additions and Corrections to FishTrace Database*.

The working group composed by J.A. González (ICCM), S. Jiménez (TFMC) and I.J. Lozano (ICCM) have created the annexed document number 3: *Additions to FishTrace Database. Specific Regional Info*.

8. Data result presentation

8.1. GIS capability

- GPS position seem not correct (due to the collection of landing specimens)
- Most of them are located on continents

8.2. Web site user needs

The discussion was about future public web interface. JRC proposed to think how to deliver spatial information (maps) and partners thought about displaying haplotyping distribution. UCM would like to include specific tools for DNA sequence analysis.

It was decided to use an existing web site Pescabase (www.pescabase.org) as model to develop a first version of the Fishtrace public web site.

9. General conclusion

This meeting has been very useful for data validation point of view especially genetics. ICCM made an overview about validation methods and the meeting was split in two groups "Genetic" and "Taxonomy + image + sampling" for more accurate discussion. Genetic group made a list of validation process that will be described in a document. The implementation of a tool to check if DNA sequence is complete (by verifying the length) when it is entered in the database was definitely accepted.

The Taxonomy and Regional Info group will also make a document to describe the way to complete the common fields relative to species. It is expected to enter not only information concerning individual specimens but also at the whole specie level. This required making an abstract of the information collected. The group proposed to add around 10 new fields for species description and 20

new items for regional information. There was a discussion of the amount of work necessary to achieve this task.

9.1. Next Fishtrace meeting

Depending of data entry progress, the next Fishtrace meeting date in Paris will be fixed from September 1st (In this case it will coupled with the Fishbase meeting) to end 2004.

9.2. Conclusions and Project Schedules:

Collecting, sampling and taxonomy tasks seem to work properly. Some partners noticed that the missing species are the harder to find and collect.

Concerning genetic partners RIVO, NRM and IFREMER have problems with Rhodopsin sequencing. UCM offers any kind of help needed to solve the problem. The group agreed in some possibilities to solve problems on Rhodopsin amplification:

- Interchange primers aliquots
- Receive DNA samples in Madrid for amplification
- Primer of Cytochrome B amplification to IFREMER for sequencing

Project is on schedule but it is urgent to begin to fill the database. The next dates to have in mind to keep the project on the path are the following:

- May 15th: The database for data input should be ready.
- End May: Ultimate database check (1 week).
- June 15th: A reasonable amount of data should be entered including species to allow a first sight of the public interface.
- End of summer (September): A first version of the public interface should be ready.

ANNEXED DOCUMENTS:

- 1) *WP7 Data Validation in Database. Guidelines for Validation Purposes*
- 2) *Additions and Corrections to FishTrace Database*
- 3) *Additions to FishTrace Database. Specific Regional Info*

WP 7 DATA VALIDATION IN DATABASE
GUIDELINES FOR VALIDATION PURPOSES

Responsible partner: ICCM (J.A. Gonzalez)

**Working Group: M. Biscoito (IMAR), P. Pruvost (MNHN), S. Jiménez (TFMC),
I.J. Lozano and J.A. González (ICCM)**

1. Partners involved

Data on Taxonomy and Sampling

NRM, ICCM-TFMC, IMAR, IFREMER-MNHN, NAGREF

Data on Biological Reference Collections

NRM, IMAR, MNHN, TFMC

Data on Molecular Genetics

UCM, NRM, RIVO, IFREMER, NAGREF

2. Objectives

Data compiled in the database are being generated by different teams of taxonomists, geneticists and curators. This important amount of valuable information requires to be checked for reliability and standardisation among the different data sets and the various groups involved.

3. The validation process

3.1. First step: early validation

Each group is responsible for checking data generated by itself in order to detect possible errors or missing data

Each genetic group is responsible for detecting possible errors related to the species identification:

- a) both samples from each putative species are different,
- b) species pretended to be the same are different, and
- c) samples supposed to be different belongs to a unique species

3.2. Second step: species assignment

Once entered in the off-line DB, sampled species have to be assigned to partners involved in order to proceed to the next step on the validation process.

The geographic proximity or familiarity criteria was accorded as a good criterion for the assignment of species to partners.

Explanation: Partners should be more familiarised with their own regional species for validation purposes of data sets on taxonomy, description, biology, images and bibliography.

(Nearly) final lists of assigned species to partners for validation purposes on taxonomy/sampling data.

PLEASE, SEND YOUR FINAL SPECIES LIST TO PARTNER ICCM

3.3. Third step: regional information

Enter relevant regional information for the species related to common names, biology, habitat, sizes characteristics, interest to fisheries, commercialisation and remarks.

Label “Common names”

Field “Common names”.

Fill the species common name(s) used in this geographical region.

Ex: Tuna, atún.

Label “Biology”

Field “Sexuality”

Fill the sexual type of the species in this geographical region by choosing from gonochoric, protogynic hermaphrodite, protoandric hermaphrodite, potentially hermaphrodite, etc. according to relevant regional bibliography.

Ex: Protogynic hermaphrodite (Lozano, 2003).

Field “Reproduction size”

First maturity length (in cm) in this geographical region.

Ex: 33 cm TL.

Ex: 17 cm FL (females), 23 cm FL (males) (González, 2004).

Field “Spawning season”

Fill the reproduction period (in a month range) of this species in this geographical region.

Ex: Abril-September (Jiménez, 2001).

Field “Food”

Fill the feeding habits and main preys categories for the species in this geographical region.

Ex: Carnivorous: crustaceans, cephalopods and fishes.

Label “Habitat”

Field “Depth range”

Fill in the known minimum and maximum depth for the species (done as a range in m) in this geographical region.

Ex: 100-400 m.

Label “Size characteristics”

Field “Common sizes”

Fill the commercial sizes regularly present in the markets (done as a range in cm) in this geographical region.

Ex: 20-40 cm TL.

Field “Capture legal size”

Fill the minimum size of capture in according to regional regulations (done in cm).

Ex: 20 cm TL.

Label “Interest to fisheries”**Field “Fisheries importance”**

Fill the species importance for the regional fisheries, aquaculture and gamefish in this geographical region.

Ex: Highly commercial. Industrial and artisanal fisheries.

Field “Fishing gears”

Fill the fishing methods used for the species in this geographical region.

Ex: Bottom trawl and trammelnets.

Label “Commercialisation”**Field “Forms of utilization”**

Fill the forms of use (whole or transformed product) for the species in this geographical region.

Ex: Marketed fresh, refrigerated, frozen and canned. Generally used in fillets or slices; eggs are also commercialised.

Field “Cooking usage”

Fill the culinary usages for the species in this geographical region.

Ex: Boiled, fried, baked, broiled.

?Field “Misidentifications”

Fill the main market misidentifications for the species in this geographical region.

Ex: It can be misidentified with other tuna fishes.

Field “Substitutions”

Fill the known market substitutions for the species in this geographical region.

Ex: Its fillets sometimes are substituted by those pelagic sharks.

Label “Conservation status and Threats”**Field “Conservation status and threats”**

Fill the conservation status and degree of threaten in this geographical region.

Ex: Vulnerable; protected species according to regional regulations. Main threats are: illegal spearing fishery and gillnets.

Label “Remarks”**Field “Remarks”**

Fill other relevant information for the species in this geographical region.

Ex: Ongoing aquaculture experiences in floating cages.

3.4. Forth step: general information

ADDITIONS AND CORRECTIONS TO FISHTRACE DATABASE

*The main aim of this research and technological development project is to undertake the **genetic identification and characterisation** of around 183 European marine fish species to guarantee the source and authenticity of fish and derived products.*

COMMON FIELDS IN THE DATABASE SPECIES TABLE

- **Bio_Visual_Exam** (max. xxx characters)
- **Diagnosis** (max. xxx characters)
- **Similar species (as separate field or inside diagnosis?)** (max. xxx characters)
- **Biology** (max. xxx characters)
- **Distribution and fisheries interest** (max. xxx characters)
- **Conservation status** (max. xxx characters)
- **Sources of information** (max. xxx characters)

DESCRIPTION OF CONTENTS AND FILLING GUIDELINES:

Bio_Visual_Exam.

In this field there will be a short popular description of the main external characters of the species, with particular emphasis on body shape, number, shape and position of fins, general colour pattern and any other striking external characteristic. Also bear in mind differences between sexes and common size.

Diagnosis

In this field there will be a scientific description of the diagnostic characters for the species. Please follow the following topics (when applicable) by the this order:

- a) Body, mouth shape and caudal region description;
- b) Fins (Dorsal, Anal, Pectoral, Pelvics);
- c) Lateral line (presence or absence, scales, scutes, etc.;
- d) Gillrakers in first gill arch;
- e) Vertebrae;
- f) Colour;
- g) Size.

Similar species

In this field include any similar species which can be misidentified, mentioning separating characters.

Biology

In this field give a description of the more relevant biological information on the species:

- a) Habitat (environment, substrata,depth range) ;
- b) Behaviour (schools, migrations,etc);
- c) Sexuality (gonochorism, hermafroditism);
- d) Reproduction (maturity size, spawning season);
- e) Food.

Distribution and fisheries interest

In this field give geographical information for the species:

- a) Climatic or biogeographical pattern;
- b) Latitudinal range; Distribution (Europe and elsewhere)
- c) Main fisheries.

Conservation status

In this field list inclusion in Red Books, CITES...

Sources of information

List relevant sources of information used: FNAM, FAO, ETI, FishBase, PescaBase, etc.

PRACTICAL EXAMPLE:

Seriola fasciata (Bloch, 1793)

Bio_Visual_Exam.

Moderately large amber jack; Body fusiform; upper profile of head convex; two dorsal fins, the first with 8 spines; anal fin much shorter than second dorsal; caudal fin forked; adults pinkish; juveniles with dark vertical bars.

Moderately large amber jack. Body fusiform. One spinous and one soft dorsal fin. Tail forked. Adults pinkish; juveniles with dark vertical bars.

Diagnosis

Body fusiform; upper profile of head convex; end of upper jaw relatively slender; grooves present on caudal peduncle; D VIII + I + 28-33; A II + I + 17-20; lateral line present and without scutes; gill rakers in first gill arch 23-26. Colour: adults dark (pinkish or violet) dorsally, sides lighter, and belly white or silvery. Juveniles smaller than about 20 cm FL with interradiated membranes of soft dorsal and anal fins lightly pigmented; juveniles about 4-26 cm FL with 7 dark body bars, irregular and broken, third to seventh extending onto the membranes of the second dorsal and anal fins, and the eighth bar, small and dark, at end of caudal peduncle. Size: to 67.5 cm FL and 4.6 Kg.

OR...

Morphology: Body fusiform; upper profile of head convex; end of upper jaw relatively slender; grooves present on caudal peduncle.

Fins (Dorsal, Anal, Pectoral, Pelvics): D VIII + I + 28-33; A II + I + 17-20

Lateral line: present, without scutes

Gillrakers in first gill arch: 23-26

Vertebrae: [Not relevant.]

Colour: adults dark (pinkish or violet) dorsally, sides lighter, and belly white or silvery. Juveniles smaller than about 20 cm FL with interradiated membranes of soft dorsal and anal fins lightly pigmented; juveniles about 4-26 cm FL with 7 dark body bars, irregular and broken, third to seventh extending onto the membranes of the second dorsal and anal fins, and the eighth bar, small and dark, at end of caudal peduncle

Size: to 67.5 cm FL and 4.6 Kg.

ADDITIONS TO FISHTRACE DATABASE – SPECIFIC REGIONAL INFO

Working Group:
J.A. González (ICCM), I.J. Lozano (ICCM) and S. Jiménez (TFMC)

The main aim of this part of the Database is to compile relevant regional information for the Fishtrace species, related to common names, biology, habitat, size characteristics, interest to fisheries, conservation status and threats, and others

Fields to be included in the Regional Info Table in DB:

Field: Species

Field: Area

Field: Label (Common names, Biology, Habitat, Size characteristics, Interest to fisheries, Commercialisation, Conservation status and Threats, and Others).

Field: Topic (for each Label, all the topics included in the box, ex., **Label** Interest to fisheries; **Topics:** Fisheries importance and Fishing gears)

Field: Remarks

Field: References

Label “Common names”

Topic “Common names”.

Fill the species common name(s) used in this geographical region.

Ex: Tuna, atún.

Label “Biology”

Topic “Sexuality”

Fill the sexual type of the species in this geographical region by choosing from gonochoric, protogynic hermaphrodite, protoandric hermaphrodite, potentially hermaphrodite, etc. according to relevant regional bibliography.

Ex: Protogynic hermaphrodite (Lozano, 2003).

Topic “Reproduction size”

First maturity length (in cm) in this geographical region.

Ex: 33 cm TL.

Ex: 17 cm FL (females), 23 cm FL (males) (González, 2004).

Topic “Spawning season”

Fill the reproduction period (in a month range) of this species in this geographical region.

Ex: Abril-September (Jiménez, 2001).

Topic “Food”

Fill the feeding habits and main preys categories for the species in this geographical region.

Ex: Carnivorous: crustaceans, cephalopods and fishes.

Label “Habitat”**Topic “Depth range”**

Fill in the known minimum and maximum depth for the species (done as a range in m) in this geographical region.

Ex: 100-400 m.

Label “Size characteristics”**Topic “Common sizes”**

Fill the commercial sizes regularly present in the markets (done as a range in cm) in this geographical region.

Ex: 20-40 cm TL.

Topic “Capture legal size”

Fill the minimum size of capture in according to regional regulations (done in cm).

Ex: 20 cm TL.

Label “Interest to fisheries”**Topic “Fisheries importance”**

Fill the species importance for the regional fisheries, aquaculture and game fishing in this geographical region.

Ex: Highly commercial. Industrial and artisanal fisheries.

Topic “Fishing gears”

Fill the fishing methods used for the species in this geographical region.

Ex: Bottom trawl and trammelnets.

Label “Commercialisation”**Topic “Forms of utilization”**

Fill the forms of use (whole or transformed product) for the species in this geographical region.

Ex: Marketed fresh, refrigerated, frozen and canned. Generally used in fillets or slices; eggs are also commercialised.

Topic “Cooking usage”

Fill the culinary usages for the species in this geographical region.

Ex: Boiled, fried, baked, broiled.

Topic “Misidentifications”

Fill the main market misidentifications for the species in this geographical region.

Ex: It can be misidentified with other tuna fishes, notably with *Thunnus albacares*. Adults of *T. albacares* are laterally more compressed, have second dorsal and anal fins very long with narrow black edged finlets, and its belly is often crossed by around 20 discontinuous almost vertical lines.

Topic “Substitutions”

Fill the known market substitutions for the species in this geographical region.

Ex: Its fillets sometimes are substituted by those from pelagic sharks.

Label “Conservation status and Threats”**Topic “Conservation status and threats”**

Fill the conservation status and degree of threat in this geographical region.

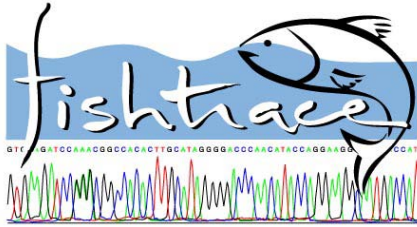
Ex: Vulnerable; protected species according to regional regulations. Main threats are: illegal spearing fishery and gillnets.

Label “Others”**Topic “Others”**

Fill other relevant information for the species in this geographical region.

Ex: Ongoing aquaculture experiences in floating cages.

Annex VI: 3rd FishTrace Plenary Meeting. Paris, 25-27 November, 2004.



FishTrace Annual Meeting

November 25 – 27, 2004.

Paris, France.

Minutes from the FishTrace Annual Meeting on Project Status, Next Actions to be taken and Project Prospects

Objectives:

- To show and reflect on the progress of all workpackages and tasks during this period of the project.
- To gather information for next Periodic Report to the Commission.
- To define Project status in relationship with scheduled plan.
- To define main aspects of the FishTrace Database Web Interface and links to other Databases.
- To define Data Validation aspects, and data submission to Databases.
- To solve problems encountered at different stages of the different workpackages and connections.
- To discuss about next actions to be taken, project prospects and new proposals in EC calls in FP6 and FP7.

Venue:

Held at Muséum National d'Histoire Naturelle. 43 rue Cuvier, Paris.

Dates:

November 25 to 27, 2004.

Participant list:

MNHN: (Host)

Guy Duhamel
Patrice Pruvost
Samuel Iglésias
Olivier Mouchel

IFREMER:

Monique Etienne
Véronique Verrez-Bagnis
Marc Jérôme

NAGREF:

Alexis Tsangridis
Grigorios Krey
Panos Leontarakis

IMAR:

Manuel Biscoito
Mafalda Freitas
Rosa Pestana

ICCM:

José A. González
Ignacio J. Lozano

UCM:

José M. Bautista
Rafael González

EC-JRC-IPSC:

Philippe Carreau

RIVO:

Hilde van Pelt

TFMC:

Sebastián Jiménez

FishBase (Invited):

Nicolas Bailly
Ma. Lourdes Palomares

Agenda:**1st DAY**

- | | |
|--|---------------|
| 1. Welcome to FishTracers | P. Pruvost |
| 2. Meeting introduction | J.M. Bautista |
| 3. Workplan progress by partner | |
| 3.1. Partner no. 1 (UCM) | R. González |
| 3.2. Partner no. 2 (EC-JRC-IPSC) | P. Carreau |
| 3.3. Partner no. 3 (NRM) | |
| 3.4. Partner no. 4 (ICCM) | I. Lozano |
| 3.5. Partner no. 5 (IFREMER) | V. Verrez |
| 3.6. Partner no. 6 (RIVO) | H. van Pelt |
| 3.7. Partner no. 7 (IMAR) | R. Pestana |
| 3.8. Partner no. 8 (TFMC) | S. Jiménez |
| 3.9. Partner no. 9 (NAGREF) | G. Krey |
| 3.10. Partner no. 10 (MNHN) | S. Iglésias |
| 4. Workpackages progress info | |
| 4.1. WP2: Sampling and taxonomy | |
| 4.2. WP3: Molecular Genetic Identification | R. González |
| 4.3. WP4: Biogeographical Molecular Haplotyping | G. Krey |
| 4.4. WP5: Biological Reference Collections | M. Biscoito |
| 4.5. WP7: Validation in Database | J.A. González |
| 4.6. WP6: Database Development | P. Carreau |
| 4.7. WP8: Implementation of www technologies | P. Carreau |
| 5. Open discussion on Workpackages progress | |

2nd DAY

- | | |
|--|--------------------------------|
| 6. Link to FishBase | N. Bailly and Ma. L. Palomares |
| 7. Database adaptation to GBIF: Darwin Core v2 | J.M. Bautista |
| 8. Project Status Summary | J.M. Bautista |
| 9. Other Links: FishGene, Museum Collection, Databases... | P. Carreau |
| 10. Database Status Summary | P. Carreau |
| 11. Data Validation Flow and Data Status Summary | J.A. González |
| 12. Submissions of sequences to GenBank (Sequin Tool) | R. González |
| 13. Special session for geneticists: Submission of sequences to GenBank | G. Krey |

3rd DAY

- | | |
|--|---------------|
| 14. Open session: Troubleshooting on Taxonomy, Genetics, Collections, Database... | |
| 15. Web Interface | P. Carreau |
| 16. Taxonomy – Collections Group Discussion | |
| 17. Genetics Group Discussion | |
| 18. Mutual agreement on FishTrace Web Interface | |
| 19. Next Period Scientific and Financial Report | J.M. Bautista |
| 20. Next FishTrace meetings | |
| 21. WP9. Dissemination: Actions for the next period security and fisheries | M. Etienne |
| 22. Open session: Discussion on dissemination and brochure | |
| 23. Follow up of FishTrace through new proposals in EC calls in FP6 and FP7 | J.M. Bautista |
| 24. Open discussion on new proposals in EC calls in FP6 and FP7 | |

Meeting Minutes:

1st DAY

1. Welcome to FishTracers

The meeting started at 12:00 with a welcoming introduction by P. Pruvost.

2. Meeting introduction

A short overview about the progress of the project during this period by J.M. Bautista, based on data from the First Annual Report (February, 2004) and the last Progress Reports sent (September, 2004). It was proposed and subsequently agreed the following list of rapporteurs, in order to draw up the meeting minutes:

1st Day, Thursday:

12:20 – 15:30	Workplan Progress info by Partner	P. Pruvost
15:30 – 18:00	Workpackages Progress Info	J. Bautista

2nd Day, Friday:

09:00 – 11:00	Project Status Summary + Troubleshooting + Discussion	R. González
14:30 – 17:00	Data Validation + Sequence Submissions	H. van Pelt

3rd Day, Saturday:

09:00 – 10:40	Database Status + Links	R. Pestana / M. Freitas
11:00 – 13:00	Web Interface + Taxonomy Discussion + Agreement	J.A. González
11:30 – 12:30	Genetics Discussion	R. González
14:30 – 16:45	WP9. Dissemination	V. Verrez
16:45 – 18:30	Follow up of FishTrace + New Proposals	G. Krey

Minutes from all rapporteurs have been compiled and edited by R. González.

3. Workplan progress by partner

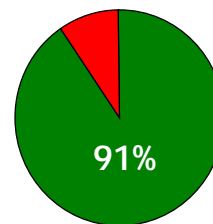
3.1. R. Gonzalez (UCM) showed the workplan progress info from UCM on Molecular Genetic Identification (WP3) and the workplan progress info on Biogeographical Molecular Haplotyping (WP4), detailing objectives achieved and results obtained during this period of the project:

WP3: Tissues samples from 260/287 species had been collected. DNA extraction from 100% of received samples 448 the total to extraction for the project is 574 tissues.

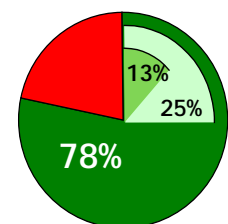
Muscle tissues received:	448 (91%)
Total DNA extractions:	448 (100%)
CytB 5' sequences:	178
CytB 3' sequences:	99 } (13%)
Rhod sequences:	142 (25%)

WP3 Species 260 (287)

Tissues 448 (574)

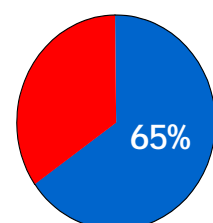
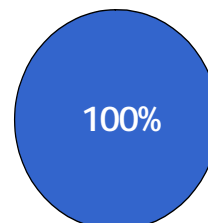


WP4 Species 6 (6)



Tissues 248 (380)

WP4: Biogeographical Molecular Haplotyping:
100% of targeted species sampled
65% tissues received



3.2. P. Carreau (EC-JRC-IPSC) presented the workplan progress info for the database. The database structure is done. A web site presentation is available on the FishTrace site. The main issue now is to have the database filled up with the information from the partners.

3.3. R. González (instead of NRM members) showed the workplan progress info from NRM, based on results from the last Progress Report sent (September 2004) and the presentation sent by M. Norén to R. González:

Sampling: 46 (of 59) species sampled
155 vouchers (entire fish in formalin)
372 tissue samples
311 photos taken

Sequencing: 64 Complete Cyt b sequences from 40 taxa
(24 taxa in duplicate = completed)
24 Complete Rhodopsin sequences from 17 taxa

Collections: No vouchers have been sent to Paris yet.
No otolith photos have been curated yet.

Problems & Solutions: We have received assurance from S. Kullander that we will get help with the otoliths and with processing/shipping of vouchers for MNHN.
We have contacted other sources for fish on the Swedish west coast.
Hopefully the sequencing problem at MSL should be solved before Christmas holidays.

3.4. I. Lozano showed the workplan progress info for ICCM on WP2:

Cantabric Sea & NW Iberian Peninsula (CS)

- Number of target species: 62
- Percentage sampled and taxonomically identified: 74%
- Species sampled for haplotyping: *M. merluccius*
- Percentage done: 17%

Western Mediterranean & Bay of Cadiz (WM)

- Number of target species: 83
- Percentage sampled and taxonomically identified: 100%
- Species sampled for haplotyping: *M. merluccius*, *M. surmuletus*, *P. erythrinus*, *P. pagrus*, *S. solea* and *M. poutassou*
- Percentage done: 100%

Canary Islands (CI)

- Number of target species: 49
- Percentage sampled and taxonomically identified: 100%
- Species sampled for haplotyping: *P. erythrinus*, *P. pagrus*, *M. surmulletus*, *M. merluccius* and *S. solea* (this species from the Saharian Bank)
- Percentage done: 100%

Extra-European Seas (EE)

- Number of target species: 39
- Percentage sampled and taxonomically identified: 100%
- Geographical origin of specimens: Eastern Atlantic (Morocco-Saharian Bank, Mauritania-Senegal, Guinea Conakry, Namibia-South Africa); North-Western Indian Ocean; South-Eastern Pacific: (Chile); North-East Atlantic (Svalbard Islands)

- Vouchers, tissues and otoliths transferred to Tenerife museum (partner TFMC): 85% done.
- Vouchers transferred to Paris and Stockholm museums (partners MNHN & NRM): 13% done (corresponding to a part from the Cantabric Sea).

3.5. V. Verrez (IFREMER) showed the workplan progress info from IFREMER:

WP2: The work had been done on 52/56 listed species + 4 « new species »: *Xiphias gladius* (swordfish 1 ind.), *Thunnus obesus* (bigeye tuna 1 ind.), *Belone belone* (garpike 5 ind.) and *Pagellus erythrinus* (common pandora 20 ind.).

4 species are not sampled from the initial list: *Salmo trutta*, *Sarda sarda*, *Mugil cephalus* and *Molva molva*.

WP3: Cyt b gene: 9 sequences in the database
» 30 complete sequences and many partial sequences

Rhodopsin gene: 18 sequences (corresponding to 9 species) in the database
» 40 complete sequences (20 species sequenced)

WP9: Short presentation at the plenary meeting of the Western European Fish Technologist's Association (WEFTA) 34th WEFTA Annual Meeting 12. – 15. September 2004 -Lübeck, Germany
M. Etienne presented FishTrace during meetings of 2 EC projects:

- Identification, classification & dissemination of organisational, technical & legal information resources dealing with traceability of fish products – Innovate actions for the fisheries sector.
- Health promoting, safe seafood of high eating quality in a consumer driven fork-to-farm concept FP6 Integrated Project (Traceability pillar).

3.6. H. van Pelt (RIVO) showed the workplan progress info from RIVO:

Species collection RIVO: - 49 species collected (original list 46 species)
 - 7 species missing
 - 10 new species added

Tissue collection: - 240 samples (01-05)
 - 60 samples for haplotyping (1-20)

DNA isolation: - 300 samples (240 +60)

Sequencing: - Rhodopsin 225
 - Cyt b 3' 240
 - Cyt b 5' 60

3.7. R. Pestana showed the workplan progress info from IMAR:

45/53 species had been sampled 35 species are completed 8 species are not found. 85 % of the work is done. 36 species are validated.

R. Pestana explained that will be necessary to be able to print the data from the data base to validate the data and that we have to discuss.

3.8. S. Jiménez showed the workplan progress info from TFMC:

The collections are completed over 64 %.

	CS	CI	WM	EE	
Vouchers	69 %	88 %	100 %	100 %	
Otoliths	68 %	82 %	100 %	97 %	
Tissues	66 %	82 %	100 %	69 %	85 %
DNA	0	0	0	0	
Average	51 %	63 %	75 %	67 %	64 %

3.9. G. Krey showed the workplan progress info from NAGREF:

Sampling, Taxonomy, Collections completion:

60 targeted species belong to 27 families and 41 genera (300 specimens in total)

55 species fully collected (90 %)

2 partly collected

2 not collected yet

290 specimens have been collected in total: taxonomically identified, photographed and preserved, measurements and counts, tissue sampling, otoliths collected in 50 species and photographed.

Database completion: Taxonomy: 41 species
Genetics: 31 species

Specimen missing from 4 species (7%): *Pegusa lascaris* (3), *Pagellus bogaraveo* (1), *Xiphias gladius* (4), *Epinephelus sp. (costae ?)* (4).

No specimen collected from 2 species (3%): *Epinephelus aeneus*, *Epinephelus marginatus*.

Additional species: *Sarpa salpa*

Molecular Genetic Id. completion:

Complete sequence for Cyt b from 33 species (2 specimen/species)

Completed sequence from a single specimen from 4 species

Total of successful Cyt b sequences = 70

Complete sequence for *rhod* from 40 species (2 specimen/species)

Completed sequence from a single specimen from 1 species

Total of successful Rhodopsin sequences = 81

Complete sequences for both Cyt b and Rhodopsin from 30 species (2 specimen/species), i.e. 50% done.

3.10. S. Iglésias showed the showed the workplan progress info for MNHN.

a. Species & specimens progress: 58 wanted in the initial list

	Wanted	Collected	
Species	58(1)	55(2)	95%
Specimens	330	301	91%

- 2 species eliminated (*Sarda sarda* & *Mugil cephalus*) because uncommon in the area.
- 1 species added (*Alosa alosa*) in replacement.
- 1 species eliminated (*Chelidonichthys capensis*) due to misidentification or taxonomic problems encountered.
- 2 species of other partners added by opportunity (*Xiphias gladius* & *Thunnus obesus*).
- 43 species are completed (5 specimens or more by species).

- 12 species uncompleted (less than 5 specimens by species).
- 3 species missing: *Atherina presbyter*, *Molva molva*; *Salmo trutta*).
- 18 specimens still waiting frozen.

b. Otoliths progress:

	Wanted	Collected	
Otoliths/species	58	19(1)	33%
Otoliths/specimens	~143	62	43%

c. Photos progress:

	Wanted	Done	
Photos/specimens	~348	295	85%
Photos/otoliths	~144	0(2)	0%

d. Environment and Species tables progress:

	Wanted	Done	
Environment tables	~95	~79	83%
Species tables	~37	5	14%

e. Voucher specimens from other institutes that will be deposited in MNHN – progress:

		Specimens waited	Specimens received
ICCM	Canary I.	?	48
RIVO	Netherlands	?	4 tanks
NRM	Sweden	?	0
NAGREF	Greece	?	0
IMAR	Madeira	?	0

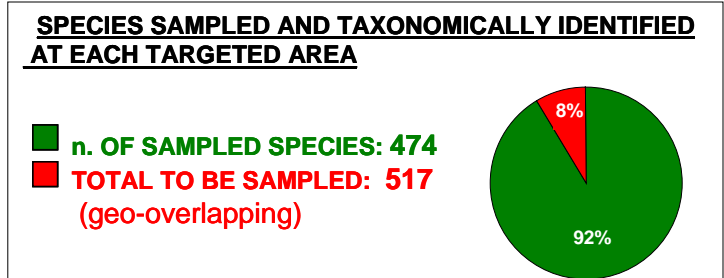
f. Haplotyping progress:

	Specimens wanted	Sp. collected
<i>Merluccius merluccius</i>	15	16
<i>Mullus surmuletus</i>	15	15
<i>Solea solea</i>	15	15
<i>Micromesitius poutassou</i>	15	20
<i>Pagellus erythrinus</i>	15	0
<i>Pagrus pagrus</i>	15	0

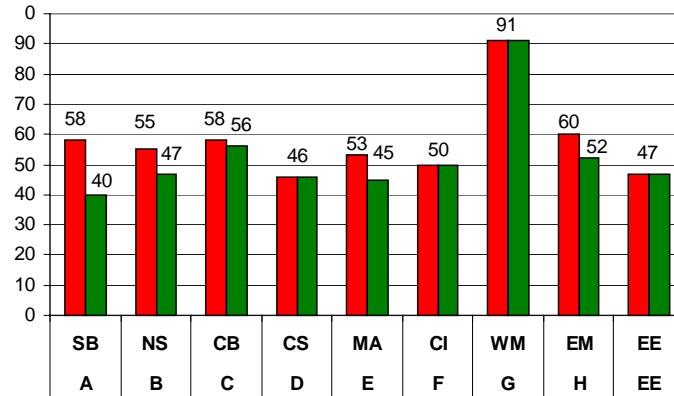
All the specimen collected from the CB area are already in the MNHN collection and are already available on the MNHN web site (www.mnhn.fr/collections/gicim).

4. Workpackages progress info

4.1. WP2 Sampling and taxonomy: R. González (instead of NRM members) showed the results compiled from the last Progress Reports (September, 2004) sent by all partners involved in this WP to the Coordinator. We can conclude that a 92% of the total number of species to be sampled is done at present (the definitive list of species must be definitively closed by March, 2005).



The following graphic shows the proportion of species sampled by geographical areas:



[Sea sampling area: A, Skagerrak and Baltic Sea; B, North Sea; C, English Channel and Bay of Biscay; D, Cantabric Sea, NW Iberian Peninsula (Galicia and Portugal) and Azores; E, Madeiran archipelago; F, Canary Islands; G, Western Mediterranean and Bay of Cadiz; H, Eastern Mediterranean (Greek Seas)].

Other outstanding parameters in this WP were presented:

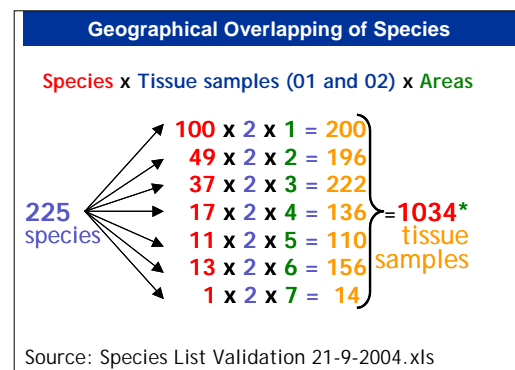
- Vouchers transferred to its curator group: 1050 / 2585
- Muscle tissues transferred to its curator group: 1642 / 2068
- Pair of otoliths transferred to its curator group: 300 / 1034
- Specific information sheets filled-in (database): 1265
- Images obtained from fish: 4132 / 2585

4.2. WP3 Molecular Genetic Identification: R. González showed WP3 progress:

Muscle tissues received:

A: Skagerrak & Baltic Sea (SB):	320 / 116
B: North Sea (NS):	235 / 110
C: English Channnel & Bay of Biscay (CB)	120 / 116
D: NW Iberian Peninsula&Cantabric Sea (CS)	58 / 92
E: Madeiran Archipelago (MA)	141 / 106
F: Canary Islands (CI)	119 / 100
G: W. Mediterrranean & Bay of Cadiz (WM)	173 / 182
H: Eastern Mediterranean (EM)	116 / 120
EE: Extra-European (EE)	61 / 94

TOTAL: 1343 / 1034*



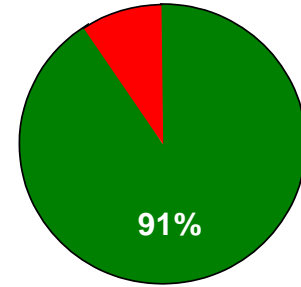
* We have collected more tissue samples than expected due to:

1. Not specified: 01 and 02 or WP4 tissue received in the progress reports.
2. New species added.
3. Tissues received from species removed from the tentative list.

We need to CHECK and UPDATE the lists of: a) **species** and b) **tissues** stored / analysed.

DNA extractions:

A: Skagerrak & Baltic Sea (SB):	70 / 116
B: North Sea (NS):	154 / 110
C: English Channel & Bay of Biscay (CB)	68 / 116
D: NW Iberian Peninsula&Cantabric Sea (CS)	58 / 92
E: Madeiran Archipelago (MA)	141 / 106
F: Canary Islands (CI)	119 / 100
G: W. Mediterranean & Bay of Cadiz (WM)	173 / 182
H: Eastern Mediterranean (EM)	100 / 120
EE: Extra-European (EE)	61 / 94

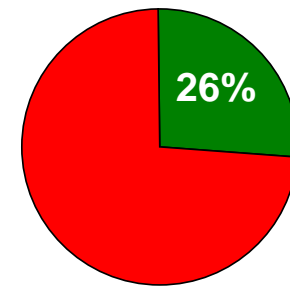


Number of DNA extractions from tissue samples by all partners at present: 944 / 1034 (**91%**)

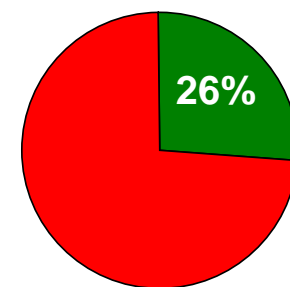
To complete the expected number of sequences for the Genetic Catalogue we need to extract DNA from 1,034 tissue samples: taking geographical overlapping of species there are 517 species x 2 tissues for genetic analysis each species = 1,034 tissue samples for DNA extraction and genetic analysis.

Complete sequences of cytochrome b:

A: Skagerrak & Baltic Sea (SB):	0 / 116
B: North Sea (NS):	60 / 110
C: English Channel & Bay of Biscay (CB)	43 / 116
D: NW Iberian Peninsula&Cantabric Sea (CS)	17 / 92
E: Madeiran Archipelago (MA)	22 / 106
F: Canary Islands (CI)	11 / 100
G: W. Mediterranean & Bay of Cadiz (WM)	39 / 182
H: Eastern Mediterranean (EM)	70 / 120
EE: Extra-European (EE)	10 / 94
TOTAL:	272 / 1034 (26%)

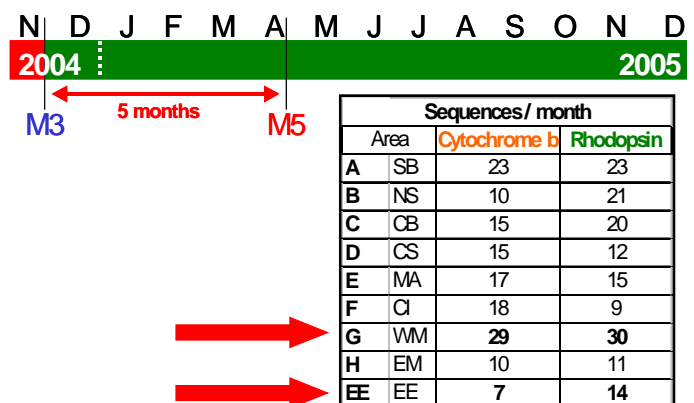
**Complete sequences of rhodopsin:**

A: Skagerrak & Baltic Sea (SB):	0 / 116
B: North Sea (NS):	5 / 110
C: English Channel & Bay of Biscay (CB)	27 / 116
D: NW Iberian Peninsula&Cantabric Sea (CS)	31 / 92
E: Madeiran Archipelago (MA)	29 / 106
F: Canary Islands (CI)	56 / 100
G: W. Mediterranean & Bay of Cadiz (WM)	32 / 182
H: Eastern Mediterranean (EM)	66 / 120
EE: Extra-European (EE)	26 / 94
TOTAL:	272 / 1034 (26%)

**Conclusions:**

We have still five months to obtain all targeted gene sequences (Milestone 5) from the complete supply of essential biological specimens for collections and genetic analyses (Milestone 3).

We have estimated the necessary number of sequences that we need to obtain by each geographical sampled area and by months to complete this workpackage at the expected date. We have to dedicate more efforts on areas G (WM) and EE (EE).



4.3. WP4 Biogeographical Molecular Haplotyping: G. Krey continued with WP4 progress info:

List of species for WP4 purposes and areas in which they are present:

Merluccius merluccius	(A,C,D,G,H)
Micromesistius poutassou	(A,B,C,D,G,H)
Mullus surmuletus	(B,C,D,E,F,G,H)
Pagrus pagrus	(C,D,E,F,G,H)
Pagrus erythrinus	(C,D,E,F,G,H)
Solea vulgaris	(A,B,C,D,F,H)

G. Krey continued explaining the protocol to follow for the comparison between populations within a species:

I. Nucleotide diversity/haplotype diversity:

Amplification and sequence of genes as described in WPs 1 and 3

II. Genetic distances between DNA sequences:

Kimura “two parameter” or Tamura and Nei distance tests.

Both tests output a corrected percentage of nucleotides for which two haplotypes are different and take into account different substitution rates between transitions and transversions. In addition Tamura and Nei distance distinguishes between transition rates between purines and pyrimidines.

III. Phylogenetic trees:

Neighbour-joining algorithm: allows for unequal rates of molecular change among branches. Branch lengths are proportional to the count of change.

The trees are calculated from the original genetic distance matrices using different software (PHYLIP, PAUP)

IV. Population structure analysis:

Hierarchical analysis of genetic variance within and between populations (groups) to estimate fixation indices (F statistics) with AMOVA (Arlequin software).

G. Krey finished his presentation showing the current status of WP4:

UCM (regions D, E, F, G):

Merluccius merluccius (region D):	17/20, only 5' sequence
Merluccius merluccius (region G):	17/20, only 5' sequence
Mullus surmuletus (region E):	16/20, only 5' sequence
Mullus surmuletus (region F):	13/20, only 5' sequence
Mullus surmuletus (region G):	18/20, only 5' sequence
Pagrus pagrus (region E):	17/20, only 5' sequence

IFREMER (region C)

Micromesistius poutassou: 20/20 entire cytb sequence (5' and 3')

NAGREF (region H):

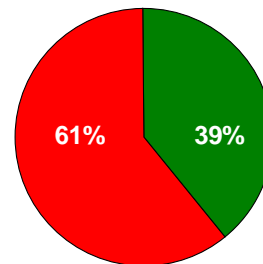
Merluccius merluccius:	17/20 entire cytb sequence (5' and 3')
Solea solea:	5/20 entire cytb sequence (5' and 3'), 5/15, only 5' sequence

4.4. WP5 Biological Reference Collections: M. Biscoito exposed the progress on WP5: The total number of vouchers, tissues and otolith samples included in collections were:

- MNHN: 104 species (49 vouchers 01 & 02) and 19 otolith samples.
- NRM: No updated information available.
- TFMC: WM - 83 species (vouchers 1 & 2), 66 pairs of otoliths, CS - 51 species (43 vouchers 1 & 2), 35 pairs of otoliths, CI - 49 species (43 vouchers 1 & 2), 24 pairs of otoliths, EE - 37 species (vouchers 1 & 2), 32 pairs of otoliths.
- IMAR: 45 species (39 complete), 17 pairs of otoliths (from voucher 2).

INCLUDED IN COLLECTION

1. Voucher specimens:	1018 / ~2585 (39%)
IMAR (MA):	118 / 265
MNHN (CB):	277 / 290
NRM (SB):	205 / 290
TFMC (CI):	83 / 250
TFMC (CS):	86 / 230
TFMC (EE):	77 / 235
TFMC (WM):	172 / 455
(NS?, EM?)	
2. Total DNA in Collections	400 / 1034
3. Muscle Tissue in Collections	2043 / 2585
4. Pair of Otoliths in Collections	381 / 1034



4.5. WP7 Validation in Database: After a short discussion among all FishTrace partners, a working group formed by: P. Pruvost (MNHN), P. Leontarakis (NAGREF), S. Jiménez (TFMC), I. Lozano (ICCM) and J.A. González (ICCM), all in charge of Taxonomy Data Validation in Database, met together to discuss this item. They have reported the agreements reached on this meeting:

MEETING REPORT ON VALIDATION SESSION plus AGREEMENTS and DEADLINES (See below: 11. Data Validation Flow and Data Status Summary)

4.6. WP6. Database Development: P. Carreau explained developments on the database. His presentation was divided in two sections : Database Status and Complains on this part.

Database Status

- Database migration on a new server: Dedicated computer for the project (perfect for future maintenance).
Load Fishbase tables concerning all European commercial species.
New password system.
FTP protected (required user IP number).
- Interface improvement.
Some bugs tracking and minor improvements (thank you to NAGREF and MNHN).
System accepts many special characters now (French Swedish Spanish).

View and delete data	
View specimens Delete data FAQ	<p>You can delete files (specimen, environment) but the system check if there is no reference somewhere in the database (for example you cannot delete an environment file if it is used in a specimen).</p>

In “view specimen” you can select the number of specimen you want to view (in case there are more than five)

- Select below -

Nb of specimens to display: 5

Search

Add a new picture
WARNING: YOU MUST LOAD FIRST THE SPECIMEN FORM BEFORE ENTERING THE SPECIMEN PICTURE.
 There are 2 step for loading an image
 1)Load the image using FTP by clicking on the links below (see how to process)
 2)Complete this form and submit it

Id_fishtrace code : AloFal-CB-01

CAUTION: only JPG format are accepted.
 Enter the picture name with the extension .jpg

Picture name : AloFal-CB-01.jpg (!Reference voucher picture name by default)

You must enter only one reference voucher picture and one reference otolith picture

Picture type : reference voucher pict (!Reference voucher picture name by default)

Remarks(Optional):

[View previous entered picture for this specimen](#)

- Reference voucher file by default in load picture:

- View the ever loaded pictures

- Public web site to upload images:
 For partners who have problems (not fix IP address)
<ftp://fish-ftp.jrc.it/fishimages/>

- Future action: Separate genetic and taxonomy upload for species.

Complains

- Generally we can say the loading process is good but ...
- Still some data quality problems: Is everything clear concerning the fields specimen dorsal and anal finray?
- Avoid -2 but don't be stopped by details:
 Time consuming: 1)search 2)find the field 3)load.
 Generate errors
 Load at the wrong place
 Delete ever loaded information
 High risk to forget to complete information
 Advice: Complete data the same week (Studies shows that once forget a lot of thing during the WE)
- Clean everything
 For example if you upload a wrong image. Please send me a Email to delete it.
 Delete non used files (for example an environment you never use because you decide to replace it by another)
- Ask for help !
 In case of doubt send an Email on the FishTrace list or me even if you think the question has been answered or is stupid
 I have less and less requests -> good sign?

4.7. WP8. Implementation of www technologies: P. Carreau continued detailing the status of this workpackage:

- 1st version of public interface online
- Pescabase like interface design
 The interface looks the same but the back office technology is totally different
 Effort to “light” the interface to increase loading connection
 The interface focus mainly on the species not specimens but I put them in the specimen tabulation
- Development of search system
 by Latin name (FishTrace list)
 using common names of the tables furnished by Fishbase (multilingual)

5. Open discussion on Workpackages progress

During this session, all partners participated in presenting the status of each WP, as well as problems encountered and the estimation of the time and tasks remaining.

WP1: There are still some minor aspects missed (part of legal aspects).

WP2: 91% covered. M. Biscoito suggested a deadline for the definitive Species List by the end of February, 2005. J.M. Bautista proposed to substitute those species that partners in charge of sampling found not available by other important species in that geographical area. This will help to built the definitive list since those not available species are a minimal part of the original list and will not affect the total number of species covered in FishTrace.

WP3: On schedule.

WP4: Some changes were taken from the beggining but almost on schedule.

WP5: On schedule but some gaps on DNA samples and vouchers exchange-storing. P. Pruvost mentioned some troubles raised from the voucher shipping procedures.

WP6: The structure and Web Interface is almost done but the data inclusion is slightly delayed. J.M. Bautista proposed to draw a protocol up with detailed steps for data inclusion on database.

WP7: Some taxonomic and genetic data is already validated. Further agreements were taken during the meeting (See below: part 11).

WP8: Undertaken and on schedule.

WP9: Next actions to be taken from next period (year) of the project (See below: part 21).

6. Link to FishBase

Possibilities and Discussion about the link Fishbase-FishTrace:

Ma. L. Palomares and N. Bailly made a introduction about the Fishbase project and suggested the following possibilities of linking our project with Fishbase:

1 - In Fishbase main page there is a field called "Links". It's easy to put FishTrace there.

2 – In the species page, in the field name "Genetic Records".

3 – In the species summary, in the field name "Occurrence". We can add in "Occurrence Records" the FishTrace code. Other possibility regarding "Occurrence" is to introduce the data through Countries (or geographical areas) in the "Occurrence Map list".

4 – Add FishTrace pictures to the ones of Fishbase.

P. Carreau indicated that the reverse link FishTrace-Fishbase is easy. In FishTrace web page there is a field called "Select latin name or enter Fishbase species code" where we can enter a Fishbase species code in order to access directly to the Fishbase summary page regarding the species in question. FishTrace must think about N. Bailly's suggestions and decide which of them should be chosen (for the next meeting). For the links might we will need specific in and other costs like personnel, maintenance of the web, etc...

7. Database adaptation to GBIF: Darwin Core v2

J.M. Bautista presented GBIF (Global Biodiversity Information Facility, www.gbif.org) and the possible adaptation of FishTrace's Database to GBIF database:

GBIF is a megascience facility aimed at making the world's biodiversity data freely and universally available via the Internet and sharing primary scientific biodiversity data for society, science and a sustainable future. It is a distributed facility, comprising a network of Participant nodes that:

- Share biodiversity data openly and freely.

- Use common standards for data and metadata.
- Encourage generation of additional content.
- Assure that data providers retain control of their own data.
- Gain access to others' data by sharing their own.

J.M. Bautista continued explaining that no special changes on FishTrace Database are necessary to link it with GBIF database but that might should be needed some infrastructure and maintenance:

- A PC outside institution firewalls
- Personnel dedicated to maintenance of Database
- etc...

8. Project Status Summary

J.M. Bautista presented the Project Status Summary concluding that although the Project is going on schedule, there are some "gaps" that must be clarified soon, and proposed to ask to the Commission for 6 months PROJECT EXTENSION, but before that it should be estimated again whether this will be necessary by May 2005.

9. Other Links: FishGene, Museum Collection, Databases...

Some possible links with FishTrace database were mentioned:

- FishBase (previously discussed)
- FishGen (this has been already linked by P. Carreau)
- Museums Databases
- Redfish
- ...

For those databases with low number of species, FishTrace would only link at species level (species in common). It was concluded that it is right to compile as more information as we can at Species level, and it implicate links with other Fish/Molecular/Collection databases.

10. Database Status Summary

P. Carreau (JRC) presented a table showing the data included on Database at present. Although all groups have started filling the database in, less information than expected has been included until now.

11. Data Validation Flow and Data Status Summary

Internal validation of data:

It was discussed how the FishTrace data should be validated before entering into the public database.

It was agreed:

- WP7 responsible will publish bimonthly reports on missing data on the web in Feb, April, June, August and Oct 2005 in a WP7 excel file.
- This excel file will be send to Philippe, to be uploaded on the web in December 2004.
- Partners have to fill in this excel file.
- The definite version should be ready in October 2005.
- There are **no** deadlines for entering data into the database.
- There is a deadline for collection of species: 28 Feb, 2005.

There was a general discussion how the data validation should be controlled. A new proposal was prepared during the meeting. H. van Pelt asked P. Pruvost if he could send this information to be incorporated in these minutes, and working group (meeting rapporteurs) formed by S. Jiménez (TFMC) and J.A. González (ICCM) have reported the agreements reached:

- **Validation of each specific specimens at each area:**
 - The scientist responsible of each partner for the information uploaded will validate this data online on each specimen table. For this purpose, P. Carreau (JRC) will modify the Specimen's Table by adding 3 fields for the signature, data and remarks from the scientist responsible.
 - An excel file has been prepared by ICCM-TFMC containing the whole target species by area, in order to control the validation data status summary at each area and to know the real progress of the dabatabase content. This file is being attached now and will be uploaded by JRC in order to make possible all checks online.
 - At this specimen level, please check carefully the Environmental Table, Specimen Table and Sample Table. Within the Specimen Table, do not forget to check the specimen's images (both voucher and otolith).
- **Validation of each target species at each area:**
 - At this species level, please check carefully the existing Regional Info Table. Also check the corresponding bibliographic references on the Bibliography Table.
 - Once all these checks were done, the taxonomist responsible of each partner will put a X on the -2 given by default in the excel file (just on the taxonomy column TA of his/her area corresponding to the target species internally validated).
 - Once all these checks were done, the curator responsible of each partner will put a X on the -2 given by default in the excel file (just on the collection column CO of his/her area corresponding to the target species internally validated). We will assume that the DNA collection field will be empty at that moment.
 - Once the internal validation agreed by geneticists was done, the geneticist responsible of each partner will put a X on the -2 given by default in the excel file (just on the genetic column GE of his/her area corresponding to the target species internally validated).
- **Validation of each assigned species:**
 - Once a given species was completed with all biogeographical crosses (X), then the scientist responsible for taxomomy/collections at each partner/group (that is, ICCM-TFMC, IMAR, NAGREF, IFREMER-MNHN and NRM) will start its own validation.
 - First of all, this validation process will consist in entering the information concerning the specific Common Fields (Species Table in part): Popular description, Diagnosis, Similar species, Biology, Distribution and interest to fisheries, Conservation status, Source of information, and Pictures & id.
 - Secondly, all information and images referred to an assigned species (including the existing Regional Info) will be compared and contrasted in order to detect any error, mistake, missing data or deviation from the specific Common Fields.
 - When necessary, the scientist responsible for this species will ask for problematic data to the involved partner(s).

- Regarding the specific images, the scientist responsible for this species will check the specimens' identity by comparing available photos, contact with partner(s) involved when any image differs, and select quality images as consensus reference pictures (both voucher and otolith).
- Once all these checks were done, the scientist responsible for this species will put a X on the -2 given by default in the excel file (just on the general TAXONOMY column corresponding to the assigned species). The general COLLECTIONS column will be completed when receiving and registering the DNA sample.
- Once a given species was completed with all biogeographical crosses (X), then the Expert Group on Genetics will start its own validation by following 3 steps: a) comparison of sequences from the same species from different geographical area; b) annotation of new changes observed, and methods employed; and c) verification of the position of each sequence on fish phylogeny.
- Once all these checks were done, the Expert Group on Genetics will put a X on the -2 given by default in the excel file (just on the general GENETICS column corresponding to the given species).
- P. Carreau (JRC) will prepare Specimens' Tables sorted by species and by area to be easily printed containing all information to be checked/compared. In order to achieve an optimal specimen comparison, we suggest to display all variables (both filled and empty) in rows and their specimen's parameters (-1 & -2 values included) in columns.
- P. Carreau (JRC) will modify the Species Table by adding 3 fields for the signature, data and remarks from the scientist responsible for taxonomy/ collections, plus 3 fields for the signature, data and remarks from the scientist responsible for molecular genetic.
- **Data Status Summary**
- The WP7 responsible (ICCM) will send a bimonthly report on Data Status Summary to the project's coordinator by the middle of February, April, June, and August 2005.
- Finally, the WP7 responsible (ICCM) will send a final report on Data Status Summary to the project's co-ordinator by the end of October 2005.

12. Submissions of sequences to GenBank (Sequin Tool)

It was agreed to submit DNA sequences to NCBI, since this is a European Initiative, in contrast to GenBank.

It was discussed how and who will submit sequences to the NCBI. It was proposed to submit data as consortium. Authors should be everyone who contributed to the sequence. Single sequences will be submitted. G. Krey proposed to enter the population study as well.

Which related data should be incorporated? It was suggested:

- Genetic variation
- Location/map
- Sex
- Species voucher number
- Tissue information

Who will enter the sequences? All validated sequences will be entered in FASTA format. Sequences will entered all together in 1-2 days before June 2006. G. Krey and J.M. Bautista will be responsible.

13. Special session for geneticists: Submission of sequences to GenBank

After a long discussion among all FishTrace geneticists involved about how to submit the sequences obtained, R. González and G. Krey have drawn a document in which the steps that we have to follow for the submission of the FishTrace DNA obtained sequences to GenBank/EMBL are detailed. This document is entitled *Submission of sequences to GenBank/EMBL: Preparing sequence files for the Sequin tool* and it is placed at the project's web page (<http://pta.jrc.cec.eu.int/>).

14. Open session: Troubleshooting on Taxonomy, Genetics, Collections, Database...

During this session, it was discussed some troubles previously mentioned (see parts 4 and 5), and suggested solutions:

Taxonomy - Collections

Vouchers sending: Several difficulties for the transport/shipping of some specimens. M. Biscoito proposed the voucher sending by ordinary mail, but in some cases it could be very expensive. It was agreed to prepare voucher exchange among institutions by March, 2005. Every voucher must carry enclosed an excel sheet with appropriate information such as "species code", "length", weight, "collecting data: coordinates" It was suggested the exchange of otoliths "in hand". P. Leontarakis and S. Iglésias added some problems found removing otoliths from some species.

Genetics

WP3: some troubles were found with the amplification of rhodopsin gene (double bands observed by M. Jérôme in electrophoresis gels). J.M. Bautista proposed to annotate any possible polymorphism or allelic variation. To solve these, it was suggested a revision of the protocols for the amplification of this gene. H. van Pelt proposed the direct amplification of the complete Cyt b gene in a single PCR. G. Krey suggested the use of PFU enzyme (Promega™) for the direct amplification of the complete Cyt b gene in a single PCR.

Database

It was mentioned possible problems linking databases (e.g. FishTrace-FishBase) if we include species that do not exist at the other databases. It was unanimously agreed that it will be better to include as much species on database as we would be able to include.

J.A. González asked about the missed location for the "geographical info" at the web interface.

M. Biscoito solved a problem found on coordinates conversions, just writing minus "-" after values.

M. Etienne proposed to add synonymous at "common names" field of database.

A hard discussion raised from the "Bibliography format" used and the format finally agreed consist of an numeric "clickable" code: e.g. FishTrace 101, 535, 1034...

15. Web Interface

During this sessions was a communal discussion on changes suggested for the web interface, as well as the bugs detected on it. Due to the complexity of this session for different expert groups (Taxonomists, Geneticists, etc), it was divided by groups and the agreements reached were subsequently presented (See below: parts 16 and 17).

16. Taxonomy – Collections Group Discussion

J.A. González (ICCM) and S. Jiménez (TFMC) sent the following report on Collections and Taxonomy plus the agreements taken and deadlines for next actions:

Taxonomical aspects (including some web interface comments):

- The final list of the target species will be closed by the end of February 2005.
- According to the fish samples availability, some target species could be replaced by others in order to accomplish the total number of species targeted at each sampling area.
- When working on several sampling areas, a balance species among areas could be valid in order to accomplish the total number of target species assigned to each partner.
- All modifications and substitutions related to the target species must be informed to the partner in charge, i.e. ICCM.
- The taxonomical status of *Spicara flexuosa* (Centracanthidae) and *Muraena augusti* (Muraenidae) was clarified: both will be considered as valid species according to FAO experts on taxonomy and publications which are being prepared on this topic.
- In order to make easier the uploading tasks, the partner JRC will activate all target species, i.e. the species on the current list delivered, a little bit different from the proposal list !!.
- The partner JRC will increase the extension of regional fields to 500 characters. Moreover, regional information for the target species must be displayed on the web.
- The use of taxonomical/identification texts by UNESCO (Fnam) and FAO, as well as the information provided by Fishbase (see specific sheets and the various items below) is highly recommended in order to fulfil in the Species Table (common fields).

Bibliography:

- M. Etienne (IFREMER) will prepare and circulate entry examples (one for classic journals and one for books) according to the IFREMER standards.
- Once discussed and finally approved the bibliographic structure, P. Carreau (JRC) will develop a Bibliography Table with the given examples and an automatic search by author, in order to enter the bibliographic references and be provided with the reference bibliographic number.
- Finally, both in Species Table and Regional Table, the bibliographic references must be entered according the above reference bibliographic number as follows (example): [101], [202], [503], [722], etc.

Transfer of biological samples to create the FishTrace Reference Collections:

- RIVO will send all its vouchers and other samples to Paris Museum (MNHN).
- NAGREF will keep specimen number 3 and send specimens numbers 1, 2 & 4 to Paris museum (MNHN) and number 5 to Stockholm museum (NRM) (the Paris museum having the priority). That includes the corresponding biological samples taken from these specimens.
- Vouchers (specimens numbers 4 & 5) deposited at Alicante (Spain) will be send to Paris museum by ICCM-TFMC with the help of MNHN for the preparation and transport of specimens by June-August 2005. Vouchers (specimens numbers 4 & 5) deposited at Las Palmas (Canary Islands,

Spain) will be prepared and send to Paris museum by ICCM and TFMC via ordinary mail. For these purposes, partners ICCM and TFMC must obtain the necessary permission of the Commission in order to modify their categories of eligible costs.

- Tissues from ICCM vouchers 4 & 5 will be transfer to MNHM and NRM respectively.
- Together with the transfer of vouchers and other biological samples, an excel file with detailed information (species, FT code, length, weight and collecting data) will be provided to the destination museums. Deadline for this task: by the end of February, once the sampling and taxonomic task were concluded.

17. Genetics Group Discussion

It was discussed the submissions of obtained sequences (See 13. Special session for geneticists: Submission of sequences to GenBank) and the Interface format for genetics purposes. For this last aspect, it was agreed:

- To remove the FishGen Link from “Species - Specimen sheets” except for those that are common species between both databases (FishTrace-FishGen).
- It was discussed on the inclusion of either a “reference” or “consensus” sequence as the representative sequence for each Species. It was agreed to define a “reference” sequence for each “Species” (a real sequence from a specimen), indicating changes observed among other specimens analysed into a new box and to add all sequences from each specimen at “Specimen” level. Attention was paid to real reference sequences vs. consensus sequences, since these latter could not reflect a real sequence from any specimen.
- It was proposed a way to export all sequences from all specimens belonging to the same species to compare them. To this action, it was recommended the viewing (computer screen interface) in GenBank format and the downloading of sequences in FASTA format. A translated protein sequence was suggested to be included at the GenBank format display.
- It was agreed to change the field “Haplotype” by “Genetic Variation” everywhere in the database. Under this term “Genetic variation” will be shown the changes observed among specimens from a reference sequence (the real sequence of a specimen taken as “reference”).
- On WP4, it was suggested to remove the “Id Haplotype” field from [Load haplotyping table](#) (database) and replace it by “Genetic Variation Id.”.

18. Mutual agreement on FishTrace Web Interface

It was agreed that four partners act as leaders for checking the web interface and suggest relevant changes to be made by March 2005. These responsables will receive suggestions from other partners on changes and bugs detected. The four partners in charge, according to the several parts of the interface, are the following:

- | | |
|---------------------------|-------------|
| - Main Menu: | UCM |
| - Identification Details: | IMAR + ICCM |
| - Genetics: | IFREMER |

A meeting on Web Interface was scheduled by March 11th, 2005 at Madrid. After the meeting, a full report on the web interface changes would be sent to P. Carreau. Prior to the meeting, all partners should send their comments on web interface to those responsables for each set of data.

19. Next Period Scientific and Financial Report

J.M. Bautista explained that according to the CONTRACT, each periodic report and the corresponding cost statements shall be submitted to the commission within two months (February, 28th) of the end of the period covered by the report (January - December, 2004).

He finished detailing the actions to be taken in order to avoid the same mistakes found in our first report:

1. Meet the deadlines in order to amend any possible mistakes on time.
2. Send just and what it is asked for, nothing more nothing less.
3. In further Reports we must meet the context, follow the guidelines and complete ALL sections (person-months, objectives, personnel, work accomplished, ...)
4. PROGRESS REPORTS are most important as they have details of each activity (sampling, taxonomy, genet, ref: Collect)
5. If in doubt, please ask the coordinator in sufficient time to be able to deal with the matter.
6. Add tables, graphics, figures, etc ... to each report as these add information and make the report look more professional.

DETAILED DOCUMENTATION TO BE TRANSFERRED FROM ALL MEMBERS (Team leaders) TO THE COORDINATOR:

All documents must be submitted to the co-ordinator before end JANUARY

Scientific Reports

STATUS OF THE INDIVIDUAL WORK PACKAGES (typically up to one page per work package: to be submitted by each workpackage leader to the coordinator).

- Summarise the main objectives of the work package for this reporting period.
- Compare the actual work accomplished against the activities planned, explaining any major deviations and the actions taken to remedy them.
- Provide an indication of resources used during the period.

CONTRIBUTION OF THE PARTICIPANTS (typically up to one page per participant)

- Describe the contribution of each participant in sufficient detail to justify its cost claim for the reporting period, noting in particular the personnel and other resources deployed on the project.
- Indicate any significant problems encountered and the actions taken to remedy them.
- Indicate any changes to the scientific team during the period.
- Report on any activity by subcontractors during the period.

Financial Reports (Cost Statements)

- Each *member* shall submit a cost statement in the format specified in part **E-1**, in duplicate and both originals signed.
- Definition of *Direct cost*, *Indirect cost*, etc, are explained on PART C, Article 17, page 27 of the Contract Model (ANNEX II).

ALL THE FORMS TO BE FILLED AND INFORMATION ARE AVAILABLE AT CORDIS WEBPAGES:

<http://www.cordis.lu/fp5/management/provisions/r-conact-pluri-4.htm>

<http://www.cordis.lu/life/src/projmgmt.htm>

MODEL CONTRACT SPECIFICATIONS:

ftp://ftp.cordis.lu/pub/fp5/docs/moc_capluri_en_199901.pdf

20. Next FishTrace meetings

Three next meetings were agreed:

- Short-term Meeting: **Madrid, March 11th, 2005 (Web Interface Resolution)**
- Intermediate Meeting: **Funchal, September, 2005 (Project progress / Status / Prospects)**
- FINAL MEETING: **Kavala, April - May, 2006**

21. WP9. Dissemination: Actions for the next period security and fisheries

- Brochure (2000 copies planned): before the Paris meeting, she sent the draft of a brochure. The last comments have to be sent to M. Etienne. The English language has to be checked. Each participant has to send to M. Etienne a "national list" containing name, address and e-mail of contacts (scientific institutes, control laboratories, technical centres, administration & regulatory bodies) for the future mailing of this brochure.
- CD-ROM (1000 copies planned): a small CD-ROM will be also realized. Its content will be the content of the brochure (in english) + different translations of the brochure (Portuguese, Spanish, French, Swedish, Dutch, Greek, Italian, German (by UCM), Danish (by who?), Lithuanian (by Philippe Carreau colleague), Polish (by Ifremer people)). Monique will sent a draft of the CD-ROM. She asked Philippe for help for the realization of the CD.
- E-mailing: suggestions have been made to send e-mails instead of sending CDs or brochures. M. Etienne will prepare a draft for this e-mail dispatching.
- Links: M. Etienne proposed also that each partner's institution will link its web site to the FishTrace database web also have a link to the FishTrace database.
- FAO and Codex Alimentarius: As she has an opportunity, M. Etienne will present the FishTrace project at FAO during a meeting at "Fishery Utilization and Marketing Service" in 14th of December.
- In order to inform the Codex Alimentarius, M. Etienne proposed to contact each national delegates of the Codex Alimentarius and the DG SANCO before the next Codex meeting (in February 2005).
- Exhibition: Regarding an eventual presentation of FishTrace at Seafood exhibition at Bruxelles in 2005, it may be possible at the booth of the RIVO. H. van Pelt will asked for.
- Symposium presentations: a presentation of FishTrace has been made by V. Verrez-Bagnis during the last WEFTA (Western European Fish Technologist's Association) meeting and by J. Bautista at the EFARO symposium.
- Brochures and CDs will be available at the next WEFTA meeting (19-22 September 2005 – Anvers).
- A presentation to FASIS symposium (Food Authenticity and Safety International Symposium) in 19-21 October 2005 in Nantes will be done by J. Bautista.
- We have not decided the 3 fish species to be putted on the database web site as examples for the dissemination.

22. Open session: Discussion on dissemination and brochure

A short mention about suggested changes on Dissemination Brochure took place on this sessions. We agreed to send M. Etienne any suggestion/change on the format, text, images, etc of the Dissemination Brochure after the meeting.

23. Follow up of FishTrace through new proposals in EC calls in FP6 and FP7

Safeguarding the continuation of FishTrace so that its structure (including the collections established, the database created, and the technological know-how generated) can be maintained functional after the completion of the project is a matter of considerable concern for the consortium members.

To that end, J. Bautista presented the possibilities, through EU funding, that could result in keeping and/or expanding the FishTrace activities and its current network structure beyond the lifetime of the project. These mainly include the submission of new proposals to calls arising either within the FP6 or the FP7.

It was discussed that such proposals could be addressing issues in either of the following three activity areas of present or forthcoming calls:

1. Policy oriented research
2. Research infrastructures
3. Food quality/safety research

Within each of the above activity areas, different instruments are available for the formulation of specific proposals, including the IP, NoE, STREP, SSA, CA and CA-I3. The nature and requirements of each instrument was explained in respect to which could better address the particular needs of the FishTrace network.

However, it was mentioned that the particular interests and priorities of the EU concerning food quality and safety issues, will probably not accommodate a proposal of the FishTrace scope. Thus, the efforts of the present consortium should be directed to obtaining further funding through proposals addressing policy oriented issues and/or proposals aiming to reinforce the present research infrastructure of the network.

Furthermore, the consortium was informed that actions that could lead to include issues related to FishTrace interests and/or to the application of the technology generated by the project in future FP6 and FP7 calls have been taken. Thus, in the most recent EFARO meeting (Lisbon, 28-31/10/2004) the need for traceability tools in both the fisheries and fisheries products sectors was argued and specific recommendations for calls including such actions were made.

24. Open discussion on new proposals in EC calls in FP6 and FP7

There was a general consent for the formulation of new proposal(s) that if successful will allow the uninterrupted continuation of the network. This implies that such proposal(s) should be prepared and submitted in the upcoming calls and if possible in calls of the FP6 already published. Concerning the latter possibility, it was mentioned that the next deadline for I3 calls (Integrated Infrastructure Initiative) is the 03-03-2005.

It was argued that this particular instrument fully covers the range of activities that the network wishes to pursue in the future, including (1) networking activities (e.g. to generate critical mass by coordinating research into new methods, concepts and technologies, while also aiming at promoting common protocols and interoperability, encouraging complementarity and, where appropriate, developing and maintaining common databases and dissemination of knowledge, training courses for potential users and foresight studies); (2) provision of access to transnational users; and (3) joint research activities (e.g. testing of new experimental techniques or methodologies).

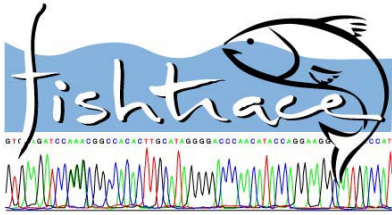
All of the consortium members present at the meeting agreed to participate in preparing and submitting a proposal for the 03-03-2005 call. A steering committee was formed, consisting of the UCM, JRC, MNHN, and NAGREF members, with the responsibility of preparing a detailed outline of the proposal by January 2005.

FOR FURTHER INFORMATION, ALL PRESENTATIONS-DOCUMENTS ARE PLACED AT THE FTP: ftp://fish-ftp.jrc.it/fishimages/Paris_ppt/ (Login and Password are the same than in Database)



FishTracers at work

Annex VII: FishTrace Database & Web Interface Meeting. Madrid, 11 March, 2005.



FishTrace Meeting on Database Web Interface

March 11th, 2005.
Madrid, Spain.

Minutes from the FishTrace Database Web Interface Meeting

Objectives:

- To find agreement positions on the definitive structure and appearance of the Database Web Interface, presenting the whole data from the FishTrace database.
- Discussion on WP9 - Deliverable 11 (Model tools for identification and differentiation of fish species).

Venue:

Held at the department of Biochemistry and Molecular Biology (IV), Veterinary Faculty, UCM. Ciudad Universitaria s/n. Madrid, Spain.

Dates:

March 11th, 2005.

Participant list:

The meeting has been attended by reponsible for each set of data (genetic, taxonomy, collections, database) and by the project coordinator. Each responsible presented the compiled suggestions from partners on each set of data:

- **Main Menu:** José M. Bautista and Rafael González (**UCM - Host**)
- **Taxonomy:** José A. González (**ICCM**) and Sebastián Jiménez (**TFMC**)
- **Collections:** Manuel Biscoito (**IMAR**)
- **Genetics:** Véronique Verrez-Bagnis (**IFREMER**)
- **Database:** Philippe Carreau (**JRC**)

Agenda:

1. Meeting introduction **J.M. Bautista**
2. Presentation of the current Interface and modifications made. **P. Carreau**
3. Discussion on Web Interface: Sampling and Taxonomy. Changes on "Taxonomy" suggested by partners. **J.A. González**
4. Discussion on Web Interface: Collections. Changes on "Collections" suggested by partners. **M. Biscoito**
5. Discussion on Web Interface: Genetics. Changes on "Genetics" suggested by partners. **V. Verrez**
6. Discussion on Web Interface: Main Menu. Common agreement on Main Menu. **R. González**
7. Common agreement on the public interface. Final report on FishTrace database web interface (first draft).
8. Discussion on *Model tools for identification and differentiation of fish species* (WP9-D11).
9. FishTrace Website design.

Minutes:

1. Meeting introduction

The meeting started at 9:00 with a welcoming introduction by J.M.Bautista.

2. Presentation of the current Interface and modifications made.

P. Carreau explained developments on the database. His presentation was divided in two sections:

- a) current status
- b) modifications made on the database web interface.

P. Carreau continued explaining that he presented FishTrace at the last meeting on DG Fish / HOMSIR projects at Brussels.

It started an open discussion on the future development and use of FishTrace database: inclusion of new species (imported, etc.), data validation by other parameters (otoliths classification), including stock identification (which interest to DG Fish), legal aspects, etc. P. Carreau concluded indicating that studies on Populations/Stocks should be considered as a priority for future development of FishTrace Database.

3. Discussion on Web Interface: Sampling and Taxonomy.

J.A. González presented the compiled suggestions on Taxonomy from all partners. In the following boxes are marked corrections on sampling and taxonomy pages of the database which also affects the Web Interface.

Corrections on Sampling and Taxonomy pages of the database presentation:

Sampling

Objectives

1. Analyse the fish specimens, the tissue and the otoliths
2. Determine and classify the species
3. Provide data about the species

Targeted Species

Around 225 marine fish species from 7 European sea areas and 1 extra-European area will be analysed:

- 57 Species from Skagerak and Baltic
- 46 Species from the North Sea
- 59 Species from the English Channel and Bay of Biscay
- 62 Species from Cantabric Sea and NW Iberian Peninsula and the Azores
- 47 Species from the Madeiran Archipelago
- 40 Species from the Canary Islands
- 85 Species from the western Mediterranean Sea
- 54 Species from the eastern Mediterranean Sea

Please click for [full list of species](#)

Numbers and lists of species must be updated by mid April 2005.

Sampling in the field

5 specimens from each area, fresh fish should be used. Steps:

1. fill out sample data record afap
2. identify species
3. mark/tag specimen
4. take specimen photo
5. measure and weigh
6. take tissue from 5 fresh specimens
7. place tissue in 15 ml 95% ethanol tubes
8. remove otoliths to marked dry tubes
9. place sampled specimens in 10% formalin formaldehyde
10. keep intact specimen frozen whole
11. complete sample data record

Sampling in the Lab

Steps:

1. distribute tissue samples as needed
2. transfer specimens 1-4 to 70% ethanol
3. take standard measurements and counts
4. X-ray
5. Record diagnostic characters relevant for taxonomy
6. distribute vouchers
7. digitise data collected

Taxonomy

Characterisation

Species characterisation will be carried out by:

- Check/verify against literature
- Basic measurements: Standard Length, Head length, Snout length, Body depth, Eye diameter, Interorbital width
- Basic Counts: -Dorsal, anal, and pectoral fin rays, Scales in lateral line, Gill rakers, Other relevant characters: Abdominal and Caudal
- Biology
- Fisheries
- Socio-economic aspects

4. Discussion on Web Interface: Collections.

M. Bischoito presented compiled suggestions on Reference Collections from all partners:

Corrections on Reference Collection page of the database presentation:

Collections

Methodology

- Two voucher specimens used to obtain mmolecular data preserved in 70% ethanol
- Replicate DNA samples to be frozen at -20°C
- Muscular tissue samples to be kept refrigerated in 70% ethanol
- Replicate DNA samples to be frozen at -20°C
- Sagittal otoliths will be adequately stored as desiccated specimens
- Inventories will be exchanged among the participating museums

More modifications on **Sampling, Taxonomy and Reference Collections** web pages, agreed during the meeting, are explained at the following report made after the meeting by J.A. González, S. Jiménez and M. Bischoito (Fig. nº 1, 2 and 3):

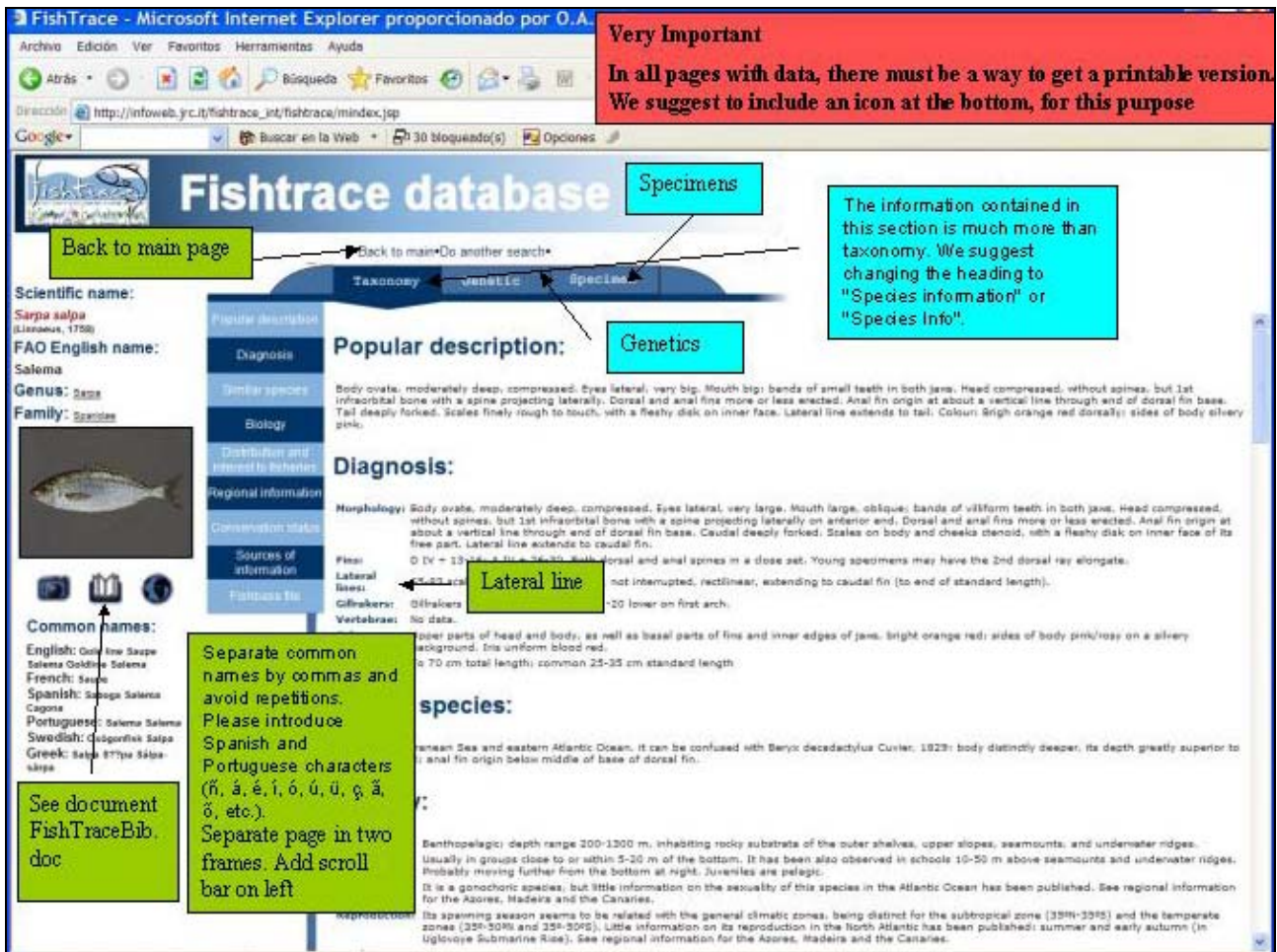


Figure nº 1. Modifications on Taxonomy web page.

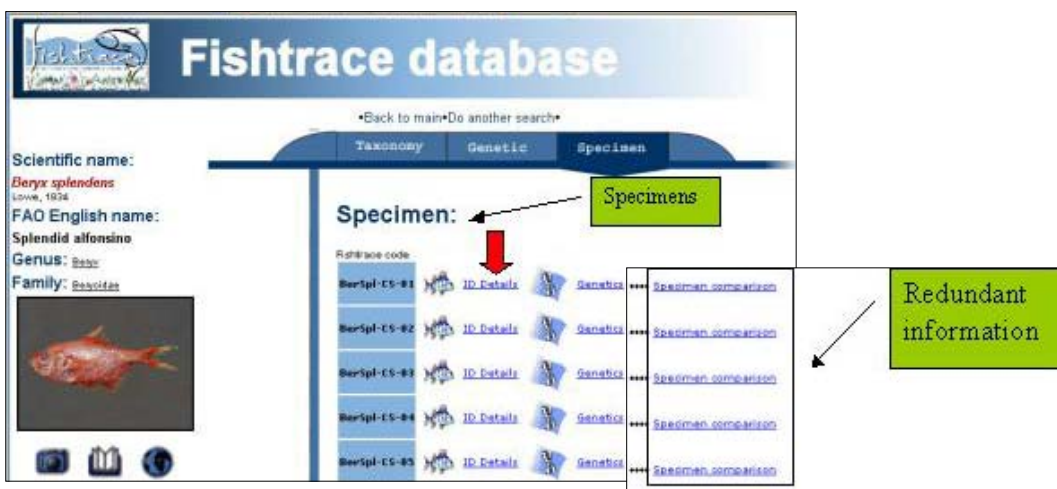


Figure nº 2. Modifications on Specimens page.

Specimen data

Fishtrace code: BerSpl-CS-03

Specimen

Fishtrace code:	BerSpl-CS-03
Species code:	1326
Identification date:	3/2/2004 0:0:0
Identified by:	Gonzalez,JA
Voucher fixation:	non relevant
Voucher preservation:	non relevant
Voucher collection institute:	non relevant
Voucher collection nb:	-1
Total length(mm):	396
Fork length(mm):	330
Standard length(mm):	210
Total weight(g):	90
Ontogenic stage:	Adult
Head length(mm):	100
Body depth(mm):	107
Sex:	Unknown
Gillrakers:	7
2nd gillrakers:	0
3rd gillrakers:	20
Dorsal fin-rays:	IV 15
2nd dorsal fin-rays:	-1
3rd dorsal fin-rays:	-1
Anal fin-rays:	IV 29
2nd anal fin-rays:	-1
Pectoral fin-rays:	17
Scales on lateral line:	-1
Otolith collection inst:	TFMC
Otolith collection nb:	EMOTO-VR/0553
Id biogeography:	ICCM_C_M_15_10_02

Collecting environment

Id biogeography:	ICCM_C_M_15_10_03
Locality:	ICES VII
Depth:	-3
Fishing method:	trammel nets
Longitude(dec deg):	-4.3666667
Latitude(dec deg):	48.5166667
Coordinate source:	other
Date of capture or purchase:	10/15/2003 0:0:0
Collector or purchaser:	Arago, S
Cruise name:	Croix Moran
Landing site:	Port of Parajay

Associated tissue sample

Fishtrace code:	BerSpl-CS-03
Tissue collection Id:	EMTEP-VR/1326
Tissue collection institute:	TFMC
Storage medium:	alcohol
Id tissue:	BerSpl-CS-03-b

General remarks:

- Remove FishTrace code from title (already in data listed) and put scientific name instead
- Sub-headings:
- Replace “Collecting environment” by “Collection data”
- Remove “Associated”
- Replace “Genetic parameter” by “Genetics”
- Delete **hour** from all date fields
- Voucher fixation: Instead of “non relevant” use “4 percent formaldehyde” or “70 percent ethanol” or “Specimen not preserved”.
- Voucher preservation: Instead of “non relevant” use “4 percent formaldehyde” or “70 percent ethanol” or “Specimen not preserved”.
- Voucher collection institute: Instead of “non relevant” use “None or no data available”
- Replace “Voucher collection nb.” by “Voucher collection No. :”
- Replace -1 by “None or no data available”
- Replace -2 by “Waiting for data”
- Replace “Otolith collection inst.” by “Otolith collection:”
- Replace “Otolith collection nb” by “Otolith collection No. :”
- Remove ID Biogeography
- Change order of Latitude and longitude. Latitude first.
- Convert Latitude and longitude to conventional notation: ex. 48° 30.6’ N and 06° 21’ W
- “Cruise / ship name:” instead of “Cruise name:”
- Replace “Tissue collection inst.” by “Tissue collection:”
- Replace “Tissue collection Id” by “Tissue collection No. :”

Gill rakers (1st gill arch)
(as label)

Upper branch:
At angle:
Lower branch:

Dorsal fin-rays:
2nd Dorsal fin-rays:
3rd Dorsal fin-rays:
Anal fin-rays:
2nd Anal fin-rays:
Pectoral fin-rays:

Please follow the same order as above: code, tissue coll., coll. No.

Very Important: In all pages with data, there must be a way to get a printable version. To include an icon at the bottom for this purpose is suggested.

Figure nº 3. Modifications on Specimen data page.

Regional Information

Information placed on the FishTrace web page should be given according to the guidelines document approved during Ispra meeting and communicated to all FishTracers.

The three different levels of information (i.e., label, topic and data) should be appear clearly differentiated. For this purpose, it is suggested:

Label name in Verdana 9.5 bold and blue.

Topic name in Verdana 7.5 bold and blue.

Information or data in Verdana 7.5 normal and black.

When Label and Topic have the same name, it should appear only the Label name (Verdana 9.5 bold and blue). Only the filled fields should appear. When no data available (empty field), the field name shouldn't appear.

Specimens Comparison

For validation purposes, it was stressed the importance of being able to print the specimens' information from every area containing data. For instance, *Bothus podas* is a target species from Western Mediterranean, Canary Islands and Madeira. In this case we must be able to obtain a printed list containing all information pertaining to the 15 specimens involved, dully sorted by area (columns) and topics (rows). Information displayed should be printable and must contain all fields/topics (both filled and empty).

Very important: all information under each topic in the same horizontal level (row).

As presented in figure nº 4, it was suggested to allow the search by one Area at the **Specimen comparison table**.

Figure nº 4. Modifications on Specimen comparison table.

Suggestions and changes to the FishTrace DATABASE LOADER:

1) Corrections and additions on FishTrace Database loader are presented in Figures nº 5 and nº 6:

Figure nº 5. Corretion on Database Loader.

Figure nº 6. Addition on Database Loader.

2) In order to validate each specific specimens at each area, at the bottom of every FishTrace Specimen Form is suggested to include the following fields (Figures n° 7 and n° 8):

Signature	Ex.: JA Gonzalez
Date	Ex.: DD-MM-YYYY
Remarks	Ex.: Lacks TFMC voucher ref.

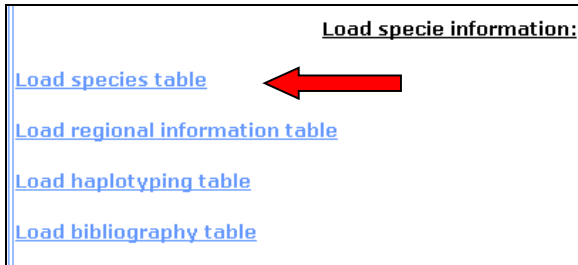


Figure n° 7. New fields to include for validation tasks.



Figure n° 8. New fields to include for validation tasks.

3) In order to validate each assigned species, at the bottom of every FishTrace Species Characteristics Form, is suggested to include the following fields (Figure n° 9):

	Taxonomy/Collections	Genetics
Signature	Ex.: JA Gonzalez	JM Bautista
Date	Ex.: DD-MM-YYYY	DD-MM-YYYY
Remarks	Ex.: Lacks TFMC voucher ref. Voucher No. 4 not yet sent	

Figure n° 9. New fields to include for validation tasks.

4) To extend the maximum number of characters to 500 to fill in the field destined to include remarks on Regional Information Tables (Figure n° 10).



Figure n° 10. Extension of the maximum number of characters to 500 of Remarks fields.

5) To include a link to the *Data Validation Excel File* on the FishTrace Database loader (Figure n° 11).



Figure n° 11. Link to the *Data Validation File*.

6) Bibliography Table (Figure nº 12). Change according to the structure adopted in Madrid (See section 7: Common agreement on the public interface: Bibliography Module).



Figure nº 12. Link to **Bibliography table**.

5. Discussion on Web Interface: Genetics.

V. Verrez presented the compiled corrections/suggestions on Genetics from all partners (in red):

a) General comments for the species page: Include **Back to main page** instead of *Back to main/ Do another search*.

b) Genetics page:

Objectives

1. Determining the nucleotide sequence of a complete mitochondrial gene (cytochrome b: 1141bp) and part of a nuclear gene (rhodopsin: 514bp).
2. Comparing of the sequences obtained
3. Detecting the intra-specific polymorphisms **degree-of-sequence-variation** in the genes analysed
4. Extracting the specific sequences of those apparent populations
5. Extracting genotypic marks of the species
6. Performing a phylogenetic analysis and population structure analysis to confirm the biogeographical distribution of the haplotypes within species

Cytochrome b:

- Use **cytochrome b** instead of cytochrome b
- Use **Cytochrome b DNA sequence**: instead of Cytochrome b DNA data:
- Erase responsible biology
- Use **EMBL accession N°**: instead of EMBL accession numb cyt b:
- It is not a consensus sequence, so erase Consensus sequence cytb (Fasta format):
- Use in all text the following abbreviations for Cytochrome b: Cyt b
- All the sequences should begin by:

```
>EngEnc FishTrace Engraulis encrasicolus cytochrome b
>EngEnc-CB-01 FishTrace Engraulis encrasicolus cytochrome b
>EngEnc-CB-02 FishTrace Engraulis encrasicolus cytochrome b
```

Rhodopsin:

- Use **rhodopsin** instead of rhodopsine.
- Erase responsible biology
- Use **EMBL accession N°**: instead of EMBL accession numb rhod.:
- It is not a consensus sequence, so erase Consensus sequence rhod. (Fasta format):
- All the sequences should begin by:

```
> EngEnc FishTrace Engraulis encrasicolus rhodopsin
> EngEnc -CB-01 FishTrace Engraulis encrasicolus rhodopsin
> EngEnc -CB-02 FishTrace Engraulis encrasicolus rhodopsin
```

If we keep a reference sequence for Cyt b and Rhodopsin it is better to indicate the genetic variation after each reference sequence for the species and not at the end of the genetics page.

Erase the **responsible biology** for the genetics variation.

It would be better to write only **Polymorphisms in cytochrome *b* sequence:** or **Polymorphisms in rhodopsin sequence:**

And after these annotations for the cytochrome *b* sequences and the rhodopsin sequences, a tool would be proposed as a "basket" to retrieve all the sequences in FASTA format of the different individuals studied for this species. This could be proposed as:

Cytochrome *b* sequences retrieval:

```
↑ EngEnc FishTrace Engraulis encrasicolus cytochrome b  
↑ EngEnc -CB-01 FishTrace Engraulis encrasicolus cytochrome b  
↑ EngEnc -CB-02 FishTrace Engraulis encrasicolus cytochrome b
```

Rhodopsin sequences retrieval:

```
↑ EngEnc FishTrace Engraulis encrasicolus rhodopsin  
↑ EngEnc-CB-01 FishTrace Engraulis encrasicolus rhodopsin  
↑ EngEnc-CB-02 FishTrace Engraulis encrasicolus rhodopsin
```

c) Specimen genetics page:

Methodology

Steps:

- Fish tissue
- Sample processing
- DNA extraction
- PCR **Sequencing**
- Sequencing
- Cloning
- Genetic distances

Is it possible to propose a click on genetics only for the specimens which have been sequenced (2/5) ?

On this page for each specimen, the DNA sequences and the amplification conditions should be mentioned as:

DNA sequence:

Cytochrome *b* sequence (Comment: sequence in Fasta format with the same label than on species page, i.e. > EngEnc-CB-01 FishTrace *Engraulis encrasicolus* cytochrome *b*)

EMBL accession N° cyt *b*:

Rhodopsin sequence (Comment: idem Cyt *b*)

EMBL accession N° rhodopsin:

Amplification conditions:

Cytochrome *b* 5' fragment

Cytochrome *b* 3' fragment

Rhodopsin

d) Specimen comparison:

- Use **Genetics** instead of *Genetic parameter*
- Replace Rhodopsine by **Rhodopsin**
- Use **EMBL accession N° cyt b:** or **EMBL accession N° Rhodopsin:**
- Replace CYTb5, CYTb3 and CYTbcomplete by **Cyt b 5'**, **Cyt b 3'** and **Cyt b complete**
- Put the Cyt b 5' above in the part of the Amplification conditions.

As there will be only 2 specimens/5 which will have genetics, the genetics information should be organized to be seen on the PC monitor (current look of this page is presented in the figure n°13).

GENETIC PARAMETER	
DNA sequence	
Cytochrome b complete cds:	agcggaaaagaaagtggggaaagggttttttgggggggyygaaaaaaaaaagaaccccccaaaagtaggggattaaaccagtagggataca atatagatatagccccattttgggggtgggaaagggttttttgggggggyygaaaaaaaaaagaaccccccaaaagtaggggattaaacc
EMBL accession numb cyt b:	testà
Rhod DNA sequence:	aaaaaaaaaa
EMBL accession numb rhod:	test
DNA collection institute:	MMF
DNA collection nb:	34
DNA extraction method:	DNA Isolation Station

Figure n°13. Current look of the the genetics information page to be corrected / modified.

6. Discussion on Web Interface: Main Menu.

The proposed main menu for the FishTrace website is structured as follows:

- Top Navigation Bar:** The Project (1), The Consortium (2), Database load (3), Map of the Web (4), Cont@ct (5).
- Central Content Area:**
 - Header:** Genetic Catalogue, Biological Reference Collections and Online Database of European Marine Fishes (7).
 - Intro:** FishTrace Database contains genetic and taxonomic data concerning more than 200 marine fish species commercialized in Europe.
 - FishTrace Database Section:**
 - Select scientific name: (11) -Select below- | 146 | View
 - Examples: *Beryx splendens* | *Sarpa salpa* | *Pomadasys incisus*
 - Search by common name: (12) Bigeye tuna | View
 - Example: Bigeye tuna
 - Advanced Search: (13) View
 - FishTrace Consortium Members (14):** Logos for RIVO, ipsc, Naturhistoriska riksmuseet, Ifremer, NAGREE, and others.
- Left Sidebar:**
 - 6 **httrace** (with fish and DNA sequence graphic)
 - 8 **Sampling and Taxonomy:** List of species, Sampling areas, Protocols.
 - 9 **Reference Collections:** Voucher location, Protocols.
 - 10 **Genetics:** Genetic catalogue, Protocols, Fish identification tools.
- Right Sidebar:**
 - 15 **Publications**
 - 16 **Meetings**
 - 17 **Events**
 - 18 **News**
 - 19 **Links:** Taxonomy, Collections, Genetics.
 - 20 **Change Language:** Spanish, French, Italian, Swedish, German, etc.
 - 21 ...
- Footer:**
 - 22 FishTrace Project is supported by the European Commission (Quality of Life and Management of Living Resources Programme)
 - Site developed by: JRC- IPSC -European Commission

Figure n° 14. New design proposed by UCM for the Main Menu including suggestions compiled by partners.

As presented in figure n° 14, suggestions compiled by partners and other corrections/modifications raised during the meeting, were taken into account for the design of a new Main Page for the network database web page. The labelled items (1 to 22) are explained below:

1. The project: Extended information about the project. It could be linked to:
<http://infoweb.jrc.it/fishtrace/web/>
2. The consortium: Extended information about partners (Institutions, Scientific responsible, addresses,...) .
3. Database loading site: Direct link to <http://infoweb.jrc.it/fishtrace/>
4. Map of the Web: It could help users to find a concrete item. Example: simulating phylogenetic tree.
5. Contact: email addresses of Scientific responsables.
6. FishTrace Logo.
7. Full Title of the project.
8. Sampling and Taxonomy Menu: Complete “open-out” menu with: List of species included in DB and its distribution, Sampling areas, map, Protocols WP2, etc.
9. Reference Collections Menu: Complete “open-out” menu with: Voucher location: Reference n°, museum, Protocols WP5, invoice forms, etc.
10. Genetics Menu: Complete “open-out” menu with: Genetic catalogue, list and status of the obtained sequences from all species included in DB, Protocols WP3, fish identification tools, etc.
11. Search by scientific name.
12. Search by common name.
13. Advanced search: search by species, family, collection, geographical area, type of sample, etc.
14. Logos of the institutions involved.
15. Publications: List of publications emanating from the project.
16. Meetings: Calendar of previous-future meetings.
17. Link to the Project Tracking & Archive web site.
18. News: Last news concerning FishTrace.
19. Links:
 - Taxonomy: FishBase, NCBI (Taxonomy), FAO, ...
 - Ref. Collections: Museums...
 - Genetics: NCBI (GenBank), EMBL,...
20. Change Language: Direct change of language of the web site by clicking on the country flags.
21. New field to add.
22. Citing EC support: agreed in the Madrid meeting.

The new design for the Main Menu proposed by UCM will be circulated among partners (in Power Point presentation format), including suggestions and changes taken from the common agreements reached at this meeting. Each slide will be followed by a short description about the additions, corrections and changes made on the Database Web Interface (Main Menu). This presentation will include an editable slide (the last one), in order to add/remove/modify it by partners if they find it necessary or if they want to change something there. Once modified, they just have to return it to UCM in order to incorporate all suggestions to the new Main Menu.

Modifications suggested by NAGREF were:

- To add “biological/ecological” at the title: ***FishTrace Database contains genetic, taxonomic and biological/ecological data concerning more than 200 marine fish species commercialized in Europe***
- More precise on the no. of species? OR more general, i.e. ***“concerning the marine fish species commercialized in Europe”***.
- Links to ***web sites of the institutes*** at point number 14.
- ***Picture gallery - Photos of species*** (at point number 21: *New fields to add.*)

Other modification suggested on the Project Page (Overview) is presented at the following box:

Overview

Partners	
• Universidad Complutense de Madrid	- UCM AD
• Joint Research Centre	- JRC
• Swedish Museum of Natural History	- SMNH NRM
• The Canarian Institute of Marine Sciences	- ICCM
• Institut Français de Recherche pour l'Exploitation de la Mer	- IFREMER
• Nederlands Instituut voor Visserij Onderzoek	- RIVO
• Instituto do Mar Portugal - Museu Municipal do Funchal (História Natural)	- IMAR- MIME
• Organismo Autónomo de Museos y Centros, Tenerife	- TFMC
• National Agricultural Research Foundation	- NAGREF RF
• Museum National d'Histoire Naturelle, Paris	- MNHN

7. Common agreement on the public interface.

It was agreed among all partners attending the meeting that reponsible for each set of data (genetic, taxonomy, collections, database) have to report all changes / additions and / or corrections done during the meeting and send to UCM which is the partner in charge of to report the final conclusions reached as well as to report this minutes.

It was also discussed the **Bibliography Module** in the FishTrace Database. Conclusions reached have been reported in the following list prepared by S. Jiménez, M. Biscoito and J.A. González:

Bibliography module description

This module should have the following characteristics:

- 1) data providers have predefined workforms (5), each one containing the necessary fields for the type of document that will be recorded;
- 2) the database will automatically assign a sequential reference number to each record;
- 3) the database is searchable by several fields;
- 4) references in output lists are formatted;
- 5) output lists are printable.

Once in the bibliography main page, the data provider should be asked to choose the type of document to record (journal article, book, book chapter, webpage, CD-ROM, etc.) from a pop-down menu. Then a specific workform would be shown on screen, with only those fields which are appropriate for the document in question.

After having submitted the record, the database should ask if another record of the same type would be submitted or not. In the first case (yes), a new blank workform will appear and the process can go on. If the second option is chosen (no), then the bibliography main page will appear on screen, granting the provider with the choice for another type of document, to edit a previously submitted record, to search the database or to leave the section.

As above mentioned, the module should have an editing tool for corrections and additions to previously entered records, accessed from the main page. Also a search tool should be available, allowing searches by author, title, region, species and provider. Each search results should be printable and references should be formatted according to a standard defined bellow.

Bibliography module structure

List of fields:

- (1) Author(s) (Analytic) [authors separated by //]
- (2) Connective phrase
- (3) Author(s) (Monographic) [authors separated by //]
- (4) Author(s) role
- (5) Title (article/chapter)
- (6) Journal title
- (7) Title (Monographic)
- (8) Year of publication
- (9) Volume/Report ID
- (10) Issue ID
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (14) Edition
- (15) URL
- (16) Version [for e-books, CD ROMs and webpages]
- (17) Series title
- (18) Series volume
- (19) Series issue
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [0001 to 9999, automatically assigned, 4 characters]

List of workforms:

Workform no. 1 – Journal article

- (1) Author(s) (Analytic) [authors separated by //]
- (5) Title (article/chapter)
- (6) Journal title
- (8) Year of publication
- (9) Volume ID
- (10) Issue ID
- (11) Page(s)
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 2 – Book

- (3) Author(s) (Monographic) [authors separated by //]
- (7) Title (Monographic)
- (8) Year of publication
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (14) Edition
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 3 – Book chapter

- (1) Author(s) (Analytic) [authors separated by //]
- (2) Connective phrase
- (3) Author(s) (Monographic) [authors separated by //]
- (4) Author(s) role
- (5) Title (article/chapter)
- (7) Title (Monographic)
- (8) Year of publication
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (14) Edition
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 4 – Report

- (1) Author(s) (Analytic) [authors separated by //]
- (5) Title (article/chapter)
- (8) Year of publication
- (9) Volume/Report ID
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (17) Series title
- (18) Series volume

- (19) Series issue
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 5 – Web Page

- (1) Author(s) (Analytic) [authors separated by //]
- (5) Title (article/chapter)
- (8) Date of publication/access
- (9) Volume/Report ID
- (13) Publisher/Editor name
- (15) URL
- (16) Version [for e-books and webpages]
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 6 – CD ROM

- (1) Author(s)
- (5) Title
- (8) Year of publication
- (9) Volume/Report ID
- (12) Place of publication
- (13) Publisher/Editor name
- (16) Version
- (17) Series title
- (20) Geographic code
- (21) Provider code
- (22) Species code
- (23) Record No. [0001 to 9999, automatically assigned, 4 characters]

Bibliography module examples

Filled workform

Workform no. 1 – Journal article

- (1) Author(s) [González, J.A./V. Rico//J.I. Santana](#)
- (5) Title [Capture of Grammicolepis brachiusculus Poey, 1873 \(Grammicolepididae\) off the Canary Islands](#)
- (6) Journal title [Scientia Marina](#)
- (8) Year of publication [2000](#)
- (9) Volume ID [64](#)
- (10) Issue ID [1](#)
- (11) Page(s) [107-109](#)
- (20) Geographic code [CI](#)
- (21) Provider code [MMF001](#)
- (22) Species code [GraBra](#)
- (23) Record No. [[1 to infinite, automatically assigned, 4 characters](#)]

Output:

[González, J.A., V. Rico, & J.I. Santana \(2000\). Capture of Grammicolepis brachiusculus Poey, 1873 \(Grammicolepididae\) off the Canary Islands. Scientia Marina, 64\(1\): 107-109.](#)

Filled workform

Workform no. 2 – Book

- (3) Author(s) Brito, A.
- (7) Title Catálogo de los Peces de las Islas Canarias
- (8) Year of publication 1991
- (11) Page(s) 230 pp.
- (12) Place of publication La Laguna
- (13) Publisher/Editor name Francisco Lemus, Ed.
- (14) Edition 1a.
- (20) Geographic code CI
- (21) Provider code MMF001
- (22) Species code CenCir/KypSec/BodScr/SpoCan/SpaAur
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output:

Brito, A. (1991). Catálogo de los Peces de las Islas Canarias. 230 pp. La Laguna. Francisco Lemus, Ed.

Workform no. 3 – Book chapter

- (1) Author(s) Robins, C.H./C.R. Robins
- (2) Connective phrase In
- (3) Author(s) Böhlke, E.B./J.E. Böhlke/M.M. Leiby/J.E. McCosker/E. Bertelsen/C.H. Robins/C.R. Robins/D.G. Smith/K.A. Tighe/J.G. Nielsen
- (4) Author(s) role (Eds.)
- (5) Title (article/chapter) Family Synphobranchidae
- (7) Title (Monographic) Fishes of the Western North Atlantic, Part VI
- (8) Year of publication 1989
- (11) Page(s) 207-253
- (12) Place of publication New Haven
- (13) Publisher/Editor name Sears Foundation for Marine Research, Yale University
- (14) Edition -
- (20) Geographic code CI/MAD/WM/CS
- (21) Provider code MMF001
- (22) Species code MurHel/MurAug/GymUni
- (23) Record No. [1 to infinite, automatically assigned, 4 character]

Output:

Robins, C.H. & C.R. Robins (1989). Family Synphobranchidae. In Fishes of the Western North Atlantic, Part VI: 207-253. Böhlke, E.B., J.E. Böhlke, M.M. Leiby, J.E. McCosker, E. Bertelsen, C.H. Robins, C.R. Robins, D.G. Smith, K.A. Tighe & J.G. Nielsen (Eds.). New Haven: Sears Foundation for Marine Research, Yale University.

Workform no. 4 – Report

- (1) Author(s) Santana, J.I./J.A. González/J.A. Quiles/M.F. Marrero/I.J. Lozano/S. Jiménez
- (5) Title Prospección con nasas camarónicas tradicionales en la zona exterior de la Reserva Marina de La Graciosa (Lanzarote)
- (8) Year of publication 2001
- (11) Page(s) 75 pp.
- (12) Place of publication Telde (Gran Canaria)
- (13) Publisher/Editor name Instituto Canario de Ciencias Marinas, Gobierno de Canarias
- (20) Geographic code CI
- (21) Provider code TFMC001
- (22) Species code SynKau/BraBra
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output:

Santana, J.I., J.A. González, J.A. Quiles, M.F. Marrero, I.J. Lozano & S. Jiménez (2001). Prospección con nasas camaroneras tradicionales en la zona exterior de la Reserva Marina de La Graciosa (Lanzarote). 75 pp. Instituto Canario de Ciencias Marinas, Gobierno de Canarias. Telde (Gran Canaria).

Workform no. 5 – Web Page

- (1) Author(s) Froese, R./D. Pauly, Eds.
- (5) Title (article/chapter) FishBase. A Global Information System on Fishes
- (8) Date of publication/access March, 2005
- (9) Volume/Report ID World Wide Web electronic publication
- (13) Publisher/Editor name Froese, R. & D. Pauly, Eds.
- (15) URL www.fishbase.org
- (16) Version 12/2004
- (20) Geographic code All
- (21) Provider code MMF001
- (22) Species code All
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output:

Froese, R. & D. Pauly, Eds. (2004). FishBase. A Global Information System on Fishes. World Wide Web electronic publication. URL: www.fishbase.org, version (12/2004). Access, March, 2005.

Filled workform**Workform no. 6 – CD ROM**

- (1) Author(s) González, J.A./J.I. Santana/J.A. Quiles/S. Jiménez/F. Hernández/A. Diez/J.M. Bautista
- (5) Title Identificación y diferenciación de teleósteos marinos en el área de influencia económica de Canarias 1FD97-1235-C04 MAR, FEDER-MCyT
- (8) Year of publication 2002
- (9) Volume/Report ID CD ROM Multimedia
- (12) Place of publication Madrid
- (13) Publisher/Editor name Instituto Canario de Ciencias Marinas, Museo de Ciencias Naturales de Tenerife y Universidad Complutense de Madrid
- (16) Version 2002 (50 especies)
- (17) Series title PesCanarias
- (20) Geographic code CI
- (21) Provider code ICCM001
- (22) Species code BooBoo/MurHel/PagPag/HeIDac
- (23) Record No. [0001 to 9999, automatically assigned, 4 characters]

Output:

González, J.A., J.I. Santana, J.A. Quiles, S. Jiménez, F. Hernández, A. Diez & J.M. Bautista (2002). Identificación y diferenciación de teleósteos marinos en el área de influencia económica de Canarias 1FD97-1235-C04 MAR, FEDER-MCyT. CD ROM Multimedia. PesCanarias. Version 2002 (50 especies).

8. Discussion on *Model tools for identification and differentiation of fish species (WP9-D11).*

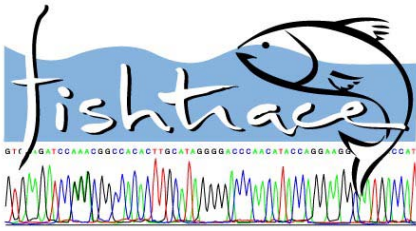
Three tools were proposed and accepted to implement at the home page of the data web site:

- **BLAST** tool to identify the species by the % of similarity
- **Phylogenetic tree** to know, at minimum, the family, or the genus and even the species
- **RFLP profiles** with commercially available restriction enzymes to provide a tool for control laboratories without sequencing the specimen.

9. FishTrace Website design:

In order to show a highly professional identity it was agreed to order the design of the FishTrace Web Interface to a website design company. **Oceanografica**, which has designed the sister Spanish web (PescaBase), has been suggested as a choice. An estimation of costs will be asked to this company, and it will be suggested the collaboration of all FishTrace partners to cover the design by applying the percentage of the total costs in proportion to the original budget assigned to each institution.

Annex VIII: 4th FishTrace Plenary Meeting. Funchal (Portugal), 24-26 October, 2005.



FishTrace Annual Meeting

October 24 – 26, 2005

Funchal, (Portugal)

MINUTES

Objectives:

- To define Project status in relationship with scheduled plan.
- To show the progress of all workpackages and to design a workplan to conclude all of them according to the Technical Annex.
- To implement the FishTrace Database Web Interface.
- To clarify and speed up Data Validation and data submission to the Database.
- To solve problems encountered at different stages of the different workpackages and their interconnections.
- To gather information for next Periodic and Final Reports to the Commission.
- To discuss about next actions to be taken, project prospects, technical developments and new proposals.
- To deal with aspects related to the publication of results obtained from the Project: Possible future scientific publications emanated from the Project, communications introducing FishTrace and Authors policy.

Venue:

Held at the Marine Biological Station of Funchal, (Portugal)

Dates:

October 24 to 26, 2005

Participant list:

IMAR (Host):

Manuel Biscoito
Mafalda Freitas
Rosa Pestana

ICCM:

José A. González
Ignacio J. Lozano

NAGREF:

Grigorios Krey
Panos Leontarakis

UCM:

José M. Bautista
Rafael González-Sevilla

IFREMER:

Véronique Verrez-Bagnis
Marc Jérôme

MNHN:

Patrice Pruvost
Samuel Iglésias

EC-JRC-IPSC:

Philippe Carreau

RIVO:

Hilde van Pelt

NRM:

Michael Norén

TFMC:

Sebastián Jiménez

Agenda

1st DAY

- | | |
|--|---------------------------|
| 1. Welcome to FishTracers | M. Biscoito |
| 2. Meeting introduction | J.M. Bautista |
| 3. Project Status Summary (Milestones: Reached. Delays) | J.M. Bautista |
| 4. Workpackages progress and expected timetable to complete them: | |
| 4.1. WP2: Sampling and Taxonomy | M. Norén |
| 4.2. WP3: Molecular Genetic Identification | J.M. Bautista |
| 4.3. WP4: Biogeographical Genetic Variation | G. Krey |
| 4.4. WP5: Biological Reference Collections | M. Biscoito |
| 4.5. WP7: Validation in Database | J. A. González |
| 4.6. WP6: Database development + WP8: Implementation of www Technologies | P. Carreau |
| 4.7. WP9: Dissemination and Applicability | V. Verrez |
| 5. Presentation of new FishTrace web page prototype | R. González-Sevilla |
| 6. Presentations of the scheduled workplan to finish by each partner: | |
| 6.1. Partner no. 1 (UCM) | R. González-Sevilla |
| 6.2. Partner no. 3 (NRM) | M. Norén |
| 6.3. Partner no. 4 (ICCM) | J. A. González |
| 6.4. Partner no. 5 (IFREMER) | M. Jérôme |
| 6.5. Partner no. 6 (RIVO) | H. van Pelt |
| 6.6. Partner no. 7 (IMAR) | R. Pestana |
| 6.7. Partner no. 8 (TFMC) | S. Jiménez |
| 6.8. Partner no. 9 (NAGREF) | G. Krey |
| 6.9. Partner no. 10 (MNHN) | S. Iglésias |
| 7. Thematic discussion on workpackages progress and expected actions to conclude: | |
| 7.1. WP2 + WP5 | Rapporteur: M Biscoito |
| 7.2. WP8 + WP9 | Rapporteur: P. Carreau |
| 7.3. WP3 + WP4 | Rapporteur: G. Krey |
| 7.4. WP6 + WP7 and an addenda to WP8 | Rapporteur: J.A. González |

2nd DAY

- | | |
|---|--------------------------|
| 8. Special session on Validation: Practical Exercise | Moderator: J.A. González |
| 9. Open discussion: Immediate loading of Database:
The importance of fill-in FishTrace | P. Carreau |
| 10. Brain storming on the new Web page prototype: suggestions | Moderator: J.M. Bautista |
| 11. WP9. Dissemination: Actions for the next period.
Suggestions for increase the presence of FishTrace | Moderator: V. Verrez |
| 12. Thematic discussions on publications | Moderator: G. Krey |

3rd DAY

- | | |
|--|------------------------|
| 13. Brain storming on technological developments from FishTrace | Moderator: H. van-Pelt |
| 14. Brain storming on the follow up of FishTrace through new proposals in EC calls in FP7 | Moderator: G. Krey |
| 15. Next Period Scientific and Financial Reports | J.M. Bautista |
| 16. Next FishTrace meetings | |
| 17. General Conclusions | |

Meeting Minutes:

1st DAY

1. Welcome to FishTracers

The meeting started at 9:00 with a welcoming introduction by M. Biscoito.

2. Meeting introduction

J.M. Bautista explained to all participants the crucial importance of the current meeting in order to finalize scheduled tasks. The aim of this meeting is to strengthen results obtained by a thorough validation and to speed up the filling-in of the Database.

It was proposed and subsequently agreed a list of rapporteurs, in order to draw up the meeting minutes. Minutes from all rapporteurs have been compiled and edited by R. González-Sevilla.

3. Project Status Summary

A short overview about the progress of the project during this period by J.M. Bautista (Milestones: Reached or Delays), based on data from the last Progress Reports sent (September, 2005). Following are the general overview presented and classified according to WP:

WP1

Reached, with the exception of Legal aspects on:

- Public access to the data base
- Patents
- Exchange and lend of biological samples

WP2

Reached, with the exception of a few species that have been replaced by other available species in the area (Figure 1).

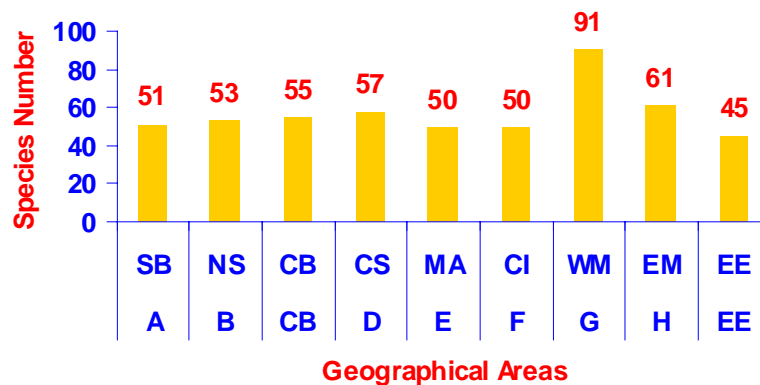


Figure 1. Number of species sampled at each geographical area.

Some extra samples have been obtained, but they are not included in the definitive list since it would be not enough time to fully proceed with them, although the consortium does not wish to waste this extra effort.

WP3

In progress. All tissues samples have been obtained, as well as all DNA extraction have been performed. 69% of definitive sequences have been obtained.

WP4

In progress, around a third of the total of samples to be analysed have not been caught yet. Population studies are also in progress.

WP5

In progress, but almost completed. Progress Indicators for WP5 show that more than 90% Vouchers, Tissues and DNA are already included in collections, and around 60% of otoliths.

WP6

In progress, but nearly completed. Definitive structure have been implemented and the “Online Loading Site” is completely operative at: <http://infoweb.jrc.it/fishtrace/>

WP7

In progress (slightly delayed!). According to recent data it seems that from the data obtained no more than 10% have been already validated. This was probably due to:

- Misunderstanding of the protocol
- Interface problems in the database
- Methodological/Systematics problems

WP8

In progress. A new web page prototype has been implemented at: http://infoweb.jrc.it/fishtrace_int/fishtrace/gb/main/index.jsp

Next development to be implemented are the *Fish identification tools* proposed during the Interface Meeting (Madrid, March 2005):

- BLAST
- RFLPs
- Phylogenetic analysis (Trees)

WP9

In progress. Main actions already done:

- Web site
- Leaflet
- Presentations at conferences / meetings

Main actions planned for the last period of the Project:

- Mini-CD (first prototype in progress)
- Scientific publications
- Links in web sites (GeneBank, FishBase, etc)

In general terms, the Project is going under scheduled plans except for some tasks actually delayed. Also it was presented the actual situation of each milestone and deliverable, as follows:

Milestones	Delivery (month)	Status
1 Standardisation of experimental procedures	4	Reached
2 Legal definition concerning the use of the FishTrace online database	6	Reached
3 The supply of essential biological specimens for collections and analyses	22	Reached
4 Regional databases on taxonomic, biological, socioeconomic and bibliographic data	24	Reached
5 Obtaining of gene sequences	32*	Delayed
6 Genetic catalogue of fishes	32	Delayed
7 Haplotyping of widespread species	36*	In progress
8 Design of a database prototype with selected cases	6	Reached
9 Compilation of the database and standardisation of data and database structure	16	Reached
10 Validation of taxonomic, biological, genetic and reference pooled data	40*	Delayed
11 validation of data for end-users	27	Delayed
12 Implementation and improvement of database interface	32	Delayed
13 Model procedures for end-users for overall detection and identification	31	Delayed
14 Establishment of reference collections	28	Delayed
15 Online networking for the supply and exchange of preserved referenced samples	31	Delayed
16 Dissemination of results	42*	In progress

Deliverables

D1: Standardised methodologies for analyses and biological collections	12	Delivered
D2: Legal rights for access of different end-users to the results.	24	Delivered
D3: Fish samples taxonomically identified to be transferred to the other institutions.	24	Delivered
D4: Catalogue (offline database) of European marine fishes.	40	In progress
D5: Protocols for molecular genetic analysis of teleost species.	24	Delivered
D6: List of identified haplotypes in European widespread teleost species.	36	In progress
D7: Protocols for exchange and access to biological reference collections.	36	In progress
D8: End-users database of European marine fishes.	40	In progress
D9: Reports on validation and missing data.	40	In progress
D10: Prototype of web database and user-interface structure.	36	In progress
D11: Model tools for identification and differentiation of fish species.	42	In progress
D12: Promoting the access to the online database and the use of reference collections.	40	In progress

Conclusions:

- 1) Most data has been obtained
- 2) Fill-in the database is the main effort to conclude successfully FishTrace
- 3) Two kind of problems could persist to supply deliverables. These problems should be detected and solved:
 - a) Small - Scientific
 - b) Bottleneck - Technical
- 4) Dissemination and applicability of the vast amount of information in the database are major issues for the future of FishTrace.

4. Workpackages progress and expected timetable to complete them:

4.1. WP2 - Sampling & Taxonomy

M. Norén informed the Consortium that this workpackage has been REACHED. Table 1 shows total number of WP2 indicators:

Table 1. Number of species sampled and other indicators for WP2 at each geographical area.

	Species	Vouchers	Tissue	Otoliths	Info Sheets	Photos	Oto images	Dissemination
IMAR	50	133	422	34	211	633	0	6
IFREMER	55	273				394		
MNHN	59	327	229	78	6	337	57	7
NAGREF	61	113	717		0		271	1
RIVO	53					240		5
ICCM-CS	57	112	281	100	281	768	8	0
ICCM-EE	45	89	186	73	186	136	9	0
ICCM-WM	91	178	434	153	434	786	18	0
ICCM-CI	50	95	224	71	224	296	7	0
NRM	51		350	14	1	304	1	2

Questions raised from this WP:

- 1) Is the species list (still) accurate? M. Norén suggested that a correction and update of the definitive list of target species is urgently needed.
- 2) Vouchers – which to deposit where? M. Norén indicated that it is very important to solve some questions remaining on the transfer of every voucher to the four Museums in charge.
- 3) Issues with identification? M. Norén raised that there are some vouchers tagged wrong and misidentification has been detected in some of the new species sampled. A few differences have been also detected with synonym common names presented in Fishbase.

4) Other issues. M. Biscoito suggested that some species should receive special treatment due to its high commercial and scientific relevance, e.g. *Solea solea*, *Spicara flexuosa*, *Muraena augusti*... Publications can be identified by those cases in which taxonomy is not correct as corroborated by genetics (specially looking into Fishbase).

4.2. WP3 - Molecular Genetic Identification

J.M. Bautista continued presenting WP3 general progress achieved and the current status of milestones and deliverables. Data compiled from the last Progress Reports indicate that:

- All samples have been obtained and all DNA have been extracted. More tissue samples than expected have been collected due to:

1. Tissue samples not specified in the progress reports (WP3: referenced 01 and 02; WP4: from 03 to 20).
2. Twelve new species added.

- 69% of definitive sequences have been obtained

- Stage with respect to the scheduled plan. At present, the progress in this WP3 is connected with several milestones and deliverables as follows:

WP3 milestones:

M5	Obtaining of gene sequences: delayed
M6	Genetic catalogue of fishes: delayed
M10	Validation of genetic data: in progress
M13	Model procedures for end-users: in progress
M16	Dissemination of genetic results: in progress

WP3 deliverables:

D4	Catalogue (offline database) of European Marine Fishes: in progress
D5	Protocols for molecular genetic analysis of teleost species: reached
D6	List of identified haplotypes in European widespread species: delayed
D9	Reports on validation and missing data: in progress
D11	Model tools for identification and differentiation of fish species: in progress

- Progress and expected timetable to complete WP3:

After three years from the beginning of the Project, more than 1500 definitive sequences from both target genes have been obtained by the Consortium, although the progression observed has not been lineal.

Once the vast amount of biological samples and most of DNA samples have been obtained, the Consortium started producing target sequences. For this reason, September 2003 should be considered as the “kick off” of DNA sequencing.

As observed in Figure 2, during this first period of the project an average of 56 sequences per month were obtained whereas this index has increased during the second period up to 71 sequences per month.

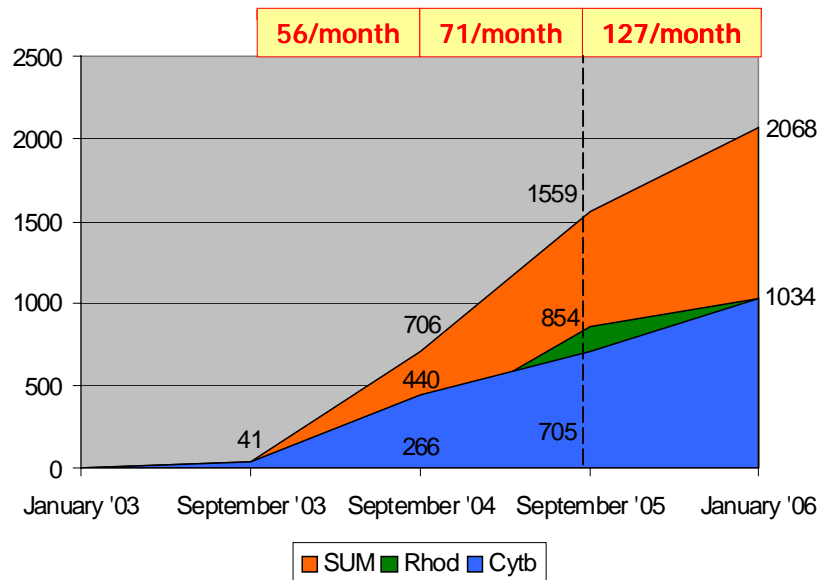


Figure 2. Estimation of sequences/month required for completion the next period.

Although the number of definitive sequences has been linealy growing up, it is still necessary an extra effort to conclude with the scheduled Genetic Catalogue (127 sequences per month are required until the end of the Project for completion!). J.M. Bautista finished his intervention summarizing the main actions to be taken and concusions for the completion of the workpackage on time:

- Actions to complete WP3:

- Obtaining gene sequences: PCR, sequencing, curating.
- Validation.
- Depositing at FishTrace database.
- Depositing at GenBank.

- Conclusions:

- The largest amount of data already obtained (Fig. 2)
- Essential amount of data to be obtained in a short period of time (Fig. 2)
- Fill-in the database is also a main effort to conclude.
- It is necessary to implement a fixed schedule to conclude.
- Increase collaboration among participating labs to solve failed sequences.

4.3. WP4 - Biogeographical Genetic Variation

Presented by G. Krey, the actual status of this workpackage indicates that around a third of the total of samples to be analysed have not been caught yet.

20 samples from each of six species at multiple geographical areas have been analyzed for genetic variation at Cyt-b results in about 660 sequences at present, which represents a 61% of the total sequences planned (Tables 2 and 3). Multiple alignments are in progress. In addition each species sequence will be compared for polymorphisms.

<u>Species</u>	<u>Geographical areas</u>	(Figs. 3 to 7)
<i>Merluccius merluccius</i>	(A, B, C, D, E, F, G, H)	
<i>Micromesistious poutassou</i>	(A, B, C, D, G, H)	
<i>Mullus surmuletus</i>	(B, C, D, E, F, G, H)	
<i>Pagrus pagrus</i>	(C, D, E, F, G, H)	
<i>Pagrus erythrinus</i>	(C, D, E, F, G, H)	
<i>Solea vulgaris</i>	(A, B, C, D, F, H, EE)	

- Protocol followed for comparisons between populations within a species:

I. Nucleotide diversity/haplotype diversity: Amplification and sequence of genes as described in WPs 1 and 3.

II. Genetic distances between DNA sequences: Kimura “two parameter” or Tamura and Nei distance tests. Both tests output a corrected percentage of nucleotides for which two haplotypes are different and take into account different substitution rates between transitions and transversions. In addition Tamura and Nei distance distinguishes between transition rates between purines and pyrimidines.

III. Phylogenetic trees: Neighbour-joining algorithm: allows for unequal rates of molecular change among branches. Branch lengths are proportional to the count of change. The trees are calculated from the original genetic distance matrices using different software (PHYLIP, PAUP)

IV. Population structure analysis. Hierarchical analysis of genetic variance within and between populations (groups) to estimate fixation indices (F statistics) with AMOVA (Arlequin software)

Table 2. WP4 Sequences: Current status

[NA= Not available; ?=Missing Data]

	Completed sequences from each Sampling Area								
	S&BS	NS	C&BB	CS	MA	CI	WM	EM	EE
<i>Merluccius merluccius</i>	20	0	18	12	0	16	4	17	NA
<i>Micromesistius poutassou</i>	0	0	20	0	NA	NA	11	13	NA
<i>Solea solea</i>	6	0	14	0	NA	0	7	17	??
<i>Pagellus erythrinus</i>	NA	NA	0	0	0	19	17	18	NA
<i>Pagrus pagrus</i>	NA	NA	0	0	18	19	19	2	NA
<i>Mullus surmuletus</i>	NA	0	0	0	8	3	9	5	NA

Table 3. Missing Data per Partner:

	<u>NRM</u>	<u>RIVO</u>	<u>IFREMER</u>	<u>UCM</u>				<u>FRI</u>
MerMer	0	20	2	8CS	4CI	20MA	16WM	3
MicPou	20	20	0	20CS	20CI		13WM	7
SolSol	14	20	6	20CS	20CI		13WM	3
PagEry	NA	NA	20	20CS	1CI	20MA	3WM	2
PagPag	NA	NA	20	20CS	1CI	2MA	1WM	18
MulSur	NA	20	20	20CS	17CI	12MA	11WM	15

G. Krey also presented some preliminary results of the population structures with the data obtained at present but taking into account that more data is in progress and should be added at the final analyses.

Figure 3. *Solea solea*:
 AMOVA
 Tamura-Nei distance
 42 different loci; Fst=0.56416
 56% of variation among populations
 P=0.0000
 Pairwise pop Fst (P<0.05):
 EM-C&BB Fst=0.68939
 EM-S&BS Fst=0.60327
 EM-WM Fst=0.24901
 C&BB-S&BS Fst=0.24428
 C&BB-WM Fst=0.60377
 S&BS-WM Fst=0.44708

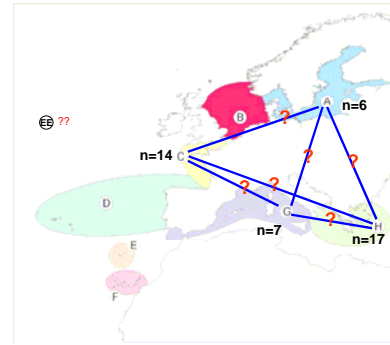


Figure 4. *Merluccius merluccius*:
 AMOVA
 Tamura-Nei distance; 72 different loci
 Fst=0.06977; P=0.0000
 7% of variation among populations
 Pairwise pop Fst (P<0.05):
 EM-S&BS Fst=0.22319
 EM-CI Fst=0.06800
 EM-CS Fst=0.11976
 C&BB-S&BS Fst=0.10381
 C&BB-CS Fst=0.05148
 S&BS-CI Fst=0.07063

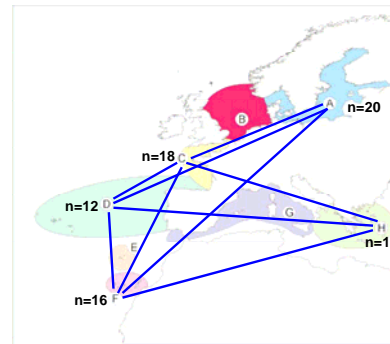


Figure 5. *Micromesistius putassou*:
 AMOVA
 Tamura-Nei distance
 37 different loci
 Fst=0.09995
 10% of variation among populations
 P=0.0098
 Pairwise pop Fst (P<0.05):
 C&BB-WM Fst=0.19100

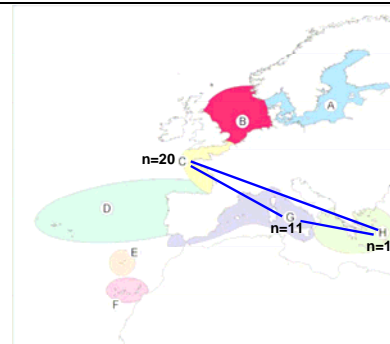


Figure 6. *Pagellus erythrinus*:
 AMOVA
 Tamura-Nei distance
 52 different loci
 Fst= -0.02189
 0% of variation among populations
 P=0.82796

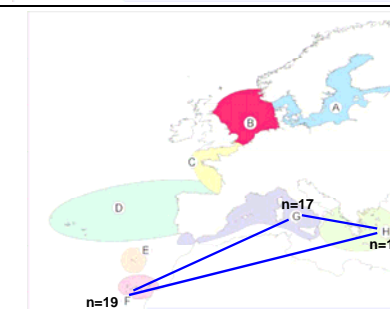
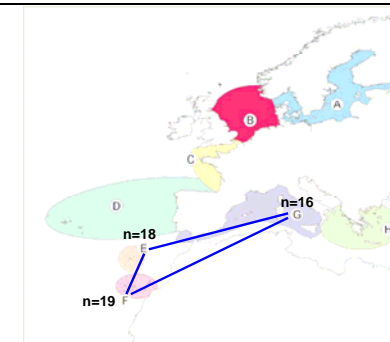


Figure 7. *Pagrus pagrus*:
 AMOVA
 Tamura-Nei distance
 24 different loci
 Fst= 0.15601
 15% of variation among populations
 P=0.0000
 Pairwise pop Fst (P<0.05):
 MA-CI Fst=0.21958
 MA-WM Fst=0.20717



? = indicated that there are not statistically significant differences between populations analysed.

4.4. WP5 - Biological Reference Collections

M. Biscoito, leader of the team in charge of WP5 responsibilities, presented the actual level of completion of the four reference collections created (Figs. 8 to 11): The actual status of the four collections is as follows:

1. Voucher specimens⁽¹⁾: > 90 %
2. Tissues⁽¹⁾: > 90 %
3. Otoliths⁽¹⁾: 50 %
4. DNA^(1,2): > 90 %

(1) Missing data from some labs

(2) DNA not transferred yet

Figure 8. NRM (Swedish Museum of Natural History)

Area A: SB

- Species: 98% done
- Otoliths: 21% done

Area B: NS

- Data to be sent!

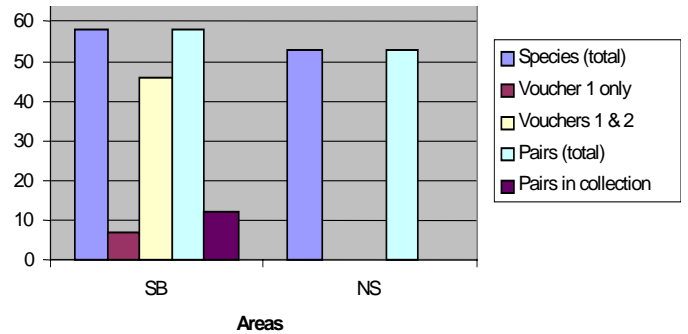


Figure 9. MNHN (French National Museum of Natural History)

Area C: CB

- Species: 93% done
- Otoliths: 34% done

Area H: EM

- Data to be sent!

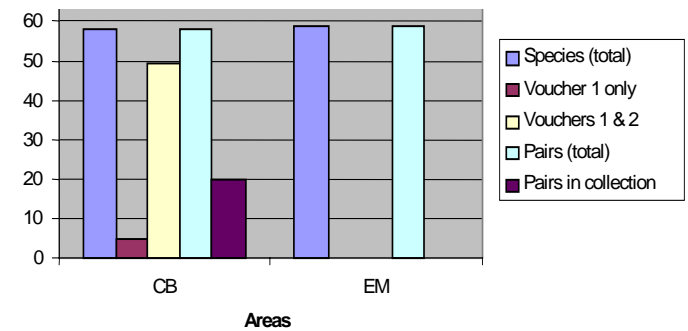


Figure 10. TFMC (Tenerife Museum of Natural History)

Areas D, F, G and EE

- Species: 100% done
- Otoliths: 92% done

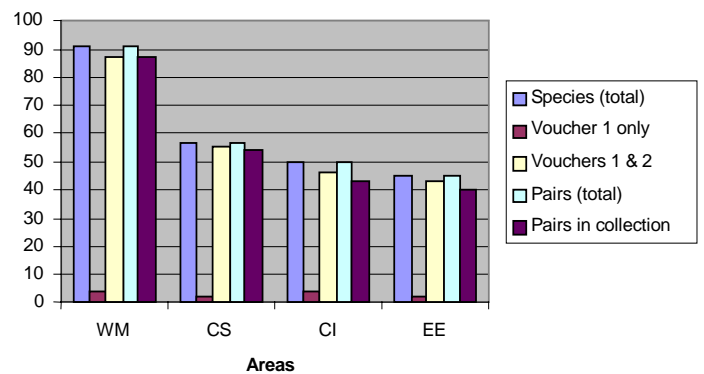
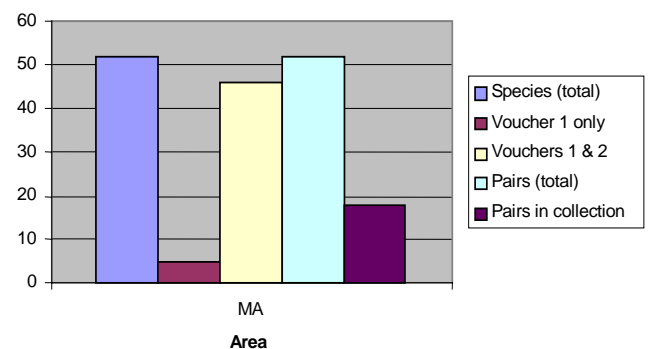


Figure 11. IMAR (Funchal Museum of Natural History)

Area E

- Species: 98% done
- Otoliths: 38% done



4.5. WP7 - Validation in Database

J.A. González presented the status of data validated in the following tables (Table 4 to 8):

“-” = missing data.

Table 4. Validation of each target species at each area (WP2-TAXONOMY, October 19th, 2005):

	Partner	Target species	Species uploaded	Species validated	% done
BS	NRM	48	1	0	0
NS	RIVO	53	-	-	-
BB	MNHN	59	-	-	-
CS	ICCM	56	11	4	7
MA	IMAR	52	22	0	0
CI	ICCM	51	8	2	4
WM	ICCM	91	20	5	5
EM	NAGREF	61	60	60	98
EE	ICCM	45	14	2	4

Table 5. Validation of each target species at each area (WP5-COLLECTIONS, October 19th, 2005):

	Partner	Target species	Species uploaded	Species validated	% done
BS	NRM	48	1	0	0
NS	MNHN	53	-	-	-
BB	MNHN	59	-	-	-
CS	TFMC	56	11	4	7
MA	IMAR	52	22	0	0
CI	TFMC	51	8	2	4
WM	TFMC	91	20	5	5
EM	MNHN	61	-	-	-
EE	TFMC	45	14	2	4

Table 6. Validation of each target species at each area (WP3-GENETICS, October 19th, 2005):

	Partner	Target species	Species uploaded	Species validated	% done
BS	NRM		1	0	0
NS	RIVO	53	-	-	-
BB	IFREMER	59	38 Rhod 33 Cytb	-	64 56
CS	UCM	56	-	18	-
MA	UCM	52	-	20	-
CI	UCM	51	4	35	-
WM	UCM	91	-	44	-
EM	NAGREF	61	48	48	79
EE	UCM	45	-	15	-

Table 7. Validation of each assigned species (WP2-TAXONOMY and WP5-COLLECTIONS, October 19th, 2005):

Group	Species assigned	Species uploaded		Species validated		% done	
		Common fields	Regional info	TAX	COLL	TAX	COLL
NRM	34	-	-	-	-	-	-
IFREMER-MNHN	36	-	-	-	-	-	-
ICCM-TFMC	81	27	19	7	7	9	9
IMAR	36	-	-	-	-	-	-
NAGREF	34	-	-	-	-	-	-
Total	221	-	-	-	-	-	-

Table 8. Validation of each assigned species (WP3-GENETICS, October 19th, 2005):

Partner	Species assigned	Species validated	% done
NRM	28	0	0
RIVO	29	-	-
IFREMER	29	0	0
UCM	105	18	-
NAGREF	30	0	0
Total	221	-	-

Also J.A. González presented a plan to complete validation according to the following table:

Table 9. Expected timetable to complete WP7

Task	Scheduled deadline	Realistic deadline
Self-validation of specimens' information in database, including texts and pictures (fish & otolith) (D8: Definitive list of refined, guaranteed data in DB)	31 DEC 2005	28 FEB 2006
Validation of the assigned species in database. <u>Warning</u> : this task depends on the work done by the other group(s) involved in a particular species (D8: Definitive list of refined, guaranteed data in DB)	31 DEC 2005	30 APR 2006
Uploading of regional information on species to database, including the corresponding bibliographic references (D8: Definitive list of refined, guaranteed data in DB)	31 DEC 2005	30 APR 2006
Last periodical and final report on validation and missing data (D9: Reports on validation and missing data)	30 APR 2006	31 MAY 2006

J.A. González concluded that the current status of WP7 can be considered DELAYED. Partners did not fill in the Excel file for validation purposes when required due to misunderstanding of the protocol and some problems in the Interface of the Database.

4.6. WP6 - Database development + WP8: Implementation of www Technologies

P. Carreau presented the main improvements produced in Database during the last period as well as for the implementation of new technologies in the frame of the Project. These new changes and technological improvements in both workpackages are listed below:

WP6

1) Real time statistical tool: The current status of the Database can be continuously checked by clicking the link placed at the "Database Loading Site" (Fig. 12)

Fishtrace database statistics

Specimens:



Nb of specimen entered (can be not completed):	990	-	-	-	-	Target ?
Nb of validated (signed) specimen:	67	-	-	-	-	Target ?
Specimen with Rhod and CytB seq completed:	154	-	-	-	-	Target ?
Number of CytB sequ completed:	173	-	-	-	-	Target ?
Number of Rhod sequ completed:	208	-	-	-	-	Target ?
Nb of NAGREF specimens:	296	rhod:	122	cyt:	102	Target ?
Nb of IF_MN specimens:	277	rhod:	77	cyt:	62	Target ?
Nb of ICCM specimens:	183	rhod:	0	cyt:	0	Target ?
Nb of RIVO specimens:	156	rhod:	0	cyt:	0	Target ?
Nb of NRM specimens:	1	rhod:	0	cyt:	0	Target ?

Species:



Total number of species entered:	34	Target	231
Total species validated (signature not null):	10	Target	231

Figure 12. Real time statistical tool output.

2) New implemented loading system: You can load information without FTP: a.Enter password; b.Use the browse to select your image; c.Select the file; d.Upload (Fig. 13)

Load with formular
[Transfer reference voucher files](#)
[Transfer reference otolith files](#)
[Transfer other files](#)

Load with FTP
[Copy here reference voucher files \(FTP account\)](#)
[Copy here reference otolith files \(FTP account\)](#)
[Copy here other files \(FTP account\)](#)

- warning: Use FTP to remove/rename files if problem c

Figure 13. Implemented loading system.

3) Visualize all in scroll list: In "Load genetic information", the scroll list contains the FishTrace code, the species and the collection number to check (Fig. 14)

If you complete a file, use the search to display all the information first.

Complete genetic information

You can't create a new FT code here. Morphology data must be entered first

FTcode | Complete name | Collection number

Id_Fishtrace code___: AbuLur-MA-02 | Abudefduf luridus | MMF 35274



Figure 14. Scroll list for select FishTrace code in "Load Genetic Information" sheet.

[Research tools](#)

For species

Select field (10 max - no double)

responsibleBiolName,
id_Haplotype,
id_Bibliography,
consensus_picture,
cons_Oto_picture,
SIMILAR_SPEC,
BIOLOGY,

View and delete data
View data
Delete data
View statistics
FAQ

Select a specie All species View

For specimens:

Fishtrace code<option>
Identification date<option>
Identified by<option>
Voucher fixation<option>
Voucher preservation<option>
Voucher collection institute<option>
Voucher collection nb<option>

Select the fields

Fields appear here

View data
-by region and by species
-by partners

Search by: region EE and specie Thunnus obesus View

Search by: institute RIVO View

Print-----Other search

voucher_fixation	det_name	det_date	voucher_preservation	voucher_col_inst	voucher_col_nb	total_length	fork_length	spec_code
formol 10pc	Iglesias.S	2003-09-12 00:00:00.0	alcohol_70deg	MNHN	2004-0752	95	90	1356
formol 10pc	Iglesias.S	2003-05-18 00:00:00.0	alcohol_70deg	MNHN	2004-0593	270	-1	1874
formol 10pc	Iglesias.S	2003-05-18 00:00:00.0	alcohol_70deg	MNHN	2004-0595	210	-1	1874
formol 10pc	Mouchel.O	2003-05-14 00:00:00.0	alcohol_70deg	MNHN	2004-0596	225	-1	1874
non relevant	Mouchel.O	2003-05-14 00:00:00.0	dry specimen	MNHN	2004-0597	220	-1	1874
formol 10pc	Iglesias.S	2003-05-18 00:00:00.0	alcohol_70deg	MNHN	2004-0637	180	170	44
formol 10pc	Iglesias.S	2003-05-18 00:00:00.0	alcohol_70deg	MNHN	2004-0639	270	250	44

Figure 15. Visualization tool. New interface to display data.

4) To transfer the page directly in Excel sheet: Save in HTML (In menu bar do “save as” and select HTML for “save as type”). Open Excel and open the HTML page you have saved (Fig. 15).

5) Database Quality: Reload all the information before complete/modify a file. Example: you have complete rhodopsine but not CytB. Search for the sample and load it. In the field Cytb there should be -2, then complete it and upload everything. Don't forget: What you see is what you get!

Case of multi sample or multi pictures: You can upload several pictures and sample for one specimen in the *Specimen loading file*:

Add a new tissue sample Upload pictures

IDFishTrace code is the link. Either in the for sample in the main loading page:

You can also directly correct and load sample form here
[Load sample table](#)

You can check the previous entered samples here:

[View previous entered samples](#)
Submit Reset

Check that the DNA sequence has no carriage return (return to next line). A sequences length validation system has been implemented.

Scheduled workplan to finish WP6:

Implementation of a BLAST standalone	-> Dec. 2006
Implementation of RFLP search system	-> April 2006
Implementation of a in-house Blast like research system (depending how the standalone system work)	-> April 2006
Test of the interface with all the system included	-> April 2006

Point to discuss: **Bibliography!**

Future of FishTrace Database:

Convince people that Europe needs a common data repository system to collect data in genetic and fisheries as FishTrace.

- Recommendation for the FP7
- Contact with DG Fish and ICES
- Meeting in Brussels with John Farnel director of DG Fish (seminar for delivering results of FP5-FP6)
- Contact with the other EU project involved in genetic (CodTrace, Homsir, MetaCod)

FishTrace have a future only if validated (also by external actors!)

- How far can we trust the data entered in FishTrace?
- How to make external people from outside trust the information?

How to establish a procedure to validate the FishTrace DB?

- Should involve external actors
- Establish a clear protocol
- Must involve new specimens (where and how to collect them)
- Compare genetic results with reality (eventually taxonomic check)

WP8

1) Interface progress:

- Made part of the modifications asked during the Madrid meeting and the last meeting
- Used as basis for the web designer
- Open an FTP link between Web designer and the server

Presently:

- Redirection page on fishtrace.com
- All the pages are on infoweb.jrc.it
- Some pages (JSP) must remain on the server to interrogate the database (Fig. 16)

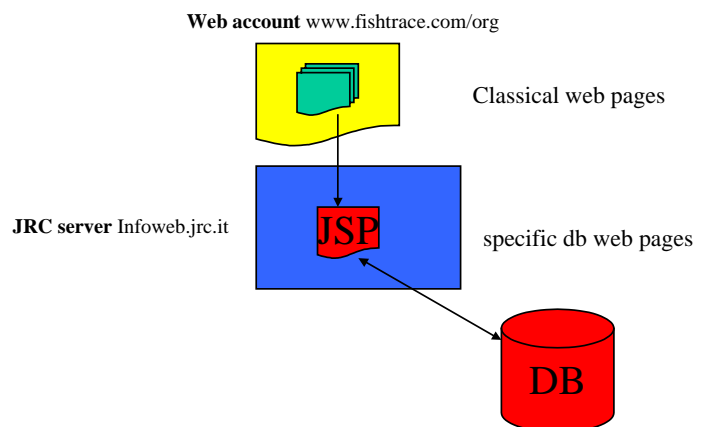


Figure 16. Database and Web link scheme

2) Future design:

- All the classical pages will be moved to fishtrace.com
- Only the page (JSP) necessary to interrogate the database will be on infoweb.jrc.it

3) BLAST investigation: It has been investigated the EU GMO bioinformatic laboratory based at ISPRA. We can use BLAST as a standalone application, but this requires to do an extraction of the database in a special format. It is interesting but not very clean. A second solution is to implement a research system in the database [Smith-Waterman like algorithm] (Fig. 17)

For RFLP there is no algorithm available (new resource E. Scalan arriving at JRC in January).

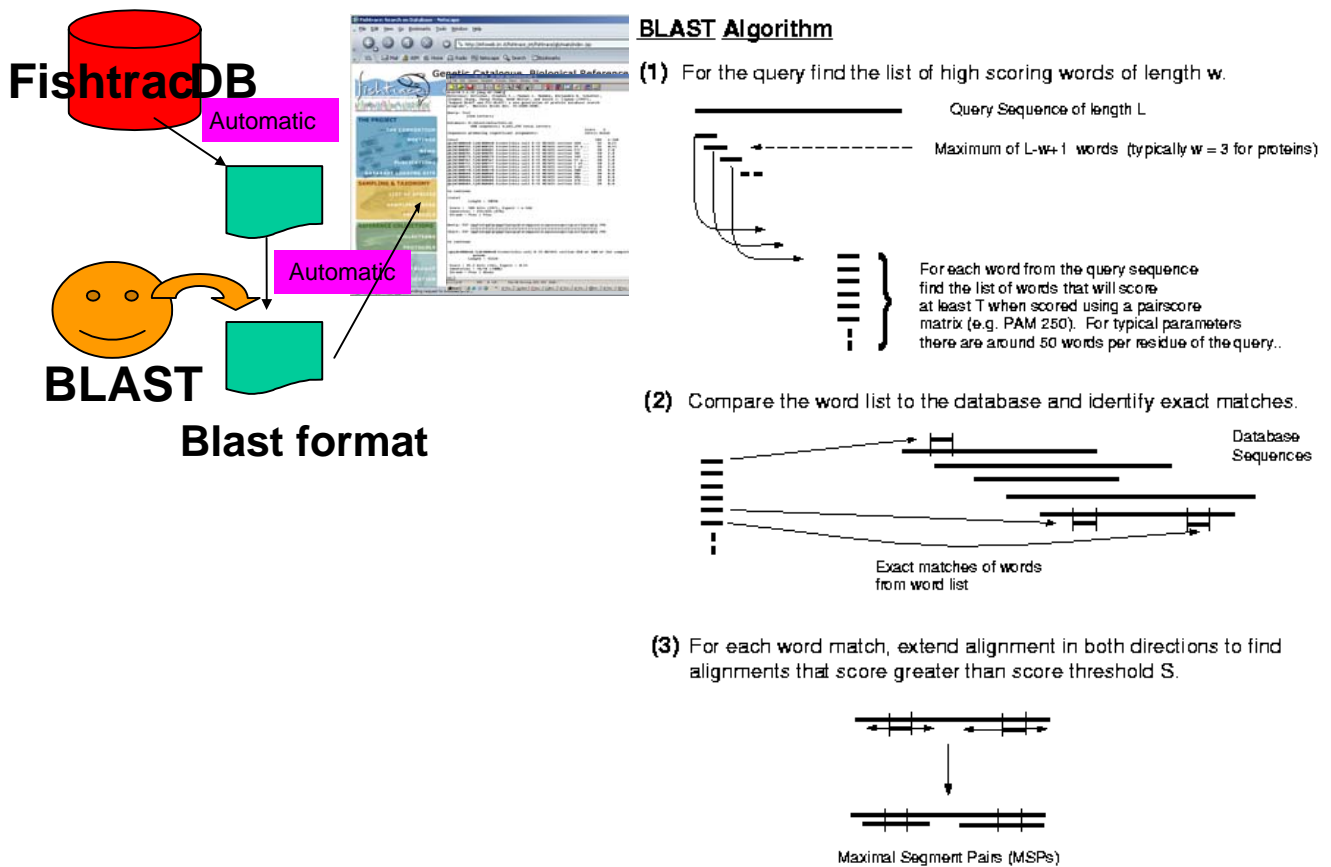


Figure 17. Scheme of the implementation of the BLAST tool on FishTrace Database.

4.7. WP9 - Dissemination + Applications

V. Verrez, partner responsible for dissemination tasks, presented the list of actions already done or in process:

1) Tasks in progress

- Web site: Presented in the next section by R. González-Sevilla.
- Leaflet: Printed in English (100 copies distributed per participant) and also in French. The PDF version will be available through the web site under the “Public Information” menu.
- MiniCD characteristics:
 1. Technology Macromedia Flash © for the animation (Fig. 18)
 2. Autorun for PC and Mac readers
 3. Logo FishTrace on the surface of the mini-CD
 4. Transparent PVC package
 5. Done in the FishTrace languages (text to be translated and sent).
 6. 1000 copies done

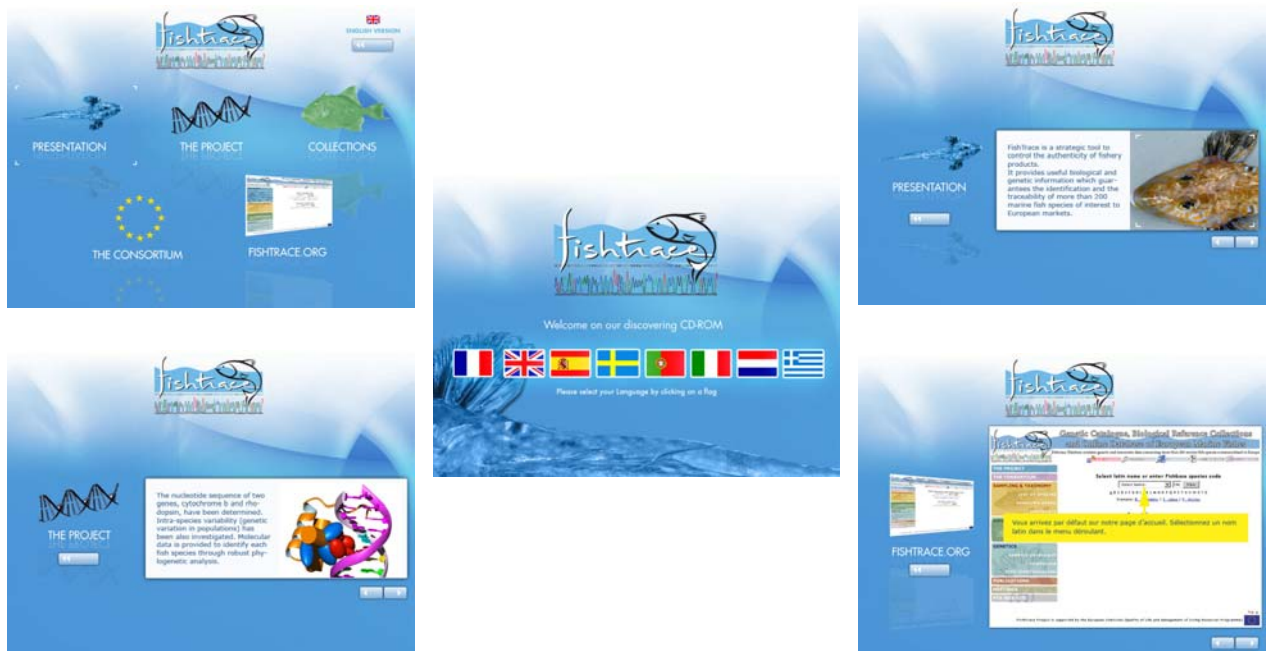


Figure 18. Mini-CD cover pages.

- Presentations at conferences / meetings:

1. Presentation of FishTrace project at The 2005 Glasgow Traceability Seminar, SECC, Thursday 19th May by M.Etienne
2. Presentation at DG Fish March 2005 by P.Carreau
3. Presentation at FISH-BOL First Initiative Meeting, held at the University of Guelph Arboretum, Ontario, Canada, June 5-8, 2005 by M.Norén
4. Petrus Artedi Tricentennial Symposium on Systematic Ichthyology, held at the Royal Swedish Academy of Sciences, Stockholm, Sweden, September 13-14, 2005 by M.Norén

5. Five contributions dealing on taxonomical aspects by partners from ICCM and TFMC at the XIII Iberian Symposium for Marine Benthos Studies, Las Palmas de Gran Canaria, Canary Islands, September 2004

- Scientific publications:

SANTANA, J.I., A.M. GARCÍA-MEDEROS, J.A. PÉREZ-PEÑALVO, R. GONZÁLEZ-CUADRADO, M. GIMENO, M.F. MARRERO, R. DOMÍNGUEZ-SEOANE, M. RABASSÓ, V.M. TUSET, M.F. HERNÁNDEZ, S. JIMÉNEZ, I.J. LOZANO & J.A. GONZÁLEZ. Grandes Serránidos comercializados en Canarias: *Epinephelus*, *Mycteroperca*, *Cephalopholis* (Osteichthyes, Serranidae).

SANTANA, J.I., R. GONZÁLEZ-CUADRADO, M. GIMENO, A.M. GARCÍA-MEDEROS, J.A. PÉREZ-PEÑALVO, M.F. MARRERO, M. RABASSÓ, R. DOMÍNGUEZ-SEOANE, V.M. TUSET, M.F. HERNÁNDEZ, S. JIMÉNEZ, I.J. LOZANO & J.A. GONZÁLEZ. Identificación y diferenciación de especies del género *Seriola* y otros Carángidos afines (*Caranx*, *Lichia*) (Osteichthyes, Carangidae) presentes en aguas de Canarias.

MARRERO, M.F., R. GONZÁLEZ-CUADRADO, M. GIMENO, J.A. PÉREZ-PEÑALVO, A.M. GARCÍA-MEDEROS, J.I. SANTANA, R. DOMÍNGUEZ-SEOANE, M. RABASSÓ, V.M. TUSET, M.F. HERNÁNDEZ, S. JIMÉNEZ, I.J. LOZANO & J.A. GONZÁLEZ. Corvinas y corvinatos oeste-africanos comercializados en Canarias: *Argyrosomus*, *Atractoscion*, *Pseudolithus* (Osteichthyes, Sciaenidae).

GIMENO, M., M.F. MARRERO, R. GONZÁLEZ-CUADRADO, A.M. GARCÍA-MEDEROS, J.A. PÉREZ-PEÑALVO, J.I. SANTANA, M. RABASSÓ, R. DOMÍNGUEZ-SEOANE, V.M. TUSET, M.F. HERNÁNDEZ, S. JIMÉNEZ, I.J. LOZANO & J.A. GONZÁLEZ. Identificación y diferenciación de lenguados (Soleidae) y otros peces planos afines (Psettodidae, Cynoglossidae) comercializados en Canarias.

JIMÉNEZ, S., J.I. SANTANA, M.F. HERNÁNDEZ, M. GIMENO, A.M. GARCÍA-MEDEROS, J.A. PÉREZ-PEÑALVO, R. GONZÁLEZ-CUADRADO, M.F. MARRERO, R. DOMÍNGUEZ-SEOANE, M. RABASSÓ, V.M. TUSET, I.J. LOZANO & J.A. GONZÁLEZ.: Murénidos comercializados y protegidos en Canarias (Osteichthyes, Anguilliformes, Muraenidae).

- Links in web sites (GeneBank, FishBase, etc)

- Next developments for FishTrace applications will be the **Fish identification tools**:

- BLAST
- RFLPs
- Phylogenetic analysis (Trees)

5. Presentation of new FishTrace web page prototype

R. González-Sevilla presented the current appearance of the FishTrace web site prototype (Fig. 19), explaining all menus and sections. The aim of this presentation was to obtain feedback from the Consortium to improve the applicability and contents of the site.

Definitive version of the site is ready at: http://infoweb.jrc.it/fishtrace_int/fishtrace/gb/main/index.jsp

Genetic Catalogue, Biological Reference Collections & Online Database of European Marine Fishes

FishTrace Database contains genetic and taxonomic data concerning more than 200 marine fish species commercialized in Europe

INTRANET LINKS LANGUAGES WEB MAP CONTACT

THE PROJECT

- THE CONSORTIUM
- MEETINGS
- NEWS
- PUBLICATIONS
- DATABASE LOADING SITE

SAMPLING & TAXONOMY

- LIST OF SPECIES
- SAMPLING AREAS
- PROTOCOLS

REFERENCE COLLECTIONS

- COLLECTIONS
- PROTOCOLS

GENETICS

- GENETIC CATALOGUE
- FISH IDENTIFICATION
- PROTOCOLS

Select Scientific Name

-Select below- View

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Examples: [B. splendens](#) | [S. salpa](#) | [P. incisus](#)

Search by common name

(ex: tuna) Bigeye tuna View

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

The FisThrace Project is supported by the European Comission

Figure 19. FishTrace web site prototype. Main page, Main Menu and Searches.

Subsequently, in the second meeting day there were raised suggestions by the Consortium through a brain storming session on the new web page prototype, described in the section n^o 10 of these Minutes.

6. Presentations of the scheduled workplan to finish by each partner:

6.1. Partner no. 1 (UCM)

The status of UCM on sequences obtained for WP3 and WP4 was presented by R. González-Sevilla. Progress achieved:

WP3

- Almost the 75% of definitive sequence obtained for cytb and rhod genes
- Timetable to complete WP3 (Fig. 20):

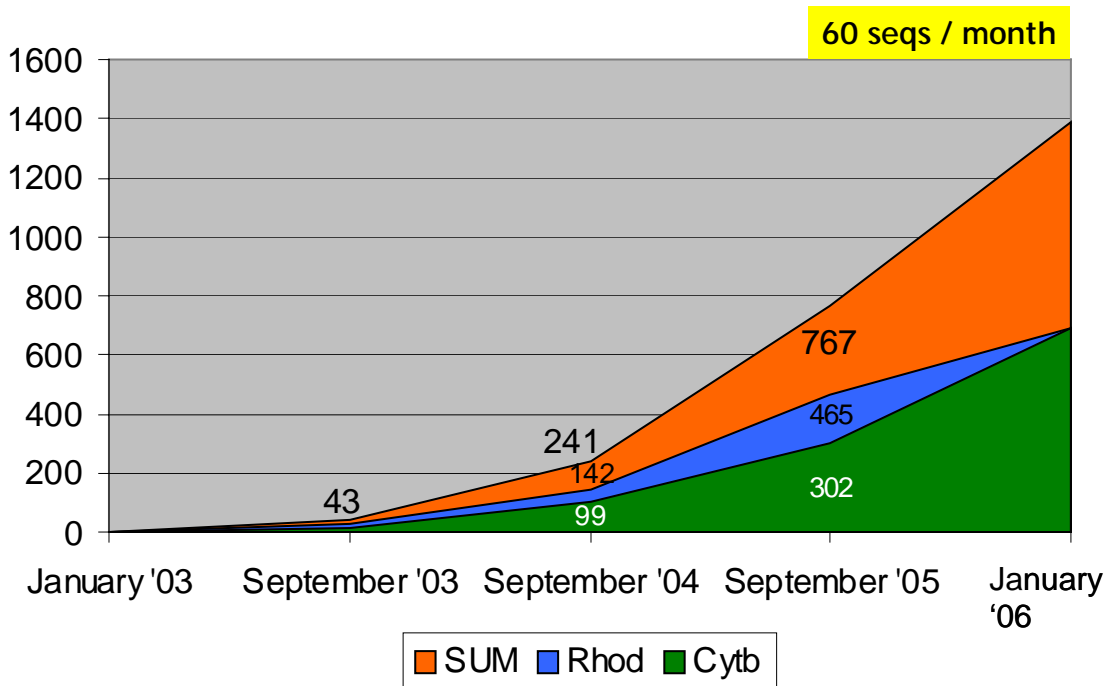


Figure 20. Progress on sequences obtained for both target genes at the UCM labs.

WP4

- 90% of the sequences have been obtained. Population analyses are in process.

Conclusion:

It is required from now three months of intensive work in the bench to:

- Complete the Genetic Catalogue: 100% of expected sequences should be obtained and corrected. (60 sequences per month are required until the end of the Project for completion!).
- Validate: around 600 sequences to complete it!
- Fill-in Database and GenBank Submissions: around 1400 sequences!
- Exchange information: alternative protocols for the amplification of problematic species.

6.2. Partner no. 3 (NRM)

M. Norén briefly showed the actual status of NRM on the different workpackages that they are in charge:

WP2 & WP5

- Species 51
- Tissue 350
- Photos 350
- Information sheets loaded: 1

WP3

- RHO 50
- CytB 148
- Species loaded in DB: 1

6.3. Partner no. 4 (ICCM)

J.A. González presented a timetable for the completion of ICCM tasks and also set deadlines to complete them:

WP2

- Transference of vouchers to Paris and Stockholm museums in co-ordination with MNHN (task to be done at Alicante, Spain): 30 NOV 2005
- Data compilation on common fields and regional information from species: 15 DEC 2005
- Informatic treatment of pictures (fish and otolith images): 15 DEC 2005
- Uploading of all information to database: 31 DEC 2005

WP6 & WP8

- Contribution to database development
- Contribution to the implementation and improvement of database interface

WP9

- Contribution to the dissemination and application of results: 30 JUN 2006

6.4. Partner no. 5 (IFREMER)

The progress achieved by IFREMER were presented by M. Jérôme as follows:

WP2

- 59 listed species
- (*Xiphias gladius* [swordfish] and *Thunnus obesus* [big eye tuna], only 1 specimen sampled)
- Haplotyping: all the 20 individuals of 5 species sampled.

WP3

Table 10. Results obtained and progress on Database loading by IFREMER.

Gene (2)	Sequence obtained (116/gene)	Species analysed (59)	Sequence Loaded in DB	Species loaded in DB
Rhodopsin gene	98	52	76	38
	Problems with sequencing / Problems with PCR <i>Gadidae (Melanogrammus aeglefinus)</i> <i>Soleidae (solea solea)</i> <i>Clupeidae (clupea harengus)</i> <i>Pleuronectidae (limanda limanda & Platichthys flesus)</i> <i>Pleuronectidae (Pleuronectes platessa & Microstomus kitt)</i>			
Cytochrome b gene	63 (92)	34 (46)	62	33
	Problems with sequencing / Problems with PCR <i>5' Congridae (Conger conger)</i> <i>Sparidae (Pagellus erythrinus)</i> <i>Moronidae (Dicentrarchus labrax)</i> <i>3' Belonidae (Belone Belone)</i> <i>Sebastidae (Helicolenus dactylopterus)</i> <i>Mugilidae (Liza ramado)</i> <i>Gadidae (Melanogrammus aeglefinus)</i> <i>Pleuronectidae (Microstomus kitt)</i> <i>Mulidae (Mullus surmuletus)</i> <i>Scombridae (Scomber japonicus)</i> <i>5'-3'</i> <i>Callionymidae (Callionymus lyra)</i> <i>Bothidae (Arnoglossus laterna)</i> <i>Gadidae (Gadiculus argenteus)</i>			

WP4

All the DNA extraction have been done on the five species. Cytochrome b sequences from 20 specimens of *Micromesistius poutassou*, 18 specimens of *Merluccius merluccius* and 14 specimens of *Solea solea* were got and sent to NAGREF. Failure on cytb 5' of *Pagellus erythrinus* and 3' of *Mullus surmulletus*.

6.5. Partner no. 6 (RIVO)

H. van Pelt showed the Consortium the progress and also the problems reached by RIVO to date:

WP2

- 27 specimen uploaded (Database)
- Missing: museum collection number 4th species

Problems:

- Abbreviation MERMER- deleted the data
- Gillraker information is entered correct?
- Database: check if the relation between specimen/tissue sample/photo is correct?

Have to be done: Enter taxomical data for 26 species in database (all collection data are already entered). Send 12 species to the museums. Collect museum catalogue numbers.

WP3

- Rhod sequences: 46 species
- Missing: *Agonus cataphractus* / *Alosa fallax* / *Glyptocephalus cynoglossus* / *Limanda limanda* / *Microstomus kitt* / *Pleuronectes platessa*

- Cytb sequences: 49 species
- Missing: *Callionymus lyra* / *maculatus* / *reticulatus*

Have to be done: Upload genetic info 47 species. Validation. Part (20 nucleotides) of some sequences of rhod are based on single strand information. Find new primers for missing sequences.

WP5

- Collected: 53 species
- Not complete:

<i>Alosa Fallax</i>	2 specimens
<i>Callionymus maculatus</i>	3 specimens
<i>Pollachius pollachius</i>	3 specimens
<i>Ciliata septentrionalis</i>	4 specimens
<i>Ammodytes tobianus</i>	recently identified as <i>Hyperoplus lanceolatus</i>

6.6. Partner no. 7 (IMAR)

R. Pestana briefly show the actual status of Sampling and also the status of the Reference Collections at IMAR (Fig. 21)

WP2 and WP5

- 98% completed (one species missed: *Bothus podas*)

WP7

- 100% Uploaded species!!!!, except the field "photos" and "species table" (Fig. 22)

Remarks: *Muraena augusti* and *Pagellus acarne* are not in FishTrace weblis.

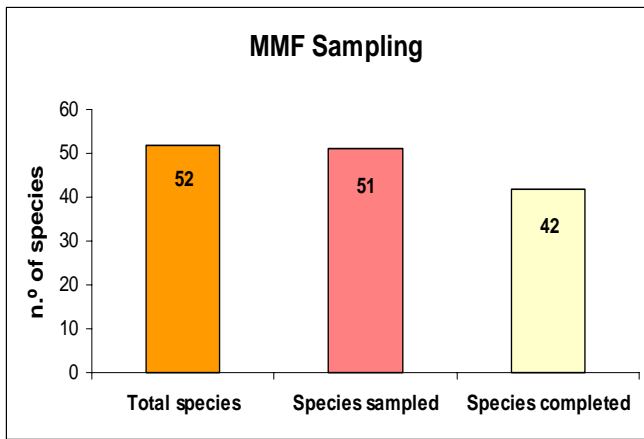


Figure 21. IMAR: Sampling current status.

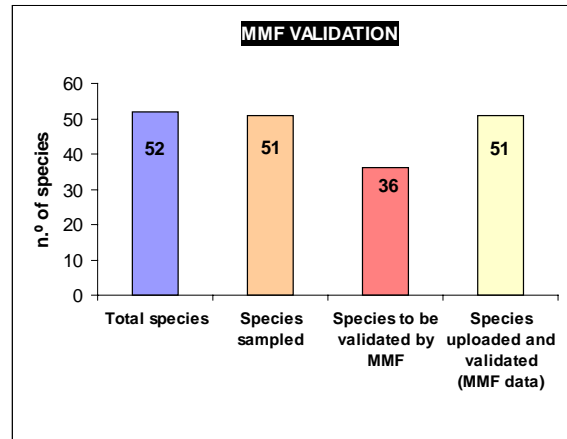


Figure 22. Data validation progress by IMAR.

6.7. Partner no. 8 (TFMC)

S. Jiménez presented a timetable for the completion of TFMC tasks and also set deadlines to complete them:

WP2

- Collaboration with ICCM in the informatic treatment of pictures (fish & otolith images) 15 DEC 2005

WP5

- Preparation of ICCM-TFMC vouchers to be sent to Paris and Stockholm museums (tasks to be done at Las Palmas and Alicante, Spain): 30 NOV 2005
- Uploading of information related to vouchers, tissues and otoliths to database: 31 DEC 2005
- Elaboration of a DNA reference collection at Tenerife museum (ca. 500 DNA samples) Two months after receiving the samples
- Uploading of DNA information to database (ca. 500 inputs to DB): Two months after receiving the samples

WP7

- Self-validation of reference collection data in database (voucher, tissue and otolith data from four areas): 31 DEC 2005
- Self-validation of reference collection data in database (DNA data from four areas): Two months after receiving the samples
- Collaboration with ICCM in elaborating the last periodical and final report on validation and missing data: 30 APR 2006

WP6 + WP8

- Contribution to database development
- Contribution to the implementation and improvement of database interface

WP9

- Contribution to the dissemination and application of results: 30 JUN 2006

6.8. Partner no. 9 (NAGREF)

G. Krey showed the Consortium the progress and also the problems reached by NAGREF to date:

WP2

- FINISHED!
- Action needed: Introduction of *Spicara flexuosa* to FishTrace species list (Table 11).

WP3

- All rhod sequences entered.
- Missing data: cytb for ConCon, BelBel, MugCep, MulSur, UraSca, ScoScr (both specimens), MulBar, SarPil, XipGla (one specimen)
- Projected completion: December, 2005.

WP4

- Missing data: 3 MerMer specimens missed, 7 MicPou, 3 SolSol, 2 PagEry, 18 PagPag, 15 MulSur.
- Projected completion: FRI sequences should be produced by December, 2005.
- Population structure analyses require data from other partners (mainly RIVO and NRM).

Species		
Targeted	Collected	Uploaded
61	60	60
Specimens		
Target	Collected	Uploaded
305	297	292
Species		
Assigned	Completed	Uploaded
61	53	53

Table 11. Results from WP2 (up), WP3 and WP4(below)

6.9. Partner no. 10 (MNHN)

S. Iglésias presented the actual level of completion of their tasks at MNHN:

WP2

- 100% Completed: More than 5 specimens deposited for several species.

WP5

- 100% Completed: More than 5 specimens deposited for several species.
- Around 300 specimens info and photos loaded in Database.
- Otoliths: For many species it is impossible to collect the otoliths without destroying the voucher specimens, so they will not be collected.
- Specimens received at MNHN: 124 from NAGREF, 238 from RIVO, 60 from the TFMC.
- Not received from NRM - IMAR.

7. Thematic discussion on progress of workpackages and expected actions to conclude:

7.1. WP2 + WP5:

M. Biscoito compiled agreements reached after this thematic discussion as follows:

WP2: In relation to the list of species: a) *Solea sp.* has been substituted by *Gymnothorax polygonius* in order to keep the total number of target species included in the sampling area D: CS. b) *Muraena augusti* and *Spicara flexuosa* will be added to the FishTrace list. c) All information concerning new species to include in the list must be sent by email to P. Carreau. A list checking will be done at the end (see the Timetable annexed).

WP5: According to the established protocol for transferring vouchers: a) NAGREF sends vouchers 1, 2 & 4 to MNHN, 5 to NRM, including otoliths and tissue samples and keeps n.3 b) NRM sends voucher n.4 to MNHN. c) Specimens pending from DNA analyses will remain in hands of collectors (this is the case of NAGREF). d) Deadline for depositing vouchers in collection at MNHN, TFMC, NRM and MMF, is January 31st, 2006.

7.2. WP8 + WP9:

P. Carreau summarized comments emerged from this thematic discussion as follows:

WP8: Concerning the web site and the Database: a) The format to be implemented for Bibliography data must follow the scheme agreed during the last Interface meeting (Madrid, March 2005). b) Database is open to include new species in the future. All information related to new species to be added must be sent to P. Carreau. c) Concerning the visualization of the data included in Database, a new format for Validation purposes must be implemented: exportation of data (contained in aligned fields) into an Excel file at *Species* and *Specimens* levels.

WP9: Agreements: a) Apply the standalone BLAST tool. b) Investigate for RFLP algorithms, enzymes selected, etc. c) To implement a tool for generating phylogenetic trees analysing user sequences entered among the whole deposited in FishTrace Database.

7.3. WP3 + WP4:

G. Krey reported the below points raised along this thematic discussion:

WP3: Regarding the difficulty found to amplify *cytb* and *rhod* genes from a few “problematic species” it was agreed: a) From each lab to circulate a list of samples reluctant to amplify at any of the fragments. b) Exchange samples and information on PCR conditions/primers between laboratories sharing the species or similar ones. c) To design new primers that could render optimal amplifications from these special species.

WP4: Agreements taken to complete WP4: a) Try to complete missed samples from geographical areas. b) Send frequently data obtained to the WP4 responsible (NAGREF). c) Repeat those amplifications where discrepancy against a reference has been detected (eg. *Solea solea* from area D:CS). d) To generate a list of samples reluctant to amplify which should circulate among partners.

WP3 + WP4: To test the new technology “Whatman FTA” for keeping DNA samples into Reference Collections.

7.4. WP6 + WP7 and an addenda to WP8:

These summary and agreements were elaborated by the following working group: S. Jiménez (TFMC), I.J. Lozano (ICCM), P. Leontarakis (NAGREF), R. Pestana (IMAR), M. Freitas (IMAR), M.J. Biscoito (IMAR), S. Iglésias (MNHN), P. Pruvost (MNHN), P. Carreau (JRC), J.A. González (ICCM), and reported by J.A. González.

WP6: DATABASE DEVELOPMENT (ALSO SEE AGREEMENTS DERIVED FROM THE THEMATIC DISCUSSION REPORTED BY P. CARREAU)

1.- Signature of responsables on both specimens' and species' tables (Database Loading Site) for validation purposes. Agreements:

A.- To include the following names (and email address) for taxonomy and collections:

IMAR: M.J.Biscoito

NAGREF: G.Krey / P.Leontarakis

MNHN: P.Pruvost / S.Iglésias

NRM: S.Kullander / M.Norén

ICCM-TFMC: J.A.González / S.Jiménez

B.- To include the following names (and email address) for genetics:

UCM: J.M.Bautista / R.G.Sevilla

RIVO: H.van Pelt

NAGREF: G.Krey / L.Favre-Krey

IFREMER: V.Verrez-Bagnis / M.Jérôme

NRM: S.Kullander / M.Norén

2.- Development of a bibliographic table in database. Agreements:

A.- To include a unique bibliographic table according to the document (FishTrace Bibliographic Guidelines) prepared by TFMC-IMAR-ICCM during Madrid meeting (March 11-12th, 2005), which are including six predefined workforms for journal article, book, book chapter, report, web page and CD ROM. The machine must list all references entered in alphabetical order.

Warning: Prior to enter a new bibliographic reference check the list of references already entered, in order to avoid duplications.

3.- What will be the format of the bibliographic references on the species tables and on the regional information table? Agreements:

A.- It will be the classical scientific format, that is author followed by year of publication in chronological order. Example: González (1995a, 1995b), Lozano (1995), Jiménez et al. (1996).

4.- When should we produce a regional information file? Agreements:

A.- When the FishTrace biometric samplings provide small discrepancies in the meristic ranges (formula of fins, gillrakers, etc.) in relation with the general information (common fields).

In this case, enter "FishTrace samplings" in the corresponding bibliographic reference field.

B.- When the regional bibliography, generally more restricted than the general one, provides with new or complementary information on the different data fields.

In this case, enter the corresponding reference citation previously entered on the bibliographic table.

5.- Display all variables from all specimens for a particular species from one to all areas, in order to make easy the validation process. Agreements:

A.- To develop a tool to produce a datafile (derived from the specimens' table) which can be exported as an excel file.

For a particular species, specimens will be listed in columns and variables in rows (both filled and empty fields) being grouped by areas.

6.- Display all regional information fields for a particular species. Agreements:

A.- To develop a tool to produce a datafile (derived from the regional info table) for a particular species which can be exported as an excel file.

WP7: DATA VALIDATION IN DATABASE

Expected TIMETABLE to complete WP7:

Task	NEW deadline
Self-validation of specimens' information in database, including texts and pictures (fish & otolith) (D8: Definitive list of refined, guaranteed data in DB)	28 FEB 06
Validation of the assigned species in database <u>Warning</u> : this task depends on the work done by the other group(s) involved in a particular species (D8: Definitive list of refined, guaranteed data in DB)	30 APR 06
Uploading of regional information on species to database, including the corresponding bibliographic references (D8: Definitive list of refined, guaranteed data in DB)	30 APR 06
Last periodical and final report on validation and missing data (D9: Reports on validation and missing data)	31 MAY 06

FIELD CHANGES IN THE TABLES:

- 1) "Sexuality" (NEW: "REPRODUCTION" – following FishBase description and contents)
- 2) "Reproduction size" (NEW: "MATURITY" – following Fishbase description and contents)
- 3) "Spawning season" (NEW: "SPAWNING" – following FishBase description and contents)
- 4) "Food" (NEW: "DIET RELATED INFO" – NOTE: Not following FishBase description and contents)

ADDENDA TO THE WP8

How the regional info for a particular species should be displayed in the website? Examples:

Regional information:

Canary Islands

Common names

Tableta

Biology

Sexuality (NEW: REPRODUCTION – following FishBase description and contents)

Beryx splendens off the Canary Islands is gonochoric. The overall sex ratio is unbalanced in favor of females. Gonzalez et al. (2003).

Reproduction size (NEW: MATURITY – following Fishbase description and contents)

In the Canaries, the length at maturity was 29.9 cm fork length (females) and 31.3 cm (males), with no significant differences between both sizes which are reached at 4 years old. Gonzalez et al. (2004). Rico et al. (2001)a, Rico et al. (2001) b

Spawning ~~season~~ (NEW: SPAWNING – following FishBase description and contents)

In the Canaries, it spawns in summer and early autumn, peaking over July-August. See field Others. Gonzalez. J.A. et al (2003). Sex and reproduction of the alfonsino Beryx splendens (Pisces).

Berycidae) from the Macaronesian archipelagos. *J. Appl. Ichthyol.* 19(2): 104-108.

Food (NEW: FEEDING HABITS – NOTE: Not following FishBase description and contents)

Carnivorous. It feeds on small fishes (myctophids, gonostomatids, sternoptychids), crustaceans (oplophorids and sergestids) and cephalopods (deepwater species). Durr, J. and J.A. Gonzalez, 2002. Feeding habits of *Beryx splendens* and *Beryx decadactylus* (Berycidae) off the Canary Islands. *Fish. Res.* 54(3): 363-374

Size characteristic

Common sizes

In the Canaries, 18-39 cm fork length (120 to 1396 g total weight). Rico et al. (2001)a, Rico et al. (2001)b.

Interest to fisheries

Fisheries importance

In the Canaries, *Beryx splendens* is a target species of a small-scale demersal fishery off the islands of Gran Canaria and El Hierro, at depths of 500-800 m. Gonzalez et al. (2003), Rico et al. (2001).

Fishing gears

Handlines, bottom droplines, and electric reels. Gonzalez et al. (2003), Rico et al. (2001).

Commercialisation

Forms of utilization

In the Canaries, it is marketed fresh. Gonzalez et al. (2003), Rico et al. (2001).

Other

Some N-S variation in the spawning season was observed in the Macaronesian islands: autumn-winter (Azores), spring (Madeira) and summer-early autumn (Canaries). Gonzalez, J.A., V. Rico, J.M. Lorenzo, S. Reis, J.G. Pajuelo, M. Afonso Dias, A. Mendonça, H.M. Krug and M.R. Pinho, 2003

Cantabric sea

Biology

Sexuality

Beryx splendens off the Azores is gonochoric. The overall sex ratio is unbalanced in favor of females. Gonzalez et al. (2003).

Reproduction size

In the Azores, the length at sexual maturity was 23.0 cm fork length both for males and females, reaching this size at 2.5 years old. Gonzalez et al. (2003), Rico et al. (2001).

Spawning season

In the Azores, it spawns during autumn and winter, peaking in January. Gonzalez et al. (2003)

Size characteristic

Common sizes

In the Azores, 14-42 cm fork length. Gonzalez et al. (2003), Rico et al. (2001).

Interest to fisheries

Fisheries importance

In the Azores, it is a target species of a traditional demersal fishery that occur off the coasts of the islands and on the various seamounts of the Azorean EZ down to 600 m. Gonzalez et al. (2003), Rico et al. (2001).

Fishing gears

Handlines and longlines. Gonzalez et al. (2003), Rico et al. (2001).

Commercialisation

Forms of utilization

In the Azores, it is marketed fresh. Gonzalez et al. (2003), Rico et al. (2001)

2nd DAY**8. Special session on Validation: Practical Exercise**

This exercise consisted in a “real” case of validation. P. Leontarakis presented himself as volunteer to carry on the validation of some data obtained by NAGREF.

RESUMED PROTOCOL FOR THE ONLINE VALIDATION PROCESS:

(Reported by the working group: J.A.González (ICCM), S.Jiménez (TFMC) & I.J.Lozano (ICCM))

STEPS TO BE DONE BY EACH PARTNER AT EACH AREA (INTERNAL VALIDATION OR SELF-VALIDATION)

STEP 1.- Enter the specimens's data (both text and pictures) for a particular species from the sampling forms.

STEP 2.- Check carefully the specimens's data: self-validation.

Every specimen table must be signed and dated by the corresponding responsible for validation purpose at each Institution.

STEP 3.-

A.- The taxonomist/curator responsible of each partner will put an X on the -2 given by default in the Control Data Validation Flow (excel file). We will assume that the DNA collection field will be empty at that moment.

B.- The geneticist responsible of each partner will put an X on the -2 given by default in the Control Data Validation Flow (excel file).

STEP 4.- When necessary/available data, enter the regional information for this particular species.

STEPS TO BE DONE BY THE PARTNER/GROUP IN CHARGE OF A PARTICULAR SPECIES (TRUE VALIDATION PROCESS)

STEP 5.- In case of assigned species, enter the information for the common fields in the species table.

STEP 6.- Check the Control Data Validation Flow (excel file). When all geographical crosses (X) are ok, proceed through the next step.

STEP 7.- Check all specimen's information from the involved areas by comparing with the specific common fields (in the species table). When a problem is detect, a feedback information flow must be started with the responsible for the involved data.

STEP 8.-

A.- The taxonomist/curator responsible for this species will put a DEFINITIVE X on the -2 given by default in the Control Data Validation Flow (excel file): just on the general TAXONOMY and COLLECTIONS columns.

B.- The Expert Group on Genetics will put a DEFINITIVE X on the -2 given by default in the Control Data Validation Flow (excel file): just on the general GENETICS column.

When a problem is detect, the responsible for genetics must be start a feedback information flow with the taxonomits involved.

9. Open discussion: Immediate loading of Database (Presentation by P. Carreau)

To fill-in the Database is a crucial action to allow all FishTrace partners to:

1. Compare info (biology, genetics, etc.)
2. Validate the data.
3. Give the right overview of the progress (remembering that filling-in the Database is the aim of the Project)
4. Solve other problems.

Therefore, to fix a final deadline for data filling, write reports and save data is urgently needed.

Advices:

1. Try to load the data regularly (not before the meeting)
2. Try to have all the info before filling a file.
3. Gain time, no research.
4. Much less chance of error (less chance of reject because of files missing)
5. Remember optional info are represented by (o)
6. Even if there shouldn't be a lot of problems remaining if you meet a problem send an email or call immediately.
7. Contact P. Carreau if you think a little improvement may make gain a lot of time.
8. Be pro-active, you have access to all the information, report immediately all errors coming from the other partners or mine (not implemented, misunderstanding)
9. Everybody is responsible to the data quality.
10. Please answer to P. Carreau emailing. If no answer I consider there are no problem.
11. Keep an eye to the statistics.
12. Consider the Database as the final step.

10. Brain storming on the new Web page prototype: suggestions

There was a big discussion on the new prototype of the FishTrace web site was largely discussed.

All corrections and suggestions raised during this session are reported in the following list:

- 1) "Minor Changes":
 - a) Redraw the map at "The Consortium" menu,
 - b) Correct grammatical mistakes (eg. "Comision", "specie", etc)
 - c) Correct email addresses at "Contact".
- 2) To simplify the process for uploading / delete pictures.
- 3) To reinforce security of the FTP server (actually, you can upload photos without login!)
- 4) Searches:
 - a) To add an "ADVANCED SEARCH" where should be possible search species by typing a "word string" (eg NCBI: *cytochrome b* "AND" *Spicara*, "NOT" *maena*).
 - b) To remove the "A,B,C,D,E,..." list for searches.

Agreements on SEARCH DATABASE:

SEARCH BY SCIENTIFIC NAME (Typing or selecting a species from the scroll down list).
 SEARCH BY COMMON NAME (No characters restrictions)
 SEARCH by BLAST (next implemented tool)
 ADVANCED SEARCH (Selecting several parameters for searches)

Agreements on MENUS:

- 1) Titles: NO CLICKEABLES (Sampling and Taxonomy, Reference Collections and Genetics)
- 2) Information on different themes (Sampling and Taxonomy, Reference Collections and Genetics) will be included into AIMS
- 3) To substitute: Objectives by AIMS
- 4) To substitute: Protocols by STANDARD PROTOCOLS and place each protocol at each thematic block (WP2 protocol at Sampling and Taxonomy block, etc)
- 5) Choose languages: Flags from all countries appears when entering the web site and then the language can be chosen.
- 6) CONTACT: to include official web addresses from all INSTITUTIONS as links
- 7) CONSORTIUM: List all names and to place a picture from the whole group
- 8) DATA VISUALIZATION: under construction

A document gathering all changes and corrections on the web site will be reported by R. González-Sevilla and subsequently sent to the webmaster. Once amended, R. González-Sevilla will inform the Consortium about it in order to take a look to the new format and find new suggestions to improve the web site.

11. WP9. Dissemination: Actions for the next period. Suggestions for increase the presence of FishTrace

V. Verrez explained the next actions to be taken in order to promote the presence of FishTrace in scientific congress, symposia, fisheries managements institutions, etc:

- **Exhibitions and congress:**
 - <http://www.tracabilite2006.com/fr/2006/intro.htm>
 - <http://www.forumlabo.com/>
 - <http://www.iufost.org/>
- **Leaflet, mini CD-ROM and e-mailing ?**
 - Printing of the leaflet in another language ?*
 - Mini-CD-ROM:*
 - how many languages ? (8 + Danish? + Lithuanian? + Polish?)*
 - File for translation send in November*
 - How to distribute leaflets and CD-ROM?*
 - E-mailing?not so easy*
- **Links with the database**
 - Partner's institution will link its web site to the FishTrace database web. Other links?*
- **Publications**
 - This section was discussed in the next section (See section 12)*
- **Protocol for the dispatching DNA or muscle extracts to control laboratories**
 - Technology Whatman FTA*
- **Promotions**
 - via the European Commission*
 - Through fish industry associations*

At the end of the V. Verrez talk on next actions for the dissemination of results, P. Carreau presented the opportunity to do a blind test on the Project data:

“EXTERNAL” Validation

- JRC want to run a project to validate a FishTrace methodology with outside partners.
- Use external labs and involve DG Fish fraud inspector. We will collect fish and send them.
- This could be a strong argument to assess the FishTrace future usability.

P. Carreau asked whether there are any question, and also what is the cost of analyses as well as what the Consortium suggest to be the best way to proceed.

This task was discussed and it was agreed to support the blind test of FishTrace protocols and methodologies by random analysis of samples directly taken from supermarkets, from different brands and to compare results obtained from our database against the consumer information shown in food packages.

This action could assess FishTrace and the already existing species information in Database. Besides, a publication on “mislabelling” and “misidentification” of fish products could arise from this action.

12. Thematic discussions on publications

The group dealt with some proposed publications arising from the Project. After discussing the topics that the publications will cover, it was agreed the following points:

- 1) The “Authors Policy” for publications directly emanated from the project: The name/s of the main author/s in any publication plus the “FishTrace Consortium” as an additional Corporate Author, will be included in all publications containing data from FishTrace.
- 2) UCM will produce a manuscript on “General Description-Applications” of FishTrace, its web site and the Database.
- 3) NAGREF offered as responsible partner for WP4 publications on “Genetic Variation Identification” and “Fish Population Studies”.
- 4) The rest of partners will find tentative publications for the dissemination from the results obtained in every workpackage of the Project. (e.g. Population studies, Fish Phylogeny, Data quality, Normalization of Databases, Technological Developments, etc.)

3rd DAY**13. Brain storming on technological developments from FishTrace**

- 1) Three tools were proposed to implement at the home page of the data web site (see the minutes from the Madrid Interface Meeting in March):
 - BLAST tool to identify the the species by the similarity percentage.
 - Phylogenetic tree to know, at minimum, the family, or the genus and even the species
 - RFLP profiles with commercially restriction enzymes (tool for control lab. without sequencing)
- 2) Diagnostic tools with cytb gene? Desing of universal primers that will cover “all fish species” and also for ichthyoplanktonic species identification.
- 3) DNA Chips: Maybe applying other project for supporting this expensive technological development.

14. Brain storming on the follow up of FishTrace through new proposals in EC calls in FP7
(Presentation and Conclusions by G. Krey)**FishTrace through new proposals****EU-funded research**

FP6 submitted proposal (Infrastructures-5)

“Prospects for integrated initiatives in the genetic identification and cataloguing of fish species in Europe (PISCES)”.

Failed due to coordination shortfalls, number of EU countries involved, limited involvement of institutions with extensive collections, taxonomic experience, and credibility.

The above need to be adequately addressed in any future proposal in the FP7.

Conclusions: There is interests among all partners to continue with FishTrace and to that end it has been agreed to undertake the effort to secure the financing of the expansion of the database through a proposal to be submitted when calls appear for the EU FP7. This however implies that activities related to FishTrace constitute a priority for FP7. Thus, it is important for the consortium to exercise efficient lobbying to relevant EU authorities as is the DG Fish. JRC as an EU center could play a significant role in lobbying and to this the proposal prepared by the JRC partners in validating the FishTrace approach on fish species identification and traceability may provide an important boost. Nevertheless, any new proposal to be submitted to the EU must answer to the criticisms on the "PISCES" proposal that was submitted in FP6. Thus, the consortium will have to be enlarged to include countries with significant fisheries and the possible involvement of additional taxonomic experts and collecting institutes (museums) must also be considered. Furthermore, the importance of solid managerial and coordination scheme for such a proposal, which requires advanced preparation and pro-active involvement of all the consortium partners was emphasized. It was also discussed that the new proposal could include the class of the chondichthies as well as applications such as identification of species in the ichthyoplankton and in stomach contents of fish species of particular interest.

The Marine Conservation Initiative (Gordon and Betty Moore Foundation)

Despite serious threats facing marine environments, the global momentum exists to reverse downward trends and ensure sustainability in marine ecosystems for future generations. The Foundation's Marine Conservation Initiative leverages the sentiments expressed in several recent global ocean reports.² These reports unanimously agree that a drastic shift in how society manages and values ocean ecosystems is necessary to create lasting protection.

As countries begin to shift away from historically ineffective management structures, they need effective and durable ocean management tools that protect ocean resources. These management tools must comprehensively address ecological, economic, and social factors.

The Marine Conservation Initiative is a 10.5-year plan to advance a suite of globally-applicable marine conservation management tools that contribute to sustainable ocean resource management.

Funding strategies for this initiative include:

- Area-based management with ecosystem considerations
- Fisheries management that aligns economic incentives with resource protection
- Conservation-minded technological innovations.

The Foundation is funding in a limited number of locations, primarily in the US and Canada, with a focus on transferring lessons learned. More specifically, the Foundation will support marine conservation in the Gulf of Maine, British Columbia, the California Current, and Fiji. In these four areas, the Initiative will be supporting area-based management with a long-term goal of comprehensive ocean zoning. The Initiative's grant-making in economic incentives (including dedicated access privileges) and technological innovations (including cleaner fishing gear, spatial planning tools, and monitoring and enforcement technology) will also be primarily in these four regions.

FAO

Fishery Resources Division (FIR) is responsible for activities leading to an improvement of the knowledge of the magnitude, distribution and potential harvest of marine and freshwater aquatic organisms used for human food, through fisheries and aquaculture and the factors affecting these magnitudes. It also ensures that this knowledge is used in the most effective way to optimize exploitation of the world's fishery resources and promote the development of national capabilities in this field in developing countries. By extension it is also concerned with the impacts of environmental change on these resources and with the culture of fish and other aquatic organisms. In addition to the various technical programmes carried out by its two Services, on Marine Resources and on Inland Water Resources and Aquaculture, the Division provides technical backstopping to the field programme in collaboration with the Field Operations Division (TCO) by assisting in project formulation and evaluation and by direct assistance to project activities, and has a significant component of direct assistance to requests from public and private sectors for specialized information. The Division also provides the Technical Secretariat for the resource oriented activities of the various FAO regional fishery bodies and collaborates with other UN and non-UN organizations on resource and environmental aspects of fisheries.

Conclusions: It is possible that other international organizations in the field of fisheries and food, such as FAO may have an interest in supporting FishTrace. V. Verrez through contacts with FAO officials will explore this possibility.

National funding agencies

Food Standards agency (UK)

Call for proposals for food authenticity (October 2005)

Is FishTrace ready enough to undertake such responsibility?

Need to prove the effectiveness of methodology/technology developed in identifying fish species in processed fish products and need more data from EE species.

Conclusions: The coordinator encouraged all partners to use the FishTrace developments to seek funds from national authorities to continue with the sampling, taxonomy and genetics in order to keep on enlarging the database, even at the regional level, till funds from EU or international sources become available.

15. Next Period Scientific and Financial Reports

J.M. Bautista explained that according to the CONTRACT, the THIRD periodic report and the corresponding cost statements shall be submitted to the commission within two months (February, 28th) of the end of the period covered by the report (January - December, 2005).

B. Mulligan (EU Scientific Officer for FishTrace), has informed us about the special conditions for this project. FishTrace duration has been extended 6 month and it implies to present this extra-work in a FINAL REPORT. Deadlines:

**3rd Periodic Report by March 1th, 2006
AND
FINAL REPORT by June, 1th, 2006**

J.M. Bautista finished detailing the actions to be taken in order to avoid mistakes:

1. Meet the deadlines in order to amend any possible mistakes on time.
2. Send just and what it is asked for, nothing more nothing less.
3. In further Reports we must meet the context, follow the guidelines and complete ALL sections (person-months, objectives, personnel, work accomplished, ...)
4. PROGRESS REPORTS are most important as they have details of each activity (sampling, taxonomy, genet, ref. collect.)
5. If in doubt, please ask the coordinator in sufficient time to be able to deal with the matter.
6. Add tables, graphics, figures, etc... to each report as these add information and make the report look more professional.

DETAILED DOCUMENTATION TO BE TRANSFERRED FROM ALL MEMBERS (TEAM LEADERS) TO THE COORDINATOR MUST BE SUBMITTED BEFORE END JANUARY

Scientific Reports

STATUS OF THE INDIVIDUAL WORK PACKAGES (typically up to one page per work package: to be submitted by each workpackage leader to the coordinator).

- Summarise the main objectives of the work package for this reporting period.
- Compare the actual work accomplished against the activities planned, explaining any major deviations and the actions taken to remedy them.
- Provide an indication of resources used during the period.

CONTRIBUTION OF THE PARTICIPANTS (typically up to one page per participant)

- Describe the contribution of each participant in sufficient detail to justify its cost claim for the reporting period, noting in particular the personnel and other resources deployed on the project.
- Indicate any significant problems encountered and the actions taken to remedy them.
- Indicate any changes to the scientific team during the period.
- Report on any activity by subcontractors during the period.

Financial Reports (Cost Statements)

- Each *member* shall submit a cost statement in the format specified in part E-1, in duplicate and both originals signed.
- Definition of *Direct cost*, *Indirect cost*, etc, are explained on PART C, Article 17, page 27 of the Contract Model (ANNEX II).

ALL THE FORMS TO BE FILLED AND INFORMATION ARE AVAILABLE AT CORDIS WEBPAGES:

<http://www.cordis.lu/fp5/management/provisions/r-conact-pluri-4.htm>

<http://www.cordis.lu/life/src/projmgmt.htm>

Expected actions to conclude and Timetable settled during the meeting
--

WP2:

Definitive List (corrections): **As soon as possible**

WP3:

1) To solve problems with some species: Circulation of "Problematic Sp. List" among WP3 labs: **December 31st 2005**

2) Exchange information and tissue samples among WP3 labs: **December 31st 2005**

WP4:

To try to obtain specimens not provided for some sp. from some Areas (PagPag, MicPou,...): **December 31st 2005**

WP5:

Deadline for depositing vouchers in collection at MNHN, TFMC, NRM and MMF. NAGREF has to send vouchers 1, 2 & 4 to MNHN, 5 to NRM, including otoliths and tissue samples and keeps no. 3. NRM sends voucher n°4 to MNHN. Specimens pending from DNA analyses will remain in hands of collectors (NAGREF): **January 31st 2006**

WP6:

1) Fill-in Database: **March 1st 2006**

2) Bibliography: A prototype created by P. Carreau, adopting the proposition made in Madrid: **March 1st 2006**

3) Complete species list at Database Loading Site: *Pagellus acarne* (not in FishTrace and *Spicara flexuosa*, *Muraena augusti*: **ASAP**

4) Visualisation for validation: Table by column exportable in Excel format with field alignment: **ASAP**

5) Signature of responsables on tables for validation purposes: **ASAP**

6) Display all variables from all specimens for a particular species: **ASAP**

7) Display all regional information fields for a particular species: **ASAP**

WP7:

1) To include Greek characters and Portuguese and Spanish Special Characters: **ASAP**

2) Self-validation of specimens' information in database, including texts and pictures (fish & otolith) **February 28th 2006**

3) Validation of the assigned species in database. Warning: this task depends on the work done by the other group(s) involved in a particular species: **April 30th 2006**

4) Uploading of regional information on species to database including the corresponding bibliographic references: **April 30th 2006**

5) Last periodical and final report on validation and missing: **May 31st 2006**

WP8 + WP9:

- 1) Implementation of the standalone BLAST (don't put Genbank n^o as reference): **Deadline not defined yet**
- 2) Investigate for RFLP and generate phylogenetic trees: **Deadline not defined yet**

DEALINES SETTLED DURING THE MEETING

DEADLINE TO FILL IN DATABASE:

MARCH 1ST 2006
20:00 h GMT

DEADLINE TO DEPOSIT SEQUENCES AT GENBANK:

MARCH 1ST 2006
20:00 h GMT
(Public release: JUNE 30th)

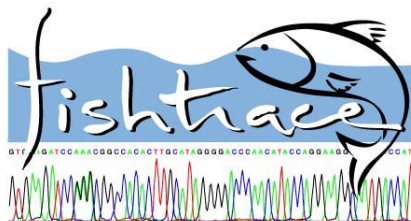
DEADLINE TO VALIDATE DATABASE:

APRIL 30TH 2006
20:00 h GMT

PROJECT FINISH JUNE 30th 2006



Annex IX: FishTrace Concluding Meeting. Kavala (Greece), May 10-12, 2006.



FishTrace Concluding Meeting

May 10th – 12th, 2006

Kavala, (Greece)

MINUTES

Objectives:

- To weight up the degree of completion of the genetic catalogue and the database and to promote any required action to definitively finish the project.
- To assess the degree of data validation in the database.
- To strengthen results obtained by a thorough validation and to speed up the filling-in of the Database.
- To improve the project www technologies and correct any bug at the database or web interface.
- To discuss on next actions for the future of the project: Technological Developments and Dissemination of the results.
- To discuss on possible future scientific publications emanated from the Project.
- To remind all partners to gather the information required for the Final Report to the European Commission.

Venue:

Held at the Fisheries Research Institute of Kavala, Greece.

Dates:

May 10th to 12th, 2006.

Participant list:

NAGREF (HOST):

Grigorios Krey
Panos Leontarakis
Laurence Favre-Krey
Alexis Tsangridis

UCM:

José M. Bautista
Rafael González

JRC:

Philippe Carreau

NRM:

Michael Norén

ICCM:

José A. González

IFREMER:

Véronique Verrez-Bagnis
Marc Jérôme

RIVO:

Hilde van Pelt

IMAR:

Manuel Biscoito
Mafalda Freitas

TFMC:

Sebastián Jiménez

MNHN:

Patrice Pruvost
Olivier Mouchel
Mélyne Hauteceur
Corinne Guchereau

Agenda:

1st DAY

May 10, 2006 - Wednesday

1. Welcome to Fishtracers and meeting introduction. G. Krey
2. Assessment of the progress achieved in FishTrace and degree of completion of each WP. J.M. Bautista
3. Degree of data validation in the database. J.A. González
4. Open discussion on actions to be taken to finish any WP not completed.
5. Open discussion on actions to be taken to complete data validation on time.
6. Short introduction and open discussion on the database. P. Carreau
7. Short introduction and open discussion on the web interface. R. González

2nd DAY

May 11, 2006 – Thursday

8. Open discussion on the web interface (continued from the first day).
9. Open discussion on technological developments from FishTrace and collaborative initiatives with other projects (v.g. Barcoding of Life).
10. Open discussion on WP9 (Dissemination). Suggestions for increase the presence of FishTrace. Thematic discussions on publications and agreements on publications, schedule and endorsement of responsibilities.

3rd DAY

May 12, 2006 - Friday

11. Final Scientific and Financial Reports. J.M. Bautista
 12. Future of the project.
-

Minutes:

1. Welcome to Fishtracers and meeting introduction.

The meeting started at 9:30 with a welcoming introduction by G. Krey.

2. Assessment of the progress achieved and degree of completion of each WP.

A short overview about the degree of completion of the project by the coordinator based on data from the last Progress Reports sent (May, 2006). J.M. Bautista gave a broad outline of the stage of completion of all workpackages and showed the updated list of current Milestones and Deliverables.

Workpackages:

- WP1: COMPLETED, but WP responsible has to produce a *final document* on Standardisation and Legal Aspects of the Project (including agreements on public access to results), for the FINAL REPORT to the Commission.
- WP2: COMPLETED. The definitive list comprise 520 species (including overlapping a each geographical area). Only a few species have been replaced by other available in the area.
- WP3: NEARLY COMPLETED. All samples have been obtained and all DNA extracted. A total number of 2068 sequences is expected and are present 89 % of them have been obtained. Only 11% are still under processing.
- WP4: NEARLY COMPLETED. 97% of the sequences have been obtained. The proper biogeographical analysis is in progress.
- WP5: NEARLY COMPLETED. Transfer of the samples has to be done.
- WP6: COMPLETED. The definitive version of the database has been obtained. Minor polishing mostly in labels has to be done for a proper display of the data at the web interface.

- WP7: IN PROGRESS. 15% of the data have been already validated. A major effort should be done at this level before June. A practical exercise was carried out. Encouragement for the completion of the validation process was a major message from the coordinator.
- WP8: IN PROGRESS. A final version is presented. Some corrections in spelling or text should be done in brief. Tools have been implemented: Blast and RFLPs. This last should be improved. The phylogenetic tool for molecular identification should be implemented.
- WP9: IN PROGRESS. This is a WP where activities will be performed even after finishing the project with the objective of promoting the database and the Web site.

Milestones: Most milestones have been completed or are nearly completed. The following milestones are finishing or in the last step of completion:

- M5.- Obtaining of gene sequences (finishing).
- M6.- Genetic catalogue of fishes (in progress).
- M7.- Haplotyping of widespread species (finishing).
- M10.- Validation of taxonomic, biological, genetic and reference pooled data (in progress).
- M11.- Validation of data for end-users (in progress).
- M13.- Model procedures for end-users for overall detection and identification (in progress).
- M15.- Online networking for the exchange of preserved referenced samples (finishing).
- M16.- Dissemination of results (in progress).

Most deliverables have been produced. The following are under way to be generated in brief:

- D4: Catalogue (offline database) of European marine fishes.
- D6: List of identified haplotypes in European widespread teleost species.
- D9: Reports on validation and missing data.
- D11: Model tools for identification and differentiation of fish species.
- D12: Promoting the access to the online database and the use of reference collections.

In general terms, the Project has been completed except for some tasks actually in progress. Practically all target data has been obtained, the database and the web interface have been produced and reference collections have been created. J.M. Bautista also explained to all participants the crucial importance of the current meeting in order to finalize scheduled tasks successfully, and now, the most important action for concluding successfully this project is to fill-in the database and then, all data should be validated. Besides, the dissemination and applicability of the vast amount of information gathered in the database are major issues for the future of FishTrace.

3. Degree of data validation in the database.

J.A. González, partner responsible for the WP7 (Validation of the data inserted in database), inform the consortium about the shortage of data validated in the database. The main problem has been the understanding of the validation process by some partners. This represents problem that must be solved in order to reach a successful end for the project. After his intervention, all participants were agreed in to repeat a "practical exercise" following the whole online validation process. P. Leontarakis carried out this practical exercise.

At the end of the exercise, the consortium agreed to differentiate two levels of validation: A first level, updating the EXCEL file used for the control of the data validated, and a second an FINAL LEVEL of VALIDATION, that will be effective once all "-2" are removed from this EXCEL file.

During this session, P. Carreau invited all participants to ask for help if someone find a problem during the validation process. There are different ways to do that: send an email, place your problem at the Discussion Board, etc.

Finally, the WP7 responsible remind to all that the best way for the validation of the data is to have a big amount of that, in order to compare reliability.

4. Open discussion on actions to be taken to finish any WP not completed (Conclusions)

Concerning WP3 and WP4: For any problem remained for PCR or sequencing it is suggested a close view to the database to know how other colleagues have resolved specific experimental difficulties.

Some problems detected are:

- Unusual STOP codons. Solution: Translate to protein sequences using the Vertebrate Mitochondrial code.
- Differences between specimens 01 and 02. Solution: Compare both with the corresponding species from other geographical areas.
- Some partners ask for the possibility of download ALL sequences in FASTA format in order to accelerate the validation process.
- The rest of problems concern amplification protocols failed and the contamination of some samples. Solution: To share samples and knowledge with other partners involved in these WPs.
- Incomplete sequences. Solution: Include the sequence as it has been obtained in the database. There will be time after the final project date to try complete data.

The coordinator encouraged the consortium to detect and circulate among all partners, every problem detected at EACH WP, as soon as possible. The use of the Discussion board with implemented email was encouraged.

5. Open discussion on actions to be taken to complete data validation on time.

After this session, the main conclusion was that all partners have to ENTER DATA on database and to COMPLETE THE SELF-VALIDATION process. Only reaching these step first, the definitive validation can be done.

6. Short introduction and open discussion on the database.

P. Carreau presented the main changes produced in Database during the last period as well as for the implementation of new technologies in the frame of the Project.

A document compiling all suggestions, corrections and new contributions to the Database will be produced.

7. Short introduction and open discussion on the web interface.

R. González presented the final version of the FishTrace web site, looking around all menus and sections. The aim of this intervention was to present the Interface to the Consortium to obtain feedback of corrections and suggestions to improve the applicability and contents of the web site. Definitive version of the site is ready at: www.fishtrace.org.

8. Open discussion on the web interface (continued from the first day).

An exercise with all participants was done to made corrections in the web interface. All suggestions, corrections and new contributions to the Database was recorded and will be implemented in due time by either by the webmaster at UCM or by Philippe at JRC. The previous outdated information on FishTrace at <http://infoweb.jrc.it/fishtrace/web/index.htm> should be removed.

9. Open discussion on technological developments from FishTrace.

- P. Carreau remind the consortium the proposal about to carry an "external validation" by other institutions. It will be like a "feedback" control of the FishTrace results in private laboratories.

- **RFLPs tool:** An agreement was done to select few of our sequences at two levels: 1) Populations (specimens) and 2) reference sequence (species). Then, the results will show only the differential diagnostic cut and will be presented in a table showing data from the several selected sequences, with the corresponding fragment sizes.

- **Phylogenetic tool:** This should be finally implemented by using the most reliable phylogenetic method for the type of sequence analysed. The tool will produce a tree and a distance matrix including the sequence loaded. A document with the technical details has been produced and sent to Philippe for implementation.

- **Morphological Tool:** Implementation of a new tool for searches in FishTrace Database by Morphological Data. A document based on Patrice Pruvost design with the technical details has been produced and sent to Philippe for implementation

...and collaborative initiatives with other projects (Conclusions).

- The coordinator and other partners involved planned to assist to the next FishBoL meeting in Paris, in order to study the possibility of link FishTrace with the Barcoding of Life project. Different levels of collaboration will be also studied before the presentation of FishTrace at this meeting.

10. Open discussion on WP9. Suggestions for increase the presence of FishTrace. Thematic discussions on publications Agreements on publications, schedule and endorsement of responsibilities.

In this session it was suggested to define the main TOPICS emerging for publications from the FishTrace project, and the RESPONSIBLE partner to write a draft on each topic at the end of the project.

The agreement reached are shown in the following table:

TENTATIVE LIST OF PAPERS FROM FishTrace AND RESPONSIBLE PARTNERS to produce the first draft

TOPIC / TENTATIVE TITLE	RESPONSIBLE PARTNER(S)
General presentation of the FishTrace Database	UCM
Morphological data and reference collections	IMAR
Database structure and data quality in FishTrace	JRC
General fish phylogeny based on cytb + Rhod vs COI	NRM + IFREMER
Genetic variation in European populations (6 species: 6 papers or just 1?)	NAGREF
Quick method to identify fish species by DNA sequencing	RIVO
New methodology for fish collections	MNHN
Species or group particularities identified in the project:	
<i>Solea</i> group + pleuronectiforms	ICCM + NAGREF
Centracanthidae (Short communication)	NAGREF + ICCM
<i>Trisopterus</i>	NAGREF + NRM
<i>Sarda sarda</i>	UCM + ICCM
<i>Alosa alosa</i> vs <i>Alosa fallax</i>	IFREMER + NRM

DEADLINE TO PRODUCE A REPORT (Contents, abstract, potential journal to submit it, authorship, etc) BY 30 SEPTEMBER 2006 (but if is prepared before this date it can start to circulate).

11. Final Scientific and Financial Reports.

The coordinator explained the guidelines to follow in order to produce the corresponding Final Scientific and Financial Reports to EC. At the end of this session R. González sent this guidelines to all partners in a PowerPoint file.

J.M. Bautista recommend the consortium to follow these instructions and to be aware and **meet the deadlines (July 18th and 31st)**.

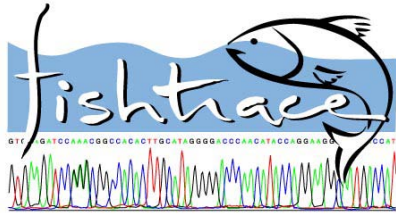
12. Future of the project.

An open discussion about the future of FishTrace was followed by several suggestions arised from different partners. The main conclusions were the following:

- G. Krey: Possibility to link with other projects: In March there will be a meeting on Fisheries of the Mediterranean and it should be a nice opportunity to find connections with other projects.
- M. Norén said that if FishTrace wants to join FishBoL, it must be a permanent infrastructure.
- J.M. Bautista would like to continue including more species (for example, including PescaBase sequences).
- J.M. Bautista will present FishTrace to some EU contacts (DG Fish, DGSanco).

Annex X: Sampling and Taxonomy (WP2) Protocol.

FISHTRACE WP2 PROTOCOLS AND PROCEDURES



Version 4, 7 June 2003

Sven O Kullander and Michael Norén

Department of Vertebrate Zoology
Swedish Museum of Natural History
POB 50007
SE-104 05 Stockholm
SWEDEN

E-mail sven.Kullander@nrm.se, michael.noren@nrm.se

The tasks of WP2 activities are:

1. Collect specimens
 2. Collect regional data on the specimens, including common names, field marks, biology, size, fisheries, and forms of use, transformation, etc.
 3. Identify species
 4. Sample muscle tissue and otoliths
 5. Transfer tissue to molecular labs and otoliths to collections
 6. Incorporate specimens in biological collections
 7. Evaluate taxonomy of each species
 8. A regional technical list of relevant publications on taxonomy, distribution, ecology and biological parameters
 9. Submit all data to FishTrace database
-

Table of Contents

3	Fish sampling protocol
3	Sampling site
3	Field sampling protocol
3	Specimen numbers
3	Specimen treatment
3	Specimen quality
4	Specimen size
4	Special conditions for large specimens
5	Photography
6	Specimen tagging
7	Tissue sampling
8	Preservation of voucher specimens
9	Otoliths
10	Identification
10	Decision on nomenclature and systematics
11	Field morphometric data
11	Laboratory morphometric data
12	Species information protocol
12	Species data
12	Basic morphology
12	Biological information
12	Common names
12	Threat status
13	Fisheries information
13	Socioeconomic information
13	Bibliography

Fish sampling protocol

Sampling site

Fish specimens may be obtained fresh in markets or by active fishing efforts.

Field sampling protocol

A protocol for field data recording is available and should be used. The protocols are archived by each partner. Copies of field protocols must accompany specimens/tissue samples/otoliths sent to reference collections/molecular laboratories.

Specimen numbers

For each geographic area is required at least 5 specimens of each target species as specified in the list of target taxa. For haplotype analysis, performed on certain taxa, a total of 15 specimens are required for each geographic area (5+15).

Specimen treatment

The five specimens should be processed minimally as follows:

- Specimen #1: Photograph
Muscle tissue: 2/3 samples, one for analysis, one for backup, one for reference collection
Voucher: send whole specimen to reference collection
- Specimen #2: Photograph
Muscle tissue: 2/3 samples, one for analysis, one for reference collection, one for backup
Otoliths: extract both sagittal otoliths, photograph right one and send to reference collection
Voucher: send whole specimen to reference collection
- Specimen #3: Photograph
Otoliths: extract both sagittal otoliths, and send to reference collection
Muscle tissue: 1/2 sample, keep it preserved
Voucher: send whole specimen to reference collection (study specimen)
- Specimen #4: Photograph
Muscle tissue: 2 samples, one for analysis, one for backup
Voucher: send whole specimen to reference collection = NRM
- Specimen #5: Photograph
Muscle tissue: 2 samples, one for analysis, one for backup
Voucher: send whole specimen to reference collection = MNHN
- Haplotyping: Specimens #6-20: Take muscle tissue; dispose of specimen as preferred

Specimen	1	2	3	4	5
Photograph	+	+	+	+	+
Tissue Analysis	+	+	-	+	+
Tissue Backup	+	+	+	+	+
Tissue Voucher	+	+	-	-	-
Otolith	-	+	+	-	-
Otolith photo	-	+	-	-	-
Voucher	+	+	+	+ (NRM)	+ (MNHN)

Specimen quality

Specimens should be sampled fresh as far as possible. Frozen specimens or specimens that have been on ice for several days are less suitable for molecular analysis. Specimens may be whole or gutted, but intact specimens are preferred. All material must be collected for this project, and old specimens cannot be used.

Specimen size

For tissue and otolith sampling, as well as for systematic analysis, adult specimens showing diagnostic marks are always preferable. For some taxa, it may be more realistic to sample young specimens, or to use adults as primary tissue source, and juveniles as backups.

Special conditions for large specimens

When large size, collection space limitation, or fishing/conservation restrictions make it expensive or difficult to preserve available specimens, e.g., of marlins or swordfish, one of three strategies is possible:

1. Take a photograph and basic measurements of the specimen, and 3 tissue samples; release the fish/return to fisherman
2. Take a photograph and basic measurements of the specimen, and 3 tissue samples; keep the head (and tail if possible) as voucher, and extract the otoliths from the head
3. Inquire with the major museums (NRM and MNHN principally), whether they may consider the transportation worthwhile, and if so, perform the usual photography and tissue and otolith sampling. The recipient will then pay the fish and the transport.

Photography

All specimens of a species must be photographed. Additional specimens may be photographed as well, to cover sexes, reproductive status and ontogenetic stages displaying the variability within a species. Specimens photographed should be as fresh as possible

Photos must be digital color images, at a resolution of 1024X768 or more. Photos must be saved in minimally compressed jpeg format (Resolution 'Fine') or in TIFF or similar non-destructive format.

Photographs must be taken of the **left** side of the specimen, except certain flatfishes. In flatfishes, the eyed side should be shown, with the gill opening down. Photographs should be taken in lateral aspect, and when relevant of other aspects (flatfishes: right and left side; anglerfishes: dorsal and ventral aspect). Fins should be spread as much as possible using alcohol-soaked cotton swabs and supports when relevant.

Photographs should be taken in ambient outdoor light, in shadow, and without flash. The background should be a neutral grey. In the case of uneven light conditions, e.g., in early morning or evening sun, the fish should be directed so that more light reaches the dorsum than the ventral parts.

The photo should show:

1. Fish plus a label showing the FishTrace specimen number and a ruler

Image files must be labeled with the specimen number plus a qualifier 'specimen' plus the appropriate file extension, e.g.,

MulSur-SB-05-specimen.jpg

Specimen tagging

All specimens should be tagged with numbered tags and the tag number indicated on the tissue sample and otolith tubes. The tag number may differ from the FishTrace specimen number. Tagging should be done before photography and sampling. Floy tags should be inserted through the lower jaw, between the dentaries, so that the locking tip is inside the mouth, and the colored tag can be hidden underneath the gill cover.

Floy Tags, and the gun required for insertion, can be ordered from Floy Tags, Seattle. Write to Betsy Conrad betsy@floytag.com for Price List and Catalog. The FD-94 is adequate for general purpose tagging.

NRM uses FD-94 tags, color WHITE, black text with sequential numbering: NRM T 12345 (The T is there to distinguish from other NRM numbers). The Pistol Grip gun is adequate for most users, be sure to add an extra needle.

Each specimen should also be assigned a unique Fishtrace code. The Fishtrace code is constructed from the first three letters of the generic name, the first three letters of the specific name, two letters denoting geographical area, two digits denoting the specimen number. Don't worry – it sounds more complicated than it is!

Species with identical first three letters in generic name and species epithet require ad hoc code, constructed based on the first differing letter.

Example: the second specimen of *Mullus surmuletus* collected in the western Mediterranean will get the following code:

MuISur-WM-127

The geographical codes are as follows:

SB	Skagerrak and Baltic Sea
NS	North Sea
CB	English Channel and Bay of Biscay: CB
CS	Cantabric Sea, NW Iberian Peninsula and Azores
MA	Madeiran archipelago
CI	Canary Islands
WM	Western Mediterranean and Bay of Cadiz
EM	Eastern Mediterranean (Greek Seas)
EE	Extra-European species

Tissue sampling

Tissue sampling should be done on fresh or fresh frozen fish/fish on ice. When handling the fish use disposable latex gloves/equivalent. Clean equipment and change scalpel blades between each specimen.

You minimally need:

Disposable clean gloves

Tissue sample tubes

Scalpel

Tweezers

Marking pen

95% pure ethanol (not denatured!)

Tissue should be removed from muscle right behind the **right** pectoral fin base (muscle from inside the head in case of large fish). Scrape away the scales, and cut a square (about 5 X 5 mm if possible). Tear or cut the sample into smaller pieces before placing in tissue sample tube, and add alcohol. Mark the tube with FishTrace code, and, if more than one sample is taken from the same specimen, **a**, **b** etc. Store sample tubes in a cool place, e.g., a fridge.

a = sequencing

b = reference collection

c = backup

'Permanent' ink markers are not reliable. It is advisable to secure ink text with adhesive transparent document tape, and to mark both the tube and the lid of the tube. It is advisable to maintain tubes in sequential order in paper/plastic boxes.

Chlorine may be used to clean instruments from DNA, but needs careful washing in ethanol.

Tissue sampling for haplotyping. About the same procedure. Vouchers beyond specimen 5 not required to be saved.

Preservation of voucher specimens

You need:

Formalin

Disposable gloves

Injection needle

Plastic glasses for protecting eyes

A plastic tray with tight lid

Nalgene bottles, 1000 ml, 2000 ml, and 4000 ml.

Fish for collections must be fixed in 10% formalin, i.e., 1 part commercial formalin (36-40% formaldehyde in solution) to 9 parts water. Water must only be fresh water, preferably distilled or deionized water.

If possible, use a flat, wide tray with tight lid for initial fixation. If you use round containers, e.g., Nalgene bottles, keep them lying on the side during fixation.

This is the procedure starting with fresh fish

1. Inject a small amount of 10% formalin in the abdominal cavity and in selected thicker muscular parts of the fish. If fish larger than 30 cm, also cut the right side abdominal wall, ca 5-10 cm, to promote entry of fixation fluid. If fish smaller than 10 cm, no injection is needed except if a herbivorous fish.
2. Dip a cotton or cloth swab in 10% or full strength formalin, and pad the fin base with one hand while raising the each fin with the other hand, to permanently keep the fins erect. Leave so for a few minutes, but keep the whole fish wet/damp during the process.
3. Place the fish in the fixation container with 10 % formalin. Keep for at least 1 week, maximum 1 month. Ensure that a volume of formalin several times that of the fish is present.
4. (Optional; you may send fish to Biological Collection while still in formalin). Rinse the fish in water for a few hours. Run in graded series of ethanol, 25-40-70%, 1 day in each.

With frozen fish, let it thaw thoroughly before fixation.

Otoliths

Otoliths can be removed before or after fixation of a specimen.

Both otoliths (right side and left side) must be obtained from two specimens.

The operation must usually be performed under a stereo dissection microscope with low magnification.

Otoliths are contained in the prootic bulla and more or less visible from the outside of the cranium in many fish species. To extract, lift the gill cover, poke the bulla with the scalpel tip, and cut a round hole on the bulla, extract the sagitta with a fine-tipped tweezer. Avoid scraping the sagitta.

Otoliths from both preserved and fresh fish should be washed lightly in water and blotted dry. With preserved fish it is important to remove all traces of formalin by more intense rinsing in water.

Otoliths should be kept individually and dry in small tubes, with an inner label stating the FishTrace specimen number, the tag number (optional), and from which side each sagitta was extracted (left or right).

Otolith image file name should be FishTrace number plus Otolith:
MulSur-SB-04-otolith.jpg

Identification

Specimens should be identified to species in the field using standard literature references. Species names should follow FishBase even when known to be incorrect (e.g., *Scophthalmus maximus* instead of *Psetta maxima*). Subspecies names should not be used.

One may extract field guides and PDA pages from FishBase to be used as field reference literature.

Decisions on nomenclature and systematics

The WP2 group will decide on the final scientific name and systematic position of each species as to be shown on the FishTrace website. The decision will be made by consensus agreement and cleared through consideration of a most recent systematic revision and the International Code of Zoological Nomenclature. Decisions will be brought before the consortium for comment before finalized.

Field Morphometric data

Field morphometric information may include the following when applicable, and should be taken from each of the five required specimens:

1. Fresh Standard Length (abbreviated SL). From the tip of the snout to the end of the hypural fan. To mm precision
2. Fresh Total Length (abbreviated TL). From the tip of the snout to the end of the caudal fin (if forked, lobes should be pressed against each other, and the length taken to the tip of the longer lobe). To mm precision
3. Fresh Fork Length (abbreviated FL). From the tip of the snout to the end of the middle caudal fin rays. To mm precision
4. Fresh weight in grammes. Note if gutted or intact

Laboratory morphometric data

Morphometric information taken in the lab may include the following when applicable, and should be taken from each specimen preserved for analysis. The data may be recorded by the recipient collection.

1. Standard Length (as above). To nearest 1/10 mm in fish smaller than 100 mm, otherwise to full mm
2. Head length. From the tip of the snout to the most distant point of the margin of the gill cover (operculum or suboperculum as the case may be). To nearest 1/10 mm
3. Body depth. Depends on systematic group. Usually from the ventral/pelvic fin base to the dorsal midline. To nearest 1/10 mm
4. Dorsal fin rays. Separate into spinous (I, II, III, etc), unbranched (i, ii, iii, etc), and branched (1, 2, 3, etc), and give separate count for each dorsal fin (usually one, in mugilids and gobies two, in some gadoids, three; occasionally no dorsal fin present).
5. Anal fin rays. As for dorsal fin
6. Gill rakers. Count ceratobranchial gill rakers only
7. Pectoral fin rays. Count
8. Lateral line scales. Count (optional)
9. Sex (when possible)

Note: This information (less Standard Length which is taken as fresh Standard Length) needs to be sampled in the field for large specimens that are not preserved.

Species information protocol

Species data

Regional and/or general information on each species will be collected as follows:

1. Basic morphology with emphasis on species diagnostic characters. This information will be compiled on a per species basis.
2. Biological information relating to habitat, reproduction, feeding, size and geographic distribution. This information will be compiled on a per region basis for each species.
3. Common names. This information will be compiled both from general sources, such as FishBase and FAO, and on a per region basis within FishTrace.
4. Threat status. This information will be compiled on a regional basis.
5. Fisheries information. This information will be compiled on a regional basis.
6. Socioeconomic information. This information will be compiled on a regional basis.
7. Market appearance of products. This information will be compiled on a regional basis.
8. Bibliography.

Basic morphology. This information will be compiled by WP2 participants and entered in a freetext database field, similar to the information in *Fishes of the Northwestern Atlantic and the Mediterranean*.

Biological information. This information will be submitted by regional groups, and should state:

- a. habitat
- b. depth range
- c. migratory behaviour
- d. foraging behaviour (prey capture method)
- e. aggregation behaviour (schooling, solitary...)
- f. sexuality (gonochorist, hermaphrodite, sex change...)
- g. spawning period
- h. spawning grounds
- i. spawning depth
- j. size at first maturity
- k. spawning type (scatterer, guarder, livebearing)
- l. litter size/egg number
- m. known average age
- n. known maximum age (per sex if possible)
- o. main prey
- p. known average size in catches
- q. known maximum size
- r. current commercial size
- s. current minimum size (as enforced by local laws)
- t. climate zone
- u. north latitude limit
- v. south latitude limit
- w. general distribution area by geographic descriptors

When information is not applicable it should be indicated as *N/A*; when not known, it should be indicated as *unknown*; when not researched it should be indicated as *not researched*; when not yet considered, leave *blank*.

Common names. This information will be submitted by regional groups. The source of the common name has to be referenced. Note that FishBase names are compiled without real quality check (e.g., most Swedish names are dubious).

Threat status. Regional information will be submitted by regional groups, and should state source of information; this will normally be a national Red List. The WP2 group will supply global IUCN threat status.

Fisheries information. This information will be submitted by regional groups and should consider, in a synoptic fashion per species:

- a. Type of fisheries
- b. Fishing methods
- c. Capture period
- d. Exploitation level
- e. Commercial interest

Socioeconomic information. This information will be submitted by regional groups and should consider, in a synoptic fashion per species:

- a. Forms of use (fresh, frozen, salted, dried, dried and salted, warm smoked, cold smoked, macerated, etc.)
- b. Transformed product before commercialization (whole, decapitated, fillet, sliced, roe only, fins, etc.)
- c. Cooking options (steamed, fried, deep-fried, grilled, raw, etc.)
- d. Typical end-consumer (industry, house-hold, subsistence, restaurants)
- e. Consumption site (local, national, exported)
- f. Known market substitutions

If possible products should be photographed as well.

Bibliography. WP2 group will select general bibliography. Regional groups should provide local references obtainable from libraries or official sources. Only literature actually seen and consulted should be considered. The format for references will be provided by the Database group.

Annex XI: Results from Molecular Genetic Procedures Standardisation.

FishTrace WP3. MOLECULAR GENETIC IDENTIFICATION

On FishTrace first meeting (Madrid; 12-14 February, 2003), all partners involved in this Workpackage agreed to standardize the protocol for DNA extraction. For this task, UCM (leader for WP3) sent muscle tissue samples from several species. Each laboratory has extracted that DNA using its own methodology and sent it back to UCM to amplify their cyt b and rhodopsin genes. The results are summarized and explained in a table below:

GROUP	EXTRACTION METHOD	CYT-B (PART "A")	CYT-B (PART "B")	RHODOPSIN
UCM	<i>Phenol / Chloroform / Prot. K</i>	✓	✓	✓
IFREMER	<i>Phenol / Chloroform</i>	✓	✓	≈
NAGREF	<i>Qiagen Dneasy Kit®</i>	✓	≈	✓
NRM	<i>QiAmp DNA mini kit®</i>	✓	✓	✓
RIVO	<i>Qiagen Dneasy Tissue Kit®</i>	✓	✓	X

✓ : Good

≈ : Weak

X : Failure

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

U.C.M.'s EXTRACTIONS and AMPLIFICATIONS:

SAMPLES:

N.	FAMILY	SPECIES	CODE
1	SPARIDAE	<i>Pagrus caeruleostictus</i>	Pcae 2
2	SCIAENIDAE	<i>Pseudolithus typus</i>	Ptyp 2
3	SCIAENIDAE	<i>Argyrosomus regius</i>	Areg 2
4	SERRANIDAE	<i>Mucteroperca fusca</i>	Mfus 2
5	SPARIDAE	<i>Boops boops</i>	Bboo 2
6	SCIAENIDAE	<i>Pseudolithus senegallus</i>	Gallus 2

DATES:

N.	INDIVIDUAL	EXTRACTION	Cyt-B "A"	Cyt-B "B"	RHODOPSIN
1	Pcae 2	04/04/00	23/01/01	23/01/01	14/11/00
2	Ptyp 2	25/01/01	21/03/01	21/03/01	14/03/01
3	Areg 2	13/09/00	02/02/01	02/02/01	20/11/00
4	Mfus 2	23/01/01	20/09/01	17/09/01	12/10/01
5	Bboo 2	08/01/02	25/06/01	25/06/01	22/03/01
6	Gallus 2	25/01/01	24/07/01	24/07/01	03/07/01

PCR MIX:

Components	Volume per reaction (µl)		Final Concentration
	RHOD	CYT-B	
10X Reaction Buffer	2.5	2.5	1X
dNTP mix (10mM of each dNTP)	1	1	0.4mM
Taq DNA polymerase (5u/µl)	0.125	0.125	1.25U/reaction
25mM MgCl ₂	2.5	2.5	2.5mM
Downstream Primer	0.5	0.25	0.5 - 0.25 ng/µl
Upstream Primer	0.5	0.25	0.5 - 0.25 ng/µl
Water MQ (to a final volume of 25µl)			

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

CYCLE PROPERTIES, PRIMERS, ALTERNATIVE STRATEGIES:

CYTOCHROME – B:

PCR PROGRAM: 95°C | 94°C | 50°C | 72°C | 72°C | 4°C
(1st and 2nd reactions) ----- ----- ----- ----- ----- -----
 7min. 30seg. 30seg. 45seg. 7min. ∞

DNA: 1^o PCR: 1µl original or 2µl [1/10] diluted.
 2^o PCR: 1 or 2µl del 1^o PCR.

PRIMERS: 1^oPCR: GluFish-F - CytBI-4R. If the compound does not work, use:
 GluFish-F - CytBI-2R or GluFish-F - CytBI-3R.

 2^oPCR: FishcytB-F - CytBI-5R. If the compound does not work, use:
 FishcytB-F - CytBI-5R.

RHODOPSIN:

1st PCR PROGRAM: 95°C | 94°C | 62°C | 72°C | 72°C | 4°C
 ----- ----- ----- ----- ----- -----
 7min. 30seg. 30seg. 30seg. 7min. ∞

2nd PCR PROGRAM: 95°C | 94°C | 56°C | 72°C | 72°C | 4°C
 ----- ----- ----- ----- ----- -----
 7min. 30seg. 30seg. 30seg. 7min. ∞

DNA: 1^o PCR: 2µl [1/10] diluted.
 2^o PCR: 1µl [1/10] 1^o PCR products diluted.

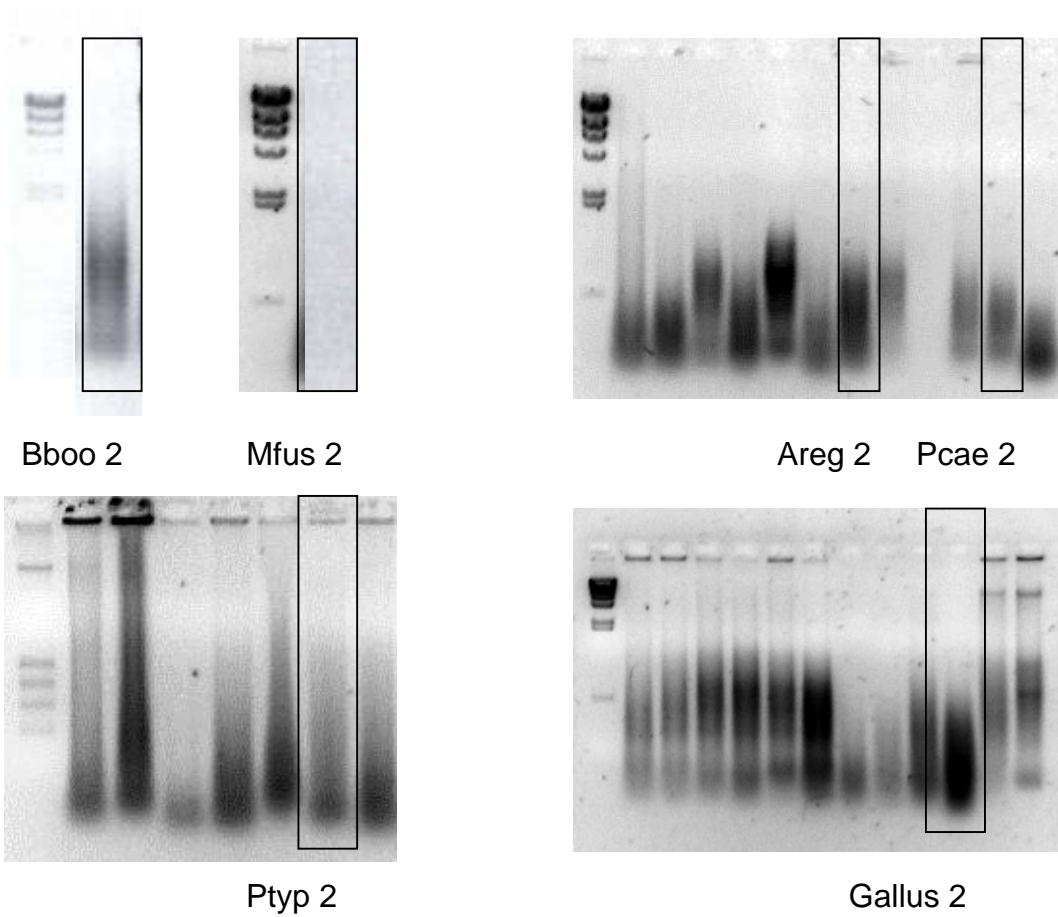
PRIMERS: 1^oPCR: Rod-F2B - Rod-R5. If the compound does not work, use:
 RHO-30F - RHO-319R.

 2^oPCR. Rod-F2w - Rod-R4n. If the compound does not work, use:
 Rod-F2x - Rod-R4n; Rod-F2y - Rod-R4n or Rod-F2 - Rod-R4.

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

UCM's DNA EXTRACTIONS

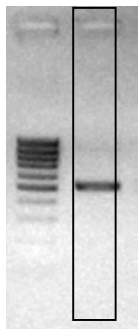
N.	CODE	SPECIES	EXTRACTION
1	Pcae 2	<i>Pagrus caeruleostictus</i>	04/04/00
2	Ptyp 2	<i>Pseudolithus typus</i>	25/01/01
3	Areg 2	<i>Argyrosomus regius</i>	13/09/00
4	Mfus 2	<i>Mucteroperca fusca</i>	23/01/01
5	Bboo 2	<i>Boops boops</i>	08/01/02
6	Gallus 2	<i>Pseudolithus senegallus</i>	25/01/01



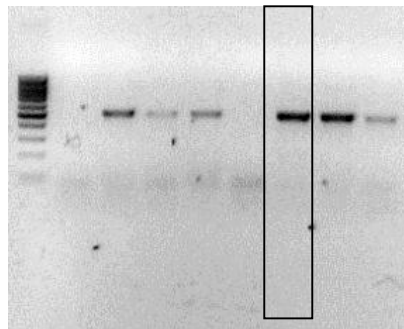
FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

UCM's amplified RHODOPSIN

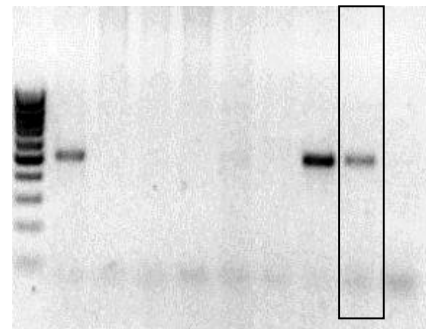
N.	CODE	SPECIES	RHODOPSIN
1	Pcae 2	<i>Pagrus caeruleostictus</i>	14/11/00
2	Ptyp 2	<i>Pseudotolithus typus</i>	14/03/01
3	Areg 2	<i>Argyrosomus regius</i>	20/11/00
4	Mfus 2	<i>Mucteroperca fusca</i>	12/10/01
5	Bboo 2	<i>Boops boops</i>	22/03/01
6	Gallus 2	<i>Pseudotolithus senegallus</i>	03/07/01



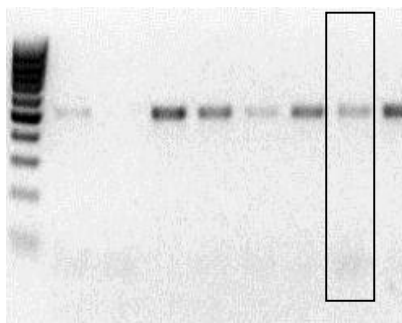
Bboo 2



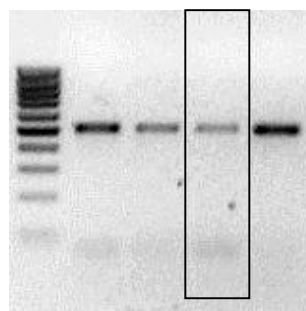
Mfus 2



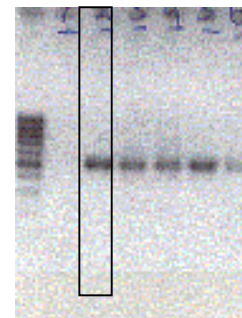
Areg 2



Pcae 2



Ptyp 2

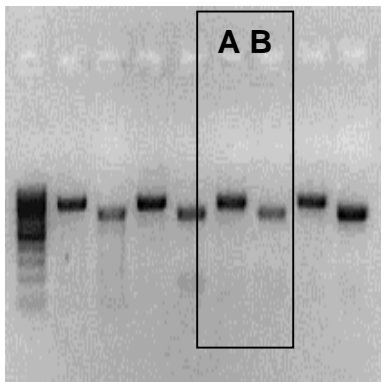


Gallus 2

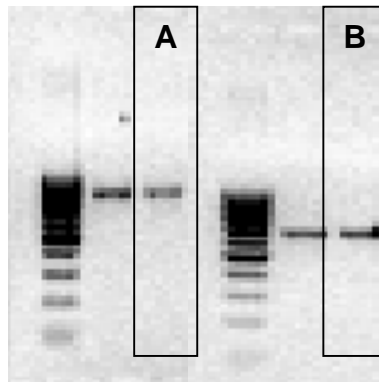
FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

UCM's amplified CITOCHROME-B

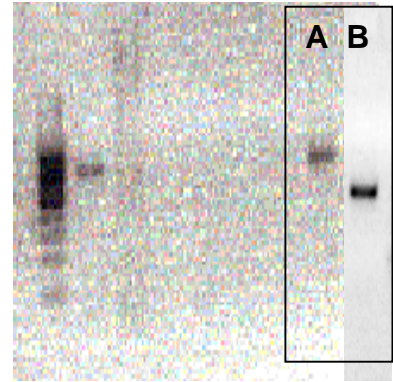
N.	CODE	SPECIES	Cyt-B "A"	Cyt-B "B"
1	Pcae 2	<i>Pagrus caeruleostictus</i>	23/01/01	23/01/01
2	Ptyp 2	<i>Pseudotolithus typus</i>	21/03/01	21/03/01
3	Areg 2	<i>Argyrosomus regius</i>	02/02/01	02/02/01
4	Mfus 2	<i>Mucteroperca fusca</i>	20/09/01	17/09/01
5	Bboo 2	<i>Boops boops</i>	25/06/01	25/06/01
6	Gallus 2	<i>Pseudotolithus senegallus</i>	24/07/01	24/07/01



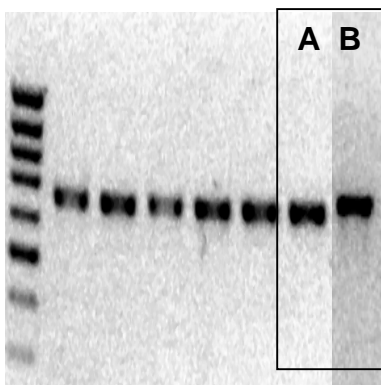
Bboo 2



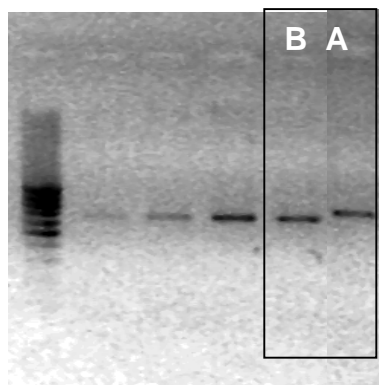
Mfus 2



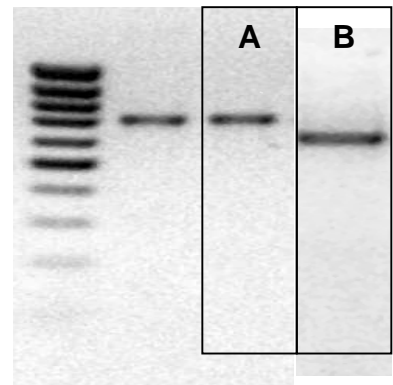
Areg 2



Pcae 2



Ptyp 2



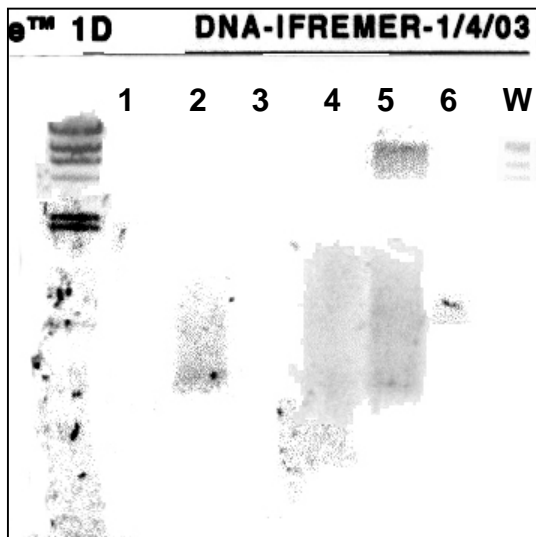
Gallus 2

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

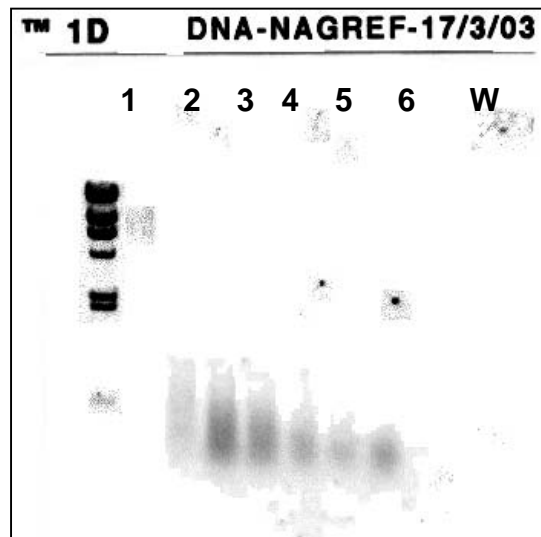
RESULTS OBTAINED BY THE REST OF THE GROUPS INVOLVED ON WP3

DNA EXTRACTIONS

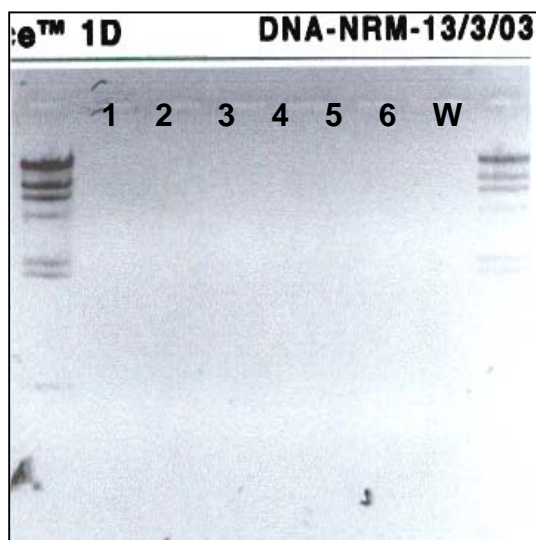
N.	CODE	SPECIES
1	Pcae 2	<i>Pagrus caeruleostictus</i>
2	Ptyp 2	<i>Pseudotolithus typus</i>
3	Areg 2	<i>Argyrosomus regius</i>
4	Mfus 2	<i>Mucteroperca fusca</i>
5	Bboo 2	<i>Boops boops</i>
6	Gallus 2	<i>Pseudotolithus senegallus</i>



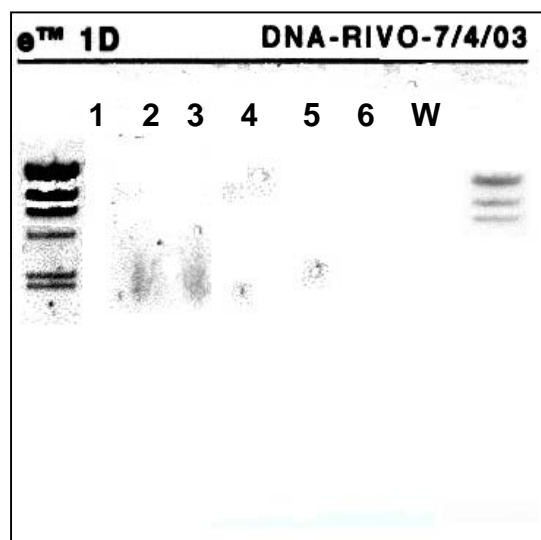
IFREMER



NAGREF



NRM

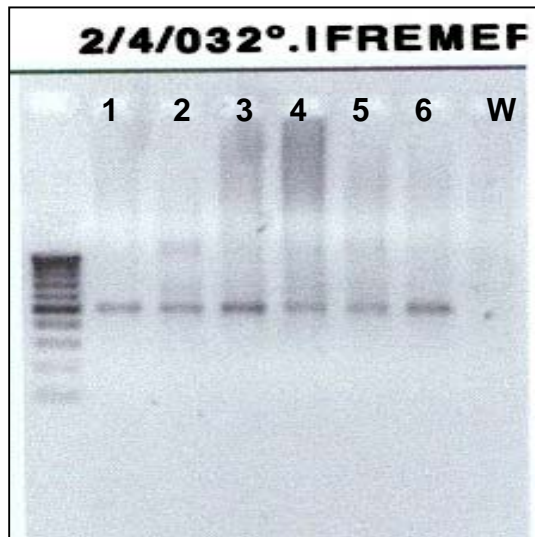


RIVO

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

RESULTS OBTAINED BY THE REST OF THE GROUPS INVOLVED ON WP3

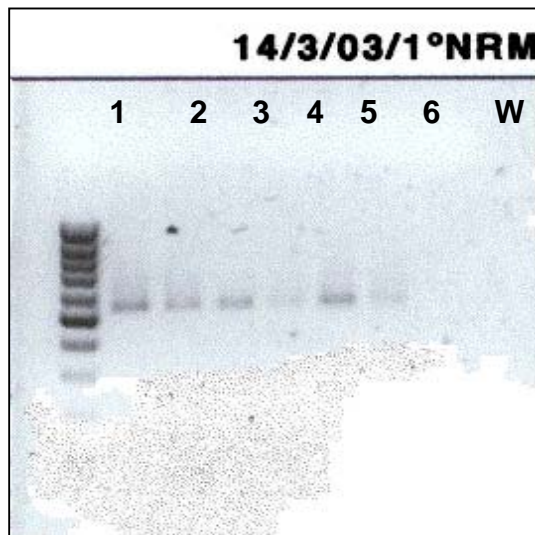
RHODOPSIN



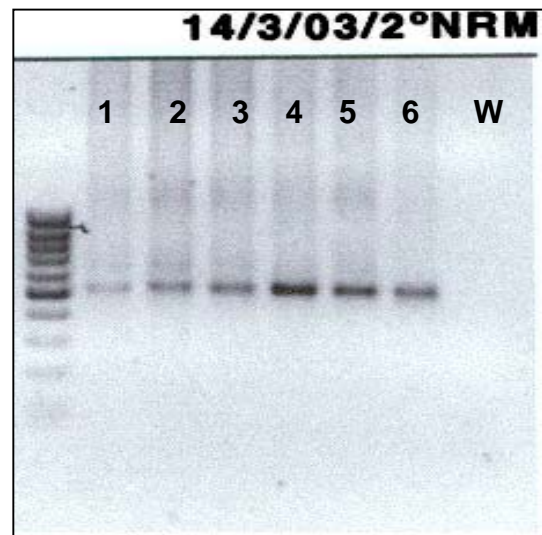
IFREMER (2nd)



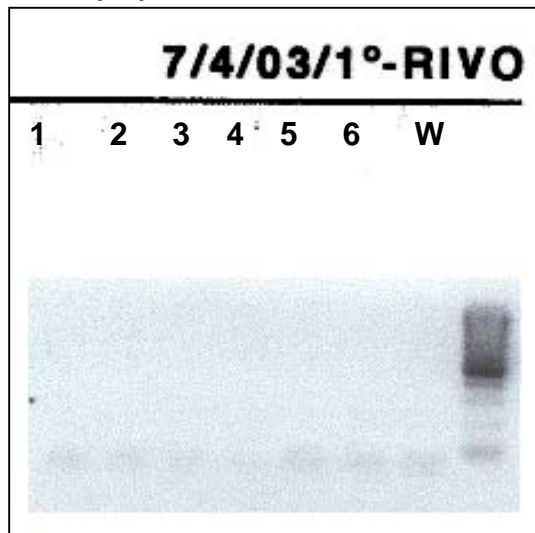
NAGREF (1st&2nd)



NRM (1st)



NRM (2nd)



RIVO (1st)

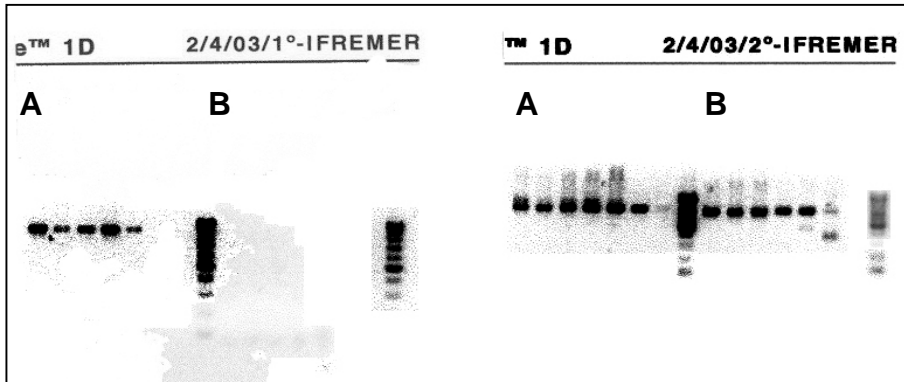


RIVO (2nd)

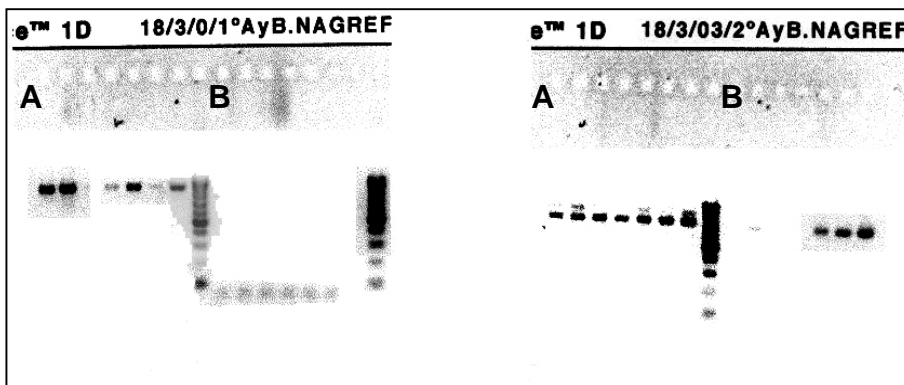
FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

RESULTS OBTAINED BY THE REST OF THE GROUPS INVOLVED ON WP3

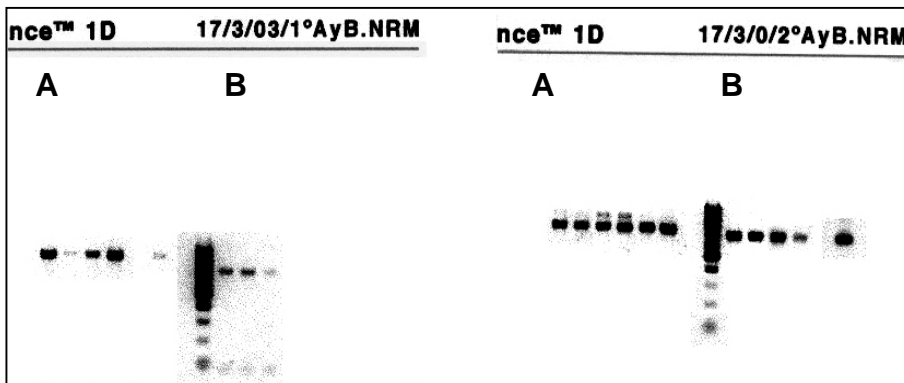
CITOCHROME-B



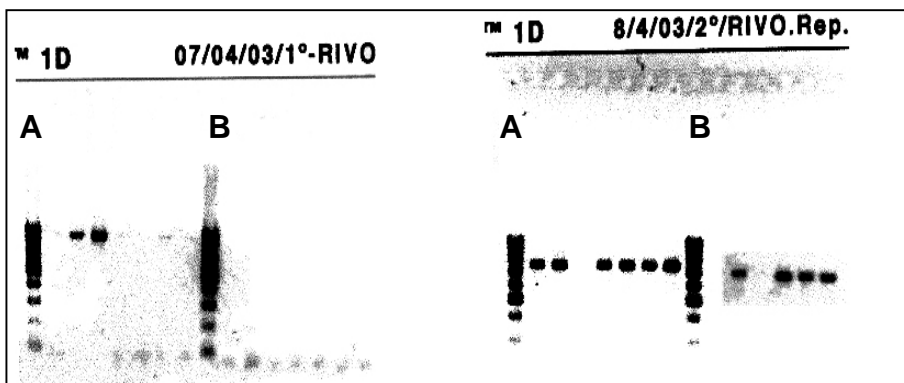
IFREMER



NAGREF



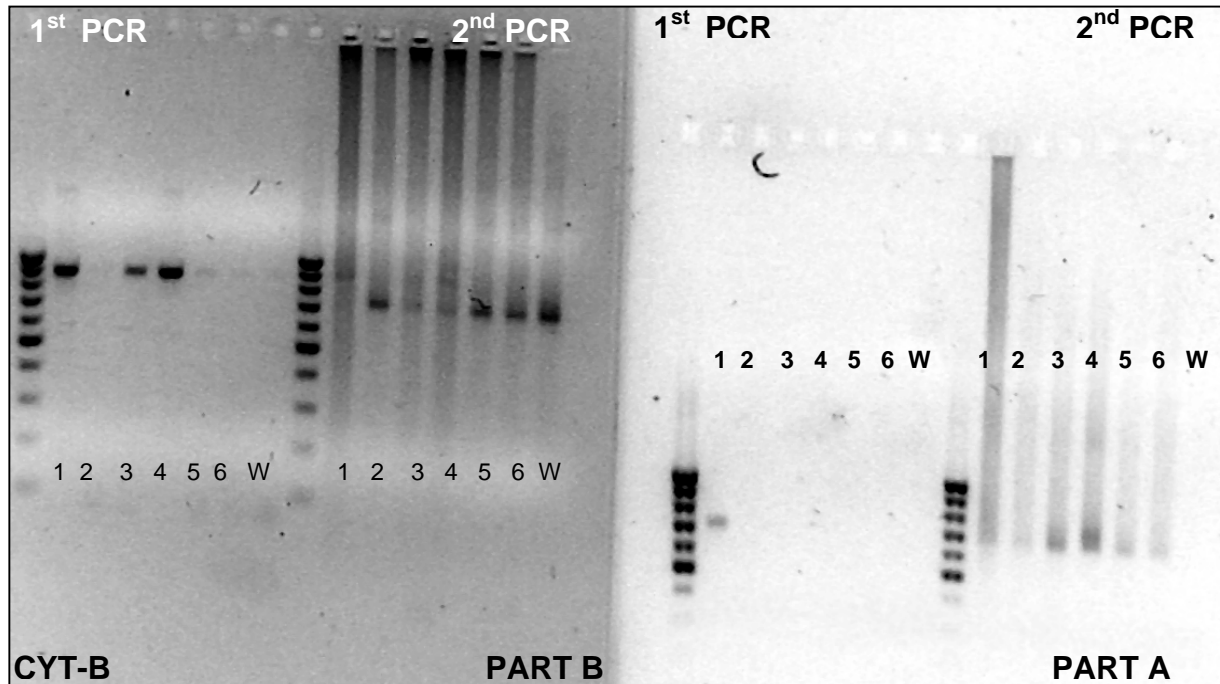
NRM



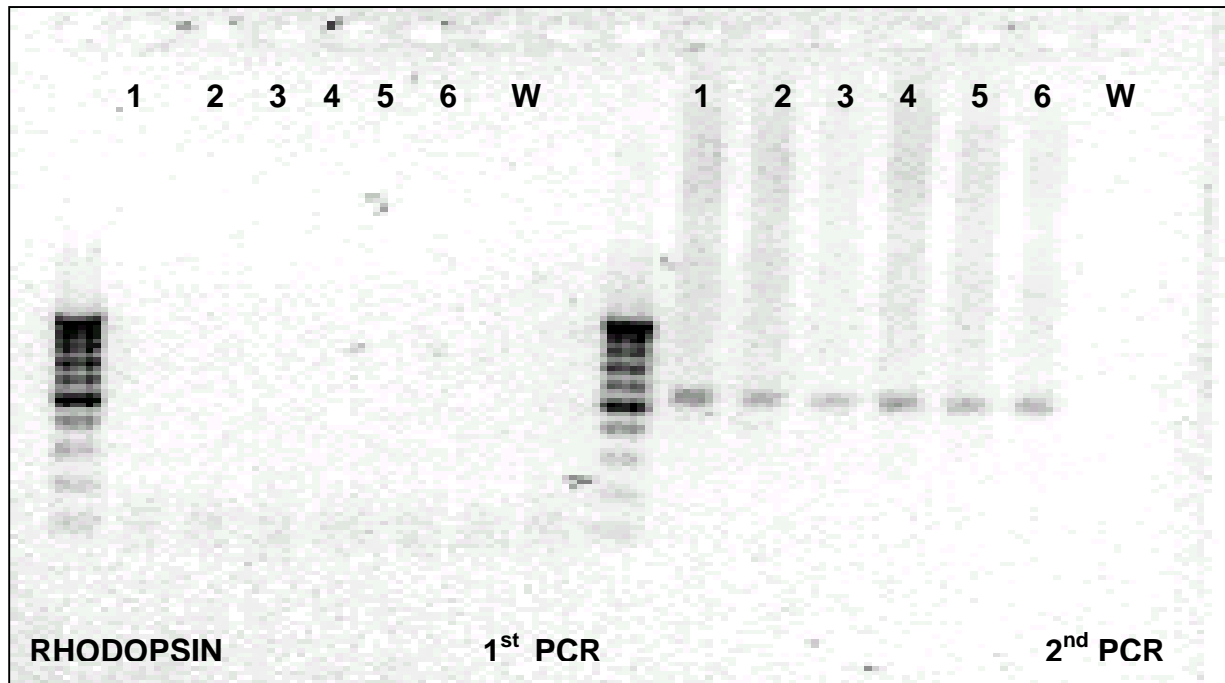
RIVO

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

NRM's ALTERNATIVE METHOD FOR PCR: Kit-Extracted DNA plus PCR Beads.



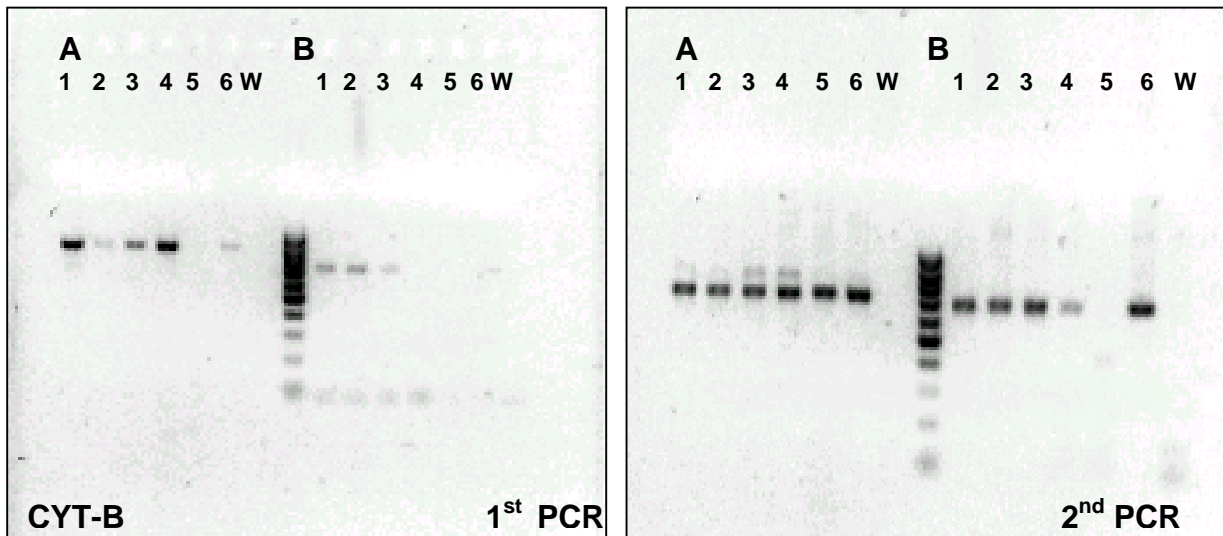
To amplify Cyt B : 2µl of the DNA sent by NRM without any dilution for the 1st PCR. 1µl from the 1st PCR in the 2nd PCR. It's necessary to **add Cl₂Mg** to get a final concentration of 2.5mM in the PCR reaction.



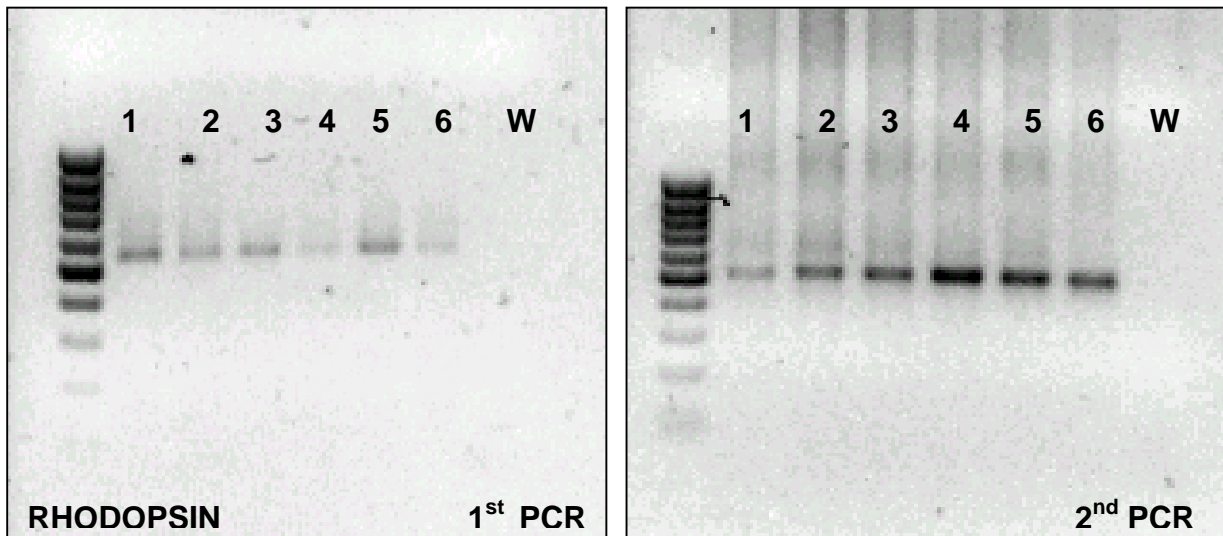
To amplify Rhodopsin: 2µl of the DNA sent by NRM without any dilution for the 1st PCR. 1µl from the 1st PCR for the 2nd PCR. Using **PCR beads** it's necessary to **add Cl₂Mg** to get a final concentration of 2.5mM in the PCR reaction.

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

NRM's KIT-EXTRACTED: DNA amplified with UCM's PCR Reagent System.



To amplify Cyt B: 2 μ l of the DNA sent by NRM, without any dilution for the 1st PCR. 1 μ l from the 1st PCR in the 2nd PCR. In this PCR should be good to try other amounts of primers to improve the specificity of the PCR products.



To amplify Rhodopsin: 2 μ l of the DNA sent by NRM, without any dilution for the 1st PCR. 1 μ l from the 1st PCR for the 2nd PCR. Probably the 2nd PCR product would be cleaner with a 1/10 dilution from the 1st PCR.

FishTrace
WP3. MOLECULAR GENETIC IDENTIFICATION

CONCLUDING REMARKS

DNA EXTRACTIONS

Under our point of view and after have checked all the results (from the extractions to the final amplification of both genes), we can conclude and verify through the pictures that all methods used for the extraction satisfy its aim correctly.

RHODOPSIN

The chosen method for the amplification of the Rodopsin gene is the UCM's reagent system (although RIVO has suffered complications with the amplification of this gene). After doing the comparison between NRM's PCR-Beads and UCM's reagent system, we opt for to still using our method, because observing the pictures of results and analysing other parameters (price, time requirements, used disposable materials, etc.), we did not note any significant advantage on PCR-Beads system.

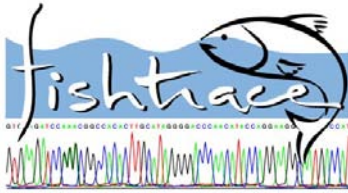
CYTOCHROME-B

The amplification of Cytochrome-B gene was more difficult (concretely of part "B"), as the pictures develop (see RIVO, NAGREF and NRM's pictures). The chosen method for this amplification (both parts: "A" and "B"), is the UCM's reagent system again.

To summarize, the best results were obtained extracting DNA with Qiagen's Kit (as well as using phenol/chloroform method), and doing PCR with UCM's reagent system.

Madrid, May 29th, 2003.

Annex XII: Molecular identification and DNA barcoding (WP3). Protocol and PCR conditions.



WP3 PROTOCOL AND PCR CONDITIONS

All PCRs are NESTED: a first PCR followed by a second PCR. The second PCR use as template the DNA generated in the first PCR. There are several alternative primers pairs for each PCR. The final mixtures must be adjusted to the figures shown in table 1.

NOMENCLATURE: Nested PCRs are named with the acronym of the amplified fragment (RHOD, CYT5, CYT3), followed by 2 digits: the first corresponding to the condition of the first PCR and the second digit corresponding to the condition of the second PCR.

Example: RHOD-12, CYT5-21, CYT3-13 ...

ORDER OF PRIORITY: We have numbered the reactions in the order that render best results. Therefore, the priority order of any fragment is as follows: 11, 12, 13, 21, ... etc.

Table 1. PCR mixtures.

Components	Volume per reaction (µl)		Final Concentration
	RHOD	CYT-B	
10X Reaction Buffer	2.5	2.5	1X
dNTP mix (10mM of each dNTP)	1	1	0.4mM
Taq DNA polymerase (5u/µl)	0.125	0.125	1.25U/reaction
25mM MgCl ₂	2.5	2.5	2.5mM
Downstream Primer	0.5	0.25	0.5 - 0.25 ng/µl
Upstream Primer	0.5	0.25	0.5 - 0.25 ng/µl
Water MQ	Fill to final volume of 25µl		

The optimal conditions for the amplification of the target fragments are as follows:

RHODOPSIN

RHOD 1st PCR	1	2	3
PRIMERS (25ng/μl)	F2B/5R (0.5ng/μl)	F2B/5R (0.5ng/μl)	30F/319R (0.5ng/μl)
Mg Cl ₂ (25mM)	2,5 mM	2.5 mM	2.5 mM
dNTPs (10mM)	0.4 mM	0.4 mM	0.4 mM
Annealing Temp. (sec) A	62	60	62
Extesion Time (°C) E	30	30	30

RHOD 2nd PCR	1	2	3	4
PRIMERS (25ng/μl)	F2W/R4N (0.5ng/μl)	F2X/R4N (0.5ng/μl)	F2W/R4N (0.5ng/μl)	F2X/R4N (0.5ng/μl)
Mg Cl ₂ (25mM)	2.5 mM	2.5 mM	2.5 mM	2.5 mM
dNTPs (10mM)	0.4 mM	0.4 mM	0.4 mM	0.4 mM
Annealing Temp. (sec) A	56	56	54	54
Extesion Time (°C) E	30	30	20	30

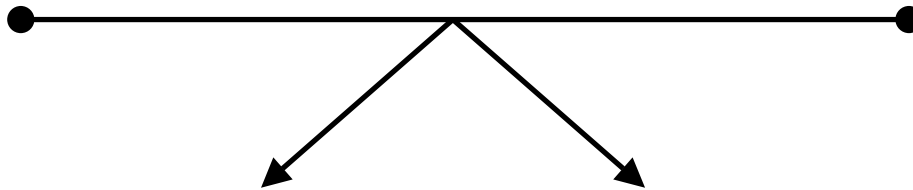
1st PCR PROGRAM:	95°C 94°C A 72°C 72°C 4°C ----- 7min. 30seg. 30seg. E 7min. ∞
2nd PCR PROGRAM:	95°C 94°C A 72°C 72°C 4°C ----- 7min. 30seg. 30seg. E 7min. ∞

CYTOCHROME-B

Notice that the first PCR for the Cytochrome-B is used as a template for the subsequent amplification of both Cyt-B fragments (5' and 3') in a second PCR.

Cyt-B complete (1st PCR)

CYT_B 5' and 3' 1 st PCR	1	2	3	4
PRIMERS (25ng/μl)	GluFish-F / THR-Fish-R (0.25 ng/μl)	FishcytB-F/ THR-Fish-R	FishcytB-F/ TrucytB-R (0.25 ng/μl)	GluFish-F/ TrucytB-R (0.25 ng/μl)
Mg Cl₂ (25mM)	2.5 mM	2.5 mM	2.5 mM	2.5 mM
dNTPs (10mM)	0.4 mM	0.4 mM	0.4 mM	0.4 mM
Annealing Temp. (sec) A	50	50	50	50
Extesion Time (°C) E	30	30	30	30



5' end Cyt-B (2nd PCR)

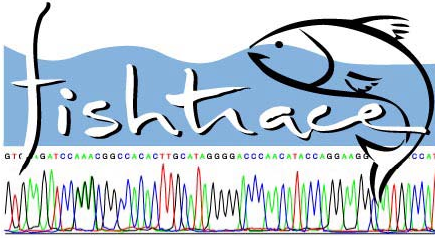
3' end Cyt-B (2nd PCR)

CYTB_5' 2 nd PCR	1	2
PRIMERS (25ng/μl)	GluFish-F/ CytBI-5R (0.25 ng/ml)	GluFish-F/ CytBI-5R (0.25 ng/ml)
Mg Cl₂ (25mM)	2.5 mM	2.5 mM
dNTPs (10mM)	0.4 mM	0.4 mM
Annealing Temp. (sec) A	50	50
Extesion Time (°C) E	30	30

CYTB_3' 2 nd PCR	1	2	3
PRIMERS (25ng/μl)	TrucytB-R/ CytBI-7F (0.25 ng/ml)	TrucytB-R/ CytBI-7F (0.25 ng/ml)	TrucytB-R/ CytBI-6F (0.25 ng/ml)
Mg Cl₂ (25mM)	2.5 mM	2.5 mM	2.5 mM
dNTPs (10mM)	0.4 mM	0.4 mM	0.4 mM
Annealing Temp. (sec) A	52	50	52
Extesion Time (°C) E	20	20	30

PCR PROGRAM:	95°C 94°C A 72°C 72°C 4°C
(1st and 2nd reactions)	----- 7min. 30seg. 30seg. E 7min. ∞

Annex XIII: Phylogenetic Validation of Sequences. Guidelines.



FishTrace WP3-WP7

PHYLOGENETIC VALIDATION OF SEQUENCES

Table of Contents

- 1. Guidelines for the installation and use of *ClustalX* v1.81 and *MEGA2.1***
- 2. Practical Exercise of Fish Phylogeny with *MEGA2.1***
- 3. Phylogenetics Reconstructions of FishTrace Sequences with *MEGA2.1***

1. Guidelines for the installation and use of *ClustalX* v1.81¹ and *MEGA2.1*²

- PC compatible software for phylogenetic analysis -

The important amount of molecular genetic information compiled throughout the duration of this project requires to be contrasted for reliability and validated among the different data sets and the various groups involved. The objectives of this workpackage (WP7) will be directed towards the validation of all the information introduced in the database.

The following flow of procedures should be completed for the validation of the molecular genetic information before to submit the definitive data generated [nucleotide sequences of the target genes *Cytochrome B* (1141bp) and *Rhodopsin* (460bp), into the database].

VALIDATION STEPS AT THREE LEVELS

1st STEP: Internal validation by **each partner**:

- 1) To compare both sequences obtained from individuals 1 and 2 (Aligning).
- 2) To register every change in the sequence.
- 3) To verify the position of each sequence on fish phylogeny.

2nd STEP: Validation by the **expert group**:

- 1) To align and compare all data submitted by each partner.
- 2) To compile the registers of changes in the sequences and in the different methods employed to obtain it.
- 3) Phylogenetic analysis (BLAST, Trees).
- 4) Submit to the database and inform to the WP7 responsible.

3rd STEP: Definitive validation by **WP7 responsible**:

- 1) Data checking (missing data, errors,...) from submissions.
- 2) Database Fields content checking from submission.
- 3) Certified Validation of every species.

This document is intended for the genetics teams for the use of the programmes in a standard way for the comparison (aligning) of the sequences obtained (*ClustalX*) and for the verification of the position of each sequence on fish phylogeny (*MEGA2.1*). Although each programme contains its own Help File, here are the guidelines to carry out the validation process of the FishTrace sequences through these programmes. We have selected these programmes due to their user-friendly interface, fast computational algorithms and useful statistical methods.

¹Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 24:4876-4882.

²Kumar,S., Tamura,K., Jakobsen,I., and Nei,M. (2001) *MEGA2: Molecular Evolutionary Genetics Analysis software*, Arizona State University, Tempe, Arizona, USA.

Guidelines for the installation and use of *ClustalX* v1.81

This is a free downloadable software. The programme and user instructions can be obtained at:

www-igbmc.u-strasbg.fr/BioInfo/

The latest version of the ClustalX programme can be also obtained by anonymous ftp:

<ftp://igbmc.u-strasbg.fr>
<ftp://embl-heidelberg.de>
<ftp://ebi.ac.uk>

Installing *ClustalX*

Supplied as executable file (*clustalx.exe*) run under MS Windows (32 bit).

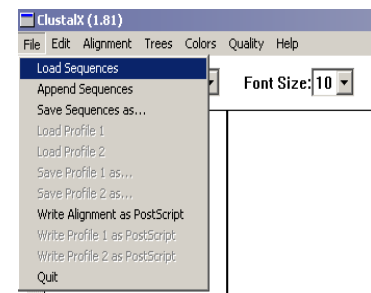
Starting *ClustalX*

ClustalX is a Windows interface for the *ClustalW* multiple sequence alignment programme. It provides an integrated environment for performing multiple sequence alignments. The sequence alignment is displayed in a window on the screen.

Sequence Input

You input sequences by using the *File* menu. Note that is possible to exchange files between *ClustalX* and *MEGA2.1*. The following seven formats are automatically recognised:

- NBRF/PIR
- EMBL/SWISSPROT
- Pearson (Fasta)
- Clustal (*.aln)
- GCG/MSF (Pileup)
- GCG9 RSF
- GDE flat file



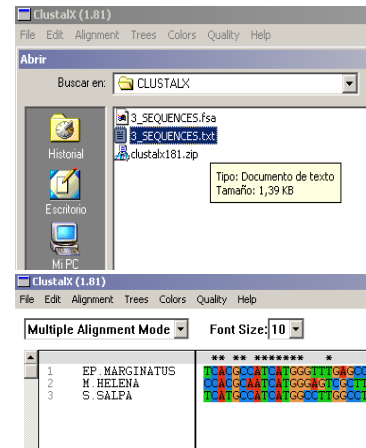
Other formats are rejected. We recommend to convert the sequences to FASTA (.fsa) format including all sequences into a single file. The FASTA format is very simple and is quite similar to the MEGA file format. It is possible to create a FASTA file using a text editor (NotePad, Word, etc) as a text file (.txt). A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater-than ">" symbol in the first column. It is recommended that all lines of text be shorter than 80 characters in length. Once the input file created has been created (NotePad, Word, etc), can be saved as plain text format as well as FASTA format (e.g. 3_sequences.txt or 3_sequences.fsa) since *ClustalX* recognizes files saved with this nomenclature.

Here is an example sequence in FASTA format:

```
>EP.MARGINATUS
TCACGCCATCATGGGTTTGAGCCTCACCTGGTTGATGGCCTTAGCTTGTTCCGTACCCCTCTTGTCGGCTGGTCTCGT
TACATTCC TGAGGGCATGCAGTGCTCATGTGGAGTCGACTACTACACACGTGCAGAGGGCTTCAATAACGAGTCCTTTG
TTATCTACATGTTCTGCTGCCACTTCATGATTCCACTGATTGTCAATTTCTTCTGCTACGGCCGTCTGCTCTGTGCTGT
CAAGGAGGCTGCTGCTGCCAGCAGGAGTCTGAGACCACCCAGAGGGCTGAGAGGGAAGTCACCCGCATGGTCGTTATC
ATGGTTATCGCTTCCTGGTGTGTTGGTGTCCCTATGCCGGTGTGGCCTGGTTTATCTTTCACACATCAAGGCTCTGAGT
TCGGACCAGTCTTCATGACCATCCCTGCCTTCTTTGCCAAGAGCTCTTCCATCTACAACCCAATG
```

Input sequence file example

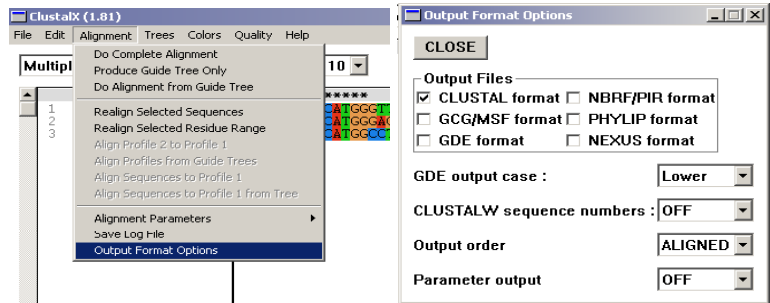
Load and analyse the sequences on *ClustalX* by clicking on "Load Sequences" and opening the example text or FASTA file: **3_sequences.fsa** or **.txt** located in the same folder as this document.



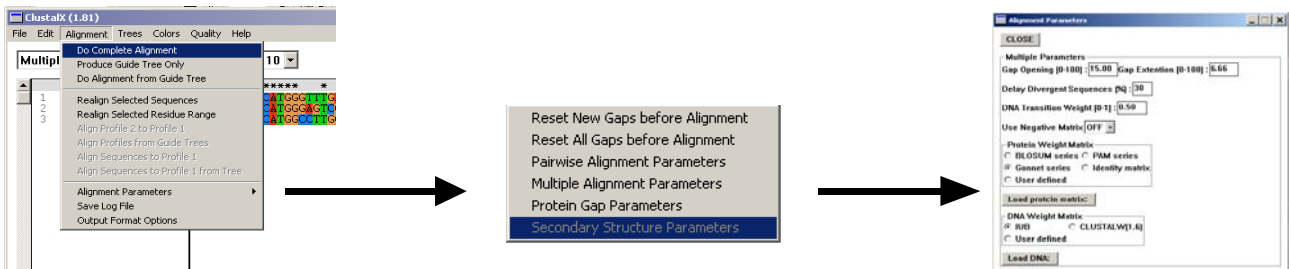
The sequences are base-colored automatically, when they are loaded. Sequences can be also colored by assigning a color to bases/residues.

Sequence Alignment

Before starting the alignment, the output file formats must be chosen as described in the figure (*ALIGNMENT* menu). Several sequence formats can be selected for the alignment output file. For use with *MEGA2.1* it is recommended Clustal format (.aln), because is easy to convert to *MEGA* files (.meg) with *MEGA2.1* software.



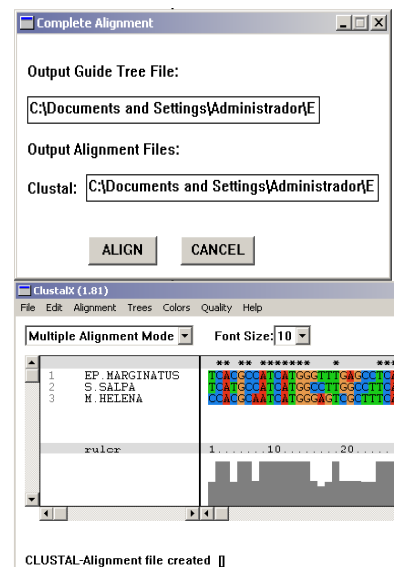
Use the *ALIGNMENT* menu to do multiple alignments. Be sure that **MULTIPLE ALIGNMENT MODE** is selected. Note that for our purpose the alignment is done using default parameters. Alignment parameters can be changed in the *ALIGNMENT PARAMETERS* menu.



Click **DO COMPLETE ALIGNMENT** in the *ALIGNMENT* menu to proceed. Select the definitive location for the output file (Hard disc, floppy, CD, etc), and press **ALIGN**.

Multiple alignments are carried out in 3 stages:

- 1) all sequences are compared to each other (pairwise alignments).
- 2) a dendrogram (like a phylogenetic tree) is constructed, describing the approximate groupings of the sequences by similarity (stored in a different file).
- 3) a final multiple alignment is calculated using the dendrogram as a guide, and a file (.aln) is created. The obtained file from the multiple alignment of the sequences with *ClustalX* is indeed the first step to make the formatted file for phylogenetic analysis with *MEGA2.1*.



Guidelines for the installation and use of *MEGA2.1*[®]

Installing *MEGA2.1*

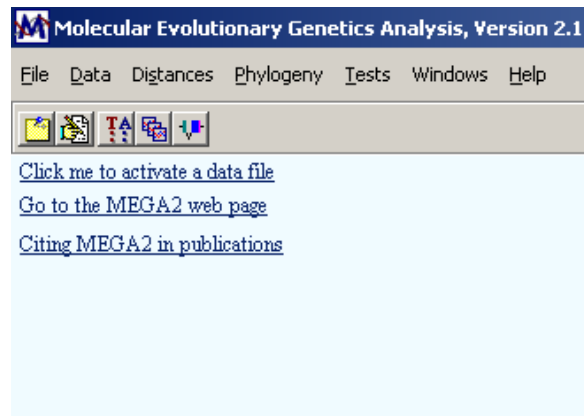
This software is also free available and the most convenient and preferred way to install *MEGA2.1* is to install it directly from the website: www.megasoftware.net where you can observe an online user manual too.


A specially designed installation programme automatically downloads *MEGA2.1* to your computer and Installs it in the location (directory) you specify during the installation procedure. If you are unable to install *MEGA2.1* directly from the website, the second option is to download the *MEGA2.1* programme as a single compressed ZIP file to your computer. You will need to use a programme to uncompress this ZIP file (e.g., WinZip software) in a temporary directory. This will produce a number of files. Click on the *setup.exe* file to install *MEGA2.1* on your computer automatically.

Starting *MEGA2.1*

This is an user-friendly programme that can be divided in three main blocks:

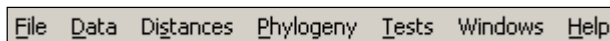
- The text editor and format converter.
- The main window.
- A toolbar with different menus for the analysis of active data: *File*, *Data*, *Distances*, *Phylogeny*, *Test*, *Windows* and *Help*.



a) The **Text File Editor and Format Converter** () included in *MEGA2.1* is useful for creating and editing ASCII text files. *Text File Editor* is also invoked automatically by *MEGA2.1* if the input data file processing modules detect errors in the data file format. In this case you should make appropriate changes and save the data file. The text editor is straightforward to use if you are familiar with programmes like Notepad. Click the mouse on the section you wish to change, type in the new text, or select text to cut, copy or paste. Your data should be formatted according to a set of rules so it can be read and interpreted correctly in *MEGA2.1* (see below under “**The MEGA format**”). All input data files are basic ASCII-text files, which may contain DNA sequence, protein sequence, evolutionary distance, or phylogenetic tree data. Most word processing packages (e.g., MicrosoftWord, WordPerfect, Notepad, WordPad) allow you to edit and save ASCII text files. These are usually marked with a *.txt* extension, but it is recommended to change this extension to *.meg* after creating the file, so you can distinguish between your data files and the other text files.

b) The main window allows to activate the data file for [Click me to activate a data file](#) analysis:

c) Once the data file is activated, the different procedures for the phylogenetic analysis can be selected in the toolbar:



The MEGA format

Common Features:

A MEGA text file can be divided in two informative regions: The *header* and the *sequence alignment*.

The Header

The first line must contain the keyword **#mega** to indicate that the data file is in the MEGA format. The data file may optionally contain a succinct description of the data (called *Title*) included in the file on the second line. *Title* and *Description* statement must be written on the line following the **#mega**. It always begins with **!Title** and ends with a semicolon ";". In addition, the data file may also contain a *Format* statement, which includes information on the type of data present in the file and some of its attributes.

Example of a common *header*:

```
#mega
!TITLE FishTrace Cyt-B Phylogeny;
!Format DataType=DNA Nseqs=11 NTaxa=11 NSites=1141 Missing=?;
!Description Cyt-B-Test;
```

The *Format* statement should be generally written after the *Title* or the *Description* statement. Writing a format statement requires knowledge of keywords used to identify different types of data and data attributes.

Keywords for the *Format* Statement (Sequence data):

Command	Setting	Remark	Example
DataType	DNA, RNA, nucleotide, protein	Specifies the type of data in the file	DataType=DNA
NSeqs	A count	Number of sequences	NSeqs=85
NTaxa	A count	Synonymous with NSeqs	NTaxa=85
NSites	A count	Number of nucleotides or amino acids	NSites=4592
Property	Exon, Intron, Coding, Noncoding, and End. specifies whether or not a domain is protein coding. Exon and Coding are synonymous, as are the Intron and Noncoding keywords. End specifies that the domain with the given name ends at this point..		Property=cyt_b
Indel	single character	Use dash (-) to identify insertion/deletions in sequence alignments	Indel=-
Identical	single character	Use period (.) to show identify with the first sequence.	Identical=.
MatchChar	single character	Synonymous with the Identical keyword.	MatchChar=.
Missing	single character	Use a question mark (?) to indicate missing data.	Missing=?
CodeTable	A name This instruction gives the name of the code table for the protein coding domains of the data		CodeTable=Standard

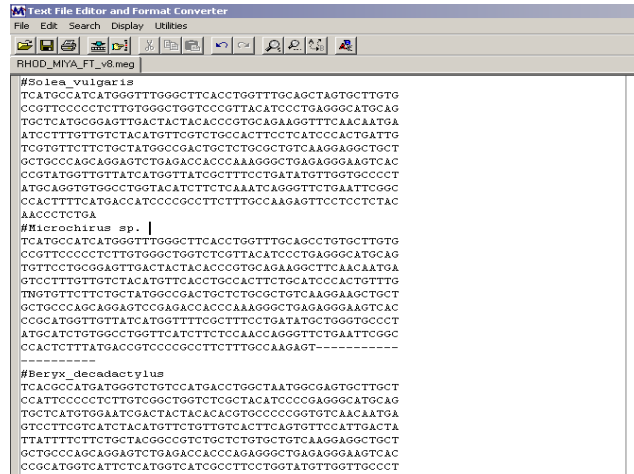
Then, the sequence data must consist of two or more sequences of equal length. All sequences must be previously aligned with *ClustalX* (*MEGA2.1* does not include an alignment programme). Nucleotide and amino acid sequences should be written in IUPAC single-letter codes. Sequences can be written in any combination of upper- and lower-case letters. Special symbols for alignment gaps, missing data, and identical sites can also be included in the sequences.

Sequence Alignment

Nucleotide or amino acid sequences should be written in IUPAC single-letter codes (*Text File Editor and Format Converter*)


Special Symbols:

Blank spaces and Tabs are frequently used to format data files, so they are simply ignored by *MEGA2.1*. ASCII characters such as the period (.), dash (-), and question mark (?), are generally used as special symbols to represent identity to the first sequence, alignment gaps, and missing data, respectively.



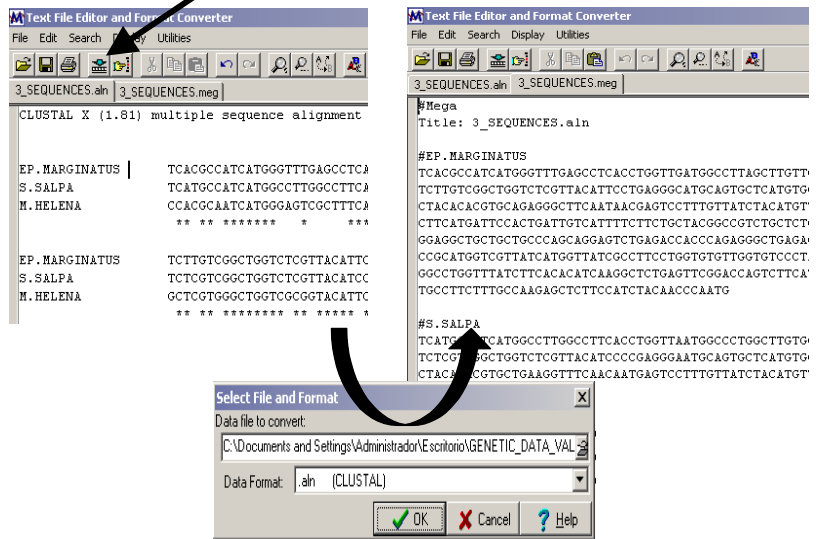
MEGA2.1 Text File Editor: Exaple of sequence alignment

Converting to Mega Format (using Text File Editor and Format Converter)

If you have your sequences aligned with *ClustalX*, you can convert your output file (.aln) into MEGA format using the *Format Converter* () as described below:

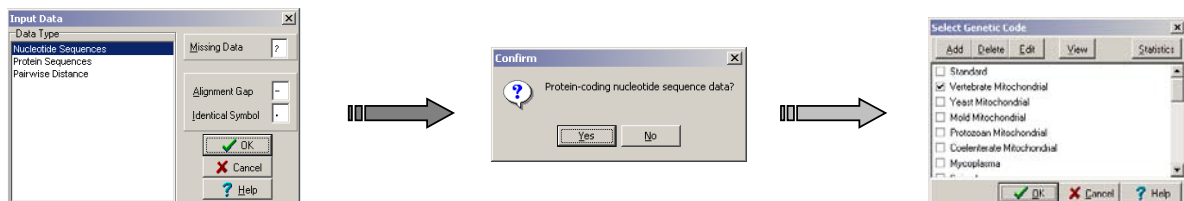
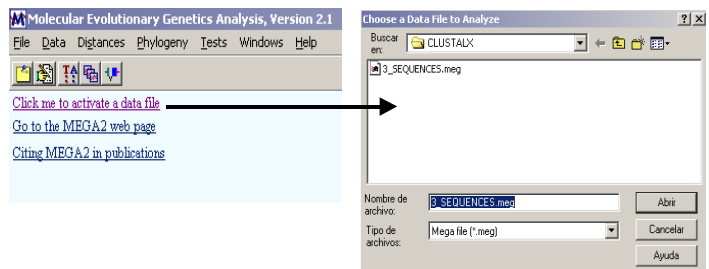
Browse and select the file to be converted (a multiple alignment file from *ClustalX*: .aln) and click **OK** to automatically convert it into MEGA format. The file obtained (.meg) has the format required for the programme to run the phylogenetic analysis under *MEGA2.1* software.

Files written in a number of popular data formats can be converted into MEGA format. *MEGA2.1* supports conversion of CLUSTAL, NEXUS (PAUP, MacClade), PHYLIP, GCG, FASTA, PIR, NBRF, MSF, IG, and XML formats.



Activating the data file

Click on the hiperlink shown and then choose the data file to be analysed on the dialog window. Before to load the file, a new dialog window appears asking for the type of input data: Nucleotide or Protein sequences), followed by the confirmation if your file contains a protein-coding nucleotide sequence data. Finally, a new window is used for the selection of the appropriate genetic code (Standard, Mitochondrial, etc).

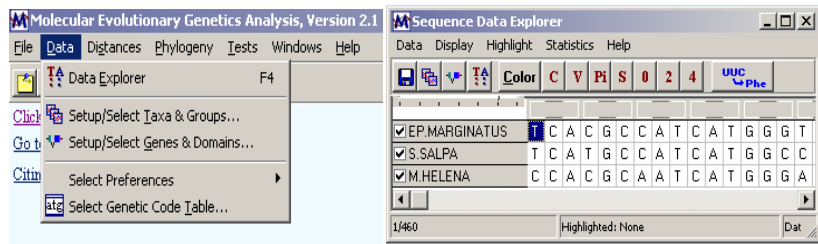


Immediately after, the programme runs the datafile and its attributes can be observed at the bottom of the main window:

Data File	C:\Documents and Settings\Administrador\Escritorio\GENETIC_DATA_VALIDATION\CLUSTALX3_SEQUENCES.meg
Title	: 3_SEQUENCES.aln
Description	
Data Type	Nucleotide
Sequences	3
Sites	460
Code Table	Vertebrate Mitochondrial

Sequence Data Explorer

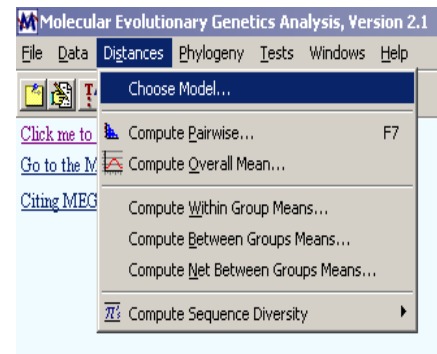
The *Sequence Data Explorer* shows the aligned sequence data. You can scroll along the alignment using the scrollbar at the bottom-right hand side of the explorer window. The *Sequence Data Explorer* provides a number of useful functionalities for exploring statistical attributes of the data and also for selecting data subsets (select-unselect taxa-groups among the rest of information of the data file for analysis, etc).



The *data explorer menu* provides access to commands for exporting the selected data subset to various formats, translating/un-translating the protein coding sequences, defining and including/excluding genes and domains, defining and including/excluding taxa and groups, etc. In studies of molecular evolution it is often necessary to know some basic statistical quantities such as nucleotide frequencies, codon frequencies, and transition/transversion ratios. You can compute these quantities in the *Sequence Data Explorer* dialog box.

Distances

When you choose to compute distances, this dialog box allows you to select the type of distance and options associated with it. Evolutionary distance between a pair of sequences is usually measured by the number of nucleotide (or amino acid) substitutions between them. Evolutionary distances are fundamental for the study of molecular evolution and are useful for phylogenetic reconstruction and estimation of divergence times. In MEGA2, most of the widely used methods for distance estimation for nucleotide and amino acid sequences are included (see this section detailed at Help File).



This dialog box is occupied by a set of overlapping pages, with each pages marked by Tabs that run across the top of the dialog box. You can go to any page by simply clicking on the Tab. Each tab page organizes a set of logically related options. Information from all the pages will be used in the requested analysis, so it is important that you examine the currently selected options in each tab before pressing OK to proceed with analysis.

Models of Substitution

In the top part, a model (method) for estimating evolutionary distance is selected by pressing the *Models* button. It will reveal a menu containing many different distance methods. The selected distance method is displayed next to the **Models** button in bold letters. Below, you may need to specify the Transition/Transversion ratio (R) and/or the Gamma parameter depending on the distance method chosen.

Substitutions to Include

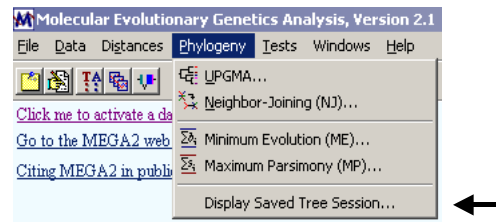
Depending on the distance method selected, the evolutionary distance can be teased apart into two or more components. The Substitutions to Include list box allows you to select one of these components. The list of components relevant to the chosen model for estimating distances is made available.

Compute

Use this to specify whether to compute Distances only or the Distances and Standard Errors. If you select the latter option, then you are given a choice on how to compute it in the *Standard Error Computation* box.

Phylogenetic analysis

Use the Phylogeny Menu to construct phylogenetic trees and also to view previously constructed trees (saved as ".mts" format by MEGA2.1).



There are numerous methods for constructing phylogenetic trees from molecular data. They can be classified into Distance methods, Parsimony methods, and Likelihood methods. You can find a sort explanation of each method at the Help File of the programme.

The following evolutionary models can be computed with MEGA2.1:

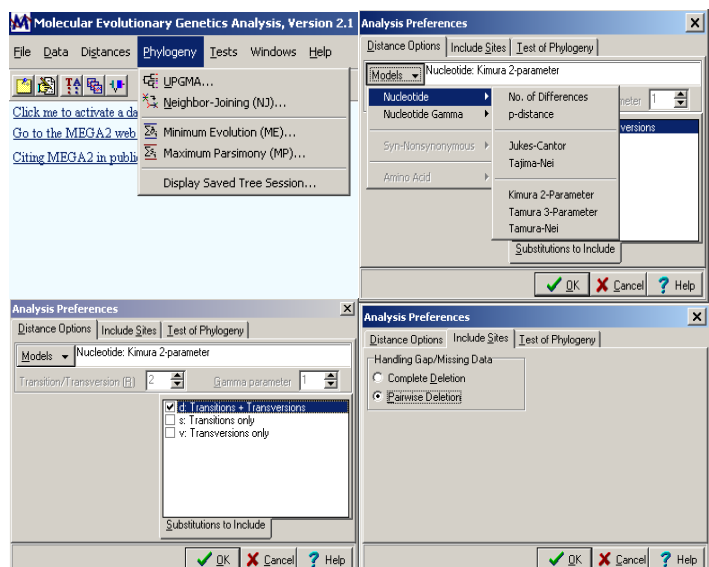
Nucleotide	Syn-Nonsynonymous	No. of differences
No. of differences	Nei-Gojobori Method	p-distance
p-distance	Modified Nei-Gojobori Method	Poisson correction
Jukes-Cantor distance (with Gamma model)	Li-Wu-Luo Method	Gamma distance
Tajima-Nei distance	Pamilo-Bianchi-Li Method	
Kimura 2-parameter distance (Gamma model)	Kumar Method	
Tamura 3-parameter distance	Amino acid	
Tamura-Nei distance (Gamma model)		

Commands from the “Phylogeny” menu

Distance options: This tab contains options for evolutionary models selection and for handling Transitions and/or Transversions analysis.

Include Sites tab: This tab contains options for handling gaps/missing data in the analysis, inclusion and exclusion of codon positions, and whether to restrict the analysis to labeled sites only (if applicable).

Handling Gaps/Missing Data: You may choose to remove all sites containing alignment gaps and missing-information sites before the calculation begins (*Complete-deletion* option).

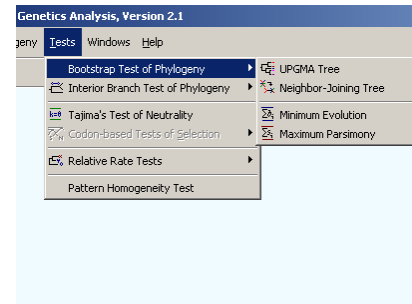


Alternatively, you may choose to retain all such sites initially, excluding them as necessary in the pairwise distance estimation (*Pairwise-deletion* option).

Codon Positions/Sites Included: In this box you may select any combination of 1st, 2nd, 3rd, and non-coding positions for analysis. It is available only if the nucleotide sequences contain protein-coding regions and you have selected nucleotide-by-nucleotide analysis.

Include only Labeled Sites: This section is enabled only if you have associated labels with some or all sites in the data.

Test of Phylogeny tab: Using this page you can select method to use for assessing the reliability of a phylogenetic tree.



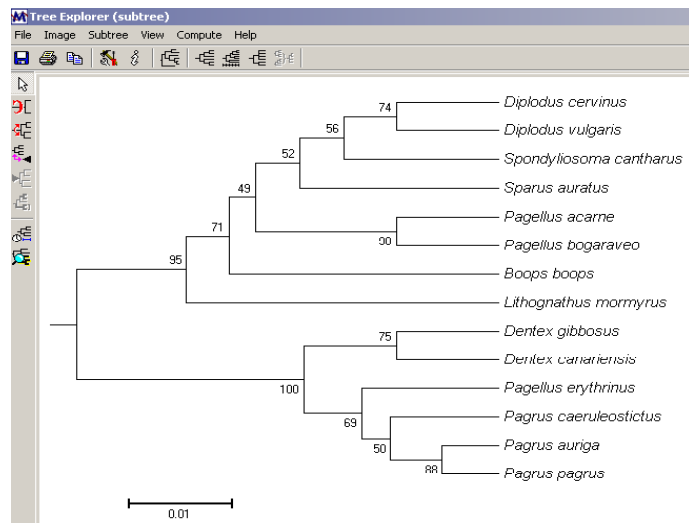
Number of Replications and Random Seed: All the tests of phylogeny included in *MEGA2.1* use the bootstrap re-sampling strategy. Therefore, you need to provide the number of replicates and a starting random seed for the chosen test (See *Help File* for further information).

Once selected the model and its options, the programme runs the phylogenetic analysis in seconds (the duration of this process depends on the amount of data, the method employed and the microprocessor chip of your computer).

Tree Explorer

The *Tree Explorer* displays the evolutionary tree based on the options used to compute or display the phylogeny.

Tree Explorer tab allows the user to manipulate aspects of the tree depending upon the style used to draw the tree. For instance, if the traditional rectangular style is used, then the user can manipulate the taxon separation distance, branch length, or tree width, in the number of pixels. This tab also contains a schematic of a tree illustrating these aspects. The *Tree Explorer* has the following items in the main menu:



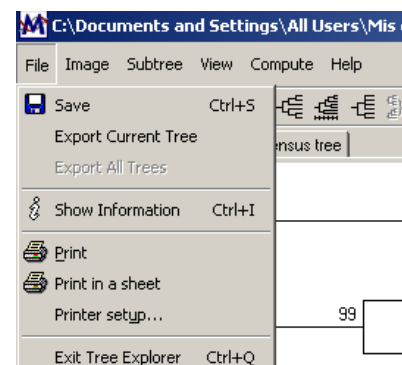
File Menu

This menu contains various file options:

Save: Brings up the Save As dialog box and saves all the information currently held by the Tree Explorer to a file in a binary format. This feature is provided to allow you to retrieve the current Tree Explorer session for tree manipulation and printing later.

Export Current Tree: Writes the topology of the current tree in the MEGA tree format to a specified file. Note that only the branching pattern is stored.

Export All Trees: Writes the topologies of all trees in the MEGA tree format to a specified file. Note that only the branching pattern is stored.



Show Information: Brings up the Information dialog box.

Print: Brings up the Print dialog box and prints the current tree in the displayed size. This means that if the displayed tree is larger than the page size, it will be printed in multiple pages.

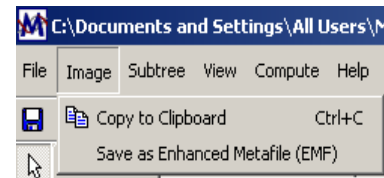
Print in a sheet: Brings up the Print dialog box and prints the current tree, after restricting the size of the printed tree to one sheet. The current tree can be printed using the button on the toolbar also.

Printer Setup: Allows the user to setup the printer.

Exit: Tree Explorer Exits.

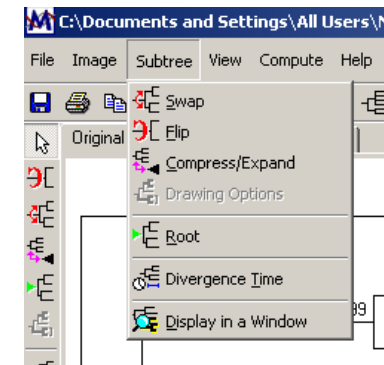
Image Menu

This contains two options, one to copy the tree image to the clipboard (this can also be done by simultaneously pressing Ctrl and 'C' keys), and the other to save the image as an enhanced windows metafile (.EMF). The second option brings up the *Save As* dialog box.



Subtree Menu

This menu contains various tree manipulation options, namely, *Swap*, *Flip* and *Compress/Uncompress*. In addition, the user can specify the root of the tree, and display a subtree (a portion of the tree defined by a given internal branch) in a separate window. This done by clicking on the corresponding items in the menu (for which there are tool buttons on the left).



Swap: Upon selecting this command if the user clicks on an interior branch, *MEGA2.1* then swaps the two subtrees defined by this branch. If the subtrees are each simply an individual taxon, then *Swap* is the same as *Flip*. The cursor may be reverted to the arrow by clicking on the arrow icon on the left hand side of the Tree Explorer.

Compress/Uncompress: The cursor changes to the 'Compress/Uncompress' icon when this command is selected. Upon selecting this command if the user clicks on an interior branch, *MEGA2.1* then prompts the user to give a name to the group that will be formed. It then compresses all the lineages defined by this branch into a solid elongated triangle, whose thickness is proportional to the number of taxa condensed. Clicking on the branch again uncompresses it.

View Menu

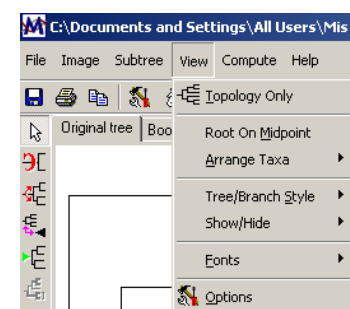
This menu brings up several viewing options:

Topology only: Displays the tree only in the form of relationships among the taxa, ignoring the branch lengths.

Root on Midpoint: This roots the tree on the midpoint of the longest path between two taxa.

Arrange Taxa: It allows you to arrange the taxa in the tree either based on the order of taxa in the input data file or to produce a tree that looks "balanced."

Tree/Branch Style: It allows you to select the display of the tree in one of three styles: *Traditional*, *Radiation*, or *Circle*. For traditional style, there are three additional options for branch display: *Rectangular*, *Straight* or *Curved*.



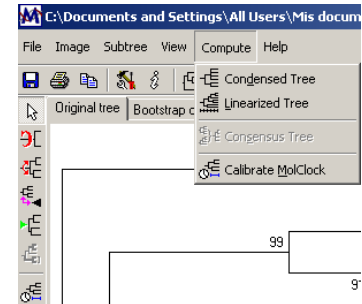
Show/Hide: It allows the user to display or hide the following information: taxon label, taxon marker, statistics (e.g., bootstrap values), branch lengths, or scale bar.

Fonts Use: it to choose the font type, size etc. for writing various information, including taxon label, statistics, scale bar etc.

Options: It brings up the Option dialog box, which provides control over various aspects of the tree drawing including individual branches, the taxon names, and the scale bar.

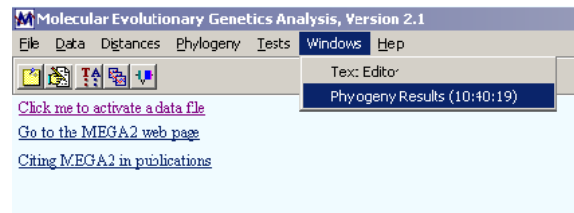
Compute Menu

Allows various computations from the tree, namely, Condensed tree, Linearized tree, and Consensus tree. Furthermore, it also allows the user to estimate the divergence time for each node using the molecular clock.



Windows Menu

Once the programme complete the computation, automatically a new *Tree Explorer* dialog box is opened. However, you can access to any of the opened windows selecting it in the command Window of the toolbar of the Main Window.



A practical exercise for the phylogenetic validation is presented on next page.

2. Practical Exercise of Fish Phylogeny with MEGA2.1

Note: All files required are placed in the same folder as this document.

1. Execute *MEGA2.1* in your PC.
2. Open the Text Editor.
3. Convert the example file (aligned sequences on *ClustalX*): ***Rhod_Phyl.aln*** to MEGA format.
4. Insert a new *header* including *title*, *description* and *format*. Example:

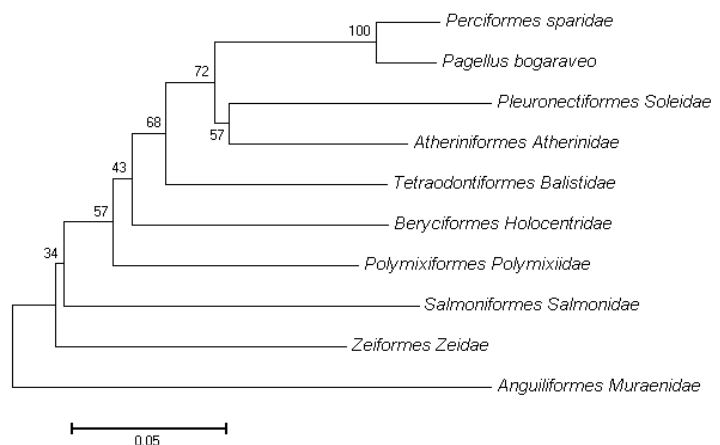
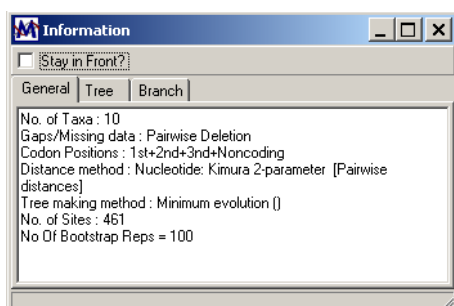
```
#mega
!TITLE FishTrace Rhod;
!Format DataType=DNA      Nseqs=9 NTaxa=9      NSites=460   CodeTable=Standard;
!Description              Phylogenetic Analysis for Genetic Data Validation;
```

5. Add a new Rhodopsin sequence to the Data File converted using the text editor (copy-paste the example shown below). Remember to increase the number of taxa (*NTaxa*) and sequences (*Nseqs*) in the *header* of the file). Example:

```
#Pagellus_bogaraveo
TCATGCCATCATGGCCTGGCCTTACCTGGACGATGGCCATGGCTTGTG
CCGCCCCCTCTCCTCGGCTGGTCTCGTTACATCCCTGAGGGAATGCAG
TGCTCATGTGGAATTGACTACTACACACGTGCTGAAGGTTTCAACAATGA
GTCCTTTGTCTATCTACATGTTTATCTGCCACTTCTCGATTCCAAGTACGA
TCGTGTTCTTCTGCTACGGCCGTCTGCTCTGTGCCGTCAAGGAGGCTGCT
GCTGCCAGCAGGAGTCTGAGACCACCCAGAGGGCTGAGAGGGAAGTCAC
CCGCATGGTTATTATGATGGTCATTGCTTTCTGGTTTGTGGCTTCCCT
ATGCCAGTGTAGCCTGGTATATCTTACACATCAAGGCTCACAGTTCGGC
CCAGTCTTCATGACCATCCCGCCTTCTTTGCCAAGAGCTCATCCATCTA
CAATCCAATG-
```

6. Save it as ***Rhod_Phyl_Chek.meg***. Once saved, active the data file: ***Rhod_Phyl_Chek.meg***.
7. Check the data executed using the Data Explorer (you can translate into protein sequences in this point. In some cases, amino acid differences give more solid information to compute the phylogenetics test).
8. Run and compare several phylogenetic methods. Check the results obtained and the location adopted by the new sequence inserted.

Results obtained after the Minimum Evolution Bootstrap Test of Phylogeny:



3. Phylogenetics Reconstructions of FishTrace Sequences with MEGA2.1

Along the duration of this project, the genetic teams obtain an important amount of nucleotide sequences of the target genes *Cytochrome B* (1141bp) and *Rhodopsin* (460bp). This information requires to be contrasted and validated for reliability. The first step of the validation process involves the alignment of these sequences and the subsequent register of all changes observed between two individuals belonging to the same fish species. Phylogenetic analysis is needed to carry out the second step of the sequence validation. In this document we show the phylogenetic analysis of several sequences (*Cytochrome B* and *Rhodopsin*) obtained from FishTrace's fish samples.

Genetic Tasks

Technical genetic protocols and standardised methodologies used in this project (treatment of fish samples received, DNA extraction, PCRs, sequencing, etc) were previously described, drawn up and placed at the webpage of the project. You can download this documents at: <http://pta.jrc.cec.eu.int> for further information.

Phylogeny Study

The sequences are aligned using *ClustalX* and phylogenetic analysis is performed with *MEGA2.1* software, following the process previously described. The phylogenetic study on *Cytochrome B* and *Rhodopsin* genes has been produced from a large amount of the acanthomorph diversity, including sequences obtained from UCM and other taken from GenBank (for reliability checking). We have obtained satisfactory results for both sets of sequence data.

Rhodopsin: We have analysed 83 sequences (45 of them obtained by UCM and the rest were taken from GenBank. See Fig. 1).

Cytochrome B: We have analysed 70 sequences (32 of them obtained by UCM and the rest were taken from GenBank) In the case of *Cytochrome B*, we used the *Translation* option of the *Sequence Data Explorer* to analyse the protein sequences because this option rendered the best trees (See Fig. 2).

Application for FishTrace Validation Purposes

This process has been designed to allow FishTrace's genetic teams to check the phylogeny position of the new sequences obtained.

1. You have to include your own sequence into the text files in MEGA format created by UCM and placed in the same folder as this document (one file for each gene):
 - ***RHOD_83_FT.meg***
 - ***CYTB_Protein_70_FT.meg***
2. Note that you have to increase the number of taxa and sequences in the *header* of each modified file (*NTaxa* and *Nseqs*).
3. Check the results obtained and pay attention to the position of the new sequence in the tree.

In the next pages are shown the conditions used and the trees obtained in the analysis of both gene sequences (*RHOD_83_FT.meg* and *CYTB_Protein_70_FT.meg*). These trees have a good correlation between taxonomy and molecular topology, and many sequences tested individually have been always located within their corresponding clades. These trees should help you to check and validate your results. Note that the different families within the Perciformes order appear scattered along Teleost phylogeny. This observation agrees with previous analysis by other authors using different genes, including complete mitochondrial genomes³.

³ Miya M, Takeshima H, Endo H, Ishiguro NB, Inoue JG, Mukai T, Satoh TP, Yamaguchi M, Kawaguchi A, Mabuchi K, Shirai SM, Nishida M. *Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences.* Mol Phylogenet Evol. 2003 Jan;26(1):121-38.

OPTIMAL CONDITIONS FOR THE ANALYSIS OF *Rhodopsin*

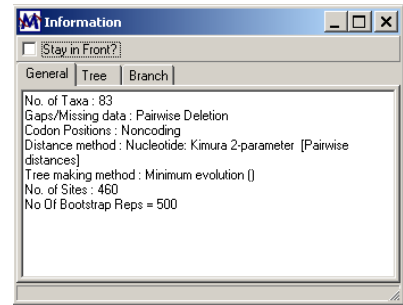
MODEL: Minimum Evolution (ME).

DISTANCE OPTIONS: Nucleotide: Kimura 2-parameter.

Ti / Tv HANDLING: Transition + Transversions.

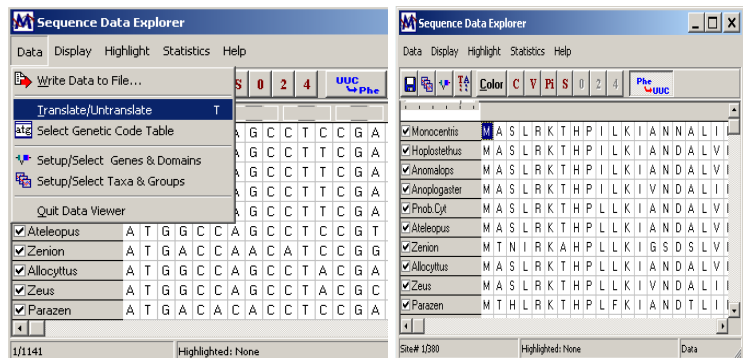
INCLUDE SITES: Gaps/Missing Data: Pairwise Deletion.

TEST OF PHYLOGENY: Bootstrap, 500 replications.

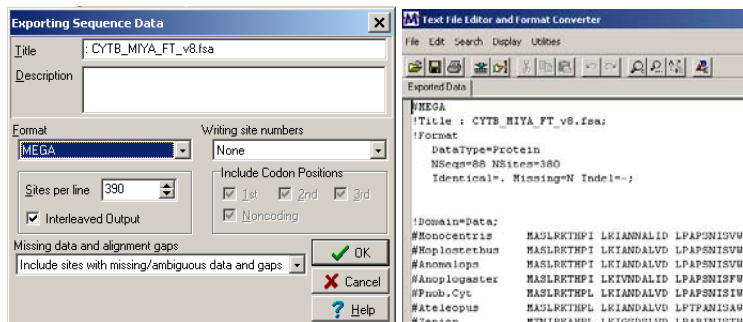


OPTIMAL CONDITIONS FOR THE ANALYSIS OF *Cytochrome B*

TRANSLATION INTO PROTEIN ALIGNMENT: The *Cytochrome B* nucleotide sequence alignment was translated into protein alignment by using the *Translate/Untranslate* option in the *Data Explorer* menu in *MEGA2.1*.



Use the *Write Data to File* option for exporting the protein alignment to the *Text Editor*. Then, the protein alignment can be saved as a new file. This new file (.meg) has the format required for phylogenetic analysis under *MEGA2.1* software.

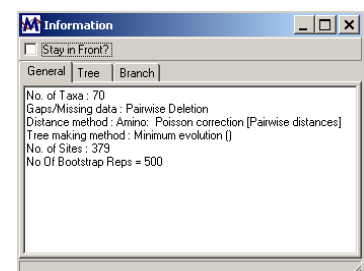


MODEL: Minimum Evolution (ME).

DISTANCE OPTIONS: Amino: Poisson correction.

INCLUDE SITES: Gaps/Missing Data: Pairwise Deletion.

TEST OF PHYLOGENY: Bootstrap, 500 replications.



Rhodopsin phylogenetic analysis

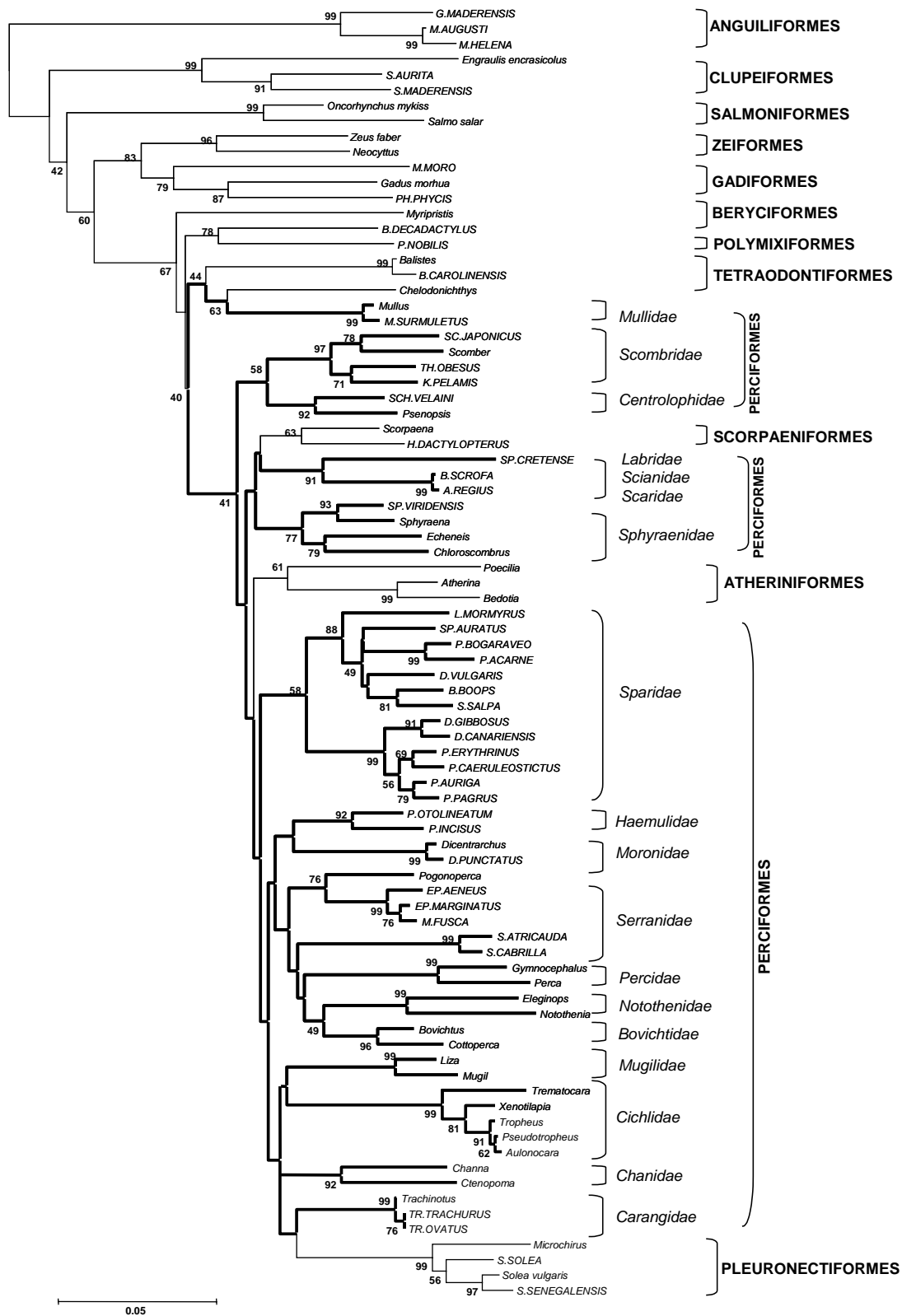


Fig 1. Minimum evolution (ME) tree from the partial Rhodopsin sequences (460bp). Taxa indicated in upper case correspond to sequences obtained in FishTrace, the rest (in lower case) were taken from GenBank to contrast the reliability of the method employed for the analysis. Bold lines represent the topology of Perciformes order in this analysis. Numbers represent bootstrap percentage from 500 replicates.

Cytochrome B phylogenetic analysis

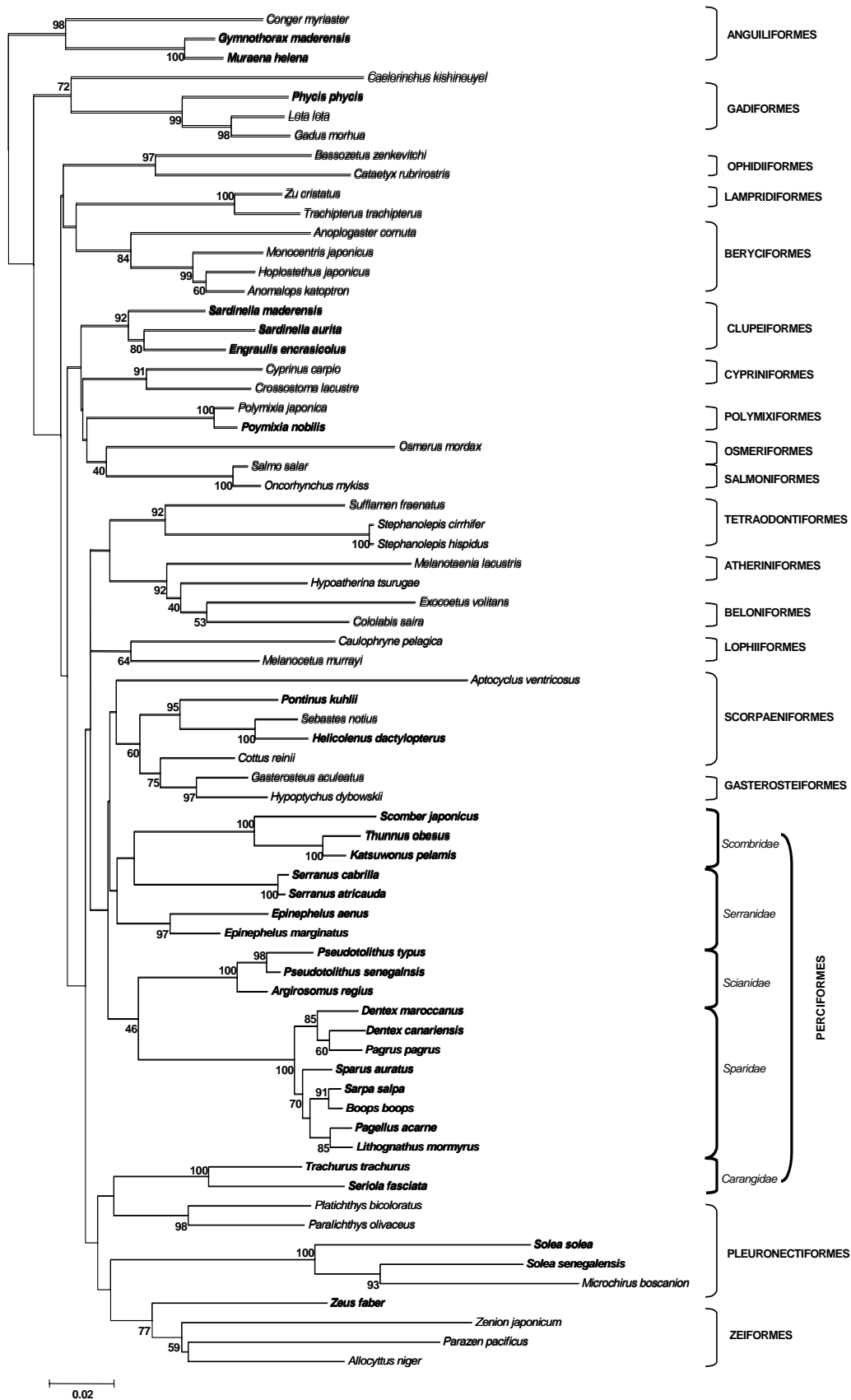
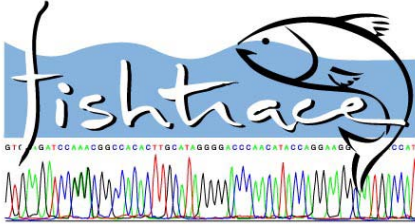


Fig 2. Minimum evolution (ME) tree from Cytochrome B protein sequences. Numbers represent bootstrap percentage from 500 replicates. Taxa indicated in bold correspond to sequences obtained in FishTrace, the rest were taken from GenBank to contrast the reliability of the method employed for the analysis.

Annex XIV: Preparing sequence files for the Sequin tool.



FishTrace

Submission of sequences to GenBank / EMBL:
Preparing sequence files for the Sequin tool

1. Convert each sequence for the species you are responsible in the **FASTA** format.
See: *Species_Lists_Validation_(genetics).xls* at <http://pta.jrc.cec.eu.int/>
2. Group all sequences from specimens of the same species (including those from the different areas) so that these sequences will be submitted as a **"Population study"** for each species.
3. The **Sequence ID** for each specimen (the first information to be included in the first line of each sequence) should be in the FishTrace specimen format (e.g. EngEnc-EM-01).
4. Sequence **modifiers** for which you should provide information for each sequence (following the sequence ID on the FASTA file) include: **1)** the scientific name of the species, **2)** the tissue type from which the DNA was isolated, **3)** the specimen-voucher code, **4)** the country of origin of the specimen, **5)** the locality that this specimen was isolated (if known), **6)** the sex of the specimen (if known), **7)** the developmental stage (if known), **8)** the name of the person that named the species (Authority), and **9)** the common name of the species (**FAO name**). The order of appearance of the modifiers is not important. Do not leave a space before or after the "=" sign! An example of the format of the modifiers is given below:

```
>DenDen-EM-01 [org=Dentex dentex] [tissue-type=muscle] [specimen-voucher=DenDen-EM-01] [country=Greece] [isolation locality=Kavala] [sex=female] [dev-stage=adult] [authority=Risso] [common=common dentex]
```

5. For many of the fish species that we are examining, the taxonomic lineage is not available in the EMBL/GENBANK databases. Please provide this information for each species in the following format:

```
Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii;  
Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio
```

6. The contact person for a set of sequences is the person responsible for the validation of the genetic data for the species as these have been assigned to each of the groups involved in the task, i.e. either José, Monique, Hilde, Mike, or Greg. Please provide also your address, tel, fax and e-mail along with the sequence files.
7. In the authors list we will include the names of all people involved in the sequencing effort as well as those providing the samples. Thus, for a species validated at NAGREF and for which sequences and/or specimens were provided by UCM, RIVO, and IFREMER, the names of all people in the above groups should be included. Please, collect and provide this information. For example, for samples and sequences provided by NAGREF the names to be listed are L. Favre-Krey, P. Leontarakis, A. Tsangridis, and G. Krey.
8. Once the sequences of a species are validated in your lab at level 1 (see the following schema), there are two possibilities:

8.1. If you are also responsible for validating at level 2 (see below), then you have to fill in a Sequin tool format and ask the partner(s) having the shared species (see: *Species_Lists_Validation_genetics.xls* at <http://pta.irc.cec.eu.int/>) for the corresponding sequences. **Copy** and **paste** the sequences from the other partners in your species file and validate the sequences according to agreed protocol (see below).

8.2. If you are not responsible for validating that species, send your Sequin tool format file to the responsible partner for validation.

<p>Level 1: Internal validation by each partner:</p> <ol style="list-style-type: none"> 1) To compare both sequences obtained from individuals 1 and 2 (Aligning) 2) To annotate every change on the sequence. 3) Verifying the position of each sequence on fish phylogeny. 	<p>Level 2: Validation by the responsible group:</p> <p>- <i>Species_Lists_Validation_genetics.xls</i> -</p> <ol style="list-style-type: none"> 1) To align and compare all data submitted by each partner. 2) To annotate every change on the sequences and methods employed. 3) Phylogenetics (BLAST, Trees,...) 4) Submit info to the WP7 responsible. 	<p>Level 3: Definitive validation by WP7 responsible:</p> <ol style="list-style-type: none"> 1) Specific data checking (errors, missing data,...) 2) Specific data compilation. 3) Specific database field validation. 4) Data collating and arranging.
--	--	--

9. Once validated, each Sequin tool file should be filled in and sent by email to **UCM** and **NAGREF**, since all submissions will be made at once by both partners in charge of **WP3** (UCM) and **WP4** (NAGREF).

Emails for sending validated files: rsevilla@vet.ucm.es (UCM) and krey@otenet.gr (NAGREF).

Sequin tool format file example:

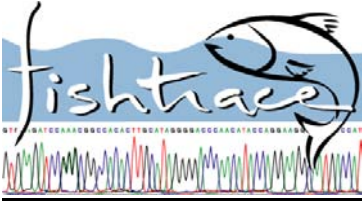
```
>DenDen_EM_01 [org=Dentex dentex] [tissue-type=muscle] [specimen-voucher=DenDen-EM-01]
[country=Greece] [isolation locality=Kavala] [sex=female] [authority=Risso] [common=common dentex]
TCATGCCATCATGGGTCTGGCCTTCACCTGGACGATGGCCTTGGCTTGTGCTGCCCCCCTCTTGTGGCTGGTCTCGTT
ACATCCCTGAGGGAATGCAGTGCTCATGTGGAATTGACTACTACACACGCTGAAGGTTTCAACAATGAGTCCTTTGTT
GTCTACATGTTACCTGTCACTTCATGACTCCACTGATCGTGTGTTCTTCTGCTACGGCCGTCTGCTCTGTGCTGCAA
GGAGGCTGTGCTGCCAGCAGGAGTCTGAGACCACCCAGAGGGCTGAGAGGGAAGTCACCCGCATGGTCGTACTIONCATGG
TCATCGCTTCCTGGTATGTTGGGTGCCCTATGCCAGTGTAGCCTGGTGGATCTTACACATCAAGGCTCAGAGTTCGGA
CCAGTCTTCATGACCATCCCGCCTCTTTGCCAAGAGCTCATCCATCTACAATCCAATG

>DenDen_CI_01 [org=Dentex dentex] [tissue-type=muscle] [specimen-voucher=DenDen-CI-01]
[country=Spain] [isolation locality=Las Palmas] [sex=female] [authority=Risso] [common=common
dentex]
TCATGCCATCATGGGTCTGGCCTTCACCTGGACGATGGCCTTGGCTTGTGCTGCCCCCCTCTTGTGGCTGGTCTCGTT
ACATCCCTGAGGGAATGCAGTGCTCATGTGGAATTGACTACTACACACGCTGAAGGTTTCAACAATGAGTCCTTTGTT
GTCTACATGTTACCTGTCACTTCATGACTCCACTGATCGTGTGTTCTTCTGCTACGGCCGTCTGCTCTGTGCTGCAA
GGAGGCTGTGCTGCCAGCAGGAGTCTGAGACCACCCAGAGGGCTGAGAGGGAAGTCACCCGCATGGTCGTACTIONCATGG
TCATCGCTTCCTGGTATGTTGGGTGCCCTATGCCAGTGTAGCCTGGTGGATCTTACACATCAAGGCTCAGAGTTCGGA
CCAGTCTTCATGACCATCCCGCCTCTTTGCCAAGAGCTCATCCATCTACAATCCAATG

>DenDen_MA_01 [org=Dentex dentex] [tissue-type=muscle] [specimen-voucher=DenDen-MA-01]
[country=Portugal] [isolation locality=Funchal] [sex=female] [authority=Risso] [common=common
dentex]
TCATGCCATCATGGGTCTGGCCTTCACCTGGACGATGGCCTTGGCTTGTGCTGCCCCCCTCTTGTGGCTGGTCTCGTT
ACATCCCTGAGGGAATGCAGTGCTCATGTGGAATTGACTACTACACACGCTGAAGGTTTCAACAATGAGTCCTTTGTT
GTCTACATGTTACCTGTCACTTCATGACTCCACTGATCGTGTGTTCTTCTGCTACGGCCGTCTGCTCTGTGCTGCAA
GGAGGCTGTGCTGCCAGCAGGAGTCTGAGACCACCCAGAGGGCTGAGAGGGAAGTCACCCGCATGGTCGTACTIONCATGG
TCATCGCTTCCTGGTATGTTGGGTGCCCTATGCCAGTGTAGCCTGGTGGATCTTACACATCAAGGCTCAGAGTTCGGA
CCAGTCTTCATGACCATCCCGCCTCTTTGCCAAGAGCTCATCCATCTACAATCCAATG

>Den_Den_05 [org=Dentex dentex] [tissue-type=muscle] [specimen-voucher=Gb-01] [country=France]
[isolation locality=Brest] [sex=female] [authority=Risso] [common=common dentex]
TCATGCCATCATGGGTCTGGCCTTCACCTGGACGATGGCCTTGGCTTGTGCTGCCCCCCTCTTGTGGCTGGTCTCGTT
ACATCCCTGAGGGAATGCAGTGCTCATGTGGAATTGACTACTACACACGCTGAAGGTTTCAACAATGAGTCCTTTGTT
GTCTACATGTTACCTGTCACTTCATGACTCCACTGATCGTGTGTTCTTCTGCTACGGCCGTCTGCTCTGTGCTGCAA
GGAGGCTGTGCTGCCAGCAGGAGTCTGAGACCACCCAGAGGGCTGAGAGGGAAGTCACCCGCATGGTCGTACTIONCATGG
TCATCGCTTCCTGGTATGTTGGGTGCCCTATGCCAGTGTAGCCTGGTGGATCTTACACATCAAGGCTCAGAGTTCGGA
CCAGTCTTCATGACCATCCCGCCTCTTTGCCAAGAGCTCATCCATCTACAATCCAATG
```

Annex XV: Reference Collections (WP5) Protocol, Loan Request Form and Invoice of Specimens.



FISHTRACE WP5 PROTOCOL

Working proposal prepared by IMAR-MMF

There will be 4 official FishTrace reference collection centres: MNHN, NRM, TFMC and MMF. Each Museum shall designate a Curator responsible for these collections (MNHN: Guy Duhamel; NRM: Sven Kullander; TFMC: Fátima Hernández and MMF: Manuel Biscoito).

In each Museum there will be four FishTrace reference collections:

- 1) Voucher specimens
- 2) Muscle tissues
- 3) Otoliths
- 4) DNA samples.

After sampling (according to particular protocol) the resulting collections shall be permanently stored in each FishTrace reference collection centre, according to the following protocol:

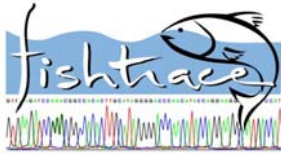
1 – Building up (during FishTrace)

- All specimens (voucher, tissue, otoliths and DNA samples) must be labelled in accordance to FishTrace WP2 protocol prior to incorporation in the respective reference collection.
- Optionally, each specimen shall get an additional registering number and label according to particular procedures of each Museum.
- Each voucher specimen shall be preserved in 70% ethanol whenever possible.
- Tissue samples shall be preserved in 95% ethanol in a dark cool place.
- Otoliths shall be stored dry (each side separate).
- DNA samples shall be kept frozen at -20°C .
- Photographs – 1 CD

2 – Long-term preservation conditions (During and post-FishTrace)

- All four Museums agree to incorporate the FishTrace collections in their own ones and therefore keep them well preserved *ad infinitum*.
- In case one of the four Museums is in need to dispose its FishTrace collections, the other three must be contacted in order to determine the final destination of the collections in question.

- 3 – Availability, exchange, loans and gifts (During and post-FishTrace)
- All Museums agree to make FishTrace collections fully available for the use of the FishTrace consortium, during the project. After the end of FishTrace, each Museum's policy applies to FishTrace collections. The Consortium will retain exclusive rights over the samples for one year after the end of FishTrace
 - Genetically validated voucher specimens and related samples (otoliths and tissues) shall remain available to members of FishTrace consortium on loan. Precautions should be taken in order to avoid damage or loss of specimens during transport.
 - Specimens 4 and 5 (see WP2 protocol) can be exchanged, loaned or donated to any FishTrace Museum, upon request.
 - During FishTrace project, the following loan conditions apply:
 - a) Loans are made between members of FishTrace consortium.
 - b) Requests shall be directed the respective Curator (Form?)
 - c) Specimens on loan must be accompanied by an Invoice (see Invoice of Specimens)
 - d) Loans are made for periods of 3 or 6 months, renewable upon request.
 - e) Loans cannot be refused to members of the FishTrace consortium. Loans to third party institutions should only be granted under exceptional conditions, after authorization of the FishTrace Coordinator and assuring the agreement of the partners. In any case, FishTrace specimens/samples must not be loaned to private/profit/commercial individuals or institutions.
 - f) The borrower is responsible for the good preservation conditions of the specimens received on loan and cannot change preservation medium without prior consent of the curator of the collection (loaner).
 - g) Dissecting, clearing and staining, cutting or any other intrusive/damaging procedure cannot be done without prior consent of the curator of the collection (loaner).
 - h) The borrower cannot transfer specimens on loan to any other individuals or institutions without prior consent of the curator of the collection (loaner). In this case a new loan contract must be done.
 - i) Each Curator shall decide the best way to send a loan (e.g. courier, air mail, by hand, etc.).



www.fishtrace.org

BIOLOGICAL COLLECTIONS LOAN REQUEST FORM

- TO: Muséum national d'Histoire naturelle in Paris (MNHN)
 Natural History Museum in Stockholm (NRM)
 Museo de Ciencias Naturales de Tenerife (TFMC)
 Museu Municipal do Funchal (História Natural) (MMF)

(use different forms for requests to more than one Museum)

The undersigned request the loan of the following FishTrace specimens for (purpose of study): _____

For a period of _____ months.

Museum collection no.	FishTrace Code	Species	Type of specimen: Voucher, backup specimen, tissue sample, otoliths, DNA sample

Add extra lines if necessary

Identification of requesting Institution

Name:

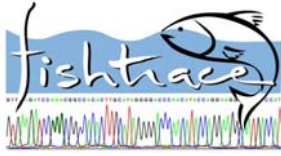
Person in charge:

Address:

E-mail:

Telephone:

Fax:



BIOLOGICAL COLLECTIONS FORM

NAME OF THE MUSEUM

INVOICE OF SPECIMENS NO.	<input type="text"/>
_____, Date	

TO:

Loan

Return

Gift

Exchange

Loan Period

Months

<i>FishTrace Code</i>	Species	Type of specimen: Voucher, backup specimen, tissue sample, otoliths, DNA sample	Preservation medium	Observations

Shipping method	Received in good order	<input type="checkbox"/>
	Except as noted	<input type="checkbox"/>
Sent by	Received by	
Signature	Signature	

MUSEUM ADDRESS

Annex XVI: Guidelines for Validation purposes (WP7) including a protocol defining format for the Validation tasks.

WP 7 DATA VALIDATION IN DATABASE

GUIDELINES FOR VALIDATION PURPOSES

Responsible partner:
ICCM (J.A. González)

Working Group:
**S. Jiménez (TFMC), I.J. Lozano (ICCM), M. Biscoito (IMAR),
P. Pruvost (MNHN) & J.A. González (ICCM)**

1. OBJECTIVES

Data compiled in the database are being generated by different teams of taxonomists, geneticists and curators. This important amount of valuable information requires to be checked for reliability and standardisation among the different data sets and the various groups involved.

2. PARTNERS/GROUPS INVOLVED

Data on Sampling, Taxonomy and Collections, 5 groups:

ICCM-TFMC, IFREMER-MNHN, IMAR, NAGREF and NRM.

Data on Molecular Genetics, 5 groups:

IFREMER, NAGREF, NRM, RIVO and UCM.

3. THE VALIDATION PROCESS

3.1. EARLY VALIDATION

Each group is responsible for checking data generated by itself in order to detect possible errors or missing data.

Each genetic group is responsible for detecting possible errors related to the species identification.

3.2. SPECIES ASSIGNMENT

As there was agreed during the Ispra Meeting, once entered in the off-line database, sampled species have to be assigned to partners/groups involved in order to proceed to the next step on the validation process.

The geographic proximity or familiarity criterion was accorded as a good criterion for the assignment of species to partners.

Explanation: Partners should be more familiarised with their own regional species for validation purposes of data sets on taxonomy, description, biology, images and bibliography.

(Nearly) final lists of assigned species to partners for validation purposes can be found here included. Two different groups of lists have been prepared in order to validate sampling-taxonomy-collections and molecular genetics data.

4. PROTOCOL DEFINING FORMAT FOR VALIDATION PURPOSES

4.1. COMMON FIELDS IN THE DATABASE SPECIES TABLE (TASKS TO BE DONE FOR EACH ASSIGNED SPECIES)

LABEL & FIELD: Popular description

Explanation

In this field there will be a short popular description of the main external characters of the species, with particular emphasis on body shape, number, shape and position of fins, general colour pattern and any other striking external characteristic. Also bear in mind differences between sexes and common size.

Example

This is a moderately large amber jack. Its body is nearly cylindrical with upper profile of head convex.

It has one dorsal fin, the anterior part with 8 spines, followed by a soft second part with a conspicuous lobe at its beginning. The base of anal fin is much shorter than that of dorsal. The caudal fin is prominently forked.

Adult specimens are usually pinkish and small juveniles show dark vertical bars.

Task

Fill carefully the appropriate popular information on body shape, number, shape and position of fins, general colour pattern and any other striking external characteristic for the assigned species in the database.

How

By using the appropriate information from FishTrace data and images, as well as from the UNESCO, FAO Catalogues and other suitable bibliographic sources.

LABEL: Diagnosis

Following the field sequence given below.

FIELD: Morphology

Explanation

In this field there will be a scientific description of the morphological characters for the species pertaining to body, mouth shape and caudal region description.

Example

Body fusiform; upper profile of head convex; end of upper jaw relatively slender; grooves present on caudal peduncle.

Task

Fill carefully the appropriate scientific information on body, mouth shape and caudal region description for the assigned species in the database.

How

By using the appropriate information from the UNESCO, FAO Catalogues and other suitable bibliographic sources, as well as taking into account the FishTrace species information (both data and images) from every sampled specimen.

FIELD: Fins

Explanation

In this field there will be a scientific formula for the dorsal (D, D(1), D(2), D(3)), anal (A, A(1), A(2)), pectoral (P) and pelvic (V, only when taxonomically relevant) fin rays, indicating the number of spines (in Roman characters) and soft rays (in Arabic characters) when present.

Examples

D XII + 13-15

A III + 13-15

P 17

V I + 5

D VIII + I + 28-33

A II + I + 17-20

P 17

V I + 5

D(1) 14-15; D(2) 18-22; D(3) 17-20

A(1) 19-23; A(2) 17-19

P 17

V 5

Task

Fill carefully the appropriate scientific consensus formula for the dorsal, anal, pectoral and pelvic fins.

How we must validate

Step 1: By compiling the fins' formulae for each species from UNESCO and FAO Catalogues and other sources.

Step 2: By checking (comparing minimum and maximum values) each specimen's fin data against the species' formula previously compiled.

Step 3: When discrepancies between data in the DB and data from main specific references are found, please contact with the partner(s) responsible for the sampling task in question, in order to clarify this deviation.

Step 4: By elaborating the final fins' formulae for each species by collating together and arranging, when appropriate, the consensus fins formulae.

Important: How to count last dorsal and anal fin ray, which sometimes is bifurcate? As one or two rays? We count them as one.

FIELD: Lateral line

Explanation

In this field there will be included information about the lateral line (presence or absence), indicating the type and number of scales, scutes and pores when present.

Example

Present, curved, with 31-36 scales and 0-3 scutes.

Task

Fill carefully the appropriate scientific consensus formula and shape for the lateral line(s), when present.

How we must validate

Step 1: By compiling the lateral line formula and shape for each species from UNESCO and FAO Catalogues and other sources.

Step 2: By checking (comparing minimum and maximum values) each specimen's lateral line data against the species' characteristics previously compiled.

Step 3: When discrepancies between data in the DB and data from main specific references are found, please contact with the partner(s) responsible for the sampling task in question, in order to clarify this deviation.

Step 4: By elaborating the final lateral line formula and shape for each species by collating together and arranging, when appropriate, the consensus lateral line formula.

FIELD: Gillrakers in first gill arch

Explanation

In this field there will be included the formula of the gillrakers in first gill arch, indicating the presence of rudiments when relevant. The format will be number of rakers in the upper branch + number of rakers in the lower branch of 1st gill arch.

Examples

10+26 or 9+1+26, or

9-11+24-27 or 9-10+1+24-27 (when we have ranges)

Task

Fill carefully the appropriate scientific consensus formula of the gillrakers in first gill arch for the species.

How we must validate

Step 1: By compiling the gillrakers formula for each species from UNESCO and FAO Catalogues and other sources.

Step 2: By checking (comparing minimum and maximum values) each specimen's gillrakers in first gill arch data against the species' characteristics previously compiled.

Step 3: When discrepancies between data in the DB and data from main specific references are found, please contact with the partner(s) responsible for the sampling task in question, in order to clarify this deviation.

Step 4: By elaborating the final gillrakers' formula for each species by collating together and arranging, when appropriate, the consensus formula for the gillrakers in first gill arch.

FIELD: Vertebrae

Explanation

In this field there will be included, when relevant, the vertebral formula.

Example

Total vertebrae: 139-143.

Task

Fill carefully the appropriate scientific information on vertebral formula for the assigned species in the database.

How

By using the appropriate information from the UNESCO, FAO Catalogues and other suitable bibliographic sources.

FIELD: Colour

Explanation

In this field there will be a full colour description of the species, with particular emphasis on the sexual chromatic dimorphism and different colour pattern related with age when existing.

Example

Adults darker (pinkish or violet) dorsally, sides lighter, and belly white or silvery. Juveniles smaller than about 20 cm FL with interradiial membranes of soft dorsal and anal fins lightly pigmented; juveniles about 4-26 cm FL with 7 dark body bars, irregular and broken, third to seventh extending onto the membranes of the second dorsal and anal fins, and the eighth bar, small and dark, at end of caudal peduncle.

Task

Fill carefully the appropriate information for the assigned species.

How we must validate

Step 1: By compiling the colour pattern for each species from UNESCO and FAO Catalogues and other sources.

Step 2: By checking the specific images with the information previously compiled.

Step 3: When any specific colour pattern differs from those of the main references, please contact with the partner(s) responsible for this sampling task in order to clarify this deviation.

Step 4: By elaborating the colour pattern taking into account whether we are dealing with fresh or post-mortem fish.

FIELD: Size

Explanation

In this field there will be an indication of the maximum size reached by the species and also the maximum weight when relevant.

Important: We prefer the standard length (SL), except for the cases in which this measurement cannot be taken.

Example

To 67.5 cm SL (or FL or TL) and 4.6 kg.

Task

Fill carefully the appropriate values for the assigned species in the database.

How

By using the appropriate size information from the UNESCO, FAO Catalogues and other suitable bibliographic sources.

LABEL & FIELD: Similar species

Explanation

In this field include any similar species which can be misidentified, mentioning separating characters. Images can also be added, if useful. [In table “Pictures”]

Example

Adult specimens can be confused with other *Seriola* species in the European seas. It differs from *S. dumerili* [figure] and *S. rivoliana* [figure] in having a narrower supramaxillar and from *S. rivoliana* and *S. carpenteri* [figure] in having a shorter lobe on the anterior part of the soft dorsal fin.

Task

Fill carefully the appropriate information on similar species.

How

By using the appropriate bibliographic information from FAO Catalogues and other suitable regional data on misidentification.

LABEL: Biology

Following the field sequence given below.

FIELD: Habitat

Explanation

In this field give a description of the more relevant habitat information (environment, substrata, depth range) on the species.

Example

Adults occur near to or on bottom at a depth of 55-130 m. Larger juveniles are pelagic or benthic in shelf waters and smaller juveniles are epipelagic in oceanic or offshore neritic waters.

Task

Fill carefully the appropriate information on habitat characteristics of the assigned species.

How

By using the appropriate bibliographic information from UNESCO, FAO Catalogues and other suitable regional sources.

FIELD: Behaviour

Explanation

In this field give a description of the more relevant behaviour information (schooling, migrations, etc.) on the species.

Example

It has an affinity for wrecks and tends to form small schools.

Task

Fill carefully the appropriate information on behaviour characteristics of the assigned species.

How

By using the appropriate bibliographic information from UNESCO, FAO Catalogues and other suitable regional sources.

FIELD: Sexuality

Explanation

In this field give a description of the more relevant information about sexuality (gonochorism versus hermaphroditism) on the species.

Examples

There is no data on sexuality.

Functional synchronic hermaphroditism.

Protandric hermaphrodite with sex reversal at 40-45 cm SL (or TL or FL).

Task

Fill carefully the appropriate information on sexuality pattern and type for the assigned species.

How

By using the appropriate bibliographic information from UNESCO, FAO Catalogues and other suitable regional sources.

FIELD: Reproduction

Explanation

In this field give a description of the more relevant information about reproduction (maturity size and spawning season) of the species.

Examples

There is no data on reproduction.

Reproduction takes place probably during summer.

Size at first maturity 23 cm SL (or TL or FL). Spawning season from March to June with a peak during April.

Task

Fill carefully the appropriate information on reproduction for the assigned species.

How

By using the appropriate bibliographic information from UNESCO, FAO Catalogues and other suitable regional sources.

FIELD: Feeding habits

Explanation

In this field give a description of the more relevant information about feeding habits and main preys on the species.

Example

Carnivorous with adults feeding on squids and fishes.

Task

Fill carefully the appropriate information on feeding habits for the assigned species.

How

By using the appropriate bibliographic information from UNESCO, FAO Catalogues and other suitable regional sources.

LABEL: Distribution and interest to fisheries

Following the field sequence given below.

FIELD: Climatic pattern

Explanation

In this field give geographical information on climatic or biogeographical pattern for the species.

Example

Subtropical, in warm waters of the Atlantic Ocean.

Task

Fill carefully the appropriate information on climatic or biogeographical pattern of the assigned species.

How

By using the appropriate bibliographic information from FishBase, UNESCO, FAO Catalogues and other suitable regional sources.

FIELD: Latitudinal range and distribution

Explanation

In this field give geographical information on latitudinal range and distribution for the species.

Example

From 42°N to 36°S in the Eastern Atlantic Ocean. Known with certainty only from the archipelagos of Madeira, Canaries and Cape Verde.

Task

Fill carefully the appropriate information on latitudinal range and distribution of the assigned species.

How

By using the appropriate bibliographic information from FishBase, UNESCO, FAO Catalogues and other suitable regional sources.

FIELD: Interest to fisheries

Explanation

In this field give geographical information on main fisheries, aquaculture and game fishing when relevant for the species.

Examples

Most probably with economic importance locally, but data are not available.

Commercial fisheries: yes; aquaculture production: yes; game fishing: yes.

Task

Fill carefully the appropriate information on interest to fisheries for the assigned species.

How

By using the appropriate bibliographic information from FishBase, UNESCO, FAO Catalogues and other suitable regional sources.

LABEL & FIELD: Conservation status

Explanation

In this field, list species' inclusion in Red Data Books, CITES ...

Example

Not in IUCN Red List.

Task

Fill carefully the appropriate information on conservation status of the assigned species.

How

By using the appropriate bibliographic information from the IUCN Red List, FishBase, UNESCO, FAO Catalogues and other suitable regional sources.

LABEL & FIELD: Sources of information

Explanation

List relevant sources of information used (FNAM, FAO, ETI, FishBase, PescaBase, etc.) using FishBase reference number.

Example

As we are dealing here with general bibliographic information, we prefer the way

[FB ref: 230](#)

Instead of

Hureau, J.-C. and Th. Monod (eds.), 1979. Check-list of the fishes of the north-eastern Atlantic and of the Mediterranean (CLOFNAM I and II). United Nations Educational Scientific and Cultural Organization (UNESCO), Paris, France. Vols 1-2. 683 p.

or

Hureau, J.-C. and Th. Monod (eds.), 1979.

Task

Fill carefully the appropriate list of sources for the assigned species.

How

By using the appropriate bibliographic references from the IUCN Red List, Fishbase, UNESCO, FAO Catalogues and other suitable regional sources.

4.2. REGIONAL FIELDS IN THE DATABASE REGIONAL INFO TABLE (TASKS TO BE DONE FOR EACH SAMPLED SPECIES)

Aim

Enter relevant regional information for the species pertaining to common names, biology, habitat, sizes, interest to fisheries, commercialisation and any other relevant remarks.

General explanation

According to the database structure, please proceed as follows:

Field: Id_regional_info

Enter the identification regional information.

THIS IDENTIFICATION CODE NEEDS CLARIFICATION IN ORDER TO FACILITATE AN EASY ACCESS AND DATA ENTERING WHICH WILL BE USED TO FILL IN THE SPECIFIC COMMON FIELDS' INFORMATION.

Field: Species

Select the species.

Field: Area

Select the FishTrace area.

Field: Label

Select one label from: Common names, Biology, Habitat, Size characteristics, Interest to fisheries, Commercialisation, Conservation status and threats, and Other.

Field: Topic

For each Label, select one topic from those included in the box, i.e., **Label** Interest to fisheries; **Topics**: Fisheries importance or Fishing gears.

Field: Remarks

Fill the appropriate information on the selected topic.

Field: Reference

Enter the complete regional reference(s) related to the information on the selected topic.

Detailed explanation and examples**Label “Common names”****Topic “Common names”**

Fill the species common name(s) used in this geographical region.

Ex: Tuna, atún.

Label “Biology”**Topic “Sexuality”**

Fill the sexual type of the species in this geographical region by choosing from gonochoric, protogynous hermaphrodite, protandrous hermaphrodite, or potentially hermaphrodite according to relevant regional bibliography.

Ex: Protogynous hermaphrodite (Lozano, 2003).

Topic “Reproduction size”

Length at first maturity (in cm) in this geographical region.

Ex: 33 cm SL.

Ex: 17 cm FL (females), 23 cm FL (males) (González, 2004).

Topic “Spawning season”

Fill the reproduction period (in a month range) of this species in this geographical region.

Ex: April-September (Jiménez, 2001).

Topic “Food”

Fill the feeding habits and main preys-categories for the species in this geographical region.

Ex: Carnivorous: crustaceans, cephalopods and fishes.

Label “Habitat”

Topic “Depth range”

Fill in the known minimum and maximum depth for the species (given as a range in m) in this geographical region.

Ex: 100-400 m.

Label “Size characteristics”

Topic “Common sizes”

Fill the commercial sizes regularly present in the markets (given as a range in cm) in this geographical region.

Ex: 20-40 cm TL.

Topic “Capture legal size”

Fill the minimum size of capture in according to regional regulations (given in cm).

Ex: 20 cm TL.

Label “Interest to fisheries”

Topic “Fisheries importance”

Fill in the species importance for the regional fisheries, aquaculture and game fishing in this geographical region.

Ex: Highly commercial. Industrial and artisanal fisheries.

Topic “Fishing gear”

Fill in the fishing methods used for the species in this geographical region.

Ex: Bottom trawl and trammelnets.

Label “Commercialisation”**Topic “Forms of utilization”**

Fill in the forms of use (whole or transformed product) in terms of human consumption for the species in this geographical region.

Ex: Marketed fresh, refrigerated, frozen and canned. Generally used in fillets or slices; eggs are also commercialised.

Topic “Cooking usage”

Fill in the culinary usages for the species in this geographical region.

Ex: Boiled, fried, baked, broiled.

Topic “Misidentifications”

Fill in the most common market misidentifications for the species in this geographical region.

Ex: It can be misidentified with other tuna fishes.

Topic “Substitutions”

Fill in the known market substitutions for the species in this geographical region.

Ex: Fillets of pelagic sharks are sometimes sold as if they belong to this species”.

Label “Conservation status and Threats”**Topic “Conservation status and threats”**

Fill in the conservation status and degree of threat in this geographical region.

Ex: Vulnerable; protected species according to regional regulations. Main threats are: illegal spearing fishery and gillnets.

Label “Other”**Topic “Other”**

Fill in other relevant information for the species in this geographical region.

Ex: Ongoing aquaculture experiences in floating cages.

How we must validate all regional information

By using the appropriate information from regional sources.

Each sampling/taxonomic group will be responsible for checking the regional data entered.

Please, bear in mind the importance of the accuracy of this regional information in order to use these data for several common fields.

4.3 REMARKS ON REFERENCE PICTURES –VOUCHER AND OTOLITH– IN THE DATABASE SPECIES TABLE (TASKS TO BE DONE FOR EACH ASSIGNED SPECIES)

Aim

Select the reference pictures (voucher and otolith) from each species.

Please, remember that the size of the pictures was established in 1024x768 pixels within the WP2 protocol.

Up to date, nothing has been established in relation to the picture resolution and scale bar addition. We are in favor of using a scale bar in all voucher specimens' photos; the consensus picture for the species can then be without scale bar.

How we must validate pictures

Step 1: Check the specimens' identity by comparing all available photos in the database among them.

Step 2: When any specific images differs from the rest of the specimens, please contact with the partner(s) involved in this sampling task in order to clarify this discrepancy.

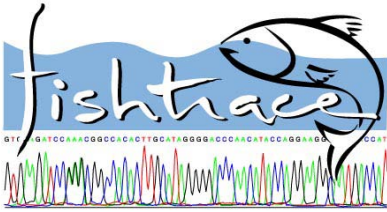
Step 3: Select quality representative images as consensus reference pictures for the species and the specific otolith.

Step 4: Other relevant photos for identification purposes could be selected by the partner/group responsible for the validation on this species. Examples: left side in soles (Soleidae), ventral side in anglerfishes (Lophiidae), sexual chromatic dimorphism in many species (Scaridae, Labridae, etc.), among others.

Step 5: The remaining specific pictures uploaded will also be maintained in the database and made accessible in an appropriate level.

WE ARE PROPOSING HERE THE ADAPTATION, IF NECESSARY, OF THE DATABASE STRUCTURE FOR THE PURPOSES INDICATED IN STEPS 4 AND 5, BEARING IN MIND THE WEB APPEARANCE AND THE BUS SPEED.

Annex XVII: Concise Protocol for the online validation process.



FishTrace WP7 RESUMED PROTOCOL FOR THE ONLINE VALIDATION PROCESS

STEPS TO BE DONE BY EACH PARTNER AT EACH AREA (INTERNAL VALIDATION OR SELF-VALIDATION)

STEP 1.- Enter the specimens's data (both text and pictures) for a particular species from the sampling forms.

STEP 2.- Check carefully the specimens's data: self-validation. Every specimen table must be signed and dated by the corresponding responsible for validation purpose at each Institution.

STEP 3.-

- 1) Visit www.fishtrace.COM
- 2) Click on: Link to project page
- 3) Click on: Database
- 4) Click on: Connection to loading site
- 5) Click on: Data validation flow document:

-Username: fishtrace
-Password: ***** (everybody knows!)

6)

6.1.) Open the "List of species for validation" file corresponding to your group and CHECK the list of validated species (Tax-Coll or Gen)

-Username: fishtrace
-Password: ***** (everybody knows!)

6.2.) Open the validation file corresponding to your group (e.g. "Control Data Validation Flow_UCM_26-10-2005.xls") and **UPDATE** it.

Don't forget to **re-upload** the document after filling! (SAVE AS...)
(e.g. "Control Data Validation Flow_UCM_19-04-2006.xls")

A.- The taxonomist/curator responsible of each partner will put an **X** on the **-2** given by default in the Control Data Validation Flow (excel file). We will assume that the DNA collection field will be empty at that moment.

B.- The geneticist responsible of each partner will put an **X** on the **-2** given by default in the Control Data Validation Flow (excel file).

STEP 4.- When necessary/available data, enter the regional information for this particular species.

STEPS TO BE DONE BY THE PARTNER/GROUP IN CHARGE OF A PARTICULAR SPECIES (TRUE VALIDATION PROCESS)

STEP 5.- In case of assigned species, enter the information for the common fields in the species table.

STEP 6.- Check the Control Data Validation Flow (excel file). When all geographical crosses (**X**) are ok, proceed through the next step.

STEP 7.- Check all specimen's information from the involved areas by comparing with the specific common fields (in the species table). When a problem is detected, a feedback information flow must be started with the responsible for the involved data.

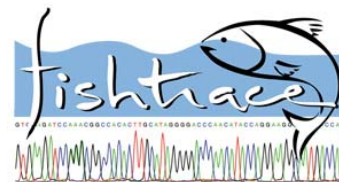
STEP 8.-

A.- The taxonomist/curator responsible for this species will put a DEFINITIVE **X** on the **-2** given by default in the Control Data Validation Flow (excel file): just on the general TAXONOMY and COLLECTIONS columns.

B.- The Expert Group on Genetics will put a DEFINITIVE **X** on the **-2** given by default in the Control Data Validation Flow (excel file): just on the general GENETICS column.

When a problem is detected, the responsible for genetics must start a feedback information flow with the taxonomists involved.

Annex XVIII: The Bibliography Module in the FishTrace Database.



The Bibliography Module in the FishTrace Database

Prepared by: S. Jiménez (TFMC), M. Biscoito (IMAR-MMF) & J.A. González (ICCM)

1 – Bibliography module description

This module should have the following characteristics:

- 1) data providers have predefined workforms (5), each one containing the necessary fields for the type of document that will be recorded;
- 2) the database will automatically assign a sequential reference number to each record;
- 3) the database is searchable by several fields;
- 4) references in output lists are formatted;
- 5) output lists are printable

Once in the bibliography main page, the data provider should be asked to choose the type of document to record (journal article, book, book chapter, webpage, CD-ROM, etc.) from a pop-down menu.

Then a specific workform would be shown on screen, with only those fields which are appropriate for the document in question.

After having submitted the record, the database should ask if another record of the same type would be submitted or not. In the first case (yes), a new blank workform will appear and the process can go on. If the second option is chosen (no), then the bibliography main page will appear on screen, granting the provider with the choice for another type of document, to edit a previously submitted record, to search the database or to leave the section.

As above mentioned, the module should have an editing tool for corrections and additions to previously entered records, accessed from the main page.

Also a search tool should be available, allowing searches by author, title, region, species and provider.

Each search results should be printable and references should be formatted according to a standard defined below.

2 – Bibliography module structure

List of fields:

- (1) Author(s) (Analytic) [authors separated by //]
- (2) Connective phrase
- (3) Author(s) (Monographic) [authors separated by //]
- (4) Author(s) role
- (5) Title (article/chapter)
- (6) Journal title
- (7) Title (Monographic)
- (8) Year of publication
- (9) Volume/Report ID
- (10) Issue ID
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (14) Edition
- (15) URL
- (16) Version [for e-books, CD ROMs and webpages]
- (17) Series title
- (18) Series volume
- (19) Series issue

- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [0001 to 9999, automatically assigned, 4 characters]

List of workforms:

Workform no. 1 – Journal article

- (1) Author(s) (Analytic) [authors separated by //]
- (5) Title (article/chapter)
- (6) Journal title
- (8) Year of publication
- (9) Volume ID
- (10) Issue ID
- (11) Page(s)
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 2 – Book

- (3) Author(s) (Monographic) [authors separated by //]
- (7) Title (Monographic)
- (8) Year of publication
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (14) Edition
- (20) Geographic code [alphanum., codes separated by /]

- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 3 – Book chapter

- (1) Author(s) (Analytic) [authors separated by //]
- (2) Connective phrase
- (3) Author(s) (Monographic) [authors separated by //]
- (4) Author(s) role
- (5) Title (article/chapter)
- (7) Title (Monographic)
- (8) Year of publication
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name
- (14) Edition
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 4 – Report

- (1) Author(s) (Analytic) [authors separated by //]
- (5) Title (article/chapter)
- (8) Year of publication
- (9) Volume/Report ID
- (11) Page(s)
- (12) Place of publication
- (13) Publisher/Editor name

- (17) Series title
- (18) Series volume
- (19) Series issue
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 5 – Web Page

- (1) Author(s) (Analytic) [authors separated by //]
- (5) Title (article/chapter)
- (8) Date of publication/access
- (9) Volume/Report ID
- (13) Publisher/Editor name
- (15) URL
- (16) Version [for e-books and webpages]
- (20) Geographic code [alphanum., codes separated by /]
- (21) Provider code [alphanumeric, 7 characters]
- (22) Species code [alphanumeric, codes separated by /]
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Workform no. 6 – CD ROM

- (1) Author(s)
- (5) Title
- (8) Year of publication
- (9) Volume/Report ID
- (12) Place of publication
- (13) Publisher/Editor name
- (16) Version

- (17) Series title
- (20) Geographic code
- (21) Provider code
- (22) Species code
- (23) Record No. [0001 to 9999, automatically assigned, 4 characters]

3 – Bibliography module examples

Filled workform

Workform no. 1 – Journal article

- (1) Author(s) [González, J.A.//V. Rico//J.I. Santana](#)
- (5) Title [Capture of Grammicolepis brachiusculus Poey, 1873 \(Grammicolepididae\) off the Canary Islands](#)
- (6) Journal title [Scientia Marina](#)
- (8) Year of publication [2000](#)
- (9) Volume ID [64](#)
- (10) Issue ID [1](#)
- (11) Page(s) [107-109](#)
- (20) Geographic code [CI](#)
- (21) Provider code [MMF001](#)
- (22) Species code [GraBra](#)
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output:

[González, J.A., V. Rico, & J.I. Santana \(2000\). Capture of Grammicolepis brachiusculus Poey, 1873 \(Grammicolepididae\) off the Canary Islands. Scientia Marina, 64\(1\): 107-109.](#)

Filled workform

Workform no. 2 – Book

- (3) Author(s) [Brito, A.](#)
- (7) Title [Catálogo de los Peces de las Islas Canarias](#)
- (8) Year of publication [1991](#)
- (11) Page(s) [230 pp.](#)
- (12) Place of publication [La Laguna](#)
- (13) Publisher/Editor name [Francisco Lemus, Ed.](#)
- (14) Edition [1a.](#)
- (20) Geographic code [CI](#)
- (21) Provider code [MMF001](#)
- (22) Species code [CenCir/KypSec/BodScr/SpoCan/SpaAur](#)
- (23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output:

[Brito, A. \(1991\). Catálogo de los Peces de las Islas Canarias. 230 pp. La Laguna. Francisco Lemus, Ed.](#)

Workform no. 3 – Book chapter

- (1) Author(s) [Robins, C.H.//C.R. Robins](#)
- (2) Connective phrase [In](#)
- (3) Author(s) [Böhlke, E.B.//J.E. Böhlke//M.M. Leiby//J.E. McCosker//E. Bertelsen//C.H. Robins//C.R. Robins//D.G. Smith//K.A. Tighe//J.G. Nielsen](#)
- (4) Author(s) role [\(Eds.\)](#)
- (5) Title (article/chapter) [Family Synphobranchidae](#)
- (7) Title (Monographic) [Fishes of the Western North Atlantic, Part VI](#)
- (8) Year of publication [1989](#)
- (11) Page(s) [207-253](#)

- (12) Place of publication [New Haven](#)
- (13) Publisher/Editor name [Sears Foundation for Marine Research, Yale University](#)
- (14) Edition -
- (20) Geographic code [CI/MAD/WM/CS](#)
- (21) Provider code [MMF001](#)
- (22) Species code [MurHel/MurAug/GymUni](#)
- (23) Record No. [1 to infinite, automatically assigned, 4 character]

Output:

[Robins, C.H. & C.R. Robins \(1989\). Family Synphobranchidae. In Fishes of the Western North Atlantic, Part VI: 207-253. Böhlke, E.B., J.E. Böhlke, M.M. Leiby, J.E. McCosker, E. Bertelsen, C.H. Robins, C.R. Robins, D.G. Smith, K.A. Tighe & J.G. Nielsen \(Eds.\). New Haven: Sears Foundation for Marine Research, Yale University.](#)

Workform no. 4 – Report

- (1) Author(s) [Santana, J.I.//J.A. González//J.A. Quiles//M.F. Marrero//I.J. Lozano//S. Jiménez](#)
- (5) Title [Prospección con nasas camaroneras tradicionales en la zona exterior de la Reserva Marina de La Graciosa \(Lanzarote\)](#)
- (8) Year of publication [2001](#)
- (11) Page(s) [75 pp.](#)
- (12) Place of publication [Telde \(Gran Canaria\)](#)
- (13) Publisher/Editor name [Instituto Canario de Ciencias Marinas, Gobierno de Canarias](#)
- (20) Geographic code [CI](#)
- (21) Provider code [TFMC001](#)

(22) Species code [SynKau/BraBra](#)

(23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output: [Santana, J.I., J.A. González, J.A. Quiles, M.F. Marrero, I.J. Lozano & S. Jiménez \(2001\). Prospección con nasas camaroneras tradicionales en la zona exterior de la Reserva Marina de La Graciosa \(Lanzarote\). 75 pp. Instituto Canario de Ciencias Marinas, Gobierno de Canarias. Telde \(Gran Canaria\).](#)

Workform no. 5 – Web Page

(1) Author(s) [Froese, R./D. Pauly, Eds.](#)

(5) Title (article/chapter) [FishBase. A Global Information System on Fishes](#)

(8) Date of publication/access [March, 2005](#)

(9) Volume/Report ID [World Wide Web electronic publication](#)

(13) Publisher/Editor name [Froese, R. & D. Pauly, Eds.](#)

(15) URL www.fishbase.org

(16) Version [12/2004](#)

(20) Geographic code [All](#)

(21) Provider code [MMF001](#)

(22) Species code [All](#)

(23) Record No. [1 to infinite, automatically assigned, 4 characters]

Output:

[Froese, R. & D. Pauly, Eds. \(2004\). FishBase. A Global Information System on Fishes. World Wide Web electronic publication. URL: www.fishbase.org, version \(12/2004\). Access, March, 2005.](#)

Filled workform

Workform no. 6 – CD ROM

(1) Author(s) [González, J.A.//J.I. Santana//J.A. Quiles//S. Jiménez//F. Hernández//A. Diez//J.M. Bautista](#)

(5) Title [Identificación y diferenciación de teleósteos marinos en el área de influencia económica de Canarias 1FD97-1235-C04 MAR, FEDER-MCyT](#)

(8) Year of publication [2002](#)

(9) Volume/Report ID [CD ROM Multimedia](#)

(12) Place of publication [Madrid](#)

(13) Publisher/Editor name [Instituto Canario de Ciencias Marinas, Museo de Ciencias Naturales de Tenerife y Universidad Complutense de Madrid](#)

(16) Version [2002 \(50 especies\)](#)

(17) Series title [PesCanarias](#)

(20) Geographic code [CI](#)

(21) Provider code [ICCM001](#)

(22) Species code [BooBoo/MurHel/PagPag/HeIDac](#)

(23) Record No. [0001 to 9999, automatically assigned, 4 characters]

Output:

[González, J.A., J.I. Santana, J.A. Quiles, S. Jiménez, F. Hernández, A. Diez & J.M. Bautista \(2002\). Identificación y diferenciación de teleósteos marinos en el área de influencia económica de Canarias 1FD97-1235-C04 MAR, FEDER-MCyT. CD ROM Multimedia. PesCanarias. Version 2002 \(50 especies\).](#)

Annex XIX: Results from the genetic population structure analysis.

**Genetic Population Structure Analysis for *Pagrus pagrus*, *Solea solea*, *Micromesistius poutassou*, *Merluccius merluccius*, *Mullus surmuletus*, and *Pagellus erythrinus*.
(FishTrace WP4)**

Analysis of molecular variance (AMOVA) for population structure based on cytochrome *b* nucleotide sequence data was performed with the ARLEQUIN v. 2.000 software (Schneider, Roessli and Excoffier, 2000, <http://anthro.unige.ch/arlequin>). The methods and the parameters used in the analysis are as indicated in **Appendices 1-6**.

Important things to note:

1. *FST P* values >0.05 indicate significant genetic differences between populations. In the Population pairwise *FST P* matrices this is indicated by a (+).
2. This analysis is still preliminary. Further genetic tests and corrections can be made to finely define population structures when all samples sequences are available.

Population (sample) names:

1. BS = Baltic Sea
2. NS = North Sea
3. CB = Bay of Biscay and English Channel
4. CS = Cantabric Sea
5. MA = Madeiran Archipelago
6. CI = Canary Islands
7. WM = Western Mediterranean
8. EM = Eastern Mediterranean

Results

1. *Solea solea*

According to the AMOVA results, this species exhibits a high percentage of variation among populations (55.2%) surpassing the variation within populations (for details of the analyses performed, see **Appendix 1**). In agreement with that, *FST P* values indicate significant differences between all samples (populations) examined. The only exception is that concerning the CB and NS samples where a significant number of individuals (6 from CB and 5 from NS) share a common haplotype. This suggests potentially important gene flow between these two proximate geographical areas. However, it is of interest to note that the NS sample contains only a four different haplotypes. This contrasts to the eight haplotypes observed in the CB and BS samples, the 11 of the WM sample, and the 13 of the EM sample. This low level of genetic variation is indicative of a homogeneous population and therefore implies a possible genetic bottleneck. Additional sampling from this area may be necessary in order to conclude on this issue.

The genetic relationship/distance of the different samples examined is depicted in Figure 1. The almost complete spatial segregation of distinct genetic populations is expected to be in agreement with the life histories of this species, at least as far as the potential mechanisms contributing to gene flow (or absence of) are concerned (benthic vs. pelagic larval stages, reproduction, feeding habits, physical barriers to migration). Therefore this analysis could profit from data concerning the biology of the species. Fisheries data could likewise contribute significantly to evaluating the genetic variation that has been observed in this study. This is a highly commercial species and, therefore, may be under intense fishing pressure in several European areas. Sampling and genetic analyses from additional European and extra-European areas could be required to further test the above hypotheses.

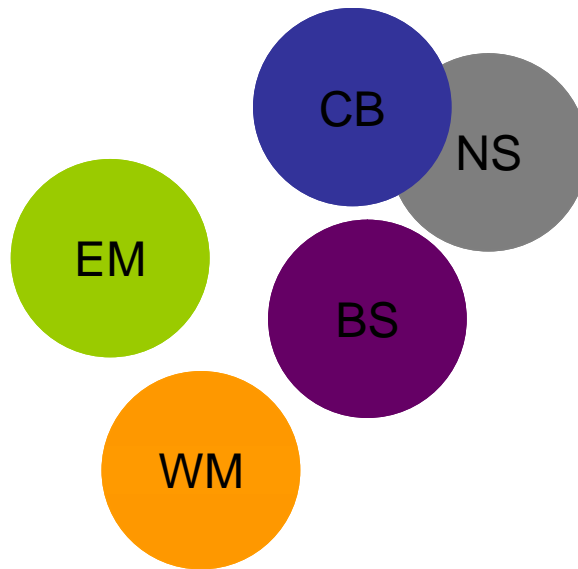


Figure 1. Genetic relations between distinct populations of *Solea solea*. Each circle represents distinct samples as indicated (n=14, 22, 20, 15, 20 for CB, EM, BS, WM, and NS, respectively). Overlapping circles indicate non-significant differences in genetic population structure, as resulted from population pairwise F_{ST} . This two-dimensional model attempts also to consider and to illustrate the relative genetic distances between populations.

2. *Meluccius merluccius*

Detailed results of the analysis of the different samples of *Meluccius merluccius* are given in **Appendix 2** and are summarized in Fig. 2. According to this analysis, the samples can be divided in two groups: those that are genetically related to the CS sample, i.e. CI, CB, BS, and WM, and those that are not, i.e. EM and NS. Furthermore, the NS sample is genetically distinct from all other samples while the EM sample is not significantly different from that of area CB ($P=0.295$).

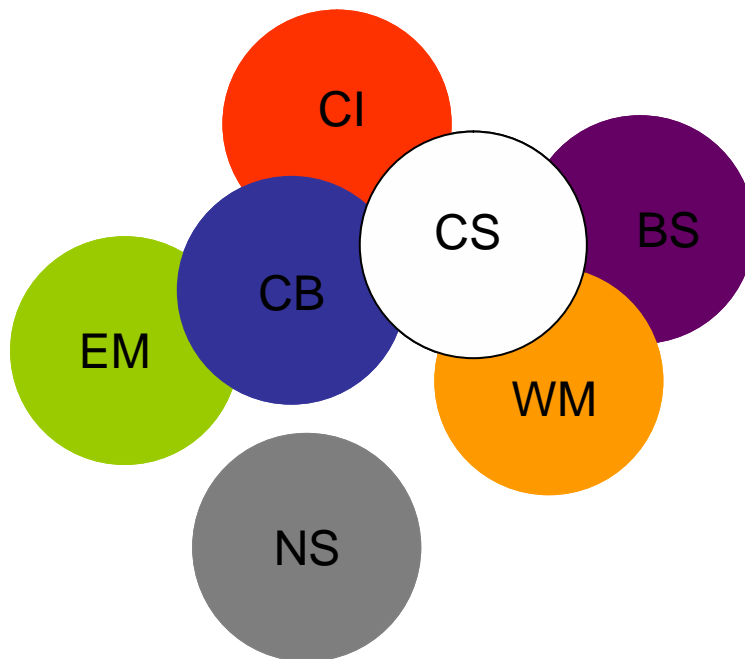


Figure 2. Genetic relations between distinct populations of *Meluccius merluccius*. Each circle represents distinct samples as indicated (n=18, 20, 18, 19, 19, 20,18 for EM, WM, CB, CI, CS, BS, and NS, respectively). Overlapping circles indicate non-significant differences in genetic population

structure, as resulted from population pairwise F_{ST} . This two-dimensional model attempts also to consider and to illustrate the relative genetic distances between populations.

This particular result is both interesting and intriguing. It could possibly be explained by assuming that a single ancient population of the species, related to the present population of CB, colonized the north-eastern part of the Mediterranean, as inferred by the presence of common haplotypes in both the CB and EM samples. More recent arrivals and expansions of *Meluccius merluccius* populations in the central Atlantic area (e.g. a CS-related population), in addition to selection pressure, may have resulted in the loss of the CB haplotypes from the populations of both the BS and WM areas and possibly in other Mediterranean and Atlantic areas that were not considered in the present study. This is consistent with 1) the presence of shared haplotypes between the CS, WM, and BS areas, 2) the presence of common haplotypes between the CS, CI, and CB areas and, 3) the absence of CI or CB haplotypes in either the BS or WM areas. According to this hypothesis CS haplotypes should be present in the southern part of the Iberian Peninsula, on both sides of the Gibraltar straights, and possibly also in the central Mediterranean area. Therefore, in order to test this hypothesis additional samples need to be obtained and analyzed from the above mentioned areas. Likewise the analysis will profit from data concerning the spawning behaviour, self-recruitment, and mechanisms of near-shore retention of larvae in this species. Furthermore, environmental factors, including past sea level changes, and present or past physical barriers such as ocean currents, which may have mixed or disrupted the populations of *Meluccius merluccius* from different geographic locations need also to be considered.

Finally, the genetic heterogeneity of the NS sample from all of the other samples examined suggests that the *Meluccius merluccius* population in this area is genetically isolated and/or is influenced by other populations of the species such as those of the North Atlantic. Additional samples from Norway, Iceland and even Canada could allow conclusions concerning this particular population.

3. *Micromesistius poutassou*

According to the analyses presented in **Appendix 3** and summarized in Figure 3, this species, as was also the case for *Meluccius merluccius* (also of the Gadidae family) exhibits an interesting genetic population structure for the areas from which samples were analyzed. However, in contrast to the significant genetic distance between the EM and CS populations of *Meluccius merluccius*, this analysis demonstrates a genetic relationship between the above two populations of absence of *Micromesistius poutassou*. Furthermore, neither of the WM or NS populations seems to be related to either of the CS, CB, or EM populations. However, the relatively low variation among populations of the species (12.68% vs. 15.39% in *Meluccius merluccius* and 55.2% in *Solea solea*) must be noted.

Concerning the NS population, the hypothesis put forth for *Meluccius merluccius* above may also be valid for this species as well. However, it is notable that the NS population includes only five different haplotypes compared to 18 in the EM sample, and 15, 14, and 14 in the WM, CB and CS samples, respectively. A similar genetic homogeneity was observed for *Solea solea* in the NS area suggesting a possible genetic bottleneck.

For the WM population it must be assumed that it has been genetically influenced by a more recent invasion of the Mediterranean basin with population(s) of the species originating possibly from the Atlantic coast of Morocco. This hypothesis could be tested by the analysis of additional samples from Morocco as well as from the North African coast within the Mediterranean. Factors that could explain the population structuring of this gregarious and erratic species and which, therefore, must be considered concern the life history of the species and environmental influences as these have been listed for the other species above.

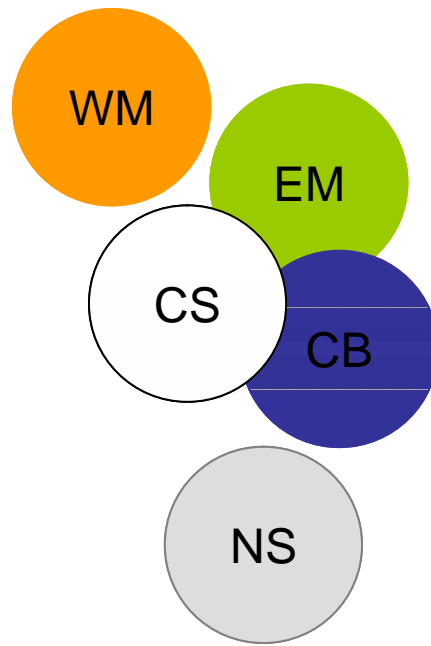


Figure 3. Genetic relations between distinct populations of *Micromesistius poutassou*. Each circle represents distinct samples as indicated (n=20, 20, 18, 20, 16 for EM, CB, WM, NS, and CS, respectively). Overlapping circles indicate non-significant differences in genetic population structure, as resulted from population pairwise F_{ST} . This two-dimensional model attempts also to consider and to illustrate the relative genetic distances between populations.

4. *Mullus surmuletus*

Detailed analyses for the structuring of the genetic populations of *Mullus surmuletus* are given in **Appendix 4**. As is also illustrated in Figure 4, all the samples examined, with the exception of NS, share a significant number of common haplotypes and thus, structuring of populations from the Atlantic (the Canary Islands) to the North-eastern Mediterranean is not evident. The MA sample appears to be significantly different from both the WM and CB ones. However, this sample contains only eight individuals. Increasing the number of sampled individuals from the MA area is likely to reveal haplotypes common to the WM and CB areas. It should also be noted that this analysis has not included the CS sample, which by virtue of its central geographic location could further support the panmictic nature of the Atlantic-Mediterranean populations.

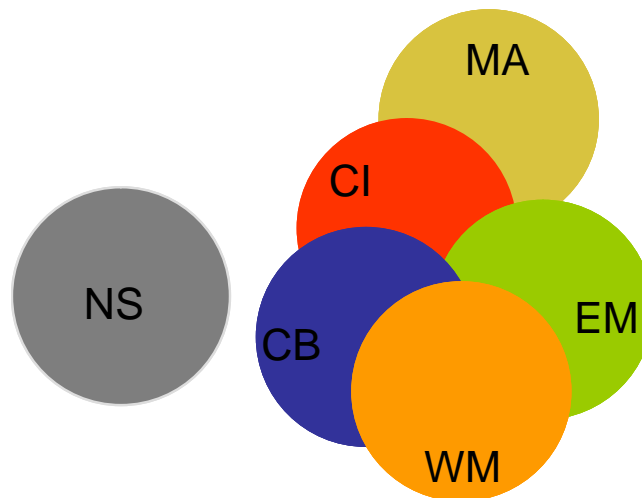


Figure 4. Genetic relations between distinct populations of *Mullus surmuletus*. Each circle represents distinct samples as indicated (n=20, 20, 19, 18, 16, 8 for EM, NS, CB, WM, CI and MA, respectively). Overlapping circles indicate non-significant differences in genetic population

structure, as resulted from population pairwise F_{ST} . This two-dimensional model attempts also to consider and to illustrate the relative genetic distances between populations.

The genetic distance of the NS population from the other samples examined suggests the existence of at least one additional genetic stock at the northern limit of repartition of this species.

5. *Pagrus pagrus*

According to the results presented in **Appendix 5** and summarized in Figure 5, the distribution of the populations of this species and consequently their structure appears sensitive to physical barriers. Thus, there is a clear separation between EM and MA and each one of these populations is different from those of CI and WM, which share a significant number of common haplotypes. The above suggests that gene flow between the CI and WM areas follows the West African coastline through the Gibraltar straights into the Mediterranean. This could be further supported by additional samples from the Moroccan coast and from both sides of the straights. Isolation of the MA population is possibly due to prohibitive depths for the biology of this species, separating the Canaries Islands from Madeira.

A possible barrier within the Mediterranean separating the WM and EM populations could be the Sicelo-Tunisian straight. However, the analysis of additional Mediterranean samples is necessary in order to demonstrate this separation.

Note that the CS sample, which could provide important information concerning the genetic relationship of the Atlantic and Mediterranean populations, is missing from this analysis.

Furthermore, the EM sample contains only 12 individuals a fact that could bias the results of the analysis.

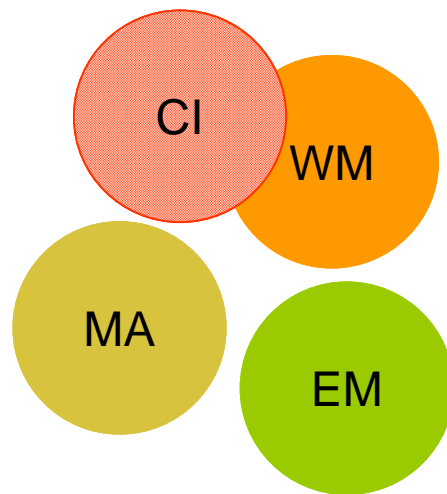


Figure 5. Genetic relations between distinct populations of *Pagrus pagrus*. Each circle represents distinct samples as indicated (n=20, 20, 20, 12 for CI, WM, MA, and EM, respectively).

Overlapping circles indicate non-significant differences in genetic population structure, as resulted from population pairwise F_{ST} . This two-dimensional model attempts also to consider and to illustrate the relative genetic distances between populations.

6. *Pagellus erythrinus*

With only three samples of *Pagellus erythrinus* analyzed (from areas CI, WM and EM, see **Appendix 6** for details), no obvious structuring of populations was observed. In fact for this species the variation within populations was at 100% with essentially each individual representing a distinct haplotype (17 haplotypes in 18 individuals in the EM sample, 18 haplotypes in 19 individuals in the CI sample, and 15 haplotypes in 17 individuals in the WM sample). This situation is indicative of a panmictic population in the areas examined due to high dispersal rates and large effective

population size. Further sampling and analysis from areas MA and CS could demonstrate the existence or absence of a genetically distinct population(s) of the species in the Atlantic.

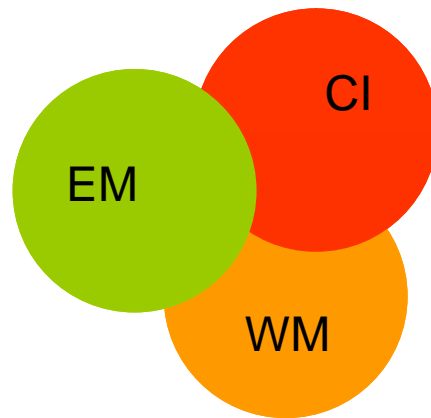


Figure 6. Genetic relations between distinct populations of *Pagellus erythrinus*. Each circle represents distinct samples as indicated (n=19, 18, 17 for CI, EM, and WM, respectively). Overlapping circles indicate non-significant differences in genetic population structure, as resulted from population pairwise F_{ST} .

APPENDIX 1

***Solea solea* Genetic Population Structure Analyses**

SolSol

[[Structure]]

```
StructureName = "Solea solea"  
NbGroups = 1  
IndividualLevel = 0  
DistMatLabel = ""  
Group={  
  "CB" 14  
  "EM" 22  
  "BS" 20  
  "WM" 15  
  "NS" 20  
}
```

Distance method: Tamura & Nei

AMOVA design and results :

Reference: Weir, B.S. and Cockerham, C.C. 1984.
Excoffier, L., Smouse, P., and Quattro, J. 1992.
Weir, B. S., 1996.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	4	133.424	1.76701 Va	55.22
Within populations	86	123.256	1.43321 Vb	44.78
Total	90	256.679	3.20022	
Fixation Index	FST :	0.55215		

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
P(rand. value = obs. value) = 0.00000
P(rand. value >= obs. value) = 0.00000+-0.00000

=====
 == Comparisons of pairs of population samples
 =====

List of labels for population samples used below:

Label	Population name
1:	CB
2:	EM
3:	BS
4:	WM
5:	NS

 Population pairwise FSTs

Distance method: Tamura & Nei

	1	2	3	4	5
1	0.00000				
2	0.71344	0.00000			
3	0.12897	0.58492	0.00000		
4	0.63655	0.27857	0.47664	0.00000	
5	0.04121	0.78661	0.20876	0.73857	0.00000

 FST P values

Number of permutations : 1023

	1	2	3	4	5
1	*				
2	0.00000+-0.0000	*			
3	0.02051+-0.0036	0.00000+-0.0000	*		
4	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*	
5	0.12891+-0.0090	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*

 Matrix of significant Fst P values
 Significance Level=0.0500

Number of permutations : 1023

	1	2	3	4	5
1		+	+	+	-
2	+		+	+	+
3	+	+		+	+
4	+	+	+		+
5	-	+	+	+	

 AMOVA design and results (average over 48 loci):

Source of variation	Sum of squares	Variance components
Among populations	160.523	2.14500
Within populations	118.609	1.37918
Total	279.132	3.52418

Average F-Statistics over all loci
 Fixation Indices
 FST : 0.60865

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.00000

Table of the Haplotype distribution through the Populations:

Haplotype:	CB (14)	EM (22)	BS (20)	WM (15)	NS (20)
Sol_01_MER	1	0	0	0	0
Sol_03_CB					
Sol_05_CB					
Sol_12_CB					
Sol_15_CB					
Sol_16_CB					
Sol_18_CB					
Sol_01_NRM					
Sol_03_NRM					
Sol_NS_01.					
Sol_NS_06.					
Sol_NS_09.					
Sol_NS_15.					
Sol_NS_20.	6	0	2	0	5
Sol_04_CB					
Sol_NS_03.					
Sol_NS_05.					
Sol_NS_08.					
Sol_NS_10.					
Sol_NS_12.					
Sol_NS_14.					
Sol_NS_18.	1	0	0	0	7
Sol_06_CB	1	0	0	0	0
Sol_07_CB	1	0	0	0	0
Sol_08_CB	1	0	0	0	0
Sol_10_CB	1	0	0	0	0
Sol_13_CB					
Sol_18_CB					
Sol_19_NRM					
Sol_21_NRM	1	0	3	0	0
Sol_14_CB					
Sol_NS_02.					
Sol_NS_11.					
Sol_NS_13.					
Sol_NS_16.	1	0	0	0	4
Sol_01_FRI					
Sol_02_FRI					
Sol_04_FRI					
sol_06_FRI					
sol_10_FRI					
sol_18_FRI					
sol_20_FRI					
Sol_12_FRI					
Sol_15_FRI					
Sol_21_FRI					
Sol_22_FRI	0	11	0	0	0
Sol_03_FRI	0	1	0	0	0
Sol_05_FRI	0	1	0	0	0
Sol_11_FRI	0	1	0	0	0
Sol_13_FRI	0	1	0	0	0
Sol_14_FRI	0	1	0	0	0
sol_07_FRI	0	1	0	0	0
sol_08_FRI	0	1	0	0	0
sol_09_FRI	0	1	0	0	0
sol_16_FRI	0	1	0	0	0
sol_17_FRI	0	1	0	0	0
Sol_19_FRI	0	1	0	0	0
Sol_02_NRM	0	0	1	0	0

Sol_04_NRM	0	0	1	0	0
Sol_05_NRM					
Sol_06_NRM					
Sol_10_NRM					
Sol_20_NRM					
Sol_12_NRM					
Sol_13_NRM					
Sol_14_NRM					
Sol_17_NRM	0	0	8	0	0
Sol_15_NRM					
Sol_11_NRM					
Sol_16_NRM	0	0	3	0	0
Sol_07_NRM	0	0	1	0	0
Sol_09_NRM	0	0	1	0	0
Sol_WM_01					
Sol_WM_12					
Sol_WM_13					
Sol_WM_18	0	0	0	4	0
Sol_WM_02	0	0	0	1	0
Sol_WM_06					
Sol_WM_20	0	0	0	2	0
Sol_WM_07	0	0	0	1	0
Sol_WM_10	0	0	0	1	0
Sol_WM_11	0	0	0	1	0
Sol_WM_14	0	0	0	1	0
Sol_WM_15	0	0	0	1	0
Sol_WM_16	0	0	0	1	0
Sol_WM_17	0	0	0	1	0
Sol_WM_19	0	0	0	1	0
Sol_NS_04.					
Sol_NS_07.					
Sol_NS_17.					
Sol_NS_19.	0	0	0	0	4

APPENDIX 2

***Meluccius merluccius* Genetic Population Structure Analyses**

MER MER

[[Structure]]

```
StructureName = "Merluccius"  
NbGroups = 1  
IndividualLevel = 0  
DistMatLabel = ""  
Group={  
  "EM" 18  
  "WM" 20  
  "CB" 18  
  "CI" 19  
  "CS" 19  
  "BS" 20  
  "NS" 18  
}
```

Distance method: Pairwise difference

AMOVA design and results :

Reference: Weir, B.S. and Cockerham, C.C. 1984.
Excoffier, L., Smouse, P., and Quattro, J. 1992.
Weir, B. S., 1996.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	6	120.753	0.82652 Va	15.39
Within populations	125	568.081	4.54465 Vb	84.61
Total	131	688.833	5.37117	
Fixation Index	FST :	0.15388		

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
P(rand. value = obs. value) = 0.00000
P(rand. value >= obs. value) = 0.00000+-0.00000

=====
 == Comparisons of pairs of population samples
 =====

List of labels for population samples used below:

Label	Population name
1:	EM
2:	WM
3:	CB
4:	CI
5:	CS
6:	BS
7:	NS

 Population pairwise FSTs

Distance method: Pairwise difference

	1	2	3	4	5	6	7
1	0.00000						
2	0.17080	0.00000					
3	0.00896	0.07007	0.00000				
4	0.07665	0.07532	0.02690	0.00000			
5	0.12944	-0.00595	0.03249	0.02983	0.00000		
6	0.21133	-0.01311	0.08788	0.08628	-0.01135	0.00000	
7	0.30213	0.33036	0.28000	0.24648	0.29916	0.35633	0.00000

 FST P values

Number of permutations : 1023

	1	2	3	4	5	6	7
1	*						
2	0.00000+-0.0000	*					
3	0.29492+-0.0160	0.01660+-0.0037	*				
4	0.01172+-0.0033	0.00391+-0.0019	0.13770+-0.0115	*			
5	0.00098+-0.0010	0.53223+-0.0153	0.11133+-0.0106	0.08496+-0.0101	*		
6	0.00000+-0.0000	0.66895+-0.0163	0.02051+-0.0041	0.00488+-0.0020	0.61426+-0.0158	*	
7	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*

 Matrix of significant Fst P values
 Significance Level=0.0500

Number of permutations : 1023

	1	2	3	4	5	6	7
1		+	-	+	+	+	+
2	+		+	+	-	-	+
3	-	+		-	-	+	+
4	+	+	-		-	+	+
5	+	-	-	-		-	+
6	+	-	+	+	-		+
7	+	+	+	+	+	+	

 AMOVA design and results (average over 72 loci):

Source of variation	Sum of squares	Variance components
Among populations	47.253	0.32366
Within populations	221.785	1.77428
Total	269.038	2.09793

Table of the Haplotype distribution through the Populations:

Haplotype:	EM (18)	WM (20)	CB (18)	CI (19)	CS (19)	BS (20)	NS (18)
MerM_EM_01							
MerM_EM_11							
MerM_CB_11							
MerM_CB_12							
MerM_CB_17							
MerM_CB_03							
MerM_17_CI							
MerM_BS_03	2	0	4	1	0	1	0
MerM_EM_02							
MerM_EM_04							
MerM_EM_06							
MerM_EM_12							
MerM_EM_19							
MerM_EM_20							
MerM_CB_18							
MerM_CB_02							
MerM_CB_06							
MerM_CB_08							
MerM_CS_12							
MerM_CS_15							
MerM_BS_10							
MerM_BS_15							
MerM_NS_02							
MerM_NS_04	6	0	4	0	2	2	2
MerM_EM_07	1	0	0	0	0	0	0
MerM_EM_08							
MerM_EM_03							
MerM_CB_01							
MerM_CB_07							
MerM_03_CI							
MerM_10_CI	2	0	2	2	0	0	0
MerM_EM_09	1	0	0	0	0	0	0
MerM_EM_10							
MerM_WM_01							
MerM_WM_03							
MerM_WM_05							
MerM_WM_06							
MerM_WM_08							
MerM_WM_09							
MerM_WM_20							
MerM_CB_13							
MerM_CB_19							
MerM_CB_04							
MerM_CB_10							
MerM_01_CI							
MerM_02_CI							
MerM_11_CI							
MerM_15_CI							
MerM_19_CI							
MerM_CS_01							
MerM_CS_02							
MerM_CS_10							
MerM_CS_13							
MerM_CS_16							
MerM_CS_17							
MerM_CS_18							
MerM_BS_01							
MerM_BS_11							
MerM_BS_16							
MerM_BS_17							
MerM_BS_18							
MerM_BS_02							
MerM_BS_20							
MerM_BS_05							
MerM_BS_07							
MerM_BS_08							
MerM_NS_01							
MerM_NS_03							
MerM_NS_05	1	7	4	5	7	10	3
MerM_EM_13	1	0	0	0	0	0	0
MerM_EM_14	1	0	0	0	0	0	0

MerM_EM_15							
MerM_06_CI	1	0	0	1	0	0	0
MerM_EM_16							
MerM_EM_17	2	0	0	0	0	0	0
MerM_WM_02							
MerM_WM_12	0	2	0	0	0	0	0
MerM_WM_04	0	1	0	0	0	0	0
MerM_WM_07	0	1	0	0	0	0	0
MerM_WM_10	0	1	0	0	0	0	0
MerM_WM_11	0	1	0	0	0	0	0
MerM_WM_13	0	1	0	0	0	0	0
MerM_WM_14							
MerM_WM_16							
MerM_CS_08	0	2	0	0	1	0	0
MerM_WM_15							
MerM_WM_17	0	2	0	0	0	0	0
MerM_WM_18	0	1	0	0	0	0	0
MerM_WM_19	0	1	0	0	0	0	0
MerM_CB_14	0	0	1	0	0	0	0
MerM_CB_15	0	0	1	0	0	0	0
MerM_CB_16							
MerM_BS_19	0	0	1	0	0	1	0
MerM_CB_05	0	0	1	0	0	0	0
MerM_04_CI							
MerM_CS_20	0	0	0	1	1	0	0
MerM_05_CI	0	0	0	1	0	0	0
MerM_08_CI							
MerM_20_CI	0	0	0	2	0	0	0
MerM_09_CI	0	0	0	1	0	0	0
MerM_12_CI	0	0	0	1	0	0	0
MerM_13_CI	0	0	0	1	0	0	0
MerM_14_CI	0	0	0	2	0	0	0
MerM_16_CI	0	0	0	1	0	0	0
MerM_18_CI	0	0	0	1	0	0	0
MerM_CS_03	0	0	0	0	1	0	0
MerM_CS_04							
MerM_CS_06	0	0	0	0	2	0	0
MerM_CS_07	0	0	0	0	1	0	0
MerM_CS_09	0	0	0	0	1	0	0
MerM_CS_11	0	0	0	0	1	0	0
MerM_CS_14	0	0	0	0	1	0	0
MerM_CS_19	0	0	0	0	1	0	0
MerM_BS_12	0	0	0	0	0	1	0
MerM_BS_13	0	0	0	0	0	1	0
MerM_BS_14	0	0	0	0	0	1	0
MerM_BS_04	0	0	0	0	0	1	0

MerM_BS_06	0	0	0	0	0	1	0
MerM_BS_09	0	0	0	0	0	1	0
MerM_NS_06	0	0	0	0	0	0	1
MerM_NS_07	0	0	0	0	0	0	1
MerM_NS_08	0	0	0	0	0	0	1
MerM_NS_09	0	0	0	0	0	0	1
MerM_NS_10							
MerM_NS_13	0	0	0	0	0	0	2
MerM_NS_11	0	0	0	0	0	0	1
MerM_NS_12	0	0	0	0	0	0	1
MerM_NS_14	0	0	0	0	0	0	1
MerM_NS_15	0	0	0	0	0	0	1
MerM_NS_16	0	0	0	0	0	0	1
MerM_NS_17	0	0	0	0	0	0	1
MerM_NS_18	0	0	0	0	0	0	1

APPENDIX 3

Micromesistius poutassou Genetic Population Structure Analyses

MIC POU

[[Structure]]

```
StructureName = "Mic Pou"  
NbGroups = 1  
IndividualLevel = 0  
DistMatLabel = ""  
Group={  
  "EM" 20  
  "CB" 20  
  "WM" 18  
  "NS" 20  
  "CS" 16  
}
```

Distance method: Tamura & Nei

AMOVA design and results :

Reference: Weir, B.S. and Cockerham, C.C. 1984.
Excoffier, L., Smouse, P., and Quattro, J. 1992.
Weir, B. S., 1996.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	4	40.659	0.39627 Va	12.68
Within populations	89	242.831	2.72844 Vb	87.32
Total	93	283.490	3.12471	
Fixation Index	FST :	0.12682		

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
P(rand. value = obs. value) = 0.00000
P(rand. value >= obs. value) = 0.00000+-0.00000

=====
 == Comparisons of pairs of population samples
 =====

List of labels for population samples used below:

Label	Population name
1:	EM
2:	CB
3:	WM
4:	NS
5:	CS

 Population pairwise FSTs

Distance method: Tamura & Nei

	1	2	3	4	5
1	0.00000				
2	-0.00471	0.00000			
3	0.12543	0.19428	0.00000		
4	0.12801	0.07442	0.38799	0.00000	
5	0.01539	0.00684	0.16068	0.12280	0.00000

 FST P values

Number of permutations : 1023

	1	2	3	4	5
1	*				
2	0.39844+-0.0123	*			
3	0.00195+-0.0014	0.00000+-0.0000	*		
4	0.00000+-0.0000	0.00195+-0.0014	0.00000+-0.0000	*	
5	0.20508+-0.0137	0.27734+-0.0159	0.00000+-0.0000	0.00000+-0.0000	*

 Matrix of significant Fst P values
 Significance Level=0.0500

Number of permutations : 1023

	1	2	3	4	5
1		-	+	+	-
2	-		+	+	-
3	+	+		+	+
4	+	+	+		+
5	-	-	+	+	

 AMOVA design and results (average over 56 loci):

Source of variation	Sum of squares	Variance components
Among populations	39.630	0.38484
Within populations	239.029	2.68572
Total	278.660	3.07056

Average F-Statistics over all loci

Fixation Indices
 FST : 0.12533

 Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.00000

Table of the Haplotype distribution through the Populations:

Haplotype:	EM (20)	CB (20)	WM (18)	NS (20)	CS (16)
Mic_01_EM					
Mic_NS_05					
Mic_NS_14					
Mic_NS_18	1	0	0	3	0
Mic_02_EM	1	0	0	0	0
Mic_03_EM	1	0	0	0	0
Mic_04_EM	1	0	0	0	0
Mic_08_EM					
Mic_11_EM					
Mic_20_EM					
Mic_02_CB					
Mic_10_CB					
Mic_15_CB					
Mic_CS_15	3	3	0	0	1
Mic_10_EM	1	0	0	0	0
Mic_13_EM	1	0	0	0	0
Mic_15_EM					
Mic_WM_09					
Mic_WM_10	1	0	2	0	0
Mic_18_EM	1	0	0	0	0
Mic_19_EM					
Mic_CS_06	1	0	0	0	1
Mic_21_EM					
Mic_09_CB	1	1	0	0	0
Mic_05_EM	1	0	0	0	0
Mic_06_EM					
Mic_12_CB	1	1	0	0	0
Mic_07_EM	1	0	0	0	0
Mic_09_EM	1	0	0	0	0
Mic_12_EM	1	0	0	0	0
Mic_16_EM	1	0	0	0	0
Mic_17_EM	1	0	0	0	0
Mic_01_CB	0	1	0	0	0
Mic_03_CB	0	1	0	0	0
Mic_04_CB	0	1	0	0	0
Mic_05_CB	0	1	0	0	0
Mic_06_CB					
Mic_20_CB					
Mic_NS_01					
Mic_NS_02					
Mic_NS_03					
Mic_NS_04					
Mic_NS_06					
Mic_NS_07					
Mic_NS_08					
Mic_NS_10					
Mic_NS_11					
Mic_NS_15					
Mic_NS_16					
Mic_NS_17					
Mic_NS_19					
Mic_NS_20	0	2	0	14	0
Mic_07_CB	0	1	0	0	0

Mic_08_CB	0	1	0	0	0
Mic_11_CB	0	1	0	0	0
Mic_13_CB					
Mic_14_CB					
Mic_17_CB	0	3	0	0	0
Mic_16_CB	0	1	0	0	0
Mic_18_CB	0	1	0	0	0
Mic_19_CB	0	1	0	0	0
Mic_WM_01	0	0	1	0	0
Mic_WM_02	0	0	1	0	0
Mic_WM_03					
Mic_WM_07	0	0	2	0	0
Mic_WM_04	0	0	1	0	0
Mic_WM_06	0	0	1	0	0
Mic_WM_08	0	0	1	0	0
Mic_WM_11	0	0	1	0	0
Mic_WM_12	0	0	1	0	0
Mic_WM_13	0	0	1	0	0
Mic_WM_14	0	0	1	0	0
Mic_WM_15	0	0	1	0	0
Mic_WM_17	0	0	1	0	0
Mic_WM_18	0	0	1	0	0
Mic_WM_19					
Mic_CS_10	0	0	1	0	1
Mic_WM_20	0	0	1	0	0
Mic_NS_09	0	0	0	1	0
Mic_NS_12	0	0	0	1	0
Mic_NS_13	0	0	0	1	0
Mic_CS_01					
Mic_CS_11	0	0	0	0	2
Mic_CS_02	0	0	0	0	1
Mic_CS_07	0	0	0	0	1
Mic_CS_08	0	0	0	0	1
Mic_CS_09	0	0	0	0	1
Mic_CS_12	0	0	0	0	1
Mic_CS_13	0	0	0	0	1
Mic_CS_14	0	0	0	0	1
Mic_CS_16	0	0	0	0	1
Mic_CS_18	0	0	0	0	1
Mic_CS_19	0	0	0	0	1
Mic_CS_20	0	0	0	0	1

APPENDIX 4

***Mullus surmuletus* Genetic Population Structure Analyses**

[[Structure]]

```

StructureName = "Mullus surmuletus"
NbGroups = 1
IndividualLevel = 0
DistMatLabel = ""
Group={
  "MA" 8
  "CI" 16
  "CB" 19
  "EM" 20
  "WM" 18
  "NS" 20
}

```

Distance method: Tamura & Nei

AMOVA design and results :

Reference: Weir, B.S. and Cockerham, C.C. 1984.
 Excoffier, L., Smouse, P., and Quattro, J. 1992.
 Weir, B. S., 1996.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	5	34.101	0.27781 Va	11.21
Within populations	95	209.135	2.20142 Vb	88.79
Total	100	243.235	2.47922	
Fixation Index	FST :	0.11205		

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.00000+-0.00000

=====
 == Comparisons of pairs of population samples
 =====

List of labels for population samples used below:

Label	Population name
1:	MA
2:	CI
3:	CB
4:	EM
5:	WM
6:	NS

 Population pairwise FSTs

Distance method: Tamura & Nei

	1	2	3	4	5	6
1	0.00000					
2	0.00451	0.00000				
3	0.18558	0.07227	0.00000			
4	0.05504	-0.01764	0.03818	0.00000		
5	0.11250	0.01970	-0.01896	-0.00106	0.00000	
6	0.52410	0.35959	0.07924	0.26485	0.15286	0.00000

 FST P values

Number of permutations : 1023

	1	2	3	4	5	6
1	*					
2	0.26367+-0.0133	*				
3	0.00781+-0.0031	0.05859+-0.0070	*			
4	0.11816+-0.0106	0.54980+-0.0161	0.10352+-0.0113	*		
5	0.04492+-0.0073	0.20703+-0.0135	0.70605+-0.0145	0.31738+-0.0129	*	
6	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0026	0.00000+-0.0000	0.00098+-0.0010	*

 Matrix of significant Fst P values
 Significance Level=0.0500

Number of permutations : 1023

	1	2	3	4	5	6
1		-	+	-	+	+
2	-		-	-	-	+
3	+	-		-	-	+
4	-	-	-		-	+
5	+	-	-	-		+
6	+	+	+	+	+	

 AMOVA design and results (average over 49 loci):

Source of variation	Sum of squares	Variance components
Among populations	28.389	0.22316
Within populations	186.918	1.96756
Total	215.307	2.19072

Average F-Statistics over all loci
 Fixation Indices
 FST : 0.10187

 Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.00000

Table of the Haplotype distribution through the Populations:

Haplotype:	MA (8)	CI (16)	CB (19)	EM (20)	WM (18)	NS (20)
MulS_MA_03						
MulS_MA_06						
MulS_MA_09						
MulS_MA_14						
MulS_MA_19						
MulS_CI_01						
MulS_CI_02						
MulS_CI_05						
MulS_CI_12						
MulS_CI_14						
MulS_CI_16						
MulS_CI_17						
MulS_CB_11						
MulS_CB_16						
MulS_EM_08						
MulS_EM_12						
MulS_EM_13						
MulS_EM_20						
MulS_WM_01						
MulS_WM_09						
MulS_WM_18	5	7	2	4	3	0
MulS_MA_11	1	0	0	0	0	0
MulS_MA_13	1	0	0	0	0	0
MulS_MA_20	1	0	0	0	0	0
MulS_CI_04	0	1	0	0	0	0
MulS_CI_06	0	1	0	0	0	0
MulS_CI_08						
MulS_EM_04						
MulS_WM_06						
MulS_WM_13						
MulS_WM_17	0	1	0	1	3	0
MulS_CI_09	0	1	0	0	0	0
MulS_CI_10						
MulS_CI_15						
MulS_EM_18	0	2	0	1	0	0
MulS_CI_11	0	1	0	0	0	0
MulS_CI_13	0	1	0	0	0	0
MulS_CI_19	0	1	0	0	0	0
MulS_CB_01	0	0	1	0	0	0
MulS_CB_02	0	0	1	0	0	0
MulS_CB_03	0	0	1	0	0	0
MulS_CB_04	0	0	1	0	0	0
MulS_CB_05	0	0	1	0	0	0
MulS_CB_06						
MulS_CB_10						
MulS_EM_05						
MulS_WM_14						
MulS_WM_19						
MulS_NS_05						
MulS_NS_06						
MulS_NS_07						
MulS_NS_08						
MulS_NS_16						
MulS_NS_17						
MulS_NS_20	0	0	2	1	2	7
MulS_CB_07	0	0	1	0	0	0
MulS_CB_08						
MulS_EM_03						
MulS_EM_06						
MulS_EM_11						

MulS_WM_16						
MulS_NS_01						
MulS_NS_02						
MulS_NS_04						
MulS_NS_09	0	0	1	3	1	4
MulS_CB_09	0	0	1	0	0	0
MulS_CB_12	0	0	1	0	0	0
MulS_CB_13	0	0	1	0	0	0
MulS_CB_14						
MulS_EM_16	0	0	1	1	0	0
MulS_CB_15	0	0	1	0	0	0
MulS_CB_17	0	0	1	0	0	0
MulS_CB_18						
MulS_WM_12	0	0	1	0	1	0
MulS_CB_19	0	0	1	0	0	0
MulS_EM_01						
MulS_EM_02	0	0	0	2	0	0
MulS_EM_07	0	0	0	1	0	0
MulS_EM_09	0	0	0	1	0	0
MulS_EM_10	0	0	0	1	0	0
MulS_EM_14	0	0	0	1	0	0
MulS_EM_15	0	0	0	1	0	0
MulS_EM_17	0	0	0	1	0	0
MulS_EM_19	0	0	0	1	0	0
MulS_WM_02	0	0	0	0	1	0
MulS_WM_03						
MulS_WM_20	0	0	0	0	2	0
MulS_WM_04	0	0	0	0	1	0
MulS_WM_07	0	0	0	0	1	0
MulS_WM_08	0	0	0	0	1	0
MulS_WM_11	0	0	0	0	1	0
MulS_WM_15	0	0	0	0	1	0
MulS_NS_03	0	0	0	0	0	1
MulS_NS_10	0	0	0	0	0	1
MulS_NS_11	0	0	0	0	0	1
MulS_NS_12	0	0	0	0	0	1
MulS_NS_13						
MulS_NS_14	0	0	0	0	0	2
MulS_NS_15	0	0	0	0	0	1
MulS_NS_18	0	0	0	0	0	1
MulS_NS_19	0	0	0	0	0	1

APPENDIX 5

Pagrus pagrus Genetic Population Structure Analyses

[[Structure]]

```

StructureName = "Pagrus pagrus"
NbGroups = 1
IndividualLevel = 0
DistMatLabel = ""
Group={
"MA" 20
"CI" 20
"WM" 20
"EM" 12
}

```

Distance method: Pairwise difference

AMOVA design and results :

Reference: Weir, B.S. and Cockerham, C.C. 1984.
 Excoffier, L., Smouse, P., and Quattro, J. 1992.
 Weir, B. S., 1996.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	3	33.461	0.48865 Va	16.53
Within populations	68	167.733	2.46667 Vb	83.47
Total	71	201.194	2.95531	
Fixation Index	FST :	0.16534		

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.00000+-0.00000

=====
 == Comparisons of pairs of population samples
 =====

List of labels for population samples used below:

Label	Population name
1:	MA
2:	CI
3:	WM
4:	EM

 Population pairwise FSTs

Distance method: Pairwise difference

	1	2	3	4
1	0.00000			
2	0.20174	0.00000		
3	0.21477	0.00617	0.00000	
4	0.21887	0.13104	0.15493	0.00000

 FST P values

Number of permutations : 1023

	1	2	3	4
1	*			
2	0.00000+-0.0000	*		
3	0.00000+-0.0000	0.29297+-0.0147	*	
4	0.00000+-0.0000	0.00098+-0.0010	0.00000+-0.0000	*

 Matrix of significant Fst P values
 Significance Level=0.0500

Number of permutations : 1023

	1	2	3	4
1		+	+	+
2	+		-	+
3	+	-		+
4	+	+	+	

 AMOVA design and results (average over 42 loci):

Source of variation	Sum of squares	Variance components
Among populations	15.017	0.22119
Within populations	72.983	1.07328
Total	88.000	1.29447

Average F-Statistics over all loci
 Fixation Indices
 FST : 0.17087

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.00000
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.00000

Table of the Haplotype distribution through the Populations:

Haplotype:	MA (20)	CI (20)	WM (20)	EM (12)
Pag_MA_01	1	0	0	0
Pag_MA_02	1	0	0	0
Pag_MA_03				
Pag_CI_01				
Pag_CI_05				
Pag_CI_07				
Pag_CI_15				
Pag_CI_17				
Pag_CI_20				
Pag_WM_02				
Pag_WM_07				
Pag_WM_09				
Pag_WM_10				
Pag_WM_11				
Pag_WM_15				
Pag_WM_16				
Pag_WM_17				
Pag_EM_04				
Pag_EM_11	1	6	8	2
Pag_MA_04	1	0	0	0
Pag_MA_05				
Pag_MA_06				
Pag_MA_07				
Pag_MA_08				
Pag_MA_09	5	0	0	0
Pag_MA_10				
Pag_MA_13				
Pag_MA_15				
Pag_MA_17				
Pag_MA_19	5	0	0	0
Pag_MA_11	1	0	0	0
Pag_MA_12	1	0	0	0
Pag_MA_14	1	0	0	0
Pag_MA_16	1	0	0	0
Pag_MA_18	1	0	0	0
Pag_MA_20	1	0	0	0
Pag_CI_02	0	1	0	0
Pag_CI_03	0	1	0	0
Pag_CI_04				
Pag_CI_08				
Pag_CI_09				
Pag_CI_13				
Pag_CI_14				
Pag_CI_16				
Pag_WM_01				
Pag_WM_04				
Pag_WM_05				
Pag_WM_12				
Pag_WM_13				
Pag_WM_14				
Pag_EM_03				
Pag_EM_21	0	6	6	2
Pag_CI_06	0	1	0	0
Pag_CI_10	0	1	0	0
Pag_CI_11	0	1	0	0
Pag_CI_12	0	1	0	0
Pag_CI_18	0	1	0	0
Pag_CI_19	0	1	0	0

Pag_WM_03	0	0	1	0
Pag_WM_06	0	0	1	0
Pag_WM_08				
Pag_WM_18	0	0	2	0
Pag_WM_19	0	0	1	0
Pag_WM_20	0	0	1	0
Pag_EM_01	0	0	0	1
Pag_EM_02	0	0	0	1
Pag_EM_05	0	0	0	1
Pag_EM_06	0	0	0	1
Pag_EM_07	0	0	0	1
Pag_EM_08	0	0	0	1
Pag_EM_10	0	0	0	1
Pag_EM_13	0	0	0	1

APPENDIX 6

Pagellus erythrinus Genetic Population Structure Analyses

Pag Ery

[[Structure]]

```

StructureName = "Pag Ery"
NbGroups = 1
IndividualLevel = 0
DistMatLabel = ""
Group={
"EM" 18
"CI" 19
"WM" 17
}

```

Distance method: Tamura & Nei

AMOVA design and results :

Reference: Weir, B.S. and Cockerham, C.C. 1984.
 Excoffier, L., Smouse, P., and Quattro, J. 1992.
 Weir, B. S., 1996.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	2	6.120	-0.10407 Va	-2.16
Within populations	51	251.492	4.93122 Vb	102.16
Total	53	257.612	4.82715	
Fixation Index	FST :	-0.02156		

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.78886
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.78886+-0.01373

=====
 == Comparisons of pairs of population samples
 =====

List of labels for population samples used below:

Label	Population name
1:	EM
2:	CI
3:	WM

 Population pairwise FSTs

Distance method: Tamura & Nei

	1	2	3
1	0.00000		
2	-0.01444	0.00000	
3	-0.02120	-0.02967	0.00000

 FST P values

Number of permutations : 1023

	1	2	3
1	*		
2	0.57910+-0.0131	*	
3	0.67871+-0.0132	0.78516+-0.0124	*

 Matrix of significant Fst P values
 Significance Level=0.0500

Number of permutations : 1023

	1	2	3
1	-	-	-
2	-	-	-
3	-	-	-

AMOVA design and results (average over 52 loci):

Source of variation	Sum of squares	Variance components
Among populations	5.860	-0.09882
Within populations	240.066	4.70717
Total	245.926	4.60835

Average F-Statistics over all loci

Fixation Indices
 FST : -0.02144

Significance tests (1023 permutations)

Va and FST : P(rand. value > obs. value) = 0.99316
 P(rand. value = obs. value) = 0.00000
 P(rand. value >= obs. value) = 0.99316

Table of the Haplotype distribution through the Populations:

Haplotype:	EM (18)	CI (19)	WM (17)
Pag_Ery_01_FRI			
Pag_Ery_11_FRI			
Pag_Ery_CI_10			
Pag_Ery_CI_12			
Pag_Ery_WM_08	2	2	1
Pag_Ery_02_FRI			
Pag_Ery_WM_14	1	0	1
Pag_Ery_03_FRI	1	0	0
Pag_Ery_04_FRI	1	0	0
Pag_Ery_05_FRI	1	0	0
Pag_Ery_06_FRI	1	0	0
Pag_Ery_07_FRI	1	0	0
Pag_Ery_08_FRI	1	0	0
Pag_Ery_09_FRI	1	0	0
Pag_Ery_10_FRI	1	0	0
Pag_Ery_12_FRI	1	0	0
Pag_Ery_13_FRI	1	0	0
Pag_Ery_15_FRI	1	0	0
Pag_Ery_17_FRI	1	0	0
Pag_Ery_18_FRI	1	0	0
Pag_Ery_19_FRI	1	0	0
Pag_Ery_20_FRI	1	0	0
Pag_Ery_CI_01	0	1	0
Pag_Ery_CI_03	0	1	0
Pag_Ery_CI_04	0	1	0
Pag_Ery_CI_05	0	1	0
Pag_Ery_CI_06	0	1	0
Pag_Ery_CI_07	0	1	0
Pag_Ery_CI_08	0	1	0
Pag_Ery_CI_09	0	1	0
Pag_Ery_CI_11	0	1	0
Pag_Ery_CI_13			
Pag_Ery_WM_06			
Pag_Ery_WM_20	0	1	2
Pag_Ery_CI_14			
Pag_Ery_WM_18	0	1	1
Pag_Ery_CI_15	0	1	0
Pag_Ery_CI_16			
Pag_Ery_WM_04			
Pag_Ery_WM_16	0	1	2
Pag_Ery_CI_17	0	1	0
Pag_Ery_CI_18			
Pag_Ery_WM_10	0	1	1
Pag_Ery_CI_19	0	1	0
Pag_Ery_CI_20	0	1	0

Pag_Ery_WM_03	0	0	1

Pag_Ery_WM_05	0	0	1

Pag_Ery_WM_07	0	0	1

Pag_Ery_WM_09	0	0	1

Pag_Ery_WM_11	0	0	1

Pag_Ery_WM_12	0	0	1

Pag_Ery_WM_13	0	0	1

Pag_Ery_WM_15	0	0	1

Pag_Ery_WM_17	0	0	1

Haplotype connectivity network analysis of *Solea solea* populations with the TCS v1.21 algorithm.

Number of sequences: 92
Length of sequences: 1141
Sequence file used: SolSol nexus

PARSIMONY PROBABILITY

For 14 step(s), P(95%) = 0.954876701413056

RUN SETTINGS

Calculated maximum connection steps at 95% = 14
Gaps treated as fifth state

Results and discussion

As with the F_{st} and AMOVA analyses the haplotype connectivity network of Figure 1 displays a well defined genetic population structure for *Solea solea* in the different geographic areas. Only the NS and the CB samples share a significant number of common haplotypes. Furthermore, the network demonstrates that the EM sample is most genetically distant from the outgroup that contains only NS and CB haplotypes. Note also that eight BS haplotypes are intermediately positioned between the bulk of the NS and WM haplotypes. Therefore, there is no clear haplotype gradient, in terms of accumulated mutations in the *cyt b* sequence, from North to South to East. Important and of interest to traceability purposes is that a single A to G transition, at position 1020 in 21 of the 22 EM individuals examined, is sufficient to genetically distinguish this population from that of WM and in fact from all populations studied.

Table of *Solea solea* HAPLOTYPES

Number of haplotypes = 40

Haplotype list:

- EM_01 : EM_02 EM_04 EM_06 EM_10 EM_12 EM_15 EM_18 EM_20 EM_21 EM_22
- EM_03 :
- EM_05 :
- EM_07 :
- EM_08 :
- EM_09 :
- EM_11 :
- EM_13 :
- EM_14 :
- EM_16 :
- EM_17 :
- EM_19 :
- NS_01 : NS_06 NS_09 NS_15 NS_20 BS_01 BS_03 CB_02 CB_04 CB_09 CB_12 CB_13 CB_14
- NS_02 : NS_11 NS_13 NS_16 CB_11
- NS_03 : NS_05 NS_08 NS_10 NS_12 NS_14 NS_18 CB_03
- NS_04 : NS_07 NS_17 NS_19
- BS_02 :
- BS_04 :
- BS_05 : BS_06 BS_09 BS_12 BS_14 BS_17 BS_18 BS_19
- BS_07 : BS_16 BS_20
- BS_08 :
- BS_10 :
- BS_11 : BS_13 BS_15 CB_10
- BS_21 :
- WM_01 : WM_07 WM_08 WM_13
- WM_02 :
- WM_03 : WM_15
- WM_04 :
- WM_05 :
- WM_06 :
- WM_09 :
- WM_10 :
- WM_11 :
- WM_12 :
- WM_14 :
- CB_01 :
- CB_05 :
- CB_06 :
- CB_07 :
- CB_08 :

OUTGROUP WEIGHTS

*** Network 1

BS_04 weight = 0.0024691358024691358
CB_01 weight = 0.0024691358024691358
CB_05 weight = 0.0024691358024691358
BS_21 weight = 0.0024691358024691358
BS_11 weight = 0.009876543209876543
BS_08 weight = 0.0024691358024691358
BS_05 weight = 0.019753086419753086
BS_02 weight = 0.0024691358024691358
NS_03 weight = 0.12345679012345678
NS_01 weight = 0.15308641975308643
NS_02 weight = 0.10864197530864197
BS_07 weight = 0.007407407407407408
BS_10 weight = 0.0024691358024691358
CB_06 weight = 0.07901234567901234
NS_04 weight = 0.08395061728395062
WM_06 weight = 0.0024691358024691358
CB_07 weight = 0.0024691358024691358
CB_08 weight = 0.0024691358024691358
WM_05 weight = 0.009876543209876543
WM_10 weight = 0.0024691358024691358
WM_14 weight = 0.0024691358024691358
WM_11 weight = 0.0024691358024691358
WM_03 weight = 0.0049382716049382715
EM_19 weight = 0.0024691358024691358
EM_14 weight = 0.0024691358024691358
EM_11 weight = 0.0024691358024691358
EM_09 weight = 0.0024691358024691358
EM_07 weight = 0.0024691358024691358
EM_03 weight = 0.0024691358024691358
EM_01 weight = 0.1037037037037037
EM_05 weight = 0.0024691358024691358
EM_08 weight = 0.0024691358024691358
EM_13 weight = 0.05925925925925926
EM_16 weight = 0.06419753086419754
WM_01 weight = 0.10864197530864197
EM_17 weight = 0.0024691358024691358
WM_02 weight = 0.0024691358024691358
WM_04 weight = 0.0024691358024691358
WM_09 weight = 0.0024691358024691358
WM_12 weight = 0.0024691358024691358
Total weight = 202.5

Biggest outgroup probability is NS_01 (0.15308641975308643)

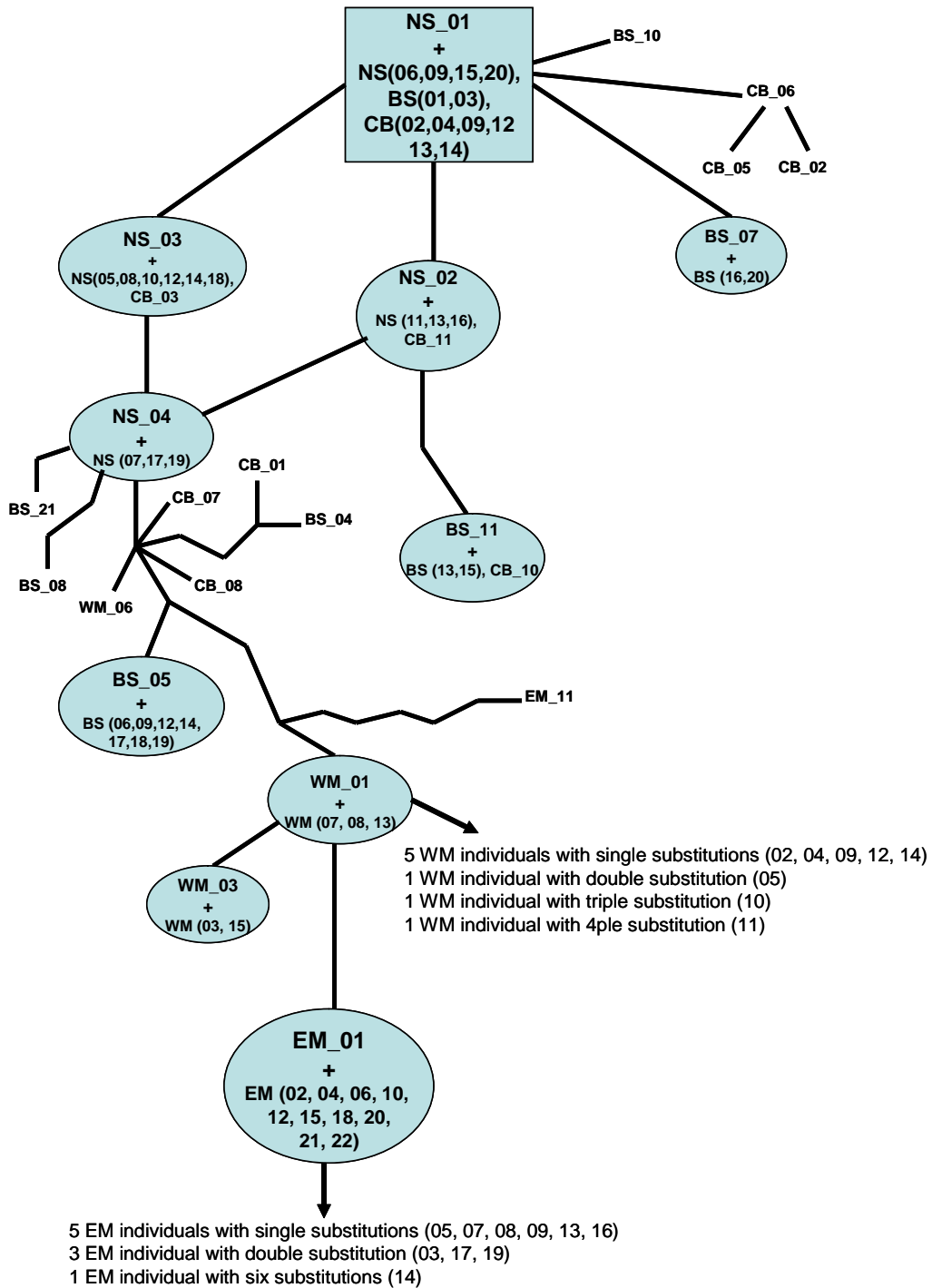


Figure 1. Connectivity network of *Solea solea* haplotypes. The rectangle and the ovals represent shared haplotypes between individuals, as indicated. The rectangle represents the biggest probability outgroup (see analyses). Each straight line in the network represents a single nucleotide substitution (e.g. there are 7 nucleotide changes between WM_01 and EM_11 in the lower half of the network; there is no relationship between line length and genetic distance).

Haplotype connectivity network analysis of of *Micromesistius poutassou* populations with the TCS v1.21 algorithm.

Number of sequences: 94
Length of sequences: 1141
Sequence file used: MicPou nexus

PARSIMONY PROBABILITY

For 14 step(s), P(95%) = 0.954876701413056

RUN SETTINGS

Calculated maximum connection steps at 95% = 14
Gaps treated as fifth state

Results and discussion

The network of Figure 2 demonstrates that the majority of the *Micromesistius poutassou* EM haplotypes contain one or two substitutions as compared to those of the CB and CS samples, explaining, thus, the results of the F_{st} analysis that did not detect significant differences in the populations of these three geographical areas. The network also demonstrates the high degree of genetic variation within all populations with the exception of that of the NS.

Table of *Micromesistius poutassou* HAPLOTYPES

Number of haplotypes = 60

Haplotype list:

- CB_01 :
- CB_02 : CB_10 CB_15 CS_12 EM_08 EM_11 EM_19
- CB_03 :
- CB_04 :
- CB_05 :
- CB_06 : CB_20 NS_01 NS_02 NS_03 NS_04 NS_06 NS_07 NS_08 NS_10 NS_11 NS_15 NS_16
NS_17 NS_19 NS_20
- CB_07 :
- CB_08 :
- CB_09 : EM_20
- CB_11 :
- CB_12 : EM_06
- CB_13 : CB_14 CB_17
- CB_16 :
- CB_18 :
- CB_19 :
- CS_01 : CS_08
- CS_02 :
- CS_03 : EM_18
- CS_04 :
- CS_05 :
- CS_06 :
- CS_07 : WM_17
- CS_09 :
- CS_10 :
- CS_11 :
- CS_13 :
- CS_14 :
- CS_15 :
- CS_16 :
- EM_01 : NS_05 NS_14 NS_18
- EM_02 :
- EM_03 :
- EM_04 :
- EM_05 :
- EM_07 :
- EM_09 :
- EM_10 :
- EM_12 :
- EM_13 :
- EM_14 : WM_08 WM_09
- EM_15 :
- EM_16 :
- EM_17 :
- NS_09 :
- NS_12 :
- NS_13 :
- WM_01 :
- WM_02 :
- WM_03 : WM_06
- WM_04 :
- WM_05 :
- WM_07 :
- WM_10 :
- WM_11 :
- WM_12 :
- WM_13 :
- WM_14 :
- WM_15 :
- WM_16 :
- WM_18 :

OUTGROUP WEIGHTS

*** Network 1

CS_02 weight = 0.0020833333333333333
CS_16 weight = 0.0083333333333333333
WM_11 weight = 0.0020833333333333333
WM_02 weight = 0.0020833333333333333
EM_15 weight = 0.0020833333333333333
EM_12 weight = 0.0020833333333333333
EM_07 weight = 0.0020833333333333333
EM_05 weight = 0.0020833333333333333
EM_04 weight = 0.0020833333333333333
CS_15 weight = 0.0020833333333333333
CS_14 weight = 0.0020833333333333333
CS_11 weight = 0.0020833333333333333
CS_10 weight = 0.0020833333333333333
CS_05 weight = 0.0020833333333333333
CS_04 weight = 0.0020833333333333333
CS_01 weight = 0.0041666666666666667
CB_16 weight = 0.0020833333333333333
CB_13 weight = 0.00625
CB_12 weight = 0.0083333333333333333
CB_08 weight = 0.0020833333333333333
CB_04 weight = 0.0708333333333333333
CB_03 weight = 0.0020833333333333333
CB_06 weight = 0.1208333333333333333
CB_05 weight = 0.0083333333333333333
WM_01 weight = 0.0791666666666666666
CB_19 weight = 0.0020833333333333333
CS_03 weight = 0.0791666666666666666
EM_01 weight = 0.0958333333333333334
EM_03 weight = 0.0020833333333333333
EM_10 weight = 0.075
NS_13 weight = 0.075
NS_09 weight = 0.0208333333333333332
CB_02 weight = 0.0416666666666666664
CB_18 weight = 0.0020833333333333333
EM_16 weight = 0.0583333333333333334
WM_13 weight = 0.0333333333333333333
NS_12 weight = 0.025
WM_10 weight = 0.0020833333333333333
WM_05 weight = 0.0083333333333333333
WM_12 weight = 0.0083333333333333333
WM_14 weight = 0.0020833333333333333
WM_07 weight = 0.0020833333333333333
WM_04 weight = 0.0020833333333333333
EM_13 weight = 0.0020833333333333333
EM_09 weight = 0.0020833333333333333
CB_01 weight = 0.0020833333333333333
CB_09 weight = 0.0083333333333333333
EM_02 weight = 0.0166666666666666666
CB_11 weight = 0.0020833333333333333
CS_07 weight = 0.0083333333333333333
EM_14 weight = 0.0333333333333333333
WM_15 weight = 0.0020833333333333333
CS_06 weight = 0.0020833333333333333
WM_03 weight = 0.0208333333333333332
CS_13 weight = 0.0020833333333333333
CB_07 weight = 0.0020833333333333333
EM_17 weight = 0.0020833333333333333
WM_16 weight = 0.0020833333333333333
CS_09 weight = 0.0020833333333333333
WM_18 weight = 0.0125
Total weight = 240.0

Biggest outgroup probability is CB_06 (0.1208333333333333333)

Haplotype connectivity network analysis of of *Merluccius merluccius* populations with the TCS v1.21 algorithm.

Number of sequences: 132
Length of sequences: 1141
Sequence file used: MerMer nexus

PARSIMONY PROBABILITY

For 14 step(s), P(95%) = 0.954876701413056

RUN SETTINGS

Calculated maximum connection steps at 95% = 14
Gaps treated as fifth state

Results and discussion

The network of Figure 3 demonstrates that a significant number of CI, CS, CB and BS individuals share a common haplotype (outgroup EM_08). Another significant cluster of a shared haplotype (EM_02) contains mostly EM and CB individuals. The above explain the results of the F_{st} analysis for the populations of this species and also demonstrate that the NS sample is more genetically distant from the other samples analyzed.

Table of *Merluccius merluccius* HAPLOTYPES

Number of haplotypes = 49

Haplotype list:

- EM_01 : EM_09 CI_16 BS_03 CB_03 CB_10 CB_11 CB_16 NS_01 WM_02
- EM_02 : EM_03 EM_04 EM_10 EM_12 EM_17 CS-11 CS-14 BS_10 BS_15 CB_02 CB_06 CB_08
CB_17 NS_03 NS_05 WM_04 WM_06
- EM_05 :
- EM_06 : CI_03 CI_09 CB_01 CB_07 WM_01
- EM_07 :
- EM_08 : CI_01 CI_02 CI_10 CI_14 CI-18 CS-01 CS-02 CS-09 CS-12 CS-15 CS-16 CS-17
BS_01 BS_02 BS_05 BS_07 BS_08 BS_11 BS_16 BS_17 BS_18 BS_20 CB_04 CB_09 CB_12 CB_18
NS_02 NS_04 NS_06 WM_03 WM_05 WM_07
- EM_11 :
- EM_13 :
- EM_14 : CI_06
- EM_15 : EM_16
- CI_04 : CS-19
- CI_05 :
- CI_07 : CI-19
- CI_08 :
- CI_11 :
- CI_12 :
- CI_13 :
- CI_15 :
- CI_17 :
- CS-03 :
- CS-04 : CS-05
- CS-06 :
- CS-07 :
- CS-08 :
- CS-10 :
- CS-13 :
- CS-18 :
- BS_04 :
- BS_06 :
- BS_09 :
- BS_12 :
- BS_13 :
- BS_14 :
- BS_19 : CB_15
- CB_05 :
- CB_13 :
- CB_14 :
- NS_07 : WM_08
- NS_08 : WM_09
- NS_09 : WM_10
- NS_10 : WM_11
- NS_11 : NS_14 WM_12 WM_15
- NS_12 : WM_13
- NS_13 : WM_14
- NS_15 : WM_16
- NS_16 : WM_17
- NS_17 : WM_18
- NS_18 : WM_19
- NS_19 : WM_20

OUTGROUP WEIGHTS

*** Network 1

EM_14 weight = 0.013333333333333334
CI_05 weight = 0.002222222222222222
CS-04 weight = 0.004444444444444444
CB_05 weight = 0.013333333333333334
CB_13 weight = 0.002222222222222222
CS-18 weight = 0.002222222222222222
CI_15 weight = 0.002222222222222222
CI_12 weight = 0.002222222222222222
EM_01 weight = 0.053333333333333334
CS-10 weight = 0.002222222222222222
BS_04 weight = 0.002222222222222222
EM_13 weight = 0.002222222222222222
EM_02 weight = 0.093333333333333334
EM_07 weight = 0.002222222222222222
EM_11 weight = 0.002222222222222222
EM_15 weight = 0.004444444444444444
EM_06 weight = 0.031111111111111111
CI_13 weight = 0.002222222222222222
CB_14 weight = 0.002222222222222222
EM_05 weight = 0.002222222222222222
CI_04 weight = 0.017777777777777778
CI_07 weight = 0.004444444444444444
BS_14 weight = 0.002222222222222222
BS_12 weight = 0.002222222222222222
BS_09 weight = 0.002222222222222222
CS-13 weight = 0.002222222222222222
CS-07 weight = 0.15555555555555556
CS-03 weight = 0.002222222222222222
CI_08 weight = 0.002222222222222222
EM_08 weight = 0.191111111111111112
CI_11 weight = 0.002222222222222222
CI_17 weight = 0.002222222222222222
CS-06 weight = 0.002222222222222222
CS-08 weight = 0.002222222222222222
BS_13 weight = 0.002222222222222222
BS_19 weight = 0.16
BS_06 weight = 0.017777777777777778
NS_13 weight = 0.008888888888888889
NS_11 weight = 0.026666666666666667
NS_09 weight = 0.017777777777777778
NS_12 weight = 0.035555555555555556
NS_15 weight = 0.008888888888888889
NS_17 weight = 0.017777777777777778
NS_16 weight = 0.004444444444444444
NS_10 weight = 0.004444444444444444
NS_07 weight = 0.026666666666666667
NS_08 weight = 0.004444444444444444
NS_19 weight = 0.026666666666666667
NS_18 weight = 0.004444444444444444
Total weight = 225.0

Biggest outgroup probability is EM_08 (0.191111111111111112)

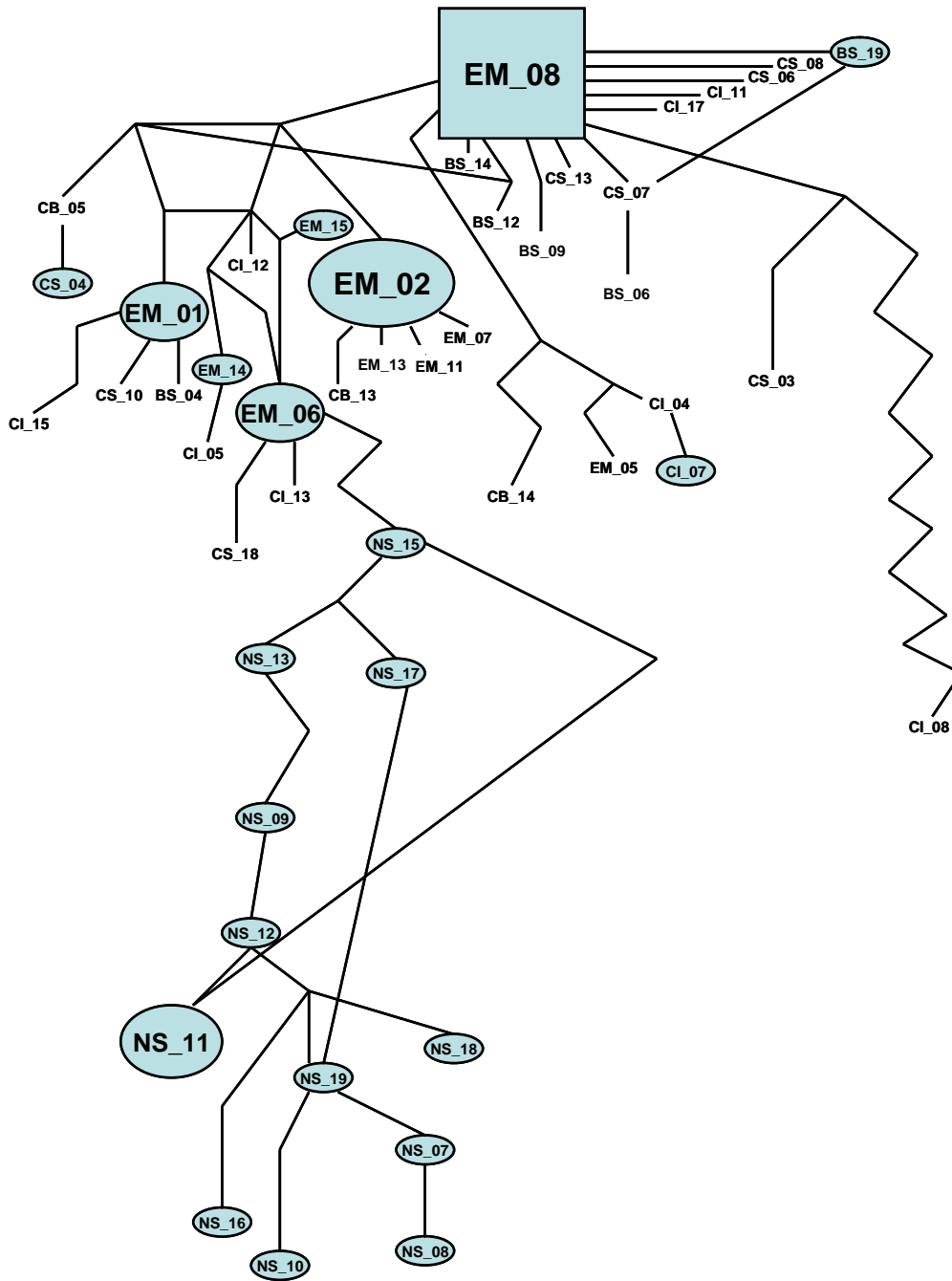


Figure 3. Connectivity network of *Merluccius merluccius* haplotypes. Symbols are as in Figure 1. For shared haplotypes see the respective table of the analyses.

Haplotype connectivity network analysis of of *Pagellus erythrinus* populations with the TCS v1.21 algorithm.

Number of sequences: 54
Length of sequences: 1141
Sequence file used: PagEry nexus

PARSIMONY PROBABILITY
For 14 step(s), P(95%) = 0.954876701413056

RUN SETTINGS
Calculated maximum connection steps at 95% = 14
Gaps treated as fifth state

Results and discussion

The network of Figure 4 demonstrates the high level of interpopulation genetic variability. A limited but nevertheless significant number of shared haplotypes, as well as clusters of haplotypes within the three populations examined, with one or two substitutions explain the results of the F_{st} analysis, i.e. the panmictic nature of the populations of the species. Notable is however, the accumulation of mutations in the individuals placed in the lower part of the network suggesting a large effective population size.

Table of *Pagellus erythrinus* HAPLOTYPES

Number of haplotypes = 43

Haplotype list:

- EM_01 : EM_11 CI_10 CI_12 WM_08
- EM_02 : WM_14
- EM_03 :
- EM_04 :
- EM_05 :
- EM_06 :
- EM_07 :
- EM_08 :
- EM_09 :
- EM_10 :
- EM_12 :
- EM_13 :
- EM_15 :
- EM_17 :
- EM_18 :
- EM_19 :
- EM_20 :
- CI_01 :
- CI_03 :
- CI_04 :
- CI_05 :
- CI_06 :
- CI_07 :
- CI_08 :
- CI_09 :
- CI_11 :
- CI_13 : WM_06 WM_20
- CI_14 : WM_18
- CI_15 :
- CI_16 : WM_04 WM_16
- CI_17 :
- CI_18 : WM_10
- CI_19 :
- CI_20 :
- WM_03 :
- WM_05 :
- WM_07 :
- WM_09 :
- WM_11 :
- WM_12 :
- WM_13 :
- WM_15 :
- WM_17 :

OUTGROUP WEIGHTS

*** Network 1

EM_20 weight = 0.0106951871657754
CI_06 weight = 0.0053475935828877
WM_07 weight = 0.0053475935828877
EM_12 weight = 0.053475935828877004
EM_04 weight = 0.0053475935828877
CI_13 weight = 0.0748663101604278
WM_09 weight = 0.0053475935828877
EM_10 weight = 0.0053475935828877
CI_14 weight = 0.06417112299465241
WM_17 weight = 0.0053475935828877
CI_08 weight = 0.0053475935828877
EM_02 weight = 0.03208556149732621
CI_11 weight = 0.0053475935828877
WM_11 weight = 0.0053475935828877
CI_20 weight = 0.0053475935828877
CI_09 weight = 0.0053475935828877
CI_07 weight = 0.0053475935828877
EM_18 weight = 0.0053475935828877
EM_17 weight = 0.0053475935828877
EM_15 weight = 0.0053475935828877
EM_13 weight = 0.0427807486631016
EM_09 weight = 0.0053475935828877
EM_07 weight = 0.0053475935828877
EM_06 weight = 0.0053475935828877
EM_03 weight = 0.0053475935828877
EM_01 weight = 0.12834224598930483
EM_05 weight = 0.06417112299465241
CI_05 weight = 0.0855614973262032
CI_16 weight = 0.0962566844919786
CI_18 weight = 0.0855614973262032
WM_03 weight = 0.0053475935828877
CI_17 weight = 0.0053475935828877
WM_05 weight = 0.0053475935828877
CI_01 weight = 0.0053475935828877
CI_04 weight = 0.0213903743315508
CI_15 weight = 0.0053475935828877
CI_03 weight = 0.0053475935828877
EM_19 weight = 0.0053475935828877
EM_08 weight = 0.0053475935828877
WM_12 weight = 0.053475935828877004
CI_19 weight = 0.0053475935828877
WM_13 weight = 0.03208556149732621
WM_15 weight = 0.0053475935828877
Total weight = 93.5
Biggest outgroup probability is EM_01 (0.12834224598930483)

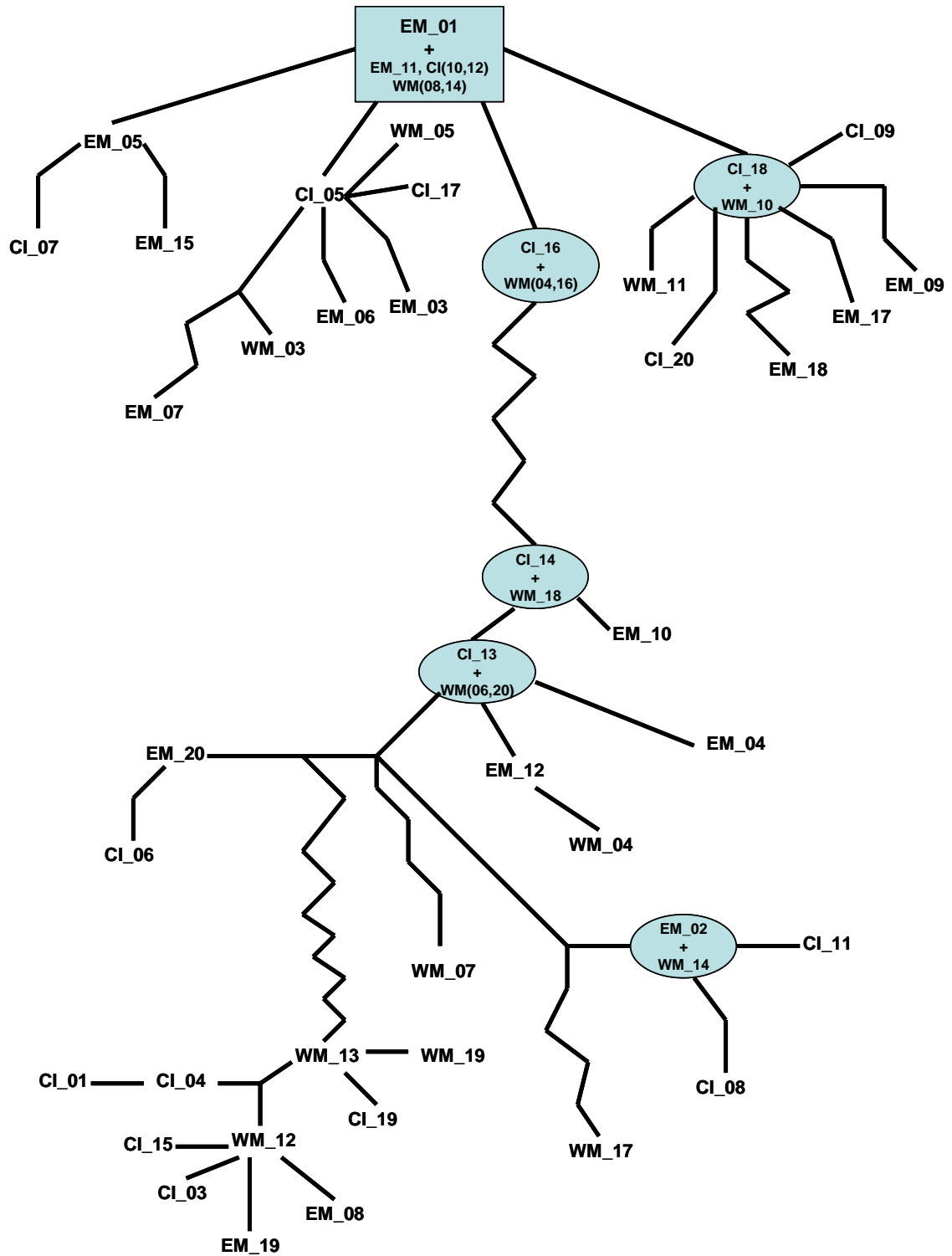


Figure 4. Connectivity network of *Pagellus erythrinus* haplotypes. Symbols are explained in Figure 1. Shared haplotypes are indicated.

Haplotype connectivity network analysis of of *Pagrus pagrus* populations with the TCS v1.21 algorithm.

Number of sequences: 72
Length of sequences: 1141
Sequence file used: PagPag nexus

PARSIMONY PROBABILITY

For 14 step(s), P(95%) = 0.954876701413056

RUN SETTINGS

Calculated maximum connection steps at 95% = 14
Gaps treated as fifth state

Results and discussion

The network of Figure 5 demonstrates that areas CI and WM share a significant number of common haplotypes. Specifically the EM_03 haplotype is shared by 6 CI and 6 WM individuals and the EM_04 haplotype is shared by 6 and 7 CI and WM individuals, respectively. In contrast of the 12 individuals of the EM sample tested, only 4 share haplotypes with the WM and CI samples (two for each of the EM_03 and EM_04 haplotypes). Furthermore, 19 out of the 20 individuals of the MA sample present ara-specific haplotypes. The above explain the results of the F_{st} analysis for the populations of this species, which suggest non-significant differences in the genetic population structure of areas CI and WM and the existence of genetically distinct populations in areas MA and EM.

Table of *Pagrus pagrus* HAPLOTYPES

Number of haplotypes = 34

Haplotype list:

- EM_01 :
- EM_02 :
- EM_03 : EM_12 CI_04 CI_08 CI_09 CI_13 CI_14 CI_16 WM_01 WM_04 WM_05
WM_12 WM_13 WM_14
- EM_04 : EM_10 CI_01 CI_05 CI_07 CI_15 CI_17 CI_20 WM_02 WM_07 WM_09
WM_10 WM_11 WM_15 WM_16 WM_17 MA_03
- EM_05 :
- EM_06 :
- EM_07 :
- EM_08 :
- EM_09 :
- EM_11 :
- CI_02 :
- CI_03 :
- CI_06 :
- CI_10 :
- CI_11 :
- CI_12 :
- CI_18 :
- CI_19 :
- WM_03 :
- WM_06 :
- WM_08 : WM_18
- WM_19 :
- WM_20 :
- MA_01 :
- MA_02 :
- MA_04 :
- MA_05 : MA_06 MA_07 MA_08 MA_09
- MA_10 : MA_13 MA_15 MA_17 MA_19
- MA_11 :
- MA_12 :
- MA_14 :
- MA_16 :
- MA_18 :
- MA_20 :

OUTGROUP WEIGHTS

*** Network 1

MA_16 weight = 0.00186219739292365
MA_11 weight = 0.00186219739292365
MA_04 weight = 0.00186219739292365
MA_02 weight = 0.00186219739292365
MA_01 weight = 0.00186219739292365
WM_19 weight = 0.00186219739292365
WM_08 weight = 0.0037243947858473
WM_06 weight = 0.00186219739292365
CI_19 weight = 0.00186219739292365
CI_06 weight = 0.07076350093109869
CI_02 weight = 0.00186219739292365
EM_06 weight = 0.00186219739292365
EM_05 weight = 0.0595903165735568
EM_03 weight = 0.1601489757914339
EM_02 weight = 0.00186219739292365
EM_01 weight = 0.00186219739292365
EM_07 weight = 0.0595903165735568
EM_04 weight = 0.1527001862197393
EM_08 weight = 0.00186219739292365
EM_11 weight = 0.00186219739292365
CI_12 weight = 0.055865921787709494
WM_20 weight = 0.00186219739292365
MA_10 weight = 0.09683426443202979
CI_03 weight = 0.07076350093109869
CI_11 weight = 0.00186219739292365
WM_03 weight = 0.07076350093109869
MA_05 weight = 0.10800744878957169
MA_18 weight = 0.00186219739292365
MA_20 weight = 0.0260707635009311
EM_09 weight = 0.00186219739292365
CI_10 weight = 0.00186219739292365
CI_18 weight = 0.00186219739292365
MA_12 weight = 0.0260707635009311
MA_14 weight = 0.00186219739292365
Total weight = 268.5
Biggest outgroup probability is EM_03 (0.1601489757914339)

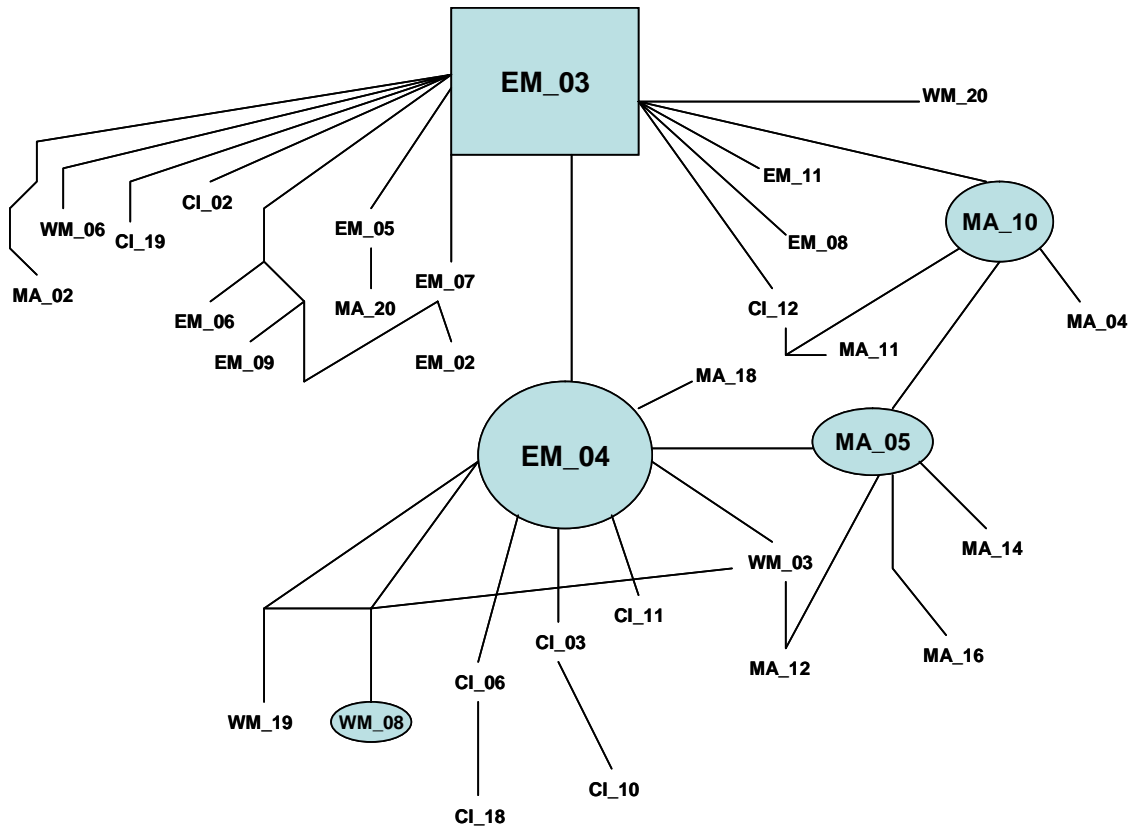


Figure 5. Connectivity network of *Pagrus pagrus* haplotypes. Symbols are as in Figure 1. For shared haplotypes see the respective table of the analyses.

Haplotype connectivity network analysis of of *Mullus surmuletus* populations with the TCS v1.21 algorithm.

Number of sequences: 101
Length of sequences: 1141
Sequence file used: MulSur nexus

PARSIMONY PROBABILITY

For 14 step(s), P(95%) = 0.954876701413056

RUN SETTINGS

Calculated maximum connection steps at 95% = 14
Gaps treated as fifth state

Results and discussion

The connectivity network of Figure 6 demonstrates that the results of the F_{st} analysis are primarily based on the haplotype cluster around the EM_08 outgroup containing individuals mostly from areas EM, WM, CI, MA, and CB. In contrast, the bulk of the NS haplotypes is located around the EM_05 outgroup. The sample with the largest number of unique haplotypes is that of area CB (12 unique haplotypes), followed by the NS sample (8 unique haplotypes). The MA sample with only 8 individuals examined appears to be related to both the CI and EM samples but more genetically distant from either of the WM or the CB samples.

Table of *Mullus surmuletus* HAPLOTYPES

Number of haplotypes = 50

Haplotype list:

- EM_01 : EM_02
- EM_03 : EM_06 EM_11 NS_01 NS_02 NS_04 NS_09 CB_08 WM_14
- EM_04 : CI_06 WM_05 WM_11 WM_15
- EM_05 : NS_05 NS_06 NS_07 NS_08 NS_16 NS_17 NS_20 CB_06 CB_10 WM_12 WM_17
- EM_07 :
- EM_08 : EM_09 EM_12 EM_13 EM_20 CB_11 CB_16 CI_01 CI_02 CI_04 CI_10 CI_12 CI_14
CI_15 MA_01 MA_02 MA_03 MA_06 MA_07 WM_01 WM_08 WM_16
- EM_10 :
- EM_14 :
- EM_15 :
- EM_16 : CB_14
- EM_17 :
- EM_18 : CI_08 CI_13
- EM_19 :
- NS_03 :
- NS_10 :
- NS_11 :
- NS_12 :
- NS_13 : NS_14
- NS_15 :
- NS_18 :
- NS_19 :
- CB_01 :
- CB_02 :
- CB_03 :
- CB_04 :
- CB_05 :
- CB_07 :
- CB_09 :
- CB_12 :
- CB_13 :
- CB_15 :
- CB_17 :
- CB_18 : WM_10
- CB_19 :
- CI_03 :
- CI_05 :
- CI_07 :
- CI_09 :
- CI_11 :
- CI_16 :
- MA_04 :
- MA_05 :
- MA_08 :
- WM_02 :
- WM_03 : WM_18
- WM_04 :
- WM_06 :
- WM_07 :
- WM_09 :
- WM_13 :

OUTGROUP WEIGHTS

*** Network 1

EM_19 weight = 0.0019801980198019802
NS_03 weight = 0.007920792079207921
WM_03 weight = 0.0039603960396039604
MA_04 weight = 0.0019801980198019802
CI_07 weight = 0.0019801980198019802
CB_15 weight = 0.0019801980198019802
CB_05 weight = 0.0019801980198019802
CB_04 weight = 0.0019801980198019802
CB_02 weight = 0.0019801980198019802
NS_19 weight = 0.007920792079207921
EM_16 weight = 0.019801980198019802
EM_15 weight = 0.0019801980198019802
EM_05 weight = 0.09504950495049505
EM_03 weight = 0.12673267326732673
EM_04 weight = 0.07920792079207921
EM_18 weight = 0.055445544554455446
NS_11 weight = 0.05148514851485148
NS_13 weight = 0.0039603960396039604
EM_10 weight = 0.0019801980198019802
NS_10 weight = 0.07128712871287128
CB_03 weight = 0.0019801980198019802
CB_07 weight = 0.0019801980198019802
CI_03 weight = 0.027722772277227723
CI_11 weight = 0.0019801980198019802
NS_15 weight = 0.05148514851485148
CB_19 weight = 0.0019801980198019802
NS_12 weight = 0.0019801980198019802
NS_18 weight = 0.007920792079207921
CB_17 weight = 0.0019801980198019802
WM_07 weight = 0.007920792079207921
WM_13 weight = 0.0019801980198019802
MA_05 weight = 0.0019801980198019802
WM_04 weight = 0.007920792079207921
CB_13 weight = 0.007920792079207921
CI_09 weight = 0.0019801980198019802
CB_01 weight = 0.0019801980198019802
CB_18 weight = 0.0039603960396039604
WM_06 weight = 0.0019801980198019802
CI_05 weight = 0.0019801980198019802
CB_09 weight = 0.0019801980198019802
EM_07 weight = 0.0019801980198019802
EM_01 weight = 0.09504950495049505
EM_08 weight = 0.1188118811881188
EM_17 weight = 0.09108910891089109
EM_14 weight = 0.0019801980198019802
CB_12 weight = 0.0019801980198019802
CI_16 weight = 0.0019801980198019802
WM_02 weight = 0.0019801980198019802
WM_09 weight = 0.0019801980198019802
MA_08 weight = 0.0019801980198019802
Total weight = 252.5

Biggest outgroup probability is EM_03 (0.12673267326732673)

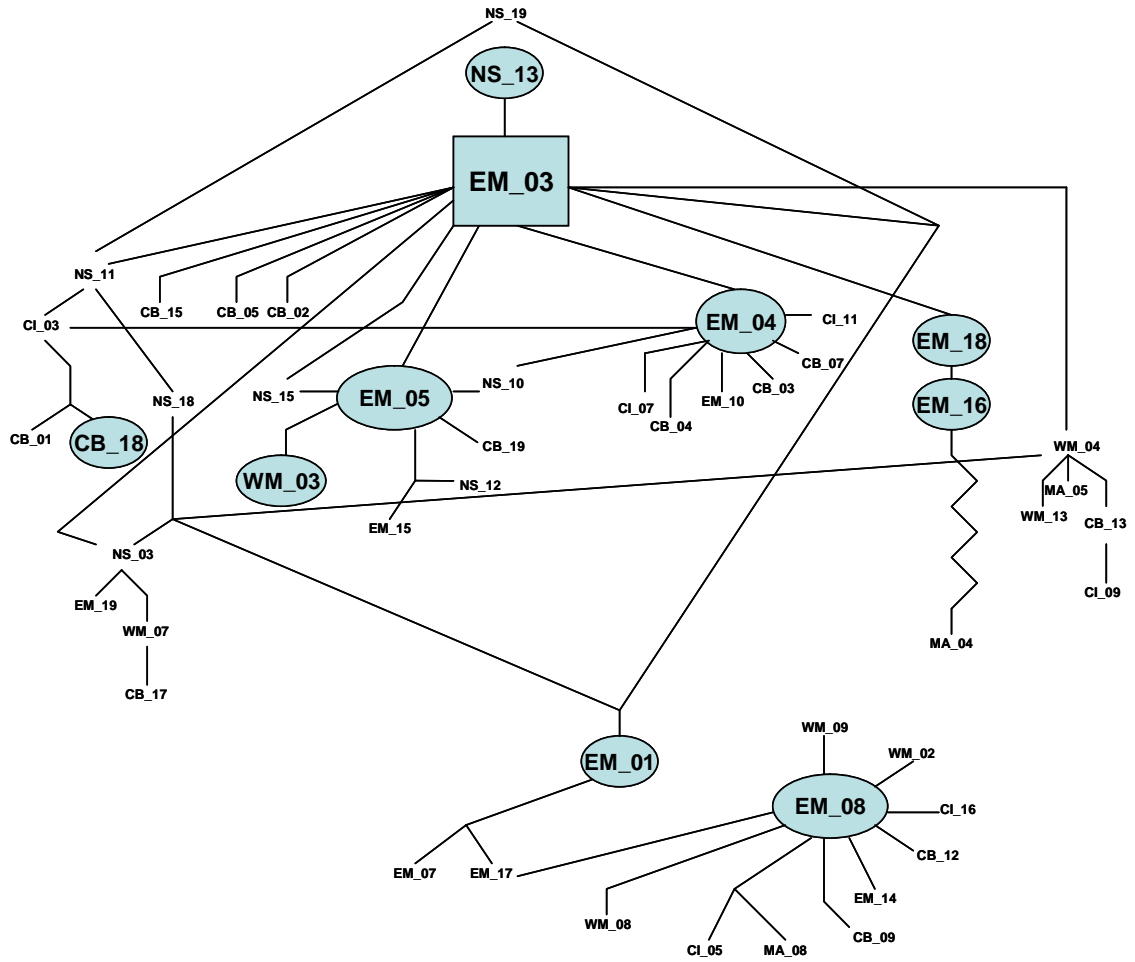


Figure 6. Connectivity network of *Mullus surmuletus* haplotypes. Symbols are as in Figure 1. For shared haplotypes see the respective table of the analyses.

Reference for the TCS algorithm:

Posada C.M. and Crandall K.A. (2000) TCS: A computer program to estimate gene genealogies. *Mol. Ecol.*, **9**, 1657-1660

Annex XXI: Final Report on Validation and Missing Data.

FishTrace Project

FINAL REPORT ON VALIDATION AND MISSING DATA

by

Partner ICCM with assistance of Partner TFMC

Canary Islands, June 28th – 29th, 2006

**THIS REPORT HAS BEEN PERMANENTLY UPDATED
UNTIL JULY 18TH AT NOON**

Workpackage 7: Validation in Database

Start date: Month 10

Completion date: Month 40

Partner responsible: 4 (ICCM)

Objectives: The data compiled in the database are being generated by different teams of taxonomists, molecular geneticists and curators. This important amount of valuable information requires to be contrasted for reliability and standardisation among the different sets of data and the various groups involved.

Deliverables:

D8. Definitive list of refined, guaranteed data in the end-users database.

D9. Periodical reports on the identification of missing data and data subjected to review and bimonthly reports on the data already validated.

Milestones:

M10. Validation of taxonomic, biological, genetic and reference collections pooled data.

M11. Validation of data for end-users.

INTRODUCTION

After the General Meeting in Paris (November 2004), a Working Group co-ordinated by ICCM with the participation of TFMC, IMAR and MNHN prepared and circulated a 16-page document with the Guidelines for Validation Purposes containing: objectives, partners/groups involved, the validation process (steps), a protocol defining format for common fields in database species table (tasks to be done for each assigned species), as well as for regional fields in the database regional info table (tasks to be done for each sampled species), and also some remarks on reference pictures (both vouchers and otoliths) in the database species table (tasks to be done for each assigned species).

Main research tasks allocated to each Partner

Sampling & Taxonomy: ICCM-TFMC, IMAR, NAGREF, IFREMER-MNHN & NRM.

Molecular Genetics: UCM, RIVO, NAGREF, IFREMER & NRM.

Reference Collections: TFMC, IMAR, MNHN & NRM.

Each “taxonomical/collection group” is responsible for checking data and images generated by itself in order to detect possible errors or missing information.

Each “genetic group” is responsible for detecting possible errors related to the species identification.

Periodical Validation Reports on Data Status Summary

According to the Paris meeting, several Bimonthly Validation Reports on Data Status Summary have been elaborated by ICCM with the assistance of TFMC. Several regional meetings (workshops) at the Canary Islands have been celebrated in order to prepare these periodical reports with the following characteristics:

First workshop on Validation: Canary Islands, April 14-16 and 28-30, 2005

Tasks done:

- Preliminary list of target species by geographical area and for the project
- Preliminary assignment of target species to partners/groups for taxonomic and reference collection purposes
- Preliminary assignment of target species to partners/groups for genetic purposes
- Preliminary Excel file to control the Data Validation Flow

Documents produced:

- Data Status Summary Validation (Preliminary Report April 05).doc
- Species Lists Validation (taxonomy-collections) 30-4-2005.xls
- Species Lists Validation (genetics) 30-4-2005.xls
- Control Data Validation Flow 30-4-2005.xls

Sent to all fishtracers on May 1st, 2005 and presented to all Partners during the General Meeting in Funchal, Portugal (October 2005).

Actions recommended:

- JRC should be update the final list of species
- JRC should be upload the file to control the data validation flow

Second workshop on Validation: Canary Islands, May 26-28, 2005

Tasks done:

- Final list of target species by geographical area and for the FishTrace project
- Definitive assignment of target species to partners/groups for taxonomic and reference collection purposes
- Definitive assignment of target species to partners/groups for genetic purposes

- Excel file controlling the Data Validation Flow

Documents produced:

- Data Status Summary Validation (Final Report May 05).doc
- Species Lists Validation (taxonomy-collections) 28-5-2005.xls
- Species Lists Validation (genetics) 28-5-2005.xls
- Control Data Validation Flow 28-5-2005.xls

Sent to all fishtracers on May 28th, 2005 and presented to all Partners during the General Meeting in Funchal.

Actions recommended:

- Check carefully the definitive assignment of species to partners/groups
- JRC should be re-upload the file to control the data validation flow

Third workshop on Validation: Canary Islands, September 27-October 3, and October 13-15, 2005

Tasks done:

- Statistics on Data Status Summary on Validation, according to the current content of the “Control Data Validation Flow” file in the database

Documents produced:

- Data Status Summary Validation (Report October 2005).doc
- Control Data Validation Flow 15-10-2005.xls

Actions recommended:

- JRC should be introduce the complete specimens tables including all information loaded for a particular species (including both filled and empty fields)
- Upload all specimen and species information ASAP enabling the validation process

Communicated to all Partners on October 24-26, 2005 during the **General Meeting in Funchal**, Portugal. This Power-point presentation contained the following information:

- WP7 Progress: Statistics on Data Status Summary on Validation (update: October 19, 2005):
 - Validation of each target species at each area: TAXONOMY
 - Validation of each target species at each area: COLLECTIONS
 - Validation of each target species at each area: GENETICS
 - Validation of each assigned species: TAXONOMY & COLLECTIONS
 - Validation of each assigned species: GENETICS
- Expected TIMETABLE to complete WP7.
- Thematic discussion on WP7 progress and expected actions to conclude (Rapporteur: J.A. González, ICCM).

Once both presentations by ICCM on WP7 progress were done and discussed among Partners, the concluding remarks and expected actions to conclude were as follows (two documents with a timetable were produced):

Resumed Protocol for the Online Validation Process. Working group: J.A.González (ICCM), S.Jiménez (TFMC) & I.J.Lozano (ICCM). A 2-page document containing the steps to be done by each partner at each area (internal validation or self-validation), and the steps to be done by the partner/group in charge of a particular species (true validation process).

Agreements on thematic discussion on progress of WP7 and expected actions to conclude. A document containing the Expected TIMETABLE to complete WP7, which was summarised in the following Table:

TASK	DEADLINE
Self-validation of specimens' information on database, including texts and pictures (both fish & otolith) (D8: Definitive list of refined, guaranteed data in DB)	28 FEB 2006
Validation of the assigned species on database <u>Warning</u> : this task depends on the work done by the other group(s) involved in a particular species (D8: Definitive list of refined, guaranteed data in DB)	30 APR 2006
Uploading of regional information on species to database, including the corresponding bibliographic references (D8: Definitive list of refined, guaranteed data in DB)	30 APR 2006
Last periodical and final report on validation and missing data (D9: Reports on validation and missing data)	31 MAY 2006

Fourth workshop on Validation: Canary Islands, November 4-6, 2005

Tasks done:

- Final list of target species by geographical area and for the FishTrace project.
- Updating of the definitive assignment of target species to partners/groups for taxonomic and reference collection purposes.
- Updating of the definitive assignment of target species to partners/groups for genetic purposes.
- Updating of the Excel file to control the Data Validation Flow by partner.

Documents produced:

- Species Lists Validation (taxonomy-collections) 06-11-2005.xls
- Species Lists Validation (genetics) 06-11-2005.xls
- Control Data Validation Flow_ALL_PARTNERS_02-11-2005.xls

Uploaded to the FishTrace website: November 2nd, 2005.

Fifth workshop on Validation: Canary Islands, December 3-4, 2005

Tasks done:

- Statistics on Data Status Summary on Validation, according to the current content of the "Control Data Validation Flow" file in the database.

Concluding General Meeting in Kavala, Greece (May 10-12, 2006)

Two Power-point presentations by ICCM were done to all Partners on May 10-12, 2006 during the Concluding General Meeting of the project, celebrated in Kavala, Greece. Both presentations contained basically the following information: WP7 Degree of Data Validation in the Database, based on nine Excel files “Control Data Validation Flow_PARTNER_date.xls” (available from website until May 5th, 2006):

- Steps to be done by each partner at each area (self-validation)
- Statistics on Data Status Summary on Validation:
 - Self-validation of target species at each area: TAXONOMY
 - Validation of target species at each area: COLLECTIONS
 - Validation of target species at each area: GENETICS
 - Validation of assigned species: TAXONOMY & COLLECTIONS
 - Validation of assigned species: GENETICS
- Expected TIMETABLE to complete WP7

TASK	NEW DEADLINE
Self-validation of specimens' information on database, including texts and pictures (both fish & otolith) (D8: Definitive list of refined, guaranteed data in DB)	(*)
Validation of the assigned species on database <u>Warning:</u> this task depends on the work done by the other group(s) involved in a particular species (D8: Definitive list of refined, guaranteed data in DB)	(*)
Uploading of regional information on species to database, including the corresponding bibliographic references (D8: Definitive list of refined, guaranteed data in DB)	(*)
Last periodical and final report on validation and missing data (D9: Reports on validation and missing data)	(*)

(*) According with the Fishtrace Concluding Meeting in Kavala (See Minutes, point 6 on Open discussion on actions to be taken to complete data validation on time), “... the main conclusion was that all partners have to ENTER DATA on database and to COMPLETE THE SELF-VALIDATION process. Only reaching these steps first, the definitive validation can be done.”

DELIVERABLES EMANATING FROM VALIDATION WORK ACCOMPLISHED

1. Definitive list of target species by geographical area and for the FishTrace project

Definitively, FishTrace consists of 220 teleostean species.

Although *Muraena augusti* Kaup, 1856 (Muraenidae) and *Spicara flexuosa* Rafinesque, 1810 (Centranchidae) are currently considered as synonyms of *Muraena helena* Linnaeus, 1758 and *Spicara maena* (Linnaeus, 1758) respectively by FishBase, fishtracers consider both *M. augusti* and *S. flexuosa* as valid species as it will be taxonomically and genetically shown in FishTrace.

Final lists of target species by geographical area include:

- Area A, Skagerrak & Baltic Sea (BS): 48 target species.
- Area B, North Sea (NS): 52 target species.
- Area C, English Channel & Bay of Biscay (CB): 59 target species.
- Area D, Cantabrian Sea & NW Iberian Peninsula (CS): 56 target species.
- Area E, Madeiran archipelago (MA): 52 target species.
- Area F, Canary Islands (CI): 51 target species.
- Area G, Western Mediterranean & Bay of Cadiz (WM): 91 target species.
- Area H, Eastern Mediterranean (Greek Seas) (EM): 61 target species.
- Area I, extra-European seas (EE): 45 target species.

2. Definitive assignment of target species to partners/groups for taxonomic and reference collection purposes

This amount of 220 target species was definitively assigned to the following partners/groups for validation purposes on taxonomy and collections:

- ICCM-TFMC: 81 species.
- IMAR: 36 species.
- NAGREF: 34 species.
- IFREMER-MNHN: 36 species.
- NRM: 33 species.

See the attached Excel file “Species Lists Validation (taxonomy-collections) 29-06-2006.xls”.

3. Definitive assignment of target species to partners/groups for genetic purposes

The same 220 target species was definitively assigned to the following partners for validation purposes on genetics:

- UCM: 105 species.
- RIVO: 29 species.
- NAGREF: 30 species.
- IFREMER: 29 species.
- NRM: 27 species.

See the attached Excel file “Species Lists Validation (genetics) 29-06-2006.xls”.

4. Excel file to control the Data Validation Flow

In accordance to the agreements of Funchal meeting, one Excel file per Partner was elaborated to control permanently, through the online website, the data validation flow.

These Excel files (“Control Data Validation Flow_PARTNER_date.xls”), together with the compilation file (“Control Data Validation Flow_ALL_PARTNERS_date.xls”), were uploaded during this General Meeting.

A **Protocol for the online Validation Process** was elaborated by Partners ICCM and TFMC and presented to the FishTrace Consortium during the Paris meeting.

Once refined and abbreviated, a new 2-page document entitled **Resumed Protocol for the online Validation Process** was prepared by the same Working Group and presented to the Consortium during the Funchal meeting with a second practical demonstration. This document differentiated the “internal validation or self validation” (with four steps, to be done by each partner at each sampling area) from the “true validation process” (with four more steps, to be done by the partner/group in charge of a particular species at each sampling area).

Finally, this **Resumed Protocol for the online Validation Process** was explained again during the Kavala meeting and a third practical demonstration was carried out.

5. Statistics on Data Status Summary on Validation

According to the current (last checking, July 18th, 2006) content of the control data validation flow files in the webpage, the status of the validation process to date can be summarised as follows:

5.1. Validation of target species by area

5.1.1. Taxonomy

Area	Partner	Target species	Species validated	% done
BS	NRM	48	38	79
NS	RIVO	52	52	100
CB	MNHN	59	0	0
CS	ICCM	56	56	100
MA	IMAR	52	0	0
CI	ICCM	51	51	100
WM	ICCM	91	91	100
EM	NAGREF	61	61	100
EE	ICCM	45	45	100

5.1.2. Reference collections

Area	Partner	Target species	Species validated	% done
BS	NRM	48	37	77
NS	MNHN	52	0	0
BB	MNHN	59	0	0
CS	TFMC	56	56	100
MA	IMAR	52	0	0
CI	TFMC	51	51	100
WM	TFMC	91	91	100
EM	MNHN	61	61	100
EE	TFMC	45	45	100

5.1.3. Genetics

Area	Partner	Target species	Species validated	% done
BS	NRM	48	43	90
NS	RIVO	52	51	98
BB	IFREMER	59	54	92
CS	UCM	56	37	66
MA	UCM	52	27	52
CI	UCM	51	36	71
WM	UCM	91	60	66
EM	NAGREF	61	61	100
EE	UCM	45	24	53

5.2. Validation of assigned species

5.2.1. Taxonomy and collections

Partner/Group	Species assigned	Species validated		% done	
		TAXO	COLL	TAXO	COLL
NRM	33	1	1	3	3
IFREMER-MNHN	36	0	0	0	0
ICCM-TFMC	81	60	60	74	74
IMAR	36	0	0	0	0
NAGREF	34	1	1	3	3
Totals	220	62	62	28	28

5.2.2. Genetics

Partner	Species assigned	Species validated	% done
NRM	27	0	0
RIVO	29	17	59
IFREMER	29	15	52
UCM	105	63	60
NAGREF	30	0	0
Totals	220	95	43

#####

Annex XXII: Control Data Validation Flow Document.

	TAXONOMY	GENETICS	COLLECTIONS	WM			CS			CI			EE			MA			BS			NS			CB			EM			AREAS
	Finished	Finished	Finished	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	
Dentex dentex	X	X	X	X	X	X				X	X	X																X	X	X	3
Dentex gibbosus	-2	X	-2													X	-2	-2													1
Dentex macrophthalmus	X	X	X							X	X	X	X	X	X																2
Dicentrarchus labrax	-2	-2	-2				X	X	X													X	X	-2	-2	-2	-2	X	X	X	4
Dicologlossa cuneata	X	X	X	X	X	X							X	X	X																2
Diplodus annularis	X	X	X	X	X	X				X	X	X															X	X	X	3	
Diplodus cervinus	-2	X	-2							X	X	X	X	X	X	X	-2	-2													3
Diplodus puntazzo	X	X	-2							X	X	X	X	X	X												X	X	X	3	
Diplodus sargus	-2	X	-2	X	X	X	X	X	X							X	-2	-2									X	X	X	4	
Diplodus vulgaris	-2	X	-2	X	X	X	X	X	X							X	-2	-2									X	X	X	4	
Echiichthys vipera	-2	X	-2																			X	X	-2						1	
Enchelycore anatina	-2	-2	-2							-2	X	X				-2	-2	-2												2	
Enchelyopus cimbrius	-2	-2	-2																-2	-2	-2	X	X	-2	X	-2	-2			3	
Engraulis encrasicolus	-2	X	-2	-2	X	X	X	X	X	-2	X	X									X	X	-2	X	-2	-2	X	X	X	6	
Epinephelus caninus	X	-2	X										-2	X	X															1	
Epinephelus costae	X	-2	X	-2	X	X							X	X	X												X	X	X	3	
Epinephelus marginatus	-2	-2	-2	X	X	X							-2	X	X	X	-2	-2									X	X	X	4	
Epinephelus tauvina	X	-2	X										-2	X	X															1	
Erythrocles monodi	X	X	X										X	X	X															1	
Euthynnus alletteratus	X	X	X	X	X	X																								1	
Gadiculus argenteus	-2	X	-2																							X	-2	-2		1	
Gadus morhua	-2	-2	-2				X	X	X				-2	X	X				X	X	X	X	X	-2						4	
Gaidropsarus biscayensis	X	X	X				X	X	X																					1	
Gaidropsarus mediterraneus	X	X	X				X	X	X																					1	
Glossanodon leioglossus	X	-2	X	-2	X	X																								1	
Glyptocephalus cynoglossus	-2	-2	-2																X	X	X	X	X	-2						2	
Gymnothorax afer	X	X	X										X	X	X															1	
Gymnothorax polygonius	X	-2	X							-2	X	X				-2	-2	-2												1	
Gymnothorax unicolor	-2	-2	-2													-2	-2	-2												1	
Halobatrachus didactylus	X	-2	X	-2	X	X																								1	
Helicolenus dactylopterus	-2	X	-2	X	X	X	-2	X	X										-2	-2	-2				X	-2	-2	X	X	X	5
Heteropriacanthus cruentatus	X	X	X							X	X	X																			1
Hippoglossoides platessoides	-2	X	-2																X	X	X	X	X	-2						2	
Hippoglossus hippoglossus	-2	X	-2																X	X	X	X	X	-2						1	
Hyperoplus lanceolatus	-2	X	-2																											2	
Katsuwonus pelamis	-2	-2	-2				-2	X	X	-2	X	X				-2	-2	-2												3	
Labrus bergylta	-2	-2	-2				-2	X	X										X	-2	-2									2	
Lepidopus caudatus	X	-2	X	-2	X	X				-2	X	X																		2	
Lepidorhombus boscii	-2	-2	-2	X	X	X	X	X	X																X	-2	-2	X	X	X	4
Lepidorhombus whiffiagonis	-2	-2	-2				X	X	X																X	-2	-2	X	X	X	3
Lepidotrigla cavillone	X	X	X	X	X	X																								1	
Lichia amia	X	-2	X	-2	X	X				X	X	X																		2	
Limanda limanda	-2	-2	-2																X	-2	-2	X	X	-2	X	-2	-2			3	
Liparis liparis	-2	X	-2																			X	X	-2						1	
Liparis montagui	-2	-2	-2																-2	-2	-2									1	
Lithognathus mormyrus	X	X	-2	X	X	X																					X	X	X	2	
Liza aurata	-2	X	-2	X	X	X	-2	X	X																	X	-2	-2			3
Liza ramado	-2	X	-2																						X	-2	-2			1	
Lophius budegassa	-2	X	-2	X	X	X	X	X	X				X	X	X										X	-2	-2	X	X	X	5
Lophius piscatorius	-2	X	-2	X	X	X	X	X	X										X	X	X	X	X	-2	X	-2	-2	X	X	X	6
Makaira nigricans	X	X	X							X	X	X																		1	
Melanogrammus aeglefinus	-2	-2	-2				-2	X	X				-2	X	X				X	X	X	X	X	-2	X	-2	-2			5	

	TAXONOMY	GENETICS	COLLECTIONS	WM			CS			CI			EE			MA			BS			NS			CB			EM			AREAS
	Finished	Finished	Finished	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	GEN	TAX	COL	
Trachurus trachurus	-2	-2	-2	-2	X	X	-2	X	X									X	X	X	X	X	-2	X	-2	-2	X	X	X	6	
Trigla lyra	X	X	-2	-2	X	X																				X	X	X	2		
Trigloporus quadricornis	-2	-2	-2															X	X	X										1	
Trisopterus esmarkii	-2	X	-2															X	X	X	X	X	-2							2	
Trisopterus luscus	-2	-2	-2				-2	X	X												X	X	-2	X	-2	-2				3	
Trisopterus minutus	-2	X	-2	X	X	X				X	X	X									X	X	-2	X	-2	-2	X	X	X	4	
Tylosurus acus	X	X	X	X	X	X				X	X	X																		2	
Umbrina canariensis	X	X	X							X	X	X																		1	
Uranoscopus scaber	X	-2	-2	X	X	X																				X	X	X	2		
Xiphias gladius	-2	X	-2	-2	X	X				-2	X	X												X	-2	-2	X	X	X	4	
Xyrichtys novacula	-2	X	-2												X	-2	-2													1	
Zeus faber	-2	X	-2	X	X	X	X	X	X	X	X	X												X	-2	-2	X	X	X	5	
Zoarces viviparus	-2	X	-2															X	X	X	X	X	-2							2	

X in red means a new target species for the area
-1 in red means a target species not studied for the area

Annex XXIII: Leaflet presenting the FishTrace project.

Collections

► The biological reference collections:

- Voucher samples of each species
- Otoliths
- Muscle samples
- DNA samples

► Reference collections network:

The specimens sampled in eight areas (from the Skagerrak & Baltic sea to Canary Islands and Mediterranean sea) have been deposited in collections at:

- Swedish Museum of Natural History, Stockholm (Sweden)
- French National Museum of Natural History, Paris (France)
- Tenerife Museum of Natural History, Tenerife (Spain)
- Funchal Museum of Natural History, Madeira (Portugal)



The Consortium

Complutense University of Madrid (UCM) (Spain)
[Coordination]
Contact: José M. Bautista (jmbau@vet.ucm.es)

Joint Research Centre, European Commission (JRC-IPSC)
(Italy)
Contact: Philippe Carreau (philippe.carreau@jrc.it)

Swedish Museum of Natural History (NRM) (Sweden)
Contact: Sven O Kullander (sven.kullander@nrm.se)

Canarian Institute of Marine Sciences (ICCM) (Spain)
Contact: José A. González (solea@iccm.rcanaria.es)

French Research Institute for the Exploitation of the Sea (Ifremer) (France)
Contacts: Monique Etienne (Monique.Etienne@ifremer.fr)
Véronique Verrez-Bagnis (Veronique.Verrez@ifremer.fr)

Netherlands Institute for Fisheries Research (RIVO)
(Netherlands)
Contact: Hilde van Pelt-Heerschap (Hilde.vanpelt@wur.nl)

Institute of Marine Research (IMAR) (Portugal)
Contact: Manuel J. Biscoito (manuel.biscoito@mail.cm-funchal.pt)

Tenerife Museum of Natural History (TFMC) (Spain)
Contact: Sebastián Jiménez
(chano@museosdetenerife.org)

National Agricultural Research Foundation (NAGREF)
(Greece)
Contact: Grigorios Krey (krey@otenet.gr)

French National Museum of Natural History (MNHN)
(France)
Contact: Guy Duhamel (duhamel@mnhn.fr)
Patrice Pruvost (pruvost@mnhn.fr)

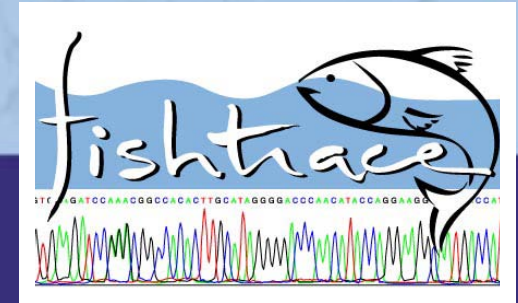


FishTrace is funded under the Quality of Life and Management of Living Resources Programme of the EU's 5th Framework Programme in Support for European Research Infrastructures

(QLRI-CT-2002-02755)



A tool for identification of fish species and traceability of fish products



FishTrace is a genetic catalogue associated to biological reference collections from more than 200 commercial marine fish. Genetic and taxonomic information is compiled in an online database.

www.fishtrace.org

Presentation

FishTrace is a strategic tool to control the authenticity of fishery products.

It provides useful biological and genetic information which guarantees the identification and the traceability of more than 200 marine fish species of interest to European markets.



► FishTrace provides information and material for a wide variety of end-users, including research, academia, biotechnology, national and transnational regulatory bodies, fisheries industry, food and transforming industry, consumers associations, control laboratories, among others

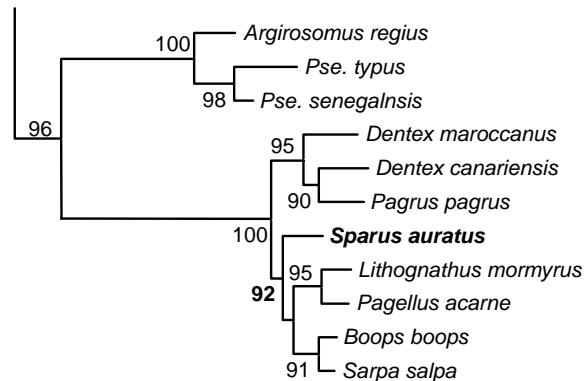
► FishTrace makes available, through a public Web site, reliable biological information and multidisciplinary methodologies for authenticity analysis

► FishTrace fulfils the consumers' demands in terms of origin and identification of fishery products

The Project

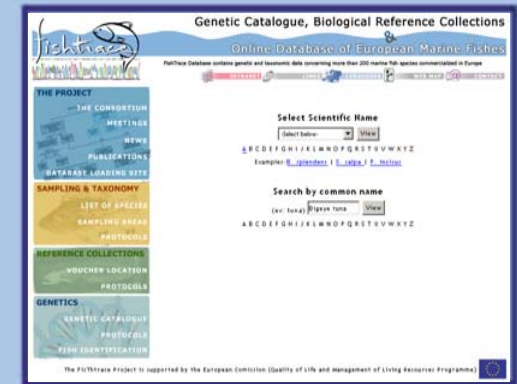
For each fish species, fresh specimens have been obtained from well-known origins. Each specimen has been identified, measured and photographed, and tissue samples have been obtained for molecular genetic analysis and for storage in reference collections.

The nucleotide sequence of two genes, cytochrome b and rhodopsin, have been determined. Intra-species variability (genetic variation in populations) has been also investigated. Molecular data is provided to identify each fish species through robust phylogenetic analysis.



Biological reference collections have been created to guarantee long-term sample preservation. Material exchange allows their use for standardization analysis and fishery products traceability.

www.fishtrace.org



The online database offers three levels of information:

► **Taxonomic and biological information on the fish species:** description, diagnosis, similar species, biology, distribution, interest to fisheries and high quality scientific pictures

► **Molecular genetic information on the fish species:** reference DNA sequence of cytochrome b and rhodopsin genes, and DNA polymorphisms from each species. Genetic data also include standardized methodologies for DNA extraction, PCR amplification and sequence analysis

► **Information about the biological collections:** access to the specimen samples stored in collections (entire fish, otoliths, extracted tissues, DNA). Sampling and collection data include sampling GIS localisation, ontogenic stage (juvenile, adult), length, weight, etc...